

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

(21) **PI0614885-9 A2**

(22) Data de Depósito: 10/08/2006  
(43) Data da Publicação: 19/04/2011  
(RPI 2102)



(51) *Int.Cl.:*

C12N 15/09  
A01H 1/00  
A01H 5/00  
C12N 5/10  
C12P 19/26

(54) Título: **PLANTA PRODUTORA DE ÁCIDO HIALURÔNICO**

(30) Prioridade Unionista: 25/08/2005 JP 2005-244192,  
21/02/2006 JP 2006-043724, 21/02/2006 JP 2006-043724, 25/08/2005  
JP 2005-244192

(73) Titular(es): TOYO BOSEKI KABUSHIKI KAISHA

(72) Inventor(es): Atsushi Sogabe, Hiroaki Kitazawa, Shigeo  
Shibatani

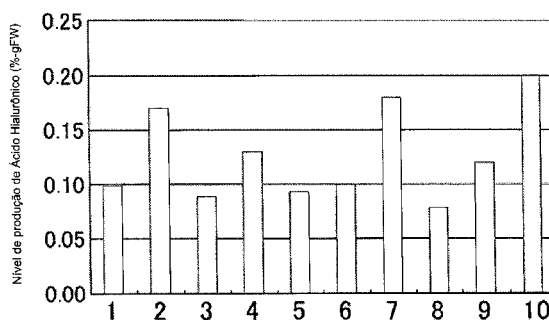
(74) Procurador(es): Dannemann, Siemsen, Bigler &  
Ipanema Moreira

(86) Pedido Internacional: PCT JP2006315817 de 10/08/2006

(87) Publicação Internacional: WO 2007/023682 de 01/03/2007

(57) Resumo: PLANTA PRODUTORA DE ÁCIDO HIALURÔNICO.

A presente invenção proporciona, aprimorando um método conhecido para produzir ácido hialurônico em uma planta, uma planta ou algumas células vegetais cultivadas as quais podem produzir ácido hialurônico em um menor custo e um rendimento maior do que antes adicional, um método para preparar as mesmas, um vetor de expressão para transformação, um método para produzir ácido hialurônico usando a planta ou as células vegetais cultivadas e semelhantes. É proporcionado o método para produzir ácido hialurônico compreendendo obter ácido hialurônico por coexpressão de uma proteína com atividade de ácido hialurônico sintase e uma proteína exógena com atividade de nucleotídeo de açúcar sintase em uma célula vegetal ou uma planta.



Relatório Descritivo da Patente de Invenção para **"PLANTA PRODUTORA DE ÁCIDO HIALURÔNICO"**.

DESCRIÇÃO

CAMPO TÉCNICO

5 A presente invenção refere-se de modo geral, a um método para produzir ácido hialurônico em plantas, células de plantas transgênicas ou plantas transgênicas tendo a capacidade de produzir ácido hialurônico, e métodos para produzir estas células transgênicas e plantas transgênicas.

ANTECEDENTES DA TÉCNICA

10 Ácido hialurônico é um glicosaminoglicano (mucopolissacarídeo) isolado do corpo vítreo de um globo ocular bovino por Meyer e Palmer em 1934 (Meyer, K. and Palmer, J. W. (1934) J. Biol. Chem., 107, 629-634). Ácido hialurônico de alto peso molecular tem sido usado para o tratamento de osteoartrite, um auxiliar cirúrgico para oftalmologia, prevenção de adesão,  
15 aceleração da cura de ferimento e semelhantes. Também foi reportado que ácido hialurônico de baixo peso molecular tem efeitos fisiologicamente ativos. Espera-se que sejam descobertos novos usos para o ácido hialurônico como um biomaterial ou em uma aplicação medicinal.

Até agora, o ácido hialurônico tem sido produzido por extração  
20 de tecidos de mamíferos ou fermentação microbiana. No entanto, o risco de contaminação com, por exemplo, encefalopatias espongiformes transmissíveis (príons) ou transmissão de vírus para humanos tem sido motivo de preocupação na extração dos tecidos de mamíferos. As células de mamíferos são caras de cultivar e manter. Necessitam de meios de crescimento caros e  
25 crescem lentamente. Entretanto, as fermentações microbianas têm problemas tais como a a necessidade de meio de crescimento contendo açúcar e instalações caras. Em *Escherichia coli*, há problemas em que proteínas não são processados, pode ser formado corpo de inclusão, proteínas são degradadas por proteases, e semelhantes. (Petrides, D. et al., (1995) Biotecnol.  
30 Bioeng., 48, 529). Quando substâncias terapêuticas são produzidas em micro-organismos, os custos de purificação se tornam extremamente caros de modo a evitar contaminação por endotoxinas.

Pelo contrário, as plantas são sistemas ideais para produzir carboidrato com baixa carga de energia, nas quais carboidratos são produzidos fotossintética-mente a partir de água e dióxido de carbono. A invenção revelada no Documento de Patente 6 mostra que o ácido hialurônico pode ser produzido introduzindo um gene de ácido hialurônico sintase em plantas ou células vegetais.

Documento de Patente 1: Pedido de Patente Japonesa Não Examinado Nº 1993-125103

Documento de Patente 2: Publicação de Patente Japonesa Não Examinada Nº 1983-056692

Documento de Patente 3: Pedido de Patente Japonesa Não Examinado Nº 1997-319579

Documento de Patente 4: Pedido de Patente Japonesa Não Examinado Nº 1994-319580

Documento de Patente 5: Pedido de Patente Japonesa Não Examinado Nº 1997-056394

Documento de Patente 6: WO 05/012529

Documento Não Patente 1: Meyer, K. and Palmer, J. W., J. Biol. Chem., 107: 629-634, 1934

Documento Não Patente 2: Petrides, D. et al., Biotechnol. Bioeng., 48: 529, 1995

### BREVE DESCRIÇÃO DOS DESENHOS

A figura 1 mostra análise SDS-PAGE de AtUGD antes e depois de purificação.

A figura 2 mostra análise SDS-PAGE de cvUGD antes e depois de purificação.

A figura 3 mostra GFAT expressado usando PROTEIOS (marca registrada).

A figura 4 mostra medição da atividade de GFAT usando o método de Reissig.

A figura 5 mostra o nível de produção de ácido hialurônico de plantas de tabaco transgênicas nas quais o gene cvHAS-cvGFAT e o gene

cvHAS são introduzidos respectivamente.

A figura 6 mostra o nível de produção de ácido hialurônico de plantas de tabaco transgênicas nas quais foram introduzidos múltiplos genes.

## 5 DESCRIÇÃO DA INVENÇÃO

### PROBLEMAS A SEREM RESOLVIDOS PELA INVENÇÃO

O objetivo primário da presente invenção é proporcionar plantas e células vegetais que podem produzir ácido hialurônico de modo mais eficaz aprimorando métodos previamente conhecidos para produzir ácido hialurônico em plantas, os métodos para produzir as mesmas, e os vetores de expressão recombinantes para esse fim.

### MEIOS PARA RESOLVER OS PROBLEMAS

Em consequência de estudo extensivo para resolver os problemas acima, o presente inventor descobriu que o ácido hialurônico é produzido extensivamente nas plantas transformando plantas e células vegetais com genes que codificam proteínas tendo atividade enzimática para produzir ácido hialurônico e genes que codificam proteínas tendo atividade enzimática de sintetização de nucleotídeos de açúcar, e estudo extensivamente adicional atingiu a presente invenção.

Isto é, a presente invenção refere-se aos seguintes itens.

1. Um método para produzir ácido hialurônico, compreendendo coexpressar uma proteína com atividade de ácido hialurônico sintase e uma proteína exógena com atividade de nucleotídeo de açúcar sintase em uma célula vegetal ou uma planta.

2. Um método para produzir ácido hialurônico, contendo as etapas de:

(1) transformar uma célula vegetal ou uma planta usando um vetor de expressão recombinante, o vetor de expressão recombinante tendo DNA codificando uma proteína com atividade de ácido hialurônico sintase e DNA codificando uma proteína com atividade de nucleotídeo de açúcar sintase sob controle de um ou mais promotores capazes de funcionamento em plantas;

- (2) cultivar um transformante obtido pela transformação; e
- (3) isolar ácido hialurônico produzido pelo transformante.

3. O método para produzir ácido hialurônico de acordo com o Item 2, em que o promotor é um promotor órgão-específico ou um tecido-específico.

4. O método para produzir ácido hialurônico de acordo com o Item 2 ou 3, em que o DNA codificando uma proteína com atividade de ácido hialurônico sintase é DNA de (a) ou (b) abaixo:

- (a) DNA consistindo em uma sequência de nucleotídeos representada pela SEQ ID NO: 1 ou 3; ou
- (b) DNA hibridizando à sequência de nucleotídeos complementar ao DNA consistindo na sequência de nucleotídeos de (a) sob condições estritas, e o DNA codificando uma proteína com atividade de ácido hialurônico sintase.

5. O método para produzir ácido hialurônico de acordo com qualquer um dos Itens 1 a 3, em que a proteína com atividade de ácido hialurônico sintase é uma proteína de (a) ou (b) abaixo:

- (a) uma proteína consistindo em uma sequência de aminoácidos representada pela SEQ ID NO: 2 ou 4; ou
- (b) uma proteína tendo a sequência de aminoácidos de (a) com um ou uns poucos aminoácidos eliminados, substituídos ou adicionados, e a proteína tendo atividade de ácido hialurônico sintase.

6. O método para produzir ácido hialurônico de acordo com qualquer um dos Itens 1 a 5, em que o nucleotídeo de açúcar é uridin-5'-difosfo(UDP)-N-acetilglucosamina e/ou ácido UDP-glucurônico.

7. O método para produzir ácido hialurônico de acordo com qualquer um dos Itens 1 a 6, em que a proteína com atividade de nucleotídeo de açúcar sintase é no mínimo uma proteína selecionada entre o grupo consistindo em glutamina:frutose-6-fosfato amidotransferase, UDP-N-acetilglucosamina difosforilase, fosfoacetil glucosamina mutase, glucosamina-6-fosfato N-acetiltransferase, glucosamina-1-fosfato N-acetiltransferase, fosfoglucomutase, N-acetilglucosamina quinase, hexoquinase, N-

acetilglucosamina acetilase, UDP-glucose desidrogenase, UDP-glucose-1-fosfato uridiltransferase, inositol oxigenase, glucuronoquinase e glucuronato-1-fosfato uridil transferase.

8. O método para produzir ácido hialurônico de acordo com qualquer um dos Itens 1 a 6, em que a proteína com atividade de nucleotídeo de açúcar sintase é no mínimo uma proteína selecionada entre o grupo consistindo em UDP-N-acetilglucosamina difosforilase, fosfoacetil glucosamina mutase, glucosamina-6-fosfato N-acetiltransferase, glucosamina-1-fosfato N-acetiltransferase, fosfoglucomutase, N-acetilglucosamina quinase, hexoquinase, N-acetilglucosamina acetilase, UDP-glucose desidrogenase, UDP-glucose-1-fosfato uridiltransferase, inositol oxigenase, glucuronoquinase e glucuronato-1-fosfato uridil transferase, e glutamina:frutose-6-fosfato amidotransferase.

9. O método para produzir ácido hialurônico de acordo com qualquer um dos Itens 1 a 6, em que uma proteína com atividade de nucleotídeo de açúcar sintase é glutamina:frutose-6-fosfato amidotransferase e/ou UDP-glucose desidrogenase.

10. O método para produzir ácido hialurônico de acordo com qualquer um dos Itens 2 a 9, em que o DNA codificando uma proteína com nucleotídeo de açúcar sintase atividade é DNA derivado de vírus chlorella e/ou *Arabidopsis thaliana*.

11. O método para produzir ácido hialurônico de acordo com qualquer um dos Itens 2 a 10, em que o DNA codificando uma proteína com atividade de nucleotídeo de açúcar sintase é DNA de (a) ou (b) abaixo:

- (a) DNA consistindo em uma sequência de nucleotídeos representada pela SEQ ID NO: 5, 7 ou 9; ou
- (b) DNA hibridizando à sequência de nucleotídeos complementar ao DNA consistindo na sequência de nucleotídeos de (a) sob condições estritas, e o referido DNA codificando uma proteína com atividade de glutamina:frutose-6-fosfato amidotransferase.

12. O método para produzir ácido hialurônico de acordo com qualquer um dos Itens 1 a 10, em que a proteína com atividade de nucleotí-

deo de açúcar sintase é uma proteína de (a) ou (b) abaixo:

- (a) uma proteína consistindo em uma sequência de aminoácidos representada pela SEQ ID NO: 6, 8 ou 10; ou
- (b) uma proteína consistindo na sequência de aminoácidos de (a) com um ou uns poucos aminoácidos eliminados, substituídos ou adicionados, e a proteína tendo atividade de glutamina:frutose-6-fosfato amidotransferase.

13. O método para produzir ácido hialurônico de acordo com qualquer um dos Itens 2 a 10, em que o DNA codificando uma proteína com atividade de nucleotídeo de açúcar sintase é DNA de (a) ou (b) abaixo:

- (a) DNA consistindo em uma sequência de nucleotídeos representada pela SEQ ID NO: 11, 13, 17, 19, ou 21; ou
- (b) DNA hibridizando à sequência de nucleotídeos complementar ao DNA consistindo na sequência de nucleotídeos de (a) sob condições estritas, e o DNA codificando uma proteína com atividade de UDP-glucose desidrogenase.

14. O método para produzir ácido hialurônico de acordo com qualquer um dos Itens 1 a 10, em que a proteína com atividade de nucleotídeo de açúcar sintase é uma proteína de (a) ou (b) abaixo:

- (a) uma proteína consistindo em uma sequência de aminoácidos representada pela SEQ ID NO: 12, 14, 16, 18, 20 ou 22; ou
- (b) uma proteína consistindo na sequência de aminoácidos de (a) com um ou uns poucos aminoácidos eliminados, substituídos ou adicionados, e a proteína tendo atividade de UDP-glucose desidrogenase.

15. O método para produzir ácido hialurônico de acordo com qualquer um dos Itens 1 a 14, em que a planta é selecionada entre o grupo consistindo em angiospermas, gimnospermas, pteridófitos e briófitas.

16. O método para produzir ácido hialurônico de acordo com o Item 3, em que os órgãos são selecionados entre o grupo consistindo em raízes, caules, tuberosidades de caules, folha, órgãos florais, raízes tuberosas, sementes e ápices dos brotos.

17. O método para produzir ácido hialurônico de acordo com o

Item 3, em que um ou mais tecidos são selecionados entre o grupo consistindo em epiderme, floema, tecidos moles, xilema e feixes vasculares.

18. Uma célula de planta transgênica ou uma planta transgênica tendo a capacidade de produzir ácido hialurônico por coexpressão de uma  
5 proteína com atividade de ácido hialurônico sintase e uma proteína exógena com atividade de nucleotídeo de açúcar sintase, ou uma progênie das mesmas, ou um órgão ou um tecido das mesmas tendo a mesma natureza que na planta.

19. Uma célula de planta transgênica ou uma planta transgênica  
10 tendo a capacidade de produzir ácido hialurônico, sendo transformada com um vetor de expressão recombinante contendo DNA codificando uma proteína com atividade de ácido hialurônico sintase e DNA codificando uma proteína com atividade de nucleotídeo de açúcar sintase sob controle de um promotor capaz de funcionamento em plantas; uma progênie tendo a mesma  
15 natureza das mesmas; ou um órgão ou um tecido das mesmas.

20. A célula de planta transgênica ou a planta transgênica, a progênie tendo a mesma natureza das mesmas, ou o órgão ou o tecido das mesmas de acordo com o Item 19, em que o promotor é um promotor órgão-específico ou um tecido-específico.

21. A célula de planta transgênica ou a planta transgênica, a progênie tendo a mesma natureza das mesmas, ou o órgão ou o tecido das mesmas de acordo com o Item 19 ou 20, em que o DNA codificando uma proteína com atividade de ácido hialurônico sintase é DNA de (a) ou (b) abaixo:

- 25 (a) DNA consistindo em uma sequência de nucleotídeos representada pela SEQ ID NO: 1 o 3; ou
- (b) DNA hibridizando à sequência de nucleotídeos complementar ao DNA consistindo na sequência de nucleotídeos de (a) sob condições estridentes, e o DNA codificando uma proteína com atividade de ácido hialurônico sintase.
- 30

22. A célula de planta transgênica ou a planta transgênica de acordo com qualquer um dos Itens 18 a 20, em que a proteína com atividade

de ácido hialurônico sintase é uma proteína de (a) ou (b) abaixo:

(a) uma proteína consistindo em uma sequência de aminoácidos representada pela SEQ ID NO: 2 ou 4; ou

(b) uma proteína consistindo na sequência de aminoácidos de (a) com um ou uns poucos aminoácidos eliminados, substituídos ou adicionados, e a proteína tendo atividade de ácido hialurônico sintase.

23. A célula de planta transgênica ou a planta transgênica, a progênie tendo a mesma natureza das mesmas, ou o órgão ou o tecido das mesmas de acordo com qualquer um dos Itens 18 a 22, em que o nucleotídeo de açúcar é UDP-N-acetilglucosamina e/ou ácido UDP-glucurônico.

24. A célula de planta transgênica ou a planta transgênica de acordo com qualquer um dos Itens 18 a 23, em que a proteína com atividade de nucleotídeo de açúcar sintase é no mínimo uma proteína selecionada entre o grupo consistindo em, glutamina:frutose-6-fosfato amidotransferase, UDP-N-acetilglucosamina difosforilase, fosfoacetil glucosamina mutase, glucosamina-6-fosfato N-acetiltransferase, glucosamina-1-fosfato N-acetiltransferase, fosfoglucomutase, N-acetilglucosamina quinase, hexoquinase, N-acetilglucosamina acetilase, UDP-glucose desidrogenase, UDP-glucose-1-fosfato uridiltransferase, inositol oxigenase, glucuronoquinase e glucuronato-1-fosfato uridil transferase.

25. A célula de planta transgênica ou a planta transgênica, a progênie tendo a mesma natureza das mesmas, ou o órgão ou o tecido das mesmas de acordo com qualquer um dos Itens 18 a 23, em que a proteína com atividade de nucleotídeo de açúcar sintase é no mínimo uma proteína selecionada entre o grupo consistindo em UDP-N-acetilglucosamina difosforilase, fosfoacetil glucosamina mutase, glucosamina-6-fosfato N-acetiltransferase, glucosamina-1-fosfato N-acetiltransferase, fosfoglucomutase, N-acetilglucosamina quinase, hexoquinase, N-acetilglucosamina acetilase, UDP-glucose desidrogenase, UDP-glucose-1-fosfato uridiltransferase, inositol oxigenase, glucuronoquinase e glucuronato-1-fosfato uridil transferase, e glutamina:frutose-6-fosfato amidotransferase.

26. A célula de planta transgênica ou a planta transgênica, a

progênie tendo a mesma natureza das mesmas, ou o órgão ou o tecido das mesmas de acordo com qualquer um dos Itens 18 a 23, em que a proteína com atividade de nucleotídeo de açúcar sintase é glutamina:frutose-6-fosfato amidotransferase e/ou UDP-glucose desidrogenase.

- 5 27. A célula de planta transgênica ou a planta transgênica, a  
progênie tendo a mesma natureza das mesmas, ou o órgão ou o tecido das  
mesmas de acordo com qualquer um dos Itens 19 a 26, em que DNA codifi-  
cando uma proteína com nucleotídeo de açúcar sintase atividade é derivado  
de vírus chlorella e/ou *Arabidopsis thaliana*.
- 10 28. A célula de planta transgênica ou a planta transgênica; a  
progênie tendo a mesma natureza das mesmas; ou o órgão ou o tecido das  
mesmas de acordo com qualquer um dos Itens 19 a 27, em que DNA codifi-  
cando uma proteína com atividade de nucleotídeo de açúcar sintase é DNA  
de (a) ou (b) abaixo:
- 15 (a) DNA consistindo em uma sequência de nucleotídeos representada pela  
SEQ ID NO: 5, 7 ou 9; ou  
(b) DNA hibridizando à sequência de nucleotídeos complementar ao DNA  
consistindo na sequência de nucleotídeos de (a) sob condições estrin-  
gentes, e o DNA codificando uma proteína com atividade de glutami-  
na:frutose-6-fosfato amidotransferase.
- 20 29. A célula de planta transgênica ou a planta transgênica, a  
progênie tendo a mesma natureza das mesmas, ou o órgão ou o tecido das  
mesmas de acordo com qualquer um dos Itens 18 a 27, em que a proteína  
com atividade de nucleotídeo de açúcar sintase é uma proteína de (a) ou (b)  
abaixo:
- 25 (a) uma proteína consistindo em uma sequência de aminoácidos represen-  
tada pela SEQ ID NO: 6, 8 ou 10; ou  
(b) uma proteína consistindo na sequência de aminoácidos de (a) com um  
ou uns poucos aminoácidos eliminados, substituídos ou adicionados, e  
a proteína tendo atividade de glutamina:frutose-6-fosfato amidotransfe-  
rase.
- 30 30. A célula de planta transgênica ou a planta transgênica, a

progênie tendo a mesma natureza das mesmas, ou o órgão ou o tecido das mesmas de acordo com qualquer um dos Itens 19 a 27, em que o DNA codificando uma proteína com atividade de nucleotídeo de açúcar sintase é DNA de (a) ou (b) abaixo:

- 5 (a) DNA consistindo em uma sequência de nucleotídeos representada pela SEQ ID NO: 11, 13, 15, 17, 19, or 21; ou
- (b) DNA hibridizando a DNA consistindo na sequência de bases complementar à sequência de bases de (a) sob condições estridentes, e o DNA codificando uma proteína com atividade de DP-glucose desidrogenase.

10

31. A célula de planta transgênica ou a planta transgênica, a progênie tendo a mesma natureza das mesmas, ou o órgão ou o tecido das mesmas de acordo com qualquer um dos Itens 18 a 27, em que a proteína com atividade de nucleotídeo de açúcar sintase é uma proteína de (a) ou (b) abaixo:

15

- (a) uma proteína consistindo em uma sequência de aminoácidos representada pela SEQ ID NO: 12, 14, 16, 18, 20 ou 22; ou
- (b) uma proteína consistindo na sequência de aminoácidos de (a) com um ou uns poucos aminoácidos eliminados, substituídos ou adicionados, e a referida proteína com atividade de UDP-glucose desidrogenase.

20

32. A célula de planta transgênica ou a planta transgênica, a progênie tendo a mesma natureza das mesmas, ou o órgão ou o tecido das mesmas de acordo com qualquer um dos Itens 18 a 31, em que a planta é qualquer planta selecionada entre o grupo consistindo em gimnospermas, pteridófitos e briófitas.

25

33. A célula de planta transgênica ou a planta transgênica; a progênie tendo a mesma natureza das mesmas; ou o órgão ou o tecido das mesmas de acordo com qualquer um dos Itens 18 a 31, em que o órgão é um ou mais órgãos selecionados entre o grupo consistindo em raízes, caules, tuberosidades de caules, folhas, órgãos florais, raízes tuberosas, sementes e ápices dos brotos.

30

34. A célula de planta transgênica ou a planta transgênica, a

progênie tendo a mesma natureza das mesmas, ou o órgão ou o tecido das mesmas de acordo com qualquer um dos Itens 18 a 31, em que o tecido é um ou mais tecidos selecionados entre o grupo consistindo em epiderme, floema, tecidos moles, xilema e feixes vasculares.

- 5                    35. Extratos vegetais obtidos da célula de planta transgênica ou da planta transgênica, da progênie tendo a mesma natureza das mesmas, ou do órgão ou do tecido das mesmas de acordo com qualquer um dos Itens 18 a 34.

- 10                   36. O extrato vegetal de acordo com o Item 35, em que o extrato vegetal contém ácido hialurônico.

37. Um vetor de expressão recombinante, compreendendo DNA codificando uma proteína com atividade de ácido hialurônico sintase e DNA codificando uma proteína com atividade de nucleotídeo de açúcar sintase sob controle de um promotor capaz de funcionamento em plantas.

- 15                   38. O vetor de expressão recombinante de acordo com o Item 37, em que o promotor é um promotor órgão-específico, ou um tecido-específico.

- 20                   39. O vetor de expressão recombinante de acordo com o Item 37 ou 38, em que o DNA codificando uma proteína com atividade de ácido hialurônico sintase é DNA de (a) ou (b) abaixo:

- (a) DNA consistindo em uma sequência de nucleotídeos representada pela SEQ ID NO: 1 o 3; ou  
(b) DNA hibridizando a DNA consistindo na sequência de nucleotídeos complementar à sequência de nucleotídeos de (a) sob condições es-  
25                   tringentes, e o DNA codificando uma proteína com ácido hialurônico a-  
tividade.

40. O vetor de expressão recombinante de acordo com o Item 37 ou 38, em que a proteína com atividade de ácido hialurônico sintase é uma proteína de (a) ou (b) abaixo:

- 30                   (a) uma proteína consistindo em uma sequência de aminoácidos represen-  
tada pela SEQ ID NO: 2 ou 4; ou  
(b) uma proteína consistindo na sequência de aminoácidos de (a) com um

ou uns poucos aminoácidos eliminados, substituídos ou adicionados, e a proteína tendo atividade de ácido hialurônico sintase.

41. Vetor de expressão recombinante de acordo com qualquer um dos Itens 37 a 40, em que o nucleotídeo de açúcar é UDP-N-acetilglucosamina e/ou ácido UDP-glucurônico.

42. Vetor de expressão recombinante de acordo com qualquer um dos Itens 37 a 41, em que a proteína com atividade de nucleotídeo de açúcar sintase é no mínimo uma proteína selecionada entre o grupo consistindo em glutamina:frutose-6-fosfato amidotransferase, UDP-N-acetilglucosamina difosforilase, fosfoacetil glucosamina mutase, glucosamina-6-fosfato N-acetiltransferase, glucosamina-1-fosfato N-acetiltransferase, fosfoglucomutase, N-acetilglucosamina quinase, hexoquinase, N-acetilglucosamina acetilase, UDP-glucose desidrogenase, UDP-glucose-1-fosfato uridiltransferase, inositol oxigenase, glucuronoquinase e glucuronato-1-fosfato uridil transferase.

43. Vetor de expressão recombinante de acordo com qualquer um dos Itens 37 a 41, em que a proteína com atividade de nucleotídeo de açúcar sintase é no mínimo uma proteína selecionada entre o grupo consistindo em UDP-N-acetilglucosamina difosforilase, fosfoacetil glucosamina mutase, glucosamina-6-fosfato N-acetiltransferase, glucosamina-1-fosfato N-acetiltransferase, fosfoglucomutase, N-acetilglucosamina quinase, hexoquinase, N-acetilglucosamina acetilase, UDP-glucose desidrogenase, UDP-glucose-1-fosfato uridiltransferase, inositol oxigenase, glucuronoquinase e glucuronato-1-fosfato uridil transferase, e glutamina:frutose-6-fosfato amidotransferase.

44. Vetor de expressão recombinante de acordo com qualquer um dos Itens 37 a 41, em que a proteína com atividade de nucleotídeo de açúcar sintase é glutamina:frutose-6-fosfato amidotransferase e/ou UDP-glucose desidrogenase.

45. Vetor de expressão recombinante de acordo com qualquer um dos Itens 37 a 41, em que o DNA codificando uma proteína com nucleotídeo de açúcar sintase atividade é DNA derivado de vírus chlorella e/ou A-

*rabidopsis thaliana*.

46. Vetor de expressão recombinante de acordo com qualquer um dos Itens 37 a 44, em que o DNA codificando uma proteína com nucleotídeo de açúcar sintase atividade é DNA de (a) ou (b) abaixo:

- 5 (a) DNA consistindo em uma sequência de nucleotídeos representada pela SEQ ID NO: 5, 7 ou 9; ou
- (b) DNA hibridizando a DNA consistindo na sequência de nucleotídeos complementar à sequência de nucleotídeos de (a) sob condições estridentes, e o referido DNA codificando uma proteína com atividade de glutamina:frutose-6-fosfato amidotransferase.
- 10

47. Vetor de expressão recombinante de acordo com qualquer um dos Itens 37 a 44, em que a proteína com atividade de nucleotídeo de açúcar sintase é uma proteína de (a) ou (b) abaixo:

- (a) uma proteína consistindo em uma sequência de aminoácidos representada pela SEQ ID NO: 6, 8 ou 10; ou
- 15 (b) uma proteína consistindo na sequência de aminoácidos de (a) com um ou uns poucos aminoácidos eliminados, substituídos ou adicionados, e a proteína tendo atividade de glutamina:frutose-6-fosfato amidotransferase.

20 48. Vetor de expressão recombinante de acordo com qualquer um dos Itens 37 a 44, em que o DNA codificando uma proteína com nucleotídeo de açúcar sintase atividade é DNA de (a) ou (b) abaixo:

- (a) DNA consistindo em uma sequência de nucleotídeos representada pela SEQ ID NO: 11, 13, 15, 17, 19, ou 21; ou
- 25 (b) DNA hibridizando a DNA consistindo na sequência de nucleotídeos complementar à sequência de nucleotídeos de (a) sob condições estridentes, e o referido DNA codificando uma proteína com atividade de UDP-glucose desidrogenase.

49. Vetor de expressão recombinante de acordo com qualquer um dos Itens 37 a 44, em que a proteína com atividade de nucleotídeo de açúcar sintase é uma proteína de (a) ou (b) abaixo:

- (a) uma proteína consistindo em uma sequência de aminoácidos represen-

tada pela SEQ ID NO: 12, 14, 16, 18, 20 ou 22; ou

- (b) uma proteína consistindo na sequência de aminoácidos de (a) com um ou uns poucos aminoácidos eliminados, substituídos ou adicionados, e a referida proteína tendo atividade de UDP-glucose desidrogenase.

5 50. Um método para produzir célula de planta transgênica ou a planta transgênica tendo a capacidade de produzir ácido hialurônico, o método compreendendo transformar uma célula vegetal ou uma planta usando qualquer vetor de acordo com os Itens 37 a 49.

10 51. Uma composição cosmética contendo ácido hialurônico como um agente ativo, em que o ácido hialurônico é obtido por qualquer um dos métodos para produzir ácido hialurônico de acordo com os Itens 1 a 17.

#### EFEITOS DA INVENÇÃO

De acordo com a presente invenção, ácido hialurônico, o qual não é produzido naturalmente em plantas, é produzido em plantas. De acordo com a presente invenção, um gene codificando uma proteína com atividade de ácido hialurônico sintase e um gene codificando uma proteína com atividade de nucleotídeo de açúcar sintase são expressados em plantas, de modo que ácido hialurônico é altamente produzido em plantas. Podem ser proporcionadas plantas ou células vegetais cultivadas capazes de produzir mais ácido hialurônico do que por métodos convencionais, um método do mesmo, e um vetor de expressão recombinante do mesmo. Portanto, a presente invenção pode proporcionar ácido hialurônico seguro produzido por plantas em baixo custo.

#### MELHOR MODO PARA REALIZAR A INVENÇÃO

25 Uma característica da presente invenção é um método para produzir ácido hialurônico, contendo coexpressão de uma proteína com atividade de ácido hialurônico sintase e uma proteína exógena com atividade de nucleotídeo de açúcar sintase em células vegetais ou plantas de modo a obter ácido hialurônico.

30 Outra característica da presente invenção é um método para produzir ácido hialurônico contendo:

- (1) transformar células vegetais ou plantas usando um vetor de expressão

recombinante que contém DNA codificando uma proteína com atividade de ácido hialurônico sintase e DNA codificando uma proteína com atividade de nucleotídeo de açúcar sintase sob o controle de um ou mais promotores capazes de funcionamento em plantas;

- 5 (2) cultivar o transformante obtido pela transformação; e
- (3) isolar ácido hialurônico produzido nos transformantes.

Ainda outra característica da presente invenção é composições cosméticas que contêm ácido hialurônico como um agente ativo, em que o ácido hialurônico é obtido pelo método para produzir ácido hialurônico acima.

Ainda outra característica da presente invenção é um vetor de expressão recombinante para produzir ácido hialurônico, em que o vetor de expressão recombinante contém DNA codificando uma proteína com atividade de ácido hialurônico sintase e DNA codificando uma proteína com atividade de nucleotídeo de açúcar sintase sob o controle de um ou mais promotores capazes de funcionamento em plantas.

Algumas características da presente invenção são células de plantas transgênicas ou plantas transgênicas, as progênies tendo a mesma natureza das mesmas, ou os órgãos ou os tecidos das mesmas, em que os transformantes obtiveram a capacidade de produzir ácido hialurônico por coexpressão de uma proteína com atividade de ácido hialurônico sintase e uma proteína exógena com atividade de nucleotídeo de açúcar sintase.

Uma característica da presente invenção é células de plantas transgênicas ou plantas transgênicas, as progênies tendo a mesma natureza das mesmas, ou os órgãos ou os tecidos das mesmas, em que os transformantes são transformadas usando um vetor de expressão recombinante contendo DNA codificando uma proteína com atividade de ácido hialurônico sintase e DNA codificando uma proteína com atividade de nucleotídeo de açúcar sintase sob o controle de um ou mais promotores capazes de funcionamento em plantas.

Outra característica da presente invenção é um método para produzir células de plantas transgênicas ou plantas transgênicas que têm a

capacidade de produzir ácido hialurônico. O método inclui transformar células vegetais ou plantas usando um vetor de expressão recombinante que contém DNA codificando uma proteína com atividade de ácido hialurônico sintase e DNA codificando uma proteína com atividade de nucleotídeo de açúcar sintase sob o controle de um ou mais promotores capazes de funcionamento em plantas.

O seguinte explica a presente invenção em detalhes.

#### Ácido Hialurônico Sintase

Na presente invenção, células vegetais ou plantas são transformadas usando DNA codificando uma proteína com atividade de ácido hialurônico sintase e DNA codificando uma proteína com atividade de nucleotídeo de açúcar sintase sob o controle de um ou mais promotores capazes de funcionamento em plantas.

Na presente invenção, uma proteína com atividade de ácido hialurônico sintase sintetiza ácido hialurônico usando ácido UDP-glucurônico e UDP-N-acetilglucosamina como substratos. O ácido hialurônico tem uma estrutura polimérica consistindo em unidades repetidas de ácido glucurônico e N-acetilglucosamina.

Na presente invenção, uma proteína com atividade de ácido hialurônico sintase, uma vez que a proteína tenha a natureza mencionada acima, não é particularmente limitada. Pode ser usada ácido hialurônico sintase (nas partes que se seguem ocasionalmente abreviada como HAS) derivada de animais, micro-organismos, vírus e semelhantes. Particularmente, pode ser usada ácido hialurônico sintase derivada de vertebrados tais como humanos, camundongos, coelhos, galinhas, gado e *Xenopus laevis*, micro-organismos tais como *Streptococcus* e *Pasteurella*, vírus tais como vírus *Chlorella* e semelhantes.

Mais especificamente, exemplos da proteína com atividade de ácido hialurônico sintase são HAS (A98R) derivada de vírus *Chlorella* PBCV-1; HAS1, HAS2 e HAS3 da ácido hialurônico sintase (hHAS) derivada de humanos; HAS1, HAS2 e HAS3 da ácido hialurônico sintase derivada de camundongo (mHAS); HAS1, HAS2 e HAS3 da ácido hialurônico sintase

derivada de galinha (gHAS); HAS2 da ácido hialurônico sintase derivada de rato (rHAS); HAS2 da ácido hialurônico sintase derivada de gado (bHAS); HAS1, HAS2 e HAS3 da da ácido hialurônico sintase derivada de *Xenopus laevis* (xHAS); a ácido hialurônico sintase derivada de *Pasteurella multocida* (pmHAS); a da ácido hialurônico sintase derivada de *Streptococcus pyogenes* (spHAS); e o gene de ácido hialurônico sintase (seHAS) derivada de *Streptococcus equisimilis*. Há vários tipos de genes de ácido hialurônico sintase (HAS) tais como HAS1, HAS2 e HAS3, no entanto, o tipo não é particularmente limitado. Pode ser usado qualquer um dos HAS descritos acima, entre os quais o HAS derivado de vírus chlorella é preferencial, HAS derivado de vírus chlorella os quais são mostrados por uma proteína consistindo em uma sequência de aminoácidos representada pela SEQ ID NO: 2 ou 4 é mais preferencial.

A proteína consistindo em uma sequência de aminoácidos representada pela SEQ ID NO: 2 ou 4 pode ser uma proteína tendo um ou uns poucos aminoácidos eliminados, substituídos ou adicionados, uma vez que a atividade de ácido hialurônico sintase não seja perdida. Por exemplo, a sequência de aminoácidos representada pela SEQ ID NO: 2 ou 4 pode ter uma eliminação de no mínimo um aminoácido, preferencialmente 1 a 10 aminoácidos, mais preferencialmente 1 a 5 aminoácidos, uma adição de no mínimo um aminoácido, preferencialmente 1 a 10 aminoácidos, mais preferencialmente 1 a 5 aminoácidos, ou uma substituição de no mínimo um aminoácido, preferencialmente 1 a 10 aminoácidos, mais preferencialmente 1 a 5 aminoácidos por outros aminoácidos. No entanto, mutações não estão limitadas às acima. As mutações referidas incluem mutações artificiais diferentes de mutações que ocorrem naturalmente. Por exemplo, é reportado que ácido hialurônico sintase derivada de *Pasteurella multocida* tem atividade de ácido hialurônico sintase mesmo se cerca de 270 aminoácidos no domínio ligado à membrana putativo e o domínio transmembrana putativo forem eliminados (Jing et al., 2000, *Glycobiology*, 10, 883-889). O número de aminoácidos mutados não é limitado, uma vez que a atividade de ácido hialurônico sintase não seja perdida. HAS pode ser uma proteína consistindo em uma parte da

sequência de aminoácidos representada pela SEQ ID NO: 2 ou 4, tendo atividade de ácido hialurônico sintase.

As atividades de ácido hialurônico sintase são determinadas como se segue. Para a reação, as amostras são incubadas em 0,2 ml de  
5 tampão Tris-HCl a 50 mM (pH 7,0) contendo ditioneitol a 1 mM, cloreto de magnésio a 20 mM, etileno glicol bis (éter b-aminoetilico)-N,N,N,N-tetraacético ácido a 1 mM, 15% de glicerol, ácido uridina-5'-difosfoglucurônico a 0,5 mM (nas partes que se seguem abreviado como UDP-GlcA), uridina-5'-difosfo-N-acetilglucosamina a 0,5 mM (nas partes que se seguem abreviado  
10 como UDP-GlcNAc abaixo), UDP-[14C]GlcA, 0,24 µM de UDP-[3H]GlcNAc a 0,1 µM, e 125 µg de ácido glucurônico por 1 hora. Depois da incubação, a reação é terminada por ebulição por 10 minutos. A mistura da reação é dividida em duas, 0,5 unidade de hialuronidase (Seikagaku Corporation) derivadas de *Streptococcus dysgalactiae* é adicionada a uma das soluções, e em  
15 seguida incubada a 30°C por 4 horas. Em seguida a solução da reação é fervida por 10 minutos para inativar a hialuronidase. A mistura da reação é fracionada por 0,5 ml usando cromatografia de coluna de Superdex Peptide HR10/30 (produzido pela Amersham Pharmacia Biotech Inc.) (elutriado: acetato de amônio a 0,2 M). Cada fração é medida por radioatividade. Em con-  
20 sequência, a atividade de ácido hialurônico pode ser determinada a partir da mistura da reação da amostra com base nas quantidades de produtos de baixo peso molecular digeridos pela hialuronidase.

A atividade de ácido hialurônico sintase também pode ser determinada usando o método de sanduíche, no qual o ácido hialurônico produzi-  
25 do é medido usando proteínas de ligação de ácido hialurônico.

De acordo com a presente invenção, o DNA codificando uma proteína com ácido hialurônico sintase e é DNA codificando uma proteína que sintetiza ácido hialurônico a partir de ácido UDP-glucurônico e UDP-N-acetilglucosamina como substratos, em que o ácido hialurônico tem uma  
30 estrutura polimérica consistindo em unidades repetidas de ácido glucurônico e N-acetilglucosamina.

De acordo com a presente invenção, o DNA codificando uma

proteína com atividade de ácido hialurônico sintase, uma vez que a proteína tenha as propriedades mencionadas acima, não é particularmente limitada. Podem ser usados genes de ácido hialurônico sintase (nas partes que se seguem ocasionalmente abreviados como HAS) derivados de animais, micro-organismos, vírus e semelhantes. Por exemplo, pode ser usado o gene de ácido hialurônico sintase derivado de vertebrados tais como humanos, camundongos, coelhos, galinhas, gado e *Xenopus laevis*, micro-organismos tais como *Streptococcus* e *Pasteurella*, e vírus tais como vírus chlorella e semelhantes.

Mais especificamente, genes HAS (A98R) derivados de vírus chlorella cepa PBCV-1; HAS1, HAS2 e HAS3 do gene de ácido hialurônico sintase derivado de humanos (hHAS); HAS1, HAS2 e HAS3 do gene de ácido hialurônico sintase derivado de camundongo (mHAS); HAS1, HAS2 e HAS3 do gene de ácido hialurônico sintase derivado de galinha (gHAS); HAS2 do gene de ácido hialurônico sintase derivado de rato (rHAS) gene; HAS2 do gene de ácido hialurônico sintase derivado de gado (bHAS); HAS1, HAS2 e HAS3 do gene de ácido hialurônico sintase derivado de *Xenopus laevis* (xHAS); o gene de ácido hialurônico sintase derivado de *Pasteurella multocida* (pmHAS); o gene de ácido hialurônico sintase derivado de *Streptococcus pyogenes* (spHAS) gene; o gene de ácido hialurônico sintase (se-HAS) derivado de *Streptococcus equisimilis* e semelhantes são incluídos. Há vários tipos de genes de ácido hialurônico sintase (HAS) tais como HAS1, HAS2 e HAS3, no entanto, o tipo não é particularmente limitado.

Pode ser usado qualquer um dos HAS descritos acima, entre os quais o gene HAS derivado de vírus chlorella é preferencial, genes HAS derivados de vírus chlorella os quais são mostrados por DNA consistindo em uma sequência de nucleotídeos representada pela SEQ ID NO: 1 ou 3 são mais preferenciais.

O DNA pode ser DNA que hibridiza com DNA consistindo na sequência de nucleotídeos complementar à sequência de nucleotídeos representada pela SEQ ID NO: 1 ou 3 sob condições estridentes, e codifica uma proteína com atividade de ácido hialurônico sintase.

O termo "condições estridentes" acima significa condições nas quais somente uma sequência de nucleotídeos codifica um polipeptídeo tendo atividade de ácido hialurônico sintase que é equivalente à de ácido hialurônico sintase tendo uma sequência de nucleotídeos representada pela SEQ ID NO: 1 ou 3 forma um híbrido com a sequência particular (que é um híbrido específico), ao passo que uma sequência de nucleotídeos que codifica um polipeptídeo tendo atividade *não-equivalente* não forma um híbrido com a sequência particular (que é um híbrido não-específico). Pessoas versadas na técnica podem escolher facilmente as condições referidas alterando as temperaturas da reação de hibridização e lavagem, e as concentrações de sais de soluções de reação de hibridização e soluções de lavagem, e semelhantes. Especificamente, um exemplo das condições estridentes da presente invenção são as condições nas quais a hibridização é realizada em 6×SSC (NaCl a 0,9 M, citrato trissódico a 0,09 M) ou 6×SSPE (NaCl a 3 M,  $\text{NaH}_2\text{PO}_4$  a 0,2 M, EDTA · 2Na a 20 mM, pH 7,4) a 42°C, e lavagem adicional é realizada usando 0,5×SSC a 42°C, no entanto, as condições estridentes não estão limitadas a semelhantes condições.

As condições estridentes são preferencialmente condições altamente estridentes. As condições altamente estridentes não são particularmente limitadas, uma vez que um gene codificando A98R não hibridize. Por exemplo, as condições altamente estridentes são condições nas quais é realizada lavagem em concentrações de sais equivalentes às de 0,1×SSC ou 0,1% de SDS a 60 °C.

#### Nucleotídeo de Açúcar

Exemplos de nucleotídeos de açúcar úteis são UDP-N-acetilglucosamina, ácido UDP-glucurônico, UDP-N-acetilgalactosamina, UDP-glucose, UDP-galactose, UDP-xilose, GDP-fucose, GDP-manose, ácido CMP-neuramínico, e semelhantes, no entanto, o nucleotídeo de açúcar não está limitado a estes. Entre estes nucleotídeos de açúcar, açúcar UDP é preferencial, UDP-N-acetilglucosamina e ácido UDP-glucurônico são mais preferenciais.

Melhorando os níveis de produção de semelhantes nucleotídeos

de açúcar, as quantidades de ácido UDP-glucurônico e UDP-N-acetilglucosamina que são substratos para ácido hialurônico sintase são aumentadas em células vegetais ou plantas. Para melhorar os níveis de produção de semelhante nucleotídeo de açúcar, uma enzima do caminho biossintético do nucleotídeo de açúcar, isto é, uma proteína com atividade de nucleotídeo de açúcar sintase é introduzida em células vegetais ou plantas.

Além disso, na presente invenção, é visto que a introdução simultânea de uma proteína com atividade de ácido hialurônico sintase em células vegetais ou plantas possibilita aumento da síntese de ácido hialurônico usando o ácido UDP-glucurônico e UDP-N-acetilglucosamina aumentados como substratos, em que o ácido hialurônico tem uma estrutura polimérica consistindo em unidades repetidas de ácido glucurônico e N-acetilglucosamina.

#### Enzimas Associadas com Caminho de Biossíntese de Nucleotídeos de Açúcar

De acordo com a presente invenção, células vegetais ou plantas são transformadas com DNA codificando uma proteína com atividade de ácido hialurônico sintase e DNA codificando uma proteína com atividade de nucleotídeo de açúcar sintase sob o controle de um ou mais promotores capazes de funcionamento em plantas.

Podem ser usadas proteínas com atividade de nucleotídeo de açúcar sintase tais como UDP-N-acetilglucosamina, ácido UDP-glucurônico, UDP-N-acetilgalactosamina, UDP-glucose, UDP-galactose, UDP-xilose, GDP-fucose, GDP-manose e ácido CMP-neuramínico. Proteínas tendo atividade de sintase dirigida para, entre estes nucleotídeos de açúcar, UDP-N-acetilglucosamina, ácido UDP-glucurônico, UDP-N-acetilgalactosamina, UDP-glucose, UDP-galactose, UDP-xilose são preferenciais. Proteínas tendo atividade de nucleotídeo de açúcar sintase dirigida para açúcares, UDP-N-acetilglucosamina e ácido UDP-glucurônico são mais preferenciais.

A proteína tendo atividade de nucleotídeo de açúcar sintase da presente invenção pode ser uma enzima catalisando as reações que são associadas com caminhos de biossíntese de nucleotídeo de açúcar. Exem-

plos das enzimas são glutamina:frutose-6-fosfato amidotransferase, UDP-N-acetilglucosamina difosforilase, fosfoacetil glucosamina mutase, glucosamina-6-fosfato N-acetiltransferase, glucosamina-1-fosfato-N-acetiltransferase, fosfoglucomutase, N-acetilglucosamina quinase, hexoquinase, N-acetilglucosamina acetilase, UDP-glucose desidrogenase, UDP-glucose-1-fosfato uridiltransferase, inositol oxigenase, glucuronoquinase, glucuronato-1-fosfato uridil transferase e semelhantes. No mínimo uma proteína selecionada entre o grupo consistindo em semelhantes proteínas com atividade de nucleotídeo de açúcar sintase pode ser expressada em células vegetais ou plantas.

Como uma proteína com atividade de nucleotídeo de açúcar sintase, no mínimo glutamina:frutose-6-fosfato amidotransferase também pode ser selecionada, e expressada em combinação com uma proteína tendo outra atividade de nucleotídeo de açúcar sintase em células vegetais ou plantas. Exemplos de outras proteínas tendo atividade de nucleotídeo de açúcar sintase diferente de glutamina:frutose-6-fosfato amidotransferase são no mínimo uma proteína de aminoácido selecionada entre o grupo consistindo em UDP-N-acetilglucosamina difosforilase, fosfoacetil glucosamina mutase, glucosamina-6-fosfato N-acetiltransferase, glucosamina-1-fosfato N-acetiltransferase, fosfoglucomutase, N-acetilglucosamina quinase, hexoquinase, N-acetilglucosamina acetilase, UDP-glucose desidrogenase, UDP-glucose-1-fosfato uridiltransferase, inositol oxigenase, glucuronoquinase, glucurônico ácido-1-fosfato uridil transferase e semelhantes.

Para obter os efeitos pretendidos da presente invenção, no mínimo, glutamina:frutose-6-fosfato amidotransferase é selecionada como uma proteína tendo a atividade de nucleotídeo de açúcar sintase, deste modo é obtido um efeito melhor do que o de métodos convencionais. Mais preferencialmente, tanto a glutamina:frutose-6-fosfato amidotransferase (nas partes que se seguem ocasionalmente abreviada como GFAT) e UDP-glucose desidrogenase (nas partes que se seguem ocasionalmente abreviada como UGD) são selecionadas. O único uso de UDP-glucose desidrogenase também mostra efeitos mais reforçados do que o de métodos convencionais.

### Glutamina: Frutose-6-Fosfato Amidotransferase

A proteína tendo atividade de glutamina:frutose-6-fosfato amidotransferase da presente invenção é uma proteína que sintetiza glucosamina-6-fosfato a partir de L-glutamina e frutose-6-fosfato como substratos.

5 A proteína tendo atividade de glutamina:frutose-6-fosfato amidotransferase da presente invenção, uma vez que a proteína tenha a natureza mencionada acima, não é particularmente limitada. Exemplos de semelhantes proteínas são GFAT derivadas de eucariotas, procariotas, vírus e semelhantes. Podem ser usadas GFAT derivadas de eucariotas tais como humanos, camundongos, milho, *Arabidopsis thaliana*, *Caenorhabditis elegans* e  
10 *Saccharomyces cerevisiae*, GFAT derivadas de procariotas tais como *Bacillus subtilis* e *Escherichia coli*, e GFAT derivadas de vírus tais como chlorella vírus, no entanto, a proteína não é limitada a estas.

A GFAT acima pode ser usada preferencialmente, entre as quais  
15 a GFAT derivada de vírus chlorella ou *Arabidopsis thaliana* é mais preferencial. A GFAT derivada de vírus chlorella apresentada por uma proteína consistindo na sequência de aminoácidos representada pela SEQ ID NO: 6 ou 8, e a GFAT derivada de *Arabidopsis thaliana* apresentada por uma proteína consistindo em uma sequência de aminoácidos representada pela SEQ ID  
20 NO: 10 são especialmente preferenciais.

A proteína também pode ser uma proteína consistindo na sequência de aminoácidos representada pela SEQ ID NO: 6, 8 ou 10, ou a proteína com um ou uns poucos aminoácidos eliminados, substituídos ou adicionados a sua sequência de aminoácidos uma vez que a atividade de glutamina:frutose-6-fosfato amidotransferase não seja perdida. Por exemplo, a  
25 sequência de aminoácidos representada pela SEQ ID NO: 6, 8 ou 10 pode ter uma eliminação de no mínimo um aminoácido, preferencialmente 1 a 10 aminoácidos, e mais preferencialmente 1 a 5 aminoácidos. A sequência de aminoácidos representada pela SEQ ID NO: 6, 8 ou 10 pode ter uma adição  
30 de no mínimo um aminoácido, preferencialmente 1 a 10 aminoácidos, mais preferencialmente 1 a 5 aminoácidos. A sequência de aminoácidos representada pela SEQ ID NO: 6, 8 ou 10 pode ter uma substituição de no mínimo

um aminoácido, preferencialmente 1 a 10 aminoácidos, mais preferencialmente 1 a 5 aminoácidos, e mais preferencialmente 1 a 5 aminoácidos, com outros aminoácidos. No entanto, mutações não estão limitadas aos exemplos mencionados acima. As referidas mutações incluem mutações artificiais diferentes de mutações que ocorrem naturalmente. O número de aminoácidos mutados, uma vez que a atividade de GFAT não seja perdida, não é particularmente limitado. Um exemplo de mutações que ocorrem naturalmente é o GFAT derivado de vírus chlorella cepa Hirosaki e tendo uma sequência de SEQ ID NO: 5 que diverge da GFAT conhecida de vírus chlorella cepa PBCV-1 e K2I no mínimo por 2% em suas sequências de aminoácidos devido a mutação. A proteína pode ser uma proteína tendo uma parte da sequência de aminoácidos representada pela SEQ ID NO: 6, 8 ou 10, e tendo atividade de glutamina:frutose-6-fosfato amidotransferase.

A atividade de glutamina:frutose-6-fosfato amidotransferase pode ser avaliada adicionando a solução enzimática à mistura da reação (pH 7,0) contendo frutose-6-fosfato (15 mM), L-glutamina (15 mM), EDTA (1 mM), DTT (1 mM) e  $\text{KH}_2\text{PO}_4$  (60 mM), e em seguida incubando a 37°C por uns poucos horas, e subsequentemente medindo a quantidade de glucosamina-6-fosfato ou ácido glutâmico produzido. Especificamente, exemplos dos métodos são o método de Reissig (J.Biol.Chem, 1955, 217(2), 959-66), o qual é uma versão modificada do método de Morgan & Elson para medir glucosamina-6-fosfato; análise enzimática (J Biochem Biophys Methods, 2004, 59(3), 201-8) para medir ácido glutâmico usando ácido glutâmico desidrogenase; e semelhantes.

DNA codificando uma proteína com atividade de glutamina:frutose-6-fosfato amidotransferase da presente invenção é DNA codificando uma proteína que tem atividade enzimática para sintetizar glucosamina-6-fosfato a partir de L-glutamina e frutose-6-fosfato como substratos.

DNA codificando uma proteína com atividade de glutamina:frutose-6-fosfato amidotransferase da presente invenção, uma vez que as proteínas tenham as propriedades mencionadas acima, não é particularmente limitado. Exemplos de semelhantes genes GFAT são genes GFAT deriva-

dos de eucariotas, procariotas, vírus e semelhantes. Podem ser usados genes GFAT derivados de eucariotas tais como humanos, camundongos, milho, *Arabidopsis thaliana*, *Caenorhabditis elegans* e *Saccharomyces cerevisiae*, genes GFAT derivados de procariotas tais como *Bacillus subtilis* e *Escherichia coli*, genes GFAT derivados de vírus tais como vírus chlorella e semelhantes. Há vários tipos de genes GFAT tais como GFAT1 e GFAT2, no entanto, o tipo não é particularmente limitado.

Os genes GFAT descritos acima podem ser usados preferencialmente, entre os quais o gene GFAT derivado de *Chlorella vírus* ou derivado de *Arabidopsis thaliana* é mais preferencial. O gene GFAT derivado de *Chlorella vírus* e consistindo na sequência de nucleotídeos representada pela SEQ ID NO: 5 ou 7, ou o gene GFAT derivado de *Arabidopsis thaliana* e consistindo na sequência de nucleotídeos representada pela SEQ ID NO: 9 é especialmente preferencial.

O DNA também pode ser DNA hibridizando com uma sequência de nucleotídeos complementar à sequência de nucleotídeos representada pela SEQ ID NO: 5, 7 ou 9 sob condições estridentes, e codificando uma proteína com atividade de GFAT.

O termo "condições estridentes" acima significa condições nas quais somente uma sequência de nucleotídeos codificando um polipeptídeo com uma atividade de glutamina: frutose-6-fosfato amidotransferase equivalente à atividade de glutamina: frutose-6-fosfato amidotransferase representada pela SEQ ID NO: 5, 7 ou 9 forma um híbrido com a sequência particular (isto é, um híbrido específico), ao passo que uma sequência de nucleotídeos codificando um polipeptídeo com atividade *não-equivalente* forma um híbrido com a particular sequência (isto é, um híbrido não-específico). Pessoas versadas na técnica podem escolher facilmente as condições referidas alterando as temperaturas para a reação de hibridização e lavagem, concentrações de sal de soluções de reação de hibridização e soluções de lavagem, e semelhantes. Um exemplo das condições estridentes da presente invenção são condições nas quais é realizada hibridização usando 6×SSC (NaCl a 0,9 M, citrato trissódico a 0,09 M) ou 6×SSPE (NaCl a 3M,  $\text{NaH}_2\text{PO}_4$  a 0,2

M, EDTA · 2Na a 20 mM, pH 7,4) a 42°C, e subsequentemente é realizada lavagem usando 0,5×SSC a 42°C, no entanto, as condições estridentes não são limitadas a estas.

As condições estridentes são preferencialmente condições altamente estridentes. As "condições altamente estridentes" são, por exemplo, condições nas quais é realizada lavagem em concentrações de sal equivalentes às de 0,1×SSC e 0,1% de SDS a 60°C.

#### UDP-Glucose Deidrogenase

A proteína tendo atividade de UDP-glucose desidrogenase da presente invenção pode ser uma proteína que sintetiza ácido UDP-glucurônico a partir de UDP-glucose como um substrato.

A proteína tendo atividade de UDP-glucose desidrogenase da presente invenção, uma vez que a proteína tenha a natureza mencionada acima, não é particularmente limitada. Exemplos das proteínas são UGD derivado de eucariotas, procariotas, vírus e semelhantes. Pode ser usado UGD derivado de eucariotas tais como UGD derivado de humanos, gado, camundongos, choupas, canas-de-açúcar e *Arabidopsis thaliana*, UGD derivado de procariotas tais como *Escherichia coli*, *Pasteurella multocida* e *Lactobacillus lactis*, e UGD derivado de vírus tal como *chlorella* vírus, no entanto, a proteína da invenção não é limitada a estes.

O UGD descrito acima pode ser usado preferencialmente, entre os quais UGD derivado de vírus *chlorella* ou *Arabidopsis thaliana* é mais preferencial. UGD derivado de vírus *chlorella* e mostrado por uma proteína consistindo em uma sequência de aminoácidos representada pela SEQ ID NO: 12 ou 14, ou UGD derivado de *Arabidopsis thaliana* mostrado por uma proteína consistindo em uma sequência de aminoácidos representada pela SEQ ID NO: 16, 18, 20 ou 22 é particularmente preferencial.

A proteína pode ser uma proteína consistindo na sequência de aminoácidos representada pela SEQ ID NO: 12, 14, 16, 18, 20, ou 22, ou uma proteína tendo uma mutação tal como uma eliminação de um ou uns poucos aminoácidos, uma substituição, ou uma adição uma vez que a atividade de UDP-glucose desidrogenase não seja perdida. Por exemplo, a se-

quência de aminoácidos representada pela SEQ ID NO: 12, 14, 16, 18, 20, ou 22 pode ter uma eliminação de no mínimo um aminoácido, preferencialmente 1 a 10 aminoácidos, e mais preferencialmente 1 a 5 aminoácidos. A sequência de aminoácidos representada pela SEQ ID NO: 12, 14, 16, 18, 20, ou 22 pode ter uma adição de no mínimo um aminoácido, preferencialmente 1 a 10 aminoácidos, e mais preferencialmente 1 a 5 aminoácidos. A sequência de aminoácidos representada pela SEQ ID NO: 12, 14, 16, 18, 20, ou 22 também pode ter uma substituição de no mínimo 1 aminoácido, preferencialmente 1 a 10 aminoácidos, e mais preferencialmente 1 a 5 aminoácidos por outros aminoácidos. No entanto, as mutações não estão limitadas às acima. As mutações referidas também incluem mutações artificiais diferentes de mutações que ocorrem naturalmente. O número de aminoácidos mutados não é particularmente limitado, uma vez que a atividade de UGD não seja perdida. A proteína também pode ser uma proteína tendo uma parte da sequência de aminoácidos representada pela SEQ ID NO: 12, 14, 16, 18, 20, ou 22, e tendo UDP-glucose desidrogenase.

A atividade de UDP-glucose desidrogenase é avaliada adicionando a solução enzimática à mistura da reação (pH 8,0) contendo UDP-glucose (4 mM), NAD<sup>+</sup> (1 mM), EDTA (1 mM) e Tris-HCl (20 mM), realizando a reação a 37°C, e subsequentemente medindo o ácido UDP-glucurônico ou NADH produzido. Especificamente, NADH pode ser medido de acordo com o relatório de Tenhaken e Thulke (Plant Physiol. 1996, 112:1127-34).

O DNA codificando uma proteína com atividade de UDP-glucose desidrogenase da presente invenção é DNA codificando uma proteína que tem atividade enzimática para sintetizar ácido UDP-glucurônico a partir de UDP-glucose como um substrato.

O DNA codificando uma proteína com atividade de UDP-glucose desidrogenase da presente invenção, uma vez que a proteína tenha a natureza mencionada acima, não é particularmente limitado. Exemplos do DNA são genes UGD derivados de eucariotas, procariotas, vírus e semelhantes. Podem ser usados genes UGD derivados de eucariotas tais como humanos, gado, camundongos, choupas, canas-de-açúcar e *Arabidopsis thaliana*, ge-

nes UGD derivados de procariotas tais como *Escherichia coli*, *Pasteurella multocida* e *Lactobacillus lactis*, e genes UGD derivados de vírus tais como vírus *chlorella* e semelhantes, no entanto, os genes UGD não estão limitados a estes.

5 Os genes UGD descritos acima podem ser usados preferencialmente, entre os quais gene UGD derivado de vírus *chlorella* ou *Arabidopsis thaliana* é mais preferencial, gene UGD derivado de vírus *chlorella* e tendo uma sequência de nucleotídeos representada pela SEQ ID NO: 11, ou 13 ou  
10 gene UGD derivado de *Arabidopsis thaliana* tendo uma sequência de nucleotídeos representada pela SEQ ID NO: 15, 17, 19 ou 21 é particularmente preferencial.

O DNA pode ser DNA hibridizando com DNA consistindo em uma sequência de nucleotídeos complementar às sequências de nucleotídeos representada pela SEQ ID NO: 11, 13, 15, 17, 19, ou 21 sob condições  
15 estridentes, e codificando uma proteína com atividade de UDP-glucose desidrogenase.

As "condições estridentes" descritas acima são condições nas quais somente uma sequência de nucleotídeos codificando um polipeptídeo com a atividade equivalente à de UDP-glucose desidrogenase representada  
20 pela SEQ ID NO: 11, 13, 15, 17, 19, ou 21 forma um híbrido particular com a sequência particular (isto é, um híbrido específico), ao passo que uma sequência de nucleotídeos codificando um polipeptídeo com atividade *não-equivalente* não forma um híbrido com a sequência particular (isto é, um híbrido não-específico). Pessoas versadas na técnica podem facilmente esco-  
25 lher as condições referidas alterando as temperaturas para reação de hibridização e lavagem, ajustando as concentrações de sal das soluções da reação de hibridização e soluções de lavagem, e semelhantes. Um exemplo de condições estridentes são condições nas quais a hibridização é realizada usando 6×SSC (NaCl a 0,9 M, citrato trissódico a 0,09 M) ou 6×SSPE (NaCl  
30 a 3 M,  $\text{NaH}_2\text{PO}_4$  a 0,2 M, EDTA a 20 mM) · 2Na, pH7,4) a 42°C, e em seguida é realizada lavagem usando 0,5×SSC. No entanto, as condições estridentes não estão limitadas a estas.

As condições estridentes são preferencialmente condições altamente estridentes. Condições altamente estridentes são, por exemplo, condições nas quais lavagem é realizada em concentrações de sal equivalentes às de 0,1×SSC ou 0,1% de SDS a 60°C.

## 5 Um Vetor de Expressão Recombinante e Transformação

Células vegetais transgênicas ou plantas transgênicas capazes de produzir ácido hialurônico, progênies, ou órgãos ou tecidos das mesmas tendo a mesma natureza das mesmas podem ser obtidas transformando hospedeiros usando um vetor de expressão recombinante. Vetor de expressão recombinante contém DNA codificando uma proteína com atividade de ácido hialurônico sintase e DNA codificando uma proteína com atividade de nucleotídeo de açúcar sintase sob o controle de um ou mais promotores capazes de funcionamento em plantas.

Uma proteína tendo atividade de ácido hialurônico sintase e uma proteína com atividade de nucleotídeo de açúcar sintase são expressadas nas células de plantas transgênicas ou nas plantas transgênicas acima, as progênies, ou os órgãos ou os tecidos das mesmas tendo a mesma natureza das mesmas.

"Uma proteína exógena com atividade de nucleotídeo de açúcar sintase" é, ao contrário de uma proteína endógena, uma proteína estranha tendo atividade de nucleotídeo de açúcar sintase e sendo introduzida recentemente em células vegetais ou plantas de fora. Um exemplo de um método para "introduzir recentemente um gene semelhante de fora" inclui transformação usando um vetor de expressão recombinante e semelhantes. "Introduzir recentemente um gene de fora" inclui transformar um gene endógeno de fora da célula usando um vetor de expressão recombinante. Uma vez que não foram encontradas proteínas com atividade de ácido hialurônico sintase em células vegetais e plantas, é óbvio que a proteína é exógena sem a descrição.

As células de plantas transgênicas ou plantas transgênicas acima, as progênies, ou os órgãos ou tecidos das mesmas tendo a mesma natureza das mesmas possibilitam alto nível de produção de ácido hialurônico.

Isto é demonstrado abaixo.

Na presente invenção, os hospedeiros significam quaisquer plantas inteiras, sementes, órgãos de plantas (por exemplo, pétalas, raízes, caules, tuberosidades de caules, folhas, órgãos florais, raízes tuberosas, sementes, ápices dos brotos e semelhantes), tecidos de plantas (por exemplo, epiderme, floema, tecidos moles, xilema, feixes vasculares e semelhantes), ou células vegetais cultivadas.

Na presente especificação, plantas significam quaisquer plantas multicelulares inclusive espermatófitos, gimnospermas, pteridófitos, briófitos, líquens e semelhantes, e incluem quaisquer plantas inteiras, sementes, órgãos de plantas (por exemplo, pétalas, raízes, caules, tuberosidades de caules, folhas, órgãos florais, raízes tuberosas, sementes, ápices dos brotos e semelhantes), tecidos de plantas (por exemplo, epiderme, floema, tecidos moles, xilema, feixes vasculares e semelhantes), ou células vegetais cultivadas.

Também pode ser produzido ácido hialurônico deste modo cultivando os transformantes resultantes da transformação, e isolando o ácido hialurônico produzido desse modo.

Vetores que são usados de modo geral para produzir células de plantas transgênicas ou plantas transgênicas podem ser usadas como vetores de expressão recombinantes.

Os vetores referidos não são particularmente limitados uma vez que os vetores compreendam uma sequência de promotor capaz de ser transcrita em células vegetais e um sítio de poliadenilação requerido para estabilização de transcriptos. Por exemplo, podem ser usados vetores tais como "pBI121", "pBI221", "pBI101" e "pIG121Hm".

Quando células vegetais cultivadas são usadas como hospedeiros, a transformação pode ser realizada introduzindo um vetor de expressão recombinante para produzir ácido hialurônico em células vegetais cultivadas pelo método de eletroporação, o método do vetor binário de *Agrobacterium* ou o método de Bombardeamento de Partículas. Vetor inclui DNA codificando uma proteína com atividade de ácido hialurônico sintase e DNA codifi-

cando uma proteína com atividade de nucleotídeo de açúcar sintase colocado sob o controle de um ou mais promotores capazes de funcionamento em plantas. As células vegetais nas quais o vetor de expressão foi introduzido são, por exemplo, selecionadas com base na resistência a antibióticos tais  
 5 como canamicina e semelhantes. As células vegetais transformadas podem ser usadas para cultura celular, cultura de tecido e cultura de órgãos. Também é possível regenerar plantas usando um método de cultura de tecido vegetal previamente conhecido.

Exemplos de células vegetais submetidas à transformação incluem células BY-2 e células T-13 derivadas de tabaco, células kurodagosun  
 10 derivadas de cenouras, células VR e células VW derivadas de uvas, células PAR, células PAP e células PAW derivadas de *Phytolacca americana* L., células T87 derivadas de *Arabidopsis thaliana*, células Asp-86, células A. per, células A. pas e células A. plo derivadas de aspargos, células Cba-1 derivadas  
 15 das de melancias, células Sly-1 derivadas de tomates, células 1-Mar derivadas de hortelã-pimenta, células CRA e células V208 derivadas de pervincas de Madagascar, células Spi-WT, células Spi-I-1 e células Spi-12F derivadas do espinafre, células Lcy-1, células LcyD6 e células LcyD7 derivadas de cucurbitáceas, células OS-1 derivadas de *Oryza sativa*, células Vma-1 derivadas  
 20 das de *Vinca rosea*, células PSB, células PSW e células PSG derivadas de gergelim, e células ZE3 derivadas de *Zinnia elegans*.

Quando células vegetais cultivadas são usadas como hospedeiros, é realizada transformação introduzindo um vetor de expressão recombinante para produzir ácido hialurônico na planta pelo método do vetor binário  
 25 de *Agrobacterium*, o método de Bombardeamento de Partículas ou o método de eletroporação em protoplastos, onde o vetor inclui DNA codificando uma proteína com atividade de ácido hialurônico sintase e DNA codificando uma proteína com atividade de nucleotídeo de açúcar sintase sob o controle de um ou mais promotores capazes de funcionamento em plantas, e em seguida isolando os tecidos tumorais, brotos, raízes filamentosas e semelhantes  
 30 resultantes da transformação.

Tecidos tumorais, brotos, raízes filamentosas e semelhantes ob-

5 tidos conforme descrito acima podem ser usados diretamente para cultura celular, cultura de tecido ou cultura de órgãos. Estes transformantes também podem ser regenerados em plantas por métodos de cultura de tecido de plantas previamente conhecidos, aplicando hormônios vegetais em concentrações apropriadas e semelhantes.

10 Para regenerar uma planta de uma célula na qual foi introduzido um gene de ácido hialurônico sintase, uma célula vegetal semelhante pode ser cultivada sobre um meio de regeneração, um meio MS livre de hormônios ou semelhante. O enraizamento da planta jovem resultante pode ser plantado no solo e cultivado. Os métodos para regeneração diferem dependendo do tipo de célula vegetal, mas é possível usar um método previamente conhecido para plantar cultura de tecido.

Por exemplo, o método de Fujimura et al. (Fujimura et al. (1955), Plant Tissue Culture Lett., Vol. 2: p. 74) pode ser usado para *Oryza sativa*, o 5 método de Shillito et al. (Shillito et al. (1989), Bio/Technology, Vol. 7: p. 581, e Gorden-Kamm (1990), Plant Cell, 2, 603) pode ser usado para milho, e o método de Visser et al. (Visser et al. (1989), Theor. Appl. Genet, Vol. 78: p 589) pode ser usado para batatas. O método de Nagata et al. (Nagata et al. (1971), Planta 99, 12) pode ser usado para tabaco, e o método de Akama et al. (Akama et al. (1992), Plant Cell Rep., Vol. 12: p. 7) pode ser usado para 20 *Arabidopsis thaliana*.

Plantas produzidas por semelhantes métodos, ou as progênie; por exemplo, plantas regeneradas de sementes, tuberosidades de caules, mudas e semelhantes) tendo a mesma natureza das mesmas, também são 25 objetos da presente invenção.

De modo a produzir uma proteína com atividade de ácido hialurônico sintase e uma proteína tendo atividade para sintetizar nucleotídeo de açúcar em plantas, e adicionalmente para produzir e acumular ou secretar ácido hialurônico, DNA codificando uma proteína com atividade de ácido hialurônico sintase e DNA codificando uma proteína com atividade de nucleotídeo de açúcar sintase são preferencialmente colocados sob o controle de 30 um ou mais promotores capazes de funcionamento em plantas, de modo

que o DNA é especificamente expressado nos tecidos ou órgãos desejados.

Para a expressão ser controlada deste modo, um promotor tecido-específico ou um promotor órgão-específico pode ser adicionalmente inserido em um vetor de expressão recombinante.

5                Se ao invés for usado um promotor de estágio específico, um gene-alvo pode ser expressado somente durante um período em particular. Portanto, a produtividade pode ser aprimorada somente durante o período em particular. Por exemplo, o uso de um promotor específico do estágio vegetativo melhora a produtividade somente durante o estágio vegetativo.

10              Exemplos de promotores órgão-específicos são promotores específicos de raízes, promotores específicos de tubérculos, promotores específicos de folhas, promotores específicos de sementes, promotores específicos de caules e semelhantes.

                 Exemplos de promotores tecido-específicos são promotores específicos de tecidos verdes e semelhantes.

15              Mais especificamente, promotores úteis incluem promotores constitutivamente de alta expressão tais como um promotor CaMV35S, o qual é um promotor do RNA do vírus do mosaico da couve-flor 35S, e semelhantes. Promotores específicos de tecidos verdes incluem, por exemplo, um

20              promotor rbs para um gene codificando a proteína de pequena subunidade de ribulose-1,5-bisfosfato carboxilase, um promotor CAB para um gene codificando a proteína de ligação de clorofila a/b, e um promotor GapA para um gene codificando a proteína de subunidade A de gliceraldeído-3-fosfato desidrogenase. Promotores específicos de sementes incluem um promotor LOX

25              do gene da lipoxigenase, um promotor Psl do gene da lectina, um promotor AmylA do gene da amilase, e semelhantes. Promotores específicos de raízes incluem um promotor A6H6H do gene da hiosciamina 6b-hidroxilase, um promotor PMT da putrescina N-metiltransferase, e semelhantes. Promotores específicos para caule incluem um promotor Sus4 da sucrose sintase, um

30              promotor de patatina para um gene codificando a glicoproteína, e semelhantes.

Também é concebível controlar a expressão de DNA codificando

uma proteína com atividade de ácido hialurônico sintase e DNA codificando uma proteína com atividade de nucleotídeo de açúcar sintase usando um promotor induzível. Exemplos dos promotores de indução são descritos abaixo.

- 5 Exemplos de promotores indutíveis incluem um PR1a promotor, o qual é um promotor de um gene relacionado com resistência a doença cujo nível de expressão é reforçado por lesão ou uma adição de ácido salicílico, e um rd29A promotor cujo nível de expressão é reforçado por secagem, baixa temperatura, alta concentração de sal, adição de ácido abscísico e seme-
- 10 lhantes. Exemplos de promotores cuja expressão é induzida por compostos usados como químicos agrícolas incluem um GST-27 promotor para um gene codificando uma subunidade de proteína de 27 KDa de glutathione-S-transferase e é induzido por protetores de herbicida, um promotor de quinase e alguns PR promotores para genes sendo induzidos por éster S-metilico
- 15 de ácido benzo(1,2,3)-tiadiazol-7-carbotiólico (BTH). Além disso, de modo a expressar de modo mais estável DNA codificando uma proteína com atividade de ácido hialurônico sintase e DNA codificando uma proteína com atividade de nucleotídeo de açúcar sintase em células vegetais, podem ser utilizados isolantes, um peptídeo de sinal pode ser adicionado para localizar a
- 20 uma proteína com atividade de ácido hialurônico sintase e uma proteína com atividade de nucleotídeo de açúcar sintase em uma organela-alvo, uma parte de ácido hialurônico sintase pode ser substituída ou eliminada, e semelhantes.

25 As plantas submetidas a transformação incluem quaisquer plantas nas quais é possível transferência de gene.

As plantas ou os os corpos das plantas da presente invenção incluem monocotiledôneas e dicotiledôneas de angiospermas, e gimnospermas. As plantas referidas incluem opcionalmente plantas úteis, particularmente plantas de colheitas, plantas de vegetais, plantas de flores e plantas

30 lenhosas.

As plantas ou os corpos das plantas da presente invenção também incluem pteridófitos e briófitas.

Exemplos de espécies de plantas úteis na presente invenção especificamente incluem plantas pertencentes às famílias das Solanáceas, Gramíneas, Crucíferas, Rosáceas, Leguminosas, Cucurbitáceas, Labiadas, Liliáceas, Quenopodiáceas, Umbelíferas, Mirtáceas, Convolvuláceas, e se-  
 5 melhantes.

Exemplos de plantas pertencentes às Solanáceas incluem as plantas pertencentes ao gênero *Nicotiana*, *Solanum*, *Datura*, *Lycopersion* ou *Petunia*, e, por exemplo, incluem tabaco, berinjelas, batatas, tomates, chili peppers, petúnias e semelhantes.

10 Exemplos de plantas pertencentes às Gramíneas incluem as plantas pertencentes ao gênero *Oryza*, *Hordeum*, *Secale*, *Saccharum*, *Echinochloa* ou *Zea*, e, por exemplo, incluem *Oryza sativa*, cevada, centeio, pilriteiro (cockspur), sorgos, milho, cana-de-açúcar e semelhantes.

15 Exemplos de plantas pertencentes às Crucíferas incluem as plantas pertencentes ao gênero *Raphanus*, *Brassica*, *Arabidopsis*, *Wasabia* ou *Capsella*, e, por exemplo, incluem rabanete japonês (*Raphanus sativus longipinnatus*), colza, *Arabidopsis thaliana*, rábano silvestre japonês, Shepherd's Purse e semelhantes.

20 Exemplos de plantas pertencentes às Rosáceas incluem as plantas pertencentes ao gênero *Orunus*, *Malus*, *Pynus*, *Fragaria* ou *Rosa*, e, por exemplo, incluem damascos japoneses, pêssegos, maçãs, peras, morango, rosas e semelhantes.

25 Exemplos de plantas pertencentes às Leguminosas incluem as plantas pertencentes ao gênero *Glycinia*, *Vigna*, *Phaseolus*, *Pisum*, *Vicia*, *Arachis*, *Trifolium*, Alfalfa ou *Medicago*, e, por exemplo, incluem feijões-soja, feijões azuki, feijões-manteiga, ervilhas, favas, amendoins, trevos e semelhantes.

30 Exemplos de plantas pertencentes às Cucurbitáceas incluem as plantas pertencentes ao gênero *Luffa*, *Cucurbita* ou *Cucumis*, e, por exemplo, incluem cucurbitáceas, abóboras-morangas, pepino, melões e semelhantes.

Exemplos de plantas pertencentes às Labiadas incluem as plan-

tas pertencentes ao gênero *Lavandula*, *Mentha* ou *Perilla*, e, por exemplo, incluem lavanda, menta, manjeriço japonês e semelhantes.

Exemplos de plantas pertencentes às Liliáceas incluem as plantas pertencentes ao gênero *Allium*, *Lilium* ou *Tulipa*, e, por exemplo, incluem  
5 cebolas galesas, alho, lírios, tulipas e semelhantes.

Exemplos de plantas pertencentes às Quenopodiáceas incluem as plantas pertencentes ao gênero *Spinacia*, e, por exemplo, incluem beterrabas, espinafre e semelhantes.

Exemplos de plantas pertencentes às Umbelíferas incluem as  
10 plantas pertencentes ao gênero *Angelica*, *Daucus*, *Cryptotaenia* ou *Apitum*, e, por exemplo, incluem shishiuídos, cenouras, ceratófilos, aipos e semelhantes.

Exemplos de plantas pertencentes às Convolvuláceas incluem as plantas pertencentes ao gênero *Ipomoea*, e, por exemplo, incluem bata-  
15 tas-doces e semelhantes.

As progênes tendo a mesma natureza que as plantas transgênicas acima, ou os órgãos ou tecidos das mesmas também são os objetos da presente invenção.

As células de plantas transgênicas as quais produzem uma proteína com atividade de ácido hialurônico sintase e uma proteína tendo uma  
20 atividade de nucleotídeo de açúcar sintase são incluídas na presente invenção. As plantas transgênicas as quais produzem uma proteína tendo atividade de ácido hialurônico sintase e uma proteína tendo atividade de nucleotídeo de açúcar sintase, as progênes tendo a mesma natureza das mesmas,  
25 ou os órgãos ou tecidos das mesmas, também são incluídos.

#### Extração de Ácido Hialurônico

Abaixo está um exemplo de métodos para isolar ou obter ácido hialurônico por coexpressão de uma proteína com atividade de ácido hialurônico sintase e uma proteína exógena tendo uma atividade de sintetizar o  
30 nucleotídeo de açúcar, e extrair ácido hialurônico das células de plantas transgênicas ou das plantas transgênicas que adquiriram a capacidade para produzir ácido hialurônico, progênes, ou os órgãos ou tecidos das mesmas

tendo a mesma natureza das mesmas.

As células de plantas transgênicas ou as plantas transgênicas são cultivadas ou desenvolvidas, o ácido hialurônico é produzido, e subsequentemente, o ácido hialurônico é opcionalmente extraído das células de plantas transgênicas ou das plantas transgênicas por um método conhecido. Os extratos são verificados para ácido hialurônico.

Por exemplo, as plantas transgênicas, as progênes tendo a mesma natureza das mesmas, os órgãos ou os tecidos das mesmas, ou semelhantes podem ser subsequentemente secos, triturados e em seguida extraídos por um solvente orgânico apropriado. O extrato contendo o ácido hialurônico é filtrado, e é obtido um filtrado contendo ácido hialurônico e nenhuma célula vegetal. Este filtrado é purificado por diafiltração para remover impurezas de baixo peso molecular. É possível separar o ácido hialurônico pela diafiltração do filtrado contendo ácido hialurônico dissolvido com água pura seguida por descartando continuamente o filtrado. Quando ácido hialurônico é usado em produtos farmacêuticos, pode ser adicionalmente realizada uma etapa de precipitação de ácidos nucleicos da solução. Esta etapa pode ser realizada, por exemplo, adicionando tensoativo de cátion tal como compostos de amônio quaternário de cloreto de cetilpiridínio.

Ácido hialurônico acumulado nas células vegetais transformadas também pode ser purificado por métodos conhecidos para a separação.

#### Uso de Ácido Hialurônico

Ácido hialurônico adquirido pela presente invenção pode ser utilizado de modo útil para composições cosméticas e farmacêuticas, ou biomateriais. Especificamente, ácido hialurônico pode ser usado como uma composição umidificante em cosméticos, um agente terapêutico para artrite, reumatismo crônico, queimaduras e cortes, ou um componente em gotas oculares.

Ácido hialurônico obtido pelo método de produção da presente invenção pode ser usado como um agente ativo para preparar composições cosméticas. Por exemplo, ácido hialurônico pode ser aplicado em formas líquidas tais como soluções aquosas, soluções de óleo, emulsões e suspen-

sões, em formas semi-sólidas tais como géis e cremes, e em formas sólidas tais como pós, grânulos, cápsulas, microcápsulas e sólidos. Ácido hialurônico pode ser preparado nestas formas usando métodos conhecidos, e preparados na formulação de loções, emulsões, géis, cremes, pomadas, emplastos, cataplasmas, aerossóis, supositórios, injeções, pós, grânulos, comprimidos, pílulas, xaropes, trociscos e semelhantes. As formulações referidas podem ser aplicadas por aplicação, por fixação, pulverizando, ingestão e semelhantes. Particularmente entre estas formulações, loções, emulsões, cremes, pomadas, emplastos, cataplasmas, aerossóis e semelhantes são adequados para aplicações na pele. Para cosméticos, ácido hialurônico pode ser usado para cosméticos para tratamento da pele tais como loções, soros, emulsões, cremes e máscaras, cosméticos para maquiagem tais como loções de base de maquiagem, cremes de maquiagem, bases em emulsões, formas de creme e unguento, batons, sombras para os olhos e blushes, cosméticos para tratamento corporal tais como cremes para as mãos, cremes para as pernas, loções para o corpo e semelhantes, essências para banho, cosméticos para tratamento oral e cosméticos para tratamento dos cabelos. Ácido hialurônico pode ser produzido em semelhantes formulações de acordo com o método geral para fabricar cosméticos.

## 20 Exemplos

Os seguintes Exemplos ilustram a presente invenção em detalhes adicionais, mas não se pretende que limitem o âmbito da invenção.

### Exemplo 1 Isolamento do Gene UGD derivado de *Arabidopsis thaliana*

Um gene de UDP-glucose desidrogenase (BT006380:AtUGD1, SEQ ID NO: 15) foi isolado de *Arabidopsis thaliana* e foi demonstrado que tem atividade (Plant J. 2000 21(6):537-46). Adicionalmente, o banco de dados NCBI BLAST (<http://www.ncbi.nlm.nih.gov/BLAST/>) mostra três tipos de genes derivados de *Arabidopsis thaliana* que são previstos para codificar UGD (AF424576:AtUGD2 (SEQ ID NO: 17), AY056200:AtUGD3 (SEQ ID NO: 19), e AY070758:AtUGD4 (SEQ ID NO: 21)). Estes quatro tipos de genes UGD foram clonados para confirmar sua atividade de UGD.

RNA foi extraído de *Arabidopsis thaliana* de acordo com o proto-

colo RNeasy (QIAGEN). Para uma reação de transcrição reversa, 2 µg de RNA total foi dissolvido em 5,5 µL de água estéril, misturado com 1 µL oligo iniciador d(T) 10 µM, e termicamente desnaturado a 70°C por 10 minutos. Depois de rápido esfriamento, a reação de transcrição reversa foi realizada usando um kit ReverTraAce (Toyobo), a 42°C por 30 minutos e a 99°C por 5 minutos.

Para amplificação por PCR dos genes UGD derivados de *Arabidopsis thaliana*, foram designados iniciadores de PCR com base nas quatro sequências de nucleotídeos no banco de dados. Sítios de clivagem de enzima de restrição que são necessários para introdução no vetor de expressão pMAL-c2 (NewEngland Biolab) foram adicionados aos iniciadores. EcoRI ou HindIII (SEQ ID NO: 23 ou 24) foram adicionados a AtUGD1; EcoRI ou PstI (SEQ ID NO: 25 ou 26) a AtUGD2; EcoRI ou PstI (SEQ ID NO: 27 ou 28) a AtUGD3; e EcoRI ou XbaI (SEQ ID NO: 29 ou 30) a AtUGD4. Foi realizada reação de PCR usando uma KOD -plus- DNA polimerase (Toyobo) e um programa de reação de 1 ciclo de 94°C por 2 minutos, 3 ciclos de 94°C por 15 segundos, 45°C por 30 segundos, e 68°C por 1 minuto, e 30 ciclos de 94°C por 15 segundos, 55°C por 30 segundos, e 68°C por 1 minuto. Um microlitro de produto de reação de transcrição reversa foi usado como modelo de DNA. Uma análise por eletroforese por gel agarose apresentou bandas amplificadas por PCR com tamanho previsto. Cada fragmento amplificado por PCR foi clivado com enzimas de restrição (AtUGD1: EcoRI e HindIII; AtUGD2 e AtUGD3: EcoRI e PstI; AtUGD4: EcoRI e XbaI), e clonado em pMAL-c2 digerido com as mesmas enzimas de restrição, usando Ligation High (Toyobo). Usando a mistura da ligação, a cepa JM109 de *Escherichia coli* foi transformada pelo método conhecido mencionado acima, e os transformantes foram aplicados a meio de ágar de Luria-Bertani (LB) (10 g/L de bactotripton, 10 g/L de extrato de levedura, 5 g/L de cloreto de sódio, 15 g/L de ágar) contendo ampicilina (50 mg/mL), e cultivado de um dia para o outro a 37°C. O plasmídeo foi extraído das colônias cultivadas de transformantes usando um método conhecido. As sequências de nucleotídeos dos fragmentos inseridos foram determinadas usando um sequenciador de DNA, e foi

confirmado que os genes amplificados foram aqueles representados pela SEQ ID NOS: 15, 17, 19, e 21. Deste modo, foram construídos os plasmídeos pMAL-c2/AtUGD1, pMAL-c2/AtUGD2, pMAL-c2/AtUGD3, e pMAL-c2/AtUGD4.

5 O acima demonstra que, além de AtUGD1, o qual já foi reportado, três tipos de genes UGD são expressados em *Arabidopsis thaliana*.

As cepas JM109 de *Escherichia coli* carregando os plasmídeos de expressão acima foram cultivadas de um dia para o outro em meio líquido LB contendo ampicilina (50 mg/mL) a 37°C. Meio LB (30 ml) contendo ampicilina (50 mg/mL) e 0,2% de glucose foi inoculado com 300 µL da pré-cultura e cultivado a 37°C por 2 horas. A temperatura foi em seguida reduzida para 25°C, foi adicionado isopropil-b-tiogalactopiranosídeo (IPTG) em uma concentração final de 0,3 mM, e foi induzida expressão das proteínas recombinantes por 24 horas.

15 As células foram recuperadas de 1 µL do meio de cultura por centrifugação e rompidas por desintegração ultrassônica para preparar um extrato bruto. Proteínas de fusão MBP foram purificadas usando MagExtractor-MBP- (Toyobo).

A figura 1 mostra a análise SDS-PAGE das proteínas de fusão MBP expressadas. Uma banda do tamanho previsto foi observada em todas as soluções enzimáticas dos clones.

A atividade de UGD das proteínas de fusão MBP obtidas foi medida de acordo com o método reportado por Tenhaken e Thulke (Plant Physiol. 1996, 112:1127-34). No método, o aumento em NADH causado por reação de UGD é detectado como um aumento em Abs340. Especificamente, 15 µL da solução enzimática (proteína de fusão MBP-UGD) foi adicionado a uma mistura de reação (pH 8.0) contendo UDP-glucose (4 mM), NAD<sup>+</sup> (1 mM), EDTA (1 mM), e Tris-HCl (20 mM); uma reação foi realizada a 37°C; e a absorvência (Abs340) da mistura da reação foi medida com o tempo. A Tabela 1 mostra a atividade enzimática das proteínas de fusão de MBP.

Tabela 1

|  |      |
|--|------|
|  | U/mg |
|--|------|

|          |      |
|----------|------|
| AtUGD1   | 2,15 |
| AtUGD2   | 1,32 |
| AtUGD3   | 1,15 |
| AtUGD4   | 0,25 |
| cvUGD-HI | 0,06 |
| cvUGD0KA | 0,14 |

Os resultados demonstram que todas as soluções enzimáticas de AtUGD1, AtUGD2, AtUGD3, e AtUGD4 têm atividade de UGD, indicando que, em *Arabidopsis thaliana*, os genes AtUGD2, AtUGD3, e AtUGD4, bem como o gene AtUGD1 já reportado, todas codificam UGD.

#### 5 Exemplo 2 Isolamento de Gene UGD Derivado de Chlorella Vírus

Para isolamento do gene UDP-glucose desidrogenase derivado de vírus chlorella (cvUGD) por PCR, os iniciadores de SEQ ID NOS: 31 e 32 foram designados com base na informação de sequência conhecida de vírus chlorella cepa PBCV-1. Para introdução no vetor de expressão pMAL-c2, os sítios EcoRI e PstI foram adicionados aos iniciadores de extremidade 5' e extremidade 3', respectivamente. Foi realizada reação de PCR usando uma KOD -plus- DNA polimerase e um programa de reação de 1 ciclo de 94°C por 2 minutos e 30 ciclos de 94°C por 15 segundos, 60°C por 30 segundos, e 68°C por 1 minuto. O DNA genômico de vírus chlorella cepa Hirosaki (C-VHI1) e cepa Kakunodate (CVKA1) foi usado como modelo de DNA. Uma análise por eletroforese por gel agarose mostrou bandas amplificadas por PCR com tamanho previsto. Cada fragmento amplificado por PCR foi clivado com EcoRI e PstI, e clonado em pMAL-c2 digerido com as mesmas enzimas de restrição, usando Ligation High. Usando a mistura de ligação, a cepa JM109 de *Escherichia coli* foi transformada de acordo com o método conhecido mencionado acima, e os transformantes foram aplicados a meio ágar LB contendo 50 mg/mL de ampicilina, e cultivados de um dia para o outro a 37°C. O plasmídeo foi extraído das colônias cultivadas dos transformantes usando um método conhecido.

As sequências de nucleotídeos dos fragmentos inseridos foram determinadas usando um sequenciador de DNA, e novos genes UGD deri-

vados das cepas CVHI1 e CVKA1 foram obtidos e denominados "gene cvUGD-HI" e "gene cvUGD-KA". As sequências de nucleotídeos dos genes cvGFAT-HI e cvUGD-KA são mostradas nas SEQ ID NOS: 11 e 13. Deste modo, foram construídos os plasmídeos pMAL-c2/cvUGD-HI e pMAL-c2/cvUGD-KA.

Proteínas de fusão de MBP foram expressadas e purificadas na mesma maneira que no Exemplo 1. A figura 2 mostra a análise SDS-PAGE das proteínas de fusão de MBP obtidas. Uma banda do tamanho previsto foi observada em todas as soluções enzimáticas dos clones.

As atividades de UGD das proteínas de fusão de MBP expressadas foram medidas na mesma maneira que no Exemplo 1. A Tabela 1 mostra a atividade enzimática das proteínas de fusão de MBP.

A medição da atividade de UGD pelo método acima demonstrou que tanto cvUGD-HI quanto cvUGD-KA codificam enzimas funcionais.

### 15 Exemplo 3 Isolamento de Gene GFAT Derivado de *Arabidopsis thaliana*

Como um gene GFAT derivado de plantas, foi reportado que um gene GFAT derivado de milho (No. de Acesso AY106905) é isolado (publicação de patente internacional N° WO 00/11192), mas não foram isolados genes GFAT de outras espécies de plantas. O banco de dados NCBI BLAST (<http://www.ncbi.nlm.nih.gov/BLAST/>) mostra um gene GFAT derivado de *Arabidopsis thaliana* (No. de Acesso NM\_113314), o qual é altamente homóloga ao gene conhecido. A sequência de nucleotídeos do gene GFAT derivado de *Arabidopsis thaliana* é mostrada na SEQ ID NO: 9 na Listagem de Sequência.

RNA foi extraído de *Arabidopsis thaliana* de acordo com o protocolo RNeasy (QIAGEN). Para uma reação de transcrição reversa, 2 µg de RNA total foi dissolvido em 5,5 µL de água estéril, misturado com 1 µL de oligo d(T) iniciador a 10 µM, e termicamente desnaturado a 70°C por 10 minutos. Depois de rápido esfriamento, a reação de transcrição reversa foi realizada usando um kit ReverTraAce (Toyobo), a 42°C por 30 minutos e a 99°C por 5 minutos.

Para amplificação por PCR do gene GFAT derivado de *Arabi-*

*dopsis thaliana*, os iniciadores de SEQ ID NOS: 33 e 34 foram designados com base nas sequências de nucleotídeos no banco de dados. Para introdução em um vetor de expressão pEU-NII livre de células (Toyobo), um sítio Sall foi adicionado ao iniciador de extremidade 3'. Foi realizada reação de

5 PCR usando uma KOD -plus- DNA polimerase e um programa de reação de 1 ciclo de 94°C por 2 minutos, 3 ciclos de 94°C por 15 segundos, 45°C por 30 segundos, e 68°C por 1 minuto, e 30 ciclos de 94°C por 15 segundos, 55°C por 30 segundos, e 68°C por 1 minuto. Um microlitro de produto de reação de transcrição reversa foi usado como modelo de DNA. Uma análise

10 por eletroforese por gel agarose mostrou bandas amplificadas por PCR com tamanho previsto. O fragmento amplificado por PCR foi clivado com uma enzima de restrição Sall, e clonado em pEU-NII digerido com enzimas de restrição EcoRV e Sall, usando Ligation High. Usando a mistura de ligação, a cepa JM109 de *Escherichia coli* foi transformada de acordo com o método

15 conhecido mencionado acima, e os transformantes foram aplicados ao meio de ágar LB contendo 50 mg/mL de ampicilina, e cultivados de um dia para o outro a 37°C. O plasmídeo foi extraído das colônias cultivadas dos transformantes usando um método conhecido. Deste modo, foi construído o plasmídeo pEU-NII/AtGFAT.

20 A sequência de nucleotídeos do fragmento inserido foi determinada usando um sequenciador de DNA, e foi confirmado que o fragmento amplificado foi o gene representado pela SEQ ID NO: 9. O acima revela que o gene GFAT é expressado em *Arabidopsis thaliana*.

Usando PROTEIOS (uma marca registrada de Toyobo), foi expressada a proteína recombinante do AtGFAT. Especificamente, foi realizada uma reação a 37°C por 4 horas usando 5 µg do plasmídeo pEU/AtGFAT como um modelo e uma T7 RNA Polimerase, para sintetizar RNAm. Depois disso, 6 µg de RNAm foi misturado com extrato de germe de trigo, e foi realizada uma reação a 26°C por 24 horas pelo método de bicamada. A mistura

25 da reação foi suspendida em tampão da amostra (Tris-HCl a 50 mM (pH 6,8), 2% de SDS, 10% de glicerol, 0,6% de b-mercaptoetanol) e fervida por 5 minutos, e em seguida a proteína expressada foi analisada por SDS-PAGE.

30

Foi observada uma banda do tamanho previsto (cerca de 75 kDa), indicando a expressão da proteína AtGFAT (figura 3).

A solução de expressão de proteína foi submetida a reação de GFAT. Especificamente, 50 µL da solução de expressão de proteína foi adicionado a uma solução (pH 7,0) contendo frutose-6-fosfato (15 mM), L-glutamina (15 mM), EDTA (1 mM), DTT (1 mM), e KH<sub>2</sub>PO<sub>4</sub> (60 mM), e reagida a 37°C por 4 horas. Glucosamina-6-fosfato na misura da reação foi medida usando o método de Reissig (J. Biol. Chem, 1955, 217 (2), 959-66), o qual é um aprimoramento do método de Morgan e Elson, para avaliar a atividade de GFAT. A figura 4 mostra os resultados. Glucosamina-6-fosfato não foi detectado em pEU-NII/DHFR usado como um controle, ao passo que um aumento de glucosamina-6-fosfato foi observado em pEU-NII/AtGFAT, demonstrando que uma enzima GFAT ativa estava presente na solução de expressão de proteína.

O acima revela que o gene AtGFAT codifica uma enzima funcional em *Arabidopsis thaliana*.

#### Exemplo 4 Isolamento de Gene GFAT Derivado de Chlorella Vírus

Primers de PCR foram preparados para isolar o gene de glutamina-frutose-6-fosfato amidotransferase derivado de vírus chlorella (cvGFAT) por PCR. Com base em informação de sequência de nucleotídeos já reportada da cepa de vírus chlorella PBCV-1, os iniciadores de SEQ ID NOS: 35 e 36 foram designados de modo a amplificar a partir de 100 bp fora da região de GFAT putativa. Foi realizado PCR usando uma KOD -plus- DNA polimerase e um programa de reação de 1 ciclo de 94°C por 2 minutos e 30 ciclos de 94°C por 15 segundos, 60°C por 30 segundos, é 68°C por 1 minuto. O DNA genômico de vírus chlorella cepas Hirosaki (CVHI1) e Kakunodate (CVKA1) foi usado como modelo de DNA. Os fragmentos amplificados por PCR foram clonados em um vetor de clonagem pPCR-Script Amp SK(+) (Stratagene). As sequências de nucleotídeos dos fragmentos inseridos foram determinadas usando um sequenciador de DNA, e novos genes GFAT derivados das cepas cvHI1 e cvKA1 foram identificados e denominados "gene cvGFAT-HI" e "gene cvGFAT-KA". As sequências de nucleotídeos dos

genes cvGFAT-HI e cvGFAT-KA são mostradas nas SEQ ID NOS: 5 e 7.

As regiões do quando de leitura aberta dos genes cvGFAT-HI e cvGFAT-KA representadas pelas SEQ ID NOS: 5 e 7 foram clonadas em vetores para síntese de proteína livre de células para expressar proteínas.

5 Foi realizada reação de PCR sob as condições mencionadas acima usando vetor de clonagem pPCR-Script Amp SK(+) contendo o gene cvGFAT-HI como um modelo e os iniciadores representados pelas SEQ ID NOS: 37 e 38. Foi realizada reação de PCR sob as condições mencionadas acima usando vetor de clonagem pPCR-Script Amp SK(+) contendo o gene  
10 cvGFAT-KA como um modelo e os iniciadores representados pelas SEQ ID NOS: 37 e 39. Para o PCR, uma KOD -plus- DNA polimerase, e um ciclo de reação de 1 ciclo de 94°C por 2 minutos e 30 ciclos de 94°C por 15 segundos, 60°C por 30 segundos, e 68°C por 1 minuto, foram usados. Cada fragmento amplificado por PCR foi clivado com XbaI e clonado nos sítios EcoRV  
15 e XbaI do vetor pEU-NII. Usando a mistura de ligação, a cepa de *Escherichia coli* DH5 foi transformada de acordo com o método conhecido mencionado acima, e os transformantes foram aplicados ao meio de ágar LB contendo 50 mg/mL de ampicilina, e cultivados de um dia para o outro a 37°C. O plasmídeo foi extraído das colônias cultivadas dos transformantes usando  
20 um método conhecido. Deste modo, foram construídos os plasmídeos pEU-NII/cvGFAT-HI e pEU-NII/cvGFAT-KA.

Usando pEU-NII/cvGFAT-HI e pEU-NII/cvGFAT-KA e PROTEIOS (marca registrada), foram expressadas proteínas na mesma maneira que no Exemplo 3. Análise SDS-PAGE das proteínas expressadas confirmaram  
25 bandas dos tamanhos previstos (cerca de 65 kDa), indicando a expressão de proteínas cvGFAT (figura 3).

A solução de expressão de proteína foi submetida a reação de GFAT na mesma maneira que no Exemplo 3. A figura 4 mostra os resultados. Glucosamina-6-fosfato não foi detectado em pEU/DHFR usado como um  
30 controle, ao passo que foi observado um aumento em glucosamina-6-fosfato em pEU/cvGFAT-HI, confirmando que uma enzima GFAT ativa estava presente na solução de expressão de proteína.

#### Exemplo 5 Clonagem de Gene HAS Derivado de Chlorella Vírus em pBI121

Para a clonagem de um gene de ácido hialurônico sintetase derivado de vírus chlorella (cvHAS, SEQ ID NO: 1) no vetor de transformação de planta (nas partes que se seguem também referido como "vetor de expressão") pBI121 (Jefferson et al., 1987, EMBO J, 6, 3901-3907), foram preparados os iniciadores representados pelas SEQ ID NOS: 40 e 41. Foi realizada reação de PCR usando DNA de plasmídeo contendo cvHAS como um modelo, uma KOD -plus-DNA polimerase, e um programa de reação de 1 ciclo de 94°C por 2 minutos, 3 ciclos de 94°C por 15 segundos, 45°C por 30 segundos, e 68°C por 1 minuto, e 30 ciclos de 94°C por 15 segundos, 55°C por 30 segundos, e 68°C por 1 minuto. Fragmento amplificado por PCR foi purificado e clivado com BamHI e DraI. Subsequentemente, cvHAS foi inserido no vetor de expressão pBI121 como se segue: pBI121 foi digerido com a enzima de restrição SacI, neutralizado (blunted) com Blunting High (Toyobo), e digerido com BamHI; e em seguida foi clonado o gene cvHAS digerido com uma enzima de restrição. Usando a mistura de ligação, a cepa de *Escherichia coli* DH5 foi transformada de acordo com o método conhecido mencionado acima, e os transformantes foram aplicados a meio de ágar LB contendo 50 mg/mL de ampicilina, e cultivados de um dia para o outro a 37°C. O plasmídeo foi extraído das colônias cultivadas dos transformantes usando um método conhecido. Deste modo, foi preparado pBI121/cvHAS (nas partes que se seguem referido algumas vezes como pBIHA) contendo cvHAS.

#### Exemplo 6 Clonagem de Gene GFAT Derivado de Chlorella Vírus em pBI121

Foi realizada reação de PCR usando os iniciadores representados pelas SEQ ID NOS: 42 e 38 e DNA de plasmídeo contendo cvGFAT-HI como um modelo. Para o PCR, KOD -plus- e um programa de reação de 1 ciclo de 94°C por 2 minutos, 2 ciclos de 94°C por 15 segundos, 50°C por 30 segundos, e 68°C por 1 minuto, e 28 ciclos de 94°C por 15 segundos, 60°C por 30 segundos, e 68°C por 1 minuto, foram usados. O fragmento amplificado por PCR foi digerido com BamHI. Subsequentemente, cvGFAT-HI foi inserido no vetor de expressão ppBI121 como se segue: pBI121 foi digerido com SacI, neutralizado com Blunting High, e digerido com BamHI, e em se-

guida o gene cvGFAT-HI digerido com as enzimas de restrição acima foi clonado. Usando a mistura de ligação, a cepa de *Escherichia coli* DH5 foi transformada de acordo com o método conhecido mencionado acima, e os transformantes foram aplicados ao meio de ágar LB contendo 50 mg/mL de ampicilina e cultivados de um dia para o outro a 37°C. O plasmídeo foi extraído das colônias cultivadas dos transformantes usando um método conhecido. Deste modo, foi preparado pBI121/cvGFAT-HI (nas partes que se seguem referido algumas vezes como pBIGF) contendo cvGFAT-HI.

#### Exemplo 7 Subclonagem do Gene GFAT Derivado de Chlorella Vírus e Gene HAS em pBluescript

pBIHA foi digerido com as enzimas de restrição PvuII e PstI, nesta ordem, e subclonado nos sítios PstI e SmaI de pBluescript para obter pBluescript/35S-cvHAS-NOS (nas partes que se seguem referido algumas vezes como pBSHA).

pBIGF foi digerido com as enzimas de restrição PvuII e SphI, nesta ordem. Depois de neutralização com Blunting High (Toyobo), pBIGF foi subclonado no sítio EcoRV de pBluescript para obter pBluescript/35S-cvGFAT-NOS (nas partes que se seguem referido algumas vezes como pBSGF).

pBSGF foi digerido com as enzimas de restrição KpnI e NotI, nesta ordem. Depois de neutralização com Blunting High (Toyobo), 35S-cvGFAT-NOS foi subclonado em pBSHA digerido previamente com SpeI, neutralização, e defosforilação, nesta ordem, para obter pBluescript/cvHAS-cvGFAT (nas partes que se seguem referido algumas vezes como pBSHG).

#### Exemplo 8 Clonagem de Gene GFAT Derivado de Chlorella Vírus e Gene HAS em pBI121

pBSHG foi digerido com as enzimas de restrição KpnI e NotI nesta ordem, e o cassete de expressão de cvHAS-cvGFAT foi clivado e neutralizado. Em seguida, cvHAS-cvGFAT foi inserido no vetor de expressão pBI121 como se segue: pBI121 foi digerido com as enzimas de restrição SacI e HindIII, e neutralizado com Blunting High, e em seguida o gene cvHAS-cvGFAT digerido com as enzimas de restrição acima foi clonado. Usando a

mistura de ligação, a cepa DH5 de *Escherichia coli* foi transformada de acordo com o método conhecido mencionado acima, e os transformantes foram aplicados a meio de ágar LB contendo 50 mg/mL de ampicilina, e cultivados de um dia para o outro a 37°C. O plasmídeo foi extraído das colônias cultivadas dos transformantes usando um método conhecido. Deste modo, foi preparado pBI121/cvHAS-cvGFAT (nas partes que se seguem referido algumas vezes como pBIHG) contendo cvHAS-cvGFAT.

#### Exemplo 9 Preparação de Células eletrocompetentes

Cinco mililitros de meio LB foi inoculado com uma única colônia de *Agrobacterium* LBA4404 (Cepa LBA4404 de *Agrobacterium tumefaciens*), e submetido a cultura em agitação de um dia para o outro a 28°C. O meio de cultura foi inoculado com 500 µL de meio LB e submetido a cultura em agitação a 28°C até a turvação a 600 nm se tornar 0,5. O meio de cultura foi colhido por centrifugação (5000 rpm, 10 min, 4°C); o sobrenadante foi removido; 500 mL de água estéril foi adicionado para suspender e lavar as células; foi realizada centrifugação (5000 rpm, 10 min, 4°C) novamente para colher as células; e o sobrenadante foi removido. Depois de realizar o procedimento acima duas vezes, os precipitados foram suspensos em 20 µL de solução resfriada de glicerol a 10%, as células foram colhidas por centrifugação (5000 rpm, 10 min, 4°C), e o sobrenadante foi removido. Os precipitados foram suspensos em 3 mL de solução resfriada de glicerol a 10%, e alíquotas de 40 µL da suspensão foram colocadas dentro de tubos de centrifuga de 1,5 mL, congeladas com nitrogênio líquido, e armazenadas a -80°C.

#### Exemplo 10 Introdução de pBIHG em Cepa LBA4404 de *Agrobacterium*

Uma suspensão obtida misturando 1 µL do plasmídeo de expressão pBIHG (200 mg/ml) com 40 µL de células eletrocompetentes de *A. tumefaciens* LBA4404 (Invitrogen) foi vertida em uma tina (distância entre os eletrodos: 1 mm) previamente esfriada sobre gelo, e foi aplicado um campo elétrico pulsátil (1,8 kV, 25 µF, 200 W). Imediatamente depois disso, 500 µL de SOC foi adicionado, e a mistura resultante foi incubada a 28°C por 3 horas. As células incubadas foram em seguida aplicadas a meio de placa LB contendo canamicina, e cultivadas a 25°C por 3 dias para obter células de

*Agrobacterium* carregando pBIHG.

Exemplo 11 Infecção de Tabaco com *Agrobacterium tumefaciens* Cepa LBA4404 contendo pBIHG

Tabaco (*Nicotiana tabacum* SR-1) foi transformado de acordo com o método de disco de folhas usando *Agrobacterium* ("Plant Biotechnology II" editado por Yasuyuki Yamada e Yoshimi Okada, Tokyo Kagaku Dojin, 1991). Discos de folhas de tabaco foram imersos por 3 minutos em um meio de cultura de *Agrobacterium* carregando pBIHG ou pBIHA previamente cultivado de um dia para o outro a 28°C em 5 µL de meio LB contendo 50 mg/L de canamicina. As células em excesso foram em seguida removidas sobre papel-filtro. Os discos de folhas foram colocados em um meio de diferenciação preparado adicionando 3% de sucrose, vitamina B5, 1 mg/L de benzilaminopurina, 1 mg/L de ácido acético de naftaleno, e 0,3% de goma gelano a sal inorgânico MS (Murashige and Skoog) (Murashige and Skoog, 1962, *Physiol. Plant.*, 15, 473) e ajustando o pH para 5,7, e foram deixadas em repouso no escuro a 28°C por 2 dias. Os discos de folhas infectados foram lavados três vezes com água estéril, e o excesso de umidade foi removido sobre papel-filtro. Os discos de folhas foram em seguida deixados em repouso no meio de diferenciação contendo canamicina (100 mg/L) e cefotaxima (250 mg/L) como antibióticos, e foi induzida a formação de calo a 25°C sob condições de 16 horas de luz. Três semanas depois de iniciar a indução, brotos morfológicamente normais foram selecionadas, cortadas em uma forma contendo caules e folhas, e transferidas para um meio de enraizamento (sal inorgânico MS, 3% de sucrose, vitamina B5 e 0,3% de goma gelano, pH 5,7) contendo canamicina (100 mg/L) e cefotaxima (250 mg/L) para induzir enraizamento a 25°C sob condições de 16 horas de luz. Depois de duas semanas, brotos enraizados foram transferidos para um meio de cultura de enraizamento recente para obter linhagens com caules e folhas em crescimento.

Exemplo 12 Quantificação de Ácido Hialurônico Produzido por Tabaco Transformado

Cerca de 100 mg de folhas de tabaco transformado obtidas pela

infecção com *Agrobacterium* descrita acima foi transferido para um tubo de 2 mL, e suspenso em 200 µL de tampão (contendo Tris-HCl a 20 mM a pH 7,5, NaCl a 0,2 M, EDTA a 1 mM, e 2-ME a 10 mM), e 400 mg de contas de aço inoxidável (diâmetro: 4,8 mm) foram adicionados. As folhas de tabaco foram pulverizadas sacudindo e agitando o tubo usando Bead Smash (Wakenyaku, BS-12) (4.000 rpm, 2 minutos). O líquido depois da pulverização foi centrifugado (15.000 rpm, 10 minutos), e o sobrenadante foi recuperado como um extrato bruto. O extrato bruto foi diluído com água e usado como uma amostra de medição. A quantificação do ácido hialurônico foi realizada usando uma lâmina de ácido hialurônico "Chugai" (Fujirebio, Inc.). A figura 5 mostra os resultados. O tabaco transformado no qual o gene cvHAS-cvGFAT tinha sido introduzido teve produtividade de ácido hialurônico significativamente aumentada comparado com o tabaco transformado no qual o gene cvHAS tinha sido introduzido.

#### 15 Exemplo 13 Clonagem de Gene UGD Derivado de Chlorella Vírus em pBI121

Foi realizada reação de PCR usando os iniciadores representados pela SEQ ID NO: 43 e SEQ ID NO: 44 e DNA de plasmídeo contendo cvUGD-HI como um modelo. Para o PCR, KOD -plus-, e um programa de reação de 1 ciclo de 94°C por 2 minutos, 2 ciclos de 94°C por 15 segundos, 50°C por 30 segundos, e 68°C por 1 minuto), e 28 ciclos de 94°C por 15 segundos, 60°C por 30 segundos, e 68°C por 1 minuto, foram usados. O fragmento amplificado por PCR foi digerido com BamHI e SacI. Subsequentemente, cvUGD-HI foi inserido no vetor de expressão pBI121 como se segue: pBI121 foi digerido com SacI e em seguida com BamHI, e o gene cvUGD-HI digerido com as enzimas de restrição mencionadas acima foi clonado em pBI121. Usando a mistura de ligação, a cepa de *Escherichia coli* DH5 foi transformada de acordo com o método conhecido mencionado acima, e o transformante foi aplicado ao meio de ágar LB contendo 50 mg/ml de ampicilina, e cultivado de um dia para o outro a 37°C. O plasmídeo foi extraído das colônias cultivadas dos transformantes usando um método conhecido. Deste modo, foi preparado cvUGD-HI contendo pBIUG.

#### 30 Exemplo 14 Subclonagem de Gene UGD Derivado de Chlorella Vírus em

pBluescript

pBIUG foi digerido com as enzimas de restrição EcoRI e HindIII, subclonado nos sítios EcoRI e HindIII de pBluescript para preparar pBSUG.

pBSUG foi digerido com SpeI e neutralizado com Blunting High para destruir o sítio SpeI. Subsequentemente, pBSHA foi digerido com NotI, neutralizado, e digerido com KpnI para cortar 35S-cvHAS-NOS. pBSUG foi digerido com XhoI, neutralizado, e digerido com KpnI; e 35S-cvHAS-NOS, o qual tinha sido cortado, foi ligado para produzir pBSHU. pBSGF foi digerido com Sall, neutralizado, e digerido com NotI para cortar 35S-cvGFAT-NOS. pBHU foi digerido com SpeI, neutralizado, e digerido com NotI, e 35S-cvGFAT-NOS, o qual tinha sido cortado, foi ligado para produzir pBSHUG.

Exemplo 15 Clonagem de Gene UGD Derivado de Chlorella Vírus, Gene GFAT, e Gene HAS, em pBI121

Usando DNA sintético, foi produzido pBI121 modificado no qual um sítio SmaI tinha sido adicionado a jusante do sítio HindIII de pBI121, e um sítio KpnI tinha sido adicionado a montante do sítio EcoRI. pBSHUG foi digerido com NotI, neutralizado, e digerido com KpnI; os cassetes de expressão de HAS, UGD e GFAT encadeados foram cortados e clonados no pBI121 modificado clivado com SmaI e KpnI. Usando a mistura de ligação, a cepa de *Escherichia coli* DH5 foi transformada de acordo com o método conhecido mencionado acima, e os transformantes foram aplicados ao meio de ágar LB contendo 50 µg/ml de ampicilina, e cultivados de um dia para o outro a 37°C. O plasmídeo foi extraído das colônias cultivadas dos transformantes usando um método conhecido. Deste modo, foram preparados cassetes de expressão de HAS, UGD e GFAT contendo pBIHUG.

Exemplo 16 Preparação de Tabaco Transformado, no qual pBIHUG tinha sido Introduzido, e Análise Quantitativa de Ácido Hialurônico

Usando o plasmídeo de expressão pBIHUG e seguindo os procedimentos mostrados nos Exemplos 10 e 11, pBIHUG foi introduzido em cepa LBA4404 de *Agrobacterium*, e discos de folhas de tabaco foram infectados com cepa LBA4404 de *Agrobacterium* contendo pBIHUG. Em consequência, foram obtidas 20 linhagens de tabaco nas quais a introdução dos

genes HAS, UGD, e GFAT tinha sido confirmada. Depois do procedimento mostrado no Exemplo 12, foram preparados extratos brutos e o ácido hialurônico foi quantificado. A figura 6 mostra os resultados. O tabaco transformado no qual os genes HAS, UGD e GFAT tinham sido introduzidos apresentou

5 produtividade de ácido hialurônico significativamente aumentada comparado com o tabaco transformado no qual o gene cvHAS tinha sido introduzido e o tabaco transformado no qual o gene cvHAS-cvGFAT tinha sido introduzido.

Exemplo 17 Análise Quantitativa de Ácido Hialurônico Produzido por Tabaco Transformado (geração T1)

10 Foram colhidas sementes do tabaco transformado do Exemplo 12, no qual o gene cvHAS-cvGFAT tinha sido introduzido, e inoculado em meio de diferenciação MS contendo canamicina (100 mg/L). Um extrato bruto foi obtido do tabaco transformado cultivado (geração T1) na mesma maneira que no Exemplo 12, e o ácido hialurônico foi quantificado, demonstrando o nível de produção de ácido hialurônico equivalente ao da geração T0.

15

APLICABILIDADE INDUSTRIAL

De acordo com a presente invenção, um gene de ácido hialurônico sintetase é expressado em uma planta, e em particular, ácido hialurônico é produzido em um alto rendimento em uma planta. A presente invenção

20 proporciona uma planta ou células vegetais cultivadas que são capazes de produzir ácido hialurônico em um maior rendimento do que a técnica anterior; um método para produzir a planta ou células vegetais cultivadas; e um vetor de expressão. Como ácido hialurônico seguro produzido em uma planta pode ser proporcionado em um baixo custo, espera-se que a presente

25 invenção contribua muito para a indústria.

LISTAGEM DE SEQUÊNCIA

<110> TOYO BOSEKI KABUSHIKI KAISHA

<120> Planta Produtora de Ácido Hialurônico

<130> 060044PC1

30 <160> 44

<170> PatentIn version 3.1

<210> 1

&lt;211&gt; 1707

&lt;212&gt; DNA

&lt;213&gt; Virus Chlorella

&lt;400&gt; 1

5 atgggtaaaa atataatcat aatgggtttcg tggtagacca tcataacttc aaatctaadc 60  
 gcggttgag gagcctctct aatcttggct ccagcaatta ctgggtatat tctacattgg 120  
 aatattgctc tctcgacaat ctggggagta tcagcttatg gtattttcgt ttttggtttt 180  
 ttccttgcac aagttttatt ttcagaactg aacaggaaac gtcttcgcaa gtggatttct 240  
 ctccagacta aggggttgaa tgatgtccgt ttggctgtga tcattgctgg ataccgcgaa 300  
 10 gatccctata tgttccaaaa gtgtctcgag tcagtgcgtg actctgacta cggtaacggt 360  
 gctcgtctca tttgtgttat tgacggcgat gacgacgctg atatgaagat gtccgatgtt 420  
 tacaagacga tctacaacga taatatcaag aagcccgagt ttgtcttggt tgagtcagac 480  
 gacaaggaag gtgaacgcat cgactctgat ttctctcgcg acatttgtgt tctccagcct 540  
 caccgtggca agagggagtg tctctatact gggttccaac ttgcaaagat ggaccccagt 600  
 15 gtcaacgccg tcgttttgat tgacagcgat actgttctcg agaaggatgc tattctggaa 660  
 gttgtatacc cacttgcatg cgatcctgag atccaagccg tcgcaggtga gtgtaagatt 720  
 tggaacacag acactctttt gagtcttctc gtcgcttggc ggtactatc tgcggtttgt 780  
 gtggagagga gtgcccagtc ttttttcagg actgttcagt gcgttggggg cccgctgggt 840  
 gcctacaaga ttgatatcat taaggagatt aaggaccctt ggatttccca gcgctttctt 900  
 20 ggtcagaagt gtacttacgg tgacgaccgc cggctaacca acgagatctt gatgcgtggt 960  
 aaaaagggtt tgttcaactc atttgctggt ggttggctg acagtcgcac caatgtgttt 1020  
 cgatacatcg ttcagcagac ccgctggagt aagtcgtggt gccgcgaaat ttggtacacc 1080  
 ctcttcgccc cgtggaagca cggtttgtct ggaatttggc tggcctttga atgtttgtat 1140  
 caaattacat acttcttctt cgtgatttac ctcttttctc gcctagccgt tgaggccgac 1200  
 25 cctcgtccc agacagccac agtgattgtg agcaccacgg ttgcattgat taagtgtggg 1260  
 tatttttcat tccgagccaa ggatattcgg gctttttact ttgtgcttta tacatttggt 1320  
 tactttttct gtatgattcc ggccagggtt actgcaatga tgacgctttg ggacattggc 1380  
 tggggtactc gcggtggaaa cgagaagcct tccgttggca cccgggtcgc tctgtgggca 1440  
 aagcaatata tcattgcata tatgtggtgg gccgcggttg ttggcgctgg agtttacagc 1500  
 30 atcgtccata actggatgtt cgattggaat tctctttctt atcgttttgc tttggttggg 1560  
 atttgttctt acattgtttt tattactatt gtgctggtga tttatttcac cggcaaaatt 1620  
 acgacttga atttcacgaa gcttcagaag gagctaactg aggatcgtgt tctgtacgat 1680

gcattctacca atgctcagtc tgtgtga

1707

<210> 2

<211> 568

<212> PRT

5 <213> Virus Chlorella

<400> 2

Met Gly Lys Asn Ile Ile Ile Met Val Ser Trp Tyr Thr Ile Ile Thr

1 5 10 15

Ser Asn Leu Ile Ala Val Gly Gly Ala Ser Leu Ile Leu Ala Pro Ala

10 20 25 30

Ile Thr Gly Tyr Ile Leu His Trp Asn Ile Ala Leu Ser Thr Ile Trp

35 40 45

Gly Val Ser Ala Tyr Gly Ile Phe Val Phe Gly Phe Phe Leu Ala Gln

50 55 60

15 Val Leu Phe Ser Glu Leu Asn Arg Lys Arg Leu Arg Lys Trp Ile Ser

65 70 75 80

Leu Arg Pro Lys Gly Trp Asn Asp Val Arg Leu Ala Val Ile Ile Ala

85 90 95

Gly Tyr Arg Glu Asp Pro Tyr Met Phe Gln Lys Cys Leu Glu Ser Val

20 100 105 110

Arg Asp Ser Asp Tyr Gly Asn Val Ala Arg Leu Ile Cys Val Ile Asp

115 120 125

Gly Asp Asp Asp Ala Asp Met Lys Met Ser Asp Val Tyr Lys Thr Ile

130 135 140

25 Tyr Asn Asp Asn Ile Lys Lys Pro Glu Phe Val Leu Cys Glu Ser Asp

145 150 155 160

Asp Lys Glu Gly Glu Arg Ile Asp Ser Asp Phe Ser Arg Asp Ile Cys

165 170 175

Val Leu Gln Pro His Arg Gly Lys Arg Glu Cys Leu Tyr Thr Gly Phe

30 180 185 190

Gln Leu Ala Lys Met Asp Pro Ser Val Asn Ala Val Val Leu Ile Asp

195 200 205

Ser Asp Thr Val Leu Glu Lys Asp Ala Ile Leu Glu Val Val Tyr Pro  
 210 215 220  
 Leu Ala Cys Asp Pro Glu Ile Gln Ala Val Ala Gly Glu Cys Lys Ile  
 225 230 235 240  
 5 Trp Asn Thr Asp Thr Leu Leu Ser Leu Leu Val Ala Trp Arg Tyr Tyr  
 245 250 255  
 Ser Ala Phe Cys Val Glu Arg Ser Ala Gln Ser Phe Phe Arg Thr Val  
 260 265 270  
 Gln Cys Val Gly Gly Pro Leu Gly Ala Tyr Lys Ile Asp Ile Ile Lys  
 10 275 280 285  
 Glu Ile Lys Asp Pro Trp Ile Ser Gln Arg Phe Leu Gly Gln Lys Cys  
 290 295 300  
 Thr Tyr Gly Asp Asp Arg Arg Leu Thr Asn Glu Ile Leu Met Arg Gly  
 305 310 315 320  
 15 Lys Lys Val Val Phe Thr Pro Phe Ala Val Gly Trp Ser Asp Ser Pro  
 325 330 335  
 Thr Asn Val Phe Arg Tyr Ile Val Gln Gln Thr Arg Trp Ser Lys Ser  
 340 345 350  
 Trp Cys Arg Glu Ile Trp Tyr Thr Leu Phe Ala Ala Trp Lys His Gly  
 20 355 360 365  
 Leu Ser Gly Ile Trp Leu Ala Phe Glu Cys Leu Tyr Gln Ile Thr Tyr  
 370 375 380  
 Phe Phe Leu Val Ile Tyr Leu Phe Ser Arg Leu Ala Val Glu Ala Asp  
 385 390 395 400  
 25 Pro Arg Ser Gln Thr Ala Thr Val Ile Val Ser Thr Thr Val Ala Leu  
 405 410 415  
 Ile Lys Cys Gly Tyr Phe Ser Phe Arg Ala Lys Asp Ile Arg Ala Phe  
 420 425 430  
 Tyr Phe Val Leu Tyr Thr Phe Val Tyr Phe Phe Cys Met Ile Pro Ala  
 30 435 440 445  
 Arg Val Thr Ala Met Met Thr Leu Trp Asp Ile Gly Trp Gly Thr Arg  
 450 455 460

Gly Gly Asn Glu Lys Pro Ser Val Gly Thr Arg Val Ala Leu Trp Ala

465 470 475 480

Lys Gln Tyr Leu Ile Ala Tyr Met Trp Trp Ala Ala Val Val Gly Ala

485 490 495

5 Gly Val Tyr Ser Ile Val His Asn Trp Met Phe Asp Trp Asn Ser Leu

500 505 510

Ser Tyr Arg Phe Ala Leu Val Gly Ile Cys Ser Tyr Ile Val Phe Ile

515 520 525

Thr Ile Val Leu Val Ile Tyr Phe Thr Gly Lys Ile Thr Thr Trp Asn

10 530 535 540

Phe Thr Lys Leu Gln Lys Glu Leu Ile Glu Asp Arg Val Leu Tyr Asp

545 550 555 560

Ala Ser Thr Asn Ala Gln Ser Val

565

15 <210> 3

<211> 1707

<212> DNA

<213> Virus Chlorella

<400> 3

20 atgggtaaaa atataatcat aatgggttcg tggtacacca tcataacttc aaatctaadc 60

gcggttggag gagcctctct aatcttggct ccagcaatta ctggatatat tctacattgg 120

aatattgctc tctcgacaat ctggggagta tcagcttatg gtattttcgt ttttggtttt 180

ttccttgcaac aagttttatt ttcagaactg aacaggaaac gtcttcgcaa atggatttct 240

ctcagacctc agggttggaa tgatgtccgt ttggctgtga tcattgctgg ataccgcaaa 300

25 gatccctata tgttccaaaa gtgtctcgag tcagtgcgtg actctgacta cggtaacgtt 360

gctcgtctca tttgtgttat tgacggcgat gacgacgtg atatgaagat gtccgatgtt 420

tacaagacga tctacaacga taatatcaag aagcccgagt ttgtcttggt tgagtcagac 480

gacaaggaag gtgaacgcac cgactctgat ttctctcgcg acatttgtgt tctccagcct 540

caccgtggca agaggagtg tctctatact ggtttccaac ttgcaaagat ggaccccgat 600

30 gtcaacgccg tcgttttgat tgacagcgat actgttctcg agaaggatgc tattctggaa 660

gttgatatacc cacttgcatt cgatcctgag atccaagccg tcgcaggtga gtgtaagatt 720

tggaacacag acactctttt gactcttctc gtcgcttggc ggtactattc tgcgttttgt 780

gtggagagga gtgcccagtc ttttttcagg actgttcagt gcgttggggg cccgctgggt 840  
 gcctacaaga ttgatatcat taaggagatt aaggaccctt ggatttccca gcgctttctt 900  
 ggtcagaagt gtacttacgg tgacgaccgc cggctaacca acgagatctt gatgcgtggt 960  
 aaaaagggtt tggttcactcc atttgcgtt gggttggtctg acagtccgac caatgtgttt 1020  
 5 cgatacatcg ttcagcagac ccgctggagt aagtcgtggt gccgcgaaat ttggtacacc 1080  
 ctctttgccg cgtggaagca cggtttgtct ggaatttggc tggcctttga atgtttgtat 1140  
 caaattacat acttcttctt cgtgatttac ctcttttctc gcctagccgt tgaagccgac 1200  
 cctcgctccc agacagccac agtgattgtg agcaccacgg ttgcattgat taagtgtggg 1260  
 tatttttcat tccgagccaa ggatattcgg gctttttact ttgtgcttta tacatttggt 1320  
 10 tactttttct gtatgattcc ggccaggggt actgcaatga tgacgctttg ggacattggc 1380  
 tggggtactc gcggtggaaa cgagaagcct tccgttgga cccgggtcgc tctgtgggca 1440  
 aagcaatata tcattgcata tatgtggtgg gccgcggtt ttggcgctgg agtttacagc 1500  
 atcgccata actggatgtt cgattggaat tctctttctt atcgttttgc tttggttggt 1560  
 atttgttctt acattgtttt tattactatt gtgctggtga tttatttcac cggcaaaatt 1620  
 15 acgacttgga atttcacgaa gcttcagaag gagctaactg aggatcgtgt tctgtacgat 1680  
 gcatctacca atgctcagtc tgtgtga 1707  
 <210> 4  
 <211> 568  
 <212> PRT  
 20 <213> Virus Chlorella  
 <400> 4  
 Met Gly Lys Asn Ile Ile Ile Met Val Ser Trp Tyr Thr Ile Ile Thr  
 1 5 10 15  
 Ser Asn Leu Ile Ala Val Gly Gly Ala Ser Leu Ile Leu Ala Pro Ala  
 25 20 25 30  
 Ile Thr Gly Tyr Ile Leu His Trp Asn Ile Ala Leu Ser Thr Ile Trp  
 35 40 45  
 Gly Val Ser Ala Tyr Gly Ile Phe Val Phe Gly Phe Phe Leu Ala Gln  
 50 55 60  
 30 Val Leu Phe Ser Glu Leu Asn Arg Lys Arg Leu Arg Lys Trp Ile Ser  
 65 70 75 80  
 Leu Arg Pro Lys Gly Trp Asn Asp Val Arg Leu Ala Val Ile Ile Ala

|    |                                                                 |     |     |
|----|-----------------------------------------------------------------|-----|-----|
|    | 85                                                              | 90  | 95  |
|    | Gly Tyr Arg Glu Asp Pro Tyr Met Phe Gln Lys Cys Leu Glu Ser Val |     |     |
|    | 100                                                             | 105 | 110 |
|    | Arg Asp Ser Asp Tyr Gly Asn Val Ala Arg Leu Ile Cys Val Ile Asp |     |     |
| 5  | 115                                                             | 120 | 125 |
|    | Gly Asp Asp Asp Ala Asp Met Lys Met Ser Asp Val Tyr Lys Thr Ile |     |     |
|    | 130                                                             | 135 | 140 |
|    | Tyr Asn Asp Asn Ile Lys Lys Pro Glu Phe Val Leu Cys Glu Ser Asp |     |     |
|    | 145                                                             | 150 | 155 |
| 10 | 160                                                             |     |     |
|    | Asp Lys Glu Gly Glu Arg Ile Asp Ser Asp Phe Ser Arg Asp Ile Cys |     |     |
|    | 165                                                             | 170 | 175 |
|    | Val Leu Gln Pro His Arg Gly Lys Arg Glu Cys Leu Tyr Thr Gly Phe |     |     |
|    | 180                                                             | 185 | 190 |
|    | Gln Leu Ala Lys Met Asp Pro Ser Val Asn Ala Val Val Leu Ile Asp |     |     |
| 15 | 195                                                             | 200 | 205 |
|    | Ser Asp Thr Val Leu Glu Lys Asp Ala Ile Leu Glu Val Val Tyr Pro |     |     |
|    | 210                                                             | 215 | 220 |
|    | Leu Ala Cys Asp Pro Glu Ile Gln Ala Val Ala Gly Glu Cys Lys Ile |     |     |
|    | 225                                                             | 230 | 235 |
| 20 | 240                                                             |     |     |
|    | Trp Asn Thr Asp Thr Leu Leu Ser Leu Leu Val Ala Trp Arg Tyr Tyr |     |     |
|    | 245                                                             | 250 | 255 |
|    | Ser Ala Phe Cys Val Glu Arg Ser Ala Gln Ser Phe Phe Arg Thr Val |     |     |
|    | 260                                                             | 265 | 270 |
|    | Gln Cys Val Gly Gly Pro Leu Gly Ala Tyr Lys Ile Asp Ile Ile Lys |     |     |
| 25 | 275                                                             | 280 | 285 |
|    | Glu Ile Lys Asp Pro Trp Ile Ser Gln Arg Phe Leu Gly Gln Lys Cys |     |     |
|    | 290                                                             | 295 | 300 |
|    | Thr Tyr Gly Asp Asp Arg Arg Leu Thr Asn Glu Ile Leu Met Arg Gly |     |     |
|    | 305                                                             | 310 | 315 |
| 30 | 320                                                             |     |     |
|    | Lys Lys Val Val Phe Thr Pro Phe Ala Val Gly Trp Ser Asp Ser Pro |     |     |
|    | 325                                                             | 330 | 335 |
|    | Thr Asn Val Phe Arg Tyr Ile Val Gln Gln Thr Arg Trp Ser Lys Ser |     |     |

|    |                                                                 |     |     |
|----|-----------------------------------------------------------------|-----|-----|
|    | 340                                                             | 345 | 350 |
|    | Trp Cys Arg Glu Ile Trp Tyr Thr Leu Phe Ala Ala Trp Lys His Gly |     |     |
|    | 355                                                             | 360 | 365 |
|    | Leu Ser Gly Ile Trp Leu Ala Phe Glu Cys Leu Tyr Gln Ile Thr Tyr |     |     |
| 5  | 370                                                             | 375 | 380 |
|    | Phe Phe Leu Val Ile Tyr Leu Phe Ser Arg Leu Ala Val Glu Ala Asp |     |     |
|    | 385                                                             | 390 | 395 |
|    | Pro Arg Ser Gln Thr Ala Thr Val Ile Val Ser Thr Thr Val Ala Leu |     |     |
|    | 405                                                             | 410 | 415 |
| 10 | Ile Lys Cys Gly Tyr Phe Ser Phe Arg Ala Lys Asp Ile Arg Ala Phe |     |     |
|    | 420                                                             | 425 | 430 |
|    | Tyr Phe Val Leu Tyr Thr Phe Val Tyr Phe Phe Cys Met Ile Pro Ala |     |     |
|    | 435                                                             | 440 | 445 |
|    | Arg Val Thr Ala Met Met Thr Leu Trp Asp Ile Gly Trp Gly Thr Arg |     |     |
| 15 | 450                                                             | 455 | 460 |
|    | Gly Gly Asn Glu Lys Pro Ser Val Gly Thr Arg Val Ala Leu Trp Ala |     |     |
|    | 465                                                             | 470 | 475 |
|    | Lys Gln Tyr Leu Ile Ala Tyr Met Trp Trp Ala Ala Val Val Gly Ala |     |     |
|    | 485                                                             | 490 | 495 |
| 20 | Gly Val Tyr Ser Ile Val His Asn Trp Met Phe Asp Trp Asn Ser Leu |     |     |
|    | 500                                                             | 505 | 510 |
|    | Ser Tyr Arg Phe Ala Leu Val Gly Ile Cys Ser Tyr Ile Val Phe Ile |     |     |
|    | 515                                                             | 520 | 525 |
|    | Thr Ile Val Leu Val Ile Tyr Phe Thr Gly Lys Ile Thr Thr Trp Asn |     |     |
| 25 | 530                                                             | 535 | 540 |
|    | Phe Thr Lys Leu Gln Lys Glu Leu Ile Glu Asp Arg Val Leu Tyr Asp |     |     |
|    | 545                                                             | 550 | 555 |
|    | Ala Ser Thr Asn Ala Gln Ser Val                                 |     | 560 |
|    | 565                                                             |     |     |
| 30 | <210> 5                                                         |     |     |
|    | <211> 1788                                                      |     |     |
|    | <212> DNA                                                       |     |     |

&lt;213&gt; Virus Chlorella

&lt;400&gt; 5

atgtgtggca tctttggagc agtgtcaaac aacaactcta tcgaggtgtc aatcaaggg 60  
 attcagaagc tagaatatcg tgggtatgat tcgtgcggta ttgcgtatac agatgggggt 120  
 5 gcgattgagc gtatacggtc ggttgacggc attgacgac tcgtaagaa aacaatcaca 180  
 gaatcatcac cagtggccat tgctcactcg cggaggagca ccactggaat tccatcagtg 240  
 gtgaacgcac atcctcatat ttctcgcgga accagtgggt gtgagtctcg tatcgcggt 300  
 gtccacaacg gtatcattga aaactatcag cagatccgaa aatatctcat caatctcggt 360  
 tatacgtttg atagtcaaac ggacacagag gtcatcgac atttgatcga ttctcagtac 420  
 10 aatgggaata tcttgacac cgtccaaatg gctgtcaagc acctgaaggg ctcttatgcc 480  
 attgcagtta tgtgtcataa agagtctggg aaaatagtcg tggcgaaaca gaagtcaccc 540  
 ctctgacttg gaatcggtc agatgggtgct tactacatcg ctccggacgt gctggcgctg 600  
 ccgacaaata aagttgttta tatttcagat ggtttctctg cagaactatc tccagggagt 660  
 atgtccattt acgatcctga tggaaatgaa gtggaatatg aagtagagga cgttgaaatg 720  
 15 gaacaaacta gtatgtctct tgataacttt gatcattaca tgattaagga aattaatgag 780  
 caaccaatca gtatcctaaa cactataaaa aataaagggt tttatgcaga aatattcggc 840  
 gatttggcgc atgaaatctt ccaaaaaata gacaacatcc tgatactggc ttgtggtaca 900  
 agttatcacg ccggtcttgt aggaaaacag tggatagaga ccattgcgag aatccccgtg 960  
 gatgttcaca tcgcgagcga atacgaacct acaattccga gagcgaacac attggtaatc 1020  
 20 actatttcac agtcgggtga aactgcggac acgatagcgg ctttgcaacg ggcccaaac 1080  
 gccgggatga ttatatacatt gtgtatttgc aattcaccaa agagcactct tgttcgcgag 1140  
 agcgttatga agtacataac gaaatgtggg tctgaggtgt cagtagcatc aacgaaggcg 1200  
 ttacctcgc aactcgtagt actgtacatg ctggcaaacg tattggcaaa taaaaccgat 1260  
 gatttgctgg gagacctccc acaggcaata gaacgggtga ttgtttgac aaatgacgaa 1320  
 25 atgaaacact gggcgagcga aatctgcaat gcgaaatctg cgatcttcct gggaagagga 1380  
 ctaaacgcac cagttgcctt tgaggagcgc ctgaagctca aagaaatctc ttacattcat 1440  
 gcagagggtc tcctgggagg tgagttgaaa catggcccc tcgcactcct tgatgacaaa 1500  
 attcctgtta tcgtaaccgt agcagatcat gcttatttgg accatatcaa agcaaatatt 1560  
 gacgaagtgc ttgcgaggaa cgttacggta tacgccatag tagaccagta tgtgaacatc 1620  
 30 gagccccagg aacgccttcg cgttgtcaag gttccgtttg tatccaaaga attttctccg 1680  
 ataattcaca ctatcccgat gcaactgctt tcgtattacg tggcaattaa gcttggggaag 1740  
 aacgttgaca aaccaaggaa tcttgcaaaa tccgtgacca ccttttaa 1788

&lt;210&gt; 6

&lt;211&gt; 595

&lt;212&gt; PRT

&lt;213&gt; Vírus Chlorella

5 &lt;400&gt; 6

Met Cys Gly Ile Phe Gly Ala Val Ser Asn Asn Asn Ser Ile Glu Val

1 5 10 15

Ser Ile Lys Gly Ile Gln Lys Leu Glu Tyr Arg Gly Tyr Asp Ser Cys

20 25 30

10 Gly Ile Ala Tyr Thr Asp Gly Gly Ala Ile Glu Arg Ile Arg Ser Val

35 40 45

Asp Gly Ile Asp Asp Leu Arg Lys Lys Thr Ile Thr Glu Ser Ser Pro

50 55 60

Val Ala Ile Ala His Ser Arg Trp Ser Thr Thr Gly Ile Pro Ser Val

15 65 70 75 80

Val Asn Ala His Pro His Ile Ser Arg Gly Thr Ser Gly Cys Glu Ser

85 90 95

Arg Ile Ala Val Val His Asn Gly Ile Ile Glu Asn Tyr Gln Gln Ile

100 105 110

20 Arg Lys Tyr Leu Ile Asn Leu Gly Tyr Thr Phe Asp Ser Gln Thr Asp

115 120 125

Thr Glu Val Ile Ala His Leu Ile Asp Ser Gln Tyr Asn Gly Asn Ile

130 135 140

Leu His Thr Val Gln Met Ala Val Lys His Leu Lys Gly Ser Tyr Ala

25 145 150 155 160

Ile Ala Val Met Cys His Lys Glu Ser Gly Lys Ile Val Val Ala Lys

165 170 175

Gln Lys Ser Pro Leu Val Leu Gly Ile Gly Ser Asp Gly Ala Tyr Tyr

180 185 190

30 Ile Ala Ser Asp Val Leu Ala Leu Pro Thr Asn Lys Val Val Tyr Ile

195 200 205

Ser Asp Gly Phe Ser Ala Glu Leu Ser Pro Gly Ser Met Ser Ile Tyr

|    |                                                                 |     |     |     |
|----|-----------------------------------------------------------------|-----|-----|-----|
|    | 210                                                             | 215 | 220 |     |
|    | Asp Pro Asp Gly Asn Glu Val Glu Tyr Glu Val Glu Asp Val Glu Met |     |     |     |
|    | 225                                                             | 230 | 235 | 240 |
|    | Glu Gln Thr Ser Met Ser Leu Asp Asn Phe Asp His Tyr Met Ile Lys |     |     |     |
| 5  | 245                                                             | 250 | 255 |     |
|    | Glu Ile Asn Glu Gln Pro Ile Ser Ile Leu Asn Thr Ile Lys Asn Lys |     |     |     |
|    | 260                                                             | 265 | 270 |     |
|    | Gly Phe Tyr Ala Glu Ile Phe Gly Asp Leu Ala His Glu Ile Phe Gln |     |     |     |
|    | 275                                                             | 280 | 285 |     |
| 10 | Lys Ile Asp Asn Ile Leu Ile Leu Ala Cys Gly Thr Ser Tyr His Ala |     |     |     |
|    | 290                                                             | 295 | 300 |     |
|    | Gly Leu Val Gly Lys Gln Trp Ile Glu Thr Ile Ala Arg Ile Pro Val |     |     |     |
|    | 305                                                             | 310 | 315 | 320 |
|    | Asp Val His Ile Ala Ser Glu Tyr Glu Pro Thr Ile Pro Arg Ala Asn |     |     |     |
| 15 | 325                                                             | 330 | 335 |     |
|    | Thr Leu Val Ile Thr Ile Ser Gln Ser Gly Glu Thr Ala Asp Thr Ile |     |     |     |
|    | 340                                                             | 345 | 350 |     |
|    | Ala Ala Leu Gln Arg Ala Gln Asn Ala Gly Met Ile Tyr Thr Leu Cys |     |     |     |
|    | 355                                                             | 360 | 365 |     |
| 20 | Ile Cys Asn Ser Pro Lys Ser Thr Leu Val Arg Glu Ser Val Met Lys |     |     |     |
|    | 370                                                             | 375 | 380 |     |
|    | Tyr Ile Thr Lys Cys Gly Ser Glu Val Ser Val Ala Ser Thr Lys Ala |     |     |     |
|    | 385                                                             | 390 | 395 | 400 |
|    | Phe Thr Ser Gln Leu Val Val Leu Tyr Met Leu Ala Asn Val Leu Ala |     |     |     |
| 25 | 405                                                             | 410 | 415 |     |
|    | Asn Lys Thr Asp Asp Leu Leu Gly Asp Leu Pro Gln Ala Ile Glu Arg |     |     |     |
|    | 420                                                             | 425 | 430 |     |
|    | Val Ile Cys Leu Thr Asn Asp Glu Met Lys His Trp Ala Asp Glu Ile |     |     |     |
|    | 435                                                             | 440 | 445 |     |
| 30 | Cys Asn Ala Lys Ser Ala Ile Phe Leu Gly Arg Gly Leu Asn Ala Pro |     |     |     |
|    | 450                                                             | 455 | 460 |     |
|    | Val Ala Phe Glu Gly Ala Leu Lys Leu Lys Glu Ile Ser Tyr Ile His |     |     |     |

465                      470                      475                      480  
 Ala Glu Gly Phe Leu Gly Gly Glu Leu Lys His Gly Pro Leu Ala Leu  
                          485                      490                      495  
 Leu Asp Asp Lys Ile Pro Val Ile Val Thr Val Ala Asp His Ala Tyr  
 5                      500                      505                      510  
 Leu Asp His Ile Lys Ala Asn Ile Asp Glu Val Leu Ala Arg Asn Val  
                          515                      520                      525  
 Thr Val Tyr Ala Ile Val Asp Gln Tyr Val Asn Ile Glu Pro Gln Glu  
                          530                      535                      540  
 10 Arg Leu Arg Val Val Lys Val Pro Phe Val Ser Lys Glu Phe Ser Pro  
                          545                      550                      555                      560  
 Ile Ile His Thr Ile Pro Met Gln Leu Leu Ser Tyr Tyr Val Ala Ile  
                          565                      570                      575  
 Lys Leu Gly Lys Asn Val Asp Lys Pro Arg Asn Leu Ala Lys Ser Val  
 15                      580                      585                      590  
 Thr Thr Phe  
                          595  
 <210> 7  
 <211> 1788  
 20 <212> DNA  
      <213> Virus Chlorella  
      <400> 7  
  
 atgtgtggca tctttggagc agtgtcaaac aacaactcta tcgaggtgtc aatcaagggt    60  
 attcagaagc tagaatatcg tgggtatgat tcgtgcggta ttgcgtatac agatgggggt    120  
 25 gcgattgagc gtatacggtc ggttgacggc attgacgatc tgcgtaagaa aacaatcaca    180  
 gaatcatcac cagtagccat tgctcactcg cgggtggagca ccaactggaat tccatcagtg    240  
 gtgaacgcac atcctcatat ttctcgcgga accagtgggt gtgagtctcg tatcgcggta    300  
 gtccacaacg gtatcattga aaactatcag cagatccgaa aatatctcat caatctcggt    360  
 tatacgtttg atagtcaaac ggacacagag gtcattgcac atttgatcga ttctcagtac    420  
 30 aatgggaata tcttgacacac cgtccaaatg gctgtcaagc acctgaaggg ctcttatgcc    480  
 attgcagtta tgtgtcataa agagtctggt aaaatagtcg tggcgaaaca gaagtcaccc    540  
 ctctgtacttg gaatcggtc agatggtgct tactacatcg ctctggacgt gctggcgctg    600

ccgacaaata aagttgttta tatttcagat gggtttctctg cagaactatc tccagggagt 660  
 atgtccattt acgatacctga tggaaatgaa gtggaatatg aagtagagga cgttgaaatg 720  
 gaacaaacta gtatgtctct ctataacttt gatcattaca tgattaagga aattaatgag 780  
 caaccaatca gtatcctaaa cactataaaa aataaagggt tctatgcaga aatattcggg 840  
 5 gatttggcgc atgaaatctt ccaaaaaata gacaacatcc tggtagtggc ttgtggtaca 900  
 agttatcacg ccgggtctcgt cgggaaacag tggatagaga ccatcgcgaa aatccccgtg 960  
 aatgttcata tcgcaagtga atacgaaccc accattccta aagcgaacac attggtaatc 1020  
 actatttcac aatcgggtga aactgcggac acgatagcgg ctttgcaacg agcccaaac 1080  
 gccgggatga ttacacact gtgtatttgc aattctccaa agagtactct agttcgcgag 1140  
 10 agcattatga agtacatcac taaatgtggt tctgaggtgt cagtagcatc aacgaaggcg 1200  
 ttacctcgc agctcgtagt actgtatatc ctggcaaacg tattggcaaa taaaaccgac 1260  
 gatttgctgg gtgagcttcc gcaagcaata gaacgggtga ttgtttgac gagcgatgaa 1320  
 atgaaacaat gggctgatga aatatgcaat gcgaaatctg cgatcttcct ggggagagga 1380  
 ctgaacgcac cagttgcttt tgagggtgcg ttgaaactca aagagatttc ttacattcat 1440  
 15 gcggagggct tcctgggagg tgagttgaaa cacgggtccc tcgcactcct tgatgacaag 1500  
 attcctgtca tcgtaactgt ggcagatcat gcttatctgg accatatcaa agcaaattt 1560  
 gacgaagtgc ttgcgaggaa cgtcacggta tatgccattg ttgaccagta tgtgaacatc 1620  
 gagccccagg aacgtcttca tatcgtcaag gttccgtttg tgtcaaaaga attttctcca 1680  
 ataattcaca ctatcccaat gcaactgctt tcgtattacg tggcaattaa gcttggaag 1740  
 20 aatgttgata aaccgaggaa tcttgcgaaa tctgtgacca ccttttaa 1788  
 <210> 8  
 <211> 595  
 <212> PRT  
 <213> Virus Chlorella  
 25 <400> 8  
 Met Cys Gly Ile Phe Gly Ala Val Ser Asn Asn Asn Ser Ile Glu Val  
 1 5 10 15  
 Ser Ile Lys Gly Ile Gln Lys Leu Glu Tyr Arg Gly Tyr Asp Ser Cys  
 20 25 30  
 30 Gly Ile Ala Tyr Thr Asp Gly Gly Ala Ile Glu Arg Ile Arg Ser Val  
 35 40 45  
 Asp Gly Ile Asp Asp Leu Arg Lys Lys Thr Ile Thr Glu Ser Ser Pro

|    |                                                                 |     |         |
|----|-----------------------------------------------------------------|-----|---------|
|    | 50                                                              | 55  | 60      |
|    | Val Ala Ile Ala His Ser Arg Trp Ser Thr Thr Gly Ile Pro Ser Val |     |         |
|    | 65                                                              | 70  | 75 80   |
|    | Val Asn Ala His Pro His Ile Ser Arg Gly Thr Ser Gly Cys Glu Ser |     |         |
| 5  | 85                                                              | 90  | 95      |
|    | Arg Ile Ala Val Val His Asn Gly Ile Ile Glu Asn Tyr Gln Gln Ile |     |         |
|    | 100                                                             | 105 | 110     |
|    | Arg Lys Tyr Leu Ile Asn Leu Gly Tyr Thr Phe Asp Ser Gln Thr Asp |     |         |
|    | 115                                                             | 120 | 125     |
| 10 | Thr Glu Val Ile Ala His Leu Ile Asp Ser Gln Tyr Asn Gly Asn Ile |     |         |
|    | 130                                                             | 135 | 140     |
|    | Leu His Thr Val Gln Met Ala Val Lys His Leu Lys Gly Ser Tyr Ala |     |         |
|    | 145                                                             | 150 | 155 160 |
|    | Ile Ala Val Met Cys His Lys Glu Ser Gly Lys Ile Val Val Ala Lys |     |         |
| 15 | 165                                                             | 170 | 175     |
|    | Gln Lys Ser Pro Leu Val Leu Gly Ile Gly Ser Asp Gly Ala Tyr Tyr |     |         |
|    | 180                                                             | 185 | 190     |
|    | Ile Ala Ser Asp Val Leu Ala Leu Pro Thr Asn Lys Val Val Tyr Ile |     |         |
|    | 195                                                             | 200 | 205     |
| 20 | Ser Asp Gly Phe Ser Ala Glu Leu Ser Pro Gly Ser Met Ser Ile Tyr |     |         |
|    | 210                                                             | 215 | 220     |
|    | Asp Pro Asp Gly Asn Glu Val Glu Tyr Glu Val Glu Asp Val Glu Met |     |         |
|    | 225                                                             | 230 | 235 240 |
|    | Glu Gln Thr Ser Met Ser Leu Tyr Asn Phe Asp His Tyr Met Ile Lys |     |         |
| 25 | 245                                                             | 250 | 255     |
|    | Glu Ile Asn Glu Gln Pro Ile Ser Ile Leu Asn Thr Ile Lys Asn Lys |     |         |
|    | 260                                                             | 265 | 270     |
|    | Gly Phe Tyr Ala Glu Ile Phe Gly Asp Leu Ala His Glu Ile Phe Gln |     |         |
|    | 275                                                             | 280 | 285     |
| 30 | Lys Ile Asp Asn Ile Leu Val Leu Ala Cys Gly Thr Ser Tyr His Ala |     |         |
|    | 290                                                             | 295 | 300     |
|    | Gly Leu Val Gly Lys Gln Trp Ile Glu Thr Ile Ala Lys Ile Pro Val |     |         |

305                      310                      315                      320  
 Asn Val His Ile Ala Ser Glu Tyr Glu Pro Thr Ile Pro Lys Ala Asn  
                          325                      330                      335  
 Thr Leu Val Ile Thr Ile Ser Gln Ser Gly Glu Thr Ala Asp Thr Ile  
 5                      340                      345                      350  
 Ala Ala Leu Gln Arg Ala Gln Asn Ala Gly Met Ile Tyr Thr Leu Cys  
                          355                      360                      365  
 Ile Cys Asn Ser Pro Lys Ser Thr Leu Val Arg Glu Ser Ile Met Lys  
                          370                      375                      380  
 10 Tyr Ile Thr Lys Cys Gly Ser Glu Val Ser Val Ala Ser Thr Lys Ala  
                          385                      390                      395                      400  
 Phe Thr Ser Gln Leu Val Val Leu Tyr Ile Leu Ala Asn Val Leu Ala  
                          405                      410                      415  
 Asn Lys Thr Asp Asp Leu Leu Gly Glu Leu Pro Gln Ala Ile Glu Arg  
 15                      420                      425                      430  
 Val Ile Cys Leu Thr Ser Asp Glu Met Lys Gln Trp Ala Asp Glu Ile  
                          435                      440                      445  
 Cys Asn Ala Lys Ser Ala Ile Phe Leu Gly Arg Gly Leu Asn Ala Pro  
                          450                      455                      460  
 20 Val Ala Phe Glu Gly Ala Leu Lys Leu Lys Glu Ile Ser Tyr Ile His  
                          465                      470                      475                      480  
 Ala Glu Gly Phe Leu Gly Gly Glu Leu Lys His Gly Pro Leu Ala Leu  
                          485                      490                      495  
 Leu Asp Asp Lys Ile Pro Val Ile Val Thr Val Ala Asp His Ala Tyr  
 25                      500                      505                      510  
 Leu Asp His Ile Lys Ala Asn Ile Asp Glu Val Leu Ala Arg Asn Val  
                          515                      520                      525  
 Thr Val Tyr Ala Ile Val Asp Gln Tyr Val Asn Ile Glu Pro Gln Glu  
                          530                      535                      540  
 30 Arg Leu His Ile Val Lys Val Pro Phe Val Ser Lys Glu Phe Ser Pro  
                          545                      550                      555                      560  
 Ile Ile His Thr Ile Pro Met Gln Leu Leu Ser Tyr Tyr Val Ala Ile

|    |                                                                    |      |     |
|----|--------------------------------------------------------------------|------|-----|
|    | 565                                                                | 570  | 575 |
|    | Lys Leu Gly Lys Asn Val Asp Lys Pro Arg Asn Leu Ala Lys Ser Val    |      |     |
|    | 580                                                                | 585  | 590 |
|    | Thr Thr Phe                                                        |      |     |
| 5  | 595                                                                |      |     |
|    | <210> 9                                                            |      |     |
|    | <211> 2034                                                         |      |     |
|    | <212> DNA                                                          |      |     |
|    | <213> <i>Arabidopsis thaliana</i>                                  |      |     |
| 10 | <400> 9                                                            |      |     |
|    | atgtgtggaa tcttcgcgta tctgaatttt cacgcgaaca aagagagacg atacattctc  | 60   |     |
|    | gatgttctct tcaatggctc tcgtcgtctt gaatacagag gctacgattc tgctggaatc  | 120  |     |
|    | gccattgata attcttctcc ttcttcttct cctctcgtgt ttcgtcaagc aggaaacatt  | 180  |     |
|    | gaatcacttg ttaattccgt taacgaagag attacgaata cggatttgaa tctagacgaa  | 240  |     |
| 15 | gttttctact ttcattgctg aattgcacat acgaggtggg ctactcatgg tgagccagct  | 300  |     |
|    | ccaaggaata gtcactctca atcttctggc cctggagatg attttttggg gggtcataat  | 360  |     |
|    | gggtttatca ctaactatga ggtattgaaa gaaacgtagg tgaggcatgg atttactttt  | 420  |     |
|    | gaatcggaca cagatactga agtaattcct aagcttgcta agtttgtttt tgacaaagct  | 480  |     |
|    | aatgaagaag gtggacaaac tgttacattc tgtgaagttg tgtttgaagt gatgaggcat  | 540  |     |
| 20 | cttgaaggag cttatgctct tatatttaaa agctggcatt atccgaatga gtttaattgcg | 600  |     |
|    | tgcaagcttg gtagtccatt gcttttaggt gttaaagagc tagatcaagg tgagagcaat  | 660  |     |
|    | agtcatgttt tccaagatgc tcaactttcta tctaagaatg accatcccaa ggagtttttc | 720  |     |
|    | ctatcaagtg atccacatgc tcttggtgag cacacaaaga aagttttggg gattgaagat  | 780  |     |
|    | ggcgaagttg tcaatctcaa ggatggaggt gtatcaatac ttaagtttga aaatgagagg  | 840  |     |
| 25 | ggaaggtgta atggtttatc gagacctgct tcagtggaac gtgccttatc tgttctagag  | 900  |     |
|    | atggaggtag agcaaataag caagggaaaa tatgatcatt acatgcaaaa ggaaatccac  | 960  |     |
|    | gagcagccag aatctttaac tactacaatg agaggccgac ttatacgcgg tggttcacgt  | 1020 |     |
|    | aaaacgaaaa ccgtcctctt aggtgggctg aaagatcacc taaagaccat aagacgcagc  | 1080 |     |
|    | cggcgtatag tttttattgg atgtgggaca agttacaatg ccgctcttgc atcaagacct  | 1140 |     |
| 30 | atccttgaag aactctctgg tataccagtc agtatggaga ttgctagtga tctatgggac  | 1200 |     |
|    | cggcaaggtc caatatacag agaagatacc gcggtgtttg tgagtcagtc tggtgaaact  | 1260 |     |
|    | gcagatacac tacttgcttt ggactatgct cgagaaaacg gtgcattatg tgtcggcata  | 1320 |     |

actaacaccg ttgggagctc catagctaga aaaacacact gtggtgtcca tataaacgca 1380  
 ggagctgaga ttggtgtcgc aagtacaaag gcatatacaa gtcagattgt ggtaatggta 1440  
 atgctagctt tagcaatagg aagtacacaca atctccagcc aaaagagacg ggaagctata 1500  
 atcgatggac tacttgattt gccgtataag gttaaggaag tactaaagct agacgatgaa 1560  
 5 atgaaagatc tcgcgcaact cttgatagac gagcagtcac tgctagtgtt tggcagagga 1620  
 tacaactacg caacagcttt agaaggagca ttaaaagtaa aagaagtagc acttatgcac 1680  
 agtgaaggaa tacttgacag agaaatgaaa catggacctt tagctttggt tgatgagaat 1740  
 ctcccatag ctgtgattgc cactcgtgat gcttgtttca gtaaacaaca atctgtgatt 1800  
 cagcaacttc acgcacgcaa agggagacta atagtaatgt gctcaaaagg tgatgctgca 1860  
 10 tcggtaagct cgagtggttc ttgtcgagct atcgaagttc ctcaagttga agattgttta 1920  
 caacctgtta ttaatatagt gccattacag ttgttggtt atcatctgac tgttttgaga 1980  
 ggtcacaatg ttgatcaacc gaggaatctg gcaaagagtg tgactactca atag 2034

<210> 10

<211> 677

15 <212> PRT

<213> *Arabidopsis thaliana*

<400> 10

Met Cys Gly Ile Phe Ala Tyr Leu Asn Phe His Ala Asn Lys Glu Arg

1 5 10 15

20 Arg Tyr Ile Leu Asp Val Leu Phe Asn Gly Leu Arg Arg Leu Glu Tyr

20 25 30

Arg Gly Tyr Asp Ser Ala Gly Ile Ala Ile Asp Asn Ser Ser Pro Ser

35 40 45

Ser Ser Pro Leu Val Phe Arg Gln Ala Gly Asn Ile Glu Ser Leu Val

25 50 55 60

Asn Ser Val Asn Glu Glu Ile Thr Asn Thr Asp Leu Asn Leu Asp Glu

65 70 75 80

Val Phe Tyr Phe His Ala Gly Ile Ala His Thr Arg Trp Ala Thr His

85 90 95

30 Gly Glu Pro Ala Pro Arg Asn Ser His Pro Gln Ser Ser Gly Pro Gly

100 105 110

Asp Asp Phe Leu Val Val His Asn Gly Val Ile Thr Asn Tyr Glu Val

|    |                 |                             |                     |
|----|-----------------|-----------------------------|---------------------|
|    | 115             | 120                         | 125                 |
|    | Leu Lys Glu Thr | Leu Val Arg His Gly Phe Thr | Phe Glu Ser Asp Thr |
|    | 130             | 135                         | 140                 |
|    | Asp Thr Glu Val | Ile Pro Lys Leu Ala Lys Phe | Val Phe Asp Lys Ala |
| 5  | 145             | 150                         | 155 160             |
|    | Asn Glu Glu Gly | Gly Gln Thr Val Thr Phe Cys | Glu Val Val Phe Glu |
|    | 165             | 170                         | 175                 |
|    | Val Met Arg His | Leu Glu Gly Ala Tyr Ala Leu | Ile Phe Lys Ser Trp |
|    | 180             | 185                         | 190                 |
| 10 | His Tyr Pro Asn | Glu Leu Ile Ala Cys Lys Leu | Gly Ser Pro Leu Leu |
|    | 195             | 200                         | 205                 |
|    | Leu Gly Val Lys | Glu Leu Asp Gln Gly Glu Ser | Asn Ser His Val Phe |
|    | 210             | 215                         | 220                 |
|    | Gln Asp Ala His | Phe Leu Ser Lys Asn Asp His | Pro Lys Glu Phe Phe |
| 15 | 225             | 230                         | 235 240             |
|    | Leu Ser Ser Asp | Pro His Ala Leu Val Glu His | Thr Lys Lys Val Leu |
|    | 245             | 250                         | 255                 |
|    | Val Ile Glu Asp | Gly Glu Val Val Asn Leu Lys | Asp Gly Gly Val Ser |
|    | 260             | 265                         | 270                 |
| 20 | Ile Leu Lys Phe | Glu Asn Glu Arg Gly Arg Cys | Asn Gly Leu Ser Arg |
|    | 275             | 280                         | 285                 |
|    | Pro Ala Ser Val | Glu Arg Ala Leu Ser Val Leu | Glu Met Glu Val Glu |
|    | 290             | 295                         | 300                 |
|    | Gln Ile Ser Lys | Gly Lys Tyr Asp His Tyr Met | Gln Lys Glu Ile His |
| 25 | 305             | 310                         | 315 320             |
|    | Glu Gln Pro Glu | Ser Leu Thr Thr Thr Met Arg | Gly Arg Leu Ile Arg |
|    | 325             | 330                         | 335                 |
|    | Gly Gly Ser Arg | Lys Thr Lys Thr Val Leu Leu | Gly Gly Leu Lys Asp |
|    | 340             | 345                         | 350                 |
| 30 | His Leu Lys Thr | Ile Arg Arg Ser Arg Arg Ile | Val Phe Ile Gly Cys |
|    | 355             | 360                         | 365                 |
|    | Gly Thr Ser Tyr | Asn Ala Ala Leu Ala Ser Arg | Pro Ile Leu Glu Glu |

|    |                                                                 |     |     |     |
|----|-----------------------------------------------------------------|-----|-----|-----|
|    | 370                                                             | 375 | 380 |     |
|    | Leu Ser Gly Ile Pro Val Ser Met Glu Ile Ala Ser Asp Leu Trp Asp |     |     |     |
|    | 385                                                             | 390 | 395 | 400 |
|    | Arg Gln Gly Pro Ile Tyr Arg Glu Asp Thr Ala Val Phe Val Ser Gln |     |     |     |
| 5  | 405                                                             | 410 | 415 |     |
|    | Ser Gly Glu Thr Ala Asp Thr Leu Leu Ala Leu Asp Tyr Ala Arg Glu |     |     |     |
|    | 420                                                             | 425 | 430 |     |
|    | Asn Gly Ala Leu Cys Val Gly Ile Thr Asn Thr Val Gly Ser Ser Ile |     |     |     |
|    | 435                                                             | 440 | 445 |     |
| 10 | Ala Arg Lys Thr His Cys Gly Val His Ile Asn Ala Gly Ala Glu Ile |     |     |     |
|    | 450                                                             | 455 | 460 |     |
|    | Gly Val Ala Ser Thr Lys Ala Tyr Thr Ser Gln Ile Val Val Met Val |     |     |     |
|    | 465                                                             | 470 | 475 | 480 |
|    | Met Leu Ala Leu Ala Ile Gly Ser Asp Thr Ile Ser Ser Gln Lys Arg |     |     |     |
| 15 | 485                                                             | 490 | 495 |     |
|    | Arg Glu Ala Ile Ile Asp Gly Leu Leu Asp Leu Pro Tyr Lys Val Lys |     |     |     |
|    | 500                                                             | 505 | 510 |     |
|    | Glu Val Leu Lys Leu Asp Asp Glu Met Lys Asp Leu Ala Gln Leu Leu |     |     |     |
|    | 515                                                             | 520 | 525 |     |
| 20 | Ile Asp Glu Gln Ser Leu Leu Val Phe Gly Arg Gly Tyr Asn Tyr Ala |     |     |     |
|    | 530                                                             | 535 | 540 |     |
|    | Thr Ala Leu Glu Gly Ala Leu Lys Val Lys Glu Val Ala Leu Met His |     |     |     |
|    | 545                                                             | 550 | 555 | 560 |
|    | Ser Glu Gly Ile Leu Ala Gly Glu Met Lys His Gly Pro Leu Ala Leu |     |     |     |
| 25 | 565                                                             | 570 | 575 |     |
|    | Val Asp Glu Asn Leu Pro Ile Ala Val Ile Ala Thr Arg Asp Ala Cys |     |     |     |
|    | 580                                                             | 585 | 590 |     |
|    | Phe Ser Lys Gln Gln Ser Val Ile Gln Gln Leu His Ala Arg Lys Gly |     |     |     |
|    | 595                                                             | 600 | 605 |     |
| 30 | Arg Leu Ile Val Met Cys Ser Lys Gly Asp Ala Ala Ser Val Ser Ser |     |     |     |
|    | 610                                                             | 615 | 620 |     |
|    | Ser Gly Ser Cys Arg Ala Ile Glu Val Pro Gln Val Glu Asp Cys Leu |     |     |     |

|    |                                                                    |      |     |     |  |
|----|--------------------------------------------------------------------|------|-----|-----|--|
|    | 625                                                                | 630  | 635 | 640 |  |
|    | Gln Pro Val Ile Asn Ile Val Pro Leu Gln Leu Leu Ala Tyr His Leu    |      |     |     |  |
|    | 645                                                                | 650  | 655 |     |  |
|    | Thr Val Leu Arg Gly His Asn Val Asp Gln Pro Arg Asn Leu Ala Lys    |      |     |     |  |
| 5  | 660                                                                | 665  | 670 |     |  |
|    | Ser Val Thr Thr Gln                                                |      |     |     |  |
|    | 675                                                                |      |     |     |  |
|    | <210> 11                                                           |      |     |     |  |
|    | <211> 1170                                                         |      |     |     |  |
| 10 | <212> DNA                                                          |      |     |     |  |
|    | <213> Virus Chlorella                                              |      |     |     |  |
|    | <400> 11                                                           |      |     |     |  |
|    | atgtcacgaa tcgcagtcgt tggtttgtgt tacgtcggaa ccgcttgtgc agtacttctt  | 60   |     |     |  |
|    | gctcaaaaaa acgaagtcac cgtgcttgat attagcgaag accgtgttca gctaatacaag | 120  |     |     |  |
| 5  | aacaagaaga gtccaatcga ggacaaggaa atcgaagagt ttctcgaaac gaaagacctg  | 180  |     |     |  |
|    | aacctgaccg cgacgactga caaggttctt gcatacgaaa acgccgaatt tgtcatcatc  | 240  |     |     |  |
|    | gcaaccccga ctgactatga cgtgggttact aggtatttta acacgaaatc tgtggaaagc | 300  |     |     |  |
|    | gttatcgggg atgtgatcaa aaatacacgg acccagccaa ctattgtgat taaatctacc  | 360  |     |     |  |
|    | atccccattg gatttggtga taagggttcgt gagcaattca actacagcaa cattatatcc | 420  |     |     |  |
| 20 | tctccggagt ttctgcgaga aggtagggca ttgtatgata atctctatcc atcgcgtatt  | 480  |     |     |  |
|    | atcgtaggag atgattcccc cattgcgctt aagttcgcaa accttctcgt tgaaggttct  | 540  |     |     |  |
|    | aaaaccccac ttgcacctgt cctgacgatg gggactcgtg aagccgaggc cgtcaaacata | 600  |     |     |  |
|    | ttctctaaca cgtatcttgc gatgcgagtt gcatatttca acgaactaga tacgtttgca  | 660  |     |     |  |
|    | ttgtctcatg gtatgagtgc gaaagaaatc attgacgggtg tgactctgga gcctcgaatt | 720  |     |     |  |
| 25 | ggtcaggggt actcaaacc ttcgttcggt tacggagctt attgcttccc aaaggatacg   | 780  |     |     |  |
|    | aagcaacttc tggctaactt tgaggggggtg cctcaaaata tcatcggggc aattgtagaa | 840  |     |     |  |
|    | tcaaatgaaa ctcgcaagga agcgattgta agtgaagtag aaaatcgttt tcccacgact  | 900  |     |     |  |
|    | gttggtgtgt ataagctcgc tgctaaagcg ggttctgata attttaggag ttctgcaatt  | 960  |     |     |  |
|    | gtagacataa tggagcgact tgcaaacagg ggttatcaca ttaagatttt cgaaccaact  | 1020 |     |     |  |
| 30 | gtggaacaat tcgaaaactt tgaagttgat aacaacctga caacatttgc gactgagagc  | 1080 |     |     |  |
|    | gatgtaatta tcgcaaacag agttcccgtt gaacatcgca ttctctttgg taaaaaattg  | 1140 |     |     |  |
|    | atcacacgtg atgtatatgg cgataactaa                                   | 1170 |     |     |  |

&lt;210&gt; 12

&lt;211&gt; 389

&lt;212&gt; PRT

&lt;213&gt; Virus Chlorella

5 &lt;400&gt; 12

Met Ser Arg Ile Ala Val Val Gly Cys Gly Tyr Val Gly Thr Ala Cys

1 5 10 15

Ala Val Leu Leu Ala Gln Lys Asn Glu Val Thr Val Leu Asp Ile Ser

20 25 30

10 Glu Asp Arg Val Gln Leu Ile Lys Asn Lys Lys Ser Pro Ile Glu Asp

35 40 45

Lys Glu Ile Glu Glu Phe Leu Glu Thr Lys Asp Leu Asn Leu Thr Ala

50 55 60

Thr Thr Asp Lys Val Leu Ala Tyr Glu Asn Ala Glu Phe Val Ile Ile

15 65 70 75 80

Ala Thr Pro Thr Asp Tyr Asp Val Val Thr Arg Tyr Phe Asn Thr Lys

85 90 95

Ser Val Glu Ser Val Ile Gly Asp Val Ile Lys Asn Thr Arg Thr Gln

100 105 110

20 Pro Thr Ile Val Ile Lys Ser Thr Ile Pro Ile Gly Phe Val Asp Lys

115 120 125

Val Arg Glu Gln Phe Asn Tyr Ser Asn Ile Ile Phe Ser Pro Glu Phe

130 135 140

Leu Arg Glu Gly Arg Ala Leu Tyr Asp Asn Leu Tyr Pro Ser Arg Ile

25 145 150 155 160

Ile Val Gly Asp Asp Ser Pro Ile Ala Leu Lys Phe Ala Asn Leu Leu

165 170 175

Val Glu Gly Ser Lys Thr Pro Leu Ala Pro Val Leu Thr Met Gly Thr

180 185 190

30 Arg Glu Ala Glu Ala Val Lys Leu Phe Ser Asn Thr Tyr Leu Ala Met

195 200 205

Arg Val Ala Tyr Phe Asn Glu Leu Asp Thr Phe Ala Leu Ser His Gly

210                      215                      220  
 Met Ser Ala Lys Glu Ile Ile Asp Gly Val Thr Leu Glu Pro Arg Ile  
 225                      230                      235                      240  
 Gly Gln Gly Tyr Ser Asn Pro Ser Phe Gly Tyr Gly Ala Tyr Cys Phe  
 5                      245                      250                      255  
 Pro Lys Asp Thr Lys Gln Leu Leu Ala Asn Phe Glu Gly Val Pro Gln  
 260                      265                      270  
 Asn Ile Ile Gly Ala Ile Val Glu Ser Asn Glu Thr Arg Lys Glu Ala  
 275                      280                      285  
 10 Ile Val Ser Glu Val Glu Asn Arg Phe Pro Thr Thr Val Gly Val Tyr  
 290                      295                      300  
 Lys Leu Ala Ala Lys Ala Gly Ser Asp Asn Phe Arg Ser Ser Ala Ile  
 305                      310                      315                      320  
 Val Asp Ile Met Glu Arg Leu Ala Asn Arg Gly Tyr His Ile Lys Ile  
 5                      325                      330                      335  
 Phe Glu Pro Thr Val Glu Gln Phe Glu Asn Phe Glu Val Asp Asn Asn  
 340                      345                      350  
 Leu Thr Thr Phe Ala Thr Glu Ser Asp Val Ile Ile Ala Asn Arg Val  
 355                      360                      365  
 20 Pro Val Glu His Arg Ile Leu Phe Gly Lys Lys Leu Ile Thr Arg Asp  
 370                      375                      380  
 Val Tyr Gly Asp Asn  
 385  
 <210> 13  
 25 <211> 1170  
 <212> DNA  
 <213> Virus Chlorella  
 <400> 13  
 atgtcacgaa tcgcagtcgt tggctgtggt tacgtcggaa ccgcttgtgc agtacttctt 60  
 30 gctcaaaaaa acgaagtcac cgtgcttgat attagtgaag accgtgttca gctaatacaag 120  
 aacaagaaga gtccaatcga ggacaaggaa atcgaagagt ttctcgaaac gaaagacctg 180  
 aacctgaccg cgacgactga caaggttctt gcatacgaaa acgccgaatt tgtcatcatc 240

gcaacccccga ctgactatga cgtgggttact aggtatttta acacgaaatc tgtggaaagc 300  
 gttatcgggg atgtgatcga aaatacacgg acccagccaa ctattgtgat taaatctacc 360  
 atccccattg gatttggtga taagggtcgt gagcaattca actacagcaa cattatatcc 420  
 tctccggagt ttctgcgca aggtagggca ttgtatgata atctctatcc atcgcgtatt 480  
 5 atcgtaggag atgattcccc cattgcgctt aagttcgcaa accttctcgt tgaaggttct 540  
 aaaactccgc ttgccctgt cctgacgatg ggaactcgcg aagccgaggc cgtcaaaacta 600  
 ttctctaaca cgtatcttgc aatgcgagtt gcatacttca acgaactaga tacattcgca 660  
 atgtctcatg gtatgaatgc gaaagaaatc attgacggtg tgactttgga gcctcgcatt 720  
 ggtcaggggt actcaaacc ttcgttcggt tatggagctt attgctttcc gaaggatacg 780  
 10 aagcaactgc tggctaattt cgagggggtg cctcaagata taatcggggc aattgtagaa 840  
 tcaaataaaa ctcgcaagga agcgattgta agtgaagtag aaaatcggtt tcccacgact 900  
 gttggtgtgt ataagctcgc tgctaaagcg ggttctgata attttagaag ttctgcaatt 960  
 gtagacataa tggagcgact tgcaaacagg ggttatcaca ttaagatttt cgaaccaact 1020  
 gtggaacaat tcgaaaactt tgaagttgat aacaacctga caacatttgc gactgatagc 1080  
 15 gatgtaatta tcgcaaacag agttcccggt gaacatcgca ttctctttgg taaaaaattg 1140  
 atcacacgtg atgtatatgg cgataactaa 1170

<210> 14

<211> 389

<212> PRT

20 <213> Virus Chlorella

<400> 14

Met Ser Arg Ile Ala Val Val Gly Cys Gly Tyr Val Gly Thr Ala Cys

1 5 10 15

Ala Val Leu Leu Ala Gln Lys Asn Glu Val Thr Val Leu Asp Ile Ser

25 20 25 30

Glu Asp Arg Val Gln Leu Ile Lys Asn Lys Lys Ser Pro Ile Glu Asp

35 40 45

Lys Glu Ile Glu Glu Phe Leu Glu Thr Lys Asp Leu Asn Leu Thr Ala

50 55 60

30 Thr Thr Asp Lys Val Leu Ala Tyr Glu Asn Ala Glu Phe Val Ile Ile

65 70 75 80

Ala Thr Pro Thr Asp Tyr Asp Val Val Thr Arg Tyr Phe Asn Thr Lys

|    |                                                                 |     |     |
|----|-----------------------------------------------------------------|-----|-----|
|    | 85                                                              | 90  | 95  |
|    | Ser Val Glu Ser Val Ile Gly Asp Val Ile Glu Asn Thr Arg Thr Gln |     |     |
|    | 100                                                             | 105 | 110 |
|    | Pro Thr Ile Val Ile Lys Ser Thr Ile Pro Ile Gly Phe Val Asp Lys |     |     |
| 5  | 115                                                             | 120 | 125 |
|    | Val Arg Glu Gln Phe Asn Tyr Ser Asn Ile Ile Phe Ser Pro Glu Phe |     |     |
|    | 130                                                             | 135 | 140 |
|    | Leu Arg Glu Gly Arg Ala Leu Tyr Asp Asn Leu Tyr Pro Ser Arg Ile |     |     |
|    | 145                                                             | 150 | 155 |
| 10 | 160                                                             |     |     |
|    | Ile Val Gly Asp Asp Ser Pro Ile Ala Leu Lys Phe Ala Asn Leu Leu |     |     |
|    | 165                                                             | 170 | 175 |
|    | Val Glu Gly Ser Lys Thr Pro Leu Ala Pro Val Leu Thr Met Gly Thr |     |     |
|    | 180                                                             | 185 | 190 |
|    | Arg Glu Ala Glu Ala Val Lys Leu Phe Ser Asn Thr Tyr Leu Ala Met |     |     |
| 15 | 195                                                             | 200 | 205 |
|    | Arg Val Ala Tyr Phe Asn Glu Leu Asp Thr Phe Ala Met Ser His Gly |     |     |
|    | 210                                                             | 215 | 220 |
|    | Met Asn Ala Lys Glu Ile Ile Asp Gly Val Thr Leu Glu Pro Arg Ile |     |     |
|    | 225                                                             | 230 | 235 |
| 20 | 240                                                             |     |     |
|    | Gly Gln Gly Tyr Ser Asn Pro Ser Phe Gly Tyr Gly Ala Tyr Cys Phe |     |     |
|    | 245                                                             | 250 | 255 |
|    | Pro Lys Asp Thr Lys Gln Leu Leu Ala Asn Phe Glu Gly Val Pro Gln |     |     |
|    | 260                                                             | 265 | 270 |
|    | Asp Ile Ile Gly Ala Ile Val Glu Ser Asn Glu Thr Arg Lys Glu Ala |     |     |
| 25 | 275                                                             | 280 | 285 |
|    | Ile Val Ser Glu Val Glu Asn Arg Phe Pro Thr Thr Val Gly Val Tyr |     |     |
|    | 290                                                             | 295 | 300 |
|    | Lys Leu Ala Ala Lys Ala Gly Ser Asp Asn Phe Arg Ser Ser Ala Ile |     |     |
|    | 305                                                             | 310 | 315 |
| 30 | 320                                                             |     |     |
|    | Val Asp Ile Met Glu Arg Leu Ala Asn Arg Gly Tyr His Ile Lys Ile |     |     |
|    | 325                                                             | 330 | 335 |
|    | Phe Glu Pro Thr Val Glu Gln Phe Glu Asn Phe Glu Val Asp Asn Asn |     |     |

340 345 350  
 Leu Thr Thr Phe Ala Thr Asp Ser Asp Val Ile Ile Ala Asn Arg Val  
 355 360 365  
 Pro Val Glu His Arg Ile Leu Phe Gly Lys Lys Leu Ile Thr Arg Asp  
 5 370 375 380  
 Val Tyr Gly Asp Asn  
 385  
 <210> 15  
 <211> 1443  
 10 <212> DNA  
 <213> *Arabidopsis thaliana*  
 <400> 15  
 atggtgaaga tctgttgtat tggagctgga tatgtaggag gaccaacaat ggcagtgatt 60  
 gcattgaaat gtccagatat tgaagtggca gttgttgata tctctgttcc tagaatcaac 120  
 15 gcttgaaca gtgatcagct tccgatttac gagccaggct ttgacgatat cgtaaagcaa 180  
 tgcagaggaa agaattcttt cttcagtact gatgtggaga aacatgtag agaagctgat 240  
 attgtctttg tctctgttaa cacaccgact aaaacgactg gtcttggagc tgggaaagct 300  
 gctgatctca cttattggga gagtgtgct cgtatgatcg cggatgtatc ggtttctgac 360  
 aagattggtt ttgagaaatc gactgttccg gtgaagacag ctgaagctat tgagaagatt 420  
 20 ttgatgcata acagtaaagg aatcaagttt cagattcttt cgaatccgga gtttcttgct 480  
 gaaggaactg ctatcgctga tctttttaac cctgaccgtg ttttgatcgg agggcgagaa 540  
 acacctgaag gattcaaagc tgttcagaca cttaaagagg tttatgctaa ttgggttcct 600  
 gaaggtcaga tcatcacaac taatctctgg tctgctgagc tttctaagtt agctgcaaat 660  
 gctttcttgg ctgagaggat ttcacagtc aatgccatgt ctgcactttg tgaatccact 720  
 25 ggtgctgatg ttactcaagt gtcttacgct gttggtactg attcaagaat cggttccaaa 780  
 ttcttgaacg ctagtgttgg attcggaggt tcttgtttcc agaaggacat tctgaatctc 840  
 gtctacatct gtcaatgcaa cggacttcca gaagtggcgg aatactggaa acaagtgatc 900  
 aagatcaacg attaccaaaa gaaccggttc gtgaacagaa tcgtgtcttc tatgttcaac 960  
 actgtctcca acaagaaggt tgcgattctt ggattcgcat tcaagaaaga cactggtgac 1020  
 30 acaagggaaa cacctgccat tgatgtgtgt aaaggtctat taggagacaa agcacagatc 1080  
 agtatctatg atcctcaagt cacagaggaa cagattcaga gagatctctc gatgaaaaag 1140  
 ttcgactggg accatcctct tcacttgacg ccaatgagtc caaccacagt gaaacaagtg 1200

```

    agtgtgactt gggacgcata tgaagctaca aaagacgcac acgcgggttg cgttttgact 1260
    gagtgggacg agtttaagtc gtttagattac cagaagatct tcgacaacat gcagaaaccg 1320
    gcttttatct tcgacggaag aaacattatg aatgttaaca agttaagaga gattgggttc 1380
    attgtttact ccattggtaa gccacttgac ccatggctca aggacatgcc tgcctttgtc 1440
5   taa 1443

    <210> 16
    <211> 480
    <212> PRT
    <213> Arabidopsis thaliana
10  <400> 16

    Met Val Lys Ile Cys Cys Ile Gly Ala Gly Tyr Val Gly Gly Pro Thr
    1           5           10           15
    Met Ala Val Ile Ala Leu Lys Cys Pro Asp Ile Glu Val Ala Val Val
           20           25           30
15  Asp Ile Ser Val Pro Arg Ile Asn Ala Trp Asn Ser Asp Gln Leu Pro
           35           40           45
    Ile Tyr Glu Pro Gly Leu Asp Asp Ile Val Lys Gln Cys Arg Gly Lys
           50           55           60
    Asn Leu Phe Phe Ser Thr Asp Val Glu Lys His Val Arg Glu Ala Asp
20  65           70           75           80
    Ile Val Phe Val Ser Val Asn Thr Pro Thr Lys Thr Thr Gly Leu Gly
           85           90           95
    Ala Gly Lys Ala Ala Asp Leu Thr Tyr Trp Glu Ser Ala Ala Arg Met
           100          105          110
25  Ile Ala Asp Val Ser Val Ser Asp Lys Ile Val Val Glu Lys Ser Thr
           115          120          125
    Val Pro Val Lys Thr Ala Glu Ala Ile Glu Lys Ile Leu Met His Asn
           130          135          140
    Ser Lys Gly Ile Lys Phe Gln Ile Leu Ser Asn Pro Glu Phe Leu Ala
30  145          150          155          160
    Glu Gly Thr Ala Ile Ala Asp Leu Phe Asn Pro Asp Arg Val Leu Ile
           165          170          175

```

Gly Gly Arg Glu Thr Pro Glu Gly Phe Lys Ala Val Gln Thr Leu Lys  
 180 185 190  
 Glu Val Tyr Ala Asn Trp Val Pro Glu Gly Gln Ile Ile Thr Thr Asn  
 195 200 205  
 5 Leu Trp Ser Ala Glu Leu Ser Lys Leu Ala Ala Asn Ala Phe Leu Ala  
 210 215 220  
 Gln Arg Ile Ser Ser Val Asn Ala Met Ser Ala Leu Cys Glu Ser Thr  
 225 230 235 240  
 Gly Ala Asp Val Thr Gln Val Ser Tyr Ala Val Gly Thr Asp Ser Arg  
 10 245 250 255  
 Ile Gly Ser Lys Phe Leu Asn Ala Ser Val Gly Phe Gly Gly Ser Cys  
 260 265 270  
 Phe Gln Lys Asp Ile Leu Asn Leu Val Tyr Ile Cys Gln Cys Asn Gly  
 275 280 285  
 15 Leu Pro Glu Val Ala Glu Tyr Trp Lys Gln Val Ile Lys Ile Asn Asp  
 290 295 300  
 Tyr Gln Lys Asn Arg Phe Val Asn Arg Ile Val Ser Ser Met Phe Asn  
 305 310 315 320  
 Thr Val Ser Asn Lys Lys Val Ala Ile Leu Gly Phe Ala Phe Lys Lys  
 20 325 330 335  
 Asp Thr Gly Asp Thr Arg Glu Thr Pro Ala Ile Asp Val Cys Lys Gly  
 340 345 350  
 Leu Leu Gly Asp Lys Ala Gln Ile Ser Ile Tyr Asp Pro Gln Val Thr  
 355 360 365  
 25 Glu Glu Gln Ile Gln Arg Asp Leu Ser Met Lys Lys Phe Asp Trp Asp  
 370 375 380  
 His Pro Leu His Leu Gln Pro Met Ser Pro Thr Thr Val Lys Gln Val  
 385 390 395 400  
 Ser Val Thr Trp Asp Ala Tyr Glu Ala Thr Lys Asp Ala His Ala Val  
 30 405 410 415  
 Cys Val Leu Thr Glu Trp Asp Glu Phe Lys Ser Leu Asp Tyr Gln Lys  
 420 425 430

Ile Phe Asp Asn Met Gln Lys Pro Ala Phe Ile Phe Asp Gly Arg Asn  
435 440 445

Ile Met Asn Val Asn Lys Leu Arg Glu Ile Gly Phe Ile Val Tyr Ser  
450 455 460

5 Ile Gly Lys Pro Leu Asp Pro Trp Leu Lys Asp Met Pro Ala Phe Val  
465 470 475 480

<210> 17  
<211> 1443  
<212> DNA

10 <213> *Arabidopsis thaliana*  
<400> 17

atggtgaaga tatgttgtat tggagctggg tatgttggtg gaccaacaat ggcagtgatt 60  
gcattgaaat gtccagacgt tgaagtagcg gttgttgata tctctgtacc acgtatcaac 120  
gcttgaaca gtgacacgct tccgatttac gagcctgggc ttgatgatgt tgtgaagcaa 180  
15 tgccgtggca agaacctttt ctttagtact gatgttgaga aacatgtag ggaagctgat 240  
attgtgtttg tttctgtcaa cacaccgact aagactagag gtcttgggtgc tggtaaagct 300  
gcggtatctta cgtactggga gagcgctgcg cgtatgatcg ctgatgtttc ggtatcggat 360  
aagattgtcg ttgagaaatc gactgttccg gttaaaacag ctgaagctat tgagaagatt 420  
ttgacacata acagtaaagg gattaagttt cagattcttt cgaatcccga gtttttggcg 480  
20 gaaggaaccg cgattaagga cctatttaac ccggaccgtg ttcttatcgg agggcgggaa 540  
acccagaag ggtttaaaagc ggtgcagact ctcaagaatg tgtatgcaca ctgggttctt 600  
gaaggccaaa tcataacaac caatctcttg tctgctgagc tgtccaagct tgcggcaaac 660  
gctttcttgg ctcaaaggat ttcacagtg aatgctatgt cggctctgtg tgaagccaca 720  
ggcgcagatg tcacgcaagt gtcttacgcg gttggtacag actcaaggat tgggtcccaag 780  
25 ttcttgaact cgagtgttgg attcgggtgg tctgtgttcc agaaggacat tctgaatctt 840  
gtctacatct gtgagtgcaa cggactcccg gaagtggcag agtactggaa gcaagtcac 900  
aagatcaatg actaccagaa gagccggttc gtgaaccgtg ttgtttcctc catgttcaac 960  
tctgtatcaa acaagaagat tgcggttctc ggtttcgcac tcaagaaaga caccggtgac 1020  
acaagggaga ctccagccat cgatgtgtgc aagggtcttt tagaagacaa agcaaggcta 1080  
30 agcatttacg acccacaagt gactgaggat cagatccaga gggatttatc catgaacaag 1140  
ttcgactggg accatcctct acatttgcag ccaatgagcc caacaacagt gaaacaagt 1200  
accgttactt gggacgcata cgaagcaact aaggacgctc acggtatctg catcatgacc 1260

gagtgggatg agttcaagaa ccttgatttc cagaagatct ttgacaacat gcagaaacca 1320  
gctttcgtgt tcgatggaag aaacattatg aatctgcaaa agctaaggga gattgggttc 1380  
attgtttact ccattggtaa gcctctcgac gactggctca aggacatgcc tgccgttgcc 1440  
taa 1443

5 <210> 18  
<211> 480  
<212> PRT  
<213> *Arabidopsis thaliana*  
<400> 18

10 Met Val Lys Ile Cys Cys Ile Gly Ala Gly Tyr Val Gly Gly Pro Thr  
1 5 10 15  
Met Ala Val Ile Ala Leu Lys Cys Pro Asp Val Glu Val Ala Val Val  
20 25 30  
Asp Ile Ser Val Pro Arg Ile Asn Ala Trp Asn Ser Asp Thr Leu Pro  
15 35 40 45  
Ile Tyr Glu Pro Gly Leu Asp Asp Val Val Lys Gln Cys Arg Gly Lys  
50 55 60  
Asn Leu Phe Phe Ser Thr Asp Val Glu Lys His Val Arg Glu Ala Asp  
65 70 75 80  
20 Ile Val Phe Val Ser Val Asn Thr Pro Thr Lys Thr Arg Gly Leu Gly  
85 90 95  
Ala Gly Lys Ala Ala Asp Leu Thr Tyr Trp Glu Ser Ala Ala Arg Met  
100 105 110  
Ile Ala Asp Val Ser Val Ser Asp Lys Ile Val Val Glu Lys Ser Thr  
25 115 120 125  
Val Pro Val Lys Thr Ala Glu Ala Ile Glu Lys Ile Leu Thr His Asn  
130 135 140  
Ser Lys Gly Ile Lys Phe Gln Ile Leu Ser Asn Pro Glu Phe Leu Ala  
145 150 155 160  
30 Glu Gly Thr Ala Ile Lys Asp Leu Phe Asn Pro Asp Arg Val Leu Ile  
165 170 175  
Gly Gly Arg Glu Thr Pro Glu Gly Phe Lys Ala Val Gln Thr Leu Lys

|    |                                                                 |     |     |
|----|-----------------------------------------------------------------|-----|-----|
|    | 180                                                             | 185 | 190 |
|    | Asn Val Tyr Ala His Trp Val Pro Glu Gly Gln Ile Ile Thr Thr Asn |     |     |
|    | 195                                                             | 200 | 205 |
|    | Leu Trp Ser Ala Glu Leu Ser Lys Leu Ala Ala Asn Ala Phe Leu Ala |     |     |
| 5  | 210                                                             | 215 | 220 |
|    | Gln Arg Ile Ser Ser Val Asn Ala Met Ser Ala Leu Cys Glu Ala Thr |     |     |
|    | 225                                                             | 230 | 235 |
|    | Gly Ala Asp Val Thr Gln Val Ser Tyr Ala Val Gly Thr Asp Ser Arg |     |     |
|    | 245                                                             | 250 | 255 |
| 10 | Ile Gly Pro Lys Phe Leu Asn Ser Ser Val Gly Phe Gly Gly Ser Cys |     |     |
|    | 260                                                             | 265 | 270 |
|    | Phe Gln Lys Asp Ile Leu Asn Leu Val Tyr Ile Cys Glu Cys Asn Gly |     |     |
|    | 275                                                             | 280 | 285 |
|    | Leu Pro Glu Val Ala Glu Tyr Trp Lys Gln Val Ile Lys Ile Asn Asp |     |     |
| 15 | 290                                                             | 295 | 300 |
|    | Tyr Gln Lys Ser Arg Phe Val Asn Arg Val Val Ser Ser Met Phe Asn |     |     |
|    | 305                                                             | 310 | 315 |
|    | Ser Val Ser Asn Lys Lys Ile Ala Val Leu Gly Phe Ala Phe Lys Lys |     |     |
|    | 325                                                             | 330 | 335 |
| 20 | Asp Thr Gly Asp Thr Arg Glu Thr Pro Ala Ile Asp Val Cys Lys Gly |     |     |
|    | 340                                                             | 345 | 350 |
|    | Leu Leu Glu Asp Lys Ala Arg Leu Ser Ile Tyr Asp Pro Gln Val Thr |     |     |
|    | 355                                                             | 360 | 365 |
|    | Glu Asp Gln Ile Gln Arg Asp Leu Ser Met Asn Lys Phe Asp Trp Asp |     |     |
| 25 | 370                                                             | 375 | 380 |
|    | His Pro Leu His Leu Gln Pro Met Ser Pro Thr Thr Val Lys Gln Val |     |     |
|    | 385                                                             | 390 | 395 |
|    | Thr Val Thr Trp Asp Ala Tyr Glu Ala Thr Lys Asp Ala His Gly Ile |     |     |
|    | 405                                                             | 410 | 415 |
| 30 | Cys Ile Met Thr Glu Trp Asp Glu Phe Lys Asn Leu Asp Phe Gln Lys |     |     |
|    | 420                                                             | 425 | 430 |
|    | Ile Phe Asp Asn Met Gln Lys Pro Ala Phe Val Phe Asp Gly Arg Asn |     |     |

|    |                                                                    |      |     |     |
|----|--------------------------------------------------------------------|------|-----|-----|
|    | 435                                                                | 440  | 445 |     |
|    | Ile Met Asn Leu Gln Lys Leu Arg Glu Ile Gly Phe Ile Val Tyr Ser    |      |     |     |
|    | 450                                                                | 455  | 460 |     |
|    | Ile Gly Lys Pro Leu Asp Asp Trp Leu Lys Asp Met Pro Ala Val Ala    |      |     |     |
| 5  | 465                                                                | 470  | 475 | 480 |
|    | <210> 19                                                           |      |     |     |
|    | <211> 1443                                                         |      |     |     |
|    | <212> DNA                                                          |      |     |     |
|    | <213> <i>Arabidopsis thaliana</i>                                  |      |     |     |
| 10 | <400> 19                                                           |      |     |     |
|    | atggtgaaga tttgctgcat tggagctgga tatgttggtg gtccaacccat ggctgtcatt | 60   |     |     |
|    | gctctaaagt gtccatctgt tgaagtagct gttgttgata tctctgtgcc aaggatcaat  | 120  |     |     |
|    | gcctggaaca gtgatcagtt accgatctat gagcctggtc ttgatgatgt cgtaagcag   | 180  |     |     |
|    | tgccgtggaa agaattctctt cttcagcacc gatgttgaga aacatgtgag agaggctgac | 240  |     |     |
| 15 | attgtttttg tgtctgtcaa caccctact aagaccctg gtcttgagc tggcaaagct     | 300  |     |     |
|    | gcggatttga cttactggga gagcgtgct cgtatgattg ccgatgttcc ggtttccgac   | 360  |     |     |
|    | aagattgttg ttgagaaatc aactgttcct gtcaaaaccg cagaggcaat tgagaagatt  | 420  |     |     |
|    | cttacacaca acagcaaagg aatcaaattc cagattctgt caaacctga gttccttgc    | 480  |     |     |
|    | gaaggaaccg ctattgaaga ctttttcatt cctgaccgtg tctcatcgg tggctcgtgaa  | 540  |     |     |
| 20 | acaactgaag gctttgcagc cgtcaaagcc ttgaaagaca tttatgccca atgggtccct  | 600  |     |     |
|    | gaagagagaa tctcaccac caatctatgg tctgccgagc tttccaagct tgcagctaat   | 660  |     |     |
|    | gccttcctag ccagagaat ctcatcagtc aatgcaatgt ccgctctctg tgaggcaact   | 720  |     |     |
|    | ggcgccaatg tctcagaggt ctcttatgct gtgggcaaag actctcgtat tgggtcccaag | 780  |     |     |
|    | ttcttgaact ctagtgttgg gttcggagga tcttgtttcc agaaagatat tctcaactta  | 840  |     |     |
| 25 | gtctacatct gcgaatgcaa cggcttacct gaagttgctg agtactggaa acaagtcac   | 900  |     |     |
|    | aagatcaacg actaccagaa aaccgattt gttaaccgca ttgtctcttc aatgtttaac   | 960  |     |     |
|    | acagtctcca acaaaaagat tgcggttctc ggcttcgctt tcaagaaaga cactggagac  | 1020 |     |     |
|    | actagagaga ctccagccat tgatgtctgc aaaggtctgt taggtgacaa ggctcgtctc  | 1080 |     |     |
|    | agcatctacg acccacaagt cactgaagag cagatccaaa gagacttaac catgaacaaa  | 1140 |     |     |
| 30 | ttcgactggg accaccact tcatctccag cccatgagcc ccaccactgt gaagcaagtc   | 1200 |     |     |
|    | tcagtcgctt gggacgcata cactgcaacc aaagacgcc acggtatctg cattttaacc   | 1260 |     |     |
|    | gagtgggacg agttcaagaa acttgatttc cagcggatct ttgagaatat gcagaaaccg  | 1320 |     |     |

gcttttggtt ttgacggtag aaacgtggtc gacgctgata aactcagggg gattggggtt 1380  
attgtttact ccattggtaa gccattggac cagtggctca aggacatgcc tgctcttgcc 1440  
taa 1443

<210> 20  
5 <211> 480  
<212> PRT  
<213> *Arabidopsis thaliana*  
<400> 20

Met Val Lys Ile Cys Cys Ile Gly Ala Gly Tyr Val Gly Gly Pro Thr  
10 1 5 10 15  
Met Ala Val Ile Ala Leu Lys Cys Pro Ser Val Glu Val Ala Val Val  
20 25 30  
Asp Ile Ser Val Pro Arg Ile Asn Ala Trp Asn Ser Asp Gln Leu Pro  
35 40 45  
15 Ile Tyr Glu Pro Gly Leu Asp Asp Val Val Lys Gln Cys Arg Gly Lys  
50 55 60  
Asn Leu Phe Phe Ser Thr Asp Val Glu Lys His Val Arg Glu Ala Asp  
65 70 75 80  
Ile Val Phe Val Ser Val Asn Thr Pro Thr Lys Thr Arg Gly Leu Gly  
20 85 90 95  
Ala Gly Lys Ala Ala Asp Leu Thr Tyr Trp Glu Ser Ala Ala Arg Met  
100 105 110  
Ile Ala Asp Val Ser Val Ser Asp Lys Ile Val Val Glu Lys Ser Thr  
115 120 125  
25 Val Pro Val Lys Thr Ala Glu Ala Ile Glu Lys Ile Leu Thr His Asn  
130 135 140  
Ser Lys Gly Ile Lys Phe Gln Ile Leu Ser Asn Pro Glu Phe Leu Ala  
145 150 155 160  
Glu Gly Thr Ala Ile Glu Asp Leu Phe Met Pro Asp Arg Val Leu Ile  
30 165 170 175  
Gly Gly Arg Glu Thr Thr Glu Gly Phe Ala Ala Val Lys Ala Leu Lys  
180 185 190

Asp Ile Tyr Ala Gln Trp Val Pro Glu Glu Arg Ile Leu Thr Thr Asn  
 195 200 205  
 Leu Trp Ser Ala Glu Leu Ser Lys Leu Ala Ala Asn Ala Phe Leu Ala  
 210 215 220  
 5 Gln Arg Ile Ser Ser Val Asn Ala Met Ser Ala Leu Cys Glu Ala Thr  
 225 230 235 240  
 Gly Ala Asn Val Ser Glu Val Ser Tyr Ala Val Gly Lys Asp Ser Arg  
 245 250 255  
 Ile Gly Pro Lys Phe Leu Asn Ser Ser Val Gly Phe Gly Gly Ser Cys  
 10 260 265 270  
 Phe Gln Lys Asp Ile Leu Asn Leu Val Tyr Ile Cys Glu Cys Asn Gly  
 275 280 285  
 Leu Pro Glu Val Ala Glu Tyr Trp Lys Gln Val Ile Lys Ile Asn Asp  
 290 295 300  
 15 Tyr Gln Lys Thr Arg Phe Val Asn Arg Ile Val Ser Ser Met Phe Asn  
 305 310 315 320  
 Thr Val Ser Asn Lys Lys Ile Ala Val Leu Gly Phe Ala Phe Lys Lys  
 325 330 335  
 Asp Thr Gly Asp Thr Arg Glu Thr Pro Ala Ile Asp Val Cys Lys Gly  
 20 340 345 350  
 Leu Leu Gly Asp Lys Ala Arg Leu Ser Ile Tyr Asp Pro Gln Val Thr  
 355 360 365  
 Glu Glu Gln Ile Gln Arg Asp Leu Thr Met Asn Lys Phe Asp Trp Asp  
 370 375 380  
 25 His Pro Leu His Leu Gln Pro Met Ser Pro Thr Thr Val Lys Gln Val  
 385 390 395 400  
 Ser Val Ala Trp Asp Ala Tyr Thr Ala Thr Lys Asp Ala His Gly Ile  
 405 410 415  
 Cys Ile Leu Thr Glu Trp Asp Glu Phe Lys Lys Leu Asp Phe Gln Arg  
 30 420 425 430  
 Ile Phe Glu Asn Met Gln Lys Pro Ala Phe Val Phe Asp Gly Arg Asn  
 435 440 445

Val Val Asp Ala Asp Lys Leu Arg Glu Ile Gly Phe Ile Val Tyr Ser

450

455

460

Ile Gly Lys Pro Leu Asp Gln Trp Leu Lys Asp Met Pro Ala Leu Ala

465

470

475

480

5

<210> 21

<211> 1446

<212> DNA

<213> *Arabidopsis thaliana*

<400> 21

10

atggtgaaga tatgctgcat aggagctggt tatgtgggtg gtccaaccat ggcggtgatg 60

gctcttaagt gtcctgagat tgaagtagtc gttgtggata tctctgaacc aaggatcaat 120

gcttggaaca gtgataggct tcctatttac gagccgggat tggaagatgt ggtgaaacaa 180

tgcagagggg aaaacctctt ctttagcaca gacgtggaga aacatgtatt tgagagtgat 240

attgtatttg tctcagttaa cactccaacc aaaacacaag gtcttggtgc tggcaaagct 300

15

gctgatctta cttactggga gagtgctgct cggatgatcg ctgatgtctc caaatctagc 360

aaaatcgttg ttgagaaatc cacggttcct gtgaggacag cagaggctat tgaaaagata 420

ctgacacata acagcaaagg catagagttt cagattctct ctaaccctga atttcttgct 480

gaggtactg caattaagga tctttataac ccagaccgtg tgttgattgg tggtagggat 540

actgcagcag ggcaaaaggc tattaaagct ttaagagatg tttatgctca ttgggttcca 600

20

gtggaacaaa tcatttgac gaacctgtgg tccgctgagc tctctaagct tgcagcaaat 660

gcattcttag ctcagaggat atcatctgtc aatgccatgt cagctctatg tgaggcaact 720

ggcgtgatg ttacacaagt tgcgcatgcc gtgggtacag atactagaat tgggtccaaag 780

ttcttgaatg ctagtgttgg ttttggtgga tcatgtttcc aaaaggacat cctaaatctt 840

atctatattt gtgaatgcaa cggttgccc gaagcagcta attactggaa acaagtcgta 900

25

aaggtgaacg actatcagaa aatacggttt gcaaaccggg ttgtttcttc aatgtttaac 960

acagtctcgg gcaagaaaat cgcgatcctc ggttttgcct tcaagaagga cacagggtgac 1020

acgagagaga ctccagcgat tgatgtttgt aacagattag ttgcagacaa ggccaagctg 1080

agcatatacg acccacaagt tcttgaagaa cagatcagaa gagatctttc catggctagg 1140

tttgactggg accaccctgt tcctcttcag cagattaaag ctgaaggat ctgagagcaa 1200

30

gtgaatgtcg tctcagatgc ttacaggca actaaagatg cgcacggcct atgtgtctta 1260

accgaatggg atgagtttaa atccttggac ttcaagaaaa tctttgacaa tatgcagaaa 1320

ccagcttttg tgttcgatgg taggaatggt gttgatgcag tgaagctcgg tgagatcggg 1380

```

ttcatcgtct actccattgg taaaccgctt gattcatggc tcaaggatat gcctgctgtg 1440
gcatga 1446

<210> 22
<211> 481
5 <212> PRT
<213> Arabidopsis thaliana
<400> 22

Met Val Lys Ile Cys Cys Ile Gly Ala Gly Tyr Val Gly Gly Pro Thr
1 5 10 15
10 Met Ala Val Met Ala Leu Lys Cys Pro Glu Ile Glu Val Val Val Val
20 25 30
Asp Ile Ser Glu Pro Arg Ile Asn Ala Trp Asn Ser Asp Arg Leu Pro
35 40 45
Ile Tyr Glu Pro Gly Leu Glu Asp Val Val Lys Gln Cys Arg Gly Lys
15 50 55 60
Asn Leu Phe Phe Ser Thr Asp Val Glu Lys His Val Phe Glu Ser Asp
65 70 75 80
Ile Val Phe Val Ser Val Asn Thr Pro Thr Lys Thr Gln Gly Leu Gly
85 90 95
20 Ala Gly Lys Ala Ala Asp Leu Thr Tyr Trp Glu Ser Ala Ala Arg Met
100 105 110
Ile Ala Asp Val Ser Lys Ser Ser Lys Ile Val Val Glu Lys Ser Thr
115 120 125
Val Pro Val Arg Thr Ala Glu Ala Ile Glu Lys Ile Leu Thr His Asn
25 130 135 140
Ser Lys Gly Ile Glu Phe Gln Ile Leu Ser Asn Pro Glu Phe Leu Ala
145 150 155 160
Glu Gly Thr Ala Ile Lys Asp Leu Tyr Asn Pro Asp Arg Val Leu Ile
165 170 175
30 Gly Gly Arg Asp Thr Ala Ala Gly Gln Lys Ala Ile Lys Ala Leu Arg
180 185 190
Asp Val Tyr Ala His Trp Val Pro Val Glu Gln Ile Ile Cys Thr Asn

```

|    |                                                                 |     |     |
|----|-----------------------------------------------------------------|-----|-----|
|    | 195                                                             | 200 | 205 |
|    | Leu Trp Ser Ala Glu Leu Ser Lys Leu Ala Ala Asn Ala Phe Leu Ala |     |     |
|    | 210                                                             | 215 | 220 |
|    | Gln Arg Ile Ser Ser Val Asn Ala Met Ser Ala Leu Cys Glu Ala Thr |     |     |
| 5  | 225                                                             | 230 | 235 |
|    | Gly Ala Asp Val Thr Gln Val Ala His Ala Val Gly Thr Asp Thr Arg |     | 240 |
|    | 245                                                             | 250 | 255 |
|    | Ile Gly Pro Lys Phe Leu Asn Ala Ser Val Gly Phe Gly Gly Ser Cys |     |     |
|    | 260                                                             | 265 | 270 |
| 10 | Phe Gln Lys Asp Ile Leu Asn Leu Ile Tyr Ile Cys Glu Cys Asn Gly |     |     |
|    | 275                                                             | 280 | 285 |
|    | Leu Pro Glu Ala Ala Asn Tyr Trp Lys Gln Val Val Lys Val Asn Asp |     |     |
|    | 290                                                             | 295 | 300 |
|    | Tyr Gln Lys Ile Arg Phe Ala Asn Arg Val Val Ser Ser Met Phe Asn |     |     |
| 15 | 305                                                             | 310 | 315 |
|    | Thr Val Ser Gly Lys Lys Ile Ala Ile Leu Gly Phe Ala Phe Lys Lys |     | 320 |
|    | 325                                                             | 330 | 335 |
|    | Asp Thr Gly Asp Thr Arg Glu Thr Pro Ala Ile Asp Val Cys Asn Arg |     |     |
|    | 340                                                             | 345 | 350 |
| 20 | Leu Val Ala Asp Lys Ala Lys Leu Ser Ile Tyr Asp Pro Gln Val Leu |     |     |
|    | 355                                                             | 360 | 365 |
|    | Glu Glu Gln Ile Arg Arg Asp Leu Ser Met Ala Arg Phe Asp Trp Asp |     |     |
|    | 370                                                             | 375 | 380 |
|    | His Pro Val Pro Leu Gln Gln Ile Lys Ala Glu Gly Ile Ser Glu Gln |     |     |
| 25 | 385                                                             | 390 | 395 |
|    | Val Asn Val Val Ser Asp Ala Tyr Glu Ala Thr Lys Asp Ala His Gly |     | 400 |
|    | 405                                                             | 410 | 415 |
|    | Leu Cys Val Leu Thr Glu Trp Asp Glu Phe Lys Ser Leu Asp Phe Lys |     |     |
|    | 420                                                             | 425 | 430 |
| 30 | Lys Ile Phe Asp Asn Met Gln Lys Pro Ala Phe Val Phe Asp Gly Arg |     |     |
|    | 435                                                             | 440 | 445 |
|    | Asn Val Val Asp Ala Val Lys Leu Arg Glu Ile Gly Phe Ile Val Tyr |     |     |

450                      455                      460  
 Ser Ile Gly Lys Pro Leu Asp Ser Trp Leu Lys Asp Met Pro Ala Val

465                      470                      475                      480

Ala

5 <210> 23

<211> 32

<212> DNA

<213> Sequência Artificial

<220>

10 <223> The sequence of designed polynucleotide described in Example 1

<400> 23

gggaattcgt gaagatctgt tgtattggag ct 32

<210> 24

<211> 33

15 <212> DNA

<213> Sequência Artificial

<220>

<223> Sequência do Polinucleotídeo Desenvolvido de acordo com o Exemplo 1

<400> 24

20 cagaagcttt tagacaaagg caggcatgtc ctt 33

<210> 25

<211> 30

<212> DNA

<213> Sequência Artificial

25 <220>

<223> Sequência do Polinucleotídeo Desenvolvido de acordo com o Exemplo 1

<400> 25

ggaattcgtg aagatatgtt gtattggagc 30

<210> 26

30 <211> 31

<212> DNA

<213> Sequência Artificial

- <220>
- <223> Sequência do Polinucleotídeo Desenvolvido de acordo com o Exemplo 1
- <400> 26
- aactgcagtt aggcaacggc aggcatgtcc t 31
- 5 <210> 27
- <211> 30
- <212> DNA
- <213> Sequência Artificial
- <220>
- 10 <223> Sequência do Polinucleotídeo Desenvolvido de acordo com o Exemplo 1
- <400> 27
- ggaattcgtg aagatttgct gcattggagc 30
- <210> 28
- <211> 31
- 15 <212> DNA
- <213> Sequência Artificial
- <220>
- <223> Sequência do Polinucleotídeo Desenvolvido de acordo com o Exemplo 1
- <400> 28
- 20 aactgcagtt aggcaagagc aggcatgtcc t 31
- <210> 29
- <211> 30
- <212> DNA
- <213> Sequência Artificial
- 25 <220>
- <223> Sequência do Polinucleotídeo Desenvolvido de acordo com o Exemplo 1
- <400> 29
- ggaattcgtg aagatatgct gcataggagc 30
- <210> 30
- 30 <211> 31
- <212> DNA
- <213> Sequência Artificial

<220>

<223> Sequência do Polinucleotídeo Desenvolvido de acordo com o Exemplo 1

<400> 30

gatctagatc atgccacagc aggcatatcc t 31

5 <210> 31

<211> 30

<212> DNA

<213> Sequência Artificial

<220>

10 <223> Sequência do Polinucleotídeo Desenvolvido de acordo com o Exemplo 2

<400> 31

ggaattctca cgaatcgagc tcgttggttg 30

<210> 32

<211> 32

15 <212> DNA

<213> Sequência Artificial

<220>

<223> Sequência do Polinucleotídeo Desenvolvido de acordo com o Exemplo 2

<400> 32

20 gactgcagtt agttatcgcc atatacatca cg 32

<210> 33

<211> 40

<212> DNA

<213> Sequência Artificial

25 <220>

<223> Sequência do Polinucleotídeo Desenvolvido de acordo com o Exemplo 3

<400> 33

gagagtcgac ctattgagta gtcacactct ttgccagatt 40

<210> 34

30 <211> 35

<212> DNA

<213> Sequência Artificial

- <220>
- <223> Sequência do Polinucleotídeo Desenvolvido de acordo com o Exemplo 3
- <400> 34
- atgtgtggaa tcttcgcgta tctgaatttt cacgc 35
- 5 <210> 35
- <211> 29
- <212> DNA
- <213> Sequência Artificial
- <220>
- 10 <223> Sequência do Polinucleotídeo Desenvolvido de acordo com o Exemplo 4
- <400> 35
- tctgtacgat gcaactacca atgctcagt 29
- <210> 36
- <211> 30
- 15 <212> DNA
- <213> Sequência Artificial
- <220>
- <223> Sequência do Polinucleotídeo Desenvolvido de acordo com o Exemplo 4
- <400> 36
- 20 tatcttacct ggggtcaaag acgaacataa 30
- <210> 37
- <211> 33
- <212> DNA
- <213> Sequência Artificial
- 25 <220>
- <223> Sequência do Polinucleotídeo Desenvolvido de acordo com o Exemplo 4
- <400> 37
- atgtgtggca tctttggagc actgtcaaac aac 33
- <210> 38
- 30 <211> 39
- <212> DNA
- <213> Sequência Artificial

<220>

<223> Sequência do Polinucleotídeo Desenvolvido de acordo com o Exemplo 4

<400> 38

aactgcagtt aaaaggtggt cacggatttt gcaagattc 39

5 <210> 39

<211> 39

<212> DNA

<213> Sequência Artificial

<220>

10 <223> Sequência do Polinucleotídeo Desenvolvido de acordo com o Exemplo 1

<400> 39

aactgcagtt aaaaggtggt cacagatttc gcaagattc 39

<210> 40

<211> 30

15 <212> DNA

<213> Sequência Artificial

<220>

<223> Sequência do Polinucleotídeo Desenvolvido de acordo com o Exemplo 5

<400> 40

20 ccggtatccat gggtaaaaat ataatacataa 30

<210> 41

<211> 30

<212> DNA

<213> Sequência Artificial

25 <220>

<223> Sequência do Polinucleotídeo Desenvolvido de acordo com o Exemplo 5

<400> 41

tatatattaaa tcacacagac tgagcattgg 30

<210> 42

30 <211> 30

<212> DNA

<213> Sequência Artificial

- <220>
- <223> Sequência do Polinucleotídeo Desenvolvido de acordo com o Exemplo 6
- <400> 42
- aaggatccga tgtgtggcat ctttgagca 30
- 5 <210> 43
- <211> 33
- <212> DNA
- <213> Sequência Artificial
- <220>
- 10 <223> Sequência do Polinucleotídeo Desenvolvido de acordo com o Exemplo 13
- <400> 43
- aaggatccat gtcacgaatc gcagtcgttg gtt 33
- <210> 44
- <211> 35
- 15 <212> DNA
- <213> Sequência Artificial
- <220>
- <223> Sequência do Polinucleotídeo Desenvolvido de acordo com o Exemplo 13
- <400> 44
- 20 ccgagctctt agttatcgcc atatacatca cgtgt 35

## REIVINDICAÇÕES

1. Método para produzir ácido hialurônico, compreendendo co-expressar uma proteína com atividade de ácido hialurônico sintase e uma proteína exógena com atividade de nucleotídeo de açúcar sintase em uma  
5 célula vegetal ou uma planta e uma proteína exógena tendo nucleotídeo de açúcar sintase atividade em uma célula vegetal ou uma planta.

2. Método para produzir ácido hialurônico, contendo as etapas de:

- 10 (1) transformar uma célula vegetal ou uma planta usando um vetor de expressão recombinante, o vetor de expressão recombinante tendo DNA codificando uma proteína com atividade de ácido hialurônico sintase e DNA codificando uma proteína com atividade de nucleotídeo de açúcar sintase, cada DNA referido estando sob controle de um promotor vegetal;
- 15 (2) cultivar um transformante obtido pela transformação; e
- (3) isolar ácido hialurônico produzido pelo transformante.

3. Método para produzir ácido hialurônico de acordo com a reivindicação 2, em que o promotor é um promotor órgão-específico ou um tecido-específico.

20 4. Método para produzir ácido hialurônico de acordo com a reivindicação 2 ou 3, em que o DNA codificando uma proteína com atividade de ácido hialurônico sintase é DNA de (a) ou (b) abaixo:

- (a) DNA tendo uma sequência de nucleotídeos representada pela SEQ ID NO: 1 o 3; ou
- 25 (b) DNA hibridizando de modo complementar à sequência de nucleotídeos de (a) sob condições estridentes, e o DNA codificando uma proteína com atividade de ácido hialurônico sintase.

5. Método para produzir ácido hialurônico de acordo com qualquer uma das reivindicações 1 a 3, em que a proteína com atividade de ácido hialurônico sintase é uma proteína de (a) ou (b):

30

- (a) uma proteína tendo uma sequência de aminoácidos mostrada pela SEQ ID NO: 2 ou 4; ou

(b) uma proteína tendo a sequência de aminoácidos de (a) com um ou uns poucos aminoácidos eliminados, substituídos ou adicionados, e a proteína tendo atividade de ácido hialurônico sintase.

6. Método para produzir ácido hialurônico de acordo com qualquer uma das reivindicações 1 a 5, em que o nucleotídeo de açúcar é uridin-5'-difosfo(UDP)-N-acetilglucosamina ou ácido UDP-glucurônico.

7. Método para produzir ácido hialurônico de acordo com qualquer uma das reivindicações 1 a 6, em que a proteína com atividade de nucleotídeo de açúcar sintase é no mínimo uma proteína selecionada entre o grupo consistindo em glutamina:frutose-6-fosfato amidotransferase, UDP-N-acetilglucosamina difosforilase, fosfoacetil glucosamina mutase, glucosamina-6-fosfato N-acetiltransferase, glucosamina-1-fosfato N-acetiltransferase, fosfoglucomutase, N-acetilglucosamina quinase, hexoquinase, N-acetilglucosamina acetilase, UDP-glucose desidrogenase, UDP-glucose-1-fosfato uridiltransferase, inositol oxigenase, glucuronoquinase e glucuronato-1-fosfato uridil transferase.

8. Método para produzir ácido hialurônico de acordo com qualquer uma das reivindicações 1 a 6, em que a proteína com atividade de nucleotídeo de açúcar sintase é no mínimo uma proteína selecionada entre o grupo consistindo em UDP-N-acetilglucosamina difosforilase, fosfoacetil glucosamina mutase, glucosamina-6-fosfato N-acetiltransferase, glucosamina-1-fosfato N-acetiltransferase, fosfoglucomutase, N-acetilglucosamina quinase, hexoquinase, N-acetilglucosamina acetilase, UDP-glucose desidrogenase, UDP-glucose-1-fosfato uridiltransferase, inositol oxigenase, glucuronoquinase e glucuronato-1-fosfato uridil transferase, e glutamina:frutose-6-fosfato amidotransferase.

9. Método para produzir ácido hialurônico de acordo com qualquer uma das reivindicações 1 a 6, em que uma proteína com atividade de nucleotídeo de açúcar sintase é glutamina:frutose-6-fosfato amidotransferase ou UDP-glucose desidrogenase.

10. Método para produzir ácido hialurônico de acordo com qualquer uma das reivindicações 2 a 9, em que DNA codificando uma proteína

com nucleotídeo de açúcar sintase atividade é DNA derivado de vírus chlorella ou Arabidopsis.

11. Método para produzir ácido hialurônico de acordo com qualquer uma das reivindicações 2 a 10, em que DNA codificando uma proteína com atividade de nucleotídeo de açúcar sintase é DNA de (a) ou (b) abaixo:

- (a) DNA tendo uma sequência de nucleotídeos mostrada pela SEQ ID NO: 5, 7 ou 9; ou
- (b) DNA hibridizando de modo complementar à sequência de nucleotídeos de (a) sob condições estridentes, e o referido DNA codificando uma proteína com atividade de glutamina:frutose-6-fosfato amidotransferase.

12. Método para produzir ácido hialurônico de acordo com qualquer uma das reivindicações 1 a 10, em que a proteína com atividade de nucleotídeo de açúcar sintase é uma proteína de (a) ou (b) abaixo:

- (a) uma proteína tendo uma sequência de aminoácidos mostrada pela SEQ ID NO: 6, 8 ou 10; ou
- (b) uma proteína tendo a sequência de aminoácidos de (a) com um ou uns poucos aminoácidos eliminados, substituídos ou adicionados, e a proteína tendo atividade de glutamina:frutose-6-fosfato amidotransferase.

13. Método para produzir ácido hialurônico de acordo com qualquer uma das reivindicações 2 a 10, em que DNA codificando uma proteína com atividade de nucleotídeo de açúcar sintase é DNA de (a) ou (b) abaixo:

- (a) DNA tendo uma sequência de nucleotídeos mostrada pela SEQ ID NO: 11, 13, 17, 19, ou 21; ou
- (b) DNA hibridizando de modo complementar à sequência de nucleotídeos de (a) sob condições estridentes, e o DNA codificando uma proteína com atividade de UDP-glucose desidrogenase.

14. Método para produzir ácido hialurônico de acordo com qualquer uma das reivindicações 1 a 10, em que a proteína com atividade de nucleotídeo de açúcar sintase é uma proteína de (a) ou (b) abaixo:

- (a) uma proteína tendo uma sequência de aminoácidos mostrada pela SEQ ID NO: 12, 14, 16, 18, 20 ou 22; ou

- (b) uma proteína tendo a sequência de aminoácidos de (a) com um ou uns poucos aminoácidos eliminados, substituídos ou adicionados, e a proteína tendo atividade de UDP-glucose desidrogenase.

15 15. Método para produzir ácido hialurônico de acordo com qualquer uma das reivindicações 1 a 14, em que uma planta é selecionada entre o grupo consistindo em angiospermas, gimnospermas, pteridófitos e briófitas.

10 16. Método para produzir ácido hialurônico de acordo com a reivindicação 3, em que um ou mais órgãos são selecionados entre o grupo consistindo em raízes, caules, tuberosidades de caules, folha, órgãos florais, raízes tuberosas, sementes e ápices dos brotos.

17. Método para produzir ácido hialurônico de acordo com a reivindicação 3, em que um ou mais tecidos são selecionados entre o grupo consistindo em epiderme, floema, tecidos moles, xilema e feixes vasculares.

15 18. Célula de planta transgênica ou uma planta transgênica tendo a capacidade de produzir ácido hialurônico por coexpressão de uma proteína com atividade de ácido hialurônico sintase e uma proteína exógena com atividade de nucleotídeo de açúcar sintase, uma progênie da mesma, ou um órgão ou um tecido da mesma.

20 19. Célula de planta transgênica ou uma planta transgênica sendo transformada com um vetor de expressão recombinante contendo DNA codificando uma proteína com atividade de ácido hialurônico sintase e DNA codificando uma proteína com atividade de nucleotídeo de açúcar sintase, cada DNA referido estando sob controle de um promotor vegetal; uma progênie tendo a mesma natureza das mesmas; ou um órgão ou um tecido das mesmas.

25 20. Célula de planta transgênica ou a planta transgênica, a progênie tendo a mesma natureza das mesmas, ou o órgão ou o tecido das mesmas de acordo com a reivindicação 19, em que o promotor é um promotor órgão-específico ou um tecido-específico.

30 21. Célula de planta transgênica ou a planta transgênica, a progênie tendo a mesma natureza das mesmas, ou o órgão ou o tecido das

mesmas de acordo com a reivindicação 19 ou 20, em que o DNA codificando uma proteína com atividade de ácido hialurônico sintase é DNA de (a) ou (b) abaixo:

- 5 (a) DNA tendo uma sequência de nucleotídeos mostrada na SEQ ID NO: 1 ou 3; ou
- (b) DNA hibridizando de modo complementar à DNA tendo a sequência de nucleotídeos de (a) sob condições estridentes, e o DNA codificando uma proteína com atividade de ácido hialurônico sintase.

22. Célula de planta transgênica ou a planta transgênica de acordo com qualquer uma das reivindicações 18 a 20, em que a proteína com atividade de ácido hialurônico sintase é uma proteína de (a) ou (b) abaixo:

- (a) uma proteína tendo uma sequência de aminoácidos mostrada pela SEQ ID NO: 2 ou 4; ou
- 15 (b) uma proteína tendo a sequência de aminoácidos de (a) com um ou uns poucos aminoácidos eliminados, substituídos ou adicionados, e a proteína tendo atividade de ácido hialurônico sintase.

23. Célula de planta transgênica ou a planta transgênica, a pro-  
gênie tendo a mesma natureza das mesmas, ou o órgão ou o tecido das  
mesmas de acordo com qualquer uma das reivindicações 18 a 22, em que o  
20 nucleotídeo de açúcar é UDP-N-acetilglucosamina ou ácido UDP-  
glucurônico.

24. Célula de planta transgênica ou a planta transgênica de a-  
cordo com qualquer uma das reivindicações 18 a 23, em que a proteína com  
atividade de nucleotídeo de açúcar sintase é no mínimo uma proteína sele-  
25 cionada entre o grupo consistindo em, glutamina:frutose-6-fosfato amido-  
transferase, UDP-N-acetilglucosamina difosforilase, fosfoacetil glucosamina  
mutase, glucosamina-6-fosfato N-acetiltransferase, glucosamina-1-fosfato N-  
acetiltransferase, fosfoglucomutase, N-acetilglucosamina quinase, hexoqui-  
nase, N-acetilglucosamina acetilase, UDP-glucose desidrogenase, UDP-  
30 glucose-1-fosfato uridiltransferase, inositol oxigenase, glucuronoquinase e  
glucuronato-1-fosfato uridil transferase.

25. Célula de planta transgênica ou a planta transgênica, a pro-

gênie tendo a mesma natureza das mesmas, ou o órgão ou o tecido das mesmas de acordo com qualquer uma das reivindicações 18 a 23, em que a proteína com atividade de nucleotídeo de açúcar sintase é no mínimo uma proteína selecionada entre o grupo consistindo em UDP-N-acetilglucosamina difosforilase, fosfoacetil glucosamina mutase, glucosamina-6-fosfato N-acetiltransferase, glucosamina-1-fosfato N-acetiltransferase, fosfoglucomutase, N-acetilglucosamina quinase, hexoquinase, N-acetilglucosamina acetilase, UDP-glucose desidrogenase, UDP-glucose-1-fosfato uridiltransferase, inositol oxigenase, glucuronoquinase e glucuronato-1-fosfato uridil transferase, e glutamina:frutose-6-fosfato amidotransferase.

26. Célula de planta transgênica ou a planta transgênica, a progênie tendo a mesma natureza das mesmas, ou o órgão ou o tecido das mesmas de acordo com qualquer uma das reivindicações 18 a 23, em que a proteína com atividade de nucleotídeo de açúcar sintase é glutamina:frutose-6-fosfato amidotransferase e/ou UDP-glucose desidrogenase.

27. Célula de planta transgênica ou a planta transgênica, a progênie tendo a mesma natureza das mesmas, ou o órgão ou o tecido das mesmas de acordo com qualquer uma das reivindicações 19 a 26, em que DNA codificando uma proteína com nucleotídeo de açúcar sintase atividade é derivado de chlorella vírus, Arabidopsis, ou vírus chlorella e Arabidopsis.

28. Célula de planta transgênica ou a planta transgênica de acordo com qualquer uma das reivindicações 19 a 27, em que DNA codificando uma proteína com atividade de nucleotídeo de açúcar sintase é DNA de (a) ou (b) abaixo:

- (a) DNA tendo uma sequência de nucleotídeos mostrada pela SEQ ID NO: 5, 7 ou 9; ou
- (b) DNA hibridizando de modo complementar à sequência de nucleotídeos de (a) sob condições estridentes, e o DNA codificando uma proteína com atividade de glutamina:frutose-6-fosfato amidotransferase;
- a progênie tendo a mesma natureza das mesmas; ou o órgão ou o tecido das mesmas.

29. Célula de planta transgênica ou a planta transgênica, a pro-

gênie tendo a mesma natureza das mesmas, ou o órgão ou o tecido das mesmas de acordo com qualquer uma das reivindicações 18 a 27, em que a proteína com atividade de nucleotídeo de açúcar sintase é uma proteína de (a) ou (b) abaixo:

- 5 (a) uma proteína tendo uma sequência de aminoácidos mostrada pela SEQ ID NO: 6, 8 ou 10; ou
- (b) uma proteína tendo a sequência de aminoácidos de (a) com um ou uns poucos aminoácidos eliminados, substituídos ou adicionados, e a proteína tendo atividade de glutamina:frutose-6-fosfato amidotransferase.

10 30. Célula de planta transgênica ou a planta transgênica, a pro-gênie tendo a mesma natureza das mesmas, ou o órgão ou o tecido das mesmas de acordo com qualquer uma das reivindicações 19 a 27, em que DNA codificando uma proteína com atividade de nucleotídeo de açúcar sintase é DNA de (a) ou (b) abaixo:

- 15 (a) DNA tendo uma sequência de nucleotídeos mostrada pela SEQ ID NO: 11, 13, 15, 17, 19, ou 21; ou
- (b) DNA hibridizando de modo complementar à sequência de nucleotídeos de (a) sob condições estridentes, e o DNA codificando uma proteína com atividade de DP-glucose desidrogenase.

20 31. Célula de planta transgênica ou a planta transgênica, a pro-gênie tendo a mesma natureza das mesmas, ou o órgão ou o tecido das mesmas de acordo com qualquer uma das reivindicações 18 a 27, em que a proteína com atividade de nucleotídeo de açúcar sintase é uma proteína de (a) ou (b) abaixo:

- 25 (a) uma proteína tendo uma sequência de aminoácidos mostrada pela SEQ ID NO: 12, 14, 16, 18, 20 ou 22; ou
- (b) uma proteína tendo a sequência de aminoácidos de (a) com um ou uns poucos aminoácidos eliminados, substituídos ou adicionados, e a referida proteína com atividade de UDP-glucose desidrogenase.

30 32. Célula de planta transgênica ou a planta transgênica, a pro-gênie tendo a mesma natureza das mesmas, ou o órgão ou o tecido das mesmas de acordo com qualquer uma das reivindicações 18 a 31, em que a

planta é qualquer planta selecionada entre o grupo consistindo em gimnospermas, pteridófitos e briófitas.

33. Célula de planta transgênica ou a planta transgênica, a progênie tendo a mesma natureza das mesmas; ou o órgão ou o tecido das  
5 mesmas de acordo com qualquer uma das reivindicações 18 a 31, em que o órgão sendo um ou mais órgãos selecionados entre o grupo consistindo em raízes, caules, tuberosidades de caules, folhas, órgãos florais, raízes tuberosas, sementes e ápices dos brotos.

34. Célula de planta transgênica ou a planta transgênica, a progênie tendo a mesma natureza das mesmas, ou o órgão ou o tecido das  
10 mesmas de acordo com qualquer uma das reivindicações 18 a 31, em que o tecido é um ou mais tecidos selecionados entre o grupo consistindo em epiderme, floema, tecidos moles, xilema e feixes vasculares.

35. Extrato vegetal obtido da célula de planta transgênica ou da  
15 planta transgênica, da progênie tendo a mesma natureza das mesmas, ou do órgão ou do tecido das mesmas como definidos em qualquer uma das reivindicações 18 a 34.

36. Extrato vegetal de acordo com a reivindicação 35, em que o extrato vegetal contém ácido hialurônico.

20 37. Vetor de expressão recombinante, compreendendo DNA codificando uma proteína com atividade de ácido hialurônico sintase e DNA codificando uma proteína com atividade de nucleotídeo de açúcar sintase, cada DNA referido estando sob controle de um promotor vegetal.

38. Vetor de expressão recombinante de acordo com a reivindicação 37, em que o promotor é um promotor órgão-específico ou um tecido-específico.  
25

39. Vetor de expressão recombinante de acordo com a reivindicação 37 ou 38, em que o DNA codificando uma proteína com atividade de ácido hialurônico sintase é DNA de (a) ou (b) abaixo:

- 30 (a) DNA tendo uma sequência de nucleotídeos mostrada na SEQ ID NO: 1 ou 3; ou  
(b) DNA hibridizando de modo complementar a sequência de nucleotídeos

de (a) sob condições estridentes, e o DNA codificando uma proteína com ácido hialurônico atividade.

40. Vetor de expressão recombinante de acordo com a reivindicação 37 ou 38, em que a proteína com atividade de ácido hialurônico sintase é uma proteína de (a) ou (b) abaixo:

(a) uma proteína tendo uma sequência de aminoácidos mostrada pela SEQ ID NO: 2 ou 4; ou

(b) uma proteína tendo a sequência de aminoácidos de (a) com um ou uns poucos aminoácidos eliminados, substituídos ou adicionados, e a proteína tendo atividade de ácido hialurônico sintase.

41. Vetor de expressão recombinante de acordo com qualquer uma das reivindicações 37 a 40, em que o nucleotídeo de açúcar é UDP-N-acetilglucosamina ou ácido UDP-glucurônico.

42. Vetor de expressão recombinante de acordo com qualquer uma das reivindicações 37 a 41, em que a proteína com atividade de nucleotídeo de açúcar sintase é no mínimo uma proteína selecionada entre o grupo consistindo em glutamina:frutose-6-fosfato amidotransferase, UDP-N-acetilglucosamina difosforilase, fosfoacetil glucosamina mutase, glucosamina-6-fosfato N-acetiltransferase, glucosamina-1-fosfato N-acetiltransferase, fosfoglucomutase, N-acetilglucosamina quinase, hexoquinase, N-acetilglucosamina acetilase, UDP-glucose desidrogenase, UDP-glucose-1-fosfato uridiltransferase, inositol oxigenase, glucuronoquinase e glucuronato-1-fosfato uridil transferase.

43. Vetor de expressão recombinante de acordo com qualquer uma das reivindicações 37 a 41, em que a proteína com atividade de nucleotídeo de açúcar sintase é no mínimo uma proteína selecionada entre o grupo consistindo em UDP-N-acetilglucosamina difosforilase, fosfoacetil glucosamina mutase, glucosamina-6-fosfato N-acetiltransferase, glucosamina-1-fosfato N-acetiltransferase, fosfoglucomutase, N-acetilglucosamina quinase, hexoquinase, N-acetilglucosamina acetilase, UDP-glucose desidrogenase, UDP-glucose-1-fosfato uridiltransferase, inositol oxigenase, glucuronoquinase e glucuronato-1-fosfato uridil transferase, e glutamina:frutose-6-fosfato

amidotransferase.

44. Vetor de expressão recombinante de acordo com qualquer uma das reivindicações 37 e 41, em que a proteína com atividade de nucleotídeo de açúcar sintase é glutamina:frutose-6-fosfato amidotransferase ou  
5 UDP-glucose desidrogenase.

45. Vetor de expressão recombinante de acordo com qualquer uma das reivindicações 37 a 41, em que o DNA codificando uma proteína com nucleotídeo de açúcar sintase atividade é DNA derivado de chlorella vírus, Arabidopsis, ou vírus chlorella e Arabidopsis.

- 10 46. Vetor de expressão recombinante de acordo com qualquer uma das reivindicações 37 a 44, em que o DNA codificando uma proteína com nucleotídeo de açúcar sintase atividade é DNA de (a) ou (b) abaixo:

- (a) DNA tendo uma sequência de nucleotídeos mostrada pela SEQ ID NO: 5, 7 ou 9; ou  
15 (b) DNA hibridizando de modo complementar à sequência de nucleotídeos de (a) sob condições estridentes, e o referido DNA codificando uma proteína com atividade de glutamina:frutose-6-fosfato amidotransferase.

- 20 47. Vetor de expressão recombinante de acordo com qualquer uma das reivindicações 37 a 44, em que a proteína com atividade de nucleotídeo de açúcar sintase é uma proteína de (a) ou (b) abaixo:

- (a) uma proteína tendo uma sequência de aminoácidos mostrada pela SEQ ID NO: 6, 8 ou 10; ou  
25 (b) uma proteína tendo a sequência de aminoácidos de (a) com um ou uns poucos aminoácidos eliminados, substituídos ou adicionados, e a proteína tendo atividade de glutamina:frutose-6-fosfato amidotransferase.

48. Vetor de expressão recombinante de acordo com qualquer uma das reivindicações 37 a 44, em que o DNA codificando uma proteína com nucleotídeo de açúcar sintase atividade é DNA de (a) ou (b) abaixo:

- 30 (a) DNA tendo uma sequência de nucleotídeos mostrada pela SEQ ID NO: 11, 13, 15, 17, 19, ou 21; ou  
(b) DNA hibridizando de modo complementar à sequência de nucleotídeos

de (a) sob condições estridentes, e o referido DNA codificando uma proteína com atividade de UDP-glucose desidrogenase.

49. Vetor de expressão recombinante de acordo com qualquer uma das reivindicações 37 a 44, em que a proteína com atividade de nucleotídeo de açúcar sintase é uma proteína de (a) ou (b) abaixo:

- (a) uma proteína tendo uma sequência de aminoácidos mostrada pela SEQ ID NO: 12, 14, 16, 18, 20 ou 22; ou
- (b) uma proteína tendo a sequência de aminoácidos de (a) com um ou uns poucos aminoácidos eliminados, substituídos ou adicionados, e a referida proteína tendo atividade de UDP-glucose desidrogenase.

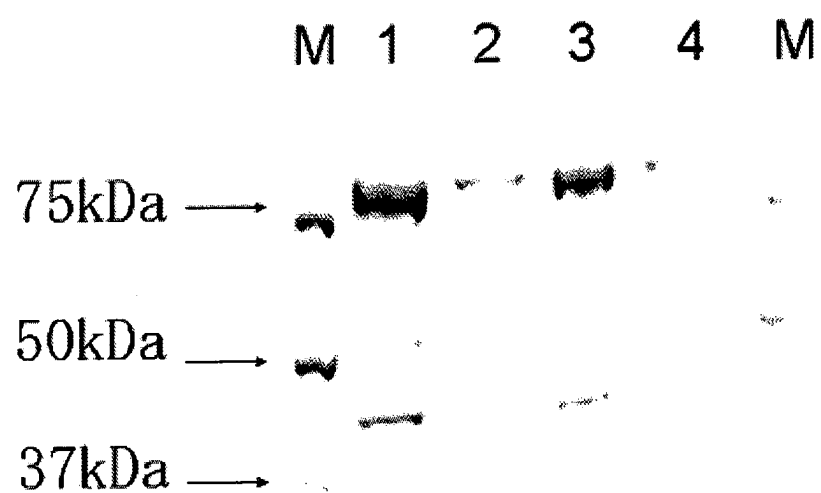
50. Método para gerar a célula de planta transgênica ou a planta transgênica compreendendo transformar uma célula vegetal ou uma planta usando qualquer vetor como definido nas reivindicações 37 a 49, em que a célula de planta transgênica ou a planta transgênica têm a capacidade de produzir ácido hialurônico.

51. Composição cosmética contendo ácido hialurônico como um agente ativo, em que o ácido hialurônico é obtido por qualquer um dos métodos para produzir ácido hialurônico como definido nas reivindicações 1 a 17.

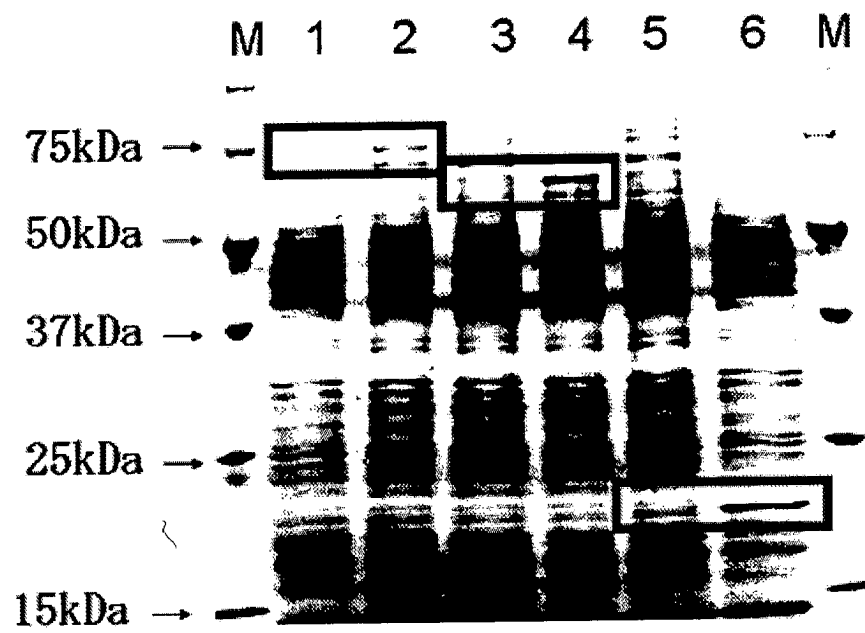
**Fig.1**

W: Padrões de peso molecular

- 1: AtUGD4 (Extrato bruto)
- 2: AtUGD4 (Proteína purificada)
- 3: AtUGD2 (Extrato bruto)
- 4: AtUGD2 (Proteína purificada)
- 5: AtUGD3 (Extrato bruto)
- 6: AtUGD3 (Proteína purificada)
- 7: AtUGD1 (Extrato bruto)
- 8: AtUGD1 (Proteína purificada)

**Fig.2**

W: Padrões de peso molecular  
1: cvUGD-HI (Extrato bruto)  
2: cvUGD-HI (Proteína purificada)  
3: cvUGD-KA (Extrato bruto)  
4: cvUGD-KA (Proteína purificada)

**Fig.3**

W: Padrões de peso molecular  
 1: AtGFAT (Antes da Reação)  
 2: AtGFAT (Depois de 24 horas de Reação)  
 3: cvGFAT-HI (Antes da Reação)  
 4: cvGFAT-HI (Depois de 24 horas de Reação)  
 5: DHFR (Antes da Reação)  
 6: DHFR (Depois de 24 horas de Reação)

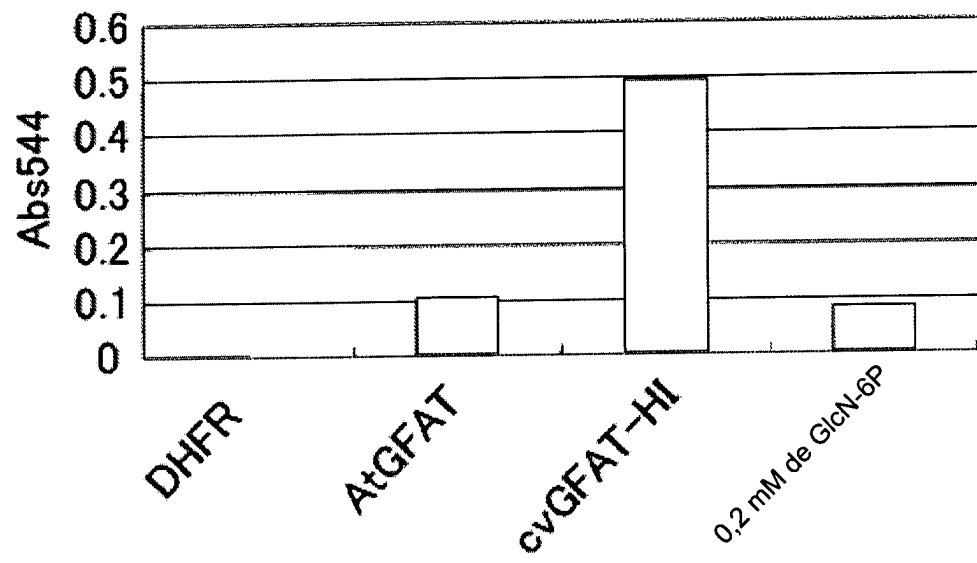
**Fig.4**

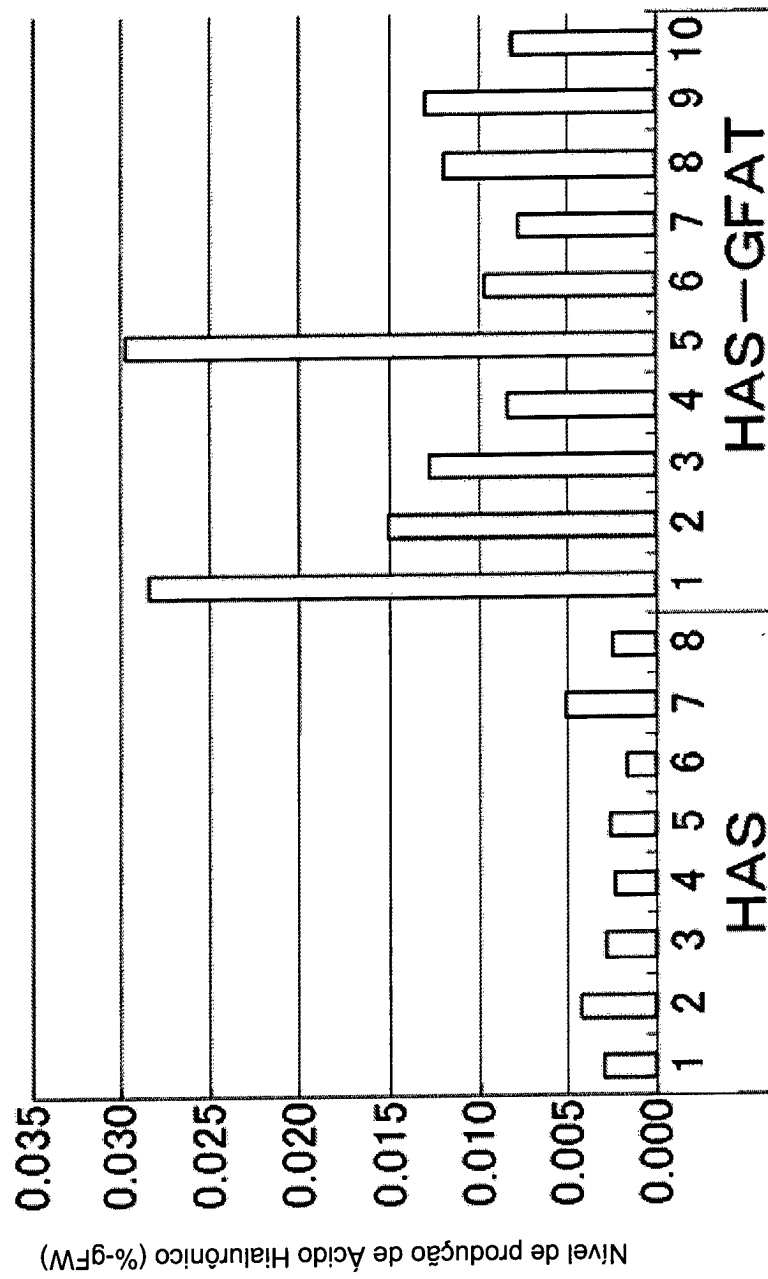
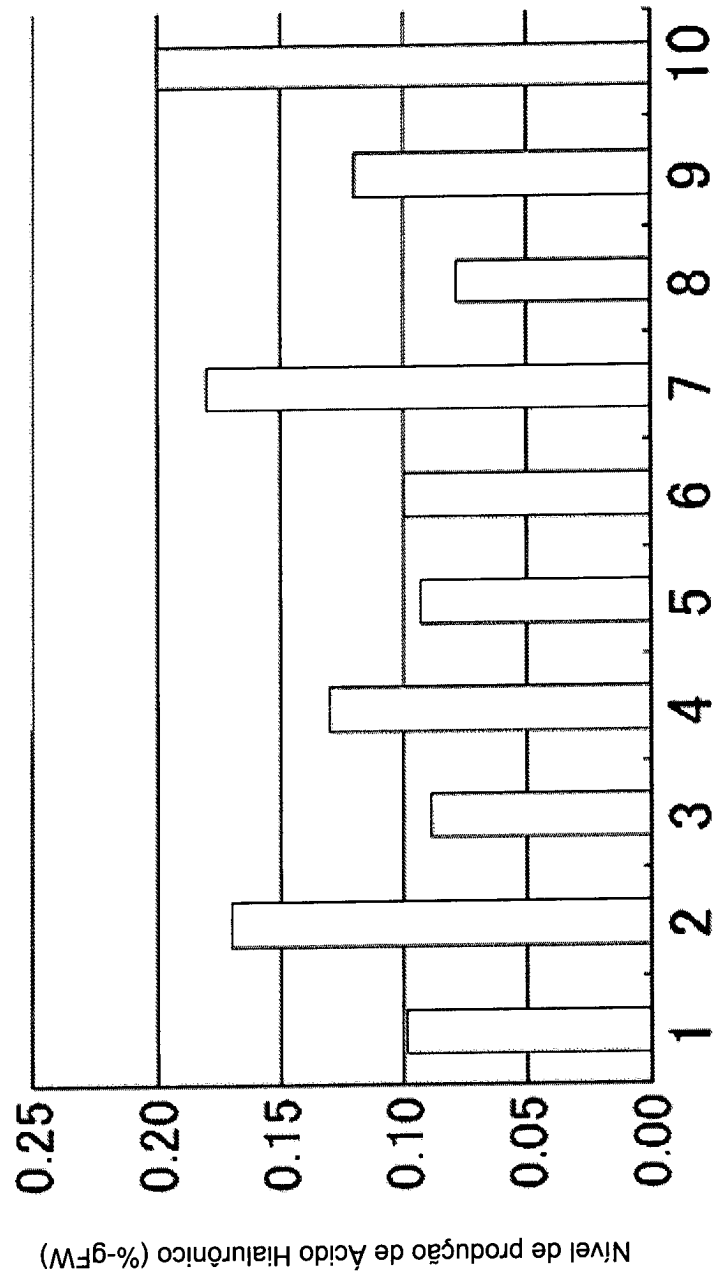
**Fig.5**

**Fig.6**

## RESUMO

Patente de Invenção: "**PLANTA PRODUTORA DE ÁCIDO HIALURÔNICO**".

A presente invenção proporciona, aprimorando um método conhecido para produzir ácido hialurônico em uma planta, uma planta ou algumas células vegetais cultivadas as quais podem produzir ácido hialurônico em um menor custo e um rendimento maior do que antes adicional, um método para preparar as mesmas, um vetor de expressão para transformação, um método para produzir ácido hialurônico usando a planta ou as células vegetais cultivadas e semelhantes. É proporcionado o método para produzir ácido hialurônico compreendendo obter ácido hialurônico por coexpressão de uma proteína com atividade de ácido hialurônico sintase e uma proteína exógena com atividade de nucleotídeo de açúcar sintase em uma célula vegetal ou uma planta.