



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

(21) **PI0709202-4 A2**

(22) Data de Depósito: 15/03/2007
(43) Data da Publicação: 28/06/2011
(RPI 2112)



★ B R P I 0 7 0 9 2 0 2 A 2 ★

(51) Int.Cl.:
C12N 15/82 2006.01
A01H 5/00 2006.01
C12Q 1/68 2006.01
C12N 9/10 2006.01

(54) Título: **SEQÜÊNCIAS DE NUCLEOTÍDEO QUE CODIFICAM NICOTIANA BETA-1,2-XILOSILTRANSFERASE**

(57) Resumo: SEQUÊNCIAS DE NUCLEOTÍDEO QUE CODIFICAM NICOTIANA BETA-1,2-XILOSILTRANSFERASE. A presente invenção refere-se a novas seqüências de nucleotídeo de 1,2-xilosiltransferase e usos das mesmas.

(30) Prioridade Unionista: 23/03/2006 EP 06 075714.3,
30/03/2006 US 60/787,551

(73) Titular(es): BAYER BIOSCIENCE N.V.

(72) Inventor(es): Koen Weterings

(74) Procurador(es): Dannemann ,Siemens, Bigler &
Ipanema Moreira

(86) Pedido Internacional: PCT EP2007002322 de 15/03/2007

(87) Publicação Internacional: WO 2007/107296de 27/09/2007

Relatório Descritivo da Patente de Invenção para "SEQÜÊNCIAS DE NUCLEOTÍDEO QUE CODIFICAM *NICOTIANA* BETA-1,2-XILOSIL-TRANSFERASE".

A presente invenção refere-se à novas seqüências de nucleotídeo de espécies e cultivares de *Nicotiana*, particularmente de *Nicotiana benthamiana* e *Nicotiana tabacum* cv. Petite Havana SR1, que codificam β -1,2-xilosiltransferase (XylT) e a seu uso para produzir plantas *Nicotiana* modificadas, particularmente plantas *Nicotiana* Bentamiana e *Nicotiana* Tabacum vc. Petite Havana SR1, as quais têm um menor nível ou um padrão alterado de N-glicanos ligadas à proteína imunogênicas, particularmente um menor nível de resíduos de beta-1,2-xilose sobre as N-glicanos ligadas à proteína, do que a contraparte plantas *Nicotiana* não-modificadas. Tais plantas *Nicotiana* podem ser obtidas através de diminuição da expressão do(s) gene(s) de XylT de *Nicotiana* endógeno, por exemplo, através de modificação da atividade do(s) gene(s) de XylT de *Nicotiana* endógeno através de troca o gene de XylT de *Nicotiana* endógeno por outro alelo do gene de XylT, o que proporciona um menor nível de resíduos de beta-1,2-xilose sobre as N-glicanos ligadas à proteína ou através de qualquer combinação dos mesmos.

Descrição da Técnica Relacionada

O uso de plantas transgênicas para a produção de proteínas recombinantes com valor adicionado, tal como anticorpos, vacinas, produtos de sangue humano, hormônios, reguladores do crescimento e similares, é descrito para oferecer muitas vantagens práticas, econômicas e de segurança comparado com sistemas mais convencionais, tais como culturas de células de inseto e animal, fungos filamentosos e bactérias (revisto por Stoger e outros, 2002; Twyman e outros, 2003; Fischer e outros, 2004).

Embora a via de síntese de proteína seja grandemente a mesma em plantas e animais, existem algumas diferenças em modificações pós-traducionais, particularmente com relação às estruturas da cadeia de glicana. Assim, proteínas humanas recombinantes derivadas de planta tendem a ter os grupos carboidrato beta(1→2)-xilose e alfa(1→3)-fucose, os quais estão ausentes em mamíferos, mas carecem dos resíduos de galactose e ácido

siálico terminais que são encontrados em muitas glicoproteínas humanas nativas (Twyman e outros, 2003).

A enzima que catalisa a transferência de xilose de UDP-xilose para manose com núcleo beta-ligado da N-glicanos ligadas à proteína é a beta-1,2-xilosiltransferase (XylT). XylT é uma enzima única às plantas e algumas espécies de animais não-vertebrados, por exemplo, em espécies *Schistosoma* (Khoo e outros, 1997) e caracol (por exemplo, Mulder e outros, 1995) e não ocorre em seres humanos ou em outros vertebrados.

Tezuka e outros (1992) caracterizaram uma XylT de sycamore (*Acer pseudoplatanus* L.).

Zeng e outros (1997) descreveram a purificação de uma XylT de microssomas de soja. Apenas uma parte do cDNA de XylT de soja foi isolada (W099/29835 AI).

Strasser e outros (2000) e o WO01/64901 descrevem o isolamento de um gene de XylT de *Arabidopsis*, a sequência de aminoácido prevista da proteína de XylT codificada e sua atividade enzimática *in vitro* e *in vivo*.

As seguintes entradas em banco de dados que identificam sequências de gene e cDNA de XylT putativas e experimentalmente demonstradas, partes das mesmas ou sequências homólogas, puderam ser identificadas: AJ627182, AJ627183 (*Nicotiana tabacum* cv. Xanthi), AM179855 (*Solanum tuberosum*), AM179856 (*Vitis vinifera*), AJ891042 (*Populus alba* x *Populus tremula*), AY302251 (*Medicago sativa*), AJ864704 (*Saccharum officinarum*), AM179857 (*Zea mays*), AM179853 (*Hordeum vulgare*), AM179854 (*Sorghum bicolor*), BD434535, AJ277603, AJ272121, AF272852, AX236965 (*Arabidopsis thaliana*), AJ621918 (*Oryza sativa*), AR359783, AR359782, AR123000, AR123001 (Soja), AJ618933 (*Physcomitrella patens*).

Strasser e outros (2004a) reportam duas abordagens para a modulação da via de N-glicosilação em plantas: primeiro, silenciamento de gene pós-transcricional foi usado para realizar bloqueios (*knock down*) na expressão de beta-1,2-N-acetil glicosaminil transferase I (GnTI), uma enzima envolvida no processamento de resíduos oligomanosídicos em um híbrido e

N-glicanos complexas em eucariotas superiores, para avaliar a influência de expressão de GnTI sobre a formação de N-glicanos complexas em *Nicotiana benthamiana*. Caracterização de perfil de N-glicano não revelou alterações significativas do padrão de N-glicano total indicando que, mesmo uma atividade residual mínima de GnTI permite a biossíntese de N-glicanos complexas em *Nicotiana benthamiana*. Eles ainda reportam que uma abordagem similar para o *knock down* de XylT resultou em uma redução significativa de N-glicanos beta-1,2-xilosiladas. Em segundo, de forma a obter uma eliminação completa de resíduos de beta-1,2-xilose e alfa-1,3-fucose de N-glicanos, plantas *Arabidopsis* com *knockout* triplo foram geradas usando linhagens com mutação por inserção. Essas plantas exibem deficiência completa de beta-1,2-xilosil transferase e alfa-1,3-fucosil transferase de núcleo ativas, carecem de N-glicanos ligadas à proteína imunogênica e sintetizam predominantemente estruturas humanizadas com resíduos de beta-N-acetil glicosamina terminais (Strasser e outros, 2004b).

Safras de folhas, tal como tabaco, são consideradas como sendo fortes candidatos para a produção comercial de proteínas recombinantes (vide, por exemplo, Twyman e outros, 2003).

O objetivo da presente invenção é proporcionar seqüências de gene e cDNA de XylT alternativas de espécies e cultivares de *Nicotiana*, particularmente de *Nicotiana benthamiana* e *Nicotiana tabacum* cv. Petite Havana SR1, as quais são melhor adequadas para modificar a expressão de XylT, em particular espécies e cultivares de *Nicotiana*.

Sumário da Invenção

Em um aspecto da invenção, é proporcionado um método para produzir uma célula de planta ou planta *Nicotiana* tendo um baixo nível de resíduos de beta-1,2-xilose sobre N-glicanos ligadas à proteína compreendendo as etapas de introdução de um gene quimérico em células de planta de uma espécie ou cultivar de *Nicotiana* para gerar células de planta transgênica, o gene quimérico compreendendo um promotor expressível em planta operavelmente ligado; uma região de DNA passível de transcrição compreendendo uma primeira região de sentido de DNA compreendendo uma

seqüência de nucleotídeo de pelo menos 19 de 20 nucleotídeos consecutivos selecionados de uma seqüência de nucleotídeo que codifica uma proteína de XylT de *Nicotiana* ou o complemento da mesma, a seqüência de nucleotídeo, de preferência, obtenível de espécies ou cultivares de *Nicotiana*, em que pelo menos 19 de 20 nucleotídeos consecutivos codificam pelo menos um aminoácido de XylT espécie ou cultivar de *Nicotiana*-específico ou selecionada de uma seqüência de nucleotídeo de um gene de XylT de *Nicotiana* ou cDNA de XylT de *Nicotiana* ou o complemento da mesma, a seqüência de nucleotídeo, de preferência, obtenível de uma espécie ou cultivar de *Nicotiana*, em que os pelo menos 19 de 20 nucleotídeos consecutivos compreendem pelo menos um nucleotídeo de XylT espécie de *Nicotiana*-específico; uma segunda região de anti-sentido de DNA compreendendo uma seqüência de nucleotídeo de pelo menos 19 nucleotídeos consecutivos os quais têm pelo menos 95% de identidade de seqüência ao complemento da primeira região de DNA; em que uma molécula de RNA transcrita a partir da região de DNA passível de transcrição é capaz de formar uma região de RNA fita simples pelo menos entre uma região de RNA transcrita a partir da primeira região de sentido de DNA e uma região de RNA transcrita a partir da segunda região de anti-sentido de DNA; e uma região de DNA compreendendo um sinal de poliadenilação e término de transcrição funcional em plantas; opcionalmente, identificação de uma célula de planta *Nicotiana* transgênica a qual tem um menor nível de resíduos de beta-1,2-xilose sobre N-glicanos ligadas à proteína do que uma célula de planta *Nicotiana* não-transformada; opcionalmente regeneração das células de planta *Nicotiana* transgênicas para obter plantas *Nicotiana* transgênicas; e opcionalmente, identificação de uma planta *Nicotiana* transgênica a qual tem um menor nível de resíduos de beta-1,2-xilose sobre N-glicanos ligadas à proteína do que uma planta *Nicotiana* não-transformada. O aminoácido ou nucleotídeo de XylT espécie ou cultivar de *Nicotiana*-específico pode ser um aminoácido ou nucleotídeo de XylT *Nicotiana benthamiana*-específico ou *Nicotiana tabacum* cv. Petite Havana SR1-específica e a espécie ou cultivar pode, de preferência, ser *Nicotiana benthamiana* ou *Nicotiana tabacum* cv. Petite Havana

SR1, respectivamente. A seqüência de nucleotídeo que codifica uma proteína de XylT de *Nicotiana* pode compreender uma seqüência de nucleotídeo que codifica a seqüência de aminoácido de SEQ ID NO.: 12 ou SEQ ID NO.: 14 ou a seqüência de aminoácido de SEQ ID NO.: 4, SEQ ID NO.: 6, SEQ ID NO.: 8 ou SEQ ID NO.: 10 e a seqüência de nucleotídeo do gene de XylT de *Nicotiana* pode compreender a seqüência de nucleotídeo de SEQ ID NO.: 11, SEQ ID NO.: 13, ou SEQ ID No. 21 ou a seqüência de nucleotídeo de SEQ ID NO.: 3, SEQ ID NO.: 5, SEQ ID NO.: 8, SEQ ID NO.: 10, ou SEQ ID NO.: 17.

É outro objetivo da invenção proporcionar um método para produzir uma célula de planta ou planta *Nicotiana* tendo um baixo nível de resíduos de beta-1,2-xilose sobre N-glicanos ligadas à proteína compreendendo as etapas de fornecimento de uma ou mais moléculas de RNA de filamento duplo às células de planta ou plantas de uma espécie ou cultivar de *Nicotiana*, em que as moléculas de RNA de filamento duplo compreendem duas fitas de RNA, uma fita de RNA consistindo essencialmente em uma seqüência de nucleotídeo de RNA de 19 de 20 a 21 nucleotídeos consecutivos selecionada de uma seqüência que codifica uma proteína de XylT de *Nicotiana* ou o complemento da mesma, a seqüência de nucleotídeo, de preferência, obténível da espécie ou cultivar de *Nicotiana*, em que 19 de 20 a 21 nucleotídeos consecutivos compreendem pelo menos um aminoácido de XylT espécie ou cultivar de *Nicotiana*-específico ou selecionada da seqüência de nucleotídeo de um gene de XylT de *Nicotiana* ou um cDNA de XylT de *Nicotiana* ou o complemento do mesmo, a seqüência de nucleotídeo, de preferência, obténível da espécie ou cultivar de *Nicotiana*, em que os 19 de 20 a 21 nucleotídeos consecutivos compreendem pelo menos um nucleotídeo de XylT espécie ou cultivar de *Nicotiana*-específico; e identificação de uma célula de planta ou planta *Nicotiana* compreendendo a molécula ou moléculas de RNA de filamento duplo as quais têm um menor nível de resíduos de beta-1,2-xilose sobre N-glicanos ligadas à proteína do que a mesma célula de planta ou planta *Nicotiana* a qual não compreende a molécula ou moléculas de RNA de filamento duplo. O RNA de filamento duplo pode ser fornecido às

células de planta ou plantas através de integração de um gene quimérico no genoma de células de planta da espécie ou cultivar de *Nicotiana* para gerar células de planta transgênicas e, opcionalmente, regeneração das células de planta para obter plantas transgênicas, o gene quimérico compreendendo uma região de DNA compreendendo pelo menos 19 de 20 nucleotídeos consecutivos selecionados de uma seqüência de nucleotídeo que codifica uma proteína de XylT de *Nicotiana* ou o complemento da mesma, a seqüência de nucleotídeo, de preferência, obténível espécie ou cultivar de *Nicotiana*, em que os 19 de 20 nucleotídeos consecutivos codificam pelo menos um aminoácido de XylT espécie ou cultivar de *Nicotiana*-específico ou selecionada da seqüência de nucleotídeo de um gene de XylT de *Nicotiana* ou um cDNA de XylT de *Nicotiana* ou o complemento da mesma, a seqüência de nucleotídeo, de preferência, obténível da espécie ou cultivar de *Nicotiana*, em que os 19 de 20 nucleotídeos consecutivos compreendem pelo menos um nucleotídeo de XylT espécie de *Nicotiana*-específico, em uma orientação sentido e/ou anti-sentido; operavelmente ligado a um promotor expressível em planta e uma região de DNA compreendendo uma sinal de poliadenilação e término de transcrição funcional em plantas. O aminoácido ou nucleotídeo de XylT espécie ou cultivar de específico *Nicotiana* pode, ser um aminoácido ou nucleotídeo de XylT *Nicotiana benthamiana*-específico ou *Nicotiana tabacum* cv. Petite Havana SR1-específico, e as espécies ou cultivar, de nicotiana podem de preferência, ser nicotiana benthamiana ou nicotiana tabacum cv. Petite Havana SR1 respectivamente. A seqüência de nitrogênio que codifica uma proteína de XylT de *Nicotiana* pode compreender uma seqüência de nucleotídeo que codifica a seqüência de aminoácido de SEQ ID NO.: 12 ou SEQ ID NO.: 14 ou a seqüência de aminoácido de SEQ ID NO.: 4, SEQ ID NO.: 6, SEQ ID NO.: 8 ou SEQ ID NO.: 10 e a seqüência de nucleotídeo do gene de XylT de *Nicotiana* pode compreender a seqüência de nucleotídeo de SEQ ID NO.: 11, SEQ ID NO.: 13 ou SEQ ID No. 21 ou a seqüência de nucleotídeo de SEQ ID NO.: 3, SEQ ID NO.: 5, SEQ ID NO.: 8, SEQ ID NO.: 10 ou SEQ ID NO.: 17.

É ainda outro objetivo da invenção proporcionar um método para

identificar um fragmento de DNA de XylT de *Nicotiana* compreendendo as etapas de fornecimento de DNA genômico ou cDNA obtível de uma espécie ou cultivar de *Nicotiana*, seleção de um meio do seguinte grupo: um fragmento de DNA compreendendo uma seqüência de nucleotídeo que codifica a seqüência de aminoácido de SEQ ID NO.: 4, SEQ ID NO.: 6, SEQ ID NO.: 8, SEQ ID NO.: 10, SEQ ID NO.: 12 ou SEQ ID NO.: 14, para uso como uma sonda; um fragmento de DNA compreendendo a seqüência de nucleotídeo de qualquer uma de SEQ ID NO.: 3, SEQ ID NO.: 5, SEQ ID NO.: 7, SEQ ID NO.: 9, SEQ ID NO.: 11, SEQ ID NO.: 13, SEQ ID NO.: 17 ou SEQ ID NO.: 21, para uso como uma sonda; um fragmento de DNA ou oligonucleotídeo compreendendo uma seqüência de nucleotídeo consistindo em entre 20 a 1513 nucleotídeos consecutivos selecionados de uma seqüência de nucleotídeo que codifica a seqüência de aminoácido de SEQ ID NO.: 4 ou SEQ ID NO.: 6 para uso como uma sonda; um fragmento de DNA ou oligonucleotídeo compreendendo uma seqüência de nucleotídeo consistindo em entre 20 a 3574 nucleotídeos consecutivos selecionados de uma seqüência de nucleotídeo que codifica a seqüência de aminoácido de SEQ ID NO.: 8, SEQ ID NO.: 10, SEQ ID NO.: 12 ou SEQ ID NO.: 14 para uso como uma sonda; um fragmento de DNA ou oligonucleotídeo compreendendo uma seqüência de nucleotídeo consistindo em entre 20 a 3574 nucleotídeos consecutivos selecionada de uma seqüência de nucleotídeo de qualquer uma de SEQ ID NO.: 3, SEQ ID NO.: 5, SEQ ID NO.: 7, SEQ ID NO.: 9, SEQ ID NO.: 11, SEQ ID NO.: 13, SEQ ID NO.: 17 ou SEQ ID NO.: 21 para uso como uma sonda; uma seqüência de oligonucleotídeo tendo uma seqüência de nucleotídeo compreendendo entre 20 a 200 nucleotídeos consecutivos selecionada de uma seqüência de nucleotídeo que codifica a seqüência de aminoácido de SEQ ID NO.: 4 ou SEQ ID NO.: 6, para uso como um iniciador em uma reação de PCR; uma seqüência de oligonucleotídeo tendo uma seqüência de nucleotídeo compreendendo entre 20 a 200 nucleotídeos consecutivos selecionada de uma seqüência de nucleotídeo que codifica a seqüência de aminoácido de SEQ ID NO.: 8, SEQ ID NO.: 10, SEQ ID NO.: 12 ou SEQ ID NO.: 14 , para uso como um iniciador em uma reação de PCR;

uma seqüência de oligonucleotídeo tendo uma seqüência de nucleotídeo compreendendo entre 20 a 200 nucleotídeos consecutivos selecionada da seqüência de nucleotídeo de qualquer uma de SEQ ID NO.: 3, SEQ ID NO.: 5, SEQ ID NO.: 7, SEQ ID NO.: 9, SEQ ID NO.: 11, SEQ ID NO.: 13, SEQ ID NO.: 17 ou SEQ ID NO.: 21, para uso como um iniciador em uma reação de PCR; ou um oligonucleotídeo tendo a seqüência de nucleotídeo de qualquer uma de SEQ ID NO.: 1, SEQ ID NO.: 2, SEQ ID NO.: 15 ou SEQ ID NO.: 16, SEQ ID NO.: 19 ou SEQ ID NO.: 20 para uso como um iniciador em uma reação de PCR; e utilização desse meio para identificar um fragmento de DNA de XylT da espécie ou cultivar de *Nicotiana* realizando uma PCR usando o DNA genômico ou o cDNA e os iniciadores ou realizando hibridização usando o DNA genômico ou cDNA e as sondas. O fragmento identificado pode, subsequente, ser isolado e usado para obter uma célula de planta ou planta *Nicotiana* tendo um baixo nível de resíduos de beta-1,2-xilose sobre N-glicanos ligadas à proteína.

A invenção também proporciona um método para identificar um alelo de XylT de *Nicotiana* correlacionado a um baixo nível de resíduos de beta-1,2-xilose sobre N-glicanos ligadas à proteína compreendendo as etapas de fornecimento de uma população, opcionalmente uma população com mutação, de diferentes linhagens de planta de uma espécie ou cultivar de *Nicotiana*; identificação, em cada linhagem de planta, da população de um alelo de XylT de *Nicotiana* de acordo com o método descrito acima; análise do nível de resíduos de beta-1,2-xilose sobre N-glicanos ligadas à proteína de cada linhagem de planta da população e identificação daquelas linhagens de planta tendo um menor nível de resíduos de beta-1,2-xilose sobre N-glicanos ligadas à proteína do que outras linhagens de planta; e correlação do baixo nível de resíduos de beta-1,2-xilose sobre N-glicanos ligadas à proteína em uma linhagem de planta à presença de um alelo de beta-1,2-xilose *Nicotiana*-específico. O alelo de XylT de *Nicotiana* pode ser introduzido em uma célula de planta ou planta *Nicotiana* de escolha para obter uma célula de planta ou planta *Nicotiana* com um baixo nível de resíduos de beta-1,2-xilose sobre N-glicanos ligadas à proteína.

É ainda outro objetivo da invenção proporcionar: um fragmento de DNA isolado que codifica uma proteína compreendendo a sequência de aminoácido SEQ ID NO.: 12 ou SEQ ID NO.: 14 ou qualquer parte da mesma que codifica pelo menos um aminoácido de XylT *Nicotiana benthamiana*-específico; um fragmento de DNA isolado compreendendo a sequência de nucleotídeo de SEQ ID NO.: 11, SEQ ID NO.: 13 ou SEQ ID NO.: 21 ou qualquer parte da mesma compreendendo pelo menos um nucleotídeo de XylT *Nicotiana benthamiana*-específico; um fragmento de DNA isolado que codifica uma proteína compreendendo a sequência de aminoácido de SEQ ID NO.: 4 ou SEQ ID NO.: 6, SEQ ID NO.: 8, SEQ ID NO.: 10 ou qualquer parte da mesma que codifica pelo menos um aminoácido de XylT *Nicotiana tabacum* cv. Petite Havana SR1-específico; um fragmento de DNA isolado compreendendo a sequência de nucleotídeo de SEQ ID NO.: 3 ou SEQ ID NO.: 5, SEQ ID NO.: 7, SEQ ID NO.: 9 ou SEQ ID NO.: 17 ou qualquer parte da mesma compreendendo pelo menos um nucleotídeo de XylT *Nicotiana tabacum* cv. Petite Havana SR1-específico.

A invenção ainda proporciona um gene quimérico compreendendo os fragmentos de DNA operavelmente ligados: um promotor expressível em planta; uma região de DNA passível de transcrição compreendendo uma primeira região de DNA compreendendo pelo menos 19 de 20 nucleotídeos consecutivos selecionados de uma sequência de nucleotídeo que codifica uma proteína de XylT de *Nicotiana* ou o complemento da mesma, em que os 19 de 20 nucleotídeos consecutivos codificam pelo menos um aminoácido de XylT espécie ou cultivar de *Nicotiana*-específico ou selecionada da sequência de nucleotídeo de um gene de XylT de *Nicotiana* ou um cDNA de XylT de *Nicotiana* ou o complemento da mesma, em que os 19 de 20 nucleotídeos consecutivos compreendem pelo menos um nucleotídeo de XylT espécie de *Nicotiana*-específico, na orientação anti-sentido; uma segunda região de DNA compreendendo pelo menos 19 de 20 nucleotídeos consecutivos selecionados de uma sequência de nucleotídeo que codifica uma proteína de XylT de *Nicotiana* ou o complemento da mesma, em que os 19 de 20 nucleotídeos consecutivos codificam pelo menos um aminoácido de XylT

espécie ou cultivar de *Nicotiana*-específico ou selecionada da sequência de nucleotídeo de um gene de XylT de *Nicotiana* ou um cDNA de XylT de *Nicotiana* ou o complemento da mesma, em que os 19 de 20 nucleotídeos consecutivos compreendem pelo menos um nucleotídeo de XylT espécie de *Nicotiana*-específico, na orientação sentido, pelo que uma molécula de RNA produzida através de transcrição da região de DNA passível de transcrição é capaz de formar uma região de DNA de filamento duplo através de emparelhamento de base pelo menos entre uma região de RNA correspondendo à primeira região de DNA e uma região de RNA correspondendo à segunda região de DNA; e uma região de DNA compreendendo a um sinal de poliadenilação e término de transcrição funcional em planta. O gene quimérico também pode compreender um promotor expressível em planta; uma região de DNA compreendendo pelo menos 19 de 20 nucleotídeos consecutivos selecionados de uma sequência de nucleotídeo que codifica uma proteína de XylT de *Nicotiana* ou o complemento da mesma, em que os 19 de 20 nucleotídeos consecutivos codificam pelo menos um aminoácido de XylT espécie ou cultivar de *Nicotiana*-específico ou selecionada da sequência de nucleotídeo de um gene de XylT de *Nicotiana* ou um cDNA de XylT de *Nicotiana* ou o complemento da mesma, em que os 19 de 20 nucleotídeos consecutivos compreendem pelo menos um nucleotídeo de XylT espécie de *Nicotiana*-específico, na orientação sentido ou anti-sentido; e uma região de DNA compreendendo um sinal de poliadenilação e término de transcrição funcional em plantas.

Células de planta *Nicotiana* compreendendo tais genes quiméricos e plantas *Nicotiana* consistindo essencialmente em tais células de planta *Nicotiana*, bem como sementes da mesma são também proporcionadas pela invenção.

A invenção também se refere ao uso de uma sequência de nucleotídeo que codifica uma proteína compreendendo a sequência de aminoácido de SEQ ID NO.: 4, SEQ ID NO.: 6, SEQ ID NO.: 8, SEQ ID NO.: 10, SEQ ID NO.: 12 ou SEQ ID NO.: 14 ou qualquer parte da mesma compreendendo pelo menos 19 de 20 nucleotídeos consecutivos que codificam pelo

menos um aminoácido de XylT espécie ou cultivar de *Nicotiana*-específico, para diminuir o nível de resíduos de beta-1,2-xilose sobre N-glicanos ligadas à proteína em uma planta *Nicotiana* ou ao uso de uma seqüência de nucleotídeo compreendendo a seqüência de nucleotídeo de SEQ ID NO.: 3, SEQ ID NO.: 5, SEQ ID NO.: 7, SEQ ID NO.: 9, SEQ ID NO.: 11, SEQ ID NO.: 13, SEQ ID NO.: 17 ou SEQ ID NO.: 21 ou qualquer parte da mesma compreendendo pelo menos 19 de 20 nucleotídeos consecutivos compreendendo pelo menos um nucleotídeo de XylT espécie ou cultivar de *Nicotiana*-específico, para diminuir o nível de resíduos de beta-1,2-xilose sobre N-glicanos ligadas à proteína em uma planta *Nicotiana*, para identificar um gene de XylT ou cDNA de XylT em uma espécie ou cultivar de *Nicotiana*, para identificar um alelo de um gene *XylT* correlacionado a um baixo nível de resíduos de beta-1,2-xilose sobre N-glicanos ligadas à proteína em uma espécie ou cultivar de *Nicotiana* ou para introduzir um alelo de um gene *XylT* correlacionado a um baixo nível de resíduos de beta-1,2-xilose sobre N-glicanos ligadas à proteína em uma espécie ou cultivar de *Nicotiana*.

Com os objetivos precedentes e outros, vantagens e características da invenção que se tornarão aqui depois evidentes, a natureza da invenção pode ser mais claramente compreendida através de referência à descrição detalhada a seguir de diferentes modalidades da invenção, às reivindicações em anexo e às figuras.

Breve Descrição das Figuras

A Figura 1 é um alinhamento de DNA global (baseado na matriz de escore linear padrão com os seguintes parâmetros: penalidade por combinação errônea = 2, penalidade por abertura de gap = 4 e penalidade por extensão de gap = 1) entre uma seqüência de cDNA de *Nicotiana tabacum* cv. Xanthi que codifica uma proteína de XylT putativa (número de acesso AJ627182; SEQ ID NO: 23) e duas diferentes seqüências de cDNA de XylT isoladas de *Nicotiana tabacum* cv. Petite Havana SR1 (SEQ ID NO: 3 e 5). Os pontos representam nucleotídeos nas seqüências de cDNA de *Nicotiana tabacum* cv. Petite Havana SR1 que são idênticas aos nucleotídeos correspondentes na seqüência de cDNA de *Nicotiana tabacum* cv. Xanthi; as li-

nhas tracejadas representam a ausência de nucleotídeos nas seqüências de cDNA de *Nicotiana tabacum* cv. Petite Havana SR1 correspondendo aos nucleotídeos na seqüência de cDNA de *Nicotiana tabacum* cv. Xanthi.

A Figura 2 é um alinhamento de proteína global (baseado na matriz de escore blossom 62) entre a proteína de XylT putativa codificada pela seqüência de cDNA de *Nicotiana tabacum* cv. Xanthi (número de acesso AJ627182; SEQ ID NO: 24) e por duas seqüências de cDNA de XylT diferentes isoladas de *Nicotiana tabacum* cv. Petite Havana SR1 (SEQ ID NO: 4 e 6). Os pontos representam aminoácidos nas seqüências de proteína de *Nicotiana tabacum* cv. Petite Havana SR1 que são idênticas aos aminoácidos correspondentes na seqüência de proteína de *Nicotiana tabacum* cv. Xanthi; as linhas tracejadas representam a ausência de aminoácidos nas seqüências de proteína de *Nicotiana tabacum* cv. Petite Havana SR1 correspondendo aos aminoácidos na seqüência de proteína de *Nicotiana tabacum* cv. Xanthi.

A Figura 3 é um alinhamento de DNA global (baseado na matriz de escore linear padrão com os seguintes parâmetros: penalidade por combinação errônea = 2, penalidade por abertura de gap = 4 e penalidade por extensão de gap = 1) entre a seqüência de DNA genômico de *Nicotiana tabacum* cv. Xanthi que codifica uma proteína de XylT putativa (número de acesso AJ627183; SEQ ID NO: 25) e duas diferentes seqüências de DNA genômico de XylT isoladas de *Nicotiana tabacum* cv. Petite Havana SR1 (SEQ ID NOs: 7 e 9) e duas diferentes seqüências de DNA genômico de XylT na isoladas de *Nicotiana benthamiana* (SEQ ID NO: 11 e 13). Os pontos representam nucleotídeos nas seqüências de DNA genômico de *Nicotiana tabacum* cv. Petite Havana SR1 que são idênticas aos nucleotídeos correspondentes na seqüência de DNA genômico de *Nicotiana tabacum* cv. Xanthi; a linha tracejada representa a ausência de nucleotídeos nas seqüências de DNA genômico de *Nicotiana tabacum* cv. Petite Havana SR1 correspondendo aos nucleotídeos na seqüência de DNA genômico de *Nicotiana tabacum* cv. Xanthi.

A Figura 4 é um alinhamento de proteína global (baseado na

matriz de escore blossom 62) entre a proteína de XylT putativa codificada pela sequência de DNA genômico de *Nicotiana tabacum* cv. Xanthi (número de acesso AJ627183; SEQ ID NO: 26) e por duas diferentes sequências de DNA genômico de XylT isoladas de *Nicotiana tabacum* cv. Petite Havana SR1 (SEQ ID NO: 8 e 10) e por duas diferentes sequências de DNA genômico isoladas de *Nicotiana benthamiana* (SEQ ID NO: 12 e 14). Os pontos representam aminoácidos nas sequências de proteína de *Nicotiana tabacum* cv. Xanthi; a linha tracejada representa a ausência de aminoácidos nas sequências de proteína de *Nicotiana tabacum* cv. Petite Havana SR1 correspondendo aos aminoácidos na sequência de proteína de *Nicotiana tabacum* cv. Xanthi.

Descrição Detalhada de Diferentes Modalidades da Invenção

A presente invenção é baseada na descoberta de que genes de XylT e cDNAs de XylT de espécies e cultivares de *Nicotiana*, particularmente *Nicotiana benthamiana* e *Nicotiana tabacum* cv. Petite Havana SR1, são excelentes sequências de nucleotídeo fonte para obter plantas dessas espécies e cultivares de *Nicotiana*, particularmente plantas *Nicotiana benthamiana* e plantas *Nicotiana tabacum* cv. Petite Havana SR1, respectivamente, tendo um baixo nível de resíduos de beta-1,2-xilose sobre N-glicanos ligadas à proteína, por exemplo, através de modificação da atividade de gene(s) endógeno(s) de XylT de *Nicotiana*, através de troca de um gene de XylT de *Nicotiana* endógeno por outro alelo do gene de XylT de *Nicotiana* o qual proporciona um baixo nível de resíduos de beta-1,2-xilose sobre N-glicanos ligadas à proteína ou através de qualquer combinação dos mesmos.

Em uma modalidade, a invenção é relacionada a um método para obtenção de uma célula de planta ou planta *Nicotiana* tendo um baixo nível de resíduos de beta-1,2-xilose sobre N-glicanos ligadas à proteína através de redução da expressão do(s) gene(s) de XylT endógeno na célula de planta ou planta *Nicotiana* através de fornecimento de uma ou mais moléculas de RNA de silenciamento às células de planta ou plantas de uma espécie ou cultivar de *Nicotiana*, em que as moléculas de RNA de silenciamento compreendem uma parte de uma sequência de nucleotídeo que codi-

fica uma proteína de XylT de *Nicotiana*, de preferência obtida da referida espécie ou cultivar de *Nicotiana*, em que a referida parte codifica pelo menos um aminoácido de XylT espécie ou cultivar de *Nicotiana*-específico ou em que as moléculas de RNA de silenciamento compreendem uma parte de uma sequência de nucleotídeo de um gene de XylT de *Nicotiana* ou um cDNA de XylT de *Nicotiana*, de preferência obtida da referida espécie ou cultivar de *Nicotiana*, em que a referida parte compreende pelo menos um nucleotídeo de XylT espécie ou cultivar de *Nicotiana*-específico.

Conforme usado aqui, "RNA de silenciamento" ou "molécula de RNA de silenciamento" se refere a qualquer molécula de RNA a qual, quando de introdução de uma célula de planta, reduz a expressão de um gene-alvo. Tal RNA de silenciamento pode, por exemplo, ser o assim denominado "RNA anti-sentido", pelo que a molécula de RNA compreende uma sequência de pelo menos 20 nucleotídeos consecutivos tendo 95% de identidade de sequência ao complemento da sequência do ácido nucléico alvo, de preferência a sequência de codificação do gene-alvo. Contudo, RNA anti-sentido também pode ser dirigido à sequências regulatórias de genes alvo, incluindo as sequências promotoras e sinais de poliadenilação e término de transcrição. RNA de silenciamento ainda inclui o assim denominado "RNA sentido", pelo que a molécula de RNA compreende uma sequência de pelo menos 20 nucleotídeos consecutivos tendo 95% de identidade de sequência à sequência do ácido nucléico alvo. Outro RNA de silenciamento pode ser "RNA não-poliadenilado" compreendendo pelo menos 20 nucleotídeos consecutivos tendo 95% de identidade de sequência ao complemento da sequência do ácido nucléico alvo, tal como descrito no WO01/12824 ou US6423885 (ambos os documentos sendo aqui incorporados por referência). Ainda outro tipo de RNA de silenciamento é uma molécula de RNA conforme descrito no WO03/076619 (aqui incorporado por referência) compreendendo pelo menos 20 nucleotídeos consecutivos tendo 95% de identidade de sequência à sequência do ácido nucléico alvo ou ao complemento da mesma e ainda compreendendo uma região grandemente fita dupla conforme descrito no WO03/076619 (incluindo regiões grande de fita dupla compreendendo um

senal de localização nuclear de um viróide do tipo viróide do tubérculo de fusso da batata ou compreendendo repetições de trinucleotídeo CUG). RNA de silenciamento também pode ser RNA de filamento duplo compreendendo uma fita sentido e anti-sentido conforme definido aqui, em que a fita sentido e anti-sentido são capazes de emparelhamento de base uma com a outra para formar uma região de RNA de filamento duplo (de preferência, os referidos pelo menos 20 nucleotídeos consecutivos do RNA sentido e anti-sentido são complementares uns aos outros). A região sentido e anti-sentido pode também estar presente dentro de uma molécula de RNA, tal como um RNA em gancho, hairpin (hpRNA) pode ser formado quando a região sendo e anti-sentido forma uma região de RNA de filamento duplo. hpRNA é bem-conhecido dentro da técnica (vide, por exemplo, WO99/53050, aqui incorporado por referência). O hpRNA pode ser classificado como hpRNA longo, tendo regiões sentido e anti-sentido longas as quais podem ser grandemente complementares, mas não precisam ser totalmente complementares (tipicamente, mais longa do que cerca de 200 bp, oscilando entre 200-1000 bp). hpRNA também pode ser menor oscilando, quanto ao tamanho, de cerca de 30 a cerca de 42 bp, mas não mais longo do que 94 bp (vide WO04/073390, aqui incorporado por referência). RNA de silenciamento pode também pode ser molécula de micro-RNA artificiais conforme descrito, por exemplo, no WO2005/052170, WO2005/047505 ou US 2005/0144667 (todos os documentos incorporados aqui por referência).

Em outra modalidade, as moléculas de RNA de silenciamento são fornecidas à célula de planta ou planta da espécie ou cultivar de *Nicotiana* através de produção de uma célula de planta ou planta transgênica da espécie ou cultivar de *Nicotiana* compreendendo um gene quimérico capaz de produzir uma molécula de RNA de silenciamento, particularmente uma molécula de RNA de filamento duplo ("dsRNA"), em que a fita de RNA complementar de tal molécula de dsRNA compreende uma parte de uma sequência de nucleotídeo que codifica uma proteína de XylT de *Nicotiana*, de preferência obtida da referida espécie ou cultivar de *Nicotiana*, em que a referida parte codifica pelo menos um aminoácido de XylT espécie ou cultivar

de *Nicotiana*-específico ou em que a fita de RNA complementar de tal molécula de dsRNA compreende uma parte de uma seqüência de nucleotídeo de um gene de XylT de *Nicotiana* ou um cDNA de XylT de *Nicotiana*, de preferência obtido da referida espécie ou cultivar de *Nicotiana*, em que a referida parte compreende pelo menos uma nucleotídeo de XylT espécie ou cultivar de *Nicotiana*-específico.

"*Nicotiana*", conforme usado aqui, inclui todas as espécies de *Nicotiana* conhecidas tais como, mas não limitado a, *Nicotiana acaulis*, *N. acuminata*, *N. africana*, *N. alata*, *N. amplexicaulis*, *N. arentsii*, *N. attenuata*, *N. benavidesii*, *N. benthamiana*, *N. bigelovii*, *N. bonariensis*, *N. cavicola*, *N. clevelandii*, *N. cordifolia*, *N. corymbosa*, *N. debneyi*, *N. excelsior*, *N. forgetiana*, *N. fragrans*, *N. glauca*, *N. glutinosa*, *N. goodspeedii*, *N. gossei*, *N. hybrid*, *N. ingulba*, *N. kawakamii*, *N. knightiana*, *N. langsдорffii*, *N. linearis*, *N. longiflora*, *N. maritima*, *N. megalosiphon*, *N. miersii*, *N. noctiflora*, *N. nudicaulis*, *N. obtusifolia*, *N. occidentalis*, *N. otophora*, *N. paniculata*, *N. pauciflora*, *N. petunioides*, *N. plumbaginifolia*, *N. quadrivalvis*, *N. raimondii*, *N. repanda*, *N. rosulata*, *N. rotundifolia*, *N. rustica*, *N. setchellii*, *N. simulans*, *N. solanifolia*, *N. spegazzinii*, *N. stocktonii*, *N. suaveolens*, *N. sylvestris*, *N. tabacum*, *N. thyrsoflora*, *N. tomentosa*, *N. tomentosiformis*, *N. trigonophylla*, *N. umbratica*, *N. undulata*, *N. velutina*, *N. wigandiioides* e *Nicotiana x sandera* e todos os cultivares conhecidos de *Nicotiana* tais como, mas não limitado a *Nicotiana tabacum*, tal como cv. Burley21, cv. Delgold, cv. Petit Havana, cv. Petit Havana SR1, cv. Samsun e cv. Xanthi.

Nicotiana tabacum, a qual é o tabaco comum, é uma espécie híbrida tetraplóide, a qual se originou das espécies diplóides *Nicotiana sylvestris* e *Nicotiana tomentosiformis*.

Um gene de XylT de *Nicotiana* ou um cDNA de XylT de *Nicotiana* se refere a uma seqüência de nucleotídeo de um gene de XylT que ocorre naturalmente em uma espécie ou cultivar de *Nicotiana* ou a um cDNA correspondendo ao mRNA de um gene de XylT que ocorre naturalmente em uma espécie ou cultivar de *Nicotiana*. Similarmente, uma proteína de XylT de *Nicotiana* se refere a uma proteína conforme ela ocorre naturalmente em

uma espécie ou cultivar de *Nicotiana*.

Exemplos de seqüências de nucleotídeo que codificam uma proteína de XylT de *Nicotiana* incluem aquelas obtidas de *Nicotiana benthamiana* que codificam a seqüência de aminoácido apresentada em SEQ ID NO.: 12 ou SEQ ID NO.: 14 e aquelas obtidas de *Nicotiana tabacum* cv. Petite Havana SR1 que codificam a seqüência de aminoácido apresentada em SEQ ID NO.: 4, SEQ ID NO.: 6, SEQ ID NO.: 8 ou SEQ ID NO.: 10.

Exemplos de seqüências de nucleotídeo que codificam um gene de XylT de *Nicotiana* incluem aquelas obtidas de *Nicotiana benthamiana* compreendendo a seqüência de nucleotídeo apresentada em SEQ ID NO.: 11, SEQ ID NO.: 13 ou SEQ ID NO.: 21 e aquelas obtidas de *Nicotiana tabacum* cv. Petite Havana SR1 compreendendo a seqüência de nucleotídeo apresentada em SEQ ID NO.: 7 ou SEQ ID NO.: 9.

Exemplos de seqüências de nucleotídeo de um cDNA de XylT de *Nicotiana* incluem aquelas obtidas de *Nicotiana tabacum* cv. Petite Havana SR1 compreendendo a seqüência de nucleotídeo apresentada em SEQ ID NO.: 3, SEQ ID NO.: 5 ou SEQ ID NO.: 17.

Contudo, será imediatamente claro para aqueles versados na técnica que as seqüências de nucleotídeo exemplificadas ou partes das mesmas podem ser usadas para identificar outras seqüências de nucleotídeo de genes de XylT de *Nicotiana* ou cDNAs de XylT de *Nicotiana* em espécies ou cultivares de *Nicotiana* e que tais seqüências de nucleotídeo ou partes das mesmas podem também ser usadas, por exemplo, para diminuir o nível de resíduos de beta-1,2-xilose sobre N-glicanos ligadas à proteína em plantas *Nicotiana*. As seqüências de nucleotídeo exemplificadas poderiam ser usadas para selecionar:

i) um fragmento de DNA compreendendo uma seqüência de nucleotídeo que codifica a seqüência de aminoácido de SEQ ID NO.: 4, SEQ ID NO.: 6, SEQ ID NO.: 8, SEQ ID NO.: 10, SEQ ID NO.: 12 ou SEQ ID NO.: 14, para uso como uma sonda;

ii) um fragmento de DNA compreendendo a seqüência de nucleotídeo de qualquer uma de SEQ ID NO.: 3, SEQ ID NO.: 5, SEQ ID NO.: 7,

SEQ ID NO.: 9, SEQ ID NO.: 11, SEQ ID NO.: 13, SEQ ID NO.: 17 ou SEQ ID NO.: 21, para uso como uma sonda;

iii) um fragmento de DNA ou oligonucleotídeo compreendendo uma sequência de nucleotídeo consistindo em entre 20 a 1513 nucleotídeos consecutivos selecionados de uma sequência de nucleotídeo que codifica a sequência de aminoácido de SEQ ID NO.: 4 ou SEQ ID NO.: 6, para uso como uma sonda;

iv) um fragmento de DNA ou oligonucleotídeo compreendendo uma sequência de nucleotídeo consistindo em entre 20 to 3574 nucleotídeos consecutivos selecionados de uma sequência de nucleotídeo que codifica a sequência de aminoácido de SEQ ID NO.: 8, SEQ ID NO.: 10, SEQ ID NO.: 12 ou SEQ ID NO.: 14 , para uso como uma sonda;

v) um fragmento de DNA ou oligonucleotídeo compreendendo uma sequência de nucleotídeo consistindo em entre 20 a 3574 nucleotídeos consecutivos selecionados de uma sequência de nucleotídeo de qualquer uma de SEQ ID NO.: 3, SEQ ID NO.: 5, SEQ ID NO.: 7, SEQ ID NO.: 9, SEQ ID NO.: 11, SEQ ID NO.: 13, SEQ ID NO.: 17 ou SEQ ID NO.: 21, para uso como uma sonda;

vi) uma sequência de oligonucleotídeo tendo uma sequência de nucleotídeo compreendendo entre 20 a 200 nucleotídeos consecutivos selecionados de uma sequência de nucleotídeo que codifica a sequência de aminoácido de SEQ ID NO.: 4 ou SEQ ID NO.: 6, para uso como um iniciador em uma reação de PCR;

vii) uma sequência de oligonucleotídeo tendo uma sequência de nucleotídeo compreendendo entre 20 a 200 nucleotídeos consecutivos selecionados de uma sequência de nucleotídeo que codifica a sequência de aminoácido de SEQ ID NO.: 8, SEQ ID NO.: 10, SEQ ID NO.: 12 ou SEQ ID NO.: 14 , para uso como um iniciador em uma reação de PCR;

viii) uma sequência de oligonucleotídeo tendo uma sequência de nucleotídeo compreendendo entre 20 a 200 nucleotídeos consecutivos selecionada da sequência de nucleotídeo de qualquer uma de SEQ ID NO.: 3, SEQ ID NO.: 5, SEQ ID NO.: 7, SEQ ID NO.: 9, SEQ ID NO.: 11, SEQ ID

NO.: 13, SEQ ID NO.: 17 ou SEQ ID NO.: 21, para uso como um iniciador em uma reação de PCR; ou

ix) um oligonucleotídeo tendo a seqüência de nucleotídeo de qualquer uma de SEQ ID NO.: 1, SEQ ID NO.: 2, SEQ ID NO.: 15 ou SEQ ID NO.: 16, SEQ ID NO.: 19 ou SEQ ID No.20 para uso como um iniciador em uma reação de PCR.

Realizando uma PCR usando DNA genômico ou cDNA de espécies ou cultivares de *Nicotiana* e os oligonucleotídeos mencionados como iniciadores ou realizando hibridização, de preferência sob condições estritas entre o DNA genômico ou cDNA de espécies e cultivares de *Nicotiana* e as sondas mencionadas, esses outros genes de XylT de *Nicotiana* ou cDNAs de XylT de *Nicotiana* ou fragmentos dos mesmos podem ser identificados e/ou isolados.

"Condições de hibridização estritas", conforme usado aqui, significa que hibridização geralmente ocorrerá se há pelo menos 95% e, de preferência, pelo menos 97% de identidade de seqüência entre a sonda e a seqüência-alvo. exemplos de condições de hibridização estrita são incubação durante a noite em uma solução compreendendo formamida a 50%, 5 x SSC (NaCl a 150 mM, citrato trissódico a 15 mM), fosfato de sódio a 50 mM (pH de 7,6), 5x solução de Denhardt, sulfato de dextrana a 10% e 20 µg/ml de DNA veículo cisalhado, desnaturado, tal como DNA de esperma de salmão, seguido por lavagem do suporte de hibridização em 0,1 x SSC a aproximadamente 65°C, por exemplo, durante cerca de 10 min. (duas vezes). Outras condições de hibridização e lavagem são bem-conhecidas e exemplificadas em Sambrook e outros, *Molecular Cloning: A Laboratory Manual*, Segunda Edição, Cold Spring Harbor, NY (1989), particularmente capítulo 11.

Um nucleotídeo de XylT espécie de *Nicotiana*-específico" ou um "nucleotídeo de XylT cultivar de *Nicotiana*-específico" se refere a um nucleotídeo de uma seqüência de nucleotídeo de um gene de XylT ou um cDNA de XylT de uma espécie ou cultivar de *Nicotiana* que difere de ou não está presente na seqüência de nucleotídeo correspondente do gene de XylT de *Ni-*

cotiana tabacum cv. Xanthi representada em SEQ ID NO: 25 ou do cDNA de XylT de *Nicotiana tabacum* cv. Xanthi representada em SEQ ID NO: 23, respectivamente.

Um "aminoácido de XylT espécie de *Nicotiana*-específico" ou um "aminoácido de XylT cultivar de *Nicotiana*-específico" se refere a um aminoácido da sequência de aminoácido de uma proteína de XylT codificada por um gene de XylT ou codificada por um cDNA de XylT de uma espécie ou cultivar de *Nicotiana* que difere de ou não está presente na sequência de aminoácido da proteína de XylT codificada pelo gene de XylT de *Nicotiana tabacum* cv. Xanthi representada em SEQ ID NO: 26 ou codificada pelo cDNA de XylT de *Nicotiana tabacum* cv. Xanthi representada em SEQ ID NO: 24, respectivamente.

Para determinar a presença de um nucleotídeo ou aminoácido de XylT espécie ou cultivar de *Nicotiana*-específico em uma sequência de nucleotídeo de um gene de XylT ou um cDNA de XylT de uma espécie ou cultivar de *Nicotiana* ou na sequência de aminoácido de uma proteína de XylT codificada por um gene de XylT ou codificada por um cDNA de XylT de uma espécie ou cultivar de *Nicotiana*, para fins da presente invenção, as sequências de nucleotídeo ou aminoácido de XylT da espécie ou cultivar de *Nicotiana* são comparadas com as sequências de nucleotídeo ou aminoácido de XylT correspondentes de *Nicotiana tabacum* cv. Xanthi através de alinhamento das sequências usando um procedimento de alinhamento global (para sequências de nucleotídeo, a matriz de escore de ausência usada é "linear-padrão" com penalidade por combinação errônea = 2, penalidade por abertura de gap = 4 e penalidade por extensão de gap = 1. Para sequências de proteína, a matriz de escore de ausência é "blossum 62"; Henikoff e Henikoff, 1992). Para realizar o alinhamento, o programa Align Plus (fornecido pela Scientific & Educational Software, EUA) pode ser usado.

Um exemplo de tal alinhamento de DNA global é o alinhamento de DNA global da sequência de cDNA de XylT de *Nicotiana tabacum* cv. Xanthi representada em SEQ ID NO: 23, com as sequências de cDNA de XylT de *Nicotiana tabacum* cv. Petite Havana SR1 representadas em SEQ

ID NO: 3 e 5, na Figura 1. exemplos de nucleotídeos de XylT *Nicotiana tabacum* cv. Petite Havana SR1–específicos baseado nesse alinhamento global de DNA incluem:

- o nucleotídeo nas posições 1041, 1323, 1332 ou 1421 de SEQ ID NO: 3,

- o nucleotídeo nas posições 62, 76, 87, 104, 117, 122, 139, 140, 148, 155, 169, 190, 199, 202, 212, 213, 216, 265, 287, 294, 316, 373, 385, 388, 430, 554, 607, 628, 643, 838, 892, 897, 898, 941, 1005, 1021, 1039 ou 1495 de SEQ ID NO: 5.

Outro exemplo de tal alinhamento de DNA global é o alinhamento de DNA global da sequência do gene de XylT de *Nicotiana tabacum* cv. Xanthi representada em SEQ ID NO: 25 com as sequências do gene de XylT de *Nicotiana tabacum* cv. Petite Havana SR1 representadas em SEQ ID NO: 7 e 9 e com as sequências do gene de XylT de *Nicotiana benthamiana* representadas em SEQ ID NO: 11 e 13, na Figura 3. exemplos de nucleotídeos de XylT de *Nicotiana tabacum* cv. Petite Havana SR1–específicos determinados baseado nesse alinhamento global de DNA incluem:

- o nucleotídeo nas posições 61, 75, 86, 100-120, 124, 137, 142, 159, 160, 168, 175, 189, 210, 219, 222, 232, 233, 236, 285, 307, 314, 336, 393, 405, 408, 450, 574, 627, 648, 663, 692, 698, 702, 721, 754, 802, 821, 842, 852, 856, 901, 903, 906, 907, 908, 917, 927, 928, 930, 931, 960, 961, 965, 974, 977, 981, 983, 986, 1001, 1019, 1027, 1029, 1034, 1068, 1073, 1099, 1120, 1129, 1144, 1154, 1158, 1181, 1193, 1208, 1212, 1228, 1230, 1239, 1275, 1313-1316, 1348, 1353, 1357, 1384, 1386, 1496, 1531, 1571, 1601, 1629, 1681, 1696, 1698, 1730, 1754, 1761, 1772, 1789, 1800, 1802, 1811, 1814, 1815, 1855-1861, 1929, 2172, 2190, 2322, 2324, 2328, 2353, 2354, 2391, 2404, 2419, 2428, 2429, 2433, 2434, 2459, 2464, 2478, 2479, 2481, 2484, 2512, 2540-2590, 2595, 2604, 2606, 2630, 2633, 2638, 2670, 2673, 2680, 2695, 2698, 2710, 2711, 2752, 2806, 2811, 2812, 2855, 2919, 2953, 2966, 3217, 3226, 3229 ou 3232 de SEQ ID NO: 7,

- o nucleotídeo nas posições 553, 606, 627, 642, 671, 677, 681, 700, 733, 781, 800, 821, 831, 835, 880, 882, 885, 886, 887, 896, 906, 907,

909, 910, 939, 940, 944, 953, 956, 960, 962, 965, 980, 998, 1006, 1008, 1013, 1047, 1052, 1078, 1099, 1108, 1123, 1133, 1137, 1160, 1172, 1187, 1191, 1207, 1209, 1218, 1254, 1292, 1293, 1294, 1295, 1327, 1332, 1336, 1363, 1365, 1475, 1510, 1550, 1580, 1608, 1660, 1675, 1677, 1709, 1733, 1740, 1751, 1768, 1779, 1781, 1790, 1793, 1794, 1834-1840, 1908, 2151, 2169, 2301, 2303, 2307, 2332, 2333, 2370, 2383, 2398, 2407, 2408, 2412, 2413, 2438, 2443, 2457, 2458, 2460, 2463, 2491, 2519-2569, 2574, 2583, 2585, 2609, 2612, 2617, 2649, 2652, 2659, 2674, 2677, 2689, 2690, 2731, 2785, 2790, 2791, 2834, 2898, 2932, 2945, 3196 ou 3205 de SEQ ID NO: 9.

Exemplos de nucleotídeos de *XylT Nicotiana benthamiana*—específicos determinados baseado nesse alinhamento global de DNA incluem:

- o nucleotídeo nas posições 71, 72, 75, 77, 86, 90, 104, 116, 147, 158, 212, 222, 246, 264, 286, 317, 321, 345, 402, 472, 479, 488, 526, 552, 612, 637, 668, 669, 670, 701, 726, 734, 742, 747, 773, 785, 795, 796, 802, 831, 871, 872, 874, 888, 897, 898, 899, 901, 902, 927, 931, 932, 1039, 1044, 1047, 1080, 1091, 1103, 1104, 1113, 1118, 1131, 1134, 1145, 1152, 1164, 1179, 1183, 1199, 1201, 1206, 1227, 1286, 1287, 1288, 1289, 1296, 1301, 1317, 1328, 1332, 1347, 1376, 1388, 1424, 1429, 1458, 1464, 1510, 1517, 1534, 1559, 1672, 1675, 1676, 1677, 1693, 1705, 1719, 1750, 1757, 1761, 1765, 1832, 1838, 1862, 1872, 1877, 1978, 2010, 2074, 2111, 2115, 2227, 2251, 2259, 2271, 2283, 2296, 2297, 2341, 2348, 2361, 2370, 2371, 2375, 2384, 2401, 2404, 2406, 2495, 2497, 2521, 2529, 2561, 2607, 2701, 2777, 2822, 2843, 2856, 2867, 3020, 3052, 3053, 3116, 3143 ou 3227 de SEQ ID NO: 11

- o nucleotídeo nas posições 77, 107, 203, 297, 312, 399, 449, 469, 489, 492, 496, 529, 538, 566, 573, 633, 661, 662, 666, 671, 683, 690, 699, 723, 763, 774, 784, 785, 791, 861, 877, 886, 887, 888, 890, 891, 920, 921, 943, 996, 1015, 1034, 1116, 1190, 1226, 1277-1280, 1282, 1287, 1331, 1343, 1360, 1386-1651, 1672, 1689, 1738, 1770, 1791, 1813, 1820, 1822, 1831, 1832, 1862, 1869, 1874, 1882, 1893, 1906, 1935, 1945, 1956, 1988, 2007, 2033, 2034, 2045, 2049, 2050, 2167, 2198, 2280, 2299, 2315, 2355, 2392, 2413, 2428, 2442, 2464, 2468, 2477, 2493, 2522, 2544, 2548, 2573,

2574, 2639, 2648, 2649, 2653, 2655, 2659, 2679, 2684, 2740, 2773, 2775, 2781, 2796, 2799, 2807, 2816, 2839, 2857, 2975, 2977, 2990, 3053, 3071, 3083, 3119, 3132, 3265, 3311 ou 3392 de SEQ ID NO: 13.

Um exemplo de tal alinhamento de proteína global é o alinhamento de proteína global da sequência de proteína de XylIT codificada pela sequência de cDNA de XylIT de *Nicotiana tabacum* cv. Petite Havana SR1 representada em SEQ ID NO: 24 com as sequências de proteína de XylIT codificadas pelas sequências de cDNA de XylIT de *Nicotiana tabacum* cv. Petite Havana SR1 representadas em SEQ ID NO: 4 e 6, na Figura 2. exemplos de aminoácidos de XylIT *Nicotiana tabacum* cv. Petite Havana SR1-específicos determinados baseado nesse alinhamento global de proteína incluem:

- o aminoácido nas posições 472 ou 502 de SEQ ID NO: 4,
- o aminoácido nas posições 20, 28, 38, 40, 46, 51, 70, 71, 95, 97, 184, 213, 298, 313, 334 ou 497 de SEQ ID NO: 6.

Outro exemplo de tal alinhamento global de proteína é o alinhamento global de proteína das sequências de proteína de XylIT codificadas pela sequência do gene de XylIT de *Nicotiana tabacum* cv. Xanthi representada em SEQ ID NO: 26 com as sequências de proteína de XylIT codificadas pelas sequências do gene de XylIT de *Nicotiana tabacum* cv. Petite Havana SR1 representadas em SEQ ID NO: 8 e 10 e com as sequências de proteína de XylIT codificadas pelas sequências do gene de XylIT de *Nicotiana benthamiana* representadas em SEQ ID NO: 12 e 14, na Figura 4. exemplos de aminoácidos de XylIT *Nicotiana tabacum* cv. Petite Havana SR1-específicos determinado baseado nesse alinhamento global de proteína incluem:

- o aminoácido nas posições 19, 27, 32-38, 44, 46, 52, 57, 76, 77, 101, 103, 190, 219, 304, 319, 340 ou 356 de SEQ ID NO: 8,
- o aminoácido nas posições 183, 212, 297, 312, 333 ou 349 de SEQ ID NO: 10.

Exemplos de aminoácidos de XylIT *Nicotiana benthamiana*-específicos determinado baseado nesse alinhamento global de proteína incluem:

- o aminoácido nas posições 22, 24, 27, 33, 37, 51, 69, 94, 104,

156, 158, 161, 174, 182, 211, 238, 297, 349 ou 414 de SEQ ID NO: 12,

- o aminoácido nas posições 1, 26, 36, 68, 99, 133, 150, 157, 166, 180, 189, 211, 218, 245, 257, 296, 327, 348 ou 392 de SEQ ID NO: 14.

A parte de uma seqüência de nucleotídeo que codifica uma proteína de XylT de *Nicotiana* e a parte de uma seqüência de nucleotídeo de um gene de XylT de *Nicotiana* ou um cDNA de XylT de *Nicotiana* compreendida dentro da molécula de RNA de silenciamento, particularmente dentro de uma fita da molécula de RNA de filamento duplo, deverá ter pelo menos 19 nucleotídeos de comprimento, mas pode variar de cerca de 19 nucleotídeos (nt) até um comprimento igual ao comprimento (em nucleotídeos) da seqüência que codifica proteína de XylT de *Nicotiana* ou da seqüência de cDNA ou gene de XylT de *Nicotiana*. O comprimento total da seqüência de nucleotídeo sentido ou anti-sentido pode, assim, ser de pelo menos 25 nt ou pelo menos cerca de 50 nt ou pelo menos cerca de 100 nt ou pelo menos cerca de 150 nt ou pelo menos cerca de 200 nt ou pelo menos cerca de 500 nt. Espera-se que não haja limite máximo quanto ao comprimento total da seqüência de nucleotídeo anti-sentido ou sentido. Contudo, por razões práticas (tais como, por exemplo, estabilidade dos genes quiméricos), espera-se que o comprimento da seqüência de nucleotídeo sentido ou anti-sentido não exceda a 5000 nt, particularmente não exceda a 2500 nt e poderia estar limitada a cerca de 1000 nt.

Será apreciado que quanto mais longo o comprimento total da parte da seqüência de nucleotídeo que codifica uma proteína de XylT de *Nicotiana* ou a parte de uma seqüência de nucleotídeo de um gene de XylT de *Nicotiana* ou um cDNA de XylT de *Nicotiana* (região sentido ou anti-sentido), menos estridentes são os requisitos para identidade de seqüência entre essas regiões e a seqüência correspondente no gene de XylT endógeno da espécie ou cultivar de *Nicotiana* que ela complementa. De preferência, o ácido nucléico de interesse terá uma identidade de seqüência de pelo menos cerca de 75% com a seqüência-alvo correspondente, particularmente pelo menos cerca de 80%, mais particularmente pelo menos cerca de 85%, muito particularmente cerca de 90%, especialmente cerca de 95%, mais especial-

mente cerca de 100%, muito especialmente será idêntica à parte correspondente da sequência-alvo ou seu complemento. Contudo, é preferido que o ácido nucléico de interesse sempre inclua uma sequência de cerca de 19 nucleotídeos consecutivos, particularmente cerca de 25 nt, mais particularmente cerca de 50 nt, especialmente cerca de 100 nt, muito especialmente cerca de 150 nt com 100% de identidade de sequência à parte correspondente do ácido nucléico de XylT alvo, em que os referidos cerca de 19 nucleotídeos consecutivos, particularmente cerca de 25 nt, mais particularmente cerca de 50 nt, especialmente cerca de 100 nt, muito especialmente cerca de 150 nt codificam pelo menos um aminoácido de XylT espécie ou cultivar de *Nicotiana*-específico ou compreendem pelo menos um nucleotídeo de XylT espécie ou cultivar de *Nicotiana*-específico.

Para fins da presente invenção, a "identidade de sequência" de duas sequências de nucleotídeo ou aminoácido relacionadas, expressa como um percentual, se refere a um número de posições nas duas sequências otimamente alinhadas as quais têm resíduos idênticos (x100) dividido pelo número de posições comparadas. Uma gap, isto é, uma posição em um alinhamento onde um resíduo está presente em uma sequência, mas não na outra, é considerada como uma posição com resíduos não idênticos. De preferência, para cálculo da identidade de sequência e designing da sequência sentido ou anti-sentido correspondente, o número de gaps deverá ser minimizado, particularmente para as sequências sentido mais curtas.

Será claro que sempre que sequências de nucleotídeo de moléculas de RNA são definidas através de referência a uma sequência de nucleotídeo das moléculas de DNA correspondentes, a timina (T) em uma sequência de nucleotídeo deverá ser substituída por uracila (U). Se referência é à moléculas de RNA ou DNA estará claro a partir do contexto da aplicação.

Foi demonstrado que o requisito mínimo para silenciamento de um gene-alvo particular é a presença, na sequência de nucleotídeo do gene quimérico de silenciamento, de uma sequência de nucleotídeo de cerca de 20-21 nucleotídeos consecutivos de comprimento correspondendo à sequência do gene-alvo, na qual pelo menos 19 dos 20-21 nucleotídeos con-

secutivos são idênticos à seqüência do gene-alvo correspondente. "19 de 20 nucleotídeos consecutivos", conforme usado aqui, se refere a uma seqüência de nucleotídeo de 20 nucleotídeos consecutivos selecionados do gene-alvo tendo um nucleotídeo erroneamente combinado.

Para silenciamento do gene de XylT endógeno de uma espécie ou cultivar de *Nicotiana*, é preferido que a seqüência de nucleotídeo de gene quimérico de silenciamento compreenda pelo menos 19 de 20-21 nucleotídeos consecutivos de uma seqüência de nucleotídeo que codifica uma proteína de XylT de *Nicotiana*, em que os referidos pelo menos 19 de 20-21 nucleotídeos consecutivos codificam pelo menos um aminoácido de XylT espécie ou cultivar de *Nicotiana*-específico ou compreenda pelo menos 19 de 20-21 nucleotídeos consecutivos de uma seqüência de nucleotídeo de um gene de XylT de *Nicotiana* ou um cDNA de XylT de *Nicotiana*, em que os referidos pelo menos 19 de 20-21 nucleotídeos consecutivos compreendem pelo menos um nucleotídeo de XylT espécie ou cultivar de *Nicotiana*-específico.

Conforme usado aqui, "uma planta *Nicotiana* tendo um baixo nível de resíduos de beta-1,2-xilose sobre N-glicanos ligadas à proteína" é uma planta (particularmente uma planta *Nicotiana* obtida de acordo com o método da invenção), na qual a atividade de XylT é diminuída ou eliminada, resultando em um menor nível de resíduos de beta-1,2-xilose sobre N-glicanos ligadas à proteína do que o nível de resíduos de beta-1,2-xilose sobre N-glicanos ligadas à proteína em uma planta *Nicotiana* de controle não tratada de acordo com os métodos da invenção ou resultando na ausência de resíduos de beta-1,2-xilose sobre N-glicanos ligadas à proteína. Uma indicação de atividade de XylT pode ser obtida através de comparação do nível de resíduos de beta-1,2-xilose presentes sobre as glicanas de proteínas da planta *Nicotiana* obtida de acordo com os métodos da invenção com o nível de resíduos de beta-1,2-xilose presentes sobre as glicanas de proteínas de uma planta *Nicotiana* de controle não tratada de acordo com os métodos da invenção. O nível de resíduos de beta-1,2-xilose sobre N-glicanos ligadas à proteína de plantas pode ser medido, por exemplo, através de análise de Western blot usando anticorpos xilose-específicos conforme descrito,

por exemplo, por Faye e outros (Analytical Biochemistry, 1993, 209: 104-108) ou através de espectrometria de massa sobre glicanas isoladas das glicoproteínas de planta usando espectronomia de massa por tempo de voo da Ionização/Dessorção de Matriz assistida por laser (MALDI-TOF-MS) conforme descrito, por exemplo, por Kolarich e Altmann (2000, Anal. Biochem. 285: 64-75) ou usando espectronomia de massa Aleatória por Cromatografia de Líquido (LC/MS/MS) conforme descrito, por exemplo, por Henriksson e outros (2003, Biochem. J. 375: 61-73).

dsRNA que codifica genes quiméricos que reduzem a expressão de XylT de *Nicotiana* de acordo com a invenção pode compreender um íntron, tal como um íntron heterólogo localizado, por exemplo, na sequência espaçadora entre as regiões de RNA sentido e anti-sentido de acordo com a descrição do WO 99/53050 (incorporado aqui por referência).

Se tornou evidente que moléculas de RNA de filamento duplo, tais como aquelas descritas acima, são clivadas em células de planta em pequenos fragmentos de RNA de cerca de 20-21 nucleotídeos, os quais servem como sequência guia na degeneração do mRNA correspondente (revisito por Baulcombe, 2004). Assim, em outra modalidade, a invenção se refere a um método para produção de uma célula de planta de *Nicotiana* ou planta tendo um baixo nível de resíduos de beta-1,2-xilose sobre N-glicanos ligadas à proteína compreendendo as etapas de:

a) fornecimento de uma ou mais moléculas de RNA de filamento duplo à células de uma planta de uma espécie ou cultivar de *Nicotiana*, em que as moléculas de RNA de filamento duplo compreendem pelo menos dois filamento de RNA, um filamento de RNA consistindo essencialmente em uma sequência de nucleotídeo de RNA de 20 a 21 nucleotídeos consecutivos selecionados de uma sequência de nucleotídeo que codifica uma proteína de XylT de *Nicotiana*, de preferência obtida da referida espécie ou cultivar de *Nicotiana*, em que os referidos 20 a 21 nucleotídeos consecutivos codificam pelo menos um aminoácido de XylT espécie ou cultivar de *Nicotiana*-específico ou um filamento de RNA consistindo essencialmente de uma sequência de RNA de 20 a 21 nucleotídeos consecutivos de uma sequência de

nucleotídeo de um gene de XylT de *Nicotiana* ou um cDNA de *Nicotiana*, de preferência obtida da referida espécie ou cultivar de *Nicotiana*, em que os referidos 20 a 21 nucleotídeos consecutivos compreendem pelo menos um nucleotídeo de XylT espécie ou cultivar de *Nicotiana*-específico; e

b) identificação de uma planta *Nicotiana* compreendendo essa molécula ou moléculas de RNA de filamento duplo as quais têm um menor nível de resíduos de beta-1,2-xilose sobre N-glicanos ligadas à proteína do que a mesma planta *Nicotiana* a qual não compreende a molécula ou moléculas de RNA de filamento duplo.

As seqüências de dsRNA de 20-21 nt de comprimento mencionadas também são geradas no curso de silenciamento RNA anti-sentido mediado ou silenciamento RNA sentido mediado convencional. Assim, em outra modalidade da invenção, um método é proporcionado para produção de uma célula de planta ou planta *Nicotiana* tendo um baixo nível de resíduos de beta-1,2-xilose sobre N-glicanos ligadas à proteína, compreendendo a etapa de fornecimento às células de uma planta da espécie ou cultivar de *Nicotiana*, de um gene quimérico compreendendo, operavelmente ligados, os seguintes fragmentos de DNA:

a) um promotor expressível em planta;

b) uma região de DNA compreendendo pelo menos 20 nucleotídeos consecutivos selecionados de uma seqüência de nucleotídeo que codifica uma proteína de XylT de *Nicotiana*, de preferência obtida da mesma espécie ou cultivar de *Nicotiana*, em que os referidos pelo menos 20 nucleotídeos consecutivos codificam pelo menos um aminoácido de XylT espécie ou cultivar de *Nicotiana*-específico ou compreendendo pelo menos 20 nucleotídeos consecutivos de uma seqüência de nucleotídeo de um gene de XylT de *Nicotiana* ou um cDNA de XylT de *Nicotiana*, de preferência obtida da referida espécie ou cultivar de *Nicotiana*, em que os referidos pelo menos 20 nucleotídeos consecutivos compreendem pelo menos um nucleotídeo de XylT espécie ou cultivar de *Nicotiana*-específico, na orientação anti-sentido ou sentido;

c) uma região de DNA compreendendo um sinal de poliadenila-

ção e de término de transcrição funcional em plantas.

As regiões de nucleotídeo anti-sentido ou sentido mencionadas podem, assim, ser de cerca de 21 nt a cerca de 5000 nt de comprimento, tal como 21nt, 40 nt, 50 nt, 100nt, 200 nt, 300nt, 500nt, 1000 nt ou mesmo cerca de 2000 nt ou mais de comprimento. Além disso, não é requerido, para fins da invenção, que a sequência de nucleotídeo da molécula de gene de XylT inibitória usada ou a região de codificação do gene quimérico seja completamente idêntica ou complementar ao gene de XylT de *Nicotiana* endógeno, a expressão do qual é objetivada para ser reduzida na célula de planta *Nicotiana*. Quanto mais longa a sequência, menos rigoroso é o requisito de identidade global de sequência. Assim, as regiões sentido ou anti-sentido podem ter uma identidade global de sequência de cerca de 40 % ou 50% ou 60 % ou 70% ou 80% ou 90 % ou 100% a uma sequência de nucleotídeo do gene de *Nicotiana* endógeno ou o complemento da mesma. Contudo, conforme mencionado, regiões sentido ou anti-sentido compreenderão, de preferência, uma sequência de nucleotídeo de 19-20 nucleotídeos consecutivos tendo cerca de 100% de identidade de sequência a uma sequência de nucleotídeo do gene de XylT, em que os referidos 19-20 nucleotídeos consecutivos codificam pelo menos um aminoácido de XylT espécie ou cultivar de *Nicotiana*-específico ou compreende pelo menos um nucleotídeo de XylT espécie ou cultivar de *Nicotiana*-específico. O trecho de cerca de 100% de identidade de sequência pode ter cerca de 50, 75 ou 100 nt.

A eficiência dos genes quiméricos acima mencionados os quais, quando transcritos, proporcionam RNA de silenciamento anti-sentido ou sentido, pode ser ainda intensificada através de inclusão de elementos de DNA os quais resultam na expressão de moléculas de RNA inibitórias de XylT anormais, não poliadeniladas. Um de tais elementos de DNA adequados para essa finalidade é uma região de DNA que codifica uma ribozima de auto-replicação, conforme descrito no WO 00/01133. A eficiência pode também ser intensificada através de fornecimento das moléculas de RNA geradas com localização nuclear ou sinais de retenção, conforme descrito no WO 03/076619.

As seqüências de nucleotídeo de XylT exemplificadas de *Nicotiana benthamiana* e de *Nicotiana tabacum* podem também ser usadas para identificar alelos de XylT em uma população de plantas de uma espécie ou cultivar de *Nicotiana* os quais são correlacionados com baixos níveis de resíduos de beta-1,2-xilose sobre N-glicanos ligadas à proteína. Tais populações de plantas de uma espécie ou cultivar de *Nicotiana* podem ser populações as quais sofreram mutação prévia. Os alelos de XylT identificados podem, então, ser introduzidos em uma linhagem de planta de uma espécie ou cultivar de *Nicotiana* de escolha usando técnicas de reprodução convencionais.

Métodos para transformar plantas *Nicotiana* também são bem-conhecidos na técnica. Transformação *Agrobacterium*-mediada de *Nicotiana* foi descrita, por exemplo, em Zambryski e outros (1983, EMBO J. 2: 2143-2150), De Block e outros (1984, EMBO J. 3(8): 1681-1689) ou Horsch e outros (Science (1985) 227: 1229-1231).

As plantas *Nicotiana* transformadas obtidas de acordo com a invenção ou as plantas *Nicotiana* obtidas tendo um baixo nível de resíduos de beta-1,2-xilose sobre N-glicanos ligadas à proteína em que o gene de XylT endógeno foi substituído por um alelo de XylT, o qual está correlacionado a um menor nível de resíduos de beta-1,2-xilose sobre N-glicanos ligadas à proteína do que o alelo de XylT original, podem ser usadas em um esquema de reprodução convencional para produzir mais plantas com as mesmas características ou introduzir o gene quimérico de acordo com a invenção em outros cultivares da mesma ou de uma espécie de planta relacionada ou em plantas híbridas. Sementes obtidas das plantas transformadas contêm os genes quiméricos da invenção como um inserto genômico estável e também são abrangidas pela invenção.

Além disso, sabe-se que a introdução de um RNA anti-sentido, sentido ou filamento duplo ou genes quiméricos de codificação podem levar a uma distribuição de fenótipos, oscilando de quase nenhuma ou muito pouca supressão da expressão do gene-alvo a uma supressão muito forte ou mesmo de 100% da expressão do gene-alvo. Contudo, aqueles versados na

técnica serão capazes de selecionar aquelas células de planta, plantas, eventos ou linhagens de planta que levam a graus desejados de silenciamento e o fenótipo desejado.

Conforme usado aqui, "compreendendo" deve ser interpretado como especificando a presença das características, inteiros, etapas ou componentes estabelecidos conforme mencionado, mas não exclui a presença ou adição de uma ou mais características, inteiros, etapas ou componentes ou grupos dos mesmos. Assim, por exemplo, um ácido nucléico ou proteína compreendendo uma seqüência de nucleotídeos ou aminoácidos pode compreender mais nucleotídeos ou aminoácidos do que aqueles realmente citados, isto é, ser incrustado em uma proteína ou ácido nucléico maior. Um gene quimérico compreendendo uma região de DNA, a qual é funcional ou estruturalmente definida, pode compreender regiões de DNA adicionais, etc.

Os exemplos não limitativos a seguir descrevem genes quiméricos para a alteração do nível de resíduos de beta-1,2-xilose sobre N-glicanos ligadas à proteína em espécies *Nicotiana*, particularmente em *Nicotiana benthamiana*, e em cultivares de *Nicotiana*, particularmente em *Nicotiana tabacum* cv. Petite Havana SR1 e a usos dos mesmos. A menos que de outro modo estabelecido nos exemplos, todas as técnicas de DNA recombinante são realizadas de acordo com protocolos-padrão conforme descrito em Sambrook e outros (1989) *Molecular Cloning: A Laboratory Manual*, Segunda Edição, Cold Spring Harbor Laboratory Press, NY e nos Volumes 1 e 2 de Ausubel e outros (1994) *Current Protocols in Molecular Biology*, Current Protocols, EUA. Materiais-padrão e métodos para trabalho molecular com plantas são descritos em *Plant Molecular Biology Labfax* (1993) por R.D.D. Croy, conjuntamente publicado pela BIOS Scientific Publications Ltda (Reino Unido) e Blackwell Scientific Publications, Reino Unido.

Por toda a descrição e exemplos, referência é feita às seguintes seqüências representadas na Listagem de Seqüência:

SEQ ID NO: 1: seqüência de nucleotídeo do oligonucleotídeo XylF4 adequado para amplificar uma parte de um gene ou cDNA de XylT de *Nicotiana*.

SEQ ID NO: 2: seqüência de nucleotídeo do oligonucleotídeo XylR4 ade-

quado para amplificar uma parte de um gene ou cDNA de XylIT de *Nicotiana*.

SEQ ID NO: 3: seqüência de cDNA parcial da variante 1 de gene de XylIT de *Nicotiana tabacum* cv. Petite Havana SR1.

SEQ ID NO: 4: seqüência de aminoácido parcial da variante 1 de proteína de XylIT de *Nicotiana tabacum* cv. Petite Havana SR1.

SEQ ID NO: 5: seqüência de cDNA parcial da variante 2 do gene de XylIT de *Nicotiana tabacum* cv. Petite Havana SR1.

SEQ ID NO: 6: seqüência de aminoácido parcial da variante 2 de proteína de XylIT de *Nicotiana tabacum* cv. Petite Havana SR1.

SEQ ID NO: 7: seqüência de nucleotídeo parcial da variante 1 do gene de XylIT de *Nicotiana tabacum* cv. Petite Havana SR1.

SEQ ID NO: 8: seqüência de aminoácido parcial da variante 1 de proteína de XylIT de *Nicotiana tabacum* cv. Petite Havana SR1.

SEQ ID NO: 9: seqüência de nucleotídeo parcial da variante 2 do gene de XylIT de *Nicotiana tabacum* cv. Petite Havana SR1.

SEQ ID NO: 10: seqüência de aminoácido parcial da variante 2 de proteína de XylIT de *Nicotiana tabacum* cv. Petite Havana SR1.

SEQ ID NO: 11: seqüência de nucleotídeo parcial da variante 1 de gene de XylIT de *Nicotiana benthamiana*.

SEQ ID NO: 12: seqüência de aminoácido parcial da variante 1 de proteína de XylIT de *Nicotiana benthamiana*.

SEQ ID NO: 13: seqüência de nucleotídeo parcial da variante 2 do gene de XylIT de *Nicotiana benthamiana*.

SEQ ID NO: 14: seqüência de aminoácido parcial da variante 2 de proteína de XylIT de *Nicotiana benthamiana*.

SEQ ID NO: 15: seqüência de nucleotídeo do oligonucleotídeo XylF8 adequado para amplificar uma parte de um gene ou cDNA de XylIT de *Nicotiana tabacum* cv. Petite Havana SR1.

SEQ ID NO: 16: seqüência de nucleotídeo do oligonucleotídeo XylR8 adequado para amplificar uma parte de um gene ou cDNA de XylIT de *Nicotiana tabacum* cv. Petite Havana SR1.

SEQ ID NO: 17: seqüência de cDNA parcial da variante 1 do gene de XylIT

de *Nicotiana tabacum* cv. Petite Havana SR1.

SEQ ID NO: 18: seqüência de nucleotídeo da região de T-DNA do vetor pTKW20.

SEQ ID NO: 19: seqüência de nucleotídeo do oligonucleotídeo XylF9 adequado para amplificar uma parte de um gene ou cDNA de XylIT de *Nicotiana benthamiana*.

SEQ ID NO: 20: seqüência de nucleotídeo do oligonucleotídeo XylR9 adequado para amplificar uma parte de um gene ou cDNA de XylIT de *Nicotiana benthamiana*.

SEQ ID NO: 21: seqüência parcial da variante 1 do gene de XylIT de *Nicotiana benthamiana*.

SEQ ID NO: 22: seqüência de nucleotídeo da região de T-DNA do vetor pTKW29.

SEQ ID NO: 23: mRNA de *Nicotiana tabacum* cv. Xanthi para beta-(1,2)-xilosil transferase putativa (número de acesso AJ627182).

SEQ ID NO: 24: beta-(1,2)-xilosil transferase putativa codificada por SEQ ID NO: 23.

SEQ ID NO: 25: gene xylt de *Nicotiana tabacum* cv. Xanthi para beta-(1,2)-xilosil transferase putativa (número de acesso AJ627183).

SEQ ID NO: 26: beta-(1,2)-xilosil transferase putativa codificada por SEQ ID NO: 25.

Exemplos

Exemplo 1: Design de iniciadores degenerados para o isolamento de cDNA de XylIT e seqüências de gene de *Nicotiana tabacum* cv. Petite Havana SR1 e *Nicotiana benthamiana*

Seqüências de oligonucleotídeo a serem usadas como iniciadores degenerados em uma amplificação por PCR de cDNA de XylIT e DNA genômico de *Nicotiana tabacum* cv. Petite Havana SR1 e *Nicotiana benthamiana* foram projetadas baseado nas seqüências de éxon de uma seqüência de DNA genômico de *Nicotiana tabacum* cv. Xanthi que codifica uma proteína de XylIT putativa (número de acesso AJ627183). O iniciador dianteiro (SEQ ID NO: 1) foi projetado com CACC em sua extremidade 5' para fins de

clonagem. Dessa forma, os seguintes iniciadores degenerados foram gerados:

XylF4: 5'-CACCTTGTTTCTCTCTTCGCTCTCAACTCAATCACTC-3'

(SEQ ID NO: 1)

XylR4: 5'-TCGATCACAACTGGAGGATCCGCATAA -3'

(SEQ ID NO: 2)

Exemplo 2: Isolamento de seqüências de cDNA de XylT de *Nicotiana tabacum* cv. Petite Havana SR1

Os iniciadores degenerados descritos no Exemplo 1 foram gerados para isolar seqüências de cDNA de XylT de *Nicotiana tabacum* cv. Petite Havana SR1:

RNA foi extraído de folhas de *Nicotiana tabacum* cv. Petite Havana SR1 usando o RNeasy Plant Mini Kit (Qiagen) de acordo com o protocolo do fabricante e usado para síntese de cDNA usando o Sistema de síntese SuperScript® First-strand para RT-PCR (Invitrogen Life Technologies) de acordo com as instruções do fabricante.

Usando o cDNA como *padrão* e um par de iniciador XylF4 / XylR4, amplificação por PCR foi realizada sob as seguintes condições: 15 seg a 94°C (desnaturação) e 3 min a 68°C durante 40 ciclos (anelamento e alongamento).

Um fragmento de DNA de cerca de 1500 pares de base foi amplificado, clonado em um vetor pENTR®/D-TOPO® (Invitrogen) e vários clones foram seqüenciados (compreendendo as seqüências de SEQ ID NO: 3 - XylTc2Nt - e SEQ ID NO: 5 - XylTc7Nt).

Um alinhamento entre uma seqüência de mRNA de *Nicotiana tabacum* cv. Xanthi que codifica uma proteína de XylT putativa (número de acesso AJ627182; SEQ ID NO: 23) e as seqüências de cDNA de XylT isoladas de *Nicotiana tabacum* cv. Petite Havana SR1 (SEQ ID NO: 3 e 5) é mostrado na Figura 1.

Um alinhamento entre a proteína de XylT putativa codificada pela seqüência de mRNA de *Nicotiana tabacum* cv. Xanthi (número de acesso AJ627182; SEQ ID NO: 24) e pelas seqüências de cDNA isoladas de *Nicotiana tabacum* cv. Petite Havana SR1 (SEQ ID NO: 4 e 6) é mostrado na Figura 2.

Exemplo 3: Isolamento de seqüências de gene de XylT de *Nicotiana tabacum* cv. Petite Havana SR1 e de *Nicotiana benthamiana*

Os iniciadores degenerados descritos no Exemplo 1 foram usados para isolar seqüências de gene de XylT de *Nicotiana tabacum* cv. Petite Havana SR1 e de *Nicotiana benthamiana*:

DNA foi extraído de folhas de *Nicotiana tabacum* cv. Petite Havana SR1 e de *Nicotiana benthamiana* baseado no protocolo descrito por Bernatzky e Tanksley (1986).

Usando o DNA genômico como *padrão* e um par de iniciador XylF4 / XylR4, amplificação por PCR foi realizada sob as seguintes condições: 15 seg a 94°C (desnaturação) e 4 min e 30 seg a 68°C durante 40 ciclos (anelamento e alongamento).

Usando o DNA genômico de *Nicotiana tabacum* cv. Petite Havana SR1 como *padrão* para a amplificação por PCR, um fragmento de DNA de cerca de 3400 pares de base foi amplificado, clonado em um vetor pENTR®/D-TOPO® (Invitrogen) e vários clones foram seqüenciados (compreendendo as seqüências de SEQ ID NO: 7 – XylTg1Nt - e SEQ ID NO: 9 – XylTg3Nt).

As seqüências de DNA genômico de XylT XylTg1Nt e XylTg3Nt compreendem duas seqüências de íntron putativas e três seqüências de éxon putativas. A localização das seqüências de íntron é:

- para XylTg1Nt: do nucleotídeo na posição 679 ao nucleotídeo na posição 1974 e do nucleotídeo na posição 2125 ao nucleotídeo na posição 2722 de SEQ ID NO: 9 e

- para XylTg3Nt: do nucleotídeo na posição 658 ao nucleotídeo na posição 1953 e do nucleotídeo na posição 2104 ao nucleotídeo na posição 2701 de SEQ ID NO: 9.

Usando o DNA genômico de *Nicotiana benthamiana* como *padrão* para a amplificação por PCR, um fragmento de DNA de entre cerca de 3300 e cerca de 3600 pares de base foi amplificado, clonado em um vetor pENTR®/D-TOPO® (Invitrogen) e vários clones foram seqüenciados (compreendendo as seqüências de SEQ ID NO: 11 – XylTg14Nb - e SEQ ID NO: 13 – XylTg19Nb).

As seqüências de DNA genômico de XylT XylTg14Nb e XylTg19Nb compreendem duas seqüências de íntron putativas e três seqüências de éxon putativas. A localização das seqüências de íntron é:

- XylTg14Nb do nucleotídeo na posição 658 ao nucleotídeo na posição 1917 e do nucleotídeo na posição 2068 ao nucleotídeo na posição 2612 de SEQ ID NO: 11 e

- XylTg19Nb do nucleotídeo na posição 649 ao nucleotídeo na posição 2194 e do nucleotídeo na posição 2345 ao nucleotídeo na posição 2888 de SEQ ID NO: 13.

Um alinhamento entre a seqüência de DNA genômico de *Nicotiana tabacum* cv. Xanthi que codifica uma proteína de XylT putativa (número de acesso AJ627183; SEQ ID NO: 25) e as seqüências de DNA genômico de XylT isoladas de *Nicotiana tabacum* cv. Petite Havana SR1 (SEQ ID NO: 7 e 9) e de *Nicotiana benthamiana* (SEQ ID NO: 11 e 13) é mostrado na Figura 3.

Um alinhamento entre a proteína de XylT putativa codificada pela seqüência de DNA genômico de *Nicotiana tabacum* cv. Xanthi (número de acesso AJ627183; SEQ ID NO: 26) e pelas seqüências de DNA genômico isoladas de *Nicotiana tabacum* cv. Petite Havana SR1 (SEQ ID NO: 8 e 10) e de *Nicotiana benthamiana* (SEQ ID NO: 12 e 14) é mostrado na Figura 4.

Exemplo 4: Construção de um vetor de T-DNA contendo um gene de silenciamento de XylT de *Nicotiana*

Fragmentos de DNA amplificados de seqüências de XylT de *Nicotiana* descritas nos exemplos 2 e 3 foram usados para construir vetores de T-DNA compreendendo um gene quimérico o qual, quando de transcrição, proporciona uma molécula de RNA compreendendo uma seqüência de DNA sentido e anti-sentido a partir do fragmento de DNA amplificado e o qual poderia se emparelhar em bases para formar uma molécula de RNA de filamento duplo. Tais genes quiméricos podem ser usados para reduzir a expressão de um gene de XylT em *Nicotiana*, particularmente em *Nicotiana tabacum* cv. Petite Havana SR1 e *Nicotiana benthamiana*.

4.1. Construção de um vetor de T-DNA compreendendo um gene de silenciamento de XylT com um fragmento de DNA amplificado de

uma sequência de XylT de *Nicotiana tabacum* cv. Petite Havana SR1.

Seqüências de oligonucleotídeo a serem usadas como iniciadores não degenerados em uma amplificação de PCR de um cDNA de XylT de *Nicotiana tabacum* cv. Petite Havana SR1 foram projetadas baseado na sequência de cDNA de *Nicotiana tabacum* cv. Petite Havana SR1 isolada no Exemplo 2. O iniciador dianteiro (SEQ ID NO: 15) foi projetado com CACC em sua extremidade 5' para fins de clonagem. Dessa forma, os seguintes iniciadores não degenerados foram gerados:

XylF8: 5'-CACCTCTCGCCTTTGGGATATGAAACT -3' (SEQ ID NO: 15)

XylR8: 5'-ACAGCTTTGGTGCTGCAGAAACT -3' (SEQ ID NO: 16)

Usando o vetor compreendendo um fragmento de DNA amplificado de uma sequência de cDNA de XylT de *Nicotiana tabacum* cv. Petite Havana SR1 conforme descrito no Exemplo 2 (SEQ ID NO: 3 - XylTc2Nt) como *padrão* e o par de iniciador XylF8 / XylR8, uma amplificação por PCR foi realizada sob as seguintes condições: 15 seg a 94°C (desnaturação), 30 seg a 56°C (anelamento) e 45 seg a 68°C (alongamento) durante 25 ciclos.

Um fragmento de DNA de cerca de 470 bp (XylTi4Nt; SEQ ID NO: 17) foi amplificado e clonado em um vetor pENTR[®]/D-TOPO[®] (Invitrogen), proporcionando o plasmídeo pKW19. O plasmídeo pKW19 foi recombinado com pHellsgate12 (Helliwell e Waterhouse, Methods (2003) 30: 289-295) usando o Gateway[®] LR Clonase[®] II (Invitrogen), proporcionando o plasmídeo pTKW20.

A sequência de T-DNA do pTKW20 (SEQ ID NO: 18), assim, compreende:

- um gene de silenciamento de XylT quimérico compreendendo:
 - o um fragmento incluindo a região promotora do transcrito 35S do Vírus Mosaico da Couve-flor (Odell e outros, 1985) (SEQ ID NO: 18 do nucleotídeo 969 ao nucleotídeo 2314)
 - o um fragmento incluindo uma parte do cDNA de XylT de *Nicotiana tabacum* cv. Petite Havana SR1 clonada na orientação sentido (SEQ ID NO: 18 do nucleotídeo 2365 ao nucleotídeo 2834)
 - o um fragmento contendo o íntron do gene de catase-1 de semente

de mamona (SEQ ID NO: 18 do nucleotídeo 2893 ao nucleotídeo 3088)

o um fragmento contendo o segundo íntron do gene de diquinase de ortofosfato de piruvato de *Flaveria trinervia* conforme descrito por Rosche e Westhoff (1995) na orientação reversa (SEQ ID NO: 18 do nucleotídeo 3130 ao nucleotídeo 3871).

o um fragmento incluindo uma parte da sequência de cDNA de XylIT de *Nicotiana tabacum* cv. Petite Havana SR1 clonada na orientação anti-sentido (SEQ ID NO: 18 do nucleotídeo 3957 ao nucleotídeo 4426).

o um fragmento incluindo a região não traduzida 3' do gene de sintase de octopina de *Agrobacterium tumefaciens* conforme descrito por De Greve e outros (1982) (SEQ ID NO: 18 do nucleotídeo 4479 ao nucleotídeo 5244).

- um gene quimérico que codifica um marcador selecionável compreendendo:

o um fragmento incluindo a região promotora do gene de sintase de nopalina de T-DNA de *Agrobacterium tumefaciens* (SEQ ID NO: 18 do nucleotídeo 5512 ao nucleotídeo 5744).

o um fragmento incluindo o gene de resistência a antibiótico *nptII* (SEQ ID NO: 18 do nucleotídeo 5745 ao nucleotídeo 6690).

o um fragmento incluindo a região não traduzida 3' do gene sintase de nopalina de T-DNA de *A. tumefaciens* (SEQ ID NO: 18 do nucleotídeo 6691 ao nucleotídeo 7396).

O vetor de T-DNA foi introduzido em *Agrobacterium tumefaciens* compreendendo um plasmídeo-Ti auxiliar.

4.2. Construção de um vetor de T-DNA compreendendo um gene de silenciamento de XylIT com um fragmento de DNA amplificado de uma sequência de XylIT de *Nicotiana benthamiana*

Seqüências de oligonucleotídeo a serem usadas como iniciadores não degenerados em uma, amplificação por PCR de uma sequência de gene XylIT de *Nicotiana benthamiana*, foram projetadas baseado na sequência de gene de *Nicotiana benthamiana* isolada no Exemplo 3. O iniciador dianteiro (SEQ ID NO: 19) foi projetado com GGCCGGATCCTCG em sua

extremidade 5' e o iniciador reservo (SEQ ID NO: 20) foi projetado com GGCCATCGATGGTACC em sua extremidade 5' para fins de clonagem.

Dessa forma, os seguintes iniciadores não degenerados foram gerados:

XylF9: 5'-GGCCGGATCCTCGAGACACAATTGGAGGAAACATGGAAAGC-3'
(SEQ ID NO: 19)

XylR9: 5'-GGCCATCGATGGTACCGGCCAGCTCTTTATGGAATCAAA -3'
(SEQ ID NO: 20)

Usando o vetor compreendendo um fragmento de DNA amplificado de uma sequência de gene de XylT de *Nicotiana benthamiana* conforme descrito no Exemplo 3 (SEQ ID NO: 11 - XylTg14Nb) como *padrão* e o par de iniciador XylF9 / XylR9, uma amplificação por PCR foi realizada sob as seguintes condições: 15 seg a 94°C (desnaturação), 30 seg a 58°C (anelamento) e 30 seg a 68°C (alongamento) durante 25 ciclos.

Um fragmento de DNA de cerca de 430 bp (XylTiNb; SEQ ID NO: 21) foi amplificado e digerido com XhoI e KpnI e com BamHI e ClaI, respectivamente. Os fragmentos digeridos com XhoI / KpnI e BamHI / ClaI foram clonados no pHANNIBAL (Helliwell e Waterhouse, 2003), digerido com XhoI / KpnI e BamHI / ClaI, proporcionando o pKW28.

O plasmídeo pKW28, assim, compreende um gene de silenciamento de XylT quimérico compreendendo:

o um fragmento incluindo a região promotora do transcrito 35S do vírus mosaico da couve-flor (Odell e outros, 1985) (SEQ ID NO: 22 do nucleotídeo 3779 ao nucleotídeo 2434)

o um fragmento incluindo uma parte da sequência do gene XylT de *Nicotiana benthamiana* clonado na orientação sentido (SEQ ID NO: 22 do nucleotídeo 2427 ao nucleotídeo 2023).

o um fragmento contendo o segundo íntron do gene de diquinase de ortofosfato de piruvato de *Flaveria trinervia* conforme descrito por Rosche e Westhoff (1995) (SEQ ID NO: 22 do nucleotídeo 1991 ao nucleotídeo 1250).

o um fragmento incluindo uma parte da sequência do gene de XylT de *Nicotiana benthamiana* clonado na orientação anti-sentido (SEQ ID NO: 22 do nucleotídeo 1211 ao nucleotídeo 807).

o um fragmento incluindo a região não traduzida 3' do gene de sintase de octopina de *Agrobacterium tumefaciens* conforme descrito por De Greve e outros (1982) (SEQ ID NO: 22 do nucleotídeo 786 ao nucleotídeo 76) entre os sítios de restrição MscI e PstI.

O plasmídeo pKW28 é digerido com MscI e PstI e o gene quimérico é introduzido entre as bordas de T-DNA de um vetor de T-DNA cortado com PstI e SmaI junto com um gene quimérico que codifica um marcador selecionado compreendendo:

o um fragmento incluindo a região promotora do gene de sintase de nopalina de T-DNA de *A. tumefaciens* T-DNA (SEQ ID NO: 22 do nucleotídeo 3854 ao nucleotídeo 4140).

o um fragmento incluindo o gene de resistência à fosfotricina *bar* (De Block e outros, 1987) (SEQ ID NO: 22 do nucleotídeo 4161 ao nucleotídeo 4712).

o um fragmento incluindo a região 3' não traduzida do gene de sintase de nopalina de T-DNA de *A. tumefaciens* (SEQ ID NO: 22 do nucleotídeo 4731 ao nucleotídeo 4991) para proporcionar o pTKW29 (seqüência do T-DNA de pTKW29 é representada em SEQ ID NO: 22).

O vetor pTKW29 é derivado do pGSC1700 (Cornelissen e Vandewiele, 1989). A cadeia principal do vetor contém os seguintes elementos genéticos:

- o núcleo do plasmídeo compreendendo a origem de replicação do plasmídeo pBR322 (Bolívar e outros, 1977) para replicação em *Escherichia coli* (ORI ColE1) e um fragmento de restrição compreendendo a origem de replicação do plasmídeo pVS1 de *Pseudomonas* (Itoh e outros, 1984) para replicação em *Agrobacterium tumefaciens* (ORI pVS1).

- um gene marcador selecionável que confere resistência à estreptomicina e espectinomicina (*aadA*) para propagação e seleção do plasmídeo em *Escherichia coli* e *Agrobacterium tumefaciens*.

- uma região de DNA consistindo em um fragmento da seqüência de codificação de fosfotransferase de neomicina do gene *nptII* do transposon *Tn903* (Oka e outros, 1981).

O vetor de T-DNA é introduzido em *Agrobacterium tumefaciens* compreendendo um plasmídeo-Ti auxiliar.

Exemplo 5: Análise de plantas transgênicas *Nicotiana* abrigando um gene de silenciamento de XylT.

Plantas *Nicotiana* foram transformadas usando as cepas de *Agrobacterium tumefaciens* descritas no Exemplo 4.

5.1. Análise de plantas transgênicas *Nicotiana tabacum* cv. Petite Havana SR1 abrigando um gene de silenciamento de XylT

Plantas *Nicotiana tabacum* cv. Petite Havana SR1 foram transformadas usando a cepa de *Agrobacterium tumefaciens* descrita no Exemplo 4.1. de acordo com o protocolo descrito em Zambryski e outros (1983). Cinquenta e duas linhagens transgênicas de *Nicotiana tabacum* compreendendo os genes quiméricos conforme descrito no Exemplo 4.1 foram obtidas.

Linhagens de planta transgênicas foram analisadas a nível molecular usando análise de Southern blot. Similarmente, as linhagens de planta foram analisadas com relação à expressão de RNA de XylT usando análise de Northern blot.

Uma indicação de atividade de XylT pode ser obtida comparando o nível de resíduos de beta-1,2-xilose presente sobre as glicanas de proteínas das linhagens transgênicas com aquele das plantas não-transformadas. O nível de resíduos de beta-1,2-xilose sobre N-glicanos ligadas à proteína de plantas pode ser medido, por exemplo, através de análise de Western blot usando anticorpos xilose-específicos conforme descrito, por exemplo, por Faye e outros (1993) ou através de espectrometria de massa sobre glicanas isoladas de glicoproteínas de planta usando Espectrometria de massa por tempo de voo da Ionização/Dessorção da matriz assistida por laser (MALDI-TOF-MS) conforme descrito, por exemplo, por Kolarich e Altmann (2000) ou usando espectrometria de massa Aleatória por Cromatografia de Líquido (LC/MS/MS) conforme descrito, por exemplo, por Henriksson e outros (2003).

5.2. Análise das plantas transgênicas *Nicotiana benthamiana* abrigando um gene de silenciamento de XylT

Similarmente, plantas *Nicotiana benthamiana* foram transformadas usando a cepa de *Agrobacterium tumefaciens* descrita no Exemplo 4.2. e a expressão do XylT e o nível de resíduos de beta-1,2-xilose presentes sobre as glicanas de proteína foram analisados conforme descrito acima.

Cinquenta e quatro linhagens transgênicas de *Nicotiana benthamiana* compreendendo os genes quiméricos descritos no Exemplo 4.2. foram obtidas após transformação do disco de folha com pTKW29.

Para determinar o nível de resíduos de beta-1,2-xilose presente sobre as glicanas de proteínas endógenas dessas linhagens de planta, proteínas de folha solúveis de cada indivíduo foram analisadas através de Western blot usando um anticorpo beta-1,2-xilose-específico.

Seis amostras mostraram uma reação muito fraca com o anticorpo e seis amostras não tinham reação detectável com o anticorpo. Para as outras amostras, o nível de reação com o anticorpo oscilava de um nível fraco ao tipo silvestre.

Para determinar o número de inserções do gene de silenciamento de XylT quimérico a partir do pTKW29, DNA genômico das linhagens de planta que mostraram reações muito fracas ou negativas ao anticorpo beta-1,2-xilose-específico foi isolado, digerido com EcoRI e analisado através de Southern blot usando uma sonda abrangendo a região do promotor 35S e uma sonda abrangendo a região de codificação do gene de resistência à fosfinotricina bar.

Nenhuma das doze linhagens de planta mostrou uma única inserção. Uma linhagem de planta continha duas inserções e era negativa para xilose usando análise por Western blot.

Para testar se esses dois genes de silenciamento de XylT quiméricos inseridos independentemente e obter plantas as quais eram negativas para xilose quando de Western blot e as quais contêm um único gene de silenciamento de XylT quimérico, a prole resultante de autofertilização da linhagem de planta contendo duas inserções foi plantada e vinte e cinco linhagens de planta foram analisadas através de análise por Western blot.

Listagem de Referência

- Baulcombe (2004). *Nature* 431: 356-363.
- Bernatzky e Tanksley (1986). *Theor. Appl. Genet.* 72: 314–321.
- Bolivar e outros (1977). *Gene* 2: 95-113.
- Cornelissen e Vandewiele (1989). *Nucleic Acids Research* 17: 19-25.
- De Block e outros (1984). *EMBO J.* 3(8): 1681-1689
- De Block e outros (1987). *EMBO J* 6: 2513.
- De Greve e outros (1982). *J. Mol. Appl. Genetics* 1 (6): 499-511.
- Faye e outros (1993). *Analytical Biochemistry* 209: 104-108.
- Fischer e outros (2004). *Curr. Opin. Plant Biol.* 7: 152–158.
- Helliwell e Waterhouse (2003). *Methods* 30: 289-295.
- Henikoff e Henikoff (1992). *Proc Natl Acad Sci USA* 89(22): 10915-10919.
- Henriksson e outros (2003). *Biochem. J.* 375: 61–73.
- Horsch e outros (1985). *Science* 227: 1229-1231.
- Itoh e outros (1984). *Plasmid* 11: 206.
- Khoo e outros (1997). *Glycobiology* 7: 663-677.
- Kolarich e Altmann (2000). *Anal. Biochem.* 285: 64-75.
- Mulder e outros (1995). *Eur. J. Biochem.* 232: 272-283.
- Odell e outros (1985). *Nature* 313: 810.
- Oka e outros (1981). *Journal de Molecular Biology*, 147, 217-226.
- Rosche e Westhoff (1995). *Plant Molecular Biology* 29 (4): 663-678.
- Stoger e outros (2002). *Curr. Opin. Biotechnol.* 13: 161–166.
- Strasser e outros (2000). *FEBS Lett.* 472: 105–108.
- Strasser e outros. (2004a). 2ª EPSO Conference, Outubro de 2004, Itália, Abstract book página 56, S036.
- Strasser e outros (2004b). *FEBS Lett.* 561: 132-136.
- Tezuka e outros (1992). *Eur J Biochem.* 203(3): 401-413.
- Twyman e outros (2003). *Trends Biotechnol.* 21: 570–578.
- Zambryski e outros (1983). *EMBO J.* 2: 2143–2150).
- Zeng e outros (1997). *J. Biol. Chem.* 272: 31340-31347.

LISTAGEM DE SEQUÊNCIA

<110> Bayer BioScience N.V.

Weterings, Koen

<120> **SEQUÊNCIAS DE NUCLEOTÍDEO QUE CODIFICAM *NICOTIANA*
BETA-1,2-XILOSILTRANSFERASE**

<130> BCS 06-2005

<160> 26

<170> PatentIn versão 3.3

<210> 1

<211> 37

<212> DNA

<213> Artificial

<220>

<223> oligonucleotídeo XylF4

<400> 1

caccttggtt ctctcttgcg tctcaactca atcaactc

37

<210> 2

<211> 27

<212> DNA

<213> Artificial

<220>

<223> oligonucleotídeo XylR4

<400> 2

tcgatcacaa ctggaggatc cgcataa

27

<210> 3

<211> 1513

<212> DNA

<213> Artificial

<220>

<223> XylTc2Nt

<220>

<221> CDS

<222> (7)..(1512)

<400> 3

cacctt gtt tct ctc ttc gct ctc aac tca atc act ctc tat ctc tac	48
Val Ser Leu Phe Ala Leu Asn Ser Ile Thr Leu Tyr Leu Tyr	
1 5 10	
ttc tct tcc cac tct gat cac ttc cgt cac aaa tcc ccc caa aac cac	96
Phe Ser Ser His Ser Asp His Phe Arg His Lys Ser Pro Gln Asn His	
15 20 25 30	
ttt cct aat acc caa aac cac tat tcc ctg tcg gaa aac cac cat gat	144
Phe Pro Asn Thr Gln Asn His Tyr Ser Leu Ser Glu Asn His His Asp	
35 40 45	
aat ttc cac tct tct gtc act tcc caa tat acc aag cct tgg cca att	192
Asn Phe His Ser Ser Val Thr Ser Gln Tyr Thr Lys Pro Trp Pro Ile	
50 55 60	
ttg ccc tcc tac ctc ccc tgg tct cag aat cct aat gtt tct ttg aga	240
Leu Pro Ser Tyr Leu Pro Trp Ser Gln Asn Pro Asn Val Ser Leu Arg	
65 70 75	
tcg tgc gag ggt tac ttc ggt aat ggg ttt act ctc aaa gtt gat ctt	288
Ser Cys Glu Gly Tyr Phe Gly Asn Gly Phe Thr Leu Lys Val Asp Leu	
80 85 90	
ctc aaa act tcg ccg gag ctt cac cag aaa ttc ggc gaa aac acc gta	336
Leu Lys Thr Ser Pro Glu Leu His Gln Lys Phe Gly Glu Asn Thr Val	
95 100 105 110	
tcc ggc gac ggc gga tgg ttt agg tgt ttt ttc agt gag act ttg cag	384
Ser Gly Asp Gly Gly Trp Phe Arg Cys Phe Phe Ser Glu Thr Leu Gln	
115 120 125	
agt tcg att tgc gag gga ggt gct ata cga atg aat ccg gac gag att	432
Ser Ser Ile Cys Glu Gly Gly Ala Ile Arg Met Asn Pro Asp Glu Ile	
130 135 140	
ttg atg tct cgt gga ggc gag aaa ttg gag tcg gtt att ggt agg agt	480
Leu Met Ser Arg Gly Gly Glu Lys Leu Glu Ser Val Ile Gly Arg Ser	
145 150 155	

gaa gat gat gag ctg ccc gtg ttc aaa aat gga gct ttt cag att aaa	528
Glu Asp Asp Glu Leu Pro Val Phe Lys Asn Gly Ala Phe Gln Ile Lys	
160 165 170	
ggt act gat aaa ctg aaa att ggg aaa aaa tta gtg gat gaa aaa atc	576
Val Thr Asp Lys Leu Lys Ile Gly Lys Lys Leu Val Asp Glu Lys Ile	
175 180 185 190	
ttg aat aaa tac tta ccg gaa ggt gca att tca agg cac act atg cgt	624
Leu Asn Lys Tyr Leu Pro Glu Gly Ala Ile Ser Arg His Thr Met Arg	
195 200 205	
gaa tta att gac tct att cag tta gtt ggc gcc gat gaa ttt cac tgt	672
Glu Leu Ile Asp Ser Ile Gln Leu Val Gly Ala Asp Glu Phe His Cys	
210 215 220	
tct gag tgg att gag gag ccg tca ctt ttg att aca cga ttt gag tat	720
Ser Glu Trp Ile Glu Glu Pro Ser Leu Leu Ile Thr Arg Phe Glu Tyr	
225 230 235	
gca aac ctt ttc cac aca gtt acc gat tgg tat agt gca tac gtg gca	768
Ala Asn Leu Phe His Thr Val Thr Asp Trp Tyr Ser Ala Tyr Val Ala	
240 245 250	
tcc agg gtt act ggc ttg ccc agt cgg cca cat ttg gtt ttt gta gat	816
Ser Arg Val Thr Gly Leu Pro Ser Arg Pro His Leu Val Phe Val Asp	
255 260 265 270	
ggc cat tgt gag aca caa ttg gag gaa aca tgg aaa gca ctc ttt tca	864
Gly His Cys Glu Thr Gln Leu Glu Glu Thr Trp Lys Ala Leu Phe Ser	
275 280 285	
agc ctc act tat gct aag aac ttt agt ggc cca gtt tgt ttc cgt cac	912
Ser Leu Thr Tyr Ala Lys Asn Phe Ser Gly Pro Val Cys Phe Arg His	
290 295 300	
gcc gtt ctc tcg cct ttg gga tat gaa act gcc ctg ttt aag gga ctg	960
Ala Val Leu Ser Pro Leu Gly Tyr Glu Thr Ala Leu Phe Lys Gly Leu	
305 310 315	
aca gaa act ata gat tgt aat gga gct tct gcc cat gat ttg tgg caa	1008
Thr Glu Thr Ile Asp Cys Asn Gly Ala Ser Ala His Asp Leu Trp Gln	

320	325	330	
aat cct gat gat aag aga act gca cgg ttg tcc gag ttt ggg gag atg			1056
Asn Pro Asp Asp Lys Arg Thr Ala Arg Leu Ser Glu Phe Gly Glu Met			
335	340	345	350
atc agg gca gcc ttt gga ttt cct gtg gat aga cag aac atc cca agg			1104
Ile Arg Ala Ala Phe Gly Phe Pro Val Asp Arg Gln Asn Ile Pro Arg			
355	360	365	
aca gtc aca ggc cct aat gtc ctc ttt gtt aga cgt gag gat tat tta			1152
Thr Val Thr Gly Pro Asn Val Leu Phe Val Arg Arg Glu Asp Tyr Leu			
370	375	380	
gct cac cca cgt cat ggt gga aag gta cag tct agg ctt agc aat gaa			1200
Ala His Pro Arg His Gly Gly Lys Val Gln Ser Arg Leu Ser Asn Glu			
385	390	395	
gag caa gta ttt gat tcc ata aag agc tgg gcc ttg aac cac tcg gag			1248
Glu Gln Val Phe Asp Ser Ile Lys Ser Trp Ala Leu Asn His Ser Glu			
400	405	410	
tgc aaa tta aat gta att aac gga ttg ttt gcc cac atg tcc atg aaa			1296
Cys Lys Leu Asn Val Ile Asn Gly Leu Phe Ala His Met Ser Met Lys			
415	420	425	430
gag caa gtt cga gca atc caa gat gca tct gtc att gtt ggt gct cat			1344
Glu Gln Val Arg Ala Ile Gln Asp Ala Ser Val Ile Val Gly Ala His			
435	440	445	
gga gca ggt cta act cac ata gtt tct gca gca cca aaa gct gta ata			1392
Gly Ala Gly Leu Thr His Ile Val Ser Ala Ala Pro Lys Ala Val Ile			
450	455	460	
cta gaa att ata agc agc gaa tat agg cac ccc cat ttt gct ctg att			1440
Leu Glu Ile Ile Ser Ser Glu Tyr Arg His Pro His Phe Ala Leu Ile			
465	470	475	
gca caa tgg aaa gga ttg gag tac cat ccc ata tat ttg gag ggg tct			1488
Ala Gln Trp Lys Gly Leu Glu Tyr His Pro Ile Tyr Leu Glu Gly Ser			
480	485	490	
tat gcg gat cct cca gtt gtg atg a			1513

Tyr Ala Asp Pro Pro Val Val Met

495

500

<210> 4

<211> 502

<212> PRT

<213> Artificial

<220>

<223> Estrutura sintética

<400> 4

Val Ser Leu Phe Ala Leu Asn Ser Ile Thr Leu Tyr Leu Tyr Phe Ser

1

5

10

15

Ser His Ser Asp His Phe Arg His Lys Ser Pro Gln Asn His Phe Pro

20

25

30

Asn Thr Gln Asn His Tyr Ser Leu Ser Glu Asn His His Asp Asn Phe

35

40

45

His Ser Ser Val Thr Ser Gln Tyr Thr Lys Pro Trp Pro Ile Leu Pro

50

55

60

Ser Tyr Leu Pro Trp Ser Gln Asn Pro Asn Val Ser Leu Arg Ser Cys

65

70

75

80

Glu Gly Tyr Phe Gly Asn Gly Phe Thr Leu Lys Val Asp Leu Leu Lys

85

90

95

Thr Ser Pro Glu Leu His Gln Lys Phe Gly Glu Asn Thr Val Ser Gly

100

105

110

Asp Gly Gly Trp Phe Arg Cys Phe Phe Ser Glu Thr Leu Gln Ser Ser

115

120

125

Ile Cys Glu Gly Gly Ala Ile Arg Met Asn Pro Asp Glu Ile Leu Met

130

135

140

Ser Arg Gly Gly Glu Lys Leu Glu Ser Val Ile Gly Arg Ser Glu Asp

145

150

155

160

Asp Glu Leu Pro Val Phe Lys Asn Gly Ala Phe Gln Ile Lys Val Thr

165

170

175

Asp Lys Leu Lys Ile Gly Lys Lys Leu Val Asp Glu Lys Ile Leu Asn
 180 185 190
 Lys Tyr Leu Pro Glu Gly Ala Ile Ser Arg His Thr Met Arg Glu Leu
 195 200 205
 Ile Asp Ser Ile Gln Leu Val Gly Ala Asp Glu Phe His Cys Ser Glu
 210 215 220
 Trp Ile Glu Glu Pro Ser Leu Leu Ile Thr Arg Phe Glu Tyr Ala Asn
 225 230 235 240
 Leu Phe His Thr Val Thr Asp Trp Tyr Ser Ala Tyr Val Ala Ser Arg
 245 250 255
 Val Thr Gly Leu Pro Ser Arg Pro His Leu Val Phe Val Asp Gly His
 260 265 270
 Cys Glu Thr Gln Leu Glu Glu Thr Trp Lys Ala Leu Phe Ser Ser Leu
 275 280 285
 Thr Tyr Ala Lys Asn Phe Ser Gly Pro Val Cys Phe Arg His Ala Val
 290 295 300
 Leu Ser Pro Leu Gly Tyr Glu Thr Ala Leu Phe Lys Gly Leu Thr Glu
 305 310 315 320
 Thr Ile Asp Cys Asn Gly Ala Ser Ala His Asp Leu Trp Gln Asn Pro
 325 330 335
 Asp Asp Lys Arg Thr Ala Arg Leu Ser Glu Phe Gly Glu Met Ile Arg
 340 345 350
 Ala Ala Phe Gly Phe Pro Val Asp Arg Gln Asn Ile Pro Arg Thr Val
 355 360 365
 Thr Gly Pro Asn Val Leu Phe Val Arg Arg Glu Asp Tyr Leu Ala His
 370 375 380
 Pro Arg His Gly Gly Lys Val Gln Ser Arg Leu Ser Asn Glu Glu Gln
 385 390 395 400
 Val Phe Asp Ser Ile Lys Ser Trp Ala Leu Asn His Ser Glu Cys Lys
 405 410 415
 Leu Asn Val Ile Asn Gly Leu Phe Ala His Met Ser Met Lys Glu Gln
 420 425 430

Val Arg Ala Ile Gln Asp Ala Ser Val Ile Val Gly Ala His Gly Ala

435

440

445

Gly Leu Thr His Ile Val Ser Ala Ala Pro Lys Ala Val Ile Leu Glu

450

455

460

Ile Ile Ser Ser Glu Tyr Arg His Pro His Phe Ala Leu Ile Ala Gln

465

470

475

480

Trp Lys Gly Leu Glu Tyr His Pro Ile Tyr Leu Glu Gly Ser Tyr Ala

485

490

495

Asp Pro Pro Val Val Met

500

<210> 5

<211> 1495

<212> DNA

<213> Artificial

<220>

<223> XylTc7Nt

<220>

<221> CDS

<222> (5)..(1495)

<400> 5

cacc ctt gtt tct ctc ttc gct ctc aac tca atc act ctc tat ctc tac 49

Leu Val Ser Leu Phe Ala Leu Asn Ser Ile Thr Leu Tyr Leu Tyr

1

5

10

15

ttc tct tcc cac cct gat cac ttc cgc cac aaa tcc cgc caa aac cac 97

Phe Ser Ser His Pro Asp His Phe Arg His Lys Ser Arg Gln Asn His

20

25

30

ttt tcc ttg tgc gaa aac cgc cat cat aat ttc cac tct tca atc act 145

Phe Ser Leu Ser Glu Asn Arg His His Asn Phe His Ser Ser Ile Thr

35

40

45

tct caa tat tcc aag cct tgg cct att ttg ccc tcc tac ctc cct tgg 193

Ser Gln Tyr Ser Lys Pro Trp Pro Ile Leu Pro Ser Tyr Leu Pro Trp

50

55

60

tct caa aac cct aat gtt gtt tgg aga tcg tgc gag ggt tac ttc ggt	241
Ser Gln Asn Pro Asn Val Val Trp Arg Ser Cys Glu Gly Tyr Phe Gly	
65 70 75	
aat ggg ttt act ctc aaa gtt gac ctt ctc aaa act tcg ccg gag ttt	289
Asn Gly Phe Thr Leu Lys Val Asp Leu Leu Lys Thr Ser Pro Glu Phe	
80 85 90 95	
cac cgg aaa ttc ggc gaa aac acc gtc tcc ggc gac ggc gga tgg ttt	337
His Arg Lys Phe Gly Glu Asn Thr Val Ser Gly Asp Gly Gly Trp Phe	
100 105 110	
agg tgt ttt ttc agt gag act ttg cag agt tcg atc tgc gag gga ggc	385
Arg Cys Phe Phe Ser Glu Thr Leu Gln Ser Ser Ile Cys Glu Gly Gly	
115 120 125	
gca ata cga atg aat ccg gac gag att ttg atg tct cgt gga ggt gag	433
Ala Ile Arg Met Asn Pro Asp Glu Ile Leu Met Ser Arg Gly Gly Glu	
130 135 140	
aaa ttg gag tcg gtt att ggt agg agt gaa gat gat gag ctg ccc gtg	481
Lys Leu Glu Ser Val Ile Gly Arg Ser Glu Asp Asp Glu Leu Pro Val	
145 150 155	
ttc aaa aat gga gct ttt cag att aaa gtt act gat aaa ctg aaa att	529
Phe Lys Asn Gly Ala Phe Gln Ile Lys Val Thr Asp Lys Leu Lys Ile	
160 165 170 175	
ggg aaa aaa tta gtg gat gaa aaa ttc ttg aat aaa tac tta ccg gaa	577
Gly Lys Lys Leu Val Asp Glu Lys Phe Leu Asn Lys Tyr Leu Pro Glu	
180 185 190	
ggg gca att tca agg cac act atg cgt gag tta att gac tct att cag	625
Gly Ala Ile Ser Arg His Thr Met Arg Glu Leu Ile Asp Ser Ile Gln	
195 200 205	
ttg gtt ggc gcc gat gat ttt cac tgt tct gag tgg att gag gag ccg	673
Leu Val Gly Ala Asp Asp Phe His Cys Ser Glu Trp Ile Glu Glu Pro	
210 215 220	
tca ctt ttg att aca cga ttt gag tat gca aac ctt ttc cac aca gtt	721

Ser Leu Leu Ile Thr Arg Phe Glu Tyr Ala Asn Leu Phe His Thr Val
 225 230 235
 acc gat tgg tat agt gca tac gtg gca tcc agg gtt act ggc ttg ccc 769
 Thr Asp Trp Tyr Ser Ala Tyr Val Ala Ser Arg Val Thr Gly Leu Pro
 240 245 250 255
 agt cgg cca cat ttg gtt ttt gta gat ggc cat tgt gag aca caa ttg 817
 Ser Arg Pro His Leu Val Phe Val Asp Gly His Cys Glu Thr Gln Leu
 260 265 270
 gag gaa aca tgg aaa gca ctt ttt tca agc ctc act tat gct aag aac 865
 Glu Glu Thr Trp Lys Ala Leu Phe Ser Ser Leu Thr Tyr Ala Lys Asn
 275 280 285
 ttt agt ggc cca gtt tgt ttc cgt cat gcc gcc ctc tcg cct ttg gga 913
 Phe Ser Gly Pro Val Cys Phe Arg His Ala Ala Leu Ser Pro Leu Gly
 290 295 300
 tat gaa act gcc ctg ttt aag gga ctg tca gaa act ata gat tgt aat 961
 Tyr Glu Thr Ala Leu Phe Lys Gly Leu Ser Glu Thr Ile Asp Cys Asn
 305 310 315
 gga gct tct gcc cat gat ttg tgg caa aat cct gat gat aag aaa act 1009
 Gly Ala Ser Ala His Asp Leu Trp Gln Asn Pro Asp Asp Lys Lys Thr
 320 325 330 335
 gca cgg ttg tcc gag ttt ggg gag atg att agg gca gcc ttt gga ttt 1057
 Ala Arg Leu Ser Glu Phe Gly Glu Met Ile Arg Ala Ala Phe Gly Phe
 340 345 350
 cct gtg gat aga cag aac atc cca agg aca gtc aca ggc cct aat gtc 1105
 Pro Val Asp Arg Gln Asn Ile Pro Arg Thr Val Thr Gly Pro Asn Val
 355 360 365
 ctc ttt gtt aga cgt gag gat tat tta gct cac cca cgt cat ggt gga 1153
 Leu Phe Val Arg Arg Glu Asp Tyr Leu Ala His Pro Arg His Gly Gly
 370 375 380
 aag gta cag tct agg ctt agc aat gaa gag caa gta ttt gat tcc ata 1201
 Lys Val Gln Ser Arg Leu Ser Asn Glu Glu Gln Val Phe Asp Ser Ile
 385 390 395

aag agc tgg gcc ttg aac cac tcg gag tgc aaa tta aat gta att aac 1249
Lys Ser Trp Ala Leu Asn His Ser Glu Cys Lys Leu Asn Val Ile Asn
400 405 410 415
gga ttg ttt gcc cac atg tcc atg aaa gag caa gtt cga gca atc caa 1297
Gly Leu Phe Ala His Met Ser Met Lys Glu Gln Val Arg Ala Ile Gln
420 425 430
gat gct tct gtc ata gtt ggt gct cat gga gca ggt cta act cac ata 1345
Asp Ala Ser Val Ile Val Gly Ala His Gly Ala Gly Leu Thr His Ile
435 440 445
gtt tct gca gca cca aaa gct gta ata cta gaa att ata agc agc gaa 1393
Val Ser Ala Ala Pro Lys Ala Val Ile Leu Glu Ile Ile Ser Ser Glu
450 455 460
tat agg cgc ccc cat ttt gct ctg att gca caa tgg aaa gga ttg gag 1441
Tyr Arg Arg Pro His Phe Ala Leu Ile Ala Gln Trp Lys Gly Leu Glu
465 470 475
tac cat ccc ata tat ttg gag ggg tct tat gcg gat cct cca gtt gtg 1489
Tyr His Pro Ile Tyr Leu Glu Gly Ser Tyr Ala Asp Pro Pro Val Val
480 485 490 495
atc gaa 1495
Ile Glu

<210> 6

<211> 497

<212> PRT

<213> Artificial

<220>

<223> Estrutura sintética

<400> 6

Leu Val Ser Leu Phe Ala Leu Asn Ser Ile Thr Leu Tyr Leu Tyr Phe
1 5 10 15
Ser Ser His Pro Asp His Phe Arg His Lys Ser Arg Gln Asn His Phe
20 25 30
Ser Leu Ser Glu Asn Arg His His Asn Phe His Ser Ser Ile Thr Ser

35	40	45
Gln Tyr Ser Lys Pro Trp	Pro Ile Leu Pro Ser Tyr Leu Pro Trp Ser	
50	55	60
Gln Asn Pro Asn Val Val Trp Arg Ser Cys Glu Gly Tyr Phe Gly Asn		
65	70	75
Gly Phe Thr Leu Lys Val Asp Leu Leu Lys Thr Ser Pro Glu Phe His		
85	90	95
Arg Lys Phe Gly Glu Asn Thr Val Ser Gly Asp Gly Gly Trp Phe Arg		
100	105	110
Cys Phe Phe Ser Glu Thr Leu Gln Ser Ser Ile Cys Glu Gly Gly Ala		
115	120	125
Ile Arg Met Asn Pro Asp Glu Ile Leu Met Ser Arg Gly Gly Glu Lys		
130	135	140
Leu Glu Ser Val Ile Gly Arg Ser Glu Asp Asp Glu Leu Pro Val Phe		
145	150	155
Lys Asn Gly Ala Phe Gln Ile Lys Val Thr Asp Lys Leu Lys Ile Gly		
165	170	175
Lys Lys Leu Val Asp Glu Lys Phe Leu Asn Lys Tyr Leu Pro Glu Gly		
180	185	190
Ala Ile Ser Arg His Thr Met Arg Glu Leu Ile Asp Ser Ile Gln Leu		
195	200	205
Val Gly Ala Asp Asp Phe His Cys Ser Glu Trp Ile Glu Glu Pro Ser		
210	215	220
Leu Leu Ile Thr Arg Phe Glu Tyr Ala Asn Leu Phe His Thr Val Thr		
225	230	235
Asp Trp Tyr Ser Ala Tyr Val Ala Ser Arg Val Thr Gly Leu Pro Ser		
245	250	255
Arg Pro His Leu Val Phe Val Asp Gly His Cys Glu Thr Gln Leu Glu		
260	265	270
Glu Thr Trp Lys Ala Leu Phe Ser Ser Leu Thr Tyr Ala Lys Asn Phe		
275	280	285
Ser Gly Pro Val Cys Phe Arg His Ala Ala Leu Ser Pro Leu Gly Tyr		

290	295	300	
Glu Thr Ala Leu Phe Lys Gly Leu Ser Glu Thr Ile Asp Cys Asn Gly			
305	310	315	320
Ala Ser Ala His Asp Leu Trp Gln Asn Pro Asp Asp Lys Lys Thr Ala			
	325	330	335
Arg Leu Ser Glu Phe Gly Glu Met Ile Arg Ala Ala Phe Gly Phe Pro			
	340	345	350
Val Asp Arg Gln Asn Ile Pro Arg Thr Val Thr Gly Pro Asn Val Leu			
	355	360	365
Phe Val Arg Arg Glu Asp Tyr Leu Ala His Pro Arg His Gly Gly Lys			
	370	375	380
Val Gln Ser Arg Leu Ser Asn Glu Glu Gln Val Phe Asp Ser Ile Lys			
385	390	395	400
Ser Trp Ala Leu Asn His Ser Glu Cys Lys Leu Asn Val Ile Asn Gly			
	405	410	415
Leu Phe Ala His Met Ser Met Lys Glu Gln Val Arg Ala Ile Gln Asp			
	420	425	430
Ala Ser Val Ile Val Gly Ala His Gly Ala Gly Leu Thr His Ile Val			
	435	440	445
Ser Ala Ala Pro Lys Ala Val Ile Leu Glu Ile Ile Ser Ser Glu Tyr			
	450	455	460
Arg Arg Pro His Phe Ala Leu Ile Ala Gln Trp Lys Gly Leu Glu Tyr			
465	470	475	480
His Pro Ile Tyr Leu Glu Gly Ser Tyr Ala Asp Pro Pro Val Val Ile			
	485	490	495
Glu			

<210> 7

<211> 3408

<212> DNA

<213> Artificial

<220>

<223> XylTg1Nt

<220>

<221> CDS

<222> (7)..(678)

<220>

<221> Intron

<222> (679)..(1974)

<220>

<221> CDS

<222> (1975)..(2124)

<220>

<221> Intron

<222> (2125)..(2722)

<220>

<221> CDS

<222> (2723)..(3406)

<400> 7

cacctt gtt tct ctc ttc gct ctc aac tca atc act ctc tat ctc tac 48

Val Ser Leu Phe Ala Leu Asn Ser Ile Thr Leu Tyr Leu Tyr

1 5 10

ttc tct tcc cac tct gat cac ttc cgt cac aaa tcc ccc caa aac cac 96

Phe Ser Ser His Ser Asp His Phe Arg His Lys Ser Pro Gln Asn His

15 20 25 30

ttt cct aat acc caa aac cac tat tcc ctg tcg gaa aac cac cat gat 144

Phe Pro Asn Thr Gln Asn His Tyr Ser Leu Ser Glu Asn His His Asp

35 40 45

aat ttc cac tct tct gtc act tcc caa tat acc aag cct tgg cca att 192

Asn Phe His Ser Ser Val Thr Ser Gln Tyr Thr Lys Pro Trp Pro Ile

50 55 60

ttg ccc tcc tac ctc ccc tgg tct cag aat cct aat gtt tct ttg aga 240

Leu Pro Ser Tyr Leu Pro Trp Ser Gln Asn Pro Asn Val Ser Leu Arg

65 70 75

tcg tgc gag ggt tac ttc ggt aat ggg ttt act ctc aaa gtt gat ctt	288
Ser Cys Glu Gly Tyr Phe Gly Asn Gly Phe Thr Leu Lys Val Asp Leu	
80 85 90	
ctc aaa act tcg ccg gag ctt cac cag aaa ttc ggc gaa aac acc gta	336
Leu Lys Thr Ser Pro Glu Leu His Gln Lys Phe Gly Glu Asn Thr Val	
95 100 105 110	
tcc ggc gac ggc gga tgg ttt agg tgt ttt ttc agt gag act ttg cag	384
Ser Gly Asp Gly Gly Trp Phe Arg Cys Phe Phe Ser Glu Thr Leu Gln	
115 120 125	
agt tcg att tgc gag gga ggt gct ata cga atg aat ccg gac gag att	432
Ser Ser Ile Cys Glu Gly Gly Ala Ile Arg Met Asn Pro Asp Glu Ile	
130 135 140	
ttg atg tct cgt gga ggc gag aaa ttg gag tcg gtt att ggt agg agt	480
Leu Met Ser Arg Gly Gly Glu Lys Leu Glu Ser Val Ile Gly Arg Ser	
145 150 155	
gaa gat gat gag ctg ccc gtg ttc aaa aat gga gct ttt cag att aaa	528
Glu Asp Asp Glu Leu Pro Val Phe Lys Asn Gly Ala Phe Gln Ile Lys	
160 165 170	
gtt act gat aaa ctg aaa att ggg aaa aaa tta gtg gat gaa aaa atc	576
Val Thr Asp Lys Leu Lys Ile Gly Lys Lys Leu Val Asp Glu Lys Ile	
175 180 185 190	
ttg aat aaa tac tta ccg gaa ggt gca att tca agg cac act atg cgt	624
Leu Asn Lys Tyr Leu Pro Glu Gly Ala Ile Ser Arg His Thr Met Arg	
195 200 205	
gaa tta att gac tct att cag tta gtt ggc gcc gat gaa ttt cac tgt	672
Glu Leu Ile Asp Ser Ile Gln Leu Val Gly Ala Asp Glu Phe His Cys	
210 215 220	
tct gag gtttagatttt gaaattttgc ttgatcttta aattaaaggt ttgaactttg	728
Ser Glu	
tgaatgttgg cagatatgga atacaataat ggattttggt tgatctgttt aatgaagatt	788
gtctagaacc tcattgttat aaatatgggt tgtttgcttc attaattaa gagcattcct	848

taaaatctcg actagatgcc agataacacc agttagttga cttttggatt atggattttt 908
 ttccatttaa tcagataaga tagtcattct taaatgtttc actaaagaat ttgtcatgat 968
 ttcagtgtat atctttaagt gtatttggaa ttttggattt ggatctagta ctgaatgggt 1028
 aactgcactt gtactcccca gtcactctggg gaggagcaac agattaaatt caaggggtga 1088
 aaagtaatac agagtcagaa attaaccaca agttggaaaa tggtaaagt atgtggtcta 1148
 agatgattac tcctataact tttgatgtct aacatggaga aagttagtgt atttatgtct 1208
 tttacttttc cttttattga ttttggtttt caaattctat caattccttt gtttgattgc 1268
 tactcagatt gaaccttaga cggagtagca atagaaaagt gaagaaaagc cattttttct 1328
 cttttcatct cttttattct gttttacaca cagaatatgg tagcatctgt ctgaactagt 1388
 taattttatt ctttaaaatt tgcataacta attcgagtaa atgccttttg aagctttagt 1448
 tgtacaactg gttgttgcatt tttgaggact atcgacttga tttgacagtg tctgatacat 1508
 ggcttgtaag ttatgaaaac ttatatctag gaagaaatcc caaccagaga tagggagctg 1568
 tcgcttggtt atgagctact ggctcaaagt tgcagtttga ccagttaatt ttagatcttc 1628
 ggagtctaatt caaattctga agcagtattg tgcactaata agaggaacac atcaaggatg 1688
 tagcactgcc aggttatgtt accttatppta ctaatgattg agaaccagct taaatgatga 1748
 caaatgggtct tagatttggt ttttacattg ctcatgactt tgggatattt ctggatcaac 1808
 attttccagt tctttatgta cttatcaaaa aattatccct gctagagcct agatgttagt 1868
 gttcaagcaa ccatgctagc ttttaaggaa gctccttctt tgattcatgc catctttccg 1928
 taatcgatgc cttacgttac tgtcattttt ctaattttca tttcag tgg att gag 1983

Trp Ile Glu

225

gag ccg tca ctt ttg att aca cga ttt gag tat gca aac ctt ttc cac 2031

Glu Pro Ser Leu Leu Ile Thr Arg Phe Glu Tyr Ala Asn Leu Phe His

230

235

240

aca gtt acc gat tgg tat agt gca tac gtg gca tcc agg gtt act ggc 2079

Thr Val Thr Asp Trp Tyr Ser Ala Tyr Val Ala Ser Arg Val Thr Gly

245

250

255

ttg ccc agt cgg cca cat ttg gtt ttt gta gat ggc cat tgt gag 2124

Leu Pro Ser Arg Pro His Leu Val Phe Val Asp Gly His Cys Glu

260

265

270

gtatgtttga aagtattgat aacgatggca tgcattgtac tgtgttatgg atgaaagaaa 2184

tgaaatcagc aattattttc tagcaggcaa tgctcttgag atgcttgtgt caaattggtc 2244

gat tcc ata aag agc tgg gcc ttg aac cac tcg gag tgc aaa tta aat 3154
Asp Ser Ile Lys Ser Trp Ala Leu Asn His Ser Glu Cys Lys Leu Asn
405 410 415
gta att aac gga ttg ttt gcc cac atg tcc atg aaa gag caa gtt cga 3202
Val Ile Asn Gly Leu Phe Ala His Met Ser Met Lys Glu Gln Val Arg
420 425 430
gca atc caa gat gca tct gtc att gtc ggc gct cat gga gca ggt cta 3250
Ala Ile Gln Asp Ala Ser Val Ile Val Gly Ala His Gly Ala Gly Leu
435 440 445 450
act cac ata gtt tct gca gca cca aaa gct gta ata cta gaa att ata 3298
Thr His Ile Val Ser Ala Ala Pro Lys Ala Val Ile Leu Glu Ile Ile
455 460 465
agc agc gaa tat agg cgc ccc cat ttt gct ctg att gca caa tgg aaa 3346
Ser Ser Glu Tyr Arg Arg Pro His Phe Ala Leu Ile Ala Gln Trp Lys
470 475 480
gga ttg gag tac cat ccc ata tat ttg gag ggg tct tat gcg gat cct 3394
Gly Leu Glu Tyr His Pro Ile Tyr Leu Glu Gly Ser Tyr Ala Asp Pro
485 490 495
cca gtt gtg atc ga 3408
Pro Val Val Ile
500

<210> 8

<211> 502

<212> PRT

<213> Artificial

<220>

<223> Estrutura sintética

<400> 8

Val Ser Leu Phe Ala Leu Asn Ser Ile Thr Leu Tyr Leu Tyr Phe Ser
1 5 10 15
Ser His Ser Asp His Phe Arg His Lys Ser Pro Gln Asn His Phe Pro

20	25	30
Asn Thr Gln Asn His Tyr Ser Leu Ser Glu Asn His His Asp Asn Phe		
35	40	45
His Ser Ser Val Thr Ser Gln Tyr Thr Lys Pro Trp Pro Ile Leu Pro		
50	55	60
Ser Tyr Leu Pro Trp Ser Gln Asn Pro Asn Val Ser Leu Arg Ser Cys		
65	70	75
Glu Gly Tyr Phe Gly Asn Gly Phe Thr Leu Lys Val Asp Leu Leu Lys		
85	90	95
Thr Ser Pro Glu Leu His Gln Lys Phe Gly Glu Asn Thr Val Ser Gly		
100	105	110
Asp Gly Gly Trp Phe Arg Cys Phe Phe Ser Glu Thr Leu Gln Ser Ser		
115	120	125
Ile Cys Glu Gly Gly Ala Ile Arg Met Asn Pro Asp Glu Ile Leu Met		
130	135	140
Ser Arg Gly Gly Glu Lys Leu Glu Ser Val Ile Gly Arg Ser Glu Asp		
145	150	155
Asp Glu Leu Pro Val Phe Lys Asn Gly Ala Phe Gln Ile Lys Val Thr		
165	170	175
Asp Lys Leu Lys Ile Gly Lys Lys Leu Val Asp Glu Lys Ile Leu Asn		
180	185	190
Lys Tyr Leu Pro Glu Gly Ala Ile Ser Arg His Thr Met Arg Glu Leu		
195	200	205
Ile Asp Ser Ile Gln Leu Val Gly Ala Asp Glu Phe His Cys Ser Glu		
210	215	220
Trp Ile Glu Glu Pro Ser Leu Leu Ile Thr Arg Phe Glu Tyr Ala Asn		
225	230	235
Leu Phe His Thr Val Thr Asp Trp Tyr Ser Ala Tyr Val Ala Ser Arg		
245	250	255
Val Thr Gly Leu Pro Ser Arg Pro His Leu Val Phe Val Asp Gly His		
260	265	270
Cys Glu Thr Gln Leu Glu Glu Thr Trp Lys Ala Leu Phe Ser Ser Leu		

275	280	285	
Thr Tyr Ala Lys Asn Phe Ser Gly Pro Val Cys Phe Arg His Ala Val			
290	295	300	
Leu Ser Pro Leu Gly Tyr Glu Thr Ala Leu Phe Lys Gly Leu Thr Glu			
305	310	315	320
Thr Ile Asp Cys Asn Gly Ala Ser Ala His Asp Leu Trp Gln Asn Pro			
	325	330	335
Asp Asp Lys Arg Thr Ala Arg Leu Ser Glu Phe Gly Glu Met Ile Arg			
	340	345	350
Ala Ala Phe Gly Phe Pro Val Asp Arg Gln Asn Ile Pro Arg Thr Val			
	355	360	365
Thr Gly Pro Asn Val Leu Phe Val Arg Arg Glu Asp Tyr Leu Ala His			
	370	375	380
Pro Arg His Gly Gly Lys Val Gln Ser Arg Leu Ser Asn Glu Glu Gln			
385	390	395	400
Val Phe Asp Ser Ile Lys Ser Trp Ala Leu Asn His Ser Glu Cys Lys			
	405	410	415
Leu Asn Val Ile Asn Gly Leu Phe Ala His Met Ser Met Lys Glu Gln			
	420	425	430
Val Arg Ala Ile Gln Asp Ala Ser Val Ile Val Gly Ala His Gly Ala			
	435	440	445
Gly Leu Thr His Ile Val Ser Ala Ala Pro Lys Ala Val Ile Leu Glu			
	450	455	460
Ile Ile Ser Ser Glu Tyr Arg Arg Pro His Phe Ala Leu Ile Ala Gln			
465	470	475	480
Trp Lys Gly Leu Glu Tyr His Pro Ile Tyr Leu Glu Gly Ser Tyr Ala			
	485	490	495
Asp Pro Pro Val Val Ile			
	500		

<210> 9

<211> 3385

<212> DNA

<213> Artificial

<220>

<223> XylTg3Nt

<220>

<221> CDS

<222> (7)..(657)

<220>

<221> Intron

<222> (658)..(1953)

<220>

<221> CDS

<222> (1954)..(2103)

<220>

<221> Intron

<222> (2104)..(2701)

<220>

<221> CDS

<222> (2702)..(3385)

<400> 9

cacctt gtt tct ctc ttc gct ctc aac tca atc act ctc tat ctc tac 48

Val Ser Leu Phe Ala Leu Asn Ser Ile Thr Leu Tyr Leu Tyr

1 5 10

ttc tct tcc cac cct gat cac ttc cgc cac aaa tcc cgc caa aac cac 96

Phe Ser Ser His Pro Asp His Phe Arg His Lys Ser Arg Gln Asn His

15 20 25 30

ttt tcc ttg tcg gaa aac cgc cat cat aat ttc cac tct tca atc act 144

Phe Ser Leu Ser Glu Asn Arg His His Asn Phe His Ser Ser Ile Thr

35 40 45

tct caa tat tcc aag cct tgg cct att ttg ccc tcc tac ctc cct tgg 192

Ser Gln Tyr Ser Lys Pro Trp Pro Ile Leu Pro Ser Tyr Leu Pro Trp

50 55 60

tct caa aac cct aat gtt gtt tgg aga tcg tgc gag ggt tac ttc ggt	240
Ser Gln Asn Pro Asn Val Val Trp Arg Ser Cys Glu Gly Tyr Phe Gly	
65 70 75	
aat ggg ttt act ctc aaa gtt gac ctt ctc aaa act tcg ccg gag ttt	288
Asn Gly Phe Thr Leu Lys Val Asp Leu Leu Lys Thr Ser Pro Glu Phe	
80 85 90	
cac cgg aaa ttc ggc gaa aac acc gtc tcc ggc gac ggc gga tgg ttt	336
His Arg Lys Phe Gly Glu Asn Thr Val Ser Gly Asp Gly Gly Trp Phe	
95 100 105 110	
agg tgt ttt ttc agt gag act ttg cag agt tcg atc tgc gag gga ggc	384
Arg Cys Phe Phe Ser Glu Thr Leu Gln Ser Ser Ile Cys Glu Gly Gly	
115 120 125	
gca ata cga atg aat ccg gac gag att ttg atg tct cgt gga ggt gag	432
Ala Ile Arg Met Asn Pro Asp Glu Ile Leu Met Ser Arg Gly Gly Glu	
130 135 140	
aaa ttg gag tcg gtt att ggt agg agt gaa gat gat gag ctg ccc gtg	480
Lys Leu Glu Ser Val Ile Gly Arg Ser Glu Asp Asp Glu Leu Pro Val	
145 150 155	
ttc aaa aat gga gct ttt cag att aaa gtt act gat aaa ctg aaa att	528
Phe Lys Asn Gly Ala Phe Gln Ile Lys Val Thr Asp Lys Leu Lys Ile	
160 165 170	
ggg aaa aaa tta gtg gat gaa aaa atc ttg aat aaa tac tta ccg gaa	576
Gly Lys Lys Leu Val Asp Glu Lys Ile Leu Asn Lys Tyr Leu Pro Glu	
175 180 185 190	
ggt gca att tca agg cac act atg cgt gaa tta att gac tct att cag	624
Gly Ala Ile Ser Arg His Thr Met Arg Glu Leu Ile Asp Ser Ile Gln	
195 200 205	
tta gtt ggc gcc gat gaa ttt cac tgt tct gag gtttagat ttt gaaattttgc	677
Leu Val Gly Ala Asp Glu Phe His Cys Ser Glu	
210 215	
ttgatcttta aattaaaggt ttgaactttg tgaatgttgg cagatatgga atacaataat	737
ggattttgtt tgatctgttt aatgaagatt gtctagaacc tcattgttat aaatatggtt	797

tgtttgcttc attaattaaa gagcattcct taaaatctcg actagatgcc agataacacc 857
 agttagtga cttttggatt atggattttt ttccatttaa tcagataaga tagtcattct 917
 taaatgtttc actaaagaat ttgtcatgat ttcagtgtat atctttaagt gtatttgga 977
 ttttggaattt ggatctagta ctgaatgggt aactgcactt gtactcccca gtcactctggg 1037
 gaggagcaac agattaaatt caaggggtga aaagtaatac agagtcagaa attaaccaca 1097
 agttggaaaa tggtaaatgt atgtggtcta agatgattac tcctataact tttgatgtct 1157
 aacatggaga aagttagtgt atttatgtct ttacttttc cctttattga ttttggtttt 1217
 caaattctat caattccttt gtttgattgc tactcagatt gaaccttaga cggagtagca 1277
 atagaaaagt gaagaaaagc ctttttttct cctttcatct ctttatttct gttttacaca 1337
 cagaatatgg tagcatctgt ctgaactagt taattttatt ccttaaaatt tgcataacta 1397
 attcgagtaa atgccttttg aagctttagt tgtacaactg gttggtgcat tttgaggact 1457
 atcgacttga tttgacagtg tctgatacat ggcttgtaag ttatgaaaac ttatatctag 1517
 gaagaaatcc caaccagaga tagggagctg tcgcttggtt atgagctact ggctcaaagt 1577
 tcgagtttga ccagttaatt ttagatcttc ggagtctaata caaattctga agcagtattg 1637
 tgcactaata agaggaacac atcaaggatg tagcactgcc aggttatggt accttattta 1697
 ctaatgattg agaaccagct taaatgatga caaatggtct tagatttggt ttttacattg 1757
 ctcatgactt tgggatattt ctggatcaac attttccagt tctttatgta cttatcaaaa 1817
 aattatccct gctagagcct agatgttagt gttcaagcaa ccatgctagc ttttaaggaa 1877
 gctccttctt tgattcatgc catctttccg taatcgatgc cttacgttac tgtcattttt 1937
 ctaattttca tttcag tgg att gag gag ccg tca ctt ttg att aca cga ttt 1989

Trp Ile Glu Glu Pro Ser Leu Leu Ile Thr Arg Phe

220

225

gag tat gca aac ctt ttc cac aca gtt acc gat tgg tat agt gca tac 2037
 Glu Tyr Ala Asn Leu Phe His Thr Val Thr Asp Trp Tyr Ser Ala Tyr

230

235

240

245

gtg gca tcc agg gtt act ggc ttg ccc agt cgg cca cat ttg gtt ttt 2085
 Val Ala Ser Arg Val Thr Gly Leu Pro Ser Arg Pro His Leu Val Phe

250

255

260

gta gat ggc cat tgt gag gtatgtttga aagtattgat aacgatggca 2133
 Val Asp Gly His Cys Glu

265

tgcatgtac tgtgttatgg atgaaagaaa tgaaatcagc aattattttc tagcaggcaa 2193

tgctcttgag atgcttgtgt caaattggtc agacttaatc ctgagtttcc atttgtttca	2253
gctttctgtt tgactgacta caataattgt cccaatacct agttgttccg gttggctcat	2313
tcttacttct atttacgtgt cactgtttct ctgaatggtc cctttgtggt gaaaagtgc	2373
tttgctatga agaaaaacta gcaaagattt catttctggg acaatttatt tttaccttac	2433
atcacgtctc ataaaattgc ttcttattgc atactttaat tcttgagag atgctttaat	2493
gtgaagaaag ttctttcact ccacttttaa ttcttgaga gatgctttaa tgtgaagaaa	2553
gttctttcac tccactactg taagcttgct gcatgaattt tacttgcca tattgggggc	2613
gtgttttgat ttatcttcaa attcatttct ttcatgtaat tcttttgagt aattttttt	2673
tctgttttct gtttgggaaa aaattcag aca caa ttg gag gaa aca tgg aaa	2725
Thr Gln Leu Glu Glu Thr Trp Lys	
270	275
gca ctc ttt tca agc ctc act tat gct aag aac ttt agt ggc cca gtt	2773
Ala Leu Phe Ser Ser Leu Thr Tyr Ala Lys Asn Phe Ser Gly Pro Val	
280	285
290	
tgt ttc cgt cac gcc gtt ctc tcg cct ttg gga tat gaa act gcc ctg	2821
Cys Phe Arg His Ala Val Leu Ser Pro Leu Gly Tyr Glu Thr Ala Leu	
295	300
305	
ttt aag gga ctg aca gaa act ata gat tgt aat gga gct tct gcc cat	2869
Phe Lys Gly Leu Thr Glu Thr Ile Asp Cys Asn Gly Ala Ser Ala His	
310	315
320	
gat ttg tgg caa aat cct gat gat aag aga act gca cgg ttg tcc gag	2917
Asp Leu Trp Gln Asn Pro Asp Asp Lys Arg Thr Ala Arg Leu Ser Glu	
325	330
335	
ttt ggg gag atg atc agg gca gcc ttt gga ttt cct gtg gat aga cag	2965
Phe Gly Glu Met Ile Arg Ala Ala Phe Gly Phe Pro Val Asp Arg Gln	
340	345
350	355
aac atc cca agg aca gtc aca ggc cct aat gtc ctc ttt gtt aga cgt	3013
Asn Ile Pro Arg Thr Val Thr Gly Pro Asn Val Leu Phe Val Arg Arg	
360	365
370	
gag gat tat tta gct cac cca cgt cat ggt gga aag gta cag tct agg	3061
Glu Asp Tyr Leu Ala His Pro Arg His Gly Gly Lys Val Gln Ser Arg	
375	380
385	

ctt agc aat gaa gag caa gta ttt gat tcc ata aag agc tgg gcc ttg 3109
 Leu Ser Asn Glu Glu Gln Val Phe Asp Ser Ile Lys Ser Trp Ala Leu
 390 395 400
 aac cac tcg gag tgc aaa tta aat gta att aac gga ttg ttt gcc cac 3157
 Asn His Ser Glu Cys Lys Leu Asn Val Ile Asn Gly Leu Phe Ala His
 405 410 415
 atg tcc atg aaa gag caa gtt cga gca atc caa gat gca tct gtc att 3205
 Met Ser Met Lys Glu Gln Val Arg Ala Ile Gln Asp Ala Ser Val Ile
 420 425 430 435
 gtt ggt gct cat gga gca ggt cta act cac ata gtt tct gca gca cca 3253
 Val Gly Ala His Gly Ala Gly Leu Thr His Ile Val Ser Ala Ala Pro
 440 445 450
 aaa gct gta ata cta gaa att ata agc agc gaa tat agg cgc ccc cat 3301
 Lys Ala Val Ile Leu Glu Ile Ile Ser Ser Glu Tyr Arg Arg Pro His
 455 460 465
 ttt gct ctg att gca caa tgg aaa gga ttg gag tac cat ccc ata tat 3349
 Phe Ala Leu Ile Ala Gln Trp Lys Gly Leu Glu Tyr His Pro Ile Tyr
 470 475 480
 ttg gag ggg tct tat gcg gat cct cca gtt gtg atc 3385
 Leu Glu Gly Ser Tyr Ala Asp Pro Pro Val Val Ile
 485 490 495

<210> 10

<211> 495

<212> PRT

<213> Artificial

<220>

<223> Estrutura sintética

<400> 10

Val Ser Leu Phe Ala Leu Asn Ser Ile Thr Leu Tyr Leu Tyr Phe Ser

1

5

10

15

Ser His Pro Asp His Phe Arg His Lys Ser Arg Gln Asn His Phe Ser

20	25	30
Leu Ser Glu Asn Arg His His Asn Phe His Ser Ser Ile Thr Ser Gln		
35	40	45
Tyr Ser Lys Pro Trp Pro Ile Leu Pro Ser Tyr Leu Pro Trp Ser Gln		
50	55	60
Asn Pro Asn Val Val Trp Arg Ser Cys Glu Gly Tyr Phe Gly Asn Gly		
65	70	75
Phe Thr Leu Lys Val Asp Leu Leu Lys Thr Ser Pro Glu Phe His Arg		
85	90	95
Lys Phe Gly Glu Asn Thr Val Ser Gly Asp Gly Gly Trp Phe Arg Cys		
100	105	110
Phe Phe Ser Glu Thr Leu Gln Ser Ser Ile Cys Glu Gly Gly Ala Ile		
115	120	125
Arg Met Asn Pro Asp Glu Ile Leu Met Ser Arg Gly Gly Glu Lys Leu		
130	135	140
Glu Ser Val Ile Gly Arg Ser Glu Asp Asp Glu Leu Pro Val Phe Lys		
145	150	155
Asn Gly Ala Phe Gln Ile Lys Val Thr Asp Lys Leu Lys Ile Gly Lys		
165	170	175
Lys Leu Val Asp Glu Lys Ile Leu Asn Lys Tyr Leu Pro Glu Gly Ala		
180	185	190
Ile Ser Arg His Thr Met Arg Glu Leu Ile Asp Ser Ile Gln Leu Val		
195	200	205
Gly Ala Asp Glu Phe His Cys Ser Glu Trp Ile Glu Glu Pro Ser Leu		
210	215	220
Leu Ile Thr Arg Phe Glu Tyr Ala Asn Leu Phe His Thr Val Thr Asp		
225	230	235
Trp Tyr Ser Ala Tyr Val Ala Ser Arg Val Thr Gly Leu Pro Ser Arg		
245	250	255
Pro His Leu Val Phe Val Asp Gly His Cys Glu Thr Gln Leu Glu Glu		
260	265	270
Thr Trp Lys Ala Leu Phe Ser Ser Leu Thr Tyr Ala Lys Asn Phe Ser		

275	280	285	
Gly Pro Val Cys Phe Arg His Ala Val Leu Ser Pro Leu Gly Tyr Glu			
290	295	300	
Thr Ala Leu Phe Lys Gly Leu Thr Glu Thr Ile Asp Cys Asn Gly Ala			
305	310	315	320
Ser Ala His Asp Leu Trp Gln Asn Pro Asp Asp Lys Arg Thr Ala Arg			
	325	330	335
Leu Ser Glu Phe Gly Glu Met Ile Arg Ala Ala Phe Gly Phe Pro Val			
	340	345	350
Asp Arg Gln Asn Ile Pro Arg Thr Val Thr Gly Pro Asn Val Leu Phe			
	355	360	365
Val Arg Arg Glu Asp Tyr Leu Ala His Pro Arg His Gly Gly Lys Val			
	370	375	380
Gln Ser Arg Leu Ser Asn Glu Glu Gln Val Phe Asp Ser Ile Lys Ser			
385	390	395	400
Trp Ala Leu Asn His Ser Glu Cys Lys Leu Asn Val Ile Asn Gly Leu			
	405	410	415
Phe Ala His Met Ser Met Lys Glu Gln Val Arg Ala Ile Gln Asp Ala			
	420	425	430
Ser Val Ile Val Gly Ala His Gly Ala Gly Leu Thr His Ile Val Ser			
	435	440	445
Ala Ala Pro Lys Ala Val Ile Leu Glu Ile Ile Ser Ser Glu Tyr Arg			
	450	455	460
Arg Pro His Phe Ala Leu Ile Ala Gln Trp Lys Gly Leu Glu Tyr His			
465	470	475	480
Pro Ile Tyr Leu Glu Gly Ser Tyr Ala Asp Pro Pro Val Val Ile			
	485	490	495

<210> 11

<211> 3298

<212> DNA

<213> Artificial

<220>

<223> XylTg14Nb

<220>

<221> CDS

<222> (7)..(657)

<220>

<221> Intron

<222> (658)..(1917)

<220>

<221> CDS

<222> (1918)..(2067)

<220>

<221> Intron

<222> (2068)..(2612)

<220>

<221> CDS

<222> (2613)..(3290)

<400> 11

cacctt gtt tct ctc ttc gct ctc aac tca atc act ctc tat ctc tac 48

Val Ser Leu Phe Ala Leu Asn Ser Ile Thr Leu Tyr Leu Tyr

1

5

10

ttc tct tcc cac cct gat cac tct cgt cgc aaa tcc ccc cag aac cac 96

Phe Ser Ser His Pro Asp His Ser Arg Arg Lys Ser Pro Gln Asn His

15

20

25

30

ttt tcc tcg tcg gaa aac cac cat cat aat ttc cac tct tca atc act 144

Phe Ser Ser Ser Glu Asn His His His Asn Phe His Ser Ser Ile Thr

35

40

45

tcc caa tat tcc agg cct tgg cct att ttg ccc tcc tac ctc cct tgg 192

Ser Gln Tyr Ser Arg Pro Trp Pro Ile Leu Pro Ser Tyr Leu Pro Trp

50

55

60

tct caa aac cct aat gtt gct tgg aga tca tgc gag ggt tac ttc ggt 240

Ser Gln Asn Pro Asn Val Ala Trp Arg Ser Cys Glu Gly Tyr Phe Gly

65

70

75

aat ggt ttt act ctc aaa gtt gat ctt ctc aaa act tcg ccg gag ctt	288
Asn Gly Phe Thr Leu Lys Val Asp Leu Leu Lys Thr Ser Pro Glu Leu	
80 85 90	
cac cgg aaa ttc ggc gaa aac acc gtc ttc gga gac ggc gga tgg ttt	336
His Arg Lys Phe Gly Glu Asn Thr Val Phe Gly Asp Gly Gly Trp Phe	
95 100 105 110	
agg tgt ttc ttc agt gag act ttg cag agt tcg atc tgc gag gga ggc	384
Arg Cys Phe Phe Ser Glu Thr Leu Gln Ser Ser Ile Cys Glu Gly Gly	
115 120 125	
gca ata cga atg aat cca gac gag att ttg atg tct cgt gga ggt gag	432
Ala Ile Arg Met Asn Pro Asp Glu Ile Leu Met Ser Arg Gly Gly Glu	
130 135 140	
aaa ttg gag tcg gtt att ggt agg agt gaa gat gat gag gtg ccc gcg	480
Lys Leu Glu Ser Val Ile Gly Arg Ser Glu Asp Asp Glu Val Pro Ala	
145 150 155	
ttc aaa act gga gct ttt cag att aaa gtt act gat aaa ctg aaa ttt	528
Phe Lys Thr Gly Ala Phe Gln Ile Lys Val Thr Asp Lys Leu Lys Phe	
160 165 170	
ggg aaa aaa tta gtg gat gaa aac ttc ttg aat aaa tac tta ccg gaa	576
Gly Lys Lys Leu Val Asp Glu Asn Phe Leu Asn Lys Tyr Leu Pro Glu	
175 180 185 190	
ggt gca att tca agg cac act atg cgt gag tta atc gac tct att cag	624
Gly Ala Ile Ser Arg His Thr Met Arg Glu Leu Ile Asp Ser Ile Gln	
195 200 205	
ttg gtt ggc gcc aat gat ttt cac tgt tct gag gttagatttt tgaaattttg	677
Leu Val Gly Ala Asn Asp Phe His Cys Ser Glu	
210 215	
tttgctcttt aaattaaagg ttggaacttt gtgaatgttg gcagatatag aatacaataa	737
tggaatttgc ttgatctgtt taatgaagat tgtctggaac ctcaatgcta taaatatttg	797
tttgtttgct tcattaatta aagagaatat cccaactaga tgccagataa caccagttag	857
ttgacttttg gatcggattg catttcattt aatcagatat ggtactcatt cttaaattgt	917
tcactaaagt atttgtcaag atttcagagt ttatatgtag gtgtatttgg aattctggat	977

ttggatctag tattgaatgg attactgaac ttgtactccc cagtcactctg gggaggagca 1037
 acagattaac ttcaagggtt gaaaagtaat actgagtcag aagttaacca cttcaacttg 1097
 gaaaattgta aatgtgtgtg gtctaagatg attactctaa cttttgaggt ctaacatgga 1157
 gaaagttagt tgatttatgc tctttacttt tccctttatt gattttggct tttaaattct 1217
 atcaattcca ttgtttgatt gctactcaaa ttgaacctta gacggagtag caatagcaaa 1277
 aagtgaagaa aggccatttt ttttctcctt tcatctctta atttccgttt tacacacaga 1337
 atatggtaga atctgtttga agcttttagt gaatagttat acaactgggt attgcatttt 1397
 gaggactatc gacttgattt gacactggac agtgtctgat acatggcttg taagttatga 1457
 gaacttctat ctaggaagaa atcccaacca gagataggga gctgtcactt ggctatgagt 1517
 tactggctca aagttcgagt ttgaccagtt aatttttagat cctcaccagg ataacattta 1577
 gagtctaadc aaattctgaa gcagtattgt gcactaataa gaggaacaca tgaaggatgt 1637
 agcactacta ggttatgtta ccttattttac taataatgac tgacaaccag ctttaattgat 1697
 gacaaatgggt cttatatttg cctttttacat tgctcatgac ttgggatatt tctgaatcag 1757
 catttttcag ttctttatgt acttatcaaa aaattatccc tgctagatgt tagtgttcaa 1817
 gcaaccatgc tagcatttaa cgaagctcct tctttgattc atgcgatctt tccgtaatct 1877
 atgccttacg ttactgtcat ttttctaatt ttcatttcag tgg att gag gag ccg 1932

Trp Ile Glu Glu Pro

220

tca ctt ttg att aca cga ttt gag tat gca aac ctt ttc cac aca att 1980
 Ser Leu Leu Ile Thr Arg Phe Glu Tyr Ala Asn Leu Phe His Thr Ile

225

230

235

acc gat tgg tat agt gca tac gtg gca tcg agg gtt act ggc ttg ccc 2028
 Thr Asp Trp Tyr Ser Ala Tyr Val Ala Ser Arg Val Thr Gly Leu Pro

240

245

250

agt cgg cca cat ttg gtt ttt gta gat ggc cat tgt gag gtatgtctga 2077
 Ser Arg Pro His Leu Val Phe Val Asp Gly His Cys Glu

255

260

265

aagtattgat aacgatggca tgcattgtac tgtcttatgg atgaaagaaa tgaaaccagc 2137
 aattattttc tagcaggcaa tgctcttgag atgcttgtgt caaattgggc agacttaac 2197
 ctgagtttcc atttgtttca gctttctgtg tgactgacta caataattgt cccgatacct 2257
 aattgttgca gttggctcat tcttatttct atttacgtgt cactgtttct ctgaatggcc 2317
 ctttgtgggtg aaaagagctt ttgatatgta aaaaaactag caaagatttc atttctggaa 2377

caatttcttt ttaccttaca tcacgtgtca taaaattgct tctaactgta tactttaatt 2437
 cttggagaga tgctttcatg tgaagaaagt tctttcactc cactactgga agcttgctgc 2497
 atgaatttta cttggccata ttggggccgt gttttgattt atcttcaa atcattttctt 2557
 catgtagttc tttcgagtaa ttttttttcc tcttttctgt ttgaaaaaat ttcag aca 2615

Thr

caa ttg gag gaa aca tgg aaa gca ctt ttt tca agc ctc act tat gct 2663
 Gln Leu Glu Glu Thr Trp Lys Ala Leu Phe Ser Ser Leu Thr Tyr Ala

270

275

280

aag aac ttt agt ggc cca gtt tgt ttc cgt cat gcc gtc ctc tcg cct 2711
 Lys Asn Phe Ser Gly Pro Val Cys Phe Arg His Ala Val Leu Ser Pro
 285 290 295 300

ttg gga tat gaa act gcc ctg ttt aag gga ctg tca gaa act ata gat 2759
 Leu Gly Tyr Glu Thr Ala Leu Phe Lys Gly Leu Ser Glu Thr Ile Asp

305

310

315

tgt aat gga gct tct gct cat gat ttg tgg caa aat cct gat gat aag 2807
 Cys Asn Gly Ala Ser Ala His Asp Leu Trp Gln Asn Pro Asp Asp Lys

320

325

330

aaa act gca cgg tta tcc gag ttt ggg gag atg atc agg gca gcc ttt 2855
 Lys Thr Ala Arg Leu Ser Glu Phe Gly Glu Met Ile Arg Ala Ala Phe

335

340

345

gga ttt cct gtt gat aga cag aac atc cca agg aca gtc aca ggc cct 2903
 Gly Phe Pro Val Asp Arg Gln Asn Ile Pro Arg Thr Val Thr Gly Pro

350

355

360

aat gtc ctc ttt gtt aga cgt gag gat tat tta gct cac cca cgt cat 2951
 Asn Val Leu Phe Val Arg Arg Glu Asp Tyr Leu Ala His Pro Arg His

365

370

375

380

ggg gga aag gta cag tct agg ctt agc aat gaa gag caa gta ttt gat 2999
 Gly Gly Lys Val Gln Ser Arg Leu Ser Asn Glu Glu Gln Val Phe Asp

385

390

395

tcc ata aag agc tgg gcc tta aac cac tcg gag tgc aaa tta aat gta 3047
 Ser Ile Lys Ser Trp Ala Leu Asn His Ser Glu Cys Lys Leu Asn Val

400	405	410	
att agt gga ttg ttt gcc cac atg tcc atg aaa gag caa gtt cga gca			3095
Ile Ser Gly Leu Phe Ala His Met Ser Met Lys Glu Gln Val Arg Ala			
415	420	425	
atc caa gat gct tct gtc att gtt ggt gct cat gga gca ggt cta acc			3143
Ile Gln Asp Ala Ser Val Ile Val Gly Ala His Gly Ala Gly Leu Thr			
430	435	440	
cac ata gtt tct gca gca cca aaa gct gta ata cta gaa att ata agc			3191
His Ile Val Ser Ala Ala Pro Lys Ala Val Ile Leu Glu Ile Ile Ser			
445	450	455	460
agc gaa tat agg cgc ccc cat ttt gct ctg att gct caa tgg aaa gga			3239
Ser Glu Tyr Arg Arg Pro His Phe Ala Leu Ile Ala Gln Trp Lys Gly			
465	470	475	
ttg gag tac cat ccc ata tat ttg gag ggg tct tat gcg gat cct cca			3287
Leu Glu Tyr His Pro Ile Tyr Leu Glu Gly Ser Tyr Ala Asp Pro Pro			
480	485	490	
gtt gtgatcga			3298
Val			

<210> 12

<211> 493

<212> PRT

<213> Artificial

<220>

<223> Estrutura sintética

<400> 12

Val Ser Leu Phe Ala Leu Asn Ser Ile Thr Leu Tyr Leu Tyr Phe Ser

1	5	10	15
---	---	----	----

Ser His Pro Asp His Ser Arg Arg Lys Ser Pro Gln Asn His Phe Ser

20	25	30
----	----	----

Ser Ser Glu Asn His His His Asn Phe His Ser Ser Ile Thr Ser Gln

35	40	45
----	----	----

Tyr Ser Arg Pro Trp Pro Ile Leu Pro Ser Tyr Leu Pro Trp Ser Gln

50

55

60

Asn Pro Asn Val Ala Trp Arg Ser Cys Glu Gly Tyr Phe Gly Asn Gly

65

70

75

80

Phe Thr Leu Lys Val Asp Leu Leu Lys Thr Ser Pro Glu Leu His Arg

85

90

95

Lys Phe Gly Glu Asn Thr Val Phe Gly Asp Gly Gly Trp Phe Arg Cys

100

105

110

Phe Phe Ser Glu Thr Leu Gln Ser Ser Ile Cys Glu Gly Gly Ala Ile

115

120

125

Arg Met Asn Pro Asp Glu Ile Leu Met Ser Arg Gly Gly Glu Lys Leu

130

135

140

Glu Ser Val Ile Gly Arg Ser Glu Asp Asp Glu Val Pro Ala Phe Lys

145

150

155

160

Thr Gly Ala Phe Gln Ile Lys Val Thr Asp Lys Leu Lys Phe Gly Lys

165

170

175

Lys Leu Val Asp Glu Asn Phe Leu Asn Lys Tyr Leu Pro Glu Gly Ala

180

185

190

Ile Ser Arg His Thr Met Arg Glu Leu Ile Asp Ser Ile Gln Leu Val

195

200

205

Gly Ala Asn Asp Phe His Cys Ser Glu Trp Ile Glu Glu Pro Ser Leu

210

215

220

Leu Ile Thr Arg Phe Glu Tyr Ala Asn Leu Phe His Thr Ile Thr Asp

225

230

235

240

Trp Tyr Ser Ala Tyr Val Ala Ser Arg Val Thr Gly Leu Pro Ser Arg

245

250

255

Pro His Leu Val Phe Val Asp Gly His Cys Glu Thr Gln Leu Glu Glu

260

265

270

Thr Trp Lys Ala Leu Phe Ser Ser Leu Thr Tyr Ala Lys Asn Phe Ser

275

280

285

Gly Pro Val Cys Phe Arg His Ala Val Leu Ser Pro Leu Gly Tyr Glu

290

295

300

Thr Ala Leu Phe Lys Gly Leu Ser Glu Thr Ile Asp Cys Asn Gly Ala
 305 310 315 320
 Ser Ala His Asp Leu Trp Gln Asn Pro Asp Asp Lys Lys Thr Ala Arg
 325 330 335
 Leu Ser Glu Phe Gly Glu Met Ile Arg Ala Ala Phe Gly Phe Pro Val
 340 345 350
 Asp Arg Gln Asn Ile Pro Arg Thr Val Thr Gly Pro Asn Val Leu Phe
 355 360 365
 Val Arg Arg Glu Asp Tyr Leu Ala His Pro Arg His Gly Gly Lys Val
 370 375 380
 Gln Ser Arg Leu Ser Asn Glu Glu Gln Val Phe Asp Ser Ile Lys Ser
 385 390 395 400
 Trp Ala Leu Asn His Ser Glu Cys Lys Leu Asn Val Ile Ser Gly Leu
 405 410 415
 Phe Ala His Met Ser Met Lys Glu Gln Val Arg Ala Ile Gln Asp Ala
 420 425 430
 Ser Val Ile Val Gly Ala His Gly Ala Gly Leu Thr His Ile Val Ser
 435 440 445
 Ala Ala Pro Lys Ala Val Ile Leu Glu Ile Ile Ser Ser Glu Tyr Arg
 450 455 460
 Arg Pro His Phe Ala Leu Ile Ala Gln Trp Lys Gly Leu Glu Tyr His
 465 470 475 480
 Pro Ile Tyr Leu Glu Gly Ser Tyr Ala Asp Pro Pro Val
 485 490

<210> 13

<211> 3574

<212> DNA

<213> Artificial

<220>

<223> XylTg19Nb

<220>

<221> CDS

<222> (7)..(648)

<220>

<221> Intron

<222> (649)..(2194)

<220>

<221> CDS

<222> (2195)..(2344)

<220>

<221> Intron

<222> (2345)..(2888)

<220>

<221> CDS

<222> (2889)..(3566)

<400> 13

cacctt gtt tct ctc ttc gct ctc aac tca atc act ctc tat ctc tac 48

Val Ser Leu Phe Ala Leu Asn Ser Ile Thr Leu Tyr Leu Tyr

1 5 10

ttc tct tcc cac cct gat cac aaa tcc ccc caa aac cac ttt tcc ttg 96

Phe Ser Ser His Pro Asp His Lys Ser Pro Gln Asn His Phe Ser Leu

15 20 25 30

tcg gaa aac cac cat cat aat ttc cac tct tca atc act tct caa tat 144

Ser Glu Asn His His His Asn Phe His Ser Ser Ile Thr Ser Gln Tyr

35 40 45

tcc aag cct tgg cct att ttg ccc tcc tac ctc cct tgg tct caa aac 192

Ser Lys Pro Trp Pro Ile Leu Pro Ser Tyr Leu Pro Trp Ser Gln Asn

50 55 60

cct aat gtt gct tgg aga tcg tgc gag ggt tac ttc ggt aat ggg ttt 240

Pro Asn Val Ala Trp Arg Ser Cys Glu Gly Tyr Phe Gly Asn Gly Phe

65 70 75

act ctc aaa gtt gac ctt ctc aaa act tcg ccg gag ttt cac cgg aaa 288

Thr Leu Lys Val Asp Leu Leu Lys Thr Ser Pro Glu Phe His Arg Lys

80	85	90	
ttc ggc gat aac acc gtc tcc ggt gac ggc gga tgg ttt agg tgt ttt			336
Phe Gly Asp Asn Thr Val Ser Gly Asp Gly Gly Trp Phe Arg Cys Phe			
95	100	105	110
ttc agt gag act ttg cag agt tcg atc tgc gag gga ggc gca ata cga			384
Phe Ser Glu Thr Leu Gln Ser Ser Ile Cys Glu Gly Gly Ala Ile Arg			
	115	120	125
atg aat ccg gac gat att ttg atg tct cgt gga ggt gag aaa ttg gag			432
Met Asn Pro Asp Asp Ile Leu Met Ser Arg Gly Gly Glu Lys Leu Glu			
	130	135	140
tcg gtt att ggt agg aat gaa gat gat gag ctg ccc atg ttc aaa aat			480
Ser Val Ile Gly Arg Asn Glu Asp Asp Glu Leu Pro Met Phe Lys Asn			
	145	150	155
gga gct ttc caa att gaa gtt act gat aaa ctg aaa att ggg aaa aaa			528
Gly Ala Phe Gln Ile Glu Val Thr Asp Lys Leu Lys Ile Gly Lys Lys			
	160	165	170
cta gtg gat aaa aaa ttc ttg aat aaa tac tta ccg gga ggt gcg att			576
Leu Val Asp Lys Lys Phe Leu Asn Lys Tyr Leu Pro Gly Gly Ala Ile			
175	180	185	190
tca agg cac act atg cgt gag tta att gac tct att cag ttg gtt ggc			624
Ser Arg His Thr Met Arg Glu Leu Ile Asp Ser Ile Gln Leu Val Gly			
	195	200	205
gcc gat gaa ttt cac tgt tct gag gttagatttt gatatttatt tgatctttaa			678
Ala Asp Glu Phe His Cys Ser Glu			
	210		
attagagggtt tgaactttgt taatgttggc agatatggaa tacaataatg gattttgttt			738
gatctgttta atgaagattg tctaaaacct caatgctata aatatttgtt tgtttgcttc			798
attaattaaa gagaatatcc cgactagatg ccagataaca ccagttagtt gacttttggg			858
ttgggttgca tttcatttaa tcagatatgg tactcattct taaatgtttc actaaagaat			918
ttgtcaagat ttcagagttt atatataggt gtatttggaa ttctggattt ggatctagta			978
ttgaatggat tactgaattt gtactcccca gtcacaggg gaggagcaat agatcgaatt			1038
caagggttga aaagtaatac tgagtcagaa attaaccact ttaacttga aacggtaaat			1098

gtatgtgttc taagatgatt attcctataa cttttgatgt ctaatatgga gaaagtgagt 1158
 tgatttatgc tttttccttt tccctttatt gatgttggtt tttaaattct atcaattcct 1218
 ttgtttgggt gctactcaaa ttgaacctta gacggagtag caatagcaaa aagtgaagaa 1278
 aggacatttt tttctccttt catctcttta tttccgtttg acatacagaa tacggtagca 1338
 tctgcctgaa gtggttaatt tcattcctta aaatttgcac aactaatatt tccgtttttg 1398
 tttttgttta tcttttccat tggcatgcca tggtattttt ggtttaggtt tacataatta 1458
 tttatgtgat ttctgatgga gttactaatg attttttggt tttgtttttg tttttttctt 1518
 ttccttttcc tgagtcgagg gtcgattgga aatagcctct ctgccctttt ggataggggt 1578
 aaggcctggg tacgtgtacc atccccagac cccactctgt gggactatac cgggtagttg 1638
 ttgttggtgt aattcgagtaaatgcctttt gaaccttttag ttgaatagtt gtacaactgg 1698
 ttgttgcatt ttgaggacta tcgacttgat ttgacacttt acatgaaaac ttttatctag 1758
 gaagaaatcc ctaccagaga tagggagctg tcgcttggtt atgagctact ggcttaaagt 1818
 ttgagtttga cctattaatt ttagatcttc accaggataa catctagagt ttaattaaat 1878
 tctcaagcag ttttttgcac taataagggg aacacatgaa ggatgtagca ctactacgtt 1938
 atgttcttta tttactattg attgacaacc agcttaaatg atgacaaatg gtcttatatt 1998
 tgctttttta cattgctcat gacttgggat atttttgaat caacattttt cggttcttta 2058
 tgtacttata aaaaaattat cctgctaga tgtagtggt caagcaacat gctagctttt 2118
 aaggaagctc cttctttgat tcatgccatc tttccgaagc cttacgtttc tgctattttt 2178
 ctaattttca tttcag tgg gtt gag gag ccg tca ctt ttg att aca cga ttt 2230

Trp Val Glu Glu Pro Ser Leu Leu Ile Thr Arg Phe

215

220

225

gag tat gca aac ctt ttc cac aca gtt acc gat tgg tat agt gca tac 2278

Glu Tyr Ala Asn Leu Phe His Thr Val Thr Asp Trp Tyr Ser Ala Tyr

230

235

240

gcg gca tcc agg gtt act ggt ttg ccc agt cgg cca aat ttg gtt ttt 2326

Ala Ala Ser Arg Val Thr Gly Leu Pro Ser Arg Pro Asn Leu Val Phe

245

250

255

gta gat ggc cat tgt gag gtatgtttga cagtattgat aacgatggca 2374

Val Asp Gly His Cys Glu

260

tgcatgttac tgtgttatgg atgaaagaaa tgaaaccatc aattattttc tagtaggcaa 2434

tgctcttaag atgcttgtgt caaattgggt agagttaatc ctaagtttcc atttgtttga 2494

gctttctgtt tgactgacta caatacttgt cccaatacct agttgttgcg gttggctcat 2554
 tcttacttct atttacgtgt cactgtttct ctgaatggc cctttgtggt gaaaagagct 2614
 tttgctatgt agaaaaacta gcaaagattt catttctgga gcaacttatt tttaccttac 2674
 atcacgtctc ataaaattgc ttctaactgt atactttaat tcttgagag atgctttcat 2734
 gtgaataaag ttctttcact ccactactgg aagcttgctg catgaaattt acttgccat 2794
 actggggccg tgttttgatt tgtcttcaaa ttcattttct tcatgtagtt ctttcgagta 2854
 atatttttct ctcttctgtt tgaaaaaat tcag aca caa ttg gag gaa aca tgg 2909
 Thr Gln Leu Glu Glu Thr Trp
 265 270
 aaa gca ctt ttt tca agc ctc act tat gct aag aac ttt agt ggc cca 2957
 Lys Ala Leu Phe Ser Ser Leu Thr Tyr Ala Lys Asn Phe Ser Gly Pro
 275 280 285
 gtt tgt ttc cgt cat gct gtc ctc tcg cct tta gga tat gaa act gcc 3005
 Val Cys Phe Arg His Ala Val Leu Ser Pro Leu Gly Tyr Glu Thr Ala
 290 295 300
 ctg ttt aag gga ctg tca gaa act ata gat tgt aat gga gct tct gct 3053
 Leu Phe Lys Gly Leu Ser Glu Thr Ile Asp Cys Asn Gly Ala Ser Ala
 305 310 315
 cat gat ttg tgg caa aag cct gat gat aaa aaa act gca cgg ttg tcc 3101
 His Asp Leu Trp Gln Lys Pro Asp Asp Lys Lys Thr Ala Arg Leu Ser
 320 325 330 335
 gag ttt ggg gag atg atc agg gca gcc ttt gga ttt cct gtg gat aga 3149
 Glu Phe Gly Glu Met Ile Arg Ala Ala Phe Gly Phe Pro Val Asp Arg
 340 345 350
 cag aac atc cca agg aca gtc aca ggc cct aat gtc ctc ttt gtt aga 3197
 Gln Asn Ile Pro Arg Thr Val Thr Gly Pro Asn Val Leu Phe Val Arg
 355 360 365
 cgt gag gat tat tta gct cac cca cgt cat ggt gga aag gta cag tct 3245
 Arg Glu Asp Tyr Leu Ala His Pro Arg His Gly Gly Lys Val Gln Ser
 370 375 380
 agg ctt agc aat gaa gag cta gta ttt gat tcc ata aag agc tgg gcc 3293
 Arg Leu Ser Asn Glu Glu Leu Val Phe Asp Ser Ile Lys Ser Trp Ala

385	390	395	
ttg aac cac tcg gag tgt aaa tta aat gta att aac gga ttg ttt gcc			3341
Leu Asn His Ser Glu Cys Lys Leu Asn Val Ile Asn Gly Leu Phe Ala			
400	405	410	415
cac atg tcc atg aaa gag caa gtt cga gca atc caa gat gct tct gtc			3389
His Met Ser Met Lys Glu Gln Val Arg Ala Ile Gln Asp Ala Ser Val			
420	425	430	
att gtt ggt gct cat gga gca ggt cta act cac ata gtt tct gca gca			3437
Ile Val Gly Ala His Gly Ala Gly Leu Thr His Ile Val Ser Ala Ala			
435	440	445	
cca aaa gct gta ata cta gaa att ata agc agc gaa tat agg cgc ccc			3485
Pro Lys Ala Val Ile Leu Glu Ile Ile Ser Ser Glu Tyr Arg Arg Pro			
450	455	460	
cat ttt gct ctg att gca caa tgg aaa gga ttg gag tac cat ccc ata			3533
His Phe Ala Leu Ile Ala Gln Trp Lys Gly Leu Glu Tyr His Pro Ile			
465	470	475	
tat ttg gag ggg tct tat gcg gat cct cca gtt gtgatcga			3574
Tyr Leu Glu Gly Ser Tyr Ala Asp Pro Pro Val			
480	485	490	

<210> 14

<211> 490

<212> PRT

<213> Artificial

<220>

<223> Estrutura sintética

<400> 14

Val Ser Leu Phe Ala Leu Asn Ser Ile Thr Leu Tyr Leu Tyr Phe Ser

1

5

10

15

Ser His Pro Asp His Lys Ser Pro Gln Asn His Phe Ser Leu Ser Glu

20

25

30

Asn His His His Asn Phe His Ser Ser Ile Thr Ser Gln Tyr Ser Lys

35

40

45

Pro Trp Pro Ile Leu Pro Ser Tyr Leu Pro Trp Ser Gln Asn Pro Asn
 50 55 60
 Val Ala Trp Arg Ser Cys Glu Gly Tyr Phe Gly Asn Gly Phe Thr Leu
 65 70 75 80
 Lys Val Asp Leu Leu Lys Thr Ser Pro Glu Phe His Arg Lys Phe Gly
 85 90 95
 Asp Asn Thr Val Ser Gly Asp Gly Gly Trp Phe Arg Cys Phe Phe Ser
 100 105 110
 Glu Thr Leu Gln Ser Ser Ile Cys Glu Gly Gly Ala Ile Arg Met Asn
 115 120 125
 Pro Asp Asp Ile Leu Met Ser Arg Gly Gly Glu Lys Leu Glu Ser Val
 130 135 140
 Ile Gly Arg Asn Glu Asp Asp Glu Leu Pro Met Phe Lys Asn Gly Ala
 145 150 155 160
 Phe Gln Ile Glu Val Thr Asp Lys Leu Lys Ile Gly Lys Lys Leu Val
 165 170 175
 Asp Lys Lys Phe Leu Asn Lys Tyr Leu Pro Gly Gly Ala Ile Ser Arg
 180 185 190
 His Thr Met Arg Glu Leu Ile Asp Ser Ile Gln Leu Val Gly Ala Asp
 195 200 205
 Glu Phe His Cys Ser Glu Trp Val Glu Glu Pro Ser Leu Leu Ile Thr
 210 215 220
 Arg Phe Glu Tyr Ala Asn Leu Phe His Thr Val Thr Asp Trp Tyr Ser
 225 230 235 240
 Ala Tyr Ala Ala Ser Arg Val Thr Gly Leu Pro Ser Arg Pro Asn Leu
 245 250 255
 Val Phe Val Asp Gly His Cys Glu Thr Gln Leu Glu Glu Thr Trp Lys
 260 265 270
 Ala Leu Phe Ser Ser Leu Thr Tyr Ala Lys Asn Phe Ser Gly Pro Val
 275 280 285
 Cys Phe Arg His Ala Val Leu Ser Pro Leu Gly Tyr Glu Thr Ala Leu
 290 295 300

Phe Lys Gly Leu Ser Glu Thr Ile Asp Cys Asn Gly Ala Ser Ala His
 305 310 315 320
 Asp Leu Trp Gln Lys Pro Asp Asp Lys Lys Thr Ala Arg Leu Ser Glu
 325 330 335
 Phe Gly Glu Met Ile Arg Ala Ala Phe Gly Phe Pro Val Asp Arg Gln
 340 345 350
 Asn Ile Pro Arg Thr Val Thr Gly Pro Asn Val Leu Phe Val Arg Arg
 355 360 365
 Glu Asp Tyr Leu Ala His Pro Arg His Gly Gly Lys Val Gln Ser Arg
 370 375 380
 Leu Ser Asn Glu Glu Leu Val Phe Asp Ser Ile Lys Ser Trp Ala Leu
 385 390 395 400
 Asn His Ser Glu Cys Lys Leu Asn Val Ile Asn Gly Leu Phe Ala His
 405 410 415
 Met Ser Met Lys Glu Gln Val Arg Ala Ile Gln Asp Ala Ser Val Ile
 420 425 430
 Val Gly Ala His Gly Ala Gly Leu Thr His Ile Val Ser Ala Ala Pro
 435 440 445
 Lys Ala Val Ile Leu Glu Ile Ile Ser Ser Glu Tyr Arg Arg Pro His
 450 455 460
 Phe Ala Leu Ile Ala Gln Trp Lys Gly Leu Glu Tyr His Pro Ile Tyr
 465 470 475 480
 Leu Glu Gly Ser Tyr Ala Asp Pro Pro Val
 485 490

<210> 15

<211> 27

<212> DNA

<213> Artificial

<220>

<223> oligonucleotídeo XylF8

<400> 15

cacctctcgc ctttgggata tgaaact

<210> 16

<211> 24

<212> DNA

<213> Artificial

<220>

<223> oligonucleotídeo XylR8

<400> 16

acagcttttg gtgctgcaga aact

24

<210> 17

<211> 473

<212> DNA

<213> Artificial

<220>

<223> XylTi4Nt

<400> 17

cacctctcgc ctttgggata tgaaactgcc ctgtttaagg gactgacaga aactatagat 60

tgtaatggag cttctgccca tgatttgtgg caaaatcctg atgataagag aactgcacgg 120

ttgtccgagt ttggggagat gatcagggca gcctttggat ttctgtgga tagacagaac 180

atcccaagga cagtcacagg ccctaattgc ctctttgtta gacgtgagga ttatttagct 240

caccacgctc atggtggaaa ggtacagtct aggcttagca atgaagagca agtatttgat 300

tccataaaga gctgggcctt gaaccactcg gaggcgaat taaatgtaat taacggattg 360

tttgcccaca tgtccatgaa agagcaagtt cgagcaatcc aagatgcac tgtcattggt 420

ggtgctcatg gagcaggtct aactcacata gtttctgcag caccaaaagc tgt 473

<210> 18

<211> 7954

<212> DNA

<213> Artificial

<220>

<223> T-DNA de pTKW20

<220>

<221> misc_feature

<222> (1)..(619)

<223> borda direita

<220>

<221> misc_feature

<222> (969)..(2314)

<223> seqüência incluindo a região promotora do transcrito 35S do vírus mosaico de Couve-Flor (Odell e outros, 1985)

<220>

<221> misc_feature

<222> (2365)..(2834)

<223> seqüência incluindo uma parte da seqüência de cDNA de XylT de *Nicotiana tabacum* cv. Petite Havana SR1 clonada na orientação sentido

<220>

<221> misc_feature

<222> (2893)..(3088)

<223> seqüência incluindo o íntron do gene de catase-1 de semente de mamona

<220>

<221> misc_feature

<222> (3130)..(3871)

<223> seqüência contendo o segundo íntron do gene de diquinase de ortofosfato de *Flaveria trinervia* conforme descrito por Rosche e Westhoff (1995) na orientação reversa

<220>

<221> misc_feature

<222> (3957)..(4426)

<223> seqüência incluindo uma parte da seqüência de cDNA de XylT de *Nicotiana tabacum* cv. Petite Havana SR1 clonada na orientação anti-sentido

<220>

<221> misc_feature

<222> (4479)..(5244)

<223> seqüência incluindo a região não traduzida 3' do gene de sintase de

octopina de *Agrobacterium tumefaciens* conforme descrito por De Greve e outros (1982)

<220>

<221> misc_feature

<222> (5512)..(5744)

<223> sequência incluindo a região promotora do gene de sintase de nopalina de T-DNA de *Agrobacterium tumefaciens*

<220>

<221> misc_feature

<222> (5745)..(6690)

<223> sequência incluindo o gene de resistência a antibiótico nptII

<220>

<221> misc_feature

<222> (6691)..(7396)

<223> sequência incluindo a região não traduzida 3' do gene de sintase de nopalina de T-DNA de *A. tumefaciens*

<220>

<221> misc_feature

<222> (7397)..(7954)

<223> borda esquerda

<400> 18

tcgacatctt gctgcgttcg gatattttcg tggagttccc gccacagacc cggattgaag	60
gcgagatcca gcaactcgcg ccagatcatc ctgtgacgga actttggcgc gtgatgactg	120
gccaggacgt cggccgaaag agcgacaagc agatcacgat ttctgacagc gtcggatttg	180
cgatcgagga tttttcgcg ctgcgctacg tccgcgaccg cgttgaggga tcaagccaca	240
gcagcccact cgaccttcta gccgaccag acgagccaag ggatcttttt ggaatgctgc	300
tccgtcgtca ggctttccga cgtttggttg gttgaacaga agtcattatc gtacggaatg	360
ccagcactcc cgaggggaac cctgtggttg gcatgcacat acaaaggac gaacggataa	420
accttttcac gcccttttaa atatccgtta ttctaataaa cgctcttttc tcttaggttt	480
acccgccaat atatcctgtc aaacactgat agtttaaact gaaggcgga aacgacaatc	540
tgatcatgag cggagaatta agggagtcac gttatgaccc ccgccgatga cgcgggacaa	600
gccgttttac gtttggaact gacagaaccg caacgattga aggagccact cagccccaat	660

acgcaaaccg cctctccccg cgcgttgccc gattcattaa tgcagctggc acgacaggtt	720
tcccgactgg aaagcgggca gtgagcgcaa cgcaattaat gtgagttagc tcaactatta	780
ggcaccaccg gctttacact ttatgcttcc ggctcgtatg ttgtgtggaa ttgtgagcgg	840
ataacaattt cacacaggaa acagctatga ccatgattac gccaaagctat ttaggtgaca	900
ctatagaata ctcaagctat gcatccaacg cgttggggagc tctcccatat cgacctgcag	960
gcggccgctc gacgaattaa ttccaatccc acaaaaatct gagcttaaca gcacagttgc	1020
tcctctcaga gcagaatcgg gtattcaaca ccctcatatc aactactacg ttgtgtataa	1080
cgggccacat gccggtatat acgatgactg gggttgtaca aaggcggcaa caaacggcgt	1140
tcccgaggtt gcacacaaga aatttgccac tattacagag gcaagagcag cagctgacgc	1200
gtacacaaca agtcagcaaa cagacaggtt gaacttcac cccaaaggag aagctcaact	1260
caagcccaag agctttgcta aggccctaac aagcccacca aagcaaaaag cccactggct	1320
cacgctagga accaaaaggc ccagcagtga tccagcccca aaagagatct cctttgcccc	1380
ggagattaca atggacgatt tcctctatct ttacgatcta ggaaggaagt tcgaaggtga	1440
aggtgacgac actatgttca ccaactgataa tgagaagggt agcctcttca atttcagaaa	1500
gaatgctgac ccacagatgg ttagagaggc ctacgcagca ggtctcatca agacgatcta	1560
cccagtaac aatctccagg agatcaaata ccttccaag aagggttaaag atgcagtcaa	1620
aagattcagg actaattgca tcaagaacac agagaaagac atatttctca agatcagaag	1680
tactattcca gtatggacga ttcaaggctt gttcataaa ccaaggcaag taatagagat	1740
tggagtctct aaaaaggtag ttctactga atctaaggcc atgcatggag tctaagattc	1800
aaatcgagga tctaacagaa ctgcgctga agactggcga acagttcata cagagtcttt	1860
tacgactcaa tgacaagaag aaaatcttcg tcaacatggt ggagcacgac actctggctc	1920
actccaaaaa tgtcaaagat acagtctcag aagaccaaag ggctattgag acttttcaac	1980
aaaggataat ttcgggaaac ctctcggat tccattgccc agctatctgt cacttcacg	2040
aaaggacagt agaaaaggaa ggtggctcct acaaatgcca tcattgcgat aaaggaaagg	2100
ctatcattca agatctctct gccgacagtg gtcccaaaga tggaccccca cccacgagga	2160
gcatcgtgga aaaagaagac gttccaacca cgtcttcaaa gcaagtggat tgatgtgaca	2220
tctccactga cgtaagggat gacgcacaat cccactatcc ttcgcaagac ccttctcta	2280
tataaggaag ttcatctcat ttggagagga cacgctcgag acaagtttgt acaaaaaagc	2340
aggctccgcg gccgccccct tcacctctcg cctttgggat atgaaactgc cctgtttaag	2400
ggactgacag aaactataga ttgtaatgga gttctgccc atgattttgt gcaaaatcct	2460
gatgataaga gaactgcacg gttgtccgag tttggggaga tgatcagggc agcctttgga	2520
tttcctgtgg atagacagaa catcccaagg acagtcacag gccctaagt cctctttgtt	2580

agacgtgagg attatttagc tcacccacgt catggtggaa aggtacagtc taggcttagc	2640
aatgaagagc aagtatttga ttccataaag agctgggcct tgaaccactc ggagtgc aaa	2700
ttaa atgtaa ttaacggatt gtttgccac atgtccatga aagagcaagt tcgagcaatc	2760
caagatgcat ctgtcattgt tgggtgctcat ggagcaggtc taactcacat agtttctgca	2820
gcacaaaaag ctgtaagggt gggcgcgccg acccagcttt cttgtacaaa gtggctctaga	2880
ggatccaagc ttgtgcagggt aaatttctag tttttctcct tcattttctt ggtaggacc	2940
cttttctctt tttatttttt tgagctttga tctttcttta aactgatcta tttttaatt	3000
gattggttat ggcgcaaata ttacatagct ttaactgata atctgattac tttatttcgt	3060
gtgtctatga tgatgatgat aactgcagcg caagcttatac gatttcgaac ccagcttccc	3120
aactgtaatc aatccaaatg taagatcaat gataacacaa tgacatgatc tatcatgtta	3180
ccttgtttat tcatgttcga ctaattcatt taattaatag tcaatccatt tagaagttaa	3240
taaaactaca agtattattt agaaattaat aagaatgttg attgaaaata atactatata	3300
aaatgataga tcttgcgctt tggtatatta gcattagatt atgttttggt acattagatt	3360
actgtttcta ttagtttgat attatttggt actttagctt gttattta attttgttta	3420
ttgataaatt acaagcagat tggaatttct aacaaaatat ttattaactt ttaactaaa	3480
atatttagta atggtataga tatttaatta tataataaac tattaatcat aaaaaatat	3540
tattttaatt tattttattct tatttttact atagtatttt atcattgata ttttaattcat	3600
caaaccagct agaattacta ttatgattaa aacaaatatt aatgctagta tatcatctta	3660
catgttcgat caaattcatt aaaaataata tacttactct caacttttat cttcttcgtc	3720
ttacacatca cttgtcatat ttttttacat tactatgttg tttatgtaa caatatattt	3780
ataaattatt ttttcacaat tataacaact atattattat aatcatacta attaacatca	3840
cttaactatt ttatactaaa agggaaaaag aaaataatta tttccttacc aagctggggt	3900
accgaattcc tcgagaccac tttgtacaag aaagctgggt cggcgcgccc acccttacag	3960
cttttggtgc tgcagaaact atgtgagtta gacctgctcc atgagacca acaatgacag	4020
atgcatcttg gattgctcga acttgctctt tcatggacat gtgggcaaac aatccgttaa	4080
ttacatttaa tttgcactcc gagtggttca aggccagct ctttatggaa tcaaatactt	4140
gctcttcatt gctaagccta gactgtacct ttccacatg acgtgggtga gctaaataat	4200
cctcacgtct aacaaagagg acattagggc ctgtgactgt ccttgggatg ttctgtctat	4260
ccacaggaaa tccaaaggct gccctgatca tctcccaaaa ctcggacaac cgtgcagttc	4320
tcttatcatc aggattttgc cacaatcat gggcagaagc tccattacaa tctatagttt	4380
ctgtcagtc cttaaacagg gcagtttcat atcccaaagg cgagaggtga agggggcggc	4440
cgcgagcct gcttttttgt acaaacttgt ctagagtcct gctttaatga gatatgcgag	4500

acgcctatga tcgcatgata tttgctttca attctgttgt gcacgttgta aaaaacctga	4560
gcatgtgtag ctcagatcct taccgccggt ttcggttcat tctaataaat atatcacccg	4620
ttactatcgt atttttatga ataataattct ccgttcaatt tactgattgt accctactac	4680
ttatatgtac aatattaaaa tgaaaacaat atattgtgct gaataaggtt atagcgacat	4740
ctatgataga gcgccacaat aacaaacaat tgcgttttat tattacaaat ccaattttta	4800
aaaaagcggc agaaccggc aaacctaaaa gactgattac ataaatctta ttcaaatttc	4860
aaaaggcccc aggggctagt atctacgaca caccgagcgg cgaactaata acgttactg	4920
aagggaactc cggttccccg ccggcgcgca tgggtgagat tccttgaagt tgagtattgg	4980
ccgtccgctc taccgaaagt tacgggcacc attcaaccgc gtccagcacg gcggccgggt	5040
aaccgacttg ctgccccgag aattatgcag ctttttttg gtgtatgtgg gccccaaatg	5100
aagtgcaggc caaaccttga cagtgcgac aaatcgttg gcgggtccag ggcaatttt	5160
gcgacaacat gtcgaggtc agcaggacct gcaggcatgc aagctagctt actagtgatg	5220
catattctat agtgtcacct aaatctgcgg ccgcactagt gatatccgc ggccatggcg	5280
gccgggagca tgcgacgtc ggccaattc gccctatagt gagtcgtatt acaattcact	5340
ggcgcgtgtt ttacaacgtc gtgactggga aaaccctggc gttacccaac ttaatgcct	5400
tgcagacat cccctttcg ccagctggcg taatagcgaa gagggccgca ccgatcgccc	5460
ttccaacag ttgcgcagcc tgaatggcga atggaaattg taaacgttaa tgggtttctg	5520
gagtttaatg agctaagcac atacgtcaga aaccattatt gcgcgttcaa aagtcgccta	5580
aggtcactat cagctagcaa atatttcttg tcaaaaatgc tccactgacg ttccataaat	5640
tcccctcggg atccaattag agtctcatat tcaactctca tccaaataat ctgcaatggc	5700
aattacctta tccgcaactt ctttacctat ttccgcccgc atccgggcag gttctccggc	5760
cgcttgggtg gagaggctat tcggctatga ctgggcacaa cagacaatcg gctgctctga	5820
tgccgccgtg ttccggctgt cagcgcaggg gcgcccgtt ctttttgtca agaccgacct	5880
gtccggtgcc ctgaatgaac tgcaggacga ggcagcggc ctatcgtggc tggccacgac	5940
gggcgttcct tgcgcagctg tgctcgacgt tgtcactgaa gcgggaaggg actggctgct	6000
attgggcgaa gtgccggggc aggatctcct gtcactcac cttgctcctg ccgagaaagt	6060
atccatcatg gctgatgcaa tgccggcgct gcatacgtt gatccggcta cctgcccatt	6120
cgaccaccaa gcgaaacatc gcacgcagcg agcacgtact cggatggaag ccggtcttgt	6180
cgatcaggat gatctggacg aagagcatca ggggctcgcg ccagccgaac tgttcgccag	6240
gctcaaggcg cgcagccccg acggcgagga tctcgtcgtg acccatggcg atgcctgctt	6300
gccgaatatc atggtggaaa atggccgctt ttctggattc atcgactgtg gccggctggg	6360
tgtggcggac cgctatcagg acatagcgtt ggctaccgct gatattgctg aagagcttgg	6420

cggcgaatgg gctgaccgct tctcgtgct ttacggatc gccgctccc attcgagcg 6480
 catcgcttc tatcgcttc ttgacgagtt cttctgagcg ggactctggg gttcgaaatg 6540
 accgaccaag cgacgcccaa cctgccatca cgagatttcg attccaccgc cgccttctat 6600
 gaaagggttg gtttcggaat cgttttccgg gacgccggct ggatgacct ccagcgcggg 6660
 gatctcatgc tggagttctt cgcccacccc gatccaacac ttacgtttgc aacgtccaag 6720
 agcaaataga ccacgaacgc cggaagggtg ccgcagcgtg tggattgct ctcaattctc 6780
 tcttgagga atgcaatgat gaatatgata ctgactatga aactttgagg gaatactgcc 6840
 tagcaccgtc acctcataac gtgcatcatg catgccctga caacatggaa catcgctatt 6900
 tttctgaaga attatgctcg ttggaggatg tcgcggcaat tgcagctatt gccaacatcg 6960
 aactaccct cagcatgca ttcataata ttattcatgc ggggaaaggc aagattaatc 7020
 caactggcaa atcatccagc gtgattggtg acttcagttc cagcgacttg attcgtttg 7080
 gtgctacca cgttttcaat aaggacgaga tgggtggagta aagaaggagt gcgtcgaagc 7140
 agatcgttca aacatttggc aataaagttt cttaagattg aatcctgttg ccggtcttgc 7200
 gatgattatc atataatttc tgttgaatta cgtaagcat gtaataatta acatgtaatg 7260
 catgacgtta tttatgagat gggtttttat gattagagtc ccgcaattat acatttaata 7320
 cgcgatagaa aacaaaatat agcgcgcaaa ctaggataaa ttatcgcgcg cgggtgtcatc 7380
 tatgttacta gatcgaatta attccaggcg gtgaagggca atcagctgtt gcccgctca 7440
 ctggtgaaaa gaaaaaccac ccagtagat taaaaacgtc cgcaatgtgt tattaagttg 7500
 tctaagcgtc aatttggtta caccacaata tatcctgcca ccagccagcc aacagctccc 7560
 cgaccggcag ctcggcacaa aatcaccact cgatacaggc agcccatcag tccgggacgg 7620
 cgtcagcggg agagccgttg taaggcggca gactttgctc atgttaccga tgctattcgg 7680
 aagaacggca actaagctgc cgggtttgaa acacggatga tctcgcgag ggtagcatgt 7740
 tgattgtaac gatgacagag cgttgctgcc tgtgatcaaa tatcatctcc ctgcagaga 7800
 tccgaattat cagccttctt attcatttct cgcttaaccg tgacaggctg tcgatcttga 7860
 gaactatgcc gacataatag gaaatcgctg gataaagccg ctgaggaagc tgagtggcgc 7920
 tatttcttta gaagtgaacg ttgacgatgt cgac 7954

<210> 19

<211> 41

<212> DNA

<213> Artificial

<220>

<223> oligonucleotídeo XylF9

<400> 19

ggccggatcc tcgagacaca attggaggaa acatggaaaag c

41

<210> 20

<211> 40

<212> DNA

<213> Artificial

<220>

<223> oligonucleotídeo XylR9

<400> 20

ggccatcgat ggtaccggcc cagctcttta tggaatcaaa

40

<210> 21

<211> 436

<212> DNA

<213> Artificial

<220>

<223> XylTiNb

<400> 21

ggccggatcc tcgagacaca attggaggaa acatggaaaag cacttttttc aagcctcact 60

tatgctaaga acttttagtgg ccagctttgt ttccgtcatg ccgtcctctc gcctttggga 120

tatgaaactg ccctgttttaa gggactgtca gaaactatag attgtaatgg agcttctgct 180

catgatttgt ggcaaaatcc tgatgataag aaaactgcac ggttatccga gtttggggag 240

atgatcaggg cagccttttg atttctgtt gatagacaga acatcccaag gacagtcaca 300

ggccctaattg tcctctttgt tagacgtgag gattatttag ctcacccacg tcatggtgga 360

aaggtagagt ctaggcttag caatgaagag caagtatttg attccataaa gagctggggc 420

ggtaccatcg atggcc 436

<210> 22

<211> 5093

<212> DNA

<213> Artificial

<220>

<223> T-DNA de pTKW29

<220>

<221> misc_feature

<222> (1)..(25)

<223> repetição de borda direita do T-DNA de *Agrobacterium tumefaciens* (Zambryski, 1988)

<220>

<221> misc_feature

<222> (76)..(786)

<223> seqüência incluindo a região não traduzida 3' do gene de sintase de octopina de *Agrobacterium tumefaciens* conforme descrito por De Greve e outros(1982) (complemento)

<220>

<221> misc_feature

<222> (807)..(1211)

<223> seqüência incluindo uma parte do gene de XylT de *Nicotiana benthamiana* clonado na orientação anti-sentido (complemento)

<220>

<221> misc_feature

<222> (1250)..(1991)

<223> seqüência contendo o segundo íntron do gene de diquinase de ortofosfato de piruvato de *Flaveria trinervia* conforme descrito por Rosche e Westhoff (1995) (complemento)

<220>

<221> misc_feature

<222> (2023)..(2427)

<223> seqüência incluindo uma parte do gene XylT de *Nicotiana benthamiana* clonado na orientação sentido (complemento)

<220>

<221> misc_feature

<222> (2434)..(3779)

<223> sequência incluindo a região promotora do transcrito 35S do vírus mosaico da couve-flor (Odell e outros, 1985) (complemento)

<220>

<221> misc_feature

<222> (3854)..(4140)

<223> sequência incluindo a região promotora do gene de sintase de nopalina de T-DNA de *A. tumefaciens*

<220>

<221> misc_feature

<222> (4161)..(4712)

<223> sequência incluindo o gene de resistência à fosfinotricina bar [De Block e outros (1987) EMBO J 6:2513]

<220>

<221> misc_feature

<222> (4731)..(4991)

<223> sequência incluindo a região não traduzida 3' do gene de sintase de nopalina de T-DNA de *A. tumefaciens*

<220>

<221> misc_feature

<222> (5069)..(5093)

<223> repetição de borda esquerda do T-DNA de *Agrobacterium tumefaciens* (Zambryski, 1988)

<400> 22

aattacaacg gtatatatcc tgccagtact cggccgctga ccgcggtacc ccggaattaa	60
gcttgcacgc ctgcaggtcc tgctgagcct cgacatgttg tcgcaaaatt cgccctggac	120
ccgccccaacg atttgtcgtc actgtcaagg tttagacctgc acttcatttg gggccccacat	180
acaccaaaaa aatgctgcat aattctcggg gcagcaagtc ggttaccggg ccgcggtgct	240
ggaccggggtt gaatggtgcc cgtaactttc ggtagagcgg acggccaata ctcaacttca	300
aggaatctca cccatgcgcg ccggcgggga accggagtgc ccttcagtga acgttattag	360
ttcgccgctc ggtgtgtcgt agatactagc ccctggggcc ttttgaaatt tgaataagat	420
ttatgtaatc agtccttttag gtttgaccgg ttctgccgct ttttttaaaa ttggatttgt	480
aataataaaa cgcaattggt tgttattgtg gcgctctatc atagatgtcg ctataaacct	540

attcagcaca atatatattggtt ttcatttttaa tattgtacat ataagtagta ggggtacaatc	600
agtaaattga acggagaata ttattcataa aaatacgata gtaacgggtg atatatcat	660
tagaatgaac cgaaaccggc ggtaaggatc tgagctacac atgctcaggt tttttacaac	720
gtgcacaaca gaattgaaag caaatatcat gcgatcatag gcgtctcgca tatctcatta	780
aagcaggact ctagaggatc ctcgagacac aattggagga aacatggaaa gcactttttt	840
caagcctcac ttatgctaag aacttttagtg gccagtttg tttccgtcat gccgtcctct	900
cgcctttggg atatgaaact gccctgttta agggactgtc agaaactata gattgtaatg	960
gagcttctgc tcatgatttg tggcaaaatc ctgatgataa gaaaactgca cggttatccg	1020
agtttgggga gatgatcagg gcagcctttg gatttcctgt tgatagacag aacatcccaa	1080
ggacagtcac aggcctaata gtcctctttg ttagacgtga ggattattta gtcacccac	1140
gtcatggttg aaagggtacag tctaggctta gcaatgaaga gcaagtattt gattccataa	1200
agagctgggc cgggtaccatc gatttcgaac ccaatttccc aactgtaatc aatccaaatg	1260
taagatcaat gataacacaa tgacatgatc tatcatgtta ccttgtttat tcatgttcga	1320
ctaattcatt taattaatag tcaatccatt tagaagttaa taaaactaca agtattattt	1380
agaaattaat aagaatgttg attgaaaata atactatata aatgataga tcttgcgctt	1440
tggtatatta gcattagatt atgttttggt acattagatt actgtttcta ttagtttgat	1500
attatttggt actttagctt gttatttaat attttgttta ttgataaatt acaagcagat	1560
tggaatttct aacaaaatat ttattaactt ttaactaaa atatttagta atggtataga	1620
tatttaatta tataataaac tattaatcat aaaaaatat tattttaatt tatttattct	1680
tatttttact atagtatttt atcattgata tttaattcat caaaccagct agaattacta	1740
ttatgattaa aacaaatatt aatgctagta tatcatctta catgttcgat caaattcatt	1800
aaaaataata tacttactct caacttttat cttcttcgtc ttacacatca cttgtcatat	1860
ttttttacat tactatgttg tttatgtaaa caatatattt ataaattatt ttttcacaat	1920
tataacaact atattattat aatcatacta attaacatca ctttaactatt ttatactaaa	1980
aggaaaaaag aaaataatta tttccttacc aattggggta ccggcccagc tctttatgga	2040
atcaaatact tgctcttcat tgctaagcct agactgtacc tttccaccat gacgtgggtg	2100
agctaaataa tcctcacgtc taacaaagag gacattaggg cctgtgactg tccttgggat	2160
gttctgtcta tcaacaggaa atccaaaggc tgccctgatc atctcccaa actcggataa	2220
ccgtgcagtt ttcttatcat caggattttg ccacaaatca tgagcagaag ctccattaca	2280
atctatagtt tctgacagtc ccttaaacag ggcagtttca tatcccaaag gcgagaggac	2340
ggcatgacgg aaacaaactg ggccactaaa gttcttagca taagtgaggc ttgaaaaaag	2400
tgctttccat gtttctcca attgtgtctc gagcgtgtcc tctccaaatg aatgaactt	2460

ccttatatag aggaaggggc ttgcgaagga tagtgggatt gtgcgtcatc ccttacgtca	2520
gtggagatgt cacatcaatc cacttgcttt gaagacgtgg ttggaacgtc ttctttttcc	2580
acgatgctcc tcgtgggtgg ggggccatct ttgggaccac tgtcggcaga gagatcttga	2640
atgatagcct ttcttttata gcaatgatgg catttgtagg agccaccttc cttttctact	2700
gtccttttga tgaagtgaca gatagctggg caatggaatc cgaggagggt tcccgaaatt	2760
atcctttgtt gaaaagtctc aatagccctt tgggtctctg agactgtatc tttgacattt	2820
ttggagtaga ccagagtgtc gtgctccacc atgttgacga agattttctt cttgtcattg	2880
agtcgtaaaa gactctgtat gaactgttcg ccagctctca cggcgagtgc tggtagatcc	2940
tcgatttgaa tcttagactc catgcatggc cttagattca gtaggaacta cttttttaga	3000
gactccaatc tctattactt gccttggttt atgaagcaag ctttgaatcg tccatactgg	3060
aatagtactt ctgatcttga gaaatatgtc tttctctgtg ttcttgatgc aattagtctt	3120
gaatcttttg actgcatctt taaccttctt gggaagggtat ttgatctcct ggagattggt	3180
actcgggtag atcgtcttga tgagacctgc tgcgtaggcc tctctaacca tctgtgggtc	3240
agcattcttt ctgaaattga agaggctaac cttctcatta tcagtgggtga acatagtgtc	3300
gtcaccttca ctttcgaact tccttcctag atcgtaaaga tagaggaaat cgtccattgt	3360
aatctccggg gcaaaggaga tctcttttgg ggctggatca ctgctgggcc ttttggttcc	3420
tagcgtgagc cagtgggctt tttgctttgg tgggcttgtt agggccttag caaagctctt	3480
gggcttgagt tgagcttctc ctttggggat gaagttcaac ctgtctgttt gctgacttgt	3540
tgtgtacgcg tcagctgctg ctcttgctc tgtaatagtg gcaaatttct tgtgtgcaac	3600
tccgggaacg ccgtttgttg ccgcctttgt acaacccag tcacgtata taccggcatg	3660
tggaccgtta tacacaacgt agtagttgat atgagggtgt tgaatacccg attctgctct	3720
gagaggagca actgtgctgt taagctcaga tttttgtggg attggaatta attcgtcgag	3780
cggccgctcg acgagctccc tatagtgagt cgtattagag gccgacttgg gggccggcca	3840
tttaaatgaa ttgatcatg agcggagaat taaggagtc acgttatgac ccccgcgat	3900
gacgcgggac aagccgtttt acgttttgaa ctgacagaac cgcaacgatt gaaggagcca	3960
ctcagccgcg ggtttctgga gtttaatgag ctaagcacat acgtcagaaa ccattattgc	4020
gcgttcaaaa gtcgcctaag gtcactatca gctagcaaat atttcttgtc aaaaatgtc	4080
cactgacgtt ccataaatc ccctcggat ccaattagag tctcatatc actctcaatc	4140
aaagatccgg cccatgatcc atggaccag aacgacgcc ggccgacatc cgccgtgcca	4200
ccgaggcgga catgccggcg gtctgcacca tcgtcaacca ctacatcgag acaagcacgg	4260
tcaacttccg taccgagccg caggaaccgc aggagtggac ggacgacctc gtccgtctgc	4320
gggagcgcta tccctggctc gtcgccgagg tggacggcga ggtcgcgggc atcgccctacg	4380

cgggcccctg gaaggcacgc aacgcctacg actggacggc cgagtcgacc gtgtacgtct 4440
 ccccccgcca ccagcggacg ggactgggct ccacgtctta caccacctg ctgaagtcce 4500
 tggaggcaca gggcttcaag agcgtggtcg ctgtcatcgg gctgcccac gacccgagcg 4560
 tgcgcatgca cgaggcgcgc ggatatgccc cccgcggcat gctgcgggcg gccggcttca 4620
 agcacgggaa ctggcatgac gtgggtttct ggcagctgga cttcagcctg ccggtaccgc 4680
 cccgtccggt cctgcccgtc accgagatct gatctcacgc gtctaggatc cgaagcagat 4740
 cgttcaaaca tttggcaata aagtttctta agattgaatc ctgttgccgg tcttgcatg 4800
 attatcatat aatttctgtt gaattacgtt aagcatgtaa taattaacat gtaatgcatg 4860
 acgttattta tgagatgggt ttttatgatt agagtccgc aattatacat ttaatacgcg 4920
 atagaaaaca aaatatagcg cgcaaactag gataaattat cgcgcgcggt gtcattctatg 4980
 ttactagatc gggaagatcc attaattcgg tactcgaggc attacggcat tacggcactc 5040
 gcgaggggtcc caattcgagc atggagccat ttacaattga atatatcctg ccg 5093

<210> 23

<211> 2318

<212> DNA

<213> *Nicotiana tabacum* cv. Xanthi

<220>

<221> 5'UTR

<222> (1)..(18)

<220>

<221> CDS

<222> (19)..(1590)

<220>

<221> 3'UTR

<222> (1592)..(2291)

<220>

<221> sinal_poliA

<222> (2292)..(2318)

<400> 23

agtcagagag agaagaag atg aac aag aaa aag ctg aaa ttt ctt gtt tct 51

Met Asn Lys Lys Lys Leu Lys Phe Leu Val Ser

ctc ttc gct ctc aac tca atc act ctc tat ctc tac ttc tct tcc cac	99
Leu Phe Ala Leu Asn Ser Ile Thr Leu Tyr Leu Tyr Phe Ser Ser His	
15 20 25	
tct gat cac ttc cgt cac aaa tcc ccc caa aac cac ttt cct aat acc	147
Ser Asp His Phe Arg His Lys Ser Pro Gln Asn His Phe Pro Asn Thr	
30 35 40	
caa aac cac tat tcc ctg tcg gaa aac cac cat gat aat ttc cac tct	195
Gln Asn His Tyr Ser Leu Ser Glu Asn His His Asp Asn Phe His Ser	
45 50 55	
tct gtc act tcc caa tat acc aag cct tgg cca att ttg ccc tcc tac	243
Ser Val Thr Ser Gln Tyr Thr Lys Pro Trp Pro Ile Leu Pro Ser Tyr	
60 65 70 75	
ctc ccc tgg tct cag aat cct aat gtt tct ttg aga tcg tgc gag ggt	291
Leu Pro Trp Ser Gln Asn Pro Asn Val Ser Leu Arg Ser Cys Glu Gly	
80 85 90	
tac ttc ggt aat ggg ttt act ctc aaa gtt gat ctt ctc aaa act tcg	339
Tyr Phe Gly Asn Gly Phe Thr Leu Lys Val Asp Leu Leu Lys Thr Ser	
95 100 105	
ccg gag ctt cac cag aaa ttc ggc gaa aac acc gta tcc ggc gac ggc	387
Pro Glu Leu His Gln Lys Phe Gly Glu Asn Thr Val Ser Gly Asp Gly	
110 115 120	
gga tgg ttt agg tgt ttt ttc agt gag act ttg cag agt tcg att tgc	435
Gly Trp Phe Arg Cys Phe Phe Ser Glu Thr Leu Gln Ser Ser Ile Cys	
125 130 135	
gag gga ggt gct ata cga atg aat ccg gac gag att ttg atg tct cgt	483
Glu Gly Gly Ala Ile Arg Met Asn Pro Asp Glu Ile Leu Met Ser Arg	
140 145 150 155	
gga ggc gag aaa ttg gag tcg gtt att ggt agg agt gaa gat gat gag	531
Gly Gly Glu Lys Leu Glu Ser Val Ile Gly Arg Ser Glu Asp Asp Glu	
160 165 170	
ctg ccc gtg ttc aaa aat gga gct ttt cag att aaa gtt act gat aaa	579
Leu Pro Val Phe Lys Asn Gly Ala Phe Gln Ile Lys Val Thr Asp Lys	

175	180	185	
ctg aaa att ggg aaa aaa tta gtg gat gaa aaa atc ttg aat aaa tac			627
Leu Lys Ile Gly Lys Lys Leu Val Asp Glu Lys Ile Leu Asn Lys Tyr			
190	195	200	
tta ccg gaa ggt gca att tca agg cac act atg cgt gaa tta att gac			675
Leu Pro Glu Gly Ala Ile Ser Arg His Thr Met Arg Glu Leu Ile Asp			
205	210	215	
tct att cag tta gtt ggc gcc gat gaa ttt cac tgt tct gag tgg att			723
Ser Ile Gln Leu Val Gly Ala Asp Glu Phe His Cys Ser Glu Trp Ile			
220	225	230	235
gag gag ccg tca ctt ttg att aca cga ttt gag tat gca aac ctt ttc			771
Glu Glu Pro Ser Leu Leu Ile Thr Arg Phe Glu Tyr Ala Asn Leu Phe			
240	245	250	
cac aca gtt acc gat tgg tat agt gca tac gtg gca tcc agg gtt act			819
His Thr Val Thr Asp Trp Tyr Ser Ala Tyr Val Ala Ser Arg Val Thr			
255	260	265	
ggc ttg ccc agt cgg cca cat ttg gtt ttt gta gat ggc cat tgt gag			867
Gly Leu Pro Ser Arg Pro His Leu Val Phe Val Asp Gly His Cys Glu			
270	275	280	
aca caa ttg gag gaa aca tgg aaa gca ctc ttt tca agc ctc act tat			915
Thr Gln Leu Glu Glu Thr Trp Lys Ala Leu Phe Ser Ser Leu Thr Tyr			
285	290	295	
gct aag aac ttt agt ggc cca gtt tgt ttc cgt cac gcc gtt ctc tcg			963
Ala Lys Asn Phe Ser Gly Pro Val Cys Phe Arg His Ala Val Leu Ser			
300	305	310	315
cct ttg gga tat gaa act gcc ctg ttt aag gga ctg aca gaa act ata			1011
Pro Leu Gly Tyr Glu Thr Ala Leu Phe Lys Gly Leu Thr Glu Thr Ile			
320	325	330	
gat tgt aat gga gct tct gcc cat gat ttg tgg caa aat cct gat gat			1059
Asp Cys Asn Gly Ala Ser Ala His Asp Leu Trp Gln Asn Pro Asp Asp			
335	340	345	
aag aga act gca cgg ttg tct gag ttt ggg gag atg atc agg gca gcc			1107

Lys Arg Thr Ala Arg Leu Ser Glu Phe Gly Glu Met Ile Arg Ala Ala
 350 355 360
 ttt gga ttt cct gtg gat aga cag aac atc cca agg aca gtc aca ggc 1155
 Phe Gly Phe Pro Val Asp Arg Gln Asn Ile Pro Arg Thr Val Thr Gly
 365 370 375
 cct aat gtc ctc ttt gtt aga cgt gag gat tat tta gct cac cca cgt 1203
 Pro Asn Val Leu Phe Val Arg Arg Glu Asp Tyr Leu Ala His Pro Arg
 380 385 390 395
 cat ggt gga aag gta cag tct agg ctt agc aat gaa gag caa gta ttt 1251
 His Gly Gly Lys Val Gln Ser Arg Leu Ser Asn Glu Glu Gln Val Phe
 400 405 410
 gat tcc ata aag agc tgg gcc ttg aac cac tcg gag tgc aaa tta aat 1299
 Asp Ser Ile Lys Ser Trp Ala Leu Asn His Ser Glu Cys Lys Leu Asn
 415 420 425
 gta att aac gga ttg ttt gcc cac atg tcc atg aaa gag caa gtt cga 1347
 Val Ile Asn Gly Leu Phe Ala His Met Ser Met Lys Glu Gln Val Arg
 430 435 440
 gca atc caa gat gct tct gtc ata gtt ggt gct cat gga gca ggt cta 1395
 Ala Ile Gln Asp Ala Ser Val Ile Val Gly Ala His Gly Ala Gly Leu
 445 450 455
 act cac ata gtt tct gca gca cca aaa gct gta ata cta gaa att ata 1443
 Thr His Ile Val Ser Ala Ala Pro Lys Ala Val Ile Leu Glu Ile Ile
 460 465 470 475
 agc agc gaa tat agg cgc ccc cat ttt gct ctg att gca caa tgg aaa 1491
 Ser Ser Glu Tyr Arg Arg Pro His Phe Ala Leu Ile Ala Gln Trp Lys
 480 485 490
 gga ttg gag tac cat ccc ata tat ttg gag ggg tct tat gcg gat cct 1539
 Gly Leu Glu Tyr His Pro Ile Tyr Leu Glu Gly Ser Tyr Ala Asp Pro
 495 500 505
 cca gtt gtg atc gac agg ctc agc agc att ttg agg agt ctt ggg tgc 1587
 Pro Val Val Ile Asp Arg Leu Ser Ser Ile Leu Arg Ser Leu Gly Cys
 510 515 520

```

taa gtccgctcga cagtttgaat agttcggctt ttctctaaaa gacggggaag      1640
gatagaggaa ttcgggggttc tggaacttgg agcctgggaa ttttgataaa tatgtttcac  1700
acgcagttct gtagtcaatg gttgcaatct aggtcctcaa tctgggtgtg ataagcttgg  1760
caatttccag cagctactaa tttattagcc cgctctgact cggttatgga ctaccagagc  1820
aatcatatca aatggaagca tggaatcctg attgtggaat ggtgagctca ttgaagagca  1880
tattctttat ggtgttgaag attacaattc acaattaaca cgtgtatgtg aaagattagg  1940
ttggtacact tacttacaat tcattgtcaa ttgtttttca ttattctcat taatgatcat  2000
aggataagaa catgagaaaa ccatccatgt tctgtgttgt tttcccatca atccggccac  2060
cctcttcctt ccttatgtag agatgatttc aacagagttt gttttgtagt tgtaacactt  2120
gcactcccag ttacagtttt gcattcgaca cattcatccc atcagatgtc aagtttaaag  2180
gcataagaca tttgacatat tgaagaagca gattaacacg aacgtcagta tgatgcttca  2240
gtgaagatat ggttgaact tgtaaccaa caaaagaaat gagactttga caaaaaaaaaa  2300
aaaaaaaaaa aaaaaaaaaa      2318

```

<210> 24

<211> 523

<212> PRT

<213> Nicotiana tabacum cv. Xanthi

<400> 24

```

Met Asn Lys Lys Lys Leu Lys Phe Leu Val Ser Leu Phe Ala Leu Asn
1           5           10           15
Ser Ile Thr Leu Tyr Leu Tyr Phe Ser Ser His Ser Asp His Phe Arg
           20           25           30
His Lys Ser Pro Gln Asn His Phe Pro Asn Thr Gln Asn His Tyr Ser
           35           40           45
Leu Ser Glu Asn His His Asp Asn Phe His Ser Ser Val Thr Ser Gln
           50           55           60
Tyr Thr Lys Pro Trp Pro Ile Leu Pro Ser Tyr Leu Pro Trp Ser Gln
65           70           75           80
Asn Pro Asn Val Ser Leu Arg Ser Cys Glu Gly Tyr Phe Gly Asn Gly
           85           90           95
Phe Thr Leu Lys Val Asp Leu Leu Lys Thr Ser Pro Glu Leu His Gln

```

100	105	110
Lys Phe Gly Glu Asn Thr Val Ser Gly Asp Gly Gly Trp Phe Arg Cys		
115	120	125
Phe Phe Ser Glu Thr Leu Gln Ser Ser Ile Cys Glu Gly Gly Ala Ile		
130	135	140
Arg Met Asn Pro Asp Glu Ile Leu Met Ser Arg Gly Gly Glu Lys Leu		
145	150	155
Glu Ser Val Ile Gly Arg Ser Glu Asp Asp Glu Leu Pro Val Phe Lys		
165	170	175
Asn Gly Ala Phe Gln Ile Lys Val Thr Asp Lys Leu Lys Ile Gly Lys		
180	185	190
Lys Leu Val Asp Glu Lys Ile Leu Asn Lys Tyr Leu Pro Glu Gly Ala		
195	200	205
Ile Ser Arg His Thr Met Arg Glu Leu Ile Asp Ser Ile Gln Leu Val		
210	215	220
Gly Ala Asp Glu Phe His Cys Ser Glu Trp Ile Glu Glu Pro Ser Leu		
225	230	235
Leu Ile Thr Arg Phe Glu Tyr Ala Asn Leu Phe His Thr Val Thr Asp		
245	250	255
Trp Tyr Ser Ala Tyr Val Ala Ser Arg Val Thr Gly Leu Pro Ser Arg		
260	265	270
Pro His Leu Val Phe Val Asp Gly His Cys Glu Thr Gln Leu Glu Glu		
275	280	285
Thr Trp Lys Ala Leu Phe Ser Ser Leu Thr Tyr Ala Lys Asn Phe Ser		
290	295	300
Gly Pro Val Cys Phe Arg His Ala Val Leu Ser Pro Leu Gly Tyr Glu		
305	310	315
Thr Ala Leu Phe Lys Gly Leu Thr Glu Thr Ile Asp Cys Asn Gly Ala		
325	330	335
Ser Ala His Asp Leu Trp Gln Asn Pro Asp Asp Lys Arg Thr Ala Arg		
340	345	350
Leu Ser Glu Phe Gly Glu Met Ile Arg Ala Ala Phe Gly Phe Pro Val		

355	360	365
Asp Arg Gln Asn Ile Pro Arg Thr Val Thr Gly Pro Asn Val Leu Phe		
370	375	380
Val Arg Arg Glu Asp Tyr Leu Ala His Pro Arg His Gly Gly Lys Val		
385	390	395
Gln Ser Arg Leu Ser Asn Glu Glu Gln Val Phe Asp Ser Ile Lys Ser		
405	410	415
Trp Ala Leu Asn His Ser Glu Cys Lys Leu Asn Val Ile Asn Gly Leu		
420	425	430
Phe Ala His Met Ser Met Lys Glu Gln Val Arg Ala Ile Gln Asp Ala		
435	440	445
Ser Val Ile Val Gly Ala His Gly Ala Gly Leu Thr His Ile Val Ser		
450	455	460
Ala Ala Pro Lys Ala Val Ile Leu Glu Ile Ile Ser Ser Glu Tyr Arg		
465	470	475
Arg Pro His Phe Ala Leu Ile Ala Gln Trp Lys Gly Leu Glu Tyr His		
485	490	495
Pro Ile Tyr Leu Glu Gly Ser Tyr Ala Asp Pro Pro Val Val Ile Asp		
500	505	510
Arg Leu Ser Ser Ile Leu Arg Ser Leu Gly Cys		

515

520

<210> 25

<211> 4973

<212> DNA

<213> Nicotiana tabacum cv. Xanthi

<220>

<221> 5'UTR

<222> (1)..(18)

<220>

<221> CDS

<222> (19)..(696)

<220>

<221> Intron

<222> (697)..(2028)

<220>

<221> CDS

<222> (2029)..(2178)

<220>

<221> Intron

<222> (2179)..(3606)

<220>

<221> CDS

<222> (3607)..(4326)

<220>

<221> 3'UTR

<222> (4330)..(4973)

<400> 25

```

agtcagagag agaagaag atg aac aag aaa aag ctg aaa att ctt gtt tct      51
      Met Asn Lys Lys Lys Leu Lys Ile Leu Val Ser
              1              5              10

ctc ttc gct ctc aac tca atc act ctc tat ctc tac ttc tct tcc cac      99
Leu Phe Ala Leu Asn Ser Ile Thr Leu Tyr Leu Tyr Phe Ser Ser His
              15              20              25

cct gat cac ttc cgc cac aaa tcc cgc caa aac cac ttt tcc ttg tcg      147
Pro Asp His Phe Arg His Lys Ser Arg Gln Asn His Phe Ser Leu Ser
              30              35              40

gaa aac cgc cat cat aat ttc cac tct tca atc act tct caa tat tcc      195
Glu Asn Arg His His Asn Phe His Ser Ser Ile Thr Ser Gln Tyr Ser
              45              50              55

aag cct tgg cct att ttg ccc tcc tac ctc cct tgg tct caa aac cct      243
Lys Pro Trp Pro Ile Leu Pro Ser Tyr Leu Pro Trp Ser Gln Asn Pro
              60              65              70              75

aat gtt gtt tgg aga tcg tgc gag ggt tac ttc ggt aat ggg ttt act      291

```

Asn Val Val Trp Arg Ser Cys Glu Gly Tyr Phe Gly Asn Gly Phe Thr	
80 85 90	
ctc aaa gtt gac ctt ctc aaa act tcg ccg gag ttt cac cgg aaa ttc	339
Leu Lys Val Asp Leu Leu Lys Thr Ser Pro Glu Phe His Arg Lys Phe	
95 100 105	
ggc gaa aac acc gtc tcc ggc gac ggc gga tgg ttt agg tgt ttt ttc	387
Gly Glu Asn Thr Val Ser Gly Asp Gly Gly Trp Phe Arg Cys Phe Phe	
110 115 120	
agt gag act ttg cag agt tcg atc tgc gag gga ggc gca ata cga atg	435
Ser Glu Thr Leu Gln Ser Ser Ile Cys Glu Gly Gly Ala Ile Arg Met	
125 130 135	
aat ccg gac gag att ttg atg tct cgt gga ggt gag aaa ttg gag tcg	483
Asn Pro Asp Glu Ile Leu Met Ser Arg Gly Gly Glu Lys Leu Glu Ser	
140 145 150 155	
gtt att ggt agg agt gaa gat gat gag ctg ccc gtg ttc aaa aat gga	531
Val Ile Gly Arg Ser Glu Asp Asp Glu Leu Pro Val Phe Lys Asn Gly	
160 165 170	
gct ttt cag att aaa gtt act gat aaa ctg aaa att ggg aaa aaa tta	579
Ala Phe Gln Ile Lys Val Thr Asp Lys Leu Lys Ile Gly Lys Lys Leu	
175 180 185	
gtg gat gaa aaa ttc ttg aat aaa tac tta ccg gaa ggt gca att tca	627
Val Asp Glu Lys Phe Leu Asn Lys Tyr Leu Pro Glu Gly Ala Ile Ser	
190 195 200	
agg cac act atg cgt gag tta att gac tct att cag ttg gtt ggc gcc	675
Arg His Thr Met Arg Glu Leu Ile Asp Ser Ile Gln Leu Val Gly Ala	
205 210 215	
gat gat ttt cac tgt tct gag gtagatttt gaattttgtt tgctctttaa	726
Asp Asp Phe His Cys Ser Glu	
220 225	
attaaaggtt taaactttgt gaatgttggc agatatggaa tacactaatg gattttgttt	786
gatctgttta atgaagattg tctagaacct caatgttata aatatggttt gggtgttca	846
ttaattaaag agaattcctt aatatcccgga ctagatgccca gataacacca gtagttgac	906

ttttggatta ttggttgcatt ttcatttgat cagataaatt gttcattctt aaatgtttca 966
 ctaaagaatt actcaagatt tcagagttta tatgtaggtg tatgtatttg gaattctgga 1026
 tttggatcta gtattgaatg gattactgaa cttgtactcc ccagtcactt ggggaggagc 1086
 aatagatcaa attcaagggt tgaaaagtaa tactgagtc gaaattaacc actttaactt 1146
 ggaaaacggt aaatgtatgt gttctaagat gggtattcct ataacttttg atgtctaata 1206
 tggagaaagt gagttgattt atgctttttc cttttccctt tattggtgtt gggttttaaa 1266
 ttctatcaat tcctttgttt gattgctact caaattgaac cttagacgga gtagcaatag 1326
 caaaaagtga aggccattct tttctcctt catctcttta tttccgtttg acatacagaa 1386
 tatggtagca tctgtctgaa gtggtaatt ttattcctta aaatttgcat aactaattcg 1446
 agtaaatgcc ttttgaagct ttagttgaat agttctacaa ctggttggtt cattttgagg 1506
 actatcgact tgatttgaca cttgacattg tctgatacat ggcttgtaag ttatgaaaac 1566
 ttttatctag gaagaaatcc caaccagaga tagggagctg tcacttggtt atgagctact 1626
 ggctcaaagt tcaagtttga ccagttaatt ttagatcttc accaggataa catttagagt 1686
 ctaatcaaatt tctgaagcag tattgtgcac taataagagg aacacatgaa ggatgtagca 1746
 ctactagggt atgttacctt atttactaat gattgacaac cagcttaaat gatgacaaat 1806
 agtcttatat ttgctttttc acattgctca tgacttgga tatttccgaa tcaacatatt 1866
 ttagttcttt atgtacttaa ctacttatca aaaaattatc cctgctagat gttagtgttc 1926
 aagcaaccat gctagctttt aaggaagctc cttctttgat tcatgccatc tttccgaaat 1986
 cgatgcctta cgttactgtc atttttctaa ttttcatttc ag tgg att gag gag 2040

Trp Ile Glu Glu

230

ccg tca ctt ttg att aca cga ttt gag tat gca aac ctt ttc cac aca 2088

Pro Ser Leu Leu Ile Thr Arg Phe Glu Tyr Ala Asn Leu Phe His Thr

235

240

245

gtt acc gat tgg tat agt gca tac gtg gca tcc agg gtt act ggc ttg 2136

Val Thr Asp Trp Tyr Ser Ala Tyr Val Ala Ser Arg Val Thr Gly Leu

250

255

260

ccc agt cgg cca cat ttg gtt ttt gta gat ggc cat tgt gag 2178

Pro Ser Arg Pro His Leu Val Phe Val Asp Gly His Cys Glu

265

270

275

gtatgtttga aagtattgat aacgatggca tgcattgtac tgtgttacgg atgaaagaaa 2238

tgaaaccagc aattattttc tagcaggcaa tgctcttgag atgcttgtgt caaattggtc 2298

agacttaatc ctgagtttcc atttgtttca gctttctgtt tgactgacta caataattgt 2358
 cccaattagg ggtgtcaatg gatattcgaa agccgactaa accgaccgaa ccgtaccgta 2418
 ccgattttta gggttctttt taagaaaccg taggttttta tataaatcta taaccgcacc 2478
 gataattagg gtaggttttt tattttataa aaataagccg aaaaaatacc gaaccgtacc 2538
 gaataaattt tacatgtgga aaatatattt atttagtaag tttaaaaata ataatgcatt 2598
 aaattttctt tgggccatgg aattatgaaa actattacaa gccacaagt aattagactc 2658
 aaaatactaa ttcctaaaac ctattatggt acttctactt aaactaagtt atttcaagta 2718
 tctttattag caagacacaa agtattctag cgattatgag caaactacaa tgtattgaat 2778
 atgtttcctt tcatatattt tagatttatt tttttgaata tttaatcttc tatagactct 2838
 attcttgagt ccagcttggt tatatctttc aactcgtgtg atttatattt tctttgcctt 2898
 tgtttgattt cttttacgtt gttgtagaat agtcgatgga tctatactct agccatcttt 2958
 ctttttcttt tttaattca tcacctttta aacagtaaaa atgtctaaag aatttttcta 3018
 agtcctataa aagaacgtat gttattgcat tctacttcta ctggtgaatt ttacatgata 3078
 ttaaaaaatt aaccgaacct taccgctacc gaagagaaac cgacatgatt gggacggttt 3138
 cgaaaagtct aattttgggt atacataata gaataaccga aaaattggta tggtaacaaat 3198
 ttataaaaat aaccggccga accgaacct tgacaccct agtccaata cctagttgtt 3258
 gcagtttgct cattcttact tctatttacg tcaactgttc tctgaatggt ccctttgtgg 3318
 tgaaaagagc ttttgctatg tagaaaaact agcaatgatt tcatagctga aacaatttat 3378
 ttttacctta catcatgtct tataaaattg cttctaactg tatactttaa ttcttgagga 3438
 gatgctttca tgtgaagaaa gttctttcac tccactactg gaagcttgcc gtatgaattt 3498
 tacttgacca tattgtggcc gtgcttgat ttatcttcaa attcattttc ttcatatagt 3558
 tctttcgagt aattcttttt tctctttttc tgtttgaaaa aaattcag aca caa ttg 3615

Thr Gln Leu

gag gaa aca tgg aaa gca ctt ttt tca agc ctc act tat gct aag aac 3663
 Glu Glu Thr Trp Lys Ala Leu Phe Ser Ser Leu Thr Tyr Ala Lys Asn
 280 285 290 295
 ttt agt ggc cca gtt tgt ttc cgt cat gcc gcc ctc tcg cct ttg gga 3711
 Phe Ser Gly Pro Val Cys Phe Arg His Ala Ala Leu Ser Pro Leu Gly
 300 305 310
 tat gaa act gcc ctg ttt aag gga ctg tca gaa act ata gat tgt aat 3759
 Tyr Glu Thr Ala Leu Phe Lys Gly Leu Ser Glu Thr Ile Asp Cys Asn

315	320	325	
gga gct tct gcc cat gat ttg tgg caa aat cct gat gat aag aaa act			3807
Gly Ala Ser Ala His Asp Leu Trp Gln Asn Pro Asp Asp Lys Lys Thr			
330	335	340	
gca cgg ttg tcc gag ttt ggg gag atg att agg gca gcc ttt aga ttt			3855
Ala Arg Leu Ser Glu Phe Gly Glu Met Ile Arg Ala Ala Phe Arg Phe			
345	350	355	
cct gtg gat aga cag aac atc cca agg aca gtc aca ggc cct aat gtc			3903
Pro Val Asp Arg Gln Asn Ile Pro Arg Thr Val Thr Gly Pro Asn Val			
360	365	370	375
ctc ttt gtt aga cgt gag gat tat tta gct cac cca cgt cat ggt gga			3951
Leu Phe Val Arg Arg Glu Asp Tyr Leu Ala His Pro Arg His Gly Gly			
380	385	390	
aag gta cag tct agg ctt agc aat gaa gag caa gta ttt gat tcc ata			3999
Lys Val Gln Ser Arg Leu Ser Asn Glu Glu Gln Val Phe Asp Ser Ile			
395	400	405	
aag agc tgg gcc ttg aac cac tcg gag tgc aaa tta aat gta att aac			4047
Lys Ser Trp Ala Leu Asn His Ser Glu Cys Lys Leu Asn Val Ile Asn			
410	415	420	
gga ttg ttt gcc cac atg tcc atg aaa gag caa gtt cga gca atc caa			4095
Gly Leu Phe Ala His Met Ser Met Lys Glu Gln Val Arg Ala Ile Gln			
425	430	435	
gat gct tct gtc ata gtt ggt gct cat gga gca ggt cta act cac ata			4143
Asp Ala Ser Val Ile Val Gly Ala His Gly Ala Gly Leu Thr His Ile			
440	445	450	455
gtt tct gca gca cca aaa gct gta ata cta gaa att ata agc agc gaa			4191
Val Ser Ala Ala Pro Lys Ala Val Ile Leu Glu Ile Ile Ser Ser Glu			
460	465	470	
tat agg cgc ccc cat ttt gct ctg att gca caa tgg aaa gga ttg gag			4239
Tyr Arg Arg Pro His Phe Ala Leu Ile Ala Gln Trp Lys Gly Leu Glu			
475	480	485	
tac cat ccc ata tat ttg gag ggg tct tat gcg gat cct cca gtt gtg			4287

Tyr His Pro Ile Tyr Leu Glu Gly Ser Tyr Ala Asp Pro Pro Val Val
 490 495 500
 atc gac aag ctc agc agc att ttg agg agt ctt ggg tgc taaatctgct 4336
 Ile Asp Lys Leu Ser Ser Ile Leu Arg Ser Leu Gly Cys
 505 510 515
 cgacagttta gttcgtcttt tctctaaaag actgggaagg atagaggaat tcggggttct 4396
 ggaacctgga gcctgggaat tgtgtaaaat atgtttcaca cgacagttcta tagtcaattg 4456
 ctgcaatctg gtgttcataa gcttggaat ttccagcagc tactaactta ttagccact 4516
 ctgactcagt tatggactac cagagcaatc atatcaaag ggagcatgga atcctgattg 4576
 tggaatggtg agctcattga agagcatatt ctttatggtg ttgaagatta cagttgacga 4636
 gtaacacgtg tatgtgaaag attaggttgt tacactttct tgcaattcat tgtcaattgt 4696
 ttttcgtcat tcttattaat gatcatagga taagaacatg agaaaaccat ccatgttctc 4756
 tgttgttttc ccatcaatct ggccaccctc tttcctcttt atgtagagat gatttcaaca 4816
 gagtttggtt tgtagttgta atacttgtag tcacagttac tgttttgcat tcatcccatc 4876
 agatgtcgaa gaagcagatt aacaagaacg tcagtatgat gtttcagtga atatatggtt 4936
 gtaacttgta accaaacaaa agaaatgaga ctttgac 4973

<210> 26

<211> 516

<212> PRT

<213> *Nicotiana tabacum* cv. Xanthi

<400> 26

Met Asn Lys Lys Lys Leu Lys Ile Leu Val Ser Leu Phe Ala Leu Asn
 1 5 10 15
 Ser Ile Thr Leu Tyr Leu Tyr Phe Ser Ser His Pro Asp His Phe Arg
 20 25 30
 His Lys Ser Arg Gln Asn His Phe Ser Leu Ser Glu Asn Arg His His
 35 40 45
 Asn Phe His Ser Ser Ile Thr Ser Gln Tyr Ser Lys Pro Trp Pro Ile
 50 55 60
 Leu Pro Ser Tyr Leu Pro Trp Ser Gln Asn Pro Asn Val Val Trp Arg
 65 70 75 80

Ser Cys Glu Gly Tyr Phe Gly Asn Gly Phe Thr Leu Lys Val Asp Leu			
85	90	95	
Leu Lys Thr Ser Pro Glu Phe His Arg Lys Phe Gly Glu Asn Thr Val			
100	105	110	
Ser Gly Asp Gly Gly Trp Phe Arg Cys Phe Phe Ser Glu Thr Leu Gln			
115	120	125	
Ser Ser Ile Cys Glu Gly Gly Ala Ile Arg Met Asn Pro Asp Glu Ile			
130	135	140	
Leu Met Ser Arg Gly Gly Glu Lys Leu Glu Ser Val Ile Gly Arg Ser			
145	150	155	160
Glu Asp Asp Glu Leu Pro Val Phe Lys Asn Gly Ala Phe Gln Ile Lys			
165	170	175	
Val Thr Asp Lys Leu Lys Ile Gly Lys Lys Leu Val Asp Glu Lys Phe			
180	185	190	
Leu Asn Lys Tyr Leu Pro Glu Gly Ala Ile Ser Arg His Thr Met Arg			
195	200	205	
Glu Leu Ile Asp Ser Ile Gln Leu Val Gly Ala Asp Asp Phe His Cys			
210	215	220	
Ser Glu Trp Ile Glu Glu Pro Ser Leu Leu Ile Thr Arg Phe Glu Tyr			
225	230	235	240
Ala Asn Leu Phe His Thr Val Thr Asp Trp Tyr Ser Ala Tyr Val Ala			
245	250	255	
Ser Arg Val Thr Gly Leu Pro Ser Arg Pro His Leu Val Phe Val Asp			
260	265	270	
Gly His Cys Glu Thr Gln Leu Glu Glu Thr Trp Lys Ala Leu Phe Ser			
275	280	285	
Ser Leu Thr Tyr Ala Lys Asn Phe Ser Gly Pro Val Cys Phe Arg His			
290	295	300	
Ala Ala Leu Ser Pro Leu Gly Tyr Glu Thr Ala Leu Phe Lys Gly Leu			
305	310	315	320
Ser Glu Thr Ile Asp Cys Asn Gly Ala Ser Ala His Asp Leu Trp Gln			
325	330	335	

Asn Pro Asp Asp Lys Lys Thr Ala Arg Leu Ser Glu Phe Gly Glu Met
 340 345 350
 Ile Arg Ala Ala Phe Arg Phe Pro Val Asp Arg Gln Asn Ile Pro Arg
 355 360 365
 Thr Val Thr Gly Pro Asn Val Leu Phe Val Arg Arg Glu Asp Tyr Leu
 370 375 380
 Ala His Pro Arg His Gly Gly Lys Val Gln Ser Arg Leu Ser Asn Glu
 385 390 395 400
 Glu Gln Val Phe Asp Ser Ile Lys Ser Trp Ala Leu Asn His Ser Glu
 405 410 415
 Cys Lys Leu Asn Val Ile Asn Gly Leu Phe Ala His Met Ser Met Lys
 420 425 430
 Glu Gln Val Arg Ala Ile Gln Asp Ala Ser Val Ile Val Gly Ala His
 435 440 445
 Gly Ala Gly Leu Thr His Ile Val Ser Ala Ala Pro Lys Ala Val Ile
 450 455 460
 Leu Glu Ile Ile Ser Ser Glu Tyr Arg Arg Pro His Phe Ala Leu Ile
 465 470 475 480
 Ala Gln Trp Lys Gly Leu Glu Tyr His Pro Ile Tyr Leu Glu Gly Ser
 485 490 495
 Tyr Ala Asp Pro Pro Val Val Ile Asp Lys Leu Ser Ser Ile Leu Arg
 500 505 510
 Ser Leu Gly Cys
 515

REIVINDICAÇÕES

1. Método para produzir uma célula de planta ou planta *Nicotiana* tendo um baixo nível de resíduos de beta-1,2-xilose sobre N-glicanos ligadas à proteína compreendendo as etapas de:

5 a) introdução de um gene quimérico em células de planta de uma espécie ou cultivar de *Nicotiana* para gerar células de planta transgênica, o referido gene quimérico compreendendo os seguintes fragmentos de DNA operavelmente ligados:

i) um promotor expressível em planta;
10 ii) uma região de DNA passível de transcrição compreendendo:

(1) uma primeira região sentido de DNA compreendendo uma seqüência de nucleotídeo de pelo menos 19 de 20 nucleotídeos consecutivos selecionados de uma seqüência de nucleotídeo que codifica uma
15 proteína de XylT de *Nicotiana* ou o complemento da mesma, a referida seqüência de nucleotídeo, de preferência, obténível da referida espécie ou cultivar de *Nicotiana*, em que os referidos pelo menos 19 de 20 nucleotídeos consecutivos codificam pelo menos um aminoácido de XylT espécie ou cultivar de *Nicotiana*-específico ou selecionado de uma seqüência de nucleotídeo de um gene de XylT de *Nicotiana* ou um cDNA de XylT de *Nicotiana* ou
20 o complemento da mesma, a referida seqüência de nucleotídeo, de preferência, obténível da referida espécie ou cultivar de *Nicotiana*, em que os referidos pelo menos 19 de 20 nucleotídeos consecutivos compreendem pelo menos um nucleotídeo de XylT espécie de *Nicotiana*-específico;

25 (2) uma segunda região de anti-sentido de DNA compreendendo uma seqüência de nucleotídeo de pelo menos 19 nucleotídeos consecutivos os quais têm pelo menos 95% de identidade de seqüência ao complemento da referida primeira região de DNA;

em que uma molécula de RNA transcrita a partir da referida região de DNA passível de transcrição é capaz de formar uma região de RNA
30 de filamento duplo pelo menos entre uma região de RNA transcrita a partir da referida primeira região de sentido de DNA e uma região de RNA transcrita

a partir da referida segunda região de anti-sentido de DNA; e

ii) uma região de DNA compreendendo um sinal de poliadenilação e término de transcrição funcional em plantas;

b) opcionalmente, identificação de uma célula de planta *Nicotiana* transgênica a qual tem um menor nível de resíduos de beta-1,2-xilose sobre N-glicanos ligadas à proteína do que uma célula de planta *Nicotiana* não-transformada;

c) opcionalmente, regeneração das referidas células de planta *Nicotiana* transgênicas para obter plantas *Nicotiana* transgênicas; e

d) opcionalmente, identificação de uma planta *Nicotiana* transgênica a qual tem um menor nível de resíduos de beta-1,2-xilose sobre N-glicanos ligadas à proteína do que uma planta *Nicotiana* não-transformada.

2. Método de acordo com a reivindicação 1 em que o referido aminoácido de XylT espécie de *Nicotiana*-específico é um aminoácido de XylT *Nicotiana benthamiana*-específico e a referida espécie *Nicotiana* é, de preferência, *Nicotiana benthamiana*.

3. Método de acordo com a reivindicação 1 ou 2 em que a referida sequência de nucleotídeo que codifica uma proteína de XylT de *Nicotiana* compreende uma sequência de nucleotídeo que codifica a sequência de aminoácido de SEQ ID NO.: 12 ou SEQ ID NO.: 14.

4. Método de acordo com uma das reivindicações 1 a 3 em que o referido nucleotídeo de *Nicotiana* espécies de *Nicotiana*-específico é um nucleotídeo de XylT *Nicotiana benthamiana*-específico e a referida espécie *Nicotiana* é, de preferência, *Nicotiana benthamiana*.

5. Método de acordo com uma das reivindicações 1 a 4 em que a referida sequência de nucleotídeo do referido gene de XylT de *Nicotiana* compreende a sequência de nucleotídeo de SEQ ID NO.: 11, SEQ ID NO.: 13 ou SEQ ID NO. 21.

6. Método de acordo com a reivindicação 1, em que o referido aminoácido de XylT cultivar de *Nicotiana*-específico é um aminoácido de XylT de *Nicotiana tabacum* cv. Petite Havana SR1-específico e o referido cultivar de *Nicotiana* é, de preferência, *Nicotiana tabacum* cv. Petite Havana

SR1.

7. Método de acordo com a reivindicação 1 ou 6, em que a referida seqüência de nucleotídeo que codifica a referida proteína de XylT de *Nicotiana* compreende uma seqüência de nucleotídeo que codifica a seqüência de aminoácido de SEQ ID NO.: 4, SEQ ID NO.: 6, SEQ ID NO.: 8 ou SEQ ID NO.: 10.

8. Método de acordo com qualquer uma das reivindicações 1, 6 ou 7, em que o referido nucleotídeo de XylT cultivar de *Nicotiana*-específico é um nucleotídeo de XylT de *Nicotiana tabacum* cv. Petite Havana SR1-específico e o referido cultivar de *Nicotiana* é, de preferência, *Nicotiana tabacum* cv. Petite Havana SR1.

9. Método de acordo com qualquer uma das reivindicações 1 ou 6 a 8, em que a referida seqüência de nucleotídeo do referido gene de XylT de *Nicotiana* ou do referido cDNA de XylT de *Nicotiana* compreende a seqüência de nucleotídeo de SEQ ID NO.: 3, SEQ ID NO.: 5, SEQ ID NO.: 8, SEQ ID NO.: 10 ou SEQ ID NO.: 17.

10. Método de acordo com qualquer uma das reivindicações 1 a 9, em que as referidas primeira e segunda regiões de DNA compreendem pelo menos 50 nucleotídeos consecutivos.

11. Método de acordo com qualquer uma das reivindicações 1 a 9, em que as referidas primeira e segunda regiões de DNA compreendem pelo menos 200 nucleotídeos consecutivos.

12. Método para produzir uma célula de planta ou planta *Nicotiana* tendo um baixo nível de resíduos de beta-1,2-xilose sobre N-glicanos ligadas à proteína compreendendo as etapas de:

a) fornecimento de uma ou mais moléculas de RNA de filamento duplo às células de planta ou plantas de uma espécie ou cultivar de *Nicotiana*, em que as moléculas de RNA de filamento duplo compreendem dois filamentos de RNA, um filamento de RNA consistindo essencialmente em uma seqüência de nucleotídeo de RNA de 19 de 20 a 21 nucleotídeos consecutivos selecionada de uma seqüência de nucleotídeo que codifica uma proteína de XylT de *Nicotiana*, ou o complemento da mesma, a referida se-

qüência de nucleotídeo, de preferência, obtenível da referida espécie ou cultivar de *Nicotiana*, em que os referidos 19 de 20 a 21 nucleotídeos consecutivos codificam pelo menos um aminoácido de XylT espécie ou cultivar de *Nicotiana*-específico ou selecionada da seqüência de nucleotídeo de um gene de XylT de *Nicotiana* ou um cDNA de XylT de *Nicotiana*, ou o complemento da mesma, a referida seqüência de nucleotídeo, de preferência, obtenível da referida espécie ou cultivar de *Nicotiana*, em que os referidos 19 de 20 a 21 nucleotídeos consecutivos compreendem pelo menos um nucleotídeo de XylT espécie ou cultivar de *Nicotiana*-específico;

- 10 b) identificação de uma célula de planta ou planta *Nicotiana* compreendendo a referida molécula ou moléculas de RNA de filamento duplo as quais têm um menor nível de resíduos de beta-1,2-xilose sobre N-glicanos ligadas à proteína do que a mesma célula de planta ou planta *Nicotiana* a qual não compreende a referida molécula ou moléculas de RNA de filamento duplo.

13. Método de acordo com a reivindicação 12, em que o referido RNA de filamento duplo é fornecido às referidas células de planta ou plantas através de integração de um gene quimérico no genoma de células de planta da referida espécie ou cultivar de *Nicotiana* para gerar células de planta transgênica e, opcionalmente, regeneração das referidas células de planta para obter plantas transgênicas, o referido gene quimérico compreendendo os seguintes fragmentos de DNA operavelmente ligados:

- a) um promotor expressível em planta;
- b) uma região de DNA compreendendo pelo menos 19 de 20 nucleotídeos consecutivos selecionados de uma seqüência de nucleotídeo que codifica uma proteína de XylT de *Nicotiana* ou o complemento da mesma, a referida seqüência de nucleotídeo, de preferência, obtenível da referida espécie ou cultivar de *Nicotiana*, em que os referidos 19 de 20 nucleotídeos consecutivos codificam pelo menos um aminoácido ou XylT espécie ou cultivar de *Nicotiana*-específico ou selecionada da seqüência de nucleotídeo de um gene de XylT de *Nicotiana* ou um cDNA de XylT de *Nicotiana* ou o complemento da mesma, a referida seqüência de nucleotídeo, de preferên-

cia, obtenível da referida espécie ou cultivar de *Nicotiana*, em que os referidos 19 de 20 nucleotídeos consecutivos compreendem pelo menos um nucleotídeo de XylT espécie de *Nicotiana*-específico, na orientação anti-sentido;

- 5 c) uma região de DNA compreendendo um sinal de poliadenilação e de término de transcrição funcional em plantas.

14. Método de acordo com a reivindicação 12, em que o referido RNA de filamento duplo é fornecido às referidas células de planta ou planta através de integração de um gene quimérico no genoma das referidas células de planta para gerar células de planta transgênica e, opcionalmente, regeneração das referidas células de planta para obter plantas transgênicas, o referido gene quimérico compreendendo os seguintes fragmentos de DNA operavelmente ligados:

- a) um promotor expressível em planta;
- 15 b) uma região de DNA compreendendo pelo menos 19 de 20 nucleotídeos consecutivos selecionados de uma seqüência de nucleotídeo que codifica uma proteína de XylT de *Nicotiana* ou o complemento da mesma, a referida seqüência de nucleotídeo, de preferência, obtenível da referida espécie ou cultivar de *Nicotiana*, em que os referidos 19 de 20 nucleotídeos consecutivos codificam pelo menos um aminoácido de XylT espécie ou cultivar de *Nicotiana*-específico ou seu complemento ou selecionada da seqüência de nucleotídeo de gene de XylT de *Nicotiana* ou um cDNA de XylT de *Nicotiana* ou o complemento da mesma, a referida seqüência de nucleotídeo, de preferência, obtenível da referida espécie ou cultivar de *Nicotiana*,
- 20 em que os referidos 19 de 20 nucleotídeos consecutivos compreendem pelo menos um nucleotídeo de XylT espécie de *Nicotiana*-específico, na orientação sentido;

- c) uma região de DNA compreendendo um sinal de poliadenilação e de término de transcrição funcional em plantas.

30 15. Método de acordo com a reivindicação 12, em que o referido RNA de filamento duplo é fornecido às referidas células de planta ou planta através de integração de um gene quimérico no genoma das referidas células

las de planta para gerar células de planta transgênica e, opcionalmente, regeneração das referidas células de planta para obter plantas transgênicas, o referido gene quimérico compreendendo os seguintes fragmentos de DNA operavelmente ligados:

- 5 a) um promotor expressível em planta;
- b) uma região de DNA passível de transcrição compreendendo:
 - i) uma primeira região de DNA compreendendo pelo menos 19 de 20 nucleotídeos consecutivos selecionados de uma seqüência de nucleotídeo que codifica uma proteína de XylIT de *Nicotiana* ou o complemento da mesma, a referida seqüência de nucleotídeo, de preferência, obtenível da referida espécie ou cultivar de *Nicotiana*, em que os referidos 19 de 20 nucleotídeos consecutivos codificam pelo menos um aminoácido de XylIT espécie ou cultivar de *Nicotiana*-específico ou selecionada da seqüência de nucleotídeo de um gene de XylIT de *Nicotiana* ou um cDNA de XylIT de *Nicotiana* ou o complemento da mesma, a referida seqüência de nucleotídeo, de preferência, obtenível da referida espécie ou cultivar de *Nicotiana*, em que os referidos 19 de 20 nucleotídeos consecutivos compreendem pelo menos um nucleotídeo de XylIT espécie de *Nicotiana*-específico, na orientação anti-sentido;
 - 20 ii) uma segunda região de DNA compreendendo pelo menos 19 de 20 nucleotídeos consecutivos selecionados de uma seqüência de nucleotídeo que codifica uma proteína de XylIT de *Nicotiana* ou o complemento da mesma, a referida seqüência de nucleotídeo, de preferência, obtenível da referida espécie ou cultivar de *Nicotiana*, em que os referidos 19 de 20 nucleotídeos consecutivos codificam pelo menos um aminoácido de XylIT espécie ou cultivar de *Nicotiana*-específico ou selecionada da seqüência de nucleotídeo de um gene de XylIT de *Nicotiana* ou um cDNA de XylIT de *Nicotiana* ou o complemento da mesma, a referida seqüência de nucleotídeo, de preferência, obtenível da referida espécie ou cultivar de *Nicotiana*, em que os referidos 19 de 20 nucleotídeos consecutivos compreendem pelo menos um nucleotídeo de XylIT espécie de *Nicotiana*-específico, na orientação sentido, pelo que uma molécula de RNA produzida através de transcrição da

referida região de DNA passível de transcrição é capaz de formar uma região de RNA de filamento duplo através de emperalhamento de base pelo menos entre uma região de RNA correspondendo à referida primeira região de DNA e uma região de RNA correspondendo à referida segunda região de RNA; e

c) uma região de DNA compreendendo um sinal de poliadenilação e de término de transcrição funcional em plantas.

16. Método de acordo com qualquer uma das reivindicações 12 a 15, em que o referido aminoácido de XylT espécie de *Nicotiana*-específico é um aminoácido de XylT *Nicotiana benthamiana*-específico e a referida espécie de *Nicotiana* é, de preferência, *Nicotiana benthamiana*.

17. Método de acordo com qualquer uma das reivindicações 12 a 16, em que a referida seqüência de nucleotídeo que codifica uma proteína de XylT de *Nicotiana* compreende uma seqüência de nucleotídeo que codifica a seqüência de aminoácido de SEQ ID NO.: 12 ou SEQ ID NO.: 14.

18. Método de acordo com qualquer uma das reivindicações 12 a 17, em que o referido nucleotídeo de XylT espécie de *Nicotiana*-específico é um nucleotídeo de XylT de *Nicotiana benthamiana*-específico e a referida espécie de *Nicotiana* é, de preferência, *Nicotiana benthamiana*.

19. Método de acordo com qualquer uma das reivindicações 12 a 18, em que a referida seqüência de nucleotídeo do referido gene de XylT de *Nicotiana* compreende a seqüência de nucleotídeo de SEQ ID NO.: 11, SEQ ID NO.: 13 ou SEQ ID No. 21.

20. Método de acordo com qualquer uma das reivindicações 12 a 15 em que o referido aminoácido de XylT cultivada de *Nicotiana*-específico é um aminoácido de XylT *Nicotiana tabacum* cv. Petite Havana SR1-específico e o referido cultivar de *Nicotiana* é, de preferência, *Nicotiana tabacum* cv. Petite Havana SR1.

21. Método de acordo com qualquer uma das reivindicações 12 a 15 ou 20 em que a referida seqüência de nucleotídeo que codifica a referida proteína de XylT de *Nicotiana* compreende uma seqüência de nucleotídeo que codifica a seqüência de aminoácido de SEQ ID NO.: 4, SEQ ID NO.:

6, SEQ ID NO.: 8 ou SEQ ID NO.: 10.

22. Método de acordo com qualquer uma das reivindicações 12 a 15, 20 ou 21 em que o referido nucleotídeo de XylT cultivar de *Nicotiana*-específico é um nucleotídeo de *Nicotiana Nicotiana tabacum* cv. Petite Havana SR1-específico e o referido cultivar de *Nicotiana* é, de preferência, *Nicotiana tabacum* cv. Petite Havana SR1.

23. Método de acordo com qualquer uma das reivindicações 12 a 15 ou 20 a 22 em que a referida seqüência de nucleotídeo do referido gene de XylT de *Nicotiana* ou o referido cDNA de XylT de *Nicotiana* compreende a seqüência de nucleotídeo de SEQ ID NO.: 3, SEQ ID NO.: 5, SEQ ID NO.: 8, SEQ ID NO.: 10 ou SEQ ID NO.: 17.

24. Método para produzir uma célula de planta ou planta 1 *Nicotiana* tendo um baixo nível de resíduos de beta-1,2-xilose sobre N-glicanos ligadas à proteína compreendendo as etapas de:

15 a) identificação de um fragmento de uma proteína de XylT que codifica uma seqüência de DNA obtível de uma primeira espécie ou cultivar de *Nicotiana* usando um meio selecionado do seguinte grupo:

i) um fragmento de DNA compreendendo uma seqüência de nucleotídeo que codifica a seqüência de aminoácido de SEQ ID NO.: 4, SEQ ID NO.: 6, SEQ ID NO.: 8, SEQ ID NO.: 10, SEQ ID NO.: 12 ou SEQ ID NO.: 14, para uso como uma sonda;

20 ii) um fragmento de DNA compreendendo a seqüência de nucleotídeo de qualquer uma de SEQ ID NO.: 3, SEQ ID NO.: 5, SEQ ID NO.: 7, SEQ ID NO.: 9, SEQ ID NO.: 11, SEQ ID NO.: 13, SEQ ID NO.: 17 ou SEQ ID NO.: 21, para uso como uma sonda;

25 iii) um fragmento de DNA ou oligonucleotídeo compreendendo uma seqüência de nucleotídeo consistindo em entre 20 a 1513 nucleotídeos consecutivos selecionados de uma seqüência de nucleotídeo que codifica a seqüência de aminoácido de SEQ ID NO.: 4 ou SEQ ID NO.: 6, para uso como uma sonda;

30 iv) um fragmento de DNA ou oligonucleotídeo compreendendo uma seqüência de nucleotídeo consistindo em entre 20 a 3574 nucleotí-

deos consecutivos selecionados de uma seqüência de nucleotídeo que codifica a seqüência de aminoácido de SEQ ID NO.: 8, SEQ ID NO.: 10, SEQ ID NO.: 12 ou SEQ ID NO.: 14 para uso como uma sonda;

5 v) um fragmento de DNA ou oligonucleotídeo compreendendo uma seqüência de nucleotídeo consistindo em entre 20 a 3574 nucleotídeos consecutivos selecionados de uma seqüência de nucleotídeo de qualquer uma de SEQ ID NO.: 3, SEQ ID NO.: 5, SEQ ID NO.: 7, SEQ ID NO.: 9, SEQ ID NO.: 11, SEQ ID NO.: 13, SEQ ID NO.: 17 ou SEQ ID NO.: 21 para uso como uma sonda;

10 vi) uma seqüência de oligonucleotídeo tendo uma seqüência de nucleotídeo compreendendo entre 20 a 200 nucleotídeos consecutivos selecionados de uma seqüência de nucleotídeo que codifica a seqüência de aminoácido de SEQ ID NO.: 4 ou SEQ ID NO.: 6, para uso como um iniciador em uma reação de PCR;

15 vii) uma seqüência de oligonucleotídeo tendo uma seqüência de nucleotídeo compreendendo entre 20 a 200 nucleotídeos consecutivos selecionados de uma seqüência de nucleotídeo que codifica a seqüência de aminoácido de SEQ ID NO.: 8, SEQ ID NO.: 10, SEQ ID NO.: 12 ou SEQ ID NO.: 14 , para uso como um iniciador em uma reação de PCR;

20 viii) uma seqüência de oligonucleotídeo tendo uma seqüência de nucleotídeo compreendendo entre 20 a 200 nucleotídeos consecutivos selecionada da seqüência de nucleotídeo de qualquer uma de SEQ ID NO.: 3, SEQ ID NO.: 5, SEQ ID NO.: 7, SEQ ID NO.: 9, SEQ ID NO.: 11, SEQ ID NO.: 13, SEQ ID NO.: 17 ou SEQ ID NO.: 21, para uso como um
25 iniciador em uma reação de PCR; ou

ix) um oligonucleotídeo tendo a seqüência de nucleotídeo de qualquer uma de SEQ ID NO.: 1, SEQ ID NO.: 2, SEQ ID NO.: 15 ou SEQ ID NO.: 16, SEQ ID NO.: 19 ou SEQ ID No.20 para uso como um iniciador em uma reação de PCR;

30 b) fornecimento de uma ou mais moléculas de RNA de filamento duplo às células de planta ou plantas das referidas primeira ou segunda espécie ou cultivar de *Nicotiana*, em que as referidas moléculas de RNA de

filamento duplo compreendem duas fitas de RNA, uma fita de RNA consistindo essencialmente em uma seqüência de nucleotídeo de RNA de 20 a 21 nucleotídeos consecutivos selecionada de uma seqüência de nucleotídeo da referida proteína de XylT que codifica um fragmento de DNA ou o complemento da mesma, em que os referidos 20 a 21 nucleotídeos consecutivos codificam pelo menos um aminoácido de XylT espécie ou cultivar de *Nicotiana*-específico, respectivamente ou em que os referidos 20 a 21 nucleotídeos consecutivos compreendem pelo menos um nucleotídeo de *Nicotiana* espécie ou cultivar de *Nicotiana*-específico, respectivamente; e

- 10 c) identificação de uma célula de planta ou planta *Nicotiana* compreendendo a referida molécula ou moléculas de RNA de filamento duplo a qual tem um menor nível de resíduos de beta-1,2-xilose sobre N-glicanos ligadas à proteína do que a mesma célula de planta ou planta de *Nicotiana*, a qual não compreende a referida molécula ou moléculas de RNA
- 15 de filamento duplo.

25. Método de acordo com a reivindicação 24, em que o fornecimento da referida molécula ou moléculas de RNA de filamento duplo é obtida através de fornecimento, às referidas células de planta ou plantas, de uma molécula ou moléculas de RNA de filamento duplo compreendendo
- 20 uma primeira seqüência de nucleotídeo de pelo menos 19 de 20 nucleotídeos consecutivos selecionada da seqüência de nucleotídeo de a referida proteína de XylT que codifica um fragmento de DNA ou o complemento da mesma, em que os referidos pelo menos 19 de 20 nucleotídeos consecutivos codificam pelo menos um aminoácido de XylT espécie ou cultivar de *Nicotiana*-específico ou em que os referidos pelo menos 19 de 20 nucleotídeos consecutivos compreendem pelo menos um nucleotídeo de XylT espécie ou cultivar de *Nicotiana*-específico e uma segunda seqüência de nucleotídeo a qual é o complemento da referida primeira seqüência de nucleotídeo.

26. Método de acordo com a reivindicação 24, em que as referidas moléculas de RNA de filamento duplo são fornecidas às referidas células de planta ou plantas através de integração de um DNA quimérico no genoma das referidas células de planta para gerar células de planta transgêni-
- 30

ca e, opcionalmente, regeneração das referidas células de planta para obter plantas transgênicas, o referido DNA quimérico compreendendo os seguintes fragmentos de DNA operavelmente ligados:

- a) um promotor expressível em planta;
- 5 b) uma região de DNA passível de transcrição compreendendo:
 - i) uma primeira região de DNA sentido compreendendo uma seqüência de nucleotídeo de pelo menos 19 de 20 nucleotídeos consecutivos selecionada da seqüência de nucleotídeo da referida proteína de XylIT que codifica um fragmento de DNA ou o complemento da mesma, em que os referidos pelo menos 19 de 20 nucleotídeos consecutivos codificam pelo menos um aminoácido de XylIT espécie ou cultivar de *Nicotiana*-específico, respectivamente ou em que os referidos pelo menos 19 de 20 nucleotídeos consecutivos compreendem pelo menos um nucleotídeo de XylIT espécie ou cultivar de *Nicotiana*-específico, respectivamente;
 - 10 ii) uma segunda região de anti-sentido de DNA compreendendo uma seqüência de nucleotídeo de pelo menos 19 nucleotídeos consecutivos a qual têm pelo menos 95% de identidade de seqüência ao complemento da referida primeira região de DNA;

em que uma molécula de RNA transcrita a partir da referida região passível de transcrição é capaz de formar uma região de RNA de filamento duplo pelo menos entre uma região de RNA transcrita a partir da referida primeira região de sentido de DNA e uma região de RNA transcrita a partir da referida segunda região de anti-sentido de DNA; e
 - c) uma região de DNA compreendendo um sinal de poliadenila-
 - 20 ção e de término de transcrição funcional em plantas.

27. Método de acordo com qualquer uma das reivindicações 1 a 26, ainda compreendendo a etapa de cruzamento da referida planta *Nicotiana* tendo um baixo nível de resíduos de beta-1,2-xilose sobre N-glicanos ligadas à proteína com outra planta *Nicotiana* para obter plantas de prole *Nicotiana* tendo um baixo nível de resíduos de beta-1,2-xilose sobre N-glicanos ligadas à proteína.

28. Método para identificar um fragmento de DNA de XylIT de

Nicotiana compreendendo as etapas de:

a) fornecimento de DNA genômico ou cDNA obtenível de uma espécie ou cultivar de *Nicotiana* ;

b) seleção de um meio do seguinte grupo:

5 i) um fragmento de DNA compreendendo uma seqüência de nucleotídeo que codifica a seqüência de aminoácido de SEQ ID NO.: 4, SEQ ID NO.: 6, SEQ ID NO.: 8, SEQ ID NO.: 10, SEQ ID NO.: 12 ou SEQ ID NO.: 14, para uso como uma sonda;

10 ii) um fragmento de DNA compreendendo a seqüência de nucleotídeo de qualquer uma de SEQ ID NO.: 3, SEQ ID NO.: 5, SEQ ID NO.: 7, SEQ ID NO.: 9, SEQ ID NO.: 11, SEQ ID NO.: 13, SEQ ID NO.: 17 ou SEQ ID NO.: 21, para uso como uma sonda;

15 iii) um fragmento de DNA ou oligonucleotídeo compreendendo uma seqüência de nucleotídeo consistindo em entre 20 a 1513 nucleotídeos consecutivos selecionados de uma seqüência de nucleotídeo que codifica a seqüência de aminoácido de SEQ ID NO.: 4 ou SEQ ID NO.: 6, para uso como uma sonda;

20 iv) um fragmento de DNA ou oligonucleotídeo compreendendo uma seqüência de nucleotídeo consistindo em entre 20 a 3574 nucleotídeos consecutivos selecionados de uma seqüência de nucleotídeo que codifica a seqüência de aminoácido de SEQ ID NO.: 8, SEQ ID NO.: 10, SEQ ID NO.: 12 ou SEQ ID NO.: 14 para uso como uma sonda;

25 v) um fragmento de DNA ou oligonucleotídeo compreendendo uma seqüência de nucleotídeo consistindo em entre 20 a 3574 nucleotídeos consecutivos selecionados de uma seqüência de nucleotídeo de qualquer uma de SEQ ID NO.: 3, SEQ ID NO.: 5, SEQ ID NO.: 7, SEQ ID NO.: 9, SEQ ID NO.: 11, SEQ ID NO.: 13, SEQ ID NO.: 17 ou SEQ ID NO.: 21 para uso como uma sonda;

30 vi) uma seqüência de oligonucleotídeo tendo uma seqüência de nucleotídeo compreendendo entre 20 a 200 nucleotídeos consecutivos selecionados de uma seqüência de nucleotídeo que codifica a seqüência de aminoácido de SEQ ID NO.: 4 ou SEQ ID NO.: 6, para uso como um inicia-

dor em uma reação de PCR;

vii) uma seqüência de oligonucleotídeo tendo uma seqüência de nucleotídeo compreendendo entre 20 a 200 nucleotídeos consecutivos selecionados de uma seqüência de nucleotídeo que codifica a seqüência de aminoácido de SEQ ID NO.: 8, SEQ ID NO.: 10, SEQ ID NO.: 12 ou SEQ ID NO.: 14 , para uso como um iniciador em uma reação de PCR;

viii) uma seqüência de oligonucleotídeo tendo uma seqüência de nucleotídeo compreendendo entre 20 a 200 nucleotídeos consecutivos selecionada da seqüência de nucleotídeo de qualquer uma de SEQ ID NO.: 3, SEQ ID NO.: 5, SEQ ID NO.: 7, SEQ ID NO.: 9, SEQ ID NO.: 11, SEQ ID NO.: 13, SEQ ID NO.: 17 ou SEQ ID NO.: 21, para uso como um iniciador em uma reação de PCR; ou

ix) um oligonucleotídeo tendo a seqüência de nucleotídeo de qualquer uma de SEQ ID NO.: 1, SEQ ID NO.: 2, SEQ ID NO.: 15 ou SEQ ID NO.: 16, SEQ ID NO.: 19 ou SEQ ID No.20 para uso como um iniciador em uma reação de PCR;

c) identificação de um fragmento de DNA de XylT da referida espécie ou cultivar de *Nicotiana* realizando uma PCR usando o referido DNA genômico ou o referido cDNA e os referidos iniciadores ou realizando hibridização usando o referido DNA genômico ou o referido cDNA e as referidas sondas.

29. Método para isolar um fragmento de DNA de XylT de *Nicotiana* compreendendo as etapas de:

a) identificação do referido fragmento de DNA de XylT de *Nicotiana* de acordo com o método como definido na reivindicação 28; e

b) isolamento do referido fragmento de DNA de XylT de *Nicotiana*.

30. Método para identificar um alelo de XylT de *Nicotiana* correlacionado a um baixo nível de resíduos de beta-1,2-xilose sobre N-glicanos ligadas à proteína compreendendo as etapas de:

(a) fornecimento de uma população, opcionalmente uma população com mutação, de diferentes linhagens de planta de uma espécie ou cul-

tivar de *Nicotiana*;

(b) identificação, em cada linhagem de planta, da referida população de um alelo de XylT de *Nicotiana* de acordo com o método como definido na reivindicação 28;

5 (c) análise do nível de resíduos de beta-1,2-xilose sobre N-glicanos ligadas à proteína de cada linhagem de planta da referida população e identificação daquelas linhagens de planta tendo um menor nível de resíduos de beta-1,2-xilose sobre N-glicanos ligadas à proteína do que outras linhagens de planta;

10 (d) correlação do baixo nível de resíduos de beta-1,2-xilose sobre N-glicanos ligadas à proteína na linhagem de planta com relação à presença de um alelo de XylT de *Nicotiana*.

31. Método para obter uma célula de planta ou planta *Nicotiana* com um baixo nível de resíduos de beta-1,2-xilose sobre N-glicanos ligadas à proteína compreendendo as etapas de:

a) identificação de um alelo de XylT de *Nicotiana* correlacionado a um baixo nível de resíduos de beta-1,2-xilose sobre N-glicanos ligadas à proteína de acordo com o método como definido na reivindicação 30;

20 b) introdução do referido alelo de XylT de *Nicotiana* em uma linhagem de planta *Nicotiana* de escolha.

32. Fragmento de DNA isolado que codifica uma proteína compreendendo a sequência de aminoácido de SEQ ID NO.: 12 ou SEQ ID NO.: 14 ou qualquer parte da mesma que codifica pelo menos um aminoácido de XylT *Nicotiana benthamiana*-específico.

25 33. Fragmento de DNA específico de acordo com a reivindicação 32, compreendendo a sequência de nucleotídeo de SEQ ID NO.: 11, SEQ ID NO.: 13 ou SEQ ID NO.: 21 ou qualquer parte da mesma compreendendo pelo menos um nucleotídeo de XylT *Nicotiana benthamiana*-específico.

30 34. Fragmento de DNA isolado que codifica uma proteína compreendendo a sequência de aminoácido de SEQ ID NO.: 4 ou SEQ ID NO.: 6, SEQ ID NO.: 8, SEQ ID NO.: 10 ou qualquer parte da mesma que codificam pelo menos um aminoácido de XylT *Nicotiana tabacum* cv. Petite Hava-

na SR1-específico.

35. Fragmento de DNA isolado de acordo com a reivindicação 34, compreendendo a seqüência de nucleotídeo de SEQ ID NO.: 3 ou SEQ ID NO.: 5, SEQ ID NO.: 7, SEQ ID NO.: 9 ou SEQ ID NO.: 17 ou qualquer
5 parte da mesma compreendendo pelo menos um nucleotídeo de XylT *Nicotiana tabacum* cv. Petite Havana SR1-específico.

36. Fragmento de DNA isolado obténível através do método como definido na reivindicação 28 que codifica pelo menos um aminoácido de XylT espécie de *Nicotiana*- ou cultivar de *Nicotiana*-específico.

10 37. Fragmento de DNA isolado de acordo com a reivindicação 36, compreendendo pelo menos um nucleotídeo de XylT espécie de *Nicotiana*- ou cultivar de *Nicotiana*-específico.

38. Gene quimérico compreendendo os seguintes fragmentos de DNA operavelmente ligados:

15 a) um promotor expressível em planta;
b) uma região de DNA passível de transcrição compreendendo:
i) uma primeira região de DNA compreendendo pelo menos
19 de 20 nucleotídeos consecutivos selecionados de uma seqüência de nucleotídeo que codifica uma proteína de XylT de *Nicotiana* ou o complemento
20 da mesma, em que os referidos 19 de 20 nucleotídeos consecutivos codificam pelo menos um aminoácido de XylT espécie ou cultivar de *Nicotiana*-específico ou selecionada da seqüência de nucleotídeo de um gene de XylT de *Nicotiana* ou um cDNA de XylT de *Nicotiana* ou o complemento da mesma, em que os referidos 19 de 20 nucleotídeos consecutivos compreendem
25 pelo menos um nucleotídeo de XylT espécie de *Nicotiana*-específico, na orientação de sentido;

ii) uma segunda região de DNA compreendendo pelo menos
19 de 20 nucleotídeos consecutivos selecionados de uma seqüência de nucleotídeo que codifica uma proteína de XylT de *Nicotiana* ou o complemento
30 da mesma, em que os referidos 19 de 20 nucleotídeos consecutivos codificam pelo menos um aminoácido de XylT espécie ou cultivar de *Nicotiana*-específico ou selecionada da seqüência de nucleotídeo de um gene de XylT

de *Nicotiana* ou um cDNA de XylT de *Nicotiana* ou o complemento da mesma, em que os referidos 19 de 20 nucleotídeos consecutivos compreendem pelo menos um nucleotídeo de XylT espécie de *Nicotiana*-específico, na orientação de sentido, pelo que uma molécula de RNA produzida através de transcrição da referida região de DNA passível de transcrição é capaz de formação de uma região de RNA de filamento duplo através de emparelhamento de base pelo menos entre uma região de RNA correspondendo à referida primeira região de DNA e uma região de RNA correspondendo à referida segunda região de RNA; e

- 10 c) uma região de DNA compreendendo um sinal de poliadenilação e de término de transcrição funcional em plantas.

39. Gene quimérico compreendendo os seguintes fragmentos de DNA operavelmente ligados:

- 15 a) um promotor expressível em planta;
- b) uma região de DNA compreendendo pelo menos 19 de 20 nucleotídeos consecutivos selecionados de uma sequência de nucleotídeo que codifica uma proteína de XylT de *Nicotiana* ou o complemento da mesma, em que os referidos 19 de 20 nucleotídeos consecutivos codificam pelo menos um aminoácido de XylT espécie ou cultivar de *Nicotiana*-específico ou selecionada da sequência de nucleotídeo de um gene de XylT de *Nicotiana* ou um cDNA de XylT de *Nicotiana* ou o complemento da mesma, em que os referidos 19 de 20 nucleotídeos consecutivos compreendem pelo menos um nucleotídeo de XylT espécie de *Nicotiana*-específico, na orientação de sentido; e

- 25 c) uma região de DNA compreendendo um sinal de poliadenilação e de término de transcrição funcional em plantas.

40. Gene quimérico compreendendo os seguintes fragmentos de DNA operavelmente ligados:

- a) um promotor expressível em planta;
- 30 b) uma região de DNA compreendendo pelo menos 19 de 20 nucleotídeos consecutivos selecionados de uma sequência de nucleotídeo que codifica uma proteína de XylT de *Nicotiana* ou o complemento da mesma,

ma, em que os referidos 19 de 20 nucleotídeos consecutivos codificam pelo menos um aminoácido de XylT espécie ou cultivar de *Nicotiana*-específico ou selecionada da sequência de nucleotídeo de um gene de XylT de *Nicotiana* ou um cDNA de XylT de *Nicotiana* ou o complemento da mesma, em

5 que os referidos 19 de 20 nucleotídeos consecutivos compreendem pelo menos um nucleotídeo de XylT espécie de *Nicotiana*-específico, na orientação de anti-sentido; e

c) uma região de DNA compreendendo um sinal de poliadenilação e de término de transcrição funcional em plantas.

10 41. Gene quimérico de acordo com qualquer uma das reivindicações 38 a 40, em que o referido aminoácido de XylT espécie de *Nicotiana*-específico é um aminoácido de XylT *Nicotiana bentamiana*-específico.

42. Gene quimérico de acordo com qualquer uma das reivindicações 38 a 41, em que a referida sequência de nucleotídeo que codifica uma

15 proteína de XylT de *Nicotiana* compreende uma sequência de nucleotídeo que codifica a sequência de aminoácido de SEQ ID NO.: 12 ou SEQ ID NO.: 14.

43. Gene quimérico de acordo com qualquer uma das reivindicações 38 a 42 em que o referido nucleotídeo de XylT espécie de *Nicotiana*-específico é um nucleotídeo de XylT *Nicotiana bentamiana*-específico.

20

44. Gene quimérico de acordo com qualquer uma das reivindicações 38 a 43 em que a referida sequência de nucleotídeo do referido gene de XylT de *Nicotiana* compreende a sequência de nucleotídeo de SEQ ID NO.: 11, SEQ ID NO.: 13 ou SEQ ID No. 21.

25 45. Gene quimérico de acordo com qualquer uma das reivindicações 38 a 40, em que o referido aminoácido de XylT cultivar de *Nicotiana*-específico é um aminoácido de XylT *Nicotiana tabacum* cv. Petite Havana SR1-específico.

46. Gene quimérico de acordo com qualquer uma das reivindicações 38 a 40 ou 45, em que a referida sequência de nucleotídeo que codifica

30 a referida proteína de XylT de *Nicotiana* compreende uma sequência de nucleotídeo que codifica a sequência de aminoácido de SEQ ID NO.: 4, SEQ

ID NO.: 6, SEQ ID NO.: 8 ou SEQ ID NO.: 10.

47. Gene quimérico de acordo com qualquer uma das reivindicações 38 a 40, 45 ou 46, em que o referido nucleotídeo de XylT cultivar de *Nicotiana*-específico é um nucleotídeo de XylT *Nicotiana tabacum* cv. Petite Havana SR1-específico.

48. Gene quimérico de acordo com qualquer uma das reivindicações 38 a 40 ou 45 a 47, em que a referida seqüência de nucleotídeo do referido gene de XylT de *Nicotiana* ou o referido cDNA de XylT de *Nicotiana* compreende a seqüência de nucleotídeo de SEQ ID NO.: 3, SEQ ID NO.: 5, SEQ ID NO.: 8, SEQ ID NO.: 10 ou SEQ ID NO.: 17.

49. Célula de planta *Nicotiana* compreendendo o gene quimérico como definido em qualquer uma das reivindicações 38 a 48.

50. Planta *Nicotiana* consistindo essencialmente nas células de planta *Nicotiana* como definidas na reivindicação 49.

51. Célula de planta ou planta *Nicotiana* obtida através do método como definido na reivindicação 31.

52. Semente de uma planta *Nicotiana* como definida na reivindicação 50 ou 51.

53. Uso de uma seqüência de nucleotídeo que codifica uma proteína compreendendo a seqüência de aminoácido de SEQ ID NO.: 4, SEQ ID NO.: 6, SEQ ID NO.: 8, SEQ ID NO.: 10, SEQ ID NO.: 12 ou SEQ ID NO.: 14 ou qualquer parte da mesma compreendendo pelo menos 19 de 20 nucleotídeos consecutivos que codificam pelo menos um aminoácido de XylT espécie ou cultivar de *Nicotiana*-específico para diminuir o nível de resíduos de beta-1,2-xilose sobre N-glicanos ligadas à proteína em uma planta *Nicotiana*.

54. Uso de uma seqüência de nucleotídeo compreendendo a seqüência de nucleotídeo de SEQ ID NO.: 3, SEQ ID NO.: 5, SEQ ID NO.: 7, SEQ ID NO.: 9, SEQ ID NO.: 11, SEQ ID NO.: 13, SEQ ID NO.: 17 ou SEQ ID NO.: 21 ou qualquer parte da mesma compreendendo pelo menos 19 de 20 nucleotídeos consecutivos compreendendo pelo menos um nucleotídeo de XylT espécie ou cultivar de *Nicotiana*-específico, para diminuir o nível de

resíduos de beta-1,2-xilose sobre N-glicanos ligadas à proteína em uma planta *Nicotiana*.

55. Uso de uma seqüência de nucleotídeo que codifica uma proteína compreendendo a seqüência de aminoácido de SEQ ID NO.: 4, SEQ ID NO.: 6, SEQ ID NO.: 8, SEQ ID NO.: 10, SEQ ID NO.: 12 ou SEQ ID NO.: 14 ou qualquer parte da mesma que codifica pelo menos um aminoácido de XylT espécie ou cultivar de *Nicotiana*-específico, para identificar um gene de XylT ou cDNA de XylT em uma espécie ou cultivar de *Nicotiana*.

56. Uso de uma seqüência de nucleotídeo compreendendo a seqüência da SEQ ID NO.: 3, SEQ ID NO.: 5, SEQ ID NO.: 7, SEQ ID NO.: 9, SEQ ID NO.: 11, SEQ ID NO.: 13, SEQ ID NO.: 17 ou SEQ ID NO.: 21, ou qualquer parte das mesmas, compreendendo pelo menos um nucleotídeo XylT espécie ou cultivar-específico de *Nicotiana*, pra identificar um gene XylT ou um CDNA XylT em uma espécie ou cultivar *Nicotiana*.

57. Uso de uma seqüência de nucleotídeo que codifica uma proteína compreendendo a seqüência de aminoácido de SEQ ID NO.: 4, SEQ ID NO.: 6, SEQ ID NO.: 8, SEQ ID NO.: 10, SEQ ID NO.: 12 ou SEQ ID NO.: 14 ou qualquer parte da mesma que codifica pelo menos um aminoácido de XylT espécie ou cultivar de *Nicotiana*-específico, para identificar um alelo de um gene *XylT* correlacionado a um baixo nível de resíduos de beta-1,2-xilose sobre N-glicanos ligadas à proteína em uma espécie ou cultivar de *Nicotiana*.

58. Uso de uma seqüência de nucleotídeo compreendendo a seqüência de nucleotídeo de SEQ ID NO.: 3, SEQ ID NO.: 5, SEQ ID NO.: 7, SEQ ID NO.: 9, SEQ ID NO.: 11, SEQ ID NO.: 13, SEQ ID NO.: 17 ou SEQ ID NO.: 21 ou qualquer parte da mesma compreendendo pelo menos um nucleotídeo de XylT espécie ou cultivar de *Nicotiana*-específico, para identificar um alelo de um gene *XylT* correlacionado a um baixo nível de resíduos de beta-1,2-xilose sobre N-glicanos ligadas à proteína em uma espécie ou cultivar de *Nicotiana*.

59. Uso de uma seqüência de nucleotídeo que codifica uma proteína compreendendo a seqüência de aminoácido de SEQ ID NO.: 4, SEQ

ID NO.: 6, SEQ ID NO.: 8, SEQ ID NO.: 10, SEQ ID NO.: 12 ou SEQ ID NO.: 14 ou qualquer parte da mesma que codificam pelo menos um aminoácido de XylT espécie ou cultivar de *Nicotiana*-específico, para introduzir um alelo de um gene *XylT* correlacionado a um baixo nível de resíduos de beta-1,2-xilose sobre N-glicanos ligadas à proteína em uma espécie ou cultivar de *Nicotiana*.

60. Uso de uma sequência de nucleotídeo compreendendo a sequência de nucleotídeo de SEQ ID NO.: 3, SEQ ID NO.: 5, SEQ ID NO.: 7, SEQ ID NO.: 9, SEQ ID NO.: 11, SEQ ID NO.: 13, SEQ ID NO.: 17 ou SEQ ID NO.: 21 ou qualquer parte da mesma compreendendo pelo menos um nucleotídeo de XylT espécie ou cultivar de *Nicotiana*-específico, para introduzir um alelo de um gene *XylT* correlacionado a um baixo nível de resíduos de beta-1,2-xilose sobre N-glicanos ligadas à proteína em uma espécie ou cultivar de *Nicotiana*.

Fig. 1

AJ627182	1	agtcagagagagaagatgaacaaagaaaagctgaaatttctgtttctctctctcgtctcaactcaatcactctcta
xy1Tc2Nt	1	-----cac-----
xy1Tc7Nt	1	-----cacc-----
AJ627182	81	tctctacttctctccactctgatcacttcggtcacaaatccccccaaaaccactttcctaataaccacaaaccactatt
xy1Tc2Nt	42	
xy1Tc7Nt	43	c.....g.....
AJ627182	161	ccctgtcggaaaaccaccatgataatttccactcttctgtcacttcccaatataccaagccttggccaattttgcccctcc
xy1Tc2Nt	122	
xy1Tc7Nt	102	..t.....g...c.....aa.....t.....t.....
AJ627182	241	tacctccctggtctcagaatcctaattgtttctttgagatcgtgcgagggttacttcggtaatgggtttactctcaaaagt
xy1Tc2Nt	202	
xy1Tc7Nt	182	t.....a.c.....gt.g.....
AJ627182	321	tgatcttctcaaaacttcgccggagcttcaccagaaaattcggcgaaaacaccgtatccggcgaggatggttttaggt
xy1Tc2Nt	282	
xy1Tc7Nt	262	..c.....t.....g.....c.....
AJ627182	401	gttttttccagtgcagactttgcagagttcgtatttcgcgaggagggtgctatacgaatgaatccgcgacgagattttgatgtct
xy1Tc2Nt	362	
xy1Tc7Nt	342	c.....c.....c.a.....
AJ627182	481	cgtggaggcgagaaaattggagtcggttatttggtaggagtgaagatgatgagctgcccggtgttcaaaaaatggagcttttca
xy1Tc2Nt	442	
xy1Tc7Nt	422	t.....
AJ627182	561	gattaaagtactgataaaactgaaaattgggaaaaaattagtggtgatgaaaaaatcttgaataaatacttaccgggaagggtg
xy1Tc2Nt	522	
xy1Tc7Nt	502	t.....
AJ627182	641	caatttcaaggcacactatgcgtgaattaattgactctattcagttagttggcgccgatgaatttcactgtttctgagtg
xy1Tc2Nt	602	
xy1Tc7Nt	582	g.....g.....t.....
AJ627182	721	attgaggagccgtcactttttgattacacgatttgagtatgcaaaccttttccacacagttaccgattgggtatagtgcata
xy1Tc2Nt	682	
xy1Tc7Nt	662	

Fig. 1 (Continuação)

AJ627182	801	cgtagcatccagggtactggcttgcacagtcgcccacatttggtttttagatggccattgtgagacacaattggagg
xy1Tc2Nt	762	
xy1Tc7Nt	742	
AJ627182	881	aaacatggaaagcactcttttcaagcctcacttatgctaagaacttttagtggcccagtttgtttccgtcacgccgttctc
xy1Tc2Nt	842	
xy1Tc7Nt	822	t.....t.....cc...
AJ627182	961	tcgcctttgggatatgaaactgcacctgtttaagggactgacagaaactatagattgtaatggagcttctgcccatgattt
xy1Tc2Nt	922	
xy1Tc7Nt	902	t.....
AJ627182	1041	gtggcaaaaatcctgatgataagagaactgcacggttgtctgtgagtttggggagatgatcagggcagccttttggtatttcctg
xy1Tc2Nt	1002	C.....
xy1Tc7Nt	982	a.....C.....t.....
AJ627182	1121	tggatagacagaacatcccaaggacagtcacaggccctaataatgtcctctttgttagacgtgaggattatttagctcaccca
xy1Tc2Nt	1082	
xy1Tc7Nt	1062	
AJ627182	1201	cgtagtggtggaagggtacagtccttaggcttagcaatgaagagcaagtatttgattccataaaagagctgggccttgaacca
xy1Tc2Nt	1162	
xy1Tc7Nt	1142	
AJ627182	1281	ctcggagtgcataaataatgtaattaacggattgtttgcccacatgtcccatgaaagagcaagtttcgagcaatcccaagatg
xy1Tc2Nt	1242	
xy1Tc7Nt	1222	
AJ627182	1361	cttctgtcatagtgtgtgctcatggagcaggtctaactcacatagtttctgcagcaccacaaaagctgtaatactagaaaatt
xy1Tc2Nt	1322	.a.....t.....
xy1Tc7Nt	1302	
AJ627182	1441	ataagcagcgaatataggcgccccattttgtctctgattgcacaatggaaaggattggagtagccatcccatatatttggga
xy1Tc2Nt	1402	a.....
xy1Tc7Nt	1382	
AJ627182	1521	ggggctcttatgcgggacccctccagttgtgatcgacaggctcagcagcatttttgaggagctcttgggtgctaaagtcgcgctcga
xy1Tc2Nt	1482	
xy1Tc7Nt	1462	a-----

Fig. 1 (Continuação)

AJ627182 xylTc2Nt xylTc7Nt	1601	cagtttgaatagttcggcttttctctaaaagacggggaaggatagaggaattcggggttcttggaacttgagcctgggaa ----- ----- -----
AJ627182 xylTc2Nt xylTc7Nt	1681	ttttgataaatatgtttcacacgcagttctgtagtcaatggttgcaatctaggtcctcaatctggtgttgataagcttgg ----- ----- -----
AJ627182 xylTc2Nt xylTc7Nt	1761	caatttcagcagctactaatttattagcccgctctgactcggttatggactaccagagcaatcatatcaaatggaagca ----- ----- -----
AJ627182 xylTc2Nt xylTc7Nt	1841	tggaaatcctgattgtggaatggtgagctcattgaagagcatattctttatggtgttgaagattacaattcacaaattaaca ----- ----- -----
AJ627182 xylTc2Nt xylTc7Nt	1921	cgtgtatgtgaaagattaggttggtacacttacttacaattcattgtccaattgttttttcattattctcattaatgatcat ----- ----- -----
AJ627182 xylTc2Nt xylTc7Nt	2001	aggataagaacatgagaaaaccatccatgttctgtgtgttttcccatcaatccggccacctcttccctccttatgtag ----- ----- -----
AJ627182 xylTc2Nt xylTc7Nt	2081 1509	agatgatttcaacagagtttgttttgtagttgtaacacttgcaactgccactccagttacagttttgcattcgacacattcatccc ----- ----- -----
AJ627182 xylTc2Nt xylTc7Nt	2161	atcagatgtcaagtttaaaggcataagacatttgacatattgaaagaagcagattaacacgaacgtcagtatgatgcttca ----- ----- -----
AJ627182 xylTc2Nt xylTc7Nt	2241	gtgaagatatggttgttaacttgaaccacaaaagaaatgagactttgacaaaaaaaaaaaaaaaaaaaaaa ----- ----- -----

Fig. 2

AJ627182	1	nmkklkflvslfalnsitlylyfsshshdfrhksqpqnfhfntqnhyslshhndfnhssvtsqytkpwpilpsylpwsq
xy1Tc2Nt	1	-----
xy1Tc7Nt	1	-----p.....r.....r.h.....i.....s
AJ627182	81	npnvlsrcegyfgngftlkvdllktspelhqkfgentvsgdggwfrcffsetlqssiceggairmpdeilmrsggekl
xy1Tc2Nt	72	vw.....f.r.....
xy1Tc7Nt	66	f.r.....
AJ627182	161	esvigrseddelpvfkngafqikvtdklkigklvdekilnkylpegaisrhtmrrelidsiqlvgadehfcsewieeps1
xy1Tc2Nt	152	f.....d.....
xy1Tc7Nt	146	f.....d.....
AJ627182	241	litrfeyanlfhtvtdwysayvasrvtglpsrphlvfvdghcetqleetwkalfssltiaknfsdpvcfrhav1splgye
xy1Tc2Nt	232	
xy1Tc7Nt	226	a.....
AJ627182	321	talfkgtetidcngasahdlwqnpddkrtarlsefgemiraaafgfpvdrqniprtvtgpnvlfvrredylahprhggkv
xy1Tc2Nt	312	s.....k.....
xy1Tc7Nt	306	s.....k.....
AJ627182	401	qsrlsneeqvfdsikswalnhsecklnvinglfahmsmkeqvraiqdasvivgahgag1thivsaapkavileiisseyr
xy1Tc2Nt	392	
xy1Tc7Nt	386	
AJ627182	481	rphfaliaqwkgleyhpiylegsyadppvvidrlsslrlslgc
xy1Tc2Nt	472	h.....m-----
xy1Tc7Nt	466	e-----

Fig. 3

AJ627183	1	agtcagagagaagaagatgaacaagaaaaagctgaaaaattcttgtttctctctctcgtctcgaactcaatcactctctcta
xylTg1Nt	1	-----cac-----
xylTg3Nt	1	-----cac-----
xylTg14Nb	1	-----cac-----
xylTg19Nb	1	-----cac-----
AJ627183	81	tctctacttctcttccaccctgatcacttccgcccacaaatcccgccaaaaaccacttt-----t
xylTg1Nt	42	t.....tg.....a.....
xylTg3Nt	42	t.....c.....cctaatacccaaaaccactat.
xylTg14Nb	42	ct..t.g.....c..g.....
xylTg19Nb	42	-----c.....
AJ627183	140	ccttgtcgaaaaaccgccatcataatttccactcttccaatcacttctccaatattccaagccttggcctattttgcccctcc
xylTg1Nt	122	..C.....a...g.....tg.....c.....a.....
xylTg3Nt	101	a.....
xylTg14Nb	101	..c.....a.....c.....g.....
xylTg19Nb	92	a.....
AJ627183	220	tacctcccttggtctcaaaaaccctaattgttggagatcgtgcgagggttacttcggtaatgggttactctctcaagt
xylTg1Nt	202	C.....g..t.....tc..t.....
xylTg3Nt	181	
xylTg14Nb	181	c.....a.....t.....
xylTg19Nb	172	c.....
AJ627183	300	tgaccttctcaaaacttcgcgaggtttccaccggaatttcggcgaaaaacaccgtctccggcgacggcgaggttttaggt
xylTg1Nt	282	...t.....c.....a.....a.....
xylTg3Nt	261	
xylTg14Nb	261	..t.....c.....t...a.....
xylTg19Nb	252	t.....t.....
AJ627183	380	gttttttcagtgagactttgcagagttcgcgatctgcgaggggagcgcaatacgaatgaatccggacgagattttgcgtct
xylTg1Nt	362	t.....t..t.....
xylTg3Nt	341	
xylTg14Nb	341	..c.....a.....
xylTg19Nb	332	t.....

Fig. 3 (Continuação)

AJ627183	460	cgtagggtgagaaattggagtcggttattggtaggagtgaagatgatgagctgccgtgttcaaaaatggagcttttca
xy1Tg1Nt	442	C.....
xy1Tg3Nt	421	
xy1Tg14Nb	421	g.....C.....
xy1Tg19Nb	412	a.....a.....C.....
AJ627183	540	gattaaagttactgataaactgaaaaattgggaaaaattagtggtgataaaattcttgaataaatacttaccggaagggtg
xy1Tg1Nt	522	a.....
xy1Tg3Nt	501	a.....
xy1Tg14Nb	501	t.....C.....
xy1Tg19Nb	492	a...g.....C.....a.....g.....
AJ627183	620	caatttcaaggcacactatgcgtgagttaattgactctattcagttggttgccgcgatgattttcactgttctgagggtt
xy1Tg1Nt	602	a.....a.....
xy1Tg3Nt	581	a.....a.....
xy1Tg14Nb	581	C.....a.....
xy1Tg19Nb	572	.g.....a.....
AJ627183	700	agatttttg--aa-ttttgtttgtctcttttaaaaggttttaaaacttttgtgaatgttggcagatatggaatacacactaatg
xy1Tg1Nt	682	--a...C...a.....g.....a.....
xy1Tg3Nt	661	--a...C...a.....g.....a.....
xy1Tg14Nb	661	tga...-.....g.....a.....
xy1Tg19Nb	652	-t-a...a...a.....g.....t.....a.....
AJ627183	777	gattttgtttgatctgtttaatgaagattgtctagaacctcaatgttataaatatg-gtttggttgcttcattaataaaa
xy1Tg1Nt	760	t.....t.....
xy1Tg3Nt	739	t.....t.....
xy1Tg14Nb	740	.a...C.....g.....C.....tt...t.....
xy1Tg19Nb	729	a.....C.....tt...t.....
AJ627183	856	gagaattccttaatatcccgactagatgccagataaacaccaggttagttgacttttggattattggttgatttcatttga
xy1Tg1Nt	839	...C.....a...t.....g.a...ttt...a.....
xy1Tg3Nt	818	.C.....a...t.....g.a...ttt...a.....
xy1Tg14Nb	820	-----a.....--cg.a.....a.....
xy1Tg19Nb	809	-----a.....--g.....a.....

AJ627183	t	cagataaaattggttcattctttaaatgtttccactaaaagaattactcaagatctcagagtttatagttaggtgtatgtattt
xylTg1Nt	936	ga.ag.
xylTg3Nt	919	ga.ag.
xylTg14Nb	898	tgg.ac
xylTg19Nb	890	tgg.ac
	879	
AJ627183	1016	ggaattctggatttgatctagtattgaatggattactgaacttgtactccccagtcacatctggaggaggaacaatatgatca
xylTg1Nt	995	t.
xylTg3Nt	974	t.
xylTg14Nb	966	
xylTg19Nb	955	
AJ627183	1096	aattcaaaggttgaaaaagtaataactgagtcagaaattaaccacctttaacttgaaaaacggtaaataatgtatgttctaaga
xylTg1Nt	1075	a.
xylTg3Nt	1054	a.
xylTg14Nb	1046	c.
xylTg19Nb	1035	
AJ627183	1176	tggttatttcctataaaccttttgatgtcctaataatggagaaagtgagttgatttatgcgtttttcccctttatcggtgt
xylTg1Nt	1152	a..C.
xylTg3Nt	1131	a..C.
xylTg14Nb	1126	--a.C.
xylTg19Nb	1114	a.....
AJ627183	1256	tggtttttaaatctcatcaattccctttgtgttgattgctactcaaaattgaaccttagacggagtagcaaatagcaaaaagt
xylTg1Nt	1232	C.
xylTg3Nt	1211	C.
xylTg14Nb	1203	C.
xylTg19Nb	1194	
AJ627183	1336	aag---gccattctcttt-ctccctttcatctctctttatttccggtttgacatacacagaatatggtagcatctgtctgaagtgg
xylTg1Nt	1310	aaa-
xylTg3Nt	1289	aaa-
xylTg14Nb	1283	...aag....t....t....t....a....t....c....a....a....a....
xylTg19Nb	1274	...aaag.a....t....t....t....t....c....c....c....c....c....

Fig. 3 (Continuação)

```

AJ627183      1411 ttaattttattccttaaaatttgcataactaat-----
xylTg1Nt      1388 .....
xylTg3Nt      1367 .....
xylTg14Nb
xylTg19Nb      1353 .....c.....attccggtttttgttttttgcattatctttccattggcatgccatgtt
-----
AJ627183
xylTg1Nt
xylTg3Nt
xylTg14Nb
xylTg19Nb      1433 atttttggtttaggtttacataattatttatgtgatttctgatggagttactaatgattttttgtttttgttttt
-----
AJ627183
xylTg1Nt
xylTg3Nt
xylTg14Nb
xylTg19Nb      1513 tttcttttcccttgagtcgagggtcgattggaaatagcctctctctgcccttttggataggggtaaggccctgggtacg
-----
AJ627183      1444 -----tcgagtaaatgcctttttgaag
xylTg1Nt      1421 -----
xylTg3Nt      1400 -----
xylTg14Nb
xylTg19Nb      1593 gtaccatccccagacccactctgtgggactataccgggtagttgttgttgaat.....c
-----
AJ627183      1465 ctttagttgaatagttctacaactggttgttgcattttgaggactatcgacttgatttgacacttgacattgtctgatac
xylTg1Nt      1442 .....g
xylTg3Nt      1421 .....g
xylTg14Nb      1360 .....a.....a.....g.....g
xylTg19Nb      1673 .....g.....t.....
-----
AJ627183      1545 atggcttgtaagtattgaaaactttttatctaggaagaaatcccaaccagagatagggagctgtcacttggttatgagcta
xylTg1Nt      1507 .....a.....g.....
xylTg3Nt      1486 .....a.....g.....
xylTg14Nb      1440 .....g.....c.....t
xylTg19Nb      1741 -----t.....g

```

Fig. 3 (Continuação)

AJ627183	1625	ctggctcaaagttcaagtttgaccagttaatttttagatcttcaccaggataaacatttttagagtctaataatcctgaagc
xy1Tg1Nt	1587	g.....-g.....
xy1Tg3Nt	1566	g.....-g.....
xy1Tg14Nb	1520	g.....c.....
xy1Tg19Nb	1807	tg.....ta.....c.....t.....c.....
AJ627183	1705	agtattgtgcactaataaaggaggaacacatgaaggatgtagcactactaggttatgtttaccttatttactaatgat---tg
xy1Tg1Nt	1652	c.....g.c.....
xy1Tg3Nt	1631	c.....g.c.....
xy1Tg14Nb	1600	g.....c.....a.gac..
xy1Tg19Nb	1887	t.....g.....c.....-t.....t.....
AJ627183	1782	acaaccagcttaaatgatgacaaaatagtccttatatttgcctttttcacatttgctcatgactt-gggatatattccgaaatcaa
xy1Tg1Nt	1729	.g.....g.....g.....t.....t.....t.g.....
xy1Tg3Nt	1708	.g.....g.....g.....t.....t.....t.g.....
xy1Tg14Nb	1680	t.....g.....c.....-t.....t.....g.....
xy1Tg19Nb	1963	g.....g.....t.....t.....t.....
AJ627183	1861	catatttttagttcttttatgtactttaactacttatcaaaaaaattatccctgctaga-----tgttagtgttcaagcaac
xy1Tg1Nt	1808	...t..cc.....-gctaga.....
xy1Tg3Nt	1787	...t..cc.....gctaga.....
xy1Tg14Nb	1758	...t..c.....-gctaga.....
xy1Tg19Nb	2042	...t...cg.....-gctaga.....
AJ627183	1934	catgctagcttttaaggaaagctccttcttttgattcatgccatctttccgaaatcgatgccttacgcttactgtcatttttc
xy1Tg1Nt	1880	t.....
xy1Tg3Nt	1859	t.....
xy1Tg14Nb	1823	a.....c.....g.....t.....t.....
xy1Tg19Nb	2106	g.....t.....t.....
AJ627183	2014	taattttcatttcagtgattgaggagcgcgtcaccttttgattacacgatttgagtatgcaaaccttttccacacagttac
xy1Tg1Nt	1960	
xy1Tg3Nt	1939	
xy1Tg14Nb	1903	a.....
xy1Tg19Nb	2180	g.....

Fig. 3 (Continuação)

AJ627183	2094	cgattggtatagtcatacgtggcatccagggttactggcttggccagtcggccacatttgggtttttgtagatggccatt
xy1Tg1Nt	2040	
xy1Tg3Nt	2019	
xy1Tg14Nb	1983	g.....
xy1Tg19Nb	2260	c.....t.....a.....
AJ627183	2174	gtgaggatgtttgaaagtattgataacgatggcatgcatgtgtactgtgttacggatgaaagaaatgaaaccagcaatta
xy1Tg1Nt	2120	t.....t.....t.....
xy1Tg3Nt	2099	t.....t.....t.....
xy1Tg14Nb	2063	c.....c.....t.....
xy1Tg19Nb	2340	c.....t.....t.....t.....
AJ627183	2254	ttttctagcaggcaatgctcttgagatgcttgtgtcaaatggtcagacttaatcctgagtttccatttgtttcagcttt
xy1Tg1Nt	2200	
xy1Tg3Nt	2179	
xy1Tg14Nb	2143	
xy1Tg19Nb	2420	t.....a.....t.....g.....a.....g.....
AJ627183	2334	ctgtttgactgactacaataattgtcccaattaggggtgtcaatgggatattcgaaagccgactaaaaccgacgaaccgta
xy1Tg1Nt	2280	
xy1Tg3Nt	2259	
xy1Tg14Nb	2223	...g.....
xy1Tg19Nb	2500	
AJ627183	2414	cogtacogatttttaggtttctttttaagaaacgtaggttttttatataaatctataaacccgaccgataattagggtagg
xy1Tg1Nt		-----
xy1Tg3Nt		-----
xy1Tg14Nb		-----
xy1Tg19Nb		-----
AJ627183	2494	ttttttattttataaaaaaagcgcgaaaaataacgcgaaccgtaccgaataaaattttacatgtggaaaaatatatttatta
xy1Tg1Nt		-----
xy1Tg3Nt		-----
xy1Tg14Nb		-----
xy1Tg19Nb		-----

Fig. 3 (Continuação)

AJ627183	2574	gtaagtttaaaaaataataatgcattaaattttctttgggccatggaattatgaaaactattacaagccaagaataatta
xy1Tg1Nt		-----
xy1Tg3Nt		-----
xy1Tg14Nb		-----
xy1Tg19Nb		-----
AJ627183	2654	gactcaaaataactaattccttaaacctattatgttacttctacttaactaagttatttcaagtatctttattagcaaga
xy1Tg1Nt		-----
xy1Tg3Nt		-----
xy1Tg14Nb		-----
xy1Tg19Nb		-----
AJ627183	2734	cacaaaagtattctctagcgattatgagcaaaactacaatgtattgaatatgtttccttccatatatattttagattatcttttt
xy1Tg1Nt		-----
xy1Tg3Nt		-----
xy1Tg14Nb		-----
xy1Tg19Nb		-----
AJ627183	2814	gaatattttaatcttctctatagactctctattcttggagtcgccagcttggtatatctttcactcgtgtgattatatttctttt
xy1Tg1Nt		-----
xy1Tg3Nt		-----
xy1Tg14Nb		-----
xy1Tg19Nb		-----
AJ627183	2894	gcctttgtttgatttcttcttaacgttggttagaatagtcgatggatctatactctagccatcttcttcttcttttttta
xy1Tg1Nt		-----
xy1Tg3Nt		-----
xy1Tg14Nb		-----
xy1Tg19Nb		-----
AJ627183	2974	attcatcaccttttaaacagtaaaaaatgtctaaagaatttttcttaagtcctctataaaaagaacgtatgtttatgcattctac
xy1Tg1Nt		-----
xy1Tg3Nt		-----
xy1Tg14Nb		-----
xy1Tg19Nb		-----

Fig. 3 (Continuação)

AJ627183	3054	ttctactggtgaattttacatgatattaaaaaattaaacccgaaccttaccgctaccgaagagaaaaccgacatgattgggac
xy1Tg1Nt		----
xy1Tg3Nt		----
xy1Tg14Nb		----
xy1Tg19Nb		----
AJ627183	3134	ggtttcgaaaagtctaattttggttatacataatagaataaaccgaaaaattggtatggtacaaaattttataaaaaataaccg
xy1Tg1Nt		----
xy1Tg3Nt		----
xy1Tg14Nb		----
xy1Tg19Nb		----
AJ627183	3214	gccgaaccgaaccattgacaccccctagtcaccaatacctagttgttgagtttgctcattcttacttctatttacgt--ca
xy1Tg1Nt	2303	-----c.g..g.....gt..
xy1Tg3Nt	2282	-----c.g..g.....gt..
xy1Tg14Nb	2246	-----g.....a.....t.....gt..
xy1Tg19Nb	2520	-----t.....g..g.....gt..
AJ627183	3292	ctgtttctctgaatggtccctttgtggtgaaaagagcttttgcctatgtagaaaaactagcaaatgatttcatagctgaaac
xy1Tg1Nt	2357	-----t.....a.....tt...gg..
xy1Tg3Nt	2336	-----t.....a.....tt...gg..
xy1Tg14Nb	2300	-----a.....a.....tt...g..
xy1Tg19Nb	2577	-----a.....a.....tt...g.g.
AJ627183	3372	aatttatttttacctacatcatgtcttataaaaattgcttctaactgtatactttaattcttggagagatgctttcatgt
xy1Tg1Nt	2437	-----c..c.....ct.t.c.....a....
xy1Tg3Nt	2416	-----c..c.....ct.t.c.....a....
xy1Tg14Nb	2379	c.....c.g.c.....
xy1Tg19Nb	2657	.c.....c..c.....
AJ627183	3452	gaagaaagtcttttccactccact-----actgga
xy1Tg1Nt	2517	tttaattcttggagagatgcttttaattgtgaagaaaagttctttccactccact....t.
xy1Tg3Nt	2496	tttaattcttggagagatgcttttaattgtgaagaaaagttctttccactccact....t.
xy1Tg14Nb	2459	
xy1Tg19Nb	2737	...t.....

Fig. 3 (Continuação)

AJ627183	3481	agcttgccgtatgaattttacttggccatattgtggccgtgctttgatttatcttcaaatcattttcttcataatagttc
xy1Tg1Nt	2597	t.c.....g.g...t.....g.a...
xy1Tg3Nt	2576	t.c.....g.g...t.....g.a...
xy1Tg14Nb	2488	t.c.....g.g...t.....g.....
xy1Tg19Nb	2766	t.c.....a.....c.g.....t.....g.....
AJ627183	3561	tttcgagtaattcttttttctctcttttctgtttga-aaaaaattcagacacaattggaggaaacatggaaagcacttttt
xy1Tg1Nt	2677	...t.....-.....t.g.....gg.....c.....
xy1Tg3Nt	2656	...t.....-.....t.g.....gg.....c.....
xy1Tg14Nb	2568	-.....-.....t.....t.....
xy1Tg19Nb	2846	-a.....--.....-.....
AJ627183	3640	tcaagcctcacttatgtctaagaacttttagtgccagtttgtttccgtcatgcccctctcgcctttgggatatgaaac
xy1Tg1Nt	2756	c.....c.....tt.....
xy1Tg3Nt	2735	c.....c.....tt.....
xy1Tg14Nb	2646	t.....t.....
xy1Tg19Nb	2922	t.....t.....a.....
AJ627183	3720	tgcctgttttaagggactgtcagaaaactatagattgttaatggagcttctgcccattgttggcaaaaatcctgatgata
xy1Tg1Nt	2836	a.....
xy1Tg3Nt	2815	a.....
xy1Tg14Nb	2726	t.....t.....
xy1Tg19Nb	3002	t.....t.....g.....
AJ627183	3800	agaaaactgcacggttgtccgagtttggggagatgattagggcagccttttagatttccctgtggatagacagaacatccca
xy1Tg1Nt	2916	...g.....c.....g.....
xy1Tg3Nt	2895	...g.....c.....g.....
xy1Tg14Nb	2806	a.....c.....g.....t.....
xy1Tg19Nb	3082	a.....c.....c.....g.....
AJ627183	3880	aggacagtcacaggccctaattgtcctctttttagacgtgaggattatttagctcaccacgtcatggtggaaagggtaca
xy1Tg1Nt	2996	
xy1Tg3Nt	2975	
xy1Tg14Nb	2886	
xy1Tg19Nb	3162	

Fig. 3 (Continuação)

AJ627183	3960	gtctaggcttagcaatgaagagcaagtagtttgcataaaagagctgggccttgaccactcggagtgcaaatataatg
xylTg1Nt	3076	
xylTg3Nt	3055	
xylTg14Nb	2966	a.....
xylTg19Nb	3242	t.....t.....
AJ627183	4040	taattaacggattgtttgcccacatgtccatgaaagagcaagttcgcagcaatccaagatgcttctgtcatagttgggtgct
xylTg1Nt	3156	a.....t.c.c.c...
xylTg3Nt	3135	a.....t.....
xylTg14Nb	3046	gt.....t.....
xylTg19Nb	3322	t.....t.....
AJ627183	4120	catggagcaggtctaactcacatagtttctgcagcaccacaaagctgtaatactagaaaattataaagcagcgaatataggcg
xylTg1Nt	3236	
xylTg3Nt	3215	c.....
xylTg14Nb	3126	
xylTg19Nb	3402	
AJ627183	4200	ccccattttgctctgattgcacaatggaaaggattggagtaccatcccatatatttggagggtctttatgcggatcctc
xylTg1Nt	3316	
xylTg3Nt	3295	
xylTg14Nb	3206	t.....
xylTg19Nb	3482	
AJ627183	4280	cagttgtgatcgacaaagctcagcagcattttgaggagctctgggtgctaaatctgctcgacagtttagttcgtcttttct
xylTg1Nt	3396	
xylTg3Nt	3375	
xylTg14Nb	3286	
xylTg19Nb	3562	
AJ627183	4360	ctaaaagactgggaaggatagagggaattcgggggttctggaaacctggagcctgggaattgtgtaaaaatatgtttcacacgc
xylTg1Nt		-----
xylTg3Nt		-----
xylTg14Nb		-----
xylTg19Nb		-----

Fig. 3 (Continuação)

AJ627183	4440	agttctatagtcaattgctgcaatctggtgttcataagcttggaatttccagcagctactaaattattagcccactctg
xy1Tg1Nt		-----
xy1Tg3Nt		-----
xy1Tg14Nb		-----
xy1Tg19Nb		-----
AJ627183	4520	actcagttatggactaccagagcaatcatatcaaatgggagcatggaatcctgattgtggaatggtgagctcattgaaga
xy1Tg1Nt		-----
xy1Tg3Nt		-----
xy1Tg14Nb		-----
xy1Tg19Nb		-----
AJ627183	4600	gcataattctttatggtgtgaagattacagttgacgagtaacacgtgtatgtgaaagattaggttgttacactttcttgc
xy1Tg1Nt		-----
xy1Tg3Nt		-----
xy1Tg14Nb		-----
xy1Tg19Nb		-----
AJ627183	4680	aattcattgtcaattgttttttcgtcattctttattaatgatcataggataagaacatgagaaaaaccatccatgttctctgt
xy1Tg1Nt		-----
xy1Tg3Nt		-----
xy1Tg14Nb		-----
xy1Tg19Nb		-----
AJ627183	4760	tgttttcccatcaatctggccaccctctttccctctttatgtagagatgatttcaacagagtttgtttttagttgtaata
xy1Tg1Nt		-----
xy1Tg3Nt		-----
xy1Tg14Nb		-----
xy1Tg19Nb		-----
AJ627183	4840	cttggtactcacagttactgttttgcattcatcccatcagatgtcgaagaagcagagattaacaagaacgtcagtatgatgtt
xy1Tg1Nt		-----
xy1Tg3Nt		-----
xy1Tg14Nb		-----
xy1Tg19Nb		-----

Fig. 3 (Continuação)

AJ627183	4920	tcagtgaatatatgggttgtaacttgtaacccaacaaaagaatgagactttgac
xylTg1Nt		-----
xylTg3Nt		-----
xylTg14Nb		-----
xylTg19Nb		-----

Fig. 4

AJ627183	1	mnkkklkilvslfalnsitlylyfsshpdhfrhksrqnhf-----slsenrhhmfhssitsqyskpwipilsypwswq
xylTg1Nt	1	-----s.....p....pntqnhy.....h.d....v...t.....
xylTg3Nt	1	-----s.....p....pntqnhy.....h.d....v...t.....
xylTg14Nb	1	-----s.....p....pntqnhy.....h.d....v...t.....
xylTg19Nb	1	-----h.....p....pntqnhy.....h.d....v...t.....
AJ627183	74	npnvvrscgyfgngftlkvdllktspefhrkfgentvsgdggwfrcffsetlqssiceggairmnpdeillmsrggekl
xylTg1Nt	72	sl.....l.q.....
xylTg3Nt	65	sl.....l.q.....
xylTg14Nb	65	a.....l.....f.....
xylTg19Nb	64	a.....d.....d.....
AJ627183	154	esvigrseddelpvfkngafqikvtdklkigkklvdekflnkylpegaisrhtmlrelidsiqlvgaddfhcsewieeps1
xylTg1Nt	152	i.....i.....e.....
xylTg3Nt	145	i.....i.....e.....
xylTg14Nb	145	v.a.t.....f.....n.....
xylTg19Nb	144	n.....e.....k.....g.....e.....v.....
AJ627183	234	litrfeyanlfhtvtdwysayvasrvtglpsrphlvfvdghcetqleetwkalfsslt yaknfsqpvfcfrhaalsplgye
xylTg1Nt	232	v.....v.....
xylTg3Nt	225	v.....v.....
xylTg14Nb	225	i.....n.....v.....
xylTg19Nb	224	a.....n.....v.....
AJ627183	314	talfkglsetidngasahdlwqnpddkktarlsefgemiraafrfpvdrqnprrtvtgpnvlfvrredylahprhgkv
xylTg1Nt	312	t.....r.....g.....
xylTg3Nt	305	t.....r.....g.....
xylTg14Nb	305	g.....g.....
xylTg19Nb	304	k.....g.....
AJ627183	394	qsrlnseeqvfdsikswalnhsecklnvngl fahmsmkeqvraiqdasvi vghagaglt hivsaaapkavileissseyr
xylTg1Nt	392	l.....
xylTg3Nt	385	s.....
xylTg14Nb	385	s.....
xylTg19Nb	384	l.....

Fig. 4 (Continuação)

AJ627183	474	rphfaliaqwkgleypylegsyadppvvidklssilrslgc
xy1Tg1Nt	472	
xy1Tg3Nt	465	
xy1Tg14Nb	465	
xy1Tg19Nb	464	

RESUMO

Patente de Invenção: "**SEQÜÊNCIAS DE NUCLEOTÍDEO QUE CODIFICAM
NICOTIANA BETA-1,2-XILOSILTRANSFERASE**".

A presente invenção refere-se a novas seqüências de nucleotídeo
5 de 1,2-xilosiltransferase e usos das mesmas.