República Federativa do Brasil
Ministério da Economia
Instituto Nacional da Propriedade Industrial

(11) BR 112015032956-0 B1

(22) Data do Depósito: 01/07/2014

(45) Data de Concessão: 13/09/2022

(54) Título: MÉTODO PARA PRODUZIR UMA PLANTA TRANSGÊNICA E CASSETE DE EXPRESSÃO

(51) Int.Cl.: C12N 15/82.

(30) Prioridade Unionista: 09/07/2013 US 61/844,248.

(73) Titular(es): BOARD OF TRUSTEES OF MICHIGAN STATE UNIVERSITY.

(72) Inventor(es): GUO-QING SONG; AARON WALWORTH; DONGYAN ZHAO.

(86) Pedido PCT: PCT US2014045115 de 01/07/2014

(87) Publicação PCT: WO 2015/006105 de 15/01/2015

(85) Data do Início da Fase Nacional: 30/12/2015

(57) Resumo: MÉTODO DE PRODUÇÃO DE PLANTAS TRANSGÊNICAS, CONJUNTO DE EXPRESSÃO, CÉLULA E PLANTA TRANSGÊNICA. São descritos métodos e composições de aumento do rendimento vegetal por meio da introdução em uma planta do domínio K de um gene de caixa MADS. A expressão do domínio K fornece plantas com desenvolvimento de flores, tamanho da planta e desenvolvimento de folhas alterados.

“MÉTODO PARA PRODUZIR UMA PLANTA TRANSGÊNICA E CASSETE DE EXPRESSÃO”

PEDIDOS RELACIONADOS

[001] O presente pedido reivindica prioridade do Pedido Provisório Norte-Americano nº 61/844.248, depositado em nove de julho de 2013, e é integralmente incorporado ao presente como referência.

ANTECEDENTES DA TÉCNICA

[002] A capacidade de produção de alimentos enfrenta uma quantidade sempre crescente de desafios, incluindo a crescente população mundial, o aquecimento global e a redução da disponibilidade de terras aráveis. Além disso, o rendimento da produção é uma característica quantitativa controlada por múltiplos genes. Os inventores do presente reconhecem a necessidade de aumento do rendimento da produção por meio de métodos diferentes do cultivo tradicional.

DESCRIÇÃO RESUMIDA DA INVENÇÃO

[003] As diversas realizações descritas no presente fornecem métodos e composições para aumento do rendimento vegetal. Segundo uma realização, é fornecido um método de aumento do rendimento em plantas com relação a uma planta controle, que compreende o aumento da expressão de domínio K de um polipeptídeo MADS ou suas variantes. Em diversas realizações, o polinucleotídeo que codifica o polipeptídeo de domínio K é descrito em SEQ ID NO: 1. Em uma realização, o polipeptídeo de domínio K é descrito em SEQ ID NO: 2. Em outras realizações, é fornecido um cassete de expressão útil para a produção de plantas transgênicas. Em ainda outras realizações, são fornecidas plantas transgênicas transformadas com um polipeptídeo de domínio K.

BREVE DESCRIÇÃO DAS FIGURAS

[004] O arquivo de pedido ou patente contém pelo menos um

desenho em cores. Cópias da presente publicação de patente ou pedido de patente com desenho(s) colorido(s) serão fornecidas pelo Escritório mediante solicitação e pagamento da taxa necessária.

[005] A Fig. 1A é uma análise filogenética de sequências de nucleotídeos de proteínas similares a *SOC1* de acordo com uma realização.

[006] A Fig. 1B é um alinhamento de múltiplas sequências do domínio K de proteínas similares a *SOC1* de acordo com uma realização.

[007] A Fig. 2 é uma imagem digital de análise de transcrição do transgene *35S-domínio VcK* e um gene actina endógeno em uma planta de fumo não transgênica (NT) e duas transgênicas T0 selecionadas aleatoriamente de acordo com a presente invenção.

[008] A Fig. 3A é uma ilustração esquemática da região de T-DNA que conduz *35S-domínio VcK* de acordo com uma realização.

[009] A Fig. 3B é uma imagem digital de plantas transgênicas T1 em comparação com uma planta do tipo selvagem (WT) não transgênica 93 dias após a germinação da semente de acordo com uma realização.

[010] A Fig. 3C é um gráfico que exibe os dias para a emergência das primeiras flores após o cultivo das sementes de acordo com uma realização.

[011] A Fig. 3D é um gráfico que exibe números de folhas quando 50% das flores no conjunto central foram abertas de acordo com uma realização.

[012] A Fig. 3E é um gráfico que exibe a altura das plantas na floração (cm) de acordo com uma realização.

[013] A Fig. 3F é um gráfico que exibe o número de vagens de sementes no conjunto de flores central.

APRESENTAÇÃO DE LISTAGEM DE SEQUÊNCIAS

[014] A Listagem de Sequências associada ao presente pedido é depositada em formato eletrônico por meio do EFS-Web e integralmente incorporada ao presente relatório descritivo como referência. O nome do arquivo

de texto que contém a Listagem de Sequências é 3000-081WO1-SequenceListing_ST25.txt. O tamanho do arquivo de texto é 9 kB e o arquivo de texto foi criado em 25 de junho de 2014.

DESCRIÇÃO DETALHADA DA INVENÇÃO

[015] Na descrição detalhada a seguir, as realizações são descritas em detalhes suficientes para permitir a sua prática para os técnicos no assunto e deve-se compreender que outras realizações podem ser utilizadas e alterações químicas e de procedimentos podem ser realizadas sem abandonar o espírito e o escopo do presente objeto. A descrição detalhada a seguir não deve, portanto, ser considerada no sentido limitador e o escopo das realizações é definido somente pelas reivindicações anexas.

[016] A expressão “janela de comparação”, da forma utilizada no presente, designa relações de sequências entre duas ou mais sequências referentes a um segmento contíguo e especificado de uma sequência de polinucleotídeos ou polipeptídeos, em que a sequência na janela de comparação pode compreender adições ou exclusões (tais como lacunas) em comparação com a sequência de referência (que não compreende adições nem exclusões) para alinhamento ideal dos dois polinucleotídeos.

[017] A expressão “promotor constitutivo”, da forma utilizada no presente, designa um promotor que é ativo na transcrição durante a maior parte, mas não todas as fases de crescimento e desenvolvimento vegetal e sob a maior parte das condições experimentais e em pelo menos um órgão ou tecido celular.

[018] A expressão “controle”, “planta controle” ou “célula de planta controle”, da forma utilizada no presente, designa um ponto de referência para medir alterações do fenótipo de uma planta ou célula de planta transgênica. O “controle” pode incluir, por exemplo, (a) célula ou planta do tipo selvagem, por exemplo, do mesmo genótipo do material de partida para a alteração que resultou na célula ou planta transgênica; (b) planta ou célula vegetal do mesmo

genótipo do material de partida, mas que foi transformada com uma construção nula (por exemplo, com uma construção que não possui efeito conhecido sobre a característica de interesse, tal como uma construção que compreende um gene marcador); (c) planta ou célula vegetal não transformada com o nucleotídeo ou gene de interesse; (d) planta ou célula vegetal que é um segregante não transformado entre a prole de uma planta ou célula vegetal transgênica; (e) planta ou célula vegetal geneticamente idêntica à planta ou célula vegetal transgênica, mas que não é exposta a condições ou estímulos que induziriam a expressão do nucleotídeo ou gene de interesse; ou (f) a própria planta ou célula vegetal transgênica, sob condições nas quais o nucleotídeo ou gene de interesse não é expresso.

[019] O termo “fragmento”, da forma utilizada no presente, designa uma parte do polinucleotídeo ou uma parte da sequência de aminoácidos. Fragmentos de um polinucleotídeo podem codificar fragmentos de proteínas que retêm a atividade biológica da proteína nativa e, portanto, regulam a transcrição.

[020] O termo “grão”, da forma utilizada no presente, designa uma semente madura produzida por agricultores comerciais para propósitos diferentes de cultivo ou reprodução da espécie.

[021] O termo “heterólogo”, da forma utilizada no presente, designa uma sequência que se origina de uma espécie exógena ou, se da mesma espécie, é modificada da sua forma nativa na sua composição, local genômico ou ambos, por meio de intervenção humana deliberada.

[022] O termo “introdução”, da forma utilizada no presente, designa sequências que apresentam a uma planta o polinucleotídeo ou polipeptídeo de forma que a sequência tenha acesso ao interior de uma célula da planta.

[023] O termo “isolado” ou “purificado”, da forma utilizada no presente, designa um polinucleotídeo ou polipeptídeo, proteína ou sua parte biologicamente ativa, removida ou livre de componentes que normalmente

acompanham ou interagem com o polinucleotídeo, polipeptídeo ou proteína da forma encontrada no seu ambiente de ocorrência natural. Portanto, uma proteína ou polinucleotídeo isolado ou purificado inclui um polinucleotídeo ou polipeptídeo produzido por meio de métodos recombinantes, removido ou livre de precursores químicos ou outras substâncias quando quimicamente sintetizado.

[024] O termo “nativo”, da forma utilizada no presente, designa sequências que contêm uma sequênica de nucleotídeos ou sequência de aminoácidos de ocorrência natural.

[025] A expressão “ligado operacionalmente”, da forma utilizada no presente, designa um cassete de expressão para expressão vegetal de uma ligação funcional entre dois ou mais elementos.

[026] O termo “ortólogo”, da forma utilizada no presente, designa um gene de polipeptídeo ou polinucleotídeo derivado de um gene ancestral comum e encontrado em diferentes espécies como resultado da especiação.

[027] O termo “sobre-expressão” ou “aumento da expressão”, da forma utilizada no presente, designa qualquer forma de expressão adicional ao nível de expressão do tipo selvagem original.

[028] A expressão “percentual de identidade de sequências”, da forma utilizada no presente, designa um valor de polinucleotídeo determinado por meio da comparação de duas sequências idealmente alinhadas ao longo de uma janela de comparação, em que a porção da sequência de polinucleotídeos na janela de comparação pode incluir adições ou exclusões (por exemplo, lacunas) em comparação com a sequência de referência (o que não inclui adições ou exclusões) para alinhamento ideal das duas sequências.

[029] O termo “planta”, da forma utilizada no presente, designa células vegetais, protoplastas vegetais, cultivos de tecidos de células vegetais dos quais as plantas podem ser regeneradas, calos de plantas, conjuntos de plantas e células vegetais que são intactas em plantas ou partes de plantas tais

como embriões, pólen, óvulos, sementes, folhas, flores, ramos, frutos, caroços, espigas, sabugos, cascas, hastes, raízes, pontas de raízes, anteras e similares.

[030] O termo “polinucleotídeo”, da forma utilizada no presente, designa polinucleotídeos que compreendem DNA. O termo inclui ribonucleotídeos e combinações de ribonucleotídeos e desoxirribonucleotídeos. Esses desoxirribonucleotídeos e ribonucleotídeos incluem moléculas de ocorrência natural e análogos sintéticos. Os polinucleotídeos também englobam todas as formas de sequências, incluindo, mas sem limitações, formas de fita simples, formas de fita dupla, grampos de cabelo, estruturas de haste e laço e similares.

[031] A expressão “sequência de referência”, da forma utilizada no presente, designa relacionamentos de sequência entre duas ou mais sequências referentes a um subconjunto ou uma sequência especificada completa como base para comparação de sequências; por exemplo, como segmento de um cDNA com comprimento total ou sequência genética, ou a sequência genética ou cDNA completo.

[032] As expressões “identidade de sequências” ou “identidade”, da forma utilizada no presente, designam relacionamentos entre dois ou mais polinucleotídeos ou polipeptídeos referentes aos resíduos nas duas sequências que são idênticos quando alinhadas para correspondência máxima ao longo de uma janela de comparação especificada.

[033] A expressão “transformação estável”, da forma utilizada no presente, designa nucleotídeos introduzidos em plantas que integram o genoma da planta e podem ser herdados pela sua prole.

[034] As expressões “condições estringentes” ou “condições de hibridização estringentes”, da forma utilizada no presente, designam condições sob as quais uma sonda (sequências de nucleotídeos) se hibridizará na sua sequência alvo em grau detectavelmente maior que em outras sequências (por

exemplo, pelo menos duas vezes sobre o fundo).

[035] A expressão “promotor específico de tecidos”, da forma utilizada no presente, designa um promotor que é ativo na transcrição em certos órgãos ou tecidos.

[036] O termo “característica”, da forma utilizada no presente, designa um fenótipo derivado de uma sequência específica ou de grupos de sequências.

[037] O termo “transgênico”, da forma utilizada no presente, designa a ocorrência de uma alteração de planta ou parte de planta, tal como a transformação ou introdução de um polinucleotídeo ou polipeptídeo, ou é uma planta ou célula vegetal descendente de planta ou célula alterada desta forma e que detém a alteração.

[038] A expressão “transformação transitória”, da forma utilizada no presente, designa um polinucleotídeo, polipeptídeo ou proteína introduzida em uma planta e não integra o genoma da planta.

[039] O termo “variantes”, da forma utilizada no presente, designa sequências que possuem sequências substancialmente similares em comparação com uma primeira sequência. Uma variante pode incluir, mas sem limitações, sua exclusão, inserção, substituição ou combinação de um ou mais nucleotídeos ou aminoácidos em um ou mais locais internos dentro do polipeptídeo ou polinucleotídeo nativo ou de referência, respectivamente.

[040] O termo “rendimento”, da forma utilizada no presente, designa uma quantidade de safra que foi colhida por unidade de área terrestre ou a quantidade de sementes geradas a partir da própria planta ou uma de suas combinações.

[041] São descritos métodos e composições de aumento do rendimento vegetal por meio de expressão da parte de domínio K de um gene, tal como gene de caixa MADS. O domínio K de um gene de caixa MADS, quando

expresso em plantas, pode aumentar o rendimento vegetal em comparação com uma planta que não expressa ou que expressa níveis mais baixos do domínio K. As moléculas podem ser obtidas por meio de qualquer método conveniente, sejam elas isoladas de plantas ou produzidas sinteticamente, por exemplo. Em uma realização, são fornecidos métodos e composições para acelerar o desenvolvimento de órgãos florais, aumentar a quantidade de flores, reduzir o tamanho das plantas, reduzir a quantidade de folhas ou suas combinações, sem prejudicar a produção de frutas ou sementes. As diversas realizações fornecem métodos e composições que resultam em aumento do rendimento de plantas ou safras aumentando a densidade vegetal no campo, reduzindo o período de crescimento das plantas ou ambos.

[042] As diversas realizações fornecem composições que incluem os polipeptídeos e polinucleotídeos de domínio K e suas variantes ou fragmentos dos genes de caixa MADS que estão envolvidos na aceleração da floração, redução do tamanho das plantas ou número de folhas ou suas combinações sem afetar o desenvolvimento das frutas ou sementes. Em uma realização, o supressor da sobre-expressão de constantes 1 (SOC) e similares codificam uma proteína vegetal com domínio MADS e domínio K. Fatores de transcrição de domínio MADS desempenham papéis essenciais na evolução de plantas em floração por meio da determinação da arquitetura floral e regulagem do desenvolvimento das flores. Genes de caixa MADS contêm um motivo de caixa MADS conservado e são frequentemente classificados em subfamílias do tipo I e tipo II e, em caixa MADS do tipo II, os genes frequentemente possuem domínios Terminais K, I e C (carbóxi). Com base nas variações nos seus domínios K (queratina) e I (de intervenção), os genes de caixa MADS do tipo II são adicionalmente divididos em subgrupos $MIKC^{nc}$ e $MIKC^{c}$ (MIKC clássico), em que este último frequentemente possui domínio K mais curto que o de $MIKC^{nc}$ e desempenha papéis específicos no modelo ABC de desenvolvimento floral. O

subgrupo MIKC^c consiste de treze conjuntos genéticos principais, dos quais doze foram identificados e extensamente estudados em *Arabidopsis thaliana*. Análises de genomas dos genes de caixa MADS do tipo MIKC^c foram relatadas em diversas espécies vegetais, incluindo *Arabidopsis*, tomate (*Solanum lycopersicon*), álamo (*Populus trichocarpa*), arroz (*Oryza sativa*), uva (*Vitis vinifera*) e pepino (*Cucumis sativus*). Dois conjuntos genéticos, *LOCAL DE FLORAÇÃO C* (*FLC*) e domínio TM3/K, são componentes importantes no processo do tempo de floração. Os genes do conjunto FLC, incluindo *FLC* e genes similares a *FLC*, são reguladores centrais da floração mediada por vernalização. *SOC1* e genes similares a *SOC1* no conjunto de domínio TM3/K são importantes integrantes do processo de floração, harmonizando sinais de floração de diversos processos. Embora tenha havido muitos relatos de sobre-expressão de genes de caixa MADS do tipo MIKC^c intactos, pouco foi feito sobre os domínios separados dessas proteínas MADS.

[043] Em uma realização, é descrito um domínio K isolado de um gene de caixa MADS (por exemplo, gene similar a *SOC1*). Em uma realização, o gene de caixa MADS é de *Vaccinium corymbosum* (mirtilo - *blueberry*) (*domínio VcK*). A sobre-expressão de *domínio VcK* isolado pode promover a floração precoce, tamanho de planta ou número de folhas reduzidos, ou suas combinações, sem reduzir nem afetar a produção de frutas ou sementes. A sequência de nucleotídeos de *domínio K* é exibida como SEQ ID NO: 1, o polipeptídeo de domínio K é exibido como SEQ ID NO: 2 e o gene de caixa MADS com comprimento total do gene *SOC1* é exibido como SEQ ID NO: 3. Em outras realizações, outros domínios K de genes de caixa MADS envolvidos em processos de floração podem ser utilizados. Em diversas realizações, pode-se utilizar *FOLHOSO* (*LFY*) (Acesso: AF466801.1), *FLOR TERMINAL 1* (*TFL1*) (Acesso: AF466804.1), *APETALA1* (*AP1*) (Acesso: NM_105581.2), similar a *LFY* (contíguo 12297), similar a *TFL1* (contíguo 02213) e similar a *AP1*.

[044] Fragmentos que são úteis como sondas de hibridização geralmente não codificam proteínas de fragmentos que retêm atividade biológica. Em diversas realizações, portanto, fragmentos de uma sequência de nucleotídeos podem variar de pelo menos cerca de dezoito nucleotídeos, cerca de vinte nucleotídeos, cerca de cinquenta nucleotídeos, cerca de cem nucleotídeos até o polinucleotídeo com comprimento total que codifica as proteínas. Um fragmento de polipeptídeo de caixa MADS pode codificar ampla variedade de aminoácidos contíguos, incluindo, mas sem limitações, pelo menos 15, 25, 30, 50, 100, 150, 200, 250, 300 ou mais aminoácidos contíguos, incluindo qualquer valor ou faixa entre eles, ou até a quantidade total de aminoácidos presentes em um domínio MADS com comprimento total ou proteína de domínio K (por exemplo, SEQ ID NO: 2). Em diversas realizações, um polinucleotídeo de domínio K ou fragmentos de domínio MADS que inclui um polinucleotídeo de domínio K úteis como sondas de hibridização, primers de PCR, geralmente não necessitam codificar uma parte biologicamente ativa de uma proteína de domínio K.

[045] Em algumas realizações, uma parte biologicamente ativa de um polipeptídeo que possui domínio K pode ser preparada por meio de isolamento de uma parte de um polinucleotídeo de domínio K que expressa a parte codificada da proteína de domínio K (por exemplo, por meio de expressão recombinante *in vitro*) e determinação da atividade da proteína codificada do domínio K. Em várias realizações, polinucleotídeos que são fragmentos de um gene de caixa MADS ou uma sequência de polinucleotídeos que possui domínio K contêm pelo menos 16, 20, 50, 75, 100, 150, 200, 250, 300, 350, 400, 450 ou 500 nucleotídeos contíguos, ou até a quantidade de nucleotídeos presentes em um domínio K com comprimento total (por exemplo, SEQ ID NO: 1).

[046] Variantes alélicas de ocorrência natural podem ser identificadas utilizando métodos de biologia molecular bem conhecidos, como,

em algumas realizações, com métodos de hibridização e reação em cadeia de polimerase (PCR) conforme detalhado abaixo. Em diversas realizações, polinucleotídeos variantes também incluem polinucleotídeos derivados sinteticamente, tais como os gerados pelo uso de mutagênese dirigida a local, mas que ainda codificam um polipeptídeo que possui domínio K capaz de regular o tempo de floração, números de flores, tamanho de plantas, número de folhas ou suas combinações, mas sem afetar o desenvolvimento de frutas ou sementes. Em diversas realizações, variantes de um polinucleotídeo específico podem ter identidade de sequências de pelo menos cerca de 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% ou mais com aquele polinucleotídeo específico, conforme determinado por programas de alinhamento de sequências e parâmetros descritos no presente.

[047] Variantes de um polinucleotídeo específico (tal como um polinucleotídeo de referência) podem também ser avaliados por meio de comparação do percentual de identidade de sequências entre o polipeptídeo codificado por um polinucleotídeo variante e o polipeptídeo codificado pelo polinucleotídeo de referência. Em diversas realizações, é descrito um polinucleotídeo isolado que codifica um polipeptídeo com um dado percentual de identidade de sequências para o polipeptídeo de SEQ ID NO: 2. O percentual de identidade de sequências entre quaisquer dois polipeptídeos pode ser calculado utilizando programas de alinhamento de sequências e parâmetros descritos em outros pontos do presente. Em algumas realizações, a janela de comparação possui comprimento de pelo menos vinte nucleotídeos contíguos e, opcionalmente, pode ter 30, 40, 50, 100, incluindo qualquer valor entre eles, ou mais. Os técnicos no assunto compreendem que, para evitar alta similaridade com uma sequência de referência devido à inclusão de intervalos na sequência de nuclelotídeos, é tipicamente introduzida uma penalidade de intervalo, que é subtraída da quantidade de coincidências. Quando qualquer dado par de

polinucleotídeos for avaliado por meio de comparação com o percentual de identidade de sequências compartilhado pelos dois polipeptídeos que codificam, o percentual de identidade de sequências entre os dois polipeptídeos codificados em algumas realizações é de pelo menos cerca de 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% ou mais.

[048] As proteínas variantes englobadas são biologicamente ativas, ou seja, elas continuam a possuir a atividade biológica desejada da proteína nativa; nomeadamente, elas regulam a transcrição e resultam em fenótipos que afetam o tempo da floração, o tamanho da planta e o número de folhas sem afetar a produção de frutas ou sementes. Em diversas realizações, essas variantes podem resultar de polimorfismo genético ou manipulação humana. Em várias realizações, variantes biologicamente ativas de domínio K apresentarão identidade de sequências de pelo menos cerca de 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% ou 99%, incluindo qualquer faixa entre eles ou mais, com o domínio *VcK* de aminoácido SEQ ID NO: 2 conforme determinado por programas de alinhamento de sequências e parâmetros descritos no presente. Em diversas realizações, uma variante biologicamente ativa de domínio K pode diferir daquela proteína em até 1-15 resíduos de aminoácidos, até 1-10, tal como 6-10, incluindo qualquer faixa entre eles, até 5, 4, 3, 2 ou mesmo 1 resíduo de aminoácido.

[049] Em diversas realizações, os polinucleotídeos podem ser alterados de diversas formas, incluindo substituições, exclusões, truncagens e inserções de aminoácidos, utilizando diversos métodos geralmente conhecidos na técnica. Em várias realizações, variantes de sequências de aminoácidos de domínio K ou fragmentos da proteína de caixa MADS podem ser preparados por meio de mutações no DNA por meio de uma série de métodos conhecidos de mutagênese e alterações de polinucleotídeos. Vide, por exemplo, Kunkel (1985),

Proc. Natl. Acad. Sci. U. S. A. 82: 488-492; Kunkel et al (1987), *Methods in Enzymol.* 154: 367-382; Patente Norte-Americana nº 4.873.192; Walker e Gaastra, eds. (1983), *Techniques in Molecular Biology* (MacMillan Publishing Company, Nova Iorque) e referências ali mencionadas. Orientações sobre as substituições de aminoácidos apropriadas que não afetam a atividade biológica da proteína de interesse podem ser encontradas no modelo de Dayhoff et al (1978), *Atlas of Protein Sequence and Structure* (Natl. Biomed. Res. Found., Washjngton DC), incorporado ao presente como referência. Em algumas realizações, são descritas substituições conservadoras, tais como o intercâmbio de um aminoácido por outro que possui propriedades similares.

[050] Em várias realizações, genes e polinucleotídeos incluem as sequências de ocorrência natural bem como formas mutantes. De forma similar, em outras realizações, as proteínas englobam proteínas de ocorrência natural, bem como suas variações e formas modificadas. Essas variantes continuarão a possuir a atividade desejada (por exemplo, ligação de transcrição que, quando sobre-expressa, pode alterar o desenvolvimento das flores, tamanho de plantas, número de plantas ou suas combinações sem afetar o desenvolvimento de frutas ou sementes). Em várias realizações, as mutações que podem ser realizadas no DNA que codifica uma variante não colocam a sequência fora da estrutura de leitura, nem criam regiões complementares que poderiam produzir estrutura de mRNA secundária. Vide DeBoer et al, Pedido de patente EP publicado nº 75.444A (1983).

[051] Em diversas realizações, as exclusões, inserções e substituições das sequências de proteína englobadas no presente deverão ainda manter as características da atividade de proteína. Quando é difícil prever o efeito exato da substituição, exclusão ou inserção antes da sua realização, entretanto, os técnicos no assunto apreciarão que o efeito será avaliado por meio de testes de seleção de rotina. Em diversas realizações, podem ser utilizados

vários métodos para testar a presença do domínio K, incluindo o monitoramento direto do nível de expressão de um gene alvo no nível de nucleotídeo ou polipeptídeo. Métodos dessa análise são conhecidos e incluem, por exemplo, Northern Blots, testes de proteção, Western Blots e testes enzimáticos ou colorimétricos. Em várias realizações, métodos de teste de modulação da atividade de transcrição podem incluir o monitoramento de alterações do fenótipo da planta. Em algumas realizações, a modulação do nível do polipeptídeo de domínio K pode resultar em alterações da formação de flores, tempo, altura da planta e número de folhas.

[052] Em diversas realizações, proteínas e polinucleotídeos variantes também englobam sequências e proteínas derivadas de um procedimento mutagênico e recombinogênico tal como comutação de DNA. Com esse procedimento, uma ou mais sequências de codificação de domínio K diferentes podem ser manipuladas para criar uma nova sequência de domínio K que possui as propriedades desejadas. Desta forma, são geradas bibliotecas de polinucleotídeos recombinantes a partir de uma população de polinucleotídeos de sequência relacionados que possuem regiões de sequência que possuem identidade de sequência substancial e podem ser recombinados de forma homóloga *in vitro* ou *in vivo*. Em diversas realizações, utilizando esta abordagem, motivos de sequência que codificam um domínio de interesse podem sofrer comutação entre o domínio K e outros domínios K de outros genes de caixa MADS conhecidos para obter um novo gene que codifica uma proteína com propriedade de interesse aprimorada, tal como tempo de floração precoce, redução da altura da planta ou redução do número de folhas ou suas combinações. Estratégias dessa comutação de DNA são conhecidas na técnica. Vide, por exemplo, Stemmer (1994), *Proc. Natl. Acad. Sci. U. S. A.* 91: 10747-10751; Stemmer (1994), *Nature* 370: 389-391; Crameri 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. U. S. A.* 94: 4504-4509; Crameri et al (1998), *Nature* 391: 288-291; e Patentes Norte-Americanas nº 5.605.793 e 5.837.458.

[053] Em algumas realizações, os polinucleotídeos podem ser utilizados para isolar sequências correspondentes de outros organismos, outras plantas ou outras plantas lenhosas. Desta forma, métodos tais como PCR, hibridização e similares podem ser utilizados para identificar essas sequências com base na sua homologia de sequências com as sequências descritas no presente. Sequências isoladas com base na sua identidade de sequências com as sequências de domínio K completas, descritas no presente ou com suas variantes e fragmentos são englobadas no presente. Em diversas realizações, essas sequências incluem sequências ortólogas das sequências descritas. Funções de ortólogos frequentemente são altamente conservadas entre as espécies. Em várias realizações, genes encontrados em espécies diferentes são considerados ortólogos quando suas sequências de nucleotídeos e/ou suas sequências de proteínas codificadas compartilham identidade de sequências de pelo menos 60%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% ou 99%, incluindo qualquer faixa ou valor entre eles ou mais. Portanto, polinucleotídeos isolados que podem causar floração precoce, redução do tamanho de plantas ou menos folhas sem afetar a produção de sementes ou frutas por planta e que codificam um domínio K e hibridizam-se sob condições estringentes nas sequências de domínio K descritas ou suas variantes e fragmentos são englobados no presente.

[054] Em abordagem de PCR, primers de oligonucleotídeos podem ser projetados para uso em reações de PCR para amplificar sequências de DNA correspondentes de cDNA ou DNA genômico extraído de qualquer planta de interesse. Métodos de projeto de primers de PCR e clonagem de PCR são geralmente conhecidos na técnica e descritos em Sambrook et al (1989), *Molecular Cloning: A Laboratory Manual* (segunda edição, Cold Spring Harbor

Laboratory Press, Plainview NY). Vide também Innis et al, eds. (1990), *PCR Protocols: A Guide to Methods and Applications* (Academic Press, Nova Iorque); Innis e Gelfand, Eds. (1995), *PCR Strategies* (Academic Press, Nova Iorque); e Innis e Gelfand, Eds. (1999), *PCR Methods Manual* (Academic Press, Nova Iorque). Em algumas realizações, métodos de PCR incluem, mas sem limitações, métodos que utilizam primers emparelhados, primers reunidos, primers específicos isolados, primers degenerados, primers específicos de genes, primers específicos de vetores, primers parcialmente emparelhados e similares.

[055] Em métodos de hibridização, utiliza-se um polinucleotídeo conhecido, no todo ou em parte, como sonda que se hibridiza seletivamente em outros polinucleotídeos correspondentes presentes em uma população de fragmentos de DNA genômicos clonados ou fragmentos de cDNA (tais como bibliotecas de cDNA ou genômicas) de um organismo selecionado. Em várias realizações, as sondas de hibridização podem ser fragmentos de DNA genômicos, fragmentos de cDNA, fragmentos de RNA ou outros oligonucleotídeos. Em ainda outras realizações, as sondas de hibridização podem ser marcadas com um grupo detectável tal como P^{32} ou qualquer outro marcador detectável. Em várias realizações, sondas de hibridização podem ser elaboradas por meio de marcação de oligonucleotídeos sintéticos com base no polinucleotídeo de domínio K. Métodos de preparação de sondas para hibridização e construção de bibliotecas genômicas e cDNA são geralmente conhecidos na técnica e descritos, por exemplo, em Sambrook et al (1989), *Molecular Cloning: A Laboratory Manual* (segunda edição, Cold Spring Harbor Laboratory Press, Plainview NY).

[056] Em várias realizações, o polinucleotídeo de domínio de caixa MADS completo ou uma ou mais de suas partes ou um polinucleotídeo que codifica um domínio K descrito no presente podem ser utilizados como primer capaz de hibridizar-se especificamente em polinucleotídeo de domínio K

correspondente e RNAs mensageiros. Para atingir hibridização específica sob uma série de condições, essas sondas incluem sequências que são exclusivas entre sequências de polinucleotídeos de domínio K. Em várias realizações, as sondas incluem pelo menos cerca de dez nucleotídeos de comprimento e pelo menos cerca de vinte nucleotídeos de comprimento. Esses pares de primers podem ser utilizados para amplificar polinucleotídeo de domínio K correspondente a partir de uma planta selecionada por meio de PCR. Este método pode ser utilizado para isolar sequências de codificação adicionais de uma planta desejada ou um teste de diagnóstico para determinar a presença de sequências de codificação em plantas. Os métodos de hibridização incluem a seleção por hibridização de bibliotecas de DNA em placas (sejam elas placas ou colônias; vide, por exemplo, Sambrook et al (1989), *Molecular Cloning: A Laboratory Manual* (segunda edição, Cold Spring Harbor Laboratory Press, Plainview NY).

[057] A hibridização dessas sequências (tais como Southern Blot e Northern BXlot) pode ser conduzida sob condições estringentes. Condições estringentes são dependentes de sequências e serão diferentes em diferentes circunstâncias. Controlando-se a estringência da hibridização, condições de lavagem ou ambas, pode-se identificar sequências alvo que são 100% complementares à sonda (sondagem homóloga). Em outros casos, as condições de estringência podem ser ajustadas para permitir alguma falha de coincidência de sequências, de forma que sejam detectados graus de similaridade mais baixos (sondagem heteróloga). Em várias realizações, uma sonda possui menos de cerca de mil nucleotídeos de comprimento, menos de quinhentos nucleotídeos de comprimento, incluindo qualquer faixa ou valor entre eles.

[058] Em várias realizações, condições estringentes serão aquelas em que a concentração de sal é de menos de cerca de 1,5 M de íons de Na, concentração de íons de Na de cerca de 0,01 a 1,0 M (ou outros sais) sob pH

7,0 a 8,3, incluindo qualquer faixa ou valor entre eles para a concentração de Na e pH e a temperatura é de pelo menos 20-35 ºC para sondas curtas (por exemplo, dez a cinquenta nucleotídeos) e pelo menos cerca de 50 a 60 ºC para sondas longas (por exemplo, mais de cinquenta nucleotídeos). Condições estringentes podem também ser atingidas com a adição de agentes desestabilizantes tais como formamida. Em várias realizações, condições de baixa estringência incluem hibridização com uma solução tampão de 30 a 35% de formamida, 1 M de NaCl, 1% de SDS (dodecil sulfato de sódio) a 37 ºC e lavagem em 1X a 2X SSC (20X SSC = 3,0 M de NaCl/0,3 M de citrato trissódico) a 50 até 55 ºC. Exemplos de condições de estringência moderada incluem hibridização em 40 a 45% de formamida, 1,0 M de NaCl, 1% de SDS a 37 ºC e lavagem em 0,5X a 1X SSC a 55 até 60 ºC. Em outras realizações, condições de alta estringência incluem hibridização em 50% de formamida, 1 M de NaCl, 1% de SDS a 37 ºC e lavagem em 0,1X SSC a 60 até 65 ºC. Em várias realizações, tampões de lavagem podem incluir cerca de 0,1% a cerca de 1% de SDS. Em diversas realizações, a duração da hibridização é de menos de cerca de 24 horas, cerca de quatro a cerca de doze horas. A duração do tempo de lavagem será pelo menos um período de tempo suficiente para atingir o equilíbrio.

[059] A especificidade é tipicamente função de lavagens após a hibridização, em que os fatores úteis são a resistência iônica e a temperatura da solução de lavagem final. Para híbridos de DNA-DNA, pode-se calcular a T_m aproximada a partir da equação de Meinkoth e Wahl (1984), *Anal. Biochem.* 138: 267-284; T_m = 81,5 ºC + 16,6 (log M) + 0,41 (% GC) – 0,61 (% form.) – 500/L; em que M é a molaridade de cátions monovalentes, % GC é o percentual de nucleotídeos de guanosina e citosina no DNA, % form é o percentual de formamida na solução de hibridização e L é o comprimento do híbrido em pares de bases. A T_m é a temperatura (sob pH e resistência iônica definidos) em que

50% de uma sequência alvo complementar hibridizam-se em uma sonda perfeitamente coincidente. A T_m é reduzida em cerca de 1 ºC para cada 1% de falta de coincidência; portanto, T_m, hibridização e/ou condições de lavagem podem ser ajustadas para hibridizar-se em sequências com a identidade desejada. Em várias realizações, caso sejam procuradas sequências com intensidade de 90%, a T_m pode ser reduzida em cerca de 10 ºC. Em várias realizações, são selecionadas condições rigorosas em cerca de 5 ºC abaixo do ponto de fusão térmica (T_m) para a sequência específica e seu complemento em pH e resistência iônica definida. Em várias realizações, condições severamente rigorosas podem empregar hibridização e/ou lavagem a 1, 2, 3 ou 4 ºC, incluindo qualquer faixa ou valor entre elas abaixo do ponto de fusão térmica (T_m). Em várias realizações, condições moderadamente rigorosas podem utilizar hibridização e/ou lavagem a 6, 7, 8, 9 ou 10 ºC menos que o ponto de fusão térmica (T_m); condições pouco rigorosas podem utilizar hibridização e/ou lavagem a 11, 12, 13, 14, 15 ou 20 ºC, incluindo qualquer faixa ou valor entre elas abaixo do ponto de fusão térmica (T_m). Utilizando a equação, composições de hibridização e lavagem e T_m desejada, os técnicos comuns no assunto compreenderão que variações das exigências de soluções de lavagem e/ou hibridização são inerentemente descritas. Caso o grau desejado de falta de coincidência resulte em T_m de menos de 45 ºC (solução aquosa) ou 32 ºC (solução de formamida), em várias realizações para aumentar a concentração de SSC, pode-se utilizar temperatura mais alta. Extensa orientação sobre a hibridização de ácidos nucleicos pode ser encontrada em Tijssen (1993), *Laboratory Techniques in Biochemistry and Molecular Biology-Hybridization with Nucleic Acid Probes*, Parte I, Capítulo 2 (Elsevier, Nova Iorque); e Ausubel et al, eds. (1995), *Current Protocols in Molecular Biology*, Capítulo 2 (Greene Publishing e Wiley-Interscience, Nova Iorque). Vide Sambrook et al (1989), *Molecular Cloning: A Laboratory Manual* (segunda edição, Cold Spring Harbor

Laboratory Press, Plainview NY).

[060] Métodos de alinhamento de sequências de proteínas ou nucleotídeos (DNA e RNA) para comparação são bem conhecidos na técnica. A determinação do percentual de identidade de sequências entre duas sequências quaisquer pode, portanto, ser realizada utilizando um algoritmo matemático. Em várias realizações, uma parte da sequência de polinucleotídeos em uma janela de comparação inclui adições ou exclusões (por exemplo, intervalos) em comparação com a sequência de referência (que não inclui adições nem exclusões) para alinhamento ideal das duas sequências.

[061] Em várias realizações, o alinhamento de sequências é realizado por meio do método de probabilidades máximas com base no modelo Tamura-Nei, utilizando o programa MEGA 5. Em várias realizações, algoritmos matemáticos são o algoritmo de Myers e Miller (1988), *CABIOS* 4: 11-17; o algoritmo de alinhamento local de Smith et al (1981), *Adv. Appl. Math.* 2: 482; o algoritmo de alinhamento global de Needleman e Wunsch (1970), *J. Mol. Biol.* 48: 443-453; o método de alinhamento de busca por local de Pearson e Lipman (1988), *Proc. Natl. Acad. Sci. U. S. A.* 85: 244-2448; o algoritmo de Karlin e Altschul (1990), *Proc. Natl. Acad. Sci. U. S. A.* 872264, modificado como em Karlin e Altschul (1993), *Proc. Natl. Acad. Sci. U. S. A.* 90: 5873-5877.

[062] Implementações computadorizadas desses algoritmos matemáticos podem ser utilizadas para comparação de sequências, para determinar a identidade de sequências. Essas implementações incluem, mas sem limitações: CLUSTAL no programa PC/Gene (disponível por meio da Intelligenetics, Mountain View, Califórnia); o programa ALIGN (Versão 2.0) e GAP, BESTFIT, BLAST, FASTA e TFASTA no Pacote de Software Genético GCG Wisconsin, Versão 10 (disponível por meio da Accelrys Inc., 9685 Scranton Road, San Diego, Califórnia, Estados Unidos). Alinhamentos utilizando esses programas podem ser realizados utilizando os parâmetros padrão. O programa

CLUSTAL é bem descrito por 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; e Pearson et al (1994), *Meth. Mol. Biol.* 24: 307-331. O programa ALIGN é baseado no algoritmo de Myers e Miller (1988), acima. Pode-se utilizar tabela de resíduos de peso PAM120, penalidade de comprimento de intervalo de 12 e penalidade de intervalo de 4 com o programa ALIGN ao comparar-se sequências de aminoácidos. Os programas BLAST de Altschul et al (1990), *J. Mol. Biol.* 215: 403 são baseados no algoritmo de Karlin e Altschul (1990), acima. Pesquisas de nucleotídeos de BLAST podem ser realizadas com o programa BLASTN, avaliação = 100, comprimento de palavra = 12, para obter sequências de nucleotídeos homólogas a uma sequência de nucleotídeos que codifica uma proteína. Pesquisas de proteínas BLAST podem ser realizadas com o programa BLASTX, avaliação = 50, comprimento de palavra = 3, para obter sequências de aminoácidos homólogas a uma proteína ou polipeptídeo.

[063] Para obter alinhamentos de intervalos para fins de comparação, BLAST com intervalos (em BLAST 2.0) pode ser utilizado conforme descrito em Altschul et al (1997), *Nucleic Acids Res.* 25: 3389. Em outros casos, pode-se utilizar PSI-BLAST (em BLAST 2.0) para realizar pesquisa com iteração que detecta relacionamentos distantes entre as moléculas. Vide Altschul et al (1997), acima. Ao utilizar BLAST, BLAST com intervalos, PSI-BLAST, pode-se utilizar os parâmetros padrão dos programas correspondentes (por exemplo, BLASTN para sequências de nucleotídeos e BLASTX para proteínas). Vide www.ncbi.nlm.nih.gov. Pode-se também realizar alinhamento manual por meio de inspeção.

[064] A menos que indicado em contrário, os valores de similaridade e identidade de sequências em pares fornecidos no presente designam o valor obtido utilizando o programa GAP, que é parte de GCG 11.0

utilizando os parâmetros a seguir: % de identidade e % de similaridade utilizando penalidade de criação de intervalos de 10, penalidade de extensão de intervalos de 1 ou qualquer de seus programas equivalentes. “Programa equivalente” destina-se a indicar qualquer programa de comparação de sequências que, para quaisquer duas sequências em questão, gere alinhamento que possui coincidências idênticas de resíduos de aminoácidos ou nucleotídeos e percentual idêntico de identidade de sequências em comparação com o alinhamento correspondente gerado por GAP.

[065] GAP utiliza o algoritmo de Needleman e Wunsch (1970), *J. Mol. Biol.* 48: 443-453, para encontrar o alinhamento de duas sequências completas que maximiza a quantidade de coincidências e minimiza a quantidade de intervalos. GAP considera todos os alinhamentos e posições de intervalo possíveis e cria o alinhamento com o maior número de bases coincidentes e menos intervalos. Ele permite o fornecimento de penalidade de criação de intervalos e penalidade de extensão de intervalos em unidades de bases coincidentes. GAP deve utilizar o número de penalidade de criação de intervalos de coincidências para cada intervalo inserido. Caso se escolha penalidade de extensão de intervalos maior que zero, GAP deve, adicionalmente, utilizar cada intervalo inserido do comprimento do intervalo multiplicado pela penalidade de extensão de intervalo. Valores de penalidade de criação de intervalo e valores de penalidade de extensão de intervalo padrão na Versão 10 do Pacote de Software Genético GCG Wisconsin para sequências de proteínas são de 8 e 2, respectivamente. Em várias realizações, para sequências de nucleotídeos, a penalidade de criação de intervalos padrão é de 50, enquanto a penalidade de extensão de intervalos padrão é 3. As penalidades de criação de intervalos e de extensão de intervalos podem ser expressas na forma de número inteiro selecionado a partir do grupo de números inteiros que incluem de 0 a 200. Portanto, em diversas realizações, as penalidades de criação de intervalos e

extensão de intervalos podem ser de 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65 ou mais.

[066] GAP apresenta um membro da família de melhores alinhamentos. Pode haver muitos membros dessa família, mas nenhum outro membro possui melhor qualidade. GAP exibe quatro números importantes para alinhamentos: qualidade, razão, identidade e similaridade. Qualidade é a medida maximizada para alinhar as sequências. Razão é a qualidade dividida pelo número de bases no segmento mais curto. Percentual de identidade é o percentual dos símbolos realmente coincidentes. Percentual de similaridade é o percentual dos símbolos que são similares. Símbolos que atravessam intervalos são ignorados. Similaridade é avaliada quando o valor de matriz de avaliação para um par de símbolos for maior ou igual a 0,5, o limite de similaridade. A matriz de avaliação utilizada na Versão 11.0 do Pacote de Software Genético Wisconsin GCG é BLOSUM62 (vide Henikoff e Henikoff (1989), *Proc. Natl. Acad. Sci. U. S. A.* 89: 10915).

[067] A menos que indicado em contrário, os valores de percentual de identidade e/ou percentual de similaridade entre diversas sequências fornecidas no presente designam o valor obtido utilizando ClustalW2 (http://www.ebi.ac.uk/Tools/msa/clustalw2/) empregando os parâmetros a seguir: para sequências de nucleotídeos utilizando abertura de intervalo de 10, extensão de intervalo de 0,2, distâncias de intervalo de 5, a matriz em peso de DNA IUB e o método de agregação de ligação de vizinhos; para sequências de aminoácidos utilizando abertura de intervalo de 10, extensão de intervalo de 0,2, distâncias de intervalo de 5, a matriz em peso de proteínas Gonnet e o método de agregação de ligação de vizinhos.

[068] ClustalW2 é um programa de alinhamento de múltiplas sequências com base na Web para DNA e proteína (http://www.ebi.ac.uk/Tools/msa/clustalw2/help/index.html) que gerará seis

arquivos de resultado, incluindo sequências de entrada, saída de ferramentas, alinhamento em formato CLUSTAL, árvore de orientação, árvore filogenética e percentual de matriz de identidade. Para realizar o alinhamento, em várias realizações, dois conjuntos de parâmetros podem ser definidos por usuários, ou seja, opções de alinhamento em pares e opções de alinhamento de múltiplas sequências. Os parâmetros selecionados para alinhamento em pares serão utilizados para estabelecer a árvore orientada para o alinhamento de múltiplas sequências, enquanto esta última calculará as melhores coincidências para as sequências de entrada e as alinhará de forma que as identidades, similaridades e diferenças possam ser visualizadas diretamente.

[069] Ao utilizar-se o percentual de identidade de sequências com referência a proteínas, reconhece-se que as posições de resíduos que não são idênticas frequentemente diferem em substituições de aminoácidos conservadores, em que os resíduos de aminoácidos são substituídos por outros resíduos de aminoácidos com propriedades químicas similares (tais como carga ou hidrofobicidade) e, portanto, não alteram as propriedades funcionais da molécula. Quando as sequências diferem em substituições conservadoras, o percentual de identidade de sequências pode ser ajustado para cima para corrigir a natureza conservadora da substituição. Afirma-se que sequências que diferem por essas substituições conservadoras possuem “similaridade de sequências” ou “similaridade”. Meios de realizar esse ajuste são bem conhecidos dos técnicos no assunto. Em várias realizações, isso envolve a avaliação de uma substituição conservadora como parcial e não coincidência completa, de forma a aumentar o percentual de identidade de sequências. Em várias realizações, portanto, em que um aminoácido idêntico recebe avaliação de 1 e substituição não conservadora recebe avaliação zero, a substituição conservadora recebe avaliação de zero a um. A avaliação de substituições conservadoras é calculada, por exemplo, conforme implementado no programa PC/GENE (Intelligenetics,

Mountain View, Califórnia). O percentual é calculado por meio de determinação da quantidade de posições nas quais ocorre o resíduo de aminoácido ou base de ácido nucleico idêntico nas duas sequências para gerar a quantidade de posições coincidentes, dividindo-se o número de posições coincidentes pela quantidade total de posições na janela de comparação e multiplicando-se o resultado por 100 para gerar o percentual de identidade de sequências.

[070] Em diversas realizações, as sequências descritas podem ser introduzidas em plantas para alterar os níveis de uma sequência de domínio K. Em outras realizações, as plantas podem possuir níveis de expressão mais altos de uma sequência de domínio K. Em várias realizações, as plantas incorporaram de forma estável ao seu genoma pelo menos um polinucleotídeo heterólogo que codifica um polipeptídeo de domínio K. Em várias realizações, sequências que podem ser utilizadas para aumentar a expressão de um polipeptídeo de domínio K incluem, mas sem limitações, a sequência descrita em SEQ ID NO: 1, suas variantes ou fragmentos. Prole, variantes e mutantes das plantas regeneradas também são incluídas dentro do escopo, desde que essas partes incluam os polinucleotídeos heterólogos ou introduzidos descritos no presente.

[071] Em diversas realizações, as plantas descritas podem possuir um fenótipo alterado. Em algumas realizações, o fenótipo alterado inclui desenvolvimento de órgãos de flores alterados, números de flores, formação e números de folhas, tamanho da planta ou suas combinações, mas nenhum efeito negativo sobre a produção de frutas ou sementes por planta. Em diversas realizações, isso resulta, portanto, em aumento de rendimento por meio de aumento da densidade de plantas no campo e redução do período de crescimento das plantas.

[072] Em várias realizações, o rendimento das plantas aumenta em cerca de 5%, 7%, 8%, 9%, cerca de 10% a 20% ou cerca de 30% a cerca de 50% em comparação com uma planta controle.

[073] Em algumas realizações, aumento do rendimento é observado como aumento do rendimento em bushels/acre. Em várias realizações, o rendimento aumenta devido a mudanças fenotípicas. Em diversas realizações, as mudanças fenotípicas incluem mudanças do tempo de floração, altura da planta, número de flores ou uma de suas combinações ao sobre-expressar *VcK*. Dependendo da espécie vegetal, os fenótipos podem variar. Em várias realizações, a expressão ectópica de *VcK* reduziu os ciclos vegetais (por exemplo, em duas a três semanas) e reduziu os tamanhos das plantas. Em diversas realizações, a sobre-expressão ectópica de *VcK* promoveu mais flores e a formação de ramos. Em ainda outras realizações, a sobre-expressão de *VcK* aumentou a quantidade de botões em floração e reduziu a necessidade de resfriamento para a floração das plantas. Em algumas realizações, o rendimento pode aumentar devido ao aumento do peso das sementes, tamanho das sementes ou formato das sementes. Em outras realizações, o rendimento pode aumentar devido à quantidade de sementes (em vagens).

[074] Podem ser utilizados diversos métodos de transformação para uma série de espécies vegetais, incluindo, mas sem limitações, monocotiledôneas e dicotiledôneas. Em várias realizações, as plantas incluem, mas sem limitações, milho (*Zea mays*), *Brassica sp.* (por exemplo, *B. napus*, *B. rapa* e *B. juncea*), particularmente as espécies de *Brassica* úteis como fontes de óleos de semente, alfafa (*Medicago sativa*), arroz (*Oryza sativa*), centeio (*Secale cereale*), sorgo (*Sorghum bicolor*, *Sorghum vulgare*), painço (por exemplo, painço pérola (*Pennisetum glaucum*), painço proso (*Panicum miliaceum*), painço rabo de raposa (*Setaria italica*), ragi (*Eleusine coracana*)), girassol (*Helianthus annuus*), açafrão (*Carthamus tinctorius*), trigo (*Triticum aestivum*), soja (*Glycine max*), fumo (*Nicotiana tabacum*), batata (*Solanum tuberosum*), amendoim (*Arachis hypogaea*), algodão (*Gossypium barbadense*, *Gossypium hirsutum*), batata doce (*Ipomoea batatus*), mandioca (*Manihot esculenta*), café (*Coffea*

spp), coco (*Cocos nucifera*), abacaxi (*Ananas comosus*), árvores cítricas (*Citrus spp*), cacau (*Theobroma cacao*), chá (*Camellia sinensis*), banana (*Musa spp*), abacate (*Persea americana*), figo (*Ficus casica*), goiaba (*Psidium guajava*), manga (*Mangifera indica*), azeitona (*Olea europaea*), mamão (*Carica papaya*), caju (*Anacardium occidentale*), macadâmia (*Macadamia integrifolia*), amêndoa (*Prunus amygdalus*), beterraba sacarina (*Beta vulgaris*), cana de açúcar (*Saccharum spp*), aveia, cevada, legumes, plantas ornamentais e coníferas.

[075] Legumes incluem tomates (*Lycopersicon esculentum*), alface (por exemplo, *Lactuca sativa*), feijão verde (*Phaseolus vulgaris*), feijão lima (*Phaseolus limensis*), ervilhas (*Lathyrus spp*), membros do gênero *Cucumis* como pepino (*C. sativus*), melão cantaloupe (*C. cantalupensis*) e melão (*C. melo*). Plantas ornamentais incluem azaleia (*Rhododendron spp*), hortênsia (*Macrophylla hydrangea*), hibisco (*Hibiscus rosasanensis*), rosas (*Rosa spp*), tulipas (*Tulipa spp*), narciso (*Narcissus spp*), petúnia (*Petunia hybrida*), cravo (*Dianthus caryophyllus*), poinsétia (*Euphorbia pulcherrima* e crisântemo.

[076] Coníferas que podem ser empregadas na prática da presente invenção incluem, por exemplo, pinheiros como pinheiro americano (*Pinus taeda*), pinus (*Pinus elliotii*), pinho Ponderosa (*Pinus ponderosa*), pinheiro da praia (*Pinus contorta*), e pinho de Monterey (*Pinus radiata*); abeto de Douglas (*Pseudotsuga menziesii*); espruce canadense (*Tsuga canadensis*); espuce branco (*Picea glauca*); sequoia (*Sequoia sempervirens*); abetos verdadeiros como abeto prateado (*Abies amabilis*) e abeto bálsamo (*Abies balsamea*); e cedros, tais como tuia (*Thuja plicata*) e cipreste de Nootka (*Chamaecyparis nootkatensis*).

[077] Em outras realizações, outras plantas de interesse incluem grãos que fornecem sementes de interesse, plantas oleaginosas e plantas leguminosas. Sementes de interesse incluem grãos como milho, trigo, cevada, arroz, sorgo, centeio e similares. Plantas oleaginosas incluem algodão, soja,

açafrão, girassol, *Brassica*, milho, alfafa, palma, coco e similares. Plantas leguminosas incluem feijão e ervilha. Feijões incluem guar, alfarroba, feno grego, soja, vagem, feijão de corda, feijão mungue, feijão lima, feijão fava, lentilhas, grão de bico etc.

[078] Em ainda outras realizações, plantas incluem frutíferas, tais como uva, laranja, limão, toronja, abacate, tâmaras, pêssego, cereja, azeitona, ameixa, coco, maçã, peras, amora preta, mirtilo, framboesa, morango, abacaxi, tomate, pepino e berinjela.

[079] Em várias realizações, os polinucleotídeos descritos empregados nos métodos e composições podem ser fornecidos em conjuntos de expressão para expressão em plantas de interesse. Em algumas realizações, o conjunto incluirá sequências reguladoras 5’ e 3’ ligadas operacionalmente a um polinucleotídeo. Em uma realização, ligação operacional entre um polinucleotídeo de interesse e uma sequência reguladora (por exemplo, promotor) é uma ligação funcional que permite a expressão do polinucleotídeo de interesse. Elementos ligados operacionalmente podem ser contíguos ou não contíguos. Quando utilizado para designar a ligação de duas regiões de codificação de proteínas, ligado operacionalmente indica que as regiões de codificação encontram-se no mesmo quadro de leitura. O conjunto pode conter adicionalmente pelo menos um gene adicional a ser cotransformado no organismo. Em outros casos, o(s) gene(s) adicional(is) pode(m) ser fornecido(s) sobre diversos conjuntos de expressão. Esse cassete de expressão pode possuir uma série de locais de restrição, locais de recombinação ou ambos para inserção do polinucleotídeo de domínio K para que fique sob a regulagem de transcrição das regiões reguladoras. Em diversas realizações, o cassete de expressão pode conter adicionalmente genes marcadores selecionáveis. Em algumas realizações, o cassete de expressão pode incluir, na direção 5’-3’ de transcrição, uma região de início de tradução e transcrição (por exemplo, um promotor), um

polinucleotídeo de domínio K e uma região de término de transcrição e tradução (por exemplo, região de término) funcional em plantas. Em várias realizações, as regiões reguladoras (por exemplo, promotores, regiões reguladoras de transcrição e regiões de término de transcrição), o polinucleotídeo de domínio K ou ambos podem ser nativos/análogos para a célula hospedeira ou entre si. Em outras realizações, as regiões reguladoras, os polinucleotídeos de domínio K ou ambos podem ser heterólogos para a célula hospedeira ou entre si. Em uma realização, um promotor ligado operacionalmente a um polinucleotídeo heterólogo é de uma espécie diferente daquela da qual o polinucleotídeo foi derivado ou, se da mesma espécie ou de espécie análoga, um ou ambos são substancialmente modificados da sua forma original e/ou local genômico ou o promotor não é o promotor nativo para o polinucleotídeo ligado operacionalmente. Em algumas realizações, um gene quimérico que possui uma sequência de codificação ligada operacionalmente a uma região de início de transcrição que é heteróloga à sequência de codificação. Em outras realizações, promotores heterólogos ou as sequências promotoras nativas podem ser utilizados para expressar as sequências. Em várias realizações, a região de término pode ser nativa com a região de início de transcrição, pode ser nativa com o polinucleotídeo de domínio K ligado operacionalmente, pode ser nativa com o hospedeiro vegetal ou pode ser derivada de outra fonte (por exemplo, externa ou heteróloga) para o promotor, o polinucleotídeo de domínio K, o hospedeiro vegetal ou qualquer de suas combinações. Regiões de término convenientes são disponíveis a partir do plasmídeo Ti de *A. tumefaciens*, tais como as regiões de término de octopino sintase e nopalino sintase. Vide também 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; e Joshi et al (1987), *Nucleic Acids*

Res. 15: 9627-9639.

[080] Quando apropriado, os polinucleotideos podem ser otimizados para aumento da expressão na planta transformada. Isso significa que os polinucleotídeos podem ser sintetizados utilizando códons preferidos de plantas para expressão aprimorada. Vide, por exemplo, Campbell e Gown (1990), *Plant Physiol.* 92: 1-11 para discussão do uso de códons preferido por hospedeiro. Métodos são disponíveis na técnica para a síntese de genes preferidos de plantas. Vide, por exemplo, as Patentes Norte-Americanas nº 5.380.831 e 5.436.391 e Murray et al (1989), *Nucleic Acids Res.* 17: 477-498, todos incorporados ao presente como referência.

[081] Sabe-se que modificações de sequência adicionais aumentam a expressão genética em um hospedeiro celular. Estas incluem a eliminação de sequências que codificam sinais de poliadenilação espúrios, sinais de local de divisão de extrons-introns, repetições de transpossons e outras sequências bem caracterizadas que podem ser prejudiciais para a expressão genética. O teor de G-C da sequência pode ser ajustado para níveis médios para um dado hospedeiro celular, conforme calculado por meio de referência a genes conhecidos expressos na célula hospedeira. Quando possível, a sequência pode ser modificada para evitar estruturas de mRNA secundárias de grampos de cabelo previstas.

[082] Os conjuntos de expressão podem conter adicionalmente sequências líderes 5'. Essas sequências líderes podem agir para aumentar a tradução. Líderes de tradução são conhecidos na técnica e incluem, por exemplo, líderes de picornavírus, líder de EMCV (região não codificadora 5' de encefalomiocardite) (Elroy-Stein et al (1989), *Proc. Natl. Acad. Sci. U. S. A.* 86: 6126-6130); líderes de potivírus, tais como líder TEV (Vírus da corrosão do fumo) (Gallie et al (1995), *Gene* 165 (2): 233-238), líder MDMV (vírus mosaico anão do milho) (*Virology* 154: 9-20) e proteína de ligação de cadeia pesada de

imunoglobulina humana (BiP) (Macejak et al (1991), *Nature* 353: 90-94); líder não traduzido do mRNA de proteína de revestimento de vírus mosaico da alfafa (RNA AMV 4) (Jobling et al (1987), *Nature* 325: 622-625); líder de vírus mosaico do fumo (TMV) (Gallie et al (1989), *Molecular Biology of RNA*, ed. Cech (Liss, Nova Iorque), págs. 237-256); e líder de vírus mosqueado clorótico do milho (MCMV) (Lommel et al (1991), *Virology* 81: 382-385). Vide também Della-Cioppa et al (1987), *Plant Physiol.* 84: 965-968.

[083] Durante a preparação do cassete de expressão, os diversos fragmentos de DNA podem ser manipulados, de forma a fornecer as sequências de DNA na orientação adequada e, conforme apropriado, no quadro de leitura adequado. Com este propósito, podem ser empregados adaptadores ou ligantes para unir os fragmentos de DNA ou outras manipulações podem ser envolvidas para fornecer locais de restrição convenientes, remoção de DNA supérfluo, remoção de locais de restrição ou similares. Com este propósito, mutagênese *in vitro*, reparo de primers, restrição, combinação, ressubstituições, tais como transições e transversões, podem estar envolvidas.

[084] Em diversas realizações, os promotores incluem o promotor nativo da sequência de polinucleotídeos de interesse. Os promotores podem ser selecionados com base no resultado desejado. Os ácidos nucleicos podem ser combinados com promotores constitutivos, preferidos de tecidos, indutíveis ou outros para expressão em plantas. Em algumas realizações, promotores constitutivos incluem o promotor central do promotor Rsyn7 e outros promotores constitutivos descritos em WO 99/43838 e na Patente Norte-Americana nº 6.072.050; o promotor CaMV 35S central (Odell et al (1985), *Nature* 313: 810-812); actina de arroz (McElroy et al (1990), *Plant Cell* 2: 163-171); ubiquitina (Christensen et al (1989), *Plant Mol. Biol.* 12: 619-632 e 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); promotor de ALS

(Patente Norte-Americana nº 5.659.026), promotor GOS2 (dePater et al (1992), *Plant J.* 2: 837-44) e similares. Outros promotores constitutivos incluem, por exemplo, as Patentes Norte-Americanas nº 5.608.149, 5.608.144, 5.604.121, 5.569.597, 5.466.785, 5.399.680, 5.268.463, 5.608.142 e 6.177.611.

[085] Em diversas realizações, promotor específico de tecido é um promotor ativo na transcrição nos órgãos ou tecido de raízes, hastes, folhas ou sementes. Promotor específico de raízes, por exemplo, é um promotor que é ativo na transcrição, predominantemente em raízes vegetais. Diversos promotores específicos de tecidos são descritos na Patente Norte-Americana nº 8.697.947.

[086] Em diversas realizações, o cassete de expressão pode também incluir um gene marcador selecionável para seleção de células transformadas. Genes marcadores selecionáveis podem ser utilizados para seleção de tecidos ou células transformadas. Em várias realizações, genes marcadores incluem genes que codificam resistência a antibióticos, tais como os que codificam neomicina fosfotransferase II (NPT II) e higromicina fosfotransferase (HPT), bem como genes que conferem resistência a compostos herbicidas, tais como glufosinato amônio, bromoxinil, imidazolinonas e 2,4-diclorofenoxiacetato (2,4-D). Em outras realizações, marcadores selecionáveis incluem marcadores fenotípicos tais como 13-galactosidase e proteínas fluorescentes tais como proteína fluorescente verde (GFP) (Su et al (2004), *Biotechnol. Bioeng.* 85: 610-9 e Fetter et al (2004), *Plant Cell* 16: 215-28), proteína fluorescente ciano (CYP) (Bolte et al (2004), *J. Cell Science* 117: 943-54 e Kato et al (2002), *Plant Physiol.* 129: 913-42) e proteína fluorescente amarela (PhiYFP® da Evrogen, vide Bolte et al (2004), *J. Cell Science* 117: 943-54). Para marcadores selecionáveis adicionais, vide, de forma geral, Yarranton (1992), *Curr. Opin. Biotech.* 3: 506-511; Christopherson et al (1992), *Proc. Natl. Acad. Sci. U. S. A.* 89: 6314-6318; Yao et al (1992), *Cell* 71: 63-72; Reznikoff

(1992), *Mol. Microbiol.* 6: 2419-2422; Barkley et al (1980) em *The Operon*, págs. 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. U. S. A.* 86: 5400-5404; Fuerst et al (1989), *Proc. Natl. Acad. Sci. U. S. A.* 86: 2549-2553; Deuschle et al (1990), *Science* 248: 480-483; Gossen (1993), Tese de Ph.D,, Universidade de Heidelberg; Reines et al (1993), *Proc. Natl. Acad. Sci. U. S. A.* 90: 1917-1921; Labow et al (1990), *Mol. Cell Biol.* 10: 3343-3356; Zambretti et al (1992), *Proc. Natl. Acad. Sci. U. S. A.* 89: 3952-3956; Baim et al (1991), *Proc. Natl. Acad. Sci. U. S. A.* 88: 5072-5076; Wyborski et al (1991), *Nucleic Acids Res.* 19: 4647-4653; Hillen e Wissman (1989), *Topics Mol. Struc. Biol.* 10: 143-162; Degenkolb et al (1991), *Antimicrob. Agents Chemother.* 35: 1591-1595; Kleinschnidt et al (1988), *Biochemistry* 27: 1094-1104; Bonin (1993), Tese de PhD, Universidade de Heidelberg; Gossen et al (1992), *Proc. Natl. Acad. Sci. U. S. A.* 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, Berlim); Gill et al (1988), *Nature* 334: 721-724. Essas descrições são incorporadas ao presente como referência. A lista acima de genes marcadores selecionáveis não se destina a ser limitadora. Qualquer gene marcador selecionável pode ser utilizado.

[087] Em diversas realizações, os polinucleotídeos descritos podem ser empilhados com qualquer combinação de sequências de polinucleotídeos de interesse para criar plantas com característica desejada. Em outras realizações, as combinações geradas podem também incluir diversas cópias de qualquer um dos polinucleotídeos de interesse. Em ainda outras realizações, os polinucleotídeos descritos podem também possuir características desejáveis para resistência a doenças ou herbicidas (por exemplo, genes de desintoxicação de fumosinina (Patente Norte-Americana nº 5.792.931); genes de resistência a doenças e avirulência (Jones et al (1994), *Science* 266: 789; Martin

et al (1993), *Science* 262: 1432; Mindrinos et al (1994), *Cell* 78: 1089); mutantes de acetolactato sintase (ALS) que geram resistência a herbicidas tais como mutações S4 e/ou Hra; inibidores de glutamino sintase tais como fosfinotricina ou basta (por exemplo, gene bar); e resistência a glifosato (gene EPSPS)); e características desejáveis para processamento ou produtos de processamento tais como alto óleo (por exemplo, Patente Norte-Americana nº 6.232.529); óleos modificados (por exemplo, genes dessaturase de ácidos graxos (Patente Norte-Americana nº 5.952.544; WO 94/11516)); amidos modificados (por exemplo, pirofosforilases de ADPG (AGPase), sintases de amido (SS), enzimas de ramificação de amido (SBE) e enzimas de desramificação de amido (SDBE); e polímeros ou bioplásticos (por exemplo, Patente Norte-Americana nº 5.602.321; betacetotiolase, sintase de póli-hidroxibutirato e acetoacetil-CoA reductase (Schubert et al (1988), *J. Bacteriol.* 170: 5837-5847) facilitam a expressão de póli-hidroxialcanoatos (PHAs)); cujas descrições são incorporadas ao presente como referência. Poder-se-á também combinar os polinucleotídeos descritos com polinucleotídeos que fornecem características agronômicas tais como a esterilidade dos machos (vide, por exemplo, Patente Norte-Americana nº 5.583.210), resistência das hastes ou características de transformação tecnológica tais como a regulagem do ciclo celular ou direcionamento genético (por exemplo, WO 99/61619, WO 00/17364 e WO 99/25821); cujas descrições são incorporadas ao presente como referência.

[088] Essas combinações reunidas podem ser criadas por meio de qualquer método, incluindo, mas sem limitações, cruzamento de plantas por meio de metodologia TopCross ou qualquer convencional, ou transformação genética. Se as sequências forem reunidas por meio de transformação genética das plantas, em várias realizações, as sequências de polinucleotídeos descritas podem ser combinadas a qualquer momento e em qualquer ordem. Em algumas realizações, uma planta transgênica que possui uma ou mais características

desejadas pode ser utilizada como alvo para introduzir características adicionais por meio de transformação subsequente. As características podem ser introduzidas simultaneamente em um protocolo de cotransformação com os polinucleotídeos descritos fornecidos por meio de qualquer combinação de conjuntos de transformação. Em uma realização, caso duas sequências sejam introduzidas, as duas sequências podem estar contidas em conjuntos de transformação separados (*trans*) ou contidas no mesmo conjunto de transformação (*cis*). A expressão das sequências pode ser dirigida pelo mesmo promotor ou por promotores diferentes. Isso pode ser combinado com qualquer combinação de outros conjuntos de supressão ou conjuntos de sobre-expressão para gerar a combinação desejada de características na planta. Reconhece-se ainda que sequências de polinucleotídeos podem ser reunidas em um local genômico desejado utilizando um sistema de recombinação específico de local. Vide, por exemplo, WO 99/25821, WO 99/25854, WO 99/25840, WO 99/25855 e WO 99/25853, todas as quais são incorporadas ao presente como referência.

[089] As diversas realizações fornecem métodos que envolvem a introdução de polinucleotídeo ou polipeptídeo em uma planta. Qualquer método apropriado para introduzir uma sequência em uma planta pode ser utilizado, de forma que o polinucleotídeo ou polipeptídeo ganhe acesso ao lado interno de pelo menos uma célula vegetal. Métodos de introdução de polinucleotídeos ou polipeptídeos em plantas são conhecidos na técnica, incluindo, mas sem limitações, métodos de transformação estáveis, métodos de transformação transitórios e métodos mediados por vírus.

[090] Protocolos de transformação, bem como protocolos de introdução de polipeptídeos ou sequências de polinucleotídeos em plantas, podem variar dependendo do tipo de planta ou célula vegetal, tais como monocotiledôneas ou dicotiledôneas, dirigidas para transformação. Métodos apropriados de introdução de polipeptídeos e polinucleotídeos em células

vegetais incluem microinjeção (Crossway et al (1986), *Biotechniques* 4: 320-334), eletroporação (Riggs et al (1986), *Proc. Natl. Acad. Sci. U. S. A.* 83: 5602-5606, transformação mediada por *Agrobacterium* (Patente Norte-Americana nº 5.563.055 e Patente Norte-Americana nº 5.981.840), transferência genética direta (Paszkowski et al (1984), *EMBO J.* 3: 2717-2722) e aceleração de partículas balística (vide, por exemplo, Patente Norte-Americana nº 4.945.050; Patente Norte-Americana nº 5.879.918; Patentes Norte-Americanas nº 5.886.244 e 5.932.782; Tomes et al (1995) em *Plant Cell, Tissue and Organ Culture: Fundamental Methods*, ed. Gamborg e Phillips (Springer-Verlag, Berlim); McCabe et al (1988), *Biotechnology* 6: 923-926); e transformação Led (WO 00/28058). Vide também Weissing et al (1988), *Ann. Rev. Genet.* 22: 421-477; Sanford et al (1987), *Particulate Science and Technology* 5: 27-37 (cebola); Christou et al (1988), *Plant Physiol.* 87: 671-674 (soja); McCabe et al (1988), *Bio/Technology* 6: 923-926 (soja); Finer e McMullen (1991), *In Vitro Cell Dev. Biol.* 27P: 175-182 (soja); Singh et al (1998), *Theor. Appl. Genet.* 96: 319-324 (soja); Datta et al (1990), *Biotechnology* 8: 736-740 (arroz); Klein et al (1988), *Proc. Natl. Acad. Sci. U. S. A.* 85: 4305-4309 (milho); Klein et al (1988), *Biotechnology* 6: 559-563 (milho); Patentes Norte-Americanas nº 5.240.855, 5.322.783 e 5.324.646; Klein et al (1988), *Plant Physiol.* 91: 440-444 (milho); Fromm et al (1990), *Biotechnology* 8: 833-839 (milho); Hooykaas-Van Slogteren et al (1984), *Nature (London)* 311: 763-764; Patente Norte-Americana nº 5.736.369 (cereais); Bytebier et al (1987), *Proc. Natl. Acad. Sci. U. S. A.* 84: 5345-5349 (Liliaceae); De Wet et al (1985), *The Experimental Manipulation of Ovule Tissues*, ed. Chapman et al (Longman, Nova Iorque), págs. 197-209 (pólen); Kaeppler et al (1990), *Plant Cell Reports* 9: 415-418 e Kaeppler et al (1992), *Theor. Appl. Genet.* 84: 560-566 (transformação mediada por Whisker); D'Halluin et al (1992), *Plant Cell* 4: 1495-1505 (eletroporação); Li et al (1993), *Plant Cell Reports* 12: 250-255 e Christou e Ford (1995), *Annals of Botany* 75: 407-413

(arroz); Osjoda et al (1996), *Nature Biotechnology* 14: 745-750 (milho via *Agrobacterium tumefaciens*); todos os quais são incorporados ao presente como referência.

[091] Em algumas realizações, as sequências de domínio K ou suas variantes podem ser fornecidas a uma planta utilizando uma série de métodos de transformação transitórios. Esses métodos de transformação transitórios incluem, mas sem limitações, a introdução do domínio K ou suas variantes diretamente na planta ou a introdução do transcrito de domínio K na planta. Esses métodos incluem, por exemplo, microinjeção ou bombardeamento de partículas. Vide, por exemplo, 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 e Hush et al (1994), *The Journal of Cell Science* 107: 775-784, todos os quais são incorporados ao presente como referência. Outros métodos incluem sistema de vetores virais e a precipitação do polinucleotídeo de forma que elimine a liberação subsequente do DNA. Desta forma, a transcrição do DNA ligado por partículas pode ocorrer, mas a frequência com que ele é liberado para integrar-se ao genoma é muito reduzida. Esses métodos incluem o uso de partículas revestidas com polietilimina (por exemplo, PEI; Sigma nº P3143).

[092] Em outras realizações, o polinucleotídeo pode ser introduzido em plantas por meio do contato de plantas com vírus ou ácidos nucleicos virais. Em algumas realizações, esses métodos envolvem a incorporação de uma construção de nucleotídeos em uma molécula de DNA ou RNA viral. Em uma realização, a sequência de domínio K ou uma de suas variantes pode ser sintetizada inicialmente como parte de uma poliproteína viral, que pode ser processada posteriormente por meio de proteólise *in vivo* ou *in vitro* para produzir a proteína recombinante desejada. Em várias realizações, os promotores também englobam promotores empregados para transcrição por polimerases de

RNA virais. Métodos de introdução de polinucleotídeos em plantas e expressão de uma proteína nelas codificada, que envolvem moléculas de DNA ou RNA virais, são conhecidos na técnica. Vide, por exemplo, as Patentes Norte-Americanas nº 5.889.191, 5.889.190, 5.866.785, 5.589.367, 5.316.931 e Porta et al (1996), *Molecular Biotechnology* 5: 209-221; incorporados ao presente como referência.

[093] Podem ser utilizados diversos métodos de inserção de polinucloeotídeos dirigidos em um local específico no genoma vegetal. Em algumas realizações, a inserção de polinucleotídeos em um local genômico desejado é atingida utilizando um sistema de recombinação específico de local. Vide, por exemplo, WO 99/25821, WO 99/25854, WO 99/25840, WO 99/25855 e WO 99/25853, todas as quais são incorporadas ao presente como referência. Em uma realização, o polinucleotídeo descrito pode estar contido em um conjunto de transferência ladeado por dois locais de recombinação não recombinogênicos. O conjunto de transferência é introduzido em uma planta que possui, incorporado de forma estável ao seu genoma, um local alvo que é ladeado por dois locais de recombinação não recombinogênicos que correspondem aos locais do conjunto de transferência. É fornecida recombinase apropriada e o conjunto de transferência é integrado ao local alvo. O polinucleotídeo descrito é, portanto, integrado em uma posição cromossômica específica no genoma vegetal.

[094] As células que foram transformadas podem ser cultivadas em plantas de acordo com formas convencionais. vide, por exemplo, McCormick et al (1986), *Plant Cell Reports* 5: 81-84. Estas plantas podem ser cultivadas em seguida e polinizadas com a mesma linhagem transformada ou com linhagens diferentes e a prole resultante que possui expressão da característica fenotípica desejada é identificada. Duas ou mais gerações podem ser cultivadas para garantir que a expressão da característica fenotípica desejada seja mantida e

herdada de forma estável e, em seguida, sementes colhidas para garantir que a expressão da característica fenotípica desejada foi atingida. Desta forma, em uma realização, podem ser elaboradas sementes transformadas (também denominadas “sementes transgênicas”) que contêm um polinucleotídeo, incorporadas de forma estável ao seu genoma. Em outras realizações, sementes transformadas ou sementes transgênicas possuem um cassete de expressão incorporado de forma estável ao seu genoma.

[095] Qualquer método ou composição que module a expressão de um produto genético alvo (por exemplo, sobre-expressão ou expressão knock-down), seja no nível de transcrição ou de tradução, ou module a atividade do produto genético alvo pode ser utilizado para atingir a sobre-expressão, concentração ou ativação do produto genético alvo. Em várias realizações, o nível aumenta em pelo menos 1%, 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% ou mais, incluindo qualquer faixa ou valor entre eles com relação a uma planta controle, parte de planta ou célula apropriada. A expressão pode ocorrer, durante e/ou após o crescimento vegetal em um estágio de desenvolvimento desejado. Em algumas realizações, os polipeptídeos descritos são modulados em plantas frutíferas. Em outras realizações, as plantas frutíferas incluem mirtilos e similares.

[096] Em várias realizações, o nível de expressão de um polipeptídeo que possui domínio K ou seu fragmento ou variante biologicamente ativa pode ser medido direta ou indiretamente. Em algumas realizações, o nível de expressão de um polipeptídeo que possui domínio K ou seu fragmento ou variante biologicamente ativa pode ser medido testando-se o nível do polipeptídeo de domínio K na planta.

[097] Em outras realizações, o nível de expressão de um polipeptídeo que possui domínio K ou seu fragmento ou variante biologicamente ativa pode ser medido por meio de medição do nível do polinucleotídeo que

codifica a proteína ou de medição da atividade do polipeptídeo de domínio K na planta. o polipeptídeo ou polinucleotídeo pode ser introduzido em uma célula vegetal. A célula vegetal que possui a sequência introduzida pode ser selecionada utilizando métodos conhecidos dos técnicos no assunto, tais como, mas sem limitações, análise Southern Blot, Northern Blot, sequenciamento de DNA, análise de PCR ou análise fenotípica. Em várias realizações, uma planta ou parte de planta alterada ou modificada pelo acima pode ser cultivada sob condições promotoras das plantas por tempo suficiente para modular a concentração e/ou a atividade de polipeptídeos descritos na planta.

[098] Em diversas realizações, o nível e/ou a atividade do polipeptídeo podem ser modulados por meio do emprego de um polinucleotídeo que não é capaz de dirigir, em uma planta transforamda, a expressão de uma proteína ou RNA. O polinucleotídeo pode, por exemplo, ser utilizado para projetar construções de polinucleotídeos que podem ser empregadas em métodos de alteração ou mutação de uma sequência de nucleotídeos genômica em um organismo. Essas construções de polinucleotídeos incluem, mas sem limitações, vetores de RNA:DNA, vetores de mutação de RNA:DNA, vetores de reparo de RNA:DNA, oligonucleotídeos de duplex misto, oligonucleotídeos de RNA:DNA autocomplementares e oligonucleobases recombinogênicas. Essas construções de nucleotídeos e métodos de uso são conhecidos na técnica. Vide as Patentes Norte-Americanas nº 5.565.350, 5.731.181, 5.756.325, 5.760.012, 5.795.972 e 5.871.984, todas as quais incorporadas ao presente como referência. Vide também WO 98/49350, WO 99/07865, WO 99/25821 e Beetham et al (1999), *Proc. Natl. Acad. Sci. U. S. A.* 96: 8774-8778, incorporados ao presente como referência.

[099] Os diversos métodos descritos no presente não dependem da incorporação de todo o polinucleotídeo no genoma, mas apenas a planta ou sua célula pode ser alterada como resultado da introdução do polinucleotídeo

em uma célula. Em várias realizações, o genoma pode ser alterado após a introdução do polinucleotídeo em uma célula. Em outras realizações, o polinucleotídeo ou qualquer de suas partes pode incorporar-se ao genoma da planta. Alterações do genoma incluem, mas sem limitações, adições, exclusões e substituições de nucleotídeos no genoma.

[100] A atividade, o nível ou ambos de um polipeptídeo de domínio K aumentam. Em diversas realizações, aumento da atividade, nível ou ambos do polipeptídeo de domínio K realizam-se por meio de fornecimento à planta de um polipeptídeo de domínio K ou seu fragmento ou variante biologicamente ativa. Conforme discutido em outros pontos do presente, muitos métodos são conhecidos na técnica para fornecer um polipeptídeo para uma planta, incluindo, mas sem limitações, introdução direta do polipeptídeo de domínio K na planta ou introdução na planta (de forma transitória ou estável) de uma construção de polinucleotídeos que codifica um polipeptídeo que possui atividade de domínio K. Em várias realizações, o nível e/ou a atividade de um polipeptídeo de domínio K podem aumentar por meio de alteração do gene que codifica o polipeptídeo de domínio K ou seu promotor. Vide, por exemplo, Kmiec, Patente Norte-Americana nº 5.565.350; Zarling et al, PCT/US93/03868. São fornecidas, portanto, plantas que sofreram mutagênese e possuem mutações em genes de domínio K, em que as mutações aumentam a expressão do gene de domínio K ou aumentam a atividade do polipeptídeo de domínio K codificado.

[101] Em diversas realizações, o nível de uma única sequência de domínio K pode ser modulado para produzir o fenótipo desejado. Em outras realizações, o nível de expressão de diversas sequências que possuem domínio K ou um de seus fragmentos ou variantes biologicamente ativos é modulado. Em ainda outras realizações, o nível da sequência de domínio K é modulado por aumento.

[102] As diversas realizações serão adicionalmente descritas por

meio de referência aos exemplos a seguir, que são oferecidos para ilustrar adicionalmente diversas realizações. Dever-se-á compreender, entretanto, que muitas variações e modificações podem ser realizadas mantendo-se ainda dentro do escopo das diversas realizações.

MATERIAIS E MÉTODOS

CLONAGEM E ANÁLISE FILOGENÉTICA DE DOMÍNIO VcK

[103] Para gerar sequências de cDNA para clonagem de domínio VcK, RNA total foi isolado de botões de flores vernalizados de um cultivo de mirtilo de arbusto alto do norte (*Vaccinium corymbosum* L.), Bluecrop, utilizando o Mini Kit Vegetal RNeasy (Qiagen, Valência CA, Estados Unidos) com modificações conforme descrito por Dhanaraj et al (2004). Após o isolamento, a amostra foi tratada com DNase RQ1 (Promega, Madison WI, Estados Unidos) de acordo com as instruções do fabricante, seguida por extração em clorofórmio:álcool isoamílico (24:1), precipitação de etanol a 100%, lavagem com etanol a 70% e, por fim, nova suspensão em água tratada com pirocarbonato de dietila (DEPC). O RNA total (0,6 µg) sofreu transcrição reversa em DNA complementar (cDNA) utilizando Transcriptase Reversa SuperScript II (Invitrogen, Carlsbad CA, Estados Unidos). Um décimo do cDNA foi utilizado em reação de amplificação por PCR utilizando Mistura Master Verde GoTaq (Promega).

[104] Sequência de EST de mirtilo "*Vaccinium_corymbosum_v1_*Contig03391", baixada do banco de dados de EST de mirtilo (http://www.vaccinium.org/) foi utilizada para projetar primers para o *domínio VcK*. PCRs abrigados foram conduzidos utilizando os primers SOCF1 e SOCR1 para a primeira rodada de amplificação e SOF1 e SOK (vide a Tabela 1) para a segunda rodada de amplificação. Os produtos de PCR foram clonados em um vetor pCR 2.1-TOPO (Invitrogen) e verificados por meio de sequenciamento.

TABELA 1

[105] Primers utilizados:

Nome do primer	Sequência	Alvo
SOCF1	SEQ ID NO: 4	Gene similar a SUPRESSOR DA SOBRE-EXPRESSÃO DE CONSTANTES 1 de mirtilo (*domínio VcK*)
SOCR1	SEQ ID NO: 5	
SOF1	SEQ ID NO: 6	
SOK	SEQ ID NO: 7	
35S-F	SEQ ID NO: 8	Parte 3' do promotor de *CaMV 35S*
NPTII-F	SEQ ID NO: 9	Gene fosfotransferase de neomicina II (*npt*II)
NPTII-R	SEQ ID NO: 10	
Actina F	SEQ ID NO: 11	*Actina*
Actina R	SEQ ID NO: 12	
Domínio BBK - F	SEQ ID NO: 13	Domínio *VcK*
Domínio BBK - R	SEQ ID NO: 14	

[106] Sequências de aminoácidos e de nucleotídeos de ortólogos de *domínio VcK* foram recuperadas do NCBI (http://blast.ncbi.nlm.nih.gov/). Sequências de aminoácidos selecionadas foram alinhadas utilizando o programa de alinhamento de sequências múltiplas Clustal W2 (Thompson et al, 1994) no EBI com parâmetros padrão (Matriz de Peso de Proteínas: Gonnet, penalidade de abertura de intervalos de 10 e penalidade de extensão de intervalo de 0,20, os intervalos mais próximos que cinco aminoácidos são penalizados, método de

formação de conjuntos de ligação de vizinhos, http://www.ebi.ac.uk/Tools/mas/clustalw2/). Realizou-se análise filogenética de sequências de nucleotídeos de genes similares a *domínio K* de dezessete plantas dicotiledôneas e uma planta monocotiledônea (*Oryza sativa*) por meio do método de probabilidade máxima com base no modelo de Tamura-Nei utilizando o programa MEGA 5 (Tamura et al, 2011). *OsMADS56* foi utilizado como grupo de saída e o ponto preto indica a proteína similar a *domínio K* de mirtilo (*similar a domínio VcK*). A variação de taxa entre os locais foi modelada utilizando distribuição Gama discreta (quatro categorias) locais sem variação (G+I). A árvore foi desenhada em escala, com o comprimento do ramo igual à quantidade de substituições por nucleotídeo.

CONSTRUÇÃO DO VETOR DE EXPRESSÃO E TRANSFORMAÇÃO DE FUMO

[107] Para realizar construção para transformação vegetal, o fragmento de 5'-*Xba*I-*domínio VcK*-*Sac*I-3' foi liberado do vetor pCR2.1-TOPO e inserido em seguida nos locais *Xba*I e *Sac*I na região de T-DNA entre o promotor 35S do vírus mosaico da couve-flor e o terminal *Nos* em PBI121, em que a região de codificação de GUS (β-glucuronidase) foi removida. O 35S-*domínio VcK* resultante foi sequenciado para confirmar a sequência de *domínio VcK* e introduzido em seguida na linhagem EHA105 de *Agrobacterium tumefaciens* (Hood et al, 1993), utilizando o método de congelamento/descongelamento.

[108] Transformação de explantes de folhas de fumo (*Nicotiana tabacum* L. cv Samsun) mediada por *Agrobacterium* com 35S-*domínio VcK* foi conduzida de acordo com Horsch et al (1985). Brotos resistentes a canamicina foram enraizadas em meio Murashige e Skoog (MS) (1962) contendo 100 mg l^{-1} de monossulfato de canamicina (Km). Plantas transgênicas de explantes de folhas separados foram consideradas eventos transgênicos independentes.

[109] Para confirmar a presença do transgene, DNA foi isolado de tecidos de folhas utilizando um método de brometo de cetiltrimetilamônio (CTAB)

(Doyle e Doyle, 1987). Dois pares de primers (Tabela 1) foram utilizados para verificação de PCR: *npt*II-F e *npt*II-R para o gene *npt*II e 35S-F (porção 3' do promotor *35S* de *CaMV*) e SOK para o gene de *domínio VcK*.

ANÁLISES DE PCR POR TRANSCRIÇÃO REVERSA (PCR RT)

[110] RNA total foi isolado de folhas jovens de plantas de fumo T0 selecionadas utilizando um Mini Kit Vegetal RNeasy (Qiagen, Valencia CA, Estados Unidos) com tratamento com DNase livre de RNase em coluna (Qiagen) de acordo com as instruções do fabricante. A transcrição reversa de RNA em cDNA foi conduzida utilizando transcriptase reversa SuperScript II (Invitrogen). O cDNA resultante, 2 µl/amostra, foi utilizado para amplificação por meio de PCR de gene de actina endógeno com 160 pares de bases e um fragmento do *domínio VcK* com 474 pares de bases. Os primers são exibidos na Tabela 1. As condições de reação para todos os pares de primers foram de 94 ºC por dois minutos, 35 ciclos de 45 segundos a 94 ºC, 60 segundos a 58 ºC e 90 segundos a 72 ºC, com extensão final por dez minutos a 72 ºC. Todos os produtos de PCR e PCR RT amplificados foram separados sobre gel de agarose a 1,0% contendo brometo de etídio, visualizados e fotografados sob luz UV.

FENOTIPIFICAÇÃO DE PLANTAS DE FUMO TRANSGÊNICAS

[111] Quando supostos transformadores T0 e plantas controle/não transgênicas cresceram até o estágio de 3-5 folhas, elas foram transplantadas do meio MS para 10 vasos de 10 cm contendo meio de vasos Suremix (Michigan Grower Products Inc., Galesburg MI, Estados Unidos) e cultivadas sob fotoperíodo de 16 horas com 30 µmol $m^{-2}s^{-1}$ de luz de lâmpadas fluorescentes brancas frias por três semanas antes de serem transplantadas para vasos de um galão. A estrutura de flores de plantas transgênicas foi comparada com plantas controle por meio de um microscópio de dissecação. Essas plantas transgênicas T0 foram autopolinizadas para obter as sementes T1 utilizadas no restante do estudo.

[112] Plantas de fumo transgênicas com pBI121 (o vetor de cadeia principal) não exibiram diferença significativa de fenótipo (por exemplo, tempo de floração, altura da planta e rendimento de sementes) em comparação com plantas do tipo selvagem. Sementes T1 e fumo não transgênico foram plantados em bandejas com 48 células contendo meio de vasos Suremix e mantidos sob fotoperíodo de 16 horas com 30 µmol $m^{-2}s^{-1}$ de luz de lâmpadas fluorescentes brancas frias enquanto a presença de transgenes foi confirmada por meio de PCR utilizando primers 35S-F (SEQ ID NO: 8) e (SEQ ID NO: 7) (Tabela 1). Mudas com tamanho similar (dez para cada evento transgênico ou controle) foram transplantadas em seguida para 10 vasos de 10 cm e cultivadas na estufa sob fotoperíodo natural (em média, 14,9 horas de luz solar natural por dia) com temperatura mínima de 23 ºC em projeto totalmente aleatorizado com irrigação e fertilização regulares.

[113] Os dias até a floração (dias entre o cultivo e o surgimento do primeiro botão de flor), número de folhas, altura da planta e o número de vagens com sementes foram registrados para as plantas T1. A quantidade de folhas sobre a haste central e a altura da planta (do nível do solo até o fundo do conjunto de flores central) foram determinadas quando 50% das flores no conjunto central estavam abertas. A quantidade de vagens com sementes no conjunto de flores central foi determinada quando as plantas terminaram a formação de botões. Conduziu-se análise estatística com PROC GLM em SAS 9.3 (SAS Institute Inc., Cary NC, Estados Unidos). A diferença de fumo do tipo selvagem foi determinada utilizando o teste de duas formas de Dunnett (α = 0,05).

ISOLAMENTO E ANÁLISE DE SEQUÊNCIAS DE DOMÍNIO *VcK*

[114] Para pesquisar um gene similar a *domínio K* em mirtilo, utilizou-se um gene similar a *domínio K* de *Vitis vinifera* (acesso GenBank GU133633.1) como consulta para BLAST contra o banco de dados de EST de mirtilo (http://www.vaccinium.org/tools/blast). Com base no valor E, a

coincidência superior foi uma sequência EST com 1189 bp (*Vaccinium_corymbosum*_v1_Contig03391), que contém um motivo de caixa MADS e um domínio K. Comparação com genes similares a *domínio K* de outras espécies vegetais indica que o Contig03391 contém a sequência completa de um gene similar a *domínio K* de mirtilo (*domínio VcK*) e é agrupado no conjunto de genes similares a *domínio K* em conjunto com *MADS1* (*Nicotiana tabacum*) e *MADS8* (*Vitis vinifera*) com base na análise filogenética (Fig. 1A).

[115] Para examinar a função do domínio K em um gene SOC1, o domínio K do *domínio VcK*, ORF com 471 bp que codifica 157 aminoácidos, foi clonado. O *domínio VcK* clonado contém três faltas de coincidência de nucleotídeos (por exemplo, posições 51, 171 e 178) do Contig03391. A falta de coincidência resultou em substituição de aminoácidos. Alinhamento de múltiplas sequências de nove sequências de aminoácidos selecionadas (por exemplo, seis de plantas lenhosas e uma de *Arabidopsis*) de ortólogos de *domínio VcK* indica que o *domínio VcK* é altamente conservado nessas espécies vegetais (Fig. 1B).

EVIDÊNCIA MOLECULAR DA SOBRE-EXPRESSÃO DE DOMÍNIO VCK

[116] PCR RT confirmou a transcrição do *domínio VcK* (Fig. 2). Em folhas jovens de plantas transgênicas (T0 selecionadas) e não transgênicas (NT), foram observadas faixas consistentes de gene *actina* endógeno; por outro lado, observou-se fragmento com 474 bp esperado do *domínio VcK* em todas as plantas transgênicas testadas, mas ele foi ausente em plantas não transformadas (Fig. 2). A Figura 2 exibe as linhas a seguir: Linhas 1 e 5: evento transgênico I5-4; Linhas 2 e 6: evento transgênico I5-6; Linhas 3 e 7: NT; Linhas 4 e 8: água como modelo; e Linha M: escada de DNA de 1 kb.

ALTERAÇÕES FENOTÍPICAS DE PLANTAS DE FUMO TRANSGÊNICAS COM DOMÍNIO VCK

[117] A expressão ectópica de *domínio VcK* não resultou em

alteração morfológica da identidade de órgãos florais, mas promoveu a floração precoce em plantas de fumo transgênicas (Figs. 3A-C). Dos 22 eventos transgênicos T0 regenerados, todos possuíam flores normais, da mesma forma que plantas T1 derivadas de cinco eventos (I5-5, I5-8, I5-16, I5-18 e I5-22) selecionadas para estudo contínuo. As plantas T1 de quatro desses eventos selecionados floresceram em 18 dias em média (variando de 15 a 21 dias, dependendo do evento) mais cedo que as plantas do tipo selvagem, que floresceram em 129 dias após o cultivo (Fig. 3B, C) .O resultado sugere que *domínio VcK* está envolvido no tempo de floração.

[118] Embora nenhuma alteração morfológica óbvia fosse observada em flores, plantas transgênicas com *domínio VcK* realmente exibiram alteração do crescimento. Todos os quatro eventos transgênicos T1 de floração precoce exibiram redução significativa da altura da planta e uma apresentou significativamente menos folhas no momento da floração em comparação com plantas do tipo selvagem (Fig. 3D-F).

[119] A expressão de 35S-*domínio VcK* não afetou significativamente a quantidade de vagens de sementes em cinco linhas de plantas T1 avaliadas (P = 0,0501), o que sugere que a expressão do 35S-*domínio VcK* apresenta pouco impacto sobre a produção de frutas ou sementes por planta.

[120] A produção de sementes de vinte plantas T0 foi comparada com plantas não transgênicas. A produção média de sementes para vinte eventos transgênicos foi de 8,02 gramas, que é 32% mais alta que plantas do tipo selvagem (6,08 g/planta), embora não aja diferença estatística entre as plantas transgênicas e plantas do tipo selvagem devido às variações entre diferentes eventos transgênicos (Tabela 2).

TABELA 2

[121] Produção de sementes de plantas de fumo transgênicas que

sobre-expressam o domínio VcK:

	Produção média de sementes por planta (grama) ± desvio padrão
Plantas T0 de vinte eventos transgênicos	8,02 ± 2,71 a
Três plantas do tipo selvagem	6,08 ± 0,42 a

[122] Dos vinte eventos transgênicos independentes de plantas de fumo, rendimentos de semente de quatorze eventos (70%) foram mais altos que a planta do tipo selvagem. Os resultados sugerem alto potencial (mais de 70% de possibilidade) de obtenção de eventos transgênicos individuais com maior produção de sementes.

EXEMPLO 2

CLONAGEM DOS HOMÓLOGOS DE *VcSOC1* COM COMPRIMENTO TOTAL

[123] Seguindo os protocolos de clonagem de *VcK* conforme descrito no Exemplo 1, os primers BOC-FX (5'-tctagaATG GTGAGA GGGAAAACCCAGA-3': SEQ ID NO: 15) e BOC-RS (5'-gagctcACTTGGGTGGCGAAACT-3'; SEQ ID NO: 16) foram utilizados para amplificação de homólogos de *VcSOC1* de cDNA de cultivar de mirtilo Bluecrop. Produtos de PCR foram clonados em um vetor pCR2.1-TOPO (Invitrogen) e verificados por meio de sequenciamento. Esses homólogos de *VcSOC1* foram identificados como *VcSOC1A* (SEQ ID NO: 17), *VcSOC1B* (SEQ ID NO: 18) e *VcSOC1D* (SEQ ID NO: 19). Os fragmentos de 5'-Xba I-VcSOC1-Sac I-3' foram liberados dos vetores pCR2.1-TOPO e inseridos em seguida nos locais *Xba*I e *Sac*I abertos na região T-DNA entre o promotor 35S do vírus mosaico da couve-flor (CaMV) e o terminal Nos em pBI121, em que a região de codificação de GUS (*β-glucuronidase*) foi removida. Os vetores resultantes foram introduzidos independentemente em linhagem EHA105 de *Agrobacterium tumefaciens* e conforme descrito no Exemplo 1. Conduziu-se transformação independente dos três homólogos de *VcSOC1* em fumo seguindo os protocolos de transformação

de *domínio VcK* conforme descrito no Exemplo 1.

ANÁLISES MOLECULARES

[124] Para avaliar a expressão dos três homólogos de *VcSOC1* em fumo, o par de primers SOF1 e SOK (conforme exibido na Tabela 1) foi utilizado para análises de PCR de transcrição reversa (PCR RT) de plantas T0. A análise de PCR RT foi conduzida conforme descrito no Exemplo 1.

FENOTIPIFICAÇÃO DE PLANTAS DE FUMO TRANSGÊNICAS

[125] A fenotipificação de plantas de fumo transgênicas transformadas com os três homólogos de VcSOC1 foi conduzida em estufa sob condições de luz natural no outono de Michigan. Foram utilizados três eventos transgênicos T0 para cada um dos homólogos de *VcSOC1*. Todas essas plantas foram positivas para PCR RT para o *VcSOC1*. Cada evento proliferou-se para obter três plantas por meio de cultivo *in vitro* de cortes entre nós sobre meio MS; plantas não transgênicas e uma planta de *domínio VcK* T1 foram utilizadas como controle. Essas plantas foram cultivadas em 10 vasos de 10 cm. Os dias para a floração (dias entre o cultivo e o surgimento do primeiro botão de flor), número de folhas, altura da planta e número de vagens de sementes foram registrados. Determinou-se a sobre-expressão ectópica por meio da presença de 657 pares de bases para *VcSOC1A*, 576 pares de bases para *VcSOC1B* e 654 pares de bases para *VcSOC1D*.

[126] A sobre-expressão ectópica dos três homólogos de *VcSOC1* (*VcSOC1A*, *VcSOC1B* e *VcSOC1D*) exibiu função similar à sobre-expressão do *domínio VcK*, tal como floração precoce (uma a três semanas antes dos controles não transgênicos) sem reduzir significativamente o número de folhas e a quantidade de vagens de sementes por planta de fumo transgênica T0.

EXEMPLO 3

[127] A construção de domínio VcK descrita no Exemplo 1 foi utilizada para sobre-expressar ectopicamente o *VcK* em plantas de mirtilo. Com

este propósito, a transformação de um cultivar de mirtilo de arbusto alto do norte Aurora foi realizada conforme relatado em Song e Sink (*Agrobacterium tumefaciens-mediated transformation of blueberry (Vaccinium corymbosum L.)*, *Plant Cell Rep.* 23: 475-484 e Song, G-Q, Sink, K. C. (2006), *Agrobacterium-mediated transformation of higbush blueberry (Vaccinium corymbosum L.) cultivars* em Wang, K., ed., *Agrobacterium protocols: Methods in Molecular Biology* 344, segunda edição, Totowa NJ: Humana Press, 37-44. Resumidamente, explantes de folhas foram cocultivados com EHA105 por seis dias sobre meio de cocultivo na presença de 100 µM de acetossiringona a 25 ºC no escuro. Conduziu-se seleção e regeneração sobre meio de regeneração (meio de plantas lenhosas de McCown modificado (WPM) contendo 4,54 µM de tidiazuron, 2,69 µM de ácido α-naftalenoacético, 10 mg l^{-1} Km e 250 mg l^{-1} cefotaxima) por doze semanas. Subcultivos de meio de seleção novo foram conduzidos em intervalos de quatro semanas. Brotos transgênicos de explantes separados foram utilizados como eventos transgênicos independentes. A proliferação de brotos resistentes a canamicina individuais foi conduzida sobre WPM contendo 4,56 µM de zeatina, 50 mg l^{-1} Km e 250 mg l^{-1} cefotaxima. Controles do tipo selvagem foram regenerados a partir de explantes de folhas não infectados e controles transgênicos foram transformados com um gene *gusA* no lugar da construção *35S-VcK*.

[128] Seis plantas para cada um dos três eventos transgênicos T0 transformados com o *VcK* foram obtidas por meio de micropropagação e cultivadas em vasos de 3,78 l (17,78 cm (altura) x 20,32 cm (diâmetro)) em estufa sob condições de luz natural. Plantas não transgênicas e plantas transgênicas transformadas com pBI121 foram utilizadas como controles. A altura das plantas, número de ramos, número de botões em floração e a data de floração foram registrados.

[129] Plantas transgênicas exibiram floração acelerada e

floresceram sob condições sem resfriamento, em que as plantas não transgênicas não floresceram. Isso sugere que a sobre-expressão do domínio *VcK* é suficiente para promover floração de mirtilos acelerada.

EXEMPLO 4

[130] A construção de domínio VcK conforme descrito no Exemplo 1 foi utilizada para sobre-expressar ectopicamente o *VcK* em plantas de petúnia. A transformação mediada por *Agrobacterium de Petunia hybrida* Mitchell foi realizada essencialmente conforme descrito por Jorgensen et al (1996). Resumidamente, explantes vegetais com 2-3 x 2-4 mm foram extirpados de mudas *in vitro* cultivadas sobre meio básico MS com meia resistência (Murashige e Skoog, 1962). Conduziu-se cocultivo por quatro dias sobre meio de regeneração (RM: MS + 4,4 l M BAP + 1,1 l M IAA) suplementado com 100 l M acetossiringona. Após seleção com 250 mg l^{-1} Timentin e 100 mg l canamicina, brotos transgênicos foram produzidos sobre RM e enraizados sobre MS. Os cultivos foram mantidos a 25 ºC sob fotoperíodo de luz de 16 horas. Vetor vazio (pGA643) e explantes não transformados foram também conduzidos através de cultivo de tecido como controles negativos. Sementes T1 foram colhidas de plantas T0 separadas.

[131] Realizou-se fenotipificação de plantas de petúnia entre janeiro e junho em estufa sob condições de luz natural em Michigan. Quinze mudas T1 de dois eventos independentes transformados com o *VcK* foram cultivadas em 10 vasos de 10 cm. A presença de *VcK* em mudas transgênicas foi confirmada por meio de PCR utilizando primers 35S-F (SEQ ID NO: 8) e SOK (SEQ ID NO: 7) (Tabela 1). Quinze plantas para cada muda transgênica e não transgênica transformada com pGA643 foram utilizadas como controles. Dias para floração (dias entre o cultivo e o aparecimento do primeiro botão de flor), número de flores, ramos e altura da haste central foram registrados (Tabela 3).

TABELA 3

[132] Efeito da expressão ectópica de *35S:VcK* sobre alterações fenotípicas em plantas de petúnia transgênicas:

Planta	Dias para a emergência das primeiras flores após o cultivo das sementes	Número de ramos da planta por planta (pesquisada em 23 de junho de 2014)	Número total de flores por planta (pesquisada em 23 de junho de 2014)
Plantas transgênicas T1 que sobre-expressam *VcK*	74	5,1	76,4
Tipo selvagem	83	2,6	48,2
Plantas controle transgênicas T1 transformadas com pGA643	82	2,5	48,3

[133] Em comparação com *Petunia hybrida* Mytchell do tipo selvagem e plantas controle transgênicas contendo pGA643, plantas de petúnia que sobre-expressam *VcK* floresceram 3-9 dias mais cedo, apresentaram mais ramos e mais flores por planta. Essas plantas transgênicas também não exibiram redução significativa da altura da planta e do número de folhas sobre a haste central.

[134] Nas diversas realizações, é fornecido um método de produção de plantas transgênicas que compreende a introdução de um polipeptídeo de domínio K codificado por uma sequência de nucleotídeos para produzir uma célula transformada. Em várias realizações, o polinucleotídeo é selecionado a partir de uma sequência de nucleotídeos descrita em SEQ ID NO: 1; uma sequência de nucleotídeos que codifica o polipeptídeo de SEQ ID NO: 2;

uma sequência de nucleotídeos que compreende identidade de sequências de pelo menos 80% para a sequência descrita em SEQ ID NO: 1; ou uma sequência de nucleotídeos que codifica uma sequência de aminoácidos que possui identidade de sequências de pelo menos 90% para a sequência descrita em SEQ ID NO: 2. Em várias realizações, a célula transformada é utilizada para gerar a planta transgênica, em que a planta transgênica aumentou a expressão do polipeptídeo de domínio K em comparação com uma planta controle. Em outras realizações, as sequências utilizadas no método incluem a introdução de SEQ ID NO: 17, SEQ ID NO: 18 ou SEQ ID NO: 19.

[135] Em uma realização, a sequência de domínio K é uma sequência de domínio K de um gene de caixa MADS.

[136] Em uma realização, a planta utilizada para produzir a planta transgênica é uma planta frutífera.

[137] Em uma realização, o método utilizado para fornecer a planta transgênica é uma planta de mirtilo. Em outras realizações, a planta é selecionada a partir de milho, arroz, sorgo, painço, soja, fumo, batata, trigo, algodão, amendoim, centeio, painço proso, painço rabo de raposa, ragi, girassol, açafrão, trigo, batata doce, mandioca, abacaxi, árvores cítricas, cacau, chá, banana, abacate, figo, cana de açúcar, aveia, cevada, legumes, plantas ornamentais e coníferas. Em ainda outras realizações, a planta é selecionada a partir de plantas frutíferas tais como uva, laranja, limão, toronja, abacate, tâmara, pêssego, cereja, azeitona, ameixa, coco, maçã, pera, amoreira preta, framboesa, morango, abacaxi, tomate, pepino e berinjela.

[138] Em várias realizações, o método inclui plantas transgênicas que compreendem partes de plantas selecionadas a partir de embrião, pólen, óvulo, semente, folha, flor, ramo, fruto, caroço, espiga, sabugo, casca, haste, raiz, pontas de raízes, antera e suas combinações.

[139] Em várias realizações, o método utilizado para produzir

plantas transgênicas é a transformação estável em plantas com domínio K e variantes do domínio K.

[140] Em várias realizações, a planta transgênica possui pelo menos uma característica selecionada a partir de tempo de floração acelerado, tamanho de planta reduzido, número de folhas reduzido e suas combinações.

[141] Em diversas realizações, o método de produção de plantas transgênicas com um polipeptídeo de domínio K nelas introduzido exibe maior rendimento em comparação com plantas controle.

[142] Em várias realizações, a planta transgênica exibe maior rendimento em pelo menos 5%, 7%, 8% ou 9%, em comparação com plantas controle. Em outras realizações, a planta transgênica exibe maior rendimento em pelo menos cerca de 10 a 20% em comparação com plantas controle. Em ainda outras realizações, a planta transgênica exibe maior rendimento em cerca de 30% a cerca de 50% em comparação com plantas controle.

[143] Em várias realizações, um cassete de expressão útil para a produção de plantas transgênicas que compreendem um polinucleotídeo que codifica um polipeptídeo de domínio K ligado operacionalmente a um promotor que dirige a expressão em plantas, em que o polinucleotídeo é selecionado a partir de:

- uma sequência de nucleotídeos descrita em SEQ ID NO: 1;
- uma sequência de nucleotídeos que codifica o polipeptídeo de SEQ ID NO: 2;
- uma sequência de nucleotídeos que compreende identidade de sequências de pelo menos 80% com a sequência descrita em SEQ ID NO: 1; e
- uma sequência de nucleotídeos que codifica uma sequência de aminoácidos que possui identidade de sequências de 90% com a sequência descrita em SEQ ID NO: 2.

[144] Em outras realizações, as sequências empregadas no cassete de expressão incluem SEQ ID NO: 3, SEQ ID NO: 17, SEQ ID NO: 18 ou SEQ ID NO: 19. Em uma realização, o cassete de expressão compreende um promotor constitutivo.

[145] Em uma realização, uma célula hospedeira isolada é transformada com o cassete de expressão no qual um polipeptídeo de domínio K ligado operacionalmente a um promotor que dirige a expressão em plantas, em que o polinucleotídeo é selecionado a partir de:

- uma sequência de nucleotídeos descrita em SEQ ID NO: 1;
- uma sequência de nucleotídeos que codifica o polipeptídeo de SEQ ID NO: 2;
- uma sequência de nucleotídeos que compreende identidade de sequências de pelo menos 80% com a sequência descrita em SEQ ID NO: 1; e
- uma sequência de nucleotídeos que codifica uma sequência de aminoácidos que possui identidade de sequências de pelo menos 90% com a sequência descrita em SEQ ID NO: 2.

[146] Em várias realizações, diversas partes de plantas podem ser transformadas com o cassete de expressão. Em ainda outras realizações, as várias partes de plantas compreendem partes de plantas selecionadas a partir de embriões, pólen, óvulos, sementes, folhas, flores, ramos, frutos, caroços, espigas, sabugos, casca, hastes, raízes, pontas de raízes, anteras e suas combinações.

[147] Em várias realizações, a planta transgênica transformada com o cassete de expressão exibe pelo menos uma característica selecionada a partir de tempo de floração celerado, redução do tamanho da planta, redução do número de folhas e suas combinações.

[148] Várias realizações fornecem uma planta transgênica que

compreende uma planta transformada com um polipeptídeo de domínio K codificado por uma sequência de nucleotídeos descrita em SEQ ID NO: 1, em que a planta transgênica apresenta aumento da expressão do polipeptídeo de domínio K em comparação com uma planta controle. Em outras realizações, a planta pode ser transformada com SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 17, SEQ ID NO: 18 ou SEQ ID NO: 19.

[149] Em uma realização, a planta transgênica é uma planta frutífera transgênica. Em outras realizações, a planta transgênica é uma planta de mirtilo. Em outras realizações, as plantas incluem uva, laranja, limão, toronja, abacate, tâmara, pêssego, cereja, azeitona, ameixa, coco, maçã, pera, amoreira preta, framboesa, morango, abacaxi, tomate, pepino e berinjela.

[150] Em ainda outras realizações, a planta transgênica é selecionada a partir de milho, arroz, sorgo, painço, soja, fumo, batata, trigo, algodão, amendoim, centeio, painço proso, painço rabo de raposa, ragi, girassol, açafrão, trigo, batata doce, mandioca, abacaxi, árvores cítricas, cacau, chá, banana, abacate, figo, cana de açúcar, aveia, cevada, legumes, plantas ornamentais e coníferas. Em várias realizações, a planta transgênica possui pelo menos uma característica alterada selecionada a partir de tempo de floração acelerado, redução do tamanho da planta, redução do número de folhas e suas combinações.

[151] Em outras realizações, a planta transgênica exibe aumento do rendimento em comparação com planta controle. Em uma realização, a planta transgênica exibe maior rendimento em comparação com plantas que não expressam o domínio K. Em ainda outras realizações, a planta transgênica exibe aumento de rendimento de pelo menos 5%, 7%, 8% ou 9% em comparação com plantas controle. Em outras realizações, a planta transgênica exibe aumento de rendimento de pelo menos cerca de 10 a 20% em comparação com plantas controle. Em ainda outras realizações, a planta transgênica exibe aumento de

rendimento de cerca de 30% a cerca de 50% em comparação com plantas controle.

[152] Todas as patentes, pedidos de patente e literatura mencionados no presente relatório descritivo são integralmente incorporados ao presente como referência. Em caso de qualquer inconsistência, prevalecerá o presente relatório descritivo, incluindo quaisquer definições do presente. Embora realizações específicas tenham sido ilustradas e descritas no presente, os técnicos comuns no assunto apreciarão que qualquer procedimento calculado para atingir o mesmo propósito pode ser substituído pelas realizações específicas exibidas. Embora o processo tenha sido discutido utilizando sequências específicas, por exemplo, podem ser utilizadas variações das sequências e de diferentes fontes. Além disso, embora o processo tenha sido discutido utilizando principalmente fumo, petúnia e mirtilo como plantas, outras plantas como soja e milho, por exemplo, podem ser utilizadas. O presente pedido destina-se a cobrir quaisquer adaptações ou variações do objeto do presente. Pretende-se declaradamente, portanto, que as realizações da presente invenção sejam limitadas apenas pelas reivindicações e seus equivalentes.

REIVINDICAÇÕES

1. MÉTODO PARA PRODUZIR UMA PLANTA TRANSGÊNICA, caracterizado por compreender:

introduzir uma sequência de nucleotídeos conforme descrita na SEQ ID NO: 1, ou as sequências degeneradas da mesma, que codifica um polipeptídeo de domínio K heterólogo de sequência de aminoácidos representada pelas posições 83-169 da SEQ ID NO: 2 em uma célula de planta para produzir uma célula transformada; e

regenerar a planta transgênica a partir da célula transformada.

2. MÉTODO, de acordo com a reivindicação 1, caracterizado pelo polipeptídeo de domínio K ser um domínio K de um polipeptídeo caixa MADS.

3. MÉTODO, de acordo com a reivindicação 1, caracterizado pela planta ser uma planta frutífera.

4. MÉTODO, de acordo com a reivindicação 3, caracterizado pela planta frutífera ser mirtilo (*blueberry*).

5. MÉTODO, de acordo com a reivindicação 3, caracterizado pela planta frutífera ser selecionda de uma uva, laranja, limão, toronja, abacate, tâmara, pêssego, cereja, azeitona, ameixa, coco, maçã, pera, amoreira preta, framboesa, morango, abacaxi, tomate, pepino ou berinjela.

6. MÉTODO, de acordo com a reivindicação 1, caracterizado pela planta compreender partes de planta selcionadas de um embrião, pólen, óvulo, semente, folha, flor, ramo, fruto, caroço, espiga, sabugo, casca, haste, raiz, ponta de raíz, antera, e combinações dos mesmos.

7. MÉTODO, de acordo com a reivindicação 1, caracterizado pela planta transgênica exibir aumento de rendimento em comparação a uma planta controle.

8. MÉTODO, de acordo com a reivindicação 1, caracterizado

pela planta ser selecionada de algodão, milho, arroz, soja, canola, beterraba sacarina, girassol, batata e cártamo.

9. CASSETE DE EXPRESSÃO, caracterizado por compreender:

um polinucleotídeo que compreende uma sequência de nucleotídeos conforme descrita na SEQ ID NO: 1, ou as sequências degeneradas da mesma, que codifica um polipeptídeo de domínio K de sequência de aminoácidos representada pelas posições 83-169 da SEQ ID NO:2, ligado operacionalmente a um promotor que dirige a expressão em uma planta,

em que o cassete de expressão compreende adicionalmente um marcador selecionável heterólogo.

10. CASSETE DE EXPRESSÃO, de acordo com a reivindicação 9, caracterizado pelo promotor ser um promotor constitutivo.

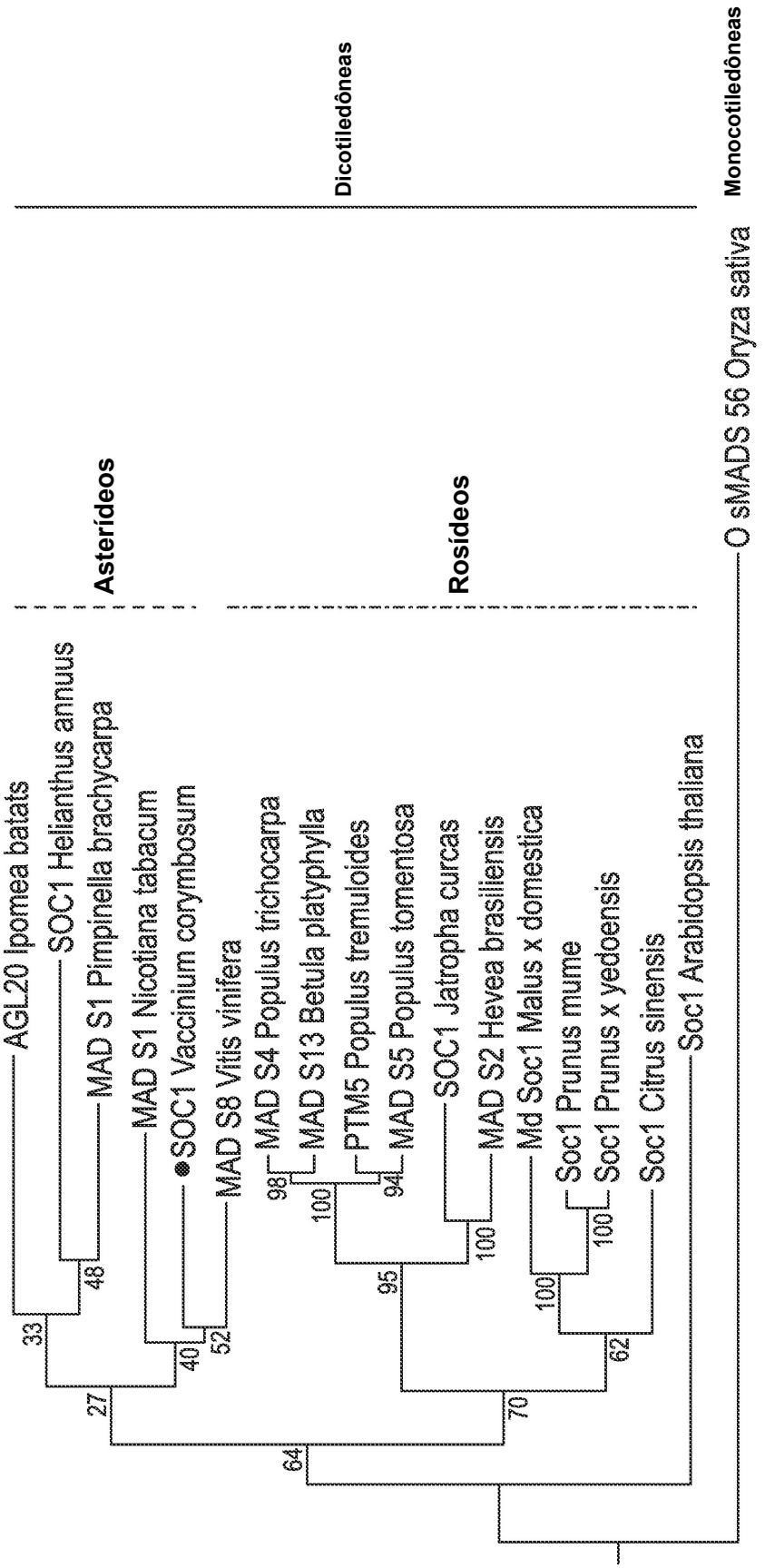

Dicotiledôneas
Monocotiledôneas
Asterídeos
Rosídeos
AGL20 Ipomea batats
SOC1 Helianthus annuus
MAD S1 Pimpinella brachycarpa
MAD S1 Nicotiana tabacum
SOC1 Vaccinium corymbosum
MAD S8 Vitis vinifera
MAD S4 Populus trichocarpa
MAD S13 Betula platyphylla
PTM5 Populus tremuloides
MAD S5 Populus tomentosa
SOC1 Jatropha curcas
MAD S2 Hevea brasiliensis
Md Soc1 Malus x domestica
Soc1 Prunus mume
Soc1 Prunus x yedoensis
Soc1 Citrus sinensis
Soc1 Arabidopsis thaliana
O sMADS 56 Oryza sativa
33
48
40
52
27
64
98
100
94
95
100
100
100
62
70

FIG. 1A

```
SOC1_Vaccinium_corymbosum       EQNMQHLKHEAADMSKKIEHLEVAKRKLLGEGLGSCTFEELIQIEQQLEQ
MADS1_Nicotiana_tabacum         EQNLQHMQHAAASLMKKIELLEESKRKLLGEGLQSCSLVELQQIEKQLER
PTM5_Populus_trichocarpoa       EQNMLQLKEEAASMIKKIEHLEVSKRKLLGECLGSCTVEELQQIEQQLER
PTM5_Populus_tremuloides        EEQNMLQLKEEAASMIKKIEHLEVSKRKLLGECLGSCTIEELQQIEQQLER
SOC1-like_Prunus-mume           DQNMQHKQESSSMMKQIELLEVSKRKLLGEGLGSCSIEELQEIEQQLER
SOC1-like_Malus_x_domestica     EQNMQHLKQEATSMMKQIELLEVSKRKLLGEGLGSCTLAELQEIEDQLEK
SOC1_Arabidopsis_thaliana       EENMQHLKYEAANMMKKIEQLEASKRKLLGEGIGTCSIEELQQIEQQLEK
SOC1_Vitis_vinifera             EHNMQHLKHEAANMAKKIELLEISKRKLLGEGLGSCSIEELQQIEQQLER
SOC1-like_Citrus_sinensis       ERYMQQLKHEIANMIEKIEHIEVSQRKLLGQDLGSRTNEELQELDDQLER
                                :. : :::   :.: ::** :* ::*****: : : :  ** :::.***:

SOC1_Vaccinium_corymbosum       SASTIRARKMQVFREQIEKLKEKEKALEAENAML
MADS1_Nicotiana_tabacum         SVSTIRARKIQVFKEQIERLKEKEKILASENAAIL
PTM5_Populus_trichocarpoa       SVSTIRARKNQVFKEQIELLRQKEKLLAAENARL
PTM5_Populus_tremuloides        SVSTIRARKNQVFKEQIELLKQKEKLLAAENARL
SOC1-like_Prunus-mume           SVSNVRARKTQVFKEQIEQLREKGKALAAENEKL
SOC1-like_Malus_x_domestica     SVYNVRARKSQVFKEQIEQLREKEKLLTAENTRL
SOC1_Arabidopsis_thaliana       SVKCIRARKTQVFKEQIEQLKQKEKALAAENEKL
SOC1_Vitis_vinifera             SVSSIRARKNQVFKEQIEQLKEKEKALAAENAML
SOC1-like_Citrus_sinensis       SLRSIRARKAQLFNEQMGQLKEKERLLLEDNARL
                                *   :**** *:*.**:  *::* :    :*  *
```

FIG. 1B

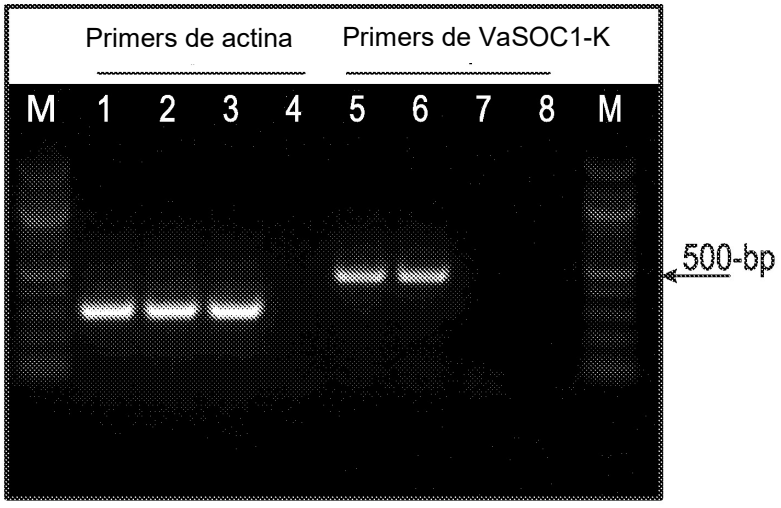


FIG. 2

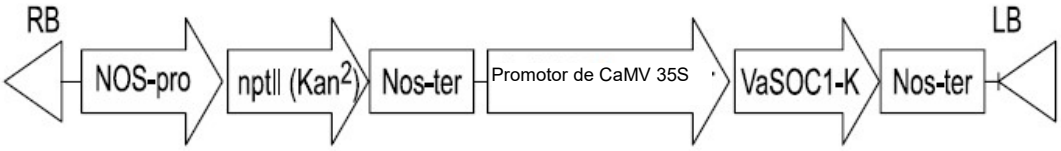

RB
NOS-pro
nptII (Kan2)
Nos-ter
Promotor de CaMV 35S
VaSOC1-K
Nos-ter
LB

FIG. 3A

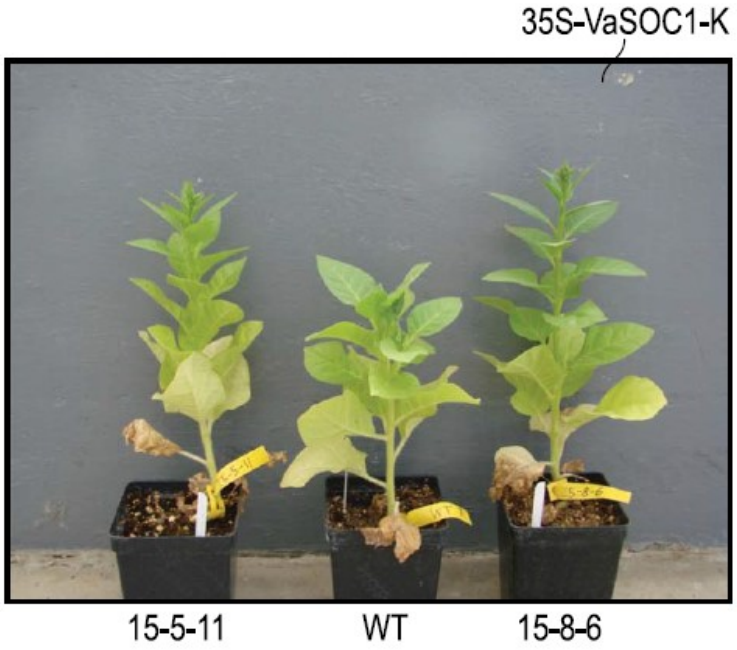

35S-VaSOC1-K
15-5-11
WT
15-8-6

FIG. 3B

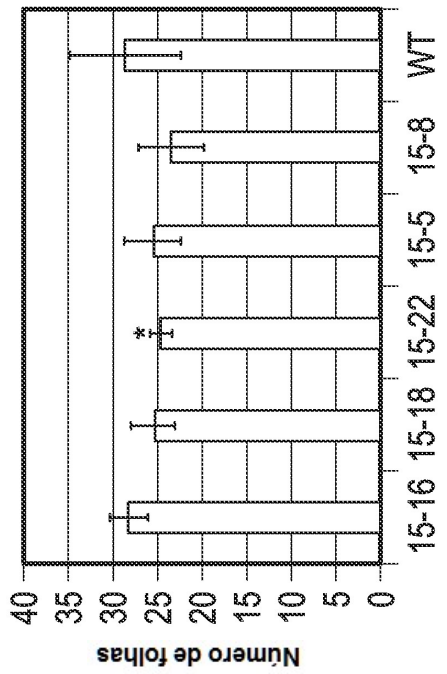


FIG. 3D

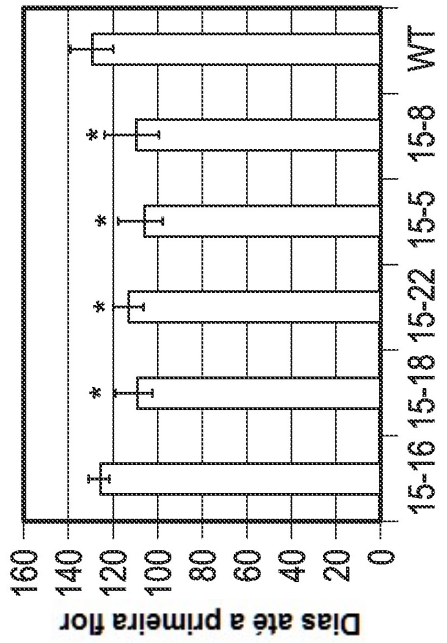


FIG. 3C

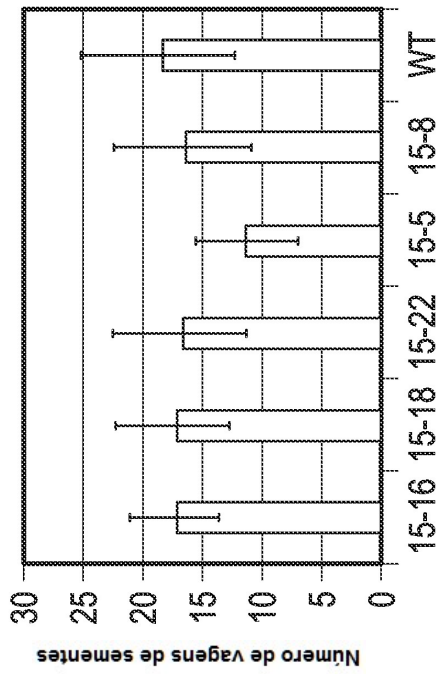


FIG. 3F

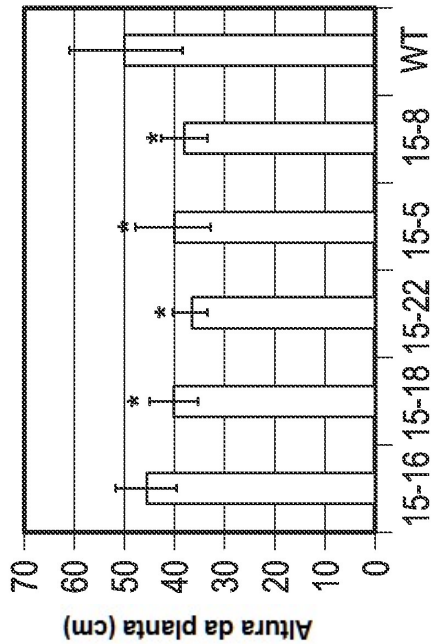


FIG. 3E