República Federativa do Brasil
Ministério da Economia
Instituto Nacional da Propriedade Industrial

(11) PI 0922641-9 B1

(22) Data do Depósito: 17/12/2009

(45) Data de Concessão: 26/10/2021

(54) Título: MÉTODO DE INTEGRAÇÃO DE UMA OU MAIS SEQUÊNCIAS DE ÁCIDO NUCLEICO EXÓGENAS NO GENOMA DE UMA CÉLULA VEGETAL

(51) Int.Cl.: C12N 15/05; C12Q 1/68.

(30) Prioridade Unionista: 17/12/2008 US 61/201946.

(73) Titular(es): DOW AGROSCIENCES LLC; SANGAMO BIOSCIENCES, INC..

(72) Inventor(es): HOLLY J.BUTLER; DAVID R.CORBIN; YANNICK DOYON; ZHIFANG GAO; VIPULA K. SHUKLA; FYODOR URNOV; SARAH E. WORDEN.

(86) Pedido PCT: PCT US2009006606 de 17/12/2009

(87) Publicação PCT: WO 2010/077319 de 08/07/2010

(85) Data do Início da Fase Nacional: 16/06/2011

(57) Resumo: INTEGRAÇÃO OBJETIVADA NO LOCUS DE Zp15. A invenção refere-se a métodos e composições para integração objetivada de uma sequência oxógena em um locus de ZP de planta, por exemplo, para a expressão de um polipeptídeo de interesse.

Relatório Descritivo da Patente de Invenção para **"MÉTODO DE INTEGRAÇÃO DE UMA OU MAIS SEQUÊNCIAS DE ÁCIDO NUCLEICO EXÓGENAS NO GENOMA DE UMA CÉLULA VEGETAL".**

DECLARAÇÃO DE DIREITOS A INVENÇÕES FEITA SOB PESQUISA CUSTEADA PELO GOVERNO FEDERAL

[001] Não aplicável.

CAMPO TÉCNICO

[002] A presente invenção está no campo de engenharia genômica de planta, particularmente integração objetivada de um transgene em um gene *Zp15* de planta.

ANTECEDENTES

[003] A biotecnologia tem emergido como uma ferramenta essencial em esforços para ir de encontro ao desafio de aumento da demanda global pela produção de alimento. Abordagens convencionais para aprimoramento da produtividade agrícola, por exemplo, rendimento intensificado ou resistência manipulada a pestes, contam com melhoramento por mutação ou introdução de novos genes nos genomas de espécies de safra por meio de transformação. Ambos os processos são inerentemente não específicos e relativamente ineficientes. Por exemplo, métodos convencionais de transformação de planta distribuem DNA exógeno que se integram no genoma em locais aleatórios. Assim, de forma a identificar e isolar linhagens transgênicas com atributos desejáveis, é necessário gerar milhares de eventos de integração aleatória única e subsequentemente realizar triagem quanto aos indivíduos desejados. Como um resultado, a manipulação convencional de traços de uma planta é trabalhosa, consome tempo e imprevisível. Além disso, a natureza aleatória dessas integrações torna difícil prever se efeitos pleiotrópicos em virtude de ruptura não intencional do genoma ocorreu. Como um resultado, a geração, isolamento e caracterização de linhagens de planta com genes ou traços manipulados têm sido um processo extremamente trabalhoso e de alto

custo, com uma baixa probabilidade de sucesso.

[004] A modificação gênica objetivada supera os desafios logísticos de práticas convencionais em sistemas de planta e, como tal, tem sido um objetivo há muito desejado, mas elusivo em pesquisa de biologia de planta básica e biotecnologia agrícola. Contudo, foi provado até recentemente que, com exceção de "objetivação gênica" via seleção de fármaco positiva-negativa em arroz ou o uso de sítios de restrição pré-manipulados, modificação genômica objetivada em todas as espécies de planta, modelo e safra, é muito difícil. Terada *et al.* (2002) *Nat Biotechnol* 20(10): 1030; Terada *et al.* (2007) *Plant Physiol* 144(2): 846; D'Halluin *et al.* (2008) *Plant Biotechnology J.* 6(1): 93.

[005] Recentemente, métodos e composições para clivagem objetivada de DNA genômico foram descritos. Tais eventos de clivagem objetivada podem ser usados, por exemplo, para induzir à mutagênese objetivada, induzir a deleções objetivadas de sequências de DNA celular e facilitar a recombinação objetivada em um *locus* cromossômico predeterminado. Vide, por exemplo, Publicações de Patente dos Estados Unidos 20030232410; 20050208489; 20050026157; 20050064474; e 20060188987 e Publicação Internacional WO 2007/014275, as divulgações das quais são incorporadas por referência em suas totalidades para todas as finalidades. A Publicação de Patente U.S. No. 20080182332 descreve o uso de nucleases do dedo de zinco (Zinc Finger Nucleases - ZFNs) não canônicas para modificação objetivada de genomas de planta e o Pedido de Patente U.S. No. 12/284.888 descreve a interação objetivada ZFN-mediada em um *locus* EPSPS de planta.

[006] Contudo, permanece uma necessidade por composições e métodos para integração objetivada estável em *loci* adicionais dentro do genoma de uma planta para estabelecimento de modificações genéticas estáveis, hereditárias na planta e sua prole.

SUMÁRIO

[007] A presente invenção proporciona métodos e composições para expressão de um ou mais produtos de uma sequência de ácido nucleico exógena (isto é, uma proteína ou uma molécula de RNA) que tenha sido integrada em um gene *Zp15* em uma célula vegetal. Conforme mostrado aqui, a integração de uma ou mais sequências exógenas em ou próximo do *locus* de *Zp15* não parece prejudicar a capacidade da planta hospedeira de regenerar, florescer ou produzir sementes e, opcionalmente, permite transmissão hereditária da(s) sequências(s) exógena(s) durante as gerações. As sequências de ácido nucleico exógenas podem compreender, por exemplo, um ou mais genes ou moléculas de cDNA ou qualquer tipo de sequência de codificação ou não codificação, bem como um ou mais elementos de controle (por exemplo, promotores). Por exemplo, genes de tolerância ao herbicida podem ser integrados nesse *locus* para produzir plantas de safra com a resistência ao herbicida desejada. Células contendo ácidos nucleicos exógenos em ou próximo do *locus* de *Zp15* também podem contribuir para o gametófito (linhagem germinativa) e, portanto, ser transmitidas para a prole em gerações subsequentes.

[008] A integração da sequência de ácido nucleico exógena em um gene *Zp15* é facilitada por clivagem de fita dupla objetivada do genoma no *locus* de *Zp15*. Clivagem é objetivada a um gene *Zp15* mediante o uso de proteínas de fusão compreendendo um domínio de ligação a DNA, tal como um domínio de ligação a DNA de meganuclease, um domínio de ligação a DNA de zíper de leucina, uma proteína do dedo de zinco (Zinc Finger Protein - ZFP) ou combinações quiméricas dos antes mencionados, o qual é manipulado para se ligar a uma sequência dentro do *locus* de *Zp15* selecionado e um domínio de clivagem ou um meio-domínio de clivagem. Tal clivagem estimula a integração de sequências de polinucleotídeo exógenas em ou próximo do sítio de clivagem. Integração de sequências exógenas pode se processar por meio de mecanismos dependentes de homologia e independentes de homologia.

[009] Em um aspecto, são divulgados aqui domínios de ligação a DNA manipulados (por exemplo, ZFPs, meganucleases ou zíperes de leucina) que se ligam a um sítio-alvo em um gene *Zp15*. O domínio de ligação a DNA pode compreender, por exemplo, qualquer um dos domínios de ligação a DNA de dedo de zinco manipulados compreendendo as hélices de reconhecimento mostradas na Tabela 1. Qualquer um dos domínios de ligação a DNA descritos aqui pode ainda compreender um domínio funcional, por exemplo, um domínio de clivagem ou meio-domínio de clivagem. Em algumas modalidades, o meio-domínio de clivagem pode ser de uma endonuclease de restrição do tipo IIS, tal como *FokI* ou *StsI.* Em outras modalidades, o domínio de clivagem pode compreender uma endonuclease "homing", por exemplo, uma endonuclease "homing" com um domínio de ligação a DNA modificado.

[0010] Em outro aspecto, são divulgadas aqui plantas ou sementes compreendendo uma sequência exógena integrada no *locus* de *Zp15*. Em determinadas modalidades, a sequência exógena é integrada no gametófito da planta.

[0011] Em outro aspecto, é divulgado aqui um método para expressão do produto de uma sequência de ácido nucleico exógena em uma célula, o método compreendendo: (a) expressão de uma primeira proteína de fusão na célula, a primeira proteína de fusão compreendendo um primeiro domínio de ligação a DNA (por exemplo, uma ZFP) e um primeiro meio-domínio de clivagem, em que o domínio de ligação a DNA tenha sido manipulado para se ligar a um primeiro sítio-alvo em um gene *Zp15* do genoma da célula; (b) expressão de uma segunda proteína de fusão na célula, a segunda proteína de fusão compreendendo um segundo domínio de ligação a DNA e um segundo meio-domínio de clivagem, em que o segundo domínio de ligação a DNA se liga a um segundo sítio-alvo no gene *Zp15* do genoma da célula, em que o segundo sítio-alvo é diferente do primeiro sítio-alvo; e (c) contato da célula com um polinucleotí-

deo compreendendo uma sequência de ácido nucleico exógena e uma primeira sequência de nucleotídeo que é homóloga à primeira sequência em um gene *Zp15*; em que ligação da primeira proteína de fusão ao primeiro sítio-alvo e ligação da segunda proteína de fusão ao segundo sítio-alvo posiciona os meio-domínios de clivagem de modo que o genoma da célula seja clivado no gene *Zp15*, desse modo, resultando em integração da sequência exógena no genoma da célula no gene *Zp15* e expressão do produto da sequência exógena.

[0012] A sequência de ácido nucleico exógena pode compreender uma sequência que codifica um ou mais polipeptídeos funcionais (por exemplo, um cDNA), com ou sem um ou mais promotores e/ou pode produzir uma ou mais sequências de RNA (por exemplo, via um ou mais cassetes de expressão de shRNA) as quais conferem traços desejáveis à planta. Tais traços incluem, mas não estão limitados a, resistência ou tolerância ao herbicida; resistência ou tolerância a insetos; resistência ou tolerância a doenças (virais, bacterianas, fúngicas, por nematoide); tolerância e/ou resistência a estresse, conforme exemplificado por resistência ou tolerância à estiagem, calor, frio, congelamento, umidade excessiva, estresse por sal; estresse oxidativo; rendimentos aumentados; teor e composição alimentícia; aparência física; esterilidade masculina; velocidade de secagem; sustentabilidade; abundância; quantidade e qualidade de amido; quantidade e qualidade de óleo; qualidade e quantidade de proteína; composição de aminoácido; e semelhantes. Naturalmente, quaisquer dois ou mais ácidos nucleicos exógenos de qualquer descrição, tais como aqueles que conferem resistência ao herbicida, insetos, doenças (virais, bacterianas, fúngicas, por nematoide) ou estiagem, esterilidade masculina, velocidade de secagem, sustentabilidade, abundância, propriedades de amido, quantidade e qualidade de óleo ou aqueles que aumentam o rendimento ou qualidade nutricional podem ser empregados, conforme desejado. Em determinadas modalidades, a sequência de ácido

nucleico compreende uma sequência que codifica uma proteína de resistência a herbicida (por exemplo, o gene *AAD-1* (Aryloxy Alkanoate Dioxygenase - Dioxigenase de Arilóxi Alcanoato), o gene *AAD-12* ou o gene de acetil transferase de fosfinotricina (Phosphinothricin Acetyl Transferase - *PAT*) e/ou fragmentos funcionais dos mesmos. Expressão da sequência integrada pode ser acionada por um promotor operativamente ligado à sequência integrada. Alternativamente, a sequência integrada é sem promotor e a transcrição é acionada pelo promotor de *Zp15* endógeno.

[0013] Em determinadas modalidades, o polinucleotídeo ainda compreende uma segunda sequência de nucleotídeo que é homóloga a uma segunda sequência no gene *Zp15*. A segunda sequência de nucleotídeo pode ser idêntica à segunda sequência no gene *Zp15*. Além disso, em modalidades compreendendo primeira e segunda sequências de nucleotídeo, a primeira sequência de nucleotídeo pode ser idêntica à primeira sequência no gene *Zp15* e a segunda sequência de nucleotídeo pode ser homóloga, mas não idêntica, a uma segunda sequência no gene *Zp15*. Em qualquer um dos métodos descritos aqui, as primeira e segunda sequências de nucleotídeo flanqueiam a sequência exógena. Em determinadas modalidades, o polinucleotídeo é um plasmídeo. Em outras modalidades, o polinucleotídeo é uma molécula de DNA linear.

[0014] Em outro aspecto, é proporcionado aqui um método para integração de uma sequência exógena no gene *Zp15* no genoma de uma célula, o método compreendendo: (a) expressão de uma primeira proteína de fusão na célula, a primeira proteína de fusão compreendendo um primeiro domínio de ligação a DNA (por exemplo, uma ZFP) e um primeiro meio-domínio de clivagem, em que o primeiro domínio de ligação a DNA foi manipulado para se ligar a um primeiro sítio-alvo no *locus* de *Zp15* no genoma da célula; (b) expressão de uma segunda proteína de fusão na célula, a segunda proteína de fusão compreendendo um segun-

do domínio de ligação a DNA (por exemplo, uma ZFP) e um segundo meio-domínio de clivagem, em que o segundo domínio de ligação a DNA se liga a um segundo sítio-alvo no *locus* de *Zp15* no genoma da célula, em que o segundo sítio-alvo é diferente do primeiro sítio-alvo; e (c) contato da célula com um polinucleotídeo compreendendo uma sequência de ácido nucleico exógena; em que ligação da primeira proteína de fusão ao primeiro sítio-alvo e ligação da segunda proteína de fusão ao segundo sítio-alvo posicionam os meio-domínios de clivagem de modo que o genoma da célula seja clivado no *locus* de *Zp15*, desse modo, resultando em integração dependente de homologia da sequência exógena no genoma da célula dentro do *locus* de *Zp15*. Em determinadas modalidades, uma sequência exógena que codifica um polipeptídeo funcional é inserida no gene *Zp15*.

[0015] Em qualquer um dos métodos descritos aqui, os primeiro e segundo meio-domínios de clivagem podem ser de uma endonuclease de restrição do tipo IIS, por exemplo, *Fok*I ou *Sts*I. Além disso, em qualquer um dos métodos descritos aqui, pelo menos uma das proteínas de fusão pode compreender uma alteração na sequência de aminoácido da interface de dimerização do meio-domínio de clivagem, por exemplo, de modo que heterodímeros obrigatórios dos meio-domínios de clivagem sejam formados.

[0016] Em qualquer um dos métodos descritos aqui, a célula vegetal pode compreender uma célula vegetal monocotiledônea ou dicotiledônea. Em determinadas modalidades, a célula vegetal é uma planta de safra, por exemplo, milho.

BREVE DESCRIÇÃO DOS DESENHOS

[0017] A **Figura 1** representa os resultados de análise de sequência exemplificativos de produtos da amplificação de *Zp15* a partir de gDNA de milho HiII derivado de células submetidas à expressão transitória do par de ZFP #25 (sítios de ligação sublinhados) e revela uma inserção

NHEJ de 6 bp (negrito) no sítio de clivagem esperado.

[0018] A **Figura 2** representa resultados de análise de sequência exemplificativos de produtos da amplificação de *Zp15* de gDNA de milho HiII derivado de células submetidas à expressão transitória do par de ZFP #24 (sítios de ligação sublinhados) e revela uma deleção de 3 bp no sítio de clivagem esperado.

[0019] A **Figura 3** é uma representação esquemática do construto designado pDAB7489.

[0020] A **Figura 4** é um esquema representando um cassete de expressão de gene de tolerância a herbicida exemplificativo que codifica um gene *AAD*.

[0021] A **Figura 5** é um esquema representando o construto designado pDAB7490.

[0022] A **Figura 6** representa o alinhamento de um evento de integração objetivada (Targeted Integration - TI) em que o fragmento doador de *Zp15* de milho do tipo silvestre (Wild Type - WT), bem como as regiões de borda 5' e 3' que unem as sequências doadoras integradas, são alinhados.

[0023] A **Figura 7** é um esquema mostrando o construto designado pDAB104101.

[0024] A **Figura 8** é um esquema representando o cassete de expressão de *PAT*.

[0025] A **Figura 9** é um esquema representando o construto designado pDAB104107.

[0026] A **Figura 10** é um esquema representando o construto designado pDAB104104.

[0027] A **Figura 11** é um esquema representando o construto designado pDAB104105.

[0028] A **Figura 12** é um esquema representando o construto designado pDAB104106.

[0029] A **Figura 13** é um esquema representando o construto designado pDAB104100.

[0030] A **Figura 14** é um esquema representando o construto designado pDAB104103.

[0031] A **Figura 15** é um esquema representando o construto designado pDAB104102.

DESCRIÇÃO DETALHADA

[0032] A presente invenção se refere a métodos e composições para integração objetivada (TI) em um gene *Zp15* de planta, o qual reside sobre o cromossomo 6 em milho. Usando proteínas de fusão compreendendo domínios de ligação a DNA (por exemplo, ZFPs, meganucleases ou zíperes de leucina) e domínios de nuclease, uma sequência inserida (doadora) pode ser operativamente ligada a um promotor exógeno ou pode ser sem promotor. Se sem promotor, transcrição do quadro de leitura aberto integrado pode ocorrer a partir do promotor do gene *Zp15* endógeno nos tecidos promotor-especificados. Uso de um doador sem promotor diminui a probabilidade de integração aleatória do doador e/ou ativação supérflua de um gene endógeno pelo promotor trazido sobre o doador.

[0033] Composições úteis para clivagem e recombinação objetivada em um gene *Zp15* incluem proteínas de fusão compreendendo um domínio de clivagem (ou um meio-domínio de clivagem) e um domínio de ligação a DNA (por exemplo, uma ZFP), polinucleotídeos que codificam essas proteínas e combinações de polipeptídeos e polinucleotídeos que codificam polipeptídeos. Um domínio de ligação de dedo de zinco pode compreender um ou mais dedos de zinco (por exemplo, 2, 3, 4, 5, 6, 7, 8, 9 ou mais dedos de zinco) e pode ser manipulado para se ligar a qualquer sequência dentro de um gene *Zp15*. A presença de tal proteína (ou proteínas) de fusão em uma célula resultará em ligação da(s) proteína(s) de fusão a seu(s) sítio(s) de ligação e clivagem dentro do gene *Zp15* en-

dógeno.

Geral

[0034] A prática dos métodos, bem como preparo e uso das composições divulgadas aqui empregam, a menos que de outro modo indicado, técnicas convencionais em biologia molecular, bioquímica, estrutura e análise de cromatina, química computacional, cultura de célula, DNA recombinante e campos relacionados, conforme dentro da capacidade no campo. Essas técnicas são totalmente explicadas na literatura. Vide, por exemplo, Sambrook *et al.* MOLECULAR CLONING: A LABORATORY MANUAL, Segunda edição, Cold Spring Harbor Laboratory Press, 1989 e Terceira edição, 2001; Ausubel *et al,* CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, New York, 1987 e atualizações periódicas; a série METHODS IN ENZYMOLOGY, Academic Press, San Diego; Wolffe, CHROMATIN STRUCTURE AND FUNCTION, Terceira edição, Academic Press, San Diego, 1998; METHODS IN ENZYMOLOGY, Vol. 304, "Chromatin" (P.M. Wassarman e A. P. Wolffe, eds.), Academic Press, San Diego, 1999; e METHODS IN MOLECULAR BIOLOGY, Vol. 119, "Chromatin Protocols" (P.B. Becker, ed.) Humana Press, Totowa, 1999.

Definições

[0035] Os termos "ácido nucleico", "polinucleotídeo" e "oligonucleotídeo" são usados permutavelmente e referem-se a um polímero de deoxirribonucleotídeo ou ribonucleotídeo, em conformação linear ou circular e na forma fita simples ou dupla. Para fins da presente invenção, esses termos não devem ser construídos como limitativos com relação ao comprimento de um polímero. Os termos podem abranger análogos conhecidos de nucleotídeos naturais, bem como nucleotídeos que são modificados nas porções de base, açúcar e/ou fosfato (por exemplo, partes principais de fosforotioato). Em geral, um análogo de um nucleotídeo em particular tem a mesma especificidade de pareamento de base; isto é, um

análogo de A realizará pareamento de base com T.

[0036] Os termos "polipeptídeo", "peptídeo" e "proteína" são usados permutavelmente para referir-se a um polímero de resíduos de aminoácido. O termo também se aplica a polímeros de aminoácido nos quais um ou mais aminoácidos são análogos químicos ou derivados modificados de um aminoácido que ocorre naturalmente correspondente.

[0037] "Ligação" se refere a uma interação não covalente, sequência-específica entre macromoléculas (por exemplo, entre uma proteína e um ácido nucleico). Nem todos os componentes de uma interação de ligação precisam ser sequência-específicos (por exemplo, contatos com resíduos de fosfato em uma parte principal de DNA), contanto que a interação como um todo seja sequência-específica. Tais interações são, em geral, caracterizadas por uma constante de dissociação (K_d) de 10^{-6} M^{-1} ou menor. "Afinidade" se refere à intensidade de ligação: a afinidade de ligação aumentada estando correlacionada a uma menor K_d.

[0038] Uma "proteína de ligação" é uma proteína que é capaz de ligação não covalentemente à outra molécula. Uma proteína de ligação pode se ligar, por exemplo, a uma molécula de DNA (uma proteína de ligação a DNA), uma molécula de RNA (uma proteína de ligação a RNA) e/ou uma molécula de proteína (uma proteína de ligação à proteína). No caso de uma proteína de ligação à proteína, ela pode se ligar a si mesma (para formar homodímeros, homotrímeros, etc.) e/ou ela pode se ligar a uma ou mais moléculas de uma proteína ou proteínas diferentes. Uma proteína de ligação pode ter mais de um tipo de atividade de ligação. Por exemplo, proteínas do dedo de zinco têm atividade de ligação a DNA, ligação a RNA e ligação à proteína.

[0039] Uma "proteína de ligação a DNA de dedo de zinco" (ou domínio de ligação) é uma proteína ou um domínio dentro de uma proteína maior que se liga ao DNA de uma maneira sequência-específica através de um ou mais dedos de zinco, os quais são regiões de sequência de

aminoácido dentro do domínio de ligação cuja estrutura é estabilizada mediante coordenação de um íon de zinco. O termo proteína de ligação a DNA de dedo de zinco é, frequentemente, abreviado como proteína de dedo de zinco ou ZFP.

[0040] Domínios de ligação de dedo de zinco podem ser "manipulados" para se ligar a uma sequência de nucleotídeo predeterminada. Exemplos não limitativos de métodos para manipulação de proteínas de dedo de zinco são design e seleção. Uma proteína de dedo de zinco projetada é uma proteína que não ocorre na natureza cujo design/composição resulta principalmente de critérios racionais. Critérios racionais para design incluem aplicação de regras de substituição e algoritmos computadorizados para processamento de informação em um banco de dados que armazena informação de designs de ZFP e dados de ligação existentes. Vide, por exemplo, Patentes US 6.140.081; 6.453.242; e 6.534.261; vide também documentos WO 98/53058; WO 98/53059; WO 98/53060; WO 02/016536 e WO 03/016496.

[0041] Uma proteína de dedo de zinco "selecionada" é uma proteína não encontrada na natureza cuja produção resulta primariamente de um processo empírico, tal como exibição de fago (“*phage display”)*, contenção para interação ou seleção de híbrido. Vide, por exemplo, documentos US 5.789.538; US 5.925.523; US 6.007.988; US 6.013.453; US 6.200.759; WO 95/19431; WO 96/06166; WO 98/53057; WO 98/54311; WO 00/27878; WO 01/60970 WO 01/88197 e WO 02/099084.

[0042] O termo "sequência" se refere a uma sequência de nucleotídeo de qualquer comprimento, a qual pode ser DNA ou RNA; pode ser linear, circular ou ramificada e pode ser fita simples ou fita dupla. O termo "sequência doadora" se refere a uma sequência de nucleotídeo que é inserida em um genoma. Uma sequência doadora pode ser de qualquer comprimento, por exemplo, entre 2 e 10.000 nucleotídeos de comprimento (ou qualquer valor inteiro entre os mesmos ou acima dos mesmos), de

preferência entre cerca de 100 e 1.000 nucleotídeos de comprimento (ou qualquer inteiro entre os mesmos), mais preferivelmente entre cerca de 200 e 500 nucleotídeos de comprimento.

[0043] Uma "sequência homóloga, não idêntica" se refere a uma primeira sequência a qual compartilha um grau de identidade de sequência com uma segunda sequência, mas cuja sequência não é idêntica àquela da segunda sequência. Por exemplo, um polinucleotídeo compreendendo a sequência do tipo silvestre de um gene mutante é homólogo e não idêntico à sequência do gene mutante. Em determinadas modalidades, o grau de homologia entre as duas sequências é suficiente para permitir recombinação homóloga entre as mesmas, utilizando mecanismos celulares normais. Duas sequências homólogas não idênticas podem ser de qualquer comprimento e seu grau de não homologia pode ser tão pequeno quanto um único nucleotídeo (por exemplo, para correção de uma mutação de ponto genética por meio de recombinação homóloga objetivada) ou tão grande quanto 10 ou mais quilobases (por exemplo, para inserção de um gene em um sítio ectópico predeterminado em um cromossomo). Dois polinucleotídeos compreendendo as sequências homólogas não idênticas não precisam ser do mesmo comprimento. Por exemplo, um polinucleotídeo exógeno (isto é, polinucleotídeo doador) de entre 20 e 10.000 nucleotídeos ou pares de nucleotídeo pode ser usado.

[0044] Técnicas para determinação de identidade de sequência de ácido nucleico e aminoácido são conhecidas no campo. Tipicamente, tais técnicas incluem determinação da sequência de nucleotídeo do mRNA para um gene e/ou determinação da sequência de aminoácido codificada pelo menos e comparação dessas sequências com uma segunda sequência de nucleotídeo ou aminoácido. Sequências genômicas podem também ser determinadas e comparadas desse modo. Em geral, identidade se refere a uma correspondência nucleotídeo-a-nucleotídeo ou aminoácido-a-aminoácido exata de duas sequências de polinucleotídeo ou

polipeptídeo, respectivamente. Duas ou mais sequências (polinucleotídeo ou aminoácido) podem ser comparadas determinando-se sua identidade percentual. A identidade percentual de duas sequências, quer sequências de ácido nucleico ou aminoácido, é o número de combinações exatas entre duas sequências alinhadas dividido pelo comprimento das sequências mais curtas e multiplicado por 100. Um alinhamento aproximado para sequências de ácido nucleico é proporcionado pelo algoritmo de homologia local de Smith e Waterman, Advances in Applied Mathematics 2: 482-489 (1981). Esse algoritmo pode ser aplicado a sequências de aminoácido usando a matriz de escore desenvolvida por Dayhoff, Atlas of Protein Sequences and Structure, M.O. Dayhoff ed., supl. 3: 353-358, National Biomedical Research Foundation, Washington, D.C., EUA e normalizada por Gribskov, Nucl. Acids Res. 14(6): 6745-6763 (1986). Uma implementação exemplificativa desse algoritmo para determinar a identidade percentual de uma sequência é fornecida pelo Genetics Computer Group (Madison, WI) no aplicativo "BestFit". Programas adequados para cálculo da identidade ou similaridade percentual entre sequências são, em geral, conhecidos na técnica, por exemplo. Outro programa de alinhamento é o BLAST, usado com os parâmetros de default. Por exemplo, o BLASTN e BLASTP podem ser usados utilizando os seguintes parâmetros de default: código genético = padrão; filtro = nenhum; fita = ambas; corte = 60; esperado = 10; Matriz = BLOSUM62; Descrições = 50 sequência; escolhido por = HIGH SCORE; Bancos de Dados = não redundante, traduções no GenBank + EMBL + DDBJ + PDB + GenBank CDS + Swiss Protein + Spupdate + PIR. Detalhes desses programas podem ser encontrados na internet. Com relação às sequências descritas aqui, a faixa de graus desejados de identidade de sequência é aproximadamente 80% a 100% e qualquer valor inteiro entre os mesmos. Tipicamente, as identidades percentuais entre as sequências são pelo menos 70-75%, de preferência 80-82%, mais preferivelmente 85-90%, ainda mais preferivelmente 92%,

mais preferivelmente 95% e, ainda mais preferivelmente, 98% de identidade de sequência.

[0045] Alternativamente, o grau de similaridade de sequência entre polinucleotídeos pode ser determinado por meio de hibridização de polinucleotídeos sob condições que permitam a formação de duplas estáveis entre regiões homólogas, seguido por digestão com nuclease(s) específica(s) para fita simples e determinação de tamanho dos fragmentos digeridos. Duas sequências de ácido nucleico ou duas de polipeptídeo são substancialmente homólogas uma à outra quando as sequências exibem pelo menos cerca de 70%-75%, de preferência 80%-82%, mais preferivelmente 85%-90%, ainda mais preferivelmente 92%, mais preferivelmente 95% e, ainda mais preferivelmente, 98% de identidade de sequência sobre um comprimento definido das moléculas, conforme determinado usando os métodos acima. Conforme usado aqui, substancialmente homólogo se refere também à sequências mostrando identidade completa a uma sequência de DNA ou polipeptídeo especificada. Sequências de DNA que são substancialmente homólogas podem ser identificadas em um experimento de hibridização de Southern, por exemplo, sob condições estringentes, conforme definido para esse sistema em particular. Definição de condições de hibridização apropriadas está dentro da capacidade no campo. Vide, por exemplo, Sambrook *et al, supra;* Nucleic Acid Hybridization: A Practical Approach, editores B.D. Hames e S.J. Higgins, (1985) Oxford; Washington, DC; IRL Press).

[0046] Hibridização seletiva de dois fragmentos de ácido nucleico pode ser determinada como segue. O grau de identidade de sequência entre duas moléculas de ácido nucleico afeta a eficiência e resistência de eventos de hibridização entre tais moléculas. Uma sequência de ácido nucleico parcialmente idêntica inibirá, pelo menos parcialmente, a hibridização de uma sequência completamente idêntica a uma molécula-alvo. Inibição de hibridização da sequência completamente idêntica pode ser

avaliada usando ensaios de hibridização que são bem-conhecidos no campo (por exemplo, Southern (DNA) blot, Northern (RNA) blot, hibridização em solução ou semelhante; vide Sambrook, *et al, Molecular Cloning: A Laboratory Manual,* Segunda Edição, (1989) Cold Spring Harbor, N.Y.). Tais ensaios podem ser conduzidos usando graus variados de seletividade, por exemplo, usando condições variáveis de baixa à alta estringência. Se condições de baixa estringência são empregadas, a ausência de ligação não específica pode ser avaliada usando uma sonda secundária que carece mesmo de um grau parcial de identidade de sequência (por exemplo, uma sonda tendo menos de cerca de 30% de identidade de sequência com a molécula-alvo) de modo que, na ausência de eventos de ligação não específica, a sonda secundária não se hibridize ao alvo.

[0047] Quando de utilização de um sistema de detecção baseado em hibridização, uma sonda de ácido nucleico é escolhida que é complementar a uma sequência de ácido nucleico de referência e, então, através de seleção de condições apropriadas, a sonda e a sequência de referência hibridizam seletivamente ou se ligam uma à outra para formar uma molécula dupla. Uma molécula de ácido nucleico que é capaz de hibridização seletiva a uma sequência de referência sob condições de hibridização moderadamente estringentes se hibridiza, tipicamente, sob condições que permitem a detecção de uma sequência de ácido nucleico-alvo de pelo menos cerca de 10-14 nucleotídeos de comprimento tendo pelo menos aproximadamente 70% de identidade de sequência com a sequência da sonda de ácido nucleico selecionada. Condições de hibridização estringentes permitem, tipicamente, a detecção de sequências de ácido nucleico-alvo de pelo menos cerca de 10-14 nucleotídeos de comprimento tendo uma identidade de sequência de mais de cerca de 90-95% com a sequência da sonda de ácido nucleico selecionada. Condições de hibridização úteis para hibridização de sonda/sequência de referência, onde a sonda e a sequência de referência têm um grau específico de identidade

de sequência, podem ser determinadas conforme é conhecido no campo (vide, por exemplo, Nucleic Acid Hvbridization: A Practical Approach, editores B.D. Hames e S.J. Higgins, (1985) Oxford; Washington, DC; IRL Press).

[0048] Condições para hibridização são bem-conhecidas por aqueles versados no campo. Estringência de hibridização se refere ao grau até o qual as condições de hibridização desfavorecem a formação de híbridos contendo nucleotídeos erroneamente correspondentes, com maior estringência correlacionada a uma menor tolerância para híbridos erroneamente correspondentes. Fatores que afetam a estringência de hibridização são bem-conhecidos por aqueles versados no campo e incluem, mas não estão limitados a, temperatura, pH, resistência iônica e concentração de solventes orgânicos tais como, por exemplo, formamida e sulfóxido de dimetila. Conforme é conhecido por aqueles versados no campo, a estringência de hibridização pode ser aumentada por maiores temperaturas, menor resistência iônica e menores concentrações de solvente.

[0049] Com relação às condições de estringência para hibridização, é bem-conhecido no campo que numerosas condições equivalentes podem ser empregadas para estabelecer uma estringência particular variando, por exemplo, os seguintes fatores: o comprimento e natureza das sequências, composição de base das várias sequências, concentrações de sais e outros componentes de hibridização, a presença ou ausência de agentes de bloqueio nas soluções de hibridização (por exemplo, sulfato de dextrana e polietileno glicol), temperatura da reação de hibridização e parâmetros de tempo, bem como variando as condições de lavagem. A seleção de um conjunto particular de condições de hibridização é selecionada seguindo métodos padrões no campo (vide, por exemplo, Sambrook, *et al.,* Molecular Cloning: A Laboratory Manual. Segunda Edição, (1989) Cold Spring Harbor, N.Y.).

[0050] "Recombinação" se refere a um processo de troca de informa-

ção genética entre dois polinucleotídeos. Para fins da presente invenção, "recombinação homóloga (Homologous Recombination - HR)" se refere à forma especializada de tal troca que ocorre, por exemplo, durante reparo de rupturas de fita dupla em células. Esse processo requer homologia de sequência de nucleotídeo, usa uma molécula "doadora" como molde para reparo de uma molécula-"alvo" (isto é, uma que experimentou a ruptura de fita dupla) e é variadamente conhecido como "conversão gênica sem cruzamento" ou "conversão gênica de trato curto" porque ele leva à transferência de informação genética do doador para o alvo. Sem desejar estar preso por qualquer teoria em particular, tal transferência pode envolver correção de correspondência errônea de DNA heteroduplex que se forma entre o alvo rompido e o doador e/ou "anelamento de fita síntese-dependente", no qual o doador é usado para ressintetizar informação genética que se tornará parte do alvo e/ou processos relacionados. Tal HR especializada frequentemente resulta em uma alteração da sequência da molécula-alvo, de modo que parte ou toda a sequência do polinucleotídeo doador é incorporada no polinucleotídeo-alvo.

[0051] "Clivagem" se refere à ruptura da parte principal covalente de uma molécula de DNA. Clivagem pode ser iniciada através de uma variedade de métodos incluindo, mas não limitado a, hidrólise enzimática ou química de uma ligação de fosfodiéster. Clivagem de fita simples e clivagem de fita dupla são possíveis e clivagem de fita dupla pode ocorrer como um resultado de dois eventos de clivagem de fita simples distintos. Clivagem de DNA pode resultar na produção de extremidades cegas ou extremidades alternadas. Em determinadas modalidades, polipeptídeos de fusão são usados para clivagem de DNA fita dupla-alvo.

[0052] Um "domínio de clivagem" compreende uma ou mais sequências polipeptídicas as quais possuem atividade catalitica para clivagem de DNA. Um domínio de clivagem pode estar contido em uma única cadeia polipeptídica ou atividade de clivagem pode resultar da associação de

dois (ou mais) polipeptídeos.

[0053] Um "meio-domínio de clivagem" é uma sequência polipeptídica a qual, em conjunto com um segundo polipeptídeo (quer idêntico ou diferente) forma um complexo tendo atividade de clivagem (de preferência atividade de clivagem de fita dupla).

[0054] "Cromatina" é a estrutura de nucleoproteína compreendendo o genoma celular. Cromatina celular compreende ácido nucleico, primariamente DNA e proteína, incluindo proteínas cromossômicas de histona e não-histona. A maioria da cromatina celular eucariota existe na forma de nucleossomas, em que um núcleo de nucleossoma compreende aproximadamente 150 pares de base de DNA associado a um octâmero compreendendo duas de cada uma das histonas H2A, H2B, H3 e H4; e DNA ligante (de comprimento variável, dependendo do organismo) se estende entre os núcleos de nucleossoma. Uma molécula de histona H1 está, em geral, associada ao DNA ligante. Para fins da presente invenção, o termo "cromatina" se destina a abranger todos os tipos de nucleoproteína celular, procariota e eucariota. A cromatina celular inclui cromatina cromossômica e episssomal.

[0055] Um "cromossomo" é um complexo de cromatina compreendendo todo ou uma parte do genoma de uma célula. O genoma de uma célula é, frequentemente, caracterizado por seu cariótipo, o qual é a coleção de todos os cromossomos que compreendem o genoma da célula. O genoma de uma célula pode compreender um ou mais cromossomos.

[0056] Um "episssoma" é um ácido nucleico de replicação, complexo de nucleoproteína ou outra estrutura compreendendo um ácido nucleico que não é parte do cariótipo cromossômico de uma célula. Exemplos de episssomas incluem plasmídeos e determinados genomas virais.

[0057] Uma "região acessível" é um local na cromatina celular no qual um sítio-alvo presente no ácido nucleico pode ser ligado por uma molécula exógena a qual reconhece o sítio-alvo. Sem desejar estar preso

por qualquer teoria em particular, acredita-se que uma região acessível é uma que não está engolfada em uma estrutura nucleossômica. A estrutura distinta de uma região acessível pode, frequentemente, ser detectada por sua sensibilidade a sondas químicas e enzimáticas, por exemplo, nucleases.

[0058] Um "sítio-alvo" ou "sequência-alvo" é uma sequência de ácido nucleico que define uma porção de um ácido nucleico à qual uma molécula de ligação se ligará, contanto que existam condições suficientes para ligação. Por exemplo, a sequência 5'-GAATTC-3' é o sítio-alvo para a endonuclease de restrição EcoRI.

[0059] Uma molécula "exógena" é uma molécula que normalmente não está presente em uma célula, mas pode ser introduzida em uma célula através de um ou mais métodos genéticos, bioquímicos ou outros. A "presença normal na célula" é determinada com relação ao estágio de desenvolvimento e condições ambientais particulares da célula. Assim, por exemplo, uma molécula que está presente apenas durante desenvolvimento embriônico do músculo é uma molécula exógena com relação a uma célula muscular em adultos. Similarmente, uma molécula induzida por choque térmico é uma molécula exógena com relação a uma célula não induzida por choque térmico. Uma molécula exógena pode compreender, por exemplo, uma sequência de codificação para qualquer polipeptídeo ou fragmento do mesmo, uma versão funcional de uma molécula endógena não funcional ou uma versão não funcional de uma molécula endógena normalmente funcional. Adicionalmente, uma molécula exógena pode compreender uma sequência de codificação de outra espécie que é um ortólogo de um gene endógeno na célula hospedeira.

[0060] Uma molécula exógena pode ser, dentre outras coisas, uma pequena molécula, tal como gerada em um processo químico combinatorial ou uma macromolécula, tal como uma proteína, ácido nucleico, carboidrato, lipídio, glicoproteína, lipoproteína, polissacarídeo, qualquer deri-

vado modificado das moléculas acima ou qualquer complexo compreendendo uma ou mais das moléculas acima. Ácidos nucleicos incluem DNA e RNA, podem ser fita simples ou dupla; podem ser lineares, ramificados ou circulares; e podem ser de qualquer comprimento. Ácidos nucleicos incluem aqueles capazes de formação de duplas, bem como ácidos nucleicos que formam tripletos. Vide, por exemplo, Patentes U.S. Nos. 5.176.996 e 5.422.251. Proteínas incluem, mas não estão limitadas a, proteínas de ligação a DNA, fatores de transcrição, fatores de remodelamento de cromatina, proteínas de ligação a DNA metilado, polimerases, metilases, demetilases, acetilases, deacetilases, quinases, fosfatases, integrases, recombinases, ligases, topoisomerases, girases e helicases.

[0061] Uma molécula exógena pode ser do mesmo tipo de molécula que uma molécula endógena, por exemplo, uma proteína ou ácido nucleico exógeno. Por exemplo, um ácido nucleico exógeno pode compreender um genoma viral infeccioso, um plasmídeo ou episssoma introduzido em uma célula ou um cromossomo que normalmente não está presente na célula. Métodos para a introdução de moléculas exógenas em células são conhecidos por aqueles versados no campo e incluem, mas não estão limitados a, transferência lipídio-mediada (isto é, lipossomas, incluindo lipídios neutros e catiônicos), eletroporação, injeção direta, fusão celular, bombardeamento de partícula, coprecipitação com fosfato de cálcio, transformação de nanopartícula, transferência mediada por DEAE-dextrana e transferência mediada por vetor viral.

[0062] Ao contrário, uma molécula "endógena" é uma que normalmente está presente em uma célula em particular em um estágio de desenvolvimento em particular sob condições ambientais particulares. Por exemplo, um ácido nucleico endógeno pode compreender um cromossomo, o genoma de uma mitocôndria, cloroplasto ou outra organela ou um ácido nucleico episssomal que ocorre naturalmente. Moléculas endógenas adicionais podem incluir proteínas, por exemplo, fatores de transcrição e

enzimas.

[0063] Conforme usado aqui, o termo "produto de um ácido nucleico exógeno" inclui produtos de polinucleotídeo e polipeptídeo, por exemplo, produtos de transcrição (polinucleotídeos, tal como RNA) e produtos de tradução (polipeptídeos).

[0064] Uma molécula de "fusão" é uma molécula na qual duas ou mais moléculas subunitárias são ligadas, por exemplo, covalentemente. As moléculas subunitárias podem ser o mesmo tipo de molécula ou podem ser tipos químicos diferentes de moléculas. Exemplos do primeiro tipo de molécula de fusão incluem, mas não estão limitados a, proteínas de fusão (por exemplo, uma fusão entre um domínio de ligação a DNA de ZFP e um domínio de clivagem) e ácidos nucleicos de fusão (por exemplo, um ácido nucleico que codifica a proteína de fusão descrita supra). Exemplos do segundo tipo de molécula de fusão incluem, mas não estão limitados a, uma fusão entre um ácido nucleico de formação de tripleto e um polipeptídeo e uma fusão entre um ligante da ranhura menor e um ácido nucleico.

[0065] Expressão de uma proteína de fusão em uma célula pode resultar de distribuição da proteína de fusão à célula ou distribuição de um polinucleotídeo que codifica a proteína de fusão a uma célula, em que o polinucleotídeo é transcrito e o transcrito é traduzido, para gerar a proteína de fusão. Trans-junção, clivagem de polipeptídeo e ligação de polipeptídeo também podem estar envolvidos em expressão de uma proteína em uma célula. Métodos para distribuição de polinucleotídeo e polipeptídeo a células são apresentados em alguma parte na presente invenção.

[0066] Um "gene", para fins da presente invenção, inclui uma região de DNA que codifica um produto gênico (vide infra), bem como todas as regiões de DNA as quais regulam a produção do produto gênico, quer tais sequências regulatórias estejam ou não adjacentes a sequências de

codificação e/ou transcritas. Consequentemente, um gene inclui, mas não está necessariamente limitado a, sequências promotoras, terminadores, sequências regulatórias de tradução, tais como sítios de ligação ribossômica e sítios de entrada ribossômica internos, intensificadores, silenciadores, isoladores, elementos limítrofes, origens de replicação, sítios de fixação de matriz e regiões de controle de *locus*.

[0067] "Expressão gênica" se refere à conversão da informação, contida em um gene, em um produto gênico. Um produto gênico pode ser o produto direto de um gene (por exemplo, mRNA, tRNA, rRNA, RNA antissenso, RNA de interferência, ribozima, RNA estrutural ou qualquer outro tipo de RNA) ou uma proteína produzida através de tradução de um mRNA. Produtos gênicos também incluem RNAs os quais são modificados através de processos tais como "capeamento", poliadenilação, metilação e edição e proteínas modificadas, por exemplo, através de metilação, acetilação, fosforilação, ubiquitinação, ADP-ribosilação, miristilação e glicosilação.

[0068] "Modulação" de expressão gênica se refere a uma alteração na atividade de um gene. Modulação de expressão pode incluir, mas não está limitada a, ativação gênica e repressão gênica.

[0069] Células de "planta" incluem, mas não estão limitadas a, células de plantas monocotiledôneas (monocots) ou dicotiledôneas (dicots). Exemplos não limitativos de monocots incluem plantas de cereal, tais como milho, arroz, cevada, aveia, trigo, sorgo, centeio, cana-de-açúcar, abacaxi, cebola, banana e coco. Exemplos não limitativo de dicots incluem tabaco, tomate, girassol, algodão, beterraba sacarínica, batata, alface, melão, soja, canola (colza) e alfafa. Células de planta podem ser de qualquer parte da planta e/ou de qualquer estágio de desenvolvimento de uma planta.

[0070] Uma "região de interesse" é qualquer região de cromatina celular tal como, por exemplo, um gene ou uma sequência de não codifica-

ção dentro ou adjacente a um gene, na qual é desejável ligar uma molécula exógena. Ligação pode ser para fins de clivagem de DNA objetivada e/ou recombinação objetivada. Uma região de interesse pode estar presente em um cromossomo, um episssoma, um genoma organelar (por exemplo, mitocondrial, cloroplasto) ou um genoma viral infeccioso, por exemplo. Uma região de interesse pode estar dentro da região de codificação de um gene, dentro de regiões de não codificação transcritas tais como, por exemplo, sequências líderes, sequências "trailer" ou íntrons ou dentro de regiões não transcritas, quer a montante ou a jusante da região de codificação. Uma região de interesse pode ser tão pequena quanto um único par de nucleotídeo ou até 2.000 pares de nucleotídeo de comprimento ou qualquer valor integral de pares de nucleotídeo.

[0071] Os termos "ligação operativa" e "operativamente ligado" (ou "operativamente ligado" são usados permutavelmente com referência a uma justaposição de dois ou mais componentes (tais como elementos de sequência) na qual os componentes estão dispostos de modo que ambos os componentes funcionam normalmente e permitem a possibilidade de que pelo menos um dos componentes possa mediar uma função que é exercida sobre pelo menos um dos outros componentes. À guisa de ilustração, uma sequência regulatória de transcrição, tal como um promotor, é operativamente ligada a uma sequência de codificação se a sequência regulatória de transcrição controla o nível de transcrição da sequência de codificação em resposta à presença ou ausência de um ou mais fatores regulatórios de transcrição. Uma sequência regulatória de transcrição é, em geral, operativamente ligada em *cis* com uma sequência de codificação, mas não precisa ser diretamente adjacente à mesma. Por exemplo, um intensificador é uma sequência regulatória de transcrição que é operativamente ligada a uma sequência de codificação, mesmo embora eles não sejam contíguos.

[0072] Com relação a polipeptídeos de fusão, o termo "operativamen-

te ligado" pode referir-se ao fato de que cada um dos componentes desempenha a mesma função em ligação ao outro componente conforme seria se não estivessem ligados. Por exemplo, com relação a um polipeptídeo de fusão no qual um domínio de ligação a DNA de ZFP é fundido a um domínio de clivagem, o domínio de ligação a DNA de ZFP e o domínio de clivagem estão em ligação operativa se, no polipeptídeo de fusão, a porção de domínio de ligação a DNA de ZFP é capaz de se ligar a seu sítio-alvo e/ou seu sítio de ligação, enquanto que o domínio de clivagem é capaz de clivar o DNA na proximidade do sítio-alvo.

[0073] Um "fragmento funcional" de uma proteína, polipeptídeo ou ácido nucleico é uma proteína, polipeptídeo ou ácido nucleico cuja sequência não é idêntica à proteína, polipeptídeo ou ácido nucleico de comprimento total, ainda que retenha a mesma função que a proteína, polipeptídeo ou ácido nucleico de comprimento total. Um fragmento funcional pode possuir mais, menos ou o mesmo número de resíduos que a molécula nativa correspondente e/ou pode conter uma ou mais substituições de aminoácido ou nucleotídeo. Métodos para determinação da função de um ácido nucleico (por exemplo, função de codificação, capacidade de hibridização a outro ácido nucleico) são bem-conhecidos no campo. Similarmente, métodos para determinação de função de proteína são bem-conhecidos. Por exemplo, a função de ligação a DNA de um polipeptídeo pode ser determinada, por exemplo, através de ensaios de ligação a filtro, desvio de mobilidade eletroforética ou imunoprecipitação. Clivagem de DNA pode ser avaliada por meio de eletroforese em gel. Vide Ausubel *et al*., supra. A capacidade de uma proteína de interagir com outra proteína pode ser determinada, por exemplo, através de coimunoprecipitação, ensaios com dois híbridos ou complementação, genética e bioquímica. Vide, por exemplo, Fields *et al.* (1989) *Nature* 340: 245-246; Patente U.S. No. 5.585.245 e documento PCT WO 98/44350.

Sítios-Alvo

[0074] Os métodos e composições divulgados incluem proteínas de fusão compreendendo um domínio de clivagem (ou um meio-domínio de clivagem) e um domínio de ligação a DNA (por exemplo, ZFP, meganuclease ou zíper de leucina), nas quais o domínio de ligação a DNA (por exemplo, domínio de dedo de zinco, meganuclease ou zíper de leucina), através de ligação a uma sequência em um *locus* de *Zp15* em uma planta, direciona a atividade do domínio de clivagem (ou meio-domínio de clivagem) para a proximidade da sequência e, consequentemente, induz à clivagem (por exemplo, uma ruptura em fita dupla) em *Zp15*. Conforme apresentado em alguma parte na presente invenção, um domínio de dedo de zinco pode ser manipulado para se ligar a virtualmente qualquer sequência desejada. Consequentemente, um ou mais domínios de ligação a DNA (por exemplo, ZFPs) podem ser manipulados para se ligar a uma ou mais sequências em um gene *Zp15* de planta. Expressão de uma proteína de fusão compreendendo um domínio de ligação a DNA (por exemplo, ZFP) e um domínio de clivagem (ou de duas proteínas de fusão, cada uma compreendendo um domínio de ligação a DNA e um meio-domínio de clivagem), em uma célula, realiza clivagem no gene *Zp15*.

[0075] Seleção de uma sequência em um *Zp15* para ligação através de um domínio de dedo de zinco (por exemplo, um sítio-alvo) pode ser realizada, por exemplo, de acordo com os métodos divulgados na Patente US cocedida No. 6.453.242 (17 de setembro de 2002), a qual também divulga métodos para design de ZFPs que se ligam a uma sequência selecionada. Será evidente para aqueles versados no campo que simples inspeção visual de uma sequência de nucleotídeo também pode ser usada para seleção de um sítio-alvo. Consequentemente, qualquer meio para seleção de sítio-alvo pode ser usado nos métodos descritos aqui.

[0076] Para domínios de ligação a DNA de ZFP, sítios-alvo são, em geral, compostos de uma pluralidade de subsítios-alvo adjacentes. Um subsítio-alvo se refere à sequência (usualmente um tripleto de nucleotí-

deo ou uma quadrupleto de nucleotídeo que pode se sobrepor por um nucleotídeo com uma quadrupleto adjacente) ligada por um dedo de zinco individual. Vide, por exemplo, documento WO 02/077227. Se a fita com a qual uma proteína de dedo de zinco faz a maioria dos contatos é designada como a "fita de reconhecimento primário" de fita-alvo ou "fita de contato primário", algumas proteínas do dedo de zinco se ligam a um tripleto de três bases na fita-alvo e uma quarta base sobre a fita não-alvo. Um sítio-alvo geralmente tem um comprimento de pelo menos 9 nucleotídeos e, consequentemente, é ligado através de um domínio de ligação de dedo de zinco compreendendo pelo menos três dedos de zinco. Contudo, ligação, por exemplo, de um domínio de ligação com 4 dedos a um sítio-alvo de 12 nucleotídeos, um domínio de ligação de 5 dedos a um sítio-alvo de 15 nucleotídeos ou um domínio de ligação de 6 dedos a um sítio-alvo de 18 nucleotídeos também é possível. Conforme será evidente, a ligação de domínios de ligação maiores (por exemplo, 7, 8, 9 ou mais dedos) a sítios-alvo maiores também é possível.

[0077] Não é necessário que um sítio-alvo seja um múltiplo de três nucleotídeos. Por exemplo, em casos nos quais interações fita-cruzada ocorrem (vide, por exemplo, Patente U.S. No. 6.453.242 e documento WO 02/077227), um ou mais dos dedos de zinco individuais de um domínio de ligação a múltiplos dedos podem se ligar a subsítios de um quadrupleto em sobreposição. Como um resultado, uma proteína com três dedos pode se ligar a uma sequência de 10 nucleotídeos, em que o décimo nucleotídeo é parte de um quadrupleto ligado por um dedo terminal, uma proteína com quatro dedos pode se ligar a uma sequência de 13 nucleotídeos, em que o décimo terceiro nucleotídeo é parte de um quadrupleto ligado por um dedo terminal, etc.

[0078] O comprimento e natureza de sequências ligantes de aminoácido entre dedos de zinco individuais em um domínio de ligação a múltiplos dedos também afetam a ligação a uma sequência-alvo. Por exemplo,

a presença de um assim denominado "ligante não canônico", "ligante longo" ou "ligante estruturado" entre dedos de zinco adjacentes em um domínio de ligação a múltiplos dedos pode permitir que esses dedos se liguem a subsítios os quais não são imediatamente adjacentes. Exemplos não limitativos de tais ligantes são descritos, por exemplo, na Patente U.S. No. 6.479.626 e documento WO 01/53480. Consequentemente, um ou mais subsítios, em um sítio-alvo para um domínio de ligação a dedo de zinco, podem ser separados uns dos outros por 1, 2, 3, 4, 5 ou mais nucleotídeos. Para proporcionar mais um exemplo, um domínio de ligação de quatro dedos pode se ligar a um sítio-alvo de 13 nucleotídeos compreendendo, em sequência, dois subsítios de 3 nucleotídeos contíguos, um nucleotídeo interveniente e dois subsítios de tripleto contíguos.

[0079] A distância entre sequências (por exemplo, sítios-alvo) se refere ao número de nucleotídeos ou pares de nucleotídeo intervenientes entre duas sequências, conforme medido a partir das bordas das sequências mais próximas umas das outras.

[0080] Em determinadas modalidades nas quais a clivagem depende da ligação de duas moléculas de fusão de domínio de dedo de zinco/meio-domínio de clivagem a sítios-alvo distintos, os dois sítios-alvo podem estar sobre fitas de DNA opostas. Em outras modalidades, ambos os sítios-alvo estão sobre a mesma fita de DNA.

Domínios de Ligação a DNA

[0081] Qualquer domínio de ligação a DNA pode ser usado nos métodos divulgados aqui. Em determinadas modalidades, o domínio de ligação a DNA compreende uma proteína do dedo de zinco. Um domínio de ligação de dedo de zinco compreende um ou mais dedos de zinco. Miller *et al.* (1985) *EMBOJ.* 4: 1609-1614; Rhodes (1993) *Scientific American* Fev.: 56-65; Patente U.S. No. 6.453.242. Os domínios de ligação de dedo de zinco descritos geralmente incluem 2, 3, 4, 5, 6 ou mesmo mais dedos de zinco.

[0082] Tipicamente, um único domínio de dedo de zinco tem cerca de 30 aminoácidos de comprimento. Estudos estruturais demonstraram que cada domínio de dedo de zinco (motivo) contém duas beta folhas (mantida em um beta giro o qual contém os dois resíduos de cisteína invariáveis) e uma alfa hélice (contendo os dois resíduos de histidina invariáveis), os quais são mantidos em uma conformação particular através de coordenação de um átomo de zinco pelas duas cisteínas e as duas histidinas.

[0083] Dedos de zinco incluem dedos de zinco C_2H_2 canônicos (isto é, aqueles nos quais o íon de zinco é coordenado por dois resíduos de cisteína e dois de histidina) e dedos de zinco não canônicos tais como, por exemplo, dedos de zinco C_3H (aqueles nos quais o íon de zinco é coordenado por três resíduos de cisteína e um resíduo de histidina) e dedos de zinco C_4 (aqueles nos quais o íon de zinco é coordenado por quatro resíduos de cisteína). Vide também documento WO 02/057293 e também Publicação de Patente U.S. No. 20080182332 com relação a ZFPs não canônicas para uso em plantas.

[0084] Um domínio de ligação de dedo de zinco manipulado pode ter uma nova especificidade de ligação, comparado com uma proteína de dedo de zinco que ocorre naturalmente. Métodos de manipulação incluem, mas não estão limitados a, design racional e vários tipos de seleção. Design racional inclui, por exemplo, uso de bancos de dados compreendendo sequências de nucleotídeo tripletos (ou quadrupletos) e sequências de aminoácido de dedo de zinco individuais, nas quais cada sequência de nucleotídeo tripleto ou quadrupleto está associada a uma ou mais sequências de aminoácido de dedos de zinco as quais se ligam à sequência tripleto ou quadrupleto em particular.

[0085] Métodos de seleção exemplificativos, incluindo exibição de fago (“*phage display”)* e sistemas com dois híbridos, são divulgados nas Patentes U.S. Nos. 5.789.538; 5.925.523; 6.007.988; 6.013.453;

6.410.248; 6.140.466; 6.200.759; e 6.242.568; bem como documentos WO 98/37186; WO 98/53057; WO 00/27878; WO 01/88197 e GB 2.338.237.

[0086] Intensificação de especificidade de ligação para domínios de dedo de zinco foi descrita, por exemplo, no documento concedido WO 02/077227.

[0087] Uma vez que um dedo de zinco individual se liga a uma sequência de três nucleotídeos (isto é, tripleto) (ou uma sequência de quatro nucleotídeos a qual pode se sobrepor, por um nucleotídeo, com o sítio de ligação de quatro nucleotídeos de um dedo de zinco adjacente), o comprimento de uma sequência à qual um domínio de ligação de dedo de zinco se liga (por exemplo, uma sequência-alvo) determinará o número de dedos de zinco em um domínio de ligação de dedo de zinco manipulado. Por exemplo, para ZFPs nas quais os motivos de dedo não se ligam a subsítios em sobreposição, uma sequência-alvo de seis nucleotídeos é ligada por domínios de ligação com dois dedos; uma sequência-alvo de nove nucleotídeos é ligada por um domínio de ligação com três dedos; uma sequência-alvo de seis nucleotídeos é ligada por um domínio de ligação com dois dedos; uma sequência-alvo de nove nucleotídeos é ligada por um domínio de ligação com três dedos, etc. Conforme mencionado aqui, sítios de ligação para dedos de zinco individuais (isto é, subsítios) em um sítio-alvo não precisam ser contíguos, mas podem ser separados por um ou vários nucleotídeos, dependendo do comprimento e natureza de sequência de aminoácido entre os dedos de zinco (isto é, os ligantes interdedo) em um domínio de ligação de múltiplos dedos.

[0088] Em um domínio de ligação de dedo de zinco a múltiplos dedos, dedos de zinco adjacentes podem ser separados por sequências ligantes de aminoácido de aproximadamente 5 aminoácidos (os assim denominados ligantes interdedo "canônicos") ou, alternativamente, por um ou mais ligantes não canônicos. Vide, por exemplo, Patentes U.S.

concedidas Nos. 6.453.242 e 6.534.261. Para domínios de dedo de zinco manipulados compreendendo mais de três dedos, inserção de ligantes interdedo mais longos (não canônicos) entre alguns dos dedos de zinco pode ser desejável em alguns casos, uma vez que isso pode aumentar a afinidade e/ou especificidade de ligação pelo domínio de ligação. Vide, por exemplo, Patente U.S. No. 6.479.626 e documento WO 01/53480. Consequentemente, domínios de ligação de dedo de zinco com múltiplos dedos podem também ser caracterizados com relação à presença e localização de ligantes interdedo não canônicos. Por exemplo, a um domínio de ligação de dedo de zinco com seis dedos compreendendo três dedos (unidos por dois ligantes interdedo canônicos), um ligante longo e três dedos adicionais (unidos por dois ligantes interdedo canônicos) são denotados uma configuração 2x3. Similarmente, a um domínio de ligação compreendendo dois dedos (com um ligante canônico entre os mesmos), um ligante longo e dois dedos adicionais (unidos por um ligante canônico) são denotados uma configuração 2x2. Uma proteína compreendendo três unidades com dois dedos (em cada uma das quais os dois dedos são unidos por um ligante canônico) e na qual cada unidade com dois dedos é unida à unidade com dois dedos adjacentes por um ligante longo é referida como uma configuração 3x2.

[0089] A presença de um ligante interdedo longo ou não canônico entre dois dedos de zinco adjacentes em um domínio de ligação a múltiplos dedos frequentemente permite que os dois dedos se liguem a subsítios os quais não são imediatamente contíguos na sequência-alvo. Consequentemente, podem haver vãos de um ou mais nucleotídeos entre subsítios em um sítio-alvo; isto é, um sítio-alvo pode conter um ou mais nucleotídeos que não são contatados por um dedo de zinco. Por exemplo, um domínio de ligação de dedo de zinco 2x2 pode se ligar a duas sequências com seis nucleotídeos separadas por um nucleotídeo, isto é, ele se liga a um sítio-alvo de 13 nucleotídeos. Vide também Moore *et al.*

(2001a) Proc. Natl. Acad. Sci. USA 98: 1432-1436; Moore *et al.* (2001b) Proc. Natl. Acad. Sci. USA 98: 1437-1441 e documento WO 01/53480.

[0090] Conforme mencionado anteriormente, um subsítio-alvo é uma sequência de três ou quatro nucleotídeos que é ligada por um único dedo de zinco. Para determinadas finalidades, uma unidade com dois dedos é denotada um "módulo de ligação". Um módulo de ligação pode ser obtido, por exemplo, selecionando dois dedos adjacentes no contexto de uma proteína de múltiplos dedos (geralmente três dedos) a qual se liga a uma sequência-alvo de seis nucleotídeos em particular. Alternativamente, módulos podem ser construídos por meio de montagem de dedos de zinco individuais. Vide também documentos WO 98/53057 e WO 01/53480.

[0091] Alternativamente, o domínio de ligação a DNA pode ser derivado de uma nuclease. Por exemplo, as sequências de reconhecimento das endonucleases e meganucleases "homing", tais como I-*Sce*I, I-*Ceu*I, VI-*Psp*I, PI-*Sce*, I-*Sce*IV, I-*Csm*I, I-*Pan*I, I-*Sce*II, I-*Ppo*I, I-*Sce*III, I-*Cre*I, I-*Tev*I, I-*Tev*II e I-*Tev*III são conhecidas. Vide também Patente U.S. No. 5.420.032; Patente U.S. No. 6.833.252; Belfort *et al*. (1997) *Nucleic Acids Res.* 25: 3379-3388; Dujon *et al.* (1989) *Gene* 82: 115-118; *Verier et al.* (1994) Nucleic Acids Res. 22, 1125-1127; Jasin (1996) *Trends Genet.* 12: 224-228; Gimble *et al.* (1996) *J. Mol. Biol.* 263: 163-180; Argast *et al.* (1998) *J. Mol. Biol.* 280: 345-353 e o catálogo da New England Biolabs. Além disso, a especificidade de ligação a DNA de endonucleases e meganucleases "homing" pode ser manipulada para se ligar a sítios-alvo não naturais. Vide, por exemplo, Chevalier *et al.* (2002) *Molec. Cell* 10: 895-905; Epinat *et al.* (2003) *Nucleic Acids Res.* 31: 2952-2962; Ashworth *et al.* (2006) *Nature* 441: 656-659; Paques *et al.* (2007) *Current Gene Therapy* 7: 49-66; Publicação de Patente U.S. No. 20070117128.

[0092] Como outra alternativa, o domínio de ligação a DNA pode ser derivado de uma proteína de zíper de leucina. Zíperes de leucina são uma classe de proteínas que estão envolvidas em interações proteína-

proteína em muitas proteínas regulatórias eucariotas que são fatores transcricionais importantes associados à expressão gênica. O zíper de leucina se refere a um motivo estrutural em comum compartilhado nesses fatores de transcrição através de vários reinos, incluindo animais, plantas, leveduras, etc. O zíper de leucina é formado por dois polipeptídeos (homodímero ou heterodímero) que se ligam a sequências de DNA específicas de uma maneira onde os resíduos de leucina são uniformemente espaçados através de uma α-hélice, de modo que os resíduos de leucina dos dois polipeptídeos terminam sobre a mesma face da hélice. A especificidade de ligação a DNA de zíperes de leucina pode ser utilizada nos domínios de ligação a DNA divulgados aqui.

[0093] Em algumas modalidades, o domínio de ligação a DNA é um domínio manipulado a partir de um efetuador TAL derivado do patógeno de planta *Xanthomonas* (vide Boch *et al,* (2009) *Science*, 29 de outubro de 2009 (10.1126/science.117881) e Moscou e Bogdanove, (2009) *Science*, 29 de outubro de 2009 (10.1126/science.1178817).

Domínios de Clivagem

[0094] Conforme mencionado acima, o domínio de ligação a DNA pode estar associado a um domínio de clivagem (nuclease). Por exemplo, endonucleases "homing" podem ser modificadas quanto à sua especificidade de ligação a DNA, ao mesmo tempo em que retêm a função de nuclease. Além disso, proteínas do dedo de zinco também podem ser fundidas a um domínio de clivagem para formar uma nuclease de dedo de zinco (ZFN). A porção de domínio de clivagem das proteínas de fusão divulgadas aqui pode ser obtida a partir de qualquer endonuclease ou exonuclease. Endonucleases exemplificativas das quais um domínio de clivagem pode ser derivado incluem, mas não estão limitadas a, endonucleases de restrição e endonucleases "homing". Vide, por exemplo, Catálogo de 2002-2003, New England Biolabs, Beverly, MA; e Belfort *et al.* (1997) *Nucleic Acids Res.* 25: 3379-3388. Enzimas adicionais as quais clivam

DNA são conhecidas (por exemplo, Nuclease SI; nuclease de feijão-da-China; DNase I pancreática; nuclease micrococica; endonuclease HO de levedura; vide também Linn *et al.* (eds.) *Nucleases,* Cold Spring Harbor Laboratory Press, 1993). Exemplos não limitativos de endonucleases e meganucleases "homing" incluem I-*Sce*I, I-*Ceu*I, VI-*Psp*I, PI-*Sce*, I-*Sce*IV, I-*Csm*I, I-*Pan*I, I-*Sce*II, I-*Ppo*I, I-*Sce*III, I-*Cre*I, I-*Tev*I, I-*Tev*II e I-*Tev*III são conhecidos. Vide também Patente U.S. No. 5.420.032; Patente U.S. No. 6.833.252; Belfort *et al.* (1997) *Nucleic Acids Res.* 25: 3379-3388; Dujon *et al.* (1989) *Gene* 82: 115-118; Perler *et al*. (1994) *Nucleic Acids Res*. 22, 1125-1127; Jasin (1996) *Trends Genet.* 12: 224-228; Gimble *et al.* (1996) *J. Mol. Biol.* 263: 163-180; Argast *et al.* (1998) *J. Mol. Biol.* 280: 345-353 e o catálogo da New England Biolabs. Uma ou mais dessas enzimas (ou fragmentos funcionais das mesmas) podem ser usadas como uma fonte de domínios de clivagem e meio-domínios de clivagem.

[0095] Endonucleases de restrição (enzimas de restrição) estão presentes em muitas espécies e são capazes de ligação sequência-específica ao DNA (em um sítio de reconhecimento) e clivagem de DNA em ou próximo do sítio de ligação. Determinadas enzimas de restrição (por exemplo, Tipo ISS) clivam o DNA em sítios removidos do sítio de reconhecimento e têm domínios de ligação e clivagem separáveis. Por exemplo, a enzima *FokI* do tipo IIS catalisa a clivagem de DNA fita dupla em 9 nucleotídeos a partir de seu sítio de reconhecimento sobre uma fita e 13 nucleotídeos a partir de seu sítio de reconhecimento sobre a outra. Vide, por exemplo, Patentes US 5.356.802; 5.436.150 e 5.487.994; bem como Li *et al.* (1992) *Proc. Natl. Acad. Sci. USA* 89: 4275-4279; Li *et al.* (1993) *Proc. Natl. Acad. Sci. USA* 90: 2764-2768; Kim *et al.* (1994a) *Proc. Natl. Acad. Sci. USA* 91: 883-887; Kim *et al.* (1994b) *J. Biol. Chem.* 269: 31, 978-31, 982. Assim, em uma modalidade, proteínas de fusão compreendem o domínio de clivagem (ou meio-domínio de clivagem) de pelo menos uma enzima de restrição do Tipo IIS e um ou mais domínios de

ligação a dedo de zinco, os quais podem ou não ser manipulados.

[0096] Uma enzima de restrição do Tipo IIS exemplificativa, cujo domínio de clivagem é separável do domínio de ligação é *Fok*I. Essa enzima em particular é ativa como um dímero. Bitinaite *et al.* (1998) *Proc. Natl. Acad. Sci. USA* 95: 10.570-10.575. Consequentemente, para fins da presente invenção, a porção da enzima *Fok*I usada nas proteínas de fusão divulgadas é considerada um meio-domínio de clivagem. Assim, para clivagem de fita dupla objetivada e/ou substituição objetivada de sequências celulares usando fusões de dedo de zinco-*Fok*I, duas proteínas de fusão, cada uma compreendendo um meio-domínio de clivagem de *Fok*I, podem ser usadas para reconstituir um domínio de clivagem catalitica-mente ativo. Alternativamente, uma única molécula polipeptídica contendo um domínio de ligação de dedo de zinco e dois meio-domínios de clivagem de *Fok*I também pode ser usada. Parâmetros para clivagem objetivada e alteração de sequência objetivada usando fusões de dedo de zinco-*Fok*I são proporcionados em alguma parte na presente invenção.

[0097] Um domínio de clivagem ou meio-domínio de clivagem pode ser qualquer porção de uma proteína que retém a atividade de clivagem ou que retém a capacidade de multimerizar (por exemplo, dimerizar) para formar um domínio de clivagem funcional.

[0098] Enzimas de restrição do Tipo IIS exemplificativas são descritas na Publicação Internacional cocedida WO 2007/014275, incorporada por referência aqui na íntegra.

[0099] Para intensificar a especificidade de clivagem, domínios de clivagem também podem ser modificados. Em determinadas modalidades, variantes do meio-domínio de clivagem são empregadas onde essas variantes minimizam ou previnem a homodimerização dos meio-domínios de clivagem. Exemplos não limitativos de tais meio-domínios de clivagem modificados são descritos em detalhes no documento WO 2007/014275,

aqui incorporado por referência na íntegra. Vide, também, Exemplos. Em determinadas modalidades, o domínio de clivagem compreende um meio-domínio de clivagem manipulado (também referido como mutantes com domínio de dimerização) que minimizam ou previnem a homodimerização são conhecidos por aqueles versados no campo e descritos, por exemplo, nas Publicações de Patente U.S. Nos. 20050064474 e 20060188987, incorporadas por referência aqui na íntegra. Resíduos de aminoácido nas posições 446, 447, 479, 483, 484, 486, 487, 490, 491, 496, 498, 499, 500, 531, 534, 537 e 538 de *Fok*I são todos alvo para influenciar a dimerização dos meio-domínios de clivagem de *Fok*I. Vide, por exemplo, Publicações de Patente U.S. Nos. 20050064474 e 20060188987; Publicação de Patente Internacional WO 07/139898; Miller *et al.* (2007) *Nat. Biotechnol.* 25(7): 778-785.

[00100] Meio-domínios de clivagem de *Fok*I manipulados adicionais que formam heterodímeros obrigatórios também podem ser usados nas ZFNs descritas aqui. Em uma modalidade, o primeiro meio-domínio de clivagem inclui mutações em resíduos de aminoácido nas posições 490 e 528 de *Fok*I e o segundo meio-domínio de clivagem inclui mutações em resíduos de aminoácido 486 e 499.

[00101] Em determinadas modalidades, o domínio de clivagem compreende dois meio-domínios de clivagem, ambos os quais são parte de um único polipeptídeo compreendendo um domínio de ligação, um primeiro meio-domínio de clivagem e um segundo meio-domínio de clivagem. Os meio-domínios de clivagem podem ter a mesma sequência de aminoácido ou diferentes sequências de aminoácido, na medida em que eles funcionem para clivar o DNA.

[00102] Em geral, duas proteínas de fusão são requeridas para clivagem se as proteínas de fusão compreendem meio-domínios de clivagem. Alternativamente, uma única proteína compreendendo dois meio-domínios de clivagem pode ser usada. Os dois meio-domínios de cliva-

gem podem ser derivados da mesma endonuclease (ou fragmentos funcionais da mesma) ou cada meio-domínio de clivagem pode ser derivado de uma endonuclease diferente (ou fragmentos funcionais das mesmas). Além disso, os sítios-alvo para as duas proteínas de fusão são, de preferência, dispostos, um com relação ao outro, de modo que ligação das duas proteínas de fusão a seus respectivos sítios-alvo coloca os meio-domínios de clivagem em uma orientação espacial um com relação ao outro, de modo a permitir que a clivagem dos meio-domínios forme um domínio de clivagem funcional, por exemplo, através de dimerização. Assim, em determinadas modalidades, as bordas próximas dos sítios-alvo são separadas por 5-8 nucleotídeos ou por 15-18 nucleotídeos. Contudo, qualquer número integral de nucleotídeos ou pares de nucleotídeo podem estar intervenientes entre dois sítios-alvo (por exemplo, de 2 a 50 nucleotídeos ou mais). Em geral, o ponto de clivagem reside entre os sítios-alvo.

Proteínas de Fusão

[00103] Métodos para design e construção de proteínas de fusão (e polinucleotídeos que codificam as mesmas) são conhecidos por aqueles versados no campo. Por exemplo, métodos para o design e construção de proteínas de fusão compreendendo domínios de ligação de DNA (por exemplo, domínios de dedo de zinco) e domínios regulatórios ou de clivagem (ou meio-domínios de clivagem) e polinucleotídeos que codificam tais proteínas de fusão são descritos nas Patentes U.S. cocedidas Nos. 6.453.242 e 6.534.261 e Publicações de Pedido de Patente U.S. 2007/0134796 e 2005/0064474; aqui incorporados por referência na íntegra. Em determinadas modalidades, polinucleotídeos que codificam as proteínas de fusão são construídos. Esses polinucleotídeos podem ser inseridos em um vetor e o vetor pode ser introduzido em uma célula (vide abaixo para divulgação adicional com relação a vetores e métodos para introdução de polinucleotídeos em células).

[00104] Em determinadas modalidades dos métodos descritos aqui,

uma nuclease de dedo de zinco compreende uma proteína de fusão compreendendo um domínio de ligação de dedo de zinco e um meio-domínio de clivagem da enzima de restrição *Fok*I e duas de tais proteínas de fusão são expressas em uma célula. Expressão de duas proteínas de fusão em uma célula pode resultar de distribuição das duas proteínas à célula; distribuição de uma proteína e um ácido nucleico que codifica uma das proteínas à célula; distribuição de dois ácidos nucleicos, cada um codificando uma das proteínas, à célula; ou através de distribuição de um único ácido nucleico, que codifica ambas as proteínas, à célula. Em modalidades adicionais, uma proteína de fusão compreende uma única cadeia polipeptídica compreendendo dois meio-domínios de clivagem e um domínio de ligação de dedo de zinco. Nesse caso, uma única proteína de fusão é expressa em uma célula e, sem desejar estar preso pela teoria, acredita-se que clive o DNA como um resultado de formação de um dímero intramolecular dos meio-domínios de clivagem.

[00105] Em determinadas modalidades, os componentes das proteínas de fusão (por exemplo, fusões de ZFP-*Fok*I) são dispostos de modo que o domínio de dedo de zinco esteja mais próximo do amino término da proteína de fusão e o meio-domínio de clivagem esteja mais próximo do carbóxi término. Isso espelha a orientação relativa do domínio de clivagem em domínios de clivagem de dimerização que ocorrem naturalmente, tais como aqueles derivados da enzima *Fok*I, nos quais o domínio de ligação a DNA está mais próximo do amino término e o meio-domínio de clivagem está mais próximo do carbóxi término. Nessas modalidades, acredita-se que dimerização dos meio-domínios de clivagem para formar uma nuclease funcional ocorra mediante ligação das proteínas de fusão a sítios sobre fitas de DNA opostas, com as extremidades 5' dos sítios de ligação estão próximas umas das outras.

[00106] Em modalidades adicionais, os componentes das proteínas de fusão (por exemplo, fusões ZFN-*Fok*I) são dispostos de modo que o

meio-domínio de clivagem esteja mais próximo do amino término da proteína de fusão e o domínio de dedo de zinco esteja mais próximo do carbóxi término. Nessas modalidades, acredita-se que dimerização dos meio-domínios de clivagem forme uma nuclease funcional através de ligação das proteínas de fusão a sítios sobre fitas de DNA opostas, com as extremidades 3' dos sítios de ligação estando próximas uma da outra.

[00107] Em outras modalidades adicionais, uma primeira proteína de fusão contém o meio-domínio de clivagem mais próximo do amino término da proteína de fusão e o domínio de dedo de zinco mais próximo do carbóxi término e uma segunda proteína de fusão está disposta de modo que o domínio de dedo de zinco esteja mais próximo do amino término da proteína de fusão e o meio-domínio de clivagem esteja mais próximo do carbóxi término. Nessas modalidades, ambas as proteínas de fusão se ligam à mesma fita de DNA, com o sítio de ligação da primeira proteína de fusão contendo o domínio de dedo de zinco mais próximo do carbóxi término localizado no lado 5' do sítio de ligação da segunda proteína de fusão contendo o domínio de dedo de zinco mais próximo do amino término.

[00108] Em determinadas modalidades das proteínas de fusão divulgadas, a sequência de aminoácido entre o domínio de dedo de zinco e o domínio de clivagem (ou meio-domínio de clivagem) é denotada o "ligante ZC". O ligante ZC deve ser distinguido dos ligantes interdedo discutidos acima. Vide, por exemplo, Publicações de Patente U.S. 20050064474A1 e 20030232410 e Publicação de Patente Internacional WO05/084190 para detalhes sobre a obtenção de ligantes ZC que otimizam a clivagem.

[00109] Em uma modalidade, a invenção proporciona uma ZFN compreendendo uma proteína de dedo de zinco tendo uma ou mais das sequências de aminoácido de hélice de reconhecimento mostradas na Tabela 1. Em outra modalidade, é proporcionado aqui um vetor de expres-

são de ZFP compreendendo uma sequência de nucleotídeo que codifica uma ZFP tendo uma ou mais hélices de reconhecimento mostradas na Tabela 1.

Integração Objetivada

[00110] Os métodos e composições divulgados podem ser usados para clivar DNA em um gene *Zp15* da cromatina celular de planta, o que facilita a integração objetivada estável de uma sequência exógena no *locus*. Conforme descrito aqui, perda de função de genes *Zp15* endógenos é bem tolerada por células de planta e sequências integradas dentro desse gene são amplamente descritas e geram plantas com modificações na linhagem germinativa para transmissão hereditária da sequência integrada. Consequentemente, o *Zp15* é um sítio desejável para integração objetivada de sequências exógenas.

[00111] Para integração objetivada em *Zp15*, um ou mais domínios de ligação a DNA (por exemplo, ZFPs) são manipulados para se ligar a um sítio-alvo em ou próximo do sítio de clivagem predeterminado e uma proteína de fusão compreendendo o domínio de ligação a DNA manipulado e um domínio de clivagem é expresso em uma célula. Quando de ligação da porção de ligação a DNA (por exemplo, dedo de zinco) da proteína de fusão ao sítio-alvo, o DNA é clivado, de preferência via uma ruptura de fita dupla, próximo do sítio-alvo pelo domínio de clivagem.

[00112] A presença de uma ruptura de fita dupla no locus de *Zp15* facilita a integração de sequências exógenas via recombinação homóloga. Assim, o polinucleotídeo compreendendo a sequência exógena a ser inserida no gene *Zp15* incluirá uma ou mais regiões de homologia com um gene *Zp15* para facilitar a recombinação homóloga.

[00113] Qualquer sequência de interesse (sequência exógena) pode ser introduzida em um *locus* de *Zp15* conforme descrito aqui. Sequências exógenas exemplificativas incluem, mas não estão limitadas a, qualquer sequência de codificação polipeptídica (por exemplo, cDNAs),

promotores, intensificadores e outras sequências regulatórias (por exemplo, sequências de RNA de interferência, cassetes de expressão de shRNA, tags de epitopo, genes marcadores, sítios de reconhecimento de enzima de clivagem, sítios e vários tipos de construtos de expressão. Tais sequências podem ser prontamente obtidas usando técnicas de biologia molecular padrões (clonagem, síntese, etc.) e/ou estão comercialmente disponíveis.

[00114] Além das moléculas de fusão descritas aqui, substituição objetivada de uma sequência genômica selecionada também envolve a introdução da sequência de substituição (ou doadora). A sequência doadora pode ser introduzida na célula antes de, concorrentemente com ou subsequentemente à expressão da(s) proteína(s) de fusão. O polinucleotídeo doador contém homologia suficiente ao *Zp15* para sustentar recombinação homóloga (ou reparo homologia-dirigido) entre o mesmo e a sequência genômica de *Zp15* com a qual ele tem homologia. Aproximadamente 25, 50, 100, 200, 500, 750, 1.000, 1.500, 2.000 nucleotídeos ou mais de homologia de sequência entre uma sequência doadora e uma genômica (ou qualquer valor integral entre 10 e 2.000 nucleotídeos ou mais) sustentarão a recombinação homóloga entre as mesmas. Em determinadas modalidades, os braços de homologia têm menos de 1.000 pares de base de comprimento. Em outras modalidades, os braços de homologia têm menos de 750 pares de base de comprimento. Vide também Pedido de Patente Provisória U.S. No. 61/124.047, a qual é incorporada aqui por referência.

[00115] Sequências doadoras podem oscilar, quanto ao comprimento, de 10 a 5.000 nucleotídeos (ou qualquer valor integral de nucleotídeos entre os mesmos) ou mais. Será prontamente evidente que a sequência doadora, tipicamente, não é idêntica à sequência genômica que ela substitui. Por exemplo, a sequência do polinucleotídeo doador pode conter uma ou mais alterações, inserções, deleções, inversões ou reestrutura-

ções de uma base com relação à sequência genômica, contanto que homologia suficiente com as sequências cromossômicas esteja presente. Alternativamente, uma sequência doadora pode conter uma sequência não homóloga flanqueada por duas regiões de homologia. Adicionalmente, sequências doadoras podem compreender uma molécula de vetor contendo sequências que não são homólogas à região de interesse na cromatina celular. Geralmente, a(s) região(ões) homóloga(s) de uma sequência doadora terá(ão) pelo menos 50% de identidade de sequência a uma sequência genômica com a qual recombinação é desejada. Em determinadas modalidades, 60%, 70%, 80%, 90%, 95%, 98%, 99% ou 99,9% de identidade de sequência estão presentes. Qualquer valor entre 1% e 100% de identidade de sequência pode estar presente, dependendo do comprimento do polinucleotídeo doador.

[00116] Uma molécula doadora pode conter várias regiões descontinuas de homologia à cromatina celular. Por exemplo, para inserção objetivada de sequências normalmente não presentes em uma região de interesse, as referidas sequências podem estar presentes em uma molécula de ácido nucleico e flanqueada por regiões de homologia a uma sequência gênica na região de interesse.

[00117] Moléculas doadoras também podem ser inseridas no *locus* de *Zp15* para servir como um reservatório para uso posterior. Por exemplo, uma molécula doadora homóloga a um gene endógeno, mas contendo uma mutação de interesse, pode ser inserida no *locus* de *Zp15*. Em seguida, ZFNs específicas ao gene endógeno podem ser introduzidas, as quais clivarão o *locus* endógeno e a molécula doadora no *locus* de *Zp15* o qual contém a mutação de interesse. O DSB resultante no genoma pode, então, se tornar o sítio de integração para a molécula doadora liberada do *locus* de *Zp15*. Dessa forma, a eficiência de integração objetivada de uma sequência doadora em qualquer região de interesse pode ser grandemente aumentada, uma vez que o método não conta com capta-

ção simultânea de ácidos nucleicos que codificam as ZFNs e aqueles das sequências doadoras.

[00118] Moléculas doadoras também podem ser inseridas no *locus* de *Zp15* para servir como um sítio-alvo para subsequentes inserções. Por exemplo, uma molécula doadora compreendida de sequências de DNA que contêm sítios de reconhecimento para designs de ZFN adicionais pode ser inserida no *locus* de *Zp15*. Subsequentemente, designs de ZFN adicionais podem ser gerados e expressos em células, de modo que a molécula doadora original seja clivada e modificada através de reparo ou recombinação homóloga. Dessa forma, integrações reiterativas de moléculas doadoras podem ocorrer no *locus* de *Zp15*.

[00119] Para simplificar os ensaios (por exemplo, hibridização, PCR, digestão com enzima de restrição) para determinação de inserção com sucesso da sequência doadora, determinadas diferenças de sequência podem estar presentes na sequência doadora quando comparado com a sequência genômica de *Zp15*. De preferência, se localizadas em uma região de codificação, tais diferenças de sequência de nucleotídeo não alterarão a sequência de aminoácido ou farão alterações de aminoácido silenciosas (isto é, alterações as quais não afetam a estrutura ou função da proteína). O polinucleotídeo doador pode, opcionalmente, conter alterações em sequências correspondendo aos sítios de ligação com domínio de ligação a DNA na região de interesse para impedir clivagem de sequências doadoras que tenham sido introduzidas na cromatina celular por meio de recombinação homóloga.

[00120] O polinucleotídeo doador pode ser DNA ou RNA, fita simples ou fita dupla e pode ser introduzido em uma célula na forma linear ou circular. Se introduzido na forma linear, as extremidades da sequência doadora podem ser protegidas (por exemplo, contra degradação exonucleolítica) através de métodos conhecidos por aqueles versados no campo. Por exemplo, um ou mais resíduos de dideoxinucleotídeo são adiciona-

dos ao 3' término de uma molécula linear e/ou oligonucleotídeos auto-complementares são ligados a uma ou ambas as extremidades. Vide, por exemplo, Chang *et al.* (1987) *Proc. Natl. Acad. Sci. USA* 84: 4959-4963; Nehls *et al*. (1996) *Science* 272: 886-889. Métodos adicionais para proteção de polinucleotídeos exógenos contra degradação incluem, mas não estão limitados a, adição de grupo(s) amino terminal(is) e o uso de ligações internucleotídeo modificadas tais como, por exemplo, resíduos de fosforotioatos, fosforamidatos e O-metil ribose ou deoxirribose.

[00121] Um polinucleotídeo pode ser introduzido em uma célula como parte de uma molécula de vetor tendo sequências adicionais tais como, por exemplo, origens de replicação, promotores e genes que codificam resistência a antibiótico. Além disso, os polinucleotídeos doadores podem ser introduzidos como ácido nucleico "nu", como um ácido nucleico em complexo com um agente, tal como uma nanopartícula, lipossoma ou poloxâmero ou pode ser distribuído por bactérias ou vírus (por exemplo, *Agrobacterium, Rhizobium sp.* NGR234, *Sinorhizoboium meliloti, Mesorhizobium loti,* vírus mosaico do tabaco, vírus X da batata, vírus mosaico da couve-flor e vírus mosaico do veio de cassava). Vide, por exemplo, Chung *et al.* (2006) *Trends Plant Sci.* 11(1): 1-4.

[00122] Parece que a presença de uma ruptura de fita dupla em uma sequência celular, acoplada à presença de uma molécula de DNA exógena tendo homologia a uma região adjacente a ou que envolve a ruptura, ativa mecanismos celulares os quais reparam a ruptura por meio de transferência de informação de sequência da molécula doadora para a sequência celular (por exemplo, genômica ou cromossômica); isto é, através de um processo de reparo homologia-dirigido, também conhecido como "conversão gênica". Os métodos da Requerente combinam, vantajosamente, as poderosas capacidades de objetivação de ZFPs manipuladas com um domínio de clivagem (ou meio-domínio de clivagem) para objetivar especificamente genes parálogos, tais como genes *Zp15*, de

modo que clivagem da sequência-alvo produz uma ruptura de fita dupla na região do genoma onde inserção de sequências exógenas é desejada.

[00123] Para alteração de uma sequência cromossômica, não é necessário que toda a sequência do doador seja copiada para o cromossomo, contanto que o bastante da sequência doadora seja copiado para realizar a alteração de sequência desejada.

[00124] A eficiência de inserção de sequências doadoras através de recombinação homóloga é inversamente relacionada à distância, no DNA celular, entre a ruptura de fita dupla e o sítio no qual recombinação é desejada. Em outras palavras, maiores eficiências de recombinação homóloga são observadas quando a ruptura de fita dupla está mais próxima do sítio no qual recombinação é desejada. Em casos nos quais um sítio preciso de recombinação não é predeterminado (por exemplo, o evento de recombinação desejado pode ocorrer sobre um intervalo de sequência genômica), o comprimento e sequência do ácido nucleico doador, junto com o(s) sítio(s) de clivagem, são selecionados para obter o evento de recombinação desejado. Em casos nos quais o evento desejado é projetado para alterar a sequência de um único par de nucleotídeo em uma sequência genômica, a cromatina celular é clivada dentro de 10.000 nucleotídeos sobre qualquer lado desse par de nucleotídeo. Em determinadas modalidades, clivagem ocorre dentro de 1,000, 500, 200, 100, 90, 80, 70, 60, 50, 40, 30, 20, 10, 5 ou 2 nucleotídeos ou qualquer valor integral entre 2 e 1.000 nucleotídeos, sobre qualquer lado do par de nucleotídeo cuja sequência tem de ser alterada.

[00125] Conforme detalhado acima, os sítios de ligação para duas proteínas de fusão, cada uma compreendendo um domínio de ligação de dedo de zinco e um meio-domínio de clivagem, podem estar localizados 5-8 ou 15-18 nucleotídeos de distância, conforme medido a partir da borda de cada sítio de ligação mais próximo do outro sítio de ligação e clivagem ocorre entre os sítios de ligação. Se clivagem ocorre em um único

sítio ou em múltiplos sítios entre os sítios de ligação não é relevante, uma vez que as sequências genômicas clivadas são substituídas pelas sequências doadoras. Assim, para alteração eficiente da sequência de um único par de nucleotídeo através de recombinação objetivada, o ponto mediano da região entre os sítios de ligação está dentro de 10.000 nucleotídeos desse par de nucleotídeo, de preferência dentro de 1.000 nucleotídeos ou 500 nucleotídeos ou 200 nucleotídeos ou 100 nucleotídeos ou 50 nucleotídeos ou 20 nucleotídeos ou 10 nucleotídeos ou 5 nucleotídeos ou 2 nucleotídeos ou um nucleotídeo ou no par de nucleotídeo de interesse.

[00126] Em determinadas modalidades, um cromossomo homólogo pode servir como o polinucleotídeo doador. Assim, por exemplo, correção de uma mutação em um heterozigoto pode ser obtida manipulando proteínas de fusão as quais se ligam a e clivam a sequência mutante sobre um cromossomo, mas não clivam a sequência do tipo silvestre sobre o cromossomo homólogo. A ruptura de fita dupla sobre o cromossomo trazendo a mutação estimula um processo de "conversão gênica" baseado em homologia no qual a sequência do tipo silvestre do cromossomo homólogo é copiada para o cromossomo clivado, assim, restaurando duas cópias da sequência do tipo silvestre.

[00127] São também proporcionados métodos e composições que podem intensificar os níveis de recombinação objetivada incluindo, mas não limitado a, o uso de fusões de domínio funcional-ZFP adicionais para ativar a expressão de genes envolvidos em recombinação homóloga tais como, por exemplo, genes de planta do grupo de epistase RAD54 (por exemplo, *AtRad54, AtRad51*) e genes cujos produtos interagem com os produtos gênicos antes mencionados. Vide, por exemplo, Klutstein M, *et al*. Genetics. Abril de 2008; 178(4): 2389-97.

[00128] Similarmente, fusões de domínio funcional-ZFP podem ser usadas, em combinação com os métodos e composições divulgados

aqui, para reprimir a expressão de genes envolvidos em união terminal não homóloga (por exemplo, Ku70/80, XRCC4, poli(ADP ribose) polimerase, DNA ligase 4). Vide, por exemplo, Riha *et al.* (2002) *EMBO* 21: 2819-2826; Freisner *et al.* (2003) *Plant J.* 34: 427-440; Chen *et al.* (1994) *European Journal of Biochemistry* 224: 135-142. Métodos para ativação e repressão de expressão gênica usando fusões entre um domínio de ligação de dedo de zinco e um domínio funcional são divulgados, por exemplo, nas Patentes US cocedidas 6.534.261; 6.824.978 e 6.933.113. Métodos de repressão adicionais incluem o uso de oligonucleotídeos antissenso e/ou pequeno RNA de interferência (siRNA ou RNAi) ou shRNAs objetivados à sequência do gene a ser reprimido.

[00129] Como uma alternativa a ou além disso, ativação de expressão de produtos gênicos envolvidos em recombinação homóloga, fusões dessa proteína (ou fragmentos funcionais da mesma) com um domínio de ligação de dedo de zinco objetivo ao *Zp15*, pode ser usada para recrutar essas proteínas (proteínas de recombinação) à região de interesse, desse modo, aumentando sua concentração local e ainda estimulando processos de recombinação homóloga. Alternativamente, um polipeptídeo envolvido em recombinação homóloga, conforme descrito acima (ou um fragmento funcional do mesmo) pode ser parte de uma proteína de fusão tripla compreendendo um domínio de ligação de dedo de zinco, um domínio de clivagem (ou meio-domínio de clivagem) e a proteína de recombinação (ou fragmento funcional da mesma). Proteínas adicionais envolvidas em conversão gênica e remodelamento de cromatina relacionado à recombinação, as quais podem ser usadas nos métodos e composições antes mencionados, incluem acetil transferases de histona (por exemplo, Esalp, Tip60), metil transferases de histona (por exemplo, Dotlp), quinases de histona e fosfatases de histona. Vide também Bhat *et al.* (1999) *Plant* 7.33: 455-469.

[00130] Aumentos adicionais na eficiência de recombinação objetivada

em células compreendendo uma molécula de fusão de dedo de zinco/nuclease e uma molécula de DNA doadora são obtidos bloqueando as células na fase G2 do ciclo celular, quando processos de reparo acionados por homologia estão maximamente ativos. Tal interrupção pode ser obtida através de uma série de formas. Por exemplo, as células podem ser tratadas, por exemplo, com fármacos, compostos e/ou pequenas moléculas as quais influenciam a progressão do ciclo celular, de forma a interromper as células na fase G2. Moléculas exemplificativas desse tipo incluem, mas não estão limitadas a, compostos os quais afetam a polimerização de microtúbulo (por exemplo, vinblastina, nocodazol, Taxol), compostos que interagem com o DNA (por exemplo, cis-platina(II), dicloreto de diamina, Cisplatina, doxorrubicina) e/ou compostos que afetam a síntese de DNA (por exemplo, timidina, hidroxiureia, L-mimosina, etoposídeo, 5-fluorouracila). Aumentos adicionais na eficiência de recombinação são obtidos mediante o uso de inibidores de deacetilase de histona (HDAC) (por exemplo, butirato de sódio, tricostatina A), os quais alteram a estrutura da cromatina para tornar o DNA genômico mais acessível à máquina de recombinação celular.

[00131] Métodos adicionais para interrupção do ciclo celular incluem superexpressão de proteínas as quais inibem a atividade das quinases CDK do ciclo celular, por exemplo, mediante introdução de um cDNA que codifica a proteína na célula ou por meio de introdução, na célula, de uma ZFP manipulada a qual ativa expressão do gene que codifica a proteína. Interrupção do ciclo celular é também obtida inibindo a atividade de ciclinas e CDKs, por exemplo, usando métodos de RNAi (por exemplo, Patente U.S. No. 6.506.559) ou mediante introdução, na célula, de uma ZFP manipulada a qual reprime a expressão de um ou mais genes envolvidos em progressão do ciclo celular tais como, por exemplo, genes de ciclina e/ou CDK. Vide, por exemplo, Patente U.S. cocedida No. 6.534.261 para métodos de síntese de proteínas do dedo de zinco manipuladas para re-

gulação de expressão gênica.

[00132] Alternativamente, em determinados casos, clivagem objetivada é conduzida na ausência de um polinucleotídeo doador (de preferência na fase S ou G2) e recombinação ocorre entre cromossomos homólogos.

Vetores de Expressão

[00133] Um ácido nucleico que codifica uma ou mais proteínas de fusão (por exemplo, ZFNs), conforme descrito aqui, pode ser clonado em um vetor para transformação em células procariotas ou eucariotas para replicação e/ou expressão. Vetores podem ser vetores procariotas, por exemplo, plasmídeos ou vetores "shuttle", vetores de inseto ou vetores eucariotas. Um ácido nucleico que codifica uma proteína de fusão também pode ser clonado em um vetor de expressão para administração a uma célula vegetal.

[00134] Para expressar as proteínas de fusão (por exemplo, ZFNs), sequências que codificam as proteínas de fusão são, tipicamente, subclonadas em um vetor de expressão que contém um promotor para dirigir a transcrição. Promotores bacterianos e eucariotas adequados são bem-conhecidos na técnica e descritos, por exemplo, em Sambrook *et al, Molecular Cloning, A Laboratory Manual* (2ª ed. 1989; 3ª ed., 2001); Kriegler, *Gene Transfer and Expression: A Laboratory Manual* (1990); e *Current Protocols in Molecular Biology* (Ausubel *et al., supra).* Sistemas de expressão bacterianos para expressão da ZFP estão disponíveis, por exemplo, em *E. coli, Bacillus sp* e *Salmonella* (Palva *et al, Gene* 22: 229-235 (1983)). Kits para tais sistemas de expressão estão comercialmente disponíveis. Sistemas de expressão para células de mamífero, levedura e células de inseto são bem-conhecidos por aqueles versados no campo e também estão comercialmente disponíveis.

[00135] O promotor usado para dirigir a expressão de um ácido nucleico que codifica proteína de fusão depende da aplicação em particular.

Por exemplo, um promotor constitutivo forte adequado para a célula hospedeira é, tipicamente, usado para expressão e purificação de proteínas de fusão.

[00136] Em contraste, quando uma proteína de fusão é administrada *in vivo* para regulação de um gene de planta (vide a seção "Distribuição de Ácido Nucleico a células de Planta" abaixo), um promotor constitutivo ou induzível é usado, dependendo do uso particular da proteína de fusão. Exemplos não limitativos de promotores de planta incluem sequências derivadas de ubiquitina-3 de *A. thaliana* (ubi-3) (Callis, *et al,* 1990, J. Biol. Chem. 265: 12486-12493); sintase de manopina de *A. Tumefaciens* (Amas) (Petolino *et al,* Patente U.S. No. 6.730.824); e/ou Vírus Mosaico do Veio de Cassava (CsVMV) (Verdaguer *et al.,* 1996, Plant Molecular Biology 31: 1129-1139). Vide também Exemplos.

[00137] Além do promotor, o vetor de expressão contém, tipicamente, uma unidade de transcrição ou cassete de expressão que contém todos os elementos adicionais requeridos para a expressão do ácido nucleico em células hospedeiras, quer procariotas ou eucariotas. Um cassete de expressão típico, assim, contém um promotor operativamente ligado, por exemplo, a uma sequência de ácido nucleico que codifica a proteína de fusão e sinais requeridos, por exemplo, para poliadenilação eficiente do transcrito, término de transcrição, sítios de ligação ribossômica ou término de tradução. Elementos adicionais do cassete podem incluir, por exemplo, intensificadores, sinais de junção heterólogos e/ou um sinal de localização nuclear (Nuclear Localization Signal - NLS).

[00138] O vetor de expressão usado em particular para transportar a informação genética para a célula é selecionado com relação ao uso pretendido das proteínas de fusão, por exemplo, expressão em plantas, animais, bactérias, fungos, protozoários, etc. (vide vetores de expressão descritos acima). Vetores de expressão bacterianos e animais padrões

são conhecidos na técnica e são descritos em detalhes, por exemplo, Publicação de Patente U.S. 20050064474A1 e Publicações de Patente Internacionais WO05/084190, WO05/014791 e WO03/080809.

[00139] Métodos padrões de transfecção podem ser usados para produzir linhagens de células bacterianas, de mamífero, levedo ou inseto que expressam grandes quantidades de proteína as quais podem, então, ser purificadas usando técnicas padrões (vide, por exemplo, Colley *et al., J. Biol. Chem.* 264: 17619-17622 (1989); *Guide to Protein Purification, in Methods in Enzymology,* vol. 182 (Deutscher, ed., 1990)). Transformação de células eucariotas e procariotas é realizada de acordo com técnicas padrões (vide, por exemplo, Morrison, *J. Bact.* 132: 349-351 (1977); Clark-Curtiss & Curtiss, *Methods in Enzymology* 101: 347-362 (Wu *et al.,* eds., 1983).

[00140] Qualquer um dos procedimentos bem-conhecidos para introdução de sequências de nucleotídeo estranhas em tais células hospedeiras pode ser usado. Esses incluem o uso de transfecção com fosfato de cálcio, polibreno, fusão de protoplasto, eletroporação, métodos de ultrassom (por exemplo, sonoporação), lipossomas, microinjeção, DNA "nu", vetores de plasmídeo, vetores virais, episssomais e integrativos e qualquer um dos outros métodos bem-conhecidos para introdução de DNA genômico, cDNA, DNA sintético clonado ou outro material genético estranho em uma célula hospedeira (vide, por exemplo, Sambrook *et al*., supra). É necessário apenas que o procedimento de engenharia genética usado em particular seja capaz de introduzir com sucesso pelo menos um gene na célula hospedeira capaz de expressão da proteína de escolha.

Distribuição de Ácido Nucleico a Células de Planta

[00141] Conforme notado acima, construtos de DNA podem ser introduzidos (por exemplo, no genoma) em um hospedeiro vegetal desejado através de uma variedade de técnicas convencionais. Para revisões de tais técnicas vide, por exemplo, Weissbach & Weissbach *Methods for*

Plant Molecular Biology (1988, Academic Press, N.Y.) Seção VIII, páginas 421-463; e Grierson & Corey, *Plant Molecular Biology* (1988, 2ª Ed.), Blackie, Londres, Capítulos 7-9.

[00142] Por exemplo, o construto de DNA pode ser introduzido diretamente no DNA genômico da célula vegetal usando técnicas tais como eletroporação e microinjeção de protoplastos de células de planta ou os construtos de DNA podem ser introduzidos diretamente no tecido de planta usando métodos biolísticos, tal como bombardeamento de partícula de DNA (vide, por exemplo, Klein *et al.* (1987) *Nature* 327: 70-73). Alternativamente, o construto de DNA pode ser introduzido na célula vegetal via transformação de nanopartícula (vide, por exemplo, Pedido de Patente US No. 12/245.685, o qual é incorporado aqui por referência em sua totalidade). Alternativamente, os construtos de DNA podem ser combinados com regiões de flanqueamento/borda de T-DNA adequadas e introduzidos em um vetor hospedeiro *Agrobacterium tumefaciens* convencional. Técnicas de transformação mediadas por *Agrobacterium tumefaciens,* incluindo desarme e uso de vetores binários, são bem descritas na literatura científica. Vide, por exemplo, Horsch *et al.* (1984) *Science* 233: 496-498 e Fraley *et al.* (1983) *Proc. Nat'l. Acad. Sci. USA* 80: 4803.

[00143] Além disso, transferência gênica pode ser obtida usando bactérias que não *Agrobacterium* ou vírus, tais como *Rhizobium sp.* NGR234, *Sinorhizoboium meliloti, Mesorhizobium loti,* vírus X de batata, vírus mosaico da couve-flor e vírus mosaico do veio de cassava e/ou vírus mosaico de tabaco. Vide, por exemplo, Chung *et al.* (2006) *Trends Plant Sci.* 11(1): 1-4.

[00144] As funções de virulência do hospedeiro *Agrobacterium tumefaciens* dirigirão a inserção de uma T-fita contendo o construto e marcador adjacente ao DNA de célula vegetal quando a célula é infectada pela bactéria usando vetor de T DNA binário (Bevan (1984) *Nuc. Acid Res.* 12: 8711-8721) ou o procedimento de cocultura (Horsch *et al.* (1985) *Science*

227: 1229-1231). Em geral, o sistema de transformação com *Agrobacterium* é usado para manipular plantas dicotiledôneas (Bevan *et al.* (1982) *Ann. Rev. Genet* 16: 357-384; Rogers *et al.* (1986) *Methods Enzymol.* 118: 627-641). O sistema de transformação de *Agrobacterium* pode também ser usado para transformar, bem como transferir, DNA para plantas e células de plantas monocotiledôneas. Vide Patente U.S. No. 5.591.616; Hernalsteen *et al.* (1984) *EMBO* 73: 3039-3041; Hooykass-Van Slogteren *et al.* (1984) *Nature* 311: 763-764; Grimsley *et al.* (1987) *Nature* 325: 1677-179; Boulton *et al.* (1989) *Plant Mol. Biol.* 12: 31-40; e Gould *et al.* (1991) *Plant Physiol.* 95: 426-434.

[00145] Métodos alternativos de transferência e transformação gênica incluem, mas não estão limitados a, transformação de protoplasto através de captação mediada por eletroporação, cálcio ou polietileno glicol (PEG) de DNA "nu" (vide Paszkowski *et al.* (1984) *EMBO J* 3: 2717-2722, Potrykus *et al.* (1985) *Molec. Gen. Genet.* 199: 169-177; Fromm *et al.* (1985) *Proc. Nat. Acad. Sci.* USA 82: 5824-5828; e Shimamoto (1989) *Nature* 338: 274-276) e eletroporação de tecidos de planta (D'Halluin *et al.* (1992) *Plant Cell* 4: 1495-1505). Métodos adicionais para transformação de célula vegetal incluem microinjeção, captação de DNA mediada por carbureto de silício (Kaeppler *et al.* (1990) *Plant Cell Reporter* 9: 415-418) e bombardeamento de microprojétil (vide Klein *et al.* (1988) *Proc. Nat. Acad. Sci. USA* 85: 4305-4309; e Gordon-Kamm *et al.* (1990) *Plant Cell* 2: 603-618).

[00146] Os métodos e composições divulgados podem ser usados para inserir sequências exógenas em um local predeterminado (por exemplo, um gene *Zp15*) em um genoma de célula vegetal. Isso é útil uma vez que expressão de um transgene introduzido em um genoma de planta depende criticamente de seu sítio de integração. Consequentemente, genes que codificam, por exemplo, tolerância a herbicida, resistência a inseto, nutrientes, antibióticos ou moléculas terapêuticas podem ser inseridos,

através de recombinação objetivada, em regiões do genoma de uma planta favoráveis à sua expressão.

[00147] Células de planta transformadas as quais são produzidas através de qualquer uma dessas técnicas de transformação podem ser cultivadas para regenerar uma planta inteira a qual possui o genótipo transformado e, assim, o fenótipo desejado. Tais técnicas de regeneração contam com manipulação de determinados fitormônios em um meio de crescimento de cultura tecidual, tipicamente contando com um marcador biocida e/ou herbicida o qual tenha sido introduzido junto com as sequências de nucleotídeo desejadas. A regeneração de planta a partir de protoplastos cultivados é descrita em Evans *et al.,* "Protoplasts Isolation and Culture" em *Handbook of Plant Cell Culture,* páginas 124-176, Macmillian Publishing Company, New York, 1983; e Binding, *Regeneration of Plants, Plant Protoplasts,* páginas 21-73, CRC Press, Boca Raton, 1985. Regeneração também pode ser obtida a partir de caule, explante, órgãos, polens, embriões de planta ou partes dos mesmos. Tais técnicas de regeneração são descritas, em geral, em Klee *et al.* (1987) *Ann. Rev. of Plant Phys.* 38: 467-486.

[00148] Ácidos nucleicos introduzidos em uma célula vegetal podem ser usados para conferir traços desejados a essencialmente qualquer planta. Uma ampla variedade de plantas e sistemas de célula vegetal podem ser manipulados quanto às características fisiológicas e agronômicas desejadas descritas aqui usando os construtos de ácido nucleico da presente invenção e os vários métodos de transformação mencionados acima. Em modalidades preferidas, plantas e células de planta-alvo para manipulação incluem, mas não estão limitadas a, aquelas plantas monocotiledôneas e dicotiledôneas, tais como safras, incluindo safras de grãos (por exemplo, trigo, milho, arroz, painço, cevada), safras de fruta (por exemplo, tomate, maçã, pêra, morango, laranja), safras de forragem (por exemplo, alfafa), safras de vegetais de raiz (por exemplo, cenoura, bata-

ta, beterrabas sacarínicas, inhame), safras de vegetais folhosos (por exemplo, alface, espinafre); plantas de floração (por exemplo, petúnia, rosa, crisântemo), coníferas e pinheiros (por exemplo, pinheiro, pinho); plantas usadas em fitoterapia (por exemplo, plantas que acumulam metal pesada); safras oleosas (por exemplo, girassol, colza) e plantas usadas para fins experimentais (por exemplo, *Arabidopsis*). Assim, os métodos e composições divulgados têm uso sobre uma ampla faixa de plantas incluindo, mas não limitado a, espécies dos gêneros Asparagus, Avena, Brassica, Citrus, Citrullus, Capsicum, Cucurbita, Daucus, Erigeron, Glycine, Gossypium, Hordeum, Lactuca, Lolium, Lycopersicon, Malus, Manihot, Nicotiana, Orychophragmus, Oryza, Persea, Phaseolus, Pisum, Pyrus, Prunus, Raphanus, Secale, Solanum, Sorghum, Triticum, Vitis, Vigna e Zea.

[00149] Aqueles versados no campo reconhecerão que, após a sequência exógena ser estavelmente incorporada em plantas transgênica e confirmada como sendo operável, ela pode ser introduzida em outras plantas através de cruzamento sexual. Qualquer uma de uma série de técnicas padrões de melhoramento pode ser usada, dependendo da espécie a ser cruzada.

[00150] Uma célula, caule, tecido de planta ou planta transformada pode ser identificado e isolado por meio de seleção ou triagem do material vegetal manipulado quanto a traços codificados pelos genes marcadores presentes sobre o DNA de transformação. Por exemplo, seleção pode ser realizada crescendo o material vegetal manipulado sobre meios contendo uma quantidade inibitória do antibiótico ou herbicida ao qual o construto gênico de transformação confere resistência. Ainda, plantas e células de planta transformadas também podem ser identificadas por meio de triagem das atividades de quaisquer genes marcadores visíveis (por exemplo, os genes de β-glucuronidase, luciferase, B ou Cl) que possam estar presentes sobre os construtos de ácido nucleico recombinante.

Tais metodologias de seleção e triagem são bem-conhecidas por aqueles versados no campo.

[00151] Métodos físicos e bioquímicos também podem ser usados para identificar transformantes de planta ou célula vegetal contendo construtos gênicos inseridos. Esses métodos incluem, mas não estão limitados a: 1) análise de Southern ou amplificação por PCR para detecção e determinação da estrutura do inserto de DNA recombinante; 2) Northern blot, proteção com SIRNase, extensão de iniciador ou amplificação por PCR com transcriptase reversa para detecção e exame dos transcritos de RNA dos construtos gênicos; 3) eletroforese em gel de proteína, técnicas de Western blot, imunoprecipitação ou imunoensaios enzima-ligados (ELISA), onde os produtos do construto gênico são proteínas. Técnicas adicionais, tais como hibridização *in situ*, coloração enzimática e imunocoloração, também podem ser usados para detectar a presença ou expressão do construto recombinante em órgãos e tecidos específicos de uma planta. Os métodos para fazer todos esses ensaios são bem-conhecidos por aqueles versados no campo.

[00152] Os efeitos de manipulação gênica usando os métodos divulgados aqui podem ser observados, por exemplo, através de Northern blots do RNA (por exemplo, mRNA) isolado dos tecidos de interesse. Tipicamente, se o mRNA está presente ou a quantidade de mRNA aumentou, pode se admitir que o transgene correspondente está sendo expresso. Outros métodos de medição de atividade do polipeptídeo codificado e/ou gene podem ser usados. Diferentes tipos de ensaios enzimáticos podem ser usados, dependendo do substrato usado e do método de detecção de aumento ou diminuição de um produto ou subproduto de reação. Além disso, os níveis de polipeptídeo expresso podem ser medidos imunoquimicamente, isto é, ELISA, RIA, ELA e outros ensaios baseados em anticorpo bem-conhecidos por aqueles versados no campo, tal como através de ensaios de detecção eletroforéticos (com coloração ou Wes-

tern blotting). Como um exemplo não limitativo, a detecção das proteínas AAD-1 e PAT usando um ensaio ELISA é descrita no Pedido de Patente U.S. No. 11/587.893, referência a qual é aqui incorporada por referência na íntegra. O transgene pode ser seletivamente expresso em alguns tecidos da planta ou em alguns estágios de desenvolvimento ou o transgene pode ser expresso em substancialmente todos os tecidos de planta, substancialmente ao longo de todo o ciclo de vida. Contudo, qualquer modo de expressão combinatorial é também aplicável.

[00153] A presente invenção também abrange sementes das plantas transgênicas descritas acima em que a semente tem o transgene ou construto gênico. A presente invenção ainda abrange a prole, clones, linhagens de célula ou células das plantas transgênicas descritas acima, em que a referida prole, clone, linhagem de célula ou célula tem o transgene ou produto gênico.

[00154] Proteínas de fusão (por exemplo, ZFNs) e vetores de expressão que codificam proteínas de fusão podem ser administrados diretamente à planta para regulação gênica, clivagem objetivada e/ou recombinação. Em determinadas modalidades, a planta contém múltiplos genes-alvo parálogos. Sabe-se que plantas podem conter múltiplos genes parálogos. Assim, uma ou mais proteínas de fusão ou vetores de expressão que codificam proteínas de fusão diferentes podem ser administrados a uma planta de forma a objetivar um ou mais genes *Zp15* na planta.

[00155] A administração de quantidades eficazes é através de qualquer uma das vias normalmente usadas para introdução de proteínas de fusão em contato com a célula vegetal a ser tratada. As ZFPs são administradas através de muitas maneiras adequadas, de preferência com veículos aceitáveis. Métodos adequados de administração de tais moduladores estão disponíveis e bem-conhecidos por aqueles versados no campo e, embora mais de uma via possa ser usada para administrar uma composição em particular, um via particular pode, frequentemente, pro-

porcionar uma reação mais imediata e mais eficaz do que outra via.

[00156] Veículos também podem ser usados e são determinados, em parte, pela composição que está sendo administrada, bem como pelo método usado para administrar a composição em particular. Consequentemente, há uma ampla variedade de formulações adequadas de veículos que estão disponíveis.

EXEMPLOS

[00157] Abaixo estão exemplos de modalidades específicas para realização da presente invenção. Os exemplos são oferecidos para fins ilustrativos apenas e não se destinam a limitar o escopo da presente invenção de qualquer forma.

[00158] Esforços foram feitos para assegurar precisão com relação aos números usados (por exemplo, quantidades, temperaturas, etc.), mas algum erro e desvio experimental pode, naturalmente, ser permitido.

Exemplo 1: Identificação e Caracterização do *Locus* Alvo de *Zp15*

[00159] Baseado em mapas genéticos publicamente disponíveis (Lawrence, C. *et al.* (2004) NAR 32: 393-397; banco de dados de milho na internet) e na sequência de genoma extraída de milho, o *locus* de *Zp15* sobre o braço curto do cromossomo 6 foi escolhido como um alvo para modificação usando ZFNs baseado em localização e informação disponível sobre o gene *Zp15*. A estrutura e sequência genômica de um gene que codifica beta zeína de 15 kD (*Zp15*) de milho foram descritas e anotadas no domínio público (Woo *et al.* (2001) *Plant Cell* 13(10): 2297-2313). A sequência para o gene *Zp15* é descrita no número de acesso ao GenBank AF371264, o qual é incorporado aqui por referência.

[00160] A sequência genômica do *Zp15* foi usada como "query" em bancos de dados de genoma TIGR e Maize GDB usando algoritmos BLAST. Várias sequências com homologia de sobreposição ao *Zp15* incluindo, mas não limitado a, dois contigs (AZM5_16782 e ZmGSStuc11-12-04.8785.1) e vários ESTs (M72708, M13507, M12147, AY103640 e

AF371264) foram identificados. Baseado na sequência desses acessos, bem como na sequência do *Zp15*, múltiplos oligonucleotídeos curtos foram projetados para uso como iniciadores de PCR usando o programa Primer3 (Rozen, S. e Skaletsky, H.J. (2000) *Primeri on the WWW for general users and for biologist programmers.* Em: Krawetz S, Misener S (eds.) Bioinformatics Methods and Protocols: Methods in Molecular Biology. Humana Press, Totowa, NJ, páginas 365-386; também disponível na internet).

[00161] Esses iniciadores incluem, mas não estão limitados a, os seguintes oligonucleotídeos na orientação para frente:

P67F 5' – CGTATGAATTCATTGACAACC – 3' (SEQ ID NO:1)
P68F 5' – ATGATCTATCTGTAAATCC – 3' (SEQ ID NO:2)
P69F 5' – CGTCATGCAACGCAACATTCC – 3' (SEQ ID NO:3)
P73F 5' – AAGAACATCACAAGTTATGC – 3' (SEQ ID NO:4)
P74F 5' – TCATGTGGATCCAAGGCATC – 3' (SEQ ID NO:5)

[00162] Além disso, os iniciadores incluem, mas não estão limitados a, os seguintes oligonucleotídeos na orientação reversa:

P70R 5' – ATGTGTGTCGTCTTACTGC – 3' (SEQ ID NO:6)
P71R 5' – CAGTAGTAGGGCGGAATG – 3' (SEQ ID NO:7)
P72R 5' – GGGCAGCTGGTACTG – 3' (SEQ ID NO:8)
P75R 5' – CTATAATCGATGTAGAGC – 3' (SEQ ID NO:9)
P76R 5' – CTATGCTTTGTCTATAGTCG – 3' (SEQ ID NO:10)

[00163] Todos os iniciadores de oligonucleotídeo foram sintetizados por e adquiridos da Integrated DNA Technologies (IDT, Coralville, LA). Amplificações de gDNA da variedade de milho Hi-II foram realizadas sobre um PCR Thermal Cycler usando 30 ng de gDNA. Um fragmento de amplificação de 2.215 bp, correspondendo ao gene *Zp15* de Hi-II, foi isolado e clonado no plasmídeo pCR2.1 (Invitrogen, Carlsbad, CA). Análise de sequência desse fragmento revelou que a estrutura genômica de *Zp15* da variedade de milho Hi-II contém dois éxons e um pequeno íntron de 31

bp (SEQ ID NO: 126). Designs para ZFN objetivada ao *Zp15* são focalizados sobre as regiões de codificação do gene *Zp15* da variedade de milho Hi-II.

Exemplo 2: Design de Nucleases de Dedo de Zinco Objetivadas ao Gene *Zp15* de Milho

[00164] De forma a montar vetores de expressão para ZFNs, um esquema de clonagem modular em etapas foi projetado o qual é aplicável para qualquer determinado par de genes que codificam ZFN selecionados do arquivo de biblioteca ou sintetizados *de novo.*

[00165] ZFNs objetivadas ao *Zp15*, primeiro, sofreram triagem usando a triagem por ensaio com levedura conforme descrito em Doyon *et al.* (2008) *Nature Biotechnology* **26(6):** 702 e Pedido de Patente U.S. No. 12/284.887. Resumidamente, todo o *locus* de *Zp15* foi introduzido no *locus* HO no genoma de levedura de forma a comparar diretamente a atividade de ZFN em diferentes sítios de ligação dentro do gene-alvo; ZFNs sofreram triagem quanto à sua capacidade de induzir a DSB no gene repórter usando um ensaio repórter (MEL1), conforme descrito em Doyon *et al.*

[00166] Baseado nos resultados desses ensaios de sistema *proxy*, foi confirmado que vários pares de ZFN testados eram capazes de induzir a DSBs dentro do *Zp15.*

[00167] Após pré-triagem de levedo, os pares de ZFN foram, então, subclonados em vetores de expressão de milho específicos. Conforme descrito na Publicação de Patente U.S. No. 20080182332, um vetor incluindo segmentos reprojetados e sintetizados de um sinal de localização nuclear (NLS) derivado do domínio de nuclease *op-2* e *FokI* de milho utilizando a tendência de códon para milho foi modificado com uma única inserção de nucleotídeo (C) a jusante do sítio *Xho* I único para criar um sítio *Sac* I extra. Um vetor similar foi modificado para incluir a sequência ribossômica 2A do vírus *Thosea asigna*. Os cassetes gênicos que codifi-

cam ORFs de proteínas do dedo de zinco individuais foram clonados em qualquer um desses vetores via sítios de restrição *Kpn* I e *BamH* I e subsequentemente os dois vetores foram combinados via sítios de restrição *Bgl* IV - *Xho* I, proporcionando um construto intermediário que continha um cassete incluindo 2 domínios de codificação de ZFN flanqueados por sítios de restrição *Nco* I e *Sac* I.

[00168] O cassete *Nco VSac* I dessa construção intermediária foi excisado via digestão com enzima de restrição e ligado na parte principal de plasmídeo pDAB3872, o qual contém um promotor do gene de *ubiquitina-1* de milho (Sharrock *et al.* (1992) *Plant Mol Biol.* 18(4): 675) e sequências terminadoras do gene de peroxidase catiônica raiz de milho preferencial (Patente US No. 7.179.902).

[00169] Os plasmídeos resultantes incluem os genes de ZFN, mais os marcadores selecionáveis relevantes para manutenção de plasmídeo e sítios de flanqueamento attL para manipulação conveniente usando o sistema GATEWAY™ da Invitrogen (Carlsbad, CA). Cada um dos construtos de ZFN gerados usando esse esquema de clonagem foi transformado em células DH5a de *E. coli* e subsequentemente mantidos sob a seleção apropriada.

[00170] A Tabela 1 mostra ZFNs objetivadas ao *Zp15* exemplificativas que foram usadas para experimentos de integração objetivada no *locus* de *Zp15*. A sequência de DNA-alvo para a ZFN é mostrada na segunda coluna (sítios-alvo de DNA indicados em letras maiúsculas; nucleotídeos não contatados indicados em letras minúsculas) e as terceira à sexta coluna mostram a sequência de aminoácido da região de reconhecimento (aminoácidos =1 a +6, com relação ao início da hélice) de cada um dos dedos de zinco (F1 a F4) na proteína. Também fornecido na primeira coluna da Tabela 1 é um número de identificação para cada proteína.

Tabela 1

Nome ZFN	Sítio-alvo	F1	F2	F3	F4
11742	cgGGGCTGCAGGGCttg tacggcgctgg (SEQ ID NO:11)	DRSHLTR (SEQ ID NO:12)	RSDNLRE (SEQ ID NO:13)	RSDVLSE (SEQ ID NO:14)	RSAHLSR (SEQ ID NO:15)
11743	gcAGGGGCAGGGCAtct gcattgcagag (SEQ ID NO:16)	QSGSLTR (SEQ ID NO:17)	RSDHLTQ (SEQ ID NO:18)	DRSHLTR (SEQ ID NO:12)	RSDHLTQ (SEQ ID NO:18)
11750	ctGAGGCAGCCGCAgtg cagcccgctgg (SEQ ID NO:19)	QSGDLTR (SEQ ID NO:20)	DRSDLSR (SEQ ID NO:21)	QSGDLTR (SEQ ID NO:20)	RSDNLTR (SEQ ID NO:22)
11753	acTCCGCGTAGGGGta cagcccgccggc (SEQ ID NO:23)	RSDHLSR (SEQ ID NO:24)	RSDNLTT (SEQ ID NO:25)	RSDDLTR (SEQ ID NO:26)	DSSDRKK (SEQ ID NO:27)
11754	acTCCGCGTAGGGGta cagcccgccggc (SEQ ID NO:23)	RSDHLSE (SEQ ID NO:29)	RNDNRKN (SEQ ID NO:30)	RSDDLTR (SEQ ID NO:26)	DSSDRKK (SEQ ID NO:27)
11755	agCCGCAGtGCAGCCc gctggcggcggc (SEQ ID NO:31)	DRSDLSR (SEQ ID NO:21)	QSSDLTR (SEQ ID NO:32)	RSDHLSE (SEQ ID NO:29)	TSSTRKT (SEQ ID NO:33)
11756	agCCGCAGtGCAGCCc gctggcggcggc (SEQ ID NO:31)	DRSDLSR (SEQ ID NO:21)	QSSDLTR (SEQ ID NO:32)	RSDHLSE (SEQ ID NO:29)	RSSTRKE (SEQ ID NO:35)
11757	ccTCAGGTACTCCGcgt aggggtacagc (SEQ ID NO:36)	RSDTLSE (SEQ ID NO:37)	ARSTRTN (SEQ ID NO:38)	QSSHLTR (SEQ ID NO:39)	QSADRTK (SEQ ID NO:40)
11758	gcCCGCTGGCGGCGgc gccctactacgc (SEQ ID NO:41)	RSDDLTR (SEQ ID NO:26)	RSDDLTR (SEQ ID NO:26)	RSDTLSA (SEQ ID NO:42)	RNQDRKT (SEQ ID NO:43)
11759	gcACTGCGGCTGCCtca ggtactccgcg (SEQ ID NO:44)	DRSDLSR (SEQ ID NO:21)	QSSDLRR (SEQ ID NO:45)	RSDDLTR (SEQ ID NO:26)	QSSDLTR (SEQ ID NO:32)
11760	cgCCGGGTGTGGGCag ccgagcgccatg (SEQ ID NO:46)	DRSHLSR (SEQ ID NO:47)	RSDALAR (SEQ ID NO:48)	QSSHLTR (SEQ ID NO:39)	RSDDRKT (SEQ ID NO:49)
11761	cgCCGGGTgTGGGCAg ccgagcgccatg (SEQ ID NO:46)	QSGSLTR (SEQ ID NO:17)	RSDHLTT (SEQ ID NO:51)	QSSHLTR (SEQ ID NO:39)	RSDDRKT (SEQ ID NO:49)
11762	gcCGGGTGTGGGCAgc cgagcgccatgt (SEQ ID NO:52)	QSGSLTR (SEQ ID NO:17)	RSDHLTT (SEQ ID NO:51)	RSDSLLR (SEQ ID NO:53)	RSDNLRE (SEQ ID NO:13)
11763	agTAGGGCGCCGCCgc cagcgggctgca (SEQ ID NO:54)	RSDNLTT (SEQ ID NO:25)	DRSDLSR (SEQ ID NO:21)	DRSHLTR (SEQ ID NO:12)	RSDNLTT (SEQ ID NO:25)
11766	tgTGGGCAGCCGAGcg ccatgttccagc (SEQ ID NO:55)	RSDNLAR (SEQ ID NO:56)	DRSDLSR (SEQ ID NO:21)	QSGSLTR (SEQ ID NO:17)	RSDHLTT (SEQ ID NO:51)
11767	tgTGGGCAGCCGAGcg ccatgttccagc (SEQ ID NO:55)	RSDNLSR (SEQ ID NO:58)	DNSTRKT (SEQ ID NO:59)	QSGSLTR (SEQ ID NO:17)	RSDHLTT (SEQ ID NO:51)
11768	cgGCGTAGTAGGGCgc cgccgccagcgg (SEQ ID NO:60)	DRSHLTR (SEQ ID NO:12)	RSDNLTT (SEQ ID NO:25)	RSDNLST (SEQ ID NO:61)	RSADLSR (SEQ ID NO:62)
11769	cgGCGTAGtAGGGCGc cgccgccagcgg (SEQ ID NO:60)	RSDDLTR (SEQ ID NO:26)	RSDHLTQ (SEQ ID NO:18)	RSDNLST (SEQ ID NO:61)	RSADLSR (SEQ ID NO:62)
11770	caGCCGCTCCGGCAac agtgctgccagc (SEQ ID NO:64)	QSGSLTR (SEQ ID NO:17)	RSDDRKT (SEQ ID NO:49)	QSSDLSR (SEQ ID NO:65)	DRSDLSR (SEQ ID NO:21)

11771	acATGGCGcTCGGCTg cccacacccggc (SEQ ID NO:66)	QSSDLSR (SEQ ID NO:65)	RNDDRKK (SEQ ID NO:67)	RSDDLTR (SEQ ID NO:26)	RSDALTQ (SEQ ID NO:68)
11772	tgGCAGCCCAGGGTctc aaccccatggc (SEQ ID NO:69)	QSSHLTR (SEQ ID NO:39)	RSDNLRE (SEQ ID NO:13)	DRSDLSR (SEQ ID NO:21)	QSSDLTR (SEQ ID NO:32)
11773	caGCTGCTGCTGCTgct gcatcagagct (SEQ ID NO:70)	QSSDLSR (SEQ ID NO:65)	QSSDLRR (SEQ ID NO:45)	QSSDLSR (SEQ ID NO:65)	QSSDLRR (SEQ ID NO:45)
11774	ctGCCCAGCTACCGcac caacccctgtg (SEQ ID NO:28)	RSDSLSA (SEQ ID NO:34)	DNSNRIK (SEQ ID NO:50)	RSDNLSE (SEQ ID NO:57)	ASKTRKN (SEQ ID NO:63)
11775	tgGTACTGGTAGAGtcc acccatggccg (SEQ ID NO:71)	RSDNLAR (SEQ ID NO:56)	QSGSLTR (SEQ ID NO:17)	RSDVLSE (SEQ ID NO:14)	QSGSLTR (SEQ ID NO:17)
11776	tgGTACTGGTAGAGtcc acccatggccg (SEQ ID NO:71)	RSDNLAR (SEQ ID NO:56)	QSGSLTR (SEQ ID NO:17)	RSDALSN (SEQ ID NO:72)	TSSARTT (SEQ ID NO:73)
11777	aaCCCCTGtGGCGTCtc cgctgccattc (SEQ ID NO:74)	DRSALSR (SEQ ID NO:75)	DRSHLAR (SEQ ID NO:76)	RSDTLSA (SEQ ID NO:42)	DRSTRTT (SEQ ID NO:77)
11778	gtGCGGTAGCTGGGca gctggtactggt (SEQ ID NO:78)	RSDHLSR (SEQ ID NO:24)	QSSDLRR (SEQ ID NO:45)	QSGALAR (SEQ ID NO:79)	RSDDLTR (SEQ ID NO:26)
11779	gtGCGGTAGCTGGGca gctggtactggt (SEQ ID NO:78)	RSDHLST (SEQ ID NO:80)	HSDTRKK (SEQ ID NO:81)	QSGALAR (SEQ ID NO:79)	RSDDLTR (SEQ ID NO:26)
11780	agGCGGGGcTTGACGa agttggaagccg (SEQ ID NO:82)	RSDSLSV (SEQ ID NO:83)	QNQHRIN (SEQ ID NO:84)	RSDHLSR (SEQ ID NO:24)	RSDDLTR (SEQ ID NO:26)
11781	aaATGGAAAAAACGcta aaattatgtgt (SEQ ID NO:85)	RSDDLSK (SEQ ID NO:86)	RNDHRKN (SEQ ID NO:87)	QRSNLVR (SEQ ID NO:88)	RSDALTQ (SEQ ID NO:68)
11782	ttGTGGTGCCAACGgga gccatgctcac (SEQ ID NO:89)	RSDTLSQ (SEQ ID NO:90)	QNATRIN (SEQ ID NO:91)	RSDALSR (SEQ ID NO:92)	RSDALAR (SEQ ID NO:48)
11783	ttGTGGTGCCAACGgga gccatgctcac (SEQ ID NO:89)	RSDTLSQ (SEQ ID NO:90)	QKATRIT (SEQ ID NO:93)	RSDALSR (SEQ ID NO:92)	RSDALAR (SEQ ID NO:48)
11784	ttGTGGTGcCAACGGga gccatgctcac (SEQ ID NO:89)	RSDHLSE (SEQ ID NO:29)	QNANRKT (SEQ ID NO:94)	RSDALSR (SEQ ID NO:92)	RSDALAR (SEQ ID NO:48)
11785	caATCACGCCGGTAgc ggggctagttat (SEQ ID NO:95)	QSGALAR (SEQ ID NO:79)	RSDDRKT (SEQ ID NO:49)	RSDTLSQ (SEQ ID NO:90)	DSSARKK (SEQ ID NO:96)

[00171] Será evidente que ZFNs podem ser prontamente inseridas em partes principais de C2H2 ou C3H e que uma variedade de sequências pode ser usada para unir a proteína de dedo de zinco e o domínio de clivagem. Vide Publicação de Patente U.S. 20080182332, particularmente Tabela 6, com relação a tais sequências, referência a qual é incorporada aqui por referência na íntegra.

Exemplo 3: Ruptura ZFN-mediada de *Zp15* em células de milho

[00172] Indução de DSB por pares de ZFN foi testada. Pares de ZFN que são capazes de produção eficiente de DSB no sítio-alvo pretendido

do gene *Zp15* endógeno foram identificados. A natureza propensa a erro de reparo DSB através de união terminal não homóloga (NHEJ), a qual é conhecida por gerar pequenas deleções/inserções de DNA no local de uma ruptura induzida por ZFN, foi utilizada para selecionar pares de ZFN os quais se ligam e clivam eficientemente o sítio-alvo do gene *Zp15* endógeno.

[00173] ZFNs foram transitoriamente expressas em células de milho cultivadas e análise de sequência do *locus* alvo no sítio de clivagem previsto foi conduzida. Por exemplo, um plasmídeo pDAB7468 que codifica o par de ZFN # 25 (11768/11766 hélices de reconhecimento mostradas acima) foi distribuído via transformação WHISKERS™-mediada em culturas de célula de milho Hi-II, conforme descrito no Pedido de Patente U.S. No. 12/001.939, o qual é incorporado aqui por referência na íntegra. Após 24 ou 72 horas de expressão transitória, a sequência-alvo de ZFN rompida resultante foi amplificada do DNA genômico isolado e clonada no vetor de plasmídeo pCR2.1. O gDNA foi submetido à digestão por restrição usando a enzima *Bsu36*I, seguido por amplificação da sequência-alvo de *Zp15* e clonagem dos produtos de PCR no vetor de plasmídeo pCR2.1. Colônias individuais do produto da amplificação clonado foram analisadas por meio de digestões por restrição do DNA de plasmídeo, seguido por eletroforese em gel de agarose (produtos da amplificação clonados que mostraram resistência à clivagem pela enzima de restrição *Bsu36*I foram considerados como contendo mutações que destroem o sítio de restrição associado ao sítio de clivagem de ZFN).

[00174] Análise de sequência direta de 192 clones revelou uma inserção de 6 bp (Figura 1). Em outro exemplo, o plasmídeo pDAB7467 que codifica o par de ZFN # 24 (11753 /11750 hélices de reconhecimento mostradas acima) foi distribuído diretamente em culturas de célula de milho e, após 24 ou 72 horas de expressão transitória, a sequência-alvo de ZFN foi amplificada e clonada no vetor de plasmídeo pCR2.1. Análise de

sequência direta de 192 clones revelou uma deleção de 3 bp (Figura 2) no sítio de clivagem preciso. A inserção e deleção descritas aqui são o início de reparo NHEJ de um DSB induzido no sítio-alvo e indica que as ZFNs 24 e 25 têm atividade de clivagem no *locus* de *Zp15* endógeno em células de milho.

[00175] O mesmo processo foi realizado usando ZFN 25 ou 28, mas antes de triagem de colônias através de digestão com enzima de restrição, 192 clones independentes foram diretamente sequenciados. Uma inserção de 6 bp foi detectada (Figura 2, superior).

[00176] Tomados juntos, esses dados demonstram que uma exposição transitória a ZFNs é insuficiente para induzir a um DSB objetivado no *locus* de *Zp15* em células de milho cultivadas.

Exemplo 4: Integração Objetivada no *locus* de *Zp15*

[00177] De forma a testar se ZFNs projetadas com atividade em *Zp15* poderiam acionar a integração de sequências exógenas, foram construídas moléculas de DNA doadoras trazendo um cassete gênico autônomo que codifica um gene de resistência a herbicida exógeno exemplificativo, *AAD-1* de *Sphingomonas herbicidovorans* (ATCC 700291). *AAD-1* codifica a enzima dioxigenase de arilóxi alcanoato e confere resistência a herbicidas de ariloxifenóxi propionato (Patentes Internacionais WO 2005/107437, WO2008141154 A2). Aqueles versados no campo apreciarão que outros ácidos nucleicos exógenos poderiam ser similarmente incorporados em moléculas doadoras de DNA incluindo, mas não limitado a, outros genes de tolerância a herbicida, tal como o gene *AAD-12* relacionado. Nesse doador de gene de tolerância a herbicida, a sequência promotora é derivada de actina de *O. sativa* (números de acesso ao GenBank S44221 e X63830) e sequências terminadoras são derivadas de lipase de Z. *mays* (Número de Acesso ao GenBank L35913).

A. Construção de Molécula de DNA Doadora

[00178] Construtos doadores contendo regiões de homologia ao *Zp15*

foram gerados como segue. Uma parte principal de plasmídeo contendo flancos de homologia para o gene *Zp15* foi manipulada para permitir a integração de qualquer sequência de DNA doadora no sítio-alvo correspondente do gene *Zp15*. A parte principal de plasmídeo exemplificada aqui se originou com o vetor de plasmídeo de base fagemídeo pBC SK(-) (3,4 kbp) (Stratagene, La Jolla, CA). Houveram quatro etapas nesse processo.

[00179] Primeiro, o plasmídeo de base foi preparado linearizando 3 μg de pBC SK(-) usando as endonucleases de restrição *Spe* I e *Sal* I (New England Biolabs, Beverly, MA). O vetor de subclonagem de 3,3 kbp digerido com *Spe* I/*Sal* I, pBC SK(-), foi excisado em gel e purificado de acordo com as orientações do fabricante usando o kit QIAQUICK Gel Extraction (QIAGEN Inc., Valencia, CA).

[00180] Segundo, os flancos de 5'- & 3'-homologia foram isolados do *Zp15* usando os seguintes iniciadores de oligonucleotídeo, que foram sintetizados pela Integrated DNA Technologies, Inc. (Coralville, IA):

Zp15HR^{Doador} NotI(205)F: 5'- GCGGCCGCATGCAAGAGCTGTTGATC -3' (SEQ ID NO:97)

Zp15HR^{Doador} MfeI(1025)R: 5'-CAATTGCCGGCGTAGTAGGGCGCCGCCGCCAGC -3' (SEQ ID NO:98)

Zp15HR^{Doador} MfeI(1025)F: 5'-CAATTGGTGTGGGCAGCCGAGCGCCATGTTCCAG - 3' (SEQ ID NO:99)

Zp15HR^{Doador} SalI(2270)R 5' - GTCGACCGATACTGATGCGGACCGTCCACCTTGTC - 3' (SEQ ID NO:100).

[00181] Reações de amplificação por PCR foram realizadas usando reagentes fornecidos com o kit LA TAQ PCR (TaKaRa Biotechnology Inc. Otsu, Shiga, Japão). O coquetel de reação de PCR consistia em: 5 uL de 10X Tampão LA PCR™ II ($Mg2^{+}$), 20 ng de molde de fita dupla (DNA genômico de milho Hi-II), 10 pmol de iniciador de oligonucleotídeo para a frente, 10 pmol de iniciador de oligonucleotídeo reverso, 8 uL de mistura de dNTP (2,5 mM cada dNTP), 33,5 uL de H_2O, 0,5 uL (2,5 unidades) de

DNA Polimerase *TaKaRa LA Taq*™, 1 gota de óleo mineral. Os iniciadores Zp15HRDoadorNotI(205)F e Zp15HRDoadorMfeI(1025)R foram usados para uma reação e os iniciadores Zp15HRDoadorMfeI(1025)F e Zp15HRDoadorSalI(2270)R foram usados para a segunda reação. Reações de PCR foram realizadas usando um Perkin-Elmer Cetus, 48-Sample DNA Thermal Cycler (Norwalk, CT) sob as seguintes condições de ciclo: 94°C, 4 min/1 ciclo; 98°C, 20 seg, 65°C, 1 min, 68°C, 1 min/30 ciclos; 72°C, 5 min/1 ciclo; 4°C/manutenção.

[00182] Quinze (15) µL de cada reação de PCR foram submetidos à eletroporação e os fragmentos amplificados foram visualizados com luz UV e os tamanhos de fragmento estimados por meio de comparação com "ladder" de DNA de 1 kbp. Clones de plasmídeo esperados foram diagnosticados pela presença de fragmentos de DNA de 825 bp para o 5'-fragmento ou 1.250 bp para o 3'-fragmento. Esses fragmentos foram excisados em gel e purificados de acordo com as orientações do fabricante usando o kit QIAquick Gel Extraction (QIAGEN Inc., Valencia, CA). Fragmentos purificados foram, então, clonados no plasmídeo pCR2.1 usando o kit TOPO TA CLONING® e transformados em células de *E. coli* quimicamente competentes ONE SHOT® TOP 10 (Invitrogen Life Technologies, Carlsbad, CA) de acordo com o protocolo do fabricante.

[00183] Colônias individuais contendo o fragmento 5' de 825 bp ou o fragmento 3' de 1.250 bp foram identificadas e confirmadas via digestão com enzima de restrição e dados de sequenciamento. Colônias contendo o fragmento 5' de 825 bp foram confirmadas via uma digestão com enzima de restrição *Mfe* I e *Not* I (New England Biolabs, Beverly, MA). Colônias contendo o fragmento 3' de 1.250 bp foram identificadas e confirmadas via digestão com enzima de restrição usando *Sal* I (New England Biolabs, Beverly, MA). Clones de plasmídeo esperados foram diagnosticados pela presença de fragmentos de DNA inseridos de 825 bp (5'-fragmento) ou 1.250 bp (3'-fragmento), além do vetor pCR®2.1 de 3,9 kbp. Reações

de sequenciamento de fita dupla dos clones de plasmídeo foram realizadas conforme descrito pelo fabricante usando o kit CEQ™ DTCS-Quick Start (Beckman-Coulter, Palo Alto, CA). As reações foram purificadas usando Performa DTR Gel Filtration Cartridges (Edge BioSystems, Gaithersburg, MD), conforme descrito pelos protocolos do fabricante. Reações de sequência foram analisadas sobre um Beckman-Coulter CEQ™ 2000 XL DNA Analysis System e caracterização de nucleotídeo realizada usando um SEQUENCHER™ versão 4.1.4 (Gene Codes Corporation, Ann Arbor, MI). As sequências de fragmentos de 5'- e 3'-homologia de *Zp15* são indicadas em SEQ ID NO: 127 e SEQ ID NO: 128.

[00184] Terceiro, o flanco de 3'-homologia foi ligado no plasmídeo de base, como segue. Clones que continham a sequência do flanco de 3'-homologia correta foram digeridos com enzimas de restrição e um fragmento de DNA foi excisado em gel e purificado usando o kit QIAquick Gel Extraction (QIAGEN Inc., Valencia, CA). Esses fragmentos foram ligados em um plasmídeo de base purificado, previamente digerido com *Spe* I/*Sal* I (vide acima), em uma proporção de vetor:inserto de 1:5 usando 500 unidades de DNA Ligase T4 (Invitrogen Life Technologies, Carlsbad, CA) em um volume de reação de 20 uL sob condições de incubação de 16 horas em um banho de água a 16°C. Cinco (5) uL da reação de ligação foram subsequentemente transformados em células de *E. coli* quimicamente competentes One SHOT® Top 10 (Invitrogen Life Technologies, Carlsbad, CA) e colocados sobre meios contendo seleção por antibiótico. Colônias putativas foram isoladas e digeridas com as enzimas de restrição *Spe* I e *Sal* I (New England Biolabs, Beverly, MA) para identificar clones os quais continham o fragmento 3' ligado.

[00185] Quarto, o flanco de 5'-homologia foi ligado no plasmídeo contendo o flanco de 3'-homologia. O plasmídeo contendo o flanco de 3'-homologia, descrito na etapa três acima, foi digerido com endonucleases de restrição *Mfe* I e *Not* I (New England Biolabs, Beverly, MA) durante 1

hora a 37°C. O plasmídeo digerido com *Not* I/ *Mfe* I contendo o flanco de 3' -homologia foi excisado em gel e purificado de acordo com as orientações do fabricante usando o kit QIAquick Gel Extraction (QIAGEN Inc., Valencia, CA).

[00186] Fragmentos isolados do doador de flanco de 5'-homologia gerados através de digestão com enzima de restrição usando as enzimas de restrição *Mfe* I e *Not* I foram produzidos e ligados com o plasmídeo contendo o flanco de 3'-homologia em uma reação de ligação de 20 μL usando uma proporção de vetor:inserto de 1:5 e 500 unidades de DNA Ligase T4 (Invitrogen Life Technologies, Carlsbad, CA). As reações de ligação foram incubadas durante 16 horas em um banho de água a 16°C.

[00187] Após a ligação, 5 μL da reação de ligação foram transformados em células DH5a™ Quimicamente Competentes MAX EFFICIENCY® (Invitrogen Life Technologies, Carlsbad, CA), conforme as recomendações do fabricante. Colônias individuais foram selecionadas e DNA de plasmídeo foi isolado e digerido com a enzima de restrição *Not* I (New England Biolabs, Beverly, MA) para identificar plasmídeos os quais continham um fragmento integrado do doador do flanco de 5'-homologia. Ao plasmídeo resultante foi fornecido o nome pDAB7489 (Figura 3).

[00188] Um cassete de expressão de gene de tolerância a herbicida compreendendo uma unidade transcricional de planta (PTU) contendo promotor, gene de tolerância a herbicida e sequências de término de poliadenilação (poliA) foi construído. A sequência promotora é derivada de actina 1 de *O. sativa* (McElroy *et al.* (1990) *Plant Cell* 2, 163-171; Acesso ao GenBank S44221 e Acesso ao GenBank X63830). O gene de tolerância a herbicida compreendia o gene *AAD-1* (dioxigenase de arilóxi alcanoato), o qual confere resistência a herbicidas de ariloxifenóxi propionato (WO 2005/107437). A versão do gene utilizada foi a versão #3, a qual inclui uma sequência códon-otimizada para expressão em plantas. As sequências terminadoras são derivadas de lipase de Z. *mays* (clone de

cDNA de lipase de milho com Número de Acesso ao GenBank L35913). Essa sequência de milho compreende a região não traduzida 3'/região de término de transcrição para o gene *AAD-1*. O cassete de expressão de gene de tolerância a herbicida é mostrado na Figura 4.

[00189] Para gerar esse cassete, os seguintes iniciadores de oligonucleotídeo foram sintetizados pela Integrated DNA Technologies, Inc. (Coralville, LA) sob condições de dessalinização padrão e diluídos com água para uma concentração de 0,125 μg/uL:

OsActAad1v.3ZmLipMfeIF 5' – CAATTGGTCATTCATATGCTTGAGAAGAG – 3' (SEQ ID NO:101)

OsActAad1v.3ZmLipMfeIR 5' – CAATTGAGCACTTAAAGATCTTTAGAAG – 3' (SEQ ID NO:102)

[00190] Reações de amplificação por PCR foram realizadas usando o kit LA TAQ PCR (*TaKaRa* Biotechnology Inc., Otsu, Shiga,, Japão). O coquetel de reação de PCR compreendia: 5 uL de 10X Tampão LA PCR™ II ($Mg2^{+}$), 20 ng de molde de fita dupla (DNA de plasmídeo pDAB3878), 10 pmol de iniciador de oligonucleotídeo para a frente, 10 pmol de iniciador de oligonucleotídeo reverso, 8 uL de mistura de dNTP (2,5 mM cada), 33,5 uL de H_2O, 0,5 uL (2,5 unidades) de DNA Polimerase *TaKaRa LA Taq*™, 1 gota de óleo mineral. As reações de PCR foram realizadas usando um Perkin-Elmer Cetus, 48-Sample DNA Thermal Cycler (Norwalk, CT) sob as seguintes condições de ciclo: 94°C, 4 min/1 ciclo; 98°C, 20 seg, 55°C, 1 min, 68°C, 3 min/30 ciclos; 72°C, 5 min/1 ciclo; 4°C/manutenção. Quinze (15) μL de cada reação de PCR foram submetidos à eletroforese a 100 V durante 1 hora em géis de agarose TAE a 1,0% suplementados com brometo de etídio a 0,5%. Fragmentos amplificados foram visualizados com luz UV e o tamanho do fragmento estimado por meio de comparação com o "ladder" de DNA de 1 kbp.

[00191] Produtos de PCR esperados foram diagnosticados pela presença de um fragmento de DNA de 2,7 kbp (*AAD-1* PTU). Esse fragmento foi excisado em gel e purificado de acordo com as orientações

do fabricante usando o kit QIAquick Gel Extraction (QIAGEN Inc., Valencia, CA). Fragmentos purificados foram, então, clonados no plasmídeo pCR2.1 usando o kit TOPO TA Cloning® (com o vetor pCR®2.1) e células de *E. coli* quimicamente competentes One Shot® TOP 10 (Invitrogen Life Technologies, Carlsbad, CA) de acordo com o protocolo do fabricante.

[00192] Colônias individuais foram selecionadas, DNA de plasmídeo foi isolado e digerido com a enzima de restrição *Mfe* I (New England Biolabs, Beverly, MA). Clones de plasmídeo esperados foram diagnosticados pela presença de um fragmento de DNA inserido de 2.674 bp (*AAD-1* PTU), além do vetor pCR®2.1 de 3,9 kbp. Reações de sequenciamento de fita dupla de clones de plasmídeo foram realizadas conforme descrito pelo fabricante usando o kit CEQ™ DTCS-Quick Start (Beckman-Coulter, Palo Alto, CA). As reações foram purificadas usando Performa DTR Gel Filtration Cartridges (Edge BioSystems, Gaithersburg, MD), conforme descrito pelos protocolos do fabricante. Reações de sequência foram analisadas sobre um Beckman-Coulter CEQ™ 2000 XL DNA Analysis System e caracterização de nucleotídeo realizada usando o SEQUENCHER versão 4.1.4 (Gene Codes Corporation, Ann Arbor, MI).

[00193] O fragmento restrito de um clone que continha a sequência de 2.674 bp correta foi excisado em gel e purificado de acordo com as orientações do fabricante usando o kit QIAquick Gel Extraction (QIAGEN Inc., Valencia, CA). Esse fragmento foi, então, combinado em uma reação de ligação com pDAB7489 purificado (parte principal de plasmídeo) o qual tinha sido digerido com a enzima de restrição *Mfe* I e subsequentemente desfosforilado. Ligação foi realizada sob as seguintes condições: proporção de vetor:inserto de 1:5 e 500 unidades de DNA Ligase T4 (Invitrogen Life Technologies, Carlsbad, CA) em um volume de reação de 20 uL sob condições de 16 horas de incubação em um banho de água a 16°C. Cinco (5) uL da reação de ligação foram subsequentemente transformados

em células *E. coli* MAX EFFICIENCY® DH5a™ quimicamente competentes (Invitrogen Life Technologies, Carlsbad, CA) e colocados em lâminas sob condições de seleção descritas pelo fabricante.

[00194] Colônias individuais foram selecionadas, DNA de plasmídeo foi isolado e digerido com a enzima de restrição *Mfe* I (New England Biolabs, Beverly, MA). Os clones de plasmídeo esperados continham fragmentos de DNA de 2.674 bp (AAD-1 PTU) e 5.413bp (vetor pDAB7489). O plasmídeo resultante foi denominado pDAB7490 (Figura 5).

[00195] Culturas de célula embriogênica da variedade de milho Hi-II (Armstrong *et al* (1991) *Maize Genet Crop Newsletter* 65: 92-93) foram geradas, mantidas e submetidas à simultânea transformação de plasmídeos que codificam ZFN24 e a molécula doadora pDAB7490. A transformação e seleção de tecido de caule e subsequente regeneração de transformantes são descritas no Pedido de Patente U.S. No. 12/001.939, referência a qual é incorporada aqui por referência na íntegra. Para orientação adicional com relação a protocolos de transformação e seleção vide Petolino *et al.* (2000) *Plant Cell Rept.* 19: 781-786. Após antese, as plantas foram autopolinizadas ou cruzadas com a variedade de milho DAS5XH751. Semente de prole resultante foi colhida e seca. A regeneração de caule em plantas de milho férteis intactas é descrita no Pedido de Patente U.S. No. 12/001.939, particularmente Exemplo 22, referência a qual é aqui incorporada por referência na íntegra.

B. Integração objetivada do cassete do gene *AAD-1* no *locus* de *Zp15*

[00196] Dos eventos tolerantes a herbicida contendo uma molécula de DNA doadora integrada que codifica um cassete gênico de tolerância a herbicida, espera-se que alguma proporção dos referidos eventos seja o produto de integração objetivada de DNA doador no sítio de ruptura de fita dupla induzida por ZFN. De forma a diferenciar esses eventos de in-

tegração objetivada daqueles derivados de integração aleatória do cassete gênico de tolerância a herbicida, uma estratégia de genotipificação baseada em PCR usando uma combinação de iniciadores de PCR genoma-específicos e subsequentemente genoma-específicos mais doador-específicos foi utilizada.

[00197] Genotipificação diferencial de integração objetivada versus aleatória do transgene *AAD-1* em todos os eventos transformados tolerantes a herbicida foi realizada usando ensaios de PCR específicos ao *locus* de *Zp15* e ao gene *AAD-1*. Nos exemplos apresentados aqui, todos os iniciadores de oligonucleotídeo foram sintetizados pela Integrated DNA Technologies, Inc. (Coralville, LA) sob condições de dessalinização padrão e diluídos com água até uma concentração de 100 μM. O conjunto a seguir de iniciadores de oligonucleotídeo para a frente e reverso foi projetado para se anelar a sequências de gDNA específicas para o gene *Zp15* alvo que reside fora dos limites das sequências de DNA doadoras:

HB501f: 5'- AAGGTCCCAAATCTGAGGCATACTGTTGCT -3' (SEQ ID NO:103)
HB502r: 5'- GAGGTCCTATGCTTTGTCTATAGTCGGCAG -3' (SEQ ID NO:104)

[00198] Um segundo conjunto de iniciadores de oligonucleotídeo para a frente e reverso foi também criado para se anelar à sequência de gDNA específica para o gene-alvo *Zp15* fora dos limites das sequências doadoras de DNA, ainda que agrupadas dentro do primeiro par:

HB503 f 5'- GGCATACTGTTGCTGCCCTGCTGGAA -3' (SEQ ID NO:105)
HB504r 5'- GACACCTATAATCGATGTAGAGCCGAAGAG -3' (SEQ ID NO:106)

[00199] Iniciadores de oligonucleotídeo para a frente e reverso foram adicionalmente criados para se anelar especificamente ao DNA doador correspondendo à região de codificação do gene de tolerância a herbicida *AAD-1*:

HB505f 5'- AGTCCACCCCAGTGATCTCAGCACCA -3' (SEQ ID NO:107)
HB506f 5'- AGTGGCTGGACAGCTATTCTCTCAAAGCGT -3' (SEQ ID NO:108)
HB507r 5'- ACGCTTTGAGAGAATAGCTGTCCAGCCACT -3' (SEQ ID NO:109)
HB508r 5'- TGGTGCTGAGATCACTGGGGTGGACT -3' (SEQ ID NO:110)

[00200] Duas reações de amplificação primária distintas foram realizadas utilizando iniciadores que se ligam à região genômica de *Zp15* e à molécula doadora, dando origem a um amplicon que abrange o limite de integração entre o genoma e o doador. A primeira reação se focalizou sobre o limite 5'- entre o genoma e o doador e usou o conjunto de iniciadores HB501f e HB507r. A segunda reação se focalizou sobre o limite 3'- entre o doador e o genoma e usou o conjunto de iniciadores HB505f e HB502r. DNA genômico foi isolado dos eventos de milho Hi-II transformados. Reações de amplificação por PCR primária foram realizadas usando reagentes fornecidos com o kit LA TAQ PCR (*TaKaRa* Biotechnology Inc., Otsu, Shiga, Japão). O coquetel de reação de PCR consistia em: 2,5 µL de 10X Tampão La Taq PCR™, 40-200 ng de molde de DNA genômico fita dupla, iniciador de oligonucleotídeo para a frente a 10 µM, iniciador de oligonucleotídeo reverso a 10 µM, 2 µL de mistura de dNTP (2,5 mM cada), 16,25 µL de H_2O, 0,25 µL (1,25 unidades) de DNA polimerase *LA Taq*™. As reações de PCR foram realizadas usando um Bio-Rad, 96-Sample DNA Engine Tetrad2, Peltier Thermal Cycler (Hercules, CA) sob as seguintes condições de ciclo: 94°C, 2 min/1 ciclo; 94°C, 30 seg, 62°C, 30 seg, 68°C, 5 min/30 ciclos; 4°C/manutenção.

[00201] Os produtos da reação de PCR primária foram subsequentemente diluídos a 1:100 em H_2O e usados como molde de DNA para duas reações de PCR secundárias distintas. As reações secundárias também utilizam iniciadores que se ligam na região genômica de *Zp15* e à molécula doadora, dando origem a um amplicon que abrange o limite de integração entre o genoma e o doador. A identidade dos iniciadores específicos determina se a amplificação está focalizada sobre o limite 5'- ou 3'-

entre o genoma e o doador. A primeira reação se focalizou sobre o limite 5'- entre o genoma e o doador e usou o conjunto de iniciadores HB503f e HB508r. A segunda reação se focalizou sobre o limite 3'- entre o doador e o genoma e usou o conjunto de iniciadores HB506f e HB504r. Ambas as reações consistiam no seguinte: 2,5 µL de 10X Tampão La Taq PCR™, 1 µL de molde [diluição a 1:50 da 1ª Reação de PCR], iniciador de oligonucleotídeo para a frente a 10 µM, iniciador de oligonucleotídeo reverso a 10 µM, 2 µL de mistura de dNTP (2,5 mM cada), 16,25 µL de H_2O, 0,25 µL (1,25 unidades) de DNA polimerase *LA Taq*™. As reações de PCR foram realizadas usando um Bio-Rad, 96-Sample DNA Engine Tetrad2, Peltier Thermal Cycler (Hercules, CA) sob as seguintes condições de ciclo: 94°C, 2 min/1 ciclo; 94°C, 0 seg, 62°C, 30 seg, 68°C, 5 min/30 ciclos; 4°C/manutenção. Fragmentos de amplificação por PCR esperados de 2.180 bp para o limite 5'- ou 2.980 bp para o limite 3'- foram observados.

[00202] Esses fragmentos foram excisados em gel e purificados de acordo com as orientações do fabricante usando o kit QIAquick Gel Extraction (QIAGEN Inc., Valencia, CA). Fragmentos purificados foram subsequentemente clonados no plasmídeo pCR2.1 usando o kit TOPO TA CLONING® (com o vetor pCR®2.1) e células de *E. coli* quimicamente competentes ONE SHOT® TOP 10 (Invitrogen Life Technologies, Carlsbad, CA) de acordo com o protocolo do fabricante.

[00203] Clones individuais foram selecionados, DNA de plasmídeo foi isolado e digerido com a enzima de restrição *Eco* RI (New England Biolabs, Beverly, MA). Clones de plasmídeo esperados foram diagnosticados pela presença dos fragmentos de DNA inseridos de tamanho apropriado, além do vetor pCR®2.1 de 3,9 kbp.

[00204] Reações de sequenciamento de fita dupla dos clones de plasmídeo foram realizadas e caracterização e alinhamento de nucleotídeo foram realizados usando o SEQUENCHER™ versão 4.1.4 (Gene Codes Corporation, Ann Arbor, MI).

[00205] Dados de sequência selecionados derivados de um evento de integração objetivado (evento #147) do cassete gênico doador de *AAD-1* inserido no gene alvo *Zp15* são mostrados no alinhamento da Figura 6.

[00206] Produtos de PCR cuja amplificação primária se focalizou sobre o limite 5' ou 3' entre o genoma e o doador foram submetidos à amplificação secundária também focalizada sobre o limite 5'- ou 3'- entre o genoma e o doador. Alinhamento de fragmentos clonados correspondendo a esses produtos de amplificação secundária com a sequência genômica de *Zp15* do tipo silvestre, bem como a sequência esperada de um evento de integração objetivada, indica claramente que a integração precisa de DNA doador no sítio-alvo ocorreu. A sequência de nucleotídeo do *locus* genômico de *Zp15*, do limite genoma/doador, sequência de nucleotídeo das regiões doadoras correspondendo a flancos de homologia ao *Zp15* e a sequência de nucleotídeo do cassete de tolerância a herbicida foram todas preservadas em múltiplos produtos de PCR clonados derivados desse evento. Portanto, esse evento representa um genoma no qual o reparo homologia-acionado de uma ruptura de fita dupla mediada por ZFN e integração objetivada de um DNA doador em um alvo gênico específico ocorreram. Na Figura 6, foram mostrados dados de alinhamento de sequência derivados de um único caule de milho transformado isolado representativo (evento #147).

[00207] Eventos transformados adicionais que representam ocorrências de integração objetivada únicas foram obtidos, demonstrando que os métodos ensinados aqui são reproduzíveis em caule de milho.

Exemplo 5: Integração objetivada de PAT no locus de *Zp15*

[00208] Como uma outra exemplificação de métodos para integração objetivada de polinucleotídeos exógenos selecionados em *loci* objetivados divulgados aqui, moléculas de DNA doadoras adicionais trazendo um cassete gênico autônomo que codifica *PAT* foram criadas e construídas para testar a integração de uma sequência doadora exógena dentro do

lócus-alvo de *Zp15* endógeno. ZFNs com atividade de clivagem específica do gene alvo *Zp15* endógeno foram empregadas para criar um DSB nesse *locus*. Moléculas de DNA doadoras trazendo um cassete gênico autônomo que codifica *PAT* de *Streptomyces viridochromogenes* e sequências de flanqueamento homólogas ao *Zp15* foram subsequentemente integradas no DSB induzido por ZFN do alvo *Zp15*. O *PAT* codifica a enzima acetil transferase de fosfinotricina e confere resistência ao composto herbicida fosfinotricina (PPT) através de acetilação (Patente U.S. No. 5.633.434). a fosfinotricina é o ingrediente ativo dos herbicidas LIBERTY, BASTA e IGNITE. A sequência de codificação de *PAT* foi construída como uma unidade de transcrição de planta (Plant Transcription Unit - PTU) e continha uma sequência promotora derivada de actina de *O. sativa* (números de acesso ao GenBank S44221 e X63830) e uma sequência terminadora derivada de lipase de Z. *mays* (número de acesso ao GenBank L35913).

A. Construção de molécula de DNA doadora

[00209] Um construto doador de *Zp15* contendo regiões de homologia ao *Zp15* foi gerado sinteticamente como segue. Uma região de homologia ao *Zp15* compreendendo os nucleotídeos 4595-5346 (sequência de homologia 5'; SEQ ID NO: 111) e nucleotídeos 21-796 (sequência de homologia 3'; SEQ ID NO: 112) a partir do pDAB7489 foi projetado. Essa região de homologia incluía sítios de clonagem *Mfe* I entre os elementos 5' e 3' - homólogos e sítios de restrição *Not* I nas extremidades 5' e 3'. Essa sequência de DNA (SEQ ID NO: 113) foi sintetizada e inserida no plasmídeo do tipo ColE1 resistente à canamicina, pMK nos sítios de clonagem *Sac* I e *Kpn* I (Gene Art Ag, Regensburg, Alemanha). O plasmídeo doador resultante foi designado pDAB104101 (Figura 7) e contém flancos de homologia para o gene *Zp15* a fim de permitir a integração de qualquer sequência de DNA de interesse no sítio-alvo correspondente do gene *Zp15*.

[00210] Um cassete de expressão de gene de tolerância a herbicida compreendendo um promotor contento PTU completo, gene de tolerância a herbicida e sequências de término de poliadenilação (poliA) foi construído. A sequência do promotor é derivada de actina 1 de *O. sativa* (McElroy *et al.* (1990) Plant Cell 2: 163-171). Acesso ao GenBank S44221 e Acesso ao GenBank X63830). O gene de tolerância a herbicida era o *PAT.* As sequências terminadoras são derivadas de lipase de *Z. mays* (clone de cDNA de lipase de milho com Número de Acesso ao GenBank L35913). Essa sequência de milho compreende a região de término de transcrição/região não traduzida 3' para o gene *PAT.* O cassete de expressão de gene *PAT* de tolerância a herbicida é mostrado na Figura 8.

[00211] O cassete do gene *PAT* foi amplificado de um plasmídeo pDAB102256 através de PCR usando iniciadores sintetizados pela Integrated DNA Technologies, Inc. (Coralville, IA):

DC001 5' – CCAGTGCAATTGGGTCATTCATATGCTTGAGAAG – 3'
(SEQ ID NO:114)

DC002 5' – CCAGTGCAATTGAATTCAGCACTTAAAGATCTTTAG – 3'
(SEQ ID NO:115)

[00212] Reações de amplificação por PCR foram realizadas usando DNA Polimerase PHUSION™ (New England Biolabs, Beverly, MA) sob as seguintes condições de ciclo: 98°C, 30 seg/1 ciclo; 98°C 10 seg, 60°C, 20 seg, 72°C, 45 seg/9 ciclos; 98°C, 10 seg, 72°C, 60 seg/24 ciclos; 72°C, 10 min/1 ciclo; 4°C/manutenção. A reação de PCR foi submetida à eletroforese em um gel de agarose TAE a 1,0%.

[00213] Produtos de PCR esperados foram diagnosticados pela presença de um fragmento de DNA de 2,3 kbp (*PAT* PUT). Esse fragmento foi excisado em gel e purificado de acordo com as orientações do fabricante usando o kit QIAquick Gel Extraction (QIAGEN Inc., Valencia, CA). Fragmentos purificados foram, então, clonados no vetor pCR-BluntII-TOPO usando o kit Zero Blunt TOPO PCR Cloning e células de *E. coli*

quimicamente competentes One Shot® TOP 10 (Invitrogen Life Technologies, Carlsbad, CA) de acordo com o protocolo do fabricante.

[00214] Colônias individuais foram selecionadas e o DNA de plasmídeo foi isolado e submetido à digestão com enzima de restrição e análise de sequência. Os insertos de *PAT* clonados foram sequenciados para demonstrar a identidade e fidelidade de sequência dos produtos de PCR clonados (SEQ ID NO: 116). Um de tais plasmídeos foi designado pDAB104107 (Figura 9) e foi subsequentemente usado como a fonte do cassete do gene *PAT* para inserção nos novos vetores doadores de homologia de *Zp15*.

[00215] O fragmento do gene *PAT* de 2,3 kbp foi recuperado do pDAB104107 através de digestão com *Mfe* I, seguido por eletroforese em gel, excisão e purificação. O plasmídeo doador de homologia de *Zp15*, pDAB104101, foi também digerido com *Mfe* I e purificado em gel.Ligação do fragmento do gene *PAT* no pDAB104101 proporcionou clones nos quais o gene *PAT* foi inserido no sítio *Mfe* I em qualquer uma de duas orientações com relação às sequências do gene *Zp15*, conforme determinado por meio de digestão com enzima de restrição diferencial. O pDAB104104 (Figura 10) compreendia o gene *PAT* inserido na mesma orientação transcricional que o gene *Zp15*. O pDAB104105 (Figura 11) compreendia o gene *PAT* inserido na orientação oposta com relação ao gene *Zp15*.

B. Construção de molécula de DNA doadora adicional

[00216] Outro construto doador contendo regiões de homologia ao *Zp15* foi gerado, no qual a sequência de homologia 3' de *Zp15* no pDAB7489 foi alterada por meio de truncamento, enquanto que a sequência de homologia 5' de *Zp15* no pDAB7489 permaneceu a mesma.

[00217] A região de homologia 3' truncada foi gerada a partir do pDAB7489 através de PCR usando iniciadores sintetizados pela Integrated DNA Technologies, Inc. (Coralville, LA):

DC003 5' – CTAATCGTCGACTCGTCAAGCCCCGCCTTTAAAT – 3'
(SEQ ID NO:117)

DC004 5' – CTAATCCAATTGGTGTGGGCAGCCGAGCG – 3'
(SEQ ID NO:118)

[00218] Reações de amplificação por PCR foram realizadas usando DNA Polimerase PHUSION HOT START (New England Biolabs, Beverly, MA) sob as seguintes condições de ciclo: 98°C, 30 s/1 ciclo; 98°C 10 seg, 72°C, 15 seg/33 ciclos; 72°C, 5 min/1 ciclo; 4°C/manutenção. A reação de PCR foi submetida à eletroforese em um gel de agarose TAE a 1,0%.

[00219] Produtos de PCR esperados foram diagnosticados pela presença de um fragmento de DNA de 0,8 kbp (homologia-3' de *Zp15*). Esse fragmento foi excisado em gel e purificado de acordo com as orientações do fabricante usando o kit QIAquick Gel Extraction (QIAGEN Inc., Valencia, CA). Fragmentos purificados foram, então, clonados no vetor pCR-BluntII-TOPO usando o kit Zero Blunt TOPO PCR Cloning e células de *E. coli* quimicamente competentes One Shot® TOP 10 (Invitrogen Life Technologies, Carlsbad, CA) de acordo com o protocolo do fabricante.

[00220] Colônias individuais foram selecionadas e submetidas a isolamento de DNA de plasmídeo e confirmação por meio de digestão com enzima de restrição. Insertos de homologia 3' de *Zp15* foram sequenciados para demonstrar a identidade e fidelidade de sequência dos produtos de PCR clonados (SEQ ID NO: 119). Um de tais plasmídeos foi designado pDAB104106 (Figura 12) e foi subsequentemente usado como a fonte da nova sequência de homologia 3' de *Zp15* para substituição no pDAB7489 a fim de criar um novo vetor doador de homologia de *Zp15*.

[00221] pDAB7489 foi digerido sequencialmente com *Mfe* I e *Sal* I e o fragmento de vetor de 4,2 kbp foi purificado em gel. O pDAB104106 foi também digerido com *Mfe* I e *Sal* I e o fragmento de 0,8 kbp compreendendo a sequência de homologia 3' de *Zp15* truncada foi purificado em gel. Ligação e transformação desses fragmentos purificados em gel proporcionaram clones nos quais a sequência de homologia 3' de *Zp15* trun-

cada foi substituída pela sequência de homologia 3' de *Zp15* original. Esse plasmídeo foi designado pDAB104100 (Figura 13) e foi usado como o recipiente para o gene *PAT.*

[00222] O gene *PAT* foi removido do pDAB104107 por meio de digestão com *Mfe* I. Após eletroforese em gel, o fragmento gênico de *PAT* de 2,3 kbp foi purificado em gel. O plasmídeo doador de homologia de *Zp15,* pDAB104100, também foi digerido com *Mfe* I e purificado em gel. Ligação do gene *PAT* no pDAB104100 proporcionou clones nos quais o gene *PAT* estava inserido no sítio *Mfe* I em qualquer uma de duas orientações com relação ao gene *Zp15*, conforme determinado por meio de digestão com enzima de restrição diferencial. O pDAB104103 (Figura 14) compreendia o gene *PAT* inserido na mesma orientação transcricional que o gene *Zp15*. O pDAB104102 (Figura 15) compreendia o gene *PAT* inserido na orientação oposta com relação ao gene *Zp15*.

C. Transformação de milho e recuperação de inserções *Zp15-PAT* objetivadas

[00223] Culturas de célula embriogênica da variedade de milho Hi-II (Armstrong *et al* (1991) *Maize Genet Crop Newsletter* 65: 92-93) foram geradas, mantidas e submetidas à transformação simultânea de plasmídeos que codificam ZFN24 e molécula doadora. Moléculas doadoras incluem aquelas descritas aqui; pDAB104102, pD AB 104103, pD AB 104104 e pD AB 104105. A transformação e seleção de tecido de caule e subsequente regeneração de transformantes são descritas no Pedido de Patente U.S. No. 12/001.939, particularmente Exemplo 19, referência a qual é aqui incorporada por referência na íntegra. Para orientação adicional com relação ao protocolo de transformação e seleção vide Petolino *et al.* (2000) *Plant Cell Rept.* 19: 781-786. Após antese, as plantas são autopolinizadas ou cruzadas com uma variedade de milho, tal como DAS5XH751. A semente de prole resultante pode ser colhida e seca e plantas dessas sementes podem ser analisadas para demonstrar a here-

ditariedade de eventos de integração objetivados. A regeneração de caule em plantas de milho férteis intactas é descrita no Pedido de Patente U.S. No. 12/001.939, particularmente Exemplo 22, referência a qual é aqui incorporada por referência na íntegra.

D. Identificação de inserções de *PAT Zp15*-objetivadas

[00224] Inserções de *PAT Zp15*-objetivadas em tecido de caule transformado são detectadas através de PCR. Molde de DNA genômico é extraído do tecido de caule via métodos bem-conhecidos e comumente usados, tais como o kit Plant DNEASY (QIAGEN Inc., Valencia, CA) ou o método de Dellaporta (Dellaporta *et al,* (1983) *Plant Mol. Biol. Rep.* 1; 19-21). Uso de iniciadores específicos para *PAT* em conjunto com os iniciadores de sequência de flanqueamento de *Zp15* já usados para detectar integração objetivada de *AAD-1* no *locus* de *Zp15* resulta na amplificação das junções de inserção objetivada de *PAT* nas regiões de homologia 5' e 3' de *Zp15*. Os iniciadores específicos para *PAT* podem ser:

DC013 5' – CAATCGTAAGCGTTCCTAGCCTTCCAG – 3'
(SEQ ID NO:120)

DC014 5' – CTGGAAGGCTAGGAACGCTTACGATTG – 3'
(SEQ ID NO:121)

[00225] Especificamente, iniciadores HB501f ou HB503f na região genômica que flanqueia o DNA doador da sequência de homologia 5' de *Zp15* são usados em conjunto com o iniciador DC013 (SEQ ID NO: 120) na região de codificação da proteína PAT para detectar junções de inserção *PAT-Zp15* 5' quando os DNAs doadores têm o gene *PAT* na orientação direta com relação ao gene *Zp15*. Da mesma forma, os iniciadores HB501f ou HB503f são usados em conjunto com o iniciador DC014 (SEQ ID NO: 121) na região de codificação da proteína PAT para detectar junções de inserção *PAT-Zp15* 5' quando os DNAs doadores têm o gene *PAT* na orientação indireta com relação ao gene *Zp15*. Para detecção de junções de inserção *PAT-Zp15* 3', os iniciadores HB502r ou HB504r são usados em conjunto com o iniciador DC013 (SEQ ID NO: 120) para de-

tectar junções de inserção quando os DNAs doadores têm o gene *PAT* na orientação indireta com relação ao gene *Zp15*. Da mesma forma, os iniciadores HB502r ou HB504r são usados em conjunto com o iniciador DC014 (SEQ ID NO: 121) para detectar junções de inserção *PAT-Zp15* 3' quando os DNAs doadores têm o gene *PAT* na orientação direta com relação ao gene *Zp15*.

[00226] Reações de amplificação por PCR foram realizadas usando DNA Polimerase PHUSION HOT START (New England Biolabs, Beverly, MA) sob as seguintes condições de ciclo: 98°C, 30 s/1 ciclo; 98°C 10 seg, 72°C, 15 seg/33 ciclos; 72°C, 5 min/1 ciclo; 4°C/manutenção. Os produtos de PCR foram decompostos e identificados usando eletroforese em gel de agarose TAE. Os tamanhos esperados de fragmento do gel para os produtos de PCR de eventos de integração objetivada *PAT - Zp15* em caule transgênico gerados usando os diferentes DNAs doadores de *PAT-Zp15* são como segue:

HB501f/HB503f + DC013 (5') = 2,6 kbp (pDAB104103), 2,6 kbp (pDAB 104104)

HB501f/HB503f + DC014 (5') = 1,6 kbp (pDAB104105), 1,7 kbp (pDAB 104102)

HB502r/HB504r + DC013 (3') = 3,2 kbp (pDAB104105), 3,2 kbp (pDAB 104102)

HB502r/HB504r + DC014 (3') = 1,2 kbp (pDAB104103), 1,2 kbp (pDAB 104104)

[00227] Os produtos de PCR compreendendo as junções de integração objetivada *PAT - Zp15* 5' e 3' são clonados e sequenciados usando métodos padrões conhecidos por aqueles versados no campo. Por exemplo, os produtos de PCR são purificados do gel de agarose e clonados no plasmídeo pCR-BluntII TOPO usando o kit TOPO BLUNT CLONING® e células de *E. coli* quimicamente competentes ONE SHOT® TOP 10 (Invitrogen Life Technologies, Carlsbad, CA) de acordo com o

protocolo do fabricante.

[00228] As junções de integração clonadas são, então, sequenciadas para demonstrar que o gene *PAT* está inserido no genoma de milho no *locus* de *Zp15* por meio de recombinação homóloga via as sequências de homologia de *Zp15* 5' e 3' que são incorporadas nos vetores de transformação doadores. Além do vetor TOPO, iniciadores específicos M13 para a frente e M13 reverso, iniciadores específicos ao cassete do gene *PAT* são usados para obter sequência completa dos clones de integração objetivados. Iniciadores de PCR SEQ ID NO: 120 e SEQ ID NO: 121, os quais são específicos à sequência de codificação da proteína PAT, também são usados como iniciadores de sequenciamento. Além disso, outros iniciadores também podem ser usados para sequenciamento. Esses incluem, mas não estão limitados a, aqueles que são específicos para o elemento promotor actina de arroz do cassete de PAT:

DC-S1 5'-CCAACTGGACAATAGTCTCCAC–3'
(SEQ ID NO:122)

DC-S2 5'–CATCGCCACTATATACATACC–3'
(SEQ ID NO:123)

[00229] e aqueles que são específicos para a sequência de codificação da proteína PAT:

DC-S3 5'-CGTCTCAATGTAATGGTTAACG-3'
(SEQ ID NO:124)

DC-S4 5'-GCCCAGCGTAAGCAATACCAG-3'
(SEQ ID NO:125)

Exemplo 6: Hereditariedade de integração objetivada de *AAD-1* no *locus* de *Zp15*

[00230] Um evento com caule transgênico trazendo o cassete do gene *AAD-1* objetivado ao *locus* de *Zp15* foi gerado (Evento 138). Plantas T_0 do Evento 138 foram regeneradas e cruzadas como fêmeas com os machos DAS5XH751. Sementes T_1 resultantes foram plantadas e plantas T_1 foram crescidas.

[00231] Plantas T_1 foram analisadas através de PCR para demonstrar a ocorrência da integração objetivada *AAD-1 - Zp15*. Reações de amplificação por PCR foram realizadas usando DNA Polimerase PHUSION HOT START (New England Biolabs, Beverly, MA) sob as seguintes condições de ciclo: 98°C, 30 s/1 ciclo; 98°C 10 seg, 72°C 15 seg/33 ciclos; 72°C, 5 min/1 ciclo; 4°C/manutenção. As reações de PCR foram analisadas através de eletroforese em géis de agarose TAE a 1,0%. DNA genômico foi extraído de uma planta Ti e usado como molde de DNA em reações de PCR com iniciadores agrupados criados para detectar a junção de integração *AAD-1 / Zp15* 5'. Reações de PCR primárias e secundárias foram conduzidas usando os mesmos iniciadores usados antes na análise do evento 147 com caule. Para PCR primária, os iniciadores HB501f e HB507r foram usados. PCR primária proporcionou uma banda no tamanho esperado de 2,2 kbp. Uma alíquota da reação de PCR primária foi diluída a 1:100 e usada como molde em PCR secundária usando iniciadores agrupados HB503f e HB508r. PCR secundária também proporcionou uma banda no tamanho esperado de 2,2 kbp.

[00232] O produto de PCR secundário de 2,2 kbp foi clonado no pCR-BluntII-TOPO usando o kit ZERO BLUNT TOPO PCR Cloning e células de *E. coli* quimicamente competentes ONE SHOT TOP 10 (Invitrogen Life Technologies, Carlsbad, CA) de acordo com o protocolo do fabricante. O DNA clonado foi sequenciado usando iniciadores específicos ao vetor de flanqueamento (M13 para a frente e M13 reverso). Descobriu-se que a sequência era idêntica àquela esperada para uma integração objetivada de *AAD-1* no *locus* de *Zp15*.

[00233] Todas as patentes, pedidos de patente e publicações mencionados aqui são incorporados por referência na íntegra para todas as finalidades.

[00234] Embora divulgação tenha sido fornecida em alguns detalhes à guisa de ilustração e exemplo para fins de clareza de compreensão, será

evidente para aqueles versados no campo que várias alterações e modificações podem ser praticadas sem se desviar do espírito ou escopo da divulgação. Consequentemente, as descrições e exemplos precedentes não deverão ser construídos como limitativos.

REIVINDICAÇÕES

1. Método de integração de uma ou mais sequências de ácido nucleico exógenas no genoma de uma célula vegetal, caracterizado pelo fato de que compreende:

transformar a célula vegetal com um ou mais polinucleotídeos que codificam uma nuclease que produz uma clivagem de fita dupla no genoma da célula vegetal em um *locus* beta zein de15kD Zp15 (Zp15), a nuclease compreendendo um domínio de ligação de dedo de zinco que se liga a um sítio alvo no *locus* Zp15, o domínio de ligação de dedo de zinco compreendendo 4 dedos de zinco ordenados e designados F1 a F4, cada dedo de zinco compreendendo uma região de reconhecimento, em que as regiões de reconhecimento do domínio de ligação de dedo de zinco e o sítio alvo são mostrados em linhas individuais da Tabela a seguir:

Sítio Alvo	**F1**	**F2**	**F3**	**F4**
cgGGGCTGCAGGGCttg-tacggcgctgg (SEQ ID NO:11)	DRSHLTR (SEQ ID NO:12)	RSDNLRE (SEQ ID NO:13)	RSDVLSE (SEQ ID NO:14)	RSAHLSR (SEQ ID NO:15)
gcAGGGGCAGGG-CAtctgcattgcagag (SEQ ID NO:16)	QSGSLTR (SEQ ID NO:17)	RSDHLTQ (SEQ ID NO:18)	DRSHLTR (SEQ ID NO:12)	RSDHLTQ (SEQ ID NO:18)
ctGAGGCAGCCGCAg-tgcagcccgctgg (SEQ ID NO:19)	QSGDLTR (SEQ ID NO:20)	DRSDLSR (SEQ ID NO:21)	QSGDLTR (SEQ ID NO:20)	RSDNLTR (SEQ ID NO:22)
acTCCGCGTAGGGGta-cagcccgccggc (SEQ ID NO:23)	RSDHLSR (SEQ ID NO:24)	RSDNLTT (SEQ ID NO:25)	RSDDLTR (SEQ ID NO:26)	DSSDRKK (SEQ ID NO:27)
acTCCGCGTAGGGGta-cagcccgccggc (SEQ ID NO:23)	RSDHLSE (SEQ ID NO:29)	RNDNRKN (SEQ ID NO:30)	RSDDLTR (SEQ ID NO:26)	DSSDRKK (SEQ ID NO:27)
agCCGCAGtGCAG-CCcgctggcggcggc (SEQ ID NO:31)	DRSDLSR (SEQ ID NO:21)	QSSDLTR (SEQ ID NO:32)	RSDHLSE (SEQ ID NO:29)	TSSTRKT (SEQ ID NO:33)
agCCGCAGtGCAG-CCcgctggcggcggc (SEQ ID NO:31)	DRSDLSR (SEQ ID NO:21)	QSSDLTR (SEQ ID NO:32)	RSDHLSE (SEQ ID NO:29)	RSSTRKE (SEQ ID NO:35)
ccTCAGGTACTCCGcg-taggggtacagc (SEQ ID NO:36)	RSDTLSE (SEQ ID NO:37)	ARSTRTN (SEQ ID NO:38)	QSSHLTR (SEQ ID NO:39)	QSADRTK (SEQ ID NO:40)
gcCCGCTGGCGGCGg-cgccctactacgc (SEQ ID NO:41)	RSDDLTR (SEQ ID NO:26)	RSDDLTR (SEQ ID NO:26)	RSDTLSA (SEQ ID NO:42)	RNQDRKT (SEQ ID NO:43)
gcACTGCGGCTGCCt-caggtactccgcg (SEQ ID NO:44)	DRSDLSR (SEQ ID NO:21)	QSSDLRR (SEQ ID NO:45)	RSDDLTR (SEQ ID NO:26)	QSSDLTR (SEQ ID NO:32)

cgCCGGGTGTGGGCa-gccgagcgccatg (SEQ ID NO:46)	DRSHLSR (SEQ ID NO:47)	RSDALAR (SEQ ID NO:48)	QSSHLTR (SEQ ID NO:39)	RSDDRKT (SEQ ID NO:49)
cgCCGGGTgTGGGCAg-ccgagcgccatg (SEQ ID NO:46)	QSGSLTR (SEQ ID NO:17)	RSDHLTT (SEQ ID NO:51)	QSSHLTR (SEQ ID NO:39)	RSDDRKT (SEQ ID NO:49)
gcCGGGTGTGGGCAg-ccgagcgccatgt (SEQ ID NO:52)	QSGSLTR (SEQ ID NO:17)	RSDHLTT (SEQ ID NO:51)	RSDSLLR (SEQ ID NO:53)	RSDNLRE (SEQ ID NO:13)
agTAGGGCGCCGCCg-ccagcgggctgca (SEQ ID NO:54)	RSDNLTT (SEQ ID NO:25)	DRSDLSR (SEQ ID NO:21)	DRSHLTR (SEQ ID NO:12)	RSDNLTT (SEQ ID NO:25)
tgTGGGCAGCCGAGcg-ccatgttccagc (SEQ ID NO:55)	RSDNLAR (SEQ ID NO:56)	DRSDLSR (SEQ ID NO:21)	QSGSLTR (SEQ ID NO:17)	RSDHLTT (SEQ ID NO:51)
tgTGGGCAGCCGAGcg-ccatgttccagc (SEQ ID NO:55)	RSDNLSR (SEQ ID NO:58)	DNSTRKT (SEQ ID NO:59)	QSGSLTR (SEQ ID NO:17)	RSDHLTT (SEQ ID NO:51)
cgGCGTAGTAGGGCg-ccgccgccagcgg (SEQ ID NO:60)	DRSHLTR (SEQ ID NO:12)	RSDNLTT (SEQ ID NO:25)	RSDNLST (SEQ ID NO:61)	RSADLSR (SEQ ID NO:62)
cgGCGTAGtAGGGCG-ccgccgccagcgg (SEQ ID NO:60)	RSDDLTR (SEQ ID NO:26)	RSDHLTQ (SEQ ID NO:18)	RSDNLST (SEQ ID NO:61)	RSADLSR (SEQ ID NO:62)
caGCCGCTCCGGCAa-cagtgctgccagc (SEQ ID NO:64)	QSGSLTR (SEQ ID NO:17)	RSDDRKT (SEQ ID NO:49)	QSSDLSR (SEQ ID NO:65)	DRSDLSR (SEQ ID NO:21)
acATGGCGcTCGGCTg-cccacacccggc (SEQ ID NO:66)	QSSDLSR (SEQ ID NO:65)	RNDDRKK (SEQ ID NO:67)	RSDDLTR (SEQ ID NO:26)	RSDALTQ (SEQ ID NO:68)
tgGCAGCCCAGGGTct-caaccccatggc (SEQ ID NO:69)	QSSHLTR (SEQ ID NO:39)	RSDNLRE (SEQ ID NO:13)	DRSDLSR (SEQ ID NO:21)	QSSDLTR (SEQ ID NO:32)
caGCTGCTGCTGCTg-ctgcatcagagct (SEQ ID NO:70)	QSSDLSR (SEQ ID NO:65)	QSSDLRR (SEQ ID NO:45)	QSSDLSR (SEQ ID NO:65)	QSSDLRR (SEQ ID NO:45)
ctGCCCAGCTACCG-caccaacccctgtg (SEQ ID NO:28)	RSDSLSA (SEQ ID NO:34)	DNSNRIK (SEQ ID NO:50)	RSDNLSE (SEQ ID NO:57)	ASKTRKN (SEQ ID NO:63)
tgGTACTGGTAGAG-tccacccatggccg (SEQ ID NO:71)	RSDNLAR (SEQ ID NO:56)	QSGSLTR (SEQ ID NO:17)	RSDVLSE (SEQ ID NO:14)	QSGSLTR (SEQ ID NO:17)
tgGTACTGGTAGAG-tccacccatggccg (SEQ ID NO:71)	RSDNLAR (SEQ ID NO:56)	QSGSLTR (SEQ ID NO:17)	RSDALSN (SEQ ID NO:72)	TSSARTT (SEQ ID NO:73)
aaCCCCTGtGGCG-TCtccgctgccattc (SEQ ID NO:74)	DRSALSR (SEQ ID NO:75)	DRSHLAR (SEQ ID NO:76)	RSDTLSA (SEQ ID NO:42)	DRSTRTT (SEQ ID NO:77)
gtGCGGTAGCTGGGca-gctggtactggt (SEQ ID NO:78)	RSDHLSR (SEQ ID NO:24)	QSSDLRR (SEQ ID NO:45)	QSGALAR (SEQ ID NO:79)	RSDDLTR (SEQ ID NO:26)
gtGCGGTAGCTGGGca-gctggtactggt (SEQ ID NO:78)	RSDHLST (SEQ ID NO:80)	HSDTRKK (SEQ ID NO:81)	QSGALAR (SEQ ID NO:79)	RSDDLTR (SEQ ID NO:26)
agGCGGGGcTTGA-CGaagttggaagccg (SEQ ID NO:82)	RSDSLSV (SEQ ID NO:83)	QNQHRIN (SEQ ID NO:84)	RSDHLSR (SEQ ID NO:24)	RSDDLTR (SEQ ID NO:26)
aaATGGAAAAAACGc-	RSDDLSK	RNDHRKN	QRSNLVR	RSDALTQ

taaaattatgtgt (SEQ ID NO:85)	(SEQ ID NO:86)	(SEQ ID NO:87)	(SEQ ID NO:88)	(SEQ ID NO:68)
ttGTGGTGCCAACGgga-gccatgctcac (SEQ ID NO:89)	RSDTLSQ (SEQ ID NO:90)	QNATRIN (SEQ ID NO:91)	RSDALSR (SEQ ID NO:92)	RSDALAR (SEQ ID NO:48)
ttGTGGTGCCAACGgga-gccatgctcac (SEQ ID NO:89)	RSDTLSQ (SEQ ID NO:90)	QKATRIT (SEQ ID NO:93)	RSDALSR (SEQ ID NO:92)	RSDALAR (SEQ ID NO:48)
ttGTGGTGcCAACGGga-gccatgctcac (SEQ ID NO:89)	RSDHLSE (SEQ ID NO:29)	QNANRKT (SEQ ID NO:94)	RSDALSR (SEQ ID NO:92)	RSDALAR (SEQ ID NO:48)
caATCACGCCGGTAg-cggggctagttat (SEQ ID NO:95)	QSGALAR (SEQ ID NO:79)	RSDDRKT (SEQ ID NO:49)	RSDTLSQ (SEQ ID NO:90)	DSSARKK (SEQ ID NO:96)

resultando assim na integração de um polinucleotídeo compreendendo uma ou mais sequências exógenas no genoma da célula no ou próximo ao *locus* Zp15, em que a planta obtida com o referido método é excluída da proteção.

2. Método, de acordo com a reivindicação 1, caracterizado pelo fato de que compreende ainda expressar um produto da uma ou mais sequências exógenas.

3. Método, de acordo com a reivindicação 1 ou 2, caracterizado pelo fato de que a clivagem de fita dupla é feita por meio de:

(a) transformar a célula vegetal com um vetor de expressão compreendendo uma sequência que codifica uma primeira proteína de fusão na célula, a primeira proteína de fusão compreendendo o primeiro domínio de ligação de dedo de zinco compreendendo 4 dedos de zinco ordenados e designados de F1 a F4, cada dedo de zinco compreendendo uma região de reconhecimento, em que o domínio de ligação de dedo de zinco é selecionado a partir do grupo que consiste em um domínio de ligação de dedo de zinco designado 11742, 11743, 11750, 11753, 11754, 11755, 11756, 11757, 11758, 11759, 11760, 11761, 11762, 11763, 11766, 11767, 11768, 11769, 11770 , 11771, 11772, 11773, 11774, 11775, 11776, 11777, 11778, 11779, 11780, 11781, 11782, 11783, 11784 e 11785 conforme estabelecido na Tabela a seguir:

ZFP designado	**F1**	**F2**	**F3**	**F4**
11742	DRSHLTR (SEQ ID NO:12)	RSDNLRE (SEQ ID NO:13)	RSDVLSE (SEQ ID NO:14)	RSAHLSR (SEQ ID NO:15)

11743	QSGSLTR (SEQ ID NO:17)	RSDHLTQ (SEQ ID NO:18)	DRSHLTR (SEQ ID NO:12)	RSDHLTQ (SEQ ID NO:18)
11750	QSGDLTR (SEQ ID NO:20)	DRSDLSR (SEQ ID NO:21)	QSGDLTR (SEQ ID NO:20)	RSDNLTR (SEQ ID NO:22)
11753	RSDHLSR (SEQ ID NO:24)	RSDNLTT (SEQ ID NO:25)	RSDDLTR (SEQ ID NO:26)	DSSDRKK (SEQ ID NO:27)
11754	RSDHLSE (SEQ ID NO:29)	RNDNRKN (SEQ ID NO:30)	RSDDLTR (SEQ ID NO:26)	DSSDRKK (SEQ ID NO:27)
11755	DRSDLSR (SEQ ID NO:21)	QSSDLTR (SEQ ID NO:32)	RSDHLSE (SEQ ID NO:29)	TSSTRKT (SEQ ID NO:33)
11756	DRSDLSR (SEQ ID NO:21)	QSSDLTR (SEQ ID NO:32)	RSDHLSE (SEQ ID NO:29)	RSSTRKE (SEQ ID NO:35)
11757	RSDTLSE (SEQ ID NO:37)	ARSTRTN (SEQ ID NO:38)	QSSHLTR (SEQ ID NO:39)	QSADRTK (SEQ ID NO:40)
11758	RSDDLTR (SEQ ID NO:26)	RSDDLTR (SEQ ID NO:26)	RSDTLSA (SEQ ID NO:42)	RNQDRKT (SEQ ID NO:43)
11759	DRSDLSR (SEQ ID NO:21)	QSSDLRR (SEQ ID NO:45)	RSDDLTR (SEQ ID NO:26)	QSSDLTR (SEQ ID NO:32)
11760	DRSHLSR (SEQ ID NO:47)	RSDALAR (SEQ ID NO:48)	QSSHLTR (SEQ ID NO:39)	RSDDRKT (SEQ ID NO:49)
11761	QSGSLTR (SEQ ID NO:17)	RSDHLTT (SEQ ID NO:51)	QSSHLTR (SEQ ID NO:39)	RSDDRKT (SEQ ID NO:49)
11762	QSGSLTR (SEQ ID NO:17)	RSDHLTT (SEQ ID NO:51)	RSDSLLR (SEQ ID NO:53)	RSDNLRE (SEQ ID NO:13)
11763	RSDNLTT (SEQ ID NO:25)	DRSDLSR (SEQ ID NO:21)	DRSHLTR (SEQ ID NO:12)	RSDNLTT (SEQ ID NO:25)
11766	RSDNLAR (SEQ ID NO:56)	DRSDLSR (SEQ ID NO:21)	QSGSLTR (SEQ ID NO:17)	RSDHLTT (SEQ ID NO:51)
11767	RSDNLSR (SEQ ID NO:58)	DNSTRKT (SEQ ID NO:59)	QSGSLTR (SEQ ID NO:17)	RSDHLTT (SEQ ID NO:51)
11768	DRSHLTR (SEQ ID NO:12)	RSDNLTT (SEQ ID NO:25)	RSDNLST (SEQ ID NO:61)	RSADLSR (SEQ ID NO:62)
11769	RSDDLTR (SEQ ID NO:26)	RSDHLTQ (SEQ ID NO:18)	RSDNLST (SEQ ID NO:61)	RSADLSR (SEQ ID NO:62)
11770	QSGSLTR (SEQ ID NO:17)	RSDDRKT (SEQ ID NO:49)	QSSDLSR (SEQ ID NO:65)	DRSDLSR (SEQ ID NO:21)
11771	QSSDLSR (SEQ ID NO:65)	RNDDRKK (SEQ ID NO:67)	RSDDLTR (SEQ ID NO:26)	RSDALTQ (SEQ ID NO:68)
11772	QSSHLTR (SEQ ID NO:39)	RSDNLRE (SEQ ID NO:13)	DRSDLSR (SEQ ID NO:21)	QSSDLTR (SEQ ID NO:32)
11773	QSSDLSR (SEQ ID NO:65)	QSSDLRR (SEQ ID NO:45)	QSSDLSR (SEQ ID NO:65)	QSSDLRR (SEQ ID NO:45)
11774	RSDSLSA (SEQ ID NO:34)	DNSNRIK (SEQ ID NO:50)	RSDNLSE (SEQ ID NO:57)	ASKTRKN (SEQ ID NO:63)
11775	RSDNLAR (SEQ ID NO:56)	QSGSLTR (SEQ ID NO:17)	RSDVLSE (SEQ ID NO:14)	QSGSLTR (SEQ ID NO:17)
11776	RSDNLAR (SEQ ID NO:56)	QSGSLTR (SEQ ID NO:17)	RSDALSN (SEQ ID NO:72)	TSSARTT (SEQ ID NO:73)
11777	DRSALSR (SEQ ID NO:75)	DRSHLAR (SEQ ID NO:76)	RSDTLSA (SEQ ID NO:42)	DRSTRTT (SEQ ID NO:77)
11778	RSDHLSR (SEQ ID NO:24)	QSSDLRR (SEQ ID NO:45)	QSGALAR (SEQ ID NO:79)	RSDDLTR (SEQ ID NO:26)
11779	RSDHLST (SEQ ID NO:80)	HSDTRKK (SEQ ID NO:81)	QSGALAR (SEQ ID NO:79)	RSDDLTR (SEQ ID NO:26)
11780	RSDSLSV (SEQ ID NO:83)	QNQHRIN (SEQ ID NO:84)	RSDHLSR (SEQ ID NO:24)	RSDDLTR (SEQ ID NO:26)
11781	RSDDLSK	RNDHRKN	QRSNLVR	RSDALTQ (SEQ

	(SEQ ID NO:86)	(SEQ ID NO:87)	(SEQ ID NO:88)	ID NO:68)
11782	RSDTLSQ (SEQ ID NO:90)	QNATRIN (SEQ ID NO:91)	RSDALSR (SEQ ID NO:92)	RSDALAR (SEQ ID NO:48)
11783	RSDTLSQ (SEQ ID NO:90)	QKATRIT (SEQ ID NO:93)	RSDALSR (SEQ ID NO:92)	RSDALAR (SEQ ID NO:48)
11784	RSDHLSE (SEQ ID NO:29)	QNANRKT (SEQ ID NO:94)	RSDALSR (SEQ ID NO:92)	RSDALAR (SEQ ID NO:48)
11785	QSGALAR (SEQ ID NO:79)	RSDDRKT (SEQ ID NO:49)	RSDTLSQ (SEQ ID NO:90)	DSSARKK (SEQ ID NO:96)

e um primeiro meio-domínio de clivagem, em que o primeiro domínio de ligação de dedo de zinco foi projetado para se ligar a um primeiro sítio alvo, conforme estabelecido na Tabela acima em ou próximo a um *locus* Zp15 no genoma da célula vegetal; e

(b) transformar a célula vegetal com um vetor de expressão compreendendo uma sequência que codifica uma segunda proteína de fusão na célula, a segunda proteína de fusão compreendendo um segundo domínio de ligação de dedo de zinco e um segundo meio domínio de clivagem, em que o segundo domínio de ligação de dedo de zinco se liga a um segundo sítio alvo no ou próximo ao *locus* Zp15 no genoma da célula vegetal, em que o segundo sítio alvo é diferente do primeiro sítio alvo; e em que a ligação da primeira proteína de fusão ao primeiro sítio alvo e a ligação da segunda proteína de fusão ao segundo sítio alvo posiciona os meios-domínios de clivagem de modo que o genoma da célula vegetal em ou próximo ao *locus* Zp15 é clivado.

4. Método, de acordo com qualquer uma das reivindicações 1 a 3, caracterizado pelo fato de que os meio-domínios de clivagem ocorrem naturalmente ou não ocorrem naturalmente.

5. Método, de acordo com a reivindicação 1 ou 2, caracterizado pelo fato de que a nuclease de dedo de zinco compreende um par de nucleases de dedo de zinco.

6. Método, de acordo com a reivindicação 1 ou 2, caracterizado pelo fato de que a uma ou mais sequências de ácido nucleico exógenas compreende uma sequência de codificação.

7. Método, de acordo com a reivindicação 6, caracterizado pelo fato de que a sequência de codificação codifica um produto que confe-

re: resistência a herbicida; tolerância a herbicida; resistência a insetos; tolerância a insetos; resistência a doenças; tolerância a doenças; rendimentos aumentados de óleo ou combinações dos mesmos.

8. Método, de acordo com qualquer uma das reivindicações 1 a 7, caracterizado pelo fato de que o polinucleotídeo compreende ainda sequências de nucleotídeo que são homólogas às sequências no *locus Zp15*.

9. Método, de acordo com a reivindicação 8, caracterizado pelo fato de que as sequências de nucleotídeo homólogas flanqueiam a sequência exógena.

10. Método, de acordo com qualquer uma das reivindicações 1 a 9, caracterizado pelo fato de que o polinucleotídeo compreende ainda um promotor.

11. Método, de acordo com qualquer uma das reivindicações 1 a 10, caracterizado pelo fato de que uma ou mais das sequências exógenas integradas são transmitidas à progênie em gerações subsequentes.

12. Método, de acordo com qualquer uma das reivindicações 1 a 11, caracterizado pelo fato de que a célula vegetal é uma célula vegetal monocotiledônea.

13. Método, de acordo com a reivindicação 12, caracterizado pelo fato de que a célula vegetal é uma célula de milho.

Zp15 genômico: GTGCAGCCCGCTGGCGGCGGCGCCCTACTACGCCG::::::GGTGTGGGCAGCCGAGCGCCATGTTCCAGCCGCTC

Lâmina 7 B10: GTGCAGCCCGCTGGCGGCGGCGCCCTACTACGCCG**GCAATT**GGTGTGGGCAGCCGAGCGCCATGTTCCAGCCGCTC (+6)

11768 11766

FIG. 1

Zp15 genômico ATGGGCGCCGGCGGGCTGTACCCCTACGCGGAGTACCTGAGGCAGCCGCAGTGCAGCCCGCTGGCGGCGGCGCCCT

Lâmina 6 C12 ATGGGCGCCGGCGGGCTGTACCCCTACGCGGAG:::CTGAGGCAGCCGCAGTGCAGCCCGCTGGCGGCGGCGCCCT (Δ3)

11753 11750

FIG. 2

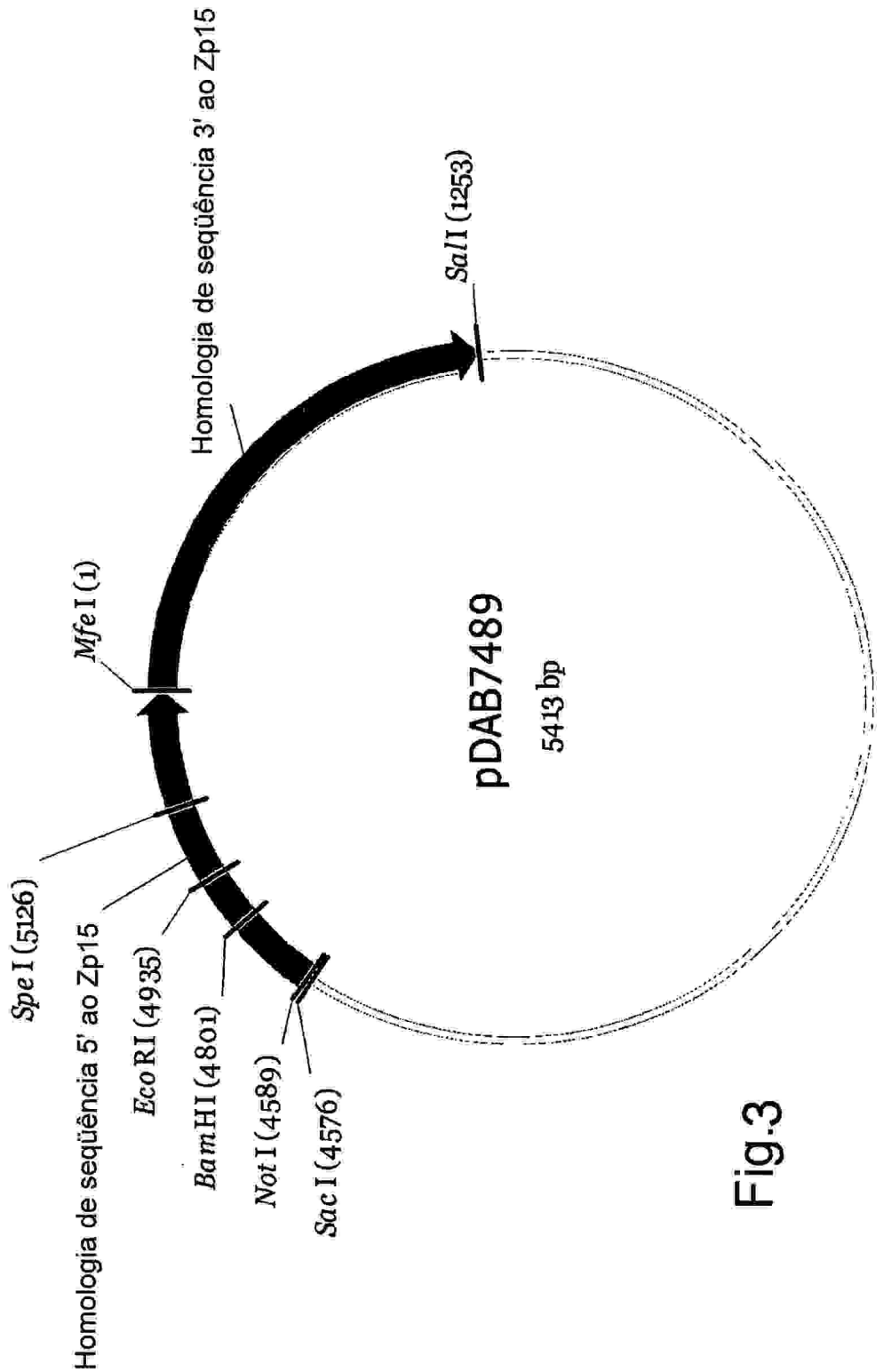

Homologia de seqüência 3' ao Zp15
SalI (1253)
MfeI (1)
pDAB7489
5413 bp
SpeI (5126)
Homologia de seqüência 5' ao Zp15
EcoRI (4935)
BamHI (4801)
NotI (4589)
SacI (4576)

Fig.3

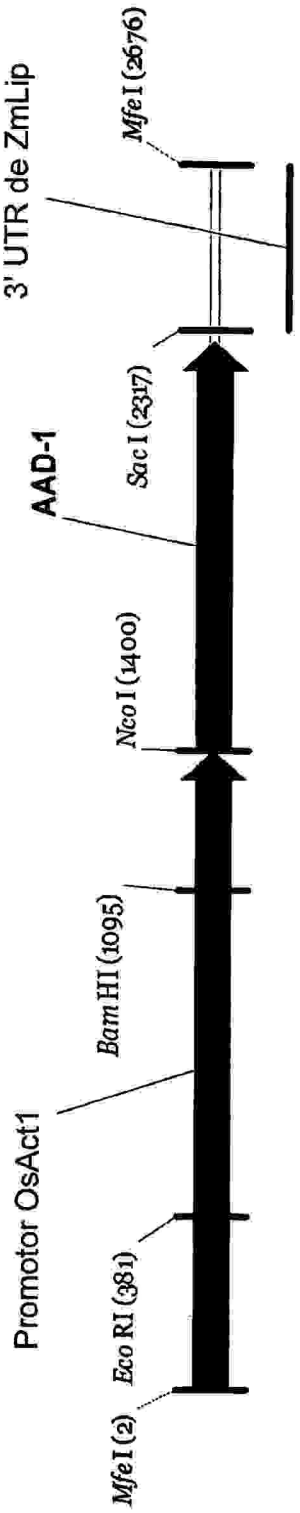

Promotor OsAct1
Mfe I (2)
Eco RI (381)
Bam HI (1095)
Nco I (1400)
AAD-1
Sac I (2317)
3' UTR de ZmLip
Mfe I (2676)
Aad-1v.3 PTU (Mfe I)
2680 bp

Fig. 4

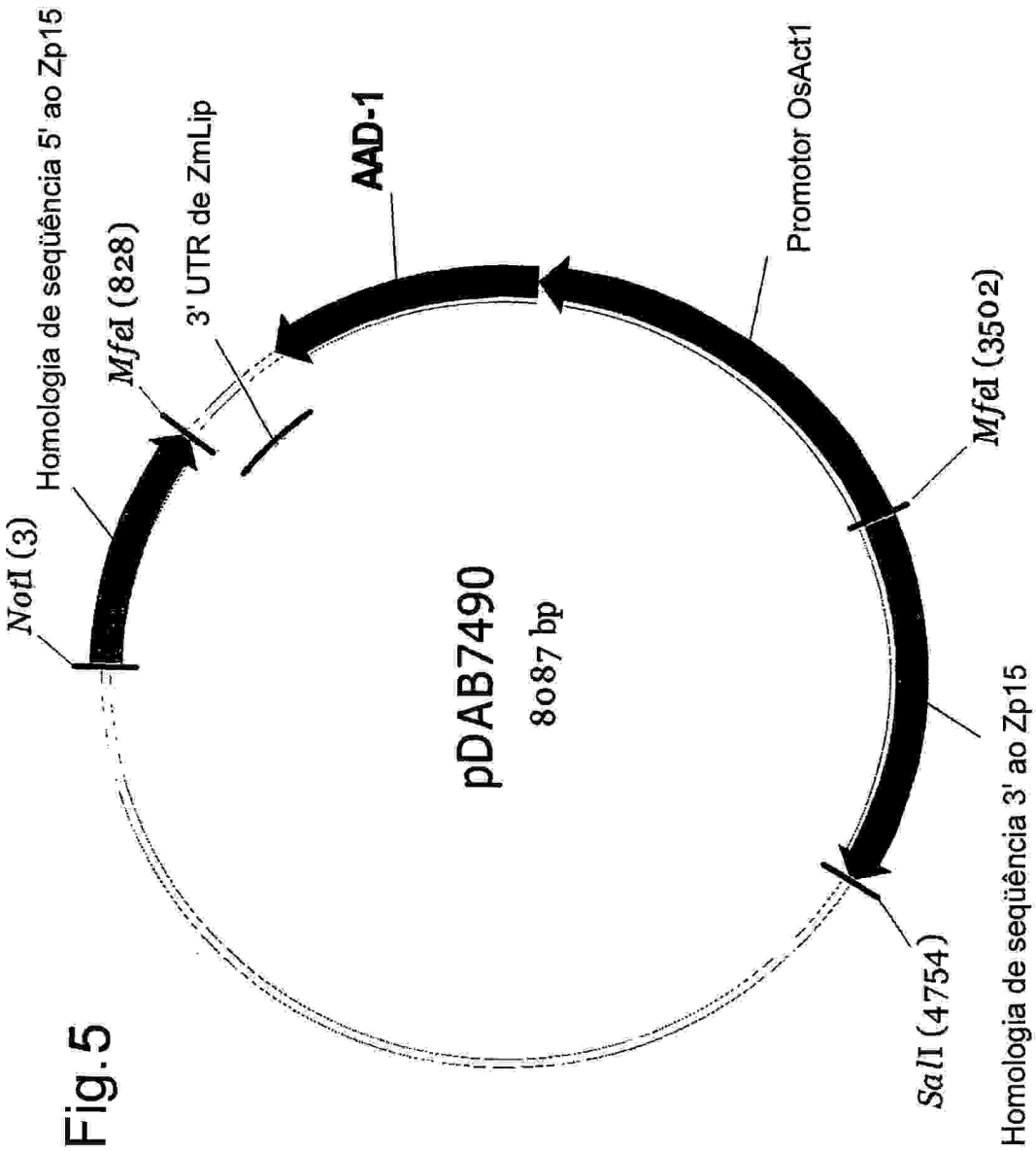
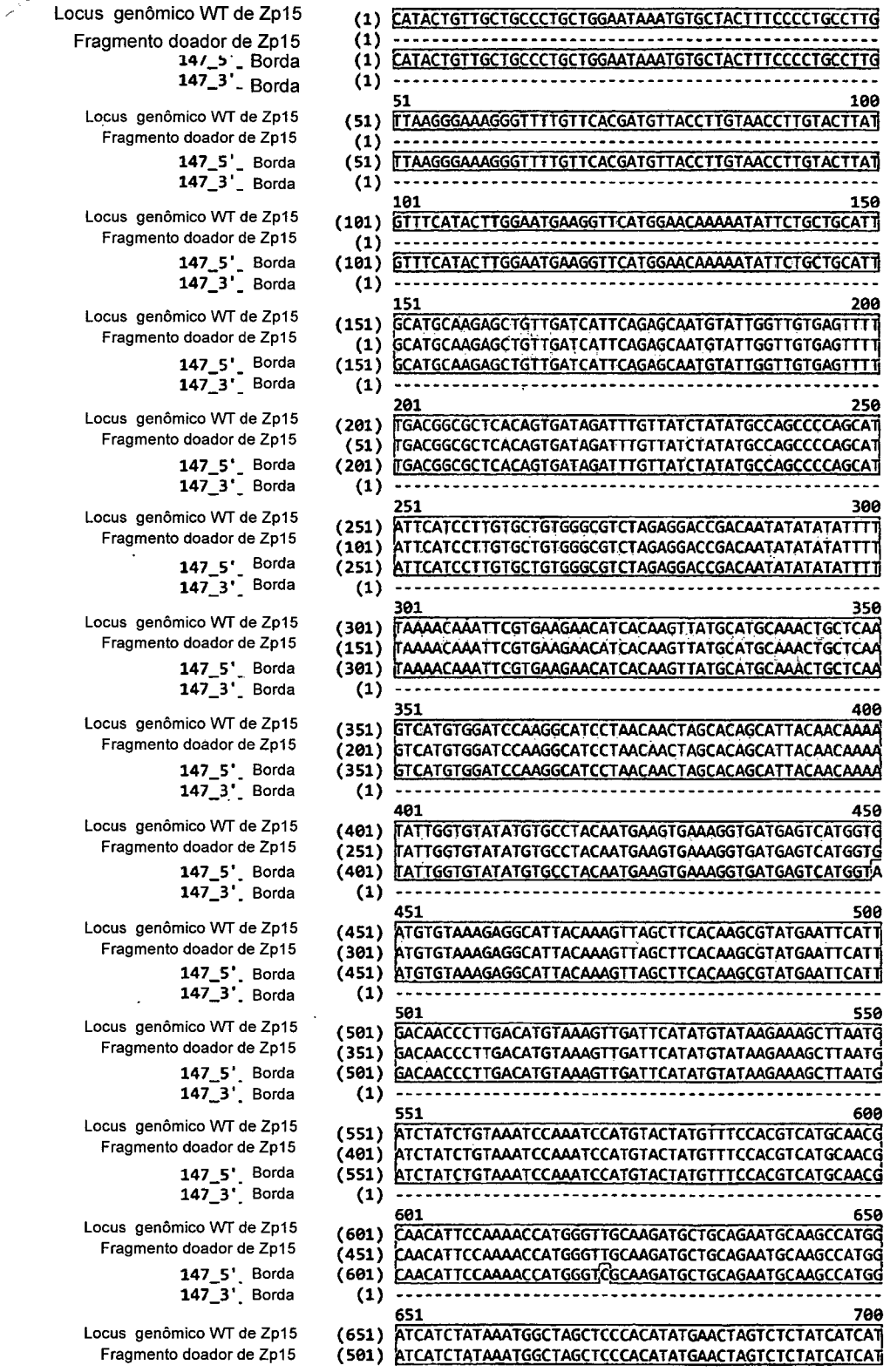
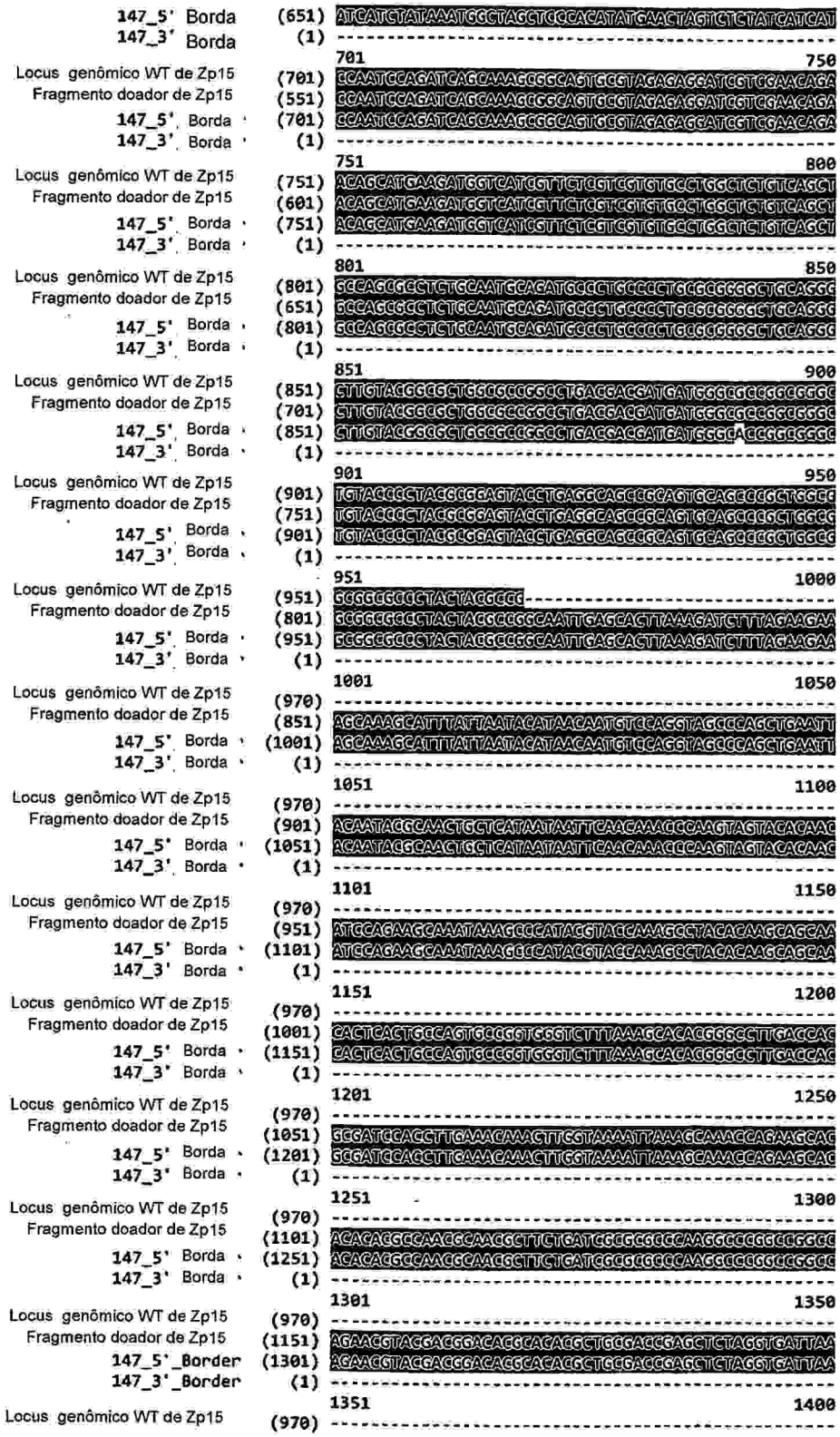
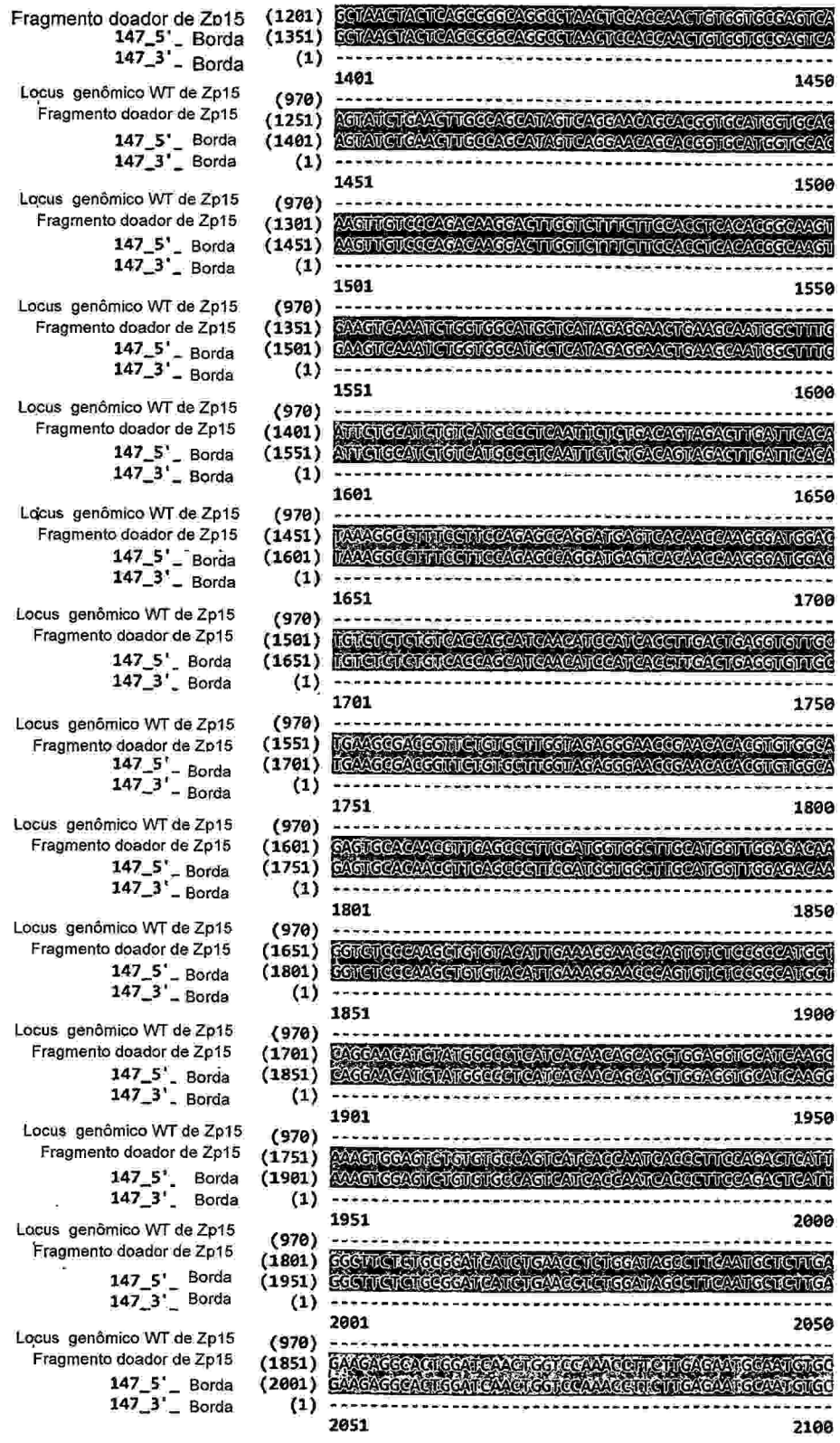

NotI (3)
Homologia de seqüência 5' ao Zp15
MfeI (828)
3' UTR de ZmLip
AAD-1
Promotor OsAct1
MfeI (3502)
Homologia de seqüência 3' ao Zp15
SalI (4754)
pDAB7490
8087 bp

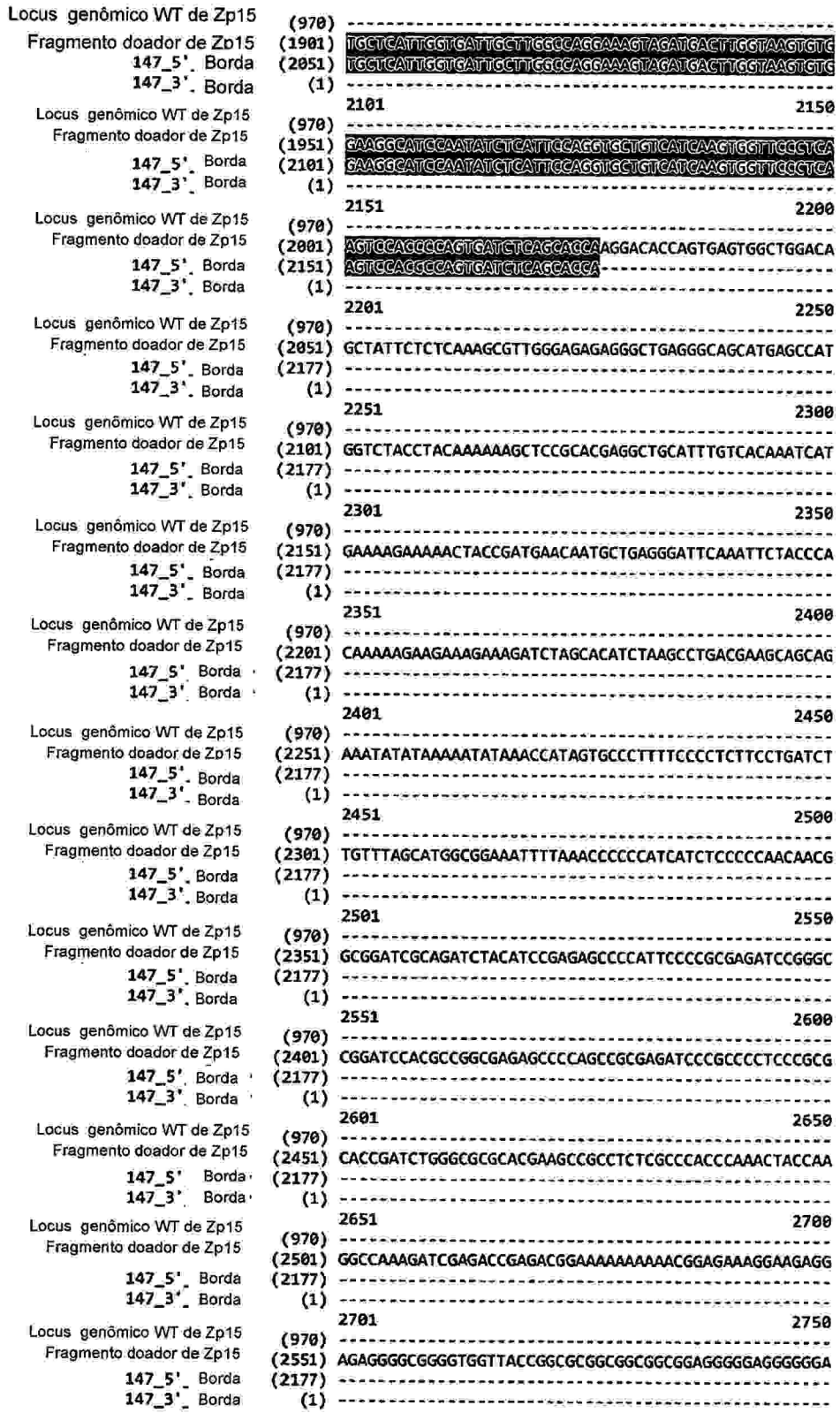
Fig.5

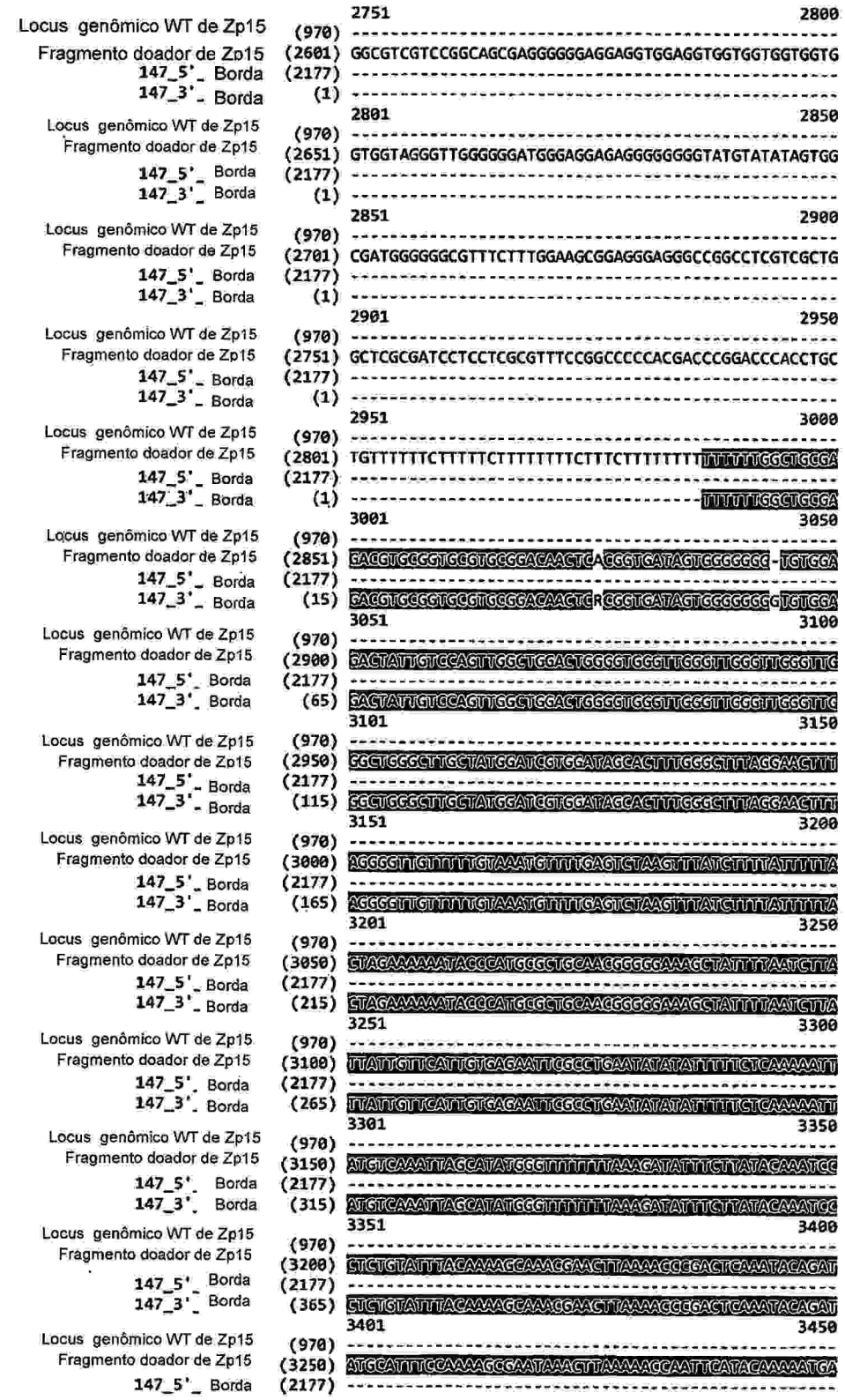
Fig. 6

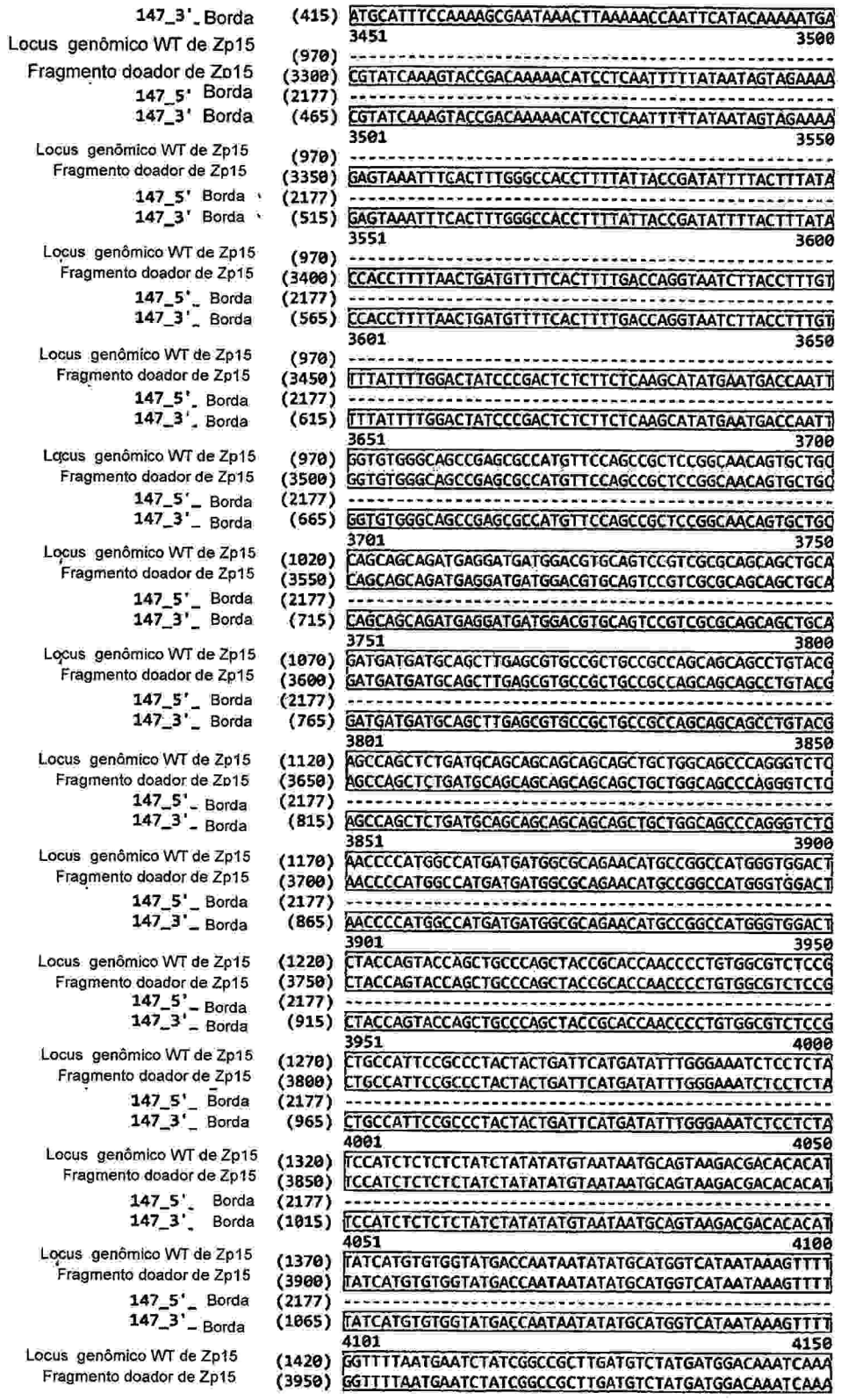

Locus genômico WT de Zp15 (1) CATACTGTTGCTGCCCTGCTGGAATAAATGTGCTACTTTCCCCTGCCTTG
Fragmento doador de Zp15 (1) ----------
147_5'_ Borda (1) CATACTGTTGCTGCCCTGCTGGAATAAATGTGCTACTTTCCCCTGCCTTG
147_3'_ Borda (1) ----------
51 100
Locus genômico WT de Zp15 (51) TTAAGGGAAAGGGTTTTGTTCACGATGTTACCTTGTAACCTTGTACTTAT
Fragmento doador de Zp15 (1) ----------
147_5'_ Borda (51) TTAAGGGAAAGGGTTTTGTTCACGATGTTACCTTGTAACCTTGTACTTAT
147_3'_ Borda (1) ----------
101 150
Locus genômico WT de Zp15 (101) GTTTCATACTTGGAATGAAGGTTCATGGAACAAAAATATTCTGCTGCATT
Fragmento doador de Zp15 (1) ----------
147_5'_ Borda (101) GTTTCATACTTGGAATGAAGGTTCATGGAACAAAAATATTCTGCTGCATT
147_3'_ Borda (1) ----------
151 200
Locus genômico WT de Zp15 (151) GCATGCAAGAGCTGTTGATCATTCAGAGCAATGTATTGGTTGTGAGTTTT
Fragmento doador de Zp15 (1) GCATGCAAGAGCTGTTGATCATTCAGAGCAATGTATTGGTTGTGAGTTTT
147_5'_ Borda (151) GCATGCAAGAGCTGTTGATCATTCAGAGCAATGTATTGGTTGTGAGTTTT
147_3'_ Borda (1) ----------
201 250
Locus genômico WT de Zp15 (201) TGACGGCGCTCACAGTGATAGATTTGTTATCTATATGCCAGCCCCAGCAT
Fragmento doador de Zp15 (51) TGACGGCGCTCACAGTGATAGATTTGTTATCTATATGCCAGCCCCAGCAT
147_5'_ Borda (201) TGACGGCGCTCACAGTGATAGATTTGTTATCTATATGCCAGCCCCAGCAT
147_3'_ Borda (1) ----------
251 300
Locus genômico WT de Zp15 (251) ATTCATCCTTGTGCTGTGGGCGTCTAGAGGACCGACAATATATATATTTT
Fragmento doador de Zp15 (101) ATTCATCCTTGTGCTGTGGGCGTCTAGAGGACCGACAATATATATATTTT
147_5'_ Borda (251) ATTCATCCTTGTGCTGTGGGCGTCTAGAGGACCGACAATATATATATTTT
147_3'_ Borda (1) ----------
301 350
Locus genômico WT de Zp15 (301) TAAAACAAATTCGTGAAGAACATCACAAGTTATGCATGCAAACTGCTCAA
Fragmento doador de Zp15 (151) TAAAACAAATTCGTGAAGAACATCACAAGTTATGCATGCAAACTGCTCAA
147_5'_ Borda (301) TAAAACAAATTCGTGAAGAACATCACAAGTTATGCATGCAAACTGCTCAA
147_3'_ Borda (1) ----------
351 400
Locus genômico WT de Zp15 (351) GTCATGTGGATCCAAGGCATCCTAACAACTAGCACAGCATTACAACAAAA
Fragmento doador de Zp15 (201) GTCATGTGGATCCAAGGCATCCTAACAACTAGCACAGCATTACAACAAAA
147_5'_ Borda (351) GTCATGTGGATCCAAGGCATCCTAACAACTAGCACAGCATTACAACAAAA
147_3'_ Borda (1) ----------
401 450
Locus genômico WT de Zp15 (401) TATTGGTGTATATGTGCCTACAATGAAGTGAAAGGTGATGAGTCATGGTG
Fragmento doador de Zp15 (251) TATTGGTGTATATGTGCCTACAATGAAGTGAAAGGTGATGAGTCATGGTG
147_5'_ Borda (401) TATTGGTGTATATGTGCCTACAATGAAGTGAAAGGTGATGAGTCATGGTA
147_3'_ Borda (1) ----------
451 500
Locus genômico WT de Zp15 (451) ATGTGTAAAGAGGCATTACAAAGTTAGCTTCACAAGCGTATGAATTCATT
Fragmento doador de Zp15 (301) ATGTGTAAAGAGGCATTACAAAGTTAGCTTCACAAGCGTATGAATTCATT
147_5'_ Borda (451) ATGTGTAAAGAGGCATTACAAAGTTAGCTTCACAAGCGTATGAATTCATT
147_3'_ Borda (1) ----------
501 550
Locus genômico WT de Zp15 (501) GACAACCCTTGACATGTAAAGTTGATTCATATGTATAAGAAAGCTTAATG
Fragmento doador de Zp15 (351) GACAACCCTTGACATGTAAAGTTGATTCATATGTATAAGAAAGCTTAATG
147_5'_ Borda (501) GACAACCCTTGACATGTAAAGTTGATTCATATGTATAAGAAAGCTTAATG
147_3'_ Borda (1) ----------
551 600
Locus genômico WT de Zp15 (551) ATCTATCTGTAAATCCAAATCCATGTACTATGTTTCCACGTCATGCAACG
Fragmento doador de Zp15 (401) ATCTATCTGTAAATCCAAATCCATGTACTATGTTTCCACGTCATGCAACG
147_5'_ Borda (551) ATCTATCTGTAAATCCAAATCCATGTACTATGTTTCCACGTCATGCAACG
147_3'_ Borda (1) ----------
601 650
Locus genômico WT de Zp15 (601) CAACATTCCAAAACCATGGGTTGCAAGATGCTGCAGAATGCAAGCCATGG
Fragmento doador de Zp15 (451) CAACATTCCAAAACCATGGGTTGCAAGATGCTGCAGAATGCAAGCCATGG
147_5'_ Borda (601) CAACATTCCAAAACCATGGGTCGCAAGATGCTGCAGAATGCAAGCCATGG
147_3'_ Borda (1) ----------
651 700
Locus genômico WT de Zp15 (651) ATCATCTATAAATGGCTAGCTCCCACATATGAACTAGTCTCTATCATCAT
Fragmento doador de Zp15 (501) ATCATCTATAAATGGCTAGCTCCCACATATGAACTAGTCTCTATCATCAT

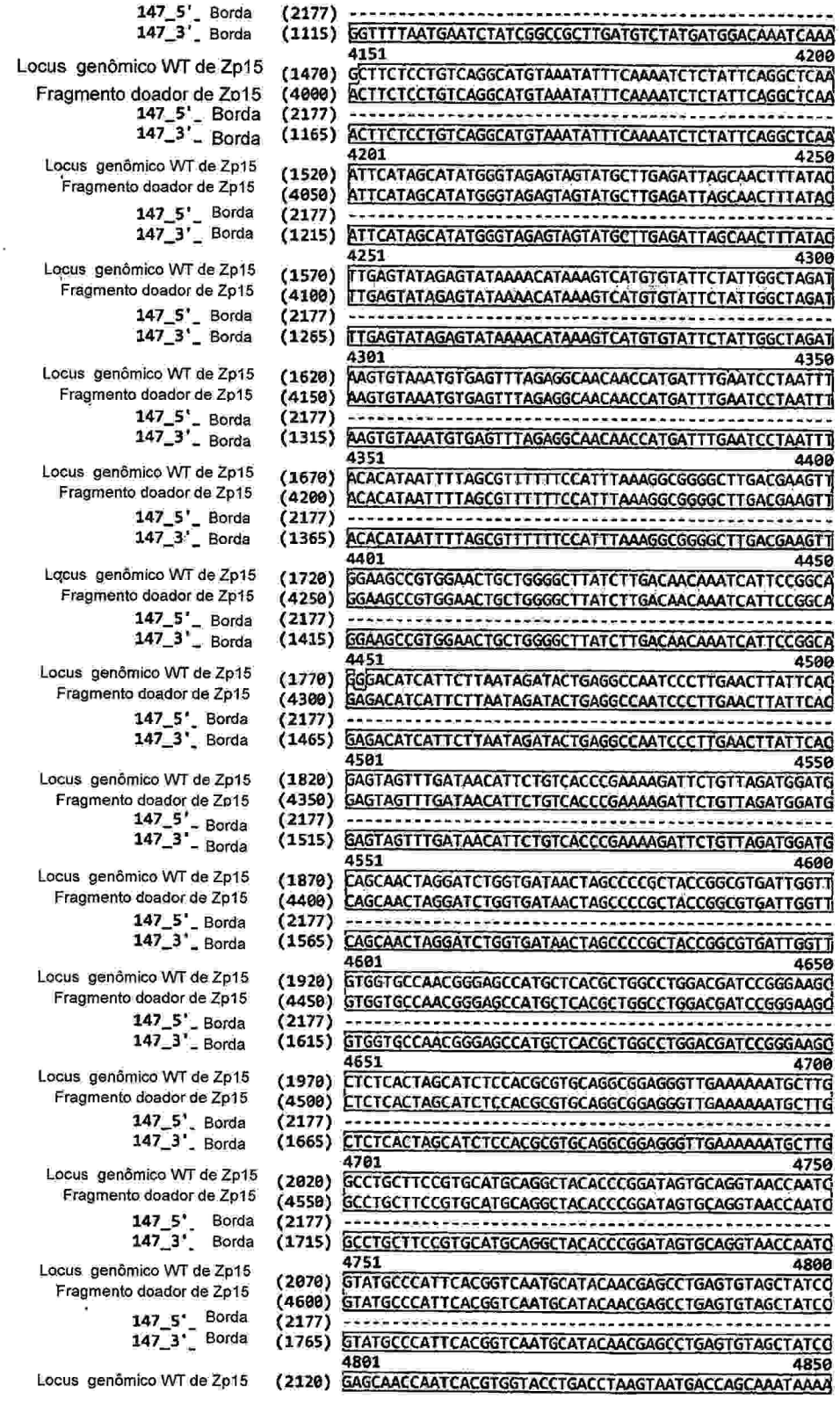

147_5' Borda (651) ATCATCTATAAATGGCTAGCTCCCACATATGAACTAGTCTCTATCATCAT
147_3' Borda (1) ----------
701 750
Locus genômico WT de Zp15 (701) CCAATCCAGATCAGCAAAGCGGCAGTGCGTAGAGAGGATCGTCGAACAGA
Fragmento doador de Zp15 (551) CCAATCCAGATCAGCAAAGCGGCAGTGCGTAGAGAGGATCGTCGAACAGA
147_5' Borda (701) CCAATCCAGATCAGCAAAGCGGCAGTGCGTAGAGAGGATCGTCGAACAGA
147_3' Borda (1) ----------
751 800
Locus genômico WT de Zp15 (751) ACAGCATGAAGATGGTCATCGTTCTCGTCGTGTGCCTGGCTCTGTCAGCT
Fragmento doador de Zp15 (601) ACAGCATGAAGATGGTCATCGTTCTCGTCGTGTGCCTGGCTCTGTCAGCT
147_5' Borda (751) ACAGCATGAAGATGGTCATCGTTCTCGTCGTGTGCCTGGCTCTGTCAGCT
147_3' Borda (1) ----------
801 850
Locus genômico WT de Zp15 (801) GCCAGCGCCTCTGCAATGCAGATGCCCTGCCCCTGCGCGGGGCTGCAGGC
Fragmento doador de Zp15 (651) GCCAGCGCCTCTGCAATGCAGATGCCCTGCCCCTGCGCGGGGCTGCAGGC
147_5' Borda (801) GCCAGCGCCTCTGCAATGCAGATGCCCTGCCCCTGCGCGGGGCTGCAGGC
147_3' Borda (1) ----------
851 900
Locus genômico WT de Zp15 (851) CTTGTACGGCGCTGGCGCCCGGCCTGACGACGATGATGGGGGCCGGCGGGC
Fragmento doador de Zp15 (701) CTTGTACGGCGCTGGCGCCCGGCCTGACGACGATGATGGGGGCCGGCGGGC
147_5' Borda (851) CTTGTACGGCGCTGGCGCCCGGCCTGACGACGATGATGGGGACCGGCGGGC
147_3' Borda (1) ----------
901 950
Locus genômico WT de Zp15 (901) TGTACCCCTACGCGGAGTACCTGAGGCAGCCGCAGTGCAGCCCCGCTGGCG
Fragmento doador de Zp15 (751) TGTACCCCTACGCGGAGTACCTGAGGCAGCCGCAGTGCAGCCCCGCTGGCG
147_5' Borda (901) TGTACCCCTACGCGGAGTACCTGAGGCAGCCGCAGTGCAGCCCCGCTGGCG
147_3' Borda (1) ----------
951 1000
Locus genômico WT de Zp15 (951) GCGGCGCCCTACTACGCCG----------
Fragmento doador de Zp15 (801) GCGGCGCCCTACTACGCCGGCAATTGAGCACTTAAAGATCTTTAGAAGAA
147_5' Borda (951) GCGGCGCCCTACTACGCCGGCAATTGAGCACTTAAAGATCTTTAGAAGAA
147_3' Borda (1) ----------
1001 1050
Locus genômico WT de Zp15 (970) ----------
Fragmento doador de Zp15 (851) AGCAAAGCATTTATTAATACATAACAATGTCCAGGTAGCCCAGCTGAATT
147_5' Borda (1001) AGCAAAGCATTTATTAATACATAACAATGTCCAGGTAGCCCAGCTGAATT
147_3' Borda (1) ----------
1051 1100
Locus genômico WT de Zp15 (970) ----------
Fragmento doador de Zp15 (901) ACAATACGCAACTGCTCATAATAATTCAACAAACCCAAGTAGTACACAAC
147_5' Borda (1051) ACAATACGCAACTGCTCATAATAATTCAACAAACCCAAGTAGTACACAAC
147_3' Borda (1) ----------
1101 1150
Locus genômico WT de Zp15 (970) ----------
Fragmento doador de Zp15 (951) ATCCAGAAGCAAATAAAGCCCATAGGTAGCAAAGCCTACACAAGCAGCAA
147_5' Borda (1101) ATCCAGAAGCAAATAAAGCCCATAGGTAGCAAAGCCTACACAAGCAGCAA
147_3' Borda (1) ----------
1151 1200
Locus genômico WT de Zp15 (970) ----------
Fragmento doador de Zp15 (1001) CACTCACTGCCAGTGCCGGTGGGTCTTTAAAGCACACGGGCCTTGACCAC
147_5' Borda (1151) CACTCACTGCCAGTGCCGGTGGGTCTTTAAAGCACACGGGCCTTGACCAC
147_3' Borda (1) ----------
1201 1250
Locus genômico WT de Zp15 (970) ----------
Fragmento doador de Zp15 (1051) GCGATCCACCTTGAAACAAACTTCGTAAAATTAAAGCAAACCAGAAGCAC
147_5' Borda (1201) GCGATCCACCTTGAAACAAACTTCGTAAAATTAAAGCAAACCAGAAGCAC
147_3' Borda (1) ----------
1251 1300
Locus genômico WT de Zp15 (970) ----------
Fragmento doador de Zp15 (1101) ACACACGCCAACGCAACGCTTCTGATCGCGCGCCCAAGGCCCGGCCGGCC
147_5' Borda (1251) ACACACGCCAACGCAACGCTTCTGATCGCGCGCCCAAGGCCCGGCCGGCC
147_3' Borda (1) ----------
1301 1350
Locus genômico WT de Zp15 (970) ----------
Fragmento doador de Zp15 (1151) AGAACGTACGACGGACACGCACACGCTGCGACCGAGCTCTAGGTGATTAA
147_5'_Border (1301) AGAACGTACGACGGACACGCACACGCTGCGACCGAGCTCTAGGTGATTAA
147_3'_Border (1) ----------
1351 1400
Locus genômico WT de Zp15 (970) ----------

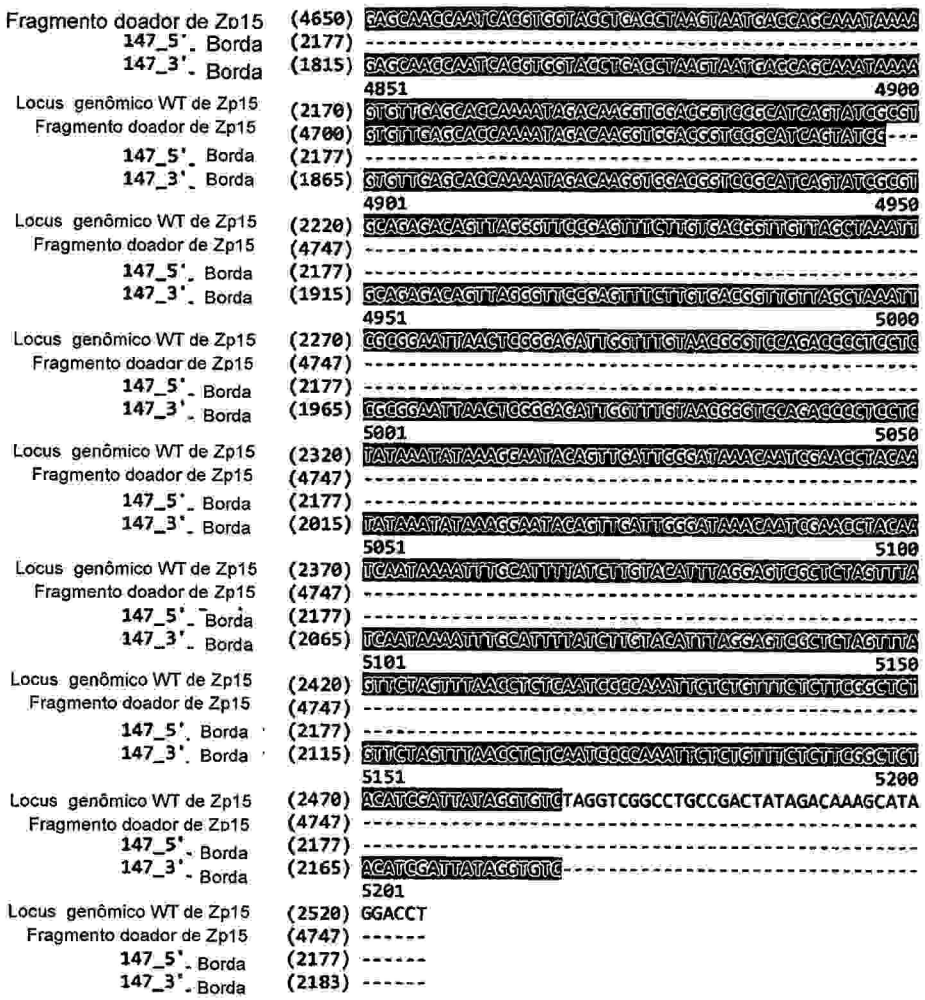

Fragmento doador de Zp15 (1201) GCTAACTACTCAGCGGGCAGGCCTAACTCCAGCAACTGTGGTGCGAGTCA
147_5'_ Borda (1351) GCTAACTACTCAGCGGGCAGGCCTAACTCCACCAACTGTGGTGCGAGTCA
147_3'_ Borda (1) ----------
1401 1450
Locus genômico WT de Zp15 (970) ----------
Fragmento doador de Zp15 (1251) AGTATCTGAACTTGCCAGCATAGTCAGGAACAGCACGGTGCATGGTGCAC
147_5'_ Borda (1401) AGTATCTGAACTTGCCAGCATAGTCAGGAACAGCACGGTGCATGGTGCAC
147_3'_ Borda (1) ----------
1451 1500
Locus genômico WT de Zp15 (970) ----------
Fragmento doador de Zp15 (1301) AAGTTGTCCCAGACAAGGACTTGGTCTTTCTTCCACCTCACACGGCAAGT
147_5'_ Borda (1451) AAGTTGTCCCAGACAAGGACTTGGTCTTTCTTCCAGCTCACACGGCAAGT
147_3'_ Borda (1) ----------
1501 1550
Locus genômico WT de Zp15 (970) ----------
Fragmento doador de Zp15 (1351) GAAGTCAAATCTCGTGGCATGCTCATAGAGGAACTGAAGCAATGGCTTTC
147_5'_ Borda (1501) GAAGTCAAATCTCGTGGCATGCTCATAGAGGAACTGAAGCAATGGCTTTC
147_3'_ Borda (1) ----------
1551 1600
Locus genômico WT de Zp15 (970) ----------
Fragmento doador de Zp15 (1401) ATTCTGCATCTGTCATGCCCTCAATTCTCTGACAGTAGACTTGATTCACA
147_5'_ Borda (1551) ATTCTGCATCTGTCATGCCCTCAATTCTCTGACAGTAGACTTGATTCACA
147_3'_ Borda (1) ----------
1601 1650
Locus genômico WT de Zp15 (970) ----------
Fragmento doador de Zp15 (1451) TAAAGGGCTTTCCTTCCAGAGCCAGGATGAGTCACAACCAAGGCATGGAG
147_5'_ Borda (1601) TAAAGGGCTTTCCTTCCAGAGCCAGGATGAGTCACAACCAAGGCATGGAG
147_3'_ Borda (1) ----------
1651 1700
Locus genômico WT de Zp15 (970) ----------
Fragmento doador de Zp15 (1501) TGTCTCTCTGTCACCAGCATCAACATCCATCACCTTGACTGAGGTGTTGC
147_5'_ Borda (1651) TGTCTCTCTGTCACCAGCATCAACATCCATCACCTTGACTGAGGTGTTGC
147_3'_ Borda (1) ----------
1701 1750
Locus genômico WT de Zp15 (970) ----------
Fragmento doador de Zp15 (1551) TGAAGCGACCGTTCTCTGCTTGCTAGAGGAACCGAACACACGTGTGGCA
147_5'_ Borda (1701) TGAAGCGACCGTTCTCTGCTTGCTAGAGGAACCGAACACACGTGTGGCA
147_3'_ Borda (1) ----------
1751 1800
Locus genômico WT de Zp15 (970) ----------
Fragmento doador de Zp15 (1601) GAGTGCACAACGTTGAGCCCTTCGATGGTGGCTTGCATGGTTGGAGACAA
147_5'_ Borda (1751) GAGTGCACAACGTTGAGCCCTTCGATGGTGGCTTGCATGGTTGGAGACAA
147_3'_ Borda (1) ----------
1801 1850
Locus genômico WT de Zp15 (970) ----------
Fragmento doador de Zp15 (1651) GGTCTGCCAAGCTGTGTACATTGAAAGGAACCCAGTGTCTCCGGCATGCT
147_5'_ Borda (1801) GGTCTCCCAAGCTGTGTACATTGAAAGGAACCCAGTGTCTCCGGCATGCT
147_3'_ Borda (1) ----------
1851 1900
Locus genômico WT de Zp15 (970) ----------
Fragmento doador de Zp15 (1701) GAGGAAGATGTATGGCCCTCATCACAACAGCAGCTGGAGGTGCATCAAGG
147_5'_ Borda (1851) GAGGAACATCTATGGCCCTCATCACAACAGCAGCTGGAGGTGCATCAAGG
147_3'_ Borda (1) ----------
1901 1950
Locus genômico WT de Zp15 (970) ----------
Fragmento doador de Zp15 (1751) AAAGTGGAGTCTGTGTGCCAGTCATCACGCAATCACCCTTCCAGACTCATT
147_5'_ Borda (1901) AAAGTGGAGTCTGTGTGCCAGTCATCACGCAATCACCCTTCCAGACTCATT
147_3'_ Borda (1) ----------
1951 2000
Locus genômico WT de Zp15 (970) ----------
Fragmento doador de Zp15 (1801) GGCTTCTCTCTGCGGCATCATCTGAACCTCTGGATAGCCTTCAATGCTCTTGA
147_5'_ Borda (1951) GGCTTCTCTCTGCGGCATCATCTGAACCTCTGGATAGCCTTCAATGCTCTTGA
147_3'_ Borda (1) ----------
2001 2050
Locus genômico WT de Zp15 (970) ----------
Fragmento doador de Zp15 (1851) GAAGAGGCACTGGATCAACTGGTCCAAACCTTCTTGAGAATGCAATGTGC
147_5'_ Borda (2001) GAAGAGGCACTGGATCAACTGGTCCAAACCTTCTTGAGAATGCAATGTGC
147_3'_ Borda (1) ----------
2051 2100


Locus genômico WT de Zp15
Fragmento doador de Zp15
147_5'_Borda
147_3'_Borda
(970) ----------
(1901) TGCTCATTGGTGATTGCTTGGCCAGGAAAGTAGATGACTTGGTAAGTGTG
(2051) TGCTCATTGGTGATTGCTTGGCCAGGAAAGTAGATGACTTGGTAAGTGTG
(1) ----------
2101 2150
(970) ----------
(1951) GAAGGCATCCAATATCTCATTCCAGGTGCTGTCATCAAGTGGTTCCCTCA
(2101) GAAGGCATCCAATATCTCATTCCAGGTGCTGTCATCAAGTGGTTCCCTCA
(1) ----------
2151 2200
(970) ----------
(2001) AGTCCACCCCAGTGATCTCAGCACCAAGGACACCAGTGAGTGGCTGGACA
(2151) AGTCCACCCCAGTGATCTCAGCACCA----------
(1) ----------
2201 2250
(970) ----------
(2051) GCTATTCTCTCAAAGCGTTGGGAGAGAGGGCTGAGGGCAGCATGAGCCAT
(2177) ----------
(1) ----------
2251 2300
(970) ----------
(2101) GGTCTACCTACAAAAAAGCTCCGCACGAGGCTGCATTTGTCACAAATCAT
(2177) ----------
(1) ----------
2301 2350
(970) ----------
(2151) GAAAAGAAAAACTACCGATGAACAATGCTGAGGGATTCAAATTCTACCCA
(2177) ----------
(1) ----------
2351 2400
(970) ----------
(2201) CAAAAAGAAGAAAGAAAGATCTAGCACATCTAAGCCTGACGAAGCAGCAG
(2177) ----------
(1) ----------
2401 2450
(970) ----------
(2251) AAATATATAAAAATATAAACCATAGTGCCCTTTTCCCCTCTTCCTGATCT
(2177) ----------
(1) ----------
2451 2500
(970) ----------
(2301) TGTTTAGCATGGCGGAAATTTTAAACCCCCCATCATCTCCCCCAACAACG
(2177) ----------
(1) ----------
2501 2550
(970) ----------
(2351) GCGGATCGCAGATCTACATCCGAGAGCCCCATTCCCCGCGAGATCCGGGC
(2177) ----------
(1) ----------
2551 2600
(970) ----------
(2401) CGGATCCACGCCGGCGAGAGCCCCAGCCGCGAGATCCCGCCCCTCCCGCG
(2177) ----------
(1) ----------
2601 2650
(970) ----------
(2451) CACCGATCTGGGCGCGCACGAAGCCGCCTCTCGCCCACCCAAACTACCAA
(2177) ----------
(1) ----------
2651 2700
(970) ----------
(2501) GGCCAAAGATCGAGACCGAGACGGAAAAAAAAAAACGGAGAAAGGAAGAGG
(2177) ----------
(1) ----------
2701 2750
(970) ----------
(2551) AGAGGGGCGGGGTGGTTACCGGCGCGGCGGCGGCGGAGGGGGAGGGGGGA
(2177) ----------
(1) ----------


Locus genômico WT de Zp15
Fragmento doador de Zp15
147_5'_ Borda
147_3'_ Borda
2751 2800
(970) ----------
(2601) GGCGTCGTCCGGCAGCGAGGGGGGAGGAGGTGGAGGTGGTGGTGGTGGTG
(2177) ----------
(1) ----------
2801 2850
(970) ----------
(2651) GTGGTAGGGTTGGGGGGATGGGAGGAGAGGGGGGGGTATGTATATAGTGG
(2177) ----------
(1) ----------
2851 2900
(970) ----------
(2701) CGATGGGGGGCGTTTCTTTGGAAGCGGAGGGAGGGCCGGCCTCGTCGCTG
(2177) ----------
(1) ----------
2901 2950
(970) ----------
(2751) GCTCGCGATCCTCCTCGCGTTTCCGGCCCCCACGACCCGGACCCACCTGC
(2177) ----------
(1) ----------
2951 3000
(970) ----------
(2801) TGTTTTTTCTTTTTCTTTTTTTTCTTTCTTTTTTTTTTTTTTTGGCTGCGA
(2177) ----------
(1) ----------TTTTTTGGCTGCGA
3001 3050
(970) ----------
(2851) GACGTGCGGTGCGTGCGGACAACTCACGGTGATAGTGGGGGGG-TGTGGA
(2177) ----------
(15) GACGTGCGGTGCGTGCGGACAACTCRCGGTGATAGTGGGGGGGGTGTGGA
3051 3100
(970) ----------
(2900) GACTATTGTCCAGTTGGCTGGACTGGGGTGGGTTGGGTTGGGTTGGGTTG
(2177) ----------
(65) GACTATTGTCCAGTTGGCTGGACTGGGGTGGGTTGGGTTGGGTTGGGTTG
3101 3150
(970) ----------
(2950) GGCTGGGCTTGCTATGGATCGTGGATAGCACTTTGGGCTTTAGGAACTTT
(2177) ----------
(115) GGCTGGGCTTGCTATGGATCGTGGATAGCACTTTGGGCTTTAGGAACTTT
3151 3200
(970) ----------
(3000) AGGGGTTGTTTTTGTAAATGTTTTGAGTCTAAGTTTATCTTTTATTTTTA
(2177) ----------
(165) AGGGGTTGTTTTTGTAAATGTTTTGAGTCTAAGTTTATCTTTTATTTTTA
3201 3250
(970) ----------
(3050) GTAGAAAAAATACGCATGGGCTGCAACGGGGGAAAGCTATTTTAATCTTA
(2177) ----------
(215) GTAGAAAAAATACGCATGGGCTGCAACGGGGGAAAGCTATTTTAATCTTA
3251 3300
(970) ----------
(3100) TTATTGTTCATTGTGAGAATTCGGCTGAATATATATTTTTCTCAAAAATT
(2177) ----------
(265) TTATTGTTCATTGTGAGAATTCGGCTGAATATATATTTTTCTCAAAAATT
3301 3350
(970) ----------
(3150) ATGTCAAATTAGCATATGGGTTTTTTTAAAGATATTTCTTATACAAATCC
(2177) ----------
(315) ATGTCAAATTAGCATATGGGTTTTTTTAAAGATATTTCTTATACAAATCC
3351 3400
(970) ----------
(3200) CTCTGTATTTACAAAAGCAAACGAACTTAAAACGCGACTCAAATACAGAT
(2177) ----------
(365) CTCTGTATTTACAAAAGCAAACGAACTTAAAACGCGACTCAAATACAGAT
3401 3450
(970) ----------
(3250) ATGCATTTGCAAAAGCGAATAAACTTAAAAACCAATTCATACAAAAATGA
(2177) ----------


147_3'_Borda (415) ATGCATTTCCAAAAGCGAATAAACTTAAAAACCAATTCATACAAAAATGA
3451 3500
Locus genômico WT de Zp15 (970) ----------
Fragmento doador de Zp15 (3300) CGTATCAAAGTACCGACAAAAACATCCTCAATTTTTATAATAGTAGAAAA
147_5' Borda (2177) ----------
147_3' Borda (465) CGTATCAAAGTACCGACAAAAACATCCTCAATTTTTATAATAGTAGAAAA
3501 3550
Locus genômico WT de Zp15 (970) ----------
Fragmento doador de Zp15 (3350) GAGTAAATTTCACTTTGGGCCACCTTTTATTACCGATATTTTACTTTATA
147_5' Borda (2177) ----------
147_3' Borda (515) GAGTAAATTTCACTTTGGGCCACCTTTTATTACCGATATTTTACTTTATA
3551 3600
Locus genômico WT de Zp15 (970) ----------
Fragmento doador de Zp15 (3400) CCACCTTTTAACTGATGTTTTCACTTTTGACCAGGTAATCTTACCTTTGT
147_5'_Borda (2177) ----------
147_3'_Borda (565) CCACCTTTTAACTGATGTTTTCACTTTTGACCAGGTAATCTTACCTTTGT
3601 3650
Locus genômico WT de Zp15 (970) ----------
Fragmento doador de Zp15 (3450) TTTATTTTGGACTATCCCGACTCTCTTCTCAAGCATATGAATGACCAATT
147_5'_Borda (2177) ----------
147_3'_Borda (615) TTTATTTTGGACTATCCCGACTCTCTTCTCAAGCATATGAATGACCAATT
3651 3700
Locus genômico WT de Zp15 (970) GGTGTGGGCAGCCGAGCGCCATGTTCCAGCCGCTCCGGCAACAGTGCTGC
Fragmento doador de Zp15 (3500) GGTGTGGGCAGCCGAGCGCCATGTTCCAGCCGCTCCGGCAACAGTGCTGC
147_5'_Borda (2177) ----------
147_3'_Borda (665) GGTGTGGGCAGCCGAGCGCCATGTTCCAGCCGCTCCGGCAACAGTGCTGC
3701 3750
Locus genômico WT de Zp15 (1020) CAGCAGCAGATGAGGATGATGGACGTGCAGTCCGTCGCGCAGCAGCTGCA
Fragmento doador de Zp15 (3550) CAGCAGCAGATGAGGATGATGGACGTGCAGTCCGTCGCGCAGCAGCTGCA
147_5'_Borda (2177) ----------
147_3'_Borda (715) CAGCAGCAGATGAGGATGATGGACGTGCAGTCCGTCGCGCAGCAGCTGCA
3751 3800
Locus genômico WT de Zp15 (1070) GATGATGATGCAGCTTGAGCGTGCCGCTGCCGCCAGCAGCAGCCTGTACG
Fragmento doador de Zp15 (3600) GATGATGATGCAGCTTGAGCGTGCCGCTGCCGCCAGCAGCAGCCTGTACG
147_5'_Borda (2177) ----------
147_3'_Borda (765) GATGATGATGCAGCTTGAGCGTGCCGCTGCCGCCAGCAGCAGCCTGTACG
3801 3850
Locus genômico WT de Zp15 (1120) AGCCAGCTCTGATGCAGCAGCAGCAGCAGCTGCTGGCAGCCCAGGGTCTC
Fragmento doador de Zp15 (3650) AGCCAGCTCTGATGCAGCAGCAGCAGCAGCTGCTGGCAGCCCAGGGTCTC
147_5'_Borda (2177) ----------
147_3'_Borda (815) AGCCAGCTCTGATGCAGCAGCAGCAGCAGCTGCTGGCAGCCCAGGGTCTC
3851 3900
Locus genômico WT de Zp15 (1170) AACCCCATGGCCATGATGATGGCGCAGAACATGCCGGCCATGGGTGGACT
Fragmento doador de Zp15 (3700) AACCCCATGGCCATGATGATGGCGCAGAACATGCCGGCCATGGGTGGACT
147_5'_Borda (2177) ----------
147_3'_Borda (865) AACCCCATGGCCATGATGATGGCGCAGAACATGCCGGCCATGGGTGGACT
3901 3950
Locus genômico WT de Zp15 (1220) CTACCAGTACCAGCTGCCCAGCTACCGCACCAACCCCTGTGGCGTCTCCG
Fragmento doador de Zp15 (3750) CTACCAGTACCAGCTGCCCAGCTACCGCACCAACCCCTGTGGCGTCTCCG
147_5'_Borda (2177) ----------
147_3'_Borda (915) CTACCAGTACCAGCTGCCCAGCTACCGCACCAACCCCTGTGGCGTCTCCG
3951 4000
Locus genômico WT de Zp15 (1270) CTGCCATTCCGCCCTACTACTGATTCATGATATTTGGGAAATCTCCTCTA
Fragmento doador de Zp15 (3800) CTGCCATTCCGCCCTACTACTGATTCATGATATTTGGGAAATCTCCTCTA
147_5'_Borda (2177) ----------
147_3'_Borda (965) CTGCCATTCCGCCCTACTACTGATTCATGATATTTGGGAAATCTCCTCTA
4001 4050
Locus genômico WT de Zp15 (1320) TCCATCTCTCTCTATCTATATATGTAATAATGCAGTAAGACGACACACAT
Fragmento doador de Zp15 (3850) TCCATCTCTCTCTATCTATATATGTAATAATGCAGTAAGACGACACACAT
147_5'_Borda (2177) ----------
147_3'_Borda (1015) TCCATCTCTCTCTATCTATATATGTAATAATGCAGTAAGACGACACACAT
4051 4100
Locus genômico WT de Zp15 (1370) TATCATGTGTGGTATGACCAATAATATATGCATGGTCATAATAAAGTTTT
Fragmento doador de Zp15 (3900) TATCATGTGTGGTATGACCAATAATATATGCATGGTCATAATAAAGTTTT
147_5'_Borda (2177) ----------
147_3'_Borda (1065) TATCATGTGTGGTATGACCAATAATATATGCATGGTCATAATAAAGTTTT
4101 4150
Locus genômico WT de Zp15 (1420) GGTTTTAATGAATCTATCGGCCGCTTGATGTCTATGATGGACAAATCAAA
Fragmento doador de Zp15 (3950) GGTTTTAATGAATCTATCGGCCGCTTGATGTCTATGATGGACAAATCAAA


147_5'_ Borda (2177) ----------
147_3'_ Borda (1115) GGTTTTAATGAATCTATCGGCCGCTTGATGTCTATGATGGACAAATCAAA
4151 4200
Locus genômico WT de Zp15 (1470) GCTTCTCCTGTCAGGCATGTAAATATTTCAAAATCTCTATTCAGGCTCAA
Fragmento doador de Zp15 (4000) ACTTCTCCTGTCAGGCATGTAAATATTTCAAAATCTCTATTCAGGCTCAA
147_5'_ Borda (2177) ----------
147_3'_ Borda (1165) ACTTCTCCTGTCAGGCATGTAAATATTTCAAAATCTCTATTCAGGCTCAA
4201 4250
Locus genômico WT de Zp15 (1520) ATTCATAGCATATGGGTAGAGTAGTATGCTTGAGATTAGCAACTTTATAC
Fragmento doador de Zp15 (4050) ATTCATAGCATATGGGTAGAGTAGTATGCTTGAGATTAGCAACTTTATAC
147_5'_ Borda (2177) ----------
147_3'_ Borda (1215) ATTCATAGCATATGGGTAGAGTAGTATGCTTGAGATTAGCAACTTTATAC
4251 4300
Locus genômico WT de Zp15 (1570) TTGAGTATAGAGTATAAAACATAAAGTCATGTGTATTCTATTGGCTAGAT
Fragmento doador de Zp15 (4100) TTGAGTATAGAGTATAAAACATAAAGTCATGTGTATTCTATTGGCTAGAT
147_5'_ Borda (2177) ----------
147_3'_ Borda (1265) TTGAGTATAGAGTATAAAACATAAAGTCATGTGTATTCTATTGGCTAGAT
4301 4350
Locus genômico WT de Zp15 (1620) AAGTGTAAATGTGAGTTTAGAGGCAACAACCATGATTTGAATCCTAATTT
Fragmento doador de Zp15 (4150) AAGTGTAAATGTGAGTTTAGAGGCAACAACCATGATTTGAATCCTAATTT
147_5'_ Borda (2177) ----------
147_3'_ Borda (1315) AAGTGTAAATGTGAGTTTAGAGGCAACAACCATGATTTGAATCCTAATTT
4351 4400
Locus genômico WT de Zp15 (1670) ACACATAATTTTAGCGTTTTTTCCATTTAAAGGCGGGGCTTGACGAAGTT
Fragmento doador de Zp15 (4200) ACACATAATTTTAGCGTTTTTTCCATTTAAAGGCGGGGCTTGACGAAGTT
147_5'_ Borda (2177) ----------
147_3'_ Borda (1365) ACACATAATTTTAGCGTTTTTTCCATTTAAAGGCGGGGCTTGACGAAGTT
4401 4450
Locus genômico WT de Zp15 (1720) GGAAGCCGTGGAACTGCTGGGGCTTATCTTGACAACAAATCATTCCGGCA
Fragmento doador de Zp15 (4250) GGAAGCCGTGGAACTGCTGGGGCTTATCTTGACAACAAATCATTCCGGCA
147_5'_ Borda (2177) ----------
147_3'_ Borda (1415) GGAAGCCGTGGAACTGCTGGGGCTTATCTTGACAACAAATCATTCCGGCA
4451 4500
Locus genômico WT de Zp15 (1770) GGGACATCATTCTTAATAGATACTGAGGCCAATCCCTTGAACTTATTCAC
Fragmento doador de Zp15 (4300) GAGACATCATTCTTAATAGATACTGAGGCCAATCCCTTGAACTTATTCAC
147_5'_ Borda (2177) ----------
147_3'_ Borda (1465) GAGACATCATTCTTAATAGATACTGAGGCCAATCCCTTGAACTTATTCAC
4501 4550
Locus genômico WT de Zp15 (1820) GAGTAGTTTGATAACATTCTGTCACCCGAAAAGATTCTGTTAGATGGATG
Fragmento doador de Zp15 (4350) GAGTAGTTTGATAACATTCTGTCACCCGAAAAGATTCTGTTAGATGGATG
147_5'_ Borda (2177) ----------
147_3'_ Borda (1515) GAGTAGTTTGATAACATTCTGTCACCCGAAAAGATTCTGTTAGATGGATG
4551 4600
Locus genômico WT de Zp15 (1870) CAGCAACTAGGATCTGGTGATAACTAGCCCCGCTACCGGCGTGATTGGTT
Fragmento doador de Zp15 (4400) CAGCAACTAGGATCTGGTGATAACTAGCCCCGCTACCGGCGTGATTGGTT
147_5'_ Borda (2177) ----------
147_3'_ Borda (1565) CAGCAACTAGGATCTGGTGATAACTAGCCCCGCTACCGGCGTGATTGGTT
4601 4650
Locus genômico WT de Zp15 (1920) GTGGTGCCAACGGGAGCCATGCTCACGCTGGCCTGGACGATCCGGGAAGC
Fragmento doador de Zp15 (4450) GTGGTGCCAACGGGAGCCATGCTCACGCTGGCCTGGACGATCCGGGAAGC
147_5'_ Borda (2177) ----------
147_3'_ Borda (1615) GTGGTGCCAACGGGAGCCATGCTCACGCTGGCCTGGACGATCCGGGAAGC
4651 4700
Locus genômico WT de Zp15 (1970) CTCTCACTAGCATCTCCACGCGTGCAGGCGGAGGGTTGAAAAAATGCTTG
Fragmento doador de Zp15 (4500) CTCTCACTAGCATCTCCACGCGTGCAGGCGGAGGGTTGAAAAAATGCTTG
147_5'_ Borda (2177) ----------
147_3'_ Borda (1665) CTCTCACTAGCATCTCCACGCGTGCAGGCGGAGGGTTGAAAAAATGCTTG
4701 4750
Locus genômico WT de Zp15 (2020) GCCTGCTTCCGTGCATGCAGGCTACACCCGGATAGTGCAGGTAACCAATG
Fragmento doador de Zp15 (4550) GCCTGCTTCCGTGCATGCAGGCTACACCCGGATAGTGCAGGTAACCAATG
147_5'_ Borda (2177) ----------
147_3'_ Borda (1715) GCCTGCTTCCGTGCATGCAGGCTACACCCGGATAGTGCAGGTAACCAATG
4751 4800
Locus genômico WT de Zp15 (2070) GTATGCCCATTCACGGTCAATGCATACAACGAGCCTGAGTGTAGCTATCG
Fragmento doador de Zp15 (4600) GTATGCCCATTCACGGTCAATGCATACAACGAGCCTGAGTGTAGCTATCG
147_5'_ Borda (2177) ----------
147_3'_ Borda (1765) GTATGCCCATTCACGGTCAATGCATACAACGAGCCTGAGTGTAGCTATCG
4801 4850
Locus genômico WT de Zp15 (2120) GAGCAACCAATCACGTGGTACCTGACCTAAGTAATGACCAGCAAATAAAA


Fragmento doador de Zp15 (4650) GAGCAAGCAATCAGGTGGTACGTGACGTAAGTAATGAGCAGCAAATAAAA
147_5'_ Borda (2177) ----------
147_3'_ Borda (1815) GAGCAAGCAATCAGGTGGTACCTGACGTAAGTAATGAGCAGCAAATAAAA
4851 4900
Locus genômico WT de Zp15 (2170) GTGTTGAGCACCAAAATAGACAAGGTGGACGGTCCGCATCAGTATCGCGT
Fragmento doador de Zp15 (4700) GTGTTGAGCACCAAAATAGACAAGGTGGACGGTCGGCATCAGTATCG---
147_5'_ Borda (2177) ----------
147_3'_ Borda (1865) GTGTTGAGCACCAAAATAGACAAGGTGGACGGTCCGCATCAGTATCGCGT
4901 4950
Locus genômico WT de Zp15 (2220) GCAGAGACAGTTAGGGTTCCGAGTTTCTTGTGACGGTTGTTAGCTAAATT
Fragmento doador de Zp15 (4747) ----------
147_5'_ Borda (2177) ----------
147_3'_ Borda (1915) GCAGAGACAGTTAGGGTTCCGAGTTTCTTGTGACGGTTGTTAGCTAAATT
4951 5000
Locus genômico WT de Zp15 (2270) GGCGGAATTAACTCGGGAGATTGGTTTGTAACGGGTCCAGACCCCTCCTC
Fragmento doador de Zp15 (4747) ----------
147_5'_ Borda (2177) ----------
147_3'_ Borda (1965) GGCGGAATTAACTCGGGAGATTGGTTTGTAACGGGTCCAGACCCCTCCTC
5001 5050
Locus genômico WT de Zp15 (2320) TATAAATATAAAGGAATACAGTTGATTGCGATAAACAATCGAACCTACAA
Fragmento doador de Zp15 (4747) ----------
147_5'_ Borda (2177) ----------
147_3'_ Borda (2015) TATAAATATAAAGGAATACAGTTGATTGCGATAAACAATCGAACCTACAA
5051 5100
Locus genômico WT de Zp15 (2370) TCAATAAAATTTGCATTTTATCTTGTACATTTAGGAGTCGCTCTAGTTTA
Fragmento doador de Zp15 (4747) ----------
147_5'_ Borda (2177) ----------
147_3'_ Borda (2065) TCAATAAAATTTGCATTTTATCTTGTACATTTAGGAGTCGCTCTAGTTTA
5101 5150
Locus genômico WT de Zp15 (2420) GTTCTAGTTTAACCTGTCAATCGCCCAAATTGTCTGTTTCTCTTCGGCTCT
Fragmento doador de Zp15 (4747) ----------
147_5'_ Borda (2177) ----------
147_3'_ Borda (2115) GTTCTAGTTTAACCTGTCAATCGCCCAAATTGTCTGTTTCTCTTCGGCTCT
5151 5200
Locus genômico WT de Zp15 (2470) ACATCGATTATAGGTGTCTAGGTCGGCCTGCCGACTATAGACAAAGCATA
Fragmento doador de Zp15 (4747) ----------
147_5'_ Borda (2177) ----------
147_3'_ Borda (2165) ACATCGATTATAGGTGTC----------
5201
Locus genômico WT de Zp15 (2520) GGACCT
Fragmento doador de Zp15 (4747) ------
147_5'_ Borda (2177) ------
147_3'_ Borda (2183) ------

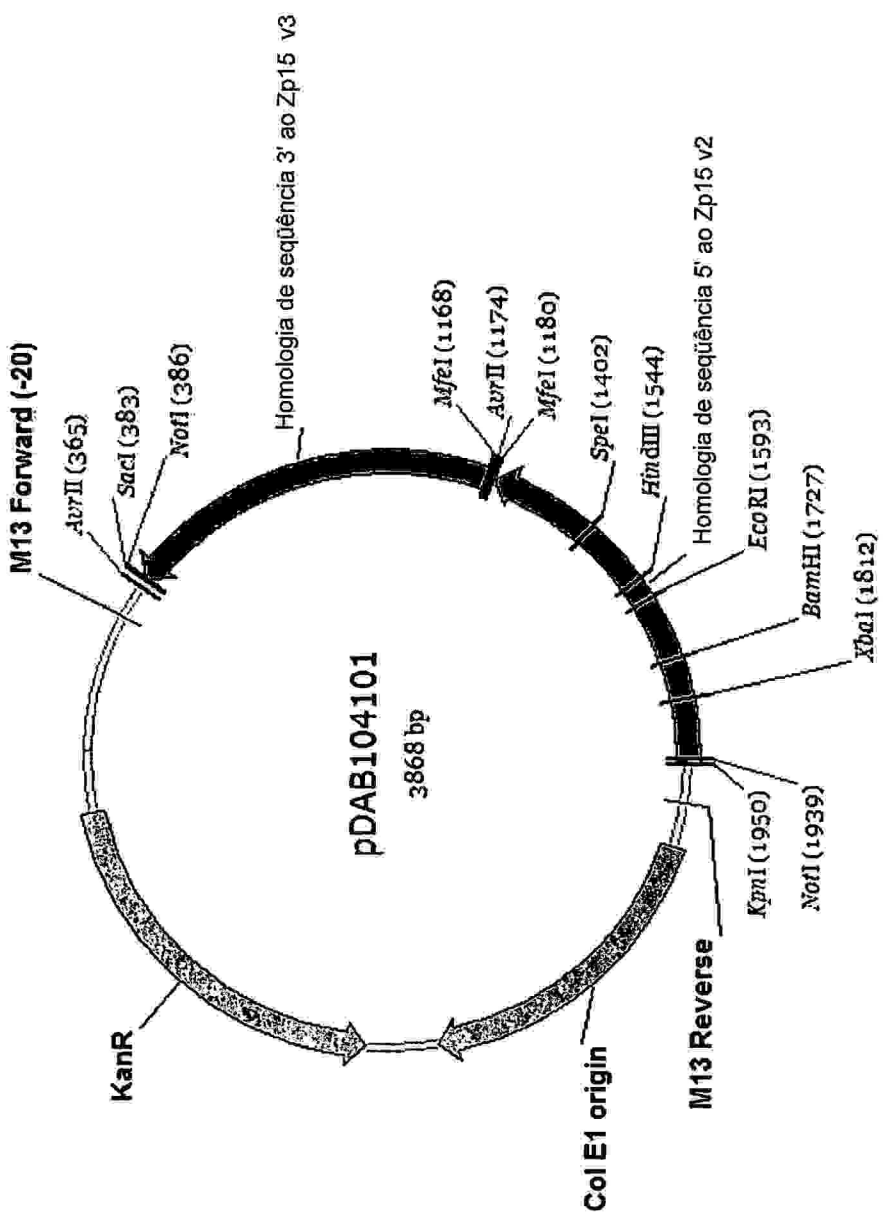

pDAB104101
3868 bp
M13 Forward (-20)
AvrII (365)
SacI (383)
NotI (386)
Homologia de seqüência 3' ao Zp15 v3
MfeI (1168)
AvrII (1174)
MfeI (1180)
SpeI (1402)
HindIII (1544)
Homologia de seqüência 5' ao Zp15 v2
EcoRI (1593)
BamHI (1727)
XbaI (1812)
KpnI (1950)
NotI (1939)
M13 Reverse
Col E1 origin
KanR

Fig. 7

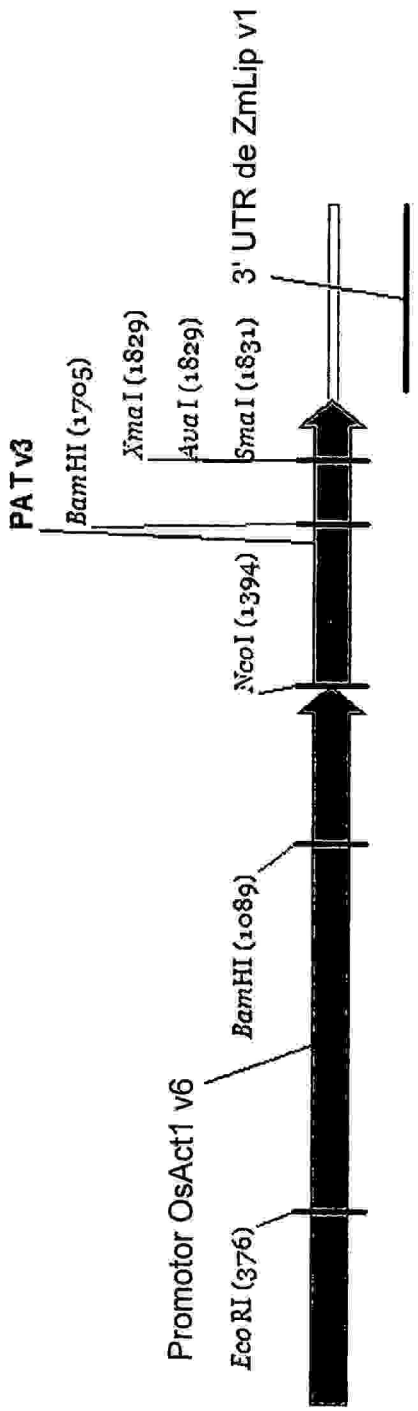

Promotor OsAct1 v6
Eco RI (376)
BamHI (1089)
NcoI (1394)
PAT v3
BamHI (1705)
XmaI (1829)
AvaI (1829)
SmaI (1831)
3' UTR de ZmLip v1
Cassete do gene PAT
2320 bp

Fig. 8

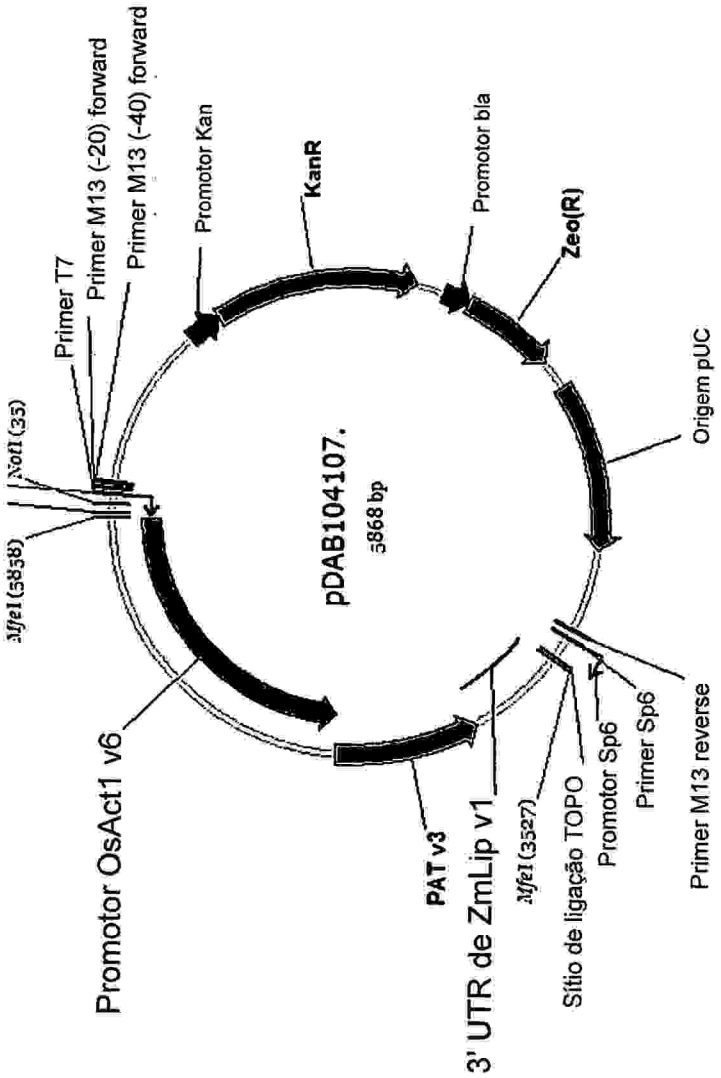

pDAB104107.
5868 bp
Promotor OsAct1 v6
NotI (35)
MfeI (5858)
Primer T7
Primer M13 (-20) forward
Primer M13 (-40) forward
Promotor Kan
KanR
Promotor bla
Zeo(R)
Origem pUC
Primer M13 reverse
Primer Sp6
Promotor Sp6
Sítio de ligação TOPO
MfeI (3527)
3' UTR de ZmLip v1
PAT v3

Fig. 9

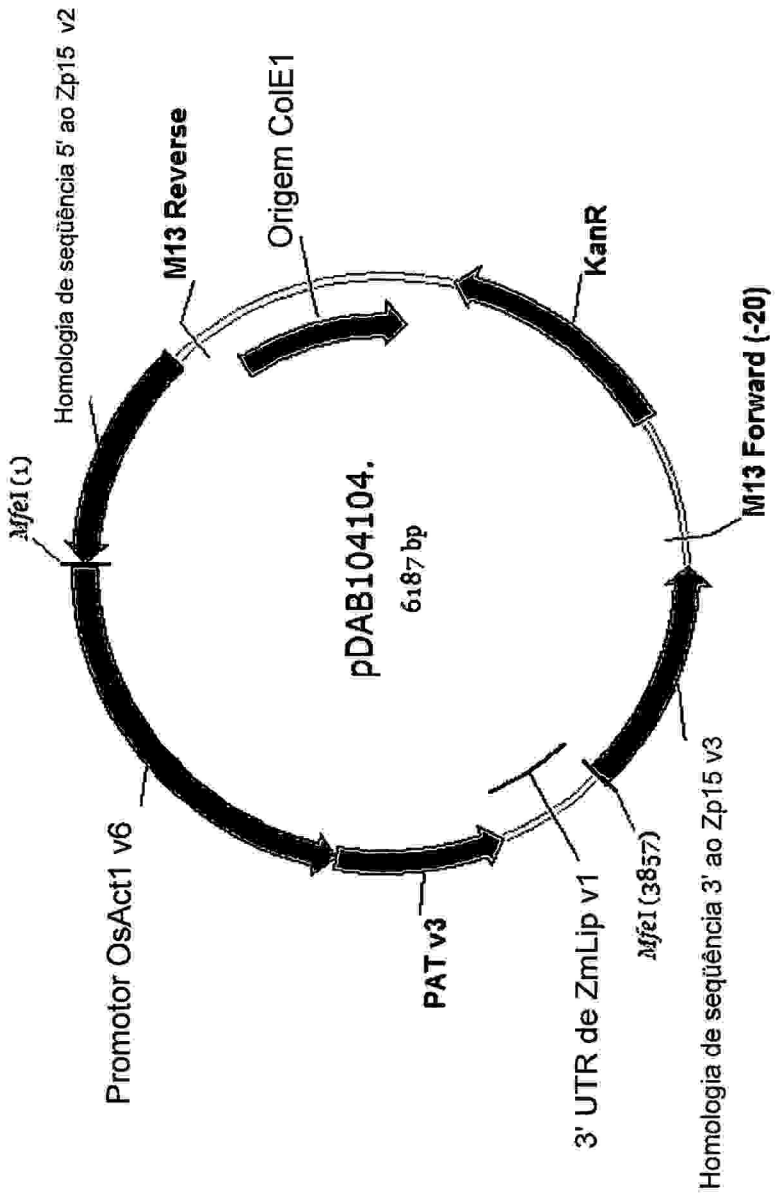

MfeI (1)
Homologia de seqüência 5' ao Zp15 v2
M13 Reverse
Origem ColE1
KanR
M13 Forward (-20)
pDAB104104.
6187 bp
Promotor OsAct1 v6
PAT v3
3' UTR de ZmLip v1
MfeI (3857)
Homologia de seqüência 3' ao Zp15 v3

Fig. 10

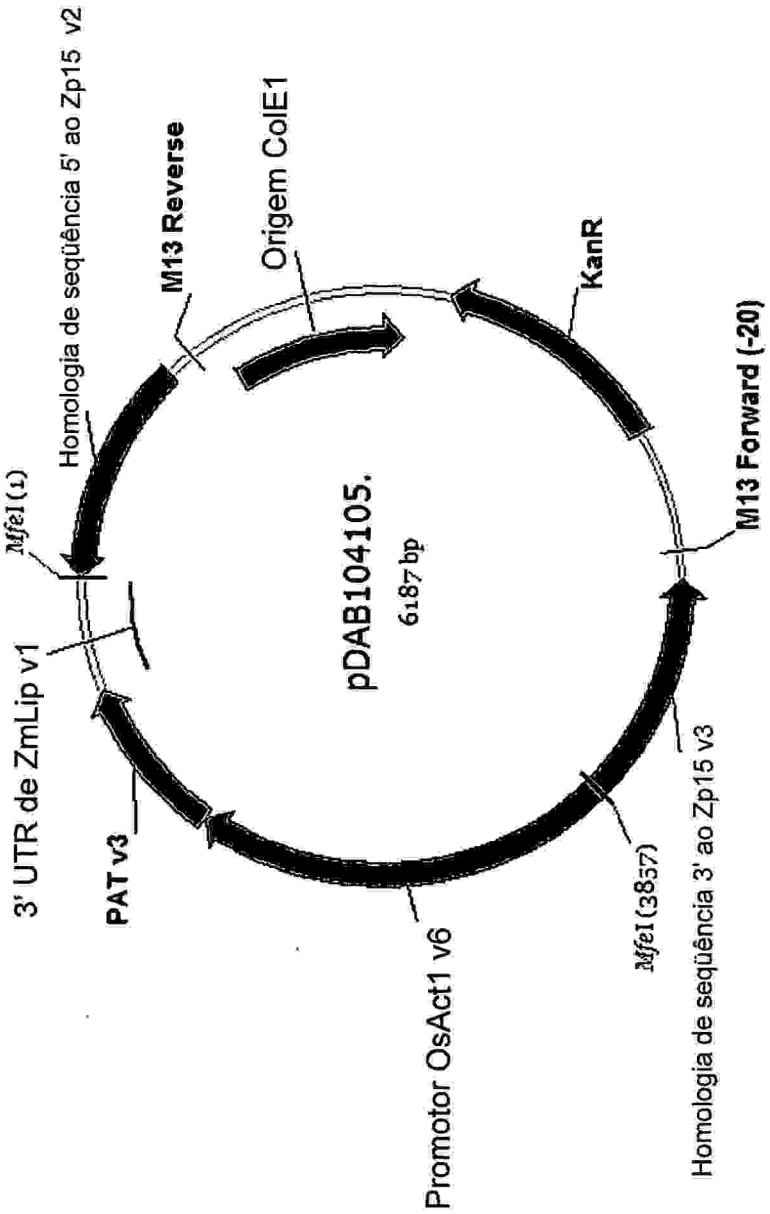

pDAB104105.
6187 bp
Homologia de seqüência 5' ao Zp15 v2
M13 Reverse
Origem ColE1
KanR
M13 Forward (-20)
Homologia de seqüência 3' ao Zp15 v3
MfeI (3857)
Promotor OsAct1 v6
PAT v3
3' UTR de ZmLip v1
MfeI (1)

Fig. 11

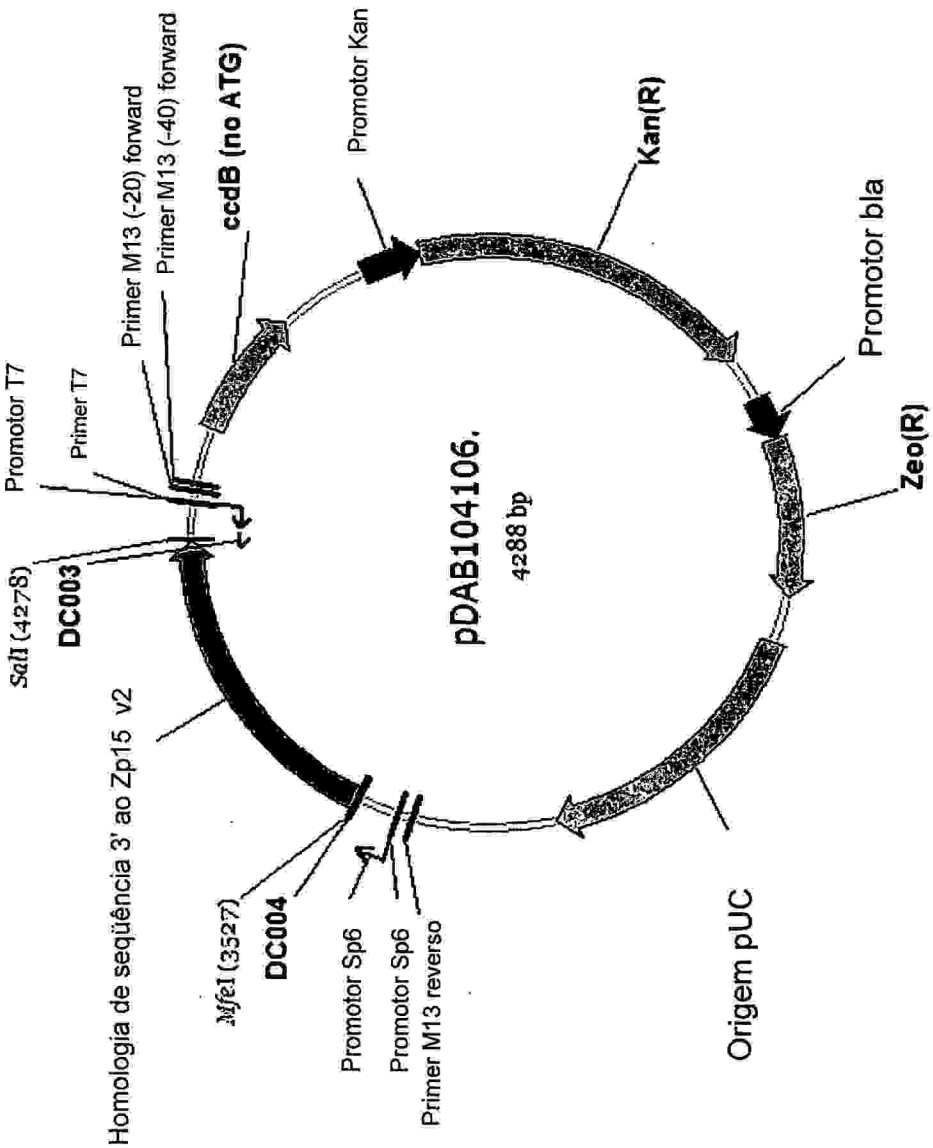

pDAB104106.
4288 bp
SalI (4278)
DC003
Promotor T7
Primer T7
Primer M13 (-20) forward
Primer M13 (-40) forward
ccdB (no ATG)
Promotor Kan
Kan(R)
Promotor bla
Zeo(R)
Origem pUC
Promotor Sp6
Promotor Sp6
Primer M13 reverso
MfeI (3527)
DC004
Homologia de seqüência 3' ao Zp15 v2

Fig.12

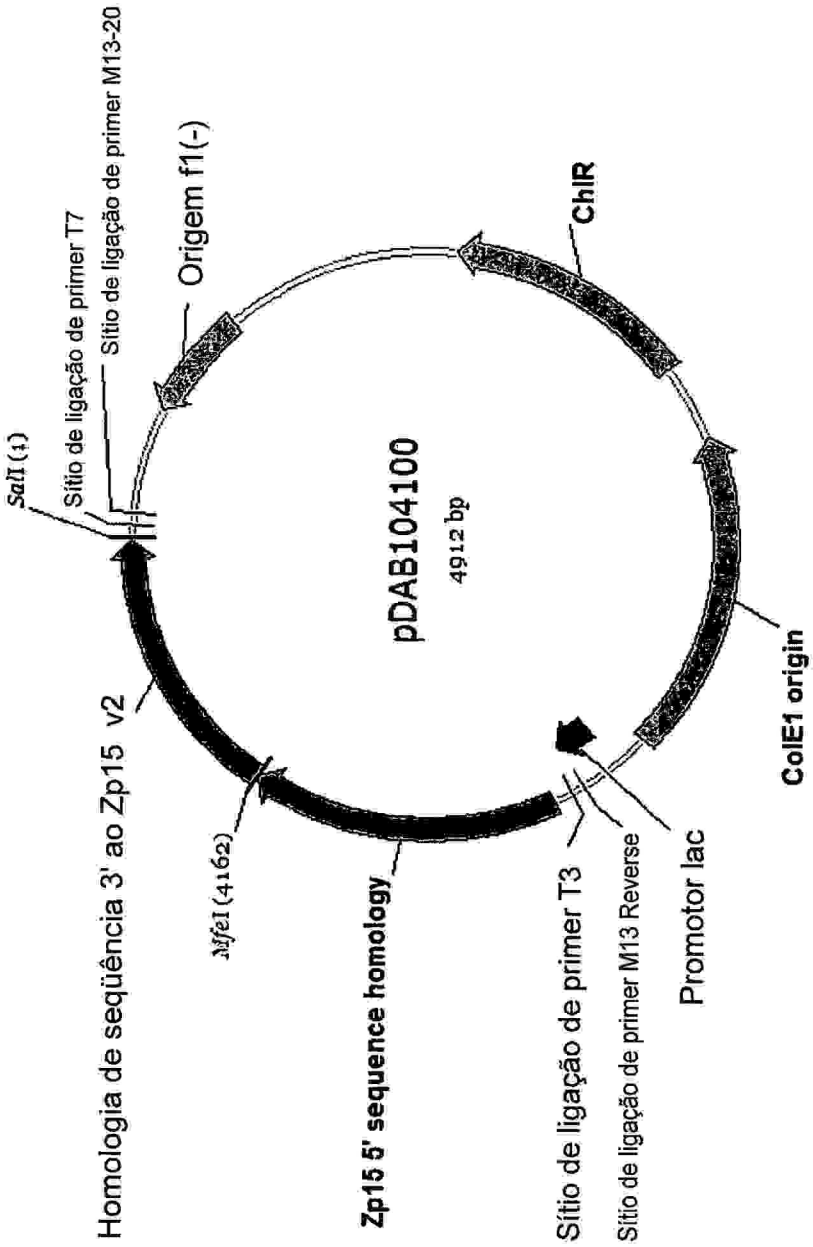

SalI (1)
Sítio de ligação de primer T7
Sítio de ligação de primer M13-20
Origem f1(-)
ChlR
pDAB104100
4912 bp
Homologia de seqüência 3' ao Zp15 v2
NsiI (4162)
Zp15 5' sequence homology
Sítio de ligação de primer T3
Sítio de ligação de primer M13 Reverse
Promotor lac
ColE1 origin

Fig. 13

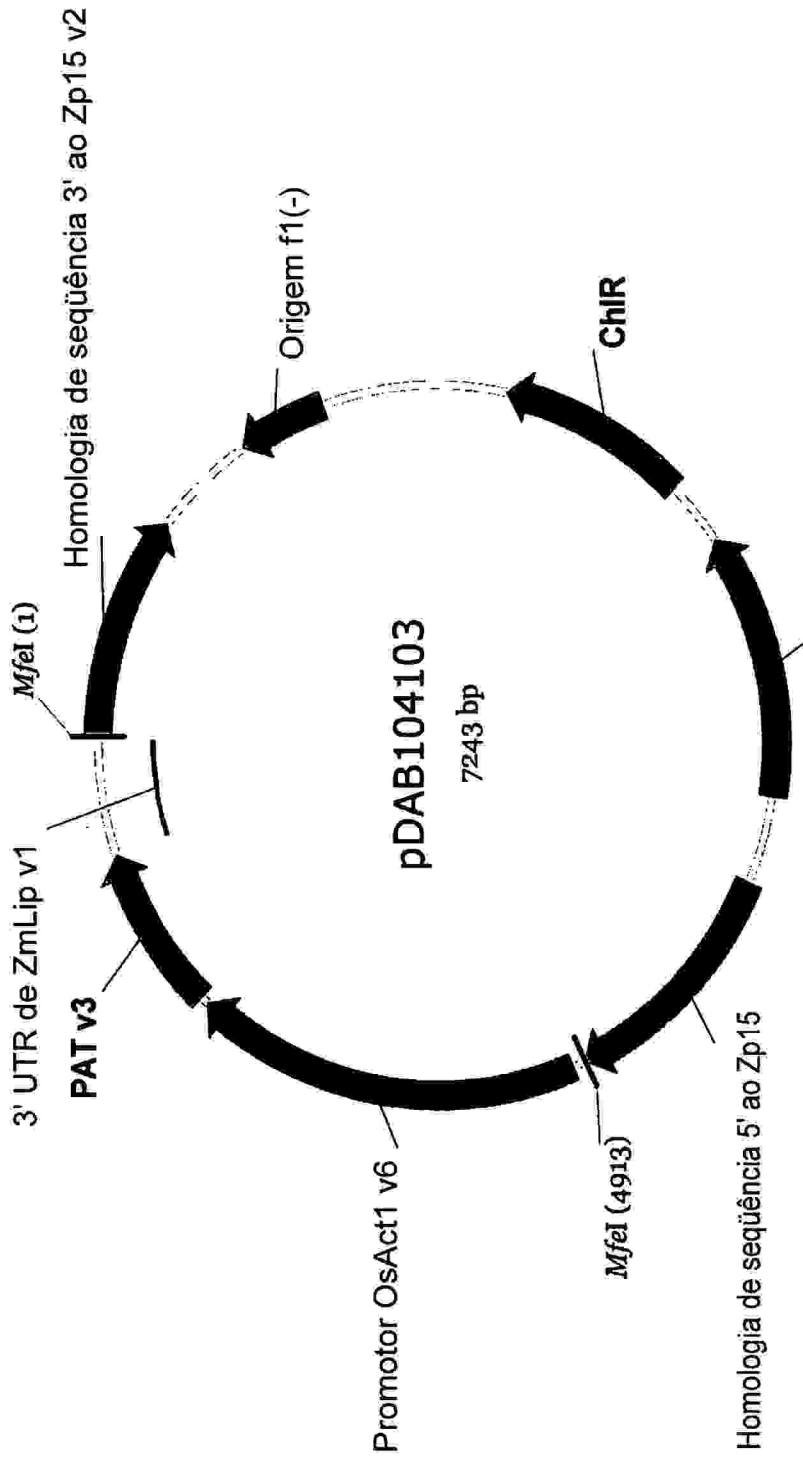

pDAB104103
7243 bp
Homologia de seqüência 3' ao Zp15 v2
Origem f1(-)
ChlR
Origem ColE1
Homologia de seqüência 5' ao Zp15
MfeI (4913)
Promotor OsAct1 v6
PAT v3
3' UTR de ZmLip v1
MfeI (1)

Fig. 14

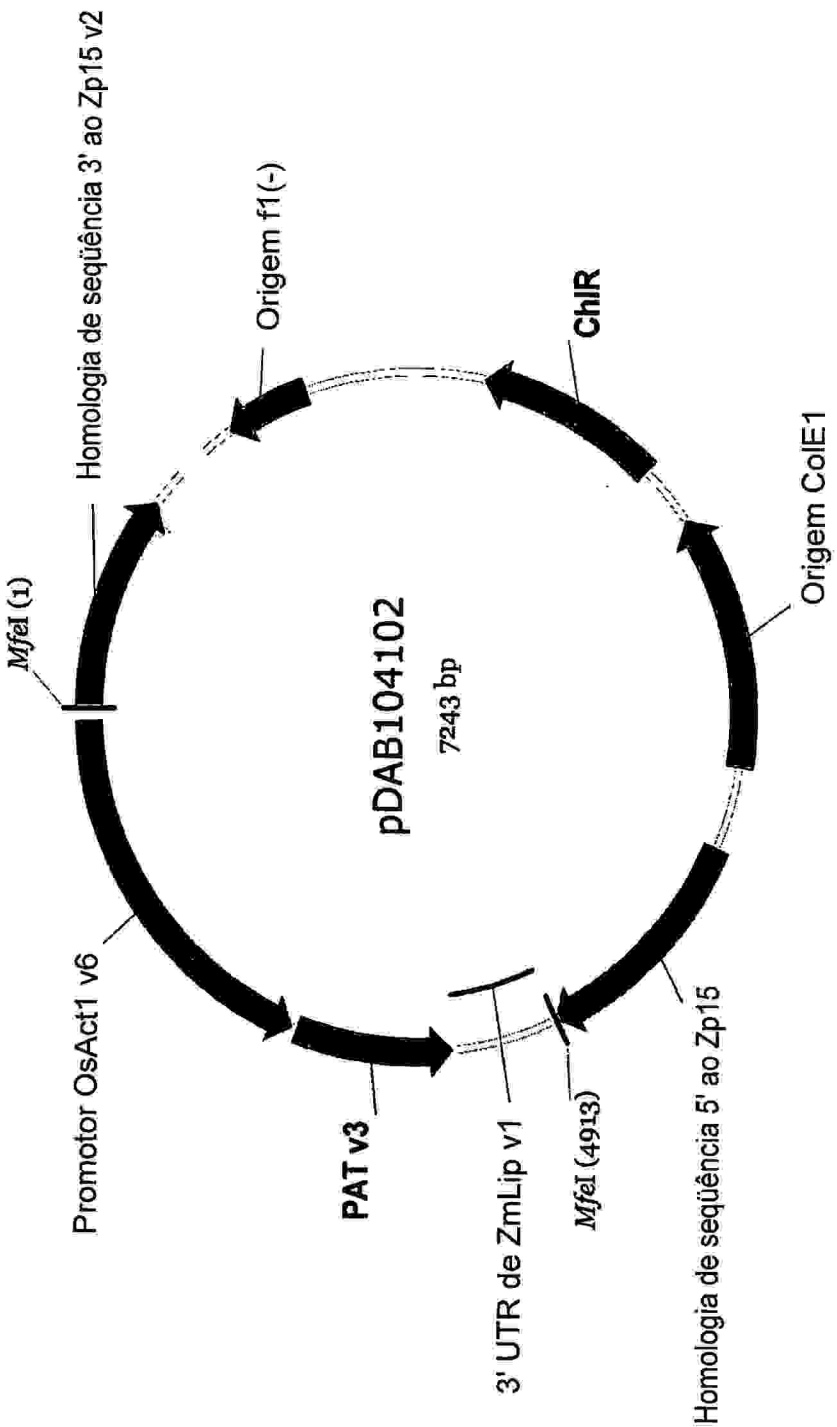

pDAB104102
7243 bp
MfeI (1)
Homologia de seqüência 3' ao Zp15 v2
Origem f1(-)
ChlR
Origem ColE1
Homologia de seqüência 5' ao Zp15
MfeI (4913)
3' UTR de ZmLip v1
PAT v3
Promotor OsAct1 v6

Fig.15