

República Federativa do Brasil
Ministério do Desenvolvimento, Indústria
e do Comércio Exterior
Instituto Nacional da Propriedade Industrial

(21) **PI0619939-9 A2**

(22) Data de Depósito: 12/12/2006
(43) Data da Publicação: 25/10/2011
(RPI 2129)



(51) *Int.Cl.:*
C12N 15/82
C12N 9/88
A01H 5/00

(54) Título: GENE QUIMÉRICO, VETOR, USO DO PROMOTOR DE UM GENE DE CISTEÍNA SINTASE CITOSSÓLICA DE PLANTA, E, MÉTODO PARA FABRICAR UMA PLANTA TRANSGÊNICA OU CÉLULA DE PLANTA

(30) Prioridade Unionista: 16/12/2005 NL pct/nl2005/050083

(73) Titular(es): Keygene N.V.

(72) Inventor(es): Ludvik Kevin Wachowski, Michiel Theodoor Jan de Both, Nicolette Elisabeth Quaadvlieg

(74) Procurador(es): Momsen, Leonardos & CIA.

(86) Pedido Internacional: PCT NL2006050314 de 12/12/2006

(87) Publicação Internacional: WO 2007/069894 de 21/06/2007

(57) Resumo: GENE QUIMÉRICO, VETOR, USO DO PROMOTOR DE UM GENE DE CISTEÍNA SINTASE CITOSSÓLICA DE PLANTA, E, MÉTODO PARA FABRICAR UMA PLANTA TRANSGÊNICA OU CÉLULA DE PLANTA. Promotores de plantas constitutivos e fortes são providos, referidos aqui como promotores de AA6, que permanecem fortes e constitutivos sob condições de estresse biótico e/ou abiótico. Também são providos células e organismos transgênicos, especialmente células de plantas e plantas, compreendendo um promotor de AA6 e métodos para expressar as seqüências de ácido nucleico em células e organismos usando promotores de AA6.

“GENE QUIMÉRICO, VETOR, USO DO PROMOTOR DE UM GENE DE CISTEÍNA SINTASE CITOSSÓLICA DE PLANTA, E, MÉTODO PARA FABRICAR UMA PLANTA TRANSGÊNICA OU CÉLULA DE PLANTA”

5 CAMPO DA INVENÇÃO

A presente invenção refere-se a uma classe de promotores de planta, constitutivos e fortes, e seus usos. Os promotores podem ser usados para expressão de proteínas homólogas ou heterólogas em plantas ou células de plantas, ou para a expressão de moléculas de ácido nucleico ativas, como RNA sentido e/ou anti-sentido. São providas seqüências de ácido nucleico tendo atividade de promotor, assim como genes quiméricos, vetores e células recombinantes (transgênicas) e organismo compreendendo os mesmos. Também são providos métodos para fabricar células transgênicas e organismos, especialmente plantas e células de plantas, compreendendo o promotor. Além disso, as proteínas de cisteína sintase citossólica e seqüências de ácido nucleico codificando as mesmas, são providas.

ANTECEDENTES DA INVENÇÃO

Um grande número de promotores de plantas são conhecidos na técnica, que são ferramentas utilizáveis para a expressão de proteínas ou peptídeos em plantas transgênicas ou células de plantas ou para o silenciamento de genes ou famílias de genes. Estes incluem promotores constitutivos, promotores indutíveis, promotores regulados no desenvolvimento, promotores preferidos de tecido ou específicos de tecido. Os exemplos de promotores constitutivos comumente usados são os seguintes: os promotores 35S ou promotores 35S melhorados (os "promotores 35S ") do vírus do mosaico da couve-flor (CaMV) de isolados

CM 1841 (Gardner *et al.*, 1981, Nucleic Acids Research 9, 2871-2887), CabbB-S (Franck *et al.*, 1980, Cell 21, 285-294) e CabbB-JI (Hull e Howell, 1987, Virology 86, 482-493); o promotor 35S descrito por Odell *et al.* (1985, Nature 313, 810-812) ou em US5164316, promotores da família de ubiquitina (por exemplo o promotor de ubiquitina de milho de Christensen *et al.*, 1992, Plant Mol. Biol. 18, 675-689, EP 0 342 926, ver também Cornejo *et al.* 1993, Plant Mol. Biol. 23, 567-581), o promotor *gos2* (de Pater *et al.*, 1992 Plant J. 2, 837-844), o promotor *emu* (Last *et al.*, 1990, Theor. Appl. Genet. 81, 581-588), promotores de actina de *Arabidopsis* como os promotores descritos por An *et al.* (1996, Plant J. 10, 107-121), promotores de actina de arroz, como os promotores descritos por Zhang *et al.* (1991, The Cell Planta 3, 1155-1165) e o promotor descrito em US 5,641,876 ou promotor de actina de arroz 2 como descrito em WO 00/70067; promotores do vírus do mosaico da veia de Cassava (WO 97/48819, Verdaguer *et al.* 1998, Plant Mol. Biol. 37, 1055-1067), a série pPLEX de promotores de vírus de retardo do crescimento do trevo subterrâneo (WO 96/06932, particularmente o promotor S7), um promotor álcool desidrogenase, por exemplo, pAdh1S (números de acesso ao GenBank X04049, X00581), o promotor do vírus do mosaico da escrofulária descrito em US 6 051 753 e em EP 426 641, promotores do gene histona, como o promotor Ph4a748 de *Arabidopsis* (Plant Mol. Biol. 8: 179-191), o promotor CoYMV (vírus do mosqueado amarelo de Commelina) (Medberry *et al.* 1992, The Cell Plant 4:185-192), ou outros.

No entanto, um inconveniente comum de promotores constitutivos conhecidos é que eles com frequência ainda mostram variação em sua atividade, com expressão regulada no órgão ou desenvolvimento ou alterações induzidas por estresse na expressão. Os presentes inventores verificaram que a atividade do promotores CaMV 35S em plantas transgênicas era sensível ao estresse abiótico, especialmente estresse por calor causado quando as plantas transgênicas foram cultivadas no campo na

Espanha. Assim, existe uma necessidade para prover novos promotores constitutivos, que são insensíveis a um ou mais estresses bióticos e/ou abióticos. Além disso, é desejável de um ponto de vista regulatório usar os promotores derivados de plantas na geração de plantas transgênicas.

5 Além disso, os promotores virais derivados de vírus capazes de infectar plantas são menos preferidos para a transformação das espécies de plantas hospedeiras, como infecção das plantas com o vírus pode causar silenciamento do promotor transgênico (Seemanpillai *et al.*, 2003, Mol Plant Microbe Interact. 16(5): 429-38; Al-Kaff *et al.*, 2000, Nat Biotechnol. 18: 10 995-9). Promotores constitutivos diferentes são também necessários para as abordagens de empilhamento de genes, como o uso de vários promotores idênticos pode resultar em silenciamento (Yang *et al.*, 2005, Plant Mol Biol. 58: 351-66).

SUMÁRIO DA INVENÇÃO

15 Provê-se uma planta transgênica ou célula de planta ou tecido de planta ou órgão compreendendo um gene quimérico integrado em seu genoma, caracterizados em que referido gene quimérico compreende um promotor constitutivo ligado operativamente a uma seqüência de ácido nucleico homóloga ou heteróloga, em que o promotor é selecionado dentre o grupo de :
20

(a) a seqüência de ácido nucleico de SEQ ID NO: 1 ou SEQ ID NO: 2, ou qualquer uma dentre SEQ ID NO: 6-9, ou a seqüência de promotor clonada no vetor pKG8135 ou pKG8137, ambos depositados junto ao CBS (Centraalbureau voor Schimmelcultures, Uppsalalaan 8, 3584 CT Utrecht, 25 Holanda em 3 de agosto de 2006 sob números de acesso CBS 120175 e CBS 120176, respectivamente);

(b) um fragmento funcional de SEQ ID NO: 1 ou SEQ ID NO: 2, ou de qualquer uma das seqüências mencionadas sob (a);

(c) a seqüência de ácido nucleico compreendendo pelo menos

70% identidade de seqüência com SEQ ID NO: 1 ou 2, ou para qualquer uma das seqüências mencionadas sob (a);

(d) um fragmento funcional de uma seqüência de ácido nucleico de (c).

5 Também é provida uma seqüência isolada de ácido nucleico tendo atividade de promotor quando introduzida em células de plantas, em que referida seqüência de ácido nucleico compreendendo uma seqüência selecionada dentre:

10 (a) a seqüência de ácido nucleico de SEQ ID NO: 1 ou SEQ ID NO: 2, ou qualquer uma dentre SEQ ID NO: 6-9, ou a seqüência de promotor clonada em vetor pKG8135 ou pKG8137, ambos depositados junto ao CBS (Centraalbureau voor Schimmelcultures, Uppsalalaan 8, 3584 CT Utrecht, Holanda em 3 de agosto de 2006 sob números de acesso CBS 120175 e CBS 120176, respectivamente);

15 (b) um fragmento de SEQ ID NO: 1 ou SEQ ID NO: 2, ou de qualquer uma das seqüências mencionadas sob (a);

(c) a seqüência de ácido nucleico compreendendo pelo menos 70% de identidade de seqüência com SEQ ID NO: 1 ou 2, ou para qualquer uma das seqüências mencionadas sob (a);

20 (d) um fragmento de uma seqüência de ácido nucleico de (c).

Vetores, gene quiméricos e células hospedeiras compreendendo as seqüências acima são também uma forma de realização da invenção.

25 Em outra forma de realização, o uso do promotor de um gene de cisteína sintase citossólica de planta para a expressão constitutiva de uma seqüência de ácido nucleico sentido e/ou anti-sentido em uma planta transgênica ou célula de planta ou tecido de planta ou órgão.

Além disso, é provido um método para fabricar uma planta transgênica ou célula de planta, compreendendo as etapas de:

(a) gerar um gene quimérico compreendendo um promotor como descrito acima, ligado operativamente a uma seqüência de ácido nucleico a ser transcrita, e opcionalmente ainda ligada a uma seqüência de ácido nucleico 3'UTR, ou um vetor compreendendo este gene quimérico;

(b) transformar uma planta ou célula de planta com referido gene quimérico ou vetor; e opcionalmente

(c) regenerar as plantas transgênicas ou células de planta.

Em ainda outra forma de realização, uma molécula de ácido nucleico isolado compreendendo a seqüência nucleotídica de SEQ ID NO: 3 ou 4 é provida.

DEFINIÇÕES GERAIS

O termo "seqüência de ácido nucleico" (ou molécula de ácido nucleico) refere-se a uma molécula de DNA ou RNA em forma de filamento único ou duplo, particularmente um DNA tendo atividade de promotor, de acordo com a invenção, ou um DNA codificando uma proteína ou fragmento de proteína. Uma "seqüência de ácido nucleico isolado" refere-se a uma seqüência de ácido nucleico que não está mais no ambiente natural do qual ela foi isolada, por exemplo, a seqüência de ácido nucleico em uma célula hospedeira bacteriana ou no genoma plastídeo ou nuclear da planta.

Os termos "proteína" ou "polipeptídeo" são usados de modo interpermutável e se referem a moléculas consistindo de uma cadeia de aminoácidos, sem referência a um modo específico de ação, tamanho, estrutura tri-dimensional ou origem. Um "fragmento" ou "porção" de uma proteína pode assim ainda ser referido como uma "proteína". Uma "proteína isolada" é usada aqui para fazer referência a uma proteína não está mais em seu ambiente natural, por exemplo in vitro ou em uma célula hospedeira de planta ou bacteriana recombinante.

O termo "gene" significa uma seqüência de DNA

compreendendo uma região (região transcrita), que é transcrita em uma molécula de RNA (por exemplo mRNA) em uma célula, ligada operativamente a regiões regulatórias de transcrição apropriadas (por exemplo um promotor). Um gene pode assim compreender várias seqüências ligadas operativamente, como um promotor, uma seqüência líder não traduzida 5' (também referida como 5' UTR, que corresponde à seqüência de mRNA transcrita a montante do códon de partida de tradução) compreendendo por exemplo seqüências envolvidas em iniciação de tradução, uma região de codificação (proteína (cDNA ou DNA genômico) e uma seqüência não traduzida 3' (também referida como região não traduzida 3', ou 3' UTR) compreendendo por exemplo sítios de terminação de transcrição e sítio de poliadenilação (como por exemplo AAUAAA ou variantes dos mesmos).

Um "gene quimérico" (ou gene recombinante) refere-se a qualquer gene, que não é normalmente encontrado na natureza em uma espécie, particularmente um gene em que uma ou mais partes da seqüência de ácido nucleico estão presentes que não são associadas uma com a outra na natureza. Por exemplo, o promotor não é associado na natureza com parte ou toda a região transcrita ou com outra região regulatória. O termo "gene quimérico" é entendido para incluir construções de expressão em que um promotor ou seqüência regulatória de transcrição é ligado operativamente a uma ou mais seqüências sentido (por exemplo seqüências de codificação) ou a uma seqüência anti-sentido ou de repetição invertida (complemento reverso do filamento sentido) (sentido e anti-sentido), pelo que o transcrito de RNA forma RNA de filamento duplo quando da transcrição).

Uma "3' UTR" ou "seqüência não traduzida 3'" (também com freqüência referido como uma região não traduzida 3' ou extremidade 3') refere-se a uma seqüência de ácido nucleico encontrada a jusante da seqüência de codificação de um gene, que compreende, por exemplo, um sítio de terminação de transcrição e (na maior parte, mas não todos os mRNAs

eucarióticos), um sinal de poliadenilação (como por exemplo AAUAAA ou variantes dos mesmos). Após terminação da transcrição, o transcrito de mRNA pode ser clivado a jusante do sinal de poliadenilação e uma cauda poli(A) pode ser adicionada, que está envolvida no transporte do mRNA para o citoplasma (onde a tradução ocorre).

"Expressão de um gene" refere-se ao processo em que uma região de DNA, que é ligada operativamente a regiões regulatórias apropriadas, particularmente um promotor, é transcrita em um RNA, que é biologicamente ativo, isto é, que é capaz de ser traduzido em uma proteína ou peptídeo biologicamente ativo (ou fragmento de peptídeo ativo) ou que é ativo sozinho (isto é, em silenciamento de gene pos-transcricional ou RNAi). Uma proteína ativa em algumas formas de realização refere-se a uma proteína tendo uma função negativa dominante devido a um domínio repressor estar presente. A sequência de codificação está preferivelmente em orientação sentido e codifica uma proteína ou peptídeo biologicamente ativo, desejado, ou um fragmento de peptídeo ativo. Nas abordagens de silenciamento de genes, a sequência de DNA está preferivelmente presente na forma de um DNA anti-sentido ou um DNA de repetição invertida, compreendendo uma sequência curta do gene alvo em anti-sentido ou em orientação sentido e anti-sentido. "Expressão ectópica" refere-se a uma expressão em um tecido em que o gene não é normalmente expressado.

Uma "sequência regulatória de transcrição" é aqui definida como uma sequência de ácido nucleico que é capaz de regular a taxa de transcrição de uma sequência de ácido nucleico ligada operativamente à sequência regulatória de transcrição. Uma sequência regulatória de transcrição como aqui definida irá assim compreender todos os elementos de sequência necessários para a iniciação de uma transcrição (elemento de promotor) para manter e para regular a transcrição, incluindo, por exemplo, atenuadores, ou melhoradores, mas também silenciadores. Apesar de

principalmente as seqüências regulatórias de transcrição a montante (5') de uma seqüência de codificação serem referidas, as seqüências regulatórias encontrada a jusante (3') de uma seqüência de codificação são também englobadas por esta definição.

5 Como usado aqui, o termo "promotor" refere-se a um fragmento de ácido nucleico que funciona para controlar a transcrição de um ou mais genes, localizados a montante (5') com relação à direção de transcrição do sítio de iniciação de transcrição do gene (o início de transcrição é referido aqui como posição +1 da seqüência e quaisquer

10 nucleotídeos a montante relativos aqui são referidos como usando números negativos) e é estruturalmente identificado pela presença de um sítio de ligação para a RNA polimerase dependente de RNA, sítios de iniciação de transcrição e quaisquer outros domínios de DNA (seqüências cis atuantes) incluindo, mas não limitados a sítios de ligação de fator de transcrição, sítios

15 de ligação de proteína repressor e ativador, e quaisquer outras seqüências de nucleotídeos bem conhecidas do versado na técnica como atuando diretamente ou indiretamente para regular a quantidade de transcrição do promotor. Os exemplos de seqüências cis-atuantes eucarióticas a montante do início de transcrição (+ 1) incluem a caixa TATA (comumente em

20 aproximadamente posição -20 a -30 do início de transcrição), a caixa CAAT (comumente em aproximadamente posição -75 com relação ao início de transcrição, elementos 5' melhorador ou silenciador etc. Um promotor "constitutivo" é um promotor que é ativo na maior parte dos tecidos (ou órgãos) sob a maior partes das condições fisiológicas ou de desenvolvimento.

25 Mais preferivelmente, um promotor constitutivo é ativo sob essencialmente todas as condições fisiológicas ou de desenvolvimento em todos os órgãos principais, como pelo menos as folhas, caules, raízes, sementes, frutos e flores. Mais preferivelmente, o promotor é ativo em todos os órgãos sob a maior parte (preferivelmente todas) as condições fisiológicas e

desenvolvimento.

Um promotor "indutível" é um promotor que é fisiologicamente (por exemplo por aplicação externa de alguns compostos) ou regulado no desenvolvimento. Um promotor "específico de tecido" é somente
5 ativo em tipos específicos de tecidos ou células. A atividade do promotor pode assim ser descrita por referência às circunstâncias sob as quais o promotor confere transcrição da sequência de ácido nucleico ligada operativamente a jusante (3') do promotor. Um "promotor que tem atividade constitutiva" ou que é "constitutivo" em uma planta ou célula de planta refere-
10 se, assim, a uma sequência de ácido nucleico que confere transcrição na planta ou células de planta na maior parte dos tecidos (ou órgãos) sob a maior parte das condições fisiológicas e de desenvolvimento. Um promotor que é "insensível a um ou mais estresses bióticos e/ou abióticos" ou cuja Atividade "não é reduzida quando exposto a uma ou mais condições de estresses
15 bióticos e/ou abióticos" refere-se a uma sequência de ácido nucleico tendo atividade de promotor sob condições fisiológicas e de desenvolvimento normais, e assim a atividade não é, ou não é de modo significativa, reduzida quantitativamente quando o estresse biótico e/ou abiótico é exercido sobre o organismo (por exemplo planta) ou células ou tecidos ou órgãos
20 compreendendo o promotor.

"Estresse" refere-se a condições ou pressões de origem física, química ou biológica atuando em uma planta ou células de planta que podem resultar em perda de rendimento e/ou perda de qualidade de uma planta, mas que não é letal para a planta.

25 "Condições de não estresse" refere-se aqui a condições sob as quais a fisiologia e desenvolvimento são normais ou ótimos.

"Estresse biótico" refere-se a estresse causado por agente biótico (vivo) como fungos, vírus, organismos como micoplasma, insetos, bactérias, nematódeos, etc (isto é, especialmente pragas de plantas e

patógenos).

"Estresse abiótico" refere-se a estresse causado por agentes abióticos (não vivos) como estresse de temperatura (frio/ congelamento, calor), salinidade (sal), vento, metais, extensão do dia (fotoperíodo), estresse de água (como uma disponibilidade de água muito pequena ou muito grande, isto é, seca, desidratação, inundação em água), etc), feridas, radiação, etc.

Como usado aqui, o termo "ligado operativamente" refere-se a uma ligação de elementos de polinucleotídeos em uma relação funcional. Um ácido nucleico é "ligado operativamente" quando ele é colocado em uma relação funcional com outra sequência de ácido nucleico. Por exemplo, um promotor, ou uma sequência regulatória da transcrição, é ligado operativamente a uma sequência de codificação se ele afetar a transcrição da sequência de codificação. Ligado operativamente significa que as sequências de DNA sendo ligadas são tipicamente contíguas e, onde necessário para unir duas regiões de codificação de proteína, contíguas e em matriz de leitura de modo a produzir uma "proteína quimérica". Uma "proteína quimérica" ou "proteína híbrida" é uma proteína composta de vários "domínios" de proteína (ou motivos) que não é encontrado como tal na natureza mas que são unidos para formar uma proteína funcional, que mostra a funcionalidade dos domínios unidos (por exemplo um domínio de ligação de DNA ou uma pressão do domínio de função levando a uma função negativa dominante). Uma proteína quimérica também pode ser uma proteína de fusão de duas ou mais proteínas de ocorrência na natureza. O termo "domínio" como usado aqui significa qualquer parte(s) ou domínio(s) da proteína com uma estrutura específica ou função que pode ser transferida para outra proteína para prover uma nova proteína híbrida com, pelo menos, a característica funcional do domínio.

O termo "peptídeo alvo" refere-se a sequências de aminoácido que marcam uma proteína em organelas intracelulares como plastídeos,

preferivelmente cloroplastos, mitocôndria, ou para o espaço extracelular (peptídeo de sinal de secreção). Uma seqüência de ácido nucleico codificando um peptídeo alvo pode ser fusionada (em matriz) para a seqüência de ácido nucleico codificando a extremidade amino terminal (extremidade N-terminal) da proteína.

Uma "construção de ácido nucleico" ou "vetor" é aqui entendido para significar uma molécula de ácido nucleico feita pelo homem resultando do uso de tecnologia de DNA recombinante e que é usada para liberar DNA exógeno em uma célula hospedeira. A estrutura dorsal do vetor pode por exemplo ser um vetor binário ou superbinário (ver por exemplo US 5 591 616, US 2002138879 e WO 95/06722), um vetor co-integrado ou um vetor de T-DNA, como bem conhecido na arte e como ainda descrito aqui, em que um gene quimérico é integrado ou, se uma seqüência / promotor regulatória de transcrição apropriada já estiver presente, somente uma seqüência de ácido nucleico desejada (por exemplo uma seqüência de codificando, uma seqüência anti-sentido ou de repetição invertida) é integrada a jusante da seqüência/ promotor regulatória de transcrição. Os vetores geralmente compreendem outros elementos genéticos para facilitar seu uso na clonagem molecular, como, por exemplo, marcadores selecionáveis, sítios de clonagem múltipla, e outros (ver abaixo).

Uma "célula hospedeira" ou uma "célula hospedeira recombinante" ou "célula transformada" são termos com referência a uma nova célula individual (ou organismo), surgindo como um resultado da introdução na referida célula de pelo menos uma molécula de ácido nucleico, especialmente compreendendo um gene quimérico compreendendo uma proteína desejada ou uma seqüência de ácido nucleico que quando da transcrição dá um RNA anti-sentido ou um RNA de repetição invertido (ou RNA grampo de cabelo) para o silenciamento de um gene alvo / família de genes. A célula hospedeira é preferivelmente uma célula de planta, mas

também pode ser uma célula bacteriana, uma célula fúngica (incluindo uma célula de levedura) etc. A célula hospedeira pode conter a construção de ácido nucleico como uma molécula de replicação extra-cromossomicamente (episomal) ou, mais preferivelmente, compreende o gene quimérico integrado no genoma nuclear ou plastídeo da célula hospedeira.

O termo "marcador selecionável" é um termo bem conhecido do versado na técnica e é usado aqui para descrever qualquer entidade genética que, quando expressada, pode ser usada para selecionar uma célula ou células contendo o marcador selecionável. Os produtos de gene de marcador selecionável conferem, por exemplo, resistência a antibiótico, ou mais preferivelmente, resistência a herbicida ou outro traço selecionável como um traço fenotípico (por exemplo uma mudança na pigmentação) ou uma exigência nutricional. O termo "repórter" é principalmente usado para fazer referência a marcadores visíveis, como proteína fluorescente verde (GFP), eGFP, luciferase, GUS e outros.

O termo "ortólogo" de um gene ou proteína refere-se aqui ao gene homólogo ou proteína encontrada em outra espécie, que tem a mesma função que o gene ou proteína, mas (geralmente) diverge em sequência do ponto de tempo quando as espécies abrigando os genes divergem (isto é, os genes evoluídos de um ancestral comum por especialização). Os ortólogos de um gene de uma espécie de planta podem ser assim identificados em outras espécies de plantas com base tanto em comparações de sequência (por exemplo com base em identidade de sequência percentual sobre a sequência completa ou sobre domínios específicos) como em análise funcional.

Os termos "homólogo" e "heterólogo" referem-se à relação entre uma sequência de ácido nucleico ou aminoácido e sua célula hospedeira ou organismo, especialmente no contexto de organismos transgênicos. Uma sequência homóloga é assim naturalmente encontrada na espécie hospedeira (por exemplo uma planta de tomate transformada com um gene de tomate),

enquanto uma sequência heteróloga não é naturalmente encontrada na célula hospedeira (por exemplo uma planta de tomate transformada com uma sequência de plantas de batata). Dependendo do contexto, o termo "homólogo" ou "homólogo" pode alternativamente fazer referência a

5 sequências que são descendentes de uma sequência ancestral comum (por exemplo eles podem ser ortólogos).

"Condições de hibridização estridentes" podem ser usadas para identificar sequências de nucleotídeos, que são substancialmente idênticas a uma dada sequência de nucleotídeos. A estridência das condições

10 de hibridização são dependentes de sequência e serão diferentes em diferentes circunstâncias. Geralmente, as condições estridentes são selecionadas para serem cerca de 5 °C menores do que o ponto de fusão térmica (T_m) para as sequências específicas em uma resistência iônica definida e pH. A T_m é a estrutura (sob a resistência iônica e pH definidos) em que 50% da sequência

15 alvo hibridiza em uma sonda perfeitamente correspondida. Tipicamente, condições estridentes serão selecionadas em que a concentração de sal é cerca de 0,02 molar a pH 7 e a temperatura é pelo menos 60 °C. A diminuição da concentração de sal e/ou aumenta da temperatura aumenta a estridência. As condições estridentes para hibridizações RNA - DNA (Northern blots

20 usando uma sonda de por exemplo 100 nt) são por exemplo as que incluem pelo menos uma lavagem em 0.2X SSC a 63°C durante 20 min, ou condições equivalentes. As condições estridentes para a hibridização for DNA-DNA (Southern blots usando uma sonda de por exemplo 100nt) são por exemplo as que incluem pelo menos uma lavagem (geralmente 2) em 0.2X SSC a uma

25 temperatura de pelo menos 50°C, geralmente cerca de 55°C, durante 20 min, ou condições equivalentes. Ver também Sambrook *et al.* (1989) e Sambrook e Russell (2001).

"Identidade de sequência" e "similaridade de sequência" podem ser determinadas por alinhamento de duas sequências de peptídeo ou

duas de nucleotídeo usando algoritmos de alinhamento global ou local, dependendo do comprimento das duas seqüências. As seqüências de comprimentos similares são preferivelmente alinhadas usando algoritmos de alinhamento globais (por exemplo Needleman Wunsch) que alinha as

5 seqüências otimamente sobre o comprimento completo, enquanto seqüências de comprimento substancialmente diferentes são preferivelmente alinhadas usando um algoritmo de alinhamento local (por exemplo Smith Waterman). Seqüências podem ser então referidas como "substancialmente idênticas" ou "essencialmente similares" quando elas (quando otimamente alinhadas por

10 exemplo os programas GAP ou BESTFIT usando parâmetros de *default*) compartilham pelo menos uma certa porcentagem mínima de identidade de seqüência (como definido abaixo). GAP usa o algoritmo de alinhamento global de Needleman e Wunsch para alinhar duas seqüências sobre seu comprimento completo (comprimento completo), maximizando o número de

15 correspondências e minimizando o número de espaços. Um alinhamento global é apropriadamente usado para determinar a identidade de seqüência quando as duas seqüências tem comprimentos similares. Geralmente, os parâmetros de default GAP são usados, com uma penalidade de criação de espaço = 50 (nucleotídeos) / 8 (proteínas) e penalidade de extensão de espaço

20 = 3 (nucleotídeos) / 2 (proteínas). Para nucleotídeos, a matriz de escore de *default* usada é nwsgapdna e para proteínas a matriz de escore de *default* é Blosum62 (Henikoff & Henikoff, 1992, PNAS 89, 915-919). Os alinhamentos de seqüência e escores para identidade de seqüência percentual podem ser determinados usando programas de computador, como GCG Wisconsin

25 Package, Versão 10.3, disponível de Accelrys Inc., 9685 Scranton Road, San Diego, CA 92121-3752 USA, ou usando o software de fonte aberta, como o programa "needle" (usando o algoritmo de Needleman Wunsch global ou "water" (usando o algoritmo local de Smith Waterman) em EmbossWIN Versão 2.10.0, usando os mesmos parâmetros como para GAP acima, ou

usando as configurações de default (tanto para 'needle' como para 'water' e ambos para alinhamentos de proteína e DNA, a penalidade de abertura Gap de *default* é 10.0 e a penalidade de extensão de espaço de *default* é 0.5; as matrizes de escore de *default* são Blossum62 para proteínas e DNABFull para DNA). Quando seqüências tem comprimentos globais substancialmente diferentes, os alinhamentos locais como usando o algoritmo de Smith Waterman, são preferidos. Alternativamente, a similaridade ou identidade percentual podem ser determinadas por busca em bases de dados públicas, usando algoritmos como FASTA, BLAST, etc.

Neste documento e em suas reivindicações, o verbo "compreender" e suas conjugações é usado em seu sentido não limitativo para significar que itens após a palavra são incluídos, mas itens não especificamente mencionados não são excluídos. Além disso, referência a um elemento pelo artigo indefinido "um" ou "uma" não exclui a possibilidade de que mais de um do elemento está presente, salvo se o contexto claramente requerer que seja um e somente um dos elementos. O artigo "indefinido "um" ou "uma" assim geralmente significa "pelo menos um". É ainda entendido que, quando fazendo referência a "seqüências" aqui, geralmente as moléculas físicas reais com um a certa seqüência de subunidades (por exemplo aminoácidos) são referidas.

Sempre que a referência a uma "planta" ou "plantas" (ou uma pluralidade de plantas) de acordo com a invenção é feita, é entendido que também as partes da planta (células, tecidos ou órgãos, sementes, partes cortadas ou coletadas, folhas, mudinhas, flores, pólen, frutas, caules, raízes, calos, protoplastos, etc), as propagações clonais ou progênie das plantas que retém as características distintas dos parentais (por exemplo a presença de um trans-gene), como semente obtida por endogamia ou cruzamento, por exemplo semente híbrida (obtida por cruzamento de duas linhagens parentais congênicas), plantas híbridas e partes de plantas derivadas das mesmas são

englobadas, salvo especificado em contrário.

DESCRIÇÃO DETALHADA

Quando testando plantas transgênicas compreendendo o promotor CaMV 35S (promotor 35S único, descrito por Franck *et al.*, 1980, Cell 21, 285-294) em tentativas no campo na Espanha, verificou-se que a atividade do promotor 35S foi influenciada por temperaturas elevadas de verão. Em folhas mais velhas, uma perda total de expressão ocorreu. Os presentes inventores, assim, iniciaram um programa para isolar promotores de plantas que tem atividade constitutiva em células de plantas e órgãos, e que são essencialmente insensíveis a um ou mais estresses bióticos e/ou abióticos, e este promotor é desejado para expressão controlada de seqüências de ácido nucleico em plantas transgênicas, especialmente em plantas que podem ser expostas a uma ou mais condições de estresse durante seu desenvolvimento ou na maturidade. Usando cDNA-AFLP, genes de promotores fortes e constitutivos foram isolados de plantas de tomate, que reterem perfis de transcrição constitutivos e fortes sob condições de estresse, como estresse de seca, estresse de frio, estresse de calor, estresse de patógeno (infecção por CMV), variação da extensão do dia, radiação (por exemplo estresse induzido por UV), estresse de água (super aguada e sub aguada), etc (ver exemplos). Ao isolar os promotores correspondentes, uma classe de promotores foi encontrada que conferiu (ou é capaz de conferir), expressão constitutiva forte em plantas. A resistência do promotor foi pelo menos equivalente e em algumas espécies de hospedeiro (tomate) significativamente maior do que a do promotor CaMV 35S. Além disso, os promotores não mostram atividade reduzida (isto é retém atividade forte e constitutiva) quando as plantas ou partes das plantas compreendendo os mesmos são submetidas a várias condições de estresses bióticos e/ou abióticos). Estes promotores são aqui referidos como os "promotores de AA6".

Seqüências de ácido nucleico , genes quiméricos e vetores

Em uma forma de realização, as seqüências de ácido nucleico isolados (preferivelmente seqüências de DNA genômico ou sintético) tendo atividade de promotor em células de plantas, são providas que mostram uma atividade transcripcional constitutiva forte em plantas e tecidos de plantas ou órgãos. Os promotores são ativos em preferivelmente todos os órgãos de plantas, mas pelo menos nos órgãos principais e a resistência do promotor é pelo menos essencialmente igual à de 35S (ou mais forte) (o promotor CaMV 35S descrito por Franck *et al.*, supra). Preferivelmente, a atividade do promotor das seqüências de ácido nucleico não é reduzida, ou pelo menos não reduzida de modo significativo, quando a planta transgênica , ou tecido de planta ou órgão, é submetida a um ou mais estresses bióticos e/ou abióticos. Uma redução significativa neste aspecto refere-se a uma redução estatisticamente significativa (quantitativa) de atividade de promotor por 1% ou mais (por exemplo 2%, 3%, 5%, 10%, etc., até 100%) comparado com a atividade em alguns tecidos ou órgãos sob condições de não estresse. Assim, preferivelmente, os promotores permanecem fortes e constitutivos (preferivelmente em todos os órgãos, especialmente pelo menos os órgãos de planta principais) sob condições de estresse.

Em uma forma de realização, um promotor constitutivo de AA6 é provido compreendendo SEQ ID NO: 1 (promotor “3kb”) ou SEQ ID NO: 2 (promotor “5kb”), ou qualquer uma dentre SEQ ID NO: 6-9 (SEQ ID NO: 6 e 8 são promotores “3kb” e SEQ ID NO: 7 e 9 são promotores “5kb” com algumas ambigüidades de nucleotídeos ou seqüências levemente diferentes comparadas com SEQ ID NO: 1 e 2), ou as seqüências de promotor clonadas em pKG8135 (CBS120175) ou pKG8137 (CBS120176), ou fragmentos ativos (funcionais) de qualquer um destes que tem atividade de promotor pelo menos em plantas, como fragmentos de pelo menos 200, 300, 400, 500, 600, 800, 900, 1000, 1200, 1500, 2000, 2500, 2800, 2900, 3000, 3500, 4000, 4500 ou mais nucleotídeos consecutivos de SEQ ID NO: 1 ou 2,

SEQ ID NO: 6-9, ou promotores dentro de pKG8135 (CBS120175) ou pKG8137 (CBS120176). "Fragmentos ativos" ou "fragmentos funcionais", ou "fragmentos tendo atividade de promotor" referem-se a fragmentos de ácido nucleico que são capazes de conferir uma transcrição constitutiva em células de plantas, órgãos e plantas, preferivelmente e, pelo menos os mesmos tecidos e órgãos como SEQ ID NO: 1 e 2, SEQ ID NO: 6-9, ou promotores dentro de pKG8135 (CBS120175) ou pKG8137 (CBS120176). Isto pode ser testados como descrito abaixo, por transformação de uma célula de planta com tal fragmento, preferivelmente ligado operativamente a um gene repórter, e testando a atividade do promotor qualitativamente (transcrição espaço-temporal) e/ou quantitativamente.

Quando fazendo referência aqui a SEQ ID NO: 1 entende-se que referência é também feita a qualquer uma de SEQ ID NO: 6 ou 8, ou a sequência presente em pKG8135 (CBS120175) e que a frase pode ser lida para fazer referência a sequências de promotor. Do mesmo modo, quando fazendo referência aqui a SEQ ID NO: 2 entende-se que referência é também feita a qualquer uma dentre SEQ ID NO: 7 ou 9, ou a sequência presente em pKG8137 (CBS120175) e que a frase pode ser lida para fazer referência a qualquer uma destas sequências de promotor.

Em uma forma de realização, a resistência dos fragmentos do promotor é quantitativamente essencialmente idêntica a ou maior do que, a da SEQ ID NO: 1 e/ou 2 (e assim também a de 35S). Como SEQ ID NO: 1 e 2 são sequências diretamente a montante do códon de iniciação de tradução (AUG em mRNA ou ATG em DNA), fragmentos ativos são preferivelmente gerados pelas deleções na extremidade 5' de SEQ ID NO: 1 e/ou 2. Os fragmentos, assim, preferivelmente compreendem pelo menos 200, 300, 400, 500, etc. nucleotídeos (como acima) da região 3' de SEQ ID NO: 1 ou 2. Obviamente, os fragmentos de DNA podem ser gerados em vários modos, por exemplo usando síntese de DNA *de novo* DNA, ou enzimas de restrição, ou

nucleases terminais, etc.

No entanto, a remoção de alguns elementos *cis* atuantes, como seqüências de melhorador) pode resultar em menor atividade do promotor. Assim, em uma forma de realização diferente, a resistência dos fragmentos do
5 promotor é quantitativamente menor do que a de SEQ ID NO: 1 e/ou 2 (e assim também de 35S). Para algumas aplicações, promotores com resistência reduzida são preferidos. Um versado pode facilmente determinar a atividade de promotores de comprimento completo e quaisquer fragmentos de promotor usando métodos bem conhecidos na técnica, e podem comparar a resistência e
10 especificidade de tecido com de 35S ou para os promotores providos aqui. Por exemplo, Medberry *et al.* (Cell Plant 1992, Vol. 4: 185-192) descrevem métodos para comparar a resistência do promotor usando testes de expressão transiente.

Os promotores de SEQ ID NO: 1 e 2, e fragmentos funcionais
15 dos mesmos, são preferivelmente insensíveis a pelo menos um (mas preferivelmente mais, o mais preferivelmente qualquer) estresse biótico e/ou abiótico ao qual a planta ou célula de planta(s), tecidos ou órgãos compreendendo o promotor podem ser expostos (ver abaixo). Assim, a atividade permanece constitutiva e forte durante a exposição a estas condições
20 de estresse.

Também são providas “variantes” dos promotores de AA6 acima e fragmentos funcionais destas variantes. Estas variantes incluem seqüências de ácido nucleico essencialmente similares a SEQ ID NO: 1 e/ou
25 SEQ ID NO: 2 (e fragmentos funcionais destas seqüências variantes, como descrito acima), e que tem atividade de promotor constitutivo, isto é, que são também capazes de prover uma transcrição constitutiva em plantas, células de plantas, tecidos ou órgãos, mais preferivelmente em pelo menos os mesmos tecidos e órgãos como SEQ ID NO: 1 e 2. Seqüências que são “essencialmente similares” a SEQ ID NO: 1 e/ou 2 são seqüências de ácido

nucleico compreendendo pelo menos cerca de 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 96, 97, 98, 99% ou mais de identidade de sequência de ácido nucleico para SEQ ID NO: 1 e/ou para SEQ ID NO: 2 (sobre o comprimento completo), usando o alinhamento aos pares de Needleman e Wunsch (programa “GAP” em GCG ou “needle” em EmbossWin, Versão 2.10.0) com penalidade de criação de espaço = 50 e penalidade de extensão de espaço = 3) e que tem atividade de promotor em plantas ou células de plantas. Em uma forma de realização preferida, a atividade das variantes (e fragmentos funcionais) é forte em todos os tecidos e órgãos, isto é, quantitativamente pelo menos tão forte (ou mais forte) do que SEQ ID NO: 1 ou 2. Em outra forma de realização, a atividade destas variantes (e fragmentos funcionais das mesmas) é insensível a um ou mais estresses bióticos e/ou abióticos, isto é, permanece forte e constitutiva.

É evidente que muitos métodos podem ser usados para identificar, sintetizar ou isolar variantes ou fragmentos funcionais das sequências de ácido nucleico aqui providas, como hibridização de ácido nucleico. A tecnologia PCR, análise *in silico* e síntese de ácido nucleico, e outros. Por exemplo, a hibridização de ácido nucleico pode ser usada para identificar sequências de DNA em outras espécies de plantas ou variedades que hibridizam para SEQ ID NO: 1 ou 2, ou para fragmentos destas, sob condições de hibridização estridentes ou moderadamente estridentes. Alternativamente, as bases de dados de sequências podem ser triadas *in silico* para sequências variante usando algoritmos conhecidos, como BLAST, FASTA, etc. Deste modo, é possível isolar sequências variantes de outras espécies de plantas ou outras variedades de tomate, ou de outros organismos no todo. São especialmente incluídos aqui os promotores de outros alelos do mesmo gene (gene de cisteína sintase citossólica e ortólogos do mesmo) encontrados em outras variedades de tomate ou em outras espécies de plantas, especialmente espécies do gênero *Solanum* como será descrito abaixo. Por

exemplo, as bibliotecas de cDNA podem ser construídas de uma ou mais espécies de plantas, uma ou mais variedades, ou tecidos diferentes de uma espécie ou variedade. As bibliotecas de cDNA podem ser triadas para cDNAs de cisteína sintase (usando por exemplo sondas ou iniciadores derivados de SEQ ID NO: 4 (cDNA de cisteína sintase de tomate) ou fragmentos ou variantes dos mesmos). Igualmente, os métodos de exibição diferencial (como cDNA-AFLP) podem ser usados para identificar transcritos de cisteína sintase, como descrito nos Exemplos. Métodos como TAIL-PCR (Liu et al. 1995, Genomics 25(3):674-81; Liu et al. 2005, Methods Mol Biol. 286:341-8), Linker-PCR, ou Inverse PCR (IPCR) podem ser usados para isolar a região regulatória de transcrição a montante do gene.

Se uma sequência de ácido nucleico (ou fragmento de variante) tem atividade de promotor constitutivo, isto é, é capaz de conferir transcrição constitutiva em todos os órgãos, se a atividade é "forte" e se a atividade biológica da sequência de ácido nucleico é insensível a pelo menos um estresse biótico e/ou abiótico (mas preferivelmente mais, o mais preferivelmente qualquer) ao qual a célula, tecido, órgão ou organismos transgênicos (especialmente planta ou célula de planta) pode ser exposto, pode ser determinado usando vários métodos. Geralmente, pode-se distinguir métodos qualitativos e métodos quantitativos. Os métodos qualitativos (como coloração GUS histológica) são usados para determinar a atividade espaço-temporal do promotor (seja o promotor ativo ou não em um determinado tecido ou órgão, ou sob algumas condições ambientais/ de desenvolvimento?), enquanto os métodos quantitativos (como testes GUS fluorométricos) também quantificam o nível de atividade, comparado com controles. Os controles apropriados são, por exemplo, plantas transformadas com vetores vazios (controle negativo) ou transformadas com construções compreendendo outros promotores, com CaMV 35S, ou plantas não transgênicas.

Para testar e opcionalmente quantificar a atividade relativa ou

absoluta, uma molécula de ácido nucleico clonada ou sintética, como SEQ ID NO: 1, 2 ou variantes da mesma, ou fragmentos de qualquer uma destas, pode ser ligada operativamente a uma sequência de ácido nucleico conhecida (por exemplo um gene repórter, como *gusA*, ou qualquer gene codificando uma proteína específica) e pode ser usado para transformar uma célula de planta usando métodos conhecidos. Em algumas formas de realização, a célula não precisa ser transformada em um modo estável, isto é, testes de expressão transiente podem ser usados (por exemplo, transfecção de protoplastos ou Agroinfiltração) para determinar se o promotor é ativo nas células, tecidos ou órgãos e em que grau o promotor está dirigindo a transcrição. A atividade do promotor pode, por exemplo, ser testada (e opcionalmente quantificada) por detecção do nível de transcritos de RNA da sequência de ácido nucleico a jusante. Isto pode ser feito usando métodos quantitativos, como por exemplo RT-PCR quantitativa, ou outros métodos à base de PCR, e outros.

Alternativamente, a proteína repórter ou a atividade da proteína repórter pode ser testada e quantificada. Por exemplo, se o gene repórter for gene *gus*, um teste GUS fluorométrico pode ser usado, como descrito nos exemplos. Deste modo, os níveis da atividade do promotor quantitativa de plantas ou células de planta transformadas mantidas sob condições fisiológicas normais (sem estresse) podem ser comparados com níveis de plantas ou células de plantas que são expostas a um ou mais estresses bióticos ou abióticos. Também, níveis de atividade relativa ou absoluta podem ser comparados com promotores de controle constitutivos, como promotor 35S, promotor duplo-25S, ou SEQ ID NO: 1 e/ou 2. Entende-se que preferivelmente níveis de atividade de promotor médio são determinados e comparados usando métodos estatísticos.

Assim, em quais tecidos ou órgãos um promotor de acordo com a invenção é ativo em um determinado tempo (atividade espaço-temporal) pode, por exemplo, ser testados por transformação de plantas ou

células de plantas com uma construção de promotor - gene repórter e analisando vários tecidos durante vários estágios de desenvolvimento para o transcrito de RNA ou proteína repórter (ou sua atividade). Um teste simples emprega, por exemplo, coloração GUS histoquímica, assim a avaliação visual de cor azul indica atividade em vários tecidos e em vários estágios de desenvolvimento.

Como já mencionado, prefere-se que a atividade do promotor seja constitutiva e preferivelmente também forte em plantas e células de plantas, especialmente nas espécies hospedeiras ou variedade em que a sequência é introduzida. A atividade constitutiva significa que o transcrito de qualquer sequência de ácido nucleico ligado operativamente ao promotor é preferivelmente produzido na maior parte dos tecidos (ou órgãos) sob a maior parte das condições fisiológicas (normal, sem estresse) e de desenvolvimento. Em uma forma de realização, os promotores de acordo com a invenção são pelo menos ativos nos tecidos e órgãos testados nos exemplos, como folhas (jovens e velhas), raízes, flores, sementes, caules (caule principal), frutas (por exemplo frutas imaturas e maduras), sementes germinadas, etc.

Preferivelmente, os promotores de acordo com a invenção provêm atividade forte, constitutiva, em todas as espécies de plantas, tanto espécies dicotiledôneas como espécies monocotiledôneas. Por exemplo, verificou-se que SEQ ID NO: 1 e SEQ ID NO: 2 provêm uma expressão forte, constitutiva, em várias espécies de plantas, como tomate, tabaco, *Brassica*, melão e alface (ver exemplos).

A resistência (atividade quantitativa) dos promotores de AA6 de acordo com a invenção (incluindo fragmentos ou variantes) em termos de sua capacidade de dirigir expressão das sequências de ácido nucleico ligadas a jusante (3') pode ser determinada quantitativamente usando vários métodos bem conhecidos. Por exemplo, a quantidade do transcrito (mRNA) pode ser quantificada usando RT-PCR quantitativo ou northern blot. Preferivelmente,

a resistência do promotor é pelo menos essencialmente igual à de CaMV 35S (Franck et al, supra) sob condições normais (sem estresse). "Forte" significa, assim, que a resistência do promotor é preferivelmente pelo menos cerca de idêntica, mas mais preferivelmente mais forte do que a de 35S sob condições normais, sem estresse. Mais preferivelmente, a atividade do promotor quantitativa média nos vários tecidos e órgãos é pelo menos equivalente à atividade do promotores de CaMV 35S, ou é pelo menos 5%, 10%, 20%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 70%, 75%, ou mais, maior do que a atividade biológica média do promotor CaMV 35S. Entende-se que o mesmo número de cópias e nível de zigosidade de transformantes deve ser comparado, por exemplo hemizigóticos ou homozigóticos para o transgene. Preferivelmente, transformantes de cópia única são identificados e comparados.

A expressão dos promotores de AA6 permanece constitutiva e forte sob condições de estresse, pelo menos sob um ou mais estresses selecionados dentre os descritos em algumas partes aqui.

Assim, a resistência dos promotores de acordo com a invenção preferivelmente permanece essencialmente inalterada, ou é pelo menos não reduzida (ou não significativamente reduzida) quando os tecidos de plantas ou órgãos ou plantas compreendendo o promotor são expostos a condições de estresse, selecionadas pelo menos dentre um ou preferivelmente vários dentre: estresse de seca, estresse de calor, estresse de água (tanto muito como pouca água), estresse de patógeno (por exemplo infecção por vírus como CMV, infecção fúngica, infecção bacteriana, etc), estresse de pragas (por exemplo alimentação de insetos), feridas, estresse de sal, estresse de radiação etc. Novamente, testes quantitativos podem ser usados para a sua determinação. Por exemplo, plantas recombinantes compreendendo o promotor podem ser transferidas de um meio de temperatura normal para um meio quente (como cerca de 27 °C até cerca de 50 °C), e a atividade do promotor em vários

tecidos pode ser comparada com a atividade nos mesmos tecidos sob as condições de temperatura normal e quente.

Variantes dos promotores de AA6 acima também incluem qualquer promotor de ácido nucleico isolado de um gene de cisteína sintase de planta, especialmente de um gene de cisteína sintase citossólica (isto é, o promotor de um gene de planta codificando uma enzima cisteína sintase, especialmente uma isoforma citossólica), que mostra atividade forte e constitutiva em plantas e é preferivelmente essencialmente insensível a um ou mais estresses bióticos e/ou abióticos. Como acima mencionado, é provável que outros genes de cisteína sintase de planta tenham promotores constitutivos e fortes, que retêm a expressão constitutiva sob uma ou mais condições de estresse. Os genes de cisteína sintase de planta são genes que codificam a enzima cisteína sintase, também referida como O-acetil-L-serina[tiol]-liase (EC 4.2.99.8). Vários genes de cisteína sintase foram clonados (ver por exemplo TC162833, referido nos exemplos), mas a utilidade de seus promotores para dirigir expressão forte, constitutiva, de seqüências homólogas ou heterólogas em plantas transgênicas não foi descrita. Os promotores destes genes de cisteína sintase conhecidos ou ainda desconhecidos podem ser assim isolados e triados para sua atividade. Por exemplo, cDNA-AFLP, outros métodos à base de PCR ou hibridização Northern podem ser usados para isolar ou identificar outros genes de cisteína sintase, especialmente os genes que produziram endogenamente mRNAs de cisteína sintase constitutivamente e também sob uma ou mais condições de estresse. Os genes de cisteína sintase com o padrão de expressão desejado são então selecionados e seu promotor pode ser clonado usando métodos conhecidos. Em uma forma de realização preferida, o promotor é obtido de um gene de cisteína sintase de uma planta pertencendo à família *Solanaceae*, como espécies do gênero *Solanum* (incluindo a espécie reclassificada *Lycopersicon*), *Nicotiana*, *Capsicum*, *Petunia*, *Coffea*, etc.

O promotor de acordo com a invenção é preferivelmente não o promotor de uma planta de ervas daninhas, como *Arabidopsis thaliana*, e é preferivelmente não o promotor do gene OASA1 de *Arabidopsis thaliana* como descrito em Gutierrez-Alcala et al. (J. of Exp. Botany 56, p24872494),
5 ou um fragmento do mesmo.

Além dos genes de cisteína sintase providos em SEQ ID NO: 4 (codificando a enzima cisteína sintase de SEQ ID NO: 5), outros genes de cisteína sintase podem ser identificados e seus promotores isolados. Vários métodos podem ser usados, como descrito acima.

10 Assim, em uma forma de realização da invenção, o promotor de um gene de cisteína sintase de planta é provido, o qual tem atividade constitutiva em plantas e/ou célula de plantas e assim a atividade não é reduzida, ou reduzida de modo significativo, quando a planta ou célula de planta, tecido ou órgão é exposto a um ou mais estresses bióticos e/ou
15 abióticos, como os descritos.

Estes promotores compreendem os seguintes: (a) qualquer promotor que é obtível a partir de uma planta e que dirigida a expressão de um gene de cisteína sintase de planta, especialmente o promotor de qualquer gene codificando uma proteína de SEQ ID NO: 5; (b) promotores dos genes
20 de planta que codificam as enzimas de cisteína sintase tendo pelo menos 30, 40, 50, 60, 70, 80, 90, 95, 99%, ou mais, de identidade de aminoácido para a SEQ ID NO: 5 (sobre o comprimento completo); (c) promotores de plantas dirigindo a expressão de seqüências de ácido nucleico, que codificam as enzimas de cisteína sintase, e pelo que as seqüências de ácido nucleico
25 compreendem pelo menos 30, 40, 50, 60, 70, 80, 90, 95, 98, 99%, ou mais de identidade de seqüência de nucleotídeos para SEQ ID NO: 4 (sobre o comprimento completo).

Em outra forma de realização, o uso de um promotor de um gene de cisteína sintase de planta para a sobre- expressão constitutiva de

seqüências de ácido nucleico homólogas ou heterólogas em uma célula ou organismo recombinante, especialmente uma célula de planta ou planta, é provido. Este uso compreende ligar operativamente o promotor a uma seqüência homóloga ou heteróloga de ácido nucleico e transformar uma
5 planta ou célula de planta, como descrito.

Apesar do foco acima ser sobre o uso dos promotores de acordo com a invenção em plantas e célula de plantas, também é uma forma de realização da invenção usar os promotores para a expressão de seqüências de ácido nucleico homólogas ou heterólogas em outras células ou organismos,
10 como em quaisquer células procarióticas ou eucarióticas ou organismos, por exemplo bactérias, fungos (incluindo leveduras, como *Pichia*, *Hansenula*, etc.), células de humanos, mamíferos, ou linhagens de células, etc.

Genes quiméricos e vetores de acordo com a invenção

Em uma forma de realização da invenção, qualquer uma das seqüências de ácido nucleico tendo atividade de promotor, são usadas para
15 fazer os genes quiméricos, e vetores compreendendo os mesmos para a transferência de gene quimérico em uma célula hospedeira e expressão de uma seqüência de ácido nucleico homóloga ou heteróloga ligada operativamente em células hospedeiras, como células, tecidos, órgãos ou
20 organismos completos derivados de célula(s) transformada(s).

As células hospedeiras são preferivelmente células de plantas. Qualquer planta pode ser um hospedeiro apropriado, com plantas monocotiledônea s ou plantas dicotiledônea s, por exemplo milho/milhete (espécies *Zea*, por exemplo *Z. mays*, *Z. diploperennis* (chapule), *Zea luxurians* (teosinte da guatemala), *Zea mays* subsp. *huehuetenangensis*
25 (teosinte San Antonio Huista), *Z. mays* subsp. *mexicana* (teosinte mexicano), *Z. mays* subsp. *parviglumis* (teosinte Balsas), *Z. perennis* (teosinte perene) e *Z. ramosa*, trigo (espécie *Triticum*), cevada (por exemplo *Hordeum vulgare*), aveia (por exemplo *Avena sativa*), sorgo (*Sorghum bicolor*), centeio (*Secale*

cereale), soja (*Glycine* spp, por exemplo *G. max*), algodão (espécie *Gossypium*, por exemplo *G. hirsutum*, *G. barbadense*), *Brassica* spp. (por exemplo *B. napus*, *B. juncea*, *B. oleracea*, *B. rapa*, etc.), girassol (*Helianthus annuus*), tabaco (espécie *Nicotiana*), alfafa (*Medicago sativa*), arroz (espécie *Oryza*, por exemplo grupo de cultivar de *O. sativa indica* ou grupo de cultivar *japonica*), gramíneas de foragem, milheto (espécie *Pennisetum*, por exemplo *P. glaucum*), espécies de árvores, espécies de vegetais, como *Lycopersicon* ssp (recentemente reclassificado como pertencendo ao gênero *Solanum*), por exemplo tomate (*L. esculentum*, sin. *Solanum lycopersicum*) como por exemplo tomato cereja, var. *cerasiforme* ou tomato comum, var. *pimpinellifolium*) ou tomate de árvore (*S. betaceum*, sin. *Cyphomandra betaceae*), batata (*Solanum tuberosum*) e outras espécies *Solanum*, como berinjela (*Solanum melongena*), pepino (*S. muricatum*), cocona (*S. sessiliflorum*) e naranjilla (*S. quitoense*); pimentas (*Capsicum annum*, *Capsicum frutescens*), ervilha (por exemplo *Pisum sativum*), feijão (por exemplo espécies *Phaseolus*), cenoura (*Daucus carota*), espécies *Lactuca* (como *Lactuca sativa*, *Lactuca indica*, *Lactuca perennis*), pepino (*Cucumis sativus*), melão (*Cucumis melo*), zucchini (*Cucurbita pepo*), abóbora (*Cucurbita maxima*, *Cucurbita pepo*, *Cucurbita mixta*), abóbora (*Cucurbita pepo*), melancia (*Citrullus lanatus* sin. *Citrullus vulgaris*), espécies de frutas carnudas (uvas, pêssegos, ameixas, morango manga, melão), espécies ornamentais (por exemplo *Rosa*, *Petunia*, *Chrysanthemum*, lírio, tulipa, espécies *Gerbera*), árvores lenhosas (por exemplo espécies de *Populus*, *Salix*, *Quercus*, *Eucalyptus*), espécies de fibras por exemplo linho (*Linum usitatissimum*) e cânhamo (*Cannabis sativa*). Em uma forma de realização, as espécies de vegetais, especialmente espécies *Solanum* (incluindo espécies *Lycopersicon*) são preferidas.

Assim, por exemplo, as espécies dos seguintes gêneros podem ser transformadas: *Cucurbita*, *Rosa*, *Vitis*, *Juglans*, *Fragaria*, *Lotus*,

Medicago, Onobrychis, Trifolium, Trigonella, Vigna, Citrus, Linum, Geranium, Manihot, Daucus, Arabidopsis, Brassica, Raphanus, Sinapis, Atropa, Capsicum, Datura, Cucumis, Hyoscyamus, Lycopersicon, Solanum, Nicotiana, Malus, Petunia, Digitalis, Majorana, Cichorium, Helianthus,
 5 *Lactuca, Bromus, Citrullus, Asparagus, Antirrhinum, Heterocallis, Nemesis, Pelargonium, Panieum, Pennisetum, Ranunculus, Senecio, Salpiglossis, Browaalia, Glycine, Pisum, Phaseolus, Gossypium, Glycine, Lolium, Festuca, Agrostis.* Uma outra preferência é para cada dentre *Cucurbita, Brassica, Lycopersicon, Solanum, Oryza* e *Zea*. Uma preferência é para cada dentre
 10 *Avena, Medicago, Capsicum, Nicotiana, Lactuca, Pisum, Cucumis, Cucurbita, Brassica, Solanum* (incluindo *Lycopersicon*), *Oryza* e *Zea*.

A construção de genes quiméricos, e vetores para introdução de genes quiméricos no genoma de células hospedeiras, é geralmente bem conhecida na técnica. Para gerar um gene quimérico, a seqüência de promotor de AA6 é ligada operativamente a outra seqüência de ácido nucleico que deve
 15 ser transcrita nas células hospedeiras, usando técnicas de biologia molecular padrões. A seqüência de promotor já pode estar presente em um vetor de modo que a seqüência de ácido nucleico, que deve ser transcrita, é simplesmente inserida no vetor a jusante da seqüência de promotor. O vetor é
 20 então usado para transformar as células hospedeiras e o gene quimérico é preferivelmente inserido no genoma nuclear ou no genoma de plastídeo, mitocondrial ou cloroplasto, de modo que a seqüência de ácido nucleico a jusante é expressada devido à atividade do promotor (por exemplo, Mc Bride *et al.*, 1995 Bio/Technology 13, 362; US 5,693, 507).

25 Um gene quimérico, assim, preferivelmente compreende um promotor de AA6 como descrito acima, ligado operativamente a uma seqüência de ácido nucleico homóloga ou heteróloga, e opcionalmente seguido por uma seqüência não traduzida 3' de ácido nucleico (3'UTR). A seqüência de ácido nucleico homóloga ou heteróloga pode ser uma seqüência

codificando uma proteína ou peptídeo, ou ela pode ser uma sequência que é transcrita em uma molécula de RNA ativa, como um RNA sentido e/ou anti-sentido RNA (dsRNA) apropriado para o silenciamento de um gene ou família de genes na célula hospedeira ou organismo.

5 O gene quimérico compreendendo promotor de AA6 pode ser inserido de modo estável em um modo covalentemente no genoma nuclear de uma célula de planta única, e a célula de planta assim transformada pode ser usada em um modo convencional para produzir uma planta transformada que tem um fenótipo alterado devido à expressão constitutiva do gene quimérico.

10 Neste aspecto, um vetor de T-DNA, compreendendo um promotor de AA6 (ou variante ou fragmento como mencionado acima) ligado operativamente a uma outra sequência de ácido nucleico, em *Agrobacterium tumefaciens* pode ser usado para transformar a célula de planta, e a seguir, uma planta transformada pode ser regenerada da célula de planta transformada usando os procedimentos descritos, por exemplo, em EP 0 116 718, EP 0 270 822, publicação PCT WO 84/02913 e pedido de patente europeia publicado EP 0 242 246 e em Gould *et al.* (1991, Plant Physiol. 95,426-434). A construção de um vetor T-DNA para transformação de planta mediada por *Agrobacterium* é bem conhecida na técnica. O vetor T-DNA
15 pode ser ou um vetor binário, como descrito em EP 0 120 561 e EP 0 120 515, ou um vetor co-integrado que pode integrar no plasmídeo Ti de *Agrobacterium* por recombinação homóloga, como descrito em EP 0 116 718.

20 Os vetores de T-DNA preferidos contém, cada, um promotor de AA6 ligado operativamente à sequência de ácido nucleico a ser transcrita entre as sequências de borda de T-DNA, ou pelo menos localizadas à esquerda da sequência de borda à direita. As sequências de borda são descritas em

Gielen *et al.* (1984, EMBO J 3,835-845). Como evidente, outros tipos de vetores podem ser usados para transformar a célula de planta,

usando procedimentos como transferência de genes direta (como descrito, por exemplo em EP 0 223 247, ou bombardeio de partículas ou micro projéteis, como descrito em US 2005/055740 e WO 2004/092345), transformação mediada por pólen (como descrito, por exemplo em EP 0 270 356 e WO 85/01856), transformação de protoplastos, como, por exemplo, descrito em US 4,684, 611, transformação mediada por vírus de plantas, transformação mediada por lipossomas (como descrito, por exemplo em US 4,536, 475), e outros métodos como os métodos descritos para transformar algumas linhagens de milho (por exemplo, US 6,140, 553; Fromm *et al.*, 1990, Bio/Technology 8, 833-839; Gordon-Kamm *et al.*, 1990, The Cell Planta 2, 603-618) e arroz (Shimamoto *et al.*, 1989, Nature 338, 274-276; Datta *et al.* 1990, Bio/Technology 8, 736-740) e o método para transformar monocótilas, geralmente (WO 92/09696). Para transformação de algodão, ver também WO 00/71733, e para transformação de arroz, ver também os métodos descritos em WO 92/09696, WO 94/00977 e WO 95/06722. Para transformação de sorgo, ver por exemplo Jeoung JM *et al.* 2002, Hereditas 137: 20-8 ou Zhao ZY *et al.* 2000, Plant Mol Biol. 44:789-98). Para transformação de tomate ou tabaco, ver também An G. *et al.*, 1986, Plant Physiol. 81: 301-305; Horsch R.B. *et al.*, 1988, Em: Plant Molecular Biology Manual A5, Dordrecht, Netherlands, Kluwer Academic Publishers. pp 1-9; Koornneef M. *et al.*, 1986, Em: Nevins D.J. e R.A. Jones, eds. Tomato Biotechnology, New York, NY, USA, Alan R. Liss, Inc. pp 169-178). Do mesmo modo, a seleção e regeneração de plantas transformadas de células transformadas é bem conhecida na técnica. Obviamente, para espécies diferentes e mesmo para diferentes variedades e cultivares de uma única espécie, protocolos são especificamente adaptados para a regeneração de transformantes em frequência elevada.

Além da transformação do genoma nuclear, também a transformação do genoma do plastídeo, preferivelmente do genoma do

cloroplasto, está incluída na invenção. Uma vantagem da transformação do genoma do plastídeo é que o risco de espalhamento do(s) transgene(s) pode ser reduzida. A transformação do genoma do plastídeo pode ser realizada como conhecido na arte, ver por exemplo Sidorov VA *et al.* 1999, Plant J.19: 209-216 ou Lutz KA *et al.* 2004, Plant J. 37(6):906-13.

A planta transformada resultante pode ser usada em esquema de criação de plantas convencional para produzir mais plantas transformadas contendo o transgene. Os transformantes de cópia única podem ser selecionados, por exemplo usando análise Southern Blot ou métodos com base em PCR ou o teste Invader® Technology (Third Wave Technologies, Inc.). As células e plantas transformadas podem ser facilmente distintas das não transformadas pela presença do gene quimérico. As seqüências do DNA de planta flanqueando o sítio de inserção do transgene também podem ser seqüenciadas, onde um método de detecção "específico do evento" pode ser desenvolvido, para uso de rotina. Ver por exemplo WO 01/41558, que descreve os kits de detecção de evento de elite (como kits de detecção de PCR) com base por exemplo na seqüência integrada e a seqüência de flanco (genômica).

Em uma forma de realização, a seqüência de ácido nucleico que deve ser transcrita, e opcionalmente traduzida (se for uma seqüência de codificação) é inserida no genoma da planta de modo que a seqüência a ser transcrita está a montante (isto é, 5') de sinais de regulação de transcrição de extremidade 3' apropriados ("extremidade 3'") (isto é, sinais de formação de transcrito e poliadenilação). Os sinais de formação de transcrito e poliadenilação incluem os do gene de nopalina sintase ("3' nos") (Depicker *et al.*, 1982 J. Molec. Appl. Genetics 1, 561-573.), o gene de octopina sintase ("3'ocs") (Gielen *et al.*, 1984, EMBO J 3, 835-845) e o gene de T-DNA 7 ("3' gene 7") (Velten e Schell, 1985, Nucleic Acids Research 13, 6981-6998), que atuam como seqüências de DNA 3' não traduzidas em células de plantas

transformadas, e outros.

Em uma forma de realização preferida, a sequência de extremidade 3' usada é a de um gene de cisteína sintase de planta, preferivelmente do gene de cisteína sintase do qual também o promotor AA6 é obtenível. As seqüências de extremidade 3' providas aqui são SEQ ID NO: 3, ou um fragmento ou variante da mesma. Uma variante de SEQ ID NO: 3 inclui seqüências de ácido nucleico compreendendo pelo menos 50, 60, 65, 70, 75, 80, 85, 90, 95, 98, 99%, ou mais, de identidade de seqüência de ácido nucleico para SEQ ID NO: 3. Fragmentos de SEQ ID NO: 3, ou de variantes de SEQ ID NO: 3, incluem seqüências de ácido nucleico compreendendo pelo menos 50, 100, 150, 200, 250, 300 ou mais nucleotídeos consecutivos de SEQ ID NO: 3 ou de uma variante de SEQ ID NO: 3. Estas seqüências de extremidade 3' também são uma forma de realização como tal, e podem ser usadas para construir qualquer gene quimérico e vetor, isto é, também em combinação com um promotor diferente.

A seqüência de ácido nucleico a ser expressada é, em uma forma de realização, uma seqüência codificando uma proteína ou peptídeo, incluindo proteínas híbridas ou peptídeos ou proteínas de fusão. A seqüência de codificação pode ser de qualquer origem, isto é, planta, fungo (incluindo levedura), animal, bacteriana, sintética, viral, etc. Ela também pode compreender uma seqüência codificando um peptídeo alvo como um peptídeo sinal de secreção ou um sinal de marcação de plastídeo. Uma seqüência de codificação também pode ser ligada em quadro a um gene codificando um marcador selecionável ou classificável, como por exemplo o gene neo (ou nptII) (EP 0 242 236) conferindo resistência a canamicina, de modo que a célula expressa uma proteína de fusão que é facilmente detectável. Apesar da região de codificação (cDNA ou DNA genômico) de qualquer gene poder ser usada, exemplos das regiões de codificação dos seguintes genes são preferivelmente ligados operativamente a um promotor de AA6 de acordo

com a invenção: 1. seqüências de repetição invertida de seqüências de ácido nucleico viral (por exemplo seqüências sentido e anti-sentido de genes de proteína de capa viral ; ver também abaixo); 2. genes de via de transdução de sinal da doença ou genes de resistência a doença; 3. genes relacionados com resposta ao estresse abiótico (por exemplo fatores de transcrição SHINE, ou genes CBF/DREB); 4. genes de biossíntese metabólito secundário, incluindo genes para a produção de produtos terapêuticos e/ou farmacologicamente importantes ou compostos industrialmente valiosos.

Obviamente, também outros genes podem ser ligados operativamente aos promotores de AA6 da invenção, como genes afetando os traços agronômicos, como genes para resistência a herbicida (por exemplo genes EPSPS, o gene bar ou PAT), genes afetando o rendimento ou traços de qualidade (por exemplo composição de proteína) e outros.

Os genes quiméricos ou vetores de acordo com a invenção também podem ser usados para transformar microorganismos, como bactérias (por exemplo *Escherichia coli*, *Pseudomonas*, *Agrobacterium*, *Bacillus*, etc.) ou fungos ou algas ou insetos, ou os genes ou vetores podem ser usados para engenheirar vírus. A transformação de bactérias com seqüências de ácido nucleico desta invenção, incorporada em um veículo de clonagem apropriado, pode ser realizada em um modo convencional, preferivelmente usando técnicas de eletroporação convencionais como descrito em Maillon *et al.* (1989, FEMS Microbiol. Letters 60, 205-210.) e WO 90/06999. Para expressão de seqüências de codificação em célula hospedeira procariótica, o uso de códons da seqüência de ácido nucleico pode ser otimizado de acordo (do mesmo modo, para expressão de seqüências de codificação em células de plantas, o uso de códons da seqüência de ácido nucleico pode ser otimizado, como conhecido) As seqüências de íntron devem ser removidas e outras adaptações para expressão ótima podem ser feitas, como conhecido.

Para obter uma expressão melhorada de uma seqüência de

ácido nucleico em plantas monocotiledôneas, como espécies de gramíneas, por exemplo milho ou arroz, um íntron, preferivelmente um íntron de monocotiledônea, pode ser adicionado ao gene quimérico. Por exemplo a inserção do íntron do gene A dh1 de milho na região regulatória 5' foi
5 mostrado como melhorando a expressão em milho (Callis *et. al.*, 1987, Genes Develop. 1: 1183-1200). Do mesmo modo, o íntron HSP70, como descrito em US 5,859, 347, pode ser usado para melhorar a expressão. Assim, um ou mais entrona pode ser opcionalmente inserido em qualquer uma das seqüências de promotor de acordo com a invenção ou no 5' UTR ou na seqüência de
10 codificação.

Em outra forma de realização, o promotor de AA6 é usado para fazer um gene quimérico e vetor para o silenciamento do gene, assim o promotor é ligado operativamente a uma seqüência de ácido nucleico sentido / anti-sentido de um gene alvo (gene endógeno ou família de genes que deve
15 ser silenciado). Em ainda outra forma de realização, o gene alvo também pode ser um gene ou família de genes de um patógeno de planta invasora, como um vírus. Por exemplo, uma seqüência de repetição invertida de um gene de proteína de capa viral pode ser usado para tornar as plantas resistentes ao vírus. Os genes de proteína de capa viral são, por exemplo, descritos em WO
20 96/21031. O "silenciamento de genes" refere-se à regulação negativa ou inibição completa de expressão de genes de um ou mais genes alvo. O uso de RNA inibitório para reduzir ou abolir a expressão de genes é bem estabelecido na arte e é o assunto de vários estudos (por exemplo Baulcombe, 1996, Cell Plant 8: 1833-1844; Stam *et al.*, 1997, Plant Journal 12: 63-82;
25 Depicker e Van Montagu, 1997, Curr. Opinion Cell Biol. 9: 373-382). Existem várias tecnologias disponíveis para obter o silenciamento de genes em plantas, como genes quiméricos que produzem RNA anti-sentido de todo ou parte do gene alvo (ver, por exemplo EP 0 140 308 B1, EP 0 240 208 B1 e EP 0 223 399 B1), ou que produzem RNA sentido (também referido como co-

supressão), ver EP 0 465 572 B1.

A abordagem de maior sucesso obtida até agora foi, no entanto, a produção de RNA tanto sentido como anti-sentido do gene alvo ("repetições invertidas"), que forma RNA de filamento duplo (dsRNA) na célula e silencia o(s) gene(s) alvo (s). Os métodos e vetores para a produção de dsRNA e silenciamento de genes foram descritos em EP 1 068 311, EP 983 370 A1, EP 1 042 462 A1, EP 1 071 762 A1 e EP 1 080 208 A1.

Um vetor de acordo com a invenção pode compreender assim um promotor de AA6 ligado operativamente a um fragmento de DNA sentido e/ou anti-sentido de um gene alvo. Os pedaços curtos (sentido e anti-sentido), da sequência de gene alvo, como pelo menos cerca de 17, 18, 19, 20, 21, 22, 23, 24 ou 25 nucleotídeos de sequência de codificação ou não codificação podem ser suficientes. As sequências mais longas são com frequência também usadas, como pelo menos cerca de 100, 200, 250, 300, 400, 500, 1000, 1500 nucleotídeos ou mais. Preferivelmente, os fragmentos sentido e anti-sentido são separados por uma sequência de espaçador como um íntron, que forma um laço (ou grampo de cabelo) quando da formação do dsRNA. Qualquer pedaço do gene alvo pode ser usado para fazer um vetor de silenciamento de gene e uma planta transgênica em que o gene alvo ou família de genes é silenciado. Um modo conveniente de gerar as construções de grampo de cabelo consiste em usar vetores genéricos como pHANNIBAL e pHELLSGATE, vetores baseados em tecnologia Gateway® (ver Wesley *et al.* 2004, *Methods Mol Biol.* 265:117-30; Wesley *et al.* 2003, *Methods Mol Biol.* 236:273-86 e Helliwell & Waterhous 2003, *Methods* 30(4):289-95.), incorporados aqui por referência.

Ao escolher sequências de ácido nucleico conservadas do gene alvo, os membros da família em uma planta hospedeira podem ser silenciados. Onde o gene alvo é um gene ou família de genes de um patógeno invasor, o gene alvo do patógeno será silenciado e a planta se tornará

resistente ao patógeno. Aqui estão englobadas também as plantas transgênicas compreendendo um promotor de AA6, ligado operativamente a um fragmento de DNA sentido e/ou anti-sentido de uma seqüência de ácido nucleico de gene alvo e demonstrando um fenótipo de silenciamento de gene alvo. O fenótipo
5 irá depender da função do gene, e pode ser uma mudança química ou molecular, visível macroscopicamente ou não visível. Estes genes quiméricos e vetores podem, assim, também ser usados para determinar ou verificar a função dos genes.

Os genes quiméricos de acordo com a invenção podem ser
10 introduzidos de modo estável no genoma hospedeiro ou podem estar presentes como uma unidade episomal.

Células transgênicas e organismos de acordo com a invenção

As células e organismos transgênicos, especialmente plantas, células de plantas, tecidos ou órgãos são providos, obteníveis pelos métodos
15 acima. Estas células e organismos são caracterizados pela presença de um gene quimérico em suas células ou genoma pela presença de um promotor de AA6 de acordo com a invenção. Além disso, o transcrito de mRNA ou da proteína traduzida, pode alterar o fenótipo das células ou organismo, por exemplo da célula de planta ou planta.

20 Apesar do promotor de AA6 ser constitutivo, a posição no genoma pode afetar a atividade do promotor e o nível de expressão do gene quimérico. Assim, transformantes (“Eventos” ou “Eventos de Transformação”) expressando níveis elevados, constitutivos da proteína ou do transcrito sentido e/ou anti-sentido (quando construções de silenciamento são
25 usadas) podem ser selecionados por, por exemplo, análise do número de cópias (Análise Southern blot), níveis de transcrito de mRNA (por exemplo análise Northern blot ou RT-PCR) ou por análise da presença e nível de proteína codificada pela seqüência de ácido nucleico (por exemplo SDS-PAGE seguido por análise Western blot; testes ELISA, testes

imunocitológico, etc). Os transformantes também podem ser testados para a estabilidade de expressão sob uma ou mais condições de estresses bióticos e/ou abióticos e os eventos que retêm uma expressão constitutiva elevada sob uma ou mais das condições desejadas, podem ser identificados e selecionados para uso posterior.

As plantas transgênicas podem ser usadas em métodos de procriação tradicionais, como cruzamento, endogamia, retro-cruzamento, etc., etc. Por endogamia dos transformantes, plantas que são homozigóticas para o transgene podem ser geradas. Os procedimentos de procriação são bem conhecidos na técnica e são descritos em livros de texto padrões de reprodução de plantas, por exemplo, Allard, R.W., Principles of Plant Breeding (1960) New York, NY, Wiley, pp 485; Simmonds, N.W., Principles of Crop Improvement (1979), London, UK, Longman, pp 408; Snee, J. et al., (1979) Tomato Breeding (p. 135-171) em: Breeding of Vegetable Crops, Mark J. Basset, (1986, editor), The Tomato crop: a scientific basis for improvement, por Atherton, J.G. & J. Rudich (editores), Plant Breeding Perspectives (1986); Fehr, Principles of Cultivar Development—Theory and Technique (1987) New York, NY, MacMillan.

As células transgênicas ou organismos também podem ser usadas em culturas de células (culturas de células de plantas, culturas de células bacterianas ou fúngicas, como culturas de leveduras, culturas de células humanas ou mamíferos, culturas de células de insetos), por exemplo para a produção em larga escala de proteínas recombinantes. Em uma forma de realização, uma cultura de células é provida, compreendendo células compreendendo um promotor de AA6 de acordo com a invenção.

Métodos e usos de acordo com a invenção

Também é provido um método para fabricar uma planta transgênica ou célula de planta, compreendendo as etapas de:

(a) gerar um gene quimérico ou um vetor compreendendo um

promotor de AA6 e/ou um AA6 3'UTR de acordo com a invenção, ligado operativamente a uma seqüência de ácido nucleico a ser expressada;

(b) transformar uma planta ou célula de planta com referido gene quimérico ou vetor; e, opcionalmente,

5 (c) regenerar a planta transgênica ou plantas.

Além disso, as plantas transgênicas podem ser identificadas que provêm uma atividade de promotor constitutivo, forte, sob condições de não estresse, e assim, a atividade do promotor permanece essencialmente não mudada (isto é, pelo menos não reduzida, ou não reduzida de modo
10 significativa) quando a planta é exposta a um ou mais estresses bióticos e/ou abióticos.

As plantas podem ser usadas em métodos agrícolas e de procriação convencionais. Particularmente, elas podem ser cultivadas em meios que submetem as plantas a uma ou mais condições de estresses bióticos
15 e/ou abióticos, sem reduzir a atividade do promotor. Assim, é provido um método para cultivar plantas transgênicas em, por exemplo, áreas com temperatura alta ou baixa, ventos fortes, elevado teor de sal, elevado teor de contaminantes do solo, níveis de água altos ou baixos, períodos climáticos de seca, áreas de elevado índice de doenças ou pressão de pragas, elevada
20 radiação, etc, sem reduzir, ou pelo menos sem reduzir de modo significativa a atividade do promotor.

SEQÜÊNCIAS

SEQ ID NO 1: "3kb" (2986 bp) promotor de AA6

SEQ ID NO 2: "5kb" (5000 bp) promotor de AA6

25 SEQ ID NO 3: 3'UTR de gene de cisteína sintase de tomate (AA6 gene)

SEQ ID NO 4: cDNA de cisteína sintase de tomate (AA6 cDNA)

SEQ ID NO 5: proteína codificada por SEQ ID NO: 4 (cisteína

sintase citossólica de tomate)

SEQ ID NO 6: “3kb” promotor de AA6 com algumas ambigüidades

5 SEQ ID NO 7: “5kb” promotor de AA6 com algumas ambigüidades

SEQ ID NO 8: “3kb” promotor de AA6 de pKG8135, como re-sequenciado e como em depósito de E. coli CBS120175)

SEQ ID NO 9: “5kb” promotor de AA6 de pKG8137, como re-sequenciado e como em depósito de E. coli CBS120176)

10 Figuras

Figura 1: Mapa de vetor de pKG1562

Figura 2: Mapa de vetor de pKG1700

Figura 3: Mapa de vetor de pKG8135, pKG8136 e pKG8137.

Os seguintes exemplos não limitativos descrevem o uso de
 15 promotores de AA6 de acordo com a invenção. Salvo especificado em
 contrário nos Exemplos, todas as técnicas de DNA recombinante foram
 realizadas de acordo com os protocolos padrões, como descrito em Sambrook
et al. (1989) *Molecular Cloning: A Laboratory Manual*, 2a. edição, Cold
 Spring Harbor Laboratory Press, e Sambrook e Russell (2001) *Molecular*
 20 *Cloning: A Laboratory Manual*, 3a. ed. Cold Spring Harbor Laboratory Press,
 NY; e em Volumes 1 e 2 de Ausubel *et al.* (1994) *Current Protocols in*
Molecular Biology, Current Protocols, USA. Os materiais e métodos padrões
 para o trabalho molecular com plantas são descritos em *Plant Molecular*
Biology Labfax (1993) por R.D.D. Croy, publicado em conjunto por BIOS
 25 Scientific Publications Ltd (UK) e Blackwell Scientific Publications, UK.

EXEMPLOS

Exemplo 1 - Materiais e Methods

Análise de exibição diferencial usando cDNA foi realizada em
Lycopersicon esculentum var. Moneyberg.

O teste de expressão sob condições de estresse foi realizado em *L. esculentum* var. Moneyberg.

Transformações usando vetores pKG8136, pKG8137 e pKG1700 foram realizadas em *Nicotiana tabacum* var. SR1 e *L. esculentum* var. RZ 52201. A transformação da planta foi realizada como descrito nas seguintes referências:

Para transformação de tomate e tabaco:

An G., B.D. Watson e C.C. Chiang. 1986. Transformation of tobacco, tomato, potato, and *Arabidopsis thaliana* using a binary Ti vector system. *Plant Physiol.* 81: 301-305.

Para transformação de tabaco : Horsch R.B., J. Fry, N. Hoffman, J. Neidermeyer, S.G. Rogers e R.T. Fraley. 1988. Leaf disc transformation. Em: *Plant Molecular Biology Manual* A5. Dordrecht, Netherlands, Kluwer Academic Publishers. pp 1-9.

Para transformação de tomate: Koornneef M., M. Jongsma, R. Weide, P. Zabel e J. Hille. 1986. Transformation of tomato. Em: Nevins D.J. e R.A. Jones, eds. *Tomato Biotechnology*. New York, NY, USA, Alan R. Liss, Inc. pp 169-178.

RACE de extremidade 5' e 3' foi realizado usando o kit de amplificação de cDNA SMART™ RACE de CLONTECH Laboratories Inc.

Clonagem de produtos PCR foi realizada com o kit Original TA Cloning® de Invitrogen BV. usando plasmídeo pCR®2.1.

O seqüenciamento de DNA foi realizado por BaseClear (P.O. Box 1336, 2302 BH, Leiden, Holanda).

Os vetores de transformação foram incorporados em *Agrobacterium tumefaciens* var. GV2260.

A análise qualitativa de atividade do promotor foi realizada usando testes GUS histológicos:

Várias partes de plantas foram incubadas durante a noite a 37

°C com agitação na presença de oxigênio atmosférico com substrato Xgluc em tampão fosfato (5-bromo-4-cloro-indolil glucuronídeo, 1 mg/ml, K₂HPO₄, 40 mM, KH₂PO₄, 10 mM, pH 7,4). As amostras foram descoloridas por lavagem repetida com etanol. As plantas não transgênicas foram usadas como

5 controles negativos.

A análise quantitativa de atividade do promotor foi realizada usando um teste GUS fluorométrico:

As amostras de proteína total foram preparadas a partir de um materiais de folha jovem; as amostras foram preparadas a partir de pedaços de

10 folhas reunidos de aproximadamente o mesmo tamanho e estágio de desenvolvimento de diferentes partes de cada planta testada. O material de folha novo foi triturado em tampão fosfato (Na₂HPO₄, 77,4 mM, NaH₂PO₄, 22,6 mM) usando contas de metal seguido por centrifugação e coleta do sobrenadante.

15 Testes GUS fluorométrico

A concentração de proteína em uma alíquota de cada sobrenadante foi medida usando o kit Nano Orange de Molecular Probes, Inc. As amostras de proteína foram diluídas para normalizar a concentração de proteína total. As alíquotas das amostras de proteína foram incubadas durante

20 a noite a 37 °C com o substrato 4-metil umbeliferil β-d-glucuronídeo (MUG) (concentração final de 1 mg/ml). As medidas de fluorescência em tempo zero e após incubação noturna foram tomadas por remoção de alíquotas de mistura de reação e adição de solução de Na₂CO₃ (concentração final de 1,1 M) para parar a reação então medindo a emissão a 460 nm causada por excitação a 355

25 nm.

Exemplo 2 - Seleção de um gene para isolamento do promotor

A análise diferencial foi realizada usando cDNA-AFLP® (Volkmuth W., *et al.*, 2003, Genome-Wide cDNA-AFLP® Analysis of the Arabidopsis Transcriptome, OMICS, 7, 2; Vos, P. e Stanssens P., 2002,

AFLP-based transcript profiling, Current Protocols in Molecular Biology, unit 25B.5., Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G. Smith, J.A., Struhl, K., John Wiley and Sons, New York; e Vos *et al.*, 1995, AFLP: a new technique for DNA fingerprinting, Nucleic Acids Research 23: 4407-4414) em cDNA isolado de materiais de planta de tomate em vários estágios de desenvolvimento (folha jovem, folha velha, caule, raiz, folha *in vitro*, caule *in vitro*, raiz *in vitro*, calo, fruto verde imaturo e maduro, fruta no estágio de romper, fruta vermelha, flores, mudas totais). Os perfis de expressão foram produzidos usando todas as combinações de iniciadores Taq1/Mse1 possíveis com dois nucleotídeos seletivos para o iniciador Taq1 e três nucleotídeos seletivos para o iniciador Mse1. Os perfis de expressão de fragmentos de genes subsequentes foram produzidos usando software registrado Keygene N.V. (Improve™) e armazenados em bases de dados. transcritos candidatos correspondendo a combinações de iniciadores Taq1/Mse1 foram selecionados, os quais mostraram uma expressão constitutiva forte em todas os tecidos amostrados.

Expressão dos genes candidatos foi então medida por cDNA-AFLP® em vários tecidos sob estresses bióticos e/ou abióticos. Plantas de tomate com três semanas de idade (Moneyberg) foram cultivadas durante 14 dias com as seguintes condições de estresse separadas:

Infecção com o vírus do mosaico do pepino (CMV), temperatura alta (37°C dia e 25°C noite), temperatura baixa (10°C dia e 5°C noite), estresse de seca (água mínima para evitar morte da planta, com 37°C dia e 25°C noite) e feridas (esmagamento das folhas com tesouras) 2 dias antes da colheita.

O seguinte material foi coletado: folhas jovens, folhas velhas, caule, raízes e brotos de flores. cDNA destes materiais foi produzido e a análise diferencial realizada usando cDNA-AFLP® com as mesmas combinações de iniciadores (e também usando combinações de iniciadores

com menos nucleotídeos seletivos) requeridos para amplificar os 13 transcritos candidatos selecionados (ver abaixo para seqüências de iniciador). Sete transcritos foram selecionados que continuaram a dar uma expressão constitutiva forte, sob todas as condições de estresse acima.

- 5 As seqüências dos iniciadores dos iniciadores (+2/+2) AFLP que amplificam o fragmento de gene da cisteína sintase (AA6) são:

Código de iniciador AFLP	Código de gene chave	seqüência
Tr13	96F06	5' GTAGACTGCGTACCGAAG 3'
M26	92N16	5' GATGAGTCCTGAGTAATT 3'

Os fragmentos de AFLP de cDNA candidato foram cortados do gel, purificados, clonados e seqüenciados. A informação de seqüência foi então usada para projetar iniciadores para RACE PCR.

- 10 Tanto a extremidade 5' como 3' SMART-RACE-PCR (Clontech) foram realizadas para produzir fragmentos de genes estendidos dos 7 genes candidatos, estes foram clonados em pCR2.1 (Invitrogen) e subsequentemente seqüenciados (BaseClear, P.O. Box 1336, 2302 BH, Leiden, Holanda). Os fragmentos de genes estendidos foram montados junto com ESTs homólogos de bases de dados públicas. Isto resultou em seqüências de cDNA de comprimento completo para os 5 genes candidatos.

- 15 Uma avaliação da contagem de cópias para os 5 genes foi realizada usando Southern blotting com DNA genômico de tomate digerido com 10 diferentes enzimas de restrição e usando os fragmentos de extremidade 5' dos cDNAs como sondas. As mesmas sondas foram também hibridizadas em uma biblioteca existente Moneyberg BAC (Keygene N.V.).

- 20 Um dos cDNAs de comprimento completo (designado IS158-53; SEQ ID NO: 4) mostrou uma homologia muito elevada para o gene de cisteína sintase citossólica (tomate|TC162833 homólogo a UP|Q9FS27 (Q9FS27). TC162833 e SEQ ID NO: 4 codificam, ambos, a mesma seqüência de aminoácido (SEQ ID NO: 5), mas diferem no nível de ácido nucleico em dois nucleotídeos. Nucleotídeo 'C' em posição 207 de SEQ ID NO: 4 é um

'T' em TC162833, enquanto nucleotídeo 'A' em posição 372 de SEQ ID NO: 4 é um 'G' em TC162833.

SEQ ID NO: 4 foi expressada constitutivamente mesmo sob condições de estresses bióticos e/ou abióticos, ela estava presente em um número de cópias baixo (aproximadamente duas cópias) no genoma do tomate e foi hibridizada bem em BAC 9 da biblioteca BAC de tomate Keygene N.V. assim o isolamento de seu promotor foi realizado.

O promotor do gene cujo cDNA de comprimento completo tinha homologia com o com gene de cisteína sintase de tomate (SEQ ID NO: 4) foi selecionado, como descrito abaixo.

Exemplo 3 – Isolamento da seqüência do promotor de AA6

Os iniciadores projetados sobre a seqüência derivada do gene de tomate designado como cisteína sintase (ver abaixo para as seqüências do iniciador) foram usados para realizar PCR "ligador" usando DNA genômico de BAC 9 como gabarito. Tanto a seqüência 5' como 3' foi obtida para os UTRs de cisteína sintase. A seqüência 3' UTR (SEQ ID NO: 3) foi obtida pelo PCR de faixa longa usando DNA genômico de tomate Moneyberg como o gabarito.

Os iniciadores designados em seqüência de cisteína sintase:

Para promotor	5'-GTTTCGATGAGGACACTCTCGC-3
e iniciador abrigado	5'-CAATTAAAGTTGCTAAGCGTCCTGA-3'
Para terminador	5'-TCAGTTACATCCTTGGCAATTCC-3'
e iniciador abrigado	5'-CAGAGAACATGACTGTGGAGCC-3'

Exemplo 4 - Construção de vetor de transformação

Dois fragmentos de DNA de promotor de AA6, um de 5000bp (SEQ ID NO: 2) e um de 2986bp (SEQ ID NO: 1), da região a montante 5' da cisteína sintase 5' foram ligados juntos, com a região de codificação de *gusA*, no plasmídeo pKG1562 (estrutura dorsal do vetor, contendo nptII dirigido pelo promotor de nopalina sintase) com ou o terminador endógeno de cisteína

sintase (SEQ ID NO: 3) ou o terminador *nos*.

As construções foram projetadas:

contendo pKG8135 (Ver figura . 3):

2986 bp da região a montante 5' de cisteína sintase (SEQ ID

5 NO: 1): *gusA* : 3'*nos*

contendo pKG8136 (Ver figura . 3):

5000 bp a região a montante 5' da cisteína sintase (SEQ ID

NO: 2): *gusA* : 3'*nos*

contendo pKG8137 (Ver figura . 3):

10 5000 bp da região a montante 5' de cisteína sintase (SEQ ID

NO: 2): *gusA* : 3' cisteína sintase (SEQ ID NO: 3)

Uma construção de controle foi usada, como a seguir (Ver figura . 2):

contendo pKG1700: promotor CaMV35S (Franck *et al.*,

15 supra): *gusA* : 3' *nos*

Ver figuras 1-3 para detalhes do mapa de vetor dos vetores de transformação usados.

Os plasmídeos pKG8135, pKG8136, pKG8137 e pKG1700 foram incorporados em *Agrobacterium tumefaciens*. Tabaco (var. SR1) e
20 tomate (var. RZ 52201) foram transformados e regenerados sob seleção de canamicina e regenerantes primários (T₀) foram cultivados até semeadura.

Exemplo 5 - Análise da expressão de transformantes

Porque a expressão de *gusA* é dirigida pelo promotor de AA6, ou o promotor de CaMV35S, nestes vetores, o nível de β -D-glucuronidase
25 (GUS) produzido nas plantas indica a eficácia do promotor. Os testes de coloração histológica para a presença de GUS foram realizados usando 5-bromo-4-cloro-indolil glucuronídeo (XGluc) como substrato.

As experiências de coloração foram realizadas em tecido de folha tanto de plantas de tabaco como de tomate T1 (obtidas após endogamia

de transformantes primários) em vários estágios de desenvolvimento desde as mudinhas até as plantas florescentes. Adicionalmente, as experiências de coloração histológica para GUS foram realizadas em uma faixa de material de linhagens de tabaco e tomate maduras, incluindo brotos de flores, flores abertas, hastes de flores, (tabaco), hastes de flores, brotos de meristema apical, folha, talos de folhas, caule principal, raízes, sementes e fruto (tomate) e também em mudinhas e sementes germinando.

Expressão de GUS foi detectada em todos os tecidos testados tanto em tabaco como tomate quando *gusA* foi dirigido ou pelo fragmento longo (EQ ID NO: 2) ou fragmento curto (SEQ ID NO: 1) do promotor de AA6 ou do promotor CaMV35S.

A comparação visual (coloração de cor) indicou que os fragmentos do promotor de AA6 apresentaram um desempenho tão bom como o do promotor de CaMV35S, tanto em tabaco como em tomate.

A análise quantitativa de expressão de GUS foi realizada em um teste fluorométrico usando 4-metil umbeliferil β -d-glucuronídeo (MUG) como substrato. Expressão de GUS dirigida pelo fragmento de 5000bp de promotor de AA6 (pKG8137) foi comparada com a expressão de GUS dirigida pelo promotor de CaMV35S (pKG1700) em tanto folhas de tabaco como de tomate.

As plantas de cópia única foram selecionadas usando Southern blotting; DNA genômico de plantas de tabaco e tomate que foram digeridas com enzima de restrição HindIII e um fragmento de *nptII* foram usadas como sonda, o processo foi repetido usando XbaI como enzima de restrição. Porque cada inserto de T-DNA contém uma cópia de cada dentre *nptII* e *gusA*, o número de cópias de *nptII* reflete o número de cópias de *gusA*. Os números de copias foram também confirmados usando um teste Invader® (Third Wave Technologies, Inc.) detectando a presença e a quantidade relativa do gene *nptII* em comparação com um controle interno de tomate ou tabaco

conhecido; a aplicação deste teste à geração de T1 permite a discriminação entre indivíduos hemizigóticos e homozigóticos.

Os extratos de proteína total foram feitos a partir de transformantes de cópia única hemizigóticos e homozigóticos pKG8137 e pKG1700 e controles de cópia zero de plantas de tabaco e tomate. O teor de proteína foi medido usando o kit NanoOrange™ (Molecular Probes) e então a concentração da proteína nas amostras foi normalizada por diluição com tampão de extração.

Após a adição de substrato MUG, a fluorescência inicial foi medida (filtro de excitação, 355 nm e filtro de emissão, 460 nm). As amostras foram então incubadas e a fluorescência devido a metilumbeliferon clivado (MU) foi medida. O aumento na fluorescência por hora de incubação foi calculada. As curvas de calibração mostrando atividade da enzima GUS (Sigma-Aldrich Chemie B.V.) em um fundo de tabaco não transgênico ou extrato de proteína total de tomate (igual ao teor de proteína total das amostras de teste) foram produzidas. As concentrações de enzima GUS (mU/mg de extrato de proteína de planta total) presente nas amostras de testes transgênicas foram calculadas usando os dados de fluorescência das amostras e curvas de calibração.

As análises estatísticas de experiências em duplicata indicaram que os níveis de expressão GUS dirigida pelo promotor de AA6 (pKG8137) e o promotor de CaMV35S não foram significativamente diferentes em tabaco, enquanto que em tomate, expressão de GUS devido ao promotor de AA6 foi de até 79% maior do que devido ao promotor de CaMV35S, em plantas de cópia única homozigóticas (conjuntos de dados mostrados nas Tabelas abaixo).

Expressão de GUS em tabaco

Tabela 1: Atividade de GUS média (expressada em mU/mg de extrato de proteína de planta total) por linhagem para plantas de tabaco

hemizigóticas com seu desvio padrão e erro padrão da média (SEM).

Linhagem/ promotor	Conjunto de dados 1			Conjunto de dados 2		
Linhagem	Média	Desvio padrão	SEM	Média	Desvio padrão	SEM
AA6-03	14250	9702	4339	26471	20251	9057
AA6-05	4107	1773	793	5411	3021	1511
AA6-06	6785	3965	1773	13608	7701	3144
AA6-07	3451	1600	716	3323	1403	627
35S-06	5151	2535	1268	10294	6224	2783
35S-09	5316	2105	1052	9303	2745	1228
35S-13	15557	18869	9435	21330	12911	5774
35S-25	11226	10219	5110	15114	7978	3568
35S-28	1425	23	10	1388	10	4
35S-29	8796	4370	1954	20911	8540	3819
Controle negativo	1425	23	10	1388	10	4
média AA6	7148	5375	2687	12203	11074	5537
média 35S	7912	8728	3563	13057	7633	3116

Tabela 2: Atividade de GUS média (expressada em mU/mg de extrato de proteína de planta total) por linhagem para plantas de tabaco homozigóticas com seu desvio padrão e erro padrão da média (SEM).

Linhagem/ Promotor	Conjunto de dados 1			Conjunto de dados 2		
Linhagem	Média	Desvio padrão	SEM	Média	Desvio padrão	SEM
AA6-03	12469	8648	3867	18877	12264	5485
AA6-05	8645	4711	2355	11223	6394	3197
AA6-06	3954	826	477	6169	2580	1154
AA6-07	3070	1018	509	6485	3832	1916
35S-06	8040	4051	2025	11133	6186	2765
35S-09	10234	2832	1416	20216	7398	3699
35S-13	14996	4704	3326	26516	17900	8005
35S-25	5841	2272	1016	17501	9293	4156
35S-28	1436	10	5	1404	19	9
35S-29	9295	3996	2307	25035	17459	10080
Controle negativo	1436	10	5	1404	19	9
média AA6	7035	5414	2707	10689	7482	3741
média 35S	8307	3131	1278	16968	10641	4344

Como pode ser visto, a atividade de promotor média do promotor de AA6 (promotor de 5 kb) é essencialmente igual à do promotor de 35S em folhas de tabaco transgênicas (tanto em plantas hemizigóticas como

homozigóticas para o transgene).

Expressão de GUS em tomate

Tabela 3: Atividade de GUS média (expressada em mU/mg de extrato de proteína de planta total)

5 Por linhagem para plantas de tomate hemizigóticas com seu desvio padrão e erro padrão da média (SEM).

Linhagem / Promotor	Conjunto de dados 1			Conjunto de dados 2		
Linhagem	Média	desvio padrão	SEM	Média	Desvio padrão	SEM
AA6-11	1126	895	400	2615	621	278
AA6-29	144	188	77	1538	739	302
AA6-30	25302	4602	2301	31004	2078	929
AA6-31	24370	4927	2463	31512	2203	1102
35S-01	5027	8483	3794	7797	9420	4213
35S-04	11731	2189	979	11479	2157	965
35S-23	3024	6272	2805	6988	9170	4585
35S-27	14310	4775	2388	15625	4109	1838
Controle negativo	0	0	0	757	875	309
médias totais AA6	12736	3129	1565	16667	1520	760
médias totais 35S	8523	5946	2973	10472	6835	3418

Tabela 4: Atividade de GUS média (expressada em mU/mg de extrato de proteína de planta total)

10 por linhagem para plantas de tomate homozigóticas com seu desvio padrão e erro padrão da média (SEM).

	Conjunto de dados 1			Conjunto de dados 2		
Linhagem	Média	Desvio padrão	SEM	Média	Desvio padrão	SEM
AA6-11	1357	844	377	2259	574	257
AA6-29	14	28	14	1279	915	458
AA6-30	35784	9382	4196	39472	6631	3828
AA6-31	36187	5942	2657	41129	2247	1125
35S-01	10496	6385	2856	11025	6425	2873
35S-04	13638	2642	1321	13569	1219	609
35S-23	4551	9103	4551	1659	434	217
35S-27	19538	2080	1201	20685	1128	564
controle negativo	0	0	0	757	875	309
médias totais AA6	18336	5713	2857	21035	3134	1567

médias 35S	totais					
		12056	6001	3001	11735	3583
						1791

Como se nota na Tabela 4, a atividade média de promotor do promotor de AA 6 (promotor de 5 kb) é significativamente maior (em conjunto de dados 1, cerca de 52% maior e no conjunto de dados 2, cerca de 79% maior do que 35S, quando tomando a atividade de 35S como padrão) em tomate comparado com a do promotor de 35S em folhas de tomate transgênico (tanto em plantas hemizigóticas como homozigóticas para o transgene).

Exemplo 6 – Atividade de promotor em outras espécies de plantas

A comparação de expressão de GUS dirigida pelos promotores de AA6 e CaMV35S em outras espécies de plantas foi testada e confirmada. Atualmente, a lista de espécies testadas com a expressão dirigida pelo promotor de AA6 e um transgene inclui (além de tabaco e tomate, como mostrado acima), alface, melão, e *Brassica*, em cada uma das quais a expressão constitutiva foi mostrada.

Exemplo 7 – Re-sequenciamento dos promotores de AA6 de “3kb” e “5kb”

As seqüências de ácido nucleico presentes em vetores pKG8135 e pKG8137 (Fig. 3) foram re-sequenciadas para verificar as seqüências e resolver as ambigüidades.

As posições ambíguas são indicadas na listagem de seqüência de SEQ ID NO: 6 (promotor “3kb”) e SEQ ID NO: 7 (promotor “5kb”), em que o também o nucleotídeo o mais provável na posição ambígua é identificado (ver também nas tabela abaixo).

Devido aos nucleotídeos ambíguos, as diferentes de seqüência mínimas existem : SEQ ID NO: 1 tem 99,6% e 99,7% de identidade de seqüência para SEQ ID NO: 8 e 6, respectivamente (usando ‘needle’, abertura Gap = 10.0, extensão de espaço =0.5; matriz DNABFull). SEQ ID NO: 2 tem

99,6% e 99,8% identidade de sequência para SEQ ID NO: 9 e 7, respectivamente (usando 'needle', abertura Gap = 10.0, extensão de espaço =0.5; Matriz DNABFull).

Ambigüidades e resolução de ambigüidades:

Posição de nucleotídeo em SEQ ID NO: 6	ambigüidade	nucleotídeo mais provável
679	W (ambigüidade A/T)	T
682	W (ambigüidade A/T)	A
683	W (ambigüidade A/T)	T
691	A inserido	A inserção parece confiável
864	Y (ambigüidade T/C)	T
1269	R (ambigüidade G/A)	A
1559	R (ambigüidade G/A)	A
1560	M (ambigüidade A/C)	C
1561	D (ambigüidade G/A/T)	T
1916	R (ambigüidade G/A)	A
2085	Y (ambigüidade T/C)	T
2101	Y (ambigüidade T/C)	C
2463	Y (ambigüidade T/C)	T
2526	A deve ser deletado	deleção de A parece confiável
2534	R (ambigüidade G/A)	G
2681	K (ambigüidade G/T)	G
2920	R (ambigüidade G/A)	A

Posição de nucleotídeo em SEQ ID NO: 7	ambigüidade	nucleotídeo mais provável
35	G inserção	G inserção parece confiável
240	Y (ambigüidade T/C)	T
768	W (ambigüidade A/T)	A
769	Y (ambigüidade T/C)	T
1558	R (ambigüidade G/A)	A
1742	Y (ambigüidade T/C)	T
2016	T inserção	T inserção parece confiável
2695	W (ambigüidade A/T)	T
2698	W (ambigüidade A/T)	A
2699	W (ambigüidade A/T)	T
2707	A inserido	A inserção parece confiável
2880	Y (ambigüidade T/C)	T
3285	R (ambigüidade G/A)	A
3575	R (ambigüidade G/A)	A
3576	M (ambigüidade A/C)	C
3577	D (ambigüidade G/A/T)	T
3932	R (ambigüidade G/A)	A
4101	Y (ambigüidade T/C)	T
4117	Y (ambigüidade T/C)	C
4479	Y (ambigüidade T/C)	T
4542	A deve ser deletado	deleção de A parece confiável

4550	R (ambigüidade G/A)	G
4697	K (ambigüidade G/T)	G
4936	R (ambigüidade G/A)	A

LISTAGEM DE SEQUÊNCIA

<110> KeyGene NV

<120> Promotor de planta constitutivo

<130> P6005193PCT1

<150> PCT/NL2005/050083

<151> 16- 12 - 2005

<160> 9

<170> PatentIn versão 3.3

<210> 1

<211> 2986

<212> DNA

<213> Lycopersicon esculentum

<400> 1

```

tgggttttcta attctctcat ctcttttctc tggtcttgta tcttatccct ttctgattat   60
cttcttttgt gttaatctct ccttagttaa tatattatta ctttttgcac gtattccaat  120
agtaacttat attcaacaat aattatatac aatagtgttt gtattagaat cgagtttatt  180
tcataaagtc tcgaaagtaa tacattactt tttttagaat aaaaaaaaaa cttgggaata  240
tggaagttca atctacgcaa aaacctatct tatgatgtca aatgaattca atgtttttac  300
tcatattgca atataaaaag aacagggtca gtgatttttt ttccttcaaa tgattgctag  360
ttaattaatg gacttggtct ttcttttgtt tcatattctt gcctttttta agaccttcca  420
aatcatcaac acaaaaacac cttagacatt attgcatccg acttaattta tatgacataa  480
aatataaagt ttttttttta aaataaaaat aaagagtttt aaattttata agaaagaaat  540
atgtatatatt aagttcatal taattataaa aacatgattt attaaaaata aatataaaaa  600
atatatcata taaattaaaa cgaatgaaat atagtatatt aggatgagaa ttcaaaggac  660
tagaatttcc aaagaaatwt gwwgtttttt atgttttatt agttgacttg ttctatgccg  720
gtcttgctcg ttcaccatat atagaccaac acttaaagta aatagtggta atttgatttg  780
acctttcaaa accttcaagg gcccaattaga cagaaacttt atattataaa agtaataatg  840
aatagtagta ataaactaat tattaccaag aatgatgaa taattgttta tatatatata  900
tagtgcaata ataattcata accggccgat cctatatatt tattataaaa tatagaagtt  960
atccaataat ttcatgtctt ttattaaaaa aattcattca ataatttatg ttttttcttt 1020
taciaaagtt aataatagaa cttatttaat taaaactaat gttactatta tgatgaaata 1080
gtcgagatct cttcattttt cataagagat tcaggattca agctgattat gtgagtgaaa 1140
tcctaaaaaa tgaaatttgt aatacgtaaa taatacgaat atcaaatacg aaacaaaaag 1200

```

aaaattaatg ttattatgat tttgaatfff aattatggta ttttaactac tgaagttgtc 1260
 tcattgaatg tatattataa ttaagcatat acaattccag aaaatatata ttattatcta 1320
 ggtgttcttt taatataata tttgagtttt atgataaaat ttgattaact tttaaaggaa 1380
 aatcttttta aaagtggca cattagattg agataaatca gaaaggacac gtcacaatct 1440
 aaatatgatt tttttaactt catctataaa tatctttcaa attatataca atgcttacgt 1500
 gaataaatgc tggcactaat gaattaatat ttgtttttta aataaaatgt tgaaattgmd 1560
 tttagaagtg aacacatgag tcgaagtgtt agaaatagtt actaataaat aatcaatgta 1620
 tttgatataa aaagtgttat taagttgttg tttgtttaga ataactcaa cattttcatt 1680
 ttatttaatt aaaaattata aatttaact atatgtataa acaaaatgag caataataaa 1740
 tatagaatga agcaaacatt atgaatttta ttttgaaaaa aaatgatgtg aattataaaa 1800
 atattatata atatttgatc aaacaaattg tttactataa gctaaagaaa tcttgatttt 1860
 aacttataaa tactttactt atcaaatacg taaaaaaatt aaaaattacc tataaatcaa 1920
 tttaatctaa cacttttagtc aaacacatta taaatatgaa cttgttgtgc cacgattata 1980
 cacatatatt tttttttatt tattgatata tcttcaaaaa taatgtatct cttaaaatat 2040
 tagtatttat taagatacgg gttcatgaaa tattctttta acttcaaatt tttatttgac 2100
 tttgaaaaca gaccctaaa aagtttaaaa ataacaaaaa ggcaacaaaa aagtttgaaa 2160
 caagaactaa aaaagttaga tgcccaaatac tgcacgataa ggtaagtcaa acaacgacaa 2220
 aatttagtaa attattgatt gtagaaaatg aagcaaatga aggtggctat ccttatataa 2280
 agctatattc agtactctcc tcaacacagt cacagtagat ttcatttctt tgtctttgtc 2340
 aaatcctagg taatttttct ctgcccctta atattttttt ttgctcttga ttgcctttta 2400
 tggaaaattg gttctctgtt catttgacaa gattgctgtt ttgttttttc ttcagtctga 2460
 ttagattttt gctctagatt cgtgggtttt acttgtaaag gtattaggga gtgagtgaga 2520
 tgggaaaatt tcratttttt tttaaaaaaa aattatggga taaaaagtga ttccgagaaa 2580
 tgggggttttg ttttttgaga ttgtatgaac tggggatgat ggattcttgg tgggtgattga 2640
 ataaaaatgg gatctttaat tattagggat aagattaaag atgctcatgt tcttctttga 2700
 atttgtttgt atgcaaagtt ttgagctttt ggggtgtttc cccctaattt agctggcaaa 2760
 tagttgtaat tgggtggaagt tggaaaagga aaagaattaa gaagaaattg atgttcattt 2820
 cagtttcttt ttatatattg taaggaacgg aatcacgggt gtcgagtgga atagatttag 2880
 agggttctca aatagttaaa tgagtagctg gtttgatgaa atttgttttg attgtataag 2940
 atttcactta gttttgctat aaatcatcgc agaatcaacc tgatac 2986

<210> 2

<211> 5000

<212> DNA

<213> *Lycopersicon esculentum*

<400> 2

```

catataaccc tgaaaattat ttaccattc ttgatcttc tttttcattt tatttgacca 60
atttataagt attatnttaa aagtttaatt ttttttcaaa tttcacgatg aattatggcc 120
aaatttctcc tagttgtggt ttgaccataa ataaaaatag aagttgttct ttactttttt 180
gtgagtaatt tgaagtgatt tttttcctaa gtttttgaaa ttcaacttta attttttlyt 240
gtttttttgt gacaaggaaa tctgtcttgc tatcttttgg gtgcatacac ggtaaaattc 300
ctgctcatat gcaataactc actctaactt taagttgaac ttaacaattt tacgattaaa 360
gtgttttcta aactcaactt catgaaaaaa ttaaaagtta ttgaggaaat tttaaaaata 420
actaacactg aataatataa ttaaactcct tagctaagt ttctctaatt acgattcata 480
gctactatta gaaggagaga agtgagcgac aatatgagag gaggagaaat acgagcgaga 540
tctgagagag atacaatcaa tttgtgtatc tgattgatac ttgtattatt gttgtaattc 600
ataataaatt gtatttaatt atatttatct attgccttgt atattgtatt tatgtgttac 660
catgtatatt tgtgtacca ttttgtattt gtatcattga ctgctatcgt tgtacgtgta 720
ttgtagaaag agagagagga gaggaatagt gagaagattg ctataawycc aaattataat 780
tacgaatagt aattctttta aattataatt atgtatagtt aagacatatt ttgtacgttt 840
atctacattt gccccaaaat tattctcctg tggactctca aagcaggaac caaatagtga 900
tggattgggc tcctagttct gaacgaattt tgttattggg cttaaagtta aggttggttg 960
gtcaggtttt tgggtttgca tatctatggg gtagtaagggt ttattctttt ctgagattc 1020
gatatttata tttgagcttt attaaatttg gattcacacc aaaaagtttt atattgaggg 1080
gggggggggg taaagtgcct cctaataaag gcaaaaacga ttccctactt agaagattca 1140
aactcaaaaa actcgggatg aagtatttac aattctatca cgttgttgat tagttagctt 1200
tatcattgta tgagggtatac ttgtgggata gaatgttaag acaatttatg aatttctttt 1260
ttgaatctag aatgtggttt gatgaacaat ggacttaaaa gaagttaatg gaaaatgaga 1320
aatcaaatth caataaaaaat gaagtthtta gtgattcttt tcatctatcg ataggtaaaa 1380
gatatatcat ttggtctaatt tcaatcccaa aagttagttt aaagggaaaa gcgggttata 1440
aagagtacat tcgtctcatt aaccaccaat gtgaactttt attattcttc aacacccatc 1500
ttattcctag tttttttagc atttgggtgtg tgaacaattt tttatttcgg aggtccraca 1560
tcgggtgaga agagttctac tctgatatca tgtgaaatta ggttttgagc taaacttaca 1620
ccaaaaagct agcttaaaaag atgaaggatt gtacaagcct tataaggaat tcacccgtct 1680
tattaagcat caatataaaa cttttgtcat tcttaaaaaa tataacttcg ataaaacat 1740
ycaaataaat atagtataca catgtacaaa gatgtatcta ctcaatgaaa tatatgagcg 1800

```

caaatgagat actatgacta ccatcatcgt aaataaaaaa acacataaag tatatatatta 1860
 aaggttggtg ttataaaagt gtcatgaacc aaactgatcc cgtaaggaca tatgggctat 1920
 tagagagttt tattagggat gcaagtctat tcactcttcg tgaaaaagga ctaggagtta 1980
 gaagagtgat tatgcaaatt attatgtcga agcttggttt tctaattctc tcatctcttt 2040
 tctctgttct tgtatcttat ccccttctga ttatcttctt ttgtgttaat ctctccttag 2100
 ttaatatatt attacttttt gcatgtattc caatagtaac ttatattcaa caataattat 2160
 atacaatagt gtttgattta gaatcgagtt tatttcataa agtctcgaaa gtaatacatt 2220
 acttttttta gaataaaaaa aaaacttggg aatatggaag ttcaatctac gcaaaaacct 2280
 attttatgat gtcaaatgaa ttcaatgttt ttactcatat tgcaatataa aaagaacagg 2340
 gctagtgatt ttttttcctt caaatgattg ctagttaatt aatggacttg ttctttcttt 2400
 tgtttcatat tcttgccctt tttaagacct tccaaatcat caacacaaaa acaccttaga 2460
 cattattgca tccgacttaa tttatatgac ataaaatata aagttttttt tttaaaataa 2520
 aaataaagag ttttaaatTT tataagaaag aaatatgtat atttaagttc atactaatta 2580
 taaaaacatg atttattaaa aataaatata aaaaatatat catataaatt aaaacgaatg 2640
 aaatatagta tattaggatg agaattcaaa ggactagaat ttccaaagaa atwtgwwgtt 2700
 ttttatgttt tattagtTga cttgttctat gccggtcttg ctcgttcacc atatatagac 2760
 caacacttaa agtaaatagt ggtaatttga tttgacctt caaaaccttc aagggccaat 2820
 tagacagaaa ctttatatta taaaagtaat aatgaatagt agtaataaac taattattac 2880
 caagaaatga tgaataattg tttatatata tatatagtgc aataataatt cataaccggc 2940
 cgatcctata tatttattat aaaatataga agttatccaa taatttcatg tcttttatta 3000
 aaaaaattca ttcaataatt tatgtttttt cttttacaaa gttaataat agaacttatt 3060
 taattaaaac taatgttact attatgatga aatagtcgag atctcttcat ttttcataag 3120
 agattcagga ttcaagtga ttatgtgagt gaaatcctaa aaaatgaaat ttgtaatacg 3180
 taaataatac gaatatcaaa tacgaaacaa aaagaaaatt aatgttatta tgattttgaa 3240
 ttttaattat ggtattttta ctactgaagt tgtctcattg aatgtatatt ataattaagc 3300
 atatacaatt ccagaaaata tatattatta tctagggtgtt cttttaatat aatatttgag 3360
 ttttatgata aaatttgatt aacttttaaa ggaaaatctt tttaaaagtg gtcacattag 3420
 attgagataa atcagaaagg acacgtcaca atctaaatat gattttttta acttcatcta 3480
 taaatatctt tcaaatata tacaatgctt acgtgaataa atgctggcac taatgaatta 3540
 atatttgttt ttaaaataaa atgttgaaat tgmddttaga agtgaacaca tgagtcgaag 3600
 tgttagaaat agttactaat aaataatcaa tgtatttgat ataaaaagtg ttattaagtt 3660

gttgttttgt tagaataact caaacattht ctttttattt aattaaaaat tataaattht 3720
 aactatatgt ataaacaaaa tgagcaataa taaatataga atgaagcaaa cattatgaat 3780
 tttattttga aaaaaaatga tgtgaattat aaaaatatta tataaattht gatcaaacaa 3840
 attgtttact ataagctaaa gaaatcttga ttttaactta taaatactth acttatcaaa 3900
 tacgtaaaaa aattaaaaat tacctataaa tcaatttaaat ctaacactth agtcaaacac 3960
 attataaata tgaacttggt gtgccacgat tatacacata tttttttttt tttttattga 4020
 tatatcctca aaaataatgt atctcttaaa atattagtat ttattaagat acgggttcat 4080
 gaaatattct ttttaactta aattttttatt tgactttgaa aacagacccc taaaaagttt 4140
 aaaaataaca aaaaggcaac aaaaaagttt gaaacaagaa ctaaaaaagt tagatgccca 4200
 aatctgcacg ataaggtaag tcaaacaacg acaaaattta gtaaattatt gattgtagaa 4260
 aatgaagcaa atgaaggtag ctatccttat ataaagctat attcagtact ctctcaaca 4320
 cagtcacagt agatttcatt tctttgtctt tgtcaaatcc taggtaattt ttctctgcc 4380
 cttaatattt ttttttgctc ttgattgctt tttatggaaa attggttctc tgttcatttg 4440
 acaagattgc tgttttgttt tttcttcagt ctgattagat ttttgctcta gattcgtggg 4500
 ttttacttgt aaaggattta gggagttagt gagatgggaa aatttcratt tttttttaa 4560
 aaaaaattat gggataaaaa gtgattccga gaaatgggtt tttgtttttt gagattgtat 4620
 gaactgggga tgatggattc ttggtggtga ttgaataaaa atgggatctt taattattag 4680
 ggataagatt aaagatgctc atgttcttct ttgaatttgt ttgtatgcaa agttttgagc 4740
 ttttggttgt tttcccccta atttagctgg caaatagttg taattggtgg aagttggaaa 4800
 aggaaaagaa ttaagaagaa attgatgttc atttcagttt ctttttatat tttgtaagga 4860
 acggaatcac cgggtgctgag tggaatagat ttagagggtt ctcaaatagt taaatgagta 4920
 gctggtttga tgaaatttgt tttgattgta taagatttca cttagttttg ctataaatca 4980
 tcgcagaatc aacctgatac 5000

<210> 3

<211> 347

<212> DNA

<213> *Lycopersicon esculentum*

<400> 3

cttgtcctt cgattactac atcatcgctc cgaatagggc ttctgatgaa gcagaatggt 60
 cttgtctaaa ccttccccctt tccattttgc tgaaactgct agaaaataag gccatctctt 120
 tcctttattg gctatgttct ttaaagtaaa aaaaaaac gttcaagtat ctttttagtgc 180
 ttttgtaatt ctactgtttc tacaacattt gaaatttata gagctgttgc ttcttgttca 240
 tgaaggtttt tgcttcttgt tccataattt tacgctacgt tgctcggact cttccaaaat 300

gtccactaat gcgtgttgga ttcttcaaaa atagtgtaac tttgaag

347

<210> 4

<211> 975

<212> DNA

<213> Lycopersicon esculentum

<400> 4

```

atggcgggggg aaaagactgg aattgccaag gatgtaactg aattgatagg taacactcct    60
ttggtatacc tgaataatgt tgtggatggg tgtgttgcac gtgttgctgc caagctggaa   120
agcatggagc catgctctag tgttaaggat aggattgggt atagtatgat tacagatgct   180
gaggagaagg gcttaatcaa acctggcgag agtgtcctca tcgaacctac aagtggaaac   240
accggtgtag gattggcatt catggctgct gctaaaggct acaaactcat cattacgatg   300
ccttcttcaa tgagtcttga gagaagaatt attctgcgtg ctttcggtgc tgagttggtg   360
cttactgacc cagcaaaagg gatgaaagg gcaatttcga aggccgaaga gataaggggc   420
aaaacaccta actcctatat tcttcagcaa tttgaaaacc ctgctaacc aaagatacac   480
tatgaaacca ctggctctga gatctggaaa ggctcaaacg ggaaagtgga tgctctagtc   540
tctggaattg gaacaggagg cacaataaca gggtcaggca agtatattgag agagcagaac   600
cccaacatca agctgtatgg cgtggaacca gttgaaagtg ctatcctttc tgggtggaaag   660
cctgggtccac ataagattca ggggattggg gctgggtttcg ttcctgggtg tttggaagtt   720
aaccttattg acgatgtagt tcagggtttca agtgatgaat ccatagaaat ggctaagctt   780
ctggcattga aggaaggatt gctagtggga atatcatctg gtgctgctgc tgctgcggca   840
attaaagttg ctaagcgtcc tgaaaatgct gggaagctca ttgttggtgt tttcccaagc   900
ttcggggagc gatatctttc ctctgtgctc ttcgaaactg tcagacgaga agcagagaac   960
atgactgtgg agcct

```

975

<210> 5

<211> 325

<212> PRT

<213> Lycopersicon esculentum

<400> 5

Met Ala Gly Glu Lys Thr Gly Ile Ala Lys Asp Val Thr Glu Leu Ile
1 5 10 15

Gly Asn Thr Pro Leu Val Tyr Leu Asn Asn Val Val Asp Gly Cys Val
20 25 30

Ala Arg Val Ala Ala Lys Leu Glu Ser Met Glu Pro Cys Ser Ser Val
35 40 45

Lys Asp Arg Ile Gly Tyr Ser Met Ile Thr Asp Ala Glu Glu Lys Gly
 50 55 60

Leu Ile Lys Pro Gly Glu Ser Val Leu Ile Glu Pro Thr Ser Gly Asn
 65 70 75 80

Thr Gly Val Gly Leu Ala Phe Met Ala Ala Ala Lys Gly Tyr Lys Leu
 85 90 95

Ile Ile Thr Met Pro Ser Ser Met Ser Leu Glu Arg Arg Ile Ile Leu
 100 105 110

Arg Ala Phe Gly Ala Glu Leu Val Leu Thr Asp Pro Ala Lys Gly Met
 115 120 125

Lys Gly Ala Ile Ser Lys Ala Glu Glu Ile Arg Gly Lys Thr Pro Asn
 130 135 140

Ser Tyr Ile Leu Gln Gln Phe Glu Asn Pro Ala Asn Pro Lys Ile His
 145 150 155 160

Tyr Glu Thr Thr Gly Pro Glu Ile Trp Lys Gly Ser Asn Gly Lys Val
 165 170 175

Asp Ala Leu Val Ser Gly Ile Gly Thr Gly Gly Thr Ile Thr Gly Ser
 180 185 190

Gly Lys Tyr Leu Arg Glu Gln Asn Pro Asn Ile Lys Leu Tyr Gly Val
 195 200 205

Glu Pro Val Glu Ser Ala Ile Leu Ser Gly Gly Lys Pro Gly Pro His
 210 215 220

Lys Ile Gln Gly Ile Gly Ala Gly Phe Val Pro Gly Val Leu Glu Val
 225 230 235 240

Asn Leu Ile Asp Asp Val Val Gln Val Ser Ser Asp Glu Ser Ile Glu
 245 250 255

Met Ala Lys Leu Leu Ala Leu Lys Glu Gly Leu Leu Val Gly Ile Ser
 260 265 270

Ser Gly Ala Ala Ala Ala Ala Ala Ile Lys Val Ala Lys Arg Pro Glu
 275 280 285

Asn Ala Gly Lys Leu Ile Val Val Val Phe Pro Ser Phe Gly Glu Arg
 290 295 300

Tyr Leu Ser Ser Val Leu Phe Glu Thr Val Arg Arg Glu Ala Glu Asn
 305 310 315 320

Met Thr Val Glu Pro
 325

<210> 6
 <211> 2987
 <212> DNA
 <213> Lycopersicon esculentum

<220>
 <221> misc_diferença
 <222> (679)..(679)
 <223> W é provavelmente T

<220>
 <221> misc_diferença
 <222> (682)..(682)
 <223> W é provavelmente A

<220>
 <221> misc_diferença
 <222> (683)..(683)
 <223> W é provavelmente T

<220>
 <221> misc_diferença
 <222> (691)..(691)
 <223> inserção de A parece confiável

<220>
 <221> misc_diferença
 <222> (864)..(864)
 <223> Y é provavelmente T

<220>
 <221> misc_diferença
 <222> (1269)..(1269)
 <223> R é provavelmente A

<220>
 <221> misc_diferença
 <222> (1559)..(1559)
 <223> R é provavelmente A

<220>
 <221> misc_diferença
 <222> (1560)..(1560)
 <223> M é provavelmente C

<220>
 <221> misc_diferença
 <222> (1561)..(1561)
 <223> D é provavelmente T

<220>
 <221> misc_diferença
 <222> (1916)..(1916)

<223> R é provavelmente A

<220>

<221> misc_diferença

<222> (2085)..(2085)

<223> Y é provavelmente T

<220>

<221> misc_diferença

<222> (2101)..(2101)

<223> Y é provavelmente C

<220>

<221> misc_diferença

<222> (2463)..(2463)

<223> Y é provavelmente T

<220>

<221> misc_diferença

<222> (2526)..(2526)

<223> A é provavelmente deletado

<220>

<221> misc_diferença

<222> (2534)..(2534)

<223> R é provavelmente G

<220>

<221> misc_diferença

<222> (2681)..(2681)

<223> K é provavelmente G

<220>

<221> misc_diferença

<222> (2920)..(2920)

<223> R é provavelmente A

<400> 6

tgggttttcta attctctcat ctcttttctc tggtcttgta tcttatccct ttctgattat 60

cttcttttgt gttaatctct ccttagttaa tatattatta ctttttgcac gtattccaat 120

agtaacttat attcaacaat aattatatac aatagtgttt gtattagaat cgagtgttatt 180

tcataaagtc tcgaaagtaa tacattactt tttttagaat aaaaaaaaaa cttgggaata 240

tggaagttca atctacgcaa aaacctatct tatgatgtca aatgaattca atgtttttac 300

tcataattgca atataaaaaag aacaggggcta gtgattttttt ttccttcaaa tgattgctag 360

ttaattaatg gacttggttct ttcttttggt tcatattctt gcctttttta agaccttcca 420

aatcatcaac acaaaaacac cttagacatt attgcatccg acttaattta tatgacataa 480

aatataaagt ttttttttta aaataaaaaat aaagagtttt aaattttata agaaagaaat 540

atgtatatatt aagttcatac taattataaa aacatgattt attaaaaata aatataaaaa 600

atatatcata taaattaaaa cgaatgaaat atagtatatt aggatgagaa ttcaaaggac 660

tagaatttcc aaagaaatwt gwwgtttttt aatgttttat tagttgactt gttctatgcc 720

ggtcttgctc gttcaccata tatagaccaa cacttaaagt aaatagtggt aatttgattt 780

gacctttcaa aaccttcaag ggccaattag acagaaactt tatattataa aagtaataat 840
 gaatagtagt aataaactaa ttaytaccaa gaaatgatga ataattgttt atatatatat 900
 atagtgcaat aataattcat aaccggccga tcctatatat ttattataaa atatagaagt 960
 tatccaataa tttcatgtct tttattaaaa aaattcattc aataatttat gttttttctt 1020
 ttacaaagtt taataataga acttatttaa ttaaaactaa tgttactatt atgatgaaat 1080
 agtcgagatc tcttcatttt tcataagaga ttcaggattc aagctgatta tgtgagtga 1140
 atcctaataa atgaaatttg taatacgtaa ataatacgaa tatcaaatac gaaacaaaaa 1200
 gaaaattaat gttattatga ttttgaattt taattatggg attttaacta ctgaagttgt 1260
 ctcatgart gtatattata attaagcata tacaattcca gaaaatatat attattatct 1320
 aggtgttctt ttaatataat atttgagttt tatgataaaa tttgattaac ttttaaagga 1380
 aaatcttttt aaaagtgggc acattagatt gagataaatc agaaaggaca cgtcacaatc 1440
 taaatatgat ttttttaact tcactataa atatctttca aattatatac aatgcttacg 1500
 tgaataaatg ctggcactaa tgaattaata tttgttttta aaataaaatg ttgaaattrm 1560
 dtttagaagt gaacacatga gtcgaagtgt tagaaatagt tactaataaa taatcaatgt 1620
 atttgatata aaaagtgtta ttaagttgtt gttttgttag aataactcaa acattttcat 1680
 tttatttaat taaaaattat aaatttaaac tatatgtata aacaaaatga gcaataataa 1740
 atatagaatg aagcaaacat tatgaatttt attttgaaaa aaaatgatgt gaattataaa 1800
 aatattatat aatatttgat caaacaatt gtttactata agctaaagaa atcttgattt 1860
 taacttataa atactttact tatcaaatac gtaaaaaaat taaaattac ctataratca 1920
 atttaatcta acactttagt caaacacatt ataaatatga acttgttgtg ccacgattat 1980
 acacatatat ttttttttat ttattgatat atcctcaaaa ataatgtatc tcttaaaata 2040
 ttagtattta ttaagatacg ggttcatgaa atattctttt aactycaaat ttttatttga 2100
 ytttgaaaac agaccctaa aaagttaaa aataacaaaa aggcaacaaa aaagtttgaa 2160
 acaagaacta aaaaagttag atgcccacaa ctgcacgata aggtaagtca aacaacgaca 2220
 aaatttagta aattattgat tgtagaaaat gaagcaaatg aaggtggcta tccttatata 2280
 aagctatatt cagtactctc ctcaacacag tcacagtaga tttcatttct ttgtctttgt 2340
 caaatcctag gtaatttttc tctgcccctt aatatttttt tttgctcttg attgcctttt 2400
 atggaaaatt ggttctctgt tcatttgaca agattgctgt tttgtttttt cttcagtctg 2460
 atyagatttt tgctctagat tcgtgggttt tacttgtaaa ggtattaggg agtgagtga 2520
 atgggaaaat ttcraatttt ttttaaaaaa aaattatggg ataaaaagtg attccgagaa 2580
 atggggtttt gttttttgag attgtatgaa ctggggatga tggattcttg gtggtgattg 2640

```

aataaaaatg ggatctttaa ttattagga taagattaa katgctcatg ttcttctttg 2700
aatttgtttg tatgcaaagt tttgagcttt tgggtgtttt cccctaatt tagctggcaa 2760
atagttgtaa ttggtggaag ttggaaaagg aaaagaatta agaagaaatt gatgttcatt 2820
tcagtttctt tttatatttt gtaaggaacg gaatcaccgg tgtcgagtgg aatagattta 2880
gagggttctc aaatagttaa atgagtagct ggtttgatgr aatttgtttt gattgtataa 2940
gatttcactt agttttgcta taaatcatcg cagaatcaac ctgatac 2987

```

```

<210> 7
<211> 5003
<212> DNA
<213> Lycopersicon esculentum

```

```

<220>
<221> misc_diferença
<222> (35)..(35)
<223> inserção de G parece confiável

```

```

<220>
<221> misc_diferença
<222> (240)..(240)
<223> Y é provavelmente T

```

```

<220>
<221> misc_diferença
<222> (768)..(769)
<223> WY é provavelmente AT

```

```

<220>
<221> misc_diferença
<222> (1558)..(1558)
<223> R é provavelmente A

```

```

<220>
<221> misc_diferença
<222> (1742)..(1742)
<223> Y é provavelmente T

```

```

<220>
<221> misc_diferença
<222> (2016)..(2016)
<223> inserção de T parece confiável

```

```

<220>
<221> misc_diferença
<222> (2695)..(2699)
<223> WTGWW é provavelmente TTGAT

```

```

<220>
<221> misc_diferença
<222> (2707)..(2707)
<223> inserção A parece confiável

```

```

<220>
<221> misc_diferença
<222> (2880)..(2880)
<223> Y é provavelmente T

```

<220>
 <221> misc_diferença
 <222> (3285)..(3285)
 <223> R é provavelmente A

<220>
 <221> misc_diferença
 <222> (3575)..(3577)
 <223> RMD é provavelmente ACT

<220>
 <221> misc_diferença
 <222> (3932)..(3932)
 <223> R é provavelmente A

<220>
 <221> misc_diferença
 <222> (4101)..(4101)
 <223> Y é provavelmente T

<220>
 <221> misc_diferença
 <222> (4117)..(4117)
 <223> Y é provavelmente C

<220>
 <221> misc_diferença
 <222> (4479)..(4479)
 <223> Y é provavelmente T

<220>
 <221> misc_diferença
 <222> (4542)..(4542)
 <223> A é provavelmente deletado

<220>
 <221> misc_diferença
 <222> (4550)..(4550)
 <223> R é provavelmente G

<220>
 <221> misc_diferença
 <222> (4697)..(4697)
 <223> K é provavelmente G

<220>
 <221> misc_diferença
 <222> (4936)..(4936)
 <223> R é provavelmente A

<400> 7
 catataaccc tgaaaattat ttaccattc ttggatctt ctttttcatt ttatttgacc 60
 aatttataag tattatttta aaagttaat tttttttcaa atttcacgat gaattatggc 120
 caaatttctc ctagttgtgt ttgaccata aataaaaata gaagttgttc ttactttttt 180
 tgtgagtaat ttgaagtgat ttttttcta agtttttgaa attcaacttt aattttttty 240
 tgtttttttg tgacaaggaa atctgtcttg ctatcttttg ggtgcataca cggtaaaatt 300
 cctgctcata tgcaataact cactctaact ttaagttgaa cttaacaatt ttacgattaa 360

agtggttttct aaactcaact tcatgaaaaa attaaaagtt attgaggaaa ttttaaaaat 420
 aactaacact gaataatata attaaactcc ttagctaattg tttctctaata tacgattcat 480
 agctactatt agaaggagag aagtgagcga caatatgaga ggaggagaaa tacgagcgag 540
 atctgagaga gatacaatca atttgtgtat ctgattgata cttgtattat tgttgttaatt 600
 cataataaat tgtatttaata tatatttatc tattgccttg tatattgtat ttatgtgtta 660
 ccatgtatat ttgtgctacc attttgtatt tgtatcattg actgctatcg ttgtacgtgt 720
 attgtagaaa gagagagagg agaggaatag tgagaagatt gctataawyc caaattataa 780
 ttacgaatag taattctttt aaattataat tatgtatagt taagacatat tttgtacgtt 840
 tatctacatt tgccccaaaa ttattctcct gtggactctc aaagcaggaa ccaaatagtg 900
 atggattggg ctctagttc tgaacgaatt ttgttattgg gcttaaagtt aaggttgggt 960
 ggtcagggtt ttgggtttgc atatctatgg ggtagtaagg tttattcttt tctcgagatt 1020
 cgatatttat atttgagctt tattaattt ggattcacac caaaaagttt tatattgagg 1080
 gggggggggg gtaaagtgtc tcctaataaa ggcaaaaacg attccctact tagaagattc 1140
 aaactcaaaa aactcgggat gaagtattta caattctatc acgttggtga ttagttagct 1200
 ttatcattgt atgaggtata cttgtgggat agaattgtta gacaatttat gaatttcttt 1260
 tttgaatcta gaatgtgggt tgatgaacaa tggacttaaa agaagttaat ggaaaatgag 1320
 aaatcaaatt tcaataaaaa tgaagttttt agtgattctt ttcattctatc gataggtaaa 1380
 agatatatca tttgggtctaa ttcaatocca aaagttagtt taaagggaaa agcgggttat 1440
 aaagagtaca ttcgtctcat taaccaccaa tgtgaacttt tattattctt caacacccat 1500
 cttatttcta gtttttttag catttggtgt gtgaacaatt ttttatttcg gaggtccrac 1560
 atcgggtgag aagagttcta ctctgatatc atgtgaaatt aggttttgag ctaaaacttac 1620
 accaaaaagc tagcttaaaa gatgaaggat tgtacaagcc ttataaggaa ttcacccgctc 1680
 ttattaagca tcaatataaa acttttgtca ttcttaacaa atataacttc gataaaacca 1740
 tycaaataaa tatagtatac acatgtacaa agatgtatct actcaatgaa atatatgagc 1800
 gcaaatgaga tactatgact accatcatcg taaataaaaa aacacataaa gtatatattt 1860
 aaaggttggg gtttataaag tgtcatgaac caaactgatc ccgtaaggac atatgggcta 1920
 ttagagagtt ttattagga tgcaagtcta ttcactcttc gtgaaaaagg actaggagtt 1980
 agaagagtga ttatgcaaat tattatgtcg aagcttgggt tttctaattc tctcatctct 2040
 tttctctgtt cttgtatctt atcccttctt gattatcttc ttttgtgtta atctctcctt 2100
 agttaatata ttattacttt ttgcatgtat tccaatagta acttatattc aacaataatt 2160
 atatacaata gtgtttgtat tagaatcgag tttatttcat aaagtctcga aagtaataca 2220

ttactttttt tagaataaaa aaaaaacttg ggaatatgga agttcaatct acgcaaaaac 2280
 ctatttttatg atgtcaaagtg aattcaatgt ttttactcat attgcaatat aaaaagaaca 2340
 gggctagtga ttttttttcc ttcaaagtgat tgctagttaa ttaatggact tgttctttct 2400
 tttgtttcat attcttgcct tttttaagac cttccaaatc atcaacacaa aaacacctta 2460
 gacattattg catccgactt aatttatatg acataaaaata taaagttttt tttttaaaat 2520
 aaaaataaag agtttttaaat tttataagaa agaaatatgt atatttaagt tcataactaat 2580
 tataaaaaca tgatttatta aaaataaata taaaaaatat atcatataaa ttaaaacgaa 2640
 tgaaatatag tatattagga tgagaattca aaggactaga atttccaaag aaatwtgwwg 2700
 ttttttaatg ttttattagt tgacttggtc tatgccggtc ttgctcgttc accatatata 2760
 gaccaacact taaagtaaat agtggttaatt tgatttgacc tttcaaaacc ttcaagggcc 2820
 aattagacag aaactttata ttataaaagt aataatgaat agtagtaata aactaattay 2880
 taccaagaaa tgatgaataa ttgtttatat atatatatag tgcaataata attcataacc 2940
 ggccgatcct atatatttat tataaaatat agaagttatc caataatttc atgtctttta 3000
 ttaaaaaaat tcattcaata atttatgttt tttcttttac aaagttaat aatagaactt 3060
 atttaattaa aactaatgtt actattatga tgaaatagtc gagatctctt cttttttcat 3120
 aagagattca ggattcaagc tgattatgtg agtgaaatcc taaaaaatga aatttgtaat 3180
 acgtaaataa tacgaatatc aaatacgaaa caaaaagaaa attaatgtta ttatgatttt 3240
 gaattttaat tatggtattt taactactga agttgtctca ttgartgtat attataatta 3300
 agcatataca attccagaaa atatataatta ttatctaggt gttcttttaa tataatatat 3360
 gagttttatg ataaaatttg attaaactttt aaaggaaaat ctttttaaaa gtggtcacat 3420
 tagattgaga taaatcagaa aggacacgtc acaatctaaa tatgattttt ttaacttcat 3480
 ctataaatat ctttcaaatt atatacaatg cttacgtgaa taaatgctgg cactaatgaa 3540
 ttaatatttg tttttaaaat aaaatgttga aattrmdttt agaagtgaac acatgagtcg 3600
 aagtgttaga aatagttact aataaataat caatgtattt gatataaaaa gtgttattaa 3660
 gttgttggtt tgttagaata actcaaacat tttcatttta ttaattaaa aattataaat 3720
 ttaaaactata tgtataaaca aaatgagcaa taataaatat agaatgaagc aaacattatg 3780
 aattttattt tgaaaaaaa tgatgtgaat tataaaaata ttatataata tttgatcaaa 3840
 caaattgttt actataagct aaagaaatct tgattttaac ttataaatac tttacttatac 3900
 aaatacgtaa aaaaattaaa aattacctat aratcaattt aatctaacac tttagtcaaa 3960
 cacattataa atatgaactt gttgtgccac gattatacac atatattttt ttttatttat 4020
 tgatatatcc tcaaaaataa tgtatctctt aaaatattag tatttattaa gatacgggtt 4080
 catgaaatat tcttttaact ycaaattttt atttgayttt gaaaacagac ccctaaaaag 4140

tttaaaaata acaaaaaggc aacaaaaaag tttgaaacaa gaactaaaaa agttagatgc 4200
 ccaaactctgc acgataaggt aagtcaaaca acgacaaaat ttagtaaatt attgattgta 4260
 gaaaatgaag caaatgaagg tggctatcct tatataaagc tatattcagt actctcctca 4320
 acacagtcac agtagatttc atttctttgt ctttgtcaaa tcctaggtaa tttttctctg 4380
 ccccttaata tttttttttg ctcttgattg ccttttatgg aaaattgggt ctctgttcat 4440
 ttgacaagat tgctgttttg ttttttcttc agtctgatya gatttttgct ctagattcgt 4500
 gggttttact tgtaaaggta ttagggagtg agtgagatgg gaaaatttcr attttttttt 4560
 aaaaaaaaaat tatgggataa aaagtgattc cgagaaatgg ggttttggtt tttgagattg 4620
 tatgaactgg ggatgatgga ttcttggtgg tgattgaata aaaatgggat ctttaattat 4680
 tagggataag attaaakatg ctcatgttct tctttgaatt tgtttgatg caaagttttg 4740
 agcttttggg tgttttcccc ctaatttagc tggcaaatag ttgtaattgg tggaagttgg 4800
 aaaaggaaaa gaattaagaa gaaattgatg ttcatttcag tttcttttta tattttgtaa 4860
 ggaacggaat caccggtgtc gagtggaata gatttagagg gttctcaa atgttaa atga 4920
 gtagctggtt tgatgraatt tgttttgatt gtataagatt tcacttagtt ttgctataaa 4980
 tcatcgaga atcaacctga tac 5003

<210> 8

<211> 2986

<212> DNA

<213> *Lycopersicon esculentum*

<400> 8

tggtttttcta attctctcat ctcttttctc tgttcttgta tcttatccct ttctgattat 60
 ctctttttgt gttaatctct ccttagttaa tatattatta ctttttgcat gtattccaat 120
 agtaacttat attcaacaat aattatatac aatagtgttt gtattagaat cgagtttatt 180
 tcataaagtc tcgaaagtaa tacattactt tttttagaat aaaaaaaaaa cttgggaata 240
 tggaagttca atctacgcaa aaacctat ttatgatgtca aatgaattca atgtttttac 300
 tcatattgca atataaaaag aacagggtca gtgatttttt ttccttcaaa tgattgctag 360
 ttaattaatg gacttggtct ttcttttggt tcatattctt gcctttttta agaccttcca 420
 aatcatcaac aaaaaaacac cttagacatt attgcatccg acttaattta tatgacataa 480
 aatataaagt ttttttttta aaataaaaat aaagagtttt aaattttata agaaagaaat 540
 atgtatat tt aagttcatat taattataaa aacatgattt attaaaaata aatataaaaa 600
 atatatcata taaattaaaa cgaatgaaat atagtatatt aggatgagaa ttcaaaggac 660
 tagaatttcc aaagaaattt gatgtttttt aatgttttat tagttgactt gttctatgcc 720
 ggtcttgctc gttcaccata tatagaccaa cacttaaagt aaatagtggg aatttgattt 780

gacctttcaa aaccttcaag ggccaattag acagaaactt tatattataa aagtaataat 840
gaatagtagt aataaactaa ttattaccaa gaaatgatga ataattgttt atatatatat 900
atagtgcaat aataattcat aaccggccga tcctatatat ttattataaa atatagaagt 960
tatccaataa tttcatgtct tttattaaaa aaattcattc aataatttat gttttttctt 1020
ttacaaagtt taataataga acttatttaa ttaaaactaa tgttactatt atgatgaaat 1080
agtcgagatc tcttcatttt tcataagaga ttcaggattc aagctgatta tgtgagtga 1140
atcctaaaaa atgaaatttg taatacgtaa ataatacgaa tatcaaatac gaaacaaaaa 1200
gaaaattaat gttattatga ttttgaattt taattatggg attttaacta ctgaagttgt 1260
ctcattgaat gtatattata attaagcata tacaattcca gaaaatatat attattatct 1320
aggtgttctt ttaatatat atttgagttt tatgataaaa tttgattaac ttttaaagga 1380
aaatcttttt aaaagtgggc acattagatt gagataaatc agaaaggaca cgtcacatc 1440
taaatatgat ttttttaact tcactataa atatctttca aattatatac aatgcttacg 1500
tgaataaatg ctggcactaa tgaattaata tttgttttta aaataaaatg ttgaaattac 1560
ttttagaagt gaacacatga gtcgaagtgt tagaaatagt tactaataaa taatcaatgt 1620
atttgatata aaaagtgtta ttaagttggt gttttgttag aataactcaa acattttcat 1680
tttatttaat taaaaattat aaatttaaac tatatgtata aacaaaatga gcaataataa 1740
atatagaatg aagcaaacat tatgaatttt attttgaaaa aaaatgatgt gaattataaa 1800
aatattatat aatatttgat caaacaaatt gtttactata agctaaagaa atcttgattt 1860
taacttataa atactttact tatcaaatac gtaaaaaaat taaaaattac ctatagatca 1920
atttaatcta acactttagt caaacacatt ataaatatga acttgttgtg ccacgattat 1980
acacatatat ttttttttat ttattgatat atcctcaaaa ataatgtatc tcttaaaata 2040
ttagtattta ttaagatacg ggttcatgaa atattctttt aacttcaaat ttttatttga 2100
ctttgaaaac agaccctaa aaagtttaaa aataacaaaa aggcaacaaa aaagtttgaa 2160
acaagaacta aaaaagttag atgccccaat ctgcacgata aggtaagtca aacaacgaca 2220
aaatttagta aattattgat tgtagaaaat gaagcaaagtg aaggtggcta tccttatata 2280
aagctatatt cagtactctc ctcaacacag tcacagtaga tttcatttct ttgtctttgt 2340
caaatcctag gtaatttttc tctgcccctt aatatttttt tttgctcttg attgcctttt 2400
atggaaaatt ggttctctgt tcatttgaca agattgctgt tttgtttttt cttcagtctg 2460
atcagatttt tgctctagat tcgtgggttt tacttgtaaa ggtattaggg agtgagtga 2520
atgggaaatt tcgatttttt tttaaaaaaa aattatggga taaaaagtga ttccgagaaa 2580
tggggttttg ttttttgaga ttgtatgaac tggggatgat ggattcttgg tgggtgattga 2640

ataaaaaatgg gatctttaat tattagggat aagattaaat atgctcatgt tcttctttga 2700
 atttgtttgt atgcaaagtt ttgagctttt ggggtgttttc cccctaattt agctggcaaa 2760
 tagttgtaat tgggtggaagt tggaaaagga aaagaattaa gaagaaattg atgttcattt 2820
 cagtttcttt ttatatatttg taaggaacgg aatcacccgg gtcgagtgga atagatttag 2880
 agggttctca aatagttaaa tgagtagctg gtttgatgga atttgttttg attgtataag 2940
 atttcactta gttttgctat aaatcatcgc agaatcaacc tgatac 2986

<210> 9

<211> 5002

<212> DNA

<213> *Lycopersicon esculentum*

<400> 9

catataaccc tgaaaattat ttaccattc tttggatctt ctttttcatt ttatttgacc 60
 aatttataag tattatttta aaagtttaat tttttttcaa atttcacgat gaattatggc 120
 caaatttctc ctagtttgtt tttgaccata aataaaaata gaagttgttc ttactttttt 180
 tgtgagtaat ttgaagtgat ttttttccta agtttttgaa attcaacttt aatttttttt 240
 tgtttttttg tgacaaggaa atctgtcttg ctatcttttg ggtgcataca cggtaaaatt 300
 cctgctcata tgcaataact cactctaact ttaagttgaa cttacaatt ttacgattaa 360
 agtgttttct aaactcaact tcatgaaaaa attaaaagtt attgaggaaa ttttaaaaat 420
 aactaacact gaataatata attaaactcc ttagctaatg tttctctaatt tacgattcat 480
 agctactatt agaaggagag aagtgagcga caatatgaga ggaggagaaa tacgagcgag 540
 atctgagaga gatacaatca atttgtgtat ctgattgata cttgtattat tgttgtaatt 600
 cataataaat tgtatttaatt tatattttatc tattgccttg tatattgtat ttatgtgtta 660
 ccatgtatat ttgtgctacc attttgtatt tgtatcattg actgctatcg ttgtacgtgt 720
 attgtagaaa gagagagagg agaggaatag tgagaagatt gctataaatc caaattataa 780
 ttacgaatag taattctttt aaattataat tatgtatagt taagacatat tttgtacgtt 840
 tatctacatt tgccccaaaa ttattctcct gtggactctc aaagcaggaa ccaaatagtg 900
 atggattggg ctccatagttc tgaacgaatt ttgttattgg gcttaaagtt aaggttgggt 960
 ggtcagggtt ttgggtttgc atatctatgg ggtagtaagg tttattcttt tctcgagatt 1020
 cgatatttat atttgagctt tattaatttt ggattcacac caaaaagttt tatattgagg 1080
 gggggggggg gtaaagtgtc tccataataa ggcaaaaacg attccctact tagaagattc 1140
 aaactcaaaa aactcgggat gaagtattta caattctatc acgttggtga ttagtttagct 1200
 ttatcattgt atgaggtata cttgtgggat agaagttaa gacaatttat gaatttcttt 1260
 tttgaatcta gaatgtgggt tgatgaacaa tggacttaaa agaagttaat ggaaaatgag 1320

aaatcaaatt tcaataaaaa tgaagttttt agtgattctt ttcattctatc gataggtaaa 1380
 agatatatca tttggtctaa ttcaatccca aaagttagtt taaagggaaa agcgggttat 1440
 aaagagtaca ttcgctctcat taaccaccaa tgtgaacttt tattattctt caacacccat 1500
 cttattccta gtttttttag catttggtgt gtgaacaatt ttttatttcg gaggtccaac 1560
 atcgggtgag aagagttcta ctctgatatc atgtgaaatt aggttttgag ctaaacttac 1620
 accaaaaagc tagcttaaaa gatgaaggat tgtacaagcc ttataaggaa ttcacccgtc 1680
 ttattaagca tcaatataaa acttttgtca ttcttaaaaca atataacttc gataaaacca 1740
 ttcaaataaa tatagtatac acatgtacaa agatgtatct actcaatgaa atatatgagc 1800
 gcaaatgaga tactatgact accatcatcg taaataaaaa aacacataaa gtatatattt 1860
 aaaggttggt gtttataaag tgtcatgaac caaactgac ccgtaaggac atatgggcta 1920
 ttagagagtt ttattaggga tgcaagtcta ttcactcttc gtgaaaaagg actaggagtt 1980
 agaagagtga ttatgcaaat tattatgtcg aagctttggt tttctaattc tctcatctct 2040
 tttctctggt cttgtatctt atccctttct gattatcttc ttttgtgtta atctctcctt 2100
 agttaatata ttattacttt ttgcatgtat tccaatagta acttatattc aacaataatt 2160
 atatacaata gtgtttgtat tagaatcgag tttatttcat aaagtctcga aagtaataca 2220
 ttactttttt tagaataaaa aaaaaacttg ggaatatgga agttcaatct acgcaaaaac 2280
 ctatttttatg atgtcaaatg aattcaatgt ttttactcat attgcaatat aaaaagaaca 2340
 gggctagtga ttttttttcc ttcaaatgat tgctagttaa ttaatggact tgttctttct 2400
 tttgtttcat attcttgcct tttttaagac cttccaaatc atcaacacaa aaacacctta 2460
 gacattattg catccgactt aatttatatg acataaaata taaagttttt tttttaaaat 2520
 aaaaataaag agttttaaat tttataagaa agaaatatgt atatttaagt tcataactaat 2580
 tataaaaaaca tgatttatta aaaataaata taaaaaatat atcatataaa ttaaaacgaa 2640
 tgaaatatag tatattagga tgagaattca aaggactaga atttccaaag aaatttgatg 2700
 tttttaatg ttttattagt tgacttgttc tatgccggtc ttgctcgttc accatatata 2760
 gaccaacact taaagtaaatt agtggttaatt tgatttgacc tttcaaaacc ttcaagggcc 2820
 aattagacag aaactttata ttataaaagt aataatgaat agtagtaata aactaattac 2880
 taccaagaaa tgatgaataa ttgtttatat atatatatag tgcaataata attcataacc 2940
 ggccgatcct atatatattat tataaaatat agaagttatc caataatttc atgtctttta 3000
 ttaaaaaaat tcattcaata atttatgttt tttcttttac aaagtttaaat aatagaactt 3060
 atttaattaa aactaatgtt actattatga tgaaatagtc gagatctctt catttttcat 3120
 aagagattca ggattcaagc tgattatgtg agtgaaatcc taaaaaatga aatttgtaat 3180
 acgtaaataa tacgaatatc aaatacgaaa caaaaagaaa attaatgtta ttatgatttt 3240

gaattttaat tatggtattt taactactga agttgtctca ttgagtgtat attataatta 3300
 agcatataca attccagaaa atatatatta ttatctaggt gttcttttaa tataatattt 3360
 gagttttatg ataaaatttg attaactttt aaaggaaaat ctttttaaaa gtggtcacat 3420
 tagattgaga taaatcagaa aggacacgtc acaatctaaa tatgattttt ttaacttcat 3480
 ctataaatat ctttcaaatt atatacaatg cttacgtgaa taaatgctgg cactaatgaa 3540
 ttaatatattg tttttaaaat aaaatggtga aattactttt agaagtgaac acatgagtcg 3600
 aagtgttaga aatagttact aataaataat caatgtattt gatataaaaa gtgttattaa 3660
 gttgttggtt tgttagaata actcaaacat tttcatttta ttttaattaaa aattataaat 3720
 ttaaactata tgtataaaca aaatgagcaa taataaatat agaatgaagc aaacattatg 3780
 aattttattt tgaaaaaaaa tgatgtgaat tataaaaaata ttatataata ttgatcaaa 3840
 caaattggtt actataagct aaagaaatct tgattttaac ttataaatac tttacttctc 3900
 aaatacgtaa aaaaattaaa aattacctat aaatcaattt aatctaacac tttagtcaaa 3960
 cacattataa atatgaactt gttgtgccac gattatacac atatattttt ttttatttat 4020
 tgatatatcc tcaaaaataa tgtatctctt aaaatattag tatttattaa gatacggggt 4080
 catgaaatat tcttttaact ccaaattttt atttgacttt gaaaacagac ccctaaaaag 4140
 tttaaaaata acaaaaaggc acaaaaaaag tttgaaacaa gaactaaaaa agttagatgc 4200
 ccaaactctgc acgataaggt aagtcaaaca acgacaaaat ttagtaaatt attgattgta 4260
 gaaaatgaag caaatgaagg tggctatcct tatataaagc tatattcagt actctcctca 4320
 acacagtcac agtagatttc atttctttgt ctttgtcaaa tcctaggtaa tttttctctg 4380
 ccccttaata tttttttttg ctcttgattg ctttttatgg aaaattgggt ctctgttcat 4440
 ttgacaagat tgctgttttg tttttcttc agtctgatta gatttttgct ctagattcgt 4500
 gggttttact tgtaaaggta ttagggagtg agtgagatgg gaaatttcga ttttttttta 4560
 aaaaaaaatt atgggataaa aagtgattcc gagaaatggg gttttgtttt ttgagattgt 4620
 atgaactggg gatgatggat tcttggtggt gattgaataa aaatgggatc ttttaattatt 4680
 aggataaga ttaaagatgc tcatgttctt ctttgaattt gtttgatgc aaagttttga 4740
 gcttttgggt gttttcccc taatttagct ggcaaatagt tgtaattggt ggaagtggga 4800
 aaaggaaaag aattaagaag aaattgatgt tcatttcagt ttctttttat attttgtaag 4860
 gaacggaatc accggtgtcg agtggaatag atttagaggg ttctcaaata gttaaagag 4920
 tagctggttt gatgaaattt gttttgattg tataagattt cacttagttt tgctataaat 4980
 catcgagaa tcaacctgat ac

TRATADO DE BUDAPESTE SOBRE O RECONHECIMENTO
INTERNACIONAL DO DEPÓSITO DE MICROORGANISMOS
PARA OS FINS DE PROCEDIMENTO DE PATENTE

FORMULÁRIO INTERNACIONAL

Keygene N.V.
Postbus 216
6700 AE Wageningen
Países Baixos

nome e endereço do depositante

RECIBO NO CASO DE UM DEPÓSITO
ORIGINAL expedido de acordo com a
Regra 7.1 pela AUTORIDADE
DEPOSITÁRIA INTERNACIONAL
identificada no final desta página

I. IDENTIFICAÇÃO DO MICROORGANISMO	
Referência de identificação dada pelo DEPOSITANTE: E. coli DH10B + pKG8137	Número de acesso dado pela AUTORIDADE DE DEPÓSITO INTERNACIONAL: CBS 120176
II. DESCRIÇÃO CIENTÍFICA E/OU DESIGNAÇÃO TAXONÔMICA PROPOSTA	
O microorganismo especificado em I acima foi acompanhado por: ☐ uma descrição científica ☐ uma designação taxonômica proposta <i>(marcar com uma cruz onde aplicável)</i>	
III. RECIBO E ACEITAÇÃO	
Esta Autoridade Internacional aceita o microorganismo identificado em I acima, que foi recebido pela mesma em 03-08-2006 (data dia-mês-ano do depósito original) ¹	
IV. RECIBO DE SOLICITAÇÃO PARA CONVERSÃO	
O microorganismo identificado em I acima foi recebido por esta Autoridade Depositária Internacional em <i>não aplicável</i> (data dia-mês-ano do depósito original) e uma solicitação para converter o depósito original em um depósito sob o Tratado de Budapeste foi recebido pela mesma em <i>não aplicável</i> (data dia-mês-ano da solicitação para conversão)	
V. AUTORIDADE DEPOSITÁRIA INTERNACIONAL	
Nome: Centraalbureau voor Schimmelcultures Endereço: Uppsalalaan 8 P.O. Box 85167 3508 AD UTRACHT Países Baixos	Assinaturas da(s) pessoa(s) competente(s) para representar a autoridade de depósito internacional ou do(s) funcionário(s) autorizado(s) Mrs. F.B. Snippe-Claus Dr. J.A. Stalpers Data (dia-mês-ano) 14-08-2006

¹ Onde se aplica a regra 6.4 (d), tal data é a data na qual o status da Autoridade Depositária Internacional foi adquirido.

BP/A/II/12
Página 24

TRATADO DE BUDAPESTE SOBRE O RECONHECIMENTO
INTERNACIONAL DO DEPÓSITO DE MICROORGANISMOS
PARA OS FINS DE PROCEDIMENTO DE PATENTE

FORMULÁRIO INTERNACIONAL

Keygene N.V.
Postbus 216
6700 AE Wageningen
Países Baixos

*nome e endereço da parte para a
qual a declaração de viabilidade
é expedida*

DECLARAÇÃO DE VIABILIDADE
Expedida de acordo com a Regra 10.2 pela
AUTORIDADE DEPOSITÁRIA
INTERNACIONAL identificada na página
seguinte

II. DEPOSITANTE	II. IDENTIFICAÇÃO DO MICROORGANISMO
Keygene N.V.	Número de acesso dado pela AUTORIDADE DE DEPÓSITO INTERNACIONAL: CBS 120176
Endereço: Postbus 216 6700 AE Wageningen Países Baixos	Data (dia-mês-ano) do depósito ou da transferência: ¹ 03-082006
III. DECLARAÇÃO DE VIABILIDADE	
A viabilidade do microorganismo identificado em II acima foi testada em 08-08-2006 ² . Nessa data (dia-mês-ano), o referido microorganismo era ☐ ³ viável ☐ ³ não mais viável	

¹ Indica a data do depósito original ou, quando um novo depósito ou uma transferência foi feita, a data relevante mais recente (data do novo depósito ou data da transferência).

² Nos casos referidos na Regra 10.2 (a) (ii) e (iii), referir-se ao teste de viabilidade mais recente.

³ Marcar com uma cruz a caixa aplicável.

IV. CONDIÇÕES NAS QUAIS A VIABILIDADE FOI REALIZADA ⁴	
V. AUTORIDADE DEPOSITÁRIA INTERNACIONAL	
Nome: Centraalbureau voor Schimmelcultures	Assinaturas da(s) pessoa(s) competente(s) para representar a autoridade de depósito internacional ou do(s) funcionário(s) autorizado(s)
Endereço: Uppsalalaan 8 P.O Box 85167 3508 AD UTRECHT Países Baixos	Mrs. F.B. Snippe-Claus Dr. J.A. Stalpers
	Data (dia-mês-ano) 14-08-2006

⁴ Preencher se a informação foi solicitada e se os resultados do teste foram negativos.

REIVINDICAÇÕES

1. Gene quimérico, caracterizado pelo fato de compreender uma seqüência de ácido nucleico tendo atividade de promotor quando introduzida em células de planta, em que a referida seqüência de ácido
5 nucléico compreende uma seqüência selecionada dentre:

(a) a seqüência de ácido nucleico de SEQ ID NO: 1 ou SEQ ID NO: 2;

(b) um fragmento de SEQ ID NO: 1 ou SEQ ID NO: 2;

(c) a seqüência de ácido nucleico compreendendo pelo menos
10 70% identidade de seqüência com SEQ ID NO: 1 ou 2;

(d) um fragmento da seqüência de ácido nucleico de (c)
e um fragmento funcional, compreendendo pelo menos 200
nucleotídeos contíguos de SEQ ID NO: 1 ou SEQ ID NO: 2, ligado
operativamente a uma seqüência homóloga ou heteróloga de ácido nucleico e
15 opcionalmente uma seqüência 3' UTR.

2. Gene quimérico de acordo com a reivindicação 1, caracterizado pelo fato de que referida seqüência 3'UTR é SEQ ID NO: 3, ou um fragmento de SEQ ID NO: 3, ou uma seqüência de ácido nucleico compreendendo pelo menos 70% identidade de ácido nucleico para SEQ ID
20 NO: 3.

3. Vetor, caracterizado pelo fato de compreender um ácido nucleico compreendendo uma seqüência selecionada dentre:

(a) a seqüência de ácido nucleico de SEQ ID NO: 1 ou SEQ ID NO: 2;

(b) um fragmento de SEQ ID NO: 1 ou SEQ ID NO: 2;

(c) a seqüência de ácido nucleico compreendendo pelo menos
70% identidade de seqüência com SEQ ID NO: 1 ou 2;

(d) um fragmento da seqüência de ácido nucleico de (c)
e um fragmento funcional, compreendendo pelo menos 200

nucleotídeos contíguos de SEQ ID NO: 1 ou SEQ ID NO: 2, ou um gene quimérico como definido na reivindicação 1 ou 2.

4. Uso do promotor de um gene de cisteína sintase citossólica de planta, caracterizado pelo fato de ser para a expressão constitutiva de uma
5 sequência de ácido nucleico sentido e/ou anti-sentido em uma planta transgênica ou célula de planta ou tecido de planta ou órgão.

5. Uso de acordo com a reivindicação 4, caracterizado pelo fato de que o gene de cisteína sintase citossólica de planta codifica a proteína de SEQ ID NO: 5.

10 6. Método para fabricar uma planta transgênica ou célula de planta, caracterizado pelo fato de compreender as etapas de:

(a) gerar um gene quimérico como definido na reivindicação 1 ou 2 ou um vetor como definido na reivindicação 3;

(b) transformar uma planta ou célula de planta com referido
15 gene quimérico ou vetor; e, opcionalmente,

(c) regenerar as plantas transgênicas.

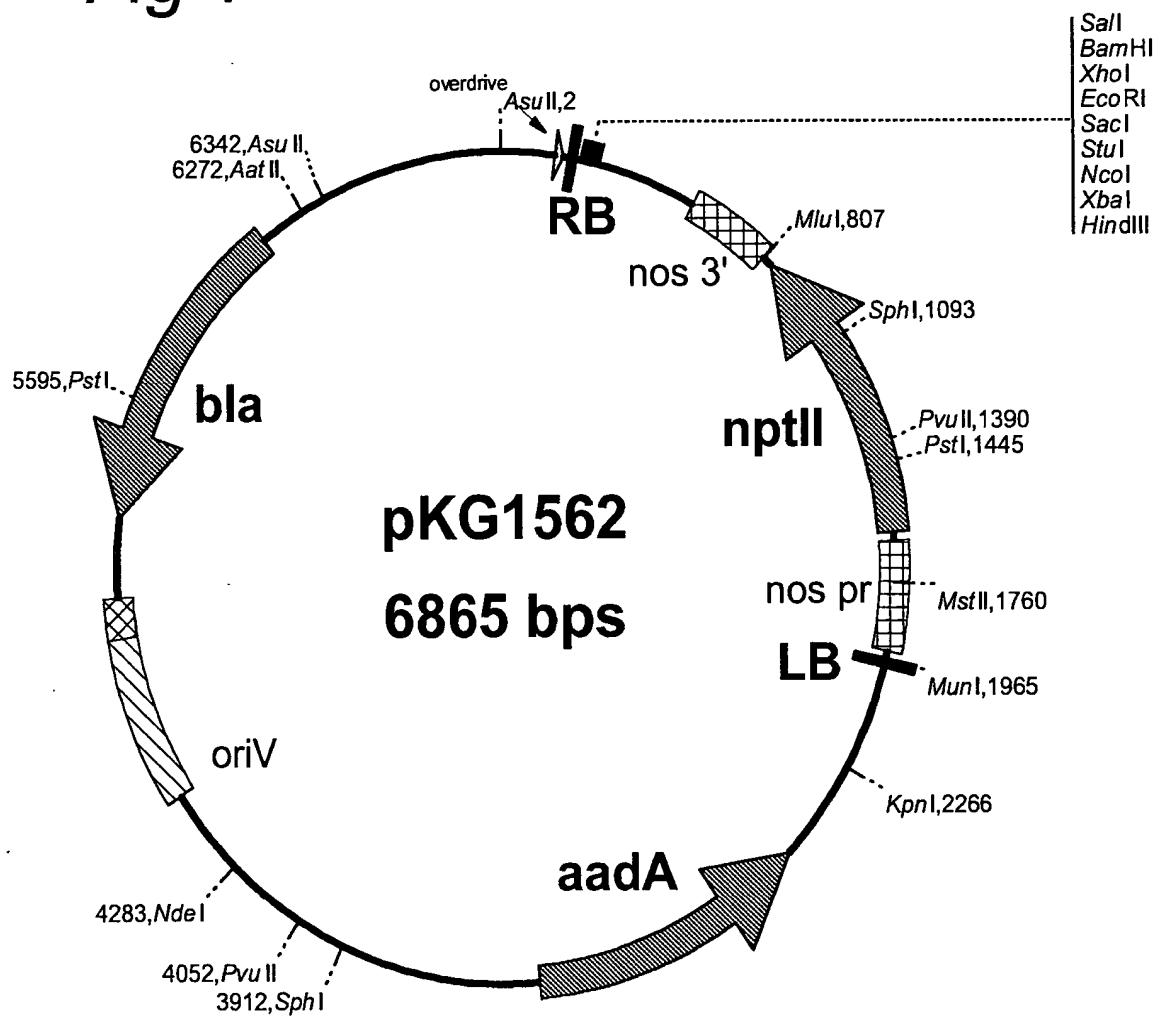
Fig 1

Fig 2

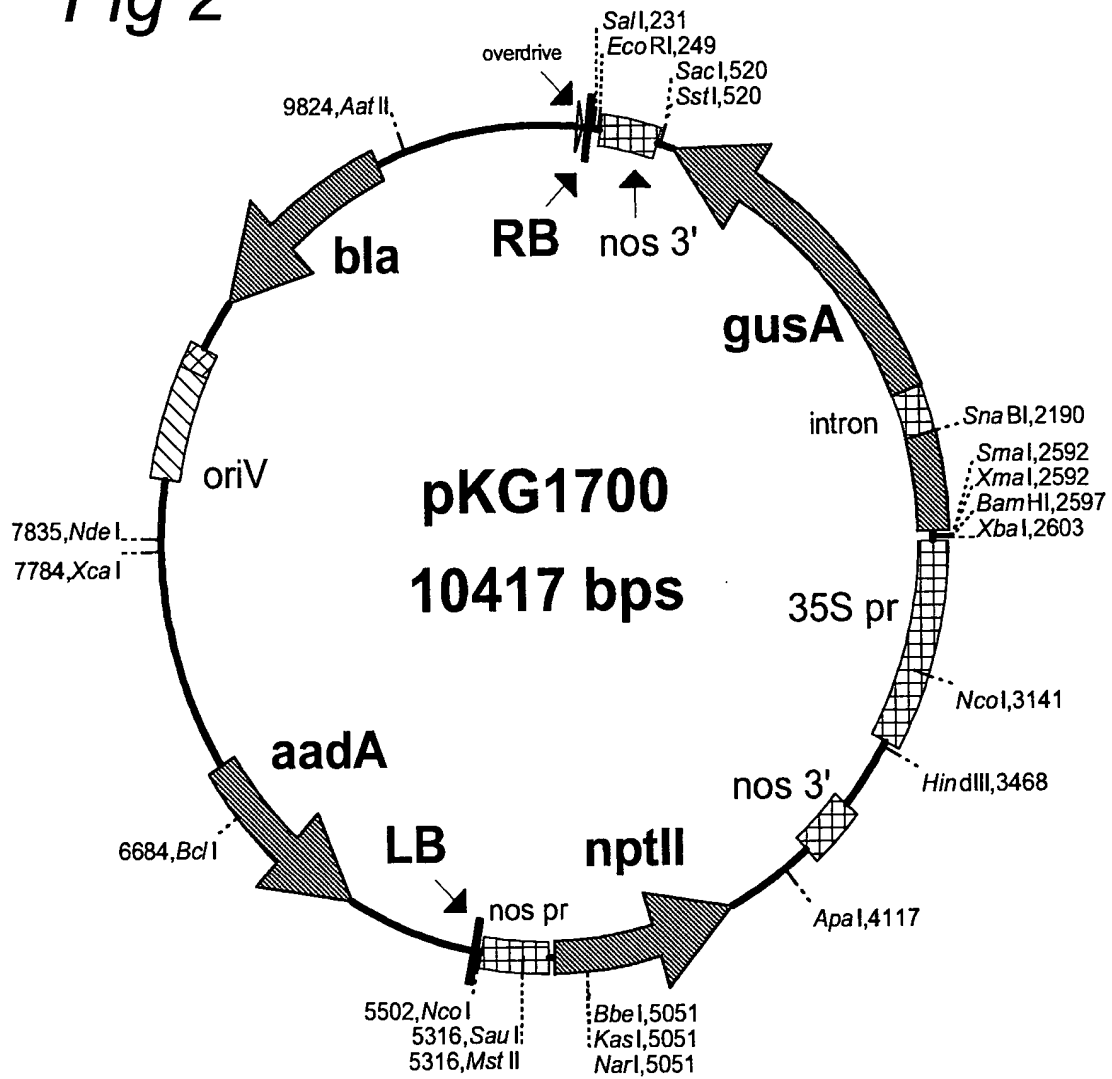
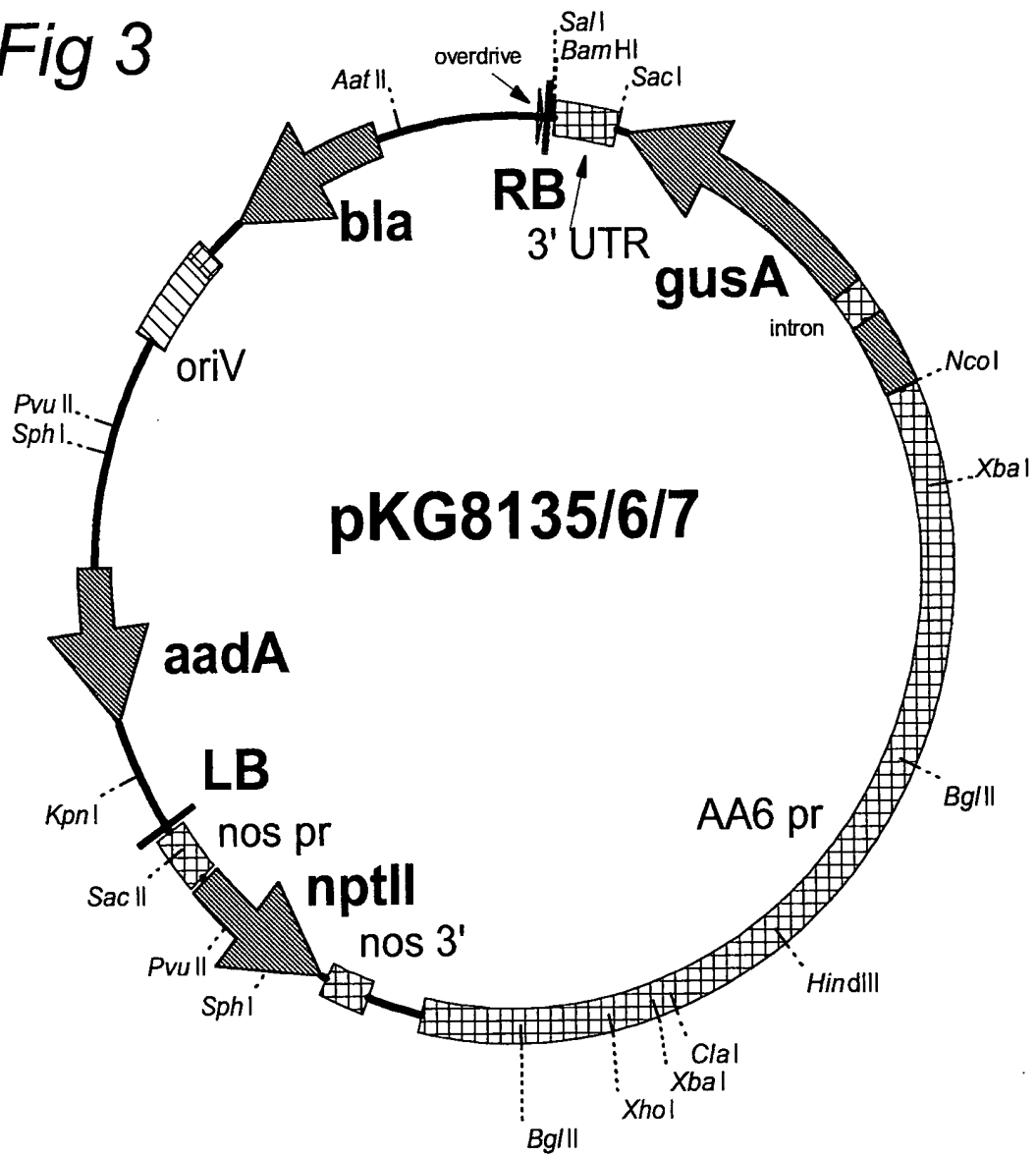


Fig 3



RESUMO

“GENE QUIMÉRICO, VETOR, USO DO PROMOTOR DE UM GENE
DE CISTEÍNA SINTASE CITOSSÓLICA DE PLANTA, E, MÉTODO
PARA FABRICAR UMA PLANTA TRANSGÊNICA OU CÉLULA DE
5 PLANTA”

Promotores de plantas constitutivos e fortes são providos,
referidos aqui como promotores de AA6, que permanecem fortes e
constitutivos sob condições de estresse biótico e/ou abiótico. Também são
providos células e organismos transgênicos, especialmente células de plantas
10 e plantas, compreendendo um promotor de AA6 e métodos para expressar as
seqüências de ácido nucleico em células e organismos usando promotores de
AA6.

BP/A/II/12

Página 14

TRATADO DE BUDAPESTE SOBRE O RECONHECIMENTO
INTERNACIONAL DO DEPÓSITO DE MICROORGANISMOS
PARA OS FINS DE PROCEDIMENTO DE PATENTE

FORMULÁRIO INTERNACIONAL

Keygene N.V.
Postbus 216
6700 AE Wageningen
Países Baixos

nome e endereço do depositante

RECIBO NO CASO DE UM DEPÓSITO
ORIGINAL expedido de acordo com a
Regra 7.1 pela AUTORIDADE
DEPOSITÁRIA INTERNACIONAL
identificada no final desta página

I. IDENTIFICAÇÃO DO MICROORGANISMO	
Referência de identificação dada pelo DEPOSITANTE: E. coli DH10B + pKG8135	Número de acesso dado pela AUTORIDADE DE DEPÓSITO INTERNACIONAL: CBS 120175
II. DESCRIÇÃO CIENTÍFICA E/OU DESIGNAÇÃO TAXONÔMICA PROPOSTA	
O microorganismo especificado em I acima foi acompanhado por: ☐ uma descrição científica ☐ uma designação taxonômica proposta <i>(marcar com uma cruz onde aplicável)</i>	
III. RECIBO E ACEITAÇÃO	
Esta Autoridade Internacional aceita o microorganismo identificado em I acima, que foi recebido pela mesma em 03-08-2006 (data dia-mês-ano do depósito original) ¹	
IV. RECIBO DE SOLICITAÇÃO PARA CONVERSÃO	
O microorganismo identificado em I acima foi recebido por esta Autoridade Depositária Internacional em <i>não aplicável</i> (data dia-mês-ano do depósito original) e uma solicitação para converter o depósito original em um depósito sob o Tratado de Budapeste foi recebido pela mesma em <i>não aplicável</i> (data dia-mês-ano da solicitação para conversão)	
V. AUTORIDADE DEPOSITÁRIA INTERNACIONAL	
Nome: Centraalbureau voor Schimmelcultures Endereço: Uppsalalaan 8 P.O. Box 85167 3508 AD UTRACHT Países Baixos	Assinaturas da(s) pessoa(s) competente(s) para representar a autoridade de depósito internacional ou do(s) funcionário(s) autorizado(s) Mrs. F.B. Snippe-Claus Dr. J.A. Stalpers Data (dia-mês-ano) 14-08-2006

¹ Onde se aplica a regra 6.4 (d), tal data é a data na qual o status da Autoridade Depositária Internacional foi adquirido.

BP/A/II/12
Página 24

TRATADO DE BUDAPESTE SOBRE O RECONHECIMENTO
INTERNACIONAL DO DEPÓSITO DE MICROORGANISMOS
PARA OS FINS DE PROCEDIMENTO DE PATENTE

FORMULÁRIO INTERNACIONAL

Keygene N.V.
Postbus 216
6700 AE Wageningen
Países Baixos

*nome e endereço da parte para a
qual a declaração de viabilidade
é expedida*

DECLARAÇÃO DE VIABILIDADE
Expedida de acordo com a Regra 10.2 pela
AUTORIDADE DEPOSITÁRIA
INTERNACIONAL identificada na página
seguinte

II. DEPOSITANTE	II. IDENTIFICAÇÃO DO MICROORGANISMO
Keygene N.V.	Número de acesso dado pela AUTORIDADE DE DEPÓSITO INTERNACIONAL: CBS 120175
Endereço: Postbus 216 6700 AE Wageningen Países Baixos	Data (dia-mês-ano) do depósito ou da transferência: ¹ 03-082006
III. DECLARAÇÃO DE VIABILIDADE	
A viabilidade do microorganismo identificado em II acima foi testada em 08-08-2006 ² . Nessa data (dia-mês-ano), o referido microorganismo era ☐ ³ viável ☐ ³ não mais viável	

¹ Indica a data do depósito original ou, quando um novo depósito ou uma transferência foi feita, a data relevante mais recente (data do novo depósito ou data da transferência).

² Nos casos referidos na Regra 10.2 (a) (ii) e (iii), referir-se ao teste de viabilidade mais recente.

³ Marcar com uma cruz a caixa aplicável.

IV. CONDIÇÕES NAS QUAIS A VIABILIDADE FOI REALIZADA ⁴	
V. AUTORIDADE DEPOSITÁRIA INTERNACIONAL	
Nome: Centraalbureau voor Schimmelcultures	Assinaturas da(s) pessoa(s) competente(s) para representar a autoridade de depósito internacional ou do(s) funcionário(s) autorizado(s)
Endereço: Uppsalalaan 8 P.O Box 85167 3508 AD UTRECHT Países Baixos	Mrs. F.B. Snippe-Claus Dr. J.A. Stalpers
	Data (dia-mês-ano) 14-08-2006

⁴ Preencher se a informação foi solicitada e se os resultados do teste foram negativos.