(11) Número de Publicação: **PT 2057178 E**

(51) Classificação Internacional:
C07H 21/02 (2013.01) ***C07H 21/04*** (2013.01)
A01N 63/00 (2013.01) ***A61K 48/00*** (2013.01)
C12N 5/00 (2013.01) ***C12N 5/02*** (2013.01)
A23K 1/165 (2013.01) ***C12N 9/16*** (2013.01)
C12N 15/82 (2013.01) ***C12N 15/55*** (2013.01)

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

(22) Data de pedido: **2007.09.21**

(30) Prioridade(s): **2006.09.21 US 846831 P**

(43) Data de publicação do pedido: **2009.05.13**

(45) Data e BPI da concessão: **2013.11.06**
010/2014

(73) Titular(es):
VERENIUM CORPORATION
3550 JOHN HOPKINS COURT SAN DIEGO, CALIFORNIA 92121 **US**

(72) Inventor(es):
GERHARD FREY **US**
BRIAN STEER **US**
MARK DYCAICO **US**
KATIE KLINE **US**
AXEL TREFZER **NL**

(74) Mandatário:
ANTÓNIO INFANTE DA CÂMARA TRIGUEIROS DE ARAGÃO
RUA DO PATROCÍNIO, Nº 94 1399-019 LISBOA **PT**

(54) Epígrafe: **FITASES, ÁCIDOS NUCLEICOS QUE AS CODIFICAM E MÉTODOS PARA AS PREPARAR E UTILIZAR**

(57) Resumo:
ESTA INVENÇÃO REFERE-SE A FITASES, AOS POLINUCLEÓTIDOS QUE AS CODIFICAM, ÀS UTILIZAÇÕES DOS POLINUCLEÓTIDOS E POLIPÉPTIDOS DA INVENÇÃO, ASSIM COMO À PRODUÇÃO E ISOLAMENTO DE TAIS POLINUCLEÓTIDOS E POLIPÉPTIDOS. EM PARTICULAR, A INVENÇÃO PROPORCIONA POLIPÉPTIDOS TENDO ACTIVIDADE DE FITASE SOB CONDIÇÕES DE ALTA TEMPERATURA E FITASES QUE RETÊM A ACTIVIDADE APÓS EXPOSIÇÃO ÀS ALTAS TEMPERATURAS. AS FITASES DA INVENÇÃO PODEM SER TERMOTOLERANTES E/OU TERMOESTÁVEIS A BAIXAS TEMPERATURAS, ALÉM DE A ALTAS TEMPERATURAS. AS FITASES DA INVENÇÃO PODEM SER UTILIZADAS EM GÉNEROS ALIMENTÍCIOS PARA MELHORAR O VALOR ALIMENTAR DE INGREDIENTES RICOS EM FITATO. AS FITASES DA INVENÇÃO PODEM SER FORMULADAS COMO ALIMENTOS OU RAÇÕES OU SUPLEMENTOS PARA, E. G., AUXILIAR NA DIGESTÃO DO FITATO. OS ALIMENTOS OU RAÇÕES DA INVENÇÃO PODEM SER NA FORMA DE PASTILHAS, LÍQUIDOS, PÓS E SEMELHANTES. NUM ASPECTO, AS FITASES DA INVENÇÃO SÃO ESTÁVEIS CONTRA DESNATURAÇÃO TÉRMICA DURANTE A FORMAÇÃO DAS PASTILHAS; E ISTO DIMINUI O CUSTO DO PRODUTO DE FITASE ENQUANTO MANTÉM A EFICÁCIA IN VIVO E A DETECÇÃO DA ACTIVIDADE NA RAÇÃO.

RESUMO

"FITASES, ÁCIDOS NUCLEICOS QUE AS CODIFICAM E MÉTODOS PARA AS PREPARAR E UTILIZAR"

Esta invenção refere-se a fitases, aos polinucleótidos que as codificam, às utilizações dos polinucleótidos e polipéptidos da invenção, assim como à produção e isolamento de tais polinucleótidos e polipéptidos. Em particular, a invenção proporciona polipéptidos tendo actividade de fitase sob condições de alta temperatura e fitases que retêm a actividade após exposição às altas temperaturas. As fitases da invenção podem ser termotolerantes e/ou termoestáveis a baixas temperaturas, além de a altas temperaturas. As fitases da invenção podem ser utilizadas em géneros alimentícios para melhorar o valor alimentar de ingredientes ricos em fitato. As fitases da invenção podem ser formuladas como alimentos ou rações ou suplementos para, *e. g.*, auxiliar na digestão do fitato. Os alimentos ou rações da invenção podem ser na forma de pastilhas, líquidos, pós e semelhantes. Num aspecto, as fitases da invenção são estáveis contra desnaturação térmica durante a formação das pastilhas; e isto diminui o custo do produto de fitase enquanto mantém a eficácia *in vivo* e a detecção da actividade na ração.

DESCRIÇÃO

"FITASES, ÁCIDOS NUCLEICOS QUE AS CODIFICAM E MÉTODOS PARA AS PREPARAR E UTILIZAR"

CAMPO DA INVENÇÃO

Esta invenção refere-se a fitases, aos polinucleótidos que as codificam, às utilizações dos polinucleótidos e polipéptidos da invenção, assim como à produção e isolamento de tais polinucleótidos e polipéptidos. Em particular, a invenção proporciona polipéptidos TENDO actividade de fitase sob condições de alta temperatura e fitases que retêm a actividade após exposição às altas temperaturas. As fitases da invenção podem ser termotolerantes e/ou termoestáveis a baixas temperaturas, além de A altas temperaturas. As fitases da invenção podem ser utilizadas em géneros alimentícios para melhorar o valor alimentar de ingredientes ricos em fitato. As fitases da invenção podem ser formuladas como alimentos ou rações ou suplementos para, *e. g.*, auxiliar na digestão do fitato. Os alimentos ou rações da invenção podem ser na forma de pastilhas, líquidos, pós e semelhantes. Num aspecto, as fitases da invenção são estáveis contra desnaturação térmica durante a formação das pastilhas; e isto diminui o custo do produto de fitase enquanto mantém a eficácia *in vivo* e a detecção da actividade na ração.

ANTECEDENTES

Os minerais são elementos essenciais para o crescimento de todos os organismos. Os minerais dietéticos podem ser derivados de muitos materiais de partida, incluindo plantas. Por exemplo, as sementes de plantas são uma fonte rica de minerais, uma vez que contêm iões que são complexados com os grupos fosfato das moléculas de ácido fítico. Estes minerais associados ao fitato podem, em alguns casos, satisfazer as necessidades dietéticas de algumas espécies de organismos de criação, tal como ruminantes com múltiplos estômagos. Desta forma, em alguns casos, os ruminantes requerem menos suplementação dietética com fosfato inorgânico e minerais, porque os microrganismos no rúmen produzem as enzimas que catalisam a conversão do fitato (mio-inositol-hexafosfato) em inositol e fosfato inorgânico. No processo, os minerais que formam complexados com fitato são libertados. Contudo, a maioria das espécies de organismos de criação é incapaz de utilizar eficientemente minerais associados com fitato. Assim, por exemplo, na produção de animais de criação monogástricos (*e. g.*, porcos, aves e peixes), a alimentação é habitualmente suplementada com minerais e/ou com substâncias antibióticas que alteram o ambiente da flora digestiva do organismo consumidor para aumentar as taxas de crescimento.

Como tal, existem muitos problemas - relacionados com a nutrição, com passos de processamento *ex vivo*, com saúde e medicina, conservação ambiental e gestão de recursos - que estão associados com uma hidrólise insuficiente do fitato em muitas aplicações. Os seguintes são exemplos não limitativos destes problemas:

1) A suplementação de dietas com minerais inorgânicos é uma despesa cara.

2) A presença de fitato não hidrolisado é indesejável e problemática em muitas aplicações *ex vivo* (*e. g.*, por provocar a presença de detritos indesejados).

3) A suplementação de dietas com antibióticos coloca uma ameaça clínica aos humanos e igualmente aos animais por aumentar a abundância de patogéneos tolerantes a antibióticos.

4) A descarga de minerais fecais não absorvidos no ambiente altera e danifica os ecossistemas de solos circundantes, de águas de viveiros de peixes e de águas superficiais, em geral.

5) A oferta nutritiva valiosa de muitos potenciais géneros alimentícios permanece significativamente inexplorada e desperdiçada.

Consequentemente, os géneros alimentícios contendo fitato requerem suplementação com nutrientes exógenos e/ou com uma fonte de actividade de fitase de modo a corrigir a sua oferta nutritiva deficiente após consumo, por um número muito grande de espécies de organismos.

Consequentemente, existe uma necessidade para meios para conseguir uma hidrólise eficaz e rentável de fitato em várias aplicações. Particularmente, existe uma necessidade para meios para optimizar a hidrólise de fitato em aplicações comerciais. Num aspecto particular, existe uma necessidade para optimizar métodos de tratamento comerciais que melhoram a oferta nutritiva de géneros alimentícios contendo fitato para consumo por humanos e animais de criação.

SUMÁRIO DA INVENÇÃO

Esta invenção proporciona polipéptidos tendo actividade de fitase, polinucleótidos que as codificam, utilizações dos polinucleótidos e polipéptidos da invenção e métodos para a produção e isolamento de tais polinucleótidos e polipéptidos. Num aspecto, a invenção proporciona polipéptidos tendo actividade de fitase sob condições de alta temperatura e fitases que retêm actividade após exposição às altas temperaturas. As fitases da invenção podem ser utilizadas em géneros alimentícios para melhorar o valor alimentar de ingredientes ricos em fitato. As fitases da invenção podem ser formuladas como alimentos ou rações ou suplementos para, *e. g.*, auxiliar na digestão do fitato. Os alimentos ou rações da invenção podem estar na forma de pastilhas, comprimidos, pílulas, líquidos, pós, sprays e semelhantes. Num aspecto, as fitases da invenção são estáveis o contra desnaturação térmica durante a formação das pastilhas; e isto diminui o custo do produto de fitase enquanto mantém a eficácia *in vivo* e a detecção da actividade na ração.

SUMÁRIO:

A invenção proporciona ácidos nucleicos isolados, sintéticos ou recombinantes compreendendo

(a) uma sequência de ácido nucleico codificando um polipéptido tendo actividade de fitase e que tem, pelo menos, 95%, 96%, 97%, 98% ou 99% de identidade de sequência com a SEQ ID N°:1, e tendo uma modificação nucleotídica em que os nucleótidos nas posições 139 a 141 da SEQ ID N°:1 são TTT ou TTC, em que o polipéptido melhorou a termostabilidade relativamente à fitase da SEQ ID N°:2. A sequência pode

opcionalmente compreender ainda, pelo menos, uma até todas as treze modificações da SEQ ID N°:1 seleccionadas de os nucleótidos nas posições 289 a 291 serem GTT, GTC, GTA ou GTG; os nucleótidos nas posições 289 a 291 serem GAA ou GAG; os nucleótidos nas posições 406 a 408 serem CAT ou CAC; os nucleótidos nas posições 475 a 477 serem GTT, GTC, GTA ou GTG; os nucleótidos nas posições 475 a 477 serem GAA ou GAG; os nucleótidos nas posições 487 a 489 serem CGT, CGC, CGA, CGG, AGA ou AGG; os nucleótidos nas posições 490 a 492 serem CGT, CGC, CGA, CGG, AGA ou AGG; os nucleótidos nas posições 502 a 504 serem CGT, CGC, CGA, CGG, AGA ou AGG; os nucleótidos nas posições 535 a 537 serem CGT, CGC, CGA, CGG, AGA ou AGG; os nucleótidos nas posições 676 a 678 serem GAT ou GAC; os nucleótidos nas posições 697 a 699 serem TGG; os nucleótidos nas posições 823 a 825 serem GTT, GTC, GTA ou GTG; os nucleótidos nas posições 865 a 867 serem GCT, GCC, GCA ou GCG; os nucleótidos nas posições 1045 a 1047 serem TAT ou TAC; e os nucleótidos nas posições 1087 a 1089 serem CCA, CCC, CCG ou CCT; ou,

(b) um polinucleótido codificando um polipéptido tendo actividade de fitase e tendo, pelo menos, 95%, 96%, 97%, 98% ou 99% ou mais de identidade de sequência com a SEQ ID N°:2 e tendo uma sequência compreendendo, pelo menos, uma modificação de resíduo de aminoácido, em que o aminoácido alanina na posição de aminoácido 47 da SEQ ID N°:2 é alterado para uma fenilalanina em que o polipéptido melhorou a termostabilidade relativamenteà fitase da SEQ ID N°:2;

(c) a sequência de ácido nucleico de (a) ou (b) codificando um polipéptido tendo uma actividade de fitase mas que carece de uma sequência de sinal ou uma sequência de proproteína, ou uma sequência promotora homóloga;

(d) o ácido nucleico de qualquer de (a) a (c) codificando um polipéptido tendo uma actividade de fitase e compreendendo

ainda uma sequência de aminoácidos heteróloga, ou o ácido nucleico de qualquer de (a) a (c) compreendendo ainda uma sequência de nucleótidos heteróloga;

(e) o ácido nucleico de (d), em que a sequência de aminoácidos heteróloga compreende ou consiste em uma sequência que codifica uma sequência de sinal heteróloga (líder), ou um Marcador ou um epitopo, ou a sequência de nucleótidos heteróloga compreende uma sequência heteróloga promotora ou de direccionamento;

(f) uma sequência de ácido nucleico totalmente complementar à sequência de ácido nucleico de qualquer de (a) a (e).

Opcionalmente as identidades de sequência são determinadas por análise com um algoritmo de comparação de sequência ou por inspecção visual. Todos estes ácidos nucleicos são "ácidos nucleicos da invenção", codificando "polipéptidos da invenção".

Num aspecto, o algoritmo de comparação de sequência é um algoritmo BLAST versão 2.2.2, em que é estabelecido um filtro como blastall -p blastp -d "nr pataa" - F F, e todas as outras opções são estabelecidas por defeito.

Num aspecto, a actividade de fitase compreende: catálise de fitato (mio-inositol-hexafosfato) em inositol e fosfato inorgânico; ou a hidrólise de fitato (mio-inositol-hexafosfato). Num aspecto, a actividade de fitase é termotolerante e, opcionalmente, o polipéptido retém uma actividade de fitase após exposição a uma temperatura na gama entre cerca de 40 °C a cerca de 70 °C, ou entre cerca de 37 °C a cerca de 100 °C, ou maior que 37 °C a cerca de 50 °C, 55 °C, 60 °C, 65 °C, 70 °C, 75 °C, 80 °C, 85 °C, 90 °C, 95 °C, 100 °C ou mais. Num aspecto, a actividade de fitase é termoestável e, opcionalmente, o polipéptido retém uma actividade de fitase a uma temperatura na

gama entre cerca de 40 °C a cerca de 70 °C, ou entre cerca de 37 °C a cerca de 100 °C, ou maior que 37 °C a cerca de 50 °C, 55 °C, 60 °C, 65 °C, 70 °C, 75 °C, 80 °C, 85 °C, 90 °C, 95 °C, 100 °C ou mais.

A invenção proporciona cassetes de expressão, vectores, veículos de clonagem, vectores de expressão, vectores de clonagem compreendendo um ácido nucleico da invenção ou contendo aí um ácido nucleico da invenção (o qual inclui ácidos nucleicos que codificam polipéptidos da invenção), em que opcionalmente o veículo de clonagem compreende ou é um vector viral, um plasmídeo, um fago, um fagemídeo, um cosmídeo, um fosmídeo, um bacteriófago ou um cromossoma artificial e, opcionalmente, o vector viral compreende um vector de adenovírus, vectores retrovirais ou um vector viral adeno-associado e, opcionalmente, a cassete de expressão, vector, veículo de clonagem, vector de expressão, vector de clonagem compreende ou está contido num cromossoma artificial bacteriano (BAC), num plasmídeo, num vector derivado do bacteriófago P1 (PAC), num cromossoma artificial de levedura (YAC), num cromossoma artificial de mamífero (MAC).

A invenção proporciona células transformadas, célula transduzidas, células hospedeiras e semelhantes, compreendendo um ácido nucleico da invenção ou contendo aí um ácido nucleico da invenção (o qual inclui ácidos nucleicos que codificam polipéptidos da invenção), ou a cassete de expressão, vector, veículo de clonagem, vector de expressão, vector de clonagem da invenção, em que opcionalmente a célula é uma célula bacteriana, uma célula de mamífero, uma célula fúngica, uma célula de levedura, uma célula de insecto ou uma célula vegetal.

Os animais não humanos transgénicos podem compreender um ácido nucleico da invenção ou contendo aí um ácido nucleico da invenção (que inclui ácidos nucleicos que codificam polipéptidos da invenção) ou a cassete de expressão, vector, veículo de clonagem, vector de expressão, vector de clonagem da invenção, em que opcionalmente o animal é um murganho, um rato, uma cabra, um coelho, uma ovelha, um porco ou uma vaca.

A invenção proporciona plantas transgénicas (incluindo partes de plantas, *e. g.*, processadas ou recolhidas, *e. g.*, folhas, caules, raízes ou frutas), ou sementes compreendendo um ácido nucleico da invenção ou contendo aí um ácido nucleico da invenção (o qual inclui ácidos nucleicos que codificam polipéptidos da invenção) ou a cassete de expressão, vector, veículo de clonagem, vector de expressão, vector de clonagem da invenção, em que opcionalmente a planta é uma planta de milho, uma planta de batata, uma planta de tomate, uma planta de trigo, uma planta oleaginosa, uma planta de colza, uma planta de soja ou uma planta de tabaco e, opcionalmente, a semente é uma semente de milho, uma semente de trigo, uma semente oleaginosa, uma semente de colza, uma semente de soja, uma semente de palma, uma semente de girassol, uma semente de sésamo, um amendoim ou uma semente da planta de tabaco. Em formas de realização alternativas, a planta ou uma semente produzida a partir de uma semente ou de uma planta da invenção, ou uma planta ou uma semente da invenção, pode incluir cultivares, por exemplo, milho, alfalfa, girassol, *Brassica,* soja, cana de açúcar, algodão, cártamo, amendoim, sorgo, trigo, aveia, centeio, milho painço, cevada, arroz, coníferas, leguminosas, *e. g.*, ervilha, feijão e soja, tubérculos/raízes amiláceas, *e. g.*, batata, batata-doce, mandioca, inhame, balizeiro e beterraba açucareira e semelhantes, ou a planta pode ser uma planta de milho, uma planta de batata, uma planta de tomate, uma planta de trigo, uma

planta oleaginosa, uma planta de colza, uma planta de soja ou uma planta de tabaco, ou uma planta forrageira e/ou de ração para um animal, ou para um animal ruminante, ou a planta pode ser ou compreender uma planta forrageira ou de ração que é feno, milho, milho painço, soja, trigo, trigo mourisco, cevada, alfalfa, centeio, uma erva anual, sorgo, erva-do-sudão, erva de veldt ou erva de buffel; ou a semente é uma semente de milho, uma semente de trigo, uma semente oleaginosa, uma semente de colza, uma semente de soja, uma semente de palma, uma semente de girassol, uma semente de sésamo, um amendoim ou uma semente de amendoim, uma semente de alfalfa, uma semente de algodão, uma semente de cártamo, uma semente de sorgo, uma semente de aveia, uma semente de centeio, uma semente de milho painço, uma semente de cevada, uma semente de arroz, uma semente de ervilha ou uma semente de planta de tabaco, ou a planta é milho, alfalfa, girassol, Brassica, soja, cana-de-açúcar, algodão, cártamo, amendoim, sorgo, trigo, aveia, centeio, milho painço, cevada, arroz, coníferas, ervilha, feijão, soja, batata, batata-doce, mandioca, inhame, balizeiro ou beterraba açucareira.

Os oligonucleótidos anti-sentido podem compreender um ácido nucleico da invenção com as modificações especificadas da sequência de par de bases nucleotídica para a SEQ ID N°:1, em que opcionalmente o oligonucleótido anti-sentido está entre cerca de 10 a 50, cerca de 20 a 60, cerca de 30 a 70, cerca de 40 a 80, cerca de 60 a 100 ou cerca de 50 a 150 bases de comprimento. A invenção também proporciona ribozimas e/ou iARN (*e. g.*, siARN ou miARN) compreendendo sequências anti-sentido da invenção.

A invenção proporciona métodos de inibição da tradução de um mensageiro de fitase numa célula compreendendo administrar à

célula ou expressar na célula um oligonucleótido anti-sentido da invenção.

A invenção proporciona polipéptidos isolados, sintéticos ou recombinantes compreendendo

(a) uma sequência de aminoácidos tendo, pelo menos, 95%, 96%, 97%, 98% ou 99% de identidade de sequência para a SEQ ID N°:2 e codificada por um polinucleótido compreendendo uma modificação nucleotídica para a SEQ ID N°:1, em que os nucleótidos nas posições 139 a 141 da SEQ ID N°:1 são TTT ou TTC, em que o polipéptido melhorou a termostabilidade da fitase da SEQ ID N°:2;
(b) uma sequência de aminoácidos tendo, pelo menos, 95%, 96%, 97%, 98% ou 99% de identidade de sequência para a SEQ ID N°:2 e tem uma modificação de aminoácido, em que o aminoácido alanina na posição de aminoácido 47 da SEQ ID N°:2 é alterado para uma fenilalanina em que o polipéptido melhorou a termostabilidade da fitase da SEQ ID N°:2. Opcionalmente, o polipéptido pode compreender, pelo menos, 1 a 13 modificações adicionais seleccionadas de: o aminoácido cisteína na posição de aminoácido 97 da SEQ ID N°:2 é alterado para uma valina; o aminoácido cisteína na posição de aminoácido 97 da SEQ ID N°:2 é alterado para um ácido glutâmico; o aminoácido treonina equivalente na posição de aminoácido 136 da SEQ ID N°:2 é alterado para uma histidina; o aminoácido asparagina na posição de aminoácido 159 da SEQ ID N°:2 é alterado para uma valina; o aminoácido asparagina na posição de aminoácido 159 da SEQ ID N°:2 é alterado para um ácido glutâmico; o aminoácido treonina equivalente na posição de aminoácido 163 da SEQ ID N°:2 é alterado para uma arginina; o aminoácido ácido aspártico na posição de aminoácido 164 da SEQ ID N°:2

é alterado para uma arginina; o aminoácido ácido glutâmico na posição de aminoácido 168 da SEQ ID N°:2 é alterado para uma arginina; o aminoácido glicina na posição de aminoácido 179 da SEQ ID N°:2 é alterado para uma arginina; o aminoácido cisteína na posição de aminoácido 226 da SEQ ID N°:2 é alterado para um ácido aspártico; o aminoácido valina na posição de aminoácido 233 da SEQ ID N°:2 é alterado para um triptofano; o aminoácido glutamina na posição de aminoácido 275 da SEQ ID N°:2 é alterado para uma valina; o aminoácido arginina na posição de aminoácido 289 da SEQ ID N°:2 é alterado para uma alanina; o aminoácido treonina na posição de aminoácido 349 da SEQ ID N°:2 é alterado para uma tirosina; e o aminoácido leucina na posição de aminoácido 363 da SEQ ID N°:2 é alterado para uma prolina,

em que os ácidos nucleicos codificam, pelo menos, um polipéptido tendo uma actividade de fitase e opcionalmente as identidades de sequência são determinadas por análise com um algoritmo de comparação de sequência ou por inspecção visual,

e o polipéptido tem uma actividade de fitase e, opcionalmente, a actividade de fitase compreende: catálise de fitato (mio-inositol-hexafosfato) em inositol e fosfato inorgânico; ou a hidrólise de fitato (mio-inositol-hexafosfato).

A invenção proporciona polipéptidos da invenção tendo actividade de fitase, cuja actividade é termotolerante e, opcionalmente, o polipéptido retém uma actividade de fitase após exposição a uma temperatura na gama desde cerca de -100 °C a cerca de -80 °C, cerca de -80 °C a cerca de -40 °C, cerca de -40 °C a cerca de -20 °C, cerca de -20 °C a cerca de 0 °C, cerca de 0 °C a cerca de 5 °C, cerca de 5 °C a cerca de 15 °C, cerca

de 15 °C a cerca de 25 °C, cerca de 25 °C a cerca de 37 °C, cerca de 37 °C a cerca de 45 °C, cerca de 45 °C a cerca de 55 °C, cerca de 55 °C a cerca de 70 °C, cerca de 70 °C a cerca de 75 °C, cerca de 75 °C a cerca de 85 °C, cerca de 85 °C a cerca de 90 °C, cerca de 90 °C a cerca de 95 °C, cerca de 95 °C a cerca de 100 °C, cerca de 100 °C a cerca de 105 °C, cerca de 105 °C a cerca de 110 °C, cerca de 110 °C a cerca de 120 °C, ou 95 °C, 96 °C, 97 °C, 98 °C, 99 °C, 100 °C, 101 °C, 102 °C, 103 °C, 104 °C, 105 °C, 106 °C, 107 °C, 108 °C, 109 °C, 110 °C, 111 °C, 112 °C, 113 °C, 114 °C, 115 °C ou mais. Os polipéptidos termotolerantes de acordo com a invenção podem reter actividade, *e. g.*, uma actividade de fitase, após exposição a uma temperatura na gama desde cerca de -100 °C a cerca de -80 °C, cerca de -80 °C cerca de a -40 °C, cerca de -40 °C a cerca de -20 °C, cerca de -20 °C a cerca de 0 °C, cerca de 0 °C a cerca de 5 °C, cerca de 5 °C a cerca de 15 °C, cerca de 15 °C a cerca de 25 °C, cerca de 25 °C a cerca de 37 °C, cerca de 37 °C a cerca de 45 °C, cerca de 45 °C a cerca de 55 °C, cerca de 55 °C a cerca de 70 °C, cerca de 70 °C a cerca de 75 °C, cerca de 75 °C a cerca de 85 °C, cerca de 85 °C a cerca de 90 °C, cerca de 90 °C a cerca de 95 °C, cerca de 95 °C a cerca de 100 °C, cerca de 100 °C a cerca de 105 °C, cerca de 105 °C a cerca de 110 °C, cerca de 110 °C a cerca de 120 °C, ou 95 °C, 96 °C, 97 °C, 98 °C, 99 °C, 100 °C, 101 °C, 102 °C, 103 °C, 104 °C, 105 °C, 106 °C, 107 °C, 108 °C, 109 °C, 110 °C, 111 °C, 112 °C, 113 °C, 114 °C, 115 °C ou mais. Em algumas formas de realização, os polipéptidos termotolerantes de acordo com a invenção retêm actividade, *e. g.*, uma actividade de fitase, após exposição a uma temperatura nas gamas descritas acima, a cerca de pH 3,0, a cerca de pH 3,5, a cerca de pH 4,0, a cerca de pH 4,5, a cerca de pH 5,0, a cerca de pH 5,5, a cerca de pH 6,0, a cerca de pH 6,5, a cerca de pH 7,0, a cerca de pH 7,5, a cerca de pH 8,0, a cerca de pH 8,5, a cerca de pH 9,0, a cerca de pH 9,5, cerca de

pH 10,0, cerca de pH 10,5, cerca de pH 11,0, cerca de pH 11,5, cerca de pH 12,0 ou mais.

A invenção proporciona polipéptidos da invenção que têm actividade de fitase, cuja actividade é termoestável. Por exemplo, um polipéptido da invenção pode ser termoestável. O polipéptido termoestável de acordo com a invenção pode reter actividade enzimática e/ou de ligação, *e. g.*, uma actividade de fitase, sob condições compreendendo uma gama de temperatura desde cerca de -100 °C a cerca de -80 °C, cerca de -80 °C a cerca de -40 °C, cerca de -40 °C a cerca de -20 °C, cerca de -20 °C a cerca de 0 °C, cerca de 0 °C a cerca de 37 °C, cerca de 0 °C a cerca de 5 °C, cerca de 5 °C a cerca de 15 °C, cerca de 15 °C a cerca de 25 °C, cerca de 25 °C a cerca de 37 °C, cerca de 37 °C a cerca de 45 °C, cerca de 45 °C a cerca de 55 °C, cerca de 55 °C a cerca de 70 °C, cerca de 70 °C a cerca de 75 °C, cerca de 75 °C a cerca de 85 °C, cerca de 85 °C a cerca de 90 °C, cerca de 90 °C a cerca de 95 °C, cerca de 95 °C a cerca de 100 °C, cerca de 100 °C a cerca de 105 °C, cerca de 105 °C a cerca de 110 °C, cerca de 110 °C a cerca de 120 °C, ou 95 °C, 96 °C, 97 °C, 98 °C, 99 °C, 100 °C, 101 °C, 102 °C, 103 °C, 104 °C, 105 °C, 106 °C, 107 °C, 108 °C, 109 °C, 110 °C, 111 °C, 112 °C, 113 °C, 114 °C, 115 °C ou mais. Os polipéptidos termoestáveis de acordo com a invenção podem reter actividade, *e. g.*, uma actividade de fitase, nas temperaturas na gama desde cerca de -100 °C a cerca de -80 °C, cerca de -80 °C a cerca de -40 °C, cerca de -40 °C a cerca de -20 °C, cerca de -20 °C a cerca de 0 °C, cerca de 0 °C a cerca de 5 °C, cerca de 5 °C a cerca de 15 °C, cerca de 15 °C a cerca de 25 °C, cerca de 25 °C a cerca de 37 °C, cerca de 37 °C a cerca de 45 °C, cerca de 45 °C a cerca de 55 °C, cerca de 55 °C a cerca de 70 °C, cerca de 70 °C a cerca de 75 °C, cerca de 75 °C a cerca de 85 °C, cerca de 85 °C a cerca de 90 °C, cerca de 90 °C a cerca de

95 °C, cerca de 95 °C a cerca de 100 °C, cerca de 100 °C a cerca de 105 °C, cerca de 105 °C a cerca de 110 °C, cerca de 110 °C a cerca de 120 °C, ou 95 °C, 96 °C, 97 °C, 98 °C, 99 °C, 100 °C, 101 °C, 102 °C, 103 °C, 104 °C, 105 °C, 106 °C, 107 °C, 108 °C, 109 °C, 110 °C, 111 °C, 112 °C, 113 °C, 114 °C, 115 °C ou mais. Em algumas formas de realização, os polipéptidos termoestáveis de acordo com a invenção retêm actividade, *e. g.*, uma actividade de fitase, a uma temperatura nas gamas descritas acima, a cerca de pH 3,0, a cerca de pH 3,5, a cerca de pH 4,0, a cerca de pH 4,5, a cerca de pH 5,0, a cerca de pH 5,5, a cerca de pH 6,0, a cerca de pH 6,5, a cerca de pH 7,0, a cerca de pH 7,5, a cerca de pH 8,0, a cerca de pH 8,5, a cerca de pH 9,0, a cerca de pH 9,5, cerca de pH 10,0, cerca de pH 10,5, cerca de pH 11,0, cerca de pH 11,5, cerca de pH 12,0 ou mais.

A invenção proporciona polipéptidos isolados, sintéticos ou recombinantes compreendendo uma sequência de aminoácidos da invenção e (a) carecendo de uma sequência de sinal (péptido líder); (b) carecendo de uma sequência de sinal (péptido líder) e compreendendo ainda uma sequência de sinal heteróloga (péptido líder); (c) a sequência de aminoácidos de (a) ou (b) e compreendendo ainda uma sequência heteróloga, em que opcionalmente a sequência heteróloga compreende um péptido de identificação.

Num aspecto, a actividade de fitase de qualquer polipéptido da invenção compreende (tem) uma actividade específica: a cerca de 37 °C na gama desde cerca de 100 a cerca de 1000 unidades por miligrama de proteína; ou desde cerca de 500 a cerca de 750 unidades por miligrama de proteína; ou a 37 °C na gama desde cerca de 500 a cerca de 1200 unidades por miligrama de proteína; ou a 37° C na gama desde cerca de 750 a cerca de 1000 unidades por miligrama de proteína. Num aspecto, a actividade de fitase

termotolerante compreende uma actividade específica após exposição a uma temperatura a cerca de 37 °C na gama desde cerca de 100 a cerca de 1000 unidades por miligrama de proteína; ou a actividade de fitase termoestável compreende uma actividade específica desde cerca de 500 a cerca de 750 unidades por miligrama de proteína; ou a actividade de fitase termoestável compreende uma actividade específica a 37 °C na gama desde cerca de 500 a cerca de 1200 unidades por miligrama de proteína; ou a actividade de fitase termoestável compreende uma actividade específica a 37 °C na gama desde cerca de 750 a cerca de 1000 unidades por miligrama de proteína. Num aspecto, a actividade de fitase termoestável compreende uma actividade específica sob condições compreendendo uma temperatura de cerca de 37 °C na gama desde cerca de 100 a cerca de 1000 unidades por miligrama de proteína; ou a actividade de fitase termoestável compreende uma actividade específica desde cerca de 500 a cerca de 750 unidades por miligrama de proteína; ou a actividade de fitase termoestável compreende uma actividade específica a 37 °C na gama desde cerca de 500 a cerca de 1200 unidades por miligrama de proteína; ou a actividade de fitase termoestável compreende uma actividade específica a 37 °C na gama desde cerca de 750 a cerca de 1000 unidades por miligrama de proteína.

Num aspecto, um polipéptido da invenção compreende, pelo menos, um local de glicosilação, em que opcionalmente a glicosilação é uma glicosilação ligada por N ou uma glicosilação ligada por O e, opcionalmente, o polipéptido é glicosilado após ter sido expresso numa levedura que é, opcionalmente, uma *P. pastoris* ou uma *S. pombe*.

Num aspecto, o polipéptido retém uma actividade de fitase sob condições compreendendo cerca de pH 6,5, pH 6, pH 5,5, pH 5, pH 4,5, pH 4,0, pH 3,5, pH 3,0 ou menor pH (mais acídico).

Alternativamente, um polipéptido da invenção pode reter uma actividade de fitase sob condições compreendendo cerca de pH 7,5, pH 8, pH 8,5, pH 9, pH 9,5, pH 10,0, pH 10,5, pH 11,0, pH 11,5, pH 12, pH 12,5 ou maior pH (mais básico).

A invenção proporciona preparações de proteína compreendendo um polipéptido da invenção, em que a preparação de proteína compreende um líquido, uma pasta, um pó, um spray, uma suspensão, uma composição/formulação liofilizada, um sólido, geltab, pílula, implante, um gel; ou uma formulação farmacêutica, um alimento, uma ração, um suplemento alimentar, um suplemento de ração, um aditivo alimentar, um aditivo de ração, um suplemento nutritivo ou um seu suplemento dietético.

Os heterodímeros podem compreender um polipéptido da invenção e num aspecto o heterodímero compreende um segundo domínio, em que opcionalmente o segundo domínio é um polipéptido e o heterodímero é uma proteína de fusão e, opcionalmente, o segundo domínio é um epitopo ou um marcador.

A invenção proporciona polipéptidos imobilizados compreendendo um polipéptido da invenção, em que o polipéptido imobilizado pode compreender um homodímero ou um heterodímero, em que opcionalmente o polipéptido é imobilizado sobre ou dentro de uma célula, uma vesícula, um lipossoma, uma película, uma membrana, um metal, uma resina, um polímero, um cerâmico, um vidro, um microelétrodo, uma partícula grafítica, um grânulo, um gel, uma placa, uma matriz, um tubo capilar, um cristal, um comprimido, um pílula, uma cápsula, um pó, um aglomerado, uma superfície ou uma estrutura porosa. As matrizes (*e. g.*, *microarrays*) podem compreender um polipéptido imobilizado, em que o polipéptido compreende o polipéptido da invenção, ou o

heterodímero, ou o ácido nucleico da invenção, ou uma sua combinação.

Os anticorpos isolados, sintéticos ou recombinantes podem-se ligar especificamente a um polipéptido da invenção ou a um polipéptido codificado pelo ácido nucleico da invenção, em que opcionalmente o anticorpo é um anticorpo monoclonal ou policlonal. Os hibridomas podem compreender o anticorpo.

A invenção proporciona métodos para produzir um polipéptido recombinante compreendendo os passos de: (a) proporcionar um ácido nucleico ligado operativamente a um promotor; em que o ácido nucleico compreende uma sequência da invenção; e (b) expressar o ácido nucleico do passo (a) sob condições que permitam a expressão do polipéptido, produzindo assim um polipéptido recombinante e, opcionalmente, o método compreende ainda a transformação de uma célula hospedeira com o ácido nucleico do passo (a), seguido pela expressão do ácido nucleico do passo (a), produzindo assim o polipéptido recombinante numa célula hospedeira transformada.

A invenção proporciona métodos para identificar um polipéptido tendo uma actividade de fitase compreendendo os seguintes passos: (a) proporcionar o polipéptido da invenção; (b) proporcionar um substrato de fitase; e (c) colocar em contacto o polipéptido ou um seu fragmento ou variante do passo (a) com o substrato do passo (b) e detectar um aumento na quantidade de substrato ou uma diminuição na quantidade de produto da reacção, em que uma diminuição na quantidade do substrato ou um aumento na quantidade do produto da reacção detecta um polipéptido tendo uma actividade de fitase.

Os métodos para identificar um substrato de fitase compreendem os seguintes passos: (a) proporcionar o polipéptido da invenção; (b) proporcionar um substrato de teste; e (c) colocar em contacto o polipéptido do passo (a) com o substrato de teste do passo (b) e detectar um aumento na quantidade de substrato ou uma diminuição na quantidade de produto da reacção, em que uma diminuição na quantidade do substrato ou um aumento na quantidade do produto da reacção identificam o substrato de teste como um substrato de fitase.

Os métodos para determinar se um composto se liga especificamente a um polipéptido compreendem os seguintes passos: (a) expressar um ácido nucleico ou um vector compreendendo o ácido nucleico sob condições permissivas para tradução do ácido nucleico num polipéptido, em que o ácido nucleico compreende uma sequência da invenção; (b) colocar em contacto o polipéptido com o composto de teste; e (c) determinar se o composto de teste se liga especificamente ao polipéptido, determinando assim que o composto se liga especificamente ao polipéptido.

Os métodos para identificar um modulador de uma actividade de fitase compreendem os seguintes passos: (a) proporcionar o polipéptido de fitase da invenção; (b) proporcionar um composto de teste; (c) colocar em contacto o polipéptido do passo (a) com o composto de teste do passo (b) e medir uma actividade de fitase, em que uma alteração na actividade de fitase medida na presença do composto de teste comparativamente à actividade na ausência do composto de teste proporciona uma determinação de que o composto de teste modula a actividade de fitase, em que opcionalmente a actividade de fitase é medida ao proporcionar um substrato de fitase e detectar um aumento na quantidade do substrato ou uma diminuição na quantidade de um produto da

reacção e, opcionalmente, uma diminuição na quantidade do substrato ou um aumento na quantidade do produto da reacção com o composto de teste em comparação à quantidade de substrato ou produto da reacção sem o composto de teste, identifica o composto de teste como um activador da actividade de fitase e, opcionalmente, um aumento na quantidade do substrato ou uma diminuição na quantidade do produto da reacção com o composto de teste em comparação à quantidade de substrato ou produto da reacção sem o composto de teste, identifica o composto de teste como um inibidor da actividade de fitase.

Os sistemas computorizados podem compreender um processador e um dispositivo de armazenamento de dados, em que o referido dispositivo de armazenamento de dados armazenou aí uma sequência de polipéptido ou uma sequência de ácido nucleico, em que a sequência de polipéptido compreende uma sequência da invenção, e o ácido nucleico compreende uma sequência da invenção e, opcionalmente, o sistema compreende ainda um algoritmo de comparação de sequências e um dispositivo de armazenamento de dados tendo, pelo menos, uma sequência de referência aí armazenada e, opcionalmente, o algoritmo de comparação de sequências compreende um programa de computador que indica polimorfismos e, opcionalmente, o sistema compreende ainda um identificador que identifica uma ou mais características na referida sequência. O meio (meios) de leitura por computador pode ter aí armazenado uma sequência de polipéptido ou uma sequência de ácido nucleico, em que a sequência de polipéptido compreende a sequência de aminoácidos da invenção, e o ácido nucleico compreende uma sequência da invenção. Os métodos para identificar uma característica numa sequência compreendem os passos de: (a) ler a sequência utilizando um programa de computador que identifica uma ou mais características numa sequência, em que a sequência compreende uma sequência de

polipéptido ou uma sequência de ácido nucleico, em que a sequência de polipéptido compreende a sequência de aminoácidos da invenção, e o ácido nucleico compreende uma sequência da invenção; e, (b) identificar uma ou mais características na sequência com o programa de computador. Os métodos para comparar uma primeira sequência a uma segunda sequência compreendem os passos de: (a) ler a primeira sequência e a segunda sequência através da utilização de um programa de computador que compara sequências, em que a primeira sequência compreende uma sequência de polipéptido ou uma sequência de ácido nucleico, em que a sequência de polipéptido compreende a sequência de aminoácidos da invenção, e o ácido nucleico compreende uma sequência da invenção; e, (b) determinar diferenças entre a primeira sequência e a segunda sequência com o programa de computador, em que opcionalmente o passo de determinar diferenças entre a primeira sequência e a segunda sequência compreende ainda o passo de identificar polimorfismos e, opcionalmente, o método compreende ainda um identificador que identifica uma ou mais características numa sequência e, opcionalmente, o método compreende ler a primeira sequência utilizando um programa de computador e identificar uma ou mais características na sequência.

A invenção proporciona métodos para hidrolisar um inositol-hexafosfato em inositol e fosfato inorgânico compreendendo os seguintes passos: (a) proporcionar um polipéptido que tem uma actividade de fitase, em que o polipéptido compreende a sequência de aminoácidos da invenção, ou um polipéptido codificado pelo ácido nucleico da invenção; (b) proporcionar uma composição compreendendo um inositol-hexafosfato; e (c) colocar em contacto o polipéptido do passo (a) com a composição do passo (b) sob condições em que o polipéptido hidrolisa o inositol-hexafosfato para produzir inositol e fosfato

inorgânico, em que opcionalmente as condições compreendem uma temperatura entre cerca de 37 °C e cerca de 70 °C, entre cerca de 50 °C e cerca de 80 °C, ou entre cerca de 60 °C e cerca de 90 °C e, opcionalmente, a composição compreende um ácido fítico.

A invenção proporciona métodos para decruagem de óleo compreendendo os seguintes passos: (a) proporcionar um polipéptido tendo uma actividade de fitase, em que o polipéptido compreende a sequência de aminoácidos da invenção, ou um polipéptido codificado pelo ácido nucleico da invenção; (b) proporcionar uma composição compreendendo um óleo vegetal; e (c) colocar em contacto o polipéptido do passo (a) e o óleo vegetal do passo (b) sob condições em que o polipéptido pode clivar uma ligação inositol-fosfato inorgânico, decruando assim o óleo vegetal.

A invenção proporciona métodos para produzir uma ração ou alimento, ou um suplemento de ração ou alimentar, ou um aditivo alimentar ou de ração, ou um suplemento nutritivo, ou um suplemento dietético, compreendendo os seguintes passos: (a) transformar uma planta, uma parte de planta ou uma célula vegetal com um polinucleótido que codifica um polipéptido da enzima fitase, em que a fitase compreende um polipéptido compreendendo a sequência de aminoácidos da invenção, ou um polipéptido codificado pelo ácido nucleico da invenção; (b) cultivar a planta, a parte de planta ou a célula vegetal sob condições em que a enzima fitase é expresso; e, (c) converter a planta, partes de planta ou célula vegetal numa composição adequada para um alimento, uma ração, um suplemento alimentar, um suplemento de ração, um aditivo alimentar, um aditivo de ração, um suplemento nutritivo ou um suplemento dietético, ou adicionar a planta cultivada, parte de planta ou célula vegetal a um alimento, uma ração, um suplemento alimentar, um suplemento

de ração, um aditivo alimentar, um aditivo de ração, um suplemento nutritivo ou um suplemento dietético, produzindo assim um alimento, uma ração, um suplemento alimentar, um suplemento de ração, um aditivo alimentar, um aditivo de ração, um suplemento nutritivo ou um suplemento dietético, em que opcionalmente o polinucleótido está contido num vector de expressão e, opcionalmente, o vector compreende uma sequência de controlo de expressão capaz de expressar o ácido nucleico numa célula vegetal e, opcionalmente, o alimento, ração, suplemento alimentar, suplemento de ração, aditivo alimentar, aditivo de ração, suplemento nutritivo ou suplemento dietético é para um animal e, opcionalmente, em que o animal é um animal monogástrico e, opcionalmente, o animal é um ruminante e, opcionalmente, o alimento, ração, suplemento alimentar, suplemento de ração, aditivo alimentar, aditivo de ração, suplemento nutritivo ou suplemento dietético, está na forma de uma matriz de distribuição, uma pastilha, um comprimido, um gel, um líquido, um spray, grão moído ou um pó. Num aspecto, a enzima fitase é glicosilada para proporcionar termotolerância ou termostabilidade nas condições de formação de pastilhas e, opcionalmente, a matriz de distribuição é formada por formação de pastilhas de uma mistura compreendendo um germe de grão e a enzima fitase para produzir uma partícula e, opcionalmente, as pastilhas são feitas sob condições compreendendo a aplicação de vapor, opcionalmente as pastilhas são feitas sob condições compreendendo a aplicação de uma temperatura em excesso de 80 °C durante cerca de 5 minutos e, opcionalmente, a pastilha compreende uma enzima fitase que compreende uma actividade específica de, pelo menos, 350 a cerca de 900 unidades por miligrama de enzima.

A invenção proporciona métodos para distribuir um suplemento à enzima fitase a um animal ou humano, o referido método

compreendendo: (a) preparar uma matriz de distribuição comestível compreendendo um veículo comestível e uma enzima fitase compreendendo um polipéptido compreendendo a sequência de aminoácidos da invenção, em que a matriz prontamente se dispersa e liberta a enzima fitase quando colocada em meio aquoso e (b) administrar a matriz de distribuição comestível enzimática ao animal ou humano, em que opcionalmente a matriz de distribuição comestível compreende um veículo comestível granulado e, opcionalmente, a matriz de distribuição comestível está na forma de pastilhas, comprimidos, géis, líquidos, sprays ou pós e, opcionalmente, o veículo comestível compreende um veículo seleccionado do grupo consistindo de germe de grão, feno, alfalfa, erva-dos-prados, casca de soja, farinha de sementes de girassol, farinha de milho, farinha de soja e farinha de trigo e, opcionalmente, o veículo comestível compreende germe de grão que é resíduo de óleo.

A invenção proporciona um alimento, ração, suplemento alimentar, suplemento de ração, aditivo alimentar, aditivo de ração, suplemento nutritivo ou suplemento dietético, para um animal ou humano, compreendendo o polipéptido da invenção, ou um homodímero ou heterodímero da invenção; em que opcionalmente o polipéptido é glicosilado e, opcionalmente, a actividade de fitase é termotolerante ou termoestável. Num aspecto, o alimento, ração, suplemento alimentar, suplemento de ração, aditivo alimentar, aditivo de ração, suplemento nutritivo ou suplemento dietético é fabricado em pastilha, pílula, comprimido, cápsula, gel, geltab, spray, pó, formulação liofilizada, formulação farmacêutica, forma líquida, como uma suspensão ou pasta, ou produzido utilizando aditivos revestidos com polímero, ou fabricado na forma de granulado, ou produzido por secagem por spray.

A invenção proporciona matriz (matrizes) de distribuição enzimática comestível ou absorvível compreendendo o polipéptido da invenção, ou um homodímero ou heterodímero da invenção; em que opcionalmente o polipéptido é glicosilado e, opcionalmente, a actividade de fitase é termotolerante ou termoestável. Num aspecto, a matriz de distribuição comestível compreende uma pastilha, ou a matriz de distribuição enzimática comestível ou absorvível é fabricada em pastilha, pílula, comprimido, cápsula, gel, geltab, spray, pó, formulação liofilizada, formulação farmacêutica, forma líquida, como uma suspensão ou uma pasta, ou produzida utilizando aditivos revestidos com polímero, ou fabricada na forma de granulado, ou produzida por secagem por spray.

A invenção proporciona pastilhas comestíveis ou absorvíveis compreendendo veículo comestível ou absorvível granulado e o polipéptido da invenção, ou um homodímero ou heterodímero da invenção; em que opcionalmente o polipéptido é glicosilado e, opcionalmente, a actividade de fitase é termotolerante ou termoestável e, opcionalmente, a pastilha é fabricada na forma de pastilha, ou como um pílula, comprimido, cápsula, gel, geltab, spray, pó, formulação liofilizada, formulação farmacêutica, forma líquida, como uma suspensão ou uma pasta, ou produzida utilizando aditivos revestidos com polímero, ou fabricada na forma de granulado, ou produzida por secagem por spray.

A invenção proporciona farinhas, *e. g.*, uma farinha de soja, compreendendo um polipéptido da invenção, ou um homodímero ou heterodímero da invenção e, opcionalmente, a farinha, *e. g.*, farinha de soja, é fabricada como uma pastilha, pílula, comprimido, cápsula, gel, geltab, pulverizador, pó, formulação liofilizada ou forma líquida.

Os métodos para aumentar a resistência de um polipéptido de fitase a inactivação enzimática num sistema digestivo de um animal compreendem glicosilar um polipéptido de fitase compreendendo o polipéptido da invenção, aumentando assim a resistência do polipéptido de fitase a inactivação enzimática num sistema digestivo de um animal e, opcionalmente, a glicosilação é glicosilação ligada por N e, opcionalmente, o polipéptido de fitase é glicosilado como resultado de expressão *in vivo* de um polinucleótido que codifica a fitase numa célula e, opcionalmente, a célula é uma célula eucariótica e, opcionalmente, a célula eucariótica é uma célula fúngica, uma célula vegetal ou uma célula de mamífero.

A invenção proporciona métodos para processar sementes de milho e sorgo compreendendo os seguintes passos: (a) proporcionar um polipéptido tendo uma actividade de fitase, em que o polipéptido compreende o polipéptido da invenção; (b) proporcionar uma composição compreendendo um água de maceração de milho ou água de maceração de sorgo; e (c) colocar em contacto o polipéptido do passo (a) e a composição do passo (b) sob condições em que o polipéptido pode clivar uma ligação inositol-fosfato inorgânico.

A invenção proporciona produtos farmacêuticos ou formulações dietéticas compreendendo um polipéptido da invenção ou heterodímero; em que opcionalmente o polipéptido está glicosilado e, opcionalmente, a formulação farmacêutica ou dietética é formulado em forma de pastilha, pílula, comprimido, cápsula, geltab, spray, pó, loção ou líquido, ou produzido utilizando aditivos revestidos com polímero, ou como um implante, ou fabricado na forma de granulado, ou produzido por secagem por spray.

A invenção proporciona composições compreendendo um polipéptido da invenção ou heterodímero; e, (b) qualquer produto como apresentado na Tabela 2, ou quaisquer das composições listadas na Tabela 1; em que opcionalmente o polipéptido está glicosilado.

A invenção proporciona uma unidade Pronta-a-Comer (MRE) de refeição autónoma, uma bebida ou um agente hidratante compreendendo um polipéptido da invenção ou heterodímero; em que opcionalmente o polipéptido está glicosilado.

A invenção proporciona composições para melhorar (retardar o progresso, tratar ou prevenir) a osteoporose para administrar a um indivíduo com a sua necessidade, uma quantidade eficaz (dosagem) de uma composição compreendendo um polipéptido da invenção ou heterodímero; em que opcionalmente o polipéptido está glicosilado.

Os aspectos da invenção são apresentados nas reivindicações anexas.

Os detalhes de um ou mais aspectos da invenção são apresentados nos desenhos anexos e na descrição abaixo. Outras características, objectivos e vantagens da invenção serão evidentes a partir da descrição e desenhos, e a partir das reivindicações.

BREVE DESCRIÇÃO DOS DESENHOS

Os seguintes desenhos são ilustrativos de aspectos da invenção e não pretendem limitar o âmbito da invenção como abrangido pelas reivindicações.

A Figura 1 é um diagrama de blocos de um sistema computorizado, como descrito em detalhe, abaixo.

A Figura 2 é um fluxograma que ilustra um aspecto de um processo 200 para comparar uma nova sequência de proteína ou nucleótidos com uma base de dados de sequências, de modo a determinar os níveis de homologia entre a nova sequência e as sequências na base de dados, como descrito em detalhe, abaixo.

A Figura 3 é um fluxograma que ilustra uma forma de realização de um processo num computador para determinar se duas sequências são homólogas, como descrito em detalhe, abaixo.

A Figura 4 é um fluxograma que ilustra um aspecto de um processo identificador para detectar a presença de uma característica numa sequência, como descrito em detalhe, abaixo.

A Figura 5 ilustra um sumário da actividade residual de polipéptidos purificados da invenção, sequências exemplificativas de mutação única da invenção, após tratamento térmico a várias temperaturas durante 30 minutos; em que a actividade de fitase é ensaiada com um substrato de fluorescência e as velocidades foram comparadas às velocidades de cada amostra não tratada correspondente; como descrito em detalhe no Exemplo 1, abaixo.

A Figura 6 ilustra um sumário da actividade residual de polipéptidos purificados da invenção compreendendo mutações de resíduo único "misturadas" (fitases contendo múltiplas mutações), após tratamento térmico num termociclador; em que a actividade de fitase é ensaiada com um substrato de fluorescência e as velocidades foram comparadas às velocidades de cada amostra não tratada correspondente; como resumido em Figura 6; como descrito em detalhe no Exemplo 1, abaixo.

A Figura 7 sumaria graficamente os dados utilizados para gerar o gráfico da Figura 6; como descrito em detalhe no Exemplo 1, abaixo.

A Figura 8 ilustra a sequência da fitase parental SEQ ID N°:2 e as modificações de sequência produzidos por mutagénese de saturação de local génico (GSSM) seleccionadas para a construção da biblioteca GeneReassembly™; como descrito em detalhe, abaixo.

A Figura 9 ilustra fitases exemplificativas tendo modificações múltiplas de resíduo relativamente à SEQ ID N°:2 parental; como aqui descrito em detalhe.

A Figura 10 ilustra fitases exemplificativas tendo modificações únicas de resíduo relativamente à SEQ ID N°:2 parental; como aqui descrito em detalhe.

A Figura 11 ilustra esquematicamente um ensaio de fitase exemplificativo da invenção utilizando o substrato de fluorescência fosfato de 4-metilumbeliferilo (MeUMB-fosfato); como descrito em detalhe no Exemplo 1, abaixo.

A Figura 12 ilustra esquematicamente um ensaio de fitase exemplificativo da invenção que utiliza o substrato de fluorescência MeUMB-fosfato; como descrito em detalhe no Exemplo 1, abaixo.

A Figura 13 ilustra esquematicamente o protocolo para um rastreio de biblioteca exemplificativo, como descrito na Figura 12, como descrito em detalhe no Exemplo 1, abaixo.

A Figura 14 ilustra um processo alcoólico exemplificativo do que pode incorporar a utilização de fitases desta invenção.

Símbolos de referência iguais nos vários desenhos indicam elementos iguais.

DESCRIÇÃO DETALHADA DA INVENÇÃO

A invenção refere-se a polipéptidos de fitase compreendendo as modificações de resíduo específicas para a SEQ ID N°:2, como descrito acima, e aos polinucleótidos que os codificam, (*e. g.*, compreendendo as modificações de pares de base específicas para a SEQ ID N°:1, como descrito acima), assim como métodos de utilização dos polinucleótidos e polipéptidos. A Figura 10 ilustra fitases exemplificativas tendo modificações únicas de resíduo relativamente à SEQ ID N°:2 parental e a Figura 9 ilustra fitases exemplificativas tendo modificações múltiplas de resíduo relativamente à SEQ ID N°:2 parental.

A actividade de fitase de polipéptidos da invenção pode abranger enzimas tendo qualquer actividade de fitase, por exemplo, enzimas capazes de catalisar a degradação de fitato, *e. g.*, a catálise do fitato (mio-inositol-hexafosfato) em

inositol e fosfato inorgânico. As fitases da invenção incluem enzimas termotolerantes e termorresistentes.

As fitases e os polinucleótidos que codificam as fitases da invenção são úteis numa variedade de processos, métodos e composições. Por exemplo, como discutido acima, uma fitase pode ser utilizada numa ração animal e em suplementos de ração, assim como em tratamentos para degradar ou remover o excesso de fitato do ambiente ou numa amostra. Outras utilizações serão evidentes aos especialistas na técnica com base nos ensinamentos aqui proporcionados, incluindo aquelas discutidos acima.

Num aspecto, moléculas de fitase da invenção – isoladamente ou em combinação com outros reagentes (incluindo, mas não limitado a, enzimas, tais como proteases, amilases e semelhantes) – são utilizadas no processamento de géneros alimentícios, *e. g.*, para a prevenção de sedimentos de milho não desejados e noutras aplicações em que é desejável a hidrólise de fitato.

Num aspecto, as moléculas de fitase da invenção são utilizadas para eliminar ou diminuir a presença de fitato não hidrolisado, especialmente nos casos em que o fitato não hidrolisado conduz a consequências problemáticas em processos *ex vivo* incluindo – mas não limitado – ao processamento de géneros alimentícios. Num aspecto, as moléculas de fitase da invenção são utilizadas em processos como descritos no documento EP0321004-B1 (Vaara *et al.*), incluindo passos no processamento de sementes de milho e sorgo pelo qual as sementes duras são embebidas em água para as amolecer. As substâncias hidrossolúveis lixiviadas durante este processo tornam-se parte de uma água de maceração de milho que é concentrada por evaporação. O ácido fítico não hidrolisado na água de maceração

de milho, maioritariamente na forma de sais de cálcio e magnésio, está associado com fósforo e deposita-se um sedimento indesejável com proteínas e iões metálicos. Este sedimento é problemático na evaporação, transporte e armazenamento da água de maceração de milho. As moléculas de fitase da invenção são utilizadas para hidrolisar este sedimento.

Num aspecto, as fitases da invenção podem proporcionar um desempenho comercial substancialmente superior do que moléculas de fitase previamente identificadas, *e. g.*, moléculas de fitase de origem fúngica. Num aspecto, as enzimas da invenção podem ter aproximadamente, 4400 U/mg, ou mais do que, aproximadamente, entre 50 a 100, ou 50 a 1000, ou 100 a 4000 U/mg de proteína.

A invenção também proporciona métodos para alterar as características de uma fitase da invenção por mutagénese e outro método, como discutido em detalhe, abaixo.

Gerar e Manipular Ácidos Nucleicos

A invenção proporciona ácidos nucleicos que codificam os polipéptidos e fitases da invenção. A invenção também proporciona cassetes de expressão, vectores, tais como vectores de expressão ou clonagem, veículos de clonagem, tais como um vector viral, um plasmídeo, um fago, um fagemídeo, um cosmídeo, um fosmídeo, um bacteriófago ou um cromossoma artificial, que podem compreender ou aí conter, um ácido nucleico da invenção.

A divulgação inclui métodos para identificar novas sequências de fitase utilizando os ácidos nucleicos da invenção. Os métodos para modificar os ácidos nucleicos da invenção

incluem, *e. g.*, remontagem de ligação sintética, sistema de evolução dirigida optimizado e/ou mutagénese de saturação.

Num aspecto, a invenção proporciona um género de ácidos nucleicos que são variantes sinteticamente produzidas da SEQ ID N°:1 parental, em que estes ácidos nucleicos da invenção têm, pelo menos, 95%, 96%, 97%, 98% ou 99% de identidade de sequência relativamente à SEQ ID N°:1 "parental" e compreendem uma específica e, pelo menos, uma, duas, três, quatro, cinco, seis, sete, oito, nove, dez, onze, doze ou todas as treze modificações adicionais de sequência de pares de base nucleotídicas para a SEQ ID N°:1, como apresentadas acima. Para referência, a SEQ ID N°:1 parental divulgada no documento U.S. 2005/246780 é:

atgaaagcga tcttaatccc atttttatct cttctgattc cgttaacccc gcaatctgca 60
ttcgctcaga gtgagccgga gctgaagctg gaaagtgtgg tgattgtcag tcgtcatggt 120
gtgcgtgctc caaccaaggc cacgcaactg atgcaggatg tcaccccaga cgcatggcca 180
acctggccgg taaaactggg tgagctgaca ccgcgcggtg gtgagctaat cgcctatctc 240
ggacattact ggcgtcagcg tctggtagcc gacggattgc tgcctaaatg tggctgcccg 300
cagtctggtc aggtcgcgat tattgctgat gtcgacgagc gtacccgtaa aacaggcgaa 360
gccttcgccg ccgggctggc acctgactgt gcaataaccg tacataccca ggcagatacg 420
tccagtcccg atccgttatt taatcctcta aaaactggcg tttgccaact ggataacgcg 480
aacgtgactg acgcgatcct cgagagggca ggagggtcaa ttgctgactt taccgggcat 540
tatcaaacgg cgtttcgcga actggaacgg gtgcttaatt ttccgcaatc aaacttgtgc 600
cttaaacgtg agaaacagga cgaaagctgt tcattaacgc aggcattacc atcggaactc 660
aaggtgagcg ccgactgtgt ctcattaacc ggtgcggtaa gcctcgcatc aatgctgacg 720
gagatatttc tcctgcaaca agcacaggga atgccggagc cggggtgggg aaggatcacc 780
gattcacacc agtggaacac cttgctaagt ttgcataacg cgcaatttga tttgctacaa 840
cgcacgccag aggttgcccg cagccgcgcc accccgttat tagatttgat caagacagcg 900
ttgacgcccc atccaccgca aaaacaggcg tatggtgtga cattacccac ttcagtgctg 960
tttatcgccg gacacgatac taatctggca aatctcggcg gcgcactgga gctcaactgg 1020
acgcttcccg gtcagccgga taacacgccg ccaggtggtg aactggtgtt tgaacgctgg 1080
cgtcggctaa gcgataacag ccagtggatt caggtttcgc tggtcttcca gactttacag 1140
cagatgcgtg ataaaacgcc gctgtcatta aatacgccgc ccggagaggt gaaactgacc 1200
ctggcaggat gtgaagagcg aaatgcgcag ggcatgtgtt cgttggcagg ttttacgcaa 1260
atcgtgaatg aagcacgcat accggcgtgc agtttgtaa 1299

Os ácidos nucleicos da invenção podem ser feitos, isolados e/ou manipulados, *e. g.*, por clonagem e expressão de bibliotecas de ADNc, amplificação de ADN genómico ou mensageiro por PCR, e semelhantes. Na prática dos métodos, os genes homólogos podem ser modificados por manipulação de um ácido nucleico molde, como aqui descrito. A invenção pode ser realizada, em conjunto, com qualquer método ou protocolo ou dispositivo conhecido na técnica, os quais estão bem descritos na literatura científica e de patente.

Técnicas Gerais

Os ácidos nucleicos utilizados para praticar esta invenção, seja ARN, iARN, ácido nucleico anti-sentido, ADNc, ADN genómico, vectores, vírus ou os seus híbridos, podem ser isolados a partir de uma variedade de fontes, manipulados geneticamente, amplificados e/ou expressos/produzidos recombinantemente. Os polipéptidos recombinantes gerados a partir destes ácidos nucleicos podem ser isolados ou clonados e testados individualmente para uma actividade desejada. Qualquer sistema de expressão recombinante pode ser utilizado, incluindo sistemas de expressão bacterianos, de mamíferos, levedura, insecto ou célula vegetal.

Alternativamente, estes ácidos nucleicos podem ser sintetizados *in vitro* por técnicas de síntese química bem conhecidas, como descrito em, *e. g.*, Adams (1983) J. Am. Chem. Soc. 105:661; Belousov (1997) Nucleic Acids Res. 25:3440-3444; Frenkel (1995) Free Radic. Biol. Med. 19:373-380; Blommers (1994) Biochemistry 33:7886-7896; Narang (1979) Meth. Enzymol. 68:90; Brown (1979) Meth. Enzymol.

68:109; Beaucage (1981) Tetra. Lett. 22:1859; Patente U.S. Nº 4458066.

As técnicas para manipulação de ácidos nucleicos, tais como, *e. g.*, subclonagem, marcação de sondas (*e. g.*, marcação de iniciador aleatória utilizando polimerase Klenow, reparação de ligação, amplificação), sequenciação, hibridação e semelhantes são bem descritas na literatura científica e de patente, ver, *e. g.*, Sambrook, ed., MOLECULAR CLONING: A LABORATORY MANUAL (2ª ED.), Vols. 1-3, Cold Spring Harbor Laboratory, (1989); CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, Ausubel, ed. John Wiley & Sons, Inc., Nova Iorque (1997); LABORATORY TECHNIQUES IN BIOCHEMISTRY AND MOLECULAR BIOLOGY: HYBRIDIZATION WITH NUCLEIC ACID PROBES, Part I. Theory and Nucleic Acid Preparation, Tijssen, ed. Elsevier, N.I. (1993).

Outros meios úteis para obter e manipular os ácidos nucleicos utilizados para praticar os métodos são a clonagem a partir de amostras genómicas e, se desejado, rastreio e reclonagem das inserções isoladas ou amplificadas a partir de, *e. g.*, clones genómicos ou clones de ADNc. As fontes de ácidos nucleicos utilizadas nos métodos incluem bibliotecas genómicas ou de ADNc contidas, *e. g.*, em cromossomas artificiais de mamífero (mAC), ver, *e. g.*, Patente U.S. Nº 5721118; 6025155; cromossomas artificiais humanos, ver, *e. g.*, Rosenfeld (1997) Nat. Genet. 15:333-335; cromossomas artificiais de levedura (YAC); cromossomas artificiais bacterianos (BAC); cromossomas artificiais P1, ver, *e. g.*, Woon (1998) Genomics 50:306-316; vectores derivados de P1 (PAC), ver, *e. g.*, Kern (1997) Biotechniques 23:120-124; cosmídeos, vírus recombinantes, fagos ou plasmídeos.

Em aspectos alternativos, as frases "ácido nucleico" ou "sequência de ácido nucleico" refere-se a um oligonucleótido, nucleótido, polinucleótido ou a um fragmento de qualquer destes, a ADN ou ARN (*e. g.*, ARNm, rARN, tARN) de origem genómica ou sintética que podem ser de cadeia simples ou cadeia dupla e podem representar uma cadeia de sentido ou anti-sentido, a ácido nucleico peptídico (PNA) ou a qualquer material de tipo ADN ou de tipo ARN, de origem natural ou sintética, incluindo, *e. g.*, iARN, ribonucleoproteínas (*e. g.*, iRNP). O termo abrange ácidos nucleicos, *i. e.*, oligonucleótidos, contendo análogos conhecidos de nucleótidos naturais. O termo abrange também estruturas de tipo ácido nucleico com estruturas sintéticas, ver, *e. g.*, Mata (1997) Toxicol. Appl. Pharmacol. 144:189-197; Strauss-Soukup (1997) Biochemistry 36:8692-8698; Samstag (1996) Antisense Nucleic Acid Drug Dev 6:153-156.

Num aspecto, os polinucleótidos recombinantes da invenção compreendem sequências adjacentes a um ácido nucleico de "estrutura" ao qual não está adjacente no seu ambiente natural. Num aspecto, os ácidos nucleicos representam 5% ou mais do número de inserções de ácido nucleico numa população de "moléculas estruturais" de ácidos nucleicos. As "moléculas estruturais" de acordo com a invenção incluem ácidos nucleicos, tais como vectores de expressão, ácidos nucleicos auto-replicantes, vírus, ácidos nucleicos de integração e outros vectores ou ácidos nucleicos utilizados para manter ou manipular uma inserção de ácido nucleico de interesse. Num aspecto, os ácidos nucleicos enriquecidos representam 15%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90% ou mais do número de inserções de ácidos nucleicos na população de moléculas estruturais recombinantes.

Num aspecto, um ácido nucleico que codifica um polipéptido da invenção é montado em fase apropriada com uma sequência líder

capaz de dirigir a secreção do polipéptido traduzido ou de um seu fragmento.

A invenção proporciona proteínas de fusão e ácidos nucleicos que as codificam. Um polipéptido invenção pode ser fundido a um péptido ou polipéptido heterólogo, tal como péptidos de identificação N-terminal que proporcionam características desejadas, tais como estabilidade aumentada ou purificação simplificada. Os péptidos e polipéptidos da invenção podem ser também sintetizados e expressos como proteínas de fusão com um ou mais domínios adicionais aí ligados para, *e. g.*, produzir um péptido mais imunogénico, para mais prontamente isolar um péptido recombinantemente sintetizado, para identificar e isolar anticorpos e célula B que expressam anticorpos e semelhantes. Os domínios que facilitam a detecção e purificação incluem, *e. g.*, péptidos quelantes de metal, tais como fragmentos de poli-histidina e módulos de histidina-triptofano que permitem a purificação em metais imobilizados, domínios de proteína A que permitem a purificação em imunoglobulinas imobilizadas e o domínio utilizado no sistema de purificação de extensão/afinidade FLAGS (Immunex Corp, Seattle WA). A inclusão de sequências ligantes cliváveis, tais como Factor Xa ou enterocinase (Invitrogen, San Diego CA) entre um domínio de purificação e o péptido ou polipéptido compreendendo o motivo para facilitar a purificação. Por exemplo, um vector de expressão pode incluir uma sequência de ácido nucleico que codifica epitopo ligada a seis resíduos de histidina, seguidos por uma tioredoxina e um local de clivagem de enterocinase (ver, *e. g.*, Williams (1995) biochemistry 34:1787-1797; Dobeli (1998) Protein Expr. Purif. 12:404-414). Os resíduos de histidina facilitam a detecção e purificação enquanto o local de clivagem de enterocinase proporciona meios para purificar o epitopo a partir do remanescente da proteína de fusão. A tecnologia

referente a vectores que codificam proteínas de fusão e a aplicação de proteínas de fusão são bem descritas na literatura científica e de patente, ver, *e. g.*, Kroll (1993) DNA Cell. Biol., 12:441-53.

Num aspecto, o termo "isolado" significa que o material é removido do seu ambiente original (*e. g.*, o ambiente natural se for de ocorrência natural). Por exemplo, um polipéptido ou polinucleótido de ocorrência natural presente num animal vivo não está isolado, mas o mesmo polinucleótido ou polipéptido, separado de alguns ou todos os materiais coexistentes no sistema natural, está isolado. Tais polinucleótidos podem ser parte de um vector e/ou tais polinucleótidos ou polipéptidos podem ser parte de uma composição e mesmo assim estar isolado uma vez que tal vector ou composição não são parte do seu ambiente natural. Num aspecto, o termo "purificado" não requer pureza absoluta; mais propriamente, é pretendido como uma definição relativa. Os ácidos nucleicos individuais obtidos a partir de uma biblioteca foram convencionalmente purificados até à homogeneidade electroforética. As sequências obtidas a partir destes clones não podiam ser obtidas directamente a partir da biblioteca ou a partir de ADN total humano. Os ácidos nucleicos purificados da invenção foram purificados a partir do remanescente do ADN genómico no organismo em, pelo menos, 10^4-10^6 vezes. Em aspectos alternativos, o termo "purificado" inclui também os ácidos nucleicos que foram purificados a partir do remanescente do ADN genómico ou a partir de outras sequências numa biblioteca ou outro ambiente em, pelo menos, uma ordem de magnitude, ou alternativamente, duas ou três ordens, ou quatro ou cinco ordens de magnitude.

Sequências de controlo transcripcional e de tradução

A invenção proporciona sequências de ácidos nucleicos (*e. g.*, ADN) da invenção ligadas operativamente a sequência(s) de controlo de expressão (*e. g.*, transcripcional ou de tradução), *e. g.*, promotores ou intensificadores, para dirigir ou modular a síntese/expressão de ARN. A sequência de controlo de expressão pode ser um vector de expressão. Os promotores bacterianos exemplificativos incluem lacI, lacZ, T3, T7, gpt, lambda PR, PL e trp. Os promotores eucarióticos exemplificativos incluem o precoce imediato de CMV, timidina cinase de HSV, SV40 precoce e tardio, LTR de retrovírus e metalotioneína I de murganho.

Os promotores adequados para expressar ou sobreexpressar um polipéptido em bactérias incluem os promotores lac ou trp de *E. coli*, o promotor lacI, o promotor lacZ, o promotor T3, o promotor T7, o promotor gpt, o promotor lambda PR, o promotor lambda PL, os promotores de operões que codificam enzimas glicolíticas, tal como a 3-fosfoglicerato cinase (PGK), e o promotor da fosfatase ácida. Os promotores eucarióticos incluem o promotor precoce imediato de CMV, o promotor da timidina cinase de HSV, promotores de choque térmico, o promotor SV40 precoce e tardio, LTR de retrovírus e o promotor de metalotioneína-I de murganho. Podem também ser utilizados outros promotores conhecidos por controlarem a expressão de genes em célula procarióticas ou eucarióticas ou nos seus vírus.

Vectores de expressão e veículos de clonagem

A invenção proporciona sistemas de expressão, *e. g.*, cassetes de expressão, vectores, veículos de clonagem e

semelhantes, compreendendo ácidos nucleicos da invenção, *e. g.*, sequência que codificam as fitases da invenção, para expressão e sobreexpressão dos polipéptidos da invenção (e ácidos nucleicos, *e. g.*, anti-sentido). Os vectores de expressão e veículos de clonagem da invenção podem compreender partículas virais, baculovírus, fagos, plasmídeos, fagemídeos, cosmídeos, fosmídeos, cromossomas artificiais bacterianos, ADN viral (*e. g.*, vaccinia, adenovírus, vírus da varíola aviária, pseudorraiva e derivados de SV40), cromossomas artificiais baseados em P1, plasmídeos de levedura, cromossomas artificiais de levedura e quaisquer outros vectores específicos para hospedeiros de interesse específicos (tais como bacilos, aspergilos e levedura). Os vectores da invenção podem incluir sequências de ADN cromossomal, não cromossomal e sintético. Um grande número vectores adequados é conhecido dos especialistas na técnica e estão comercialmente disponíveis. Os vectores exemplificativos incluem: bacterianos: vectores pQE (Qiagen), plasmídeos pBluescript, vectores pNH, vectores lambda-ZAP (Stratagene); ptrc99a, pKK223-3, pDR540, pRIT2T (Pharmacia); Eucarióticos: pXT1, pSG5 (Stratagene), pSVK3, pBPV, pMSG, pSVLSV40 (Pharmacia). Contudo, qualquer outro plasmídeo ou outro vector pode ser utilizado desde que seja replicável e viável no hospedeiro. Podem ser utilizados com a presente invenção vectores com número baixo de cópias ou com número elevado de cópias.

Como exemplos representativos de vectores de expressão que podem ser utilizados, podem ser mencionadas partículas virais, baculovírus, fagos, plasmídeos, fagemídeos, cosmídeos, fosmídeos, cromossomas artificiais bacterianos, ADN viral (*e. g.*, vaccinia, adenovírus, vírus da varíola aviária, pseudorraiva e derivados de SV40), cromossomas artificiais baseados em P1, plasmídeos de levedura, cromossomas artificiais

de levedura e quaisquer outros vectores específicos para hospedeiros de interesse específicos (tais como bacilos, aspergilos e levedura). Assim, por exemplo, o ADN pode ser incluído em qualquer de uma variedade de vectores de expressão para expressar um polipéptido. Tais vectores incluem sequências de ADN cromossomal, não cromossomal e sintético. Um grande número de vectores adequados é conhecido dos especialistas na técnica e estão comercialmente disponíveis. Os seguintes vectores são proporcionados a título de exemplo; Bacteriano: vectores pQE (Qiagen), plasmídeos pBluescript, vectores pNH, vectores lambda-ZAP (Stratagene); ptrc99a, pKK223-3, pDR540, pRIT2T (Pharmacia); Eucarióticos: pXT1, pSG5 (Stratagene), pSVK3, pBPV, pMSG, pSVLSV40 (Pharmacia). Contudo, qualquer outro plasmídeo ou outro vector pode ser utilizado desde que seja replicável e viável no hospedeiro. Podem ser utilizados com a presente invenção vectores com baixo número de cópias ou com elevado número de cópias.

Um vector exemplificativo para utilização na presente invenção contém uma origem de replicação de factor f. O factor f (ou factor de fertilidade) em *E. coli* é um plasmídeo que resulta em transferência de alta frequência de si próprio durante a conjugação e transferência de menor frequência do próprio cromossoma bacteriano. Um aspecto utiliza vectores de clonagem, referidos como “fosmídeos” ou vectores de cromossomas artificiais bacterianos (BAC). Estes são derivados do factor f de *E. coli* que é capaz de integrar estavelmente segmentos grandes de ADN genómico. Quando integrado com ADN de uma amostra ambiental mista não cultivada, isto torna possível conseguir grandes fragmentos genómicos na forma de uma “biblioteca de ADN ambiental” estável.

Outro tipo de vector para utilização na presente invenção é um vector cosmídeo. Os vectores cosmídeos foram concebidos originalmente para clonar e propagar grandes segmentos de ADN genómico. A clonagem em vectores cosmídeos é descrita em detalhe em "Molecular Cloning: A laboratory Manual" (Sambrook *et al.*, 1989).

A sequência de ADN no vector de expressão está ligada operativamente a uma ou mais sequências apropriadas de controlo de expressão (promotor) para dirigir a síntese de ARN. Os promotores bacterianos particularmente nomeados incluem *lacI*, *lacZ*, *T3*, *T7*, *gpt*, *lambda* P_R, P_L e *trp*. Os promotores eucarióticos incluem imediato precoce de CMV, timidina cinase de HSV, precoce e tardio de SV40, LTR de retrovírus e metalotioneína-I de murganho. A selecção do vector e promotor apropriado está no âmbito de um técnico com conhecimento geral na matéria. O vector de expressão contém também um local de ligação a ribossoma para iniciação da tradução e um terminador de transcrição. O vector pode também incluir sequências apropriadas para amplificar a expressão. As regiões promotoras podem ser seleccionadas de qualquer gene desejado utilizando vectores CAT (cloranfenicol transferase) ou outros vectores com marcadores de selecção. Além disso, os vectores de expressão podem conter um ou mais genes marcadores de selecção para proporcionar uma característica fenotípica para selecção de célula hospedeiras transformadas, tais como di-hidrofolato redutase ou resistência à neomicina para a cultura de células eucarióticas, ou resistência à tetraciclina ou ampicilina em *E. coli*.

Num aspecto, as cassetes de expressão da invenção compreendem uma sequência da invenção e uma sequência de nucleótidos que são capazes de afectar a expressão de um gene

estrutural (*i. e.*, uma sequência codificante de proteína, tal como uma fitase da invenção) num hospedeiro compatível com tais sequências. As cassetes de expressão incluem, pelo menos, um promotor ligado operativamente com a sequência codificante do polipéptido; e, opcionalmente, com outras sequências, *e. g.*, sinais de terminação de transcrição. Factores adicionais necessários ou úteis para efectuar a expressão podem também ser utilizados, *e. g.*, intensificadores. Num aspecto, "ligado operativamente", como aqui utilizado, refere-se à ligação de um promotor a montante de uma sequência de ADN de modo que o promotor medeie a transcrição da sequência de ADN. Assim, as cassetes de expressão incluem também plasmídeos, vectores de expressão, vírus recombinantes, qualquer forma de vector recombinante de "ADN nu" e semelhantes. Num aspecto, um "vector" compreende um ácido nucleico que pode infectar, transfectar, transiente ou permanentemente transduzir uma célula. Um vector pode ser um ácido nucleico nu ou um ácido nucleico complexado com proteína ou lípido. O vector compreende opcionalmente ácidos nucleicos e/ou proteínas virais ou bacterianas e/ou membranas (*e. g.*, uma membrana celular, um envelope lipídico viral, etc.). Num aspecto, os vectores incluem, mas não estão limitados a, replicões (*e. g.*, replicões de ARN, bacteriófagos) aos quais podem ser ligados fragmentos de ADN e tornarem-se replicados. Num aspecto, os vectores incluem, mas não estão limitados, a ARN, ADN ou ARN circular ou linear auto-replicante autónomo (*e. g.*, plasmídeos, vírus e semelhantes, ver, *e. g.*, Patente U.S. N° 5217879) e incluem plasmídeos de expressão e de não expressão. Nos casos em que uma cultura celular ou microrganismo recombinante é descrito como hospedando um "vector de expressão" isto inclui ADN circular e linear extra-cromossomal e ADN que foi incorporado no(s) cromossoma(s) do hospedeiro. Nos casos em que um vector é mantido por uma célula hospedeira, o vector pode

ser estavelmente replicado pelas células durante a mitose como uma estrutura autónoma ou é incorporado no genoma do hospedeiro.

O vector de expressão pode compreender um promotor, um local de ligação a ribossoma para iniciação da tradução e um terminador de transcrição. O vector pode também incluir sequências apropriadas para amplificar a expressão. Os vectores de expressão de mamífero podem compreender uma origem de replicação, quaisquer locais necessários de ligação a ribossoma, um local de poliadenilação, locais dadores e aceitadores de *splice*, sequências de terminação de transcrição e sequências não transcritas que flanqueiam 5'. Em alguns aspectos, as sequências de ADN derivadas dos locais de *splice* e de poliadenilação de SV40 podem ser utilizadas para proporcionar os elementos genéticos não transcritos requeridos.

Num aspecto, os vectores de expressão contêm um ou mais genes marcadores de selecção para permitir a selecção de células hospedeiras contendo o vector. Tais marcadores de selecção incluem genes que codificam a di-hidrofolato redutase ou genes que conferem resistência à neomicina para culturas celulares eucarióticas, genes que conferem resistência à tetraciclina ou ampicilina em *E. coli* e o gene TRP1 de *S. cerevisiae*. As regiões promotoras podem ser seleccionadas de qualquer gene desejado utilizando vectores de cloranfenicol transferase (CAT) ou outros vectores com marcadores de selecção.

Os vectores para expressar o polipéptido ou um seu fragmento em células eucarióticas podem também conter intensificadores para aumentar os níveis de expressão. Os intensificadores são elementos de actuação cis de ADN, habitualmente desde cerca de 10 a cerca de 300 pb de comprimento que actuam num promotor para aumentar a sua transcrição. Os exemplos incluem o intensificador

SV40 no lado tardio da origem de replicação pb 100 a 270, o intensificador do promotor precoce de citomegalovírus, o intensificador de polioma no lado tardio da origem de replicação e os intensificadores de adenovírus.

Uma sequência de ADN pode ser introduzida num vector através de uma variedade de processos. Em geral, a sequência de ADN é ligada à posição desejada no vector, após digestão da inserção e do vector com as endonucleases de restrição apropriadas. Alternativamente, podem ser ligadas extremidades rombas na inserção e no vector. Uma variedade de técnicas de clonagem é conhecida na técnica, *e. g.*, como descrito em Ausubel e Sambrook. Considera-se que tais processos e outros estão no âmbito dos especialistas na técnica.

O vector pode ser na forma de um plasmídeo, uma partícula viral ou um fago. Outros vectores incluem sequências de ADN cromossomal, não cromossomal e sintético, derivados de SV40; plasmídeos bacterianos, ADN fágico, baculovírus, plasmídeos de levedura, vectores derivados de combinações de plasmídeos e de ADN fágico, ADN viral, tais como de vaccinia, adenovírus, vírus da varíola aviária e pseudorraiva. Uma variedade de vectores de clonagem e de expressão para utilização com hospedeiros procarióticas e eucarióticas é descrito por, *e. g.*, Sambrook.

Vectores bacterianos particulares que podem ser utilizados incluem os plasmídeos comercialmente disponíveis compreendendo elementos genéticos do bem conhecido vector de clonagem pBR322 (ATCC 37017), pKK223-3 (Pharmacia Fine Chemicals, Uppsala, Suécia), GEM1 (Promega Biotec, Madison, WI, EUA) pQE70, pQE60, pQE-9 (Qiagen), pD10, psiX174 pBluescript II KS, pNH8A, pNH16a, pNH18A, pNH46A (Stratagene), ptrc99a, pKK223-3, pKK233-3, pDR540, pRIT5 (Pharmacia), pKK232-8 e pCM7. Vectores

eucarióticos particulares incluem pSV2CAT, pOG44, pXT1, pSG (Stratagene) pSVK3, pBPV, pMSG e pSVL (Pharmacia). Contudo, qualquer outro vector pode ser utilizado desde que seja replicável e viável na célula hospedeira.

Células hospedeiras e célula transformadas

A invenção também proporciona uma célula transformada compreendendo uma sequência de ácido nucleico da invenção, *e. g.*, uma sequência codificando uma fitase da invenção, ou compreendendo uma cassete de expressão, vector, veículo de clonagem, vector de expressão, vector de clonagem da invenção. A célula hospedeira pode ser quaisquer células hospedeiras familiares dos especialistas na técnica, incluindo células procarióticas, células eucarióticas, tais como células bacterianas, células fúngicas, células de levedura, células de mamífero, células de insecto ou células vegetais. As células bacterianas exemplificativas incluem quaisquer espécies dentro dos géneros *Escherichia, Bacillus, Streptomyces, Salmonella, Pseudomonas* e *Staphylococcus,* incluindo, *e. g., Escherichia coli, Lactococcus lactis, Bacillus subtilis, Bacillus cereus, Salmonella typhimurium, Pseudomonas fluorescens.* As células fúngicas exemplificativas incluem quaisquer espécies de *Aspergillus.* As células de levedura exemplificativas incluem quaisquer espécies *de Pichia, Saccharomyces, Schizosaccharomyces* ou *Schwanniomyces,* incluindo *Pichia pastoris, Saccharomyces cerevisiae* ou *Schizosaccharomyces pombe.* Células de insecto exemplificativas incluem quaisquer espécies de *Spodoptera* ou *Drosophila,* incluindo *Drosophila S2* e *Spodoptera Sf9.* Células de insecto exemplificativas incluem *Drosophila* S2 e *Spodoptera* Sf9. Células de levedura exemplificativas incluem *Pichia pastoris,*

Saccharomyces cerevisiae ou *Schizosaccharomyces pombe*. Células animais exemplificativas incluem CHO, COS ou melanoma de Bowes ou qualquer linha celular de murganho ou humana. A selecção de um hospedeiro apropriado está dentro das capacidades dos especialistas na técnica.

O vector pode ser introduzido nas células hospedeiras utilizando qualquer de uma variedade de técnicas, incluindo transformação, transfecção, transdução, infecção viral, pistolas génicas ou transferência génica mediada por Ti. Os métodos particulares incluem transfecção de fosfato de cálcio, transfecção mediada por DEAE-Dextrano, lipofecção ou electroporação (Davis, L., Dibner, M., Battey, I., Basic Methods in Molecular Biology, (1986)).

Sempre que apropriado, as células hospedeiras manipuladas podem ser cultivadas em meio nutriente convencional modificado como apropriado para activar promotores, seleccionar transformantes ou amplificar os genes da invenção. Após a transformação de uma estirpe hospedeira adequada e crescimento da estirpe hospedeira a uma densidade celular apropriada, o promotor seleccionado pode ser induzido por meios apropriados (*e. g.*, desvio de temperatura ou indução química) e as células podem ser cultivadas por um período adicional para permitir que produzam o polipéptido desejado ou um seu fragmento.

As células podem ser recolhidas por centrifugação, quebradas por meios químicos ou físicos e o extracto em bruto resultante é retido para purificação adicional. As células microbianas utilizadas para expressão de proteínas podem ser quebradas através de qualquer método conveniente, incluindo ciclos de congelamento-descongelamento, sonicação, quebra mecânica ou utilização de agentes de lise celular. Tais métodos são bem

conhecidos dos especialistas na técnica. O polipéptido expresso ou um seu fragmento pode ser recuperado e purificado a partir de culturas celular recombinantes por métodos incluindo precipitação com sulfato de amónio ou ethanol, extracção ácida, cromatografia de permuta aniónica ou catiónica, cromatografia de fosfocelulose, cromatografia de interacção hidrofóbica, cromatografia de afinidade, cromatografia de hidroxilapatite e cromatografia de lecitina. Podem ser utilizados passos de re-enrolamento proteico, como necessário, para completar a configuração do polipéptido. Se desejado, pode ser utilizada cromatografia líquida de elevado desempenho (HPLC) para passos finais de purificação.

As construções em células hospedeiras podem ser utilizadas de um modo convencional para produzir o produto génico codificado pela sequência recombinante. Dependendo do hospedeiro utilizado num processo de produção recombinante, os polipéptidos produzidos por células hospedeiras contendo o vector podem ser glicosilados ou pode ser não glicosilados. Os polipéptidos da invenção podem ou não também incluir um resíduo de aminoácido de metionina inicial.

Podem também ser utilizados sistemas de tradução sem célula para produzir um polipéptido da invenção. Os sistemas de tradução sem célula podem utilizar mARN transcritos a partir de uma construção de ADN compreendendo um promotor ligado operativamente a um ácido nucleico codificando o polipéptido ou um seu fragmento. Em alguns aspectos, a construção de ADN pode ser linearizada antes de conduzir uma reacção de transcrição *in vitro*. O ARNm transcrito é depois incubado com um extracto de tradução sem célula apropriado, tal como um extracto de reticulócito de coelho, para produzir o polipéptido desejado ou um seu fragmento.

Os vectores de expressão podem conter um ou mais genes marcadores de selecção para proporcionar uma característica fenotípica para selecção de células hospedeiras transformadas, tais como di-hidrofolato redutase ou resistência à neomicina, para cultura de células eucarióticas ou como resistência à tetraciclina ou ampicilina em *E. coli*.

Os ácidos nucleicos da invenção podem ser expressos ou sobreexpressos em qualquer sistema de expressão *in vitro* ou *in vivo*. Quaisquer sistemas de cultura celular podem ser utilizados para expressar ou sobreexpressar proteína recombinante, incluindo culturas bacterianas, de insecto, de levedura, fúngicas ou de mamíferos. A sobreexpressão pode ser realizada por escolha apropriada de promotores, intensificadores, vectores (*e. g.*, utilização de vectores de replicão, vectores dicistrónicos (ver, *e. g.*, Gurtu (1996) Biochem. Biophys. Res. Commun. 229:295-8)), meios, sistemas de cultura e semelhantes. Num aspecto, a amplificação génica utilizando marcadores de selecção, *e. g.*, glutamina sintetase (ver, *e. g.*, Sanders (1987) Dev. Biol. Stand. 66:55-63), em sistemas celulares são utilizados para sobreexpressar os polipéptidos da invenção.

Podem ser utilizados vários sistemas de cultura de células de mamífero para expressar a proteína recombinante, os exemplos de sistemas de expressão de mamífero incluem as linhas COS-7 de fibroblastos de rim de macaco, descritas em "SV40-transformed simian cells support the replication of early SV40 mutantes" (Gluzman, 1981) e outras linhas celulares capazes de expressar um vector compatível, por exemplo, as linhas celulares C127, 3T3, CHO, heLa e BHK. Os vectores de expressão de mamífero irão compreender uma origem de replicação, um promotor e

intensificador adequado e também quaisquer locais necessários de ligação a ribossoma, local de poliadenilação, locais de dador e aceitador de *splice*, sequência de terminação de transcrição e sequências não transcritas que flanqueiam 5'. As sequências de ADN derivadas do *splice* de SV40 e locais de poliadenilação podem ser utilizados para proporcionar aos elementos genéticos não transcritos requeridos.

As células hospedeiras contendo os polinucleótidos de interesse podem ser cultivadas em meios nutrientes convencionais modificados como apropriado para activar promotores, seleccionar transformantes ou amplificar genes. As condições de cultura, tais como a temperatura, pH e semelhantes, são aquelas previamente utilizadas com a célula hospedeira seleccionada para a expressão e será evidente ao técnico com conhecimento geral na matéria. Os clones que são identificados como tendo a actividade enzimática especificada podem depois ser sequenciados para identificar a sequência do polinucleótido que codifica uma enzima tendo actividade aumentada.

Amplificação de ácidos Nucleicos

Na prática da invenção, os ácidos nucleicos codificando os polipéptidos da invenção ou ácidos nucleicos modificados, podem ser reproduzidos através, *e. g.*, de amplificação. A invenção proporciona pares de sequências iniciadoras de amplificação para amplificar ácidos nucleicos que codificam polipéptidos com uma actividade de fitase ou suas sub-sequências, em que os pares de iniciadores são capazes de amplificar sequências de ácidos nucleico incluindo a SEQ ID N°:1 exemplificativa, e as modificações de sequência específicas apresentadas acima. Um especialista na técnica pode conceber pares de sequências

iniciadoras de amplificação para qualquer parte ou tamanho total destas sequências; por exemplo:

A SEQ ID N°:1 "parental" é

ATGAAAGCGATCTTAATCCCATTTTTATCTCTTCTGATTCCGTTAACCCCGCAATCTG
CATTCGCTCAGAGTGAGCCGGAGCTGAAGCTGGAAAGTGTGGTGATTGTCAGTCGT
CATGGTGTGCGTGCTCCAACCAAGGCCACGCAACTGATGCAGGATGTCACCCCAGA
CGCATGGCCAACCTGGCCGGTAAAACTGGGTGAGCTGACACCGCGCGGTGGTGAGC
TAATCGCCTATCTCGGACATTACTGGCGTCAGCGTCTGGTAGCCGACGGATTGCTGC
CTAAATGTGGCTGCCCGCAGTCTGGTCAGGTCGCGATTATTGCTGATGTCGACGAGC
GTACCCGTAAAACAGGCGAAGCCTTCGCCGCCGGGCTGGCACCTGACTGTGCAATA
ACCGTACATACCCAGGCAGATACGTCCAGTCCCGATCCGTTATTTAATCCTCTAAAA
ACTGGCGTTTGCCAACTGGATAACGCGAACGTGACTGACGCGATCCTCGAGAGGGC
AGGAGGGTCAATTGCTGACTTTACCGGGCATTATCAAACGGCGTTTCGCGAACTGG
AACGGGTGCTTAATTTTCCGCAATCAAACTTGTGCCTTAAACGTGAGAAACAGGAC
GAAAGCTGTTCATTAACGCAGGCATTACCATCGGAACTCAAGGTGAGCGCCGACTG
TGTCTCATTAACCGGTGCGGTAAGCCTCGCATCAATGCTGACGGAGATATTTCTCCT
GCAACAAGCACAGGGAATGCCGGAGCCGGGGTGGGGAAGGATCACCGATTCACAC
CAGTGGAACACCTTGCTAAGTTTGCATAACGCGCAATTTGATTTGCTACAACGCACG
CCAGAGGTTGCCCGCAGCCGCGCCACCCCGTTATTAGATTTGATCAAGACAGCGTTG
ACGCCCCATCCACCGCAAAAACAGGCGTATGGTGTGACATTACCCACTTCAGTGCT
GTTTATCGCCGGACACGATACTAATCTGGCAAATCTCGGCGGCGCACTGGAGCTCA
ACTGGACGCTTCCCGGTCAGCCGGATAACACGCCGCCAGGTGGTGAACTGGTGTTT
GAACGCTGGCGTCGGCTAAGCGATAACAGCCAGTGGATTCAGGTTTCGCTGGTCTTC
CAGACTTTACAGCAGATGCGTGATAAAACGCCGCTGTCATTAAATACGCCGCCCGG
AGAGGTGAAACTGACCCTGGCAGGATGTGAAGAGCGAAATGCGCAGGGCATGTGTT
CGTTGGCAGGTTTTACGCAAATCGTGAATGAAGCACGCATACCGGCGTGCAGTTTG
AGATCTCATCTA

Assim, um par de sequências iniciadoras de amplificação para amplificar esta sequência parental ou uma das sequências exemplificativas da invenção tendo, pelo menos, uma das modificações de sequência específicas aqui apresentadas, podem ser os resíduos 1 a 21 da SEQ ID N°:1 (*i. e.*, ATGAAAGCGATCTTAATCCCA) e a cadeia complementar dos últimos 21 resíduos da SEQ ID N°:1 (*i. e.*, a cadeia complementar de TGCAGTTTGAGATCTCATCTA).

Reacções de amplificação podem também ser utilizadas para quantificar a quantidade de ácido nucleico numa amostra (tal como a quantidade de mensageiro numa amostra celular), marcar o

ácido nucleico (*e. g.*, para o aplicar numa matriz ou transferência), detectar o ácido nucleico ou quantificar a quantidade de um ácido nucleico específico numa amostra. Num aspecto da invenção, o mensageiro isolado de uma célula ou o ADNc de uma biblioteca é amplificado. O especialista pode seleccionar e conceber iniciadores oligonucleotídicos de amplificação adequados. Os métodos de amplificação são também bem conhecidos na técnica e incluem, *e. g.*, reacção em cadeia da polimerase, PCR (ver, *e. g.*, PCR PROTOCOLS, A GUIDE TO METHODS AND APPLICATIONS, ed. Innis, Academic Press, N.I. (1990) e PCR STRATEGIES (1995), ed. Innis, Academic Press, Inc., N.I., reacção em cadeia da ligase (RCL) (ver, *e. g.*, Wu (1989) Genomics 4:560; Landegren (1988) Science 241:1077; Barringer (1990) Gene 89:117); amplificação de transcrição (ver, *e. g.*, Kwoh (1989) Proc. Natl. Acad. Sci. USA 86:1173); e replicação de sequência auto-sustentada (ver, *e. g.*, Guatelli (1990) Proc. Natl. Acad. Sci. USA 87:1874); Amplificação de Q-beta replicase (ver, *e. g.*, Smith (1997) J. Clin. Microbiol. 35:1477-1491), ensaio automatizado de amplificação de Q-beta replicase (ver, *e. g.*, Burg (1996) Mol. Cell. Probes 10:257-271) e outras técnicas mediadas por ARN polimerase (*e. g.*, NASBA, Cangene, Mississauga, Ontário); ver também Berger (1987) Methods Enzymol. 152:307-316; Sambrook; Ausubel; Patente U.S. N° 4683195 e 4683202; Sooknanan (1995) Biotechnology 13:563-564.

Determinação do grau de identidade de sequência

A invenção proporciona um ácido nucleico isolado, sintético ou recombinante compreendendo uma sequência de ácido nucleico que tem, pelo menos, 95%, 96%, 97%, 98%, 99%, ou mais de identidade de sequência para a SEQ ID N°:1 e incluindo as modificações especificamente enumeradas para a SEQ ID N°:1,

discutidas acima. Num aspecto, a extensão da identidade de sequência (homologia) pode ser determinada utilizando qualquer programa de computador e parâmetros associados, incluindo aqueles aqui descritos, tais como BLAST 2.2.2. ou FASTA versão 3.0t78, com os parâmetros de defeito.

As sequências homólogas incluem também sequências de ARN em uridinas substituem timinas nas sequências de ácidos nucleicos. As sequências homólogas podem ser obtidas utilizando qualquer dos processos aqui descritos ou podem resultar da correcção de um erro de sequenciação.

Vários programas de comparação de sequências aqui identificados são utilizados neste aspecto da invenção. As identidades (homologias) de sequências de proteína e/ou ácidos nucleicos podem ser avaliadas utilizando qualquer de uma variedade de algoritmos e programas de comparação de sequências conhecidos na técnica. Tais algoritmos e programas incluem, mas não estão limitados a, TBLASTN, BLASTP, FASTA, TFASTA e CLUSTALW (Pearson e Lipman, Proc. Natl. Acad. Sci. USA 85(8):2444-2448, 1988; Altschul *et al.*, J. Mol. Biol. 215(3):403-410, 1990; Thompson *et al.*, Nucleic Acids Res. 22(2):4673-4680, 1994; Higgins *et al.*, Methods Enzymol. 266:383-402, 1996; Altschul *et al.*, J. Mol. Biol. 215(3):403-410, 1990; Altschul *et al.*, Nature Genetics 3:266-272, 1993.

A homologia ou a identidade pode ser medida utilizando software de análise de sequências (*e. g.*, Sequence Analysis Software Package do Genetics Computer Group, University of Wisconsin Biotechnology Center, 1710 University Avenue, Madison, WI 53705). Tal software emparelha sequências semelhantes por atribuição de graus de homologia a várias deleções, substituições e outras modificações. Os termos “homologia” e

“identidade” no contexto de duas ou mais sequências de ácidos nucleicos ou polipéptidos, refere-se a duas ou mais sequências ou sub-sequências que são iguais ou têm uma percentagem específica de resíduos de aminoácidos ou nucleótidos que são iguais quando comparadas e alinhadas para correspondência máxima ao longo de uma janela de comparação ou região designada, como medido utilizando qualquer número de algoritmos de comparação de sequências ou por alinhamento manual e inspecção visual. Para a comparação de sequências, uma sequência pode actuar como uma sequência de referência (uma sequência exemplificativa da invenção) com a qual são comparadas sequências de teste. Ao utilizar um algoritmo de comparação de sequências, as sequências de referência e teste dão entrada num computador, coordenadas de sub-sequência são designadas, se necessário, e os parâmetros do programa de algoritmo de sequências são designados. Os parâmetros do programa de defeito podem ser utilizados ou parâmetros alternativos podem ser designados. O algoritmo de comparação de sequências calcula depois a percentagem de identidades de sequência para as sequências de teste relativamente à sequência de referência, com base nos parâmetros do programa.

Uma “janela de comparação”, como aqui utilizada, inclui referência a um segmento de qualquer número de resíduos contíguos. Por exemplo, em aspectos alternativos da invenção, os resíduos contíguos variando desde qualquer coisa como 20 até ao tamanho total de sequências exemplificativas da invenção são comparadas com uma sequência de referência do mesmo número de posições contíguas, depois das duas sequências serem optimamente alinhadas. Se a sequência de referência tiver a identidade de sequência requerida para sequências exemplificativas da invenção, *e. g.*, 98% de identidade de sequência para a SEQ ID N°:1, SEQ ID N°:2 e tiver uma das modificações de

sequência específicas assinaladas acima, a sequência está dentro do âmbito da invenção. Em formas de realização alternativas, as sub-sequências variando desde cerca de 20 a 600, cerca de 50 a 200 e cerca de 100 a 150 são comparadas com uma sequência de referência do mesmo número de posições contíguas, depois das duas sequências serem optimamente alinhadas. Os métodos de alinhamento de sequências para comparação são bem conhecidos na técnica. O alinhamento óptimo de sequências para comparação pode ser conduzido, *e. g.*, pelo algoritmo de homologia local de Smith & waterman, Adv. Appl. Math. 2:482, 1981, pelo algoritmo de alinhamento de homologia de Needleman & Wunsch, J. Mol. Biol. 48:443, 1970, pelo método de pesquisa de semelhança de Person & Lipman, Proc. Nat'l. Acad. Sci. USA 85:2444, 1988, por implementações computorizadas destes algoritmos (GAP, BESTFIT, FASTA e TFASTA no Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr., Madison, WI) ou por alinhamento manual e inspecção visual. Outros algoritmos para determinar a homologia ou a identidade incluem, por exemplo, para além do programa BLAST (Basic Local Alignment Search Tool no National Center for Biological Information), ALIGN, AMAS (Analysis of Multiply Aligned Sequences), AMPS (Protein Multiple Sequence Alignment), ASSET (Aligned Segment Statistical Evaluation Tool), BANDS, BESTSCOR, BIOSCAN (Biological Sequence Comparative Analysis Node), BLIMPS (BLocks IMProved Searcher), FASTA, Intervals & Points, BMB, CLUSTAL V, CLUSTAL W, CONSENSUS, LCONSENSUS, WCONSENSUS, algoritmo de Smith-Waterman, DARWIN, algoritmo de Las Vegas, FNAT (Forced Nucleotide Alignment Tool), Framealign, Framesearch, DYNAMIC, FILTER, FSAP (Fristensky Sequence Analysis Package), GAP (Global Alignment Program), GENAL, GIBBS, GenQuest, ISSC (Sensitive Sequence Comparison), LALIGN (Local Sequence Alignment), LCP (Local Content Program), MACAW (Multiple Alignment Construction & Analysis Workbench), MAP (Multiple Alignment Program), MBLKP, MBLKN, PIMA

(Pattern-Induced Multi-sequence Alignment), SAGA (Sequence Alignment by Genetic Algorithm) e WHAT-IF. Tais programas de alinhamento podem também ser utilizados para rastrear bases de dados de genoma para identificar sequências de polinucleótidos tendo sequências substancialmente idênticas. Uma variedade de bases de dados de genoma está disponível, por exemplo, uma porção substancial do genoma humano está disponível como parte do Projecto de Sequenciação do Genoma Humano (Gibbs, 1995). Diversos genomas foram sequenciados, *e. g.*, *M. genitalium* (Fraser *et al.*, 1995), *M. jannaschii* (Bult *et al.*, 1996), *H. influenzae* (Fleischmann *et al.*, 1995), *E. coli* (Blattner *et al.*, 1997) e levedura (*S. cerevisiae*) (Mewes *et al.*, 1997) e *D. melanogaster* (Adams *et al.*, 2000). Foi também feito progresso significativo na sequenciação de genomas de organismos modelo, tais como murganho, *C. elegans* e *Arabadopsis sp.* Bases de dados contendo informação genómica anotada com alguma informação funcional são mantidas por diferentes organizações e estão acessíveis pela Internet.

Os algoritmos BLAST, BLAST 2.0 e BLAST 2.2.2 são também utilizados para realizar a invenção. São descritos, *e. g.*, em Altschul (1977) Nuc. Acids Res. 25:3389-3402; Altschul (1990) J. Mol. Biol. 215:403-410. O software para realizar análises de BLAST está disponível publicamente através do National Center for Biotechnology Information. Este algoritmo envolve primeiro identificar pares de sequências de elevada pontuação (HSP) através de identificação de curtas palavras de comprimento W na sequência de pesquisa, que emparelham ou satisfazem alguma pontuação T de limite avaliado positivamente, quando alinhadas com uma palavra do mesmo comprimento numa sequência da base de dados. T é referido como o limite de pontuação de palavra da vizinhança (Altschul (1990) supra). Estes acertos de palavra da vizinhança inicial actuam como sementes para iniciar pesquisas

para encontrar HSP mais longos que as contêm. Os acertos de palavra são prolongados em ambos os sentidos ao longo de cada sequência, tanto quanto possível, para aumentar a pontuação de alinhamento cumulativa. As pontuações cumulativas são calculadas utilizando, para sequências de nucleótidos, os parâmetros M (pontuação de recompensa para um par de resíduos condizentes; sempre > 0). Para sequências de aminoácidos, uma matriz de pontuação é utilizada para calcular a pontuação cumulativa. A extensão dos acertos de palavra em cada sentido é parada quando: a pontuação cumulativa de alinhamento cai pela quantidade X a partir do seu valor máximo obtido; a pontuação cumulativa vai para zero ou abaixo, devido à acumulação de uma ou mais alinhamentos de resíduo com pontuação negativa; ou o fim de qualquer das sequências é alcançado. Os parâmetros W, T e X do algoritmo BLAST determinam a sensibilidade e a velocidade de alinhamento. O programa BLASTN (para sequências de nucleótidos) utiliza como defeito um comprimento de palavra (W) de 11, uma expectativa (E) de 10, M=5, N= -4 e uma comparação de ambas as cadeias. Para sequências de aminoácidos, o programa BLASTP utiliza como defeito um comprimento de palavra de 3 e expectativas (E) de 10 e os alinhamentos da matriz de pontuação BLOSUM62 (ver Henikoff & Henikoff (1989) Proc. Natl. Acad. Sci. USA 89:10915) (B) de 50, expectativa (E) de 10, M=5, N = -4 e uma comparação de ambas as cadeias. O algoritmo BLAST também realiza uma análise estatística da semelhança entre duas sequências (ver, *e. g.*, Karlin & Altschul (1993) Proc. Natl. Acad. Sci. USA 90:5873). Uma medida de semelhança proporcionada pelo algoritmo BLAST é a menor probabilidade de soma (P(N)), que proporciona uma indicação da probabilidade através da qual um emparelhamento entre duas sequências de nucleótidos ou de aminoácidos ocorreria por sorte. Por exemplo, um ácido nucleico é considerado semelhante a uma sequência de referência se a menor probabilidade de soma numa comparação do ácido nucleico de

teste com o ácido nucleico de referência for menor do que cerca de 0,2, menor do que cerca de 0,01 ou menor do que cerca de 0,001. Num aspecto, as homologias de sequências de proteína e de ácidos nucleicos são avaliadas utilizando Basic Local Alignment Search Tool ("BLAST"). Por exemplo, podem ser utilizados cinco programas BLAST específicos para realizar a seguinte tarefa: (1) BLASTP e BLAST3 comparam uma sequência de aminoácidos de pesquisa contra uma base de dados de sequências de proteínas; (2) BLASTN compara uma sequência de nucleótidos de pesquisa contra uma base de dados de sequências de nucleótidos; (3) BLASTX compara os produtos de tradução conceptuais de seis grelhas de uma sequência de nucleótidos de pesquisa (ambas as cadeias) contra uma base de dados de sequências de proteínas; (4) TBLASTN compara uma sequência da proteína de pesquisa contra uma base de dados de sequências de nucleótidos traduzida em todas as seis grelhas de leitura (ambas as cadeias); e (5) TBLASTX compara as traduções das seis grelhas de uma sequência de nucleótidos de pesquisa contra as traduções de seis grelhas de uma base de dados de sequências de nucleótidos. Os programas BLAST identificam sequências homólogas através da identificação de segmentos semelhantes, os quais são aqui referidos como "pares de segmentos de elevada pontuação" entre uma sequência de aminoácidos ou de ácidos nucleicos de pesquisa e um sequência de teste que pode ser obtida de uma base de dados de sequências de proteína ou de ácido nucleico. Os pares de segmentos de elevada pontuação podem ser identificados (*i. e.*, alinhados) por meio de uma matriz de pontuação, muitas das quais são conhecidas na técnica. Uma matriz de pontuação exemplificativa utilizada é a matriz BLOSUM62 (Gonnet *et al.*, Science 256:1443-1445, 1992; Henikoffand Henikoff, Proteins 17:49-61, 1993). Alternativamente, as matrizes PAM ou PAM250 podem ser utilizadas (ver, *e. g.*, Schwartz e Dayhoff, eds., 1978, Matrices for Detecting Distance Relationships: Atlas of

Protein Sequence and Structure, Washington: National Biomedical Research Foundation).

Num aspecto, para determinar se um ácido nucleico tem a identidade de sequência requerida dentro do âmbito da invenção, são utilizados os programas BLAST 2.2.2 da NCBI. Opções de defeito para blastp. Existem cerca de 38 opções de ajuste no programa BLAST 2.2.2. Neste aspecto exemplificativo da invenção, todos os valores de defeito são utilizados à excepção do ajuste de filtro por defeito (*i. e.*, todos os parâmetros ajustados por defeito excepto o filtro que é ajustado para OFF); no seu lugar é utilizado um ajuste "-F F", o qual desliga o filtro. A utilização de filtro por defeito resulta frequentemente em violações de Karlin-Altschul devido ao curto comprimento da sequência.

Os valores de defeito utilizados neste aspecto exemplificativo da invenção incluem:

"Filtro para baixa complexidade: ON

> Tamanho de Palavra: 3
> Matriz: Blosum62
> Custos de Lacuna: Existência:11

> Extensão:1"

"Filtro para a complexidade baixa: ON

Outros ajustes de defeito são: filtro para baixa complexidade OFF, tamanho de palavra de 3 para proteína, matriz BLOSUM62, penalidade de existência de lacuna de -11 e uma penalidade de extensão de lacuna de -1.

Sistemas computorizados e produtos de programas de computador

Para determinar e identificar identidades de sequência, homologias estruturais, motivos e semelhantes *in silico,* a sequência da invenção pode ser armazenada, gravada e manipulada em qualquer meio que possa ser lido e acedido por um computador. Os motivos que podem ser detectados utilizando os programas acima incluem sequências que codificam fechos de leucina, motivos de hélice-volta-hélice, locais de glicosilação, locais de ubiquitinação, hélices alfa e folhas beta, sequências de sinal que codificam péptidos de sinal que direccionam a secreção das proteínas codificadas, sequências implicadas na regulação de transcrição, tais como homeoboxes, regiões acídicas, locais activos enzimaticamente, locais de ligação de substrato e locais de clivagem enzimática.

Os computadores, sistemas computorizados, meios de leitura por computador, produtos de programas de computador e semelhantes podem ter aí gravadas ou armazenadas as sequências de ácidos nucleicos e polipéptidos da invenção, *e. g.*, , as sequências exemplificativas da invenção. Como aqui utilizadas, as palavras “gravado” e “armazenado” referem-se a um processo para armazenar informação num meio computorizado. Um especialista na técnica pode prontamente adoptar quaisquer métodos conhecidos para gravar informação num meio de leitura por computador para gerar produtos fabricados compreendendo uma ou mais das sequências de ácidos nucleicos e/ou de polipéptidos da invenção. Como aqui utilizadas, o termo “computador”, e as expressões “programa de computador” e “processador” são

utilizadas nos seus contextos gerais mais amplos e incorporam todos esses dispositivos, como aqui descritos.

Um meio de leitura por computador pode ter aí gravado, pelo menos, uma sequência de ácidos nucleicos e/ou polipéptidos da invenção. Os meios de leitura por computador incluem meios de leitura magnética, meios de leitura óptica, meios de leitura electrónica e meios magnéticos/ópticos. Por exemplo, os meios de leitura por computador podem ser um disco rígido, um disco flexível, uma fita magnética, CD-ROM, Disco Versátil Digital (DVD), Memória de Acesso Aleatório (RAM) ou Memória Apenas de Leitura (ROM), assim como outros tipos de outros meios conhecidos dos especialistas na técnica.

Os sistemas (*e. g.*, sistemas baseados na Internet), particularmente sistemas computorizados, podem armazenar e manipular as sequências e a informação de sequência aqui descrita. Um Exemplo de um sistema **100** computorizado é ilustrado na forma de diagrama de blocos na Figura 1. Como aqui utilizado, um "sistema computorizado" refere-se aos componentes de hardware, aos componentes de software e aos componentes de armazenamento de dados utilizados para analisar uma sequência de nucleótidos ou de polipéptidos da invenção. O sistema **100** computorizado pode incluir um processador para processar, aceder e manipular os dados de sequência. O processador **105** pode ser qualquer tipo bem conhecido de unidade de processamento central, tal como, por exemplo, o Pentium III da Intel Corporation ou processador semelhante da Sun, Motorola, Compaq, AMD ou International Business Machines. O sistema **100** computorizado é um sistema para uso geral que compreende o processador **105** e um ou mais componentes **110** de armazenamento de dados internos para armazenar dados e um ou mais dispositivos de recuperação de

dados para recuperar os dados armazenados nos componentes de armazenamento de dados. Um especialista na técnica entende prontamente que qualquer dos sistemas computorizados actualmente disponíveis é adequado.

O sistema **100** computorizado pode incluir um processador **105** conectado a um barramento que é conectado a uma memória **115** principal (pode ser implementado como RAM) e um ou mais dispositivos **110** de armazenamento de dados internos, tais como um disco rígido e/ou outros meios de leitura por computador tendo dados aí gravados. O sistema **100** computorizado pode ainda incluir um ou mais dispositivos **118** de recuperação de dados para ler os dados armazenados nos dispositivos **110** de armazenamento de dados internos.

O dispositivo **118** de recuperação de dados pode representar, por exemplo, uma unidade de disco flexível, uma unidade de disco compacto, uma unidade de fita magnética ou um modem capaz de conexão a um sistema remoto de armazenamento de dados de (*e. g.*, através da Internet), etc. Em algumas formas de realização, o dispositivo **110** de armazenamento de dados internos é um meio de leitura por computador removível, tais como um disco flexível, um disco compacto, uma fita magnética, etc., contendo aí controlo lógico e/ou dados gravados. O sistema **100** computorizado pode, vantajosamente, incluir ou ser programado através de software apropriado para ler o controlo lógico e/ou os dados do componente de armazenamento de dados, uma vez introduzidos no dispositivo de recuperação de dados.

O sistema **100** computorizado inclui um visor **120** que é utilizado para apresentar resultados a um utilizador do computador. Deve ser também assinalado que o sistema **100**

computorizado pode ser ligado a outros sistemas **125a-c** computorizados numa rede ou numa rede de área alargada para proporcionar acesso centralizado ao sistema **100** computorizado. O software para aceder e processar as sequências de nucleótidos ou aminoácidos da invenção pode residir na memória **115** principal durante a execução.

O sistema **100** computorizado pode compreender ainda um algoritmo de comparação de sequências para comparar uma sequência de ácido nucleico da invenção. O algoritmo e as sequências podem ser armazenadas num meio de leitura por computador. Um "algoritmo de comparação de sequências" refere-se a um ou mais programas que são implementados (localmente ou remotamente) no sistema **100** computorizado para comparar uma sequência de nucleótidos com outras sequências de nucleótidos e/ou compostos armazenados dentro dos meios de armazenamento de dados. Por exemplo, o algoritmo de comparação de sequências pode comparar as sequências de nucleótidos da invenção com a sequência SEQ ID N°:1 "parental" e/ou SEQ ID N°:2, armazenada num meio de leitura por computador para referenciar as sequências armazenadas num meio de leitura por computador para identificar homologias ou motivos estruturais.

Os parâmetros utilizados com os algoritmos acima podem ser adaptados dependendo do comprimento da sequência e do grau de homologia estudados. Em alguns aspectos, os parâmetros podem ser os parâmetros de defeito utilizados pelos algoritmos na ausência de instruções do utilizador. A Figura 2 é um fluxograma que ilustra um aspecto de um processo **200** para comparar uma nova sequência de proteína ou nucleótidos com uma base de dados de sequências, de modo a determinar os níveis de homologia entre a nova sequência e as sequências na base de dados. A base de dados

de sequências pode ser uma base de dados privada armazenada no sistema **100** computorizado, ou uma base de dados pública, tal como GENBANK, o qual está disponível através da Internet. O processo **200** começa num estado **201** inicial e depois move-se para um estado **202** em que a nova sequência a ser comparativa é armazenada numa memória num sistema **100** computorizado. Como discutido acima, a memória pode ser qualquer tipo de memória, incluindo RAM ou um dispositivo de armazenamento interno.

O processo **200** depois move-se para um estado **204** em que uma base de dados de sequências é aberta para análise e comparação. O processo **200** depois move-se para um estado **206** em que a primeira sequência armazenada na base de dados é lida numa memória no computador. Uma comparação é depois realizada num estado **210** para determinar se a primeira sequência é a mesma que a segunda sequência. É importante assinalar que este passo não está limitado a realizar uma comparação exacta entre a nova sequência e a primeira sequência na base de dados. Os métodos bem conhecidos são conhecidos dos especialistas na técnica para comparar duas sequências de nucleótidos ou de proteína, mesmo se não forem idênticas. Por exemplo, lacunas podem ser introduzidas numa sequência de modo a aumentar o nível de homologia entre as duas sequências testadas. Os parâmetros que controlam se as lacunas ou outras características são introduzidas numa sequência durante a comparação são normalmente introduzidos pelo utilizador do sistema computorizado.

Assim que uma comparação de duas sequências for realizada no estado **210**, é feita uma determinação num estado **210** de decisão se as duas sequências são a mesma. Naturalmente, o termo “mesma” não está limitado às sequências que são absolutamente idênticas. As sequências que estão dentro dos parâmetros de homologia

introduzidas pelo utilizador serão marcadas como “mesma” no processo **200**. Se for feita uma determinação que as duas sequências são a mesma, o processo **200** move-se para um estado **214** em que o nome da sequência da base de dados é apresentado ao utilizador. Este estado notifica o utilizador que a sequência com o nome indicado cumpre os constrangimentos de homologia que foram introduzidos. Uma vez apresentado o nome da sequência armazenada ao utilizador, o processo **200** move-se para um estado **218** de decisão em que é feita uma determinação se existem mais sequências na base de dados. Se não existirem mais sequências na base de dados, então processo **200** termina num estado **220** final. Contudo, se existirem mais sequências na base de dados, então o processo **200** move-se para um estado **224** em que um indicador é movido para a sequência seguinte na base de dados, de modo que possa ser comparada à nova sequência. Deste modo, a nova sequência é alinhada e comparada com cada sequência na base de dados.

Deve ser assinalado que se for feita uma determinação no estado **212** de decisão que as sequências não eram homólogas, então o processo **200** move-se imediatamente para o estado **218** de decisão de modo a determinar se está disponível alguma outra sequência na base de dados para a comparação. Desta forma, um aspecto da invenção é um sistema computorizado compreendendo um processador, um dispositivo de armazenamento de dados que aí armazena uma sequência de ácidos nucleicos da invenção e um comparador de sequência para conduzir a comparação. O comparador de sequência pode indicar um nível de homologia entre as sequências comparadas ou identificar motivos estruturais ou pode identificar motivos estruturais nas sequências as quais são comparadas com estes códigos de ácidos nucleicos e códigos de polipéptidos.

A Figura 3 é um fluxograma que ilustra uma forma de realização de um processo **250** num computador para determinar se duas sequências são homólogas. O processo **250** começa num estado **252** inicial e move-se depois para um estado **254** em que uma primeira sequência a ser comparada é armazenada numa memória. A segunda sequência a ser comparada é depois armazenada numa memória num estado **256**. O processo **250** move-se depois para um estado **260** em que o primeiro caracter na primeira sequência é lido e depois para um estado **262** em que o primeiro caracter da segunda sequência é lido. Deve ser compreendido que se a sequência for uma sequência de nucleótidos, então o caracter seria normalmente A, T, C, G ou U. Se a sequência for uma sequência de proteína, então pode ser um código de aminoácido de letra única de modo que a primeira e a segunda sequência possam ser facilmente comparadas. É feita então uma determinação num estado **264** de decisão se os dois caracteres são o mesmo. Se forem o mesmo, então o processo **250** move-se para um estado **268** em que os caracteres seguintes na primeira e segunda sequência são lidos. É feito depois uma determinação se os caracteres seguintes forem o mesmo. Se forem, então o processo **250** continua este ciclo até que dois caracteres não sejam o mesmo. Se for feita uma determinação de que os dois caracteres seguintes não são o mesmo, o processo **250** move-se para um estado **274** de decisão para determinar se existem mais caracteres em ambas as sequências para ler. Se não existirem mais caracteres para ler, então o processo **250** move-se para um estado **276** em que o nível de homologia entre a primeira e a segunda sequência é indicado ao utilizador. O nível de homologia é determinado através de cálculo da proporção de caracteres entre as sequências que eram as mesmas para um número total de sequências na primeira sequência. Assim, se cada caracter numa primeira sequência de

100 nucleótidos estiver alinhada com cada caracter numa segunda sequência, o nível de homologia é 100%.

Alternativamente, o programa de computador pode comparar uma sequência de referência com uma sequência da invenção para determinar se as sequências diferem numa ou mais posições. O programa pode registar o comprimento e a identidade de nucleótidos ou resíduos de aminoácidos introduzidos, suprimidos ou substituídos relativamente à sequência de referência ou da invenção. O programa de computador pode ser um programa que determina se uma sequência de referência contém um polimorfismo de nucleótido único (SNP) relativamente a uma sequência da invenção, ou se uma sequência da invenção compreende um SNP de uma sequência conhecida. Assim, em alguns aspectos, o programa de computador é um programa que identifica SNP. O método pode ser implementado pelos sistemas computorizados descritos acima e o método ilustrado na Figura 3. O método pode ser realizado através de leitura de uma sequência da invenção e das sequências de referência através da utilização do programa de computador e identificação das diferenças com o programa de computador.

O sistema baseado em computador pode compreender um identificador para identificar características num ácido nucleico ou polipéptido da invenção. Um "identificador" refere-se a um ou mais programas que identificam determinadas características numa sequência de ácido nucleico. Por exemplo, um identificador pode compreender um programa que identifica uma grelha de leitura aberta (ORF) numa sequência de ácidos nucleicos. A Figura 4 é um fluxograma que ilustra um aspecto de um processo **300** de identificador para detectar a presença de uma característica numa sequência. O processo **300** começa num estado **302** inicial e move-se depois para um estado **304** em que uma primeira sequência que é para ser verificada para

características é armazenada numa memória **115** no sistema **100** computorizado. O processo **300** move-se depois para um estado **306** em que uma base de dados de características de sequência é aberta. Tal base de dados incluiria uma lista de atributos de cada característica, em conjunto, com o nome da característica. Por exemplo, um nome da característica poderia ser "Codão de Iniciação" e o atributo seria "ATG". Outro Exemplo seria o nome de característica "Caixa TAATAA" e o atributo da característica seria "TAATAA". Um Exemplo de tal base de dados é produzido pelo University of Wisconsin Genetics Computer Group. Alternativamente, as características podem ser motivos estruturais do polipéptido, tais como hélices alfa, folhas beta, ou motivos funcionais do polipéptido, tais como locais activos enzimaticamente, motivos hélice-volta-hélice ou outros motivos conhecidos dos especialistas na técnica. Uma vez que a base de dados de características é aberta no estado **306**, o processo **300** move-se para um estado **308** em que a primeira característica é lida da base de dados. Uma comparação do atributo da primeira característica com a primeira sequência é depois feita num estado **310**. É feito depois uma determinação num estado **316** de decisão se o atributo da característica foi encontrado na primeira sequência. Se o atributo for encontrado, então o processo **300** move-se para um estado **318** em que o nome da característica encontrada é indicado ao utilizador. O processo **300** move-se depois para a um estado **320** de decisão em que é feita uma determinação se existem na base de dados características de movimento. Se não existirem mais características, então o processo **300** termina num estado **324** final. Contudo, se mais características existirem na base de dados, então o processo **300** lê a característica seguinte da sequência num estado **326** e volta ao estado **310** em que o atributo da característica seguinte é comparado contra a primeira

sequência. Se o atributo da característica não for encontrado na primeira sequência no estado **316** de decisão, o processo **300** move-se directamente para o estado **320** de decisão de modo a determinar se existem mais características na base de dados. Um programa de computador pode identificar grelhas de leitura aberta (ORF).

Um polipéptido ou uma sequência de ácidos nucleicos da invenção pode ser armazenado e manipulado numa variedade de programas de processamento de dados numa variedade de formatos. Por exemplo, uma sequência pode ser armazenada como texto num ficheiro de processamento de texto, tais como Microsoft WORD ou WORDPERFECT ou como um ficheiro ASCII numa variedade de programas de base de dados familiares aos especialistas na técnica, tal como DB2, SYBASE ou ORACLE. Além disso, muitos programas de computador e bases de dados podem ser utilizados como algoritmos de comparação de sequências, identificadores ou fontes de sequências de nucleótidos de referência ou sequências de polipéptidos a serem comparados a uma sequência de ácidos nucleicos da invenção. Os programas e as bases de dados utilizadas incluem, mas não estão limitadas: MacPattem (EMBL), DiscoveryBase (Molecular Applications Group), GeneMine (Molecular Applications Group), Look (Molecular Applications Group), MacLook (Molecular Applications Group), BLAST e BLAST2 (NCBI), BLASTN e BLASTX (Altschul *et al.*, J. Mol. Biol. 215: 403, 1990), FASTA (Pearson and Lipman, Proc. Natl. Acad. Sci. USA, 85: 2444, 1988), FASTDB (Brutlag *et al.* Comp. App. Biosci. 6:237-245, 1990), Catalyst (Molecular Simulations Inc.), Catalyst/SHAPE (Molecular Simulations Inc.), Cerius2.DBAccess (Molecular Simulations Inc.), HypoGen (Molecular Simulations Inc.), Insight II, (Molecular Simulations Inc.), Discover (Molecular Simulations Inc.), CHARMm (Molecular Simulations Inc.), Felix (Molecular Simulations Inc.), DelPhi, (Molecular

Simulations Inc.), QuanteMM, (Molecular Simulations Inc.), Homology (Molecular Simulations Inc.), Modeler (Molecular Simulations Inc.), ISIS (Molecular Simulations Inc.), Quanta/Protein Design (Molecular Simulations Inc.), WebLab (Molecular Simulations Inc.), WebLab Diversity Explorer (Molecular Simulations Inc.), Gene Explorer (Molecular Simulations Inc.), SeqFold (Molecular Simulations Inc.), a base de dados MDL Available Chemicals Directory database, a base de dados MDL Drug Data Report, a base de dados Comprehensive Medicinal Chemistry, a base de dados Derwent's World Drug Index, a base de dados BioByteMasterFile, a base de dados Genbank e a base de dados Genseqn. Muitas outras bases de dados e programas seriam evidentes a um especialista na técnica dada a presente divulgação.

Os motivos que podem ser detectados utilizando os programas acima incluem sequências que codificam fechos de leucina, motivos de hélice-volta-hélice, locais de glicosilação, locais de ubiquitinação, hélices alfa e folhas beta, sequências de sinal que codificam péptidos de sinal que dirigem a secreção das proteínas codificadas, sequências implicadas na regulação da transcrição, tais como homeoboxes, regiões acídicas, locais activos enzimaticamente, locais de ligação de substrato e locais de clivagem enzimática.

Inibição da Expressão de uma Fitase

A invenção proporciona ainda ácidos nucleicos complementares (*e. g.*, sequências anti-sentido) às sequências de ácidos nucleicos da invenção, incluindo ácidos nucleicos compreendendo anti-sentido, iARN, ribozimas. As sequências anti-sentido são capazes de inibir o transporte, *splice* ou transcrição de genes

que codificam fitases. A inibição pode ser efectuada através do direccionamento ao ADN genómico ou ARN mensageiro. A transcrição ou função do ácido nucleico direccionado pode ser inibida, por exemplo, através de hibridação e/ou clivagem. Um conjunto particularmente útil de inibidores proporcionados pela presente invenção inclui oligonucleótidos que são capazes de se ligarem ao gene ou mensageiro da fitase de ligação, em qualquer dos casos prevenindo ou inibindo a produção ou função da enzima fitase. A associação pode ser através de hibridação específica de sequência. Uma outra classe útil de inibidores inclui oligonucleótidos que provocam inactivação ou clivagem do mensageiro da fitase. O oligonucleótido pode ter actividade enzimática que provoca essa clivagem, tal como ribozimas. O oligonucleótido pode ser modificado quimicamente ou conjugado a uma enzima ou composição capaz de clivar o ácido nucleico complementar. Pode-se rastrear um agrupamento de muitos desses oligonucleótidos diferentes para aqueles com a actividade desejada.

Oligonucleótidos Anti-sentido

Os oligonucleótidos anti-sentido podem compreender as novas modificações da sequência de fitase da invenção, em que estes oligonucleótidos anti-sentido são capazes de se ligarem ao mensageiro da fitase o que pode inibir a actividade da fitase por direccionamento ao ARNm. As estratégias para conceber oligonucleótidos anti-sentido estão bem descritas na literatura científica e de patente, e o especialista na técnica pode conceber tais oligonucleótidos de fitase utilizando os novos reagentes da invenção. Por exemplo, os protocolos de progressão no gene/mapeamento de ARN para o rastreio de oligonucleótidos anti-sentido eficazes são bem conhecidos na técnica, ver, *e. g.*,

Ho (2000) Methods Enzymol. 314:168-183, o qual descreve um ensaio de mapeamento de ARN que é baseado em técnicas moleculares convencionais para proporcionar um método fácil e seguro para a selecção de potentes sequências anti-sentido. Ver também Smith (2000) Eur. J. Pharm. Sci. 11:191-198.

Os ácidos nucleicos de ocorrência natural são utilizados como oligonucleótidos anti-sentido. Os oligonucleótidos anti-sentido podem ser de qualquer comprimento; por exemplo, em aspectos alternativos, os oligonucleótidos anti-sentido estão entre cerca de 5 a 100, cerca de 10 a 80, cerca de 15 a 60, cerca de 18 a 40. O comprimento óptimo pode ser determinado por rastreio de rotina. Os oligonucleótidos anti-sentido podem estar presentes em qualquer concentração. A concentração óptima pode ser determinada por rastreio de rotina. Uma vasta variedade de nucleótidos sintéticos, de ocorrência não natural e análogos de ácidos nucleicos são conhecidos o que pode abordar este potencial problema. Por exemplo, podem ser utilizados ácidos nucleicos peptídicos (PNA) contendo estruturas não iónicas, tal como unidades de N-(2-aminoetil)glicina. Podem também ser utilizados oligonucleótidos anti-sentido que têm ligações fosforotioato, como descrito nos documentos WO 97/03211; WO 96/39154; Mata (1997) Toxicol Appl Pharmacol 144:189-197; Antisense Therapeutics, ed. Agarwal (Humana Press, Totowa, N.J., 1996). Os oligonucleótidos anti-sentido tendo análogos sintéticos de estrutura de ADN proporcionados pela da invenção podem também incluir ácidos nucleicos de fosforo-ditioato, metilfosfonato, fosforamidate, fosfotriéster alquílico, sulfamato, 3'-tioacetal, metileno(metilimino), 3'-N-carbamato e morfolinocarbamato, como descrito acima.

Pode ser utilizada metodologia de química combinatória para criar vastos números de oligonucleótidos que podem ser

rapidamente rastreados para oligonucleótidos específicos tendo afinidades de ligação e especificidades apropriadas para qualquer alvo, tal como as sequências de fitase de sentido e anti-sentido da invenção (ver, *e. g.*, Gold (1995) J. of Biol. Chem. 270:13581-13584).

Ribozimas Inibidoras

As ribozimas podem compreender as novas modificações de sequência de fitase da invenção, em que as ribozimas são capazes de se ligarem ao mensageiro da fitase que pode inibir a actividade da enzima fitase por direccionamento ao ARNm. As estratégias para conceber ribozimas e seleccionar a sequência anti-sentido específica de fitase para direccionamento estão bem descritas na literatura científica e de patente, e o especialista na técnica pode conceber tais ribozimas utilizando os novos reagentes da invenção. As Ribozimas actuam por ligação a um ARN alvo através da porção de ligação a ARN alvo de uma ribozima que é mantida em íntima proximidade com uma porção enzimática do ARN que cliva o ARN alvo. Assim, a ribozima reconhece e liga-se a um ARN alvo através de emparelhamento de bases complementares e uma vez ligado ao local correcto, actua enzimaticamente para clivar e inactivar o ARN alvo. A clivagem de um ARN alvo em tal modo destruirá a sua capacidade para dirigir a síntese de uma proteína codificada se a clivagem ocorrer na sequência codificante. Depois de uma ribozima ter ligado e clivado o seu alvo de ARN, é tipicamente libertada desse ARN e pode assim ligar e clivar repetidamente novos alvos.

Em algumas circunstâncias, a natureza enzimática de uma ribozima pode ser vantajosa relativamente a outras tecnologias, tal como tecnologia anti-sentido (onde uma molécula de ácido

nucleico simplesmente se liga a um alvo de ácido nucleico para bloquear a sua transcrição, tradução ou associação com outra molécula) uma vez que a concentração eficaz da ribozima necessária para efectuar um tratamento terapêutico pode ser mais baixa do que aquela de um oligonucleótido anti-sentido. Esta potencial vantagem reflecte a capacidade da ribozima para actuar enzimaticamente. Assim, uma única molécula de ribozima é capaz de clivar muitas moléculas de ARN alvo. Além disso, uma ribozima é tipicamente um inibidor altamente específico, com a especificidade de inibição a depender não apenas do mecanismo de emparelhamento de bases de ligação, mas também no mecanismo através da qual a molécula inibe a expressão de ARN a que ele se liga. Isto é, a inibição é provocada pela clivagem do alvo de ARN e assim a especificidade é definida como a razão da taxa de clivagem do ARN direccionado relativamente à taxa de clivagem do ARN não direccionado. Este mecanismo de clivagem é dependente de factores adicionais àqueles envolvidos no emparelhamento de bases. Assim, a especificidade de acção de uma ribozima pode ser maior do que aquela de oligonucleótido anti-sentido a ligar-se ao mesmo local de ARN. A molécula de ARN enzimática ribozima pode ser formada num motivo cabeça-de-martelo, mas pode também ser formado no motivo de um gancho-de-cabelo, vírus delta da hepatite, intrão do grupo I ou um ARN de tipo RNaseP (em associação com uma sequência guia de ARN). Exemplos de tais motivos de cabeça-de-martelo são descritos por Rossi (1992) Aids Research and Human Retroviruses 8:183; motivos de gancho-de-cabelo por Hampel (1989) Biochemistry 28:4929, e Hampel (1990) Nuc. Acids Res. 18:299; o motivo do vírus delta da hepatite por Perrotta (1992) Biochemistry 31:16; o motivo RNaseP por Guerrier-Takada (1983) Cell 35:849 e o intrão do grupo I por Cech na Pat. U.S. N° 4987071. A recitação destes motivos específicos não pretende ser limitativa; os especialistas na técnica reconhecerão que uma molécula de ARN enzimático desta

invenção tem um local de ligação ao substrato específico, complementar a uma ou mais das regiões de ARN do gene alvo e tem a sequência de nucleótidos dentro ou à volta desse local de ligação ao substrato que proporciona uma actividade de clivagem de ARN à molécula.

ARN de Interferência (ARNi)

Uma molécula inibidora de ARN, uma denominada molécula de "ARNi" pode compreender uma sequência de enzima da invenção. A molécula de ARNi compreende uma molécula de ARN de cadeia dupla (dsARN). A molécula de ARNi, *e. g.*, siARN e/ou miARN, pode inibir a expressão do gene da enzima fitase. Num aspecto, a molécula de ARNi, *e. g.*, siARN e/ou miARN, é cerca de 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25 ou mais nucleótidos de cadeia dupla de comprimento.

O ARNi pode entrar uma célula e provocar a degradação de um ARN de cadeia simples (ssARN) de sequências semelhantes ou idênticas, incluindo mARN endógenos. Quando uma célula é exposta ao ARN de cadeia dupla (dsARN), o ARNm do gene homólogo é seletivamente degradado através de um processo denominado ARN de interferência (ARNi). Um mecanismo básico possível na base do ARNi é a quebra de um ARN de cadeia dupla (dsARN) que emparelha com uma sequência génica específica, em curtas partes denominadas pequenos ARN de interferência, as quais desencadeiam a degradação do ARNm que emparelha com a sua sequência. Num aspecto, os ARNi da invenção são utilizados em terapêutica de silenciamento génico, ver, *e. g.*, Shuey (2002) Drug Discov. Today 7:1040-1046. Num aspecto, a invenção proporciona métodos para degradar seletivamente o ARN utilizando as moléculas de ARNi, *e. g.*, siARN e/ou miARN. Num aspecto, o ARN micro-inibidor

(miARN) inibe a tradução e o siARN inibe a transcrição. O processo pode ser praticado *in vitro, ex vivo* ou *in vivo.* Num aspecto, as moléculas de ARNi podem ser utilizadas para gerar uma mutação de perda-de-função numa célula, um órgão ou um animal. Os métodos para fazer e utilizar moléculas de ARNi, *e. g.*, siARN e/ou miARN, para degradar seletivamente o ARN são bem conhecidos na técnica, ver, *e. g.*, Patentes U.S. N° 6506559; 6511824; 6515109; 6489127.

Modificação de Ácidos Nucleicos

Os métodos para produzir variantes dos ácidos nucleicos da invenção, *e. g.*, aqueles que codificam uma enzima fitase, podem ser repetidos ou utilizados em várias combinações para produzir as enzimas fitase que têm uma actividade alterada ou diferente ou uma estabilidade alterada ou diferente que uma fitase codificada pelo ácido nucleico molde. Estes métodos podem também ser repetidos ou utilizados em várias combinações, *e. g.*, para produzir variações na expressão do gene/mensageiro, tradução do mensageiro ou estabilidade do mensageiro. Num outro aspecto, a composição genética de uma célula é alterada, *e. g.*, por modificação de um gene homólogo *ex vivo,* seguido pela sua reintrodução na célula.

Os métodos para alterar as características de uma fitase da invenção incluem mutagénese e outros métodos, incluindo evolução dirigida, *e. g.*, as abordagens de propriedade industrial da Diversa Corporation (San Diego, CA); *e. g.*, DirectEvolution; (ver, *e. g.*, Patente U.S. N° 5830696; Mutagénese de saturação de Local Génico (GSSM) (ver, *e. g.*, Patentes U.S. N° 6171820 e 6579258), Montagen Génica Mediada por Exonuclease em Evolução Dirigida (ver, *e. g.*, Patentes U.S. N° 6361974 e 6352842),

Selecção de Extremidade em Evolução Dirigida (ver, *e. g.*, Patentes U.S. Nº 6358709 e 6238884), Arranjo Combinatório de Síntese Baseada em Recombinação (ver, *e. g.*, Patentes U.S. Nº 5965408 e 6440668, e Patente Australiana Nº AU724521) e Evolução Dirigida de Enzimas Termofílicas (ver, *e. g.*, Patentes U.S. Nº 5830696 e 6335179).

Num aspecto, as características de uma fitase são modificadas através de um protocolo de DirectEvolution compreendendo: a) submeter um ou mais moldes moleculares, *e. g.*, os ácidos nucleicos de fitase da invenção, a mutagénese para produzir novas moléculas e b) seleccionar entre estas espécies descendentes de novas moléculas com características mais desejáveis. O poder da evolução dirigida depende da escolha inicial de moldes iniciais (*e. g.*, SEQ ID Nº:1 "parental" ou qualquer sequência desta invenção), assim como no(s) processo(s) de mutagénese escolhidos e no(s) processo(s) de rastreio utilizado(s). Assim, invenção proporciona novas fontes altamente activas, fisiologicamente eficazes e económicas de actividade de fitase, incluindo novas fitases que: a) têm actividades superiores sob uma ou mais aplicações específicas, tal como o fabrico a alta temperatura de géneros alimentícios e são assim úteis para optimizar estas aplicações específicas; b) são úteis como moldes para evolução dirigida para conseguir novas moléculas ainda mais melhoradas; e c) são úteis como ferramentas para identificação de moléculas adicionais relacionadas por meios como abordagens baseadas em hibridação.

Um ácido nucleico da invenção pode ser alterado através de quaisquer meios. Por exemplo, métodos aleatórios ou estocásticos, ou métodos não estocásticos ou "evolução dirigida". Os métodos para a mutação aleatória de genes são bem conhecidos na técnica, ver, *e. g.*, Patente U.S. Nº 5830696. Por

exemplo, mutagénios podem ser utilizados aleatoriamente para mutar um gene. Os mutagénios incluem, *e. g.*, luz ultravioleta ou radiação gama, ou um mutagénio químico, *e. g.*, mitomicina, ácido nitroso, psoralenos fotoactivados, isoladamente ou em combinação, para induzir rupturas de ADN susceptíveis a reparação por recombinação. Outros mutagénios químicos incluem, por exemplo, bissulfito de sódio, ácido nitroso, hidroxilamina, hidrazina ou ácido fórmico. Outros mutagénios são análogos de precursores nucleotídicos, *e. g.*, nitrosoguanidina, 5-bromouracilo, 2-aminopurina ou acridinado. Estes agentes podem ser adicionados a uma reacção de PCR em vez do precursor nucleotídico mutando assim a sequência. Agentes intercalantes, tais como proflavina, acriflavina, quinacrina e semelhantes podem também ser utilizados.

Qualquer técnica de biologia molecular pode ser utilizada, *e. g.*, mutagénese aleatória de PCR, ver, *e. g.*, Rice (1992) Proc. Natl. Acad. Sci. USA 89:5467-5471; ou mutagénese combinatória de múltipla cassete, ver, *e. g.*, Crameri (1995) Biotechniques 18:194-196. Alternativamente, os ácidos nucleicos, *e. g.*, genes, podem ser remontados após fragmentação aleatória ou "estocástica", ver, *e. g.*, Patentes U.S. Nº 6291242; 6287862; 6287861; 5955358; 5830721; 5824514; 5811238; 5605793. Em aspectos alternativos, as modificações, adições ou deleções são introduzidas por PCR propenso ao erro, arranjo combinatório, mutagénese dirigida por oligonucleótido, PCR de montagem, mutagénese de PCR sexual, mutagénese *in vivo*, mutagénese de cassete, mutagénese de montagem recursiva, mutagénese de montagem exponencial, mutagénese específica de local, remontagem génica, mutagénese de saturação de local génico (GSSM), remontagem de ligação sintética (SLR), recombinação, recombinação de sequência recursiva, mutagénese de ADN modificada por fosfotioato, mutagénese de molde contendo

uracilo, mutagénese de cadeia dupla aberta, mutagénese de reparação de desemparelhamento pontual, mutagénese de estirpe hospedeira com deficiência de reparação, mutagénese química, mutagénese radiogénica, mutagénese de deleção, mutagénese de selecção de restrição, mutagénese de purificação de restrição, síntese de gene artificial, mutagénese de montagem, criação de multímero de ácido nucleico quimérico e/ou uma combinação destes e de outros métodos.

As seguintes publicações descrevem uma variedade de processos de recombinação recursiva e/ou métodos que podem ser incorporados nos métodos da invenção: Stemmer (1999) "Molecular breeding of viruses for targeting and other clinical properties" Tumor Targeting 4:1-4; Ness (1999) Nature Biotechnology 17:893-896; Chang (1999) "Evolution of a cytokine using DNA family shuffling" Nature Biotechnology 17:793-797; Minshull (1999) "Protein evolution by molecular breeding" Current Opinion in Chemical Biology 3:284-290; Christians (1999) "Directed evolution of thymidine kinase for AZT phosphorylation using DNA family shuffling" Nature Biotechnology 17:259-264; Crameri (1998) "DNA shuffling of a family of genes from diverse species accelerates directed evolution" Nature 391:288-291; Crameri (1997) "Molecular evolution of an arsenate detoxification pathway by DNA shuffling," Nature Biotechnology 15:436-438; Zhang (1997) "Directed evolution of an effective fucosidase from a galactosidase by DNA shuffling and screening" Proc. Natl. Acad. Sci. USA 94:4504-4509; Patten *et al.* (1997) "Applications of DNA Shuffling to Pharmaceuticals and Vaccines" Current Opinion in Biotechnology 8:724-733; Crameri *et al.* (1996) "Construction and evolution of antibody-phage libraries by DNA shuffling" Nature Medicine 2:100-103; Crameri *et al.* (1996) "Improved green fluorescent protein by molecular evolution using DNA shuffling" Nature Biotechnology 14:315-

319; Gates *et al.* (1996) "Affinity selective isolation of ligands from peptide libraries through display on a lac repressor "headpiece dimer" Journal of Molecular Biology 255:373-386; Stemmer (1996) "Sexual PCR and Assembly PCR" Em: The Encyclopedia of Molecular Biology. VCH Publishers, Nova Iorque. pp.447-457; Crameri e Stemmer (1995) "Combinatorial multiple cassette mutagenesis creates all the permutations of mutant and wildtype cassettes" BioTechniques 18:194-195; Stemmer *et al.* (1995) "Single-step assembly of a gene and entire plasmid form large numbers of oligodeoxyribonucleotides" Gene, 164:49-53; Stemmer (1995) "The Evolution of Molecular Computation" Science 270: 1510; Stemmer (1995) "Searching Sequence Space" Bio/Technology 13:549-553; Stemmer (1994) "Rapid evolution of a protein in vitro by DNA shuffling" Nature 370:389-391; e Stemmer (1994) "DNA shuffling by random fragmentation and reassembly: In vitro recombination for molecular evolution." Proc. Natl. Acad. Sci. USA 91:10747-10751.

Os métodos mutacionais para gerar diversidade incluem, por exemplo, mutagénese dirigida (Ling *et al.* (1997) "Approaches to DNA mutagenesis: an overview" Anal Biochem. 254(2): 157-178; Dale *et al.* (1996) "Oligonucleotide-directed random mutagenesis using the phosphorothioate method" Methods Mol. Biol. 57:369-374; Smith (1985) "In vitro mutagenesis" Ann. Rev. Genet. 19:423-462; Botstein & Shortle (1985) "Strategies and applications of in vitro mutagenesis" Science 229:1193-1201; Carter (1986) "Site-directed mutagenesis" Biochem. J. 237:1-7; e Kunkel (1987) "The efficiency of oligonucleotide directed mutagenesis" em Nucleic Acids & Molecular Biology (Eckstein, F. e Lilley, D. M. J. eds., Springer Verlag, Berlim)); mutagénese utilizando moldes contendo uracilo (Kunkel (1985) "Rapid and efficient site-specific mutagenesis without phenotypic selection" Proc. Natl. Acad. Sci. USA 82:488-

492; Kunkel *et al.* (1987) "Rapid and efficient site-specific mutagenesis without phenotypic selection" Methods in Enzymol. 154, 367-382; e Bass *et al.* (1988) "Mutant Trp repressors with new DNA-binding specificities" Science 242:240-245); mutagénese dirigida por oligonucleótidos (Methods in Enzymol. 100: 468-500 (1983); Methods in Enzymol. 154: 329-350 (1987); Zoller & Smith (1982) "Oligonucleotide-directed mutagenesis using M13-derived vectors: an efficient and general procedure for the production of point mutations in any DNA fragment" Nucleic Acids Res. 10:6487-6500; Zoller & Smith (1983) "Oligonucleotide-directed mutagenesis of DNA fragments cloned into M13 vectors" Methods in Enzymol. 100:468-500; e Zoller & Smith (1987) "Oligonucleotide-directed mutagenesis: a simple method using two oligonucleotide primers and a single-stranded DNA template" Methods in Enzymol. 154:329-350); mutagénese de ADN modificado com fosforotioato (Taylor *et al.* (1985) "The use of phosphorothioate-modified DNA in restriction enzyme reactions to prepare nicked DNA" Nucl. Acids Res. 13: 8749-8764; Taylor *et al.* (1985) "The rapid generation of oligonucleotide-directed mutations at high frequency using phosphorothioate-modified DNA" Nucl. Acids Res. 13: 8765-8787 (1985); Nakamaye (1986) "Inhibition of restriction endonuclease Nci I cleavage by phosphorothioate groups and its application to oligonucleotide-directed mutagenesis" Nucl. Acids Res. 14: 9679-9698; Sayers *et al.* (1988) "Y-T Exonucleases in phosphorothioate-based oligonucleotide-directed mutagenesis" Nucl. Acids Res. 16:791-802; e Sayers *et al.* (1988) "Strand specific cleavage of phosphorothioate-containing DNA by reaction with restriction endonucleases in the presence of ethidium bromide" Nucl. Acids Res. 16: 803-814); mutagénese utilizando ADN de dupla cadeia aberta (Kramer *et al.* (1984) "The gapped duplex DNA approach to oligonucleotide-directed mutation construction" Nucl. Acids Res. 12: 9441-9456; Kramer & Fritz (1987) Methods in Enzymol. "Oligonucleotide-directed

construction of mutations via gapped duplex DNA" 154:350-367; Kramer *et al.* (1988) "Improved enzymatic in vitro reactions in the gapped duplex DNA approach to oligonucleotide-directed construction of mutations" Nucl. Acids Res. 16: 7207; e Fritz *et al.* (1988) "Oligonucleotide-directed construction of mutations: a gapped duplex DNA procedure without enzymatic reactions in vitro" Nucl. Acids Res. 16: 6987-6999).

Protocolos adicionais incluem reparação de desemparelhamento pontual (Kramer (1984) "Point Mismatch Repair" Cell 38:879-887), mutagénese utilizando estirpes hospedeiras deficientes na reparação (Carter *et al.* (1985) "Improved oligonucleotide site-directed mutagenesis using M13 vectors" Nucl. Acids Res. 13: 4431-4443; e Carter (1987) "Improved oligonucleotide-directed mutagenesis using M13 vectors" Methods in Enzymol. 154: 382-403), mutagénese de deleção (Eghtedarzadeh (1986) "Use of oligonucleotides to generate large deletions" Nucl. Acids Res. 14: 5115), selecção por restrição e selecção por restrição e purificação por restrição (Wells *et al.* (1986) "Importance of hydrogen-bond formation in stabilizing the transition state of subtilisin" Phil. Trans. R. Soc. Lond. A 317: 415-423), mutagénese por síntese génica total (Nambiar *et al.* (1984) "Total synthesis and cloning of a gene coding for the ribonuclease S protein" Science 223: 1299-1301; Sakamar e Khorana (1988) "Total synthesis and expression of a gene for the a-subunit of bovine rod outer segment guanine nucleotide-binding protein (transducin)" Nucl. Acids Res. 14: 6361-6372; Wells *et-al.* (1985) "Cassette mutagenesis: an efficient method for generation of multiple mutations at defined sites" Gene 34:315-323; e Grundstrom *et al.* (1985) "Oligonucleotide-directed mutagenesis by microscale 'shot-gun' gene synthesis" Nucl. Acids Res. 13: 3305-3316), reparação de quebra de dupla cadeia (Mandecki (1986); Arnold (1993) "Protein engineering for unusual

environments" Current Opinion in Biotechnology 4:450-455. "Oligonucleotide-directed double-strand break repair in plasmids of Escherichia coli: a method for site-specific mutagenesis" Proc. Natl. Acad. Sci. USA 83:7177-7181).

Detalhes adicionais ou protocolos alternativos em muitos dos métodos acima podem ser encontrados em Methods of Enzymology volume 154, o qual também descreve controlos úteis para resolução de problemas com vários métodos de mutagénese. Ver também Patente U.S. N° 5605793 de Stemmer *et al.* (22 Sep., 1998) "Methods for Generating Polynucleotides having Desired Characteristics by Iterative Selection and Recombination"; Pat. U.S. N° 5830721 de Stemmer *et al.* (Nov. 3, 1998), "DNA Mutagenesis by Random Fragmentation and Reassembly"; Pat. U.S. N° 5834252 de Stemmer, *et al.* (Nov. 10, 1998) "End-Complementary Polymerase Reaction"; Pat. U.S. N° 5837458 de Minshull, *et al.* (Nov. 17, 1998), "Methods and Compositions for Cellular and Metabolic Engineering"; documento WO 95/22625, Stemmer e Crameri, "Mutagenesis by Random Fragmentation and Reassembly"; documento WO 96/33207 por Stemmer e Lipschutz "End Complementary Polymerase Chain Reaction"; documento WO 97/20078 por Stemmer e Crameri "Methods for Generating Polynucleotides having Desired Characteristics by Iterative Selection and Recombination"; documento WO 97/35966 por Minshull e Stemmer, "Methods and Compositions for Cellular and Metabolic Engineering"; documento WO 99/41402 por Punnonen *et al.* "Targeting of Genetic Vaccine Vectors"; documento WO 99/41383 por Punnonen *et al.* "Antigen Library ImLmunization"; documento WO 99/41369 por Punnonen *et al.* "Genetic Vaccine Vector Engineering"; documento WO 99/41368 por Punnonen *et al.* "Optimization of Immunomodulatory Properties of Genetic Vaccines"; documento EP 752008 por Stemmer and Crameri, "DNA Mutagenesis by Random Fragmentation and Reassembly"; documento EP 0932670 por Stemmer

"Evolving Cellular DNA Uptake by Recursive Sequence Recombination"; documento WO 99/23107 por Stemmer *et al.*, "Modification of Virus Tropism and Host Range by Viral Genome Shuffling"; documento WO 99/21979 por Apt *et al.*, "Human Papillomavirus Vectors"; documento WO 98/31837 por del Cardayre *et al.* "Evolution of Whole Cells and Organisms by Recursive Sequence Recombination"; documento WO 98/27230 por Patten and Stemmer, "Methods and Compositions for Polypeptide Engineering"; documento WO 98/27230 por Stemmer *et al.*, "Methods for Optimization of Gene Therapy by Recursive Sequence Shuffling and Selection", documento WO 00/00632, "Methods for Generating Highly Diverse Libraries", documento WO 00/09679, "Methods for Obtaining in Vitro Recombined Polynucleotide Sequence Banks and Resulting Sequences", documento WO 98/42832 por Arnold *et al.*, "Recombination of Polynucleotide Sequences Using Random or Defined Primers", documento WO 99/29902 por Arnold *et al.*, "Method for Creating Polynucleotide and Polypeptide Sequences", documento WO 98/41653 por Vind, "An in Vitro Method for Construction of a DNA Library", documento WO 98/41622 por Borchert *et al.*, "Method for Constructing a Library Using DNA Shuffling" e documento WO 98/42727 por Pati and Zarling, "Sequence Alterations using Homologous Recombination."

Pedidos U.S. proporcionam detalhes adicionais ou protocolos alternativos relativamente a vários métodos de criação de diversidade, incluindo "SHUFFLING OF CODON ALTERED GENES" por Patten *et al.* apresentado em 28 Set., 1999, (Patente U.S. N° Ser. 09/407800); "EVOLUTION OF WHOLE CELLS AND ORGANISMS BY RECURSIVE SEQUENCE RECOMBINATION" por del Cardayre *et al.*, apresentado em 15 Jul., 1998 (Patente U.S N° Ser. 09/166188) e 15 Jul., 1999 (Patente U.S. N° Ser. 09/354,922); "OLIGONUCLEOTIDE MEDIATED NUCLEIC ACID RECOMBINATION" por Crameri *et al.*, apresentado em 28 Set., 1999 (Patente U.S. N°

Ser. 09/408392) e "OLIGONUCLEOTIDE MEDIATED NUCLEIC ACID RECOMBINATION" por Crameri *et al.*, apresentado em 18 Jan., 2000 (documento PCT/LTS00/01203); "USE OF CODON-VARIED OLIGONUCLEOTIDE SYNTHESIS FOR SYNTHETIC SHUFFLING" por Welch *et al.*, apresentado em 28 Sep., 1999 (Patente U.S. N° Ser. 09/408393); "METHODS FOR MAKING CHARACTER STRINGS, POLYNUCLEOTIDES & POLYPEPTIDES HAVING DESIRED CHARACTERISTICS" por Selifonov *et al.*, apresentado em 18 Jan., 2000, (documento PCT/US00/01202) e, *e. g.*, "METHODS FOR MAKING CHARACTER STRINGS, POLYNUCLEOTIDES & POLYPEPTIDES HAVING DESIRED CHARACTERISTICS" por Selifonov *et al.*, apresentado em 18 Jul., 2000 (Patente U.S. N° Ser. 09/618579); "METHODS OF POPULATING DATA STRUCTURES FOR USE IN EVOLUTIONARY SIMULATIONS" por Selifonov e Stemmer, apresentado em 18 Jan., 2000 (documento PCT/US00/01138); e "SINGLE-STRANDED NUCLEIC ACID TEMPLATE-MEDIATED RECOMBINATION AND NUCLEIC ACID FRAGMENT ISOLATION" por Affholter, apresentado em 6 Set., 2000 (Patente U.S. N° Ser. 09/656,549).

Os métodos não estocásticos ou de "evolução dirigida" incluem, *e. g.*, , mutagénese de saturação de local génico (GSSM), remontagem de ligação sintética (SLR) ou uma sua combinação são utilizados para modificar os ácidos nucleicos da invenção para produzir fitases com propriedades novas ou alteradas (*e. g.*, actividade sob condições altamente acídicas ou alcalinas, altas ou baixas temperaturas, e semelhantes). Os polipéptidos codificados pelos ácidos nucleicos modificados podem ser rastreados para uma actividade antes de testar para uma fitase ou outra actividade. Qualquer protocolo ou modalidade de teste pode ser utilizado, *e. g.*, utilizando uma plataforma de matriz capilar. Ver, *e. g.*, Patentes U.S. N° 6361974; 6280926; 5939250.

Mutagénese de saturação ou GSSM

Os métodos para fazer enzimas utilizando mutagénese de Saturação de Local Génico, ou GSSM, são aqui descritos e também nas Patentes U.S. N° 6171820 e 6579258.

Num aspecto, iniciadores de codão contendo uma sequência degenerada N,N,G/T são utilizados para introduzir mutações pontuais num polinucleótido, *e. g.*, uma enzima fitase ou um anticorpo da invenção, para produzir um conjunto de polipéptidos descendentes em que uma gama completa de substituições de aminoácido único é representada em cada posição de aminoácido, *e. g.*, um resíduo de aminoácido num sítio activo da enzima ou no local de ligação do ligando direccionado para ser modificado. Estes oligonucleótidos podem compreender uma primeira sequência homóloga contígua, um sequência degenerada N,N,G/T e, opcionalmente, uma segunda sequência homóloga. Os produtos de tradução de descendência a jusante da utilização de tais oligonucleótidos incluem todas as alterações de aminoácidos possíveis em cada local de aminoácido ao longo do polipéptido, porque a degenerescência da sequência N,N,G/T inclui codões para todos os 20 aminoácidos. Num aspecto, um tal oligonucleótido degenerado (compreendendo, *e. g.*, uma cassete degenerada N,N,G/T) é utilizado para submeter cada codão original num molde polinucleotídico parental a uma gama completa de substituições de codão. Num outro aspecto, pelo menos, duas cassetes degeneradas são utilizadas - no mesmo oligonucleótido ou não, para submeter, pelo menos, dois codões originais num molde polinucleotídico parental a uma gama completa de substituições de codão. Por exemplo, mais de uma sequência N,N,G/T pode estar contida num oligonucleótido para introduzir mutações de aminoácidos em mais de um local. Esta pluralidade de sequências N,N,G/T pode ser directamente contígua, ou separada por uma ou

mais sequências nucleotídicas adicionais. Num outro aspecto, oligonucleótidos úteis para introduzir adições e deleções podem ser utilizados isoladamente ou em combinação com codões contendo uma sequência N,N,G/T, para introduzir qualquer combinação ou permutação de adições, deleções e/ou substituições de aminoácidos.

Num aspecto, mutagénese simultânea em duas ou mais posições de aminoácidos contíguas é feita utilizando um oligonucleótido que contenha tripletos contíguos N,N,G/T, *i. e.*, uma sequência degenerada de (N,N,G/T)n. Num outro aspecto, são utilizadas cassetes degeneradas tendo menos degenerescência do que a sequência N,N,G/T. Por exemplo, pode ser desejável em alguns casos utilizar (*e. g.*, num oligonucleótido) uma sequência degenerada de tripleto compreendendo apenas um N, em que o referido N pode estar na primeira, segunda ou terceira posição do tripleto. Quaisquer outras bases incluindo quaisquer suas combinações e permutações podem ser utilizadas nas duas posições remanescentes do tripleto. Alternativamente, pode ser desejável em alguns casos utilizar (*e. g.*, num oligo) uma sequência degenerada de tripleto N,N,N.

Num aspecto, a utilização de tripletos degeneradas (*e. g.*, tripletos N,N,G/T) permite a produção sistemática e fácil de uma gama completa de aminoácidos naturais possíveis (para um total de 20 aminoácidos) em todas as posições de aminoácido num polipéptido (em aspectos alternativos, os métodos incluem também a produção de menos do que todas as substituições possíveis por posição de resíduo de aminoácido ou codão). Por exemplo, para um polipéptido de 100 aminoácidos, podem ser produzidas 2000 espécies distintas (*i. e.*, 20 aminoácidos possíveis por posição X 100 posições de aminoácidos). Com a utilização de um oligonucleótido ou de um conjunto de oligonucleótidos contendo

um tripleto degenerado N,N,G/T, 32 sequências individuais podem codificar para todos os 20 aminoácidos naturais possíveis. Assim, num vaso reaccional em que uma sequência polinucleotídica parental é submetida a mutagénese de saturação utilizando, pelo menos, um tal oligonucleótido, são produzidos 32 polinucleótidos descendentes distintos que codificam 20 polipéptidos distintos. Pelo contrário, a utilização de um oligonucleótido não degenerado em mutagénese dirigida conduz apenas a um produto polipeptídico descendente por vaso reaccional. Os oligonucleótidos não degenerados podem opcionalmente ser utilizados em combinação com iniciadores degenerados divulgados; por exemplo, os oligonucleótidos não degenerados podem ser utilizados para produzir mutações pontuais específicas que num polinucleótido de trabalho. Isto proporciona meios para produzir mutações pontuais silenciosas específicas, mutação pontuais que conduzem a alterações de aminoácidos correspondentes e mutações pontuais que provocam a produção de codões de terminação e a correspondente expressão de fragmentos polipeptídicos.

Num aspecto, cada vaso reaccional de mutagénese de saturação contém polinucleótidos que codificam, pelo menos, 20 moléculas de polipéptido descendente (*e. g.*, enzima fitase) de modo que todos os 20 aminoácidos naturais estão representados na posição de aminoácido específica que corresponde à posição de codão mutada no polinucleótido parental (outros aspectos utilizam menos do que todas as 20 combinações naturais). Os 32 x polipéptidos degenerados descendentes gerados a partir de cada vaso reaccional de mutagénese de saturação podem ser submetidos a amplificação clonal (*e. g.*, clonados num hospedeiro adequado, *e. g.*, hospedeiro *E. coli*, utilizando, *e. g.*, um vector de expressão) e submetidos a rastreio de expressão. Quando um polipéptido individual descendente for identificado por rastreio por apresentar uma alteração favorável na propriedade (quando

comparada ao polipéptido parental, tal como a actividade aumentada da hidrólise de glucano sob condições alcalinas ou acídicas), este pode ser sequenciado para identificar a substituição de aminoácido correspondentemente favorável aí contida.

Num aspecto, após mutação de todas as posições de aminoácido num polipéptido parental utilizando mutagénese de saturação, como aqui divulgada, as alterações de aminoácidos favoráveis podem ser identificadas em mais de uma posição de aminoácido. Um ou mais novas moléculas descendentes podem ser produzidas que contêm uma combinação de todas ou parte destas substituições de aminoácidos favoráveis. Por exemplo, se 2 alterações de aminoácidos favoráveis específicas forem identificadas em cada das 3 posições de aminoácido num polipéptido, as permutações incluem 3 possibilidades em cada posição (nenhuma alteração do aminoácido original e cada das duas alterações favoráveis) e 3 posições. Assim, existem 3 x 3 x 3 ou 27 possibilidades totais, incluindo 7 que foram previamente examinadas - 6 mutações pontuais únicas (*i. e.*, 2 em cada das três posições) e nenhuma alteração em qualquer posição.

Ainda noutro aspecto, a mutagénese de saturação de local pode ser utilizada em conjunto com arranjo combinatório, quimerização, recombinação e outros processos de mutação, em conjunto com rastreio. Esta invenção proporciona a utilização de qualquer processo(s) de mutação, incluindo mutagénese de saturação, de um modo iterativo. Numa exemplificação, a utilização iterativa de qualquer processo(s) de mutação é utilizada em combinação com rastreio.

Os métodos também proporcionam a utilização de iniciadores de codão registados (contendo uma sequência degenerada N,N,N)

para introduzir mutações pontuais num polinucleótido, de modo a gerar um conjunto de polipéptidos descendentes em que uma gama completa de substituições únicas de aminoácidos é representada em cada posição de aminoácido (Mutagénese de Saturação de Local Génico (GSSM)). Os oligos utilizados são compreendidos contiguamente de uma primeira sequência homóloga, uma sequência degenerada N,N,N e num aspecto, mas não necessariamente, numa segunda sequência homóloga. Os produtos de tradução descendentes a jusante a partir da utilização de tais oligos incluem todas as alterações de aminoácidos possíveis em cada local de aminoácido ao longo do polipéptido, porque a degenerescência da sequência N,N,N inclui codões para todos os 20 aminoácidos.

Num aspecto, um tal oligo degenerado (compreendendo uma cassete degenerada N,N,N) é utilizado para submeter cada codão original num molde polinucleotídico parental a uma gama completa de substituições de codão. Num outro aspecto, pelo menos, duas cassetes degeneradas N,N,N são utilizadas - no mesmo oligo ou não, para submeter, pelo menos, dois codões originais num molde polinucleotídico parental a uma gama completa de substituições de codão. Assim, mais de uma sequência N,N,N pode estar contida num oligo para introduzir mutações de aminoácidos em mais de um local. Esta pluralidade de sequências N,N,N pode estar directamente contígua ou separada por uma ou mais sequências nucleotídicas adicionais. Num outro aspecto, os oligos úteis para introduzir adições e deleções podem ser utilizados isoladamente ou em combinação com os codões contendo uma sequência N,N,N para introduzir qualquer combinação ou permutação de adições, deleções e/ou substituições de aminoácidos.

Numa exemplificação particular, é possível mutar simultaneamente duas ou mais posições de aminoácidos contíguas

utilizando um oligo que contenha tripletos N,N,N contíguos, *i. e.*, uma sequência degenerada $(N,N,N)_n$.

Num outro aspecto, os métodos proporcionam a utilização de cassetes degeneradas que têm menos degenerescência do que a sequência N,N,N. Por exemplo, pode ser desejável em alguns casos utilizar (*e. g.*, num oligo) uma sequência degenerada de tripleto compreendendo apenas um N, em que o N pode ser na primeira, segunda ou terceira posição do tripleto. Quaisquer outras bases incluindo quaisquer suas combinações e permutações podem ser utilizadas nas duas posições remanescentes do tripleto. Alternativamente, pode ser desejável em alguns casos utilizar (*e. g.*, num oligo) uma sequência degenerada de tripleto N,N,N, N,N,G/T ou uma sequência de tripleto N,N,G/C.

É entendido, contudo, que a utilização de um tripleto degenerado (tal como uma sequência de tripleto N,N,G/T ou N,N,G/C) como divulgado é vantajosa por diversas razões. Isto proporciona meios para gerar sistematicamente e de forma razoavelmente fácil a substituição da gama completa de aminoácidos possíveis (para um total de 20 aminoácidos) em todas as posições de aminoácido num polipéptido. Assim, para um polipéptido de 100 aminoácidos, proporciona um modo para produzir sistematicamente e de forma razoavelmente fácil 2000 espécies distintas (*i. e.*, 20 aminoácidos possíveis por posição vezes 100 posições de aminoácidos). É entendido que é proporcionado, através da utilização de um oligo contendo um sequência degenerada de tripleto N,N,G/T ou N,N,G/C, 32 sequências individuais que codificam para 20 aminoácidos possíveis. Assim, num vaso reaccional em que uma sequência polinucleotídica parental é submetida mutagénese de saturação utilizando um tal oligo, são produzidos 32 polinucleótidos descendentes distintos que codificam 20 polipéptidos distintos.

Pelo contrário, a utilização de um oligo não degenerado em mutagénese dirigida conduz apenas a um produto polipeptídico descendente por vaso reaccional.

Oligos não degenerados podem ser opcionalmente utilizados em combinação com iniciadores degenerados divulgados. É entendido que em algumas situações é vantajoso utilizar oligos não degenerados para produzir mutações pontuais específicas num polinucleótido de trabalho. Isto proporciona meios para produzir mutações pontuais silenciosas específicas, mutações pontuais que conduzem a alterações de aminoácidos correspondentes e mutações pontuais que provocam a produção de codões de terminação e a correspondente expressão de fragmentos de polipéptido.

Assim, num aspecto desta invenção, cada vaso reaccional de mutagénese de saturação contém polinucleótidos qe codificando, pelo menos, 20 moléculas de polipéptido descendente de modo que todos os 20 aminoácidos estão representados na posição de aminoácido específica que corresponde à posição de codão mutada no polinucleótido parental. Os 32 polipéptidos degenerados descendentes produzidos a partir de cada vaso reaccional de mutagénese de saturação podem ser submetidos a amplificação clonal (*e. g.*, clonados num hospedeiro *E. coli* adequado utilizando um vector de expressão) e submetidos a rastreio de expressão. Quando um polipéptido individual descendente for identificado por rastreio por apresentar uma alteração favorável na propriedade (quando comparada ao polipéptido parental), pode ser sequenciado para identificar a substituição de aminoácido correspondentemente favorável aí contida.

É entendido que após mutação de todas as posições de aminoácido num polipéptido parental utilizando mutagénese de saturação como aqui divulgada, as alterações de aminoácidos

favoráveis podem ser identificadas em mais de uma posição de aminoácido. Uma ou mais novas moléculas descendentes podem ser produzidas que contêm uma combinação de todas ou de uma parte destas substituições de aminoácidos favoráveis. Por exemplo, se 2 alterações de aminoácidos favoráveis específicas forem identificadas em cada de 3 posições de aminoácido num polipéptido, as permutações incluem 3 possibilidades em cada posição (nenhuma alteração do aminoácido original e cada de duas alterações favoráveis) e 3 posições. Assim, existem 3 x 3 x 3 ou 27 possibilidades totais, incluindo 7 que foram previamente examinadas - 6 mutações pontuais únicas (*i. e.*, 2 em cada das três posições) e nenhuma alteração em qualquer posição.

Assim, numa exemplificação não limitante, a mutagénese de saturação é utilizada em combinação com processos de mutação adicionais, tal como o processo em que dois ou mais polinucleótidos relacionados são introduzidos numa célula hospedeira adequada de modo que seja produzido um polinucleótido híbrido por recombinação e rearranjo redutor.

Além disso ao realizar mutagénese ao longo da sequência inteira de um gene, a mutagénese pode ser utilizada para substituir cada de qualquer número de bases numa sequência de polinucleótidos, o número de bases para ser mutado é, num aspecto, todos os números inteiros de 15 a 100000. Assim, em vez de mutar cada posição ao longo de uma molécula, pode-se submeter todas ou de um número discreto de bases (num aspecto, um subconjunto que totaliza de 15 a 100000) a mutagénese. Num aspecto, um nucleótido separado é utilizado para mutar cada posição ou grupo de posições ao longo de uma sequência polinucleotídica. Um grupo de 3 posições a ser mutada pode ser um codão. As mutações podem ser introduzidas utilizando um iniciador mutagénico, contendo uma cassete heteróloga, também

referida como uma cassete mutagénica. Cassetes exemplificativas podem ter de 1 a 500 bases. Cada posição de nucleótido em tais cassetes heterólogas é N, A, C, G, T, A/C, A/G, A/T, C/G, C/T, G/T, C/G/T, A/G/T, A/C/T, A/C/G ou E, em que E é qualquer que não é A, C, G ou T (E pode ser referida como a um oligo desenhador).

Num sentido geral, a mutagénese de saturação é compreendida de mutação de um conjunto completo de cassetes mutagénicas (em que num aspecto cada cassete tem cerca de 1-500 bases de comprimento) numa sequência polinucleotídica definida para ser mutada (em que num aspecto a sequência a ser mutada tem desde cerca de 15 a 100000 bases de comprimento). Assim, um grupo de mutações (variandom de 1 a 100 mutações) é introduzido em cada cassete a ser mutada. Um grupo de mutações a ser introduzido numa cassete pode ser diferente ou igual a um segundo grupo de mutações a ser introduzido numa segunda cassete durante a aplicação de um ciclo de mutagénese de saturação. Tais grupos são exemplificados por deleções, adições, grupos de codões particulares e grupos de cassetes nucleotídicas particulares.

As sequências definidas a serem mutadas incluem um gene inteiro, via, ADNc, uma grelha de leitura aberta inteira (ORF) e promotor, intensificador, repressor/transactivador, origem de replicação, intrão, operador inteiro ou qualquer grupo polinucleotídico funcional. Geralmente, as "sequências definidas" para este objectivo podem ser qualquer polinucleótido que sejam uma sequência polinucleotídica de 15 bases e sequências polinucleotídicas de comprimentos entre 15 bases e 15000 bases (esta invenção refere especificamente cada número inteiro entre esses números). As considerações de escolha de grupos de codões incluem tipos de aminoácidos codificados por uma cassete mutagénica degenerada.

Numa exemplificação, um grupo de mutações que pode ser introduzido numa cassete mutagénica, esta invenção especificamente proporciona substituições de codão degenerado (utilizando oligos degenerados) que codificam 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 e 20 aminoácidos em cada posição e uma biblioteca de polipéptidos assim codificados.

Remontagem de Ligação Sintética (SLR)

A divulgação proporciona um sistema de modificação génica não estocástico denominado "remontagem de ligação sintética" ou simplesmente "SLR", um processo de "evolução dirigida" para gerar polipéptidos, *e. g.*, enzimas fitase da invenção, com propriedades novas ou alteradas. A SLR é um método de ligação de fragmentos oligonucleotídicos, em conjunto, não estocasticamente. Este método difere do arranjo combinatório oligonucleotídico estocástico no sentido dos blocos construtivos de ácidos nucleicos não serem arranjados em combinação, concatenados ou quimerizados aleatoriamente, mas em vez disso são montados não estocasticamente. Ver, *e. g.*, Patentes U.S. Nº 6773900; 6740506; 6713282; 6635449; 6605449; 6537776.

Num aspecto, a SLR compreende os seguintes passos: (a) proporcionar um polinucleótido molde, em que o polinucleótido molde compreende a sequência que codifica um gene homólogo; (b) proporcionar uma pluralidade de polinucleótidos de bloco construtivo, em que os polinucleótidos de bloco construtivo são concebidos para remontarem de forma cruzada com o polinucleótido molde numa sequência predeterminada e um polinucleótido de bloco construtivo compreende uma sequência que é uma variante do gene homólogo e uma sequência homóloga ao

polinucleótido molde que flanqueia a sequência variante; (c) combinar um polinucleótido de bloco construtivo com um polinucleótido molde de modo que o polinucleótido de bloco construtivo se remonte de forma cruzada com o polinucleótido molde para produzir polinucleótidos compreendendo variações de sequências génicas homólogas.

A SLR não depende da presença de níveis elevados de homologia entre polinucleótidos para ser rearranjada. Assim, este método pode ser utilizado para produzir não estocasticamente bibliotecas (ou conjuntos) de moléculas descendentes compreendendo mais de 10^{100} quimeras diferentes. A SLR pode ser utilizada para produzir bibliotecas compreendendo mais de 10^{1000} quimeras descendentes diferentes. Assim, os métodos não estocásticos produzem um conjunto de moléculas de ácidos nucleicos quiméricas finalizadas atingindo uma ordem total de montagem que é escolhida pela concepção. Este método inclui os passos de produzir por concepção uma pluralidade de blocos construtivos de ácidos nucleicos específicos tendo extremidades ligáveis mutuamente compatíveis úteis, e montar estes blocos construtivos de ácidos nucleicos, de modo que seja obtida uma ordem de montagem global concebida. As extremidades ligáveis mutuamente compatíveis dos blocos construtivos de ácidos nucleicos a ser montados, são consideradas como sendo “úteis” para este tipo de montagem ordenada se permitirem aos blocos construtivos serem acoplados em ordens predeterminadas. Assim, a ordem global de montagem em que os blocos construtivos de ácidos nucleicos podem ser acoplados, é especificada pela concepção das extremidades ligáveis. Se for para ser utilizado mais de um passo de montagem, então a ordem global de montagem em que os blocos construtivos de ácidos nucleicos podem ser acoplados é especificada também através da ordem sequencial do passo de montagem. Num aspecto, as partes construtivas emparelhadas são

tratadas com uma enzima, tal como uma ligase (*e. g.*, T4 ADN ligase), para conseguir ligação covalente das partes construtivas. Num aspecto, um método não estocástico denominado remontagem de ligação sintética (SLR), o qual está de certa forma relacionado com arranjo combinatório estocástico, garante que os blocos construtivos de ácidos nucleicos não são arranjados em combinação ou concatenados ou quimerizados aleatoriamente, mas que em vez disso são montados não estocasticamente, pode ser utilizado para criar variantes.

O método de SLR não depende da presença de um nível elevado de homologia entre polinucleótidos a serem arranjados em combinação. A invenção pode ser utilizada para produzir não estocasticamente bibliotecas (ou conjuntos) de moléculas descendentes compreendendo mais de 10^{100} quimeras diferentes. Concebivelmente, a SLR pode ser mesmo utilizada para produzir bibliotecas compreendendo mais de 10^{1000} quimeras descendentes diferentes.

Um método não estocástico produz um conjunto de moléculas de ácidos nucleicos quiméricas finalizadas tendo uma ordem global de montagem que é escolhida por concepção, cujo método compreende passos de produzir por concepção uma pluralidade de blocos construtivos de ácidos nucleicos específicos tendo extremidades ligáveis mutuamente compatíveis úteis e montar estes blocos construtivos de ácidos nucleicos, de modo que seja obtida uma ordem de montagem global concebida.

As extremidades ligáveis mutuamente compatíveis dos blocos construtivos de ácidos nucleicos a serem montados são consideradas como sendo “úteis” para este tipo de montagem ordenada se permitirem que os blocos construtivos sejam acoplados em ordens predeterminadas. Assim, num aspecto, a ordem

global de montagem em que os blocos construtivos de ácidos nucleicos pode ser acoplados é especificada pela concepção das extremidades ligáveis e, se for para ser utilizado mais de um passo de montagem, então a ordem global de montagem em que os blocos construtivos de ácidos nucleicos podem ser acoplados é especificada também pela ordem sequencial do(s) passo(s) de montagem. As partes construtivas emparelhadas podem ser tratadas com uma enzima, tal como uma ligase (*e. g.*, T4 ADN ligase) para conseguir a ligação covalente das partes construtivas.

Num outro aspecto, a concepção de blocos construtivos de ácidos nucleicos é obtida após análise das sequências de um conjunto de moldes de ácidos nucleicos progenitores que servem como uma base para produzir um conjunto de descendentes de moléculas de ácidos nucleicos quiméricas finalizadas. Estes moldes de ácidos nucleicos progenitores servem assim como uma fonte de informação de sequência que ajuda na concepção dos blocos construtivos de ácidos nucleicos que são para serem mutados, *i. e.*, quimerizados ou arranjados em combinação.

Numa exemplificação, o método proporciona a quimerização de uma família de genes relacionados e da sua família codificada de produtos relacionados. Num exemplo particular, os produtos codificados são enzimas. As enzimas e os polipéptidos para utilização na invenção podem ser mutados de acordo com os métodos aqui descritos.

As sequências de uma pluralidade de moldes de ácidos nucleicos progenitores podem ser alinhadas de modo a seleccionar um ou mais pontos de demarcação, cujos pontos de demarcação estão localizados numa área de homologia. Os pontos de demarcação podem ser utilizados para delinear as fronteiras dos blocos construtivos de ácidos nucleicos a serem produzidos.

Assim, os pontos de demarcação identificados e seleccionados nas moléculas progenitoras servem como potenciais pontos de quimerização na montagem das moléculas descendentes.

Tipicamente, um ponto de demarcação útil é uma área de homologia (compreendida de, pelo menos, uma base nucleotídica homóloga) partilhada por, pelo menos, dois moldes progenitores, mas o ponto de demarcação pode ser uma área de homologia que é partilhada por, pelo menos, metade dos moldes progenitores, pelo menos, dois terços dos moldes progenitores, pelo menos, três quartos dos moldes progenitores ou quase todos os moldes progenitores. Num aspecto, um ponto de demarcação útil é uma área de homologia que é partilhada por todos os moldes progenitores.

Num aspecto, o processo de remontagem de ligação é realizado exaustivamente de modo a produzir uma biblioteca exaustiva. Por outras palavras, todas as combinações ordenadas possíveis dos blocos construtivos de ácidos nucleicos são representadas no conjunto de moléculas de ácidos nucleicos quiméricas finalizadas. Ao mesmo tempo, a ordem de montagem (*i. e.*, a ordem de montagem de cada bloco construtivo na sequência 5' para 3 de cada ácido nucleico quimérico finalizado) em cada combinação é por concepção (ou não estocástico). Devido à natureza não estocástica do método, a possibilidade de produtos secundários não desejados é extremamente reduzida.

Noutro aspecto, o método proporciona que, o processo de remontagem de ligação seja realizado sistematicamente, por exemplo, de modo a gerar uma biblioteca sistematicamente compartimentalizada, com compartimentos que podem ser rastreados sistematicamente, *e. g.*, um a um. Por outras palavras, através da utilização selectiva e sensata de blocos construtivos de

ácidos nucleicos específicos, acoplada com a utilização selectiva e sensata de reacções de montagem com passos sequenciais, pode ser obtida uma concepção experimental em que conjuntos específicos de produtos descendentes são feitos em cada de diversos vasos reaccionais. Isto permite que seja realizado um processo sistemático de examinação e selecção. Assim, permite que um número potencialmente muito grande de moléculas descendentes seja examinado sistematicamente em menores grupos.

Devido à sua capacidade para realizar quimerizações de um modo que é altamente flexível, contudo também exaustiva e sistemática, particularmente quando existe um baixo nível de homologia entre as moléculas progenitoras, os presentes métodos proporcionam a produção de uma biblioteca (ou conjunto) compreendida de um grande número de moléculas descendentes. Devido à natureza não estocástica da presente invenção de remontagem de ligação, as moléculas descendentes produzidas podem compreender uma biblioteca de moléculas de ácidos nucleicos quiméricas finalizadas tendo uma ordem global de montagem que é escolhida por concepção. Num aspecto particular, uma tal biblioteca gerada é compreendida de mais de 10^3 até mais de 10^{1000} espécies moleculares descendentes diferentes.

Num aspecto, um conjunto de moléculas de ácidos nucleicos quiméricas finalizadas, produzidas como descrito é compreendido de um polinucleótido codificando um polipéptido. De acordo com um aspecto, este polinucleótido é um gene, o qual pode ser um gene feito pelo homem. De acordo com um outro aspecto, este polinucleótido é uma via génica, a qual pode ser uma via génica feita pelo homem. Um ou mais genes feitos pelo homem produzidos podem ser incorporados numa via génica feita pelo homem, tal

como uma via operável num organismo eucariótico (incluindo uma planta).

Noutra exemplificação, a natureza sintética do passo em que os blocos construtivos são produzidos permite a concepção e introdução de nucleótidos (*e. g.*, um ou mais nucleótidos que podem ser, por exemplo, codões ou intrões ou sequências reguladoras) que podem posteriormente ser opcionalmente removidos num processo *in vitro* (*e. g.*, por mutagénese) ou num processo *in vivo* (*e. g.*, por utilização da capacidade de *splice* do gene de um organismo hospedeiro). É entendido que em muitos casos a introdução destes nucleótidos pode, também, ser desejável por muitas outras razões para além do potencial benefício de criação de um ponto de demarcação útil.

Um ácido nucleico de bloco construtivo pode ser utilizado para introduzir um intrão. Os intrões funcionais podem ser introduzidos num gene da invenção feito pelo homem. Intrões funcionais podem ser introduzidos numa via génica feita pelo homem. Os polinucleótidos quiméricos que são genes feitos pelo homem contêm um (ou mais) intrão(ões) introduzido(s) artificialmente.

Desta forma, a divulgação também proporciona a produção de um polinucleótido quimérico que é uma via génica feita pelo homem contendo um (ou mais) intrão(ões) introduzido(s) artificialmente. Num aspecto, o(s) intrão(ões) introduzido(s) artificialmente é(são) funcional(ais) em uma ou mais células hospedeiras para *splice* génico de mesmo modo que intrões de ocorrência natural servem funcionalmente no *splice* génico. Um processo é para produzir polinucleótidos feitos pelo homem contendo intrão, para serem introduzidos em organismos hospedeiros para recombinação e/ou *splice*.

Um gene feito pelo homem produzido pode servir também como um substrato para recombinação com outro ácido nucleico. Do mesmo modo, uma via génica feita pelo homem produzida pode servir também como um substrato para recombinação com outro ácido nucleico. Num aspecto, a recombinação é facilitada ou ocorre em áreas de homologia entre o gene feito pelo homem contendo intrão e um ácido nucleico que serve como um parceiro de recombinação. Num aspecto, o parceiro de recombinação pode também ser um ácido nucleico produzido pela invenção, incluindo um gene feito pelo homem ou uma via génica feita pelo homem. A recombinação pode ser facilitada ou pode ocorrer em áreas de homologia que existem em um ou mais intrões introduzidos artificialmente no gene feito pelo homem.

O método feito pelo homem de remontagem de ligação utiliza uma pluralidade de blocos construtivos de ácidos nucleicos, cada um dos quais pode ter duas extremidades ligáveis. As duas extremidades ligáveis em cada bloco construtivo de ácido nucleico podem ser duas extremidades rombas (*i. e.*, cada uma tendo uma saliência de zero nucleótidos) ou uma extremidade romba e uma saliente, ou duas salientes.

Uma saliência útil para este objectivo pode ser uma saliência 3' ou uma saliência 5'. Assim, um bloco construtivo de ácido nucleico pode ter uma saliência 3' ou alternativamente uma saliência 5' ou alternativamente duas saliências 3' ou alternativamente duas saliências 5'. A ordem global em que os blocos construtivos de ácidos nucleicos são montados para formar uma molécula de ácido nucleico quimérica finalizada é determinada pela concepção experimental proposta e não é aleatória.

Num aspecto, um bloco construtivo de ácido nucleico é gerado por síntese química de dois ácidos nucleicos de cadeia simples (também referidos como oligos de cadeia simples) e colocando-os em contacto para permitir que emparelhem para formar um bloco construtivo de ácido nucleico de cadeia dupla.

Um bloco construtivo de ácido nucleico de cadeia dupla pode ser de tamanho variável. Os tamanhos destes blocos construtivos podem ser pequenos ou grandes. Os tamanhos exemplificativos para blocos construtivos variam de 1 par de bases (não incluindo quaisquer saliências) a 100000 pares de bases (não incluindo quaisquer saliências). Outras gamas de tamanho são também proporcionadas, a qual tem limites inferiores de desde 1 pb a 10000 pb (incluindo todos valores inteiros entre estes) e limites superiores desde 2 pb a 100000 pb (incluindo todos valores inteiros entre estes).

Existem muitos métodos através dos quais um bloco construtivo de ácido nucleico de cadeia dupla pode ser produzido que sejam úteis para a invenção; e estes são conhecidos na técnica e pode ser prontamente realizados por um especialista.

De acordo com um aspecto, um bloco construtivo de ácido nucleico de cadeia dupla é gerado através primeiro de produzir dois ácidos nucleicos de cadeia simples e permitindo que estes emparelhem para formar um bloco construtivo de ácido nucleico de cadeia dupla. As duas cadeias de um bloco construtivo de ácido nucleico de cadeia dupla podem ser complementares em cada nucleótido à excepção daqueles que formam uma saliência; não contendo assim quaisquer falsos emparelhamentos, para além de quaisquer saliências. De acordo com um outro aspecto, as duas cadeias de um bloco construtivo de ácido nucleico de cadeia dupla são complementares em menos do que todos os nucleótidos à

excepção daqueles que formam uma saliência. Assim, de acordo com este aspecto, um bloco construtivo de ácido nucleico de cadeia dupla pode ser utilizado para introduzir degenerescência de codão. Num aspecto, a degenerescência de codão é introduzida utilizando mutagénese de saturação de local aqui descrita, utilizando uma ou mais cassetes N,N,G/T ou, alternativamente utilizando uma ou mais cassetes N,N,N.

O método de recombinação *in vivo* pode ser realizado de forma cega num agrupamento de híbridos ou alelos desconhecidos de um polinucleótido ou sequência específica. Contudo, não é necessário saber a sequência de ADN ou ARN concreta do polinucleótido específico.

A abordagem de utilizando de recombinação dentro de uma população mista de genes pode ser útil para a produção de quaisquer proteínas úteis, por exemplo, interleucina I, anticorpos, tPA e hormona do crescimento. Esta abordagem pode ser utilizada para produzir proteínas que têm a especificidade ou a actividade alterada. A abordagem pode também ser útil para produção de sequências híbridas de ácidos nucleicos, por exemplo, regiões promotoras, intrões, exões, sequências intensificadoras, regiões 3' não traduzidas ou regiões 5' não traduzidas de genes. Assim esta abordagem pode ser utilizada para produzir genes tendo velocidades de expressão aumentadas. Esta abordagem pode também ser útil no estudo de sequências de ADN repetitivas. Finalmente, esta abordagem pode ser útil para mutar ribozimas ou aptâmeros.

Num aspecto, as variantes dos polinucleótidos e polipéptidos aqui descritos são obtidos por utilização de ciclos repetidos de rearranjo redutor, recombinação e selecção que permitem a

evolução molecular dirigida de sequências lineares altamente complexas, tais como ADN, ARN ou proteínas por recombinação.

O arranjo combinatório *in vivo* das moléculas é útil a proporcionar variantes e pode ser realizado utilizando a propriedade natural de células para recombinarem multimêros. Embora a recombinação *in vivo* tenha proporcionado a maior via natural para diversidade molecular, a recombinação genética permanece um processo relativamente complexo que envolve 1) o reconhecimento de homologias; 2) clivagem de cadeias, invasão de cadeias e passos metabólicos que conduzem à produção de quiasma recombinante; e finalmente 3) a resolução do quiasma em moléculas recombinadas discretas. A formação de quiasma requer o reconhecimento de sequências homólogas.

Noutro aspecto, a divulgação inclui um método para produzir um polinucleótido híbrido a partir de, pelo menos, um primeiro polinucleótido e de um segundo polinucleótido. Isto pode ser utilizado para produzir um polinucleótido híbrido por introdução de, pelo menos, um primeiro polinucleótido e um segundo polinucleótido que partilha, pelo menos, uma região de homologia de sequência parcial (*e. g.*, SEQ ID N°:1) numa célula hospedeira adequada. As regiões de homologia de sequência parcial promovem processos que resultam em reorganização de sequência produzindo um polinucleótido híbrido. A expressão “polinucleótido híbrido”, como aqui utilizada, é qualquer sequência de nucleótidos que resulte do método da presente invenção e contenha a sequência de, pelo menos, duas sequências polinucleotídicas originais. Tais polinucleótidos híbridos podem resultar de eventos de recombinação intermolecular que promovem a integração de sequências entre moléculas de ADN. Além disso, tais polinucleótidos híbridos podem resultar de processos de rearranjo redutor intramolecular que utilizam sequências

repetidas para alterar uma sequência de nucleótidos dentro de uma molécula de ADN.

A divulgação proporciona métodos para produzir polinucleótidos híbridos que podem codificar polipéptidos híbridos biologicamente activos (*e. g.*, uma fitase híbrida). Num aspecto, os polinucleótidos originais codificam polipéptidos biologicamente activos. O método produz novos polipéptidos híbridos através da utilização de processos celulares que integram a sequência dos polinucleótidos originais de modo que o polinucleótido híbrido resultante codifique um polipéptido que demonstra actividades derivadas dos polipéptidos biologicamente activos originais. Por exemplo, os polinucleótidos originais podem codificar uma enzima particular de diferentes microrganismos. Uma enzima codificada por um primeiro polinucleótido de um organismo ou variante pode, por exemplo, funcionar eficazmente sob uma condição ambiental particular, *e. g.*, elevada salinidade. Uma enzima codificada por um segundo polinucleótido de um organismo ou variante diferente, pode funcionar eficazmente sob uma condição ambiental diferente, tal como temperaturas muito altas. Um polinucleótido híbrido contendo sequências dos primeiros e segundos polinucleótidos originais podem codificar uma enzima que apresenta características de ambas as enzimas codificadas pelos polinucleótidos originais. Assim, a enzima codificada pelo polinucleótido híbrido pode funcionar eficazmente sob condições ambientais partilhadas por cada das enzimas codificadas pelos primeiros e segundos polinucleótidos, *e. g.*, elevada salinidade e temperaturas extremas.

Para além dos vários métodos descritos acima, são conhecidos na técnica vários métodos que podem ser utilizados para obter polinucleótidos híbridos com propriedades enzimáticas

aumentadas. Os seguintes exemplos ilustram a utilização de tais processos para obter enzimas termoestáveis ou termotolerantes por mutagénese de um polinucleótido codificando uma enzima de tipo selvagem de interesse.

Por exemplo, num aspecto, a invenção utiliza métodos como descritos por M. Lehmann *et al.* (em Biochimica et Biophysica Acta 1543:408-415, 2000) que descreve uma "abordagem de consenso", em que o alinhamento de sequências de fitases fúngicas homólogas foi utilizado para calcular uma sequência de aminoácidos de fitase de consenso. Após construção do gene de consenso correspondente, expressão recombinante e purificação, a fitase recombinante obtida apresentou uma temperatura de desenrolamento (Tm) 15-22 °C mais alta do que aquela de todas as fitases parentais utilizadas na concepção. Mutagénese dirigida do gene codificando a proteína recombinante foi utilizada para aumentar mais o valor de Tm para 90,4 °C. O efeito termoestabilizante foi atribuído a uma combinação de múltiplas trocas de aminoácidos que foram distribuídas sobre toda a sequência da proteína e que afectou principalmente os resíduos expostos à superfície.

Os métodos para obter uma enzima com propriedades térmicas aumentadas são descritos por L. Jermutus *et al.* (J. of biotechnology 85:15-24, 2001). Nesta abordagem as interacções iónicas e as ligações de hidrogénio na superfície da fitase de *Aspergillus terreus* foram primeiro restauradas para corresponder aquelas presentes na enzima homóloga, mas mais termostável, de *A. niger*. Depois a totalidade dos elementos estruturais secundários foram substituídos na mesma região e com base na estrutura cristalina da fitase de *A. niger*. A substituição de uma α-hélice na superfície da fitase de *A. terreus* pelo trecho correspondente da fitase de *A. niger* resultou numa enzima

quimérica baseada na estrutura (proteína de fusão) com termostabilidade melhorada e actividade enzimática inalterada.

L. Giver *et al.* (Proc. Natl. Acad. Sci. USA 95:12809-12813, 1998), descrevem um processo em que seis gerações de mutagénese aleatória introduzidas durante PCR mutagénico de um polinucleótido codificando a *p*-nitrobenzilo esterase de *Bacillus subtilis*, seguida por recombinação *in vitro* baseada no método de Stemmer resultaram numa esterase recombinante com termostabilidade aumentada (aumento de mais de 14 °C na Tm) sem comprometer a actividade catalítica a temperaturas mais baixas.

C. Vetriani *et al.* (Proc. Natl. Acad. Sci. USA 95:12300-12305, 1998) descrevem um processo através do qual modelação baseada em homologia e comparação estrutural directa das glutamato desidrogenases hexaméricas dos hipertermófilos *Pyrococcus furiosus* e *Thermococcus litoralis*, com temperaturas óptimas de crescimento de 100 °C e 88 °C, respectivamente, foram utilizadas para determinar características termoestabilizantes chave. Uma rede de pares iónicos inter-subunidades observada como sendo substancialmente reduzida na enzima menos estável foi alterada por mutagénese de dois resíduos aí presentes para restaurar as interacções encontradas na enzima mais estável. Embora cada das mutações simples tivesse efeitos adversos na termostabilidade, com ambas as mutações em simultâneo foi observada uma melhoria de quatro vezes na estabilidade a 104 °C relativamente à enzima de tipo selvagem.

A. Tomschy *et al.* (Protein Science 9:1304-1311, 2000), descrevem um processo que utiliza a estrutura cristalina da fitase de *Aspergillus niger* (à resolução de 2,5 angstroms) para especificar todos os locais activos da enzima. Um alinhamento de múltiplas sequências de aminoácidos foi depois utilizado para

identificar os resíduos não conservados do sítio activo que se possam correlacionar com uma dada propriedade favorável de interesse. Utilizando esta abordagem, Gln27 da fitase de *A. fumigatus*, que difere de Leu27 de *A. niger*, foi identificado como estando provavelmente envolvido na ligação e/ou libertação do substrato e responsável pela actividade específica mais baixa da fitase de *A. fumigatus* (26,5 vs. 196 6 U/mg de proteína a pH 5,0). Mutagénese dirigida de Gln27 da fitase de *A. fumigatus* para leu aumentou a actividade específica da enzima mutante para 92,1 U/mg de proteína.

Plantas e Sementes Transgénicas

A invenção proporciona plantas e sementes transgénicas compreendendo um ácido nucleico, um polipéptido, uma cassete de expressão, mecanismo de clonagem ou vector da invenção, ou uma célula transfectada ou transformada da invenção. A invenção também proporciona produtos vegetais, *e. g.*, óleos, sementes, folhas, extractos e semelhantes, compreendendo um ácido nucleico e/ou um polipéptido da invenção. A planta transgénica pode ser dicotiledónea (uma dicot) ou monocotiledónea (uma monocot). A invenção também proporciona métodos para fazer e utilizar estas plantas e sementes transgénicas. A planta ou célula vegetal transgénica que expressa um polipéptido da presente invenção pode ser construída de acordo com qualquer método conhecido na técnica. Ver, por exemplo, Patente U.S. N° 6309872.

A expressão recombinante ou sobreexpressão das moléculas de fitase da invenção pode ser conseguida em combinação com uma ou mais moléculas adicionais como, por exemplo, outras enzimas. Esta abordagem é útil para produzir produtos de combinação, tal como uma planta ou uma parte de planta que contenha as presentes

moléculas de fitase, assim como uma ou mais moléculas adicionais. As moléculas de fitase desta invenção e as moléculas adicionais podem ser utilizadas num tratamento de combinação. As moléculas expressas recombinantemente resultantes podem ser utilizadas na forma homogeneizada e/ou purificada ou alternativamente na forma relativamente não purificada (*e. g.*, como partes vegetais consumíveis que são úteis quando misturadas com outros géneros alimentícios para catalisar a degradação do fitato).

Num aspecto particular, a presente invenção proporciona a expressão de fitase em plantas ou órgãos vegetais transgénicos e métodos para a sua produção. As construções de expressão de ADN são proporcionadas para a transformação de plantas com um gene que codifica uma fitase sob controlo de sequências reguladoras que são capazes de dirigir a expressão da fitase. Estas sequências reguladores incluem sequências capazes de dirigir a transcrição em plantas, constitutivamente ou de modo específico de fase e/ou tecido.

O modo de expressão depende, em parte, da utilização da planta ou das suas partes. As plantas e órgãos vegetais transgénicos, proporcionados pela presente invenção, podem ser aplicados a uma variedade de processos industriais directamente, *e. g.*, em rações para animais ou, alternativamente, a fitase expressa pode ser extraída e, se desejado, purificada antes da aplicação. Alternativamente, a planta hospedeira ou a parte vegetal recombinante podem ser utilizados directamente. Num aspecto particular, a presente invenção proporciona métodos para catalisar reacções de hidrólise de fitato utilizando sementes contendo quantidades aumentadas de fitase. O método envolve colocar em contacto sementes transgénicas de tipo não selvagem, *e. g.*, numa forma moída ou mascada, com o substrato contendo

fitato e permitir que as enzimas nas sementes aumentem a velocidade da reacção. Ao adicionar directamente as sementes a um substrato contendo fitato, à invenção proporciona uma solução para o processo dispendioso e problemático de extrair e purificar a enzima. Numa exemplificação, a presente invenção proporciona métodos de tratamento através dos quais um organismo que carece de um fornecimento suficiente de uma enzima é administrado a enzima na forma de sementes contendo quantidades aumentadas da enzima. Num aspecto, o momento da administração da enzima a um organismo é coordenado com o consumo de um género alimentício contendo fitato.

A expressão de fitase em plantas pode ser realizada por uma variedade de meios. Especificamente, por exemplo, estão disponíveis tecnologias para transformar um grande número de espécies de plantas, incluindo espécies dicotiledóneas (*e. g.*, tabaco, batata, tomate, petúnia, *Brassica*) e espécies monocot. Adicionalmente, por exemplo, estão disponíveis estratégias para a expressão de genes estranhos em plantas. Ainda adicionalmente, foram identificadas sequências reguladoras de genes de plantas que são úteis para a construção de genes quiméricos que podem ser funcionalmente expressos em plantas e em células vegetais (*e. g.*, Klee (1987) Ann. Rev. of Plant Phys. 38:467-486; Clark *et al.* (1990) Virology Dez; 179(2):640-7; Smith *et al.* (1990) Mol. Gen. Genet. Dez; 224(3):477-81.

A introdução de construções génicas em plantas pode ser realizada utilizando diversas tecnologias incluindo transformação com *Agrobacterium tumefaciens* ou *Agrobacterium rhizogenes.* Exemplos não limitantes de tecidos vegetais que podem ser assim transformados incluem protoplastos, microsporos ou pólen, e explantes, tais como folhas, caules, raízes, hipocótilos e cótilos. Além disso, o ADN pode ser introduzido

directamente em protoplastos e em células ou tecidos vegetais por microinjecção, electroporação, bombardeamento de partículas e incorporação directa de ADN.

As proteínas podem ser produzidas em plantas através de uma variedade de sistemas de expressão. Por exemplo, a utilização de um promotor constitutivo, tal como o promotor 35S do vírus do mosaico da couve-flor (Guilley *et al.*, 1982) é útil para a acumulação da proteína expressa em virtualmente todos os órgãos vegetais transgénicos. Alternativamente, a utilização de promotores que são altamente específicos de tecido e/ou específicos de fase é útil para esta invenção (Higgins, 1984; Shotwell, 1989) de modo a direccionar a expressão para tecidos desejados e/ou para uma fase desejada do desenvolvimento. A invenção também utiliza protocolos para expressão em plantas de moléculas de fitase da presente invenção, como divulgado, por exemplo, na Patente U.S. N° 5770413 (Van Ooijen *et al.*) e Patente U.S. N° 5593963 (Van Ooijen *et al.*), as quais ensinam a utilização de fitases fúngicas.

Modificação de Sequências Codificantes e Sequências Adjacentes

A expressão transgénica em plantas de genes derivados de fontes heterólogas pode envolver a modificação daqueles genes para conseguir e optimizar a sua expressão em plantas. Em particular, as ORF bacterianas que codificam enzimas separadas, mas que são codificadas pelo mesmo transcripto no micróbio nativo, são melhor expressas em plantas em transcriptos separados. Assim, num aspecto, para conseguir isto, cada ORF microbiana é isolada individualmente e clonada dentro de uma cassete que proporciona uma sequência promotora da planta na

extremidade 5’ da ORF e um terminador transcripcional de planta na extremidade 3’ da ORF. A sequência de ORF isolada pode incluir o codão de iniciação ATG e o codão de terminação STOP, mas pode incluir sequência adicional além do ATG de iniciação e o codão de STOP. Além disso, a ORF pode ser truncada, mas reter ainda a actividade requerida; para ORF particularmente longas, as versões truncadas que retêm actividade podem ser preferíveis para a expressão em organismos transgénicos. “Promotores de plantas” e “terminadores transcricionais de planta” que podem ser utilizados para realizar esta invenção incluem quaisquer promotores e/ou terminadores transcricionais que operam dentro de células vegetais. Isto inclui promotores e terminadores de transcrição que podem ser derivados de fontes não vegetais, tal como vírus (um exemplo é o vírus do mosaico da couve-flor).

Em alguns casos, não é requerida a modificação de sequências codificantes da ORF e sequências adjacentes. É suficiente isolar um fragmento contendo a ORF de interesse e introduzi-lo a jusante de um promotor vegetal. Por exemplo, Gaffney *et al.* (Science 261: 754-756 (1993)) expressou o gene *nahG* de *Pseudomonas* em plantas transgénicas sob o controlo do promotor 35S de CaMV e do terminador *tml* de CaMV com sucesso, sem modificação da sequência codificante e com nucleótidos do gene de *Pseudomonas* a montante do ATG ainda ligado, e com os nucleótidos a jusante de codão STOP ainda ligados à ORF *nahG*. De um modo preferido, deve ser deixada a menor sequência microbiana adjacente possível ligada a montante do ATG e a jusante do codão STOP. Na prática, tal construção pode depender da disponibilidade de locais de restrição.

Em outros casos, a expressão de genes derivados de fontes microbianas pode proporcionar problemas na expressão. Estes problemas foram bem caracterizados na técnica e são

particularmente comuns com genes derivados de determinadas fontes microbianas. Estes problemas podem-se aplicar à sequência de nucleótidos desta invenção e a modificação destes genes pode ser realizada utilizando técnicas agora bem conhecidas no campo técnico. Os seguintes problemas podem ser encontrados:

Utilização de Codão

A invenção proporciona ácidos nucleicos tendo codões modificados para utilização em plantas; em alguns casos a utilização preferida de codão em plantas difere da utilização preferida de codão em determinados microrganismos. A comparação da utilização de codões dentro de uma ORF microbiana clonada para utilização em genes de planta (e em particular em genes da planta alvo) permitirá uma identificação dos codões dentro da ORF que, de um modo preferido, devem ser mudados. Tipicamente, a evolução das plantas tendeu para uma forte preferência dos nucleótidos C e G na terceira posição de base das monocotiledóneas, visto que as dicotiledóneas utilizam frequentemente os nucleótidos A ou T nesta posição. Ao modificar um gene para incorporar uma utilização de codão preferida para uma espécie transgénica alvo particular, muitas dos problemas descritos abaixo para o teor de GC/AT e *splice* ilegítimo serão superados.

Teor de GC/AT

A invenção proporciona ácidos nucleicos que têm o seu teor de GC modificado, *e. g.*, para utilização em plantas; os genes de plantas têm tipicamente um teor de GC superior a 35%. As sequências de ORF que são ricas em nucleótidos A e T podem provocar diversos problemas em plantas. Primeiro, pensa-se que os motivos ATTTA provocam destabilização dos mensageiros e são encontrados na extremidade 3' de muitos mARN de vida curta. Segundo, pensa-se que a ocorrência de sinais de poliadenilação, tal como AATAAA, em posições inapropriadas dentro do mensageiro provoca truncamento prematuro da transcrição. Além disso, as monocotiledóneas podem reconhecer sequências ricas em AT como locais de *splice* (ver abaixo).

Sequências Adjacentes à Metionina de Iniciação

A invenção proporciona ácidos nucleicos que têm nucleótidos adjacentes ao ATG modificado e/ou adicionado; as plantas diferem de microrganismos uma vez que os seus mensageiros não possuem um local definido de ligação a ribossoma. Em vez disso, pensa-se que os ribossomas se ligam à extremidade 5' do mensageiro e fazem um varrimento até ao primeiro ATG disponível em que começa a tradução. Não obstante, pensa-se que existe uma preferência para determinados nucleótidos adjacentes ao ATG e que a expressão microbiana de genes pode ser aumentada por inclusão de um iniciador de tradução de consenso eucariótico no ATG. A Clontech (catálogo de 1993/1994, página 210, aqui incorporado por referência) sugeriu uma sequência como um iniciador de tradução de consenso para a expressão do gene *uidA* de *E. coli* em plantas. Além disso, Joshi (N.A.R. 15: 6643-6653 (1987)) comparou muitas sequências de planta adjacentes ao ATG e sugere

outra sequência de consenso. Nas situações em que são encontradas dificuldades na expressão da ORF microbiana em plantas, a inclusão de uma destas sequências no ATG de iniciação pode melhorar a tradução. Nesses casos, os últimos três nucleótidos de consenso não podem ser apropriados para inclusão na sequência modificada devido à sua modificação do segundo resíduo AA. Em alguns aspectos, as sequências preferidas adjacentes à metionina de iniciação podem diferir entre espécies diferentes de plantas. Uma análise a 14 genes de milho localizados na base de dados GenBank proporcionou os seguintes resultados:

Posição Antes do ATG de Iniciação em 14 Genes de Milho:

	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1
C	3	8	4	6	2	5	6	0	10	7
T	3	0	3	4	3	2	1	1	1	0
A	2	3	1	4	3	2	3	7	2	3
G	6	3	6	0	6	5	4	6	1	5

Esta análise pode ser feita para a espécie desejada de planta em que a sequência de nucleótidos vai ser incorporada e a sequência adjacente ao ATG modificada para incorporar os nucleótidos preferidos.

Remoção de Locais de *Splice* Ilegítimos

A invenção proporciona ácidos nucleicos tendo locais de *splice* ilegítimos modificados ou removidos ou funcionalmente "inactivados"; os genes clonados de fontes não vegetais e não optimizados para a expressão em plantas podem também conter motivos que podem ser reconhecidos em plantas como locais de *splice* 3' ou 5' e são clivados, produzindo assim mensageiros

truncados ou suprimidos. Estes locais podem ser removidos utilizando as técnicas bem conhecidas no campo técnico.

As técnicas para modificação de sequências codificantes e de sequências adjacentes são bem conhecidos na técnica. Nos casos em que a expressão inicial de uma ORF microbiana é baixa e é considerada como apropriada para fazer alterações à sequência como descrito acima, então a construção de genes sintéticos pode ser realizada de acordo com métodos bem conhecidos na técnica. Estes são, por exemplo, descritos nas divulgações de patente publicadas EP 0385962 (da Monsanto), EP 0359472 (da Lubrizol) e WO 93/07278 (da Ciba-Geigy). Na maioria dos casos é preferido ensaiar a expressão das construções génicas utilizando protocolos de ensaio transiente (que são bem conhecidos na técnica) antes da sua transferência a plantas transgénicas.

Promotores de Plantas

As composições da invenção podem conter sequências de ácidos nucleicos, *e. g.*, promotores, *e. g.*, para transformação e expressão numa planta de interesse. As sequências de ácidos nucleicos podem estar presentes em construções de ADN ou em cassetes de expressão. Os ácidos nucleicos da invenção podem ser ou compreender “cassetes de expressão”, incluindo qualquer molécula de ácido nucleico capaz de dirigir a expressão de uma sequência de nucleótidos particular numa célula hospedeira apropriada compreendendo um promotor ligado operativamente à sequência de nucleótidos de interesse que está ligada operativamente a sinais de terminação.

As composições (*e. g.*, sequências de ácidos nucleicos) da invenção também podem compreender sequências requeridas para a

tradução apropriada de sequências de nucleótidos. A região codificante habitualmente codifica para uma proteína de interesse mas pode também codificar para um ARN funcional de interesse, por exemplo, ARN anti-sentido ou ARN não traduzido, na direcção de sentido ou anti-sentido. A cassete de expressão compreendendo a sequência de nucleótidos de interesse pode ser quimérica, significando que, pelo menos, um dos seus componentes é heterólogo relativamente a, pelo menos, um dos seus outros componentes. A cassete de expressão pode também ser uma que é de ocorrência natural, mas que foi obtida de forma recombinante útil para a expressão heteróloga. Tipicamente, contudo, a cassete de expressão é heteróloga relativamente ao hospedeiro, *i. e.*, a sequência de ADN particular da cassete de expressão não ocorre naturalmente na célula hospedeira e deve ter sido introduzida na célula hospedeira ou num antepassado da célula hospedeira através de um evento de transformação. A expressão da sequência de nucleótidos na cassete de expressão pode estar sob controlo de um promotor constitutivo ou de um promotor induzível que inicie a transcrição apenas quando a célula hospedeira é exposta a algum estímulo externo particular. Adicionalmente, o promotor pode também ser específico para um tecido ou órgão ou fase particular do desenvolvimento.

A presente invenção abrange a transformação de plantas com as cassetes de expressão capazes de expressar polinucleótidos. A cassete de expressão incluirá no sentido 5'-3' de transcrição, de uma região de iniciação de transcrição e tradução (*i. e.*, um promotor) e um polinucleótido de interesse. A cassete de expressão pode opcionalmente compreender uma região de terminação de transcrição e tradução (*i. e.*, região de terminação) funcional em plantas. Em algumas formas de realização, a cassete de expressão compreende um gene marcador de selecção para permitir o rastreio para transformantes

estáveis. As construções de expressão da invenção podem também compreender uma sequência líder e/ou uma sequência que permite a expressão induzível do polinucleótido de interesse. Ver, Guo *et al*. (2003) Plant J. 34:383-92 e Chen *et al*. (2003) Plant J. 36:731-40 para exemplos de sequências que permitem a expressão induzível.

As sequências reguladoras da construção de expressão estão operativamente ligadas ao polinucleótido de interesse. Por "operativamente ligada" entende-se uma ligação funcional entre um promotor e uma segunda sequência, em que a sequência promotora inicia e medeia a transcrição da sequência de ADN que corresponde à segunda sequência. Geralmente, operativamente ligada significa que as sequências de nucleótidos ligadas são contíguas.

Qualquer promotor capaz de dirigir a expressão na planta de interesse pode ser utilizado para realizar a invenção. O promotor pode ser nativo ou análogo ou estranho ou heterólogo ao hospedeiro vegetal. Os termos "heterólogo" e "exógeno", quando aqui utilizados, referem-se a uma sequência de ácidos nucleicos (*e. g.*, uma sequência de ADN ou ARN) ou a um gene, referem-se a uma sequência que é originária de uma fonte estranha à célula hospedeira particular ou, se da mesma fonte, for modificada da sua forma original. Assim, um gene heterólogo numa célula hospedeira inclui um gene que seja endógeno à célula hospedeira particular mas foi modificado. Os termos incluem também múltiplas cópias de ocorrência não natural de uma sequência de ADN natural. Assim, os termos referem-se a um segmento de ADN que é estranho ou heterólogo à célula, ou homólogo à célula mas numa posição dentro do ácido nucleico da célula hospedeira em que o elemento não é habitualmente encontrado. Os segmentos exógenos de ADN são expressos para produzir polipéptidos

exógenos. Em formas de realização alternativas, uma sequência de ácidos nucleicos "homóloga" (*e. g.*, ADN) é uma sequência de ácidos nucleicos (*e. g.*, ADN ou ARN) associada naturalmente com uma célula hospedeira em que é introduzida.

A escolha dos promotores para serem incluídos depende de diversos factores, incluindo, mas não limitados a, eficiência, selectabilidade, induzibilidade, nível desejado de expressão e expressão preferencial de célula ou tecido. É uma matéria de rotina para um especialista na técnica, modular a expressão de uma sequência através de seleccionar e posicionar apropriadamente promotores e outras regiões reguladoras relativamente a essa sequência.

Alguns promotores adequados apenas iniciam a transcrição, ou predominantemente, em determinados tipos de células. Assim, como aqui utilizado, um promotor preferencial de tipo de célula ou tecido é um que dirige a expressão, de um modo preferido, no tecido alvo, mas pode também conduzir a alguma expressão noutros tipos de células ou tecidos. Os métodos para identificar e caracterizar regiões promotoras no ADN genómico de plantas incluem, por exemplo, aqueles descritos nas seguintes referências: Jordano, *et al.*, Plant Cell, 1:855-866 (1989); Bustos, *et al.*, Plant Cell, 1:839-854 (1989); Green, *et al.*, EMBO J. 7, 4035-4044 (1988); Meier, *et al.*, Plant Cell, 3, 309-316 (1991); e Zhang, *et al.*, Plant Physiology 110: 1069-1079 (1996).

Diversos promotores e/ou genes regulados de forma preferida em tecidos foram descritos em plantas. Alguns genes preferidos de tecidos descritos incluem os genes que codificam as proteínas de armazenamento de semente (tais como napina, cruciferina, beta-conglicinina e faseolina, prolaminas, glutelinas,

globulinas e zeínas) ou proteínas de corpo oleoso (tal como oleosina) ou genes envolvidos na biossíntese de ácidos gordos (incluindo proteína transportadora de acilo, estearoil-ACP desaturase e desaturases de ácido gordos (fad 2-1)) e outros genes expressos durante o desenvolvimento do embrião (tal como Bce4, ver, por exemplo, documento EP 255378 e Kridl *et al.*, (1991) Seed Science Research, 1:209).

Os exemplos de promotores específicos de tecido que foram descritos, incluem a lecitina (Vodkin, Prog. Clin. Biol. Res., 138;87 (1983); Lindstrom *et al.*, (1990) Der. Genet., 11:160), álcool desidrogenase I de milho (Dennis *et al.*, Nucleic Acids Res., 12:3983 (1984)), complexo de captura de luz de milho (ver, *e. g.*, Simpson, (1986) Science, 233:34; Bansal (1992) Proc. Natl. Acad. Sci. USA 89:3654), proteína de choque térmico de milho (ver, *e. g.*, Odell *et al.*, (1985) Nature, 313:810); subunidade pequena de RuBP carboxilase de ervilha (ver, *e. g.*, Poulsen *et al.*, (1986) Mol. Gen. Genet., 205:193-200; Cashmore *et al.*, (1983) Gen. Eng. of Plants, Plenum Press, Nova Iorque, 29-38); manopina Sintase de plasmídeo Ti (ver, *e. g.*, Langridge *et al.*, (1989) Proc. Natl. Acad. Sci. USA, 86:3219-3223), nopalina Sintase de plasmídeo Ti (Langridge *et al.*, (1989) Proc. Natl. Acad. Sci. USA, 86:3219-3223), chalcone isomerase de petúnia (ver, *e. g.*, van Tunen (1988) EMBO J. 7:1257); proteína 1 rica em glicina de feijão (ver, *e. g.*, Keller (1989) Genes Dev. 3:1639); 35s truncado de CaMV (ver, *e. g.*, Odell (1985) Nature 313:810); patatina de batata (ver, *e. g.*, Wenzler (1989) Plant Mol. Biol. 13:347); célula de raiz (ver, *e. g.*, Yamamoto (1990) Nucleic Acids Res. 18:7449); zeína de milho (ver, *e. g.*, Reina (1990) Nucleic Acids Res. 18:6425; Lopes *et al.* (1995) Mol. Gen. Genet. 247: 603-613; Kriz (1987) Mol. Gen. Genet. 207:90; Wandelt (1989) Nucleic Acids Res., 17:2354; Langridge

(1983) Cell, 34:1015; Reina (1990) Nucleic Acids Res., 18:7449); promotor ADP-gpp (ver, *e. g.*, Patente U.S. Nº 7102057); globulina-1 (ver, *e. g.*, Belanger (1991) Genetics 129:863); α-globulina (Sunilkumar, *et al.* (2002), Transgenic Res. 11:347-359); α-tubulina; cab (ver, *e. g.*, Sullivan (1989) Mol. Gen. Genet., 215:431); PEPCase (ver, *e. g.*, Hudspeth & Grula, (1989) Plant Molec. Biol., 12:579-589); Promotores associados ao complexo do gene R (chandler *et al.*, (1989) Plant Cell, 1:1175); promotor da vicilina de ervilha (Czako *et al.*, (1992) Mol. Gen. Genet., 235:33; Pat. U.S. Nº 5625136); promotor GTL1 (Takaiwa *et al.* (1991) Plant Mol. Biol. 16 (1), 49-58); promotores da chalcone sintase (Franken *et al.*, (1991) EMBO J., 10:2605); promotor GY1 (Sims & Goldburg (1989) Nuc. Acids Res. 17(11) 4368) e semelhantes.

A invenção pode utilizar promotores preferidos de fruta, incluindo qualquer classe de promotores preferidos de fruta, *e. g.*, como expressos em ou durante antítese durante o desenvolvimento da fruta, pelo menos, no início do amadurecimento, *e. g.*, como discutido no documento U.S. 4943674. O promotor para o gene da poligalacturonase está activo no amadurecimento da fruta. A invenção pode utilizar o gene da poligalacturonase como descrito, *e. g.*, na patente U.S. Nº 4535060, Patente U.S. Nº 4769061, Patente U.S. Nº 4801590 e Patente U.S. Nº 5107065.

A invenção pode utilizar quaisquer promotores preferidos de tecido, incluindo aqueles que dirigem a expressão em células de folha, após lesão da folha (por exemplo, a partir de insectos mascadores), em tubérculos (por exemplo, promotor do gene da patatina) e em células de fibra (um exemplo de uma proteína de célula de fibra regulada pelo desenvolvimento é E6 (John & Crow (1992) PNAS 89:5769-5773). O gene E6 é o mais activo na fibra,

embora níveis baixos de transcriptos sejam encontrados na folha, óvulos e flor.

A invenção pode utilizar promotores activos no tecido fotossintético, *e. g.*, de modo conduzir a transcrição em tecidos verdes, tais como folhas e caules, os quais são adequados quando dirigem a expressão, apenas ou predominantemente, em tais tecidos. Alternativamente, a invenção pode utilizar promotores que conferem expressão constitutivamente ao longo da planta, ou diferencialmente relativamente aos tecidos verdes, ou diferencialmente relativamente à fase de desenvolvimento do tecido verde em que ocorre a expressão, ou em resposta a estímulos externos.

Promotores exemplificativos utilizados para praticar esta invenção incluem os promotores da ribulose-1,5-bisfosfato carboxilase (RbcS), tais como o promotor RbcS de larício oriental (*Larix laricina*), o promotor cab6 de pinheiro (Yamamoto *et al*. (1994) Plant Cell Physiol. 35:773-778), o promotor do gene Cab-1 de trigo (Fejes *et al*. (1990) Plant Mol. Biol. 15:921-932), o promotor Cab-1 de espinafre (Lubberstedt *et al*. (1994) Plant Physiol. 104:997-1006), o promotor cab1R de arroz (Luan *et al*. (1992) Plant Cell 4:971-981), o promotor da piruvato ortofosfato dicinase (PPDK) de milho (Matsuoka *et al*. (1993) Proc Acad natl Sci USA 90:9586-9590), promotor de Lhcb1*2 de tabaco (Cerdan *et al*. (1997) Plant Mol. Biol. 33:245-255), o promotor do simporte sacarose-H+ SUC2 de *Arabidopsis thaliana* (Truernit *et al*. (1995) Planta 196:564-570) e promotores de proteínas de membranas tilacóides de espinafre (psaD, psaF, psaE, PC, FNR, atpC, atpD, cab, rbcS). Outros promotores que dirigem a transcrição em caules, folhas e tecido verde são descritos na publicação da Patente U.S. Nº 2007/0006346.

Em algumas formas de realização, a especificidade de tecido de alguns promotores "preferidos de tecido" pode não ser absoluta e podem ser testados genes repórter, tais como Gus ou a proteína fluorescente verde, proteína fluorescente ciana, proteína fluorescente amarela ou proteína fluorescente vermelha. Pode-se também obter expressão preferida de tecido com expressão "não estanque" através de uma combinação de diferentes promotores preferidos de tecido. Outros promotores preferidos de tecido podem ser isolados por um especialista na técnica (ver documento U.S. 5589379).

Num aspecto, os promotores de planta que podem ser induzíveis após exposição a hormonas vegetais, tal como auxinas, podem ser utilizados para expressar os ácidos nucleicos da invenção. Por exemplo, a invenção pode utilizar o fragmento do promotor de elementos E1 de resposta a auxina (AuxRE) em soja (*Glycine max L.*) (Liu (1997) Plant Physiol. 115:397-407); o promotor GST6 responsivo a auxina de *Arabidopsis* (também responsivo ao ácido salicílico e ao peróxido de hidrogénio) (Chen (1996) Plant J. 10: 955-966); o promotor parC induzível por auxina de tabaco (Sakai (1996) 37:906-913); um elemento de resposta à biotina vegetal (Streit (1997) Mol. Plant Microbe Interact. 10:933-937); e o promotor responsivo à hormona de stress ácido abscísico (Sheen (1996) Science 274:1900-1902).

Os ácidos nucleicos da invenção podem também ser ligados operativamente a promotores de plantas que são induzíveis após exposição a reagentes químicos que podem ser aplicados à planta, tais como herbicidas ou antibióticos. Por exemplo, os sistemas de expressão génica que são activados na presença de um ligando químico, incluindo etanol, podem ser encontrados nos documentos WO 96/27673; WO 93/01294; WO 94/03619; WO 02/061102. O promotor In2-2 de milho, activado por fitoprotectores herbicidas de

benzenossulfonamida pode ser utilizado (De Veylder (1997) Plant Cell Physiol. 38:568-577); a aplicação de diferentes fitoprotectores herbicidas induz distintos perfis de expressão génica, incluindo expressão na raiz, hidatodes e meristema apical de rebento. A sequência codificante pode estar sob controlo, *e. g.*, de um promotor induzível por tetraciclina, *e. g.*, como descrito com plantas transgénicas de tabaco contendo o gene da arginina descarboxilase de *Avena sativa* L. (aveia) (Masgrau (1997) Plant J. 11:465-473); estrogénio, tais como, o receptor de ecdisona (documento WO 01/52620) ou um elemento responsivo a ácido salicílico (Stange (1997) Plant J. 11:1315-1324). Utilizando promotores induzidos quimicamente (*e. g.*, por hormonas ou pesticidas), *i. e.*, promotor responsivo a um químico que pode ser aplicado à planta transgénica no campo, a expressão de um polipéptido da invenção pode ser induzida numa fase particular do desenvolvimento da planta.

Promotores constitutivos exemplificativos que podem ser utilizados para praticar esta invenção e que foram descritos, incluem a actina 1 de arroz (Wang *et al.* (1992) Mol. Cell. Biol., 12:3399; Patente U.S. Nº 5641876); outras isoformas da actina (McElroy *et al.* (1990) Plant Cell 2: 163-171 e McElroy *et al.* (1991) Mol. Gen. Genet. 231: 150-160); 35S de CaMV (Odell *et al.* (1985) Nature, 313:810); 19S de CaMV (Lawton *et al.* (1987) Plant Mol. Biol. 9:315-324; Patente U.S. Nº 5639949); NOS (Ebert *et al.* (1987) PNAS USA 84:5745-5749); Adh (walker *et al.* (1987) PNAS USA 84:6624-6628), sacarose sintase (Yang & Russell (1990) PNAS USA 87:4144-4148); e os promotores da ubiquitina (*e. g.*, girassol - Binet *et al.* (1991) Plant Science 79: 87-94; milho - Christensen *et al.* (1989) Plant Molec. Biol. 12: 619-632; e Arabidopsis - Callis *et al.*, J. Biol. Chem. (1990)

265:12486-12493; e Norris *et al.*, Plant Mol. Biol. (1993) 21:895-906.

Qualquer terminador de transcrição pode ser utilizado para praticar esta invenção, *e. g.*, pode ser utilizado em vectores, cassetes de expressão e semelhantes. Estes são responsáveis pela terminação da transcrição para além do transgene e correcção da poliadenilação do ARNm. A região de terminação pode ser nativa com a região de iniciação de transcrição, pode ser nativa com a sequência de ADN ligada operativamente de interesse, pode ser nativa com a planta hospedeira ou pode ser derivada de outra fonte (*i. e.*, estranho ou heterólogo ao promotor, à sequência de ADN de interesse, à planta hospedeira ou a qualquer sua combinação). Os terminadores transcricionais apropriados são aqueles que são conhecidos por funcionarem em plantas e incluem o terminador 35S de CAMV, o terminador tml, o terminador da nopalina sintase e o terminador E9 de rbcs de ervilha. Estes podem ser utilizados em monocotiledóneas e dicotiledóneas. Além disso, pode ser utilizado o terminador de transcrição nativo de um gene.

A invenção pode utilizar qualquer sequência para aumentar a expressão génica dentro da unidade de transcrição; e estas sequências podem ser utilizadas em conjunto com os genes desta invenção para aumentar a sua expressão em plantas transgénicas. Por exemplo, foi demonstrado que várias sequências de intrão aumentam a expressão, particularmente em células monocotiledóneas. Por exemplo, verificou-se que os intrões do gene Adhl de milho aumentam significativamente a expressão do gene de tipo selvagem sob o seu promotor aparentado quando introduzidos em célula de milho. Uma variedade de sequências líder não traduzidos derivadas de vírus são também conhecidas por aumentarem a expressão e estas são particularmente eficazes

em células dicotiledóneas. Especificamente, foi demonstrado que as sequências líder do Vírus do Mosaico do Tabaco (TMV, a "sequência W"), Vírus de Mancha Clorótica de milho (MCMV) e Vírus do Mosaico da Alfalfa (AMV) são eficazes no aumento da expressão (*e. g.*, Gallie *et al*. Nucl. Acids Res. 15: 8693-8711 (1987); Skuzeski *et al*. Plant Molec. Biol. 15: 65-79 (1990)).

Direccionamento do Produto Génico Dentro da Célula

Qualquer mecanismo para direccionamento de produtos génicos, *e. g.*, em plantas, pode ser utilizado para praticar esta invenção e tais mecanismos são conhecidos por existirem em plantas e as sequências que controlam o funcionamento destes mecanismos foram caracterizados com algum detalhe. Foram caracterizadas sequências que provocam o direccionamento de produtos génicos a outros compartimentos celulares. As sequências do terminal amino podem ser responsáveis por direccionamento de uma proteína de interesse a qualquer compartimento celular, tais como, um vacúolo, mitocôndria, peroxissoma, corpo proteico, reticulo endoplasmático, cloroplasto, grânulo de amido, amiloplasto, apoplasto ou parede celular de uma planta (*e. g.*, Unger *et al*. Plant Molec. Biol. 13: 411-418 (1989); Rogers *et al*. (1985) Proc. Natl. Acad. Sci. USA 82: 6512-651; Patente U.S. N° 7102057; documento WO 2005/096704). Opcionalmente, a sequência de sinal pode ser uma sequência de sinal do N-terminal de waxy, uma sequência de sinal do N-terminal de γ-zeína, um domínio de ligação a amido, um domínio de ligação a amido C-terminal, uma sequência de direccionamento ao cloroplasto, a qual importa a proteína madura para o cloroplasto (Comai *et al*. (1988) J. Biol. Chem. 263: 15104-15109; van den Broeck, *et al*. (1985) Nature 313: 358-363; Patente U.S. N° 5639949) ou uma sequência de sinal de secreção

de células de aleurona (Koehler & Ho, Plant Cell 2: 769-783 (1990)). Adicionalmente, as sequências amino terminais em conjunto com sequência carboxilo terminais são responsáveis pelo direccionamento vacuolar de produtos génicos (Shinshi *et al*. (1990) Plant Molec. Biol. 14: 357-368).

Num aspecto, a sequência de sinal seleccionada deve incluir o local de clivagem conhecido e a fusão construída deve tomar em consideração quaisquer aminoácidos após o local ou locais de clivagem, as quais são requeridos para clivagem. Em alguns casos este requisito pode ser satisfeito através da adição de um pequeno número de aminoácidos entre o local de clivagem e o ATG do transgene ou, alternativamente, substituição de alguns aminoácidos dentro da sequência do transgene. Estas técnicas de construção são bem conhecidas na técnica e são igualmente aplicáveis a qualquer compartimento celular.

Num aspecto, os mecanismos descritos acima para direccionamento celular podem ser utilizados não apenas em conjunto com os seus promotores aparentados, mas também em conjunto com promotores heterólogos para atingir um objectivo de direccionamento celular específico sob regulação de transcrição de um promotor que tenha um perfil de expressão diferente daquele do promotor do qual deriva o sinal de direccionamento.

Em suma, uma variedade de meios podem ser utilizados para realizar esta invenção, incluindo quaisquer meios para realizar a expressão recombinante de fitase numa planta transgénica, semente, órgão ou qualquer parte de planta. Tais plantas e partes de planta transgénicas são úteis como fontes de fitase recombinantemente expressa, as quais podem ser adicionadas directamente a fontes contendo fitato. Alternativamente, a

fitase recombinante expressa em planta pode ser extraída para fora da fonte vegetal e, se desejado, purificada antes de ser colocada em contacto com o substrato de fitase.

No contexto da presente invenção, as plantas que podem ser seleccionadas (utilizadas para realizar esta invenção) incluem, mas não estão limitadas a, cultivares que produzem flores comestíveis, tais como couve-flor (*Brassica oleracea*), alcachofra (*Cynara scolymus*), frutas, tais como maçã (*Malus, e. g., domesticus*), banana (*Musa, e. g., acuminata*), frutos de baga (tal como groselha, *Ribes, e. g., rubrum*), cerejas (tal como cereja doce, *Prunus, e. g., avium*), pepino (*Cucumis, e. g., sativus*), uva (*Vitis, e. g., vinifera*), *limão* (*Citrus limon*), melão (*Cucumis melo*), frutos secos (tais como noz, *Juglans, e. g., regia;* amendoim, *Arachis hypogeae*), laranja (*Citrus, e. g., maxima*), pêssego (*Prunus, e. g., persica*), pera (*Pyra, e. g., communis*), ameixa (*Prunus, e. g., domestica*), morango (*Fragaria, e. g., moschata*), tomate (*Lycopersicon, e. g., esculentum*), folhas, tais como alfalfa (*Medicago, e. g., sativa*), couves (*Brassica, e. g., oleracea*), endívias (*Cichoreum, e. g., endivia*), alho francês (*Allium, e. g., porrum*), alface (*Lactuca, e. g., sativa*), espinafre (*Spinacia, e. g., oleraceae*), tabaco (Nicotiana, *e. g., tabacum*), raízes, tais como araruta (*Maranta, e. g., arundinacea*), beterraba (*Beta, e. g., vulgaris*), cenoura (*Daucus, e. g., carota*), mandioca (*Manihot, e. g., esculenta*), nabo (*Brassica, e. g., rapa*), rabanete (*Raphanus, e. g., sativus*), inhame (*Dioscorea, e. g., esculenta*), batata-doce (*Ipomoea batatas*) e sementes, tais como feijão (Phaseolus, *e. g., vulgaris*), ervilha (*Pisum, e. g., sativum*), soja (*Glycin, e. g.,* max), trigo (*Triticum, e. g., aestivum*), cevada (*Hordeum, e. g., vulgare*), milho (*Zea, e. g., mays*), arroz (*Oryza, e. g., sativa*), colza (*Brassica*

Napus), milho painço (*Panicum L.*), girassol (*Helianthus annus*), aveia (*Avena sativa*), tubérculos, tais como couve-rábano (*Brassica, e. g., oleraceae*), batata (*Solanum, e. g., tuberosum*) e semelhantes.

Num aspecto, os ácidos nucleicos e polipéptidos da invenção são expressos ou introduzidos em qualquer planta ou semente. As plantas transgénicas da invenção podem ser dicotiledóneas ou monocotiledóneas. Os exemplos de plantas transgénicas monocot da invenção são ervas, tais como erva do prado (erva azul, *Poa*), erva de forragem, tais como festuca, azevém, erva temperadas, tais como *Agrostis,* e cereais, *e. g.*, trigo, aveia, centeio, cevada, arroz, sorgo e milho. Os exemplos de plantas transgénicas dicot da invenção são o tabaco, leguminosas, tais como tremoceiro, batata, beterraba açucareira, ervilha, feijão e soja, e plantas crucíferas (família *Brassicaceae*), tais como couve-flor, colza e o organismo modelo intimamente relacionado *Arabidopsis thaliana*. Assim, as plantas e sementes transgénicas da invenção incluem uma vasta gama de plantas, incluindo, mas não limitadas a, espécies dos géneros *Anacardium, Arachis, Asparagus, Atropa, Avena, Brassica, Citrus, Citrullus, Capsicum, Carthamus, Cocos, Coffea, Cucumis, Cucurbita, Daucus, Elaeis, Fragaria, Glycine, Gossypium, Helianthus, Heterocallis, Hordeum, Hyoscyamus, Lactuca, Linum, Lolium, Lupinus, Lycopersicon, Malus, Manihot, Majorana, Medicago, Nicotiana, Olea, Oryza, Panieum, Pannisetum, Persea, Phaseolus, Pistachia, Pisum, Pyrus, Prunus, Raphanus, Ricinus, Secale, Senecio, Sinapis, Solanum, Sorghum,Theobromus, Trigonella, Triticum, Vicia, Vitis, Vigna* e *Zea*.

Em formas de realização alternativas, os ácidos nucleicos da invenção são expressos em plantas que contêm células de fibra,

incluindo, *e. g.*, algodão, mafumeira (Kapok, *Ceiba pentandra*), salgueiro do deserto, arbusto creosoto, *winterfat*, balsa, rami, quenafe, cânhamo, cânhamo Rosella, juta, sisal, abacá e linho. Em formas de realização alternativas, as plantas transgénicas da invenção podem ser membros do género *Gossypium,* incluindo membros de quaisquer espécies de *Gossypium*, tais como *G. arboreum; G. Herbaceum, G. barbadense* e *G. hirsutum.*

Plantas adicionais, assim como sistemas de expressão não de plantas, podem ser utilizados para realizar esta invenção. A escolha da espécie de planta é principalmente determinada pela utilização pretendida da planta ou das suas partes e da facilidade de transformação da espécie vegetal.

Diversas técnicas estão disponíveis para introdução da construção de expressão contendo a sequência de ADN codificando a fitase nas plantas alvo. As técnicas para transformar uma vasta variedade de espécies de plantas superiores são bem conhecidas e descritas na literatura técnica e científica. Ver, *e. g.*, Weising (1988) Ann. Rev. Genet. 22:421-477; Patente U.S. Nº 5750870. Tais técnicas podem também incluir, mas não estão limitadas, à transformação de protoplastos utilizando o método de cálcio/polietilenoglicol, electroporação e microinjecção ou bombardeamento de partículas (revestidas) (Potrykus, 1990). Para além destes denominados métodos directos de transformação de ADN, estão extensamente disponíveis sistemas de transformação que envolvem vectores, tais como vectores virais (*e. g.*, do Vírus do Mosaico de Couve-flor (CaMV) e vectores bacterianos (*e. g.*, do género *Agrobacterium*) (Potrykus, 1990). Após selecção e/ou rastreio, os protoplastos, células ou partes de plantas que foram transformados são regenerados em plantas inteiras, utilizando métodos conhecidos na técnica (Horsch *et al.*, 1985).

A escolha das técnicas de transformação e/ou regeneração não é crítica para esta invenção.

Os ácidos nucleicos e as construções de expressão da invenção podem ser introduzidos numa célula vegetal por quaisquer meios. Em aspectos alternativos para realizar esta invenção, o termo “introduzir” no contexto de um polinucleótido, por exemplo, uma construção nucleotídica de interesse, pretende designar a apresentação à planta do polinucleótido de tal modo que o polinucleótido ganha acesso ao interior de uma célula vegetal. Nos casos em que é para ser introduzido mais de um polinucleótido, estes polinucleótidos podem ser montados como parte de uma única construção nucleotídica, ou como construções nucleotídicas separadas, e podem estar localizadas no mesmo ou em diferentes vectores de transformação. Desta forma, estes polinucleótidos podem ser introduzidos na célula hospedeira de interesse num único evento de transformação, em eventos separados de transformação ou, por exemplo, em plantas, como parte de um protocolo de cruzamento. Os métodos da invenção não dependem de um método particular para introduzir um ou mais polinucleótidos numa planta, apenas que o ou os polinucleótidos ganham acesso ao interior de, pelo menos, uma célula vegetal. Os métodos para introduzir polinucleótidos em plantas são conhecidos na técnica incluindo, mas não limitados a métodos de transformação transiente, métodos de transformação estável e métodos mediados por vírus.

A “transformação transiente” pode ser utilizada para praticar esta invenção e em alguns aspectos no contexto de um polinucleótido pretende designar que um polinucleótido é introduzido na planta e não se integra no genoma da planta. “Introduzir estavelmente” ou “introduzido estavelmente” no contexto de um polinucleótido introduzido numa planta pode ser

utilizado para realizar esta invenção e, em alguns aspectos, pretende designar que o polinucleótido introduzido é incorporado estavelmente no genoma da planta e assim a planta é transformada estavelmente com o polinucleótido.

"Transformação estável" ou "transformado estavelmente" no contexto de um polinucleótido introduzido numa planta pode ser utilizado para realizar esta invenção e, em alguns aspectos, pretende designar que um polinucleótido, por exemplo, uma construção nucleotídica aqui descrita, é introduzido numa planta, integrada no genoma da planta e é capaz de ser herdada através da sua descendência, mais particularmente, através da descendência de múltiplas gerações sucessivas. A introdução no genoma de uma planta desejada pode ser tal que a enzima é regulada por elementos de controlo transcricionais ou de tradução endógenos. As técnicas de transformação para ambas as monocotiledóneas e as dicotiledóneas são bem conhecidas no campo técnico.

Os ácidos nucleicos da invenção podem ser utilizados para conferir características desejadas essencialmente em qualquer planta. Numa forma de realização, a enzima da invenção pode ser expressa de tal modo que a enzima não entrará em contacto com o seu substrato até desejado. Por exemplo, uma enzima da invenção pode ser direccionada e retida no retículo endoplasmático de uma célula vegetal. A retenção da enzima, no retículo endoplasmático da célula, irá prevenir a enzima de entrar em contacto com o seu substrato. A enzima e o substrato podem depois ser colocados em contacto através de quaisquer meios capazes de quebrar a arquitectura subcelular, tais como, trituração, moagem, aquecimento e semelhantes. Ver, documentos WO 98/11235, WO 2003/18766 e WO 2005/096704.

Genes marcadores de selecção podem ser adicionados à construção génica para identificar células ou tecidos vegetais que integraram com sucesso o transgene. Isto pode ser necessário porque conseguir a incorporação e expressão de genes em células vegetais é um evento raro, ocorrendo apenas numa pequena percentagem de tecidos ou células alvejadas. Os genes marcadores de selecção codificam proteínas que proporcionam resistência a agentes que são normalmente tóxicos para as plantas, tais como antibióticos ou herbicidas. Apenas as células vegetais que integraram o gene marcador de selecção sobreviverão quando cultivadas num meio contendo o antibiótico ou herbicida apropriado. Os marcadores de selecção utilizados rotineiramente em transformação incluem o gene nptll que confere resistência à canamicina e antibióticos relacionados (Messing & Vierra. Gene 19: 259-268 (1982); Bevan *et al.*, Nature 304:184-187 (1983)), o gene bar que confere resistência ao herbicida fosfinotricina (White *et al.*, Nucl. Acids Res 18: 1062 (1990), Spencer *et al.* Theor. Appl. Genet 79: 625-631 (1990)), o gene hph que confere resistência ao antibiótico higromicina (Blochinger & Diggelmann, Mol Cell Biol 4: 2929-2931), o gene dhfr que confere resistência ao metotrexato (Bourouis *et al.*, EMBO J. 2(7): 1099-1104 (1983)), o gene EPSPS que confere resistência ao glifosato (Pat. U.S. N° 4940935 e 5188642),

Alternativamente, material vegetal transgénico pode ser identificado através de um sistema de selecção positiva, tal como o sistema que utiliza o gene da manose-6-fosfato isomerase que proporciona a capacidade para metabolizar a manose (Pat. U.S. N° 5767378 e 5994629).

Num aspecto, fazer plantas ou sementes transgénicas compreende a incorporação de sequências da invenção e, opcionalmente, de genes marcadores numa construção de expressão

alvo (*e. g.*, um plasmídeo), em conjunto com posicionamento do promotor e das sequências terminadoras. Isto pode envolver transferir o gene modificado para a planta com um método adequado. Uma ou mais das sequências da invenção podem ser combinadas com sequências que conferem resistência a insectos, doenças, secas, rendimento aumentado, melhoram a qualidade nutritiva do grão, melhoram o rendimento de etanol e semelhantes.

Por exemplo, uma construção pode ser introduzida directamente no ADN genómico da célula vegetal utilizando técnicas, tais como electroporação e microinjecção de protoplastos de células vegetais, ou as construções podem ser introduzidas directamente no tecido vegetal utilizando métodos balísticos, tal como bombardeamento de partículas de ADN. Por exemplo, ver, *e. g.*, Christou (1997) Plant Mol. Biol,35:197-203; Pawlowski (1996) Mol. Biotechnol. 6:17-30; Klein (1987) Nature 327:70-73; Takumi (1997) Genes Genet. Syst. 72:63-69, discutem a utilização de bombardeamento de partículas para introduzir transgenes no trigo; e Adam (1997) supra, para utilização de bombardeamento de partículas para introduzir YAC em células vegetais. Por exemplo, Rinehart (1997) supra, utilizou bombardeamento de partículas para produzir plantas transgénicas de algodão. O dispositivo para acelerar partículas é descrito na Pat. U.S. Nº 5015580; e o instrumento comercialmente disponível de aceleração de partículas PDS-2000 da BioRad (Biolística); ver também, John, Patente U.S. Nº 5608148; e Ellis, Patente U.S. Nº 5681730, as quais descrevem a transformação mediada por partículas de gimnospérmicas.

Num aspecto, os protoplastos podem ser imobilizados e injectados com ácidos nucleicos, *e. g.*, uma construção de expressão. Embora a regeneração de plantas a partir de

protoplastos não seja fácil com cereais, a regeneração de plantas é possível em leguminosas utilizando embriogénese somática a partir de protoplastos derivados de *callus*. Os tecidos organizados podem ser transformados com ADN nu utilizando a técnica da pistola de genes, em que o ADN é revestido em microprojécteis de tungsténio, disparos com 1/100 do tamanho das células, que transportam o ADN profundamente nas células e organelos. O tecido transformado é depois induzido para regenerar, habitualmente por embriogénese somática. Esta técnica tem sido bem sucedida em diversas espécies de cereais incluindo milho e arroz.

Os ácidos nucleicos, *e. g.*, construções de expressão, podem também ser introduzidos dentro das células vegetais utilizando vírus recombinantes. As células vegetais podem ser transformadas utilizando vectores virais, tal como, *e. g.*, vectores derivados do vírus do mosaico do tabaco (Rouwendal (1997) Plant Mol. Biol. 33:989-999), ver Porta (1996) "Use of viral replicons for the expression of genes in plants", Mol. Biotechnol. 5:209-221.

Alternativamente, os ácidos nucleicos, *e. g.*, uma construção de expressão, podem ser combinados com regiões flanqueadoras de T-ADN adequadas e introduzidos num vector hospedeiro convencional de *Agrobacterium tumefaciens*. As funções virulentas do hospedeiro *Agrobacterium tumefaciens* irão direccionar a inserção da construção e do marcador adjacente no ADN da célula vegetal quando a célula for infectada com as bactérias. As técnicas de transformação mediadas por *Agrobacterium tumefaciens*, incluindo desarme e utilização de vectores binários, estão bem descritas na literatura científica. Ver, *e. g.*, Horsch (1984) Science 233:496-498; Fraley (1983) Proc. Natl. Acad. Sci. USA 80:4803 (1983); Gene Transfer to Plants, Potrykus, ed. (Springerlag, Berlin 1995). O ADN numa célula de

A. tumefaciens está contido no cromossoma bacteriano, assim como numa outra estrutura conhecida como um plasmídeo Ti (indutor de tumor). O plasmídeo Ti contém um trecho de ADN denominado T-ADN (~20 kb de comprimento) que é transferido à célula vegetal no processo de infecção e uma série de genes vir (virulência) que dirigem o processo de infecção. *A. tumefaciens* pode apenas infectar uma planta através de feridas: quando uma raiz ou um caule de planta é ferido emite determinados sinais químicos, em resposta aos quais os genes vir *de A. tumefaciens* tornam-se activos e dirigem uma série de eventos necessários para transferência do T-ADN do plasmídeo Ti para o cromossoma da planta. O T-ADN entra depois na célula vegetal através da ferida. Uma possibilidade é que o T-ADN espera até que o ADN da planta esteja a ser replicado ou transcrito, depois introduz-se no ADN exposto da planta. De modo a utilizar *A. tumefaciens* como um vector de transgene, a secção indutora de tumor de T-ADN tem que ser removida, enquanto retém as regiões marginais de T-ADN e os genes vir. O transgene é depois introduzido entre as regiões marginais de T-ADN, onde é transferido para a célula vegetal e torna-se integrado nos cromossomas da planta.

A invenção proporciona a transformação de plantas monocotiledóneas utilizando os ácidos nucleicos da invenção, incluindo cereais importantes, ver Hiei (1997) Plant Mol. Biol. 35:205-218. Ver também, *e. g.*, Horsch, Science (1984) 233:496; Fraley (1983) Proc. Natl. Acad. Sci USA 80:4803; Thykjaer (1997) supra; Park (1996) Plant Mol. Biol. 32:1135-1148, discutem a integração de T-ADN no ADN genómico. Ver também D'Halluin, Patente U.S. Nº 5712135, a qual descreve um processo para a integração estável de um ADN compreendendo um gene que é funcional numa célula de um cereal ou de outra planta monocotiledónea.

Num aspecto, o terceiro passo pode envolver selecção e regeneração de plantas inteiras capazes de transmitir o gene alvo incorporado à geração seguinte. Tais técnicas de regeneração confiam na manipulação de determinadas fito-hormonas num meio de crescimento de cultura de tecidos, confiando tipicamente num marcador biocida e/ou herbicida que foi introduzido em conjunto com as sequências nucleotídicas desejadas. A regeneração de plantas a partir de protoplastos cultivados é descrita em Evans *et al.*, Protoplasts Isolation and Culture, Handbook of Plant Cell Culture, pp. 124-176, MacMillilan Publishing Company, Nova Iorque, 1983; e Binding, Regeneration of Plants, Plant Protoplasts, pp. 21-73, CRC Press, Boca Raton, 1985. A regeneração pode também ser obtida a partir de *callus*, explantes, órgãos de plantas ou suas partes. Tais técnicas de regeneração são descritas genericamente em Klee (1987) Ann. Rev. of Plant Phys. 38:467-486. Para obter plantas inteiras a partir de tecidos transgénicos, tal como embriões imaturos, estes podem ser cultivados sob condições ambientais controladas numa série de meios contendo nutrientes e hormonas, um processo conhecido como cultura de tecidos. Assim que as plantas inteiras são geradas e produzem semente, a avaliação da descendente começa.

Num aspecto, depois da cassete de expressão estar incorporada estavelmente em plantas transgénicas, esta pode ser introduzida noutras plantas por cruzamento sexual. Qualquer de uma variedade de técnicas de cruzamento convencionais pode ser utilizada, dependendo das espécies a serem cruzadas. Ver, por exemplo, Welsh J. R., Fundamentals of Plant Genetics and Breeding, John Wiley & Sons, NI (1981); Crop Breeding, Wood D. R. (Ed.) American Society of Agronomy Madison, Wis. (1983); Mayo O., The Theory of Plant Breeding, Segunda Edição, Clarendon Press, Oxford (1987); Singh, D. P., Breeding for Resistance to

Diseases and Insect Pests, Springer-Verlag, NI (1986); e Wricke e Weber, Quantitative Genetics and Selection Plant Breeding, Walter de Gruyter and Co., Berlim (1986).

Num aspecto, uma vez que a expressão transgénica dos ácidos nucleicos da invenção conduz a alterações fenotípicas, as plantas compreendendo os ácidos nucleicos recombinantes da invenção podem ser cruzados sexualmente com uma segunda planta para obter um produto final. Assim, a semente da invenção pode ser derivada de um cruzamento entre duas plantas transgénicas da invenção, ou um cruzamento entre uma planta da invenção e uma outra planta. Os efeitos desejados (*e. g.*, expressão dos polipéptidos da invenção para produzir uma planta em que o comportamento de florescência está alterado) podem ser aumentados quando ambas as plantas parentais expressam os polipéptidos (*e. g.*, fitase) da invenção. Os efeitos desejados podem ser passados a gerações futuras de plantas por meios de propagação convencionais.

Para dicots, um sistema de vector binário pode ser utilizado (Hoekema *et al.*, 1983; Ep 0120516 Schilperoort *et al.*). Por exemplo, estirpes de *Agrobacterium* podem ser utilizadas que contêm um plasmídeo *vir* com os genes de virulência e um plasmídeo compatível contendo a construção génica a ser transferida. Este vector pode replicar em *E. coli* e em *Agrobacterium* e é derivado do vector binário Bin19 (Bevan, 1984) que está alterado em detalhes que não são relevantes para esta invenção. Os vectores binários como utilizados neste exemplo contêm entre as sequências marginais esquerda e direito do T-ADN, um gene NPTII idêntico que codifica a resistência à canamicina (Bevan, 1984) e um local de clonagem múltipla para clonar nas construções génicas requeridas.

A transformação e regeneração de cultivares monocotiledóneos podem ser praticados utilizar qualquer método ou processo convencional; as monocots são susceptíveis à transformação e plantas transgénicas férteis podem ser regeneradas a partir de células transformadas. Num aspecto alternativo, as monocots são transformadas por transformação com *Agrobacterium*.

Num aspecto, plantas transgénicas de arroz podem ser obtidas utilizando o gene hph bacteriano, a qual codifica a resistência à higromicina, como um marcador de selecção; o gene pode ser introduzido por electroporação. Num aspecto, plantas transgénicas de milho podem ser obtidas por introdução do gene bar de *Streptomyces hygroscopicus*, a qual codifica a fosfinotricina acetiltransferase (uma enzima que inactiva o herbicida fosfinotricina), em células embriogénicas de uma cultura em suspensão de milho por bombardeamento de micropartículas. Num aspecto, material genético pode ser introduzido em protoplastos de aleurona de cultivares monocot, tais como trigo e cevada. Num aspecto, as plantas de trigo são regeneradas a partir de cultura em suspensão embriogénica através de selecção apenas de tecidos de *callus* embriogénicos compactos e nodulares envelhecidos para o estabelecimento das culturas em suspensão embriogénicas. Num aspecto, a combinação com sistemas de transformação para estas cultivares permite a aplicação da presente invenção a monocots. Estes métodos e outros métodos podem também ser aplicados para a transformação e regeneração de dicots.

Na prática desta invenção, a expressão da construção de fitase envolve tais detalhes como transcrição do gene por polimerases vegetais, tradução de ARNm, etc. que são conhecidos dos especialistas no campo técnico de técnicas de ADN recombinante. Alguns detalhes relevantes para realizar algumas

formas de realização desta invenção são aqui discutidos. Sequências reguladoras que são conhecidas ou que se verifica que provocam a expressão de fitase, podem ser utilizadas na presente invenção. A escolha das sequências reguladoras utilizadas pode depender do cultivar alvo e/ou do órgão alvo de interesse. Tais sequências reguladoras podem ser obtidas a partir de plantas ou vírus de plantas, ou podem ser sintetizadas quimicamente. Tais sequências reguladoras são promotores activos em dirigir a transcrição em plantas, constitutivamente ou específica de fase e/ou tecido, dependendo da utilização da planta ou das suas partes. Estes promotores incluem, mas não estão limitados a, promotores que apresentam expressão constitutiva, tal como o promotor 35S do vírus do mosaico da couve-flor (CaMV) (Guilley *et al.*, 1982), aqueles para expressão específica de folha, tal como o promotor do gene da subunidade pequena da ribulose bisfosfato carboxilase (Coruzzi *et al.*, 1984), aqueles para expressão específica de raiz, tal como o promotor do gene da glutamina sintase (Tingey *et al.*, 1987), aqueles para a expressão específica de semente, tal como o promotor da cruciferina A de *Brassica napus* (Ryan *et al.*, 1989), aqueles para a expressão específica de tubérculos, tal como o promotor da patatina de classe-I de batata (Koster-Topfer *et al.*, 1989; Wenzler *et al.*, 1989) ou aqueles para a expressão específica de fruta, tal como o promotor da poligalacturonase (PG) de tomate (Bird *et al.*, 1988).

Outras sequências reguladoras, tais como sequências terminadoras e sinais de poliadenilação, podem ser utilizadas para realizar esta invenção e podem incluir uma tal sequência que funciona como tal em plantas, cuja escolha está dentro do nível do especialista na técnica. Um exemplo de tais sequências é a região 3' flanqueadora do gene da nopalina sintase (NOS) de *Agrobacterium tumefaciens* (Bevan, *supra*). As sequências

reguladoras podem também incluir sequências intensificadoras, tais como encontradas no promotor 35S de CaMV e sequências estabilizadoras de ARNm, tal como a sequência líder do Vírus do mosaico de alfalfa (AlMV) ARN4 (Brederode *et al.*, 1980) ou quaisquer outras sequências que funcionam de um modo semelhante.

Em algumas formas de realização, uma fitase da invenção é expressa num ambiente que permita a estabilidade da proteína expressa. A escolha de compartimentos celulares, tais como citosol, retículo endoplasmático, vacúolo, corpos proteicos ou espaço periplasmático, pode ser utilizada na presente invenção para criar um tal ambiente estável, dependendo dos parâmetros biofísicos da fitase. Tais parâmetros incluem, mas não estão limitados a pH óptimo, sensibilidade a proteases ou sensibilidade à molaridade do compartimento preferido.

Em algumas formas de realização, uma fitase da invenção é expressa no citoplasma; em algum aspecto, para obter expressão no citoplasma da célula, a enzima expressa não deve conter um péptido de sinal secretor ou qualquer outra sequência alvo. Para expressão em cloroplastos e mitocôndrias, a enzima expressa deve conter um péptido específico denominado péptido de trânsito para a importação nestes organelos. São conhecidas sequências de direccionamento que podem ser ligadas à enzima de interesse de modo a conseguir isto (Smeekens *et al.*, 1990; van den Broeck *et al.*, 1985; Wolter *et al.*, 1988). Se a actividade da enzima for desejada nos vacúolos, um péptido de sinal secretor tem que estar presente, assim como uma sequência de direccionamento específica que dirija a enzima para estes vacúolos (Tague *et al.*, 1990). O mesmo é verdadeiro para os corpos proteicos em sementes. A sequência de ADN codificando a enzima de interesse deve ser modificada de tal modo que a enzima possa exercer a sua acção no local desejado na célula.

Em algumas formas de realização, para conseguir expressão extracelular da fitase, a construção de expressão da presente invenção utiliza uma sequência de sinal secretora. Embora sequências de sinal que sejam homólogas (nativas) à espécie da planta hospedeira possam ser preferidas, sequências de sinal heterólogas, *i. e.*, aquelas originárias de outras espécies de plantas ou de origem microbiana, podem também ser utilizadas. Tais sequências de sinal são conhecidas dos especialistas na técnica. As sequências de sinal apropriadas que podem ser utilizadas dentro do contexto da presente invenção são divulgadas em Blobel *et al.*, 1979; Von Heijne, 1986; Garcia *et al.*, 1987; Sijmons *et al.*, 1990; Ng *et al.*, 1994; e Powers *et al.*, 1996).

Em algumas formas de realização, todas as partes das construções de ADN relevantes (promotores, sequências reguladoras, secretoras, estabilizadoras, de direccionamento ou de terminação) da presente invenção são modificadas, se desejadas, para afectar as suas características de controlo utilizando os métodos conhecidos dos especialistas na técnica. As plantas contendo fitase obtidas através da presente invenção podem ser utilizadas para obter plantas ou órgãos de planta com ainda níveis maia elevados de fitase. Por exemplo, pode ser possível obter tais plantas ou órgãos de planta através da utilização de técnicas de variação somoclonal ou técnicas de reprodução cruzada. Tais técnicas são bem conhecidas dos especialistas no campo técnico.

Num aspecto, a presente invenção proporciona um método (e seus produtos) para obter um sistema de sobreexpressão altamente eficiente para a fitase e outras moléculas. Num aspecto, a invenção proporciona um método (e os seus produtos) para obter

um sistema de sobreexpressão altamente eficiente para a fitase e fosfatase ácida pH 2,5 em *Trichoderma*. Este sistema resulta em composições enzimáticas tendo utilidade particular na indústria de rações animais. Detalhes adicionais relativamente a esta abordagem estão na literatura publicada e/ou são conhecidos do especialista na técnica. Num exemplo particular não limitante, tal literatura publicamente disponível inclui o documento EP 0659215 (WO 9403612 A1) (Nevalainen *et al.*), embora estas referências não ensinem as moléculas inventivas do presente pedido.

Em algumas formas de realização, a invenção utiliza rearranjo *in vivo*, que pode estar focado em processos "intermoleculares" colectivamente referidos como "recombinação", que em bactéria pode ser um fenómeno "dependente de RecA". A invenção pode-se basear em processos de recombinação de uma célula hospedeira para recombinar e rearranjar sequências, ou na capacidade das células para mediarem processos redutores para diminuírem a complexidade de sequências quasi-repetidas numa célula por deleção. Este processo de "rearranjo redutor" ocorre através de um processo "intramolecular" independente de RecA.

Deste modo, num outro aspecto da invenção, polinucleótidos variantes podem ser produzidos pelo processo de rearranjo redutor. O método envolve a geração de construções contendo sequências consecutivas (sequências codificadas no original), a sua inserção num vector apropriado e a sua subsequente introdução numa célula hospedeira apropriada. O rearranjo das identidades moleculares individuais ocorre através de processos combinatórios entre as sequências consecutivas na construção que possuem regiões de homologia ou entre unidades quasi-repetidas. O processo de rearranjo recombina e/ou reduz a complexidade e extensão das sequências repetidas e resulta na produção de novas

espécies moleculares. Vários tratamentos podem ser aplicados para aumentar a taxa de rearranjo. Estes podem incluir tratamento com luz ultravioleta ou químicos que danificam o ADN, e/ou a utilização de linhas celulares hospedeiras que apresentam níveis aumentados de "instabilidade genética". Assim, o processo de rearranjo pode envolver recombinação homóloga ou a propriedade natural de sequências quasi-repetidas para dirigirem a sua própria evolução.

A invenção pode utilizar sequências repetidas ou "quasi-repetidas"; estas sequências podem desempenhar um papel na instabilidade genética. Na presente invenção, "quasi-repetidas" são repetições que não estão restritas à sua estrutura unitária original. As unidades quasi-repetidas podem ser apresentadas como uma matriz de sequências numa construção; unidades consecutivas de sequências semelhantes. Uma vez ligadas, as junções entre as sequências consecutivas tornam-se essencialmente invisíveis e a natureza quasi-repetitiva da construção resultante é agora contínua ao nível molecular. O processo de deleção que a célula realiza para reduzir a complexidade da construção resultante opera entre as sequências quasi-repetidas. As unidades quasi-repetidas proporcionam um repertório praticamente ilimitado de moldes sobre os quais podem ocorrer eventos de deslizamento. As construções contendo as quasi-repetições proporcionam assim efectivamente elasticidade molecular suficiente que eventos de deleção (e potencialmente inserção) podem ocorrer virtualmente em qualquer lugar dentro das unidades quasi-repetitivas.

Em alguns aspectos, quando as sequências quasi-repetidas estão todas ligadas na mesma orientação, por exemplo, cabeça com cauda ou vice-versa, a célula não consegue distinguir unidades individuais. Consequentemente, o processo redutor pode ocorrer

ao longo das sequências. Pelo contrário, quando por exemplo, as unidades são apresentadas cabeça com cabeça, em vez de cabeça com cauda, a inversão delineia os pontos finais da unidade adjacente de modo que a formação de deleção irá favorecer a perda de unidades discretas. Assim, num aspecto da invenção, as sequências estão na mesma orientação. A orientação aleatória de sequências quasi-repetidas resultará na perda de eficiência de rearranjo, enquanto a orientação consistente das sequências irá oferecer a eficiência mais elevada. Contudo, embora ter menos das sequências contíguas na mesma orientação diminua a eficiência, pode ainda proporcionar elasticidade suficiente para a recuperação eficaz de novas moléculas. As construções podem ser feitas com as sequências quasi-repetidas na mesma orientação para permitir uma eficiência mais elevada.

As sequências podem ser montadas numa orientação cabeça com cauda utilizando qualquer de uma variedade de métodos, incluindo os seguintes:

(a) Podem ser utilizados iniciadores que incluem uma cabeça poli-A e uma cauda poli-T, as quais quando tornados de cadeia simples, proporcionam orientação. Isto é realizado tendo as primeiras bases dos iniciadores feitos de ARN e desta forma facilmente removidos com RNAse H.

(b) Podem ser utilizados iniciadores que incluem locais únicos de clivagem de restrição. Seriam requeridos múltiplos locais, uma bateria de sequências únicas e passos repetidos de síntese e ligação.

(c) As poucas bases internas do iniciador podem ser tioladas e uma exonuclease utilizada para produzir moléculas com cauda correctamente colocada.

Em alguns aspectos, a recuperação das sequências rearranjadas baseia-se na identificação de vectores de clonagem com um RI reduzido. As sequências codificantes rearranjadas podem ser depois recuperadas por amplificação. Os produtos são re-clonados e expressos. A recuperação de vectores de clonagem com RI reduzido pode ser efectuado por:

1) A utilização vectores apenas mantidos estavelmente quando a construção é reduzida em complexidade;
2) A recuperação física de vectores encurtados através de processos físicos. Neste caso, o vector de clonagem é recuperado utilizando processos de isolamento de plasmídeos convencionais e fraccionados pelo tamanho num gel de agarose ou numa coluna com um baixo limite de peso molecular utilizando processos convencionais;
3) A recuperação de vectores contendo genes interrompidos os quais podem ser seleccionados quando o tamanho da inserção diminui; e
4) A utilização de técnicas directas de selecção com um vector de expressão e o rastreio apropriado.

As sequências codificantes (por exemplo, genes) de organismos relacionados podem ser utilizadas para praticar esta invenção, e pode demonstrar um grau elevado de homologia e codificam produtos proteicos totalmente diversos. Estes tipos de sequências são particularmente úteis na presente invenção como quasi-repetições. Contudo, embora os protocolos exemplificativos discutidos abaixo demonstrem o rearranjo de sequências codificantes originais quase idênticas (quasi-repetições), este processo não está limitado a tais repetições quase idênticas.

Uma vez formadas, as construções podem ou não ser fraccionadas por tamanho num gel de agarose de acordo com os

protocolos publicados, introduzidas num vector de clonagem e transfectadas numa célula hospedeira apropriada. As células são depois propagadas e o "rearranjo redutor" é efectuado. A velocidade do processo de rearranjo redutor pode ser estimulada através da introdução de lesão de ADN, se desejado. É irrelevante se a redução em RI é mediada por formação de deleção entre sequências repetidas por um mecanismo "intramolecular" ou mediada por eventos de tipo recombinação através de mecanismos "intermoleculares". O resultado final é um rearranjo das moléculas em todas as combinações possíveis.

Num aspecto, os métodos desta invenção compreendem o passo adicional de rastrear os membros da biblioteca do agrupamento rearranjado em combinação para identificar membros individuais da biblioteca rearranjados em combinação que têm a capacidade para se ligarem ou então para interagirem, ou catalisarem uma reacção particular (*e. g.*, tal como catalisar a hidrólise de um fitato).

Num aspecto, os polipéptidos que são identificados a partir de tais bibliotecas podem ser utilizados para fins terapêuticos, diagnósticos, de investigação e relacionados (*e. g.*, catalisadores, solutos para aumentar a osmolaridade de uma solução aquosa e semelhantes) e/ou podem ser submetidos a um ou mais ciclos adicionais de arranjo combinatório e/ou selecção.

Noutro aspecto, antes ou durante a recombinação ou rearranjo, os polinucleótidos da invenção ou os polinucleótidos gerados pelo método aqui descrito são submetidos a agentes ou processos que promovem a introdução de mutações nos polinucleótidos originais. A introdução de tais mutações aumentaria a diversidade de polinucleótidos híbridos resultantes e de polipéptidos aí codificados. Os agentes ou processos que

promovem a mutagénese incluem, mas não estão limitados: (+)-CC-1065 ou um análogo artificial tal como (+)-CC-1065-(N3-Adenina, ver Sun e Hurley, 1992); um aducto de 4'-fluro-4-aminobifenilo N-acetilado ou desacetilado capaz de inibir a síntese de ADN (ver, por exemplo, van de Poll *et al.*, 1992); ou um aducto de 4-aminobifenilo N-acetilado ou desacetilado capaz de inibir a síntese de ADN (ver também, van de Poll *et al.*, 1992, pp. 751-758); crómio trivalente, um sal de crómio trivalente, um aducto de ADN de hidrocarboneto aromático policíclico ("PAH") capaz de inibir a replicação de ADN, tais como 7-bromometil-benz[a]antraceno ("BMA"), tris(2,3-dibromopropil)fosfato ("Tris-BP"), 1,2-dibromo-3-cloropropano ("DBCP"), 2-bromoacroleína (2BA), benzo[a]pireno-7,8-di-hidrodiol-9-10-epóxido ("BPDE"), um sal de halogéneo de platina(II), N-hidroxi-2-amino-3-metilimidazo[4,5-f]-quinolina ("N-hidroxi-IQ") e N-hidroxi-2-amino-1-metil-6-fenilimidazo[4,5-*f*]-piridina ("N-hidroxi-PhIP"). Os meios exemplificativos para retardar ou parar a amplificação de PCR consistem em luz UV (+)-CC-1065 e (+)-CC-1065-(N3-Adenina). Os meios particularmente abrangidos são aductos de ADN ou polinucleótidos compreendendo os aductos de ADN dos polinucleótidos ou do agrupamento de polinucleótidos, que podem ser libertados ou removidos por um processo incluindo aquecimento da solução compreendendo os polinucleótidos antes de processamento adicional.

Noutro aspecto, a invenção é dirigida a um método de produzir proteínas recombinantes tendo actividade biológica através de tratamento de uma amostra compreendendo polinucleótidos de molde cadeia dupla que codificam uma proteína de *tipo selvagem* sob condições de acordo com a invenção que proporcionam a produção de polinucleótidos híbridos ou rearranjados.

A invenção também proporciona a utilização de iniciadores de codão registados (contendo uma sequência degenerada N,N,G/T) para introduzir mutações pontuais num polinucleótido, de modo a produzir um conjunto de polipéptidos descendentes em que uma gama completa de substituições únicas de aminoácidos é representada em cada posição de aminoácido (Mutagénese de Saturação de Local Génico (GSSM)). Os oligos utilizados são compreendidos contiguamente de uma primeira sequência homóloga, uma sequência degenerada N,N,G/T e opcionalmente uma segunda sequência homóloga. Os produtos de tradução descendentes a jusante a partir da utilização de tais oligos incluem todas as alterações de aminoácidos possíveis em cada local de aminoácido ao longo do polipéptido, porque a degenerescência da sequência N,N,G/T inclui codões para todos os 20 aminoácidos.

Num aspecto, um tal oligo degenerado (compreendendo uma cassete degenerada N,N,G/T) é utilizado para submeter cada codão original num molde polinucleotídico parental a uma gama completa de substituições de codão. Num outro aspecto, pelo menos, duas cassetes degeneradas N,N,G/T são utilizadas - no mesmo oligo ou não, para submeter, pelo menos, dois codões originais num molde polinucleotídico parental a uma gama completa de substituições de codão. Assim, mais de uma sequência N,N,G/T pode estar contida num oligo para introduzir mutações de aminoácidos em mais de um local. Esta pluralidade de sequências N,N,G/T pode estar directamente contígua ou separada por uma ou mais sequências nucleotídicas adicionais. Num outro aspecto, os oligos úteis para introduzir adições e deleções podem ser utilizados isoladamente ou em combinação com os codões contendo uma sequência N,N,G/T para introduzir qualquer combinação ou permutação de adições, deleções e/ou substituições de aminoácidos.

Num aspecto, é possível mutar simultaneamente duas ou mais posições de aminoácidos contíguas utilizando um oligo que contenha tripletos N,N,G/T contíguos, *i. e.*, uma sequência degenerada $(N,N,G/T)_n$.

Num outro aspecto, a presente invenção proporciona a utilização de cassetes degeneradas tendo menos degenerescência do que a sequência N,N,G/T. Por exemplo, pode ser desejável em alguns casos utilizar (*e. g.*, num oligo) uma sequência degenerada de tripleto compreendendo apenas um N, em que o referido N pode ser na primeira, segunda ou terceira posição do tripleto. Quaisquer outras bases incluindo quaisquer suas combinações e permutações podem ser utilizadas nas duas posições remanescentes do tripleto. Alternativamente, pode ser desejável em alguns casos utilizar (*e. g.*, num oligo) uma sequência degenerada de tripleto N,N,N ou uma sequência de tripleto N,N,G/C.

É entendido, contudo, que a utilização de um tripleto degenerado (tais como uma sequência de tripleto N,N,G/T ou N,N,G/C) como divulgado na presente invenção é vantajosa por diversas razões. Isto proporciona meios para produzir sistematicamente e de forma razoavelmente fácil a substituição da gama completa de aminoácidos possíveis (para um total de 20 aminoácidos) em todas as posições de aminoácido num polipéptido. Assim, para um polipéptido de 100 aminoácidos, proporciona um modo para gerar sistematicamente e de forma razoavelmente fácil 2000 espécies distintas (*i. e.*, 20 aminoácidos possíveis por posição vezes 100 posições de aminoácidos). É entendido que é proporcionado, através da utilização de um oligo contendo um sequência degenerada de tripleto N,N,G/T ou N,N,G/C, 32 sequências individuais codificando para 20 aminoácidos possíveis. Assim, num vaso

reaccional em que uma sequência polinucleotídica parental é submetida mutagénese de saturação utilizando um tal oligo, são produzidos 32 polinucleótidos descendentes distintos que codificam 20 polipéptidos distintos. Pelo contrário, a utilização de um oligo não degenerado em mutagénese dirigida conduz apenas a um produto polipeptídico descendente por vaso reaccional.

Esta invenção também proporciona a utilização de oligos não degenerados que podem ser opcionalmente utilizados em combinação com iniciadores degenerados divulgados. É entendido que em algumas situações é vantajoso utilizar oligos não degenerados para produzir mutações pontuais específicas num polinucleótido de trabalho. Isto proporciona meios para produzir mutações pontuais silenciosas específicas, mutações pontuais que conduzem a alterações de aminoácidos correspondentes e mutações pontuais que provocam a produção de codões de terminação e a correspondente expressão de fragmentos de polipéptido.

Assim, num aspecto desta invenção, cada vaso reaccional de mutagénese de saturação contém polinucleótidos codificando, pelo menos, 20 moléculas de polipéptido descendente de modo que todos os 20 aminoácidos estão representados na posição de aminoácido específica que corresponde à posição de codão mutada no polinucleótido parental. Os 32 polipéptidos degenerados descendentes produzidos a partir de cada vaso reaccional de mutagénese de saturação podem ser submetidos a amplificação clonal (*e. g.*, clonados num hospedeiro *E. coli* adequado utilizando um vector de expressão) e submetidos a rastreio de expressão. Quando um polipéptido individual descendente for identificado por rastreio por apresentar uma alteração favorável na propriedade (quando comparada ao polipéptido parental), pode

ser sequenciado para identificar a substituição de aminoácido correspondentemente favorável aí contida.

É entendido que após mutação de todas as posições de aminoácido num polipéptido parental utilizando mutagénese de saturação (GSSM) como aqui divulgada, as alterações de aminoácidos favoráveis podem ser identificadas em mais de uma posição de aminoácido. Uma ou mais novas moléculas descendentes podem ser produzidas que contêm uma combinação de todas ou de uma parte destas substituições de aminoácidos favoráveis. Por exemplo, se 2 alterações de aminoácidos favoráveis específicas forem identificadas em cada de 3 posições de aminoácido num polipéptido, as permutações incluem 3 possibilidades em cada posição (nenhuma alteração do aminoácido original e cada de duas alterações favoráveis) e 3 posições. Assim, existem 3 x 3 x 3 ou 27 possibilidades totais, incluindo 7 que foram previamente examinadas - 6 mutações pontuais únicas (*i. e.*, 2 em cada das três posições) e nenhuma alteração em qualquer posição.

Ainda noutro aspecto, a mutagénese de saturação de local pode ser utilizada em conjunto com arranjo combinatório, quimerização, recombinação e outros processos de mutação, em conjunto com rastreio. Esta invenção proporciona a utilização de qualquer processo(s) de mutação, incluindo mutagénese de saturação, de um modo iterativo. Numa exemplificação, a utilização iterativa de qualquer processo(s) de mutação é utilizada em combinação com rastreio.

Assim, numa exemplificação não limitativo, os polinucleótidos e polipéptidos da invenção podem ser derivados por mutagénese de saturação de local génico (GSSM) em combinação com processos adicionais de mutação, tal como processos em que dois ou mais polinucleótidos relacionados são introduzidos numa

célula hospedeira adequada de modo que um polinucleótido híbrido seja produzido por recombinação e rearranjo redutor.

Além disso ao realizar mutagénese ao longo da sequência inteira de um gene, a mutagénese pode ser utilizada para substituir cada de qualquer número de bases numa sequência de polinucleótidos, o número de bases para ser mutado é, num aspecto, todos os números inteiros de 15 a 100000. Assim, em vez de mutar cada posição ao longo de uma molécula, pode-se submeter todas ou de um número discreto de bases (num aspecto, um subconjunto que totaliza de 15 a 100000) a mutagénese. Num aspecto, um nucleótido separado é utilizado para mutar cada posição ou grupo de posições ao longo de uma sequência polinucleotídica. Um grupo de 3 posições a ser mutada pode ser um codão. As mutações podem ser introduzidas utilizando um iniciador mutagénico, contendo uma cassete heteróloga, também referida como uma cassete mutagénica. Cassetes exemplificativas podem ter de 1 a 500 bases. Cada posição de nucleótido em tais cassetes heterólogas é N, A, C, G, T, A/C, A/G, A/T, C/G, C/T, G/T, C/G/T, A/G/T, A/C/T, A/C/G ou E, em que E é qualquer que não é A, C, G ou T (E pode ser referida como a um oligo desenhador).

Num sentido geral, a mutagénese de saturação é compreendida de mutação de um conjunto completo de cassetes mutagénicas (em que num aspecto cada cassete tem cerca de 1-500 bases de comprimento) numa sequência polinucleotídica definida para ser mutada (em que num aspecto a sequência a ser mutada tem desde cerca de 15 a 100000 bases de comprimento). Assim, um grupo de mutações (variando de 1 a 100 mutações) é introduzido em cada cassete a ser mutada. Um grupo de mutações a ser introduzido numa cassete pode ser diferente ou igual a um segundo grupo de mutações a ser introduzido numa segunda cassete durante a

aplicação de um ciclo de mutagénese de saturação. Tais grupos são exemplificados por deleções, adições, grupos de codões particulares e grupos de cassetes nucleotídicas particulares.

As sequências definidas a serem mutadas incluem um gene inteiro, via, ADNc, uma grelha de leitura aberta inteira (ORF) e promotor, intensificador, repressor/transactivador, origem de replicação, intrão, operador inteiro ou qualquer grupo polinucleotídico funcional. Geralmente, as "sequências definidas" para este objectivo podem ser qualquer polinucleótido que sejam uma sequência polinucleotídica de 15 bases e sequências polinucleotídicas de comprimentos entre 15 bases e 15000 bases (esta invenção refere especificamente cada número inteiro entre esses números). As considerações de escolha de grupos de codões incluem tipos de aminoácidos codificados por uma cassete mutagénica degenerada.

Numa exemplificação, um grupo de mutações que pode ser introduzido numa cassete mutagénica, esta invenção especificamente proporciona substituições de codão degenerado (utilizando oligos degenerados) codificando 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 e 20 aminoácidos em cada posição e uma biblioteca de polipéptidos assim codificados.

Em aspectos alternativos, os ácidos nucleicos da invenção compreendem ADN, incluindo ADNc, ADN genómico e ADN sintético. O ADN pode ser de cadeia dupla ou cadeia simples e, se for de cadeia simples, pode ser a cadeia codificante ou a cadeia não codificante (anti-sentido). Alternativamente, os ácidos nucleicos da invenção podem compreender ARN.

Como discutido mais detalhadamente abaixo, as sequências de ácidos nucleicos isoladas da invenção podem ser utilizadas para preparar os polipéptidos da invenção.

Desta forma, um outro aspecto da invenção é uma sequência de ácidos nucleicos isolada codificando um polipéptido da invenção. Uma sequência de ácidos nucleicos da invenção pode compreender sequências codificantes adicionais, tais como sequências líder ou sequências de proproteína e sequências não codificantes, tais como intrões ou sequências 5' e/ou 3' não codificantes da sequência codificante. Assim, como aqui utilizado, a expressão "polinucleótido codificando um polipéptido" abrange um polinucleótido que inclui apenas a sequência codificante para o polipéptido, assim como um polinucleótido que inclui sequências codificantes e/ou não codificantes adicionais.

Alternativamente, as sequências de ácidos nucleicos da invenção podem ser mutadas utilizando técnicas convencionais, tal como mutagénese dirigida ou outras técnicas familiares aos especialistas na técnica, para introduzir alterações silenciosas no polinucleótido da invenção. Como aqui utilizado, "alterações silenciosas" incluem, por exemplo, alterações que não alteram a sequência de aminoácidos codificada pelo polinucleótido. Tais alterações podem ser desejáveis de modo a aumentar o nível do polipéptido produzido por células hospedeiras contendo um vector que codifica o polipéptido, através de introdução de codões ou pares de codões que ocorrem frequentemente no organismo hospedeiro.

A invenção também se refere a polinucleótidos que têm alterações nucleotídicas que resultam em substituições, adições, deleções, fusões e truncamentos de aminoácidos nos polipéptidos da invenção. Tais alterações nucleotídicas podem ser

introduzidas, utilizando técnicas, tais como mutagénese dirigida, mutagénese química aleatória, deleção por exonuclease III e outras técnicas de ADN recombinante.

Sempre que necessário, as condições que permitem que a sonda hibride especificamente com sequências complementares podem ser determinadas através de colocação da sonda em contacto com as sequências complementares de amostras que se sabem conter as sequências complementares, assim como sequências de controlo que não contêm a sequência complementar. As condições de hibridação, tais como a concentração de sal do tampão de hibridação, a concentração de formamida do tampão de hibridação ou a temperatura de hibridação, podem ser variadas para identificar condições que permitam que a sonda hibride especificamente com ácidos nucleicos complementares.

Se a amostra contiver o organismo a partir do qual o ácido nucleico foi isolado, é então detectada hibridação específica da sonda. A hibridação pode ser detectada por marcação da sonda com um agente detectável, tal como um isótopo radioactivo, um corante fluorescente ou uma enzima capaz de catalisar a formação de um produto detectável.

Muitos métodos para utilização de sondas marcadas para detectar a presença de ácidos nucleicos complementares numa amostra são familiares dos especialistas na técnica. Estes incluem Transferência de Southern, Transferência de Northern, processos de hibridação de colónia e transferências por gota. Os protocolos para cada um destes processos são proporcionados em Ausubel *et al*. Current Protocols in Molecular Biology, John Wiley 503 Sons, Inc. 1997 e Sambrook *et al*., Molecular Cloning: A Laboratory Manual, 2ª Ed., Cold Spring Harbor Laboratory Press, 1989.

Alternativamente, mais de uma sonda (pelo menos uma das quais é capaz de hibridar especificamente com quaisquer sequências complementares que estão presentes na amostra de ácido nucleico) pode ser utilizada numa reacção de amplificação para determinar se a amostra contém um organismo contendo uma sequência de ácidos nucleicos da invenção (*e. g.*, um organismo a partir do qual foi isolado o ácido nucleico). Tipicamente, as sondas compreendem oligonucleótidos. Num aspecto, a reacção de amplificação pode compreender uma reacção de PCR. Os protocolos de PCR são descritos em Ausubel e Sambrook, *supra*. Alternativamente, a amplificação pode compreender uma reacção em cadeia da ligase, 3SR, ou uma reacção de deslocamento de cadeia. (ver Barany, F., "The Ligase Chain Reaction in a PCR World", PCR Methods and Applications 1:5-16, 1991; E. Fahy *et al.*, "Self-sustained Sequence Replication (3SR): An Isothermal Transcription-based Amplification System Alternative to PCR", PCR Methods and Applications 1:25-33, 1991; e Walker G.T. *et al.*, "Strand Displacement Amplification-an Isothermal in vitro DNA Amplification Technique", Nucleic Acid Research 20:1691-1696, 1992). Em tais processos, os ácidos nucleicos na amostra são colocados em contacto com as sondas, a reacção de amplificação é realizada e qualquer produto de amplificação resultante é detectado. O produto de amplificação pode ser detectado por realização de electroforese em gel nos produtos da reacção e por coloração do gel com um intercalador, tal como brometo de etídio. Alternativamente, uma ou mais das sondas podem ser marcadas com um isótopo radioactivo e a presença de um produto de amplificação radioactivo pode ser detectada através de autorradiografia após electroforese em gel.

Sondas derivadas de sequências perto das extremidades de uma sequência da invenção podem também ser utilizadas em processos

de progressão no cromossoma para identificar clones contendo sequências genómicas localizadas adjacentemente às sequências de ácidos nucleicos como apresentados acima. Tais métodos permitem o isolamento dos genes que codificam proteínas adicionais do organismo hospedeiro.

Uma sequência de ácidos nucleicos da invenção pode ser utilizada como uma sonda para identificar e isolar ácidos nucleicos relacionados. Em alguns aspectos, os ácidos nucleicos relacionados podem ser ADNc ou ADN genómico de organismos diferentes daquele onde o ácido nucleico foi isolado. Por exemplo, os outros organismos podem ser organismos relacionados. Em tais processos, uma amostra de ácido nucleico é colocada em contacto com a sonda sob condições que permitem a sonda hibridar especificamente com as sequências relacionadas. A hibridação da sonda com ácidos nucleicos do organismo relacionado é depois detectada utilizando qualquer dos métodos descritos acima.

Em reacções de hibridação de ácidos nucleicos, as condições utilizadas para conseguir um nível particular de restringência vai variar dependendo da natureza dos ácidos nucleicos hibridados. Por exemplo, o comprimento, grau de complementaridade, composição da sequência de nucleótidos (*e. g.*, teor de GC vs. AT) e tipo de ácido nucleico (*e. g.*, ADN vs. ARN) das regiões de hibridação dos ácidos nucleicos podem ser considerados na selecção das condições de hibridação. Uma consideração adicional é se um dos ácidos nucleicos está imobilizado, por exemplo, num filtro.

A hibridação pode ser realizada sob condições de restringência baixa, restringência moderada ou restringência elevada. Como um exemplo de hibridação de ácidos nucleicos, uma membrana polimérica contendo ácidos nucleicos desnaturados

imobilizados é pré-hibridada primeiro durante 30 minutos a 45 °C numa solução consistindo de 0,9 M de NaCl, 50 mM de NaH_2PO_4, pH 7,0, 5,0 mM de Na_2EDTA, 0,5% de SDS, 10X reagente de Denhardt e 0,5 mg/mL de ácido polirriboadenílico. Aproximadamente 2 x 10^7 cpm (actividade específica de 4-9 x 10^8 cpm/ug) de sonda oligonucleotídica marcada na extremidade com ^{32}P é depois adicionada à solução. Após 12-16 horas de incubação, a membrana é lavada durante 30 minutos à temperatura ambiente em 1X SET (150 mM de NaCl, 20 mM de cloridrato de Tris, pH 7,8, 1 mM de Na_2EDTA) contendo 0,5% de SDS, seguida por uma lavagem de 30 minutos em 1X SET fresco a Tm-10 °C para a sonda oligonucleotídica. A membrana é depois exposta a película autorradiográfica para detecção de sinais de hibridação.

Por variação da restringência das condições de hibridação utilizadas para identificar ácidos nucleicos, tais como ADNc ou ADN genómico, os quais hibridam com a sonda detectável, os ácidos nucleicos que têm níveis diferentes de homologia com a sonda podem ser identificados e isolados. A restringência pode ser variada por realização da hibridação em temperaturas variadas abaixo das temperaturas de fusão das sondas. A temperatura de fusão, T_m, é a temperatura (sob força iónica e pH definidos) em que 50% da sequência alvo hibrida com uma sonda perfeitamente complementar. Condições muito restringentes são seleccionadas para serem iguais ou cerca de 5 °C inferiores à T_m para uma sonda particular. A temperatura de fusão da sonda pode ser calculada utilizando as seguintes fórmulas: Para sondas entre 14 e 70 nucleótidos de comprimento, a temperatura de fusão (T_m) é calculado utilizando a fórmula: T_m = 81,5 + 16,6(log[Na+])+ 0,41(fracção G+C)-(600/N), em que N é o comprimento da sonda. Se a hibridação for realizada numa solução contendo formamida, a temperatura de fusão pode ser calculada utilizando a equação: T_m = 81,5+16,6(log[Na+])+0,41(fracção G+C)-

(0,63% de formamida)- (600/N), em que N é o comprimento da sonda. A pré-hibridação pode ser realizada em 6X SSC, 5X reagente de Denhardt, 0,5% de SDS, 100 µg/mL de ADN de esperma de salmão fragmentado desnaturado ou 6X SSC, 5X reagente de Denhardt, 0,5% de SDS, 100 µg/mL de ADN de esperma de salmão fragmentado desnaturado, 50% de formamida. As fórmulas para as soluções SSC e de Denhardt são encontradas, *e. g.*, em Sambrook *et al.*, *supra*.

A hibridação é conduzida por adição da sonda detectável às soluções de pré-hibridação listadas acima. Nos casos em que a sonda compreende ADN de cadeia dupla, este é desnaturado antes da adição à solução de hibridação. O filtro é colocado em contacto com a solução de hibridação durante um período de tempo suficiente para permitir que a sonda hibride com os ADNc ou ADN genómicos contendo sequências complementar a isso ou homólogas a isso. Para sondas com mais de 200 nucleótidos de comprimento, a hibridação pode ser realizada a 15-25 °C abaixo da Tm. Para sonda mais curtas, tais como sondas oligonucleotídicas, a hibridação pode ser conduzida a 5-10 °C abaixo da T_m. Tipicamente, para hibridações em 6X SSC, a hibridação é conduzida a aproximadamente 68 °C. Geralmente, para hibridações em soluções contendo 50% de formamida, a hibridação é conduzida a aproximadamente 42 °C. Todas as hibridações anteriores são consideradas como sendo sob condições de restringência elevada.

Após a hibridação, o filtro é lavado para remover qualquer sonda detectável ligada não especificamente. A restringência utilizada para lavar os filtros pode também ser variada dependendo da natureza dos ácidos nucleicos hibridados, comprimento dos ácidos nucleicos hibridados, grau de complementaridade, composição da sequência de nucleótidos (*e. g.*, teor de GC vs. AT) e tipo de ácido nucleico (*e. g.*, ADN

vs. ARN). Os exemplos de lavagens em condições de restringência progressivamente mais elevadas são como se seguem: 2X SSC, 0,1% de SDS à temperatura ambiente durante 15 minutos (restringência baixa); 0,1X SSC, 0,5% de SDS à temperatura ambiente durante 30 minutos a 1 hora (restringência moderada); 0,1X SSC, 0,5% de SDS durante 15 a 30 minutos entre a temperatura de hibridação e 68 °C (restringência elevada); e 0,15 M de NaCl durante 15 minutos a 72 °C (restringência muito elevada). Uma lavagem final de restringência baixa pode ser conduzida em 0,1X SSC à temperatura ambiente. Os exemplos acima são meramente ilustrativos de um conjunto de condições que podem ser utilizadas para lavar filtros. Uma especialista na técnica sabe que existem numerosos protocolos para lavagens com diferentes restringências. Alguns outros exemplos são apresentados abaixo.

Os ácidos nucleicos que hibridaram com a sonda podem ser identificados por autorradiografia ou outras técnicas convencionais.

O processo acima pode ser modificado para identificar os ácidos nucleicos que têm níveis menores de homologia com a sequência da sonda. Por exemplo, para obter ácidos nucleicos de menor homologia com a sonda detectável, podem ser utilizadas condições menos restringentes. Por exemplo, a temperatura de hibridação pode ser diminuída em incrementos de 5 °C de 68 °C a 42 °C num tampão de hibridação tendo uma concentração de Na^+ de aproximadamente, 1 M. Após hibridação, o filtro pode ser lavado com 2X SSC, 0,5% de SDS na temperatura de hibridação. Estas condições são consideradas como sendo condições “moderadas” acima de 50 °C e condições “baixas” abaixo de 50 °C. Um exemplo específico de condições de hibridação “moderada” é quando a hibridação acima é conduzida a 55 °C. Um exemplo específico de

condições de hibridação de “restringência baixa” é quando a hibridação acima é conduzida a 45 °C.

Alternativamente, a hibridação pode ser realizada em tampões, tal como 6X SSC, contendo formamida a uma temperatura de 42 °C. Neste caso, a concentração de formamida no tampão de hibridação pode ser reduzida incrementos de 5% desde 50% a 0% para identificar clones tendo níveis menores de homologia com a sonda. Após hibridação, o filtro pode ser lavado com 6X SSC, 0,5% de SDS a 50 °C. Estas condições são consideradas como sendo condições “moderadas” acima de 25% de formamida e condições “baixas” abaixo de 25% de formamida. Um exemplo específico de condições de hibridação “moderadas” é quando a hibridação acima é conduzida a 30% de formamida. Um exemplo específico de condições de hibridação de “restringência baixa” é quando a hibridação acima é conduzida a 10% de formamida.

Por exemplo, os métodos anteriores podem ser utilizados para isolar ácidos nucleicos tendo uma sequência com, pelo menos, cerca de 99%, pelo menos, 98%, pelo menos, 97% ou, pelo menos, 95% de homologia com uma sequência de ácidos nucleicos como apresentada na SEQ ID N°:1, sequências substancialmente idênticas a isso, ou fragmentos compreendendo, pelo menos, cerca de 10, 15, 20, 25, 30, 35, 40, 50, 75, 100, 150, 200, 300, 400, ou 500 bases consecutivas da mesma, e as sequências complementares a quaisquer das sequências anteriores. A homologia pode ser medida utilizando um algoritmo de alinhamento. Por exemplo, os polinucleótidos homólogos podem ter uma sequência codificante que é uma variante alélica de ocorrência natural de uma das sequências codificantes aqui descritas. Tais variantes alélicas podem ter uma substituição, deleção ou adição de um ou mais nucleótidos quando comparadas a

uma sequência de ácidos nucleicos como apresentado na SEQ ID N°:1 ou as suas sequências complementares.

Adicionalmente, os processos acima podem ser utilizados para isolar ácidos nucleicos codificando polipéptidos tendo, pelo menos, cerca de 99% ou, pelo menos, 95% de homologia com um polipéptido que tem uma sequência como apresentada na SEQ ID N°:2, sequências substancialmente idênticas a isso, ou fragmentos compreendendo, pelo menos, 5, 10, 15, 20, 25, 30, 35, 40, 50, 75, 100 ou 150 aminoácidos consecutivos da mesma como determinado utilizando um algoritmo de alinhamento de sequências (*e. g.*, tal como o algoritmo FASTA versão 3.0t78 com os parâmetros de defeito).

Um outro aspecto da invenção é um polipéptido isolado ou purificado como discutido acima. Tais polipéptidos podem ser obtidos através de introdução de um ácido nucleico codificand o polipéptido num vector de modo que a sequência codificante esteja ligada operativamente a uma sequência capaz de dirigir a expressão do polipéptido codificado numa célula hospedeira adequada. Por exemplo, o vector de expressão pode compreender um promotor, um local de ligação a ribossoma para iniciação da tradução e um terminador de transcrição. O vector pode também incluir sequências apropriadas para amplificar a expressão.

Os promotores adequados para expressar o polipéptido ou um seu fragmento em bactérias incluem os promotores *lac* ou *trp* de *E. coli*, o promotor *lacI*, o promotor *lacZ*, o promotor *T3*, o promotor T7, o promotor gpt, o promotor *lambda* P_R, o promotor *lambda* P_L, promotores de operões que codificam enzimas glicolíticas, tais como 3-fosfoglicerato cinase (PGK) e o promotor da fosfatase ácida. Os promotores fúngicos incluem o promotor do factor ∀. Os promotores eucarióticos incluem o

promotor precoce imediato de CMV, o promotor da timidina cinase de HSV, promotores de choque térmico, o promotor precoce e tardio de SV40, LTR de retrovírus e o promotor de metalotioneína-I de murganho. Outros promotores conhecidos por controlarem a expressão de genes em células procarióticas ou eucarióticas ou nos seus vírus podem também ser utilizados.

Os vectores de expressão de mamíferos podem também compreender uma origem de replicação, quaisquer locais de ligação a ribossoma necessários, um local de poliadenilação, locais dadores e aceitadores de *splice*, sequência de terminação de transcrição e sequências não transcritas flanqueadoras 5'. Em alguns aspectos, as sequências de ADN derivadas dos locais de *splice* e de poliadenilação de SV40 podem ser utilizadas para proporcionar os elementos genéticos não transcritos requeridos.

Os vectores para expressar o polipéptido ou um seu fragmento em células eucarióticas podem também conter intensificadores para aumentar níveis de expressão. Os intensificadores são elementos de actuação cis de ADN, habitualmente desde cerca de 10 a cerca de 300 pb de comprimento que actuam num promotor para aumentarem a sua transcrição. Os exemplos incluem o intensificador de SV40 no lado tardio dos pb 100 a 270 da origem de replicação, o intensificador do promotor precoce de citomegalovírus, o intensificador de polioma no lado tardio da origem de replicação e os intensificadores de adenovírus.

Além disso, os vectores de expressão contêm tipicamente um ou mais genes marcadores de selecção para permitirem a selecção de células hospedeiras contendo o vector. Tais marcadores de selecção incluem genes que codificam a di-hidrofolato redutase ou genes que conferem resistência à neomicina para cultura de células eucarióticas, genes que conferem resistência à

tetraciclina ou ampicilina em *E. coli*, e o gene *TRP1* de *S. cerevisiae*.

O ADN da sonda utilizado para isolar selectivamente o ADN alvo de interesse do ADN derivado de, pelo menos, um microrganismo pode ser uma sequência de região codificante de tamanho total ou uma sequência parcial de região codificante de ADN para uma enzima de actividade conhecida. A biblioteca de ADN original pode ser sondada utilizando misturas de sondas compreendendo, pelo menos, uma porção da sequência de ADN que codifica uma enzima tendo actividade enzimática especificada. Estas sondas ou bibliotecas de sondas podem ser de cadeia simples e o ADN microbiano que é sondado pode ser convertido na forma de cadeia simples. As sondas que são adequadas são aquelas que são derivadas de ADN que codifica enzimas tendo actividade semelhante ou idêntica à actividade enzimática especificada que é para ser rastreada.

O ADN da sonda pode ter, pelo menos, cerca de 10 bases ou, pelo menos, 15 bases. Num aspecto, a região codificante inteira pode ser utilizada como uma sonda. As condições para a hibridação em que o ADN alvo é isolado selectivamente através da utilização de, pelo menos, uma sonda de ADN, serão concebidas para proporcionar uma restringência de hibridação de, pelo menos, cerca de 50% de identidade de sequência, mais particularmente, uma restringência que proporciona uma identidade de sequência de, pelo menos, cerca de 70%.

O ADN da sonda pode ser "marcado" com um parceiro de um par de ligação específico (*i. e.*, um ligando) e o outro parceiro do par é ligado a uma matriz sólida para proporcionar facilidade de separação do alvo da sua fonte. O ligando e o parceiro de ligação específico podem ser seleccionados, em qualquer

orientação, do seguinte: (1) um antigénio ou hapteno e um anticorpo ou um seu fragmento específico de ligação; (2) biotina ou iminobiotina e avidina ou estreptavidina; (3) um açúcar e uma lecitina específica para este; (4) uma enzima e um inibidor para esta; (5) um apoenzima e um cofactor; (6) oligonucleótidos homopoliméricos complementares; e (7) uma hormona e um receptor para esta. A fase sólida pode ser seleccionada de: (1) uma superfície de vidro ou polimérica; (2) uma coluna empacotada de esferas poliméricas; e (3) partículas magnéticas ou paramagnéticas.

A sequência de ADN apropriada pode ser introduzida no vector através de uma variedade de processos. Em geral, a sequência de ADN está ligada à posição desejada no vector, após digestão da inserção e do vector com endonucleases de restrição apropriadas. Alternativamente, extremidades rombas na inserção e no vector podem ser ligadas. Uma variedade de técnicas de clonagem é divulgada em Ausubel *et al*. Current Protocols in Molecular Biology, John Wiley and Sons, Inc. 1997 e Sambrook *et al*., Molecular Cloning: A Laboratory Manual, 2ª Ed., Cold Spring Harbor Laboratory Press, 1989. Tais processos e outros consideram-se dentro do âmbito dos especialistas na técnica.

O vector pode ser, por exemplo, na forma de um plasmídeo, uma partícula viral ou um fago. Outros vectores incluem sequências de ADN cromossomal, não cromossomal e sintético, derivados de SV40; plasmídeos bacterianos, ADN fágico, baculovírus, plasmídeos de levedura, vectores derivados de combinações de plasmídeos e de ADN fágico, ADN viral tal como vaccinia, adenovírus, vírus da varíola aviária e pseudorraiva. Uma variedade de vectores de clonagem e expressão para utilização com hospedeiros procarióticos e eucarióticos são

descritos por Sambrook, *et al.*, Molecular Cloning: A Laboratory Manual, Segunda Edição, Cold Spring Harbor, N.I., (1989).

Os vectores bacterianos particulares que podem ser utilizados, incluem os plasmídeos comercialmente disponíveis compreendendo elementos genéticos do vector de clonagem bem conhecido pBR322 (ATCC 37017), pKK223-3 (Pharmacia Fine Chemicals, Uppsala, Suécia), GEM1 (Promega Biotec, Madison, WI, EUA) pQE70, pQE60, pQE-9 (Qiagen), pD10, psiX174 pBluescript II KS, pNH8A, pNH16a, pNH18A, pNH46A (Stratagene), ptrc99a, pKK223-3, pKK233-3, pDR540, pRIT5 (Pharmacia), pKK232-8 e pCM7. Os vectores eucarióticos particulares incluem pSV2CAT, pOG44, pXT1, pSG (Stratagene) pSVK3, pBPV, pMSG e pSVL (Pharmacia). Contudo, qualquer outro vector pode ser utilizado desde que seja replicável e viável na célula hospedeira.

A célula hospedeira pode ser qualquer das células hospedeiras familiares aos especialistas na técnica, incluindo células procarióticas, células eucarióticas, células de mamíferos, células de insecto ou células vegetais. Como exemplos representativos de hospedeiros apropriados, podem ser mencionados: células bacterianas, tais como *E. coli*, *Streptomyces*, *Bacillus subtilis*, *Bacillus cereus*, *Salmonella typhimurium* e várias espécies dentro do género *Pseudomonas*, *Streptomyces* e *Staphylococcus*, células fúngicas, tal como *Aspergillus*, levedura, tais como quaisquer espécies de *Pichia*, *Saccharomyces*, *Schizosaccharomyces*, *Schwanniomyces*, incluindo *Pichia pastoris*, *Saccharomyces cerevisiae*, ou *Schizosaccharomyces pombe*, células de insecto, tais como *Drosophila S2* e *Spodoptera Sf9*, células animais, tais como CHO, COS ou melanoma de Bowes, e adenovírus. A selecção de um hospedeiro apropriado está dentro das capacidades dos especialistas na técnica.

O vector pode ser introduzido nas células hospedeiras utilizando qualquer de uma variedade de técnicas, incluindo transformação, transfecção, transdução, infecção viral, pistolas génicas ou transferência génica mediada por Ti. Os métodos particulares incluem transfecção com fosfato de cálcio, transfecção mediada por DEAE-Dextrano, lipofecção ou electroporação (Davis, L., Dibner, M., Battey, I., Basic Methods in Molecular Biology, (1986)).

Sempre que apropriado, as células hospedeiras manipuladas podem ser cultivadas em meios nutrientes convencionais modificados como apropriado para activar promotores, seleccionar transformantes ou amplificar os genes da invenção. Após transformação de uma estirpe hospedeira adequada e crescimento da estirpe hospedeira até uma densidade celular apropriada, o promotor seleccionado pode ser induzido por meios apropriados (*e. g.*, deslocamento de temperatura ou indução química) e as células podem ser cultivadas durante um período adicional para permitir que produzam o polipéptido desejado ou um seu fragmento.

As células são tipicamente recolhidas por centrifugação, quebradas por meios químicos ou físicos e o extracto em bruto resultante é retido para purificação adicional. As células microbianas utilizadas para expressão de proteínas podem ser quebradas através de qualquer método conveniente, incluindo ciclos de congelamento-descongelamento, sonicação, quebra mecânica ou utilização de agentes de lise celular. Tais métodos são bem conhecidos dos especialistas na técnica. O polipéptido expresso ou um seu fragmento pode ser recuperado e purificado a partir de culturas celulares recombinantes por métodos incluindo precipitação com sulfato de amónio ou etanol, extracção ácida,

cromatografia de permuta iónica ou catiónica, cromatografia de fosfocelulose, cromatografia de interacção hidrofóbica, cromatografia de afinidade, cromatografia de hidroxilapatite e cromatografia de lecitina. Podem ser utilizados passos de re-enrolamento de proteína, como necessário, para completar a configuração do polipéptido. Se desejado, pode ser utilizada cromatografia líquida de elevado desempenho (HPLC) para passos finais de purificação.

Vários sistemas de cultura de células de mamífero podem também ser utilizados para expressar a proteína recombinante. Os exemplos de sistemas de expressão de mamífero incluem as linhas COS-7 de fibroblastos de rim de macaco (descritas por Gluzman, Cell, 23:175, 1981) e outras linhas celulares capazes de expressar proteínas a partir de um vector compatível, tais como as linhas celulares C127, 3T3, CHO, HeLa e BHK.

As construções em células hospedeiras podem ser utilizadas de um modo convencional para produzir o produto génico codificado pela sequência recombinante. Dependendo do hospedeiro utilizado num processo de produção recombinante, os polipéptidos produzidos por células hospedeiras contendo o vector podem ser glicosilados ou podem ser não glicosilados. Os polipéptidos da invenção podem ou não também incluir um resíduo de aminoácido de metionina inicial. Detalhes adicionais que se referem à expressão recombinante de proteínas estão disponíveis para os especialistas na técnica. Por exemplo, Protein Expression: A Practical Approach (Practical Approach Series por S. J. Higgins (Editor), B. D. Hames (Editor) (Julho 1999) Oxford University Press; ISBN: 0199636249 proporciona ampla orientação para os especialistas na técnica para a expressão de proteínas numa vasta variedade de organismos.

Alternativamente, os polipéptidos da invenção podem ser produzidos sinteticamente por sintetizadores de péptidos convencionais. Em outros aspectos, os fragmentos ou porções dos polipéptidos podem ser utilizados para produzir o polipéptido de tamanho total correspondente por síntese peptídica; deste modo, os fragmentos podem ser utilizados como intermediários para produzir os polipéptidos de tamanho total.

Como conhecido dos especialistas na técnica, as sequências de ácidos nucleicos da invenção podem ser optimizadas para expressão numa variedade de organismos. Num aspecto, as sequências da invenção são optimizadas para utilização de codão num organismo de interesse, *e. g.*, um fungo, tal como *S. cerevisiae* ou uma bactéria, tal como *E. coli*. A optimização de sequências de ácidos nucleicos para fins de utilização de codão é bem entendida na técnica para se referir à selecção de um codão particular favorecido por um organismo para codificar um aminoácido particular. As tabelas de utilização de codão optimizada são conhecidas para muitos organismos. Por exemplo, ver Transfer RNA in Protein Synthesis por Dolph L. Hatfield, Byeong J. Lee, Robert M. Pirtle (Editor) (Julho 1992) CRC Press; ISBN: 0849356989. Assim, a invenção também inclui ácidos nucleicos da invenção adaptados para a utilização de codão de um organismo.

A expressão optimizada de sequências de ácidos nucleicos da invenção também se refere a mutagénese dirigida ou aleatória de um ácido nucleico para conduzir a expressão aumentada da proteína codificada. A mutagénese dos ácidos nucleicos da invenção pode directa ou indirectamente proporcionar um rendimento aumentado da proteína expressa. A título de exemplo não limitativo, as técnicas de mutagénese aqui descritas podem ser utilizadas para efectuar a mutação da região 5' não

traduzida, região 3’ não traduzida ou região codificante de um ácido nucleico, cuja mutação pode resultar em estabilidade aumentada ao nível de ARN ou proteína, resultando assim num rendimento aumentado de proteína.

Os sistemas de tradução sem células podem também ser utilizados para produzir um dos polipéptidos da invenção, utilizando ARNm transcritos a partir de uma construção de ADN compreendendo um promotor ligado operativamente a um ácido nucleico que codifica o polipéptido ou um seu fragmento. Em alguns aspectos, a construção de ADN pode ser linearizada antes de ser conduzida uma reacção de transcrição *in vitro*. O ARNm transcrito é depois incubado com um extracto de tradução sem células apropriado, tal como um extracto de reticulócito de coelho, para produzir o polipéptido desejado ou um seu fragmento.

A invenção também se refere a variantes dos polipéptidos da invenção. O termo “variante” inclui derivados ou análogos destes polipéptidos. Em particular, as variantes podem diferir na sequência de aminoácidos dos polipéptidos da invenção, e de sequências substancialmente idênticas a estas, por uma ou mais substituições, adições, deleções, fusões e truncamentos, que podem estar presentes em qualquer combinação.

As variantes podem ser de ocorrência natural ou criadas *in vitro*. Em particular, tais variantes podem ser criadas utilizando técnicas de engenharia genética, tais como mutagénese dirigida, mutagénese química aleatória, processos de deleção por Exonuclease III e técnicas de clonagem convencionais. Alternativamente, tais variantes, fragmentos, análogos ou derivados podem ser criados utilizando processos de modificação ou síntese química.

Outros métodos para fazer variantes são também familiares aos especialistas na técnica. Estes incluem processos em que sequências de ácidos nucleicos obtidas a partir de isolados naturais são modificadas para produzir ácidos nucleicos que codificam polipéptidos que têm as características que aumentam o seu valor em aplicações industriais ou laboratoriais. Em tais processos, um grande número de sequências variantes tendo uma ou mais diferenças de nucleótido relativamente à sequência obtida a partir do isolado natural são produzidas e caracterizadas. Tipicamente, estas diferenças de nucleótidos resultam em alterações de aminoácidos relativamente aos polipéptidos codificados pelos ácidos nucleicos de isolados naturais.

Por exemplo, as variantes podem ser criadas utilizando PCR propensa ao erro. Na PCR propensa ao erro, a PCR é a realizada sob condições em que a fidelidade da ADN polimerase é baixa, de modo que é obtida uma taxa elevada de mutações pontuais ao longo da totalidade do comprimento do produto de PCR. A PCR propensa ao erro é descrita em Leung, D.W., *et al.*, Technique, 1:11-15, 1989) e Caldwell, R. C. e Joyce G.F., PCR Methods Applic., 2:28-33, 1992. Resumidamente, em tais processos, os ácidos nucleicos a serem mutados são misturados com iniciadores de PCR, tampão de reacção, $MgCl_2$, $MnCl_2$, Taq polimerase e uma concentração apropriada de dNTP para conseguir uma taxa elevada de mutações pontuais ao longo da totalidade do comprimento do produto de PCR. Por exemplo, a reacção pode ser realizada utilizando 20 fmoles de ácido nucleico para ser mutado, 30 pmole de cada iniciador de PCR, tampão de reacção compreendendo 50 mM de KCl, 10 mM de Tris HCl (pH 8,3) e 0,01% de gelatina, 7 mM de $MgCl_2$, 0,5 mM de $MnCl_2$, 5 unidades de Taq polimerase, 0,2 mM de dGTP, 0,2 mM de dATP, 1 mM de dCTP e 1 mM de dTTP. A PCR pode ser realizada durante 30 ciclos de 94 °C durante 1 minuto, 45 °C

durante 1 minuto e 72 °C durante 1 min. Contudo, deve ser entendido que estes parâmetros podem ser alterados como apropriado. Os ácidos nucleicos mutados são clonados num vector apropriado e as actividades dos polipéptidos codificados pelos ácidos nucleicos mutados são avaliadas.

As variantes podem também ser criadas utilizando mutagénese dirigida por oligonucleótido para produzir mutações específicas de local em qualquer ADN clonado de interesse. A mutagénese por oligonucleótido é descrita em Reidhaar-O'son, J.F. e Sauer, R.T., *et al.*, Science, 241:53-57, 1988. Resumidamente, em tais processos uma pluralidade de oligonucleótidos de cadeia dupla que contêm uma ou mais mutações a serem introduzidas no ADN clonado são sintetizados e introduzidos no ADN clonado a ser mutado. Os clones contendo o ADN mutado são recuperados e as actividades dos polipéptidos que codificam são avaliadas.

Um outro método para produzir variantes é PCR de montagem. A PCR de montagem envolve a montagem de um produto de PCR a partir de uma mistura de pequenos fragmentos de ADN. Um grande número de diferentes reacções de PCR ocorre em paralelo no mesmo frasquinho, com os produtos de uma reacção a iniciarem os produtos de uma outra reacção. A PCR de montagem é descrita no Pedido de Patente U.S. N° de série 08/677112 apresentado em 9 de Julho de 1996, pendente, intitulado, Método de "DNA Shuffling with Polynucleotides Produced by Blocking or interrupting a Synthesis or Amplification Process".

Ainda um outro método de produzir variantes é a mutagénese por PCR sexual. Na mutagénese por PCR sexual, recombinação homóloga forçada ocorre entre moléculas de ADN de sequência de ADN diferente mas altamente relacionada *in vitro,* como resultado de fragmentação aleatória da molécula de ADN baseada em

homologia de sequência, seguida por fixação do cruzamento por extensão de iniciador numa reacção de PCR. A mutagénese por PCR sexual é descrita em Stemmer, W.P., PNAS, USA, 91:10747-10751, 1994. Resumidamente, em tais processos uma pluralidade de ácidos nucleicos a serem recombinados são digeridos com DNase para gerar fragmentos que têm um tamanho médio de 50-200 nucleótidos. Os fragmentos de tamanho médio desejado são purificados e ressuspensos numa mistura de PCR. A PCR é conduzida sob condições que facilitam a recombinação entre os fragmentos de ácidos nucleicos. Por exemplo, a PCR pode ser realizada por ressuspensão dos fragmentos purificados a uma concentração de 10-30 ng/:1 numa solução de 0,2 mM de cada dNTP, 2,2 mM de $mgCl_2$, 50 mM de KCl, 10 mM de Tris HCl pH 9,0 e 0,1% de Triton X-100, 2,5 unidades de Taq polimerase por 100:1 da mistura de reacção é adicionado e a PCR é realizada utilizando o seguinte regime: 94 °C durante 60 segundos, 94 °C durante 30 segundos, 50-55 °C durante 30 segundos, 72 °C durante 30 segundos (30-45 vezes) e 72 °C durante 5 minutos. Contudo, deve ser entendido esse estes parâmetros podem ser alterados como apropriado. Em alguns aspectos, oligonucleótidos podem ser incluídos nas reacções de PCR. Em outros aspectos, o fragmento Klenow da ADN polimerase I pode ser utilizado num primeiro conjunto de reacções de PCR e a Taq polimerase pode ser utilizada num subsequente conjunto de reacções de PCR. As sequências recombinantes são isoladas e as actividades dos polipéptidos que codificam são avaliadas.

As variantes podem também ser criadas através de mutagénese *in vivo*. Em alguns aspectos, mutações aleatórias numa sequência de interesse são geradas através de propagação da sequência de interesse numa estirpe bacteriana, como uma estirpe de *E. coli*, que contém mutações em uma ou mais das vias de reparação de ADN. Tal estirpe "mutadora" tem uma taxa mais elevada de mutação aleatória do que aquela de *tipo selvagem* parental. A propagação

de ADN numa destas estirpes irá produzir eventualmente mutações aleatórias dentro do ADN. As estirpes mutadoras adequadas para utilização em mutagénese *in vivo* são descritas na Publicação de PCT N° WO 91/16427, publicada em 31 de Outubro de 1991, intitulada "Methods for Phenotype Creation from Multiple Gene Populations".

Variantes podem também ser produzidas utilizando mutagénese de cassete. Na mutagénese de cassete, uma pequena região de uma molécula de ADN de cadeia dupla é substituída com uma "cassete" de oligonucleótido sintético que difere da sequência nativa. O oligonucleótido contém frequentemente a sequência nativa totalmente e/ou parcialmente aleatorizada.

A mutagénese de conjunto recursivo pode também ser utilizada para produzir variantes. A mutagénese de conjunto recursivo é um algoritmo para engenharia de proteína (mutagénese de proteína) desenvolvido para produzir populações diversas de mutantes relacionados fenotipicamente, cujos membros diferem na sequência de aminoácidos. Este método utiliza um mecanismo de feedback para controlar ciclos sucessivos de mutagénese combinatória de cassete. A mutagénese de conjunto recursivo é descrita em Arkin, A.P. e Youvan, C.C., PNAS, USA, 89:7811-7815, 1992.

Em alguns aspectos, as variantes são criadas utilizando mutagénese de conjunto exponencial. A mutagénese de conjunto exponencial é um processo para produzir bibliotecas combinatórias com uma percentagem elevada de mutantes únicos e funcionais, em que pequenos grupos de resíduos são aleatorizados em paralelo para identificar, em cada posição alterada, os aminoácidos que conduzem a proteínas funcionais. A mutagénese de conjunto exponencial é descrita em Delegrave, S. e Youvan, D.C., Biotechnol. Res., 11:1548-1552, 1993. As mutagéneses aleatórias

e dirigidas são descritas em Arnold, F.H., Current Opinion in Biotechnology, 4:450-455, 1993.

Em alguns aspectos, as variantes são criadas utilizando processos de arranjo combinatório em que porções de uma pluralidade de ácidos nucleicos codificando polipéptidos distintos são fundidas em conjunto para criar sequências de ácidos nucleicos quiméricos que codificam polipéptidos quiméricos como descrito no Pedido de Patente U.S. Nº de série 08/677112 apresentado em 9 de Julho de 1996, pendente, intitulado "Method of DNA Shuffling with Polynucleotides Produced by Blocking or interrupting a Synthesis or Amplification Process" e no Pedido de Patente U.S. Nº de série 08/651568 apresentado em 22 de Maio de 1996, intitulado "Combinatorial Enzyme Development".

As variantes dos polipéptidos da invenção podem ser variantes em que um ou mais dos resíduos de aminoácidos dos polipéptidos da invenção são substituídos com resíduos de aminoácidos conservados ou não conservados (*e. g.*, um resíduo de aminoácido conservado) e tal resíduo de aminoácido substituído pode ou não ser um codificado pelo código genético.

As substituições conservadoras são aquelas que substituem um dado aminoácido num polipéptido por outro aminoácido de características semelhantes. São vistas tipicamente como substituições conservadoras as seguintes substituições: substituições de um aminoácido alifático, tais como Ala, Val, Leu e Ile com um outro aminoácido alifático; substituição de um Ser com uma Thr ou vice-versa; substituição de um resíduo acídico, tais como Asp e Glu com outro resíduo acídico; substituição de um resíduo que contém um grupo amida, tais como Asn e Gln, com outro resíduo que contém um grupo amida; troca de

um resíduo básico, tais como Lys e Arg com outro resíduo básico; e substituição de um resíduo aromático, tais como Phe, Tyr com outro resíduo aromático.

Outras variantes são aquelas em que um ou mais dos resíduos de aminoácidos dos polipéptidos da invenção incluem um grupo substituinte.

Ainda outras variantes são aquelas em que o polipéptido está associado com outro composto, tal como um composto para aumentar a semi-vida do polipéptido (por exemplo, polietilenoglicol).

Variantes adicionais são aquelas em que aminoácidos adicionais são fundidos ao polipéptido, tais como uma sequência líder, uma sequência secretora, uma sequência de proproteína ou uma sequência que facilita a purificação, enriquecimento ou estabilização do polipéptido. Em alguns aspectos, os derivados e os análogos retêm a mesma função ou actividade biológica que os polipéptidos da invenção e podem incluir uma proproteína, de modo que o fragmento, derivados ou análogo pode ser activado por clivagem da porção proproteína para produzir um polipéptido activo.

Optimizar codões para conseguir níveis elevados de expressão proteica em células hospedeiras

A invenção proporciona métodos para modificar ácidos nucleicos que codificam fitases para modificar a utilização de codão. Num aspecto, a invenção proporciona métodos para modificar codões num ácido nucleico codificando uma fitase para aumentar ou diminuir a sua expressão numa célula hospedeira. A invenção também proporciona ácidos nucleicos codificando uma

fitase modificada para aumentar a sua expressão numa célula hospedeira, enzimas fitases assim modificadas e métodos para fazer as enzimas fitases modificadas. O método compreende identificar um codão "não preferido" ou "menos preferido" em ácido nucleico codificando uma fitase e substituir um ou mais destes codões não preferidas ou menos preferidos com um "codão preferido" codificando o mesmo aminoácido que o codão substituído e, pelo menos, um codão não preferido ou menos preferido no ácido nucleico foi substituído por um codão preferido que codifica o mesmo aminoácido. Um codão preferido é um codão sobre-representado em sequências codificantes em genes na célula hospedeira e um codão não preferido ou menos preferido é um codão sob-representado em sequências codificantes nos genes na célula hospedeira.

As células hospedeiras para expressar os ácidos nucleicos, cassetes de expressão e vectores da invenção incluem bactérias, leveduras, fungos, células vegetais, células de insecto e células de mamífero. Assim, a invenção proporciona métodos para optimizar a utilização de codão em todas estas células, ácidos nucleicos de codão alterado e polipéptidos feitos pelos ácidos nucleicos de codão alterado. As células hospedeiras exemplificativas incluem bactérias Gram negativas, tais como *Escherichia coli* e *Pseudomonas fluorescens;* bactérias Gram positivas, tais como *Lactobacillus gasseri, Lactococcus lactis, Lactococcus cremoris, Bacillus subtilis.* As células hospedeiras exemplificativas incluem também organismos eucarióticos, *e. g.*, várias leveduras, tais como *Saccharomyces sp.*, incluindo *Saccharomyces cerevisiae, Schizosaccharomyces pombe, Pichia pastoris* e *Kluyveromyces lactis, Hansenula polimorpha, Aspergillus niger* e células de mamífero e linhas celulares, e célula de insecto e linhas celulares. Assim, a invenção inclui

também ácidos nucleicos e polipéptidos optimizados para a expressão nestes organismos e espécies.

Por exemplo, os codões de um ácido nucleico codificando uma fitase isolada de uma célula bacteriana são modificados de modo que o ácido nucleico é optimamente expresso numa célula bacteriana diferente da bactéria da qual a fitase foi derivada, um levedura, um fungo, uma célula vegetal, uma célula de insecto ou uma célula de mamífero. Os métodos para optimizar codões são bem conhecidos na técnica, ver, *e. g.*, Patente U.S. N° 5795737; Baca (2000) Int. J. Parasitol. 30:113-118; Hale (1998) Protein Expr. Purif. 12:185-188; Narum (2001) Infect. Immun. 69:7250-7253. Ver também Narum (2001) infect. Immun. 69:7250-7253, que descreve a optimização de codões em sistemas de murganho; Outchkourov (2002) Protein Expr. Purif. 24:18-24, que descreve a optimização de codões em levedura; Feng (2000) Biochemistry 39:15399-15409, que descreve a optimização de codões em *E. coli;* Humphreys (2000) Protein Expr. Purif. 20:252-264, que descreve a optimização da utilização de codão que afecta a secreção em *E. coli.*

Animais não humanos transgénicos

Os animais não humanos transgénicos podem compreender um ácido nucleico, um polipéptido, uma cassete de expressão ou um vector ou uma célula transfectada ou transformada da invenção. Os animais não humanos transgénicos podem ser, *e. g.*, cabras, coelhos, carneiros, porcos, vacas, ratos e murganhos, compreendendo os ácidos nucleicos da invenção. Estes animais podem ser utilizados, *e. g.*, como modelos *in vivo* para estudar a actividade de fitase ou como modelos para rastrear para moduladores da actividade de fitase *in vivo.* As sequências

codificantes para que os polipéptidos a serem expressos nos animais não humanos transgénicos podem ser concebidos para serem constitutivos ou sob o controlo de factores reguladores transcricionais induzíveis ou específicos de tecido, específicos de desenvolvimento. Os animais não humanos transgénicos podem ser concebidos e gerados utilizando qualquer método conhecido na técnica; ver, *e. g.*, Patente U.S. Nº 6211428; 6187992; 6156952; 6118044; 6111166; 6107541; 5959171; 5922854; 5892070; 5880327; 5891698; 5639940; 5573933; 5387742; 5087571, que descrevem descrevendo a preparação e utilização de células e ovos transformados e murganhos, ratos, coelhos, carneiros, porcos e vacas transgénicos. Ver também, *e. g.*, Pollock (1999) J. Immunol. Methods 231:147-157, que descreve a produção de proteínas recombinantes no leite de animais leiteiros transgénicos; Baguisi (1999) Nat. Biotechnol. 17:456-461, que demonstra a produção de cabras transgénicas. A Patente U.S. Nº 6211428 descreve a preparação e utilização de mamíferos não humanos transgénicos que expressam nos seus cérebros uma construção de ácido nucleico compreendendo uma sequência de ADN. A Patente U.S. Nº 5387742 descreve a injecção de sequências de ADN recombinantes ou sintéticas clonadas em ovos de murganho fertilizados, implantação dos ovos injectados em fêmeas pseudo-grávidas e crescimento de murganhos transgénicos cujas células expressam as proteínas relacionadas com a patologia da doença de Alzheimer. A Patente U.S. Nº 6187992 descreve a preparação e utilização de murganhos transgénicos cujo genoma compreende uma interrupção do gene que codifica a proteína precursora amilóide (APP).

Os "animais com gene inactivado" podem também ser utilizados. Por exemplo, num aspecto, os animais transgénicos ou modificados compreendem um "animal com gene inactivado", *e. g.*,

um “murganho com gene inactivado”, manipulado para não expressar ou para ser incapaz de expressar uma fitase.

Noutro aspecto, são proporcionados organismos não humanos transgénicos que contêm uma sequência heteróloga que codifica uma fitase da invenção (*e. g.*, as modificações de sequência enumeradas especificamente da SEQ ID N°:2). Vários métodos para fazer os animais transgénicos podem ser utilizados. Em termos gerais, três de tais métodos podem ser utilizados. Num de tal método, um embrião na fase pronuclear (um “embrião de célula única”) é recolhido a partir de uma fêmea e o transgene é microinjectado no embrião, caso em que o transgene será integrado cromossomicamente nas células germinais e células somáticas do animal maduro resultante. Noutro tal método, células estaminais embrionárias são isoladas e o transgene aí incorporado por electroporação, transfecção de plasmídeo ou microinjecção, seguida por reintrodução das células estaminais no embrião onde colonizam e contribuem para a linha germinal. Métodos para microinjecção de espécies de mamíferos são descritos na Pat. U.S. N° 4873191.

Ainda noutro método exemplificativo, células embrionárias são infectadas com um retrovírus contendo o transgene, pelo que as células germinais do embrião têm o transgene aí integrado cromossomicamente. Quando os animais a serem feitos transgénicos são aves, porque os ovos fertilizados de aves avançam para divisão celular durante as primeiras vinte horas no oviducto, a microinjecção no pronúcleo do ovo fertilizado é problemática devido à inacessibilidade do pronúcleo. Deste modo, dos métodos para fazer animais transgénicos descritos genericamente acima, a infecção por retrovírus é preferida para espécie aviárias, por exemplo, como descrito na Pat. U.S. N° 5162215. Se a micro-injecção for para ser utilizada com espécies aviárias, contudo,

pode ser utilizado um processo publicado por Love *et al.*, (Biotechnol., 12, jan 1994), em que o embrião é obtido a partir de uma galinha sacrificada, aproximadamente, duas horas e meia após a postura do último ovo, o transgene é microinjectado no citoplasma do disco germinal e o embrião é cultivado numa casca hospedeira até à maturidade. Quando os animais a serem feitos transgénicos são bovinos ou porcinos, a microinjecção pode ser dificultada pela opacidade dos ovos fazendo assim com que os núcleos sejam difíceis de identificar através de microscopia de contraste por interferência diferencial tradicional. Para superar este problema, os ovos podem ser primeiro centrifugados para segregar os pronúcleos para melhor visualização.

Num aspecto, os "animais não humanos" incluem animais bovinos, porcinos, ovinos e aviários (*e. g.*, vaca, porco, carneiro, galinha). Os "animais não humanos transgénicos" da invenção são produzidos por introdução de "transgenes" na linha germinal do animal não humano. As células alvo embrionárias em várias fases de desenvolvimento são utilizadas para introduzir transgenes. São utilizados métodos diferentes, dependendo da fase de desenvolvimento da célula alvo embrionária. O zigoto é melhor alvo para a micro-injecção. A utilização de zigotos como alvo para a transferência génica tem uma importante vantagem uma vez que na maioria dos casos o ADN injectado será incorporado no gene hospedeiro antes da primeira clivagem (Brinster *et al.*, Proc. Natl. Acad. Sci. USA 82:4438-4442, 1985). Consequentemente, todas as células do animal não humano transgénico irão conter o transgene incorporado. Isto também será reflectido genericamente na transmissão eficiente do transgene à prole do animal uma vez que 50% das células germinais irá abrigar o transgene.

Num aspecto, o termo "transgénico" é utilizado para descrever um animal que inclui material genético exógeno dentro de todas as suas células. Um animal "transgénico" pode ser produzido por reprodução cruzada de dois animais quiméricos que incluem material genético exógeno dentro das células utilizadas na reprodução. Vinte e cinco porcento da prole resultante será transgénica, *i. e.*, animais que incluem o material genético exógeno dentro de todas as suas células em ambos os alelos, 50% dos animais resultantes incluirá o material genético exógeno dentro de um alelo e 25% não incluirá nenhum material genético exógeno.

Num aspecto, um método de microinjecção é utilizado. O transgene é digerido e purificado livre de qualquer ADN de vector, *e. g.*, por electroforese em gel. Num aspecto, o transgene inclui um promotor operativamente associado que interage com proteínas celulares envolvidas em transcrição, resultando no final em expressão constitutiva. Os promotores úteis nesta consideração incluem aqueles de citomegalovírus (CMV), vírus da leucemia de Moloney (MLV) e vírus do herpes, assim como os aqueles de genes que codificam a metalotionina, actina esquelética, P-enolpiruvate carboxilase (PEPCK), fosfoglicerato (PGK), DHFR e timidina cinase. Os promotores para repetições terminais longas virais (LTR), tal como do vírus do sarcoma de Rous são também utilizados. Quando os animais a serem feitos transgénicos são aviários, os promotores preferidos podem incluir aqueles para o gene da b-globina de galinha, gene da lisozima de galinha e o vírus da leucose aviária. As construções úteis no transfecção de plasmídeos de célula estaminais embrionárias irão empregar elementos reguladores adicionais bem conhecidos na técnica, tais como elementos intensificadores para estimular a transcrição, aceitadores de *splice*, sinais de

terminação e poliadenilação e locais de ligação a ribossoma para permitir a tradução.

A infecção retroviral pode também ser utilizada para introduzir transgenes num animal não humano, como descrito acima. O embrião não humano em desenvolvimento pode ser cultivado *in vitro* até à fase de blastocisto. Durante este tempo, os blastómeros podem ser alvos para infecção retroviral (Jaenich, R., Proc. Natl. Acad. Sci. USA 73:1260-1264, 1976). A infecção eficiente dos blastómeros é obtida por tratamento enzimático para remover a zona pelúcida (Hogan, *et al.* (1986) em Manipulating the Mouse Embryo, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.I.). O sistema do vector viral utilizado para introduzir o transgene é tipicamente um retrovírus de replicação defeituosa que contém o transgene (Jahner, *et al.*, Proc. Natl. Acad. Sci. USA 82: 6927-6931, 1985; van der Putten, *et al.*, Proc. Natl. Acad. Sci. USA 82: 6148-6152, 1985). A transfecção é fácil e eficientemente obtida através do cultivo dos blastómeros numa monocamada de células produtoras de vírus (van der Putten, *supra;* Stewart, *et al.*, EMBO J. 6: 383-388, 1987). Alternativamente, a infecção pode ser realizada numa fase posterior. O vírus ou células produtoras de vírus podem ser injectados no blastocélio (D. Jahner *et al.*, Nature 298: 623-628, 1982). A maioria dos animais receptores será mosaico relativamente ao transgene uma vez que a incorporação ocorre apenas num subconjunto das células que formam o animal não humano transgénico. Além disso, o animal receptor pode conter várias inserções retrovirais do transgene em posições diferentes no genoma que habitualmente irão segregar na prole. Além disso, é também possível introduzir transgenes na linha germinal, embora com baixa eficiência, por infecção retroviral intrauterina do embrião em gestação média (D. Jahner *et al.*, *supra*).

Um terceiro tipo de célula alvo para a introdução de transgenes é a célula estaminal embrionária (ES). As células ES são obtidas a partir de embriões pré-implantados cultivados *in vitro* e fundidos com embriões (M. J. Evans *et al.*, Nature 292:154-156, 1981; M. O. Bradley *et al.*, Nature 309:255-258, 1984; Gossler, *et al.*, Proc. Natl. Acad. Sci. USA 83:9065-9069, 1986; e Robertson *et al.*, Nature 322:445-448, 1986). Os transgenes podem ser introduzidos eficientemente nas células ES por transfecção de ADN ou por transdução mediada por retrovírus. Tais células ES transformadas podem ser posteriormente combinadas com blastocistos de um animal não humano. As células ES posteriormente colonizam o embrião e contribuem para a linha germinal do animal quimérico resultante. (para revisão ver Jaenisch, R., Science 240:1468-1474, 1988).

Num aspecto, "transformada" significa uma célula dentro da qual (ou numa sua antepassada) foi introduzida, através de técnicas de ácidos nucleicos recombinantes, uma molécula de ácido nucleico heterólogo. "Heterólogo" refere-se a uma sequência de ácidos nucleicos que é originária de outra espécie ou está modificada da sua forma original ou da forma primeiro expressa na célula.

Num aspecto, o "transgene" significa qualquer porção de ADN que é introduzida numa célula por artifício e torna-se parte de genoma do organismo (*i. e.*, integrado estavelmente ou como um elemento extracromossomal estável) que se desenvolve a partir dessa célula. Tal transgene pode incluir um gene que é parcial ou totalmente heterólogo (*i. e.*, estranho) ao organismo transgénico, ou pode representar um gene homólogo para um gene endógeno do organismo. Incluído dentro desta definição está um transgene criado ao proporcionar uma sequência de ARN que é

transcrita em ADN e depois incorporada no genoma. Os transgenes da invenção incluem sequências de ADN codificando fitases ou polipéptidos tendo actividade de fitase e incluem polinucleótidos, os quais podem ser expressos num animal não humano transgénico. Como aqui utilizado, o termo "transgénico" inclui adicionalmente qualquer organismo cujo genoma for alterado através de manipulação *in vitro* do embrião inicial ou do ovo fertilizado ou por qualquer tecnologia transgénica para induzir uma inactivação de gene específica. A expressão "inactivação de gene", como aqui utilizada, refere-se à interrupção direccionada de um gene *in vivo* com perda completa da função que foi conseguida por qualquer tecnologia transgénica familiar aos especialistas na técnica. Num aspecto, os animais transgénicos tendo inactivações de genes são aqueles em que o gene alvo foi tornado não funcional por uma inserção dirigida ao gene a ser tornado não funcional por recombinação homóloga.

Num aspecto, o termo "transgénico" inclui qualquer tecnologia transgénica familiar aos especialistas na técnica que pode produzir um organismo contendo um transgene introduzido ou um organismo em que um gene endógeno foi tornado não funcional ou "inactivado".

O transgene a ser utilizado na realização da presente invenção é uma sequência de ADN compreendendo uma sequência codificante para uma fitase ou um polipéptido tendo actividade de fitase. Num aspecto, um polinucleótido tendo uma sequência como apresentada na SEQ ID N°:1 ou uma sequência codificando um polipéptido que tem uma sequência como apresentada na SEQ ID N°:2 é o transgene, tal como o termo é aqui definido. Sempre que apropriado, sequências de ADN codificando proteínas tendo actividade de fitase, mas diferem na sequência de ácidos nucleicos devido à degenerescência do código genético, podem

também ser aqui utilizadas, assim como podem formas truncadas, variantes alélicas e homólogos inter-espécies.

Num aspecto, após um embrião ser microinjectado, colonizado com células estaminais embrionárias transfectadas ou infectado com um retrovírus contendo o transgene (excepto para a prática em espécies aviárias que é abordada noutra parte deste documento), o embrião é implantado no oviducto de uma fêmea pseudo-grávida. A consequente descendência é testada para a incorporação do transgene por análise de transferência de Southern de amostras de sangue ou tecido, utilizando sondas específicas para o transgene. PCR é particularmente útil nesta fase. A descendência positiva (G0) é reproduzida por cruzamento para produzir prole (G1) que é analisada para expressão do transgene por análise de transferência de Northern de amostras de tecido.

Num aspecto, os métodos compreendem o aumento da incorporação de fósforo no animal transgénico e/ou diminuição da quantidade de *polltant* nas fezes do organismo transgénico em cerca de 15%, cerca de 20% ou cerca de 20%, a cerca de 50% ou mais.

Num aspecto, os animais contemplados são aqueles animais habitualmente considerados como animais domesticos, incluindo animais de estimação (*e. g.*, espécies caninas, felinas, aviárias, etc.) e aqueles úteis para o processamento de alimentos, *i. e.*, aviários, tais como galinhas e perus para a produção de carne e ovos, ovinos tal como cordeiro, bovinos, tais como gado produtor de carne e vacas leiteiras, peixes e porcinos. Num aspecto, estes animais são referidos como "transgénicos" quando tal animal teve uma sequência de ADN heterólogo, ou uma ou mais sequências de ADN adicionais

normalmente endógenas ao animal (referidas colectivamente aqui como “transgenes”) integradas cromossomicamente nas células germinais do animal. O animal transgénico (incluindo a sua descendência) terá também o transgene integrado fortuitamente nos cromossomas de células somáticas.

Metodologias de Rastreio e dispositivos de Monitorização “Em-linha”

Na prática dos métodos da invenção, uma variedade de dispositivos e metodologias podem ser utilizados em conjunto com os polipéptidos e ácidos nucleicos da invenção, *e. g.*, para rastrear polipéptidos para actividade de fitase, para rastrear compostos como potenciais moduladores da actividade (*e. g.*, potenciação ou inibição da actividade enzimática), para anticorpos que se ligam a um polipéptido da invenção, para ácidos nucleicos que hibridam com um ácido nucleico da invenção e semelhantes.

Suportes Sólidos para Enzimas Imobilizadas

As enzimas fitase, os seus fragmentos e os ácidos nucleicos que codificam as enzimas e fragmentos podem ser fixados a um suporte sólido. Isto é frequentemente económico e eficiente na utilização das fitases em processos industriais. Por exemplo, um consórcio ou cocktail de enzimas fitase (ou dos seus fragmentos activos), que são utilizadas numa reacção química específica, pode ser ligado a um suporte sólido e mergulhado num vaso reaccional de processo. A reacção enzimática pode ocorrer. Depois, o suporte sólido pode ser removido para fora do vaso reaccional, em conjunto com as enzimas aí fixadas, para

utilização repetida. Numa forma de realização da invenção, um ácido nucleico isolado da invenção é afixado a um suporte sólido. Numa outra forma de realização da invenção, o suporte sólido é seleccionado do grupo de um gel, uma resina, um polímero, um cerâmico, um vidro, um microelétrodo e qualquer sua combinação.

Por exemplo, Os suportes sólidos úteis nesta invenção incluem géis. Alguns exemplos de géis incluem sepharose, gelatinA, glutaraldeído, glutaraldeído tratado com quitosano, albumina-glutaraldeído, quitosano-Xantano, gel toyopearl (gel polimérico), alginato, alginato-polilisina, carragenina, agarose, glioxil-agarose, agarose magnética, dextrano-agarose, hidrogel de poli(sulfonato de carbamoílo), hidrogel de BSA-PEG, álcool polivinílico fosforilado (PVA), monoaminoetil-N-aminoetilo (MANA), amino ou qualquer sua combinação.

Um outro suporte sólido útil na presente invenção é resinas ou polímeros. Alguns exemplos de resinas ou polímeros incluem celulose, acrilamida, nylon, rayon, poliéster, resina de permuta aniónica, AMBERLITE™ XAD-7, AMBERLITE™ XAD-8, AMBERLITE™ IRA-94, AMBERLITE™ IRC-50, polivinilo, poliacrílico, polimetacrilato ou qualquer sua combinação. Um outro tipo de suporte sólido útil na presente invenção é a cerâmica. Alguns exemplos incluem cerâmica não porosa, cerâmica porosa, SiO_2, Al_2O_3. Um outro tipo de suporte sólido útil na presente invenção é o vidro. Alguns exemplos incluem vidro não poroso, vidro poroso, vidro de aminopropilo ou qualquer sua combinação. Um outro tipo de suporte sólido que pode ser utilizado é um microelétrodo. Um exemplo é uma magnetite revestida com polietilenoimina. Partículas grafíticas podem ser utilizadas como um suporte sólido. Um outro exemplo de um suporte sólido é uma célula, tal como um glóbulo vermelho.

Métodos de imobilização

Existem muitos métodos que são conhecidos de um especialista na técnica para imobilizar enzimas ou os seus fragmentos, ou ácidos nucleicos, num suporte sólido. Alguns exemplos de tais métodos incluem, *e. g.*, produção de gota electrostática, meios electroquímicos, por adsorção, por ligação covalente, por reticulação, por uma reacção ou processo químico, por encapsulação, por captura, por alginato de cálcio ou por poli(metacrilato de 2-hidroxietilo). Métodos semelhantes são descritos em Methods in Enzymology, Immobilized Enzymes and Cells, Part C. 1987. Academic Press. Editado por S. P. Colowick e N. O. Kaplan. Volume 136; e Immobilization of Enzymes and Cells. 1997. Humana Press. Editado por G. F. Bickerstaff. Series: Methods in Biotechnology, Editado por J. M. Walker.

Matrizes Capilares

Matrizes capilares, tal como a GIGAMATRIX™, Diversa Corporation, San Diego, CA, podem ser utilizadas nos métodos da invenção. Os ácidos nucleicos ou polipéptidos da invenção podem ser imobilizados ou aplicados a uma matriz, incluindo matrizes capilares. As matrizes podem ser utilizadas para rastrear ou monitorizar bibliotecas de composições (*e. g.*, moléculas pequenas, anticorpos, ácidos nucleicos, etc.) para a sua capacidade de se ligarem ou modularem a actividade de um ácido nucleico ou um polipéptido da invenção. As matrizes capilares proporcionam um outro sistema para fixar e rastrear amostras. Por exemplo, um dispositivo de rastreio de amostras pode incluir uma pluralidade de capilares formados numa matriz de capilares adjacentes, em que cada capilar compreende, pelo menos, uma

parede que define um lúmen para reter uma amostra. O dispositivo pode ainda incluir material intersticial disposto entre capilares adjacentes na matriz e uma ou mais indicação de referência formada dentro do material intersticial. Um capilar para rastrear uma amostra, em que o capilar está adaptado para ser ligado a uma matriz de capilares, pode incluir uma primeira parede que define um lúmen para reter a amostra e uma segunda parede formada de um material filtrante, para filtrar energia de excitação proporcionada ao lúmen para excitar a amostra.

Um polipéptido ou ácido nucleico, *e. g.*, um ligando, pode ser introduzido num primeiro componente em, pelo menos, uma porção de um capilar de uma matriz capilar. Cada capilar da matriz capilar pode compreender, pelo menos, uma parede que define um lúmen para reter o primeiro componente. Uma bolha de ar pode ser introduzida no capilar atrás do primeiro componente. Um segundo componente pode ser introduzido no capilar, em que o segundo componente é separado do primeiro componente através da bolha de ar. Uma amostra de interesse pode ser introduzida como um primeiro líquido marcado com uma partícula detectável num capilar de uma matriz capilar, em que cada capilar da matriz capilar compreende, pelo menos, uma parede que define um lúmen para reter o primeiro líquido e a partícula detectável, e em que a pelo menos uma parede é revestida de um material de ligação para ligar a partícula detectável à pelo menos uma parede. O método pode ainda incluir remover o primeiro líquido do tubo capilar, em que a partícula detectável ligada é mantida dentro do capilar e introduzir um segundo líquido no tubo capilar.

A matriz capilar pode incluir uma pluralidade de capilares individuais compreendendo, pelo menos, uma parede exterior que define um lúmen. A parede exterior do capilar pode ser uma ou mais paredes fundidas em conjunto. De modo semelhante, a parede

pode definir um lúmen que é cilíndrico, quadrado, sextavado ou qualquer outra forma geométrica desde que as paredes formem um lúmen para retenção de um líquido ou amostra. Os capilares da matriz capilar podem ser fixos em conjunto em íntima proximidade para formar uma estrutura planar. Os capilares podem ser ligados em conjunto, ao serem fundidos (*e. g.*, quando os capilares são feitos de vidro), colados, ligados ou agarrados lado-a-lado. A matriz capilar pode ser formada de qualquer número de capilares individuais, por exemplo, um da gama de 100 a 4000000 capilares. Uma matriz capilar pode formar uma placa de microtitulação tendo cerca de 100000 ou mais capilares individuais ligados em conjunto.

Matrizes ou "BioChips"

Os ácidos nucleicos ou polipéptidos da invenção podem ser imobilizados ou aplicados a uma matriz. As matrizes podem ser utilizadas para rastrear ou monitorizar bibliotecas de composições (*e. g.*, moléculas pequenas, anticorpos, ácidos nucleicos, etc.) para a sua capacidade de ligação ou modulação da actividade de um ácido nucleico ou um polipéptido da invenção. Por exemplo, num aspecto da invenção, um parâmetro monitorizado é a expressão de transcrito de um gene de fitase. Um ou mais ou todos os transcritos de uma célula podem ser medidos por hibridação de uma amostra compreendendo transcritos da célula ou ácidos nucleicos representativos ou complementares a transcritos de uma célula, por hibridação a ácidos nucleicos imobilizados numa matriz ou "biochip". Ao utilizar uma "matriz" de ácidos nucleicos num microchip, alguns ou todos os transcritos de uma célula podem ser quantificados simultaneamente. Alternativamente, matrizes compreendendo ácidos nucleicos genómicos podem também ser utilizados para determinar

o genótipo de uma estirpe recentemente manipulada feita pelos métodos da invenção. “Matrizes de polipéptidos” podem também ser utilizadas para quantificar simultaneamente uma pluralidade de proteínas.

Em aspectos alternativos, “matrizes” ou “micromatrizes” ou “biochips” ou “chips” da invenção compreendem uma pluralidade de elementos alvo para além de um ácido nucleico e/ou um polipéptido ou péptido da invenção; cada elemento alvo pode compreender uma quantidade definida de um ou mais polipéptidos (incluindo anticorpos) ou ácidos nucleicos imobilizados sobre numa área definida de uma superfície de substrato, como discutido mais detalhadamente abaixo.

A presente invenção pode ser praticada com qualquer “matriz” conhecida, também referida como “micromatriz” ou “matriz de ácido nucleico” ou “matriz de polipéptido” ou “matriz de anticorpo” ou “biochip” ou suas variações. As matrizes são genericamente uma pluralidade de “pontos” ou “elementos alvo”, cada elemento alvo compreendendo uma quantidade definida de uma ou mais moléculas biológicas, *e. g.*, oligonucleótidos, imobilizados sobre uma área definida de uma superfície de substrato para ligação específica a uma molécula de amostra, *e. g.*, transcritos de ARNm.

Na prática dos métodos da invenção, qualquer matriz e/ou método conhecido para fazer e utilizar matrizes pode ser utilizado na totalidade ou em parte, ou as suas variações, como descrito, por exemplo, nas patentes U.S. N° 6277628; 6277489; 6261776; 6258606; 6054270; 6048695; 6045996; 6022963; 6013440; 5965452; 5959098; 5856174; 5830645; 5770456; 5632957; 5556752; 5143854; 5807522; 5800992; 5744305; 5700637; 5556752; 5434049; ver também, *e. g.*, documentos WO 99/51773; WO 99/09217;

WO 97/46313; WO 96/17958; ver também, *e. g.*, Johnston (1998) Curr. Biol. 8:r171-r174; Schummer (1997) Biotechniques 23:1087-1092; Kern (1997) Biotechniques 23:120-124; Genes, Chromosomes & Cancer 20:399-407; Bowtell (1999) Nature Genetics Supp. 21:25-32. Ver também Pedidos de Patente U.S. Publicados N° 20010018642; 20010019827; 20010016322; 20010014449; 20010014448; 20010012537; 20010008765.

Polipéptidos e péptidos

A invenção proporciona polipéptidos isolados, sintéticos ou recombinantes tendo uma sequência de aminoácidos com, pelo menos, 95%, 96%, 97%, 98% ou 99% de identidade de sequência para a SEQ ID N°:2 e codificados por um polinucleótido compreendendo uma específica e opcionalmente, pelo menos, uma, duas, três, quatro, cinco, seis, sete, oito, nove, dez, onze, doze ou todas as treze modificações específicas adicionais de sequência de pares de bases nucleotídicos para a SEQ ID N°:1, descritos acima, ou tendo uma sequência de aminoácidos compreendendo uma específica e opcionalmente, pelo menos, uma, duas, três, quatro, cinco, seis, sete, oito, nove, dez, onze, doze ou todas as treze modificações adicionais de resíduos de aminoácidos para a SEQ ID N°:2, descritos acima. Para referência, a SEQ ID N°:2 "parental" sinteticamente produzida é:

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

Lys Leu Gly Glu Leu Thr Pro Arg Gly Gly Glu Leu Ile Ala Tyr Leu
65 70 75 80
Gly His Tyr Trp Arg Gln Arg Leu Val Ala Asp Gly Leu Leu Pro Lys
85 90 95
Cys Gly Cys Pro Gln Ser Gly Gln Val Ala Ile Ile Ala Asp Val Asp
100 105 110
Glu Arg Thr Arg Lys Thr Gly Glu Ala Phe Ala Ala Gly Leu Ala Pro
115 120 125
Asp Cys Ala Ile Thr Val His Thr Gln Ala Asp Thr Ser Ser Pro Asp
130 135 140
Pro Leu Phe Asn Pro Leu Lys Thr Gly Val Cys Gln Leu Asp Asn Ala
145 150 155 160
Asn Val Thr Asp Ala Ile Leu Glu Arg Ala Gly Gly Ser Ile Ala Asp
165 170 175
Phe Thr Gly His Tyr Gln Thr Ala Phe Arg Glu Leu Glu Arg Val Leu
180 185 190
Asn Phe Pro Gln Ser Asn Leu Cys Leu Lys Arg Glu Lys Gln Asp Glu
195 200 205
Ser Cys Ser Leu Thr Gln Ala Leu Pro Ser Glu Leu Lys Val Ser Ala
210 215 220
Asp Cys Val Ser Leu Thr Gly Ala Val Ser Leu Ala Ser Met Leu Thr
225 230 235 240
Glu Ile Phe Leu Leu Gln Gln Ala Gln Gly Met Pro Glu Pro Gly Trp
245 250 255
Gly Arg Ile Thr Asp Ser His Gln Trp Asn Thr Leu Leu Ser Leu His
260 265 270
Asn Ala Gln Phe Asp Leu Leu Gln Arg Thr Pro Glu Val Ala Arg Ser
275 280 285
Arg Ala Thr Pro Leu Leu Asp Leu Ile Lys Thr Ala Leu Thr Pro His
290 295 300
Pro Pro Gln Lys Gln Ala Tyr Gly Val Thr Leu Pro Thr Ser Val Leu
305 310 315 320
Phe Ile Ala Gly His Asp Thr Asn Leu Ala Asn Leu Gly Gly Ala Leu
325 330 335
Glu Leu Asn Trp Thr Leu Pro Gly Gln Pro Asp Asn Thr Pro Pro Gly
340 345 350
Gly Glu Leu Val Phe Glu Arg Trp Arg Arg Leu Ser Asp Asn Ser Gln
355 360 365
Trp Ile Gln Val Ser Leu Val Phe Gln Thr Leu Gln Gln Met Arg Asp
370 375 380
Lys Thr Pro Leu Ser Leu Asn Thr Pro Pro Gly Glu Val Lys Leu Thr
385 390 395 400
Leu Ala Gly Cys Glu Glu Arg Asn Ala Gln Gly Met Cys Ser Leu Ala
405 410 415
Gly Phe Thr Gln Ile Val Asn Glu Ala Arg Ile Pro Ala Cys Ser Leu
420 425 430

como divulgada no documento US2005/246780.

A sequência de fitase parental SEQ ID N°:2, codificada por, *e. g.*, SEQ ID N°:1, que mostra as modificações de sequência geradas por mutagénese de saturação de local génico (GSSM) seleccionados para construção da biblioteca GeneReassembly™ é mostrada na Figura 8.

Num aspecto, os polipéptidos e péptidos da invenção têm actividade de fitase. Em aspectos alternativos, também podem ser úteis como, *e. g.*, sondas de marcação, antigénios, toleragénios, motivos, sítios activos de fitase.

Em aspectos alternativos, os polipéptidos e péptidos da invenção são sintéticos ou são polipéptidos produzidos recombinantemente. Os péptidos e proteínas podem ser recombinantemente expressos *in vitro* ou *in vivo*. Os péptidos e polipéptidos da invenção podem ser feitos e isolados utilizando qualquer método conhecido na técnica. Os polipéptidos e péptidos da invenção podem ser também sintetizados, totalmente ou em parte, utilizando métodos químicos bem conhecidos na técnica. Ver, *e. g.*, Caruthers (1980) Nucleic Acids Res. Symp. Ser. 215-223; Horn (1980) Nucleic Acids Res. Symp. Ser. 225-232; Banga, A.K., Therapeutic Peptides and Proteins, Formulation, Processing and Delivery Systems (1995) Technomic Publishing Co., Lancaster, PA. Por exemplo, a síntese peptídica pode ser realizada utilizando várias técnicas de fase sólida (ver, *e. g.*, Roberge (1995) Science 269:202; Merrifield (1997) Methods Enzymol. 289:3□ 13) e síntese automatizada pode ser realizada, *e. g.*, utilizando o Sintetizador de Péptidos ABI 431A (Perkin Elmer) de acordo com as instruções fornecidas pelo fabricante.

As enzimas e polinucleótidos da presente invenção podem ser proporcionados numa forma isolada ou purificados até à homogeneidade. O polipéptido de fitase da invenção pode ser obtido utilizando qualquer um de diversos métodos convencionais. Por exemplo, os polipéptidos de fitase podem ser produzidos num sistema de expressão recombinante convencional (como aqui descrito), sintetizados quimicamente (embora um tanto limitados a pequenos fragmentos peptídicos de fitase) ou purificados a partir de organismos em que são naturalmente expressos. Os métodos de expressão recombinante úteis incluem hospedeiros mamíferos, hospedeiros microbianos e hospedeiros vegetais.

Em aspectos alternativos, os polipéptidos e péptidos da invenção compreendem "aminoácidos" ou "sequências de aminoácidos" que são oligopéptidos, péptidos, polipéptidos ou sequências de proteína, ou alternativamente, são fragmentos, porções ou subunidades de alguns destes, e moléculas sintéticas ou de ocorrência natural.

Em aspectos alternativos, os polipéptidos ou proteínas "recombinantes" da invenção incluem (referem-se) polipéptidos ou proteínas produzidas por técnicas de ADN recombinante; *e. g.*, produzidos a partir de células transformadas por uma construção de ADN exógeno que codifica o polipéptido ou proteína desejada. Os ácidos nucleicos (incluindo oligonucleótidos), polipéptidos ou proteínas "sintéticos" da invenção incluem aqueles preparados por síntese química, como aqui descrito em detalhe. Em aspectos alternativos, os polipéptidos ou proteínas da invenção compreendem aminoácidos ligados uns aos outros através de ligações peptídicas ou ligações peptídicas modificadas, *i. e.*, isoésteres peptídicos, e podem conter aminoácidos modificados diferentes dos 20 aminoácidos codificados geneticamente. Os polipéptidos podem ser modificados através de processos

naturais, tal como processamento pós-tradução, ou por técnicas de modificação química que são bem conhecidas no campo técnico. As modificações podem ocorrer em qualquer lugar no polipéptido, incluindo na estrutura peptídica, nas cadeias laterais dos aminoácidos e nos terminais amino ou carboxilo. Deve ser entendido que o mesmo tipo de modificação pode estar presente no mesmo ou em graus variáveis em diversos locais num dado polipéptido. Também um dado polipéptido pode ter muitos tipos de modificações, por exemplo, acetilação, acilação, ADP-ribosilação, amidação, ligação covalente de flavina, ligação covalente de uma unidade hémica, ligação covalente de um nucleótido ou derivado de nucleótido, ligação covalente de um lípido ou derivado de lípido, ligação covalente de um fosfatidilinositol, ciclização por reticulação, formação de ligação persulfureto, desmetilação, formação de reticulações covalentes, formação de cisteína, formação de piroglutamato, formilação, gama-carboxilação, glicosilação, formação de âncora de GPI, hidroxilação, iodinação, metilação, miristoilação, oxidação, pegilação, processamento proteolítico, fosforilação, prenilação, racemização, selenoilação, sulfatação e adição mediada por transferência de ARN de amino ácidos para proteína, tal como arginilação.

Em aspectos alternativos, os polipéptidos ou as proteínas “sintéticas” são aquelas preparadas por síntese química. Os métodos de síntese peptídica química de fase sólida podem também ser utilizados para sintetizar o polipéptido ou fragmentos da invenção. Tal método é conhecido na técnica desde o início dos anos de 1960 (Merrifield, R. B., J. Am. Chem. Soc., 85:2149-2154, 1963) (ver também Stewart, J. M. and Young, J. D., Solid Phase Peptide Synthesis, 2ª ed., Pierce Chemical Co., Rockford, III., pp. 11-12)) e têm sido recentemente utilizados em kits de desenho e síntese peptídica laboratorial comercialmente

disponíveis (Cambridge Research Biochemicals). Tais kits laboratoriais comercialmente disponíveis utilizam habitualmente os ensinamentos de H. M. Geysen *et al.* Proc. Natl. Acad. Sci., USA, 81:3998 (1984) e proporcionam a síntese de péptidos nas pontas de um multitude de "varas" ou "pinos" todos os quais estão conectados a uma única placa. Quando tal sistema é utilizado, uma placa de varas ou pinos é invertida e introduzida numa segunda placa de poços ou reservatórios correspondentes, que contêm soluções para ligar ou ancorar um aminoácido apropriado às pontas das varas ou pinos. Ao repetir um tal passo de processo, *i. e.*, invertendo e introduzindo as pontas das varas e pinos em soluções apropriadas, os aminoácidos são construídos em péptidos desejados. Além disso, um número de sistemas de síntese peptídica de FMOC está disponível. Por exemplo, a montagem de um polipéptido ou fragmento pode ser realizada num suporte sólido utilizando um sintetizador de péptidos automatizado modelo 431A da Applied Biosystemas, Inc. Tal equipamento proporciona pronto acesso aos péptidos da invenção, por síntese directa ou por síntese de uma série de fragmentos que podem ser acoplados utilizando outras técnicas conhecidas.

Em aspectos alternativos, os péptidos e polipéptidos da invenção são glicosilados. A glicosilação pode ser adicionada pós-tradução quimicamente ou por mecanismos biossintéticos celulares, em que o último incorpora a utilização de motivos de glicosilação conhecidos, os quais podem ser nativos à sequência ou podem ser adicionados como um péptido ou adicionados na sequência codificante de ácidos nucleicos. A glicosilação pode ser ligada por O ou ligada por N, ou uma sua combinação.

Em aspectos alternativos, os péptidos e polipéptidos da invenção, como acima definidos, compreendem formas "miméticas" e

“peptidomiméticas”, em parte ou na totalidade. Num aspecto, os termos “mimético” e “peptidomimético” refere-se a um composto químico sintético que tem substancialmente as mesmas características estruturais e/ou funcionais dos polipéptidos da invenção. O mimético pode ser composto totalmente de análogos sintéticos, não naturais de aminoácidos ou é uma molécula quimérica de aminoácidos parcialmente naturais e de análogos parcialmente não naturais de aminoácidos. O mimético pode também incorporar qualquer quantidade de substituições conservadoras de aminoácidos naturais, desde que tais substituições também não alteram substancialmente a estrutura e/ou actividade do mimético. Como com os polipéptidos da invenção que são variantes conservadores, a experimentação de rotina determinará se um mimético está dentro do âmbito da invenção, *i. e.*, que a sua estrutura e/ou função não está substancialmente alterada. Assim, num aspecto, uma composição mimética está dentro do âmbito da invenção se tiver uma actividade de fitase.

Em aspectos alternativos, os péptidos e polipéptidos da invenção têm sequências compreendendo a modificação específica para a SEQ ID N°:2, como definida acima, e também substituições conservadoras que podem ou não modificar a actividade, *e. g.*, actividade enzimática. Em aspectos alternativos, as substituições conservadoras são aquelas que substituem um dado aminoácido num polipéptido por outro aminoácido de características semelhantes. Em aspectos alternativos, as substituições conservadoras são as seguintes substituições: substituições de um aminoácido alifático, tais como Ala, Val, Leu e Ile com um outro aminoácido alifático; substituição de um Ser com uma Thr ou vice-versa; substituição de um resíduo acídico, tais como Asp e Glu com outro resíduo acídico; substituição de um resíduo que contém um grupo amida, tais como Asn e Gln, com outro resíduo que contém um grupo amida; troca de

um resíduo básico, tais como Lys e Arg com outro resíduo básico; e substituição de um resíduo aromático, tais como Phe, Tyr com outro resíduo aromático.

As composições miméticas de polipéptido da invenção podem conter qualquer combinação de componentes estruturais não naturais. Num aspecto alternativo, as composições miméticas da invenção incluem um ou todos os seguintes três grupos estruturais: a) grupos de ligação de resíduo à excepção das ligações de amida natural ("ligação peptídica"); b) resíduos não naturais no lugar de resíduos de aminoácidos de ocorrência natural; ou c) resíduos que induzem mimetização estrutural secundária, *i. e.*, para induzir ou estabilizar uma estrutura secundária, *e. g.*, uma conformação em volta beta, volta gama, folha beta, hélice alfa e semelhantes. Por exemplo, um polipéptido da invenção pode ser caracterizado como um mimético quando todos ou alguns dos seus resíduos são ligados por meios químicos à excepção de ligações peptídicas naturais. Os resíduos peptidomiméticos individuais podem ser ligados por ligações peptídicas, outras ligações químicas ou meios de acoplamento, tais como, *e. g.*, glutaraldeído, ésteres de N-hidroxi-succinimida, maleimidas bifuncionais, N,N'-diciclo-hexil-carbodiimida (DCC) ou N,N'-diisopropilcarbodiimida (DIC). Os grupos de ligação que podem ser uma alternativa às ligações tradicionais de amida ("ligação peptídica") incluem, *e. g.*, cetometileno (*e. g.*, -C(=O)-CH2- para -C(=O)-NH-), aminometileno (CH2-NH), etileno, olefina (CH=CH), éter (CH2-O), tioéter (CH2-S), tetrazole (CN4-), tiazole, retroamida, tioamida ou éster (ver, *e. g.*, Spatola (1983) em Chemistry and Biochemistry of Amino Acids, Peptides and Proteins, Vol. 7, pp 267-357, "Peptide Backbone Modifications", Marcell Dekker, NI).

Um polipéptido da invenção pode também ser caracterizado como um mimético por conter todos ou alguns resíduos não naturais no lugar dos resíduos de aminoácidos de ocorrência natural. Os resíduos não naturais são bem descritos na literatura científica e de patente; algumas composições não naturais exemplificativas úteis como miméticas de resíduos de aminoácidos naturais e orientações são descritas abaixo. Os miméticos de aminoácidos aromáticos podem ser gerados por substituição por, *e. g.*, D- ou L- nafilalanina; D- ou L-fenilglicina; D- ou tieneilalanina L-2; D- ou L-1, -2, -3 ou 4-pireneilalanina; D ou L-3 tieneilalanina; D- ou L-(2-piridinil)-alanina; D- ou L-(3-piridinil)-alanina; D- ou L-(2-pirazinil)-alanina; D- ou L-(4-isopropil)-fenilglicina; D-(trifluorometil)-fenilglicina; D-(trifluorometil)-fenilalanina; D-fluoro-fenilalanina; D- ou L-p-bifenilfenilalanina; K ou L-p-metoxi-bifenilfenilalanina; D- ou L-2-indole(alquil)alaninas; e D- ou L-alquilalaninas, em que alquil pode ser metilo, etilo, propilo, hexil, butilo, pentilo, isopropilo, iso-butilo, sec-isotilo, isopentilo substituídos ou não substituídos, ou um aminoácido não acídico. Os anéis aromáticos de um aminoácido não natural incluem, *e. g.*, anéis aromáticos de tiazolilo, tiofenilo, pirazolilo, benzimidazolilo, naftilo, furanilo, pirrolilo e piridilo.

Os miméticos de aminoácidos acídicos podem ser gerados através de substituição por, *e. g.*, aminoácidos não de carboxilato enquanto é mantida uma carga negativa; (fosfono)alanina; treonina sulfatada. Os grupos laterais de carboxilo (*e. g.*, aspartilo ou glutamilo) podem também ser seletivamente modificados por reacção com carbodiimidas (R'-N-C-N-R'), tais como, *e. g.*, 1-ciclo-hexil-3(2-morfolinil-(4-etil)-carbodiimida ou 1-etil-3(4-azonia-4,4-dimetolpentil)-carbodiimida. Aspartilo ou glutamilo pode também ser convertido

resíduos de asparaginilo e glutaminilo por reacção com iões de amónio. Os miméticos de aminoácidos básicos podem ser produzidos através de substituição com, *e. g.*, (além de lisina e arginina) os aminoácidos ornitina, citrulina ou ácido (guanidino)-acético ou ácido (guanidino)alquil-acético, em que alquilo é definido acima. Os derivados de nitrilo (*e. g.*, contendo a unidade CN no lugar de COOH) podem ser substituídos para asparagina ou glutamina. Os resíduos de asparaginilo e glutaminilo podem ser desaminados aos resíduos correspondentes de aspartilo ou glutamilo. Os miméticos de resíduos de arginina podem ser gerados por reacção de arginilo com, *e. g.*, um ou mais reagentes convencionais, incluindo, *e. g.*, fenilglioxal, 2,3-butanodiona, 1,2-ciclo-hexanodiona ou ninidrina, que para estes reagentes pode ser preferido utilizar condições alcalinas. Os miméticos de resíduos de tirosina podem ser produzidos por reacção de tirosilo com, *e. g.*, compostos aromáticos de diazónio ou tetranitrometano. N-acetilimidizole e tetranitrometano podem ser utilizados para formar espécies de O-acetil-tirosilo e derivados 3-nitro, respectivamente. Os miméticos de resíduos de cisteína podem ser produzidos por reacção de resíduos de cisteinilo com, *e. g.*, alfa-haloacetatos, tais como ácido 2-cloroacético ou cloroacetamida e correspondentes aminas; para produzir derivados de carboximetilo ou carboxiamidometilo. Os miméticos de resíduos de cisteína podem também ser produzidos por reacção de resíduos de cisteinilo com, *e. g.*, bromo-trifluoroacetona, ácido alfa-bromo-beta-(5-imidozoil)propiónico; fosfato de cloroacetilo, N-alquil-maleimidas, dissulfureto de 3-nitro-2-piridilo; dissulfureto de metil-2-piridilo; p-cloromercuribenzoato; 2-cloromercuri-4-nitrofenol; ou cloro-7-nitrobenzo-oxa-1,3-diazole. Os miméticos de lisina podem ser produzidos (e os resíduos amino terminais podem ser alterados) por reacção de lisinilo com, *e. g.*, succínico ou outros anidridos de ácido carboxílico. A lisina e outros miméticos de resíduos contendo

alfa-amino podem também ser gerados por reacção com imidoésteres, tais como metilpicolinimidato, fosfato de piridoxal, piridoxal, cloroboro-hidreto, ácido trinitro-benzenossulfónico, O-metilisoureia, 2,4-pentanodiona e reacções catalizadas por transamidase com glioxilato. Os miméticos de metionina podem ser gerados por reacção com, *e. g.*, sulfóxido de metionina. Os miméticos de prolina incluem, *e. g.*, ácido pipecólico, ácido tiazolidina-carboxílico, 3- ou 4-hidroxiprolina, desidroprolina, 3- ou 4-metilprolina ou 3,3,dimetilprolina. Os miméticos de resíduos de histidina podem ser gerados por reacção de histidilo com, *e. g.*, dietilprocarbonato ou brometo de para-bromofenacilo. Outros miméticos incluem, *e. g.*, aqueles produzidos por hidroxilação da prolina e lisina; fosforilação dos grupos hidroxilo de resíduos de serilo ou treonilo; metilação dos grupos alfa-amino de lisina, arginina e histidina; acetilação da amina N-terminal; metilação de resíduos de amida da cadeia principal ou substituição com N-metil-aminoácidos; ou amidação de grupos carboxilo do C-terminal.

Um resíduo, *e. g.*, um aminoácido, de um polipéptido da invenção pode também ser substituído por um aminoácido (ou resíduo peptidomimético) de quiralidade oposta. Assim, qualquer aminoácido de ocorrência natural na L-configuração (que pode também ser referido como R ou S, dependendo da estrutura da entidade química) pode ser substituído com o aminoácido do mesmo tipo de estrutura química ou um peptidomimético, mas da quiralidade oposta, referido como o aminoácido D, mas também pode ser referida como a forma R ou S.

A invenção também proporciona métodos para modificar os polipéptidos da invenção através de processos naturais, tal como processamento pós-tradução (*e. g.*, fosforilação, acilação,

etc.), ou técnicas de modificação química, e os polipéptidos modificados resultantes. As modificações podem ocorrer em qualquer lugar no polipéptido, incluindo a estrutura peptídica, as cadeias laterais dos aminoácidos e os terminais amino ou carboxilo. Deve ser entendido que o mesmo tipo de modificação pode estar presente no mesmo ou vários graus em diversos locais num dado polipéptido. Também, um dado polipéptido pode ter muitos tipos de modificações. As modificações incluem acetilação, acilação, ADP-ribosilação, amidação, ligação covalente de flavina, ligação covalente de uma unidade hémica, ligação covalente de um nucleótido ou derivado de nucleótido, ligação covalente de um lípido ou derivados de lípido, ligação covalente de um fosfatidilinositol, ciclização por reticulação, formação de ligação dissulfureto, desmetilação, formação de reticulações covalentes, formação de cisteína, formação de piroglutamato, formilação, gama-carboxilação, glicosilação, formação de âncora de GPI, hidroxilação, iodinação, metilação, miristoilação, oxidação, pegilação, processamento proteolítico, fosforilação, prenilação, racemização, selenoilação, sulfatação e adição mediada por transferência de ARN de amino ácidos para proteína, tal como arginilação. Ver, *e. g.*, Creighton, T.E., Proteins - Structure and Molecular Properties 2ª Ed., W.H. Freeman and Company, Nova Iorque (1993); Posttranslational Covalent Modification of Proteins, B.C. Johnson, Ed., Academic Press, Nova Iorque, pp. 1-12 (1983).

Os métodos de síntese peptídica química de fase sólida podem também ser utilizados para sintetizar o polipéptido ou fragmentos da invenção. Tal método é conhecido na técnica desde o início dos anos de 1960 (Merrifield, R. B., J. Am. Chem. Soc., 85:2149-2154, 1963) (ver também Stewart, J. M. e Young, J. D., Solid Phase Peptide Synthesis, 2ª ed., Pierce Chemical Co., Rockford, III., pp. 11-12)) e têm sido recentemente utilizados

em kits de concepção e síntese peptídica laboratorial comercialmente disponíveis (Cambridge Research Biochemicals). Tais kits laboratoriais comercialmente disponíveis utilizam habitualmente os ensinamentos de H. M. Geysen *et al.*, Proc. Natl. Acad. Sci., USA, 81:3998 (1984) e proporcionam a síntese de péptidos nas pontas de um multitude de "varas" ou "pinos" todos os quais estão conectados a uma única placa. Quando tal sistema é utilizado, uma placa de varas ou pinos é invertida e introduzida numa segunda placa de poços ou reservatórios correspondentes, que contêm soluções para ligar ou ancorar um aminoácido apropriado às pontas das varas ou pinos. Ao repetir um tal passo de processo, *i. e.*, invertendo e introduzindo as pontas das varas e pinos em soluções apropriadas, os aminoácidos são construídos em péptidos desejados. Além disso, um número de sistemas de síntese peptídica de FMOC está disponível. Por exemplo, a montagem de um polipéptido ou fragmento pode ser realizada num suporte sólido utilizando um sintetizador de péptidos automatizado modelo 431A da Applied Biosystemas, Inc. Tal equipamento proporciona pronto acesso aos péptidos da invenção, por síntese directa ou por síntese de uma série de fragmentos que podem ser acoplados utilizando outras técnicas conhecidas.

Num aspecto, os péptidos e polipéptidos da invenção têm sequências compreendendo a modificação específica para a SEQ ID N°:2, como definida acima, e também sequências de aminoácidos "substancialmente idênticas", *i. e.*, uma sequência que difere por uma ou mais substituições, deleções ou inserções de aminoácidos conservadoras ou não conservadoras, particularmente quando tal substituição ocorre num local que não é o sítio activo da molécula e desde que o polipéptido retenha essencialmente as suas propriedades funcionais. Num aspecto, os péptidos e polipéptidos da invenção têm sequências compreendendo

a modificação específica para a SEQ ID N°:2, como definida acima, e substituições de aminoácidos conservadoras que substituem um aminoácido por outra da mesma classe, por exemplo, substituição de um aminoácido hidrofóbico, tais como isoleucina, valina, leucina ou metionina, por outro, ou substituição de um aminoácido polar por outro, tais como substituição de arginina para lisina, ácido glutâmico para ácido aspártico ou glutamina para asparagina. Num aspecto, um ou mais dos aminoácidos podem ser removidos, por exemplo, a partir de um polipéptido de fitase da invenção para resultar em modificação da estrutura do polipéptido sem alterar significativamente a sua actividade biológica ou, alternativamente, para alterar propositadamente de forma significativa a sua actividade biológica. Por exemplo, aminoácidos de terminal amino ou carboxilo que são requeridos ou alternativamente não são requeridos para a actividade biológica de fitase podem ser removidos e/ou adicionados. As sequências de polipéptido modificadas da invenção podem ser ensaiadas para a actividade biológica de fitase através de qualquer variedade de métodos, incluindo colocar em contacto a sequência de polipéptido modificada com um substrato de fitase e determinar se o polipéptido modificado diminui a quantidade de substrato específico no ensaio ou aumenta os subprodutos da reacção enzimática de um polipéptido funcional de fitase com o substrato.

Um outro aspecto da invenção compreende polipéptidos que têm cerca de 95%, 96%, 97%, 98%, 99% ou mais de identidade da sequência para a SEQ ID N°:2 e têm as modificações de sequência enumeradas específicas, como acima (apresentado) discutido.

Estas variantes de sequência de aminoácidos da invenção podem ser caracterizadas por uma natureza predeterminada da variação, uma característica que os separada de uma forma de

ocorrência natural, *e. g.*, uma variação alélica ou interespécies de uma sequência de fitase. Num aspecto, as variantes da invenção apresentam a mesma actividade biológica qualitativa que o análogo de ocorrência natural. Alternativamente, as variantes podem ser seleccionadas para terem características modificadas. Num aspecto, embora o local ou região para introduzir uma variação de sequência de aminoácido seja predeterminado, a mutação não necessita por si mesmo ser predeterminada. Por exemplo, de modo a optimizar o desempenho de uma mutação num dado local, pode ser conduzida mutagénese aleatória no codão ou região alvo e as variantes de fitase expressas rastreadas para a combinação óptima da actividade desejada. As técnicas para fazer mutações de substituição em locais predeterminados em ADN que tem uma sequência conhecida, são bem conhecidas, como aqui discutido, por exemplo, mutagénese do iniciador M13 e mutagénese por PCR. O rastreio dos mutantes pode ser feito utilizando, *e. g.*, ensaios de catálise de fitato (mio-inositol-hexafosfato) em inositol e fosfato inorgânico; ou a hidrólise de fitato (mio-inositol-hexafosfato). Em aspectos alternativos, as substituições de aminoácido podem ser em resíduos únicos; as inserções podem ser na ordem desde cerca de 1 a 20 aminoácidos, embora possam ser feitas inserções consideravelmente maiores. As deleções podem variar desde cerca de 1 a cerca de 20, 30, 40, 50, 60, 70 resíduos ou mais. Para obter um derivado final com as propriedades óptimas, podem ser utilizadas substituições, deleções, inserções ou qualquer sua combinação. Geralmente, estas alterações são feitas em poucos aminoácidos para minimizar a alteração da molécula. Contudo, alterações maiores podem ser toleradas em determinadas circunstâncias.

Os polipéptidos da invenção podem ser obtidos através de processos de enriquecimento ou purificação bioquímica. A sequência de potenciais polipéptidos homólogos ou fragmentos

pode ser determinada por digestão proteolítica, electroforese em gel e/ou microssequenciação.

Um outro aspecto da divulgação é um ensaio para identificar fragmentos ou variantes dos polipéptidos da invenção. Os polipéptidos da invenção podem ser utilizados para catalisar reacções bioquímicas para indicar que o referido fragmento ou variante retém actividade enzimática de polipéptidos da invenção.

Um ensaio exemplificativo para determinar se os fragmentos de variantes retêm a actividade enzimática dos polipéptidos da invenção compreende os passos de: colocar em contacto o fragmento ou variante de polipéptido com uma molécula de substrato sob condições que permitem que o fragmento ou variante de polipéptido funcione, e detectar uma diminuição no nível de substrato ou um aumento no nível do produto da reacção específico da reacção entre o polipéptido e o substrato.

Os polipéptidos da invenção podem ser utilizados para catalisar reacções bioquímicas. De acordo com um aspecto da invenção, é proporcionado um processo para utilizar um polipéptido da invenção como uma fitase.

A invenção proporciona fitases não tendo sequências de sinal (também denominados péptidos de sinal (SP) ou péptidos líder) ou estão modificadas, ou sequências de sinal heterólogas. Os polipéptidos da invenção podem também não ter domínios prepro modificados ou heterólogos ou estão modificados, e/ou domínios catalíticos (CD). Os SP modificados ou heterólogos, domínios prepro e/ou CD incorporados num polipéptido da invenção podem ser parte de uma proteína de fusão, *e. g.*, como um domínio heterólogo numa proteína quimérica, ou adicionados por um agente

de ligação química. Por exemplo, uma enzima da invenção pode compreender um SP heterólogo e/ou prepro num vector, *e. g.*, um vector da série pPIC (Invitrogen, Carlsbad, CA).

Adicionalmente, os polipéptidos da invenção podem ainda compreender sequências heterólogas, sequências de outras fitases ou de fontes não de fitase, ou sequências inteiramente sintéticas. Assim, num aspecto, um ácido nucleico da invenção compreende uma sequência codificante para uma sequência de sinal (SP) endógena, modificada ou heterólogo, domínio prepro e/ou domínio catalítico (CD) e uma sequência heteróloga (*i. e.*, uma sequência associada não naturalmente com uma sequência de sinal (SP), domínio prepro e/ou domínio catalítico (CD) da invenção). A sequência heteróloga pode estar na extremidade terminal 3', extremidade terminal 5' e/ou em ambas da SP, domínio prepro e/ou sequência codificante do CD.

Os métodos para identificar as sequências do domínio "prepro" e as sequências de sinal são bem conhecidas na técnica, ver, *e. g.*, van de Ven (1993) Crit. Rev. Oncog. 4(2):115-136. Por exemplo, para identificar uma sequência prepro, a proteína é purificado a partir do espaço extracelular e a sequência proteica do N-terminal é determinada e comparada à forma não processada. Vários métodos de reconhecimento de sequência de sinal são conhecidos dos especialistas na técnica. Por exemplo, num aspecto, os péptidos de sinal para utilização com os polipéptidos da invenção são identificados por um método referido como SignalP. O SignalP utiliza uma rede neuronal combinada que reconhece os péptidos de sinal e os seus locais de clivagem; ver, *e. g.*, Nielsen (1997) "Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites" Protein Engineering 10:1-6.

A invenção proporciona enzimas fitase em que a estrutura do esqueleto polipeptídico, estrutura secundária ou terciária, *e. g.*, uma estrutura de alfa-helicoidal ou folha beta, foram modificados. Num aspecto, a carga ou a hidrofobicidade foram modificadas. Num aspecto, o volume de uma cadeia lateral foi modificado. Alterações substanciais na função ou identidade imunológica são feitas através de selecção de substituições que são menos conservadoras. Por exemplo, podem ser feitas substituições que afectam mais significativamente: a estrutura do esqueleto polipeptídico na área da alteração, por exemplo, uma estrutura alfa-helicoidal ou folha beta; uma carga ou um local hidrofóbico da molécula que pode estar num sítio activo; ou uma cadeia lateral. A invenção proporciona substituições no polipéptido da invenção em que (a) um resíduo hidrofílico, *e. g.*, serilo ou treonilo, é substituído para (ou por) um resíduo hidrofóbico, *e. g.*, leucilo, isoleucilo, fenilalanilo, valilo ou alanilo; (b) uma cisteína ou uma prolina são substituídas para (ou por) qualquer outro resíduo; (c) um resíduo que tem uma cadeia lateral electropositivo, *e. g.*, lisilo, arginilo ou histidilo é substituído para (ou por) um resíduo electronegativo, *e. g.*, glutamilo ou aspartilo; ou (d) um resíduo que tem uma cadeia lateral volumosa, *e. g.*, fenilalanina é substituído para (ou por) um não tendo uma cadeia lateral, *e. g.* glicina. As variantes podem apresentar a mesma actividade biológica qualitativa (*i. e.*, actividade enzimática de fitase) embora as variantes possam ser seleccionadas para modificar as características da fitase como necessário.

Num aspecto, as fitases da invenção compreendem epitopos ou marcadores de purificação, sequências de sinal ou outras sequências de fusão, etc. Num aspecto, as fitases da invenção podem ser fundidas a um péptido aleatório para formar um polipéptido de fusão. Por “fundido” ou “ligado operativamente”

entende-se aqui que o péptido aleatório e a fitase estão ligados em conjunto , de modo a minimizar a perturbação da estabilidade da actividade de fitase. O polipéptido de fusão (ou polinucleótido de fusão que codifica o polipéptido de fusão) pode também compreender mais componentes, incluindo múltiplos péptidos em múltiplas ansas.

Num aspecto, as fitases da invenção são polipéptidos quiméricos, *e. g.*, compreendendo SP heterólogos, moléculas de ligação a hidratos de carbono, domínios catalíticos da enzima fitase, ligantes e/ou domínios catalíticos não de fitase. A invenção proporciona meios para gerar polipéptidos quiméricos que podem codificar polipéptidos híbridos biologicamente activos (*e. g.*, enzimas híbridas de fitase). Num aspecto, os polinucleótidos originais codificam polipéptidos biologicamente activos. O método da invenção produz novos polipéptidos híbridos que utilizam através da utilização de processos celulares que integram a sequência dos polinucleótidos originais de modo que o polinucleótido híbrido resultante codifique um polipéptido que demonstra actividades derivadas dos polipéptidos biologicamente activos originais. Por exemplo, os polinucleótidos originais podem codificar uma enzima particular de diferentes microrganismos. Uma enzima codificada por um primeiro polinucleótido de um organismo ou variante pode, por exemplo, funcionar eficazmente sob um condição ambiental particular, *e. g.*, elevada salinidade. Uma enzima codificada por um segundo polinucleótido de um organismo ou variante diferente pode funcionar eficazmente sob uma condição ambiental diferente, tal como temperaturas extremamente altas. Um polinucleótido híbrido contendo sequências dos primeiros e segundos polinucleótidos originais pode codificar uma enzima que apresenta características de ambas as enzimas codificadas pelos polinucleótidos originais. Assim, a enzima codificada pelo

polinucleótido híbrido pode funcionar eficazmente sob condições ambientais partilhadas por cada uma das enzimas codificadas pelos primeiros e segundos polinucleótidos, *e. g.*, elevada salinidade e temperaturas extremas.

Assim, um polipéptido híbrido que resulta deste método da invenção pode apresentar actividade enzimática especializada não apresentada nas enzimas originais. Por exemplo, após recombinação e/ou rearranjo redutor de polinucleótidos que codificam a enzima fitase, o polipéptido híbrido resultante codificado por um polinucleótido híbrido pode ser rastreado para actividades enzimáticas especializadas, *e. g.*, actividades de hidrolase, peptidase, fosforilase, etc., obtidas de cada uma das enzimas originais. Assim, por exemplo, o polipéptido híbrido pode ser rastreado para verificar essas funcionalidades químicas que distinguem o polipéptido híbrido dos polipéptidos parentais originais, tais como temperatura, pH ou concentração de sal, às quais o polipéptido híbrido funciona.

Um polipéptido híbrido que resulta do método da invenção pode apresentar actividade enzimática especializada não apresentada nas enzimas originais. Por exemplo, após recombinação e/ou rearranjo redutor de polinucleótidos que codificam actividades de hidrolase, o polipéptido híbrido resultante codificado por um polinucleótido híbrido pode ser rastreado para actividades de hidrolase especializadas obtidas de cada uma das enzimas originais, *i. e.*, o tipo de ligação sobre a qual a hidrolase actua e a temperatura em que a hidrólase funciona. Assim, por exemplo, uma fitase pode ser rastreada para verificar essas funcionalidades químicas que distinguem a fitase híbrida das fitases originais, tais como: (a) amida (ligações peptídicas), *i. e.*, proteases; (b) ligações éster, *i. e.*, esterases e lipases; (c) acetais, *i. e.*,

glicosidases e, por exemplo, a temperatura, pH ou concentração de sal em que o polipéptido híbrido funciona.

Num aspecto, na invenção refere-se a um método para produzir um polipéptido híbrido biologicamente activo e rastrear tal polipéptido para actividade aumentada através de:

(1) introduzir, pelo menos, um primeiro polinucleótido em ligação operável e um segundo polinucleótido em ligação operável, os referidos, pelo menos um, primeiro polinucleótido e segundo polinucleótido partilham, pelo menos, uma região de homologia de sequência parcial, numa célula hospedeira adequada;
(2) cultivar a célula hospedeira sob condições que promovem a reorganização de sequências resultando num polinucleótido híbrido em ligação operável;
(3) expressar um polipéptido híbrido codificado pelo polinucleótido híbrido;
(4) rastrear o polipéptido híbrido sob condições que promovem a identificação da actividade biológica aumentada; e
(5) isolar um polinucleótido codificando o polipéptido híbrido.

Os métodos para rastrear para várias actividades enzimáticas são conhecidos dos especialistas na técnica e são discutidos ao longo da presente descrição. Tais métodos podem ser utilizados ao isolar os polipéptidos e polinucleótidos da invenção.

Num aspecto, a presente invenção proporciona um método (e seus produtos) para produzir formulação líquidas aquosas estabilizadas tendo actividade de fitase que apresentam resistência aumentada à inactivação térmica da actividade

enzimática e as quais retêm a sua actividade de fitase durante períodos prolongados de armazenamento. As formulações líquidas são estabilizadas por meio da adição de ureia e/ou um poliol, tais como sorbitol e glicerol como agente estabilizador. São também proporcionadas preparações de ração para animais monogástricos e métodos para a sua produção que resultam da utilização de tais formulações líquidas aquosas estabilizadas. Detalhes adicionais a respeito desta abordagem estão na literatura publicada e/ou são conhecidos do especialista na técnica. Numa exemplificação particular não limitativo, tal literatura publicamente disponível inclui o documento EP 0626010 (documento WO 9316175 A1) (Barendse *et al.*), embora as referências na literatura publicamente disponível não ensinem as moléculas inventivas do presente pedido.

Anticorpos e métodos de rastreio baseados em anticorpos

Anticorpos isolados, sintéticos ou recombinantes podem-se ligar especificamente a uma fitase da invenção. Estes anticorpos podem ser utilizados para isolar, identificar ou quantificar as fitases da invenção ou polipéptidos relacionados. Estes anticorpos podem ser utilizados para inibir a actividade de uma enzima da invenção. Estes anticorpos podem ser utilizados para isolar polipéptidos relacionados com aqueles da invenção, *e. g.*, enzimas fitase relacionadas.

Os anticorpos podem compreender um péptido ou polipéptido derivado, modelado com base ou substancialmente codificado por um gene de imunoglobulina ou genes de imunoglobulinas, ou os seus fragmentos, capazes de se ligarem especificamente a um antigénio ou epitopo, ver, *e. g.*, Fundamental Immunology, Terceira Edição, W.E. Paul, ed., Raven Press, N.I.

(1993); Wilson (1994) J. Immunol. Methods 175:267-273; Yarmush (1992) J. Biochem. Biophys. Methods 25:85-97. O termo anticorpo inclui porções de ligação a antigénio, *i. e.*, "locais de ligação a antigénio" (*e. g.*, fragmentos, sub-sequências, regiões determinantes de complementaridade (CDR)) que retêm a capacidade de ligação ao antigénio, incluindo (i) um fragmento Fab, um fragmento monovalente consistindo nos domínios VL, VH, CL e CH1; (ii) um fragmento F(ab')2, um fragmento bivalente compreendendo dois fragmentos Fab ligados por uma ligação persulfureto na região de charneira; (iii) um fragmento Fd consistindo nos domínios VH e CH1; (iv) um fragmento Fv consistindo nos domínios VL e VH de um único braço de um anticorpo, (v) um fragmento dAb (Ward *et al.*, (1989) Nature 341:544-546), o qual consiste num domínio VH; e (vi) uma região determinante de complementaridade (CDR) isolada. Os anticorpos de cadeia única estão também incluídos por referência no termo "anticorpo".

Os anticorpos podem ser utilizados em imunoprecipitação, coloração (*e. g.*, FACS), colunas de imunoafinidade e semelhantes. Se desejado, as sequências de ácidos nucleicos codificando para antigénios específicos podem ser produzidas por imunização, seguidas por isolamento do polipéptido ou ácido nucleico, amplificação ou clonagem e imobilização do polipéptido numa matriz da invenção. Alternativamente, podem ser utilizados métodos para modificar a estrutura de um anticorpo produzido por uma célula para ser modificado, *e. g.*, a afinidade de um anticorpo pode ser aumentada ou diminuída. Além disso, a capacidade para fazer ou modificar anticorpos pode ser um fenótipo manipulado numa célula através dos métodos da invenção.

Os métodos de imunização, produção e isolamento de anticorpos (policlonais e monoclonais) são conhecidos dos especialistas na técnica e descritos na literatura científica e

de patente, ver, *e. g.*, Coligan, CURRENT PROTOCOLS IN IMMUNOLOGY, Wiley/Greene, NI (1991); Stites (eds.) BASIC AND CLINICAL IMMUNOLOGY (7th ed.) Lange Medical Publications, Los Altos, CA ("Stites"); Goding, MONOCLONAL ANTIBODIES: PRINCIPLES AND PRACTICE (2d ed.) Academic Press, Nova Iorque, NI (1986); Kohler (1975) Nature 256:495; Harlow (1988) ANTIBODIES, A LABORATORY MANUAL, Cold Spring Harbor Publications, Nova Iorque. Os anticorpos podem também ser produzidos *in vitro*, *e. g.*, utilizando bibliotecas de apresentação de fagos expressando locais de ligação de anticorpo recombinante, além dos métodos tradicionais *in vivo* utilizando animais. Ver, *e. g.*, Hoogenboom (1997) Trends Biotechnol. 15:62-70; Katz (1997) Annu. Rev. Biophys. Biomol. Struct. 26:27-45.

Os polipéptidos podem ser utilizados para produzir anticorpos que se ligam especificamente aos polipéptidos da invenção. Os anticorpos resultantes podem ser utilizados em processos de cromatografia de imunoafinidade para isolar ou purificar o polipéptido ou para determinar se o polipéptido está presente numa amostra biológica. Em tais processos, uma preparação de proteína, tais como um extracto ou uma amostra biológica é colocada em contacto com um anticorpo capaz de se ligar especificamente a um dos polipéptidos da invenção.

Em processos de imunoafinidade, o anticorpo é ligado a um suporte sólido, tais como um grânulo ou outra matriz de coluna. A preparação de proteína é colocada em contacto com o anticorpo sob condições em que o anticorpo se liga especificamente a um dos polipéptidos da invenção. Após uma lavagem para remover as proteínas ligadas não especificamente, os polipéptidos ligados especificamente são eluídos.

A capacidade de proteínas numa amostra biológica de se ligarem ao anticorpo pode ser determinada utilizando qualquer um de uma variedade de processos familiares dos especialistas na técnica. Por exemplo, a ligação pode ser determinada por marcação do anticorpo com um marcador detectável, tais como um agente fluorescente, um marcador enzimático ou um radioisótopo. Alternativamente, a ligação do anticorpo à amostra pode ser detectada utilizando um anticorpo secundário tendo aí um marcador detectável. Os ensaios particulares incluem ensaios de ELISA, ensaios de sanduíche, radioimunoensaios, Transferências de Western.

Os anticorpos policlonais gerados contra os polipéptidos da invenção podem ser obtidos por injecção directa dos polipéptidos num animal ou por administração dos polipéptidos a um animal, por exemplo, um não humano. O anticorpo assim obtido irá depois ligar-se ao próprio polipéptido. Deste modo, mesmo uma sequência codificando apenas um fragmento do polipéptido pode ser utilizada para produzir anticorpos que se podem ligar à totalidade do polipéptido nativo. Tais anticorpos podem depois ser utilizados para isolar o polipéptido de células expressando esse polipéptido.

Para a preparação de anticorpos monoclonais, pode ser utilizada qualquer técnica que proporcione anticorpos produzidos por culturas contínuas de linhas celulares. Os exemplos incluem a técnica do hibridoma, a técnica do trioma, a técnica do hibridoma de célula B humana e a técnica de EBV-hibridoma (ver, *e. g.*, Cole (1985) em Monoclonal Antibodies and Cancer Therapy, Alan R. Liss, Inc., pp. 77-96).

As técnicas descritas para a produção de anticorpos de cadeia única (ver, *e. g.*, patente U.S. N° 4946778) podem ser

adaptadas para produzir anticorpos de cadeia única para os polipéptidos da invenção. Alternativamente, murganhos transgénicos podem ser utilizados para expressar anticorpos humanizados para estes polipéptidos ou os seus fragmentos.

Os anticorpos gerados contra os polipéptidos da invenção podem ser utilizados no rastreio de polipéptidos semelhantes de outros organismos e amostras. Em tais técnicas, os polipéptidos do organismo são colocados em contacto com o anticorpo e esses polipéptidos que se ligam especificamente ao anticorpo são detectados. Quaisquer dos processos descritos acima podem ser utilizados para detectar a ligação de anticorpo.

Kits

Os kits compreendem as composições, *e. g.*, ácidos nucleicos, cassetes de expressão, vectores, células, polipéptidos (*e. g.*, fitases) e/ou anticorpos da invenção. Os kits podem também conter instruções que ensinam as metodologias e utilizações industriais da invenção, como aqui descrito.

Os polipéptidos da invenção podem também ser utilizados para produzir anticorpos que se ligam especificamente aos polipéptidos ou fragmentos enzimáticos. Os anticorpos resultantes podem ser utilizados em processos de cromatografia de imunoafinidade para isolar ou purificar o polipéptido ou para determinar se o polipéptido está presente numa amostra biológica. Em tais processos, uma preparação de proteína, tais como um extracto ou uma amostra biológica, é colocada em contacto com um anticorpo capaz de se ligar especificamente a um de um polipéptido da SEQ ID N°:2, a sequências substancialmente idênticas a isso ou a fragmentos das sequências anteriores.

Em processos de imunoafinidade, o anticorpo é ligado a um suporte sólido, tais como um grânulo ou outra matriz de coluna. A preparação de proteína é colocado em contacto com o anticorpo sob condições em que o anticorpo se liga especificamente a um dos polipéptidos da SEQ ID N°:2, a sequências substancialmente idênticas a isso ou aos seus fragmentos. Após uma lavagem para remover as proteínas ligadas não especificamente, os polipéptidos especificamente ligados são eluídos.

As sequências de polinucleótidos, sequências de polipéptidos, variantes e os seus mutantes isolados podem ser medidos para retenção da actividade biológica característica da enzima da presente invenção, por exemplo, num ensaio para detectar actividade enzimática de fitase (Food Chemicals Codex, 4ª Ed.). Tais enzimas incluem formas truncadas de fitase e variantes, tais como variantes de deleção e inserção da sequência de polipéptido como apresentada na SEQ ID N°:2 modificada como aqui descrito. Estas fitases têm termotolerância. Isto é, a fitase tem uma actividade específica residual de cerca de 90% após tratamento a 70 °C durante 30 minutos e cerca de 50% após tratamento a 75 °C durante 30 minutos. A termotolerância das fitases da invenção é vantajosa na utilização da enzima como um aditivo de ração, uma vez que a ração pode ser moldada, granulada ou preparada em pastilhas a uma alta temperatura.

Por exemplo, num aspecto, a invenção proporciona uma matriz de distribuição enzimática em pastilhas comestíveis e método de utilização para distribuição de fitase a um animal, por exemplo como um suplemento nutritivo. A matriz de distribuição enzimática liberta prontamente uma enzima fitase, tal como uma tendo a sequência de aminoácidos da SEQ ID N°:2 modificada, como

aqui descrito ou, pelo menos, 30 dos seus aminoácidos contíguos, em meio aquoso, tal como, por exemplo, o suco digestivo de um animal. A matriz de distribuição enzimática da invenção é preparada a partir de um veículo comestível granulado seleccionado de tais componentes como germe de grão que é desperdício de óleo, feno, alfalfa, erva-dos-prados, casca de soja, farinha de sementes de girassol, farinha de trigo e semelhantes, que dispersam prontamente a enzima recombinante aí contida em meio aquoso. Em utilização, a matriz de distribuição enzimática em pastilhas comestíveis é administrada a um animal para distribuição da fitase ao animal. Os substratos à base de grão adequados podem compreender ou ser derivados de qualquer grão comestível adequado, tais como trigo, milho, soja, sorgo, alfalfa, cevada e semelhantes. Um substrato à base de grão exemplificativo é um substrato à base de milho. O substrato pode ser derivado de qualquer parte adequada do grão, *e. g.*, um germe de grão, aprovado para utilização em ração animal, tal como o germe de milho que é obtido num processo de moagem húmido ou seco. O germe de grão pode compreender desperdício de germe, que é germe de grão do qual foi removido óleo, tais como por prensagem ou extracção com hexano ou outro solvente. Alternativamente, o germe de grão é bagaço extraído, isto é, o óleo foi removido por prensagem.

A matriz de distribuição enzimática da invenção está na forma de partículas, pastilhas ou grânulos plurais discretos. Por “grânulos” entende-se partículas que são comprimidas ou compactadas, tais como através de formação de pastilhas, extrusão, ou compactação semelhante para remover água da matriz. Tal compressão ou compactação das partículas também promove coesão intra-partículas das partículas. Por exemplo, os grânulos podem ser preparados através da formação de pastilhas de substrato à base de grão num moinho de pastilhas. As pastilhas

assim preparadas estão moídas ou desintegradas num tamanho de grânulo adequado para utilização como um adjuvante na ração animal. Uma vez que a matriz é ela própria aprovada para utilização em ração animal, esta pode ser utilizada como um diluente para a distribuição de enzimas na ração animal.

A matriz de distribuição enzimática pode estar na forma de grânulos tendo um tamanho de grânulo variando desde cerca de 4 até cerca de 400 de malha (USS); ou cerca de 8 a cerca de 80 de malha; ou cerca de 14 a cerca de 20 de malha. Se o germe de grão for desperdício de extracção com solvente, a utilização de um agente lubrificante, tal como óleo de milho pode ser necessário na preparação das pastilhas, mas tal agente lubrificante habitualmente não é necessário se o germe for extraído por prensagem. Em outros aspectos da invenção, a matriz é preparada através de outros processos de compactação ou compressão, tais como, por exemplo, por extrusão do substrato à base de grão através de uma fieira e moagem do extrudido a um tamanho de grânulo adequado.

A matriz de distribuição enzimática pode ainda incluir um componente de polissacárido como um agente de coesividade para aumentar a coesividade dos grânulos da matriz. Pensa-se que o agente de coesividade proporciona grupos hidroxilo adicionais, que aumentam a ligação entre as proteínas de grão dentro do grânulo da matriz. Pensa -se ainda que os grupos hidroxilo adicionais funcionam por aumento das ligações de hidrogénio das proteínas para o amido e para outras proteínas. O agente de coesividade pode estar presente numa qualquer quantidade adequada para aumentar a coesividade dos grânulos da matriz de distribuição enzimática. Os agentes de coesividade adequados incluem uma ou mais dextrinas, maltodextrinas, amidos, tal como amido de milho, farinhas, celulósicos, hemicelulósicos e

semelhantes. Por exemplo, a percentagem de germe de grão e de agente de coesividade na matriz (não incluindo a enzima) é 78% de farinha de germe de milho e 20% em peso de amido de milho.

Uma vez que a matriz de libertação de enzima da invenção é feita de materiais biodegradáveis, a matriz pode ser sujeita a deterioração, tal como por bolores. Para prevenir ou inibir tais bolores, a matriz pode incluir um inibidor de bolores, tal como um sal de propionato, quo qual pode estar presente numa qualquer quantidade suficiente para inibir o bolor da matriz de libertação de enzima, proporcionando assim uma matriz de distribuição numa formulação estável que não requer refrigeração.

A enzima fitase contida na matriz de distribuição enzimática da invenção e nos métodos é, num aspecto, uma fitase termotolerante, como aqui descrita, para resistir a inactivação de fitase durante o fabrico onde podem ser utilizadas temperaturas elevadas e/ou vapor para preparar a matriz de distribuição enzimática em pastilhas. Durante a digestão de ração contendo a matriz de distribuição enzimática da invenção, os sucos digestivos aquosos causarão a libertação da enzima activa. Outros tipos de enzima termotolerantes e suplementos nutritivos que são termotolerantes podem também ser incorporados na matriz de distribuição para libertação sob qualquer tipo de condições aquosas.

Pode ser aplicado um revestimento às partículas da matriz enzimática da invenção para fins muito diferentes, tais como para adicionar um sabor ou suplemento nutricional à ração animal, para retardar a libertação de suplementos da ração animal e enzimas em condições gástricas, e semelhantes. Ou o revestimento pode ser aplicado para atingir um objectivo

funcional, por exemplo, sempre que for desejável retardar a libertação da enzima a partir das partículas da matriz ou para controlar as condições sob as quais a enzima será libertada. A composição do material de revestimento pode ser tal que é degradado selectivamente por um agente a que é susceptível (tais como calor, ácido ou base, enzimas ou outros químicos). Alternativamente, dois ou mais revestimentos susceptíveis a diferentes desses agentes de degradação podem ser aplicados consecutivamente às partículas da matriz.

A invenção é também dirigida a um processo para preparar uma matriz de libertação de enzima. De acordo com a invenção, o processo compreende proporcionar partículas plurais discretas de um substrato à base de grão num tamanho de partícula adequado para utilização como uma matriz de libertação de enzima, em que as partículas compreendem uma enzima fitase da invenção. O processo pode inclui compactação ou compressão de partículas da matriz de libertação de enzima em grânulos, o qual pode ser realizado por formação de pastilhas. O inibidor de bolor e o agente de coesividade, quando utilizados, podem ser adicionados em qualquer momento adequado e podem ser misturados com o substrato à base de grão nas proporções desejadas antes da formação de pastilhas do substrato à base de grão. O teor de humidade no alimentador do moinho de pastilhas pode ser nas gamas apresentadas acima relativamente ao teor de humidade no produto final, ou cerca de 14% a 15%, ou cerca de 10% a 20%. A humidade pode ser adicionada à ração na forma de uma preparação aquosa da enzima para ajustar a ração a este teor de humidade. A temperatura no moinho de pastilhas é ajustada a cerca de 82 °C com vapor. O moinho de pastilhas pode ser operado sob quaisquer condições que proporcionem trabalho suficiente à ração para proporcionar pastilhas. O próprio processo de formação de

pastilhas é um processo rentável para remover a água da composição contendo enzima.

Num aspecto, o moinho de pastilhas é operado com uma fieira de 3-50 mm (1/8 polegada por 2 polegadas) à pressão de 45 kg/min (100 1b./min.) a 82 °C para proporcionar as pastilhas, as quais são depois desintegradas num desintegrador de moinho de pastilhas para proporcionar partículas plurais discretas tendo um tamanho de partícula capaz de passar através de um peneiro de malha 8, mas ser retido num peneiro de malha 20.

As fitases termotolerantes aqui descritas podem ter temperaturas óptimas altas e podem ter resistência ao calor ou tolerância ao calor alta. Assim, as fitases da invenção podem realizar reacções enzimáticas a temperaturas normalmente consideradas acima da óptima. As fitases da invenção podem também realizar reacções enzimáticas após serem expostas a altas temperaturas (sendo a termotolerância a capacidade para reter actividade enzimática a temperaturas em que a fitase de tipo selvagem é activa após ser previamente exposta a altas temperaturas, mesmo se a alta temperatura puder inactivar ou diminuir a actividade da enzima, ver também a definição de termotolerância, acima). O gene que codifica a fitase de acordo com a presente invenção é utilizado na preparação de fitases (*e. g.*, utilizando GSSM como aqui descrito) tendo características diferentes daquelas da fitase da SEQ ID N°:2 (em termos de pH óptimo, temperatura óptima, resistência ao calor, estabilidade a solventes, actividade específica, afinidade para o substrato, capacidade de secreção, taxa da tradução, controlo de transcrição e semelhantes). Além disso, os polinucleótidos da invenção podem ser utilizados para o rastreio de fitases variantes preparadas pelos métodos aqui descritos, para determinar aquelas tendo uma actividade desejada, tais como

termostabilidade ou termotolerância melhorada ou modificada. Por exemplo, a Patente U.S. N° 5830732 descreve um ensaio de rastreio para determinar a termotolerância de uma fitase.

Um exemplo *in vitro* de tal ensaio de rastreio é o seguinte ensaio para a detecção da actividade de fitase: A actividade de fitase pode ser medida por incubação de 150 µL da preparação de enzima com 600 µL de fitato de sódio a 2 mM em tampão Tris HCl a 100 mM, pH 7,5, suplementado com $CaCl_2$ a 1 mM durante 30 minutos a 37 °C. Após incubação a reacção é parada por adição de 750 µL de ácido tricloroacético a 5%. O fosfato libertado foi medido espectrofotometricamente a 700 nm contra fosfato após ter sido adicionado 1500 µL de reagente de cor (4 volumes de molibdato de amónio a 1,5% em ácido sulfúrico a 5,5% e 1 volume de sulfato ferroso a 2,7%; Shimizu, 1992). Uma unidade de actividade da enzima é definida como a quantidade de enzima requerida para libertar um µmol de Pi por minuto sob condições de ensaio. A actividade específica pode ser expressa em unidades de actividade da enzima por mg de proteína. A enzima da presente invenção tem actividade enzimática relativamente à hidrólise de fitato a inositol e fosfato livre.

Num aspecto, a presente invenção proporciona um método para hidrolisar fitato compreendido de colocar em contacto o fitato com uma ou mais novas moléculas de fitase aqui divulgadas (*e. g.*, proteínas tendo as modificações específicas da SEQ ID N°:2). Desta forma, a invenção proporciona um método para catalisar a hidrólise de fitato a inositol e fosfato livre com libertação de minerais do complexo de ácido fítico. O método inclui colocar em contacto um substrato de fitato com uma quantidade eficaz de degradação de uma enzima da invenção. O termo quantidade "eficaz de degradação" refere-se à quantidade de enzima que é requerida para degradar, pelo menos, 50% do

fitato, em comparação ao fitato não colocado em contacto com a enzima. 80% do fitato pode ser degradado.

Noutro aspecto, a invenção proporciona um método para hidrolisar ligações fosfo-mono-éster no fitato. O método inclui a administração de uma quantidade eficaz de moléculas de fitase da invenção, para produzir inositol e fosfato livre. Num aspecto, uma quantidade "eficaz" refere-se à quantidade de enzima que é requerida para hidrolisar, pelo menos, 50% das ligações fosfo-mono-éster, em comparação ao fitato não colocado em contacto com a enzima. Num aspecto, pelo menos, 80% das ligações são hidrolisadas.

Num aspecto particular, quando desejado, as moléculas de fitase podem ser utilizadas em combinação com outros reagentes, tal como outros catalisadores; de modo a efectuar alterações químicas (*e. g.*, hidrólise) nas moléculas de fitato e/ou noutras moléculas da(s) fonte(s) de substrato. De acordo com este aspecto, as moléculas de fitase e o(s) reagente(s) adicional(ais) não se inibirão uns aos outros. As moléculas de fitase e o(s) reagente(s) adicional(ais) têm um efeito aditivo global, ou alternativamente, as moléculas de fitase e o(s) reagente(s) adicional(ais) podem ter um efeito sinérgico global.

As fontes relevantes de moléculas de fitato de substrato incluem géneros alimentícios, potenciais géneros alimentícios, subprodutos de géneros alimentícios (subprodutos *in vitro* e subprodutos *in vivo*, *e. g.*, produtos de reacção *ex vivo* e produtos de excrementos de animais), precursores de géneros alimentícios e qualquer outra fonte material de fitato.

Num aspecto não limitante, a fitase recombinante pode ser consumida por organismos e reter actividade após consumo. Noutra

exemplificação, as abordagens transgénicas podem ser utilizadas para conseguir a expressão da fitase recombinante - *e. g.*, de um modo controlado (estão disponíveis métodos para controlar a expressão de moléculas transgénicas em modos específicos de tempo e específicos de tecido).

Num aspecto, a actividade de fitase no material fonte (*e. g.*, uma fonte vegetal transgénica ou um hospedeiro procariótico recombinante) pode ser aumentada após consumo; este aumento na actividade pode ocorrer, por exemplo, após conversão de uma molécula de fitase precursora em pró-forma numa enzima significativamente mais activa numa forma mais madura, em que a referida conversão pode resultar, por exemplo, da ingestão e digestão da fonte de fitase. A hidrólise do substrato fitato pode ocorrer em qualquer altura após colocação em contacto da fitase com o fitato; por exemplo, isto pode ocorrer antes da ingestão ou após ingestão ou tanto antes como depois da ingestão do substrato ou da enzima ou de ambos. É adicionalmente entendido que o substrato fitato pode ser colocado em contacto com - para além da fitase - um ou mais reagentes adicionais, tal como outra enzima, a qual pode também ser aplicada directamente ou após purificação do seu material fonte.

É entendido que o(s) material(ais) fonte de fitase pode(m) ser colocado(s) em contacto directamente com o(s) material(ais) fonte de fitato; *e. g.*, após moagem ou trituração, *in vitro* ou *in vivo,* de cada uma ou de ambas as fontes de fitase e fontes de fitato. Alternativamente, a enzima fitase pode ser purificada do(s) material(ais) fonte ou o substrato fitato pode purificado do(s) material(ais) fonte ou a enzima fitase e o substrato fitato podem ser purificados do(s) material(ais) fonte antes do contacto da enzima fitase com o substrato fitato. É entendido que uma combinação de reagentes purificados e não purificados -

incluindo enzima(s) ou substrato(s) ou ambos - pode ser utilizada.

É entendido que mais do que um material fonte pode ser utilizado como uma fonte de actividade de fitase. Isto é útil como uma forma para conseguir uma libertação temporizada do(s) reagente(s) a partir do(s) material(ais) fonte, em que a libertação a partir dos diferentes reagentes dos seus materiais fonte ocorre diferencialmente, por exemplo, uma vez que materiais fonte ingeridos são digeridos *in vivo* ou uma vez que os materiais fonte são processados em aplicações *in vitro*. A utilização de mais de um material fonte da actividade de fitase é também útil para se obter actividades de fitase sob uma gama de condições e as suas flutuações, as quais podem ser encontradas - tal como uma gama de valores de pH, temperaturas, salinidades e intervalos de tempo - por exemplo, durante diferentes passos de processamento da uma aplicação. A utilização de diferentes materiais fonte é também útil de modo a obter diferentes reagentes, como exemplificado por uma ou mais formas ou isómeros de fitase e/ou fitato e/ou outros materiais.

É entendido que um único material fonte, tal como uma espécie vegetal transgénica (ou as suas partes de planta), pode ser um material fonte de fitase e fitato; e que enzimas e substratos podem ser diferencialmente compartimentalizados dentro da referida única fonte - *e. g.*, segregada vs. não segregada, diferencialmente expressa e/ou tendo abundâncias diferenciais em diferentes partes da planta ou órgãos ou tecidos ou compartimentos subcelulares dentro da mesmo parte de planta ou órgão ou tecido. A purificação das moléculas de fitase aí contidas pode compreender isolamento e/ou processamento adicional de uma ou mais partes desejáveis de planta ou órgãos ou tecidos ou compartimentos subcelulares.

Um método para catalisar reacções *in vivo* e/ou *in vitro* utilizando sementes contendo quantidades aumentadas de enzimas compreende a adição de sementes transgénicas, não de tipo selvagem, *e. g.*, numa forma moída, a uma mistura reaccional e permitir que as enzimas nas sementes aumentem a velocidade da reacção. Ao adicionar directamente as sementes à mistura reaccional, o método proporciona uma solução ao processo mais caro e mais incómodo de extrair e purificar a enzima. São também proporcionados métodos de tratamento em que um organismo que carece de uma fonte suficiente de uma enzima é-lhe administrada a enzima na forma de sementes de uma ou mais espécies de plantas, *e. g.*, espécies vegetais transgénicas, contendo quantidades aumentadas da enzima. Detalhes adicionais relativamente a esta abordagem estão na literatura publicada e/ou são conhecidos do especialista na técnica. Num exemplo particular não limitativo, tal literatura publicamente disponível inclui a patente U.S. N° 5543576 (van Ooijen *et al.*) e a Patente U.S. N° 5714474 (van Ooijen *et al.)*, embora estas referências não divulguem as moléculas inventivas do presente pedido e, em vez disso, ensinem a utilização de fitases fúngicas.

Num aspecto, as presentes moléculas de fitase são úteis para gerar formas vivas recombinantes do sistema digestivo (ou micróbios ou flora) e para a administração das referidas formas vivas recombinantes do sistema digestivo a animais. A administração pode ser opcionalmente realizada isoladamente ou em combinação com outras enzimas e/ou com outras formas vivas que podem proporcionar actividade enzimática num sistema digestivo, em que as referidas outras enzimas e referidas formas vivas podem ser recombinantes ou não. Por exemplo, a

administração pode ser realizada em combinação com bactérias xilanolíticas.

Num aspecto, a presente invenção proporciona um método para embeber sementes de milho ou sorgo em água morna contendo dióxido de enxofre na presença de uma preparação de enzima compreendendo uma ou mais enzimas que degrada a fitina, *e. g.*, em tal quantidade que a fitina presente no milho ou sorgo é substancialmente degradada. A preparação de enzima pode compreender fitase e/ou fosfatase ácida e opcionalmente outras enzimas que degradam material vegetal. O tempo de embebimento pode ser 12 a 18 horas. O embebimento pode ser interrompido por um passo de moagem intermédio, reduzindo o tempo de embebimento. Num aspecto, as sementes de milho ou sorgo são embebidas em água morna contendo dióxido de enxofre na presença de uma preparação de enzima incluindo uma ou mais enzimas que degrada a fitina, tais como fitase e fosfatases ácidas, para eliminar ou reduzir muito o ácido fítico e os sais de ácido fítico. Detalhes adicionais relativamente a esta abordagem estão na literatura publicadada e/ou são conhecidos do especialista na técnica, *e. g.*, Patente U.S. N° 4914029, (Caransa *et al.*) e documento EP 0321004 (Vaara *et al.*).

Num aspecto, a presente divulgação proporciona um método para obter uma massa de pão tendo propriedades físicas desejáveis, tais como não adesividade e elasticidade e um produto de pão de qualidade superior, tal como um volume específico, compreendendo a adição de moléculas de fitase à massa de pão. Num aspecto, as moléculas de fitase da presente invenção são adicionadas a uma preparação de massa de pão de trabalho que é subsequentemente formada e cozida. Detalhes adicionais relativamente a esta abordagem estão na literatura

publicadada e/ou são conhecidos do especialista na técnica, por exemplo, documento JP 03076529 (Hara *et al.*).

Num aspecto, a presente divulgação proporciona um método para produzir géneros alimentícios de soja melhorados. A soja é combinada com moléculas de fitase da presente invenção para remover ácido fítico da soja, produzindo assim géneros alimentícios de soja que estão melhorados no seu fornecimento de nutrientes vestigiais para organismos consumidores e na sua digestibilidade de proteínas. Num aspecto, na produção de leite de soja, as moléculas de fitase da presente invenção são adicionadas ou colocadas em contacto com soja de modo a reduzir o teor de ácido fítico. Num exemplo não limitativo, o processo de aplicação pode ser acelerado através de agitação do leite de soja em conjunto com a enzima sob aquecimento ou por condução de uma reacção de tipo mistura num recipiente de agitação utilizando uma enzima imobilizada. Detalhes adicionais relativamente a esta abordagem estão na literatura publicada e/ou são conhecidos do especialista na técnica, por exemplo, documento JP 59166049 (Kamikubo *et al.*).

Num aspecto, a presente divulgação proporciona um método para produzir um produto de mistura para água de beber ou ração animal em forma fluida e que compreende a utilização de misturas minerais e misturas de vitaminas e, também, novas moléculas de fitase da presente invenção. Num aspecto, é obtida uma mistura correctamente composta e doseada de nutrientes necessários para o organismo consumidor sem qualquer risco de precipitação e destruição de minerais/vitaminas importante, enquanto ao mesmo tempo é feita utilização óptima do fosfato ligado à fitina na ração. Detalhes adicionais relativamente a esta abordagem estão na literatura publicada e/ou são conhecidos do especialista na técnica, *e. g.*, documento EP 0772978 (Bendixen *et al.*).

É entendido que as moléculas de fitase da presente invenção podem também ser utilizadas para produzir outros géneros alimentícios bebíveis (ou bebidas) alcoólicos e não alcoólicos com base na utilização de bolores e/ou grãos e/ou outras plantas. Estes géneros alimentícios bebíveis incluem licores, vinhos, bebidas alcoólicas mistas (*e. g.*, *wine coolers*, outros cafés alcoólicos, tal como *Irish coffees*, etc.), cervejas, bebidas de tipo cerveja, sumos, extractos, homogenatos e purés. Num aspecto, as moléculas de fitase presentemente divulgadas são utilizadas para produzir versões transgénicas de bolores e/ou grãos e/ou outras plantas úteis para a produção de tais géneros alimentícios bebíveis. Noutro aspecto, as moléculas de fitase presentemente divulgadas são utilizadas como ingredientes adicionais no processo de fabrico e/ou no conteúdo final de tais géneros alimentícios bebíveis. Detalhes adicionais relativamente a esta abordagem estão na literatura publicada e/ou são conhecidos do especialista na técnica.

Num aspecto, a presente invenção proporciona meios para obter *sake* refinado tendo uma quantidade reduzida de fitina e um teor aumentado de inositol. Tal *sake* pode ter - através de efeitos directos e/ou psicogénicos - uma acção preventiva na doença hepática, arteriosclerose e noutras doenças. Num aspecto, um *sake* é produzido a partir de arroz *koji* por multiplicação de um bolor de arroz *koji* tendo actividade elevada de fitase como uma matéria-prima. É entendido que as moléculas de fitase da presente invenção pode ser utilizadas para produzir um bolor útil com actividade aumentada (*e. g.*, um bolor transgénico) e/ou adicionado exogenamente para aumentar os efeitos de um bolor *koji*. A estirpe é adicionada a arroz fervido e o *koji* é produzido por um processo convencional. Numa exemplificação, o *koji* preparado é utilizado, o arroz completo é preparado em duas

fases e o *sake* é produzido à temperatura de *sake* constante de 15 °C para produzir o pretendido *sake* refinado tendo uma quantidade reduzida de fitina e uma quantidade aumentada de inositol. Detalhes adicionais relativamente a esta abordagem estão na literatura publicada e/ou são conhecidos do especialista na técnica, por exemplo, documentos JP 06153896 (Soga *et al.)* e JP 06070749 (Soga *et al.*).

Num aspecto, a presente divulgação proporciona um método para obter um absorvente capaz de promover a absorção de minerais, incluindo cálcio, ingeridos sem ser digeridos por sucos gástricos ou sucos intestinais com baixo custo. Num aspecto, o absorvente mineral contém um hidrolisado parcial de ácido fítico como um ingrediente activo. Um hidrolisado parcial do ácido fítico pode ser produzido por hidrólise de ácido fítico ou os seus sais utilizando novas moléculas de fitase da presente invenção. O tratamento com as moléculas de fitase pode ocorrer isoladamente e/ou num tratamento de combinação (para inibir ou aumentar o efeito final) e é seguido por inibição da hidrólise dentro de uma gama de modo a não libertar todos os radicais de fosfato. Detalhes adicionais relativamente a esta abordagem estão na literatura publicada e/ou são conhecidos do especialista na técnica, *e. g.*, documento JP 04270296 (Hoshino).

Num aspecto, a presente invenção proporciona um método (e seus produtos) para produzir uma composição enzimática que tem uma actividade de hidrólise do fitato aditiva ou sinérgica; a referida composição compreende novas moléculas de fitase da presente invenção e um ou mais reagentes adicionais para conseguir uma composição que seja útil para um tratamento de combinação. Num aspecto, o tratamento de combinação da presente invenção é conseguido através da utilização de, pelo menos, duas fitases de diferentes especificidades de posição, *i. e.*,

quaisquer combinações de 1-, 2-, 3-, 4-, 5- e 6-fitases. Ao combinar fitases de diferentes especificidades de posição, é obtido um efeito aditivo ou sinérgico. As composições tais como alimentos e rações ou aditivos de alimentos e rações compreendendo tais fitases em combinação estão também incluídas nesta invenção, assim como processos para a sua preparação. Detalhes adicionais relativamente a esta abordagem estão na literatura publicada e/ou são conhecidos do especialista na técnica, *e. g.*, documento WO 30681 (Ohmann *et al.*).

Noutro aspecto, o tratamento de combinação da presente invenção é conseguido por utilização de uma fosfatase ácida que tem actividade de hidrólise de fitato a um pH de 2,5, numa razão baixa que corresponde a um perfil da actividade de pH 2,5:5,0 desde cerca de 0,1:1,0 a 10:1, ou desde cerca de 0,5:1,0 a 5:1, ou desde cerca de 0,8:1,0 a 3:1, ou desde cerca de 0,8:1,0 a 2:1. A composição enzimática pode apresentar uma eficiência sinérgica de hidrólise de fitato mais elevada através de tratamento térmico. A composição enzimática é útil no tratamento de géneros alimentícios (bebíveis e alimentos sólidos, produtos de ração e forragem) para melhorar a hidrólise de fitato. Detalhes adicionais ou protocolos alternativos relativamente a esta abordagem estão na literatura publicada e/ou são conhecidos do especialista na técnica, *e. g.*, Patente U.S. N° 5554399 (Vanderbeke *et al.*) e Patente U.S. N° 5443979 (Vanderbeke *et al.*), que ensina a utilização de fitases fúngicas (em particular A*spergillus*).

Noutro aspecto, a presente invenção proporciona um método (e seus produtos) para produzir uma composição compreendendo a presente e nova enzima actuando no fitato em combinação com uma ou mais enzimas adicionais actuando em polissacáridos. Tais polissacáridos podem ser seleccionados do grupo consistindo de

arabinanas, fructanas, fucanas, galactanas, galacturonanas, glucanas, mananas, xilanas, levana, fucoidana, carragenina, galactocarolose, pectina, ácido péctico, amilose, pululano, glicogénio, amilopectina, celulose, carboxilmetilcelulose, hidroxipropilmetilcelulose, dextrano, pustulano, quitina, agarose, queratano, condroitina, dermatano, ácido hialurónico, ácido algínico, e polissacáridos contendo, pelo menos, uma aldose, cetose, ácido ou amina seleccionada do grupo consistindo de eritrose, treose, ribose, arabinose, xilose, lixose, alose, altrose, glucose, manose, gulose, idose, galactose, talose, eritrulose, ribulose, xilulose, psicose, frutose, sorbose, tagatose, ácido glucurónico, ácido glucónico, ácido glucárico, ácido galacturónico, ácido manurónico, glucosamina, galactosamina e ácido neuramínico.

Num aspecto, a presente invenção proporciona um método (e seus produtos) para produzir uma composição tendo uma actividade de hidrólise do fitato sinérgica compreendendo uma ou mais novas moléculas de fitase da presente invenção, uma celulase (pode também incluir uma xilanase), opcionalmente uma protease e, opcionalmente, um ou mais reagentes adicionais. Em aspectos alternativos, tais tratamentos de combinação são úteis no tratamento de géneros alimentícios, produtos de madeira, tal como produtos de papel, e como soluções e sólidos de limpeza.

Num aspecto, as fitases da invenção são úteis em combinação com componentes de celulose. É conhecido que as celulases de muitas bactérias celulolíticas estão organizadas em complexos multi-enzimáticos discretos, denominados celulossomas. As múltiplas subunidades de celulossomas são compostas de numerosos domínios funcionais, os quais interagem uns com os outros e com o substrato celulósico. Uma destas subunidades compreende uma nova classe distintiva de polipéptidos estruturais não

catalíticos, que integra selectivamente as várias subunidades de celulase e xilanase no complexo coesivo. A aplicação inteligente de híbridos de celulossoma e construções quiméricas de domínios celulossomais deve permitir uma melhor utilização da biomassaa celulósica e pode oferecer uma vasta gama de novas aplicações em investigação, medicina e indústria.

Num aspecto, as fitases da invenção são úteis - isoladamente ou em tratamentos de combinação - nas áreas de bioprodução de pasta e bio-branqueamento onde é desejada uma redução na utilização de químicos ambientalmente prejudiciais, tradicionalmente utilizados de indústria de papel e pasta. O tratamento de água residuais representa uma outra vasta área de aplicação em que foi demonstrado que as enzimas biológicas são eficazes não apenas na remoção da cor, mas também na bioconversão de substâncias potencialmente nocivas em subprodutos úteis.

Num aspecto, as fitases da invenção são úteis para produzir formas vivas que podem proporcionar, pelo menos, uma actividade enzimática - isoladamente ou em tratamentos de combinação - no tratamento de sistemas digestivos de organismos. Os organismos particularmente relevantes a serem tratados incluem organismos não ruminantes, embora os organismos ruminantes também possam beneficiar de tal tratamento. Especificamente, é entendido que esta abordagem pode ser realizadaa isoladamente ou em combinação com outras moléculas biológicas (por exemplo, xilanases) para gerar um hospedeiro recombinante que expresse uma pluralidade de moléculas biológicas. É também entendido que a administração das presentes moléculas de fitase e/ou hospedeiros recombinantes que expressam as presentes moléculas de fitase pode ser realizada isoladamente ou em combinação com outras moléculas biológicas, e/ou formas vivas que podem proporcionar actividade enzimática

num sistema digestivo - em que as referidas outras enzimas e referidas formas vivas podem ser recombinantes ou não. Por exemplo, a administração pode ser realizada em combinação com bactérias xilanolíticas.

Por exemplo, além do fitato, muitos organismos são também incapazes de digerir adequadamente hemiceluloses. As hemiceluloses ou os xilanos são os principais componentes (35%) de materiais vegetais. Para animais ruminantes, cerca de 50% dos xilanos dietéticos são degradados, mas apenas pequenas quantidades de xilanos são degradadas no intestino inferior de animais e humanos não ruminantes. No rúmen, as principais espécies xilanolíticas são *Butyrivibrio fibrisolvens* e *Bacteroides ruminicola*. No cólon humano, as subespécies "a" de *Bacteroides ovatus* e *Bacteroides fragilis* são as principais bactérias xilanolíticas. Os xilanos são quimicamente complexos e a sua degradação requer múltiplas enzimas. A expressão destas enzimas por bactérias intestinais varia muito entre espécies. *Butyrivibrio fibrisolvens* produz xilanases extracelulares mas as espécies Bacteróides têm actividade de xilanase ligada à célula. A caracterização bioquímica de enzimas xilanolíticas de bactérias intestinais não foi totalmente realizada. Um gene da xilosidase foi clonado de *B. fibrosolvens*. Os dados de hibridações de ADN utilizando um gene de xilanase clonado de *B. fibrisolvens* indicam que este gene pode estar presente noutras estirpes de *B. fibrisolvens*. Uma xilanase clonada de *Bact. ruminicola* foi transferida e altamente expressa em *Bact. fragilis* e *Bact. Uniformis*. Os genes da arabinosidase e xilosidase de *Bact. ovatus* foram clonados e ambas as actividades parecem ser catalisadas por uma única enzima bifuncional nova.

Num aspecto, as fitases da invenção são úteis para 1) transferir para um hospedeiro adequado (tal como *Bact. fragilis*

ou *Bact. uniformis*); 2) conseguir expressão adequada num hospedeiro recombinante resultante; e 3) administrar os referidos organismos hospedeiros recombinantes para melhorar a capacidade dos organismos tratados para degradar fitato. A investigação continuada nas áreas da genética e bioquímica irá proporcionar conhecimento e compreensão para manipulação da digestão ao nível intestinal e melhorar a compreensão da digestão de fibra no cólon.

Detalhes adicionais ou protocolos alternativos relativamente a esta abordagem estão na literatura publicada e/ou são conhecidos do especialista na técnica, por exemplo, a invenção pode utilizar processos como descritos na Patente U.S. N° 5624678 (Bedford *et al.*), Patente U.S. N° 5683911 (Bodie *et al.*), Patente U.S. N° 5720971 (Beauchemin *et al.*), Patente U.S. N° 5759840 (Sung *et al.*), Patente U.S. N° 5770012 (Cooper), Patente U.S. N° 5786316 (Baeck *et al.*), Patente U.S. N° 5817500 (Hansen *et al.*).

A presente invenção ensina que as moléculas de fitase da presente invenção podem ser adicionadas ao(s) reagente(s) divulgado(s) de modo a obter preparações tendo uma actividade adicional de fitase. Num aspecto, o(s) reagente(s) e as moléculas adicionais de fitase não se inibirão entre si. Num aspecto, o(s) reagente(s) e as moléculas adicionais de fitase podem ter um efeito aditivo global. Num aspecto, o(s) reagente(s) e as moléculas adicionais de fitase podem ter um efeito sinérgico global.

Num aspecto, a presente invenção proporciona um método (e os seus produtos) para aumentar a utilização do fósforo de fitato e tratamento e prevenção da discondroplasia tibial em animais, particularmente aves, por administração aos animais de uma

composição de ração contendo um derivado hidroxilado da vitamina D_3. O derivado da vitamina D_3 pode ser administrado a animais em rações contendo níveis reduzidos de cálcio e fósforo para aumentar a utilização do fósforo de fitato. Desta forma, o derivado da vitamina D_3 pode ser administrado em combinação com novas moléculas de fitase da presente invenção para um aumento adicional da utilização do fósforo de fitato. Detalhes adicionais ou protocolos alternativos relativamente a esta abordagem estão na literatura publicada e/ou são conhecidos do especialista na técnica, *e. g.*, Patente U.S. N° 5516525 (Edwards *et al.*) e Patente U.S. N° 5366736 (Edwards *et al.*), Patente U.S. N° 5316770 (Edwards *et al.*).

Num aspecto, a presente invenção proporciona um método (e seus produtos) para obter géneros alimentícios que 1) compreendem fitina que é facilmente absorvida e utilizada numa forma de inositol num corpo de um organismo; 2) que é capaz de reduzir o fósforo em excrementos; e 3) que é, desta forma, útil para melhorar a poluição ambiental. Os referidos géneros alimentícios são compreendidos de uma mistura de uma grão contendo fitina, um microrganismo produtor de ácido láctico e uma nova molécula de fitase da presente invenção. Num aspecto, os referidos géneros alimentícios são produzidos por combinação de um grão contendo fitina (*e. g.*, farelo de arroz) com um grupo microbiano eficaz tendo uma propriedade acidofílica, produção de ácido láctico, sem produzirem ácido butírico, sem patogenicidade e uma fitase. Os exemplos de um grupo microbiano eficaz incluem, *e. g.*, *Estreptomyces sp.* (American Type Culture Collection N° ATCC 3004) pertencentes ao grupo dos actinomicetes e *Lactobacillus sp.* (IFO 3070) pertencentes ao grupo dos lactobacilos.

Uma quantidade exemplificativa de adição de um grupo microbiano eficaz é 0,2% em peso em termos do peso corporal bacteriano com base num material de grão. Num aspecto, a quantidade da adição de fitase é cerca de 1-2% em peso com base na fitina no material de grão. Detalhes adicionais ou protocolos alternativos relativamente a esta abordagem estão na literatura publicada e/ou são conhecidos do especialista na técnica, *e. g.*, documento JP 08205785 (Akahori *et al.*).

Num aspecto, a presente invenção proporciona um método para melhorar a solubilidade de proteínas vegetais. Mais especificamente, a invenção refere-se a métodos para a solubilização de proteínas em fontes de proteína vegetal, cujos métodos compreendem tratar fonte de proteína vegetal com uma quantidade eficiente de uma ou mais enzimas fitase da invenção e tratar a fonte de proteína vegetal com uma quantidade eficiente de uma ou mais enzimas proteolíticas. Noutro aspecto, a invenção proporciona aditivos de ração animal compreendendo uma fitase da invenção e uma ou mais enzimas proteolíticas. Detalhes adicionais ou protocolos alternativos relativamente a esta abordagem estão na literatura publicada e/ou são conhecidos do especialista na técnica, *e. g.*, documento EP 0756457 (WO 9528850 A1) (Nielsen e knap).

Num aspecto, a presente invenção proporciona um método para produzir uma preparação de proteína vegetal compreendendo dispersar materiais fonte de proteína vegetal em água a um pH na gama de 2 a 6 e misturar aí moléculas de fitase da presente invenção. o extracto acídico contendo proteína solúvel é separado e seco para produzir uma proteína sólida de carácter desejável. Uma ou mais proteases podem também ser utilizadas para melhorar as características da proteína. Detalhes adicionais ou protocolos alternativos relativamente a esta

abordagem estão na literatura publicada e/ou são conhecidos do especialista na técnica, *e. g.*, Patente U.S. Nº 3966971.

Num aspecto, a presente divulgação proporciona um método (e os seus produtos) para activar fósforo inerte no solo e/ou adubo, para melhorar a taxa de utilização de um composto azotado e para suprimir a propagação de bolores patogénicos por adição de três reagentes, fitase, saponina e quitosano, ao adubo.

Num aspecto, o método pode compreender tratar o adubo através de 1) adicionar microrganismos contendo fitase no meio, *e. g.*, hospedeiros recombinantes que sobreexpressam novas moléculas de fitase da presente invenção, por exemplo, a 100 mL de meio/100 kg de adubo húmido; 2) alternativamente também adicionar uma fonte vegetal contendo fitase – tal como farelo de trigo – *e. g.*, a 0,2 a 1 kg/100 kg de adubo húmido; 3) adicionar uma fonte contendo saponina – tal como turfa, artemísia e plantas yucca – *e. g.*, a 0,5 a 3,0 g/kg; 4) adicionar materiais contendo quitosano – tais como cascas pulverizadas de camarões, caranguejos, etc. – *e. g.*, a 100 a 300 g/kg de adubo húmido.

Num aspecto, são utilizadas fontes recombinantes dos três reagentes, fitase, saponina e quitosano. Detalhes adicionais ou protocolos alternativos relativamente a esta abordagem estão na literatura publicada e/ou são conhecidos do especialista na técnica, *e. g.*, documento JP 07277865 (Toya Taisuke).

Em alguns exemplos pode ser vantajoso distribuir e expressar localmente uma sequência de fitase da invenção (*e. g.*, dentro de um tipo particular de tecido ou célula). Por exemplo, a expressão local de uma fitase ou enzima digestiva no intestino de um animal ajudarão na digestão e incorporação, por exemplo, de fitato e fósforo, respectivamente. A sequência de ácido

nucleico pode ser distribuída directamente às glândulas, tecidos e células salivares e/ou às células epiteliais de revestimento do intestino, por exemplo. Tais métodos de distribuição são conhecidos na técnica e incluem electroporação, vectores virais e incorporação directa de ADN. Qualquer polipéptido que tem actividade de fitase pode ser utilizado nos métodos da invenção (*e. g.*, aqueles especialmente descritos na subsecção 6.3.18, assim como aqueles descritos noutras secções da invenção).

Por exemplo, as construções de ácidos nucleicos da presente invenção irão compreender moléculas de ácidos nucleicos numa forma adequada para incorporação em células alvo dentro de um tecido hospedeiro. Os ácidos nucleicos podem estar na forma de moléculas de ADN ou ARN nu, em que as moléculas podem compreender um ou mais genes estruturais, um ou mais genes reguladores, cadeias anti-sentido, cadeias capazes de formação de hélice tripla ou semelhantes. Geralmente, a construção de ácidos nucleicos incluirá, pelo menos, um gene estrutural sob controlo de transcrição e tradução de uma região reguladora adequada. Mais habitualmente, as construções de ácidos nucleicos da presente invenção irão compreender ácidos nucleicos incorporados num veículo de distribuição para melhorar a eficiência de transfecção, em que o veículo de distribuição será disperso dentro de partículas maiores compreendendo um material excipiente hidrofílico seco.

Um desses veículos de distribuição compreendem vectores virais, tais como retrovírus, adenovírus e vírus adeno-associados, os quais foram inactivados para prevenir auto-replicação mas que mantêm a capacidade viral nativa para se ligarem a uma célula alvo hospedeira, distribuir material genético no citoplasma da célula alvo hospedeira e promover a expressão de genes estruturais ou outros que foram incorporados

na partícula. Os vectores retrovírus adequados para transferência génica mediada são descritos em Kahn *et al*. (1992) Circ. Res. 71:1508-1517. Uma distribuição génica de adenovírus adequada é descrita em Rosenfeld *et al*. (1991) Science 252:431-434. Sistemas de distribuição retrovirais e adenovirais são descritos em Friedman (1989) Science 244:1275-1281.

Um segundo tipo de veículo de distribuição de ácidos nucleicos compreende vesículas lipossomais de transfecção, incluindo construções lipossomais aniónicas e catiónicas. A utilização de lipossomas aniónicos requer que os ácidos nucleicos sejam capturados dentro do lipossoma. Os lipossomas catiónicos não requerem a captura dos ácidos nucleicos e, em vez disso, podem ser formados por simples mistura dos ácidos nucleicos e lipossomas. Os lipossomas catiónicos ligam-se avidamente às moléculas de ácidos nucleicos carregadas negativamente, incluindo ADN e ARN, para produzir complexos que têm razoável eficiência de transfecção em muitos tipos de células. Ver, Farhood *et al*. (1992) Biochem. Biophys. Acta. 1111:239-246. Um material exemplificativo para formar vesículas lipossomais é a lipofectina que é composta de uma mistura equimolar de dioleilfosfatidiletanolamina (DOPE) e dioleiloxipropil-trietilamónio (DOTMA), como descrito em Felgner e Ringold (1989) Nature 337:387-388.

É também possível combinar estes dois tipos de sistemas de distribuição. Por exemplo, Kahn *et al*. (1992), *supra*., ensina que um vector retroviral pode ser combinado numa vesícula catiónica de DEAE-dextrano para aumentar ainda a eficiência de transformação. É também possível incorporar proteínas nucleares em vesículas de distribuição viral e/ou lipossomal para melhorar ainda mais as eficiências de transfecção. Ver, Kaneda *et al*. (1989) Science 243:375-378.

Noutro aspecto, é proporcionado um auxiliar digestivo contendo uma enzima como o único ingrediente activo ou em combinação com um ou mais outros agentes e/ou enzimas. A utilização de enzimas e de outros agentes em auxiliares digestivos de animais de criação ou animais domésticos não apenas melhora a saúde e a esperança de vida do animal, mas também ajuda a aumentar a saúde dos animais de criação e a produção de géneros alimentícios a partir de animais de criação.

A invenção pode também utilizar rações para animais de criação (*e. g.*, determinadas rações para aves) que são altamente suplementadas com numerosos minerais (*e. g.*, fósforo inorgânico), enzimas, factores de crescimento, fármacos e outros agentes para distribuição aos animais de criação. Estes suplementos substituem muitas das calorias e nutrientes naturais presentes no grão, por exemplo. Ao reduzir ou eliminar o suplemento de fósforo inorgânico e outros suplementos (*e. g.*, sais de minerais vestigiais, factores de crescimento, enzimas, antibióticos) da própria ração, a ração é capaz de conter mais nutrientes e energia. Desta forma, a dieta remanescente conteria mais energia utilizável. Por exemplo, as dietas de farinha de semente oleaginosa contêm habitualmente cerca de 3200 kcal de energia metabolizável por kg de dieta e os sais minerais não fornecem energia metabolizável. A remoção dos minerais desnecessários e a substituição com grão aumenta assim a energia utilizável na dieta. Assim, a invenção é diferenciada relativamente a ração contendo fitase habitualmente utilizada. Por exemplo, num aspecto, é utilizado um material biocompatível que é resistente à digestão pelo tracto gastrointestinal de um organismo.

Em muitos organismos, incluindo, por exemplo, aves domésticas ou pássaros como, por exemplo, galinhas, perus, gansos, patos, papagaios, pavões, avestruzes, faisões, codornizes, pombos, emas, quivis, mergulhões, caturras, catatua, canários, pinguins, flamingos e pombas, o tracto digestivo inclui uma moela que armazena e utiliza objectos biocompatíveis duros (*e. g.*, rochas e conchas de mariscos) para ajudar na digestão de sementes ou outros alimentos consumidos por uma ave. Um tracto digestivo típico desta família geral de organismos, inclui o esófago que contém uma bolsa, chamado papo, em que o alimento é armazenado durante um breve período de tempo. A partir do papo, o alimento desce ao verdadeiro estômago, ou *proventriculus*, onde o ácido clorídrico e a pepsina começam o processo de digestão. Em seguida, o alimento move-se para a moela, que é de forma oval e paredes grossas com músculos poderosos. A principal função da moela é triturar ou moer partículas alimentares - um processo é ajudado pela ave engolir pequenas quantidades de cascalho fino ou saibro. A partir da moela, o alimento move-se para o duodeno. O intestino delgado das aves é semelhante ao dos mamíferos. Existem duas bolsas cegas ou *ceca*, cerca de 100-150 mm (4-6 polegadas) de comprimento na junção do intestino delgado e grosso. O intestino grosso é curto consistindo essencialmente do recto com cerca de 75-100 mm (3-4 polegadas) de comprimento. O recto esvazia na *cloaca* e as fezes são excretadas através do ânus.

Os objectos biocompatíveis duros consumidos (ou de outro modo introduzidos) e presentes na moela proporcionam um vector útil para a distribuição de vários agentes enzimáticos, químicos, terapêuticos e antibióticos. Estas substâncias duras têm uma duração de algumas horas a alguns dias e passam após um período de tempo. Desta forma, a invenção proporciona auxiliares dietéticos modificados impregnados (*e. g.*, matriz e membranas

impregnadas) revestidos, para distribuição de agentes digestivos ou terapêuticos úteis a um organismo. Tais auxiliares dietéticos incluem objectos que são ingeridos tipicamente por um organismo para ajudar na digestão dentro da moela (*e. g.*, areia ou saibro). A invenção proporciona objectos biocompatíveis que têm aí revestidos ou impregnados agentes úteis como um auxiliar digestivo para um organismo ou para a distribuição de um agente ou químico terapêutico ou medicinal.

Num aspecto, a invenção proporciona um auxiliar dietético, tendo uma composição biocompatível concebida para a libertação de um agente que ajuda na digestão, em que a composição biocompatível é concebida para consumo oral e libertação no tracto digestivo (*e. g.*, moela) de um organismo. "Biocompatíveis" significa que a substância, após contacto com um organismo hospedeiro (*e. g.*, um ave), não induz uma resposta prejudicial suficiente para resultar na rejeição da substância ou que torne a substância inoperável. Tal inoperabilidade pode ocorrer, por exemplo, por formação de uma estrutura fibrótica em torno da substância que limita a difusão de agentes impregnados ao organismo hospedeiro ou uma substância que resulte num aumento na mortalidade ou morbidade no organismo devido a toxicidade ou infecção. Uma substância biocompatível pode ser não biodegradável ou biodegradável. Num aspecto, a composição biocompatível é resistente a degradação ou digestão pelo tracto gastrointestinal. Noutro aspecto, a composição biocompatível tem a consistência de uma rocha ou pedra.

Um material não biodegradável útil na invenção é um que permita ligação ou impregnação de um agente dietético. Tais materiais não biodegradáveis não limitativos incluem, por exemplo, termoplásticos, tais como acrílico, modacrílico, poliamida, policarbonato, poliéster, polietileno, polipropileno,

poliestireno, polissulfona, polietersulfona e fluoreto de polivinilideno. Os elastómeros são também materiais úteis e incluem, por exemplo, poliamida, poliéster, polietileno, polipropileno, poliestireno, poliuretano, álcool polivinílico e silicone (*e. g.*, à base de silicone ou contendo sílica). A invenção proporciona que a composição biocompatível possa conter uma pluralidade de tais materiais, os quais podem ser, *e. g.*, misturados ou depositados para formar misturas, copolímeros ou as suas combinações.

Num aspecto, um material "biodegradável" significa que a composição irá erodir ou degradar *in vivo* para formar espécies químicas menores. A degradação pode ocorrer, por exemplo, através de processos enzimáticos, químicos ou físicos. Os materiais biodegradáveis adequados contemplados para utilização na invenção incluem, mas não estão limitados a, poli(lactido)s, poli(glicolido)s, poli(acido láctico)s, poli(ácido glicólico)s, polianidros, poliortoésteres, polieterésteres, policaprolactona, poliesteramidas, policarbonato, policianoacrilato, poliuretanos, poliacrilato e semelhantes. Tais materiais podem ser misturados ou depositados para formar misturas, copolímeros ou as suas combinações.

Num aspecto, uma variedade de diferentes substâncias biocompatíveis da invenção podem ser administradas ao animal e ingeridas sequencialmente ou, caso contrário, proporcionadas ao mesmo organismo simultaneamente ou em várias combinações (*e. g.*, um material antes do outro). Além disso, as substâncias biocompatíveis da invenção podem ser concebidas para passagem lenta através do tracto digestivo. Por exemplo, as substâncias grandes ou gordas tendem a mover-se mais lentamente através do tracto digestivo, desta forma, pode ser utilizado um material biocompatível tendo um tamanho grande para prevenir a passagem

rápida no tracto digestivo. Tais substâncias grandes podem ser uma combinação de substâncias não biodegradáveis e biodegradáveis. Por exemplo, uma substância pequena não biodegradável pode ser abrangida por uma substância biodegradável da invenção de modo que durante um período de tempo a porção biodegradável será degradada permitindo que a porção não biodegradável passe através do tracto digestivo. Além disso, é reconhecido que qualquer número de aromatizantes podem ser proporcionados a uma substância biocompatível da invenção para ajudar no consumo.

Qualquer número de agentes isoladamente ou em combinação com outros agentes pode ser revestido nas substâncias biocompatíveis da invenção, incluindo polipéptidos (*e. g.*, enzimas, anticorpos, citocinas ou pequenas moléculas terapêuticas) e antibióticos, por exemplo. Os exemplos de agentes úteis particulares são listados na Tabela 1 e 2, abaixo. Está também contemplado que as células pode ser encapsuladas no material biocompatível da invenção e utilizado para distribuir enzimas ou agentes terapêuticos. Por exemplo, podem ser concebidas substâncias porosas tendo poros suficientemente grandes para que as células cresçam dentro e através e que estes materiais porosos possam ser depois levados para o tracto digestivo. Por exemplo, a substância biocompatível da invenção pode compreender uma pluralidade de ambientes microflorais (*e. g.*, diferente porosidade, pH, etc.) que proporcionam suporte para uma pluralidade de tipos de células. As células podem ser manipuladas geneticamente para distribuir um fármaco, enzima ou químico particular ao organismo. As células podem ser eucarióticas ou procarióticas.

Tabela 1

Classe de Tratamento	Químico	Descrição
Antibióticos	Amoxicilina e sua Combinação Mastox Injecção (Amoxicilina e cloxacilina)	Tratamento contra doenças bacterianas causadas por bactérias Gram + e Gram –
	Ampicilina e sua Combinação Biolox Injecção (Ampicilina e cloxacilina)	Tratamento contra doenças bacterianas causadas por bactérias Gram + e Gram –
	Nitrofurazona + Ureia Bolus de Nefrea	Tratamento de infecções genitais
	Trimetoprima + Sulfametoxazole Bolus de Trizole	Tratamento de infecções do tracto respiratório, infecções do tracto gastrointestinal, infecções urogenitais.
	Metronidazole e Furazolidona Bolus de Metofur	Tratamento de doenças bacterianas e protozoárias.
	Ftalilsulfatiazole, pectina e caulino Pectolina Bolus Suspensão	Tratamento de diarreia bacteriana e não específica, disenteria bacilar e paratifose.

(continuação)

Classe de Tratamento	Químico	Descrição
Anti-helmínticos	Unguento de Ectoparasiticida Germex (Hexacloreto de gama Benzeno, Hemissulfato de Proflavina e Cetrimida) Endoparasiticidas > Albendazole e a sua Combinação Alben (Albendazole) Suspensão (Albendazole 2,5%) Suspensão Plus (Albendazole 5%) Bolus Forte (Albendazole 1,5 Gm.) Comprimido (Albendazole 600 Mg.) Pó (Albendazole 5%, 15%)	Ectoparasiticida e antisséptico Prevenção e tratamento de infestações de lombrigas, ténias e fascíola
	Alpraz (Albendazole e Praziquantel) Comprimido	Prevenção e tratamento de infestações de lombrigas e ténias em caninos e felinos.
	Oxiclozanida e a sua Combinação Clozan (Oxiclozanida) bolus, Suspensão	Prevenção e tratamento de infestações de fascíola

(continuação)

Classe de Tratamento	Químico	Descrição
	Tetzan (Oxiclozanida e Tetramisole HCl) Bolus, Suspensão	Prevenção e tratamento de infestações de lombrigas e fascíola
	Fluzan (Oxiclozanida e Levamisole HCl) Bolus, Suspensão	Prevenção e tratamento de infestações de lombrigas e imunidade crescente
	Levamisole Injecção de Nemasol Pó Wormnil	Prevenção e tratamento de infestações de lombrigas e imunidade crescente.
	Fenbendazole Fenzole Comprimido (Fenbendazole 150 Mg.) Bolus (Fenbendazole 1,5 Gm.) Pó (Fenbendazole 2,5% p/p)	Prevenção e tratamento de infestações de lombrigas e ténias
Tónicos	Complexo de vitamina B, aminoácidos e Extracto de fígado Injecção de Heptogen	Tratamento de anorexia, hepatite, debilidade, convulsões nevrálgicas Emagrecimento e crescimento atrofiado.

(continuação)

Classe de Tratamento	Químico	Descrição
	Levulinato de Cálcio com Vit.B_{12} e Vit D_3 Injecção de Hilactina	Prevenção e tratamento de hipocalcemia, terapia de apoio em estados de doença (especialmente hipotermia) e tratamento de fases iniciais de raquitismo.
Suplementos para Ração Animal	Minerais essenciais, Selénio e Vitamina E Bolus de Ginolactina	Tratamento de anestro que causa infertilidade e ovulação retardada em animais leiteiros e cavalos.
	Minerais essenciais, Vitamina E e Iodo Pó De Hilactina	Infertilidade, lactação imprópria, imunidade diminuída, crescimento atrofiado e debilidade.
	Electrólitos Essenciais Com Vitamina C Electra - Pó C	Diarreia, desidratação, antes e após transporte, em temperaturas extremas (elevadas ou baixas) e outros estados de stress.

(continuação)

Classe de Tratamento	Químico	Descrição
	Pyrenox Plus (Diclofenac Sódico + Paracetamol) Bolus, Injecção.	Tratamento de mastite, febre pós cirúrgica, dor e inflamação, prolapso do útero, claudicação e artrite.

Tabela 2. Formulações terapêuticas

Produto	Descrição
Acutrim® (fenilpropanolamina)	Comprimidos supressores do apetite uma vez por dia.
Infusor Baxter®	Para distribuição intravenosa controlada de anticoagulantes, antibióticos, agentes quimioterapêuticos e outros fármacos extensamente utilizados.
Catapres-TTS® (sistema terapêutico transdérmico de clonidina)	Sistema transdérmico uma vez por semana para tratamento da hipertensão.
Covera HS3 (cloridrato de verapamil)	Comprimidos de libertação perlongada de início controlado (COER-24) uma vez por dia para tratamento de hipertensão e angina de peito.
DynaCirc CR® (isradipina)	Comprimidos de libertação prolongada uma vez por dia para tratamento da hipertensão.

(continuação)

Produto	Descrição
Efidac 24® (maleato de clorfeniramina)	Comprimidos de libertação prolongada uma vez por dia para o alívio de sintomas de alergia.
Estraderm® (sistema transdérmico de estradiol)	Sistema transdérmico duas vezes por semana para tratamento de determinados sintomas pós-menopáusicos e prevenir a osteoporose
Glucotrol XL® (glipizida)	Comprimidos de libertação prolongada uma vez por dia, utilizados como um adjuvante à dieta para o controlo de hiperglicemia em doentes com diabetes mellitus não dependentes de insulina.
IVOMEC SR® Bolus (ivermectina)	Sistema de distribuição ruminal para controlo sazonal de principais parasitas internos e externos no gado.
Minipress XL® (prazosina)	Comprimidos de libertação prolongada uma vez por dia para tratamento da hipertensão.
NicoDerm® CQ™ (sistema transdérmico de nicotina)	Sistema transdérmico utilizado uma vez por dia como auxiliar à cessação de tabagismo para alívio de sintomas da retirada de nicotina.
Procardia XL® (nifedipina)	Comprimidos de libertação prolongada uma vez por dia para tratamento da angina e hipertensão.
Sudafed® 24 Horas (pseudoefedrina)	Descongestionante nasal uma vez por dia para alívio de constipações, sinusite, febre dos fenos e outras alergias respiratórias.

(continuação)

Produto	Descrição
Transderm-Nitro® (sistema transdérmico de nitroglicerina)	Sistema transdérmico uma vez por dia para prevenção da angina de peito devido a doença arterial coronária.
Transderm Scop® (sistema transdérmico de escopolamina)	Sistema transdérmico para a prevenção de náuseas e vómitos associados com a cinetose.
Volmax (albuterol)	Comprimidos de libertação prolongada para o alívio de broncospasmos em doentes com doença obstrutiva reversível das vias aéreas.
Actisite®	(cloridrato de tetraciclina) fibra peridental utilizada como um adjuvante para raspagem e polimento radicular para redução das bolsas peridentais e hemorragia pós-sondagem em doentes com periodontite do adulto.
ALZET®	Bombas osmóticas para investigação laboratorial.
Amphotec® (complexo de colesteril-sulfato de anfotericina B para injecção)	AMPHOTEC® é um tratamento fungicida para aspergilose invasiva em doentes em que deterioração renal ou toxicidade inaceitável impossibilitam a utilização de anfotericina B em doses eficazes e em doentes com aspergilose invasiva onde terapia prévia com anfotericina B falhou.
BiCitra® (citrato de sódio e ácido cítrico)	Agente alcalinizante utilizado naqueles estados em que é desejável a manutenção a longo prazo de urina alcalina.

(continuação)

Produto	Descrição
Ditropan® (cloreto de oxibutinina)	Para alívio de sintomas de instabilidade da bexiga associada a bexiga neurogénica não inibida ou neurogénica de reflexo (*i. e.*, urgência, frequência, perdas urinarias, incontinência de urgência, disúria).
Ditropan® XL (cloreto de oxibutinina)	é um comprimido de libertação controlada uma vez por dia, indicado para tratamento da bexiga hiperactiva com sintomas de incontinência urinária de urgência, urgência e frequência.
DOXIL® (injecção lipossomal de doxorrubicina HCl)	
Duragesic® (sistema transdérmico de fentanilo) CII	Sistema transdérmico de 72 horas para gestão da dor crónica em doentes que requerem analgesia opióide contínua para a dor que não pode ser controlada com meios menores, tais como combinações de acetaminofeno-opióide, analgésicos não esteróides ou dosagem PRN com opióides de curta acção.
Elmiron® (polisulfato sódico de pentosano)	Indicado para o alívio da dor ou desconforto da bexiga associado com cistite intersticial.
ENACT AirWatch™	Um sistema de monitorização e gestão da asma.

(continuação)

Produto	Descrição
Ethyl® (amifostina)	Indicado para reduzir a toxicidade renal cumulativa associada com a administração repetida de cisplatina em doentes com cancro do ovário avançado ou cancro das células não pequenas do pulmão. Indicado para reduzir a incidência de xerostomia moderada a grave em doentes submetidos a tratamento de radiação pós-operatório para o cancro da cabeça e garganta, em que a área de tratamento com radiação inclui uma porção substancial de glândulas parótidas.
Mycelex® Troche (clotrimazole)	Para o tratamento local de candidíase orofaríngea. Indicado também profilacticamente para reduzir a incidência de candidíase orofaríngea em doentes imunocomprometidos por estados que incluem quimioterapia, radioterapia ou terapia esteróide utilizada no tratamento de leucemia, tumores sólidos ou transplantação renal.
Neutra-Phos® (fosfato de potássio e sódio)	um suplemento diatético /nutricional

(continuação)

Produto	Descrição
PolyCitra® -K Solução Oral e PolyCitra® -K Cristais (citrato de potássio e ácido cítrico)	Agente alcalinizante útil naqueles estados em que é desejável a manutenção a longo prazo de urina alcalina, tais como em doentes com cálculos de ácido úrico e cistina do tracto urinário, especialmente quando a administração de sais de sódio for indesejável ou contra-indicada
PolyCitra® -K Xarope e LC (tricitratos)	Agente alcalinizante útil naqueles estados em que é desejável a manutenção a longo prazo de urina alcalina, tais como em doentes com cálculos de ácido úrico e cistina do tracto urinário.
Progestaseet® (progesterona)	Sistema contraceptivo intra-uterino de progesterona
Testoderm® Testoderm® com Adesivo e Testoderm® TTS CIII	Sistema transdérmico de testosterona Os produtos Testoderm® são indicados para terapia de substituição em homens para estados associados com uma deficiência ou ausência de testosterona endógena: (1) hipogonadismo primário (congénito ou adquirido) ou (2) hipogonadismo hipogonadotrófico (congénito ou adquirido).
Viadur™ (implante de acetato de leuprolida)	Implante anual para tratamento paliativo do cancro da próstata

Determinados agentes podem ser concebidos para se tornarem activos ou inactivos sob determinadas condições (*e. g.*, determinados pH, presença de um agente activador, etc.). Além disso, pode ser vantajoso utilizar pro-enzimas nas composições da invenção. Por exemplo, uma pro-enzima pode ser activada por uma protease (*e. g.*, uma protease salivar que esteja presente no tracto digestivo ou seja introduzida artificialmente no tracto digestivo de um organismo). Está contemplado que os agentes distribuídos pelas composições biocompatíveis da invenção são activados ou inactivados pela adição de um agente activador que pode ser ingerido ou, caso contrário, distribuído ao organismo. Um outro mecanismo para controlo do agente no tracto digestivo é um agente sensível ao ambiente que é activado no compartimento digestivo apropriado. Por exemplo, um agente pode estar inactivo a pH baixo, mas activo a pH neutro. Desta forma, o agente estaria inactivo no intestino, mas activo no tracto intestinal. Alternativamente, o agente pode torna-se activo em resposta à presença de um factor específico de microrganismo (*e. g.*, microrganismos presentes no intestino).

Num aspecto, os benefícios potenciais da presente invenção incluem, por exemplo, (1) redução ou possível eliminação da necessidade para suplementos minerais (*e. g.*, suplementos de fósforo inorgânico), enzimas ou fármacos terapêuticos para animais (incluindo peixe) a partir da ração ou grão diário aumentando assim a quantidade de calorias e nutrientes presentes na ração e (2) saúde e crescimento aumentados de animais de criação e não domésticos, incluindo, por exemplo, aves aviárias, animais porcinos, bovinos, equinos, caninos e felinos.

Um grande número de enzimas pode ser utilizado nos métodos e composições da presente invenção além das fitases da invenção. Estas enzimas incluem enzimas necessárias para digestão

apropriada de alimentos consumidos, ou para metabolismo, activação ou derivação apropriada de químicos, profármacos ou outros agentes ou compostos distribuídos ao animal através do tracto digestivo. Os exemplos de enzimas que podem ser distribuídas ou incorporadas nas composições da invenção, incluem, por exemplo, enzimas enriquecedoras de ração seleccionadas do grupo consistindo de α-galactosidases, β-galactosidases, em particular lactases, fitases, β-glucanases, em particular endo-β-1,4-glucanases e endo-β-1,3(4)-glucanases, celulases, xilosidases, galactanases, em particular arabinogalactano-endo-1,4-β-galactosidases e arabinogalactan-endo-1,3-β-galactosidases, endoglucanases, em particular endo-1,2-β-glucanase, endo-1,3-α-glucanase e endo-1,3-β-glucanase, enzimas que degradam a pectina, em particular pectinases, pectinesterases, liases de pectina, poligalacturonases, arabinanases, ramnogalacturonases, ramnogalacturonano acetilo esterases, ramnogalacturonano-α-ramnosidase, liases de pectato e α-galacturonisidases, mananases, β-manosidases, manano acetilo esterases, xilano acetilo esterases, proteases, xilanases, arabinoxilanases e enzimas lipolíticas, tais como lipases, fitases e cutinases. Podem ser utilizadas nos métodos e composições da invenção fitases além das fitases tendo uma sequência de aminoácidos como apresentada em SEQ ID N°:2.

Num aspecto, a enzima utilizada nas composições (*e. g.*, um auxiliar dietético) da presente invenção é uma enzima fitase que é estável ao calor e termorresistente e catalisa a hidrólise enzimática de fitato, *i. e.*, a enzima é capaz de renaturar e readquirir actividade após um período breve (*i. e.*, 5 a 30 segundos) ou mais longo, por exemplo, minutos ou horas, de exposição a temperaturas acima de 50 °C.

Uma “ração” e “alimento”, respectivamente, significa qualquer dieta, refeição natural ou artificial, ou semelhantes, ou componentes de tais farinhas pretendidas ou adequadas para serem comidas, tomadas, digeridas, por um animal e um humano, respectivamente. “Auxiliar dietético”, como aqui utilizado, refere-se, por exemplo, a uma composição contendo agentes que proporcionam um agente terapêutico ou digestivo a um animal ou organismo. Um “auxiliar dietético” tipicamente não é uma fonte da entrada calórica para um organismo, por outras palavras, um auxiliar dietético tipicamente não é uma fonte de energia para o organismo, mas em vez disso é uma composição que é tomada para além da “ração” ou do “alimento” típico.

Em vários aspectos da invenção, são proporcionadas composições de ração que compreendem uma proteína recombinante de fitase tendo, pelo menos, trinta aminoácidos contíguos de uma proteína que tem uma sequência de aminoácidos da SEQ ID N°:2 modificada, como acima mencionado; e um género alimentício contendo fitato. Como será conhecido dos especialistas na técnica, tais composições podem ser preparadas numa variedade de modos, incluindo mas não limitada a, em forma de pastilha com ou sem aditivos revestidos poliméricos, em forma de granulado e por secagem por spray. A título de exemplo não limitativo, os ensinamentos na técnica dirigidos à preparação de ração incluem as Publicações Internacionais N° WO0070034 A1, WO0100042 A1, WO0104279 A1, WO0125411 A1, WO0125412 A1 e o documento EP 107334À.

Um agente ou enzima (*e. g.*, uma fitase) pode exercer o seu efeito *in vitro* ou *in vivo*, *i. e.*, antes da entrada ou no estômago ou moela do organismo, respectivamente. Também é possível uma acção combinada.

Embora qualquer enzima possa ser incorporada num auxiliar dietético, é feita aqui referência à fitase como uma exemplificação dos métodos e composições da invenção. Um auxiliar dietético da invenção inclui uma enzima (*e. g.*, uma fitase). Geralmente, um auxiliar dietético contendo uma composição de fitase é líquido ou seco.

As composições líquidas não necessitam conter mais nada para além da enzima (*e. g.*, uma fitase), de um modo preferido, numa forma altamente purificada. Geralmente, contudo, é também adicionado um estabilizador, tais como glicerol, sorbitol ou monopropilenoglicol. A composição líquida pode também compreender outros aditivos, tais como sais, açúcares, conservantes, agentes de ajuste de pH, proteínas, fitato (um substrato de fitase). As composições líquidas típicas são suspensões aquosas ou à base de óleo based. As composições líquidas podem ser adicionadas a uma composição biocompatível para libertação lenta. De um modo preferido, a enzima é adicionada a uma composição de auxiliar dietético que seja um material biocompatível (*e. g.*, biodegradável ou não biodegradável) e inclua a adição de células recombinantes, por exemplo, em microesferas porosas.

As composições secas podem ser composições secas por pulverização, caso esse em que a composição não necessita conter nada mais do que a enzima numa forma seca. Geralmente, contudo, as composições secas são denominadas granulados que podem ser prontamente misturadas com componentes de ração ou alimento ou, de um modo mais preferido, formam um componente de uma pré-mistura. O tamanho de partícula do granulado enzimático é, de um modo preferido, compatível com aquele de outros componentes da mistura. Isto proporciona meios seguros e convenientes de incorporar enzimas na ração animal. Os

granulados da invenção podem ser biocompatíveis ou podem ser granulados biocompatíveis que não são biodegradáveis.

Os granulados de aglomeração da invenção revestidos com uma enzima podem ser preparados utilizando técnica de aglomeração num misturador de elevado cisalhamento. Os granulados de absorção são preparados ao ter núcleos de um material veículo para absorver/para ser revestido pela enzima. Num aspecto, o material veículo é um material não biodegradável biocompatível que simule o papel das pedras ou saibro na moela de um animal. Os materiais típicos de enchimento utilizados em técnicas de aglomeração incluem sais, tal como sulfato dissódico. Outros enchimentos são o caulino, talco, aluminossilicato de magnésio e fibras de celulose. Opcionalmente, aglutinantes, tal como dextrinas, são incluídos também em granulados de aglomeração. Os materiais veículo podem ser quaisquer materiais biocompatíveis, incluindo materiais biodegradáveis e não biodegradáveis (*e. g.*, rochas, pedras, cerâmicos, vários polímeros). Num aspecto, os granulados são revestidos com uma mistura de revestimento. Tal mistura compreende agentes de revestimento, *e. g.*, agentes de revestimento hidrofóbico, tais como óleo de palma hidrogenado e sebo bovino e, se desejado, outros aditivos, tais como carbonato de cálcio ou caulino.

Num aspecto, as composições auxiliares dietéticas (*e. g.*, composições auxiliares dietéticas de fitase) podem conter outros substituintes, tais como agentes de coloração, compostos aromáticos, estabilizadores, vitaminas, minerais, outras enzimas de enriquecimento de ração ou alimento. Num aspecto, um aditivo utilizado numa composição da invenção compreende um ou mais compostos, tais como vitaminas, minerais ou enzimas enriquecedoras de ração e veículos e/ou excipientes adequados.

Num aspecto, as composições auxiliares dietéticas da invenção compreendem ainda uma quantidade eficaz de uma ou mais enzimas enriquecedoras de ração, em particular, enzimas enriquecedoras de ração seleccionadas do grupo consistindo de α-galactosidases, β-galactosidases, em particular lactases, outras fitases, β-glucanases, em particular endo-β-1,4-glucanases e endo-β-1,3(4)-glucanases, celulases, xilosidases, galactanases, em particular arabinogalactano endo-1,4-β-galactosidases e arabinogalactano-endo-1,3-β-galactosidases, endoglucanases, em particular endo-1,2-β-glucanase, endo-1,3-α-glucanase e endo-1,3-β-glucanase, enzima que degradam a pectina, em particular pectinases, pectinesterases, liases de pectina, poligalacturonases, arabinanases, ramnogalacturonases, ramnogalacturonano acetilo esterases, ramnogalacturonan-α-ramnosidase, liases de pectato e α-galacturonisidases, mananases, β-manósidos, manana acetilo esterases, xilano acetilo esterases, proteases, xilanases, arabinoxilanases e enzimas lipolíticas, tais como lipases, fitases e cutinases.

O auxiliar dietético animal da invenção é suplementado ao animal mono-gástrico antes ou simultaneamente com a dieta. Num aspecto, o auxiliar dietético da invenção é suplementado ao animal mono-gástrico simultaneamente com a dieta. Noutro aspecto, o auxiliar dietético é adicionado à dieta na forma de um granulado ou de um líquido estabilizado.

Uma quantidade eficaz de uma enzima num auxiliar dietético da invenção é desde cerca de 10-20000; desde cerca de 10 a 15000, desde cerca de 10 a 10000, desde cerca de 100 a 5000 ou desde cerca de 100 a cerca de 2,000 FYT/kg de auxiliar dietético.

Os exemplos não limitantes de outras utilizações específicas de fitase da invenção são no processamento de soja e no fabrico de inositol ou os seus derivados.

A invenção também se refere a um método para reduzir os níveis de fitato nas fezes animais, em que o animal é alimentado com uma auxiliar dietético contendo uma quantidade eficaz de fitase da invenção. Como referido no início do presente pedido, um seu importante efeito é reduzir a poluição de fosfato do ambiente.

Noutro aspecto, o auxiliar dietético é um veículo magnético. Por exemplo, um veículo magnético contendo uma enzima (*e. g.*, uma fitase) distribuída dentro, sobre ou através de um veículo magnético (*e. g.*, um grânulo magnético poroso) que pode ser distribuído sobre uma área elevada em fitato e recolhido por magnetes após um período de tempo. Tais distribuição e recolha de grânulos reduzem adicionalmente a poluição e permitem a reutilização dos grânulos. Além disso, a utilização de tais grânulos magnéticos *in vivo* permite a localização do auxiliar dietético num ponto no tracto digestivo onde, por exemplo, a actividade de fitase pode ser realizada. Por exemplo, um auxiliar dietético da invenção contendo enzimas digestivas (*e. g.*, uma fitase) pode ser localizado na moela do animal por justaposição de um magnete perto d moela do animal após o animal consumir um auxiliar dietético de veículos magnéticos. O magnete pode ser removido após um período de tempo que permita o auxiliar dietético passar através do tracto digestivo. Além disso, os veículos magnéticos são adequados para remoção do organismo após sacrifício ou para auxiliar na recolha.

Quando o auxiliar dietético é uma partícula porosa, tais partículas são tipicamente impregnadas por uma substância com a

qual é desejado libertar lentamente para formar uma partícula de libertação lenta. Tais partículas de libertação lenta podem ser não apenas preparadas por impregnação de partículas porosas com a substância que é desejada à libertação, mas também por primeiro dissolver a substância desejada na primeira fase de dispersão. Neste caso, as partículas de libertação lenta preparadas pelo método em que a substância para ser libertada é primeiro dissolvida na primeira fase de dispersão está também dentro do âmbito e espírito da invenção. As partículas ocas porosas podem, por exemplo, ser impregnadas por uma substância de libertação lenta, tais como um medicamento, químico agrícola ou enzima. Em particular, quando as partículas ocas porosas impregnadas por uma enzima são feitas de polímeros biodegradáveis, as partículas elas mesmas podem ser utilizadas como um químico agrícola ou fertilizante, e não têm efeitos adversos no ambiente. Num aspecto, as partículas porosas são magnéticas na natureza.

As partículas ocas porosas podem ser utilizadas como um suporte de biorreactor, em particular um suporte de enzima. Deste modo, é vantajoso preparar o auxiliar dietético utilizando um método de libertação lenta, por exemplo, por encapsulação da enzima do agente numa microvesícula, tal como um lipossoma, a partir do qual a dose é libertada durante o período de diversos dias, de um modo preferido, entre cerca de 3 a 20 dias. Alternativamente, o agente (*e. g.*, uma enzima) pode ser formulado para libertação lenta, tal como a incorporação num polímero de libertação lento a partir do qual a dosagem do agente (*e. g.*, enzima) é lentamente libertada durante o período de diversos dias, por exemplo, de 2 a 30 dias e pode durar a vida do animal.

Num aspecto, os lipossomas da invenção são derivados de fosfolípidos ou outras substâncias lipídicas. Os lipossomas são formados por cristais líquidos hidratados mono- ou multilamelares que são dispersos num meio aquoso. Pode ser utilizado qualquer lípido não tóxico metabolizável e fisiologicamente aceitável capaz de formar lipossomas. As composições da invenção na forma de lipossoma podem conter estabilizadores, conservantes, excipientes e semelhantes, para além do agente. Alguns lípidos exemplificativos são os fosfolípidos e as fosfatidilcolinas (lecitinas), naturais e sintéticas. Os métodos para formar lipossomas são conhecidos na técnica. Ver, por exemplo, Prescott, Ed., Methods in Cell Biology, Volume XIV, Academic Press, Nova Iorque, N.I. (1976), p. 33 et seq.

Está também dentro do âmbito da invenção a utilização de uma fitase da invenção durante a preparação de alimentos ou preparações de ração ou aditivos, *i. e.*, a fitase exerce a sua actividade de fitase durante apenas o fabrico e não é activa no produto final de alimento ou ração. Este aspecto é relevante, por exemplo, na preparação de massa de pão e cozedura. Desta forma, a fitase ou levedura recombinante que expressa a fitase pode ser impregnada dentro, sobre ou através de veículos magnéticos, distribuídos na massa de pão ou alimento, e recuperados por magnetes.

O auxiliar dietético da invenção pode ser administrado isoladamente a animais num veículo biocompatível (*e. g.*, biodegradável ou não biodegradável) ou em combinação com outros agentes aditivos de digestão. O auxiliar dietético da invenção pode assim ser prontamente administrado como uma cobertura superficial ou através da sua mistura directamente na ração animal ou proporcionado separadamente da ração, por dosagem oral

separada, por injecção ou meios transdérmicos ou em combinação com outros compostos comestíveis relacionados com o crescimento, senda as proporções de cada um dos compostos em combinação dependentes do organismo particular ou problema que está a ser abordado e do grau de resposta desejado. Deve ser entendido que a dosagem dietética específica administrada em cada caso será ajustada de acordo com os compostos específicos a serem administrados, o problema a ser tratado, o estado do indivíduo e de outros factos relevantes que podem modificar a actividade do ingrediente eficaz ou a resposta do indivíduo, como é bem conhecido dos especialistas na técnica. Em geral, pode ser utilizada uma dose única diária ou dosagens diárias divididas, como é bem conhecido na técnica.

Se administradas separadamente da ração animal, podem ser preparadas formas do auxiliar dietético através da sua combinação com veículos comestíveis farmaceuticamente aceitáveis não tóxicos para fazer formulações de libertação imediata ou libertação lenta, como é bem conhecido na técnica. Tais veículos comestíveis podem ser sólidos ou líquidos, tais como, por exemplo, amido de milho, lactose, sacarose, flocos de soja, óleo de amendoim, azeite, óleo de sésamo e propilenoglicol. Se um veículo sólido for utilizado na forma de dosagem dos compostos, estes podem ser comprimidos, cápsulas, pós, trociscos ou pastilhas ou coberturas superficiais como formas micro-dispersíveis. Se um veículo líquido for utilizado, cápsulas moles de gelatina ou xarope ou suspensões, emulsões ou soluções líquidas podem ser a forma de dosagem. As formas de dosagem podem também conter adjuvantes, tais como agentes conservantes, estabilizadores, molhante ou emulsionante, promotores de solução, etc. Podem também conter outras substâncias terapeuticamente valiosas. Um processo para preparar um veículo comestível granulado a alta temperatura para

libertação da enzima quando ingerido, é descrito no Pedido de Patente U.S copendente N° de série 09/910579, apresentado em 20 de Julho de 2001.

Em formas de realização alternativas, as vantagens significativas da invenção podem incluir

1) facilidade de fabrico das composições biocompatíveis carregadas com ingrediente activo;
2) versatilidade, uma vez que se refere com a classe de polímeros e/ou ingredientes activos que podem ser utilizados;
3) rendimentos e eficiências de carregamento mais elevadas; e 4) a provisão de formulações de libertação sustida que libertam agentes activos intactos activos *in vivo*, proporcionando assim libertação controlada de um agente activo durante um período de tempo prolongado. Numa forma de realização, uma vantagem pode ser devida à distribuição local de agente dentro do tracto digestivo (*e. g.*, a moela) do organismo. Num aspecto, a frase "contida dentro" refere-se a um método para formular um agente numa composição útil para libertação controlada, durante um período de tempo prolongado do agente.

Em formas de realização alternativas das composições de libertação sustida ou libertação lenta da invenção, é utilizada uma quantidade eficaz de um agente (*e. g.*, uma enzima ou antibiótico). Num aspecto, a libertação sustida ou libertação lenta refere-se à libertação gradual de um agente a partir de um material biocompatível, durante um período de tempo prolongado. A libertação sustida pode ser contínua ou descontínua, linear ou não linear, e pode ser realizada utilizando uma ou mais composições biodegradáveis ou não biodegradáveis, cargas de

fármacos, selecção de excipientes ou outras modificações. Contudo, é para ser reconhecido que pode ser desejável proporcionar uma composição de libertação "rápida" que proporcione libertação rápida uma vez consumida pelo organismo. É também para ser entendido que a "libertação" não significa necessariamente que o agente é libertado do veículo biocompatível. Em vez disso, num aspecto, a libertação lenta abrange a activação lenta ou a activação contínua de um agente presente na composição biocompatível. Por exemplo, uma fitase não necessita ser libertada da composição biocompatível para ser eficaz. Neste aspecto, a fitase é imobilizada na composição biocompatível.

A ração animal pode ser qualquer refeição orgânica contendo proteína normalmente utilizada para satisfazer os requisitos dietéticos de animais. Muitas dessas farinhas contendo proteína são principalmente compostas tipicamente de farinha de milho, soja ou uma mistura de farinha de milho/soja. Por exemplo, os produtos típicos comercialmente disponíveis dados a aves de capoeira incluem Egg Maker Complete, um produto alimentar para aves de capoeira da Land O'Lakes AG Services, assim como Country Game and Turkey Grower um produto da Agwa, Inc. (ver também The Emu Farmer's Handbook por Phillip Minnaar e Maria Minnaar). Ambos estes produtos comercialmente disponíveis são exemplos típicos de rações animais com as quais o presente auxiliar dietético e/ou a fitase enzimática podem ser incorporados para reduzir ou eliminar a quantidade de entrada suplementar de fósforo, zinco, manganês e ferro requerida em tais composições.

A invenção proporciona novas formulações e suplementos dietéticos e aditivos, e métodos para suplementação dietética de determinadas dietas, *e. g.*, dieta de Atkins, dieta vegetariana, dieta macrobiótica, dieta vegan ou dietas regionais, *e. g.*,

dietas do mundo em desenvolvimento. Os alimentos associados com determinadas dietas electivas, tais como as dietas de Atkins, vegetarianas, macrobióticas, vegan ou regionais (por exemplo, dietas do mundo em desenvolvimento) enfatizam determinadas categorias de alimentos, tais como proteínas e gorduras, soja, etc., ou baseiam-se em cultivares indígenas, *e. g.*, cereais, arroz, feijões e semelhantes como contribuintes substanciais ou únicos para a nutrição individual. Muitas destes cultivares com base em cereal tê níveis elevados (3 a 10 vezes) de ácido fítico. Os produtos alimentares processados, tal como hidrolisado de proteína de soja e outros parecem reter níveis elevados de ácido fítico e a sua inclusão como uma fonte de proteína para barras nutrientes, pós e outros alimentos ou suplementos alimentares e ingredientes, aumenta a carga de ácido fítico sentida pelos indivíduos que praticam estas dietas.

Prevenindo e invertendo a perda óssea

A invenção também proporciona novas formulações farmacêuticas e dietéticas para serem utilizadas como suplementos e aditivos, e métodos para suplementação dietética, compreendendo fitases, *e. g.*, qualquer fitase, incluindo uma fitase da invenção, para indivíduos predispostos a perda óssea, indivíduos com perda óssea e indivíduos com determinados estados clínicos, *e. g.*, osteoporose, caquexia e tratamentos médicos, tais como quimioterapias, as quias podem comprometer a incorporação apropriada ou a utilização de nutrientes essenciais. Os métodos e composições da invenção podem ser utilizados isoladamente ou em combinação com outros suplementos ou regimes de tratamento, incluindo com medicações e semelhantes. Por exemplo, as formulações, suplementos dietéticos e métodos para a suplementação dietética podem ser administrados

com outros suplementos dietéticos ou medicações para o tratamento ou prevenção da osteoporose, *e. g.*, com vitamina D3 e/ou cálcio (que foram provados na prevenção da perda óssea). Num aspecto, a invenção proporciona uma formulação compreendendo uma fitase, *e. g.*, qualquer fitase ou uma fitase da invenção, e vitamina D3 e/ou cálcio. Num aspecto, a invenção proporciona uma formulação compreendendo uma fitase, *e. g.*, qualquer fitase ou uma fitase da invenção, para prevenir a perda óssea. Num aspecto, a invenção proporciona uma formulação compreendendo uma fitase da invenção, para inverter a perda óssea.

A formulação pode ser na forma de uma composição farmacêutica, ou pode ser um aditivo para um produto farmacêutico, cada um dos quais pode ser em forma líquida, sólida, pó, loção, pulverizador ou aerossol. As composições farmacêuticas e as formulações da invenção para administração oral podem ser formuladas utilizando veículos farmaceuticamente aceitáveis bem conhecidos na técnica, em dosagens apropriadas e adequadas. Tais veículos permitem os produtos farmacêuticos de ser formulados em forma de dosagem unitária como comprimidos, pílulas, pós, drageias, cápsulas, líquidos, pastilhas, géis, xaropes, pastas, suspensões, etc., adequadas para ingestão pelo doente. As preparações farmacêuticas para utilização oral podem ser formuladas como um excipiente sólido, opcionalmente, moendo uma mistura resultante e processando a mistura de grânulos, após adicionar compostos adicionais adequados, se desejado, para obter comprimidos ou núcleos de drageia. Os excipientes sólidos adequados são enchimentos de hidratos de carbono ou proteína incluindo, *e. g.*, açúcares, incluindo lactose, sacarose, manitol ou sorbitol; amido de milho, trigo, arroz, batata ou outras plantas; celulose, tais como metilcelulose, hidroxipropilmetil-celulose ou carboxi-metilcelulose de sódio; e gomas incluindo arábica e tragacanto; e proteínas, *e. g.*, gelatina e colagénio.

Podem ser adicionados agentes de desintegração e solubilizantes, tais como polivinilpirrolidona reticulada, ágar, ácido algínico ou um seu sal, como alginato de sódio.

A invenção proporciona suspensões aquosas compreendendo uma fitase da invenção em mistura com excipientes adequados para o fabrico de suspensões aquosas. Tais excipientes incluem um agente de suspensão, tais como carboximetilcelulose de sódio, metilcelulose, hidroxipropilmetilcelulose, alginato de sódio, polivinilpirrolidona, goma tragacanto e goma de acácia, e agentes molhantes e de dispersão, tais como uma fosfatida de ocorrência natural (*e. g.*, lecitina), um produto de condensação de óxido de alquileno com um ácido gordo (*e. g.*, estearato de polioxietileno), um produto de condensação de óxido de etileno com um álcool alifático de cadeia longa (*e. g.*, heptadecaetileno-oxicetanol), um produto de condensação de óxido de etileno com um éster parcial derivado de ácido gordo e um anidrido hexitol (*e. g.*, mono-oleato de polioxietileno e sorbitol) ou um produto de condensação de óxido de etileno com um éster parcial derivado de ácido gordo e um anidrido hexitol (*e. g.*, mono-oleato de polioxietileno e sorbitano). A suspensão aquosa pode também conter um ou mais conservantes, tais como p-hidroxibenzoato de etilo ou n-propilo, um ou mais agentes corantes, um ou mais agentes aromatizantes e um ou mais agentes edulcorantes, tais como sacarose, aspartame ou sacarina. As formulações podem ser ajustadas para a osmolaridade.

O regime de dosagem toma também em consideração parâmetros farmacocinéticos bem conhecidos na técnica, *i. e.*, a taxa de absorção dos agentes activos, biodisponibilidade, metabolismo, eliminação e semelhantes (ver, *e. g.*, Hidalgo-Aragones (1996) J. Steroid Biochem. Mol. Biol. 58:611-617; Groning (1996) Pharmazie 51:337-341; Fotherby (1996) Contraception 54:59-69; Johnson

(1995) J. Pharm. Sci. 84:1144-1146; Rohatagi (1995) Pharmazie 50:610-613; Brophy (1983) Eur. J. Clin. Pharmacol. 24:103-108; a última edição de Remington, The Science and Practice of Pharmacy 20ª Ed. Lippincott Williams & Wilkins). O estado da técnica permite que o clínico determine o regime de dosagem para cada doente individual, agente activo e doença ou estado tratado. As orientações proporcionadas para composições semelhantes utilizadas como produtos farmacêuticos podem ser utilizadas para determinar o regime de dosagem, *i. e.*, programação de dose e níveis de dosagem, administrados na prática dos métodos da invenção (*e. g.*, revertendo a perda óssea ou prevenindo a perda óssea) são apropriadas e correctas.

Suplementos físicos ao treino

A invenção também proporciona novos suplementos e aditivos dietéticos, e métodos para a sua utilização, compreendendo fitases da invenção para os indivíduos que se submetem a treino físico atlético intenso ou outro, *e. g.*, treino para soldados. O treino atlético e hiperdispendioso pode esgotar nutrientes essenciais e requer suplementação dietética. Estas dietas e condições têm em comum uma falta de micronutrientes essenciais, tais como metais (K, Ca, Fe, Zn, Mn, Se) e iões (PO_4) necessários para nutrição óptima. As dietas ricas em ácido fítico exacerbam este problema e podem também conduzir a estados crónicos e agudos que resultam da dependência voluntária ou economicamente reforçada em dietas ricas em alimentos com ácido fítico elevado.

Por exemplo, os indivíduos que seguem várias dietas baixas em hidratos de carbono (“low carb”) são, frequentemente, flagelados com cãibras musculares, *e. g.*, músculo da perna. O conselho típico para isto é adicionar mais potássio, cálcio e

outros nutrientes à sua dieta. A invenção proporciona composições para suplementação dietética, auxiliares dietéticos e suplementos e métodos para suplementação dietética para enriquecer a nutrição, caso contrário comprometida, através da mobilização de macro e micronutrientes utilizando a suplementação de fitase à dieta (por utilização de uma fitase da invenção).

Num aspecto da invenção, a utilização de uma fitase da invenção é optimizada para demonstrar perfis termo-lábeis ou de estabilidade de pH que a tornarão adequada para adição directa ao alimento e processo de suplementação e/ou demonstra estabilidade e actividade aumentadas no tracto gastrointestinal humano ou animal.

A invenção também proporciona novos suplementos e aditivos dietéticos, e métodos para a sua utilização, compreendendo fitases da invenção para indivíduos que se submetem a suplementação mineral. A suplementação mineral para pessoas que consomem alimentos com elevado teor de ácido fítico pode de facto exacerbar problemas com disponibilidade de nutrientes. As referências da literatura sugerem que complexos de ácido fítico, cálcio e zinco são muito mais insolúveis que complexos de ácido fítico e cálcio. As pessoas tomam frequentemente suplementos multiminerais. A adição de fitase a um esquema planeado para combinar suplementos minerais na presença de alimentos com elevado ácido fítico pode tornar estes suplementos muito mais eficazes.

Em aspectos alternativos, as composições e métodos da invenção compreendendo uma fitase da invenção, são utilizados como suplementos ou aditivos para

- Programas de perda de peso que limitam a entrada de grupos de alimentos particulares; dietas vegetarianas, macrobióticas ou vegan que limitam ou impossibilitam o consumo de carnes, vegetais da família das solanáceas, pães, etc. e outras dietas focadas no consumo de frutos secos,
- Suplemento específico para indivíduos em dietas "low carb" ricas em alimentos com elevado ácido fítico para aliviar os sintomas fisiológicos baseados em incorporação mineral reduzida,
- Regimes de treino atlético que pretendem aumentar o desempenho através do consumo dietético, incluindo regimes de treino militar,
- Dietas hospitalares à medida de necessidades específicas de doentes comprometidos no consumo ou restringidos a grupos de alimentos
- Dietas de cereais e leguminosas pobres em micronutrientes no mundo em desenvolvimento,
- Programas de almoço escolar.

A invenção também proporciona kits compreendendo composições da invenção compreendendo uma fitase da invenção e instruções na incorporação da composição ou método da invenção nestas dietas. Os kits podem compreender embalagem, rótulos, folhetos informativos e semelhantes.

Num aspecto, a invenção proporciona uma fitase da invenção formulada ou optimizada para (*e. g.*, sequência optimizada para) produção, processamento ou passagem através do sistema humano ou animal, *e. g.*, tracto digestivo. A enzima fitase pode ser optimizada utilizando formulações alternativas.

Alternativamente, uma enzima fitase da invenção pode ser optimizada por manipulação da sua sequência, *e. g.*, utilizando, por exemplo, evolução dirigida, PCR propenso ao erro, arranjo combinatório, mutagénese dirigida por oligonucleótido, PCR de montagem, mutagénese por PCR sexual, em mutagénese *in vivo*, mutagénese por cassete, mutagénese de montagem recursiva, mutagénese de montagem exponencial, mutagénese específica de local, remontagem de ligação, GSSM™ e qualquer sua combinação, para reter a actividade durante o processamento, ingestão e dentro do intestino humano.

As composições (*e. g.*, formulações dietéticas compreendendo uma enzima fitase da invenção) podem ser distribuídas de uma variedade de modos para proporcionar eficácia dietética. Por exemplo, a invenção proporciona composições (*e. g.*, formulações dietéticas ou aditivos compreendendo uma enzima fitase da invenção) e métodos compreendendo a utilização de:

- Em suplementos de alimentos empacotados, tais como comprimidos mastigáveis ou barras nutritivas,
- Como um produto liofilizada disponível para hidratação antes da ingestão,
- Co-empacotado com produtos dietéticos, *e. g.*, produto processado de soja ou vendido como uma formulação com hidrolisado de proteína de soja e outras fracções de processamento a partir de alimentos inteiros que são vendidos como ingredientes à indústria de alimentos processados,
- Em produtos comerciais cozidos,
- Spray para cereais de pequeno almoço,
- Formulações administradas com Spray (*e. g.*, Spray nasal),
- Como um produto transgénico expresso em cultivares indígenas, *i. e.*, em cereais e leguminosas (*e. g.*, como um

produto transgénico de um microrganismo, tal como uma bactéria)

- Como um organismo transgénico, *e. g.*, um microrganismo; por exemplo, um humano ou animal é alimentado com bactérias ou outro microrganismo capaz de fazer (e, numa forma de realização alternativa, segregar) uma fitase recombinante da invenção, após o ingestão ou implantação, *e. g.*, no intestino do humano ou animal.

Os produtos contendo fitase e métodos da invenção podem ser marcados como enriquecidos com nutrientes, compatíveis com nutrientes ou, caso contrário, assinalados pela capacidade para enriquecer o desempenho nutritivo e aliviar vários sintomas associados com a deficiência de nutrientes.

Os produtos contendo fitase e métodos da invenção são utilizados para mitigar os efeitos anti-nutritivos do fitato, que quelata importantes minerais dietéticos, tais como zinco, cobre, ferro, magnésio, estanho e cálcio. Desta forma, os produtos contendo fitase e métodos da invenção são utilizados como suplementos dietéticos para prevenir a precipitação de enzima e proteínas de ligação a metal em alimentos ingeridos. Num aspecto, os produtos contendo fitase e métodos da invenção são utilizados para mitigar os efeitos anti-nutritivos do fitato em dietas humanas, em particular, aquelas ricas em leguminosas e cereais, para aumentar a biodisponibilidade mineral. Num aspecto, uma fitase num suplemento dietético da invenção catalisa a remoção hidrolítica parcial ou completa do ortofosfato de um fitato, em que a hidrólise completa do fitato resulta na produção de 1 molécula de inositol e 6 moléculas de fosfato inorgânico.

Os produtos contendo fitase e métodos da invenção são aplicáveis à dieta de humanos e numerosos animais, incluindo aves de capoeira e peixes. Por exemplo, os produtos de suplementação dietética contendo fitase e os métodos de suplementação dietética da invenção podem ser praticados com espécies comercialmente significativas, *e. g.*, porcos, gado bovino, carneiros, cabras, roedores de laboratório (ratos, murganhos, hámsteres e gerbilos), animais com pêlo, tais como vison e raposa, e animais de jardim zoológico, tais como macacos e símios, assim como mamíferos domésticos, tais como gatos e cães. As espécies aviárias típicas comercialmente significativas incluem galinhas, perús, patos, gansos, faisões, emas, avestruzes, mergulhões, quivi, pombas, papagaios, caturras, catatuas, canários, pinguins, flamingos e codornizes. Os peixes comercialmente cultivados, tais como a truta também beneficiariam do auxiliar dietético aqui divulgado. Outros peixes que podem beneficiar incluem, por exemplo, peixe (especialmente num ambiente de aquário ou aquacultura, *e. g.*, peixes tropicais), pimpão e outras carpas, peixe-gato, truta, salmão, tubarão, raia, solhas, linguado, tilápia, *medaka*, *guppies*, *mollies*, azevia, espadas, peixe-zebra e verdemão ornamentais.

Os produtos contendo fitase e métodos da invenção são também utilizados em vários ágares, géis, meios e em soluções utilizadas para cultivar tecidos e/ou células. A inconsistência dos hidrolisados de soja pode ser um problema enfrentado quando se utiliza cultura de tecidos e/ou células. Num aspecto, os produtos contendo fitase e métodos da invenção são utilizados como aditivos de meios de cultura de células ou como tratamentos, *e. g.*, para aumentar o rendimento da cultura de células e a consistência do desempenho. Num aspecto, a invenção

proporciona hidrolisados para cultura de células compreendendo fitases, *e. g.*, fitases da invenção.

Num aspecto, para proporcionar um produto consistente, a invenção proporciona métodos para fazer hidrolisados, suplementos ou outros aditivos para a cultura de células compreendendo fitases por utilização de biomarcadores de fitase. Por exemplo, o método compreende "pontuar" ou "marcar" diversas moléculas de fitase em lotes de hidrolisado, suplemento ou outro aditivo, e depois misturar os lotes no hidrolisado, suplemento ou outro aditivo para conseguir um perfil consistente de biomarcador. Num aspecto, o desempenho da cultura com cada lote é medido num mini-biorreactor(es) e o desempenho com cada biomarcador e lote é correlacionado. Num aspecto, uma mistura é feita para gerar um produto de desempenho mais elevado que seja consistente ou melhor do que a média. Num aspecto, tiorredoxina (TRX) é adicionada para aumentar a biodisponibilidade de muitas proteínas por eliminação da estrutura secundária provocada pelas ligações persulfureto. Num aspecto, as proteases são adicionadas também aos hidrolisados, suplementos ou aditivos da invenção. As proteases podem ser "marcadas" ou a sua qualidade controlada com outros biomarcadores (como com fitase, como discutido acima) para dirigir o processo de mistura.

Num aspecto, a invenção proporciona métodos para adicionar fitases a grãos para proporciona um produto consistente utilizando uma "pontuação" de biomarcador ou processo de controlo de qualidade análogo àquele descrito acima para os hidrolisados, suplementos ou aditivos da invenção.

Dietas enriquecidas em enzimas para maior eficiência guerreira e moral

Num aspecto, a invenção proporciona novos suplementos e aditivos dietéticos e métodos para suplementação dietética compreendendo fitases da invenção para dietas enriquecidas em enzimas para maior eficiência e guerreira moral. Num aspecto, estas composições de suplementação dietética da invenção trabalham, *in situ*, para aumentar a energia, vigor e moral num formato estável, facilmente utilizável e desejável enquanto limita o desperdício alimentar.

Num aspecto, estes composições de suplementos dietéticos e métodos da invenção aborda o desafio operacional militar compreendendo distribuição eficiente de nutrientes e a saúde, moral e eficácia operacional associada dos soldados. A invenção proporciona enzimas optimizadas pra funcionar eficientemente num intestino humano. Estas enzimas podem aumentar a extracção de nutrientes e a geração de energia, assim como prolongar a manutenção da suficiência nutritiva e saciedade individual.

Para além de fitases, outras enzimas, *e. g.*, amilases, xilanases, proteases, lipases, são utilizadas para praticar as composições de suplemento dietético e métodos da invenção. Num aspecto, a invenção proporciona formulações, suplementos alimentares, alimentos, unidade Pronto-a-Comer (MRE) de refeição autónoma, bebidas, agentes hidratantes e semelhantes, compreendendo fitase, *e. g.*, uma fitase da invenção, e outra enzima, *e. g.*, amilases, xilanases, proteases, lipases ou uma sua combinação. Quando ingeridas com alimento, foi demonstrado que estas enzimas aumentam a libertação de nutrientes críticos, *e. g.*, fósforo, metais e iões essenciais, aminoácidos e açúcares. Além disso, a co-ingestão destas enzimas aumenta a

mecânica gastrointestinal e a absorção por despolimerização da celulose, hemicelulose e amido derivados de plantas. Este documento propõe o desenvolvimento destas enzimas como suplementos a dietas militares para proporcionar utilização melhorada de nutrientes para guerreiros.

Num aspecto, o suplemento alimentar da invenção provoca a libertação de fosfato essencial a partir de fitato normalmente anti-nutritivo, derivado de plantas, para aumentar o rendimento energético alimentar e a deposição de $caPO_4$ ósseo. Num aspecto, as fitases e outras potenciais enzimas de suplementação nutritiva podem suportar o pH intestinal e actividades de proteases endógenas.

Num aspecto, a invenção proporciona suplementos de enzima a rações, bebidas, alimentos, MRE, agentes hidratantes e semelhantes, para melhorar significativamente o valor nutritivo, digestibilidade e teor energético de refeições militares (ou qualquer refeição, incluindo refeição de consumidor geral e produtos de suplementação dietética) servidas a soldados em treino, batalha ou qualquer situação stressante. O suplemento pode ser formulado para facilidade de utilização e transporte pessoal (em ou com MRE, agentes hidratantes, etc.). Num aspecto, a suplemento de enzima não comprometerá a aparência, gosto e/ou consistência do alimento. Num aspecto, o produto melhora a saúde e aumenta o vigor dos soldados.

Em aspectos alternativos, a suplemento de enzima é distribuído numa variedade de modos para proporcionar eficácia dietética; por exemplo, a invenção proporciona fitases, incluindo fitases da invenção e, em alguns aspectos, enzimas adicionais, em:

- Suplementos empacotados alimentares ou de bebidas, tais como MRE, rações, kits de sobrevivência, agentes hidratantes, comprimidos mastigáveis ou barras nutritivas;
- Como um produto liofilizado (*e. g.*, um pó) disponível para hidratação antes da ingestão;
- Co-empacotados com produtos dietéticos, alimentos, bebidas, *e. g.*, um produto processado de soja ou uma formulação com hidrolisado de proteína de soja e outras fracções de processamento a partir de alimentos inteiros que são vendidos como ingredientes para a indústria de alimentos processados;
- Em bens cozidos;
- Spray para cereais;
- Formulações, tais como comprimidos, *geltabs*, cápsulas, Sprays e semelhantes.

Num aspecto, as composições e métodos da invenção proporcionam suplementação nutritiva que rapidamente liberta calorias e macro- e micronutrientes de refeições ingeridas. Num aspecto, as composições e métodos da invenção proporcionam energia e força corporal a indivíduos em situações stressantes, *e. g.*, envolvendo híper-esforço e períodos descontínuos de privação. Num aspecto, as composições e métodos da invenção proporcionam enzimas optimizadas e formuladas para trabalhar eficazmente no intestino humano, enquanto mantém a estabilidade, estabilidade em armazenamento e transportabilidade num ambiente desejado, *e. g.*, num cenário militar.

Num aspecto, as composições e métodos da invenção proporcionam formulações para características de gosto aumentado, capacidade de dissolução, mastigação e eficiência de transporte pessoal do produto. Num aspecto, as composições e métodos da invenção compreendem ainda outros componentes, tais

como potássio, glucose, $CaCl_2$. O $CaCl_2$ na formulação pode combinar-se com fosfato libertado e, por sua vez, aumentar a deposição óssea e aumento de peso. Num aspecto, as composições e métodos da invenção compreendem ainda formulações de outras enzimas, tais como proteases, celulases, hemicelulases, para digestão de proteína, celulose e hemicelulose, respectivamente. Estas enzimas podem melhorar a disponibilidade de proteína e amido e ainda aumentam a absorção do ferro a partir de muitos alimentos ricos em ferro.

Num aspecto, as composições e métodos da invenção compreendem ainda enzimas para hidrolisar alimentos derivados de material vegetal, que é rico nos polímeros com base em glucose e xilose, celulose, hemicelulose e amido, assim como no polímero de aminoácidos, proteína. Num aspecto, as composições e métodos da invenção facilitam a hidrólise de materiais poliméricos nos alimentos; *i. e.*, para facilitar a digestão completa de polímeros em monómeros, *e. g.*, polissacáridos em açúcares monoméricos, ou proteínas em resíduos de aminoácidos. Assim, neste aspecto, as composições e métodos da invenção permitem um alimento, bebida ou ração para realizar o seu valor calórico e nutritivo total. Num aspecto, a suplementação de enzima compreende a utilização de enzimas estáveis, *e. g.*, hidrolases de vários tipos, celulases, hemicelulases, amilases, lipases, amidases, proteases e outras enzimas. Num aspecto, enzimas utilizadas nas composições e métodos da invenção podem suportar condições ambientais do intestino, *i. e.*, estabilidade a baixo pH e na presença de proteases gástricas.

Utilizações industriais de fitases

Além daquelas descritas acima, a invenção proporciona novas utilizações industriais para fitases da invenção.

Redução de poluição de fosfato no ambiente

Num aspecto, a divulgação proporciona composições compreendendo fitases (incluindo as fitases da invenção) para adição a pilhas de resíduos ou fezes para converter ácido fítico "ambiental". Num aspecto, isto serve os objectivos de reduzir a poluição e aumentar a disponibilidade nutriente. A divulgação também proporciona composições e métodos para adicionar uma fitase ao solo, massas de água naturais ou artificiais (*e. g.*, lagos, lagoas, poços, lagoas de fezes e semelhantes), esgoto municipal, qualquer efluente de esgoto e semelhantes. Como descrito acima, as composições e métodos são para reduzir níveis de fitato em resíduos ou esgoto, *e. g.*, fezes animais, em que o animal é alimentado com um auxiliar dietético contendo uma quantidade eficaz de uma fitase da invenção. Um aplicação exemplificativa das composições da invenção é para reduzir a poluição de fosfato no ambiente. Assim, as composições e métodos da invenção podem ser utilizados em qualquer aplicação que reduz a poluição por degradação de ácidos fíticos.

Aplicações de agricultura e crescimento de plantas

Num aspecto, a divulgação proporciona composições compreendendo fitases (incluindo as fitases da invenção) e métodos para aplicações de cultivo ou outras aplicações de crescimento de plantas, *e. g.*, adicionar fitases a fertilizantes

ou aditivos alimentares de plantas (*e. g.*, MIRACLEGROW™) para plantas, *e. g.*, plantas de interior. Na utilização de composições da invenção para aplicações de cultivo, os utilizadores incluem agricultores orgânicos. As composições da invenção podem ser utilizadas para adicionar fitases a qualquer solo deficiente em fósforo ou que necessita de fósforo suplementar para um cultivar ou aplicação particular. Porque a libertação de fósforo ajuda as plantas a crescerem, as composições da invenção podem ser utilizadas para adicionar fitases a qualquer coisa que tenha nele algas ou material vegetal.

Produtos de fabrico

A divulgação proporciona uma variedade de produtos de fabrico compreendendo uma ou mais fitases desta invenção. Por exemplo, num aspecto, a divulgação proporciona composições compreendendo fitases (incluindo as fitases da invenção) e métodos para aplicações cosméticas, *e. g.*, champôs, loções ou produtos vegetais contendo sabões.

Num aspecto, a invenção proporciona composições compreendendo fitases (incluindo as fitases da invenção) e métodos para imobilizar a fitase. Num aspecto, a fitase imobilizada actua como um mecanismo de libertação controlada. Por exemplo, num aspecto, a divulgação proporciona formulações de libertação controlada (libertação temporizada) de fitases para aplicação a solo, *e. g.*, argila, para plantas interiores, etc. Num aspecto, as fitases são imobilizadas em grânulos, *e. g.*, grânulos de polysorb. Estes grânulos podem ser distribuídos ao solo, *e. g.*, para plantas agrícolas ou interiores. Noutro aspecto, formulações de libertação controlada

(libertação temporizada) de fitases da invenção são utilizadas em suplementos e aditivos dietéticos.

Biocombustíveis e conversão de biomassa

A divulgação proporciona métodos para fazer combustíveis, *e. g.*, biocombustíveis, compreendendo a utilização de uma ou mais fitases desta invenção; incluindo proporcionar combustíveis, *e. g.*, biocombustíveis, compreendendo um ou mais fitases desta invenção. A divulgação proporciona métodos para conversão de biomassa compreendendo a utilização de uma ou mais fitases desta invenção.

Num aspecto, a divulgação proporciona composições compreendendo fitases da invenção e métodos para as utilizar num processo de produção de álcool ou fermentação, *e. g.*, produção de etanol. Por exemplo, as composições da invenção podem ser utilizadas para proporcionar alternativas eficazes e sustentáveis ou adjuvantes para utilização de produtos baseados em petróleo, *e. g.*, como uma mistura de bioetanol e gasolina.

A divulgação proporciona organismos que expressam enzimas da invenção para participação em ciclos químicos que envolvem a conversão de biomassa natural. Além disso, a combinação de fitase (uma enzima desta invenção) com uma ou mais enzimas de degradação de amido, tais como amilase ou glucoamilase, melhora a produção de etanol a partir de amido. A divulgação proporciona métodos para descobrir e implementar as mais eficazes das enzimas para permitir estas importantes novas “conversões de biomassa” e processos industriais energéticos alternativos.

Conversão de biomassa e produção de biocombustíveis limpos

A divulgação proporciona polipéptidos, incluindo enzimas do (fitases da invenção) e anticorpos, e métodos para o processamento de uma biomassa ou qualquer material lenhocelulósico (*e. g.*, qualquer composição compreendendo uma celulose, hemicelulose e lenhina), a um combustível (*e. g.*, um bioetanol, biopropanol, biobutanol, biopropanol, biometanol, biodiesel), para além de rações, alimentos e químicos. Por exemplo, num aspecto, uma enzima da invenção degrada ácido fítico (fitato) indigerível numa biomassa (*e. g.*, um material lenhocelulósico, uma grão ou uma semente oleaginosa) para libertar um fósforo digerível; assim, num exemplo, as fitases desta invenção são utilizados para tratar ou pré-tratar uma biomassa.

Assim, as composições da invenção podem ser utilizadas na produção e/ou processamento de biocombustíveis, *e. g.*, para proporcionar alternativas eficazes e sustentáveis e/ou adjuvantes à utilização de produtos baseados em petróleo; por exemplo, as composições da invenção podem ser utilizadas com uma mistura de enzimas para produzir um biocombustível - tais como biometanol, bioetanol, biopropanol, biobutanol, biodiesel e semelhantes; que pode ser adicionado a um combustível diesel, uma gasolina, um querosene e semelhantes. A divulgação proporciona organismos que expressam enzimas da invenção para participação em ciclos químicos que envolvem a conversão natural de biomassa. Num aspecto, as enzimas para a conversão são utilizadas em conjuntos enzimáticos para o processamento eficiente de biomassa em conjunto com a despolimerização de polímeros de polissacáridos, celulósicos e/ou hemicelulósicos a unidades metabolizáveis (*e. g.*, fermentáveis) de carbono. A divulgação proporciona métodos para identificar e implementar as

mais eficazes das enzimas para permitir estas importantes novas "conversões de biomassa" e processos industriais energéticos alternativos.

As composições da invenção podem ser utilizadas para proporcionar alternativas eficazes e sustentáveis ou adjuvantes para utilização de produtos baseados em petróleo, *e. g.*, como uma mistura de bioetanol, biopropanol, biobutanol, biopropanol, biometanol e/ou biodiesel e gasolina. A invenção proporciona organismos que expressam enzimas da invenção para participação em ciclos químicos que envolvem a conversão natural de biomassa. A divulgação proporciona métodos para descobrir e implementar as mais eficazes das enzimas para permitir estas importantes novas "conversões de biomassa" e processos industriais energéticos alternativos.

São proporcionados métodos, enzimas e misturas de enzimas ou "cocktails" para processar um material, *e. g.*, um material de biomassa, *e. g.*, composições compreendendo um oligossacárido, um oligómero de arabinoxilano, uma lenhina, uma lenhocelulose, um xilano, um glucano, uma celulose e/ou um açúcar fermentável; *e. g.*, incluindo métodos compreendendo colocar em contacto a composição com um polipéptido da invenção, ou um polipéptido codificado por um ácido nucleico da invenção, em que, opcionalmente, o material é derivado de um cultivar agrícola (*e. g.*, trigo, cevada, batatas, *switchgrass*, madeira de álamo), é um subproduto de um alimento ou uma produção de ração, é um produto residual lenhocelulósico ou é um resíduo vegetal ou um produto residual de papel ou resíduo de papel e, opcionalmente, o resíduo vegetal caules, folhas, vagens, cascas, milho ou espigas de milho, forragem de milho, fibra de milho, feno, palha (*e. g.*, palha de arroz ou palha de trigo), bagaço de cana, polpa de beterraba açucareira, polpa de citrinos e cascas de citrinos,

madeira, desbastes de madeira, cavacos de madeira, pasta de madeira, desperdício de pasta, desperdício de madeira, aparas e serradura de madeira, desperdícios de construção e/ou demolição e restos (*e. g.*, madeira, aparas de madeira e serradura) e, opcionalmente, o desperdício de papel compreende papel de fotocópia rejeitado ou utilizado, papel de impressora de computador, papel de caderno, papel de bloco, papel de máquina de escrever, jornais, revistas, cartão e materiais de empacotamento baseados em papel, e materiais de papel reciclado. Além disso, podem ser utilizados resíduos urbanos, *e. g.*, a fracção de papel de resíduos sólidos municipais, resíduos de madeira municipais e resíduos verdes municipais, em conjunto com outros materiais contendo açúcar, amido e/ou celulose. Em formas de realização alternativas, o processamento do material, *e. g.*, o material de biomassa, produz um bioálcool, *e. g.*, um biodiesel, bioetanol, biometanol, biobutanol ou biopropanol.

Alternativamente, o polipéptido da invenção pode ser expresso no material vegetal de biomassa ou na própria matéria-prima.

Os métodos incluem também pegar no material lenhocelulósico convertido (processado por enzimas da invenção) e torná-lo num combustível (*e. g.*, bioálcool, *e. g.*, um bioetanol, biometanol, biobutanol ou biopropanol, ou biodiesel) por fermentação e/ou por síntese química. Num aspecto, os açúcares produzidos são fermentados e/ou os produtos não fermentáveis são gaseificados.

Os métodos também incluem algas conversoras, óleos vegetais virgens, óleos vegetais residuais, gorduras animais e gorduras (*e. g.*, sebo, banha e gordura amarela) ou esgotos, utilizando enzimas da invenção e transformá-los num combustível (*e. g.*, um bioálcool, *e. g.*, bioetanol, biometanol, biobutanol ou

biopropanol, ou biodiesel) por fermentação e/ou síntese química ou conversão.

As enzimas da invenção (incluindo, por exemplo, organismos, tais como microrganismos, *e. g.*, fungos, leveduras ou bactérias, que fazem e em alguns aspectos segregam enzimas recombinantes da invenção) podem ser utilizadas dentro ou incluídas/integradas em qualquer fase de qualquer processo de conversão de biomassa, *e. g.*, em qualquer passo, diversos passos, ou incluídos em todos os passos, ou em todos os seguintes métodos de processos de conversão de biomassa, ou todas estas alternativas de biocombustível:

- Combustão directa: a queima do material por calor directo e é a tecnologia mais simples de biomassa; pode ser muito económica se existir uma fonte de biomassa próxima.

- Pirólise: é a degradação térmica de biomassa por calor na ausência de oxigénio. Num aspecto, a biomassa é aquecida a uma temperatura entre cerca de 425-750 °C (800 e 1400 graus Fahrenheit), mas nenhum oxigénio é introduzido para suportar a combustão, resultando na criação de gás, óleo combustível e carvão.

- Gaseificação: a biomassa pode ser utilizada para produzir metano através de aquecimento ou digestão anaeróbica. O gás de síntese, uma mistura de monóxido de carbono e hidrogénio, pode ser derivado a partir de biomassa.

- Gás de Aterro: é produzido pela deterioração (digestão anaeróbica) do lixo enterrado em aterros. Quando o resíduo orgânico se decompõe, este produz gás consistindo de

aproximadamente 50% de metano, o principal componente do gás natural.

- Digestão anaeróbica: converte matéria orgânica numa mistura de metano, o principal componente do gás natural, e dióxido de carbono. Num aspecto, biomassa, tais como águas residuais (esgoto), fezes ou resíduos de processamento de alimentos, é misturada com água e alimentada num tanque digestor sem ar.

- Fermentação

  - Fermentação Alcoólica: o álcool combustível é produzido por conversão de massa celulósica e/ou amido em açúcar, fermentar o açúcar em álcool, depois separar a mistura de água e álcool por destilação. Rações, tais como cultivares dedicados (*e. g.*, milho, trigo, cevada, batatas, *switchgrass*, *Miscanthus*, madeira de álamo), resíduos e desperdícios agrícolas (*e. g.*, palha de arroz, forragem de milho, palha de trigo, bagaço de cana, cascas de arroz, fibra de milho, polpa de beterraba açucareira, polpa de citrinos e cascas de citrinos), desperdícios de silvicultura (*e. g.*, desbastes de folhosas e resinosas, resíduos de folhosas e resinosas a partir de operações madeireiras, aparas de madeira e serradura), resíduos urbanos (*e. g.*, fracção de papel de resíduos sólidos municipais, resíduos de madeira municipais, resíduos verdes municipais), resíduos de madeira (*e. g.*, resíduos de serração, resíduos de pasta de papel, resíduos de construção, resíduos de demolição, aparas de madeira e serradura) e papel usado ou outros materiais contendo açúcar, amido e/ou celulose podem ser convertidos em açúcares e depois em álcool por fermentação com levedura.

Alternativamente, os materiais contendo açúcares podem ser convertidos directamente em álcool por fermentação.

- Transesterificação: Uma reacção exemplificativa para converter óleo em biodiesel é denominada transesterificação. O processo d transesterificação reage um álcool (como metanol) com óleos triglicéridos contidos em óleos vegetais, gorduras animais ou gorduras recicladas, formando ésteres alquílicos de ácido gordos (biodiesel) e glicerina. A reacção requer calor e um catalisador de base forte, tal como hidróxido de sódio ou hidróxido de potássio.

- Biodiesel: O biodiesel é uma mistura de ésteres alquílicos de ácido gordos feitos a partir de óleos vegetais, gorduras animais ou gorduras recicladas. O Biodiesel pode ser utilizado como um combustível para veículos na sua forma pura, mas é habitualmente utilizado como um aditivo de petróleo diesel para reduzir níveis de partículas, monóxido de carbono, hidrocarbonetos e tóxicos aéreos partir de veículos a diesel.

- Hidrólise: inclui hidrólise de um composto, *e. g.*, uma biomassa, tal como um material lenhocelulósico, catalisada utilizando uma enzima da presente invenção.

- Cogeração: é a produção simultânea de mais de uma forma de energia utilizando um único combustível e instalação. Num aspecto, a cogeração de biomassa tem um crescimento mais potencial do que a geração de biomassa isoladamente porque a cogeração produz calor e electricidade.

Num aspecto, os polipéptidos da invenção podem ser utilizados em conjunto com outras enzimas, *e. g.*, hidrolases ou enzima que têm actividade celulolíticas, *e. g.*, uma glucanase, endoglucanase, manase e/ou outra enzima, para produzir um combustível, tal como um bioálcool, *e. g.*, um bioetanol, biometanol, biobutanol ou biopropanol, ou biodiesel, a partir de qualquer material orgânico, *e. g.*, uma biomassa, tal como composições derivadas de plantas e animais, incluindo qualquer cultivar agrícola ou outra matéria-prima renovável, um resíduo agrícola ou um resíduo animal, os componentes orgânicos de resíduos municipais e industriais, ou resíduos ou restos de construção ou demolição, ou microrganismos, tais como algas ou levedura.

Num aspecto, os polipéptidos da invenção são utilizados nos processos para converter biomassa lenhocelulósica num combustível (*e. g.*, um bioálcool, *e. g.*, um bioetanol, biometanol, biobutanol ou biopropanol, ou biodiesel), ou, caso contrário, são utilizados em processos para hidrolisar ou digerir biomateriais de modo que possam ser utilizados como um combustível (*e. g.*, um bioálcool, *e. g.*, um bioetanol, biometanol, biobutanol ou biopropanol, ou biodiesel) ou para tornar mais fácil que a biomassa seja processada num combustível.

Num aspecto alternativo, os polipéptidos da invenção, incluindo a mistura de enzimas ou “cocktails” da invenção, são utilizados em processos para um processo de transesterificação que reage um álcool (como etanol, propanol, butanol, propanol, metanol) com um óleo triglicérido contido num óleo vegetal, gordura animal ou gorduras recicladas, formando ésteres alquílicos de ácido gordos (biodiesel) e glicerina. Num aspecto, o biodiesel é feito a partir de óleo de soja ou óleos de cozinha

reciclados. As gorduras animais, outros óleos vegetais e outros óleos reciclados podem também ser utilizados para produzir o biodiesel, dependendo dos seus custos e disponibilidade. Noutro aspecto, misturas de todos os tipos de gorduras e óleos são utilizadas para produzir um combustível de biodiesel.

As enzimas da invenção, incluindo a mistura de enzimas ou "cocktails" da invenção, podem também ser utilizados na refinação da glicerina. O subproduto glicerina contém catalisador e sabões não reagidos que são neutralizados com um ácido. A água e o álcool são removidos para produzir 50% a 80% de glicerina em bruto. Os contaminantes remanescentes incluem gorduras e óleos não reagidos, que podem ser processados utilizando os polipéptidos da invenção. Nas grandes instalações de biodiesel da invenção, a glicerina pode ser ainda purificada, *e. g.*, até pureza de 99% ou mais, para as indústrias farmacêuticas e cosméticas.

Os combustíveis (incluindo bioálcoois, tais como bioetanóis, biometanóis, biobutanóis ou biopropanóis, ou biodiesels) feitos utilizando os polipéptidos da invenção, incluindo a mistura de enzimas ou "cocktails", podem ser utilizados com oxigenados de combustível para melhorar as características de combustão. A adição de oxigénio resulta numa combustão mais completa, a qual reduz as emissões de monóxido de carbono. Isto é outro benefício ambiental da substituição de combustíveis de petróleo com biocombustíveis (*e. g.*, um combustível da invenção). Um biocombustível feito utilizando as composições desta invenção pode ser misturado com gasolina para formar uma mistura E10 (cerca de 5% a 10% de etanol e cerca de 90% a 95% de gasolina), mas pode ser utilizado em concentrações mais elevadas, tal como E85 ou na sua forma pura. Um biocombustível feito utilizando as composições e/ou métodos desta invenção pode ser misturado com

diesel de petróleo para formar uma mistura B20 (20% de biodiesel e 80% de diesel de petróleo), embora outros níveis de mistura possam ser utilizados até B 100 (biodiesel puro).

A divulgação também proporciona processos utilizando enzimas desta invenção para fazer biocombustíveis (incluindo bioálcoois, tais como bioetanóis, biometanóis, biobutanóis ou biopropanóis, ou biodiesels) a partir de composições compreendendo uma biomassa, *e. g.*, uma fonte derivada de plantas, tal como uma biomassa lenhocelulósica. O material de biomassa pode ser obtido a partir de cultivares agrícolas, como um subproduto da produção de alimentos ou rações, ou como produtos residuais, incluindo produtos residuais lenhocelulósicos, tais como resíduos vegetais, papel residual ou resíduos ou restos de construção e/ou demolição. Os exemplos de fontes vegetais ou de resíduos vegetais adequadas para tratamento com polipéptidos da invenção incluem laminária, algas, grãos, sementes, caules, folhas, cascas, vagens, espigas de milho, forragem de milho, palha, cana-de-açúcar, bagaço de cana-de-açúcar, ervas (*e. g.*, erva-indiana, tal como *Sorghastrum mitans*; ou *switchgrass*, *e. g.*, espécies *de Panicum*, tal como *Panicum virgatum*) e semelhantes, assim como madeira, cavacos de madeira, pasta de madeira e serradura. Os exemplos de resíduos de papel adequados para tratamento com polipéptidos da invenção incluem papel de fotocópia rejeitado, papel de impressora de computador, papel de caderno, papel de bloco, papel de máquina de escrever e semelhantes, assim como jornais, revistas, cartão e materiais de empacotamento baseados em papel. Os exemplos de resíduos e restos de construção e demolição incluem madeira, entulho de madeira, aparas de madeira e serradura.

Numa forma de realização, as enzimas, incluindo a mistura de enzimas ou “cocktails” da invenção podem ser utilizadas em

conjunto com meios mais "tradicionais" de produzir etanol, metanol, propanol, butanol, propanol e/ou diesel a partir de biomassa, *e. g.*, como métodos compreendendo hidrolisar materiais lenhocelulósicos ao submeter material lenhocelulósico seco num reactor a um catalisador compreendido de uma solução diluída de um ácido forte e um sal metálico; isto pode diminuir a energia de activação, ou a temperatura, da hidrólise da celulose para obter rendimentos mais elevados de açúcar; ver, *e. g.*, Patentes U.S. Nº 6660506 e 6423145.

Outro método exemplificativo que incorpora a utilização das enzimas da invenção, incluindo a mistura de enzimas ou "cocktails" da invenção, compreende hidrolisar uma biomassa, incluindo qualquer material lenhocelulósico, *e. g.*, contendo hemicelulose, celulose e lenhina, ou qualquer outro polissacárido que possa ser hidrolisado, ao submeter o material a um passo da hidrólise de primeira fase num meio aquoso a uma temperatura e pressão escolhidos para efectuar primeiro a despolimerização da hemicelulose sem grande despolimerização da celulose em glucose. Este passo resulta numa suspensão em que a fase aquosa líquida contém monossacáridos dissolvidos resultantes da despolimerização da hemicelulose e uma fase sólida contendo celulose e lenhina. Um passo de hidrólise de segunda fase pode compreender condições tais que, pelo menos, uma importante porção da celulose é despolimerizada, resultando tal passo numa fase aquosa líquida contendo produtos de despolimerização dissolvidos/solúveis de celulose. Ver, *e. g.*, Patente U.S. Nº 5536325. As enzimas da invenção (incluindo as misturas da invenção ou "cocktails" de enzima) são adicionadas em qualquer fase deste processo exemplificativo.

Outro método exemplificativo que incorpora a utilização de enzimas da invenção, incluindo a mistura de enzimas ou

“cocktails” da invenção, compreende o processamento de um material de biomassa contendo lenhocelulose através de uma ou mais fases de hidrólise ácida diluída com cerca de 0,4% a 2% de ácido forte; e tratamento de um componente lenhocelulósico sólido não reagido do material de biomassa hidrolisado ácido por deslenhificação alcalina para produzir precursores para termoplásticos biodegradáveis e derivados. Ver, *e. g.*, Patente U.S. Nº 6409841. As enzimas da invenção podem ser adicionadas em qualquer fase deste processo exemplificativo.

Um outro método exemplificativo que incorpora a utilização das enzimas da invenção, incluindo a mistura de enzimas ou “cocktails” da invenção, compreende pré-hidrolisar material lenhocelulósico num reactor de pré-hidrólise; adicionar um líquido acídico ao material lenhocelulósico sólido para fazer uma mistura; aquecer a mistura à temperatura da reacção; manter a temperatura da reacção durante tempo suficiente para fraccionar o material lenhocelulósico numa porção solubilizada contendo, pelo menos, cerca de 20% da lenhina do material lenhocelulósico e uma fracção sólida contendo celulose; remover uma porção solubilizada a partir da fracção sólida à temperatura da reacção ou perto desta, em que a celulose na fracção sólida é tornada mais susceptível à digestão enzimática; e recuperar uma porção solubilizada. Ver, *e. g.*, Patente U.S. Nº 5705369. As enzimas da invenção podem ser adicionadas em qualquer fase deste processo exemplificativo.

A divulgação proporciona métodos para fazer composições combustíveis para motor (*e. g.*, para motores de ignição com faísca) com base em hidrocarbonetos líquidos misturados com um álcool de grau de combustível feito utilizando uma enzima da invenção. Num aspecto, os combustíveis feitos por utilização de uma enzima da invenção compreendem, *e. g.*, misturas etanólicas

de gás líquido de carvão ou gás líquido natural. Num aspecto, um co-solvente é 2-metiltetra-hidrofurano (MTHF) derivado de biomassa. Ver, *e. g.*, Patente U.S. Nº 6712866.

Num aspecto, os métodos da divulgação para a degradação enzimática de lenhocelulose, *e. g.*, para a produção de biocombustíveis (incluindo bioálcoois, tais como bioetanóis, biometanóis, biobutanóis ou biopropanóis, ou biodiesels) a partir de material lenhocelulósico, pode também compreender a utilização de tratamento ultrassónico do material de biomassa; ver, *e. g.*, Patente U.S. Nº 6333181.

Noutro aspecto, os métodos da divulgação para produzir biocombustíveis (incluindo bioálcoois, tais como bioetanóis, biometanóis, biobutanóis ou biopropanóis, ou biodiesels) a partir de um substrato celulósico compreende proporcionar uma mistura reaccional na forma de uma suspensão compreendendo substrato celulósico, uma enzima desta invenção e um agente de fermentação (*e. g.*, dentro de uma vaso reaccional, tal como um biorreactor semi-contínuo alimentado a sólidos) e a mistura reaccional é feita reagir sob condições suficientes para iniciar e manter uma reacção de fermentação (como descrito, *e. g.*, no Pedido de Pat. U.S. Nº 20060014260). Num aspecto, os cálculos experimentais ou teóricos podem determinar uma frequência de alimentação óptima. Num aspecto, quantidades adicionais do substrato celulósico e da enzima são proporcionadas no vaso reaccional num intervalo de acordo com a frequência de alimentação optimizada.

Um processo exemplificativo para fazer biocombustíveis (incluindo bioálcoois, tais como bioetanóis, biometanóis, biobutanóis ou biopropanóis, ou biodiesels) é descrito nas Publicações de Pedidos de Pat. U.S. Nº 20050069998; 20020164730;

e num aspecto compreende fases de moagem da biomassa lenhocelulósica (*e. g.*, até um tamanho de 15-30 mm), submeter o produto obtido a pré-tratamento de explosão de vapor (*e. g.*, a uma temperatura de 190-230 °C) durante entre 1 e 10 minutos num reactor; recolher o material pré-tratado num ciclone ou produto de fabrico relacionado; e separar as fracções líquida e sólida por filtração num filtro-prensa, introduzir a fracção sólida num depósito de fermentação e adicionar uma ou mais enzimas da invenção, *e. g.*, uma enzima celulase e/ou beta-glucosidase (*e. g.*, dissolvida em tampão citrato pH 4,8).

Um outro processo exemplificativo para fazer biocombustíveis (incluindo bioálcoois, tais como bioetanóis, biometanóis, biobutanóis ou biopropanóis, ou biodieseiss) compreendendo bioetanóis, biometanóis, biobutanóis ou biopropanóis utilizando enzimas da invenção compreende pré-tratar um material inicial compreendendo uma matéria-prima lenhocelulósica compreendendo, pelo menos, hemicelulose e celulose. Num aspecto, o material inicial compreende batatas, soja (colza), cevada, centeio, milho, aveia, trigo, beterrabas ou cana-de-açúcar ou um componente ou resíduo ou subproduto da produção de alimento ou ração. O material inicial ("matéria-prima") é feito reagir em condições que quebram a estrutura da fibra vegetal para efectuar, pelo menos, uma hidrólise parcial da hemicelulose e celulose. As condições de quebra compreendem, *e. g.*, submeter o material inicial a uma temperatura média de 180 °C a 270 °C, a pH 0,5 a 2,5, durante um período de cerca de 5 segundos a 60 minutos; ou temperatura de 220 °C a 270 °C, a pH 0,5 a 2,5, durante um período de 5 segundos a 120 segundos, ou equivalente. Isto produz um matéria-prima com acessibilidade aumentada para ser digerida por uma enzima, *e. g.*, uma enzima celulase da invenção. Patente U.S. N° 6090595.

As condições exemplificativas para utilizar enzimas da invenção na hidrólise do material lenhocelulósico incluem reacções a temperaturas entre cerca de 30 °C e 48 °C e/ou um pH entre cerca de 4,0 e 6,0. Outras condições exemplificativas incluem uma temperatura entre cerca de 30 °C e 60 °C e um pH entre cerca de 4,0 e 8,0.

Podem ser utilizadas glucanases (ou celulases), mananases, xilanases, amilases, xantanases e/ou glicosidases, *e. g.*, celobio-hidrolases, mananases e/ou beta-glucosidases na conversão de biomassa em combustíveis e na produção de etanol, *e. g.*, como descrito nos Pedidos PCT N° WO0043496 e WO8100857. Podem ser utilizadas glucanases (ou celulases), mananases, xilanases, amilases, xantanases e/ou glicosidases, *e. g.*, celobio-hidrolases, mananases e/ou beta-glucosidases em combinação com a fitase (*e. g.*, enzima da invenção) para produzir açúcares fermentáveis e biomassa contendo glucano que podem ser convertidos em combustível etanol. Podem ser utilizadas amilases, glucoamilases, pulanases, glucoisomerase, alfa-glucosidase e semelhantes, em combinação com a fitase (*e. g.*, enzima da invenção) para converter amido em açúcares ou etanol fermentável. Ver, por favor, Pedido PCT N° WO2005/096804.

Processamento de grão seco de destilação

Noutro aspecto, as enzimas da invenção podem ser utilizadas para tratar/processar "solúveis secos de destilação (DDS)", "grãos secos de destilação (DDS)", "solúveis de destilação condensados(CDS)", "grãos húmidos de destilação (DWG)" e "grãos secos de destilação com solúveis (DDGS)"; os grãos secos de destilação podem ser um subproduto cereal de um processo de destilação e podem incluir solúveis. Estes processos podem

compreender subprodutos vegetais de moagem a seco , *e. g.*, para aplicações de ração, *e. g.*, para aves domésticas, bovinos, suínos e outros animais de criação. Assim, as enzimas da invenção podem ser utilizadas para tratar/processar grãos, *e. g.*, cereais, que são subprodutos de qualquer processo de destilação, incluindo processos que utilizam qualquer fonte de grão, por exemplo, as fontes tradicionais de fabricantes de cerveja ou, alternativamente, de uma instalação produtora de etanol (fábrica, moinho ou semelhantes). As enzimas da invenção podem ser utilizadas para tratar/processar mosto a secar de destilarias; este mosto pode ser utilizado subsequentemente para uma variedade de fins, *e. g.*, como forragem para animais de criação, especialmente ruminantes; assim a divulgação proporciona métodos para processar forragem para animais de criação, tal como ruminantes, e forragem processada por enzima compreendendo fitases desta invenção.

As fitases desta invenção podem ser utilizados isoladamente ou com outras enzimas para processar "solúveis secos de destilação (DDS)", "grãos secos de destilação (DDS)", "solúveis de destilação condensados (CDS)", "grãos húmidos de destilação (DWG)" e "grãos secos de destilação com solúveis (DDGS)". Por exemplo, as fitases desta invenção podem ser utilizadas em qualquer passo de um processo de produto alcoólico como ilustrado na Figura 14. As fitases desta invenção podem ser utilizadas para aumentar a biodisponibilidade do fósforo em qualquer biocombustível, ou potencial biocombustível, incluindo fósforo encontrado em "solúveis secos de destilação (DDS)", "grãos secos de destilação (DDS)", "solúveis de destilação condensados (CDS)", "grãos húmidos de destilação (DWG)" e "grãos secos de destilação com solúveis (DDGS)" (ver, *e. g.*, C. Martinez Amezcua, 2004 Poultry Science 83:971-976).

Bebidas alcoólicas ou produção de álcool bebível

As fitases desta invenção podem também ser utilizadas no processamento de grãos secos de destilação para a produção de álcool - álcool como em "bebidas alcoólicas", *e. g.*, produção de cerveja ou whiskey (para além da utilização no processamento de biomassa para fazer biocombustíveis). As fitases desta invenção podem ser utilizadas em instalações de etanol, *e. g.*, para processar grãos, tal como milho. Os grãos secos de destilação podem ser feitos através primeiro de moagem de um grão (*e. g.*, milho) até uma consistência grosseira e sua adição a água quente. Após arrefecer, a levedura é adicionada e a mistura fermenta durante diversos dias a uma semana. Os sólidos remanescentes após fermentação são as grãos de destilação. As fitases desta invenção podem ser utilizadas em qualquer passo deste processo.

Formulações

A invenção proporciona novas formulações compreendendo fitases, *e. g.*, como aquelas aqui descritas, e formulações para fitases, incluindo formulações que incluem as novas fitases da invenção. As fitases da invenção podem ser utilizadas ou formuladas isoladamente ou como uma mistura de fitases ou fitases e outras enzimas, tais como xilanases, celulases, proteases, lipases, amílases, ou enzimas redox, tais como lacases, peroxidases, catalases, oxidases ou redutase. Podem ser formuladas numa forma sólida, tais como um pó, uma preparação liofilizada, um grânulo, um comprimido, uma barra, um cristal, uma cápsula, um pílula, uma pastilha ou numa forma líquida, tais como uma solução aquosa, um aerossol, um gel, uma pasta, uma

suspensão, uma emulsão aquosa, um creme, uma cápsula ou numa suspensão vesicular ou micelar. As formulações da invenção podem compreender qualquer ou uma combinação dos seguintes ingredientes: polióis, tais como um polietilenoglicol, um polivinilálcool, um glicerol, um açúcar, tais como uma sacarose, um sorbitol, uma trealose, uma glucose, uma frutose, uma maltose, uma manose, um agente de gelificação, tais como uma goma de guar, uma carragenina, um alginato, dextranas, derivados celulósicos, uma pectina, um sal, tais como um cloreto de sódio, um sulfato de sódio, um sulfato de amónio, um cloreto de cálcio, um cloreto de magnésio, um cloreto de zinco, um sulfato de zinco, um sal de um ácido gordo e um derivado de ácido gordo, um quelante de metal, tais como um EDTA, um EGTA, um citrato de sódio, um agente antimicrobiano, tais como um ácido gordo ou um derivado de ácido gordo, um parabeno, um sorbato, um benzoato, um composto modulador adicional para bloquear o impacto de uma enzima, tais como uma protease, uma proteína volumosa, tais como uma BSA, um hidrolisado de trigo, um composto de borato, um aminoácido ou um péptido, um composto modulador apropriado de pH ou temperatura, um emulsionante, tal como um detergente não iónico e/ou iónico, um agente redox, tais como uma cistina/cisteína, um glutationo, um glutationo oxidado, um composto reduzido ou antioxidante, tal como um ácido ascórbico, uma cera ou óleo, ou um dispersante. Reticulação e modificação proteica, tais como pegilação, modificação de ácido gordo, glicosilação podem também ser utilizados para melhorar a estabilidade da enzima.

Medição de Parâmetros Metabólicos

Os métodos da divulgação envolvem evolução de célula inteira ou engenharia de célula inteira, de uma célula para desenvolver

uma nova estirpe celular tendo um novo fenótipo através de modificação da composição genética da célula, em que a composição genética é modificada por adição à célula de um ácido nucleico da invenção. Para detectar o novo fenótipo, pelo menos, um parâmetro metabólico de uma célula modificada é monitorizado na célula numa escala temporal de "tempo real" ou "em linha". Num aspecto, uma pluralidade de célula, tal como uma cultura celular, é monitorizada em "tempo real" ou "em linha". Num aspecto, uma pluralidade de parâmetros metabólicos é monitorizada em "tempo real" ou "em linha".

A análise de fluxo metabólico (MFA) é baseada num contexto bioquímico conhecido. Uma matriz metabólica independentemente linear é construída com base na lei de conservação de massa e na hipótese de pseudo-estado estacionário (PSSH) em metabolitos intracelulares. Na prática dos métodos, são estabelecidas redes metabólicas, incluindo:

- identificar todos as vias de substratos, produtos e metabolitos intermediários
- identificar todas as reacções químicas que interconvertem os metabolitos da via, a estequiometria das reacções da via,
- identificar todas as enzimas que catalisam as reacções, as cinéticas reaccionais das enzimas,
- as interacções reguladoras entre os componentes da via, *e. g.*, interacções alostéricas, interacções enzima-enzima, etc.
- compartimentalização intracelular das enzimas ou qualquer outra organização supramolecular das enzimas e
- a presença de quaisquer gradientes de concentração dos metabolitos, enzimas ou moléculas efectoras ou barreiras de difusão ao seu movimento.

Uma vez construída a rede metabólica para uma dada estirpe, a apresentação matemática por noção de matriz pode ser introduzida para estimar os fluxos metabólicos intracelulares se os dados em linha do metaboloma estiverem disponíveis.

O fenótipo metabólico baseia-se nas alterações de toda a rede metabólica dentro de uma célula. O fenótipo metabólico baseia-se na alteração da utilização da via relativamente a condições ambientais, regulação genética, estado de desenvolvimento e genótipo, etc. Num aspecto, após o cálculo em linha da MFA, o comportamento dinâmico das células, seu fenótipo e outras propriedades são analisadas por investigação da utilização da via. Por exemplo, se a fonte de glucose for aumentada e o oxigénio diminuído durante a fermentação de levedura, a utilização da via respiratório será reduzida e/ou parada, e a utilização das vias fermentativas dominará. O controlo do estado fisiológico de culturas celulares tornar-se-á possível após análise da via. Os métodos podem ajudar a determinar como manipular a fermentação por determinação de como alterar a fonte de substrato, temperatura, utilização de indutores, etc. para controlar o estado fisiológico das células para se moverem num sentido desejável. Na prática dos métodos, os resultados da MFA podem também ser comparados com os dados do transcritoma e proteoma para conceber experiências e protocolos para engenharia metabólica ou arranjo combinatório génico, etc.

Na prática dos métodos, qualquer novo ou modificado fenótipo pode ser conferido e detectado, incluindo características novas ou melhoradas na célula. Qualquer aspecto de metabolismo ou crescimento pode ser monitorizado.

Monitorização da expressão de um transcrito de ARNm

Num aspecto, o fenótipo manipulado compreende aumentar ou diminuir a expressão de um transcrito de ARNm ou gerar novos transcritos numa célula. O transcrito de ARNm, ou mensageiro, pode ser detectado e quantificado através de qualquer método conhecido na técnica, incluindo, *e. g.*, transferência de Northern, reacções quantitativas de amplificação, hibridação em matrizes e semelhantes. As reacções quantitativas de amplificação incluem, *e. g.*, PCR quantitativa, incluindo, *e. g.*, reacção em cadeia da polimerase quantitativa de transcrição reversa ou RT-PCR; RT-PCR quantitativa em tempo real, ou "RT-PCR cinético em tempo real" (ver, *e. g.*, kreuzer (2001) Br. J. Haematol. 114:313-318; Xia (2001) Transplantation 72:907-914).

Num aspecto, o fenótipo manipulado é gerado por inactivação da expressão de um gene homólogo. A sequência codificante do gene ou um ou mais elementos de controlo transcricionais podem ser inactivados, *e. g.*, promotores ou intensificadores. Assim, a expressão de um transcrito pode ser totalmente abolida ou apenas diminuída.

Num aspecto, o fenótipo manipulado compreende o aumento da expressão de um gene homólogo. Este pode ser realizado por inactivação de um elemento de controlo negativo, incluindo um elemento regulador transcricional que actua em *cis*- ou *trans*- ou mutação de um elemento de controlo positivo.

Como discutido abaixo em detalhe, um ou mais, ou todos os transcritos de uma célula podem ser medidos por hibridação de uma amostra compreendendo transcritos de uma célula, ou ácidos nucleicos representativos ou complementares aos transcritos de

uma célula, por hibridação com ácidos nucleicos imobilizados numa matriz.

Monitorização da expressão de polipéptidos, péptidos e aminoácidos

Num aspecto, o fenótipo manipulado compreende o aumento ou diminuição da expressão de um polipéptido ou gerar novos polipéptidos numa célula. Os polipéptidos, péptidos e aminoácidos podem ser detectados e quantificados através de qualquer método conhecido na técnica, incluindo, *e. g.*, ressonância magnética nuclear (RMN), espectrofotometria, radiografia (radiomarcação de proteína), electroforese, electroforese capilar, cromatografia líquida de elevado desempenho (HPLC), cromatografia de camada fina (TLC), cromatografia de hiperdifusão, vários métodos imunológicos, *e. g.*, imunoprecipitação, imunodifusão, imuno-electroforese, radioimunoensaios (RIA), imunoensaio de adsorção ligado à enzima (ELISA), ensaios imuno-fluorescentes, electroforese em gel (*e. g.*, SDS-PAGE), coloração com anticorpos, separação de células activadas por fluorescência (FACS), espectrometria de massa de pirólise, espectrometria de infravermelho com transformadas de Fourier, espectrometria de Raman, GC-MS e espectrometrias de massa de electrosspray LC e de electrosspray tandem cap-LC e semelhantes. As novas bioactividades também podem ser rastreadas utilizando métodos ou as suas variações, descritos na Patente U.S. Nº 6057103. Além disso, como discutido abaixo em detalhe, um ou mais, ou todos os polipéptidos de uma célula podem ser medidos utilizando uma matriz proteica.

A marcação fraccionária de ^{13}C dirigida biossinteticamente de aminoácidos proteicos pode ser monitorizada por alimentação

de uma mistura de compostos de fonte de carbono não marcados e marcados uniformemente com C^{13} numa rede de biorreacção. A análise do perfil de marcação resultante permite uma caracterização detalhada da topologia da rede e a determinação das razões do fluxo metabólico dos aminoácidos; ver, *e. g.*, Szyperski (1999) Metab. Eng. 1:189-197.

Os seguintes exemplos são pretendidos para ilustrar, mas não limitar, a invenção. Embora os processos descritos nos exemplos sejam típicos daqueles que podem ser utilizados para realizar determinados aspectos da invenção, outros processos conhecidos dos especialistas na técnica podem também ser utilizados.

EXEMPLOS

Exemplo 1: Caracterização da actividade de fitases exemplificativas da invenção

Este Exemplo descreve a caracterização da actividade de fitase de polipéptidos da invenção, que são modificações de sequência (as denominadas fitases "evoluídas") de fitase parental da SEQ ID N°:2 e um ensaio exemplificativo da actividade de fitase. Este ensaio de actividade de fitase pode ser utilizado para determinar se um polipéptido tem actividade suficiente para estar dentro do âmbito da invenção reivindicada.

Após geração dos polipéptidos da invenção por expressão das sequências de ácidos nucleicos modificadas por GSSM da invenção, os polipéptidos de fitase "evoluídos" - apenas espécies exemplificativas de mutação simples de resíduo neste estudo - foram purificados e depois tratados termicamente (pH 7,0, 0,01% de Tween) a várias temperaturas durante 30 minutos. Após o passo

de tratamento térmico, as amostras (20 uL) foram ensaiadas com o substrato de fluorescência (180 uL) 4 mM de DiFMUP a pH 5,5. As velocidades foram comparadas às velocidades de cada amostra não tratada correspondente. Os resultados são ilustrados na Figura 5.

Noutro estudo, as fitases "evoluídas" da invenção compreendendo mutações simples "misturadas", *i. e.*, fitases contendo múltiplas mutações, foram cultivadas de um dia para o outro em LBCARB100™ (LBcarb100) a 30 °C. 100 uL de cada cultura (mutante misturado) foram tratados termicamente num termociclador de 72 °C a 100 °C. 20 uL da cultura tratada termicamente foram misturados com 180 uL de 4 mM de DiFMUP a pH 5,5. As velocidades foram comparadas às velocidades de cada amostra não tratada correspondente; como resumido na Figura 6. A Tabela ilustrada na Figura 7 resume graficamente os dados (arredondados à décima mais próximo) utilizados para gerar o gráfico da Figura 6.

A amostra 1 a 21 corresponde às mutações simples "misturadas" da SEQ ID N°:2 parental, como ilustrado na tabela da Figura 9; assinale-se que a fitase "evoluída" número 10 tem uma modificação de resíduo de sequência (da SEQ ID N°:2) que não foi introduzida por GSSM; esta mutação foi introduzida por acaso aleatório e pode não ter qualquer relevância para a estabilidade térmica desta fitase exemplificativa da invenção. Também, assinale-se que as fitases exemplificativas 19, 20 e 21 marcadas com ** têm marcadores (6xHis) de histidina no C-terminal (-RSHHHHHH). A Figura 9 ilustra as fitases exemplificativas tendo modificações múltiplas de resíduo relativamente à SEQ ID N°:2 parental; como aqui descrito em detalhe. A Figura 10 ilustra as fitases exemplificativas tendo modificações únicas de

resíduo relativamente à SEQ ID N°:2 parental; como aqui descrito em detalhe.

A Figura 11 ilustra esquematicamente um ensaio exemplificativo de fitase da invenção utilizando o substrato de fluorescência 4-metilumbeliferil-fosfato (MeUMB-fosfato, cuja estrutura é também ilustrada): (i) a fitase é desafiada pelo calor durante 20 minutos, 72 °C, pH 4,5; e a 80 °C a pH fisiológico (pH 7,4); e (ii) a actividade residual é testada a 37 °C, pH 4,5:

- Medir a actividade residual a pH elevado e baixo,
- Calcular a actividade residual relativamente ao tipo selvagem após tratamento térmico.

A Figura 12 ilustra esquematicamente um outro ensaio exemplificativo de fitase da invenção utilizando também o substrato de fluorescência MeUMB-fosfato: (i) a fitase é desafiada pelo calor durante 30 minutos, 86 °C, pH 5,5; e (ii) a actividade residual é testada a 37 °C, pH 4,5:

- Medir a actividade residual relativamente à variante 6X de controlo,
- Seleccionar acertos e re-ensaiar a restringência mais elevada para seleccionar variantes superiores.

Este ensaio foi utilizado para rastrear bibliotecas de variantes de GSSM (da SEQ ID N°:1), por ensaio da actividade de fitase dos polipéptidos que codificam. A Figura 13 ilustra esquematicamente o protocolo para este rastreio de biblioteca (como descrito na Figura 12), em que o tamanho de biblioteca rastreado é 24576 variantes.

Lisboa, 8 de Janeiro de 2014

REIVINDICAÇÕES

1. Ácido nucleico isolado, sintético ou recombinante, compreendendo

(a) uma sequência de ácidos nucleicos codificando um polipéptido tendo uma actividade de fitase e tendo, pelo menos, 95%, 96%, 97%, 98% ou 99% ou mais de identidade de sequência para a SEQ ID N°:1, e tem uma modificação nucleotídica em que os nucleótidos nas posições 139 a 141 da SEQ ID N°: 1 estão alterados para TTT ou TTC, em que o polipéptido tem termostabilidade melhorada relativamente à fitase da SEQ ID N°:2;

(b) um ácido nucleico codificando um polipéptido tendo uma actividade de fitase e tendo, pelo menos, 95%, 96%, 97%, 98% ou 99% ou mais de identidade de sequência para a SEQ ID N°:2, e tem uma modificação de aminoácido em que o aminoácido alanina na posição de aminoácido 47 da SEQ ID N°:2 é alterado para uma fenilalanina, em que o polipéptido tem termostabilidade melhorada relativamente à fitase da SEQ ID N°:2;

(c) a sequência de ácidos nucleicos de (a) ou (b) codificando um polipéptido tendo uma actividade de fitase mas que carece de: uma sequência de sinal ou uma sequência de proproteína, ou uma sequência promotora homóloga;

(d) o ácido nucleico de qualquer de (a) a (c) codificando um polipéptido tendo uma actividade de fitase e compreendendo ainda uma sequência de

aminoácidos heteróloga, ou o ácido nucleico de qualquer de (a) a (c) compreendendo ainda uma sequência de nucleótidos heteróloga;

(e) o ácido nucleico de (d), em que a sequência de aminoácidos heteróloga compreende ou consiste numa sequência codificando uma sequência de sinal heteróloga (líder), ou um marcador ou um epitopo, ou a sequência de nucleótidos heteróloga compreende um promotor heterólogo ou codificando uma sequência de direccionamento; ou

(f) uma sequência de ácidos nucleicos totalmente complementar à sequência de ácidos nucleicos de qualquer de (a) a (e).

2. Ácido nucleico isolado, sintético ou recombinante, da reivindicação 1, em que o ácido nucleico compreende ainda, pelo menos, uma, duas, três, quatro, cinco, seis, sete, oito, nove, dez, onze, doze ou todas as treze das modificações de sequência de pares de bases nucleotídicos para a SEQ ID Nº:1, seleccionado do grupo consistindo de:

os nucleótidos nas posições 289 a 291 são GTT, GTC, GTA ou GTG;
os nucleótidos nas posições 289 a 291 são GAA ou GAG;
os nucleótidos nas posições 406 a 408 são CAT ou CAC;
os nucleótidos nas posições 475 a 477 são GTT, GTC, GTA ou GTG;
os nucleótidos nas posições 475 a 477 são GAA ou GAG;
os nucleótidos nas posições 487 a 489 são CGT, CGC, CGA, CGG, AGA ou AGG;

os nucleótidos nas posições 490 a 492 são CGT, CGC, CGA, CGG, AGA ou AGG;
os nucleótidos nas posições 502 a 504 são CGT, CGC, CGA, CGG, AGA ou AGG;
os nucleótidos nas posições 535 a 537 são CGT, CGC, CGA, CGG, AGA ou AGG;
os nucleótidos nas posições 676 a 678 são GAT ou GAC;
os nucleótidos nas posições 697 a 699 são TGG;
os nucleótidos nas posições 823 a 825 são GTT, GTC, GTA ou GTG;
os nucleótidos nas posições 865 a 867 são GCT, GCC, GCA ou GCG;
os nucleótidos nas posições 1045 a 1047 são TAT ou TAC; e
os nucleótidos nas posições 1087 a 1089 são CCA, CCC, CCG ou CCT,

em que o polipéptido tem termostabilidade melhorada relativamente à fitase da SEQ ID N°:2.

3. Cassete de expressão, vector, veículo de clonagem, vector de expressão, vector de clonagem (i) compreendendo o ácido nucleico da reivindicação 1 ou 2, ou contendo aí um ácido nucleico da reivindicação 1 ou 2; ou (ii) a cassete de expressão, vector, veículo de clonagem, vector de expressão, vector de clonagem de (i), em que o veículo de clonagem compreende ou é um vector viral, um plasmídeo, um fago, um fagemídeo, um cosmídeo, um fosmídeo, um bacteriófago ou um cromossoma sintético e, opcionalmente, o vector viral compreende um vector de adenovírus, vectores retrovirais ou um vector viral adeno-associado e, opcionalmente, a cassete de expressão, vector, veículo de clonagem, vector de expressão, vector de clonagem compreende ou está contido num

cromossoma sintético bacteriano (BAC), um plasmídeo, um vector de bacteriófago derivado de P1 (PAC), um cromossoma sintético de levedura (YAC), um cromossoma sintético de mamífero (MAC).

4. Célula transformada ou célula hospedeira (i) compreendendo: o ácido nucleico da reivindicação 1 ou 2, ou a cassete de expressão, vector, veículo de clonagem, vector de expressão, vector de clonagem da reivindicação 3; ou (ii) a célula transformada ou célula hospedeira de (i), em que a célula é uma célula bacteriana, uma célula de mamífero, uma célula fúngica, uma célula de levedura, uma célula de insecto ou uma célula vegetal; (iii) a célula transformada ou célula hospedeira de (ii), em que a célula bacteriana é qualquer espécie dentro dos géneros *Escherichia, Bacillus, Streptomyces, Salmonella, Pseudomonas* ou *Staphylococcus*, ou *Escherichia coli, Lactococcus lactis, Bacillus subtilis, Bacillus cereus, Salmonella typhimurium* ou *Pseudomonas fluorescens.*

5. Planta ou semente transgénica (i) compreendendo: o ácido nucleico da reivindicação 1 ou 2, ou a cassete de expressão, vector, veículo de clonagem, vector de expressão, vector de clonagem da reivindicação 3, ou a célula transformada ou célula hospedeira da reivindicação 4; ou (ii) planta ou semente transgénica de (i),
e, opcionalmente, em que a planta é uma planta de milho, uma planta de batata, uma planta de tomate, uma planta de trigo, uma planta oleaginosa, uma planta de colza, uma planta de soja ou uma planta de tabaco, ou uma planta forrageira e/ou de ração para um animal, ou para um animal ruminante,
e, opcionalmente, em que a planta forrageira ou de ração é ou compreende feno, milho, milho painço, soja, trigo, trigo

mourisco, cevada, alfalfa, centeio, uma erva anual, sorgo, erva-do-sudão, erva de veldt ou erva de buffel; ou a semente é uma semente de milho, uma semente de trigo, uma oleaginosa, uma semente de colza, uma semente de soja, uma semente de palma, uma semente de girassol, uma semente de sésamo, um amendoim ou uma semente de amendoim, uma semente de alfalfa, uma semente de algodão, uma semente de cártamo, uma semente de sorgo, uma semente de aveia, uma semente de centeio, uma semente de milho painço, uma semente de cevada, uma semente de arroz, uma semente de ervilha, ou uma semente de planta de tabaco,

e, opcionalmente, a planta é milho, alfalfa, girassol, Brassica, soja, cana-de-açúcar, algodão, cártamo, amendoim, sorgo, trigo, aveia, centeio, milho painço, cevada, arroz, coníferas, ervilha, feijão, soja, batata, batata-doce, mandioca, inhame, balizeiro ou beterraba açucareira.

6. Polipéptido isolado, sintético ou recombinante, de fitase compreendendo

(a) uma sequência de aminoácidos tendo, pelo menos, 95%, 96%, 97%, 98% ou 99% de identidade de sequência para a SEQ ID Nº:2 e codificada por um ácido nucleico compreendendo uma modificação nucleotídica, em que os nucleótidos nas posições 139 a 141 da SEQ ID Nº: 1 são TTT ou TTC, em que o polipéptido tem termostabilidade melhorada relativamente à fitase da SEQ ID Nº:2;

(b) uma sequência de aminoácidos tendo, pelo menos, 95%, 96%, 97%, 98% ou 99% de identidade de sequência para a SEQ ID Nº:2 e tem uma modificação de aminoácido em que o aminoácido alanina na posição de aminoácido 47 da SEQ ID Nº:2 é alterado para uma fenilalanina, em que

o polipéptido tem termostabilidade melhorada relativamente à fitase da SEQ ID N°:2;

(c) o polipéptido de (a) ou (b) tendo uma actividade de fitase mas carece de: uma sequência de sinal ou uma sequência de proproteína;

(d) o polipéptido de qualquer de (a) a (c) tendo uma actividade de fitase e compreendendo ainda uma sequência heteróloga;

(e) o polipéptido de (d), em que a sequência de aminoácidos heteróloga compreende ou consiste numa sequência de sinal heteróloga (líder), ou um marcador ou um epitopo, ou a sequência heteróloga compreende um péptido de identificação ou uma sequência de direccionamento; ou

(f) o polipéptido de qualquer de (a) a (e), em que (i) o polipéptido está glicosilado, ou o polipéptido compreende, pelo menos, um local de glicosilação, (ii) o polipéptido (i) em que a glicosilação é uma glicosilação ligada por N ou uma glicosilação ligada por O; (iii) o polipéptido de (i) ou (ii) em que o polipéptido está glicosilado após ter sido expresso numa célula de levedura; ou (iv) o polipéptido (iii) em que a célula de levedura é uma de *P. pastoris* ou uma de *S. pombe*.

7. Polipéptido isolado, sintético ou recombinante, da reivindicação 6, em que
a sequência de aminoácidos compreende ainda, pelo menos, uma, duas, três, quatro, cinco, seis, sete, oito, nove, dez,

onze, doze, ou todas as treze das modificações de resíduo de aminoácidos para a SEQ ID N°:2, seleccionado do grupo consistindo emde

o aminoácido cisteína na posição de aminoácido 97 da SEQ ID N°:2 é alterado para uma valina;
o aminoácido cisteína na posição de aminoácido 97 da SEQ ID N°:2 é alterado para um ácido glutâmico;
o aminoácido treonina equivalente na posição de aminoácido 136 da SEQ ID N°:2 é alterado para uma histidina;
o aminoácido asparagina na posição de aminoácido 159 da SEQ ID N°:2 é alterado para uma valina;
o aminoácido asparagina na posição de aminoácido 159 da SEQ ID N°:2 é alterado para um ácido glutâmico;
o aminoácido treonina equivalente na posição de aminoácido 163 da SEQ ID N°:2 é alterado para uma arginina;
o aminoácido ácido aspártico na posição de aminoácido 164 da SEQ ID N°:2 é alterado para uma arginina;
o aminoácido ácido glutâmico na posição de aminoácido 168 da SEQ ID N°:2 é alterado para uma arginina;
o aminoácido glicina na posição de aminoácido 179 da SEQ ID N°:2 é alterado para uma arginina;
o aminoácido cisteína na posição de aminoácido 226 da SEQ ID N°:2 é alterado para um ácido aspártico;
o aminoácido valina na posição de aminoácido 233 da SEQ ID N°:2 é alterado para um triptofano;
o aminoácido glutamina na posição de aminoácido 275 da SEQ ID N°:2 é alterado para uma valina;
o aminoácido arginina na posição de aminoácido 289 da SEQ ID N°:2 é alterado para uma alanina;

o aminoácido treonina na posição de aminoácido 349 da SEQ ID N°:2 é alterado para uma tirosina; e
o aminoácido leucina na posição de aminoácido 363 da SEQ ID N°:2 é alterado para uma prolina,

em que o polipéptido tem termostabilidade melhorada relativamente à fitase da SEQ ID N°:2.

8. Método para produzir um polipéptido recombinante da reivindicação 6 ou reivindicação 7 compreendendo os passos de: (a) proporcionar um ácido nucleico da reivindicação 1 ou 2; e (b) expressar o ácido nucleico do passo (a) sob condições que permitem a expressão do polipéptido, produzindo assim um polipéptido recombinante,
e, opcionalmente, o método compreende ainda a transformação de uma célula hospedeira com o ácido nucleico do passo (a) seguido pela expressão do ácido nucleico do passo (a), produzindo assim o polipéptido recombinante numa célula hospedeira transformada.

9. Método para hidrolisar um inositol-hexafosfato em inositol e fosfato inorgânico compreendendo:

(i) (a) proporcionar um polipéptido tendo uma actividade de fitase, em que o polipéptido compreende a sequência de aminoácidos da reivindicação 6 ou 7, ou um polipéptido codificado pelo ácido nucleico da reivindicação 1 ou 2; (b) proporcionar uma composição compreendendo um inositol-hexafosfato; e (c) colocar em contacto o polipéptido do passo (a) com a composição do passo (b) sob condições em que o polipéptido hidrolisa o inositol-hexafosfato para produzir inositol e fosfato inorgânico;

(ii) o método de (i), em que as condições compreendem uma temperatura entre cerca de 37 °C e cerca de 70 °C, entre cerca de 50 °C e cerca de 80 °C, ou entre cerca de 60 °C e cerca de 90 °C; ou
(iii) o método de (i) ou (ii), em que a composição compreende um ácido fítico.

10. Método para a decruagem de óleo compreendendo: (a) proporcionar um polipéptido tendo uma actividade de fitase, em que o polipéptido compreende a sequência de aminoácidos da reivindicação 6 ou 7, ou um polipéptido codificado pelo ácido nucleico da reivindicação 1 ou 2; (b) proporcionar uma composição compreendendo um óleo vegetal; e (c) colocar em contacto o polipéptido do passo (a) e o óleo vegetal do passo (b) sob condições em que o polipéptido pode clivar uma ligação inositol-fosfato inorgânico, conduzindo assim à decruagem do óleo vegetal.

11. Método para produzir uma matriz de distribuição enzimática em pastilhas comestíveis, compreendendo:

(a) proporcionar o polipéptido da reivindicação 6 ou 7, ou um polipéptido codificado pelo ácido nucleico da reivindicação 1 ou 2; (b) proporcionar um grão, um germe de grão ou um substrato à base de grão; e (c) combinar o polipéptido de (a) e o grão, germe de grão ou substrato à base de grão de (b) para formar uma matriz de distribuição enzimática em pastilhas comestíveis, em que a matriz de distribuição comestível é formada por formação de pastilhas para produzir uma partícula e, opcionalmente, as pastilhas são feitas sob condições compreendendo a aplicação de vapor, opcionalmente as pastilhas são feitas sob condições

compreendendo a aplicação de uma temperatura superior a 80 °C durante cerca de 5 minutos e, opcionalmente, a pastilha compreende uma enzima fitase que compreende uma actividade específica de, pelo menos, 350 a cerca de 900 unidades por miligrama de enzima.

12. Método não terapêutico para distribuir um suplemento de enzima fitase a um animal ou humano, o referido método compreendendo:

(a) preparar uma matriz de distribuição em pastilhas comestíveis num método de acordo com a reivindicação 11 e
(b) administrar a matriz de distribuição enzimática comestível ao animal ou humano.

13. Método para processar sementes de milho e sorgo compreendendo (a) proporcionar um polipéptido tendo uma actividade de fitase, em que o polipéptido compreende o polipéptido da reivindicação 6 ou 7; (b) proporcionar uma composição compreendendo uma água de maceração de milho ou água de maceração de sorgo, e (c) colocar em contacto o polipéptido do passo (a) e a composição do passo (b) sob condições em que o polipéptido pode clivar uma ligação inositol-fosfato inorgânico.

14. Composição para melhorar (retardar o progresso, tratar ou prevenir) a osteoporose compreendendo o polipéptido da reivindicação 6 ou 7; em que, opcionalmente, o polipéptido está glicosilado e em que a actividade de fitase é termotolerante ou termoestável.

15. Preparação de proteína compreendendo o polipéptido da reivindicação 6 ou 7, em que a preparação de proteína compreende um líquido, uma pasta, um pó, um spray, uma suspensão, uma composição/formulação liofilizada, um sólido, um *geltab*, uma pílula, um implante, um gel, ou uma formulação farmacêutica, um alimento ou uma ração ou um seu suplemento.

16. Composição compreendendo o polipéptido da reivindicação 6 ou 7, em que a composição é seleccionada de uma unidade Pronto-a-Comer (MRE) de refeição autónoma, uma bebida, um agente hidratante, um comprimido mastigável, uma barra nutritiva, um produto cozido, um cereal de pequeno almoço, um produto farmacêutico, uma formulação dietética, uma refeição de soja, um alimento, uma ração e uma matriz de distribuição enzimática em pastilhas comestíveis.

Lisboa, 8 de Janeiro de 2014

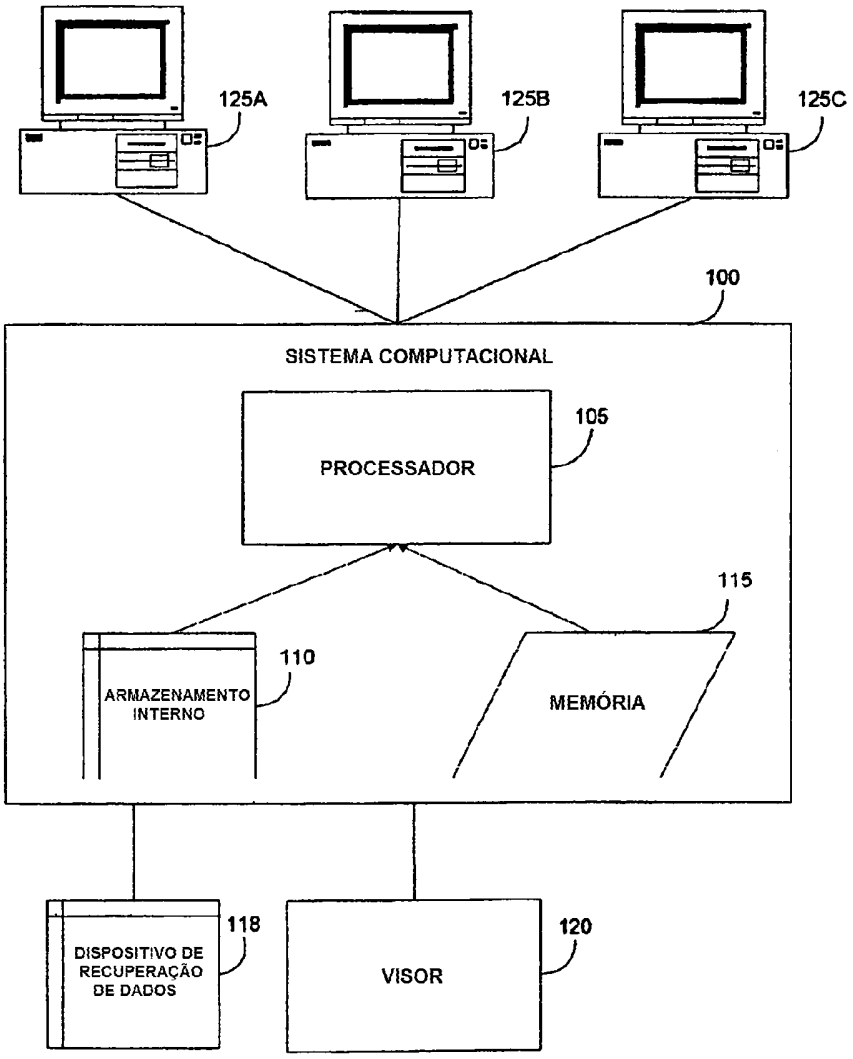

125A
125B
125C
100
SISTEMA COMPUTACIONAL
PROCESSADOR
105
115
ARMAZENAMENTO INTERNO
110
MEMÓRIA
DISPOSITIVO DE RECUPERAÇÃO DE DADOS
118
VISOR
120

FIGURA 1

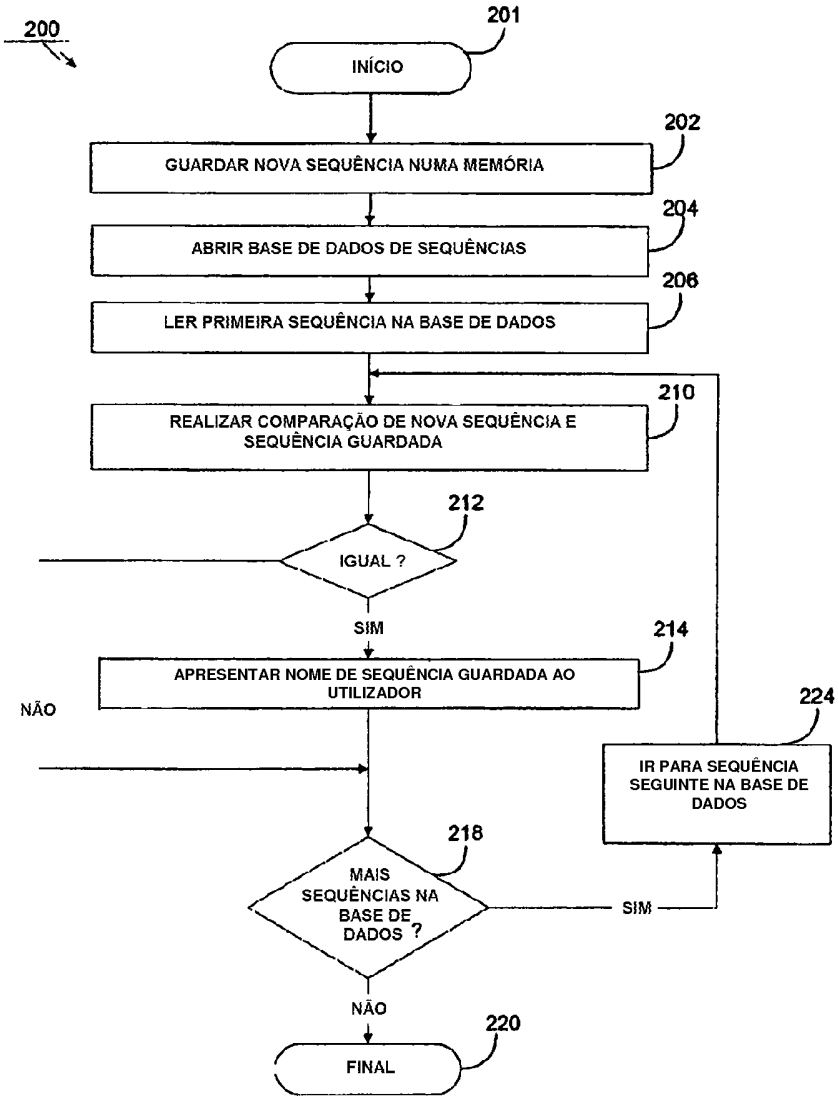

200
201
INÍCIO
202
GUARDAR NOVA SEQUÊNCIA NUMA MEMÓRIA
204
ABRIR BASE DE DADOS DE SEQUÊNCIAS
206
LER PRIMEIRA SEQUÊNCIA NA BASE DE DADOS
210
REALIZAR COMPARAÇÃO DE NOVA SEQUÊNCIA E SEQUÊNCIA GUARDADA
212
IGUAL ?
SIM
214
APRESENTAR NOME DE SEQUÊNCIA GUARDADA AO UTILIZADOR
NÃO
224
IR PARA SEQUÊNCIA SEGUINTE NA BASE DE DADOS
218
MAIS SEQUÊNCIAS NA BASE DE DADOS ?
SIM
NÃO
220
FINAL

FIGURA 2

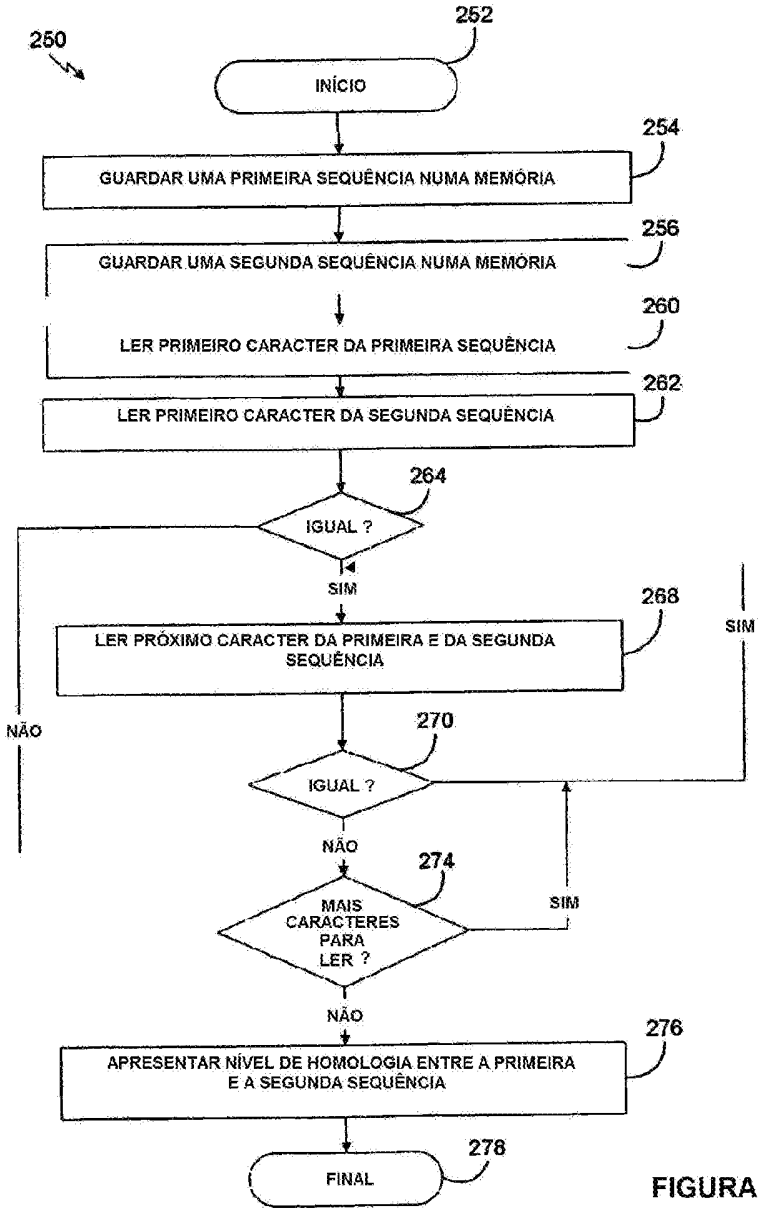

250
252
INÍCIO
254
GUARDAR UMA PRIMEIRA SEQUÊNCIA NUMA MEMÓRIA
256
GUARDAR UMA SEGUNDA SEQUÊNCIA NUMA MEMÓRIA
260
LER PRIMEIRO CARACTER DA PRIMEIRA SEQUÊNCIA
262
LER PRIMEIRO CARACTER DA SEGUNDA SEQUÊNCIA
264
IGUAL ?
NÃO
SIM
268
LER PRÓXIMO CARACTER DA PRIMEIRA E DA SEGUNDA SEQUÊNCIA
SIM
270
IGUAL ?
NÃO
274
MAIS CARACTERES PARA LER ?
SIM
NÃO
276
APRESENTAR NÍVEL DE HOMOLOGIA ENTRE A PRIMEIRA E A SEGUNDA SEQUÊNCIA
278
FINAL

FIGURA 3

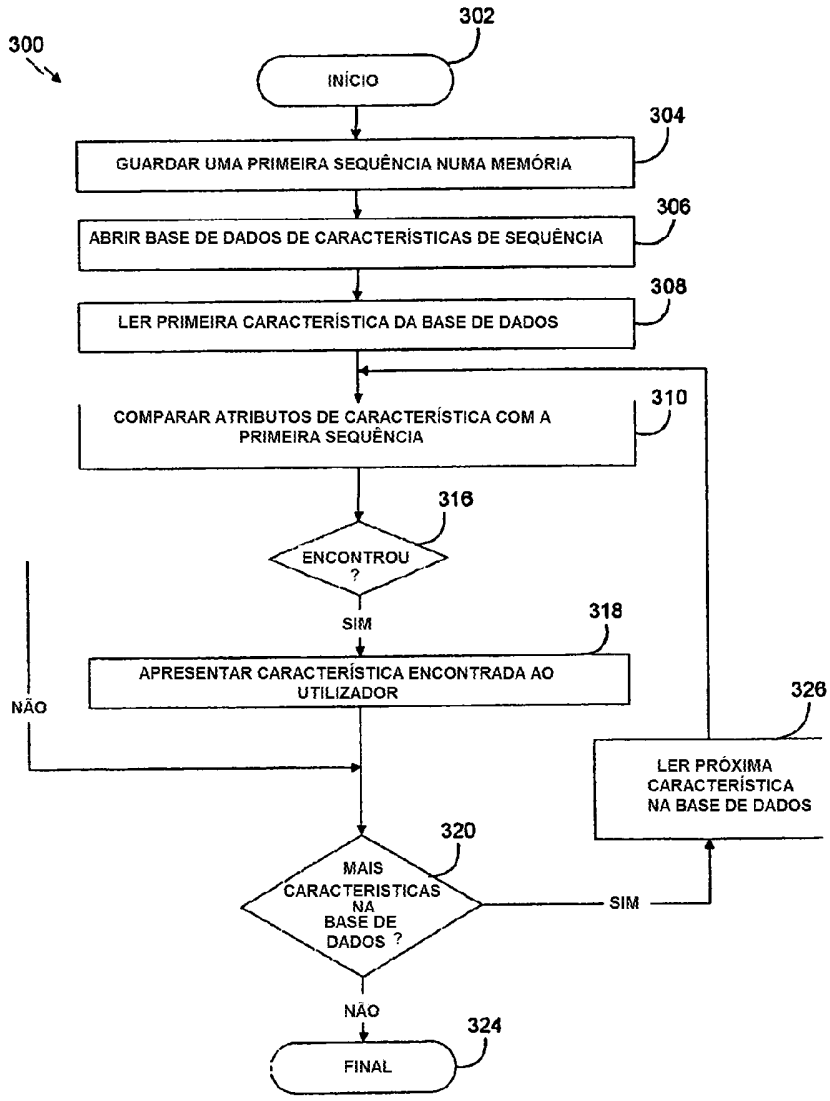

300
302
INÍCIO
304
GUARDAR UMA PRIMEIRA SEQUÊNCIA NUMA MEMÓRIA
306
ABRIR BASE DE DADOS DE CARACTERÍSTICAS DE SEQUÊNCIA
308
LER PRIMEIRA CARACTERÍSTICA DA BASE DE DADOS
310
COMPARAR ATRIBUTOS DE CARACTERÍSTICA COM A PRIMEIRA SEQUÊNCIA
316
ENCONTROU ?
SIM
318
APRESENTAR CARACTERÍSTICA ENCONTRADA AO UTILIZADOR
NÃO
326
LER PRÓXIMA CARACTERÍSTICA NA BASE DE DADOS
320
MAIS CARACTERISTICAS NA BASE DE DADOS ?
SIM
NÃO
324
FINAL

FIGURA 4

Figura 5

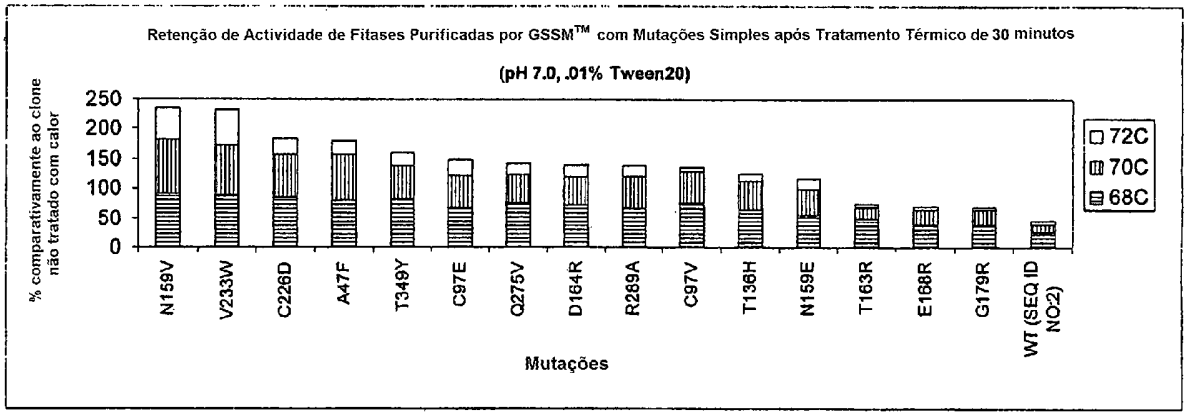
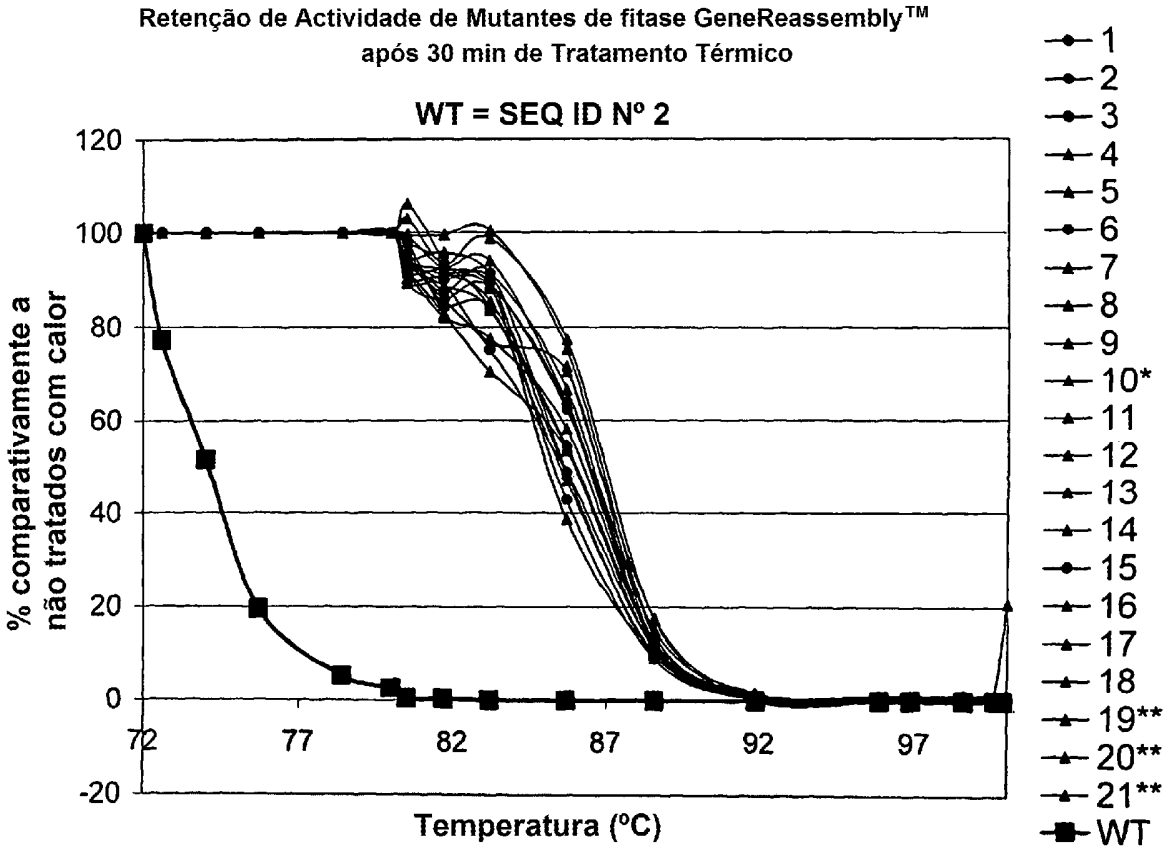

Retenção de Actividade de Fitases Purificadas por GSSM™ com Mutações Simples após Tratamento Térmico de 30 minutos
(pH 7.0, .01% Tween20)
% comparativamente ao clone não tratado com calor
250
200
150
100
50
0
72C
70C
68C
N159V
V233W
C226D
A47F
T348Y
C97E
Q275V
D164R
R269A
C97V
T136H
N159E
T163R
E168R
G179R
WT (SEQ ID NO:2)
Mutações

Figura 6


Retenção de Actividade de Mutantes de fitase GeneReassembly™
após 30 min de Tratamento Térmico
WT = SEQ ID Nº 2
% comparativamente a não tratados com calor
120
100
80
60
40
20
0
-20
72
77
82
87
92
97
Temperatura (ºC)
1
2
3
4
5
6
7
8
9
10*
11
12
13
14
15
16
17
18
19**
20**
21**
WT

	Temperatura (°C)																
Fitase Nº	72	72.6	74	75.7	78.4	80	80.5	81.7	83.2	85.7	88.6	91.9	95.9	96.9	98.6	99.6	100
1	100	100	100	100	100	100	94.7	90.0	90.2	42.9	9.1	0.5	0.7	0.5	0.7	0.5	0.4
2	100	100	100	100	100	100	94.2	85.7	75.3	48.9	11.8	0.9	0.6	0.5	0.8		0.6
3	100	100	100	100	100	100	94.5	87.3	91.5	48.8				0.3	0.3		
4	100	100	100	100	100	100	94.3	92.5	94.2	63.7	11.7	1.2	0.8	0.8	0.8	0.8	0.9
5	100	100	100	100	100	100	90.3	90.8	88.9	63.3	12.1	1.1	1.0	0.9	0.9	0.8	1.0
6	100	100	100	100	100	100	98.0	93.5	84.2	54.5	12.8	1.3	0.0	0.8	0.6	0.6	1.0
7	100	100	100	100	100	100	94.6	87.3		71.7	17.6	0.7	0.2	0.2	0.1	0.3	0.2
8	100	100	100	100	100	100	99.5	99.5	100.5	75.3	14.4	0.5	0.1	0.1	0.1	0.1	0.5
9	100	100	100	100	100	100	91.8	82.2	77.5	58.2	12.2	1.4	1.1	1.0			1.2
10*	100	100	100	100	100	100	94.6	95.8	91.7	66.6	11.7	0.5	0.2	0.3	0.5	0.4	0.4
11	100	100	100	100	100	100	96.6	82.8		64.5	15.2	1.6	1.1	1.1	1.2	1.2	1.2
12	100	100	100	100	100	100	93.2	88.8	83.6	53.8	11.5	0.3	0.3	0.3	0.3	0.3	0.4
13	100	100	100	100	100	100	92.4	82.7	70.6	53.6	11.0	0.8	0.4	0.4		0.5	20.8
14	100	100	100	100	100	100	103.0	93.4	98.8	77.4	17.3	0.2	0.2	0.1	0.1	0.2	0.1
15	100	100	100	100	100	100	91.7	92.2	89.4	62.2	12.9		0.2	0.3	-0.2	0.2	0.2
16	100	100	100	100	100	100	93.2	91.4	89.2	38.7	10.6	0.3	0.2	0.2	0.2	0.1	0.2
17	100	100	100	100	100	100	98.0	84.7	84.3	47.0	10.1	0.2	0.2	0.2		0.2	0.2
18	100	100	100	100	100	100	89.2	85.8	88.2	48.5	9.4	0.7	0.8	0.4	0.5	0.4	0.2
19**	100	100	100	100	100	100	99.2	92.1	85.5	54.4	11.0	0.4	0.2	0.2	0.2	0.3	0.2
20**	100	100	100	100	100	100	89.4	88.8	76.6	70.6	12.2	0.3	0.2		0.2	0.2	0.2
21**	100	100	100	100	100	100	106.1	93.9		63.8	12.7	1.0	0.6	0.6	0.6	0.5	0.6
WT (SEQ ID Nº 2)	100	77.4	51.7	19.5	5.29	2.6	0.4	0.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Figura 7

Figura 8

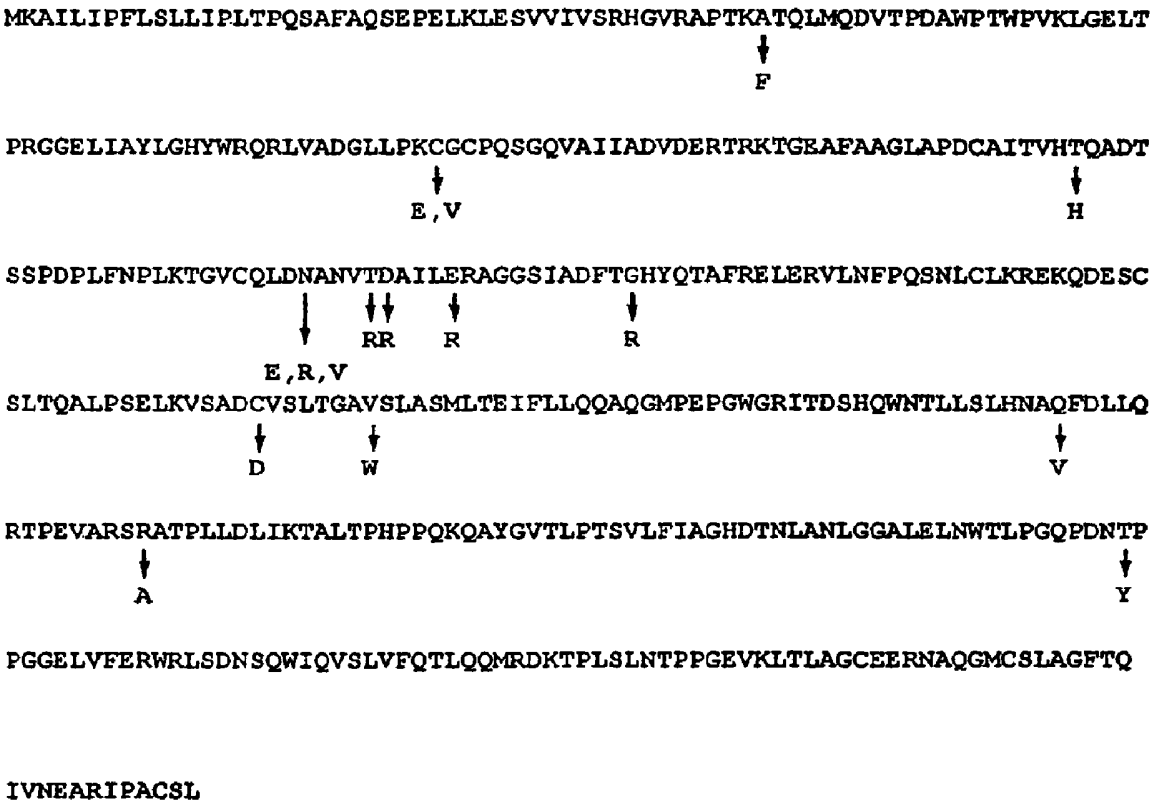

MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAWPTWPVKLGELT
F
PRGGELIAYLGHYWRQRLVADGLLPKCGCPQSGQVAIIADVDERTRKTGEAFAAGLAPDCAITVHTQADT
E,V
H
SSPDPLFNPLKTGVCQLDNANVTDAILERAGGSIADFTGHYQTAFRELERVLNFPQSNLCLKREKQDESC
RR
R
R
E,R,V
SLTQALPSELKVSADCVSLTGAVSLASMLTEIFLLQQAQGMPEPGWGRITDSHQWNTLLSLHNAQFDLLQ
D
W
V
RTPEVARSRATPLLDLIKTALTPHPPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTP
A
Y
PGGELVFERWRLSDNSQWIQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ
IVNEARIPACSL

Figura 9

Fitase evoluída Nº	Nº Mutações	Mutações Possíveis															
		A47F	C97V	C97E	T136H	N159V	N159E	T163R	D164R	E168R	G179R	C226D	V233W	Q275V	R289A	T349Y	L363P
1	10X	A47F		C97E	T136H	N159V					G179R	C226D	V233W	Q275V	R289A	T349Y	
2	10X	A47F		C97E		N159V		T163R	D164R		G179R	C226D	V233W	Q275V		T349Y	
3	10X	A47F			T136H	N159V		T163R	D164R		G179R	C226D	V233W	Q275V		T349Y	
4	12X	A47F			T136H	N159V		T163R	D164R	E168R	G179R	C226D	V233W	Q275V	R289A	T349Y	
5	12X	A47F	C97V		T136H	N159V		T163R		E168R	G179R	C226D	V233W	Q275V	R289A	T349Y	
6	12X	A47F	C97V		T136H	N159V		T163R	D164R	E168R	G179R	C226D	V233W	Q275V		T349Y	
7	11X	A47F			T136H	N159V		T163R	D164R	E168R	G179R	C226D	V233W	Q275V		T349Y	
8	11X	A47F		C97E	T136H	N159V		T163R	D164R		G179R	C226D	V233W	Q275V		T349Y	
9	12X	A47F		C97E	T136H	N159V		T163R	D164R	E168R	G179R		V233W	Q275V	R289A	T349Y	
10*	12X	A47F		C97E	T136H		N159E	T163R	D164R		G179R		V233W	Q275V	R289A	T349Y	L363P
11	11X	A47F	C97V		T136H	N159V		T163R	D164R		G179R		V233W	Q275V	R289A	T349Y	
12	13X	A47F		C97E	T136H	N159V		T163R	D164R	E168R	G179R	C226D	V233W	Q275V	R289A	T349Y	
13	10X	A47F		C97E		N159V			D164R		G179R	C226D	V233W	Q275V	R289A	T349Y	
14	12X	A47F		C97E		N159V		T163R	D164R	E168R	G179R	C226D	V233W	Q275V	R289A	T349Y	
15	13X	A47F	C97V		T136H		N159E	T163R	D164R	E168R	G179R	C226D	V233W	Q275V	R289A	T349Y	
16	10X	A47F		C97E	T136H	N159V		T163R			G179R	C226D	V233W	Q275V		T349Y	
17	12X	A47F		C97E	T136H	N159V		T163R	D164R	E168R	G179R	C226D	V233W	Q275V		T349Y	
18	11X	A47F			T136H		N159E	T163R	D164R		G179R	C226D	V233W	Q275V	R289A	T349Y	
19**	9X	A47F			T136H	N159V			D164R		G179R	C226D	V233W	Q275V		T349Y	
20**	10X	A47F	C97V		T136H	N159V			D164R		G179R	C226D	V233W	Q275V		T349Y	
21**	9X	A47F	C97V		T136H	N159V			D164R		G179R		V233W	Q275V		T349Y	
22	6X	A47F			T136H	N159V						C226D	V233W			T349Y	

* Possui uma evolução adicional (L363P) não introduzida por GSSM™. Esta mutação foi introduzida por acaso aleatório e pode ou não ter qualquer relevância para a estabilidade térmica

** Possui um marcador Histidina (RSHHHHHH) C- terminal adicional

Figura 10

Alteração GSSM™	Amino Ácido Original	Amino Ácido Alterado para	Local de Mutação de Codão	Posições de Nucleótido de Codão Alterado na SEQ ID Nº 1	Codão Original	Codão Alterado para	Todos os Possíveis Codões para o Aminoácido Mutado
A47F	A	F	47	139-141	GCC	TTT	TTT, TTC
C97V	C	V	97	289-291	TGT	GTT	GTT, GTC, GTA, GTG
C97E	C	E	97	289-291	TGT	GAG	GAA, GAG
T136H	T	H	136	406-408	ACC	CAT	CAT, CAC
N159V	N	V	159	475-477	AAC	GTC or GTG	GTT, GTC, GTA, GTG
N159E	N	E	159	475-477	AAC	GAG	GAA, GAG
T163R	T	R	163	487-489	ACT	CGG	CGT, CGC, CGA, CGG, AGA, AGG
D164R	D	R	164	490-492	GAC	CGG	CGT, CGC, CGA, CGG, AGA, AGG
E168R	E	R	168	502-504	GAG	CGG	CGT, CGC, CGA, CGG, AGA, AGG
G179R	G	R	179	535-537	GGG	CGG	CGT, CGC, CGA, CGG, AGA, AGG
C226D	C	D	226	676-678	TGT	GAT	GAT, GAC
V233W	V	W	233	697-699	GTA	TGG	TGG
Q275V	Q	V	275	823-825	CAA	GTG	GTT, GTC, GTA, GTG
R289A	R	A	289	865-867	CGC	GCT	GCT, GCC, GCA, GCG
T349Y	T	Y	349	1045-1047	ACG	TAT	TAT, TAC
L363P	L	P	363	1087-1089	CTA	CCA	CCA, CCC, CCG, CCT

Figura 11

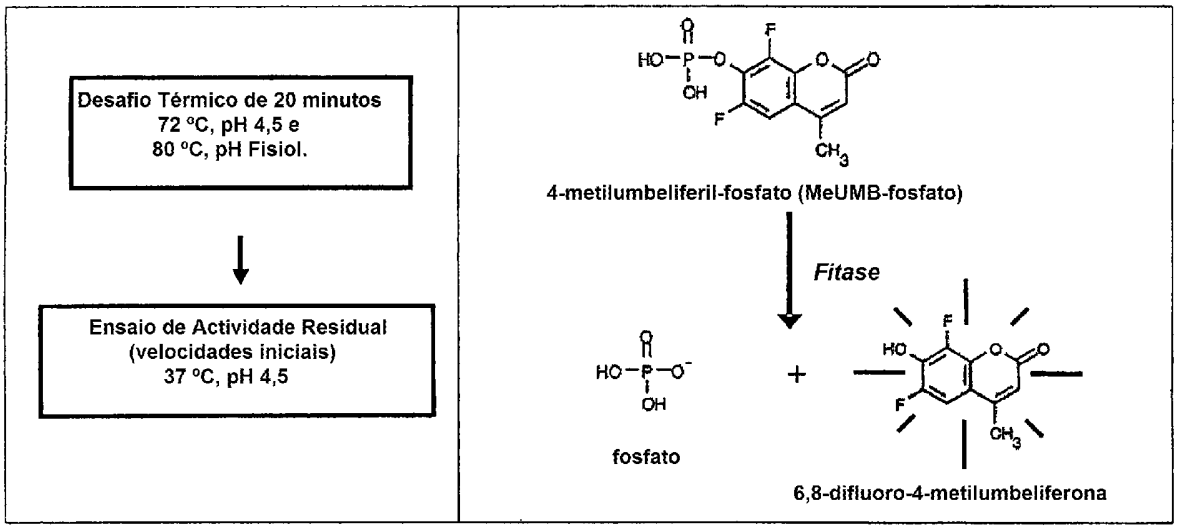

Desafio Térmico de 20 minutos
72 °C, pH 4,5 e
80 °C, pH Fisiol.
Ensaio de Actividade Residual
(velocidades iniciais)
37 °C, pH 4,5
4-metilumbeliferil-fosfato (MeUMB-fosfato)
Fitase
fosfato
6,8-difluoro-4-metilumbeliferona

Figura 12

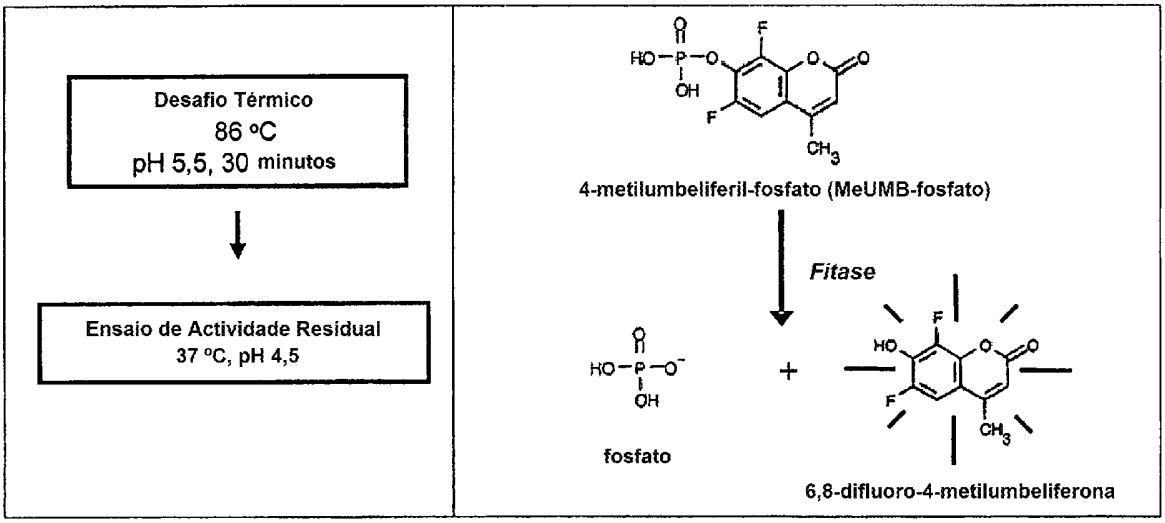
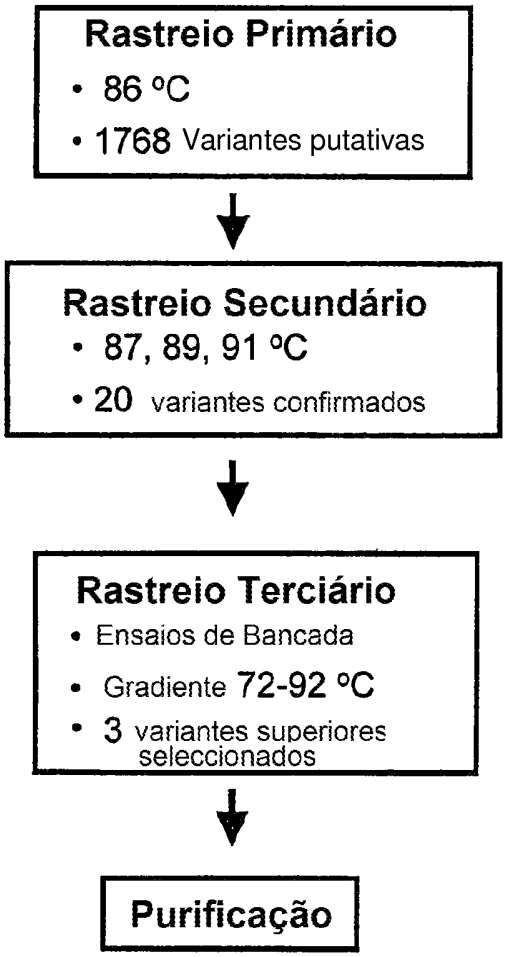

Desafio Térmico
86 °C
pH 5,5, 30 minutos
Ensaio de Actividade Residual
37 °C, pH 4,5
4-metilumbeliferil-fosfato (MeUMB-fosfato)
Fitase
fosfato
6,8-difluoro-4-metilumbeliferona

Figura 13


Rastreio Primário
• 86 °C
• 1768 Variantes putativas
Rastreio Secundário
• 87, 89, 91 °C
• 20 variantes confirmados
Rastreio Terciário
• Ensaios de Bancada
• Gradiente 72-92 °C
• 3 variantes superiores seleccionados
Purificação

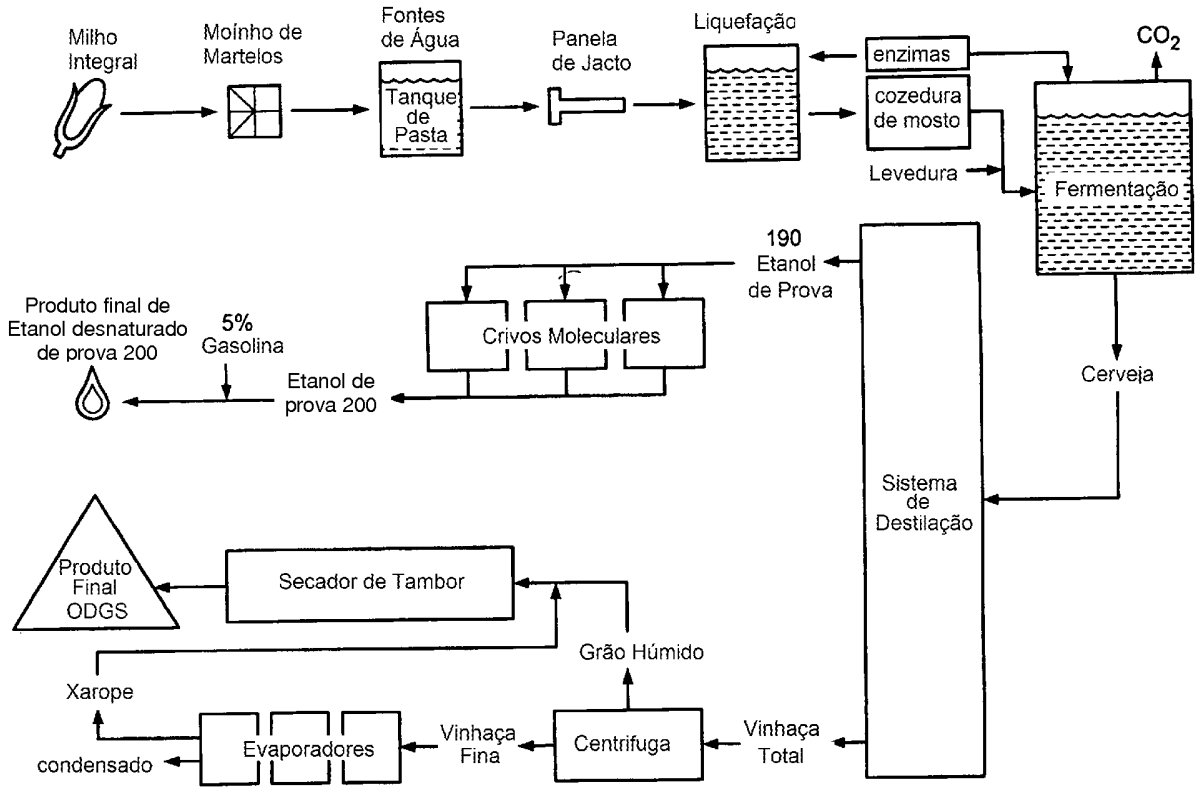

Milho Integral
Moinho de Martelos
Fontes de Água
Tanque de Pasta
Panela de Jacto
Liquefação
enzimas
cozedura de mosto
Levedura
CO_2
Fermentação
190 Etanol de Prova
Crivos Moleculares
Produto final de Etanol desnaturado de prova 200
5% Gasolina
Etanol de prova 200
Cerveja
Sistema de Destilação
Produto Final ODGS
Secador de Tambor
Grão Húmido
Xarope
condensado
Evaporadores
Vinhaça Fina
Centrifuga
Vinhaça Total

FIG. 14