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

(11) **PI 0208851-7 B1**

* B R P I 0 2 0 8 8 5 1 B 1 *

(22) Data de Depósito: 24/05/2002

(45) Data da Concessão: 21/07/2015
(RPI 2324)

(54) Título: FITASES RECOMBINANTES, PREPARAÇÃO DE PROTEÍNA, HETERODÍMERO, POLIPEPTÍDEO IMOBILIZADO, ARRANJO, SUPLEMENTO ALIMENTAR, MATRIZ DE LIBERAÇÃO DE ENZIMA COMESTÍVEL E PÉLETE COMESTÍVEL COMPREENDENDO AS MESMAS, BEM COMO MÉTODOS DE PRODUÇÃO DAS REFERIDAS ENZIMAS E PARA HIDROLISAR UM INOSITOL-HEXAFOSFATO, DESGRUDADURA DE ÓLEO E PROCESSAMENTO DE GRÃOS

(51) Int.Cl.: C11D12/00; C12P21/06; C12N9/16; C12N15/00; C12N5/00; C12N1/20; C07H21/04

(30) Prioridade Unionista: 24/05/2001 US 09/866,379

(73) Titular(es): Diversa Corporation, Subverday, Inc., Verenium Corporation

(72) Inventor(es): Eileen O'Donoghue, Eric J. Mathur, James B. Garrett, Jay M. Short, Keith Kretz, Kevin A. Gray, Nelson R. Barton

Relatório Descritivo da Patente de Invenção para **"FITASES RECOMBINANTES, PREPARAÇÃO DE PROTEÍNA, HETERODÍMERO, POLIPEPTÍDEO IMOBILIZADO, ARRANJO, SUPLEMENTO ALIMENTAR, MATRIZ DE LIBERAÇÃO DE ENZIMA COMESTÍVEL E PÉLETE COMESTÍVEL COMPREENDENDO AS MESMAS, BEM COMO MÉTODOS DE PRODUÇÃO DAS REFERIDAS ENZIMAS E PARA HIDROLISAR UM INOSITOL-HEXAFOSFATO, DESGRUDADURA DE ÓLEO E PROCESSAMENTO DE GRÃOS".**

REFERÊNCIA CRUZADA AOS PEDIDOS RELACIONADOS

Este pedido de patente é uma continuação-em-parte do Pedido de Patente U. S. Série No. (USSN) 09/866.379, depositado em 24 de maio de 2001.

CAMPO DA INVENÇÃO

Esta invenção refere-se aos polinucleotídeos recentemente produzidos, polipeptídeos codificados por tais polinucleotídeos, usos de tais polinucleotídeos e polipeptídeos, como também a produção e isolamento de tais polinucleotídeos e polipeptídeos. Em particular, a invenção fornece polipeptídeos tendo atividade de fitase, por exemplo, SEQ ID Nº: 1. A invenção fornece enzimas de fitase isoladas e recombinantes. As fitases podem ser produzidas por modificação do appA do tipo selvagem de E. coli ou elas podem ser produzidas de células hospedeiras recombinantes. As fitases da invenção podem ser usadas para ajudar na digestão de fitato. As fitases da invenção podem ser usadas em gêneros alimentícios para melhorar o valor alimentício de ingredientes ricos em fitato. As fitases da invenção podem ser termotolerantes e/ou termoestáveis. Também fornecido são métodos para obter um polinucleotídeo variante codificando uma fitase e para obter uma fitase com termoestabilidade ou termotolerância em temperaturas altas ou baixas.

FUNDAMENTO

Minerais são elementos essenciais para o desenvolvimento de todos os organismos. Minerais dietéticos podem ser derivados de muitos materiais-fontes, incluindo plantas. Por exemplo, sementes de plantas são

Segue-se folha 1a

uma fonte rica de minerais uma vez que elas contêm íons que são complexados com os grupos de fosfato das moléculas de ácido fítico. Estes minerais associados ao fitato podem, em alguns casos, satisfazer as necessidades dietéticas de algumas espécies de organismos criados no campo, como ruminantes com multiestômago. Conseqüentemente, em alguns casos, os

Segue-se folha 2

ruminantes requerem suplementação menos dietética com fosfato inorgânico e minerais porque os microorganismos no rume produzem enzimas que catalisam a conversão de fitato (mio-inositol-hexafosfato) em inositol e fosfato inorgânico. No processo, minerais que foram complexados com fitato são liberados. A maioria das espécies de organismos cultivados em campo, porém, não pode eficazmente utilizar minerais associados ao fitato. Desse modo, por exemplo, na produção de animais de criação de animais monogástricos (por exemplo, porcos, pássaros e peixe), a alimentação é comumente suplementada com minerais e/ou com substâncias antibióticas que alteram o ambiente da flora digestiva do organismo em consumo para intensificar as taxas de crescimento.

Como tal, há muitos fardos problemáticos - relacionados à nutrição, etapas de processamento *ex vivo*, saúde e medicina, conservação ambiental e administração de recursos - que estão associados com uma hidrólise insuficiente de fitato em muitas aplicações. O seguinte são exemplos não limitativos destes problemas:

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

2) A presença de fitato não-hidrolisado é indesejável e problemática em muitas aplicações *ex vivo* (por exemplo causando a presença de lama indesejada).

3) A suplementação de dietas com antibióticos propõe uma ameaça médica ao homem e animais semelhantes aumentando a abundância de patógenos tolerantes a antibióticos.

4) A descarga de minerais fecais não-absorvidos no ambiente rompem e danificam os ecossistemas de solos circunvizinhos, águas de criação de peixes e águas superficiais em geral.

5) Os valiosos oferecimentos nutricionais de muitos gêneros alimentícios potenciais permanecem significativamente inexplorados e desperdiçados.

Muitas plantas potencialmente nutritivas, incluindo particularmente suas sementes, contêm quantidades apreciáveis de nutrientes, por

exemplo fosfato, que estão associados com fitato de uma maneira de modo que estes nutrientes não estão livremente disponíveis no consumo. A indisponibilidade destes nutrientes é pelo menos parcialmente superada por alguns organismos, incluindo vacas e outros ruminantes tendo uma capacidade digestiva suficiente - largamente derivada da presença de formas de vida simbiôntica em seus tratos digestivos - para hidrolisar fitato e liberar os nutrientes associados. Porém, a maioria das espécies de animais do campo, incluindo porcos, peixe, galinhas, perus, como também outros organismos não-ruminantes incluindo o homem, não ineficazes de liberar estes nutrientes eficazmente após ingestão.

Por conseguinte, gêneros alimentícios contendo fitato requerem suplementação com nutrientes exógenos e/ou com uma fonte de atividade de fitase para melhorar seus oferecimentos nutricionais deficientes sob consumo por um número muito grande de espécies de organismos.

Em ainda outro aspecto, a presença de fitato não-hidrolisado conduz às conseqüências problemáticas em processos *ex vivo* incluindo - mas não limitado - o processamento de gêneros alimentícios. Em uma mera exemplificação, como descrito em EP0321004-B1 (Vaara et al.), há uma etapa no processamento de grãos de milho e de sorgo por meio do qual os grãos duros são macerados em água para amaciá-los. As substâncias solúveis em água que se lixiviam durante este processo se tornam parte de um líquido de maceração de milho que é concentrado por evaporação. Ácido fítico não-hidrolisado no líquido de maceração do milho, em grande parte na forma de sais de cálcio e magnésio, é associado com fósforo e deposita uma lama indesejável com proteínas e íons de metal. Esta lama é problemática na evaporação, transporte e armazenamento do líquido de maceração do milho.

A suplementação de dietas com substâncias antibióticas tem muitos resultados benéficos na produção de animais de criação. Por exemplo, além de seu papel como meios profiláticos para repelir doença, a administração de antibióticos exógenos foi mostrada aumentar as taxas de crescimento acima de 3-5 %. O mecanismo desta ação também pode envolver -

em parte - uma alteração no ambiente da flora digestiva de animais do campo, resultando em um equilíbrio microfloral que está mais ideal para absorção de nutriente.

Porém, um efeito negativo significativo associado com a overdose de antibióticos é o perigo de criar um repositório de cepas microbianas patogênicas resistentes a antibióticos patogênicas. Este perigo é iminente, e a elevação de patógenos resistentes a drogas em seres humanos já esteve ligada ao uso de antibióticos em animais de criação. Por exemplo, Avoparcin, o antibiótico usado nas alimentações de animais, foi proibido em muitos lugares em 1997, e os animais estão sendo agora dado outro antibiótico, virginiamicina, que é bem parecido à nova droga Synercid, usada para substituir vancomicina em seres humanos. Porém, estudos já mostraram que alguns enterococos em animais de fazenda são resistentes à Synercid. Por conseguinte, conseqüências indesejadas de tolerância, como aquelas já vistas com Avoparcin e vancomicina, são prováveis reocorrer não importa que novos antibióticos são usados como profiláticos gerais para animais do campo. Conseqüentemente, investigadores estão pedindo controles mais rígidos no uso da droga na indústria.

Os aumentos nas taxas de crescimento alcançados em animais criados com gêneros alimentícios suplementados com as moléculas de fitase imediatamente divulgadas equiparam - se não exceder - a aqueles alcançados usando antibióticos como, por exemplo, Avoparcin. Conseqüentemente, as moléculas de fitase imediatamente descritas - ou sozinhas ou em combinação com outros reagentes (incluindo mas não limitado a enzimas, incluindo proteases) - não só são úteis nesta aplicação (por exemplo, por aumentar a taxa de crescimento de animais do campo) mas também em outras aplicações onde hidrólise de fitato é desejável.

Uma conseqüência ambiental é que o consumo de gêneros alimentícios contendo fitato por qualquer espécie de organismo que é deficiente em fitase - independente se os gêneros alimentícios são suplementados com minerais - conduz à poluição fecal resultante da excreção dos minerais não-absorvidos. Esta poluição não só tem um impacto negativo no ha-

bitat imediato mas conseqüentemente também nas águas circunvizinhas. As alterações ambientais ocorrem primariamente na base da cadeia alimentar, e portanto tem o potencial para penetrar para cima e em todo um ecossistema para efetuar dano permanente e catastrófico - particularmente após anos de poluição ininterrupta. Este problema tem o potencial de se manifestar em qualquer área onde o processamento de fitato concentrado ocorre - incluindo etapas de processamento *in vivo* (por exemplo por animais em áreas de produção de animais de criação, pisos de zoológicos, refúgios da vida selvagem, etc.) e *in vitro* (por exemplo em moenda a úmido de milho comercial, processos de maceração de cereais e outros).

A decisão de usar moléculas de fitase adicionadas de forma exógena - ou para substituir completamente ou aumentar o uso de minerais e/ou antibióticos administrados de modo exógeno - por fim necessita passar por um teste de viabilidade financeira e eficácia de custo pelo usuário cujo sustento depende da aplicação pertinente, como produção de animais de criação.

Por conseguinte, há uma necessidade por meios para alcançar hidrólise de fitado eficiente e de custo eficaz em várias aplicações. Particularmente, há uma necessidade por meios para otimizar a hidrólise de fitato em aplicações comerciais. Em um aspecto particular, há uma necessidade de otimizar métodos de tratamento comerciais que melhoram os oferecimentos nutricionais de gêneros alimentícios contendo fitato para consumo por seres humanos e animais do campo.

Fitato ocorre como uma fonte de fósforo armazenado em virtualmente todas alimentações vegetais (Graf (Ed.), 1986). Ácido fítico forma uma parte normal da semente em cereais e legumes. Ele funciona para ligar minerais dietéticos que são essenciais à nova planta enquanto ela emerge da semente. Quando os grupos de fosfato de ácido fítico forem removidos pela fitase da enzima da semente, a capacidade de ligar íons de metal é perdida e os minerais ficam disponíveis à planta. Em grãos de alimentação de animais de criação, os minerais de traço ligados através do ácido fítico encontram-se largamente indisponíveis para absorção por animais mono-

gástricos, que carecem de atividade de fitase.

Embora alguma hidrólise de fitato ocorra no cólon, a maior parte do fitato passa pelo trato gastrintestinal dos animais monogástricos e é excretada no adubo contribuindo para problemas de poluição de fosfato fecal em áreas de intensa produção de animais de criação. Fósforo inorgânico liberado no cólon tem um valor nutricional apreciavelmente diminuído em animais de criação porque o fósforo inorgânico é principalmente absorvido - se não de modo virtual exclusivamente - no intestino delgado. Desse modo, uma quantidade apreciável dos minerais dietéticos nutricionalmente importantes em fitato encontra-se indisponível para animais monogástricos.

Em suma, nutrientes associados ao fitato são compreendidos não só de fosfato que é covalentemente ligado ao fitato, mas também outros minerais que são quelados pelo fitato também. Além disso, ao ingerir, o fitato não-hidrolisado pode também encontrar e tornar-se associado aos minerais adicionais. A quelação dos minerais pode inibir a atividade das enzimas às quais estes minerais servem como co-fatores.

A conversão de fitato em inositol e fósforo inorgânico pode ser catalisada por enzimas microbianas referidas amplamente como fitases. As fitases, como fitase Nº EC 3.1.3.8, são capazes de catalisar a hidrólise de mio-inositol-hexafosfato em D-mio-inositol 1,2,4,5,6-pentafosfato e ortofosfato. Certas fitases fúngicas, segundo informações, hidrolisam o pentafosfato de inositol em fosfatos tetra-, tri- e mais baixos. Por exemplo, fitases A. Ficuum, segundo informações, produzem misturas de di e monofosfato de mio-inositol (Ullah, 1988). Microorganismos produtores de fitase são compreendidos de bactérias como Bacilo subtilis (Powar e Jagannathan, 1982) e Pseudomonas (Cosgrove, 1970); leveduras como Sacchoromyces cerevisiae (Nayini e Markakis, 1984); e fungos como Aspergillus terreus (Yamada et al., 1968).

Ácido fosfatases são enzimas que cataliticamente hidrolisam uma ampla variedade de ésteres de fosfato e usualmente apresentam ótima de pH abaixo de 6,0 (Igarashi e Hollander, 1968). Por exemplo, enzimas de Nº EC 3.1.3.2 catalisam a hidrólise de monoésteres ortofosfóricos em pro-

dutos de ortofosfato. Uma ácido fosfatase foi purificada, segundo informações, de A. ficuum. A forma desglicosilada do ácido fosfatase tem um peso molecular aparente de 32,6 kDa (Ullah et al., 1987).

Fitase e ácido fosfatases menos específicas são produzidas pelo fungo Aspergillus ficuum como enzimas extracelulares (Shieh et al., 1969)., segundo informações, Ullah purificou uma fitase de A. ficuum do tipo selvagem que tinha um peso molecular aparente de 61,7 kDA (em SDS-PAGE; conforme corrigido para glicosilação); ótima de pH em pH 2,5 e pH 5,5; um Km de cerca de 40 μm; e uma atividade específica de cerca de 50 U/mg (Ullah, 1988). Pedido de Patente do PCT WO 91/05053 também descreve, segundo informações, isolamento e clonagem molecular de uma fitase de Aspergillus ficuum com ótima de pH em pH 2,5 e pH 5,5, um Km de cerca de 250 μm e atividade específica de cerca de 100 U/mg de proteína. Resumidamente, a atividade específica previamente citada para estas enzimas microbianas informadas foi aproximadamente na faixa de 50-100 U/mg de proteína.

A possibilidade de usar micróbios capazes de produzir fitase como um aditivo alimentício para animais monogástricos foi previamente informado (Patente U. S. No. 3.297.548 Shieh e Ware; Nelson et al., 1971). A eficácia de custo deste método tem sido uma limitação principal para esta e outras aplicações comerciais. Portanto, as moléculas de fitase melhoradas são altamente desejáveis.

Fitases microbianas também podem ser, segundo informações, úteis para produzir alimentação de animais de certos processos industriais, por exemplo, produtos residuais de trigo e milho. Em um aspecto, o processo de moenda a úmido de milho produz gluténs vendidos como alimentações de animais. A adição de fitase pode melhorar o valor nutricional do produto de alimentação, segundo informações. Por exemplo, o uso de enzimas de fitases fúngicas e condições de processo (t~50°C e pH $\simeq$ 5,5) foram previamente informados em (por exemplo EP 0 321 004). Brevemente, no processamento de farinha de soja usando métodos tradicionais de maceração, isto é, métodos sem a adição de enzima de fitase exógena, a presença de fitato

não-hidrolisado, segundo informações, rende farinha e sobras inadequadas para as alimentações usadas na criação de peixe, aves e outros não-ruminantes como também bezerros alimentados no leite. Fitase é, segundo informações, útil para melhorar o valor nutricional e comercial deste material de soja de alta proteína (ver Finase Enzymes por Alko, Rajamäki, Finlândia). Uma combinação de fitase fúngica e uma forma de ácido fosfatase de ótima de pH 2,5 A. niger foi usado por Alko, Ltd como um suplemento de alimentação de animais em seus produtos degradativos de ácido fítico Finas F e Finase S. Porém, a eficácia de custo deste método permaneceu uma limitação principal ao uso mais difundido. Desse modo uma fonte eficaz no custo de fitase grandemente intensificaria o valor das farinhas de feijão-soja como uma alimentação de animais (Shieh et al., 1969).

Para solucionar os problemas descritos, foi proposto o tratamento de gêneros alimentícios com enzimas de fitase exógena, mas este método não foi otimizado por completo, particularmente com respeito à viabilidade e eficiência de custo. Esta otimização requer a consideração que uma faixa extensiva de aplicações exista, particularmente para produção em grande escala. Por exemplo, há uma faixa extensiva de gêneros alimentícios, métodos de preparação destes e espécies de organismos recipientes.

Em uma exemplificação particular, é apreciado que a fabricação de péletes de alimentação de peixe requer exposição de ingredientes a altas temperaturas e/ou pressão para produzir péletes que não dissolvam e/ou degradem prematuramente (por exemplo antes do consumo) na sujeição à água. Desse modo, seria desejável para este processo industrial obter enzimas aditivas que são estáveis sob condições altas de temperatura e/ou pressão. Conseqüentemente é apreciado que as fitases distintas podem ser diferentemente preferíveis ou ideais para aplicações distintas.

É reconhecido, além disso, que um modo importante para otimizar um processo enzimático é através da modificação e melhoria da enzima catalítica pivotal. Por exemplo, uma planta transgênica pode ser formada que é compreendida de um sistema de expressão para expressar uma molécula de fitase. É apreciado que tentando melhorar os fatores que não estão

relacionados diretamente com a atividade da própria molécula expressa, como o nível de expressão, apenas um nível finito - e potencialmente insuficiente - de otimização pode ser de forma máxima alcançado. Conseqüentemente, também há uma necessidade de obter moléculas com características melhoradas.

SUMÁRIO DA INVENÇÃO

A invenção fornece um ácido nucléico isolado ou recombinante compreendendo uma seqüência de ácidos nucléicos tendo pelo menos 98 % de identidade de seqüência à SEQ ID Nº: 1 em uma região de pelo menos cerca de 100 resíduos, em que os ácidos nucléicos codificam pelo menos um polipeptídeo tendo uma atividade de fitase e as identidades das seqüências são determinadas por análise com um algoritmo de comparação de seqüência ou por uma inspeção visual. Em modalidades alternativas, a seqüência de ácidos nucléicos tem pelo menos 98 % de identidade de seqüência à SEQ ID Nº: 1 em uma região de pelo menos cerca de 50 resíduos, 100 resíduos, 150 resíduos, 200 resíduos, 250 resíduos, 300 resíduos, 350 resíduos, 400 resíduos, 450 resíduos, 500 resíduos, 550 resíduos, 600 resíduos, 700 resíduos, 800 resíduos, 900 resíduos, 1000 resíduos, 1200 resíduos ou 1300 resíduos.

Em modalidades alternativas a seqüência de ácidos nucléicos tem pelo menos 98 %, 98,5 %, 99 % ou 99,5 % de identidade de seqüência à SEQ ID Nº: 1 em uma região de pelo menos cerca de 50 resíduos, 100 resíduos, 150 resíduos, 200 resíduos, 250 resíduos, 300 resíduos, 350 resíduos, 400 resíduos, 450 resíduos, 500 resíduos, 550 resíduos, 600 resíduos, 700 resíduos, 800 resíduos, 900 resíduos, 1000 resíduos, 1200 resíduos ou 1300 resíduos. A seqüência de ácidos nucléicos pode ter uma seqüência como exposto na SEQ ID Nº: 1. A seqüência de ácidos nucléicos pode codificar um polipeptídeo tendo uma seqüência como exposto na SEQ ID Nº: 2.

Em um aspecto, o algoritmo de comparação de seqüência é um algoritmo de BLAST versão 2.2.2. O ajuste de filtração pode ser ajustado para blastall -p blastp - d "nr pataa", e todas as outras opções são ajustadas em default.

Em um aspecto, a atividade de fitase compreende catálise de fitato (mio-inositol-hexafosfato) em inositol e fosfato inorgânico, ou equivalente. A atividade de fitase pode compreender a hidrólise de fitato (mio-inositol-hexafosfato).

Em um aspecto, a atividade de fitase pode ser termoestável ou termotolerante. O polipeptídeo pode reter uma atividade de fitase sob condições compreendendo uma faixa de temperatura de entre cerca de 40°C a cerca de 70°C. O polipeptídeo pode reter uma atividade de fitase após exposição a uma temperatura na faixa de mais que 37°C a cerca de 90°C. O polipeptídeo pode reter uma atividade de fitase após exposição a uma temperatura na faixa de mais que 37°C a cerca de 50°C.

A invenção fornece um ácido nucléico isolado ou recombinante compreendendo uma seqüência que hibrida sob condições rigorosas com uma seqüência de ácidos nucléicos como exposto na SEQ ID Nº: 1, em que o ácido nucléico codifica um polipeptídeo tendo uma atividade de fitase. O ácido nucléico é pelo menos cerca de 50 resíduos, 100 resíduos, 150 resíduos, 200 resíduos, 250 resíduos, 300 resíduos, 350 resíduos, 400 resíduos, 450 resíduos, 500 resíduos, 550 resíduos, 600 resíduos, 700 resíduos, 800 resíduos, 900 resíduos, 1000 resíduos, 1200 resíduos ou 1300 resíduos no comprimento. Em um aspecto, as condições rigorosas incluem uma etapa de lavagem compreendendo uma lavagem em 0,2X SSC em uma temperatura de cerca de 65°C durante cerca de 15 minutos.

A invenção fornece uma sonda de ácido nucléico para identificar um ácido nucléico codificando um polipeptídeo com uma atividade de fitase, em que a sonda compreende pelo menos 10 bases sucessivas de uma seqüência selecionada de um grupo que consiste em uma seqüência como exposto na SEQ ID Nº: 1, em que a sonda identifica o ácido nucléico por ligação ou hibridação. A sonda pode compreender um oligonucleotídeo compreendendo pelo menos cerca de 10 a 50, cerca de 20 a 60, cerca de 30 a 70, cerca de 40 a 80 ou cerca de 60 a 100 bases sucessivas de uma seqüência como exposto na SEQ ID Nº: 1.

A invenção fornece uma sonda de ácido nucléico para identificar

um ácido nucléico codificando um polipeptídeo com uma atividade de fitase, em que a sonda compreende uma seqüência de ácidos nucléicos tendo pelo menos 98 % de identidade de seqüência à SEQ ID Nº: 1 em uma região de pelo menos cerca de 100 resíduos, em que as identidades das seqüências são determinadas por análise com um algoritmo de comparação de seqüência ou através de inspeção visual. A sonda pode compreender um oligonucleotídeo compreendendo pelo menos cerca de 10 a 50, cerca de 20 a 60, cerca de 30 a 70, cerca de 40 a 80 ou cerca de 60 a 100 bases sucessivas de uma seqüência de ácidos nucléicos como exposto na SEQ ID Nº: 1. A sonda pode compreender uma seqüência de ácidos nucléicos tendo pelo menos 99 % de identidade de seqüência a uma seqüência de ácidos nucléicos como exposto na SEQ ID Nº: 1. A sonda pode compreender um subconjunto de uma seqüência como exposto na SEQ ID Nº: 1.

A invenção fornece um par de seqüências de iniciadores de amplificação para amplificar um ácido nucléico codificando um polipeptídeo com uma atividade de fitase, em que o par de iniciadores é capaz de amplificar uma seqüência de ácidos nucléicos como exposto na SEQ ID Nº: 1. Em um aspecto, cada membro das seqüências de iniciadores de amplificação compreende um oligonucleotídeo compreendendo pelo menos cerca de 10 a 50 bases sucessivas da seqüência.

A invenção fornece um método de amplificar um ácido nucléico codificando um polipeptídeo com uma atividade de fitase compreendendo amplificação de um ácido nucléico modelo com um par de seqüências de iniciadores de amplificação capaz de amplificar uma seqüência de ácidos nucléicos conforme exposto na SEQ NO: 1.

A invenção fornece um cassete de expressão compreendendo um ácido nucléico da invenção, por exemplo, uma seqüência de ácidos nucléicos pelo menos 98 % de identidade de seqüência à SEQ ID Nº: 1 em uma região de pelo menos cerca de 100 resíduos, em que as identidades das seqüências são determinadas por análise com um algoritmo de comparação de seqüência ou através de inspeção visual, ou, um ácido nucléico que hibrida sob condições rigorosas com uma seqüência de ácidos nucléicos

como exposto na SEQ ID Nº: 1, ou uma subseqüência desta.

A invenção fornece um vetor compreendendo um ácido nucléico da invenção, por exemplo, uma seqüência de ácidos nucléicos pelo menos 98 % de identidade de seqüência à SEQ ID Nº: 1 em uma região de pelo menos cerca de 100 resíduos, em que as identidades das seqüências são determinadas por análise com um algoritmo de comparação de seqüência ou através de inspeção visual, ou, um ácido nucléico que hibrida sob condições rigorosas com uma seqüência de ácidos nucléicos como exposto na SEQ ID Nº: 1, ou uma subseqüência desta.

A invenção fornece um veículo de clonagem compreendendo um vetor da invenção ou um ácido nucléico da invenção, em que o veículo de clonagem compreende um vetor virótico, um plasmídeo, um fago, um fagemídeo, um cosmídeo, um fosmídeo, um bacteriófago ou um cromossomo artificial. O vetor virótico pode compreender um vetor de adenovírus, vetores retrovirótico ou um vetor virótico adeno-associado. O vetor virótico pode compreender um cromossomo artificial bacteriano (BAC), um plasmídeo, um vetor derivado de bacteriófago P1 (PAC), um cromossomo artificial de levedura (YAC), um cromossomo artificial mamífero (MAC).

A invenção fornece uma célula transformada compreendendo um vetor da invenção ou um ácido nucléico da invenção. O vetor pode compreender uma seqüência de ácidos nucléicos pelo menos 98 % de identidade de seqüência à SEQ ID Nº: 1 em uma região de pelo menos cerca de 100 resíduos, em que as identidades das seqüências são determinadas por análise com um algoritmo de comparação de seqüência ou através de inspeção visual, ou, um ácido nucléico que hibrida sob condições rigorosas com uma seqüência de ácidos nucléicos como exposto na SEQ ID Nº: 1, ou uma subseqüência desta.

A invenção fornece uma célula transformada compreendendo um vetor da invenção ou um ácido nucléico da invenção. O ácido nucléico pode compreender uma seqüência de ácidos nucléicos pelo menos 98 % de identidade de seqüência à SEQ ID Nº: 1 em uma região de pelo menos cerca de 100 resíduos, em que as identidades das seqüências são determina-

das por análise com um algoritmo de comparação de seqüência ou através de inspeção visual, ou o ácido nucléico hibrida sob condições rigorosas com uma seqüência de ácidos nucléicos como exposto na SEQ ID Nº: 1, ou uma subseqüência desta. Em aspectos alternativos, a célula é uma célula bacteriana, uma célula mamífera, uma célula fúngica, uma célula de levedura, uma célula de inseto ou uma célula vegetal.

A invenção fornece um animal não-humano transgênico compreendendo um vetor da invenção ou um ácido nucléico da invenção. Em um aspecto, o ácido nucléico compreende pelo menos 98 % de identidade de seqüência à SEQ ID Nº: 1 em uma região de pelo menos cerca de 100 resíduos, em que as identidades das seqüências são determinadas por análise com um algoritmo de comparação de seqüência ou através de inspeção visual, ou, compreende um ácido nucléico que hibrida sob condições rigorosas com uma seqüência de ácidos nucléicos como exposto na SEQ ID Nº: 1, ou uma subseqüência desta. Em aspectos alternativos, animal não-humano transgênico é um camundongo, uma cabra ou um porco.

A invenção fornece uma planta transgênica compreendendo um vetor da invenção ou um ácido nucléico da invenção. Em um aspecto, a seqüência de ácidos nucléicos tem pelo menos 98 % de identidade de seqüência à SEQ ID Nº: 1 em uma região de pelo menos cerca de 100 resíduos, em que as identidades das seqüências são determinadas por análise com um algoritmo de comparação de seqüência ou através de inspeção visual, ou o ácido nucléico hibrida sob condições rigorosas com uma seqüência de ácidos nucléicos como exposto na SEQ ID Nº: 1, ou uma subseqüência desta. Em aspectos alternativos, a planta é uma planta de milho, uma planta de batata, uma planta de tomate, uma planta de trigo, uma planta de azeitona, uma planta de colza, uma planta de feijão-soja ou uma planta de tabaco.

A invenção fornece uma semente transgênica compreendendo um vetor da invenção ou um ácido nucléico da invenção. Em um aspecto, a seqüência de ácidos nucléicos tem pelo menos 98 % de identidade de seqüência à SEQ ID Nº: 1 em uma região de pelo menos cerca de 100 resíduos, em que as identidades das seqüências são determinadas por análise

com um algoritmo de comparação de seqüência ou através de inspeção visual, ou o ácido nucléico hibrida sob condições rigorosas com uma seqüência de ácidos nucléicos como exposto na SEQ ID Nº: 1, ou uma subseqüência desta. Em aspectos alternativos, a semente é uma semente de milho, um grão de trigo, uma azeitona, uma colza, uma semente de feijão-soja, um grão de palma, uma semente de girassol, uma semente de gergelim, um amendoim ou uma semente de planta de tabaco.

A invenção fornece um oligonucleotídeo anti-sentido compreendendo uma seqüência de ácidos nucléicos complementar ou capaz de hibridar sob condições rigorosas com uma seqüência de ácidos nucléicos pelo menos 98 % de identidade de seqüência à SEQ ID Nº: 1 em uma região de pelo menos cerca de 100 resíduos, em que as identidades das seqüências são determinadas por análise com um algoritmo de comparação de seqüência ou através de inspeção visual, ou, um ácido nucléico que hibrida sob condições rigorosas com uma seqüência de ácidos nucléicos como exposto na SEQ ID Nº: 1, ou uma subseqüência desta, ou, uma seqüência de ácidos nucléicos como exposto na SEQ ID Nº: 1. O oligonucleotídeo anti-sentido pode ser entre cerca de 10 a 50, cerca de 20 a 60, cerca de 30 a 70, cerca de 40 a 80 ou cerca de 60 a 100 bases no comprimento.

A invenção fornece um método de inibir a tradução de uma mensagem de fitase em uma célula compreendendo administrar à célula ou expressar na célula um oligonucleotídeo anti-sentido compreendendo uma seqüência de ácidos nucléicos complementar ou capaz de hibridar sob condições rigorosas com uma seqüência de ácidos nucléicos pelo menos 98 % de identidade de seqüência à SEQ ID Nº: 1 em uma região de pelo menos cerca de 100 resíduos, em que as identidades das seqüências são determinadas por análise com um algoritmo de comparação de seqüência ou através de inspeção visual, ou, um ácido nucléico que hibrida sob condições rigorosas com uma seqüência de ácidos nucléicos como exposto na SEQ ID Nº: 1, ou uma subseqüência desta.

A invenção fornece um polipeptídeo isolado ou recombinante compreendendo uma seqüência de aminoácidos tendo pelo menos 98 % de

identidade de seqüência à SEQ ID Nº: 2 em uma região de pelo menos cerca de 100 resíduos, ou, um polipeptídeo codificado por um ácido nucléico compreendendo uma seqüência: (i) tendo pelo menos 98 % de identidade de seqüência à SEQ ID Nº: 1 em uma região de pelo menos cerca de 100 resíduos, em que as identidades das seqüências são determinadas por análise com um algoritmo de comparação de seqüência ou através de inspeção visual, ou, (ii) que hibrida sob condições rigorosas com um ácido nucléico como exposto na SEQ ID Nº: 1. Em um aspecto, o polipeptídeo tem uma atividade de fitase. A atividade de fitase pode compreender a hidrólise de fitato (mio-inositol-hexafosfato).

Em um aspecto, o polipeptídeo isolado ou recombinante tem um fenótipo termotolerante, isto é, sua atividade de fitase é termotolerante. Em um aspecto, o polipeptídeo isolado ou recombinante tem um fenótipo termoestável, isto é, sua atividade de fitase é termoestável.

Em aspectos alternativos, o polipeptídeo isolado ou recombinante (seqüência de aminoácidos) da invenção tem pelo menos 98 % de identidade de seqüência à SEQ ID Nº: 2 em uma região de pelo menos cerca de 100 resíduos, em que as identidades das seqüências são determinadas por análise com um algoritmo de comparação de seqüência ou por uma inspeção visual. Em modalidades alternativas, a seqüência de aminoácidos tem pelo menos 98 % de identidade de seqüência à SEQ ID Nº: 2 em uma região de pelo menos cerca de 50 resíduos, 100 resíduos, 150 resíduos, 200 resíduos, 250 resíduos, 300 resíduos, 350 resíduos, 400 resíduos ou 435 resíduos.

Em aspectos alternativos, o polipeptídeo isolado ou recombinante (seqüência de aminoácidos) da invenção tem pelo menos 98 %, 98,5 %, 99 % ou 99,5 % de identidade de seqüência à SEQ ID Nº: 2 em uma região de pelo menos cerca de 50 resíduos, 100 resíduos, 150 resíduos, 200 resíduos, 250 resíduos, 300 resíduos, 350 resíduos, 400 resíduos ou 435. O polipeptídeo isolado ou recombinante (seqüência de aminoácidos) da invenção pode ter uma seqüência como exposto na SEQ ID Nº: 2. O polipeptídeo pode ser codificado por um ácido nucléico tendo uma seqüência como ex-

posto na SEQ ID Nº: 1. Em um aspecto, o algoritmo de comparação de seqüência é um algoritmo de BLAST versão 2.2.2. O ajuste de filtração pode ser ajustado para blastall -p blastp -d "nr pataa" -F F, e todas as outras opções são ajustadas em default.

O polipeptídeo isolado ou recombinante (seqüência de aminoácidos) da invenção pode ter uma atividade de fitase. Em um aspecto, a atividade de fitase compreende catálise de fitato (mio-inositol-hexafosfato) com inositol e fosfato inorgânico, ou equivalente. A atividade de fitase pode compreender a hidrólise de fitato (mio-inositol-hexafosfato).

A invenção fornece um polipeptídeo isolado ou recombinante, em que o polipeptídeo tem uma atividade de fitase e carece de uma seqüência de sinal e compreende uma seqüência de aminoácidos da invenção, por exemplo, uma seqüência tendo pelo menos 98 % de identidade de seqüência à SEQ ID Nº: 2 em uma região de pelo menos cerca de 100 resíduos, ou, um polipeptídeo codificado por um ácido nucléico compreendendo uma seqüência: (i) tendo pelo menos 98 % de identidade de seqüência à SEQ ID Nº: 1 em uma região de pelo menos cerca de 100 resíduos, em que as identidades das seqüências são determinadas por análise com um algoritmo de comparação de seqüência ou através de inspeção visual, ou, (ii) que hibrida sob condições rigorosas com um ácido nucléico como exposto na SEQ ID Nº: 1. Em aspectos alternativos, a atividade de fitase compreende uma termoestabilidade quando aquecida para uma temperatura na faixa de cerca de 37°C a cerca de 50°C, cerca de 50°C a cerca de 70°C ou aproximadamente 70°C a cerca de 90°C. A atividade de fitase termoestável pode compreender uma atividade específica a cerca de 37°C na faixa de cerca de 100 a cerca de 1000 unidades por miligrama de proteína. A atividade de fitase termoestável pode compreender uma atividade específica de cerca de 500 a cerca de 750 unidades por miligrama de proteína. A atividade de fitase termoestável pode compreender uma atividade específica a 37°C na faixa de cerca de 500 a cerca de 1200 unidades por miligrama de proteína. A atividade de fitase termoestável pode compreender uma atividade específica a 37°C na faixa de cerca de 750 a cerca de 1000 unidades por miligrama de proteína. A ati-

vidade de fitase pode ser termotolerante após ser aquecida para uma temperatura elevada na faixa de cerca de 37°C a cerca de 90°C, ou, após ser aquecida para uma temperatura na faixa de cerca de 37°C a cerca de 70°C. A termotolerância pode compreender retenção de pelo menos metade da atividade específica da fitase a 37°C após ser aquecida para temperatura elevada. A termotolerância pode compreender retenção de atividade específica a 37°C na faixa de cerca de 500 a cerca de 1200 unidades por miligrama de proteína após ser aquecida para temperatura elevada. A fitase pode compreender pelo menos um sítio de glicosilação. A glicosilação pode ser uma ou mais glicosilações N-ligadas ou uma ou mais glicosilações N-ligadas ou uma combinação destas. A fitase pode ser glicosilada *in vitro* ou *in vivo*, por exemplo, após ser expressada em uma célula, por exemplo, células eucarióticas, por exemplo, uma célula de levedura, por exemplo, P. pastoris ou um S. pombe, ou uma célula de inseto, ou uma célula mamífera, por exemplo, uma célula humana. Em um aspecto, o polipeptídeo retém uma atividade de fitase sob condições acídicas, por exemplo, condições compreendendo cerca de pH 5, ou, sob condições compreendendo cerca de pH 4,5.

A invenção fornece uma preparação de proteína compreendendo um polipeptídeo da invenção, em que a preparação de proteína compreende um líquido, um sólido ou um gel.

A invenção fornece um heterodímero compreendendo um polipeptídeo da invenção e um segundo domínio. Em um aspecto, o segundo domínio é um polipeptídeo e o heterodímero é uma proteína de fusão. O segundo domínio pode ser um epítopo ou um marcador ou uma combinação destes.

A invenção fornece um polipeptídeo imobilizado tendo uma atividade de fitase, em que o polipeptídeo compreende uma seqüência de um polipeptídeos (aminoácidos) da invenção ou um heterodímero ou proteína de fusão da invenção. Em um aspecto, a fitase é imobilizada em uma célula, um metal, uma resina, um polímero, uma cerâmica, um vidro, um microeletrodo, uma partícula grafítica, uma conta, um gel, uma placa, um arranjo ou um tubo capilar.

A invenção fornece um arranjo compreendendo um polipeptídeo imobilizado da invenção ou um heterodímero ou proteína de fusão da invenção, ou um ácido nucléico da invenção.

A invenção fornece um anticorpo isolado ou recombinante que especificamente liga a um polipeptídeo da invenção ou a um polipeptídeo codificado por um ácido nucléico da invenção. O anticorpo pode ser um anticorpo monoclonal ou um policlonal. A invenção fornece um hibridoma compreendendo um anticorpo que especificamente liga a um polipeptídeo da invenção ou a um polipeptídeo codificado por um ácido nucléico da invenção.

A invenção fornece um suplemento alimentar para um animal compreendendo um polipeptídeo da invenção ou um polipeptídeo codificado por um ácido nucléico da invenção. O polipeptídeo no suplemento alimentar pode ser glicosilado.

A invenção fornece uma matriz de liberação de enzima comestível compreendendo um polipeptídeo da invenção ou um polipeptídeo codificado por um ácido nucléico da invenção, em que o polipeptídeo compreende uma atividade de fitase. A matriz de liberação de enzima comestível pode compreender um pélete, um comprimido ou uma pílula. O polipeptídeo da invenção na matriz de liberação de enzima comestível pode ser glicosilado. O polipeptídeo da invenção na matriz de liberação de enzima comestível pode ser termotolerante e/ou termoestável.

A invenção fornece um pélete comestível compreendendo um veículo comestível granulado e um polipeptídeo da invenção ou um polipeptídeo codificado por um ácido nucléico da invenção, em que o polipeptídeo compreende uma atividade de fitase.

A invenção fornece uma composição de alimentação compreendendo um gênero alimentício compreendendo uma proteína de fitase recombinante tendo pelo menos trinta aminoácidos contíguos de uma seqüência de aminoácidos como exposto na SEQ ID Nº: 2 ou uma variação conservadora desta, e um veículo comestível. A proteína de fitase no gênero alimentício pode ser glicosilada. A proteína de fitase no gênero alimentício pode ser

termotolerante e/ou termoestável. O gênero alimentício pode ser fabricado na forma de pélete, pílula ou comprimido. O gênero alimentício pode ser produzido usando aditivos revestidos com polímero. O gênero alimentício pode ser fabricado na forma granulada. O gênero alimentício pode ser produzido mediante secagem por pulverização.

A invenção fornece uma farinha de feijão-soja compreendendo um polipeptídeo da invenção ou um polipeptídeo codificado por um ácido nucléico da invenção, em que o polipeptídeo compreende uma atividade de fitase.

A invenção fornece um método de isolar ou identificar um polipeptídeo com atividade de fitase compreendendo as etapas de: (a) fornecer um anticorpo da invenção; (b) fornecer uma amostra compreendendo polipeptídeos; e (c) contatar a amostra da etapa (b) com o anticorpo da etapa (a) sob condições em que o anticorpo especificamente pode ligar ao polipeptídeo, assim isolando ou identificando uma fitase.

A invenção fornece um método de produzir um anticorpo de antifitase compreendendo administrar a um animal não-humano um ácido nucléico da invenção, ou um polipeptídeo da invenção, em uma quantidade suficiente para gerar uma resposta imune humoral, assim produzindo um anticorpo de antifitase.

A invenção fornece um método de produzir um polipeptídeo recombinante compreendendo as etapas de: (a) fornecer um ácido nucléico operavelmente ligado a um promotor; em que o ácido nucléico compreende uma seqüência da invenção; e (b) expressar o ácido nucléico da etapa (a) sob condições que permitem expressão do polipeptídeo, assim produzindo um polipeptídeo recombinante. O método pode também compreender transformar uma célula hospedeira com o ácido nucléico da etapa (a) seguido expressando o ácido nucléico da etapa (a), assim produzindo um polipeptídeo recombinante em uma célula transformada.

A invenção fornece um método para identificar um polipeptídeo tendo uma atividade de fitase compreendendo as etapas a seguir: (a) fornecer um polipeptídeo da invenção ou um polipeptídeo codificado por um ácido

nucléico tendo uma seqüência da invenção; (b) fornecer um substrato de fitase; e (c) contatar o polipeptídeo ou um fragmento ou variante deste da etapa (a) com o substrato da etapa (b) e detectar um aumento na quantidade de substrato ou uma diminuição na quantidade de produto de reação, em que uma diminuição na quantidade do substrato ou um aumento na quantidade do produto de reação detecta um polipeptídeo tendo uma atividade de fitase.

A invenção fornece um método para identificar um substrato de fitase compreendendo as etapas a seguir: (a) fornecer um polipeptídeo da invenção ou um polipeptídeo codificado por um ácido nucléico tendo uma seqüência da invenção; (b) fornecer um substrato de teste; e (c) contatar o polipeptídeo da etapa (a) com o substrato de teste da etapa (b) e detectar um aumento na quantidade de substrato ou uma diminuição na quantidade de produto de reação, em que uma diminuição na quantidade do substrato ou um aumento na quantidade do produto de reação identifica o substrato de teste como um substrato de fitase.

A invenção fornece um método de determinar se um composto especificamente liga a um polipeptídeo compreendendo as etapas a seguir: (a) expressar um ácido nucléico ou um vetor compreendendo o ácido nucléico sob condições rigorosas para tradução do ácido nucléico para um polipeptídeo, em que o ácido nucléico tem uma seqüência da invenção, ou, fornecer um polipeptídeo da invenção; (b) contatar o polipeptídeo com o composto de teste; e (c) determinar se o composto de teste especificamente liga ao polipeptídeo, assim determinando que o composto especificamente liga ao polipeptídeo.

A invenção fornece um método para identificar um modulador de uma atividade de fitase compreendendo as etapas a seguir: (a) fornecer um polipeptídeo de fitase da invenção ou um polipeptídeo de fitase codificado por um ácido nucléico da invenção; (b) fornecer um composto de teste; (c) contatar o polipeptídeo da etapa (a) com o composto de teste da etapa (b) e medir uma atividade da fitase, em que uma alteração na atividade de fitase medida na presença do composto de teste comparada à atividade na ausên-

cia do composto de teste fornece uma determinação que o composto de teste modula a atividade de fitase. Em um aspecto, a atividade de fitase é medida fornecendo um substrato de fitase e detectando um aumento na quantidade do substrato ou uma diminuição na quantidade de um produto de reação. Em um aspecto, uma diminuição na quantidade do substrato ou um aumento na quantidade do produto de reação com o composto de teste é comparada à quantidade de substrato ou produto de reação sem o composto de teste para identificar o composto de teste como um ativador de atividade de fitase. Em um aspecto, um aumento na quantidade do substrato ou uma diminuição na quantidade do produto de reação com o composto de teste é comparado à quantidade de substrato ou produto de reação sem o composto de teste para identificar o composto de teste como um inibidor de atividade de fitase.

A invenção fornece um sistema de computador compreendendo um processador e um dispositivo de armazenamento de dados em que o dito dispositivo de armazenamento de dados tem nele armazenada uma seqüência de polipeptídeos ou uma seqüência de ácidos nucléicos, em que a seqüência de polipeptídeos compreende uma seqüência de polipeptídeos da invenção, ou subseqüência desta, e o ácido nucléico compreende uma seqüência de ácidos nucléicos da invenção ou subseqüência desta. O sistema de computador pode também compreender um algoritmo de comparação de seqüência e um dispositivo de armazenamento de dados tendo pelo menos uma seqüência de referência nele armazenada. O algoritmo de comparação de seqüência pode compreender um programa de computador que indica polimorfismos. O sistema de computador pode também compreender um identificador que identifica uma ou mais características na dita seqüência.

A invenção fornece um meio legível por computador tendo nele armazenada uma seqüência de polipeptídeos ou uma seqüência de ácidos nucléicos, em que a seqüência de polipeptídeos compreende uma seqüência de polipeptídeos da invenção, ou subseqüência desta, e o ácido nucléico compreende uma seqüência de ácidos nucléicos da invenção ou subseqüência desta.

A invenção fornece um método para identificar uma característica em uma seqüência compreendendo as etapas de: (a) ler a seqüência usando um programa de computador que identifica uma ou mais características em uma seqüência, em que a seqüência compreende uma seqüência de polipeptídeos ou uma seqüência de ácidos nucléicos, em que a seqüência de polipeptídeos compreende a seqüência da invenção ou subseqüência desta, e o ácido nucléico compreende uma seqüência da invenção ou subseqüência desta, e (b) identificar uma ou mais características na seqüência com o programa de computador.

A invenção fornece um método para comparar uma primeira seqüência a uma segunda seqüência compreendendo as etapas de: (a) ler a primeira seqüência e a segunda seqüência através de um programa de computador que compara as seqüências, em que a primeira seqüência compreende uma seqüência de polipeptídeos ou uma seqüência de ácidos nucléicos, em que a seqüência de polipeptídeos compreende a seqüência da invenção, ou subseqüência desta, e o ácido nucléico compreende uma seqüência da invenção ou subseqüência desta, e (b) determinar a diferença entre a primeira seqüência e a segunda seqüência com o programa de computador. Em um aspecto, a etapa de determinar as diferenças entre a primeira seqüência e a segunda seqüência pode também compreender a etapa de identificar polimorfismos. O método pode também compreender um identificador que identifica uma ou mais características em uma seqüência. O método pode também compreender ler a primeira seqüência usando um programa de computador e identificar uma ou mais características na seqüência.

A invenção fornece um método para isolar ou restabelecer um ácido nucléico codificando um polipeptídeo com uma atividade de fitase de uma amostra ambiental compreendendo as etapas de: (a) fornecer um par de seqüências de iniciadores de amplificação para amplificar um ácido nucléico codificando um polipeptídeo com uma atividade de fitase, em que o par de iniciadores é capaz de amplificar SEQ ID Nº: 1, ou uma subseqüência desta; (b) isolar um ácido nucléico da amostra ambiental ou tratar a amostra

ambiental de modo que ácido nucléico na amostra seja acessível para hibridação com o par de iniciadores de amplificação; e, (c) combinar o ácido nucléico da etapa (b) com o par de iniciadores de amplificação da etapa (a) e amplificar o ácido nucléico da amostra ambiental, assim isolando ou restabelecendo um ácido nucléico codificando um polipeptídeo com uma atividade de fitase de uma amostra ambiental. Em um aspecto, cada membro das seqüências de iniciadores de amplificação pode compreender um oligonucleotídeo compreendendo pelo menos cerca de 10 a 50 bases sucessivas de uma seqüência como exposto na SEQ ID Nº: 1. A invenção fornece um método para isolar ou restabelecer um ácido nucléico codificando um polipeptídeo com uma atividade de fitase de uma amostra ambiental compreendendo as etapas de: (a) fornecer uma sonda de polinucleotídeo compreendendo uma seqüência da invenção, ou uma subseqüência desta; (b) isolar um ácido nucléico da amostra ambiental ou tratar a amostra ambiental de modo que ácido nucléico na amostra seja acessível para hibridação com uma sonda de polinucleotídeo da etapa (a); (c) combinar o ácido nucléico isolado ou a amostra ambiental tratada da etapa (b) com a sonda de polinucleotídeo da etapa (a); e (d) isolar um ácido nucléico que especificamente hibrida com a sonda de polinucleotídeo da etapa (a), assim isolando ou restabelecendo um ácido nucléico codificando um polipeptídeo com uma atividade de fitase de uma amostra do solo. A amostra ambiental pode compreender uma amostra de água, uma amostra líquida, uma amostra de solo, uma amostra de ar ou uma amostra biológica. A amostra biológica pode ser derivada de uma célula bacteriana, uma célula protozoária, uma célula de inseto, uma célula de levedura, uma célula vegetal, uma célula fúngica ou uma célula mamífera.

A invenção fornece um método de gerar uma variante de um ácido nucléico codificando uma fitase compreendendo as etapas de: (a) fornecer um ácido nucléico modelo compreendendo uma seqüência de ácidos nucléicos da invenção; e (b) modificar, excluir ou adicionar um ou mais nucleotídeos na seqüência modelo, ou uma combinação destes, para gerar uma variante do ácido nucléico modelo. O método pode também compreender expressar o ácido nucléico variante para gerar um polipeptídeo de fitase

variante. As modificações, adições ou deleções podem ser introduzidas por PCR propensa a erro, embaralhamento, mutagênese dirigida a oligonucleotídeo, PCR de montagem, mutagênese de PCR sexual, mutagênese *in vivo*, mutagênese de cassete, mutagênese de conjunto recursiva, mutagênese de conjunto exponencial, mutagênese loco-específica, remontagem genética, mutagênese saturada no sítio genético (GSSM), remontagem de ligação sintética (SLR) e/ou uma combinação destas. As modificações, adições ou deleções são introduzidas por recombinação, recombinação de seqüência recursiva, mutagênese de DNA modificado por fosfotioato, mutagênese de modelo contendo uracila, mutagênese duplex com intervalos, mutagênese de reparo de equívoco de ponto, mutagênese de cepa hospedeira deficiente de reparo, mutagênese química, mutagênese radiogênica, mutagênese de deleção, mutagênese de restrição-seleção, mutagênese de restrição-purificação, síntese de gene artificial, mutagênese de conjunto, criação multímera de ácido nucléico quimérico e/ou uma combinação destas. Em aspectos alternativos, as modificações, adições ou deleções são introduzidas por PCR propensa a erro, embaralhamento, através de mutagênese dirigida a oligonucleotídeo, através de PCR de montagem, através de mutagênese de PCR sexual, mutagênese *in vivo*, mutagênese de cassete, através de mutagênese de conjunto recursiva, através de mutagênese de conjunto exponencial, mutagênese loco-específica, através de remontagem genética, através de remontagem de ligação sintética (SLR) e/ou através de mutagênese saturada no sítio genético (GSSM).

Em um aspecto, método é iterativamente repetido até que uma fitase tendo uma atividade alterada ou diferente ou uma estabilidade alterada ou diferente daquela de uma fitase codificada pelo ácido nucléico modelo seja produzida. O polipeptídeo de fitase variante pode ser termotolerante, em que a fitase retém alguma atividade após ser exposta a uma temperatura elevada. O polipeptídeo de fitase variante pode ter glicosilação aumentada quando comparado à fitase codificada por um ácido nucléico modelo. O polipeptídeo de fitase variante pode ter uma atividade de fitase sob uma temperatura alta, em que a fitase codificada pelo ácido nucléico modelo não é ativa

sob a temperatura alta. Em um aspecto, método é iterativamente repetido até que uma seqüência codificadora de fitase tendo um uso de códon alterado daquela do ácido nucléico modelo seja produzida. Em um aspecto, o método é iterativamente repetido até que um gene de fitase tendo nível mais alto ou mais baixo de expressão de mensagem ou estabilidade daquele do ácido nucléico modelo seja produzido.

A invenção fornece um método para modificar os códons em um ácido nucléico codificando uma fitase para aumentar sua expressão em uma célula hospedeira, o método compreendendo (a) fornecer um ácido nucléico codificando uma fitase compreendendo uma seqüência de ácidos nucléicos da invenção; e, (b) identificar um códon não-preferido ou um menos preferido no ácido nucléico da etapa (a) e substitui-lo por um códon preferido ou de forma neutra usado que codifica o mesmo aminoácido que o códon substituído, em que um códon preferido é um códon sobre-representado nas seqüências de codificação nos genes na célula hospedeira e um códon não-preferido ou menos preferido é um códon sub-representado nas seqüências de codificação nos genes na célula hospedeira, assim modificando o ácido nucléico para aumentar sua expressão em uma célula hospedeira.

A invenção fornece um método para modificar os códons em um ácido nucléico codificando uma fitase, o método compreendendo (a) fornecer um ácido nucléico codificando uma fitase compreendendo uma seqüência de ácidos nucléicos da invenção; e, (b) identificar um códon no ácido nucléico da etapa (a) e substitui-lo por um códon diferente que codifica o mesmo aminoácido que o códon substituído, assim modificando códon em um ácido nucléico codificando uma fitase.

A invenção fornece um método para modificar os códons em um ácido nucléico codificando uma fitase para aumentar sua expressão em uma célula hospedeira, o método compreendendo (a) fornecer um ácido nucléico codificando uma fitase compreendendo uma seqüência de ácidos nucléicos da invenção; e, (b) identificar um códon não-preferido ou um menos preferido no ácido nucléico da etapa (a) e substitui-lo por um códon preferido ou de forma neutra usado que codifica o mesmo aminoácido que o códon substi-

tuído, em que um códon preferido é um códon sobre-representado nas seqüências de codificação nos genes na célula hospedeira e um códon não-preferido ou menos preferido é um códon sub-representado nas seqüências de codificação nos genes na célula hospedeira, assim modificando o ácido nucléico para aumentar sua expressão em uma célula hospedeira.

A invenção fornece um método para modificar um códon em um ácido nucléico codificando uma fitase para diminuir sua expressão em uma célula hospedeira, o método compreendendo (a) fornecer um ácido nucléico codificando uma fitase compreendendo uma seqüência da invenção; e (b) identificar pelo menos um códon preferido no ácido nucléico da etapa (a) e substitui-lo por um códon não-preferido ou menos preferido que codifica o mesmo aminoácido que o códon substituído, em que um códon preferido é um códon sobre-representado nas seqüências de codificação nos genes em uma célula hospedeira e um códon não-preferido ou menos preferido é um códon sub-representado nas seqüências de codificação nos genes na célula hospedeira, assim modificando o ácido nucléico para diminuir sua expressão em uma célula hospedeira. Em aspectos alternativos, a célula hospedeira é uma célula bacteriana, uma célula fúngica, uma célula de inseto, uma célula de levedura, uma célula vegetal ou uma célula mamífera, por exemplo, uma célula humana.

A invenção fornece um método para produzir uma biblioteca de ácidos nucléicos codificando uma pluralidade de sítios ativos de fitase modificada ou sítios de ligação ao substrato, em que os sítios ativos modificados ou sítios de ligação ao substrato são derivados de um primeiro ácido nucléico compreendendo uma seqüência codificando um primeiro sítio ativo ou um primeiro sítio de ligação ao substrato, o método compreendendo: (a) fornecer um primeiro ácido nucléico codificando um primeiro sítio ativo ou primeiro sítio de ligação ao substrato, em que a primeira seqüência de ácidos nucléicos compreende uma seqüência que hibrida sob condições rigorosas com uma seqüência como exposto na SEQ ID Nº: 1, e o ácido nucléico codifica um sítio ativo de fitase ou um sítio de ligação ao substrato de fitase; (b) fornecer um conjunto de oligonucleotídeos mutagênicos que codificam varian-

tes de aminoácido de ocorrência natural em uma pluralidade de códons alvejados no primeiro ácido nucléico; e, (c) usar o conjunto de oligonucleotídeos mutagênicos para gerar um conjunto de ácidos nucléicos de variantes de codificação de sítio ativo ou de sítio de ligação ao substrato que codificam uma faixa de variações de aminoácido em cada códon do aminoácido que foi submetido à mutagênese, assim produzindo uma biblioteca de ácidos nucléicos codificando uma pluralidade de sítios ativos de fitase modificada ou sítios de ligação ao substrato. Em um aspecto, o método compreende submeter à mutagênese o primeiro ácido nucléico da etapa (a) por um método compreendendo um sistema de evolução direcionada otimizada, ou, um método compreendendo mutagênese de saturação de sítio genético (GSSM), ou, um método compreendendo uma remontagem de ligação sintética (SLR). O método pode também compreender transformar o primeiro ácido nucléico da etapa (a) ou variantes por um método compreendendo PCR propensa a erro, embaralhamento, mutagênese dirigida a oligonucleotídeo, PCR de montagem, mutagênese de PCR sexual, mutagênese *in vivo*, mutagênese de cassete, mutagênese de conjunto recursiva, mutagênese de conjunto exponencial, mutagênese loco-específica, remontagem genética, mutagênese saturada no sítio genético (GSSM), remontagem de ligação sintética (SLR) e uma combinação destes. O método pode também compreender transformar o primeiro ácido nucléico da etapa (a) ou variantes por um método compreendendo recombinação, recombinação de seqüência recursiva, mutagênese de DNA modificado por fosfotioato, mutagênese de modelo contendo uracila, mutagênese duplex com intervalos, mutagênese de reparo de equívoco de ponto, mutagênese de cepa hospedeira deficiente de reparo, mutagênese química, mutagênese radiogênica, mutagênese de deleção, mutagênese de restrição-seleção, mutagênese de restrição-purificação, síntese de gene artificial, mutagênese de conjunto, criação multímera de ácido nucléico quimérico e uma combinação destes.

A invenção fornece um método de produzir uma molécula pequena compreendendo as etapas de: (a) fornecer uma pluralidade de enzimas biossintéticas capazes de sintetizar ou modificar uma molécula peque-

na, em que uma das enzimas compreende uma enzima de fitase codificada por um ácido nucléico compreendendo uma seqüência da invenção; (b) fornecer um substrato para pelo menos uma das enzimas da etapa (a); e (c) reagir o substrato da etapa (b) com as enzimas sob condições que facilitam uma pluralidade de reações biocatalíticas gerar uma molécula pequena por uma série de reações biocatalíticas.

A invenção fornece um método para modificar uma molécula pequena compreendendo as etapas: (a) fornecer uma enzima de fitase codificada por um ácido nucléico compreendendo uma seqüência da invenção; (b) fornecer uma molécula pequena; e (c) reagir a enzima da etapa (a) com a molécula pequena da etapa (b) sob condições que facilitam uma reação enzimática catalisada pela enzima de fitase, assim modificando uma molécula pequena por uma reação enzimática de fitase. O método pode compreender uma pluralidade de substratos de moléculas pequenas para a enzima da etapa (a), assim gerando uma biblioteca de moléculas pequenas modificadas produzidas por pelo menos uma reação enzimática catalisada pela enzima de fitase. O método pode compreender uma pluralidade de enzimas adicionais sob condições que facilitam uma pluralidade de reações biocatalíticas pelas enzimas formar uma biblioteca de moléculas pequenas modificadas produzidas pela pluralidade de reações enzimáticas. O método pode também compreender a etapa de testar a biblioteca para determinar se uma molécula pequena modificada particular que apresenta uma atividade desejada está presente dentro da biblioteca.

Em um aspecto, o método compreende a etapa de testar a biblioteca pelas etapas também compreendendo sistematicamente eliminar todas menos uma das reações biocatalíticas usadas para produzir uma porção da pluralidade das moléculas pequenas modificadas dentro da biblioteca testando a porção da molécula pequena modificada pela presença ou ausência da molécula pequena modificada particular com uma atividade desejada, e identificar pelo menos uma reação biocatalítica específica que produz a molécula pequena modificada particular de atividade desejada.

A invenção fornece um método para determinar um fragmento

funcional de uma enzima de fitase compreendendo as etapas de: (a) fornecer uma enzima de fitase, em que a enzima compreende uma seqüência de aminoácidos da invenção, ou, é codificada por um ácido nucléico tendo uma seqüência da invenção; e (b) excluir uma pluralidade de resíduos de aminoácido da seqüência da etapa (a) e testar a subseqüência restante pela uma atividade de fitase, assim determinando um fragmento funcional de uma enzima de fitase. A atividade de fitase pode ser medida fornecendo um substrato de fitase e detectando um aumento na quantidade do substrato ou uma diminuição na quantidade de um produto de reação. Em um aspecto, uma diminuição na quantidade de um substrato de enzima ou um aumento na quantidade do produto de reação com o composto de teste é comparada à quantidade de substrato ou produto de reação sem o composto de teste para identificar o composto de teste como um ativador de atividade de fitase.

A invenção fornece um método para engenharia de célula total de fenótipos novos ou modificados usando análise de fluxo metabólico em tempo real, o método compreendendo as seguintes etapas: (a) produzir uma célula modificada modificando a composição genética de uma célula, em que a composição genética é modificada por adição à célula de um ácido nucléico compreendendo uma seqüência da invenção; (b) cultivar a célula modificada para gerar uma pluralidade de células modificadas; (c) medir pelo menos um parâmetro metabólico da célula monitorando a cultura de célula da etapa (b) em tempo real; e, (d) analisar os dados da etapa (c) para determinar se o parâmetro medido difere de uma medição comparável em uma célula inalterada sob condições similares, assim identificando um fenótipo criado na célula usando análise de fluxo metabólico em tempo real. A composição genética da célula pode ser modificada por um método compreendendo deleção de uma seqüência ou modificação de uma seqüência na célula, ou, golpeando a expressão de um gene. O método pode também compreender selecionar uma célula compreendendo um fenótipo recentemente criado. O método pode também compreender cultivar a célula selecionada, assim gerando uma cepa de célula nova compreendendo um fenótipo recentemente criado.

A invenção fornece um método para hidrolisar um inositol-hexafosfato para inositol e fosfato inorgânico compreendendo as etapas a seguir: (a) fornecer um polipeptídeo tendo uma atividade de fitase, em que o polipeptídeo compreende uma seqüência de aminoácidos da invenção, ou, um polipeptídeo codificado por um ácido nucléico tendo uma seqüência da invenção; (b) fornecer uma composição compreendendo um inositol-hexafosfato; e (c) contatar o polipeptídeo da etapa (a) com a composição da etapa (b) sob condições em que o polipeptídeo hidrolisa no inositol-hexafosfato para produzir o inositol e fosfato inorgânico. As condições podem compreender uma temperatura de entre cerca de 37°C e cerca de 70°C. A composição pode compreender um ácido fítico.

A invenção fornece um método para desgrudadura de óleo compreendendo as etapas a seguir: (a) fornecer um polipeptídeo tendo uma atividade de fitase, em que o polipeptídeo compreende uma seqüência de aminoácidos da invenção, ou, um polipeptídeo codificado por um ácido nucléico tendo uma seqüência da invenção; (b) fornecer uma composição compreendendo um óleo vegetal; e (c) contatar o polipeptídeo da etapa (a) e o óleo vegetal da etapa (b) sob condições em que o polipeptídeo pode clivar uma ligação de inositol-fosfato inorgânico, assim desgrudando o óleo.

A invenção fornece um método para produzir uma alimentação de animais compreendendo as etapas a seguir: (a) transformando uma planta, parte de planta ou célula vegetal com um polinucleotídeo codificando um polipeptídeo de enzima de fitase, em que a fitase compreende pelo menos trinta aminoácidos contíguos de uma seqüência da invenção, ou um polipeptídeo codificado por um ácido nucléico tendo uma seqüência da invenção, ou um polipeptídeo tendo uma seqüência como exposto na SEQ ID Nº: 2; (b) cultivar a planta, parte de planta ou célula vegetal sob condições em que a enzima de fitase é expressada; e (c) converter a planta, partes da planta ou célula vegetal em uma composição adequada para alimentação para um animal, ou adicionar a planta, parte de planta ou célula vegetal cultivada a uma alimentação de animais, assim produzindo uma alimentação de animais. O polinucleotídeo pode compreender um vetor de expressão, em

que o vetor compreende uma seqüência de controle de expressão capaz de expressar o ácido nucléico em uma célula vegetal. O animal pode ser um animal monogástrico, por exemplo, um ruminante.

A invenção fornece um método para liberar um suplemento de enzima de fitase a um animal, o dito método compreendendo: (a) preparar uma matriz de liberação comestível compreendendo um veículo comestível e uma enzima de fitase, em que a matriz facilmente dispersa e libera a enzima de fitase quando colocada em meios aquosos, e (b) administrar a matriz de liberação de enzima comestível ao animal. A matriz de liberação comestível pode compreender um veículo comestível granulado. A matriz de liberação comestível pode ser na forma de péletes, pílulas, comprimidos, e outros. O veículo comestível pode compreender um veículo selecionado do grupo que consiste em germe de grão, feno, alfafa, capim-rabo-de-rato, casca de soja, farinha de semente de girassol e farinha de trigo. O veículo comestível pode compreender germe de grão sem óleo. A fitase pode compreender pelo menos trinta aminoácidos contíguos de uma seqüência da invenção, ou um polipeptídeo codificado por um ácido nucléico tendo uma seqüência da invenção, ou um polipeptídeo tendo uma seqüência como exposto na SEQ ID Nº: 2. A enzima de fitase pode ser glicosilada para fornecer termotolerância ou termoestabilidade em várias condições, por exemplo, a condições de peletização. A matriz de liberação pode ser formada peletizando uma mistura compreendendo um germe de grão e a enzima de fitase para render uma partícula. Os péletes podem ser feitos sob condições compreendendo aplicação de vapor. Os péletes podem ser feitos sob condições compreendendo aplicação de uma temperatura mais de 80°C durante cerca de 5 minutos. O pélete pode compreender uma enzima de fitase que compreende uma atividade específica pelo menos 350 a cerca de 900 unidades por miligrama de enzima.

A invenção fornece um método de aumentar a resistência de um polipeptídeo de fitase à inativação enzimática em um sistema digestivo de um animal, o método compreendendo glicosilar um polipeptídeo de fitase, em que a fitase compreende pelo menos trinta aminoácidos contíguos de

uma seqüência da invenção, ou um polipeptídeo codificado por um ácido nucléico tendo uma seqüência da invenção, ou um polipeptídeo tendo uma seqüência como exposto na SEQ ID Nº: 2, assim aumentando a resistência do polipeptídeo de fitase à inativação enzimática em um sistema digestivo de um animal. A glicosilação pode ser glicosilação N-ligada e/ou glicosilação O-ligada. O polipeptídeo de fitase pode ser glicosilada como um resultado de expressão *in vitro*, ou expressão *in vivo* de um polinucleotídeo codificando a fitase em uma célula. A célula pode ser uma célula eucariótica, como uma célula fúngica, uma célula vegetal, uma célula de inseto ou uma célula mamífera.

A invenção fornece um método de gerar ou aumentar a termotolerância ou termoestabilidade de um polipeptídeo de fitase, o método compreendendo glicosilar um polipeptídeo de fitase, em que a fitase compreende pelo menos trinta aminoácidos contíguos de uma seqüência da invenção, ou um polipeptídeo codificado por um ácido nucléico tendo uma seqüência da invenção, ou um polipeptídeo tendo uma seqüência como exposto na SEQ ID Nº: 2, assim aumentando a termotolerância ou termoestabilidade do polipeptídeo de fitase. A atividade específica de fitase pode ser termoestável ou em uma temperatura termotolerante na faixa de mais que cerca de 37°C a cerca de 90°C.

A invenção fornece um método para processamento de grãos de milho e de sorgo compreendendo as etapas a seguir: (a) fornecer um polipeptídeo tendo uma atividade de fitase, em que o polipeptídeo compreende uma seqüência de aminoácidos da invenção, ou, um polipeptídeo codificado por um ácido nucléico tendo uma seqüência da invenção; (b) fornecer uma composição compreendendo um líquido de maceração de milho ou um líquido de maceração de sorgo; e (c) contatar o polipeptídeo da etapa (a) e a composição da etapa (b) sob condições em que o polipeptídeo pode clivar uma ligação de inositol-fosfato inorgânico.

A invenção fornece um método para sobreexpressar uma fitase recombinante em uma célula compreendendo expressar um vetor compreendendo um ácido nucléico da invenção, por exemplo, um ácido nucléico

compreendendo uma seqüência de ácidos nucléicos pelo menos 98 % de identidade de seqüência à SEQ ID Nº: 1 em uma região de pelo menos cerca de 100 resíduos, em que as identidades das seqüências são determinadas por análise com um algoritmo de comparação de seqüência ou através de inspeção visual, ou, um ácido nucléico que hibrida sob condições rigorosas com uma seqüência de ácidos nucléicos como exposto na SEQ ID Nº: 1, ou uma subseqüência desta. A sobreexpressão podem ser realizada por quaisquer meios, por exemplo, uso de um promotor de atividade alta, um vetor distrônico ou por amplificação do gene do vetor.

A invenção fornece métodos para gerar uma fitase variante tendo uma atividade desejada obtendo um ácido nucléico compreendendo uma seqüência de polinucleotídeos selecionado de uma seqüência codificando uma enzima de fitase, uma seqüência substancialmente idêntica a esta, uma seqüência complementar a esta, e um fragmento compreendendo pelo menos 30 nucleotídeos sucessivos desta, e modificar um ou mais nucleotídeos na dita seqüência para outro nucleotídeo, excluir um ou mais nucleotídeos na dita seqüência, ou adicionar um ou mais nucleotídeos à dita seqüência. Por um tal método, um polinucleotídeo variante é obtido que codifica uma enzima de fitase modificada tendo uma atividade desejada, como termoestabilidade intensificada ou termotolerância intensificada.

Em ainda outro aspecto, a invenção fornece métodos para liberar um suplemento de fitase a um animal preparando uma matriz de liberação de enzima comestível na forma de péletes compreendendo um veículo comestível granulado e uma enzima de fitase recombinante termotolerante, em que as partículas facilmente dispersam na enzima de fitase nela contidas em meios aquosos, e administrando a matriz de liberação de enzima comestível ao animal.

Em ainda outro aspecto, a invenção fornece métodos para aumentar a resistência de um polipeptídeo de fitase à inativação enzimática em um sistema digestivo de um animal compreendendo glicosilar um polipeptídeo de fitase tendo uma seqüência de aminoácidos como exposto na SEQ ID Nº: 2, ou uma variação conservadora desta, assim aumentando a resis-

tência do polipeptídeo de fitase à inativação enzimática no sistema digestivo de um animal.

Em outro aspecto, a invenção fornece métodos para utilizar fitase como um suplemento nutricional nas dietas de animais preparando um suplemento nutricional contendo uma enzima de fitase recombinante compreendendo pelo menos trinta aminoácidos contíguos da SEQ ID Nº: 2, e administrando o suplemento nutricional a um animal para aumentar a utilização de fitato contido na comida ingerida pelo animal.

Em ainda outro aspecto, a invenção fornece métodos de aumentar a termotolerância de um polipeptídeo de fitase, o método compreendendo glicosilar um polipeptídeo de fitase, ou uma variação conservadora deste, para assim aumentar a termotolerância do polipeptídeo de fitase.

Em ainda outro aspecto, a invenção fornece um polipeptídeo de fitase termoestável, o dito polipeptídeo de fitase termotolerante sendo glicosilado para fornecer atividade específica aumentada do polipeptídeo de fitase após exposição a uma temperatura na faixa de mais que 37°C a cerca de 90°C comparado a um polipeptídeo de fitase substancialmente não-glicosilado correspondente.

Em ainda outro aspecto, a composição de alimentação da invenção compreende uma proteína de fitase recombinante tendo pelo menos trinta aminoácidos contíguos da seqüência de aminoácidos exposto na SEQ ID Nº: 2, ou uma variação conservadora desta, e um gênero alimentício contendo fitato.

Os detalhes de um ou mais aspectos da invenção estão expostos nos desenhos em anexo e na descrição abaixo. Outras características, objetivos e vantagens da invenção serão óbvios da descrição e desenhos, e das reivindicações.

Todas as publicações, patentes, pedidos de patente, seqüências do GenBank e depósitos da ATCC, aqui citados são por este expressamente incorporados por referência para todos os propósitos.

BREVE DESCRIÇÃO DOS DESENHOS

Os desenhos a seguir são ilustrativos dos aspectos da invenção

e não são intencionados a limitar o escopo da invenção conforme abrangido pelas reivindicações.

Figura 1A mostra a seqüência de nucleotídeos para uma fitase modificada (SEQ ID Nº: 1).

Figura 1B mostra a seqüência de aminoácidos para uma fitase modificada (SEQ ID Nº: 2).

Figuras 2A e 2B mostram o perfil de pH e temperatura e dados de estabilidade para a enzima de fitase da presente invenção (SEQ ID Nº: 2). OD a 700nm é indicada no eixo Y dos gráficos na Figura 2. Temperatura ou pH é indicado no eixo X dos gráficos.

Figura 3 mostra um gráfico com os resultados de um ensaio de tolerância térmica entre SEQ ID Nº: 4 (fitase do tipo selvagem) e SEQ ID Nº: 2 (fitase modificada). O gráfico mostra atividade de fitase residual em um fluido intestinal gástrico simulado (SGF) com pepsina. O percentual das atividades residuais (com base nas taxas iniciais) da fitase recombinante da invenção digerido *in vitro* (SEQ ID Nº: 2) expressada em vários hospedeiros de expressão foi representado nos gráficos versus tempo. A fitase foi expressada em E. coli (não-glicosilada), como também em S. pombe. e P. pastoris (glicosilados).

Figura 4 mostra um gráfico ilustrando o percentual da atividade residual da fitase K12SGF e da fitase da SEQ ID de NO: 2 (não-glicosilada) sob condições de digestibilidade simuladas usando pepsina como um fluido intestinal gástrico simulado.

Figura 5 é uma tabela mostrando dados de experimentos designados para determinar a meia vida relativa de fitase de E. coli, P. pastoris, e S. cerevisiae após exposição à pepsina como um fluido intestinal gástrico simulado.

Figura 6 apresenta a seqüência de aminoácidos de uma enzima de fitase com os sítios de glicosilação prognosticados em negrito.

Figura 7 apresenta em formato de tabela os resultados obtidos da análise em um Gel de Tris-Glicina a 12 % de proteína de fitase de P. pastoris e S. cerevisiae digerida com O-glicosidase e Endo H.

Figura 8A apresenta uma descrição esquemática das etapas para mapeamento de peptídeo máximo.

Figura 8B apresenta uma descrição esquemática das etapas para mapeamento de peptídeo de glicosilação.

Figura 9A apresenta a seqüência de aminoácidos da fitase da SEQ ID Nº: 2 expressada em P. pastoris com os resíduos glicosilados como determinado experimentalmente indicado em negrito na seqüência parcial da SEQ ID Nº: 4.

Figura 9B apresenta a seqüência de aminoácidos da fitase da SEQ ID Nº: 2 expressada em S. cerevisiae com os resíduos glicosilados como determinado experimentalmente indicado em negrito na seqüência parcial da SEQ ID Nº: 4.

Figura 10 apresenta um sumário dos resultados de glicosilação que mapeia para a fitase de Figuras 9A e 9B com uma seqüência parcial da SEQ ID Nº: 4.

Figura 11 mostra um gráfico com resultados de um ensaio de tolerância térmica para expressão de fitase modificada (SEQ ID Nº: 2) em várias células hospedeiras.

Figura 12 é um gráfico ilustrando atividade residual da fitase da SEQ ID Nº: 2 após exposição a ensaio *in vitro* de digestibilidade usando um fluido intestinal gástrico simulado (SGF) com pepsina. O percentual das atividades residuais (com base nas taxas iniciais) é mostrado para expressão em E. coli (não-glicosilada), como também P. pastoris e S. pombe (glicosilada).

Figura 13 mostra a seqüência de nucleotídeos codificando a fitase de appA de E. coli do tipo selvagem (SEQ ID Nº: 3).

Figura 14 mostra as seqüências de aminoácidos para o polipeptídeo de fitase de appA de E. coli do tipo selvagem (SEQ ID Nº: 4).

Figura 15 é um diagrama de bloco de um sistema de computador, como descrito em detalhes, abaixo.

Figura 16 é um fluxograma que ilustra um aspecto de um processo 200 para comparar uma nova seqüência de nucleotídeos ou de pro-

teínas com um banco de dados das seqüências para determinar os níveis de homologia entre a nova seqüência e as seqüências no banco de dados, como descrito em detalhes, abaixo.

Figura 17 é um fluxograma que ilustra uma modalidade de um processo em um computador para determinar se duas seqüências são homólogas, como descrito em detalhes, abaixo.

Figura 18 é um fluxograma que ilustra um aspecto de um processo identificador para detectar a presença de uma característica em uma seqüência, como descrito em detalhes, abaixo.

DESCRIÇÃO DETALHADA DA INVENÇÃO

A invenção refere-se aos polipeptídeos de fitase (por exemplo, SEQ ID Nº: 2) e polinucleotídeos (por exemplo, SEQ ID Nº: 1) codificando-os como também métodos de uso dos polinucleotídeos e polipeptídeos. A terminologia "fitase" abrange enzimas tendo alguma atividade de fitase, por exemplo, enzimas capazes de catalisar a degradação de fitato, por exemplo, a catálise de fitato (mio-inositol-hexafosfato) com inositol e fosfato inorgânico. As fitases da invenção incluem enzimas termotolerantes e termorresistentes.

As fitases e polinucleotídeos codificando as fitases da invenção são úteis em vários processos, métodos e composições. Por exemplo, como debatido acima, uma fitase pode ser usada na alimentação de animais, e suplementos alimentares como também em tratamentos para degradar ou remover fitato de excesso do ambiente ou uma amostra. Outros usos serão óbvios àqueles versados na técnica com base nos ensinos fornecidos aqui, incluindo aqueles debatidos acima.

Em um aspecto, as moléculas de fitase da invenção - ou sozinhas ou em combinação com outros reagentes (incluindo mas não limitado a enzimas, incluindo proteases) - são usadas no processamento de gêneros alimentícios, por exemplo, para prevenção do lama de milho não desejado, e em outras aplicações onde hidrólise de fitato é desejável. Em um aspecto, as moléculas de fitase da invenção são usadas para eliminar ou diminuir a presença de fitato não-hidrolisado, especialmente onde fitato não-hidrolisado

conduz às conseqüências problemáticas em processos *ex vivo* incluindo - mas não limitado ao - processamento de gêneros alimentícios. Em um aspecto, as moléculas de fitase da invenção são usadas nos procedimentos como descrito em EP0321004-B1 (Vaara et al.), etapas incluindo no processamento de grãos de milho e de sorgo por meio do qual os grãos duros são macerados em água para os amaciar. As substâncias solúveis em água que lixiviam durante este processo se tornam parte de um líquido de maceração de milho que é concentrado por evaporação. Ácido fítico não-hidrolisado no líquido de maceração do milho, em grande parte na forma de cálcio e sais de magnésio, é associado com fósforo e deposita uma lama indesejável com proteínas e íons de metal. Desse modo, a lama é problemática na evaporação, transporte e armazenamento do líquido de maceração do milho. As moléculas de fitase da invenção são usadas para hidrolisar esta lama.

As moléculas de fitase da invenção fornecem desempenho comercial substancialmente superior que moléculas de fitase previamente identificadas, por exemplo moléculas de fitase de origem fúngica.

A atividade de fitase das enzimas da invenção pode ser aproximadamente 4400 U/mg. Esta corresponde a cerca de uma melhoria de 40 vezes ou mais na atividade das enzimas microbianas previamente informadas foi aproximadamente na faixa de 50-100 U/mg de proteína.

A invenção também fornece métodos para alterar as características de uma fitase da invenção por mutagênese e outro método, incluindo evolução direcionada, por exemplo, métodos de propriedade da Diversa Corporation (por exemplo, DirectEvolution®). Estes métodos são também elaborados na Patente U. S. No. 5.830.696. Em resumo, DirectEvolution® compreende: a) a sujeição de um ou mais modelos moleculares, por exemplo, os ácidos nucléicos de fitase da invenção, mutagênese para gerar novas moléculas, e b) a seleção entre estes espécies da progênie de novas moléculas com características mais desejáveis.

O poder de evolução direcionada depende da escolha inicial dos modelos de partida (por exemplo, SEQ ID Nº: 1), como também do(s) processo(s) de mutagênese escolhido(s) e do(s) processos(s) de triagem usa-

do(s). Desse modo, a invenção fornece novas fontes altamente ativas, fisiologicamente eficazes e econômicas de atividade de fitase, incluindo novas fitases que: a) têm atividades superiores sob uma ou mais aplicações específicas, como fabricação em temperatura alta de gêneros alimentícios, e são desse modo útil para otimizar estas aplicações específicas; b) são úteis como modelos para evolução direcionada para alcançar até mesmo novas moléculas melhoradas; e c) são úteis como ferramentas para a identificação de moléculas relacionadas adicionais através de meios como métodos com base em hibridação.

DEFINIÇÕES

O termo "anticorpo" inclui um peptídeo ou polipeptídeo derivado de, modelado após ou substancialmente codificado por um gene de imunoglobulina ou genes de imunoglobulina, ou fragmentos destes, capazes de especificamente ligar um antígeno ou epítopo, ver, por exemplo 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 de antígeno, isto é, "sítios de ligação ao antígeno," (e. g., fragmentos, subseqüência, regiões de determinação de complementaridade (CDRs)) que retêm capacidade de ligar antígeno, incluindo (i) um fragmento de Fab, um fragmento monovalente que consiste nos domínios de VL, VH, CL e CH1; (ii) um fragmento de F(ab')2, um fragmento bivalente compreendendo dois fragmentos de Fab ligados por uma ligação em ponte de dissulfeto na região da dobradiça; (iii) um fragmento de Fd que consiste nos domínios de VH e de CH1; (iv) um fragmento de Fv que consiste nos domínios de VL e de VH de um braço simples de um anticorpo, (v) um fragmento de dAb (Ward et al., (1989) Nature 341:544-546) que consiste em um domínio de VH; e (vi) uma região de determinação de complementaridade isolada (CDR). Os anticorpos de cadeia simples também são incluídos por referência no termo "anticorpo."

Os termos "arranjo" ou "microarranjo" ou "biochip" ou "chip" como aqui usados são uma pluralidade de elementos alvo, cada elemento alvo compreendendo uma quantidade definida de um ou mais polipeptídeos

(incluindo anticorpos) ou ácidos nucléicos imobilizados sobre uma área definida de uma superfície de substrato, como debatido em também detalhe, abaixo.

Como aqui usado, os termos "computador," "programa de computador" e "processador" são usados em seus contextos gerais mais amplos e incorporam todos tais dispositivos, como descrito em detalhes, abaixo.

Uma "seqüência de codificação de" ou uma "seqüência codifica" um polipeptídeo ou proteína particular, é uma seqüência de ácidos nucléicos que é transcrita e traduzida em um polipeptídeo ou proteína quando colocada sob o controle de seqüências reguladoras apropriadas.

O termo "cassete de expressão" como aqui usado refere-se a uma seqüência de nucleotídeos que é capaz de afetar a expressão de um gene estrutural (isto é, uma seqüência de codificação de proteína, como uma fitase da invenção) em um hospedeiro compatível com tais seqüências. Os cassetes de expressão incluem pelo menos um promotor operavelmente ligado com a seqüência de codificação de polipeptídeo; e, opcionalmente, com outras seqüências, por exemplo, sinais de terminação de transcrição. Fatores adicionais necessários ou úteis para realizar a expressão também podem ser usados, por exemplo, intensificadores. "Operavelmente ligado" como aqui usado adiante refere-se a ligação de um promotor de uma seqüência de DNA de modo que o promotor medeia a transcrição da seqüência de DNA. Desse modo, os cassetes de expressão também incluem plasmídeos, vetores de expressão, vírus recombinantes, qualquer forma de vetor de DNA recombinante "nu" e outros. Um "vetor" compreende um ácido nucléico que pode infetar, transfeccionar, transiente ou permanentemente transduzir uma célula. Será reconhecido que um vetor pode ser um ácido nucléico nu, ou um ácido nucléico complexado com proteína ou lipídio. O vetor opcionalmente compreende ácidos nucléicos viróticos ou bacterianos e/ou proteínas, e/ou membranas (por exemplo, uma membrana de célula, um envelope de lipídio virótico, etc.). Vetores incluem, mas não são limitados a réplicons (por exemplo, réplicons de RNA, bacteriófagos) aos quais fragmentos de DNA podem ser ligados e se tornarem replicados. Vetores desse

modo incluem, mas não são limitados a RNA, DNA ou RNA circular ou linear de auto-replicação autônomo (por exemplo, plasmídeos, vírus, e outros, ver, por exemplo, Patente U. S. No. 5.217.879), e incluem o plasmídeo de expressão e de não-expressão. Onde um microorganismo recombinante ou cultura de célula é descrito como hospedeiro de um "vetor de expressão" este inclui tanto DNA circular e linear extracromossômico quanto DNA que foi incorporado no(s) cromossomo(s) do hospedeiro. Onde um vetor estiver sendo mantido por uma célula hospedeira, o vetor ou pode ser estavelmente replicado pelas células durante mitose como uma estrutura autônoma, ou é incorporado dentro do genoma do hospedeiro.

As frases "ácido nucléico" ou "seqüência de ácidos nucléicos" como aqui usadas refere-se a um oligonucleotídeo, nucleotídeo, polinucleotídeo, ou a um fragmento de qualquer um destes, para DNA ou RNA (por exemplo, mRNA, rRNA, tRNA) de origem genômica ou sintética que pode ser unifilamentado ou bifilamentado e pode representar um filamento sentido ou anti-sentido, para o ácido nucléico de peptídeo (PNA), ou para qualquer material semelhante a DNA ou semelhante a RNA, natural ou sintético na origem, incluindo, por exemplo, iRNA, ribonucleoproteínas (por exemplo, iRNPs). O termo abrange ácidos nucléicos, isto é, oligonucleotídeos, contendo análogos conhecidos de nucleotídeos naturais. O termo também abrange estruturas semelhantes a ácido nucléico com cadeias principais sintéticas, ver por exemplo, Mata (1997) Toxicol. Appl. Pharmacol. 144:189-197; Strauss-Soukup (1997) Biochemistry 36:8692-8698; Samstag (1996) Antisentido Nucleic Acid Drog Dev 6:153-156.

"Aminoácido" ou "seqüência de aminoácidos" como aqui usado refere-se a uma seqüência de oligopeptídeos, peptídeos, polipeptídeos ou de proteínas, ou a um fragmento, porção ou subunidade de qualquer um destes, e às moléculas de ocorrência natural ou sintética.

Como aqui usado, o termo "isolado" significa que o material é removido de seu ambiente original (por exemplo, o ambiente natural se for de ocorrência natural). Por exemplo, um polinucleotídeo de ocorrência natural ou polipeptídeo presente em um animal vivo não é isolado, mas o mesmo

polinucleotídeo ou polipeptídeo, separado de alguns ou todos os materiais coexistentes no sistema natural, é isolado. Tais polinucleotídeos podem fazer parte de um vetor e/ou tais polinucleotídeos ou polipeptídeos podem fazer parte de uma composição, e ainda ser isolados naquele tal vetor ou composição que não faz parte de seu ambiente natural. Como aqui usado, um material ou composição isolada também podem ser uma composição "purificada", isto é, não requer pureza absoluta; de preferência, é intencionado como uma definição relativa. Os ácidos nucléicos individuais obtidos de uma biblioteca podem ser convencionalmente purificados para homogeneidade eletroforética. Em aspectos alternativos, a invenção fornece ácidos nucléicos que foram purificados de DNA genômico ou de outras seqüências em uma biblioteca ou outro ambiente por pelo menos um, dois, três, quatro, cinco ou mais ordens de magnitude.

Como aqui usado, o termo "recombinantes" significa que o ácido nucléico é adjacente a um ácido nucléico da "cadeia principal" ao qual não é adjacente em seu ambiente natural. Em um aspecto, os ácidos nucléicos representam 5 % ou mais do número de inserções de ácido nucléico em uma população de "moléculas da cadeia principal" de ácido nucléico. "Moléculas da cadeia principal" de acordo com a invenção incluem ácidos nucléicos como vetores de expressão, ácidos nucléicos de auto-replicação, vírus, ácidos nucléicos de integração, e outros vetores ou ácidos nucléicos mantidos ou manipulados de uma inserção de ácido nucléico de interesse. Em um aspecto, os ácidos nucléicos enriquecidos representam 15 %, 20 %, 30 %, 40 %, 50 %, 60 %, 70 %, 80 %, 90 % ou mais do número de inserções de ácido nucléico na população de moléculas recombinantes da cadeia principal. Polipeptídeos ou proteínas "recombinantes" referem-se aos polipeptídeos ou proteínas produzidas por técnicas de DNA recombinante; por exemplo, produzidas de células transformadas por um constructo de DNA exógeno codificando o polipeptídeo ou proteína desejada. Polipeptídeos ou proteínas "sintéticas" são aquelas preparadas por síntese química, como descrito em também detalhe, abaixo.

Uma seqüência promotora é "operavelmente ligada a" uma se-

qüência de codificação quando polimerase de RNA que inicia a transcrição no promotor transcrever a seqüência de codificação no mRNA, como debatido também, abaixo.

"Oligonucleotídeo" refere-se ou a um polideoxinucleotídeo unifilamentado ou polideoxinucleotídeo bifilamentado complementar que pode ser quimicamente sintetizado. Tais oligonucleotídeos sintéticos não têm nenhum 5' fosfato e desse modo não se ligam a outro oligonucleotídeo sem adicionar um fosfato com um ATP na presença de uma cinase. Um oligonucleotídeo sintético ligará a um fragmento que não foi desfosforilado.

A frase "substancialmente idênticas" no contexto de dois ácidos nucléicos ou polipeptídeos, refere-se a duas ou mais seqüências tendo pelo menos 50 %, 60 %, 70 %, 75 %, 80 %, 85 %, 90 %, 95 %, 98 % ou 99 % de identidade (seqüência) de resíduo de nucleotídeo ou de aminoácido, quando comparadas e alinhadas para correspondência máxima, como medido usando qualquer algoritmo de comparação de seqüência conhecido, como debatido em detalhes abaixo, ou através de inspeção visual. Em aspectos alternativos, a invenção fornece seqüências de ácidos nucléicos e polipeptídeos tendo identidade substancial a uma seqüência exemplar da invenção, por exemplo, SEQ ID Nº: 1, SEQ ID Nº: 2, em uma região de pelo menos cerca de 100 resíduos, 150 resíduos, 200 resíduos, 300 resíduos, 400 resíduos, ou uma região que varia de entre cerca de 50 resíduos para o comprimento total do ácido nucléico ou polipeptídeos. As seqüências de ácidos nucléicos da invenção podem ser substancialmente idênticas no comprimento inteiro de uma região de codificação de polipeptídeos.

Adicionalmente uma seqüência de aminoácidos "substancialmente idêntica" é uma seqüência que difere de uma seqüência de referência por uma ou mais substituições, deleção ou inserções de aminoácidos mais conservadores ou não-conservadores, particularmente quando uma tal substituição ocorre em um sítio que não é o sítio ativo da molécula, e contanto que o polipeptídeo essencialmente retenha suas propriedades funcionais. Por exemplo, uma substituição de aminoácido conservadora substitui um aminoácido por outro da mesma classe (por exemplo, substituição de um

aminoácido hidrofóbico, como isoleucina, valina, leucina ou metionina, por outro, ou substituição de um aminoácido polar por outro, como substituição de arginina por lisina, ácido glutâmico por ácido aspártico ou glutamina por asparagina). Um ou mais aminoácidos podem ser excluídos, por exemplo, de um polipeptídeo de fitase, resultando na modificação da estrutura do polipeptídeo, sem significativamente alterar sua atividade biológica. Por exemplo, aminoácidos amino- ou carboxil-terminais que não são requeridos para atividade biológica de fitase podem ser removidos. As seqüências de polipeptídeos modificadas da invenção podem ser ensaiadas para atividade biológica de fitase por qualquer número de métodos, incluindo contando a seqüência de polipeptídeos modificada com um substrato de fitase e determinando se o polipeptídeo modificado diminui a quantidade de substrato específico no ensaio ou aumenta os bioprodutos da reação enzimática de uma fitase funcional com o substrato, como debatido também, abaixo.

"Hibridação" refere-se ao processo pelo que um filamento de ácido nucléico se une a um filamento complementar através de emparelhamento de bases. As reações de hibridação podem ser sensíveis e seletivas de forma que uma seqüência particular de interesse pode ser identificada até mesmo em amostras nas quais ela está presente em baixas concentrações. Condições adequadamente rigorosas podem ser definidas, por exemplo, pelas concentrações de sal ou formamida nas soluções de pré-hibridação e hibridação, ou pela temperatura de hibridação, e são bem-conhecidas na técnica. Por exemplo, a severidade pode ser aumentada reduzindo a concentração de sal, aumentando a concentração de formamida ou elevando a temperatura de hibridação, alterando o tempo de hibridação, como descrito em detalhes, abaixo. Em aspectos alternativos, os ácidos nucléicos da invenção são definidos por sua capacidade de hibridar sob várias condições de severidade (por exemplo, alta, média e baixa), como exposto aqui.

O termo "variante" refere-se aos polinucleotídeos ou polipeptídeos da invenção modificados em um ou mais pares básicos, códons, íntrons, éxons ou resíduos de aminoácido (respectivamente) embora ainda retenha a atividade biológica de uma fitase da invenção. Variantes podem

ser produzidas por qualquer número de meios incluindo métodos como, por exemplo, PCR propensa a erro, embaralhamento, mutagênese dirigida a oligonucleotídeo, PCR de montagem, mutagênese de PCR sexual, mutagênese *in vivo*, mutagênese de cassete, mutagênese de conjunto recursiva, mutagênese de conjunto exponencial, mutagênese loco-específica, remontagem genética, GSSM e qualquer combinação destes. Técnicas para produzir fitases variantes tendo atividade em um pH ou temperatura, por exemplo, que é diferente de uma fitase do tipo selvagem, são aqui incluídas.

O termo "mutagênese de saturação" ou "GSSM" inclui um método que usa iniciadores de oligonucleotídeos degenerados para introduzir mutações de ponto dentro de um polinucleotídeo, como descrito em detalhes, abaixo.

O termo "sistema de evolução direcionada otimizada" ou "evolução direcionada otimizada" inclui um método para remontar fragmentos de seqüências de ácidos nucléicos relacionadas, por exemplo, genes relacionados, e explicado em detalhes, abaixo.

O termo "remontagem de ligação sintética" ou "SLR" inclui um método de ligar fragmentos de oligonucleotídeo em uma forma não-estotástica, e explicou em detalhes, abaixo.

As frases "ácido nucléico" ou "seqüência de ácidos nucléicos" como aqui usadas referem-se a um oligonucleotídeo, nucleotídeo, polinucleotídeo ou a um fragmento de qualquer um destes, para DNA ou RNA de origem genômica ou sintética que pode ser unifilamentado ou bifilamentado e pode representar um filamento sentido ou anti-sentido, ácido nucléico de peptídeo (PNA) ou qualquer material semelhante a DNA ou semelhante a RNA, natural ou sintético na origem. Em um aspecto, uma "seqüência de ácidos nucléicos" da invenção inclui, por exemplo, uma seqüência codificando um polipeptídeo como exposto na SEQ ID Nº: 2, e variantes desta. Em outro aspecto, uma "seqüência de ácidos nucléicos" da invenção inclui, por exemplo, uma seqüência como exposto na SEQ ID Nº: 1, seqüências complementares a esta, fragmentos das seqüências anteriores e variantes destas.

Uma "seqüência de codificação" ou uma "seqüência de nucleotídeos codificando" um polipeptídeo ou proteína particular, é uma seqüência de ácidos nucléicos que é transcrita e traduzida em um polipeptídeo ou proteína quando colocada sob o controle das seqüências reguladoras apropriadas.

O termo "gene" significa o segmento de DNA envolvido na produção de uma cadeia de polipeptídeos; inclui regiões precedentes e seguindo a região de codificação (líder e de rastro) como também, onde aplicável, seqüências intervenientes (íntrons) entre os segmentos de codificação individuais (éxons).

"Aminoácido" ou "seqüência de aminoácidos" como aqui usado refere-se a uma seqüência de oligopeptídeos, peptídeos, polipeptídeos ou de proteínas, ou um fragmento, porção ou subunidade de qualquer um destes, e às moléculas de ocorrência natural ou sintética. Em um aspecto, uma "seqüência de aminoácidos" ou "seqüência de polipeptídeos" da invenção inclui, por exemplo, uma seqüência como exposto na SEQ ID Nº: 2, fragmentos da seqüência anterior e variantes desta. Em outro aspecto, uma "seqüência de aminoácidos" da invenção inclui, por exemplo, uma seqüência codificada por um polinucleotídeo tendo uma seqüência como exposto na SEQ ID Nº: 1, seqüências complementares a esta, fragmentos das seqüências anteriores e variantes destas.

O termo "polipeptídeo" como aqui usado, refere-se aos aminoácidos ligados um ao outro através de ligações de peptídeo ou ligações de peptídeo modificado, isto é, isósteres de peptídeo, e podem conter aminoácidos modificados diferentes dos 20 aminoácidos gene-codificados. Os polipeptídeos podem ser modificados por quaisquer processos naturais, como processamento pós-translacional, ou por técnicas de modificação química que são bem-conhecidas na técnica. Modificações podem ocorrer em qualquer lugar no polipeptídeo, incluindo na cadeia principal do peptídeo, nas cadeias laterais de aminoácido e no término amino ou carboxila. Será apreciado que o mesmo tipo de modificação pode estar presente nos mesmos ou variados graus nos vários sítios em um polipeptídeo dado. Também um poli-

peptídeo dado pode ter muitos tipos de modificações. Modificações incluem acetilação, acilação, PAD-ribosilação, amidação, ligação covalente de flavina, ligação covalente de uma hemi-metade, ligação covalente de um nucleotídeo ou derivado de nucleotídeo, ligação covalente de um lipídio ou derivado de lipídio, ligação covalente de um fosfatidilinositol, ciclização de reticulação, formação de ligação de dissulfeto, desmetilação, formação de reticulações covalentes, formação de cisteína, formação de piroglutamato, formilação, faixa-carboxilação, glicosilação, formação de âncora de GPI, hidroxilação, iodação, metilação, miristoliação, oxidação, tratamento com PEG, processamento proteolítico, fosforilação, prenilação, racemização, selenoilação, sulfatização e adição mediada por tradução de RNA de aminoácidos para proteína como arginilação. (Ver Proteins - Structure and Molecular Properties 2ª Ed., T. E. Creighton, W. H. Freeman and Company, Nova Iorque (1993); Posttranslational Colvante Modification of Proteins, B. C. Johnson, Ed., Academic Press, Nova Iorque, pp. 1-12 (1983)).

Como aqui usado, o termo "isolado" significa que o material é removido de seu ambiente original (por exemplo, o ambiente natural se for de ocorrência natural). Por exemplo, um polinucleotídeo de ocorrência natural ou polipeptídeo presente em um animal vivo não é isolado, mas o mesmo polinucleotídeo ou polipeptídeo, separado de alguns ou todos os materiais coexistentes no sistema natural, é isolado. Tais polinucleotídeos podem fazer parte de um vetor e/ou tais polinucleotídeos ou polipeptídeos podem fazer parte de uma composição, e ainda ser isolados naquele tal vetor ou composição que não faz parte de seu ambiente natural.

Como aqui usado, o termo "purificado" não requer pureza absoluta; de preferência, é intencionado como uma definição relativa. Os ácidos nucléicos individuais obtidos de uma biblioteca foram convencionalmente purificados à homogeneidade eletroforética. As seqüências destes clones não puderam ser obtidas diretamente ou da biblioteca ou do DNA humano total. Os ácidos nucléicos purificados da invenção foram purificados do resto do DNA genômico no organismo por pelo menos 10^4-10^6 vezes. Porém, o termo "purificado" também inclui ácidos nucléicos que foram purificados do

resto do DNA genômico ou de outras seqüências em uma biblioteca ou outro ambiente por pelo menos uma ordem de magnitude, tipicamente duas ou três ordens, e mais tipicamente quatro ou cinco ordens de magnitude.

Como aqui usado, o termo "recombinante" significa que o ácido nucléico é adjacente ao ácido nucléico da "cadeia principal" ao qual não se encontra adjacente em seu ambiente natural. Adicionalmente, ser "enriquecidos" os ácidos nucléicos representarão 5 % ou mais do número de inserções de ácido nucléico em uma população de moléculas da cadeia principal de ácido nucléico. Moléculas da cadeia principal de acordo com a invenção incluem ácidos nucléicos como vetores de expressão, ácidos nucléicos de auto-reprodução, vírus, ácidos nucléicos integrantes, e outros vetores ou ácidos nucléicos usados para manter ou manipular uma inserção de ácido nucléico de interesse. Tipicamente, os ácidos nucléicos enriquecidos representam 15 % ou mais do número de inserções de ácido nucléico na população de moléculas recombinantes da cadeia principal. Mais tipicamente, os ácidos nucléicos enriquecidos representam 50 % ou mais do número de inserções de ácido nucléico na população de moléculas recombinantes da cadeia principal. Em um aspecto, os ácidos nucléicos enriquecidos representam 90 % ou mais do número de inserções de ácido nucléico na população de moléculas recombinantes da cadeia principal.

Polipeptídeos ou proteínas "recombinantes" referem-se aos polipeptídeos ou proteínas produzidas por técnicas de DNA recombinantes; isto é, produzidas de células transformadas por um constructo de DNA exógeno codificando o polipeptídeo ou proteína desejada. Polipeptídeos ou proteína "sintéticos" são aqueles preparados por síntese química. Métodos de síntese química de peptídeo de fase sólida podem ser também usados para sintetizar o polipeptídeo ou fragmentos da invenção. Tal método foi conhecido na técnica desde os anos 60 (Merrifield, R. B., J. Am. Chem. Soc., 85:2149-2154, 1963) (também ver Stewart, J. M. e Young, J. D., Solid Phase Peptide Synthesis, 2ª ed., Pierce Chemical Co., Rockford, Ill., pp. 11-12)) e foi recentemente empregado em kits de modelo e síntese de peptídeo de laboratório comercialmente disponíveis (Cambridge Research Biochemicals). Tais

kits de laboratório comercialmente disponíveis em geral utilizaram os ensinamentos de H. M. Geysen et al, Proc. Natl. Acad. Sci., USA, 81:3998 (1984) e forneceram sintetização de peptídeos nas pontas de uma multidão de "bastões" ou "pinos" todos destes estão conectados a uma única placa. Quando um tal sistema é utilizado, uma placa de bastões ou pinos é invertida e inserida em uma segunda placa de cavidades ou reservatórios correspondentes, que contêm soluções para ligar ou ancorar um aminoácido apropriado às pontas do pino ou bastão. Repetindo uma tal etapa de processo, isto é, invertendo e inserindo as pontas do bastão e do pino nas soluções apropriadas, aminoácidos são construídos em peptídeos desejados. Além disso, vários sistemas de síntese de peptídeo de FMOC disponíveis estão disponíveis. Por exemplo, uma montagem de um polipeptídeo ou fragmento pode ser realizada em um suporte sólido usando um sintetizador de peptídeo automatizado da Applied Biosystems, Inc., Modelo 431A. Tal equipamento fornece acesso fácil aos peptídeos da invenção, ou através da síntese direta ou pela síntese de uma série de fragmentos que podem ser acoplados usando outras técnicas conhecidas.

Uma seqüência promotora está "operavelmente ligada a" uma seqüência de codificação quando polimerase de RNA que inicia a transcrição no promotor transcrever a seqüência de codificação dentro do mRNA.

"Plasmídeos" são designados por uma letra minúscula p precedida e/ou seguida de letras maiúsculas e/ou números. Os plasmídeos de partida ou são aqui comercialmente disponíveis, publicamente disponíveis em uma base não-restrita, ou podem ser construídos dos plasmídeos disponíveis de acordo com os procedimentos publicados. Além disso, os plasmídeos equivalentes para aqueles descritos aqui são conhecidos na técnica e serão óbvios ao versado na técnica.

"Digestão" de DNA refere-se à clivagem catalítica do DNA com uma enzima de restrição que apenas age em certas seqüências no DNA. As várias enzimas de restrição usadas aqui estão comercialmente disponíveis e suas condições de reação, co-fatores e outros requerimentos foram usados como é conhecido ao versado na técnica. Tipicamente para propósitos ana-

líticos, 1 μg de plasmídeo ou fragmento de DNA é usado com cerca de 2 unidades de enzima em cerca de 20 μl de solução tampão. Tipicamente para o propósito de isolar os fragmentos de DNA para construção do plasmídeo, são digeridos 5 a 50 μg de DNA com 20 a 250 unidades de enzima em um volume maior. Tampões apropriados e quantidades de substrato para as enzimas de restrição particulares são especificados pelo fabricante. O tempo de incubação de cerca de 1 hora a 37°C é normalmente usado, mas pode variar de acordo com as instruções do vendedor. Após digestão, a eletroforese em gel pode ser executada para isolar o fragmento desejado.

"Oligonucleotídeo" refere-se ou a um polideoxinucleotídeo unifilamentado ou polideoxinucleotídeo bifilamentado complementar que pode ser quimicamente sintetizado. Tais oligonucleotídeos sintéticos não têm nenhum 5' fosfato e desse modo não ligarão a outro oligonucleotídeo sem adicionar um fosfato com um ATP na presença de uma cinase. Um oligonucleotídeo sintético ligará a um fragmento que não foi desfosforilado.

A frase "substancialmente idêntica" no contexto de duas seqüências de ácidos nucléicos ou polipeptídeos, refere-se a duas ou mais seqüências tendo pelo menos 60 %, 70 %, 80 %, e em alguns aspectos 90-95 % de identidade de nucleotídeo ou resíduo de aminoácido, quando comparadas e alinhadas para correspondência máxima, como medido usando um dos algoritmos de comparação de seqüência conhecidos ou através de inspeção visual. Tipicamente, a identidade substancial existe em uma região de pelo menos cerca de 100 resíduos, e mais comumente as seqüências são substancialmente idênticas pelo menos cerca de 150-200 resíduos. Em alguns aspectos, as seqüências são substancialmente idênticas no comprimento inteiro das regiões de codificação.

O termo "cerca de" é aqui usado para significar "aproximadamente," ou "em torno de," ou "por volta de," ou "na região de." Quando o termo "cerca de" é usado junto com uma faixa numérica, ele modifica aquela faixa estendendo os limites para acima e para baixo dos valores numéricos expostos. Em geral, o termo "cerca de" é aqui usado para modificar um valor numérico para cima e para baixo do valor declarado por uma variância de 20

por cento.

Adicionalmente, uma seqüência de aminoácidos "substancialmente idêntica" é uma seqüência que difere de uma seqüência de referência por uma ou mais substituições, deleções ou inserções de aminoácido conservadoras ou não-conservadoras, particularmente quando uma tal substituição ocorre em um sítio que não é o sítio ativo da molécula, e contanto que o polipeptídeo essencialmente retenha suas propriedades funcionais. Por exemplo, uma substituição de aminoácido conservadora substitui um aminoácido por outro da mesma classe (por exemplo, substituição de um aminoácido hidrofóbico, como isoleucina, valina, leucina, ou metionina, por outro, ou substituição de um aminoácido polar por outro, como substituição de arginina por lisina, ácido glutâmico por ácido aspártico ou glutamina por asparagina). Um ou mais aminoácidos podem ser excluídos, por exemplo, de um polipeptídeo de fitase, resultando na modificação da estrutura do polipeptídeo, sem significativamente alterar sua atividade biológica. Por exemplo, aminoácidos amino- ou carboxil-terminais podem ser removidos que não são requeridos para atividade biológica de fitase. Seqüências de polipeptídeos modificadas da invenção podem ser ensaiadas para atividade biológica de fitase por qualquer número de métodos, incluindo contatar a seqüência de polipeptídeos modificada com um substrato de fitase e determinar se o polipeptídeo modificado diminui a quantidade de substrato específico no ensaio ou aumenta os bioprodutos da reação enzimática de um polipeptídeo de fitase funcional com o substrato.

"Fragmentos" como aqui usado são uma porção de uma proteína de ocorrência natural ou recombinante onde podem existir pelo menos duas conformações diferentes. Fragmentos podem ter a mesma ou substancialmente a mesma seqüência de aminoácidos como a proteína de ocorrência natural. "Substancialmente a mesma" significa que uma seqüência de aminoácidos é em grande parte, mas não completamente, a mesma, mas retém pelo menos uma atividade funcional da seqüência à qual está relacionada. Em geral, duas seqüências de aminoácidos são "substancialmente as mesmas" ou "substancialmente homólogas" se elas são pelo menos cerca

de 70, mas mais tipicamente cerca de 85 % ou mais idênticas. Também são incluídos fragmentos tendo estruturas tridimensionais diferentes como a proteína de ocorrência natural. Um exemplo deste, é uma molécula de "pró-forma", como uma pró-proteína de baixa atividade que pode ser modificada através de clivagem para produzir uma enzima matura com atividade significativamente mais alta.

"Hibridação" refere-se ao processo pelo qual um filamento de ácido nucléico une a um filamento complementar através de emparelhamento de bases. Reações de hibridação podem ser sensíveis e seletivas de forma que uma seqüência particular de interesse pode ser identificada até mesmo nas amostras em que ela está presente em baixas concentrações. Condições adequadamente rigorosas podem ser definidas, por exemplo, pelas concentrações de sal ou formamida nas soluções de pré-hibridação e de hibridação, ou pela temperatura de hibridação, e são bem-conhecidas na técnica. Em particular, a severidade pode ser aumentada reduzindo a concentração de sal, aumentando a concentração de formamida ou elevando a temperatura de hibridação.

Por exemplo, hibridação sob condições de severidade alta pode ocorrer em cerca de 50 % de formamida a cerca de 37°C a 42°C. Hibridação pode ocorrer sob condições de severidade reduzidas em cerca de 35 % a 25 % de formamida a cerca de 30°C a 35°C. Em particular, a hibridação pode ocorrer sob condições de severidade alta a 42°C em 50 % de formamida, 5X SSPE, 0,3 % de SDS e 200 ng/ml de DNA de esperma de salmão cisalhado e desnaturado. A hibridação pode ocorrer sob condições de severidade reduzidas como descrito acima, mas em 35 % de formamida em uma temperatura reduzida de 35°C. A faixa de temperatura que corresponde a um nível particular de severidade pode ser também estreitada calculando a razão de purina para pirimidina do ácido nucléico de interesse e ajustando a temperatura conseqüentemente. Variações nas faixas e condições acima são bem-conhecidas na técnica.

O termo "variante" refere-se aos polinucleotídeos ou polipeptídeos da invenção modificados em um ou mais pares de base, códons, ín-

trons, éxons ou resíduos de aminoácido (respectivamente) embora ainda retenha a atividade biológica de uma fitase da invenção. Variantes podem ser produzidas por qualquer número de meios incluindo métodos como, por exemplo, PCR propensa a erro, embaralhamento, mutagênese dirigida a oligonucleotídeo, PCR de montagem, mutagênese de PCR sexual, mutagênese *in vivo*, mutagênese de cassete, mutagênese de conjunto recursiva, mutagênese de conjunto exponencial, mutagênese sítio específica, remontagem de ligação, GSSM e qualquer combinação destes.

Os termos "termoestável" e "termoestabilidade" como aqui usados com referência a uma enzima significam a capacidade da enzima de operar em temperaturas aumentadas, por exemplo ter atividade específica comparável a 70°C e a 85°C em um pH comum. Uma enzima "termoestável" manterá muita ou toda sua atividade em uma temperatura aumentada ou pode ser mais ativa em uma temperatura aumentada que a sua temperatura normal (por exemplo, temperatura ambiente) ou sua temperatura ideal antes da mutagênese para obter termoestabilidade intensificada.

Os termos "termotolerante" e "termotolerância" como aqui usados com referência a uma enzima significam a capacidade da enzima de normalmente operar após exposição em temperatura alta, embora a temperatura alta possa desativar a enzima temporariamente.

GERAÇÃO E MANIPULAÇÃO DOS ÁCIDOS NUCLÉICOS

A invenção fornece ácidos nucléicos, incluindo cassetes de expressão como vetores de expressão, codificando os polipeptídeos e fitases da invenção. A invenção também inclui métodos para descobrir novas seqüências de fitase usando os ácidos nucléicos da invenção. Também fornecidos são métodos para modificar os ácidos nucléicos da invenção, por exemplo, por remontagem de ligação sintética, sistema de evolução direcionada otimizada e/ou mutagênese de saturação.

Os ácidos nucléicos da invenção podem ser feitos, isolados e/ou manipulados, por exemplo, por clonagem e expressão de bibliotecas de cDNA, amplificação de DNA mensageiro ou genômico por PCR e similares. Praticando os métodos da invenção, genes homólogos podem ser modifica-

dos manipulando um ácido nucléico modelo, como descrito aqui. A invenção pode ser praticada junto com qualquer método ou protocolo ou meios conhecidos na técnica, que são bem descritos na literatura científica e de patentes.

TÉCNICAS GERAIS

Os ácidos nucléicos usados para praticar esta invenção, ou RNA, iRNA, ácido nucléico de anti-sentido, cDNA, DNA genômico, vetores, vírus ou híbridos destes, podem ser isolados de uma variedade de fontes, geneticamente criadas, amplificadas e/ou expressadas/recombinantemente geradas. Os polipeptídeos recombinantes gerados destes ácidos nucléicos podem ser individualmente isolados ou clonados e testados por uma atividade desejada. Qualquer sistema de expressão recombinante pode ser usado, incluindo sistemas de expressão bacteriana, mamífera, de levedura, de inseto ou de planta.

Alternativamente, estes ácidos nucléicos podem ser sintetizados *in vitro* através de técnicas de síntese química bem-conhecidas, como descrito em, por exemplo, 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 4.458.066.

Técnicas para a manipulação de ácidos nucléicos, como, por exemplo, subclonagem, sondas de marcação (por exemplo, marcação de iniciador aleatório usando polimerase de Klenow, tradução de entalhe, amplificação), seqüenciação, hibridação e similares, são bem descritas na literatura científica e de patentes, ver, por exemplo, Sambrook, ed., MOLECULAR CLONING: A LABORATORY MANUAL (2ND 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, Parte I. Theory and Nucleic Acid Preparation, Tijssen, ed. Elsevier, N. I.

(1993).

Outros meios úteis de obter e manipular ácidos nucléicos usados para praticar os métodos da invenção são clonagem das amostras genômicas, e, se desejado, triar e reclonar inserções isoladas ou amplificadas de, por exemplo, clones genômicos ou clones de cDNA. Fontes de ácido nucléico usadas nos métodos da invenção incluem bibliotecas genômicas ou de cDNA contido, por exemplo, nos cromossomos artificiais mamíferos (MACS), ver, por exemplo, Patentes U. S. Nºs 5.721.118; 6.025.155; cromossomos artificiais humanos, ver, por exemplo, Rosenfeld (1997) Nat. Genet. 15:333-335; cromossomos artificiais de levedura (YAC); cromossomos artificiais bacterianos (BAC); cromossomos artificiais de P1, ver, por exemplo, Woon (1998) Genomics 50:306-316; vetores P1-derivados (PACs), ver, por exemplo, Kern (1997) Biotechniques 23:120-124; cosmídeos, vírus recombinantes, fagos ou plasmídeos.

Em um aspecto, um ácido nucléico codificando um polipeptídeo da invenção é montado em fase apropriada com uma seqüência líder capaz de direcionar secreção do polipeptídeo transferido ou fragmento deste.

A invenção fornece proteínas de fusão e ácidos nucléicos que os codificam. Um polipeptídeo da invenção pode ser fundido em um peptídeo ou polipeptídeo heterólogo, como peptídeos de identificação N-terminais que dão características desejadas, como estabilidade aumentada ou purificação simplificada. Os peptídeos e polipeptídeos da invenção podem também ser sintetizados e expressados como proteínas de fusão com um ou mais domínios adicionais a eles ligadas para, por exemplo, produzir um peptídeo mais imunogênico, para mais facilmente isolar um peptídeo recombinantemente sintetizado, para identificar e isolar anticorpos e células B de expressão de anticorpos, e similares. A detecção e purificação que facilitam os domínios incluem, por exemplo, peptídeos quelantes de metal como tratos de poliistidina e módulos de histidina-triptofano que permitem purificação em metais imobilizados, domínios de proteína A que permitem purificação em imunoglobulina imobilizada e o domínio utilizado no sistema de purificação de extensão/afinidade de FLAGS (Immunex Corp, Seattle WA). A inclusão de

umas seqüências ligantes cliváveis como Fator Xa ou enterocinase (Invitrogen, San Diego CA) entre um domínio de purificação e o peptídeo ou polipeptídeo compreendendo o motivo facilita a purificação. Por exemplo, um vetor de expressão pode incluir uma seqüência de ácidos nucléicos de codificação de epítopo ligada a seis resíduos de histidina seguidos por uma tioredoxina e um sítio de clivagem de enterocinase (ver por exemplo, 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 sítio de clivagem de enterocinase fornece um meio para purificar o epítopo do resto da proteína de fusão. Tecnologia relacionada aos vetores que codificam proteínas de fusão e aplicação de proteínas de fusão é bem descrita na literatura científica e de patente, ver por exemplo, Kroll (1993) DNA Cell. Biol., 12:441-53.

SEQÜÊNCIAS DE CONTROLE TRANSCRICIONAL E TRADUÇÃO

A invenção fornece seqüências de ácidos nucléicos (por exemplo, DNA) da invenção operativamente ligadas à(s) seqüência(s) de controle de expressão (por exemplo, transcricional ou translacional). por exemplo, promotores ou intensificadores, para direcionar ou modular síntese /expressão de RNA. A seqüência de controle de expressão pode ser um vetor de expressão. Promotores bacterianos exemplares incluem lacI, lacZ, T3, T7, gpt, PR lambda, PL e trp. Promotores eucarióticos exemplares incluem CMV precoce imediato, timidina cinase de HSV, SV40 precoce e tardio, LTRs de retrovírus e metalotioneína de camundongo I.

Promotores adequados para expressar, ou sobreexpressar, um polipeptídeo em bactérias incluem os promotores de lac ou trp de E. coli, o promotor de lacI, o promotor de lacZ, o promotor de T3, o promotor de T7, o promotor de gpt, o promotor de lambda PR, o promotor de lambda PL, promotores de óperons que codificam enzimas glicolíticas como 3-fosfoglicerato cinase (PGK) e o promotor de ácido fosfatase. Promotores eucarióticos incluem o promotor de CMV precoce imediato, o promotor de timidina cinase de HSV, promotores de choque térmico, o promotor de SV40 precoce e tardio, LTRs de retrovírus e o promotor de metalotioneína I de camundongo.

Outros promotores conhecidos para controlar expressão de genes em células procarióticas ou eucarióticas ou seus vírus podem ser também usados.

VETORES DE EXPRESSÃO E VEÍCULOS DE CLONAGEM

A invenção fornece vetores de expressão e veículos de clonagem compreendendo ácidos nucléicos da invenção, por exemplo, seqüências que codificam as fitases da invenção para expressão e sobreexpressão dos polipeptídeos da invenção (e ácidos nucléicos, por exemplo, antisentido). Vetores de expressão e veículos de clonagem da invenção podem compreender partículas viróticas, baculovírus, fago, plasmídeos, fagomídeos, cosmídeos, fosmídeos, cromossomos artificiais bacterianos, DNA virótico (por exemplo, vacínia, adenovírus, vírus do epitelioma contagioso, pseudo-raiva e derivados de SV40), cromossomos artificiais com base em P1, plasmídeos de levedura, cromossomos artificiais de levedura, e qualquer outro vetor específico para hospedeiros específicos de interesse (como bacilo, Aspergilo e levedura). Os vetores da invenção podem incluir seqüências de DNA cromossômicas, não-cromossômicas e sintéticas. Grandes números de vetores adequados são conhecidos por aqueles de habilidade na técnica, e estão comercialmente disponíveis. Vetores exemplares incluem: bacterianos: vetores de pQE (Qiagen), plasmídeos de pBluescript, vetores de pNH, (vetores lambda-ZAP (Stratagene); ptrc99a, pKK223-3, pDR540, pRIT2T (Pharmacia); eucarióticos: pXT1, pSG5 (Stratagene), pSVK3, pBPV, pMSG, pSVLSV40 (Pharmacia). Porém, qualquer outro plasmídeo ou outro vetor pode ser usado tanto quanto eles forem replicáveis e viáveis no hospedeiro. Vetores de baixo número de cópias ou de alto número de cópias podem ser empregados com a presente invenção.

O vetor de expressão pode compreender um promotor, um sítio de ligação de ribossomo para iniciação da tradução e um terminador de transcrição. O vetor também pode incluir seqüências apropriadas para amplificar a expressão. Os vetores de expressão de mamíferos podem compreender uma origem de replicação, quaisquer sítios de ligação ao ribossomo necessários, um sítio de poliadenilação, doador de junção e sítios de aceptores, seqüências de terminação transcricional, e seqüências não-transcritas

de flanqueamento 5'. Em alguns aspectos, as seqüências de DNA derivadas da junção de SV40 e sítios de poliadenilação podem ser usados para fornecer os elementos genéticos não-transcritos requeridos.

Em um aspecto, os vetores de expressão contêm um ou mais genes marcadores selecionáveis para permitir a seleção de células hospedeiras contendo o vetor. Tais marcadores selecionáveis incluem genes que codificam diidrofolato reductase ou genes que conferem resistência à neomicina para a cultura de células eucarióticas, genes que conferem resistência à tetraciclina ou à ampicilina em E. coli e os genes de TRP1 de S. Cerevisiae. As regiões de promotor podem ser selecionadas de qualquer gene desejado usando vetores de cloranfenicol transferase (CAT) ou outros vetores com marcadores selecionáveis.

Os vetores para expressar o polipeptídeo ou fragmento deste em células eucarióticas podem também conter intensificadores para aumentar os níveis de expressão. Intensificadores são elementos de ação cis de DNA, usualmente de cerca de 10 a cerca de 300 bp no comprimento que agem em um promotor para aumentar sua transcrição. Exemplos incluem o intensificador de SV40 no último lado da origem de replicação bp 100 a 270, o intensificador de promotor precoce de citomegalovírus, o intensificador de polioma no último lado da origem de replicação e os intensificadores de adenovírus.

Uma seqüência de DNA pode ser inserida em um vetor por uma variedade de procedimentos. Em geral, a seqüência de DNA é ligada à posição desejada no vetor seguindo digestão da inserção e o vetor com endonucleases de restrição apropriadas. Alternativamente, as extremidades cegas tanto na inserção quanto no vetor podem ser ligadas. Uma variedade de técnicas de clonagem é conhecida na técnica, por exemplo, como descrito em Ausubel e Sambrook. Tais procedimentos e outros são julgados estar dentro do escopo daqueles versados na técnica.

O vetor pode ser na forma de um plasmídeo, uma partícula virótica ou um fago. Outros vetores incluem seqüências de DNA cromossômicas, não-cromossômicas e sintéticas, derivados de SV40; plasmídeos bacte-

rianos, DNA de fago, baculovírus, plasmídeos de levedura, vetores derivados de combinações de plasmídeo e DNA de fago, DNA virótico como vacínia, adenovírus, vírus do epitelioma contagioso e pseudo-raiva. Uma variedade de vetores de clonagem e expressão para o uso com hospedeiros procarióticos e eucarióticos é descrita, por exemplo, por Sambrook.

Vetores bacterianos particulares que podem ser usados incluem os plasmídeos comercialmente disponíveis compreendendo elementos genéticos do vetor de clonagem bem-conhecido pBR322 (ATCC 37017), pKK223-3 (Pharmacia Fine Chemicals, Uppsala, Suécia), GEM1 (Promega Biotec, Madison, WI, USA) 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. Vetores eucarióticos particulares incluem pSV2CAT, pOG44, pXTI, pSG (Stratagene) pSVK3, pBPV, pMSG e pSVL (Pharmacia). Porém, qualquer outro vetor pode ser usado contanto que seja replicável e viável na célula hospedeira.

CÉLULAS HOSPEDEIRAS E CÉLULAS TRANSFORMADAS

A invenção também fornece uma célula transformada compreendendo uma seqüência de ácidos nucléicos da invenção, por exemplo, uma seqüência codificando uma fitase da invenção, um vetor da invenção. A célula hospedeira pode ser quaisquer das células hospedeiras familiares a aqueles versados na técnica, incluindo células procarióticas, células eucarióticas, como células bacterianas, células fúngicas, células de levedura, células mamíferas, células de inseto ou células vegetais. Células bacterianas exemplares incluem *E. coli, Streptomyces, Bacillus subtilis, Salmonella typhimurium* e várias espécies dentro dos gêneros *Pseudomonas*, *Streptomyces* e *Staphylococcus*. Células de inseto exemplares incluem *Drosophila S2* e *Spodoptera Sf9*. Células animais exemplares incluem CHO, COS ou melanoma de Bowes ou qualquer linhagem de células de camundongo ou humanas. A seleção de um hospedeiro apropriado está dentro das habilidades daqueles versados na técnica.

O vetor pode ser introduzido nas células hospedeiras usando

qualquer de uma variedade de técnicas, incluindo transformação, transfecção, transdução, infecção virótica, disparos de genes ou tradução de gene Ti-mediada. Métodos particulares incluem transfecção de fosfato de cálcio, transfecção mediada por DEAE-dextrana, lipofecção ou eletroporação (Davis, L., Dibner, M., Battey, I., Basic Methods in Molecular Biology, (1986)).

Onde apropriado, as células hospedeiras criadas podem ser cultivadas em meios nutrientes convencionais modificados como apropriado para ativar os promotores, selecionando os transformantes ou amplificando os genes da invenção. A transformação a seguir de uma cepa hospedeira adequada e desenvolvimento da cepa hospedeira para uma densidade de célula apropriada, o promotor selecionado pode ser induzido através de meios apropriados (por exemplo, alteração de temperatura ou indução química) e as células podem ser cultivadas durante um período adicional para lhes permitir produzir o polipeptídeo desejado ou fragmento deste.

As células podem ser colhidas por centrifugação, rompimentos por meios físicos ou químicos e o extrato bruto resultante é retido para purificação adicional. Células microbianas empregadas para expressão das proteínas podem ser rompidas por qualquer método conveniente, incluindo ciclos de congelamento-descongelameto, sonicação, rompimento mecânico ou uso de agentes de lise de células. Tais métodos são bem-conhecidos àqueles versados na técnica. O polipeptídeo expresso ou fragmento deste ou pode ser restabelecido e purificado de culturas de células recombinantes através de métodos incluindo precipitação de sulfato de amônio ou de etanol, extração de ácido, cromatografia de permuta de ânion ou cátion, cromatografia de fosfocelulose, cromatografia de interação hidrofóbica, cromatografia de afinidade, cromatografia de hidioxilapatite e cromatografia de lectina. As etapas de redobragem de proteína podem ser usadas, quando necessário, completando a configuração do polipeptídeo. Se desejado, cromatografia líquida de alto desempenho (HPLC) pode ser empregada para etapas de purificação finais.

Também podem ser empregados vários sistemas de cultura de células mamíferas para expressar, ou sobreexpressar a proteína recombi-

nante. Exemplos de sistemas de expressão mamíferos incluem as linhagens de COS-7 de fibroblastos de rim de macaco e outras linhagens de células capazes de expressar proteínas de um vetor compatível, como as linhagens de células C127, 3T3, CHO, HeLa e BHK.

Os constructos em células hospedeiras podem ser usados de uma maneira convencional para produzir o produto de gene codificado pela seqüência recombinante. Dependendo do hospedeiro empregado em um procedimento de produção recombinante, os polipeptídeos produzidos por células hospedeiras contendo o vetor podem ser glicosilados ou não-glicosilados. Os polipeptídeos da invenção podem ou não também incluir um resíduo de aminoácido de metionina inicial.

Os sistemas de tradução livre de células também podem ser empregados para produzir um polipeptídeo da invenção. Sistemas de tradução livre de células podem usar mRNAs transcritos de um constructo de DNA compreendendo um promotor operavelmente ligado a um ácido nucléico codificando o polipeptídeo ou fragmento deste. Em alguns aspectos, o constructo de DNA pode ser linearizado antes conduzindo uma reação de transcrição *in vitro*. O mRNA transcrito é depois incubado com um extrato de tradução livre de células apropriado, como um extrato de reticulócito de coelho, para produzir o polipeptídeo desejado ou fragmento deste.

Os vetores de expressão podem conter um ou mais genes marcadores selecionáveis para fornecer uma característica fenotípica para seleção de células hospedeiras transformadas como diidrofolato reductase ou resistência à neomicina por cultura de células eucarióticas, ou como resistência à tetraciclina ou ampicilina em E. coli.

AMPLIFICAÇÃO DE ÁCIDOS NUCLÉICOS

Na prática da invenção, ácidos nucléicos codificando os polipeptídeos da invenção, ou ácidos nucléicos modificados, podem ser reproduzidos, por exemplo, por amplificação. A invenção fornece pares de seqüência de iniciadores de amplificação para amplificar os ácidos nucléicos codificando polipeptídeos com uma atividade de fitase onde os pares de iniciador são capazes de amplificar seqüências de ácidos nucléicos incluindo a

SEQ ID Nº: 1 exemplar, ou uma subseqüência desta. Aqueles versados na técnica pode designar os pares de seqüência de iniciadores de amplificação para qualquer parte ou comprimento total destas seqüências; por exemplo:

A SEQ ID Nº: 1 exemplar é

ATGAAAGCGATCTTAATCCCATTTTTATCTCTTCTGATTCCGTTAACCCCG
CAATCTGCATTCGCTCAGAGTGAGCCGGAGCTGAAGCTGGAAAGTGTGG
TGATTGTCAGTCGTCATGGTGTGCGTGCTCCAACCAAGGCCACGCAACT
GATGCAGGATGTCACCCCAGACGCATGGCCAACCTGGCCGGTAAAACTG
GGTGAGCTGACACCGCGCGGTGGTGAGCTAATCGCCTATCTCGGACATT
ACTGGCGTCAGCGTCTGGTAGCCGACGGATTGCTGCCTAAATGTGGCTG
CCCGCAGTCTGGTCAGGTCGCGATTATTGCTGATGTCGACGAGCGTACC
CGTAAAACAGGCGAAGCCTTCGCCGCCGGGCTGGCACCTGACTGTGCAA
TAACCGTACATACCCAGGCAGATACGTCCAGTCCCGATCCGTTATTTAAT
CCTCTAAAAACTGGCGTTTGCCAACTGGATAACGCGAACGTGACTGACG
CGATCCTCGAGAGGGCAGGAGGGTCAATTGCTGACTTTACCGGGCATTA
TCAAACGGCGTTTCGCGAACTGGAACGGGTGCTTAATTTTCCGCAATCAA
ACTTGTGCCTTAAACGTGAGAAACAGGACGAAAGCTGTTCATTAACGCAG
GCATTACCATCGGAACTCAAGGTGAGCGCCGACTGTGTCTCATTAACCG
GTGCGGTAAGCCTCGCATCAATGCTGACGGAGATATTTCTCCTGCAACAA
GCACAGGGAATGCCGGAGCCGGGGTGGGGAAGGATCACCGATTCACAC
CAGTGGAACACCTTGCTAAGTTTGCATAACGCGCAATTTGATTTGCTACAA
CGCACGCCAGAGGTTGCCCGCAGCCGCGCCACCCCGTTATTAGATTTGA
TCAAGACAGCGTTGACGCCCCATCCACCGCAAAAACAGGCGTATGGTGT
GACATTACCCACTTCAGTGCTGTTTATCGCCGGACACGATACTAATCTGG
CAAATCTCGGCGGCGCACTGGAGCTCAACTGGACGCTTCCCGGTCAGCC
GGATAACACGCCGCCAGGTGGTGAACTGGTGTTTGAACGCTGGCGTCGG
CTAAGCGATAACAGCCAGTGGATTCAGGTTTCGCTGGTCTTCCAGACTTT
ACAGCAGATGCGTGATAAAACGCCGCTGTCATTAAATACGCCGCCCGGA
GAGGTGAAACTGACCCTGGCAGGATGTGAAGAGCGAAATGCGCAGGGC
ATGTGTTCGTTGGCAGGTTTTACGCAAATCGTGAATGAAGCACGCATACC
GGCGTGCAGTTTGAGATCTCATCTA

Desse modo um par de seqüências de iniciadores de amplifica-

ção exemplar é resíduos 1 a 21 da SEQ ID Nº: 1 (isto é, ATGAAAGCGATCTTAATCCCA) e o filamento complementar dos últimos 21 resíduos da SEQ ID Nº: 1 (isto é, o filamento complementar de TGCAGTTTGAGATCTCATCTA).

As reações de amplificação também podem ser usadas para quantificar a quantidade de ácido nucléico em uma amostra (como a quantidade de mensagem em uma amostra de célula), marcar o ácido nucléico (por exemplo, aplicá-lo a um arranjo ou uma macha), detectar o ácido nucléico, ou quantificar a quantidade de um ácido nucléico específico em uma amostra. Em um aspecto da invenção, a mensagem isolada de uma célula ou uma biblioteca de cDNA é amplificada. O versado na técnica pode selecionar e designar os iniciadores de amplificação de oligonucleotídeo adequados. Métodos de amplificação também são bem-conhecidos na técnica, e incluem, por exemplo, reação em cadeia de polimerase, PCR (ver, por exemplo, 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., reação em cadeia de lipase (LCR) (ver, por exemplo, Wu (1989) Genomics 4:560; Landegren (1988) Science 241:1077; Barringer (1990) Gene 89:117); amplificação de transcrição (ver, por exemplo, Kwoh (1989) Proc. Natl. Acad. Sci. USA 86:1173); e, replicação de seqüência autocontínua (ver, por exemplo, Guatelli (1990) Proc. Natl. Acad. Sci. USA 87:1874); amplificação de Q-beta replicase (ver, por exemplo, Smith (1997) J. Clin. Microbiol. 35:1477-1491), ensaio de amplificação de Q-beta replicase automatizado (ver, por exemplo, Burg (1996) Mol. Cell. Probes 10:257-271) e outras técnicas mediadas por RNA polimerase (por exemplo, NASBA, Cangene, Mississauga, Ontário); também ver Berger (1987) Methods Enzymol. 152:307-316; Sambrook; Ausubel; Patentes U. S. Nºs 4.683.195 e 4.683.202; Sooknanan (1995) Biotechnology 13:563-564.

DETERMINAÇÃO DO GRAU DA IDENTIDADE DA SEQÜÊNCIA

A invenção fornece um ácido nucléico isolado ou recombinante compreendendo uma seqüência de ácidos nucléicos tendo pelo menos 98 %

de identidade de seqüência à SEQ ID Nº: 1 em uma região de pelo menos cerca de 100 resíduos, em que os ácidos nucléicos codificam pelo menos um polipeptídeo tendo uma atividade de fitase e as identidades das seqüências são determinadas por análise com um algoritmo de comparação de seqüência ou por uma inspeção visual. Em modalidades alternativas a seqüência de ácidos nucléicos tem pelo menos 98 %, 98,5 %, 99 % ou 99,5 % de identidade de seqüência à SEQ ID Nº: 1 em uma região de pelo menos cerca de 50 resíduos, 100 resíduos, 150 resíduos, 200 resíduos, 250 resíduos, 300 resíduos, 350 resíduos, 400 resíduos, 450 resíduos, 500 resíduos, 550 resíduos, 600 resíduos, 700 resíduos, 800 resíduos, 900 resíduos, 1000 resíduos, 1200 resíduos ou 1300 resíduos. A seqüência de ácidos nucléicos pode ter uma seqüência como exposto na SEQ ID Nº: 1. Em um aspecto, a extensão de identidade de seqüência (homologia) pode ser determinada usando qualquer programa de computador e parâmetros associados, incluindo aqueles aqui descritos, como BLAST 2.2.2. ou FASTA versão 3.0t78, com os parâmetros padrão.

Seqüências homólogas também incluem seqüências de RNA em que as uridinas substituem as timinas nas seqüências de ácidos nucléicos. As seqüências homólogas podem ser obtidas usando quaisquer dos procedimentos aqui descritos ou podem ser o resultado da correção de um erro de seqüenciação. Será apreciado que as seqüências de ácidos nucléicos como aqui expostas podem ser representadas no formato de caractere simples tradicional (ver, por exemplo, Stryer, Lubert. Biochemistry, 3ª Ed., W. H. Freeman & Co., Nova Iorque) ou em qualquer outro formato que registra a identidade dos nucleotídeos em uma seqüência.

Vários programas de comparação de seqüência aqui identificados são usados neste aspecto da invenção. Identidades das seqüências de proteína e/ou ácidos nucléicos (homologias) podem ser avaliadas usando qualquer da variedade de algoritmos de comparação de seqüência e programas conhecidos na técnica. Tais algoritmos e programas incluem, mas não são 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.

Homologia ou identidade pode ser medida usando software de análise de seqüência (por exemplo, Pacote de Software de Análise de Seqüência do Genetics Computer Group, Universidade do Centro de Biotecnologia de Wisconsin, avenida Universidade, nº 1710, Madison, WI 53705). Tal software combina seqüências similares atribuindo graus de homologia às várias deleções, substituições e outras modificações. Os termos "homologia" e "identidade" no contexto de duas ou mais seqüências de ácidos nucléicos ou polipeptídeos, referem-se a duas ou mais seqüências ou subseqüência que são as mesmas ou têm uma porcentagem especificada de resíduos de aminoácidos ou nucleotídeos que são os mesmos quando comparados e alinhados para correspondência máxima em uma janela de comparação ou região designada quando medida usando qualquer número de algoritmos de comparação de seqüência ou por alinhamento manual e inspeção visual. Para comparação das seqüências, uma seqüência pode agir como uma seqüência de referência (uma seqüência exemplar SEQ ID Nº: 1, SEQ ID Nº: 2) à qual as seqüências de teste são comparadas. Quando usar um algoritmo de comparação de seqüência, as seqüências de teste e de referência são entradas em um computador, coordenadas de subseqüência são designadas, se necessário, e parâmetros de programa de algoritmo de seqüência são designados. Os parâmetros de programa padrão podem ser usados ou parâmetros alternativos podem ser designados. O algoritmo de comparação de seqüência depois calcula o percentual de identidade das seqüências para as seqüências de teste relativo à seqüência de referência, com base nos parâmetros do programa.

Uma "janela de comparação", como aqui usado, inclui referência a um segmento de qualquer um dos vários resíduos contíguos. Por exemplo, em aspectos alternativos da invenção, os resíduos contíguos que variam em qualquer lugar de 20 ao comprimento total das seqüências exemplares SEQ

ID Nº: 1, SEQ ID Nº: 2 são comparados a uma seqüência de referência do mesmo número de posições contíguas após as duas seqüências estarem otimamente alinhadas. Se a seqüência de referência tem a identidade de seqüência requerida à SEQ ID Nº: 1, SEQ ID Nº: 2, por exemplo, 98 % de identidade de seqüência à SEQ ID Nº: 1, SEQ ID Nº: 2, aquela seqüência está dentro do escopo da invenção. Em modalidades alternativas, subseqüências que variam de cerca de 20 a 600, cerca de 50 a 200 e cerca de 100 a 150 são comparadas a uma seqüência de referência do mesmo número de posições contíguas após as duas seqüências estarem otimamente alinhadas. Métodos de alinhamento de seqüências para comparação são bem-conhecidos na técnica. Ótimo alinhamento das seqüências para comparação pode ser conduzido, por exemplo, 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, pela pesquisa por método de similaridade de Person & Lipman, Proc. Natl. Acad. Sci. USA 85:2444, 1988, por implementações computadorizadas destes algoritmos (GAP, BESTFIT, FASTA e TFASTA no Pacote de Software de Genética de Wisconsin, Genetics Computer Group, 575 Science Dr., Madison, WI), ou por alinhamento manual e inspeção visual. Outros algoritmos para determinar homologia ou identidade incluem, por exemplo, além de um programa de BLAST (Basic Local Alignment Search Tool no Centro Nacional para Informação Biológica), 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 Comparation), LALIGN (Local Sequence Alignment), LCP (Local Content Program),

MACAW (Multiple Alignment Construction & Analysis Workbench), MAP (Multiple Alignment Program), MBLKP, MBLKN, PIMA (Pattern-Induce Multi-sequence Alignment), SAGA (Sequence Alignment by Genetic Algorithm) e WHAT-IF. Tais programas de alinhamento podem ser usados para triar bancos de dados de genoma para identificar seqüências de polinucleotídeos tendo seqüências substancialmente idênticas. Por exemplo, vários bancos de dados de genoma estão disponíveis, uma porção substancial do genoma humano está disponível como parte do Projeto de Seqüenciação de Genoma Humano (Gibbs, 1995). Vários genomas foram seqüenciados, por exemplo, *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). Progresso significativo também foi feito na seqüenciação dos genomas de organismo modelo, como camundongo, *C. elegans*, e *Arabadopsis sp.*. Bancos de dados contendo informação genômica anotada com alguma informação funcional são mantidos por organização diferente, e são acessíveis por meio da Internet.

Algoritmos de BLAST, BLAST 2.0 e BLAST 2.2.2 são também usados para praticar a invenção. Eles são descritos, por exemplo, em Altschul (1977) Nuc. Acids Res. 25:3389-3402; Altschul (1990) J. Mol. Biol. 215:403-410. Software para executar análises de BLAST está publicamente disponível através do Centro Nacional para Informação de Biotecnologia. Este algoritmo envolve primeiro identificar os pares de seqüência de alta escalação (HSPs) para identificar palavras curtas de comprimento W na seqüência de consulta, que combina ou satisfaz algum escore T de limite de valor positivo quando alinhados com uma palavra do mesmo comprimento em uma seqüência do banco de dados. T está referido como o limite de escore de palavra vizinha (Altschul (1990) supra). Estes acertos de palavras vizinhas iniciais agem como sementes para iniciar pesquisas para encontrar HSPs mais longos os contendo. Os acertos de palavra são entendidos em ambas direções ao longo de cada seqüência até o escore de alinhamento cumulativo poder ser aumentado. Os escores cumulativos são calculados

usando, para seqüências de nucleotídeos, os parâmetros M (escore de recompensa para um par de resíduos combinando; sempre >0). Para as seqüências de aminoácidos, uma matriz de escore é usada para calcular o escore cumulativo. Extensão dos acertos da palavra em cada direção é detida quando: o escore de alinhamento cumulativo cai pela quantidade X de seu valor alcançado máximo; o escore cumulativo tornar-se zero ou menos, devido à acumulação de um ou mais alinhamentos de resíduo de escore negativo; ou a extremidade de qualquer seqüência é alcançada. Os parâmetros de algoritmo de BLAST W T, e X determinam a sensibilidade e velocidade do alinhamento. O programa de BLASTN (para seqüências de nucleotídeos) usa como padrão um comprimento de palavra (W) de 11, uma expectativa (E) de 10, M=5, N=-4 e uma comparação de ambos os filamentos. Para seqüências de aminoácidos, o programa de BLATP usa como padrão um comprimento de palavra de 3, e expectativas (E) de 10, e a matriz de contagem de BLOSLTM62 (ver Henikoff & Henikoff (1989) Proc. Natl. Acad. Sci. USA 89:10915) alinhamentos (B) de 50, expectativa (E) de 10, M=5, N = -4, e uma comparação de ambos os filamentos. O algoritmo de BLAST também executa uma análise estatística da similaridade entre duas seqüências (ver, por exemplo, Karlin & Altschul (1993) Proc. Natl. Acad. Sci. USA 90:5873). Uma medida de similaridade fornecida por algoritmo de BLAST é a menor probabilidade de soma (P(N)) que fornece uma indicação da probabilidade pela qual uma partida entre duas seqüências de nucleotídeos ou de aminoácidos ocorreria por casualidade. Por exemplo, um ácido nucléico é considerado similar a uma seqüência de referência se a menor probabilidade de soma em uma comparação do ácido nucléico de teste ao ácido nucléico de referência for menor que cerca de 0,2, menor que cerca de 0,01 ou menor que cerca de 0,001. Em um aspecto, as homologias de seqüência de ácidos nucléicos e de proteínas são avaliadas usando Basic Local Alignment Search Tool ("BLAST"). Por exemplo, cinco programas de BLAST específicos podem ser usados para executar a tarefa a seguir: (1) BLASTP e BLAST3 comparam uma seqüência de consulta de aminoácido junto a um banco de dados de seqüência de proteínas; (2) BLASTN compara uma seqüência de

consulta de nucleotídeo junto a um banco de dados de seqüência de nucleotídeos; (3) BLASTX compara os produtos de traduções conceituais de seis estruturas de uma seqüência de nucleotídeos de consulta (ambos os filamentos) junto a um banco de dados de seqüência de proteínas; (4) TBLASTN compara uma seqüência de proteína de consulta junto a um banco de dados de seqüência de nucleotídeos transferida em todas as seis estruturas de leitura (ambos os filamentos); e, (5) TBLASTX compara as traduções de seis estruturas de uma seqüência de consulta de nucleotídeo com as traduções de seis estruturas de um banco de dados de seqüência de nucleotídeo. Os programas de BLAST identificam seqüências homólogas identificando segmentos similares, que são aqui referidos como "pares de segmentos de escore alto," entre uma seqüência de aminoácidos ou de ácidos nucléicos de consulta e uma seqüência de teste que podem ser obtidas de uma proteína ou banco de dados de seqüência de ácidos nucléicos. Pares de segmento de escore alto podem ser identificados (isto é, alinhados) por meio de uma matriz de contagem, muitas destas são conhecidas na técnica. Uma matriz de contagem exemplar usada é a matriz de BLOSUM62 (Gonnet et al., Science 256:1443-1445, 1992; Henikoff e Henikoff, Proteins 17:49-61, 1993). Alternativamente, matrizes de PAM ou PAM250 podem ser usadas (ver, por exemplo, Schwartz e Dayhoff, eds., 1978, Matrices for Detecting Distance Relationships: Atlas of Protein Sequence and Structure, Washington: Fundação de Pesquisa Biomédica Nacional).

Em um aspecto da invenção, para determinar se um ácido nucléico tem a identidade de seqüência requerida para estar dentro do escopo da invenção, os programas de NCBI BLAST 2.2.2 são usados. Opções padrão para blastp. Há aproximadamente 38 opções de colocação no programa de BLAST 2.2.2. Neste aspecto exemplar da invenção, todos os valores padrão são usados com a exceção para o ajuste de filtração padrão (isto é, todos os parâmetros ajustados para padrão exceto filtração que é ajustada para OFF); em seu lugar um ajustamento de "- F F" é usado, que incapacita a filtração. Uso de filtração padrão freqüentemente resulta em violações de Karlin-Altschul devido à extensão curta da seqüência.

Os valores padrão usados neste aspecto exemplar da invenção incluem:

"Filtro para baixa complexidade: ON

> Tamanho de Palavra: 3

> Matriz: Blosum62

> Custos de intervalos: Existência: 11

> Extensão: 1"

"Filtro para baixa complexidade: ON

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

Um ajuste de programa de NCBI BLAST 2.2.2 exemplar é determinado no Exemplo 1, abaixo. Observar que a opção "-W" padrão em 0. Isto significa que, se não-ajustado, o tamanho de palavra padrão em 3 para as proteínas e 11 para nucleotídeos.

SISTEMAS DE COMPUTADOR E PRODUTOS DE PROGRAMA DE COMPUTADOR

Para determinar e identificar as identidades de seqüência, homologias estruturais, motivos e similares in silico a seqüência da invenção pode ser armazenada, registrada e manipulada em qualquer meio que pode ser lido e pode ser acessado por um computador. Conseqüentemente, a invenção fornece computadores, sistemas de computador, meios legíveis por computador, produtos de programas de computação e similares neles registradas ou armazenadas as seqüências de ácidos nucléicos e polipeptídeos da invenção, por exemplo, as seqüências exemplares SEQ ID Nº: 1, SEQ ID Nº: 2. Como aqui usado, as palavras "registrada" e "armazenada" referem-se a um processo para armazenar informação em um meio de computador. Um versado na técnica pode facilmente adotar qualquer método conhecido para gravar informação em um meio legível por computador para gerar fabricações compreendendo uma ou mais das seqüências de ácidos nucléicos e/ou de polipeptídeos da invenção.

Outro aspecto da invenção é um meio legível por computador tendo nele registrada pelo menos uma seqüência de ácido nucléico e/ou de polipeptídeos da invenção. Meios legíveis por computador incluem meios magneticamente legíveis, meios opticamente legíveis, meios eletronicamente legíveis e meios magnéticos/ópticos. Por exemplo, os meios legíveis por computador podem ser um disco rígido, um disquete, uma fita magnética, CD-ROM, Disco Versátil Digital (DVD), Memória de Acesso Aleatório (RAM) ou Memória Exclusiva de Leitura (ROM) como também outros tipos de outros meios conhecidos àqueles versados na técnica.

Aspectos da invenção incluem sistemas (por exemplo, sistemas com base na Internet), particularmente sistemas de computador que armazenam e manipulam as seqüências e informação de seqüências aqui descritas. Um exemplo de um sistema de computador 100 é ilustrado na forma de diagrama de bloco na Figura 15. Como aqui usado, um "sistema de computador" refere-se aos componentes de hardware, componentes de software e componentes de armazenamento de dados usados para analisar uma seqüência de nucleotídeos ou de polipeptídeos da invenção. O sistema de computador 100 pode incluir um processador para processamento, acessando e manipulando os dados da seqüência. O processador 105 pode ser qualquer tipo bem-conhecido de unidade de processamento central, como, por exemplo, o Pentium III da Intel Corporation, ou processador similar de Sun, Motorola, Compaq, AMD ou International Business Machines. O sistema de computador 100 é um sistema de propósito geral compreendendo o processador 105 e um ou mais componentes de armazenamento de dados internos 110 para armazenar dados, e um ou mais dispositivos de recuperação de dados para recuperar os dados armazenados nos componentes de armazenamento de dados. Um versado na técnica pode facilmente apreciar que qualquer um dos sistemas de computador atualmente disponíveis é adequado.

Em um aspecto, o sistema de computador 100 inclui um processador 105 conectado a um barramento que é conectado a uma memória principal 115 (pode ser implementada como RAM) e um ou mais dispositivos

de armazenamento de dados internos 110, como um disco rígido e/ou outro meios legíveis por computador tendo os dados nele gravados. O sistema de computador 100 pode também incluir um ou mais dispositivos de recuperação de dados 118 para ler os dados armazenados nos dispositivos de armazenamento de dados internos 110.

Os dispositivos de recuperação de dados 118 podem representar, por exemplo, uma unidade de disco, um drive de disco a laser, um drive de fita magnética, ou um modem capaz de conexão a um sistema de armazenamento de dados remoto (por exemplo, por meio da Internet) etc. Em algumas modalidades, o dispositivo de armazenamento de dados interno 110 é um meio legível por computador removível como um disquete, um disco a laser, uma fita magnética, etc. contendo nele gravado lógica e/ou dados de controle. O sistema de computador 100 pode vantajosamente incluir ou ser programado através de software apropriado para ler a lógica e/ou os dados de controle do componente de armazenamento de dados uma vez inserido nos dispositivos de recuperação de dados.

O sistema de computador 100 inclui uma tela 120 que é usada para exibir para um usuário do computador. Deve ser também observado que o sistema de computador 100 pode ser ligado a outros sistemas de computador 125a-c em uma rede ou rede de longa distância para fornecer acesso centralizado ao sistema de computador 100. Software para acessar e processar as seqüências de nucleotídeos ou de aminoácidos da invenção pode residir na memória principal 115 durante a execução.

Em alguns aspectos, o sistema de computador 100 também pode compreender um algoritmo de comparação de seqüência para comparar uma seqüência de ácidos nucléicos da invenção. O algoritmo e a(s) seqüência(s) podem ser armazenados em um meio legível por computador. Um "algoritmo de comparação de seqüência" refere-se a um ou mais programas que são implementados (local ou remotamente) no sistema de computador 100 para comparar uma seqüência de nucleotídeos com outras seqüências de nucleotídeos e/ou compostos armazenados dentro dos dispositivos de armazenamento de dados. Por exemplo, o algoritmo de comparação

de seqüência pode comparar as seqüências de nucleotídeos de uma seqüência exemplar SEQ ID Nº: 1, SEQ ID Nº: 2, armazenada em um meio legível por computador às seqüências de referência armazenadas em um meio legível para o computador identificar homologias ou motivos estruturais.

Os parâmetros usados podem ser adaptados com os algoritmos acima dependendo do comprimento da seqüência e do grau de homologia estudado. Em alguns aspectos, os parâmetros podem ser os parâmetros padrão usados pelos algoritmos na ausência de instruções do usuário. Figura 16 é um fluxograma que ilustra um aspecto de um processo 200 para comparar uma nova seqüência de nucleotídeo ou de proteína com um banco de dados de seqüências para determinar os níveis de homologia entre a nova seqüência e as seqüências no banco de dados. O banco de dados de seqüências pode ser um banco de dados privado armazenado dentro do sistema de computador 100, ou um banco de dados público como GENBANK que está disponível através da Internet. O processo 200 inicia em um estado de início 201 e depois se move para um estado 202 em que a nova seqüência a ser comparada é armazenada em uma memória em um sistema de computador 100. Como debatido acima, a memória pode ser qualquer tipo de memória, incluindo RAM ou um dispositivo de armazenamento interno.

O processo 200 depois se move para um estado 204 em que um banco de dados de seqüências é aberto para análise e comparação. O processo 200 depois se move para um estado 206 em que a primeira seqüência armazenada no banco de dados é lida em uma memória no computador. Uma comparação é depois executada em um estado 210 para determinar se a primeira seqüência é igual à segunda seqüência. É importante observar que esta etapa não é limitada a executar uma comparação exata entre a nova seqüência e a primeira seqüência no banco de dados. Métodos bem-conhecidos são conhecidos por aqueles versados na técnica para comparar duas seqüências de nucleotídeos ou de proteína, até mesmo se elas não são idênticas. Por exemplo, intervalos podem ser introduzidos em uma se-

qüência para elevar o nível de homologia entre as duas seqüências testadas. Os parâmetros que controlam, se são introduzidos intervalos ou outras características em uma seqüência durante a comparação normalmente, são introduzidos pelo usuário do sistema de computador.

Uma vez uma comparação das duas seqüências foi executada no estado 210, uma determinação é feita em um estado de decisão 210 se as duas seqüências forem as mesmas. Claro que, o termo "mesmas" não é limitado às seqüências que são absolutamente idênticas. Seqüências que estão dentro dos parâmetros de homologia inseridos pelo usuário serão marcadas como "mesmas" no processo 200. Se uma determinação for feita que as duas seqüências são as mesmas, o processo 200 se move para um estado 214 em que o nome da seqüência do banco de dados é exibido para o usuário. Este estado notifica o usuário que a seqüência com o nome exibido cumpre os requerimentos de homologia aos quais foram introduzidos. Uma vez o nome da seqüência armazenada é exibido para o usuário, o processo 200 se move pare um estado de decisão 218 em que uma determinação é feita se mais seqüências existem no banco de dados. Se mais nenhuma seqüência existe no banco de dados, então o processo 200 termina em um estado de finalização 220. Porém, se mais seqüências existem no banco de dados, então o processo 200 se move para um estado 224 em que um meio de apontamento é movido para a próxima seqüência no banco de dados de forma que ela pode ser comparada à nova seqüência. Desta maneira, a nova seqüência é alinhada e comparada com toda seqüência no banco de dados.

Deve ser observado que se uma determinação tivesse sido feita no estado de decisão 212 que as seqüências não eram homólogas, então o processo 200 moveria imediatamente para o estado de decisão 218 para determinar se qualquer outra seqüência estaria disponível no banco de dados para comparação. Conseqüentemente, um aspecto da invenção é um sistema de computador compreendendo um processador, um dispositivo de armazenamento de dados tendo nele armazenada uma seqüência de ácidos nucléicos da invenção e um comparador de seqüência para conduzir a com-

paração. O comparador de seqüência pode indicar um nível de homologia entre as seqüências comparadas ou pode identificar motivos estruturais, ou pode identificar motivos estruturais nas seqüências que são comparadas a estes códigos de ácido nucléico e códigos de polipeptídeos.

Figura 17 é um fluxograma que ilustra uma modalidade de um processo 250 em um computador para determinar se duas seqüências são homólogas. O processo 250 inicia em um estado de inicialização 252 e depois se move para um estado 254 em que uma primeira seqüência a ser comparada é armazenada em uma memória. A segunda seqüência a ser comparada é depois armazenada em uma memória em um estado 256. O processo 250 depois se move para um estado 260 em que o primeiro caractere na primeira seqüência é lido e depois para um estado 262 em que o primeiro caractere da segunda seqüência é lido. Deve ser entendido que se a seqüência for uma seqüência de nucleotídeos, então o caractere normalmente é A, T, C, G ou U. Se a seqüência for uma seqüência de proteína, então ela pode ser um código de aminoácido de letra minúscula de modo que a primeira e as seqüências da seqüência podem ser facilmente comparadas. Uma determinação é depois feita em um estado de decisão 264 se os dois caracteres forem os mesmos. Se eles forem os mesmos, então o processo 250 se move para um estado 268 em que são lidos os próximos caracteres nas primeira e segunda seqüências. Uma determinação é depois feita se os próximos caracteres forem os mesmos. Se eles forem, então o processo 250 continua este laço até que dois caracteres sejam desiguais. Se uma determinação for feita que os próximos dois caracteres não são os mesmos, o processo 250 se move para um estado de decisão 274 para determinar se há mais algum caractere de qualquer seqüência para ler. Se não há mais nenhum caractere para ler, então o processo 250 se move para um estado 276 em que o nível de homologia entre as primeira e segunda seqüências é exibido para o usuário. O nível de homologia é determinado calculando a proporção de caracteres entre as seqüências que foram as mesmas fora do número total de seqüências na primeira seqüência. Desse modo, se cada caractere em uma primeira seqüência de 100 nucleotídeos

alinhar com cada caractere em uma segunda seqüência, o nível de homologia é 100 %.

Alternativamente, o programa de computador pode comparar uma seqüência de referência a uma seqüência da invenção para determinar se as seqüências diferem-se em uma ou mais posições. O programa pode registrar o comprimento e identidade de nucleotídeos ou resíduos de aminoácido inseridos, excluídos ou substituídos com respeito à seqüência da referência ou da invenção. O programa de computador pode ser um programa que determina se uma seqüência de referência contém um polimorfismo de nucleotídeo simples (SNP) com respeito a uma seqüência da invenção, ou, se uma seqüência da invenção compreende um SNP de uma seqüência conhecida. Desse modo, em alguns aspectos, o programa de computador é um programa que identifica SNPs. O método pode ser implementado pelos sistemas de computador acima descritos e pelo método ilustrado na Figura 17. O método pode ser executado lendo uma seqüência da invenção e as seqüências de referência por meio do uso do programa de computador e identificando as diferenças com o programa de computador.

Em outros aspectos o sistema com base em computador compreende um identificador para identificar características dentro de um ácido nucléico ou polipeptídeo da invenção. Um "identificador" refere-se a um ou mais programas que identificam certas características dentro de uma seqüência de ácidos nucléicos. Por exemplo, um identificador pode compreender um programa que identifica uma estrutura de leitura aberta (ORF) em uma seqüência de ácidos nucléicos. Figura 18 é um fluxograma que ilustra um aspecto de um processo identificador 300 para detectar a presença de uma característica em uma seqüência. O processo 300 inicia em um estado de inicialização 302 e depois se move para um estado 304 em que uma primeira seqüência que será conferida nas características é armazenada a uma memória 115 no sistema de computador 100. O processo 300 depois se move para um estado 306 em que um banco de dados de características de seqüência é aberto. Um tal banco de dados inlcui uma lista de cada um dos atributos da característica junto ao nome da característica. Por exemplo, um

nome da característica pode ser "Códon de Iniciação" e o atributo seria "ATG." Outro exemplo pode ser o nome da característica "Caixa TAATAA" e o atributo da característica seria "TAATAA." Um exemplo de um tal banco de dados é produzido pela University of Wisconsin Genetics Computer Group. Alternativamente, as características podem ser motivos de polipeptídeos estruturais como hélices alfa, lâminas betas, ou motivos de polipeptídeos funcionais como sítios ativos enzimáticos, motivos de hélice-volta-hélice ou outros motivos conhecidos àqueles versados na técnica. Uma vez o banco de dados de características é aberto no estado 306, o processo 300 se move para um estado 308 em que a primeira característica é lida do banco de dados. Uma comparação do atributo da primeira característica com a primeira seqüência é depois feita em um estado 310. Uma determinação é depois feita em um estado de decisão 316 se o atributo da característica foi encontrado na primeira seqüência. Se o atributo foi encontrado, então o processo 300 se move para um estado 318 em que o nome da característica encontrada é exibido para o usuário. O processo 300 depois move-se para um estado de decisão 320 em que uma determinação é feita se as características existem no banco de dados. Se nenhuma mais característica existe, então o processo 300 termina em um estado de finalização 324. Porém, se mais características existem no banco de dados, então o processo 300 lê a próxima característica de seqüência em um estado 326 e retorna para o estado 310 em que o atributo da próxima característica é comparado com a primeira seqüência. Se o atributo da característica não é encontrado na primeira seqüência no estado de decisão 316, o processo 300 se move diretamente para o estado de decisão 320 para determinar se existe mais alguma característica no banco de dados. Desse modo, em um aspecto, a invenção fornece um programa de computador que identifica estruturas de leitura aberta (ORFs).

Uma seqüência de polipeptídeos ou de ácidos nucléicos da invenção pode ser armazenada e manipulada em uma variedade de programas processadores de dados em uma variedade de formatos. Por exemplo, uma seqüência pode ser armazenada como texto em um arquivo processa-

dor de textos, como MicrosoftWORD ou WORDPERFECT ou como um arquivo ASCII em uma variedade de programas de banco de dados familiares àqueles versados na técnica, como DB2, SYBASE ou ORACLE. Além disso, muitos programas de computação e bancos de dados como algoritmos de comparação de seqüência, identificadores ou fontes de seqüências de nucleotídeos de referência ou seqüências de polipeptídeos ser comparadas a uma seqüência de ácidos nucléicos da invenção podem ser usados. Os programas e bancos de dados usados para praticar a invenção incluem, mas não são limitados a: MacPattern (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 e 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.), SegFold (Molecular Simulations Inc.), o banco de dados de MDL Available Chemicals Directory, o banco de dados de MDL Drug Data Report, o banco de dados de Comprehensive Medicinal Chemistry, o banco de dados de World Drug Index de Derwent, o banco de dados de BioByte-MasterFile, o banco de dados de Genbank e o banco de dados de Genseqn. Muitos outros programas e bancos de dados podem ser óbvios àlguém versado na técnica dada a presente descrição.

Os motivos que podem ser detectados usando os programas acima incluem seqüências que codificam zíperes de leucina, motivos de hé-

lice-volta-hélice, sítios de glicosilação, sítios de ubiquitinação, hélices alfa e lâminas betas, seqüências sinalizadoras que codificam peptídeos de sinal que direcionam a secreção das proteínas codificadas, seqüências implicadas na regulação de transcrição como homeocaixas, alongamentos acídicos, sítios ativos enzimáticos, sítio de ligação aos substratos e sítios de clivagem enzimática.

HIBRIDAÇÃO DE ÁCIDOS NUCLÉICOS

A invenção fornece ácidos nucléicos isolados ou recombinantes que hibridizam sob condições rigorosas com uma seqüência exemplar da invenção, por exemplo, uma seqüência como exposta na SEQ ID Nº: 1, ou um ácido nucléico que codifica um polipeptídeo compreendendo uma seqüência como exposto na SEQ ID Nº: 2. As condições rigorosas podem ser condições altamente rigorosas, condições rigorosas médias, condições rigorosas baixas, incluindo as condições de severidade alta e reduzidas aqui descritas. Em modalidades alternativas, os ácidos nucléicos da invenção como definidos por sua capacidade de hibridar-se sob condições rigorosas podem ser entre cerca de cinco resíduos e o comprimento total da molécula da SEQ ID Nº: 1; por exemplo, eles podem ser pelo menos 5, 10, 15, 20, 25, 30, 35, 40, 50, 55, 60, 65, 70, 75, 80, 90, 100, 150, 200, 250, 300, 350, 400 resíduos no comprimento. Ácidos nucléicos mais curtos que comprimento total também estão incluídos. Estes ácidos nucléicos são úteis como, por exemplo, sondas de hibridação, sondas de marcação, sondas de oligonucleotídeo de PCR, seqüências de iRNA, de anti-sentido ou codificando os peptídeos de ligação de anticorpo (epítopo), motivos, sítios ativos e similares.

Em um aspecto, os ácidos nucléicos da invenção são definidos por sua capacidade de hibridar sob condições de alta severidade compreendendo de cerca de 50 % de formamida a cerca de 37°C a 42°C. Em um aspecto, os ácidos nucléicos da invenção são definidos por sua capacidade de hibridar sob condições de severidade reduzida compreendendo cerca de 35 % a 25 % de formamida a cerca de 30°C a 35°C. Alternativamente, os ácidos nucléicos da invenção são definidos por sua capacidade de hibridar sob

condições de alta severidade compreendendo a 42°C em 50 % de formamida, 5X SSPE, 0,3 % de SDS, e um ácido nucléico de bloqueio de seqüência repetitiva, como cot-1 ou DNA de esperma de salmão (por exemplo, 200 n/ml de DNA de esperma de salmão cisalhado e desnaturado). Em um aspecto, os ácidos nucléicos da invenção são definidos por sua capacidade de hibridar sob condições de severidade reduzida compreendendo 35 % de formamida em uma temperatura reduzida de 35°C.

Seguindo hibridação, o filtro pode ser lavado com 6X SSC, 0,5 % de SDS a 50°C. É considerado que estas condições são "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 em 30 % de formamida. Um exemplo específico de condições de hibridação de "baixa severidade" é quando a hibridação acima é conduzida em 10 % de formamida.

A faixa de temperatura que corresponde a um nível particular de severidade pode ser também estreitada calculando a razão de purina para pirimidina do ácido nucléico de interesse e ajustando a temperatura conseqüentemente. Os ácidos nucléicos da invenção também são definidos por sua capacidade de hibridar sob condições de severidade alta, média e baixa como exposto em Ausubel e Sambrook. Variações nas faixas e condições acima são bem-conhecidas na técnica. As condições de hibridação são também debatidas abaixo.

SONDAS DE OLIGONUCLEOTÍDEOS E MÉTODOS PARA USÁ-LAS

A invenção também fornece sondas de ácido nucléico para identificar os ácidos nucléicos codificando um polipeptídeo com uma atividade de fitase. Em um aspecto, a sonda compreende pelo menos 10 bases sucessivas de uma seqüência como exposta na SEQ ID Nº: 1. Alternativamente, uma sonda da invenção pode ser pelo menos cerca de 5, 6, 7, 8 ou 9 a cerca de 40, cerca de 10 a 50, cerca de 20 a 60, cerca de 30 a 70, bases sucessivas de uma seqüência como exposto na SEQ ID Nº: 1. As sondas identificam um ácido nucléico por ligação ou hibridação. As sondas podem ser usadas em arranjos da invenção, ver debate abaixo, incluindo, por exemplo,

arranjos capilares. As sondas da invenção também podem ser usadas para isolar outros ácidos nucléicos ou polipeptídeos.

As sondas da invenção podem ser usadas para determinar se uma amostra biológica, como uma amostra do solo, contém um organismo tendo uma seqüência de ácidos nucléicos da invenção ou um organismo dos quais o ácido nucléico foi obtido. Em tais procedimentos, uma amostra biológica potencialmente abriga o organismo do qual o ácido nucléico foi isolado é obtido e os ácidos nucléicos da amostra são obtidos. Os ácidos nucléicos são contatados com a sonda sob condições que permitem a sonda especificamente hibridar com qualquer seqüência complementar presente na amostra. Onde necessário, as condições que permitem a sonda especificamente hibridar com as seqüências complementares podem ser determinadas colocando a sonda em contato com as seqüências complementares das amostras conhecidas possuir a seqüência complementar, como também seqüências de controle que não possuem a seqüência complementar. Condições de hibridação, 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 as condições que permitem a sonda especificamente hibridar com os ácidos nucléicos complementares (ver debate sobre condições de hibridação específicas).

Se a amostra contém o organismo do qual o ácido nucléico foi isolado, a hibridação específica da sonda é depois detectada. A hibridação pode ser detectada marcando a sonda com um agente detectável como um isótopo radioativo, um corante fluorescente ou uma enzima capaz de catalisar a formação de um produto detectável. Muitos métodos de usar as sondas marcadas para detectar a presença de ácidos nucléicos complementares em uma amostra são familiares àqueles versados na técnica. Estes incluem Southern blot, Northern blot, procedimentos de hibridação de colônia e manchas de ponto. Os protocolos para cada um destes procedimentos são fornecidos em Ausubel e Sambrook.

Alternativamente, mais de uma sonda (pelo menos uma destas é capaz de especificamente hibridar com qualquer seqüência complementar

que está presente na amostra de ácido nucléico), pode ser usada em uma reação de amplificação para determinar se a amostra contém um organismo contendo uma seqüência de ácidos nucléicos da invenção (por exemplo, um organismo do qual o ácido nucléico foi isolado). Em um aspecto, as sondas compreendem oligonucleotídeos. Em um aspecto, a reação de amplificação pode compreender uma reação de PCR. Protocolos de PCR são descritos em Ausubel e Sambrook (ver debate sobre reações de amplificação). Em tais procedimentos, os ácidos nucléicos são contatados na amostra com as sondas, a reação de amplificação é executada e qualquer produto de amplificação resultante é detectado. O produto de amplificação pode ser detectado executando eletroforese em gel nos produtos de reação e manchando o gel com um intercalador como brometo de etídeo. Alternativamente, uma ou mais das sondas podem ser marcadas com um isótopo radioativo e a presença de um produto de amplificação radioativo pode ser detectada através de auto-radiografia após eletroforese em gel.

Sondas derivadas das seqüências próximas das extremidades 3' ou 5' também podem ser usadas de uma seqüência de ácidos nucléicos da invenção em procedimentos de passeio pelos cromossomos para identificar clones que contêm seqüências adicionais, por exemplo, genômicas. Tais métodos permitem o isolamento dos genes que codificam proteínas adicionais de interesse do organismo hospedeiro.

Em um aspecto, as seqüências de ácidos nucléicos da invenção são usadas como sondas para identificar e isolar ácidos nucléicos relacionados. Em alguns aspectos, os ácidos nucléicos relacionados assim identificados podem ser cDNAs ou os DNAs genômicos de organismos diferentes daquele que o ácido nucléico da invenção foi primeiro isolado. Em tais procedimentos, uma amostra de ácido nucléico é contatada com a sonda sob condições que permitem a sonda especificamente hibridar com as seqüências relacionadas. Hibridação da sonda com os ácidos nucléicos do organismo relacionado é depois detectada usando quaisquer dos métodos acima descritos.

Em reações de hibridação de ácido nucléico, as condições usa-

das para alcançar um nível particular de severidade variarão, dependendo da natureza dos ácidos nucléicos que são hibridados. Por exemplo, o comprimento, grau de complementaridade, composição da seqüência de nucleotídeos (por exemplo, GC V. teor de AT) e tipo de ácido nucléico (por exemplo, RNA V. DNA) das regiões de hibridação dos ácidos nucléicos podem ser considerados na seleção das condições de hibridação. Uma consideração adicional é se um dos ácidos nucléicos for imobilizado, por exemplo, em um filtro. A hibridação pode ser realizada sob condições de baixa severidade, severidade moderada ou alta severidade. Uma membrana de polímero contendo os ácidos nucléicos desnaturados imobilizados é primeiro pré-hibridada durante 30 minutos a 45°C em uma solução que consiste em 0,9 M NaCl como um exemplo de hibridação de ácido nucléico, 50 mM NaH2PO4, pH 7,0, 5,0 mM Na2EDTA, 0,5 % de SDS, 10X Denhardt, e 0,5 mg/ml de ácido polirriboadenílico. Aproximadamente 2 X 107 cpm (atividade específica 4-9 X 108 cpm/ug) de sonda de oligonucleotídeo marcada na extremidade com ^{32}P é depois adicionada à solução. Após 12-16 horas de incubação, a membrana é lavada durante 30 minutos em temperatura ambiente (TA) em 1X SET (150 mM NaCl, 20 mM cloridrato de Tris, pH 7,8, 1 mM Na2EDTA) contendo 0,5 % de SDS, seguido por uma lavagem de 30 minutos em 1X SET fresco a Tm-10°C para a sonda de oligonucleotídeo. A membrana é depois exposta à película auto-radiográfica para detecção de sinais de hibridação.

Variando a severidade das condições de hibridação usadas para identificar os ácidos nucléicos, como cDNAs ou DNAs, genômicos que hibridizam com a sonda detectável, os ácidos nucléicos tendo níveis diferentes de homologia quanto à sonda podem ser identificados e isolados. A severidade pode ser variada conduzindo a hibridação em temperaturas variadas abaixo das temperaturas de derretimento das sondas. A temperatura de derretimento, Tm, é a temperatura (sob força iônica definida e pH) na qual 50 % da seqüência alvo hibridizam com uma sonda perfeitamente complementar. Condições muito rigorosas são selecionadas para ser iguais a ou cerca de 5°C inferior à Tm para uma sonda particular. A temperatura de derretimento

da sonda pode ser calculada usando as fórmulas exemplares a seguir. Para sondas entre 14 e 70 nucleotídeos no comprimento a temperatura de derretimento (Tm) é calculada usando a fórmula: Tm=81,5+16,6(log[Na+])+0,41(fração G+C)-(600/N) onde N é o comprimento da sonda.

Se a hibridação for realizada em uma solução contendo formamida, a temperatura de derretimento pode ser calculada usando a equação: Tm=81,5+16,6(log[Na+])+0,41(fração G+C)-(0,63 % de formamida)-(600/N) onde N é o comprimento da sonda. Pré-hibridação pode ser realizada em 6X SSC, 5X o reagente de Denhardt, 0,5 % de SDS, 100 μg de DNA de esperma de salmão fragmentado desnaturado ou 6X SSC, 5X o reagente de Denhardt, 0,5 % de SDS, 100 μg de DNA de esperma de salmão fragmentado desnaturado, 50 % de formamida. Fórmulas para SSC e outras soluções de Denhardt estão listadas, por exemplo, em Sambrook.

Hibridação é conduzida adicionando a sonda detectável às soluções de pré-hibridação acima listadas. Onde a sonda compreender o DNA bifilamentado, ele é desnaturado antes da adição à solução de hibridação. O filtro é contatado com a solução de hibridação durante um período suficiente de tempo para permitir a sonda hibridar com os cDNAs ou DNAs genômicos contendo as seqüências complementares a estes ou homólogas a estes. Para sondas maiores de 200 nucleotídeos no comprimento, a hibridação pode ser realizada a 15-25°C abaixo da Tm. Para sondas mais curtas, como sondas de oligonucleotídeo, a hibridação pode ser conduzida a 5-10°C abaixo da Tm. Em um aspecto, hibridações em 6X SSC são conduzidas a cerca de 68°C. Em um aspecto, hibridações em 50 % de formamida contendo soluções são conduzidas a cerca de 42°C. Todas as hibridações anteriores podem ser consideradas estar sob condições de alta severidade.

Seguindo a hibridação, o filtro é lavado para remover qualquer sonda detectável não-especificamente ligada. A severidade usada para lavar os filtros pode ser variada, dependendo também da natureza dos ácidos nucléicos que são hibridados, do comprimento dos ácidos nucléicos que são hibridados, do grau de complementaridade, da composição da seqüência de

nucleotídeos (por exemplo, GC V. teor de AT) e do tipo de ácido nucléico (por exemplo, RNA V. DNA). Exemplos de lavagens de condição de severidade progressivamente mais altas são como segue: 2X SSC, 0,1 % de SDS em temperatura ambiente durante 15 minutos (baixa severidade); 0,1X SSC, 0,5 % de SDS em temperatura ambiente durante 30 minutos a 1 hora (severidade moderada); 0,1X SSC, 0,5 % de SDS durante 15 a 30 minutos entre a temperatura de hibridação e 68°C (severidade alta); e 0,15M NaCl durante 15 minutos a 72°C (severidade muito alta). Uma lavagem final de baixa severidade pode ser conduzida em 0,1X SSC em temperatura ambiente. Os exemplos acima são meramente ilustrativos de um conjunto de condições que podem ser usadas para lavar os filtros. Aquele versado na técnica saberia que há numerosas receitas para lavagens de diferentes severidades.

Os ácidos nucléicos que hibridizaram com a sonda podem ser identificados por auto-radiografia ou outras técnicas convencionais. O procedimento acima pode ser modificado para identificar os ácidos nucléicos tendo níveis decrescentes de homologia à seqüência de sonda. Por exemplo, para obter ácidos nucléicos de homologia decrescente à sonda detectável, podem ser usadas condições menos rigorosas. Por exemplo, a temperatura de hibridação pode ser diminuída em incrementos de 5°C de 68°C a 42°C em um tampão de hibridação tendo uma concentração de Na+ de cerca de 1M. Seguindo a hibridação, o filtro pode ser lavado com 2X SSC, 0,5 % de SDS na temperatura de hibridação. É considerado que estas condições são condições "moderadas" acima de 50°C e condições "baixas" abaixo de 50°C. Um exemplo de condições de hibridação "moderadas" é quando a hibridação acima é conduzida a 55°C. Um exemplo de condições de hibridação de "baixa severidade" é quando a hibridação acima é conduzida a 45°C.

Alternativamente, a hibridação pode ser realizada em tampões, como 6X SSC, contendo formamida em uma temperatura de 42°C. Neste caso, a concentração de formamida no tampão de hibridação pode ser reduzida em incrementos de 5 % de 50 % a 0 % para identificar clones tendo níveis decrescentes de homologia à sonda. Seguindo a hibridação, o filtro pode ser lavado com 6X SSC, 0,5 % de SDS a 50°C. É considerado que

estas condições são 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 "baixa severidade" é quando a hibridação acima é conduzida a 10 % de formamida.

Estas sondas e métodos da invenção podem ser usados para isolar ácidos nucléicos tendo uma seqüência com pelo menos cerca de 99 %, 98 %, 97 %, pelo menos 95 %, pelo menos 90 %, pelo menos 85 %, pelo menos 80 %, pelo menos 75 %, pelo menos 70 %, pelo menos 65 %, pelo menos 60 %, pelo menos 55 %, ou pelo menos 50 % de homologia a uma seqüência de ácidos nucléicos da invenção compreendendo pelo menos cerca de 10, 15, 20, 25, 30, 35, 40, 50, 75, 100, 150, 200, 250, 300, 350, 400 ou 500 bases sucessivas destes, e as seqüências complementares a esta. A homologia pode ser medida usando um algoritmo de alinhamento, como debatido aqui. Por exemplo, os polinucleotídeos homólogos podem ter uma seqüência de codificação que é uma variante alélica de ocorrência natural de uma das seqüências de codificação aqui descritas. Tais variantes alélicas podem ter uma substituição, deleção ou adição de um ou mais nucleotídeos quando comparados a um ácido nucléico da invenção.

Adicionalmente, as sondas e métodos da invenção podem ser usados para isolar ácidos nucléicos codificando polipeptídeos tendo pelo menos cerca de 99 %, pelo menos 95 %, pelo menos 90 %, pelo menos 85 %, pelo menos 80 %, pelo menos 75 %, pelo menos 70 %, pelo menos 65 %, pelo menos 60 %, pelo menos 55 %, ou pelo menos 50 % de identidade de seqüência (homologia) a um polipeptídeo da invenção compreendendo pelo menos 5, 10, 15, 20, 25, 30, 35, 40, 50, 75, 100 ou 150 aminoácidos sucessivos desta como determinado usando um algoritmo de alinhamento de seqüência (por exemplo, como algoritmo de FASTA versão 3.0t78 com os parâmetros padrão, ou um programa BLAST 2.2.2 com ajustes exemplares como aqui expostos).

INIBIÇÃO DE EXPRESSÃO DE UMA FITASE

A invenção também fornece para ácidos nucléicos complementares (por exemplo, seqüências anti-sentido) às seqüências de ácidos nucléicos da invenção. Seqüências anti-sentido são capazes de inibir o transporte, junção ou transcrição dos genes de codificação da fitase. A inibição pode ser realizada através do alvejamento do DNA genômico ou RNA mensageiro. A transcrição ou função do ácido nucléico alvejado pode ser inibida, por exemplo, por hibridação e/ou clivagem. Um conjunto de inibidores particularmente úteis fornecido pela presente invenção inclui oligonucleotídeos que são capazes ou de ligar gene ou mensagem de fitase, em ambos os casos prevenindo ou inibindo a produção ou função de enzima de fitase. A associação pode ser, entretanto hibridação de seqüência específica. Outra classe útil de inibidores inclui oligonucleotídeos que causam inativação ou clivagem de mensagem de fitase. O oligonucleotídeo pode ter atividade da enzima que causa tal clivagem, como ribozimas. O oligonucleotídeo pode ser quimicamente modificado ou conjugado com uma enzima ou composição capaz de clivar o ácido nucléico complementar. Pode-se triar um "pool" de muitos tais oligonucleotídeos diferentes para aqueles com a atividade desejada.

OLIGONUCLEOTÍDEOS ANTI-SENTIDO

A invenção fornece oligonucleotídeos anti-sentido capazes de ligar mensagem de fitase que pode inibir atividade de fitase mediante alvejamento do mRNA. Estratégias para designar oligonucleotídeos anti-sentido são bem-conhecidas na literatura científica e de patente, e o versado na técnica pode designar tais oligonucleotídeos de fitase usando os novos reagentes da invenção. Por exemplo, protocolos de passeio pelo gene/mapeamento de RNA para triar por oligonucleotídeos anti-sentido eficazes são bem-conhecidos na técnica, ver, por exemplo, Ho (2000) Methods Enzymol. 314:168-183, descrevendo um ensaio de mapeamento de RNA que é com base nas técnicas moleculares padrão para prover um método fácil e seguro para seleção de seqüência anti-sentido potente. Também ver Smith (2000) Euro. J. Pharm. Sci. 11:191-198.

Ácidos nucléicos de ocorrência natural são usados como oligonucleotídeos anti-sentido. Os oligonucleotídeos anti-sentido podem ser de qualquer comprimento; por exemplo, em aspectos alternativos, os oligonucleotídeos anti-sentido são entre cerca de 5 a 100, cerca de 10 a 80, cerca de 15 a 60, cerca de 18 a 40. O comprimento ideal pode ser determinado através de triagem rotineira. Os oligonucleotídeos anti-sentido podem estar presentes em qualquer concentração. A concentração ideal pode ser determinada através de triagem rotineira. Uma ampla variedade de análogos de nucleotídeo e de ácido nucléico sintéticos, de ocorrência natural e não-natural é conhecida que pode dirigir-se a este problema potencial. Por exemplo, ácidos nucléicos de peptídeo (PNAs) contendo cadeias principais não-iônicas, como unidades de N-(2-aminoetil)glicina podem ser usados. Também podem ser usados oligonucleotídeos anti-sentido tendo ligações de fosforotioato, como descrito em WO 97/03211; WO 96/39154; Mata (1997) Toxicol Appl Pharmacol 144:189-197; Antisense Therapeutics, ed. Agarwal (Humana Press, Totowa, N. J., 1996). Oligonucleotídeos anti-sentido tendo análogos de cadeia principal de DNA sintético fornecidos pela invenção também podem incluir ácidos nucléicos de fosforo-ditioato, metilfosfonato, fosforamidato, fosfotriéster de alquila, sulfamato, 3'-tioacetal, metileno(metilimino), 3'-N-carbamato e carbamato de morfolino, como descrito acima.

Metodologia química combinatória pode ser usada para criar vastos números de oligonucleotídeos que podem ser triados rapidamente por oligonucleotídeos específicos tendo afinidades e especificidades de ligação apropriadas para qualquer alvo, como as seqüências de fitase sense e anti-sentido da invenção (ver, por exemplo, Gold (1995) J. of Biol. Chem. 270:13581-13584).

RIBOZIMAS INIBITÓRIAS

A invenção fornece ribozimas capazes de ligar mensagem de fitase que pode inibir atividade da enzima de fitase mediante alvejamento de mRNA. Estratégias para designar ribozimas e selecionar a seqüência anti-sentido fitase-específica por alvejamento são bem descritas na literatura ci-

entífica e de patente, e o versado pode designar tais ribozimas usando os novos reagentes da invenção. Ribozimas agem ligando a um RNA alvo através da porção ligadora de RNA alvo de uma ribozima que é mantida em proximidade íntima a uma porção enzimática do RNA que cliva o RNA alvo. Desse modo, a ribozima reconhece e liga um RNA alvo através de emparelhamento de bases complementares, e uma vez ligada ao sítio correto, age enzimaticamente para clivar e inativar o RNA alvo. A clivagem de um RNA alvo em uma tal maneira destruirá sua capacidade de direcionar a síntese de uma proteína codificada se a clivagem ocorrer na seqüência de codificação. Após uma ribozima ter sido ligada e clivado seu RNA alvo, ela é tipicamente liberada daquele RNA e assim pode ligar e clivar novos alvos repetidamente.

Em algumas circunstâncias, a natureza enzimática de uma ribozima pode ser vantajosa em outras tecnologias, como tecnologia anti-sentido (onde uma molécula de ácido nucléico simplesmente liga a um alvo de ácido nucléico para bloquear sua transcrição, tradução ou associação à outra molécula) como a concentração eficaz de ribozima necessária para realizar um tratamento terapêutico pode ser mais baixa que a de um oligonucleotídeo anti-sentido. Esta vantagem potencial reflete a capacidade da ribozima de agir enzimaticamente. Desse modo, uma única molécula de ribozima pode clivar muitas moléculas de RNA alvo. Além disso, uma ribozima é tipicamente um inibidor altamente específico, com a especificidade de inibição que não só depende do mecanismo de emparelhamento de bases de ligação, mas também do mecanismo pelo qual a molécula inibe a expressão do RNA ao qual ela liga. Ou seja, a inibição é causada por clivagem do RNA alvo e assim a especificidade é definida como a razão da taxa de clivagem do RNA alvejado para a taxa de clivagem de RNA não-alvejado. Este mecanismo de clivagem é dependente dos fatores adicionais além daqueles envolvidos no emparelhamento de bases. Desse modo, a especificidade de ação de uma ribozima pode ser maior que a do oligonucleotídeo anti-sentido que liga o mesmo sítio de RNA. A molécula de RNA de ribozima enzimática pode ser formada em um motivo de cabeça de martelo, mas também pode ser formada no motivo de um grampo, vírus de hepatite delta, grupo íntron I

ou RNA semelhante a RNaseP (em associação com uma seqüência guia de RNA). Exemplos de tais motivos de cabeça de martelo são descritos por Rossi (1992) Aids Research and Human Retrovirus 8:183; motivos de grampo por Hampel (1989) Biochemistry 28:4929, e Hampel (1990) Nuc. Acids Res. 18:299; o motivo de vírus de hepatite delta por Perrotta (1992) Biochemistry 31:16; o motivo de RNaseP por Guerrier-Takada (1983) Cell 35:849; e o grupo de íntron I por Cech Patente U.S. Nº 4.987.071. A recitação destes motivos específicos não é intencionada ser limitativa; aqueles versados na técnica reconhecerão que uma molécula de RNA enzimática desta invenção tem um sítio de ligação ao substrato específico complementar a uma ou mais das regiões de RNA do gene alvo, e tem seqüência de nucleotídeos dentro ou adjacente ao sítio de ligação ao substrato que dá uma atividade de clivagem de RNA à molécula.

MODIFICAÇÃO DOS ÁCIDOS NUCLÉICOS

A invenção fornece métodos de variantes geradoras dos ácidos nucléicos da invenção, por exemplo, aqueles que codificam uma enzima de fitase. Estes métodos podem ser repetidos ou usados em várias combinações para gerar enzimas de fitase tendo uma atividade alterada ou diferente ou uma estabilidade alterada ou diferente daquela de uma fitase codificada pelo ácido nucléico modelo. Estes métodos também podem ser repetidos ou usados em várias combinações, por exemplo, para gerar variações na expressão de gene/mensagem, tradução de mensagem ou estabilidade de mensagem. Em outro aspecto, a composição genética de uma célula é alterada, por exemplo, por modificação de um gene homólogo *ex vivo*, seguido por sua reinserção na célula.

Um ácido nucléico da invenção pode ser alterado por quaisquer meios. Por exemplo, métodos aleatórios ou estocásticos, ou, métodos não-estocásticos ou de "evolução direcionada".

Os métodos para mutação aleatória de genes são bem-conhecidos na técnica, ver, por exemplo, Patente U. S. Nº 5.830.696. Por exemplo, mutagênios podem ser usados para transformar um gene randomicamente. Mutagênios incluem, por exemplo, luz ultravioleta ou irradiação

faixa, ou um mutágeno químico, por exemplo, mitomicina, ácido nitroso, psoralens fotoativados, sozinhos ou em combinação, para induzir rompimentos de DNA amenizáveis por reparo por meio de recombinação. Por exemplo, outros mutagênios químicos incluem bissulfeto de sódio, ácido nitroso, hidroxilamina, hidrazina ou ácido fórmico. Outros mutagênios são análogos de precursores de nucleotídeo, por exemplo, nitrosoguanidina, 5-bromouracila, 2-aminopurina ou acridina. Estes agentes podem ser adicionados a uma reação de PCR no lugar do precursor de nucleotídeo assim transformando a seqüência. Agentes de intercalação como proflavina, acriflavina, quinacrina e similares podem ser também usados.

Qualquer técnica em biologia molecular pode ser usada, por exemplo, mutagênese de PCR aleatória, ver, por exemplo, Rice (1992) Proc. Natl. Acad. Sci. USA 89:5467-5471; ou, mutagênese de cassete múltiplo combinatória, ver, por exemplo, Crameri (1995) Biotechniques 18:194-196. Alternativamente, ácidos nucléicos, por exemplo, genes, podem ser remontados após fragmentação aleatória ou "estocástica," ver, por exemplo, Patentes U.S. Nºs 6.291.242; 6.287.862; 6.287.861; 5.955.358; 5.830.721; 5.824.514; 5.811.238; 5.605.793. Em aspectos alternativos, as modificações, adições ou deleções são introduzidas por PCR propensa a erro, embaralhamento, mutagênese dirigida a oligonucleotídeo, PCR de montagem, mutagênese de PCR sexual, mutagênese *in vivo*, mutagênese de cassete, mutagênese de conjunto recursiva, mutagênese de conjunto exponencial, mutagênese sítio específica, remontagem genética, mutagênese saturada no sítio genético (GSSM), remontagem de ligação sintética (SLR), recombinação, recombinação de seqüência recursiva, mutagênese de DNA modificado por fosfotioato, mutagênese de modelo contendo uracila, mutagênese dúplex com intervalos, mutagênese de reparo de equívoco de ponto, mutagênese de cepa hospedeira deficiente de reparo, mutagênese química, mutagênese radiogênica, mutagênese de deleção, mutagênese de restrição-seleção, mutagênese de restrição-purificação, síntese de gene artificial, mutagênese de conjunto, criação multímera de ácido nucléico quimérica, e/ou uma combinação destes e outros métodos.

As publicações a seguir descrevem uma variedade de procedimentos e/ou métodos de recombinação recursiva 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 8:724733; 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 shu-

ffling by random fragmentation and reassembly: *In vitro* recombination for molecular evolution". Proc. Natl. Aca. Sci. USA 91:10747-10751.

Métodos mutacionais de gerar diversidade incluem, por exemplo, mutagênese sítio direcionada (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)); mutagênese usando modelos contendo uracila (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); mutagênese oligonucleotídeo-dirigida (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); mutagênese de DNA fosforotioate-modificado (Taylor et al. (1985) "The use of phosphorothioate-modified DNA in restriction enzyme reactions to prepare nicked DNA" Nucl. Acids Res. 13: 87498764; 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 res-

triction 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); mutagênese usando DNA dúplex com intervalos (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 usados nos métodos da invenção incluem reparo de equívoco de ponto (Kramer (1984) "Point Mismatch Repair" Cell 38:879-887), mutagênese usando cepas de hospedeiro deficientes de reparo (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), mutagênese de deleção (Eghtedarzadeh (1986) "Use of oligonucleotides to generate large deletions" Nucl. Acids Res. 14: 5115), restrição-seleção e restrição-seleção e restrição-purificação (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), mutagênese por síntese de gene 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), reparo de rompimento de filamento duplo (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 sobre muitos destes métodos acima podem ser encontrados em Methods in Enzymology Volume 154, que também descreve controles úteis para diagnosticar problemas com vários métodos de mutagênese. Ver também Patente U. S. Nos. 5.605.793 de Stemmer (fev. 25, 1997), "Methods for *In vitro* Recombination"; Pat. U. S. No. 5.811.238 de Stemmer et al. (set. 22, 1998) "Methods for Generating Polynucleotides having Desired Characteristics by Iterative Selection and Recombination"; Pat. U. S. No. 5.830.721 de Stemmer et al. (nov. 3, 1998), "DNA Mutagenesis by Random Fragmentation and Reassembly"; Pat. U. S. No. 5.834.252 de Stemmer, et al. (nov. 10, 1998) "End-Complementary Polymerase Reaction"; Pat. U. S. No. 5.837.458 de Minshull, et al. (nov. 17, 1998), "Methods and Compositions for Cellular and Metabolic Engineering"; WO 95/22625, Stemmer e Crameri, "Mutagenesis by Random Fragmentation and Reassembly"; WO 96/33207 por Stemmer e Lipschutz "End Complementary Polymerase Chain Reaction"; WO 97/20078 por Stemmer e Crameri "Methods for Generating Polynucleotides having Desired Characteristics by Iterative Selection and Recombination"; WO 97/35966 por Minshull e Stemmer, "Methods and Compositions for Cellular and Metabolic Engineering"; WO 99/41402 por Punnonen et al. "Targeting of Genetic Vaccine Vectors"; WO 99/41383 por Punnonen et al. "Antigen Library Immmunization"; WO 99/41369 por Punnonen et al. "Genetic Vaccine Vector Engineering"; WO 99/41368 por Punnonen et al. "Optimization of Immunomodulatory Properties of Genetic Vaccines"; EP 752008 por Stemmer e Crameri, "DNA Mutagenesis by Random Fragmentation and Reassembly," EP 0932670 por Stemmer "Evolving Cellular DNA Uptake by Recursive Se-

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

Certos pedidos de patente U. S. fornecem detalhes adicionais com relação aos vários métodos de gerar diversidade adicionais, incluindo "SHUFFLING OF CODON ALTERED GENES" por Patten et al. depositado em 28 de setembro, 1999, (U. S. Ser. No. 09/407.800); "EVOLUTION OF WHOLE CELLS AND ORGANISMS BY RECURSIVE SEQUENCE RECOMBINATION" por del Cardayre et al., depositado em 15 de julho, 1998 (U. S. Ser. No. 09/166.188) e em 15 de julho, 1999 (U. S. Ser. No. 09/354.922); "OLIGONUCLEOTIDE MEDIATED NUCLEIC ACID RECOMBINATION" por Crameri et al., depositado em 28 de setembro, 1999 (U. S. Ser. No. 09/408.392) e "OLIGONUCLEOTIDE MEDIATED NUCLEIC ACID RECOMBINATION" por Crameri et al., depositado em 18 de janeiro, 2000 (PCT/US00/01203); "USE OF CODON-VARIED OLIGONUCLEOTIDE SYNTHESIS FOR SYNTHETIC SHUFFLING" por Welch et al., depositado em 28 de setembro, 1999 (U. S. Ser. No. 09/408.393), "METHODS FOR MAKING CHARACTER STRINGS, POLYNUCLEOTIDES &

POLYPEPTIDES HAVING DESIRED CHARACTERISTICS" por Selifonov et al., depositado em 18 de janeiro, 2000, (PCT/US00/01202) e, por exemplo "METHODS FOR MAKING CHARACTER STRINGS, POLYNUCLEOTIDES & POLYPEPTIDES HAVING DESIRED CHARACTERISTICS" por Selifonov et al., depositado em 18 de julho, 2000 (U. S. Ser. No. 09/618.579), "METHODS OF POPULATING DATA STRUCTURES FOR USE IN EVOLUTIONARY SIMULATIONS" por Selifonov e Stemmer, depositado em 18 de janeiro, 2000 (PCT/US00/01138); e "SINGLE-STRANDED NUCLEIC ACID TEMPLATE-MEDIATED RECOMBINATION AND NUCLEIC ACID FRAGMENT ISOLATION" por Affholter, depositado em 6 de setembro, 2000 (U. S. Ser. No. 09/656.549).

Métodos não-estocásticos ou de "evolução dirigida" incluem, por exemplo, mutagênese de saturação (GSSM), remontagem de ligação sintética (SLR), ou uma combinação destes, são usados para modificar os ácidos nucléicos da invenção para gerar fitases com propriedades novas ou alteradas (por exemplo, atividade sob condições altamente acídicas ou alcalinas, temperaturas altas e outras). Os polipeptídeos codificados pelos ácidos nucléicos modificados podem ser triados por uma atividade antes de testar uma atividade de fitase ou outra. Qualquer modalidade de prova ou protocolo pode ser usada, por exemplo, usando uma plataforma de arranjo capilar. Ver, por exemplo, Patente U. S. Nos. 6.280.926; 5.939.250.

MUTAGÊNESE DE SATURAÇÃO OU GSSM

Em um aspecto da invenção, modificação não-estocástica de gene, um "processo de evolução direcionada", é usado para gerar fitases com propriedades novas ou alteradas. Variações deste método foram denominadas "mutagênese de saturação de sítio genético," "mutagênese de saturação de sítio," "mutagênese de saturação" ou simplesmente "GSSM". Ela pode ser usada em combinação com outros processos de metagênese. Ver, por exemplo, Patente U. S. Nos. 6.171.820; 6.238.884. Em um aspecto, GSSM compreende fornecer um polinucleotídeo modelo e uma pluralidade de oligonucleotídeos, em que cada oligonucleotídeo compreende uma seqüência homóloga ao polinucleotídeo modelo, assim alvejando uma seqüên-

cia específica do polinucleotídeo modelo, e uma seqüência que é uma variante do gene homólogo; gerando polinucleotídeos das progênies compreendendo variações de seqüência não-estocástica mediante replicação do polinucleotídeo modelo com os oligonucleotídeos, assim gerando polinucleotídeos compreendendo variações de seqüência de gene homólogas.

REMONTAGEM DE LIGAÇÃO SINTÉTICA (SLR)

A invenção fornece um sistema de modificação não-estocástica de gene denominado "remontagem de ligação sintética," ou simplesmente "SLR," um "processo de evolução direcionada," para gerar fitases com propriedades novas ou alteradas. SLR é um método de ligar fragmentos de oligonucleotídeos entre si não-estocasticamente. Este método difere do embaralhamento estocástico de oligonucleotídeos em que os blocos de construção de ácido nucléico não são embaralhados, concatenados ou quimerizados fortuitamente, mas de preferência são montados não-estocasticamente. Ver, por exemplo, Pedido de Patente U. S. Série No. (USSN) 09/332.835 intitulado "Remontagem de Ligação Sintética em Evolução Direta" e depositado em 14 de junho de 1999 ("USSN 09/332.835"). Em um aspecto, SLR compreende as etapas a seguir: (a) fornecer um polinucleotídeo modelo, em que o polinucleotídeo modelo compreende seqüência codificando um gene homólogo; (b) fornecer uma pluralidade de blocos de construção de polinucleotídeo, em que os blocos de construção de polinucleotídeo são designados para remontagem cruzada com o polinucleotídeo modelo em uma seqüência predeterminada, e um polinucleotídeo do bloco de construção compreende uma seqüência que é uma variante do gene homólogo e uma seqüência homóloga ao polinucleotídeo modelo que flanqueia a seqüência variante; (c) combinar um polinucleotídeo de bloco de construção com um polinucleotídeo modelo de modo que o polinucleotídeo de bloco de construção remonta cruzado com o polinucleotídeo modelo para gerar polinucleotídeos compreendendo variações de seqüência de gene homóloga.

SLR não depende da presença de níveis altos de homologia entre os polinucleotídeos para ser rearranjado. Desse modo, este método pode ser usado para gerar não-estocasticamente bibliotecas (ou conjuntos) de

moléculas da progênie compreendidas de mais de 10^{100} quimeras diferentes. SLR pode ser usada para gerar bibliotecas compreendidas de mais de 10^{1000} quimeras da progênie diferentes. Desse modo, aspectos da presente invenção incluem métodos não-estocásticos de produzir um conjunto de moléculas de ácido nucléico quiméricas finalizada aparando uma ordem de montagem geral que é selecionado por modelo. Este método inclui as etapas de gerar por modelo uma pluralidade de blocos de construção de ácido nucléico específicos tendo terminações ligáveis mutuamente compatíveis úteis, e montando estes blocos de construção de ácido nucléico, de modo que uma ordem de montagem geral designada é alcançada. As terminações ligáveis mutuamente compatíveis dos blocos de construção de ácido nucléico a serem montados são consideradas serem "úteis" para este tipo de conjunto ordenado se eles permitem que os blocos de construção sejam acoplados em ordens predeterminadas. Desse modo, a ordem de montagem geral em que os blocos de construção de ácido nucléico podem ser acoplados é especificada pelo modelo das terminações ligáveis. Se mais de uma etapa de conjunto for usada, então a ordem de montagem geral em que os blocos de construção de ácido nucléico podem ser acoplados também é especificada pela ordem seqüencial da(s) etapa(s) de montagem. Em um aspecto, os pedaços de construção tratados a calor são tratados com uma enzima, como uma ligase (por exemplo DNA ligase de T4), para alcançar ligação covalente dos pedaços de construção. Em um aspecto, um método não-estocástico é denominado remontagem de ligação sintética (SLR), que de alguma forma está relacionado ao embaralhamento estocástico, a menos que os blocos de construção de ácido nucléico não sejam embaralhados ou concatenados ou quimerizados fortuitamente, mas de preferência serem montados não-estocasticamnete, podem ser usados para criar variantes.

O método de SLR não depende da presença de um nível alto de homologia entre os polinucleotídeos a ser embaralhados. A invenção pode serem usada não-estocasticamente para gerar bibliotecas (ou conjuntos) de moléculas da progênie compreendidas de mais de 10^{100} quimeras diferentes. Concebivelmente, SLR pode ser até mesmo usada para gerar bibliotecas

compreendidas de mais de 10^{1000} quimeras da progênie diferentes.

Desse modo, em um aspecto, a invenção fornece um método não-estocástico de produzir um conjunto de moléculas de ácido nucléico quiméricas finalizadas tendo uma ordem de montagem geral que é escolhida por modelo cujo método é compreendido das etapas de gerar por modelo uma pluralidade de blocos de construção de ácido nucléico específicos tendo terminações ligáveis mutuamente compatíveis úteis, e montar estes blocos de construção de ácido nucléico, de modo que uma ordem de montagem geral designada seja alcançada.

As terminações ligáveis mutuamente compatíveis dos blocos de construção de ácido nucléico a serem montados são consideradas ser "úteis" para este tipo de conjunto ordenada se elas permitirem os blocos de construção ser acoplados em ordens predeterminadas. Desse modo, em um aspecto, a ordem de montagem geral em que os blocos de construção de ácido nucléico podem ser acoplados é especificada pelas terminações ligáveis modelos e, se mais de uma etapa de conjunto for usada, então a ordem de montagem geral em que os blocos de construção de ácido nucléico podem ser acoplados também é especificada pela ordem seqüencial da(s) etapa(s) da montagem. Em um aspecto da invenção, os pedaços de construção tratados a calor são tratados com uma enzima, como uma ligase (por exemplo, DNA ligase de T4) para alcançar ligação covalente dos pedaços de construção.

Em um outro aspecto, o modelo de blocos de construção de ácido nucléico é obtido na análise das seqüências de um conjunto de modelos de ácido nucléico progenitores que serve como uma base para produzir um conjunto da progênie de moléculas de ácido nucléico quiméricas finalizadas. Estes modelos de ácido nucléico progenitores desse modo servem como uma fonte de informação da seqüência que ajuda no modelo dos blocos de construção de ácido nucléico que são submetidos à mutagênese, isto é quimerizados ou embaralhados.

Em uma exemplificação, a invenção fornece a quimerização de uma família de genes relacionados e sua família codificada dos produtos

relacionados. Em uma exemplificação particular, os produtos codificados são enzimas. As enzimas e polipeptídeos para uso na invenção podem ser submetidos à mutagênese de acordo com os métodos aqui descritos.

Desse modo de acordo com um aspecto da invenção, as seqüências de uma pluralidade de modelos de ácido nucléico progenitores são alinhadas para selecionar um ou mais pontos de demarcação, cujos pontos de demarcação podem estar localizados em uma área de homologia. Os pontos de demarcação podem ser usados para delinear os limites dos blocos de construção de ácido nucléico a serem gerados. Desse modo, os pontos de demarcação identificados e selecionados nas moléculas progenitoras servem como pontos de quimerização potenciais no conjunto das moléculas da progênie.

Tipicamente um ponto de demarcação útil é uma área de homologia (compreendida de pelo menos uma base de nucleotídeo homóloga) compartilhada por pelo menos dois modelos progenitores, mas o ponto de demarcação pode ser uma área de homologia pela qual é compartilhada pelo menos metade dos modelos progenitores, pelo menos dois terços dos modelos progenitores, pelo menos três quarto dos modelos progenitores ou quase todos os modelos progenitores. Em um aspecto, um ponto de demarcação útil é uma área de homologia que é compartilhada por todos os modelos progenitores.

Em um aspecto, o processo de remontagem de ligação é executado exaustivamente para gerar uma biblioteca exaustiva. Em outras palavras, todas possíveis combinações ordenadas dos blocos de construção de ácido nucléico são representadas no conjunto de moléculas de ácido nucléico quiméricas finalizadas. Ao mesmo tempo, a ordem de montagem (isto é a ordem de montagem de cada bloco de construção nas seqüências 5' a 3' de cada ácido nucléico quimérico finalizado) em cada combinação é por modelo (ou não-estocástica). Por causa da natureza não-estocástica do método, a possibilidade de produtos laterais não desejados é muito reduzida.

Em outro aspecto, o método fornece que, o processo de remontagem de ligação é executado sistematicamente, por exemplo, para gerar

uma biblioteca sistematicamente compartimentada, com compartimentos que podem ser triados sistematicamente, por exemplo, um por um. Em outras palavras a invenção fornece, através do uso seletivo e judicioso de blocos de construção de ácido nucléico, acoplado ao uso seletivo e judicioso de reações de montagem seqüencialmente em etapas, um modelo experimental pode ser alcançado onde conjuntos específicos de produtos da progênie são feitos em cada de vários vasos de reação. Isto permite executar um exame sistemático e procedimento de triagem. Desse modo, ela permite um número potencialmente muito grande de moléculas da progênie ser examinado sistematicamente em grupos menores.

Por causa de sua capacidade de executar quimerizações de uma maneira que ainda é altamente flexível, embora exaustivo e sistemático também, particularmente quando houver um nível baixo de homologia entre as moléculas progenitoras, a presente invenção fornece a geração de uma biblioteca (ou conjunto) compreendida de um número grande de moléculas da progênie. Por causa da natureza não-estocástica da invenção de remontagem de ligação imediata, as moléculas da progênie geradas podem compreender uma biblioteca de moléculas de ácido nucléico quiméricas finalizadas tendo uma ordem de montagem geral que é selecionada por modelo. Em um aspecto particular, uma tal biblioteca gerada é compreendida de mais de 10^3 a mais de 10^{1000} espécies moleculares diferentes da progênie.

Em um aspecto, um conjunto de moléculas de ácido nucléico quiméricas finalizadas, produzidas como descrito, é compreendido de um polinucleotídeo codificando um polipeptídeo. De acordo com um aspecto, este polinucleotídeo é um gene que pode ser um gene artificial. De acordo com outro aspecto, este polinucleotídeo é uma via de gene, que pode ser uma via de gene feita pelo homem. A invenção fornece um ou mais genes artificiais gerados pela invenção que podem ser incorporados em uma via de gene artificial, como via operável em um organismo eucariótico (incluindo uma planta).

Em outra exemplificação, a natureza sintética da etapa em que os blocos de construção são gerados permite a modelagem e introdução dos

nucleotídeos (por exemplo, um ou mais nucleotídeos que podem ser, por exemplo, códon ou íntrons ou seqüências reguladoras) que podem ser depois opcionalmente removidos em um processo *in vitro* (por exemplo, através de mutagênese) ou em um processo *in vivo* (por exemplo, utilizando a capacidade de junção de gene de um organismo hospedeiro). É apreciado que em muitas circunstâncias a introdução destes nucleotídeos pode ser também desejável por muitas outras razões além do benefício potencial de criar um ponto de demarcação útil.

Desse modo, de acordo com outro aspecto, a invenção fornece um bloco de construção de ácido nucléico que pode ser usado para introduzir um íntron. Desse modo, a invenção fornece íntrons funcionais que podem ser introduzidos em um gene artificial da invenção. A invenção também fornece íntrons funcionais que podem ser introduzidos em uma via de gene artificial da invenção. Conseqüentemente, a invenção fornece a geração de um polinucleotídeo quimérico que é um gene artificial contendo um (ou mais) íntron(s) artificialmente introduzido(s).

Conseqüentemente, a invenção também fornece a geração de um polinucleotídeo quimérico que é uma via de gene artificial que contém um (ou mais) íntron artificialmente introduzidos(s). Em um aspecto, o(s) íntron(s) artificialmente introduzido(s) é(são) funcional(is) em uma ou mais células hospedeiras para a junção do gene grandemente de modo que os íntrons de ocorrência natural servem funcionalmente na junção do gene. A invenção fornece um processo de produzir polinucleotídeos contendo íntrons artificiais a serem introduzidos em organismos hospedeiros para recombinação e/ou junção.

Um gene artificial produzido usando a invenção também pode servir como um substrato para recombinação com outro ácido nucléico. Igualmente, uma via de gene artificial produzida usando a invenção também pode servir como um substrato para recombinação com outro ácido nucléico. Em um aspecto, a recombinação é facilitada, ou ocorre, em áreas de homologia entre o gene contendo íntron artificial e um ácido nucléico com servidor como um par de recombinação. Em um aspecto, o par de recombinação

pode ser também um ácido nucléico gerado pela invenção, incluindo um gene artificial ou uma via de gene artificial. Recombinação pode ser facilitada, ou pode ocorrer nas áreas de homologia que existe em um (ou mais) íntron(s) artificialmente introduzido(s) no gene artificial.

O método de remontagem de ligação sintética da invenção utiliza uma pluralidade de blocos de construção de ácido nucléico cada um destes pode ter duas terminações ligáveis. As duas terminações ligáveis em cada bloco de construção de ácido nucléico pode ser duas extremidades cegas (isto é cada uma tendo uma extensão de zero nucleotídeo), ou uma extremidade cega e uma extensão, ou duas extensões.

Uma extensão útil para este propósito pode ser uma extensão 3' ou uma extensão 5'. Desse modo, um bloco de construção de ácido nucléico pode ter uma extensão 3' ou alternativamente uma extensão 5' ou alternativamente duas extensões 3' ou alternativamente duas extensões 5'. A ordem geral em que os blocos de construção de ácido nucléico são montados para formar uma molécula de ácido nucléico quimérico finalizada é determinada por modelo experimental proposital e não é fortuito.

Em um aspecto, um bloco de construção de ácido nucléico é gerado por síntese química de dois ácidos nucléicos unifilamentados (também referido como oligos unifilamentados) e os contatando para lhes permitir tratar a calor para formar um bloco de construção de ácido nucléico bifilamentado.

Um bloco de construção de ácido nucléico bifilamentado pode ser de tamanho variável. Os tamanhos destes blocos de construção podem ser pequenos ou grandes. Tamanhos exemplares para os blocos de construção variam de 1 par de base (não incluindo a extensão) para 100.000 pares de base (não incluindo a extensão). Também são fornecidas outras faixas de tamanho tendo limites mais baixos de 1 bp para 10.000 bp (incluindo todo valor de número inteiro nos intervalos), e limites superiores de 2 bp para 100.000 bp (incluindo todo valor de número inteiro nos intervalos).

Muitos métodos existem pelos quais um bloco de construção de ácido nucléico bifilamentado pode ser gerado que é útil para a invenção; e

estes são conhecidos na técnica e podem ser facilmente executados pelo versado.

De acordo com um aspecto, um bloco de construção de ácido nucléico bifilamentado é gerado primeiro gerando dois ácidos nucléicos unifilamentados e lhes permitindo tratar a calor para formar um bloco de construção de ácido nucléico bifilamentado. Os dois filamentos de um bloco de construção de ácido nucléico bifilamentado podem ser complementares a cada nucleotídeo exceto qualquer um que forma uma extensão; desse modo não contendo nenhuma desigualdade, além da(s) extensão(ões). De acordo com outro aspecto, os dois filamentos de um bloco de construção de ácido nucléico bifilamentado são complementares nada menos que todo nucleotídeo exceto qualquer um que forma uma extensão. Desse modo, de acordo com este aspecto, um bloco de construção de ácido nucléico bifilamentado pode ser usado para introduzir degeneração de códon. Em um aspecto, a degeneração de códon é introduzida usando a mutagênese de saturação de sítio aqui descrita, usando um ou mais cassetes de N,N,G/T ou alternativamente usando um ou mais cassetes de N,N,N.

O método *in vivo* de recombinação da invenção pode ser executado inconscientemente em um banco comum de híbridos ou alelos desconhecidos de um polinucleotídeo específico ou seqüência. Porém, não é necessário saber a seqüência real de DNA ou de RNA do polinucleotídeo específico.

O método de usar recombinação dentro de uma população embaralhada de genes pode ser útil para a geração de qualquer proteína útil, por exemplo, interleucina I, anticorpos, tPA e hormônio de crescimento. Este método pode ser usado para gerar proteínas tendo especificidade ou atividade alterada. O método também pode ser útil para a geração de seqüências de ácidos nucléicos híbridas, por exemplo, regiões de promotor, íntrons, éxons, seqüências intensificadora, 31 regiões não-transferidas ou 51 regiões não-transferidas de genes. Desse modo, este método pode ser usado para gerar genes tendo taxas de expressão aumentadas. Este método também pode ser útil no estudo de seqüências repetitivas de DNA. Por fim, este mé-

todo pode ser útil para transformar ribozimas ou aptâmeros.

Em um aspecto, variantes dos polinucleotídeos e polipeptídeos aqui descritas são obtidas pelo uso de ciclos repetidos de reclassificação redutiva, recombinação e seleção que permitem a evolução molecular direcionada de seqüências lineares altamente complexas, como DNA, RNA ou proteínas através de recombinação.

Embaralhamento *in vivo* de moléculas é útil para fornecer variantes e pode ser executada utilizando a propriedade natural das células para recombinar os multímeros. Embora a recombinação *in vivo* tenha fornecido a rota natural principal para diversidade molecular, a recombinação genética permanece um processo relativamente complexo que envolve 1) o reconhecimento das homologias; 2) clivagem do filamento, invasão do filamento e etapas metabólicas que conduzem à produção de quiasma recombinante; e por fim 3) a resolução do quiasma em moléculas recombinadas distintas. A formação do quiasma requer o reconhecimento das seqüências homólogas.

Em um outro aspecto, a invenção inclui um método para produzir um polinucleotídeo híbrido de pelo menos um primeiro polinucleotídeo e um segundo polinucleotídeo. A invenção pode ser usada para produzir um polinucleotídeo híbrido introduzindo pelo menos um primeiro polinucleotídeo e um segundo polinucleotídeo que compartilha pelo menos uma região de homologia de seqüência parcial (por exemplo, SEQ ID Nº: 1) em uma célula hospedeira adequada. As regiões de homologia de seqüência parcial promovem processos que resultam na reorganização da seqüência produzindo um polinucleotídeo híbrido. O termo "polinucleotídeo híbrido", como aqui usado, é qualquer seqüência de nucleotídeos que é o resultado do método da presente invenção e contém seqüência de pelo menos duas seqüências de polinucleotídeos originais. Tais polinucleotídeos híbridos podem ser o resultado de eventos de recombinação intermolecular que promovem integração de seqüência entre as moléculas de DNA. Além disso, tais polinucleotídeos híbridos podem ser o resultado de processos de reclassificação intermolecular redutiva que utiliza seqüências repetidas para alterar uma seqüência de nucleotídeos dentro de uma molécula de DNA.

A invenção fornece métodos para gerar polinucleotídeos híbridos que podem codificar polipeptídeos híbridos biologicamente ativos (por exemplo, uma fitase híbrida). Em um aspecto, os polinucleotídeos originais codificam polipeptídeos biologicamente ativos. O método da invenção produz novos polipeptídeos híbridos utilizando processos celulares que integram a seqüência dos polinucleotídeos originais de modo que o polinucleotídeo híbrido resultante codifica um polipeptídeo que demonstra atividades derivadas dos polipeptídeos originais biologicamente ativos. Por exemplo, os polinucleotídeos originais podem codificar uma enzima particular de microorganismos diferentes. Uma enzima codificada por um primeiro polinucleotídeo de um organismo ou variante pode, por exemplo, funcionar eficazmente sob uma condição ambiental particular, por exemplo, salinidade alta. Uma enzima codificada por um segundo polinucleotídeo de um organismo diferente ou variante pode funcionar eficazmente sob uma condição ambiental diferente, como temperaturas extremamente altas. Um polinucleotídeo híbrido contendo seqüências dos primeiro e segundo polinucleotídeos originais pode codificar uma enzima que apresenta características de ambas as enzimas codificadas pelos polinucleotídeos originais. Desse modo, a enzima codificada pelo polinucleotídeo híbrido pode funcionar eficazmente sob condições ambientais compartilhadas por cada uma das enzimas codificadas pelos primeiro e segundo polinucleotídeos, por exemplo, salinidade alta e temperaturas extremas.

Além dos vários métodos acima descritos, vários métodos são conhecidos na técnica que podem ser usados para obter polinucleotídeos híbridos com propriedades enzimáticas intensificadas. Os exemplos a seguir ilustram o uso de tais procedimentos para obter enzimas termoestáveis ou termotolerantes por mutagênese de um polinucleotídeo codificando uma enzima do tipo selvagem de interesse.

Por exemplo, M. Lehmann et al. (em Biochimica et Biophysica Acta 1543:408-415, 2000) descreve um "método consensual" em que o alinhamento da seqüência de fitases fúngicas homólogas foi usado para calcular uma seqüência consensual de aminoácidos de fitase. Na construção

do gene consensual correspondente, expressão recombinante e purificação, a fitase recombinante obtida apresentou uma temperatura de não-dobramento (Tm) 15-22°C mais alta que aquela de todas as fitases de origem usadas no modelo. Mutagênese sítio-direcionada da proteína recombinante de codificação do gene foi usada também para aumentar o valor de Tm para 90,4°C. O efeito termoestabilizante foi atribuído a uma combinação de permutas de aminoácidos múltiplas que foram distribuídas na seqüência inteira da proteína e principalmente resíduos afetados na superfície.

Outro método para obter uma enzima com propriedades térmicas intensificadas é descrito por L. Jermutus et al. (J. Of Biotecnology 85:15-24, 2001). Neste método, interações iônicas e ligações de hidrogênio na superfície de fitase de Aspergillus terreus foram primeiro restabelecidas para corresponder àqueles presentes no homólogo, mas mais enzima termoestável de A. niger. Depois os elementos estruturais secundários inteiros foram substituídos na mesma região e com base na estrutura cristalina de fitase de A. niger. A substituição de uma α-hélice na superfície de fitase de A. terreus pela extensão correspondente da fitase de A. niger resultou em uma enzima quimérica com base na estrutura (proteína de fusão) com termoestabilidade melhorada e atividade enzimática inalterada.

Ainda outro método é ilustrado por L. Giver et al. (Proc. Natl. Acad. Sci. USA 95:12809-12813, 1998), em que seis gerações de mutagênese aleatória introduzidas durante PCR mutagênica de um polinucleotídeo codificando p-nitrobenzil esterase de Bacillus subtilisp seguido por recombinação *in vitro* com base no método de Stemmer resultou em uma esterase recombinante com termoestabilidade aumentada (maior que aumento de 14°C na Tm) sem comprometer a atividade catalítica em temperaturas mais baixas.

C. Vetriani et al. (Proc. Natl. Acad. Sci USA 95:12300-12305, 1998) descreve um procedimento pelo qual modelação com base na homologia e comparação de estrutura direta das glutamato desidrogenases hexaméricas dos hipertermófilos Pyrococcus furiosus e Thermococcus litoralis, com ótimas temperaturas de desenvolvimento de 100°C e 88°C, respectiva-

mente, foram usados para determinar as características de termoestabilização fundamentais. Uma rede de par de íon de intersubunidade observada ser substancialmente reduzida na enzima menos estável foi alterada por mutagênese de dois resíduos dela para restabelecer as interações encontradas na enzima mais estável. Embora qualquer mutação simples tivesse tido efeitos adversos na termoestabilidade, com ambas as mutações no lugar, uma melhoria de quatro vezes da estabilidade a 104°C na enzima do tipo selvagem foi observada.

A. Tomschy et al. (Protein Science 9:1304-1311, 2000) descreve um procedimento que utiliza a estrutura cristalina de fitase de Aspergillus Niger (em resolução de 2,5 angströms) para especificar todos os sítios ativos da enzima. Um alinhamento de seqüência de aminoácidos múltiplo foi depois usado para identificar os resíduos de sítio ativo não-conservado que podem correlatar uma certa propriedade favorável de interesse. Usando este método, Gln27 de fitase de A. fumigatus, que diferiu de Leu27 de A. niger, foi identificado provavelmente como estar envolvido na ligação e/ou liberação de substrato e responsável pela atividade específica mais baixa da fitase de A. fumigatus (26,5 versus 196 6 U/mg de proteína a pH 5,0). Mutagênese sítio-direcionada de Gln27 de fitase de A. fumigatus para Leu aumentou a atividade específica da enzima mutante para 92,1 U/mg de proteína.

Em um aspecto, a invenção atual fornece um método (e produtos deste) para produzir formulações líquidas aquosas estabilizadas tendo atividade de fitase que apresenta resistência aumentada para aquecer à inativação por calor da atividade da enzima e retendo sua atividade de fitase durante períodos prolongados de armazenamento. As formulações líquidas são estabilizadas por meio da adição de uréia e/ou um poliol como sorbitol e glicerol como agente estabilizante. Também fornecido são preparações de alimentação de animais monogástricos e métodos para a produção destes que resultam do uso de tais formulações líquidas aquosas estabilizadas. Detalhes adicionais relativos a este método estão na literatura pública e/ou são conhecidos ao versado. Em uma exemplificação não-limitativa particular, tal literatura publicamente disponível inclui EP 0626010 (WO 9316175 A1)

(Barendse et al.), embora referências na literatura publicamente disponível não ensinem as moléculas inventivas do presente pedido de patente.

Enzimas codificadas por polinucleotídeos originais incluem, mas não são limitadas, hidrolases e fitases. Um polipeptídeo híbrido que é o resultado do método da invenção pode exibir atividade de enzima especializada não mostrada nas enzimas originais. Por exemplo, seguindo a recombinação e/ou reclassificação redutiva de polinucleotídeos codificando atividades de hidrolase, o polipeptídeo híbrido resultante codificado por um polinucleotídeo híbrido pode ser triado por atividades de hidrolases especializadas obtidas de cada uma das enzimas originais, i. e., o tipo de ligação na qual a hidrolase age e a temperatura na qual a hidrolase funciona. Por exemplo, desse modo a hidrolase pode ser triada para averiguar quais funcionalidades químicas que distinguem a hidrolase híbrida das hidrolases originais, como: (a) amida (ligação peptídica), isto é, protease; (b) ligações de éster, isto é, esterases e lipases; (c) acetais, isto é, por exemplo, glicosidase e, por exemplo, a temperatura, pH ou concentração de sal em que os polipeptídeos híbridos funcionam.

As fontes dos polinucleotídeos originais podem ser isoladas de organismos individuais ("isolados"), coletas de organismos que foram cultivados em meios definidos ("culturas de enriquecimento"), ou, organismos não-cultivados ("amostras ambientais"). O uso de um método independente da cultura para derivar polinucleotídeos codificando novas bioatividades de amostras ambientais, como termoestabilidade ou termotolerância, podem ser usados para acessar recursos inexplorados de biodiversidade.

"Bibliotecas ambientais" são geradas de amostras ambientais e representam os genomas coletivos de organismos de ocorrência natural arquivados nos vetores de clonagem que podem ser propagados em hospedeiros procarióticos adequados. Porque o DNA clonado é inicialmente extraído diretamente das amostras ambientais, as bibliotecas não são limitadas à fração pequena de procariotas que podem ser cultivados em cultura pura. Adicionalmente, uma normalização do DNA ambiental presente nestas amostras pode permitir representação mais igual do DNA de todas as espé-

cies presentes na amostra original. Isto pode dramaticamente aumentar a eficiência de encontrar genes interessantes de constituintes secundários da amostra que pode ser sub-representada por várias ordens de magnitude comparada às espécies dominantes.

Por exemplo, bibliotecas de genes geradas de um ou mais microorganismos não-cultivados são triadas por uma atividade de interesse. As vias potenciais que codificam moléculas bioativas de interesse são primeiro capturadas em células procarióticas na forma de bibliotecas de expressão de gene. Os polinucleotídeos codificando atividades de interesse são isolados de tais bibliotecas e introduzidos em uma célula hospedeira. A célula hospedeira é cultivada sob condições que promovem recombinação e/ou reclassificação redutiva criando biomoléculas potencialmente ativas com atividades novas ou intensificadas.

Os microorganismos dos quais o polinucleotídeo pode ser preparado incluem microorganismos procarióticos, como Xanthobacter, Eubacteria e Archaebacteria, e microorganismos eucarióticos inferiores como fungos, algumas algas e protozoários. Os polinucleotídeos podem ser isolados das amostras ambientais em cujo caso o ácido nucléico pode ser restabelecido sem cultivo de um organismo ou pode ser restabelecido de um ou mais organismos cultivados. Em um aspecto, tais microorganismos podem ser extremófilos, como hipertermófilos, psicrófilos, psicrotrófilos, halófilos, barófilos e acidófilos. Polinucleotídeos codificando enzimas isoladas de microorganismos extremofílicos podem ser usados. Tais enzimas podem funcionar em temperaturas acima de 100°C em fontes térmicas terrestres e aberturas térmicas em alto mar, em temperaturas abaixo de 0°C em águas árticas, no ambiente de sal saturado do Mar Morto, em valores de pH por volta de 0 em depósitos de carvão e fontes geotérmicas ricas em enxofre, ou em valores de pH maior que 11 em lama de esgoto. Por exemplo, várias esterases e lipases clonadas e expressadas de organismos extremófilos mostram alta atividade ao longo de uma faixa extensiva de temperaturas e pHs.

Os polinucleotídeos selecionados e isolados como anteriormente descritos são introduzidos em uma célula hospedeira adequada. Uma célula

hospedeira adequada é qualquer célula que é capaz de promover recombinação e/ou reclassificação redutiva. Os polinucleotídeos selecionados podem estar em um vetor incluindo seqüências de controle apropriadas. A célula hospedeira pode ser uma célula eucariótica mais alta, como uma célula mamífera, ou uma célula eucariótica mais baixa, como uma célula de levedura, ou a célula hospedeira pode ser uma célula procariótica, como uma célula bacteriana. A introdução da construção na célula hospedeira pode ser realizada através de transfecção de fosfato de cálcio, transfecção mediada por DEAE-dextrana ou eletroporação (Davis et al., 1986).

Como exemplos representativos de hospedeiros apropriados, pode ser mencionado: células bacterianas, como E. coli, Streptomyces, Salmonela typhimurium; células fúngicas, como levedura; células de inseto como Drosophila S2 e Spodoptera Sf9; células animais como CHO, COS ou melanoma de Bowes; adenovírus; e células vegetais. A seleção de um hospedeiro apropriado é julgada estar dentro do escopo daqueles versados na técnica dos ensinamentos aqui.

A maioria dos compostos bioativos correntemente em uso é derivada de microorganismos do solo. Muitos micróbios que habitam no solo e outras comunidades ecológicas complexas produzem uma variedade de compostos que aumentam sua capacidade para sobreviver e proliferar. Em geral, acredita-se que estes compostos sejam dispensáveis para o desenvolvimento do organismo e sejam sintetizados com a ajuda de genes envolvidos no metabolismo intermediário daí seu nome - "metabólitos secundários." Metabólitos secundários, em geral, são os produtos de vias biossintéticas complexas e são usualmente derivados de precursores celulares comuns. Metabólitos secundários que influenciam o desenvolvimento ou sobrevivência de outros organismos são conhecidos como compostos "bioativos" e servem como componentes fundamentais do arsenal de defesa química de tanto micro quanto macro-organismos. O homem tem explorado estes compostos para uso como antibióticos, antiinfectivos e outros compostos bioativos com atividade contra uma faixa larga de patógenos procarióticos e eucarióticos. Aproximadamente 6.000 compostos bioativos de ori-

gem microbiana foram caracterizados, com mais de 60 % produzidos pelas bactérias do solo gram-positivas do gênero Streptomyces. (Barnes et al., Proc. Nat. Acad. Sci. U.S.A, 91, 1994).

Triagem de hibridação usando filtros de alta densidade ou uso de bateia provou ser um método eficiente para detectar homólogos de vias contendo genes de interesse para descobrir novas moléculas bioativas que não podem ter nenhuma contraparte conhecida. Uma vez um polinucleotídeo de interesse é enriquecido em uma biblioteca de clones pode ser desejável triar por uma atividade. Por exemplo, pode ser desejável triar a expressão de estruturas de anel de molécula pequena ou "cadeias principais." Porque os genes que codificam estas estruturas policíclicas podem freqüentemente ser expressados em E. coli, a cadeia principal de molécula pequena pode ser produzida, até mesmo se em uma forma inativa. Bioatividade é conferida ao transferir a molécula ou via para um hospedeiro apropriado que expressa a glicosilação requerida e genes de metilação que podem modificar ou "decorar" a estrutura para sua forma ativa. Desse modo, até mesmo se os compostos do anel inativo, recombinantemente expressados em E. coli forem detectados para identificar os clones que são depois transportados para um hospedeiro metabolicamente rico, como Streptomyces (por exemplo, Streptomyces diversae ou venezuelae) para produção subseqüente da molécula bioativa. Deve ser entendido que o E. coli pode produzir moléculas pequenas ativas e em certas circunstâncias pode ser desejável transportar os clones para um hospedeiro metabolicamente rico para "decoração" da estrutura, mas não requerido. O uso de sistemas robóticos de alto processamento permite a triagem de centenas de milhares de clones em arranjos multiplexados em discos de microtitulação.

Os ácidos nucléicos da invenção podem ser expressados, ou sobreexpressados, em qualquer sistema de expressão *in vitro* ou *in vivo*. Qualquer sistema de cultura de células pode ser empregado para expressar, ou sobreexpressa, a proteína recombinante, incluindo culturas bacterianas, de insetos, de levedura, fúngicas ou mamíferas. A sobreexpressão pode ser realizada por escolha apropriada de promotores, intensificadores, vetores

(por exemplo, uso de vetores de réplicon, vetores distrônicos (ver, por exemplo, Gurtu (1996) Biochem. Biophys. Res. Commun. 229:295-8)), meios, sistemas de cultura e outros. Em um aspecto, a amplificação do gene usando marcadores de seleção, por exemplo, glutamina sintetase (ver, por exemplo, Sanders (1987) Dev. Biol. Stand. 66:55-63), em sistemas celulares é usada para sobreexpressar os polipeptídeos da invenção.

Vários sistemas de cultura de células mamíferas podem ser empregados para expressar proteína recombinante, exemplos de sistemas de expressão mamífera incluem as linhagens de COS-7 de fibroblastos de rim de macaco, descritas em "SV40-transformed simian cells support the replication of early SV40 mutants" (Gluzman, 1981), e outras linhagens de células capazes de expressar um vetor compatível, por exemplo, as linhagens de células de C127, 3T3, CHO, HeLa e de BHK. Os vetores de expressão mamífera compreenderão uma origem de replicação, um promotor e intensificador adequado, e também se necessário sítios de ligação ao ribossoma, sítio de poliadenilação, sítios doadores e aceptores de junção, seqüências de terminação transcricional e seqüências não-transcritas de flanqueamento 5'. Seqüências de DNA derivadas da junção SV40 e os sítios de poliadenilação podem ser usados para fornecer os elementos genéticos não-transcritos requeridos.

As células hospedeiras contendo os polinucleotídeos de interesse podem ser cultivadas em meios nutrientes convencionais modificados quando apropriado para ativar os promotores, selecionando transformantes ou genes de amplificação. As condições de cultura, como temperatura, pH e similares, são aquelas previamente usadas com a célula hospedeira selecionada para expressão, e será óbvio ao artesão normalmente versado. Os clones que são identificados como tendo a atividade de enzima especificada podem ser depois seqüenciados para identificar a seqüência de polinucleotídeos codificando uma enzima tendo a atividade intensificada.

As enzimas e polinucleotídeos da presente invenção podem ser fornecidos em uma forma isolada ou podem ser purificados para homogeneidade. O polipeptídeo de fitase da invenção pode ser obtido usando

quaisquer de vários métodos padrão. Por exemplo, podem ser produzidos polipeptídeos de fitase em um sistema de expressão recombinante padrão (como descrito aqui), quimicamente sintetizado (embora um pouco limitado aos fragmentos de peptídeo de fitase pequenos) ou purificados de organismos em que eles são expressados naturalmente. Métodos de expressão recombinantes úteis incluem os hospedeiros mamíferos, hospedeiros microbianos e hospedeiros vegetais.

A expressão recombinante ou sobreexpressão das moléculas de fitase da invenção pode ser alcançada em combinação com uma ou mais moléculas adicionais como, por exemplo, outras enzimas. Este método é útil para produzir produtos de combinação, como uma planta ou parte de planta contendo as moléculas de fitase imediatas como também uma ou mais moléculas adicionais - podem ser usadas as moléculas de fitase e as moléculas adicionais em um tratamento de combinação. As moléculas recombinantemente espessadas resultantes podem ser usadas na forma homogeneizada e/ou purificada ou alternativamente na forma relativamente não purificada (por exemplo como partes vegetais de artigo de consumo que são úteis quando embaralhadas com outros gêneros alimentícios para catalisar a degradação de fitato).

A presente invenção fornece uma enzima recombinante expressa em um hospedeiro. A presente invenção fornece uma enzima substancialmente pura de fitase. Desse modo, uma enzima da presente invenção pode ser uma enzima recombinante, uma enzima natural ou uma enzima sintética, ou uma enzima recombinante.

Em um aspecto particular, a presente invenção fornece para a expressão de fitase em plantas ou órgãos de plantas transgênicas e métodos para sua produção. Constructos de expressão de DNA são fornecidos para a transformação das plantas com um gene que codifica fitase sob o controle de seqüências reguladoras que são capazes de direcionar a expressão da fitase. Estas seqüências reguladoras incluem seqüências capazes de direcionar transcrição em plantas, ou constitutivamente, ou em estágio e/ou maneiras tecido-específicas.

A maneira de expressão depende, em parte, do uso da planta ou partes desta. As plantas e órgãos de plantas transgênicas fornecidos pela presente invenção podem ser aplicados a uma variedade de processos industriais ou diretamente, por exemplo em alimentações de animal ou alternativamente, a fitase expressada pode ser extraída e se desejado, purificada antes da aplicação. Alternativamente, a planta hospedeira recombinante ou parte da planta pode ser usada diretamente. Em um aspecto particular, a presente invenção fornece métodos de catalisar as reações de hidrolisação de fitase usando sementes contendo quantidades intensificadas de fitase. O método envolve contatar as sementes transgênicas, do tipo não-selvagem, por exemplo, em uma forma moída ou mastigada, com substrato contendo fitase e deixar as enzimas nas sementes para aumentar a taxa de reação. Adicionando diretamente as sementes a um substrato contendo fitase, a invenção fornece uma solução ao processo dispendioso e problemático de extrair e purificar a enzima. Em uma exemplificação, a presente invenção fornece métodos de tratamento por meio dos quais um organismo que necessita de um fornecimento suficiente de uma enzima é administrado à enzima na forma de sementes contendo quantidades intensificadas da enzima. Em um aspecto, o tempo da administração da enzima a um organismo é coordenado quanto ao consumo de um gênero alimentício contendo fitase.

A expressão de fitase em plantas pode ser alcançada por uma variedade de meios. Especificamente, por exemplo, tecnologias estão disponíveis para transformar um número grande de espécies de plantas, incluindo espécies dicotiledôneas (por exemplo tabaco, batata, tomate, Petúnia, Brassica). Adicionalmente, por exemplo, estratégias para a expressão de genes estranhos em plantas estão disponíveis. Adicionalmente ainda, foram identificadas seqüências reguladoras de genes de planta que são úteis para a construção de genes quiméricos que podem ser funcionalmente expressados em plantas e em células vegetais (por exemplo Klee et al., 1987; Clark et al., 1990; Smith et al., 1990).

A introdução de construções de gene em plantas pode ser alcançada usando várias tecnologias incluindo transformação com Agrobacte-

rium tumefaciens ou Agrobacterium rhizogenes. Exemplos não-limitativos de tecidos vegetais que podem ser transformados assim incluem protoplastos, microsporos ou pólen, e explantes como folhas, talos, raízes, hipocótilos e cotilédones. Além disso, DNA pode ser introduzido diretamente nos protoplastos e células ou tecidos vegetais por microinjeção, eletroporação, bombardeio de partícula e absorção de DNA direta.

As proteínas podem ser produzidas em plantas por uma variedade de sistemas de expressão. Por exemplo, o uso de um promotor constitutivo como o promotor 35S de Vírus do Mosaico de Couve-flor (Guilley et al., 1982) é útil para a acumulação da proteína expressada em virtualmente todos órgãos da planta transgênica. Alternativamente, o uso de promotores que são altamente tecido-específicos e/ou estágio-específicos é útil para esta invenção (Higgins, 1984; Shotwell, 1989) para influenciar a expressão para os tecidos desejados e/ou para um estágio desejado de desenvolvimento. Outros detalhes pertinentes à expressão em plantas das moléculas de fitase da invenção atual são divulgados, por exemplo, na Patente U. S. No. 5.770.413 (Van Ooijen et al.) e Patente U. S. No. 5.593.963 (Van Ooijen et al.), embora estas referências não ensinam as moléculas inventivas do presente pedido de patente e ao contrário ensinam o uso de fitases fúngicas.

Em suma, é relevante a esta invenção que uma variedade de meios para alcançar a expressão recombinante de fitase em uma planta transgênica ou parte de planta pode ser usada. Tais plantas transgênicas e partes da planta são úteis como fontes de fitase recombinantemente expressada que pode ser adicionada diretamente às fontes contendo fitase. Alternativamente, a fitase expressada em planta recombinante pode ser extraída da fonte de planta e, se desejado, purificada antes de contatar o substrato de fitase.

Dentro do contexto da presente invenção, plantas a serem selecionadas incluem, mas não são limitadas às plantações que produzem flores comestíveis como couve-flor (Brassica oleracea), alcachofra (Cynara scolymus), frutas como maçã (Malus, por exemplo domesticus), banana (Musa, por exemplo acuminata), framboesa (como a groselha, Ribes, por exemplo

rubrum), cerejas (como a cereja doce, Prunus por exemplo avium), pepino (Cucumis, por exemplo sativus), uva (Vitis, por exemplo vinifera), limão (Citrus limon), melão (Cucumis melo), nozes (como a noz, Juglans, por exemplo regia; amendoim, Arachis hypogeae), laranja (Citrus, por exemplo maxima), pêssego (Prunus, por exemplo persica), pêra (Pyra, por exemplo communis), ameixa (Prunus, por exemplo domestica), morango (Fragaria, por exemplo moschata), tomate (Lycopersicon, por exemplo esculentum), folhas, como alfafa (Medicago, por exemplo sativa), repolhos (por exemplo Brassica oleracea), endívia (Cichoreum, por exemplo endivia), alho-poró (Allium, por exemplo porrum), alface (Lactuca, por exemplo sativa), espinafre (Spinacia, por exemplo oleraceae), tabaco (Nicotiana, por exemplo tabacum), raízes, como araruta (Maranta, por exemplo arundinacea), beterraba (Beta, por exemplo vulgaris), cenoura (Daucus, por exemplo carota), mandioca (Manihot, por exemplo esculenta), nabo (Brassica, por exemplo rapa), rabanete (Raphanus, por exemplo sativus), inhame (Dioscorea, por exemplo esculenta), batata-doce (Ipomea batatas) e sementes, como feijão (Phaseolus, por exemplo vulgaris), ervilha (Pisum, por exemplo sativium), feijão-soja (Glycin, por exemplo max), trigo (Triticum, por exemplo aestivum), cevada (Hordeum, por exemplo vulgare), milho (Zea, por exemplo mays), arroz (Oryza, por exemplo sativa), colza (Brassica napus), painço (Panicum L.), girassol (Helianthus annus), aveias (Avena sativa), tubérculos, como couve-rábano (Brassica, por exemplo oleraceae), batata (Solanum, por exemplo tuberosum) e similares.

É entendido que planta adicional como também sistemas de expressão de não-planta podem ser usados dentro do contexto desta invenção. A escolha das espécies vegetais é primariamente determinada pelo uso intencionado da planta ou partes desta e a amenabilidade das espécies de plantas para transformação.

Várias técnicas estão disponíveis para a introdução do constructo de expressão contendo a seqüência de DNA de codificação de fitase dentro das plantas alvo. Tais técnicas incluem mas não são limitadas à transformação de protoplastos usando o método de cálcio/polietileno glicol,

eletroporação e microinjeção ou bombardeio de partículas (revestidas) (Potrykus, 1990). Além destes assim chamados métodos de transformação de DNA diretos, sistemas de transformação que envolvem vetores estão extensamente disponíveis, como vetores viróticos (por exemplo do Vírus do Mosaico de Couve-flor (CaMV) e vetores bacterianos (por exemplo do gênero Agrobacterium) (Potrykus, 1990). Após seleção e/ou triagem, os protoplastos, células ou partes da planta que foram transformadas em plantas inteiras, podem ser regenerados usando métodos conhecidos na técnica (Horsch et al., 1985). A escolha das técnicas de transformação e/ou regeneração não é crítica para esta invenção.

Para dicotiledôneas, um sistema de vetor binário pode ser usado (Hoekema et al., 1983; EP 0120516 Schilperoort et al.). Por exemplo, cepas de Agrobacterium podem ser usadas que contém um plasmídeo vir com os genes de virulência e um plasmídeo compatível contendo o constructo de gene a ser transferida. Este vetor pode reproduzir em tanto E. coli quanto em Agrobacterium, e é derivado do vetor binário Bin19 (Bevan, 1984) que é alterado nos detalhes que não são relevantes a esta invenção. Os vetores binários quando usados neste exemplo contêm entre as seqüências da borda da esquerda e da direita do T-DNA, um gene de NPTII idêntico codificando para resistência de canamicina (Bevan, 1984) e um sítio de clonagem múltiplo para clonar nos constructos de gene requeridos.

A transformação e regeneração de plantações monocotiledôneas não são um procedimento padrão. Porém, recente progresso científico mostra que a princípio monocotiledôneas são amenas à transformação e aquelas plantas transgênicas férteis podem ser regeneradas das células transformadas. O desenvolvimento de sistemas de cultura de tecido reprodutíveis para estas plantações, junto com os métodos poderosos para introdução de material genético em células vegetais facilitou a transformação. Agora os métodos de escolha para transformação de monocotiledôneas são bombardeio de microprojétil de explantes ou células de suspensão, e absorção de DNA direta ou eletroporação de protoplastos. Por exemplo, plantas de arroz transgênicas foram prosperamente obtidas usando o gene de hph

bacteriano, codificando resistência de higromicina, como um marcador de seleção. O gene foi introduzido através de eletroporação (Shimamoto et al., 1993). Plantas de milho transgênicas foram obtidas introduzindo o gene de barra Streptomyces hygroscopicus, que codifica fosfinotricin acetiltransferase (uma enzima que inativa o fosfinotricina do herbicida), em células embriogênicas de uma cultura de suspensão de milho através de bombardeio de micropartícula (Gordon-Kamm et al., 1990). A introdução do material genético nos protoplastos de aleurona de outras plantações monocotiledôneas como trigo e cevada foi informada (Lee et al., 1989). Plantas de trigo foram regeneradas de cultura de suspensão embriogênica selecionando apenas o calo compacto e nodular embriogênico envelhecido para o estabelecimento das culturas de suspensão embriogênica (Vasil et al., 1972: Vasil et al., 1974). A combinação com sistemas de transformação para estas plantações permite a aplicação da presente invenção a monocotiledôneas. Estes métodos também podem ser aplicados para a transformação e regeneração de dicotiledôneas.

A expressão do constructo de fitase envolve tais detalhes como transcrição do gene por polimerases de planta, tradução de mRNA, etc. que são conhecidos às pessoas versadas na técnica das técnicas de DNA recombinante. Apenas detalhes relevantes para a compreensão adequada desta invenção são debatidos abaixo. Seqüências reguladoras que são conhecidas ou são observadas causar expressão de fitase podem ser usadas na presente invenção. A escolha das seqüências reguladoras usadas depende da plantação alvo e/ou órgão alvo de interesse. Tais seqüências reguladoras podem ser obtidas de plantas ou vírus de planta, ou podem ser quimicamente sintetizadas. Tais seqüências reguladoras são promotores ativos no direcionamento de transcrição em plantas, ou constitutivamente ou estágio e/ou tecido específicos, dependendo do uso da planta ou partes desta. Estes promotores incluem, mas não são limitados a promotores mostrando expressão constitutiva, como o promotor 35S de Vírus do Mosaico de Couve-flor (CaMV) (Guilley et al., 1982), aqueles para expressão folha-específica, como o promotor da ribulose bisfosfato carboxilase de gene de

subunidade pequena (Coruzzi et al., 1984), aqueles para expressão raiz-específica, como o promotor do gene de glutamina sintase (Tingey et al., 1987), aqueles para expressão semente-específica, como o promotor de cruciferina A de Brassica napus (Ryan et al., 1989), aqueles para expressão tubérculo-específica, como o promotor de patatina da classe I de batata (Koster-Topfer et al., 1989; Wenzler et al., 1989) ou aqueles para expressão fruta-específica, como o promotor de poligalacturonase (PG) de tomate (Bird et al., 1988).

Outras seqüências reguladoras como seqüências terminadoras e sinais de poliadenilação incluem qualquer tal seqüência que funciona como em plantas, a escolha desta está dentro do nível do versado. Um exemplo de tais seqüências é a região de flanqueamento 3' do gene de nopalina sintase (nos) de Agrobacterium tumefaciens (Bevan, supra). As seqüências reguladoras também podem incluir seqüências intensificadoras, como encontradas no promotor 35S de CaMV, e seqüências estabilizadoras de mRNA como a seqüência líder de RNA4 de Vírus do Mosaico de Alfaifa (AlMV) (Brederode et al., 1980) ou qualquer outra seqüência que funciona de uma maneira igual.

A fitase deve ser expressada em um ambiente que permite estabilidade da proteína expressada. A escolha de compartimentos celulares, como citosol, retículo endoplásmico, vacúolo, corpo de proteína ou espaço periplásmico podem ser usados na presente invenção para criar um tal ambiente estável, dependendo dos parâmetros biofísicos da fitase. Tais parâmetros incluem, mas não são limitados a pH-ótimo, sensibilidade à protease ou sensibilidade à molaridade do compartimento preferido.

Para obter expressão no citoplasma da célula, a enzima expressada não deve conter um peptídeo sinal secretório ou qualquer outra seqüência alvo. Para expressão em cloroplastos e mitocôndrias a enzima expressada deve conter peptídeo específico assim chamado de trânsito para importação nestas organelas. As seqüências de alvejamento que podem ser ligadas à enzima de interesse para alcançar esta são conhecidas (Smeekens et al., 1990; van der Broeck et al., 1985; Wolter et al., 1988). Se a ati-

vidade da enzima for desejada nos vacúolos, um peptídeo sinal secretório tem que estar presente, como também uma seqüência de alvejamento específica que direciona a enzima para estes vacúolos (Tague et al., 1990). O mesmo é verdade para os corpos de proteína nas sementes. A seqüência de DNA codificando a enzima de interesse deve ser modificada em um tal modo que a enzima pode exercer sua ação na localização desejada na célula.

Para alcançar expressão extracelular da fitase, o constructo de expressão da presente invenção utiliza uma seqüência sinal secretória. Embora seqüências sinais que são homólogas (nativas) às espécies de hospedeiro da planta podem ser preferidas, seqüências sinais heterólogas, isto é, aquelas que originam de outras espécies de planta ou de origem microbiana, podem ser usadas também. Tais seqüências sinais são conhecidas àqueles versados na técnica. As seqüências sinais apropriadas que podem ser usadas 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.

Todas as partes dos constructos de DNA relevantes (promotores, seqüências reguladoras, secretórias, estabilizantes, alvejantes ou de terminação) da presente invenção podem ser modificadas, se desejado, para afetar suas características de controle usando métodos conhecidos àqueles versados na técnica. É mostrado que plantas contendo fitase obtidas por meio da presente invenção podem ser usadas para obter plantas ou órgãos de planta com ainda níveis de fitase mais altos. Por exemplo, pode ser possível obter tal planta ou órgãos de planta pelo uso de técnicas de variação somoclonal ou através de técnicas de hibridação. Tais técnicas são bem-conhecidas àqueles versados na técnica.

Em um aspecto, a invenção atual fornece um método (e produtos deste) de alcançar um sistema de sobreexpressão altamente eficiente para fitase e outras moléculas. Em um aspecto, a invenção fornece um método (e produtos deste) de alcançar um sistema de sobreexpressão altamente eficiente para fitase e ácido fosfatase de pH 2,5 em Trichoderma. Este sistema resulta em composições de enzima tendo utilidade particular na

indústria de alimentação de animais. Detalhes adicionais relativos a este método estão na literatura pública e/ou são conhecidos ao versado. Em uma exemplificação não-limitativa particular, tal literatura publicamente disponível inclui EP 0659215 (WO 9403612 A1) (Nevalainen et al.), embora estas referências não ensinem as moléculas inventivas do presente pedido de patente.

Em outro aspecto, métodos podem ser usados para gerar novos polinucleotídeos que codificam vias bioquímicas de um ou mais operons ou agrupamentos de genes ou porções destes. Por exemplo, bactérias e muitos eucariotas têm um mecanismo coordenado para regular genes cujos produtos estão envolvidos em processos relacionados. Os genes são agrupados, em estruturas referidas como "agrupamentos de gene," em um cromossomo simples ou imediatamente adjacente a um outro e são transcritos juntos sob o controle de uma seqüência reguladora simples, incluindo um promotor simples que inicia a transcrição do agrupamento inteiro. Desse modo, um agrupamento de genes é um grupo de genes adjacentes que são idênticos ou relacionados, usualmente quanto a sua função. Um exemplo de uma via bioquímica codificada por agrupamentos de genes é policetídeos. Policetídeos são moléculas que são uma fonte extremamente rica de bioatividades, incluindo antibióticos (como tetraciclinas e eritromicina), agentes anti-câncer (daunomicina), imunossupressores (FK506 e rapamicina) e produtos veterinários (monensina). Muitos policetídeos (produzidos através de policetídeo sintases) são valiosos como agentes terapêuticos. Policetídeo sintases são enzimas multifuncionais que catalisam a biogênese de uma variedade enorme de cadeias de carbono que diferem no comprimento e padrões de funcionalidade e ciclização. Os genes de policetídeo sintase entram no agrupamento de genes e pelo menos um tipo (designado tipo I) de policetídeo sintases têm genes de tamanho grandes e enzimas, complicando a manipulação genética e estudos *in vitro* destes genes/proteínas.

DNA de agrupamento de genes pode ser isolado de diferentes organismos e ligado dentro dos vetores, particularmente vetores que contêm expressão de seqüências reguladoras que podem controlar e regular a produção de uma proteína detectável ou atividade de arranjo proteína-

relacionada dos agrupamentos de genes ligados. Uso de vetores que têm uma capacidade excepcionalmente grande para introdução de DNA exógeno é particularmente apropriado para uso com tal agrupamento de genes e é descrito aqui por via de exemplo para incluir o fator f (ou fator de fertilidade) de E. coli. Este fator f de E. coli é um plasmídeo que afeta a tradução de alta freqüência dela própria durante conjugação e é ideal para alcançar e estavelmente propagar fragmentos de DNA grandes, como agrupamentos de genes de amostras microbianas embaralhadas. Uma vez ligado em um vetor apropriado, dois ou mais vetores contendo diferentes agrupamentos de genes de fitase podem ser introduzidos em uma célula hospedeira adequada. As regiões de homologia de seqüência parcial compartilhadas pelos agrupamentos de gene promoverão processos que resultam na reorganização da seqüência resultando em um agrupamento de genes híbridos. O novo agrupamento de genes híbridos pode depois ser triado por atividades intensificadas não encontradas nos agrupamentos de genes originais.

Portanto, em um aspecto, a invenção diz respeito a um método para produzir um polipeptídeo híbrido biologicamente ativo e triar um tal polipeptídeo por atividade intensificada:

(1) introduzindo pelo menos um primeiro polinucleotídeo em ligação operável e um segundo polinucleotídeo em ligação operável, o dito pelo menos primeiro polinucleotídeo e segundo polinucleotídeo compartilhando pelo menos uma região de homologia de seqüência parcial, em uma célula hospedeira adequada;

(2) cultivar a célula hospedeira sob condições que promovem reorganização de seqüência resultando em um polinucleotídeo híbrido em ligação operável;

(3) expressar um polipeptídeo híbrido codificado pelo polinucleotídeo híbrido;

(4) triar o polipeptídeo híbrido sob condições que promovem identificação de atividade biológica intensificada; e

(5) isolar um polinucleotídeo codificando o polipeptídeo híbrido.

Os métodos para triar várias atividades de enzima são conheci-

dos àqueles de habilidade na técnica e são debatidos ao longo do presente relatório descritivo. Tais métodos podem ser empregados para isolar os polipeptídeos e polinucleotídeos da invenção.

Como exemplos representativos de vetores de expressão que podem ser usados podem ser mencionados partículas viróticas, baculovírus, fago, plasmídeos, fagemídeos, cosmídeos, fosmídeos, cromossomos bacterianos artificiais, DNA virótico (por exemplo, vacínia, adenovírus, vírus do epitelioma contagioso, pseudo-raiva e derivados de SV40), cromossomos artificiais com base em P1, plasmídeos de levedura, cromossomos artificiais de levedura e qualquer outro vetor específico para hospedeiros específicos de interesse (como bacilo, Aspergillus e levedura). Por exemplo, desse modo o DNA pode ser incluído em qualquer um de uma variedade de vetores de expressão para expressar um polipeptídeo. Tais vetores incluem seqüências de DNA cromossômicas, não-cromossômicas e sintéticas. Números grandes de vetores adequados são conhecidos àqueles de habilidade na técnica, e estão comercialmente disponíveis. Os vetores a seguir são fornecidos por via de exemplo; Bacteriano: vetores de pQE (Qiagen), plasmídeo de pBluescript, vetores de pNH, (vetores de lambda-ZAP (Stratagene)); ptrc99a, pKK223-3, pDR540, pRIT2T (Pharmacia); Eucariótico: pXT1, pSG5 (Stratagene), pSVK3, pBPV, pMSG, pSVLSV40 (Pharmacia). Porém, qualquer outro plasmídeo ou outro vetor pode ser usado desde que sejam replicáveis e viáveis no hospedeiro. Vetores de baixo número de cópias ou de número de cópias alto podem ser empregados com a presente invenção.

Um vetor exemplar para uso na presente invenção contém uma replicação de origem de fator f. O fator f (ou fator de fertilidade) em E. coli é um plasmídeo que realiza tradução de alta freqüência de si mesmo durante conjugação e tradução menos freqüente do próprio cromossomo bacteriano. Um aspecto usa vetores de clonagem, referidos como "fosmídeos" ou vetores de cromossomo artificial bacteriano (BAC). Estes são derivados de fator f de E. coli que é capaz de estavelmente integrar segmentos grandes de DNA genômico. Quando integrado com DNA de uma amostra ambiental não cultivada embaralhada, isto o torna possível alcançar fragmentos genômicos

grandes na forma de uma "biblioteca de DNA ambiental" estável.

Outro tipo de vetor para uso na presente invenção é um vetor de cosmídeo. Vetores de cosmídeo foram originalmente designados para clonar e propagar segmentos grandes de DNA genômico. A clonagem dentro de vetores de cosmídeo é descrita em detalhes em "MOLECULAR CLONING: A LABORATORY MANUAL" (Sambrook et al., 1989).

A seqüência de DNA no vetor de expressão é operativamente ligada a uma(s) seqüência(s) de controle de expressão apropriada(s) (promotor) para direcionar síntese de RNA. Os promotores bacterianos particulares nomeados incluem lacI, lacZ, T3, T7, gpt, lambda P_R, P_L e trp. Promotores eucarióticos incluem CMV precoce imediato, timidina cinase de HSV, SV40 precoce e tardio, LTRs de retrovírus e metalotioneína I de camundongo. A seleção do vetor e promotor apropriados está bem dentro do nível de capacidade usual na técnica. O vetor de expressão também contém um sítio de ligação de ribossomo para iniciação de tradução e um terminador de transcrição. O vetor também pode incluir seqüências apropriadas para amplificar expressão. Regiões de promotor de qualquer gene desejado podem ser selecionadas usando vetores de CAT (cloranfenicol transferase) ou outros vetores com marcadores selecionáveis. Além disso, os vetores de expressão podem conter um ou mais genes marcadores selecionáveis para fornecer uma característica fenotípica para seleção de células hospedeiras transformadas como diidrofolato reductase ou resistência à neomicina por cultura de células eucarióticas, ou resistência à tetraciclina ou ampicilina em E. coli.

Reclassificação *in vivo* é coletivamente focalizada em "processos intermoleculares" referidos como "recombinação" que em bactérias, em geral é visto como um "fenômeno RecA-dependente." A invenção pode contar com processos de recombinação de uma célula hospedeira para recombinar e reclassificar as seqüências, ou a capacidade das células para mediar processos redutivos para diminuir a complexidade das seqüências quase-repetidas na célula através de deleção. Este processo de "reclassificação redutiva" ocorre por um "processo intramolecular" RecA-

independente.

Portanto, em outro aspecto da invenção, polinucleotídeos variantes podem ser gerados pelo processo de reclassificação redutiva. O método envolve a geração de constructos contendo seqüências sucessivas (seqüências de codificação originais), sua inserção em um vetor apropriado e sua introdução subseqüente em uma célula hospedeira apropriada. A reclassificação das identidades moleculares individuais ocorre através de processos combinatórios entre as seqüências sucessivas nos constructos possuindo regiões de homologia, ou entre unidades quase-repetidas. O processo de reclassificação recombina e/ou reduz a complexidade e extensão das seqüências repetidas, e resulta na produção de novas espécies moleculares. Vários tratamentos para intensificar a taxa de reclassificação podem ser aplicados. Estes podem incluir tratamento com luz ultravioleta, ou químicas de danificação de DNA e/ou o uso de linhagem de células hospedeiras que apresenta, níveis intensificados de "instabilidade genética". Desse modo o processo de reclassificação pode envolver recombinação homóloga ou a propriedade natural das seqüências quase-repetidas para direcionar sua própria evolução.

Seqüências repetidas ou "quase-repetidas" representam um papel na instabilidade genética. Na presente invenção, "quase-repetições" são repetições que não são restritas a sua estrutura de unidade original. Unidades quase-repetidas podem estar presentes como um arranjo de seqüências em um constructo; unidades sucessivas de seqüências similares. Uma vez ligadas, as junções entre as seqüências sucessivas ficam essencialmente invisíveis e a natureza quase-repetitiva do constructo resultante é agora contínua no nível molecular. O processo de deleção que a célula executa para reduzir a complexidade do constructo resultante opera entre as seqüências quase-repetidas. As unidades quase-repetidas fornecem um repertório praticamente infinito de modelos nos quais os eventos de deslizamento podem ocorrer. Os constructos contendo as quase-repetições desse modo eficazmente fornecem elasticidade molecular suficiente que eventos de deleção (e potencialmente de inserção) podem ocorrer virtualmente em

qualquer lugar dentro das unidades quase-repetitivas.

Quando as seqüências quase-repetidas forem todas ligadas na mesma orientação, por exemplo, cabeça a cauda ou vice-versa, a célula não pode distinguir unidades individuais. Por conseguinte, o processo redutivo pode ocorrer ao longo das seqüências. Em contraste, quando, por exemplo, as unidades estão presentes de cabeça a cabeça, ao invés cabeça a cauda, a inversão delineia os pontos terminais da unidade adjacente de modo que a formação de deleção favorecerá a perda de unidades distintas. Assim, em um aspecto da invenção as seqüências estão na mesma orientação. A orientação aleatória de seqüências quase-repetidas resultará na perda da eficiência de reclassificação, enquanto a orientação consistente das seqüências oferecerá a eficiência mais alta. Porém, embora tendo poucas das seqüências contíguas na mesma orientação diminui a eficiência, elas ainda podem prover elasticidade suficiente pela recuperação eficaz de novas moléculas. Os constructos podem ser feitos com as seqüências quase-repetidas na mesma orientação para permitir eficiência mais alta.

As seqüências podem ser montadas em uma orientação cabeça para cauda usando qualquer de uma variedade de métodos, incluindo os seguintes:

(a) Iniciadores incluindo uma cabeça poli-A e cauda poli-T que quando feitos unifilamentados fornecem orientação, podem ser utilizados. Isto é realizado tendo as primeiras bases dos iniciadores feitas de RNA e conseqüentemente facilmente o RNAse H removida.

(b) Iniciadores incluindo sítios de clivagem de restrição únicos podem ser utilizados. Sítios múltiplos, uma bateria de seqüências únicas, e síntese repetida e etapas de ligação podem ser requeridas.

(c) As poucas bases internas do iniciador podem ser tioladas e uma exonuclease usada para produzir moléculas corretamente enfileiradas.

A recuperação das seqüências reclassificadas conta com a identificação de vetores de clonagem com um RI reduzido. As seqüências de codificação reclassificadas podem ser depois restabelecidas por amplificação. Os produtos são re-clonados e são expressados. A recuperação de

vetores de clonagem com RI reduzido pode ser realizada por:

1) O uso de vetores apenas estavelmente mantidos quando o constructo é reduzido na complexidade;

2) A recuperação física de vetores encurtados através de procedimentos físicos. Neste caso, o vetor de clonagem é restabelecido usando procedimentos de isolamento de plasmídeo padrão e fracionado no tamanho ou em um gel de agarose, ou coluna com um corte de baixo peso molecular utilizando procedimentos padrão;

3) A recuperação de vetores contendo genes suspensos que podem ser selecionados quando o tamanho da inserção diminui; e

4) O uso de técnicas de seleção diretas com um vetor de expressão e a seleção apropriada.

Seqüências de codificação (por exemplo, genes) de organismos relacionados podem demonstrar um grau alto de homologia e codificar produtos de proteína bastante diversos. Estes tipos de seqüências são particularmente úteis na presente invenção como quase-repetições. Porém, embora os exemplos ilustrados abaixo demonstrem a reclassificação das seqüências de codificação originais quase idênticas (quase-repetições), este processo não é limitado a tais repetições quase idênticas.

O exemplo a seguir demonstra um método da invenção. Seqüências de ácido nucléico codificando (quase-repetições) derivadas de três espécies únicas são descritas. Cada seqüência codifica uma proteína com um conjunto distinto de propriedades. Cada uma das seqüências diferem-se por um único ou alguns pares de base em uma posição única na seqüência que é designada "A", "B" e "C." As seqüências quase-repetidas são separadas ou coletivamente amplificadas e ligadas em conjuntos fortuitos de modo que todas as possíveis permutações e combinações estejam disponíveis na população de moléculas ligadas. O número de unidades de quase-repetição pode ser controlado pelas condições de conjunto. O número comum de unidades quase-repetidas em um constructo é definido como o índice repetitivo (RI).

Uma vez formados, os constructos podem ou não ser fraciona-

dos no tamanho em um gel de agarose de acordo com os protocolos publicados, inseridos em um vetor de clonagem e transfeccionados em uma célula hospedeira apropriada. As células são depois propagadas e "reclassificação redutiva" é realizada. A taxa do processo de reclassificação redutiva pode ser estimulada pela introdução de dano de DNA, se desejado. Se a redução em RI é mediada através de formação de deleção entre as seqüências repetidas por um "mecanismo intramolecular", ou mediada por eventos semelhantes à recombinação através de mecanismos "intermoleculares" é imaterial. O resultado final é uma reclassificação das moléculas em todas as possíveis combinações.

Opcionalmente, o método compreende a etapa adicional de triar os membros de biblioteca do fundo comum embaralhado para identificar os membros da biblioteca embaralhados individuais tendo a capacidade de ligar ou do contrário interagir, ou catalisar uma reação particular (por exemplo, como catalisar a hidrólise de um fitato).

Os polipeptídeos que são identificados de tais bibliotecas podem ser usados para propósitos terapêuticos, de diagnóstico, de pesquisa e relacionados (por exemplo, catalisadores, solutos para osmolaridade crescente de uma solução aquosa e outros), e/ou podem ser submetidos a um ou mais ciclos adicionais de embaralhamento e/ou seleção.

Em outro aspecto, antes ou durante a recombinação ou reclassificação, os polinucleotídeos da invenção ou polinucleotídeos gerados pelo método aqui descrito podem ser submetidos aos agentes ou processos que promovem a introdução de mutações nos polinucleotídeos originais. A introdução de tais mutações aumentaria a diversidade dos polinucleotídeos híbridos resultantes e dos polipeptídeos deles codificados. Os agentes ou processos que promovem mutagênese podem incluir, mas não são limitados: (+)-CC-1065, ou um análogo sintético como (+)-CC-1065-(N3-adenina, ver Sun e Hurley, 1992); um aduto N-acetilado ou 4'-flúor-4-aminobifenil-desacetialdo capaz de inibir síntese de DNA (ver, por exemplo, van de Poll et al., 1992); ou um aduto N-acetilado ou 4-aminobifenil-desacetilado capaz de inibir síntese de DNA (também ver, van de Poll et al., 1992, pp. 751-758);

cromo trivalente, um sal de cromo trivalente, um aduto de DNA de hidrocarboneto aromático policíclico ("PAH") capaz de inibir replicação de DNA, 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-diidrodiol-9-10-epóxido ("BPDE"), um sal de halogênio de platina(II), N-hidróxi-2-amino-3-metilimidazo[4,5-f]-quinolina ("N-hidróxi-I.Q") e N-hidróxi-2-amino-1-metil-6-fenilimidazo[4,5-f]-piridina ("N-hidróxi-PhIP"). Uns meios exemplares para reduzir ou deter a amplificação de PCR consistem em luz UV (+)-CC-1065 e (+)-CC-1065-(N3-Adenina). Meios particularmente abrangidos são adutos de DNA ou polinucleotídeos compreendendo os adutos de DNA dos polinucleotídeos ou fundo comum de polinucleotídeos, que podem ser liberados ou removidos por um processo incluindo aquecer a solução compreendendo os polinucleotídeos antes de processamento adicional.

Em outro aspecto, a invenção é direcionada a um método de produzir proteínas recombinantes tendo atividade biológica tratando uma amostra compreendendo polinucleotídeos modelos bifilamentados codificando uma proteína do tipo selvagem sob condições de acordo com a invenção que fornecem a produção de polinucleotídeos híbridos ou reclassificados.

A invenção também fornece o uso de iniciadores de códon proprietário (contendo uma seqüência de N,N,G/T degenerada) para introduzir mutações de ponto dentro de um polinucleotídeo, para gerar um conjunto de polipeptídeos da progênie em que uma faixa total de substituições de aminoácidos simples é representada em cada posição de aminoácido (mutagênese saturada no sítio genético (GSSM)). Os oligos usados são compreendidos de forma contígua de uma primeira seqüência homóloga, uma seqüência de N,N,G/T degenerada, e opcionalmente uma segunda seqüência homóloga. Produtos translacionais da progênie a jusante do uso de tais oligos incluem todas as possíveis alterações em cada sítio de aminoácido ao longo do polipeptídeo, porque a degeneração da seqüência de N,N,G/T inclui códon para todos os 20 aminoácidos.

Em um aspecto, um tal oligo degenerado (compreendido de um

cassete de N,N,G/T degenerado) é usado para submeter cada códon original em um modelo de polinucleotídeo parental a uma faixa total de substituições de códons. Em outro aspecto, pelo menos dois cassetes de N,N,G/T degenerados são usados - ou no mesmo oligo ou não, para submeter pelo menos dois códons originais em um modelo de polinucleotídeo parental a uma faixa total de substituições de códons. Desse modo, mais de uma seqüência de N,N,G/T pode ser contida em um oligo para introduzir mutações de aminoácidos em mais de um sítio. Esta pluralidade de seqüências de N,N,G/T pode ser diretamente contígua, ou separada por uma ou mais seqüências de nucleotídeos adicionais. Em outro aspecto, oligos úteis para introduzir adições e deleção ou podem ser usados sozinho ou em combinação com os códons contendo uma seqüência de N,N,G/T, para introduzir qualquer combinação ou permutação de adições, deleção, e/ou substituições de aminoácidos.

Em um aspecto, é possível simultaneamente submeter à mutagênese duas ou mais posições de aminoácidos contíguas usando um oligo contendo os tripletos de N,N,G/T contíguos, isto é uma seqüência degenerada (N,N,G/T)n.

Em outro aspecto, a presente invenção fornece o uso de cassetes degenerados tendo menos degeneração que a seqüência de N,N,G/T. Por exemplo, pode ser desejável em algumas circunstâncias usar (por exemplo em um oligo) uma seqüência tripleto degenerada compreendida de apenas um N onde o dito N pode estar na primeira, segunda ou terceira posição do tripleto. Qualquer outra base que inclui qualquer combinação e permutações desta pode ser usada no permanecendo duas posições do tripleto. Alternativamente, pode ser desejável em algumas circunstâncias usar (por exemplo, em um oligo) uma seqüência do tripleto de N,N,N degenerada, ou uma seqüência do tripleto N,N, G/C.

Porém, é apreciado que o uso de um tripleto degenerado (como N,N,G/T ou uma seqüência do tripleto N,N,G/C) como divulgado na invenção atual é vantajoso por várias razões. Em um aspecto, esta invenção fornece uns meios para sistemática e razoavelmente fácil gerar a substituição da faixa de completa de possíveis aminoácidos (para um total de 20 aminoáci-

dos) em cada e toda posição de aminoácidos em um polipeptídeo. Desse modo, para 100 polipeptídeos de aminoácido, a invenção fornece um modo sistemático e razoavelmente fácil para gerar 2000 espécies distintas (isto é, 20 possíveis aminoácidos por posição vezes 100 posições de aminoácidos). É apreciado que é fornecido, através do uso de um oligo contendo um N,N,G/T ou uma seqüência do tripleto degenerado N,N,G/C, 32 seqüências individuais codificando 20 possíveis aminoácidos. Desse modo, em um vaso de reação em que uma seqüência de polinucleotídeo parental é submetida à mutagênese de saturação usando um tal oligo, são gerados 32 polinucleotídeos distintos da progênie que codificam 20 polipeptídeos distintos. Em contraste, o uso de um oligo não-degenerado em mutagênese sítio-direcionada leva a apenas um produto de polipeptídeos da progênie por vaso de reação.

Esta invenção também fornece o uso de oligos não-degenerados, que podem opcionalmente ser usados em combinação com iniciadores degenerados divulgados. É apreciado que em algumas situações, é vantajoso usar oligos não-degenerados para gerar mutações de ponto-específicas em um polinucleotídeo de funcionamento. Isto fornece um meio para gerar mutações de ponto silencioso-específicas, mutações de ponto que conduzem às alterações de aminoácidos correspondentes e mutações de ponto que causam a geração de códons terminais e a expressão correspondente de fragmentos de polipeptídeos.

Desse modo, em um aspecto, cada vaso de reação de mutagênese de saturação contém polinucleotídeos codificando pelo menos 20 moléculas de polipeptídeos da progênie de modo que todos os 20 aminoácidos são representados em uma posição de aminoácido específico que corresponde à posição do códon submetido à mutagênese no polinucleotídeo parental. Os polipeptídeos da progênie degenerada 32 vezes gerados de cada vaso de reação de mutagênese de saturação podem ser submetidos à amplificação clonal (por exemplo, clonado em um hospedeiro adequado de E. coli usando um vetor de expressão) e submetido à triagem de expressão. Quando um polipeptídeo da progênie individual é identificado mediante triagem para exibir uma alteração favorável na propriedade (quando comparado

ao polipeptídeo parental), ele pode ser seqüenciado para identificar a substituição de aminoácido correspondentemente favorável nele contido.

É apreciado que ao transformar cada e toda posição de aminoácido em um polipeptídeo parental usando mutagênese de saturação como aqui divulgado, alterações de aminoácido favoráveis podem ser identificadas em mais de uma posição de aminoácido. Uma ou mais moléculas da progênie novas podem ser geradas que contém uma combinação de toda ou parte destas substituições de aminoácido favoráveis. Por exemplo, se são identificadas 2 alterações de aminoácido favoráveis específicas em cada uma de 3 posições de aminoácido em um polipeptídeo, as permutações incluem 3 possibilidades em cada posição (nenhuma alteração do aminoácido original, e cada de duas alterações favoráveis) e 3 posições. Desse modo, há 3 x 3 x 3 ou 27 possibilidades totais, incluindo 7 que foram previamente examinadas - 6 mutações de ponto simples (isto é, 2 em cada uma das três posições) e nenhuma alteração em qualquer posição.

Em ainda outro aspecto, mutagênese de saturação de sítio pode ser usada junto com embaralhamento, quimerização, recombinação e outros processos de metagênese, junto com triagem. Esta invenção fornece o uso de qualquer/quaisquer do(s) processos(s) de metagênese, incluindo mutagênese de saturação, de uma maneira iterativa. Em uma exemplificação, o uso iterativo de qualquer/quaisquer do(s) processo(s) de metagênese é usado em combinação com triagem.

Desse modo, em uma exemplificação não-limitativa, os polinucleotídeos (por exemplo, SEQ ID Nº: 1) e polipeptídeos (por exemplo, SEQ ID Nº: 2) da invenção podem ser derivados através de mutagênese de saturação em combinação com processos adicionais de mutagênese, como processo onde dois ou mais polinucleotídeos relacionados são introduzidos em uma célula hospedeira adequada de modo que um polinucleotídeo híbrido é gerado por recombinação e reclassificação redutiva.

Além de executar mutagênese ao longo da seqüência inteira de um gene, pode ser usada mutagênese para substituir cada uma de qualquer número de bases em uma seqüência de polinucleotídeos, em que o número

de bases a ser submetido à mutagênese pode ser todo número inteiro de 15 a 100.000. Desse modo, em vez de submeter à mutagênese cada posição ao longo de uma molécula, pode-se submeter cada ou um número distinto de bases (pode ser uma totalização de subconjunto de 15 a 100.000) à mutagênese. Um nucleotídeo separado pode ser usado para transformar cada posição ou grupo de posições ao longo de uma seqüência de polinucleotídeos. Um grupo de 3 posições para ser submetido à mutagênese pode ser um códon. As mutações podem ser introduzidas usando um iniciador mutagênico, contendo um cassete heterólogo, também referido como um cassete mutagênico. Cassetes exemplares podem ter de 1 a 500 bases. Cada posição de nucleotídeo em tais cassetes heterólogos 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 onde E é qualquer base que não seja A, C, G ou T (E pode ser referido como um oligo modelo).

Em um sentido geral, mutagênese de saturação é compreendida de submissão à metagênese de um conjunto completo de cassetes mutagênicos (em que cada cassete pode ser cerca de 1-500 bases no comprimento) na seqüência de polinucleotídeos definida a ser submetida à mutagênese (em que a seqüência a ser submetida à mutagênese pode ser de cerca de 15 a 100.000 bases no comprimento). Desse modo, um grupo de mutações (variando de 1 a 100 mutações) é introduzido em cada cassete a ser submetido à mutagênese. Um agrupamento de mutações a ser introduzido em um cassete pode ser diferente ou o mesmo de um segundo agrupamento de mutações a ser introduzido em um segundo cassete durante a aplicação de um ciclo de mutagênese de saturação. Tais agrupamentos são exemplificados através de deleção, adições, agrupamentos de códons particulares, e agrupamentos de cassetes de nucleotídeo particulares.

Seqüências definidas a serem submetidas à mutagênese incluem um gene total, via, cDNA, uma estrutura de leitura aberta total (ORF) e o promotor total, intensificador, repressor/transativador, origem de replicação, íntron, operador ou qualquer grupo funcional de polinucleotídeo. Em geral, umas "seqüências definidas" para este propósito podem ser qualquer polinucleotídeos que uma seqüência de polinucleotídeos de 15 bases, e seqüên-

cias de polinucleotídeos de comprimento entre 15 bases e 15.000 bases (esta invenção especificamente nomeia cada número inteiro nos intervalos). Considerações na seleção dos agrupamentos de códon incluem tipos de aminoácidos codificados por um cassete mutagênico degenerado.

Em um aspecto, um agrupamento de mutações que pode ser introduzido em um cassete mutagênico, esta invenção especificamente fornece para substituições de códons degenerados (usando oligos degenerados) que codificam para 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 polipeptídeos assim codificada.

Um aspecto da invenção é um ácido nucléico isolado compreendendo SEQ ID Nº: 1, seqüências substancialmente idênticas a esta, seqüências complementares a esta ou um fragmento compreendendo pelo menos 10, 15, 20, 25, 30, 35, 40, 50, 75, 100, 150, 200, 300, 400 ou 500 bases sucessivas de uma das seqüências da SEQ ID Nº: 1. Os ácidos nucléicos isolados podem compreender DNA, incluindo cDNA, DNA genômico, e o DNA sintético. O DNA pode ser bifilamentado ou unifilamentado, e se unifilamentado pode ser o filamento de codificação ou o filamento de não-codificação (anti-sentido). Alternativamente, os ácidos nucléicos isolados podem compreender RNA.

Como debatido em mais detalhe abaixo, as seqüências de ácidos nucléicos isoladas da invenção podem ser usadas preparar para o polipeptídeo da SEQ ID Nº: 2, e seqüências substancialmente idênticas a esta, ou fragmentos compreendendo pelo menos 5, 10, 15, 20, 25, 30, 35, 40, 50, 75, 100 ou 150 aminoácidos sucessivos de um dos polipeptídeos da SEQ ID Nº: 2, e seqüências substancialmente idênticas a esta.

Conseqüentemente, outro aspecto da invenção é uma seqüência de ácidos nucléicos isolada que codifica um dos polipeptídeos da SEQ ID Nº: 2, seqüências substancialmente idênticas a esta, ou fragmentos compreendendo pelo menos 5, 10, 15, 20, 25, 30, 35, 40, 50, 75, 100 ou 150 aminoácidos sucessivos de um dos polipeptídeos da SEQ ID Nº: 2. As seqüências de codificação destes ácidos nucléicos podem ser idênticas a uma das

seqüências de codificação da SEQ ID Nº: 1, ou um fragmento desta, ou podem ser seqüências de codificação diferentes que codificam um dos polipeptídeos da SEQ ID Nº: 2, e seqüências substancialmente idênticas a esta, e fragmentos tendo pelo menos 5, 10, 15, 20, 25, 30, 35, 40, 50, 75, 100 ou 150 aminoácidos sucessivos de um dos polipeptídeos da SEQ ID Nº: 2 como resultado da redundância ou degeneração do código genético. O código genético é bem-conhecido àqueles versados na técnica e é obtido, por exemplo, na página 214 de B. Lewin, Genes VI, Oxford Universidade Press, 1997.

A seqüência de ácidos nucléicos isolada que codifica um dos polipeptídeos da SEQ ID Nº: 2, e seqüências substancialmente idênticas a esta, pode incluir, mas não é limitada a apenas uma seqüência de codificação de uma da SEQ ID Nº: 1, e seqüências substancialmente idênticas a esta, e seqüências de codificação adicionais, como seqüências líderes ou seqüências de proproteína e seqüências de não-codificação, como íntrons ou seqüências de não-codificação 5' e/ou 3' da seqüência de codificação. Desse modo, como aqui usado, o termo "polinucleotídeo codificando um polipeptídeo" abrange um polinucleotídeo que inclui a seqüência de codificação apenas para o polipeptídeo como também um polinucleotídeo que inclui seqüência de codificação e/ou de não-codificação adicional(is).

Alternativamente, as seqüências de ácidos nucléicos da invenção podem ser submetidas à mutagênese usando técnicas convencionais, como mutagênese sítio direcionada, ou outras técnicas familiares àqueles versados na técnica, para introduzir alterações silenciosas no polinucleotídeo da SEQ ID Nº: 1, e seqüências substancialmente idênticas a esta. Como aqui usado, "alterações silenciosas" incluem, por exemplo, alterações que não alteram a seqüência de aminoácidos codificada pelo polinucleotídeo. Tais alterações podem ser desejáveis para aumentar o nível do polipeptídeo produzido por células hospedeiras contendo um vetor que codifica o polipeptídeo introduzindo códon ou pares de códon que freqüentemente ocorrem no organismo hospedeiro.

A invenção também diz respeito aos polinucleotídeos tendo alterações de nucleotídeo que resultam em substituições, adições, deleção, fu-

sões e mutilações de aminoácidos, nos polipeptídeos da invenção (por exemplo, SEQ ID Nº: 2). Podem ser introduzidas tais alterações de nucleotídeo usando técnicas como mutagênese sitio-direcionada, mutagênese química aleatória, deleção de exonuclease III, e outras técnicas de DNA recombinante. Alternativamente, tais alterações de nucleotídeo podem ser variantes alélicas de ocorrência natural que são isoladas para identificar as seqüências de ácidos nucléicos que especificamente hibridam com as sondas compreendendo pelo menos 10, 15, 20, 25, 30, 35, 40, 50, 75, 100, 150, 200, 300, 400 ou 500 bases sucessivas de uma das seqüências da SEQ ID Nº: 1, e seqüências substancialmente idênticas a esta, (ou as seqüências complementares a esta), sob condições de severidade alta, moderada ou baixa como aqui fornecida.

Os ácidos nucléicos isolados da SEQ ID Nº: 1, seqüências substancialmente idênticas a esta, seqüências complementares ou um fragmento compreendendo pelo menos 10, 15, 20, 25, 30, 35, 40, 50, 75, 100, 150, 200, 300, 400 ou 500 bases sucessivas de uma das seqüências anteriores, podem ser também usados como sondas para determinar se uma amostra biológica, como uma amostra do solo, contém um organismo tendo uma seqüência de ácidos nucléicos da invenção ou um organismo do qual o ácido nucléico foi obtido. Em tais procedimentos, uma amostra biológica que abriga o organismo do qual o ácido nucléico foi potencialmente isolado é obtida e os ácidos nucléicos da amostra são obtidos. Os ácidos nucléicos são contatados com a sonda sob condições que permitem a sonda especificamente hibridar com qualquer seqüência complementar que está nela presente.

Onde necessário, condições que permitem a sonda especificamente hibridar com seqüências complementares podem ser determinadas colocando a sonda em contato com seqüências complementares de amostras conhecidas para conter a seqüência complementar como também seqüências de controle que não contêm a seqüência complementar. Condições de hibridação, como a concentração de sal do tampão de hibridação, a concentração de formamida do tampão de hibridação ou a temperatura de hibri-

dação, podem ser variadas para identificar as condições que permitem a sonda especificamente hibridar com os ácidos nucléicos complementares.

Se a amostra contiver o organismo do qual o ácido nucléico foi isolado, a hibridação específica da sonda é então detectada. Hibridação pode ser detectada marcando a sonda com um agente detectável como um isótopo radioativo, uma tintura fluorescente ou uma enzima capaz de catalisar a formação de um produto detectável.

Muitos métodos para usar as sondas marcadas para detectar a presença de ácidos nucléicos complementares em uma amostra são familiares àqueles versados na técnica. Estes incluem Southern blots, Northern blots, procedimentos de hibridação de colônia, e mancha de ponto. Os protocolos para cada um destes procedimentos são fornecidos em Ausubel et al. Current Protocols in Molecular Biology, John Wiley 503 Sons, Inc., 1997 e Sambrook et al., Molecular Cloning: A Laboratory Manual, 2d Ed., Cold Spring Harbor Laboratory Press, 1989.

Alternativamente, mais de uma sonda (pelo menos uma das quais é capaz de especificamente hibridar com qualquer seqüência complementar que está presente na amostra de ácido nucléico), pode ser usada em uma reação de amplificação para determinar se a amostra contém um organismo contendo uma seqüência de ácidos nucléicos da invenção (por exemplo, um organismo do qual o ácido nucléico foi isolado). Tipicamente, as sondas compreendem oligonucleotídeos. Em um aspecto, a reação de amplificação pode compreender uma reação de PCR. Os protocolos de PCR são descritos em Ausubel e Sambrook, supra. Alternativamente, a amplificação pode compreender uma reação em cadeia de ligase, 3SR, ou reação de deslocamento de filamento. (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 procedimentos, os ácidos nucléicos na

amostra são contatados com as sondas, a reação de amplificação é executada, e qualquer produto de amplificação resultante é detectado. O produto de amplificação pode ser detectado executando eletroforese em gel nos produtos de reação e manchando o gel com um intercalador como brometo de etídio. Alternativamente, uma ou mais das sondas podem ser marcadas com um isótopo radioativo e a presença de um produto de amplificação radioativo pode ser detectada através de auto-radiografia após eletroforese em gel.

Sondas derivadas das seqüências próximas às extremidades de uma seqüência como exposto na SEQ ID Nº: 1, e seqüências substancialmente idênticas a esta, também podem ser usadas em procedimentos de passeio pelos cromossomos para identificar clones que contêm seqüências genômicas localizadas adjacentes às seqüências de ácidos nucléicos como exposto acima. Tais métodos permitem o isolamento de genes que codificam proteínas adicionais do organismo hospedeiro.

Uma seqüência de ácidos nucléicos isolada conforme exposto na SEQ ID Nº: 1, seqüências substancialmente idênticas a esta, seqüências complementares a esta ou um fragmento compreendendo pelo menos 10, 15, 20, 25, 30, 35, 40, 50, 75,100,150, 200, 300, 400 ou 500 bases sucessivas de uma das seqüências anteriores podem ser usados como sondas para identificar e isolar os ácidos nucléicos relacionados. Em alguns aspectos, os ácidos nucléicos relacionados podem ser cDNAs ou os DNAs genômicos de organismos diferentes do qual o ácido nucléico foi isolado. Por exemplo, os outros organismos podem estar relacionados aos organismos. Em tais procedimentos, uma amostra de ácido nucléico é contatada com a sonda sob condições que permitem a sonda especificamente hibridar com as seqüências relacionadas. Hibridação da sonda com ácidos nucléicos do organismo relacionado é depois detectada usando quaisquer dos métodos acima descritos.

Em reações de hibridação de ácido nucléico, as condições usadas para alcançar um nível particular de severidade variarão, dependendo da natureza dos ácidos nucléicos que são hibridados. Por exemplo, o comprimento, grau de complementaridade, composição de seqüência de nucleo-

tídeos (por exemplo, GC V. teor de AT) e o tipo de ácido nucléico (por exemplo, RNA v. DNA) das regiões de hibridação dos ácidos nucléicos podem ser considerados na seleção das condições de hibridação. Uma consideração adicional é se um dos ácidos nucléicos for imobilizado, por exemplo, em um filtro.

Hibridação pode ser realizada sob condições de baixa severidade, severidade moderada ou alta severidade. Como um exemplo de hibridação de ácido nucléico, uma membrana de polímero contendo ácidos nucléicos desnaturados imobilizados é primeiro pré-hidrolisada durante 30 minutos a 45°C em uma solução que consiste em 0,9 M NaCl, 50 mM NaH_2PO_4, pH 7,0, 5,0 mM Na_2EDTA, 0,5 % de SDS, 10X do Denhardt e 0,5 mg/ml de ácido polirriboadenílico. Aproximadamente 2 x 10^7 cpm (atividade específica 4-9 x 108 cpm/ug) de sonda de oligonucleotídeo marcada na extremidade com ^{32}P são depois adicionadas à solução. Após 12-16 horas de incubação, a membrana é lavada durante 30 minutos em temperatura ambiente em 1X SET (150 mM NaCl, 20 mM cloridrato de Tris, pH 7,8, 1 mM Na_2EDTA) contendo 0,5 % de SDS, seguido por uma lavagem de 30 minutos em 1X SET fresco a Tm-10°C para a sonda de oligonucleotídeo. A membrana é depois exposta à película auto-radiográfica para detecção de sinais de hibridação.

Variando a severidade das condições de hibridação usadas para identificar os ácidos nucléicos, como cDNAs ou DNAs genômicos, que hibridizam com sonda detectável, ácidos nucléicos tendo níveis diferentes de homologia para a sonda podem ser identificados e isolados. A severidade pode ser variada conduzindo a hibridação em temperaturas variadas abaixo das temperaturas de derretimento das sondas. A temperatura de derretimento, T_m, é a temperatura (sob força iônica definida e pH) na qual 50 % da seqüência alvo cruzam com uma sonda perfeitamente complementar. Condições muito rigorosas são selecionadas para ser iguais a ou aproximadamente 5°C abaixo da T_m para uma sonda particular. A temperatura de derretimento da sonda pode ser calculada usando as fórmulas a seguir: Para sondas entre 14 e 70 nucleotídeos no comprimento a temperatura de derretimento (T_m) é calculada usando a fórmula: $T_m = 81,5 + 16,6(\log[Na+])+0,41$

(fração G+C) - (600/N), onde N é o comprimento da sonda. Se a hibridação for realizada em uma solução contendo formamida, a temperatura de derretimento pode ser calculada usando a equação: T_m = 81,5 + 16,6(log[Na+])+0,41 (fração G+C) - (0,63 % de formamida) - (600/N), onde N é o comprimento da sonda. Pré-hibridação pode ser realizada em 6X SSC, 5X o reagente de Denhardt, 0,5 % de SDS, 100 µg de DNA de esperma de salmão fragmentado desnaturado ou 6X SSC, 5X o reagente de Denhardt, 0,5 % de SDS, 100 µg DNA de esperma de salmão fragmentado desnaturado, 50 % de formamida. As fórmulas para SSC e as soluções de Denhardt podem ser encontradas, por exemplo, em Sambrook et al., supra.

A hibridação é conduzida adicionando a sonda detectável às soluções de pré-hibridação acima listadas. Onde a sonda compreender DNA bifilamentado, ela é desnaturada antes da adição à solução de hibridação. O filtro é contactado com a solução de hibridação durante um período suficiente de tempo para permitir a sonda hibridar com cDNAs ou DNAs genômicos que contêm seqüências complementares a esta ou homólogas a estas. Para sondas de mais de 200 nucleotídeos no comprimento, a hibridação pode ser realizada a 15-25°C abaixo da Tm. Para sondas mais curtas, como sondas de oligonucleotídeo, a hibridação pode ser conduzida a 5-10°C abaixo da Tm. Tipicamente, para hibridações em 6X SSC, a hibridação é conduzida a cerca de 68°C. Usualmente, para hibridações em 50 % de soluções contendo formamida, a hibridação é conduzida a cerca de 42°C. É considerado que todas as hibridações anteriores estão sob condições de alta severidade.

Seguindo a hibridação, o filtro é lavado para não-especificamente remover qualquer sonda detectável ligada. A severidade usada para lavar os filtros pode também ser variada dependendo da natureza dos ácidos nucléicos sendo hibridados, do comprimento dos ácidos nucléicos sendo hibridados, do grau de complementaridade, da composição da seqüência de nucleotídeos (por exemplo, GC V. teor de AT) e do tipo de ácido nucléico (por exemplo, RNA V. DNA). Exemplos de lavagens de condição de severidade progressivamente mais alta são como segue: 2X SSC, 0,1 % de SDS

em temperatura ambiente durante 15 minutos (baixa severidade); 0,1X SSC, 0,5 % de SDS em temperatura ambiente durante 30 minutos a 1 hora (severidade moderada); 0,1X SSC, 0,5 % de SDS durante 15 a 30 minutos entre a temperatura de hibridação e 68°C (severidade alta); e 0,15M NaCl durante 15 minutos a 72°C (severidade muito alta). Uma lavagem de baixa severidade final pode ser conduzida em 0,1X SSC em temperatura ambiente. Os exemplos acima são meramente ilustrativos de um conjunto de condições que pode ser usado para lavar os filtros. Alguém versado na técnica saberia que há numerosas receitas para lavagens de diferentes severidades. Alguns outros exemplos são dados abaixo.

Ácidos nucléicos que foram hibridados com a sonda podem ser identificados por auto-radiografia ou outras técnicas convencionais.

O procedimento acima pode ser modificado para identificar ácidos nucléicos tendo níveis decrescentes de homologia para a seqüência de sonda. Por exemplo, para obter ácidos nucléicos de homologia decrescente para sonda detectável, condições menos rigorosas podem ser usadas. Por exemplo, a temperatura de hibridação pode ser diminuída em incrementos de 5°C de 68°C a 42°C em um tampão de hibridação tendo uma concentração de Na+ de cerca de 1 M. Seguindo a hibridação, o filtro pode ser lavado com 2X SSC, 0,5 % de SDS na temperatura de hibridação. É considerado que estas condições são "condições moderadas" acima de 50°C e condições "baixas" abaixo de 50°C. Um exemplo específico de "condições de hibridação moderadas" é quando a hibridação acima é conduzida a 55°C. Um exemplo específico de condições de hibridação de "baixa severidade" é quando a hibridação acima é conduzida a 45°C.

Alternativamente, a hibridação pode ser realizada em tampões, como 6X SSC, contendo em uma temperatura de formamida de 42°C. Neste caso, a concentração de formamida no tampão de hibridação pode ser reduzida em 5 % incrementos de 50 % a 0 % para identificar clones tendo níveis decrescentes de homologia para a sonda. Seguindo a hibridação, o filtro pode ser lavado com 6X SSC, 0,5 % de SDS a 50°C. É considerado que estas condições são 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 "baixa severidade" é quando a hibridação acima é conduzida a 10 % de formamida.

Por exemplo, os métodos precedentes podem ser usados para isolar ácidos nucléicos tendo uma seqüência com pelo menos cerca de 99 %, pelo menos 98 %, pelo menos 97 %, pelo menos 95 %, pelo menos 90 %, ou pelo menos 80 % de homologia a uma seqüência de ácidos nucléicos como exposto na SEQ ID Nº: 1, seqüências substancialmente idênticas a esta, ou fragmentos compreendendo pelo menos cerca de 10, 15, 20, 25, 30, 35, 40, 50, 75, 100, 150, 200, 300, 400, ou 500 bases sucessivas desta, e as seqüências complementares a quaisquer das seqüências anteriores. Homologia pode ser medida usando um algoritmo de alinhamento. Por exemplo, os polinucleotídeos homólogos podem ter uma seqüência de codificação que é uma variante alélica de ocorrência natural de uma das seqüências de codificação aqui descritas. Tais variantes alélicas podem ter uma substituição, deleção ou adição de um ou mais nucleotídeos quando comparadas a uma seqüência de ácidos nucléicos como exposto na SEQ ID Nº: 1, ou seqüências complementares a esta.

Adicionalmente, os procedimentos acima podem ser usados para isolar ácidos nucléicos codificando polipeptídeos tendo pelo menos cerca de 99 %, pelo menos 95 %, pelo menos 90 %, pelo menos 85 %, pelo menos 80 % ou pelo menos 70 % de homologia a um polipeptídeo tendo uma seqüência como exposto na SEQ ID Nº: 2, seqüências substancialmente idênticas a esta, ou fragmentos compreendendo pelo menos 5, 10, 15, 20, 25, 30, 35, 40, 50, 75, 100, ou 150 aminoácidos sucessivos desta como determinado usando um algoritmo de alinhamento de seqüência (por exemplo, como o algoritmo de FASTA versão 3.0t78 com os parâmetros predefinidos).

Outro aspecto da invenção é um polipeptídeo isolado ou purificado compreendendo uma seqüência como exposto na SEQ ID Nº: 1, se-

qüências substancialmente idênticas a esta, ou fragmentos compreendendo pelo menos cerca de 5, 10, 15, 20, 25, 30, 35, 40, 50, 75, 100, ou 150 aminoácidos sucessivos desta. Como debatido acima, tais polipeptídeos podem ser obtidos inserindo um ácido nucléico codificando o polipeptídeo em um vetor de modo que a seqüência de codificação é operavelmente ligada a uma seqüência capaz de dirigir a expressão do polipeptídeo codificada em uma célula hospedeira adequada. Por exemplo, o vetor de expressão pode compreender um promotor, um sítio de ligação de ribossomo para iniciação de tradução e um terminador de transcrição. O vetor também pode incluir seqüências apropriadas para amplificar a expressão.

Promotores adequado para expressar o polipeptídeo ou fragmento deste em bactérias incluem os promotores de lac ou trp de E. coli, o promotor de lacI, o promotor de lacZ, o promotor de T3, o promotor de T7, o promotor de gpt, o promotor de lambda PR, o promotor de lambda PL, os promotores de operons que codificam enzimas glicolíticas como 3-fosfoglicerato cinase (PGK) e o promotor de ácido fosfatase. Os promotores fúngicos incluem um promotor de fator α. Promotores eucarióticos incluem o promotor de CMV precoce imediato, o promotor de timidina cinase de HSV, promotores de choque térmico, o promotor de SV40 precoce e recente, LTRs de retrovírus e o promotor de metalotioneína I de camundongo. Outros promotores conhecidos para controlar a expressão dos genes em células procarióticas ou eucarióticas, ou seus vírus, também podem ser usadas.

Vetores de expressão mamíferos também podem compreender uma origem de replicação, quaisquer sítios de ligação ao ribossomo necessários, um sítio de poliadenilação, sítios doadores e aceptores de junção, seqüências de terminação transcricional e seqüências não-transcritas de flanqueamento 5'. Em alguns aspectos, as seqüências de DNA derivadas dos sítios de junção de SV40 e de poliadenilação podem ser usados para fornecer os elementos genéticos não-transcritos requeridos.

Vetores para expressar o polipeptídeo ou fragmento deste em células eucarióticas podem conter também intensificadores para aumentar os níveis de expressão. Intensificadores são elementos de ação cis de DNA,

usualmente de cerca de 10 a cerca de 300 bp no comprimento que agem em um promotor para aumentar sua transcrição. Exemplos incluem o intensificador de SV40 no lado tardio da origem de replicação bp 100 a 270, o intensificador de promotor citomegalovírus precoce, o intensificador de polioma no lado tardio da origem de replicação, e os intensificadores de adenovírus.

Além disso, os vetores de expressão contêm um ou mais genes marcadores selecionáveis tipicamente para permitir seleção de células hospedeiras contendo o vetor. Tais marcadores selecionáveis incluem genes que codificam diidrofolato reductase ou genes que conferem resistência à neomicina por cultura de células eucarióticas, genes que conferem resistência à tetraciclina ou ampicilina em E. coli, e o gene de TRP1 de S. cerevisiae.

Após as bibliotecas de expressão serem geradas, a etapa adicional de "usar biopanning" tais bibliotecas antes de triar célula ordenada pode estar incluída. O procedimento de "usar biopanning" refere-se a um processo para identificar clones tendo uma atividade biológica especificada triando para homologia de seqüência em uma biblioteca de clones preparada por (i) seletivamente isolar o DNA alvo, de DNA derivado de pelo menos um microorganismo, através do uso de pelo menos uma sonda de DNA compreendendo pelo menos uma porção de uma seqüência de DNA codificando uma biologia tendo atividade biológica especificada; e (ii) opcionalmente transformar um hospedeiro com o DNA alvo isolado para produzir uma biblioteca de clones que são triados para a atividade biológica especificada.

O DNA da sonda usado para seletivamente isolar o DNA alvo de interesse do DNA derivado de pelo menos um microorganismo pode ser uma seqüência de região de codificação de uma seqüência de região de codificação de corpo total ou parcial de DNA para uma enzima de atividade conhecida. A biblioteca de DNA original pode ser sondada usando misturas de sondas compreendendo pelo menos uma porção da seqüência de DNA codificando uma enzima tendo a atividade de enzima especificada. Estas sondas ou bibliotecas de sonda podem ser unifilamentadas e o DNA microbiano que é sondado pode ser convertido em forma unifilamentada. As sondas que

são adequadas são aquelas derivadas de enzimas de codificação de DNA tendo uma atividade similar ou idêntica à atividade de enzima especificada que será triada.

O DNA da sonda pode ser pelo menos cerca de 10 bases ou pelo menos 15 bases. Em um aspecto, a região de codificação inteira pode ser empregada como uma sonda. Condições para a hibridação em que DNA alvo é seletivamente isolado pelo uso de pelo menos uma sonda de DNA serão designadas para fornecer uma severidade de hibridação de pelo menos cerca de 50 % de identidade de seqüência, mais particularmente uma severidade que fornece para uma identidade de seqüência de pelo menos cerca de 70 %.

Em reações de hibridação de ácido nucléico, as condições usadas para alcançar um nível particular de severidade variarão, dependendo da natureza dos ácidos nucléicos sendo hibridados. Por exemplo, o comprimento, grau de complementaridade, composição de seqüência de nucleotídeos (por exemplo, GC V. teor de A) e o tipo de ácido nucléico (por exemplo, RNA V. DNA) das regiões de hibridação dos ácidos nucléicos podem ser considerados na seleção das condições de hibridação. Uma consideração adicional é se um dos ácidos nucléicos for imobilizado, por exemplo, em um filtro.

Um exemplo de condições de severidade progressivamente mais altas é como segue: 2X SSC/0,1 % de SDS a cerca de temperatura ambiente (condições de hibridação); 0,2X SSC/0,1 % de SDS a cerca de temperatura ambiente (baixas condições de severidade); 0,2X SSC/0,1 % de SDS a cerca de 42°C (condições de severidade moderada); e 0,1X SSC a cerca de 68°C (condições de severidade alta). Lavagem pode ser realizada usando apenas uma destas condições, por exemplo, condições de severidade alta, ou cada uma das condições pode ser usada, por exemplo, durante 10-15 minutos cada, na ordem listada acima, repetindo uma ou outra ou todas das etapas listadas. Porém, como acima mencionado, condições ideais variarão, dependendo da reação de hibridação particular envolvida, e podem ser determinadas empiricamente.

Técnicas de hibridação para sondar uma biblioteca de DNA microbiano para isolar DNA alvo de interesse potencial são bem-conhecidas na técnica e quaisquer daquelas que são descritas na literatura são adequadas para uso aqui, particularmente aquelas que usam um DNA de sonda ligado de fase sólida, direta ou indiretamente ligado, para facilitar a separação do resto do DNA derivado dos microorganismos.

A DNA da sonda pode ser "marcado" com um membro de um par de ligação específico (isto é um ligando) e o outro membro do par é ligado a uma matriz sólida para fornecer facilidade de separação do alvo de sua fonte. O ligando e membro de ligação específico podem ser selecionados, em qualquer orientação, dos seguintes: (1) um antígeno ou hapteno e um anticorpo ou fragmento de ligação específico deste; (2) biotina ou iminobiotina e avidina ou estreptavidina; (3) um açúcar e uma lectina específica para este; (4) uma enzima e um inibidor para esta; (5) uma apoenzima e co-fator; (6) oligonucleotídeos homopoliméricos complementares; e (7) um hormônio e um receptor para este. A fase sólida pode ser selecionada de: (1) uma superfície vítrea ou polimérica; (2) uma coluna acumulada de contas poliméricas; e (3) partículas magnéticas ou paramagnéticas.

Também, é opcional mas desejável executar uma amplificação do DNA alvo que foi isolado. Neste aspecto o DNA alvo foi separado do DNA da sonda após isolamento. Ele é depois amplificado antes de ser usado para transformar os hospedeiros. O DNA bifilamentado selecionado para incluir como pelo menos uma porção deste uma seqüência de DNA predeterminada podendo ser unifilamentado, submetido à amplificação e re-tratado a calor para fornecer números amplificados de DNA bifilamentado selecionado. Numerosas metodologias de amplificação são bem-conhecidas agora na técnica.

O DNA selecionado é depois usado para preparar uma biblioteca para triar transformando um organismo adequado. Hospedeiros, por exemplo, aqueles especificamente aqui identificados, são transformados por introdução artificial dos vetores contendo o DNA alvo através de inoculação sob condições conducentes para tal transformação. As bibliotecas resultantes de

clones transformados são depois triadas por clones que apresentam atividade para a enzima de interesse.

Tendo preparado uma multiplicidade de clones de DNA seletivamente isolado de um organismo, tais clones são triados por uma atividade de enzima específica e para identificar os clones tendo as características de enzima especificada.

A triagem por atividade de enzima pode ser realizada em clones de expressão individuais ou pode ser realizada inicialmente em um embaralhamento de clones de expressão para averiguar se ou não o embaralhamento tem uma ou mais atividades de enzima especificada. Se a mistura tiver uma atividade de enzima especificada, então os clones individuais podem ser retriados utilizando uma máquina de FACS para tal atividade de enzima ou para uma atividade mais específica. Alternativamente, técnicas de encapsulação como microgotículas em gel podem ser empregadas para localizar clones múltiplos em uma localização a ser triada em uma máquina de FACS para clones de expressão positiva dentro do grupo de clones, que pode ser depois fragmentado em clones individuais sendo triados novamente em uma máquina de FACS para identificar clones individuais positivos. Por exemplo, desse modo se uma mistura de clone tiver atividade de hidrolase, então os clones individuais podem ser restabelecidos e triados utilizando uma máquina de FACS para determinar quais de tais clones tem atividade de hidrolase. Como aqui usado, "biblioteca de inserção pequena" significa uma biblioteca de genes contendo clones com inserções de ácido nucléico de tamanho pequeno aleatórios de até aproximadamente 5000 pares de bases. Como aqui usado, "biblioteca de inserção grande" significa uma biblioteca de genes contendo clones com inserções de ácido nucléico de tamanho grande aleatórios de cerca de 5000 até vários cem mil pares de bases ou mais.

Como descrito com respeito a um dos aspectos acima, a invenção fornece um processo para triagem de atividade de enzima de clones contendo o DNA selecionado derivado de um microorganismo cujo processo inclui: triar uma biblioteca por atividade de enzima especificada, a dita bibli-

oteca incluindo uma pluralidade de clones, ditos clones tendo sido preparados restabelecendo o DNA genômico de um DNA selecionado do microorganismo, cujo DNA é selecionado por hibridação para pelo menos uma seqüência de DNA que é toda ou uma porção de uma seqüência de DNA codificando uma enzima tendo a atividade especificada; e transformar um hospedeiro com o DNA selecionado para produzir clones que são triados pela atividade de enzima especificada.

Em um aspecto, uma biblioteca de DNA derivada de um microorganismo é submetida a um procedimento de seleção para selecionar DNA deste que hibrida com uma ou mais seqüências de DNA de sonda que são todas ou uma porção de uma seqüência de DNA codificando uma enzima tendo a atividade de enzima especificada: (a) rendendo a população de DNA genômico bifilamentada em uma população de DNA unifilamentada; (b) contatar a população de DNA unifilamentada de (a) com a sonda de DNA ligada a um ligando sob condições permissivas de hibridação para produzir um complexo bifilamentado de sonda e membros da população de DNA genômico que hibridam com este; (c) contatar o complexo bifilamentado de (b) com um membro de ligação de fase sólida específico para o dito ligando para produzir um complexo de fase sólida; (d) separar o complexo de fase sólida da população de DNA unifilamentada de (b); (e) liberar da sonda os membros da população genômica que foram ligados à sonda ligada de fase sólida; (f) formar o DNA bifilamentado dos membros da população genômica de (e); (g) introduzir o DNA bifilamentado de (f) em um hospedeiro adequado para formar uma biblioteca contendo uma pluralidade de clones que contêm o DNA selecionado; e (h) triar a biblioteca para a atividade de enzima especificada.

Em outro aspecto, o processo inclui uma pré-seleção para restabelecer o DNA incluindo seqüências sinais ou de secreção. Desta maneira é possível selecionar da população de DNA genômica por hibridação como anteriormente descrito apenas DNA incluindo uma seqüência sinal ou de secreção. Os parágrafos a seguir descrevem o protocolo para este aspecto da invenção, a natureza e função de seqüências sinais de secreção em geral

e uma aplicação exemplar específica de tais seqüências para um ensaio ou processo de seleção.

Um aspecto deste aspecto também compreende, após (a) mas antes de (b) acima, as etapas de: (ai) contatar a população de DNA unifilamentada de (a) com uma sonda de oligonucleotídeo ligada ao ligando que é complementar a uma seqüência sinal de secreção única para uma classe dada de proteínas sob condições permissivas de hibridação para formar um complexo bifilamentado; (aii) contatar o complexo bifilamentado de (ai) com um membro de ligação específico de fase sólida para dito ligando para produzir um complexo de fase sólida; (aiii) separar o complexo de fase sólida da população de DNA unifilamentada de (a); (aiv) liberar os membros da população genômica que foram ligados a dita sonda ligada de fase sólida; e (av) separar a sonda ligada de fase sólida dos membros da população genômica que foram a ela ligados.

O DNA que foi selecionado e isolado para incluir uma seqüência sinal é depois submetido ao procedimento de seleção anteriormente descrito para selecionar e isolar DNA que liga a uma ou mais seqüências de DNA de sonda derivadas de DNA codificando uma(s) enzima(s) tendo a atividade de enzima especificada. Este e outros procedimentos que podem ser usados para praticar a invenção são descritos, por exemplo, na Patente U.S. No. 6.368.798 e 6.054.267.

Usos de biopannings *in vivo* podem ser executados utilizando umas máquinas com base em FACS e não-ópticas (por exemplo, magnéticas). As bibliotecas de genes complexas são construídas com vetores contendo elementos que estabilizam o RNA transcrito. Por exemplo, a inclusão de seqüências que resultam em estruturas secundárias como grampos que são designados para flanquear as regiões transcritas do RNA serviria para intensificar sua estabilidade, desse modo aumentando sua meia vida dentro da célula. As moléculas de sonda usadas no processo de uso de biopanning consistem em oligonucleotídeos marcados com moléculas repórteres que apenas fluorescem ao ligar a sonda a uma molécula alvo. Estas sondas são introduzidas nas células recombinantes da biblioteca usando um de vários

métodos de transformação. As moléculas de sonda ligam ao mRNA alvo transcrito que resulta em moléculas heterodúplex de DNA/RNA. A ligação da sonda a um alvo renderá um sinal fluorescente que é detectado e classificado pela máquina de FACS durante o processo de triagem.

Em alguns aspectos, o ácido nucléico codificando um dos polipeptídeos da SEQ ID Nº: 2, seqüências substancialmente idênticas a esta, ou fragmentos compreendendo pelo menos cerca de 5, 10, 15, 20, 25, 30, 35, 40, 50, 75, 100, ou 150 aminoácidos sucessivos desta é montado em fase apropriada com uma seqüência líder capaz de direcionar a secreção do polipeptídeo transferido ou fragmento deste. Opcionalmente, o ácido nucléico codifica um polipeptídeo de fusão em que um dos polipeptídeos da SEQ ID Nº: 2, seqüências substancialmente idênticas a esta, ou fragmentos compreendendo pelo menos 5, 10, 15, 20, 25, 30, 35, 40, 50, 75, 100, ou 150 aminoácidos sucessivos desta, é fundido em peptídeos ou polipeptídeos heterólogos, como peptídeos N-terminais de identificação que dão características desejadas, como estabilidade aumentada ou purificação simplificada.

A seqüência de DNA apropriada pode ser inserida no vetor por uma variedade de procedimentos. Em geral, a seqüência de DNA é ligada à posição desejada no vetor seguindo digestão da inserção e o vetor com endonucleases de restrição apropriadas. Alternativamente, extremidades cegas tanto na inserção quanto no vetor podem ser ligadas. É divulgada uma variedade de técnicas de clonagem 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. Tais procedimentos e outros são julgados estar dentro do escopo daqueles versados na técnica.

Por exemplo, o vetor pode ser na forma de um plasmídeo, uma partícula virótica ou um fago. Outros vetores incluem seqüências de DNA cromossômicas, não-cromossômicas e sintéticas derivadas de SV40; plasmídeos bacterianos, DNA de fago, baculovírus, plasmídeos de levedura, vetores derivados de combinações de plasmídeo e DNA de fago, DNA virótico como vacínia, adenovírus, vírus do epitelioma contagioso ("fowl pox vi-

rus") e pseudo-raiva. Uma variedade de vetores de clonagem e de expressão para uso com hospedeiros procarióticos e eucarióticos são descritos por Sambrook, et al., Molecular Cloning: A Laboratory Manual, Segunda Edição, Cold Spring Harbor, N. I., (1989).

Vetores bacterianos particulares que podem ser usados incluem os plasmídeos comercialmente disponíveis compreendendo elementos genéticos do vetor de clonagem pBR322 bem-conhecido (ATCC 37017), pKK223-3 (Pharmacia Fine Chemicals, Uppsala, Suécia), GEM1 (Promega Biotec, Madison, WI, USA) 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. Vetores eucarióticos particulares incluem pSV2CAT, pOG44, pXTI, pSG (Stratagene) pSVK3, pBPV, pMSG, e pSVL (Pharmacia). Porém, qualquer outro vetor pode ser usado contanto que seja replicável e viável na célula hospedeira.

A célula hospedeira pode ser quaisquer das células hospedeiras familiares àqueles versados na técnica, incluindo células procarióticas, células eucarióticas, células mamíferas, células de inseto, ou células vegetais. Como exemplos representativos de hospedeiros apropriados, podem ser mencionados: células bacterianas, como E. coli, *Streptomyces*, *Bacillus subtilis*, *Salmonella typhimurium* e várias espécies dentro dos gêneros *Pseudomonas, Streptomyces* e *Staphylococcus*, células fúngicas, como levedura, células de inseto como *Drosophila* S2 e *Spodoptera* Sf9, células animais como CHO, COS ou melanoma de Bowes, e adenovírus. A seleção de um hospedeiro apropriado está dentro das habilidades daqueles versados na técnica.

O vetor pode ser introduzido nas células hospedeiras usando qualquer de uma variedade de técnicas, incluindo transformação, transfecção, transdução, infecção virótica, disparos de genes ou tradução de gene mediada por Ti. Métodos particulares incluem transfecção de fosfato de cálcio, transfecção mediada por DEAE-Dextrana, lipofecção ou eletroporação (Davis, L., Dibner, M., Battey, I., Basic Methods in Molecular Biology,

(1986)).

Onde apropriado, as células hospedeiras criadas podem ser cultivadas em meios nutrientes convencionais modificados como apropriado para ativar os promotores, selecionar transformantes ou amplificar os genes da invenção. Seguindo a transformação de uma cepa hospedeira adequada e desenvolvimento da cepa hospedeira para uma densidade de célula apropriada, o promotor selecionado pode ser induzido através de meios apropriados (por exemplo, alteração de temperatura ou indução química) e as células podem ser cultivadas durante um período adicional para lhes permitir produzir o polipeptídeo desejado ou fragmento deste.

As células são tipicamente colhidas por centrifugação, rompidas por meios físicos ou químicos e o extrato bruto resultante é retido para purificação adicional. As células microbianas empregadas para expressão das proteínas podem ser rompidas por qualquer método conveniente, incluindo ciclo de congelamento-desconglemento, sonicação, rompimento mecânico ou uso de agentes de lise celular. Tais métodos são bem-conhecidos àqueles versados na técnica. O polipeptídeo expresso ou fragmento deste pode restabelecer e ser purificado de culturas de células recombinantes através de métodos incluindo precipitação de sulfato de amônio ou etanol, extração de ácido, cromatografia de permuta de ânion ou de cátion, cromatografia de fosfocelulose, cromatografia de interação hidrofóbica, cromatografia de afinidade, cromatografia de hidroxilapatita e cromatografia de lectina. Etapas de redobragem de proteína podem ser usadas, quando necessário, no complemento da configuração do polipeptídeo. Se desejado, cromatografia líquida de alto desempenho (HPLC) pode ser empregada para etapas de purificação final.

Também podem ser empregados vários sistemas de cultura de célula mamíferas para expressar a proteína recombinante. Exemplos de sistemas de expressão mamíferos incluem as linhagens de COS-7 de fibroblastos de rim de macaco (descritas por Gluzman, Cell, 23:175, 1981), e outras linhagem de células capazes de expressar proteínas de um vetor compatível, como as linhagens de células de C127, 3T3, CHO, HeLa e BHK.

Os constructos podem ser usados em células hospedeiras de uma maneira convencional para produzir o produto de gene codificado pela seqüência recombinante. Dependendo do hospedeiro empregado em um procedimento de produção recombinante, os polipeptídeos produzidos por células hospedeiras que contêm o vetor podem ser glicosiladas ou não-glicosiladas. Polipeptídeos da invenção podem ou não incluir um resíduo de aminoácido de metionina inicial também. Detalhes adicionais relativas à expressão recombinante de proteínas estão disponíveis àqueles versados na técnica. Por exemplo, Protein Expression: A Practical Approach (Practical Approach Series por S. J. Higgins (Editor), B. D. Hames (Editor) (julho de 1999) Oxford University Press; ISBN: 0199636249) fornece ampla orientação para o aqueles versados na técnica para a expressão de proteínas em uma ampla variedade de organismos.

Alternativamente, os polipeptídeos da SEQ ID Nº: 2, seqüências substancialmente idênticas a esta, ou fragmentos compreendendo pelo menos 5, 10, 15, 20, 25, 30, 35, 40, 50, 75, 100, ou 150 aminoácidos sucessivos desta, podem ser produzidos sinteticamente através de sintetizadores de peptídeo convencionais. Em outros aspectos, fragmentos ou porções dos polipeptídeos podem ser empregados para produzir o polipeptídeo de corpo total correspondente através de síntese de peptídeo; portanto, os fragmentos podem ser empregados como intermediários para produzir os polipeptídeos de corpo total.

Como conhecido por aqueles versados na técnica, as seqüências de ácidos nucléicos da invenção podem ser otimizadas para expressão em uma variedade de organismos. Em um aspecto, seqüências da invenção são otimizadas para uso de códon em um organismo de interesse, por exemplo, um fungo como S. cerevisiae ou uma bactéria como E. coli. A otimização das seqüências de ácidos nucléicos para o propósito de uso de códon é bem entendida na técnica para referir à seleção de um códon particular favorecido por um organismo para codificar um aminoácido particular. Tabelas de uso de códons otimizados são conhecidas por muitos organismos. Por exemplo, ver Transfer RNA in Protein Synthesis por Dolph L. Hatfi-

eld, Byeong J. Lee, Robert M., Pirtle (Editor) (julho de 1992) CRC Press; ISBN: 0849356989. Desse modo, a invenção também inclui ácidos nucléicos da invenção adaptados para o uso de códons de um organismo.

Expressão otimizada das seqüências de ácidos nucléicos da invenção também refere-se à mutagênese direcionada ou aleatória de um ácido nucléico para realizar expressão aumentada da proteína codificada. A mutagênese dos ácidos nucléicos da invenção pode direta ou indiretamente fornecer um rendimento aumentado de proteína expressada. Por via de exemplo não-limitativo, técnicas de mutagênese aqui descritas podem ser utilizadas para realizar mutação das região não-transladadas 5', região não-transladada 3' ou região de codificação de um ácido nucléico, a mutação desta pode resultar em estabilidade aumentada para o nível de RNA ou de proteína, assim resultando em um rendimento aumentado de proteína.

Sistemas de tradução livre de células podem ser também empregados para produzir um dos polipeptídeos da SEQ ID Nº: 2, seqüências substancialmente idênticas a esta, ou fragmentos compreendendo pelo menos 5, 10, 15, 20, 25, 30, 35, 40, 50, 75, 100, ou 150 aminoácidos sucessivos desta, usando mRNAs transcritos de um constructo de DNA compreendendo um promotor operavelmente ligado a um ácido nucléico codificando o polipeptídeo ou fragmento deste. Em alguns aspectos, o constructo de DNA pode ser linearizado antes de conduzir uma reação de transcrição *in vitro*. O mRNA transcrito é depois incubado com um extrato de tradução livre de célula apropriado, como um extrato de reticulócito de coelho, para produzir o polipeptídeo desejado ou fragmento deste.

A invenção também diz respeito às variantes dos polipeptídeos da SEQ ID Nº: 2, seqüências substancialmente idênticas a esta, ou fragmentos compreendendo pelo menos 5, 10, 15, 20, 25, 30, 35, 40, 50, 75, 100, e 150 aminoácidos sucessivos desta. O termo "variante" inclui derivados ou análogos destes polipeptídeos. Em particular, as variantes podem diferir em seqüência de aminoácidos dos polipeptídeos da SEQ ID Nº: 2, e seqüências substancialmente idênticas a esta, por uma ou mais substituições, adições, deleção, fusões e mutilações que podem estar presente em

qualquer combinação.

As variantes podem ser de ocorrência natural ou criadas *in vitro*. Em particular, tais variantes podem ser criadas usando técnicas de engenharia genética como mutagênese locodirecionada, mutagênese química aleatória, procedimentos de deleção de Exonuclease III e técnicas de clonagem padrão. Alternativamente, tais variantes, fragmentos, análogos, ou derivados podem ser criados usando procedimentos de síntese ou modificação química.

Outros métodos de criar variantes também estão familiares àqueles versados na técnica. Estes incluem procedimentos em que as seqüências de ácidos nucléicos obtidas de isolado natural são modificadas para gerar ácidos nucléicos codificando polipeptídeos tendo características que intensificam seu valor em aplicações industriais ou de laboratório. Em tais procedimentos, um número grande de seqüências variantes, tendo uma ou mais diferenças de nucleotídeos com respeito à seqüência obtida do isolado natural, é gerado e caracterizado. Tipicamente, estas diferenças de nucleotídeos resultam na alteração do aminoácido com respeito aos polipeptídeos codificados pelos ácidos nucléicos do isolado natural.

Por exemplo, variantes podem ser criadas usando PCR propensa a erro. Em PCR propensa a erro, a PCR é executada sob condições onde a fidelidade de transcrição da polimerase de DNA é baixa, de modo que uma taxa alta de mutações de ponto é obtida ao longo do comprimento inteiro do produto de PCR. PCR propensa a erro é descrita em Leung, D. W., et al., Tecnhique, 1:11-15, 1989 e Caldwell, R. C. e Joyce G. F., Methods Applic., 2:28-33, 1992. Brevemente, em tais procedimentos, os ácidos nucléicos a serem submetidos à mutagênese, são embaralhados com iniciadores de PCR, tampão de reação, MgCl2, MnCl2, Taq polimerase e uma concentração apropriada de dNTPs para alcançar uma taxa alta de mutação de ponto ao longo do comprimento inteiro do produto de PCR. Por exemplo, a reação pode ser executada usando 20 fmoles de ácido nucléico a serem submetidos à mutagênese, 30 pmol de cada iniciador de PCR, um tampão de reação compreendendo 50 mM KCl, 10 mM Tris HCl (pH 8,3) e 0,01% de gelatina,

7mM $MgCl_2$, 0,5 mM MnCl2, 5 unidades de Taq polimerase, 0,2 mM dGTP, 0,2 nM dATP 1 mM dCTP, e 1 mM dTTP. PCR pode ser executada em 30 ciclos de 94°C durante 1 min, 45°C a 1 min e 72°C durante 1 min. Porém, será apreciado que estes parâmetros podem ser variados quando apropriado. Os ácidos nucléicos submetidos à mutagênese são clonados em um vetor apropriado e as atividades dos polipeptídeos codificados pelos ácidos nucléicos submetidos à mutagênese é avaliada.

Variantes podem também ser criadas usando mutagênese direcionada por oligonucleotídeo para gerar mutações sítio-específicas em qualquer DNA clonado de interesse. Mutagênese de oligonucleotídeo é descrita em Reidhaar-Olson, J. F. e Sauer, R. T., et al., Scicence, 241:53-57, 1988. Brevemente, em tais procedimentos uma pluralidade de oligonucleotídeos bifilamentados carregando uma ou mais mutações, a serem introduzidas no DNA clonado, é sintetizada e inserida no DNA clonado a ser submetido à mutagênese. Os clones contendo o DNA submetidos à mutagênese são restabelecidos e as atividades dos polipeptídeos que eles codificam são avaliadas.

Outro método para gerar variantes é PCR de montagem. PCR de montagem envolve a montagem de um produto de PCR de uma mistura de fragmentos de DNA pequenos. Um número grande de reações de PCR diferente ocorre em paralelo no mesmo frasco, com os produtos de um iniciador de reação dos produtos de outra reação. PCR de montagem é descrita no Pedido de Patente U.S. Pendente Serie No. 08/677.112 depositado em 9 de julho de 1996, intitulado, "Método de Embaralhamento de DNA com Polinucleotídeos produzidos por Bloqueio ou Interrupção de um Processo de Síntese ou Amplificação."

Ainda outro método de gerar variantes é mutagênese de PCR sexual. Em mutagênese de PCR sexual, recombinação homóloga forçada ocorre entre moléculas de DNA de diferentes seqüências de DNA mas altamente relacionadas *in vitro*, como resultado de fragmentação aleatória da molécula de DNA com base na homologia da seqüência, seguido por fixação da passagem através da extensão do iniciador em uma reação de PCR.

Mutagênese de PCR sexual é descrita em Stemmer, W. P., PNAS, USA, 91:10747-10751, 1994. Brevemente, em tais procedimentos uma pluralidade de ácidos nucléicos a serem recombinados com DNase é digerida para gerar fragmentos tendo um tamanho médio de 50-200 nucleotídeos. Os fragmentos do tamanho médio desejado são purificados e ressuspensos em uma mistura de PCR. PCR é conduzida sob condições que facilitam recombinação entre os fragmentos de ácido nucléico. Por exemplo, PCR pode ser executada ressuspendendo os fragmentos purificados em uma concentração de 10-30ng/μl em uma solução de 0,2 mM de cada dNTP, 2,2 mM MgCl2, 50 mM KCl, 10 mM Tris HCl, pH 9,0, e 0,1% Triton X-100. 2,5 unidades de Taq polimerase por 100 μl da mistura de reação são adicionadas e PCR é executada usando o regime a seguir: 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. Porém, será apreciado que estes parâmetros podem ser variados quando apropriado. Em alguns aspectos, podem ser incluídos oligonucleotídeos nas reações de PCR. Em outros aspectos, o fragmento de Klenow de DNA polimerase I pode ser usado em um primeiro grupo de reações de PCR e Taq polimerase pode ser usada em um grupo subseqüente de reações de PCR. As seqüências recombinantes são isoladas e as atividades dos polipeptídeos que eles codificam são avaliadas.

As variantes também podem ser criadas por mutagênese *in vivo*. Em alguns aspectos, mutações aleatórias em uma seqüência de interesse são geradas propagando a seqüência de interesse em uma cepa bacteriana, como uma cepa de E. coli que carrega mutações em uma ou mais das vias de reparo de DNA. Tais cepas "mutantes" têm uma taxa de mutação aleatória mais alta que aquela de um parente do tipo selvagem. Propagando o DNA em um destas cepas eventualmente gerará mutações aleatórias dentro do DNA. As cepas mutantes adequadas para o uso em mutagênese *in vivo* são descritas na Publicação do PCT No. WO 91/16427, publicado em 31 de outubro de 1991, intitulado "Métodos para Criação de Fenótipo de Múltiplas Populações de Gene."

Também podem ser geradas variantes usando mutagênese de

cassete. Em mutagênese de cassete uma região pequena de uma molécula de DNA bifilamentado é substituída por um "cassete" de oligonucleotídeo sintético que difere da seqüência nativa. O oligonucleotídeo freqüentemente contém seqüência nativa completamente e/ou parcialmente randomizada.

Mutagênese de conjunto recursiva pode ser também usada para gerar variantes. Mutagênese de conjunto recursiva é um algoritmo para engenharia de proteína (mutagênese de proteína) desenvolvida para produzir populações diversas de mutantes fenotipicamente relacionados cujos membros diferem na seqüência de aminoácidos. Este método usa um mecanismo de avaliação para controlar círculos sucessivos de mutagênese de cassete combinatória. Mutagênese de conjunto recursiva é descrita em Arkin, A. P. e Youvan, D. C., PNAS, USA, 89:7811-7815,1992.

Em alguns aspectos, as variantes são criadas usando mutagênese de conjunto exponencial. Mutagênese de conjunto exponencial é um processo para gerar bibliotecas combinatórias com uma porcentagem alta de mutantes únicos e funcionais, em que grupos pequenos de resíduos são randomizados em paralelo para identificar, em cada posição alterada, os aminoácidos que conduzem às proteínas funcionais. Mutagênese de conjunto exponencial é descrita em Delegrave, S., e Youvan, D. C., Biotechnol. Res., 11:1548-1552, 1993. Mutagêneses aleatória e locodirecionada são descritas em Arnold, F. H., Current Opinion in Biotechnology, 4:450-455, 1993.

Em alguns aspectos, as variantes são criadas usando procedimentos de embaralhamento em que porções de uma pluralidade de ácidos nucléicos codificando polipeptídeos distintos são fundidos para criar seqüências quiméricas de ácidos nucléicos que codificam polipeptídeos quiméricos como descrito no Pedido de Patente U.S. Pendente Serie No. 08/677.112 depositado em 9 de julho de 1996, intitulou, "Method of DNA Shuffling With Polynucleotides Produced by Blocking or interrupting a Synthesis or Simplification Process", e o Pedido de Patente U.S. Pendente Série No. 08/651.568 depositado em 22 de maio de 1996, intitulou, "Combinatorial Enzyme Development".

As variantes dos polipeptídeos da SEQ ID Nº: 2 podem ser variantes em que um ou mais dos resíduos de aminoácido dos polipeptídeos da SEQ ID Nº: 2 são substituídos por um resíduo de aminoácido conservado ou não-conservado (por exemplo, um resíduo de aminoácido conservado) e tal resíduo de aminoácido substituído pode ou não ser codificado pelo código genético.

As substituições conservadoras são aquelas que substituem um certo aminoácido em um polipeptídeo por outro aminoácido de características iguais. Tipicamente visto como substituições conservadoras são as substituições a seguir: substituições de um aminoácido alifático como o Ala, Val, Leu e Ile por outro aminoácido alifático; substituição de um Ser por um Thr ou vice-versa; substituição de um resíduo acídico como Asp e Glu por outro resíduo acídico; substituição de um resíduo carregando um grupo amida, como Asn e Gln, por outro resíduo carregando um grupo amida; permuta de um resíduo básico como Lys e Arg por outro resíduo básico; e substituição de um resíduo aromático como Phe, Tyr por outro resíduo aromático.

Outras variantes são aquelas em que um ou mais dos resíduos de aminoácido dos polipeptídeos da SEQ ID Nº: 2 incluem um grupo substituinte.

Ainda outras variantes são aquelas em que o polipeptídeo está associado com outro composto, como um composto para aumentar a meia-vida do polipeptídeo (por exemplo, polietileno glicol).

Variantes adicionais são aquelas em que aminoácidos adicionais são fundidos com o polipeptídeo, como uma seqüência líder, uma seqüência secretória, uma seqüência de proproteína ou uma seqüência que facilitam purificação, enriquecimento ou estabilização do polipeptídeo. Em alguns aspectos, os fragmentos, derivados e análogos retêm a mesma função ou atividade biológica como os polipeptídeos da SEQ ID Nº: 2, e seqüências substancialmente idênticas a esta. Em outros aspectos, o fragmento, derivado, ou análogo inclui uma proproteína, de modo que o fragmento, derivado, ou análogo pode ser ativado por clivagem da porção de proproteína para produzir um polipeptídeo ativo.

OTIMIZAÇÃO DE CÓDONS PARA ALCANÇAR NÍVEIS ALTOS DE EXPRESSÃO DE PROTEÍNA EM CÉLULAS HOSPEDEIRAS

A invenção fornece métodos para modificar ácidos nucléicos de codificação de fitase para modificar uso de códons. Em um aspecto, a invenção fornece métodos para modificar os códons em um ácido nucléico codificando uma fitase para aumentar ou diminuir sua expressão em uma célula hospedeira. A invenção também fornece ácidos nucléicos codificando uma fitase modificada para aumentar sua expressão em uma célula hospedeira, enzimas de fitase assim modificadas e métodos de produzir as enzimas de fitase modificada. O método compreende identificar um códon "não-preferido" ou um "menos preferido" no ácido nucléico de codificação de fitase e substituir um ou mais destes códon não-preferidos ou menos preferidos por um códon "preferido" codificando o mesmo aminoácido que o códon substituído e pelo menos um códon não-preferido ou menos preferido no ácido nucléico foi modificado por um códon preferido codificando o mesmo aminoácido. Um códon preferido é um códon sobre-representado nas seqüências de codificação nos genes na célula hospedeira e um códon não-preferido ou menos preferido é um códon sub-representado nas seqüências de codificação nos genes na célula hospedeira.

Células hospedeiras para expressar os ácidos nucléicos, cassetes de expressão e vetores da invenção incluem bactérias, levedura, fungos, células vegetais, células de inseto e células mamíferas. Desse modo, a invenção fornece métodos para otimizar o uso de códon em todas estas células, ácidos nucléicos alterados no códon e polipeptídeos feitos pelos ácidos nucléicos alterados no códon. Células hospedeiras exemplares incluem bactérias gram-negativas, como *Escherichia coli* e *Pseudomonas fluorescens*; bactérias gram-positivas, como *Streptomyces* diversa, *Lactobacillus gasseri*, *Lactococcus lactis*, *Lactococcus cremoris*, *Bacillus subtilis*. Células hospedeiras exemplares também incluem organismos eucarióticos, por exemplo, várias leveduras, como *Saccharomyces sp.*, incluindo Saccharomyces cerevisiae, Schizosaccharomyces pombe, *Pichia pastoris* e *Kluyveromyces lactis*, *Hansenula polymorpha*, *Aspergillus niger* e células e linha-

gens de células mamíferas e células e linhagens de células de inseto. Desse modo, a invenção também inclui ácidos nucléicos e polipeptídeos otimizados para expressão nestes organismos e espécies.

Por exemplo, os códons de um ácido nucléico são modificados codificando uma fitase isolada de uma célula bacteriana de modo que o ácido nucléico é expresso otimamente em uma célula bacteriana diferente das bactérias das quais a fitase foi derivado, uma levedura, um fungo, uma célula vegetal, uma célula de inseto ou uma célula mamífera. Métodos para otimizar códons são bem-conhecidos na técnica, ver, por exemplo, Patente U.S. No. 5.795.737; Baca (2000) Int. J. Parasitol. 30:113-118; Hale (1998) Protein Expr. Purif. 12:185-188; Narum (2001) Infect. Immun. 69:7250-7253. Também ver Narum (2001) Infect. Immun. 69:7250-7253, descrevendo otimização de códons em sistemas de camundongo; Outchkourov (2002) Protein Expr. Purif. 24:18-24, descrevendo otimizar códons em levedura; Feng (2000) Biochemistry 39:15399-15409, descrevendo otimizar códons em E. coli; Humphreys (2000) Protein Expr. Purif. 20:252-264, descrevendo otimizar o uso de códons que afeta a secreção em E. coli.

ANIMAIS NÃO-HUMANOS TRANSGÊNICOS

A invenção fornece animais não-humanos transgênicos compreendendo um ácido nucléico, um polipeptídeo, um cassete ou vetor de expressão ou uma célula transfectada ou transformada da invenção. Os animais não-humanos transgênicos podem ser, por exemplo, cabras, coelhos, ovelhas, porcos, vacas, ratos e camundongos, compreendendo os ácidos nucléicos da invenção. Estes animais podem ser usados, por exemplo, como modelos *in vivo* para estudar atividade de fitase, ou, como modelos triar para moduladores de atividade de fitase *in vivo*. As seqüências de codificação para os polipeptídeos a serem expressos nos animais não-humanos transgênicos podem ser designadas sendo constitutivas, ou, sob o controle de fatores reguladores transcripcionais tecido-específicos, desenvolvimento-específicos ou induzíveis. Animais não-humanos transgênicos podem ser designados e gerados usando qualquer método conhecido na técnica; ver, por exemplo, Patente U.S. Nos. 6.211.428; 6.187.992; 6.156.952; 6.118.044;

6.111.166; 6.107.541; 5.959.171; 5.922.854; 5.892.070; 5.880.327; 5.891.698; 5.639.940; 5.573.933; 5.387.742; 5.087.571, que descrevem a criação e o uso de células e ovos transformados e camundongos, ratos, coelhos, ovelhas, porcos e vacas transgênicos. Também ver, por exemplo, Pollock (1999) J. Immunol. Methods 231:147 157, que descreve a produção de proteínas recombinantes no leite de animais leiteiros transgênicos; Baguisi (1999) Nat. Biotechnol. 17:456-461, demonstrando a produção de cabras transgênicas. Patente U.S. No. 6.211.428, descreve fabricação e uso de mamíferos não-humanos transgênicos que expressam em seus cérebros um constructo de ácido nucléico compreendendo uma seqüência de DNA. Patente U.S. No. 5.387.742, descreve injetar seqüências de DNA recombinantes ou sintéticas clonadas em ovos de camundongo fertilizados, implantando os ovos injetados em fêmeas pseudoprenhas, e cultivando para chegar a camundongos transgênicos cujas células expressam proteínas relacionadas à patologia da doença de Alzheimer. Patente U.S. No. 6.187.992, descreve fabricação e uso de um camundongo transgênico cujo genoma compreende um rompimento do gene que codifica proteína precursora amilóide (APP).

"Animais golpeados" também podem ser usados para praticar os métodos da invenção. Por exemplo, em um aspecto, os animais transgênicos ou modificados da invenção compreendem um "animal golpeado," por exemplo, um "camundongo golpeado," desenvolvido para não expressar ou não poder expressar uma fitase.

METODOLOGIAS DE TRIAGEM E DISPOSITIVOS DE MONITORAMENTO "ON-LINE"

Na prática dos métodos da invenção, uma variedade de aparelho e metodologias pode ser usada para junto com os polipeptídeos e ácidos nucléicos da invenção, por exemplo, triar polipeptídeos para atividade de fitase, triar compostos como moduladores potenciais de atividade (por exemplo, potenciação ou inibição de atividade de enzima), para anticorpos que ligam a um polipeptídeo da invenção, para ácidos nucléicos que hibridizam com um ácido nucléico da invenção, e outros.

SUPORTES SÓLIDOS DE ENZIMA IMOBILIZADA

As enzimas de fitase, fragmentos destas e ácidos nucléicos codificando as enzimas e fragmentos podem ser presos a um suporte sólido. Isto é freqüentemente econômico e eficiente no uso das fitases em processos industriais. Por exemplo, uma associação ou coquetel de enzimas de fitase (ou fragmentos ativos destas) sendo usado em uma reação química específica pode ser ligada a um suporte sólido e pode ser a embebido em um barril do processo. A reação enzimática pode ocorrer. Depois, o suporte sólido pode ser tirado do barril, junto com as enzimas presas a este, para uso repetido. Em uma modalidade da invenção, um ácido nucléico isolado da invenção é anexado a um suporte sólido. Em outra modalidade da invenção, o suporte sólido é selecionado do grupo de um gel, uma resina, um polímero, uma cerâmica, um vidro, um microeletrodo e qualquer combinação destes.

Por exemplo, suportes sólidos úteis nesta invenção incluem géis. Alguns exemplos de géis incluem Sepharose, gelatina, glutaraldeído, glutaraldeído tratado com quitosana, glutaraldeído de albumina, quitosana-Xantana, gel de tolu ("toyopearl gel") (gel de polímero), alginato, alginato-polilisina, carragenina, agarose, agarose de glioxila, agarose magnética, dextrana-agarose, poli(Sulfonato de Carbamoíla) hidrogel, hidrogel de BSA-PEG, álcool polivinílico fosforilado (PVA), monoaminoetil-N-aminoetila (MANA), amino, ou qualquer combinação destes.

Outro suporte sólido útil na presente invenção são resinas ou polímeros. Alguns exemplos de resinas ou polímeros incluem celulose, acrilamida, náilon, raiom, poliéster, resina de permuta de ânion, AMBERLITE® XAD-7, AMBERLITE® XAD-8, AMBERLITE® IRA 94, AMBERLITE® IRC-50, polivinila, poliacrílico, polimetacrilato, ou qualquer combinação destes. Outro tipo de suporte sólido útil na presente invenção é cerâmica. Alguns exemplos incluem cerâmica não-porosa, cerâmica porosa, SiO2, Al_2O_3. Outro tipo de suporte sólido útil na presente invenção é de vidro. Alguns exemplos incluem vidro não-poroso, vidro poroso, vidro de aminopropila ou qualquer combinação destes. Outro tipo de suporte sólido que pode ser usado é um microele-

trodo. Um exemplo é uma magnetita revestida de polietilenoimina.

As partículas grafítica podem ser usadas como um suporte sólido. Outro exemplo de um suporte sólido é uma célula, como uma hemácia.

MÉTODOS DE IMOBILIZAÇÃO

Há muitos métodos que podem ser conhecidos a alguém de versado na técnica para imobilizar enzimas ou fragmentos destas, ou ácidos nucléicos, sobre um suporte sólido. Alguns exemplos de tais métodos incluem, por exemplo, geração de gotícula eletrostática, meios eletroquímicos, por meio de adsorção, por meio de ligação covalente, por meio de reticulação, por meio de uma reação ou processo químico, por meio de encapsulação, por meio de capturação, por meio de alginato de cálcio ou por meio de poli(metacrilato de 2-hidroxietila). Métodos semelhantes são descritos em Methods in Enzymology, Immobilized Enzymes and Cells, Parte 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. Séries: Methods in Biotechnology, Editado por J. M. Walker.

ARRANJOS CAPILARES

Arranjos capilares, como o GIGAMATRIX®, Diversa Corporation, San Diego, CA, podem ser usados nos métodos da invenção. Ácidos nucléicos ou polipeptídeos da invenção podem ser imobilizados ou aplicados a um arranjo, incluindo arranjos capilares. Os arranjos podem ser usados para triar ou monitorar bibliotecas de composições (por exemplo, as moléculas pequenas, anticorpos, ácidos nucléicos, etc.) por sua capacidade de ligar ou modular a atividade de um ácido nucléico ou um polipeptídeo da invenção. Arranjos capilares fornecem outro de sistema para segurar e triar amostras. Por exemplo, um aparelho de triagem de amostra pode incluir uma pluralidade de capilares formados em um arranjo de capilares adjacentes, em que cada capilar compreende pelo menos uma parede que define um lúmen para reter uma amostra. O aparelho pode também incluir material intersticial disposto entre capilares adjacentes no arranjo, e um ou mais índices de referência formados dentro do material intersticial. Um capilar para triar uma amostra, em que o capilar é adaptado por estar ligado em um arranjo de ca-

pilares, pode incluir uma primeira parede definindo um lúmen para reter a amostra, e uma segunda parede formada de um material de filtração, para filtrar energia de excitação fornecida ao lúmen para excitar a amostra.

Um polipeptídeo ou ácido nucléico, por exemplo, um ligando, pode ser introduzido em um primeiro componente em pelo menos uma porção de um capilar de um arranjo capilar. Cada capilar do arranjo 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 pela bolha de ar. Uma amostra de interesse pode ser introduzida como um primeiro líquido marcado com uma partícula detectável em um capilar de um arranjo capilar, em que cada capilar do arranjo 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 com um material de ligação para ligar a partícula detectável a pelo menos uma parede. O método pode também inclui 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.

O arranjo 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 entre si. Similarmente, a parede pode definir um lúmen que é cilíndrico, quadrado, hexagonal ou qualquer outra forma geométrica tão longa quanto as paredes que formam um lúmen para retenção de um líquido ou amostra. Os capilares do arranjo capilar podem ser unidos em proximidade íntima para formar uma estrutura plana. Os capilares podem ser ligados entre si, sendo fundidos (por exemplo, onde os capilares são feitos de vidro), colados, ligados ou presos por grampo de lado-a-lado. O arranjo capilar pode ser formado de qualquer número de capilares individuais, por exemplo, uma faixa de 100 a 4.000.000 capilares. Um arranjo capilar pode formar uma placa de microtitulação tendo cerca de 100.000 ou mais capilares individuais ligados

entre si.

ARRANJOS, OU "BIOCHIPS"

Os ácidos nucléicos ou polipeptídeos da invenção podem ser imobilizados ou aplicados a um arranjo. Os arranjos podem ser usados para triar ou monitorar bibliotecas de composições (por exemplo, as moléculas pequenas, anticorpos, ácidos nucléicos, etc.) por sua capacidade de ligar ou modular a atividade de um ácido nucléico ou um polipeptídeo da invenção. Por exemplo, em um aspecto da invenção, um parâmetro monitorado é expressão de transcrição de um gene de fitase. Uma ou mais, ou, todas as transcrição de uma célula podem ser medidas por hibridação de uma amostra compreendendo transcrição da célula, ou, ácidos nucléicos representativos ou complementares às transcrição de uma célula, por hibridação com ácidos nucléicos imobilizados em um arranjo, ou "biofragmento." Usando um "arranjo" de ácidos nucléicos em um microfragmento, algumas ou todas as transcrição de uma célula podem ser quantificadas simultaneamente. Alternativamente, também podem ser usados arranjos compreendendo ácido nucléico genômico para determinar o genótipo de uma cepa recentemente criada feita pelos métodos da invenção. "Arranjos de polipeptídeos" também podem ser usados para simultaneamente quantificar uma pluralidade de proteínas.

A presente invenção pode ser praticada com qualquer "arranjo" conhecido, também referido como um "microarranho" ou ""arranjo de ácido nucléico" ou "arranjo de polipeptídeos" ou "arranjo de anticorpo" ou "biofragmento," ou variação destes. Arranjos são genericamente uma pluralidade de de "pontos" ou "elementos alvo," cada elemento alvo compreendendo uma quantidade definida de uma ou mais moléculas biológicas, por exemplo, oligonucleotídeos, imobilizados sobre uma área definida de uma superfície de substrato para ligação específica a uma molécula de amostra, por exemplo, transcrição de mRNA.

Na prática dos métodos da invenção, qualquer arranjo e/ou método conhecido(s) de produzir e usar arranjos podem ser incorporados ao todo ou em parte, ou variações destes, como descrito, por exemplo, na Pa-

tente U.S. Nos. 6.277.628; 6.277.489; 6.261.776; 6.258.606; 6.054.270; 6.048.695; 6.045.996; 6.022.963; 6.013.440; 5.965.452; 5.959.098; 5.856.174; 5.830.645; 5.770.456; 5.632.957; 5.556.752; 5.143.854; 5.807.522; 5.800.992; 5.744.305; 5.700.637; 5.556.752; 5.434.049; também ver, por exemplo, WO 99/51773; WO 99/09217; WO 97/46313; WO 96/17958; também ver, por exemplo, Johnston (1998) Curr. Biol. 8:R171-R174; Schummer (1997) Biotechniques 23:1087-1092; Kern (1997) Biotechniques 23:120-124; Solinas-Toldo (1997) Genes, Chromosomes & Cancer 20:399-407; Bowtell (1999) Nature Genetics Supp. 21:2532. Ver t os pedidos de patente U.S. Publicados Nos. 20010018642; 20010019827; 20010016322; 20010014449; 20010014448; 20010012537; 20010008765.

POLIPEPTÍDEOS E PEPTÍDEOS

A invenção fornece polipeptídeos isolados ou recombinantes tendo uma identidade de seqüência a uma seqüência exemplar da invenção, por exemplo, SEQ ID Nº: 2. Como debatido acima, a identidade pode ser no comprimento total do polipeptídeo, ou, a identidade pode ser em uma região de pelo menos cerca de 50, 77, 100, 150, 200, 250, 300 ou mais resíduos (para o comprimento total do polipeptídeo). Polipeptídeos da invenção também podem ser mais curtos que o comprimento total dos polipeptídeos exemplares (por exemplo, SEQ ID Nº: 2). Em modalidade alternativa, a invenção fornece polipeptídeos (peptídeos, fragmentos) variando no tamanho entre cerca de 5 e o comprimento total de um polipeptídeo, por exemplo, uma fitase; tamanhos exemplares sendo de cerca de 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 100, 125, 150, 175, 200, 250, 300, 350, 400 ou mais resíduos, por exemplo, resíduos contíguos das fitases exemplares da SEQ ID Nº: 2. Peptídeos da invenção podem ser úteis como, por exemplo, sondas de marcação, antígenos, tolerágenos, motivos, sítios ativos de fitase.

Os polipeptídeos e peptídeos da invenção podem ser isolados de fontes naturais, podem ser sintéticos ou podem ser polipeptídeos recombinantemente gerados. Os peptídeos e proteínas podem ser recombinantemente expressos *in vitro* ou *in vivo*. Os peptídeos e polipeptídeos da inven-

ção podem ser feitos e isolados usando qualquer método conhecido na técnica. Os polipeptídeos e peptídeos da invenção também podem ser sintetizados, ao todo ou em parte, usando métodos químicos bem-conhecidos na técnica. Ver por exemplo, Caruthers (1980) Nucleic Acids Res. Symp. Ser. 215-223; Horn (1980) Nucleic Acids Res. Symp. Ser. 225-232; Banga, A. K., Therapeutic Peptides Tand Proteins, Formulation, Processing and Delivery Systems (1995) Technomic Publishing Co., Lancaster, PA. Por exemplo, a síntese de peptídeo pode ser executada usando várias técnicas de fase sólida (ver por exemplo, Roberge (1995) Science 269:202; Merrifield (1997) Methods Enzymol. 289:3 13) e síntese automatizada pode ser alcançada, por exemplo, usando o Sintetizador de Peptídeo ABI 431A (Perkin Elmer) de acordo com as instruções fornecidas pelo fabricante.

Os peptídeos e polipeptídeos da invenção também podem ser glicosilados. A glicosilação pode ser adicionada pós-translacional ou quimicamente ou através de mecanismos biossintéticos celulares, em que o último incorpora o uso de motivos de glicosilação conhecidos, que podem ser nativos para seqüência ou podem ser adicionados como um peptídeo ou ser adicionados na seqüência de codificação de ácido nucléico. A glicosilação pode ser O-ligada ou N-ligada, ou, uma combinação destas.

Os peptídeos e polipeptídeos da invenção, como acima definidos, incluem todas as formas "miméticas" e "peptidomiméticas." Os termos "mimético" e "peptidomimético" referem-se a um composto químico sintético tendo substancialmente as mesmas características estruturais e/ou funcionais dos polipeptídeos da invenção. O mimético ou pode ser composto completamente de análogos sintéticos, não-naturais de aminoácidos, ou é uma molécula quimérica dos aminoácidos de peptídeo em parte naturais e análogos em parte não-naturais de aminoácidos. O mimético também pode incorporar qualquer quantidade de substituições conservadoras de aminoácido natural contanto que tais substituições também não alterem substancialmente a estrutura e/ou atividade mimética. Quanto aos polipeptídeos da invenção que são variantes conservadoras, rotina de experimentação determinará se um mimético está dentro do escopo da invenção, isto é, se sua es-

trutura e/ou função não é alterada substancialmente. Desse modo, em um aspecto, uma composição mimética está dentro do escopo da invenção se tiver uma atividade de fitase.

Composições miméticas de polipeptídeos da invenção podem conter qualquer combinação de componentes estruturais não-naturais. Em aspecto alternativo, as composições miméticas da invenção incluem um ou todos dos seguintes três grupos estruturais: a) grupos de ligação de resíduo diferente de ligações ligadas a amida natural ("ligação de peptídeo); b) resíduos não-naturais no lugar de resíduos de aminoácido de ocorrência natural; ou c) resíduos que induzem imitação estrutural secundária, isto é, para induzir ou estabilizar uma estrutura secundária, por exemplo, uma volta beta, volta faixa, folha beta, conformação de hélice alfa e outras. Por exemplo, um polipeptídeo da invenção pode ser caracterizado como um mimético quando todos ou alguns de seus resíduos são unidos por diferentes meios químicos de ligações de peptídeo natural. Resíduos peptidomiméticos individuais podem ser unidos por ligações de peptídeo, outras ligações químicas ou meios de acoplamento, como, por exemplo, glutaraldeído, ésteres de N-hidroxissucinimida, maleimidas bifuncionais, N,N'-dicicloexil-carbodiimida (DCC) ou N,N'-diisopropilcarbodiimida (DIC). Grupos de ligação que podem ser uma alternativa para ligações ligadas à amida tradicionais ("ligação peptídica") incluem, por exemplo, cetometileno (por exemplo, - C(=O)CH2- para -C(=O)-NH-), aminometileno (CH2-NH), etileno, olefina (CH=CH), éter (CH2-O), tioéter (CH2-S), tetrazol (CN4-), tiazol, retroamida, tiomida ou éster (ver, por exemplo, Spatola (1983) em Chemistry and Biochemistry of Amino Acids, Peptides and Proteins, Vol. 7, pp 267-357, "Pepide Backbone Modifications," Marcell Dekker, N. Y.).

Um polipeptídeo da invenção também pode ser caracterizado como um mimético contendo todos ou alguns resíduos não-naturais no lugar de resíduos de aminoácido de ocorrência natural. Resíduos não-naturais são bem descritos na literatura científica e de patente; algumas composições não-naturais exemplares úteis como miméticos de resíduos de aminoácido naturais e normas são descritos abaixo. Miméticos de aminoácidos aromáti-

cos podem ser gerados substituindo, por exemplo, por D ou L-nafilalanina; D ou L-fenilglicina; D ou L-2-tienoilalanina; D ou L-1, 2, 3, ou 4-pirenoilalanina; D ou L-3-tienoilalanina; 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-p-flúor-fenilalanina; D ou L-p-bifenilfenilalanina; K ou L-p-metóxi-bifenilfenilalanina; D ou L-2-indol(alquil)alaninas; e D ou L-alquilalaninas onde alquila pode ser substituída ou insubstituída por metila, etila, propila, hexila, butila, pentila, isopropila, iso-butila, sec-isotila, iso-pentila ou uns aminoácidos não-acídicos. Anéis aromáticos de um aminoácido não-natural incluem, por exemplo, anéis aromáticos de tiazolila, tiofenila, pirazolila, benzimidazolila, naftila, furanila, pirrolila e piridila.

Miméticos de aminoácidos acídicos podem ser gerados através de substituição, por exemplo, por aminoácidos de não-carboxilato ao mesmo tempo mantendo uma carga negativa; (fosfono)alanina; treonina sulfatada. Grupos carboxila laterais (por exemplo, aspartila ou glutamila) também podem ser seletivamente modificados através de reação com carbodiimidas (R'-N-C-N-R') como, por exemplo, 1-cicloexil-3(2-morfolinil-(4-etil)carbodiimida ou 1-etil-3(4-azonia-4,4-dimetolpentil)carbodiimida. Aspartila ou glutamila também podem ser convertidos em resíduos de asparaginila e glutaminila através de reação com íons de amônio. Miméticos de aminoácidos básicos podem ser gerados através de substituição por, por exemplo, (além de lisina e arginina) os aminoácidos ornitina, citrulina ou ácido (guanidino)-acético, ou ácido guanidino)alquil-acético onde alquila é acima definida. Derivado de nitrila (por exemplo, contendo a metade de CN no lugar de COOH) pode ser substituído por asparagina ou glutamina.

Resíduos de asparaginila e glutaminila podem ser desaminados para os resíduos de aspartila ou glutamila correspondentes. Miméticos de resíduo de arginina podem ser gerados reagindo arginila com, por exemplo, um ou mais reagentes convencionais, incluindo, por exemplo, fenilglioxal, 2,3-butanodiona, 1,2-ciclo-hexanodiona, ou ninhidrin, preferivelmente sob condições alcalinas. Miméticos de resíduos de tirosina podem ser gerados

reagindo tirosila com, por exemplo, compostos de diazônio aromáticos ou tetranitrometano. N-acetilimidizol e tetranitrometano podem ser usados para formar espécies de tirosila de O-acetila e derivados de 3-nitro, respectivamente. Os miméticos de resíduos de cisteína podem ser gerados reagindo resíduos de cisteinila com, por exemplo, alfa-haloacetatos como ácido 2-cloroacético ou cloroacetamida e aminas correspondentes; para dar derivados de carboximetila ou carboxiamidometila. Miméticos de resíduos de cisteína também podem ser gerados reagindo resíduos de cisteinila com, por exemplo, bromo-trifluoroacetona, ácido alfa-bromo-beta-(5-imidozoíla)propiônico; fosfato de cloroacetila, N-alquilmaleimidas, dissulfeto de 3-nitro-2-piridila; dissulfeto de 2-piridil metila; p-cloromercuriobenzoato; 2-cloromercúrio-4-nitrofenol; ou, cloro-7-nitrobenzo-oxa-1,3-diazol. Miméticos de lisina podem ser gerados (e resíduos amino-terminais podem ser alterados) reagindo lisinila com, por exemplo, anidridos de ácido sucínico ou outros de ácido carboxílico. Lisina e outros miméticos de resíduos contendo alfa-amino também podem ser gerados através de reação com imidoésteres, como picolinimidato de metila, fosfato de piridoxal, piridoxal, cloroboroidreto, ácido de trinitro-benzenossulfônico, O-metilisouréia, 2,4-pentanodiona e reações catalisadas por transamidase com glioxilato. Os miméticos de metionina podem ser gerados através de reação com, por exemplo, sulfóxido de metionina. Miméticos de prolina incluem, por exemplo, ácido pipecólico, ácido carboxílico de tiazolidina, 3 ou 4-prolina de hidróxi, deidroprolina, 3 ou 4-metilprolina, ou 3,3,-dimetilprolina. Miméticos de resíduos de histidina podem ser gerados reagindo lustidila com, por exemplo, dietilprocarbonato ou brometo de para-bromofenacila. Outros miméticos incluem, por exemplo, aqueles gerados por hidroxilação de prolina e lisina; fosforilação dos grupos hidroxila de resíduos serila ou treonila; metilação dos grupos alfa-amino de lisina, arginina e histidina; acetilação da amina N-terminal; metilação de resíduos de amida da estrutura principal ou substituição por aminoácidos de N-metila; ou amidaçao de grupos carboxila C-terminais.

Um resíduo, por exemplo, um aminoácido, de um polipeptídeo da invenção também pode ser substituído por um aminoácido (ou resíduo

peptidomimético) da quiralidade oposta. Desse modo, qualquer aminoácido de ocorrência natural na configuração L (que também pode ser referido como o R ou S, dependendo da estrutura da entidade química) pode ser substituído pelo aminoácido do mesmo tipo estrutural químico ou um peptidomimético, mas da quiralidade oposta, referida como o aminoácido D, mas também pode ser referido como a forma R ou S.

A invenção também fornece métodos para modificar os polipeptídeos da invenção por quaisquer processos naturais, como processamento pós-translacional (por exemplo, fosforilação, acilação, etc), ou através de técnicas de modificação químicas, e os polipeptídeos modificados resultantes. Modificações podem ocorrer em qualquer lugar no polipeptídeo, incluindo a estrutura principal de peptídeo, as cadeias laterais de aminoácido e os términos amino ou carboxila. Será apreciado que o mesmo tipo de modificação pode estar presente nos mesmos graus ou variados em vários sítios em um certo polipeptídeo. Também um certo polipeptídeo pode ter muitos tipos de modificações. Modificações incluem acetilação, acilação, ribosilação de PAD, amidação, ligação covalente de flavina, ligação covalente de uma semi-metade, ligação covalente de um nucleotídeo ou derivado de nucleotídeo, ligação covalente de um lipídio ou derivado de lipídio, ligação covalente de um fosfatidilinositol, ciclização de reticulação, formação de ligação de dissulfeto, desmetilação, formação de reticulações covalentes, formação de cisteína, formação de piroglutamato, formilação, faixa-carboxilação, glicosilação, formação de âncora de GPI, hidroxilação, iodação, metilação, miristoliação, oxidação, tratamento com PEG, processamento proteolítico, fosforilação, prenilação, racemização, selenoilação, sulfatização e tradução de RNA mediada por adição de aminoácidos a proteína como arginilação. Ver, por exemplo, 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).

Métodos de síntese química de peptídeo de fase sólida podem ser usados para sintetizar o polipeptídeo ou fragmentos da invenção. Tal

método foi conhecido na técnica desde os anos 60 (Merrifield, R., B., J. Am. Chem. Soc., 85:2149-2154, 1963) (também ver Stewart, J., M. e Young J. D., Solid Phase Peptide Synthesis, 2ª Ed., Pierce Chemical Cia., Rockford, Ill., pp. 11-12) e foi recentemente empregado em kits de modelo e síntese de peptídeo de laboratório comercialmente disponíveis (Cambridge Research Biochemicals). Tais kits de laboratório comercialmente disponíveis em geral utilizaram os ensinamentos de H. M. Geysen et al, Proc. Natl. Acad. Sci., USA, 81:3998 (1984) e fornecem sintetização de peptídeos nas pontas de uma multidão de "bastões" ou "pinos" todos destes estão conectados a uma única placa. Quando um tal sistema é utilizado, uma placa de bastões ou pinos é invertida e inserida em uma segunda placa de cavidades ou reservatórios correspondentes, que contêm soluções para ligar ou ancorar um aminoácido apropriado às pontas do pino ou bastão. Repetindo uma tal etapa de processo, isto é, invertendo e inserindo as pontas do bastão e do pino nas soluções apropriadas, aminoácidos são construídos em peptídeos desejados. Além disso, vários sistemas de síntese de peptídeo de FMOC disponíveis estão disponíveis. Por exemplo, uma montagem de um polipeptídeo ou fragmento pode ser realizada em um suporte sólido usando um sintetizador de peptídeo automatizado da Applied Biosystems, Inc., Modelo 431A®. Tal equipamento fornece acesso fácil aos peptídeos da invenção, ou através da síntese direta ou pela síntese de uma série de fragmentos que podem ser acoplados usando outras técnicas conhecidas.

Outro aspecto da invenção é polipeptídeos ou fragmentos destes tendo pelo menos cerca de 70%, pelo menos cerca de 80%, pelo menos cerca de 85%, pelo menos cerca de 90%, pelo menos cerca de 95%, ou mais que cerca de 95% de homologia a um dos polipeptídeos da SEQ ID Nº: 2, seqüências substancialmente idênticas a esta, ou um fragmento compreendendo pelo menos 5, 10, 15, 20, 25, 30, 35, 40, 50, 75, 100, ou 150 aminoácidos sucessivos desta. Homologia pode ser determinada usando quaisquer dos programas acima descritos que alinha os polipeptídeos ou os fragmentos sendo comparados e determina a extensão de identidade do aminoácido ou similaridade entre eles. Será apreciado que "homologia" de amino-

ácido inclui substituições de aminoácido conservadoras como aquelas descritas acima.

Os polipeptídeos ou fragmentos tendo homologia a um dos polipeptídeos da SEQ ID Nº: 2, seqüências substancialmente idênticas a esta, ou um fragmento compreendendo pelo menos cerca de 5, 10, 15, 20, 25, 30, 35, 40, 50, 75, 100, ou 150 aminoácidos sucessivos desta, podem ser obtidos isolando os ácidos nucléicos que os codificam usando as técnicas acima descritas.

Alternativamente, os polipeptídeos homólogos ou fragmentos podem ser obtidos através de procedimentos de enriquecimento ou purificação bioquímicos. A seqüência de polipeptídeos potencialmente homólogos ou fragmentos pode ser determinada por digestão proteolítica, eletroforese em gel e/ou microsseqüenciação. A seqüência do polipeptídeo homólogo prospectivo ou fragmento pode ser comparada a um dos polipeptídeos da SEQ ID Nº: 2, seqüências substancialmente idênticas a esta, ou um fragmento compreendendo pelo menos cerca de 5, 10,15, 20, 25, 30, 35, 40, 50, 75, 100, ou 150 aminoácidos sucessivos desta usando quaisquer dos programas aqui descritos.

Outro aspecto da invenção é um ensaio para identificar fragmentos ou variantes da SEQ ID Nº: 2, ou seqüências substancialmente idênticas a esta, que retém a função enzimática dos polipeptídeos da SEQ ID Nº: 2 e seqüências substancialmente idênticas a esta. Por exemplo os fragmentos ou variantes dos polipeptídeos, podem ser usados para catalisar reações bioquímicas que indicam que o dito fragmento ou variante retém a atividade enzimática dos polipeptídeos na SEQ ID Nº: 2.

O ensaio para determinar se os fragmentos das variantes retêm a atividade enzimática dos polipeptídeos da SEQ ID Nº: 2, e seqüências substancialmente idênticas a esta, inclui as etapas de: contatar o fragmento de polipeptídeos ou variante com uma molécula de substrato sob condições que permitem o fragmento de polipeptídeos ou variante funcionar, e ou detectar uma diminuição no nível de substrato ou um aumento no nível do produto de reação específico da reação entre o polipeptídeo e substrato.

Os polipeptídeos da SEQ ID Nº: 2, seqüências substancialmente idênticas a esta, ou fragmentos compreendendo pelo menos 5,10, 15, 20, 25, 30, 35, 40, 50, 75, 100, ou 150 aminoácidos sucessivos desta, podem ser usados em uma variedade de aplicações. Por exemplo, os polipeptídeos ou fragmentos destes podem ser usados para catalisar reações bioquímicas. De acordo com um aspecto da invenção, é provido um processo para utilizar um polipeptídeo tendo SEQ ID Nº: 2, e seqüências substancialmente idênticas a esta, ou polinucleotídeos codificando tais polipeptídeos para hidrolisar haloalcanos. Em tais procedimentos, uma substância que contém um composto de haloalcano é contatada com um dos polipeptídeos da SEQ ID Nº: 2, seqüências substancialmente idênticas a esta, sob condições que facilitam a hidrólise do composto.

ANTICORPOS E MÉTODOS DE TRIAGEM COM BASE EM ANTICORPOS

A invenção fornece anticorpos isolados ou recombinantes que especificamente ligam a uma fitase da invenção. Estes anticorpos podem ser usados para isolar, identificar ou quantificar as fitases da invenção ou polipeptídeos relacionados. Estes anticorpos podem ser usados para inibir a atividade de uma enzima da invenção. Estes anticorpos podem ser usados para isolar polipeptídeos relacionados àqueles da invenção, por exemplo, enzimas de fitase relacionadas.

Os anticorpos podem ser usados em imunoprecipitação, manchamento (por exemplo, FACS), colunas de imunoafinidade e outros. Se desejado, as seqüências de ácidos nucléicos codificando para antígenos específicos podem ser geradas por imunização seguida por isolamento do polipeptídeo ou ácido nucléico, amplificação ou clonagem e imobilização do polipeptídeo em um arranjo da invenção. Alternativamente, os métodos da invenção podem ser usados para modificar a estrutura de um anticorpo produzido por uma célula a ser modificada, por exemplo, a afinidade de um anticorpo pode ser aumentada ou diminuída. Além disso, a capacidade de fazer ou modificar anticorpos pode ser um fenótipo criado em uma célula pelos métodos da invenção.

Os métodos de imunização, produção e isolamento de anticor-

pos (policlonais e monoclonais) são conhecidos a aqueles de versado na técnica e descritos na literatura científica e de patente, ver, por exemplo, Coligan, CURRENT PROTOCOLS IN IMMUNOLOGY, Wiley/Greene, NI (1991); Stites (eds.) BASIC AND CLINICAL IMMUNOLOGY (7ª ed.) Lange Medical Publications, Los Altos, CA ("Stites"); Goding, MONOCLONAL ANTIBODIES: PRINCIPLES AND PRACTICE (2ª ed.) Academic Press, Nova Iorque, NI (1986); Kohler (1975) Nature 256:495; Harlow (1988) ANTIBODIES, A LABORATORY MANUAL, Cold Springer Harbor Publications, Nova Iorque. Os anticorpos também podem ser gerados *in vitro*, por exemplo, usando sítio de ligação de anticorpo recombinante que expressa bibliotecas de exibição de fago, além dos métodos *in vivo* tradicionais usando animais. Ver, por exemplo, Hoogenboom (1997) Trends Biotechnol. 15:62-70; Katz (1997) Annu. Rev. Biophys. Biomol. Struct. 26:27-45.

Os polipeptídeos podem ser usados para gerar anticorpos que especificamente ligam aos polipeptídeos da invenção. Os anticorpos resultantes podem ser usados em procedimentos de cromatografia de imunoafinidade para isolar ou purificar o polipeptídeo ou determinar se o polipeptídeo está presente em uma amostra biológica. Em tais procedimentos, uma preparação de proteína, como um extrato, ou uma amostra biológica é contatada com um anticorpo capaz de especificamente ligar a um dos polipeptídeos da invenção.

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

A capacidade de proteínas em uma amostra biológica de ligar ao anticorpo pode ser determinada usando qualquer de uma variedade de procedimentos familiares àqueles versados na técnica. Por exemplo, ligação pode ser determinada marcando o anticorpo com um rótulo detectável como agente fluorescente, um rótulo enzimática ou um radioisótopo. Alternativa-

mente, a ligação do anticorpo à amostra pode ser detectada usando um anticorpo secundário tendo nele uma tal marcação detectável. Ensaios particulares incluem ensaios de ELISA, ensaios em sanduíche, radioimunoensaios e Manchação Ocidental.

Os anticorpos policlonais gerados junto com os polipeptídeos da invenção podem ser obtidos por injeção direta dos polipeptídeos em um animal ou administrando os polipeptídeos a um animal, por exemplo, um não-humano. O anticorpo assim obtido depois ligará o próprio polipeptídeo. Desta maneira, até mesmo uma seqüência codificando apenas um fragmento dos polipeptídeos pode ser usada para gerar anticorpos que podem ligar ao polipeptídeo nativo inteiro. Tais anticorpos podem ser depois usados para isolar o polipeptídeo de células que expressam aquele polipeptídeo.

Para preparação de anticorpos monoclonais, pode ser usada qualquer técnica que fornece anticorpos produzidos por culturas de linhagem de células contínuas. Exemplos incluem a técnica de hibridoma, a técnica de trioma, a técnica de hibridoma de células B humanas e a técnica de hibridoma de EBV (ver, por exemplo, Cole (1985) em Monoclonal Antibodies and Câncer Therapy, Alan R. Liss, Inc., pp. 77-96).

As técnicas descritas para a produção de anticorpos de cadeia simples (ver, por exemplo, Patente U.S. No. 4.946.778) podem ser adaptadas para produzir anticorpos de cadeia simples para os polipeptídeos da invenção. Alternativamente, camundongos transgênicos podem ser usados para expressar anticorpos humanizados a estes polipeptídeos ou fragmentos destes.

Anticorpos gerados junto com os polipeptídeos da invenção podem ser usados para triar pelos polipeptídeos similares de outros organismos e amostras. Em tais técnicas, os polipeptídeos do organismo são contatados com o anticorpo e aqueles polipeptídeos que especificamente ligam ao anticorpo são detectados. Quaisquer dos procedimentos acima descritos podem ser usados para detectar a ligação do anticorpo.

KITS

A invenção fornece kits compreendendo as composições, por

exemplo, ácidos nucléicos, cassetes de expressão, vetores, células, polipeptídeos (por exemplo, fitases) e/ou anticorpos da invenção. Os kits também podem conter material instrutivo que ensina as metodologias e usos industriais da invenção, como descrito aqui.

Os polipeptídeos da SEQ ID Nº: 2, seqüências substancialmente idênticas a esta, ou fragmentos compreendendo pelo menos 5, 10, 15, 20, 25, 30, 35, 40, 50, 75, 100, ou 150 aminoácidos sucessivos desta, também podem ser usados para gerar anticorpos que especificamente ligam aos polipeptídeos de enzima ou fragmentos. Os anticorpos resultantes podem ser usados nos procedimentos de cromatografia de imunoafinidade para isolar ou purificar o polipeptídeo ou determinar se o polipeptídeo está presente em uma amostra biológica. Em tais procedimentos, uma preparação de proteína, como um extrato, ou uma amostra biológica é contatada com um anticorpo capaz de especificamente ligar a um de um polipeptídeo da SEQ ID Nº: 2, seqüências substancialmente idênticas a esta, ou fragmentos das seqüências anteriores.

Em procedimentos de imunoafinidade, o anticorpo é ligado a um suporte sólido, como uma conta ou outra matriz de coluna. A preparação de proteína é colocada em contato com o anticorpo sob condições em que o anticorpo especificamente liga a um dos polipeptídeos da SEQ ID Nº: 2, seqüências substancialmente idênticas a esta, ou fragmento destes. Após uma lavagem para não-especificamente remover as proteínas ligadas, os polipeptídeos especificamente ligados são eluídos.

A capacidade das proteínas em uma amostra biológica de ligar ao anticorpo pode ser determinada usando qualquer de uma variedade de procedimentos familiares àqueles versados na técnica. Por exemplo, a ligação pode ser determinada marcando o anticorpo com uma marcação detectável como agente fluorescente, uma marcação enzimática ou um radioisótopo. Alternativamente, a ligação do anticorpo à amostra pode ser detectada usando um anticorpo secundário tendo nele uma tal marcação detectável. Ensaios particulares incluem ensaios de ELISA, ensaios em sanduíche, radioimunoensaios e Manchas Ocidentais.

Os anticorpos policlonais gerados com os polipeptídeos da SEQ ID Nº: 2, e seqüências substancialmente idênticas a esta, ou fragmentos compreendendo pelo menos 5, 10, 15, 20, 25, 30, 35, 40, 50, 75, 100, ou 150 aminoácidos sucessivos desta, podem ser obtidos por injeção direta dos polipeptídeos em um animal ou administrando os polipeptídeos a um animal, por exemplo, um não-humano. O anticorpo assim obtido depois liga ao próprio polipeptídeo. Até mesmo desta maneira, uma seqüência codificando apenas um fragmento do polipeptídeo pode ser usada para gerar anticorpos que podem ligar ao polipeptídeo nativo inteiro. Tais anticorpos podem depois ser usados para isolar o polipeptídeo de células que expressam aquele polipeptídeo.

Para preparação de anticorpos monoclonais, qualquer técnica que fornece anticorpos produzidos por culturas de linhagem de célula contínuas pode ser usada. Exemplos incluem a técnica de hibridoma (Kohler e Milstein, Nature, 256:495-497, 1975), a técnica de trioma, a técnica de hibridoma de células B humanas (Kozbor et al., Immunol. Today 4:72. 1983), e a técnica de hibridoma de EBV (Cole, et al., 1985, em Monoclonal Antibodies and Cancer Therapy, Alan R. Liss, Inc., pp. 77-96).

Técnicas descritas para a produção de anticorpos de cadeia simples (Patente U.S. No. 4.946.778) podem ser adaptadas para produzir anticorpos de cadeia simples para os polipeptídeos, por exemplo, da SEQ ID Nº: 2, e fragmentos desta. Alternativamente, camundongos transgênicos podem ser usados para expressar anticorpos humanizados para estes polipeptídeos ou fragmentos.

Os anticorpos gerados com um polipeptídeo da SEQ ID Nº: 2, seqüências substancialmente idênticas a esta, ou fragmentos compreendendo pelo menos 5, 10, 15, 20, 25, 30, 35, 40, 50, 75, 100, ou 150 aminoácidos sucessivos desta, podem ser usados para triar por polipeptídeos similares de outros organismos e amostras. Em tais técnicas, polipeptídeos do organismo são contatados com o anticorpo e aqueles polipeptídeos que especificamente ligam ao anticorpo são detectados. Quaisquer dos procedimentos acima descritos podem ser usados para detectar ligação de anticorpo. Um

tal ensaio de triagem é descrito em "Methods for Measruting Cellulase Activities", Methods in Enzymology, Vol 160, pp. 87-116.

Como aqui usado o termo "seqüência de ácidos nucléicos exposta na SEQ ID Nº: 1" abrange uma seqüência de ácidos nucléicos como exposto na SEQ ID Nº: 1, uma seqüência substancialmente idêntica a esta, fragmentos de qualquer uma ou mais das seqüências anteriores, seqüências de nucleotídeos homólogos à SEQ ID Nº: 1, ou homólogo ao fragmentos da SEQ ID Nº: 1, e seqüências complementares a todas as seqüências precedentes. Os fragmentos incluem porções da SEQ ID Nº: 1 compreendendo pelo menos 10, 15, 20, 25, 30, 35, 40, 50, 75, 100, 150, 200, 300, 400, ou 500 nucleotídeos sucessivos da SEQ ID Nº: 1, e seqüências substancialmente idênticas a esta. Seqüências homólogas e fragmentos da SEQ ID Nº: 1, e seqüências substancialmente idênticas a esta, referem-se a uma seqüência tendo pelo menos 99%, 98%, 97%, 96%, 95%, 90%, 85%, 80%, 75% ou 70% de homologia a estas seqüências. Homologia pode ser determinada usando quaisquer dos programas de computação e parâmetros aqui descritos, incluindo FASTA versão 3.0t78 com os parâmetros predefinidos. Seqüências homólogas também incluem seqüências de RNA em que as uridinas substituem as timinas nas seqüências de ácidos nucléicos como exposto SEQ ID Nº: 1. As seqüências homólogas podem ser obtidas usando quaisquer dos procedimentos aqui descritos ou podem ser o resultado da correção de um erro de seqüenciação. Será apreciado que as seqüências de ácidos nucléicos da invenção podem ser representadas no formato de caractere simples tradicional (Ver o encarte de Stryer, Lubert. Biochemistry, 3ª edição. W. H. Freeman and Co., Nova Iorque.) ou em qualquer outro formato que registra a identidade dos nucleotídeos em uma seqüência.

Como aqui usado o termo "uma seqüência de polipeptídeos exposta na SEQ ID Nº: 2" abrange a seqüência de polipeptídeos como exposto na SEQ ID Nº: 2, seqüências substancialmente idênticas a esta que são codificadas por uma seqüência como exposto na SEQ ID Nº: 1, seqüências de polipeptídeos homólogas aos polipeptídeos da SEQ ID Nº: 2, e seqüências substancialmente idênticas a esta, ou fragmentos de quaisquer das seqüên-

cias precedentes. Seqüências de polipeptídeos homólogas referem-se a uma seqüência de polipeptídeos tendo pelo menos 99%, 98%, 97%, 96%, 95%, 90%, 85%, 80%, 75% ou 70% de homologia a uma das seqüências de polipeptídeos da invenção. Homologia pode ser determinada usando quaisquer dos programas de computação e parâmetros aqui descritos, incluindo FASTA versão 3.0t78 com os parâmetros predefinidos ou com quaisquer parâmetros modificados. As seqüências homólogas podem ser obtidas usando quaisquer dos procedimentos aqui descritos ou podem ser o resultado da correção de um erro de seqüenciação. Os fragmentos de polipeptídeos compreendem pelo menos 5, 10, 15, 20, 25, 30, 35, 40, 50, 75, 100, ou 150 aminoácidos sucessivos dos polipeptídeos da SEQ ID Nº: 2, e seqüências substancialmente idênticas a esta. Será apreciado que os polipeptídeos da invenção podem ser representados no formato de caractere simples tradicional ou formato de carta três (Ver o encarte de Starrier, Lubert. Biochemistry, 3ª edição. W. H. Freeman and Co, Nova Iorque.) ou em qualquer outro formato que relaciona à identidade dos polipeptídeos em uma seqüência.

Motivos que podem ser detectados usando os programas acima incluem seqüências que codificam zíperes de leucina, motivos de hélice-volta-hélice, sítios de glicosilação, sítios de ubiquitinação, hélices alfa, e lâminas betas, seqüências sinais que codificam peptídeos de sinal que direcionam a secreção das proteínas codificadas, seqüências implicadas na sobre-regulação de transcrição como homeocaixas, extensões acídicas, sítios ativos enzimáticos, sítio de ligação ao substratos e sítios de clivagem enzimática.

As seqüências de polinucleotídeos isoladas, seqüência de polipeptídeos, variantes e mutantes destas podem ser medidas para retenção de característica de atividade biológica para a enzima da presente invenção, por exemplo, em um ensaio para detectar atividade de fitase enzimática (Food Chemical Codex, 4ª Ed.). Tais enzimas incluem formas truncadas de fitase, e variantes como variantes de deleção e de inserção da seqüência de polipeptídeos como exposto na SEQ ID Nº: 2. Estas fitases têm termotolerância. Ou seja, a fitase tem uma atividade específica residual de cerca de

90% após tratamento a 70°C, durante 30 minutos e aproximadamente 50% após tratamento a 75°C, durante 30 minutos. A termotolerância das fitases da invenção é vantajosa usando a enzima como um aditivo de alimentação uma vez que a alimentação pode ser moldada, granulada ou peletizada em uma temperatura alta.

Por exemplo, em um aspecto, a invenção fornece uma matriz de liberação de enzima peletizada comestível e método de uso para liberação de fitase para um animal, por exemplo como um suplemento nutricional. A matriz de liberação da enzima facilmente libera uma enzima de fitase, como uma tendo a seqüência de aminoácidos da SEQ ID Nº: 2 ou pelo menos 30 aminoácidos contíguos desta, em meios aquosos, como, por exemplo, o fluido digestivo de um animal. A matriz de liberação de enzima da invenção é preparada de um veículo comestível granulado selecionado de tais componentes como germe de grão sem óleo, feno, alfafa, capim-rabo-de-rato, casca de soja, farinha de semente de girassol, farinha de trigo e outros, que facilmente dispersam a enzima recombinante nele contida em meios aquosos. Em uso, a matriz de liberação de enzima peletizada comestível é administrada a um animal para liberação de fitase para o animal. Substratos com base em grão adequados podem compreender ou derivar de qualquer grão comestível adequado, como trigo, milho, soja, sorgo, alfafa, cevada e outros. Um substrato com base em grão exemplar é um substrato com base em milho. O substrato pode ser derivado de qualquer parte adequada do grão, por exemplo, um germe de grão, aprovado para uso em alimentação de animais, como germe de milho que é obtido em um processo de moenda a úmido ou seco. O germe de grão pode compreender germe gasto que é um germe de grão do qual óleo foi expelido, como por compressão ou hexano ou outra extração de solvente. Alternativamente, o germe de grão é extraído por expelidor, ou seja, o óleo foi removido comprimindo.

A matriz de liberação de enzima da invenção é na forma de partículas, péletes ou grânulos plurais distintos. Por "grânulos" é significado partículas que estão comprimidas ou compactadas, como por uma peletização, extrusão, ou compactação similar para remover água da matriz. Tal

compressão ou também compactação das partículas promove coesão intra-partículas das partículas. Por exemplo, os grânulos podem ser preparados peletizando o substrato com base no grão em um moinho de pélete. As péletes assim preparadas são moídas ou esmigalhadas para um tamanho de grânulo adequado para uso como um adjuvante na alimentação de animais. Considerando que a matriz seja aprovada para uso em alimentação de animais, ela pode ser usada como um diluente para liberação das enzimas na alimentação de animais.

A matriz de liberação de enzima pode ser na forma de grânulos tendo um tamanho de grânulo que varia de cerca de 4 a cerca de 400 malhas (USS); ou cerca de 8 a cerca de 80 malhas; ou cerca de 14 a cerca de 20 malhas. Se o germe de grão for gasto por meio de extração por solvente, o uso de um agente de lubricidade como óleo de milho pode ser necessário no peletizador, mas um tal agente de lubricidade comum não é necessário se o germe for extraído por expelidor. Em outros aspectos da invenção, a matriz é preparada por outros processos de compactação ou compressão como, por exemplo, por extrusão do substrato com base em grão através de uma matriz e moagem do extrusado para um tamanho de grânulo adequado.

A matriz de liberação de enzima pode também incluir um componente de polissacarídeo como um agente de coesão para intensificar a coesão dos grânulos da matriz. Acredita-se que o agente de coesão forneça grupos hidroxila adicionais que intensificam a ligação entre as proteínas do grão dentro do grânulo da matriz. É também acreditado que os grupos hidroxila adicionais assim funcionam intensificando a ligação de hidrogênio das proteínas ao amido e a outras proteínas. O agente de coesão pode estar presente em qualquer quantidade adequada para intensificar a coesão dos grânulos da matriz de liberação de enzima. Agentes coesivos adequados incluem uma ou mais de dextrinas, maltodextrinas, amidos, como amido de milho, farinhas, celulósicos, semi-celulósicos e outros. Por exemplo, a porcentagem de germe de grão e do agente de coesão na matriz (não incluindo a enzima) é 78% de farinha de germe de milho e 20% em peso de amido de milho.

Porque a matriz de liberação de enzima da invenção é feita de materiais biodegradáveis, a matriz pode estar sujeita ao desperdiçamento, como pela moldagem. Para prevenir ou inibir tal moldagem, a matriz pode incluir um inibidor de molde, como um sal de propionato que pode estar presente em qualquer quantidade suficiente para inibir a moldagem da matriz de liberação de enzima desse modo fornecendo uma matriz de liberação em uma formulação estável que não requer refrigeração.

A enzima de fitase contida na matriz de liberação de enzima da invenção e métodos é um aspecto de uma fitase termotolerante, como aqui descrita, para resistir à inativação da fitase durante fabricação onde podem ser empregadas temperaturas elevadas e/ou vapor para preparar a matriz de liberação de enzima peletizada. Durante a digestão de alimentação contendo a matriz de liberação de enzima da invenção, fluidos digestivos aquosos causarão liberação da enzima ativa. Outros tipos de enzimas termotolerantes e suplementos nutricionais que são termoestáveis podem ser também incorporados na matriz de liberação para liberação sob qualquer tipo de condições aquosas.

Um revestimento pode ser aplicado às partículas de matriz de enzima da invenção para muitos diferentes propósitos, como para adicionar um sabor ou suplemento nutricional à alimentação de animais, para retardar a liberação dos suplementos e enzimas da alimentação de animais em condições gástricas, e outros. Ou, o revestimento pode ser aplicado para alcançar um objetivo funcional, por exemplo, sempre que for desejável reduzir a liberação da enzima das partículas das matrizes ou para controlar as condições sob as quais a enzima será liberada. A composição do material de revestimento pode ser de modo que seja seletivamente fragmentada por um agente ao qual ela é suscetível (como calor, ácido ou base, enzimas ou outras químicas). Alternativamente, dois ou mais revestimentos suscetíveis a tais diferentes agentes de fragmentação podem ser consecutivamente aplicados às partículas da matriz.

A invenção também é direcionada a um processo para preparar uma matriz de liberação de enzima. De acordo com a invenção, o processo

compreende fornecer partículas plurais distintas de um substrato com base em grão em um tamanho de partícula adequado para uso como uma matriz de liberação de enzima, em que as partículas compreendem uma enzima de fitase codificada pela SEQ ID Nº: 2 ou pelo menos 30 aminoácidos sucessivos desta. O processo pode incluir compactação ou compressão das partículas da matriz de liberação de enzima em grânulos que podem ser realizadas mediante peletização. O inibidor de molde e agente de coesão, quando usados, podem ser adicionados em qualquer momento adequado, e podem ser misturados com o substrato com base em grão nas proporções desejadas antes da peletização do substrato com base em grão. O teor de umidade na alimentação do moinho de pélete pode ser nas faixas acima expostas com respeito ao teor de umidade no produto acabado, e preferivelmente é cerca de 14-15%. A umidade pode ser adicionada à carga de alimentação na forma de uma preparação aquosa da enzima trazer a carga de alimentação para este teor de umidade. A temperatura no moinho de pélete pode ser trazida para cerca de 82°C com vapor. O moinho de pélete pode ser operado sob qualquer condição para transmitir trabalho suficiente à carga de alimentação para fornecer as péletes. O processo de peletização em si é um processo eficaz no custo para remover água da composição contendo enzima.

Em um aspecto, o moinho de pélete é operado com uma matriz de 0,31 cm (1/8 in) por 5,08 cm (2 in) a 45,35 kg/min (100 lb/min). pressão a 82°C. Para fornecer péletes que depois são esmigalhados em um esmigalhador de moinho de pélete para fornecer partículas plurais distintas tendo um tamanho de partícula capaz de atravessar uma tela de 8 malha mas ser retido em uma tela de 20 malha.

As fitases termotolerantes aqui descritas podem ter temperaturas ótimas altas e podem ter alta resistência ao calor ou tolerância ao calor. Desse modo, as fitases da invenção podem ser efetuadas em temperaturas de reações enzimáticas normalmente consideradas acima de ótima. As fitases da invenção também podem realizar reações enzimáticas após serem expostas às altas temperaturas (termotolerância sendo a capacidade de reter atividade enzimática em temperaturas onde a fitase do tipo selvagem é

ativa após previamente ser exposta a temperaturas altas, até mesmo se a temperatura alta puder inativar ou diminuir a atividade da enzima, também ver definição de termotolerância, acima). O gene que codifica a fitase de acordo com a presente invenção (por exemplo, como exposto na SEQ ID Nº: 1) pode ser usado na preparação das fitases (por exemplo usando GSSM como aqui descrito) tendo características diferentes daquelas da fitase da SEQ ID Nº: 2 (em termos de pH ótimo, temperatura ótima, resistência ao calor, estabilidade a solventes, atividade específica, afinidade ao substrato, capacidade de secreção, taxa de tradução, controle de transcrição e outros). Além disso, o polinucleotídeo da SEQ ID Nº: 1 pode ser empregado para triar as fitases variantes preparadas pelos métodos aqui descritos para determinar aquelas tendo uma atividade desejada, como termoestabilidade ou termotolerância melhorada ou modificada. Por exemplo, Patente U.S. No. 5.830.732, descreve um ensaio de triagem para determinar a termotolerância de uma fitase.

Um exemplo *in vitro* de um tal ensaio de triagem é o ensaio a seguir é para a detecção da atividade de fitase: atividade de fitase pode ser medida incubando 150 µl da preparação da enzima com 600 µl de 2 mM fitato de sódio em 100 mM tampão de Tris HCl, pH 7,5, suplementado com 1 mM CaCl2 durante 30 minutos a 37°C. Após incubação, a reação é interrompida adicionando 750 µl de 5% de ácido tricloroacético. Fosfato liberado foi medido com fosfato padrão espectrofotometricamente a 700 nm após adicionar 1500 µl do reagente de cor (4 volumes de 1,5% de molibdato de amônio em 5,5% de ácido sulfúrico e 1 volume de 2,7% de sulfato ferroso; Shimizu, 1992). Uma unidade da atividade da enzima é definida como a quantidade de enzima requerida para liberar um µmol Pi por min sob condições de ensaio. A atividade específica pode ser expressa em unidades de atividade de enzima por mg de proteína. A enzima da presente invenção tem atividade enzimática com respeito à hidrólise de fitato com inositol e fosfato livre.

Em um aspecto, a invenção atual fornece um método de hidrolisar fitato compreendendo contatar o fitato com uma ou mais das novas mo-

léculas de fitase aqui descritos (por exemplo, SEQ ID Nº: 2). Conseqüentemente, a invenção fornece um método para catalisar a hidrólise de fitato com inositol e fosfato livre com liberação de minerais do complexo de ácido fítico. O método inclui contatar um substrato de fitato com uma quantidade eficaz degradante de uma enzima da invenção, como a enzima mostrada na SEQ ID Nº: 2. O termo "quantidade eficaz degradante" refere-se à quantidade de enzima que é requerida para degradar pelo menos 50% do fitato, quando comparado ao fitato não contatado com a enzima. 80% do fitato pode ser degradado.

Em outro aspecto, a invenção fornece um método para hidrolisar ligações de fosfo-mono-éster em fitato. O método inclui administrar uma quantidade eficaz de moléculas de fitase da invenção (por exemplo, SEQ ID Nº: 2), para render inositol e fosfato livre. Uma "quantidade" eficaz" refere-se à quantidade de enzima que é requerida para hidrolisar pelo menos 50% das ligações de fosfo-mono-éster, quando comparado ao fitato não contatado com a enzima. Em um aspecto, pelo menos 80% das ligações são hidrolisadas.

Em um aspecto particular, quando desejado, as moléculas de fitase podem ser usadas em combinação com outros reagentes, como outros catalisadores; para realizar alterações químicas (por exemplo hidrólise) nas moléculas de fitato e/ou em outras moléculas da(s) fonte(s) de substrato. De acordo com este aspecto, as moléculas de fitase e o(s) reagente(s) adicional(is) não inibirão um ao outro. As moléculas de fitase e o(s) reagente(s) adicional(is) podem ter um efeito aditivo geral, ou, alternativamente, as moléculas de fitase e o(s) reagente(s) adicional(is) pode(m) ter um efeito sinergístico geral.

Fontes relevantes das moléculas de fitato do substrato incluem gêneros alimentícios, gêneros alimentícios potenciais, subprodutos de gêneros alimentícios (tanto subprodutos *in vitro* quanto subprodutos *in vivo*, por exemplo produtos *ex vivo* de reação e produtos excrementais de animal), precursores de gêneros alimentícios, e qualquer outra fonte material de fitato.

Em um aspecto não-limitativo, a fitase recombinante pode ser consumida por organismos e reter as atividades sob consumo. Em outra exemplificação, métodos transgênicos podem ser usados para alcançar expressão da fitase recombinante - por exemplo, em uma forma controlada (métodos estão disponíveis para controlar expressão de moléculas transgênicas de formas tempo-específicas e tecido-específicas).

Em um aspecto, a atividade de fitase no material de fonte (por exemplo uma fonte de planta transgênica ou um hospedeiro procariótico recombinante) pode ser aumentada no consumo; este aumento na atividade pode ocorrer, por exemplo, na conversão de uma molécula de fitase precursora em pró-forma a uma enzima significativamente mais ativa em uma forma mais madura onde a dita conversão pode resultar, por exemplo, da ingestão e digestão da fonte de fitase. Hidrólise do substrato de fitato pode ocorrer qualquer hora ao contatar a fitase com o fitato; por exemplo, isto pode ocorrer antes da ingestão ou após ingestão ou tanto antes quanto após a ingestão do substrato ou da enzima ou ambos. É apreciado adicionalmente que o substrato de fitato pode ser contatado com - além da fitase - um ou mais reagentes adicionais, como outra enzima que também tinha sido aplicada ou diretamente ou após purificação de seu material de fonte.

É apreciado que o(s) material(is) de fonte da fitase pode ser contatado diretamente com o(s) material(is) de fonte do fitato; por exemplo em moagem ou mastigação *in vitro* ou *in vivo* de qualquer uma ou ambas a(s) fontes(s) de fitase e a(s) fonte(s) de fitato. Alternativamente a enzima de fitase pode ser purificada longe do(s) material(is) de fonte, ou o substrato de fitato pode ser purificado longe do(s) material(is) de fonte, ou a enzima de fitase e o substrato de fitato podem ser purificados longe do(s) material(is) de fonte antes do contato da enzima de fitase com o substrato de fitato. É apreciado que uma combinação de reagentes purificados e não-purificados - incluindo enzima(s) ou substrato(s) ou ambos - pode ser usada.

É apreciado que mais de um material de fonte pode ser usado como uma fonte de atividade de fitase. Este é útil como um modo para alcançar uma liberação em intervalos do(s) reagente(s) do(s) material(is) de

fonte, onde a liberação dos reagentes diferentes de seus materiais de fonte ocorre diferentemente, por exemplo como materiais de fonte ingeridos são digeridos *in vivo* ou como materiais de fonte são processados em aplicações *in vitro*. O uso de mais de um material de fonte de atividade de fitase também é útil para obter atividades de fitase sob uma faixa de condições e flutuações destes, que podem ser encontradas - como uma faixa de valores de pH, temperaturas, salinidades e intervalos de tempo - por exemplo durante as etapas de diferentes processamentos de uma aplicação. O uso de materiais de fonte diferentes também é útil para obter reagentes diferentes, como exemplificado por uma ou mais formas ou isômeros de fitase e/ou fitato e/ou outros materiais.

É apreciado que um material de fonte simples, uma tal espécie de planta transgênica (ou partes de planta desta), pode ser um material de fonte de fitase e fitato; e que enzimas e substratos podem ser diferentemente compartimentados dentro da dita fonte simples - por exemplo segregada versus não-segregada, diferentemente expressa e/ou ter abundâncias diferenciais em diferentes partes da planta ou órgãos ou tecidos ou em compartimentos subcelulares dentro da mesma parte da planta ou órgão ou tecido. Purificação das moléculas de fitase nela contidas pode compreender isolando e/ou também processando uma ou mais partes da planta desejáveis ou órgãos ou tecidos ou compartimentos subcelulares.

Em um aspecto particular, esta invenção fornece um método de catalisar reações *in vivo* e/ou *in vitro* usando sementes que contêm quantidades intensificadas de enzimas. O método compreende adicionar sementes do tipo transgênico, não-selvagem, por exemplo, em uma forma moída, a uma mistura de reação e deixar as enzimas nas sementes aumentar a taxa de reação. Adicionando as sementes diretamente à mistura de reação o método fornece uma solução para o processo mais caro e incômodo de extrair e purificar a enzima. Também são fornecidos métodos de tratamento por meio dos quais um organismo carecendo de uma provisão suficiente de uma enzima é administrado à enzima na forma de sementes de uma ou mais espécies de planta, por exemplo, espécies de planta transgênica, contendo

quantidades intensificadas da enzima. Detalhes adicionais relativos a este método estão na literatura pública e/ou são conhecidos ao versado. Em uma exemplificação não-limitativa particular, tal literatura publicamente disponível inclui Patente U.S. No. 5.543.576 (Van Ooijen et al.) e Patente U.S. No. 5.714.474 (Van Ooijen et al.), embora estas referências não ensinem as moléculas inventivas do presente pedido de patente e ao invés ensinam o uso de fitases fúngicas.

Em um aspecto não-limitativo particular, as moléculas de fitase imediatas são úteis para gerar formas de vida do sistema digestivo recombinante (ou micróbios ou flora) e para a administração das ditas formas de vida do sistema digestivo recombinante aos animais. Administração pode ser opcionalmente executada sozinha ou em combinação com outras enzimas e/ou com outras formas de vida que podem fornecer atividade enzimática em um sistema digestivo, onde as ditas outras enzimas e as ditas formas de vida podem ser recombinantes ou outras. Por exemplo, a administração pode ser executada em combinação com bactérias xilanolíticas.

Em um aspecto não-limitativo, a presente invenção fornece um método para macerar grãos de milho ou sorgo em água morna contendo dióxido de enxofre na presença de uma preparação de enzima compreendendo uma ou mais enzimas degradantes de fitina, por exemplo, em uma tal quantidade que a presente fitina no milho ou sorgo é substancialmente degradada. A preparação da enzima pode compreender fitase e/ou ácido fosfatase e opcionalmente outro material de enzimas de planta degradantes. O tempo de maceração pode ser 12 a 18 horas. As macerações podem ser interrompidas por uma etapa de moagem intermediária, reduzindo o tempo de maceração. Em um aspecto, os grãos de milho ou sorgo são macerados em água morna contendo dióxido de enxofre na presença de uma preparação de enzima que inclui uma ou mais enzimas degradantes de fitina, como fitase e ácido fosfatases, para eliminar ou grandemente reduzir ácido fítico e os sais de ácido fítico. Detalhes adicionais relativos a este método estão na literatura pública e/ou são conhecidos ao versado. Em uma exemplificação, tal literatura publicamente disponível inclui Patente U.S. No. 4.914.029 (Ca-

ransa et al.) e EP 0321004 (Vaara et al.), embora estas referências não ensinem as moléculas inventivas do presente pedido de patente.

Em um aspecto não-limitativo, a presente invenção fornece um método para obter uma massa de pão tendo propriedades físicas desejáveis como não-viscosidade e elasticidade e um produto de pão de qualidade superior como um volume específico compreendendo adicionar as moléculas de fitase à massa de pão. Em um aspecto, as moléculas de fitase da invenção atual são adicionadas a uma preparação em andamento de massa de pão que é subseqüentemente formada e assada. Detalhes adicionais com relação a esse método estão na literatura pública e/ou são conhecidos ao versado. Em uma exemplificação, tal literatura publicamente disponível inclui JP 03076529 (Hara et al.), embora esta referência não ensine as moléculas de fitase inventivas do presente pedido de patente.

Em um aspecto não-limitativo, a presente invenção fornece um método para produzir gêneros alimentícios de feijão-soja melhorados. São combinados feijões-sojas com as moléculas de fitase da invenção atual para remover ácido fítico dos feijões-sojas, desse modo produzindo os gêneros alimentícios de feijão-soja que são melhorados em seu fornecimento de nutrientes de traço essenciais para consumir organismos e em sua digestibilidade de proteínas. Em um aspecto, na produção de leite de feijão-soja, as moléculas de fitase da invenção atual são adicionadas ou colocadas em contato com os feijões-sojas para reduzir o teor de ácido fítico. Em uma exemplificação não-limitativa, o processo de aplicação pode ser acelerado agitando o leite de feijão-soja junto com a enzima sob aquecimento ou conduzindo uma reação do tipo mistura em um recipiente de agitação usando uma enzima imobilizada. Detalhes adicionais relativos a este método estão na literatura pública e/ou são conhecidos ao versado. Em uma exemplificação não-limitativa particular, tal literatura publicamente disponível inclui JP 59166049 (Kamikubo et al.), embora esta referência não ensine as moléculas inventivas do presente pedido de patente.

Em um aspecto, a invenção atual fornece um método de produzir um produto de admissão para água potável ou alimentação de animais

em forma de fluido, e compreendendo usar as misturas minerais e misturas de vitaminas, e também novas moléculas de fitase da invenção atual. Em um aspecto, é alcançada uma mistura corretamente dosada e misturada de nutrientes necessários para o organismo consumindo sem qualquer risco de precipitação e destruição dos minerais/vitaminas importantes, embora ao mesmo tempo seja feita a utilização ótima do fosfato ligado à fitina na alimentação. Detalhes adicionais relativos a este método estão na literatura pública e/ou são conhecidos ao versado. Em uma exemplificação, tal literatura publicamente disponível inclui EP 0772978 (Bendixen et al.), embora esta referência não ensine as moléculas inventivas do presente pedido de patente.

É apreciado que as moléculas de fitase da invenção atual também possam ser usadas para produzir outros gêneros alimentícios bebíveis alcoólicos e não-alcoólicos (ou bebidas) com base no uso de moldes e/ou em grãos e/ou em outras plantas. Estes gêneros alimentícios bebíveis incluem licores, vinhos, bebidas alcoólicas misturadas (por exemplo refresco de vinho, outros cafés alcoólicos como cafés irlandeses, etc.), cervejas, quase-cervejas, sucos, extratos, homogenatos e purês. Em um aspecto, as moléculas de fitase imediatamente divulgadas são usadas para gerar versões transgênicas de moldes e/ou grãos e/ou outras plantas úteis para a produção de tais gêneros alimentícios bebíveis. Em outro aspecto, as moléculas de fitase imediatamente divulgadas são usadas como ingredientes adicionais no processo industrial e/ou no teor final de tais gêneros alimentícios bebíveis. Detalhes adicionais relativos a este método estão na literatura pública e/ou são conhecidos ao versado. Porém - devido à novidade da invenção atual - referências na literatura publicamente disponível não ensinam as moléculas inventivas divulgadas imediatamente.

Em outra exemplificação não-limitativa, a presente invenção fornece meios para obterem objetivo refinado tendo uma quantidade reduzida de fitina e um teor aumentado de inositol. Um tal objetivo pode ter - através de efeitos diretos e/ou psicogênicos - uma ação preventiva em doença hepática, arteriosclerose e outras doenças. Em um aspecto, um objetivo é pro-

duzido de arroz Koji multiplicando um molde de arroz Koji tendo atividade de fitase alta como um material bruto. É apreciado que as moléculas de fitase da invenção atual podem ser usadas para produzir um molde útil com atividade intensificada (por exemplo, um molde transgênico) e/ou exogenamente adicionadas para aumentar os efeitos de um molde de Koji. A cepa é adicionada ao arroz fervido e Koji é produzido por um procedimento convencional. Em uma exemplificação, o Koji preparado é usado, o arroz inteiro é preparado em dois estágios e o objetivo é produzido em temperatura de Objetivo constante de 15°C para dar o Objetivo refinado objetivo tendo uma quantidade reduzida de fitina e uma quantidade aumentada de inositol. Detalhes adicionais relativos a este método estão na literatura pública e/ou são conhecidos ao versado. Em uma exemplificação, tal literatura publicamente disponível inclui JP 06153896 (Soga et al.) e JP 06070749 (Soga et al.), embora estas referências não ensinem as moléculas inventivas do presente pedido de patente.

Em um aspecto não-limitativo, a presente invenção fornece um método para obter um absorbefaciente capaz de promover a absorção de minerais incluindo cálcio ingerido sem ser digeridos por sucos gástricos ou sucos intestinais a um baixo custo. Em um aspecto, o absorbefaciente mineral contém um hidrolisado parcial de ácido fítico como um ingrediente ativo. Um hidrolisado parcial do ácido fítico pode ser produzido através de hidrólise do ácido fítico ou seus sais usando as novas moléculas de fitase da invenção atual. O tratamento com as moléculas de fitase pode ocorrer ou um em tratamento sozinho e/ou um de combinação (para inibir ou aumentar o efeito final), e é seguido inibindo a hidrólise dentro de uma faixa para não liberar todos os radicais de fosfato. Detalhes adicionais relativos a este método estão na literatura pública e/ou são conhecidos ao versado. Em uma exemplificação não-limitativa particular, tal literatura publicamente disponível inclui JP 04270296 (Hoshino), embora referência na literatura publicamente disponível não ensine as moléculas inventivas do presente pedido de patente.

Em um aspecto não-limitativo, a presente invenção fornece um método (e seus produtos) para produzir uma composição de enzima tendo

um aditivo ou preferivelmente uma atividade de hidrolização de fitato sinergística; a dita composição compreende as novas moléculas de fitase da invenção atual e um ou mais reagentes adicionais para alcançar uma composição que é útil para um tratamento de combinação. Em um aspecto, o tratamento de combinação da presente invenção é alcançado com o uso de pelo menos duas fitases de especificidade de posição diferente, isto é qualquer combinação de 1, 2, 3, 4, 5 e 6-fitases. Combinando as fitases de especificidade de posição diferente um efeito aditivo ou sinergístico é obtido. Composições como comida e alimentação ou comida e aditivo de alimentação compreendendo tais fitases em combinação também estão incluídas nesta invenção como também seus processos para a preparação. Detalhes adicionais relativos a este método estão na literatura pública e/ou são conhecidos ao versado. Em uma exemplificação, tal literatura publicamente disponível inclui WO9 830681 (Ohmann et al.), embora referências na literatura publicamente disponível não ensinem o uso das moléculas inventivas do presente pedido de patente.

Em outro aspecto, o tratamento de combinação da presente invenção é alcançado com o uso de um ácido fosfatase tendo atividade de hidrolização do fitato a um pH de 2,5, em uma razão baixa que corresponde a um pH 2,5:5,0 perfil de atividade de cerca de 0,1:1,0 a 10:1, preferivelmente de cerca de 0,5:1,0 a 5:1, ou de cerca de 0,8:1,0 a 3:1, ou de cerca de 0,8:1,0 a 2:1. A composição de enzima preferivelmente apresenta uma eficiência de hidrolização de fitato sinergética mais alta mediante tratamento térmico. A composição de enzima é útil no tratamento de gêneros alimentícios (comida bebível e sólida, alimentação e produtos de alimentação) para melhorar a hidrólise de fitato. Detalhes adicionais relativos a este método estão na literatura pública e/ou são conhecidos ao versado. Em uma exemplificação não-limitativa particular, tal literatura publicamente disponível inclui Patente U.S. No. 5.554.399 (Vanderbeke et al.) e Patente U.S. No. 5.443.979 (Vanderbeke et al.) que ambas ensinam o uso de fitases fúngicas (em particular Aspergillus).

Em um aspecto não-limitativo, a presente invenção fornece um

método (e seus produtos) para produzir uma composição compreendida da nova enzima de ação de fitato imediata em combinação com uma ou mais enzimas adicionais que agem nos polissacarídeos. Tais polissacarídeos podem ser selecionados do grupo que consiste em arabinanas, fructanas, fucanas, galactanas, galacturonanas, glucanas, mananas, xilanas, levan, fucoidan, carragenina, galactocarolose, pectina, ácido péctico, amilose, pululan, glicogênio, amilopectina, celulose, carboxilmetilcelulose, hidroxipropilmetilcelulose, dextrano, pustulan, quitina, agarose, ceratan, condroitina, dermatan, ácido hialurônico, ácido algínico e polissacarídeos contendo pelo menos uma aldose, cetose, ácido ou amina selecionados do grupo que consiste em eritrose, treose, ribose, arabinose, xilose, lixose, alose, altrose, glicose, manose, gulose, idose, galactose, talose, eritrulose, ribulose, xilulose, psicose, frutose, sorbose, tagatose, ácido glicurônico, ácido glicônico, ácido glucárico, ácido galacturônico, ácido manurônico, glicosamina, galactosamina e ácido neuramínico.

Em um aspecto particular, a presente invenção fornece um método (e seus produtos) para produzir uma composição tendo uma atividade de hidrolização de fitato sinergística compreendendo uma ou mais novas moléculas de fitase da invenção atual, uma celulase (também pode 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 gêneros alimentícios, produtos de madeira, como produtos de papel e como soluções e sólidos de limpeza.

Em uma exemplificação não-limitativa, as moléculas de fitase imediatas são úteis em combinação com componentes celulósicos. É conhecido que celulases de muitas bactérias celulolíticas são organizadas em complexos de multienzima distintos, denominados celulossomas. As múltiplas subunidades de celulossomas são compostas de numerosos domínios funcionais que interagem entre si e com o substrato celulósico. Uma destas subunidades compreende uma classe nova distintiva de polipeptídeos de formação de estrutura não-catalítica, que seletivamente integra as várias subunidades de celulase e xilanase no complexo coesivo. Aplicação inteli-

gente de híbridos de celulossoma e construções quiméricas de domínios celulossômicos deve permitir melhor uso de biomassa celulósica e pode oferecer uma faixa extensiva de novas aplicações em pesquisa, medicina e indústria.

Em outra exemplificação não-limitativa, as moléculas de fitase imediatas são úteis - ou sozinhas ou em tratamentos de combinação - em áreas de biopolpação e bioalvejamento, onde uma redução no uso de químicas ambientalmente prejudiciais tradicionalmente usadas na indústria de polpa e papel é desejada. Tratamento de água residual representa outra área de aplicação vasta onde enzimas biológicas foram mostradas não só ser eficazes na remoção da cor mas também na bioconversão de substâncias potencialmente nocivas em bioprodutos úteis.

Em outra exemplificação não-limitativa, as moléculas de fitase imediatas são úteis para gerar formas de vida que podem fornecer pelo menos uma atividade enzimática - ou sozinho ou em tratamentos de combinação - no tratamento de sistemas digestivos de organismos. Particularmente organismos importantes a ser tratados incluem organismos não-ruminantes, embora organismos ruminantes também podem beneficiar de tal tratamento. Especificamente, é apreciado que este método pode ser executado sozinho ou em combinação com outras moléculas biológicas (por exemplo, xilanases) para gerar um hospedeiro recombinante expressando uma pluralidade de moléculas biológicas. Também é apreciado que a administração das moléculas de fitase imediatas e/ou hospedeiros recombinantes que expressam as moléculas de fitase imediatas ou pode ser executada sozinha ou em combinação com outras moléculas biológicas, e/ou formas de vida que podem fornecer atividades enzimáticas em um sistema digestivo - onde as ditas outras enzimas e as ditas formas de vida podem ser recombinantes ou outras. Por exemplo, a administração pode ser executada em combinação com bactérias xilanolíticas.

Por exemplo, além de fitato, muitos organismos não podem também digerir hemiceluloses adequadamente. Hemiceluloses ou xilanas são componentes principais (35%) de materiais de planta. Para animais rumi-

nantes, cerca de 50% de xilanas dietéticas são degradadas, mas apenas quantidades pequenas de xilanas são degradadas no intestino inferior de animais não-ruminantes e do homem. No rume, as espécies xilanolíticas principais são Butyrivibrio fibrisolvens e Bacteroides ruminicola. No cólon humano, as subespécies "a" Bacteroides ovatus e Bacteroides fragilis são bactérias xilanolíticas principais. As xilanas são quimicamente complexas, e sua degradação requer enzimas múltiplas. Expressão destas enzimas através de bactérias do intestino varia grandemente entre as espécies. Butyrivibrio Fibrisolvens produz xilanases extracelulares mas espécies Bacteroides têm atividade de xilanase ligada à célula. Caracterização bioquímica de enzimas xilanolíticas das bactérias do intestino não foi completamente terminada. Um gene de xilosidase foi clonado de B. fibrosolvens 113. Os dados das hibridizações de DNA usando um gene de xilanase clonado de B. fibrisolvens 49 indicam que este gene pode estar presente em outras cepas de B. fibrisolvens. Uma xilanase clonada de Bact. ruminicola foi transferida e altamente expressa em Bact. fragilis e Bact. uniformis. Genes de arabinosidase e xilosidase de Bact. ovatus foram clonados e ambas as atividades parecem ser catalisadas por uma nova enzima simples, bifuncional.

Conseqüentemente, é apreciado que as moléculas de fitase presentes são úteis para 1) transferir em um hospedeiro adequado (como Bact. fragilis ou Bact. uniformis); 2) alcançar expressão adequada em um hospedeiro recombinante resultante; e 3) administrar o dito hospedeiro recombinante em organismos para melhorar a capacidade dos organismos tratados de degradar fitato. Pesquisa continuada em áreas genéticas e bioquímicas proverá conhecimento e discernimento para manipulação de digestão no nível de intestino e entendimento melhorado de digestão de fibra colônica.

Detalhes adicionais relativos a este método estão na literatura pública e/ou são conhecidos ao versado. Em uma exemplificação não-limitativa particular, tal literatura publicamente disponível inclui Patente U.S. No. 5.624.678 (Bedford et al.), Patente U.S. No. 5.683.911 (Bodie et al.), Patente U.S. No. 5.720.971 (Beauchemin et al.), Patente U.S. No. 5.759.840 (Sung et al.), Patente U.S. No. 5.770.012 (Cooper), Patente U.S. No.

5.786.316 (Baeck et al.), Patente U.S. No. 5.817.500 (Hansen et al.), e artigos de jornal (Jeffries, 1996; Prade, 1996; Bayer et al., 1994; Duarte et al., 1994; Hespell e Whitehead, 1990; Wong et al., 1988), embora estas referências não ensinem as moléculas de fitase inventivas do presente pedido de patente, nem nenhum deles ensina a adição de moléculas de fitase na produção de gêneros alimentícios, produtos de madeira, como produtos de papel, e como soluções e sólidos de limpeza. Em contraste, a invenção atual ensina que as moléculas de fitase - por exemplo, as moléculas de fitase da invenção atual - podem ser adicionadas ao(s) reagente(s) divulgado(s) para obter preparações tendo uma atividade de fitase adicional. Os reagente(s) e as moléculas de fitase adicionais podem não inibir um ao outro. O(s) reagente(s) e as moléculas de fitase adicionais podem ter um efeito aditivo geral. O(s) reagente(s) e as moléculas de fitase adicionais podem ter um efeito sinergístico geral.

Em um aspecto não-limitativo, a presente invenção fornece um método (e seus produtos) para intensificação de utilização de fósforo de fitato e tratamento e prevenção de discondroplasia tibial em animais, particularmente avícula, administrando aos animais uma composição de alimentação contendo um derivado de vitamina D3 hidroxilada. O derivado de vitamina D3 pode ser administrado aos animais na alimentação contendo níveis reduzidos de cálcio e fósforo para intensificação de utilização de fósforo de fitato. Conseqüentemente, o derivado de vitamina D3 pode ser administrado em combinação com as novas moléculas de fitase da invenção atual para também intensificação da utilização de fósforo de fitato. Detalhes adicionais relativos a este método estão na literatura pública e/ou são conhecidos ao versado. Em uma exemplificação não-limitativa particular, tal literatura publicamente disponível inclui Patente U.S. No. 5.516.525 (Edwards et al.) e Patente U.S. No. 5.366.736 (Edwards et al.), Patente U.S. No. 5.316.770 (Edwards et al.) embora esta referência não ensine as moléculas inventivas do presente pedido de patente.

Em um aspecto não-limitativo, a presente invenção fornece um método (e seus produtos) para obter gêneros alimentícios que 1) compreen-

de fitina que é facilmente absorvida e utilizada em uma forma de inositol em um corpo de um organismo; 2) que é capaz de reduzir fósforo em substância excrementar; e 3) que é conseqüentemente útil para melhorar a poluição ambiental. O dito gênero alimentício é compreendido de uma admistão de um grão contendo fitina, um microorganismo de produção de ácido láctico e uma nova molécula de fitase da invenção atual. Em um aspecto, o dito gênero alimentício é produzido através da composição de uma fitina - contendo grão (preferivelmente, por exemplo farelo de trigo de arroz) com um grupo microbiano eficaz tendo uma propriedade acidofílica, produzindo ácido láctico, sem produzir ácido butírico, livre de patogenicidade, e uma fitase. Exemplos de um grupo microbiano eficaz incluem por exemplo *Streptomyces* sp. (American Type Culture Collection No. ATCC 3004) pertencentes ao grupo de actinomices e Lactobacillus sp. (IFO 3070) pertencente ao grupo de lactobacilos. Também, uma quantidade preferível de adição de um grupo microbiano eficaz é 0,2% em peso em termos do peso de corpo bacteriano com base em um material de grão. Além disso, a quantidade da adição da fitase é preferivelmente 1-2% em peso com base na fitina no material de grão. Detalhes adicionais relativos a este método estão na literatura pública e/ou são conhecidos ao versado. Em uma exemplificação não-limitativa particular, tal literatura publicamente disponível inclui JP 08205785 (Akahori et al.), embora referências na literatura publicamente disponível não ensinem as moléculas inventivas do presente pedido de patente.

Em um aspecto não-limitativo, a presente invenção fornece um método para melhorar a solubilidade de proteínas vegetais. Mais especificamente, a invenção diz respeito aos métodos para a solubilização de proteínas em fontes de proteína vegetais cujos métodos compreendem tratar a fonte de proteína vegetal com uma quantidade eficiente de uma ou mais enzimas de fitase - incluindo moléculas de fitase da invenção atual - e tratar a fonte de proteína vegetal com uma quantidade eficiente de uma ou mais enzimas proteolíticas. Em outro aspecto, a invenção fornece aditivos de alimentação de animais compreendendo uma fitase e uma ou mais enzimas proteolíticas. Detalhes adicionais relativos a este método estão na literatura

pública e/ou são conhecidos ao versado. Em uma exemplificação não-limitativa particular, tal literatura publicamente disponível inclui EP 0756457 (WO 9528850 A1) (Nielsen e Knap), embora referências na literatura publicamente disponível não ensinam as moléculas inventivas do presente pedido de patente.

Em um aspecto não-limitativo, a presente invenção fornece um método de produzir uma preparação de proteína vegetal compreendendo dispersar os materiais de fonte da proteína vegetal em água a um pH na faixa de 2 a 6 e misturar com eles as moléculas de fitase da invenção atual. O extrato acídico contendo proteína solúvel é separado e seco para render uma proteína sólida de caráter desejável. Uma ou mais proteases também podem ser usadas para melhorar as características da proteína. Detalhes adicionais relativos a este método estão na literatura pública e/ou são conhecidos ao versado. Em uma exemplificação não-limitativa particular, tal literatura publicamente disponível inclui Patente U.S. No. 3.966.971 (Morehouse et al.), embora referências na literatura publicamente disponível não ensinem as moléculas inventivas do presente pedido de patente.

Em um aspecto não-limitativo, a presente invenção fornece um método (e produtos destes) para ativar fósforo inerte em terra e/ou composto, para melhorar a taxa de utilização de um composto de nitrogênio, e suprir propagação de moldes patogênicos adicionando três reagentes, fitase, saponina e quitosana, ao composto. Em um aspecto não-limitativo o método pode compreender tratar o composto 1) adicionando microorganismos contendo fitase em meios - preferivelmente os hospedeiros recombinantes que superexpressam as novas moléculas de fitase da invenção atual - por exemplo a 100 ml de meio/100 kg de composto úmido; 2) também adicionando alternativamente uma fonte de planta contendo fitase - como farelo de trigo - por exemplo a 0,2 a 1 kg/100 kg de composto úmido; 3) adicionando uma fonte contendo saponina - como turfa, flor-de-diana e planta de iúca - por exemplo a 0,5 a 3,0 g/kg; 4) adicionando materiais contendo quitosana - como conchas pulverizadas de camarões, caranguejos, etc. - por exemplo a 100 a 300g/kg de composto úmido. Em outro aspecto não-limitativo, fontes

recombinantes compreendendo os três reagentes, fitase, saponina e quitosana, são usadas. Detalhes adicionais relativos a este método estão na literatura pública e/ou são conhecidos ao versado. Em uma exemplificação não-limitativa particular, tal literatura publicamente disponível inclui JP 07277865 (Toya Taisuke), embora referências na literatura publicamente disponível não ensinem as moléculas inventivas do presente pedido de patente.

Os fragmentos do gene de comprimento total da presente invenção podem ser usados como uma sonda de hibridação com um cDNA ou uma biblioteca genômica para isolar o DNA de comprimento total e isolar outros DNAs tendo uma similaridade de seqüência alta para o gene ou atividade biológica similar. Sondas deste tipo têm pelo menos 10, preferivelmente pelo menos 15 e até mesmo mais preferivelmente pelo menos 30 bases e podem conter, por exemplo, pelo menos 50 ou mais bases. A sonda também pode ser usada para identificar um clone de DNA que corresponde a uma transcrição de comprimento total e um clone genômico ou clones contendo o gene completo que inclui regiões reguladoras e promotoras, éxons e íntrons.

Em outro aspecto, são fornecidos organismos não-humanos transgênicos que contém uma seqüência heteróloga que codifica uma fitase da invenção (por exemplo, SEQ ID Nº: 2). Vários métodos para fazer os animais transgênicos da invenção em questão podem ser empregados. Em geral, três tais métodos podem ser empregados. Em um tal método, um embrião no estágio pró-nuclear (um "embrião unicelular") é colhido de uma fêmea e o transgene é microinjetado dentro do embrião, em cujo caso o transgene será cromossomicamente integrado às células germinais e às células somáticas do animal maduro resultante. Em outro tal método, células do tronco embrionário são isoladas e o transgene nelas incorporado por eletroporação, transfecção de plasmídeo ou microinjeção, seguido por reintrodução das células do tronco no embrião onde elas colonizam e contribuem para a linhagem genética. Métodos para microinjeção de espécies mamíferas são descritos na Patente U.S. No. 4.873.191. Em ainda outro tal método, células embrionárias são infectadas com um retrovírus contendo o transgene

por meio do qual as células germinais do embrião têm o transgene cromossomicamente nelas integrado. Quando os animais a serem feitos transgênicos forem aves, porque óvulos fertilizados aviários em geral passam por divisão celular durante as primeiras vinte horas no oviduto, microinjeção dentro do pró-núcleo do ovo fertilizado é problemática devido à inacessibilidade do pró-núcleo. Portanto, dos métodos para fazer animais transgênicos no geral acima descritos, infecção de retrovírus é preferida para espécies aviárias, por exemplo como descrito na Pat. U.S. No. 5.162,215. Se microinjeção for para ser usada com espécies aviárias, porém, um procedimento publicado por Love et al., (Biotechnol., 12 de janeiro de 1994) pode ser utilizado por meio do qual o embrião é obtido de uma galinha sacrificada aproximadamente duas horas e meia após a colocação do ovo botado, o transgene é microinjetado no citoplasma do disco germinal e o embrião é cultivado em uma concha de hospedeiro até maturidade. Quando os animais a serem feitos transgênicos são bovinos ou porcinos, microinjeção pode ser impedida pela opacidade dos óvulos assim tornando difícil de identificar os grãos através de microscopia de contraste de interferência diferencial tradicional. Para superar este problema, os óvulos podem ser centrifugados para segregar os pró-núcleos para melhor visualização primeiro.

Os "animais não-humanos" da invenção são animais bovinos, porcinos, ovinos e aviários (por exemplo, vaca, porco, ovelha, galinha). Os "animais não-humanos transgênicos" da invenção são produzidos introduzindo os "transgenes" dentro da linha germinal do animal não-humano. As células alvo embrionárias podem ser usadas em vários estágios desenvolventes para introduzir os transgenes. Diferentes métodos são usados dependendo do estágio de desenvolvimento da célula alvo embrionária. O zigoto é o melhor alvo para microinjeção. O uso de zigotos como é alvo para tradução de gene tem uma vantagem principal em que na maioria dos casos o DNA injetado será incorporado no gene de hospedeiro antes da primeira clivagem (Brinster et al., Proc. Natl. Acad. Sci USA 82:4438-4442, 1985). Como uma conseqüência, todas as células do animal transgênico não-humano levará o transgene incorporado. Isto em geral será refletido também

na transmissão eficiente do transgene na descendência do fundador uma vez que 50% das células germinais abrigarão o transgene.

O termo "transgênico" é usado para descrever um animal que inclui material genético exógeno dentro de todas as suas células. Um animal "transgênico" pode ser produzido hibridando dois animais quiméricos incluindo material genético exógeno dentro de células usadas na reprodução. Vinte e cinco percentual da descendência resultante será transgênico, isto é, animais incluindo o material genético exógeno dentro de todas de suas células em ambos os alelos, 50% dos animais resultantes incluirá o material genético exógeno dentro de um alelo e 25% não incluirão nenhum material genético exógeno.

No método de microinjeção útil na prática da invenção em questão, o transgene é digerido e purificado livre de qualquer DNA de vetor, por exemplo, através de eletroforese em gel. É preferido que o transgene inclua um promotor operativamente associado que interage com as proteínas celulares envolvidas na transcrição, resultando por fim na expressão constitutiva. Promotores úteis nesta consideração incluem aqueles do citomegalovírus (CMV), vírus de leucemia de Moloney (MLV) e vírus do herpes, como também aqueles dos genes que codificam metalotionina, actina do esqueleto, P-enolpiruvato carboxilase (PEPCK), fosfoglicerato (PGK), DHFR e timidina cinase. Promotores para repetições viróticas terminais longas (LTRs) tal como Vírus do Sarcoma de Rous também podem ser empregados. Quando os animais a serem feitos transgênicos forem os promotores aviários, os preferidos incluem aqueles para o gene de □-globina de galinha, gene de lisozima de galinha, e vírus de leucose aviária. Construções úteis em transfecção de plasmídeo de células de tronco embrionário empregarão elementos regulador adicionais bem-conhecidos na técnica como elementos intensificadores para estimular transcrição, aceptores de junção, sinais de terminação e poliadenilação e sítios de ligação ao ribossomo para permitir a tradução.

Infecção retrovirótica também pode ser usada para introduzir transgene em um animal não-humano, como descrito acima. O embrião não-

humano em desenvolvimento pode ser cultivado *in vitro* para o estágio de blastocisto. Durante este tempo, os blastômeros podem ser alvos para infecção retrovirótica (Jaenich, R., Proc. Natl. Acad Sci. USA 73:1260-1264, 1976). Infecção eficiente dos blastômeros é obtida através de tratamento enzimático para remover os pelúcidos da zona (Hogan, et al. (1986) em Manipulating the Mouse Embryo, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N. Y.). O sistema de vetor virótico para introduzir o transgene é tipicamente um retrovírus defeituoso na replicação que carrega 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). Transfecção é fácil e eficazmente obtida cultivando os blastômeros em uma monocamada de células produtoras de vírus (Van der Putten, supra; Stewart, et al., EMBO J 6: 383-388, 1987). Alternativamente, a infecção pode ser executada em um estágio posterior. Vírus ou células produtoras de vírus podem ser injetadas no blastocele (D. Jahner et al., Nature 298: 623-628, 1982). A maioria dos fundadores será mosaico para o transgene uma vez que a incorporação apenas ocorre em um subconjunto das células que formaram o animal não-humano transgênico. Também, o fundador pode conter várias inserções retrovirótica do transgene em diferentes posições no genoma que em geral segregarão na descendência. Além disso, também é possível introduzir transgenes na linhagem genética, embora com baixa eficiência, através de infecção retrovirótica intra-uterina do embrião de meia-gestação (D. Jahner et al., supra).

Um terceiro tipo de célula alvo para introdução do transgene é a célula do tronco embrionário (ES). As células de ES são obtidas de embriões de pré-implantação cultivados *in vitro* e fundidos com os 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 eficazmente introduzidos nas células de ES através de transfecção de DNA ou através de transdução mediada por retrovírus. Tais células de ES transformadas podem ser combinadas depois disso com blastocistos de um

animal não-humano. As células de ES depois disso colonizam o embrião e contribuem para a linhagem genética do animal quimérico resultante. (Para revisão ver Jaenisch, R., Science 240:1468-1474, 1988).

"Transformadas" significa uma célula dentro da qual (ou dentro de um antecessor desta) foi introduzida, por meio de técnicas de ácido nucléico recombinante, uma molécula de ácido nucléico heterólogo. "Heterólogo" refere-se a uma seqüência de ácido nucléico que origina ou de outras espécies ou é modificada de sua forma original ou a forma primariamente expressa na célula.

"Transgene" quer dizer qualquer pedaço de DNA que é inserido por artifício dentro de uma célula, e se torna parte do genoma do organismo (isto é, ou estavelmente integrado ou como um elemento extracromossomicamente estável) que desenvolve daquela célula. Um tal transgene pode incluir um gene que é em parte ou completamente heterólogo (isto é, estranho) ao organismo transgênico, ou pode representar um gene homólogo a um gene endógeno do organismo. Incluído dentro desta definição está um transgene criado mediante o fornecimento de uma seqüência de RNA que é transcrita dentro do DNA e depois incorporada no genoma. Os transgenes da invenção incluem seqüências de DNA codificando fitases tendo atividade de fitase ou polipeptídeos, e inclui polinucleotídeos que podem ser expressos em um animal não-humano transgênico. O termo "transgênico" como aqui usado adicionalmente inclui qualquer organismo cujo genoma foi alterado por manipulação *in vitro* do embrião precoce ou ovo fertilizado ou por qualquer tecnologia transgênica para induzir um golpe de gene específico. O termo "golpe de gene" como aqui usado, refere-se ao rompimento alvejado de um de gene *in vivo* com perda completa da função que foi alcançada por qualquer tecnologia transgênica familiar para aqueles na técnica. Em um aspecto, animais transgênicos tendo golpes de gene são aqueles em que o gene alvo foi feito não-funcional por uma inserção alvejada para o gene a ser feito não-funcional através de recombinação homóloga. Como aqui usado, o termo "transgênico" inclui qualquer tecnologia transgênica familiar para aqueles na técnica que podem produzir um organismo que carrega um

transgene introduzido ou uma em que um gene endógeno foi feito não-funcional ou "golpeado."

O transgene a ser usado na prática da invenção em questão é uma seqüência de DNA compreendendo uma seqüência codificando para uma fitase ou um polipeptídeo tendo atividade de fitase. Em um aspecto, um polinucleotídeo tendo uma seqüência como exposto na SEQ ID Nº: 1 ou uma seqüência codificando um polipeptídeo tendo uma seqüência como exposto na SEQ ID Nº: 2 é o transgene como o termo é aqui definido. Onde apropriado, as seqüências de DNA que codificam as proteínas tendo atividade de fitase mas diferem-se na seqüência de ácido nucléico devido à degeneração do código genético também podem ser aqui usadas, como as formas truncadas, variantes alélicas e homólogos de interespécies.

Após um embrião ter sido microinjetado, colonizado com células do tronco embrionário transfectadas ou infetadas com um retrovírus contendo o transgene (com exceção da prática da invenção em questão em espécies aviárias que são dirigidas em outro lugar aqui) o embrião é implantado no oviduto de uma fêmea pseudoprenha. A progênie conseqüente é testada para incorporação do transgene por análise de Southern blot de amostras de sangue ou tecido usando sondas específicas por transgene. PCR é particularmente útil nesta consideração. Progênie positiva (G0) é hibridada para produzir descendência (G1) que é analisada para expressão do transgene por análise de Northern blot das amostras de tecido.

Desse modo, a presente invenção inclui métodos para aumentar a absorção de fósforo no animal transgênico e/ou diminuir a quantidade de poluente no adubo do organismo transgênico em cerca de 15%, cerca de 20% ou cerca de 20%, a cerca de 50%.

Os animais contemplados para uso na prática da invenção em questão em geral são aqueles animais considerados animais domesticados incluindo animais de estimação (por exemplo, espécies caninas, felinas, aviárias etc.) e aqueles úteis para o processamento de matérias-primas de comida, isto é, avícula como galinha e peru para produção de carne e chocação de ovos, ovino como cordeiro, bovino como gado vacum para engorda e

vacas leiteiras, piscino e porcino. Para propósitos da invenção em questão, estes animais são referidos como "transgênicos" quando tal animal tiver uma seqüência de DNA heteróloga, ou uma ou mais seqüências de DNA adicionais normalmente endógenas ao animal (coletivamente referida aqui como "transgenes") cromossomicamente integradas às células germinais do animal. O animal transgênico (incluindo sua progênie) também terá o transgene fortuitamente integrado nos cromossomos de células somáticas.

Em algumas circunstâncias pode ser vantajoso liberar e expressar uma seqüência de fitase da invenção localmente (por exemplo, dentro de um tecido ou tipo de célula particular). Por exemplo, expressão local de uma fitase ou enzima digestiva no intestino de um animal ajudará na digestão e absorção, por exemplo, de fitato e fósforo, respectivamente. A seqüência de ácidos nucléicos pode ser liberada diretamente às glândulas salivais, tecido e células e/ou para células epiteliais revestindo o intestino, por exemplo. Tais métodos de liberação são conhecidos na técnica e incluem eletroporação, vetores viróticos e absorção de DNA direta. Qualquer polipeptídeo tendo atividade de fitase pode ser utilizado nos métodos da invenção (por exemplo, aqueles especificamente descritos nesta subseção 6.3.18, como também aqueles descritos em outras seções da invenção).

Por exemplo, uma construção de ácido nucléico da presente invenção compreenderá moléculas de ácido nucléico em uma forma adequada para absorção em células alvo dentro de um tecido hospedeiro. Os ácidos nucléicos podem ser na forma de DNA nu ou moléculas de RNA onde as moléculas podem compreender um ou mais genes estruturais, um ou mais genes reguladores, filamentos anti-sentido, filamentos capazes de formação de tríplex, ou outros. Comumente, o constructo de ácido nucléico incluirá pelo menos um gene estrutural sob o controle transcricional e translacional de uma região reguladora adequada. Mais usualmente, os constructos de ácido nucléico da presente invenção compreenderá ácidos nucléicos incorporados em um veículo de liberação para melhorar eficiência de transfecção, em que o veículo de liberação será dispersado dentro de partículas maiores compreendendo um material de excipiente hidrofílico seco.

Tais veículos de liberação compreendem vetores viróticos, como retrovírus, adenovírus e vírus adeno-associado que foram inativados para prevenir auto-replicação mas que mantêm a capacidade virótica nativa para ligar uma célula hospedeira alvo, liberar material genético no citoplasma da célula hospedeira alvo e promover expressão de genes estruturais ou outros que foram incorporados na partícula. Vetores de retrovírus adequados para tradução de gene mediada são descritos em Kahn et al. (1992) Circ. Res. 71:1508-1517. Uma liberação de gene de adenovírus adequada é descrita em Rosenfeld et al. (1991) Science 252:431 434. Ambos sistemas de liberação retrovirótico e adenovirótico são descritos em Friedman (1989) Science 244:1275-1281.

Um segundo tipo de veículo de liberação de ácido nucléico compreende vesículas de transfecção lipossômica, incluindo constructos lipossômicos aniônicos e catiônicos. O uso de lipossomas aniônicos requer que os ácidos nucléicos estejam dentro do lipossoma. Lipossomas catiônicos não requerem capturação do ácido nucléico e ao invés podem ser formados simplesmente misturando os ácidos nucléicos e lipossomas. Os lipossomas catiônicos avidamente ligam-se às moléculas de ácido nucléico negativamente carregadas, incluindo DNA e RNA, para render os complexos que dão eficiência de transfecção razoável em muitos tipos de célula. Ver, Farhood et al. (1992) Biochem. Biophys. Acta. 1111:239-246. Um material exemplar para formar vesículas lipossômicas é lipofectina que é composta de uma mistura equimolar de etanolamina de dioleilfosfatidila (DOPE) e dioleiloxipropil-trietilamônio (DOTMA), como descrito em Felgner e Ringold (1989) Nature 337:387-388.

Também é possível combinar estes dois tipos de sistemas de liberação. Por exemplo, Kahn et al. (1992), supra, ensina que um vetor de retrovírus pode ser combinado em uma vesícula de DEAE-dextrano catiônica para também intensificar a eficiência de transformação. Também é possível incorporar as proteínas nucleares em vesículas de liberação viróticas e/ou lipossômicas para até mesmo também melhorar as eficiências de transfecção. Ver, Kaneda et al. (1989) Science 243:375-378.

Em outro aspecto, um auxiliar digestivo contendo uma enzima como o único ingrediente ativo ou em combinação com um ou mais outros agentes e/ou enzimas é fornecido. O uso de enzimas e outros agentes auxiliares digestivos de animais de criação ou animais domesticados, não só melhora a saúde do animal e probabilidade de vida mas também ajuda a aumentar a saúde de animais de criação e na produção de gêneros alimentícios de animais de criação.

Atualmente, alguns tipos de alimentação de animais de criação (por exemplo, certa alimentação de avícula) são altamente suplementares com numerosos minerais (por exemplo, fósforo inorgânico), enzimas, fatores de desenvolvimento, drogas e outros agentes para liberação para os animais de criação. Estes suplementos substituem muitas das calorias e nutrientes naturais presentes no grão, por exemplo.

Reduzindo ou eliminando o suplemento fosforoso inorgânico e outros suplementos (por exemplo, sais minerais de traço, fatores de desenvolvimento, enzimas, antibióticos) da própria alimentação, a alimentação é capaz de portar mais nutriente e energia. Conseqüentemente, a dieta restante pode conter energia mais utilizável. Por exemplo, dietas de farinha de grão-semente oleaginosa em geral contêm cerca de energia metabolizável de 3.200 kcal por quilograma da dieta, e sais minerais não fornecem nenhuma energia metabolizável. A remoção dos minerais desnecessários e substituição por grão portanto aumenta a energia utilizável na dieta. Desse modo, a invenção é diferenciada em alimentação contendo fitase comumente usada. Por exemplo, em um aspecto, um material biocompatível é usado que é resistente à digestão pelo trato gastrintestinal de um organismo.

Em muitos organismos, incluindo, por exemplo, avícula ou pássaros como, por exemplo, galinhas, perus, gansos, patos, papagaios, pavões, avestruzes, faisões, codorniz, pombos, ema, quiuí, mergulhões, cacatiel, cacatua, canários, pingüins, flamingos e pomba, que o trato digestivo inclui uma moela que armazena e usa objetos bio-compatíveis duros (por exemplo, pedras e conchas de marisco) para ajudar na digestão das sementes ou outra alimentação consumida por um pássaro. Um trato digestivo

típico desta família geral de organismos, inclui o esôfago que contém uma bolsa, chamada um papo, onde é armazenada a comida durante um breve período de tempo. Do papo, a comida se move para o verdadeiro estômago, ou proventrículo onde ácido clorídrico e pepsina iniciam o processo de digestão. Em seguida, a comida passa para a moela que é de formato oval e grosso cercado com músculos poderosos. A função principal da moela é moer ou esmagar as partículas de comida - um processo que é auxiliado pelo pássaro que engole quantidades pequenas de pedregulho fino ou areia. Da moela, a comida passa para o duodeno. O intestino delgado dos pássaros é similar ao dos mamíferos. Há duas bolsas cegas ou "ceca", de cerca de 10,16-15,24 cm (4-6 polegadas) no comprimento na junção do intestino delgado e grosso. O intestino grosso é curto, embora consistindo principalmente no reto aproximadamente 7,62-10,16 cm (3-4 polegadas) no comprimento. O reto esvazia-se na cloaca e as fezes são excretadas através da abertura.

Objetos biocompatíveis duros consumidos (ou do contrário introduzidos) e presentes na moela fornecem um vetor útil para liberação de vários agentes enzimáticos, químicos, terapêuticos e antibióticos. Estas substâncias duras têm um período de vida de algumas horas a alguns dias e são passadas após um período de tempo. Conseqüentemente, a invenção fornece auxiliares dietéticos revestidos, impregnados (por exemplo, matrizes e membranas impregnadas) modificados para liberação de agentes digestivos ou terapêuticos úteis para um organismo. Tais auxiliares dietéticos incluem objetos que são ingeridos tipicamente por um organismo para ajudar na digestão dentro da moela (por exemplo, pedras ou areia). A invenção fornece objetos biocompatíveis que têm neles revestidos ou impregnados agentes úteis como um auxiliar digestivo para um organismo ou para a liberação de um agente terapêutico ou medicinal ou químico.

Em um aspecto, a invenção fornece um auxiliar dietético, embora tendo uma composição biocompatível designada para liberação de um agente que ajuda na digestão, em que a composição biocompatível é designada para consumo oral e se libera no trato digestivo (por exemplo, a mo-

ela) de um organismo. "Biocompatíveis" significa que a substância, em contato com um organismo hospedeiro (por exemplo, um pássaro), não suscita uma resposta prejudicial suficiente para resultar na rejeição da substância ou tornar a substância inoperável. Por exemplo, tal inoperabilidade pode ocorrer por exemplo, por formação de uma estrutura fibrótica ao redor da substância que limita a difusão de agentes nela impregnados ao organismo hospedeiro ou uma substância que resulta em um aumento na mortalidade ou morbidez no organismo devido à toxicidade ou infecção. Uma substância biocompatível pode ser não-biodegradável ou biodegradável. Em um aspecto, a composição biocompátivel é resistente à degradação ou digestão pelo trato gastrintestinal. Em outro aspecto, a composição biocompátivel tem a consistência de uma rocha ou pedra.

Um material não-biodegradável útil na invenção é um que permite ligação ou impregnação de um agente dietético. Por exemplo, tais materiais não-biodegradáveis não-limitativos incluem termoplásticos, como acrílico, modacrílico, poliamida, policarbonato, poliéster, polietileno, polipropileno, poliestireno, polissulfona, polietersulfona e fluoreto de polivinilideno. Elastômeros também são materiais úteis e incluem, por exemplo, poliamida, poliéster, polietileno, polipropileno, poliestireno, poliuretano, álcool de polivinila e silicone (por exemplo, silicone com base ou contendo sílica). A invenção fornece que a composição biocompatível pode conter uma pluralidade de tais materiais que podem ser por exemplo, misturados ou em camadas para formar misturas, copolímeros ou combinações destes.

Como aqui usado, um material "biodegradável" significa que a composição corroerá ou degradará *in vivo* para formar espécies químicas menores. Por exemplo, a degradação pode ocorrer através de processos enzimáticos, químicos ou físicos. Materiais biodegradáveis adequados contemplados para uso na invenção incluem, mas não são limitados a, poli(lactídeo)s, poli(glicolídeo)s, poli(ácido láctico)s, poli(ácido glicólico)s, polianidridos, poliortoésteres, polieterésteres, policaprolactona, poliesteramidas, policarbonato, policianoacrilato, poliuretano, poliacrilato e outros. Tais materiais podem ser misturados ou colocados em camadas para formar

misturas, copolímeros ou combinações destes.

É contemplado que um número diferente de substâncias biocompatíveis pode ser ingerido ou do contrário fornecido simultaneamente ao mesmo organismo, ou em várias combinações (por exemplo, um material antes do outro). Além disso, a substância biocompátivel pode ser designada para passagem lenta através do trato digestivo. Por exemplo, substâncias grandes ou gordurosas tendem a mover mais lentamente pelo trato digestivo, conseqüentemente, um material biocompatível tendo um tamanho grande para prevenir transcurso rápido no trato digestivo pode ser usado. Tais substâncias grandes podem ser uma combinação de substâncias não-biodegradáveis e biodegradáveis. Por exemplo, uma substância não-biodegradável pequena pode ser abrangida durante um certo tempo por uma substância biodegradável de modo que a porção biodegradável, que permite a porção não-biodegradável atravessar o trato digestivo será degradada. Além disso, é reconhecido que qualquer número de aromatizantes pode ser fornecido à substância biocompatível para ajudar no consumo.

Qualquer número de agentes sozinhos ou em combinação com outros agentes pode ser revestido na substância biocompatível incluindo polipeptídeo (por exemplo, enzimas, anticorpos, citocinas ou moléculas pequenas terapêuticas), e antibióticos, por exemplo. São listados exemplos de agentes úteis particulares na Tabela 1 e 2, abaixo. Também é contemplado que podem ser encapsuladas células no material biocompatível da invenção e usado para liberar as enzimas ou terapêuticos. Por exemplo, substâncias porosas podem ser elaboradas tendo poros grandes o bastante para as células crescerem e que estes materiais porosos possam ser depois levados para o trato digestivo. Por exemplo, a substância biocompatível pode ser compreendida de uma pluralidade de ambientes microflorais (por exemplo, porosidade diferente, pH etc.) que fornecem suporte para uma pluralidade de tipos de célula. As células podem ser geneticamente criadas para liberar uma droga particular, enzima ou química ao organismo. As células podem ser eucarióticas ou procarióticas.

TABELA 1

Classe de Tratamento	Química	Descrição
Antibióticos	Amoxicilina e Sua Combinação Injeção de Mastox (Amoxicilina e Cloxacilina)	Tratamento Contra Doenças Bacterianas Causadas Por Bactérias Gram + e Gram -
	Ampicilina e Sua Combinação Injeção de Biolox (Ampicilina e Cloxacilina)	Tratamento Contra Doenças Bacterianas Causadas Por Bactérias Gram + e Gram -.
	Nitrofurazona + Uréia Bolo de Nefrea	Tratamento De Infecções Genitais
	Trimetoprim + Sulfamedoxazol Bolo de Trizol	Tratamento De Infecções do Trato Respiratório, Infecções do Trato Gastrintestinal, Infecções Urino-Genitais.
	Metronidazol e Furazolidona Bolo de Metofur	Tratamento De Doenças Bacterianas E Protozoárias.
	Ftalilsulfatiazol, Pectina e Caulim Pectolin Bolo Suspensão	Tratamento De Diarréia Bacteriana E Não-Específica, Disenteria Bacilária E Pole de Bezerro.
Anti-helmínticos	Ectoparasiticida Ungüento de Germex (Hexacloreto de Faixa Benzeno, Hemissulfato de Proflavina e Cetrimida)	Ectoparasiticida e Anti-séptico

Continuação

Classe de Tratamento	Química	Descrição
	Endoparasiticidas > Albendazol e Sua Combinação Alben (Albendazol) Suspensão (Albendazol 2,5%) Mais Suspensão (Albendazol 5%) Bolo Forte (Albendazol 1,5 Gm.) Comprimido (Albendazol 600 Mg) Pó (Albendazol 5%, 15%)	Prevenção E Tratamento De Infestações de Verme Redondo, Solitária e Trematódeo
	Alpraz (Albendazol e Praziquantel) Comprimido	Prevenção E Tratamento De Infestação de Verme Redondo e Solitária Em Caninos e Felinos.
	Oxiclozanida e Sua Combinação Clozan (Oxyclozanida) Bolo, Suspensão	Prevenção e Tratamento De Prevenção de Infestações de Trematódeo
	Tetzan (Oxiclozanida e Tetramisol Hcl) Bolo, Suspensão	Prevenção e Tratamento De Infestações de Verme Redondo e Trematódeo
	Fluzan (Oxiclozanida e Levamisol Hcl) Bolo, Suspensão	Prevenção e Tratamento De Infestações de Verme Redondo e Imunidade Crescente
	Levamisol Injeção de Nemasol Pó de Wormnil	Prevenção e Tratamento De Infestações de Verme Redondo e Imunidade Crescente

Continuação

Classe de Tratamento	Química	Descrição
	Fenbendazol Fenzol Comprimido (Fenbendazol 150 Mg.) Bolo (Fenbendazol 1,5 Gm.) Pó (Fenbendazol 2,5% P/P)	Prevenção E Tratamento de Infestações de Verme Redondo e Solitária
Tônicos	Complexo de Vitamina B, Aminoácidos e Extrato do Fígado Injeção de Heptógeno	Tratamento De Anorexia, Hepatite, Debilidade, Emaciação de Convulsões Neurálgicas e Desenvolvimento Raquítico.
	Levulinato de Cálcio Com Vit.B_{12} e Vit.D_3 Injeção de Hilactina	Prevenção e tratamento de hipocalcemia, terapia de suporte em condições de doença (especialmente hipotermia) e tratamento de estágios precoces de raquitismos.
Suplementos de Alimentação de Animais	Minerais Essenciais, Selênio e Vitamina E Bolo de Ginolactina	Tratamento De Infertilidade Causadora por Anoestrus e Procriação Repetida Em Animais de Leiteria e Cavalos.
	Minerais Essenciais, Vitamina E e Iodo Pó de Hilactina	Infertilidade, Lactação Imprópria, Imunidade Diminuída, Desenvolvimento Raquítico e Debilidade.
	Eletrólitos Essenciais Com Vitamina C Pó de Electra - C	Diarréia, Desidratação, Antes e Após Transporte, Em temperaturas Extremas (Alta Ou Baixa) e outras Condições de tensão
	Pirenox Plus (Sódio de Diclofenac + Paracetamol), Bolo, Injeção	Tratamento De Mastite, Pirexia, Dor e Inflamação Pós-Cirúrgica, Pró-lapso De Útero, Coxeadura e Artrite.

TABELA 2

FORMULAÇÕES TERAPÊUTICAS

Produto	Descrição
Acutrim® (fenilpropanolamina)	Comprimidos supressores de apetite uma vez ao dia.
O Infusor de Baxter®	Para liberação intravenosa controlada de anticoagulantes, antibióticos, agentes quimioterapêuticos e outras drogas amplamente usadas.
Catapres-TTS® (sistema terapêutico transdérmico de clonidina)	Sistema transdérmico uma vez por semana para o tratamento de hipertensão.
Covera HS® (cloridrato de verapamila)	Comprimidos de início controlado e liberação estendida uma vez ao dia (COER-24) para o tratamento de hipertensão e angina de peito.
DynaCirc CR® (isradipina)	Comprimidos de liberação estendida uma vez ao dia para o tratamento de hipertensão.
Efidac 24® (maleato de clorfeniramina)	Comprimidos de liberação estendida uma vez ao dia para o alívio de sintomas de alergia.
Estraderm® (sistema transdérmico de estradiol)	Sistema transdérmico duas vezes por semana para tratar certos sintomas de pós-menopausa e prevenir osteoporose
Glucotrol XL® (glipizida)	Comprimidos de liberação estendida uma vez ao dia usados como um adjunto à dieta para o controle de hiperglicemia em pacientes com diabete melito não dependente de insulina.
Bolo de IVOMEC SR® (ivermectina)	Sistema de liberação ruminal para controle de longa temporada dos principais parasitas internos e externos em gado.
Minipress XL® (prazosina)	Comprimidos de liberação estendida uma vez ao dia para o tratamento de hipertensão.
NicoDerm® CQ® (sistema transdérmico de nicotina)	Sistema transdérmico usado uma vez ao dia como um auxiliar para parar de fumar para alívio de sintomas de retirada de nicotina
Procardia XL® (nifedipina)	Comprimidos de liberação estendida uma vez ao dia para o tratamento de angina e hipertensão.
Sudafed® 24 Horas (pseudofedroma)	Descongestionante nasal uma vez ao dia para alívio de resfriados, sinusite, febre do feno e outras alergias.

Continuação

Produto	Descrição
Transderm-Nitro® (sistema transdérmico de nitroglicerina)	Sistema transdérmico uma vez ao dia para a prevenção de angina de peito devido à doença de artéria coronária.
Transderm Scop® (sistema transdérmico de escopolamin)	Sistema transdérmico para a prevenção de náusea e vômito associado ao enjôo por movimento.
Volmax (albuterol)	Comprimidos de liberação estendida para alívio de broncoespasmo em pacientes com doença das vias aéreas obstrutivas reversíveis.
Actisite®	(cloridrato de tetraciclina) Fibra periodental usada como um adjunto para escamar e aplanar raízes para redução de profundidade da cavidade e sangramento na sondagem em pacientes com periodontite adulta.
ALZET®	Bombas osmóticas para pesquisa de laboratório.
Amphotec® (complexo de sulfato de colesterila de anfotericina B para injeção)	AMPHOTEC® é um tratamento fúngico para aspergilose invasiva em pacientes onde insuficiência renal ou toxicidade inaceitável impede uso de anfotericina B em doses eficazes e em pacientes com aspergilose invasiva onde a terapia com anfotericina B anterior falhou.
BiCitra® (citrato de sódio e ácido cítrico)	Agente alcalinizante usado naquelas condições onde manutenção a longo prazo de urina alcalina é desejável.
Ditropan® (cloreto de oxibutinina)	Para o alívio de sintomas de instabilidade da bexiga associada à bexiga neurogênica não inibida ou neurogênica atônica (isto é de urgência, freqüência, vazamento urinário, incontinência incontrolável, disúria).
Ditropan® XL (cloreto de oxibutinina)	É um comprimido de liberação controlada uma vez ao dia indicado para o tratamento de bexiga superativa com sintomas de incontinência urinária incontrolável, de urgência e de freqüência

Continuação

Produto	Descrição
DOXIL® (injeção de lipossoma de HCl de doxorrubicina) Duragesic® (sistema transdérmico de fentanila) CII	Sistema transdérmico de 72-horas para administração de dor crônica em pacientes que requerem analgesia contínua de opióides para dor que não pode ser controlada por meios mais amenos como combinações de acetaminofeno-opióide, analgésicos não-esteroidais, ou dosagem de PRN com opióides de ação curta.
Elmiron® (polissulfato de sódio de pentosan)	Indicado para o alívio de dor da bexiga ou desconforto associado com cistites intersticiais.
ENACT AirWatch™	Um sistema de monitoramento e administração de asma
Ethyol® (amifostina)	Indicado para reduzir a toxicidade renal cumulativa associada com administração repetida de cisplatina em pacientes com câncer ovariano avançado ou câncer do pulmão de células não-pequenas. Indicado para reduzir a incidência de xerostomia moderada a severa em pacientes que passam por tratamento de radiação de pós-operatório para câncer de cabeça e de pescoço, onde a porta de radiação inclui uma porção substancial das glândulas parótidas.
Troche de Mycelex® (clotrimazol)	Para o tratamento local de candidíase orofaríngea. Também indicado profilaticamente para reduzir a incidência de candidíase orofaríngea em pacientes imunocomprometidos por condições que incluem quimioterapia, radioterapia, ou terapia de esteróides utilizadas no tratamento de leucemia, tumores sólidos, ou transplante renal.
Neutra-Phos® (fosfato de potássio e de sódio)	um suplemento dietético/nutricional
Solução Oral de PolyCitra® -K e Cristais de PolyCitra® -K (Citrato de potássio e ácido cítrico)	Agente alcalinizante útil naquelas condições onde a manutenção a longo prazo de uma urina alcalina é desejável, como em pacientes com ácido úrico e cálculos de cistina do trato urinário, especialmente quando a administração de sais de sódio é indesejável ou contra-indicado

Continuação

Produto	Descrição
Xarope de PolyCitra® -K e LC (tricitratos)	Agente alcalinizante útil naquelas condições onde manutenção a longo prazo de uma urina alcalina é desejável, como em pacientes com ácido úrico e cálculos de cistina do trato urinário.
Progestasert® (progesterona)	Sistemas Contraceptivos Intrauterinos 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 condições associadas com uma deficiência ou ausência de testosterona endógena: (1) hipogonadismo Primário (congênito ou adquirido) ou (2) hipogonadismo Hipogonadotrópico (congênito ou adquirido).
Viadur™ (implante de acetato de leuprolida)	Implante uma vez ao ano para o tratamento paliativo de câncer prostático.

Certos agentes podem ser designados para tornarem-se ativos ou inativos sob certas condições (por exemplo, em certos pH's, na presença de um agente de ativação, etc.). Além disso, pode ser vantajoso usar as pró-enzimas nas composições da invenção. Por exemplo, pró-enzimas podem ser ativadas por uma protease (por exemplo, uma protease salival que está presente no trato digestivo ou é introduzida artificialmente no trato digestivo de um organismo). É contemplado que os agentes liberados pelas composições biocompatíveis da invenção são ativados ou inativados pela adição de um agente de ativação que pode ser ingerido pelo, ou do contrário liberado do, organismo. Outro mecanismo para controle do agente no trato digestivo é um agente sensível ao ambiente que é ativado no próprio compartimento digestivo. Por exemplo, um agente pode ser inativo em pH baixo mas ativo em pH neutro. Conseqüentemente, o agente pode ser inativo dentro do intestino mas ativo no trato intestinal. Alternativamente, o agente pode tornase ativo com respeito à presença de um microorganismo de fator específico (por exemplo, microorganismos presentes no intestino).

Conseqüentemente, os benefícios potenciais da presente invenção incluem, por exemplo, (1) redução ou possível eliminação da necessida-

de de suplementos minerais (por exemplo, suplementos fosforosos inorgânicos), enzimas, ou drogas terapêuticas para animal (incluindo peixe) da alimentação ou grão diário assim aumentando a quantidade de calorias e nutrientes presentes na alimentação, e (2) aumento da saúde e desenvolvimento dos animais incluindo domésticos e não-domésticos, por exemplo, animais avículas, porcinos, bovinos, eqüinos, caninos e felinos.

Um grande número de enzimas pode ser usado nos métodos e composições da presente invenção além das fitases da invenção. Estas enzimas incluem enzimas necessárias para própria digestão das comidas consumidas, ou para o próprio metabolismo, ativação ou derivação das químicas, pró-drogas ou outros agentes ou compostos liberados para o animal por meio do trato digestivo. Exemplos de enzimas que podem ser liberadas ou incorporadas nas composições da invenção, incluem, por exemplo, enzimas intensificadoras da alimentação selecionadas do grupo que consiste em 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 arabinogalactana endo-1,4-β-galactosidases e arabinogalactana 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 de pectina degradantes, em particular pectinases, pectinesterases, pectina liases, poligalacturonases, arabinanases, rhamnogalacturonases, rhamnogalacturonana acetil esterases, rhamnogalacturonana-α-rhamnosidase, pectato liases e α-galacturonisidases, mananases, β-manosidases, manana acetil esterases, xilana acetil esterases, proteases, xilanases, arabinoxilanases e enzimas lipolíticas como lipases, fitases e cutinases. Fitases, além das fitases tendo uma seqüência de aminoácidos como exposto na SEQ ID Nº: 2, podem ser usadas nos métodos e composições da invenção.

Em um aspecto, a enzima usada nas composições (por exemplo, um auxiliar dietético) da presente invenção é uma enzima de fitase que é estável sob aquecimento e é resistente ao calor e catalisa a hidrólise enzimática de fitato, isto é, a enzima é capaz de satisfazer biologicamente e

recuperar a atividade após um breve período (isto é, 5 a 30 segundos), ou mais longo, por exemplo, minutos ou horas, sob exposição às temperaturas acima de 50º C.

Uma "alimentação" e uma "comida", respectivamente, significa qualquer dieta natural ou artificial, refeição ou outros ou componentes de tais refeições intencionados ou adequados para serem comidos, tomados, digeridos, por um animal e um ser humano, respectivamente. "Auxiliar dietéticos", como aqui usado, denota, por exemplo, uma composição contendo agentes que fornecem um agente terapêutico ou digestivo a um animal ou organismo. Um "auxiliar dietético," tipicamente não é uma fonte de absorção calórica para um organismo, em outras palavras, um auxiliar dietético não é tipicamente uma fonte de energia para o organismo, mas de preferência é uma composição que é comida além da "alimentação" ou "comida" típica.

Em vários aspectos da invenção, composição da alimentação é fornecida compreendendo uma proteína de fitase recombinante tendo pelo menos trinta aminoácidos contíguos de uma proteína tendo uma seqüência de aminoácidos da SEQ ID Nº: 2; e um gênero alimentício contendo fitato. Como será conhecido por aqueles versados na técnica, tais composições podem ser preparadas de vários modos, incluindo mas não limitadas, na forma de pélete com ou sem aditivos revestidos com polímero, na forma granulada, e por secagem por pulverização. Por via de exemplo não-limitativo, os ensinamentos na técnica direcionados à preparação de alimentação incluem as Publicações Internacionais Nos. WO0070034 A1, WO0100042 A1, WO0104279 A1, WO0125411 A1, WO0125412 A1 e EP 1073342A.

Um agente ou enzima (por exemplo, uma fitase) pode exercer seu efeito *in vitro* ou *in vivo*, isto é antes de entrar ou no estômago ou na moela do organismo, respectivamente. Uma ação combinada é também possível.

Embora qualquer enzima possa ser incorporada em um auxiliar dietético, referência é aqui feita à fitase como uma exemplificação dos métodos e composições da invenção. Um auxiliar dietético da invenção inclui uma

enzima (por exemplo, uma fitase). Em geral, um auxiliar dietético contendo uma composição de fitase é líquido ou seco.

Composições líquidas não necessitam conter qualquer coisa mais que a enzima (por exemplo uma fitase), preferivelmente em uma forma altamente purificada. Porém, usualmente um estabilizante como glicerol, sorbitol ou mono propileno glicol é também adicionado. A composição líquida também pode compreender outros aditivos, como sais, açúcares, conservantes, agentes de ajuste de pH, proteínas, fitato (um substrato de fitase). Composições líquidas típicas são pastas fluidas aquosas ou com base em óleo. As composições líquidas podem ser adicionadas a uma composição biocompatível para liberação lenta. Preferivelmente a enzima é adicionada a uma composição auxiliar dietética que é um material biocompatível (por exemplo, biodegradável ou não-biodegradável) e inclui a adição de células recombinantes, por exemplo, em microcontas porosas.

As composições secas podem ser composições secas de pulverização, em cujo caso a composição não necessita conter nada mais que a enzima em uma forma seca. Porém, as composições usualmente secas são denominadas granulados que podem ser facilmente misturados com uma comida ou componentes de alimentação, ou mais preferivelmente, formam um componente de uma pré-mistura. O tamanho de partícula dos granulados de enzima preferivelmente é compatível com que o dos outros componentes da mistura. Isto fornece um meio seguro e conveniente de incorporar enzimas na alimentação de animais. Preferivelmente os granulados são biocompatíveis e mais preferivelmente estes granulados biocompatíveis são não-biodegradáveis.

Os granulados de aglomeração revestidos por uma enzima podem ser preparados usando técnica de aglomeração em um misturador de alto cisalhamento. Os granulados de absorção são preparados tendo grãos de um material veículo para absorver/ser revestido pela enzima. Preferivelmente o material veículo é um material biocompatível não-biodegradável que simula o papel de pedras ou areia na moela de um animal. Os materiais enchedores típicos usados nas técnicas de aglomeração incluem sais, como

sulfato de dissódio. Outros enchedores são caulim, talco, silicato de alumínio de magnésio e fibras de celulose. Opcionalmente, também são incluídos aglutinantes como dextrinas nos granulados de aglomeração. Os materiais veículos podem ser qualquer material biocompatível que inclui materiais biodegradáveis e não-biodegradáveis (por exemplo, rochas, pedras, cerâmicas, vários polímeros). Opcionalmente, os granulados são revestidos com uma mistura de revestimento. Tal mistura compreende os agentes de revestimento, preferivelmente agentes de revestimento hidrofóbicos, como óleo de palma hidrogenado e sebo de carne de boi, e, se desejado, outros aditivos, como carbonato de cálcio ou caulim.

Adicionalmente, as composições de auxiliar dietético (por exemplo, composições de auxiliar dietético de fitase) podem conter outros substituintes como agentes de coloração, compostos de aroma, estabilizantes, vitaminas, minerais, outras enzimas intensificadoras de alimentação ou comida etc. Um aditivo típico usualmente compreende um ou mais compostos como vitaminas, minerais ou enzimas intensificadoras de alimentação e os veículos e/ou excipientes adequados.

Em um aspecto, as composições de auxiliar dietético da invenção adicionalmente compreendem uma quantidade eficaz de uma ou mais enzimas intensificadoras de alimentação, em particular enzimas intensificadoras de alimentação selecionadas do grupo que consiste em 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 arabinogalactana 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 degradantes de pectina, em particular pectinases, pectinesterases, pectina liases, poligalacturonases, arabinanases, rhamnogalacturonases, rhamnogalacturonana acetil esterases, rhamnogalacturonana-α-rhamnosidase, pectate liases e α-galacturonisidases, mananases, β-manosidases, manana acetil esterases, xilana acetil esterases, proteases, xilanases, arabinoxilanases e enzimas

lipolíticas como lipases, fitases e cutinases.

O auxiliar dietético de animais da invenção é suplementado ao animal monogástrico antes ou simultaneamente com a dieta. Em um aspecto, o auxiliar dietético da invenção é suplementado simultaneamente ao animal monogástrico com a dieta. Em outro aspecto, o auxiliar dietético é adicionado à dieta na forma de um granulado ou um líquido estabilizado.

Uma quantidade eficaz de uma enzima em um auxiliar dietético da invenção é de cerca de 10-20.000; de cerca de 10 a 15.000, de cerca de 10 a 10.000, de cerca de 100 a 5.000, ou de cerca de 100 a cerca de 2.000 FYT/kg de auxiliar dietético.

Exemplos não-limitativos de outros usos específicos da fitase da invenção estão no processamento de soja e na fabricação de inositol ou seus derivados.

A invenção também diz respeito a um método para reduzir níveis de fitato em adubo de animal, em que o animal é alimentado com um auxiliar dietético contendo uma quantidade eficaz da fitase da invenção. Como declarado no princípio do presente pedido de patente, um efeito importante disto é reduzir a poluição de fosfato do ambiente.

Em outro aspecto, o auxiliar dietético é um veículo magnético. Por exemplo, um veículo magnético contendo uma enzima (por exemplo, uma fitase) distribuída em um, sobre um ou através de um veículo magnético (por exemplo, uma conta magnética porosa), pode ser distribuído em uma área alta em fitato e pode ser colhido através de ímãs após um período de tempo. Tal distribuição e recoletagem de contas reduzem a poluição adicional e permitem o uso de novo das contas. Além disso, o uso de tais contas magnéticas *in vivo* permite a localização do auxiliar dietético em um ponto no trato digestivo onde, por exemplo, a atividade de fitase pode ser realizada. Por exemplo, um auxiliar dietético da invenção contendo enzimas digestivas (por exemplo, uma fitase) pode ser localizado na moela do animal através de justaposição de um ímã próximo à moela do animal após o animal consumir um auxiliar dietético de veículos magnéticos. O ímã pode ser removido após um período de tempo que permite o auxiliar dietético atravessar o trato di-

gestivo. Além disso, os veículos magnéticos são adequados para remoção do organismo após sacrifício ou ajuda na coleta.

Quando o auxiliar dietético for uma partícula porosa, tais partículas são tipicamente impregnadas por uma substância com a qual é desejado liberar lentamente para formar uma partícula de liberação lenta. Tais partículas de liberação lentas não só podem ser preparadas impregnando as partículas porosas com a substância que é desejado liberar, mas também dissolvendo a substância desejada primeiro na primeira fase de dispersão. Neste caso, as partículas de liberação lentas preparadas pelo método em que a substância a ser liberada é dissolvida primeiro na primeira fase de dispersão também está dentro do escopo e espírito da invenção. Por exemplo, as partículas ocas porosas podem ser impregnadas por uma substância de liberação lenta como um medicamento, química agrícola ou enzima. Em particular, quando as partículas ocas porosas impregnadas por uma enzima são feitas de polímeros biodegradáveis, as partículas em si podem ser usadas como uma química agrícola ou fertilizante, e elas não têm nenhum efeito adverso no ambiente. Em um aspecto as partículas porosas são magnéticas na natureza.

As partículas ocas porosas podem ser usadas como um suporte de biorreator, em particular um suporte de enzima. Portanto, é vantajoso preparar o auxiliar dietético utilizando um método de uma liberação lenta, por exemplo encapsulando a enzima do agente em uma microvesícula, como um lipossoma, da qual a dose é liberada no curso de vários dias, preferivelmente entre cerca de 3 a 20 dias. Alternativamente, o agente (por exemplo, uma enzima) pode ser formulado para liberação lenta, como incorporação em um polímero de liberação lenta do qual a dosagem de agente (por exemplo, enzima) é lentamente liberada no curso de vários dias, por exemplo de 2 a 30 dias e pode variar até a vida do animal.

Como é conhecido na técnica, lipossomas são em geral derivados de fosfolipídeos ou outras substâncias de lipídio. Lipossomas são formados por cristais líquidos hidratados mono ou multilamelares são dispersos que em um meio aquoso. Qualquer lipídio não-tóxico, fisiologicamente acei-

tável e metabolizável capaz de formar lipossomas pode ser usado. As presentes composições em forma de lipossoma podem conter estabilizantes, preservativos, excipientes e outros além do agente. Alguns lipídios preferidos são os fosfolipídeos e as colinas de fosfatidila (lecitinas), tanto naturais quanto sintéticos. Os métodos para formar lipossomas são conhecidos na técnica. Por exemplo, ver Prescott, Ed., Methods in Cell Biology, Volume XIV, Academic Press, Nova Iorque, N. I. (1976), pág. 33 et seq.

Também dentro do escopo da invenção é o uso de uma fitase da invenção durante a preparação da comida ou preparações de alimentação ou aditivos, isto é, a fitase exerce sua atividade de fitase apenas durante a fabricação e não é ativa na comida ou produto de alimentação final. Este aspecto é por exemplo relevante ao fazer e assar a massa. Conseqüentemente, a fitase ou levedura recombinante que expressa fitase pode ser impregnada dentro de uns, sobre uns ou através de uns veículos magnéticos, distribuídos na massa ou meio de comida, e recolhidos através de ímãs.

O auxiliar dietético da invenção pode ser administrado sozinho a animais em um veículo biocompatível (por exemplo, um biodegradável ou não-biodegradável) ou em combinação com outros agentes aditivos de digestão. O auxiliar dietético desta invenção pode ser facilmente administrado como uma cobertura cozinhando-o ou misturando-o diretamente na alimentação de animais ou fornecido separado da alimentação, através de dosagem oral separada, por injeção ou através de meios transdérmicos ou em combinação com outros compostos comestíveis relacionados ao desenvolvimento, as proporções de cada um dos compostos na combinação que são dependentes do organismo particular ou do problema sendo dirigido e do grau de resposta desejado. Deve ser entendido que a dosagem dietética específica administrada em qualquer caso dado será ajustada de acordo com os compostos específicos que são administrados, o problema sendo tratado, a condição do indivíduo e os outros fatos pertinentes que podem modificar a atividade do ingrediente eficaz ou a resposta do indivíduo, como é bem-conhecido por aqueles versados na técnica. Em geral, ou uma única dose diária ou dosagens diárias divididas podem ser empregadas, como é

bem-conhecido na técnica.

Se administradas separadamente da alimentação de animais, as formas do auxiliar dietético podem ser preparadas combinando-as com veículos comestíveis farmaceuticamente aceitável para produzir formulações ou de liberação imediata ou de liberação lenta, como é bem-conhecido na técnica. Tais veículos comestíveis podem ser sólidos ou líquidos como, por exemplo, amido de milho, lactose, sacarose, flocos de soja, óleo de amendoim, azeite de oliva, óleo de gergelim e propileno glicol. Se um veículo sólido for usado, a forma de dosagem dos compostos pode ser em comprimidos, cápsulas, pós, trociscos ou pastilhas ou cobertura como formas microdispersáveis. Se um veículo líquido for usado, cápsulas de gelatina macias, ou xarope ou suspensões líquidas, emulsões ou soluções podem ser a forma de dosagem. As formas de dosagem também podem conter adjuvantes, como preservativos, estabilizantes, agentes umectantes ou emsulficantes, promotores de solução, etc. Elas também podem conter outras substâncias terapeuticamente valiosas. Um processo para preparar um veículo comestível granulado em temperatura alta para liberação da enzima quando ingerida é descrito no Pedido de Patente U.S. Co-pendente Série No. 09/910.579, depositado em 20 de julho de 2001.

Desse modo, umas vantagens significativas da invenção incluem por exemplo, 1) facilidade de fabricação do ingrediente ativo carregado de composições biocompatíveis; 2) versatilidade relacionada à classe de polímeros e/ou ingredientes ativos que podem ser utilizados; 3) rendimentos mais altos e eficiências de carregamento; e 4) a provisão de formulações de liberação contínua que liberam agentes ativos, intatos *in vivo*, desse modo fornecendo liberação controlada de um agente ativo em um período estendido de tempo. Além disso, outra vantagem é devido à liberação local do agente no trato digestivo (por exemplo, a moela) do organismo. Como aqui usado a frase "contido dentro" denota um método para formular um agente em uma composição útil para liberação controlada, em um período estendido de tempo do agente.

Nas composições de liberação contínua ou de liberação lenta da

invenção, uma quantidade eficaz de um agente (por exemplo, uma enzima ou antibiótico) será utilizada. Como aqui usado, liberação contínua ou liberação lenta refere-se à liberação gradual de um agente de um material biocompatível, em um período estendido de tempo. A liberação contínua pode ser contínua ou descontínua, linear ou não-linear, e esta pode ser realizada usando uma ou mais composições biodegradáveis ou não-biodegradáveis, carregamentos de droga, seleção de excipientes, ou outras modificações. Porém, será reconhecido que pode ser desejável fornecer uma composição de liberação "rápida" que fornece liberação rápida uma vez consumida pelo organismo. Também será entendido que "liberação" necessariamente não significa que o agente é liberado do veículo biocompatível. De preferência em um aspecto, a liberação lenta abrange ativação lenta ou ativação ininterrupta de um agente presente na composição biocompatível. Por exemplo, uma fitase não necessita ser liberada da composição biocompatível para ser eficaz. Neste aspecto, a fitase é imobilizada na composição biocompatível.

A alimentação de animais pode ser qualquer farinha orgânica contendo proteína empregada para normalmente satisfazer para os requerimentos dietéticos dos animais. Muitas de tais farinhas contendo proteína são tipicamente compostas primariamente de milho, farinha de feijão-soja ou uma mistura de farinha de milho/feijão-soja. Por exemplo, produtos comercialmente disponíveis típicos alimentados em galinhas incluem Egg Maker Complete, um produto de alimentação de avícula de Land O'Lakes AG Services, como também Country Game and Turkey Grower um produto de Agwa, Inc., (também ver The Emu Farmer's Handbook por Phillip Minnaar e Maria Minnaar). Ambos destes produtos comercialmente disponíveis são exemplos típicos de alimentações de animais com qual o auxiliar dietético e/ou a fitase de enzima presente(s) pode(m) ser incorporado(s) para reduzir ou eliminar a quantidade de absorção suplementar de fósforo, zinco, manganês e ferro requerida em tais composições.

A presente invenção é aplicável à dieta de numerosos animais que são aqui definidos como incluindo mamíferos (incluindo seres humanos), ave e peixe. Em particular, a dieta pode ser empregada com mamíferos co-

mercialmente significativos como porcos, gado, ovelha, cabras, roedores de laboratório (ratos, camundongos, hamsters e gerbilos), animais de pele como visom e raposa, e animais de zoológico como macacos e chipanzés, como também mamíferos domésticos como gatos e cachorros. Espécies aviárias comercialmente significativas típicas incluem galinhas, perus, patos, gansos, faisões, ema, avestruz, mergulhões, quiuí, pombas, papagaios, cockatiel, cacatua, canários, pingüins, flamingos e codorniz. Peixes comercialmente cultivados como truta também podem ser beneficiados dos auxiliares dietéticos divulgadas aqui. Outros peixes que podem beneficiar-se incluem, por exemplo, peixe (especialmente em um aquário ou ambiente de aquacultura, por exemplo, peixe tropical), peixe-dourado e outra carpa ornamental, peixe-gato, truta, salmão, tubarão, arraia, solha, linguado, tilapia, medaka, guppy, molinésia, plati, espada, paulistinha e loach.

MEDIÇÕES DE PARÂMETROS METABÓLICOS

Os métodos da invenção envolvem evolução de célula inteira, ou engenharia de célula inteira, de uma célula para desenvolver uma cepa de novas células tendo um novo fenótipo modificando a composição genética das células onde a composição genética é modificada por adição à célula de um ácido nucléico da invenção. Para detectar o novo fenótipo, pelo menos um parâmetro metabólico de uma célula modificada, é monitorado na célula em uma estrutura de "tempo real" ou "on-line". Em um aspecto, uma pluralidade de células é monitorada, como uma cultura de células, em "tempo real" ou "on-line". Em um aspecto, uma pluralidade de parâmetros metabólicos é monitorada em "tempo real" ou "on-line".

A análise de fluxo metabólica (MFA) é com base em uma estrutura bioquímica conhecida. Um matriz linearmente metabólica independente é construída com base na lei de conservação de massa e na hipótese de estado pseudopronto (PSSH) nos metabólitos intracelulares. Praticando os métodos da invenção, redes metabólicas são estabelecidas, incluindo a:

- identidade de todos os substratos, produtos e metabólitos intermediários da via,
- identidade de todas as reações químicas interconvertendo os metabólitos

da via, a estequiometria das reações da via,

- a identidade de todas as enzimas catalisando as reações, a cinética das reações enzimáticas,

- as interações reguladoras entre os componentes da via, por exemplo, interações alostéricas, interações de enzima-enzima etc,

- a compartimentalização intracelular de enzimas ou qualquer outra organização supramolecular das enzimas, e,

- a presença de quaisquer gradientes de concentração de metabólitos, enzimas ou moléculas efetoras ou barreiras de difusão ao seu movimento.

Uma vez que a rede metabólica para uma certa cepa é construída, a apresentação matemática por noção de matriz pode ser introduzida para estimar os fluxos metabólicos intracelulares se os dados de metabolismo on-line estiverem disponíveis.

O fenótipo metabólico conta com as alterações da rede metabólica inteira dentro de uma célula. O fenótipo metabólico conta com a alteração de utilização da via com respeito às condições ambientais, regulação genética, estado desenvolvente e ao genótipo, etc. Em um aspecto dos métodos da invenção, após o cálculo de MFA on-line, o comportamento dinâmico das células, seu fenótipo e outras propriedades são analisadas investigando a utilização da via. Por exemplo, se o fornecimento de glicose for aumentado e o oxigênio diminuído durante a fermentação de levedura, a utilização das vias respiratórias será reduzida e/ou parada, e a utilização das vias fermentativas dominará.

O controle do estado fisiológico das culturas celulares será possível após a análise da via. Os métodos da invenção podem ajudar a determinar como manipular a fermentação determinando como alterar o fornecimento do substrato, temperatura, uso de indutores, etc. para controle do estado fisiológico das células para remover direção desejável. Praticando os métodos da invenção, os resultados de MFA podem ser comparados também com os dados de transcriptoma e proteoma para projetar experimentos e protocolos por engenharia metabólica ou embaralhamento de gene, etc.

Praticando os métodos da invenção, qualquer fenótipo novo ou

modificado pode ser conferido e detectado, incluindo características novas ou melhoradas na célula. Qualquer aspecto de metabolismo ou desenvolvimento pode ser monitorado.

MONITORAMENTO DA EXPRESSÃO DE UMA TRANSCRIÇÃO DE mRNA

Em um aspecto da invenção, o fenótipo criado compreende aumentar ou diminuir a expressão de uma transcrição de mRNA ou gerar novas transcrições em uma célula. A transcrição de mRNA, ou mensagem pode ser detectada e quantificada por qualquer método conhecido na técnica, incluindo, por exemplo, Northern Blots, reações de amplificação quantitativas, hibridação com arranjos e outros. Reações de amplificação quantitativas incluem, por exemplo, PCR quantitativa, incluindo, por exemplo, reação em cadeia de polimerase de transcrição inversa quantitativa, ou RT-PCR; RT-PCR em tempo real quantitativa, ou " RT-PCR cinética em tempo real " (ver, por exemplo, Kreuzer (2001) Br. J. Haematol. 114:313-318; Xia (2001) Transplantation 72:907-914).

Em um aspecto da invenção, o fenótipo criado é gerado golpeando a expressão de um gene homólogo. A seqüência de codificação do gene ou um ou mais elementos transcripcionais de controle podem ser golpeados, por exemplo, intensificadores de promotores. Desse modo, a expressão de uma transcrição pode ser completamente removida ou apenas diminuída.

Em um aspecto da invenção, o fenótipo criado compreende aumentar a expressão de um gene homólogo. Isto pode ser realizado golpeando um elemento de controle negativo, incluindo um elemento regulador transcricional que age em cis ou trans, ou, mutagênese de um elemento de controle positivo.

Como debatido em detalhes abaixo, uma ou mais, ou, todas as transcrições de uma célula podem ser medidas por hibridação de uma amostra compreendendo transcrições da célula, ou, ácidos nucléicos representativos ou complementares às transcrições de uma célula, por hibridação com ácidos nucléicos imobilizados em um arranjo.

MONITORAMENTO DE EXPRESSÃO DE POLIPEPTÍDEOS, PEPTÍDEOS E AMINOÁCIDOS

Em um aspecto da invenção, o fenótipo criado compreende aumentar ou diminuir a expressão de um polipeptídeo ou gerar novos polipeptídeos em uma célula. Os polipeptídeos, peptídeos e aminoácidos podem ser detectados e quantificados por qualquer método conhecido na técnica, incluindo, por exemplo, ressonância magnética nuclear (RMN), espectrofotometria, radiografia (radiomarcação de proteína), eletroforese, eletroforese capilar, cromatografia líquida de alto desempenho (HPLC), cromatografia de camada fina (TLC), cromatografia de hiperdifusão, vários métodos imunológicos, por exemplo imunoprecipitação, imunodifusão, imunoeletroforese, radioimunoensaios (RIAs), ensaios imunossorventes ligados à enzima (ELISAs), ensaios imunofluorescentes, eletroforese em gel (por exemplo, SDS-PAGE), manchação com anticorpos, separador de célula ativada fluorescente (FACS), espectrometria de massa por pirólise, Espectrometria Infravermelha de Transformação de Fourier, espectrometria de Raman, GC-MS e Eletropulverização de LC e espectrometrias de massa por eletroporação em tandém de LC de revestimento e outros. Novas bioatividades também podem ser triadas usando os métodos, ou variações destes, descritos na Patente U.S. No. 6.057.103. Além disso, como debatido em detalhes abaixo, um ou mais, ou, todos os polipeptídeos de uma célula podem ser medidos usando um arranjo de proteína.

Marcação com ^{13}C fracionária biossinteticamente direcionada de aminoácidos proteinogênicos pode ser monitorada alimentando uma mistura de compostos uniformemente marcados com ^{13}C e desmarcados com composto de fonte de carbono dentro de uma rede de biorreação. Análise do padrão de marcação resultante permite tanto uma caracterização inclusiva da topologia da rede quanto a determinação de razões de fluxo metabólico dos aminoácidos; ver, por exemplo, Szyperski (1999) Metab. Eng. 1:189-197.

Os exemplos a seguir são intencionados ilustrar, mas não limitar, a invenção. Embora os procedimentos descritos nos exemplos sejam típicos

daqueles que podem ser usados para realizar certos aspectos da invenção, outros procedimentos conhecidos àqueles versados na técnica também podem ser usados.

EXEMPLO 1

O gene de appA de fita do tipo selvagem de E. coli (cepa K12) (SEQ ID Nº: 3) (Figura 13) que codifica uma fitase do tipo selvagem (SEQ ID Nº: 4) (Figura 14), foi usado para preparar um polinucleotídeo modificado por GSSM. O polinucleotídeo modificado tendo uma seqüência como exposto na SEQ ID Nº: 1 (Figura 1A) codifica uma fitase não-glicosilada (SEQ ID Nº: 2) (Figura 1B). Especificamente GSSM foi empregado para encontrar mutações de ponto simples que intensificaram a termotolerância do appA de K12 de E. coli. Oito polinucleotídeos variantes que continham mutações de ponto que intensificaram a termotolerância foram identificados. Estas oito mutações foram combinadas em uma única proteína como mostrado nas Figuras 8A e 8B.

Os polinucleotídeos do tipo selvagem e mutagenizados foram expressados em E. coli e purificados à homogeneidade. No ensaio de tolerância térmica, 100 µL de 0,01 mg/ml de proteína em 100 mM MOPS / pH 7,0 foram aquecidos a 37°C, 50°C, 60°C, 70°C, 80°C ou 90°C em um termociclizador de pesquisa de RJ. Na conclusão dos 5 minutos na temperatura, as amostras foram esfriadas para 4°C e incubadas em gelo. Um ensaio de atividade foi operado usando 40 µL da solução de enzima em 1,5 ml de 100 mM NaOAc / 4 mM fitato / pH 4,5 a 37°C. Alíquotas de 60 µL foram retiradas em intervalos de 2 minutos e adicionadas a 60 µL da solução desenvolvente/interruptora de cor do ensaio de TNO, que é conhecido na técnica como o padrão de indústria para detectar fosfato em uma solução, como descrito em A. J. Engelen et al. ("Related Articles Simples and rapid determination of phytase activity," J. AOAC Int. maio-junho de 1994 77(3):760-4). Claramente, a enzima modificada, SEQ ID Nº: 2, contendo 8 alterações de aminoácidos quando comparada à enzima de appA do tipo selvagem de E. coli, é tolerante em temperaturas maiores que as da enzima do tipo selvagem. (ver Figura 3)

EXEMPLO 2

ESTABILIDADE DE ENZIMA DE FITASE EM CONDIÇÕES DE DIGESTIBILIDADE SIMULADAS

O presente exemplo mostra o efeito de um fluido intestinal gástrico simulado em digestão de fitase glicosilada e não-glicosilada da SEQ ID Nº: 2. O percentual das atividades residuais (com base nas taxas iniciais) do K12 de E. coli *in vitro* digeriu e a fitase não-glicosilada da SEQ ID Nº: 2 foi representado em gráfico versus o tempo. Uma concentração padrão de fluido intestinal gástrico simulado (SGIF) contendo 2 mg/ml de NaCl, 6 M HCl e 3,2 mg/ml de pepsina foi preparada como descrito. O pH da solução era cerca de 1,4 e não foi ajustado. O ensaio de digestibilidade *in vitro* foi executado adicionando 1:4 (vol:vol) de fitase à solução de digestão e incubando imediatamente a 37°C para iniciar a reação de digestão.

Alíquotas da mistura de reação de digestão foram removidas em vários intervalos de tempo e ensaiadas para atividade de fitase residual usando o ensaio de TNO. Cada um dos ensaios foi executado pelo menos duas vezes. Uma curva exponencial com a equação y = Ae-kt foi ajustada aos dados. As meia-vidas das proteínas foram determinadas usando a equação t 1/2 = ln 2 / k. A meia-vida da fita de E. coli de K12 foi apenas 2,7 ± 0,2 minuto enquanto a fitase não-glicosilada da SEQ ID Nº: 2 teve uma meia-vida de 8,4 ± 1,1 minuto. Portanto, as mutações na fitase de K12 de E. coli do tipo selvagem intensificaram a estabilidade da enzima sob condições de digestibilidade simuladas *in vitro*. Ver Figura 4.

EXEMPLO 3

GLICOSILAÇÃO ESTABILIZA FITASE PARA DIGESTÃO DE PEPSINA

Experimentos foram conduzidos para avaliar o efeito de glicosilação na meia-vida de atividade de enzima da fitase exposta à pepsina, usada como um fluido gástrico intestinal simulado uma vez que pepsina é um dos componentes principais do fluido gástrico intestinal. Os resultados dos estudos que examinam a meia-vida da fitase exposta à pepsina são apresentados na Figura 5. Estes resultados indicaram que formas glicosiladas de fitase têm meia-vida mais longa que as formas não-glicosiladas da enzima.

Análise de computador fornece um meio de prognosticar resíduos de aminoácidos putativos que são pós-translacionalmente modificados através de glicosilação. A prognosticação dos sítios glicosilados da fitase foi terminada usando o programa de Prognosticação de Modificação Pós-translacional na ampla rede mundial no endereço expasy.ch. A identificação dos peptídeos glicosilados foi mapeada pelo programa PeptideMass no mesmo website. Os sítios de glicosilação prognosticados para fitase estão presentes na Figura 6.

Estudos foram depois empreendidos para determinar o tipo de glicosilação na fitase expressa em *Pichia pastoris* e S. cerevisiae. Após purificação da proteína dos respectivos organismos, as cadeias O-glicosiladas putativas foram removidas da proteína por adição de 1 mU de O-glicosidase (Roche Molecular Biochemicals, Alemanha) a 50 µg de fitase em um tampão contendo 20 mM Tris, pH 7,5 seguido durante a noite por incubação a 37°C. As cadeias N-glicosiladas foram removidas adicionando 50 mU de Endoglicosidase H (Roche Molecular Biochemicals, Alemanha) a 50 µg de fitase em um tampão contendo 50 mM fosfato de sódio, pH 6,5 e incubados durante a noite a 37°C. Após digestão, 1 µg da proteína foi analisada em um Gel de Tris-glicina a 12% (Invitrogen, San Diego, CA). Os resultados estão resumidos na Figura 7 no formato de tabela.

As proteínas foram depois analisadas através de análise espectral de massa para mapeamento de peptídeo máximo (Figura 8A) e mapeamento de glicosilação (Figura 8B) (dados não-mostrados). Para este experimento, todas as proteínas necessitam ser desnaturadas, reduzidas e alquiladas. Brevemente, volume igual de 8 M uréia (Sigma, MI) foi adicionado à solução de fitase e incubado a 37°C durante 30 min. Para reduzir a proteína, DTT recentemente feito (10 mg/ml) (Sigma, MI) foi adicionado a esta mistura a uma concentração final de 0,04 mg/ml seguido por uma incubação a 37°C durante 30 minutos. Em seguida, 20 mg/ml de Iodoacetamida (Sigma, MI) foram adicionados à mistura de proteína reduzida a uma concentração final de 20 µg/mL e incubados a 37°C durante 30 min para alquilação.

Após a proteína de fitase ser desnaturada, reduzida e alquilada,

a proteína foi depois submetida à diálise em um tampão contendo 34 mM NaCl e 0,08 N HCl. Pepsina (5-20 mg/ml) foi adicionada para digerir a fitase durante a noite a 37°C. A digestão completa da proteína pode ser analisada através de SDS-PAGE.

Fragmentos de fitase digeridos por pepsina foram carregados em uma coluna de Con A (Pharmacia Biotech, Piscataway, NJ) em um tampão contendo 20 mM Tris, pH 7,4, 0,5 M NaCl, 1 mM CaCl2, 1 mM MnCl2, e 1 mM MgCl2. A coluna foi lavada extensivamente com o mesmo tampão. Os peptídeos glicosilados foram eluídos usando tampão de Tris a 20 mM, pH 7,5 contendo 0,5 M D-Metilmanosido.

Para análise espectral de massa de MALDI, dois tipos de matrizes foram usados nestes experimentos para análise de peptídeos ou proteína. Ácido 3,5-dimetoxi-4-hidroxicinâmico (10 mg/ml) dissolvido em 49,9% de água, 50% de metanol e 0,1% de TFA foi usado para análise de proteína. Ácido alfa-ciano-4-hidroxicinâmico (10 mg/ml) dissolvido em 50% de metanol, 49,9% de etanol e 0,1% de TFA foi usado para análise de peptídeo. Para aplicar em uma ponta da sonda de aço, 1 µL de amostra foi bem misturado com 1 µL de solução de matriz. As amostras misturadas com matriz foram secas com ar na sonda e analisadas em um instrumento de STR de Voyager-DE (PE Biosystems, Foster City, CA).

Sítios de glicosilação para fitase de S. cerevisiae e P. pastoris estão apresentados nas Figuras 9A e 9B. Os resultados destes estudos estão resumidos na Figura 10.

EXEMPLO 4

COMPARAÇÕES DE HOSPEDEIRO DE EXPRESSÃO

A construção de DNA de GSSM da SEQ ID Nº: 1 foi inserida em E. coli, P. pastoris e S. pombe para expressão. As proteínas expressadas foram purificadas a homogeneidade. No ensaio de tolerância térmica, 100 uL de 0,01 mg/ml de proteína em 100 mM de MOPS, pH 7,0 foram aquecidos na temperatura de incubação indicada como mostrado na Figura 11 durante 5 minutos em um termociclizado de pesquisa de RJ. Ao concluir em 5 minutos em temperatura, as amostras foram esfriadas para 4°C e incubadas em

gelo. Um ensaio de atividade foi operado usando 40 uL da solução de enzima em 1,46 ml de 100 mM NaOAc / 4 mM fitato / pH 4,5 a 37°C. Alíquotas de 60 uL foram retiradas em intervalos de 2 minutos e adicionadas a 60 uL da solução desenvolvente/interruptora de cor do ensaio de TNO. (Ver Figura 11).

EXEMPLO 5

O percentual das atividades residuais (com base nas taxas iniciais) da fitase recombinante digerida *in vitro* (SEQ ID Nº: 2) expressada em E. coli (não-glicosilada), como também em S. pombe e P. pastoris (glicosilados) foi representado em gráfico versus tempo. Uma concentração padrão de fluido gástrico simulado contendo 2 mg/ml NaCl, 6 M HCl e 3,2 mg/mL de pepsina foi preparada como descrito no S.O.P. O pH da solução foi cerca de 1,4 e não foi ajustado. O ensaio de digestibilidade *in vitro* foi executado adicionando 1:4 (vol:vol) da fitase à solução de digestão e incubando imediatamente a 37°C para iniciar a reação de digestão. Alíquotas da mistura de reação de digestão foram removidas em vários intervalos de tempo e ensaiadas para atividade de fitase residual usando o ensaio de TNO. Cada um dos ensaios foi executado em triplicata. Uma curva exponencial com a equação y = Ae-kt foi ajustada aos dados. As meio-vidas das proteínas foram determinadas usando a equação t ½ = ln 2 / k. A meia-vida da fitase não-glicosilada da SEQ ID Nº: 2 expressada em E. coli foi 8,4 ± 1.1 minuto enquanto a fitase glicosilada expressada em S. pombe tive uma meia-vida de 10,4 ± 0,9 minuto e a mesma fitase expressada em P. pastoris tive uma meia-vida de 29,2 ± 6,7 minutos. Portanto, a glicosilação da fitase da SEQ ID de NO: 2 intensificou a estabilidade da enzima sob condições de digestibilidade simulada *in vitro*. (ver Figura 12).

EXEMPLO 5

Para testar a tolerância térmica da invenção a fitase modificada por GSSM (SEQ ID Nº: 2) quando expressada em E. coli, P. pastoris e S. pombe, as amostras foram aquecidas a 37°C, 50°C, 60°C, 70°C 80°C e 90°C durante 5 minutos e depois submetidas ao ensaio atividade-específico. A fitase de K12 do tipo selvagem (SEQ ID Nº: 3) expressada em E. coli foi

usada como o controle. A faixa de atividade específica a 37°C quando medida em unidades por miligrama de enzima foi medida a pH 4,5 de acordo com o protocolo de TNO acima descrito. A Tabela 3 abaixo resume os resultados da prova de atividade do teste de tolerância térmica/atividade específica.

TABELA 3

TOLERÂNCIA TÉRMICA VERSUS ATIVIDADE ESPECÍFICA DA FITASE DA SEQ ID Nº: 2 EXPRESSA EM E. COLI E LEVEDURA

faixa de temperatura	Faixa de atividade especifica a 37°C (Unidades por miligrama de proteína)			
	Fitase de K12 do tipo selvagem	SEQ ID NO: 2 fitase expressa em		
		E. coli	Pichia Pastoris	Schizosaccharomyces pombe
37°C-50°C	1000-1200	1200	1200	1200
50°C-70°C	0-1000	525-1200	750-1200	1000-1200
70°C-90°C	0	100-500	350-750	610-1000

Os resultados destes testes mostram que as enzimas de fitase glicosilada obtidas através de expressão em P. pastoris e S. pombe exibem tolerância superior à exposição em temperaturas acima de 37°C quando comparadas àquela enzima do tipo selvagem e também tolerância térmica intensificada quando comparadas à fitas modificada por GSSM, mas não-glicosilada. Além disso, expressão da enzima em S. pombe conferiram a maior tolerância térmica, com retenção de pelo menos metade (610 - 100 unidades por miligrama de proteína) da atividade da enzima específica após exposição a temperaturas de 70°C a 90°C. Em contraste, a enzima do tipo selvagem reteve zero de atividade específica nesta faixa de temperatura, e até mesmo a fitase modificada por GSSM expressa em E. coli (não-glicosilada) (SEQ ID Nº: 2) reteve apenas 100-500 unidades de atividade de enzima por miligrama de enzima após exposição a temperatura na faixa de 70°C a 90°C durante 5 minutos. Portanto, a glicosilação da fitase da SEQ ID Nº: 2 também intensificou a tolerância térmica e atividade específica da enzima após exposição a temperatura elevada.

LITERATURA CITADA

(Os ensinamentos de todas as referências citadas neste pedido

de patente foram aqui incorporados por referência na sua totalidade a menos que do contrário indicado).

Association of Official Analytical Chemists: Official Methods of Analysis. Association of Official Analytical Chemists, Washington, DC., 1970.

Ausubel FM, et al. Current Protocols in Molecular Biology. Greene Publishing Assoc., Media, PA. ©1987, ©1989, ©1992.

Barnes WM: PCR amplification of up to 35-kb DNA with high fidelity and high yield from lambda bacteriophage templates. Proceedings of the National Academy of Sciences, USA 91(6):2216-2220, 1994.

Bayer EA, Morag E, Lamed R: The cellulosome--a treasure-trove for biotechnology. Trends Biotechnol 12(9):379-86, (Sep) 1994.

Bevan M: Binary Agrobacterium vectors for plant transformation. Nucleic Acids Research 12(22):8711-21, 1984.

Bird et al. Plant Mol Biol 11:651, 1988.

Blobel G, Walter P, Chang CN, Goldman BM, Erickson AH, Lingappa VR: Translocation of proteins across membranes: the signal hypothesis and beyond. Symp Soc Exp Biol 33:9-36, 1979.

Brederode FT, Koper-Zawrthoff EC, Bol JF: Complete nucleotide sequence of alfalfa mosaic virus RNA 4. Nucleic Acids Research 8(10):2213-23, 1980.

Clark WG, Register JC 3d, Nejidat A, Eichholtz DA, Sanders PR, Fraley RT, Beachy RN: Tissue-specific expression of the TMV coat protein in transgenic tobacco plants affects the level of coat protein-mediated virus protection. Virology 179(2):640-7, (dezembro) 1990.

Cole, et al.: Monoclonal Antibodies and Cancer Therapy. A.R. Liss, Nova Iorque. ©1985.

Coligan JE, et al.: Current Protocols in Immunology. J. Wiley and Sons, Nova Iorque. ©1996.

Coruzzi G, Broglie R, Edwards C, Chua NH: Tissue-specific and light-regulated expression of a pea nuclear gene encoding the small subunit of ribulose-1,5-bisphosphate carboxylase. EMBO J 3(8):1671-9, 1984.

Cosgrove DJ: Inositol phosphate phosphatases of microbiologi-

cal origin. Inositol phosphate intermediates in the dephosphorylation of the hexaphosphates of myo-inositol, scyllo-inositol, and D-chino-inositol by a bacterial (Pseudomonas sp.) phytase. Aust J Biol Sci 23(6):1207-1220, 1970.

Dassa E, Cahu M, Desjoyaux-Cherel B, Boquet PL: The acid phosphatase with optimum pH of 2.5 of *Escherichia coli*. Physiological and Biochemical study. J Biol Chem 257(12):6669-76, (25 de junho) 1982.

Davis LG, et al. Basic Methods in Molecular Biology. Elsevier, Nova lorque, ©1986.

Duarte JC, Costa-Ferreira M: Aspergilli and lignocellulosics: enzymology and biotechnological applications. FEMS Microbiol Rev 13(2-3):377-86, (março) 1994.

Food Chemicals Codex, 4ª Edição. Committee on Food Chemicals Codex, Food and Nutrition Board, Institute of Medicine, National Academy of Sciences. Published: National Academy Press, Washington, DC, ©1996.

Garcia PD, Ghrayeb J, Inouye M, Walter P: Wild type and mutant signal peptides of *Escherichia coli* outer membrane lipoprotein interact with equal efficiency with mammalian signal recognition particle. J Biol Chem 262(20):9463-8, (15 de julho) 1987.

Gluzman Y: SV40-transformed simian cells support the replication of early SV40 mutants. Cell 23(1):175-182, 1981.

Goeddel DV, Shepard HM, Yelverton E, Leung D, Crea R, Sloma A, Pestka S: Synthesis of human fibroblast interferon by E. coli. Nucleic Acids Research 8(18):4057-4074, 1980.

Gordon-Kanun WJ, Spencer TM, Mangano ML, Adams TR, Daines RJ, Start WG, O'Brien JV, Chambers SA, Adams Jr. WR, Willets NG, Rice TB, Mackey CJ, Krueger RW, Kausch AP, Lemaux PG. Plant Cell 2:603, 1990.

Graf E: Phytic Acid: Chemistry and Applications. Pilatus Press, Minneapolis. 1986.

Greiner R, Haller E, Konietzny U, Jany KD: Purification and characterization of a phytase from Klebsiella terrigena. Arch Biochem Biophys

341(2):201-6, (15 de maio) 1997.

Greiner R, Konietzny U: Construction of a bioreactor to produce special breakdown products of phytate. J Biotechnol 48(1-2):153-9, (18 de julho) 1996.

Greiner R, Konietzny U, Jany KD: Purification and characterization of two phytases from *Escherichia coli*. Arch Biochem Biophys 303(1):107-13, (15 de maio) 1993.

Guilley H, Dudley RK, Jonard G, Balazs E, Richards KE: Transcription of Cauliflower mosaic virus DNA: detection of promoter sequences, and characterization of transcripts. Cell 30(3):763-73, 1982.

Hespell RB, Whitehead TR: Physiology and genetics of xylan degradation by gastrointestinal tract bacteria. J Dairy Sci 73(10):3013-22, (outubro) 1990.

Hoekema A, Hirsch PR, Hooykaas PJJ, Schilperoort RA. Nature 303:179, 1983.

Horsch RB, Fry JE, Hoffmann NL, Eichholtz D, Rogers SG, Fraley RT. Science 227:1229, 1985.

Igarashi M, Hollander VP: Acid phosphatase from rat liver. Purification, crystallization, and properties. J Biol Chem 243(23):6084-9, (Dec. 10) 1968.

International Union of Biochemistry and Molecular Biology, Nomenclature Committee: Enzyme nomenclature 1992: recommendations of the Nomenclature Committee of the International Union of Biochemistry and Molecular Biology on the nomenclature and classification of enzymes / prepared for NC-IUBMB por Edwin C. Webb. Academic Press, c1992.

Jeffries TW: Biochemistry and genetics of microbial xylanases. Curr Opin Biotechnol 7(3):337-42, (junho) 1996.

Jermutus L, Tessier M, Pasamontes L, van Loon APGM, Lehmann M: Structure based chimeric enzymes as an alternative to directed enzyme evolution: phytase as a test case. J. Biotechnology 85:15-24, 2001.

Klee HJ, Muskopf YM, Gasser CS: Cloning of an Arabidopsis thaliana gene encoding 5-enolpyruvylshikirnate-3-phosphate synthase: se-

quence analysis and manipulation to obtain glyphosate-tolerant plants. Mol Gen Genet 210(3):437-42, (dezembro) 1987.

Kohler G, Milstein C: Continuous cultures of fused cells secreting antibody of predefined specificity. Nature 256(5517):495-497,1975.

Koster-Topfer M, Frommer WB, Rocha-Sosa M, Rosahl S, Schell J, Willmitzer L: A class II patatin promoter is under developmental control in both transgenic potato and tobacco plants. Mol Gen Genet 219(3):390-6, (novembro) 1989.

Kozbor. Immunology Today 4:72, 1983.

Lee B, et al.: Transient gene expression in aleurone protoplasts isolated from developing caryopses of barley and wheat. Plant Mol Biol 13(1):21-9, 1989.

Lehmann M, Lopez-Ulibarri R, Loch C, Viarouge C, Wyss M and van Loon, APGM. Protein Science 9:1866-1872, 2000.

National Research Council: Nutrient Requirements of Poultry (9ª Edição Revisada). National Academy Press, Washington, DC, 1994.

Nayini NR, et al.: Lebensmittel Wissenschaft und Technologie 17:24-26, 1984.

NCBI, National Library of Medicine. National Institutes of Health: BLAST Sequence Similarity Searching (website = endereço do website ncbi.nlm.nih.gov).

Nelson TS, Shieh TR, Wodzinski RJ, Wwere JH: Effect of supplemental phytase on the utilization of phytate phosphorus by chicks. J Nutr 101(10):1289-1293, 1971.

Ng DT, Walter P: Protein translocation across the endoplasrnic reticulum. Curr Opin Cell Biol 6(4):510-6, (agosto), 1994.

Potrykus I: Gene transfer methods for plants and cell cultures. Ciba Found Syrup 154:198-208; debate 208-12, 1990.

Powar VK, et al.: Purification and properties of phytate-specific phosphatase from *Bacillus subtilis*. J Bacteriol 151(3):1102-1108, 1982.

Powers T, et al.: The nascent polypeptide-associated complex modulates interactions between the signal recognition particle and the ribo-

some. Curr Biol 6(3):331-8, (1º de março), 1996.

Prade RA: Xylanases: from biology to biotechnology. Biotechnol Genet Eng Rev;13:101-31, 1996.

Ryan AJ, et al.: Genomic sequence of a 12S seed storage protein from oilseed rape (Brassica napus c.v. jet neuf). Nucl Acids Res 17(9):3584, 1989.

Saiki RK, Gelfand DH, Stoffel S, Scharf SJ, Higuchi R, Horn GT, Mullis KB, Erlich HA: Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science 239(4839):487-491, 1988.

Sambrook J, Fritsch EF, Maniatis T: Molecular Cloning: A Laboratory Manual, Cold Spring Habor Press, Cold Spring Habor, NY, ©1989.

SAS: Statistics In: SAS User's Guide (1984 ed.). SAS Institute, Cwerey, NC, 1984.

Schoner FJ, Hope PP, Schwarz G, Wiesche H: Comparative effects of microbial phytase and inorganic phosphorus on performance and retention of phosphorus, calcium, and crude ash in broilers. J Anim Physiol Anim Nutr 66:248, 1991.

Schoner FJ, Hope PP, Schwarz G, Wiesche H: Effects of microbial phytase and inorganic phosphate in broiler chicken: Performance and mineral retention at various calcium levels. J Anim Physiol Anim Nutr 69:235,1993.

Shieh TR, et al.: Regulation of the formation of acid phosphatases by inorganic phosphate in Aspergillus ficuum. J Bacteriol 100(3):1161-5, (dezembro) 1969.

Shimamoto K, Miyazaki C, Hashimoto H, Izawa T, Itoh K, Terada R, Inagaki Y, Iida S: Trans-activation and stable integration of the maize transposable element Ds cotransfected with the Ac transposase gene in transgenic rice plants. Mol Gen Genet 239(3):354-60, (junho) 1993.

Shimizu M: Bioscience, Biotechnology, and Biochemistry 56:1266-1269, 1992.

Sijmons PC, Dekker BM, Schrammeijer B, Verwoerd TC, van den Elzen PJ, Hoekema A: Production of correctly processed human serum

albumin in transgenic plants. Biotechnology (NY) 8(3):217-21, 1990.

Simons PC, Versteegh HA, Jongbloed AW, Kemme PA, Slump P, Bos KD, Wolters MG, Beudeker RF, Verschoor GJ: Improvement of phosphorus availability by microbial phytase in broilers and pigs. Br J Nutr 64(2):525-540, 1990.

Smeekens S, Weisbeek P, Robinson C: Protein transport into and within chloroplasts. Trends Biochem Sci 15(2):73-6, 1990.

Smith AG, et al.: Identification and characterization of stamen- and tapetum-specific genes from tomato. Mol Gen Genet 222(1):9-16, (junho) 1990.

Tague BW, et al.: A short domain of the plant vacuolar protein phytohemagglutinin targets invertase to the yeast vacuole. Plant Cell 2(6):533-46, (junho) 1990.

Tingey SV, et al.: Glutamine synthetase genes of pea encode distinct polypeptides which were differentially expressed in leaves, roots and nodules. EMBO J 6(1):1-9, 1987.

Tomschy A, et al.. Optimization of the catalytic properties of Aspergillus Fuinigatus phytase based on the three-dimensional structure. Protein Science 9:1304-1311, 2000.

Ullah AH: Production, rapid purification and catalytic characterization of extracellular phytase from Aspergillus ficuum. Prep Biochem 18(4):443-458, 1988.

Ullah AH, Gibson DM: Extracellular phytase (E.C. 3.1.3.8) from Aspergillus ficuum NRRL 3135: purification and characterization. Prep Biochem 17(1):63-91, 1987.

Van den Broeck G, et al.: Targeting of a foreign protein to chloroplasts by fusion to the transit peptide from the small subunit of ribulose 1,5-bisphosphate carboxylase. Nature 313(6001):358-63, 1985.

Vasil IK, Vasil V: Totipotency and embryogenesis in plant cell and tissue cultures. *In vitro* 8(3):117-27, (novembro-dezembro) 1972.

Vasil V, Vasil IK: Regeneration of tobacco and petunia plants from protoplasts and culture of corn protoplasts. *In vitro* 10:83-96, (julho-

agosto) 1974.

Von Heijne G: Towards a comparative anatomy of N-terminal topogenic protein sequences. J Mol Biol 189(I):239-42, 1986.

Walter P, Blobel G. Biochem Soc Symp 47:183, 1986.

Wenzler H, et al.: Sucrose-regulated expression of a chimeric potato tuber gene in leaves of transgenic tobacco plants. Plant Mol Biol 13(4):347-54, 1989.

Wolter FP, et al. rbcS genes in Solanum tuberosum: conservation of transit peptide and exon shuffling during evolution. Proc Natl Acad Sci USA 85(3):846-50, (fevereiro) 1988.

Wong KK, et al.: Multiplicity of beta-1,4-xylanase in microorganisms: functions and applications. Microbiol Rev 52(3):305-17, (setembro) 1988.

Wyss M, et al.. Biochemical Characterization of Fungal Phytases (myo-inositol Hexakisphosphate Phosphohydrolases): Catalytic Properties. Applied and Environmental Microbiology fevereiro 1999: 367-373.

Yamada K, et al.: Agricultural and Biological Chemistry 32:1275-1282, 1968.

Patente U.S. No. 3.297.548; depositada em 28 de julho, 1964; publicada em 10 de janeiro de 1967. Wwere JH, Bluff L, Shieh TK: Preparation of acid phytase.

Patente U.S. No. 4.946.778; depositada em 19 de janeiro de 1989; publicada em 7 de agosto de 1990. Ladner RC, Bird RE, Hardman K: Single polypeptide chain binding molecules.

Patente U.S. No. 5.830.732, depositada em 3 de julho de 1995, publicada em 3 de novembro de 1998. Mochizuki D, Tokuda J, Shimada M e Tawaki S. Phytase.

EP0 120.516; depositada em 21 de fevereito de 1984; publicada em 3 de outubro de 1984. Schilperoort RA, et al.: A process of the incorporation of foreign DNA into the genome of dicotyledonous plants; Agrobacterium tumefaciens bacteria and a process for the production thereof; plants and plant cells with modified genetic properties; a process for the preparation.

EP0 321.004; depositada em 28 de outubro de 1988; publicada em 22 de janeiro de 1992. Vaara T, et al., A process for steeping cereals with a new enzyme preparation.

IPN WO 91/05053; depositado em 27 de setembro de 1990; publicada em 18 de abril de 1991. VanGorcom R, et al.: Cloning and expression of microbial phytase.

Plant Cell Culture Protocols (Methods in Molecular Biology (Cloth), 111) por Robert D. Hall (Editor) (março de 1999) Humana Press; ISBN: 0896035492

Plant Molecular Biology: Essential Techniques por P. Jones (Editor), J. M. Sutton (Editor), Mark Sutton (Contribuidor) (25 de setembro de 1997) John Wiley & Son Ltd; ISBN: 0471972681

Plant Biochemistry and Molecular Biology por Hans-Walter Heldt (abril de 1998) Oxford University Press; ISBN: 019850179X

Biochemistry and Molecular Biology of Plants por Bob B. Buchanan (Editor), Wilhelm Gruissem (Editor), Russell L. Jones (julho 2000) Amer Society of Plant; ISBN: 0943088372

Monoclonal Antibodies: A Manual of Techniques por Heddy Zola (setembro de 1987) CRC Press; ISBN: 0849364760

Immunochemistry in Practice por Robin Thorpe (Contribuidor), Alan P. Johnstone 3ª ed (15 de janeiro de 1996) Blackwell Science Inc; ISBN: 0865426333

REIVINDICAÇÕES

1. Polipeptídeo recombinante possuindo atividade fitase, caracterizado pelo fato de que

compreende uma sequência de aminoácidos de SEQ ID Nº: 2, ou

é codificado por uma sequência de ácido nucleico de SEQ ID Nº: 1, em que a fitase é glicosilada após ser expressa em *P. pastoris* ou *S. pombe* e possui atividade fitase específica de 350 a 1000 unidades por miligrama de proteína após exposição a temperaturas de 70ºC a 90ºC.

2. Polipeptídeo de acordo com a reivindicação 1, caracterizado pelo fato de que a atividade de fitase compreende a hidrólise de fitato (mio-inositol-hexafosfato).

3. Polipeptídeo de acordo com a reivindicação 1, caracterizado pelo fato de que a fitase compreende pelo menos um sítio de glicosilação.

4. Polipeptídeo de acordo com a reivindicação 3, caracterizado pelo fato de que a glicosilação é uma glicosilação N-ligada.

5. Preparação de proteína, caracterizada pelo fato de que compreende um polipeptídeo, como definido na reivindicação 1, em que a preparação de proteína compreende um líquido, um sólido ou um gel.

6. Heterodímero, caracterizado pelo fato de que compreende um polipeptídeo, como definido na reivindicação 1 e um segundo domínio.

7. Heterodímero de acordo com a reivindicação 6, caracterizado pelo fato de que o segundo domínio é um polipeptídeo e o heterodímero é uma proteína de fusão.

8. Heterodímero de acordo com a reivindicação 7, caracterizado pelo fato de que o segundo domínio é um epítopo.

9. Heterodímero de acordo com a reivindicação 6, caracterizado pelo fato de que o segundo domínio é um marcador.

10. Polipeptídeo imobilizado tendo uma atividade de fitase, caracterizado pelo fato de que o polipeptídeo compreende uma sequência, como definida na reivindicação 1.

11. Polipeptídeo imobilizado de acordo com a reivindicação 10,

caracterizado pelo fato de que a fitase é imobilizada em uma célula, um metal, uma resina, um polímero, uma cerâmica, um vidro, um microeletrodo, uma partícula grafítica, uma conta, um gel, uma placa, um arranjo ou um tubo capilar.

12. Arranjo, caracterizado pelo fato de que compreende um polipeptídeo imobilizado, como definido na reivindicação 1.

13. Suplemento alimentar para um animal, pélete comestível, matriz de liberação de enzima comestível, caracterizado pelo fato de que compreende um polipeptídeo, como definido na reivindicação 1.

14. Matriz de liberação de enzima comestível, caracterizada pelo fato de que compreende o polipeptídeo possuindo atividade fitase, como definido na reivindicação 1.

15. Matriz de liberação de enzima comestível de acordo com a reivindicação 14, caracterizada pelo fato de que compreende um pélete.

16. Pélete comestível, caracterizado pelo fato de que compreende um veículo comestível granulado e um polipeptídeo, como definido na reivindicação 1.

17. Método de produzir um polipeptídeo recombinante, como definido na reivindicação 1, caracterizado pelo fato de que compreende as etapas de:

(a) fornecer um ácido nucleico compreendendo uma sequência que hibrida sob condições rigorosas com uma sequência de ácidos nucleicos como exposta na SEQ ID Nº: 1, em que o ácido nucleico codifica um polipeptídeo tendo uma atividade de fitase, operavelmente ligado a um promotor; e

(b) expressar o ácido nucleico da etapa (a) sob condições que permitem expressão do polipeptídeo, assim produzindo um polipeptídeo recombinante.

18. Método de acordo com a reivindicação 17, caracterizado pelo fato de que compreende ainda transformar uma célula hospedeira com o ácido nucleico da etapa (a) seguido pela expressão do ácido nucleico da etapa (a), produzindo assim um polipeptídeo recombinante em uma célula

transformada.

19. Método para hidrolisar um inositol-hexafosfato para inositol e fosfato inorgânico, caracterizado pelo fato de que compreende as etapas a seguir:

(a) fornecer um polipeptídeo tendo uma atividade de fitase, em que o polipeptídeo compreende uma sequência de aminoácidos, como definida na reivindicação 1;

(b) fornecer uma composição compreendendo um inositol-hexafosfato; e

(c) contatar o polipeptídeo da etapa (a) com a composição da etapa (b) sob condições em que o polipeptídeo hidrolisa com o inositol-hexafosfato para produzir o inositol e fosfato inorgânico.

20. Método para desgrudadura de óleo, caracterizado pelo fato de que compreende as etapas a seguir:

(a) fornecer um polipeptídeo tendo uma atividade de fitase, em que o polipeptídeo compreende uma sequência de aminoácidos, como definida na reivindicação 1;

(b) fornecer uma composição compreendendo um óleo vegetal; e

(c) contatar o polipeptídeo da etapa (a) e o óleo vegetal da etapa (b) sob condições em que o polipeptídeo pode clivar uma ligação de inositol-fosfato inorgânico, assim desgrudando o óleo.

21. Método para processamento de grãos de milho e de sorgo, caracterizado pelo fato de que compreende as etapas a seguir:

(a) fornecer um polipeptídeo tendo uma atividade de fitase, em que o polipeptídeo compreende uma sequência de aminoácidos, como definida na reivindicação 1;

(b) fornecer uma composição compreendendo um líquido de maceração de milho ou um líquido de maceração de sorgo; e

(c) contatar o polipeptídeo da etapa (a) e a composição da etapa (b) sob condições em que o polipeptídeo pode clivar uma ligação de inositol-fosfato inorgânico.

FIG. 1A

Seqüência de DNA (SEQ ID NO:1)

ATGAAAGCGATCTTAATCCCATTTTTATCTCTTCTGATTCCGTTAACCCCGCAATCTGCATTCGCTCAGAG
TGAGCCGGAGCTGAAGCTGGAAAGTGTGGTGATTGTCAGTCGTCATGGTGTGCGTGCTCCAACCAAGGC
CACGCAACTGATGCAGGATGTCACCCCAGACGCATGGCCAACCTGGCCGGTAAAACTGGGTGAGCTGACA
CCGCGCGGTGGTGAGCTAATCGCCTATCTCGGACATTACTGGCGTCAGCGTCTGGTAGCCGACGGATTG
CTGCCTAAATGTGGCTGCCCGCAGTCTGGTCAGGTCGCGATTATTGCTGATGTCGACGAGCGTACCCGTA
AAACAGGCGAAGCCTTCGCCGCCGGGCTGGCACCTGACTGTGCAATAACCGTACATACCCAGGCAGATAC
GTCCAGTCCCGATCCGTTATTTAATCCTCTAAAAACTGGCGTTTGCCAACTGGATAACGCGAACGTGACT
GACGCGATCCTCGAGAGGGCAGGAGGGTCAATTGCTGACTTTACCGGGCATTATCAAACGGCGTTTCGC
GAACTGGAACGGGTGCTTAATTTTCCGCAATCAAACTTGTGCCTTAAACGTGAGAAACAGGACGAAAGCT
GTTCATTAACGCAGGCATTACCATCGGAACTCAAGGTGAGCGCCGACTGTGTCTCATTAACCGGTGCGGT
AAGCCTCGCATCAATGCTGACGGAGATATTTCTCCTGCAACAAGCACAGGGAATGCCGGAGCCGGGGTG
GGGAAGGATCACCGATTCACACCAGTGGAACACCTTGCTAAGTTTGCATAACGCGCAATTTGATTTGCTA
CAACGCACGCCAGAGGTTGCCCGCAGCCGCGCCACCCCGTTATTAGATTTGATCAAGACAGCGTTGACGC
CCCATCCACCGCAAAAACAGGCGTATGGTGTGACATTACCCACTTCAGTGCTGTTTATCGCCGGACACGA
TACTAATCTGGCAAATCTCGGCGGCGCACTGGAGCTCAACTGGACGCTTCCCGGTCAGCCGGATAACACG
CCGCCAGGTGGTGAACTGGTGTTTGAACGCTGGCGTCGGCTAAGCGATAACAGCCAGTGGATTCAGGTT
TCGCTGGTCTTCCAGACTTTACAGCAGATGCGTGATAAAACGCCGCTGTCATTAAATACGCCGCCCGGAG
AGGTGAAACTGACCCTGGCAGGATGTGAAGAGCGAAATGCGCAGGGCATGTGTTCGTTGGCAGGTTTTA
CGCAAATCGTGAATGAAGCACGCATACCGGCGTGCAGTTTGAGATCTCATCTA

FIG. 1B

Seqüência de proteína (SEQ ID NO:1)

MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDV
TPDAWPTWPVKLGELTPRGGELIAYLGHYWRQRLVADGLLPKCGCPQSGQV
AIIADVDERTRKTGEAFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLD
NANVTDAILERAGGSIADFTGHYQTAFRELERVLNFPQSNLCLKREKQDESCS
LTQALPSELKVSADCVSLTGAVSLASMLTEIFLLQQAQGMPEPGWGRITDSHQ
WNTLLSLHNAQFDLLQRTPEVARSRATPLLDLIKTALTPHPPQKQAYGVTLPT
SVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERWRRLSDNSQW
IQVSLVFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQIV
NEARIPACSLRSHL

FIG. 2A

Perfil de pH/temperatura e estabilidafde

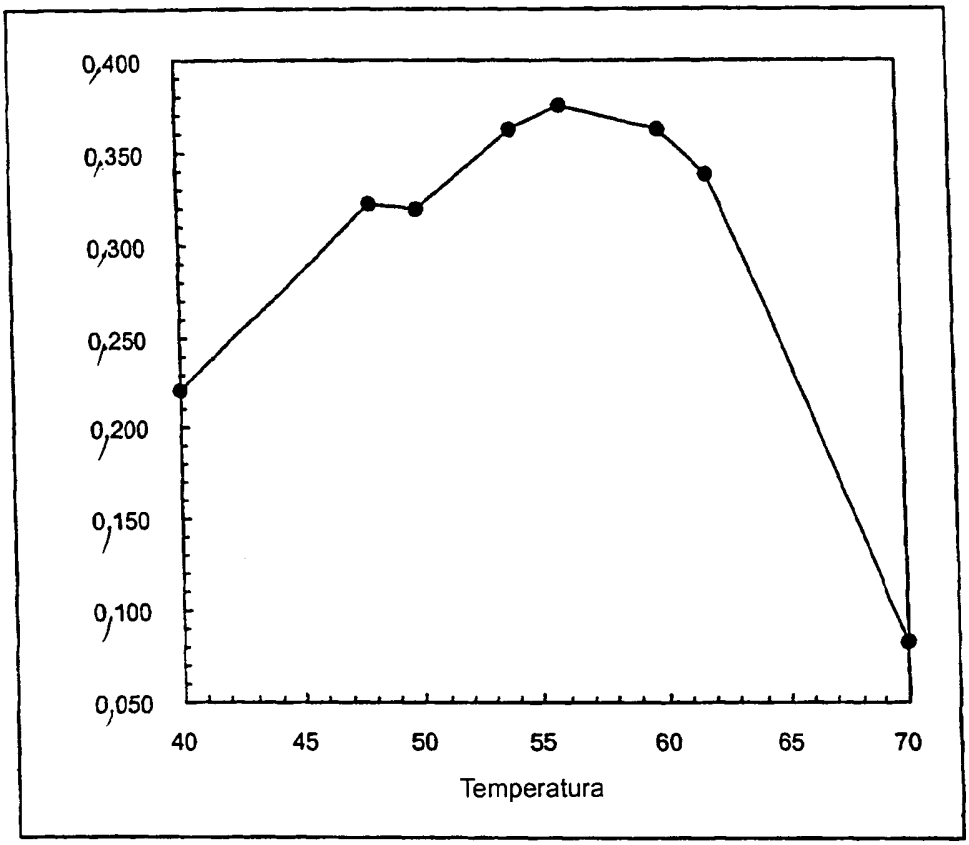

FIG. 2B

Perfil de pH/temperatura e estabilidafde

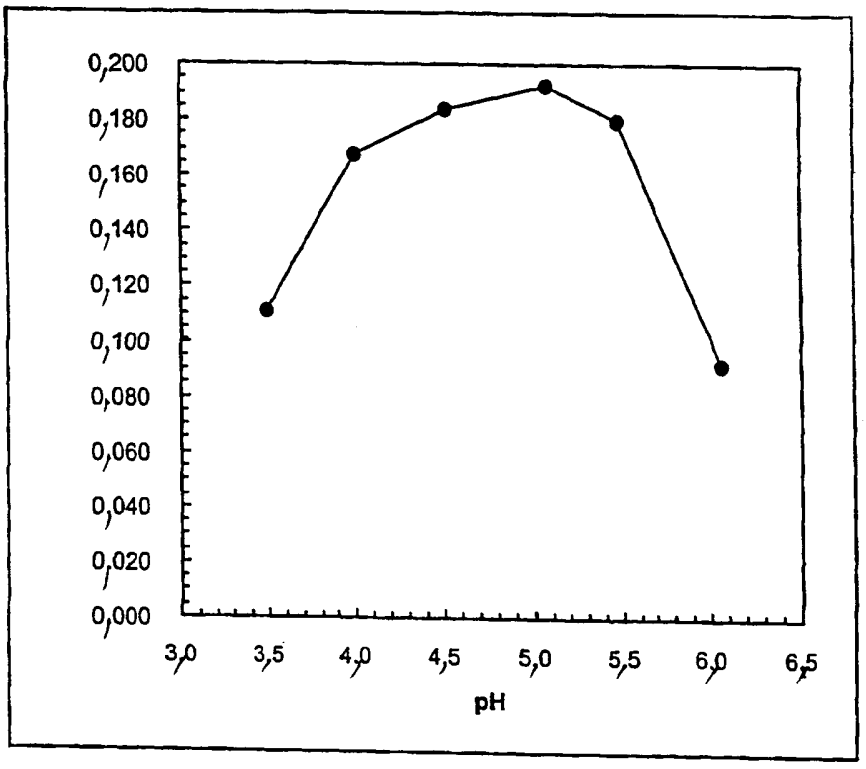

FIG. 3

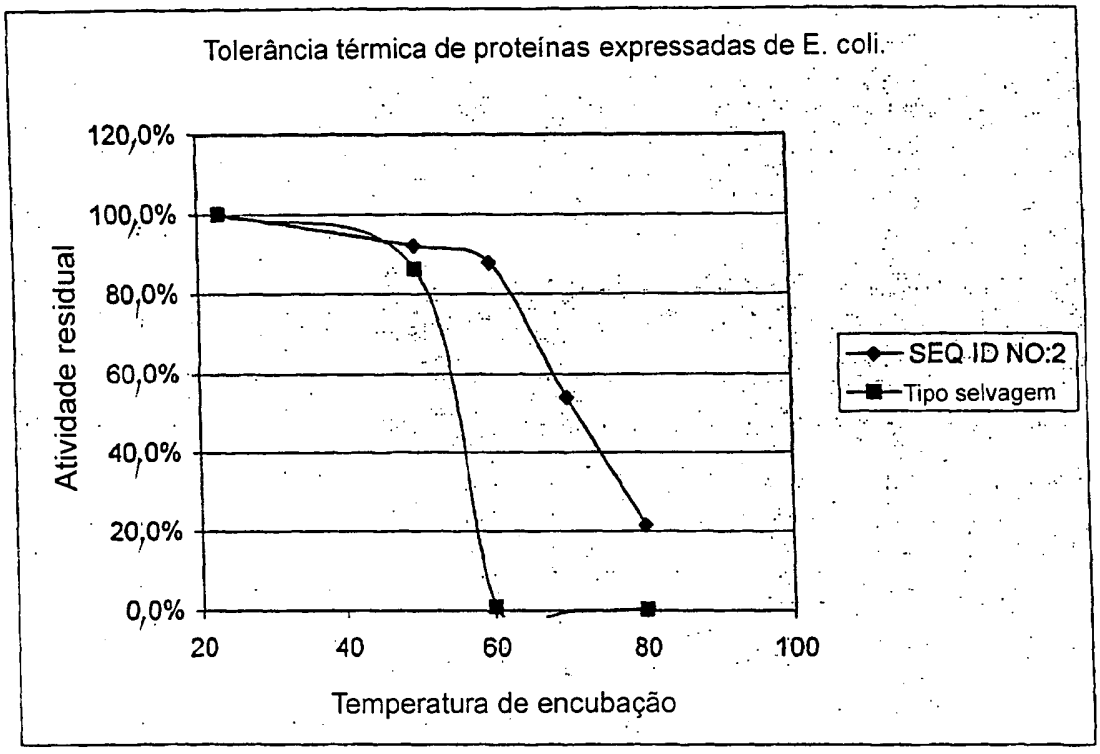
Tolerância térmica de proteínas expressadas de E. coli.
120,0%
100,0%
80,0%
60,0%
40,0%
20,0%
0,0%
Atividade residual
20
40
60
80
100
Temperatura de encubação
SEQ ID NO:2
Tipo selvagem

FIG. 4

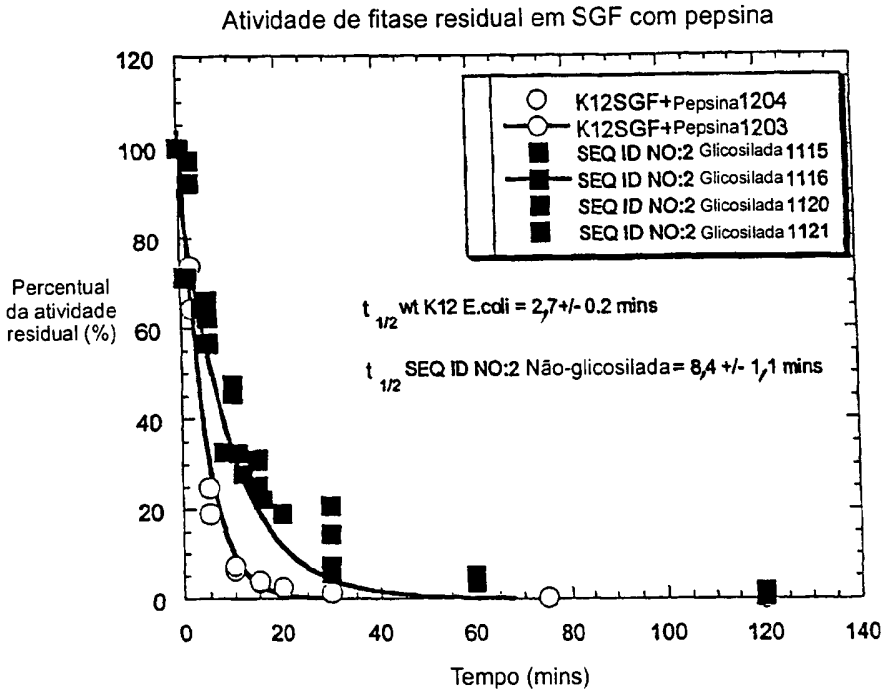
Atividade de fitase residual em SGF com pepsina
Percentual da atividade residual (%)
120
100
80
60
40
20
0
0
20
40
60
80
100
120
140
Tempo (mins)
K12SGF+Pepsina1204
K12SGF+Pepsina1203
SEQ ID NO:2 Glicosilada 1115
SEQ ID NO:2 Glicosilada 1116
SEQ ID NO:2 Glicosilada 1120
SEQ ID NO:2 Glicosilada 1121
t 1/2 wt K12 E.coli = 2,7+/- 0.2 mins
t 1/2 SEQ ID NO:2 Não-glicosilada= 8,4 +/- 1,1 mins

FIG. 5

- T ½ de fitase na digestão de pepsina

Hospedeiro	**T ½ (min)**
E. Coli	~8
Pichia	~10
S. Cerevisiae	~25

FIG. 6

Sétios de glicosilação prognosticados

QSEPELKLES VVIVSRHGVR APTKATQLMQ DV**T**PDAWP**T**W PVKLGELTPR GGELIAYLGH

YWRQRLVADG LLPKCGCPQS GQVAIIADVD ERTRKTGEAF AAGLAPDCAI TVH**T**QAD**T**S**S**

PDPLFNPLKT GVCQLDNA**N**V TDAILERAGG SIADFTGHYQ TAFRELERVL NFPQSNLCLK

REKQDESCSL TQALPSELKV SADCVSLTGA VSLASMLTEI FLLQQAQGMP EPGWGRITDS

HQWNTLLSLH NAQFDLLQRT PEVARSRATP LLDLIK**T**ALT PHPPQKQAYG VTLPTSVLFI

AGHDTNLANL GGALEL**N**WTL PGQPDNTPPG GELVFERWRR LSDNSQWIQV SLVFQTLQQM

RDKTPLSLNT PPGEVKLTLA GCEERNAQGM CSLAGFTQIV NEARIPACSL

FIG. 7

Conclusão de SDS-page

	Hospedeiro	S. Cerevisiae
0-glicosilação	Não	Não
N-glicosilação	Sim	Sim
Forma N-glicosilação	1 domina	2

FIG. 8A

Mapeamento de peptídeo máximo

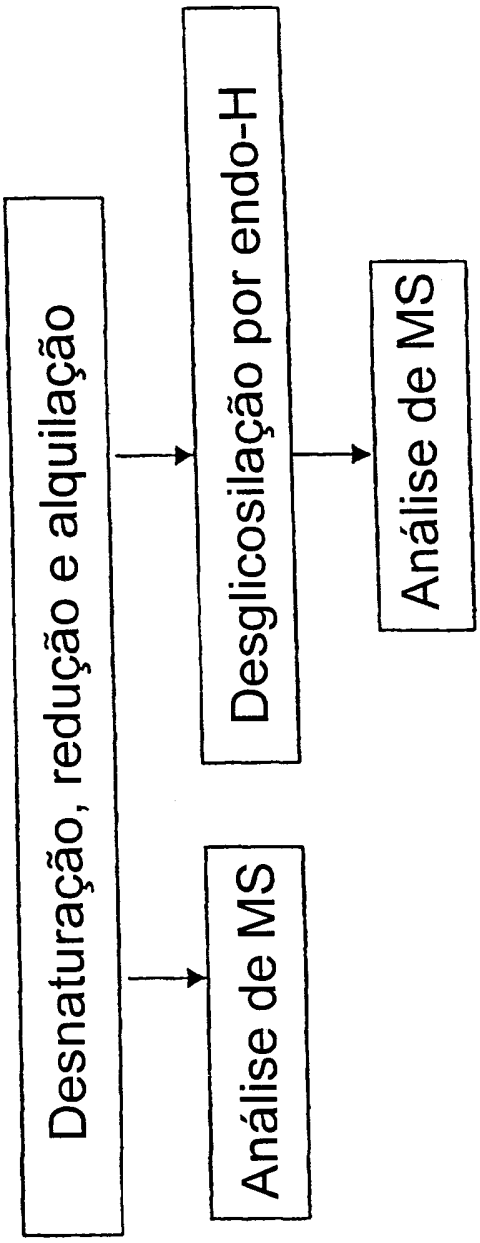

FIG. 8B

Mapeamento de glicosilação

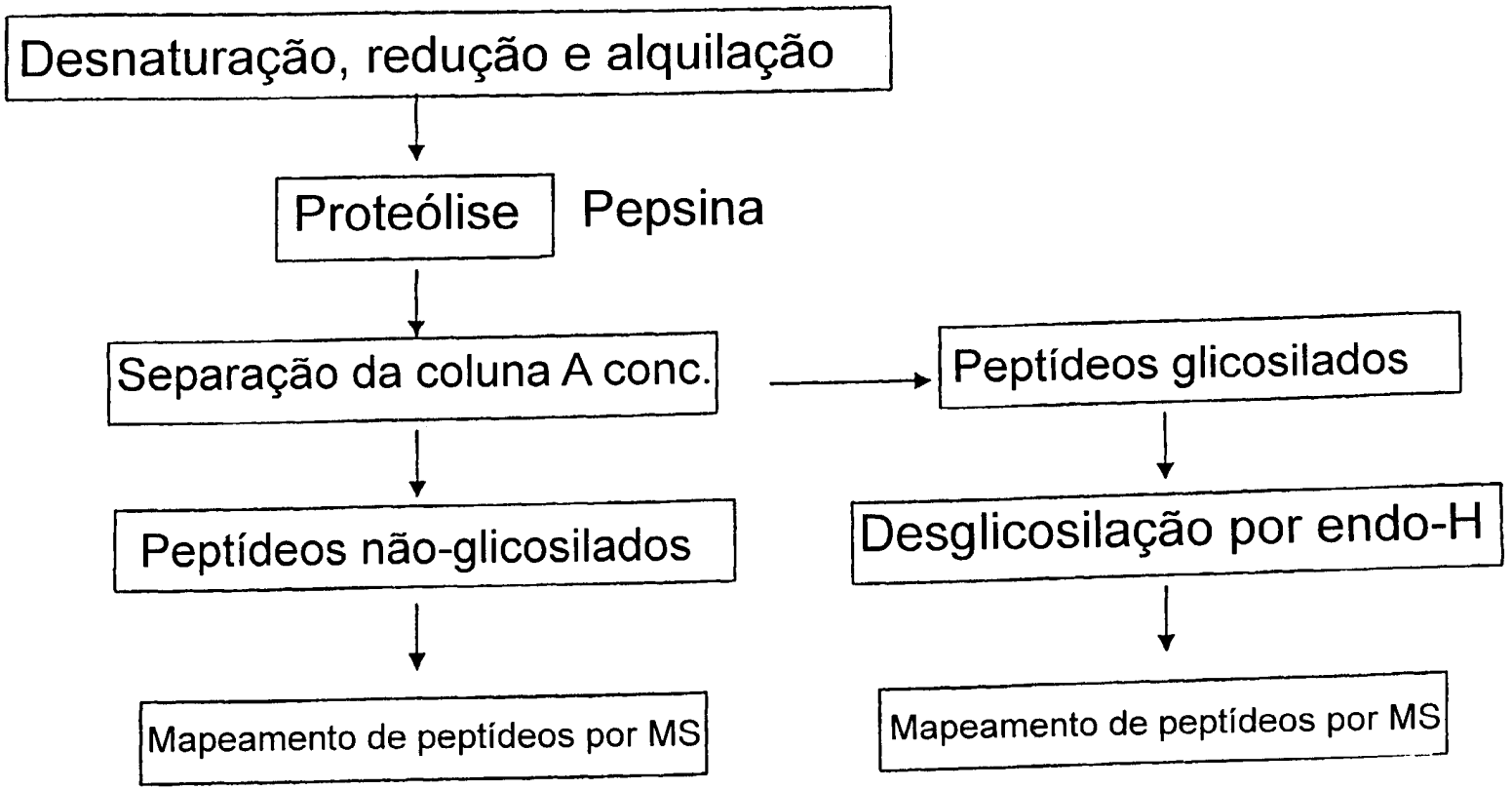

FIG. 9A

Sítios de glicosilação identificados para fitase de pichia

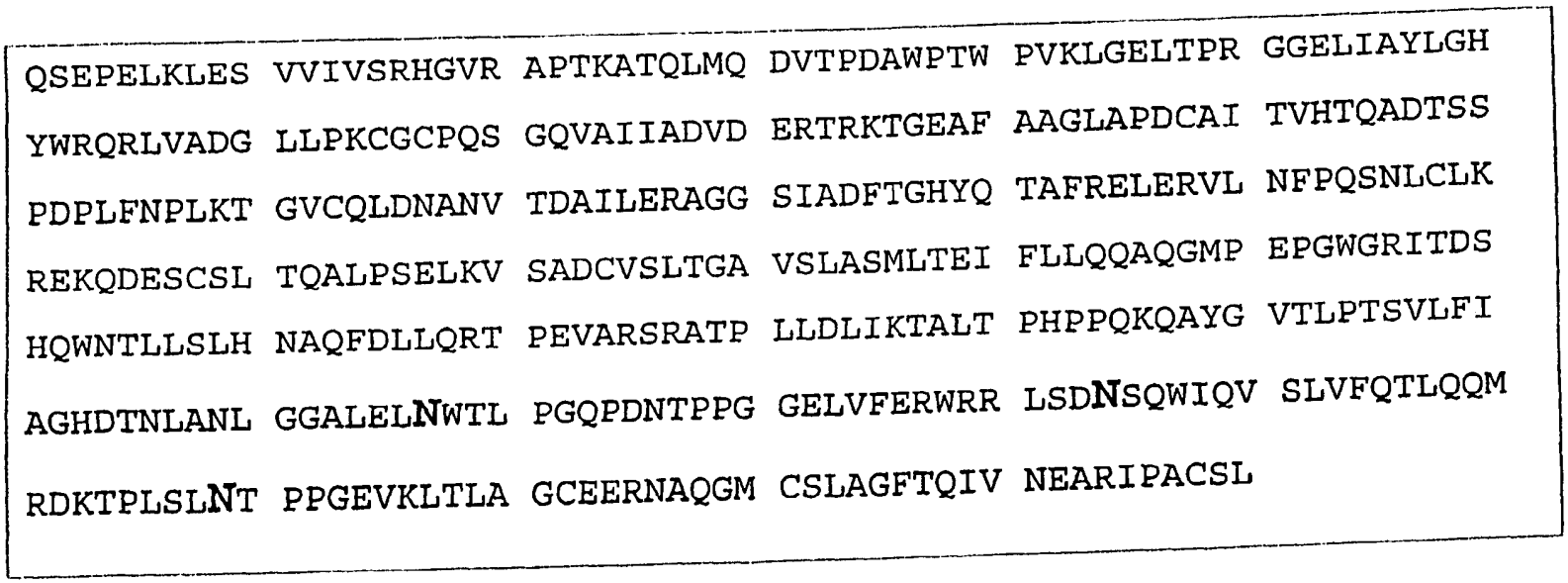

FIG. 9B

Sítios de glicosilação identificados para fitase de S. Cerevisiae

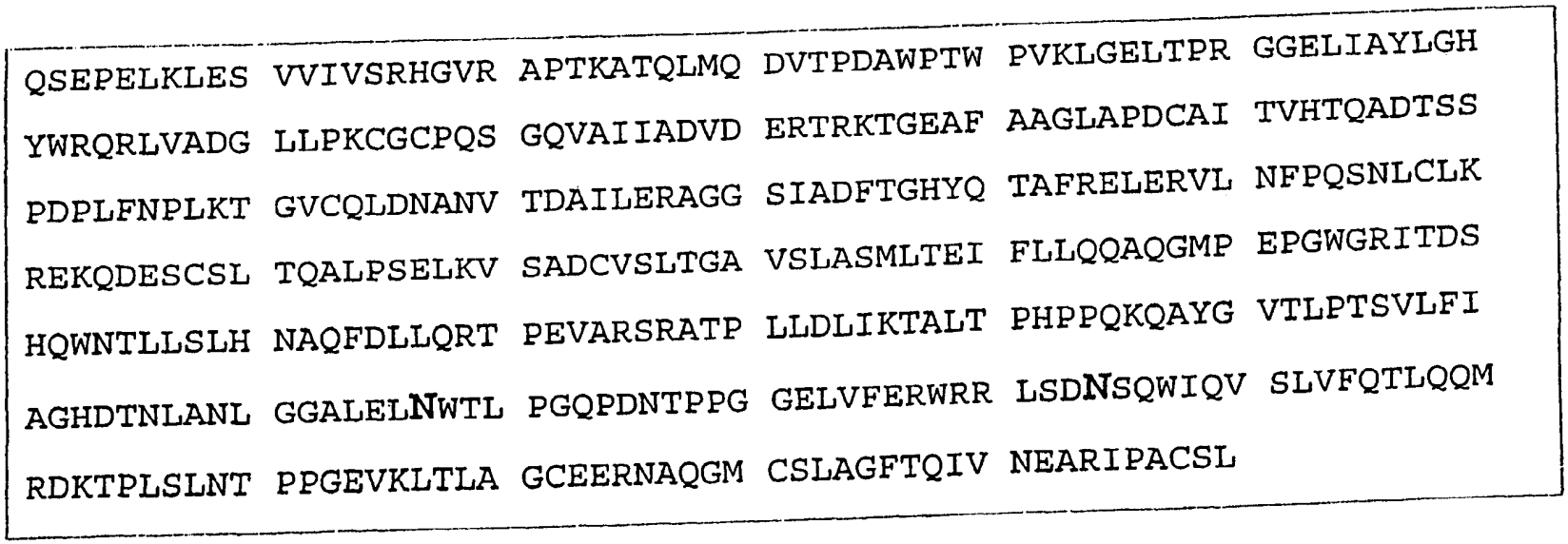

FIG.10

Resumo

- Fitase em pichia
  - Existem três formas de fitase expressadas em pichia. Cada uma delas contém uma glico-cadeia simples no aminoácido #317 e #344.
- Fitase em S. Cerevisiae
  - Existem três formas de fitase expressadas em S.cerevisiae A primeira forma duas glico-cadeiaa simples que estta ligada no aminoácido #317 e #344.

FIG. 11

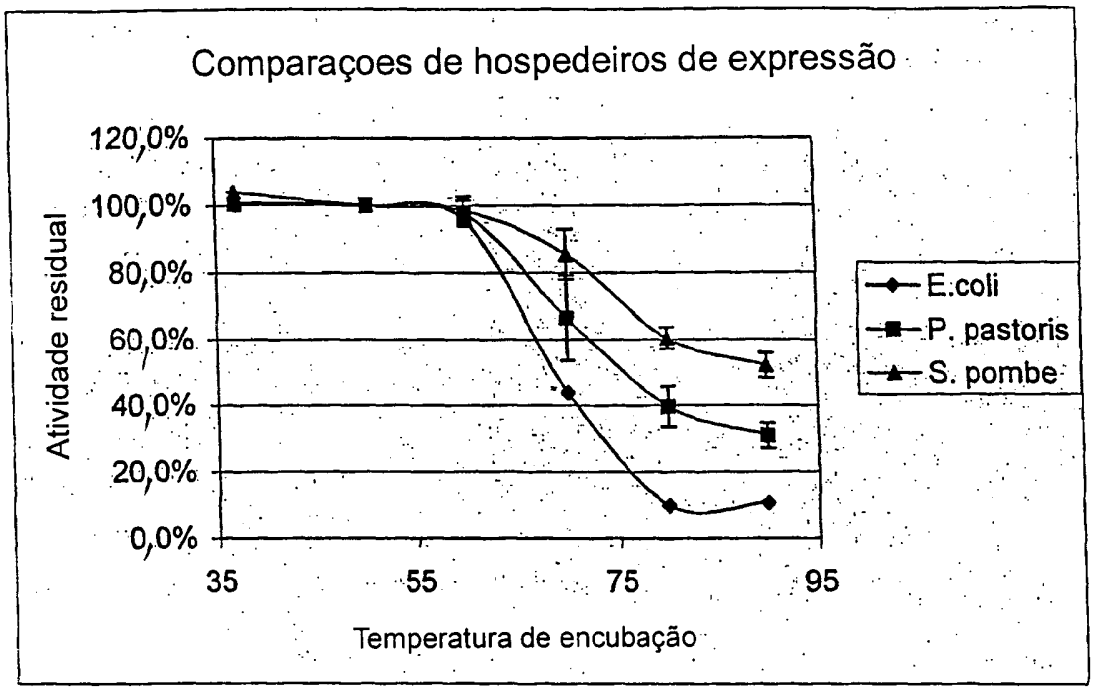
Comparaçoes de hospedeiros de expressão
Atividade residual
120,0%
100,0%
80,0%
60,0%
40,0%
20,0%
0,0%
35
55
75
95
Temperatura de encubação
E.coli
P. pastoris
S. pombe

FIG. 12

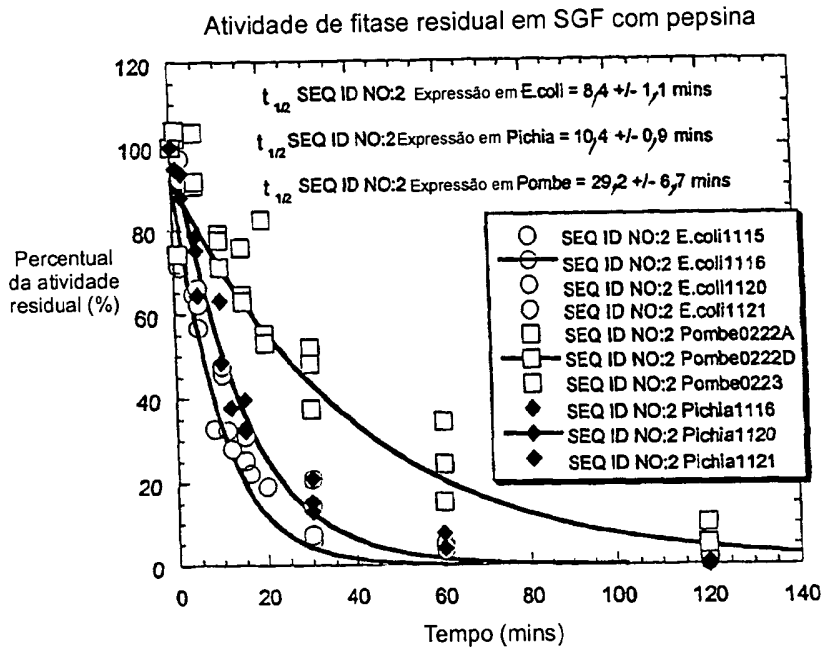
Atividade de fitase residual em SGF com pepsina
$t_{1/2}$ SEQ ID NO:2 Expressão em E.coli = 8,4 +/- 1,1 mins
$t_{1/2}$ SEQ ID NO:2 Expressão em Pichia = 10,4 +/- 0,9 mins
$t_{1/2}$ SEQ ID NO:2 Expressão em Pombe = 29,2 +/- 6,7 mins
SEQ ID NO:2 E.coli1115
SEQ ID NO:2 E.coli1116
SEQ ID NO:2 E.coli1120
SEQ ID NO:2 E.coli1121
SEQ ID NO:2 Pombe0222A
SEQ ID NO:2 Pombe0222D
SEQ ID NO:2 Pombe0223
SEQ ID NO:2 Pichia1116
SEQ ID NO:2 Pichia1120
SEQ ID NO:2 Pichia1121
Percentual da atividade residual (%)
120
100
80
60
40
20
0
0
20
40
60
80
100
120
140
Tempo (mins)

FIG. 13

Appa de E. coli (número de acesso do GenBank) (SEQ ID NO:3)

1 taaggagcag aaacaatgtg gtatttactt tggttcgtcg gcattttgtt gatgtgttcg
61 ctctccaccc ttgtgttggt atggctggac ccgcgtctga aaagttaacg aacgtaggcc
121 tgatgcggcg cattagcatc gcatcaggca atcaataatg tcagatatga aaagcggaaa
181 catatcgatg aaagcgatct taatcccatt tttatctctt ctgattccgt taaccccgca
241 atctgcattc gctcagagtg agccggagct gaagctggaa agtgtggtga ttgtcagtcg
301 tcatggtgtg cgtgctccaa ccaaggccac gcaactgatg caggatgtca ccccagacgc
361 atggccaacc tggccggtaa aactgggttg gctgacaccg cgnggtggtg agctaatcgc
421 ctatctcgga cattaccaac gccagcgtct ggtagccgac ggattgctgg cgaaaaaggg
481 ctgcccgcag tctggtcagg tcgcgattat tgctgatgtc gacgagcgta cccgtaaaac
541 aggcgaagcc ttcgccgccg ggctggcacc tgactgtgca ataaccgtac atacccaggc
601 agatacgtcc agtcccgatc cgttatttaa tcctctaaaa actggcgttt gccaactgga
661 taacgcgaac gtgactgacg cgatcctcag cagggcagga gggtcaattg ctgactttac
721 cgggcatcgg caaacggcgt ttcgcgaact ggaacgggtg cttaattttc cgcaatcaaa
781 cttgtgcctt aaacgtgaga aacaggacga aagctgttca ttaacgcagg cattaccatc
841 ggaactcaag gtgagcgccg acaatgtctc attaaccggt gcggtaagcc tcgcatcaat
901 gctgacggag atatttctcc tgcaacaagc acagggaatg ccggagccgg ggtggggaag
961 gatcaccgat tcacaccagt ggaacacctt gctaagtttg cataacgcgc aattttattt
1021 gctacaacgc acgccagagg ttgcccgcag ccgcgccacc ccgttattag atttgatcaa
1081 gacagcgttg acgccccatc caccgcaaaa acaggcgtat ggtgtgacat tacccacttc
1141 agtgctgttt atcgccggac acgatactaa tctggcaaat ctcggcggcg cactggagct
1201 caactggacg cttcccggtc agccggataa cacgccgcca ggtggtgaac tggtgtttga
1261 acgctggcgt cggctaagcg ataacagcca gtggattcag gtttcgctgg tcttccagac
1321 tttacagcag atgcgtgata aaacgccgct gtcattaaat acgccgcccg gagaggtgaa
1381 actgaccctg gcaggatgtg aagagcgaaa tgcgcagggc atgtgttcgt tggcaggttt
1441 tacgcaaatc gtgaatgaag cacgcatacc ggcgtgcagt ttgtaatgca taaaaaagag
1501 cattcagtta cctgaatgct ctgaggctga tgacaaacga agaactgtct aatgcgtaga
1561 ccggaaaagg cgttcacgcc gcatccggcc actttcagtt ttcctctttc tcggagtaac
1621 tataaccgta atagttatag ccgtaactgt aagcggtgct ggcgcgttta atcacaccat
1681 tgaggatagc gcctttaata ttgacgcctg cctgttccag acgctgcatt gacaaactca
1741 cctctttggc ggtgttcaag ccaaaacgcg caaccagcag gctggtgcca acagaacgcc
1801 ccacgaccgc ggcatcactc accgccagca tcggcggcgt atcgacaatc accagatcgt
1861 aatggtcgtt cgcccattcc agtaattgac gcatccgatc g

FIG. 14

Seqüência de aminóácido para fitase de appA de E. coli (tipo selvagem) **(SEQ ID NO:4)**

MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVT
PDAWPTWPVKLGWLTPRGGELIAYLGHYQRQRLVADGLLAKKGCPQSGQVA
IIADVDERTRKTGEAFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNA
NVTDAILSRAGGSIADFTGHRQTAFRELERVLNFPQSNLCLKREKQDESCSLTQ
ALPSELKVSADNVSLTGAVSLASMLTEIFLLQQAQGMPEPGWGRITDSHQWNT
LLSLHNAQFYLLQRTPEVARSRATPLLDLIKTALTPHPPQKQAYGVTLPTSVLFI
AGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERWRRLSDNSQWIQVSL
VFQTLQQMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQIVNEARI
PACSL

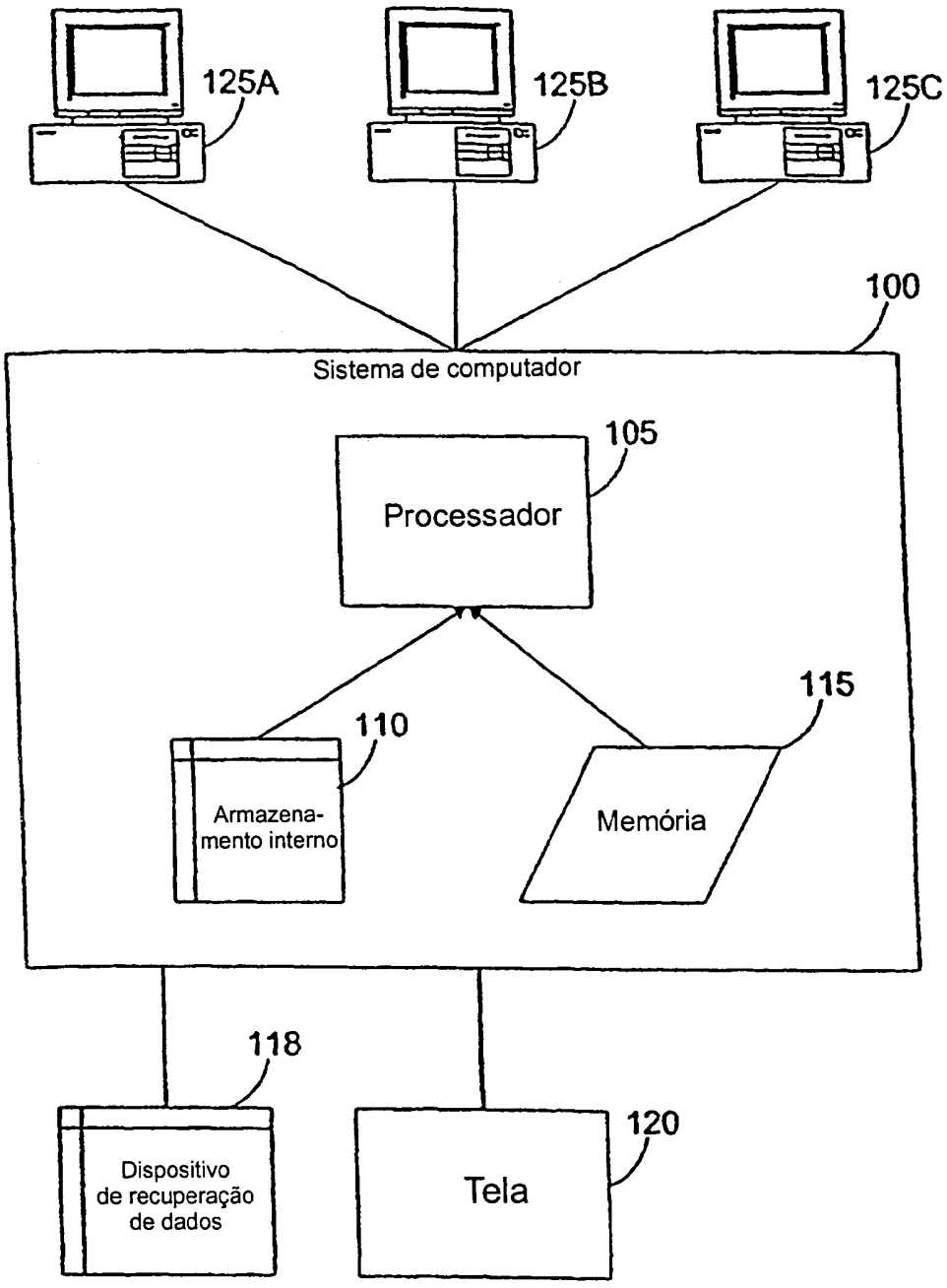
125A
125B
125C
100
Sistema de computador
105
Processador
110
Armazena-
mento interno
115
Memória
118
Dispositivo
de recuperação
de dados
120
Tela

FIG. 15

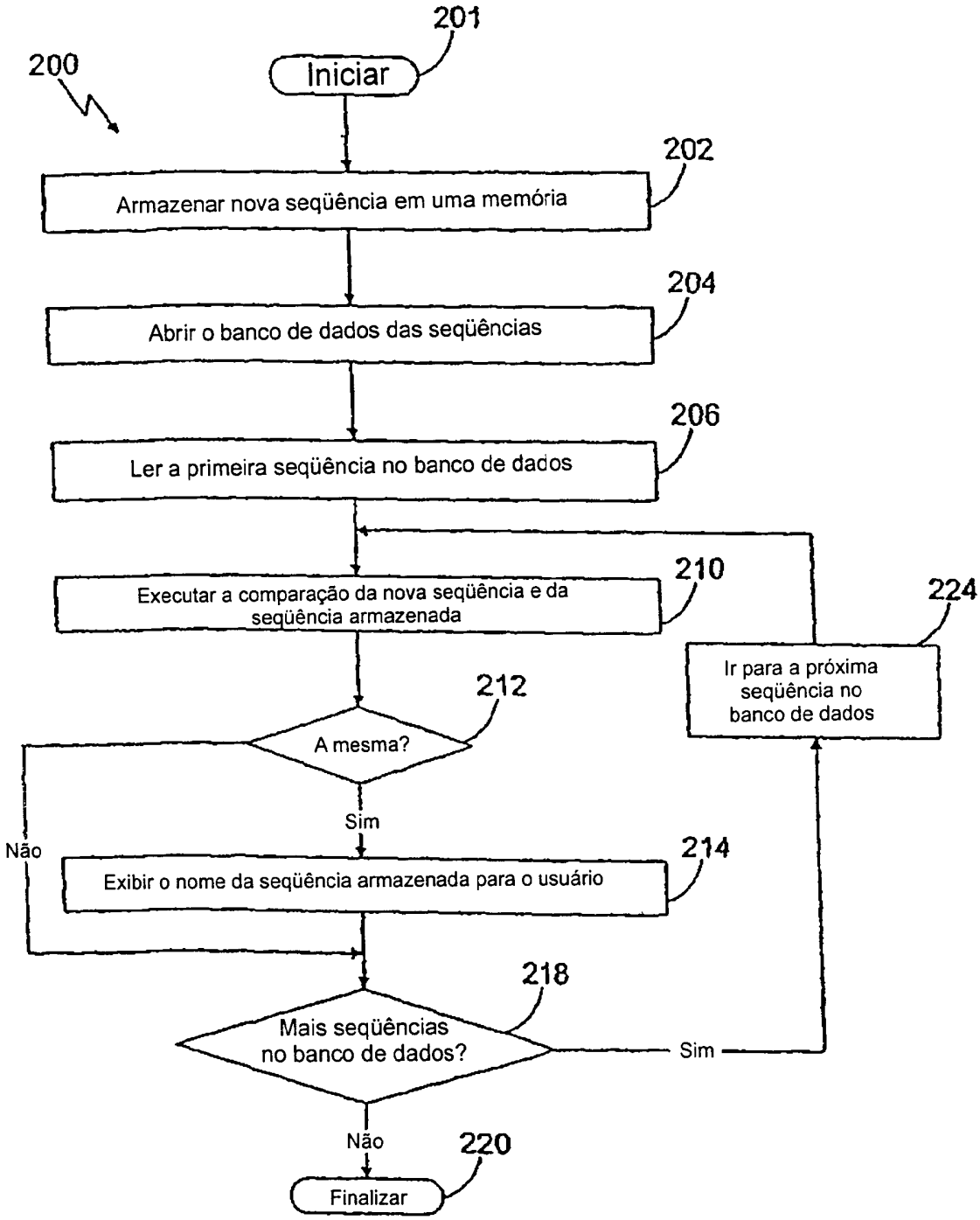
200
201
Iniciar
202
Armazenar nova seqüência em uma memória
204
Abrir o banco de dados das seqüências
206
Ler a primeira seqüência no banco de dados
210
Executar a comparação da nova seqüência e da seqüência armazenada
212
A mesma?
Não
Sim
214
Exibir o nome da seqüência armazenada para o usuário
218
Mais seqüências no banco de dados?
Sim
Não
220
Finalizar
224
Ir para a próxima seqüência no banco de dados

FIG. 16

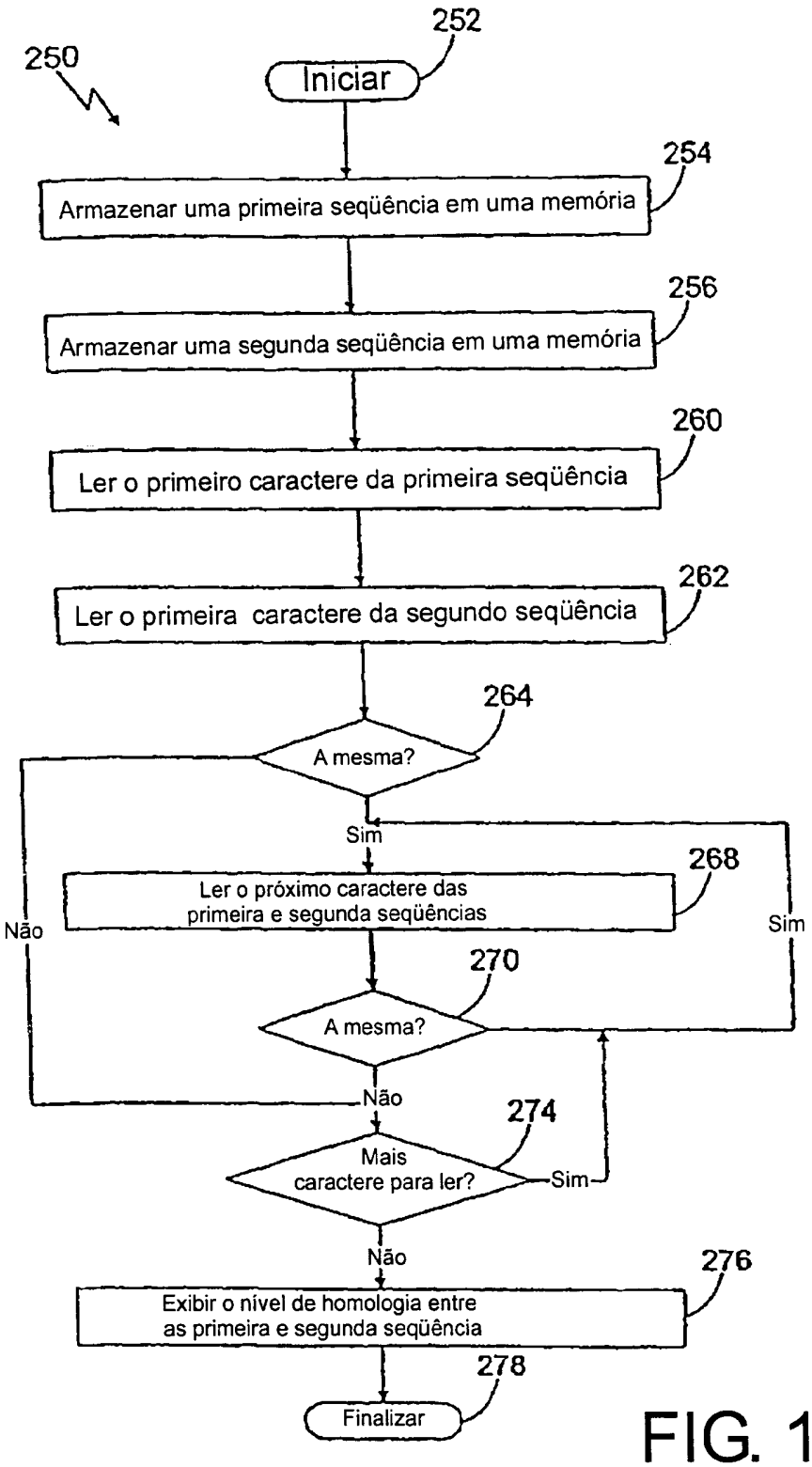
250
252
Iniciar
254
Armazenar uma primeira seqüência em uma memória
256
Armazenar uma segunda seqüência em uma memória
260
Ler o primeiro caractere da primeira seqüência
262
Ler o primeira caractere da segundo seqüência
264
A mesma?
Sim
Não
268
Ler o próximo caractere das primeira e segunda seqüências
Sim
270
A mesma?
Não
274
Mais caractere para ler?
Sim
Não
276
Exibir o nível de homologia entre as primeira e segunda seqüência
278
Finalizar

FIG. 17

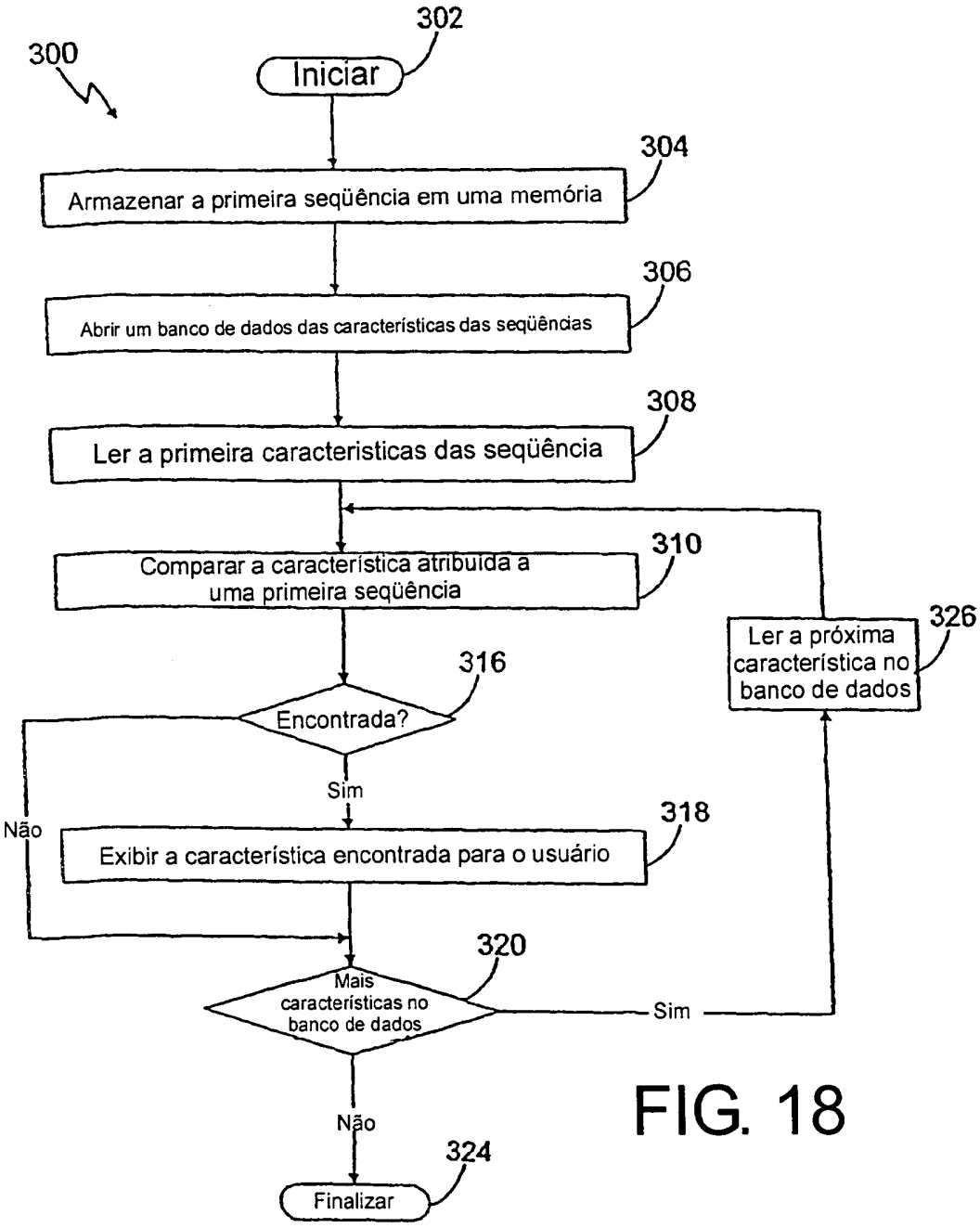
300
302
Iniciar
304
Armazenar a primeira seqüência em uma memória
306
Abrir um banco de dados das características das seqüências
308
Ler a primeira caracteristicas das seqüência
310
Comparar a característica atribuída a uma primeira seqüência
316
Encontrada?
Sim
Não
318
Exibir a característica encontrada para o usuário
320
Mais características no banco de dados
Sim
Não
324
Finalizar
326
Ler a próxima característica no banco de dados

FIG. 18

RESUMO

Patente de Invenção: **"FITASES RECOMBINANTES, PREPARAÇÃO DE PROTEÍNA, HETERODÍMERO, POLIPEPTÍDEO IMOBILIZADO, ARRANJO, SUPLEMENTO ALIMENTAR, MATRIZ DE LIBERAÇÃO DE ENZIMA COMESTÍVEL E PÉLETE COMESTÍVEL COMPREENDENDO AS MESMAS, BEM COMO MÉTODOS DE PRODUÇÃO DAS REFERIDAS ENZIMAS E PARA HIDROLISAR UM INOSITOL-HEXAFOSFATO, DESGRUDADURA DE ÓLEO E PROCESSAMENTO DE GRÃOS".**

A invenção fornece enzimas de fitase isoladas e recombinantes. Em um aspecto, as fitases são produzidas por modificação do appA do tipo selvagem de *E. coli*. A enzima pode ser produzida de células hospedeiras recombinantes. As fitases da invenção podem ser usadas para ajudar na digestão de fitato onde desejado. Em particular, as fitases da invenção podem ser usadas em gêneros alimentícios para melhorar o valor alimentício de ingredientes ricos em fitato. As fitases da invenção podem ser termotolerantes e/ou termoestáveis. Também fornecido são métodos para obter um polinucleotídeos variante codificando uma fitase e para obter uma fitase com termoestabilidade ou termotolerância em temperaturas altas ou baixas.

LISTAGEM DE SEQUÊNCIAS

<110> Verenium Corporation

<120> FITASES, ÁCIDOS NUCLÉICOS CODIFICANDO-AS E MÉTODOS PARA PRODUZIR E USÁ-LAS.

<130> DIVER1370-9

<140> PI0208851-7
<141> 24/05/2002

<150> US 09/866,379
<151> 24/05/2001

<150> US 09/580,515
<151> 25/05/2000

<150> US 09/318,528
<151> 25/05/1999

<150> US 09/291,931
<151> 13/04/1999

<150> US 09/259,214
<151> 01/03/1999

<150> US 08/910,798
<151> 13/08/1997

<160> 4

<170> PatentIn versão 3.1

<210> 1
<211> 1308
<212> DNA
<213> Sequência Artificial

<220>
<223> Enzima Fitase Modificada

<400> 1
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 agtttgagat ctcatcta 1308

<210> 2
<211> 436
<212> PRT
<213> Sequência Artificial

<220>

<223> Enzima Fitase Modificada

<400> 2

```
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

Arg Ser His Leu
435

<210> 3
<211> 1901
<212> DNA
<213> Escherichia coli fitase appA

<220>
<221> misc_feature
<222> (403)..(403)
<223> n é qualquer nucleotídeo

<400> 3
taaggagcag aaacaatgtg gtatttactt tggttcgtcg gcattttgtt gatgtgttcg 60

ctctccaccc ttgtgttggt atggctggac ccgcgtctga aaagttaacg aacgtaggcc 120

tgatgcggcg cattagcatc gcatcaggca atcaataatg tcagatatga aaagcggaaa 180

catatcgatg aaagcgatct taatcccatt tttatctctt ctgattccgt taaccccgca 240

atctgcattc gctcagagtg agccggagct gaagctggaa agtgtggtga ttgtcagtcg 300

tcatggtgtg cgtgctccaa ccaaggccac gcaactgatg caggatgtca ccccagacgc 360

atggccaacc tggccggtaa aactgggttg gctgacaccg cgnggtggtg agctaatcgc 420

ctatctcgga cattaccaac gccagcgtct ggtagccgac ggattgctgg cgaaaaaggg 480

ctgcccgcag tctggtcagg tcgcgattat tgctgatgtc gacgagcgta cccgtaaaac 540

aggcgaagcc ttcgccgccg ggctggcacc tgactgtgca ataaccgtac atacccaggc 600

agatacgtcc agtcccgatc cgttatttaa tcctctaaaa actggcgttt gccaactgga 660

taacgcgaac gtgactgacg cgatcctcag cagggcagga gggtcaattg ctgactttac 720

cgggcatcgg caaacggcgt ttcgcgaact ggaacgggtg cttaattttc cgcaatcaaa 780

cttgtgcctt aaacgtgaga aacaggacga aagctgttca ttaacgcagg cattaccatc 840

ggaactcaag gtgagcgccg acaatgtctc attaaccggt gcggtaagcc tcgcatcaat 900

gctgacggag atatttctcc tgcaacaagc acagggaatg ccggagccgg ggtggggaag 960

gatcaccgat tcacaccagt ggaacacctt gctaagtttg cataacgcgc aattttattt 1020

gctacaacgc acgccagagg ttgcccgcag ccgcgccacc ccgttattag atttgatcaa 1080

gacagcgttg acgccccatc caccgcaaaa acaggcgtat ggtgtgacat tacccacttc 1140

agtgctgttt atcgccggac acgatactaa tctggcaaat ctcggcggcg cactggagct 1200

caactggacg cttcccggtc agccggataa cacgccgcca ggtggtgaac tggtgtttga 1260

acgctggcgt cggctaagcg ataacagcca gtggattcag gtttcgctgg tcttccagac 1320

tttacagcag atgcgtgata aaacgccgct gtcattaaat acgccgcccg gagaggtgaa 1380

actgaccctg gcaggatgtg aagagcgaaa tgcgcagggc atgtgttcgt tggcaggttt 1440

tacgcaaatc gtgaatgaag cacgcatacc ggcgtgcagt ttgtaatgca taaaaaagag 1500

cattcagtta cctgaatgct ctgaggctga tgacaaacga agaactgtct aatgcgtaga 1560

ccggaaaagg cgttcacgcc gcatccggcc actttcagtt ttcctctttc tcggagtaac 1620

tataaccgta atagttatag ccgtaactgt aagcggtgct ggcgcgttta atcacaccat 1680

tgaggatagc gcctttaata ttgacgcctg cctgttccag acgctgcatt gacaaactca 1740

cctctttggc ggtgttcaag ccaaaacgcg caaccagcag gctggtgcca acagaacgcc 1800

ccacgaccgc ggcatcactc accgccagca tcggcggcgt atcgacaatc accagatcgt 1860

aatggtcgtt cgcccattcc agtaattgac gcatccgatc g 1901

<210> 4
<211> 432
<212> PRT
<213> Escherichia coli fitase appA

<400> 4

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 Trp Leu Thr Pro Arg Gly Gly Glu Leu Ile Ala Tyr Leu
65 70 75 80

Gly His Tyr Gln Arg Gln Arg Leu Val Ala Asp Gly Leu Leu Ala Lys
85 90 95

Lys 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 Ser Arg Ala Gly Gly Ser Ile Ala Asp
                165                 170                 175

Phe Thr Gly His Arg 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 Asn 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 Tyr 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
```

Este anexo apresenta o código de controle da listagem de sequências biológicas de que trata a Resolução INPI 228 de 11/11/2009:

Código de Controle

Campo 1



47880E43E84F0F77

Campo 2



A91F4B28EB12DE7D

Outras Informações:

- Nome do Arquivo: P123635 19627978_1.txt
- Data de Geração do Código: 31-10-2013
- Hora de Geração do Código: 13:57:29
- Código de Controle:
    - Campo 1: 47880E43E84F0F77
    - Campo 2: A91F4B28EB12DE7D