

(11) Número de Publicação: **PT 1786908 E**

(51) Classificação Internacional:

**C12N 15/82** (2007.10) **C12N 15/54** (2007.10)

**C12N 9/12** (2007.10) **A01H 5/10** (2007.10)

**A21D 2/18** (2007.10) **A23L 1/10** (2007.10)

**A23L 1/522** (2007.10) **C08B 30/00** (2007.10)

(12) **FASCÍCULO DE PATENTE DE INVENÇÃO**

(22) Data de pedido: **2005.08.18**

(30) Prioridade(s): **2004.08.18 EP 04090316**  
**2004.08.18 US 602454 P**

(43) Data de publicação do pedido: **2007.05.23**

(45) Data e BPI da concessão: **2010.03.03**  
**082/2010**

(73) Titular(es):

**BAYER CROPSCIENCE AG**

**ALFRED-NOBEL-STRASSE 50 40789 MONHEIM**  
**DE**

(72) Inventor(es):

**CLAUS FROHBERG** **DE**

(74) Mandatário:

**MANUEL ANTÓNIO DURÃES DA CONCEIÇÃO ROCHA**  
**AV LIBERDADE, Nº. 69 1250-148 LISBOA**

**PT**

(54) Epígrafe: **PLANTAS COM UM AUMENTO DA ACTIVIDADE DA ENZIMA DE FOSFORILAÇÃO DO AMIDO R3 NOS PLASTÍDEOS**

(57) Resumo:

**RESUMO****"PLANTAS COM UM AUMENTO DA ACTIVIDADE DA ENZIMA DE  
FOSFORILAÇÃO DO AMIDO R3 NOS PLASTÍDEOS"**

A presente invenção refere-se a células vegetais e plantas que são geneticamente modificadas, conduzindo a modificação genética ao aumento da actividade de uma proteína R3 de fosforilação do amido em comparação com as células vegetais de tipo selvagem ou plantas de tipo selvagem que não foram geneticamente modificadas. Além disso, a presente invenção refere-se a meios e métodos de geração dessas células vegetais e plantas. As células vegetais e plantas deste tipo sintetizam um amido modificado. Por conseguinte, a presente invenção refere-se também ao amido sintetizado pelas células vegetais e plantas de acordo com a presente invenção, métodos para a produção deste amido e produção dos derivados de amido deste amido modificado, bem como farinhas que contêm amidos de acordo com a presente invenção. Mais ainda, a presente invenção refere-se também a ácidos nucleicos codificadores de proteínas R3 de fosforilação de amidos e a vectores, células hospedeiras, células vegetais e plantas que incluem tais moléculas de ácido nucleico. Além disso, a presente invenção refere-se igualmente a proteínas R3 com uma actividade de fosforilação de amido.

**DESCRIÇÃO****"PLANTAS COM UM AUMENTO DA ACTIVIDADE DA ENZIMA DE  
FOSFORILAÇÃO DO AMIDO R3 NOS PLASTÍDEOS"**

A presente invenção refere-se a células vegetais e plantas que são geneticamente modificadas, conduzindo a modificação genética ao aumento da actividade de uma proteína R3 de fosforilação do amido em comparação com as células vegetais de tipo selvagem ou plantas de tipo selvagem que não foram geneticamente modificadas. Além disso, a presente invenção refere-se a composições e métodos de geração dessas células vegetais e plantas. As células vegetais e plantas deste tipo sintetizam um amido modificado. Por conseguinte, a presente invenção refere-se também ao amido sintetizado pelas células vegetais e plantas de acordo com a presente invenção, métodos para a produção deste amido e produção dos derivados de amido deste amido modificado, bem como farinhas que contêm amidos de acordo com a presente invenção. Mais ainda, a presente invenção refere-se também a ácidos nucleicos codificadores de proteínas R3 de fosforilação de amidos e a vectores, células hospedeiras, células vegetais e plantas que incluem tais moléculas de ácido nucleico. Além disso, a presente invenção refere-se igualmente a proteínas R3 com uma actividade de fosforilação de amido.

Considerando a importância cada vez maior actualmente associada aos constituintes vegetais como matérias primas renováveis, uma das funções da investigação biotecnológica

consiste em tentar adaptar estas matérias primas vegetais aos requisitos da indústria transformadora. De forma a tornar possível a utilização de matérias primas renováveis no maior número possível de campos de aplicação, é adicionalmente necessário chegar a muitas substâncias diversas.

O polissacárido amido é constituído por unidades quimicamente uniformes, as moléculas de glucose, mas constitui, no entanto, uma mistura complexa de diferentes formas de moléculas que exibem diferenças relativamente ao grau de polimerização e ramificação e, portanto, diferem muito entre si quanto às suas características físico-químicas. Distingue-se entre a amilose do amido, um polímero essencialmente não ramificado constituído por unidades de glucose com ligações alfa-1,4-glicosídicas, e a amilopectina do amido, um polímero ramificado, no qual as ramificações derivam da ocorrência de ligações adicionais alfa-1,6-glicosídicas. Uma outra diferença essencial entre a amilose e a amilopectina reside no peso molecular. Enquanto a amilose possui, conforme a origem do amido, um peso molecular de  $5 \times 10^5$ - $10^6$  Da, o da amilopectina situa-se entre  $10^7$  e  $10^8$  Da. As duas macromoléculas podem ser distinguidas pelo seu peso molecular e diferentes características físico-químicas, que podem ser facilmente tornadas visíveis pelas respectivas características de ligação do iodo diferentes.

A amilose tem sido, desde há muito, considerada como um polímero linear, constituído por monómeros de alfa-D-glucose com ligações alfa-1,4-glicosídicas. Em estudos mais recentes, contudo, foi demonstrada a presença de pontos de ramificação alfa-1,6-glicosídicos (cerca de 0,1 %)

(Hizukuri e Takagi, Carbohydr. Res. 134, (1984), 1-10; Takeda et al., Carbohydr. Res. 132, (1984), 83-92).

As características funcionais do amido, tal como, por exemplo, a solubilidade, o comportamento de retrogradação, a capacidade de ligação da água, características de formação de película, viscosidade, propriedades de gelificação, estabilidade no ciclo congelação-descongelação, resistência da gelificação, dimensão da partícula de amido dos amidos e outras características são afectadas pela proporção amilose/amilopectina, o peso molecular, o padrão da distribuição da cadeia lateral, a concentração de iões, o teor de lípidos e proteínas, o tamanho médio de partícula do amido, morfologia da partícula de amido, etc. As características funcionais do amido são também afectadas pelo teor de fosfato, um componente do amido que não é um hidrato de carbono. Distingue-se, aqui, entre o fosfato com ligação covalente às moléculas de glucose sob a forma de monoésteres (posteriormente designados fosfato de amido) e fosfatos sob a forma de fosfolípidos associados ao amido.

O conteúdo em fosfato de amido varia em função da variedade da planta. Assim, por exemplo, alguns mutantes de amido sintetizam um amido com um conteúdo maior de fosfato de amido (amido ceroso 0,002 % e milho com alto teor de amilose 0,013 %), enquanto as espécies convencionais de milho apenas possuem vestígios de fosfato de amido. Encontram-se, igualmente, pequenas quantidades de fosfato de amido no trigo (0,001 %) enquanto que na aveia e sorgo não se detectou fosfato de amido. Do mesmo modo, encontra-se menos fosfato de amido em mutantes do arroz (amido ceroso 0,003 %) do que nas espécies convencionais de arroz

(0,013 %). Detectaram-se quantidades significativas de fosfato de amido em plantas sintetizadoras de amido bulbosas ou com raízes tuberosas, tal como por exemplo, a tapioca (0,008 %), batata doce (0,011 %), araruta (0,021 %) ou batata (0,089 %). As percentagens acima indicadas para a concentração de fosfato de amido são baseadas em cada caso no peso seco do amido e foram determinadas por Jane et al. (1996, *Cereal Foods World* 41 (11), 827-832).

O fosfato de amido pode existir sob a forma de monoésteres na posição C-2, C-3 ou C-6 dos monómeros de glucose polimerizados (Takeda e Hizukuri, 1971, *Starch/Stärke* 23, 267-272). A distribuição do fosfato no amido sintetizado por plantas é normalmente diferenciada pelo facto de que aproximadamente 30 % a 40 % dos resíduos de fosfato têm uma ligação covalente na posição C-3 e aproximadamente 60 % a 70 % dos resíduos fosfato têm uma ligação covalente na posição C-6 das moléculas de glucose (Blennow et al., *Int. J. of Biological Macromolecules* 27, 211-218). Blennow et al. (2000, *Carbohydrate Polymers* 41, 163-174) determinou uma concentração de fosfato de amido ligada na posição C-6 das moléculas de glucose para vários amidos, tais como amido de batata (entre 7,8 e 33,5 nMol por mg de amido, conforme a espécie), amido de várias espécies de Curcuma (entre 1,8 e 63 nMol por mg), amido de tapioca (2,5 nMol por mg de amido), amido de arroz (1,0 nMol por mg de amido), amido de feijão mungo (3,5 nMol por mg de amido) e amido de sorgo (0,9 nMol por mg de amido). Nenhum destes autores detectou a existência de fosfato de amido ligado na posição C-6 em amido de **cevada** e em amido de diferentes mutantes cerosos de milho. Não se estabeleceu ainda qualquer ligação entre o genotipo de uma planta e a

concentração de fosfato de amido (Jane et al., 1996, *Cereal Foods World* 41 (11), 827-832). Actualmente não é possível influenciar a concentração de fosfato de amido em plantas através de métodos de produção de plantas.

Foi apenas descrita uma proteína que medeia a introdução de ligações covalentes de resíduos de fosfato em moléculas de glucose do amido. Esta proteína possui a actividade enzimática de uma alfa-glucano-água-diquinase (GWD, E.C.: 2.7.9.4) (Ritte et al., 2002, *PNAS* 99, 7166-7171), é frequentemente designada R1 na literatura científica e está ligada aos grânulos de amido do amido de armazenagem em tubérculos de batata (Lorberth et al., 1998, *Nature Biotechnology* 16, 473-477). Considerando que o amido de armazenamento é sintetizado em plastídeos (amiloplastos) e a proteína R1 é ligada a estas partículas de amido, mas codificada por ácidos nucleicos localizados no núcleo, a sequência de aminoácidos é facultada com um peptídeo sinal para o plastídico (transit) (Lorberth et al., 1996, *Nature Biotechnology* 16, 473-477). Na reacção catalisada por R1, os materiais de partida alfa-1-glucano (amido), trifosfato de adenosina (ATP) e água são convertidos nos produtos fosfato de glucano (fosfato de amido), monofosfato e monofosfato de adenosina. Em simultâneo, o resíduo gama-fosfato do ATP é transferido para a água e o resíduo beta-fosfato do ATP é transferido para o glucano (amido). A R1 transfere o resíduo beta-fosfato *in vitro* do ATP para as posições C-6 e C-3 das moléculas de glucose de alfa-1,4-glucanos. A proporção de fosfato C-6 para fosfato C-3 que é obtida através da reacção *in vitro* corresponde à proporção que existe no amido isolado a partir de plantas (Ritte et al., 2002, *PNAS* 99, 7166-7171). Como cerca de 70 % do

fosfato de amido presente no amido de batata se encontra na posição C-6 e cerca de 30 % na posição C-3 dos monómeros de glucose do amido, quer dizer que R1 fosforila de preferência a posição C-6 das moléculas de glucose. Além disso, demonstrou-se com R1, entre outros aspectos, que a utilização de amilopectina de milho permite a fosforilação R1 de alfa-1,4-glucanos, que não contém ainda fosfato com ligação covalente (Ritte et al., 2002, PNAS 99, 7166-7171), isto é R1 é capaz de introduzir fosfato de novo em alfa-1,4-glucanos.

As sequências de ácidos nucleicos e sequências de aminoácidos correspondentes, codificadoras de uma proteína R1 encontram-se descritas a partir de diferentes espécies, tal como a batata (WO 97 11188, GenBank Acc.: AY027522, Y09533), trigo (WO 00 77229, patente norte-americana nº 6,462,256, GenBank Ace: AAN93923, GenBank Ace: AR236165), arroz (GenBank Ace: AAR61445, GenBank Ace: AR400814), milho (GenBank Ace: AAR61444, GenBank Ace: AR400813), soja (GenBank Ace: AAR61446, GenBank Ace: AR400815), limão (GenBank Acc.: AY094062) e *Arabidopsis* (GenBank Acc.: AF312027).

Na WO 02 34923 descreve-se plantas do trigo com uma actividade acrescida de uma proteína R1 em resultado da sobreexpressão de um gene R1 da batata. Em comparação com as plantas tipo selvagem, nas quais não foi detectado qualquer fosfato de amido, estas plantas sintetizam um amido com quantidades significativas de fosfato de amido na posição C-6 das moléculas de glucose.

Não foram descritas outras proteínas que catalisem uma reacção que introduz grupos de fosfato com ligação

covalente em amido, assim como enzimas conhecidas que introduzam preferencialmente grupos de fosfato na posição C-3 e/ou posição C-2 das moléculas de glucose do amido. À parte do aumento do teor de fosfato de amido em plantas, também não existem vias para influenciar especificamente a fosforilação do amido em plantas, de modificar a distribuição do fosfato no amido sintetizado por plantas e/ou aumentar mais o teor de fosfato de amido.

O objectivo da presente invenção consiste, assim, no fornecimento de mais possibilidades de geração de amidos modificados com teor de fosfato acrescido e/ou proporcionar células vegetais e/ou plantas que sintetizam este amido modificado, bem como métodos e meios para a geração destas plantas e/ou células vegetais.

Este objectivo é alcançado nas formas de utilização designadas nas reivindicações

Deste modo, a presente invenção refere-se a células vegetais geneticamente modificadas e plantas geneticamente modificadas, em que os plastídeos presentes em tais células vegetais ou plantas incluem uma proteína R3.

A presente invenção refere-se igualmente a células vegetais e plantas geneticamente modificadas, em que os plastídeos presentes em tais células vegetais ou plantas apresentam uma actividade acrescida de, pelo menos, uma proteína R3 comparativamente aos plastídeos presentes em células vegetais do tipo selvagem que não foram geneticamente modificadas.

Um primeiro aspecto da presente invenção refere-se a células vegetais e plantas que são geneticamente modificadas, conduzindo a modificação genética à presença

de uma proteína R3 ou a um aumento da actividade de, pelo menos, uma proteína R3 nos plastídeos presentes nestas células vegetais ou plantas em comparação com os plastídeos presentes nas células vegetais de tipo selvagem ou plantas de tipo selvagem correspondentes que não foram geneticamente modificadas. A modificação genética pode ser tomar a forma de qualquer modificação genética conducente a um aumento da actividade de pelo menos uma proteína R3 nos plastídeos presentes em células vegetais geneticamente modificadas ou plantas geneticamente modificadas em comparação com as células vegetais do tipo selvagem e plantas do tipo selvagem correspondentes que não foram geneticamente modificadas.

No contexto da presente invenção, o termo "célula vegetais de tipo selvagem" significa que as células vegetais em questão foram utilizadas como material de partida para a produção das células vegetais de acordo com a invenção, isto é, a sua informação genética, independentemente da modificação genética introduzida, corresponde à de uma célula vegetal de acordo com a presente invenção.

No contexto da presente invenção, o termo "planta de tipo selvagem" significa que as plantas em questão foram utilizadas como material de partida para a produção das plantas de acordo com a invenção, isto é, a sua informação genética, independentemente da modificação genética introduzida, corresponde à de uma planta de acordo com a presente invenção.

No contexto da presente invenção, o termo "correspondente" significa que, em comparação com vários

objectos, os objectos em questão que são comparados entre si foram mantidos nas mesmas condições. No contexto da presente invenção, o termo "correspondentes", no contexto das células vegetais do tipo selvagens e plantas do tipo selvagens, significa que as células vegetais ou plantas que são comparadas entre si foram produzidas em condições de cultura idênticas e têm uma idade (cultura) idêntica.

No contexto da presente invenção, o termo "plastídeo" é entendido como organelos celulares que existem apenas nas células vegetais e os quais são circundadas por uma bicamada e possuem o seu próprio sistema genético (plastoma, aparelho de transcrição e aparelho de tradução). Os plastídeos, os quais são preferidos no contexto da presente invenção, são cromoplastos e leucoplastos, sendo especialmente preferidos os cloroplastos e com particular preferência amiloplastos.

No âmbito da presente invenção, o termo "actividade acrescida de pelo menos uma proteína R3" significa um aumento da expressão de genes endógenos codificadores das proteínas R3 e/ou um aumento da quantidade de proteína R3 nas células e/ou um aumento da actividade enzimática das proteínas R3 nas células.

O aumento na expressão pode ser determinada, por exemplo por meio da medição da quantidade de transcrições codificadoras das proteínas R3, por exemplo, através da análise *Northern blot* ou por RT-PCR. Neste caso, um aumento significa de preferência um aumento da quantidade de transcrições em comparação com as células correspondentes que não foram geneticamente modificadas de pelo menos 50%, em particular pelo menos 70%, de preferência pelo menos 85%

e com particular preferência pelo menos 100%. Um aumento na quantidade de transcrições codificadoras da proteína R3 significa igualmente que as plantas ou células vegetais que não possuem quantidade detectável de transcrições codificadoras da proteína R3 incluem, após modificação genética de acordo com a invenção, quantidades detectáveis de transcrições codificadoras da proteína R3.

O aumento da quantidade da proteína R3, a qual implica uma actividade acrescida destas proteínas nas células vegetais em questão, pode ser determinado, por exemplo, através de métodos imunológicos tal como a análise de Western blot, ELISA (ensaio imuno-enzimático) ou RIA (radio imuno ensaio) Neste caso, um aumento significa de preferência um aumento da quantidade da proteína R3 em comparação com as células correspondentes que não foram geneticamente modificadas de pelo menos 50%, em particular pelo menos 70%, de preferência pelo menos 85% e com particular preferência pelo menos 100%. Um aumento da quantidade de uma proteína R3, significa também que as plantas ou células vegetais que não possuem actividade detectável de uma proteína R3 possuem uma quantidade detectável de uma proteína R3 após modificação genética de acordo com a presente invenção.

Métodos para aumentar anticorpos que reagem especificamente a determinadas proteínas, isto é, que se ligam especificamente à citada proteína, são conhecidos dos especialistas (ver, por exemplo, Lottspeich e Zorbas (Eds.), 1998, Bioanalytik, Spektrum Akad. Verlag, Heidelberg, Berlin, ISBN 3-8274-0041-4). Estes anticorpos são aumentados comercialmente por algumas companhias (por exemplo, Eurogentec, Bélgica).

Pode detectar-se a localização de uma proteína R3 em plastídeos de plantas geneticamente modificadas através de fraccionamento subcelular de células vegetais, onde se podem obter plastídeos intactos. Métodos para fraccionar compartimentos subcelulares são conhecidos dos especialistas e descritos na literatura (ver, por exemplo, Heldt, 1996, "Pflanzenbiochemie", Spektrum Akad. Verl., ISBN 3-8274-0103-8; Plant Biochemistry, 1997, Dey e Harborne Ed., Academic Press, ISBN 0-12-214674-3) O material de partida que pode ser empregue para isolar plastídeos inclui partes de plantas (por exemplo, folhas) ou protoplastos. Se as partes de plantas forem empregues como material de partida, as células são fragmentadas primeiro pelas partes das plantas que são cortadas em pequenas secções e subsequentemente divididas num misturador. O homogenato celular daí resultante é subsequentemente sujeito a centrifugação.

Um outro método para preparação de um homogenato celular que inclua plastídeos começa pela preparação de protoplastos que são subsequentemente fragmentados. Esta fragmentação ocorre pela suspensão do protoplasto que é comprimido contra numa malha com uma dimensão de malha menor que o tamanho dos protoplastos. O homogenato celular daí resultante é subsequentemente sujeito a centrifugação. O último método mencionado é particularmente apropriado para a preparação do homogenato celular de plantas de cereais, cujo homogenato inclui plastídeos intactos

O homogenato celular pode ser purificado recorrendo a centrifugação diferencial ou centrifugação de densidade.

No caso de centrifugação diferencial, os homogenatos celulares são centrifugados num meio cuja densidade é inferior ao dos organelos celulares. A taxa de sedimentação dos organelos individuais depende principalmente do tamanho da partícula. Os organelos individuais são separados efectuando-se uma pluralidade de centrifugações, uma a seguir à outra, com uma velocidade crescente. Após cada centrifugação efectuada, os organelos em questão encontram-se sedimentados.

No caso de centrifugação de equilíbrio, os organelos são separados de acordo com a densidade num meio de centrifugação que consiste em meios diferentes de várias densidades, as quais são colocadas umas em cima das outras. A densidade do meio individual diminui de baixo para cima. A centrifugação é realizada até que todas as partículas do homogenato celular terem alcançado a zona correspondente à sua densidade. É descrito o isolamento de amiloplastos intactos, *inter alia*, para a batata (Wischmann et al., 1999, Plant Physiology 119, 455-462), milho (Echeveria et al., 1988, Plant Physiology 86, 786-792) e trigo Tetlow et al., 1993, Planta 189, 597-600), ervilha (Smith et al., 1990, Planta 180, 517-523).

Para se determinar a pureza de uma preparação de organelos, podem medir-se determinadas enzimas de principais, as quais são específicas dos organelos em questão, em fracções individuais de organelos. As enzimas principais para determinados organelos ou compartimentos subcelulares são conhecidas dos especialistas e mencionadas *inter alia* em Strasburger, 1999, "Lehrbuch der Botanik" [Manual de Botânica], 34ª Edição, Spektrum Akad. Verl., ISBN 3-8274-0779-6. A fracção dos compartimentos

subcelulares que incluem plastídeos pode ser analisada quanto à presença ou actividade acrescida de uma proteína R3. Esta análise pode ser efectuada, por exemplo, através de métodos imunológicos ou por detecção da actividade na fracção dos plastídeos.

No âmbito da presente invenção, o termo "proteína R3" pressupõe significar uma proteína que transfere um resíduo fosfato de um nucleosídeo trifosfato para amido.

As sequências de aminoácidos codificadoras das proteínas R3 contêm um domínio fosfo-histidina. Os domínios fosfo-histidina são descritos, por exemplo, por Tien-Shin Yu et al. (2001, *Plant Cell* 13, 1907-1918). O domínio fosfo-histidina da protein R3 *Arabidopsis thaliana* é apresentado em SEQ ID NO 5.

Comparativamente a uma proteína R1, as proteínas R3 não têm peptídeo sinal plastídico nas sequências aminoácidas que as (naturalmente) codificam, isto é, as proteínas R3 codificadoras de ácidos nucleicos endógenos de plantas (ácidos nucleicos que ocorrem naturalmente no genoma da planta) estão localizadas fora dos plastídeos.

Durante a catálise de uma reacção de fosforilação de amido F dá lugar a uma proteína R3 fosforilada como produto intermediário, por meio do qual é ligado covalentemente um resíduo fosfato do ATP a um aminoácido da proteína R3. O produto intermediário resulta da autofosforilação da proteína R3, isto é, a própria proteína R3 catalisa a reacção que conduz ao produto intermediário. A autofosforilação fosforila preferencialmente um resíduo de histidina da sequência de aminoácidos codificadora de uma proteína R3, com particular preferência um resíduo de

histidina que faça parte de um domínio fosfo-histidina (Tien-Shin Yu et al., 2001 , Plant Cell 13, 1907-1918).

Uma vez que, para além da R1, não foram descritas enzimas que fosforilam o amido, não foi possível identificar o aumento da concentração de fosfato de amido em plantas com a ajuda de outras enzimas. Esta operação só é possível através da utilização de uma proteína de acordo com a presente invenção ou uma molécula de ácido nucleico de acordo com a invenção para modificação genética de plantas. Uma vez que, para além da R1, não foram descritas enzimas que fosforilam o amido, não foi possível identificar o aumento da concentração de fosfato de amido em plantas com a ajuda de outras enzimas.

No contexto da presente invenção, o termo "fosfato de amido" pressupõe grupos fosfato com ligação covalente às moléculas de glucose do amido.

A actividade de uma proteína R3 pode ser identificada, por exemplo, por incubação de uma proteína R3 *in vitro* utilizando nucleosídeo trifosfatos que incluem um resíduo de fosfato marcado (nucleosídeo trifosfato marcado). Nucleosídeos trifosfato marcados radioactivamente são preferencialmente utilizados para este fim.

Resíduos de fosfato marcado que foram incorporados em amido por uma proteína R3, podem ser identificados, por exemplo, através da separação do amido marcado (por exemplo, através da centrifugação, de precipitação de etanol, filtração, métodos de cromatografia, etc) do resíduo da mistura reaccional e subsequente detecção do resíduo de fosfato marcado na fracção amido. Neste contexto, os resíduos de fosfato marcados ligados na

fracção amido podem ser identificado, por exemplo por meio de determinação da quantidade de radioactividade existente na fracção amido (por exemplo, por meio de um contador de cintilações). São abaixo descritos, em Métodos Gerais Item 3, os métodos de detecção possíveis de uma proteína que necessite de amido como substrato para uma reacção de fosforilação. As posições dos átomos de carbono (C-2, C-3 ou C-6) dos monómeros da glucose no amido, preferencialmente fosforilados por uma proteína R3, podem ser determinados, por exemplo, pela análise dos amidos F que foram fosforilados por uma proteína tal como descrito em Ritte et al. (2002, PNAS 99, 7166-7171). Com este objectivo, um amido fosforilado por uma proteína é hidrolisado utilizando um ácido e posteriormente analisado por meio de cromatografia de permuta de aniões.

De preferência, o amido fosforilado por uma proteína R3 é analisado preferencialmente por meio de RMN 31-P ou 33P, de modo a estabelecer as posições dos átomos de carbono (C-2, C-3 ou C-6) dos monómeros de glucose no amido F que são fosforiladas. Um método particularmente preferido de identificação das posições dos átomos C de uma molécula de glucose de um amido, que são fosforiladas por uma reacção catalisada por uma proteína R3, é descrito mais à frente, em Métodos Gerais, ponto 4.

A sequência de aminoácidos revelada na SEQ ID NO 2 codifica a proteína R3 da *Arabidopsis thaliana*. A sequência de aminoácidos revelada na SEQ ID NO 4 codifica uma proteína R3 de *Arabidopsis thaliana*, a sequência de aminoácidos revelada na SEQ ID NO 2 em fusão (traducional) com um peptídeo sinal plastídico, dando lugar a uma sequência de aminoácidos recombinante que codifica uma

proteína R3, a qual, para além da sua sequência de aminoácidos codificadora, tem um peptídeo sinal plastídico.

Numa outra concretização da presente invenção, as sequências de aminoácidos codificadoras de proteínas R3 possuem uma identidade de pelo menos 60%, em particular de pelo menos 70%, de preferência de pelo menos 80% e particularmente preferido pelo menos 90% e especialmente preferido pelo menos 95% com a sequência especificada na SEQ ID N° 2 ou SEQ ID N° 4.

Numa outra concretização da presente invenção, a proteína R3 possui um domínio fosfo-histidina (Tien-Shin Yu et al., 2001, Plant Cell 13, 1907-1918). Sequências de aminoácidos codificadoras da proteína R3 que possui um domínio fosfo-histidina com pelo menos 85%, de preferência pelo menos 88%, com especial preferência pelo menos 91% e com particular preferência pelo menos 94% e com especial preferência pelo menos 97% de identidade com a sequência de aminoácidos do domínio fosfo-histidina da proteína R3 de *Arabidopsis thaliana* da SEQ ID NO 5.

Uma outra concretização da presente invenção refere-se a uma célula vegetal geneticamente modificada de acordo com a invenção ou a uma planta geneticamente modificada de acordo com a invenção, em que a modificação genética consiste na introdução de pelo menos uma molécula de ácido nucleico estranha no genoma da célula vegetal ou no genoma da planta.

Neste contexto, o termo "modificação genética" significa a introdução de moléculas de ácido nucleico exógenas homólogas e/ou heterólogas no genoma de uma célula vegetal ou no genoma de uma planta, em que a dita

introdução destas moléculas conduz a um aumento da actividade de uma proteína R3. As células vegetais, de acordo com a invenção, ou as plantas, de acordo com a invenção, são modificadas no que se refere à sua informação genética pela introdução de uma molécula de ácido nucleico estranha. A presença ou a expressão de uma molécula de ácido nucleico conduz a uma modificação fenotípica. Neste caso modificação "fenotípica" significa de preferência uma modificação mensurável de uma ou mais funções das células. Por exemplo, as células vegetais geneticamente modificadas, de acordo com a invenção e as plantas geneticamente modificadas de acordo com a invenção exibem um aumento da actividade de uma proteína R3 devido à presença ou expressão da molécula de ácido nucleico introduzida.

No contexto da presente invenção, o termo "molécula de ácido nucleico estranha" pressupõe significar uma molécula que ou não ocorre naturalmente nas células vegetais de tipo selvagem ou não ocorre naturalmente na disposição espacial concreta em células vegetais de tipo selvagem ou que se encontra localizada no ponto no genoma da célula vegetal de tipo selvagem onde não ocorre naturalmente. De preferência, a molécula de ácido nucleico estranha é uma molécula recombinante que consiste em diferentes elementos, não ocorrendo esta combinação ou disposição espacial específica naturalmente em células vegetais.

Em princípio, uma molécula de ácido nucleico estranha pode ser qualquer molécula de ácido nucleico que provoque um aumento da actividade de uma proteína R3 na célula vegetal ou planta. No contexto da presente invenção, o termo "genoma" pressupõe significar a totalidade do material genético presente na célula vegetal. O

especialista tem conhecimento que não apenas o núcleo, mas também outros compartimentos (por exemplo, plastídeos, mitocôndria) incluem material genético.

Numa outra concretização, as células vegetais de acordo com a presente invenção e as plantas de acordo com a presente invenção compreendem a molécula de ácido nucleico estranha codificadora de uma proteína R3, preferencialmente uma proteína R3 *Arabidopsis thaliana*.

Noutra concretização, a molécula de ácido nucleico estranha codifica uma proteína R3 com a sequência de aminoácidos especificada na SEQ ID N° 2 ou na SEQ ID N° 4.

Existe um grande número de técnicas disponíveis para a introdução de ADN na célula hospedeira vegetal. Estas técnicas incluem a transformação de células vegetais com T-ADN, utilizando *Agrobacterium tumefaciens* ou *Agrobacterium rhizogenes* como meio de transformação, a fusão de protoplastos, injeção, electroporação de ADN, introdução de ADN por meio de uma abordagem biolística bem como outras possibilidades. A utilização da transformação por meio de agrobacteria de células vegetais tem sido objecto de um estudo intensivo e foi devidamente descrita em EP 120516; Hoekema, IN: *The Binary Plant Vector System* Offsetdrukkerij Kanters B.V. Alblasserdam (1985), Chapter V; Fraley et al., *Crit. Rev. Plant Sci.* 4, 1-46 e in An et al. *EMBO J.* 4, (1985), 277-287. Relativamente à transformação da batata, ver, por exemplo, Rocha-Sosa et al., *EMBO J.* 8, (1989), 29-33).

A transformação de plantas monocotiledóneas por meio de vectores baseados na transformação de *Agrobacterium* foi igualmente descrita (Chan et al., *Plant MoI. Biol.* 22,

(1993), 491-506; Hiei et al., *Plant J.* 6, (1994) 271-282; Deng et al., *Science in China* 33, (1990), 28-34; Wilmink et al., *Plant Cell Reports* 11, (1992), 76-80; May et al., *Bio/Technology* 13, (1995), 486-492; Conner and Domisse, *Int. J. Plant Sci.* 153 (1992), 550-555; Ritchie et al., *Transgenic Res.* 2, (1993), 252-265). Um sistema alternativo à transformação de plantas monocotiledóneas consiste na transformação por meio da abordagem biolística (Wan and Lemaux, *Plant Physiol.* 104, (1994), 37-48; Vasil et al., *Bio/Technology* 11 (1993), 1553-1558; Ritala et al., *Plant Mol. Biol.* 24, (1994), 317-325; Spencer et al., *Theor. Appl. Genet.* 79, (1990), 625-631 ), transformação de protoplastos, electroporação de células parcialmente permeabilizadas e a introdução de ADN por meio de fibras de vidro. Em particular, tem sido muitas vezes descrita na literatura a transformação de milho (cf. por exemplo WO95/06128, EP0513849, EP0465875, EP0292435; Fromm et al., *Biotechnology* 8, (1990), 833- 844; Gordon-Kamm et al., *Plant Cell* 2, (1990), 603-618; Koziel et al., *Biotechnology* 11 (1993), 194-200; Moroc et al., *Theor. Appl. Genet.* 80, (1990), 721-726).

Esta transformação com êxito de outros tipos de cereal foi também já descrita, por exemplo quanto à cevada (Wan and Lemaux, ver acima; Ritala et al., ver acima; Krens et al., *Nature* 296, (1982), 72-74) e quanto ao trigo (Nehra et al., *Plant J.* 5, (1994), 285-297; Becker et al., 1994, *Plant Journal* 5, 299-307). Todos os métodos apresentados anteriormente são adequados no âmbito da presente invenção.

Entre outros aspectos, as células vegetais e plantas que foram geneticamente modificadas através da introdução de uma proteína R3 podem ser distinguidas das células

vegetais de tipo selvagem e plantas de tipo selvagem respectivamente, na medida em que contém uma molécula de ácido nucleico estranha que não ocorre naturalmente nas células vegetais de tipo selvagem ou plantas de tipo selvagem, ou na medida em que uma molécula destas se encontra presente integrada num ponto do genoma da célula vegetal de acordo com a invenção ou no genoma da planta de acordo com a invenção onde não ocorre nas células vegetais de tipo selvagem ou nas plantas de tipo selvagem, isto é, num ambiente genómico diferente. Além disso, as células vegetais de acordo com a invenção e as plantas de acordo com a invenção podem ser distinguidas das células vegetais de tipo selvagem e plantas de tipo selvagem respectivamente por conterem pelo menos uma cópia da molécula de ácido nucleico integrada de forma estável dentro do seu genoma, possivelmente para além das cópias naturais desta molécula nas células vegetais de tipo selvagem ou plantas de tipo selvagem. Se a(s) molécula(s) de ácido nucleico estranha(s), introduzidas nas células vegetais de acordo com a invenção ou nas plantas de acordo com a invenção é/são cópias adicionais das moléculas que já ocorrem naturalmente nas células vegetais de tipo selvagem ou nas plantas de tipo selvagem respectivamente, então as células vegetais de acordo com a invenção e as plantas de acordo com a invenção podem ser distinguidas das células vegetais de tipo selvagem ou das plantas de tipo selvagem respectivamente, em particular na medida em que esta cópia adicional ou estas cópias adicionais está/estão em pontos do genoma onde não ocorre(m) nas células vegetais de tipo selvagem ou plantas de tipo selvagem. Este facto pode ser verificado, por exemplo, com ajuda de uma análise *Southern*

*blot*. Além disso, as células vegetais de acordo com a invenção e as plantas de acordo com a invenção podem ser distinguidas das células de plantas de tipo selvagem e plantas de tipo selvagem respectivamente, preferencialmente por pelo menos uma das seguintes características: Se uma molécula de ácido nucleico estranha que foi introduzida é heteróloga no que se refere à célula vegetal ou planta, então as células vegetais de acordo com a invenção ou as plantas de acordo com a invenção têm transcrições das moléculas de ácido nucleico introduzidas. Este facto pode ser verificado, por exemplo, por análise *Northern blot* ou por RT-PCR (Reverse Transcription Polymerase Chain Reaction - reacção em cadeia de polimerase - transcrição reversa). As células vegetais de acordo com a invenção e as plantas de acordo com a invenção, que expressam uma transcrição antisense e/ou ARNi podem ser detectadas, por exemplo com ajuda de sondas de ácido nucleico específicas, complementares ao ARN (que ocorre naturalmente na célula vegetal), que codifica a proteína. De preferência, as células vegetais de acordo com a invenção e as plantas de acordo com a invenção contêm uma proteína que é codificada por uma molécula de ácido nucleico introduzida. Este facto pode ser verificado, por exemplo, por métodos imunológicos, em particular pela análise *Western blot*.

Se a molécula de ácido nucleico exógena que foi introduzida é homóloga no que se refere à célula vegetal ou planta, então as células vegetais de acordo com a invenção ou as plantas de acordo com a invenção podem ser diferenciadas das células vegetais de tipo selvagem ou plantas de tipo selvagem respectivamente, por exemplo, com base na expressão adicional das moléculas de ácido nucleico

estranhas introduzida. De preferência, as células vegetais de acordo com a invenção e as plantas de acordo com a invenção contêm transcrições das moléculas de ácido nucleico estranhas. Este aspecto pode ser detectado, por exemplo, por análise de *Northern blot* ou com ajuda da análise conhecida como PCR quantitativo.

Noutra concretização, as células vegetais de acordo com a invenção e as plantas de acordo com a invenção assumem a forma de, respectivamente, células vegetais transgénicas e plantas transgénicas. Numa outra concretização, a presente invenção refere-se a células vegetais de acordo com a invenção e plantas de acordo com a invenção, em que a molécula de ácido nucleico estranha é seleccionada do grupo constituído por

a) moléculas de ácido nucleico que codificam uma proteína com a sequência de aminoácidos especificada na SEQ ID N° 2 ou na SEQ ID N° 4;

b) moléculas de ácido nucleico codificadoras de uma proteína que inclui a sequência de aminoácidos, codificada pela inserção no plasmídeo pIR138-210 ou a inserção no plasmídeo pIR139-210;

c) moléculas de ácido nucleico que codificam uma proteína com pelo menos 60 % de identidade com a sequência de aminoácidos especificada na SEQ ID N° 2 ou na SEQ ID N° 4;

d) moléculas de ácido nucleico codificadoras de uma proteína cuja sequência tem uma identidade de pelo menos 60% com a sequência de aminoácidos, codificada pela região codificadora da inserção no plasmídeo pIR138-210 ou pela inserção no plasmídeo pIR139-210;

e) moléculas de ácido nucleico que incluem a sequência de nucleótidos da SEQ ID N° 1 ou da SEQ ID N° 3 ou uma sequência complementar;

f) moléculas de ácido nucleico que compreendem a sequência de nucleótidos ou a inserção presente no plasmídeo pIR138-210 ou no plasmídeo pIR139-210;

g) moléculas de ácido nucleico que têm uma identidade de pelo menos 60% com as sequências de ácido nucleico descritas em a), b), e) ou f);

h) moléculas de ácido nucleico que hibridam com pelo menos uma cadeia das moléculas de ácido nucleico descritas em a), b), d), e) ou f) sob condições restritivas;

i) moléculas de ácido nucleico cuja sequência de nucleótidos se desvia da sequência de moléculas de ácido nucleico mencionadas em a), b), e) ou f), em resultado da degenerescência do código genético e

j) moléculas de ácido nucleico que constituem fragmentos, variantes alélicas e/ou derivados das moléculas de ácido nucleico descritas em a), b), c), d), e), f), g), h) ou i).

A sequência de aminoácidos da SEQ ID NO 2 codifica uma proteína R3 de *Arabidopsis thaliana*. As proteínas codificadas pelas diferentes variantes das moléculas de ácido nucleico de acordo com a presente invenção partilham certas características. Estas podem incluir, por exemplo, actividade biológica, peso molecular, reactividade imunológica, conformação e semelhantes e propriedades físicas tal como, por exemplo, o comportamento de migração em electroforese em gel, comportamento cromatográfico,

coeficientes de sedimentação, solubilidade, propriedades espectroscópicas, estabilidade, nível óptimo de pH, temperatura óptima e semelhantes. O peso molecular da proteína R3 de *Arabidopsis thaliana* deduzida da sequência de aminoácidos da SEQ ID NO 2 é de aproximadamente 131 kDa. O peso molecular de uma proteína de acordo com a invenção deduzido da sequência de aminoácidos codificadora encontra-se, assim, no intervalo entre 120 kDa a 145 kDa, de preferência entre 120 kDa e 140 kDa, com especial preferência entre 125 kDa e 140 kDa, particularmente preferencial entre 130 kDa e 135 kDa. A sequência de aminoácidos da SEQ ID NO 2 que codifica uma proteína R3 de *Arabidopsis thaliana* inclui um domínio fosfo-histidina. Por conseguinte, uma proteína R3 de acordo com a invenção compreende preferencialmente um domínio fosfo-histidina que possui pelo menos 85%, de preferência pelo menos 88%, com especial preferência pelo menos 91% e com particular preferência pelo menos 95% de identidade com o domínio fosfo-histidina da SEQ ID NO 5.

A presente invenção refere-se a moléculas de ácido nucleico que codificam uma proteína com a actividade enzimática de acordo com a invenção de uma proteína R3, possuindo a proteína R3 codificada pelo menos 70%, de preferência pelo menos 80%, com especial preferência pelo menos 90% e com particular preferência pelo menos 95% de identidade com as sequências de aminoácidos reveladas na SEQ ID NO 2.

Uma estirpe *Escherichia coli* compreendendo o plasmídeo pIR138-210, compreendendo um cADN codificador de uma proteína R3 de *Arabidopsis thaliana*, que está em fusão traducional com uma sequência de ácidos nucleicos

codificadora do peptídeo de sinal plastídeo do gene *dul 1* de *Oryza sativa* foi depositada na Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH, Mascheroder Weg 1b, 38124 Braunschweig, Alemanha, em conformidade com as disposições do Tratado de Budapest de 21 de Julho de 2004 com o número DSM 16587. A sequência de aminoácidos apresentada na SEQ ID N° 4 pode ser derivada da região codificadora da sequência de cADN integrada no plasmídeo pIR138-210 e codifica uma proteína R3 de *Arabidopsis thaliana*, que possui um peptídeo de sinal plastídeo. Por conseguinte, a presente invenção refere-se a moléculas de ácido nucleico, codificadoras de uma proteína com a actividade enzimática de uma proteína R3, que compreende a sequência de aminoácidos codificada pela inserção no plasmídeo pIR138-210, possuindo a proteína codificada pelo menos 70%, de preferência pelo menos 80%, com especial preferência 90% e com particular preferência 95% de identidade com a sequência de aminoácidos que pode ser deduzida da inserção em pIR138-210.

O plasmídeo pIR139-210, que compreende o cADN codificador de uma proteína R3 de *Arabidopsis thaliana* foi depositado junta da Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH, Mascheroder Weg 1b, 38124 Braunschweig, Alemanha, em conformidade com as disposições do Tratado de Budapest de 10 de Agosto de 2004, com o número DSM 16645. A sequência de aminoácidos apresentada na SEQ ID N° 2 pode ser derivada da região codificadora da sequência de cADN integrada no plasmídeo pIR139-210 e codifica uma proteína R3 de *Arabidopsis thaliana*. Por conseguinte, a presente invenção refere-se também a moléculas de ácido nucleico, codificadoras de uma proteína

com a actividade enzimática de uma proteína R3, que compreende a sequência de aminoácidos codificada pela inserção no plasmídeo pIR139-210, possuindo a proteína codificada pelo menos 70%, de preferência pelo menos 80%, com especial preferência 90% e com particular preferência 95% de identidade com a sequência de aminoácidos que pode ser deduzida da inserção em pIR139-210.

A sequência de ácido nucleico da SEQ ID NO 1 é uma sequência de cADN que compreende a região codificadora de uma proteína R3 de *Arabidopsis thaliana*.

Por conseguinte, a presente invenção refere-se também a moléculas de ácido nucleico que codificam uma proteína R3 e incluem a região codificadora das sequências de nucleótidos da SEQ ID N° 1 ou sequências complementares desta, moléculas de ácido nucleico, que incluem a região codificadora da sequência de nucleótidos da inserção contida no plasmídeo pIR139-210 e moléculas de ácido nucleico que possuem uma identidade de pelo menos 70%, de preferência pelo menos 80%, particularmente preferido pelo menos 90% e especialmente preferido pelo menos 95% com as moléculas de ácido nucleico mencionadas. A sequência de ácidos nucleicos da SEQ ID NO 3 é uma sequência de cADN que inclui a região codificadora de uma proteína R3 de *Arabidopsis thaliana*, que está em fusão traducional com a sequência de ácido nucleico codificadora do peptídeo de sinal plastídeo do gene *dul I* de *Oryza sativa*. A presente invenção refere-se também a moléculas de ácido nucleico que codificam uma proteína R3 e incluem a região codificadora das sequências de nucleótidos da SEQ ID N° 3 ou sequências complementares desta, moléculas de ácido nucleico, que incluem a região codificadora da sequência de nucleótidos

da inserção contida no plasmídeo pIR138-210 e moléculas de ácido nucleico que possuem uma identidade de pelo menos 70%, de preferência pelo menos 80%, particularmente preferido pelo menos 90% e especialmente preefrido pelo menos 95% com as moléculas de ácido nucleico mencionadas.

O especialista pode isolar as sequências homólogas de outras espécies vegetais com auxílio da informação de sequência das moléculas de ácido nucleico de acordo com a invenção ou com auxílio de uma molécula de ácido nucleico de acordo com a invenção. Esta operação pode ser efectuada, por exemplo com a ajuda de métodos convencionais, tal como o exame de bibliotecas de cADN ou bibliotecas genómicas com sondas de hibridação adequadas. Os especialistas na técnica sabem que podem também ser isoladas sequências homólogas com a ajuda de oligonucleótidos (degenerados) e utilização de métodos à base de PCR. O exame de bases de dados, tal como as disponibilizadas, por exemplo, pela EMBL (<http://www.ebi.ac.uk/Tools/index.htm>) ou NCBI (National Center for Biotechnology Information, <http://www.ncbi.nlm.nih.gov/>), pode também servir para identificar sequências homólogas que codificam proteínas R3.

Nestes exames, introduz-se uma ou várias sequências, realizando o que é conhecido por pesquisa. Esta sequência de pesquisa é então comparada por meio de programas informáticos estatísticos com sequências que estão contidas nas bases de dados seleccionadas. Estas pesquisas da base de dados (por exemplo pesquisas blast e fasta) são conhecidas dos especialistas na técnica e podem ser efectuadas por vários fornecedores. Caso for efectuada uma pesquisa da base de dados, por exemplo junto da NCBI

(National Center for Biotechnology Information, <http://www.ncbi.nlm.nih.gov/>), devem ser utilizadas as definições padrão que são especificadas para esta pesquisa comparativa específica. Para comparações de sequências de proteínas (blastp), estas definições são as seguintes: Limit entrez = not activated; Filter = low complexity activated; Expect value = 10; word size = 3; Matrix = BLOSUM62; Gap costs: Existence = 11, Extension = 1. Deverão utilizar-se os seguintes parâmetros para a comparação de sequências de ácido nucleico (blastn): Limit entrez = not activated; Filter = low complexity activated; Expect value = 10; word size = 11.

Por exemplo, no caso de uma pesquisa da base de dados deste género, as sequências descritas na presente invenção podem ser utilizadas como sequência de pesquisa, para identificar mais moléculas de ácido nucleico e/ou proteínas codificadoras de uma proteína R3. Com a ajuda dos métodos descritos é ainda possível identificar e/ou isolar moléculas de ácido nucleico de acordo com a invenção que hibridam com a sequência especificada na SEQ ID N° 1 e que codificam uma proteína R3. Para os objectivos da presente invenção, o termo "hibridação" significa hibridação nas condições de hibridação convencionais, de preferência em condições restritivas, por exemplo, tal como descritas em Sambrock et al. (Molecular Cloning, A Laboratory Manual, 3ª edição (2001) Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.I.). ISBN: 0879695773, Ausubel et al., Short Protocols in Molecular Biology, John Wiley & Sons; 5ª edição (2002), ISBN: 0471250929). Com particular preferência, "hibridação" significa hibridação nas seguintes condições: Tampão de hibridação: 2×SSC;

10×solução de Denhardt (Ficoll 400+PEG+BSA; proporção 1:1:1); 0,1% SDS; 5 mM EDTA; 50 mM Na<sub>2</sub>HPO<sub>4</sub>; 250 µg/ml ADN de esperma de arenque; 50 µg/ml tARN; ou 25 M tampão de fosfato de sódio pH 7,2; 1 mM EDTA; 7% SDS

Tampão de hibridação: T = 65° a 68°C

Tampão de lavagem: 0,1×SSC; 0,1% SDS

Temperatura de lavagem: T=65 a 68 °C.

Em princípio, as moléculas de ácido nucleico que hibridam com molécula de ácido nucleico de acordo com a invenção podem ser derivaads de qualquer espécie vegetal. As moléculas de ácido nucleico que hibridam com as moléculas de acordo com a invenção podem, por exemplo, ser isoladas a partir de bibliotecas genómicas ou bibliotecas de cADN. A identificação e isolamento de moléculas de ácido nucleico deste tipo podem ser efectuadas de acordo com a invenção, ou partes destas moléculas ou os complementos reversos destas moléculas, por exemplo por meio de hibridação de acordo com os métodos padrão (ver, por exemplo Sambrook et al., Molecular Cloning, A Laboratory Manual, 3ª edição Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.I. ISBN: 0879695773, Ausubel et al., Short Protocols in Molecular Biology, John Wiley & Sons 5ª edição (2002), ISBN: 0471250929) ou por amplificação por meio de PCR. Exemplos de sondas de hibridação que podem ser utilizadas são as moléculas de ácido nucleico com precisamente ou essencialmente a sequência de nucleótidos indicada na SEQ ID NO 1, ou partes destas sequências. Estes fragmentos utilizados como sondas de hibridação podem também ser fragmentos sintéticos ou oligonucleótidos que foram produzidos utilizando técnicas de síntese

estabelecidas e cuja sequência corresponde essencialmente à de uma molécula de ácido nucleico de acordo com a presente invenção. Assim que os genes que hibridam com as sequências de ácido nucleico de acordo com a invenção foram identificados e isolados, dever ser sequenciados e as características das proteínas codificadas por esta sequência devem ser analisadas para confirmar se é uma proteína R3. As comparações de homologia ao nível da sequência de ácido nucleico ou de aminoácidos e uma determinação da actividade enzimática são métodos particularmente adequados para este fim. A actividade de uma proteína R3 pode ser determinada, por exemplo como anteriormente descrito em Métodos Gerais, Ponto 2. As moléculas que hibridam com as moléculas de ácido nucleico de acordo com a invenção incluem em particular fragmentos, derivados e variantes alélicas das moléculas de ácido nucleico de acordo com a invenção que codificam uma proteína R3 de plantas, em particular uma proteína R3 de plantas de *Arabidopsis thaliana*. No contexto da presente invenção, o termo "derivado" significa que as sequências destas moléculas divergem em uma ou mais do que uma posições das sequências das moléculas de ácido nucleico descritas anteriormente e possuem um alto grau de identidade com estas sequências. Os desvios das moléculas de ácido nucleico descritas anteriormente, pode ter ocorrido, por exemplo devido a eliminação, adição, substituição, inserção ou recombinação.

No contexto da presente invenção, o termo "identidade" significa uma identidade de sequência em todo o comprimento da região codificadora (sequência integral) de pelo menos 70%, em particular uma identidade de pelo menos 80%, de

preferência maior que 85%, particularmente preferível maior que 90% e especialmente de pelo menos 95%. No contexto da presente invenção, o termo "identidade" pressupõe significar o número de aminoácidos/nucleótidos correspondente (identidade) a outras proteínas/ácidos nucleicos expressos em percentagem. De preferência, a identidade é determinada por comparação da SEQ ID NO 2, no caso de aminoácidos, ou SEQ ID NO 1, no caso de ácidos nucleicos, com outras proteínas/ácidos nucleicos, com ajuda de programas informáticos. Se as sequências que são comparadas entre si possuem comprimentos diferentes, a identidade deve ser determinada de forma a que o número de aminoácidos que têm a sequência mais curta em comum com a sequência mais longa determine o quociente de percentagem da identidade. De preferência, a identidade é determinada por meio do programa informático ClustalW, que é bem conhecido e disponível ao público (Thompson et al., *Nucleic Acids Research* 22 (1994), 4673-4680). ClustalW é disponibilizado ao público por Julie Thompson (Thompson@EMBL-Heidelberg.DE) e Toby Gibson (Gibson@EMBL-Heidelberg.DE), European Molecular Biology Laboratory, Meyerhofstrasse 1, D 69117 Heidelberg, Alemanha. Igualmente, ClustalW pode ainda ser descarregado de vários *sites* na Internet, incluindo IGBMC (Institut de Génétique et de Biologie Moléculaire et Cellulaire, B.P.163, 67404 Illkirch Cedex, France; <ftp://ftp-igbmc.u-strasbg.fr/pub/>) e o EBI (<ftp://ftp.ebi.ac.uk/pub/software/>) bem como em todos os *sites* na Internet espelhados do EBI (European Bioinformatics Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SD, Reino Unido). Para determinar a identidade entre proteínas de acordo com a invenção e

outras proteínas é preferida a versão 1.8 do programa informático ClustalW. Devem ser definidos os seguintes parâmetros: KTUPLE=I , TOPDIAG=5, WINDOW=5, PAIRGAP=3, GAOPEN=IO, GAPEXTEND=0.05, GAPDIST=8, MAXDIV=40, MATRIX=GONNET, ENDGAPS(OFF), NOPGAP, NOHGAP. Para determinar a identidade entre, por exemplo, sequência de nucleótidos das moléculas de ácido nucleico de acordo com a invenção e a sequência de nucleótidos de outras moléculas de ácido nucleico utiliza-se de preferência, a versão 1.8 do programa informático ClustalW. Devem ser definidos os seguintes parâmetros: KTUPLE=2, TOPDIAGS=4, PAIRGAP=5, DNAMATRIXMUB, GAOPEN=IO, GAPEXT=5, MAXDIV=40, TRANSITIONS: imponderado Além disso, identidade significa que existe equivalência funcional e/ou estrutural entre as moléculas de ácido nucleico em questão ou as proteínas codificadas por estas. As moléculas de ácido nucleico que são homólogas às moléculas descritas anteriormente e que constituem derivados destas moléculas assumem geralmente a forma de variações destas moléculas, que constituem modificações que executam a mesma função biológica. Podem assumir a forma de variações naturais, por exemplo, sequências de outras espécies vegetais ou mutações, sendo possível que estas mutações tenham ocorrido naturalmente ou tenham sido introduzidas por mutagénese. Além disso, as variações podem assumir a forma de sequências geradas sinteticamente. As variantes alélicas tanto podem ser variantes naturais como variantes produzidas sinteticamente ou variantes produzidas por técnicas de ADN recombinante. Constituem uma forma especial de derivados, por exemplo, as moléculas de ácido nucleico que derivam das moléculas de ácido nucleico de

acordo com a invenção devido a degeneração do código genético.

As proteínas codificadas pelos diferentes derivados das moléculas de ácido nucleico de acordo com a presente invenção partilham certas características. Estas podem incluir, por exemplo, actividade biológica, especificidade do substrato, peso molecular, reactividade imunológica, conformação e semelhantes e propriedades físicas tal como, por exemplo, o comportamento de migração em electroforese em gel, comportamento cromatográfico, coeficientes de sedimentação, solubilidade, propriedades espectroscópicas, estabilidade, nível óptimo de pH, temperatura óptima e semelhantes. As propriedades preferidas de uma proteína R3 foram já pormenorizadas e aplicam-se analogamente neste caso.

As moléculas de ácido nucleico de acordo com a invenção podem ser quaisquer moléculas de ácido nucleico, em particular moléculas de ADN ou ARN, por exemplo cADN, ADN genómico, mARN, etc. Podem ser moléculas naturais ou moléculas produzidas por método de recombinação ou métodos de síntese química. Podem ser moléculas de cadeia simples contendo a cadeia codificadora ou não codificadora, como moléculas de cadeia dupla. A further embodiment of the present invention relates to plant cells according to the invention and plants according to the invention, the foreign nucleic acid molecule being selected from the group consisting of

a) moléculas T-ADN que conduzem a um aumento da expressão do gene R3 devido à integração no genoma da planta (endereçamento de activação T-ADN);

b) moléculas de ADN que incluem transposões, que conduzem a um aumento da expressão do gene R3 (endereçamento de activação transposição) devido à integração no genoma da planta;

c) moléculas de ADN que codificam uma proteína R3 e que estão ligadas a sequências reguladoras que garantem a transcrição em células vegetais e conduzem a um aumento numa actividade da proteína R3 na célula,

d) moléculas de ácido nucleico introduzidas por meio de mutagenese *in vivo* e que conduzem a uma mutação ou uma inserção de uma sequência heteróloga em pelo menos um gene endógeno, codificador de uma proteína R3, em que a mutação ou inserção provoca um aumento da expressão de um gene codificador de uma proteína R3.

No contexto da presente invenção, podem também ser produzidas células vegetais de acordo com a invenção e plantas de acordo com a invenção recorrendo à técnica conhecida por mutagenese de inserção (artigo de resumo: Thorneycroft et al., 2001, Journal of experimental Botany 52 (361), 1593-1601). No contexto da presente invenção, a mutagenese de inserção pressupõe significar particularmente a inserção de transposões ou ADN dito de transferência (T-ADN) num gene ou na proximidade de um gene codificador de uma proteína R3, em resultado do que aumenta a actividade de uma proteína R3 na célula em questão. Neste contexto, os transposões podem ambos ser aqueles que ocorrem naturalmente na célula (transposões endógenos) e ainda os que não ocorrem naturalmente na célula mencionada mas são introduzidos na célula por meio de métodos de recombinação, tal como transformação da célula, por exemplo (transposões

heterólogos). O especialista está familiarizado com a modificação da expressão de genes por meio de transposições. Encontra-se uma sinopse da utilização de transposições endógenos e heterólogos como ferramentas em biotecnologia vegetal em Ramachandran e Sundaresan (2001, *Plant Physiology and Biochemistry* 39, 234-252). A mutagénese de inserção T-ADN baseia-se no facto de certos segmentos (T-ADN) de plasmídeos Ti de *Agrobacterium* poderem ser integradas no genoma de células vegetais. O ponto de integração no cromossoma da planta não é definido, mas pode ocorrer em qualquer lugar. Se o T-ADN se integrar numa parte do cromossoma ou na vizinhança de uma parte do cromossoma que constitui uma função genética, então este facto pode conduzir a um aumento da expressão genética e também a uma alteração na actividade de uma proteína codificada pelo gene em questão. Neste caso, as sequências inseridas no genoma (particularmente transposições ou T-ADN) distinguem-se devido ao facto de conterem sequências que conduzem a uma activação das sequências reguladoras de um gene ("*activatin tagging*").

As células vegetais e plantas de acordo com a invenção podem ser produzidas por meio do método designado "*activatin tagging*" (ver, por exemplo, Walden et al., *Plant J.* (1991), 281-288; Walden et al., *Plant Mol. Biol.* 26 (1994), 1521-1528). Este método baseia-se na activação de promotores endógenos por meio de sequências intensificadoras, tal como o intensificador do promotor <sup>35S</sup> ARN do vírus do mosaico da couve-flor ou o intensificador da octopina-sintase.

No contexto da presente invenção, o termo "*T-ADN activation tagging*" pressupõe significar um fragmento de T-

ADN que contém sequências intensificadoras e que conduz a um aumento da actividade de pelo menos uma proteína R3 por integração no genoma de uma célula vegetal.

No contexto da presente invenção, o termo "activation tagging transposição" pressupõe significar um transposição que contém sequências intensificadoras e que conduz a um aumento da actividade de pelo menos uma proteína R3 por integração no genoma de uma célula vegetal.

Numa concretização adicional, moléculas de ADN de acordo com a invenção, que codificam uma proteína R3 e que estão ligadas a sequências reguladoras que iniciam a transcrição em células vegetais (promotores) e conduzem a um aumento numa actividade da proteína R3 na célula. Neste contexto, as moléculas de ácido nucleico de acordo com a invenção estão presentes na orientação "sense" relativamente às sequências reguladoras.

Para expressar moléculas de ácido nucleico que codificam uma proteína R3, estas moléculas estão ligadas preferencialmente a sequências de ADN reguladoras, que garantem a transcrição em células vegetais. Em particular, estas incluem promotores. Em geral, qualquer promotor que está activo em células vegetais é elegível para expressão. Neste contexto, o promotor pode ser escolhido, de forma que a expressão tem lugar de forma constitutiva ou apenas num determinado tecido, numa determinada etapa do desenvolvimento da planta ou num momento determinado por influências externas. O promotor pode ser homólogo ou heterólogo, ambos relativamente à planta e relativamente à molécula de ácido nucleico. Exemplos de promotores adequados são o promotor do vírus do mosaico da couve-flor

<sup>35</sup>S ARN e o promotor da ubiquitina do milho para expressão constitutiva, o promotor patatina B33 (Rocha-Sosa et al., EMBO J. 8 (1989), 23-29) para expressão específica de tuberosas em batatas ou um promotor que apenas garante a expressão em tecidos com actividade de fotossíntese, por exemplo o promotor ST-LS1 (Stockhaus et al., Proc. Natl. Acad. Sci. USA 84 (1987), 7943-7947; Stockhaus et al., EMBO J. 8 (1989), 2445-2451) ou para a expressão específica do endoesperma do promotor HMG do trigo, o promotor USP, o promotor faseolina, promotores dos genes da zeína do milho (Pedersen et al., Cell 29 (1982), 1015-1026; Quatroccio et al., Plant Mol. Biol. 15 (1990), 81-93), promotor glutelina (Leisy et al., Plant Mol. Biol. 14 (1990), 41-50; Zheng et al., Plant J. 4 (1993), 357-366; Yoshihara et al., FEBS Lett. 383 (1996), 213-218) ou o promotor Shrunken-1 (Werr et al., EMBO J. 4 (1985), 1373-1380). No entanto, podem também ser utilizados promotores que só são activados num momento determinado por influências externas (ver, por exemplo WO 9307279). Os promotores que podem revestir-se de interesse específicos neste contexto podem ser promotores de proteínas de choque térmico, as quais são simples de induzir. Além disso, podem ser utilizados promotores específicos da semente, tal como o promotor USP de *Vicia faba*, que garante a expressão específica da semente em *Vicia faba* e outras plantas (Fiedler et al., Plant Mol. Biol. 22 (1993), 669-679; Baumlein et al., Mol. Gen. Genet. 225 (1991), 459-467). Além disso, pode estar presente uma sequência de terminação (sinal de poliadenilação) que pode ser utilizada para adicionar uma cauda poli-A à transcrição. A cauda poli-A é apontada como possuindo uma função na estabilização das transcrições. São descritos

elementos deste tipo na literatura (cf. Gielen et al., EMBO J. 8 (1989), 23-29) e podem ser trocados facultativamente.

Podem também estar presentes sequências intrão entre o promotor e a região codificadora. Estas sequências intrão podem conduzir à expressão estável e a uma expressão aumentada em plantas (Callis et al., 1987, Genes Devel. 1 , 1183-1200; Luehrsen, and Walbot, 1991 , Mol. Gen. Genet. 225, 81-93; Rethmeier, et al., 1997; Plant Journal. 12(4):895- 899; Rose and Beliakoff, 2000, Plant Physiol. 122 (2), 535-542; Vasil et al., 1989, Plant Physiol. 91 , 1575-1579; XU et al., 2003, Science in China Series C vol. 46 No. 6, 561-569). As sequências intrão adequadas são, por exemplo, o primeiro intrão do gene *sh1* do milho, o primeiro intrão do gene da poliubiquitina 1 de milho, o primeiro intrão do gene EPSPS do arroz ou um dos dois primeiros intrões do gene PAT1 da *Arabidopsis*. As sequências intrão adequadas são conhecidas do especialista e extensamente descritas na literatura.

Além disso, as células vegetais de acordo com a invenção e plantas de acordo com a invenção podem ser geradas pelo método conhecido por "activação *in situ*", em que a modificação genética que foi introduzida acarreta uma modificação das sequências reguladoras dos genes R3 endógenos, o que conduz a uma expressão intensificada dos genes R3. A activação de um gene R3 é efectuada preferencialmente opr mutagénese "in vivo" de um promotor ou de sequências intensificadoras de um gene R3 endógeno. Neste contexto, um promotor ou uma sequência intensificadora, podem ser alterados através de mutagénese de forma a que a mutação produzida conduza a uma expressão acrescida de um gene R3 em células vegetais de acordo com a

invenção ou plantas de acordo com a invenção, em comparação com a expressão de um R3 em células vegetais de tipo selvagem ou plantas tipo selvagem. A mutação num promotor ou sequência intensificadora pode também conduzir à expressão de genes R3 em células vegetais de acordo com a invenção ou plantas de acordo com a invenção numa altura em que não seriam expressos em células vegetais tipo selvagem ou plantas de tipo selvagem. No contexto da presente invenção, o termo "gene R3" pressupõe significar uma molécula de ácido nucleico (cADN, DNA) que codifica uma proteína R3, de preferência uma proteína R3 de *Arabidopsis thaliana*.

Durante o processo conhecido por "mutagénesis *in vivo*" introduz-se um oligonucleótido híbrido ARN/ADN ("quimeroplasto") em células vegetais por meio de transformação de células vegetais (Kipp, P. B. et al., sessão poster no "5th International Congress of Plant Molecular Biology", 21 a 27 Setembro de 1997, Singapore; R. A. Dixon and C. J. Arntzen, relatório do encontro "Metabolic Engineering in Transgenic Plants", Keystone Symposia, Copper Mountain, Colo., USA, TIBTECH 15, (1997), 441-447; pedido de patente internacional WO 9515972; Kren et al., *Hepatology* 25, (1997), 1462-1468; Cole-Strauss et al., *Science* 273, (1996), 1386-1389; Beetham et al., 1999, *PNAS* 96, 8774-8778). Uma parte dos componentes do oligonucleótido ARN-ADN é homóloga de uma sequência de ácido nucleico de um gene R3 endógeno, mas, em comparação com a sequência de ácido nucleico de um gene R3 endógeno possui uma mutação ou contém uma região heteróloga que é cercada por regiões homólogas. Em resultado das regiões homólogas do oligonucleótido ARN-ADN e da molécula de ácido

nucleico endógena sujeitos a emparelhamento de bases, seguido de recombinação homóloga, a mutação ou a região heteróloga contida no componente ADN do oligonucleótido ARN-ADN pode ser transferida para o genoma de uma célula vegetal. Esta operação conduz a um aumento da actividade de uma ou mais do que uma proteína R3.

Todos esses métodos baseiam-se na introdução de uma molécula de ácido nucleico estranha no genoma de uma célula vegetal ou planta e, por conseguinte, são basicamente adequados para a produção de células vegetais de acordo com a invenção e plantas de acordo com a invenção.

Numa concretização preferida da presente invenção, as moléculas de ácido nucleico estranhas de acordo com a invenção assumem a forma de moléculas de ácido nucleico recombinantes, estando as moléculas de ácido nucleico de acordo com a invenção, codificadoras de uma proteína R3 ligadas a um peptídeo de sinal de tal forma que a transcrição e tradução das moléculas de ácido nucleico recombinantes mencionadas dão origem a uma proteína que possui um peptídeo de sinal plastídeo.

A presente invenção refere-se ainda, por conseguinte, a células vegetais geneticamente modificadas de acordo com a invenção ou plantas geneticamente modificadas de acordo com a invenção, em que a molécula de ácido nucleico estranha introduzida na célula vegetal ou planta assume a forma de uma molécula de ácido nucleico recombinante, sendo a molécula de ácido nucleico estranha uma molécula de ácido nucleico recombinante em que uma sequência de ácido nucleico codificadora da proteína R3 está fundida a sequências de ácido nucleico codificadoras de um peptídeo

de sinal plastídeo, de forma que a moléculade ácido nucleico recombinante codifique uma proteína R3 que possui um peptídeo de sinal plastídeo para além da sua sequência de aminoácidos codificadora.

No contexto da presente invenção, a expressão "molécula de ácido nucleico recombinante" subentende significar uma molécula de ácido nucleico que, para além das moléculas de ácido nucleico de acordo com a invenção, codificadoras de uma proteína R3, incluem sequências adicionais que não estão naturalmente presentes numa combinação em que estão presentes em ácidos nucleicos recombinantes de acordo com a invenção. Neste contexto, a adição das sequências mencionadas podem ser quaisquer sequências, de preferência com a forma de sequências reguladoras (promotoras, sinais de terminação, intensificadoras), com particular preferência a forma de sequências reguladora que estão activas em tecidos vegetais, com particular preferência sequências reguladoras que estão activas no tecido de plantas armazenadoras de amido. Sequências adicionais preferidas são as sequências que codificam peptídeos de sinal (sequências de sinal) que, em resultado da fusão traducional com sequências codificadoras da proteína, acarreta a translocação das proteínas de acordo com a invenção em compartimentos subcelulares (vacúolos, plastídeos, mitocôndrias, núcleos, retículos endoplasmáticos, parede celular). São especialmente importantes no contexto da presente invenção as sequências de sinal que codificam peptídeos de sinal plastídeos. Os especialistas na técnica conhecem métodos de produção de moléculas de ácido nucleico recombinantes de acordo com a invenção e incluem métodos de recombinação ou molecular-

biológicos, tal como, por exemplo, a união de moléculas de ácido nucleico por ligação, recombinação genética ou a síntese *de novo* de moléculas de ácido nucleico (ver, por exemplo Sambrook et al., *Molecular Cloning, A Laboratory Manual*, 3ª edição (2001) Cold Spring Harbour Laboratory Press, Cold Spring Harbour, N.Y. ISBN: 0879695773, Ausubel et al., *Short Protocols in Molecular Biology*, John Wiley & Sons 5ª edição (2002), ISBN: 0471250929).

No contexto da presente invenção, o termo "peptídeo de sinal plastídeo" subentende significar um peptídeo que acarreta a translocação de uma proteína para o lúmen de plastídeos quando o peptídeo de sinal faz parte da sequência de aminoácidos de uma proteína.

No que se refere à expressão em plantas das moléculas de ácido nucleico de acordo com a invenção, é possível, em princípio, que a proteína sintetizada possa estar localizada em qualquer compartimento da célula vegetal. Para conseguir a localização num compartimento específico, a região codificadora deve, se apropriado, estar ligada com sequências de ADN que garantem a localização no compartimento em questão. Estas sequências são conhecidas (ver, por exemplo, Braun, *EMBO J.* 11 (1992), 3219-3227; Sonnewald, *Plant J.* 1 (1991), 95-106; Rocha-Sosa, *EMBO J.* 8 (1989), 23-29, Neuhaus and Rodgers, 1998, *Plant Molecular Biology* 38: 127-144). Neste caso, as sequências de sinal são ligadas à região codificadora da proteína de tal forma que as sequências de ácido nucleico fundidas do peptídeo de sinal e as codificadoras da forma proteica, em caso de transcrição, uma estrutura de leitura aberta (fusão traducional). Deste modo, após a tradução do ARN formado pela transcrição das sequências de ácido nucleico fundidas,

é formada uma proteína que, para além da região codificadora da proteína, possui um peptídeo de sinal. Em princípio, qualquer sequência de sinal que garanta a importação de proteínas com codificação nuclear em plastídeos pode ser utilizada como sequência de sinal plastídica. O especialista sabe como obter as sequências de sinal plastídicas. Deste modo, por exemplo, é possível utilizar programas informáticos (1999, Emanuelsson et al., *Protein Science* 8, 978-984) para analisar sequências proteicas codificadoras para detectar a presença de uma sequência de sinal plastídica. Os programas informáticos que executam estas análises estão ambos disponíveis no comércio e ao público na internet (por exemplo, <http://www.cbs.dtu.dk/services/ChloroP/>). Por exemplo, é possível utilizar a sequência de sinal plastídica da ferredoxina do espinafre:NADP<sup>+</sup> oxidoreductase (FNR). Este sinal inclui a região não traduzida 5' e a sequência do peptídeo transit de flanco do cADN da proteína plastídica da ferredoxina:NADP<sup>+</sup> oxidoreductase (FNR) do espinafre (nucleótidos -171 a +165; Jansen et al., *Current Genetics* 13 (1988), 517-522). Além disso, é possível utilizar, por exemplo, o peptídeo transit da proteína cerosa do milho mais os 34 primeiros aminoácidos da proteína cerosa do milho madura (Klösgen et al., *Mol Gen Genet.* 217 (1989), 155-161) como sequência de sinal plastídeo. Além disso, é também possível utilizar o peptídeo transit da proteína cerosa do milho (ver acima) sem os 34 primeiros aminoácidos da proteína cerosa madura. A utilização das seguintes sequências de sinal plastídeo é também possível: sequência de sinal da subunidade pequena da ribulose bifosfato carboxilase (Wolter et al., *Proc. Natl. Acad.*

Sci. EUA 85 (1988), 846-850; Nawrath et al., Proc. Natl. Acad. Sci. EUA 91 (1994), 12760-12764); sequência de sinal de malato NADP desidrogenase (Gallardo et al., Planta 197 (1995), 324-332); sequência de sinal glutathiona reductase (Creissen et al., Plant J. 8 (1995), 167-175), sequência de sinal de EPSPS (US 5,188,642). É preferível utilizar a sequência de sinal do gene *dul* do arroz A molécula de ácido nucleico recombinante pode incluir uma ou mais, de preferência duas, sequências de sinal. As sequências de sinal podem estar presentes imediatamente umas a seguir às outras, no sentido da transcrição ou então podem ser separadas umas das outras pelo que é conhecido como espaçador. As sequências de sinal plastídeas que incluem duas sequências de sinal encontram-se descritas, por exemplo, na EP 0508909 e EP 0924299, US 5,510,471).

Surpreendentemente, constatou-se que as células vegetais de acordo com a invenção e plantas de acordo com a invenção sintetizam um amido modificado em comparação com amido das células vegetais de tipo selvagem ou plantas de tipo selvagem correspondentes que não foram geneticamente modificadas.

As células vegetais de acordo com a invenção e plantas de acordo com a invenção sintetizam um amido modificado cujas características físico-químicas, em particular cujo teor de fosfato de amido está modificado em comparação com o amido sintetizado em células vegetais de tipo selvagem ou plantas de tipo selvagem, sendo mais adequado para aplicações específicas. A presente invenção inclui, assim, células vegetais de acordo com a invenção e plantas de acordo com a invenção sintetizam um amido modificado em comparação com amido das células vegetais de tipo selvagem

ou plantas de tipo selvagem correspondentes que não foram geneticamente modificadas.

No contexto da presente invenção, o termo "amido modificado" significa que o amido sofreu alterações das características físico-químicas em comparação com amido não modificado, obtido de células vegetais de tipo selvagem ou plantas de tipo selvagem correspondentes.

Noutra concretização da presente invenção, as células vegetais de acordo com a invenção ou plantas de acordo com a invenção sintetizam um amido cujo teor de fosfato de amido é aumentado em comparação com o amido isolado a partir de células vegetais de tipo selvagem ou plantas de tipo selvagem correspondentes.

Foi descrita uma variedade de métodos para a determinação da quantidade de fosfato de amido. De preferência podem ser utilizados os métodos descritos por Ritte et al. (2000, *Starch/Starke* 52, 179-185) para a determinação da quantidade de fosfato de amido. A quantidade de fosfato de amido é efectuada de preferência por meio de  $^{31}\text{P}$ -NMR, utilizando os métodos descritos por Kasemusuwan e Jane (1996, *Cereal Chemistry* 73, 702-707).

A invenção refere-se ainda a plantas geneticamente modificadas, compreendendo células vegetais de acordo com a invenção. Estas plantas podem ser geradas a partir de células vegetais de acordo com a invenção por regeneração.

Em princípio, as plantas de acordo com a invenção podem ser plantas de qualquer espécie vegetal, isto é tanto monocotiledóneas como dicotiledóneas. De preferência são plantas úteis, isto é plantas que são cultivadas pelo homem

para fins alimentares ou técnicos, em particular industriais.

Noutra concretização, a planta de acordo com a invenção é uma planta armazenadora de amido. No contexto da presente invenção, o termo "plantas armazenadoras de amido" significa todas as plantas com partes que contêm um amido armazenado, tal como, por exemplo milho, arroz, trigo, centeio, aveia, cevada, mandioca, batata, sagú, feijão mungo, ervilha ou sorgo.

No contexto da presente invenção, o termo "batateira" ou "batata" significa espécies vegetais do genoma *Solanum*, em particular espécies produtoras de tuberosas do género *Solanum* e especialmente *Solanum tuberosum*.

No contexto da presente invenção, o termo "planta do trigo" significa espécies vegetais do genoma *Triticum* ou plantas resultantes de hibridações com plantas do género *Triticum*, particularmente espécies vegetais do género *Triticum* que são utilizadas em agricultura para fins comerciais ou plantas originárias de hibridações com plantas do género *Triticum*, com particular preferência *Triticum aestivum*.

No contexto da presente invenção, o termo "planta do milho" significa espécies vegetais do género *Zea*, em particular espécies vegetais do género *Zea*, que são utilizadas na agricultura para fins comerciais, com particular preferência para *Zea mais*.

Noutra outra concretização, a presente invenção refere-se a planta armazenadoras de amido da família de acordo com a invenção, de preferência batateiras, com especial preferência plantas da família (sistemática) das poáceas.

Estas plantas assumem especialmente de preferência a forma de plantas do milho ou trigo.

A presente invenção refere-se também a material de propagação das plantas de acordo com a invenção, contendo uma célula vegetal da invenção.

Neste caso, o termo "material de propagação" inclui os constituintes da planta que são adequados para a produção de descendência por meios vegetativos ou geradores. Os que são adequados à propagação de plantas são as estacas, culturas de calos, rizomas, tubérculos. Outro material de propagação inclui, por exemplo, frutos, sementes, rebentos, protoplastos, culturas de células, etc. O material de propagação é de preferência tubérculos e com especial preferência grãos com endoesperma. Noutra concretização, a presente invenção refere-se a partes de plantas que podem ser colhidas de plantas de acordo com a invenção, tal como frutos, raízes armazenadoras, raízes, flores, botões, rebentos ou hastes, de preferência sementes, grãos ou tubérculos, em que estas partes que podem ser colhidas contém células vegetais de acordo com a invenção.

A presente invenção refere-se também a um método de geração de plantas geneticamente modificadas de acordo com a invenção, em que

a) uma célula vegetal é geneticamente modificada, conduzindo a modificação genética ao aumento da actividade de uma proteína R3 em comparação com as células vegetais de tipo selvagem correspondentes, que não foram geneticamente modificadas.

b) uma planta é regenerada a partir de células vegetais da etapa a);

c) e, se apropriado, são produzidas mais plantas com ajuda das plantas de acordo com a etapa b).

Relativamente à modificação genética introduzida na célula vegetal de acordo com a etapa a), esta poderá em princípio assumir a forma de qualquer tipo de modificação que conduza a um aumento da actividade de uma proteína R3. As plantas podem ser regeneradas de acordo com a etapa b) recorrendo a métodos conhecidos dos especialistas na técnica (por exemplo descrito em "Plant Cell Culture Protocols", 1999, ed. Por R. D. Hall, Humana Press, ISBN 0-89603-549-2).

A geração de mais plantas de acordo com a etapa c) do método de acordo com a invenção pode ser efectuada, por exemplo, por meio de propagação vegetativa (por exemplo utilizando estacas, tubérculos ou por meio de cultura de calos e regeneração de plantas inteiras) ou por sexuada. Neste contexto, propagação sexuada ocorre preferencialmente em condições controladas, isto é plantas seleccionadas com características particulares são hibridadas e propagadas entre si. A selecção ocorre preferencialmente de modo a que outras plantas obtidas de acordo com a etapa c) incluam a modificação genética que foi introduzida na etapa a).

Numa concretização preferida do método de acordo com a invenção para a geração de uma planta geneticamente modificada, a modificação genética de uma célula vegetal na etapa a) conduz a um aumento da actividade de uma proteína R3 em plastídeos da célula vegetal mencionada, em comparação com plastídeos de células vegetais de tipo selvagem que não foram geneticamente modificadas correspondentes.

Noutra concretização do método de acordo com a invenção, a modificação genética consiste na introdução no genoma da célula vegetal de uma molécula de ácido nucleico de acordo com a invenção ou uma molécula de ácido nucleico recombinantes de acordo com a invenção, conduzindo a presença da molécula de ácido nucleico estranha mencionada ou molécula de ácido nucleico recombinante mencionada a uma maior actividade de uma proteína R3 na célula.

Noutra concretização do método de acordo com a invenção, a modificação genética consiste na introdução de uma molécula de ácido nucleico estranha no genoma da célula vegetal, codificando a molécula de ácido nucleico estranha uma proteína R3.

Noutra concretização do método de acordo com a invenção, a modificação genética consiste na introdução de uma molécula de ácido nucleico estranha no genoma da célula vegetal sendo a molécula de ácido nucleico estranha uma molécula de ácido nucleico recombinante, em que uma sequência de ácido nucleico codificadora da proteína R3 está fundida a uma segunda sequência de ácido nucleico codificadoras de um peptídeo de sinal plastídeo, de forma que a moléculade ácido nucleico recombinante codifique uma proteína R3 que possui um peptídeo de sinal plastídeo para além da sua sequência de aminoácidos codificadora.

Noutra concretização, o método de acordo com a invenção é utilizado para a geração de uma planta geneticamente modificada de acordo com a invenção para a geração de plantas armazenadora de amido.

Noutra concretização, o método de acordo com a invenção é utilizado para gerar plantas da batata, do milho ou trigo de acordo com a invenção.

Noutra concretização do método de acordo com a invenção, a molécula de ácido nucleico estranha e é selecionado do grupo constituído por

a) moléculas de ácido nucleico que codificam uma proteína com a sequência de aminoácidos especificada na SEQ ID N° 2 ou na SEQ ID N° 4;

b) moléculas de ácido nucleico codificadoras de uma proteína que inclui a sequência de aminoácidos, codificada pela inserção no plasmídeo pIR138-210 ou a inserção no plasmídeo pIR139-210;

c) moléculas de ácido nucleico que codificam uma proteína com pelo menos 60 % de identidade com a sequência de aminoácidos especificada na SEQ ID N° 2 ou na SEQ ID N° 4;

d) moléculas de ácido nucleico codificadoras de uma proteína cuja sequência tem uma identidade de pelo menos 60% com a sequência de aminoácidos, codificada pela inserção no plasmídeo pIR139-210;

e) moléculas de ácido nucleico que incluem a sequência de nucleótidos da SEQ ID N° 1 ou da SEQ ID N° 3 ou uma sequência complementar;

f) moléculas de ácido nucleic que compreendem a sequência de nucleótidos da inserção presente no plasmídeo pIR138-210 ou no plasmídeo pIR139-210;

g) moléculas de ácido nucleico que têm uma identidade de pelo menos 70% com as sequências de ácido nucleico descritas em a), b), e) ou f);

h) moléculas de ácido nucleico que hibridam com pelo menos uma cadeia das moléculas de ácido nucleico descritas em a), b), d), e) ou f) sob condições restritivas;

i) moléculas de ácido nucleico cuja sequência de nucleótidos se desvia da sequência de moléculas de ácido nucleico mencionadas em a), b), e) ou f), em resultado da degenerescência do código genético e

j) moléculas de ácido nucleico que constituem fragmentos, variantes alélicas e/ou derivados das moléculas de ácido nucleico descritas em a), b), c), d), e), f), g), h) ou i).

Noutra concretização do método de acordo com a invenção, a molécula de ácido nucleico estranha é seleccionada do grupo constituído por

a) moléculas T-ADN que conduzem a um aumento da expressão do gene R3 devido à integração no genoma da planta (endereço de activação T-ADN);

b) moléculas de ADN que incluem transposões, que conduzem a um aumento da expressão do gene R3 (endereço de activação transposição) devido à integração no genoma da planta;

c) moléculas de ADN que codificam uma proteína R3 e que estão ligadas a sequências reguladoras que garantem (iniciam) a transcrição em células vegetais e conduzem a um aumento numa actividade da proteína R3 na célula,

d) moléculas de ácido nucleico introduzidas por meio de mutagênese *in vivo* e que conduzem a uma mutação ou uma inserção de uma sequência heteróloga em pelo menos um gene endógeno R3, em que a mutação ou inserção provoca um aumento da expressão de um gene R3.

Noutra concretização do método de acordo com a invenção, a molécula de ácido nucleico de acordo com a invenção é uma molécula de ácido nucleico recombinante, sendo a molécula de ácido nucleico estranha uma molécula de ácido nucleico recombinante, em que uma sequência de ácido nucleico codificadora da proteína R3 está fundida a sequências de ácido nucleico codificadoras de um peptídeo de sinal plastídeo, de forma que a molécula de ácido nucleico recombinante codifique uma proteína R3 que possui um peptídeo de sinal plastídeo para além da sua sequência de aminoácidos codificadora.

Noutra concretização, a presente invenção refere-se a um método de acordo com a invenção, em que a planta geneticamente modificada sintetiza um amido que é modificado em comparação com plantas de tipo selvagem que não foram geneticamente modificadas.

Noutra concretização do método de acordo com a presente invenção, as plantas de acordo com a invenção sintetizam um amido cujo teor de fosfato de amido é aumentado em comparação com o amido isolado a partir de células vegetais de tipo selvagem ou plantas de tipo selvagem correspondentes.

A presente invenção refere-se ainda a plantas que podem ser obtidas por meio de métodos de acordo com a invenção.

Surpreendentemente constatou-se que o amido isolado a partir de células vegetais de acordo com a invenção e plantas de acordo com a invenção, possuidoras de uma actividade aumentada de uma proteína R3 nos plastídeos presentes, sintetizam um amido modificado. Em particular, a maior quantidade de fosfato de amido nos amidos de acordo com a invenção proporciona aos amidos características surpreendentes e vantajosas. Em resultado da maior quantidade de fosfato de amido, os amidos de acordo com a invenção suportam uma maior quantidade de grupos com cargas que afectam significativamente as características funcionais do amido. O amido que suporta os grupos funcionais carregados é particularmente aplicável na indústria do papel, onde é utilizado para o revestimento do papel. O papel revestido com moléculas carregadas, que possui adicionalmente boas propriedades adesivas é particularmente adequado para a absorção de cores, tal como, tais como, tinta, tintas de impressão e outras.

A presente invenção refere-se também a amidos modificados que são obtidos a partir de células vegetais de acordo com a invenção e plantas de acordo com a invenção, de material de propagação de acordo com a invenção ou de partes de plantas que se podem colher de acordo com a invenção.

Noutra concretização, a presente invenção refere-se a amido modificado de acordo com a invenção de planta armazenadoras de amido, de preferência da batateira, com especial preferência de plantas armazenadoras de amido da família (sistemática) das poáceas, com particular preferência das plantas do milho e do trigo.

A presente invenção refere-se ainda a um método de produção de um amido modificado, incluindo a etapa de extracção do amido a partir de uma célula vegetal de acordo com a invenção e plantas de acordo com a invenção, de material de propagação de acordo com a invenção destas plantas e/ou de partes de plantas que se podem colher desta planta de acordo com a invenção, de preferência de partes armazenadoras de amido de acordo com a invenção, desta planta. De preferência, este método inclui ainda a etapa de colheita das plantas cultivadas ou partes de plantas e/ou o material de propagação destas plantas antes da extração do amido e ainda, com particular preferência, a etapa de cultivo das plantas de acordo com a invenção antes da colheita.

Os métodos de extracção do amido a partir de plantas ou de plantas armazenadoras do amido são conhecidos dos especialistas na técnica. Além disso, os métodos de extracção do amido a partir de diferentes plantas armazenadoras de amido por exemplo em *Starch: Chemistry and Technology* (Editores: Whistler, BeMiller and Paschall (1994), 2ª edição, Academic Press Inc. London Ltd; ISBN 0-12-746270-8; ver por exemplo capítulo XII, página 412-468: *Maize and Sorghum Starches: Manufacture*; por Watson; capítulo XII, página 469-479: *Tapioca, Arrowroot and Sago Starches: Manufacture*; por Corbishley e Miller; capítulo XIV, página 479-490: *Potato starch: Manufacture and Uses*; por Mitch; capítulo XV, página 491 a 506: *Wheat starch: Manufacture, Modification and Uses*; por Knight e Oson; e capítulo XVI, página 507 a 528: *Rice starch: Manufacture and Uses*; por Rohmer e Klem; *Maize starch: Eckhoff et al., Cereal Chem.* 73 (1996), 54-57, a extracção do amido de

milho a uma escala industrial é geralmente efectuada pelo processo designado "moagem por via húmida"). Os dispositivos vulgarmente utilizados em processos de extracção de amido a partir de material vegetal são separadores, decantadores, hidrociclones, pulverizadores de secagem e secadores de leito fluidizado.

No contexto da presente invenção, o termo "partes armazenadoras de amido" significa as partes das plantas onde se armazena amido como depósito de sobrevivência por longos períodos, em contraste com o amido transitório da folha. As partes de plantas armazenadoras de amido preferidas são, por exemplo, tubérculos, raízes armazenadoras e grãos, são particularmente preferidos os grãos contendo um endosperma, em especial são particularmente preferidos os grãos que contêm um endoesperma de plantas do milho ou do trigo.

O amido modificado obido por um método de acordo com a invenção, para a produção de amido modificado é também um assunto da presente invenção.

Noutra concretização da presente invenção, os amidos modificados de acordo com a invenção são amidos nativos.

No contexto da presente invenção o termo "amido nativo" significa que o amido é isolado de plantas de acordo com a invenção, plantas que se podem colher de acordo com a invenção, partes armazenadoras de amido de acordo com a invenção ou material de propagação vegetal de acordo com a invenção por meio de métodos conhecidos dos especialistas na técnica. Outro assunto objecto da presente invenção reside na utilização de células vegetais de acordo com a

invenção e plantas de acordo com a invenção para a produção de amido modificado.

Um especialista na técnica sabe que as características do amido podem ser alteradas por meio de derivação térmica, química, enzimática ou mecânica, por exemplo. Os amidos derivados são particularmente adequados para diferentes aplicações no sector alimentar e/ou não alimentar. Os amidos de acordo com a invenção são mais adequados como substância de partida para a produção de amidos derivados do que os amidos convencionais, uma vez que têm um teor mais elevado de grupos funcionais reactivos devido a um maior teor de fosfato de amido.

Por conseguinte, a presente invenção refere-se também à produção de um amido derivado, em que se deriva posteriormente amido modificado de acordo com a invenção.

No contacto da presente invenção, o termo "amido derivado" pressupõe significar um amido modificado de acordo com a invenção, cujas características foram alteradas depois de isolamento das células vegetais por meio de métodos químicos, enzimáticos, térmicos ou mecânicos.

Noutra concretização da presente invenção, o amido derivado de acordo com a invenção é um amido que foi submetido a tratamento térmico e/ou ácido.

Noutra concretização, os amidos derivados são éteres de amido, em particular alquil éter de amido, éteres O-alílicos, éteres hidroxialquílicos, éteres O-carboximéticos, éteres de amido contendo azoto, éteres de amido contendo fosfato ou éteres de amido contendo enxofre.

Noutra concretização, os amidos derivados são amidos reticulados.

Noutra concretização, os amidos derivados são polímeros de enxerto de amidos. Noutra concretização, os amidos derivados são amidos oxidados.

Noutra concretização, os amidos derivados são ésteres de amido, em particular ésteres de amido que foram introduzidos no amido utilizando ácidos orgânicos. Com particular preferência estes são fosfato, nitrato, sulfato, xantato, acetato ou citrato de amido. •

Os amidos derivados de acordo com a invenção são adequados para diferentes aplicações na indústria farmacêutica e no sector alimentar e/ou não alimentar. Os métodos de produção de amidos derivados de acordo com a invenção são conhecidos dos especialistas na técnica e são adequadamente descritos na literatura geral. Encontra-se um resumo da produção de amidos derivados, por exemplo em "Corn, Chemistry and Technology" (1987, eds. Watson und Ramstad, capítulo 16, 479-499). O amido derivado obido por um método de acordo com a invenção, para a produção de amido derivado é também um assunto da presente invenção. A utilização de amidos derivados de acordo com a invenção para a produção de amido derivado é também um assunto da presente invenção.

As partes armazenadoras de amido de plantas são frequentemente transformadas em farinhas. Os exemplos de partes de plantas a partir das quais são processadas as farinhas, por exemplo, tubérculos da batateira e grãos de plantas cerealíferas. Para produzir farinhas a partir das plantas cerealíferas, moi-se e peneira-se os grãos que

contém endosperma destas plantas. O amido é o principal componente do endosperma. No caso doutras plantas que não contêm endosperma, tendo antes outras partes que armazenam amido, tal como tubérculos ou raízes, a farinha é frequentemente produzida por meio da fragmentação, secagem e depois moagem dos órgãos de armazenamento em questão. O amido do endosperma ou o amido contido em partes de armazenamento de amido de plantas é um importante constituinte da farinha que é produzida a partir das partes da planta em questão. As características das farinhas são, assim, também afectadas pelo amido presente na farinha em questão. As células vegetais de acordo com a invenção e plantas de acordo com a invenção sintetizam um amido modificado em comparação com amido das células vegetais de tipo selvagem ou plantas de tipo selvagem correspondentes que não foram geneticamente modificadas. As farinhas produzidas a partir de células vegetais de acordo com a invenção, plantas de acordo com a invenção, material de propagação de acordo com a invenção ou de partes de plantas que se podem colher de acordo com a invenção têm, por conseguinte, características modificadas. As características das farinhas podem também ser afectadas pela mistura de amido com farinhas ou mistura de farinhas com características diferentes.

Por conseguinte, a presente invenção refere-se ainda a farinhas que contêm um amido de acordo com a invenção.

A presente invenção refere-se também a farinhas que são processadas a partir de células vegetais de acordo com a invenção, plantas de acordo com a invenção, partes de plantas armazenadoras de amido de acordo com a invenção de material de propagação de acordo com a invenção ou de

partes de plantas que se podem colher de acordo com a invenção. Partes armazenadoras de amido preferidas de plantas de acordo com a invenção são tubérculos, raízes armazenadoras e um grão que contenha endosperma. Preferencialmente, os tubérculos são tubérculos da batateira e os grãos são grãos de plantas da família (sistemática) das poáceas, com particular preferência os grãos são originários das plantas do milho ou trigo.

No contexto da presente invenção, o termo "farinha" significa um pó obtido da moagem de partes de plantas. Se necessário, as partes das plantas são secas antes da moagem e reduzidas e/ou peneiradas depois da moagem.

Em resultado do amido com um teor de fosfato modificado presente, as farinhas de acordo com a invenção distinguem-se particularmente pela sua maior capacidade de ligação de água. Este facto é desejado, por exemplo, para várias aplicações no processamento de farinhas na indústria alimentar, em particular a produção de produtos de panificação.

A presente invenção refere-se ainda ao método de produção de farinhas, incluindo a etapa de moagem de células vegetais de acordo com a invenção, plantas de acordo com a invenção, partes de plantas de acordo com a invenção, partes de plantas armazenadoras de amido de acordo com a invenção, de material de propagação de acordo com a invenção ou material que se pode colher de acordo com a invenção.

As farinhas podem ser produzidas por meio de moagem de partes armazenadoras de amido, de acordo com a invenção. O especialista na técnica sabe como produzir farinhas. De

preferência, um método de produção de farinhas inclui ainda a etapa de colheita das plantas cultivadas ou partes de plantas e/ou o material de propagação ou as partes armazenadoras de amido destas plantas antes da moagem e com particular preferência a etapa de cultivo das plantas de acordo com a invenção antes da colheita.

No contexto da presente invenção, o termo "partes de plantas" deve pressupor significar todas as partes de uma planta, que representa uma planta completa como constituintes na sua totalidade. Partes de plantas são, por exemplo, rebentos, folhas, rizomas, raízes, beterraba, tubérculo, vagens, sementes ou grãos.

Numa outra concretização da presente invenção, o método de produção de farinhas inclui o processamento de plantas de acordo com a invenção, partes de plantas armazenadoras de amido de acordo com a invenção, de material de propagação de acordo com a invenção ou de material de acordo com a invenção que se pode colher antes da moagem. Neste contexto, o processamento pode ser, por exemplo um tratamento térmico e/ou secagem. O tratamento térmico seguido de uma secagem do material tratado termicamente é utilizado, por exemplo, para a produção de farinhas a partir de raízes armazenadoras ou tubérculos, tal como, por exemplo tubérculos de batata, antes do que tem lugar a moagem. Plantas de acordo com a presente invenção, partes de plantas armazenadoras de amido de acordo com a invenção, de material de propagação de acordo com a invenção ou de material de acordo com a invenção que se pode colher antes da moagem constitui igualmente um processamento para os fins da presente invenção. A remoção de tecido vegetal, tal como, por exemplo, as espigas do grão, antes da moagem

representa também um processamento antes da moagem em termos da presente invenção.

Uma outra concretização da presente invenção inclui o método de produção de farinhas após a moagem de um produto do processo de moenda. Por exemplo, o material moído pode ser peneirado após a moagem, por exemplo para produzir tipos diferentes de farinhas.

A presente invenção refere-se ainda à utilização de células vegetais de acordo com a invenção e plantas de acordo com a invenção geneticamente modificadas para a produção de farinhas. Outro objectivo da presente invenção consiste em proporcionar meios, tal como, por exemplo moléculas de ADN, para a geração de células vegetais de acordo com a invenção e plantas de acordo com a invenção, que sintetizam um amido modificado em comparação com células vegetais de tipo selvagem modificadas ou plantas de tipo selvagem que não foram geneticamente modificadas.

Assim, a presente invenção refere-se também a moléculas de ácido nucleico codificadoras de uma proteína com a actividade enzimática de uma proteína R3, seleccionadas do grupo constituído por

a) moléculas de ácido nucleico que codificam uma proteína com a sequência de aminoácidos especificada na SEQ ID N° 2 ou na SEQ ID N° 4;

b) moléculas de ácido nucleico codificadoras de uma proteína que inclui a sequência de aminoácidos, codificada pela inserção no plasmídeo pIR138-210 ou a inserção no plasmídeo pIR139-210;

c) moléculas de ácido nucleico que codificam uma proteína com pelo menos 60 % de identidade com a sequência de aminoácidos especificada na SEQ ID N° 2 ou na SEQ ID N° 4;

d) moléculas de ácido nucleico codificadoras de uma proteína cuja sequência tem uma identidade de pelo menos 60% com a sequência de aminoácidos, codificada pela inserção no plasmídeo pIR139-210;

e) moléculas de ácido nucleico que incluem a sequência de nucleótidos da SEQ ID N° 1 ou da SEQ ID N° 3 ou uma sequência complementar;

f) moléculas de ácido nucleic que compreendem a sequência de nucleótidos da inserção presente no plasmídeo pIR138-210 ou no plasmídeo pIR139-210;

g) moléculas de ácido nucleico que têm uma identidade de pelo menos 70% com as sequências de ácido nucleico descritas em a), b), e) ou f);

h) moléculas de ácido nucleico que hibridam com pelo menos uma cadeia das moléculas de ácido nucleico descritas em a), b), e) ou f) sob condições restritivas;

i) moléculas de ácido nucleic cuja sequência de nucleótidos se desvia da sequência de moléculas de ácido nucleico mencionadas em a), b), e) ou f), em resultado da degenerescência do código genético e

j) moléculas de ácido nucleico que constituem fragmentos, variantes alélicas e/ou derivados das moléculas de ácido nucleico descritas em a), b), c), d), e), f), g), h) ou i). As moléculas de ácido nucleico de acordo com a

invenção podem ser derivadas, em princípio, de qualquer planta, de preferência de *Arabidopsis thaliana*.

A presente invenção refere-se ainda a moléculas de ácido nucleico com pelo menos 21, de preferência mais de 50 e preferivelmente mais de 200 nucleótidos de comprimento, moléculas de ácido nucleico essas que hibridam especificamente com pelo menos uma molécula de ácido nucleico de acordo com a invenção. No presente contexto, hibridar especificamente significa que estas moléculas hibridam com moléculas de ácido nucleico que codificam uma proteína de acordo com a invenção mas não com moléculas de ácido nucleico que codificam outras proteínas. Em particular, a invenção refere-se às moléculas de ácido nucleico que hibridam com transcrições de moléculas de ácido nucleico de acordo com a invenção e, por conseguinte, podem impedir a tradução destas últimas. Estas moléculas de ácido nucleico que hibridam especificamente com moléculas de ácido nucleico de acordo com a invenção podem ser, por exemplo, componentes de construções antisense, construções ARNi, construções de co-supressão ou ribossomas ou podem ser utilizadas como iniciadores para a amplificação por meio de PCR.

A invenção refere-se ainda a moléculas de ácido nucleico recombinantes compreendendo uma molécula de ácido nucleico de acordo com a invenção.

A presente invenção refere-se preferivelmente a uma molécula de ácido nucleico em que está fundida uma sequência de ácido nucleico codificadora de uma proteína R3 às sequências de ácido nucleico que codificam um peptídeo de sinal plastídeo, de forma que a molécula de ácido

nucleico recombinante codifica uma proteína R3 que, para além da respectiva sequência de aminoácidos codificadora, compreende um peptídeo de sinal plastídeo.

Uma outra concretização de moléculas de ácido nucleico recombinantes de acordo com a invenção é constituída por vectores, em particular plasmídeos, cosmídeos, vírus, bacteriófagos e outros vectores comuns em engenharia genética que contêm as moléculas de ácido nucleico de acordo com a invenção descritos anteriormente. Noutra concretização, as moléculas de ácido nucleico de acordo com a invenção, contidas nos vectores são ligadas com sequências reguladoras que iniciam a expressão em células procariotas ou eucariotas. Em simultâneo, o termo "expressão" pode significar transcrição tal como transcrição e tradução. Neste caso, as moléculas de ácido nucleico de acordo com a invenção podem estar presentes na orientação sense e/ou antisense em relação às sequências reguladoras.

As sequências reguladoras para expressão em organismos procariotas, por exemplo *E.coli* e em organismos eucariotas são extensamente descritas na literatura, em particular aquelas para expressão em levedura, tal como *Saccharomyces cerevisiae*, por exemplo. Encontra-se um resumo dos diferentes sistemas de expressão para proteínas em diferentes organismos hospedeiros, por exemplo em *Methods in Enzymology* 153 (1987), 383516 e em Bitter et al. (*Methods in Enzymology* 153 (1987), 516-544).

Um outro assunto da presente invenção é uma célula hospedeira, em particular uma célula procariota ou eucariota que é geneticamente modificada com uma molécula

de ácido nucleico de acordo com a invenção e/ou com um vector de acordo com a invenção, bem como células derivadas de células hospedeiras deste tipo e que contêm a modificação genética de acordo com a invenção.

As células hospedeiras podem ser bactérias (por exemplo *E.coli*, bactérias do género *Agrobacterium* em particular *Agrobacterium tumefaciens* ou *Agrobacterium rhizogenes*) ou células fúngicas (por exemplo levedura, em particular *S. cerevisiae*, *Agaricus*, in particular *Agaricus bisporus*, *Aspergillus*, *Trichoderma*) bem como células vegetais ou animais. No presente contexto, o termo "transformado" significa que as células de acordo com a invenção são geneticamente modificadas com uma molécula de ácido nucleico de acordo com a invenção na medida em que contêm pelo menos uma molécula de ácido nucleico de acordo com a presente invenção para além do seu genoma natural. Esta molécula de ácido nucleico de acordo com a invenção pode encontrar-se presente livremente na célula, facultativamente como molécula auto-replicante, ou pode estar integrada de forma estável no genoma da célula hospedeira. As células hospedeiras são preferencialmente microorganismos. Para os fins do presente pedido de patente, estas são encaradas como significando todas as bactérias e todos os protistas (por exemplo, fungos, em particular leveduras e algas) como definido, por exemplo, em Schlegel "Allgemeine Mikrobiologie" (Georg Thieme Verlag (1985), 1-2).

As células hospedeiras de acordo com a invenção são preferivelmente células vegetais. Em princípio, estas podem ser células vegetais de qualquer espécie vegetal desejada, isto é tanto monocotiledóneas como dicotiledóneas. De

preferência são plantas úteis, isto é plantas que são cultivadas pelo homem para fins alimentares ou técnicos, em particular industriais. A presente invenção refere-se de preferência a células vegetais e plantas de plantas armazenadoras de amido (milho, arroz, trigo, centeio, aveia, cevada, mandioca, batata, sagú, feijão mungo, ervilha ou sorgo), de preferência células vegetais da batateira, de preferência células vegetais de plantas da família (sistemática) das poáceas, sendo especialmente preferidas células vegetais de plantas do milho ou do trigo. A presente invenção refere-se também a composições contendo uma molécula de ácido nucleico de acordo com a invenção, uma molécula de ácido nucleico recombinante de acordo com a invenção ou um vector de acordo com a invenção. São preferidas as composições de acordo com a presente invenção contendo uma molécula de ácido nucleico de acordo com a invenção, uma molécula de ácido nucleico recombinante de acordo com a invenção ou um vector de acordo com a invenção e uma célula hospedeira. A célula hospedeira assume preferivelmente a forma de uma célula vegetal, com especial preferência a forma de uma célula de planta do milho ou de trigo. Nestas composições, a molécula de ácido nucleico recombinante de acordo com a invenção existe fora da célula hospedeira, isto é está localizada fora da parte interior da célula hospedeira que está envolvida pela membrana citoplasmática.

Um outro aspecto das composições de acordo com a invenção refere-se a composições que podem ser utilizadas para gerar células hospedeiras de acordo com a invenção, de preferência para a geração de células vegetais de acordo com a invenção. Este aspecto tem preferivelmente a forma de

uma composição contendo uma molécula de ácido nucleico de acordo com a invenção, uma molécula de ácido nucleico recombinante de acordo com a invenção ou um vector de acordo com a invenção e um veículo biolístico que é adequado para a introdução de uma molécula de ácido nucleico de acordo com a invenção numa célula hospedeira. Os veículos biolísticos preferidos são partículas de tungsténio, ouro ou materiais sintéticos. Uma outra concretização de composições de acordo com a presente invenção refere-se a composições contendo uma molécula de ácido nucleico de acordo com a invenção, uma molécula de ácido nucleico recombinante de acordo com a invenção ou um vector de acordo com a invenção e uma célula vegetal e um meio de cultura sintético. Para além das moléculas de ácido nucleico de acordo com a invenção, as células vegetais e o meio de cultura sintético, estas composições compreende preferivelmente também polietilenoglicol (PEG). Nestas composições, a molécula de ácido nucleico recombinante de acordo com a invenção existe fora da célula vegetal, isto é está localizada fora da parte interior da célula vegetal que está envolvida pela membrana citoplasmática. Os especialistas na técnica conhecem meios de cultura sintéticos que são adequados para a cultura e/ou transformação de células vegetais, que são adequadamente descritos na literatura. Existem também muitos meios de cultura sintéticos diferentes disponíveis no comércio especializado (por exemplo DUCHEFA Biochemie B.V., Bélgica).

Uma outra concretização das composições de acordo com a invenção refere-se a composições que são utilizadas para identificar os ácidos nucleicos de acordo com a invenção.

Para além de uma molécula de ácido nucleico de acordo com a invenção, uma molécula de ácido nucleico recombinante de acordo com a invenção ou um vector de acordo com a invenção estas composições incluem preferivelmente outras moléculas de ácido nucleico, em particular moléculas de ácido nucleico de origem vegetal que podem estar presentes sob a forma de ADN genómico, mRNA ou clones no que é conhecido como bibliotecas de ADN. São preferidas as bibliotecas de ADN que existem em cosmídeos, fagemídeos, YACs ou BACs. As bibliotecas de ADN podem incluir tanto ADN genómico como cADN. As moléculas de ácido nucleico de acordo com a invenção, as moléculas de ácido nucleico recombinantes de acordo com a invenção ou vectores de acordo com a invenção são preferivelmente empregues como sondas de hibridação nestas composições.

Um outro assunto objecto da presente invenção consiste em proteínas seleccionada do grupo constituído por

a) proteínas que compreendem a sequência de aminoácidos especificada na SEQ ID N° 2 ou na SEQ ID N° 4;

b) proteínas que estão codificadas na região codificadora do ADN inserido no plasmídeo pIR138-210 ou no plasmídeo pIR139-210 ou

c) proteínas com pelo menos 60 % de identidade com a sequência de aminoácidos das proteínas citadas em a) ou b).

A presente invenção refere-se ainda a uma proteína de acordo com a invenção com um peso molecular entre 120 kDa a 145 kDa, de preferência entre 120 kDa e 140 kDa, com especial preferência entre 125 kDa e 140 kDa, particularmente preferencial entre 130 kDa e 135 kDa, derivada da sequência de aminoácidos.

Uma outra concretização da presente invenção refere-se a uma proteína de acordo com a invenção que tem um domínio fosfo-histidina.

Noutra concretização, a presente invenção refere-se a proteínas com actividade de fosforilação do amido, em que a proteína codificada tem pelo menos 70%, de preferência pelo menos 80%, preferivelmente pelo menos 90% e com especial preferência pelo menos 95% de identidade com a sequência de aminoácidos da SEQ ID NO 2 ou SEQ ID NO 4 ou com a sequência de aminoácido de uma proteína R3, codificada pela inserção no plasmídeo pIR138-210 ou pela inserção no plasmídeo pIR139-210.

Uma outra concretização da presente invenção refere-se a uma proteína de acordo com a invenção em que a sequência de aminoácidos que codifica a proteína tem um domínio fosfo-histidina. Preferivelmente, a proteína de acordo com a invenção possui um domínio fosfo-histidina que possui pelo menos 85%, de preferência pelo menos 88%, com especial preferência pelo menos 91% e com particular preferência pelo menos 94% e com especial preferência pelo menos 97% de identidade com a sequência de aminoácidos da SEQ ID NO 5.

Uma concretização preferida do processo de acordo com a invenção refere-se a proteínas recombinantes que compreendem uma sequência de aminoácidos codificadora da proteína R3, que está fundida às sequências de aminoácidos que codificam um peptídeo de sinal plastídeo.

No contexto da presente invenção, o termo "proteína recombinante" subentende significar uma proteína que, para além das sequências de aminoácidos codificadoras de uma proteína R3 de acordo com a invenção, compreende sequências

de aminoácidos adicionais, em que a combinação das sequências de aminoácidos codificadoras da proteína R3 e das sequências de aminoácidos adicionais mencionadas não existe naturalmente na combinação em que existe nas proteínas recombinantes de acordo com a invenção. Neste contexto, as sequências de aminoácidos adicionais mencionadas podem ser quaisquer sequências de aminoácidos. Preferivelmente assumem a forma de sequências de aminoácidos que codificam peptídeos de sinal que, em resultado da fusão traducional com sequências de aminoácidos codificadoras da proteína R3, acarreta a translocação das proteínas em questão para compartimentos subcelulares (vacúolos, plastídeos, mitocôndrias, núcleos, retículos endoplasmáticos, parede celular). São especialmente importantes no contexto da presente invenção as sequências de aminoácidos que codificam peptídeos de sinal plastídeos. Os métodos para a geração de proteínas recombinantes de acordo com a invenção são conhecidos do especialista e incluem métodos recombinantes, tal como, tais como a transcrição e tradução de moléculas de ácido nucleico recombinantes que codificam proteínas recombinantes.

As fontes de sequências de aminoácidos que codificam peptídeos de sinal plastídeos já foram mencionadas anteriormente no contexto do termo "molécula de ácido nucleico recombinante".

Noutra concretização adicional da presente invenção refere-se a uma proteína de acordo com a invenção, sendo a proteína derivada da planta *Arabidopsis*.

Noutra concretização, a invenção refere-se também a proteínas que são codificadas por moléculas de ácido nucleico de acordo com a invenção.

#### Descrição das sequências

SEQ ID NO 1: Sequência de ácido nucleico contendo a região codificadora da proteína R3 de *Arabidopsis thaliana*. A sequência é inserida no plasmídeo pLR139- 210. SEQ ID NO 2: Sequência de aminoácidos codificadora da proteína R3 de *Arabidopsis thaliana*. Esta sequência pode ser derivada da sequência de ácido nucleico apresentada na SEQ ID NO 1. SEQ ID NO 3: Sequência de ácido nucleico compreendendo a região codificadora da proteína R3 *Arabidopsis thaliana* em fusão traducional com a sequência de sinal codificadora do peptídeo de sinal para plastídeo do gene *dil I* de *Oryza sativa*. Esta sequência é inserida no plasmídeo pLR138-210. SEQ ID NO 4: Sequência de aminoácidos codificadora da proteína de fusão que consiste na sequência de aminoácidos codificadora da proteína R3 de *Arabidopsis thaliana* e a sequência de aminoácidos codificadora do peptídeo de sinal para plastídeo do gene *dil I* (TrEMBL Ace. No.: Q621D6) de *Oryza sativa*. Esta sequência pode ser derivada da sequência de ácido nucleico apresentada na SEQ ID NO 3. SEQ ID NO 5: Sequência de aminoácidos do domínio fosfo-histidina que está presente na sequência de aminoácidos codificadora da proteína R3 de *Arabidopsis thaliana*.

O teor de todas as publicações mencionadas e sequências citadas por números de acesso a base de dados são incorporados na descrição por referência.

#### Métodos gerais

O texto seguinte irá descrever métodos que podem ser aplicados na implementação dos métodos de acordo com a invenção. Estes métodos constituem concretizações específicas da presente invenção, mas não limitam a presente invenção a esses métodos. O especialista na técnica sabe que a invenção pode ser implementada da mesma forma por meio da modificação dos métodos descritos e/ou substituição de partes individuais dos métodos por partes alternativas dos métodos.

1. Expressão recombinante de uma proteína de fosforilação do amido identificada

a) Geração de um vector de expressão bacteriana contendo um cADN que codifica uma proteína de fosforilação do amido. O cADN codificador de uma proteína de fosforilação do amido pode ser amplificada por meio de uma reacção em cadeia de polimerase (PCR), por exemplo, utilizando mARN ou poli-A-plus-mARN de tecidos vegetais como modelo. Com este fim, utiliza-se primeiro uma transcriptase reversa para a geração de uma cadeia de cADN, que é complementar de um mARN, que codifica uma proteína de fosforilação do amido, antes de a cadeia de cADN em questão ser amplificada por meio de polimerase de ADN. Os ditos "kits" contendo substâncias, enzimas e instruções para realização de reacções PCR estão disponíveis no comércio (por exemplo SuperScript™ One-Step RT-PCR System, Invitrogen, Prod. No.: 10928-034). O cADN que foi amplificado e que codifica uma proteína de fosforilação do amido pode subsequentemente ser clonado num vector de expressão bacteriana, por exemplo pDEST™17 (Invitrogen). pDEST™17 contém o promotor T7, que é utilizado para iniciar a transcrição da T7-ARN-polimerase. Além disso, o vector de

expressão pDEST™17 contém uma sequência de Shine-Dalgarno na direcção 5' do promotor T7 seguida de um codão de partida (ATG) e por uma etiqueta designada His. Esta etiqueta His consiste em seis codões consecutivos, que codificam cada um o aminoácido histidina e estão localizados na estrutura de leitura aberta do codão de partida mencionado. Um cADN codificador de uma proteína de fosforilação do amido é clonado em pDEST™17 de forma a gerar uma fusão de traducional entre os codões para o codão de partida, a etiqueta His e o cADN codificador de uma proteína de fosforilação de amido. Após a transcrição, iniciada no promotor T7 e subsequente tradução, obtém-se uma proteína de fosforilação do amido que contém aminoácidos adicionais, contendo a etiqueta His no respectivo terminal N. No entanto, outros vectores que são adequados para a expressão em microrganismos podem também ser utilizados para a expressão de uma proteína de fosforilação do amido. Os vectores de expressar e estirpes de expressão equivalentes são conhecidos dos especialistas e existem igualmente no mercado combinações adequadas.

b) Geração de clones de expressão em *Escherichia Coli*

Em primeiro lugar, transforma-se uma estirpe de *E.coli* apropriada, de transformação competente, que codifica em termos cromossomais uma polimerase T7-ARN, com o plasmídeo de expressão produzido como se descreve na etapa a) e a estirpe é depois incubada de um dia para o outro a 30° C em meio nutriente de cultura solidificado com agar. As estirpes de expressão adequadas são, por exemplo, estirpes BL21 (Invitrogen Prod. No.: C6010-03, que codificam em termos cromossomais uma polimerase T7-ARN sob o controlo de um promotor induzível com IPTG (lacZ)). As colónias de

bacterianas, que constituem o resultado da etapa de transformação, podem ser analisadas por métodos conhecidos dos especialistas para detectar a presença do plasmídeo de expressão desejado que compreende cADN, que codifica a proteína de fosforilação do amido. Obtêm-se assim os clones de expressão.

c) Expressão de uma proteína de fosforilação do amido em *Escherichia Coli*

Primeiro é preparada uma pré-cultura. Com este objectivo, inocula-se um clone de expressão obtido de acordo com a etapa b) em 30 ml de Terrific Broth (meio TB) contendo um antibiótico, para a selecção em presença do plasmídeo de expressão e a cultura é incubada de um dia para o outro a 30° C, com agitação (250 rpm). É depois produzida uma cultura principal para a expressão de uma proteína de fosforilação do amido. Com este objectivo, balões de Erlenmeyer de 1 litro, contendo cada um 300 ml de meio TB, pré-aquecido a 30° C e um antibiótico para selecção da presença do plasmídeo de expressão são inoculados, cada um com 10 ml de um pré-cultura apropriada e incubados a 30° C com agitação (250 rpm) até se atingir uma densidade óptica (medida a um comprimento de onda de 600 nm; OD<sub>600</sub>) de cerca de 0,8. Se, for utilizado um plasmídeo de expressão para expressar uma proteína de fosforilação do amido, plasmídeo esse em que a expressão da proteína de fosforilação do amido é iniciada por meio de um sistema induzível (por exemplo, o vector de expressão pDEST™17 em estirpes BL21 de *E. coli*, induzível por meio de IPTG), o indutor em questão (por exemplo IPTG) é adicionado quando é atingido um OD<sub>600</sub> de cerca de 0,8 na cultura principal. Após adição do indutor, a cultura

principal é incubada a 30° C, com agitação (250 rpm), até ser atingido um OD<sub>600</sub> de cerca de 1,8. Seguidamente, a cultura principal é depois arrefecida durante 30 minutos em gelo antes de as células da cultura principal são então separadas do meio de cultura por centrifugação (10 minutos a 4.000xg e 4 °C).

## 2. Purificação de uma Proteína de Fosforilação do Amido

a) Fragmentação de células que expressam uma proteína de fosforilação do amido

As células obtidas após centrifugação na etapa c), Ponto 1 dos Métodos Gerais, são ressuspensas em tampão de lise, adicionando-se aproximadamente 4 ml de tampão de lise a aproximadamente 1 g de células. As células ressuspensas são então incubadas durante 30 minutos em gelo antes de serem fragmentadas com ajuda de uma sonda ultrassónica (Baudelin Sonoplus UW 2070, Baudelin electronic, Berlin, definições: ciclo 6, 70%, 1 minuto) com refrigeração contínua por meio de gelo. É necessário cuidado para garantir que a suspensão de células não aquece demasiado durante o tratamento ultrassónico. A suspensão obtida após o tratamento ultrassónico é centrifugada (12 minutos a 20000 xg, 4° C) e o sobrenadante obtido após a centrifugação é filtrado utilizando um filtro com uma dimensão de poro de 45 µm.

b) Purificação da proteína de fosforilação do amido

Se a proteína de fosforilação do amido, expressa em células de *E.coli* for uma proteína de fusão com uma etiqueta His, a purificação pode ocorrer com ajuda de iões de níquel, aos quais a etiqueta His se liga com maior afinidade. Com este fim, 25 ml do filtrado obtido na etapa

d) é tratado com 1 ml de suspensão de Ni-agarose (Qiagen, Prod. No.: 30210) e a mistura é incubada durante 1 hora em gelo.

Seguidamente, a mistura de suspensão de Ni-agarose e filtrado é seguidamente passada por uma coluna de polistireno (Pierce, Prod. No.: 29920). Elimina-se o produto que passou pela coluna. A coluna é lavada primeiro por meio da aplicação de 8 ml de tampão de lise, sendo o produto que passou pela a coluna novamente posto de lado. A proteína de fosforilação do amido é depois eluida por aplicação fraccionada na coluna de duas porções de 1 ml de tampão E1 cada, seguido de uma porção de 1 ml de tampão E2 e depois cinco porções de 1 ml de coluna E3 cada. O produto que passou obtido durante a aplicação da fracção individual dos tampões de eluição em questão (tampões E1, E2 e E3) é recolhido em fracções separadas. As aliquotas destas fracções são posteriormente analisadas por meio de electroforese em gel de acrilamida SDS desnaturante, seguida de coloração com azul de Coomassie. As fracções que contém a proteína de fosforilação do amido em quantidade suficiente e com pureza satisfatória são combinadas e concentradas a 4° C, com ajuda de filtração pressurizada. A filtração pressurizada pode ser efectuada, por exemplo com ajuda de um Amicon cell (Amicon Ultrafiltration Cell, Model 8010, Prod. No.: 5121) utilizando uma membrana Diaflo PM30 (Millipore, Prod. No.: 13212) a 4° C. No entanto, podem também ser empregues outros métodos conhecidos dos especialistas para a etapa de concentração.

c) Composição dos Tampões Utilizados

Tampão de lise: 50 mM HEPES

300 mM NaCl  
10 mM Imidazole  
pH 8,0 (ajuste com NaOH)

1 mg/ml Lisosima (adicionar imediatamente antes de utilizar o tampão)

¼ de comprimido por 10 ml de inibidores de protease completamente isentos de EDTA (produto Roche No.: 1873580) (adicionar imediatamente antes de utilizar o tampão)

Tampão de eluição E1: 50 mM HEPES  
300 mM NaCl  
75 mM Imidazole  
pH 8,0 (ajuste com NaOH)

Tampão de eluição E2: 50 mM HEPES  
300 mM NaCl  
75 mM imidazole  
pH 8,0 (ajuste com NaOH)

Tampão de eluição E3: 50 mM HEPES  
300 mM NaCl  
250 mM Imidazole  
pH 8,0 (ajuste com NaOH)

3. Método de detecção da actividade fosforiladora do amido de uma proteína

a) Incubação de proteínas com amido e/ou amido não fosforilado

Para detectar se uma proteína possui uma actividade de fosforilação do amido, as proteínas de ensaio podem ser

incubadas com amido e com trifosfatos de nucleósido marcados radioactivamente. Com este objectivo, aproximadamente 5 mg de amido em conjunto com a proteína de ensaio (0,01 a 5,0 por mg de amido empregue) são incubados com agitação à temperatura ambiente durante 10 minutos até 30 minutos em 500  $\mu$ l de tampão de fosforilação. Seguidamente, a reacção é parada com a adição de SDS até atingir uma concentração de 2% (peso/volume). Os grânulos de amido presentes na mistura reaccional são separados por centrifugação (1 minuto, 13000 xg) e lavados uma vez com 900  $\mu$ l de uma solução SDS a 2% e quatro vezes com 900  $\mu$ l de uma solução trifosfato de nucleósido 2 mM cada. Cada etapa de lavagem é efectuada durante 15 minutos à temperatura ambiente, com agitação. Após cada etapa de lavagem, os grânulos de amido são separados do tampão de lavagem em questão por centrifugação (1 min, 13.000 xg). Durante uma experiência para detectar a actividade de fosforilação do amido de uma proteína, outras misturas reaccionais, que não contêm a proteína ou contêm a proteína inactivada, mas que de resto são tratadas da mesma forma que as preparações reaccionais descritas, devem ser adicionalmente processadas para assumirem o que é conhecido como controlo.

b) Determinação da quantidade de resíduos fosfato incorporados no amido em resultado da actividade enzimática

Os grânulos de amido obtidos na etapa a) podem ser analisados quanto à presença de resíduos de fosfato com marcação radioactiva. Com este objectivo, o amido em questão é re-suspenso em 100  $\mu$ l de água cada e tratado com 3 ml de *cocktail* de cintilação em cada caso (por exemplo, Ready Safe™, BECKMANN Coulter) e seguidamente analisado com

ajuda de um contador de cintilações (por exemplo, LS 6500 Multi-Purpose Scintillation Counter, BECKMANN COULTER™).

d) Composição dos Tampões Utilizados

Tampão de fosforilação: 50 mM HEPES-KOH, pH 7,5

1 mM EDTA

6 mM MgCl<sub>2</sub>

0,01 a 0,5 mM ATP

0,2 a 2 µCi por ml <sup>33</sup>P-ATP aleatorizado

4. Identificação das posições átomo C das moléculas de glucose de um alfa-1,4-glucano, no qual são introduzidos resíduos fosfato por uma proteína de fosforilação do amido.

As posições átomo C das moléculas de glucose de um alfa-1,4-glucano que são fosforiladas por uma proteína, podem ser detectadas submetendo os glucanos obtidos, que foram fosforilados *in vitro*, por meio de uma proteína específica, a hidrólise, subsequente separação dos monómeros de glucose obtidos após a hidrólise, seguido de medição do fosfato incorporado por uma proteína apropriada em certas fracções das moléculas de glucose.

a) Hidrólise Total dos alfa-1-glucanos

Suspensões aquosas contendo alfa-1,4-glucanos são centrifugadas, os *pellets* sedimentados são depois re-suspenso em HCL 0,7 M (Baker, nível analítico) e incubadas durante 2 horas a 95 °C, com agitação. Quando completa a incubação, as amostras são rapidamente arrefecidas e centrifugadas (por exemplo 2 minutos 10000 xg). O sobrenadante obtido é transferido para um reactor novo e neutralizado com a adição de NaOH 2 M (Baker nível

analítico). Se restar um *pellet*, este é re-suspenso em 100 µl de água e a quantidade de fosfato marcado presente é determinada para controlo. O sobrenadante neutralizado é depois centrifugado através de um filtro de 10 kDa. Medindo uma alíquota do filtrado obtido, é determinada a quantidade de fosfato marcado no filtrado, por exemplo, com ajuda de um contador de cintilações.

b) Fraccionamento dos hidrolisados e determinação das posições do átomo C fosforilado.

Os filtrados hidrolíticos neutralizados, obtidos na etapa a) (aproximadamente 3000 cpm quando é utilizado ATP com marcação radioactiva) podem ser separados com ajuda de, por exemplo, cromatografia de permuta de aniões de alta pressão (HPAE). Para que o filtrado neutralizado atinja o volume desejado para HPAE pode ser diluído com H<sub>2</sub>O. Além disso, adicionam-se glucose-6-fosfato (cerca de 0,15 M) e glucose3-fosfato (cerca de 0,3 mM) a cada filtrado em questão para controlo interno. Pode ser executada uma separação por meio de HPAE, por exemplo com ajuda de um sistema Dionex DX 600 Bio Lc, utilizando uma coluna CarboPac PA 100 (com pré-coluna apropriada) e um detector amperométrico pulsante (ED 50). Antes da injeção da amostra, a coluna é enxaguada primeiro durante 10 minutos com 99% de eluente C e 1% de eluente D. Seguidamente, injecta-se 600 µl do volume da amostra em cada caso. A amostra é eluída nas condições seguintes:

Caudal: 1 ml por minuto

Gradiente: Gradiente linear de 0 minutos a 30 minutos

Eluente C:      Eluente D:

0 minutos 99% 1%  
30 minutos 0% 100%  
35 minutos 0% 100%

Passagem terminada

Os produtos da hidrólise eluídos da coluna são recolhidos em fracções individuais de 1 ml cada. Dado que, em cada caso, foram adicionados glucose-3-fosfato não marcado (Ritte et al. 2002, PNAS 99, 7166-7171) e glucose-6-fosfato não marcado (Sigma, Prod. No.: G7879) às amostras injectadas dos produtos da hidrólise como padrões internos, as fracções que contêm glucose-3-fosfato ou glucose-6-fosfato podem ser determinadas por meio de detecção amperométrica pulsante. Medindo a quantidade de fosfatos marcados nas fracções individuais e comparando depois com as fracções que contêm glucose-3-fosfato ou glucose-6-fosfato, é possível determinar as fracções em que o glucose-6-fosfato marcado ou o glucose-3-fosfato marcado está presente. A quantidade de fosfato marcado na fracção em questão é determinada. Agora é possível determinar a posição do átomo C que é preferencialmente fosforilado por uma enzima de fosforilação alfa-1,4-glucano com base nas proporções das quantidades, medidas para fosfato marcado, de glucose-3-fosfato em comparação com glucose-6-fosfato nos hidrolisados individuais.

c) Tampões utilizados

Eluente C: 100 mM NaOH

Eluente D: 100 mM NaOH

500 mM acetato de sódio

## 5. Transformação de plantas do arroz

Plantas do arroz foram transformadas de acordo com os métodos descritos por Hiei et al. (1994, Plant Journal 6(2), 271-282).

## 6. Determinação do Teor de Fosfato de Amido

### a) Determinação do Teor de Fosfato C-6 (C-6-P)

No amido, as posições C2, C3 e C6 das unidades de glucose podem ser fosforiladas. Para determinar o teor de C6-F no amido, hidrolisa-se 50 mg de amido em 500 µl de HCl 0,7 M durante 4 horas a 95 °C. Em seguida, as misturas são centrifugadas durante 10 minutos a 15.500 g e os sobrenadantes são removidos. 7 µl dos sobrenadantes são misturados com 193 µl de tampão de imidazole (100 mM de imidazole, pH 7,4; 5 mM MgCl<sub>2</sub>, 1 mM EDTA e 0,4 mM NAD). A medição é efectuada no fotómetro a 340 nm. Depois de estabelecer uma absorção de base, a reacção enzimática foi iniciada por meio de adição de 2 unidades de glucose-6-fosfato desidrogenase (de *Leuconostoc mesenteroides*, Boehringer Mannheim). A alteração na absorção é directamente proporcional à concentração do teor de G-6-P do amido.

### b) Determinação do Teor de Fosfato Total

A determinação do teor de fosfato total é efectuada de acordo com o método de Ames (Methods in Enzymology VIII, (1966), 115-118).

Aproximadamente 50 mg de amido são misturados com 30 µl de solução de nitrato de magnésio etanólica e a mistura é incinerada durante três horas a 500° C em forno de mufla. O resíduo é tratado com 300 µl de ácido clorídrico a 0,5 M e

incubado durante 30 min a 60° C. Depois, perfaz-se uma porção de aliquota até 300 µl de ácido clorídrico a 0,5 M, adiciona-se à mistura de 100 µl de ácido ascórbico a 10% e 600 µl de molibdato de amónio a 0,42% em 2 M de ácido sulfúrico e incuba-se durante 20 minutos a 45 °C.

c) Determinação do Teor de Fosfato C-6 e fosfato C-3

Para a determinação do teor de fosfato que está ligado na posição C-6 e na posição C-3 das moléculas de glucose de um alfa-1-glucano, os glucanos em questão podem ser sujeitos a uma hidrólise total e depois separados por meio de HPAE. As quantidades de glucose6-fosfato e glucose3-fosfato podem ser determinadas por integração das áreas pico individuais, obtidas depois da separação com HPEA. Com a comparação das áreas pico obtida de glucose-6-fosfato e glucose-3-fosfato em amostras desconhecidas com as áreas pico que são obtidas com quantidades conhecidas de glucose-6-fosfato e glucose-3-fosfato, depois da separação com HPEA, é possível determinar a quantidade de glucose-6-fosfato e glucose.3-fosfato nas amostras em análise.

Exemplos

1. Informação sobre os vectores e plasmídeos

Geração do vector de expressão ME5/6

O pGSV71 é um derivado do plasmídeo pGSV7, que provém do vector intermédio pGSV1. O pGSV71 é um derivado do pGSC1700, cuja construção foi descrita por Cornelissen and Vanderwiele (Nucleic Acid Research 17, (1989), 19-25). O pGSV1 foi obtido a partir do pGSC1700 por eliminação do gene resistente à carbenicilina e eliminação das sequências do T-DNA da região TL-DNA do plasmídeo pTiB6S3. O pGSV7 contém a origem da replicação do plasmídeo pBR322 (Bolivar

et al., Gene 2, (1977), 95-113), assim como a origem da replicação do plasmídeo pVSI *Pseudomonas* (Itoh et al., Plasmid 11, (1984), 206). O pGSV7 contém também o gene marcador selecionável *aadA*, a partir de transposões Tn1331 de *Klebsiella pneumoniae*, que oferece resistência contra os antibióticos spectinomomicina e estreptomicina (Tolmasky, Plasmid 244 (3), (1990), 218-226; Tolmasky e Crosa, Plasmid 29(1), (1993), 31-40). O plasmídeo pGSV71 foi obtido pela clonagem de um gene *bar* quimérico entre as regiões limite de pGSV7. O gene *bar* quimérico contém a sequência promotora do vírus mosaico da couve-flor para iniciar a transcrição (Odell et al., Nature 313, (1985), 180), o gene *bar* do *Streptomyces hygrosopicus* (Thompson et al., 1987, EMBO J. 6, 2519-2523) e a região 3'-não traduzida do gene da sintetase nopalina do T-ADN de pTiT37 para terminar a transcrição e a poliadenilação. O gene *bar* confere uma tolerância contra o herbicida glufosinato-amônio. O T-ADN contém a sequência do flanco direito do TL-ADN do plasmídeo pTiB6S3 (Gielen et al., EMBO J. 3, (1984), 835-846) na posição 198-222. Uma sequência poliligante encontra-se localizada entre os nucleótidos 223-249. Os nucleótidos 250-1634 contêm a região promotora p35S3 do vírus mosaico da couve-flor (Odell et al., c.f. acima). A sequência codificadora do gene da resistência à fosfotricina de *Streptomyces hygrosopicus* (*bar*) (Thompson et al. 1987, c.f. acima) é disposta entre os nucleótidos 1635-2186. Os dois codões terminais no terminal 5' do gene de tipo selvagem *bar* foram substituídos pelos codões ATG e GAC. Uma sequência poliligante encontra-se localizada entre os nucleótidos 2187-2205. O fragmento TaqI de 260 bp do terminal 3' não traduzido do gene da sintetase nopalina

(3'nos) do t-ADN do plasmídeo pTiT37 (Depicker et al., J. Mol. Appl. Genet. 1, (1982), 561-573) está localizado entre os nucleótidos 2206 e 2465. Os nucleótidos 2466-2519 contêm uma sequência poliligante. A sequência do flanco esquerdo do pTiB6S3 TL-ADN (Gielen et al., EMBO J. 3, (1984), 835-846) encontra-se localizado na posição 2520-2544. O vector pGSV71 foi depois cortado com a enzima PstI e embotado. O promotor B33 e a cassete ocs foram depois excisados do vector pB33-Kan sob a forma de um fragmento EcoRI-HindIII, tornado embotado por preenchimento das extremidades e inserido no vector pGSV71, que tinha sido cortado com PstI e embotado. O vector resultante (ME4/6) foi utilizado como vector de partida para a construção de ME5/6. Introduziu-se um oligonucleótido contendo locais de clivagem EcoRI, PacI, SpeI, SrfI, SpeI, NotI, PacI e EcoRI no local de clivagem PstI do vector ME4/6, localizado entre o promotor B33 e o elemento ocs, duplicando o local de clivagem PstI. O vector de expressão resultante foi designado ME5/6.

#### Geração do vector de expressão pML80

Então, o fragmento BamHI de ME5-6 foi trocado por um produto PCR que tinha sido prolongado em alguns sítios de clivagem, mas que de resto era idêntico, dando origem ao plasmídeo pUL1-17. O promotor B33, presente no pUKL1-17, foi excisado utilizando as enzimas de restrição HindIII e PstI e o vector foi embotado e religado, dando origem ao vector pML18-56. Este vector foi aberto com MunI e PstI e foi inserido um MCS (local de clonagem múltiplo) com extremidades adesivas adequadas e que tinha sido sintetizado por dois oligonucleótidos emparelhados (GAG CTC CTA GGC TCG AGT TAA CAC TAG TAA GCT TAA TTA AGA TAT CAT TTA CA and AAT TGT AAA TGA TAT CTT AAT TAA GCT TAC TAG TGT TAA

CTC GAG CCT AGG AGC TCT GCA). O plasmídeo resultante pML72-129 foi cortado com a enzima de restrição *Bam*HI, a saliência dos locais de clivagem foram preenchidas com ajuda de polimerase Klenow e as extremidades embotadas resultantes foram ligadas. Esta operação deu origem a um local de clivagem de restrição *Cla* I. O plasmídeo resultante foi designado pIR91-123. O promotor ubiquitina do milho foi excisado de pML8-52 por digestão de restrição com as enzimas *Sal* I e *Not* I e embotado e o fragmento resultante foi clonado no plasmídeo *Eco* RV-cut pIR91-123. O plasmídeo resultante foi designado pML80-123.

Geração do vector pML8-52: O pMCS5 (MobiTec, Prod. No.: pMCS5) foi cortado com *Bgl* II e *Bam*HI e foi religado. O plasmídeo resultante foi designado pML4-52. O terminador nos do gene nopalina sintetase de *Agrobacterium tumefaciens* (Depicker et al., J. MoI. Appl. Genet. 1, (1982), 561-573), que tinha sido amplificado com ajuda dos oligonucleótidos PML9 (ACT TCT GCA GCG GCC GCG ATC GTT CAA ACA TTT GGC AAT AAA GTT TC) e PML10 (TCT AAG CTT GGC GCC GCT AGC AGA TCT GAT CTA GTA ACA TAG ATG ACA CC) foi clonado neste vector e nos locais de clivagem de restrição *Pst* I e *Hind* III. O plasmídeo resultante foi designado pML4-nos. Um fragmento *Pst* I de 1986 pares básicos de comprimento que contém o promotor do gene da poliubiquitina a partir do milho (Genbank Ace: 94464, Christensen et al., 1992, Plant MoI. Biol. 18: 675-689) e o primeiro intrão do mesmo gene que tinha sido truncado por digestão com *Cla* I e religação foi clonado no vector mencionado. O plasmídeo resultante foi designado pML8-52.

2. Isolamento das sequências codificadoras da proteína R3 de *Arabidopsis thaliana*

A sequência codificadora da proteína R3 de *Arabidopsis thaliana* foi isolada por amplificação, por meio de RT-PCR de várias sequências sobrepostas, a começar no mRNA isolado a partir de folhas de plantas de *Arabidopsis thaliana* com aproximadamente três semanas de idade (ecotipo C24) e subsequente clonagem sequencial dos fragmentos amplificados. Foram utilizados os seguintes iniciadores:

Nome	Sequência	F1	AAGCTGCCATGGCAACCTCTAAATC
R2	AAAACTCGAGCTTAAACTTGGGGTCTAGCTTGGA		
R4	CTCAGAGCCATTGGAGCACTCA		
F4	CATTGCAAACAAAAGCGGTCCTTG		

Com este objetivo, foi sintetizada uma estirpe de cADN por meio da transcriptase inversa Revert Aid H Minus First Strand cDNA Synthesis Kit (Fermentas Prod. No: EP0451), que foi então amplificada com polimerase de ADN (Platinum Taq DNA Polymerase High Fidelity; Invitrogen Prod No.: 11304). Um fragmento de ADN com aprox. 3,2 kb foi amplificado com 300 ng de um ARN poli-A-plus, com ajuda dos iniciadores F1 e R4 nas seguintes condições de reacção:

Síntese da primeira cadeia: Foram utilizadas as condições e tampão especificados pelo fabricante. A mistura reaccional para a síntese da primeira cadeia continha ainda as seguintes substâncias:

300 ng	ARN poli-A-plus
0,5 µM	Iniciador R2

Juntou-se água à mistura reaccional até atingir um volume de 12 µl, incubou-se durante 5 minutos a 70 °C e depois arrefeceu-se em gelo. Seguidamente, adicionou-se 5x

tampão de reacção (4 µl), inibidor RNase (1 µl) e mistura dNTP (2 µl de 10 mM) e incubou-se a mistura reaccional durante 5 minutos a 37 °C. Após a adição da transcriptase inversa (1 µl), teve lugar a síntese de cADN durante 60 minutos a 42 °C, antes de a reacção ser parada por meio de aquecimento durante 5 minutos a 70 °C.

Condições de amplificação do fragmento R3 de 3,2 kb por meio de PCR:

2 µL da mistura reaccional da síntese da primeira cadeia

0,6 µM      Iniciador F1

0,6 µM      Iniciador R4

0,4 mM      Mistura dNTP

3 mM    MgSO<sub>4</sub>

1x tampão de reacção

1,25 unidades de Platinum Taq DNA Polymerase High Fidelity

Condições de reacção:

Etapa 1      95 °C      2 min

Etapa 2      95 °C      45 seg

Etapa 3      61,7 °C    45 seg

Etapa 4      68 °C      3 min 30 seg + 1 seg/ciclo

Repetir 35 vezes a partir da etapa 2

Depois 4 °C

Depois de separada a mistura reaccional PCR por meio de um TBE-gel de agarose a 0,8 %, o fragmento obtido foi

excisado utilizando uma lâmina cirúrgica e o ADN foi isolado por meio de kit de extracção gel (Quiagen; Prod. No.: 28704). Seguidamente, o fragmento que tinha sido isolado foi cortado com as endonucleases de restrição *Nco I* e *Eco RI* e clonado no vector de clonagem pMCS5 (MobiTec, Prod. No.: pMCS5) que tinha sido cortado com as mesmas endonucleases de restrição. O plasmídeo resultante foi designado pRS79-191. A análise da sequência do produto PCR pRS79-191, que tinha sido clonado no plasmídeo pRS79-191, revelou que o fragmento continha uma inserção que resultou na interrupção da sequência de codificação. A comparação correspondente com a sequência correspondente da base de dados NCBI e uma análise da sequência inserida sugeriu que a inserção poderia ser um intrão que não tinha sido excisado. Dado que não havia a certeza de se poder sujeitar este intrão putativo a *splicing* em plantas transgénicas, as etapas de clonagem seguintes envolveram uma troca nesta região por uma sequência "isenta de inserções", restaurando-se assim a estrutura de leitura contínua da proteína R3 de *Arabidopsis thaliana*.

Para complementar a sequência de codificação da proteína R3 de *Arabidopsis thaliana*, a região 3' em falta da sequência de codificação anteriormente mencionada foi amplificada nas condições de reacção pormenorizadas mais adiante, a começar em 300 ng de ARN poli-A-plus, com ajuda dos iniciadores F4 e R2:

O cADN foi sintetizado como anteriormente descrito, exceptuando ter-se utilizado um oligo dT (parte do kit RT) em vez de um iniciador R2.

A amplificação foi efectuada nas seguintes condições:

2 µL da mistura reaccional da síntese da primeira cadeia

0,3 µM      Iniciador F4

0,3 µM      Iniciador R2

0,4 mM      Mistura dNTP

2 mM    MgSO<sub>4</sub>

1x tampão de reacção

1,25 unidades de Platinum Taq DNA Polymerase High Fidelity

Condições de reacção:

Etapa 1      95 °C      2 min

Etapa 2      95 °C      30 seg

Etapa 3      61 °C      1 min

Etapa 4      68 °C      1 min 45 sec

Repetir 40 vezes a partir da etapa 2

Depois 4 °C

Depois de separada a mistura PCR por meio de um TBE-gel de agarose a 0,8 %, o fragmento de 1,7 kb resultante foi excisado do gel e o ADN foi isolado por meio de kit de extracção gel (Quiagen; Prod. No.: 28704). Depois do tratamento com as endonucleases de restrição *Xho I* e *Spe I*, o plasmídeo foi clonado no plasmídeo pRS79-191, que tinha sido cortado com as mesmas endonucleases de restrição. O plasmídeo resultante foi designado pRS80-227.

3. Sequências de fusão tradicionais que codificam uma proteína R3 com sequências que codificam um peptídeo de sinal plastídeo

Primeiro, a sequência codificadora do peptídeo de sinal plastídeo de *Oryza sativa* (variedade M202) gene *dull-1* foi amplificada por meio de PCR utilizando um clone de cADN (pRS55-132) como modelo, utilizando os iniciadores SP-Mac fwd (5'-CGA TAT CTG CAG GTT TTG GCA ATG GAG AT-3') e SP-Mac-rev (5'-CTT CGA ACA ATT CTG CTT CCA G-3') e foi clonada no vector pCR 2.1 (Invitrogen. Prod. No.: K2000-01). O plasmídeo resultante foi designado plC328-215. O clone de cADN pRS55-132 foi isolado por meio de hibridação de uma biblioteca de cADN fágica, com uma sonda radioactiva que compreende as sequências codificadoras de uma proteína de milho *dul I* (EMBL Accession: T01265; Plant Cell 10; 399-412 (1998)). Utilizando o plasmídeo plC328-215 como modelo, amplificou-se o peptídeo de sinal plastídeo do gene *dul I* de *Oryza sativa* foi amplificado com ajuda dos iniciadores Os\_dul-1\_SP-Pac (5'-GAC CTT AAT TAA GGT TTT GGC AAT GGA GAT G-3') e Os\_dul-1\_SP\_Asu (5'-GCC AGG TTT CGA ACA ATT CTG CTT CGA-3'). O iniciador 5' (Os\_dul-1\_SP-Pac) incluía um local de clivagem de restrição *Pac I* e o iniciador 3' (Os\_dul-1\_SP\_Asu) incluía um local de clivagem de restrição *Bst BI*. A amplificação foi efectuada nas seguintes condições:

10 ng plC328-215

0,2  $\mu$ M        Iniciador OS\_dul1\_SP-Pac

0,2  $\mu$ M        Iniciador Os\_dul1\_SP-Asu

0,4 mM        Mistura dNTP

3 mM    MgSO<sub>4</sub>

1x tampão de reacção

1,25 unidades de Platinum Taq DNA Polymerase High Fidelity

Condições de reacção:

Etapa 1	95 °C	4 min
Etapa 2	95 °C	30 seg
Etapa 3	60 °C	30 seg
Etapa 4	68 °C	20 seg

Repetir 40 vezes a partir da etapa 2

Depois 4 °C

Depois da digestão com o fragmento PCR resultante com as endonucleases de restrição *Pac I* e *Est BI*, o fragmento PCR foi clonado no vector pMCS5 (MobiTec, Prod. No.: pMCS5) que tinha sido cortado com as mesmas endonucleases de restrição. O plasmídeo resultante foi designado pAH41-227. Seguidamente, a sequência que codifica a proteína R3 de *Arabidopsis thaliana* foi amplificada por meio de PCR, utilizando o plasmídeo pRS80-191 como modelo e os iniciadores At\_R3\_Asu (5'-AAT TGT TCG AAA CCT GGC AAC CTC TAA ATC CC-3') e R2 (c.f. acima). A sequência do iniciador 5' foi seleccionada de forma a que o codão de partida da sequência, que codifica a proteína R3, não está presente na sequência amplificada por meio de PCR, tendo sido introduzido um local de clivagem de restrição *Est BI*. A sequência R2 do iniciador 3' compreendia um local de clivagem de restrição *Xho I*.

A amplificação foi efectuada nas seguintes condições:

20 ng pRS80-227

0,3 µM Iniciador At\_R3\_Asu

0,3 µM Iniciador R2

0,4 mM Mistura dNTP

3 mM MgSO<sub>4</sub>

1x tampão de reacção

1,25 unidades de Platinum Taq DNA Polymerase High Fidelity

Condições de reacção

Etapa 1 95° C 2 min

Etapa 2 95° C 50 seg

Etapa 3 60° C 50 seg

Etapa 4 68° C 4 min

Repetir 40 vezes a partir da etapa 2

Depois 4 °C

O produto PCR resultante foi cortado com as endonucleases *Bst* *BI* e *Xho* *I* e clonado no plasmídeo pAH41-227 que tinha sido cortado com *Bst* *BI* e *Sal* *I*, o que deu origem ao plasmídeo AH42-227. Deste modo, foi obtida uma sequência em que o peptídeo de sinal plastídeo do gene *dul* *I* de *Oryza sativa* e a região codificadora da proteína R3 de *Arabidopsis thaliana* foram uma estrutura de leitura aberta partilhada, de forma que, depois da transcrição e tradução, obtém-se uma proteína R3 fundida traducionalmente que possui um peptídeo de sinal plastídeo.

#### 4. Geração do vector de expressão vegetal pAH43-227

A sequência codificadora da proteína R3 de *Arabidopsis thaliana*, que foi fundida à sequência codificadora do peptídeo de sinal plastídeo do gene *dul* *I* de *Oryza sativa* foi excisada do plasmídeo pAH42-227 com ajuda das endonucleases de restrição *Pac* *I* e *Avr* *II* e clonada no

plasmídeo pML80-123 que tinha sido cortado com as mesmas nucleases de restrição.

5. Geração do vector de expressão vegetal plR138-210, na qual foram removidas sequências não codificadoras da região codificadora da proteína R3

O fragmento que está presente na região codificadora da proteína R3 *Arabidopsis thaliana* e que compreende sequências não codificadoras foi removido do plasmídeo pAH43-227. Com este objectivo, uma sequência parcial da região codificadora da proteína R3 de *Arabidopsis thaliana* foi amplificada por meio de RT-PCR, a começar no mRNA isolado a partir de folhas de plantas de *Arabidopsis thaliana* com aproximadamente três semanas de idade (ecotipo C24).

Foram utilizados os seguintes iniciadores:

Nome	Sequência
F1	AAGCTGCCATGGCAACCTCTAAATC
R5	CCAAGATATTTTGCAGTAGGCTGT

A síntese da primeira cadeia foi efectuada como anteriormente descrito (exemplo, ponto 2). O fragmento desejado foi amplificado nas seguintes condições:

2 µL da mistura reaccional da síntese da primeira cadeia

0,3 µM	Iniciador F1
0,3 µM	Iniciador R5
0,4 mM	Mistura dNTP
3 mM	MgSO <sub>4</sub>

1x tampão de reacção

1,25 unidades de Platinum Taq DNA Polymerase High Fidelity

Condições de reacção

Etapa 1	95° C	2 min
Etapa 2	95° C	1 min
Etapa 3	61° C	1 min
Etapa 4	68° C	2 min 20 sec

Repetir 40 vezes a partir da etapa 2

Depois 4 °C

Uma vez que a inserção presente nos plasmídeos pRS79-191, pRS80-227, AH42-227 e AH43-227 (inirão putativo) possui um local de clivagem de restrição *Ssp I*, a sequência parcial da região codificadora da proteína R3 de *Arabidopsis thaliana*, que foi amplificada por meio de PCR, foi primeiro digerida com a endonuclease de restrição *Ssp I* e os fragmentos por digerir (sem inserção) foram depois purificados por meio de um gel de agarose. Em seguida, o fragmento purificado foi cortado com as endonucleases de restrição *Bam HI* e *Spe I* e trocado pelo fragmento correspondente *Bam HI/Spe I* (compreendendo, entre outras, as sequências não codificadoras) do plasmídeo pAH43-227. O plasmídeo resultante foi designado R138-210.

6. Geração do vector de expressão vegetal pIR139-210, no qual a sequência de codificação da proteína R3 não possui qualquer peptídeo de sinal plastídeo

Utilizando o plasmídeo pIR 138-210 como modelo, amplifiquei-se uma sequência parcial codificadora da região

5' da proteína R3 de *Arabidopsis thaliana*, por meio de PCR com ajuda dos iniciadores IR138-F1 (5'-AAT TAA CTT AAT TAA TCC ATG GCA ACC TCT AAA TCC CAA CAA T-3') e IR138-R2 (5'-AAT TAC TTG GAT CCC GCA ACT CAA GGA TCA CTA CAT ATG CA-3'). O iniciador 5' (IR138-F1) incluía um local de clivagem de restrição *Pac I* e a sequência para o codão de partida original da proteína R3. O iniciador 3' (IR138-R2) compreendia um local de reconhecimento da endonuclease de restrição *BamHI*. A amplificação foi efectuada nas seguintes condições:

50 ng plasmídeo IR138-210

0,5 µM        Iniciador IR138-F1

0,5 µM        Iniciador I R138-R2

2 mM    MgSO<sub>4</sub>

0,3 mM        Mistura dNTP

1 x    tampão de reacção

1,5 unidades de Platinum Taq DNA Polymerase High Fidelity

Condições de reacção

Etapa 1        95 °C        2 min

Etapa 2        94 °C        30 seg

Etapa 3        60 °C        20 seg

Etapa 4        68 °C        20 seg

Repetir 30 vezes a partir da etapa 2

Depois 4 °C

O fragmento de 247 bp amplificado foi cortado com as endonucleases de restrição *Pac I* e *Bam HI* e foi substituído

pelo fragmento *Pac I/Bam HI*, que está presente no plasmídeo pIR138-210 e que compreende o peptídeo de sinal plastídeo do gene *dull 1* de *Oryza sativa* e parte da sequência 5' que codifica a proteína R3 de *Arabidopsis thaliana*. O plasmídeo resultante foi designado pIR139-210. A sequência codificadora da proteína R3 de *Arabidopsis thaliana* inclui o codão de partida que originalmente é do cADN, em contraste com a sequência que compreende no plasmídeo pIR138-210 e AH43- 227.

#### 7. Transformação de plantas do arroz

Plantas do arroz (variedade M202) foram transformadas com o plasmídeo pIRI 38-210 ou com o plasmídeo pIR139-201 por meio de agrobactérias (compreendendo o plasmídeo pIR138-210 ou pIR239-210) utilizando o método descrito por Hiei et al. (1994, *Plant Journal* 6(2), 271-282).

#### 8. Análise de plantas do arroz transgênicas

a) Plantas do arroz que foram transformadas com o plasmídeo pIR138-210 e que expressaram uma proteína R3 de *Arabidopsis thaliana* em fusão com um peptídeo de sinal plastídeo. Foi utilizada análise RT-PCR quantitativa para a identificação de plantas com expressão de mRNA codificador da proteína R3 de *Arabidopsis thaliana*. As plantas que, em comparação com as plantas de tipo selvagem correspondente, compreende uma quantidade detectável de mRNA codificador da proteína R3 de *Arabidopsis thaliana* foram cultivadas em estufa. Os grãos dessas plantas foram colhidos. O amido isolado desses grãos maduros revelou um teor mais elevado em fosfato que estava ligado covalentemente ao amido em questão, em comparação com o amido isolado dos grãos de plantas de tipo selvagem correspondentes.

b) Plantas do arroz que foram transformadas com o plasmídeo pIR139-210 e expressam uma proteína R3 de *Arabidopsis thaliana*

Utilizando a análise quantitativa RT-PCR, foi possível identificar as plantas que expressam o mRNA codificador da proteína R3 de *Arabidopsis thaliana*. As plantas que, em comparação com as plantas de tipo selvagem correspondente, compreende uma quantidade detectável de mRNA codificador da proteína R3 de *Arabidopsis thaliana* foram cultivadas em estufa. Os grãos dessas plantas foram colhidos. O amido isolado desses grãos maduros não revelou um teor significativamente mais elevado em fosfato que estava ligado covalentemente ao amido em questão, em comparação com o amido isolado dos grãos de plantas de tipo selvagem correspondentes.

<110> Bayer CropScience GmbH

<120> Plantas com um aumento da actividade de uma enzima de fosforilação do amido nos plastídeos

<130> BCS 04-5008-PCT

<150> EP 04090316

<151> 2001-08-18

<150> US 60/602,454

<151> 2004-08-18

<160> 4

<170> PatentIn version 3.1

<210> 1

<211> 3837

<212> ADN

<213> *Arabidopsis thaliana*

<400> 1

atggcaacct	ctaaatccca	acaattccag	ctaattgaag	ggatggagct	tcagatcacc	60
gtcactggat	tgccgaatgg	gagcagtgta	agagctgagt	ttcacctgaa	gaactgcact	120
cgcgcatgga	ttcttcattg	gggttgatt	taccaaggaa	ataaccattg	gtatatccca	180
tctgaacatt	cctcgaagca	aggtgcattg	cagactacct	ttgtgaagag	tggggatgca	240
tatgtagtga	tccttgagtt	gcgggatcca	agggtacgag	caattgaatt	tgttttgaag	300
gatggtagcc	ataacaggtg	gttgagacag	cataatggaa	acttccgtgt	tgagattccc	360
tggaatgatc	tccatgctca	tcatcggata	ccgaagactc	tgatagaaag	aagagcacat	420
aagatatggg	accggaaggg	acgaccacaa	agctctgcac	gtgaacaaca	gatagactat	480
gacaatgcag	taagagagct	ccacgccgag	ctcgttagag	gaatatcttt	agacgaactt	540
caggctaact	ctacagtacc	agtgagagaaa	gaagtaacca	gtgagccaca	tcacacaatg	600
atccaatcat	atcgccggag	gcatgatggt	cagaaatggt	tacagaaata	tacggaacca	660

atcaacagaa	gtggaagtgt	gaaaagttca	gcccttgag	aactctcca	gagatctgtg	720
ggccaagaaa	atctagtttc	acagaaaagc	tttcatgtca	gaaactacga	gatcacggtc	780
ctccaaaggg	atgtaaggg	agattgtcgc	ttatggattg	ccacgaacat	ggcaggtcca	840
acagttctcc	attggggagt	cgcaaagtca	tctgcaggag	agtggttgat	accaccacct	900
gacgtgttac	ctgagaaatc	aaaatttgtt	cacggagcat	gtcaaacaca	atctaccgat	960
atgtcgagca	gagaacactc	ttatcagttt	attgatataa	atctaaaacg	aggtggtttt	1020
gttggatcc	aatttgaat	atggtcggga	ggctactggg	tgaacaataa	cggagccaac	1080
ttcgtttaa	acctgaagtc	agcagacagt	accagtggta	agcttgacgt	ggatgaaaag	1140
tatgttctca	aatggttact	tgatgaaata	tctgagcgag	agaaagaggc	tgagagatca	1200
ttgatgcaca	ggttcaacat	tgcaacagag	ttgaccgagc	gttgtaagga	tgaaggagaa	1260
ggtggatgta	ttggtataat	ggtgtggatg	agattcatgg	ccaccagaca	tctcacctgg	1320
aataagaact	ataacgtgaa	acctcgggaa	attagtgaag	cactagaaag	attaccaat	1380
ttgatggaga	aaatatattt	gcagcaacca	aataagagag	aaattgtgag	actaactatg	1440
gcacttgtgg	gtcgtggagg	tcaaggtgat	gttgggcaga	gaatccgtga	cgaaatcctt	1500
gttattcaga	gacataacca	ctgcaaatcc	ggcatgatgg	aagagtggca	ccaaaagttg	1560
cacaacaaca	gtagcgaga	tgatgtgata	atctgtgagg	ctctcttgaa	ctatgtgaga	1620
tctgacttca	ggatcgatgc	atactggcaa	acactacaga	ccaatggtct	cacaaaagaa	1680
aggcttgcaa	gttatgaccg	ccccatagta	tcagagcctc	gtttcagaag	tgattctaaa	1740
gaaggactca	tccgtgacct	tacaatgtac	ttgaaaacat	taaaggcagt	tcattcaggt	1800
gcagaccttg	agtctgctat	tgatacgttc	ctttctccat	ctaagggtca	tcattgcttt	1860
gctgtcaatg	gtttatcacc	aaaattgcag	gatctactga	atctagttaa	gaggcttgtt	1920
cgcgaagaga	acactgagcc	cctaatagag	aagttagttg	atgctcgcat	tcagttacac	1980
cctgcactac	gagcacctcg	cacgagggca	aaagatttac	tatttttggg	cattgctgtg	2040
gagtcatggt	ttaaaacaac	aatagagaag	agactcatct	ctttaaactt	caataacca	2100
ccgaaatta	tatatgtcat	ctgctgggtg	cttgagaatc	tgtgcttata	tatagttaac	2160
aatgaagaaa	tcataattctg	tacaaaggat	tggtaccgtg	tcagcgaggc	ttacagacct	2220
catgatgttc	agtgggcatt	gcaaacaaaa	gcggtccttg	accgtttaca	actagtactt	2280
gctgatagat	gtcagcatta	ttttacaata	atacagccta	ctacaaaata	tcttgggtcaa	2340
ctgttacgtg	ttgacaagca	tgggattgat	gttttcaactg	aagaggttat	aagagcaggg	2400
ccaggagccg	ttttatcaac	tcttgtaaac	agatttgatc	cttctctaag	gaaaatcgcg	2460
aatttaggct	gttggcaggt	tatcagcctg	gctgatgcat	atgggtttgt	ggtttgtgtg	2520
aatgagttaa	ttgttgttca	gaacaaatc	tactcgaagc	caactgtaat	tattgcaagt	2580
aaagtcacag	gagaagagga	gatccctgct	ggtgtttag	ccgtgctgac	tcctagatg	2640
attgatgtct	tgtctcatgt	atctattaga	gcaagaaca	gcaagatatg	ctttgctaca	2700

tgcttcgadc	agaatgttct	cagcaatctg	aagtcgaagg	aaggaagagc	aatatctatt	2760
catacaaagt	ctactggttt	agttatcagt	gatggcaaca	actctggtgt	ctctgttcgt	2820
catattttta	tttctctgt	tccacgggga	gtgatctcta	aggggaagaa	gttttgtggc	2880
cactatgtga	tttcatctaa	ggaattcaca	gatgaaaggg	ttggctcaaa	atcatacaat	2940
ataaaatttc	tacgtgaaag	agttccatca	tggatcaaga	tacctacctc	agctgctctt	3000
ccattcggaa	catttgagaa	tatactctca	gatgattcca	ataaggatgt	agcacgcagg	3060
atttctgtcc	ttaaagattc	tcttaacaga	ggagacctga	caaaacttaa	gtccattcaa	3120
gaagctatct	tacaaatgag	tgctccaatg	gctctgagaa	atgaactgat	cacaaaattg	3180
agaagtgaaa	gaatgcctta	tcttgggtgat	gaatcaggct	ggaaccgctc	ctgggtggca	3240
attaagaagg	tctgggcttc	aaagtggaac	gagagagctt	atgtcagttg	caaaaaaat	3300
aagcttgatc	atgatgcggt	ctgcatggct	gtgctgattc	aagaagtcac	ttgtggcgat	3360
tatgctttcg	tcattcacac	aaacaatccg	gtctctgggt	actcttcaga	aatatacaca	3420
gagattgtga	agggtttggg	agagaccttg	gttggagcat	atccaggacg	agcaatgagc	3480
ttcatcacca	agaaaacaaa	cctcaagtct	ccaaccgtga	tcagttacc	aagtaagagg	3540
ataggctctg	attctaaacc	ctcaatcata	ttcagatcag	attcaaacaa	cgaggatctc	3600
gaaggcaacg	caggcgttgg	actttacgac	agtgtgataa	tggatgaagc	agaggaagta	3660
gtgggtggatt	actcaagggg	gccactaata	atggacaaat	cctttcgagt	gcgtctattc	3720
tcagcgattg	cagaagctgg	aaatgtgata	gagtcaatct	atggttgtcc	tcaagacatt	3780
gaagggtgtt	tcaaaggtgg	acatatctac	atcgtccaag	ctagacccca	agtttaa	3837

<210> 2

<211> 3987

<212> ADN

<213> Sequência artificial

<220>

<230> Sequência contendo a região codificadora da sequência de sinal plastídeo do gene *dul-I* de *Oryza sativa* fundido ao cADN codificador do gene R3 de *Arabidopsis thaliana*

<220>

<221> CDS

<222> (1)..(3984)

<223> sequência codificadora da sequência artificial

<220>

<221> sig\_peptide

<222> (1)..(147)

<223> sequência codificadora do peptídeo de sinal  
plastídeo do gene *dul I* de *Oryza sativa*

<220>

<221> mat\_peptide

<222> (154)..()

<223> sequência codificadora do peptídeo maturo do gene  
R3 (GWD-3) de *Arabidopsis thaliana* sem a metionina do codão  
de partida natural

<400> 2

atg gag atg gct ctc cgg cct caa agc ctt cta tgc cct cgg agc agg	48
Met Glu Met Ala Leu Arg Pro Gln Ser Leu Leu Cys Pro Arg Ser Arg	
-50 -45 -40	
ctg aag gtc gtc atc cgg cca gct agc agt gcc agc ggt ggt ggc ctc	96
Leu Lys Val Val Ile Arg Pro Ala Ser Ser Ala Ser Gly Gly Gly Leu	
-35 -30 -25 -20	
gcg cag tat ttc tta atg acc agg aga tat act gga agc aga att gtt	144
Ala Gln Tyr Phe Leu Met Thr Arg Arg Tyr Thr Gly Ser Arg Ile Val	
-15 -10 -5	
cga aac ctg gca acc tct aaa tcc caa caa ttc cag cta att gaa ggg	192
Arg Asn Leu Ala Thr Ser Lys Ser Gln Gln Phe Gln Leu Ile Glu Gly	
-1 1 5 10	
atg gag ctt cag atc acc gtc act gga ttg ccg aat ggg agc agt gta	240
Met Glu Leu Gln Ile Thr Val Thr Gly Leu Pro Asn Gly Ser Ser Val	
15 20 25	
aga gct gag ttt cac ctg aag aac tgc act cgc gca tgg att ctt cat	288
Arg Ala Glu Phe His Leu Lys Asn Cys Thr Arg Ala Trp Ile Leu His	
30 35 40 45	
tgg ggt tgt att tac caa gga aat aac cat tgg tat atc cca tct gaa	336
Trp Gly Cys Ile Tyr Gln Gly Asn Asn His Trp Tyr Ile Pro Ser Glu	
50 55 60	
cat tcc tcg aag caa ggt gca ttg cag act acc ttt gtg aag agt ggg	384
His Ser Ser Lys Gln Gly Ala Leu Gln Thr Thr Phe Val Lys Ser Gly	
65 70 75	
gat gca tat gta gtg atc ctt gag ttg ccg gat cca agg gta cgc gca	432
Asp Ala Tyr Val Val Ile Leu Glu Leu Arg Asp Pro Arg Val Arg Ala	
80 85 90	
att gaa ttt gtt ttg aaa gat ggt agc cat aac agg tgg ttg aga cag	480
Ile Glu Phe Val Leu Lys Asp Gly Ser His Asn Arg Trp Leu Arg Gln	
95 100 105	
cat aat gga aac ttc cgt gtt gag att ccc tgg aat gat ctc cat gct	528
His Asn Gly Asn Phe Arg Val Glu Ile Pro Trp Asn Asp Leu His Ala	
110 115 120 125	
cat cat cgg ata ccg aag act ctg ata gaa aga aga gca cat aag ata	576
His His Arg Ile Pro Lys Thr Leu Ile Glu Arg Arg Ala His Lys Ile	
130 135 140	
tgg gac cgg aag gga cga cca caa agc tct gca cgt gaa caa cag ata	624
Trp Asp Arg Lys Gly Arg Pro Gln Ser Ser Ala Arg Glu Gln Gln Ile	

145					150					155						
gac Asp	tat Tyr	gac Asp 160	aat Asn	gca Ala	gta Val	aga Arg	gag Glu 165	ctc Leu	cac His	gcc Ala	gag Glu	ctc Leu 170	gct Ala	aga Arg	gga Gly	672
ata Ile	tct Ser 175	tta Leu	gac Asp	gaa Glu	ctt Leu	cag Gln 180	gct Ala	aac Asn	tct Ser	aca Thr	gta Val 185	cca Pro	gtg Val	gag Glu	aaa Lys	720
gaa Glu 190	gta Val	acc Thr	agt Ser	gag Glu	cca Pro 195	cat His	cac His	aca Thr	atg Met	atc Ile 200	caa Gln	tca Ser	tat Tyr	cgc Arg	cgg Arg 205	768
agg Arg	cat His	gat Asp	gtt Val	cag Gln 210	aaa Lys	tgg Trp	tta Leu	cag Gln	aaa Lys 215	tat Tyr	acg Thr	gaa Glu	cca Pro	atc Ile 220	aac Asn	816
aga Arg	agt Ser	gga Gly	agt Ser 225	gtg Val	aaa Lys	agt Ser	tca Ser	gcc Ala 230	ctt Leu	gca Ala	gaa Glu	ctc Leu	tcc Ser 235	aag Lys	aga Arg	864
tct Ser	gtg Val	ggc Gly 240	caa Gln	gaa Glu	aat Asn	cta Leu	gtt Val 245	tca Ser	cag Gln	aaa Lys	agc Ser	ttt Phe 250	cat His	gtc Val	aga Arg	912
aac Asn	tac Tyr 255	gag Glu	atc Ile	acg Thr	gtc Val	ctc Leu 260	caa Gln	agg Arg	gat Asp	gtt Val	aag Lys 265	gga Gly	gat Asp	tgt Cys	cgc Arg	960
tta Leu 270	tgg Trp	att Ile	gcc Ala	acg Thr	aac Asn 275	atg Met	gca Ala	ggt Gly	cca Pro	aca Thr 280	gtt Val	ctc Leu	cat His	tgg Trp	gga Gly 285	1008
gtc Val	gca Ala	aag Lys	tca Ser 290	tct Ser	gca Ala	gga Gly	gag Glu	tgg Trp	ttg Leu 295	ata Ile	cca Pro	cca Pro	cct Pro	gac Asp 300	gtg Val	1056
tta Leu	cct Pro	gag Glu	aaa Lys 305	tca Ser	aaa Lys	ttt Phe	gtt Val	cac His 310	gga Gly	gca Ala	tgt Cys	caa Gln	aca Thr 315	caa Gln	ttt Phe	1104
acc Thr	gat Asp	atg Met 320	tcg Ser	agc Ser	aga Arg	gaa Glu	cac His 325	tct Ser	tat Tyr	cag Gln	ttt Phe 330	att Ile	gat Asp	ata Ile	aat Asn	1152
tta Leu	aaa Lys 335	cga Arg	ggt Gly	ggt Gly	ttt Phe 340	ggt Val 340	ggt Gly	atc Ile	caa Gln	ttt Phe 345	gta Val 345	ata Ile	tgg Trp	tcg Ser	gga Gly	1200
ggc Gly 350	tac Tyr	tgg Trp	gtg Val	aac Asn	aat Asn 355	aac Asn	gga Gly	gcc Ala	aac Asn	ttc Phe 360	gtt Val	gta Val	aac Asn	ctg Leu 365	aag Lys	1248
tca Ser	gca Ala	gac Asp	agt Ser	acc Thr 370	agt Ser	ggt Gly	aag Lys	ctt Leu	gac Asp 375	gtg Val	gat Asp	gaa Glu	aag Lys	tat Tyr 380	gtt Val	1296
ctc Leu	aaa Lys	tgg Trp	tta Leu 385	ctt Leu	gat Asp	gaa Glu	ata Ile	tct Ser 390	gag Glu	cga Arg	gag Glu	aaa Lys	gag Glu 395	gct Ala	gag Glu	1344
aga Arg	tca Ser	ttg Leu 400	atg Met	cac His	agg Arg	ttc Phe	aac Asn 405	att Ile	gca Ala	aca Thr	gag Glu	ttg Leu 410	acc Thr	gag Glu	cgt Arg	1392
tgt Cys	aag Lys 415	gat Asp	gaa Glu	gga Gly	gaa Glu	ggt Gly 420	gga Gly	tgt Cys	att Ile	ggt Gly	ata Ile 425	atg Met	gtg Val	tgg Trp	atg Met	1440

aga	ttc	atg	gcc	acc	aga	cat	ctc	acc	tgg	aat	aag	aac	tat	aac	gtg	1488
Arg	Phe	Met	Ala	Thr	Arg	His	Leu	Thr	Trp	Asn	Lys	Asn	Tyr	Asn	Val	
430					435					440					445	
aaa	cct	cgg	gaa	att	agt	gaa	gca	cta	gaa	aga	ttc	acc	aat	ttg	atg	1536
Lys	Pro	Arg	Glu	Ile	Ser	Glu	Ala	Leu	Glu	Arg	Phe	Thr	Asn	Leu	Met	
				450					455					460		
gag	aaa	ata	tat	ttg	cag	caa	cca	aat	aag	aga	gaa	att	gtg	aga	cta	1584
Glu	Lys	Ile	Tyr	Leu	Gln	Gln	Pro	Asn	Lys	Arg	Glu	Ile	Val	Arg	Leu	
			465					470					475			
act	atg	gca	ctt	gtg	ggg	cgt	gga	ggg	caa	ggt	gat	ggt	ggg	cag	aga	1632
Thr	Met	Ala	Leu	Val	Gly	Arg	Gly	Gly	Gln	Gly	Asp	Val	Gly	Gln	Arg	
		480				485						490				
atc	cgt	gac	gaa	atc	ctt	gtt	att	cag	aga	cat	aac	cac	tgc	aaa	tcc	1680
Ile	Arg	Asp	Glu	Ile	Leu	Val	Ile	Gln	Arg	His	Asn	His	Cys	Lys	Ser	
	495					500					505					
ggc	atg	atg	gaa	gag	tgg	cac	caa	aag	ttg	cac	aac	aac	agt	agc	gca	1728
Gly	Met	Met	Glu	Glu	Trp	His	Gln	Lys	Leu	His	Asn	Asn	Ser	Ser	Ala	
510					515					520					525	
gat	gat	gtg	ata	att	tgt	gag	gct	ctc	ttg	aac	tat	gtg	aga	tct	gac	1776
Asp	Asp	Val	Ile	Ile	Cys	Glu	Ala	Leu	Leu	Asn	Tyr	Val	Arg	Ser	Asp	
			530						535					540		
ttc	agg	atc	gat	gca	tac	tgg	caa	aca	cta	cag	acc	aat	ggt	ctc	aca	1824
Phe	Arg	Ile	Asp	Ala	Tyr	Trp	Gln	Thr	Leu	Gln	Thr	Asn	Gly	Leu	Thr	
			545					550					555			
aaa	gaa	agg	ctt	gca	agt	tat	gac	cgc	ccc	ata	gta	tca	gag	cct	cgt	1872
Lys	Glu	Arg	Leu	Ala	Ser	Tyr	Asp	Arg	Pro	Ile	Val	Ser	Glu	Pro	Arg	
		560					565					570				
ttc	aga	agt	gat	tct	aaa	gaa	gga	ctc	atc	cgt	gac	ctt	aca	atg	tac	1920
Phe	Arg	Ser	Asp	Ser	Lys	Glu	Gly	Leu	Ile	Arg	Asp	Leu	Thr	Met	Tyr	
	575					580				585						
ttg	aaa	aca	tta	aag	gca	gtt	cat	tca	ggt	gca	gac	ctt	gag	tct	gct	1968
Leu	Lys	Thr	Leu	Lys	Ala	Val	His	Ser	Gly	Ala	Asp	Leu	Glu	Ser	Ala	
590					595					600					605	
att	gat	acg	ttc	ctt	tct	cca	tct	aag	ggt	cat	cat	gtc	ttt	gct	gtc	2016
Ile	Asp	Thr	Phe	Leu	Ser	Pro	Ser	Lys	Gly	His	His	Val	Phe	Ala	Val	
				610					615					620		
aat	ggt	tta	tca	cca	aaa	ttg	cag	gat	cta	ctg	aat	tta	ggt	aag	agg	2064
Asn	Gly	Leu	Ser	Pro	Lys	Leu	Gln	Asp	Leu	Leu	Asn	Leu	Val	Lys	Arg	
			625					630					635			
ctt	ggt	cgc	gaa	gag	aac	act	gag	ccc	cta	ata	gag	aag	tta	ggt	gat	2112
Leu	Val	Arg	Glu	Glu	Asn	Thr	Glu	Pro	Leu	Ile	Glu	Lys	Leu	Val	Asp	
		640					645					650				
gct	cgc	att	cag	tta	cac	ctt	gca	cta	cga	gca	cct	cgc	acg	agg	gca	2160
Ala	Arg	Ile	Gln	Leu	His	Pro	Ala	Leu	Arg	Ala	Pro	Arg	Thr	Arg	Ala	
	655					660					665					
aaa	gat	tta	cta	ttt	ttg	gac	att	gcg	ttg	gag	tca	tgt	ttt	aaa	aca	2208
Lys	Asp	Leu	Leu	Phe	Leu	Asp	Ile	Ala	Leu	Glu	Ser	Cys	Phe	Lys	Thr	
670					675					680					685	
aca	ata	gag	aag	aga	ctc	atc	tct	tta	aac	ttc	aat	aac	cca	ccg	gaa	2256
Thr	Ile	Glu	Lys	Arg	Leu	Ile	Ser	Leu	Asn	Phe	Asn	Asn	Pro	Pro	Glu	
				690					695					700		

att Ile	ata Ile	tat Tyr	gtc Val 705	atc Ile	tgc Cys	gtg Val	gtg Val	ctt Leu 710	gag Glu	aat Asn	ctg Leu	tgc Cys 715	tta Leu 715	tct Ser	ata Ile	2304
gtt Val	aac Asn	aat Asn 720	gaa Glu	gaa Glu	atc Ile	ata Ile	ttc Phe 725	tgt Cys	aca Thr	aag Lys	gat Asp	tgg Trp 730	tac Tyr	cgf Arg	gtc Val	2352
agc Ser	gag Glu 735	gct Ala	tac Tyr	aga Arg	cct Pro	cat His 740	gat Asp	ggt Val	cag Gln	tgg Trp	gca Ala 745	ttg Leu	caa Gln	aca Thr	aaa Lys	2400
gcg Ala 750	gtc Val	ctt Leu	gac Asp	cgf Arg	tta Leu 755	caa Gln	cta Leu	gta Val	ctt Leu	gct Ala 760	gat Asp	aga Arg	tgt Cys	cag Gln	cat His 765	2448
tat Tyr	ttt Phe	aca Thr	ata Ile	ata Ile 770	cag Gln	cct Pro	act Thr	aca Thr	aaa Lys 775	tat Tyr	ctt Leu	ggt Gly	caa Gln	ctg Leu 780	tta Leu	2496
cgf Arg	ggt Val	gac Asp	aag Lys 785	cat His	ggg Gly	att Ile	gat Asp	ggt Val 790	ttc Phe	act Thr	gaa Glu	gag Glu	ggt Val 795	ata Ile	aga Arg	2544
gca Ala	ggg Gly	cca Pro 800	gga Gly	gcc Ala	ggt Val	tta Leu	tca Ser 805	act Thr	ctt Leu	gta Val	aac Asn	aga Arg 810	ttt Phe	gat Asp	cct Pro	2592
tct Ser	cta Leu 815	agg Arg	aaa Lys	atc Ile	gcg Ala	aat Asn 820	tta Leu	ggc Gly	tgt Cys	tgg Trp	cag Gln 825	ggt Val	atc Ile	agc Ser	ctg Leu	2640
gct Ala 830	gat Asp	gca Ala	tat Tyr	ggg Gly	ttt Phe 835	gtg Val	ggt Val	tgt Cys	gtg Val	aat Asn 840	gag Glu	tta Leu	att Ile	ggt Val	ggt Val 845	2688
cag Gln	aac Asn	aaa Lys	ttc Phe	tac Tyr 850	tgc Ser	aag Lys	cca Pro	act Thr	gta Val 855	att Ile	att Ile	gca Ala	agt Ser 860	aaa Lys	gtc Val	2736
aca Thr	gga Gly	gaa Glu	gag Glu 865	gag Glu	atc Ile	cct Pro	gct Ala	ggt Gly 870	ggt Val	gta Val	gcc Ala	gtg Val	ctg Leu 875	act Thr	cct Pro	2784
agt Ser	atg Met	att Ile 880	gat Asp	gtc Val	ttg Leu	tct Ser	cat His 885	gta Val	tct Ser	att Ile	aga Arg	gca Ala 890	aga Arg	aac Asn	agc Ser	2832
aag Lys	ata Ile 895	tgc Cys	ttt Phe	gct Ala	aca Thr	tgc Cys 900	ttc Phe	gat Asp	cag Gln	aat Asn	ggt Val 905	ctc Leu	agc Ser	aat Asn	ctg Leu	2880
aag Lys 910	tcg Ser	aag Lys	gaa Glu	gga Gly	aga Arg 915	gca Ala	ata Ile	tct Ser	att Ile	cat His 920	aca Thr	aag Lys	tct Ser	act Thr	ggt Gly 925	2928
tta Leu	ggt Val	atc Ile	agt Ser	gat Asp 930	ggc Gly	aac Asn	aac Asn	tct Ser	ggt Gly 935	gtc Val	tct Ser	ggt Val	cgf Arg	cat His 940	att Ile	2976
ttt Phe	att Ile	tcc Ser	tct Ser 945	ggt Val	cca Pro	cgg Arg	gga Gly	gtg Val 950	atc Ile	tct Ser	aag Lys	ggg Gly	aag Lys 955	aag Lys	ttt Phe	3024
tgt Cys	ggc Gly	cac His 960	tat Tyr	gtg Val	att Ile	tca Ser	tct Ser 965	aag Lys	gaa Glu	ttc Phe	aca Thr	gat Asp 970	gaa Glu	agg Arg	ggt Val	3072
ggc Gly	tca Val	aaa Lys	tca Val	tac Val	aat Asn	ata Ile	aaa Lys	ttt Phe	cta Val	cgf Arg	gaa Glu	aga Arg	ggt Val	cca Pro	tca Val	3120

Gly	Ser	Lys	Ser	Tyr	Asn	Ile	Lys	Phe	Leu	Arg	Glu	Arg	Val	Pro	Ser		
975						980					985						
tgg	atc	aag	ata	cct	acc	tca	gct	gct	ctt	cca	ttc	gga	aca	ttt	gag		3168
Trp	Ile	Lys	Ile	Pro	Thr	Ser	Ala	Ala	Leu	Pro	Phe	Gly	Thr	Phe	Glu		
990					995					1000					1005		
aat	ata	ctc	tca	gat	gat	tcc	aat	aag	gat	gta	gca	cgc	agg	att			3213
Asn	Ile	Leu	Ser	Asp	Asp	Ser	Asn	Lys	Asp	Val	Ala	Arg	Arg	Ile			
				1010					1015					1020			
tct	gtc	ctt	aaa	gat	tct	ctt	aac	aga	gga	gac	ctg	aca	aaa	ctt			3258
Ser	Val	Leu	Lys	Asp	Ser	Leu	Asn	Arg	Gly	Asp	Leu	Thr	Lys	Leu			
				1025					1030					1035			
aag	tcc	att	caa	gaa	gct	atc	tta	caa	atg	agt	gct	cca	atg	gct			3303
Lys	Ser	Ile	Gln	Glu	Ala	Ile	Leu	Gln	Met	Ser	Ala	Pro	Met	Ala			
				1040					1045					1050			
ctg	aga	aat	gaa	ctg	atc	aca	aaa	ttg	aga	agt	gaa	aga	atg	cct			3348
Leu	Arg	Asn	Glu	Leu	Ile	Thr	Lys	Leu	Arg	Ser	Glu	Arg	Met	Pro			
				1055					1060					1065			
tat	ctt	ggt	gat	gaa	tca	ggc	tgg	aac	cgc	tcc	tgg	gtg	gca	att			3393
Tyr	Leu	Gly	Asp	Glu	Ser	Gly	Trp	Asn	Arg	Ser	Trp	Val	Ala	Ile			
				1070					1075					1080			
aag	aag	gtc	tgg	gct	tca	aag	tgg	aac	gag	aga	gct	tat	gtc	agt			3438
Lys	Lys	Val	Trp	Ala	Ser	Lys	Trp	Asn	Glu	Arg	Ala	Tyr	Val	Ser			
				1085					1090					1095			
tgc	aaa	aaa	aat	aag	ctt	gat	cat	gat	gcg	gtc	tgc	atg	gct	gtg			3483
Cys	Lys	Lys	Asn	Lys	Leu	Asp	His	Asp	Ala	Val	Cys	Met	Ala	Val			
				1100					1105					1110			
ctg	att	caa	gaa	gtc	att	tgt	ggc	gat	tat	gct	ttc	gtc	att	cac			3528
Leu	Ile	Gln	Glu	Val	Ile	Cys	Gly	Asp	Tyr	Ala	Phe	Val	Ile	His			
				1115					1120					1125			
aca	aac	aat	ccg	gtc	tct	ggt	gac	tct	tca	gaa	ata	tac	aca	gag			3573
Thr	Asn	Asn	Pro	Val	Ser	Gly	Asp	Ser	Ser	Glu	Ile	Tyr	Thr	Glu			
				1130					1135					1140			
att	gtg	aag	ggt	ttg	gga	gag	acc	ttg	gtt	gga	gca	tat	cca	gga			3618
Ile	Val	Lys	Gly	Leu	Gly	Glu	Thr	Leu	Val	Gly	Ala	Tyr	Pro	Gly			
				1145					1150					1155			
cga	gca	atg	agc	ttc	atc	acc	aag	aaa	aca	aac	ctc	aag	tct	cca			3663
Arg	Ala	Met	Ser	Phe	Ile	Thr	Lys	Lys	Thr	Asn	Leu	Lys	Ser	Pro			
				1160					1165					1170			
acc	gtg	atc	agt	tac	cca	agt	aag	agg	ata	ggt	ctg	tat	tct	aaa			3708
Thr	Val	Ile	Ser	Tyr	Pro	Ser	Lys	Arg	Ile	Gly	Leu	Tyr	Ser	Lys			
				1175					1180					1185			
ccc	tca	atc	ata	ttc	aga	tca	gat	tca	aac	aac	gag	gat	ctc	gaa			3753
Pro	Ser	Ile	Ile	Phe	Arg	Ser	Asp	Ser	Asn	Asn	Glu	Asp	Leu	Glu			
				1190					1195					1200			
ggc	aac	gca	ggc	gtt	gga	ctt	tac	gac	agt	gtg	ata	atg	gat	gaa			3798
Gly	Asn	Ala	Gly	Val	Gly	Leu	Tyr	Asp	Ser	Val	Ile	Met	Asp	Glu			
				1205					1210					1215			
gca	gag	gaa	gta	gtg	gtg	gat	tac	tca	agg	gag	cca	cta	ata	atg			3843
Ala	Glu	Glu	Val	Val	Val	Asp	Tyr	Ser	Arg	Glu	Pro	Leu	Ile	Met			
				1220					1225					1230			
gac	aaa	tcc	ttt	cga	gtg	cgt	cta	ttc	tca	gcg	att	gca	gaa	gct			3888
Asp	Lys	Ser	Phe	Arg	Val	Arg	Leu	Phe	Ser	Ala	Ile	Ala	Glu	Ala			

	1235		1240		1245	
gga aat gtg ata gag tca atc tat ggt tgt cct caa gac att gaa						3933
Gly Asn Val Ile Glu Ser Ile Tyr Gly Cys Pro Gln Asp Ile Glu						
	1250		1255		1260	
ggt gtt gtc aaa ggt gga cat atc tac atc gtc caa gct aga ccc						3978
Gly Val Val Lys Gly Gly His Ile Tyr Ile Val Gln Ala Arg Pro						
	1265		1270		1275	
caa gtt taa						3987
Gln Val						

<210> 3

<211> 1328

<212> PRT

<213> Sequência artificial

<220>

<230> Sequência contendo a região codificadora da sequência de sinal plastídeo do gene *dul-I* de *Oryza sativa* fundido ao cADN codificador do gene R3 de *Arabidopsis thaliana*

<400> 3

Met Glu Met Ala Leu Arg Pro Gln Ser Leu Leu Cys Pro Arg Ser Arg  
 -50 -45 -40

Leu Lys Val Val Ile Arg Pro Ala Ser Ser Ala Ser Gly Gly Gly Leu  
 -35 -30 -25 -20

Ala Gln Tyr Phe Leu Met Thr Arg Arg Tyr Thr Gly Ser Arg Ile Val  
 -15 -10 -5

Arg Asn Leu Ala Thr Ser Lys Ser Gln Gln Phe Gln Leu Ile Glu Gly  
 -1 1 5 10

Met Glu Leu Gln Ile Thr Val Thr Gly Leu Pro Asn Gly Ser Ser Val  
 15 20 25

Arg Ala Glu Phe His Leu Lys Asn Cys Thr Arg Ala Trp Ile Leu His  
 30 35 40 45

Trp Gly Cys Ile Tyr Gln Gly Asn Asn His Trp Tyr Ile Pro Ser Glu  
 50 55 60

His Ser Ser Lys Gln Gly Ala Leu Gln Thr Thr Phe Val Lys Ser Gly  
 65 70 75

Asp Ala Tyr Val Val Ile Leu Glu Leu Arg Asp Pro Arg Val Arg Ala  
 80 85 90

Ile Glu Phe Val Leu Lys Asp Gly Ser His Asn Arg Trp Leu Arg Gln  
 95 100 105

His Asn Gly Asn Phe Arg Val Glu Ile Pro Trp Asn Asp Leu His Ala  
 110 115 120 125

His His Arg Ile Pro Lys Thr Leu Ile Glu Arg Arg Ala His Lys Ile  
 130 135 140

Trp Asp Arg Lys Gly Arg Pro Gln Ser Ser Ala Arg Glu Gln Gln Ile  
 145 150 155

Asp Tyr Asp Asn Ala Val Arg Glu Leu His Ala Glu Leu Ala Arg Gly  
 160 165 170

Ile Ser Leu Asp Glu Leu Gln Ala Asn Ser Thr Val Pro Val Glu Lys  
 175 180 185

Glu Val Thr Ser Glu Pro His His Thr Met Ile Gln Ser Tyr Arg Arg  
 190 195 200 205

Arg His Asp Val Gln Lys Trp Leu Gln Lys Tyr Thr Glu Pro Ile Asn  
 210 215 220

Arg Ser Gly Ser Val Lys Ser Ser Ala Leu Ala Glu Leu Ser Lys Arg  
 225 230 235

Ser Val Gly Gln Glu Asn Leu Val Ser Gln Lys Ser Phe His Val Arg  
 240 245 250

Asn Tyr Glu Ile Thr Val Leu Gln Arg Asp Val Lys Gly Asp Cys Arg  
 255 260 265

Leu Trp Ile Ala Thr Asn Met Ala Gly Pro Thr Val Leu His Trp Gly  
 270 275 280 285

Val Ala Lys Ser Ser Ala Gly Glu Trp Leu Ile Pro Pro Pro Asp Val  
 290 295 300

Leu Pro Glu Lys Ser Lys Phe Val His Gly Ala Cys Gln Thr Gln Phe  
 305 310 315

Thr Asp Met Ser Ser Arg Glu His Ser Tyr Gln Phe Ile Asp Ile Asn  
 320 325 330

Leu Lys Arg Gly Gly Phe Val Gly Ile Gln Phe Val Ile Trp Ser Gly  
 335 340 345

Gly Tyr Trp Val Asn Asn Asn Gly Ala Asn Phe Val Val Asn Leu Lys  
 350 355 360 365

Ser Ala Asp Ser Thr Ser Gly Lys Leu Asp Val Asp Glu Lys Tyr Val



Ala Arg Ile Gln Leu His Pro Ala Leu Arg Ala Pro Arg Thr Arg Ala  
 655 660 665

Lys Asp Leu Leu Phe Leu Asp Ile Ala Leu Glu Ser Cys Phe Lys Thr  
 670 675 680 685

Thr Ile Glu Lys Arg Leu Ile Ser Leu Asn Phe Asn Asn Pro Pro Glu  
 690 695 700

Ile Ile Tyr Val Ile Cys Val Val Leu Glu Asn Leu Cys Leu Ser Ile  
 705 710 715

Val Asn Asn Glu Glu Ile Ile Phe Cys Thr Lys Asp Trp Tyr Arg Val  
 720 725 730

Ser Glu Ala Tyr Arg Pro His Asp Val Gln Trp Ala Leu Gln Thr Lys  
 735 740 745

Ala Val Leu Asp Arg Leu Gln Leu Val Leu Ala Asp Arg Cys Gln His  
 750 755 760 765

Tyr Phe Thr Ile Ile Gln Pro Thr Thr Lys Tyr Leu Gly Gln Leu Leu  
 770 775 780

Arg Val Asp Lys His Gly Ile Asp Val Phe Thr Glu Glu Val Ile Arg  
 785 790 795

Ala Gly Pro Gly Ala Val Leu Ser Thr Leu Val Asn Arg Phe Asp Pro  
 800 805 810

Ser Leu Arg Lys Ile Ala Asn Leu Gly Cys Trp Gln Val Ile Ser Leu  
 815 820 825

Ala Asp Ala Tyr Gly Phe Val Val Cys Val Asn Glu Leu Ile Val Val  
 830 835 840 845

Gln Asn Lys Phe Tyr Ser Lys Pro Thr Val Ile Ile Ala Ser Lys Val  
 850 855 860

Thr Gly Glu Glu Glu Ile Pro Ala Gly Val Val Ala Val Leu Thr Pro  
 865 870 875

Ser Met Ile Asp Val Leu Ser His Val Ser Ile Arg Ala Arg Asn Ser  
 880 885 890

Lys Ile Cys Phe Ala Thr Cys Phe Asp Gln Asn Val Leu Ser Asn Leu  
 895 900 905

Lys Ser Lys Glu Gly Arg Ala Ile Ser Ile His Thr Lys Ser Thr Gly  
 910 915 920 925



Pro Ser Ile Ile Phe Arg Ser Asp Ser Asn Asn Glu Asp Leu Glu  
 1190 1195 1200

Gly Asn Ala Gly Val Gly Leu Tyr Asp Ser Val Ile Met Asp Glu  
 1205 1210 1215

Ala Glu Glu Val Val Val Asp Tyr Ser Arg Glu Pro Leu Ile Met  
 1220 1225 1230

Asp Lys Ser Phe Arg Val Arg Leu Phe Ser Ala Ile Ala Glu Ala  
 1235 1240 1245

Gly Asn Val Ile Glu Ser Ile Tyr Gly Cys Pro Gln Asp Ile Glu  
 1250 1255 1260

Gly Val Val Lys Gly Gly His Ile Tyr Ile Val Gln Ala Arg Pro  
 1265 1270 1275

Gln Val

<210> 4

<211> 12

<212> PRT

<213> *Arabidopsis thaliana*

<400> 4

Ile Asp Val Leu Ser His Val Ser Ile Arg Ala Arg  
 1 5 10

**REFERÊNCIAS CITADAS NA DESCRIÇÃO**

Esta lista das referências citadas pelo requerente serve apenas de interesse ao leitor. Não faz parte do documento da patente europeia. Apesar da compilação cuidadosa das referências, os erros ou as omissões não podem ser excluídos e o IEP rejeita toda a responsabilidade a este respeito.

Documentos de patente citados na descrição:

WO 9711188 A [0010]

WO 0077229 A [0010]

US 6462256 B [0010]

WO 0234923 A [0011]

EP 120516 A [0054]

WO 950612 A [0055]

EP 0513849 A [0055]

EP 0465875 A[0055]

EP 0292435 A[0055]

WO 9307279 A [0082]

WO 9515972 A [0087]

US 5188642 A [0093]

EP 0508909 A[0093]

EP 0924299 A[0093]

US 5510471 A [0093]

EP 04090316 A[0231]

US 60602454 B[0231]

Literatura, não relacionada com patentes, citada na descrição:

Hizukuri; Takagi. Carbohydr. Res., 1984, vol. 134, 1-10 [0005]

Takeda et al. Carbohydr. Res., 1984, vol. 132, 83-92 [0005]

Jane et al. Cereal Foods World, 1996, vol. 41 (11), 827-832 [0007] [0008]

Takeda; Hizukuri. Starch/Stärke, 1971, vol. 23, 267-272 [0008]

Blennow et al. Int. J. of Biological Macromolecules, 2000, vol. 27, 211-218 [0008]

Blennow et al. Carbohydrate Polymers, 2000, vol. 41, 163-174 [0008]

Ritte et al. PNAS, 2002, vol. 99, 7166-7171 [0009] [0043] [0192]

Lorberth et al. Nature Biotechnology, 1998, vol. 16, 473-477 [0009]

Lorberth et al. Nature Biotechnology, 1996, vol. 16, 473-477 [0009]

Bioanalytik. Spektrum Akad. Verlag, 1998 [0027]

Heidt. Pflanzenbiochemie. Spektrum Akad. Verl, 1996 [0028]

Plant Biochemistry. Academic Press, 1997 [0028]

- Wischmann et al. *Plant Physiology*, 1999, vol. 119, 455-462 [0033]
- Echeveria et al. *Plant Physiology*, 1988, vol. 86, 786-792 [0033]
- Tetlow et al. *Planta*, 1993, vol. 189, 597-600 [0033]
- Smith et al. *Planta*, 1990, vol. 180, 517-523 [0033]
- Strasburger. *Lehrbuch der Botanik*. Spektrum Akad. Verl, 1999 [0034]
- Tien-Shin Yu et al. *Plant Cell*, 2001, vol. 13, 1907-1918 [0036] [0038] [0047]
- Hoekema, IN. *The Binary Plant Vector System* Offsetdrukkerij Kanters B.V. Alblasterdam. 1985 [0054]
- Fraley et al. *Crit. Rev. Plant Sei.*, vol. 4,1 -46 [0054]
- An et al. *EMBO J.*, 1985, vol. 4, 277-287 [0054]
- Rocha-Sosa et al. *EMBO J*, 1989, vol. 8, 29-33
- Chan et al. *Plant Mol. Biol.*, 1993, vol. 22, 491 -506
- Hiei et al. *Plant J.*, 1994, vol. 6, 271 -282 [0055]
- Deng et al. *Science in China*, 1990, vol. 33, 28-34 [0055]
- Wilmink et al. *Plant Cell Reports*, 1992, vol. 11, 76-80 [0055]
- May et al. *Bio/Technology*, 1995, vol. 13, 486-492 [0055]
- Conner; Domisse. *Int. J. Plant Sci.*, 1992, vol. 153, 550-555 [0055]

- Ritchie et al. *Transgenic Res.*, 1993, vol. 2, 252-265  
[0055]
- Wan ; Lemaux. *Plant Physiol.*, 1994, vol. 104, 37-48  
[0055]
- Vasil et al. *Bio/Technology*, 1993, vol. 11, 1553-1558  
[0055]
- Ritala et al. *Plant Mol. Biol.*, 1994, vol. 24, 317-325  
[0055]
- Spencer et al. *Theor. Appl. Genet.*, 1990, vol. 79, 625-631 [0055]
- Fromm et al. *Biotechnology*, 1990, vol. 8, 833-844  
[0055]
- Gordon-Kamm et al. *Plant Cell*, 1990, vol. 2, 603-618  
[0055]
- Kozziel et al. *Biotechnology*, 1993, vol. 11, 194-200
- Moroc et al. *Theor. Appl. Genet.*, 1990, vol. 80, 721-726  
[0055]
- Krens et al. *Nature*, 1982, vol. 296, 72-74 [0056]
- Nehra et al. *Plant J.*, 1994, vol. 5, 285-297 [0056]
- Becker et al. *Plant Journal*, 1994, vol. 5, 299-307
- Sambrook et al. *Molecular Cloning, A Laboratory Manual*.  
Cold Spring Harbour Laboratory Press, 2001 [0071] [0072]  
[0091]
- Ausubel et al. *Short Protocols in Molecular Biology*.  
John Wiley & Sons, 2002 [0071] [0072] [0091]
- Thompson et al. *Nucleic Acids Research*, 1994, vol. 22,  
4673-4680 [0073]

Thomeycroft et al. *Journal of experimental Botany*, 2001, vol. 52 (361), 1593-1601 [0077]

Ramachandran ; Sundaresan. *Plant Physiology and Biochemistry*, 2001, vol. 39, 234-252 [0077]

Walden et al. *Plant J.*, 1991, 281 -288 [0078]

Walden et al. *Plant Mol. Biol.*, 1994, vol. 26, 1521-1528 [0078]

Rocha-Sosa et al. *EMBO J.*, 1989, vol. 8, 23-29 [0082]

Stockhaus et al. *Proc. Natl. Acad. Sci. USA*, 1987, vol. 84, 7943-7947 [0082]

Stockhaus et al. *EMBO J.*, 1989, vol. 8, 2445-2451 [0082]

Pedersen et al. *Cell*, 1982, vol. 29,1015-1026 [0082]

Quatroccio et al. *Plant Mol. Biol.*, 1990, vol. 15, 81-93 [0082]

Leisy et al. *Plant Mol. Biol.*, 1990, vol. 14, 41-50 [0082]

Zheng et al. *Plant J.*, 1993, vol. 4, 357-366 [0082]

Yoshihara et al. *FEBS Lett.*, 1996, vol. 383,213-218 [0082]

Werr et al. *EMBO J.*, 1985, vol. 4,1373-1380 [0082]

Fiedler et al. *Plant Mol. Biol.*, 1993, vol. 22, 669-679 [0082]

Baumlein et al. *Mol. Gen. Genet.*, 1991, vol. 225, 459-467 [0082]

Gielen et al. *EMBO J.*, 1989, vol. 8, 23-29 [0083]

Callis et al. *Genes Devel.*, 1987, vol. 1, 1183-1200  
[0084]

Luehrsen ; Walbot. *Mol. Gen. Genet.*, 1991, vol. 225,  
81-93 [0084]

Rethmeier et al. *Plant Journal*, 1997, vol. 12 (4), 895-  
899 [0084]

Rose ; Beliakoff. *Plant Physiol.*, 2000, vol. 122 (2),  
535-542 [0084]

Vasil et al. *Plant Physiol.*, 1989, vol. 91, 1575-1579  
[0084]

XU et al. *Science in China Series C*, 2003, vol. 46 (6),  
561-569 [0084]

Kipp, P.B. et al. Poster Session at the "5th Interna-  
tional Congress of Plant Molecular Biology, 21 September  
1997 [0087]

Metabolic Engineering in Transgenic Plants. R.A. Dixon;  
C.J. Amtzen. *Keystone Symposia. Tl- BTECH*, 1997, vol. 15,  
441-447 [0087]

Kren et al. *Hepatology*, 1997, vol. 25, 1462-1468 [0087]

Cole-Strauss et al. *Science*, 1996, vol. 273, 1386-1389  
[0087]

Beetham et al. *PNAS*, 1999, vol. 96, 8774-8778 [0087]

Braun. *EMBO J.*, 1992, vol. 11, 3219-3227 [0093]

Sonnewald. *Plant J.*, 1991, vol. 1, 95-106 [0093]

Rocha-Sosa. *EMBO J.*, 1989, vol. 8, 23-29 [0093]

Neuhaus ; Rodgers. *Plant Molecular Biology*, 1998, vol.  
38, 127-144 [0093]

Emanuelsson et al. *Protein Science*, 1999, vol. 8, 978-984 [0093]

Jansen et al. *Current Genetics*, 1988, vol. 13, 517-522 [0093]

Klôsgen et al. *Mol Gen Genet.*, 1989, vol. 217, 155-161 [0093]

Wolter et al. *Proc. Natl. Acad. Sci. USA*, 1988, vol. 85, 846-850 [0093]

Nawrath et al. *Proc. Natl. Acad. Sci. USA*, 1994, vol. 91, 12760-12764 [0093]

Gallardo et al. *Planta*, 1995, vol. 197, 324-332 [0093]

Creissen et al. *Plant J.*, 1995, vol. 8, 167-175 [0093]

Ritte et al. *Starch/Starke*, 2000, vol. 52, 179-185 [0099]

Kasemusuwan ; Jane. *Cereal Chemistry*, 1996, vol. 73, 702-707 [0099]

*Plant Cell Culture Protocols*. Humana Press, 1999 [0113]

*Starch: Chemistry and Technology*. Academic Press Inc. London Ltd [0131]

maize and sorghum starches: production. Watson. *STARCH : CHEMISTRY AND TECHNOLOGY*. 412-468 [0131]

tapioca, arrowroot and sago starches: production. Corbishley ; Miller. *STARCH : CHEMISTRY AND TECHNOLOGY*. 469-479 [0131]

potato starch: production and uses. Mitch. *STARCH : CHEMISTRY AND TECHNOLOGY*. 479-490 [0131]

wheat starch: production, modification and uses.  
Knight; Oson. STARCH: CHEMISTRY AND TECHNOLOGY. 491-506  
[0131]

rice starch: production and uses. Rohmer ; Klem. STARCH  
: CHEMISTRY AND TECHNOLOGY. 507-528 [0131]

Eckhoff et al. Cereal Chem., 1996, vol. 73, 54-57  
[0131]

Com, Chemistry and Technology. 1987, 479-499 [0142]

Methods in Enzymology, 1987, vol. 153, 383-516 [0161]

Bitter et al. Methods in Enzymology, 1987, vol. 153,  
516-544 [0161]

Schlegel. Allgemeine Mikrobiologie. Georg Thieme  
Verlag, 1985, 1-2 [0163]

Hiei et al. Plant Journal, 1994, vol. 6 (2), 271-282  
[0194] [0228]

Ames. Methods in Enzymology VIII, 1966, 115-118 [0196]

Comelissen ; Vanderwiele. Nucleic Acid Research, 1989,  
vol. 17, 19-25 [0198]

Bolivar et al. Gene, 1977, vol. 2, 95-113 [0198]

Itoh et al. Plasmid, 1984, vol. 11, 206 [0198]

Tolmasky. Plasmid, 1990, vol. 24 (3), 218-226 [0198]

Tolmasky ; Crosa. Plasmid, 1993, vol. 29 (1), 31 -40  
[0198]

Odell et al. Nature, 1985, vol. 313, 180 [0198]

Thompson et al. Embo J., 1987, vol. 6, 2519-2523 [0198]

Gielen et al. EMBO 1, 1984, vol. 3, 835-846 [0198]

Depicker et al. J. Mol. Appl. Genet., 1982, vol. 1,  
561-573 [0198] [0200]

Christensen et al. Plant Mol. Biol., 1992, vol. 18,  
675-689 [0200]

Plant Cell, 1998, vol. 10, 399-412 [0213]

Lisboa, 23/04/2010

**REIVINDICAÇÕES**

1. Célula vegetal geneticamente modificada, em que os plastídeos presentes na dita célula vegetal incluem uma proteína que transfere um resíduo fosfato de um nucleosídeo trifosfato para o amido, em que a modificação genética consiste na introdução da molécula de ácido nucleico exógena no genoma da célula vegetal, onde a molécula de ácido nucleico exógena é seleccionada a partir do grupo constituído por
  - a) moléculas de ácido nucleico que codificam uma proteína com a sequência de aminoácidos especificada na SEQ ID N° 3;
  - b) moléculas de ácido nucleico que codificam uma proteína que inclui a sequência de aminoácidos, codificada pela inserção no plasmídeo DSM 16587 ou pela inserção no plasmídeo DSM 16645;
  - c) moléculas de ácido nucleico que codificam uma proteína cuja sequência tem pelo menos 60 % de identidade com a sequência de aminoácidos especificada na SEQ ID N° 3;
  - d) moléculas de ácido nucleico que codificam uma proteína cuja sequência tem uma identidade de pelo menos 60% com a sequência de aminoácidos codificada pela inserção no plasmídeo DSM 16587 ou pela inserção no plasmídeo DSM 16645;
  - e) moléculas de ácido nucleico que incluem a sequência nucleotídica especificada na SEQ ID N° 1 ou na SEQ ID N° 2 ou uma sequência complementar;

- f) moléculas de ácido nucleico que compreendem a sequência nucleotídica da inserção presente no plasmídeo DSM 16587 ou no plasmídeo DSM 16587;
- g) moléculas de ácido nucleico que têm uma identidade de pelo menos 60% com as sequências de ácido nucleico descritas em a), b), e) ou f);
- h) moléculas de ácido nucleico que hibridam com pelo menos uma cadeia das moléculas de ácido nucleico descritas em a), b), e) ou f), sob condições restritivas;
- i) moléculas de ácido nucleico cuja sequência nucleotídica se desvia da sequência de moléculas de ácido nucleico mencionadas em a), b), e) ou f), em resultado da degenerescência do código genético e
- j) moléculas de ácido nucleico que constituem fragmentos, variantes alélicas e/ou derivados das moléculas de ácido nucleico descritas em a), b), c), d), e), f), g), h) ou i)

e em que a molécula de ácido nucleico exógena é uma molécula de ácido nucleico recombinante compreendendo sequências adicionais que não se encontram naturalmente presentes numa combinação em que se encontram presentes nos ácidos nucleicos recombinantes, nos quais a molécula de ácido nucleico codificadora da proteína que transfere um resíduo fosfato a partir de um nucleosídeo trifosfato para o amido está fundida às sequências de ácido nucleico que codificam um peptídeo de sinal para o plastídeo, de forma a que a molécula de ácido nucleico recombinante codifica uma proteína que transfere um resíduo fosfato a partir de um nucleosídeo trifosfato para o amido e que possui uma

sequência de sinal para o plastídeo para além da sua sequência de aminoácidos.

2. Célula vegetal de acordo com a reivindicação 1 que sintetiza um amido modificado em comparação com as células vegetais de tipo selvagem não modificadas geneticamente, em que o amido modificado inclui um teor de fosfato de amido aumentado e/ou uma distribuição de fosfato modificada.
3. Planta que inclui células vegetais geneticamente modificadas de acordo com qualquer uma das reivindicações 1 e 2.
4. Planta de acordo com a reivindicação 3 que é uma planta armazenadora de amido.
5. Material de propagação de uma planta de acordo com a reivindicação 3 ou 4, compreendendo uma célula vegetal de acordo com qualquer uma das reivindicações 1 ou 2.
6. Parte que pode ser colhida de uma planta de acordo com a reivindicação 3 ou 4, compreendendo uma célula vegetal de acordo com qualquer uma das reivindicações 1 ou 2.
7. Método de geração de uma planta geneticamente modificada em que
  - a) uma célula vegetal é transformada, conduzindo a transformação a um aumento da actividade de pelo menos uma proteína que codifica uma proteína que

transfere um resíduo fosfato a partir de um nucleosídeo trifosfato para o amido em comparação com células vegetais correspondentes não modificadas geneticamente, de tipo selvagem, e que consiste na introdução de uma molécula de ácido nucleico exógena codificadora de uma proteína que transfere um resíduo fosfato a partir de um nucleosídeo trifosfato para o amido seleccionada a partir do grupo constituído por I

- i) moléculas de ácido nucleico que codificam uma proteína com a sequência de aminoácidos especificada na SEQ ID N° 3;
- ii) moléculas de ácido nucleico que codificam uma proteína que inclui a sequência de aminoácidos, codificada pela inserção em DSM 16587 ou pela inserção em DSM 16645;
- iii) moléculas de ácido nucleico que codificam uma proteína cuja sequência de aminoácidos tem pelo menos 60 % de identidade com a sequência de aminoácidos especificada na SEQ ID N° 3;
- iv) moléculas de ácido nucleico que codificam uma proteína cuja sequência tem uma identidade de pelo menos 60% com a sequência de aminoácidos, codificada pela inserção no plasmídeo DSM 16587 ou pela inserção no plasmídeo DSM 16645;
- v) moléculas de ácido nucleico que incluem a sequência nucleotídica especificada na SEQ ID N° 1 ou na SEQ ID N° 2 ou uma sequência complementar;

- vi) moléculas de ácido nucleico que compreendem a sequência nucleotídica da inserção presente no plasmídeo DSM 16587 ou no plasmídeo DSM 16645;
- vii) moléculas de ácido nucleico que têm uma identidade de pelo menos 70% com as sequências de ácido nucleico descritas em i), ii), v) ou vi);
- viii) moléculas de ácido nucleico que hibridam com pelo menos uma cadeia das moléculas de ácido nucleico descritas em i), ii), v) ou vi) sob condições restritivas;
- iix) moléculas de ácido nucleico cuja sequência nucleotídica se desvia da sequência de moléculas de ácido nucleico mencionadas em i), ii), v) ou vi), em resultado da degenerescência do código genético e
- ix) moléculas de ácido nucleico que constituem fragmentos, variantes alélicas e/ou derivados das moléculas de ácido nucleico descritas em i), ii), iii), iv), v), vi), vii), viii) ou iix),

no genoma da célula vegetal, em que a dita molécula de ácido nucleico exógena é uma molécula de ácido nucleico recombinante compreendendo sequências adicionais que não se encontram naturalmente presentes numa combinação em que se encontram presentes nos ácidos nucleicos recombinantes, nos quais a molécula de ácido nucleico codificadora da proteína que transfere um resíduo fosfato a partir de um nucleosídeo trifosfato para o amido está fundida às sequências de ácido nucleico que codificam um peptídeo de sinal para o plastídeo, de forma a que a molécula de ácido nucleico recombinante codifica uma proteína que transfere um resíduo

fosfato a partir de um nucleosídeo trifosfato para o amido e que possui uma sequência de sinal para o plastídeo para além da sua sequência de aminoácidos,

b) uma planta é regenerada a partir de células vegetais da etapa a) e

c) se apropriado, são produzidas mais plantas com ajuda das plantas de acordo com a etapa b).

8. Método de produção de amido modificado, compreendendo a etapa de extração do amido a partir de uma célula vegetal de acordo com o reivindicado em qualquer uma das reivindicações 1 ou 2 ou de extração do amido de uma planta de acordo com o reivindicado na reivindicação 3 ou 4.

9. Método de produção de um amido modificado, compreendendo a etapa de extração do amido de partes de planta que podem ser colhidas de acordo com o reivindicado na reivindicação 6.

10. Método de produção de farinhas compreendendo a etapa de moagem de partes de plantas de acordo com o reivindicado na reivindicação 3 ou 4 ou de material de propagação de acordo com o reivindicado na reivindicação 5 ou material que pode ser colhido de acordo com o reivindicado na reivindicação 6.

11. Utilização de uma célula vegetal geneticamente modificada de acordo com o reivindicado em qualquer uma das reivindicações 1 ou 2 ou de uma planta de acordo

com o reivindicado na reivindicação 3 ou 4 para a produção de uma farinha.

12. Molécula de ácido nucleico codificadora de uma proteína com a actividade enzimática de uma proteína que transfere um resíduo fosfato a partir de um nucleosídeo trifosfato para amido seleccionada a partir do grupo constituído por

- a) molécula de ácido nucleico que codifica uma proteína com a sequência de aminoácidos especificada na SEQ ID N° 3
- b) molécula de ácido nucleico que codifica uma proteína cuja sequência de aminoácidos tem pelo menos 60 % de identidade com a sequência de aminoácidos especificada na SEQ ID N° 3;
- c) molécula de ácido nucleico que inclui a sequência nucleotídica especificada na SEQ ID N° 1 ou na SEQ ID N° 2 ou uma sequência complementar destas sequências;
- d) molécula de ácido nucleico que tem uma identidade de pelo menos 70% com as sequências de ácido nucleico descritas em a) ou c);
- e) molécula de ácido nucleico que hibrida com pelo menos uma cadeia da molécula de ácido nucleico descrita em a) ou c) sob condições restritivas;
- f) molécula de ácido nucleic cuja sequência nucleotídica se desvia da sequência de moléculas de ácido nucleico mencionadas em a) ou c), em resultado da degenerescência do código genético e

- g) molécula de ácido nucleico que constitui um fragmento, variante alélica e/ou derivado das moléculas de ácido nucleico descritas em a), b), c), d), e) ou f),

e em que a molécula de ácido nucleico mencionada é uma molécula de ácido nucleico recombinante compreendendo sequências adicionais que não se encontram naturalmente presentes numa combinação em que se encontram presentes no ácido nucleico recombinante, em que a molécula de ácido nucleico codificadora da proteína que transfere um resíduo fosfato a partir de um nucleosídeo trifosfato para o amido está fundida às sequências de ácido nucleico que codificam um peptídeo de sinal para o plastídeo.

13. Vector compreendendo uma molécula de ácido nucleico de acordo com o reivindicado na reivindicação 12.
14. Vector de acordo com a reivindicação 13, em que a molécula de ácido nucleico está ligada a sequências reguladoras que iniciam a transcrição em células procariotas e eucariotas.
15. Célula hospedeira que compreende uma molécula de ácido nucleico de acordo com a reivindicação 12 ou um vector de acordo com a reivindicação 13 ou 14.
16. Composição compreendendo uma molécula de ácido nucleico de acordo com a reivindicação 12 ou um vector de acordo com a reivindicação 13 ou 14.

17. Utilização de uma composição de acordo com a reivindicação 16 para a identificação de células vegetais, em comparação com células vegetais de tipo selvagem, não modificadas geneticamente, que revela um aumento na actividade de uma proteína que transfere um resíduo fosfato a partir de nucleosídeo trifosfato para o amido.
  
18. Proteína com actividade de fosforilação do amido seleccionada a partir do grupo constituído por
  - a) uma proteína que inclui uma sequência de aminoácidos especificada na SEQ ID N° 3,
  - b) uma proteína com pelo menos 60 % de identidade com a sequência de aminoácidos das proteínas citadas em a) fundida às sequências de aminoácidos que codificam um peptídeo de sinal para o plastídeo.
  
19. Proteína recombinante que inclui sequências de aminoácidos codificadoras de uma proteína de acordo com a reivindicação 18, fundida às sequências de aminoácidos que codificam um peptídeo de sinal para o plastídeo.
  
20. Proteína de acordo com a reivindicação 18 ou 19 que possui um domínio fosfo-histidina.