Innovation, Sciences et Développement économique Canada
Office de la Propriété Intellectuelle du Canada

Innovation, Science and Economic Development Canada
Canadian Intellectual Property Office

CA 2884162 C 2020/12/29

(11)(21) 2 884 162

(12) BREVET CANADIEN
CANADIAN PATENT

(13) C

(86) Date de dépôt PCT/PCT Filing Date: 2013/09/05
(87) Date publication PCT/PCT Publication Date: 2014/03/13
(45) Date de délivrance/Issue Date: 2020/12/29
(85) Entrée phase nationale/National Entry: 2015/03/05
(86) N° demande PCT/PCT Application No.: US 2013/058267
(87) N° publication PCT/PCT Publication No.: 2014/039684
(30) Priorités/Priorities: 2012/09/07 (US61/697,854); 2013/05/07 (US61/820,260)

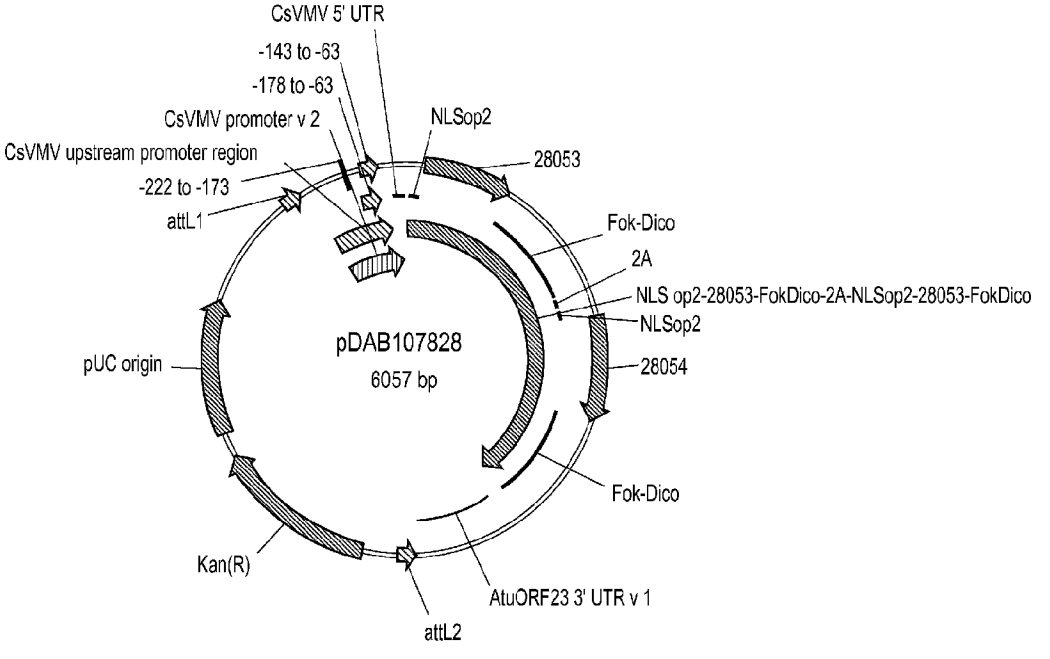
(51) Cl.Int./Int.Cl. *C12N 9/22* (2006.01), *C12N 15/00* (2006.01), *C12N 15/09* (2006.01), *C12N 15/53* (2006.01), *C12N 15/63* (2006.01), *C12N 15/82* (2006.01), *C12N 5/04* (2006.01), *C12N 5/10* (2006.01), *C12N 9/02* (2006.01)

(72) Inventeurs/Inventors:
COGAN, NOEL, US;
FORSTER, JOHN, US;
HAYDEN, MATTHEW, US;
SAWBRIDGE, TIM, US;
...

(73) Propriétaires/Owners:
DOW AGROSCIENCES LLC, US; ...

(54) Titre : LOCI DE PERFORMANCE FAD3 ET PROTEINES SE LIANT A UN SITE SPECIFIQUE CIBLE CORRESPONDANTES POUVANT INDUIRE DES CASSURES CIBLEES

(54) Title: FAD3 PERFORMANCE LOCI AND CORRESPONDING TARGET SITE SPECIFIC BINDING PROTEINS CAPABLE OF INDUCING TARGETED BREAKS



(57) Abrégé/Abstract:
A method of gene editing or gene stacking within a FAD3 loci by cleaving, in a site directed manner, a location in a FAD3 gene in a cell, to generate a break in the FAD3 gene and then ligating into the break a nucleic acid molecule associated with one or more traits of interest is disclosed.

CA 2884162 C 2020/12/29

(11)(21) 2 884 162

(13) C

(72) Inventeurs(suite)/Inventors(continued): SPANGENBERG, GERMAN, US; WEBB, STEVEN R., US; GUPTA, MANJU, US; AINLEY, WILLIAM MICHAEL, US; HENRY, MATTHEW J., US; MILLER, JEFFREY C., US; GUSCHIN, DMITRY Y., US

(73) Propriétaires(suite)/Owners(continued):SANGAMO THERAPEUTICS, INC., US

(74) Agent: SMART & BIGGAR LLP

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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
13 March 2014 (13.03.2014)

WIPO | PCT

(10) International Publication Number
WO 2014/039684 A1

(51) International Patent Classification:
C12N 5/10 (2006.01)

(21) International Application Number:
PCT/US2013/058267

(22) International Filing Date:
5 September 2013 (05.09.2013)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/697,854 7 September 2012 (07.09.2012) US
61/820,260 7 May 2013 (07.05.2013) US

(71) Applicants: **DOW AGROSCIENCES LLC** [US/US]; 9330 Zionsville Road, Indianapolis, Indiana 46268-1054 (US). **SANGAMO BIOSCIENCES, INC.** [US/US]; Point Richmond Tech Center, 501 Canal Blvd., Suite A100, Richmond, California 94804 (US).

(72) Inventors: **COGAN, Noel**; c/o Dow Agrosciences LLC, 9330 Zionsville Road, Indianapolis, Indiana 46268-1054 (US). **FORSTER, John**; c/o Dow Agrosciences LLC, 9330 Zionsville Road, Indianapolis, Indiana 46268-1054 (US). **HAYDEN, Matthew**; c/o Dow Agrosciences LLC, 9330 Zionsville Road, Indianapolis, Indiana 46268-1054 (US). **SAWBRIDGE, Tim**; c/o Dow Agrosciences LLC, 9330 Zionsville Road, Indianapolis, Indiana 46268-1054 (US). **SPANGENBERG, German**; c/o Dow Agrosciences LLC, 9330 Zionsville Road, Indianapolis, Indiana 46268-1054 (US). **WEBB, Steven R.**; c/o Dow Agrosciences LLC, 9330 Zionsville Road, Indianapolis, Indiana 46268-1054 (US). **GUPTA, Manju**; c/o Dow Agrosciences LLC, 9330 Zionsville Road, Indianapolis, Indiana 46268-1054 (US). **AINLEY, William Michael**; c/o Dow Agrosciences LLC, 9330 Zionsville Road, Indianapolis, Indiana 46268-1054 (US). **HENRY, Matthew J.**; c/o Dow Agrosciences LLC, 9330 Zionsville Road, Indianapolis, Indiana 46268-1054 (US). **MILLER, Jeffrey C.**; c/o Sangamo BioSciences, Inc., Point Richmond Tech Center, 501 Canal Blvd., Suite A100, Richmond, California 94804 (US). **GUSCHIN, Dmitry Y.**; c/o Sangamo BioSciences, Inc., Point Richmond Tech Center, 501 Canal Blvd., Suite A100, Richmond, California 94804 (US).

(74) Agent: **PASTERNAK, Dahna S.**; Pasternak Patent Law, 1900 Embarcadero Rd., Suite 211, Palo Alto, California 94303 (US).

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

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ,

[Continued on next page]

(54) Title: FAD3 PERFORMANCE LOCI AND CORRESPONDING TARGET SITE SPECIFIC BINDING PROTEINS CAPABLE OF INDUCING TARGETED BREAKS

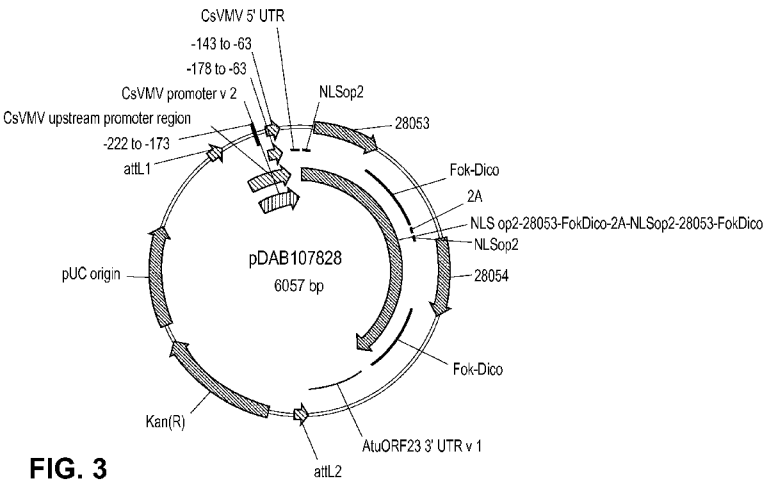


FIG. 3

(57) Abstract: A method of gene editing or gene stacking within a FAD3 loci by cleaving, in a site directed manner, a location in a FAD3 gene in a cell, to generate a break in the FAD3 gene and then ligating into the break a nucleic acid molecule associated with one or more traits of interest is disclosed.

WO 2014/039684 A1

WO 2014/039684 A1

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))*

FAD3 PERFORMANCE LOCI AND CORRESPONDING TARGET SITE SPECIFIC BINDING PROTEINS CAPABLE OF INDUCING TARGETED BREAKS

CROSS -REFERENCE TO RELATED APPLICATIONS

[0001] The present application claims priority to the benefit of U.S. Provisional Patent Application No. 61/697,854, filed September 7, 2012, and to U.S. Provisional Patent Application No. 61/820,260, filed on May 7, 2013.

FIELD OF THE DISCLOSURE

[0002] The present disclosure relates generally to compositions and methods for use in recombinant plant technology (for example, for generating a transgenic plant). More specifically, the present disclosure relates to plant cells and plants including loci within their genomes that may be used for the site-specific introduction of any nucleic acid of interest.

BACKGROUND

[0003] Many plants are genetically transformed with exogenous nucleic acids (*e.g.*, transgenes) to introduce desirable traits, for example, to improve agricultural value. Examples of improvements in agricultural value that can be achieved through genetic transformation include: improved nutritional quality, increased yield, pest or disease resistance, drought and stress tolerance, improved horticultural quality (*e.g.*, improved pigmentation and/or growth), herbicide resistance, production of industrially useful compounds and/or materials from the plant, and/or production of pharmaceuticals. The introduction of cloned genes into plant cells and recovery of stable fertile transgenic plants can be used to make a genetic modification of a plant stable through multiple generations, and thereby allow the genetic engineering of a crop plant.

[0004] In methods for genetic transformation and transgenic plant production, exogenous DNA is typically randomly introduced into the nuclear or plastid DNA of a eukaryotic plant cell, followed by isolation of cells containing integrated exogenous DNA, and subsequent regeneration of a stably transformed plant. Transgenic plants were typically generated by *Agrobacterium*-mediated transformation technology. Successes with these techniques spurred the development of other methods to introduce a nucleic acid molecule of

interest into the genome of a plant, such as PEG-mediated DNA uptake in protoplasts, microprojectile bombardment, and silicon whisker-mediated transformation.

[0005] In all of these plant transformation methods, however, the exogenous nucleic acids incorporated in the plant genome are integrated randomly in the genome of the plant cell, and in unpredictable copy number. Terada et al. (2002) *Nat Biotechnol* 20(10):1030; Terada et al. (2007) *Plant Physiol* 144(2):846; D'Halluin et al. (2008) *Plant Biotechnology J.* 6(1):93. For example, the transgenes are frequently integrated in the form of sequence repeats, either of the whole transgene or of parts thereof. Such a complex integration pattern commonly adversely impacts the expression level of the integrated nucleic acid (*e.g.*, by destruction of transcribed RNA through post-transcriptional gene silencing mechanisms, or by inducing methylation of the integrated DNA). Also, the location of the integration site commonly influences the level of expression of the integrated nucleic acid. Moreover, the integration of the exogenous DNA may have a disruptive effect on the region of the genome where the integration occurs, and thereby influence or disturb the normal function of that target region to produce undesirable side-effects. The combination of factors including the foregoing results in a wide variation in the level of expression of transgene or exogenous DNA (and overall agronomic quality) between different transgenic plant cell and plant lines, even those created by the same methods. Because the integration is random, these effects are not able to be controlled by the practitioner while he or she attempts to produce a new plant with desirable characteristics.

[0006] The foregoing considerations necessitate that, whenever the effects of introducing a particular exogenous nucleic acid into a plant is investigated, a large number of transgenic plant lines must be generated and analyzed in order to obtain significant results. Likewise, in the generation of a transgenic plant containing a particular integrated nucleic acid so as to provide the transgenic plant with a desired phenotype, a large population of independently created transgenic plant lines must be created to allow the selection of a plant line with optimal expression of the nucleic acid, and with minimal or no side-effects on the overall phenotype and performance of the transgenic plant. These practical considerations take on added importance in transgenic plants created by inserting multiple exogenous nucleic acids (*i.e.*, gene stacking). In such plants, phenomena such as post-transcriptional gene silencing may be amplified.

[0007] Several methods have been developed in an effort to control transgene insertion in plants. *See, e.g.*, Kumar and Fladung (2001) Trends Plant Sci. 6:155-9. These methods rely on homologous recombination-based transgene integration, which has been successfully applied

both in prokaryotes and lower eukaryotes. Paszkowski *et al.* (1988) EMBO J. 7:4021-6. However, until recently in plants, the predominant mechanism for transgene integration has been based on illegitimate recombination, which involves little homology between recombining DNA strands. A major challenge in this area is therefore the detection and selective generation of rare homologous recombination events, which are masked by far more efficient integration events *via* illegitimate recombination. Moreover, even if the selective generation and detection of targeted homologous recombination events is achieved, the event must be targeted to a desirable location in the host genome in order to realize the maximum benefit of this strategy.

[0008] For example, an assumed benefit of targeted genetic transformation is the reduction in event-to-event variability of transgene expression, as compared to transformation events that are obtained from random integration. A further assumed benefit is a significant reduction in the number of events required to screen introduced nucleic acids, sort transformation constructs, and produce events that contribute to desirable overall characteristics in the resulting transgenic plant. A critical factor required to realize these benefits is the identification of specific locations in the genome where transgene performance is consistent, and if possible, where adverse effects on the host plant are eliminated or minimized.

[0009] Recently, methods and compositions for targeted cleavage of genomic DNA have been described. Such targeted cleavage events can be used, for example, to induce targeted mutagenesis, induce targeted deletions of cellular DNA sequences, and facilitate targeted recombination and integration at a predetermined chromosomal locus. *See*, for example, Urnov *et al.* (2010) *Nature* 435(7042):646-51; United States Patent Publications 20030232410; 20050208489; 20050026157; 20050064474; 20060188987; 20090263900; 20090117617; 20100047805; 20110207221; 20110301073; 2011089775; 20110239315; 20110145940; and International Publication WO 2007/014275.

Cleavage can occur through the use of specific nucleases such as engineered zinc finger nucleases (ZFN), transcription-activator like effector nucleases (TALENs), or using the CRISPR/Cas system with an engineered crRNA/tracr RNA ('single guide RNA') to guide specific cleavage. U.S. Patent Publication No. 20080182332 describes the use of non-canonical zinc finger nucleases (ZFNs) for targeted modification of plant genomes; U.S. Patent Publication No. 20090205083 describes ZFN-mediated targeted modification of a plant EPSPS locus; U.S. Patent Publication No. 20100199389 describes targeted modification of a plant *Zp15* locus and U.S. Patent Publication No. 20110167521 describes targeted modification of plant genes involved in fatty acid biosynthesis. In addition,

Moehle *et al.* (2007) *Proc. Natl. Acad, Sci. USA* 104(9):3055-3060 describes using designed ZFNs for targeted gene addition at a specified locus. U.S. Patent Publication 20110041195 describes methods of making homozygous diploid organisms.

[0010] However, there remains a need for compositions and methods for modifying and/or modulating expression of *FAD3* genes in plants, including generation of plants with targeted insertions of desired transgenes at the *FAD3* locus.

BRIEF SUMMARY OF THE DISCLOSURE

[0011] The present disclosure describes compositions and methods for modulating expression of FAD3 genes (e.g., in plants, algae, and fungi) and the use of these loci as sites for the targeted integration of a nucleic acid sequence of interest (e.g., an exogenous nucleic acid sequence) into a host cell. In some embodiments, a host cell may contain one or more genomes with one or more FAD3 sequences (*e.g.*, homeologues and/or paralogs), any or all of which may be selectively modified and/or disrupted. In specific examples, the present disclosure describes FAD3A, FAD3A', FAD3C' and/or FAD3C genes, as well as corresponding homeologues or paralogs, in *Brassica napus* (*i.e.*, *B. napus* line, DH12075) and their use as loci for targeted integration of a nucleic acid sequence of interest. As described herein, though FAD3 genes are involved in fatty acid biosynthesis in the host, their modification or disruption (*e.g.*, by integration of an exogenous nucleic acid in the FAD3 coding sequence) unexpectedly may have no or minimal adverse effects on the resultant host organism.

[0012] Also described herein is the use of one or more particular FAD3 loci in tandem with a polypeptide capable of effecting cleavage and/or integration of specific nucleic acid sequences within the FAD3 loci. Examples of the use of FAD3 loci in tandem with a polypeptide capable of effecting cleavage and/or integration of the FAD3 loci include a polypeptide selected from the group consisting of zinc finger proteins, meganucleases, TAL domains, TALENs, RNA-guided CRISPR-Cas9, recombinases, leucine zippers, CRISPr/Cas and others known to those in the art. Particular examples include a chimeric ("fusion") protein comprising a site-specific DNA binding domain polypeptide and cleavage domain polypeptide (e.g., a nuclease), such as a ZFN protein comprising a zinc-finger polypeptide and a FokI nuclease polypeptide. For example, described herein is a demonstration of the *in vitro* and *in vivo* efficacy and specificity of particular ZFNs designed to bind and induce double stranded breaks in FAD3A, FAD 3A', FAD3A'', FAD3C. FAD3C', FAD3C'', and in combinations thereof without cleaving corresponding homeologues or paralogs. In some embodiments,

particular FAD3 loci may be used with any of the foregoing polypeptides to effect site-specific integration of a nucleic acid of interest that is subsequently expressed in the host while having a minimal adverse impact on the agronomic performance of the host.

[0013] In certain aspects, described herein are polypeptides comprising a DNA-binding domain that specifically binds to a FAD3 gene. In some embodiments such a polypeptide may also comprise a nuclease (cleavage) domain or half-domain (*e.g.*, a ZFN, a recombinase, a transposase, or a homing endonuclease, including a homing endonuclease with a modified DNA-binding domain, TAL domains, TALENs, RNA-guided CRISPR-Cas9), and/or a ligase domain, such that the polypeptide may induce a targeted double-stranded break, and/or facilitate recombination of a nucleic acid of interest at the site of the break. In particular embodiments, a DNA-binding domain that targets a FAD3 locus may be a DNA-cleaving functional domain. The foregoing polypeptides may be used in some embodiments to introduce an exogenous nucleic acid into the genome of a host organism exhibiting homologous recombination (e.g., a plant or animal species) at one or more FAD3 loci.(e.g., a plant or animal species) at one or more FAD3 loci. In certain embodiments, the DNA-binding domains comprise a zinc finger protein with one or more zinc fingers (*e.g.*, 2, 3, 4, 5, 6, 7, 8, 9 or more zinc fingers), and can which is engineered (non-naturally occuring) to bind to any sequence within a *FAD3* gene. Any of the zinc finger proteins described herein may bind to a target site within the coding sequence of the target gene or within adjacent sequences (*e.g.*, promoter or other expression elements). In certain embodiments, the zinc finger protein binds to a target site in an *FAD3* gene, for example, as shown in Table 4. The recognition helix regions of exemplary *FAD3*-binding zinc fingers are shown in Table 3. One or more of the component zinc finger binding domains of the zinc finger protein can be a canonical (C2H2) zinc finger or a non-canonical (*e.g.*, C3H) zinc finger (*e.g.*, the N-terminal and/or C-terminal zinc finger can be a non-canonical finger).

[0014] Also described herein are methods for disrupting or editing a FAD3 gene. Additionally described herein are genetically modified host organisms (e.g., transgenic plants) produced by methods according to embodiments of the invention. In particular examples, a transgenic organism produced by a method according to an embodiment of the invention may be, without limitation, algae, a fungus, a monocotyledonous plant, a dicotyledonous plant, etc.

[0014a] In an embodiment, there is provided a method for modifying the genome of a cell, the method comprising: cleaving, in a site-specific manner, a target site in a FAD3 gene in a cell, to thereby generate a break in the FAD3 gene, the FAD3 gene comprising any of SEQ ID NOs:20-23, SEQ ID NOs: 25-38, SEQ ID NOs: 40-45, SEQ ID NO:47 and SEQ ID NO: 49; wherein the FAD3 gene is modified at or near any of SEQ ID NOs: 20-23, SEQ ID NOs: 25-28, SEQ ID NOs: 40-45, SEQ ID NO:47 and SEQ ID NO: 49 following cleavage, and wherein the method further comprises integrating into the break a nucleic acid sequence of interest, wherein the cleaving in a site-specific manner comprises introducing one or more nucleases comprising a DNA-binding domain and a cleavage domain or cleavage half-domain into the cell or a polynucleotide encoding the nuclease, wherein the nuclease binds with specificity to the target site and cleaves at or near the target site to thereby generate the break.

[0014b] In an embodiment, there is provided a cell modified according to the method as described herein.

[0014c] In an embodiment, there is provided a site specific zinc finger nuclease that cleaves at or near a nucleic acid target site selected from the group consisting of SEQ ID NOs: 20-23, SEQ ID NOs: 25-38, SEQ ID NOs: 40-45, SEQ ID NO: 47 and SEQ ID NO: 49, wherein the zinc finger nuclease comprises from three to six zinc finger domains, each zinc finger domain comprising a recognition helix region, wherein the zinc finger protein comprises the recognition helix regions ordered and shown in a single row of the following Table:

ZFP	sF1	F2	F3	F4	F5	F6
27961	RSDNLAR (SEQ ID NO:116)	QKKDRSY (SEQ ID NO:117)	RSDNLAR (SEQ ID NO:116)	QRGNRNT (SEQ ID NO:119)	RSDHLSR (SEQ ID NO:120)	RNQDRTN (SEQ ID NO:121)
27962	DRSNLSR (SEQ ID NO:122)	RQDSRSQ (SEQ ID NO:123)	QSSDLSR (SEQ ID NO:124)	DRSALAR (SEQ ID NO:125)	TSGSLTR (SEQ ID NO:126)	N/A
27973	QSSDLSR (SEQ ID NO:124)	AASNRSK (SEQ ID NO:128)	TSGSLSR (SEQ ID NO:129)	RSDALAR (SEQ ID NO:130)	RSDVLST (SEQ ID NO:131)	WGRLRKL (SEQ ID NO:132)

27974	ERGTLAR (SEQ ID NO:133)	RSDDLTR (SEQ ID NO:134)	RSDHLSA (SEQ ID NO:135)	QHGALQT (SEQ ID NO:136)	TSGNLTR (SEQ ID NO:137)	QSGHLSR (SEQ ID NO:138)
27987	TSGSLTR (SEQ ID NO:126)	RSDHLSQ (SEQ ID NO:140)	CTRNRWR (SEQ ID NO:141)	RSDNLSE (SEQ ID NO:142)	ASKTRKN (SEQ ID NO:143)	N/A
27990	TSGSLSR (SEQ ID NO:129)	TSSNRAV (SEQ ID NO:145)	TSGNLTR (SEQ ID NO:137)	DRSALAR (SEQ ID NO:125)	RSDVLSE (SEQ ID NO:148)	RNFSLTM (SEQ ID NO:149)
27991	QSGDLTR (SEQ ID NO:150)	TSGSLSR (SEQ ID NO:129)	QSGNLAR (SEQ ID NO:152)	TSGSLSR (SEQ ID NO:129)	QSGSLTR (SEQ ID NO:154)	N/A
27992	DRSHLAR (SEQ ID NO:155)	TSGSLSR (SEQ ID NO:129)	TSSNRAV (SEQ ID NO:145)	TSGNLTR (SEQ ID NO:137)	DRSALAR (SEQ ID NO:125)	N/A
28004	QSGNLAR (SEQ ID NO:152)	HLGNLKT (SEQ ID NO:161)	RSDHLSQ (SEQ ID NO:140)	TARLLKL (SEQ ID NO:163)	QSGNLAR (SEQ ID NO:152)	QTSHLPQ (SEQ ID NO:165)
28005	RSDNLSV (SEQ ID NO:166)	TSGHLSR (SEQ ID NO:167)	TSGSLTR (SEQ ID NO:126)	RSDALST (SEQ ID NO:169)	DRSTRTK (SEQ ID NO:170)	N/A
28021	QNAHRKT (SEQ ID NO:171)	TSGNLTR (SEQ ID NO:137)	LKQMLAV (SEQ ID NO:173)	RSDNLSR (SEQ ID NO:174)	DNSNRKT (SEQ ID NO:175)	N/A
28022	RSDNLSV (SEQ ID NO:166)	QNANRIT (SEQ ID NO:177)	TSGSLSR (SEQ ID NO:129)	QSSVRNS (SEQ ID NO:179)	DRSALAR (SEQ ID NO:125)	N/A
28023	RSDNLSR (SEQ ID NO:174)	DNSNRKT (SEQ ID NO:175)	DRSNLTR (SEQ ID NO:183)	RSDVLSE (SEQ ID NO:148)	TRNGLKY (SEQ ID NO:185)	N/A
28024	RSDALAR (SEQ ID NO:130)	RSDVLSE (SEQ ID NO:148)	RSSDRTK (SEQ ID NO:188)	RSDNLSV (SEQ ID NO:166)	QNANRIT (SEQ ID NO:177)	N/A

28025	QSSDLSR (SEQ ID NO:124)	QSTHRNA (SEQ ID NO:192)	RSDNLAR (SEQ ID NO:116)	QRGNRNT (SEQ ID NO:119)	RSDHLSR (SEQ ID NO:120)	RNQDRTN (SEQ ID NO:121)
28026	DRSNLSR (SEQ ID NO:122)	RQDSRSQ (SEQ ID NO:123)	QSSDLSR (SEQ ID NO:124)	DRSALAR (SEQ ID NO:125)	TSGSLTR (SEQ ID NO:126)	N/A
28035	QSSDLSR (SEQ ID NO:124)	AASNRSK (SEQ ID NO:128)	TSGSLSR (SEQ ID NO:129)	RSDALAR (SEQ ID NO:130)	RSDTLSQ (SEQ ID NO:206)	QRDHRIK (SEQ ID NO:207)
28036	RSDDLTR (SEQ ID NO:134)	QSSDLRR (SEQ ID NO:209)	RSDHLSA (SEQ ID NO:135)	QHGALQT (SEQ ID NO:136)	TSGNLTR (SEQ ID NO:137)	QSGHLSR (SEQ ID NO:138)
28039	TSGSLSR (SEQ ID NO:129)	RSDALAR (SEQ ID NO:130)	RSDTLSQ (SEQ ID NO:206)	QRDHRIK (SEQ ID NO:207)	TSGNLTR (SEQ ID NO:137)	DRGDLRK (SEQ ID NO:219)
28040	DSSDRKK (SEQ ID NO:220)	TSGNLTR (SEQ ID NO:137)	DNYNRAK (SEQ ID NO:222)	DRSHLTR (SEQ ID NO:223)	RSDNLTT (SEQ ID NO:224)	N/A
28051	RSDNLSN (SEQ ID NO:225)	TSSSRIN (SEQ ID NO:226)	RSDNLSE (SEQ ID NO:142)	ASKTRKN (SEQ ID NO:143)	RSDALTQ (SEQ ID NO:229)	N/A
28052	RSDTLST (SEQ ID NO:230)	DRSSRIK (SEQ ID NO:231)	RSDDLSK (SEQ ID NO:232)	DNSNRIK (SEQ ID NO:233)	N/A	N/A
28053	QSSDLSR (SEQ ID NO:124)	QAGNLSK (SEQ ID NO:235)	QSGDLTR (SEQ ID NO:150)	TSGSLSR (SEQ ID NO:129)	QSGNLAR (SEQ ID NO:152)	N/A
28054	TSGSLSR (SEQ ID NO:129)	LRQTLRD (SEQ ID NO:240)	TSGNLTR (SEQ ID NO:137)	DRSALAR (SEQ ID NO:125)	RSDVLSE (SEQ ID NO:148)	RNFSLTM (SEQ ID NO:149)
28055	QSGDLTR (SEQ ID NO:150)	TSGSLSR (SEQ ID NO:129)	QSGNLAR (SEQ ID NO:152)	TSGSLSR (SEQ ID NO:129)	QSGSLTR (SEQ ID NO:154)	N/A

28056	DRSHLAR (SEQ ID NO:155)	TSGSLSR (SEQ ID NO:129)	LRQTLRD (SEQ ID NO:240)	TSGNLTR (SEQ ID NO:137)	DRSALAR (SEQ ID NO:125)	N/A

[0014d] In an embodiment, there is provided use of a seed comprising the cell as described herein to generate a transgenic plant.

[0014e] In an embodiment, there is provided use of a plant comprising the cell as described herein for producing seed.

[0014f] In an embodiment, there is provided use of a plant comprising the cell as described herein to grow a crop.

[0015] The foregoing and other features will become more apparent from the following detailed description of several embodiments, which proceeds with reference to the accompanying figures.

BRIEF DESCRIPTION OF THE FIGURES

[0016] **Figure 1, panels A to T**, show sequence alignment of FAD3 gene sequences (SEQ ID NOs:7-12), generated using AlignX®.

[0017] **Figure 2** shows a phylogenetic tree of FAD3 gene sequences generated using Jalview v 2.3 based on neighbour joining distances. The labeled sequences correspond as follows: FAD3A'/A" is described throughout this application as FAD3A'; Haplotype2 is described throughout the application as FAD3C'; Haplotype 1 is described throughout the application as FAD3C"; and, Haplotype 3 is described throughout the application as FAD3A".

[0018] **Figure 3** shows a plasmid map of pDAB107828.

[0019] **Figure 4** shows a plasmid map of pDAB107829.

[0020] **Figure 5** shows a plasmid map of pDAS000271.

[0021] **Figure 6** shows a plasmid map of pDAS000272.

[0022] **Figure 7** shows a plasmid map of pDAS000273.

[0023] **Figure 8** shows a plasmid map of pDAS000274.

[0024] **Figure 9** shows a plasmid map of pDAS000275.

[0025] **Figure 10** shows a plasmid map of pDAS000031.

[0026] **Figure 11** shows a plasmid map of pDAS000036.

[0027] **Figure 12** shows a plasmid map of pDAS000037.

[0028] **Figure 13** shows a plasmid map of pDAB107827.

[0029] **Figure 14** shows a plasmid map of pDAB107828.

[0030] **Figure 15** shows a plasmid map of pDAS000340.

[0031] **Figure 16** shows a plasmid map of pDAS000341.

[0032] **Figure 17** shows a plasmid map of pDAS000342.

[0033] **Figure 18** shows a plasmid map of pDAS000343.

[0034] **Figure 19** is a schematic which shows the locations of the primers and their position relative to the start and stop codong of Fad3C. Panel A shows the location of the primer sites for the wiltd type Fad3C locus. Panel B shows the location of the primer sites to confirm donor integration, and the possible orientations by which the donor could integrate within the Fad3C locus.

[0035] **Figure 20, panels A and B,** show sequence alignments after modification with the indicated ZFNs and donor plasmids. Figure 20A shows a sequence alignment amplified from the junction of the tGFP cassette of pDAS000341 with Fad3C at the double strand break as recognized by ZFN 28051-2A-28052. The ":" indicates the deletions located at the cut sites.

SEQ ID NO:300 to SEQ ID NO:313 are shown in the alignment. Figure 20B shows a sequence alignment amplified from the junction of the tGFP cassette of pDAS000343 with Fad3C at the double strand break as recognized by ZFN 28051-2A-28052 and ZFN 28053-2A-28054. The ":" indicates the deletions located at the cut sites. SEQ ID NO:314 to SEQ ID NO:327 are shown in the alignment.

[0036] **Figure 21, panels A and B,** show sequence alignments of sequencesamplified from the junction of the hph cassette of pDAS000340 with FAD3C at the double strand break as recognized by ZFN 28051-2A-28052. "Sample" is a unique identifier for each plant that was assayed. The ":" indicates the deletions located at the cut sites. The sequences shown in Figure 21A are for the 5' junction and the sequences shown in the Figure 21B are for the 3' junction. SEQ ID NO:368 to SEQ ID NO:375 are shown in the alignment of Figure 21A. SEQ ID NO:376 to SEQ ID NO:377 are shown in the alignment of Figure 21B.

[0037] **Figure 22** shows a sequence alignment of sequences amplified from the junction of the hph cassette of pDAS000342 with FAD3C at the double strand break as recognized by ZFN 28053-2A-28054. "Sample" is a unique identifier for each plant that was assayed. The ":" indicates the deletions located at the cut sites. The sequences shown in the Figure 22 are for the 3' junction. SEQ ID NO:378 to SEQ ID NO:379 are shown in the alignment.

[0038] **Figure 23, panels A and B,** show a sequence alignment for sequences amplified from the junction of the hph cassette of pDAS000340 with FAD3C at the double strand break as recognized by ZFN 28051-2A-28052. The ":" indicates the deletions located at the cut sites. The sequences shown in Figure 23A are for the 5' junction and the sequences shown in the box (B) are for the 3' junction. SEQ ID NO:328 to SEQ ID NO:334 are shown in the alignment of Figure 23A. SEQ ID NO:335 to SEQ ID NO:342 are shown in the alignment of Figure 23B.

[0039] **Figure 24, panels A and B,** shows a sequence alignment of sequences amplified from the junction of the hph cassette of pDAS000342 with FAD3C at the double strand break as recognized by ZFN 28053-2A-28054. The ":" indicates the deletions located at the cut sites. The sequences shown in Figure 24A are for the 5' junction and the sequences shown in Figure 24B) are for the 3' junction. SEQ ID NO:343 to SEQ ID NO:346 are shown in the alignment of Figure 24A. SEQ ID NO:347 to SEQ ID NO:351 are shown in the alignment of Figure 24B.

SEQUENCE LISTING

[0040] The nucleic acid sequences are shown using standard letter abbreviations for nucleotide bases, as defined in 37 C.F.R. § 1.822. Only one strand of each nucleic acid sequence is shown, but the complementary strand is understood to be included by any reference to the displayed strand.

DETAILED DESCRIPTION

I. Overview of several embodiments

[0041] Embodiments of the invention establish an approach for targeted integration of exogenous nucleic acids (*e.g.*, transgenes) in a host genome without greatly adversely impacting other phenotypes of the host beyond those affected by the integrated nucleic acid. Some embodiments may be used for "stacking" multiple nucleic acids in a single host genome. Such an approach uses the development and deployment of four inter-connected technologies: targeting technologies allowing the introduction of double stranded breaks in specific genomic DNA locations (*see, e.g.*, Puchta *et al.* (1993) Nucleic Acids Res. 21:5034-40; Siebert and Puchta (2002) Plant Cell 14:1121-31; D'Halluin et al. (2008) Plant Biotechnol. J. 6(1):93-102; Cai *et al.* (2009) Plant Mol. Biol. 69(6):699-709; Shukla *et al.* (2009) Nature 459(7245):437-41); Shan *et al.* (2103) Nature Biotechnol. 31:686-680; Le *et al.* (2013) Nature Biotechnol 31: 688-691; Nekrasov *et al.* (2013) Nature Biotechnol. 31:691-693, Ainely *et al.* (2013) Plant Biotechnol. J. (On Line 19 Aug); delivery technologies allowing the delivery of an optimized exogenous (donor) nucleic acid (Bibikova *et al.* (2003) Science 300(5620):764); integration technologies involving modification of the host genes (located either in the homologous recombination or NHEJ pathways) so as to increase the HDR or NHEJ frequencies for targeted donor DNA integration; analytical tools to enrich and characterize targeted integration events; and specific desired host genomic locations ("performance loci") that are genetically well-defined and that support stable gene expression across generations without greatly adversely affecting the transformed host organism. *See, also,* United States Patent Publications 20030232410; 20050208489; 20050026157; 20050064474; 20060188987; 20090263900; 20090117617; 20100047805; 20110207221; 20110301073; 2011089775; 20110239315; 20110145940; 20080182332; 20090205083; 20100199389; 20110167521. For example, in plants, a performance locus is a locus where the negative impact on the agronomic or quality properties of a transgenic plant wherein a transgene has been inserted at the locus is negligible or non-existent.

[0042] Embodiments described herein take advantage of the unexpected finding that plant *FAD3* genes are performance loci for the targeted insertion of exogenous nucleic acids (*e.g.*, gene(s); non-coding DNA sequences, such as an Engineered Landing Pads (ELPs) (U.S. Application 12/011,735) and Engineered Transgene Insertion Platform (ETIP) (pending U.S. Application No: 61/697882); and plant transformation unit(s)). The ubiquitous nature of *FAD3* loci in plants, and evidence that alteration or knock-out of *FAD3* in canola, corn, sunflower, wheat, cotton, and soybean does not carry an agronomic or quality penalty, identifies *FAD3* loci as a broad class of performance loci across commercially-relevant plant species.

[0043] Some embodiments utilize site-specific double-stranded DNA cleavage at a *FAD3* locus, for example, resulting from the delivery and expression of a target-site specific DNA recognition and cleavage protein. In specific examples, such a *FAD3*-specific DNA recognition and cleavage protein may be, for example and without limitation, a ZFN; a TALEN; RNA-guided CRISPR-Cas9 system, a recombinase (*e.g.*, Cre, Hin, RecA, Tre, and FLP recombinases); a meganuclease, and an engineered protein derived from any of the foregoing or their equivalents. Cleavage may also be effected using the CRISPR/Cas system with an engineered crRNA/tracr RNA ('single guide RNA') to guide specific cleavage. In some embodiments, such a double-strand break may be repaired via integration of a donor nucleic acid at the cleavage site within the *FAD3* performance locus, for example, by Homology Directed Repair (HDR) or Non-Homologous End Joining (NHEJ).

[0044] This disclosure exemplifies the utility of FAD3 loci as performance loci, for example, by describing the FAD3A or 3C locus in canola (*Brassica napus*), and corresponding FAD3-specific ZFNs that may be utilized to integrate an exogenous nucleic acid at the FAD3Aor 3C locus.

[0045] Embodiments of the present invention address many unsolved problems in the art. For example, the selectivity of the targeted integration approach described herein may reduce or eliminate the necessity of repeated field trials required for elimination of unwanted transgenic events, which trials are costly due to the resources involved and the burdensome regulatory requirements in this area. Furthermore, the targeted DNA insertion approaches described herein may be particularly beneficial in the process of transgene stacking.

[0046] Although the native nucleotide sequence at an endogenous FAD3 locus may be used to directly target a nucleic acid of interest, in some embodiments, a nucleic acid may first be targeted to at least one FAD3 locus of the host, such that the integration of further nucleic acid molecules of interest into the host is facilitated. In other examples, nucleotide sequences

that are not homologous to native sequences of the host organism (*e.g.*, essentially randomly generated nucleic acid sequences) that flank a DNA recognition site (*e.g.*, zinc finger recognition sites) may be utilized.

II. Terms

[0047] As used in this application, including the claims, terms in the singular and the singular forms, "a," "an," and "the," for example, include plural referents, unless the content clearly dictates otherwise. Thus, for example, a reference to "plant," "the plant," or "a plant" also refers to a plurality of plants. Furthermore, depending on the context, use of the term, "plant," may also refer to genetically-similar or identical progeny of that plant. Similarly, the term, "nucleic acid," may refer to many copies of a nucleic acid molecule. Likewise, the term, "probe," may refer to many similar or identical probe molecules.

[0048] Numeric ranges are inclusive of the numbers defining the range, and expressly include each integer and non-integer fraction within the defined range. Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art.

[0049] In order to facilitate review of the various embodiments described in this disclosure, the following explanation of specific terms is provided:

[0050] Isolated: An "isolated" biological component (such as a nucleic acid or protein) has been substantially separated, produced apart from, or purified away from other biological components in the cell of the organism in which the component naturally occurs (*i.e.*, other chromosomal and extra-chromosomal DNA and RNA, and proteins), while effecting a chemical or functional change in the component (*e.g.*, a nucleic acid may be isolated from a chromosome by breaking chemical bonds connecting the nucleic acid to the remaining DNA in the chromosome). Nucleic acid molecules and proteins that have been "isolated" include nucleic acid molecules and proteins purified by standard purification methods. The term also embraces nucleic acids and proteins prepared by recombinant expression in a host cell, as well as chemically-synthesized nucleic acid molecules, proteins, and peptides.

[0051] Cross: As used herein in regard to plants, the term "cross" or "crossed" refers to the fusion of gametes via pollination to produce progeny (*e.g.*, cells, seeds, and plants). This term encompasses both sexual crosses (*i.e.*, the pollination of one plant by another) and selfing (*i.e.*, self-pollination, for example, using pollen and ovule from the same plant).

[0052] Backcrossing: Backcrossing methods may be used to introduce a nucleic acid sequence into a plant. This technique has been widely used for decades to introduce new traits into plants. Jensen, N., Ed. Plant Breeding Methodology, John Wiley & Sons, Inc., 1988. In a typical backcross protocol, the original variety of interest (recurrent parent) is crossed to a second variety (non-recurrent parent) that carries a nucleic acid sequence of interest to be transferred. The resulting progeny from this cross are then crossed again to the recurrent parent, and the process is repeated until a plant is obtained wherein essentially all of the desired morphological and physiological characteristics of the recurrent plant are recovered in the converted plant, in addition to the transferred nucleic acid sequence from the non-recurrent parent.

[0053] Introgression: As used herein, the term "introgression" refers to the transmission of an allele (or modified allele comprising an exogenous nucleic acid) into a genetic background at a particular locus. In some embodiments, introgression of a specific allele at the locus may occur by transmitting the allele to at least one progeny via a sexual cross between two parents of the same species, where at least one of the parents has the specific allele form in its genome. Progeny comprising the specific allele may be repeatedly backcrossed to a line having a desired genetic background. Backcross progeny may be selected for the specific allele form, so as to produce a new variety wherein the specific allele form has been fixed in the genetic background. In some embodiments, introgression of a specific allele may occur by recombination between two donor genomes (*e.g.*, in a fused protoplast), where at least one of the donor genomes has the specific allele form in its genome. Introgression may involve transmission of a specific allele form that may be, for example and without limitation, a disrupted or modified allele; a transgene; a PTU; and an ELP.

[0054] Germplasm: As used herein, the term "germplasm" refers to genetic material of or from an individual plant, a group of plants (*e.g.*, a plant line, variety, and family), and a clone derived from a plant or group of plants. A germplasm may be part of an organism or cell, or it may be separate (*e.g.*, isolated) from the organism or cell. In general, germplasm provides genetic material with a specific molecular makeup that is the basis for hereditary qualities of the plant. As used herein, "germplasm" refers to cells of a specific plant; seed; tissue of the specific plant (*e.g.*, tissue from which new plants may be grown); and non-seed parts of the specific plant (*e.g.*, leaf, stem, pollen, and cells). As used herein, the term "germplasm" is synonymous with "genetic material," and it may be used to refer to seed (or other plant material) from which a plant may be propagated. A "germplasm bank" may refer to an organized collection of different

seed or other genetic material (wherein each genotype is uniquely identified) from which a known cultivar may be cultivated, and from which a new cultivar may be generated.

[0055] Gene: As used herein, the term "gene" (or "genetic element") may refer to a heritable genomic DNA sequence with functional significance. A gene may be a native nucleic acid, or a nucleic acid that has been integrated into the genome. The term "gene" may also be used to refer to, for example and without limitation, a cDNA and/or an mRNA encoded by a heritable genomic DNA sequence.

[0056] Nucleic acid molecule: As used herein, the term "nucleic acid molecule" may refer to a polymeric form of nucleotides (*i.e.*, ribonucleotides, deoxyribonucleotides, and/or a modified form of either of the foregoing). A "nucleic acid molecule" as used herein is synonymous with "nucleic acid" and "polynucleotide." The term includes both sense and anti-sense strands of RNA, cDNA, genomic DNA, and synthetic forms and mixed polymers thereof. The term includes any topological conformation, including single-stranded, double-stranded, partially duplexed, triplexed, hairpinned, circular, and padlocked conformations. A nucleic acid molecule can include either or both of naturally-occurring and modified nucleotides. Such nucleotides may be linked together by naturally-occurring and/or non-naturally-occurring nucleotide linkages.

[0057] Nucleic acid molecules may be modified chemically or biochemically, or may contain derivatized nucleotide bases, as will be readily appreciated by those of skill in the art. Such modifications include, for example and without limitation: labels; methylation; substitution of one or more of the naturally-occurring nucleotides with an analog; and inter-nucleotide modifications (*e.g.*, uncharged linkages, for example, methyl phosphonates, phosphotriesters, phosphoramidates, and carbamates; charged linkages, for example, phosphorothioates and phosphorodithioates; pendent moieties, for example, peptides; intercalators, for example, acridine and psoralen; chelators; alkylators; and modified linkages, for example, alpha anomeric nucleic acids).

[0058] Exogenous: An "exogenous" molecule is a molecule that is not native to a specified system (*e.g.*, a germplasm, variety, elite variety, and/or plant) with respect to nucleotide sequence and /or genomic location (*i.e.*, locus) for a polynucleotide (and with respect to amino acid sequence and/or cellular localization for a polypeptide). In embodiments, exogenous or heterologous polynucleotides or polypeptides may be molecules that have been artificially supplied to a biological system (*e.g.*, a plant cell, a plant gene, a particular plant species or variety, and/or a plant chromosome) and are not native to that particular biological

system. Thus, the designation of a nucleic acid as "exogenous" may indicate that the nucleic acid originated from a source other than a naturally-occurring source, or it may indicate that the nucleic acid has a non-natural configuration, genetic location, or arrangement of elements.

[0059] In contrast, for example, a "native" or "endogenous" nucleic acid is a nucleic acid (*e.g.*, a gene) that does not contain a nucleic acid element other than those normally present in the chromosome or other genetic material on which the nucleic acid is normally found in nature. An endogenous gene transcript is encoded by a nucleotide sequence at its natural chromosomal locus, and is not artificially supplied to the cell.

[0060] Operably linked: A first nucleic acid sequence is operably linked with a second nucleic acid sequence when the first nucleic acid sequence is in a functional relationship with the second nucleic acid sequence. For instance, a promoter is operably linked with a coding sequence when the promoter affects the transcription or expression of the coding sequence. When recombinantly produced, operably linked nucleic acid sequences are generally contiguous and, where necessary to join two protein-coding regions, in the same reading frame. However, elements need not be contiguous to be operably linked.

[0061] Promoter: A promoter is a region of DNA that generally is located upstream (towards the 5' region) of a nucleic acid that enhances transcription of the nucleic acid. Promoters permit the proper activation or repression of the nucleic acid(s) with which they are operably linked. A promoter contains specific sequences that are recognized by transcription factors. These factors bind to the promoter DNA sequences and result in the recruitment of RNA polymerase, the enzyme that synthesizes the RNA from the coding region of the nucleic acid. Transformed: A vector "transforms" or "transduces" a cell when it transfers nucleic acid molecules into the cell. A cell is "transformed" by a nucleic acid molecule when the nucleic acid molecule becomes stably replicated by the cell, either by incorporation of the nucleic acid molecule into the cellular genome or by episomal replication. As used herein, the term "transformation" encompasses all techniques by which a nucleic acid molecule can be introduced into a cell. Examples include, but are not limited to: transfection with viral vectors; transformation with plasmid vectors; electroporation (Fromm *et al.* (1986) Nature 319:791-3); lipofection (Felgner *et al.* (1987) Proc. Natl. Acad. Sci. USA 84:7413-7); microinjection (Mueller *et al.* (1978) Cell 15:579-85); *Agrobacterium*-mediated transfer (Fraley *et al.* (1983) Proc. Natl. Acad. Sci. USA 80:4803-7); direct DNA uptake; and microprojectile bombardment (Klein *et al.* (1987) Nature 327:70).

[0062] Introduced: As used herein, the term "introduced," when referring to translocation of an exogenous nucleic acid into a cell, refers to the incorporation of the nucleic acid into the cell using any methodology available in the art. This term encompasses nucleic acid introduction methods including, for example and without limitation, transfection; transformation; and transduction.

[0063] Transgene: As used herein, the term "transgene" refers to an exogenous nucleic acid coding sequence of interest. For example, a transgene may encode an industrially or pharmaceutically useful compound, or an expression product that contributes to a desirable agricultural trait (*e.g.*, herbicide resistance or pest resistance). In a further example, a transgene may be an antisense nucleic acid, wherein expression of the antisense nucleic acid inhibits expression of a target nucleic acid sequence. A transgene may comprise regulatory sequences operably linked to the transgene (*e.g.*, a promoter). In some embodiments, a nucleic acid molecule of interest to be introduced by site-specific targeting at a *FAD3* locus is a transgene. However, in other embodiments, a nucleic acid molecule of interest may be a PTU, an ELP, an ETIP, or an endogenous nucleic acid sequence (*e.g.*, wherein additional, exogenous genomic copies of the endogenous nucleic acid sequence are desired).

[0064] Elements can also include DNA that encodes for a structural RNA, such as shRNA. Such RNA can modify exogenous or endogenous genes including, but not limited to affecting postings or conferring herbicide resistance.

[0065] Recombinant: As used herein, the term "recombinant" refers to a material (*e.g.*, nucleic acid, gene, polynucleotide, and/or polypeptide) that has been altered by human intervention. For example, the arrangement of the parts or elements of a recombinant molecule may not be a native arrangement, and/or the primary sequence of the recombinant molecule may have been changed from its native sequence, *e.g.*, to optimize its expression and/or activity. A material may be altered to produce a recombinant material within or removed from its natural environment or state. As one example, an open reading frame of a nucleic acid is recombinant if the nucleotide sequence of the open reading frame has been removed from it natural context and cloned into an artificial nucleic acid molecule (*e.g.*, a vector). Protocols and reagents to produce recombinant molecules (*e.g.*, recombinant nucleic acids) are common in the art, and their use is routine. The term "recombinant" may also refer herein to a cell or organism that comprises recombinant material (*e.g.*, a plant and/or plant cell that comprises a recombinant nucleic acid). In some examples, a recombinant organism is a transgenic organism.

[0066] Vector: As used herein, the term "vector" refers to a polynucleotide or other molecule that is capable of transferring at least one nucleic acid segment(s) into a cell. A vector may optionally comprise components/elements that mediate vector maintenance and/or enable its intended use (*e.g.*, sequences necessary for replication, genes imparting drug or antibiotic resistance, a multiple cloning site, and/or operably linked promoter/enhancer elements that enable the expression of a cloned gene). Vectors may be derived, for example, from plasmids, bacteriophages, or plant or animal viruses. A "cloning vector," "shuttle vector," or "subcloning vector" generally comprises operably linked elements to facilitate cloning or subcloning steps (*e.g.*, a multiple cloning site containing multiple restriction endonuclease sites).

[0067] Expression Vector: The term "expression vector," as used herein, refers to a vector comprising operably linked polynucleotide sequences that may facilitate expression of a coding sequence in a particular host organism. For example, a bacterial expression vector may facilitate expression of a coding sequence in a bacterium. Likewise, a plant expression vector may facilitate expression of a coding sequence in a plant cell. Polynucleotide sequences that facilitate expression in prokaryotes may include, for example and without limitation, a promoter; an operator; and a ribosome binding site. Eukaryotic expression vectors (*e.g.*, a plant expression vector) may comprise, for example, promoters; enhancers; termination signals; and polyadenylation signals (and other sequences) that are generally different from those used in prokaryotic expression vectors.

[0068] Sequence identity: The term "sequence identity" or "identity," as used herein in the context of two nucleic acid or polypeptide sequences, refers to the residues in the two sequences that are the same when aligned for maximum correspondence over a specified comparison window. A value of sequence identity may be determined by comparing two optimally aligned sequences (*e.g.*, nucleic acid sequences and amino acid sequences) over a comparison window, wherein the portion of the sequence in the comparison window may comprise additions or deletions (*i.e.*, gaps) as compared to the reference sequence (which does not comprise additions or deletions) for optimal alignment of the two sequences. The sequence identity is calculated as a percentage by determining the number of positions at which the identical nucleotide or amino acid residue occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the comparison window, and multiplying the result by 100 to yield the percentage of sequence identity.

[0069] Methods for aligning sequences for comparison are well-known in the art. Various programs and alignment algorithms are described in, for example: Smith and Waterman (1981) Adv. Appl. Math. 2:482; Needleman and Wunsch (1970) J. Mol. Biol. 48:443; Pearson and Lipman (1988) Proc. Natl. Acad. Sci. U.S.A. 85:2444; Higgins and Sharp (1988) Gene 73:237-44; Higgins and Sharp (1989) CABIOS 5:151-3; Corpet *et al.* (1988) Nucleic Acids Res. 16:10881-90; Huang *et al.* (1992) Comp. Appl. Biosci. 8:155-65; Pearson *et al.* (1994) Methods Mol. Biol. 24:307-31; Tatiana *et al.* (1999) FEMS Microbiol. Lett. 174:247-50. A detailed consideration of sequence alignment methods and homology calculations can be found in Altschul *et al.* (1990) J. Mol. Biol. 215:403-10.

[0070] The National Center for Biotechnology Information (NCBI) Basic Local Alignment Search Tool (BLAST™; Altschul *et al.* (1990)) may be used to align sequences, and it is available from several sources, including the National Center for Biotechnology Information (Bethesda, MD), and on the internet, for use in connection with several sequence analysis programs. A description of how to determine sequence identity using this program is available on the internet under the "help" section for BLAST™. For comparisons of nucleic acid sequences, the "Blast 2 sequences" function of the BLAST™ (Blastn) program may be employed using the default parameters. Nucleic acid sequences with greater similarity to the reference sequences will show increasing percentage identity when assessed by this method.

[0071] As used herein, the term "substantially identical" may refer to nucleotide sequences that are more than 80% identical. For example, a substantially identical nucleotide sequence may be at least 85%, at least 86%; at least 87%; at least 88%; at least 89%; at least 90%; at least 91%; at least 92%; at least 93%; at least 94%; at least 95%; at least 96%; at least 97%; at least 98%; at least 99%; or at least 99.5% identical to the reference sequence.

[0072] Locus: As used herein, the term "locus" refers to a position on a genome that corresponds to a measurable characteristic (*e.g.*, a trait). In some embodiments, a locus of particular interest is the genomic position of a FAD3 gene, where disruption of the gene reduces or eliminates expression of the mRNA transcribed from the wild-type gene. A locus may be defined by a probe that hybridizes to a unique nucleotide sequence contained within the locus either during Southern hybridization or PCR.

[0073] Marker: As used herein, a "marker" refers to a gene or nucleotide sequence that can be used to identify plants that are likely to have a particular allele and/or exhibit a particular trait or phenotype. A marker may be described as a variation at a given genomic locus. A genetic marker may be a short DNA sequence, such as a sequence surrounding a single base-pair

change (single nucleotide polymorphism, or "SNP"), or a long sequence, for example, a minisatellite/simple sequence repeat ("SSR"). A "marker allele" refers to the version of the marker that is present in a particular plant. The term marker as used herein may refer to a cloned segment of plant chromosomal DNA (*e.g.*, a segment comprising a FAD3 locus, or a modified and/or disrupted *FAD3* locus), and may also or alternatively refer to a DNA molecule that is complementary to a cloned segment of plant chromosomal DNA. As is recognized by those of ordinary skill in the art, the process of obtaining additional, contiguous nucleotide sequence for inclusion in a marker may be repeated nearly indefinitely (limited only by the length of the chromosome), thereby identifying additional markers along the chromosome. Any and all of the above-described varieties of markers may be used in some embodiments of the present invention.

[0074] In some embodiments, the presence of a transgene or marker (which are characterized by a "target" sequence) in a germplasm may be detected through the use of a nucleic acid probe; *e.g.*, an oligonucleotide. A probe may be a DNA molecule or an RNA molecule. An oligonucleotide probe may be prepared synthetically or by cloning. Suitable cloning vectors are well-known to those of skill in the art. RNA probes can be synthesized by means known in the art, for example, using a DNA molecule template.

[0075] An oligonucleotide probe may be labeled or unlabeled. A wide variety of techniques exist for labeling nucleic acid molecules, including, for example and without limitation, radiolabeling by nick translation; random priming; and tailing with terminal deoxytransferase, where the nucleotides employed are labeled, for example, with radioactive ^{32}P. Other labels which may be used include, for example and without limitation, fluorophores; enzymes; enzyme substrates; enzyme cofactors; and enzyme inhibitors. Alternatively, the use of a label that provides a detectable signal, by itself or in conjunction with other reactive agents, may be replaced by ligands to which receptors bind, where the receptors are labeled (for example, by the above-indicated labels) to provide detectable signals, either by themselves, or in conjunction with other reagents. *See*, *e.g.*, Leary *et al.* (1983) Proc. Natl. Acad. Sci. USA 80:4045-9.

[0076] A probe may be an exact copy of a transgene or marker to be detected. A probe may also be a nucleic acid molecule comprising, or consisting of, a nucleotide sequence that is substantially identical to a cloned segment of chromosomal DNA comprising the transgene or marker to be detected. A probe may further comprise additional nucleic acid sequences, for example, promoters; transcription signals; and/or vector sequences.

[0077] A probe may contain all or a portion of the target nucleotide sequence and additional, contiguous nucleotide sequence from the genome. This is referred to herein as a "contiguous probe." The additional, contiguous nucleotide sequence is referred to as "upstream" or "downstream" of the original target, depending on whether the contiguous nucleotide sequence from the chromosome is on the 5' or the 3' side of the original marker, as conventionally understood. A probe may also contain a nucleotide sequence that is not contiguous to that of the original target; this probe is referred to herein as a "non-contiguous probe." The sequence of the non-contiguous probe may be located sufficiently close to the sequence of the original target on the chromosome so that the non-contiguous probe is linked to the original marker or transgene.

[0078] In some embodiments, a probe is a nucleic acid molecule that is "specifically hybridizable" or "specifically complementary" to an exact copy of the target to be detected. "Specifically hybridizable" and "specifically complementary" are terms that indicate a sufficient degree of complementarity, such that stable and specific binding occurs between the nucleic acid molecule and the target. A nucleic acid molecule need not be 100% complementary to its target sequence to be specifically hybridizable. A nucleic acid molecule is specifically hybridizable when there is a sufficient degree of complementarity to avoid non-specific binding of the nucleic acid to non-target sequences under conditions where specific binding is desired, for example, under stringent hybridization conditions.

[0079] Hybridization conditions resulting in particular degrees of stringency will vary depending upon the nature of the hybridization method of choice and the composition and length of the hybridizing nucleic acid sequences. Generally, the temperature of hybridization and the ionic strength (especially the Na^+ and/or Mg^{++} concentration) of the hybridization buffer will determine the stringency of hybridization, though wash times also influence stringency. Calculations regarding hybridization conditions required for attaining particular degrees of stringency are known to those of ordinary skill in the art, and are discussed, for example, in Sambrook *et al.* (ed.) Molecular Cloning: A Laboratory Manual, 2nd ed., vol. 1-3, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989, chapters 9 and 11; and Hames and Higgins (eds.) Nucleic Acid Hybridization, IRL Press, Oxford, 1985. Further detailed instruction and guidance with regard to the hybridization of nucleic acids may be found, for example, in Tijssen, "Overview of principles of hybridization and the strategy of nucleic acid probe assays," in Laboratory Techniques in Biochemistry and Molecular Biology- Hybridization with Nucleic Acid Probes, Part I, Chapter 2, Elsevier, NY, 1993; and Ausubel *et al.*, Eds.,

Current Protocols in Molecular Biology, Chapter 2, Greene Publishing and Wiley-Interscience, NY, 1995.

[0080] As used herein, "stringent conditions" encompass conditions under which hybridization will only occur if there is less than 25% mismatch between the hybridization molecule and the DNA target. "Stringent conditions" include further particular levels of stringency. Thus, as used herein, "moderate stringency" conditions are those under which molecules with more than 25% sequence mismatch will not hybridize; conditions of "medium stringency" are those under which molecules with more than 15% mismatch will not hybridize; and conditions of "high stringency" are those under which sequences with more than 10% mismatch will not hybridize. Conditions of "very high stringency" are those under which sequences with more than 6% mismatch will not hybridize.

[0081] In particular embodiments, stringent conditions are hybridization at 65°C in 6x saline-sodium citrate (SSC) buffer, 5x Denhardt's solution, 0.5% SDS, and 100 μg sheared salmon testes DNA, followed by 15-30 minute sequential washes at 65°C in 2x SSC buffer and 0.5% SDS, followed by 1x SSC buffer and 0.5% SDS, and finally 0.2x SSC buffer and 0.5% SDS.

[0082] Linkage (dis)equilibrium: As used herein, the term "linkage equilibrium" refers to the situation where a marker and a second nucleic acid (*e.g.*, transgene, PTU, and second marker) independently segregate; *i.e.*, the marker and the second nucleic acid sort randomly among progeny. Nucleic acids that show linkage equilibrium are considered unlinked (whether or not they lie on the same chromosome). As used herein, the term "linkage disequilibrium" refers to the situation where a marker and a second nucleic acid segregate in a non-random manner; *i.e.*, the nucleic acids have a recombination frequency of less than 50% (and thus by definition, are separated by less than 50 cM on the same linkage group). In some examples, nucleic acids that show linkage disequilibrium are considered linked.

[0083] Linked, tightly linked, and extremely tightly linked: As used herein, linkage between a marker and a second nucleic acid (*e.g.*, transgene, PTU, and second marker) may refer to the phenomenon in which nucleic acids on a chromosome show a measurable probability of being passed on together to individuals in the next generation. Thus, linkage of one marker to a second nucleic acid may be measured and/or expressed as a recombination frequency. The closer two nucleic acids are to each other, the closer to "1" this probability becomes. Thus, the term "linked" may refer to one or more genes or markers that are passed together with a second nucleic acid with a probability greater than 0.5 (which is expected from

independent assortment where markers/genes are located on different chromosomes). When the presence of a gene (*e.g.*, a transgene) contributes to a phenotype in an individual, markers that are linked to the gene may be said to be linked to the phenotype. Thus, the term "linked" may refer to a relationship between a marker and a gene, or between a marker and a phenotype.

[0084] A relative genetic distance (determined by crossing over frequencies and measured in centimorgans (cM)) is generally proportional to the physical distance (measured in base pairs) that two linked markers or genes are separated from each other on a chromosome. One centimorgan is defined as the distance between two genetic markers that show a 1% recombination frequency (*i.e.*, a crossing-over event occurs between the two markers once in every 100 cell divisions). In general, the closer one marker is to another marker or gene (whether the distance between them is measured in terms of genetic distance or physical distance,) the more tightly they are linked. Because chromosomal distance is approximately proportional to the frequency of recombination events between traits, there is an approximate physical distance that correlates with recombination frequency. This correlation is generally known or readily determinable across the major crop plants (Helentjaris and Burr (eds.) (1989) Development and Application of Molecular Markers to Problems in Plant Genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY; Gresshoff (ed.) (1994) Plant Genome Analysis. CRC Press, Boca Raton, FL; Lander *et al.* (1987) Genomics 1:174-81; Tanksley *et al.* (1988) "Molecular mapping of plant chromosomes," In Chromosome Structure and Function. Gustafson and Appels (eds.) Plenum Press, NY, pp. 157-73) and many other organisms. For example, 1 cM corresponds to about 2.5-3.0 kb in yeast, about 140 kb in *Arabidopsis*, about 400 kb in sunflower, and about 350 kb in *Eucalyptus*.

[0085] The term "linked" may refer herein to one or more nucleic acids that show a recombination frequency of less than 50% (*i.e.*, less than 50 cM). For example, "linked" nucleic acids may recombine with a frequency of about 45% or less, about 40% or less, about 35% or less, about 30% or less, about 25% or less, about 20% or less, about 15% or less, and about 10% or less. The physical distances between such nucleic acids on the same chromosome (nucleic acids on different chromosomes are expected to be in linkage equilibrium) that correspond to the foregoing recombination frequencies depend on the host genome, and may be easily calculated as set forth, *supra*.

[0086] As used herein, the term "tightly-linked" may refer to one or more nucleic acids that show a recombination frequency of about 20% or less (*i.e.*, about 20 cM or less). For example, "tightly linked" nucleic acids may recombine with a frequency of 22% or less, about

18% or less, about 16% or less, about 14% or less, about 12% or less, about 10% or less, about 8% or less, about 6% or less, about 4% or less, and about 2% or less.

[0087] As used herein, the term "extremely tightly-linked" may refer to one or more nucleic acids that show a recombination frequency of about 10% or less (*i.e.*, about 10 cM or less). For example, "extremely tightly linked" nucleic acids may recombine with a frequency of 11% or less, about 9% or less, about 8% or less, about 7% or less, about 6% or less, about 5% or less, about 4% or less, about 3% or less, about 2% or less, and about 1% or less.

[0088] The closer a particular nucleic acid is to a gene that encodes a polypeptide that contributes to a particular phenotype (whether measured in terms of genetic or physical distance), the more tightly-linked is the particular nucleic acid to the phenotype. In view of the foregoing, it will be appreciated that nucleic acids linked to a particular gene or phenotype include those nucleic acids that are tightly linked, and those nucleic acids that are extremely tightly linked, to the gene or phenotype. In some embodiments, the closer a particular nucleic acid is to a *FAD3* locus (*e.g.*, a modified or disrupted *FAD3* locus), whether measured in terms of genetic or physical distance, the more tightly-linked is the particular nucleic acid to any trait/phenotype conferred by an exogenous nucleic acid integrated at the *FAD3* locus (or to a wild-type *FAD3* phenotype in the case of an unmodified locus). Thus, genetic markers that are linked, tightly linked, and/or extremely tightly linked to a *FAD3* locus comprising an integrated exogenous nucleic acid may be useful in an MAS program to identify organisms (*e.g.*, plants and plant varieties) comprising the integrated nucleic acid, to identify organisms comprising a phenotype conferred by the integrated nucleic acid, and to breed such an integrated nucleic acid and/or a phenotype conferred by the integrated nucleic acid into other compatible organisms.

[0089] Marker-assisted breeding: As used herein, the term "marker-assisted breeding" may refer to an approach to breeding plants directly for one or more trait(s) (*e.g.*, a polygenic trait). In current practice, plant breeders attempt to identify easily detectable traits, such as flower color, seed coat appearance, or isozyme variants that are linked to an agronomically desired trait. The plant breeders then follow the agronomic trait in the segregating, breeding populations by following the segregation of the easily detectable trait. However, there are very few of these linkage relationships between traits of interest and easily detectable traits available for use in plant breeding. In some embodiments of the invention, marker-assisted breeding comprises identifying one or more genetic markers (*e.g.*, SNP, isozyme, and/or SSR markers) that are linked to a *FAD3* locus wherein an exogenous nucleic acid contributing to a trait of interest has been integrated, and following the trait of interest in a segregating, breeding

population by following the segregation of the one or more genetic markers. In some examples, the segregation of the one or more genetic markers may be determined utilizing a probe for the one or more genetic markers by assaying a genetic sample from a progeny plant for the presence of the one or more genetic markers. Marker-assisted breeding provides a time- and cost-efficient process for improvement of plant varieties.

[0090] Trait or phenotype: The terms "trait" and "phenotype" are used interchangeably herein. For the purposes of the present disclosure, traits of particular interest include agronomically important traits, as may be expressed, for example, in a crop plant, and the production of transgene expression products from a targeted integration event. The term "molecular phenotype" may refer to a phenotype that is detectable at the level of a population of (one or more) molecules. In some examples, the molecular phenotype may only be detectable at the molecular level. The detectable molecules of the phenotype may be nucleic acids (*e.g.*, genomic DNA or RNA); proteins; and/or metabolites. For example, a molecular phenotype may be an expression profile for one or more gene products (*e.g.*, at a specific stage of plant development, or in response to an environmental condition or stress).

[0091] Quantitative Trait Locus: Traits that are continuously varying due to genetic (additive, dominant, and epistatic) and environmental influences are commonly referred to as "quantitative traits." Quantitative traits may be distinguished from "qualitative," or "discrete," traits on the basis of two factors; environmental influences on gene expression that produce a continuous distribution of phenotypes, and the complex segregation pattern produced by multigenic inheritance. The identification of one or more regions of the genome linked to the expression of a quantitative trait defines such regions as Quantitative Trait Loci ("QTL").

[0092] Plant: As used herein, the term "plant" may refer to a whole plant, a cell or tissue culture derived from a plant, and/or any part of any of the foregoing. Thus, the term "plant" encompasses, for example and without limitation, whole plants; plant components and/or organs (*e.g.*, leaves, stems, and roots); plant tissue; seed; and a plant cell. A plant cell may be, for example and without limitation, a cell in and/or of a plant, a cell isolated from a plant, and a cell obtained through culturing of a cell isolated from a plant.

[0093] A "transgenic plant" is a plant comprising within at least one of its cells an exogenous polynucleotide. The term "transgenic" is used herein to refer to any cell, cell line, callus, tissue, plant part, or plant, the genotype of which has been altered by the presence of a exogenous nucleic acid. Thus, this term encompasses transgenic organisms and cells that have been initially altered to comprise the exogenous polynucleotide, and those organisms and cells

created by crosses or asexual propagation of the initial transgenic organism or cell. The term "transgenic," as used herein, does not encompass genome (chromosomal or extra-chromosomal) alternations introduced by conventional plant breeding methods (*e.g.*, crosses of only non-transgenic organisms) or by naturally-occurring events (*e.g.*, random cross-fertilization, non-recombinant viral infection, non-recombinant bacterial transformation, non-recombinant transposition, and spontaneous mutation).

[0094] A plant "line," "variety," or "strain" is a group of individual plants having the same parentage. Plants of a line generally are inbred to some degree, and are generally homozygous and homogeneous at most genetic loci (*e.g.*, a *FAD3* locus). A "subline" may refer to an inbred subset of descendents from a common progenitor that are genetically distinct from other similarly inbred subsets descended from the same progenitor. In some embodiments, a "subline" may be produced by inbreeding seed from an individual transgenic plant selected at the F_3 to F_5 generation until the residual segregating loci are homozygous across most or all loci.

[0095] A "binding protein" is a protein that is able to bind to another molecule. A binding protein can bind to, for example, a DNA molecule (a DNA-binding protein), an RNA molecule (an RNA-binding protein) and/or a protein molecule (a protein-binding protein). In the case of a protein-binding protein, it can bind to itself (to form homodimers, homotrimers, etc.) and/or it can bind to one or more molecules of a different protein or proteins. A binding protein can have more than one type of binding activity. For example, zinc finger proteins have DNA-binding, RNA-binding and protein-binding activity.

[0096] A "zinc finger DNA binding protein" (or binding domain) is a protein, or a domain within a larger protein, that binds DNA in a sequence-specific manner through one or more zinc fingers, which are regions of amino acid sequence within the binding domain whose structure is stabilized through coordination of a zinc ion. The term zinc finger DNA binding protein is often abbreviated as zinc finger protein or ZFP.

[0097] A "TALE DNA binding domain" or "TALE" is a polypeptide comprising one or more TALE repeat domains/units. The repeat domains are involved in binding of the TALE to its cognate target DNA sequence. A single "repeat unit" (also referred to as a "repeat") is typically 33-35 amino acids in length and exhibits at least some sequence homology with other TALE repeat sequences within a naturally occurring TALE protein.

[0098] Zinc finger and TALE binding domains can be "engineered" to bind to a predetermined nucleotide sequence, for example via engineering (altering one or more amino acids) of the recognition helix region of a naturally occurring zinc finger or TALE protein.

Therefore, engineered DNA binding proteins (zinc fingers or TALEs) are proteins that are non-naturally occurring. Non-limiting examples of methods for engineering DNA-binding proteins are design and selection. A designed DNA binding protein is a protein not occurring in nature whose design/composition results principally from rational criteria. Rational criteria for design include application of substitution rules and computerized algorithms for processing information in a database storing information of existing ZFP and/or TALE designs and binding data. See, for example, US Patents 6,140,081; 6,453,242; and 6,534,261; see also WO 98/53058; WO 98/53059; WO 98/53060; WO 02/016536 and WO 03/016496 and U.S. Publication No. 20110301073.

[0099] A "selected" zinc finger protein or TALE is a protein not found in nature whose production results primarily from an empirical process such as phage display, interaction trap or hybrid selection. See e.g., US 5,789,538; US 5,925,523; US 6,007,988; US 6,013,453; US 6,200,759; WO 95/19431; WO 96/06166; WO 98/53057; WO 98/54311; WO 00/27878; WO 01/60970 WO 01/88197, WO 02/099084 and U.S. Publication No. 20110301073.

[0100] "Cleavage" refers to the breakage of the covalent backbone of a DNA molecule. Cleavage can be initiated by a variety of methods including, but not limited to, enzymatic or chemical hydrolysis of a phosphodiester bond. Both single-stranded cleavage and double-stranded cleavage are possible, and double-stranded cleavage can occur as a result of two distinct single-stranded cleavage events. DNA cleavage can result in the production of either blunt ends or staggered ends. In certain embodiments, fusion polypeptides are used for targeted double-stranded DNA cleavage.

[0101] A "cleavage half-domain" is a polypeptide sequence which, in conjunction with a second polypeptide (either identical or different) forms a complex having cleavage activity (preferably double-strand cleavage activity). The terms "first and second cleavage half-domains;" "+ and – cleavage half-domains" and "right and left cleavage half-domains" are used interchangeably to refer to pairs of cleavage half-domains that dimerize.

[0102] An "engineered cleavage half-domain" is a cleavage half-domain that has been modified so as to form obligate heterodimers with another cleavage half-domain (e.g., another engineered cleavage half-domain). See, also, U.S. Patent Publication Nos. 2005/0064474, 20070218528, 2008/0131962 and 2011/0201055.

[0103] Means for generating a double strand DNA break: As used herein, the term "means for generating a double strand DNA break" is intended to invoke the special claiming

provisions authorized by Congress in 35 U.S.C. § 112, sixth paragraph. Specifically, a "means for generating a double strand DNA break" refers to a molecular structure that is capable of cleaving both strands of a double-stranded DNA molecule. Such structures include polypeptide domains comprised within many known nuclease proteins, for example, the FokI nuclease domain, the catalytic domain is selected from the group consisting of proteins Mmel, Colicin-E7 (CEA7_ECOLX), Colicin-E9, APFL, EndA, Endo I (END1_EC0LI), Human Endo G (NUCG_HUMAN), Bovine Endo G (NUCG_BOVIN), R.HinPII, I-BasI, I-BmoI, I-HmuI, I-TevI, I-TevII, I-TevIII, I-TwoI, R.MspI, R.MvaI, NucA, NucM, Vvn, Vvn_CLS, Staphylococcal nuclease (NUC_STAAU), Staphylococcal nuclease (NUC_STAHY), Micrococcal nuclease (NUC_SHIFL), Endonuclease yncB, Endodeoxyribonuclease I (ENRN_BPT7), Metnase, Nb.BsrDI, BsrDI A, Nt.BspD6I (R.BspD6I large subunit), ss.BspD6I (R.BspD6I small subunit), R.PIeI, MlyI, AlwI, Mval269I, BsrI, BsmI, Nb.BtsCI, Nt.BtsCI, RI.BtsI, R2.BtsI, BbvCI subunit 1, BbvCI subunit 2, BpulOI alpha subunit, BpulOI beta subunit, BmrI, BfiI, I-CreI, hExoI (EX01JHUMAN), Yeast ExoI (EX01_YEAST), E.coli ExoI, Human TREX2, Mouse TREX1, Human TREX1, Bovine TREX1, Rat TREX1, Human DNA2, Yeast DNA2 (DNA2_YEAST).

[0104] Means for repairing a double strand DNA break: As used herein, the term "means for repairing a double strand DNA break" is also intended to invoke the special claiming provisions authorized by Congress in 35 U.S.C. § 112, sixth paragraph. Specifically, a "means for repairing a double strand DNA break" refers to a molecular structure that is capable of facilitating/catalyzing the joining of the ends of double-stranded DNA molecules, for example, by joining ends generated by cleaving a single double-stranded DNA molecule, or by joining one end generated by cleaving a single double-stranded DNA molecule with the end of an exogenous double-stranded DNA molecule. Such structures include polypeptide domains comprised within many known ligase proteins, for example, Cre recombinase. In some examples, the same molecular structure may serve as both a means for generating a double strand DNA break and a means for repairing a double strand DNA break, where the same structure facilitates both the cleavage and repair of double-stranded DNA molecules (*e.g.*, Hin recombinase).

[0105] The induction of the site specific double stranded breaks in the genome induces the host plant cell DNA repair pathway which resolves the double stranded break through homology-directed repair (HDR) or non-homologous end-joining (NHEJ) repair. In plants, the scientific literature reports that precise gene or donor DNA integration into native genomic or at pre-engineered locations have involved incoming donor DNA construct(s) that comprise varying

amounts of sequence homologous to the sequences flanking the targeted double stranded break. The integration of such donors into the specific target locus presumably has relied on the HDR pathway. Exclusively relying on the HDR approach for gene targeting in plants can have limitations due to reports that the HDR repair pathway is not the dominate DNA repair pathway when compared to NHEJ. The published plant scientific literature utilizing target specific DNA breaks (ZFN, TALeNs, or Engineered Meganucleases, etc.) the NHEJ pathway has been reported as the method to introduce specific point mutations (insertions, or deletions) into the geneome. Here we report that site specific double stranded breaks (induced by ZFN, TALeNs, etc.) in the presents of various donor DNA design with homology regions of 0 to <10 bp can be specifically inserted at targeted break via the NHEJ repair pathway in plants. A variety of different DNA donor designs with zero homology to small 1 – 10 bp of ranging from linear to circular, single stranded to double stranded can be targeted to specific locations using the NHEJ pathway. NHEJ based donor DNA plant genome targeting can be based on "sticky end capture", where the targeted double stranded break in the genome generated by *Fok*1 (or other Type II endonuclease domains) and the corresponding sticky ends are on the NHEJ donor DNA designs. The sticky ends donor DNA can be delivered directly to the cell as linear donor DNA with predefined overhangs. An alternative approach is to produce the donor DNA sticky ends *in vivo* by co-delivering the host target ZFN and a circular DNA donor molecule that contains at least one ZFN recognition site that is identical to the target recognition site. Expression of at least one ZFN cuts the host genomic DNA (native or pre-engineered) and the circular donor DNA to produce sticky ends that are resolved using the hosts NHEJ repair pathway.

[0106] It is possible to have one or more ZFN cuts sites on the donor molecule (a single ZFN cut site to linearize the entire donor molecule, 2 of the same ZFN sites to release a smaller donor DNA fragment or 2 different ZFN sites to release a fragment from the donor and a corresponding fragment from the host genomic DNA (DNA replacement).

[0107] Thus, the donor polynucleotide can be DNA or RNA, single-stranded and/or double-stranded and can be introduced into a cell in linear or circular form. *See, e.g.,* U.S. Patent Publication Nos. 20100047805 and 20110207221. In certain, embodiments of the present invention may also include linear exogenous (donor) nucleic acid(s), compositions comprising these nucleic acids and methods of making and using these linear donor molecules. In certain embodiments, the linear donor molecule stably persists in the cell into which it is introduced. In other embodiments, the linear donor molecule is modified to resist exonucleolytic cleavage, for example by placing one or more phosphorothioate phosphodiester bonds between one or more

base pairs on the ends of the donor molecule. The linear exogenous nucleic acid may also include single stranded specific DNA.

IV. FAD3 Performance Loci

[0108] The loci designated FAD3 (fatty acid desaturase 3) are included in QTLs involved in the inheritance of the complex multigenic trait of fatty acid content in plants. FAD3 encodes the enzyme responsible for the desaturation of linoleic acid (18:2) to linolenic acid (C18:3). Tanhuanpaa *et al.* (1998) Mol. Breed. 4:543-50; Schierholt *et al.* (2001) Crop Sci. 41:1444-9.

[0109] Within the plant oil biosynthetic pathway the fatty acid desaturases (FADs) play a key role in plant lipid biosynthesis and their activity significantly influences the fatty acid composition. FADs are abundant in plants, and expression analysis suggested that FAD mRNAs are produced in over-abundance. Furthermore, FAD genes are expressed in various, tissues, and cell types, as well as subcellular compartments including the plastid and endoplasmic reticulum.

[0110] The fatty acid composition of plants, and the performance of oils produced therefrom in many applications, is determined by the relative concentrations of the major fatty acid constituents; oleic, linoleic, and linolenic (C18:3). The concentrations of these fatty acids are predominantly regulated by the function of the enzymes FAD2 and FAD3. Oleic acid is converted to linoleic acid and linolenic acid in plants according to the scheme:

$$\text{C18:0} \rightarrow \text{C18:1} \xrightarrow{\text{FAD2}} \text{C18:2} \xrightarrow{\text{FAD3}} \text{C18:3}$$

[0111] *FAD3* genes have been identified in major plant and algal species including but not limited to maize, soybean, cotton, *Arabidopsis*, wheat, forage grasses, rice, sunflower and *Brassica*, and modification of *FAD3* expression leads to altered fatty acid profiles in such organisms. Furthermore, plants comprising modified *FAD3* genes have been commercialized, and disruption of a *FAD3* gene has been shown to be able to improve the nutritional and functional properties of oil produced by a host plant without an agronomic penalty to the host plant. For example, canola and sunflower varieties that have been commercialized under the Nexera® brand (Dow AgroSciences, LLC) are characterized by a higher oleic acid, lower linoleic aced, and lower linolenic acid (and lower saturated fatty acid) composition, when compared to wild-type canola and sunflower profiles.The dominant canola species grown in

Europe, North America, and Australia is *Brassica napus*, a polyploid *Brassica* species considered to have arisen from the hybridization of *B. oleracea* (having a diploid C genome) and *B. rapa* (having a diploid A genome). Cytogenetic investigation revealed the AA and CC genomes show a degree of relatedness as being partially homologous to one another. Both the A and C genomes contain a high percentage of homeologous and/or paralogous genes. Thus, it is thought that the AA and CC genomes are derived from a common ancestor genome. Prakash and Hinata (1980) Opera Botanica 55:1-57. Although the genomes of both progenitor species are technically classified as diploids, these genomes contain a high percentage of regions that are duplicative of one another. Song *et al.* (1991) Theor. Appl. Genet. 82:296-304. A detailed organelle and nuclear RFLP analysis revealed that the AA genome of *B. rapa* contributed ten chromosomes to *B. napus*, while *B. oleracea* contributed nine chromosomes from its CC genome as the maternal donor. Song *et al.* (1992) Genome 35:992-1001. Through the number of genome duplications in both ancestral genomes, as well as the high percentage of similarity between the A, B and C genomes, there have arisen several copies of *FAD2* and *FAD3* genes. As a practical matter, this fact makes breeding canola with modified and/or disrupted copies of these genes challenging in order to produce a particular fatty acid profile.

[0112] All of the known functional gene copies of *FAD3* in canola are located on linkage group N4 of the A genome. Scheffler *et al.* (1997) TAG 94(5):583-91; Schierholt *et al.* (2000) TAG 101(5-6):897-901. More recently, a high oleic trait in canola has been associated with a modified and disrupted FAD3 gene that is located on the A genome. U.S. Patent Application Publication No. US 2006/0248611 A1; Hu *et al.* (2006) "Identification and Mapping of FAD2 and FAD3 Mutations and Development of Allele-specific Markers for High Oleic and Low Linolenic Acid Contents in Canola (*Brassica napus* L.)," Plant & Animal Genomes XIV Conference, January 14-18, 2006, San Diego, CA. An inactivating FAD3 allele contributes to the control of oleic acid content by reducing the desaturation of linoleic acid to linolenic acid. This high oleic acid and FAD3 trait was identified in a B. napus variety (DMS100) that has a characteristic oleic acid content of about 77%. See, U.S. Publication No. 20060248611. Further, genetic markers have been developed to assist the introgression of the *Fad3* and high oleic acid trait into canola.

[0113] FAD3 loci may be modified and/or disrupted in a plant without detrimentally affecting the value of the plant, and for many purposes, with an actual increase in its value, including alteration of FAD3 expression, alteration of oil content/ratios and/or integration and expression of desired transgenes. Furthermore, according to the ubiquitous nature of FAD loci

in plants, FAD3 loci may be modified and/or disrupted without detriment for at least some purposes in many species, including, for example and without limitation: canola; soybean; maize; wheat; forage grasses; *brassica* sp.; rice, tomatoes, barley; oats; sorghum; cotton; and sunflower, as well as fungi and algae. Embodiments of the invention include FAD3 loci, and the use thereof as performance loci for integration of exogenous nucleic acids. In examples, a FAD3 locus exhibits at least one of several features that have been found to be desirable within the context of its use as a performance locus, including, for example and without limitation: that there is an approximately consistent level of expression during the life cycle of the host organism; and surprisingly, that insertion of donor DNA at a FAD3 locus does not induce a quality or fitness penalty on the host.

[0114] In some embodiments of the present invention, at least one FAD3 locus (e.g., a FAD3A and/or FAD3C locus) is used as a target site for the site-specific integration of an exogenous nucleic acid (e.g., a nucleic acid comprising a nucleotide sequence encoding a polypeptide of interest). In particular embodiments, integration of the exogenous nucleic acid results in a modified locus. For example, integration of the exogenous nucleic acid may modify the locus so as to produce a disrupted (*i.e.*, inactivated) FAD3 gene.

[0115] In some embodiments, a FAD3 locus may comprise a nucleotide sequence that is specifically hybridizable to the complement of a nucleotide sequence selected from the group consisting of SEQ ID NOs: 20-23, SEQ ID NOs: 25-38, SEQ ID NOs: 40-45, SEQ ID NO: 47 and SEQ ID NO: 49. For example, a *FAD3* locus may comprise a nucleotide sequence selected from the group consisting of SEQ ID NOs: 20-23, SEQ ID NOs: 25-38, SEQ ID NOs: 40-45, SEQ ID NO: 47 and SEQ ID NO: 49. In some embodiments, a *FAD3* locus may comprise a nucleotide sequence that is substantially identical to a nucleotide sequence selected from the group consisting of SEQ ID NOs: 20-23, SEQ ID NOs: 25-38, SEQ ID NOs: 40-45, SEQ ID NO: 47 and SEQ ID NO: 49. For example, in some embodiments, a *FAD3* locus is a *FAD3* homologue (*e.g.*, an ortholog or a paralog) that comprises a nucleotide sequence that is at least about 85% identical to a nucleotide sequence selected from the group consisting of SEQ ID NOs: 20-23, SEQ ID NOs: 25-38, SEQ ID NOs: 40-45, SEQ ID NO: 47 and SEQ ID NO: 49. A *FAD3* homologue may comprise a nucleotide sequence that is, for example and without limitation: at least 80%; at least 85%; at least about 90%; at least about 91%; at least about 92%; at least about 93%; at least about 94%; at least about 95%; at least about 96%; at least about 97%; at least about 98%; at least about 99%; at least about 99.5%; 99.6%, 99.7%, 99.8% and/or at least about 99.9% identical to a nucleotide sequence selected from the group consisting of

SEQ ID NOs: 20-23, SEQ ID NOs: 25-38, SEQ ID NOs: 40-45, SEQ ID NO: 47 and SEQ ID NO: 49. Such a *FAD3* homologue may be readily identified and isolated from any complete or partial genome readily available to those of skill in the art for a variety of organisms.

IV. Targeted integration of a nucleic acid at a FAD3 locus

[0116] Site-specific integration of an exogenous nucleic acid at a *FAD3* locus may be accomplished by any technique known to those of skill in the art. In some embodiments, integration of an exogenous nucleic acid at a FAD3 locus comprises contacting a cell (e.g., an isolated cell or a cell in a tissue or organism) with a nucleic acid molecule comprising the exogenous nucleic acid. In examples, such a nucleic acid molecule may comprise nucleotide sequences flanking the exogenous nucleic acid that facilitate homologous recombination between the nucleic acid molecule and at least one FAD3 locus. In particular examples, the nucleotide sequences flanking the exogenous nucleic acid that facilitate homologous recombination may be complementary to endogenous nucleotides of the *FAD3* locus. In particular examples, the nucleotide sequences flanking the exogenous nucleic acid that facilitate homologous recombination may be complementary to previously integrated exogenous nucleotides. In some embodiments, a plurality of exogenous nucleic acids may be integrated at one *FAD3* locus, such as in gene stacking.

[0117] Integration of a nucleic acid at a *FAD3* locus may be facilitated (*e.g.*, catalyzed) in some embodiments by endogenous cellular machinery of a host cell, such as, for example and without limitation, endogenous DNA and endogenous recombinase enzymes. In some embodiments, integration of a nucleic acid at a *FAD3* locus may be facilitated by one or more factors (e.g., polypeptides) that are provided to a host cell. For example, nuclease(s), recombinase(s), and/or ligase polypeptides may be provided (either independently or as part of a chimeric polypeptide) by contacting the polypeptides with the host cell, or by expressing the polypeptides within the host cell. Accordingly, in some examples, a nucleic acid comprising a nucleotide sequence encoding at least one nuclease, recombinase, and/or ligase polypeptide may be introduced into the host cell, either concurrently or sequentially with a nucleic acid to be integrated site-specifically at a *FAD3* locus, wherein the at least one nuclease, recombinase, and/or ligase polypeptide is expressed from the nucleotide sequence in the host cell.

A. DNA-binding polypeptides

[0118] In some embodiments, site-specific integration may be accomplished by utilizing factors that are capable of recognizing and binding to particular nucleotide sequences, for example, in the genome of a host organism. For instance, many proteins comprise polypeptide domains that are capable of recognizing and binding to DNA in a site-specific manner. A DNA sequence that is recognized by a DNA-binding polypeptide may be referred to as a "target" sequence. Polypeptide domains that are capable of recognizing and binding to DNA in a site-specific manner generally fold correctly and function independently to bind DNA in a site-specific manner, even when expressed in a polypeptide other than the protein from which the domain was originally isolated. Similarly, target sequences for recognition and binding by DNA-binding polypeptides are generally able to be recognized and bound by such polypeptides, even when present in large DNA structures (*e.g.*, a chromosome), particularly when the site where the target sequence is located is one known to be accessible to soluble cellular proteins (*e.g.*, a gene).

[0119] While DNA-binding polypeptides identified from proteins that exist in nature typically bind to a discrete nucleotide sequence or motif (*e.g.*, a consensus recognition sequence), methods exist and are known in the art for modifying many such DNA-binding polypeptides to recognize a different nucleotide sequence or motif. DNA-binding polypeptides include, for example and without limitation: zinc finger DNA-binding domains; leucine zippers; UPA DNA-binding domains; GAL4; TAL; LexA; a Tet repressor; LacR; and a steroid hormone receptor.

[0120] In some examples, a DNA-binding polypeptide is a zinc finger. Individual zinc finger motifs can be designed to target and bind specifically to any of a large range of DNA sites. Canonical Cys_2His_2 (as well as non-canonical Cys_3His) zinc finger polypeptides bind DNA by inserting an α-helix into the major groove of the target DNA double helix. Recognition of DNA by a zinc finger is modular; each finger contacts primarily three consecutive base pairs in the target, and a few key residues in the polypeptide mediate recognition. By including multiple zinc finger DNA-binding domains in a targeting endonuclease, the DNA-binding specificity of the targeting endonuclease may be further increased (and hence the specificity of any gene regulatory effects conferred thereby may also be increased). *See, e.g.*, Urnov *et al.* (2005) Nature 435:646-51. Thus, one or more zinc finger DNA-binding polypeptides may be engineered and utilized such that a targeting endonuclease introduced into a host cell interacts with a DNA sequence that is unique within the genome of the host cell.

[0121] Preferably, the zinc finger protein is non-naturally occurring in that it is engineered to bind to a target site of choice. See, for example, See, for example, Beerli et al. (2002) Nature Biotechnol. 20:135-141; Pabo et al. (2001) Ann. Rev. Biochem. 70:313-340; Isalan et al. (2001) Nature Biotechnol. 19:656-660; Segal et al. (2001) Curr. Opin. Biotechnol. 12:632-637; Choo et al. (2000) Curr. Opin. Struct. Biol. 10:411-416; U.S. Patent Nos. 6,453,242; 6,534,261; 6,599,692; 6,503,717; 6,689,558; 7,030,215; 6,794,136; 7,067,317; 7,262,054; 7,070,934; 7,361,635; 7,253,273; and U.S. Patent Publication Nos. 2005/0064474; 2007/0218528; 2005/0267061.

[0122] An engineered zinc finger binding domain can have a novel binding specificity, compared to a naturally-occurring zinc finger protein. Engineering methods include, but are not limited to, rational design and various types of selection. Rational design includes, for example, using databases comprising triplet (or quadruplet) nucleotide sequences and individual zinc finger amino acid sequences, in which each triplet or quadruplet nucleotide sequence is associated with one or more amino acid sequences of zinc fingers which bind the particular triplet or quadruplet sequence. See, for example, co-owned U.S. Patents 6,453,242 and 6,534,261.

[0123] Exemplary selection methods, including phage display and two-hybrid systems, are disclosed in US Patents 5,789,538; 5,925,523; 6,007,988; 6,013,453; 6,410,248; 6,140,466; 6,200,759; and 6,242,568; as well as WO 98/37186; WO 98/53057; WO 00/27878; WO 01/88197 and GB 2,338,237. In addition, enhancement of binding specificity for zinc finger binding domains has been described, for example, in co-owned WO 02/077227.

[0124] In addition, as disclosed in these and other references, zinc finger domains and/or multi-fingered zinc finger proteins may be linked together using any suitable linker sequences, including for example, linkers of 5 or more amino acids in length. See, also, U.S. Patent Nos. 6,479,626; 6,903,185; and 7,153,949 for exemplary linker sequences 6 or more amino acids in length. The proteins described herein may include any combination of suitable linkers between the individual zinc fingers of the protein.

[0125] Selection of target sites; ZFPs and methods for design and construction of fusion proteins (and polynucleotides encoding same) are known to those of skill in the art and described in detail in U.S. Patent Nos. 6,140,0815; 789,538; 6,453,242; 6,534,261; 5,925,523; 6,007,988; 6,013,453; 6,200,759; WO 95/19431; WO 96/06166; WO 98/53057; WO 98/54311; WO 00/27878; WO 01/60970 WO 01/88197; WO 02/099084; WO 98/53058; WO 98/53059; WO 98/53060; WO 02/016536 and WO 03/016496.

[0126] In addition, as disclosed in these and other references, zinc finger domains and/or multi-fingered zinc finger proteins may be linked together using any suitable linker sequences, including for example, linkers of 5 or more amino acids in length. See, also, U.S. Patent Nos. 6,479,626; 6,903,185; and 7,153,949 for exemplary linker sequences 6 or more amino acids in length. The proteins described herein may include any combination of suitable linkers between the individual zinc fingers of the protein.

[0127] In some examples, a DNA-binding polypeptide is a DNA-binding domain from GAL4. GAL4 is a modular transactivator in *Saccharomyces cerevisiae*, but it also operates as a transactivator in many other organisms. *See, e.g.*, Sadowski *et al.* (1988) Nature 335:563-4. In this regulatory system, the expression of genes encoding enzymes of the galactose metabolic pathway in *S. cerevisiae* is stringently regulated by the available carbon source. Johnston (1987) Microbiol. Rev. 51:458-76. Transcriptional control of these metabolic enzymes is mediated by the interaction between the positive regulatory protein, GAL4, and a 17 bp symmetrical DNA sequence to which GAL4 specifically binds (the UAS).

[0128] Native GAL4 includes 881 amino acid residues, with a molecular weight of 99 kDa. GAL4 comprises functionally autonomous domains, the combined activities of which account for activity of GAL4 *in vivo*. Ma and Ptashne (1987) Cell 48:847-53); Brent and Ptashne (1985) Cell 43(3 Pt 2):729-36. The N-terminal 65 amino acids of GAL4 comprise the GAL4 DNA-binding domain. Keegan *et al.* (1986) Science 231:699-704; Johnston (1987) Nature 328:353-5. Sequence-specific binding requires the presence of a divalent cation coordinated by 6 Cys residues present in the DNA binding domain. The coordinated cation-containing domain interacts with and recognizes a conserved CCG triplet at each end of the 17 bp UAS *via* direct contacts with the major groove of the DNA helix. Marmorstein *et al.* (1992) Nature 356:408-14. The DNA-binding function of the protein positions C-terminal transcriptional activating domains in the vicinity of the promoter, such that the activating domains can direct transcription.

[0129] Additional DNA-binding polypeptides that may be utilized in certain embodiments include, for example and without limitation, a binding sequence from a AVRBS3-inducible gene; a consensus binding sequence from a AVRBS3-inducible gene or synthetic binding sequence engineered therefrom (*e.g.*, UPA DNA-binding domain); TAL; LexA (*see, e.g.*, Brent & Ptashne (1985), *supra*); LacR (*see, e.g.*, Labow *et al.* (1990) Mol. Cell. Biol. 10:3343-56; Baim *et al.* (1991) Proc. Natl. Acad. Sci. USA 88(12):5072-6); a steroid hormone receptor (Ellliston *et al.* (1990) J. Biol. Chem. 265:11517-121); the Tet repressor (U.S. Patent

6,271,341) and a mutated Tet repressor that binds to a *tet* operator sequence in the presence, but not the absence, of tetracycline (Tc); the DNA-binding domain of NF-κB; and components of the regulatory system described in Wang *et al.* (1994) Proc. Natl. Acad. Sci. USA 91(17):8180-4, which utilizes a fusion of GAL4, a hormone receptor, and VP16.

[0130] In certain embodiments, the DNA-binding domain of one or more of the nucleases used in the methods and compositions described herein comprises a naturally occurring or engineered (non-naturally occurring) TAL effector DNA binding domain. *See, e.g.,* U.S. Patent Publication No. 20110301073. The plant pathogenic bacteria of the genus Xanthomonas are known to cause many diseases in important crop plants. Pathogenicity of Xanthomonas depends on a conserved type III secretion (T3S) system which injects more than 25 different effector proteins into the plant cell. Among these injected proteins are transcription activator-like (TAL) effectors which mimic plant transcriptional activators and manipulate the plant transcriptome (see Kay et al (2007) Science 318:648-651). These proteins contain a DNA binding domain and a transcriptional activation domain. One of the most well characterized TAL-effectors is AvrBs3 from Xanthomonas campestgris pv. Vesicatoria (see Bonas et al (1989) Mol Gen Genet 218: 127-136 and WO2010079430). TAL-effectors contain a centralized domain of tandem repeats, each repeat containing approximately 34 amino acids, which are key to the DNA binding specificity of these proteins. In addition, they contain a nuclear localization sequence and an acidic transcriptional activation domain (for a review see Schornack S, et al (2006) J Plant Physiol 163(3): 256-272). In addition, in the phytopathogenic bacteria Ralstonia solanacearum two genes, designated brg11 and hpx17 have been found that are homologous to the AvrBs3 family of Xanthomonas in the R. solanacearum biovar 1 strain GMI1000 and in the biovar 4 strain RS1000 (See Heuer et al (2007) Appl and Envir Micro 73(13): 4379-4384). These genes are 98.9% identical in nucleotide sequence to each other but differ by a deletion of 1,575 bp in the repeat domain of hpx17. However, both gene products have less than 40% sequence identity with AvrBs3 family proteins of Xanthomonas. See, e.g., U.S. Patent Nos., 8,420,782 and 8,440,431 and U.S. Patent Publication No. 20110301073.

[0131] In other embodiments, the nuclease comprises a CRISPR/Cas system. The CRISPR (clustered regularly interspaced short palindromic repeats) locus, which encodes RNA components of the system, and the cas (CRISPR-associated) locus, which encodes proteins (Jansen *et al.*, 2002. *Mol. Microbiol.* 43: 1565-1575; Makarova *et al.*, 2002. *Nucleic Acids Res.* 30: 482-496; Makarova *et al.*, 2006. *Biol. Direct* 1: 7; Haft *et al.*, 2005. *PLoS Comput. Biol.* 1:

e60) make up the gene sequences of the CRISPR/Cas nuclease system. CRISPR loci in microbial hosts contain a combination of CRISPR-associated (Cas) genes as well as non-coding RNA elements capable of programming the specificity of the CRISPR-mediated nucleic acid cleavage.

[0132] The Type II CRISPR is one of the most well characterized systems and carries out targeted DNA double-strand break in four sequential steps. First, two non-coding RNA, the pre-crRNA array and tracrRNA, are transcribed from the CRISPR locus. Second, tracrRNA hybridizes to the repeat regions of the pre-crRNA and mediates the processing of pre-crRNA into mature crRNAs containing individual spacer sequences. Third, the mature crRNA:tracrRNA complex directs Cas9 to the target DNA via Wastson-Crick base-pairing between the spacer on the crRNA and the protospacer on the target DNA next to the protospacer adjacent motif (PAM), an additional requirement for target recognition. Finally, Cas9 mediates cleavage of target DNA to create a double-stranded break within the protospacer. Activity of the CRISPR/Cas system comprises of three steps: (i) insertion of alien DNA sequences into the CRISPR array to prevent future attacks, in a process called 'adaptation', (ii) expression of the relevant proteins, as well as expression and processing of the array, followed by (iii) RNA-mediated interference with the alien nucleic acid. Thus, in the bacterial cell, several of the so-called 'Cas' proteins are involved with the natural function of the CRISPR/Cas system and serve roles in functions such as insertion of the alien DNA etc.

[0133] In certain embodiments, Cas protein may be a "functional derivative" of a naturally occurring Cas protein. A "functional derivative" of a native sequence polypeptide is a compound having a qualitative biological property in common with a native sequence polypeptide. "Functional derivatives" include, but are not limited to, fragments of a native sequence and derivatives of a native sequence polypeptide and its fragments, provided that they have a biological activity in common with a corresponding native sequence polypeptide. A biological activity contemplated herein is the ability of the functional derivative to hydrolyze a DNA substrate into fragments. The term "derivative" encompasses both amino acid sequence variants of polypeptide, covalent modifications, and fusions thereof. Suitable derivatives of a Cas polypeptide or a fragment thereof include but are not limited to mutants, fusions, covalent modifications of Cas protein or a fragment thereof. Cas protein, which includes Cas protein or a fragment thereof, as well as derivatives of Cas protein or a fragment thereof, may be obtainable from a cell or synthesized chemically or by a combination of these two procedures. The cell may be a cell that naturally produces Cas protein, or a cell that naturally produces Cas protein and is

genetically engineered to produce the endogenous Cas protein at a higher expression level or to produce a Cas protein from an exogenously introduced nucleic acid, which nucleic acid encodes a Cas that is same or different from the endogenous Cas. In some case, the cell does not naturally produce Cas protein and is genetically engineered to produce a Cas protein.

[0134] In particular embodiments, a DNA-binding polypeptide specifically recognizes and binds to a target nucleotide sequence comprised within a genomic nucleic acid of a host organism. Any number of discrete instances of the target nucleotide sequence may be found in the host genome in some examples. The target nucleotide sequence may be rare within the genome of the organism (*e.g.*, fewer than about 10, about 9, about 8, about 7, about 6, about 5, about 4, about 3, about 2, or about 1 copy(ies) of the target sequence may exist in the genome). For example, the target nucleotide sequence may be located at a unique site within the genome of the organism. Target nucleotide sequences may be, for example and without limitation, randomly dispersed throughout the genome with respect to one another; located in different linkage groups in the genome; located in the same linkage group; located on different chromosomes; located on the same chromosome; located in the genome at sites that are expressed under similar conditions in the organism (*e.g.*, under the control of the same, or substantially functionally identical, regulatory factors); and located closely to one another in the genome (*e.g.*, target sequences may be comprised within nucleic acids integrated as concatemers at genomic loci).

B. Targeting endonucleases

[0135] In particular embodiments, a DNA-binding polypeptide that specifically recognizes and binds to a target nucleotide sequence may be comprised within a chimeric polypeptide, so as to confer specific binding to the target sequence upon the chimeric polypeptide. In examples, such a chimeric polypeptide may comprise, for example and without limitation, nuclease, recombinase, and/or ligase polypeptides, as these polypeptides are described above. Chimeric polypeptides comprising a DNA-binding polypeptide and a nuclease, recombinase, and/or ligase polypeptide may also comprise other functional polypeptide motifs and/or domains, such as for example and without limitation: a spacer sequence positioned between the functional polypeptides in the chimeric protein; a leader peptide; a peptide that targets the fusion protein to an organelle (*e.g.*, the nucleus); polypeptides that are cleaved by a cellular enzyme; peptide tags (*e.g.*, Myc, His, etc.); and other amino acid sequences that do not interfere with the function of the chimeric polypeptide.

[0136] Functional polypeptides (*e.g.*, DNA-binding polypeptides and nuclease polypeptides) in a chimeric polypeptide may be operatively linked. In some embodiments, functional polypeptides of a chimeric polypeptide may be operatively linked by their expression from a single polynucleotide encoding at least the functional polypeptides ligated to each other in-frame, so as to create a chimeric gene encoding a chimeric protein. In alternative embodiments, the functional polypeptides of a chimeric polypeptide may be operatively linked by other means, such as by cross-linkage of independently expressed polypeptides.

[0137] In some embodiments, a DNA-binding polypeptide that specifically recognizes and binds to a target nucleotide sequence may be comprised within a natural isolated protein (or mutant thereof), wherein the natural isolated protein or mutant thereof also comprises a nuclease polypeptide (and may also comprise a recombinase and/or ligase polypeptide). Examples of such isolated proteins include TALENs, recombinases (*e.g.*, Cre, Hin, Tre, and FLP recombinase), RNA-guided CRISPR-Cas9, and meganucleases.

[0138] As used herein, the term "targeting endonuclease" refers to natural or engineered isolated proteins and mutants thereof that comprise a DNA-binding polypeptide and a nuclease polypeptide, as well as to chimeric polypeptides comprising a DNA-binding polypeptide and a nuclease. Any targeting endonuclease comprising a DNA-binding polypeptide that specifically recognizes and binds to a target nucleotide sequence comprised within a *FAD3* locus (*e.g.*, either because the target sequence is comprised within the native sequence at the locus, or because the target sequence has been introduced into the locus, for example, by recombination) may be utilized in certain embodiments.

[0139] Some examples of chimeric polypeptides that may be useful in particular embodiments of the invention include, without limitation, combinations of the following polypeptides: zinc finger DNA-binding polypeptides; a FokI nuclease polypeptide; TALE domains; leucine zippers; transcription factor DNA-binding motifs; and DNA recognition and/or cleavage domains isolated from, for example and without limitation, a TALEN, a recombinase (*e.g.*, Cre, Hin, RecA, Tre, and FLP recombinases), RNA-guided CRISPR-Cas9, a meganuclease; and others known to those in the art. Particular examples include a chimeric protein comprising a site-specific DNA binding polypeptide and a nuclease polypeptide. Chimeric polypeptides may be engineered by methods known to those of skill in the art to alter the recognition sequence of a DNA-binding polypeptide comprised within the chimeric polypeptide, so as to target the chimeric polypeptide to a particular nucleotide sequence of interest.

[0140] In certain embodiments, the chimeric polypeptide comprises a DNA-binding domain (*e.g.*, zinc finger, TAL-effector domain, etc.) and a nuclease (cleavage) domain. The cleavage domain may be heterologous to the DNA-binding domain, for example a zinc finger DNA-binding domain and a cleavage domain from a nuclease or a TALEN DNA-binding domain and a cleavage domain, or meganuclease DNA-binding domain and cleavage domain from a different nuclease. Heterologous cleavage domains can be obtained from any endonuclease or exonuclease. Exemplary endonucleases from which a cleavage domain can be derived include, but are not limited to, restriction endonucleases and homing endonucleases. *See*, for example, 2002-2003 Catalogue, New England Biolabs, Beverly, MA; and Belfort *et al.* (1997) *Nucleic Acids Res.* 25:3379-3388. Additional enzymes which cleave DNA are known (*e.g.*, S1 Nuclease; mung bean nuclease; pancreatic DNase I; micrococcal nuclease; yeast HO endonuclease; *see* also Linn *et al.* (eds.) *Nucleases*, Cold Spring Harbor Laboratory Press,1993). One or more of these enzymes (or functional fragments thereof) can be used as a source of cleavage domains and cleavage half-domains.

[0141] Similarly, a cleavage half-domain can be derived from any nuclease or portion thereof, as set forth above, that requires dimerization for cleavage activity. In general, two fusion proteins are required for cleavage if the fusion proteins comprise cleavage half-domains. Alternatively, a single protein comprising two cleavage half-domains can be used. The two cleavage half-domains can be derived from the same endonuclease (or functional fragments thereof), or each cleavage half-domain can be derived from a different endonuclease (or functional fragments thereof). In addition, the target sites for the two fusion proteins are preferably disposed, with respect to each other, such that binding of the two fusion proteins to their respective target sites places the cleavage half-domains in a spatial orientation to each other that allows the cleavage half-domains to form a functional cleavage domain, *e.g.*, by dimerizing. Thus, in certain embodiments, the near edges of the target sites are separated by 5-8 nucleotides or by 15-18 nucleotides. However any integral number of nucleotides or nucleotide pairs can intervene between two target sites (*e.g.*, from 2 to 50 nucleotide pairs or more). In general, the site of cleavage lies between the target sites.

[0142] Restriction endonucleases (restriction enzymes) are present in many species and are capable of sequence-specific binding to DNA (at a recognition site), and cleaving DNA at or near the site of binding, for example, such that one or more exogenous sequences (donors/trangsenes) are integrated at or near the binding (target) sites. Certain restriction enzymes (*e.g.*, Type IIS) cleave DNA at sites removed from the recognition site and have

separable binding and cleavage domains. For example, the Type IIS enzyme *Fok* I catalyzes double-stranded cleavage of DNA, at 9 nucleotides from its recognition site on one strand and 13 nucleotides from its recognition site on the other. *See*, for example, U.S. Patents 5,356,802; 5,436,150 and 5,487,994; as well as Li *et al.* (1992) *Proc. Natl. Acad. Sci. USA* **89**:4275-4279; Li *et al.* (1993) *Proc. Natl. Acad. Sci. USA* **90**:2764-2768; Kim *et al.* (1994a) *Proc. Natl. Acad. Sci. USA* **91**:883-887; Kim *et al.* (1994b) *J. Biol. Chem.* 269:31,978-31,982. Thus, in one embodiment, fusion proteins comprise the cleavage domain (or cleavage half-domain) from at least one Type IIS restriction enzyme and one or more zinc finger binding domains, which may or may not be engineered.

[0143] An exemplary Type IIS restriction enzyme, whose cleavage domain is separable from the binding domain, is *Fok* I. This particular enzyme is active as a dimer. Bitinaite *et al.* (1998) *Proc. Natl. Acad. Sci. USA* 95: 10,570-10,575. Accordingly, for the purposes of the present disclosure, the portion of the *Fok* I enzyme used in the disclosed fusion proteins is considered a cleavage half-domain. Thus, for targeted double-stranded cleavage and/or targeted replacement of cellular sequences using zinc finger-*Fok* I fusions, two fusion proteins, each comprising a *Fok*I cleavage half-domain, can be used to reconstitute a catalytically active cleavage domain. Alternatively, a single polypeptide molecule containing a DNA binding domain and two *Fok*I cleavage half-domains can also be used.

[0144] A cleavage domain or cleavage half-domain can be any portion of a protein that retains cleavage activity, or that retains the ability to multimerize (*e.g.*, dimerize) to form a functional cleavage domain.

[0145] Exemplary Type IIS restriction enzymes are described in U.S. Patent Publication No. 20070134796. Additional restriction enzymes also contain separable binding and cleavage domains, and these are contemplated by the present disclosure. *See*, for example, Roberts *et al.* (2003) *Nucleic Acids Res.* **31**:418-420.

[0146] In certain embodiments, the cleavage domain comprises one or more engineered cleavage half-domain (also referred to as dimerization domain mutants) that minimize or prevent homodimerization, as described, for example, in U.S. Patent Publication Nos. 20050064474; 20060188987 and 20080131962. Amino acid residues at positions 446, 447, 479, 483, 484, 486, 487, 490, 491, 496, 498, 499, 500, 531, 534, 537, and 538 of *Fok* I are all targets for influencing dimerization of the *Fok* I cleavage half-domains.

[0147] Exemplary engineered cleavage half-domains of *Fok* I that form obligate heterodimers include a pair in which a first cleavage half-domain includes mutations at amino acid residues at positions 490 and 538 of *Fok* I and a second cleavage half-domain includes mutations at amino acid residues 486 and 499.

[0148] Thus, in one embodiment, a mutation at 490 replaces Glu (E) with Lys (K); the mutation at 538 replaces Iso (I) with Lys (K); the mutation at 486 replaced Gln (Q) with Glu (E); and the mutation at position 499 replaces Iso (I) with Lys (K). Specifically, the engineered cleavage half-domains described herein were prepared by mutating positions 490 (E→K) and 538 (I→K) in one cleavage half-domain to produce an engineered cleavage half-domain designated "E490K:I538K" and by mutating positions 486 (Q→E) and 499 (I→L) in another cleavage half-domain to produce an engineered cleavage half-domain designated "Q486E:I499L". The engineered cleavage half-domains described herein are obligate heterodimer mutants in which aberrant cleavage is minimized or abolished. *See, e.g.,* U.S. Patent Publication No. 2008/0131962.

[0149] In certain embodiments, the engineered cleavage half-domain comprises mutations at positions 486, 499 and 496 (numbered relative to wild-type FokI), for instance mutations that replace the wild type Gln (Q) residue at position 486 with a Glu (E) residue, the wild type Iso (I) residue at position 499 with a Leu (L) residue and the wild-type Asn (N) residue at position 496 with an Asp (D) or Glu (E) residue (also referred to as a "ELD" and "ELE" domains, respectively). In other embodiments, the engineered cleavage half-domain comprises mutations at positions 490, 538 and 537 (numbered relative to wild-type FokI), for instance mutations that replace the wild type Glu (E) residue at position 490 with a Lys (K) residue, the wild type Iso (I) residue at position 538 with a Lys (K) residue, and the wild-type His (H) residue at position 537 with a Lys (K) residue or a Arg (R) residue (also referred to as "KKK" and "KKR" domains, respectively). In other embodiments, the engineered cleavage half-domain comprises mutations at positions 490 and 537 (numbered relative to wild-type FokI), for instance mutations that replace the wild type Glu (E) residue at position 490 with a Lys (K) residue and the wild-type His (H) residue at position 537 with a Lys (K) residue or a Arg (R) residue (also referred to as "KIK" and "KIR" domains, respectively). (See US Patent Publication No. 20110201055). Engineered cleavage half-domains described herein can be prepared using any suitable method, for example, by site-directed mutagenesis of wild-type

cleavage half-domains (*Fok* I) as described in U.S. Patent Publication Nos. 20050064474; 20080131962; and 20110201055.

[0150] Alternatively, nucleases may be assembled *in vivo* at the nucleic acid target site using so-called "split-enzyme" technology (*see e.g.* U.S. Patent Publication No. 20090068164). Components of such split enzymes may be expressed either on separate expression constructs, or can be linked in one open reading frame where the individual components are separated, for example, by a self-cleaving 2A peptide or IRES sequence. Components may be individual zinc finger binding domains or domains of a meganuclease nucleic acid binding domain.

C. Zinc finger nucleases

[0151] In specific embodiments, a chimeric polypeptide is a custom-designed zinc finger nuclease (ZFN) that may be designed to deliver a targeted site-specific double-strand DNA break into which an exogenous nucleic acid, or donor DNA, may be integrated (*See* co-owned US Patent publication 20100257638). ZFNs are chimeric polypeptides containing a non-specific cleavage domain from a restriction endonuclease (for example, *FokI*) and a zinc finger DNA-binding domain polypeptide. *See, e.g.*, Huang *et al.* (1996) J. Protein Chem. 15:481-9; Kim *et al.* (1997a) Proc. Natl. Acad. Sci. USA 94:3616-20; Kim *et al.* (1996) Proc. Natl. Acad. Sci. USA 93:1156-60; Kim *et al.* (1994) Proc Natl. Acad. Sci. USA 91:883-7; Kim *et al.* (1997b) Proc. Natl. Acad. Sci. USA 94:12875-9; Kim *et al.* (1997c) Gene 203:43-9; Kim *et al.* (1998) Biol. Chem. 379:489-95; Nahon and Raveh (1998) Nucleic Acids Res. 26:1233-9; Smith *et al.* (1999) Nucleic Acids Res. 27:674-81. In some embodiments, the ZFNs comprise non-canonical zinc finger DNA binding domains (see co-owned US Patent publication 20080182332). The *FokI* restriction endonuclease must dimerize *via* the nuclease domain in order to cleave DNA and introduce a double-strand break. Consequently, ZFNs containing a nuclease domain from such an endonuclease also require dimerization of the nuclease domain in order to cleave target DNA. Mani *et al.* (2005) Biochem. Biophys. Res. Commun. 334:1191-7; Smith *et al.* (2000) Nucleic Acids Res. 28:3361-9. Dimerization of the ZFN can be facilitated by two adjacent, oppositely oriented DNA-binding sites. *Id.*

[0152] The flexibility and specificity of the ZFN system provides a level of control previously unachievable by known recombinase-mediated gene editing strategies. As one example, ZFNs can be easily engineered, for example, to recognize specific nucleic acid sequences. Wu *et al.* (2007) Cell. Mol. Life Sci. 64:2933-44 (*See*, US Patent Publications

20090205083, 20110189775, 20110167521 and 20100199389). Randomization of the codons for zinc finger recognition residues allows the selection of new fingers that have high affinity for arbitrarily chosen DNA sequences.

Furthermore, zinc fingers are natural DNA-binding molecules, and engineered zinc fingers have been shown to act on their designed targets in living cells. Thus, nucleases based on zinc fingers are targetable to specific but arbitrary recognition sites.

[0153] In particular examples, a method for the site-specific integration of an exogenous nucleic acid into at least one *FAD3* performance locus of a host comprises introducing into a cell of the host a ZFN, wherein the ZFN recognizes and binds to a target nucleotide sequence, wherein the target nucleotide sequence is comprised within at least one *FAD3* locus of the host. In certain examples, the target nucleotide sequence is not comprised within the genome of the host at any other position than the at least one *FAD3* locus. For example, a DNA-binding polypeptide of the ZFN may be engineered to recognize and bind to a target nucleotide sequence identified within the at least one *FAD3* locus (*e.g.*, by sequencing the *FAD3* locus). A method for the site-specific integration of an exogenous nucleic acid into at least one *FAD3* performance locus of a host that comprises introducing into a cell of the host a ZFN may also comprise introducing into the cell an exogenous nucleic acid, wherein recombination of the exogenous nucleic acid into a nucleic acid of the host comprising the at least one *FAD3* locus is facilitated by site-specific recognition and binding of the ZFN to the target sequence (and subsequent cleavage of the nucleic acid comprising the *FAD3* locus).

VI. Exogenous nucleic acids for integration at a FAD3 locus

[0154] Embodiments of the invention may include one or more nucleic acids selected from the group consisting of: an exogenous nucleic acid for site-specific integration in at least one *FAD3* locus, for example and without limitation, a PTU, ELP, ETIP or an ORF; a nucleic acid comprising a nucleotide sequence encoding a targeting endonuclease; and a vector comprising at least one of either or both of the foregoing. Thus, particular nucleic acids for use in some embodiments include nucleotide sequences encoding a polypeptide, structural nucleotide sequences, and/or DNA-binding polypeptide recognition and binding sites.

A. Exogenous nucleic acid molecules for site-specific integration

[0155] As noted above, insertion of an exogenous sequence (also called a "donor sequence" or "donor" or "transgene") is provided, for example for expression of a polypeptide,

correction of a mutant gene or for increased expression of a wild-type gene. It will be readily apparent that the donor sequence is typically not identical to the genomic sequence where it is placed. A donor sequence can contain a non-homologous sequence flanked by two regions of homology to allow for efficient HDR at the location of interest. Additionally, donor sequences can comprise a vector molecule containing sequences that are not homologous to the region of interest in cellular chromatin. A donor molecule can contain several, discontinuous regions of homology to cellular chromatin. For example, for targeted insertion of sequences not normally present in a region of interest, said sequences can be present in a donor nucleic acid molecule and flanked by regions of homology to sequence in the region of interest.

[0156] The donor polynucleotide can be DNA or RNA, single-stranded or double-stranded and can be introduced into a cell in linear or circular form. *See, e.g.,* U.S. Patent Publication Nos. 20100047805, 20110281361, 20110207221 and U.S. Application No. 13/889,162. If introduced in linear form, the ends of the donor sequence can be protected (*e.g.*, from exonucleolytic degradation) by methods known to those of skill in the art. For example, one or more dideoxynucleotide residues are added to the 3' terminus of a linear molecule and/or self-complementary oligonucleotides are ligated to one or both ends. *See*, for example, Chang *et al.* (1987) *Proc. Natl. Acad. Sci. USA* 84:4959-4963; Nehls *et al.* (1996) *Science* 272:886-889. Additional methods for protecting exogenous polynucleotides from degradation include, but are not limited to, addition of terminal amino group(s) and the use of modified internucleotide linkages such as, for example, phosphorothioates, phosphoramidates, and O-methyl ribose or deoxyribose residues.

[0157] A polynucleotide can be introduced into a cell as part of a vector molecule having additional sequences such as, for example, replication origins, promoters and genes encoding antibiotic resistance. Moreover, donor polynucleotides can be introduced as naked nucleic acid, as nucleic acid complexed with an agent such as a liposome or poloxamer, or can be delivered by viruses (*e.g.*, adenovirus, AAV, herpesvirus, retrovirus, lentivirus and integrase defective lentivirus (IDLV)).

[0158] The donor is generally integrated so that its expression is driven by the endogenous promoter at the integration site, namely the promoter that drives expression of the endogenous gene into which the donor is integrated (*e.g., FAD3*). However, it will be apparent that the donor may comprise a promoter and/or enhancer, for example a constitutive promoter or an inducible or tissue specific promoter.

[0159] Furthermore, although not required for expression, exogenous sequences may also include transcriptional or translational regulatory sequences, for example, promoters, enhancers, insulators, internal ribosome entry sites, sequences encoding 2A peptides and/or polyadenylation signals.

[0160] Exogenous nucleic acids that may be integrated in a site-specific manner into at least one *FAD3* locus, so as to modify the *FAD3* locus, in embodiments include, for example and without limitation, nucleic acids comprising a nucleotide sequence encoding a polypeptide of interest; nucleic acids comprising an agronomic gene; nucleic acids comprising a nucleotide sequence encoding an RNAi molecule; or nucleic acids that disrupt the *FAD3* gene.

[0161] In some embodiments, an exogenous nucleic acid is integrated at a *FAD3* locus, so as to modify the *FAD3* locus, wherein the nucleic acid comprises an agronomic gene or nucleotide sequence encoding a polypeptide of interest, such that the agronomic gene or nucleotide sequence is expressed in the host from the *FAD3* locus. In some examples, the polypeptide of interest (*e.g.*, a foreign protein) is expressed from a nucleotide sequence encoding the polypeptide of interest in commercial quantities. In such examples, the polypeptide of interest may be extracted from the host cell, tissue, or biomass. In some embodiments, the host is a plant, and plant material provided for commercial production of a polypeptide of interest may be a plant, plant part, plant tissue, or plant cell. In some examples, the plant part may be plant seed. Protein extraction from a plant biomass may be accomplished by known methods which are discussed, for example, in Heney and Orr (1981) Anal. Biochem. 114:92-6.

[0162] Likewise, agronomic genes may be expressed in transformed plant cells, plants, and/or their progeny. For example, a plant may be genetically engineered *via* methods of particular embodiments to express various phenotypes of agronomic interest from at least one *FAD3* locus.

[0163] In some embodiments, nucleic acids comprising an agronomic gene or nucleotide sequence encoding a polypeptide of interest may include, for example and without limitation: a gene that confers resistance to a pests or disease (*See*, *e.g.*, Jones *et al.* (1994) Science 266:789 (cloning of the tomato Cf-9 gene for resistance to *Cladosporium fulvum*); Martin *et al.* (1993) Science 262:1432; Mindrinos *et al.* (1994) Cell 78:1089 (*RSP2* gene for resistance to *Pseudomonas syringae*); PCT International Patent Publication No. WO 96/30517 (resistance to soybean cyst nematode); PCT International Patent Publication No. WO 93/19181); a gene that encodes a *Bacillus thuringiensis* protein, a derivative thereof, or a synthetic polypeptide modeled thereon (*See*, *e.g.*, Geiser *et al.* (1986) Gene 48:109 (cloning and

nucleotide sequence of a Bt δ-endotoxin gene; moreover, DNA molecules encoding δ-endotoxin genes can be purchased from American Type Culture Collection (Manassas, VA), for example, under ATCC Accession Nos. 40098; 67136; 31995; and 31998)); a gene that encodes a lectin (*See*, *e.g.*, Van Damme *et al.* (1994) Plant Molec. Biol. 24:25 (nucleotide sequences of several *Clivia miniata* mannose-binding lectin genes)); a gene that encodes a vitamin-binding protein, *e.g.*, avidin (*See* PCT International Patent Publication No. US93/06487 (use of avidin and avidin homologues as larvicides against insect pests)); a gene that encodes an enzyme inhibitor, *e.g.*, a protease, proteinase inhibitor, or amylase inhibitor (*See*, *e.g.*, Abe *et al.* (1987) J. Biol. Chem. 262:16793 (nucleotide sequence of rice cysteine proteinase inhibitor); Huub *et al.* (1993) Plant Molec. Biol. 21:985 (nucleotide sequence of cDNA encoding tobacco proteinase inhibitor I); Sumitani *et al.* (1993) Biosci. Biotech. Biochem. 57:1243 (nucleotide sequence of *Streptomyces nitrosporeus* alpha-amylase inhibitor) and U.S. Patent 5,494,813); a gene encoding an insect-specific hormone or pheromone, *e.g.*, an ecdysteroid or juvenile hormone, a variant thereof, a mimetic based thereon, or an antagonist or agonist thereof (*See*, *e.g.*, Hammock *et al.* (1990) Nature 344:458 (baculovirus expression of cloned juvenile hormone esterase, an inactivator of juvenile hormone)); a gene encoding an insect-specific peptide or neuropeptide that, upon expression, disrupts the physiology of the affected pest (*See*, *e.g.*, Regan (1994) J. Biol. Chem. 269:9 (expression cloning yields DNA coding for insect diuretic hormone receptor); Pratt *et al.* (1989) Biochem. Biophys. Res. Comm. 163:1243 (an allostatin in *Diploptera puntata*); and U.S. Patent 5,266,317 (genes encoding insect-specific, paralytic neurotoxins)); a gene encoding an insect-specific venom produced in nature by a snake, a wasp, or other organism (*See*, *e.g.*, Pang *et al.* (1992) Gene 116:165 (heterologous expression in plants of a gene coding for a scorpion insectotoxic peptide)); a gene encoding an enzyme responsible for a hyperaccumulation of a monoterpene, a sesquiterpene, a steroid, hydroxamic acid, a phenylpropanoid derivative or other molecule with insecticidal activity; a gene encoding an enzyme involved in the modification, including the post-translational modification, of a biologically active molecule, *e.g.*, a glycolytic enzyme, a proteolytic enzyme, a lipolytic enzyme, a nuclease, a cyclase, a transaminase, an esterase, a hydrolase, a phosphatase, a kinase, a phosphorylase, a polymerase, an elastase, a chitinase, or a glucanase, whether natural or synthetic (*See*, *e.g.*, PCT International Patent Publication No. WO 93/02197 (nucleotide sequence of a callase gene); moreover, DNA molecules containing chitinase-encoding sequences can be obtained, for example, from the ATCC, under Accession Nos. 39637 and 67152; Kramer *et al.* (1993) Insect Biochem. Molec. Biol. 23:691 (nucleotide sequence of a cDNA encoding tobacco hornworm chitinase); and

Kawalleck *et al.* (1993) Plant Molec. Biol. 21:673 (nucleotide sequence of the parsley ubi4-2 polyubiquitin gene)); a gene encoding a molecule that stimulates signal transduction (*See*, *e.g.*, Botella *et al.* (1994) Plant Molec. Biol. 24:757 (nucleotide sequences for mung bean calmodulin cDNA clones); and Griess *et al.* (1994) Plant Physiol. 104:1467 (nucleotide sequence of a maize calmodulin cDNA clone)); a gene that encodes a hydrophobic moment peptide (*See*, *e.g.*, PCT International Patent Publication No. WO 95/16776 (peptide derivatives of Tachyplesin which inhibit fungal plant pathogens); and PCT International Patent Publication No. WO 95/18855 (synthetic antimicrobial peptides that confer disease resistance)); a gene that encodes a membrane permease, a channel former, or a channel blocker (*See*, *e.g.*, Jaynes *et al.* (1993) Plant Sci 89:43 (heterologous expression of a cecropin-β lytic peptide analog to render transgenic tobacco plants resistant to *Pseudomonas solanacearum*)); a gene that encodes a viral-invasive protein or complex toxin derived therefrom (*See*, *e.g.*, Beachy *et al.* (1990) Ann. rev. Phytopathol. 28:451); a gene that encodes an insect-specific antibody or immunotoxin derived therefrom (*See*, *e.g.*, Taylor *et al.*, Abstract #497, Seventh Int'l Symposium on Molecular Plant-Microbe Interactions (Edinburgh, Scotland) (1994) (enzymatic inactivation in transgenic tobacco via production of single-chain antibody fragments)); a gene encoding a virus-specific antibody (*See*, *e.g.*, Tavladoraki *et al.* (1993) Nature 366:469 (transgenic plants expressing recombinant antibody genes are protected from virus attack)); a gene encoding a developmental-arrestive protein produced in nature by a pathogen or a parasite (*See*, *e.g.*, Lamb *et al.* (1992) Bio/Technology 10:1436 (fungal endo α-1,4-D-polygalacturonases facilitate fungal colonization and plant nutrient release by solubilizing plant cell wall homo-α-1,4-D-galacturonase); Toubart *et al.* (1992) Plant J. 2:367 (cloning and characterization of a gene which encodes a bean endopolygalacturonase-inhibiting protein)); a gene encoding a developmental-arrestive protein produced in nature by a plant (*See*, *e.g.*, Logemann *et al.* (1992) Bio/Technology 10:305 (transgenic plants expressing the barley ribosome-inactivating gene have an increased resistance to fungal disease)).

[0164] In some embodiments, nucleic acids comprising an agronomic gene or nucleotide sequence encoding a polypeptide of interest may also and/or alternatively include, for example and without limitation: genes that confer resistance to an herbicide, such as an herbicide that inhibits the growing point or meristem, for example, an imidazolinone or a sulfonylurea (exemplary genes in this category encode mutant ALS and AHAS enzymes, as described, for example, by Lee *et al.* (1988) EMBO J. 7:1241, and Miki *et al.* (1990) Theor. Appl. Genet. 80:449, respectively); glyphosate resistance as conferred by, *e.g.*, mutant 5-

enolpyruvylshikimate-3-phosphate synthase (EPSPs) genes (*via* the introduction of recombinant nucleic acids and/or various forms of *in vivo* mutagenesis of native EPSPs genes (including but not limited to CP4, DMMG, and DGT-28); *aroA* genes and glyphosate acetyl transferase (GAT) genes, respectively); other phosphono compounds, such as glufosinate phosphinothricin acetyl transferase (PAT) genes from *Streptomyces* species, including *Streptomyces hygroscopicus* and *Streptomyces viridichromogenes*); and pyridinoxy or phenoxy proprionic acids and cyclohexones (ACCase inhibitor-encoding genes). *See, e.g.*, U.S. Patents 4,940,835 and 6,248,876 (nucleotide sequences of forms of EPSPs which can confer glyphosate resistance to a plant). A DNA molecule encoding a mutant aroA gene can be obtained under ATCC accession number 39256. *See also* U.S. Pat. No. 4,769,061 (nucleotide sequence of a mutant aroA gene). European patent application No. 0 333 033 and U.S. Pat. No. 4,975,374 disclose nucleotide sequences of glutamine synthetase genes, which may confer resistance to herbicides such as L-phosphinothricin. Nucleotide sequences of exemplary PAT genes are provided in European application No. 0 242 246, and DeGreef *et al.* (1989) Bio/Technology 7:61 (production of transgenic plants that express chimeric bar genes coding for PAT activity). Exemplary of genes conferring resistance to phenoxy proprionic acids and cyclohexones, such as sethoxydim and haloxyfop, include the Acc1-S1, Acc1-S2 and Acc1-S3 genes described by Marshall *et al.* (1992) Theor. Appl. Genet. 83:435. GAT genes capable of conferring glyphosate resistance are described, for example, in WO 2005012515. Genes conferring resistance to 2,4-D, phenoxyproprionic acid and pyridyloxy auxin herbicides are described, for example, in WO 2005107437 and WO 2007053482.

[0165] Nucleic acids comprising an agronomic gene or nucleotide sequence encoding a polypeptide of interest may also include, for example and without limitation: a gene conferring resistance to an herbicide that inhibits photosynthesis, such as a triazine (psbA and gs+ genes) or a benzonitrile (nitrilase gene). *See, e.g.*, Przibila *et al.* (1991) Plant Cell 3:169 (transformation of *Chlamydomonas* with plasmids encoding mutant psbA genes). Nucleotide sequences for nitrilase genes are disclosed in U.S. Patent 4,810,648, and DNA molecules containing these genes are available under ATCC Accession Nos. 53435; 67441; and 67442. *See also* Hayes *et al.* (1992) Biochem. J. 285:173 (cloning and expression of DNA coding for a glutathione S-transferase).

[0166] In some embodiments, nucleic acids comprising an agronomic gene or nucleotide sequence encoding a polypeptide of interest may also and/or alternatively include, genes that confer or contribute to a value-added trait, for example and without limitation:

modified fatty acid metabolism, *e.g.*, by transforming a plant with an antisense gene of stearyl-ACP desaturase to increase stearic acid content of the plant (*See*, *e.g.*, Knultzon *et al.* (1992) Proc. Natl. Acad. Sci. U.S.A. 89:2624); decreased phytate content, *e.g.*, introduction of a phytase-encoding gene may enhance breakdown of phytate, adding more free phosphate to the transformed plant (*See*, *e.g.*, Van Hartingsveldt *et al.* (1993) Gene 127:87 (nucleotide sequence of an *Aspergillus niger* phytase gene); a gene may be introduced to reduce phytate content- in maize, for example, this may be accomplished by cloning and then reintroducing DNA associated with the single allele which may be responsible for maize mutants characterized by low levels of phytic acid (*See* Raboy *et al.* (1990) Maydica 35:383)); and modified carbohydrate composition effected, *e.g.*, by transforming plants with a gene encoding an enzyme that alters the branching pattern of starch (*See*, *e.g.*, Shiroza *et al.* (1988) J. Bacteol. 170:810 (nucleotide sequence of *Streptococcus* mutant fructosyltransferase gene); Steinmetz *et al.* (1985) Mol. Gen. Genet. 20:220 (levansucrase gene); Pen *et al.* (1992) Bio/Technology 10:292 (α-amylase); Elliot *et al.* (1993) Plant Molec. Biol. 21:515 (nucleotide sequences of tomato invertase genes); Sogaard *et al.* (1993) J. Biol. Chem. 268:22480 (barley α-amylase gene); and Fisher *et al.* (1993) Plant Physiol. 102:1045 (maize endosperm starch branching enzyme II)).

[0167] In some embodiments, an exogenous nucleic acid is integrated at a *FAD3* locus, so as to modify the *FAD3* locus, wherein the nucleic acid comprises a PTU or ELP, such that, for example, the subsequent site-specific integration of a second exogenous nucleic acid at the site of the PTU or ELP is facilitated. *See, also,* U.S. Application No. 13/889,162.

[0168] Targeting endonuclease-mediated integration of a nucleic acid molecule of interest into a plant genome *via* targeted integration requires delivery of targeting endonucleases or targeting endonuclease-encoding nucleic acid molecules, followed by expression of a functional targeting endonuclease protein in the host. An exogenous nucleic acid is preferably also be present in the host cell at the same time as the targeting endonuclease is delivered or expressed therein, such that functional targeting endonuclease protein induces double-stranded breaks at the target site(s) in the at least one *FAD3* locus, which are then repaired, for example *via* homology-driven integration of the exogenous nucleic acid into the locus. One skilled in the art may envision that expression of a functional targeting endonuclease protein may be achieved by several methods, including, but not limited to, transgenesis of a targeting endonuclease-encoding construct, and transient expression of a targeting endonuclease-encoding construct. In both these cases, expression of a functional targeting endonuclease protein and delivery of an

exogenous nucleic acid in the host cell may be simultaneously achieved in order to drive targeted integration at a *FAD3* locus.

[0169] A particular advantage obtained in embodiments utilizing ZFNs as targeting endonucleases, is that the requirement for dimerization of cleavage domains of chimeric zinc finger nucleases imparts a high level of sequence, and hence cleavage, specificity. Since each set of three fingers binds nine consecutive base pairs, two chimeric nucleases effectively demand an 18 bp target if each zinc finger domain has perfect specificity. Any given sequence of this length is predicted to be unique within a single genome (assuming approximately 10^9 bp). Bibikova *et al.* (2001) Mol. Cell. Biol. 21(1):289-97; Wu *et al.* (2007), *supra.* Furthermore, additional fingers can provide enhanced specificity, Beerli *et al.* (1998) Proc. Natl. Acad. Sci. USA 95:14628-33; Kim and Pabo (1998) Proc. Natl. Acad. Sci. USA 95:2812-7; Liu *et al.* (1997) Proc. Natl. Acad. Sci. USA 94:5525-30, so the number of zinc fingers in each DNA-binding domain may be increased to provide even further specificity. For example, specificity may be further increased by using a pair of 4-, 5-, 6- or more finger ZFNs that recognize a 24 bp sequence. Urnov *et al.* (2005) Nature 435:646-51. Thus, ZFNs may be used such that a recognition sequence is introduced into the host plant genome is unique within the genome.

B. Nucleic acid molecules comprising a nucleotide sequence encoding a targeting endonuclease

[0170] In some embodiments, a nucleotide sequence encoding a targeting endonuclease may be engineered by manipulation (*e.g.*, ligation) of native nucleotide sequences encoding polypeptides comprised within the targeting endonuclease. For example, the nucleotide sequence of a gene encoding a protein comprising a DNA-binding polypeptide may be inspected to identify the nucleotide sequence of the gene that corresponds to the DNA-binding polypeptide, and that nucleotide sequence may be used as an element of a nucleotide sequence encoding a targeting endonuclease comprising the DNA-binding polypeptide. Alternatively, the amino acid sequence of a targeting endonuclease may be used to deduce a nucleotide sequence encoding the targeting endonuclease, for example, according to the degeneracy of the genetic code.

[0171] In exemplary nucleic acid molecules comprising a nucleotide sequence encoding a targeting endonuclease, the last codon of a first polynucleotide sequence encoding a nuclease polypeptide, and the first codon of a second polynucleotide sequence encoding a DNA-binding polypeptide, may be separated by any number of nucleotide triplets, *e.g.*, without coding for an

intron or a "STOP." Likewise, the last codon of a nucleotide sequence encoding a first polynucleotide sequence encoding a DNA-binding polypeptide, and the first codon of a second polynucleotide sequence encoding a nuclease polypeptide, may be separated by any number of nucleotide triplets. In these and further embodiments, the last codon of the last (*i.e.*, most 3' in the nucleic acid sequence) of a first polynucleotide sequence encoding a nuclease polypeptide, and a second polynucleotide sequence encoding a DNA-binding polypeptide, may be fused in phase-register with the first codon of a further polynucleotide coding sequence directly contiguous thereto, or separated therefrom by no more than a short peptide sequence, such as that encoded by a synthetic nucleotide linker (*e.g.*, a nucleotide linker that may have been used to achieve the fusion). Examples of such further polynucleotide sequences include, for example and without limitation, tags, targeting peptides, and enzymatic cleavage sites. Likewise, the first codon of the most 5' (in the nucleic acid sequence) of the first and second polynucleotide sequences may be fused in phase-register with the last codon of a further polynucleotide coding sequence directly contiguous thereto, or separated therefrom by no more than a short peptide sequence.

[0172] A sequence separating polynucleotide sequences encoding functional polypeptides in a targeting endonuclease (*e.g.*, a DNA-binding polypeptide and a nuclease polypeptide) may, for example, consist of any sequence, such that the amino acid sequence encoded is not likely to significantly alter the translation of the targeting endonuclease. Due to the autonomous nature of known nuclease polypeptides and known DNA-binding polypeptides, intervening sequences will not in examples interfere with the respective functions of these structures.

C. Vectors and expression constructs

[0173] In some embodiments, at least one nucleic acid molecule(s) comprising at least one exogenous polynucleotide sequence encoding a polypeptide of interest, and/or a targeting endonuclease, may be introduced into a cell, tissue, or organism for expression therein. For example, a nucleic acid molecule comprising a polynucleotide sequence encoding a targeting endonuclease that specifically recognizes a nucleotide sequence comprised within at least one FAD3 locus may be introduced into a cell for expression of the targeting endonuclease, and a nucleic acid molecule comprising a polynucleotide sequence encoding a polypeptide of interest may be introduced into the cell, such that the polynucleotide sequence encoding the polypeptide of interest is integrated into the at least one *FAD3* locus, *e.g.*, by homologous recombination

following introduction of a double strand break at the locus by the expressed targeting endonuclease, and the polypeptide of interest is expressed from the integrated polynucleotide sequence.

[0174] In some embodiments, a nucleic acid molecule such as one of the foregoing may, for example, be a vector system including, for example and without limitation, a linear plasmid, or a closed circular plasmid. In particular examples, the vector may be an expression vector. Nucleic acid sequences according to particular embodiments may, for example, be integrated into a vector, such that the nucleic acid sequence is operably linked to one or more regulatory sequences. Many vectors are available for this purpose, and selection of the particular vector may depend, for example, on the size of the nucleic acid to be inserted into the vector, the particular host cell to be transformed with the vector, and/or the amount of any encoded polypeptide that is desired to be expressed. A vector typically contains various components, the identity of which depend on a function of the vector (*e.g.*, amplification of DNA or expression of DNA), and the particular host cell(s) with which the vector is compatible.

[0175] In some embodiments, a regulatory sequence operably linked to one or more coding sequence(s) may be a promoter sequence that functions in a host cell, such as a bacterial cell, algal cell, fungal cell, or plant cell, wherein the nucleic acid molecule is to be amplified or expressed. Some embodiments may include a plant transformation vector that comprises a nucleotide sequence comprising at least one regulatory sequence operably linked to one or more nucleotide sequence(s) encoding a polypeptide of interest or a targeting endonuclease, wherein the one or more nucleotide sequence(s) may be expressed, under the control of the regulatory sequence(s), in a plant cell, tissue, or organism to produce the polypeptide of interest or the targeting endonuclease.

[0176] Promoters suitable for use in nucleic acid molecules according to some embodiments include those that are inducible, tissue-specific, viral, synthetic, or constitutive, all of which are well known in the art. Non-limiting examples of promoters that may be useful in embodiments of the invention are provided by: U.S. Patent Nos. 6,437,217 (maize RS81 promoter); 5,641,876 (rice actin promoter); 6,426,446 (maize RS324 promoter); 6,429,362 (maize PR-1 promoter); 6,232,526 (maize A3 promoter); 6,177,611 (constitutive maize promoters); 5,322,938, 5,352,605, 5,359,142, and 5,530,196 (35S promoter); 6,433,252 (maize L3 oleosin promoter); 6,429,357 (rice actin 2 promoter, and rice actin 2 intron); 6,294,714 (light-inducible promoters); 6,140,078 (salt-inducible promoters); 6,252,138 (pathogen-inducible promoters); 6,175,060 (phosphorous deficiency-inducible promoters); 6,388,170 (bidirectional

promoters); 6,635,806 (gamma-coixin promoter); 5,447,858 (soybean heat shock promoter); and U.S. Patent Application Serial No. 09/757,089 (maize chloroplast aldolase promoter).

[0177] Additional exemplary promoters include the nopaline synthase (NOS) promoter (Ebert *et al.* (1987) Proc. Natl. Acad. Sci. USA 84(16):5745-9); the octopine synthase (OCS) promoter (which is carried on tumor-inducing plasmids of *Agrobacterium tumefaciens*); the caulimovirus promoters such as the cauliflower mosaic virus (CaMV) 19S promoter (Lawton *et al.* (1987) Plant Mol. Biol. 9:315-24); the CaMV 35S promoter (Odell *et al.* (1985) Nature 313:810-2; the figwort mosaic virus 35S-promoter (Walker *et al.* (1987) Proc. Natl. Acad. Sci. USA 84(19):6624-8); the sucrose synthase promoter (Yang and Russell (1990) Proc. Natl. Acad. Sci. USA 87:4144-8); the R gene complex promoter (Chandler *et al.* (1989) Plant Cell 1:1175-83); the chlorophyll a/b binding protein gene promoter; CaMV35S (U.S. Patent Nos. 5,322,938, 5,352,605, 5,359,142, and 5,530,196); FMV35S (U.S. Patent Nos. 6,051,753, and 5,378,619); a PC1SV promoter (U.S. Patent No. 5,850,019); the SCP1 promoter (U.S. Patent No. 6,677,503); and AGRtu.nos promoters (GenBank Accession No. V00087; Depicker *et al.* (1982) J. Mol. Appl. Genet. 1:561-73; Bevan *et al.* (1983) Nature 304:184-7).

[0178] In particular embodiments, nucleic acid molecules may comprise a tissue-specific promoter. A tissue-specific promoter is a nucleotide sequence that directs a higher level of transcription of an operably linked nucleotide sequence in the tissue for which the promoter is specific, relative to the other tissues of the organism. Examples of tissue-specific promoters include, without limitation: tapetum-specific promoters; anther-specific promoters; pollen-specific promoters (*See, e.g.*, U.S. Patent No. 7,141,424, and International PCT Publication No. WO 99/042587); ovule-specific promoters; (See, e.g., U.S. Patent Application No. 2001/047525 A1); fruit-specific promoters (*See, e.g.*, U.S. Patent Nos. 4,943,674, and 5,753,475); and seed-specific promoters (*See, e.g.*, U.S. Patent Nos. 5,420,034, and 5,608,152). In some embodiments, a developmental stage-specific promoter (*e.g.*, a promoter active at a later stage in development) may be used.

[0179] Additional regulatory sequences that may in some embodiments be operably linked to a nucleic acid molecule include 5' UTRs located between a promoter sequence and a coding sequence that function as a translation leader sequence. The translation leader sequence is present in the fully-processed mRNA, and it may affect processing of the primary transcript, and/or RNA stability. Examples of translation leader sequences include maize and petunia heat shock protein leaders (U.S. Patent No. 5,362,865), plant virus coat protein leaders, plant rubisco leaders, and others. *See, e.g.*, Turner and Foster (1995) Molecular Biotech. 3(3):225-36. Non-

limiting examples of 5' UTRs are provided by: GmHsp (U.S. Patent No. 5,659,122); PhDnaK (U.S. Patent No. 5,362,865); AtAnt1; TEV (Carrington and Freed (1990) J. Virol. 64:1590-7); and AGRtunos (GenBank Accession No. V00087; and Bevan *et al.* (1983), *supra*).

[0180] Additional regulatory sequences that may in some embodiments be operably linked to a nucleic acid molecule also include 3' non-translated sequences, 3' transcription termination regions, or poly-adenylation regions. These are genetic elements located downstream of a nucleotide sequence, and include polynucleotides that provide polyadenylation signal, and/or other regulatory signals capable of affecting transcription or mRNA processing. The polyadenylation signal functions in plants to cause the addition of polyadenylate nucleotides to the 3' end of the mRNA precursor. The polyadenylation sequence can be derived from a variety of plant genes, or from T-DNA genes. A non-limiting example of a 3' transcription termination region is the nopaline synthase 3' region (nos 3'; Fraley *et al.* (1983) Proc. Natl. Acad. Sci. USA 80:4803-7). An example of the use of different 3' nontranslated regions is provided in Ingelbrecht *et al.* (1989) Plant Cell 1:671-80. Non-limiting examples of polyadenylation signals include one from a *Pisum sativum* RbcS2 gene (Ps.RbcS2-E9; Coruzzi *et al.* (1984) EMBO J. 3:1671-9) and AGRtu.nos (GenBank Accession No. E01312).

[0181] Additional information regarding regulatory sequences that may be useful in particular embodiments is described, for example, in Goeddel (1990) "Gene Expression Technology," Methods Enzymol. 185, Academic Press, San Diego, CA.

[0182] A recombinant nucleic acid molecule or vector may comprise a selectable marker that confers a selectable phenotype on a transformed cell, such as a plant cell. Selectable markers may also be used to select for cells or organisms that comprise a nucleic acid molecule comprising the selectable marker. A marker may encode biocide resistance, antibiotic resistance (*e.g.*, kanamycin, Geneticin (G418), bleomycin, and hygromycin), or herbicide resistance (*e.g.*, glyphosate). Examples of selectable markers include, but are not limited to: a *neo* gene that confers kanamycin resistance and can be selected for using, *e.g.*, kanamycin and G418; a *bar* gene that confers bialaphos resistance; a mutant EPSP synthase gene that confers glyphosate resistance; a nitrilase gene that confers resistance to bromoxynil; a mutant acetolactate synthase gene (ALS) that confers imidazolinone or sulfonylurea resistance; and a methotrexate-resistant DHFR gene. Multiple selectable markers are available that confer resistance to chemical agents including, for example and without limitation, ampicillin; bleomycin; chloramphenicol; gentamycin; hygromycin; kanamycin; lincomycin; methotrexate; phosphinothricin; puromycin;

spectinomycin; rifampicin; streptomycin; and tetracycline. Examples of such selectable markers are illustrated in, *e.g.*, U.S. Patents 5,550,318; 5,633,435; 5,780,708 and 6,118,047.

[0183] A nucleic acid molecule or vector may also or alternatively include a screenable marker. Screenable markers may be used to monitor expression. Exemplary screenable markers include a β-glucuronidase or uidA gene (*GUS*) which encodes an enzyme for which various chromogenic substrates are known (Jefferson *et al.* (1987) Plant Mol. Biol. Rep. 5:387-405); an R-locus gene, which encodes a product that regulates the production of anthocyanin pigments (red color) in plant tissues (Dellaporta *et al.* (1988) "Molecular cloning of the maize R-nj allele by transposon tagging with Ac." In 18th Stadler Genetics Symposium, P. Gustafson and R. Appels, eds., Plenum, NY (pp. 263-82); a β-lactamase gene (Sutcliffe *et al.* (1978) Proc. Natl. Acad. Sci. USA 75:3737-41); a gene which encodes an enzyme for which various chromogenic substrates are known (*e.g.*, PADAC, a chromogenic cephalosporin); a luciferase gene (Ow *et al.* (1986) Science 234:856-9); a *xylE* gene that encodes a catechol dioxygenase that converts chromogenic catechols (Zukowski *et al.* (1983) Gene 46(2-3):247-55); an amylase gene (Ikatu *et al.* (1990) Bio/Technol. 8:241-2); a tyrosinase gene which encodes an enzyme capable of oxidizing tyrosine to DOPA and dopaquinone, which in turn condenses to melanin (Katz *et al.* (1983) J. Gen. Microbiol. 129:2703-14); and an α-galactosidase.

[0184] All of the nucleotide sequences that encode, for example, a particular polypeptide of interest or a particular targeting endonuclease, will be immediately recognizable by those of skill in the art. The degeneracy of the genetic code provides a finite number of coding sequences for a particular amino acid sequence. The selection of a particular sequence to encode a polypeptide according to embodiments of the invention is within the discretion of the practitioner. Different coding sequences may be desirable in different applications.

[0185] In some embodiments, it may be desirable to modify the nucleotides of a nucleic acid, for example, to enhance expression of a polynucleotide sequence comprised within the nucleic acid in a particular host. The genetic code is redundant with 64 possible codons, but most organisms preferentially use a subset of these codons. The codons that are utilized most often in a species are called optimal codons, and those not utilized very often are classified as rare or low-usage codons. Zhang *et al.* (1991) Gene 105:61-72. Codons may be substituted to reflect the preferred codon usage of a particular host in a process sometimes referred to as "codon optimization." Optimized coding sequences containing codons preferred by a particular prokaryotic or eukaryotic host may be prepared, for example, to increase the rate of translation

or to produce recombinant RNA transcripts having desirable properties (*e.g.*, a longer half-life, as compared with transcripts produced from a non-optimized sequence).

[0186] Nucleic acids may be introduced into a host cell in embodiments of the invention by any method known to those of skill in the art, including, for example and without limitation: by transformation of protoplasts (*See, e.g.*, U.S. Patent 5,508,184); by desiccation/inhibition-mediated DNA uptake (*See, e.g.*, Potrykus *et al.* (1985) Mol. Gen. Genet. 199:183-8); by electroporation (*See, e.g.*, U.S. Patent 5,384,253); by agitation with silicon carbide fibers (*See, e.g.*, U.S. Patents 5,302,523 and 5,464,765); by *Agrobacterium*-mediated transformation (*See, e.g.*, U.S. Patents 5,563,055, 5,591,616, 5,693,512, 5,824,877, 5,981,840, and 6,384,301); and by acceleration of DNA-coated particles (*See, e.g.*, U.S. Patents 5,015,580, 5,550,318, 5,538,880, 6,160,208, 6,399,861, and 6,403,865). Through the application of techniques such as these, the cells of virtually any species may be stably transformed. In some embodiments, transforming DNA is integrated into the genome of the host cell. In the case of multicellular species, transgenic cells may be regenerated into a transgenic organism. Any of these techniques may be used to produce a transgenic plant, for example, comprising one or more nucleic acid sequences of the invention in the genome of the transgenic plant.

[0187] The most widely-utilized method for introducing an expression vector into plants is based on the natural transformation system of *Agrobacterium*. *A. tumefaciens* and *A. rhizogenes* are plant pathogenic soil bacteria that genetically transform plant cells. The T_i and R_i plasmids of *A. tumefaciens* and *A. rhizogenes*, respectively, carry genes responsible for genetic transformation of the plant. The T_i (tumor-inducing)-plasmids contain a large segment, known as T-DNA, which is transferred to transformed plants. Another segment of the T_i plasmid, the vir region, is responsible for T-DNA transfer. The T-DNA region is bordered by left-hand and right-hand borders that are each composed of terminal repeated nucleotide sequences. In some modified binary vectors, the tumor-inducing genes have been deleted, and the functions of the vir region are utilized to transfer foreign DNA bordered by the T-DNA border sequences. The T-region may also contain, for example, a selectable marker for efficient recovery of transgenic plants and cells, and a multiple cloning site for inserting sequences for transfer such as a nucleic acid encoding a fusion protein of the invention.

[0188] Thus, in some embodiments, a plant transformation vector is derived from a T_i plasmid of *A. tumefaciens* (*See, e.g.*, U.S. Patent Nos. 4,536,475, 4,693,977, 4,886,937, and 5,501,967; and European Patent EP 0 122 791) or a R_i plasmid of *A. rhizogenes*. Additional plant transformation vectors include, for example and without limitation, those described by

Herrera-Estrella *et al.* (1983) Nature 303:209-13; Bevan *et al.* (1983), *supra*; Klee *et al.* (1985) Bio/Technol. 3:637-42; and in European Patent EP 0 120 516, and those derived from any of the foregoing. Other bacteria, such as *Sinorhizobium*, *Rhizobium*, and *Mesorhizobium*, that naturally interact with plants can be modified to mediate gene transfer to a number of diverse plants. These plant-associated symbiotic bacteria can be made competent for gene transfer by acquisition of both a disarmed T_i plasmid and a suitable binary vector.

[0189] After providing exogenous DNA to recipient cells, transformed cells are generally identified for further culturing and plant regeneration. In order to improve the ability to identify transformed cells, one may desire to employ a selectable or screenable marker gene, as previously set forth, with the vector used to generate the transformant. In the case where a selectable marker is used, transformed cells are identified within the potentially transformed cell population by exposing the cells to a selective agent or agents. In the case where a screenable marker is used, cells may be screened for the desired marker gene trait.

[0190] Cells that survive the exposure to the selective agent, or cells that have been scored positive in a screening assay, may be cultured in media that supports regeneration of plants. In some embodiments, any suitable plant tissue culture media (*e.g.*, MS and N6 media) may be modified by including further substances, such as growth regulators. Tissue may be maintained on a basic media with growth regulators until sufficient tissue is available to begin plant regeneration efforts, or following repeated rounds of manual selection, until the morphology of the tissue is suitable for regeneration (*e.g.*, at least 2 weeks), then transferred to media conducive to shoot formation. Cultures are transferred periodically until sufficient shoot formation has occurred. Once shoots are formed, they are transferred to media conducive to root formation. Once sufficient roots are formed, plants can be transferred to soil for further growth and maturity.

[0191] To confirm the presence of a nucleic acid molecule of interest (for example, a nucleotide sequence encoding a polypeptide comprising at least one fusion protein of the invention) in a regenerating plant, a variety of assays may be performed. Such assays include, for example: molecular biological assays, such as Southern and Northern blotting, PCR, and nucleic acid sequencing; biochemical assays, such as detecting the presence of a protein product, *e.g.*, by immunological means (ELISA and/or Western blots) or by enzymatic function; plant part assays, such as leaf or root assays; and analysis of the phenotype of the whole regenerated plant.

[0192] Integration events may be analyzed, for example, by PCR amplification using, *e.g.*, oligonucleotide primers that are specific for a nucleotide sequence of interest. PCR genotyping is understood to include, but not be limited to, polymerase-chain reaction (PCR) amplification of genomic DNA derived from isolated host plant tissue predicted to contain a nucleic acid molecule of interest integrated into the genome, followed by standard cloning and sequence analysis of PCR amplification products. Methods of PCR genotyping have been well described (*see, e.g.*, Rios, G. *et al.* (2002) Plant J. 32:243-53), and may be applied to genomic DNA derived from any plant species or tissue type, including cell cultures.

[0193] A transgenic plant formed using *Agrobacterium*-dependent transformation methods typically contains a single to multiple copies of recombinant DNA. The single recombinant DNA sequence is referred to as a "transgenic event" or "integration event." Such transgenic plants are heterozygous for the inserted DNA sequence. In some embodiments, a transgenic plant homozygous with respect to a transgene may be obtained by sexually mating (selfing) an independent segregant transgenic plant that contains a single exogenous gene sequence to itself, for example, an F_0 plant, to produce F_1 seed. One fourth of the F_1 seed produced will be homozygous with respect to the transgene. Germinating F_1 seed results in plants that can be tested for heterozygosity, typically using a SNP assay or a thermal amplification assay that allows for the distinction between heterozygotes and homozygotes (*i.e.*, a zygosity assay).

[0194] In addition to direct transformation of a plant or plant cell with a nucleic acid molecule in some embodiments, transgenic plants may be prepared in particular embodiments by crossing a first plant having at least one transgenic event with a second plant lacking such an event. For example, a nucleic acid comprising at least one modified *FAD3* locus, wherein an exogenous nucleic acid has been integrated in a site-specific manner, may be introduced into a first plant line that is amenable to transformation, to produce a transgenic plant, which transgenic plant may be crossed with a second plant line to introgress the at least one modified *FAD3* locus (and therefore the exogenous nucleic acid) into the second plant line.

[0195] To confirm the presence of a nucleic acid molecule of interest in regenerating plants, a variety of assays may be performed. Such assays include, for example: molecular biological assays, such as Southern and Northern blotting and PCR; biochemical assays, such as detecting the presence of a protein product, *e.g.*, by immunological means (ELISA and/or Western blots) or by enzymatic function; plant part assays, such as leaf or root assays; and analysis of the phenotype of the whole regenerated plant.

[0196] Targeted integration events may be screened, for example, by PCR amplification using, *e.g.*, oligonucleotide primers specific for nucleic acid molecules of interest. PCR genotyping is understood to include, but not be limited to, polymerase-chain reaction (PCR) amplification of genomic DNA derived from isolated host plant callus tissue predicted to contain a nucleic acid molecule of interest integrated into the genome, followed by standard cloning and sequence analysis of PCR amplification products. Methods of PCR genotyping have been well described (for example, Rios, G. *et al.* (2002) Plant J. 32:243-53) and may be applied to genomic DNA derived from any plant species or tissue type, including cell cultures. Combinations of oligonucleotide primers that bind to both target sequence and introduced sequence may be used sequentially or multiplexed in PCR amplification reactions. Oligonucleotide primers designed to anneal to the target site, introduced nucleic acid sequences, and/or combinations of the two are feasible. Thus, PCR genotyping strategies may include (but are not limited to) amplification of specific sequences in the plant genome, amplification of multiple specific sequences in the plant genome, amplification of non-specific sequences in the plant genome, or combinations thereof. One skilled in the art may devise additional combinations of primers and amplification reactions to interrogate the genome. For example, a set of forward and reverse oligonucleotide primers may be designed to anneal to nucleic acid sequence(s) specific for the target outside the boundaries of the introduced nucleic acid sequence.

[0197] Forward and reverse oligonucleotide primers may be designed to anneal specifically to an introduced nucleic acid molecule of interest, for example, at a sequence corresponding to a coding region within the nucleic acid molecule of interest, or other parts of the nucleic acid molecule of interest. These primers may be used in conjunction with the primers described above. Oligonucleotide primers may be synthesized according to a desired sequence, and are commercially available (*e.g.*, from Integrated DNA Technologies, Inc., Coralville, IA). Amplification may be followed by cloning and sequencing, or by direct sequence analysis of amplification products. One skilled in the art might envision alternative methods for analysis of amplification products generated during PCR genotyping. In one embodiment, oligonucleotide primers specific for the gene target are employed in PCR amplifications.

VI. Transgenic plants and plant materials comprising a nucleic acid integrated at a FAD3 performance locus

[0198] In some embodiments, a transgenic plant is provided, wherein the plant comprises a plant cell comprising at least one modified (*e.g.*, *FAD3* locus, disrupted and/or targeted integration of an exogenous sequence) FAD3 locus. In particular embodiments, such a plant may be produced by transformation of a plant tissue or plant cell, and regeneration of a whole plant. In further embodiments, such a plant may be obtained through introduction of an exogenous nucleic acid at the at least one *FAD3* locus in a site-specific manner, or through introgression of the modified *FAD3* locus into a germplasm. Plant materials comprising such a plant cell are also provided. Such a plant material may be obtained from a plant comprising the plant cell.

[0199] A transgenic plant or plant material comprising a plant cell comprising at least one modified *FAD3* locus may in some embodiments exhibit one or more of the following characteristics: expression of a targeting endonuclease in a cell of the plant; expression of a polypeptide of interest in a cell of the plant (or in a plastid therein); expression of a targeting endonuclease in the nucleus of a cell of the plant; localization of a targeting endonuclease in a cell of the plant; integration at a *FAD3* locus in the genome of a cell of the plant; integration of a nucleotide sequence encoding a polypeptide of interest or an agronomic gene at a *FAD3* locus in the genome of a cell of the plant; and/or the presence of an RNA transcript corresponding to a coding sequence integrated at a *FAD3* locus in the genome of a cell of the plant. Such a plant may additionally have one or more desirable traits, including, for example and without limitation, those resulting from the expression of an endogenous or transgenic nucleotide sequence, the expression of which is regulated by a polypeptide of interest or an agronomic gene integrated at a *FAD3* locus in the genome of a cell of the plant; resistance to insects, other pests, and disease-causing agents; tolerances to herbicides; enhanced stability, yield, or shelf-life; environmental tolerances; pharmaceutical production; industrial product production; and nutritional enhancements.

[0200] A transgenic plant according to the invention may be any plant capable of being transformed with a nucleic acid that is subsequently integrated in at least one *FAD3* locus according to methods described herein. Accordingly, the plant may be a dicot or monocot. Non-limiting examples of dicotyledonous plants usable in the present methods include *Arabidopsis*, alfalfa, beans, broccoli, cabbage, canola, carrot, cauliflower, celery, Chinese cabbage, cotton, cucumber, eggplant, lettuce, melon, pea, pepper, peanut, potato, pumpkin,

radish, rapeseed, spinach, soybean, squash, sugarbeet, sunflower, tobacco, tomato, and watermelon. Non-limiting examples of monocotyledonous plants usable in the present methods include corn, barley, onion, rice, sorghum, wheat, rye, millet, sugarcane, oat, triticale, switchgrass, and turfgrass. Transgenic plants according to the invention may be used or cultivated in any manner.

[0201] Some embodiments also provide commodity products produced from transgenic plants of the invention. Commodity products include, for example and without limitation: food products, meals, oils, or crushed or whole grains or seeds of a plant comprising one or more nucleotide sequences integrated in at least one *FAD3* locus. The detection of one or more such nucleotide sequences in one or more commodity or commodity products is *de facto* evidence that the commodity or commodity product was at least in part produced from a transgenic plant produced according to an embodiment of the invention. In some embodiments, a transgenic plant or seed comprising a plant cell comprising at least one modified *FAD3* locus may comprise at least one other transgenic event in its genome, including without limitation: a transgenic event from which is transcribed an RNAi molecule; a gene encoding an insecticidal protein (*e.g.*, a *Bacillus thuringiensis* insecticidal protein); an herbicide tolerance gene (*e.g.*, a gene providing tolerance to glyphosate); and a gene contributing to a desirable phenotype in the transgenic plant (*e.g.*, increased yield, altered fatty acid metabolism, or restoration of cytoplasmic male sterility).

[0202] A transgenic plant comprising a plant cell comprising at least one modified *FAD3* locus may have one or more desirable traits. Such traits can include, for example: resistance to insects, other pests, and disease-causing agents; tolerances to herbicides; enhanced stability, yield, or shelf-life; environmental tolerances; pharmaceutical production; industrial product production; and nutritional enhancements. The desirable traits may be conferred by one or more nucleic acid molecules integrated by targeted recombination at the *FAD3* locus that are expressed in the plant exhibiting the desirable traits. Thus, in some embodiments, the desired trait can be due to the presence of a transgene(s) in the plant, which is introduced into the genome of the plant at the site of at least one modified *FAD3* locus. In an additional embodiment, the desirable trait can be obtained through conventional breeding, which trait may be conferred by one or more nucleic acid molecules integrated by targeted recombination at the at least one modified *FAD3* locus.

[0203] Transgenic plants according to the invention may be used or cultivated in any manner, wherein presence of at least one modified *FAD3* locus is desirable. Accordingly, a

plant may be engineered to, *inter alia*, have one or more desired traits, by being transformed with nucleic acid molecules that are subsequently integrated in a site-specific manner in at least one *FAD3* locus according to the invention, and cropped and cultivated by any method known to those of skill in the art.

VII. Marker-assisted breeding of transgenic plants comprising a nucleic acid integrated at a FAD3 performance locus

[0204] Molecular markers that are linked (*e.g.*, tightly-linked) to *Fad2* and *Fad3*, in *Brasicca* spp. are provided. For example, DNA segments containing sequences involved in the HO trait (*FAD3*) are identified. These segments are located around and between markers that are linked (*e.g.*, tightly-linked) to the mutant alleles in a genomic linkage group. Thus, nucleic acid molecules comprising a mutant *FAD3* gene having an inactivating mutation are also provided. The segments identified, and the markers thereof, are included in the present subject matter, in part, by their position in linkage groups in the *B. napus* genome.

[0205] The references discussed herein are provided solely for their disclosure prior to the filing date of the present application. Nothing herein is to be construed as an admission that the inventors are not entitled to antedate such disclosure by virtue of prior invention. The following examples are provided to illustrate certain particular features and/or embodiments. The examples should not be construed to limit the disclosure to the particular features or embodiments exemplified.

EXAMPLES

EXAMPLE 1: IDENTIFICATION OF FAD3 TARGET SEQUENCES FROM A BACTERIAL ARTIFICIAL CHROMOSOME LIBRARY

BAC LIBRARY CONSTRUCTION

[0206] A Bacterial Artificial Chromosome (BAC) library was sourced from a commercial vendor (Amplicon Express, Pullman, WA). The BAC library included 110,592 BAC clones containing high molecular weight genomic DNA (gDNA) fragments isolated from *Brassica napus* L. var. DH10275. The gDNA was digested with either the *BamHI* or *HindIII*

restriction enzyme. Isolated gDNA fragments of about 135 Kbp were ligated into the pCC1BAC vector (Epicentre, Madison, WI) and transformed into *Escherichia coli* str. DH10B (Invitrogen). The BAC library was made up of an even number of BAC clones that were constructed using the two different restriction enzymes. As such, the *Hind III* constructed BAC library was contained in 144 individual 384-well plates. Likewise, the *BamHI* constructed BAC library was contained in 144 individual 384-well plates. A total of 110,592 BAC clones were isolated and arrayed into 288 individual 384-well plates. Each of the 288 individual 384 well plates were provided by the vendor as a single DNA extraction for rapid PCR based screening. The resulting BAC library covers approximately 15 Gbp of gDNA, which corresponds to a 12-fold genome coverage of *Brassica napus* L. var. DH10275genome (estimate of the *Brassica napus* L. genome is ca. 1.132 Gbp as described in Johnston *et al.* (2005) Annals of Botany 95:229-235).

SEQUENCE ANALYSIS OF FAD3 CODING SEQUENCES ISOLATED FROM THE BAC LIBRARY

[0207] The constructed BAC library was used to isolate FAD3 gene coding sequences. Sequencing experiments were conducted to identify the specific gene sequences of six FAD3 gene homeologoues and paralogs from *Brassica napus* L. var. DH10275.

[0208] The FAD3 gene sequence was initially identified within the model species *Arabidopsis thaliana*. The gene sequence is listed in Genbank as Locus Tag: At2g29980. Comparative genomic relationships between the model plant species *Arabidopsis thaliana* and the diploid *Brassica rapa,* one of the progenitors of the tetraploid *Brassica napus,* have been previously described. (Schranz *et al.* (2006) Trends in Plant Science 11(11):535-542). With specific relation to the FAD gene the comparative analysis predicted that 3-4 copies of the gene may occur within the diploid *Brassica* genome. Additional genetic mapping studies were completed by Scheffler *et al.* (1997) Theoretical and Applied Genetics 94; 583-591. The results of these genetic mapping studies indicated that six copies of the FAD3 gene were present in *Brassica napus*.

[0209] Previous sequencing efforts focused on the FAD3 genes from *Brassica napus* had identified and genetically mapped both A and C genome specific copies (Hu *et al.*, (2006) Theoretical and Applied Genetics, 113(3): 497-507). A collection of EST sequences from seed specific cDNA libraries had previously been constructed and sequenced from the plant line DH12075 by Andrew Sharpe of Agriculture and Agri-food Canada, 107 Science Place, Saskatoon, Saskatchewan. As a collection of ESTs from the doubled haploid canola plant

DH12075 full length gene sequences were not available, moreover the indications of sequence quality and confidence of correctly called nucleotides was also not available. Consequently, sequence variation between different FAD gene sequence reads could not be unequivocally attributed to different gene copies of the various homeologues and paralogs of the FAD3 gene family, nor was the genomic sequence available. However, when a combined sequence analysis was performed with the ESTs as well as the two FAD3A and FAD3C full length gene sequences described in Hu *et al.*, (2006), ESTs that matched both of the genes were identified along with an additional 4 haplotypes. As a result, a total of six unique haplotypes of FAD3 were identified. Following the assembly of all available data for the various FAD3 haplotypes, high levels of exon sequence divergence in exon 1 was identified. The divergence of the FAD3 sequence in exon 1 was identified as an opportunity which could be utilized for the design of gene/allele specific PCR primers. In addition, exons were identified that were either minimally differentiated between haplotypes (*e.g.*, exons 5, 6, 7 and 8 had 1-3 bp that varied between FAD3A and FAD3C) or that were devoid of sequence variation (*e.g.*, exons 2 and 3).

[0210] Sequencing analysis of the BAC library which was constructed from *B. napus* L. var. DH12075 resulted in the isolation of six BAC sequences (SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, and SEQ ID NO:6) from which the coding sequences for the FAD3A (SEQ ID NO:7), FAD3A' (SEQ ID NO:8), FAD3A'' (SEQ ID NO:9), FAD3C (SEQ ID NO:10), FAD3C'' (SEQ ID NO:11), and FAD3C' (SEQ ID NO:12) genes were determined. The FAD3A, FAD3A', FAD3A'', FAD3C, FAD3C'', and FAD3C' gene sequences were identified and genetically mapped.

[0211] Sequence analysis of the six FAD3 genes was conducted using a sequence alignment program and a neighbor-joining tree using percentage of identity. The sequence alignment was made via the AlignX® program from the Vector NTI Advance 11.0 computer program (Life Technologies, Carlsbad, CA) and is shown in Figure 1. AlignX® uses a modified Clustal W algorithm to generate multiple sequence alignments of either protein or nucleic acid sequences for similarity comparisons and for annotation. The neighbour-joining tree was created with Jalview v2.3® software and is shown in Figure 2. (Waterhouse *et al.* (2009) Bioinformatics 25 (9) 1189-1191). The contigs identified as containing FAD3 genes were used as BLASTn queries against a database of *Arabidopsis thaliana* genes. The region of each of the 6 contigs containing the FAD3 gene was identified through comparison to the *Arabidopsis thaliana* FAD3 gene (Genbank Accession No: At2g29980). The FAD3 contigs were then orientated such that all FAD3 genes were in the 5' to 3' orientation. FAD3 contigs were trimmed

to contain as many as 2 upstream (5') and 1 downstream (3') *Arabidopsis thaliana* genes where possible. Once orientated the complete coding region of the FAD3 genes were extracted from each contig and used to generate a Neighbour joining tree to display the relationship between the different FAD3 gene family members. The 6 FAD3 family members were aligned into 3 pairs of FAD3 genes (Figure 2).

PCR BASED SCREENING

[0212] A cohort of PCR primers were designed to screen the aforementioned BAC library. The primers were designed as either universal primers, which would amplify all members of the gene family, or as gene specific primers for targeted allele amplification. The PCR primers were designed to be 20 bp long (+/- 1bp) and contain a G/C content of 50% (+/- 8%). Table 1 lists the primers which were designed and synthesized. The clones of the BAC library were pooled and screened via the Polymerase Chain Reaction (PCR).

Table 1: Primer sequences used for PCR amplification of FAD3 sequences

Primer Name:	**SEQ ID NO:**	**Sequence:**
D_uni_F3_F1	SEQ ID NO:13	GAATAAGCCATCGGACACAC
D_spec_F3_F2	SEQ ID NO:14	ATGCGAACGGAGACGAAAGG
D_spec_F3_F3	SEQ ID NO:15	TGTTAACGGAGATTCCGGTG
D_spec_F3_F4	SEQ ID NO:16	GTAGCAATGTGAACGGAGAT
D_uni_F3_R1	SEQ ID NO:17	CAGTGTATCTGAGCATCCG
D_spec_F3_R2	SEQ ID NO:18	GTGGCCGAGTACGAAGATAG
D_spec_F3_R3	SEQ ID NO:19	CAGTAGAGTGGCCAGAGGA

[0213] Two different sets of conditions were used for the polymerase chain reactions (PCR). The first series of PCR reactions contained: 1X PCR buffer (containing dNTPs); 1.5 mM $MgCl_2$; 200 μM of 0.25 U Immolase® DNA polymerase (Bioline, London, UK); 250 nM of each primer; and, about 5-10 ng template DNA. A second series of PCR reactions were developed for the amplification of genomic DNA and contained: 5-10 ng of genomic DNA, 1X PCR buffer, 2 mM dNTPs, 0.4 μM forward and reverse primer, and 0.25 U Immolase® DNA polymerase (Bioline, London, UK). Reagents were pooled into a final volume of 13 μL and amplified using an MJ PTC200® thermocycler (BioRad, Hercules, CA) or an ABI 9700 Gene Amp System® (Life Technologies, Carlsbad, CA). PCR based screening of specific plates was conducted using a 4 dimension screening approach based on the screening system described by Bryan *et al* (Scottish Crops Research Institute annual report: 2001-2002) with the above described PCR conditions. Following PCR based screening of pooled BAC libraries; the amplified PCR product was sequenced using a direct Sanger sequencing method. The amplified products were purified with ethanol, sodium acetate and EDTA following the BigDye® v3.1 protocol (Applied Biosystems) and electrophoresis was performed on an ABI3730xl® automated capillary electrophoresis platform.

[0214] Following PCR based screening and confirmational Sanger sequencing, a collection of plates were identified that contained the various different FAD3 gene family members. A total of six unique FAD3 homeologous and paralogous gene sequences were identified (Table 2). A total of two plates per each FAD3 gene sequence were chosen to undergo plate screening to identify the specific well and clone within the plate that contained the FAD3 gene (Table 2). The specific wells were identified for both of the plates and an individual clone was selected for each of the FAD3 gene family members (Table 2).

Table 2: Identification of the BAC clone plates that provided positive reaction with the detailed PCR primer combinations, along with two plate identities that were taken forward for clone identification within the plate

Gene Name	Primer Sets	Positive Plate Pools	Chosen Plates
FAD3A (FAD3A-1)	F2+R2	16, 231	Plate 16 Plate 231
FAD3C	F4+R2	18, 27, 136, 178, 211, 232	Plate 18 Plate 27

FAD3C" (Haplotype1)	F4+R2, F4+R3, F3+R3	23, 44, 53, 56, 77, 116, 158, 199, 209, 278, 280, 282, 283, 284, 286	Plate 44 Plate 199
FAD3A' (FAD3A'/FAD3A")	F4+R2	52, 121, 139	Plate 121 Plate 139
FAD3C' (Haplotype2)	F4+R2	144, 188, 235	Plate 144 Plate 188
FAD3A" (Haplotype3)	F4+R3 and F3+R3	69, 105, 106, 229, 242, 247, 248	Plate 69 Plate 106

[0215] The single BAC clone, for each identified FAD gene family member, was further analysed via sequencing. The DNA was isolated for the BAC clone and was prepared for sequencing using a Large Construct kit® (Qiagen, Valencia, CA) following the manufacturer's instructions. The extracted BAC DNA was prepared for sequencing using GS-FLX Titanium Technology® (Roche, Indianapolis, IN) following manufacturer's instructions. Sequencing reactions were performed using a physically sectored GS-FLX TI Pico-titer plate® with the BACs pooled in pairs for optimal data output. The BACs were combined in pairs where the FAD2 gene was paired with a FAD3 gene. All generated sequence data was assembled by Newbler v2.0.01.14® (454 Life Sciences, Branford, CT). The assembled contigs were manually assessed for the presence of the corresponding FAD gene using Sequencher v3.7® (GeneCodes, Ann Arbor, MI).

[0216] After the full genomic sequence of all six FAD3 genes had been identified and fully characterized, zinc finger nucleases were designed to bind to the sequences for each specific gene family member.

EXAMPLE 2: DESIGN OF ZINC FINGER BINDING DOMAINS SPECIFIC TO FAD3 GENES

[0217] Zinc finger proteins directed against DNA sequences encoding various functional sequences of the FAD3 gene locus were designed as previously described. *See, e.g.,* Urnov *et al.* (2005) Nature 435:646-651. Exemplary target sequence and recognition helices are shown in Table 3 (recognition helix regions designs) and Table 4 (target sites). In Table 4,

nucleotides in the target site that are contacted by the ZFP recognition helices are indicated in uppercase letters; non-contacted nucleotides indicated in lowercase. Zinc finger nuclease (ZFN) target sites were designed to bind seven target sites of FAD3. The FAD3 zinc finger designs were incorporated into zinc finger expression vectors encoding a protein having at least one finger with a CCHC structure. *See*, U.S. Patent Publication No. 2008/0182332. In particular, the last finger in each protein had a CCHC backbone for the recognition helix. The non-canonical zinc finger-encoding sequences were fused to the nuclease domain of the type IIS restriction enzyme *Fok*I (amino acids 384-579 of the sequence of Wah *et al.*, (1998) Proc. Natl. Acad. Sci. USA 95:10564-10569) via a four amino acid ZC linker and an *opaque-2* nuclear localization signal derived from *Zea mays* to form FAD3 zinc-finger nucleases (ZFNs). Expression of the fusion proteins was driven by a relatively strong constitutive promoter such as a promoter derived from the Cassava Vein Mosaic Virus (CsVMV) promoter and flanked by the *Agrobacterium tumefaciens* ORF23 3'UnTranslated Region (AtuORF23 3'UTR v1). The self-hydrolyzing 2A encoding nucleotide sequence from *Thosea asigna* virus (Szymczak *et al.*, 2004) was added between the two ZFNs that were cloned into the construct. Exemplary vectors are described below.

[0218] The optimal zinc fingers were verified for cleavage activity using a budding yeast based system previously shown to identify active nucleases. *See, e.g.*, U.S. Patent Publication No. 20090111119; Doyon *et al.* (2008) *Nat Biotechnol.* 26:702–708; Geurts *et al.* (2009) *Science* 325:433. Zinc fingers for the various functional domains were selected for *in-vivo* use. Of the numerous ZFNs that were designed, produced and tested to bind to the putative FAD genomic polynucleotide target sites, fifteen ZFNs were identified as having *in vivo* activity at high levels, and selected for further experimentation. These ZFNs were characterized as being capable of efficiently binding and cleaving the unique FAD3 genomic polynucleotide target sites *in planta*.

Table 3: FAD3 Zinc Finger Designs

ZFP	sF1	F2	F3	F4	F5	F6
27961	RSDNLAR (SEQ ID NO:116)	QKKDRSY (SEQ ID	RSDNLAR (SEQ ID NO:116)	QRGNRNT (SEQ ID NO:119)	RSDHLSR (SEQ ID NO:120)	RNQDRTN (SEQ ID NO:121)

		NO:117)				
27962	DRSNLSR (SEQ ID NO:122)	RQDSRSQ (SEQ ID NO:123)	QSSDLSR (SEQ ID NO:124)	DRSALAR (SEQ ID NO:125)	TSGSLTR (SEQ ID NO:126)	N/A
27973	QSSDLSR (SEQ ID NO:124)	AASNRSK (SEQ ID NO:128)	TSGSLSR (SEQ ID NO:129)	RSDALAR (SEQ ID NO:130)	RSDVLST (SEQ ID NO:131)	WGRLRKL (SEQ ID NO:132)
27974	ERGTLAR (SEQ ID NO:133)	RSDDLTR (SEQ ID NO:134)	RSDHLSA (SEQ ID NO:135)	QHGALQT (SEQ ID NO:136)	TSGNLTR (SEQ ID NO:137)	QSGHLSR (SEQ ID NO:138)
27987	TSGSLTR (SEQ ID NO:126)	RSDHLSQ (SEQ ID NO:140)	CTRNRWR (SEQ ID NO:141)	RSDNLSE (SEQ ID NO:142)	ASKTRKN (SEQ ID NO:143)	N/A
27990	TSGSLSR (SEQ ID NO:129)	TSSNRAV (SEQ ID NO:145)	TSGNLTR (SEQ ID NO:137)	DRSALAR (SEQ ID NO:125)	RSDVLSE (SEQ ID NO:148)	RNFSLTM (SEQ ID NO:149)
27991	QSGDLTR (SEQ ID NO:150)	TSGSLSR (SEQ ID NO:129)	QSGNLAR (SEQ ID NO:152)	TSGSLSR (SEQ ID NO:129)	QSGSLTR (SEQ ID NO:154)	N/A
27992	DRSHLAR (SEQ ID NO:155)	TSGSLSR (SEQ ID NO:129)	TSSNRAV (SEQ ID NO:145)	TSGNLTR (SEQ ID NO:137)	DRSALAR (SEQ ID NO:125)	N/A
28004	QSGNLAR (SEQ ID NO:152)	HLGNLKT (SEQ ID NO:161)	RSDHLSQ (SEQ ID NO:140)	TARLLKL (SEQ ID NO:163)	QSGNLAR (SEQ ID NO:152)	QTSHLPQ (SEQ ID NO:165)
28005	RSDNLSV (SEQ ID	TSGHLSR (SEQ ID	TSGSLTR (SEQ ID	RSDALST (SEQ ID	DRSTRTK (SEQ ID	N/A

	NO:166)	NO:167)	NO:126)	NO:169)	NO:170)	
28021	QNAHRKT (SEQ ID NO:171)	TSGNLTR (SEQ ID NO:137)	LKQMLAV (SEQ ID NO:173)	RSDNLSR (SEQ ID NO:174)	DNSNRKT (SEQ ID NO:175)	N/A
28022	RSDNLSV (SEQ ID NO:166)	QNANRIT (SEQ ID NO:177)	TSGSLSR (SEQ ID NO:129)	QSSVRNS (SEQ ID NO:179)	DRSALAR (SEQ ID NO:125)	N/A
28023	RSDNLSR (SEQ ID NO:174)	DNSNRKT (SEQ ID NO:175)	DRSNLTR (SEQ ID NO:183)	RSDVLSE (SEQ ID NO:148)	TRNGLKY (SEQ ID NO:185)	N/A
28024	RSDALAR (SEQ ID NO:130)	RSDVLSE (SEQ ID NO:148)	RSSDRTK (SEQ ID NO:188)	RSDNLSV (SEQ ID NO:166)	QNANRIT (SEQ ID NO:177)	N/A
28025	QSSDLSR (SEQ ID NO:124)	QSTHRNA (SEQ ID NO:192)	RSDNLAR (SEQ ID NO:116)	QRGNRNT (SEQ ID NO:119)	RSDHLSR (SEQ ID NO:120)	RNQDRTN (SEQ ID NO:121)
28026	DRSNLSR (SEQ ID NO:122)	RQDSRSQ (SEQ ID NO:123)	QSSDLSR (SEQ ID NO:124)	DRSALAR (SEQ ID NO:125)	TSGSLTR (SEQ ID NO:126)	N/A
28035	QSSDLSR (SEQ ID NO:124)	AASNRSK (SEQ ID NO:128)	TSGSLSR (SEQ ID NO:129)	RSDALAR (SEQ ID NO:130)	RSDTLSQ (SEQ ID NO:206)	QRDHRIK (SEQ ID NO:207)
28036	RSDDLTR (SEQ ID NO:134)	QSSDLRR (SEQ ID NO:209)	RSDHLSA (SEQ ID NO:135)	QHGALQT (SEQ ID NO:136)	TSGNLTR (SEQ ID NO:137)	QSGHLSR (SEQ ID NO:138)
28039	TSGSLSR (SEQ ID	RSDALAR (SEQ ID	RSDTLSQ (SEQ ID	QRDHRIK (SEQ ID	TSGNLTR (SEQ ID	DRGDLRK (SEQ ID

	NO:129)	NO:130)	NO:206)	NO:207)	NO:137)	NO:219)
28040	DSSDRKK (SEQ ID NO:220)	TSGNLTR (SEQ ID NO:137)	DNYNRAK (SEQ ID NO:222)	DRSHLTR (SEQ ID NO:223)	RSDNLTT (SEQ ID NO:224)	N/A
28051	RSDNLSN (SEQ ID NO:225)	TSSSRIN (SEQ ID NO:226)	RSDNLSE (SEQ ID NO:142)	ASKTRKN (SEQ ID NO:143)	RSDALTQ (SEQ ID NO:229)	N/A
28052	RSDTLST (SEQ ID NO:230)	DRSSRIK (SEQ ID NO:231)	RSDDLSK (SEQ ID NO:232)	DNSNRIK (SEQ ID NO:233)	N/A	N/A
28053	QSSDLSR (SEQ ID NO:124)	QAGNLSK (SEQ ID NO:235)	QSGDLTR (SEQ ID NO:150)	TSGSLSR (SEQ ID NO:129)	QSGNLAR (SEQ ID NO:152)	N/A
28054	TSGSLSR (SEQ ID NO:129)	LRQTLRD (SEQ ID NO:240)	TSGNLTR (SEQ ID NO:137)	DRSALAR (SEQ ID NO:125)	RSDVLSE (SEQ ID NO:148)	RNFSLTM (SEQ ID NO:149)
28055	QSGDLTR (SEQ ID NO:150)	TSGSLSR (SEQ ID NO:129)	QSGNLAR (SEQ ID NO:152)	TSGSLSR (SEQ ID NO:129)	QSGSLTR (SEQ ID NO:154)	N/A
28056	DRSHLAR (SEQ ID NO:155)	TSGSLSR (SEQ ID NO:129)	LRQTLRD (SEQ ID NO:240)	TSGNLTR (SEQ ID NO:137)	DRSALAR (SEQ ID NO:125)	N/A

Table 4: Target Sites of FAD3 Zinc Fingers

ZFP	Target Site (5' to 3')	SEQ ID NO:
27961	cgCCGGAGAAAGAGAGAGAGctttgagg	SEQ ID NO:20
27962	tgGTTGTCGCTATGGACcagcgtagcaa	SEQ ID NO:21
27969	tcTCCGTTcGCATTGcTACGCTggtcca	SEQ ID NO:22
27970	gaAAGGTTtGATCCGAGCGCAcaaccac	SEQ ID NO:23
27973	tcTCCGTTcGCATTGcTACGCTggtcca	SEQ ID NO:22
27974	tcGGAGATATAAGGGCGGCCattcctaa	SEQ ID NO:25
27987	taGCCCAGAACAGGGTTccttgggcggc	SEQ ID NO:26
27988	ctTCGTACTCGGCCACGactggtaattt	SEQ ID NO:27
27989	ttGAAGTTGCAaTAAGCTttctctcgct	SEQ ID NO:28
27990	acTTGCTGGTCGATCATGTTggccactc	SEQ ID NO:29
27991	aaGTAGTTGAAGTTGCAataagctttct	SEQ ID NO:30
27992	tgGTCGATCATGTTGGCcactcttgttt	SEQ ID NO:31
28004	aaCGAGAATGAAGGAATGAAgaatatga	SEQ ID NO:32
28005	atACCATGGTTGGTAAGtcatttatttt	SEQ ID NO:33
28021	ccAACGAGgAATGATAGAtaaacaagag	SEQ ID NO:34
28022	caGTCACAGTTcTAAAAGtctatggtgt	SEQ ID NO:35
28023	tgTGACTGGACcAACGAGgaatgataga	SEQ ID NO:36
28024	tcTAAAAGTCTATGGTGttccttacatt	SEQ ID NO:37
28025	cgCCGGAGAAAGAGAGAGCTttgaggga	SEQ ID NO:38

28026	tgGTTGTCGCTATGGACcagcgtagcaa	SEQ ID NO:21
28035	ctTAAACGGTGGTTgTGCGCTcggatca	SEQ ID NO:40
28036	tcGGAGATATAAGGGCTGCGattcctaa	SEQ ID NO:41
28039	tcTCCGATctTAAACGGTGGTTgtgcgc	SEQ ID NO:42
28040	atAAGGGCTGCGATTCCtaagcattgtt	SEQ ID NO:43
28051	agATGGCCCAGAAAAGGgttccttgggc	SEQ ID NO:44
28052	cgTACTCGGCCACGactggtaatttaat	SEQ ID NO:45
28053	ttGAAGTTGCAaTAAGCTttctctcgct	SEQ ID NO:28
28054	acTTGCTGGTCGATCGTGTTggccactc	SEQ ID NO:47
28055	aaGTAGTTGAAGTTGCAataagctttct	SEQ ID NO:30
28056	tgGTCGATCGTGTTGGCcactcttgttt	SEQ ID NO:49

EXAMPLE 3: EVALUATION OF ZINC FINGER NUCLEASE CLEAVAGE OF FAD3 GENES

CONSTRUCT ASSEMBLY

[0219] Plasmid vectors containing ZFN expression constructs of the exemplary zinc finger nucleases, which were identified using the yeast assay, as described in Example 2, were designed and completed using skills and techniques commonly known in the art. Each zinc finger-encoding sequence was fused to a sequence encoding an opaque-2 nuclear localization signal (Maddaloni *et al.* (1989) *Nuc. Acids Res.* 17(18):7532), that was positioned upstream of the zinc finger nuclease.

[0220] Next, the opaque-2 nuclear localization signal::zinc finger nuclease fusion sequence was paired with the complementary opaque-2 nuclear localization signal::zinc finger nuclease fusion sequence. As such, each construct included a single open reading frame comprised of two opaque-2 nuclear localization signal::zinc finger nuclease fusion sequences separated by the 2A sequence from *Thosea asigna* virus (Mattion *et al.* (1996) *J. Virol.* 70:8124-

8127). Expression of the fusion proteins was driven by a relatively strong constitutive promoter such as a promoter derived from the Cassava Vein Mosaic Virus (CsVMV) promoter and flanked by the *Agrobacterium tumefaciens* ORF23 3'UnTranslated Region (AtuORF23 3'UTR).

[0221] The vectors were assembled using the IN-FUSION™ Advantage Technology (Clontech, Mountain View, CA). Restriction endonucleases were obtained from New England BioLabs (NEB; Ipswich, MA) and T4 DNA Ligase (Invitrogen) was used for DNA ligation. Plasmid preparations were performed using NUCLEOSPIN® Plasmid Kit (Macherey-Nagel Inc., Bethlehem, PA) or the Plasmid Midi Kit (Qiagen) following the instructions of the suppliers. DNA fragments were isolated using QIAquick Gel Extraction Kit™ (Qiagen) after agarose Tris-acetate gel electrophoresis. Colonies of all assembled plasmids were initially screened by restriction digestion of miniprep DNA. Plasmid DNA of selected clones was sequenced by a commercial sequencing vendor (Eurofins MWG Operon, Huntsville, AL). Sequence data were assembled and analyzed using the SEQUENCHER™ software (Gene Codes Corp., Ann Arbor, MI). Before delivery to *B. napus* protoplasts, Plasmid DNA was prepared from cultures of *E. coli* using the Pure Yield Plasmid Maxiprep System® (Promega Corporation, Madison, WI) or Plasmid Maxi Kit® (Qiagen, Valencia, CA) following the instructions of the suppliers.

[0222] The resulting eleven plasmid constructs; pDAB107824 (ZFNs 28025-2A-28026), pDAB107815 (ZFNs 27961-2A-27962), pDAB107816 (ZFNs 27969-2A-27970), pDAB107817 (ZFNs 27973-2A-27974), pDAB107825 (ZFNs 28035-2A-28036), pDAB107826 (ZFNs 28039-2A-28040), pDAB107818 (ZFNs 27987-2A-27988), pDAB107827 (ZFNs 28051-2A-28052), pDAB107821 (ZFNs 28004-2A-28005), pDAB107819 (ZFNs 27989-2A-27990), pDAB107828 (ZFNs 28053-2A-28054) (Figure 3), pDAB107829 (ZFNs 28055-2A-28056) (Figure 4), pDAB107820 (ZFNs 27991-2A-27992), pDAB107822 (ZFNs 28021-2A-28022) and pDAB107823 (ZFNs 28023-2A-28024) were confirmed via restriction enzyme digestion and via DNA sequencing.

PREPARATION OF DNA FOR TRANSFECTION

[0223] Plasmid DNA of the above described vectors was sterilized by precipitation, washed in 100% (v/v) ethanol, and dried in a laminar flow hood. The DNA pellet was suspended in 30 µL of sterile double-distilled water at a final concentration of 0.7 µg/µl for transfection into protoplast cells as described below. The preparation of the plasmid DNA was undertaken to result in supercoiled plasmid DNA for transient transfection and linearized plasmid DNA for stable transfection. The addition of carrier DNA (e.g. fish-sperm DNA) to the

transforming plasmid was not required for the transient transfection of protoplast cells. For transient studies about 30 μg of plasmid DNA per 10^6 protoplasts was used per transformation.

TRANSFECTION

[0224] Transfection of *Brassica napus* L. var. DH10275 was completed as described in Spangenberg *et al.*, (1986) Plant Physiology 66: 1-8, the media formulations are described in Spangenberg G. and Protrykus I. (1995) Polyethylene Glycol-Mediated Direct Gene Transfer in Tobacco Protoplasts. In: *Gene Transfer to Plants*. (Protrykus I. and Spangenberg G. Eds.) Springer-Verlag, Berlin. *Brassica napus* seeds were surface sterilized in 70% ethanol. The seeds were immersed in 12 mL of the 70% ethanol solution and mixed by gently rocking the cocktail for 10 minutes. The 70% ethanol solution was removed by decanting the solution and exchanged with a seed sterilization solution of 1% w/v calcium hypochlorite and 0.1% v/v Tween-20. The seeds were immersed in the seed sterilization solution and mixed by gently rocking the cocktail for 25 minutes. The seed sterilization solution was decanted and the sterilized seeds were rinsed three times in 50 mL of sterile water. Finally, the seeds were transferred to a sterile 80 mm Whatman filter paper disc® (Fisher-Scientific, St. Louis, MO) that had been laid within a Petri dish and the seeds were lightly saturated with sterile water. The Petri dish was sealed with Parafilm® (Fisher-Scientific, St. Louis, MO) and the plates were incubated at 25°C under complete darkness for one to two days. After signs of seedling emergence were observed from the seeds, the seedlings were transferred to Petri dish containing solidified GEM medium to encourage further seed germination. The seedlings were incubated on the GEM medium at 25°C for four to five days.

[0225] A volume of liquid PS medium (about 10 mL) was decanted into a sterile Petri dish. Using sterile forceps and a scalpel, an aerial portion of the four to five day old seedling in the 4-leaf stage of growth and development, was removed and discarded. Hypocotyl segments in lengths of 20-40 mm were determined to produce the highest population of small, cytoplasmic-rich protoplasts. The hypocotyl segments were aseptically excised and transferred to liquid PS medium. The excised hypocotyl segments were grouped together and cut transversely into 5-10 mm segments. Next, the hypocotyl segments were transferred to fresh PS medium and incubated at room temperature for 1 hour. The plasmolysed hypocotyls were transferred to a Petri dish containing enzyme solution. Care was taken to immerse all of the hypocotyl segments into the solution. The Petri dishes were sealed with Parafilm® and incubated overnight for sixteen to eighteen hours at 20 - 22°C with gentle rocking.

[0226] Protoplast cells were released from the hypocotyl segments. The overnight hypocotyl digests were gently agitated to release protoplasts into the enzyme solution. The Petri dish was angled slightly to aid the transfer of the digesting suspension of enzyme solution and plant debris. Using a 10 mL pipette the digesting suspension was transferred to a sterilized protoplast filtration (a filter of 100 micron mesh) unit to further separate the protoplasts from the plant debris. The filtration unit was tapped gently to release the excess liquid that had been caught in the sieve. The protoplast suspension, about 8 to 9 mL, was gently mixed and distributed into 14 mL sterile plastic round-bottomed centrifuge tubes. Each suspension was overlaid with 1.5 mL of W5 solution. The W5 solution was carefully dispensed over the protoplast suspension at an angle and dispensed drop-by-drop with minimal agitation. The addition of the W5 solution to the protoplast suspension resulted in the production of a protoplast rich interface. This interface was collected using a pipette. Next, the collected protoplasts were transferred into a new 14 mL centrifuge tube, and gently mixed. The yield or obtained protoplasts were determined using a haemocytometer to determine the number of protoplasts per milliliter. The method was repeated, wherein leaf tissue was digested to produce mesophyll protoplasts.

[0227] Next, W5 solution was added to a volume of 10 mL and the protoplasts were pelleted at 70g, before removing the W5 solution. The remaining protoplast suspension was resuspended by gentle shaking. Each tube containing the protoplast suspension was filled with 5 mL of W5 solution and incubated at room temperature from one to four hours. The protoplast suspensions were pelleted at 70g, and all of the W5 solution was removed. Next, 300 μL of transformation buffer was added to each of the pelleted protoplast suspensions which contained the isolated protoplasts. To each of the tubes, 10 μg of plasmid DNA was added to the protoplast suspensions. The plasmid DNA included the zinc finger nuclease constructs described above. Next, 300 μL of pre-warmed PEG 4000 solution was added to the protoplast suspension and the tubes were gently tapped. The protoplast suspensions and transformation mixture was allowed to incubate at room temperature for fifteen minutes without any agitation. An additional 10 mL of W5 solution was added to each tube in sequential aliquots of 1 mL, 1 mL, 1 mL, 2 mL, 2 mL, and 3 mL with gentle inversion of the tubes between each addition of W5 solution. The protoplasts were pelleted by spinning in a centrifuge at 70g. All of the W5 solution was removed leaving a pure protoplast suspension.

[0228] Next, 0.5 mL of K3 medium was added to the pelleted protoplast cells and the cells were resuspended. The resuspended protoplast cells were placed in the center of a Petri

dish and 5 mL of K3 and 0.6 mL Sea Plaque™ agarose (Cambrex, East Rutherford, NJ) in a 1:1 concentration. The Petri dishes were shaken in a single gentle swirling motion and left to incubate for 20-30 minutes at room temperature. The Petri dishes were sealed with Parafilm® and the protoplasts were cultured for twenty-four hours in complete darkness. After the incubation in darkness, the Petri dishes were cultured for six days in dim light (5 μMol m^{-2} s^{-1} of Osram L36 W/21 Lumilux white tubes). After the culture step, a sterile spatula was used to divide the agarose containing the protoplasts into quadrants. The separated quadrants were placed into a 250 mL plastic culture vessel containing 20 mL of A medium and incubated on a rotary shaker at 80 rpm and 1.25 cm throw at 24 °C in continuous dim light for 14 days and then analyzed to determine the level of activity of each ZFN construct.

GENOMIC DNA ISOLATION FROM CANOLA PROTOPLASTS

[0229] Transfected protoplasts were supplied in individual 1.5 or 2.0 mL microfuge tubes. The cells were pelleted at the base of the tube in a buffer solution. DNA extraction was carried out by snap freezing the cells in liquid nitrogen followed by freeze drying the cells, for about 48 hours in a Labconco Freezone 4.5® (Labconco, Kansas City, MO) at -40 °C and about 133 x 10^{-3} mBar pressure. The lyophilized cells were subjected to DNA extraction using the DNeasy® (QIAGEN, Carlsbad, CA) plant kit following manufactures instructions, with the exception that tissue disruption was not required and the protoplast cells were added directly to the lysis buffer.

TESTING OF FAD3A and FAD3C ZFN's FOR GENOMIC DNA SEQUENCE CLEAVAGE IN CANOLA PROTOPLASTS

[0230] The design of the ZFN target sites within the FAD3A and FAD3C gene locus were clustered, so that multiple pairs of ZFN were design to overlapping target sites. The clustering of ZFN target sites enabled PCR primers to be designed that would amplify the flanking genomic sequence from all FAD3A and FAD3C gene family members within a 100 bp window so as to encompass all of the overlapping ZFN target sites. As such, the Illumina short read sequence technology could be used to assess the integrity of the target ZFN site of the transfected protoplasts. In addition, the PCR primers designed needed to include specific nucleotide bases that would attribute sequence reads to the specific gene member of the FAD3A and FAD3C gene family. Therefore, all of the PCR primers would be required to bind 5-10 nucleotides away from any ZFN target cut site as non-homologous end joining (NHEJ) activity is known to cause small deletions that could remove a priming site to inhibit amplification and therefore distort the assessment of NHEJ activity.

[0231] Primers were designed to bind to all of the ZFN target loci for the FAD3A and FAD3C gene families (Table 5) and were empirically tested for amplification of all gene family members through Sanger based sequencing of PCR amplification products. In several instances primers could not be developed that would distinguish all gene family members (Table 6), however in all instances the target gene sequences of FAD3A or FAD3C, could be distinguished. Following PCR primer design custom DNA barcode sequences were incorporated into the PCR primers that were used to distinguish the different ZFN target loci and identify specific sequence reads to a transfection and ZFN (Tables 5 and 6).

Table 5: Amplification performance of the designed PCR primers on the FAD3 gene families. An "X" indicates gene copy detection specificity, grey shading and an "+" indicates that at the specific locus in question the sequence reads designed by the two primers were unable to be distinguished and an "N/A" indicates that the locus was unable to be amplified from those specific gene copies

ZFN Locus	FAD Gene Copy					
	FAD3A	FAD3C	FAD3A'	FAD3C'	FAD3A''	FAD3C''
Locus 1	X	X	X	X	X	X
Locus 2	X	X	X	X	N/A	X
Locus 3	X	X	+	+	X	X
Locus 4	X	X	X	X	+	+
Locus 5	X	X	N/A	N/A	N/A	N/A
Locus 6	X	X	X	X	X	X
Locus 7	X	X	X	X	X	X

Table 6: Primer sequences designed for FAD3 ZFN assessment of activity. Primers include custom barcodes, along with both requisite illumina adaptor sequences for construction of illumina library for sequencing-by-synthesis analysis. Purchased primer was the sum of all three columns presented

Locus ID	**SEQ ID NO:**	**Illumina Adaptor Primer Sequence *Barcode* Locus Primer**
FAD3_ZFN_Locus1A_F3	50	ACACTCTTTCCCTACACGACGCTCTTCCGATCT *ACGTA* CCTTTCTTCACCACATTYCA
FAD3_ZFN_Loc	50	ACACTCTTTCCCTACACGACGCTCTTCCGATCT *CGTAC*

us1B_F3		CCTTTCTTCACCACATTYCA
FAD3_ZFN_Loc us2C_F1	50	ACACTCTTTCCCTACACGACGCTCTTCCGATCT*CTGAC* GATGGTTGTCGCTATGGACC
FAD3_ZFN_Loc us3D_F1	50	ACACTCTTTCCCTACACGACGCTCTTCCGATCT*TGACTC* GAAAGGTTTGATCCRAGCG
FAD3_ZFN_Loc us3E_F1	50	ACACTCTTTCCCTACACGACGCTCTTCCGATCT*GACTG* CGAAAGGTTTGATCCRAGCG
FAD3_ZFN_Loc us3F_F1	50	ACACTCTTTCCCTACACGACGCTCTTCCGATCT*ACTGAC* GAAAGGTTTGATCCRAGCG
FAD3_ZFN_Loc us4G_F1	50	ACACTCTTTCCCTACACGACGCTCTTCCGATCT*GCTAG* CCGTGTATTTTGATAGCTGGTTC
FAD3_ZFN_Loc us4H_F1	50	ACACTCTTTCCCTACACGACGCTCTTCCGATCT*CTAGCC* CGTGTATTTTGATAGCTGGTTC
FAD3_ZFN_Loc us5J_F1	50	ACACTCTTTCCCTACACGACGCTCTTCCGATCT*TAGCTG* GAGCTTCTCAGACATTCCTCT
FAD3_ZFN_Loc us6K_F1	50	ACACTCTTTCCCTACACGACGCTCTTCCGATCT*TCAGTG* TTTATTTGCCCCAAGCGAGAG
FAD3_ZFN_Loc us6L_F1	50	ACACTCTTTCCCTACACGACGCTCTTCCGATCT*CAGTC* GTTTATTTGCCCCAAGCGAGAG
FAD3_ZFN_Loc us6M_F1	50	ACACTCTTTCCCTACACGACGCTCTTCCGATCT*AGTCAG* TTTATTTGCCCCAAGCGAGAG
FAD3_ZFN_Loc us6N_F1	50	ACACTCTTTCCCTACACGACGCTCTTCCGATCT*GTCAG* GTTTATTTGCCCCAAGCGAGAG
FAD3_ZFN_Loc us7P_F3	50	ACACTCTTTCCCTACACGACGCTCTTCCGATCT*GTACG* ACTTCAACTACTTGCTGGTCSAT
FAD3_ZFN_Loc us7Q_F3	50	ACACTCTTTCCCTACACGACGCTCTTCCGATCT*TACGTA* CTTCAACTACTTGCTGGTCSAT
FAD3_ZFN_Loc us1A_R1	65	CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT*A CGTA*CGTTCACATTGSTRCGYTGG
FAD3_ZFN_Loc us1B_R1	65	CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT*C GTAC*CGTTCACATTGSTRCGYTGG

FAD3_ZFN_Locus2C_R1	65	CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCTC *TGAC*CCGATCTTAAACGGYGGTTGT
FAD3_ZFN_Locus3D_R1	65	CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCTT *GACT*TAGCTCATGGATCTCAAAGGACT
FAD3_ZFN_Locus3E_R1	65	CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCTG *ACTG*TAGCTCATGGATCTCAAAGGACT
FAD3_ZFN_Locus3F_R1	65	CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCTA *CTGA*TAGCTCATGGATCTCAAAGGACT
FAD3_ZFN_Locus4G_R_uni	65	CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCTG *CTAG*TTAAATTACCAGTCGTGGCC
FAD3_ZFN_Locus4H_R_uni	65	CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCTC *TAGC*TTAAATTACCAGTCGTGGCC
FAD3_ZFN_Locus5J_R2	65	CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCTT *AGCT*CTTTTTTCTTCGATKCTAAAGATT
FAD3_ZFN_Locus6K_R1	65	CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCTT *CAGT*CTGTGACTGGACCAACGAGG
FAD3_ZFN_Locus6L_R1	65	CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCTC *AGTC*CTGTGACTGGACCAACGAGG
FAD3_ZFN_Locus6M_R1	65	CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCTA *GTCA*CTGTGACTGGACCAACGAGG
FAD3_ZFN_Locus6N_R1	65	CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCTG *TCAG*CTGTGACTGGACCAACGAGG
FAD3_ZFN_Locus7P_R1	65	CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCTG *TACG*ACTTACAATGTAAGGAACRCCRTA
FAD3_ZFN_Locus7Q_R1	65	CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCTT *ACGT* ACTTACAATGTAAGGAACRCCRTA

[0232] Following DNA extraction of canola protoplasts transfected with the ZFN, PCR amplification of the target ZFN loci was performed to generate the requisite loci specific DNA molecules in the correct format for Illumina based sequencing by synthesis technology. Each assay was optimised to work on 25 ng starting DNA (about 12,500 cell equivalents of the

Brassica napus genome). Multiple reactions were performed, per sample to provide the coverage required to assess NHEJ efficiency and specificity at the appropriate level, about sixteen PCR reactions equivalent to 200,000 copies of the *Brassica napus* genome taken from individual protoplasts. PCR amplification master-mixes were made for all samples to be tested with the same assay and one reaction, performed in triplicate, was assayed using a quantitative PCR method that was used to determine the optimal number of cycles to perform on the target tissue, to ensure that PCR amplification had not become reagent limited and was still in an exponential amplification stage. The experimentation with the necessary negative control reactions was performed in 96 well format using a MX3000P thermocycler® (Stratagene, LaJolla, CA).

[0233] From the output gathered from the quantitative PCR platform, the relative increase in fluorescence was plotted from cycle-to-cycle and the cycle number was determined per assay that would deliver sufficient amplification, while not allowing the reaction to become reagent limited, in an attempt to reduce over cycling and the amplification of common transcripts or molecules. The unused master mix, remained on ice until the quantitative PCR analysis was concluded and the cycle number determined and was then aliquoted into the desired number of reaction tubes (about 16 per ZFN assay) and the PCR reaction was performed.

[0234] Following amplification, samples for a single ZFN locus were pooled together and 200 µL of pooled product per ZFN was cleaned using the MinElute PCR purification kit® (Qiagen) following manufacturer's instructions. To enable the sample to be sequenced using the Illumina short read technology additional paired end primers were required to be attached by amplification onto the generated fragments. This was achieved by PCR amplification using primers that would be, in part complementary to the sequence added in the first round of amplification, but also contain the paired end sequence required. The optimal number of PCR cycles to perform, that would add the paired end sequences without over amplifying common fragments to the template was again determined using a sample pass through a quantitative PCR cycle analysis, as described previously.

[0235] Following PCR amplification, the generated product was cleaned using a MinElute column® (Qiagen) following manufacturer's instructions and was resolved on a 2.5% agarose gel. DNA fragments visualised using Syber® Safe (Life Technologies, Carlsbad, CA) as bands of the correct size were gel extracted to remove any residual PCR generated primer-dimer or other spurious fragments, the DNA was extracted from the gel slice using a MinElute gel

extraction kit® (Qiagen) following manufacturer's instructions. After completion of the gel extraction an additional clean up of the DNA was performed using AMPure magnetic beads® (Beckman-Coulter, Brea, CA) with a DNA to bead ratio of 1:1.7. The DNA was then assessed for concentration using a quantitative PCR based library quantification kit for Illumina sequencing (KAPA) with a 1/40,000 and a 1/80,000 dilution and with the reaction being performed in triplicate. Based on the quantitative PCR results the DNA was diluted to a standard concentration of 2 nM and all libraries were combined for DNA sequencing. The samples were prepared for sequencing using a cBot cluster generation kit® (Illumina, San Diego, CA) and were sequenced on an Illumina GA2x® with 100 bp paired-end sequencing reads following manufacturer's instructions.

METHOD OF DATA ANALYSIS FOR DETECTION OF NON-HOMOLOGOUS END JOINING AT TARGET ZINC FINGER SITES

[0236] Following completion of the sequencing reaction and primary data calling performed using the Illumina bioinformatic pipeline for base calling, full analysis was performed to identify deleted bases at the target ZFN site in each instance. A custom PERL script was designed to extract and sort barcodes from DNA sequences computationally following a list of input sequences. The barcode had to match the reference sequence at a Phred score of greater than 30 to be accepted, to reduce misattributing sequence reads. After the sequence reads had been binned into the different barcode groups that had been used, a quality filter was passed across all sequences. The quality filter was a second custom developed PERL script. Sequence reads were excluded if there were more than three bases called as "N", or if the median Phred score was less than 20, or if there were 3 consecutive bases with a Phred score of less than 20, or if the sequence read was shorter than 40 bp in length. The remaining sequences were merged where both of the paired sequence reads were available using the NextGENe® (SoftGenetics, State College, PA) package. The remaining merged sequence reads were then reduced to a collection of unique sequence reads using a third custom PERL script with a count of the number of redundant sequences that had been identified recorded on the end of the remaining sequence identifier. The unique sequence reads were then aligned to the FAD3 reference sequence using the NextGENe® software that created a gapped FASTA aligned file.

[0237] Using the gapped FASTA file a conversion of the gapped base position number to the input reference was performed using a fourth custom PERL script. This enabled bases that discriminate the different gene family members (either homoeologous or paralogous sequence variation between the different gene family members) to be identified in the assembled data.

Once the conversion of base numbering had been performed it was possible to generate haplotype reports for each unique sequence reads and assign the reads to specific gene family members. Once the reads had been grouped by gene a 10 bp window was identified and assessed that surrounded the ZFN target site. The number of sequences with deletions was recorded per gene along with the number of missing bases.

[0238] The data was then graphically displayed as a multiple line graph, with the number of sequences with 1 through 10 bases deleted at the target ZFN site per 10,000 sequence reads. This analysis was performed for all ZFN transfections along with control transfections. In several instances, repeats in the native DNA sequence lead to an increase in sequencing error in the target ZFN site, such an error can be commonly seen as an increase in the prevalence of single base deletions that were reported in all samples, both transfected with ZFN or controls.

[0239] From these results highest level of ZFN activity at a FAD3A and FAD3C target site was observed as determined by the greater activity of NHEJ. The ZFNs which were encoded on plasmid pDAB107828 (*i.e.,* ZFN28053 and 28054) and pDAB107829 (*i.e.,* ZFN28055 and 28056) were selected for *in planta* targeting of an Engineered Transgene Integration Platform (ETIP) given its characteristics of significant genomic DNA cleavage activity and minimal non-target activity.

EXAMPLE 4: DNA CONSTRUCTS FOR ENGINEERED TRANSGENE INTEGRATION PLATFORM (ETIP) CANOLA PLANT LINES

[0240] The plasmid vector constructs described below were built using methods and techniques commonly known by one with skill in the art. The application of specific reagents and techniques described within this paragraph are readily known by those with skill in the art, and could be readily interchanged with other reagents and techniques to achieve the desired purpose of building plasmid vector constructs. The restriction endonucleases were obtained from New England BioLabs (NEB; Ipswich, MA). Ligations were completed with T4 DNA Ligase (Invitrogen, Carlsbad, CA). Gateway reactions were performed using GATEWAY® LR CLONASE® enzyme mix (Invitrogen) for assembling one entry vector into a single destination vector. IN-FUSION™ reactions were performed using IN-FUSION™ Advantage Technology (Clontech, Mountain View, CA) for assembling one entry vector into a single destination vector Plasmid preparations were performed using NUCLEOSPIN® Plasmid Kit (Macherey-Nagel Inc., Bethlehem, PA) or the Plasmid Midi Kit® (Qiagen) following the instructions of the suppliers. DNA fragments were isolated using QIAquick Gel Extraction Kit™ (Qiagen) after

agarose Tris-acetate gel electrophoresis. Colonies of all assembled plasmids were initially screened by restriction digestion of miniprep DNA. Plasmid DNA of selected clones was sequenced by a commercial sequencing vendor (Eurofins MWG Operon, Huntsville, AL). Sequence data were assembled and analyzed using the SEQUENCHER™ software (Gene Codes Corp., Ann Arbor, MI).

CONTROL VECTORS

[0241] A control vector was used to develop a Fluorescence Activated Cell Sorting (FACS) cell based sorting method. Standard cloning methods were used in the construction of a control vector, pDAS000031 (Figure 10: T-strand insert as SEQ ID NO:85) including two gene expression cassettes. The first gene expression cassette contained the Cauliflower mosaic virus 19s promoter (CaMV 19S promoter; Shillito, *et al.*, (1985) *Bio/Technology* 3; 1099-1103):: hygromycin resistance gene (hph(HygR);US Patent No. 4,727,028) :: and the *Agrobacterium tumefaciens* Open Reading Frame 1 3'UnTranslated Region (AtORF1 terminator; Huang *et al.*, (1990) *J. Bacteriol.* 1990 *172*:1814-1822). The second gene expression cassette contained the *Arabidopsis thaliana* Ubiquitin 10 promoter (AtUbi10 promoter; Callis, *et al.*, (1990) *J. Biol. Chem.*, 265: 12486-12493):: dsRED (dsRED(D); US Patent No. 6,852,849) and an intron from Arabidopsis (intron #1; GenBank: AB025639.1) :: *Agrobacterium tumefaciens* Open Reading Frame 23 3'UnTranslated Region (AtORF23 terminator; US Patent No. 5,428,147) as an in-frame fusion with a *trans* orientation (*e.g.*, head to head orientation). The plasmid vector was assembled using the IN-FUSION™ Advantage Technology (Clontech, Mountain View, CA).

EXAMPLE 5: GENERATION OF ETIP CANOLA PLANT LINES

TRANSFORMATION OF *Brassica napus*

[0242] The ETIP constructs for the FAD3A and FAD3C site specific construct (pDAS000271 – pDAS000275) and accompanying ZFN (pDAB107828 and 107829) and the control the DS-Red control construct (pDAS000031) are previously described in Example 4. These binary vectors are transformed into *Agrobacterium tumefaciens* strain GV3101: PM90. Transformation of *Brassica napus* protoplast cells is completed using the transfection protocol described in Example 3 with some modification.

[0243] The modifications to the protocol include the use of sodium alginate instead of Sea Plaque™ agarose. The transfection experiments in which both the ZFN construct and the ETIP construct are co-delivered into *Brassica napus* protoplast cells are completed at DNA

concentrations comprising a 5:1 molar ratio of plasmid DNA. The other ETIP and control plasmid constructs are transformed at concentrations of 30 μg of plasmid DNA.

[0244] Additional modifications to the protocol include the propagation of whole plants from the transformed protoplast cells in medium containing 1.5 mg/mL of hygromycin. The propagation of whole plants requires that the A medium is replaced every two weeks and the growth of the protoplast-derived colonies is monitored. After the protoplast-derived colonies grow to approximately 2-3 mm in diameter, the colonies are transferred into individual wells of a 12-well Costar® plate (Fisher Scientific, St. Louis, MO) containing solidified MS morpho medium. The plates are incubated for one to two weeks at 24 °C under continuous dim light until the calli proliferate to a size of 8-10 mm in diameter. After the protoplast cells reach a diameter of 1-2 cm in diameter, the protoplast cells are transferred to individual 250 mL culture vessels containing MS morpho medium. The vessels are incubated at 24 °C under 16 h light (20 μMol m^{-2} s^{-1} of Osram L36 W/21 Lumilux white tubes) and 8 h dark conditions. Within one to two weeks, multiple shoots are visible. The shoots are transferred into 250 mL culture vessels containing MS medium after they reach a length of 3-4 cm. The 250 mL culture vessels are incubated at 24 °C under 16 h light (20 μMol m^{-2} s^{-1} of Osram L36 W/21 Lumilux white tubes) and 8h dark conditions. The shoots are maintained in the culture vessels until they develop into plantlets at which time they are transferred to a greenhouse to grow to maturity.

EXAMPLE 6: MOLECULAR CONFIRMATION OF INTEGRATION OF T-DNAS CONTAINING ETIPS IN CANOLA

[0245] Genomic DNA is extracted from leaf tissue of all putative transgenic plants using a DNeasy 96 Plant DNA extraction kit™ or a DNeasy Plant Mini Kit™ (Qiagen). The genomic DNA from each plant is analyzed by PCR using primers designed to amplify *vir*C from pTiC58 Forward (SEQ ID NO:88 CGAGAACTTGGCAATTCC) and pTiC58 Reverse (SEQ ID NO:89 TGGCGATTCTGAGATTCC) to test for persistence of *A.tumfaciens,* primers designed to amplify actin from *B. napus*; Actin Forward (SEQ ID NO:90 GACTCATCGTACTCTCCCTTCG) and Actin Reverse (SEQ ID NO:91 GACTCATCGTACTCTCCCTTCG) to check the quality of the genomic DNA. Primers are designed to amplify the *hph* gene; HPH Forward (SEQ ID NO:92 TGTTGGTGGAAGAGGATACG) and HPH Reverse (SEQ ID NO:93 ATCAGCAGCAGCGATAGC) encoded by the ETIP. Plants that do not give a product from

*vir*C primers, and that produce amplicons of the correct size when amplified with primers to actin and *hph* are confirmed as transgenic.

[0246] A second screen is completed, where gDNA from each transgenic plant is analysed by PCR using five sets of primers designed to amplify the binary vector outside of the T-DNA region [(1F SEQ ID NO:94 ATGTCCACTGGGTTCGTGCC; 1R SEQ ID NO:95 GAAGGGAACTTATCCGGTCC) (2F SEQ ID NO:96 TGCGCTGCCATTCTCCAAAT; 2R SE ID NO:97 ACCGAGCTCGAATTCAATTC) (3F SEQ ID NO:98 CCTGCATTCGGTTAAACACC; 3R SEQ ID NO:99 CCATCTGGCTTCTGCCTTGC) (4F SEQ ID NO:100 ATTCCGATCCCCAGGGCAGT; 4R SEQ ID NO:101 GCCAACGTTGCAGCCTTGCT) (5F SEQ ID NO:102 GCCCTGGGATGTTGTTAAGT; 5R SEQ ID NO:103 GTAACTTAGGACTTGTGCGA)]. Plants from which PCR products of the correct and expected size are amplified with primer sets 3 and 4 are considered to have backbone integration.

[0247] DNA from plants with no backbone integration is purified from 20 g of leaf tissue using a modified CTAB method (Maguire *et al,*. (1994) Plant Molecular Biology Reporter , 12(2): 106-109). The isolated gDNA is digested with several restriction enzymes and 10 μg of gDNA is separated by electrophoresis on an agarose gel and transferred to membrane using a standard Southern blotting protocol. Membranes are probed using the DIG Easy Hyb System™ (Roche, South San Francisco, CA) following the manufacturer's instructions. Probes to each expression cassette to the ELP and to an endogenous control gene, actin, are amplified from the ETIP construct using the following primers: (IPT-F SEQ ID NO:104 TCTCTACCTTGATGATCGG; IPT-R SEQ ID NO:105 AACATCTGCTTAACTCTGGC; dsRED-F SEQ ID NO:106 ATGGCTTCATCTGAGAACG; dsRED-R SEQ ID NO:107 TTCCGTATTGGAATTGAGG; PAT-F SEQ ID NO:108 TTGCTTAAGTCTATGGAGGCG; PAT-R SEQ ID NO:109 TGGGTAACTGGCCTAACTGG; ELP-F SEQ ID NO:110 ATGATATGTAGACATAGTGGG; ELP-R SEQ ID NO:111 AGGGTGTAAGGTACTAGCC; Hph-F SEQ ID NO:112 TGTTGGTGGAAGAGGATACG; Hph-R SEQ ID NO:113, ATCAGCAGCAGCGATAGC; actin-F SEQ ID NO:114 GTGGAGAAGAACTACGAGCTACCC; actin-R SEQ ID NO:115 GACTCATCGTACTCTCCCTTCG).

[0248] The ETIP sequence is amplified and sequenced from all plants containing only a single copy of the ETIP. The sequence of each T-DNA insert is analyzed by direct sequencing of PCR products using the ABI3730xI™ (Applied Biosystems, Life Teachnologies). The T-DNA

insert was amplified from genomic DNA, using Phusion Hot Start II Polymerase™ (Finnzymes, Thermo Fisher Scientific). The amplification reactions of the T-DNA are completed with multiple primer pairs to amplify overlapping sequences of approximately 2 Kbp in length. Each PCR product is sequenced with multiple primers to ensure complete coverage. The PCR reactions are treated with shrimp alkaline phosphatase and exonuclease I (Applied Biosystems, Life Technologies) to inactivate excess primer prior to the sequencing PCR reaction. The sequences flanking the T-DNA insert of each single copy ETIP line are identified by digestion of purified genomic DNA with eight restriction endonucleases separately followed by ligation of double-stranded adapters specific for the overhangs created by the restriction endonucleases. Following this ligation step a PCR is performed with a biotinylated primer to either the 3' or 5' end of the ETIP and a primer to each adapter. The PCR products are captures and cleaned on Ampure Solid Phase Reversible Immobilization (SPRI) beads™ (Agencourt Bioscience Corporation, Beckman Coulter Company). A nested PCR is performed and all products are sequenced using ABI Sanger Sequencing and Big Dye Terminator v3.1 cycle™ sequencing protocol (Applied Biosystems, Life Technologies). Sequence data are assembled and analyzed using the SEQUENCHER™ software (Gene Codes Corp., Ann Arbor, MI).

RESULTS OF ETIP TRANSGENIC CANOLA TRANSFORMED WITH ZINC FINGER NUCLEASE AND PDAS000271-PDAS000275 ETIP CONSTRUCTS

[0249] The transgenic *Brassica napus* events which are produced via transformation of ETIP and ZFN constructs result in the integration of a single copy, full length T-strand insertion of the ETIP polynucleotide sequence from pDAS000273 or pDAS275 within the FAD3A locus, and from pDAS000271, pDAS000272 or pDAS000274 into the FAD3C locus. Three to four events are fully characterized and confirmed to contain the integrated ETIP. The confirmation is completed using an in-out PCR amplification method, and further validated via Southern blot. The selected T_0 events are grown to the T_1 stage of development. The T_1 plants are rescreened to determine the zygosity of the integrated T-strand. Screened events are categorized as homozygous, hemizygous, or null.

[0250] The homozygous events are used to produce protoplasts via the previously described method. The protoplasts are subsequently co-transformed with a ZFN that is designed to target a zinc finger binding site which is incorporated within the ETIP sequence and a donor plasmid which shares homology with specific regions of the ETIP. The ZFN cleaves the ETIP locus and the donor plasmid is integrated within the genome of *Brassica napus* cells via homology directed repair. As a result of the integration of the donor plasmid, the partial *DS-red*

transgene is repaired to a full length *DS-red* transgene. The expression of the now fully operational *DS-red* transgene is used to sort protoplast cells with a FACS method. Putative transgenic plants are sorted using the FACS method described in Example 7 and the isolated protoplasts are regenerated into mature plants. The integration of the donor plasmid is confirmed within the ETIP-targeted plants using molecular confirmation methods. As such, the ETIP locus serves as a site-specific locus for gene targeted integration of a donor polynucleotide sequence.

EXAMPLE 7: FACS BASED SORTING OF PROTOPLAST CELLS

[0251] *Brassica napus* protoplasts that were transfected with the DS-Red control construct, pDAS000031, were sorted via FACS-mediated cell sorting using a BD Biosciences Influx-Cell sorter™ (San Jose, CA). The protoplast cells were isolated and transfected as described in Example 3. After the cells had been transfected with pDAS000031, the cells were sorted using the FACS sorter with the conditions described in Table 7.

Table 7: Conditions used for sorting protoplast cells transfected with pDAS000031

Parameters	
Drop frequency	6.1 KHz
Nozzle diameter	200μm
Sheath pressure	4 psi
Recovery media	W5 media
Culture conditions	Bead type culture using sea-plaque agarose and sodium alginate
Sort criteria	Sorting based on chlorophyll autofluorescence, reporter gene expression (Ds-Red)
Sort recovery (%)	50-75
Viability post sorting (%)	>95

[0252] The protoplasts which expressed the *DS-red* transgene were sorted and isolated. The FACS isolated protoplasts were counted using the sorter. About $1x10^5$ to $1.8x10^5$ of cells were placed in a well of a 24-well micro titer plate on the first day after the FACS isolation. The cells were transferred to a bead culture for 5 to 20 days. Similar conditions were tested, wherein about $1x10^4$ of cells were placed in a well of a 2 or 4-well micro titer plate on the second day after the FACS isolation. The various conditions that were tested resulted in the recovery of cells at a viability or 95 – 98% of the total isolated protoplast cells. The FACS sorted protoplast cells were transferred to a bead culture for 3 – 20 days. The FACS sorted protoplast cells were regenerated into plants on media which contained 1.5 mg/mL of hygromycin using the above described protocol. The putative transgenic plants were confirmed to contain an intact T-strand insert from pDAS000031 via molecular confirmation protocols.

[0253] The FACS sorting method is directly applicable to screen any fluorescent transgene sequence and is used to isolate a proportion of *Brassica napus* protoplast cells that are targeted with a fluorescent transgene via homology mediated repair within a specific site in the ETIP region within a genomic locus.

EXAMPLE 8: TARGETED INTEGRATION INTO AND DISRUPTION OF *BRASSICA NAPUS* OMEGA-3 FATTY ACID DESATURASE (*FAD3*) VIA NHEJ

SELECTION OF ZINC FINGER BINDING DOMAINS SPECIFIC TO FAD3C AND FAD3A

[0254] The transcribed regions for homoeologous Fad3 genes were identified and characterized, zinc finger nucleases that were designed to bind and cleave these sites for NHEJ-mediated targeting of a donor sequence. Zinc finger proteins (ZFPs) directed against DNA sequences from homeologues of Fad3 sequences were designed and tested as described above. From the ZFNs showing on-target activity, two zinc finger proteins were selected that cut the Fad3 target at high efficiency: ZFP 28051-2A-28052 recognizes SEQ ID NO:255 5'-gcccaaggaacCCTTTTCTGGGCCATcttcgTACTCGGCCACGactggtaatttaat -3' and was shown to specifically bind and cleave the Fad3C genomic locus. Likewise Zinc finger protein 28053-2A-28054 recognizes SEQ ID NO:256 5'-agcgagagaaAGCTTAtTGCAACTTCaactacTTGCTGGTCGATCGTGTTggccactc -3' and was shown to specifically bind and cleave the Fad3A and Fad3C genomic locus. Exemplary target sites are shown in Table 8; nucleotides in the target site that are contacted by the ZFP recognition helices are indicated in uppercase letters; non-contact nucleotides are indicated in lowercase. Nucleotides in copies of Fad3 that differ from Fad3C are identified by underlining.

Nucleotides in the target sites that are contacted by the ZFP recognition helices are shown in Table 8.

Table 8: Zinc Finger Protein Binding Sites specific to Fad3C (28051-2A-28052) or Fad3A and Fad3C (28053-2A-28054)

28051-2A-28052	*SEQ ID NO:*	*SEQ ID NO:257 gcccaaggaacCCTTTTCTGGGCCATct* *SEQ ID NO:45 cgTACTCGGCCACGactggtaatttaat*
Fad3C	259	GCCCAAGGAACCCTTTTCTGGGCCATCTTCGTACTCGGC CACGACTGGTAATTTAAT
Fad3A	260	GCCCAAGGAACCCTGTTCTGGGCTATCTTCGTACTCGGC CACGACTGGTAATTTAAT
Fad3C'	261	GCCCAAGGAACCCTTTTCTGGGCCATCTTCGTCCTCGGC CACGACTGGTAAAGTTTC
Fad3A'	261	GCCCAAGGAACCCTTTTCTGGGCCATCTTCGTCCTCGGC CACGACTGGTAAAGTTTC
Fad3A"	263	GCCCAAGGAACCCTTTTCTGGGCCATCTTCGTTCTTGGCC ACGACTGGTAAATTAAA
Fad3C"	263	GCCCAAGGAACCCTTTTCTGGGCCATCTTCGTTCTTGGCC ACGACTGGTAAATTAAA

28053-2A-28054	*SEQ ID NO:*	*SEQ ID NO:265 agcgagagaaAGCTTAtTGCAACTTCaa* *SEQ ID NO:47 acTTGCTGGTCGATCGTGTTggccactc*
Fad3C	256	AGCGAGAGAAAGCTTATTGCAACTTCAACTACTTGCTGG TCGATCGTGTTGGCCACTC
Fad3A	268	AGCGAGAGAAAGCTTATTGCAACTTCAACTACTTGCTGG TCGATCATGTTGGCCACTC
Fad3C'	269	AGCGAGAGAAAGCTTATTGCAACTTCAACTACTTGCTGG TCCATAATGTTGGCCATTC
Fad3A'	270	AGCGAGAGAAAGCTTATTGCAACTTCGACTACTTGCTGG TCCATAATGTTGGCAATTC
Fad3A"	271	AGCGAGAGGAAGCTTATTGCAACTTCAACAACTTGCTGG TCCATAATGTTGGCCACTC

Fad3C"	272	AGCGAGAGGAAGCTTATTGCAACTTCAACTACTTGCTGG TCCATAATGTTGGCCACTC

DESIGN AND CONSTRUCTION OF EXPRESSION VECTORS ENCODING ZINC FINGER NUCLEASES SPECIFIC TO FAD3C AND FAD3A

[0255] The Fad3 zinc finger designs were incorporated into zinc finger expression vectors encoding a protein having at least one finger with a CCHC structure (U.S. Patent Publication No. 2008/0182332). In particular, the last finger in each protein had a CCHC backbone for the recognition helix. The non-canonical, zinc finger-encoding-sequences were fused to the nuclease domain of the type IIS restriction enzyme FokI (amino acids 384-579 of the sequence of Wah et al., (1998) Proc. Natl. Acad. Sci. USA 95:10564-10569) via a four amino acid ZC linker and a sop2 nuclear localization signal. The self-hydrolyzing 2A encoding nucleotide sequence from Thosea asigna virus (Szymczak et al., 2004) was added between the two ZFN fusion proteins. Expression of the ZFNs was driven by the strong constitutive promoter and 5' untranslated region (UTR) from Cassava Vein Mosaic Virus (Verdaguer et al, Plant Molecular Biology 1996, 31(6); 1129-1139) and flanked by the 3' UTR (including the transcriptional terminator and polyadenylation site) from open reading frame 23 (ORF23) of Agrobacterium tumefaciens pTi15955 (Barker et al., Plant Molecular Biology 1983, 2(6); 335-50).

[0256] The vectors were assembled using the IN-FUSION™ Advantage Technology (Clontech, Mountain View, CA). Restriction endonucleases were obtained from New England BioLabs (NEB; Ipswich, MA) and T4 DNA Ligase (Invitrogen) was used for DNA ligation. Plasmid preparations were performed using NUCLEOSPIN® Plasmid Kit (Macherey-Nagel Inc., Bethlehem, PA) or the Plasmid Midi Kit™ (Qiagen) following the instructions of the suppliers. DNA fragments were isolated using QIAquick Gel Extraction Kit™ (Qiagen) after agarose Tris-acetate gel electrophoresis. Colonies of assembled plasmids were initially screened by restriction digestion of miniprep DNA. Plasmid DNA of selected clones was sequenced by a commercial sequencing vendor (Eurofins MWG Operon, Huntsville, AL). Sequence data were assembled and analyzed using the SEQUENCHER™ software (Gene Codes, Ann Arbor, MI). The resulting plasmid constructs: pDAB107827 (ZFN 28051-2A-28052, Figure 13, SEQ ID NO:273) and pDAB107828 (ZFN 28053-2A-28054, Figure 14, SEQ ID NO:274) were confirmed via restriction enzyme digestion and via DNA sequencing.

DESIGN AND CONSTRUCTION OF "DONOR" VECTORS FOR NHEJ-DIRECTED DNA REPAIR

[0257] Two strategies of integration of DNA into Fad3 were undertaken; gene splicing, where an expression cassette was integrated into a single ZFN-induced double-stranded break and gene-editing where a portion of the gene was removed by the use of two ZFN-induced double-stranded breaks and an expression cassette was inserted to repair the gap.

[0258] For each integration method, gene splicing or gene-editing, two vectors were constructed. The first encoded a turboGFP (tGFP) gene expression cassette and the second encoded a gene expression cassette to confer resistance to the antibiotic hygromycin. The tGFP expression cassette included the promoter, 5' untranslated region and intron from the Arabidopsis thaliana polyubiquitin 10 (UBQ10) gene (Norris et al, Plant Molecular Biology 1993, 21(5), 895-906) followed by the tGFP coding sequence (Evrogen, Moscow, Russia). The tGFP coding sequence was codon-optimised for expression in dicot plants and the 3' untranslated region (UTR) comprising the transcriptional terminator and polyadenylation site of open reading frame 23 (ORF23) of A. tumefaciens pTi15955 (Barker et al, Plant Molecular Biology 1983, 2(6), 335-50). The hygromycin resistance gene expression cassette included the 19S promoter including a 5' UTR from cauliflower mosaic virus (CaMV) (Cook and Penon Plant Molecular Biology 1990 14(3), 391-405) followed by the hygromycin phosphotransferase (hph) gene (Kaster et al Nucleic Acids Research 1983 11 (19), 6895-6911). The hph gene was codon-optimised for expression in dicotyledonous plants and was flanked by a 3'UTR comprising the transcriptional terminator and polyadenylation site of Open Reading Frame 1 (ORF1) of A. tumefaciens pTi15955 (Barker et al, Plant Molecular Biology 1983, 2(6), 335-50). Both cassettes were synthesized by a commercial gene synthesis vendor (GeneArt, Life Technologies, Regensberg, Germany).

[0259] Vectors for the gene splicing experiments were constructed by cloning two tandem copies of the ZFN recognition sequence targeted by the ZFN encoded in the vector pDAB10782. Vectors for the gene editing experiments were constructed by cloning one copy of each of the ZFN recognition sequences targeted by the ZFNs encoded in the vectors pDAB107827 and pDAB107828. In both cases the two ZFN recognition sequences were separated by the recognition sequences for BamHI and NotI restriction endonucleases. The tGFP and HPH cassettes were cloned into the BamHI and NotI sites of each vector resulting in four "donor" vectors: pDAS000340 (hygromycin-resistant gene-splicing donor: SEQ ID NO:275, Figure 15), pDAS000341 (tGFP reporter gene splicing donor: SEQ ID NO:276, Figure 16) ,

pDAS00342 (hygromycin-resistant gene-editing donor: SEQ ID NO:277, Figure 17) and pDAS000343 (tGFP reporter gene editing donor: SEQ ID NO:278, Figure 18).

[0260] Colonies of the assembled plasmids were initially screened by restriction endonuclease digestion of DNA purified from overnight cultures of E.coli. Restriction endonucleases were obtained from New England BioLabs™ (NEB, Ipswich, MA) and Promega™ (Promega Corporation, WI). Plasmid preparations were performed using the QIAprep Spin Miniprep Kit™ (Qiagen, Hilden, Germany) or the Pure Yield Plasmid Maxiprep System™ (Promega Corporation, WI) following the instructions of the suppliers. After the restriction fragments were confirmed by agarose gel electrophoresis of resulting fragments, plasmid DNA of selected clones were sequenced using ABI Sanger Sequencing and Big Dye Terminator v3.1™ cycle sequencing protocol (Applied Biosystems, Life Technologies). Sequence data were assembled and analyzed using the Sequencher™ software (Gene Codes, Ann Arbor, MI).

MAINTENANCE OF PLANT MATERIAL FOR PROTOPLAST ISOLATION

[0261] Mesophyll derived protoplasts were isolated from three-week old sterile shoot cultures of Brassica napus (DH10275). The corresponding seeds were germinated following the methods herein described. The seeds were surface-sterilized using 70% ethanol for 1 minute and gently shaken followed by 3-4 rinses in sterile double-distilled water. The seeds were subsequently sterilized using 20% bleach and 10 µl of Tween 20. The seeds were further treated with the bleach on a table top shaker at approximately 100 RPM, for 15 minutes followed by 3-4 rinses in sterile double-distilled water, seeds were carefully transferred to a sterile filter paper to remove the excess moisture and plated on seed germination medium (½ strength MS/B5 Vitamins + 1% sucrose + 0.8% Agar; pH 5.8.

[0262] Approximately, 50-60 mL of media was poured into each Petri™ dish (15 X 100 mm) and the plates were placed with a slight angle using a support. Approximately 50 seeds were placed on each plate. The plates were incubated upright at 22°C in 16h/d light (20 µmol m-2 s-1) for 6 days. Hypocotyl segments of 0.5 cm size were dissected from the six day old seedlings and cultured on shoot induction medium (MS/B5 Vitamins + 3% sucrose + 500 mg/L MES + BAP (13µm) + Zeatin (5µm) + Silver Nitrate (5mg/L) + 0.8% Agar (pH 5.8). The medium was poured into a 100 x 20 mm sterile PETRI™ dish, approximately 20 explants were placed on the medium per plate. Shoot meristems that appeared after 3-4 weeks were transferred to shoot elongation medium (MS/B5 Vitamins + 2% sucrose + 500mg/L MES + BAP (2µm) + GA-3 (0.1µm) + 0.8% Agar (pH 5.8) and poured in 250 mL culture vessels) and the cultures

were maintained in this medium for 4 weeks with one round of sub-culturing in between. Shoots of 2-3cm height were then transferred to root initiation media (1/2 strength MS/B5 Vitamins + 1% sucrose + 500mg/L MES + IBA (2.5μm) + 0.6% Agar (pH 5.8) and poured in 700 mL culture vessels) for root development. Rooted shoots were sub-cultured in fresh root initiation media at 3-4 weeks intervals as stem cuttings for two-three rounds before use. The cultures were maintained throughout at 22°C in 16h/d light (30 μmol m-2 s-1).

ISOLATION AND PURIFICATION OF MESOPHYLL PROTOPLASTS

[0263] In vitro grown DH12075 Brassica napus plants were used as the explant source for isolating mesophyll protoplasts. To isolate the protoplasts, the 3rd to 4th upper fully expanded leaves from 3 - 4 weeks old plantlets were cut with a sharp scalpel into small strips (0.5 to 1 mm) for protoplast isolation. Enzymatic digestion was carried out by treating 250-500 mg of leaf material with 25 mL of digestion buffer (1.2% (w/v) Cellulase "Onozuka™" R10 and 0.2% (w/v) Macerozyme® R10 dissolved in K4 media (Spangenberg et al., 1998)). The PETRI™ dish containing the leaf material and digestion buffer was sealed with Parafilm™ and incubated at room temperature for 12 to 15 h in darkness. After overnight incubation the digests were filtered through a BD® cell strainer (mesh size 70μm). Protoplast suspensions (5-6 mL) collected in a 14 mL round bottomed tube was over layered with 1 mL of W5 washing buffer (154 mM NaCl, 125 mM CaCl2, 5 mM KCl and 5 mM glucose; pH 5.8 Menzel et al. (1981)).

[0264] The protoplast suspensions were further centrifuged at 400 RPM for 10 min. After centrifugation, protoplasts that floated in the interphase were withdrawn and washed by centrifugation using 10 mL of W5 buffer at 400 RPM for 10 min. After the final wash, isolated protoplasts were resuspended at a density of 1X 106 protoplasts per mL of W5 buffer and incubated for 1 hour before transfections.

ASSESSMENT OF PROTOPLAST YIELD AND VIABILITY

[0265] Protoplasts yield was assessed using a haemocytometer following the method of Sambrook and Russell, (2006). The cell viability was tested using 400mg/L of Evans blue stain dissolved in 0.5 M of mannitol as described by Huang et al. (1996) with few minor modifications to the protocol.

PEG 4000 MEDIATED DNA DELIVERY

[0266] Before delivery to B. napus protoplasts, plasmid DNA of each donor and ZFN construct was prepared from cultures of E. coli using the Pure Yield Plasmid Maxiprep System® (Promega Corporation, Madison, WI) following the instructions of the suppliers. Aliquots of donor and ZFN plasmid DNA were prepared in three molar ratios: 1:1 (30 μg of each plasmid),

5:1 (donor plasmid to ZFN plasmid to a total of 30 μg of plasmid DNA) and 10:1 (donor plasmid to ZFN plasmid to a total of 30 μg of plasmid DNA). Additionally, donor-only and ZFN-only aliquots (30 μg) were prepared as controls. The amounts of DNA delivered to the *B. napus* protoplasts via the PEG4000 mediated transformation are summarized in Table 9.

Table 9: Quantities of ZFN and donor DNA delivered to protoplasts

	Molar Ratio of plasmid DNA	*Total quantity of DNA (μg) delivered to 1 million protoplasts*
Splicing	Donor plasmid only	30
	ZFN plasmid only (pDAB107827)	30
	1:1 Donor:ZFN	60
	5:1 Donor:ZFN	30
	10: Donor:ZFN	30
Editing	Donor plasmid only	30
	1:1: ZFN plasmids (pDAB107827 and pDAB107828)	30
	1:1:1 Donor:ZFN:ZFN	90
	5:1:1 Donor:ZFN:ZFN	30
	10:1:1 Donor:ZFN:ZFN	30

[0267] Each aliquot of plasmid DNA was applied to one million protoplasts (viability ≥95) suspended in 100μl of transformation buffer (15mM MgCl2, 0.1% (w/v) morpholinoethanesulphonic acid (MES) and 0.5 M Mannitol; pH 5.8) followed by 150 μl of PEG solution (40% (w/v) PEG 4000 in 0.4 M Mannitol and 0.1 M Ca (NO3)2 (pH 6-7) Spangenberg and Potrykus (1995). After 10-15 min of incubation at room temperature, 5 mL of W5 buffer was added in a drop wise manner and the protoplasts were gently mixed. Another 5 mL of W5 buffer was added as a slow stream to the protoplasts suspension. Protoplasts were mixed gently and centrifuged at 400 RPM for 10 min and the W5 supernatant was removed carefully leaving behind the protoplasts in the form of a pellet. Transfected protoplasts were then incubated in 1 mL of W5 buffer at room temperature until they were embedded in bead type

cultures. The transfected protoplasts were embedded following the sodium alginate method as described below.

CULTURING OF MESOPHYLL DERIVED PROTOPLASTS TO RECOVER VIABLE MICROCALLI

[0268] Before embedding within the medium, the transfected protoplasts were centrifuged at 400 RPM for 10 minutes and the W5 buffer was carefully removed. The protoplasts were then resuspended in 1.0 mL of 0.5 M Mannitol and incubated on ice. To the protoplast solution, an equal volume of 1.0 % sodium alginate was added and mixed gently. The protoplasts suspension was incubated in ice until it was embedded. Bead forming solution (0.4 M Mannitol + 50mM CaCl2 (pH 5.8)) was transferred to a sterile six well plate (3-4 mL per well) using a serological pipette. Exactly 1.0 mL of the protoplasts suspension was added in a drop wise manner using a 1 mL pipette into the bead forming solution and each transfected sample (ca. 5 x 105 protoplasts) was embedded per well. The protoplasts suspension was incubated for 1-2 hours at room temperature to form sodium alginate beads. After the incubation period the bead forming solution was carefully removed and replaced with 4-5 mL of 1: 2 mixture of K3+H:A media (Spangenberg et al 1998) supplemented with 1.5 mg/L of hygromycin. The protoplasts were cultured for 3-4 weeks in darkness at 22°C in a shaker (50 RPM). After 3-4 weeks the resistant microcalli (0.5-1.0mm) were released by treating with depolymerisation buffer (0.3 M Mannitol + 20mM Sodium Citrate (pH 5.8)). After removing the liquid media, 3-4 mL of depolymerisation buffer was added to each well containing the bead-type cultures and incubated at room temperature for 2 hours. Using a sterile forceps the beads were gently mixed to enhance the efficient release of the microcalli. Next a sterile 1.0 mL pipette was used to gently mix gelling agent that was released in the depolymerisation buffer and subsequently removed. The microcalli was washed twice using 5 mL of liquid A media and the microcalli was resuspended in sufficient quantity of liquid A (50 mL of liquid A was used for one mL of the settled cell volume (SCV: this was measured after transferring all the released microcalli to a sterile 50 or 15 mL falcon tube and allowed to settle down for 5 min)). After mixing the microcalli uniformly, 0.5 mL of the microcalli suspended in the liquid A media was transferred to B1 media (MS/MS Vitamins + 3.5 % Sucrose + 500mg/L MES + BAP (5μm) + NAA (5μm) + 2, 4-D (5μm) + 1.5 mg/L hygromycin + 0.7 % Agarose Type I (pH 6.0) and poured in 100 x 20 mm sterile PETRI™ dish) and using 1-2 mL of additional liquid A media the microcalli was distributed uniformly in the B1 media and the excess liquid A media was carefully removed from each plate. The plates were sealed using a micropore tape which

enhanced the embryo maturation. The cultures were maintained at 22°C in 16h/d light (30 µmol m-2 s-1).

PROLIFERATION AND REGENERATION OF SHOOTS FROM MESOPHYLL DERIVED PROTOPLASTS

[0269] Hygromycin resistant colonies were picked from B1 media (microcalli derived from both SA and SP methods) after 2-3 weeks of incubation and transferred to B2 media (MS/MS Vitamins + 3.0 % Sucrose + 500mg/L MES + 500mg/L PVP + 5mg/L Silver nitrate + 5mg/L 2i P + NAA (0.5µm) + GA-3 (0.3µm) + 1.5 mg/L Hygromycin + 0.7 % Agarose Type I (pH 5.8) and poured in 100 x 20 mm sterile PETRI™ dish). Approximately 25-30 calli were placed per plate and the plates were sealed using Parafilm™ and incubated at 22°C in 16h/d light (30 µmol m-2 s-1). Hygromycin resistant colonies were subsequently recovered after 5-6 rounds of sub-culturing in B2 media at two weeks interval. The number of calli per plate was reduced to 12-15 after a third round of sub-culturing. Shoot primordias that appear after 10-12 weeks were carefully recovered along with the residual calli and transferred to shoot elongation medium (MS/B5 Vitamins + 2% sucrose + 500mg/L MES + BAP (2µm) + GA-3 (0.1µm) + 300mg/L Timentin + 1.5 mg/L Hygromycin + 0.8% Agar (pH 5.8) and poured in 250 mL culture vessels). The shoots that survive after 2- 3 rounds of Hygromycin selection were transferred to rooting media (1/2 strength MS/B5 Vitamins + 1% sucrose + 500mg/L MES + IBA (2.5µm) + 1.5 mg/L Hygromycin + 0.6% Agar (pH 5.8) and poured in 700 mL culture vessels).

ISOLATION OF GENOMIC DNA FROM MESOPHYLL PROTOPLASTS

[0270] Transfected protoplasts were transferred from the 3 cm PETRI™ dish to a 2 mL microfuge tube. The cells were pelleted by centrifugation at 70 g and the supernatant was removed. To maximize the recovery of transfected protoplasts, the PETRI™ dish was rinsed three times with 1 mL of wash buffer. Each rinse was performed by swirling the wash buffer in the PETRI™ dish for 1 minute, followed by transfer of the liquid to the same 2 mL microfuge tube. At the end of each rinse, the cells were pelleted by centrifugation at 70 g and the supernatant was removed. The pelleted protoplasts were snap frozen in liquid nitrogen before freeze drying for 24 h in a Labconco Freezone 4.5® (Labconco, Kansas City, MO) at -40°C and 133 x 10-3 mBar pressure. The lyophilized cells were subjected to DNA extraction using the DNeasy® Plant DNA Extraction Mini kit (Qiagen) following the manufacturer's instructions, with the exception that tissue disruption was not required and the protoplast cells were added directly to the lysis buffer.

ISOLATION OF GENOMIC DNA FROM CALLUS TISSUE

[0271] Individual calli was snap frozen in liquid nitrogen before freeze drying for 24 h in a Labconco Freezone 4.5® (Labconco, Kansas City, MO) at -40°C and 133 x 10-3 mBar pressure. The lyophilized calli was subjected to DNA extraction using the DNeasy® Plant DNA Extraction Maxi kit (Qiagen, Hilden, Germany) following the manufacturer's instructions.

ISOLATION OF GENOMIC DNA FROM LEAF TISSUE

[0272] Thirty (30) mg of young leaf tissue from regenerated plants was snap frozen in liquid nitrogen before freeze drying for 24 h in a Labconco Freezone 4.5® (Labconco, Kansas City, MO) at -40°C and 133 x 10-3 mBar pressure. The lyophilized calli was subjected to DNA extraction using the DNeasy® Plant DNA Extraction Maxi kit (Qiagen, Hilden, Germany) following the manufacturer's instructions.

PCR ASSAYS OF GENOMIC DNA FOR NHEJ-MEDIATED SPLICING AND EDITING OF FAD3C

[0273] Detection of integration of donor DNA to the Fad3C gene of B. napus was done by a series of PCR where at least one primer was specific to the Fad3C locus (Table 10) and a second primer specific to either the promoter or terminator of the gfp cassette (Table 10 and Figure 19A). Specificity was obtained by designing oligonucleotides where the last base pair aligned to a SNP that differentiated the Fad3C genomic sequence from the other copies of Fad3 genes and included a phosphorothioate internucleotide linkage before this base pair as indicated by an asterisk [*]. This design, used in combination with a polymerase having proofreading activity, directed specific amplification of each Fad3C or Fad3A allele and excluded other Fad3 copies as noted. Each primer set was empirically tested for amplification of the correct gene copies through Sanger-based sequencing of the PCR amplification products obtained from wild type *B. napus*.

Table 10: Oligonucleotide sequences used to detect integration of DNA into ZFN-induced double-stranded breaks

	Primer Name	*Primer Sequence*	*SEQ ID NO:*	*Specificity*
1	FAD3CNHEJ-L4-F2	gattcctaagcattgttgggt*c	279	Fad3C only
2	FAD3CNHEJ-L4-R2	gaaaatctcatatcgaacgtgcg*t	280	Fad3C only

3	FAD3CNHEJ-L6-F1	cgcttaccctctctatctggta*a	281	Does not amplify Fad3C' or Fad3C"
4	FAD3CNHEJ-L6-R2	ccttgcctctgtaccaaggca*g	282	Fad3C only
5	19SPNHEJ-R2	gtgtgtgggaatcttatcttcgg	283	n/a
6	AtORF1NHEJ-F1	caagtcaggtattatagtccaagca	284	n/a
7	AtUbiNHEJ-R1	caagaatatcctgatccgttgac	285	n/a
8	AtORF23tNHEJ-F1	tggcagttgaaatactcaaacc	286	n/a
9	FAD3aCNHEJ-L4-F1	gtcctttgagatccatgagcta*t	287	Fad3A only
10	FAD3aCNHEJ-L4-F2	gattcctaagcattgttgggt*a	288	Fad3A only
11	FAD3aNHEJ-L4-R1	tgcgttcaagaaatcaaagac*a	289	Fad3A only
12	FAD3aNHEJ-L4-R2	gaaaatctcatatcgaacgtgcg*g	290	Fad3A only
13	FAD3aNHEJ-L6-F1	tctggtaaatcctaattcct*c	291	Fad3A only
14	FAD3aNHEJ-L6-R2	ccttgcctctgtaccaaggca*a	292	Fad3A only
15	FAD3aNHEJ-L6-R1	cttgcctctgtaccaaggcaactt*c	293	Excludes Fad3C

* Indicates phosphorothioate internucleotide linkages to direct specific amplification (with proofreading polymerase) of Fad3C or Fad3A to exclusion of other copies of Fad3 as noted. Each primer set was empirically tested for amplification of the correct gene copies by Sanger-based sequencing of the PCR amplification products obtained from wild type *B. napus*.

DETECTION OF GENE ADDITION TO FAD3C BY NON-HOMOLOGOUS END JOINING IN PROTOPLASTS

[0274] Genomic DNA was extracted from protoplast pools (one million protoplast per pool) to which donor DNA encoding a functional tGFP reporter cassette (pDAS000341 or pDAS000343), ZFN DNA (pDAB107827 or pDAB107828) or a mixture of donor and ZFN DNA had been delivered twenty-four hours earlier. Quantities of DNA delivered for transformation are described above. PCR products were cloned into plasmid vectors. The genomic editing occurs independently in each cell giving rise to a variety of different insertion events, by cloning into a plasmid vector, each genomic edit can be sequenced without ambiguity. Several clones were sequenced on an ABI3730XL® automated capillary electrophoresis platform. Analysis of gene sequences was done using Sequencher SOFTWARE v5.0™ (GeneCodes, Ann Arbor, MI).

[0275] Evidence of gene addition to the Fad3C locus by editing or splicing was provided by amplification of both the 5' and 3' Fad3C-cassette junctions from genomic DNA extracted from protoplasts using the primers described in Table 10. Products of PCR amplification with primers "FAD3CNHEJ-L4-F2" and "AtUbiNHEJ-R1" was completed to amplify the 5' junction of tGFP cassette and Fad3C. PCR amplification with primers "FAD3CNHEJ-L4-R2" and "AtORF23tNHEJ-F1" was completed to amplify the 3' junction of tGFP cassette and Fad3C. PCR amplification with primers "FAD3CNHEJ-L4-F2" and "FAD3CNHEJ-L4-R2" was completed to amplify across the double strand breaks induced by ZFN 28051-2A-28052. No amplification was observed from protoplasts to which ZFN plasmid or donor plasmid alone had been delivered. All junction sequences were indicative of insertion of the tGFP cassette at the Fad3C locus via an NHEJ-mediated repair pathway. Deletions of varying lengths from either or both the genome and the cassette were observed as well as the addition of sequences derived from the vector backbones (either from the donor or ZFN) being inserted between the genome and the cassette (Figure 20A and Figure 20B).

DETECTION OF GENE ADDITION TO FAD3C BY NON-HOMOLOGOUS END JOINING IN CALLUS TISSUE REGENERATED FROM PROTOPLASTS

[0276] Further evidence of splicing and editing of the Fad3C locus was obtained from callus tissue regenerated from protoplasts on selection (1.5 mg/L hygromycin, as described above) to which donor DNA encoding an hph cassette (pDAS000340 or pDAS000342), ZFN DNA only (pDAB107827 or pDAB107828) or donor and ZFN DNA had been delivered

(quantities of DNA delivered are given in Table 9). DNA was extracted from approximately 80 calli for each ratio, except editing 1:1:1, for which no calli survived, four weeks after protoplast transfection.

[0277] Integration of the hph cassette into the B. napus genome (fwat Fad3C or randomly) was confirmed by Taqman™ qPCR using primers (SEQ ID NO:294; F - 5' CTTACATGCTTAGGATCGGACTTG 3', SEQ ID NO:295; R - 5' AGTTCCAGCACCAGATCTAACG 3') and probe (SEQ ID NO:296; 5' CCCTGAGCCCAAGCAGCATCATCG 3') specific to the hph gene. These primer-probe pairs were used in a duplex reaction with primers (SEQ ID NO:297; F - 5' CGGAGAGGGCGTGGAAGG 3', SEQ ID NO:298; R - 5' TTCGATTTGCTACAGCGTCAAC 3') and probe (SEQ ID NO:299; 5' AGGCACCATCGCAGGCTTCGCT 3') specific to the B. napus high mobility group protein I/I (HMG I/Y), which is present as a single copy on the A genome (Weng et al., 2004, Plant Molecular Biology Reporter). Amplification was performed on a C1000 thermal cycler with the CFX96 or CF384 real-time PCR detection system™ (BioRad, Hercules, CA). Results were analyzed using the CFX Manager™ (BioRad) software package. Relative quantification was calculated according to the 2-ΔΔCt method (Livak and Schmittgen, 2001), which provided an estimation of the number of copies of hph cassette inserted into the genome.

[0278] Evidence of NHEJ-mediated splicing and editing of Fad3C was obtained by conducting PCR assays with one primer specific to Fad3C and a second primer specific to either the promoter or terminator of the hph cassette (Table 9 and Figure 19B). Due to limited quantities of DNA obtained from callus tissue, only integration in the sense orientation was assayed. PCR products were gel-purified using QiaQuick MiniElute PCR Purification kit™ (Qiagen) and sequenced using a direct Sanger sequencing method. The sequencing products were purified with ethanol, sodium acetate and EDTA following the BigDye® v3.1 protocol (Applied Biosystems) and sequenced and analysed as above.

[0279] The numbers of calli containing the donor cassette in each experiment are given in Table 11. Evidence of donor gene addition to the Fad3C locus by editing and/or splicing was provided by PCR amplification (with primers shown in Table 10) across the ZFN cut sites and both the 5' and 3' Fad3C-hph cassette junctions. PCR amplification of the genomic DNA isolated from callus tissue recovered from control protoplasts which were transformed with only the hph plasmid (pDAS000340 and pDAS000342) or only the ZFN plasmid (pDAB107827 and pDAB107828) did not result in the production of PCR amplification products.

[0280] The PCR amplicons produced from the amplification of the 5' and 3' Fad3C-hph cassette junctions were purified from the agarose gel and sequenced to confirm specificity of the integration within the Fad3C genomic locus. The results of the sequencing analysis of the PCR products indicated that each isolated callus which was generated from an individually transformed protoplast only produced a single PCR amplification product and did not contain cells of mixed genotypes.

[0281] In NHEJ-mediated integration of donor sequences within the Fad3C genomic locus experiments the frequency of addition to the target locus (as defined by any part of the donor DNA vector being amplified from the target locus) was 42%, 46% and 32% for the DNA concentrations of 1:1, 5:1, and 10:1 (Donor DNA: ZFN DNA), respectively. See, Table 12. The frequency of on-target splicing was determined by assaying whether both cassette junctions were amplifiable and from the sequencing of the PCR products. These results verified that the cassette was inserted at the target locus in the correct orientation. The frequency of integration was calculated as 4%, 3% and 3% for the 1:1, 5:1 and 10:1 of Donor plasmid DNA: ZFN plasmid DNA concentrations, respectively. In gene editing experiments the frequency of addition to the target locus defined by any part of the donor DNA vector being amplified from the target locus, was 66% and 65% for the 5:1:1 and 10:1:1 of Donor plasmid DNA: ZFN plasmid DNA concentrations, respectively. See, Table 13. The frequency of on-target editing, was determined by both cassette junctions being amplifiable and producing a sequence of PCR products. These results verified that the cassette was inserted at the target locus in the correct orientation at frequencies of 3% and 6% for the 5:1:1 and 10:1:1 of Donor plasmid DNA: ZFN plasmid DNA concentrations, respectively. As observed in the protoplast assays, the base pairs were either deleted or additional bases were inserted between the genome and the cassette as a result of the cleavage of the genomic locus by the ZFN (Figures 21-22).

[0282] In certain instances the PCR products resulted in an addition of nucleotide sequences within the target locus, no PCR product, or a larger PCR product than observed in wild-type samples. These results which were produced from the PCR amplification using primers flanking the cut site indicated that the locus had been disrupted in both pairs of chromosomes (Figures 21-22). In some of the instances more than one band was amplified at the splice junctions (Figures 21-22) indicating that different insertions had occurred independently in each copy of the genome.

Table 11: Number of calli positive for presence of *hph* after four weeks on selection

Vectors delivered	Molar Ratio of Donor DNA: ZFN DNA	Number of calli sampled	Number of calli positive for *hph* after four weeks on selection
pDAS000340 DAB107827	1:1	88	76
	5:1	88	35
	10:1	87	37
pDAS000342 DAB107827 DAB107828	1:1:1	-	-
	5:1:1	80	38
	10:1:1	79	52

Table 12: Number of calli with *hph* inserted by splicing at FadC locus at the DSB induced by ZFN28051-2A-28052

Vectors delivered	Molar Ratio of Donor DNA: ZFN DNA	Number of calli positive for *hph* after four weeks on selection	Number of calli from which at least one splicing border amplified	Number calli from which at least one perfect* border amplified	Number of calli from which both splicing borders amplified
pDAS000340 + DAB107827	1:1	76	32	0	3
	5:1	35	16	0	1
	10:1	37	12	0	1

* number base pairs deleted or additional base pairs inserted at cut site

Table 13: Number of calli with *hph* inserted by editing at FadC locus at the cut sites induced by by ZFN28051-2A-28052 and ZFN28053-2A-28054

Vectors delivered	Molar Ratio of Donor DNA: ZFN DNA	Number of calli positive for *hph* after four weeks on selection	Number of calli from which at least one splicing border amplified	Number calli from which at least one perfect* border amplified	Number of calli from which both editing borders amplified
pDAS000342 + DAB107827 + DAB107828	5:1:1	38	25	2	1
	10:1:1	52	34	2	3

* number base pairs deleted or additional base pairs inserted at cut site

DETECTION OF GENE ADDITION TO FAD3C BY NON-HOMOLOGOUS END JOINING IN PLANTS

[0283] DNA was extracted from plants that were regenerated from protoplasts and transferred to potting medium (as described above). The majority of plants recovered were estimated to contain only 1-2 copies of the hph cassette encoded in the donor DNA. Plants were analyzed with the same suite of assays described for callus tissue as well as with assays to determine if the cassette had inserted in an antisense orientation or a donor integration at the Fad3A locus.

Table 14: Estimated copy number of plants regenerated from protoplasts. For each ratio three transfections of one million protoplasts were performed

Vectors delivered	Molar Ratio of Donor DNA: ZFN DNA	# plants with 1-2 copies *hph*	# plants with 3-4 copies *hph*	# plants with 5 or more copies *hph*
pDAS000340 DAB107827	1:1	37	16	34
	5:1	18	14	30
	10:1	16	13	18
pDAS000342 DAB107827 DAB107828	1:1:1	0	1	1
	5:1:1	22	14	18
	10:1:1	23	11	27
Total	---	116	69	128

[0284] The frequency of on-target splicing for the linear donor design constructs, where the hph cassette was inserted into Fad3C in either direction, was 51%, 32% and 56% for Donor DNA: ZFN DNA at concentrations of 1:1, 5:1 and 10:1, respectively (Table 15). Of these results, 35% 32% and 50% (1:1, 5:1 and 10:1) were inserted in the forward orientation (Table 15).

[0285] The frequency of on-target editing, where the hph cassette was inserted into Fad3C in either direction, replacing the area from locus 4 to locus 6, was 2% and 0% for Donor DNA: ZFN DNA: ZFN DNA at concentrations of 5:1:1 and 10:1:1, respectively (Table 16). In addition, when both ZFNs were delivered at 5:1:1, 2% and spliced into locus 4 and 10% spliced into locus 6 and when both ZFNs were delivered at 10:1:1 10% and spliced into locus 4 and 15% spliced into locus 6. The PCR amplicons were obtained and sequenced to determine the insert junction sequences. The resulting sequences for specifically labeled plants are described in Table17.

Table 15: Number of plants with *hph* inserted by splicing at FadC locus at the DSB induced by ZFN28051-2A-28052

Ratio	Forward Orientation			Reverse Orientation			Both Orientations (Forward & Reverse)		Total		Events Tested Positive for HPH
	One 5' In-	One 3' In-	Both 5' &	One 5' In-	One 3' In-	Both 5' &	One 5' or 3'	Both 5' & 3'	On-	Off-target	
1:1	3	2	2	5	3	4	17	4	40	47	87
5:1	8	2	1	-	-	-	3	-	14	48	62
10:1	8	2	1	2	-	2	9	2	26	21	47

Table 16: Number of plants with *hph* inserted by editing at FadC locus at the cut sites induced by ZFN28051-2A-28052 and ZFN28053-2A-28054

RATIO	Forward Orientation			Reverse Orientation			Both Orientations (Forward & Reverse)		Total		Events Tested Positive for HPH
	One border 5' In-out	One border 3' In-out	Both borders 5' & 3'	One border 5' In-out	One border 3' In-out	Both borders 5' & 3'	One border either 5' or 3'	Both borders 5' & 3'	On-target	Off-target	
1:1	3	2	2	5	3	4	17	4	40	47	87
5:1	8	2	1	-	-	-	3	-	14	48	62
10:1	8	2	1	2	-	2	9	2	26	21	47

Table 17: Plant details of single copy *hph*, target inserted at Fad3C locus at the cut sites induced by ZFN28051-2A-28052 and ZFN28053-2A-28054

Plant barcode	PCR/ Sequence information	Sequence ID Number
349711	Locus 4 upstream	353
349685	Locus 4 upstream	354
346258	Locus 4 upstream	355
348918	Locus 4 upstream	356
359900	Locus 4 upstream	357
346125	Locus 4 upstream	358
348919	Locus 4 upstream	359
349215c	Locus 4 upstream	360
349216c	Locus 4 upstream	361
346102	Locus 4 downstream	362

346175	Locus 4 downstream	363
345888	Locus 6 downstream	364
356731	Locus 6 downstream	365
346128	Locus 4 downstream antisense orientation	366
347359	Locus 6 upstream antisense orientation	367

[0286] The frequency of on-target splicing, where the *hph* cassette was inserted into Fad3C in either direction for the circular donor, was 51%, 32% and 56% for 1:1, 5:1 and 10:1 respectively (Table 18; Figure 23). Of these 35% 32% and 50% (1:1, 5:1 and 10:1) were inserted the forward orientation (Table 18).

[0287] The frequency of on-target editing, where the *hph* cassette was inserted into Fad3C in either direction, replacing the area from locus 4 to locus 6, was 2% and 0% for 5:1: and 10:1:1 respectively (Table 19; Figure 24). In addition, when both ZFNs were delivered at 5:1:1, 2% and spliced into locus 4 and 10% spliced into locus 6 and when both ZFNs were delivered at 10:1:1 10% and spliced into locus 4 and 15% spliced into locus 6.

Table 18: Number of plants with *hph* inserted by splicing at FadC locus at the DSB induced by ZFN28051-2A-28052

Vectors delivered	Molar Ratio of Donor DNA: ZFN DNA	Number of plants analysed (positive for *hph)*	Number of plants from which at least one splicing border amplified (forward/reverse/either)	Number of plants from which both splicing borders amplified (forward/reverse/either)
pDAS000340 + DAB107827	1:1	60	21/23/31	4/7/8
	5:1	37	12/4/12	3/1/3
	10:1	46	23/12/26	4/4/7

* no base pairs deleted or additional base pairs inserted at cut site

Table 19: Number of plants with *hph* inserted by editing at FadC locus at the cut sites induced by ZFN28051-2A-28052 and ZFN28053-2A-28054

Vectors delivered	Molar Ratio of Donor	Number of plants analysed (positive for	Number of plants from which at least	Number of plants from which both

	DNA: ZFN DNA	*hph)*	one splicing border amplified (forward/reverse/either)	editing borders amplified (forward/reverse/either)
pDAS000342 + DAB107827 + DAB107828	5:1:1	39	17/11/24	0/1/1
	10:1:1	63	27/27/34	0/0/0

* no base pairs deleted or additional base pairs inserted at cut site

TARGETED INTEGRATION OF BRASSICA NAPUS OMEGA-3 FATTY ACID DESATURASE VIA HDR

[0288] The donor vectors containing the tGFP and HPH cassettes are modified to include 1kb of FAD3 upstream and downstream donor sequences. The FAD3 upstream and downstream donor sequences are 100% identical to the native FAD3 sequence and are obtained from the FAD3 zinc finger binding site; GCCCAAGGAACCCTTTTCTGGGCCATCTTCGTACTCGGCCACGACTGGTAATTTAAT (SEQ ID NO:255) or AGCGAGAGAAAGCTTATTGCAACTTCAACTACTTGCTGGTCGATCGTGTTGGCCACTC (SEQ ID NO:256). The resulting four "donor" vectors are similar to pDAS000340 (hygromycin-resistant gene-splicing donor), pDAS000341 (tGFP reporter gene splicing donor), pDAS00342 (hygromycin-resistant gene-editing donor) and pDAS000343 (tGFP reporter gene editing donor), wherein the only modification is the inclusion of 1Kb of FAD3 genomic upstream and downstream sequences. The zinc finger nuclease plasmids (pDAB107827 and pDAB107828) previously described for NHEJ mediated integration are used for the HDR mediated integration.

TRANSFORMATION OF *BRASSICA NAPUS*

[0289] Mesophyll derived protoplasts are isolated and prepared from *Brassica napus* (DH10275) plants as described above. The protoplasts are transformed with purified plasmid DNA. Aliquots of donor and ZFN plasmid DNA are prepared in three molar ratios: 1:1 (30 μg

of each plasmid), 5:1 (donor plasmid to ZFN plasmid to a total of 30 μg of plasmid DNA) and 10:1 (donor plasmid to ZFN plasmid to a total of 30 μg of plasmid DNA). Additionally, donor-only and ZFN-only aliquots (30 μg) are prepared as controls. The amounts of DNA delivered to the *B. napus* protoplasts via a PEG4000 mediated transformation are summarized in Table 20. The transformed protoplast cells are cultured as previously described, wherein the selection medium is glufosinate selection medium, and putative transformants are assayed via qPCR analysis for transgene insertions.

Table 20: Quantities of ZFN and donor DNA delivered to protoplasts

	Molar Ratio of plasmid DNA	***Total quantity of DNA (μg) delivered to 1 million protoplasts***
Splicing	Donor plasmid only	30
	ZFN plasmid only	30
	1:1 Donor:ZFN	60
	5:1 Donor:ZFN	30
	10: Donor:ZFN	30
Editing	Donor plasmid only	30
	1:1: ZFN plasmids	30
	1:1:1 Donor:ZFN:ZFN	90
	5:1:1 Donor:ZFN:ZFN	30
	10:1:1 Donor:ZFN:ZFN	30

DETECTION OF GENE ADDITION TO *FAD3* BY HDR IN PROTOPLASTS

[0290] Genomic DNA is extracted from protoplast pools (one million protoplast per pool) to which donor DNA encoding a functional reporter cassette or selectable marker cassette, ZFN DNA or a mixture of donor and ZFN DNA are delivered twenty-four hours earlier. Quantities of DNA delivered for transformation are described above. PCR products are cloned into plasmid vectors. The genomic editing occurs independently in each cell giving rise to a variety of different insertion events, by cloning into a plasmid vector, each genomic edit can be sequenced without ambiguity. Several clones are sequenced on an ABI3730XL® automated

capillary electrophoresis platform. Analysis of gene sequences is done using SEQUENCHER SOFTWARE V5.0™ (GeneCodes, Ann Arbor, MI).

[0291] Evidence of gene addition to the FAD3 locus by editing or splicing is provided by amplification of both the 5' and 3' FAD3-cassette junctions from genomic DNA extracted from protoplasts. No amplification is observed from protoplasts to which ZFN plasmid or donor plasmid alone had been delivered. All junction sequences are indicative of insertion of the cassette at the FAD3 locus via an HDR-mediated repair pathway. Deletions of varying lengths from either or both the genome and the cassette are observed as well as the addition of sequences derived from the vector backbones (either from the donor or ZFN) being inserted between the genome and the cassette.

DETECTION OF GENE ADDITION TO *FAD3* BY HDR IN CALLUS TISSUE REGENERATED FROM PROTOPLASTS

[0292] Further evidence of splicing and editing of the FAD3 locus was obtained from callus tissue regenerated from protoplasts on selection to which donor DNA encoding a cassette, ZFN DNA only, or donor and ZFN DNA are delivered. DNA is extracted from approximately 80 calli for each ratio.

[0293] Integration of the cassette into the *B. napus* genome is confirmed by TAQMAN™ qPCR using primer and probes specific to the donor insert and the genomic flanking sequences. Relative quantification is calculated according to the $2^{-\Delta\Delta Ct}$ method (Livak and Schmittgen, 2001), which provided an estimation of the number of copies of cassette inserted into the genome. Evidence of NHEJ-mediated splicing and editing of FAD3 is obtained by conducting PCR assays with one primer specific to FAD3 and a second primer specific to either the promoter or terminator of the cassette. PCR products are gel-purified using QIAQUICK MINIELUTE PCR PURIFICATION KIT™ (Qiagen) and sequenced using a direct Sanger sequencing method. The sequencing products are purified with ethanol, sodium acetate and EDTA following the BIGDYE® v3.1 protocol (Applied Biosystems) and sequenced and analyzed as above.

[0294] The numbers of calli containing the donor cassette in each experiment are determined. Evidence of donor gene addition to the FAD3 locus by editing and/or splicing is provided by PCR amplification across the ZFN cut sites and both the 5' and 3' FAD3-cassette junctions. PCR amplification of the genomic DNA isolated from callus tissue recovered from control protoplasts which are transformed with only the plasmid or only the ZFN plasmid do not result in the production of PCR amplification products.

[0295] The PCR amplicons produced from the amplification of the 5' and 3' FAD3-cassette junctions are purified from the agarose gel and sequenced to confirm specificity of the integration within the FAD3 genomic locus. The results of the sequencing analysis of the PCR products indicate that each isolated callus which is generated from an individually transformed protoplast only produce a single PCR amplification product and do not contain cells of mixed genotypes.

DETECTION OF GENE ADDITION TO *FAD3* BY HDR IN PLANTS

[0296] DNA is extracted from plants that are regenerated from protoplasts and transferred to potting medium. The majority of plants recovered are estimated to contain only 1-2 copies of the cassette encoded in the donor DNA. Plants are analyzed with the same suite of assays described for callus tissue as well as with assays to determine if the cassette had inserted in the FAD3 locus.

[0297] The frequency of on-target splicing, where the cassette is inserted into FAD3 locus is determined using the PCR assays described above. The amplicon bands obtained are sequenced to determine the flanking sequences. Additionally, plants are screened for off-target insertions to determine the frequency of integration of the cassette at sites other than FAD3.

EXAMPLE 9: TARGETED INTEGRATION OF *Brassica napus* OMEGA-3 FATTY ACID DESATURASE (FAD3) WITH AN AGRONOMICALLY IMPORTANT GENE

[0298] Constructs containing the DGT-28 transgene (International Patent Application No. WO/2013/116700) that confers resistance to the herbicide glyphosate are designed and built for integration within the FAD3 genomic loci of *Brassica napus*. The constructs and associated zinc finger nuclease constructs (*e.g.*, (pDAB107827 and pDAB107828)) are transformed into *Brassica napus* cells as previously described above. Transformants are identified and confirmed via molecular confirmation assays as previously described. The FAD3 chromosomal integrants, comprising an integrated *dgt-28* transgene are isolated. The integration of the *dgt-28* transgene within the FAD3 locus is exemplified via NHEJ mediated integration and HDR mediated integration. The integration within the FAD3 locus can be directed into the FAD3 endogenous sequence or into the previouslt described ETIP (pDAS000271 – pDAS000275) that is stably integrated within the FAD3 locus. The integration within the FAD3 locus via an NHEJ mediated mechanism can be made using linearized donor or circular donor DNA designs. Transformed DGT-28 *Brassica napus* events are obtained and

tested for robust expression of the DGT-28 and the subsequent resistance to the herbicide glyphosate.

[0299] While certain exemplary embodiments have been described herein, those of ordinary skill in the art will recognize and appreciate that many additions, deletions, and modifications to the exemplary embodiments may be made without departing from the scope of the following claims. In addition, features from one embodiment may be combined with features of another embodiment.

SEQUENCE LISTING IN ELECTRONIC FORM

In accordance with Section 111(1) of the Patent Rules, this description contains a sequence listing in electronic form in ASCII text format (file: 54964-3 Seq 02-MAR-15 v1.txt).

A copy of the sequence listing in electronic form is available from the Canadian Intellectual Property Office.

CLAIMS:

1. A method for modifying the genome of a cell, the method comprising:

cleaving, in a site-specific manner, a target site in a FAD3 gene in a cell, to thereby generate a break in the FAD3 gene, the FAD3 gene comprising any of SEQ ID NOs:20-23, SEQ ID NOs: 25-38, SEQ ID NOs: 40-45, SEQ ID NO:47 and SEQ ID NO: 49;

wherein the FAD3 gene is modified at or near any of SEQ ID NOs: 20-23, SEQ ID NOs: 25-28, SEQ ID NOs: 40-45, SEQ ID NO:47 and SEQ ID NO: 49 following cleavage, and wherein the method further comprises integrating into the break a nucleic acid sequence of interest,

wherein the cleaving in a site-specific manner comprises introducing one or more nucleases comprising a DNA-binding domain and a cleavage domain or cleavage half-domain into the cell or a polynucleotide encoding the nuclease, wherein the nuclease binds with specificity to the target site and cleaves at or near the target site to thereby generate the break.

2. The method according to claim 1, wherein the FAD3 gene is a FAD3A, FAD3A', FAD3A", FAD3C, FAD3C" and/or a FAD3C' gene.

3. The method according to claim 1 or claim 2, wherein the DNA-binding domain is selected from the group consisting of a meganuclease DNA-binding domain, a leucine zipper DNA-binding domain, a transcription activator-like (TAL) DNA-binding domain, a RNA-guided CRISPR-Cas9, a recombinase, a zinc finger protein DNA-binding domain, and chimeric combinations of any one of the foregoing.

4. The method according to claim 3, wherein the cleavage domain or cleavage half-domain is selected from the group consisting of a cleavage half-domain from a type IIS restriction endonuclease, a cleavage half-domain from FokI endonuclease, a cleavage half-domain from StsI endonuclease, and a homing endonuclease.

5. The method according to claim 3, wherein the nuclease is a zinc finger nuclease.

6. The method according to claim 5, wherein the zinc finger nuclease comprises from three to six zinc finger domains, each zinc finger domain comprising a recognition helix region, wherein the

zinc finger nuclease comprises the recognition helix regions ordered and shown in a single row of Table 3.

7. The method according to claim 1 or claim 2, wherein the cleaving in a site specific manner is specific for some but not all copies of FAD3A, FAD3A', FAD3A", FAD3C, FAD3C" and/or FAD3C'.

8. The method according to claim 1 or claim 2, wherein the target site is selected from the group consisting of SEQ ID NOs: 20-23, SEQ ID NOs: 25-38, SEQ ID NOs: 40-45, SEQ ID NO: 47 and SEQ ID NO: 49.

9. The method according to claim 1 or claim 2, wherein the cell is a plant cell.

10. The method according to claim 9, wherein the plant cell is a monocot plant cell or a dicot plant cell.

11. The method according to claim 10, wherein the plant cell is selected from the group consisting of *Brassica sp.*; *Brassica napus*; *Brassica rapa*; *Brassica juencea*; *Brassica oleracea*; *Brassica nigra*; *Zea sp.*; *Zea mays*; *Glycine sp.*; *Glycine max*; *Triticum sp*; *Triticum aestivum*; *Oryza sp*; *Oryza sativa*; *Triticae sp.*; *Triticae triticum*; *Heliantheae sp.*; *Heliantheae helianthus*; *Gossypium sp.*; *Gossypium hirsutum*; and *Hordeum vulgar*.

12. The method according to claim 1 or claim 2, wherein the nucleic acid sequence of interest is selected from the group consisting of a sequence comprising a DNA-binding domain binding target site, one or more insecticidal resistance genes, one or more herbicide tolerance genes, one or more nitrogen use efficiency genes, one or more water use efficiency genes, one or more nutritional quality genes, one or more DNA binding genes, one or more selectable marker genes and combinations thereof.

13. A cell modified according to the method of any one of claims 1 to 12.

14. The cell according to claim 13, wherein the cell is a transgenic cell comprising a nucleotide sequence of interest integrated into one or more copies of a FAD3A, FAD3A', FAD3A", FAD3C, FAD3C" and/or a FAD3C' gene.

15. The cell of claim 14, wherein the nucleotide sequence is heterologous or homologous to the cell.

16. The cell of claim 15, wherein the homologous sequence comprises at least one single nucleotide polymorphism.

17. A site specific zinc finger nuclease that cleaves at or near a nucleic acid target site selected from the group consisting of SEQ ID NOs: 20-23, SEQ ID NOs: 25-38, SEQ ID NOs: 40-45, SEQ ID NO: 47 and SEQ ID NO: 49, wherein the zinc finger nuclease comprises from three to six zinc finger domains, each zinc finger domain comprising a recognition helix region, wherein the zinc finger protein comprises the recognition helix regions ordered and shown in a single row of the following Table:

ZFP	sF1	F2	F3	F4	F5	F6
27961	RSDNLAR (SEQ ID NO:116)	QKKDRSY (SEQ ID NO:117)	RSDNLAR (SEQ ID NO:116)	QRGNRNT (SEQ ID NO:119)	RSDHLSR (SEQ ID NO:120)	RNQDRTN (SEQ ID NO:121)
27962	DRSNLSR (SEQ ID NO:122)	RQDSRSQ (SEQ ID NO:123)	QSSDLSR (SEQ ID NO:124)	DRSALAR (SEQ ID NO:125)	TSGSLTR (SEQ ID NO:126)	N/A
27973	QSSDLSR (SEQ ID NO:124)	AASNRSK (SEQ ID NO:128)	TSGSLSR (SEQ ID NO:129)	RSDALAR (SEQ ID NO:130)	RSDVLST (SEQ ID NO:131)	WGRLRKL (SEQ ID NO:132)
27974	ERGTLAR (SEQ ID NO:133)	RSDDLTR (SEQ ID NO:134)	RSDHLSA (SEQ ID NO:135)	QHGALQT (SEQ ID NO:136)	TSGNLTR (SEQ ID NO:137)	QSGHLSR (SEQ ID NO:138)
27987	TSGSLTR (SEQ ID NO:126)	RSDHLSQ (SEQ ID NO:140)	CTRNRWR (SEQ ID NO:141)	RSDNLSE (SEQ ID NO:142)	ASKTRKN (SEQ ID NO:143)	N/A
27990	TSGSLSR (SEQ ID NO:129)	TSSNRAV (SEQ ID NO:145)	TSGNLTR (SEQ ID NO:137)	DRSALAR (SEQ ID NO:125)	RSDVLSE (SEQ ID NO:148)	RNFSLTM (SEQ ID NO:149)

27991	QSGDLTR (SEQ ID NO:150)	TSGSLSR (SEQ ID NO:129)	QSGNLAR (SEQ ID NO:152)	TSGSLSR (SEQ ID NO:129)	QSGSLTR (SEQ ID NO:154)	N/A
27992	DRSHLAR (SEQ ID NO:155)	TSGSLSR (SEQ ID NO:129)	TSSNRAV (SEQ ID NO:145)	TSGNLTR (SEQ ID NO:137)	DRSALAR (SEQ ID NO:125)	N/A
28004	QSGNLAR (SEQ ID NO:152)	HLGNLKT (SEQ ID NO:161)	RSDHLSQ (SEQ ID NO:140)	TARLLKL (SEQ ID NO:163)	QSGNLAR (SEQ ID NO:152)	QTSHLPQ (SEQ ID NO:165)
28005	RSDNLSV (SEQ ID NO:166)	TSGHLSR (SEQ ID NO:167)	TSGSLTR (SEQ ID NO:126)	RSDALST (SEQ ID NO:169)	DRSTRTK (SEQ ID NO:170)	N/A
28021	QNAHRKT (SEQ ID NO:171)	TSGNLTR (SEQ ID NO:137)	LKQMLAV (SEQ ID NO:173)	RSDNLSR (SEQ ID NO:174)	DNSNRKT (SEQ ID NO:175)	N/A
28022	RSDNLSV (SEQ ID NO:166)	QNANRIT (SEQ ID NO:177)	TSGSLSR (SEQ ID NO:129)	QSSVRNS (SEQ ID NO:179)	DRSALAR (SEQ ID NO:125)	N/A
28023	RSDNLSR (SEQ ID NO:174)	DNSNRKT (SEQ ID NO:175)	DRSNLTR (SEQ ID NO:183)	RSDVLSE (SEQ ID NO:148)	TRNGLKY (SEQ ID NO:185)	N/A
28024	RSDALAR (SEQ ID NO:130)	RSDVLSE (SEQ ID NO:148)	RSSDRTK (SEQ ID NO:188)	RSDNLSV (SEQ ID NO:166)	QNANRIT (SEQ ID NO:177)	N/A
28025	QSSDLSR (SEQ ID NO:124)	QSTHRNA (SEQ ID NO:192)	RSDNLAR (SEQ ID NO:116)	QRGNRNT (SEQ ID NO:119)	RSDHLSR (SEQ ID NO:120)	RNQDRTN (SEQ ID NO:121)
28026	DRSNLSR (SEQ ID NO:122)	RQDSRSQ (SEQ ID NO:123)	QSSDLSR (SEQ ID NO:124)	DRSALAR (SEQ ID NO:125)	TSGSLTR (SEQ ID NO:126)	N/A
28035	QSSDLSR (SEQ ID NO:124)	AASNRSK (SEQ ID NO:128)	TSGSLSR (SEQ ID NO:129)	RSDALAR (SEQ ID NO:130)	RSDTLSQ (SEQ ID NO:206)	QRDHRIK (SEQ ID NO:207)

28036	RSDDLTR (SEQ ID NO:134)	QSSDLRR (SEQ ID NO:209)	RSDHLSA (SEQ ID NO:135)	QHGALQT (SEQ ID NO:136)	TSGNLTR (SEQ ID NO:137)	QSGHLSR (SEQ ID NO:138)
28039	TSGSLSR (SEQ ID NO:129)	RSDALAR (SEQ ID NO:130)	RSDTLSQ (SEQ ID NO:206)	QRDHRIK (SEQ ID NO:207)	TSGNLTR (SEQ ID NO:137)	DRGDLRK (SEQ ID NO:219)
28040	DSSDRKK (SEQ ID NO:220)	TSGNLTR (SEQ ID NO:137)	DNYNRAK (SEQ ID NO:222)	DRSHLTR (SEQ ID NO:223)	RSDNLTT (SEQ ID NO:224)	N/A
28051	RSDNLSN (SEQ ID NO:225)	TSSSRIN (SEQ ID NO:226)	RSDNLSE (SEQ ID NO:142)	ASKTRKN (SEQ ID NO:143)	RSDALTQ (SEQ ID NO:229)	N/A
28052	RSDTLST (SEQ ID NO:230)	DRSSRIK (SEQ ID NO:231)	RSDDLSK (SEQ ID NO:232)	DNSNRIK (SEQ ID NO:233)	N/A	N/A
28053	QSSDLSR (SEQ ID NO:124)	QAGNLSK (SEQ ID NO:235)	QSGDLTR (SEQ ID NO:150)	TSGSLSR (SEQ ID NO:129)	QSGNLAR (SEQ ID NO:152)	N/A
28054	TSGSLSR (SEQ ID NO:129)	LRQTLRD (SEQ ID NO:240)	TSGNLTR (SEQ ID NO:137)	DRSALAR (SEQ ID NO:125)	RSDVLSE (SEQ ID NO:148)	RNFSLTM (SEQ ID NO:149)
28055	QSGDLTR (SEQ ID NO:150)	TSGSLSR (SEQ ID NO:129)	QSGNLAR (SEQ ID NO:152)	TSGSLSR (SEQ ID NO:129)	QSGSLTR (SEQ ID NO:154)	N/A
28056	DRSHLAR (SEQ ID NO:155)	TSGSLSR (SEQ ID NO:129)	LRQTLRD (SEQ ID NO:240)	TSGNLTR (SEQ ID NO:137)	DRSALAR (SEQ ID NO:125)	N/A

18. Use of a seed comprising the cell as defined in any one of claims 13-16 to generate a transgenic plant.

19. Use of a plant comprising the cell as defined in any one of claims 13-16 for producing seed.

20. Use of a plant comprising the cell as defined in any one of claims 13-16 to grow a crop.

		1 … 40
FAD3A (SEQ ID NO:7)	(1)	CATCAGACCCTTTCTTCACCACATTTCACTCAGAGCCCAC
FAD3A' (SEQ ID NO:8)	(1)	CATCGAACCCTTTCTTCACCACATTCCACTTCCCACACTC
FAD3C' (SEQ ID NO:12)	(1)	CATCGAACCCTTTCTTCACCACATTCCAGTTCCCACACTT
FAD3A'' (SEQ ID NO:9)	(1)	CATCAAAC-CTTTCTTCACCACATTTCACTGAAAGGCCAC
FAD3C'' (SEQ ID NO:11)	(1)	CATCAAAC-CTTTATTCACCACATTTCACTGAAAGGCCAC
FAD3C (SEQ ID NO:10)	(1)	CATCAAA--CTCTCTCCACCACATTTCACTCAGAGCCCAC
		41 … 80
FAD3A (SEQ ID NO:7)	(41)	ACAGTTTTAG------AGAGAGAGAGAAACATCCCTCAAA
FAD3A' (SEQ ID NO:8)	(41)	TCTTTTTTTTTGAATTATAGAGAGAGAATCCTCCTCCAAA
FAD3C' (SEQ ID NO:12)	(41)	TCTTTTTTTT-GAATTATAGAGAGAGAATCTTCCTCCAAA
FAD3A'' (SEQ ID NO:9)	(40)	ACATCT----------AGAGAGAGA--AACTTCGTCCAAA
FAD3C'' (SEQ ID NO:11)	(40)	ACATCT----------AGAGAGAGA--AACTTCGTCCAAA
FAD3C (SEQ ID NO:10)	(39)	ACAGTTTTAG------AGAGAGAGA--AACATCCCTCAAA
		81 … 120
FAD3A (SEQ ID NO:7)	(75)	GCTCTCTCTCTTTCTCCGGCGATGGTTGTCGCTATGGACC
FAD3A' (SEQ ID NO:8)	(81)	TCTCTCTCTCTC----CCAGGATGGTTGTTGCTATGGACC
FAD3C' (SEQ ID NO:12)	(80)	TCTCTCTCTCTCTCTCCCAGGATGGTTGTTGCTATGGACC
FAD3A'' (SEQ ID NO:9)	(68)	TCTCTCTC------TCCAGCAATGGTTGTTGCTATGGACC
FAD3C'' (SEQ ID NO:11)	(68)	TCTCTCTC------TCCAGCGATGGTTGTTGCTATGGACC
FAD3C (SEQ ID NO:10)	(71)	GCTCTCTC--TTTCTCCGGCGATGGTTGTCGCTATGGACC
		121 … 160
FAD3A (SEQ ID NO:7)	(115)	AGCGTAGCAATGCGAACGGAGA------------------
FAD3A' (SEQ ID NO:8)	(117)	AACGCACCAATGTGAACGGAGATGCCGGTGCCCGGAAGGA
FAD3C' (SEQ ID NO:12)	(120)	AACGCACCAATGTGAACGAAGATGCCGGTGCCCGGAAGGA
FAD3A'' (SEQ ID NO:9)	(102)	AGCGCAGCAATGTTAACGGAGATTCCGGTGCCCGGAAGGA
FAD3C'' (SEQ ID NO:11)	(102)	AGCGCAGCAATGTTAACGGAGATTCCGGTGCCCGGAAGGA
FAD3C (SEQ ID NO:10)	(109)	AGCGTAGCAATGTGAACGGAGATTCC---------AAGGA
		161 … 200
FAD3A (SEQ ID NO:7)	(137)	CGAAAGGTTTGATCCGAGCGCACAACCACCGTTCAAGATC
FAD3A' (SEQ ID NO:8)	(157)	AGAAGGGTTTGATCCGAGCGCACAACCGCCGTTTAAGATC
FAD3C' (SEQ ID NO:12)	(160)	AGAAGGGTTTGATCCGAGCGCACAACCGCCGTTTAAGATC
FAD3A'' (SEQ ID NO:9)	(142)	AGAAGGGTTTGATCCAAGCGAACAACCACCGTTTAAGATC
FAD3C'' (SEQ ID NO:11)	(142)	AGAAGGGTTTGATCCAAGCGCACAACCACCGTTTAAGATC
FAD3C (SEQ ID NO:10)	(140)	CGAAAGGTTTGATCCGAGCGCACAACCACCGTTTAAGATC
		201 … 240
FAD3A (SEQ ID NO:7)	(177)	GGAGATATAAGGGCGGCCATTCCTAAGCATTGTTGGGTAA
FAD3A' (SEQ ID NO:8)	(197)	GGGGACATAAGGGCTGCGATTCCTAAGCATTGTTGGGTGA
FAD3C' (SEQ ID NO:12)	(200)	GGGGACATAAGGGCTGCGATTCCTAAGCATTGTTGGGTGA

FIG. 1A

```
FAD3A'' (SEQ ID NO:9)   (182) GGAGATATCAGGGCGGCGATTCCTAAGCATTGTTGGGTGA
FAD3C'' (SEQ ID NO:11)  (182) GGAGATATAAGGGCGGCGATTCCTAAGCATTGCTGGGTGA
  FAD3C (SEQ ID NO:10)  (180) GGAGATATAAGGGCTGCGATTCCTAAGCATTGTTGGGTCA
                              241                                     280
   FAD3A (SEQ ID NO:7)  (217) AGAGTCCTTTGAGATCCATGAGCTATGTCGCCAGAGACAT
  FAD3A' (SEQ ID NO:8)  (237) AAAGTCCTTTGAGATCTATGAGCTACGTAGCCAGAGACAT
 FAD3C' (SEQ ID NO:12)  (240) AAAGTCCTTTGAGATCTATGAGCTACGTAGCCAGAGACAT
 FAD3A'' (SEQ ID NO:9)  (222) AGAGTCCTTTGAGATCTATGAGCTACGTCGCCAGAGACAT
FAD3C'' (SEQ ID NO:11)  (222) AGAGTCCTTTGAGATCTATGAGCTACGTCGCCAGAGACAT
  FAD3C (SEQ ID NO:10)  (220) AGAGTCCTTTGAGATCCATGAGCTACGTCGCGAGAGACAT
                              281                                     320
   FAD3A (SEQ ID NO:7)  (257) TTTCGCCGTCGTGGCTCTTGCCGTCGCCGCCGTGTATTTT
  FAD3A' (SEQ ID NO:8)  (277) TTGTGCCGTCGCGGCTTTGGCCATTGCCGCCGTGTATTTT
 FAD3C' (SEQ ID NO:12)  (280) TTGTGCCGTCGCTGCTTTGGCCATTGCCGCCGTGTATTTT
 FAD3A'' (SEQ ID NO:9)  (262) TTTCGCCGTCGCGGCTCTGGCCATGGCCGCCGTGTATTTT
```

FIG. 1A (CONT.)

```
FAD3C'' (SEQ ID NO:11)  (262) TTTCGCCGTCGCGGCTCTGGCCATGGCCGCCGTGTATTTT
  FAD3C (SEQ ID NO:10)  (260) TTTCTCCGTCGTGGCTCTGGCCGTCGCCGCCGTGTATTTT
                              321                                     360
   FAD3A (SEQ ID NO:7)  (297) GATAGCTGGTTCTTTTGGCCTCTTTATTGGGCCGCCCAAG
  FAD3A' (SEQ ID NO:8)  (317) GATAGCTGGTTCCTCTGTCCTCTCTATTGGGTCGCCCAAG
 FAD3C' (SEQ ID NO:12)  (320) GATAGCTGGTTCCTCTGGCCTCTCTATTGGGTCGCCCAAG
 FAD3A'' (SEQ ID NO:9)  (302) GATAGCTGGTTCCTCTGGCCACTCTACTGGGTTGCCCAAG
FAD3C'' (SEQ ID NO:11)  (302) GATAGCTGGTTCCTCTGGCCACTCTACTGGGTTGCCCAAG
  FAD3C (SEQ ID NO:10)  (300) GATAGCTGGTTCTTCTGGCCTCTTTATTGGGCCGCCCAAG
                              361                                     400
   FAD3A (SEQ ID NO:7)  (337) GAACCCTGTTCTGGGCTATCTTCGTACTCGGCCACGACTG
  FAD3A' (SEQ ID NO:8)  (357) GAACCCTTTTCTGGGCCATCTTCGTCCTCGGCCACGACTG
 FAD3C' (SEQ ID NO:12)  (360) GAACCCTTTTCTGGGCCATCTTCGTCCTCGGCCACGACTG
 FAD3A'' (SEQ ID NO:9)  (342) GAACCCTTTTCTGGGCCATCTTCGTTCTTGGCCACGACTG
FAD3C'' (SEQ ID NO:11)  (342) GAACCCTTTTCTGGGCCATCTTCGTTCTTGGCCACGACTG
  FAD3C (SEQ ID NO:10)  (340) GAACCCTTTTCTGGGCCATCTTCGTACTCGGCCACGACTG
                              401                                     440
   FAD3A (SEQ ID NO:7)  (377) GTAATTTAATTTT----------TCTTTCAACTTCTTAA
  FAD3A' (SEQ ID NO:8)  (397) GTAA----AGTTT--------------------------
 FAD3C' (SEQ ID NO:12)  (400) GTAA----AGTTT--------------------------
 FAD3A'' (SEQ ID NO:9)  (382) GTAAATTAAATTT----------------------TCTG
FAD3C'' (SEQ ID NO:11)  (382) GTAAATTAAATTT----------------------TCAG
  FAD3C (SEQ ID NO:10)  (380) GTAATTTAATTTCAATTTATTTTTTCTTCAACTTCTTAA
                              441                                     480
```

FIG. 1B

```
FAD3A (SEQ ID NO:7)    (406) TTTTGATATGTTTATATGTTTTTTTCGTTTTTTGCATTGT
  FAD3A' (SEQ ID NO:8)    (406) CTTCCAT----------------------TTTGCATTGC
 FAD3C' (SEQ ID NO:12)    (409) CTTCCAT----------------------TTTGCATTGC
 FAD3A'' (SEQ ID NO:9)    (399) TTTTAAT-------TATTTTGACT-CTTTTTGTTCAATTT
FAD3C'' (SEQ ID NO:11)    (399) TTTTAAT-------TATTTTGTCT-CTTTTTGTTCAATTT
  FAD3C (SEQ ID NO:10)    (420) TTTTGATATGTTTATATGTTTTTT-CGTTTTTTGCATCGT
                                  481                                 520
   FAD3A (SEQ ID NO:7)    (446) CTTTGATTTCTTGACCGTACGTTCGATATGAGATTTTC--
  FAD3A' (SEQ ID NO:8)    (423) ATCG-ATTTATTGAATGCACGTTCTACGAGT-ATTGTTTG
 FAD3C' (SEQ ID NO:12)    (426) ATCG-ATTTATTGAATGCACGTTCTATGAGT-ATTGT---
 FAD3A'' (SEQ ID NO:9)    (431) ATTA-ATTTCTTGAATGCACGTTCGATGAGT-ATCGTCGT
FAD3C'' (SEQ ID NO:11)    (431) ATTA-ATTTCTTGAATGCACGTTCGATGAGT-ATCGTC--
  FAD3C (SEQ ID NO:10)    (459) CTTTGATTTCTTGAACGCACGTTCGATATGAGATTTTC--
                                  521                                 560
   FAD3A (SEQ ID NO:7)    (484) --ACTGACTTCAAGATTTGATTCTCTTCAGGTTTACTTTT
  FAD3A' (SEQ ID NO:8)    (461) TCAGTTACTTCGTAAAATGATTCTTTTGATGTTCATTTTT
 FAD3C' (SEQ ID NO:12)    (461) -CAGT-ACTTTATGAATTGATTCTTTTGATGTTCATTTTT
 FAD3A'' (SEQ ID NO:9)    (469) -CACTGACTTCAAGATTTAATTCTTTTGAGGTT-ACCTTT
FAD3C'' (SEQ ID NO:11)    (467) --ACTGACTTCAAGATTTAATTCTTTTGAGGTT-ACTTTT
  FAD3C (SEQ ID NO:10)    (497) --ACTGACTTCAAGATTTGATTCTCTTCAGGTTTACTTTT
                                  561                                 600
   FAD3A (SEQ ID NO:7)    (522) TTCAATTTTAATTATTATGTTCATCCAATTTGGCCTATTT
  FAD3A' (SEQ ID NO:8)    (501) TGAAGATCTAAG-ATTT--------------------TTT
 FAD3C' (SEQ ID NO:12)    (499) TGAAGATCTAAG-ATTT--------------------TTT
 FAD3A'' (SEQ ID NO:9)    (507) T-CATGTTCAATTATTA---------AA--------AAAT
FAD3C'' (SEQ ID NO:11)    (504) T-CATGTTTAATTATTA---------AA--------AAAT
  FAD3C (SEQ ID NO:10)    (535) AAAAAAAAAAATTATTATGTTCACCCAAATTGGCCTATTT
                                  601                                 640
   FAD3A (SEQ ID NO:7)    (562) TAAAAGCAAAAGGGGATCTAAGATTTTTAATTCTTTGTT
  FAD3A' (SEQ ID NO:8)    (520) T---------------TTT-AGATTTTCT-TTTTAAATCA
 FAD3C' (SEQ ID NO:12)    (518) T---------------TTTTAGATTTTCT-TTTTAAATCA
```

FIG. 1B (CONT.)

```
FAD3A'' (SEQ ID NO:9)   (529) AAAATAAAATATAGGATCTAAGATTTTT--TTCTTCATCA
FAD3C'' (SEQ ID NO:11)   (526) AAAAGAAAATATAGGATCTAAGATTTTT--TTCTTCATCA
  FAD3C (SEQ ID NO:10)   (575) TAAAAGCAAAAGGGGATCTAAGATTTTTAATTCTTCTCTT
                               641                                     680
   FAD3A (SEQ ID NO:7)   (602) TTTTTTTGGT---------------TCTTTTTCATCAG-T
  FAD3A' (SEQ ID NO:8)   (543) TTGTTCCACCACCA-------------CCTTTCATCGG-T
 FAD3C' (SEQ ID NO:12)   (542) TTGTTCCACCACC----------------TTTCATCGG-T
 FAD3A'' (SEQ ID NO:9)   (567) --GTTCAAGCA-------------TCATCACTCATCAG-T
FAD3C'' (SEQ ID NO:11)   (564) ATGTTCAAGCA-------------TCGTCACTCATCAG-T
  FAD3C (SEQ ID NO:10)   (615) TTTCAGTCGTAACACTGCTAACTTTTTTTTTTGATCAAAT
                               681                                     720
   FAD3A (SEQ ID NO:7)   (626) CGTAACACTC-------CTAACTAAACATCTTTTTCTTTC
  FAD3A' (SEQ ID NO:8)   (569) CGTACGACTC----GTTACAACACCACATCTT--TATTTT
 FAD3C' (SEQ ID NO:12)   (565) CGTACGACTC----GTTACAAAACCACATCTT--TATTTT
 FAD3A'' (SEQ ID NO:9)   (591) CGTAAGACTC-------GTAACAAAATATCTT---CTTTT
FAD3C'' (SEQ ID NO:11)   (590) CGTCAGACTC-------GTAACAAAATATCTT---CTTTT
  FAD3C (SEQ ID NO:10)   (655) CGTAACACTCATAAGTCCTAACTAAACATCTTTTTCTTTC
                               721                                     760
   FAD3A (SEQ ID NO:7)   (659) CTATAATTATTGTTGTTTCCGCGTTTTATGGATCTACGTT
  FAD3A' (SEQ ID NO:8)   (603) CTATAATTACTACTGCTTCCGCATTTTATGGATCTCTCAA
 FAD3C' (SEQ ID NO:12)   (599) CTATAATTACGACTGCTTCCGCATTTTATGGATCTCTCAA
 FAD3A'' (SEQ ID NO:9)   (621) CTATAATTAATATTATTCCGCATTTAATGGATCTACGTT
FAD3C'' (SEQ ID NO:11)   (620) CTATAATTAATATTATTCCGCATTTTATGGATCTACGTT
  FAD3C (SEQ ID NO:10)   (695) CTATAATTATTGTTGGTTCCGCATTTTATGGATCTACGTT
                               761                                     800
   FAD3A (SEQ ID NO:7)   (699) T-GAAATTTTCAA----------------TAAAAC---
  FAD3A' (SEQ ID NO:8)   (643) CTTATAATTAAAG----------------TATAATATC
 FAD3C' (SEQ ID NO:12)   (639) CTTATAATTAAAG----------------TATAAAATC
 FAD3A'' (SEQ ID NO:9)   (661) TTGATGTTCTCAAATTTTGTTTCTCTTTCTCTAGATCCCC
FAD3C'' (SEQ ID NO:11)   (660) TTGATGTTCTCAATTTTGTTTCTCTTTCTCTAGATCCCC
  FAD3C (SEQ ID NO:10)   (735) T-GAAAGTTTCAA----------------TAAAAC---
                               801                                     840
   FAD3A (SEQ ID NO:7)   (717) ---ACATTTTATTGTT-TTCT-GTA----ACAATTT----
  FAD3A' (SEQ ID NO:8)   (665) AAGAATATCTATTATTTTTCTTAAACAAGA-AAGAT----
 FAD3C' (SEQ ID NO:12)   (661) AAGAATATCTATTGTTTTTCTAAAACAAGA-AAGAT----
 FAD3A'' (SEQ ID NO:9)   (701) GGAACTTTTAATTATAATTATAGTATAGTATAATATCAAG
FAD3C'' (SEQ ID NO:11)   (700) GGAACTTTTAATTATAATTATAGTATAGTATAATATCAAG
  FAD3C (SEQ ID NO:10)   (753) ---ACATTTTATTGTT-TGAAAGTA----ACAATAT----
                               841                                     880
```

FIG. 1C

```
   FAD3A (SEQ ID NO:7)   (744) --AAT-TACTGTTTATTGGTTC----------------
  FAD3A' (SEQ ID NO:8)   (700) --AAT--ATTGTTTCTTTGTTA----------------
 FAD3C' (SEQ ID NO:12)   (696) --AAT--ATTGTTTCTTTGTTA----------------
 FAD3A'' (SEQ ID NO:9)   (741) AAAATATACTGTTTATTTTTTTGGCAACAAATATATTAC
FAD3C'' (SEQ ID NO:11)   (740) AAAATATACTGTTTATTTTTTT-GGCAACAAATATATT--
  FAD3C (SEQ ID NO:10)   (781) --AAT-TACTGTATATTGATTC----------------
                               881                                 920
   FAD3A (SEQ ID NO:7)   (763) ----TTTT----------------A-----------ATTA
  FAD3A' (SEQ ID NO:8)   (718) ------------------------------------TTT
 FAD3C' (SEQ ID NO:12)   (714) ------------------------------------TTT
 FAD3A'' (SEQ ID NO:9)   (781) TCTTGTTTCTTTGACAAGAAAAAAATATATTGTTTTTTTC
FAD3C'' (SEQ ID NO:11)   (777) ----GTTT-TTTGACAAGAAAAA--TATATTGTTTTTTTC
  FAD3C (SEQ ID NO:10)   (800) ----TTTT----------------A-----------ATTA
                               921                                 960
   FAD3A (SEQ ID NO:7)   (772) TTGTGTGT-TGTTCCAATCTATTTCGAAATATAGTCATG
  FAD3A' (SEQ ID NO:8)   (721) TGGTGTAT---TTCCAATCTA-TTTCGAGATTTAGAAATG
```

FIG. 1C (CONT.)

```
FAD3C' (SEQ ID NO:12)   (717) TGGTGTAT---T-CCAATCTA-TTTCGAGATTTAGAAATG
 FAD3A'' (SEQ ID NO:9)   (821) TTCTTTTTGTGTTTCCAATCTATTTTCGAGATTTAGACAAG
FAD3C'' (SEQ ID NO:11)   (810) TTCTTTTTGTGTTTCCAATCTATTTT-GTGATTTAGACAAG
  FAD3C (SEQ ID NO:10)   (809) TTGTGTGT-TGTTCCAATCTACTTTCGAAATATAGTCATG
                                961                                  1000
   FAD3A (SEQ ID NO:7)   (811) TGACACGTCATATTCTATTTTTGTTACCTTGTTGAAACGT
  FAD3A' (SEQ ID NO:8)   (757) TGACACGTCAT-------------TACCTTGTTGAAGTGT
 FAD3C' (SEQ ID NO:12)   (752) TGTCACGTCAT-------------TACCTTGTTGAAGCTT
 FAD3A'' (SEQ ID NO:9)   (861) TGACACGTCATATACCGGATTTGTTACCTTGTTAAAGAGT
FAD3C'' (SEQ ID NO:11)   (849) TGACACGTCATATACCGGATTTGTTACCTTGTTAAAGAGT
  FAD3C (SEQ ID NO:10)   (848) TGACACGTCATATTCTATTTTTGTTACCTTGTTGGAACGT
                                1001                                 1040
   FAD3A (SEQ ID NO:7)   (851) TTG----------AATTGAGGAAAGTTCAGTTAACATTGT
  FAD3A' (SEQ ID NO:8)   (784) TTA------AAACAAACATGGAAAGTTTAAATAA-ATAGT
 FAD3C' (SEQ ID NO:12)   (779) TTA------AAACAAACATGGAAAGTTTAAATAA-ATAGT
 FAD3A'' (SEQ ID NO:9)   (901) TTGGGTTAAAACAAATGTAGAAAAGTTAAAATAA-ATTGT
FAD3C'' (SEQ ID NO:11)   (889) TTGAGTTAAAACAAATGTAGAAAAGTTAAAATAA-ATTGT
  FAD3C (SEQ ID NO:10)   (888) TTG----------AATTGAGTAAAGTTTAATTAACATTGT
                                1041                                 1080
   FAD3A (SEQ ID NO:7)   (881) GCAATAAATGATAAA-TGTGTTT----------ATGAT
  FAD3A' (SEQ ID NO:8)   (817) GCAATAAATGATATA-TATGTAT--ATGATGAATAATGAT
 FAD3C' (SEQ ID NO:12)   (812) GCAATAAATGATATACTATATTT--ACGATGAATAATGAT
 FAD3A'' (SEQ ID NO:9)   (940) GCAATAAATGATAAA-TACGTTTTTATGTTAAACAATGAT
FAD3C'' (SEQ ID NO:11)   (928) GCACTAAATGATAAA-TACGTTTTTATGTTAAATAATGAT
  FAD3C (SEQ ID NO:10)   (918) GCAATAAATGATAAA-CATGTTT----------ATGAT
                                1081                                 1120
   FAD3A (SEQ ID NO:7)   (908) GTAAAATTTCATTTGAATAATA-CAGTGGACATGGGAGCT
  FAD3A' (SEQ ID NO:8)   (854) GTGAAA-TATAATTGAATAATGGCAGTGGACATGGGAGTT
 FAD3C' (SEQ ID NO:12)   (850) GTGAAA-TATAATTGAATAATGGCAGTGGACATGTGAGTT
 FAD3A'' (SEQ ID NO:9)   (979) GTGAAAATAAAATTGAATAATGGCAGTGGACATGGGAGTT
FAD3C'' (SEQ ID NO:11)   (967) GTGAAAATAAAATTGAATAATGGCAGTGGACATGGGAGTT
  FAD3C (SEQ ID NO:10)   (945) GTAAAATTCAATTTGAATAATA-CAGTGGACATGGGAGCT
                                1121                                 1160
   FAD3A (SEQ ID NO:7)   (947) TCTCAGACATTCCTCTTCTGAATACTGCGGTTGGTCATAT
  FAD3A' (SEQ ID NO:8)   (893) TCTCAGACATTCCTCTGCTGAATAGTGTGGTTGGCCATAT
 FAD3C' (SEQ ID NO:12)   (889) TCTCAGACATTCCTCTGCTGAATAGCGTGGTTGGCCATAT
 FAD3A'' (SEQ ID NO:9)  (1019) TTTCAGACATTCCTCTGCTGAACAGTGTGGTTGGTCACAT
FAD3C'' (SEQ ID NO:11)  (1007) TCTCAGACATTCCTCTGCTGAACAGTGTGGTTGGTCACAT
  FAD3C (SEQ ID NO:10)   (984) TCTCAGACATTCCTCTTCTGAATACTGCGGTTGGTCATAT
```

FIG. 1D

```
                                    1161                                1200
   FAD3A (SEQ ID NO:7)    (987) TCTTCATTCCTTCATTCTCGTTCCATACCATGGTTGGTAA
  FAD3A' (SEQ ID NO:8)    (933) TCTTCATTCCTTCATCCTCGTTCCTTACCATGGTTGGTAA
 FAD3C' (SEQ ID NO:12)    (929) TCTTCATTCCTTCATCCTCGTTCCTTACCATGGTTGGTAA
 FAD3A'' (SEQ ID NO:9)   (1059) TCTTCATTCATTCATCCTCGTTCCTTACCATGGTTGGTAA
FAD3C'' (SEQ ID NO:11)   (1047) TCTTCATTCATTCATCCTCGTTCCTTACCATGGTTGGTAA
  FAD3C (SEQ ID NO:10)   (1024) TCTTCATTCCTTCATTCTCGTTCCATACCATGGTTGGTAA
                                    1201                                1240
   FAD3A (SEQ ID NO:7)   (1027) GTCATTTATTTTAACTTCTTTTTCATGCAAA---TTTA
  FAD3A' (SEQ ID NO:8)    (973) GTCAGCTTATC--AACC-CTTTTT--ACTAT-ATTATTAA
 FAD3C' (SEQ ID NO:12)    (969) GTCAACTTATT--AACC-CTTTTT--ATTATTATTATTAA
 FAD3A'' (SEQ ID NO:9)   (1099) GTCATTTATT--AAC---TATTTCCATGTAAACTATTAG
FAD3C'' (SEQ ID NO:11)   (1087) GTCATTTATT--AAC---TATTTCCATGTAAATTATTAG
  FAD3C (SEQ ID NO:10)   (1064) GTCATTTATTTAAACATCTTTTT-CATGCAAA---TTTA
                                    1241                                1280
   FAD3A (SEQ ID NO:7)   (1063) TTCTTGTTTTCGTATTTCTTACATTTTCCTT-GTCATTCT
```

FIG. 1D (CONT.)

```
 FAD3A' (SEQ ID NO:8)   (1007) TTATTAAACTTGCATTTGT-ATACTT-----GGTGCAAGT
 FAD3C' (SEQ ID NO:12)   (1004) TTATTAAACTTTCATTTGTTATACTTTTTTGGTTTAAAT
 FAD3A'' (SEQ ID NO:9)   (1133) TACTTGTTTTCGTATTTCTTACATTTTCGTTTGTCATTCT
FAD3C'' (SEQ ID NO:11)   (1121) TACTTGTTTTCGTATTTCTTACATTTTCGTTTGTTATTCT
  FAD3C (SEQ ID NO:10)   (1099) TTCTTGTTTTCGTATTTCTTACATTTTCCTT-GTCATTCT
                                    1281                                1320
   FAD3A (SEQ ID NO:7)   (1102) T----GGTGCA-TGTTAGCAAACAGTAATCTGA--TAACT
  FAD3A' (SEQ ID NO:8)    (1041) TGGTAAATGTAATCTGATAACTGAA-AATCTAT--TCATT
 FAD3C' (SEQ ID NO:12)    (1044) -GTTAAATGAATTACTTGGTGCAAG-AATCTAT--TCATT
 FAD3A'' (SEQ ID NO:9)   (1173) TCTTGGGTGCA-TGCTAGCAAACTGTAATCAGTATTAACT
FAD3C'' (SEQ ID NO:11)   (1161) T---GGGTGCAATGCTAGGAAACTGTAATCAGTATTAACT
  FAD3C (SEQ ID NO:10)   (1138) T----GGTGCA-TGTTAGCAAACTGTAATCTGA--TAACT
                                    1321                                1360
   FAD3A (SEQ ID NO:7)   (1135) GAAAA----------TATATTAATT-------------TT
  FAD3A' (SEQ ID NO:8)    (1078) GCTCGTTCTA-----TTTTTTTTTTGGCT-AGAGACAATT
 FAD3C' (SEQ ID NO:12)    (1080) GCTCGTTCT------TTTTTTTTTTGGCT-AGAGCCAATT
 FAD3A'' (SEQ ID NO:9)   (1212) GGGAACTACCAACTGTTTTTTTTTTGCTAGAGTAGCAATT
FAD3C'' (SEQ ID NO:11)   (1198) GGAAGCTACCAACT-TTTTTTTGTTGCTAGAGTAGCAATT
  FAD3C (SEQ ID NO:10)   (1171) GAAAA----------TATATTAATT-------------TT
                                    1361                                1400
   FAD3A (SEQ ID NO:7)   (1152) TCATAGTAAAATAA-----------------TGCATGTG
  FAD3A' (SEQ ID NO:8)    (1112) TTATAATTAAATAATGCATGTGAGAATATGACTATTTATG
 FAD3C' (SEQ ID NO:12)    (1113) TTATAATTAAATAATGCATGTGAAAGTATGACTATATATG
```

FIG. 1E

```
FAD3A'' (SEQ ID NO:9)  (1252) TTATAATTAAATAAGAATCCTATTA--AACAATGCATGTG
FAD3C'' (SEQ ID NO:11) (1237) TTATAATTAAATAAGAATCCTATTA--AACAATGCATGTG
  FAD3C (SEQ ID NO:10) (1188) CCATAGTAAAATAA----------------TGCATGTG
                               1401                                1440
   FAD3A (SEQ ID NO:7) (1174) ACTAAAAGCA--------------TCAAAA--------TC
  FAD3A' (SEQ ID NO:8) (1152) TGAGGTAGCTTTTCTTATTCCTGTCGAAAAGCATCAAATC
 FAD3C' (SEQ ID NO:12) (1153) TGAGGTAGCTTTTCTTATTCTTGACGAAAAGCATCGAATC
 FAD3A'' (SEQ ID NO:9) (1290) ACAATATGAGGTTGCTTTT-CTGTTCAAAA----CAAATC
FAD3C'' (SEQ ID NO:11) (1275) ACTATATGAGGTTGCTTTTTCTGTTCAAAAGCATCAAATC
  FAD3C (SEQ ID NO:10) (1210) ACTAAAAGCA--------------TCAAAA--------TC
                               1441                                1480
   FAD3A (SEQ ID NO:7) (1192) TTTAGCATCGAAGAAAAAAGAA-CCAAACTTTTATTT--A
  FAD3A' (SEQ ID NO:8) (1192) TTTAGCAACGAAGGAAAAAGGAATCAAATTTTTTATT-AA
 FAD3C' (SEQ ID NO:12) (1193) TTTAGCAACGAAGGAAAAAGGAATCAAAACTTTTATT-AA
 FAD3A'' (SEQ ID NO:9) (1325) TTTAGAAGCCAATGAAAAAGAATCCAAAACTTTTTTTAA
FAD3C'' (SEQ ID NO:11) (1315) TTTAGCAGCCAATGAAAAAGAATCCAAACCTTTTCTT-AA
  FAD3C (SEQ ID NO:10) (1228) TTTAGCATCGAAGAAAAAAGAA-CCAAACTTTTATT--A
                               1481                                1520
   FAD3A (SEQ ID NO:7) (1229) ATGCTATGGGCCTATTTATGG--------TCCA-------
  FAD3A' (SEQ ID NO:8) (1231) ATGCAATGGGTCTATGTCTTGG-------TCATTAGTTTT
 FAD3C' (SEQ ID NO:12) (1232) ATGCAATGGGCCTATATCT-GG-------TCATTAGTATT
 FAD3A'' (SEQ ID NO:9) (1365) ATGATATGCGCCTATCTATTGGTCCTGACTCCTGAGTTTT
FAD3C'' (SEQ ID NO:11) (1354) ATGATATGCGCCTATCTATGG--------TCCTGAGTTTT
  FAD3C (SEQ ID NO:10) (1265) ATGCTATGGGCCTATTTATGG--------TCCA-------
                               1521                                1560
   FAD3A (SEQ ID NO:7) (1254) --------A--TTAGCTATTATCATATGAC-ATGTCCTTG
  FAD3A' (SEQ ID NO:8) (1264) TTGCATATAATTTATTTATATTTTTTTCTTAACAGCAGCT
 FAD3C' (SEQ ID NO:12) (1264) TTGAATATAATTTATTTATAATTTTTTTTGAACAACAGCT
 FAD3A'' (SEQ ID NO:9) (1405) CTTACTTTC--TTAAGTATAATTAGATTTTGATTTTTTTT
FAD3C'' (SEQ ID NO:11) (1386) CTTAGTTCA--TTAAGTATAATTAGATTTTGATTTTTTTT
  FAD3C (SEQ ID NO:10) (1290) --------A--TTAGCTATTATCATATGAC-ATGTCCTTG
                               1561                                1600
```

FIG. 1E (CONT.)

```
   FAD3A (SEQ ID NO:7) (1283) AA--------TAAATTAATGT-A----------TAAGTTT
  FAD3A' (SEQ ID NO:8) (1304) AATTTAATTATAATTAAATATTCATTTTATAAATAATATT
 FAD3C' (SEQ ID NO:12) (1304) AATTTATTTATAATTAAATATTCATTTTATAAATAATATT
 FAD3A'' (SEQ ID NO:9) (1443) TATAGGTTT-TCACT-ATTGTTATTTGTTTACATCAGCTT
FAD3C'' (SEQ ID NO:11) (1424) TA--GGTTT-TCACTTATTGTTATTTGTTTACATCAGCTT
  FAD3C (SEQ ID NO:10) (1319) AA--------TAAATTAATGT-AGCTTCATATGTGAGTTT
                               1601                                1640
```

FIG. 1F

```
FAD3A (SEQ ID NO:7)  (1304) AATAT----------------------AATATTTAT--A
  FAD3A' (SEQ ID NO:8)  (1344) AGACCAATTATTAAAGGTTAGATATTTTAAGAATTATTCA
 FAD3C' (SEQ ID NO:12)  (1344) AAACCAATTATTAAAGGTTAGATATTTGAAGAATTATTCA
 FAD3A'' (SEQ ID NO:9)  (1481) CAGATATCTTCGAAA-------------AAGATTTAC--A
FAD3C'' (SEQ ID NO:11)  (1461) CAAACATCTTCGAAA-------------AAGACTTAC--A
  FAD3C (SEQ ID NO:10)  (1350) AAT----------------------AATATTTAT--A
                               1641                                1680
   FAD3A (SEQ ID NO:7)  (1319) TATATTTGTTT---------TAATGGCTTAT---TTTA-T
  FAD3A' (SEQ ID NO:8)  (1384) TGACTTTGTTTATTGGAA-----CTCCTTTTATCTTTTAA
 FAD3C' (SEQ ID NO:12)  (1384) TGACTTTGTTTATTGGGAAATTACTCCTTTTATCTTTTAT
 FAD3A'' (SEQ ID NO:9)  (1506) TGCATCAATTTCATGAGGATTTATAGTTTTCT-TTTACT
FAD3C'' (SEQ ID NO:11)  (1486) TGCATCAATTTCCTGAGGATTTATAGTTTT---TTTACT
  FAD3C (SEQ ID NO:10)  (1363) TATTTTTGTTT---------TAATGGCTTAT---TTTA-T
                               1681                                1720
   FAD3A (SEQ ID NO:7)  (1346) TGTTA-------AATGGATAC-----ATCAGCTTGAAATA
  FAD3A' (SEQ ID NO:8)  (1419) TCTTTT---CTATTTCTCCATTTTTTAATAATGAGAAACTG
 FAD3C' (SEQ ID NO:12)  (1424) TCTTTT---CTATTTCTCTATTTTTTAATATTGAGAAACTG
 FAD3A'' (SEQ ID NO:9)  (1545) TATTTCCGACACAATGTTTAGTAGTAAAAAGCATTAAATG
FAD3C'' (SEQ ID NO:11)  (1523) TATTTCTG-CACAATGTTTATTAGTAAAAAGCATCAAATG
  FAD3C (SEQ ID NO:10)  (1390) TGTTA-------AATGGATAC-----ATCAGCTTGAAATG
                               1721                                1760
   FAD3A (SEQ ID NO:7)  (1374) TCT----------ACGAACAT-GCATCATTTTCCTAGAT
  FAD3A' (SEQ ID NO:8)  (1456) ACTTCAAATCTCCAATAAAGATGGTCTTATGTAGTAACAG
 FAD3C' (SEQ ID NO:12)  (1461) ACTTCAAACCTCCAATAAAAATGGTTTCCTGTAGTAACAT
 FAD3A'' (SEQ ID NO:9)  (1585) TTTTTTTG-CTCAAAAAAAAA-GAATGGGATTGTTAGAG
FAD3C'' (SEQ ID NO:11)  (1562) TTTTTTTG-CTCAAAAAAA---GAATGGGATTGTTAGAG
  FAD3C (SEQ ID NO:10)  (1418) TCT----------ACGAACAT-GCATCATTTTCCTAGAT
                               1761                                1800
   FAD3A (SEQ ID NO:7)  (1402) A---CATTTGTTTGTTGCTCAAAAAATGAATAACGTAGTT
  FAD3A' (SEQ ID NO:8)  (1496) TA-TAATTTTTTGTTTGGTAAATGTAACATCATCTTCAAA
 FAD3C' (SEQ ID NO:12)  (1501) CA-TAATTTTTTGTTTGGTAAATGTAACATCATCTTCAAA
 FAD3A'' (SEQ ID NO:9)  (1623) CACTCTATTGTTAGTTGTTCAATAAATATACCAACTAAAA
FAD3C'' (SEQ ID NO:11)  (1598) CACTCTATTGTTAGTTGTTCAATAAATATATCAACTAAAA
  FAD3C (SEQ ID NO:10)  (1446) A---CACTTGTTTGTTGCTCAAAAA-TGAATAACTTAGTT
                               1801                                1840
   FAD3A (SEQ ID NO:7)  (1439) AAAC-------------------GAGTGAGA---------
  FAD3A' (SEQ ID NO:8)  (1535) TATCTTTGAAAATAGACTTACATGCATTATTTTGCTGCGA
 FAD3C' (SEQ ID NO:12)  (1540) TATCTTTGAAAATAGACTTACATGCATTATTTTGCTGCGA
 FAD3A'' (SEQ ID NO:9)  (1663) AAACAAAATAAATATA---AAATGAGTGAGATTGTTAAAT
FAD3C'' (SEQ ID NO:11)  (1638) AAACAAAATAAATATA---AAATGAGTGAGATTGTTAAAT
  FAD3C (SEQ ID NO:10)  (1482) AAAC-------------------GAGTGAGCATGTTCTAT
```

FIG. 1F (CONT.)

```
                                1841                                1880
   FAD3A (SEQ ID NO:7)  (1451) ----------------------TTCTTAG------------
  FAD3A' (SEQ ID NO:8)  (1575) CATTATTGTCACTTATTCCTGGCAATAAAT-TAGTTTATT
 FAD3C' (SEQ ID NO:12)  (1580) CATTATTGTAACTTATTCCTGGCAATAAAAATAATTTATT
 FAD3A'' (SEQ ID NO:9)  (1700) CATTATAGAGACAATTTCATTTTCACAAAAATAAATAAAT
FAD3C'' (SEQ ID NO:11)  (1675) CATTATAGAGACAATTTCATTTTCACAAAAATAAATAAAT
  FAD3C (SEQ ID NO:10)  (1503) GGGG----------------TTTCTTAGAGCATGATTATT
                                1881                                1920
```

FIG. 1F (CONT.)

```
FAD3A (SEQ ID NO:7)  (1458) ----------------------------------------
  FAD3A' (SEQ ID NO:8)  (1614) ACTG-AACTTTTTTTGGTCAATTTATTACTAGTAACTTT
 FAD3C' (SEQ ID NO:12)  (1620) ACTGGAAACTATTTTTGGTCAATTTATTACTAGTAACTTA
 FAD3A'' (SEQ ID NO:9)  (1740) ACAT--AACTTTTTATAATTGGGGTTTGCAGGAGAATAAG
FAD3C'' (SEQ ID NO:11)  (1715) ACAT--AACTTTTG-TAATTGGGGTTTGCAGGAGAATAAG
  FAD3C (SEQ ID NO:10)  (1527) GAGA--AGTTCCTA-GAGTGAGGTTCTTACCGGAATATAA
                                1921                                1960
   FAD3A (SEQ ID NO:7)  (1458) ----------------------------------------
  FAD3A' (SEQ ID NO:8)  (1653) AAACTTAAAAGAGTGAGATTGTTTGATCAAAAAAAAT---
 FAD3C' (SEQ ID NO:12)  (1660) AAACTTAAAAGAGTGAGATTGTTTGATCAAAAAAAAAGAG
 FAD3A'' (SEQ ID NO:9)  (1778) CCATCGGACACACCACCAGAACCATGGCCATGTTGAAAAC
FAD3C'' (SEQ ID NO:11)  (1752) CCATCGGACACACCACCAGAACCATGGCCATGTTGAAAAC
  FAD3C (SEQ ID NO:10)  (1564) GAATCTATCTCTTAACTTTTAACTAAAAAAATTAAGAACC
                                1961                                2000
   FAD3A (SEQ ID NO:7)  (1458) ----------------------------------------
  FAD3A' (SEQ ID NO:8)  (1690) ---AAAAATAGAGTGAGATAGTTAGAATCTGCCATGAAAG
 FAD3C' (SEQ ID NO:12)  (1700) AAAAAAAATAGAGTGAGATTGTTAGAATCTGCCATGAAAG
 FAD3A'' (SEQ ID NO:9)  (1818) GACGAGTCTTGGGTTCCGGTAATCTTTCCTACTCTCGTAG
FAD3C'' (SEQ ID NO:11)  (1792) GACGAGTCTTGGGTTCCGGTAATCTTTCCTACTCTCATTG
  FAD3C (SEQ ID NO:10)  (1604) GGCTTTTAAAACTCGTATTTAAGAACCGTTTTTTAGTTTT
                                2001                                2040
   FAD3A (SEQ ID NO:7)  (1458) ----------------------------------------
  FAD3A' (SEQ ID NO:8)  (1727) CAACACTATATAG--------------------------
 FAD3C' (SEQ ID NO:12)  (1740) CAACACTATATAGGTGATGATTGGTTCGACTGTGGCCGTA
 FAD3A'' (SEQ ID NO:9)  (1858) TTTCTCTTGTCTTTTATTTATTTGTTTGTTTTTCGGAATT
FAD3C'' (SEQ ID NO:11)  (1832) TTTCTCTTGTCTTTTATTTATTTGTTCTTTTTTGGGAATT
  FAD3C (SEQ ID NO:10)  (1644) TTTAGTTAAAAATCAAGAGACGAGTTCTTATATTCCGCTA
                                2041                                2080
   FAD3A (SEQ ID NO:7)  (1458) ----------------------------------------
  FAD3A' (SEQ ID NO:8)  (1740) ----------------------------------------
 FAD3C' (SEQ ID NO:12)  (1780) GAATTTTAGCTGTAGATAAATTGGTTGTAGTTGTAAAGTT
 FAD3A'' (SEQ ID NO:9)  (1898) TATTCTTA--TGTC--TATGTTCTTAGGATTCTATATGTT
FAD3C'' (SEQ ID NO:11)  (1872) CATTCTTA--TGTC--TAAGTTCTTATGATTATTGAAGTT
```

FIG. 1G

```
FAD3C (SEQ ID NO:10)  (1684) AGAACTCC--ACCC--TGAGAACTTCTCAATAATCATGCT
                                2081                                2120
   FAD3A (SEQ ID NO:7)  (1458) ------------------------------------
  FAD3A' (SEQ ID NO:8)  (1740) ------------------------------------
 FAD3C' (SEQ ID NO:12)  (1820) GTTACTGTT-GATTATTTTTGCAGAGACTTTTGCTGTAGT
 FAD3A'' (SEQ ID NO:9)  (1934) TATTTTATTAGTTTATGTTTTCAGTCTGAGGTCA-GACCG
FAD3C'' (SEQ ID NO:11)  (1908) CTTAAGGTGGGGTTCTTAACGGAATATGAGAACCTGTCTC
  FAD3C (SEQ ID NO:10)  (1720) CTTAGTGCTCTAAGAAGGGTCCTTAACAAAATAT------
                                2121                                2160
   FAD3A (SEQ ID NO:7)  (1458) ------------------------------------
  FAD3A' (SEQ ID NO:8)  (1740) ------------------------------------
 FAD3C' (SEQ ID NO:12)  (1859) TAAATTTGTTGTAGCTGTAAGCTATAGGCTGCAGATATTT
 FAD3A'' (SEQ ID NO:9)  (1973) ACCACTTGTCAG-------ATCTGTTTTCTAGCTGT--AG
FAD3C'' (SEQ ID NO:11)  (1948) TTAACTTTTAACTAAAA-AAGCTAAGAACCAGCTTTTAAA
  FAD3C (SEQ ID NO:10)  (1754) --TAATAATAAG-------ATATAGTGTGGGCCCAA----
                                2161                                2200
   FAD3A (SEQ ID NO:7)  (1458) ------------------------------------
  FAD3A' (SEQ ID NO:8)  (1740) ------------------------------------
 FAD3C' (SEQ ID NO:12)  (1899) TAAAATAAAATATGTAAAATATGTGATGCATGTATATATA
 FAD3A'' (SEQ ID NO:9)  (2004) TAAAA--------AACAA-TTTGCAAGTGTAATAGTTCAG
FAD3C'' (SEQ ID NO:11)  (1987) TAAGAGTTTTATGAACACGTTCTTAATTTTTTTTAGTTAAA
  FAD3C (SEQ ID NO:10)  (1781) -AAAA--------AACAAAAAACCGGTTACAAAAGTCGCG
                                2201                                2240
```

FIG. 1G (CONT.)

```
  FAD3A (SEQ ID NO:7)  (1458) ------------------------------------
  FAD3A' (SEQ ID NO:8)  (1740) ------------------------------------
 FAD3C' (SEQ ID NO:12)  (1939) AAATAATTATTATTTTTATCACTTAAAAT-AATTTATATT
 FAD3A'' (SEQ ID NO:9)  (2035) CATAATTGATCTTGTT-------------AGAGCAT-TT
FAD3C'' (SEQ ID NO:11)  (2027) AGTTAAGAAACGGGTTCTTATATTCCGCTAAGAACCTCTT
  FAD3C (SEQ ID NO:10)  (1812) AAAGAAGGATCGATTT-------------TGGTCTTTTA
                                2241                                2280
   FAD3A (SEQ ID NO:7)  (1458) ------------------------------------
  FAD3A' (SEQ ID NO:8)  (1740) ------------------------------------
 FAD3C' (SEQ ID NO:12)  (1978) AATATTTTTTAAAATTATCAAAGTTTACTGTTATTTAAAA
 FAD3A'' (SEQ ID NO:9)  (2060) CCAAAA-----CAA-------------------------
FAD3C'' (SEQ ID NO:11)  (2067) CCTAAAAACCCCAATAATCATACTC--CTAGGATTCTATA
  FAD3C (SEQ ID NO:10)  (1838) CTTGTA---------------------------------
                                2281                                2320
   FAD3A (SEQ ID NO:7)  (1458) ------------------------------------
  FAD3A' (SEQ ID NO:8)  (1740) ------------------------------------
```

FIG. 1H

```
 FAD3C' (SEQ ID NO:12)  (2018) TGTGATATGTAAATAATCTATATTATTTAAAATATTTCAA
 FAD3A'' (SEQ ID NO:9)  (2069) ------ACTTTATAATTTTAAATATACAGT-TT-------
FAD3C'' (SEQ ID NO:11)  (2105) TGTT-TATTTTATTAGTTTATGTTTTCAGTCTGAGGTCAG
  FAD3C (SEQ ID NO:10)  (1844) ------CTGTTTGTGGATCCCACTGGTGGT----------
                                2321                                2360
   FAD3A (SEQ ID NO:7)  (1458) ------------------------------------
  FAD3A' (SEQ ID NO:8)  (1740) ------------------------------------
 FAD3C' (SEQ ID NO:12)  (2058) TAATTTAAAAGCACCCAAAATTAGAGTAAAATATTTATAG
 FAD3A'' (SEQ ID NO:9)  (2095) ---------TT--------TGTTCTCT---------AAAA
FAD3C'' (SEQ ID NO:11)  (2144) ACCGGCCACTTGTCAGATCTGTTTTCTAGCTGTAGTAAAA
  FAD3C (SEQ ID NO:10)  (1868) ------------------GGTCCGCG---------ATTG
                                2361                                2400
   FAD3A (SEQ ID NO:7)  (1458) ------------------------------------
  FAD3A' (SEQ ID NO:8)  (1740) ------------------------------------
 FAD3C' (SEQ ID NO:12)  (2098) ATGTTTTTTTATTATGATTATCTTATT--TATTTAATATT
 FAD3A'' (SEQ ID NO:9)  (2109) AAGAATTT--------AAAAATT---------TTAAAGTT
FAD3C'' (SEQ ID NO:11)  (2184) AACAATTTGCAAGTGTAATAGTTCAGCGGTAATTAATGTT
  FAD3C (SEQ ID NO:10)  (1880) GTTTCTTT--------TTTAATT------TAATTTATTTT
                                2401                                2440
   FAD3A (SEQ ID NO:7)  (1458) ------------------------------------
  FAD3A' (SEQ ID NO:8)  (1740) ------------------------------------
 FAD3C' (SEQ ID NO:12)  (2136) ATAGATATTTTTTGTTCTTACAGTTTCTACAGCTTATAAA
 FAD3A'' (SEQ ID NO:9)  (2132) TGAGGGACGA----------------AACTTCAAATT
FAD3C'' (SEQ ID NO:11)  (2224) CTCGGATCTATCTCAAAAAAAAATTTTATAACTTCAAATA
  FAD3C (SEQ ID NO:10)  (1906) TTTAATCGGA----------------GAAAAAAATTA
                                2441                                2480
   FAD3A (SEQ ID NO:7)  (1458) ------------------------------------
  FAD3A' (SEQ ID NO:8)  (1740) ------------------------------------
 FAD3C' (SEQ ID NO:12)  (2176) TGAAAGATGTAAGTTGTTTAACTAAAATACATAAGAA---
 FAD3A'' (SEQ ID NO:9)  (2153) TGAAC----------TTTCACTACTCAACTTC-AAATTT
FAD3C'' (SEQ ID NO:11)  (2264) TAAAGATTTTTTTGTTTTTCAAAAATGAACTTCGAAACTT
  FAD3C (SEQ ID NO:10)  (1927) AGAAA---------C----CAAAAAACAGTTTT-----AA
                                2481                                2520
   FAD3A (SEQ ID NO:7)  (1458) ------------------------------------
  FAD3A' (SEQ ID NO:8)  (1740) ------------------------------------
 FAD3C' (SEQ ID NO:12)  (2213) -AAATGTTTGGTTTTTTTTTTGCTGTAGCTTTATTTTTAA
 FAD3A'' (SEQ ID NO:9)  (2181) GAAATTTCATCTTTTTTATTTACATTTTGATCATTATAAT
FAD3C'' (SEQ ID NO:11)  (2304) CAAATTTGAAGTTTTTTTTTTGCATTTTGATCATTATAAT
  FAD3C (SEQ ID NO:10)  (1949) TCATGGCCTCATGTTGGGGTTGAGTTTTATATTCTGATAA
                                2521                                2560
```

FIG. 1H (CONT.)

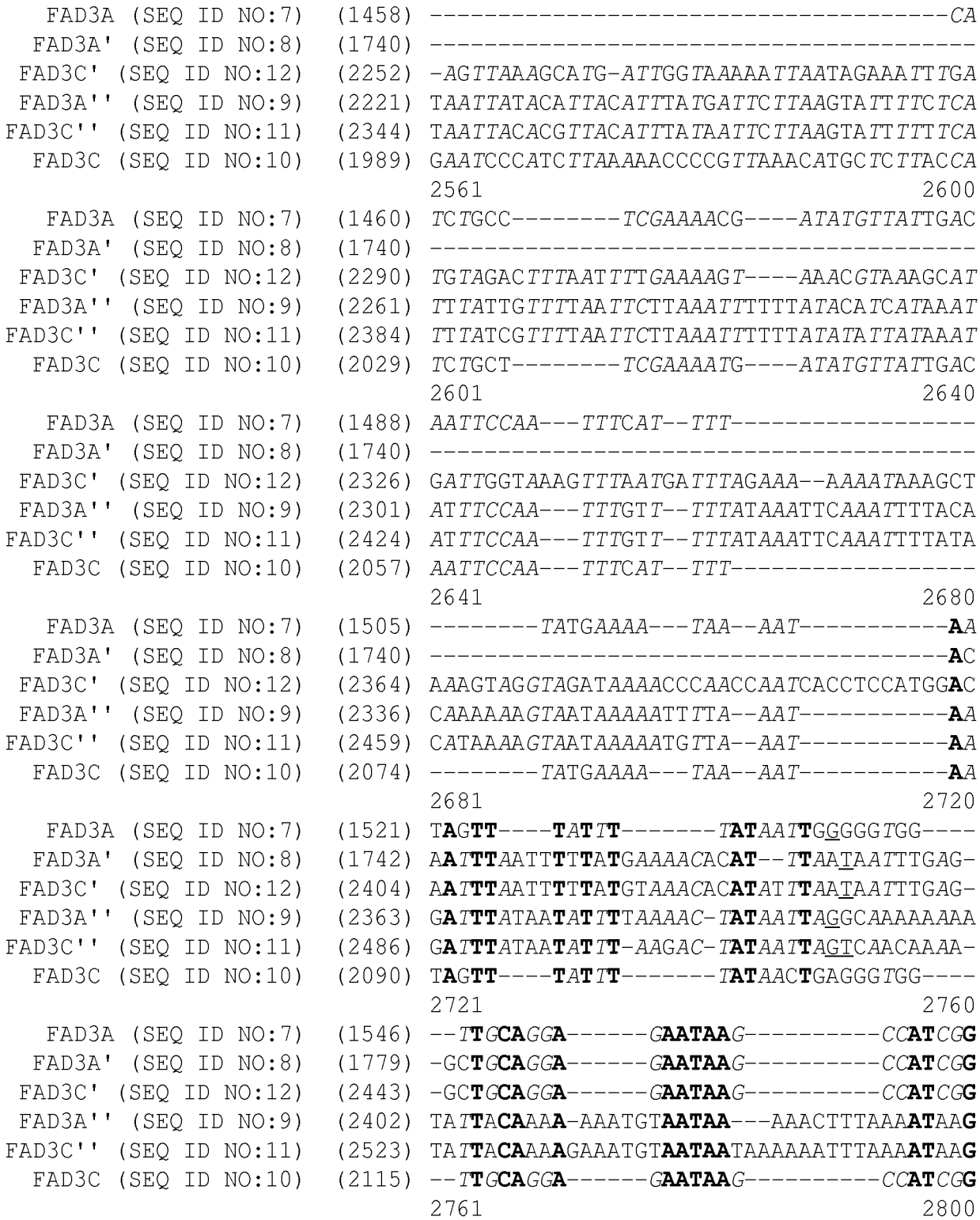

FAD3A (SEQ ID NO:7) (1458)
FAD3A' (SEQ ID NO:8) (1740)
FAD3C' (SEQ ID NO:12) (2252)
FAD3A'' (SEQ ID NO:9) (2221)
FAD3C'' (SEQ ID NO:11) (2344)
FAD3C (SEQ ID NO:10) (1989)
2561 2600
FAD3A (SEQ ID NO:7) (1460)
FAD3A' (SEQ ID NO:8) (1740)
FAD3C' (SEQ ID NO:12) (2290)
FAD3A'' (SEQ ID NO:9) (2261)
FAD3C'' (SEQ ID NO:11) (2384)
FAD3C (SEQ ID NO:10) (2029)
2601 2640
FAD3A (SEQ ID NO:7) (1488)
FAD3A' (SEQ ID NO:8) (1740)
FAD3C' (SEQ ID NO:12) (2326)
FAD3A'' (SEQ ID NO:9) (2301)
FAD3C'' (SEQ ID NO:11) (2424)
FAD3C (SEQ ID NO:10) (2057)
2641 2680
FAD3A (SEQ ID NO:7) (1505)
FAD3A' (SEQ ID NO:8) (1740)
FAD3C' (SEQ ID NO:12) (2364)
FAD3A'' (SEQ ID NO:9) (2336)
FAD3C'' (SEQ ID NO:11) (2459)
FAD3C (SEQ ID NO:10) (2074)
2681 2720
FAD3A (SEQ ID NO:7) (1521)
FAD3A' (SEQ ID NO:8) (1742)
FAD3C' (SEQ ID NO:12) (2404)
FAD3A'' (SEQ ID NO:9) (2363)
FAD3C'' (SEQ ID NO:11) (2486)
FAD3C (SEQ ID NO:10) (2090)
2721 2760
FAD3A (SEQ ID NO:7) (1546)
FAD3A' (SEQ ID NO:8) (1779)
FAD3C' (SEQ ID NO:12) (2443)
FAD3A'' (SEQ ID NO:9) (2402)
FAD3C'' (SEQ ID NO:11) (2523)
FAD3C (SEQ ID NO:10) (2115)
2761 2800

FIG. 1I

```
FAD3A (SEQ ID NO:7)  (1568) ACACACCAC--CAGAACCATGGCCATGTTGAAA----ACG
  FAD3A' (SEQ ID NO:8)  (1802) ACACACCAC--CAGAACCATGGCCATGTTGAAA----ACG
 FAD3C' (SEQ ID NO:12)  (2466) ACACACCAC--CAGAACCATGGCCATGTTGAAA----ACG
 FAD3A'' (SEQ ID NO:9)  (2438) ATATATCAAGACATAATTATTAGAAATTTAAATATTATA
FAD3C'' (SEQ ID NO:11)  (2563) ATACATGAAGACATAACTATTAGAAAATTTAAATATTATA
  FAD3C (SEQ ID NO:10)  (2137) ACACACCAC--CAGAACCATGGCCATGTTGAAA----ACG
                              2801                                  2840
   FAD3A (SEQ ID NO:7)  (1602) ACGAGTCTTGGGTTCCGGTAA------TC-----CCCCTC
  FAD3A' (SEQ ID NO:8)  (1836) ACGAGTCTTGGGTTCCGGTAACATT--TC-----CCTCTT
 FAD3C' (SEQ ID NO:12)  (2500) ACGAGTCTTGGGTTCCGGTAACATT--TC-----CCTCTT
 FAD3A'' (SEQ ID NO:9)  (2478) ACAATATTAATAATCTGGTAATTTGCTCCAAAACCTCAA
FAD3C'' (SEQ ID NO:11)  (2603) ACAATACTAATAATCTGGTAATTTGCTCTGGAACCTCTA
  FAD3C (SEQ ID NO:10)  (2171) ACGAGTCTTGGGTTCCGGTAA------TCTTTC-CCTCTC
                              2841                                  2880
```

FIG. 1I (CONT.)

```
   FAD3A (SEQ ID NO:7)  (1631) TCATT--------------------ATTTTTTTT-----
  FAD3A' (SEQ ID NO:8)  (1869) TAATA---------ATT------TCTATTTTTCT------
 FAD3C' (SEQ ID NO:12)  (2533) TAATA---------ATT------TCTATTTTTCTT--T--
 FAD3A'' (SEQ ID NO:9)  (2518) AATTTCTAAATTATTGTCCAAACAAATTT-GTTTAACCG
FAD3C'' (SEQ ID NO:11)  (2643) AATT--------ATTGTCTAAACAATTTTGTGTAACCG
  FAD3C (SEQ ID NO:10)  (2204) TCAT--------------------ATTTTTTTT-----
                              2881                                  2920
   FAD3A (SEQ ID NO:7)  (1645) --------------------TCTTTTTTTGAAAC------
  FAD3A' (SEQ ID NO:8)  (1888) ---------GTCAAAATAATTAGTTTTTCGAAATTTGAGG
 FAD3C' (SEQ ID NO:12)  (2554) ---------GTCAAAATAATTTGTTTTTCGAAATTTGAGG
 FAD3A'' (SEQ ID NO:9)  (2557) AATATGGAGCATTACAAAAATAATTTTTATGGAATAGTGTG
FAD3C'' (SEQ ID NO:11)  (2675) AAGATGGAGCATTACGAAAATAATTTTTATGAAATAATATG
  FAD3C (SEQ ID NO:10)  (2217) --------------------CTTTTTTTTGAAAT------
                              2921                                  2960
   FAD3A (SEQ ID NO:7)  (1659) ------------------T--CTTTCATTTTAATTTTCT-
  FAD3A' (SEQ ID NO:8)  (1919) CCAGAACGACCACTTGTCAA-ATTTGATT-TTTAGCTGTA
 FAD3C' (SEQ ID NO:12)  (2585) CCAGAACGACCACTTGTCAG-ATTTGATT-TCTAGCTGTA
 FAD3A'' (SEQ ID NO:9)  (2597) GTATTTTGCTTGTAGTT-AATATTTAATTATGTATTTCTA
FAD3C'' (SEQ ID NO:11)  (2715) GTATTTTGCTTCTAGTTTAATATTTAATTATATATTTCTA
  FAD3C (SEQ ID NO:10)  (2231) ------------------T--CTTTCATTTTAATTTTCT-
                              2961                                  3000
   FAD3A (SEQ ID NO:7)  (1678) --TAGAATTCTATGTATTTA-----------TTTTAATCA
  FAD3A' (SEQ ID NO:8)  (1957) GTAAAAACAGTTTGCTAGTGTCACAGTTAACCGGTAATTG
 FAD3C' (SEQ ID NO:12)  (2623) GTAAAAACAGTTTGCTAGTGTCACAGTTAACCGGTAATTG
 FAD3A'' (SEQ ID NO:9)  (2636) TTTATAATTTTATATATTTAATGTAAGATTTTTTTAATTA
```

FIG. 1J

```
FAD3C'' (SEQ ID NO:11)  (2755) TTTATAATTTTATATATTTAATGTAAATTTTTATTAATTA
  FAD3C (SEQ ID NO:10)  (2250) --TAGGATTCTATGTATTTA-----------TTTTAATCA
                                3001                                 3040
   FAD3A (SEQ ID NO:7)  (1705) ATCCT----------------------------------
  FAD3A' (SEQ ID NO:8)  (1997) ATTCTTTTTAACGATTTATAGAAGTAACATTTTTGTAAAA
 FAD3C' (SEQ ID NO:12)  (2663) ATTCTTTTTAGCGATTTATAGAAGTAACATTTTTGTAAAA
 FAD3A'' (SEQ ID NO:9)  (2676) ATATTACTGTAATATTTTTATATATGTACTAGTTATTTAT
FAD3C'' (SEQ ID NO:11)  (2795) ATATTACTGTAATATTTTTATATATGTGCTAGTTATTTAT
  FAD3C (SEQ ID NO:10)  (2277) ATCCT----------------------------------
                                3041                                 3080
   FAD3A (SEQ ID NO:7)  (1710) -----TTTT-------------------------------
  FAD3A' (SEQ ID NO:8)  (2037) TAAAATATACATTATGGTATGTGACAACGGACCACGCTTA
 FAD3C' (SEQ ID NO:12)  (2703) TAAAATATACATAATAGTATGTGACAACGGACCACGCCTA
 FAD3A'' (SEQ ID NO:9)  (2716) AAAAGTTTT-ATAGATTTGTATTAGTTATAACAAAAATAA
FAD3C'' (SEQ ID NO:11)  (2835) AATTTTTTTTATGGATTTATATTAG----ACCATGATTAA
  FAD3C (SEQ ID NO:10)  (2282) -----TTTT-------------------------------
                                3081                                 3120
   FAD3A (SEQ ID NO:7)  (1714) ----------------------------C-------CAGTG
  FAD3A' (SEQ ID NO:8)  (2077) TTTGTATTGGTGAATCTTTTAATTAC-TC--CCT-CCAAT
 FAD3C' (SEQ ID NO:12)  (2743) TTTGTATCGGTGAATCTTCTAATTAC-TT--CCT-CCGAT
 FAD3A'' (SEQ ID NO:9)  (2755) GGATCATTGTGTAAAATACAAATAATTTTGAAATTACGTT
FAD3C'' (SEQ ID NO:11)  (2871) CCCGGAGTTCTTAGAGTG-----------GAGTTTTAGTT
  FAD3C (SEQ ID NO:10)  (2286) ----------------------------C-------CAGTT
                                3121                                 3160
   FAD3A (SEQ ID NO:7)  (1720) TGAGGCTTG-------------------------------
  FAD3A' (SEQ ID NO:8)  (2113) TTATTTTAGTTGCAGATTTAGATTTATGCACATAGATTAA
 FAD3C' (SEQ ID NO:12)  (2779) TTATTTTAGTTACAGTTTTAGATTTATACACATAGATTAC
 FAD3A'' (SEQ ID NO:9)  (2795) TAAAGTTTTGGTTATGAAAAAAATACTTTGAAACTTTAAA
FAD3C'' (SEQ ID NO:11)  (2900) AAACGTT----------AAGAAACAGTTTCTTAACTTCCG
  FAD3C (SEQ ID NO:10)  (2292) TGAGGCTAG-------------------------------
```

FIG. 1J (CONT.)

```
                                  3161                                 3200
   FAD3A (SEQ ID NO:7)  (1729) ----------G---ACGACCACTTGTCAGATTTGTCG--
  FAD3A' (SEQ ID NO:8)  (2153) TAAAAATA-----TTTTGCACATTTTCAAAATAAAAACAC
 FAD3C' (SEQ ID NO:12)  (2819) AAAAAATAAAATATTTTGTCCATTTTTAAATAAAAACAT
 FAD3A'' (SEQ ID NO:9)  (2835) TTTAGAGTTTTGCAAACTTTAAAATGTTAGATAGATAGTT
FAD3C'' (SEQ ID NO:11)  (2930) GTAAGAACC---CCATCCTAAGAATCCCAGGTTAATC---
  FAD3C (SEQ ID NO:10)  (2301) ----------G---ACGACCACTTGTCAGATTTGTCG--
                                  3201                                 3240
   FAD3A (SEQ ID NO:7)  (1753) -------T-------TTAGCTGTAG---------------
  FAD3A' (SEQ ID NO:8)  (2188) CATTAC-TTATACAACTAACCATATTTCAACCAATAAAAA
 FAD3C' (SEQ ID NO:12)  (2859) CACTAA-TTATACACCTAACAATATTTTAACCAATAAAAA
 FAD3A'' (SEQ ID NO:9)  (2875) TTTTTGGAGATGCATTTAGTGGTTATGGTAGTAACTCAGA
FAD3C'' (SEQ ID NO:11)  (2964) ---------ATGCTCTTAGTTATAA----------CAAA
  FAD3C (SEQ ID NO:10)  (2325) -------T-------TTAGCTGTAG---------------
                                  3241                                 3280
   FAD3A (SEQ ID NO:7)  (1764) ------------------------------------------
  FAD3A' (SEQ ID NO:8)  (2227) --TAAATTAGAAAATATTATTTATAAATTTTGTATTGAAA
 FAD3C' (SEQ ID NO:12)  (2898) A-TAAACTAGAAAATATTATTCATAATTTTTACATTGAAA
 FAD3A'' (SEQ ID NO:9)  (2915) AAATGAAAAATCTATACTTTTATACTCCCTCCGTTTTTTA
FAD3C'' (SEQ ID NO:11)  (2984) TAAGGATCATTGTGTAA----------------------A
  FAD3C (SEQ ID NO:10)  (2336) ------------------------------------------
                                  3281                                 3320
   FAD3A (SEQ ID NO:7)  (1764) ------------------------------------------
  FAD3A' (SEQ ID NO:8)  (2265) TTATAAAATAATACTTATTTTAAAACGAAATT------AA
 FAD3C' (SEQ ID NO:12)  (2937) TTATAAAACGATACTTATTTTAAAACAAAATTTT----AA
 FAD3A'' (SEQ ID NO:9)  (2955) ATATAAGTCGTTTTACAGTTATACACGTAGATTAAGAAAA
FAD3C'' (SEQ ID NO:11)  (3002) ATACAAATAATTTTGAAGTTATGTTTGAAGTTTG------
  FAD3C (SEQ ID NO:10)  (2336) ------------------------------------------
                                  3321                                 3360
   FAD3A (SEQ ID NO:7)  (1764) ----------------------T----AAACAACTG---
  FAD3A' (SEQ ID NO:8)  (2299) TTTACAACGACAATTAAACTGAAACGGAAAGAAATTATTA
 FAD3C' (SEQ ID NO:12)  (2973) TTTACAACGACAATTAAATTGAAACGGAAGAAGTTTATTA
 FAD3A'' (SEQ ID NO:9)  (2995) CCATTAATTTCTTATATTTTCTAGACAAAAACATCATTAA
FAD3C'' (SEQ ID NO:11)  (3036) ----------------TTTTC--GAAGAAAACCACTTTGA
  FAD3C (SEQ ID NO:10)  (2336) ----------------------T----AAACAACTG---
                                  3361                                 3400
   FAD3A (SEQ ID NO:7)  (1774) --ATTTA--------------------------------
  FAD3A' (SEQ ID NO:8)  (2339) ATACTTAATTAAAGAGTTTTT-------AGAAAAATTGAA
 FAD3C' (SEQ ID NO:12)  (3013) TTACTTAATTAAAGAGTTTTTT-----TAAAAAAAATGAA
 FAD3A'' (SEQ ID NO:9)  (3035) TTATTTACCTAACCACAATTCAACCAATATAAAAATAGAA
FAD3C'' (SEQ ID NO:11)  (3058) AACTTTA----------------AATTTAGAGT---AA
```

FIG. 1K

```
FAD3C (SEQ ID NO:10)  (2346) --ATTTA----------------------------------
                              3401                                 3440
   FAD3A (SEQ ID NO:7)  (1779) ----------------------------AATTGTTTATGG
  FAD3A' (SEQ ID NO:8)  (2372) AGACATGTTTATGCGAAACTCATGTGAAAGTCTTGAAAT
 FAD3C' (SEQ ID NO:12)  (3048) AGACATGTTTATGCGAAACTCATGTGAAAGTCTTCAAAT
 FAD3A'' (SEQ ID NO:9)  (3075) GATATATTACCATTGGTCATACAACATTAATTATTAATAA
FAD3C'' (SEQ ID NO:11)  (3077) ACTCTATT-----------TAGAG----AGTTTTTTTTAG
  FAD3C (SEQ ID NO:10)  (2351) ----------------------------AATTGTTTATAG
                              3441                                 3480
   FAD3A (SEQ ID NO:7)  (1791) ----TACT---------------------------------
  FAD3A' (SEQ ID NO:8)  (2412) AATAGATTTTGGTATAAATATTTCAAATTTTCTT------
 FAD3C' (SEQ ID NO:12)  (3088) AAAATATTTTGGTATAAATTTTTCAAATTTTCA-------
 FAD3A'' (SEQ ID NO:9)  (3115) ATTTTACATAG-AAAACCGAAAACGACATATAATTTGGAA
FAD3C'' (SEQ ID NO:11)  (3102) AGGTTACGCAGTAACTCAGAAAATGA--------------
  FAD3C (SEQ ID NO:10)  (2363) ----TACT---------------------------------
                              3481                                 3520
```

FIG. 1K (CONT.)

```
   FAD3A (SEQ ID NO:7)  (1795) ----------------------------------------
  FAD3A' (SEQ ID NO:8)  (2446) -----AAAATAATAATTATATATTAATATAAT--------
 FAD3C' (SEQ ID NO:12)  (3121) -----AAAATAATAATTATAAATTAATATAATATAAT---
 FAD3A'' (SEQ ID NO:9)  (3154) CAAAAAAATTTCTCTAAAACGACTTATATTAAAAAACGGA
FAD3C'' (SEQ ID NO:11)  (3128) ----AAAATCTAT--------ACTTTTAT-----------
  FAD3C (SEQ ID NO:10)  (2367) ----------------------------------------
                              3521                                 3560
   FAD3A (SEQ ID NO:7)  (1795) ----G---TAGTTAACTTTAACAACGGGCCACTTATATTC
  FAD3A' (SEQ ID NO:8)  (2473) ----TTGTGATAAAATCTCGTCAAAAACTCACTAATGCAA
 FAD3C' (SEQ ID NO:12)  (3153) ----TTGTGATAAAATCTCGTCAAAAACTCACTAATGCAA
 FAD3A'' (SEQ ID NO:9)  (3194) GGGAGTAGTACCTAACTTTAACGATGGACCACTTATATTC
FAD3C'' (SEQ ID NO:11)  (3145) ------AGTACCTAACTTTATCGATGGACCACTTATATTC
  FAD3C (SEQ ID NO:10)  (2367) ----G---TAGTTAACTTTAACAACGGACCACTTATATTC
                              3561                                 3600
   FAD3A (SEQ ID NO:7)  (1828) GAGCCATTGG-CATAAAATGATT-CTTCTCGAAATTCGTT
  FAD3A' (SEQ ID NO:8)  (2509) ATGCTTTTAT-TTTGAATTTCTTACTCCTCTAAATGCATT
 FAD3C' (SEQ ID NO:12)  (3189) ATGCTTTTATATTTGAGTTTCTTACTCCTCTAAATGCATT
 FAD3A'' (SEQ ID NO:9)  (3234) GAGTCCTTAG-CATAAAATGATT-CTCCTCGAAATCCGTT
FAD3C'' (SEQ ID NO:11)  (3179) GAGTCCTTAG-CATAACATGATT-CTCCTCGAAATCCGTT
  FAD3C (SEQ ID NO:10)  (2400) GAGCCATTGG-CATAAAATGATT-CTTCTCGAAATTCGTT
                              3601                                 3640
   FAD3A (SEQ ID NO:7)  (1866) TACTTTTCT--TAGTATT-TTT-------CAGTTTTGTAG
  FAD3A' (SEQ ID NO:8)  (2548) TACTTTTATACTAATATTATTTTCTTTCTCTAATTTGGCG
```

FIG. 1L

```
FAD3C' (SEQ ID NO:12)  (3229) TACTTTTATACTATTATTTATTTTCTTTCTCTAATTTGGTG
 FAD3A'' (SEQ ID NO:9)  (3272) TACTTTCTT--CATTATT-TTTTCCTTTTCAGTTTTGGCG
FAD3C'' (SEQ ID NO:11)  (3217) TACTTTCTT--CGTTATT-TTTTCCTTTTCAGTTTTGGCG
  FAD3C (SEQ ID NO:10)  (2438) TACTTTTCT--TAGTATT-TTT-------CAATTTTGGAG
                                3641                                 3680
   FAD3A (SEQ ID NO:7)  (1896) TTTACGTAGAACTAAT------AA-----AAAG-------
  FAD3A' (SEQ ID NO:8)  (2588) TTT-CGTAATAGTTTG--TCTGTATTTTGAAAACTA----
 FAD3C' (SEQ ID NO:12)  (3269) TTTTCGTAATAGTTTG--CCTGTGTTTTGAAAACTA----
 FAD3A'' (SEQ ID NO:9)  (3309) TTTTCGTAATACTTTTGTCTTCAATCTTGAAAGCTATTAG
FAD3C'' (SEQ ID NO:11)  (3254) TTTTCGTAATACTTTTGTCTGCAATCTTGAAAGCTATTAG
  FAD3C (SEQ ID NO:10)  (2468) TTTACGTAGAACTAAT------AA-----AAAG-------
                                3681                                 3720
   FAD3A (SEQ ID NO:7)  (1918) -AAAAAAACTTATAAACACACC------------------
  FAD3A' (SEQ ID NO:8)  (2621) -ACAAAAAATAATAAAAACACAA---------AGCTTATAA
 FAD3C' (SEQ ID NO:12)  (3303) -ACAAAAAATAATAAAAACACAA---------AGTTTATAA
 FAD3A'' (SEQ ID NO:9)  (3349) TATAAAAACTTATAAACACATCACATGCAATGAATTAATA
FAD3C'' (SEQ ID NO:11)  (3294) TATAAAA-CTTATAAACACAT----------GAATTAATA
  FAD3C (SEQ ID NO:10)  (2490) -AAAA--ACTTATAAACACACC------------------
                                3721                                 3760
   FAD3A (SEQ ID NO:7)  (1939) ----------------------ACATGCAATGAATA---
  FAD3A' (SEQ ID NO:8)  (2651) ---ACACAT------------A-GCATGCAATGAATATG-
 FAD3C' (SEQ ID NO:12)  (3333) ---ACACAT------------A-GCATGCAATGAAT----
 FAD3A'' (SEQ ID NO:9)  (3389) CGAATACATAACCAGAATGACAAATTTTCAATGAATATTT
FAD3C'' (SEQ ID NO:11)  (3323) CGAATACATAACCAGAATGACAAATTTTCAATGAATATTT
  FAD3C (SEQ ID NO:10)  (2509) ----------------------ACATGCAATGAATA---
                                3761                                 3800
   FAD3A (SEQ ID NO:7)  (1953) AATTCGAATATATAA----CCATACTGTTAAA--------
  FAD3A' (SEQ ID NO:8)  (2674) TACGAATATATATACCAATACATA-TCTAAGTACTATTTT
 FAD3C' (SEQ ID NO:12)  (3353) -------ATATATATCAATACATA-TCTAAGTACTATTTT
 FAD3A'' (SEQ ID NO:9)  (3429) AATACCAGTAAGTACTACTCCGTAATAGTAATAGTAATAG
FAD3C'' (SEQ ID NO:11)  (3363) AATACTAGTAAGTACTACTCCGTAATAGTAAT-----TAG
  FAD3C (SEQ ID NO:10)  (2523) AATTCGAATATATAA----CCATACTGTTAAA--------
```

FIG. 1L (CONT.)

```
                                3801                                    3840
   FAD3A (SEQ ID NO:7)  (1981) ---TATTAAT--------------------T----AA---
  FAD3A' (SEQ ID NO:8)  (2713) TCCAAGTACT---T--------------AATCTTGATTAC
 FAD3C' (SEQ ID NO:12)  (3385) TGCAAGTACT---T--------------AATCTTGATTAC
 FAD3A'' (SEQ ID NO:9)  (3469) TCATATTAATTTTTTTTTGTCATCAAACAAACAGTAATAG
FAD3C'' (SEQ ID NO:11)  (3398) TAATAGTAAT---------------------AGTAATAG
  FAD3C (SEQ ID NO:10)  (2551) ---TATTAAT--------------------T----TA---
                                3841                                    3880
   FAD3A (SEQ ID NO:7)  (1991) -CATTTTAATCTTAATTTTGCATTCCAGTTGCCAGAAAAA
  FAD3A' (SEQ ID NO:8)  (2736) TAAAATTCATTTTAATTGTTCCTTTCAGTTACCAGAAAGG
 FAD3C' (SEQ ID NO:12)  (3408) TAAAATTCATTTTAATTGTTCCTTTCAGTTACCAGAAAAG
 FAD3A'' (SEQ ID NO:9)  (3509) TAATATTAATTATAATTATGTATTTCAGTTGCCAGAAAAG
FAD3C'' (SEQ ID NO:11)  (3416) TCATATTAATTATAATTATGTATTTCAGTTGCCAGAAAAG
  FAD3C (SEQ ID NO:10)  (2561) -CATTTTAATCTTAATTTTGCATTCCAGTTGCCAGAAAAA
                                3881                                    3920
   FAD3A (SEQ ID NO:7)  (2030) TTATACAAGAATTTGTCCCACAGTACACGGATGCTCAGAT
  FAD3A' (SEQ ID NO:8)  (2776) TTATACAAGAATTTACCCCACAGTACTCGGATGCTCAGAT
 FAD3C' (SEQ ID NO:12)  (3448) TTATACAAGATTTTACCCCACAGTACTCGGATGCTCAGAT
 FAD3A'' (SEQ ID NO:9)  (3549) TTGTACAAGAACTTGCCCCATAGTACTCGGATGCTCAGAT
FAD3C'' (SEQ ID NO:11)  (3456) TTGTACAAGAACTTGCCCCATAGTACTCGGATGCTCAGAT
  FAD3C (SEQ ID NO:10)  (2600) TTATACAAGAATTTGTCCCACAGTACACGGATGCTCAGAT
                                3921                                    3960
   FAD3A (SEQ ID NO:7)  (2070) ACACTGTCCCTCTCCCCATGCTCGCTTACCCTCTCTATCT
  FAD3A' (SEQ ID NO:8)  (2816) ACACTGTCCCTCTGCCCATGCTCGCTTACCCGATCTATCT
 FAD3C' (SEQ ID NO:12)  (3488) ACACTGTCCCTCTGCCCATGCTCGCTTACCCGATCTATCT
 FAD3A'' (SEQ ID NO:9)  (3589) ACACTGTTCCTCTGCCCATGCTCGCTTACCCGATCTATCT
FAD3C'' (SEQ ID NO:11)  (3496) ACACTGTCCCTCTGCCCATGCTCGCTTACCCGATCTATCT
  FAD3C (SEQ ID NO:10)  (2640) ACACTGTCCCTCTCCCCATGCTCGCTTACCCTCTCTATCT
                                3961                                    4000
   FAD3A (SEQ ID NO:7)  (2110) GGTAAATCCTAATTCCTCATTTTTCTTCCTGATTATAATT
  FAD3A' (SEQ ID NO:8)  (2856) GGTAT-------------TTTTAATTCCTAAAATTTACT
 FAD3C' (SEQ ID NO:12)  (3528) GGTAT-------------TTTTAATTCCTAAAACTTACC
 FAD3A'' (SEQ ID NO:9)  (3629) GGTAAAAAAAAA-TACAATTTCAATTTTTTCTTAAAATT
FAD3C'' (SEQ ID NO:11)  (3536) GGTAAAAAAA--TACAATTTCTATTTTTT-CTTAAAATT
  FAD3C (SEQ ID NO:10)  (2680) GGTAAATCCTAATTCCTAATTTTTCTTCCTGATTATAATT
                                4001                                    4040
   FAD3A (SEQ ID NO:7)  (2150) ACAATTTTGAATTTTTAGATTTTGAGTATTAA--CTAAAT
  FAD3A' (SEQ ID NO:8)  (2883) ACAAGT----CATTTTAGAC--TGTGTTTTAA--AACAAT
 FAD3C' (SEQ ID NO:12)  (3555) ACAATT----CATTTTAGAT--TGTGTTTTAA--AACAAT
 FAD3A'' (SEQ ID NO:9)  (3668) ACAAAT----GGTTTTATATTTTGAGTTTTAAGCCAATAT
FAD3C'' (SEQ ID NO:11)  (3573) ACAAAT----GATTTTATATTTTGAGTTTTAAGCCAATAT
```

FIG. 1M

```
FAD3C (SEQ ID NO:10)  (2720) ACAAITTTGAATTTTTAGATTTTGAGTATTAA--CTAAAT
                                    4041                                4080
   FAD3A (SEQ ID NO:7)  (2188) ATAAATTAAATTTGTTTGGGGATGA-CTACAGTGGTACAG
  FAD3A' (SEQ ID NO:8)  (2915) ATAA-TTATTTTTG-TTTGGTTTTA-CTGCAGTGGTACAG
 FAD3C' (SEQ ID NO:12)  (3587) ATAAATTATTTTTTCTTTGGTTTTA-CTGCAGTGGTACAG
 FAD3A'' (SEQ ID NO:9)  (3704) ATAAATTAATTTTGATTGGATTTTAACTACAGTGGTACAG
FAD3C'' (SEQ ID NO:11)  (3609) ATAAATTAATTTTGATTGGATTTTAACTACAGTGGTACAG
  FAD3C (SEQ ID NO:10)  (2758) ATAAATTAAATTTGTTTGGGGATGA-CTACAGTGGTACAG
                                    4081                                4120
   FAD3A (SEQ ID NO:7)  (2227) AAGTCCTGGTAAAGAAGGGTCACATTATAACCCATACAGT
  FAD3A' (SEQ ID NO:8)  (2952) AAGTCCTGGAAAAGAAGGGTCACATTTTAACCCATACAGT
 FAD3C' (SEQ ID NO:12)  (3626) AAGTCCTGGAAAAGAAGGGTCACATTTTAACCCATACAGT
 FAD3A'' (SEQ ID NO:9)  (3744) AAGTCCTGGAAAAGAAGGGTCACATTTTAACCCATACAGT
FAD3C'' (SEQ ID NO:11)  (3649) AAGTCCTGGAAAAGAAGGGTCACATTTTAACCCATACAGT
  FAD3C (SEQ ID NO:10)  (2797) AAGTCCTGGTAAAGAAGGGTCACATTATAACCCATACAGT
                                    4121                                4160
```

FIG. 1M (CONT.)

```
   FAD3A (SEQ ID NO:7)  (2267) AGTTTATTTGCCCCAAGCGAGAGAAGCTTATTGCAACTT
  FAD3A' (SEQ ID NO:8)  (2992) GGTTTATTTGCTCCAAGCGAGAGAAGCTTATTGCAACTT
 FAD3C' (SEQ ID NO:12)  (3666) GGTTTATTTGCTCCAAGCGAGAGAAGCTTATTGCAACTT
 FAD3A'' (SEQ ID NO:9)  (3784) AGTTTATTTGCTCCAAGCGAGAGGAAGCTTATTGCAACTT
FAD3C'' (SEQ ID NO:11)  (3689) AGTTTATTTGCTCCAAGCGAGAGGAAGCTTATTGCAACTT
  FAD3C (SEQ ID NO:10)  (2837) AGTTTATTTGCCCCAAGCGAGAGAAGCTTATTGCAACTT
                                4161                                 4200
   FAD3A (SEQ ID NO:7)  (2307) CAACTACTTGCTGGTCGATCATGTTGGCCACTCTTGTTTA
  FAD3A' (SEQ ID NO:8)  (3032) CGACTACTTGCTGGTCCATAATGTTGGCAATTCTTATCTG
 FAD3C' (SEQ ID NO:12)  (3706) CAACTACTTGCTGGTCCATAATGTTGGCCATTCTTATCTG
 FAD3A'' (SEQ ID NO:9)  (3824) CAACAACTTGCTGGTCCATAATGTTGGCCACTCTTGTTTA
FAD3C'' (SEQ ID NO:11)  (3729) CAACTACTTGCTGGTCCATAATGTTGGCCACTCTTGTTTA
  FAD3C (SEQ ID NO:10)  (2877) CAACTACTTGCTGGTCGATCGTGTTGGCCACTCTTGTTTA
                                4201                                 4240
   FAD3A (SEQ ID NO:7)  (2347) TCTATCATTCCTCGTTGGTCCAGTCACAGTTCTAAAAGTC
  FAD3A' (SEQ ID NO:8)  (3072) TCTTTCCTTCCTCGTTGGTCCAGTCACAGTTCTCAAAGTA
 FAD3C' (SEQ ID NO:12)  (3746) TCTTTCCTTCCTCGTTGGTCCAGTCACAGTTCTCAAAGTA
 FAD3A'' (SEQ ID NO:9)  (3864) TCTATCGTTCCTCGTTGGTCCAGTCACAGTTCTCAAAGTC
FAD3C'' (SEQ ID NO:11)  (3769) TCTATCGTTCCTCGTTGATCCAGTCACAGTTCTCAAAGTC
  FAD3C (SEQ ID NO:10)  (2917) TCTATCATTCCTCGTTGGTCCAGTCACAGTTCTAAAAGTC
                                4241                                 4280
   FAD3A (SEQ ID NO:7)  (2387) TATGGTGTTCCTTACATTGTAAGTTTCATA-TATTTC---
  FAD3A' (SEQ ID NO:8)  (3112) TACGGTGTTCCTTACATTGTAAGTTTCTTAGTATATCATA
 FAD3C' (SEQ ID NO:12)  (3786) TACGGTGTTCCTTACATCGTAAGTTTCTTAGTATATCATA
 FAD3A'' (SEQ ID NO:9)  (3904) TATGGTGTTCCTTACATTGTAAGTTTCACA-TATTATTAC
FAD3C'' (SEQ ID NO:11)  (3809) TATGGCGTTCCTTACATTGTAAGTTTCACA-TATTATTAC
  FAD3C (SEQ ID NO:10)  (2957) TATGGTGTTCCTTACATTGTAAGTTTCATA-TATTTC---
                                4281                                 4320
   FAD3A (SEQ ID NO:7)  (2423) ------ATTATTATATCATTGCTAATATA---------AT
  FAD3A' (SEQ ID NO:8)  (3152) AAGGGTATATATTTATTATTCAATATATATACTATATGAT
 FAD3C' (SEQ ID NO:12)  (3826) AAGGGTATATATTTATTATTCAATATATATACTATATGAT
 FAD3A'' (SEQ ID NO:9)  (3943) AAGAG-ATTTATATATTATTAATAATAAA---------TT
FAD3C'' (SEQ ID NO:11)  (3848) AAGAA-ATTTATATATTATTAATAATAAA---------TT
  FAD3C (SEQ ID NO:10)  (2993) ------TTTATTATATCATTGCTAATATA---------AT
                                4321                                 4360
   FAD3A (SEQ ID NO:7)  (2448) TTGTTTTTGACATAAA-GTTTTGGAAAAATTTCAGATCTT
  FAD3A' (SEQ ID NO:8)  (3192) TTGTTTTTGTCATATA-TTTTTG--AAATATTCAGATCTT
 FAD3C' (SEQ ID NO:12)  (3866) TTGTTTTTGTCATAAA-CTTTTG--AAAT--TCAGATCTT
 FAD3A'' (SEQ ID NO:9)  (3973) TGTTTTTTGACATAAA-GTTTTGGAAAATTTTCAGATCTT
FAD3C'' (SEQ ID NO:11)  (3878) TGTTTTTTGACATAAG-GGTTTGGAAAATTTTCAGATCTT
```

FIG. 1N

```
FAD3C (SEQ ID NO:10)  (3018) TTGTTTTTGACATAAAAGTTTGGAAAAATTTCAGATCTT
                                    4361                          4400
   FAD3A (SEQ ID NO:7)  (2487) TGTAATGTGGTTGGACGCTGTCACGTACTTGCATCATCAT
  FAD3A' (SEQ ID NO:8)  (3229) TGTGATGTGGTTGGACGCTGTCACTTACTTGCATCACCAT
 FAD3C' (SEQ ID NO:12)  (3901) TGTGATGTGGTTGGACGCTGTCACTTACTTGCATCACCAT
 FAD3A'' (SEQ ID NO:9)  (4012) TGTAATGTGGTTGGACGCTGTCACGTACTTGCATCATCAT
FAD3C'' (SEQ ID NO:11)  (3917) TGTGATGTGGTTGGACGCTGTCACGTACTTGCATCATCAT
  FAD3C (SEQ ID NO:10)  (3058) TGTAATGTGGTTGGACGCTGTCACGTACTTGCATCATCAT
                                    4401                          4440
   FAD3A (SEQ ID NO:7)  (2527) GGTCACGATGATAAGTTGCCTTGGTACAGAGGCAAGGTAA
  FAD3A' (SEQ ID NO:8)  (3269) GGTCATGATGAGAAGTTGCCTTGGTACAGAGGCAAGGTAA
 FAD3C' (SEQ ID NO:12)  (3941) GGTCATGATGAGAAGTTGCCTTGGTACAGAGGCAAGGTAA
 FAD3A'' (SEQ ID NO:9)  (4052) GGTCACGATGAGAAGTTGCCTTGGTACAGAGGCAAGGTAA
FAD3C'' (SEQ ID NO:11)  (3957) GGTCACGATGAGAAGTTGCCTTGGTACAGAGGCAAGGTAA
  FAD3C (SEQ ID NO:10)  (3098) GGTCACGATGATAAGCTGCCTTGGTACAGAGGCAAGGTAA
```

FIG. 1N (CONT.)

```
                                      4441                                 4480
   FAD3A (SEQ ID NO:7)  (2567) GTAGATCAACATT--------AATTTATAA---------G
  FAD3A' (SEQ ID NO:8)  (3309) TTAAATTAACTATTACAA--GTATTTTAC----------A
 FAD3C' (SEQ ID NO:12)  (3981) TTAAATTAACTCCTAGGT--GATTTTCCCGTGCTCATGTA
 FAD3A'' (SEQ ID NO:9)  (4092) ATAAATCAATTTTTAAAAAGAAATGTACAG---------A
FAD3C'' (SEQ ID NO:11)  (3997) TTAAATCAATTTTTAAAAAGAAATGTACAG---------A
  FAD3C (SEQ ID NO:10)  (3138) GTAGATCAACATT--------A-TTTATAA---------G
                                      4481                                 4520
   FAD3A (SEQ ID NO:7)  (2590) AAGCAACAATGATTAGTAT-TTGATTAATCTA-AATTATT
  FAD3A' (SEQ ID NO:8)  (3337) AAAAACTAATGATTAGTATATTTGATTAATCTTAATTCTT
 FAD3C' (SEQ ID NO:12)  (4019) CGGATATAAATATTTCTAAAGTAAATATACTATAATAATT
 FAD3A'' (SEQ ID NO:9)  (4123) AAGCAATAATGGTTAGTA--TTGATTAATCTT-AATTTTT
FAD3C'' (SEQ ID NO:11)  (4028) AAGCAATAATGGTTAGTA--TTGATTAATCTT-AATTTTT
  FAD3C (SEQ ID NO:10)  (3160) AAGCAATAATGATTAGTAG-TTGAATAATCTG-AATTTTT
                                      4521                                 4560
   FAD3A (SEQ ID NO:7)  (2628) GATGTTTTGTGTACAATAATAGGAATGGAGTTATTTACGT
  FAD3A' (SEQ ID NO:8)  (3377) GATGTTTTGTGATTAATAATAGGAATGGAGTTACTTACGT
 FAD3C' (SEQ ID NO:12)  (4059) AATTGTTATTTATTTTTAATTTTAAATTAGTTTATAATTT
 FAD3A'' (SEQ ID NO:9)  (4160) GATGTTTTGCATACAATAATAGGAATGGAGTTATTTACGT
FAD3C'' (SEQ ID NO:11)  (4065) GATGTTTTGCATACAATAATAGGAATGGAGTTATTTACGT
  FAD3C (SEQ ID NO:10)  (3198) GATGTTTT-TGTACAATAATAGGAATGGAGTTATTTACGT
                                      4561                                 4600
   FAD3A (SEQ ID NO:7)  (2668) GGAGGATTAACAACTATTGATAGAG-----ATTACGG-GA
  FAD3A' (SEQ ID NO:8)  (3417) GGAGGATTAACAACTATTGATAGAG-----ATTACGG-AA
 FAD3C' (SEQ ID NO:12)  (4099) GTATGCATGATTTATATTAATAAAATTTATATTACTTTAA
 FAD3A'' (SEQ ID NO:9)  (4200) GGAGGATTAACAACTATTGATAGAG-----ATTACGG-AA
FAD3C'' (SEQ ID NO:11)  (4105) GGAGGATTAACAACTATTGATAGAG-----ATTACGG-AA
  FAD3C (SEQ ID NO:10)  (3237) GGAGGATTAACAACTGTTGATAGAG-----ATTACGG-GA
                                      4601                                 4640
   FAD3A (SEQ ID NO:7)  (2702) TCTTCAACAACATTCATCACGATATTGGAACTCACGTGAT
  FAD3A' (SEQ ID NO:8)  (3451) TTTTCAACAACATTCATCACGACATTGGAACTCACGTGAT
 FAD3C' (SEQ ID NO:12)  (4139) TTATAAATATGATTT-TATATATGTTATATCTAATCGGTT
 FAD3A'' (SEQ ID NO:9)  (4234) TCTTCAACAACATCCATCACGACATTGGAACTCACGTGAT
FAD3C'' (SEQ ID NO:11)  (4139) TCTTCAACAACATCCATCACGACATTGGAACTCACGTGAT
  FAD3C (SEQ ID NO:10)  (3271) TCTTCAACAACATTCATCACGATATTGGAACTCACGTGAT
                                      4641                                 4680
   FAD3A (SEQ ID NO:7)  (2742) CCATCATCTTTTCCCACAAATCCCTCACTATCACTTGGTT
  FAD3A' (SEQ ID NO:8)  (3491) CCATCATCTTTTCCCACAAATCCCTCACTATCACTTGGTC
 FAD3C' (SEQ ID NO:12)  (4178) TTGTTGTTTTTACAGTCGATTTAGT---TATCATTGGGT
 FAD3A'' (SEQ ID NO:9)  (4274) CCATCATCTTTTCCCACAAATCCCTCACTATCACTTGGTC
FAD3C'' (SEQ ID NO:11)  (4179) CCATCATCTTTTCCCACAAATCCCTCACTATCACTTGGTC
```

FIG. 1O

```
  FAD3C (SEQ ID NO:10)  (3311) CCATCATCTTTTCCCACAAATCCCTCACTATCACTTGGTC
                               4681                                  4720
   FAD3A (SEQ ID NO:7)  (2782) GATGCCGTGAGTGATCTCGCT----CTCTCTC---TAGTT
  FAD3A' (SEQ ID NO:8)  (3531) GATGCTGTGAGTCATCTCACTCTCTGGCTAC------TTT
 FAD3C' (SEQ ID NO:12)  (4215) -AAATTGGATTGCATCTCAGAATTCAACTGTAATATTTTT
 FAD3A'' (SEQ ID NO:9)  (4314) GATGCGGTGAGTGATCTAGCTTTCTCTCTCTCTC---TAGTT
FAD3C'' (SEQ ID NO:11)  (4219) GATGCCGTGAGTGATCTAGCTTTCTCTCTCTC---TAGTT
  FAD3C (SEQ ID NO:10)  (3351) GATGCCGTGAGTGATCTCGCT----CTCTCTC---TAGTT
                               4721                                  4760
   FAD3A (SEQ ID NO:7)  (2815) TCATTTGATTAAAA--TTAAAGGGTGATTAATTACTAAAT
  FAD3A' (SEQ ID NO:8)  (3565) CATCAAAACCATTTGATTAAAGGGTGATTAATTACTAATG
 FAD3C' (SEQ ID NO:12)  (4254) TATTTTAACTATAT--TAAAATTTTGATTAATTTCTTATT
 FAD3A'' (SEQ ID NO:9)  (4351) TCATTTGATTAAA-------TG-GTGATTAATTACTAATT
FAD3C'' (SEQ ID NO:11)  (4256) TCATTTGATTAAA-------TG-GTGATTAATTACTAATT
  FAD3C (SEQ ID NO:10)  (3384) TCATTTGATTATA---TTAAAGGGTGATTAATTACTAAAT
                               4761                                  4800
```

FIG. 1O (CONT.)

```
   FAD3A (SEQ ID NO:7)  (2853) TAGTGATCTTAATTAATGATATGCG-ACAGACGAAATCAG
  FAD3A' (SEQ ID NO:8)  (3605) TAGTGATTTTA-ACAAATGGAATGTGACAGACAAAAGCAG
 FAD3C' (SEQ ID NO:12)  (4292) T--TCATTT-----AGGTGGTTGTTGTCTTAGAACTT---
 FAD3A'' (SEQ ID NO:9)  (4383) TA--------A-TTAATGAATTGTGGACAGACGAGAGCAG
FAD3C'' (SEQ ID NO:11)  (4288) TA--------A-TTAATGAATTGTGGACAGACGAGAGCAG
  FAD3C (SEQ ID NO:10)  (3421) TAGTGATCTTAATTAATGACATGCG-ACAGACGAAAGCAG
                               4801                                  4840
   FAD3A (SEQ ID NO:7)  (2892) CTAAACATGTGTTGGGAAGATACTACAGAGAACCAAAGAC
  FAD3A' (SEQ ID NO:8)  (3644) CTAAACATGTGTTGGGAAGATACTACAGAGAACCAAAGAC
 FAD3C' (SEQ ID NO:12)  (4322) -TAAATATATTTTATAAAGATTATGTATAACTTAATATAT
 FAD3A'' (SEQ ID NO:9)  (4414) CTAAACATGTGTTAGGAAGATACTACAGAGAGCCGAAGAC
FAD3C'' (SEQ ID NO:11)  (4319) CTAAACATGTGTTAGGAAGATACTACAGAGAGCCGAAGAC
  FAD3C (SEQ ID NO:10)  (3460) CTAAACATGTGTTGGGAAGATACTACAGAGAACCAAAGAC
                               4841                                  4880
   FAD3A (SEQ ID NO:7)  (2932) GTCAGGAGC----AAT--ACCGATCCACTTGGTGGAAAGT
  FAD3A' (SEQ ID NO:8)  (3684) GTCAGGAGC----AAT--ACCGATCCACTTGGTGGAGAGT
 FAD3C' (SEQ ID NO:12)  (4361) ATATTGTGCTTAAAATGAAATAAAAAATAAAATAAAGTGT
 FAD3A'' (SEQ ID NO:9)  (4454) GTCAGGAGC----AAT--ACCGATTCACTTGGTGGAGAGT
FAD3C'' (SEQ ID NO:11)  (4359) GTCAGGAGC----AAT--ACCGATTCACTTGGTGGAGAGT
  FAD3C (SEQ ID NO:10)  (3500) GTCAGGAGC----AAT--ACCGATCCACTTAGTGGAAAGT
                               4881                                  4920
   FAD3A (SEQ ID NO:7)  (2966) TTGGTGGCAAGTATTAAGAAAGATCATTACGTCAGTGACA
  FAD3A' (SEQ ID NO:8)  (3718) TTGGTAGCAAGTATTAAGAAAGATCATTACGTCAGTGACA
```

FIG. 1P

```
 FAD3C' (SEQ ID NO:12)  (4401) CTGATTCTAAATTACATAAATTAATATAACGATAAT-ATT
 FAD3A'' (SEQ ID NO:9)  (4488) TTGGTCGCAAGTATTAAAAAAGATCATTACGTCAGTGACA
FAD3C'' (SEQ ID NO:11)  (4393) TTGGTCGCAAGTATTAAAAAAGATCATTACGTCAGTGACA
  FAD3C (SEQ ID NO:10)  (3534) TTGGTGGCAAGTATTAAGAAAGATCATTACGTCAGTGACA
                               4921                                    4960
   FAD3A (SEQ ID NO:7)  (3006) CTG--GTGATATTGTCTTCTACG---AGACAGATCCAGAT
  FAD3A' (SEQ ID NO:8)  (3758) CTG--GTGACATTGTCTTCTACG---AGACTGATCCAGAT
 FAD3C' (SEQ ID NO:12)  (4440) CTGAAGTCTCATGCATATATATATATAAATTTTACAAAAG
 FAD3A'' (SEQ ID NO:9)  (4528) CTG--GTGATATTGTCTTCTACG---AGACAGATCCAGAT
FAD3C'' (SEQ ID NO:11)  (4433) CTG--GTGATATTGTCTTCTACG---AGACAGATCCAGAT
  FAD3C (SEQ ID NO:10)  (3574) CTG--GTGATATTGTCTTCTACG---AGACAGATCCAGAT
                               4961                                    5000
   FAD3A (SEQ ID NO:7)  (3041) CTCTACGTT-TATGCTTCTGACAA-ATCCAAAATCAACTA
  FAD3A' (SEQ ID NO:8)  (3793) CTCTACGTT-TATGCTTCTGTCAA-ATCGAAAATCAATTA
 FAD3C' (SEQ ID NO:12)  (4480) AACTAAATTGTAACATTTGGTTAATATTTTACAGTAATTA
 FAD3A'' (SEQ ID NO:9)  (4563) CTCTACGTT-TATGCTTCGGACAA-ATCTAAAATCAATTA
FAD3C'' (SEQ ID NO:11)  (4468) CTCTACGTT-TATGCTTCTGACAA-ATCTAAAATCAATTA
  FAD3C (SEQ ID NO:10)  (3609) CTCTACGTT-TATGCTTCTGACAA-ATCCAAAATCAATTA
                               5001                                    5040
   FAD3A (SEQ ID NO:7)  (3079) ACCTTTCTTCCTAGCTCTATTTAG----------GAATAA
  FAD3A' (SEQ ID NO:8)  (3831) AACTTTCTTCCCCCTTTTGTTTAGCACTATTATGAATAA
 FAD3C' (SEQ ID NO:12)  (4520) AAATATTTTATAAATTCTAAATA---ACT-TTATGTATTT
 FAD3A'' (SEQ ID NO:9)  (4601) ACTTTTCTTCCTAGCTCTATT-AG----------GAATAA
FAD3C'' (SEQ ID NO:11)  (4506) ACTTTTCTTCCTAGCTCTATT-AG----------GAATAA
  FAD3C (SEQ ID NO:10)  (3647) ATCTTTCTTCCTAGCTCTATTTAG----------GAATAA
                               5041                                    5080
   FAD3A (SEQ ID NO:7)  (3109) AACAGTCCTTTGGTTTTTACTTATTTCTGGTTGTTTTTAA
  FAD3A' (SEQ ID NO:8)  (3871) A--CCAGTTTTTTTT---ACTTATATATTGTTGTTTTTAA
 FAD3C' (SEQ ID NO:12)  (4556) A--ATTTATTGAATGGAAACTGAAATTTATTTTTAAATAAT
 FAD3A'' (SEQ ID NO:9)  (4630) A-CACTCCTTCTCTTTT-ACTTATTTGTTTCTGCTTT-AA
FAD3C'' (SEQ ID NO:11)  (4535) A-CACTCCTTCTCTTTT-ACTTATTTGTTTCTGCTTT-AA
  FAD3C (SEQ ID NO:10)  (3677) AACACTCCTTTGGTTTT-ACTTATTTCTGGTTGTTTTTAA
                               5081                                    5120
```

FIG. 1P (CONT.)

```
 FAD3A (SEQ ID NO:7)  (3149) GTTAAA--TGTACTCGTGAAACTTTTTTTA-ATTAAATGT
  FAD3A' (SEQ ID NO:8)  (3906) GTTAAAAATGTACTCGTGAAACTCTTCTTAATTTAGATAT
 FAD3C' (SEQ ID NO:12)  (4594) CTTAAAAATGAAAACATATTTGCTTTGGTATTTTGCTTAT
 FAD3A'' (SEQ ID NO:9)  (4667) GTTTAAAATGTACTCGTGAAACCTTTTT---TATTAATGT
FAD3C'' (SEQ ID NO:11)  (4572) GTTTAAAATGTACTCGTGAAACCTTTTTTT-TATTAATGT
  FAD3C (SEQ ID NO:10)  (3716) GTTAAAAATGTACTCGTGAAACTTTTTTT-ATTAAATGT
                               5121                                    5160
   FAD3A (SEQ ID NO:7)  (3186) ATTTACATT-------ACAAATC----AAGTTTTGTTCG
```

FIG. 1Q

```
  FAD3A' (SEQ ID NO:8)  (3946) TATTCCATT------TACA--CTGAAAAACATACAATTTC
 FAD3C' (SEQ ID NO:12)  (4634) GGTTCCATTAAGTTCTACAAACATAAAAACATAACATTTA
 FAD3A'' (SEQ ID NO:9)  (4704) ATTTACGTT-------ACAAAAAGTGGAAGTTTT-GTTAT
FAD3C'' (SEQ ID NO:11)  (4611) ATTTACGTT-------ACAAAAAGTGGAAGTTTT-GTTAT
  FAD3C (SEQ ID NO:10)  (3755) ATTTACATT-------ACAAATCGTAAAAGTTTTTGTTCG
                                5161                                 5200
   FAD3A (SEQ ID NO:7)  (3215) TTTTCTTTATGTTTTTAGTTACAA---TA---AATAAAG-
  FAD3A' (SEQ ID NO:8)  (3978) AAAGGT-TGAAAAGAAAGACAAAATTTTCT---AGAATGA
 FAD3C' (SEQ ID NO:12)  (4674) AAAACTGTGATTATTTTGTAACTATTTGATCAAACAATGA
 FAD3A'' (SEQ ID NO:9)  (4736) CTTTTTCTCTAGTTGCAATCAAAAGG--------------
FAD3C'' (SEQ ID NO:11)  (4643) CTTTTTCTCTGGTTGCAATCAAAAGG--------------
  FAD3C (SEQ ID NO:10)  (3788) TTTTCTCTATGTTTTTAGTTACAAACTTAC--AATCAAAA
                                5201                                 5240
   FAD3A (SEQ ID NO:7)  (3248) ------------------------------------------
  FAD3A' (SEQ ID NO:8)  (4014) C-----------------------------------------
 FAD3C' (SEQ ID NO:12)  (4714) TTATTTTTTAATTTTAATTTTAGTTTTTTAATAACTCTTA
 FAD3A'' (SEQ ID NO:9)  (4762) ------------------------------------------
FAD3C'' (SEQ ID NO:11)  (4669) ------------------------------------------
  FAD3C (SEQ ID NO:10)  (3826) AG----------------------------------------
                                5241                                 5280
   FAD3A (SEQ ID NO:7)  (3248) ------------------------------------------
  FAD3A' (SEQ ID NO:8)  (4015) ------------------------------------------
 FAD3C' (SEQ ID NO:12)  (4754) AAAATAAGCAGTGAACAAAAGTGAGATTGTATTTGAAATT
 FAD3A'' (SEQ ID NO:9)  (4762) ------------------------------------------
FAD3C'' (SEQ ID NO:11)  (4669) ------------------------------------------
  FAD3C (SEQ ID NO:10)  (3828) ------------------------------------------
                                5281                                 5320
   FAD3A (SEQ ID NO:7)  (3248) ------------------------------------------
  FAD3A' (SEQ ID NO:8)  (4015) ------------------------------------------
 FAD3C' (SEQ ID NO:12)  (4794) AATATTATACAAGTAAAATATAATTTTTTAAGTTTATAAA
 FAD3A'' (SEQ ID NO:9)  (4762) ------------------------------------------
FAD3C'' (SEQ ID NO:11)  (4669) ------------------------------------------
  FAD3C (SEQ ID NO:10)  (3828) ------------------------------------------
                                5321                                 5360
   FAD3A (SEQ ID NO:7)  (3248) ------------------------------------------
  FAD3A' (SEQ ID NO:8)  (4015) ------------------------------------------
 FAD3C' (SEQ ID NO:12)  (4834) AAAATTCCTTTTTATTATATGTATATGTTTTTTTGGAAAA
 FAD3A'' (SEQ ID NO:9)  (4762) ------------------------------------------
FAD3C'' (SEQ ID NO:11)  (4669) ------------------------------------------
  FAD3C (SEQ ID NO:10)  (3828) ------------------------------------------
                                5361                                 5400
```

FIG. 1Q (CONT.)

```
FAD3A (SEQ ID NO:7)  (3248) ----------------------------------------
  FAD3A' (SEQ ID NO:8)  (4015) ----------------------------------------
 FAD3C' (SEQ ID NO:12)  (4874) TTTTAAAAAGGAAACTAAATAAAAAAATAAATAATAGTAT
  FAD3A'' (SEQ ID NO:9)  (4762) ----------------------------------------
FAD3C'' (SEQ ID NO:11)  (4669) ----------------------------------------
  FAD3C (SEQ ID NO:10)  (3828) ----------------------------------------
                               5401                                 5440
```

FIG. 1Q (CONT.)

```
   FAD3A (SEQ ID NO:7)  (3248) ----------------------------------------
  FAD3A' (SEQ ID NO:8)  (4015) ----------------------------------------
 FAD3C' (SEQ ID NO:12)  (4914) TTTAAATGTAATATTTTTAATTCATTAAGTGTATTAGTGT
  FAD3A'' (SEQ ID NO:9)  (4762) ----------------------------------------
FAD3C'' (SEQ ID NO:11)  (4669) ----------------------------------------
  FAD3C (SEQ ID NO:10)  (3828) ----------------------------------------
                               5441                                 5480
   FAD3A (SEQ ID NO:7)  (3248) ----------------------------------------
  FAD3A' (SEQ ID NO:8)  (4015) ----------------------------------------
 FAD3C' (SEQ ID NO:12)  (4954) AATCAACTATCGTGAGAGTTAACGTGAGAGCGATACATAG
  FAD3A'' (SEQ ID NO:9)  (4762) ----------------------------------------
FAD3C'' (SEQ ID NO:11)  (4669) ----------------------------------------
  FAD3C (SEQ ID NO:10)  (3828) ----------------------------------------
                               5481                                 5520
   FAD3A (SEQ ID NO:7)  (3248) ----------------------------------------
  FAD3A' (SEQ ID NO:8)  (4015) ----------------------------------------
 FAD3C' (SEQ ID NO:12)  (4994) AAAACCGACTTCTCAAATAATATTTTATAGAGATTACGAT
  FAD3A'' (SEQ ID NO:9)  (4762) ----------------------------------------
FAD3C'' (SEQ ID NO:11)  (4669) ----------------------------------------
  FAD3C (SEQ ID NO:10)  (3828) ----------------------------------------
                               5521                                 5560
   FAD3A (SEQ ID NO:7)  (3248) ----------------------------------------
  FAD3A' (SEQ ID NO:8)  (4015) ----------------------------------------
 FAD3C' (SEQ ID NO:12)  (5034) GTTTCACAAAAAAAAATTATTAGTATTTGATTAATCTTAA
  FAD3A'' (SEQ ID NO:9)  (4762) ----------------------------------------
FAD3C'' (SEQ ID NO:11)  (4669) ----------------------------------------
  FAD3C (SEQ ID NO:10)  (3828) ----------------------------------------
                               5561                                 5600
   FAD3A (SEQ ID NO:7)  (3248) ----------------------------------------
  FAD3A' (SEQ ID NO:8)  (4015) ----------------------------------------
 FAD3C' (SEQ ID NO:12)  (5074) TTCTTGATGTTTTGTGATTAATAATAGGAATGGAGTTACT
  FAD3A'' (SEQ ID NO:9)  (4762) ----------------------------------------
```

FIG. 1R

```
FAD3C'' (SEQ ID NO:11)  (4669) ----------------------------------------
  FAD3C (SEQ ID NO:10)  (3828) ----------------------------------------
                                5601                                5640
   FAD3A (SEQ ID NO:7)  (3248) ----------------------------------------
  FAD3A' (SEQ ID NO:8)  (4015) ----------------------------------------
 FAD3C' (SEQ ID NO:12)  (5114) TACGTGGAGGATTAACAACTATTGATAGAGATTACGGAAT
 FAD3A'' (SEQ ID NO:9)  (4762) ----------------------------------------
FAD3C'' (SEQ ID NO:11)  (4669) ----------------------------------------
  FAD3C (SEQ ID NO:10)  (3828) ----------------------------------------
                                5641                                5680
   FAD3A (SEQ ID NO:7)  (3248) ----------------------------------------
  FAD3A' (SEQ ID NO:8)  (4015) ----------------------------------------
 FAD3C' (SEQ ID NO:12)  (5154) TTTCAACAACATTCATCACGACATTGGAACTCACGTGATC
 FAD3A'' (SEQ ID NO:9)  (4762) ----------------------------------------
FAD3C'' (SEQ ID NO:11)  (4669) ----------------------------------------
  FAD3C (SEQ ID NO:10)  (3828) ----------------------------------------
                                5681                                5720
   FAD3A (SEQ ID NO:7)  (3248) ----------------------------------------
  FAD3A' (SEQ ID NO:8)  (4015) ----------------------------------------
 FAD3C' (SEQ ID NO:12)  (5194) CATCATCTTTTCCCACAAATCCCTCACTATCACTTGGTCG
 FAD3A'' (SEQ ID NO:9)  (4762) ----------------------------------------
FAD3C'' (SEQ ID NO:11)  (4669) ----------------------------------------
  FAD3C (SEQ ID NO:10)  (3828) ----------------------------------------
                               5721                                5760
```

FIG. 1R (CONT.)

```
 FAD3A (SEQ ID NO:7)  (3248) ----------------------------------------
  FAD3A' (SEQ ID NO:8)  (4015) ----------------------------------------
 FAD3C' (SEQ ID NO:12)  (5234) ATGCTGTGAGTCATCTCACTCTCTCGCTACTTTCATCTAA
 FAD3A'' (SEQ ID NO:9)  (4762) ----------------------------------------
FAD3C'' (SEQ ID NO:11)  (4669) ----------------------------------------
  FAD3C (SEQ ID NO:10)  (3828) ----------------------------------------
                                5761                                5800
   FAD3A (SEQ ID NO:7)  (3248) ----------------------------------------
  FAD3A' (SEQ ID NO:8)  (4015) ----------------------------------------
 FAD3C' (SEQ ID NO:12)  (5274) ACCATTTCATTAAAGGGTGATTAATTACTAATGTACTGAT
 FAD3A'' (SEQ ID NO:9)  (4762) ----------------------------------------
FAD3C'' (SEQ ID NO:11)  (4669) ----------------------------------------
  FAD3C (SEQ ID NO:10)  (3828) ----------------------------------------
                                5801                                5840
   FAD3A (SEQ ID NO:7)  (3248) ----------------------------------------
  FAD3A' (SEQ ID NO:8)  (4015) ----------------------------------------
 FAD3C' (SEQ ID NO:12)  (5314) TTTAACAAATGGAATGTGACAGACAAAAGCAGCTAAACAT
```

FIG. 1S

```
FAD3A'' (SEQ ID NO:9)  (4762) ----------------------------------------
FAD3C'' (SEQ ID NO:11)  (4669) ----------------------------------------
  FAD3C (SEQ ID NO:10)  (3828) ----------------------------------------
                                5841                                5880
   FAD3A (SEQ ID NO:7)  (3248) ----------------------------------------
  FAD3A' (SEQ ID NO:8)  (4015) ----------------------------------------
 FAD3C' (SEQ ID NO:12)  (5354) GCGTTGGGAAGATACTACAGAGAACCGAAGACGTCAGGAG
 FAD3A'' (SEQ ID NO:9)  (4762) ----------------------------------------
FAD3C'' (SEQ ID NO:11)  (4669) ----------------------------------------
  FAD3C (SEQ ID NO:10)  (3828) ----------------------------------------
                                5881                                5920
   FAD3A (SEQ ID NO:7)  (3248) ----------------------------------------
  FAD3A' (SEQ ID NO:8)  (4015) ----------------------------------------
 FAD3C' (SEQ ID NO:12)  (5394) CAATACCGATCCACTTGGTGGAGAGTTTGGTAGCAAGTAT
 FAD3A'' (SEQ ID NO:9)  (4762) ----------------------------------------
FAD3C'' (SEQ ID NO:11)  (4669) ----------------------------------------
  FAD3C (SEQ ID NO:10)  (3828) ----------------------------------------
                                5921                                5960
   FAD3A (SEQ ID NO:7)  (3248) ----------------------------------------
  FAD3A' (SEQ ID NO:8)  (4015) ----------------------------------------
 FAD3C' (SEQ ID NO:12)  (5434) TAAGAAAGATCATTACGTCAGTGACACCGGTGACATTGTC
 FAD3A'' (SEQ ID NO:9)  (4762) ----------------------------------------
FAD3C'' (SEQ ID NO:11)  (4669) ----------------------------------------
  FAD3C (SEQ ID NO:10)  (3828) ----------------------------------------
                                5961                                6000
   FAD3A (SEQ ID NO:7)  (3248) ----------------------------------------
  FAD3A' (SEQ ID NO:8)  (4015) ----------------------------------------
 FAD3C' (SEQ ID NO:12)  (5474) TTCTACGAGACTGATCCAGATCTCTACGTTTATGCTTCTG
 FAD3A'' (SEQ ID NO:9)  (4762) ----------------------------------------
FAD3C'' (SEQ ID NO:11)  (4669) ----------------------------------------
  FAD3C (SEQ ID NO:10)  (3828) ----------------------------------------
                                6001                                6040
   FAD3A (SEQ ID NO:7)  (3248) ----------------------------------------
  FAD3A' (SEQ ID NO:8)  (4015) ----------------------------------------
 FAD3C' (SEQ ID NO:12)  (5514) TCAAATCGAAAATCAATTAAACTTTCTTCCCCCTTTTTGT
 FAD3A'' (SEQ ID NO:9)  (4762) ----------------------------------------
FAD3C'' (SEQ ID NO:11)  (4669) ----------------------------------------
  FAD3C (SEQ ID NO:10)  (3828) ----------------------------------------
                                6041                                6080
```

FIG. 1S (CONT.)

```
   FAD3A (SEQ ID NO:7)  (3248) ----------------------------------------
  FAD3A' (SEQ ID NO:8)  (4015) ----------------------------------------
 FAD3C' (SEQ ID NO:12)  (5554) TTAGCCCTATTATGAATAAACCAGTCTTTTTTCACTTATT
 FAD3A'' (SEQ ID NO:9)  (4762) ----------------------------------------
FAD3C'' (SEQ ID NO:11)  (4669) ----------------------------------------
  FAD3C (SEQ ID NO:10)  (3828) ----------------------------------------
                                6081                                6120
   FAD3A (SEQ ID NO:7)  (3248) ----------------------------------------
  FAD3A' (SEQ ID NO:8)  (4015) ----------------------------------------
 FAD3C' (SEQ ID NO:12)  (5594) TATTGGTGTTTTTAAGTTAAAAATGTACTCGTGAAACTCT
 FAD3A'' (SEQ ID NO:9)  (4762) ----------------------------------------
FAD3C'' (SEQ ID NO:11)  (4669) ----------------------------------------
  FAD3C (SEQ ID NO:10)  (3828) ----------------------------------------
                                6121                                6160
   FAD3A (SEQ ID NO:7)  (3248) ----------------------------------------
  FAD3A' (SEQ ID NO:8)  (4015) ----------------------------------------
 FAD3C' (SEQ ID NO:12)  (5634) TCTTTTATTATTAATCCATTTATACACTGAAAAACATACA
 FAD3A'' (SEQ ID NO:9)  (4762) ----------------------------------------
FAD3C'' (SEQ ID NO:11)  (4669) ----------------------------------------
  FAD3C (SEQ ID NO:10)  (3828) ----------------------------------------
                                6161                                6200
   FAD3A (SEQ ID NO:7)  (3248) ----------------------------------------
  FAD3A' (SEQ ID NO:8)  (4015) ----------------------------------------
 FAD3C' (SEQ ID NO:12)  (5674) ATTTCAAAGGTTAAAAAGAAAAATAAATTTTCTAGACTGA
 FAD3A'' (SEQ ID NO:9)  (4762) ----------------------------------------
FAD3C'' (SEQ ID NO:11)  (4669) ----------------------------------------
  FAD3C (SEQ ID NO:10)  (3828) ----------------------------------------
                                6201
   FAD3A (SEQ ID NO:7)  (3248) -
  FAD3A' (SEQ ID NO:8)  (4015) -
 FAD3C' (SEQ ID NO:12)  (5714) C
 FAD3A'' (SEQ ID NO:9)  (4762) -
FAD3C'' (SEQ ID NO:11)  (4669) -
  FAD3C (SEQ ID NO:10)  (3828) -
```

FIG. 1T (CONT.)

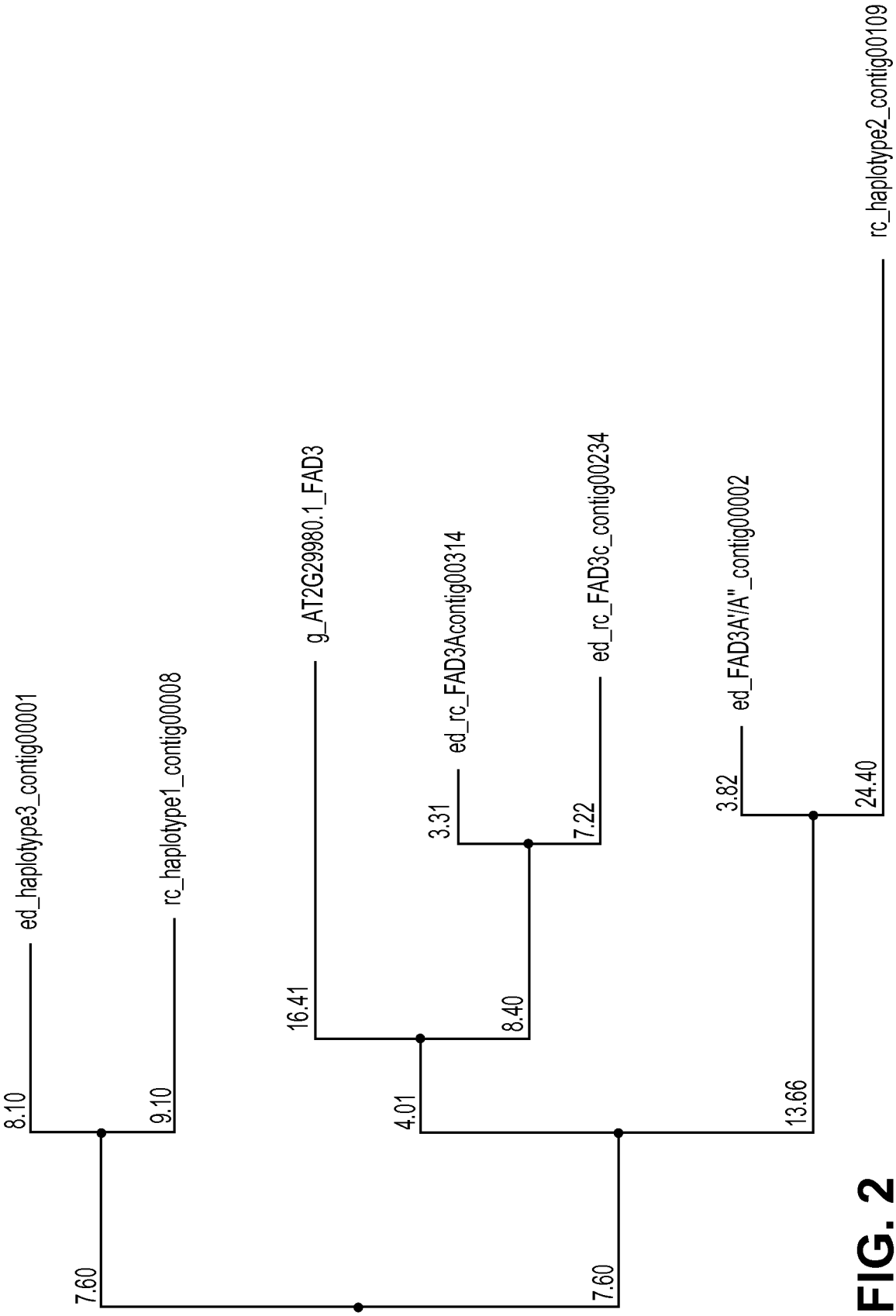

ed_haplotype3_contig00001
rc_haplotype1_contig00008
g_AT2G29980.1_FAD3
ed_rc_FAD3Acontig00314
ed_rc_FAD3c_contig00234
ed_FAD3A'/A"_contig00002
rc_haplotype2_contig00109
8.10
9.10
16.41
3.31
8.40
7.22
3.82
24.40
4.01
13.66
7.60
7.60

FIG. 2

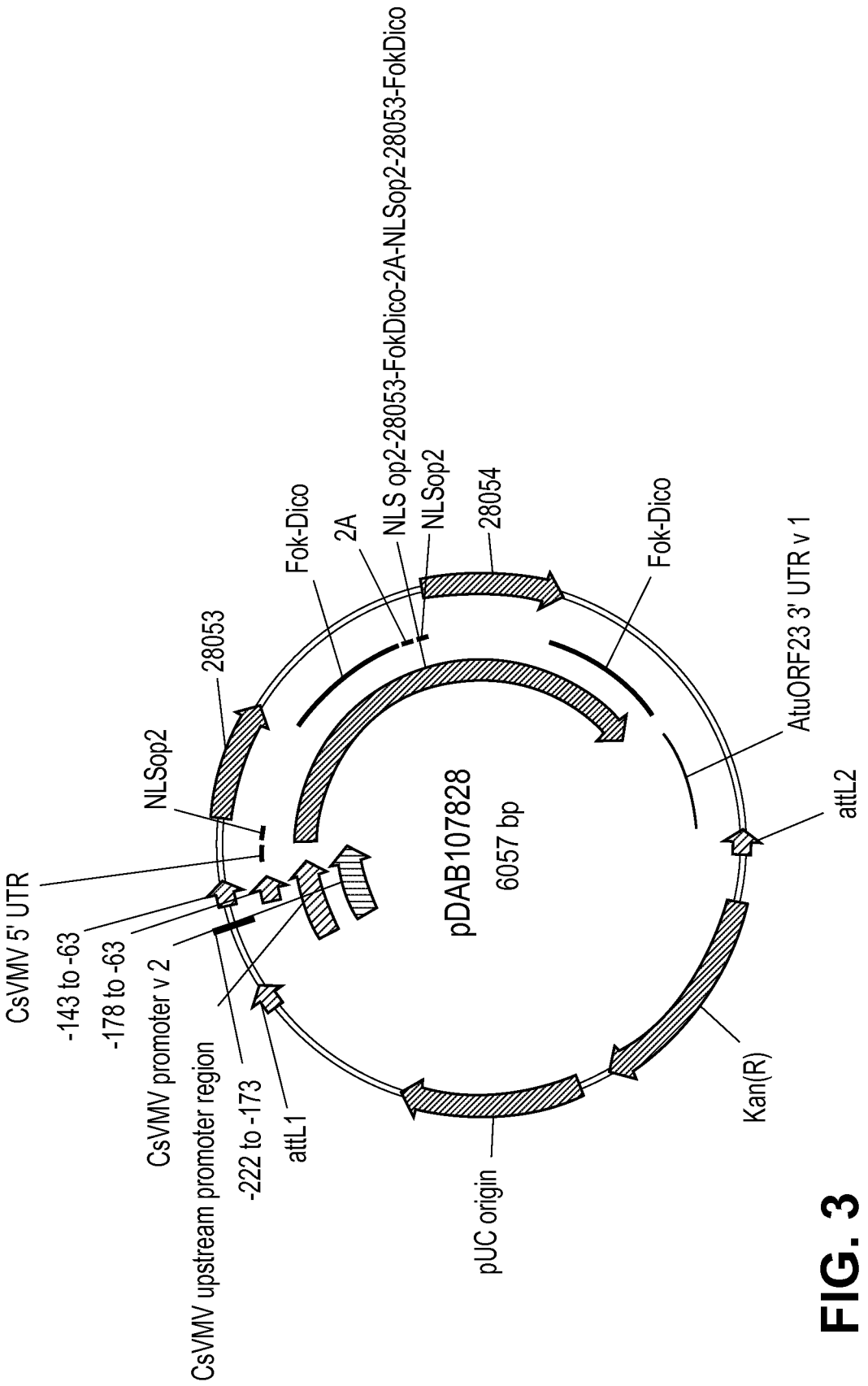

pDAB107828
6057 bp
CsVMV 5' UTR
-143 to -63
-178 to -63
CsVMV promoter v 2
CsVMV upstream promoter region
-222 to -173
attL1
NLSop2
28053
Fok-Dico
2A
NLS op2-28053-FokDico-2A-NLSop2-28053-FokDico
NLSop2
28054
Fok-Dico
AtuORF23 3' UTR v 1
attL2
Kan(R)
pUC origin

FIG. 3

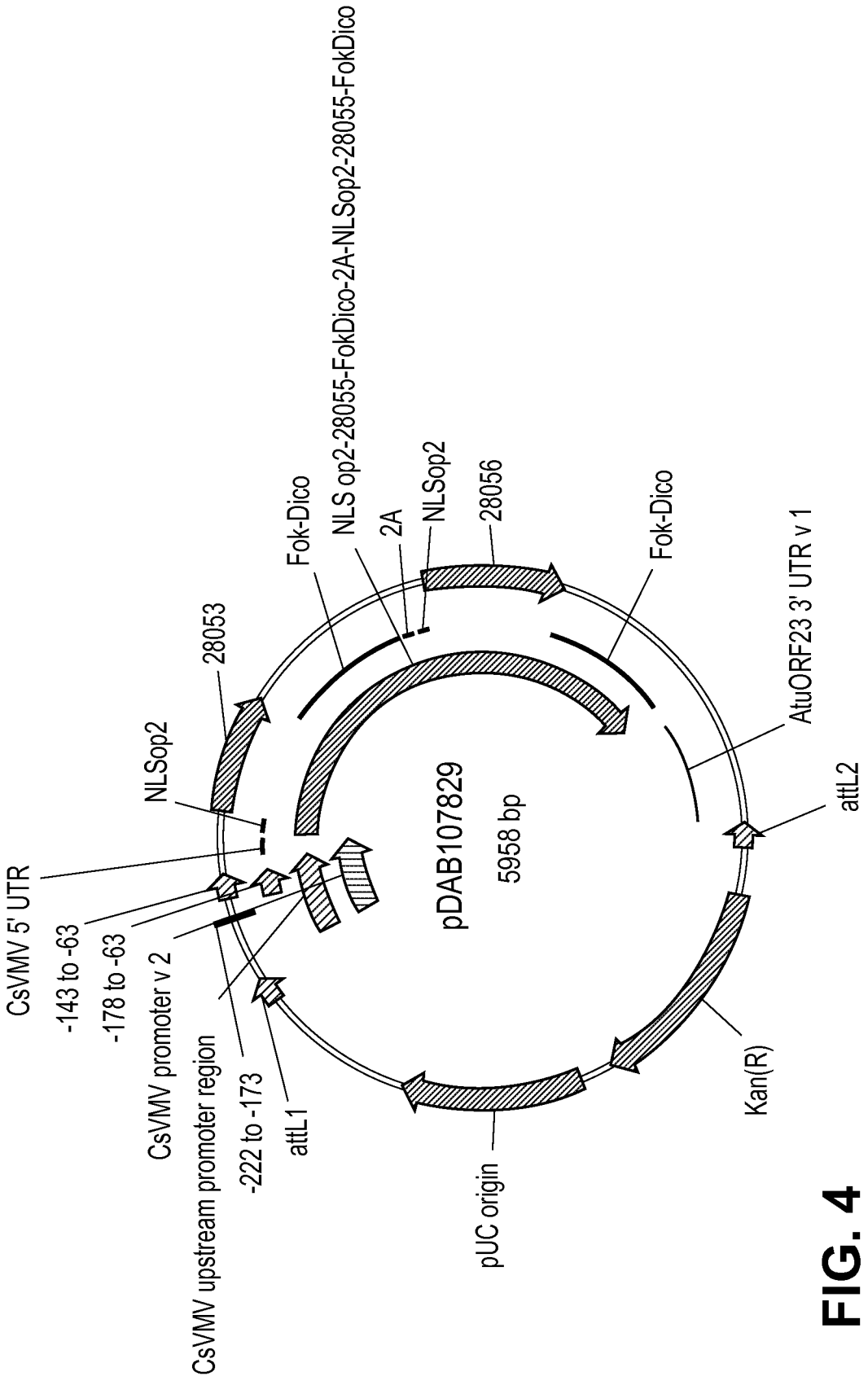

pDAB107829
5958 bp
CsVMV 5' UTR
-143 to -63
-178 to -63
CsVMV promoter v 2
CsVMV upstream promoter region
-222 to -173
attL1
NLSop2
28053
Fok-Dico
NLS op2-28055-FokDico-2A-NLSop2-28055-FokDico
2A
NLSop2
28056
Fok-Dico
AtuORF23 3' UTR v 1
attL2
Kan(R)
pUC origin

FIG. 4

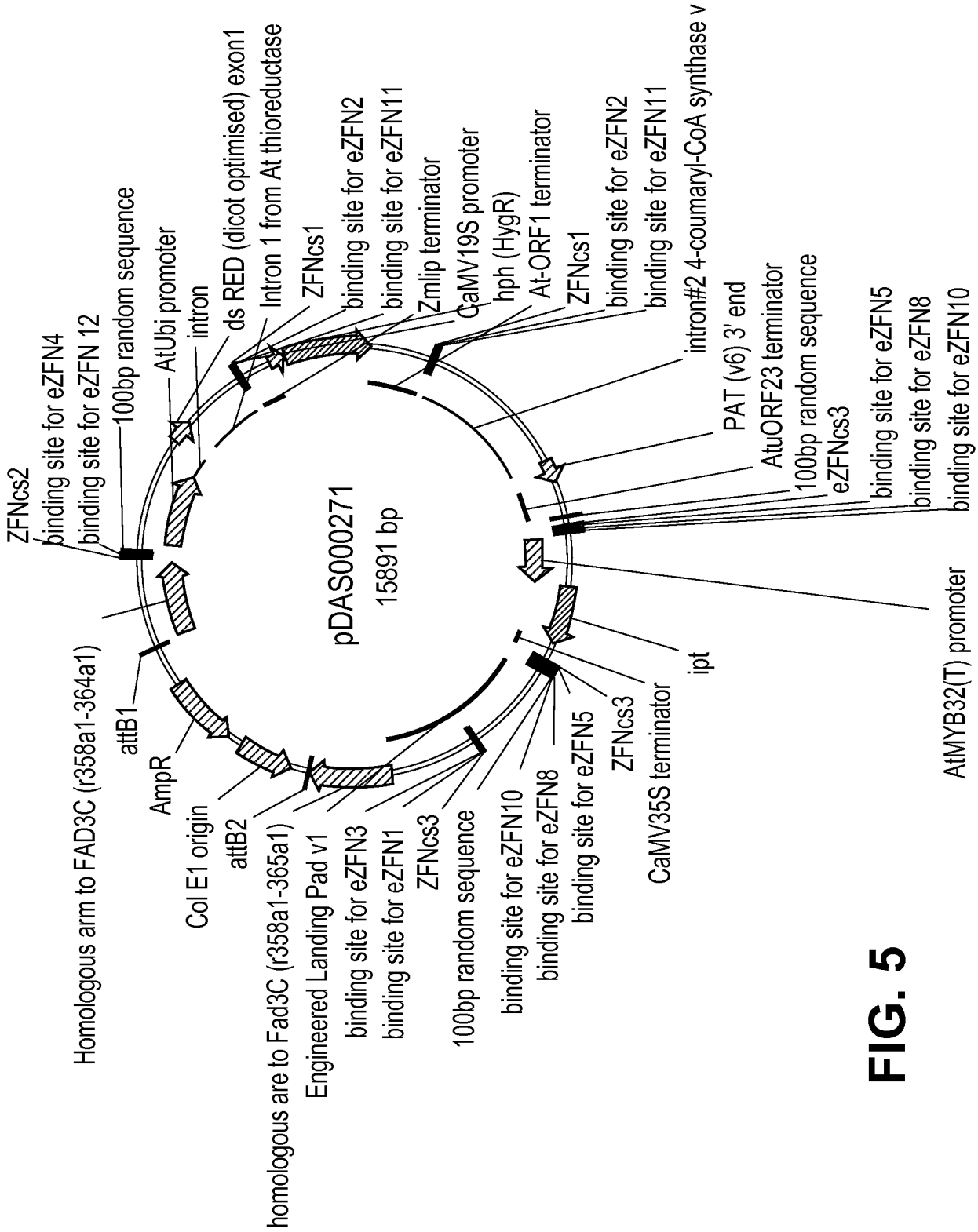

pDAS000271
15891 bp
ZFNcs2
binding site for eZFN4
binding site for eZFN 12
100bp random sequence
AtUbi promoter
intron
ds RED (dicot optimised) exon1
Intron 1 from At thioreductase
ZFNcs1
binding site for eZFN2
binding site for eZFN11
Zmlip terminator
CaMV19S promoter
hph (HygR)
At-ORF1 terminator
ZFNcs1
binding site for eZFN2
binding site for eZFN11
intron#2 4-coumaryl-CoA synthase v
PAT (v6) 3' end
AtuORF23 terminator
100bp random sequence
eZFNcs3
binding site for eZFN5
binding site for eZFN8
binding site for eZFN10
AtMYB32(T) promoter
ipt
CaMV35S terminator
ZFNcs3
binding site for eZFN5
binding site for eZFN8
binding site for eZFN10
100bp random sequence
ZFNcs3
binding site for eZFN1
binding site for eZFN3
Engineered Landing Pad v1
homologous are to Fad3C (r358a1-365a1)
attB2
Col E1 origin
AmpR
attB1
Homologous arm to FAD3C (r358a1-364a1)

FIG. 5

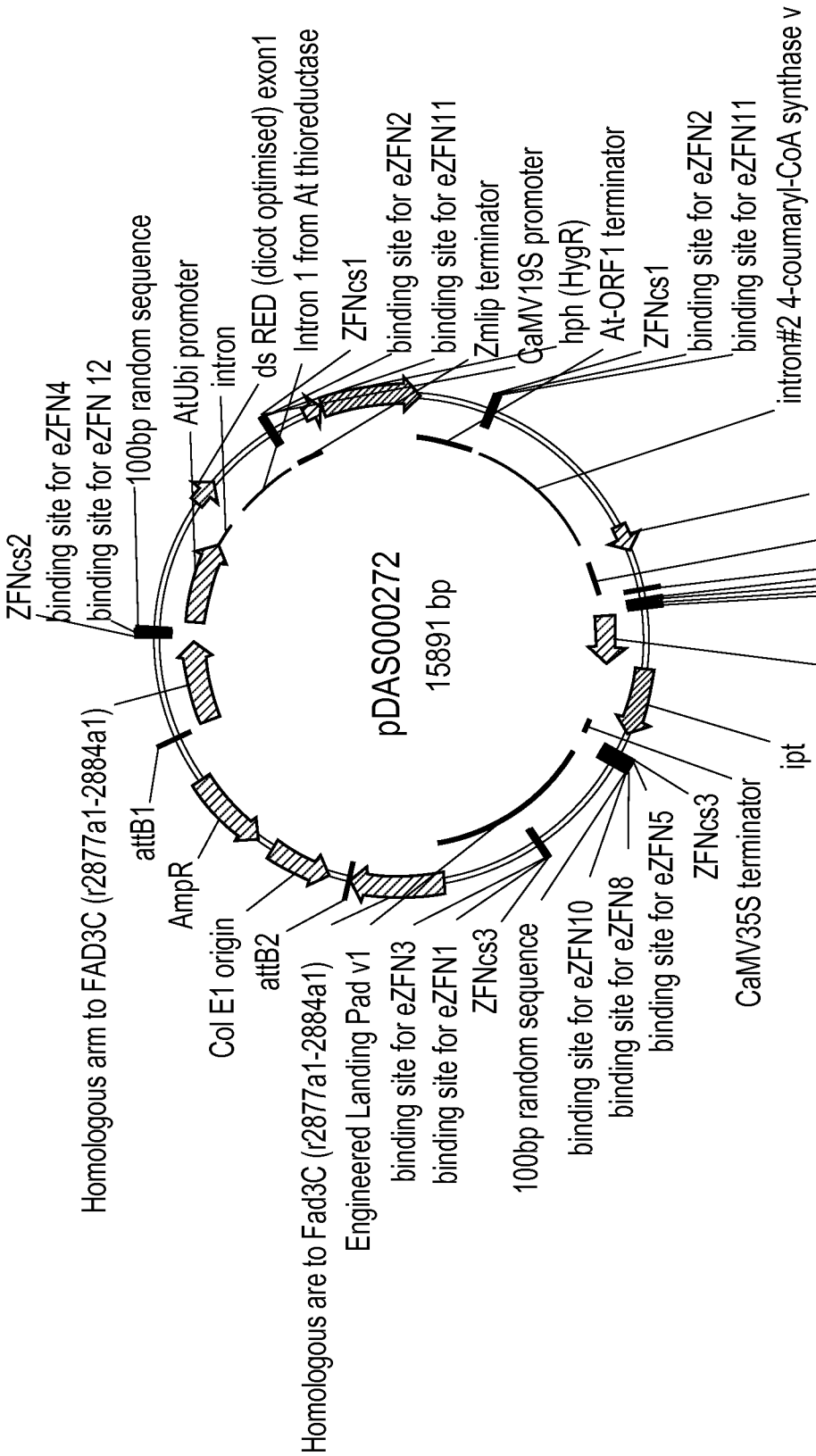

pDAS000272
15891 bp
ZFNcs2
binding site for eZFN4
binding site for eZFN 12
100bp random sequence
AtUbi promoter
intron
ds RED (dicot optimised) exon1
Intron 1 from At thioreductase
ZFNcs1
binding site for eZFN2
binding site for eZFN11
Zmlip terminator
CaMV19S promoter
hph (HygR)
At-ORF1 terminator
ZFNcs1
binding site for eZFN2
binding site for eZFN11
intron#2 4-coumaryl-CoA synthase v
PAT (v6) 3' end
AtuORF23 terminator
100bp random sequence
eZFNcs3
binding site for eZFN5
binding site for eZFN8
binding site for eZFN10
AtMYB32(T) promoter
ipt
CaMV35S terminator
ZFNcs3
binding site for eZFN5
binding site for eZFN8
binding site for eZFN10
100bp random sequence
ZFNcs3
binding site for eZFN1
binding site for eZFN3
Engineered Landing Pad v1
Homologous are to Fad3C (r2877a1-2884a1)
attB2
Col E1 origin
AmpR
attB1
Homologous arm to FAD3C (r2877a1-2884a1)

FIG. 6

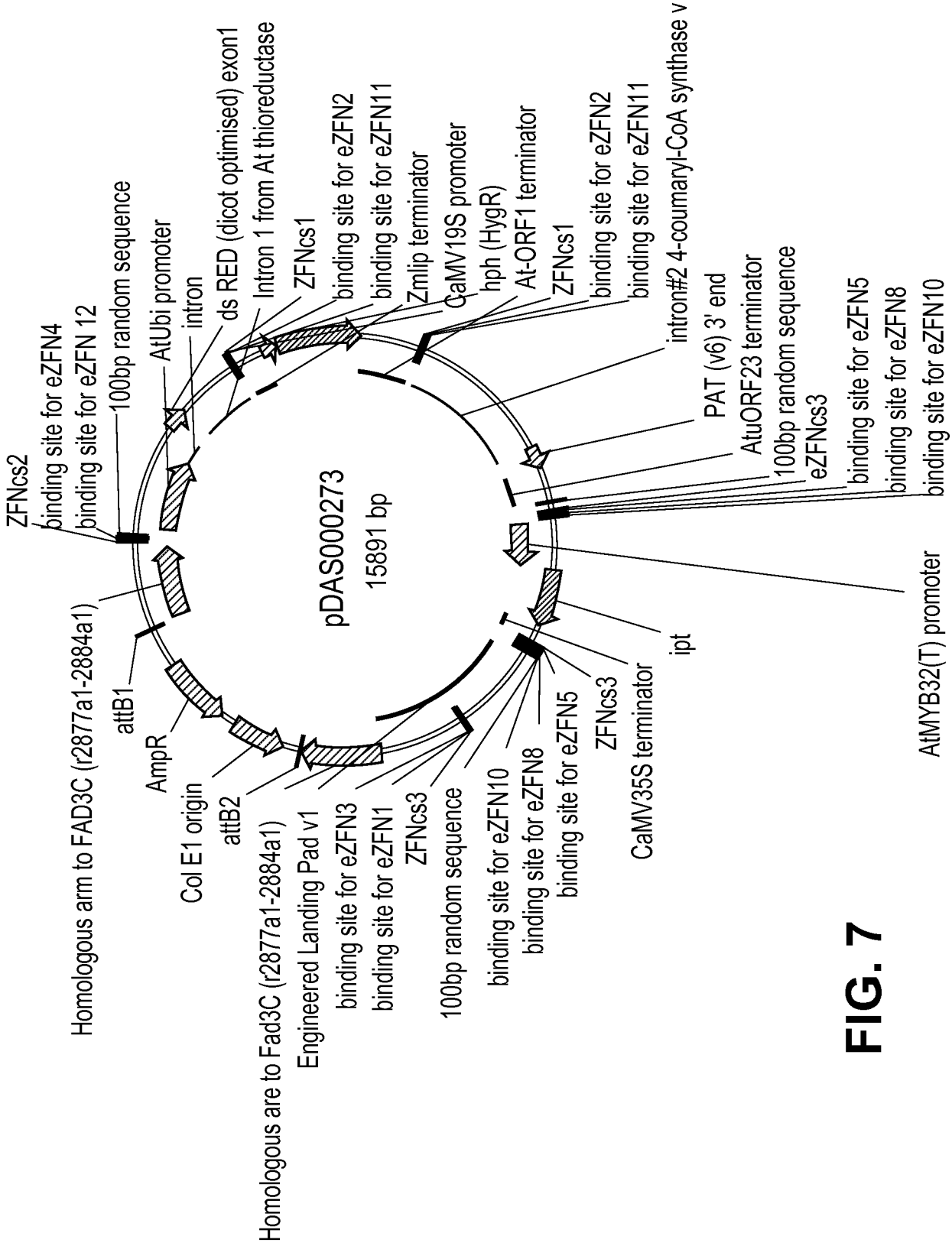

pDAS000273
15891 bp
ZFNcs2
binding site for eZFN4
binding site for eZFN 12
100bp random sequence
AtUbi promoter
intron
ds RED (dicot optimised) exon1
Intron 1 from At thioreductase
ZFNcs1
binding site for eZFN2
binding site for eZFN11
Zmlip terminator
CaMV19S promoter
hph (HygR)
At-ORF1 terminator
ZFNcs1
binding site for eZFN2
binding site for eZFN11
intron#2 4-coumaryl-CoA synthase v
PAT (v6) 3' end
AtuORF23 terminator
100bp random sequence
eZFNcs3
binding site for eZFN5
binding site for eZFN8
binding site for eZFN10
AtMYB32(T) promoter
ipt
CaMV35S terminator
ZFNcs3
binding site for eZFN5
binding site for eZFN8
binding site for eZFN10
100bp random sequence
ZFNcs3
binding site for eZFN1
binding site for eZFN3
Engineered Landing Pad v1
Homologous are to Fad3C (r2877a1-2884a1)
attB2
Col E1 origin
AmpR
attB1
Homologous arm to FAD3C (r2877a1-2884a1)

FIG. 7

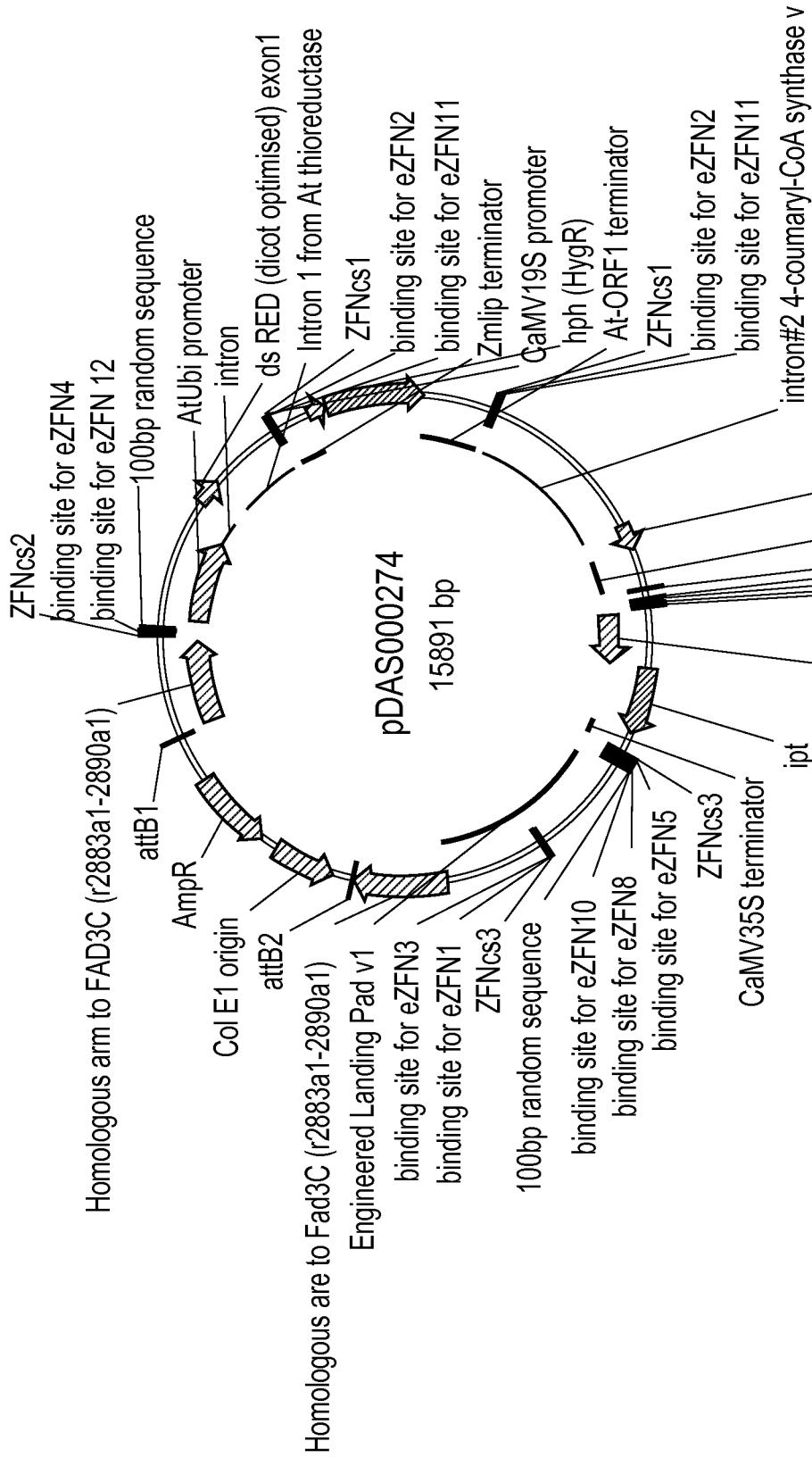

pDAS000274
15891 bp
ZFNcs2
binding site for eZFN4
binding site for eZFN 12
100bp random sequence
AtUbi promoter
intron
ds RED (dicot optimised) exon1
Intron 1 from At thioreductase
ZFNcs1
binding site for eZFN2
binding site for eZFN11
Zmlip terminator
CaMV19S promoter
hph (HygR)
At-ORF1 terminator
ZFNcs1
binding site for eZFN2
binding site for eZFN11
intron#2 4-coumaryl-CoA synthase v
PAT (v6) 3' end
AtuORF23 terminator
100bp random sequence
eZFNcs3
binding site for eZFN5
binding site for eZFN8
binding site for eZFN10
AtMYB32(T) promoter
ipt
CaMV35S terminator
ZFNcs3
binding site for eZFN5
binding site for eZFN8
binding site for eZFN10
100bp random sequence
ZFNcs3
binding site for eZFN1
binding site for eZFN3
Engineered Landing Pad v1
Homologous are to Fad3C (r2883a1-2890a1)
attB2
Col E1 origin
AmpR
attB1
Homologous arm to FAD3C (r2883a1-2890a1)

FIG. 8

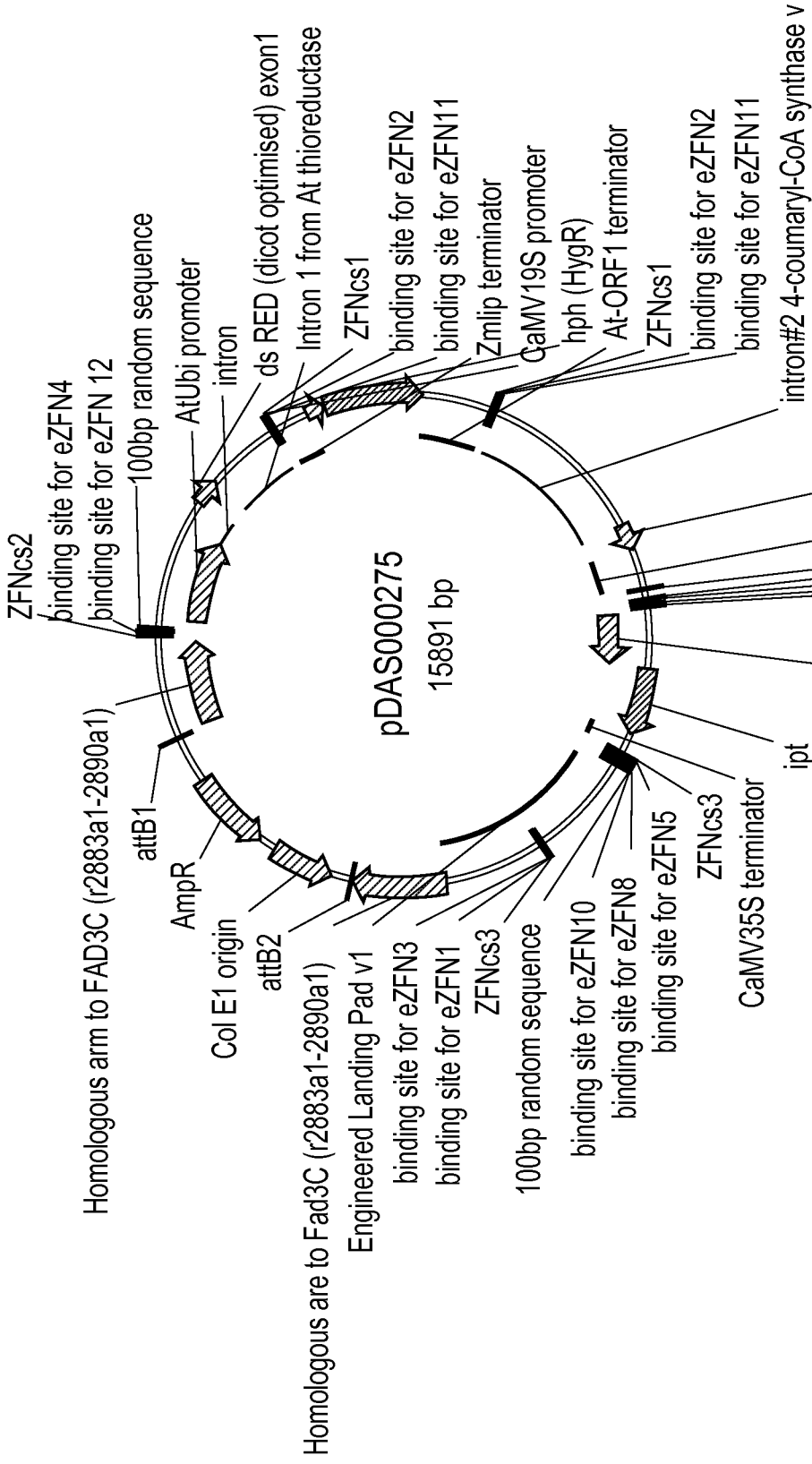

pDAS000275
15891 bp
ZFNcs2
binding site for eZFN4
binding site for eZFN 12
100bp random sequence
AtUbi promoter
intron
ds RED (dicot optimised) exon1
Intron 1 from At thioreductase
ZFNcs1
binding site for eZFN2
binding site for eZFN11
Zmlip terminator
CaMV19S promoter
hph (HygR)
At-ORF1 terminator
ZFNcs1
binding site for eZFN2
binding site for eZFN11
intron#2 4-coumaryl-CoA synthase v
PAT (v6) 3' end
AtuORF23 terminator
100bp random sequence
eZFNcs3
binding site for eZFN5
binding site for eZFN8
binding site for eZFN10
AtMYB32(T) promoter
ipt
CaMV35S terminator
ZFNcs3
binding site for eZFN5
binding site for eZFN8
binding site for eZFN10
100bp random sequence
ZFNcs3
binding site for eZFN1
binding site for eZFN3
Engineered Landing Pad v1
Homologous are to Fad3C (r2883a1-2890a1)
attB2
Col E1 origin
AmpR
attB1
Homologous arm to FAD3C (r2883a1-2890a1)

FIG. 9

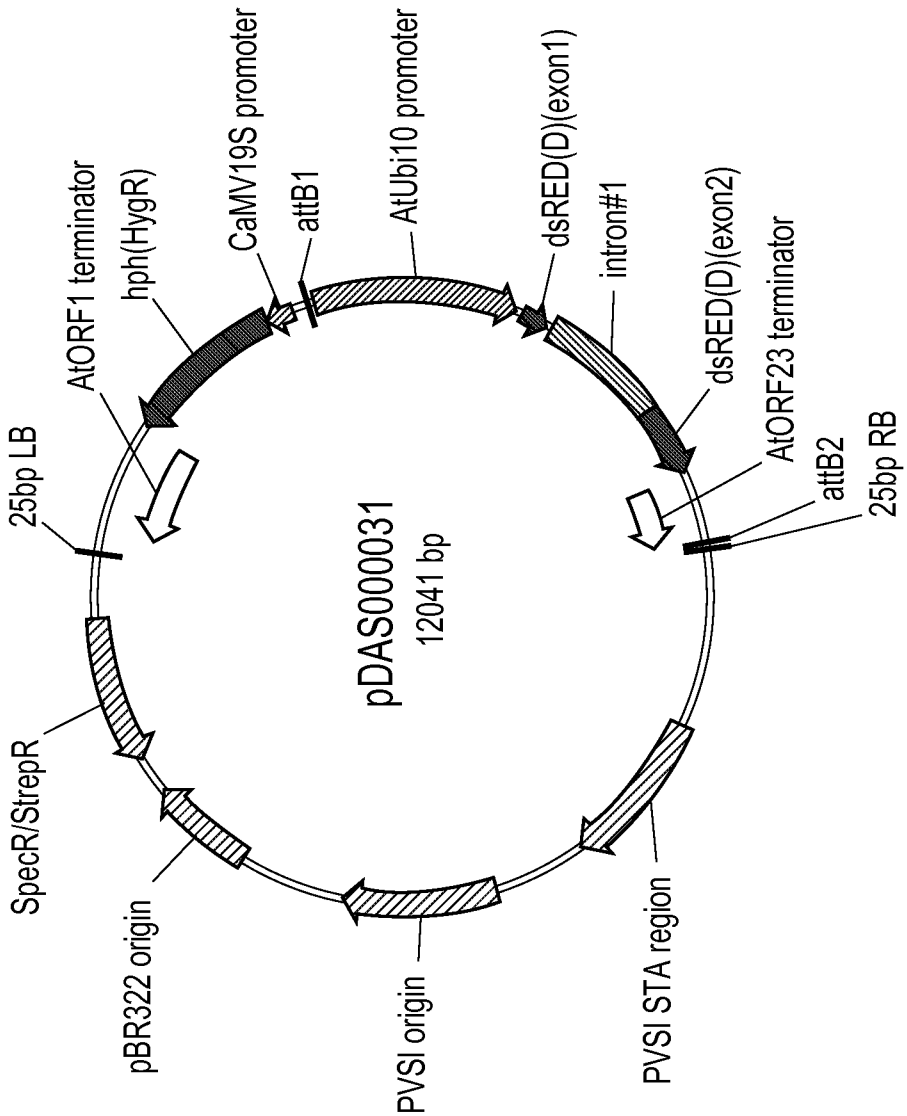

pDAS000031
12041 bp
25bp LB
AtORF1 terminator
hph(HygR)
CaMV19S promoter
attB1
AtUbi10 promoter
dsRED(D)(exon1)
intron#1
dsRED(D)(exon2)
AtORF23 terminator
attB2
25bp RB
SpecR/StrepR
pBR322 origin
PVSI origin
PVSI STA region

FIG. 10

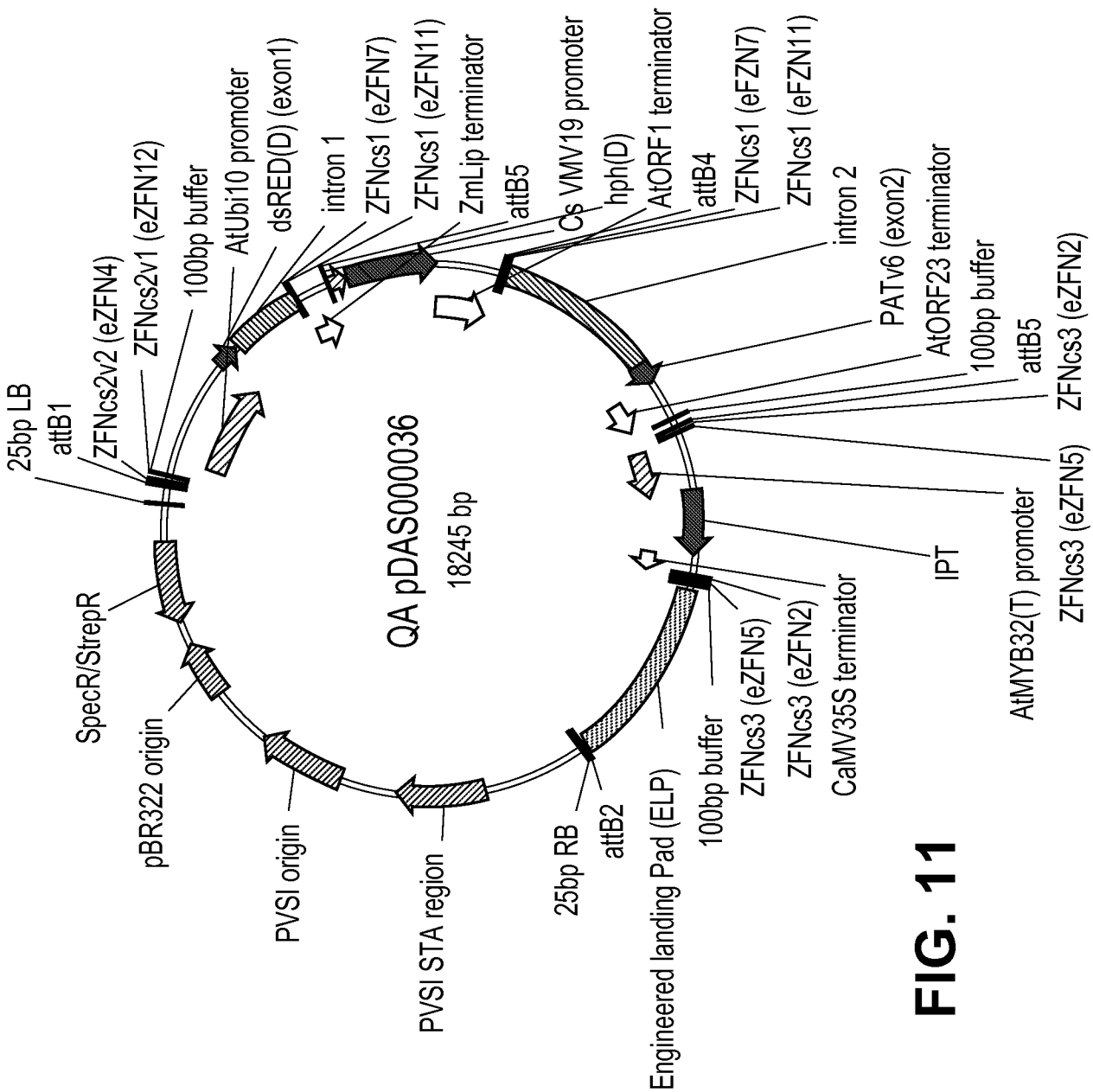

QA pDAS000036
18245 bp
25bp LB
attB1
ZFNcs2v2 (eZFN4)
ZFNcs2v1 (eZFN12)
100bp buffer
AtUbi10 promoter
dsRED(D) (exon1)
intron 1
ZFNcs1 (eZFN7)
ZFNcs1 (eZFN11)
ZmLip terminator
attB5
Cs VMV19 promoter
hph(D)
AtORF1 terminator
attB4
ZFNcs1 (eFZN7)
ZFNcs1 (eFZN11)
intron 2
PATv6 (exon2)
AtORF23 terminator
100bp buffer
attB5
ZFNcs3 (eZFN2)
ZFNcs3 (eZFN5)
AtMYB32(T) promoter
IPT
CaMV35S terminator
ZFNcs3 (eZFN2)
ZFNcs3 (eZFN5)
100bp buffer
Engineered landing Pad (ELP)
attB2
25bp RB
PVSI STA region
PVSI origin
pBR322 origin
SpecR/StrepR

FIG. 11

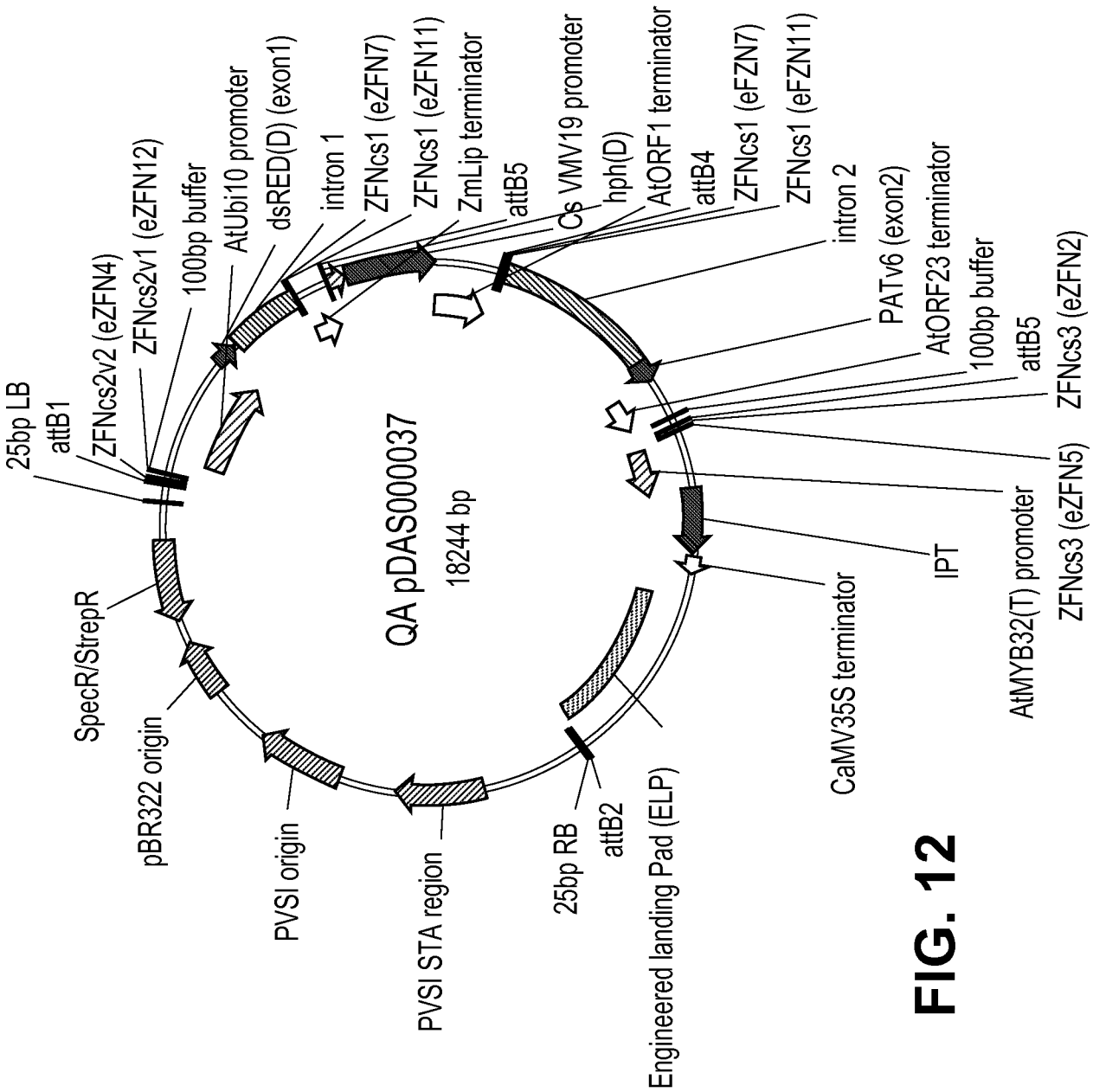

QA pDAS000037
18244 bp
25bp LB
attB1
ZFNcs2v2 (eZFN4)
ZFNcs2v1 (eZFN12)
100bp buffer
AtUbi10 promoter
dsRED(D) (exon1)
intron 1
ZFNcs1 (eZFN7)
ZFNcs1 (eZFN11)
ZmLip terminator
attB5
Cs VMV19 promoter
hph(D)
AtORF1 terminator
attB4
ZFNcs1 (eFZN7)
ZFNcs1 (eFZN11)
intron 2
PATv6 (exon2)
AtORF23 terminator
100bp buffer
attB5
ZFNcs3 (eZFN2)
ZFNcs3 (eZFN5)
AtMYB32(T) promoter
IPT
CaMV35S terminator
Engineered landing Pad (ELP)
attB2
25bp RB
PVSI STA region
PVSI origin
pBR322 origin
SpecR/StrepR

FIG. 12

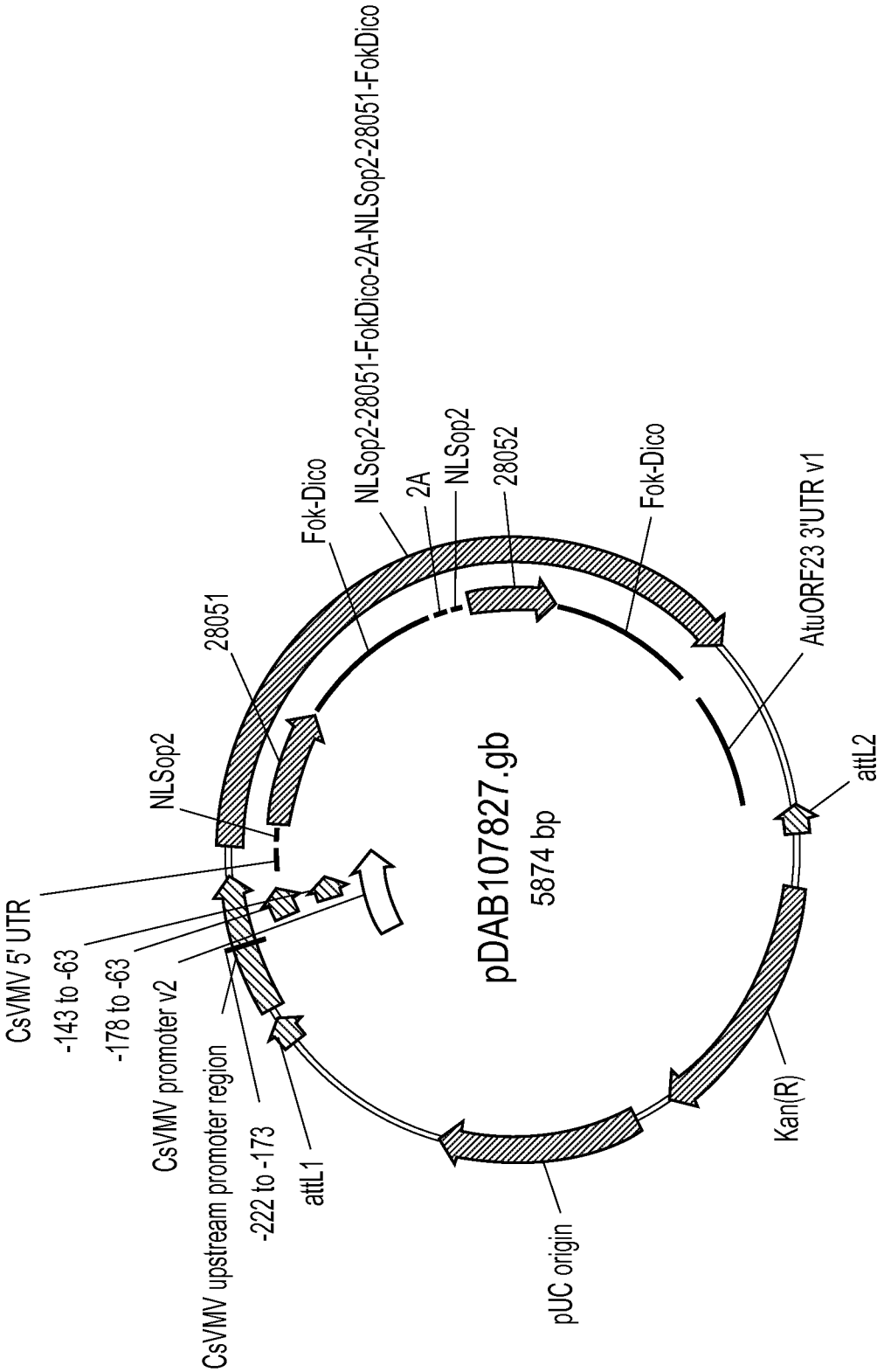

pDAB107827.gb
5874 bp
Fok-Dico
NLSop2-28051-FokDico-2A-NLSop2-28051-FokDico
2A
NLSop2
28052
Fok-Dico
AtuORF23 3'UTR v1
attL2
Kan(R)
pUC origin
attL1
-222 to -173
CsVMV upstream promoter region
CsVMV promoter v2
-178 to -63
-143 to -63
CsVMV 5' UTR
NLSop2
28051

FIG. 13

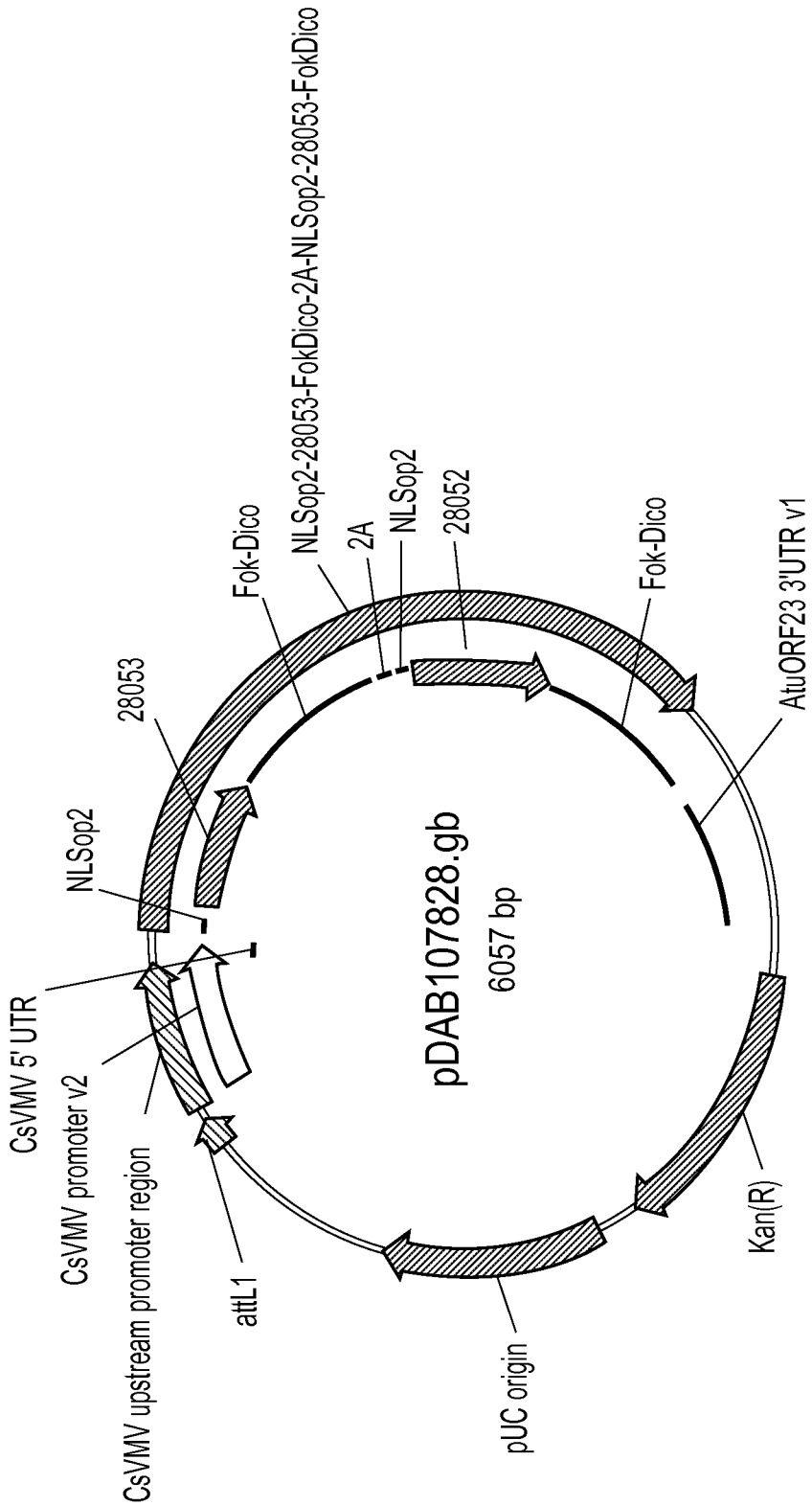

pDAB107828.gb
6057 bp
Fok-Dico
NLSop2-28053-FokDico-2A-NLSop2-28053-FokDico
2A
NLSop2
28052
Fok-Dico
AtuORF23 3'UTR v1
28053
NLSop2
CsVMV 5' UTR
CsVMV promoter v2
CsVMV upstream promoter region
attL1
pUC origin
Kan(R)

FIG. 14

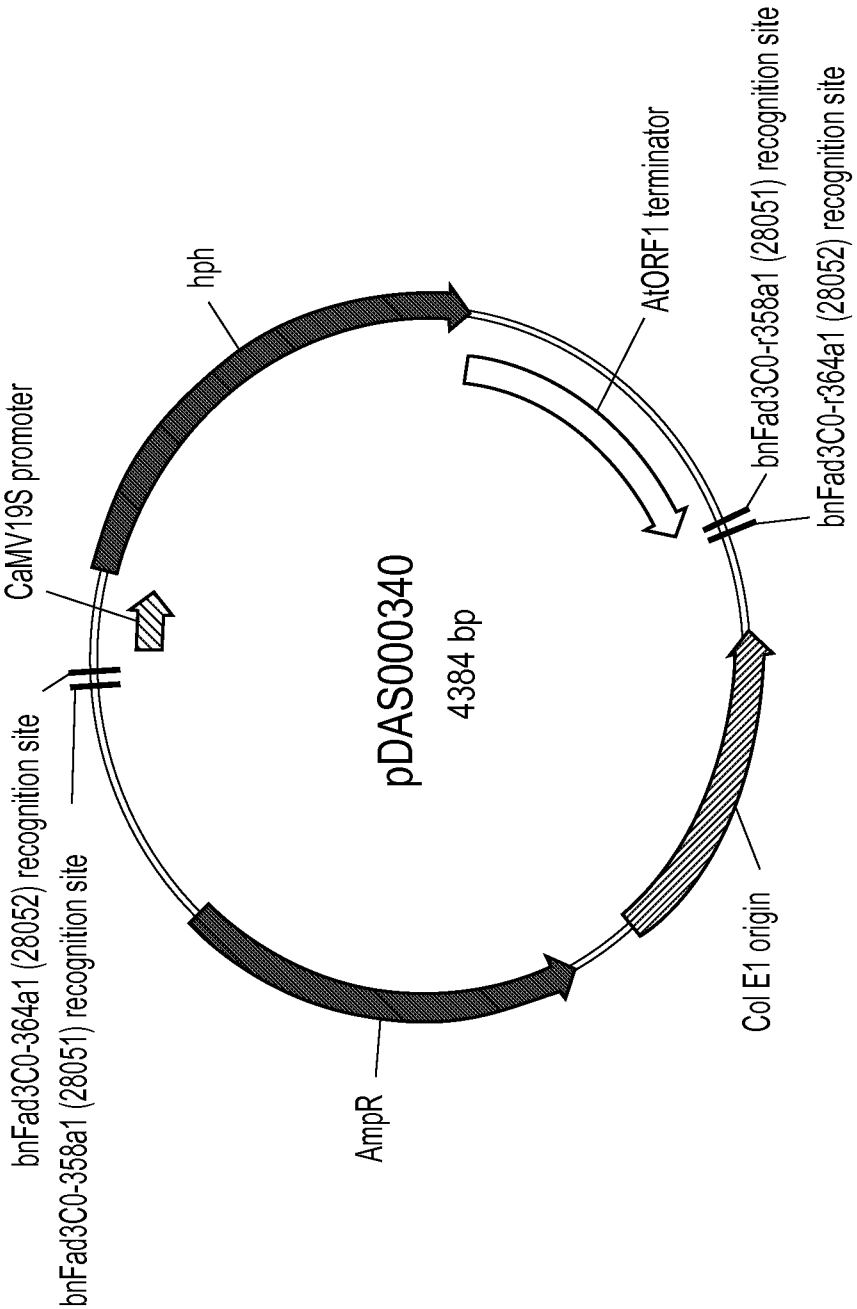

hph
AtORF1 terminator
bnFad3C0-r358a1 (28051) recognition site
bnFad3C0-r364a1 (28052) recognition site
CaMV19S promoter
pDAS000340
4384 bp
bnFad3C0-364a1 (28052) recognition site
bnFad3C0-358a1 (28051) recognition site
AmpR
Col E1 origin

FIG. 15

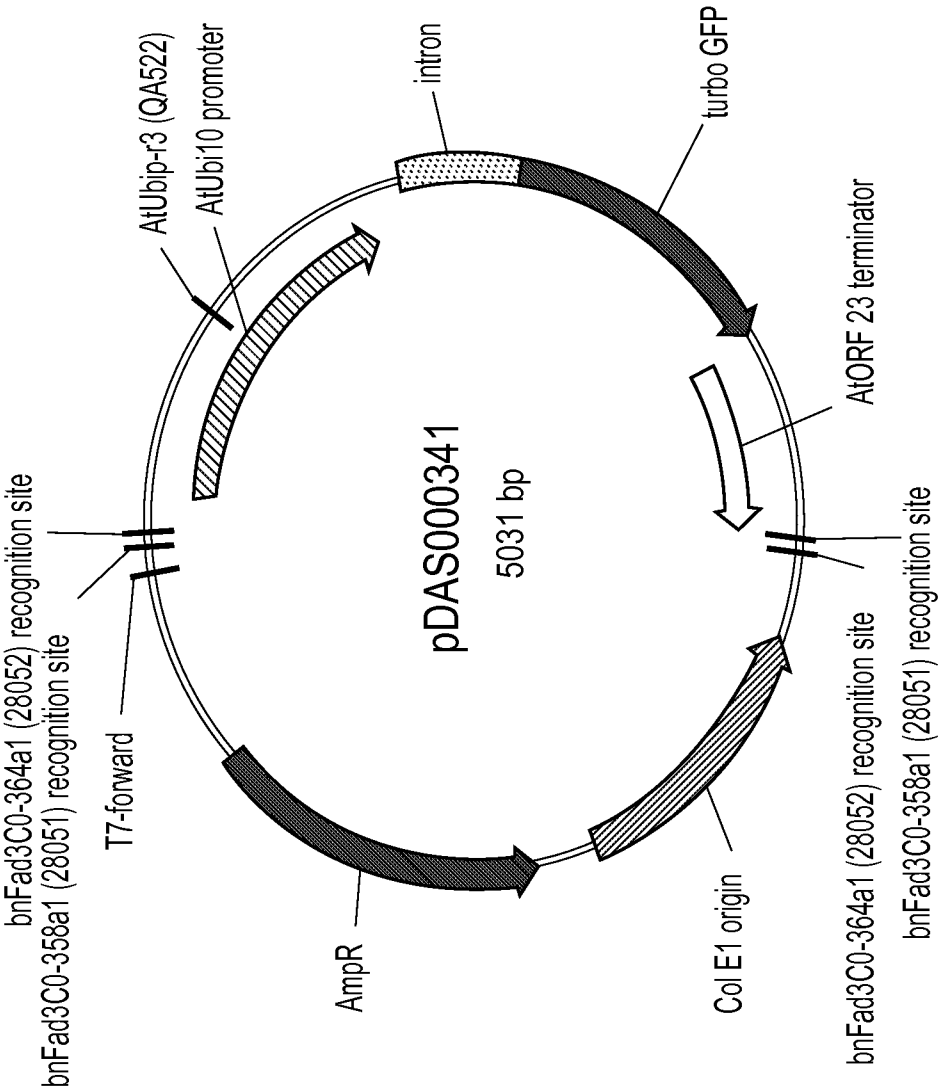

pDAS000341
5031 bp
AtUbip-r3 (QA522)
AtUbi10 promoter
intron
turbo GFP
AtORF 23 terminator
bnFad3C0-364a1 (28052) recognition site
bnFad3C0-358a1 (28051) recognition site
T7-forward
AmpR
Col E1 origin
bnFad3C0-364a1 (28052) recognition site
bnFad3C0-358a1 (28051) recognition site

FIG. 16

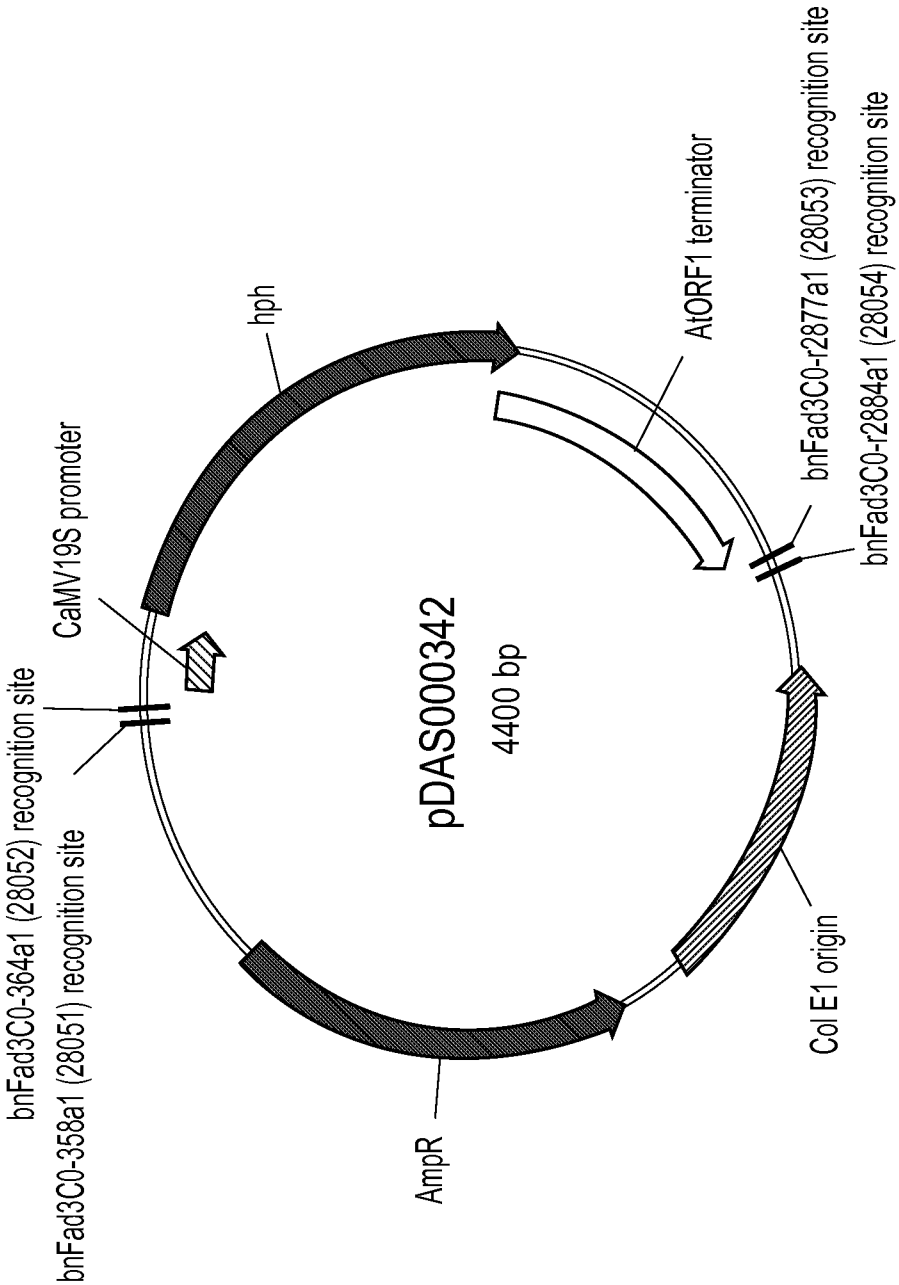

pDAS000342
4400 bp
hph
AtORF1 terminator
bnFad3C0-r2877a1 (28053) recognition site
bnFad3C0-r2884a1 (28054) recognition site
Col E1 origin
AmpR
bnFad3C0-364a1 (28052) recognition site
bnFad3C0-358a1 (28051) recognition site
CaMV19S promoter

FIG. 17

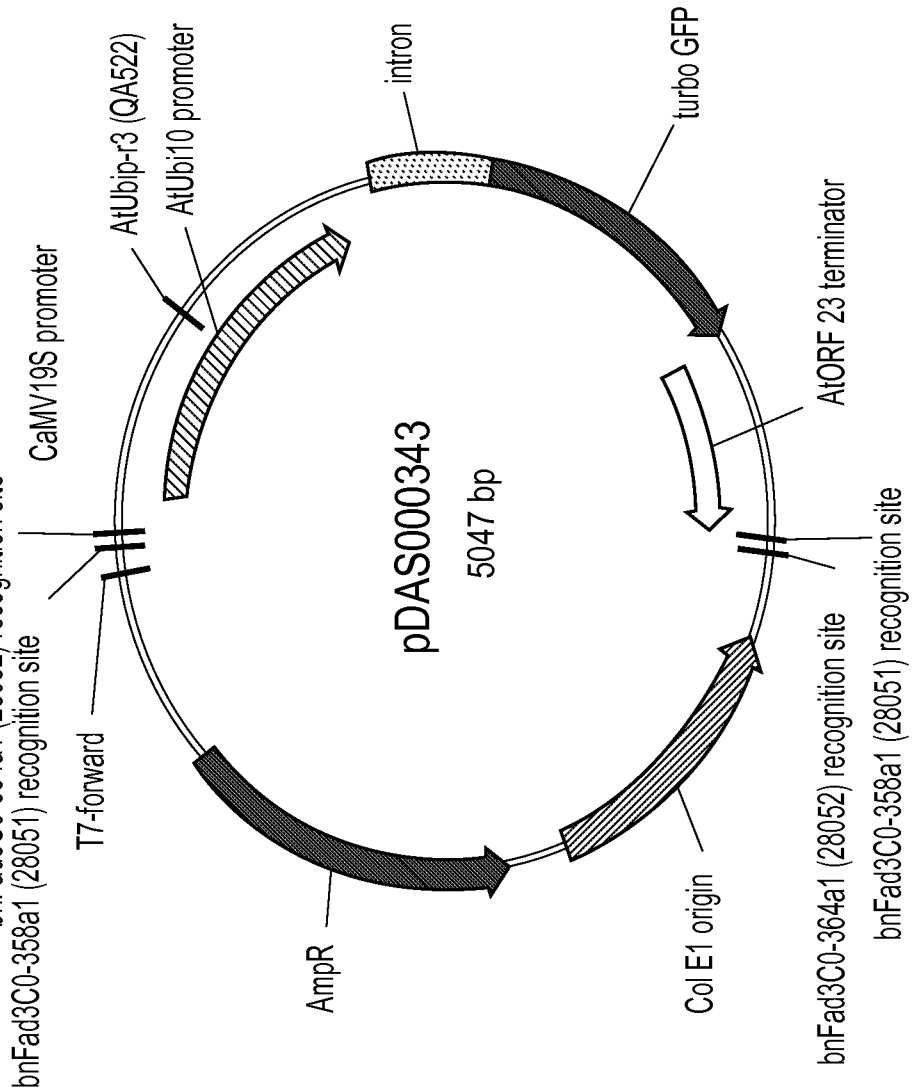
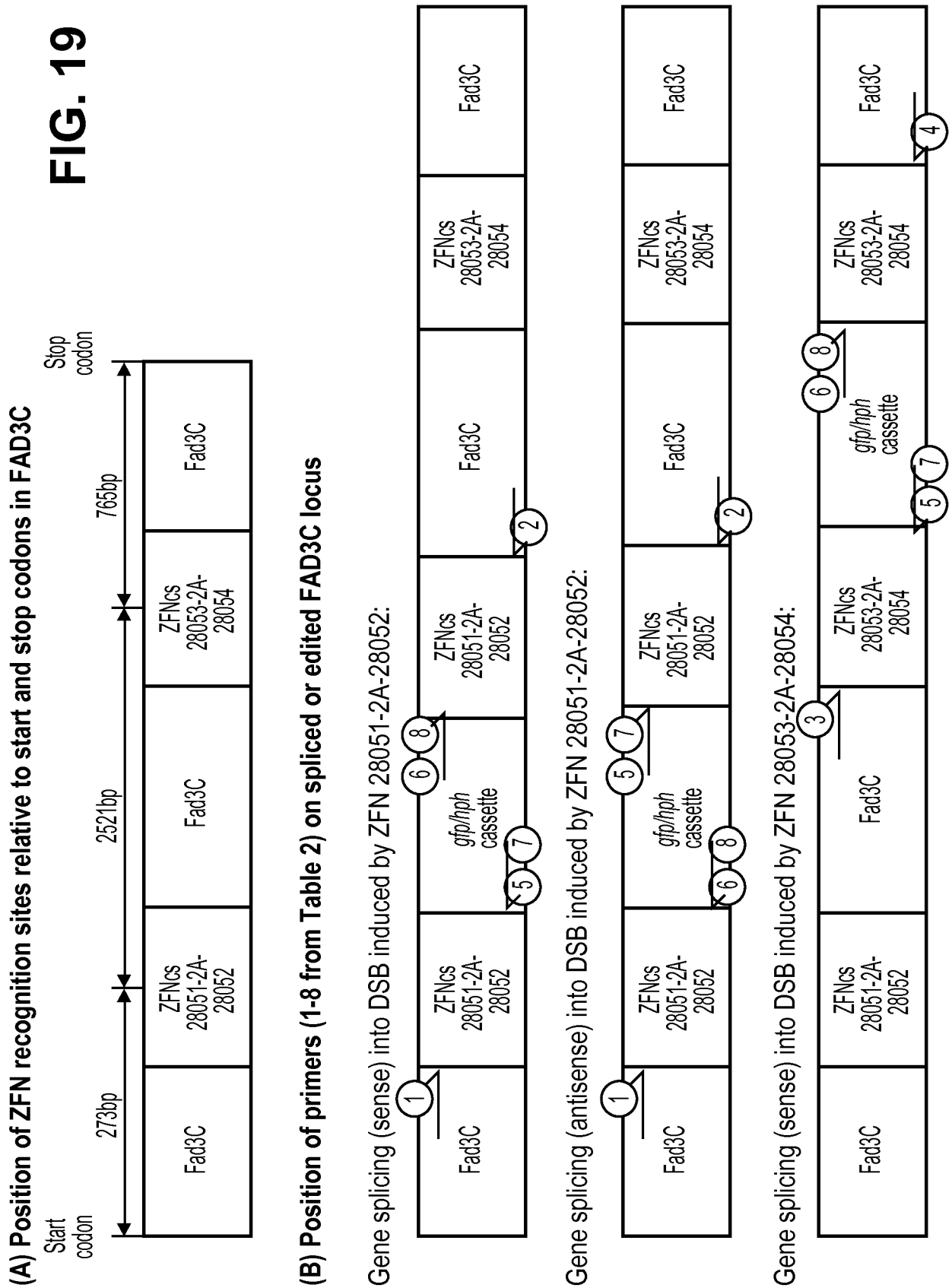

pDAS000343
5047 bp
bnFad3C0-364a1 (28052) recognition site
bnFad3C0-358a1 (28051) recognition site
CaMV19S promoter
AtUbip-r3 (QA522)
AtUbi10 promoter
intron
turbo GFP
AtORF 23 terminator
bnFad3C0-364a1 (28052) recognition site
bnFad3C0-358a1 (28051) recognition site
Col E1 origin
AmpR
T7-forward

FIG. 18

FIG. 19


(A) Position of ZFN recognition sites relative to start and stop codons in FAD3C
Start codon
Stop codon
273bp
2521bp
765bp
Fad3C
ZFNcs 28051-2A-28052
Fad3C
ZFNcs 28053-2A-28054
Fad3C
(B) Position of primers (1-8 from Table 2) on spliced or edited FAD3C locus
Gene splicing (sense) into DSB induced by ZFN 28051-2A-28052:
Gene splicing (antisense) into DSB induced by ZFN 28051-2A-28052:
Gene splicing (sense) into DSB induced by ZFN 28053-2A-28054:
gfp/hph cassette

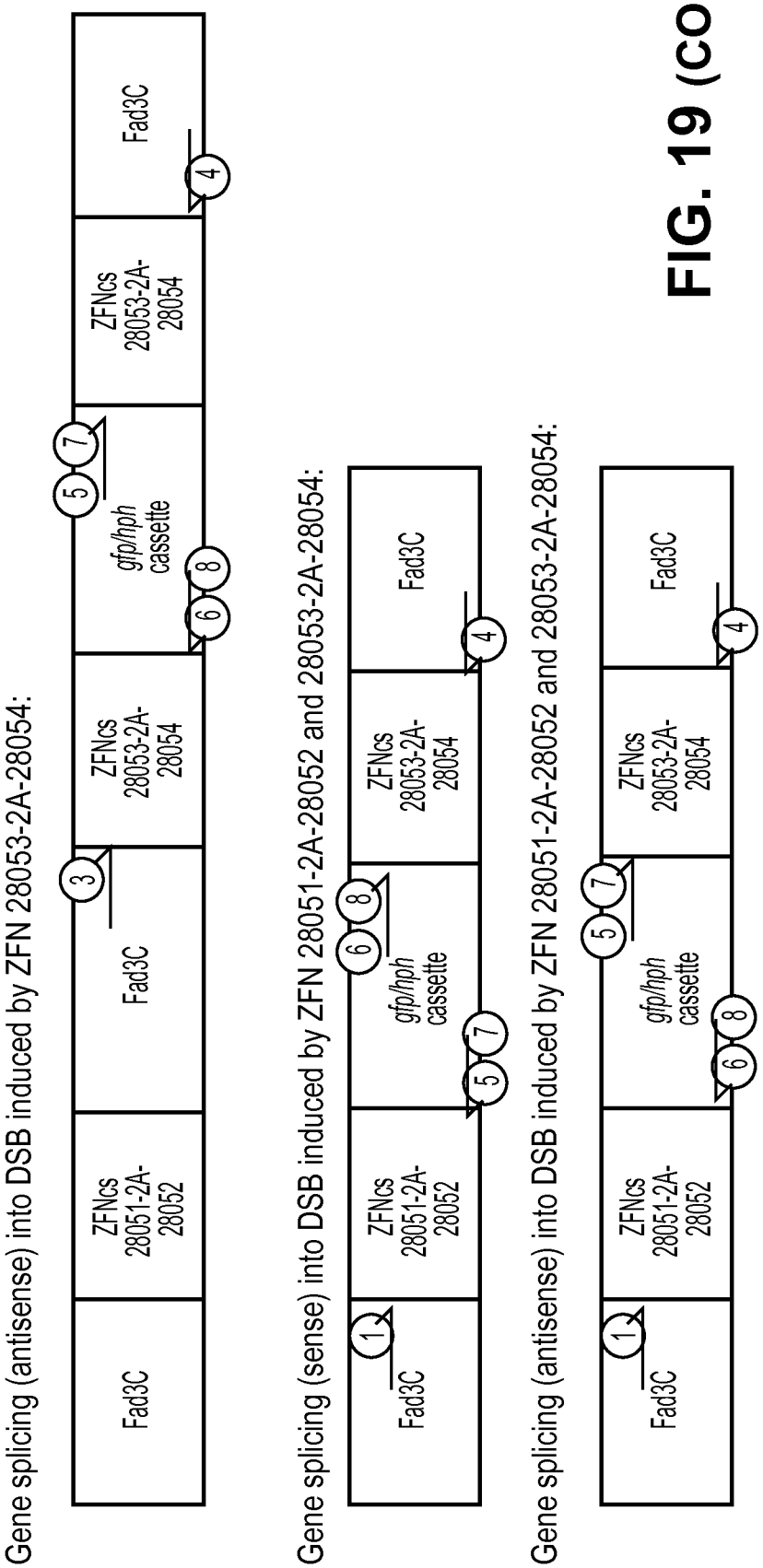

Gene splicing (antisense) into DSB induced by ZFN 28053-2A-28054:
Fad3C
ZFNcs 28051-2A-28052
Fad3C
ZFNcs 28053-2A-28054
gfp/hph cassette
ZFNcs 28053-2A-28054
Fad3C
Gene splicing (sense) into DSB induced by ZFN 28051-2A-28052 and 28053-2A-28054:
Fad3C
ZFNcs 28051-2A-28052
gfp/hph cassette
ZFNcs 28053-2A-28054
Fad3C
Gene splicing (antisense) into DSB induced by ZFN 28051-2A-28052 and 28053-2A-28054:
Fad3C
ZFNcs 28051-2A-28052
gfp/hph cassette
ZFNcs 28053-2A-28054
Fad3C

FIG. 19 (CONT.)

(A) Sequences amplified from the junction of the tGFP cassette from pDAS000341 with Fad3C at the DSB recognized by ZFN 28051-2A-28052

":" indicates deletions at cut-site

5' junction of tGFP cassette with FadC

Fad3	ZFN recognition site 28051	Inserted Bases	ZFN recognition site 28052	AtUbi10p
TTCTGGCCTCTTTATTGGGCCGCCCAAGGAACCCTTTTCTGGGCCATCT			TCGTACTCGGCCACGACTGGTAATTTAATGGATCCACTAGTAA	
TTCTGGCCTCTTTATTGGGCCGCCCAAGGAACCCTTTTCTGGGCCATCT			TCGTACTCGGCCACGACTGGTAATTTAATGGATCCACTAGTAA	
TTCTGGCCTCTTTATTGGGCCGCCCAAGGAACCCTTTTCTGGGCCATC:		CAGTCGTGGCCGAGTACGAAGATGGCCCAGA	:::TACTCGGCCACGACTGGTAATTTAATGGATCCACTAGTAA	
TTCTGGCCTCTTTATTGGGCCGCCCAAGGAACCCTTTTCTGGGCCATC:			::GTACTCGGCCACGACTGGTAATTTAATGGATCCACTAGTAA	
TTCTGGCCTCTTTATTGGGCCGCCCAAGGAACCCTTTTCTGGG::::::		TATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGCGTGCACGAAC	:CGTACTCGGCCACGACTGGTAATTTAATGGATCCACTAGTAA	
TTCTGGCCTCTTTATTGGGCCGCCCAAGGAACCCTTTTCTGGGCCA:::			::::::::::::::GACTGGTAATTTAATGGATCCACTAGTAA	

3' junction of tGFP cassette with FadC

AtuOrf23t	ZFN recognition site 28051	Inserted Bases	ZFN recognition site 28052	Fad3C
TCCAAGGTTGCGGCCGCGCCCAAGGAACCCTTTTCTGGGCCATCT			TCGTACTCGGCCACGACTGGTAATTTAATTTTCAATTTATTT	
TCCAAGGTTGCGGCCGCGCCCAAGGAACCCTTTTCTGGGCCAT::			:::TACTCGGCCACGACTGGTAATTTAATTTTCAATTTATTT	
TCCAAGGTTGCGGCCGCGCCCAAGGAACCCTTTTCTGGGCCAT:T			:::TACTCGGCCACGACTGGTAATTTAATTTTCAATTTATTT	
:::::78 bases deleted ::::::::::::::::::::::			:CGTACTCGGCCACGACTGGTAATTTAATTTTCAATTTATTT	
TCCAAGGTTGCGGCCGCGCCCAAGGAACCCTTTTCTGGGCCATCT			TCGTACTCGGCCACGACTGGTAATTTAATTTTCAATTTATTT	
TCCAAGGTTGCGGCCGCGCCCAAGGAACCCTTTTCTGGGCCATCTTC			::::::::::::::TGGTAATTTAATTTTCAATTTATTTTT	
TCCAAGGTTGCGGCCGCGCCCAAGGAACCCTTTTCTGG::::::		TAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGA	TCGTACTCGGCCACGACTGGTAATTTAATTTTCAATTTATTT	
TCCAAGGTTGCGGCCGCGCCCAAGGAACCCTTTTCTGGGCCATCTTACGAGCGTAATGGCTGGCCTGTTGAACAAGTCTGGAAAGAAATGCATAAACATATCCCAGCCACGACT			:::::::::::::::::GGTAATTTAATTTTCAATTTATTT	

SUBSTITUTE SHEET (RULE 26)

FIG. 20A

(B) Sequences amplified from the junction of the tGFP cassette from pDAS000343 with Fad3C at the DSBs recognized by ZFNs 28051-2A-28052 and 28053-2A-28054. ":" indicates deletions at cut-site

Fad3	ZFN recognition site 28052	Inserted Bases	ZFN recognition site 28052	AtUbi10p
TAGTTTATTTGCCCCAAGCGAGAGAAAGCTTATTGCAACTTCAAC			**TCGTACTCGGCCACGACTGGTAATTTAATGGATCCACTAGTAA**	
TAGTTTATTTGCCCCAAGCGAGAGAAAGCTTATTGCAACTTCAACT			::::::::::::::82 bases deleted::::::::::::::	
TAGTTTATTTGCCCCAAGCGAGAGAAAGCTTATTGCAACTTCAA::		CG	::::::::::::::68 bases deleted::::::::::::::	
TAGTTTATTTGCCCCAAGCGAGAGAAAGCTTATTGCAACTTCAACT			TCGTACTCGGCCACGACTGGTAATTTAATGGATCCACTAGTAA	
TAGTTTATTTGCCCCAAGCGAGAGAAAGCTTATTGCAACTTCAACT		AT	::GTACTCGGCCACGACTGGTAATTTAATGGATCCACTAGTAA	
TAGTTTATTTGCCCCAAGCGAGAGAAAGCTTATTGCAACTTCA:::			:::TACTCGGCCACGACTGGTAATTTAATGGATCCACTAGTAA	
::::::::121 bases deleted::::::::::::::::::::			::::::::::::::::::AGGTAATTTAATGGATCCACTAGTAA	

AtuOrf23t	ZFN recognition site 28053	Inserted Bases	ZFN recognition site 28054	Fad3C
TCCAAGGTTGCGGCCGCAGCGAGAGAAAGCTTATTGCAACTTCAAC			**TACTTGCTGGTCGATCGTGTTGGCCACTCTTGTTTATCTATCA**	
TCCAAGGTTGCGGCCGCAGCGAGAGAAAGCTTATTGCAACTTCA::			:ACTTGCTGGTCGATCGTGTTGGCCACTCTTGTTTATCTATCA	
TCCAAGGTTGCGGCCGC::::::::::::::::::::::::::::	GCGCCGACCCAGCTTTCTTGTACAAAGTTGGCATTATAAGAAAGCATTGCTTATCAATTTGTTGCAACGAACAGGTCACTATCAGTCAAA		:ACTTGCTGGTCGATCGTGTTGGCCACTCTTGTTTATCTATCA	
TCCAAGGTTGCGGCCGCAGCGAGAGAAAGCTTATTGCAAA::::::		CTTC	:ACTTGCTGGTCGATCGTGTTGGCCACTCTTGTTTATCTATCA	
TCCAAGGTTGCGGCCGCAGCGAGAGAAAGCTTATTGCAACTTCA::		GATAAAAGTTGCTCGCCTGTGTGGGTGTGGATGCT	:ACTTGCTGGTCGATCGTGTTGGCCACTCTTGTTTATCTATCA	
TCCAAGGTTGCGGCCGCAGCGAGAGAAAGCTTATTGCAACTTCAAC		TACAC	TACTTGCTGGTCGATCGTGTTGGCCACTCTTGTTTATCTATCA	
TCCAAGGTTGCGGCCGCAGCGAGAGAAAGCTTATTGCAACTTCAAC			TACTTGCTGGTCGATCGTGTTGGCCACTCTTGTTTATCTATCA	

SUBSTITUTE SHEET (RULE 26)

FIG. 20B

(A)

Sample	Fad3	ZFN recognition site 28051	# of Extra Bases Inserted	ZFN recognition site 28052	CaMV19sp
	TTCTGGCCTCTTTATTGGGCCGCCCAAGGAACCCTTTCTGGGCCATCT			TCGTACTCGGCCACGACTGGTAATTTAATGGATCCAACCGACAACCACTT	
349711	TTCTGGCCTCTTTATTGGGCCGCCCAAGGAACCCTTTCTGGGCCAT::		442	:::TACTCGGCCACGACTGGTAATTTAATGGATCCAACCGACAACCACTT	
349215c	TTCTGGCCTCTTTATTGGGCCGCCCAAGGAACCCTTTCTGGGCCATC:		406	TCGTACTCGGCCACGACTGGTAATTTAATGGATCCAACCGACAACCACTT	
349216c	TTCTGGCCTCTTTATTGGGCCGCCCAAGGAACCCTTTCTGGGCCATC:		406	TCGTACTCGGCCACGACTGGTAATTTAATGGATCCAACCGACAACCACTT	
349685	TTCTGGCCTCTTTATTGGGCCGCCCAAGGAACCCTTTCTGGGCCAT ::			:::TACTCGGCCACGACTGGTAATTTAATGGATCCAACCGACAACCACTT	
346258	TTCTGGCCTCTTTATTGGGCCGCCCAAGGAACCCTTTCTGGGCCATC:		435	:::TACTCGGCCACGACTGGTAATTTAATGGATCCAACCGACAACCACTT	
348918	TTCTGGCCTCTTTATTGGGCCGCCCAAGGAACCCTTTCTGGGCC::::		378	:::TACTCGGCCACGACTGGTAATTTAATGGATCCAACCGACAACCACTT	
359900	TTCTGGCCTCTTTATTGGGCCGCCCAAGGAACCCTTTCTGGGCCATCT			:::TACTCGGCCACGACTGGTAATTTAATGGATCCAACCGACAACCACTT	
346125	TTCTGGCCTCTTTATTGGGCCGCCCAAGGAACCCTTTCTGGGCCA::T		62	:::TACTCGGCCACGACTGGTAATTTAATGGATCCAACCGACAACCACTT	
348919	TTCTGGCCTCTTTATTGGGCCGCCCAAGGAACCCTTTCTGGGCCA:::		378	:::TACTCGGCCACGACTGGTAATTTAATGGATCCAACCGACAACCACTT	

(B)

Sample	AtuORF1-t-	ZFN recognition site 28051	# of Extra Bases Inserted	ZFN recognition site 28052	Fad3C
	GTAATACATAGCGGCCGCGCCCAAGGAACCCTTTCTGGGCCATCT			TCGTACTCGGCCACGACTGGTAATTTAATTTCAATTTATTTTTCTTCAACTTCTTA	
346175	GTAATACATAGCGGCCGCGCCCAAGGAACCCTTTCTGGGCCATCT			:::::::::GCCACGACTGGTAATTTAATTTCAATTTATTTTTCTTCAACTTCTTA	
346102	**GTAATACATAGCGGCCGCGCCCAAGGAACCCTTTCTGGGCCATCT**			TCGTACTCGGCCACGACTGGTAATTTAATTTCAATTTATTTTTCTTCAACTTCTTA	

FIG. 21

(B)

Sample	AtuORF1-t-	ZFN recognition site 28053	# of Extra Bases Inserted	ZFN recognition site 28054	Fad3C
345888	GTAATACATAGCGGCCGCAGCGAGAGAAAGCTTATTGCAACTTCAAC		137	TACTTGCTGGTCGATCGTGTTGGCCACTCTTGTTTATCTATCATTCCTCGTTGGTC	
356731	GTAATACATAGCGGCCGCAGCGAGAGAAAGCTTATTGCAACTTCAAC			TACTTGCTGGTCGATCGTGTTGGCCACTCGGTACCTGGAGCACAAGACTGGCCTCA	

FIG. 22

(A)

Sample	Fad3	ZFN recognition site 28051	# of ExtraBases Inserted	ZFN recognition site 28052	CaMV19sp
	TTCTGGCCTCTTTATTGGGCCGCCCAAGGAACCCTTTTCTGGGCCATCT			TCGTACTCGGCCACGACTGGTAATTTAATGGATCCAACCGACAACCACTT	
1:1 #5	TTCTGGCCTCTTTATTGGGCCGCCCAAGGAACCCTTT::::::::::::		3	:::TACTCGGCCACGACTGGTAATTTAATGGATCCAACCGACAACCACTT	
1:1 #39	TTCTGGCCTCTTTATTGGGCCGCCCAAGGAACCCTTTTCTGG:::::::		153	TCGTACTCGGCCACGACTGGTAATTTAATGGATCCAACCGACAACCACTT	
1:1 #46	TTCTGGCCTCTTTATTGGGCCGCCCAAGGAACCCTTTTCTGG:::::::			:::::::::::::::::::::::136bp missing::::::::::::::::::::::::	
5:1 #63	TTCTGGCCTCTTTATTGGGCCGCCCAAGGAACCCTTTTCTGGGCCATCT		370	:::::::CGGCCACGACTGGTAATTTAATGGATCCAACCGACAACCACTT	
10:1 #16	TTCTGGCCTCTTTATTGGGCCGCCCAAGGAACCCTTTTCTGGGCCAT::		36	TCGTACTCGGCCACGACTGGTAATTTAATGGATCCAACCGACAACCACTT	
10:1 #64	TTCTGGCCTCTTTATTGGGCCGCCCAAGGAACCCTTTTCTGGGCCATCT		78	:::::::::::::::::::::52bp missing::::::::::::::::::::::	
10:1 #66	::::::::::::::::::51bp missing::::::::::::::::::::			:::::::::::::::::::::254bp missing::::::::::::::::::::::	

(B)

Sample	AtuORF1-t-	ZFN recognition site 28051	# of ExtraBases Inserted	ZFN recognition site 28052	Fad3C
	GTAATACATAGCGGCCGCGCCCAAGGAACCCTTTTCTGGGCCATCT			TCGTACTCGGCCACGACTGGTAATTTAATTTTTCTTCAATTTATTTTTTCTTCAACTTCTTA	
1:1 #5	GTAATACATAGCGGCCGCGCCCAAGGAACCCTTTTCTGGGCCAT::		4	:::::::::GCCACGACTGGTAATTTAATTTTCAATTTATTTTTTCTTCAACTTCTTA	
1:1 #39	GTAATACATAGCGGCCGCGCCCAAGGAACCCTTTTCTGGGCCATCT		112	:::::::::::::::::GACTGGTAATTTAATTTTCAATTTATTTTTTCTTCAACTTCTTA	
1:1 #46	GTAATACATAGCGGCCGCGCCCAAGGAACCCTTTTCTGGGCCATCT		112	:::::::::::::::::GACTGGTAATTTAATTTTCAATTTATTTTTTCTTCAACTTCTTA	
5:1 #63	GTAATACATAGCGGCCGCGCCCAAGGAACCCTTTTCTGGGCCAT::		234	::GTACTCGGCCACGACTGGTAATTTAATTTTCAATTTATTTTTTCTTCAACTTCTTA	
10:1 #16	::1655bp missing (possible double insertion)::			::GTACTCGGCCACGACTGGTAATTTAATTTTCAATTTATTTTTTCTTCAACTTCTTA	
10:1 #64	GTAATACATAGCGGCCGCGCCCAA::::::::::::::::::::::		9	:::TACTCGGCCACGACTGGTAATTTAATTTTCAATTTATTTTTTCTTCAACTTCTTA	
10:1 #66	TGTAATACATAGCGGCCGCGCCCAAGGAACCCTTTACTCGGCCA::		7	::::::::::::::::::::::TAATTTAATTTTCAATTTATTTTTTCTTCAACTTCTTA	

FIG. 23

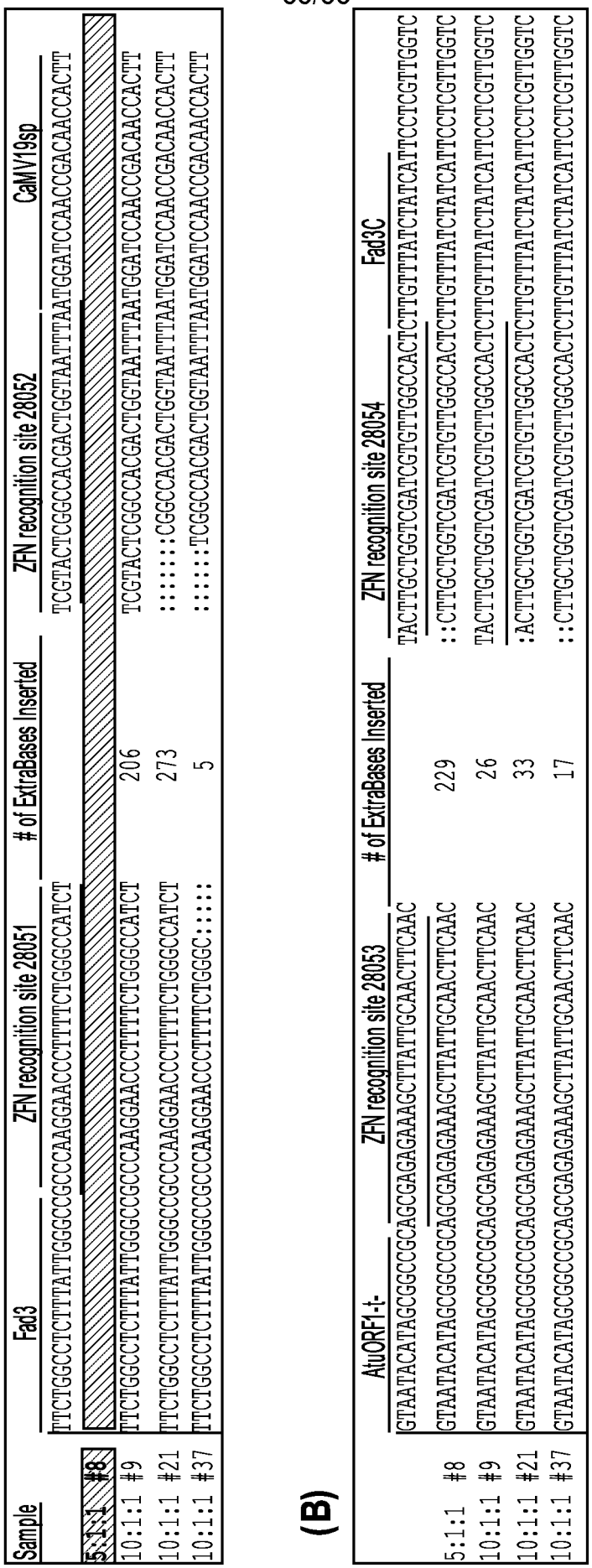
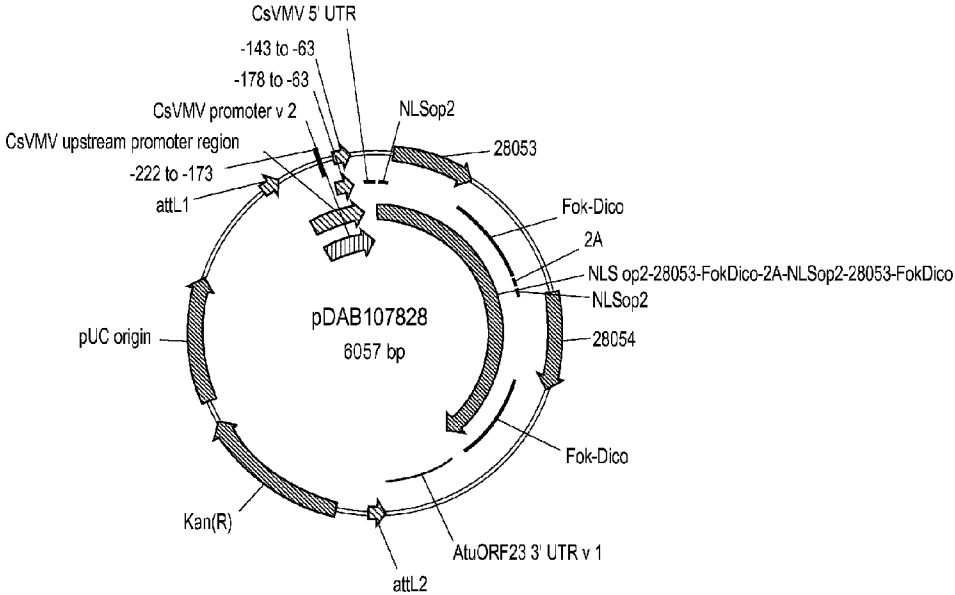

(A)
Sample
Fad3
ZFN recognition site 28051
# of ExtraBases Inserted
ZFN recognition site 28052
CaMV19sp
TTCTGGCCTCTTTATTGGGCCGCCCAAGGAACCCTTTTCTGGGCCATCT
TCGTACTCGGCCACGACTGGTAATTTAATGGATCCAACCGACAACCACTT
5:1:1 #8
10:1:1 #9
206
10:1:1 #21
273
10:1:1 #37
5
(B)
AtuORF1-t-
ZFN recognition site 28053
# of ExtraBases Inserted
ZFN recognition site 28054
Fad3C
GTAATACATAGCGGCCGCAGCGAGAGAAAGCTTATTGCAACTTCAAC
TACTTGCTGGTCGATCGTGTTGGCCACTCTTGTTTATCTATCATTCCTCGTTGGTC
5:1:1 #8
229
::CTTGCTGGTCGATCGTGTTGGCCACTCTTGTTTATCTATCATTCCTCGTTGGTC
10:1:1 #9
26
TACTTGCTGGTCGATCGTGTTGGCCACTCTTGTTTATCTATCATTCCTCGTTGGTC
10:1:1 #21
33
:ACTTGCTGGTCGATCGTGTTGGCCACTCTTGTTTATCTATCATTCCTCGTTGGTC
10:1:1 #37
17
::CTTGCTGGTCGATCGTGTTGGCCACTCTTGTTTATCTATCATTCCTCGTTGGTC

FIG. 24


CsVMV 5' UTR
-143 to -63
-178 to -63
CsVMV promoter v 2
CsVMV upstream promoter region
-222 to -173
attL1
NLSop2
28053
Fok-Dico
2A
NLS op2-28053-FokDico-2A-NLSop2-28053-FokDico
NLSop2
28054
Fok-Dico
AtuORF23 3' UTR v 1
attL2
Kan(R)
pUC origin
pDAB107828
8057 bp