



(86) Date de dépôt PCT/PCT Filing Date: 2014/07/01  
(87) Date publication PCT/PCT Publication Date: 2015/01/15  
(45) Date de délivrance/Issue Date: 2021/01/12  
(85) Entrée phase nationale/National Entry: 2015/12/30  
(86) N° demande PCT/PCT Application No.: US 2014/045115  
(87) N° publication PCT/PCT Publication No.: 2015/006105  
(30) Priorité/Priority: 2013/07/09 (US61/844,248)

(51) Cl.Int./Int.Cl. *C12N 15/82* (2006.01)  
(72) Inventeurs/Inventors:  
SONG, GUO-QING, US;  
WALWORTH, AARON, US;  
ZHAO, DONGYAN, US  
(73) Propriétaire/Owner:  
BOARD OF TRUSTEES OF MICHIGAN STATE  
UNIVERSITY, US  
(74) Agent: RIDOUT & MAYBEE LLP

(54) Titre : PLANTES TRANSGENIQUES COMPORTANT UN DOMAINE K, PROCEDES ET CASSETTES  
D'EXPRESSION S'Y RAPPORTANT  
(54) Title: TRANSGENIC PLANTS PRODUCED WITH A K-DOMAIN, AND METHODS AND EXPRESSION CASSETTES  
RELATED THERETO

(57) **Abrégé/Abstract:**

Methods and compositions of improving plant yield by introducing into a plant the K-domain of a MADS box gene are disclosed. The expression of the K-domain provides plants with altered flower development, plant size and leaf development.

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

(19) World Intellectual Property  
Organization  
International Bureau



(10) International Publication Number  
**WO 2015/006105 A1**

(43) International Publication Date  
15 January 2015 (15.01.2015)

- (51) **International Patent Classification:**  
C12N 15/82 (2006.01)
- (21) **International Application Number:**  
PCT/US2014/045115
- (22) **International Filing Date:**  
1 July 2014 (01.07.2014)
- (25) **Filing Language:** English
- (26) **Publication Language:** English
- (30) **Priority Data:**  
61/844,248 9 July 2013 (09.07.2013) US
- (71) **Applicant:** BOARD OF TRUSTEES OF MICHIGAN STATE UNIVERSITY [US/US]; 450 Administration Building, East Lansing, Michigan 48824-1046 (US).
- (72) **Inventors:** SONG, Guo-Qing; Department of Horticulture, Michigan State University, 1066 Bogue Street, Plant & Soil Sciences Building, East Lansing, Michigan 48823 (US). WALWORTH, Aaron; School of Packaging, Michigan State University, 448 Wilson Road, Rm 130, East Lansing, Michigan 48824 (US). ZHAO, Dongyan; 12010A 31st Avenue NE, Seattle, Washington 98125 (US).
- (74) **Agent:** CLARK, Barbara; Clark IP Law, PLC, 1631 Crestwood Circle, Ames, Iowa 50010 (US).
- (81) **Designated States** (*unless otherwise indicated, for every kind of national protection available*): AE, AG, AL, AM,

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

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

**Declarations under Rule 4.17:**

— *of inventorship (Rule 4.17(iv))*

**Published:**

— *with international search report (Art. 21(3))*

— *with sequence listing part of description (Rule 5.2(a))*

(54) **Title:** TRANSGENIC PLANTS PRODUCED WITH A K-DOMAIN, AND METHODS AND EXPRESSION CASSETTES RELATED THERETO

(57) **Abstract:** Methods and compositions of improving plant yield by introducing into a plant the K-domain of a MADS box gene are disclosed. The expression of the K-domain provides plants with altered flower development, plant size and leaf development.



WO 2015/006105 A1

**TRANSGENIC PLANTS PRODUCED WITH A K-DOMAIN, AND METHODS AND  
EXPRESSION CASSETTES RELATED THERETO**

**[0001]**                      Background

**[0002]**        Food production capacity is faced with an ever-growing number of challenges, including a growing world population, warming globe, and dwindling availability of arable land. In addition, crop yield is a multiple gene-controlled, quantitative trait. The inventors recognize the need for improving crop yield through methods other than traditional breeding.

Summary

**[0003]**        The various embodiments described herein provide methods and compositions for increasing plant yield. According to one embodiment a method of increasing yield in a plant relative to a control plant is provided, comprising increasing expression of a K-domain of a MADS polypeptide or variants thereof. In various embodiments the polynucleotide encoding the K-domain polypeptide is set forth in SEQ ID NO: 1. In one embodiment, the K-domain polypeptide is set forth in SEQ ID NO: 2. In other embodiments, an expression cassette useful for producing a transgenic plant is provided. In still other embodiments, transgenic plants transformed with a K-domain polypeptide are provided.

**[0003a]**                      According to one aspect, we disclose herein a method of producing a transgenic plant comprising:

introducing a biologically active K-domain polypeptide encoded by a nucleotide sequence to produce a transformed cell, wherein the nucleotide sequence is:

a nucleotide sequence set forth in SEQ ID NO: 1;

a nucleotide sequence encoding the polypeptide of SEQ ID NO: 2;

a nucleotide sequence comprising at least 80% sequence identity to the sequence set forth in SEQ ID NO: 1; or

a nucleotide sequence encoding an amino acid sequence having at least 90% sequence identity to the sequence set forth in SEQ ID NO: 2; and

generating the transgenic plant from the transformed cell, wherein the transgenic plant has increased expression of the K-domain polypeptide as compared to a control plant, wherein the % sequence identity is as determinable by a Basic Local Alignment Search Tool (BLAST) program.

**[0003b]** We further disclose an expression cassette comprising a polynucleotide encoding a K-domain polypeptide operably linked to a promoter that drives expression in a plant, wherein the polynucleotide has:

- a nucleotide sequence set forth in SEQ ID NO: 1;
- a nucleotide sequence encoding the polypeptide of SEQ ID NO: 2;
- a nucleotide sequence comprising at least 80% sequence identity to the sequence set forth in SEQ ID NO: 1; or
- a nucleotide sequence encoding an amino acid sequence having at least 90% sequence identity to the sequence set forth in SEQ ID NO: 2, wherein the % sequence identity is as determinable by a BLAST program.

**[0003c]** We further disclose a cell of a transgenic plant, the cell comprising a K-domain polypeptide encoded by a nucleotide sequence set forth in SEQ ID NO: 1, wherein the transgenic plant has increased expression of the K-domain polypeptide compared to a control plant.

#### Brief Description of Drawings

**[0004]** The patent or application file contains at least one drawing executed in color. Copies of this patent or patent application publication with color drawing(s) will be provided by the Office upon request and payment of the necessary fee.

**[0005]** FIG. 1A is a phylogenetic analysis of nucleotide sequences of selected *SOC1*-like proteins according to an embodiment.

**[0006]** FIG. 1B is a multiple sequence alignment of the K-domain of *SOC1*-like proteins according to an embodiment.

**[0007]** FIG. 2 is a digital image of transcriptional analysis of the transgene *35S-VcK-domain* and an endogenous actin gene in one non transgenic (NT) and two randomly selected T0 transgenic tobacco plants according to an embodiment.

- [0008] FIG. 3A is a schematic illustration of the T-DNA region carrying the 35S-VcK-domain according to an embodiment.
- [0009] FIG. 3B is a digital image of T1 transgenic plants compared with a nontransgenic wild type (WT) plant 93 days after seed germination according to an embodiment.
- [0010] FIG. 3C is a graph showing days to emergence of first flowers after the seeds were sown according to an embodiment.
- [0011] FIG. 3D is a graph showing leaf numbers when 50% of flowers in central cluster were open according to an embodiment.
- [0012] FIG. 3E is a graph showing plant height at flowering (cm) according to an embodiment.
- [0013] FIG. 3F is a graph showing number of seed pods in central flower cluster.

#### Submission of Sequence Listing

[0014] The Sequence Listing associated with this application is filed in electronic format via EFS-Web. The name of the text file containing the Sequence Listing is 3000-081W01-SequenceListingST25.txt. The size of the text file is 9 KB, and the text file was created on June 25, 2014.

#### Detailed Description

[0015] In the following detailed description, embodiments are described in sufficient detail to enable those skilled in the art to practice them, and it is to be understood that other embodiments may be utilized and that chemical and procedural changes may be made without departing from the spirit and scope of the present subject matter. The following detailed description is, therefore, not to be taken in a limiting sense, and the scope of embodiments is defined only by the appended claims.

[0016] The term, "comparison window" as used herein refers to sequence relationships between two or more sequences refers to a contiguous and specified segment of a polynucleotide or polypeptide sequence, wherein the sequence in the comparison window may comprise additions or deletions (e.g., gaps) compared to the reference sequence (which does not comprise additions or deletions) for optimal alignment of the two polynucleotides.

[0017] The term, "constitutive promoter" as used herein refers to a promoter that is active in transcription during most, but not all plant growth and development phases and under most environmental conditions and in at least one cell, tissue or organ

[0018] The term "control" or "control plant" or "control plant cell" as used herein refers to a reference point for measuring changes in phenotype of a transgenic plant or plant cell. The

“control” may include, for example (a) a wild-type plant or cell, e.g., of the same genotype as the starting material for the alteration which resulted in the transgenic plant or cell; (b) a plant or plant cell of the same genotype as the starting material but which has been transformed with a null construct (e.g., with a construct which has no known effect on the trait of interest, such as a construct comprising a marker gene); (c) a plant or plant cell not transformed with the nucleotide or gene of interest; (d) a plant or plant cell which is a non-transformed segregant among progeny of a transgenic plant or plant cell; (e) a plant or plant cell genetically identical to the transgenic plant or plant cell but which is not exposed to conditions or stimuli that would induce the expression of the nucleotide or gene of interest; or (f) the transgenic plant or plant cell itself, under conditions in which the nucleotide or gene of interest is not expressed.

[0019] The term “fragment” as used herein refers to a portion of the polynucleotide or a portion of the amino acid sequence. Fragments of a polynucleotide can encode protein fragments that retain the biological activity of the native protein and hence regulate transcription.

[0020] The term “grain” as used herein refers to a mature seed produced by commercial growers for purposes other than growing or reproducing the species.

[0021] The term, “heterologous” as used herein refers to a sequence that originates from a foreign species, or, if from the same species, is modified from its native form in composition, genomic locus or both by deliberate human intervention.

[0022] The term “introducing” as used herein refers to sequences presenting to a plant the polynucleotide or polypeptide in such a manner that the sequence gains access to the interior of a cell of a plant.

[0023] The term “isolated” or “purified” as used herein refers to a polynucleotide or polypeptide, protein, or biologically active portion thereof, removed or free from components that normally accompany or interact with the polynucleotide, polypeptide or protein as found in its naturally occurring environment. Thus, an isolated or purified polynucleotide or protein includes a polynucleotide or polypeptide produced by recombinant techniques, or removed from or free of chemical precursors or other chemicals when chemically synthesized.

[0024] The term “native” as used herein refers to sequences containing a naturally occurring nucleotide sequence or amino acid sequence.

[0025] The term “operably linked” as used herein refers to an expression cassette for plant expression of a functional linkage between two or more elements.

[0026] The term “ortholog” as used herein refers to a polynucleotide or polypeptide gene derived from a common ancestral gene and which are found in different species as a result of speciation.

[0027] The term “overexpression” or “increased expression” as used herein refers to any form of expression that is additional to the original, wild-type expression level.

[0028] The term “percentage of sequence identity” as used herein refers to a polynucleotide’s value determined by comparing two optimally aligned sequences over a comparison window, wherein the portion of the polynucleotide sequence in the comparison window may include additions or deletions (e.g., gaps) as compared to the reference sequence (which does not include additions or deletions) for optimal alignment of the two sequences.

[0029] The term “plant” as used herein refers to plant cells, plant protoplasts, plant cell tissue cultures from which plants can be regenerated, plant calli, plant clumps, and plant cells that are intact in plants or parts of plants such as embryos, pollen, ovules, seeds, leaves, flowers, branches, fruit, kernels, ears, cobs, husks, stalks, roots, root tips, anthers, and the like.

[0030] The term “polynucleotide” as used herein refers to polynucleotides comprising DNA. The term includes ribonucleotides and combinations of ribonucleotides and deoxyribonucleotides. Such deoxyribonucleotides and ribonucleotides include both naturally occurring molecules and synthetic analogues. The polynucleotides also encompass all forms of sequences including, but not limited to, single-stranded forms, double-stranded forms, hairpins, stem-and-loop structures, and the like.

[0031] The term “reference sequence” as used herein refers to sequence relationships between two or more sequences refers to a subset or the entirety of a specified sequence as a basis for sequence comparison; for example, as a segment of a full-length cDNA or gene sequence, or the complete cDNA or gene sequence.

[0032] The terms “sequence identity” or “identity” as used herein refers to relationships between two or more polynucleotides or polypeptides refer to the residues in the two sequences that are the same when aligned for maximum correspondence over a specified comparison window.

[0033] The term, “stable transformation” as used herein refers to nucleotides introduced into plants that integrate into the genome of the plant and is capable of being inherited by the progeny thereof.

[0034] The terms “stringent conditions” or “stringent hybridization conditions” as used herein refers to conditions under which a probe (nucleotide sequences) will hybridize to its target sequence to a detectably greater degree than to other sequences (e.g., at least 2-fold over background).

[0035] The term “tissue-specific promoter” as used herein refers to a promoter that is active in transcription in certain organs or tissues.

[0036] The term “trait” as used herein refers to a phenotype derived from a particular sequence or groups of sequences.

[0037] The term “transgenic” as used herein refers to a plant or plant part alteration, such as transformation or introduction of a polynucleotide or polypeptide, has occurred, or is a plant or plant cell which is descended from a plant or cell so altered and which has the alteration.

[0038] The terms “transient transformation” as used herein refers to a polynucleotide, polypeptide or protein introduced into a plant and does not integrate into the genome of the plant.

[0039] The term “variants” as used herein refers to sequences having substantially similar sequences as compared to a first sequence. A variant can include, but is not limited to, a deletion, insertion, substitution or combination thereof of one or more nucleotides or amino acids at one or more internal sites within the native or reference polynucleotide or polypeptide respectively.

[0040] The term “yield” as used herein refers to a crop amount that was harvested per unit of land area or the amount of seeds generated from the plant itself or a combination thereof.

[0041] Disclosed are methods and compositions of increasing plant yield by expressing the K-domain portion of a gene, such as a MADS-box gene. The K-domain of a MADS-box gene when expressed in plants can increase plant yield compared to a plant not expressing or expressing at lower levels of the K-domain. The molecules may be obtained by any convenient method, whether isolated from a plant or synthetically produced, for example. In one embodiment, methods and compositions are provided to accelerate floral organ development, increase flower numbers, reduce plant size, and reduce leaf numbers, or combinations thereof, without negatively affecting seed or fruit production. The various embodiments provide for methods and compositions which result in improved plant or crop yield by increasing plant density in the field, by reducing plant growth period, or both.

[0042] The various embodiments provide for compositions that include the K-domain polynucleotides and polypeptides and variants thereof, or fragment of the MADS-box genes that are involved in accelerating flowering, reducing plant size or leaf numbers or combinations thereof without affecting seed or fruit development. In one embodiment, the *SUPPRESSOR OF OVEREXPRESSION OF CONSTANS 1 (SOC1)* and *SOC1*-likes encodes a plant protein with MADS domain and K-domain. MADS-domain transcription factors play essential roles in the evolution of flowering plants through determining floral architecture and regulating flower development. MADS box genes contain a conserved MADS box motif and are often classified into type I and type II subfamilies, in which type II MADS box genes often have conserved K-, I-, and C (Carboxy terminal)-domains. Based on the variations in their K (Keratin)-domains and I (Intervening) domains, the type II MADS box genes are further divided into MIKC<sup>nc</sup> and



WO 2015/006105

PCT/US2014/045115

MIKC<sup>c</sup> (classical MIKC) subgroups, where the latter often have a shorter K domain than that of the MIKC<sup>nc</sup> and play specific roles in the ABC model of floral development. MIKC<sup>c</sup>-subgroup consists of 13 major gene clades, of which 12 have been identified and extensively studied in *Arabidopsis thaliana*. Genome-wide analyses of the MIKC<sup>c</sup>-type MADS box genes have been reported in several plant species, including *Arabidopsis*, tomato (*Solanum lycopersicon*), poplar (*Populus trichocarpa*), rice (*Oryza sativa*), grapevine (*Vitis vinifera*), and cucumber (*Cucumis sativus*). Two gene clades, *FLOWERING LOCUS C* (*FLC*) and TM3/K-domain, are major components in the flowering-time pathway. The *FLC* clade genes, including *FLC* and *FLC*-like genes, are central regulators of vernalization-mediated flowering. *SOC1* and *SOC1*-like genes in the TM3/K-domain clade are major flowering pathway integrators, harmonizing flowering signals from multiple pathways. Although there have been many reports on overexpression of intact MIKC<sup>c</sup>-type MADS box genes, little has been done on the separate domains of these MADS proteins.

[0043] In one embodiment, an isolated K-domain of a MADS-box gene (e.g., *SOC1*-like gene) is disclosed. In one embodiment, the MADS-box gene is from a *Vaccinium corymbosum* (blueberry) (*VcK-domain*). Overexpression of *VcK-domain* alone can promote early flowering, reduced plant size or leaf number or combinations thereof without reducing or affecting seed or fruit production. The *K-domain* nucleotide sequence is shown as SEQ ID NO: 1, the *K-domain* polypeptide is shown as SEQ ID NO: 2 and the full length MADS-box gene of the *SOC1* gene is shown as SEQ ID NO: 3. In other embodiments, other K-domains from MADS-box genes involved in flowering pathways may be used. In various embodiments, *LEAFY* (*LFY*) (Accession: AF466801.1), *TERMINAL FLOWER 1* (*TFL1*) (Accession: AF466804.1), *APETALA1* (*API*) (Accession: NM\_105581.2), *LFY*-like (contig12297), *TFL1*-like (contig02213), and *API*-like may be used.

[0044] Fragments that are useful as hybridization probes generally do not encode fragment proteins retaining biological activity. Thus, in various embodiments, fragments of a nucleotide sequence may range from at least about 18 nucleotides, about 20 nucleotides, about 50 nucleotides, about 100 nucleotides and up to the full-length polynucleotide encoding the proteins. A fragment of a MADS-box polypeptide can encode a wide number of contiguous amino acids, including, but not limited to, at least 15, 25, 30, 50, 100, 150, 200, 250, 300 or more contiguous amino acids, including any value or range therebetween, or otherwise up to the total number of amino acids present in a full-length MADS domain, or K-domain protein (e.g., SEQ ID NO: 2). In various embodiments, a K-domain polynucleotide or fragments of a MADS domain that includes a K-domain polynucleotide that are useful as hybridization probes, PCR primers, generally need not encode a biologically active portion of a K-domain protein.

WO 2015/006105

PCT/US2014/045115

**[0045]** In some embodiments, a biologically active portion of a polypeptide having a K-domain, can be prepared by isolating a portion of a K-domain polynucleotide, expressing the encoded portion of the K-domain protein (e.g., by recombinant expression in vitro), and assessing the activity of the encoded portion of the K-domain. In various embodiments, polynucleotides that are fragments of a MADS-box gene or a polynucleotide sequence having a K domain have at least 16, 20, 50, 75, 100, 150, 200, 250, 300, 350, 400, 450, 500, contiguous nucleotides or up to the number of nucleotides present in a full-length K-domain (e.g., SEQ ID NO: 1).

**[0046]** Naturally occurring allelic variants can be identified with the use of well-known molecular biology techniques, as, in some embodiments, with polymerase chain reaction (PCR) and hybridization techniques as outlined below. In various embodiments, variant polynucleotides also include synthetically derived polynucleotides, such as those generated, by using site-directed mutagenesis but which still encode a polypeptide having a K-domain that is capable of regulating flowering time, flower numbers, plant size, leaf number or combinations thereof but without affecting seed or fruit development. In various embodiments, variants of a particular polynucleotide may have at least about 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or more sequence identity to that particular polynucleotide as determined by sequence alignment programs and parameters described elsewhere herein.

**[0047]** Variants of a particular polynucleotide (e.g., a reference polynucleotide) can also be evaluated by comparison of the percent sequence identity between the polypeptide encoded by a variant polynucleotide and the polypeptide encoded by the reference polynucleotide. In various embodiments an isolated polynucleotide that encodes a polypeptide with a given percent sequence identity to the polypeptide of SEQ ID NO:2 is disclosed. Percent sequence identity between any two polypeptides can be calculated using sequence alignment programs and parameters described elsewhere herein. In some embodiments, the comparison window is at least 20 contiguous nucleotides in length, and optionally can be 30, 40, 50, 100, including any value therebetween, or longer. Those skilled in the art understand that to avoid a high similarity to a reference sequence due to inclusion of gaps in the polynucleotide sequence a gap penalty is typically introduced and is subtracted from the number of matches. Where any given pair of polynucleotides is evaluated by comparison of the percent sequence identity shared by the two polypeptides they encode, the percent sequence identity between the two encoded polypeptides in some embodiments are at least about 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or more sequence identity.

**[0048]** Variant proteins encompassed are biologically active, that is, they continue to possess the desired biological activity of the native protein, namely they regulate transcription and result in phenotypes affecting flowering time, plant size and leaf number without affecting seed or fruit production. In various embodiments, such variants may result from genetic polymorphism or from human manipulation. In various embodiments, biologically active variants of a K-domain will have at least about 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% including any range therebetween or more sequence identity to the amino acid SEQ ID NO: 2 *VcK-domain* as determined by sequence alignment programs and parameters described elsewhere herein. In various embodiments, a biologically active variant of a K-domain may differ from that protein by as few as 1-15 amino acid residues, as few as 1-10, such as 6-10, including any range therebetween, to as few as 5, as few as 4, 3, 2 or even 1 amino acid residue.

**[0049]** In various embodiments, the polynucleotides may be altered in various ways including amino acid substitutions, deletions, truncations, and insertions using a variety of methods generally known in the art. In various embodiments, amino acid sequence variants of K-domain or fragments of the MADS-box protein can be prepared by mutations in the DNA by a variety of known methods for mutagenesis and polynucleotide alterations. See, for example, Kunkel (1985) *Proc Natl Acad Sci USA* 82:488-492; Kunkel, et al., (1987) *Methods in Enzymol* 154:367-382; U.S. Pat. No. 4,873,192; Walker and Gaastra, eds. (1983) *Techniques in Molecular Biology* (MacMillan Publishing Company, New York) and the references cited therein. Guidance as to appropriate amino acid substitutions that do not affect biological activity of the protein of interest may be found in the model of Dayhoff, et al., (1978) *Atlas of Protein Sequence and Structure* (Natl. Biomed. Res. Found. Washington, D.C.). In some embodiments, conservative substitutions, such as exchanging one amino acid with another having similar properties are disclosed.

**[0050]** In various embodiments, genes and polynucleotides include both the naturally occurring sequences as well as mutant forms. Likewise, in other embodiments the proteins encompass naturally occurring proteins as well as variations and modified forms thereof. Such variants will continue to possess the desired activity (e.g., transcription binding that when overexpressed may alter flower development, plant size, leaf number or combinations thereof without affecting seed or fruit development). In various embodiments, the mutations that may be made in the DNA encoding a variant do not place the sequence out of reading frame and do not create complementary regions that could produce secondary mRNA structure. See, DeBoer et al., EP Patent Application Publication Number 75,444A (1983).

**[0051]** In various embodiments, the deletions, insertions, and substitutions of the protein sequences encompassed herein should still retain the characteristics of the protein activity. However, when it is difficult to predict the exact effect of the substitution, deletion, or insertion in advance of doing so, one skilled in the art will appreciate that the effect will be evaluated by routine screening assays. In various embodiments, various methods can be used to assay for the presence of the K-domain, including, directly monitoring the level of expression of a target gene at the nucleotide or polypeptide level. Methods for such an analysis are known and include, for example, Northern blots, protection assays, Western blots, enzymatic or colorimetric assays. In various embodiments, methods to assay for a modulation of transcriptional activity can include monitoring for an alteration in the phenotype of the plant. In some embodiments, modulating the level of the K-domain polypeptide can result in alterations in flower formation, timing, plant height and leaf number.

**[0052]** In various embodiments, variant polynucleotides and proteins also encompass sequences and proteins derived from a mutagenic and recombinogenic procedure such as DNA shuffling. With such a procedure, one or more different K-domain coding sequences can be manipulated to create a new K-domain sequence possessing the desired properties. In this manner, libraries of recombinant polynucleotides are generated from a population of related sequence polynucleotides having sequence regions that have substantial sequence identity and can be homologous recombined in vitro or in vivo. In various embodiments, using this approach, sequence motifs encoding a domain of interest may be shuffled between the K-domain and other K-domains of other known MADS-box genes to obtain a new gene coding for a protein with an improved property of interest, such as an early flowering time, reduced plant height or reduced leaf number or combinations thereof. Strategies for such DNA shuffling are known in the art. See, for example, Stemmer (1994) *Proc Natl Acad Sci USA* 91:10747-10751; Stemmer (1994) *Nature* 370:389-391; Cramer, et al, (1997) *Nature Biotech* 15:436-438; Moore, et al., (1997) *J Mol Biol* 272:336-347; Zhang, et al., (1997) *Proc Natl Acad Sci USA* 94:4504-4509; Cramer, et al., (1998) *Nature* 391:288-291; and U.S. Pat. Nos. 5,605,793 and 5,837,458.

**[0053]** In some embodiments, the polynucleotides can be used to isolate corresponding sequences from other organisms, other plants, or other woody plants. In this manner, methods such as PCR, hybridization, and the like can be used to identify such sequences based on their sequence homology to the sequences set forth herein. Sequences isolated based on their sequence identity to the entire K-domain sequences, set forth herein or to variants and fragments thereof are encompassed herein. In various embodiments, such sequences include sequences that are orthologs of the disclosed sequences. Functions of orthologs are often highly conserved among species. In various embodiments, genes found in different species are considered

WO 2015/006105

PCT/US2014/045115

orthologs when their nucleotide sequences and/or their encoded protein sequences share at least 60%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, including any range or value therebetween or greater sequence identity. Thus, isolated polynucleotides that can cause early flowering, reduced plant size or fewer leaves without affecting seed or fruit production per plant and that encodes for a K-domain and which hybridize under stringent conditions to the disclosed K-domain sequences, or to variants or fragments thereof, are encompassed herein.

[0054] In a PCR approach, oligonucleotide primers can be designed for use in PCR reactions to amplify corresponding DNA sequences from cDNA or genomic DNA extracted from any plant of interest. Methods for designing PCR primers and PCR cloning are generally known in the art and are disclosed in Sambrook, et al., (1989) *Molecular Cloning: A Laboratory Manual* (2d ed., Cold Spring Harbor Laboratory Press, Plainview, N.Y.). See also, Innis, et al., eds. (1990) *PCR Protocols: A Guide to Methods and Applications* (Academic Press, New York); Innis and Gelfand, eds. (1995) *PCR Strategies* (Academic Press, New York); and Innis and Gelfand, eds. (1999) *PCR Methods Manual* (Academic Press, New York). In some embodiments, methods of PCR include, but are not limited to, methods using paired primers, nested primers, single specific primers, degenerate primers, gene-specific primers, vector-specific primers, partially-mismatched primers, and the like.

[0055] In hybridization techniques, all or part of a known polynucleotide is used as a probe that selectively hybridizes to other corresponding polynucleotides present in a population of cloned genomic DNA fragments or cDNA fragments (e.g., genomic or cDNA libraries) from a chosen organism. In various embodiments, the hybridization probes may be genomic DNA fragments, cDNA fragments, RNA fragments, or other oligonucleotides. In still other embodiments, the hybridization probes may be labeled with a detectable group such as  $^{32}\text{P}$ , or any other detectable marker. In various embodiments, probes for hybridization can be made by labeling synthetic oligonucleotides based on the K-domain polynucleotide. Methods for preparation of probes for hybridization and for construction of cDNA and genomic libraries are generally known in the art and are disclosed, for example in Sambrook, et al., (1989) *Molecular Cloning: A Laboratory Manual* (2d ed., Cold Spring Harbor Laboratory Press, Plainview, N.Y.).

[0056] In various embodiments, the entire MADS-box domain polynucleotide or one or more portions thereof or a polynucleotide encoding a K domain disclosed herein, may be used as a primer capable of specifically hybridizing to corresponding K-domain polynucleotide and messenger RNAs. To achieve specific hybridization under a variety of conditions, such probes include sequences that are unique among K-domain polynucleotide sequences. In various embodiments, the probes include at least about 10 nucleotides in length, at least about 20

WO 2015/006105

PCT/US2014/045115

nucleotides in length. Such primer pairs may be used to amplify corresponding K-domain polynucleotide from a chosen plant by PCR. This technique may be used to isolate additional coding sequences from a desired plant or as a diagnostic assay to determine the presence of coding sequences in a plant. Hybridization techniques include hybridization screening of plated DNA libraries (either plaques or colonies; see, for example, Sambrook, et al., (1989) Molecular Cloning: A Laboratory Manual (2d ed., Cold Spring Harbor Laboratory Press, Plainview, N.Y.).

[0057] Hybridization of such sequences (e.g., Southern blot and Northern blot) may be carried out under stringent conditions. Stringent conditions are sequence-dependent and will be different in different circumstances. By controlling the stringency of the hybridization, washing conditions or both, target sequences that are 100% complementary to the probe can be identified (homologous probing). In other cases, stringency conditions can be adjusted to allow some mismatching in sequences so that lower degrees of similarity are detected (heterologous probing). In various embodiments, a probe is less than about 1000 nucleotides in length, less than 500 nucleotides in length, including any range or value therebetween.

[0058] In various embodiments, stringent conditions will be those in which the salt concentration is less than about 1.5 M Na ion, about 0.01 to 1.0 M Na ion concentration (or other salts) at pH 7.0 to 8.3 including any range or value therebetween for both Na concentration and pH, and the temperature is at least 20-35°C for short probes (e.g., 10 to 50 nucleotides) and at least about 50 to 60°C for long probes (e.g., greater than 50 nucleotides). Stringent conditions may also be achieved with the addition of destabilizing agents such as formamide. In various embodiments, low stringency conditions include hybridization with a buffer solution of 30 to 35% formamide, 1 M NaCl, 1% SDS (sodium dodecyl sulphate) at 37°C, and a wash in 1X to 2X SSC (20 X SSC=3.0 M NaCl/0.3 M trisodium citrate) at 50 to 55°C. Exemplary moderate stringency conditions include hybridization in 40 to 45% formamide, 1.0 M NaCl, 1% SDS at 37°C, and a wash in 0.5X to 1X SSC at 55 to 60°C. In other embodiments high stringency conditions include hybridization in 50% formamide, 1 M NaCl, 1% SDS at 37°C, and a wash in 0.1 X SSC at 60 to 65°C. In various embodiments, wash buffers may include about 0.1% to about 1% SDS. In various embodiments, duration of hybridization is less than about 24 hours, about 4 to about 12 hours. The duration of the wash time will be at least a length of time sufficient to reach equilibrium.

[0059] Specificity is typically the function of post-hybridization washes, the useful factors being the ionic strength and temperature of the final wash solution. For DNA-DNA hybrids, the  $T_m$  can be approximated from the equation of Meinkoth and Wahl (1984) Anal. Biochem. 138:267-284:  $T_m = 81.5^\circ\text{C} + 16.6 (\log M) + 0.41 (\% \text{ GC}) - 0.61 (\% \text{ form}) - 500/L$ ; where M is the molarity of monovalent cations, % GC is the percentage of guanosine and cytosine nucleotides in

the DNA, % form is the percentage of formamide in the hybridization solution, and L is the length of the hybrid in base pairs. The  $T_m$  is the temperature (under defined ionic strength and pH) at which 50% of a complementary target sequence hybridizes to a perfectly matched probe.  $T_m$  is reduced by about 1°C for each 1% of mismatching; thus,  $T_m$ , hybridization, and/or wash conditions can be adjusted to hybridize to sequences of the desired identity. In various embodiments, if sequences with  $\geq 90\%$  identity are sought, the  $T_m$  can be decreased by about 10°C. In various embodiments, stringent conditions are selected to be about 5°C lower than the thermal melting point ( $T_m$ ) for the specific sequence and its complement at a defined ionic strength and pH. In various embodiments, severely stringent conditions can utilize a hybridization and/or wash at 1, 2, 3 or 4°C including any range or value therebetween lower than the thermal melting point ( $T_m$ ); In various embodiments, moderately stringent conditions can utilize a hybridization and/or wash at 6, 7, 8, 9 or 10°C lower than the thermal melting point ( $T_m$ ); low stringency conditions can utilize a hybridization and/or wash at 11, 12, 13, 14, 15 or 20°C including any range or value therebetween lower than the thermal melting point ( $T_m$ ). Using the equation, hybridization and wash compositions, and desired  $T_m$ , those of ordinary skill will understand that variations in the stringency of hybridization and/or wash solutions are inherently described. If the desired degree of mismatching results in a  $T_m$  of less than 45°C (aqueous solution) or 32°C (formamide solution), in various embodiments to increase the SSC concentration so that a higher temperature can be used. An extensive guide to the hybridization of nucleic acids may be found in Tijssen (1993) *Laboratory Techniques in Biochemistry and Molecular Biology-Hybridization with Nucleic Acid Probes*, Part I, Chapter 2 (Elsevier, New York); and Ausubel, et al., eds. (1995) *Current Protocols in Molecular Biology*, Chapter 2 (Greene Publishing and Wiley-Interscience, New York). See, Sambrook, et al., (1989) *Molecular Cloning: A Laboratory Manual* (2d ed., Cold Spring Harbor Laboratory Press, Plainview, N.Y.).

**[0060]** Methods of alignment of nucleotide (DNA and RNA) or protein sequences for comparison are well known in the art. Thus, the determination of percent sequence identity between any two sequences can be accomplished using a mathematical algorithm. In various embodiments, a portion of the polynucleotide sequence in a comparison window includes additions or deletions (e.g., gaps) as compared to the reference sequence (which does not include additions or deletions) for optimal alignment of the two sequences.

**[0061]** In various embodiments, the sequence alignment is by Maximum Likelihood method based on the Tamura-Nei model using the MEGA 5 program. In various embodiments, mathematical algorithms are the algorithm of Myers and Miller (1988) *CABIOS* 4:11-17; the local alignment algorithm of Smith, et al., (1981) *Adv Appl Math* 2:482; the global alignment algorithm of Needleman and Wunsch (1970) *J Mol Biol* 48:443-453; the search-for-local

alignment method of Pearson and Lipman (1988) *Proc Natl Acad Sci* 85:2444-2448; the algorithm of Karlin and Altschul (1990) *Proc Natl Acad Sci USA* 87:2264, modified as in Karlin and Altschul (1993) *Proc Natl Acad Sci USA* 90:5873-5877.

[0062] Computer implementations of these mathematical algorithms can be utilized for comparison of sequences to determine sequence identity. Such implementations include, but are not limited to: CLUSTAL in the PC/Gene program (available from Intelligenetics, Mountain View, Calif.); the ALIGN program (Version 2.0) and GAP, BESTFIT, BLAST, FASTA, and TFASTA in the GCG Wisconsin Genetics Software Package, Version 10 (available from Accelrys Inc., 9685 Scranton Road, San Diego, Calif., USA). Alignments using these programs can be performed using the default parameters. The CLUSTAL program is well described by Higgins, et al., (1988) *Gene* 73:237-244 (1988); Higgins, et al., (1989) *CABIOS* 5:151-153; Corpet, et al., (1988) *Nucleic Acids Res* 16:10881-90; Huang, et al., (1992) *CABIOS* 8:155-65; and Pearson, et al., (1994) *Meth Mol Biol* 24:307-331. The ALIGN program is based on the algorithm of Myers and Miller (1988) *supra*. A PAM120 weight residue table, a gap length penalty of 12, and a gap penalty of 4 can be used with the ALIGN program when comparing amino acid sequences. The BLAST programs of Altschul, et al., (1990) *J Mol Biol* 215:403 are based on the algorithm of Karlin and Altschul (1990) *supra*. BLAST nucleotide searches can be performed with the BLASTN program, score=100, word length=12, to obtain nucleotide sequences homologous to a nucleotide sequence encoding a protein. BLAST protein searches can be performed with the BLASTX program, score=50, word length=3, to obtain amino acid sequences homologous to a protein or polypeptide.

[0063] To obtain gapped alignments for comparison purposes, Gapped BLAST (in BLAST 2.0) can be utilized as described in Altschul, et al., (1997) *Nucleic Acids Res* 25:3389. In other cases, PSI-BLAST (in BLAST 2.0) can be used to perform an iterated search that detects distant relationships between molecules. See, Altschul, et al., (1997) *supra*. When utilizing BLAST, Gapped BLAST, PSI-BLAST, the default parameters of the respective programs (e.g., BLASTN for nucleotide sequences, BLASTX for proteins) can be used. See, [www.ncbi.nlm.nih.gov](http://www.ncbi.nlm.nih.gov). Alignment may also be performed manually by inspection.

[0064] Unless otherwise stated, pairwise sequence identity/similarity values provided herein refer to the value obtained using the GAP program, which is part of GCG 11.0 using the following parameters: % identity and % similarity using gap creation penalty of 10, gap extension penalty of 1; or any equivalent program thereof. By "equivalent program" is intended any sequence comparison program that, for any two sequences in question, generates an alignment having identical nucleotide or amino acid residue matches and an identical percent sequence identity when compared to the corresponding alignment generated by GAP.



WO 2015/006105

PCT/US2014/045115

[0065] GAP uses the algorithm of Needleman and Wunsch (1970) J Mol Biol 48:443-453, to find the alignment of two complete sequences that maximizes the number of matches and minimizes the number of gaps. GAP considers all possible alignments and gap positions and creates the alignment with the largest number of matched bases and the fewest gaps. It allows for the provision of a gap creation penalty and a gap extension penalty in units of matched bases. GAP must make a profit of gap creation penalty number of matches for each gap it inserts. If a gap extension penalty greater than zero is chosen, GAP must, in addition, make a profit for each gap inserted of the length of the gap times the gap extension penalty. Default gap creation penalty values and gap extension penalty values in Version 10 of the GCG Wisconsin Genetics Software Package for protein sequences are 8 and 2, respectively. In various embodiments, for nucleotide sequences the default gap creation penalty is 50 while the default gap extension penalty is 3. The gap creation and gap extension penalties can be expressed as an integer selected from the group of integers that include from 0 to 200. Thus, in various embodiments, the gap creation and gap extension penalties can be 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65 or greater.

[0066] GAP presents one member of the family of best alignments. There may be many members of this family, but no other member has a better quality. GAP displays four figures of merit for alignments: Quality, Ratio, Identity, and Similarity. The Quality is the metric maximized in order to align the sequences. Ratio is the quality divided by the number of bases in the shorter segment. Percent Identity is the percent of the symbols that actually match. Percent Similarity is the percent of the symbols that are similar. Symbols that are across from gaps are ignored. A similarity is scored when the scoring matrix value for a pair of symbols is greater than or equal to 0.5, the similarity threshold. The scoring matrix used in Version 11.0 of the GCG Wisconsin Genetics Software Package is BLOSUM62 (see, Henikoff and Henikoff (1989) Proc Natl Acad Sci USA 89:10915).

[0067] Unless otherwise stated, percent identity and/or percent similarity values among multiple sequences provided herein refer to the value obtained using ClustalW2 (<http://www.ebi.ac.uk/Tools/msa/clustalw2/>) using the following parameters: for nucleotide sequences using gap open of 10, gap extension of 0.2, gap distances of 5, the IUB DNA weight matrix, and the clustering method of Neighbor-joining; for amino acid sequences using gap open of 10, gap extension of 0.2, gap distances of 5, the Gonnet protein weight matrix, and the clustering method of Neighbor-joining.

[0068] ClustalW2 is a web-based multiple sequence alignment program for DNA and protein (<http://www.ebi.ac.uk/Tools/msa/clustalw2/help/index.html>) which will generate six result files, including input sequences, tool output, alignment in CLUSTAL format, guide tree, phylogenetic

WO 2015/006105

PCT/US2014/045115

tree, and percent identity matrix. To perform the alignment, in various embodiments, two sets of parameters can be set by users, i.e., pairwise alignment options and multiple sequence alignment options. Parameters selected for pairwise alignment will be used to establish the guided tree for the multiple sequence alignment whereas the latter will calculate the best matches for the input sequences and line them up so that identities, similarities, and differences can be visualized directly.

[0069] When percentage of sequence identity is used in reference to proteins it is recognized that residue positions which are not identical often differ by conservative amino acid substitutions, where amino acid residues are substituted for other amino acid residues with similar chemical properties (e.g., charge or hydrophobicity) and therefore do not change the functional properties of the molecule. When sequences differ in conservative substitutions, the percent sequence identity may be adjusted upwards to correct for the conservative nature of the substitution. Sequences that differ by such conservative substitutions are said to have “sequence similarity” or “similarity”. Means for making this adjustment are well known to those of skill in the art. In various embodiments, this involves scoring a conservative substitution as a partial rather than a full mismatch, thereby increasing the percentage sequence identity. Thus, in various embodiments, where an identical amino acid is given a score of 1 and a non-conservative substitution is given a score of zero, a conservative substitution is given a score between zero and 1. The scoring of conservative substitutions is calculated, e.g., as implemented in the program PC/GENE (Intelligenetics, Mountain View, Calif.). The percentage is calculated by determining the number of positions at which the identical nucleic acid base 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 window of comparison, and multiplying the result by 100 to yield the percentage of sequence identity.

[0070] In various embodiments, the disclosed sequences may be introduced into plants to alter levels of a K-domain sequence. In other embodiments, plants may have increased expression levels of a K-domain sequence. In various embodiments, the plants have stably incorporated into their genome at least one heterologous polynucleotide encoding a K-domain polypeptide. In various embodiments, sequences that can be used to increase expression of a K-domain polypeptide include, but are not limited to, the sequence set forth in SEQ ID NO: 1 or variants or fragments thereof. Progeny, variants, and mutants of the regenerated plants are also included within the scope, provided that these parts include the introduced or heterologous polynucleotides disclosed herein.

[0071] In various embodiments, the disclosed plants can have an altered phenotype. In some embodiments the altered phenotype includes altered flower organ development, flower numbers,

WO 2015/006105

PCT/US2014/045115

leaf numbers and formation, plant size or combinations thereof, but no negative effect on seed or fruit production per plant. In various embodiments, this thereby results in improved yield by increasing plant density in the field and reducing the plant growing period.

[0072] In various embodiments, the plant yield is increased by about 5%, 7%, 8%, 9%, or about 10% to 20 % or about 30% to about 50% compared to a control plant.

In some embodiments, an increase in yield is seen as an increase in bushels/acre yield. In various embodiments, yield is increased because of phenotypic changes. In various embodiments, the phenotypic changes include change in flowering time, plant height, flower number or a combination thereof when overexpressing *VcK*. Depending on the plant species, the phenotypes can vary. In various embodiments, ectopic expression of the *VcK* shortened plant cycles (e.g. 2-3 weeks) and reduced plant sizes. In various embodiments, ectopic overexpression of the *VcK* promoted more flowers and branch formation. In still other embodiments, overexpressing the *VcK* increased the number of flowering buds, reduced the amount of chilling requirement for plant flowering. In some embodiments, yield may be increased because of an increase in seed weight, seed size or seed shape. In other embodiments, yield may be increased because of the number of seed (in pods).

[0073] Various transformation techniques for a variety of plant species, including, but not limited to, monocots and dicots may be used. In various embodiments, plants include, but are not limited to, corn (*Zea mays*), Brassica sp. (e.g., *B. napus*, *B. rapa*, *B. juncea*), particularly those Brassica species useful as sources of seed oil, alfalfa (*Medicago sativa*), rice (*Oryza sativa*), rye (*Secale cereale*), sorghum (*Sorghum bicolor*, *Sorghum vulgare*), millet (e.g., pearl millet (*Pennisetum glaucum*), proso millet (*Panicum miliaceum*), foxtail millet (*Setaria italica*), finger millet (*Eleusine coracana*)), sunflower (*Helianthus annuus*), safflower (*Carthamus tinctorius*), wheat (*Triticum aestivum*), soybean (*Glycine max*), tobacco (*Nicotiana tabacum*), potato (*Solanum tuberosum*), peanuts (*Arachis hypogaea*), cotton (*Gossypium barbadense*, *Gossypium hirsutum*), sweet potato (*Ipomoea batatas*), cassava (*Manihot esculenta*), coffee (*Coffea spp.*), coconut (*Cocos nucifera*), pineapple (*Ananas comosus*), citrus trees (*Citrus spp.*), cocoa (*Theobroma cacao*), tea (*Camellia sinensis*), banana (*Musa spp.*), avocado (*Persea americana*), fig (*Ficus casica*), guava (*Psidium guajava*), mango (*Mangifera indica*), olive (*Olea europaea*), papaya (*Carica papaya*), cashew (*Anacardium occidentale*), macadamia (*Macadamia integrifolia*), almond (*Prunus amygdalus*), sugar beets (*Beta vulgaris*), sugarcane (*Saccharum spp.*), oats, barley, vegetables, ornamentals, and conifers.

[0074] Vegetables include tomatoes (*Lycopersicon esculentum*), lettuce (e.g., *Lactuca sativa*), green beans (*Phaseolus vulgaris*), lima beans (*Phaseolus limensis*), peas (*Lathyrus spp.*), and members of the genus *Cucumis* such as cucumber (*C. sativus*), cantaloupe (*C.*

*cantalupensis*), and musk melon (*C. melo*). Ornamentals include azalea (*Rhododendron spp.*), hydrangea (*Macrophylla hydrangea*), hibiscus (*Hibiscus rosasanensis*), roses (*Rosa spp.*), tulips (*Tulipa spp.*), daffodils (*Narcissus spp.*), petunias (*Petunia hybrida*), carnation (*Dianthus caryophyllus*), poinsettia (*Euphorbia pulcherrima*), and chrysanthemum.

Conifers that may be employed in practicing the present invention include, for example, pines such as loblolly pine (*Pinus taeda*), slash pine (*Pinus effiotii*), ponderosa pine (*Pinus ponderosa*), lodgepole pine (*Pinus contorta*), and Monterey pine (*Pinus radiata*); Douglas-fir (*Pseudotsuga menziesii*); Western hemlock (*Tsuga canadensis*); Sitka spruce (*Picea glauca*); redwood (*Sequoia sempervirens*); true firs such as silver fir (*Abies amabilis*) and balsam fir (*Abies balsamea*); and cedars such as Western red cedar (*Thuja plicata*) and Alaska yellow-cedar (*Chamaecyparis nootkatensis*).

**[0075]** In other embodiments, other plants of interest include grain plants that provide seeds of interest, oil-seed plants, and leguminous plants. Seeds of interest include grain seeds, such as corn, wheat, barley, rice, sorghum, rye, and the like. Oil-seed plants include cotton, soybean, safflower, sunflower, Brassica, maize, alfalfa, palm, coconut, and the like. Leguminous plants include beans and peas. Beans include guar, locust bean, fenugreek, soybean, garden beans, cowpea, mungbean, lima bean, fava bean, lentils, chickpea, and the like.

**[0076]** In still other embodiments, plants include fruit-bearing plants such as grape, orange, lemon, grapefruit, avocado, date, peach, cherry, olive, plum, coconut, apple, pear, blackberry, blueberry, raspberry, strawberry, pineapple, tomato, cucumber and eggplant plants.

**[0077]** In various embodiments, the disclosed polynucleotides employed in the methods and compositions can be provided in expression cassettes for expression in a plant of interest. In some embodiments, the cassette will include 5' and 3' regulatory sequences operably linked to a polynucleotide. In an embodiment, an operable linkage between a polynucleotide of interest and a regulatory sequence (e.g., a promoter) is a functional link that allows for expression of the polynucleotide of interest. Operably linked elements may be contiguous or non-contiguous. When used to refer to the joining of two protein coding regions, by operably linked is intended that the coding regions are in the same reading frame. The cassette may additionally contain at least one additional gene to be cotransformed into the organism. In other cases, the additional gene(s) can be provided on multiple expression cassettes. Such an expression cassette may be provided with a plurality of restriction sites, recombination sites or both for insertion of the K-domain polynucleotide to be under the transcriptional regulation of the regulatory regions. In various embodiments, the expression cassette may additionally contain selectable marker genes. In some embodiments, the expression cassette can include in the 5'-3' direction of transcription, a transcriptional and translational initiation region (e.g., a promoter), a K-domain polynucleotide,

and a transcriptional and translational termination region (*e.g.*, termination region) functional in plants. In various embodiments, the regulatory regions (*e.g.*, promoters, transcriptional regulatory regions, and translational termination regions), the K-domain polynucleotide or both may be native/analogous to the host cell or to each other. In other embodiments, the regulatory regions, the K-domain polynucleotides or both may be heterologous to the host cell or to each other. In one embodiment, a promoter operably linked to a heterologous polynucleotide is from a species different from the species from which the polynucleotide was derived, or, if from the same/analogous species, one or both are substantially modified from their original form and/or genomic locus, or the promoter is not the native promoter for the operably linked polynucleotide. In some embodiments, a chimeric gene having a coding sequence operably linked to a transcription initiation region that is heterologous to the coding sequence. In other embodiments, heterologous promoters or the native promoter sequences may be used to express the sequences. In various embodiments, the termination region may be native with the transcriptional initiation region, may be native with the operably linked K-domain polynucleotide, may be native with the plant host, or may be derived from another source (*e.g.*, foreign or heterologous) to the promoter, the K-domain polynucleotide, the plant host, or any combination thereof. Convenient termination regions are available from the Ti-plasmid of *A. tumefaciens*, such as the octopine synthase and nopaline synthase termination regions. See also, Guerineau, et al., (1991) *Mol Gen Genet* 262:141-144; Proudfoot (1991) *Cell* 64:671-674; Sanfacon, et al., (1991) *Genes Dev* 5:141-149; Mogen, et al., (1990) *Plant Cell* 2:1261-1272; Munroe, et al., (1990) *Gene* 91:151-158; Ballas, et al., (1989) *Nucleic Acids Res* 17:7891-7903; and Joshi, et al., (1987) *Nucleic Acids Res* 15:9627-9639.

**[0078]** Where appropriate, the polynucleotides may be optimized for increased expression in the transformed plant. That is, the polynucleotides can be synthesized using plant-preferred codons for improved expression. See, for example, Campbell and Gown (1990) *Plant Physiol* 92:1-11 for a discussion of host-preferred codon usage. Methods are available in the art for synthesizing plant-preferred genes. See, for example, U.S. Pat. Nos. 5,380,831 and 5,436,391, and Murray, et al., (1989) *Nucleic Acids Res* 17:477-498.

**[0079]** Additional sequence modifications are known to enhance gene expression in a cellular host. These include elimination of sequences encoding spurious polyadenylation signals, exon-intron splice site signals, transposon repeats, and other such well sequences that may be deleterious to gene expression. The G-C content of the sequence may be adjusted to levels average for a given cellular host, as calculated by reference to known genes expressed in the host cell. When possible, the sequence may be modified to avoid predicted hairpin secondary mRNA structures.

WO 2015/006105

PCT/US2014/045115

**[0080]** The expression cassettes may additionally contain 5' leader sequences. Such leader sequences can act to enhance translation. Translation leaders are known in the art and include, for example picornavirus leaders, EMCV leader (Encephalomyocarditis 5' noncoding region) (Elroy-Stein, et al., (1989) Proc Natl Acad Sci USA 86:6126-6130); potyvirus leaders, for example, TEV leader (Tobacco Etch Virus) (Gallie, et al., (1995) Gene 165(2):233-238), MDMV leader (Maize Dwarf Mosaic Virus) (Virology 154:9-20), and human immunoglobulin heavy-chain binding protein (BiP) (Macejak, et al., (1991) Nature 353:90-94); untranslated leader from the coat protein mRNA of alfalfa mosaic virus (AMV RNA 4) (Jobling, et al., (1987) Nature 325:622-625); tobacco mosaic virus leader (TMV) (Gallic, et al., (1989) in Molecular Biology of RNA, ed. Cech (Liss, New York), pp. 237-256); and maize chlorotic mottle virus leader (MCMV) (Lommel, et al., (1991) Virology 81:382-385). See also, Della-Cioppa, et al., (1987) Plant Physiol 84:965-968.

**[0081]** In preparing the expression cassette, the various DNA fragments may be manipulated, so as to provide for the DNA sequences in the proper orientation and, as appropriate, in the proper reading frame. Toward this end, adapters or linkers may be employed to join the DNA fragments or other manipulations may be involved to provide for convenient restriction sites, removal of superfluous DNA, removal of restriction sites, or the like. For this purpose, in vitro mutagenesis, primer repair, restriction, annealing, resubstitutions, *e.g.*, transitions and transversions, may be involved.

**[0082]** In various embodiments, promoters include the native promoter of the polynucleotide sequence of interest. The promoters can be selected based on the desired outcome. The nucleic acids can be combined with constitutive, tissue-preferred, inducible or other promoters for expression in plants. In some embodiments, constitutive promoters include the core promoter of the Rsyn7 promoter and other constitutive promoters disclosed in WO 99/43838 and U.S. Pat. No. 6,072,050; the core CaMV 35S promoter (Odell, et al., (1985) Nature 313:810-812); rice actin (McElroy, et al., (1990) Plant Cell 2:163-171); ubiquitin (Christensen, et al., (1989) Plant Mol Biol 12:619-632 and Christensen, et al., (1992) Plant Mol Biol 18:675-689); pEMU (Last, et al., (1991) Theor Appl Genet 81:581-588); MAS (Velten, et al., (1984) EMBO J 3:2723-2730); ALS promoter (U.S. Pat. No. 5,659,026), GOS2 promoter (dePater, et al., (1992) Plant J 2:837-44), and the like. Other constitutive promoters include, for example, U.S. Pat. Nos. 5,608,149; 5,608,144; 5,604,121; 5,569,597; 5,466,785; 5,399,680; 5,268,463; 5,608,142; and 6,177,611.

**[0083]** In various embodiments, a tissue-specific promoter is a promoter that is active in transcription in the organs or tissue of roots, stem, leaves, or seeds. For example, a root-specific promoter is a promoter that is active in transcription predominantly in plant roots. Various tissue-specific promoters are disclosed in US. Patent No. 8,697,947.

**[0084]** In various embodiments, the expression cassette can also include a selectable marker gene for the selection of transformed cells. Selectable marker genes may be used for the selection of transformed cells or tissues. In various embodiments, marker genes include genes encoding antibiotic resistance, such as those encoding neomycin phosphotransferase II (NPT II) and hygromycin phosphotransferase (HPT), as well as genes conferring resistance to herbicidal compounds, such as glufosinate ammonium, bromoxynil, imidazolinones, and 2,4-dichlorophenoxyacetate (2,4-D). In other embodiments, selectable markers include phenotypic markers such as  $\beta$ -galactosidase and fluorescent proteins such as green fluorescent protein (GFP) (Su, et al., (2004) *Biotechnol Bioeng* 85:610-9 and Fetter, et al., (2004) *Plant Cell* 16:215-28), cyan fluorescent protein (CYP) (Bolte, et al., (2004) *J Cell Science* 117:943-54 and Kato, et al., (2002) *Plant Physiol* 129:913-42), and yellow fluorescent protein (PhiYFP.TM. from Evrogen, see, Bolte, et al., (2004) *J Cell Science* 117:943-54). For additional selectable markers, see generally, Yarranton (1992) *Curr Opin Biotech* 3:506-511; Christopherson, et al., (1992) *Proc Natl Acad Sci USA* 89:6314-6318; Yao, et al., (1992) *Cell* 71:63-72; Reznikoff (1992) *Mol Microbiol* 6:2419-2422; Barkley, et al., (1980) in *The Operon*, pp. 177-220; Hu, et al., (1987) *Cell* 48:555-566; Brown, et al., (1987) *Cell* 49:603-612; Figge, et al., (1988) *Cell* 52:713-722; Deuschle, et al., (1989) *Proc Natl Acad Sci USA* 86:5400-5404; Fuerst, et al., (1989) *Proc Natl Acad Sci USA* 86:2549-2553; Deuschle, et al., (1990) *Science* 248:480-483; Gossen (1993) Ph.D. Thesis, University of Heidelberg; Reines, et al., (1993) *Proc Natl Acad Sci USA* 90:1917-1921; Labow, et al., (1990) *Mol Cell Biol* 10:3343-3356; Zambretti, et al., (1992) *Proc Natl Acad Sci USA* 89:3952-3956; Baim, et al., (1991) *Proc Natl Acad Sci USA* 88:5072-5076; Wyborski, et al., (1991) *Nucleic Acids Res* 19:4647-4653; Hillenand-Wissman (1989) *Topics Mol Struc Biol* 10:143-162; Degenkolb, et al., (1991) *Antimicrob Agents Chemother* 35:1591-1595; Kleinschmidt, et al., (1988) *Biochemistry* 27:1094-1104; Bonin (1993) Ph.D. Thesis, University of Heidelberg; Gossen, et al., (1992) *Proc Natl Acad Sci USA* 89:5547-5551; Oliva, et al., (1992) *Antimicrob Agents Chemother* 36:913-919; Hlavka, et al., (1985) *Handbook of Experimental Pharmacology*, Vol. 78 (Springer-Verlag, Berlin); Gill, et al., (1988) *Nature* 334:721-724. The above list of selectable marker genes is not meant to be limiting. Any selectable marker gene can be used.

**[0085]** In various embodiments, the disclosed polynucleotides can be stacked with any combination of polynucleotide sequences of interest to create plants with a desired trait. In other embodiments, the combinations generated can also include multiple copies of any one of the polynucleotides of interest. In still other embodiments, the disclosed polynucleotides can also be stacked with traits desirable for disease or herbicide resistance (e.g., fumonisin detoxification genes (U.S. Pat. No. 5,792,931); avirulence and disease resistance genes (Jones, et al., (1994)

Science 266:789; Martin, et al., (1993) Science 262:1432; Mindrinos, et al., (1994) Cell 78:1089); acetolactate synthase (ALS) mutants that lead to herbicide resistance such as the S4 and/or Hra mutations; inhibitors of glutamine synthase such as phosphinothricin or basta (*e.g.*, bar gene); and glyphosate resistance (EPSPS gene)); and traits desirable for processing or process products such as high oil (*e.g.*, U.S. Pat. No. 6,232,529); modified oils (*e.g.*, fatty acid desaturase genes (U.S. Pat. No. 5,952,544; WO 94/11516)); modified starches (*e.g.*, ADPG pyrophosphorylases (AGPase), starch synthases (SS), starch branching enzymes (SBE), and starch debranching enzymes (SDBE)); and polymers or bioplastics (*e.g.*, U.S. Pat. No. 5,602,321; beta-ketothiolase, polyhydroxybutyrate synthase, and acetoacetyl-CoA reductase (Schubert, et al., (1988) J Bacteriol 170:5837-5847) facilitate expression of polyhydro xyalkanoates (PHAs)). One could also combine the disclosed polynucleotides with polynucleotides providing agronomic traits such as male sterility (*e.g.*, see U.S. Pat. No. 5,583,210), stalk strength, or transformation technology traits such as cell cycle regulation or gene targeting (*e.g.*, WO 99/61619, WO 00/17364, and WO 99/25821).

**[0086]** These stacked combinations can be created by any method including, but not limited to, cross-breeding plants by any conventional or TopCross methodology, or genetic transformation. If the sequences are stacked by genetically transforming the plants, in various embodiments the disclosed polynucleotide sequences can be combined at any time and in any order. In some embodiments, a transgenic plant having one or more desired traits can be used as the target to introduce further traits by subsequent transformation. The traits can be introduced simultaneously in a co-transformation protocol with the disclosed polynucleotides provided by any combination of transformation cassettes. In one embodiment, if two sequences will be introduced, the two sequences can be contained in separate transformation cassettes (*trans*) or contained on the same transformation cassette (*cis*). Expression of the sequences can be driven by the same promoter or by different promoters. This may be combined with any combination of other suppression cassettes or overexpression cassettes to generate the desired combination of traits in the plant. It is further recognized that polynucleotide sequences can be stacked at a desired genomic location using a site-specific recombination system. See, for example, WO99/25821, WO99/25854, WO99/25840, WO99/25855, and WO99/25853.

**[0087]** The various embodiments provide for methods that involve introducing a polynucleotide or polypeptide into a plant. Any suitable method for introducing a sequence into a plant may be used such that the polynucleotide or polypeptides gains access to the interior of at



least one plant cell. Methods for introducing polynucleotide or polypeptides into plants are known in the art including, but not limited to, stable transformation methods, transient transformation methods, and virus-mediated methods.

**[0088]** Transformation protocols as well as protocols for introducing polypeptides or polynucleotide sequences into plants may vary depending on the type of plant or plant cell, *e.g.*, monocot or dicot, targeted for transformation. Suitable methods of introducing polypeptides and polynucleotides into plant cells include microinjection (Crossway, et al., (1986) *Biotechniques* 4:320-334), electroporation (Riggs, et al., (1986) *Proc Natl Acad Sci USA* 83:5602-5606, *Agrobacterium*-mediated transformation (U.S. Pat. No. 5,563,055 and U.S. Pat. No. 5,981,840), direct gene transfer (Paszkowski, et al., (1984) *EMBO J* 3:2717-2722), and ballistic particle acceleration (see, for example, U.S. Pat. No. 4,945,050; U.S. Pat. No. 5,879,918; U.S. Pat. Nos. 5,886,244; and 5,932,782; Tomes, et al., (1995) in *Plant Cell, Tissue, and Organ Culture: Fundamental Methods*, ed. Gamborg and Phillips (Springer-Verlag, Berlin); McCabe, et al., (1988) *Biotechnology* 6:923-926); and *Led* transformation (WO 00/28058). Also see, Weissinger, et al., (1988) *Ann Rev Genet* 22:421-477; Sanford, et al., (1987) *Particulate Science and Technology* 5:27-37 (onion); Christou, et al., (1988) *Plant Physiol* 87:671-674 (soybean); McCabe, et al., (1988) *Bio/Technology* 6:923-926 (soybean); Finer and McMullen (1991) *In Vitro Cell Dev Biol* 27P:175-182 (soybean); Singh, et al., (1998) *Theor Appl Genet* 96:319-324 (soybean); Datta, et al., (1990) *Biotechnology* 8:736-740 (rice); Klein, et al., (1988) *Proc Natl Acad Sci USA* 85:4305-4309 (maize); Klein, et al., (1988) *Biotechnology* 6:559-563 (maize); U.S. Pat. Nos. 5,240,855; 5,322,783; and 5,324,646; Klein, et al., (1988) *Plant Physiol* 91:440-444 (maize); Fromm, et al., (1990) *Biotechnology* 8:833-839 (maize); Hooykaas-Van Slogteren, et al., (1984) *Nature (London)* 311:763-764; U.S. Pat. No. 5,736,369 (cereals); Bytebier, et al., (1987) *Proc Natl Acad Sci USA* 84:5345-5349 (*Liliaceae*); De Wet, et al., (1985) in *The Experimental Manipulation of Ovule Tissues*, ed. Chapman, et al., (Longman, New York), pp. 197-209 (pollen); Kaeppler, et al., (1990) *Plant Cell Reports* 9:415-418 and Kaeppler, et al., (1992) *Theor Appl Genet* 84:560-566 (whisker-mediated transformation); D'Halluin, et al., (1992) *Plant Cell* 4:1495-1505 (electroporation); Li, et al., (1993) *Plant Cell Reports* 12:250-255 and Christou and Ford (1995) *Annals of Botany* 75:407-413 (rice); Osjoda, et al., (1996) *Nature Biotechnology* 14:745-750 (maize via *Agrobacterium tumefaciens*).

**[0089]** In some embodiments, the K-domain sequences or variants thereof can be provided to a plant using a variety of transient transformation methods. Such transient transformation methods include, but are not limited to, the introduction of the K-domain or variants thereof directly into the plant or the introduction of the K-domain transcript into the plant. Such

methods include, for example, microinjection or particle bombardment. See, for example, Crossway, et al., (1986) *Mol Gen Genet* 202:179-185; Nomura, et al., (1986) *Plant Sci* 44:53-58; Hepler, et al., (1994) *Proc Natl Acad Sci* 91:2176-2180 and Hush, et al., (1994) *The Journal of Cell Science* 107:775-784. Other techniques include viral vector system and the precipitation of the polynucleotide in a manner that precludes subsequent release of the DNA. Thus, the transcription from the particle-bound DNA can occur, but the frequency with which it is released to become integrated into the genome is greatly reduced. Such methods include the use particles coated with polyethylimine (e.g., PEI; Sigma #P3143).

**[0090]** In other embodiments, the polynucleotide may be introduced into plants by contacting plants with a virus or viral nucleic acids. In some embodiments, such methods involve incorporating a nucleotide construct within a viral DNA or RNA molecule. In one embodiment the K-domain sequence or a variant thereof may be initially synthesized as part of a viral polyprotein, which later may be processed by proteolysis in vivo or in vitro to produce the desired recombinant protein. In various embodiments, the promoters also encompass promoters utilized for transcription by viral RNA polymerases. Methods for introducing polynucleotides into plants and expressing a protein encoded therein, involving viral DNA or RNA molecules, are known in the art. See, for example, U.S. Pat. Nos. 5,889,191, 5,889,190, 5,866,785, 5,589,367, 5,316,931, and Porta, et al., (1996) *Molecular Biotechnology* 5:209-221.

**[0091]** A variety of methods for the targeted polynucleotide insertion at a specific location in the plant genome may be used. In some embodiments, the polynucleotide insertion at a desired genomic location is achieved using a site-specific recombination system. See, for example, WO99/25821, WO99/25854, WO99/25840, WO99/25855, and WO99/25853. In one embodiment, the disclosed polynucleotide can be contained in a transfer cassette flanked by two non-recombinogenic recombination sites. The transfer cassette is introduced into a plant having stably incorporated into its genome a target site which is flanked by two non-recombinogenic recombination sites that correspond to the sites of the transfer cassette. An appropriate recombinase is provided and the transfer cassette is integrated at the target site. The disclosed polynucleotide is thereby integrated at a specific chromosomal position in the plant genome.

**[0092]** The cells that have been transformed may be grown into plants in accordance with conventional ways. See, for example, McCormick, et al., (1986) *Plant Cell Reports* 5:81-84. These plants may then be grown, and either pollinated with the same transformed strain or different strains, and the resulting progeny having expression of the desired phenotypic

characteristic identified. Two or more generations may be grown to ensure that expression of the desired phenotypic characteristic is stably maintained and inherited and then seeds harvested to ensure expression of the desired phenotypic characteristic has been achieved. In this manner, in one embodiment, transformed seed (also referred to as “transgenic seed”) having a polynucleotide, stably incorporated into their genome may be made. In other embodiments, transformed seed or transgenic seed have an expression cassette stably incorporated into their genome.

**[0093]** Any method or composition that modulates expression of a target gene product (*e.g.*, overexpression or knock-down expression), either at the level of transcription or translation, or modulates the activity of the target gene product can be used to achieve overexpression, concentration, activation of the target gene product. In various embodiments, the level is increased by at least 1%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% or greater including any range or value therebetween relative to an appropriate control plant, plant part, or cell. Expression may occur during subsequent to plant growth at a desired development stage or both. In some embodiments, the disclosed polypeptides are modulated in fruit bearing plants. In other embodiments, the fruit bearing plants include blueberries and the like.

**[0094]** In various embodiments, the expression level of a polypeptide having a K domain or a biologically active variant or fragment thereof may be measured directly or indirectly. In some embodiments, the expression level of a polypeptide having a K domain or a biologically active variant or fragment thereof may be measured by assaying for the level of the K-domain polypeptide in the plant.

**[0095]** In other embodiments, the expression level of a polypeptide having a K domain or a biologically active variant or fragment thereof may be measured by measuring the level of the polynucleotide encoding the protein or by measuring the activity of the K-domain polypeptide in the plant. The polypeptide or the polynucleotide may be introduced into a plant cell. The plant cell having the introduced sequence may be selected using methods known to those of skill in the art such as, but not limited to, Southern blot analysis, Northern blot, DNA sequencing, PCR analysis, or phenotypic analysis. In various embodiments, a plant or plant part altered or modified by the foregoing may be grown under plant-promoting conditions for a time sufficient to modulate the concentration and/or activity of disclosed polypeptides in the plant.

**[0096]** In various embodiments, the level and/or activity of the polypeptide may be modulated by employing a polynucleotide that is not capable of directing, in a transformed plant, the expression of a protein or an RNA. For example, the polynucleotide may be used to design polynucleotide constructs that can be employed in methods for altering or mutating a genomic nucleotide sequence in an organism. Such polynucleotide constructs include, but are not limited

to, RNA:DNA vectors, RNA:DNA mutational vectors, RNA:DNA repair vectors, mixed-duplex oligonucleotides, self-complementary RNA:DNA oligonucleotides, and recombinogenic oligonucleobases. Such nucleotide constructs and methods of use are known in the art. See, U.S. Pat. Nos. 5,565,350; 5,731,181; 5,756,325; 5,760,012; 5,795,972; and 5,871,984. See also, WO 98/49350, WO 99/07865, WO 99/25821, and Beetham, et al., (1999) Proc Natl Acad Sci USA 96:8774-8778.

**[0097]** The various methods described herein do not depend on the incorporation of the entire polynucleotide into the genome, only that the plant or cell thereof may be altered as a result of the introduction of the polynucleotide into a cell. In various embodiments, the genome may be altered following the introduction of the polynucleotide into a cell. In other embodiments, the polynucleotide, or any part thereof, may incorporate into the genome of the plant. Alterations to the genome include, but are not limited to, additions, deletions, and substitutions of nucleotides into the genome.

**[0098]** The activity, level or both of a K-domain polypeptide is increased. In various embodiments, an increase in the activity, level or both of the K-domain polypeptide is by providing to the plant a K-domain polypeptide or a biologically active variant or fragment thereof. As discussed elsewhere herein, many methods are known in the art for providing a polypeptide to a plant including, but not limited to, direct introduction of the K-domain polypeptide into the plant or introducing into the plant (transiently or stably) a polynucleotide construct encoding a polypeptide having K-domain activity. In various embodiments, the level and/or activity of a K-domain polypeptide may be increased by altering the gene encoding the K-domain polypeptide or its promoter. See, e.g., Kmiec, U.S. Pat. No. 5,565,350; Zarling, et al., PCT/US93/03868. Therefore, mutagenized plants that carry mutations in K-domain genes, where the mutations increase expression of the K-domain gene or increase the activity of the encoded K-domain polypeptide, are provided.

**[0099]** In various embodiments, the level of a single K-domain sequence can be modulated to produce the desired phenotype. In other embodiments, the level of expression of multiple sequences having a K-domain or a biologically active variant or fragment thereof are modulated. In still other embodiments, the level of the K-domain sequence is modulated by an increase. **[00100]** The various embodiments will be further described by reference to the following examples, which are offered to further illustrate various embodiments. It should be understood, however, that many variations and modifications may be made while remaining within the scope of the various embodiments.

## EXAMPLE 1

Materials and methodsCloning and phylogenetic analysis of VcK-domain

[00101] To generate cDNA sequences for cloning of *VcK-domain*, total RNA was isolated from vernalized flower buds from a northern highbush blueberry (*Vaccinium corymbosum* L.) cultivar, Bluecrop, using the RNeasy Plant Mini Kit (Qiagen, Valencia, CA, USA) with modifications as described by Dhanaraj *et al.* (2004). Following isolation, the sample was treated with RQ1 DNase (Promega, Madison, WI, USA) according to the manufacturer's instructions, followed by chloroform:isoamyl alcohol (24:1) extraction, 100% ethanol precipitation, 70% ethanol wash, and finally resuspended in diethylpyrocarbonate (DEPC)-treated water. Total RNA (0.6 µg) was reverse transcribed into complementary DNA (cDNA) using SuperScript II Reverse Transcriptase (Invitrogen, Carlsbad, CA, USA). One-tenth of the cDNA was used in one PCR amplification reaction using GoTaq Green Master Mix (Promega).

[00102] Blueberry EST sequence "*Vaccinium corymbosum\_v1\_Contig03391*", downloaded from the blueberry EST database (<http://www.vaccinium.org/>) was used to design primers for *VcK-domain*. Nested PCRs were conducted using primers SOCF1 and SOCR1 for the 1<sup>st</sup> round of amplification and SOF1 and SOK (see Table 1) for the 2<sup>nd</sup> round of amplification. PCR products were cloned into a pCR 2.1-TOPO vector (Invitrogen) and verified by sequencing.

Table 1. Primers used.

Primer name	Sequence	Target
SOCF1	SEQ ID NO: 4	Blueberry <i>SUPPRESSOR OF OVEREXPRESSION OF CONSTANS 1</i> -like gene ( <i>VcK-domain</i> )
SOCR1	SEQ ID NO: 5	
SOF1:	SEQ ID NO: 6	
SOK:	SEQ ID NO: 7	
35S-F	SEQ ID NO: 8	3' portion of the <i>CaMV 35S</i> promoter
NPTII-F	SEQ ID NO: 9	Neomycin phosphotransferase II gene ( <i>nptII</i> )
NPTII-R	SEQ ID NO: 10	
Actin-F	SEQ ID NO: 11	<i>Actin</i>
Actin-R	SEQ ID NO: 12	
BBK-domain - F	SEQ ID NO: 13	<i>VcK-domain</i>

---

BBK- SEQ ID NO: 14  
domain - R

---

**[00103]** Amino acid and nucleotide sequences of *VcK-domain* orthologues were retrieved from the NCBI (<http://blast.ncbi.nlm.nih.gov/>). Selected amino acid sequences were aligned using the ClustalW2 multiple sequence alignment program (Thompson et al. 1994) at the EBI with default parameters (Protein Weight Matrix: Gonnet, gap open penalty of 10 and gap extension penalty of 0.20, gaps that are closer than 5 amino acids are penalized, clustering method of Neighbour-joining, <http://www.ebi.ac.uk/Tools/msa/clustalw2/>). Phylogenetic analysis of nucleotide sequences of *K-domain*-like genes from 17 dicot plants and one monocot plant (*Oryza sativa*) by Maximum Likelihood method based on the Tamura-Nei model using the MEGA 5 program (Tamura et al. 2011) was performed. *OsMADS56* was used as an outgroup and the black dot denotes the *K-domain*-like protein from blueberry (*VcK-domain-like*). Rate variation among sites was modeled using a discrete Gamma distribution (4 categories) with invariant sites (G+I). The tree was drawn to scale, with the branch length equal to the number of substitutions per nucleotide.

#### Construction of the expression vector and transformation of tobacco

**[00104]** To make a construct for plant transformation, the 5'-*Xba*I- *VcK-domain* -*Sac*I-3' fragment was released from the pCR2.1-TOPO vector and then inserted into the opened *Xba*I and *Sac*I sites in the T-DNA region between the cauliflower mosaic virus (CaMV) 35S promoter and the *Nos* terminator in pBI121, where the GUS ( $\beta$ -glucuronidase) coding region was removed. The resulting 35S-*VcK-domain* was sequenced to confirm the *VcK-domain* sequence and subsequently introduced into *Agrobacterium tumefaciens* strain EHA105 (Hood et al. 1993) using the freeze-thaw method.

**[00105]** *Agrobacterium*-mediated transformation of tobacco (*Nicotiana tabacum* L. cv Samsun) leaf explants with 35S-*VcK-domain* was conducted according to Horsch et al. (1985). Kanamycin-resistant shoots were rooted on Murashige and Skoog medium (MS) (1962) containing 100 mg L<sup>-1</sup> kanamycin monosulfate (Km). Transgenic plants from separate leaf explants were considered as independent transgenic events.

**[00106]** To confirm the presence of the transgene, DNA was isolated from leaf tissues using a cetyltrimethylammonium bromide (CTAB) method (Doyle and Doyle, 1987). Two pairs of primers (Table 1) were used for PCR verification: *nptII*-F and *nptII*-R for the *nptII* gene, and 35S-F (3' portion of the *CaMV* 35S promoter) and SOK for the *VcK-domain* gene.

#### Reverse-transcription PCR (RT-PCR) analyses

[00107] Total RNA was isolated from young leaves of selected T0 tobacco plants using an RNeasy Plant Mini Kit (Qiagen, Valencia, CA, USA) with on-column RNase-free DNase (Qiagen) treatment according to the manufacturer's instructions. Reverse-transcription of RNA to cDNA was carried out using SuperScript II reverse transcriptase (Invitrogen).

The resulting cDNA, 2  $\mu$ l/sample, was used for PCR amplification of a 160 base pair endogenous *actin* gene and a 474-base pair fragment of the *VcK-domain*. The primers are shown in table 1. The reaction conditions for all primer pairs were 94°C for 2 min, 35 cycles of 45 sec at 94°C, 60 sec at 58°C and 90 sec at 72°C, with a final 10 min extension at 72°C. All amplified PCR and RT-PCR products were separated on 1.0% agarose gel containing ethidium bromide, visualized and photographed under UV light.

#### Phenotyping of transgenic tobacco plants

[00108] When T0 putative transformants and nontransgenic/control plants grew to the stage of 3-5 leaves, they were transplanted from the MS medium to 10  $\times$  10 cm pots containing Suremix potting media (Michigan Grower Products Inc, Galesburg, MI, USA) and grown under a 16-hour photoperiod with 30  $\mu$ mol m<sup>-2</sup>s<sup>-1</sup> light from cool white fluorescent tubes for 3 weeks before being repotted into 1-gallon pots. Flower structure of transgenic plants was compared with control plants through a dissecting microscope. These T0 transgenic plants were self-pollinated to obtain the T1 seeds used in the remainder of the study.

[00109] Transgenic tobacco plants with pBI121 (the backbone vector) showed no significant difference in phenotype (*e.g.*, flowering time, plant height, and seed yield) compared to wild type plants. T1 seeds and nontransgenic tobacco were planted in 48-cell trays containing Suremix potting media and maintained under a 16-hour photoperiod with 30  $\mu$ mol m<sup>-2</sup>s<sup>-1</sup> light from cool white fluorescent tubes while transgene presence was confirmed by PCR using 35S-F (SEQ ID NO: 8) and SOK (SEQ ID NO: 7) primers (Table 1). Seedlings of similar size (10 for each transgenic event or control) were then transplanted into 10  $\times$  10 cm pots and grown in the greenhouse under natural photoperiod (an average of 14.9 hrs of natural sunlight per day) with a minimum temperature of 23 °C in a completely randomized design with regular irrigation and fertilization.

[00110] Days to flowering (days between sowing and appearance of the first flower bud), leaf number, plant height, and number of seed pods were recorded for the T1 plants. The number of leaves on the central stem and plant height (from soil level to the bottom of the central flower cluster) were determined when 50% of flowers in the central cluster were open. Number of seed pods in the central flower cluster was determined when plants finished blooming. Statistical

analysis was conducted with PROC GLM in SAS 9.3 (SAS Institute Inc., Cary, NC, USA). Difference from wild type tobacco was determined using Dunnett's two-tailed test ( $\alpha = 0.05$ ).

#### Isolation and sequence analysis of *VcK*-domain

[00111] To search for a *K*-domain-like gene in blueberry, a *Vitis vinifera* *K*-domain-like gene (GenBank accession GU133633.1) was used as the query to BLAST against the blueberry EST database (<http://www.vaccinium.org/tools/blast>). Based on E-value, the top hit was an 1189-bp EST sequence (*Vaccinium\_corymbosum\_v1\_Contig03391*), which contains a MADS-box motif and a K domain. Comparison with *K*-domain-like genes from other plant species indicates that the Contig03391 contains the full sequence of a blueberry *K*-domain-like gene (*VcK*-domain) and is grouped in the clade of *K*-domain-like genes along with *MADS1* (*Nicotiana tabacum*) and *MADS8* (*Vitis vinifera*) based on the phylogentic analysis (Fig. 1A).

[00112] To investigate the function of the K-domain in a SOC1 gene, the K-domain of the *VcK*-domain, a 471-bp ORF encoding 157 amino acids was cloned. The cloned *VcK*-domain has three nucleotide mismatches (e.g., positions 51, 171, and 178) from that of the Contig03391. One mismatch resulted in an amino acid substitution. Multiple sequence alignment of nine selected amino acid sequences (e.g., six from woody plants and one from *Arabidopsis*) of *VcK*-domain orthologues indicates that the *VcK*-domain is highly conserved in these plant species (Fig. 1B).

#### Molecular evidence of the overexpression of *VcK*-domain

[00113] RT-PCR confirmed transcription of the *VcK*-domain (Fig. 2). In young leaves of both transgenic (selected T0) and non-transgenic (NT) plants, the consistent bands of an endogenous *actin* gene were observed; in contrast, an expected 474-bp fragment of the *VcK*-domain was observed in all of the transgenic plants tested but was absent in non-transformed plants (Fig. 2). Figure 2 shows the following lanes; Lanes 1 & 5: transgenic event I5-4; Lanes 2 & 6: transgenic event I5-6; Lanes 3 & 7: NT; Lanes 4 & 8: water as template; and lane M: 1-kb DNA ladder.

#### Phenotype changes of *VcK*-domain transgenic tobacco plants

[00114] Ectopic expression of *VcK*-domain did not result in any morphological changes in floral organ identity, but promoted early flowering in transgenic tobacco plants (Fig. 3A-C). Of the 22 T0 transgenic events regenerated, all had normal flowers, as did T1 plants derived from five events (I5-5, I5-8, I5-16, I5-18, and I5-22) chosen for continued study. The T1 plants from four of these selected events flowered an average of 18 days (ranging from 15 to 21 days, depending on event) earlier than wild type plants, which flowered 129 days after sowing (Fig. 3B, C). The result suggests that *VcK*-domain is involved in flower timing.

[00115] Although no obvious morphological changes were observed in flowers, *VcK*-domain transgenic plants did exhibit altered growth. All four early flowering T1 transgenic events



showed significant decreases in plant height and one had significantly fewer leaves at the time of flowering compared to wild type plants (Fig. 3D-F).

[00116] Expression of 35S-*VcK-domain* did not significantly affect the number of seed pods in five lines of T1 plants evaluated ( $P = 0.0501$ ), suggesting that expression of the 35S-*VcK-domain* has little impact on seed or fruit production per plant.

[00117] Seed production of 20 T0 plants was compared with non-transgenic plants. The average seed production for 20 transgenic events was 8.02 grams, which is 32% higher than wild type plants (6.08 g/plant), although there is no statistical difference between transgenic plants and wild type plants due to the variations among different transgenic events (Table 2).

Table 2. Seed production of transgenic tobacco plants overexpressing the VcK domain

	Average seed production per plant (gram) $\pm$ STDEV
T0 plants of 20 transgenic events	8.02 $\pm$ 2.71 a
Three wild type plants	6.08 $\pm$ 0.42 a

[00118] Of the 20 independent transgenic events of tobacco plants, seed yields of 14 events (70%) were higher than the wild type plant. The results suggest a high potential (more than 70% chances) of obtaining individual transgenic events with an increased seed production.

## EXAMPLE 2

### Cloning the full length *VcSOC1* homologues

[00119] Following the protocols for the *VcK* cloning as described in Example 1, primers BOC-FX (5'-tctagaATG GTG AGA GGG AAA ACC CAGA-3'; SEQ ID NO:15) and BOC-RS (5'-gagctcACTTGGGTGGCGAAACT-3'; SEQ ID NO:16) were used for amplification of *VcSOC1* homologues from cDNA of blueberry cultivar Bluecrop. PCR products were cloned into a pCR2.1-TOPO vector (Invitrogen) and verified by sequencing. Three *VcSOC1* homologues were identified as *VcSOC1A* (SEQ ID NO: 17), *VcSOC1B* (SEQ ID NO: 18) and *VcSOC1D* (SEQ ID NO: 19). The 5'-Xba I-VcSOC1-Sac I-3' fragments were released from the pCR2.1-TOPO vectors and then inserted into the opened *Xba*I and *Sac*I sites in the T-DNA region between the cauliflower mosaic virus (CaMV) 35S promoter and the Nos terminator in pBI121, where the GUS ( $\beta$ -glucuronidase) coding region was removed. The resulting vectors were independently introduced into *Agrobacterium tumefaciens* strain EHA105 and as described in Example 1. Independent transformation of the three *VcSOC1* homologues into tobacco was carried out following the protocols for the *VcK-domain* transformation as described in Example 1.

### Molecular analyses

To evaluate the expression of the three *VcSOCI* homologues in tobacco, the SOF1 and SOK primer pair (as shown in Table 1) were used for reverse-transcription PCR (RT-PCR) analyses of T0 plants. The RT-PCR analysis was conducted as described in Example 1.

#### Phenotyping of transgenic tobacco plants

**[00120]** Phenotyping of transgenic tobacco plants transformed with the three *VcSOCI* homologues was conducted in the greenhouse under natural light conditions in the Fall of Michigan. Three T0 transgenic events for each of the *VcSOCI* homologues were used. These plants were all RT-PCR positive for the *VcSOCI*. Each event was proliferated to obtain three plants by in vitro culturing internode cuttings on MS medium; non-transgenic plants and one T1 *VcK-domain* plant were used as control. These plants were grown in 10 x 10 cm pots. Days to flowering (days between sowing and appearance of the first flower bud), leaf number, plant height, and number of seed pods were recorded. Ectopic overexpression was determined by the presence of a 657 base pair for *VcSOCI A*, 576 base pair for *VcSOCI B*, and 654 base pair *VcSOCI D*.

**[00121]** Ectopic overexpression of the three *VcSOCI* homologues (*VcSOCI A*, *VcSOCI B*, and *VcSOCI D*) showed similar function as overexpression of the *VcK-domain*, such as early flowering (1-3 weeks earlier than non-transgenic controls) without significantly reducing the leaf number and the number of seed pods per T0 transgenic tobacco plant.

#### EXAMPLE 3

**[00122]** The *VcK-domain* construct as described in Example 1 was used to ectopically overexpress the *VcK* in blueberry plants. To this end, transformation of a northern highbush blueberry cultivar Aurora was performed as reported in Song and Sink, (*Agrobacterium tumefaciens*-mediated transformation of blueberry (*Vaccinium corymbosum* L.). Plant Cell Rep. 23: 475-484 and Song G-Q, Sink KC (2006) *Agrobacterium*-mediated transformation of highbush blueberry (*Vaccinium corymbosum* L.) cultivars. In: Wang K, ed. *Agrobacterium protocols: Methods in molecular biology* 344. 2nd edn. Totowa NJ: Humana Press, 37-44. Briefly, leaf explants were co-cultivated with EHA105 for 6 days on co-cultivation medium in the presence of 100  $\mu$ M acetosyringone at 25°C in the dark. Selection and regeneration was carried out on regeneration medium (modified McCown's Woody Plant Medium (WPM) containing 4.54  $\mu$ M thidiazuron, 2.69  $\mu$ M  $\alpha$ -naphthaleneacetic acid, 10 mg L<sup>-1</sup>Km, and 250 mg L<sup>-1</sup> cefotaxime) for 12 weeks. Subcultures to fresh selection medium were conducted at four-week intervals. Transgenic shoots from separate explants were used as independent transgenic

events. Proliferation of individual kanamycin-resistant shoots was conducted on WPM containing 4.56  $\mu$ M zeatin, 50 mg L<sup>-1</sup> Km, and 250 mg L<sup>-1</sup> cefotaxime. Wild type controls were regenerated from non-infected leaf explants and transgenic controls were transformed with a *gusA* gene instead of the 35S-*VcK* construct.

[00123] Six plants for each of the three T0 transgenic events transformed with the *VcK* were obtained through micropropagation and they were grown in 3.78-L [17.78 cm (height) x 20.32 cm (diameter)] pots in the greenhouse under natural light conditions. Non-transgenic plants and transgenic plants transformed with pBI121 were used as controls. Plant height, number of branches, number of flowering buds, and date of flowering were recorded.

[00124] Transgenic plants showed accelerated flowering and flowered under no-chilling conditions, where non-transgenic plants did not flower. This suggests overexpressing the *VcK*-domain is sufficient to promote accelerated blueberry flowering.

#### Example 4

[00125] The *VcK*-domain construct as described in Example 1 was used to ectopically overexpress the *VcK* in petunia plants. *Agrobacterium*-mediated transformation of *Petunia hybrida* Mitchell was performed essentially as described by Jorgensen et al. (1996). Briefly, leaf explants, 2–3 × 2–4 mm, were excised from *in vitro* seedlings grown on half-strength MS basal medium (Murashige and Skoog 1962). Co-cultivation was carried out for 4 days on regeneration medium (RM: MS + 4.4 l M BAP + 1.1 l M IAA) supplemented with 100 l M acetosyringone. Following selection with 250 mg l<sup>-1</sup> Timentin and 100 mg l kanamycin, transgenic shoots were produced on RM and rooted on MS. The cultures were maintained at 25 °C under a 16-h light photoperiod. Empty vector (pGA643) and non-transformed explants were also carried through tissue culture as negative controls. T1 seeds were harvested from separate T0 plants.

[00126] Phenotyping of petunia plants was performed between January and June in the greenhouse under natural light conditions in Michigan. Fifteen T1 seedlings of two independent events transformed with the *VcK* were grown in 10 × 10 cm pots. *VcK* presence in transgenic seedlings was confirmed by PCR using 35S-F (SEQ ID NO: 8) and SOK (SEQ ID NO: 7) primers (Table 1). Fifteen plants for each non-transgenic and transgenic seedlings transformed with pGA643, were used as controls. Days to flowering (days between sowing and appearance of the first flower bud), number of flowers, branches and height of the central stem were recorded. (Table 3)

Table 3. Effect of ectopic expression of 35S:*VcK* on phenotypic changes in T1 transgenic petunia plants

Plant	Days to emergence of first flowers after the seeds were sown	Number of plant branches per plant (investigated on June 23 <sup>rd</sup> , 2014)	Number of total flowers per plant (investigated on June 23 <sup>rd</sup> , 2014)
T1 transgenic plants overexpressing <i>VcK</i>	74	5.1	76.4
Wild-type	83	2.6	48.2
T1 transgenic control plants transformed with pGA643	82	2.5	48.3

**[00127]** Compared to wild-type *Petunia hybrida* Mitchell and transgenic control plants containing pGA643, *VcK* overexpressing petunia plants flowered 3-9 days earlier, had more branches and more flowers per plant. These transgenic plants also did not show a significant decrease in plant height and the leaf number on the central stem.

**[00128]** In the various embodiments a method of producing a transgenic plant is provided comprising introducing a K-domain polypeptide encoded by a nucleotide sequence to produce a transformed cell. In various embodiments, the polynucleotide is selected from a nucleotide sequence set forth in SEQ ID NO: 1; a nucleotide sequence encoding the polypeptide of SEQ ID NO: 2; a nucleotide sequence comprising at least 80% sequence identity to the sequence set forth in SEQ ID NO: 1; or a nucleotide sequence encoding an amino acid sequence having at least 90% sequence identity to the sequence set forth in SEQ ID NO: 2. In various embodiments the transformed cell is used to generate the transgenic plant, wherein the transgenic plant has increased expression of the K-domain polypeptide as compared to a control plant. In other embodiments, the sequences used in the method include introducing SEQ ID NO: 17, SEQ ID NO: 18 or SEQ ID NO: 19.

**[00129]** In one embodiment, the K-domain sequence is a K-domain sequence of a MADS box gene.

**[00130]** In one embodiment, the plant used to produce the transgenic plant is a fruit-bearing plant.

**[00131]** In one embodiment, the method used to provide the transgenic plant is a blueberry plant. In other embodiments, the plant is selected from corn, rice, sorghum, millet, soybean, tobacco, potato, wheat, cotton, peanut, rye, proso millet, foxtail millet, finger millet, sunflower, safflower, wheat, sweet potato, cassava, pineapple, citrus trees, cocoa, tea, banana, avocado, fig,

sugarcane, oats, barley, vegetables, ornamentals, and conifers. In still other embodiments, the plant is selected from fruit-bearing plants such as grape, orange, lemon, grapefruit, avocado, date, peach, cherry, olive, plum, coconut, apple, pear, blackberry, raspberry, strawberry, pineapple, tomato, cucumber and eggplant.

**[00132]** In various embodiments the method includes transgenic plants that comprise plant parts selected from an embryo, pollen, ovule, seed, leaf, flower, branch, fruit, kernel, ear, cob, husk, stalk, root, root tips, anther, and combinations thereof.

**[00133]** In various embodiments, the method used to produce a transgenic plant is by stably transforming into a plant with a K-domain and variants of the K-domain.

**[00134]** In various embodiments, the transgenic plant has at least one characteristic selected from accelerated flowering time, reduced plant size, reduced leaf number, and combinations thereof.

**[00135]** In various embodiments, the method of producing transgenic plant with a K-domain polypeptide introduced into it exhibits increased yield compared to a control plant.

**[00136]** In various embodiments, the transgenic plant exhibits increased yield by at least 5%, 7%, 8% or 9%, compared to a control plant. In other embodiments, the transgenic plant exhibits increased yield by at least about 10 to 20 % compared to a control plant. In still other embodiments, the transgenic plant exhibits increased yield by about 30% to about 50% compared to a control plant.

**[00137]** In various embodiments, an expression cassette useful for producing a transgenic plant comprising a polynucleotide encoding a K-domain polypeptide operably linked to a promoter that drives expression in a plant, wherein the polynucleotide is selected from:

a nucleotide sequence set forth in SEQ ID NO: 1;

a nucleotide sequence encoding the polypeptide of SEQ ID NO: 2;

a nucleotide sequence comprising at least 80% sequence identity to the sequence set forth in SEQ ID NO: 1; and

a nucleotide sequence encoding an amino acid sequence having at least 90% sequence identity to the sequence set forth in SEQ ID NO: 2.

**[00138]** In other embodiments, the sequences used in the expression cassette include SEQ ID NO: 3, SEQ ID NO: 17, SEQ ID NO: 18 or SEQ ID NO: 19. In one embodiment, the expression cassette comprises a constitutive promoter.

**[00139]** In one embodiment, an isolated host cell is transformed with the expression cassette in which a K-domain polypeptide operably linked to a promoter that drives expression in a plant, wherein the polynucleotide is selected from

a nucleotide sequence set forth in SEQ ID NO: 1;

a nucleotide sequence encoding the polypeptide of SEQ ID NO: 2;  
a nucleotide sequence comprising at least 80% sequence identity to the  
sequence set forth in SEQ ID NO: 1; and

a nucleotide sequence encoding an amino acid sequence having at least  
90% sequence identity to the sequence set forth in SEQ ID NO: 2.

**[00140]** In various embodiments, various plant parts may be transformed with the expression cassette. In still other embodiments, the various plant parts comprises plant parts selected from an embryo, pollen, ovule, seed, leaf, flower, branch, fruit, kernel, ear, cob, husk, stalk, root, root tips, anther, and combinations thereof.

**[00141]** In various embodiments, the transgenic plant transformed with the expression cassette exhibits has at least one characteristic selected from accelerated flowering time, reduced plant size, reduced leaf number, and combinations thereof.

**[00142]** Various embodiments provide a transgenic plant comprising a plant transformed with a K-domain polypeptide encoded by a nucleotide sequence set forth in SEQ ID NO:1, wherein the transgenic plant has increased expression of the K-domain polypeptide compared to a control plant. In other embodiments, the plant may be transformed with SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:17, SEQ ID NO:18 or SEQ ID NO:19.

**[00143]** In one embodiment, the transgenic plant is a transgenic fruit-bearing plant. In other embodiments the transgenic plant is a blueberry plant. In other embodiments, the plants include grape, orange, lemon, grapefruit, avocado, date, peach, cherry, olive, plum, coconut, apple, pear, blackberry, raspberry, strawberry, pineapple, tomato, cucumber and eggplant.

**[00144]** In still other embodiments, the transgenic plant is selected from corn, rice, sorghum, millet, soybean, tobacco, potato, wheat, cotton, peanut, rye, proso millet, foxtail millet, finger millet, sunflower, safflower, wheat, sweet potato, cassava, pineapple, citrus trees, cocoa, tea, banana, avocado, fig, sugarcane, oats, barley, vegetables, ornamentals, and conifers. In various embodiments, the transgenic plant has at least one altered characteristic selected from accelerated flowering time, reduced plant size, reduced leaf number, and combinations thereof.

**[00145]** In other embodiments, the transgenic plant exhibits increased yield compared to a control plant. In one embodiment, the transgenic plant exhibits increased yield compared to a plant not expressing the K-domain. In still other embodiments, the transgenic plant exhibits an increased yield by at least 5%, 7%, 8% or 9%, compared to a control plant. In other embodiments, the transgenic plant exhibits increased yield by at least about 10 to 20 % compared to a control plant. In still other embodiments, the transgenic plant exhibits increased yield by about 30% to about 50% compared to a control plant.

**[00146]** In the case of any inconsistencies, the present disclosure, including any definitions herein will prevail over all patents, patent applications and literature cited in the specification. Although specific embodiments have been illustrated and described herein, it will be appreciated by those of ordinary skill in the art that any procedure that is calculated to achieve the same purpose may be substituted for the specific embodiments shown. For example, although the process has been discussed using particular sequences, variations on the sequences and from different sources may be used. Additionally, although the process has been discussed using primarily tobacco, petunia and blueberry as the plant, other plants, such as soybean, corn, for example, may be used. This application is intended to cover any adaptations or variations of the present subject matter. Therefore, it is manifestly intended that embodiments of this invention be limited only by the claims and the equivalents thereof.

WHAT IS CLAIMED IS:

1. A method of producing a transgenic plant by transformation comprising:  
introducing a biologically active K-domain polypeptide encoded by a nucleotide sequence to produce a transformed cell, wherein the nucleotide sequence is:  
the nucleotide sequence set forth in SEQ ID NO: 1;  
a nucleotide sequence encoding the polypeptide of SEQ ID NO: 2;  
a nucleotide sequence comprising at least 95% sequence identity to the full-length sequence set forth in SEQ ID NO: 1; or  
a nucleotide sequence encoding an amino acid sequence having at least 95% sequence identity to the full-length sequence set forth in SEQ ID NO: 2; and  
generating the transgenic plant from the transformed cell, wherein the transgenic plant has increased expression of the K-domain polypeptide as compared to a control plant.
2. The method of claim 1 wherein the K-domain polypeptide is a K-domain of a MADS box polypeptide.
3. The method of claim 1 wherein the plant is a fruit-bearing plant.
4. The method of claim 3 wherein the fruit-bearing plant is blueberry.
5. The method of claim 3 wherein the fruit-bearing plant is grape, orange, lemon, grapefruit, avocado, date, peach, cherry, olive, plum, coconut, apple, pear, blackberry, raspberry, strawberry, pineapple, tomato, cucumber or eggplant.
6. The method of claim 1 wherein the plant comprises a plant part which is one or more of an embryo, pollen, ovule, seed, leaf, flower, branch, fruit, kernel, ear, cob, husk, stalk, root, root tip, or an anther.
7. The method of claim 1 wherein the nucleotide sequence encoding the K-domain polypeptide is stably transformed into the plant.
8. The method of claim 7 wherein the transgenic plant has at least one characteristic which is one or more of accelerated flowering time, reduced plant size, or reduced leaf



number, relative to a control plant.

9. The method of any one of claims 1-8 wherein the transgenic plant exhibits increased seed yield compared to a control plant.

10. An isolated expression cassette for producing a transgenic plant, comprising a polynucleotide encoding a K-domain polypeptide operably linked to a promoter that drives expression in the transgenic plant, wherein the polynucleotide has:

the nucleotide sequence set forth in SEQ ID NO: 1;

a nucleotide sequence encoding the polypeptide of SEQ ID NO: 2;

a nucleotide sequence comprising at least 95% sequence identity to the full length sequence set forth in SEQ ID NO: 1; or

a nucleotide sequence encoding an amino acid sequence having at least 95% sequence identity to the full-length sequence set forth in SEQ ID NO: 2.

11. The isolated expression cassette of claim 10 wherein the promoter is a constitutive promoter.

12. A bacterial or plant host cell transformed with the expression cassette of claim 11.

13. The isolated expression cassette of claim 10 wherein the transgenic plant comprises embryo, pollen, ovule, seed, leaf, flower, branch, fruit, kernel, ear, cob, husk, stalk, root, root tips, anther, or combinations thereof.

14. The isolated expression cassette of claim 13 wherein the transgenic plant has at least one characteristic of accelerated flowering time, reduced plant size, reduced leaf number, or combinations thereof relative to a control plant.

15. The expression cassette of any one of claims 10, 11, 13 and 14, wherein the plant exhibits increased seed yield compared to a control plant.

16. The method of claim 1 wherein the plant is cotton, maize, soybean, or sugar beet.

17. The expression cassette of claim 10 wherein the plant is cotton, maize, soybean or

sugar beet.

18. A transgenic plant cell produced according to the method of any one of claims 1 to 9 or 16, wherein the transgenic plant cell comprises a K-domain polypeptide encoded by: a nucleotide sequence comprising at least 95% sequence identity to the full-length sequence set forth in SEQ ID NO: 1, or a nucleotide sequence encoding an amino acid sequence having at least 95% sequence identity to the full-length sequence set forth in SEQ ID NO: 2.

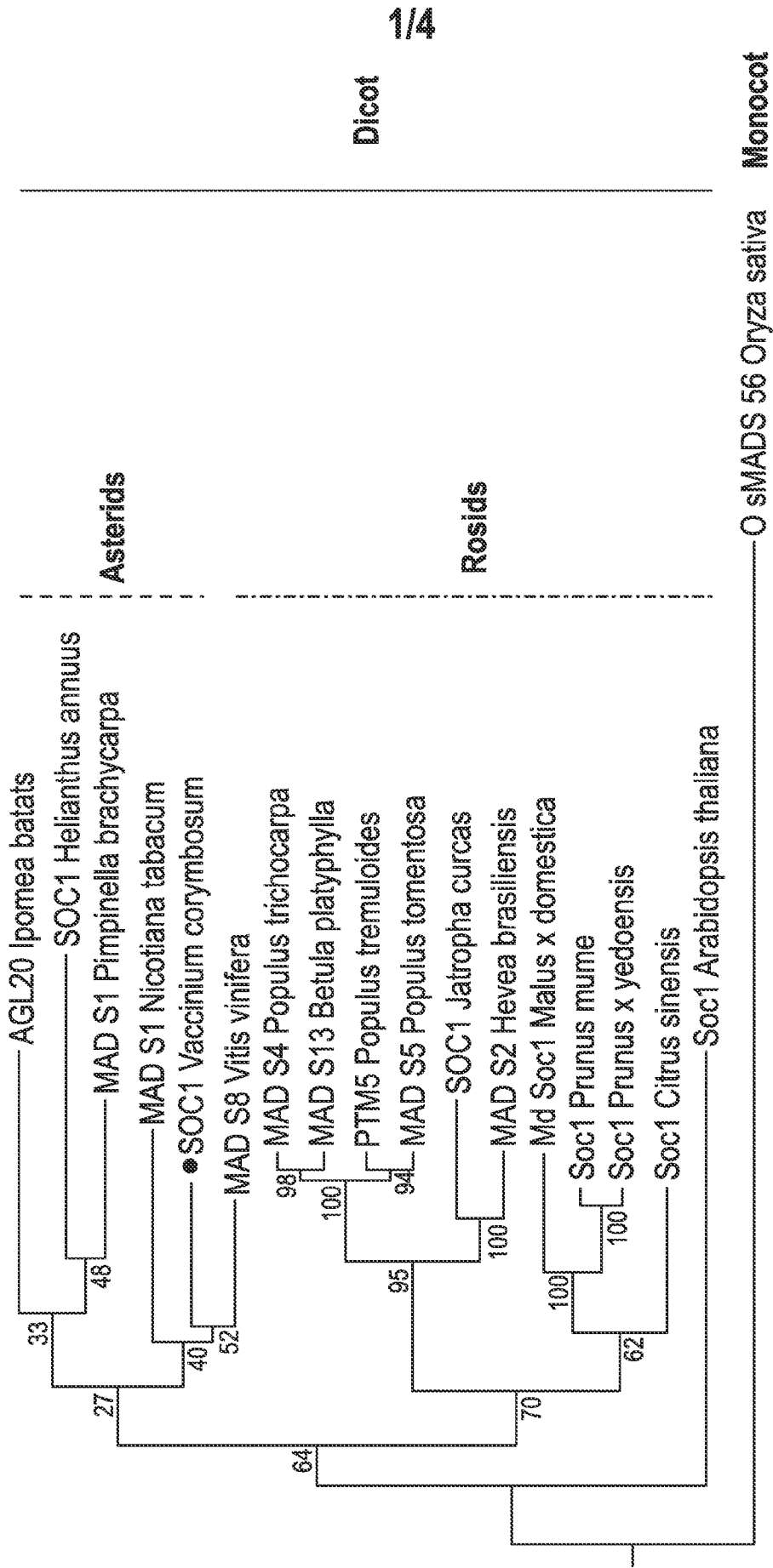


FIG. 1A

2/4

SOC1_Vaccinium_corymbosum	EQNMQHLKHEAADMSKKIEHLEVAKRKLLGEGLGSCCTFEELIQIEQQLEQ
MADS1_Nicotiana_tabacum	EQNLQHMQHAAASLMKKIELLEESKRKLLGEGLGSCSLVELQQIEKQLER
PTM5_Populus_trichocarpa	EQNMLQLKEEAASMIKKIEHLEVSKRKLLGEGCLGSCCTVEELQQIEQQLER
PTM5_Populus_tremuloides	EEQNMLQLKEEAASMIKKIEHLEVSKRKLLGEGCLGSCCTIEELQQIEQQLER
SOC1-like_Prunus-mume	DQNMQHKQESSMMKQIELLEVSKRKLLGEGLGSCSIEELQEIEQQLER
SOC1-like_Malus_x_domestica	EQNMQHLKQEATSMMKQIELLEVSKRKLLGEGLGSCSLAELQEIEDQLEK
SOC1_Arabidopsis_thaliana	EENMQHLKYEAANMMKKIEQLEASKRKLLGEGIGTCSIEELQQIEQQLEK
SOC1_Vitis_vinifera	EHNMQHLKHEAANMAKKIELLEISKRKLLGEGLGSCSIEELQQIEQQLER
SOC1-like_Citrus_sinensis	ERYMQQLKHEIANMIEKIEHIEVSQRKLLGQDLGSRTNEELQELDDQLER
	:. : ::: :.: :*: * :*:*****: : : : ** :*:***:
SOC1_Vaccinium_corymbosum	SASTIRARKMQVFREQIEKLKEKEKALEAENAML
MADS1_Nicotiana_tabacum	SVSTIRARKIQVFKEQIERLKEKEKILASENAAIL
PTM5_Populus_trichocarpa	SVSTIRARKNQVFKEQIELLRQKEKLLAAENARL
PTM5_Populus_tremuloides	SVSTIRARKNQVFKEQIELLKQKEKLLAAENARL
SOC1-like_Prunus-mume	SVSNVRARKTQVFKEQIEQLREKKGALAAENEKL
SOC1-like_Malus_x_domestica	SVYNVRARKSQVFKEQIEQLREKEKLLTAENTRL
SOC1_Arabidopsis_thaliana	SVKCIARKTQVFKEQIEQLKQKEKALAAENEKL
SOC1_Vitis_vinifera	SVSSIRARKNQVFKEQIEQLKEKEKALAAENAML
SOC1-like_Citrus_sinensis	SLRSIRARKAQLFNEQMGQLKEKERLLLEDNARL
	* :***** *:*.**: *::* : :* *

FIG. 1B

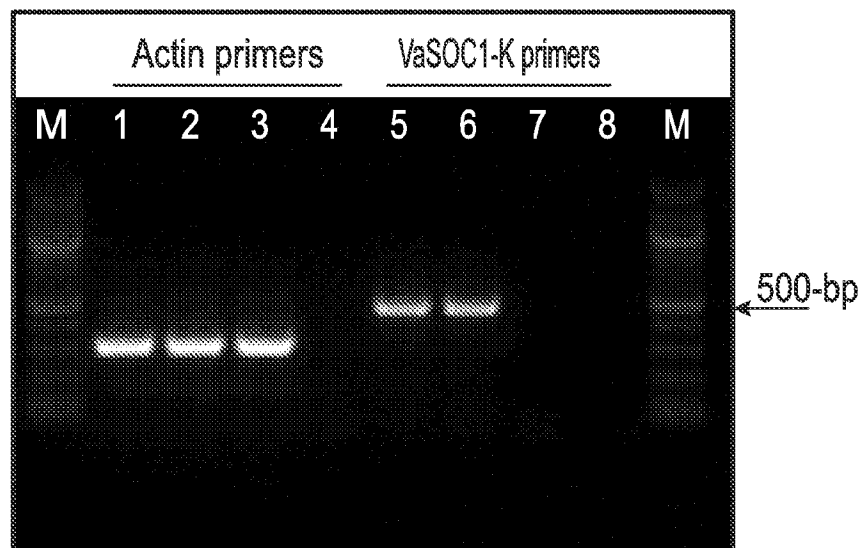


FIG. 2

3/4

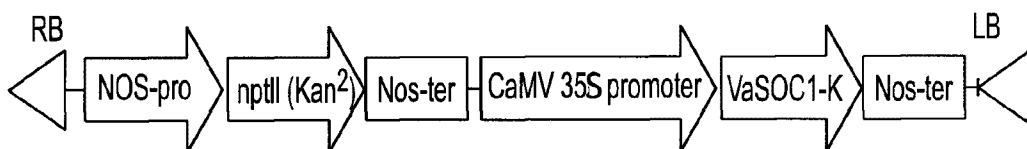


FIG. 3A

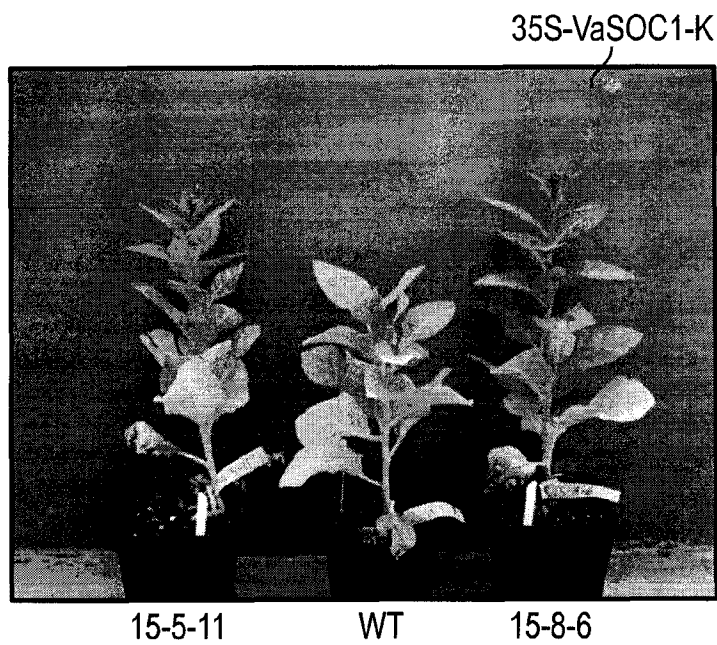


FIG. 3B

4/4

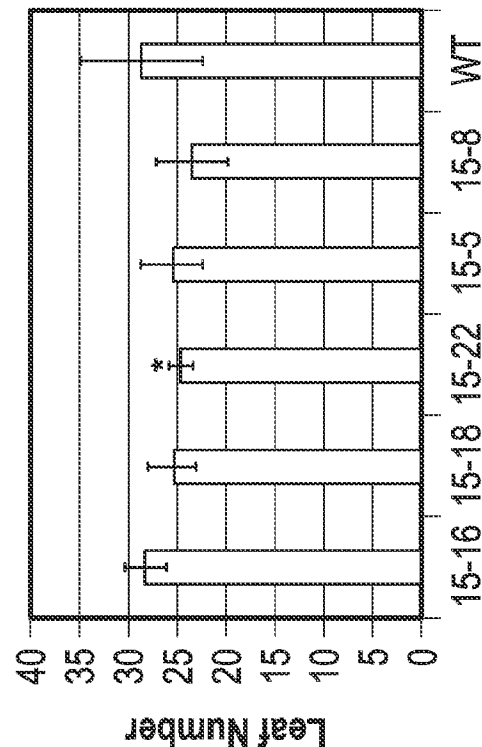


FIG. 3D

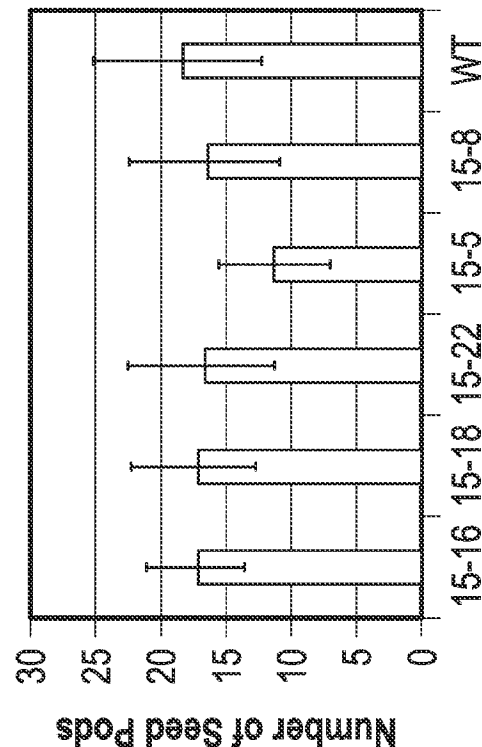


FIG. 3F

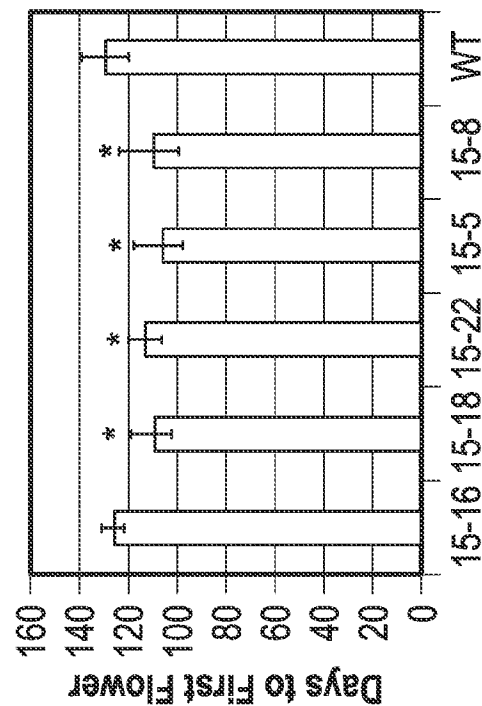


FIG. 3C

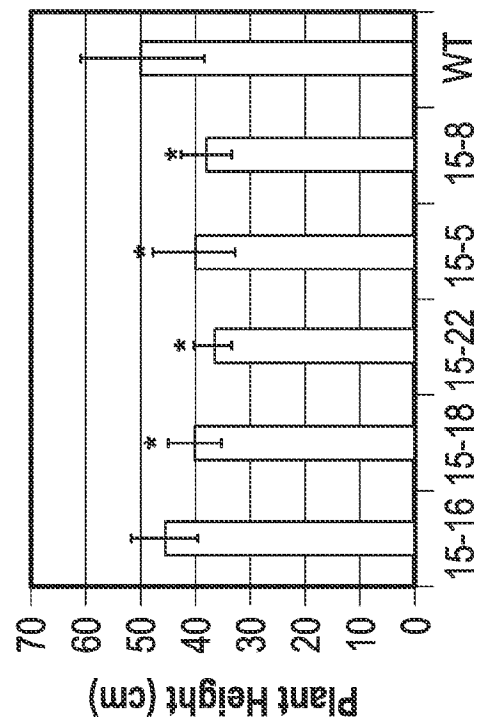


FIG. 3E