

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

(21) **PI 0712398-1 A2**

(22) Data de Depósito: 17/05/2007  
(43) Data da Publicação: 16/10/2012  
(RPI 2180)



(51) *Int.Cl.:*  
C12N 15/82  
A01H 5/00

**(54) Título:** MINICROMOSSOMO ARTIFICIAL DE PLANTA, PLANTA DE MILHO, POLINUCLEOTÍDEO ISOLADO, CONSTRUÇÃO RECOMBINANTE, PLANTA TRANSGÊNICA DE MILHO E MÉTODO PARA PRODUZIR UMA PLANTA TRANSGÊNICA DE MILHO COMPREENDENDO UM MINICROMOSSOMO ARTIFICIAL DE PLANTA TENDO CENTRÔMERO FUNCIONAL

**(57) Resumo:** MINICROMOSSOMO ARTIFICIAL DE PLANTA, PLANTA DE MILHO, POLINUCLEOTÍDEO ISOLADO, CONSTRUÇÃO RECOMBINANTE, PLANTA TRANSGÊNICA DE MILHO E MÉTODO PARA PRODUZIR UM PLANTA TRANSGÊNICA DE MILHO COMPREENDENDO UM MINICROMOSSOMO ARTIFICIAL DE PLANTA TENDO UM CENTRÔMERO FUNCIONAL. São descritos minicromossomos artificiais de planta compreendendo um centrômero funcional o qual se liga especificamente à proteína centromérica C (CENPC) e métodos para produzir tais minicromossomos.

**(30) Prioridade Unionista:** 17/05/2006 US 60/801,004

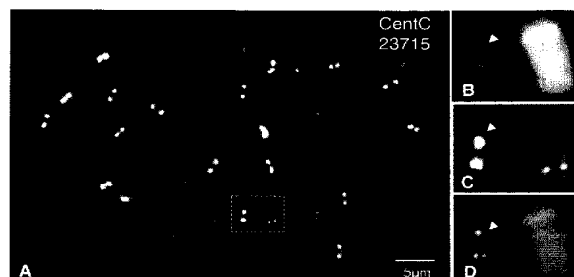
**(73) Titular(es):** Pioneer Hi-Bred International

**(72) Inventor(es):** Chengcang Wu, Evgeni Ananiev, Mark A. Chamberlim, Sergei Svitashhev, William J. Gordon-Kamm

**(74) Procurador(es):** Pedro Frankovsky Barroso

**(86) Pedido Internacional:** PCT US2007069139 de 17/05/2007

**(87) Publicação Internacional:** WO 2007/137114de 29/11/2007



MINICROMOSSOMO ARTIFICIAL DE PLANTA, PLANTA DE MILHO,  
POLINUCLEOTÍDEO ISOLADO, CONSTRUÇÃO RECOMBINANTE, PLANTA  
TRANSGÊNICA DE MILHO E MÉTODO PARA PRODUZIR UMA PLANTA  
TRANSGÊNICA DE MILHO COMPREENDENDO UM MINICROMOSSOMO  
5 ARTIFICIAL DE PLANTA TENDO UM CENTRÔMERO FUNCIONAL

#### CAMPO DA INVENÇÃO

Esta invenção é no campo de biotecnologia vegetal, em particular, ela pertence à minicromossomos artificiais e aos métodos de fazer tais minicromossomos em uma planta.

10

#### FUNDAMENTOS DA INVENÇÃO

Recentes avanços na engenharia de cromossomos têm feito ser possível alterar o genoma da planta, então, alterando seu fenótipo. Quando um transgene é integrado  
15 dentro do genoma da planta, geralmente é de caráter aleatório e em um número de cópia imprevisível. Assim sendo, esforços de pesquisa têm sido direcionados em direção ao melhor controle da integração do transgene.

20 Dada esta necessidade, pesquisadores têm imaginado se a resposta pode ser encontrada no uso de minicromossomos artificiais. Estas são moléculas de DNA linear ou circular, feitas pelo homem construídas de elementos de seqüência de DNA agindo de forma cis que provê replicação e  
25 particionamento dos minicromossomos construídos.

Acredita-se que a produção de cromossomos artificiais possa reduzir ou eliminar algumas questões associadas com integrações genômicas ao acaso dentro de um cromossomo de  
30 plana nativo, por exemplo, uma ligação de arraste para associação do transgene com o material genômico da planta hospedeira. Cromossomos artificiais podem também prover

meios para liberar 10-100 vezes mais genes do que vetores de transformação padrões, e para prover grandes segmentos cromossômicos para complementação e/ou clonagem baseada em mapa.

5

Três componentes têm sido identificados para replicação, estabilidade e manutenção/herança de cromossomos artificiais: (i) seqüências de replicação autônomas que funcionam como origem de replicação; (ii) 10 telômeros que funcionam para estabilizar e manter a extremidade de cromossomos lineares; e, (iii) centrômeros que são o sítio de cinetócoro reunido para a segregação cromossômica própria na mitose e meiose. Centrômeros isolados de organismo unicelulares, tais como levedura, não 15 funcionam em eucariotos desenvolvidos.

A patente americana US 5.270.201, emitida por Richards *et al.* em 14 de dezembro de 1993, descreve cromossomos de planta artificiais baseados em telômeros e, opcionalmente, 20 um centrômero.

A patente americana US 7.119.250, emitida por Luo *et al.* em 10 de outubro de 2006, descreve composições de centrômeros vegetais.

25

A patente americana US 7.132.240, emitida por Richards *et al.* em 7 de novembro de 2006, descreve um método para isolar um centrômero metilado de DNA potencialmente de qualquer centrômero em um organismo.

A patente americana US 7.193.128, emitida por Copenhaver *et al.* em 20 de março de 2007, descreve um método para gerar ou aumentar rendimento de cultivares usando seqüências de ácido nucléico de centrômeros de  
5 planta.

O pedido PCT tendo número de publicação WO 2007/030510 que foi publicado em 15 de março de 2007 descreve métodos de fazer plantas transformadas com minicromossomos  
10 autônomos.

#### **SUMÁRIO DA INVENÇÃO**

A presente invenção diz respeito a um minicromossomo de planta compreendendo um centrômero funcional contendo:  
15 (a) pelo menos duas matrizes de repetições em tandem de CentC em uma orientação invertida onde a primeira matriz compreende pelo menos cinquenta cópias de CentC e a segunda matriz compreende pelo menos cinquenta cópias de CentC; e,  
(b) pelo menos uma cópia de um elemento retrotransponível,  
20 onde o elemento retrotransponível está situado entre a primeira e a segunda matriz.

Em uma segunda concretização, um minicromossomo artificial de planta da invenção compreende um elemento  
25 retrotransponível selecionado do grupo consistindo de CentA, CRM1, e CRM2.

Em uma terceira concretização, o minicromossomo artificial de planta da invenção também compreende pelo

menos um telômero funcional.

Em uma quarta concretização, o centrômero funcional, compreendido pelo minicromossomo artificial de planta, se  
5 liga especificamente à proteína C centromérica (CENPC).

Em uma quinta concretização, uma planta de milho pode compreender qualquer dos minicromossomos artificiais da invenção.

10

Em uma sexta concretização, a presente invenção diz respeito a um minicromossomo artificial de planta compreendendo um centrômero funcional, onde o centrômero se liga especificamente à proteína centromérica C (CENPC).

15

Em uma sétima concretização, a invenção diz respeito a um polinucleotídeo isolado compreendendo: (a) pelo menos duas matrizes de repetição em tandem de CentC em uma orientação invertida onde a primeira matriz compreende pelo  
20 menos dez cópias de CentC e a segunda matriz compreende pelo menos dez cópias de CentC; e, (b) pelo menos uma cópia de um elemento retrotransponível, onde o elemento retrotransponível está situado entre a primeira e a segunda matriz.

25

Em uma oitava concretização, o polinucleotídeo isolado da invenção compreende um elemento retrotransponível que é selecionado do grupo consistindo de CentA, CRM1, e CRM2.

Em uma nona concretização, a invenção diz respeito a um polinucleotídeo isolado compreendendo: (a) pelo menos uma matriz de repetições em tandem de CentC, a matriz compreendendo pelo menos 10 cópias de CentC; e, (b) pelo menos uma cópia de um elemento retrotransponível selecionado do grupo consistindo de CentA, CRM1, e CRM2.

Em uma décima concretização, a invenção diz respeito a um polinucleotídeo isolado compreendendo: (a) pelo menos uma matriz de repetições em tandem de CentC, a matriz compreendendo pelo menos 10 cópias de CentC; e, (b) pelo menos uma cópia de cada um dos CentA, CRM1, e CRM2.

Em uma décima primeira concretização, a invenção diz respeito a uma construção recombinante compreendendo qualquer dos polinucleotídeos isolados da invenção bem como planta de milho transgênica compreendendo tais construções recombinantes.

Em uma décima segunda concretização, a invenção diz respeito a um método para fazer uma planta transgênica de milho compreendendo um minicromossomo artificial de planta tendo um centrômero funcional sendo o método compreendendo:

(a) contactar pelo menos uma célula vegetal de milho com uma mistura compreendendo uma construção recombinante da invenção;

(b) identificar pelo menos uma célula vegetal de milho da etapa (a) compreendendo um minicromossomo artificial de

planta tendo um centrômero funcional; e

(c) regenerar uma planta de milho fértil da célula vegetal de milho da etapa (b) onde a dita planta de milho  
5 compreende um minicromossomo de planta artificial tendo um centrômero funcional. A mistura pode também compreender um polinucleotídeo codificando um polipeptídeo para estimular o crescimento celular onde o polipeptídeo é selecionado do grupo consistindo de um wuschel, a baby boom, a RepA, ou  
10 uma Lecl.

#### **BREVE DESCRIÇÃO DAS FIGURAS**

Esta patente ou pedido de patente contém pelo menos um desenho executado em cor. Cópias desta patente ou pedido de  
15 patente com desenhos coloridos serão providos pelo escritório mediante um pagamento da taxa necessária.

A invenção pode ser melhor entendida através da seguinte descrição detalhada e desenhos e listagem de  
20 seqüências acompanhantes, que formam uma parte deste pedido.

**Figura 1.** Hibridização fluorescente *in situ* (FISH) em uma distribuição cromossômico mitótica de calos  
25 embriogênicos de milho da transformação Hi-II CMC3 grupo 1 evento #14. Calos foram derivados de embriões imaturos transformados com grupo 1 do clone BAC linearizado readaptado com Tn5-3. Núcleos na prometáfase (esquerda) e metáfase (direita) mostram os 20 cromossomos nativos mais 1

minicromossomo (setas e insertos). Ambos minicromossomos são positivos para o CentC (verde - colorido; branco - escala de cinza) específica para repetição centrômero e a sonda de marcador único 23715 (vermelho - colorido; branco - escala de cinza) específica para transformação da construção, com ambos os CentC e 23715 sendo essencialmente co-localizados no minicromossomo (insertos).

**Figura 2.** FISH em uma distribuição cromossômico mitótica de calos embriogênicos de milho da transformação Hi-II CMC3 grupo 1 evento #14. Calos foram derivados de embriões imaturos transformados com grupo 1 do clone BAC linearizado readaptado com Tn5-3. Paineis A mostra um núcleo em metáfase mostrando os 20 cromossomos nativos mais 2 minicromossomos (quadro). Ambos os minicromossomos são positivos para o CentC (verde - colorido; branco - escala de cinza) específica para repetição centrômero e a sonda marcadora única 23715 (vermelho - colorido; branco - escala de cinza) específica para a transformação da construção. Painéis B-D têm uma ampliação do quadro mostrando os minicromossomos (seta) com: B - DAPI apenas; C - DAPI + sonda 23715 (vermelho - colorido; branco - escala de cinza); e D - DAPI + sonda CentC (verde - colorido; branco - escala de cinza).

25

**Figura 3.** Imunofluorescência em uma distribuição cromossômico mitótica de calos embriogênicos de milho da transformação Hi-II CMC3 grupo 1 evento #14. Calos foram derivados de embriões imaturos transformados com grupo 1 do



clone BAC linearizado readaptado com Tn5-3. Painel A mostra um núcleo em metáfase mostrando os 20 cromossomos nativos mais 1 minicromossomo (seta). Todos os centrômeros dos cromossomos nativos e minicromossomos são positivos para a proteína centromérica C, CENPC (vermelho - colorido; branco - escala de cinza), uma proteína específica de centrômero/cinetócoro. Painéis B-C mostram ampliação do minicromossomo com: B - DAPI apenas; e C - DAPI + CENPC (vermelho - colorido; branco - escala de cinza). A morfologia e imunolocalização de CENPC indicam que o minicromossomo é composto de duas cromátides irmãs cada uma tendo um centrômero funcional.

**Figura 4.** Painel A - Imunofluorescência em uma distribuição cromossômica mitótica de calos embriogênicos de milho da transformação Hi-II CMC3 grupo 1 evento #14. Calos foram derivados de embriões imaturos transformados com grupo 1 do clone BAC linearizado readaptado com Tn5-3. A separação das cromátides irmãs dos cromossomos nativos e o minicromossomo (quadro) foram observados durante a anáfase. Todos os centrômeros dos cromossomos nativos e o minicromossomo são positivos para a proteína centromérica C, CENPC (vermelho - colorido; branco - escala de cinza) uma proteína específica de centrômero/cinetócoro. Painel B é uma imagem amplificada do quadro em A, mostrando a separação das cromátides irmãs do minicromossomo (seta dupla) indicando que o minicromossomo, como cromossomos normais, pode segregar durante a mitose.

**Figura 5.** FISH em uma distribuição cromossômico mitótica de calos embriogênicos de milho da transformação Hi-II CMC3 grupo 3 evento #12. Calos foram derivados de embriões imaturos transformados com grupo 3 do clone BAC linearizado readaptado com Tn5-3. Um núcleo metafásico tetra-aneuploide (39 cromossomos, perdendo uma cópia do cromossomo 6) mostrando os cromossomos nativos mais 1 minicromossomo (seta) é apresentado. O minicromossomo é positivo para a CentC (verde - colorido; branco - escala de cinza) específica para repetição centrômero e a sonda marcadora única 23715 (vermelho - colorido; branco - escala de cinza) específica para a transformação da construção. Painéis B-D são uma amplificação da área do quadro mostrando o minicromossomo (ponta da seta) e um cromossomo nativo com: A - DAPI apenas; B - DAPI + sonda CentC (verde - colorido; branco - escala de cinza); e D - DAPI + sonda 23715 (vermelho - colorido; branco - escala de cinza). Localização bipolar de repetições CentC como revelado por coloração FISH no minicromossomo indicam que ele é composto de duas cromátides irmãs similares aquelas observadas nos cromossomos nativos.

**Figura 6.** FISH em uma distribuição cromossômico mitótica de calos embriogênicos de milho da transformação Hi-II CMC3 grupo 3 evento #12. Calos foram derivados de embriões imaturos transformados com grupo 3 do clone BAC linearizado readaptado com Tn5-3. Painel A - cromossomos tetra-aneuploides (39 cromossomos, perdendo uma cópia do cromossomo 6) de um núcleo metafásico mostrando os

cromossomos nativos mais 2 minicromossomos (setas). Minicromossomos são positivos para ambos CentC (verde - colorido; branco - escala de cinza) específica para repetição centrômero e a sonda marcadora única 23715 (vermelho - colorido; branco - escala de cinza) específica para a transformação da construção. Painel B é uma amplificação da imagem dos 2 minicromossomos mostrando a variação em abundância das repetições CentC e o marcador único 23715.

10

**Figura 7.** FISH em uma distribuição cromossômico mitótica de calos embriogênicos de milho da transformação Hi-II CMC3 grupo 3 evento #12. Calos foram derivados de embriões imaturos transformados com grupo 3 do clone BAC linearizado readaptado com Tn5-3. Painel A - cromossomos tetra-aneuploides (39 cromossomos, perdendo uma cópia do cromossomo 6) de um núcleo mostrando a separação das cromátides irmãs dos cromossomos nativos e dois minicromossomos (quadro) no início da anáfase. As cromátides irmãs de ambos os minicromossomos são positivas para o CentC (verde - colorido; branco - escala de cinza) específica para repetição centrômero e a sonda marcadora única 23715 (vermelho - colorido; branco - escala de cinza) específica para a transformação da construção. Painéis B-C são uma ampliação das imagens dos 2 minicromossomos (seta dupla) mostrando: B - DAPI + sonda CentC (verde - colorido; branco - escala de cinza); e C - DAPI + sonda 23715 (vermelho - colorido; branco - escala de cinza). Separação das cromátides irmãs do minicromossomo na anáfase sugere a

15

20

25

presença de centrômeros funcionais, permitindo a segregação durante a mitose.

**Figura 8.** Imunofluorescência em uma distribuição cromossômica mitótica de calos embriogênicos de milho da transformação Hi-II CMC3 grupo 3 evento #12. Calos foram derivados de embriões imaturos transformados com grupo 3 do clone BAC linearizado readaptado com Tn5-3. Painei A mostra um núcleo metafásico tetra-aneuplóide (39 cromossomos, perdendo uma cópia do cromossomo 6) mostrando 39 cromossomos nativos mais 2 minicromossomos (setas). Todos os centrômeros dos cromossomos nativos e os minicromossomos são positivos para a proteína centromérica C, CENPC (vermelho - colorido; branco - escala de cinza) uma proteína específica de centrômero/cinetócoro. Painéis B-C são imagens amplificadas dos minicromossomos. O padrão de imunolocalização de CENPC, dois loci por minicromossomo, indica que o minicromossomo é composto de duas cromátides irmãs e cada uma tem um centrômero funcional hábil para formar um complexo cinetócoro.

**Figura 9.** FISH em uma distribuição cromossômica mitótica da extremidade da raiz de uma planta regenerada de um evento de transformação de milho Hi-II. Plantas foram derivadas de embriões imaturos transformados com bacm.pk128.j21 linearizados readaptado com Tn5-3. Painei A mostra um núcleo metafásico aneuplóide apresentando 19 cromossomos nativos mais 1 minicromossomo (seta). O minicromossomo é positivo para o CentC (verde - colorido;

branco - escala de cinza) específica para repetição centrômero e a sonda marcadora única 23715 (vermelho - colorido; branco - escala de cinza) específica para a transformação da construção. Painéis B-D são ampliações do minicromossomo com: B - DAPI apenas; C - DAPI + sonda CentC (verde - colorido; branco - escala de cinza); e D - DAPI + sonda 23715 (vermelho - colorido; branco - escala de cinza).

10 **Figura 10.** Imunofluorescência em uma distribuição cromossômico mitótica da extremidade da raiz de uma planta regenerada de um evento de transformação de milho Hi-II. Plantas foram derivadas de embriões imaturos transformados com bacm.pk128.j21 linearizados readaptado com Tn5-3.

15 Painel A mostra um núcleo metafásico aneuplóide apresentando 19 cromossomos nativos mais 1 minicromossomo (seta). Todos os centrômeros dos cromossomos nativos e os minicromossomos são positivos para a proteína centromérica C, CENPC (vermelho - colorido; branco - escala de cinza),

20 uma proteína específica de centrômero/cinetócoro. Painéis B-C são ampliações do minicromossomo com: B - DAPI apenas; e C - DAPI + CENPC. O padrão de imunolocalização CENPC, dois loci por minicromossomo, indica que o minicromossomo é composto de duas cromátides irmãs e cada

25 uma tem um centrômero funcional hábil para formar um complexo cinetócoro.

**Figura 11.** Fina estrutura de centrômeros de milho revelados por fibra-FISH. Quatro repetições centroméricas,

CentC (verde - colorido; branco - escala de cinza) e uma soma de CentA, CRM1, e CRM2 (vermelho - colorido; verde - escala em cinza) foram usadas em FISH multicolorido em fibras de DNA estendidas de linhagens adicionais de aveia-  
5 milho contendo cromossomos individuais de milho. Isto revelou estiramentos de hibridização megabase-alongada, que são únicas para cada cromossomo.

**Figura 12.** Modelo de um centrômero de milho.  
10 Organização centromérica é mostrada usando nomenclatura repetida centromérica de milho. Arranjos ininterruptos de CentC podem ser compostos de várias centenas a milhares de elementos repetidos. Outros elementos retrotransponíveis específicos de centrômero de milho tais como CentA, CRM1,  
15 e/ou CRM2 podem ser integrados dentro de um arranjo CentC, entre si, e/ou nele mesmo em regiões centroméricas. Em adição aos retrotransposons específicos de centrômeros, outros retrotransposons podem ser integrados no arranjo, em elementos como CentA, CentC, CRM1, e CRM2, e/ou nele mesmo  
20 para formar insertos que interrompam arranjos repetidos em tandem CentC. Esta figura mostra um modelo de organização dos elementos CentC de milho (extremidade da seta) formando dois arranjos de repetições cabeça-para-cauda em tandem. Os arranjos CentC podem ser encontrados em uma orientação  
25 invertida para formar um grande segmento do DNA centromérico. Fibra-FISH alongada com FISH nos cromossomos na anáfase meiótica e análises de hibridização-blot de segmentos de DNA centroméricos clonados indicam que regiões com alta densidade de todas as quatro repetições

centroméricas (CentC, CRM1, CentA, e CRM2) estão envolvidas na formação do cinetócoro.

**Figura 13.** Reequipamento e conversão *in vitro* de um clone BAC em um minicromossomo artificial linear. Clone BAC de DNA é readaptado com transposon Tn5-3 feito como de costume compreendendo gene de resistência à ampicilina (AP<sup>r</sup>), origem de replicação (ori), marcadores de seleção (MO-PAT) e visual (DS-RED2) sob o promotor da ubiquitina (UBI1ZM PRO), seqüências teloméricas (TEL) na orientação reversa separadas por um gene de resistência à canamicina (gene KAN<sup>r</sup>), e sítios para enzimas de restrição da fonte I-Ppo I, I-Ceu I, e PI-Sce I. Suporte ME para extremidade em mosaico de transposon. Digestão da construção BAC com enzimas de restrição da fonte I-Ceu I converte um BAC circular em uma molécula de DNA linear flanqueada com seqüências teloméricas.

**Figura 14.** Núcleo metafásico de calos de CMC3 grupo 1 evento #14 marcado para os elementos centrômero e telômeros. Análises FISH foram feitas usando sondas marcadas fluorescentemente para a repetição específica para centrômero CentC (verde - colorido; branco - escala em cinza) e repetições específicas para telômeros telo-31 (vermelho - colorido; branco - escala em cinza). Localização dessas sondas é notada para um cromossomo nativo, CentC é indicada por asteriscos (\*), e telo-31 indicada por setas duplas. Painéis B-E mostram uma amplificação do minicromossomo. Painel B - DAPI + Cent C +

telo31 (verde/vermelho - colorido; branco - escala em cinza); C - DAPI apenas; D - DAPI + sonda CentC (verde - colorido; branco - escala em cinza); e E - DAPI + sonda 23715 (vermelho - colorido; branco - escala em cinza). O padrão da hibridização telo-31 sugere que o minicromossomo (seta) tem telômeros funcionais similares aos cromossomos nativos.

**Figura 15.** Núcleo metafásico de calos CMC3 subgrupo 1.3 evento #27 marcados para os elementos de centrômero e telômeros. Análises FISH foram feitas usando sondas marcadas fluorescentemente para repetições específicas de centrômero CentC (verde - colorido; branco - escala em cinza) e repetições específicas para telômeros telo-31 (vermelho - colorido; branco - escala em cinza). Localização dessas sondas é notada para um cromossomo nativo, CentC é indicada por asteriscos (\*), e telo-31 indicada por setas duplas. Painéis B-E mostram amplificação do minicromossomo. Painel B - DAPI + Cent C + telo-31 (verde/vermelho - colorido; branco - escala em cinza); C - DAPI apenas; D - DAPI + sonda CentC (verde - colorido; branco - escala em cinza); e E - DAPI + sonda 23715 (vermelho - colorido; branco - escala em cinza). O padrão da hibridização telo-31 sugere que o minicromossomo (seta) tem telômeros funcionais similares a cromossomos nativos.

#### DESCRIÇÃO DETALHADA DA INVENÇÃO

A revelação de cada referência descrita aqui é por



meio desta incorporada por referência em sua totalidade.

Como usado aqui nas reivindicações anexadas, as formas singulares "um", "uma", "o" e "a" incluem referências no plural a menos que o contexto claramente indique de outra  
5 forma. Então, por exemplo, a referência para "uma planta" inclui uma pluralidade de tais plantas, a referência para "uma célula" inclui uma ou mais células e equivalentes das mesmas conhecidas para os especialistas no estado da  
10 técnica, e assim por diante.

No contexto desta revelação, um número de termos e abreviações é usado. As seguintes definições são providas.

15 "Fase aberta de leitura" é abreviada como ORF.

"American Type Culture Collection" é abreviada como ATCC.

20 O termo "minicromossomo artificial de planta" como usado aqui refere-se a qualquer cromossomo criado artificialmente compreendendo um centrômero e telômeros que possuam propriedades comparáveis com aquelas de um cromossomo nativo, tais como replicação e segregação  
25 durante a mitose e meiose e conseqüentemente autonomia e transmissibilidade na divisão celular. Os termos minicromossomo artificial, minicromossomo, e cromossomo artificial são usados alternadamente aqui.

O termo "centrômero funcional" refere-se a uma região de ligação ao fuso de um cromossomo eucariótico que funciona de uma maneira comparável a centrômeros em um cromossomo nativo. É a região mais condensada e contraída de um cromossomo, para a qual a fibra do fuso é fixada durante a mitose. Durante a mitose em uma planta ou célula animal típica, cada cromossomo divide-se longitudinalmente em dois cromossomos irmãos que eventualmente separam e migram para pólos opostos do fuso mitótico. No início da mitose, quando os cromossomos irmãos estiverem divididos, mas estiverem ainda pareados, cada cromossomo fixa ao fuso em um ponto específico ao longo de seu comprimento. Esse ponto é referido como o centrômero ou região de fixação do fuso. Centrômeros são compostos de DNA altamente repetitivo, isto é, seqüências de DNA que estão presentes em um genoma em muitas cópias.

O termo "arranjo" refere-se a um arranjo ordenado de elementos.

20

O termo "repetição em tandem" refere-se a múltiplas cópias da mesma seqüência base na mesma orientação. Então, essas são cópias de seqüências de nucleotídeos, que são repetidas uma e outra vez um número de vezes em tandem, por exemplo, ao longo do cromossomo. Qualquer arranjo de repetições em tandem pode compreender múltiplas cópias de um simples elemento, ou pode ter pelo menos um outro elemento intercalado dentro do arranjo, ou dentro de um elemento do arranjo.

O termo "orientação invertida" refere-se a duas ou mais cópias da mesma seqüência presente de uma forma invertida.

5 Os termos "elemento retrotransponível" e "retrotransposon" são usados alternadamente aqui e referem-se a um elemento genético que transpõe para um novo local no DNA através de fazer primeiro uma cópia de RNA dele mesmo, depois fazer uma cópia de DNA deste RNA com uma  
10 transcriptase reversa, e então inserir a cópia de DNA dentro do DNA alvo. Retrotransposons são elementos genéticos que podem amplificar-se em um genoma e são componentes ubíquos do DNA de muitos organismos eucariotas. Eles são uma subclasse de transposon. Eles são  
15 particularmente abundantes em plantas, onde eles são freqüentemente um componente principal do DNA nuclear.

O termo "telômero funcional" refere-se a estruturas encontradas nas extremidades dos cromossomos nas células de  
20 eucariotos. Telômeros funcionam através da proteção das extremidades dos cromossomos de recombinação, fusão com outros cromossomos, ou degradação por nucleases. Eles permitem às células distinguir entre quebras aleatórias de DNA e extremidades de cromossomo. Eles também desempenham  
25 uma significativa função na determinação do número de vezes que uma célula normal pode dividir. Um telômero é uma região de DNA altamente repetitiva na extremidade de um cromossomo linear que funciona como um tampão disponível. Cada vez que cromossomos eucariotos lineares são replicados

durante a fase S o complexo DNA polimerase é incapaz de replicar todo o caminho até a extremidade do cromossomo; se não fosse pelos telômeros, isso poderia rapidamente resultar na perda da informação genética vital, que é  
5 necessária para sustentar uma atividade celular.

Como usado aqui, o termo "ácido nucléico" significa um polinucleotídeo e inclui polímeros de fita simples ou dupla de bases de deoxiribonucleotídeo ou ribonucleotídeo. Ácidos  
10 nucléicos podem também incluir fragmentos e nucleotídeos modificados. Então, os termos "polinucleotídeo", "seqüência de ácido nucléico", "seqüência de nucleotídeo" ou "fragmento de ácido nucléico" são usados de forma substituível para indicar um polímero de RNA ou DNA que é  
15 fita simples ou dupla, opcionalmente contendo bases de nucleotídeo sintéticas, não naturais ou alteradas. Nucleotídeos (usualmente encontrados na sua forma 5'-monofosfato) são referidos por sua designação de letras simples como segue: "A" para adenosina ou deoxiadenosina  
20 (para RNA ou DNA, respectivamente), "C" para citosina ou deoxicitosina, "G" para guanosina ou deoxiguanosina, "U" para uridina, "T" para deoxitimidina, "R" para purinas (A ou G), "Y" para pirimidinas (C ou T), "K" para G ou T, "H" para A ou C ou T, "I" para inosina, e "N" para qualquer  
25 nucleotídeo.

Os termos "subfragmento que é funcionalmente equivalente" e "subfragmento funcionalmente equivalente" são usados alternadamente aqui. Esses termos referem-se a

uma porção ou subsequência de um fragmento de ácido nucléico isolado onde a habilidade para alterar a expressão gênica ou produzir um certo fenótipo é retida se ou não o fragmento ou subfragmento codificar uma enzima ativa. Por exemplo, o fragmento ou subfragmento pode ser usado na construção de genes quiméricos para produzir o fenótipo desejado em uma planta transformada. Genes quiméricos podem ser construídos para uso na supressão através de ligação de um fragmento de ácido nucléico ou subfragmento dos mesmos, se ou não ele codificar uma enzima ativa, na orientação senso ou antisenso com relação à seqüência promotora vegetal.

O termo "domínio conservado" ou "motivo" significa um grupo de aminoácidos conservados em posições específicas ao longo de uma seqüência alinhada de proteínas relacionadas evolutivamente. Enquanto aminoácidos em outras posições podem variar entre proteínas homólogas, aminoácidos que são altamente conservados em posições específicas indicam aminoácidos que são essenciais na estrutura, estabilidade, ou atividade de uma proteína. Porque eles são identificados pelo seu alto grau de conservação em seqüências alinhadas de uma família de proteínas homólogas, eles podem ser usados como identificadores, ou "assinaturas", para determinar se uma proteína com uma seqüência recentemente determinada pertence a uma família de proteínas previamente identificada.

Os termos "homologia", "homólogo", "substancialmente

similar", "substancialmente idêntico", e "substancialmente correspondente" são usados alternadamente aqui. Eles referem-se a fragmentos de ácido nucléico onde mudanças em uma ou mais bases de nucleotídeos não afetam a habilidade do fragmento de ácido nucléico de mediar a expressão gênica ou produzir um certo fenótipo. Esses termos também referem-se a modificações dos fragmentos de ácido nucléico da presente invenção tais como deleção ou inserção de um ou mais nucleotídeos que não alteram substancialmente as propriedades funcionais do fragmento de ácido nucléico resultante relativo ao fragmento não modificado inicial. Esses termos também referem-se a seqüências de aminoácidos, polipeptídeos, ou fragmentos de peptídeos com ou sem modificações, deleções, inserções, ou substituições que não alteram substancialmente as propriedades funcionais relacionadas a uma seqüência não modificada inicial. É, portanto, entendido, como os especialistas no estado da técnica irão apreciar, que a invenção engloba mais do que as seqüências exemplares específicas.

20

Mais ainda, um especialista na área reconhece que seqüências de ácido nucléico substancialmente similares englobadas por esta invenção são também definidas por sua habilidade de hibridizar (sob condições moderadamente estridentes, pex., 0.5X SSC, 0.1% SDS, 60 °C) com as seqüências exemplificadas aqui, ou para qualquer porção das seqüências de nucleotídeo reveladas aqui e que são funcionalmente equivalentes a qualquer das seqüências de ácido nucléico reveladas aqui. Condições de estringência

25

podem ser ajustadas para rastrear fragmentos moderadamente similares, tais como seqüências homólogas de organismos distancialmente relacionados, para fragmentos altamente similares, tais como genes que duplicam enzimas funcionais de organismos intimamente relacionados. Lavagens pós-hibridização determinam condições de estringência.

O termo "hibridiza seletivamente" inclui referência a hibridização, sob condições estringentes de hibridização, de uma seqüência de ácido nucléico para uma seqüência alvo de ácido nucléico especificada a um grau altamente detectável (pex., pelo menos 2 vezes sobre o background) do que sua hibridização a seqüências de ácido nucléico não alvo e a exclusão substancial de ácidos nucléico não alvo. Seqüências hibridizando seletivamente tipicamente têm cerca de pelo menos 80% de identidade de seqüência, ou 90% de identidade de seqüência, até e incluindo 100% de identidade de seqüência (i.e., totalmente complementar) umas com as outras.

20

O termo "condições estringentes" ou "condições estringentes de hibridização" faz referência a condições sobre as quais uma sonda hibridizará seletivamente com sua seqüência alvo. Condições estringentes são seqüência-dependente e serão diferentes em diferentes circunstâncias. Através do controle da estringência de hibridização e/ou condições de lavagem, seqüências alvo podem ser identificadas que sejam 100% complementares à sonda (marcação homóloga). Alternativamente, condições de

estringência podem ser ajustadas para permitir algum mal emparelhamento nas seqüências de modo que baixos graus de similaridade sejam detectados (marcação heteróloga). Geralmente, uma sonda tem menos do que cerca de 1000  
5 nucleotídeos em comprimento, opcionalmente menos do que 500 nucleotídeos em comprimento.

Tipicamente, condições estringentes serão aquelas onde a concentração de sal é menor do que cerca de 1.5 M de íon Na, tipicamente cerca de 0.01 a 1.0 M de concentração de  
10 íon Na (ou outros sais) em pH 7.0 a 8.3 e a temperatura é pelo menos cerca de 30 °C para sondas curtas (pex., 10 a 50 nucleotídeos) e pelo menos cerca de 60 °C para sondas longas (pex., maior do que 50 nucleotídeos). Condições  
15 estringentes podem também ser alcançadas com a adição de agentes desestabilizadores tais como formamida. Baixas condições de estringência exemplares incluem hibridização com uma solução tampão de 30 a 35% de formamida, 1 M de NaCl, SDS 1% (dodecil sulfato de sódio) a 37 °C, e uma  
20 lavagem em 1X a 2X SSC (20X SSC = 3.0 M NaCl/0.3 M de citrato de trisódio) de 50 a 55 °C. Condições de estringência moderadas exemplares incluem hibridização em formamida de 40 a 45%, 1 M de NaCl, SDS 1% a 37 °C, e uma lavagem em 0.5X a 1X de SSC de 55 a 60 °C. Condições de  
25 alta estringência exemplares incluem hibridização em formamida 50%, 1 M de NaCl, SDS 1% a 37 °C, e uma lavagem em SSC 0.1X de 60 a 65 °C.

Especificidade é tipicamente a função de lavagens pós



hibridização, os fatores críticos sendo a força iônica e a temperatura da solução de lavagem final. Para híbridos DNA-DNA, a  $T_m$  pode ser aproximada da equação de Meinkoth et al. ((1984) Anal Biochem 138:267-284):  $T_m = 81.5 \text{ }^\circ\text{C} + 16.6 (\log M) + 0.41 (\%GC) - 0.61 (\% \text{ form}) - 500/L$ ; onde M é a molaridade de cátions monovalentes, %GC é a porcentagem dos nucleotídeos guanosina e citosina no DNA, % form é a porcentagem de formamida na solução de hibridização, e L é o comprimento do híbrido em pares de base. A  $T_m$  é a temperatura (sob força iônica e pH definidos) onde 50% de uma seqüência alvo complementar hibridiza com uma sonda pareada perfeitamente.  $T_m$  é reduzida em cerca de  $1^\circ\text{C}$  para cada 1% de mal pareamento; então,  $T_m$ , condições de hibridização e/ou lavagem podem ser ajustadas para hibridizar com seqüências de identidade desejada. Por exemplo, se seqüências com  $\geq 90\%$  de identidade são procuradas, a  $T_m$  pode ser diminuída  $10 \text{ }^\circ\text{C}$ . Geralmente, condições estringentes são selecionadas para serem cerca de  $5 \text{ }^\circ\text{C}$  mais baixas do que o ponto de fusão termal ( $T_m$ ) para a seqüência específica e seu complemento em uma força iônica e pH definidos. No entanto, condições severamente estringentes podem utilizar uma hibridização e/ou lavagem a 1, 2, 3, ou  $4 \text{ }^\circ\text{C}$  mais baixos do que o ponto de fusão termal ( $T_m$ ); condições moderadamente estringentes podem utilizar uma hibridização e/ou lavagem a 6, 7, 8, 9, ou  $10 \text{ }^\circ\text{C}$  mais baixos do que o ponto de fusão termal ( $T_m$ ); condições mais baixas de estringência podem utilizar uma hibridização e/ou lavagem a 11, 12, 13, 14, 15, ou  $20 \text{ }^\circ\text{C}$  mais baixos do que o ponto de fusão termal ( $T_m$ ). Usando a equação, composições

de hibridização e lavagem, e  $T_m$  desejadas, os especialistas na área entenderão que variações na estringência das soluções de hibridização e/ou lavagens são inerentemente descritas. Se o grau desejado de mal pareamento desejado resulta em uma  $T_m$  menor do que 45 °C (solução aquosa) ou 32 °C (solução de formamida) é preferível aumentar a concentração de SSC de tal modo que uma temperatura maior possa ser utilizada. Um guia extensivo para a hibridização de ácidos nucleicos é encontrado em Tijssen, *Laboratory*  
5 *Techniques in Biochemistry and Molecular Biology--Hybridization with Nucleic Acid Probes*, Part I, Chapter 2 "Overview of principles of hybridization and the strategy of nucleic acid probe assays", Elsevier, New York (1993); e  
10 *Current Protocols in Molecular Biology*, Chapter 2, Ausubel et al., Eds., Greene Publishing and Wiley-Interscience, New York (1995). Condições de hibridização e/ou lavagem podem ser aplicadas por pelo menos 10, 30, 60, 90, 120, ou 240 minutos.

20 O termo "identidade de seqüência" ou "identidade" no contexto de seqüências de ácido nucleico ou polipeptídeos referem-se a bases de ácidos nucleicos ou resíduos de aminoácidos em duas seqüências que são as mesmas quando alinhadas para um máximo de correspondência sobre uma  
25 janela de comparação especificada. O termo "porcentagem de identidade de seqüência" refere-se ao valor determinado pela comparação de duas seqüências otimamente alinhadas sobre uma janela de comparação, onde a porção da seqüência de polinucleotídeo ou polipeptídeo na janela de comparação

pode compreender adições ou deleções (i.e., gaps) quando comparada com a seqüência referência (que não compreende adições ou deleções) para um alinhamento ótimo das duas seqüências. A porcentagem é calculada pela determinação do número de posições onde a base de ácido nucléico idêntica ou resíduo de aminoácido idêntico ocorre em ambas as seqüências para produzir o número de posições pareadas, dividindo o número de posições pareadas pelo número total de posições na janela de comparação e multiplicando os resultados por 100 para produzir a porcentagem de identidade de seqüência. Exemplos úteis de porcentagem de identidade de seqüência incluem, mas não estão limitados a, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, ou 95%, ou qualquer porcentagem inteira de 50% a 100%. Essas identidades podem ser determinadas usando qualquer dos programas descritos aqui.

Alinhamentos de seqüência e cálculos de porcentagem de identidade ou similaridade podem ser determinados usando uma variedade de métodos de comparação designados para detectar seqüências homólogas incluindo, mas não limitado ao, programa MegAlign™ do grupo de computação de bioinformática LASERGENE (DNASTAR Inc., Madison, WI). Dentro do contexto deste pedido é entendido que onde o programa de computador de análise de seqüências é utilizado para análise, os resultados das análises serão baseados nos "valores padrões" do programa referenciado, a menos que seja especificado de outra forma. Como usado aqui "valores padrões" significarão qualquer grupo de valores ou

parâmetros que originalmente carregam com o programa de computador quando da primeira inicialização. O termo "método de alinhamento Clustal V" corresponde ao método de alinhamento classificado como Clustal V (descrito por Higgins & Sharp (1989) CABIOS 5:151-153; Higgins et al. (1992) Comput Appl Biosci 8:189-191) e encontrado no programa MegAlign™ do grupo de computação de bioinformática LASERGENE (DNASTAR Inc., Madison, WI). Para múltiplos alinhamentos, os valores padrões correspondem ao Gap Penalty = 10 E Gap Length Penalty = 10. Parâmetros padrões para alinhamentos emparelhados e cálculo da porcentagem de identidade das seqüências de proteína usando o método Clustal são KTUPLE=1, GAP PENALTY=3, WINDOW=5 e DIAGONALS SAVED=5. Para ácidos nucleicos esses parâmetros são KTUPLE=2, GAP PENALTY=5, WINDOW=4 e DIAGONALS SAVED=4. Depois do alinhamento de seqüências usando o programa Clustal V, é possível obter uma "porcentagem de identidade" através da observação da tabela "distâncias de seqüências" no mesmo programa. O termo "método de alinhamento Clustal W" corresponde ao método de alinhamento classificado Clustal W (descrito por Higgins & Sharp (1989) CABIOS 5:151-153; Higgins et al. (1992) Comput Appl Biosci 8:189-191) e encontrado no programa MegAlign™ v6.1 do grupo de computação de bioinformática LASERGENE (DNASTAR Inc., Madison, WI). Parâmetros padrões para alinhamento múltiplo são GAP PENALTY=10, GAP LENGTH PENALTY=0.2, Delay Divergen Seqs(%)=30, DNA Transition Weight=0.5, Protein Weight Matrix=Gonnet Series, DNA Weight Matrix=IUB. Após o alinhamento das seqüências usando o programa Clustal W, é

possível obter uma "porcentagem de identidade" através da observação da tabela "distâncias de seqüência" no mesmo programa. O termo "método de alinhamento BLASTN" é um algoritmo provido pelo Centro Nacional para Informação Biotecnológica (NCBI) para comparar seqüências de nucleotídeo usando parâmetros padrões.

É bem conhecido por um especialista no estado da técnica que muitos níveis de identidade de seqüência são úteis na identificação de polipeptídeos, de outras espécies, onde tais polipeptídeos têm a mesma ou similar função ou atividade. Exemplos úteis de porcentagem de identidades incluem, mas não estão limitados, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, ou 95%, ou qualquer porcentagem inteira de 50% a 100%. De fato, qualquer identidade de aminoácido inteiro de 50% a 100% pode ser útil na descrição da presente invenção, tais como 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% ou 99%.

O termo "gene" refere-se a um fragmento de ácido nucléico que expressa uma proteína específica, incluindo seqüências regulatórias precedentes (seqüências 5' não codificadoras) e seguintes (seqüências 3' não codificadoras) à seqüência codificadora. "Gene nativo" refere-se a um gene como encontrado na natureza com suas próprias seqüências regulatórias. "Gene quimérico" refere-se a qualquer gene

que não é um gene nativo, compreendendo seqüências regulatórias e codificadoras que não são encontradas juntas na natureza. Portanto, um gene quimérico pode compreender seqüências regulatórias e seqüências codificadoras que são derivadas de diferentes origens, ou seqüências regulatórias e seqüências codificadoras derivadas da mesma origem, mas arranjas de uma maneira diferente do que aquela encontrada na natureza. Um gene "exógeno" refere-se a um gene normalmente não encontrado no organismo hospedeiro, mas que é introduzido dentro do organismo hospedeiro através de transferência gênica. Genes exógenos podem compreender genes nativos inseridos dentro de um organismo não nativo, ou genes quiméricos. Um "transgene" é um gene que tem sido introduzido dentro do genoma através do procedimento de transformação.

O termo "genoma" como é aplicado a células vegetais engloba não apenas DNA cromossômico encontrado dentro do núcleo, como DNA de organelas encontrado dentro de componentes subcelulares (pex., mitocôndria, ou plastídeo) da célula.

Um "gene códon-otimizado" ou "gene códon-preferencial" é um gene tendo sua freqüência de uso de códon designada para imitar a freqüência do uso preferencial de códon da célula hospedeira.

Um "alelo" é uma das várias formas alternativas de um gene ocupando um dado locus em um cromossomo. Quando todos

os alelos presentes em um dado locus em um cromossomo são os mesmos, aquela planta é homozigota naquele locus. Se os alelos presentes em um dado locus em um cromossomo diferem, aquela planta é heterozigota naquele locus.

5

O termo "seqüência codificadora" refere-se a uma seqüência de polinucleotídeo que codifica para uma seqüência de aminoácido específica. "Seqüências regulatórias" referem-se a seqüências de nucleotídeo localizadas acima (seqüências 5' não codificadoras), dentro, ou abaixo (seqüências 3' não codificadoras) de uma seqüência codificadora, e que influencia a transcrição, processamento ou estabilidade de RNA, ou tradução da seqüência codificadora associada. Seqüências regulatórias podem incluir, mas não estão limitadas a: promotores, seqüências líderes de tradução, íntrons, seqüências de reconhecimento de poliadenilação, sítios de processamento de RNA, sítios de ligação efetores e estruturas de haste-volta.

20

O termo "promotor" refere-se a uma seqüência de DNA capaz de controlar a expressão de uma seqüência codificadora ou RNA funcional. A seqüência promotora consiste de elementos acima proximais ou mais distais, os elementos tardios frequentemente referenciados como "potenciadores". Portanto, um "potenciador" é uma seqüência de DNA que pode estimular atividade do promotor, e pode ser um elemento inato do promotor ou um elemento heterólogo inserido para potencializar o nível ou tecido-

especificidade de um promotor. Promotores podem ser derivados na sua totalidade de um gene nativo, ou ser compostos de diferentes elementos derivados de diferentes promotores encontrados na natureza, ou ainda compreender 5 segmentos de DNA sintéticos. É entendido pelos especialistas na área que promotores diferentes podem dirigir a expressão de um gene em diferentes tecidos ou tipos celulares, ou em estágios diferentes do desenvolvimento, ou em resposta a diferentes condições ambientais. É ainda reconhecido que desde que na maioria dos casos os limites das seqüências regulatórias não tenham sido completamente definidos, fragmentos de DNA de alguma variação podem ter atividade promotora idêntica. Promotores que possibilitam a um gene ser expresso na maioria dos 15 tipos celulares na maioria das vezes são comumente referenciados como "promotores constitutivos". Novos promotores de vários tipos úteis em células vegetais estão constantemente sendo descobertos; numerosos exemplos podem ser encontrados na compilação de Okamura & Goldberg (1989) 20 *Biochemistry of Plants* 15:1-82.

O termo "seqüência líder de tradução" refere-se a uma seqüência de polinucleotídeo localizada entre a seqüência do promotor de um gene e a seqüência codificadora. A 25 seqüência líder de tradução está presente no mRNA totalmente processado acima da seqüência de início da tradução. A seqüência líder de tradução pode afetar o processamento do transcrito primário para mRNA, estabilidade do mRNA ou eficiência de tradução. Exemplos de



seqüência líderes de tradução têm sido descritas (Turner & Foster (1995) Mol Biotechnol 3:225-236).

Os termos "seqüências 3' não codificadoras",  
5 "terminador de transcrição" ou "seqüências de terminação"  
referem-se a seqüências de DNA localizadas abaixo de uma  
seqüência codificadora e inclui seqüências de  
reconhecimento de poliadenilação e outras seqüências  
10 codificando sinais regulatórios capazes de afetar o  
processamento de mRNA ou expressão gênica. O sinal de  
poliadenilação é usualmente caracterizado por afetar a  
adição de tratos de ácido poliadenílico para a extremidade  
3' do precursor do mRNA. O uso de diferentes seqüências 3'  
não codificadoras é exemplificado por Ingelbrecht *et al.*  
15 (1989) Plant Cell 1:671-680.

O termo "transcrito de RNA" refere-se ao produto  
resultante da transcrição catalisada pela RNA polimerase de  
uma seqüência de DNA. Quando o transcrito de RNA é uma  
20 cópia complementar perfeita da seqüência de DNA, ele é  
referenciado como transcrito primário. Um transcrito de RNA  
é referido como RNA maduro quando ele é uma seqüência de  
RNA derivada de um processamento pós transcricional do  
transcrito primário. "RNA mensageiro" ou "mRNA" refere-se  
25 ao RNA que está sem íntrons e que pode ser traduzido em  
proteína pela célula. "cDNA" refere-se a um DNA que é  
complementar a, e sintetizado de, a mRNA molde usando a  
enzima transcriptase reversa. O cDNA pode ser de fita  
simples ou convertido na forma de fita dupla usando o

fragmento Klenow da DNA polimerase I. RNA "senso" refere-se ao transcrito de RNA que inclui o mRNA e pode ser traduzido em proteína dentro de uma célula ou *in vitro*. "RNA antisenso" refere-se a um transcrito de RNA que é complementar ao todo ou parte de um transcrito primário alvo ou mRNA, e que bloqueia a expressão de um gene alvo (patente americana US 5.107.065). A complementariedade de um RNA antisenso pode ser com qualquer parte do transcrito gênico específico, *i.e.*, na seqüência 5' não codificadora, seqüência 3' não codificadora, íntrons, ou a seqüência codificadora. "RNA funcional" refere-se ao RNA antisenso, RNA ribozima, ou outro RNA que não pode ser traduzido mas até o momento tem um efeito nos processos celulares. Os termos "complemento" e "complemento reverso" são usados de forma substituível aqui com relação aos transcritos de mRNA, e são meios para definir o RNA antisenso da mensagem.

O termo "ligado operacionalmente" refere-se à associação de seqüências de ácido nucléico em um fragmento de ácido nucléico simples tanto que a função de uma é regulada pela outra. Por exemplo, um promotor é operacionalmente ligado com uma seqüência codificadora quando ele é capaz de regular a expressão daquela seqüência codificadora (*i.e.*, a seqüência codificadora está sob controle transcricional do promotor). Seqüências codificadoras podem estar operacionalmente ligadas a seqüências regulatórias na orientação senso ou antisenso. Em outro exemplo, as regiões de RNA complementares da invenção podem estar operacionalmente ligadas, diretamente

ou indiretamente cada, 5' para o mRNA alvo, ou 3' para mRNA alvo, ou dentro do RNAm alvo, ou uma primeira região complementar é 5' e seu complemento é 3' para o RNAm alvo.

5 DNA recombinante padrão e técnicas de clonagem molecular usadas aqui são bem conhecidas no estado da técnica e são descritas mais completamente em Sambrook et al. *Molecular Cloning: A Laboratory Manual*; Cold Spring Harbor Laboratory: Cold Spring Harbor, NY (1989). Métodos de transformação são bem conhecidos no estado da técnica para os especialistas na área e são descritos *infra*.

15 "PCR" ou "reação de polimerase em cadeia" é uma técnica para síntese de segmentos específicos de DNA e consiste de uma série de repetitivos ciclos de desnaturação, anelamento, e extensão. Tipicamente, um DNA de fita dupla é desnaturado por calor, os dois iniciadores complementares às extremidades 3' do segmento alvo são anelados a baixa temperatura, e então estendido em uma temperatura intermediária. Um grupo dessas três etapas consecutivas é referido como um "ciclo".

25 O termo "recombinante" refere-se a uma combinação artificial de dois segmentos separados de outra forma da seqüência, pex., através de síntese química ou através de manipulação de segmentos isolados de ácidos nucleicos através de técnicas de engenharia genética.

Os termos "plasmídeo", "vetor" e "cassete" referem-se

a um elemento extra-cromossômico freqüentemente carregando genes que não são parte do metabolismo central da célula, e usualmente na forma de fragmentos de DNA de fita dupla circular. Tais elementos podem ser seqüências replicando autonomamente, seqüências integrando genomas, seqüências de nucleotídeo ou fago, lineares ou circulares, de um DNA ou RNA simples ou fita dupla, derivadas de qualquer origem, onde um número de seqüências de nucleotídeos tem sido unidas ou recombinadas dentro de uma única construção que é capaz de introduzir um fragmento de promotor e seqüência de DNA para um produto gênico selecionado adiante com seqüência 3' não traduzida apropriada dentro da célula. "Cassete de transformação" refere-se a um vetor específico contendo um gene exógeno e tendo elementos em adição ao gene exógeno que facilitam a transformação de uma célula hospedeira particular. "Cassete de expressão" refere-se a um vetor específico contendo um gene exógeno e tendo elementos em adição ao gene exógeno que permitem a expressão daquele gene em um hospedeiro exógeno.

20

Os termos "construção recombinante", "construção de expressão", "construção quimérica", "construção", e "construção de DNA recombinante" são usadas de forma substituível aqui. Uma construção recombinante compreende uma combinação artificial de fragmentos de ácido nucléico, pex., seqüência regulatórias e codificadoras que não são encontradas juntas na natureza. Por exemplo, uma construção quimérica pode compreender seqüências regulatórias e seqüências codificadoras que são derivadas de diferentes

origens, ou seqüências regulatórias e seqüências codificadoras derivadas da mesma origem, mas arranjadas de uma maneira diferente daquela encontrada na natureza. Tal construção pode ser usada sozinha ou pode ser usada em conjunto com um vetor. Se um vetor é utilizado, então a escolha do vetor é dependente do método que será usado para transformar as células hospedeiras como são bem conhecidos por especialistas na área. Por exemplo, um vetor plasmidial pode ser usado. Um especialista na área é bem informado dos elementos genéticos que devem estar presentes no vetor a fim de transformar com sucesso, selecionar e propagar as células hospedeiras compreendendo qualquer dos fragmentos de ácido nucléico isolados da invenção. O especialista na área também reconhecerá que diferentes eventos de transformação independentes podem resultar em diferentes níveis e padrões de expressão (Jones *et al.* (1985) *EMBO J* 4:2411-2418; De Almeida *et al.* (1989) *Mol Gen Genet* 218:78-86), e então que múltiplos eventos devem ser rastreados a fim de obter linhagens apresentando o nível de expressão e padrão desejado. Tal rastreamento pode ser acompanhado por análises de Southern de DNA, análises de Northern da expressão do mRNA, análises de immunoblotting da expressão de proteína, ou análises fenotípicas, entre outras.

25

O termo "expressão", como usado aqui, refere-se à produção de um produto final funcional (*pex.*, um mRNA ou uma proteína, no precursor ou forma madura).

O termo "introduzido" significa prover um ácido nucléico (*pex.*, construção de expressão) ou proteína dentro de uma célula. Introduzido faz referência à incorporação de um ácido nucléico dentro de uma célula eucariota ou procariota onde o ácido nucléico pode ser incorporado dentro do genoma da célula, e faz referência à provisão transiente de um ácido nucléico ou proteína na célula. Introduzido faz referência a métodos de transformação transientes ou estáveis, bem como cruzamento sexual. Então, "introduzido" no contexto de inserir um fragmento de ácido nucléico (*pex.*, uma construção de DNA recombinante/construção de expressão) em uma célula, significa "transfecção" ou "transformação" ou "transdução" e faz referência à incorporação de um fragmento de ácido nucléico em uma célula eucariota ou procariota onde o fragmento de ácido nucléico pode ser incorporado dentro do genoma da célula (*pex.*, cromossomo, plasmídeo, plastídio ou DNA mitocondrial), convertido em um replicon autônomo, ou expresso transientemente (*pex.*, mRNA transfectado).

20

O termo proteína "madura" refere-se a um polipeptídeo processado pós transcricionalmente (*i.e.*, a partir dos quais qualquer pré ou pró peptídeos presentes no produto primário de tradução tenha sido removido). Proteína "precursora" refere-se a um produto primário de tradução de mRNA (*i.e.*, com pré e pró peptídeos ainda presentes). Pré e pró peptídeos podem ser, mas não estão limitados a, sinais de localização intracelulares.

25

O termo "transformação estável" refere-se a transferência de um fragmento de ácido nucléico para dentro do genoma de um organismo hospedeiro, incluindo ambos os genomas nuclear e de organelas, resultando em uma herança  
5 estável geneticamente. Em contraste, "transformação transiente" refere-se à transferência de um fragmento de ácido nucléico para dentro do núcleo, ou uma organela contendo DNA, de um organismo hospedeiro resultando em expressão gênica sem integração ou herança estável.  
10 Organismos hospedeiros contendo os fragmentos de ácido nucléico transformado são referenciados como organismos "transgênicos".

O termo "transgênico" refere-se a uma planta ou uma  
15 célula, que compreende dentro de seu genoma um polinucleotídeo heterólogo. Preferencialmente, o polinucleotídeo heterólogo é estavelmente integrado dentro do genoma de tal modo que o polinucleotídeo é transferido para gerações sucessivas. O polinucleotídeo heterólogo pode  
20 ser integrado dentro do genoma sozinho ou como parte de uma construção de expressão. Transgênico é usado aqui para incluir qualquer célula, linhagem celular, calo, tecido, parte de planta ou planta, onde o genótipo do mesmo tenha sido alterado pela presença de ácidos nucléicos heterólogos  
25 incluindo aqueles transgênicos inicialmente alterados bem como aqueles criados através de cruzamento sexual ou propagação assexuada transgênica inicial. O termo "transgênico" como usado aqui não engloba a alteração do genoma (cromossômico ou extra cromossômico) através de

métodos convencionais de melhoramento genético de planta ou através de eventos ocorrendo naturalmente tais como fertilização cruzada ao acaso, infecção viral não recombinante, transformação bacteriana não recombinante, 5 transposição não recombinante, ou mutação espontânea.

O termo "planta" refere-se a plantas inteiras, órgãos vegetais, tecidos vegetais, sementes, células vegetais, sementes e progênies das mesmas. Células vegetais incluem, 10 sem limitação, células de sementes, culturas em suspensão, embriões, regiões meristemáticas, tecido de calos, folhas, raízes, brotos, gametófitos, esporófitos, pólen e micrósporos. Partes de planta incluem tecidos diferenciados e não diferenciados incluindo, mas não limitado aos 15 seguintes: raízes, caules, brotos, folhas, pólen, sementes, tecido tumoral e várias formas de células e culturas (pex., células simples, protoplastos, embriões e tecido de calo). O tecido vegetal pode estar na planta ou em um órgão vegetal, tecido ou cultura celular. O termo "órgão vegetal" refere-se 20 se ao tecido vegetal ou a um grupo de tecidos que constituem uma parte de uma planta distinta morfológica e funcionalmente. O termo "genoma" refere-se ao seguinte: (1) o complemento inteiro do material genético (genes e seqüências não codificadoras) que está 25 presente em cada célula de um organismo, ou vírus ou organela; e/ou (2) um grupo completo de cromossomos herdados como uma unidade (haplóide) de um parental. "Progênie" compreende qualquer geração subsequente de uma planta.



A presente invenção diz respeito a um minicromossomo artificial de uma planta compreendendo um centrômero funcional contendo: (a) pelo menos dois arranjos de repetições em tandem de CentC em uma orientação invertida onde o primeiro arranjo compreende pelo menos 50 cópias de CentC e o segundo arranjo compreende pelo menos 50 cópias de CentC; e, (b) pelo menos uma cópia de um elemento retrotransponível, onde o elemento retrotransponível está situado entre o primeiro e o segundo arranjo. Preferencialmente, o elemento retrotransponível é selecionado do grupo consistindo de CentA, CRM1, e CRM2.

O cromossomo artificial compreende um centrômero funcional tendo arranjos de repetições em tandem de CentC. Cada arranjo de repetições CentC pode compreender pelo menos 30, 40, 50, 60, 70, 80, 90, 100, 120, 140, 150, 160, 180, 200, 220, 240, 250, 260, 280, 300, 320, 340, 360, 380, 400, 450, ou 500 cópias de CentC. Além do mais, cada arranjo de repetições em tandem de CentC pode ser interrompido por outro elemento seqüência, incluindo mas não limitado a retrotransposon, que é inserido entre cópias de CentC, ou dentro de um elemento CentC, ou dentro de um retrotransposon, ou qualquer outro elemento de seqüência no arranjo. Retrotransposons incluem, mas não estão limitados a, CentA, CRM1, CRM2.

Na maioria dos eucariotos o centrômero, que é o sítio para formação do cinetócoro e fuso fixado nos cromossomos, é embebido na heterocromatina. Cromossomos de *S. cerevisiae*

carecem de seqüências satélites e têm pequenos centrômeros precisamente localizados que especificam fixação do fuso com ~125 bp de DNA (Blackburn & Szostak (1984) Ann Rev Biochem 53:163-194).

5

No entanto, centrômeros de outras linhagens de fungo incluem arranjos de repetições mais similares para aqueles encontrados em plantas e animais (Fishel et al. (1988) Mol Cell Biol 8:754-763). Em eucariotos desenvolvidos, estudos citológicos e bioquímicos demonstraram uma associação física entre DNAs satélites repetidos tandemicamente, regiões de centrômeros, e proteínas específicas associadas centrômero (Henikoff et al. (2001) Science 293:1098-1102; Yu & Dawe (2000) J Cell Biol 151:131-142).

15

Apesar da ausência dos motivos de seqüência universais, a maioria das repetições satélites centroméricas tem notavelmente comprimento de unidade similar entre organismos, por exemplo, a unidade satélite básica é 171 bp em primatas, 186 bp no peixe *Sparus aurata*, e 155 bp no inseto *Chironomus pallidivittatus* (Henikoff et al. (2001) Science 293:1098-1102).

20

Centrômeros vegetais possuem repetições de comprimento de unidade similares, por exemplo, repetições de 156 bp em milho (Ananiev et al. (1998) Proc Natl Acad Sci USA 95:13073-13078), repetições de 168 bp em arroz (Dong et al. (1998) Proc Natl Acad Sci USA 95:8135-8140), e repetições de 180 bp em *Arabidopsis* (Copenhaver (2003) Chromosome Res

25

11:255-262). Em *Arabidopsis*, centrômeros contém tipicamente 2.8-4 Mb tratos de repetições em tandem 178 bp de seqüências satélites (Hall et al. (2004) *Curr Opin Plant Biol* 7:108-114). Em milho, um centrômero de cromossomo B 5 supranumerário totalmente funcional contém cerca de 500 kb de repetições em tandem, onde deleções parciais reduzem transmissão (Alfenito & Birchler (1993) *Genetics* 135:589-597). Extensões de cromossomo de milho contendo o cromossomo B supranumerário foram hibridizadas com sondas 10 de vários elementos repetitivos incluindo CentC, CRM, e CentA, que localizaram as regiões centroméricas nos cromossomos A. Esses elementos repetitivos, predominantemente encontrados próximos aos centrômeros do cromossomo A, hibridizaram com muitos sítios distintos do 15 centrômero no cromossomo B (Lamb et al. (2005) *Chromosoma* 113:337-349).

Pelo menos dois exemplos divergiram da regra geral da formação do centrômero na base do DNA satélite 20 centromérico. Primeiro, o funcionamento aparentemente normal dos centrômeros estranhos em híbridos somáticos ou em introgressão de linhagens cevada-milho (Ananiev et al. (1998) *Proc Natl Acad Sci USA* 95:13073-13078) é indicativo da conservação da função do centrômero e correspondentes 25 complexos de proteínas. Todas as proteínas centroméricas que auxiliam a função de um centrômero estranho compreendendo DNA satélite centromérico não relacionado são aparentemente providas pelo hospedeiro (Jin et al. (2004) *Plant Cell* 16:571-81). Segundo, neocentrômeros são uma nova

classe de centrômeros baseados em DNA não repetitivo recentemente descritos em humanos e *Drosophila* (Williams (1998) Nat Genet 18:30-37; Choo (1997) Am J Hum Genet 61:1225-1233). Encontrados como derivativos de cromossomos normais resultantes de múltiplos rearranjos cromossômicos, neocentrômeros são formados em regiões aparentemente de DNA eucromáticas livres das repetições tipicamente associadas com a função do centrômero. Cromossomos com neocentrômeros tem estabilidade mitótica ou meiótica variável.

10

A natureza e o funcionamento do centrômero ainda não estão completamente entendidas e requerem análises adicionais. Até o momento, a maioria dos cromossomos artificiais têm centrômeros funcionais baseados dos DNAs satélites centroméricos nativos. É possível que repetições knob, tais como 180 bp e 350 bp (TRI), possam ser usadas como um componente de um neocentrômero. Foi mostrado que alguns knobs poderiam adquirir função de centrômero em cromossomos de milho meióticos, esses neocentrômeros compreendiam 180 bp e 350 bp repetições em tandem apenas. O estudo de neocentrômeros em humanos e organismos menores, têm desvendado um fenômeno insuspeito a respeito da natureza dinâmica do DNA centromérico (Choo et al. (1997) Am J Hum Genet. 61:1225-33). No centro desse fenômeno, não parece haver requerimento de seqüência de DNA específica para função do centrômero; além disso, uma variedade de seqüências que podem responder a uma influência epigenética apropriada parece prover essa função.

25

Extensivas caracterizações de seqüências de centrômero vieram de estudos em leveduras, por exemplo *S. cerevisiae* e *S. pombe*, e têm definido elementos e organização de centrômero em levedura funcionais. Por exemplo, em 5 centrômeros de *S. cerevisiae* a estrutura e função das três regiões essenciais, CDEI, CDEII, e CDEIII, totalizando apenas 125 bp, ou 0.006 - 0.06% de cada cromossomo foram descritas (Carbon et al. (1990) *New Biologist* 2:10-19; Bloom (1993) *Cell* 73:621-624).

10

Centrômeros de *S. pombe* estão entre 40-100 kb e consistem de elementos repetitivos que compreendem 1-3% de cada cromossomo (Baum et al. (1994) *Mol Cell Biol* 5:747-761). Estudos subsequentes demonstraram que menos do que 15 1/3 do centrômero de *S. pombe* nativo é suficiente para a função do centrômero (Baum et al. (1994) *Mol Cell Biol* 5:747-761). Em *S. pombe*, foi mostrado que uma região repetida invertida foi essencial para a função centromérica, mas nem o núcleo central nem o braço da 20 repetição invertida sozinha conferiram função. Deleção de uma porção de seqüências repetidas que flanqueiam o núcleo central não tiveram efeito nas funções de segregação mitótica, ou segregação meiótica de um minicromossomo para a progênie haplóide, mas drasticamente prejudicada a 25 manutenção mediada pelo centrômero de cromátides irmãs de homólogos pareando na meiose I. Existe variabilidade significativa entre cada um dos três cromossomos diferentes em *S. pombe*, e o centrômero de qualquer cromossomo particular pode conter variabilidade significativa através

de diferentes linhagens de *S. pombe*. No entanto, o motivo estrutural de DNA básico, nominalmente, as repetições invertidas, é um parâmetro comum do centrômero de *S. pombe* (Clarke et al. (1993) Cold Spring Harb Symp Quant Biol 58:687-695).

Centrômeros de eucariotos desenvolvidos são menos caracterizados. Fragmentos de DNA que hibridizam com regiões centroméricas em eucariotos desenvolvidos têm sido identificados, entretanto geralmente pouco é conhecido sobre a estrutura, organização, e/ou funcionalidade dessas seqüências. No entanto, arroz é uma exceção por causa de seus tamanhos de centrômeros diferentes. Embora alguns cromossomos de arroz tenham centrômeros similares em tamanho àqueles em outras espécies (>1 Mb), os centrômeros de vários cromossomos são surpreendentemente pequenos e podem estar totalmente cobertos pelos contigs BAC construídos usando técnicas padrões. Seqüenciamento completo do centrômero e arroz 4 e 8 revelaram a presença de blocos invertidos de repetições em tandem centroméricas dentro do segmento cromossômico considerado o centrômero, similar à estrutura repetida invertida observada em leveduras (Zhang et al. (2004) Nucl Acids Res 32:2023-2030; Wu et al. (2004) Plant Cell 16:967-976).

25

Em muitos casos sondas para repetições em centrômero correlacionam com localização no centrômero ambos citologicamente e geneticamente, com muitas dessas seqüências presentes como elementos satélite repetidos em

tandem e seqüências repetidas dispersas em arranjos de variando de 300 - 5000 kb em comprimento (Willard (1990) Trends Genet 6:410-416). Hibridização *in situ* tem mostrado que a repetição satélite alfoíde de 171 bp está presente em  
5 cada centrômero humano (Tyler-Smith *et al.* (1993) Curr Biol 390-397). Se essas repetições constituem centrômeros funcionais não está ainda determinado, e parece que outro DNA genômico é necessário para conferir hereditariedade ao DNA. Transfecção de linhagens celulares com satélites  
10 alfoídes produziu novos cromossomos, no entanto esses novos cromossomos também contém DNA hospedeiro, que poderiam contribuir para atividade do centrômero (Haaf *et al.* (1992) Cell 70:681-696; Willard (1997) Nat Genet 15:345-354). Adicionalmente, os novos cromossomos podem mostrar DNA  
15 alfoíde distribuídos ao longo de seu comprimento total tendo apenas uma constrictão centromérica, indicando que um bloco de DNA alfoíde pode ser insuficiente para conferir função de centrômero.

20 Caracterização genética de centrômeros de plantas tem usado análises de segregação de fragmentos de cromossomo, incluindo análises de linhagens trissômicas carregando um fragmento telocêntrico marcado geneticamente (*pex.*, Koornneef (1983) Genetica 62:33-40). Elementos repetitivos  
25 de centrômero de planta que são geneticamente (Richards *et al.* (1991)) ou fisicamente (Alfenito *et al.* (1993) Genetics 135:589-597; Maluszynska *et al.* (1991) Plant J 1:159-166) ligados a um centrômero têm sido identificados, no entanto a importância dessas seqüências em relação à função do

centrômero não tem sido totalmente caracterizada funcionalmente.

Estudos citológicos em *Arabidopsis thaliana* têm correlacionado estrutura de centrômero com seqüências repetidas. Marcações com um agente de ligação ao DNA fluorescente não específico, tal como 4',6-diamidino-2-fenilindole (DAPI), permite visualização de domínios de cromatina centroméricas em cromossomos na metáfase. Uma hibridização fluorescente *in situ* (FISH) marca para seqüências repetidas de 180 bp pALI co-localizadas com a assinatura DAPI próxima aos centrômeros de todos os cinco cromossomos de *Arabidopsis* (Maluszynska et al. (1991) Plant J 1:159-166; Martinez-Zapater et al. (1986) Mol Gen Genet 204:417-423). Um papel funcional para pALI foi proposto, no entanto estudos mais recentes não têm detectado esta seqüência próxima aos centrômeros de espécies intimamente relacionadas a *Arabidopsis* (Maluszynska et al. (1993) Ann Botany 71:479-484). Uma espécie testada, *A. pumila* acredita-se ser um afidiplóide derivado de um cruzamento de *A. thaliana* com outro parente próximo (Maluszynska et al. (1991) Plant J 1:159-166; Price et al. (1995) in *Arabidopsis*, Somerville & Meyerowitz (eds) Cold Spring Harbor Press, NY). Outras seqüência repetitiva, pAt12, mapeia geneticamente para dentro de 5cM do centrômero do cromossomo 1, e a região central do cromossomo 5 (Richards et al. (1991) Nucl Acids Res 19:3351-3357), mas seu papel na função do centrômero continua a ser estabelecida.



Regiões centroméricas vegetais são compostas predominantemente de repetições centroméricas específicas, retrotransposons centroméricos, e um pouco de outros elementos repetitivos que são mais dispersos ao longo do genoma da planta. Por exemplo, repetições centroméricas tais como CentO e CRR são conhecidas de arroz. Quatro elementos repetitivos de centrômero têm sido descritos em milho: CentA, CentC, CRM1, e CRM2 (SEQ ID NOS: 1-4). Em milho, o primeiro elemento específico de centrômero repetido em tandem descoberto foi CentC (Ananiev *et al.* (1998) Proc Natl Acad Sci USA 95:13073-13078). CentC forma múltiplos arranjos em tandem de comprimento variável, com alguns arranjos em tandem compreendendo até mil cópias de repetição CentC. A repetição em tandem CentC interage com a proteína CENH3 no nucleossoma centromérico.

O elemento específico de centrômero de milho, CentA, parece ser um retrotransposon baseado na sua estrutura e propriedades (Ananiev *et al.* (1998) Proc Natl Acad Sci USA 95:13073-13078; GenBank AF078917). Outro retrotransposon específico de centrômero altamente conservado de milho, CRM2, foi encontrado em 2003 (Nagaki *et al.* (2003) Genetics 163:759-770; GenBank AY129008). Um quarto retrotransposon específico de centrômero, CRM1 (SEQ ID NO: 3), foi identificado através de análises comparativas de seqüências de DNA publicadas de dois clones BAC centroméricos de milho (Nagaki *et al.* (2003) Genetics 163:759-770) e seqüências de DNA genômico proprietárias de milho (Ananiev (2005) não publicado). Alguma homologia pode ser detectada entre os

elementos repetidos centroméricos das espécies proximamente relacionadas, tais como sorgo e cana-de-açúcar (Miller et al. (1998) Genetics 150:1615-1623; Nagaki et al. (1998) Chromosome Res 6:295-302; Zwick et al. (2000) Am J Bot 5 87:1757-1764); e milho e arroz (Ananiev et al. (1998) Proc Natl Acad Sci USA 95:13073-13078); Cheng et al. (2002) Plant Cell 14:1691-1704).

Em adição, centrômeros de planta contém abundantes 10 retrotransposons (CR), em cereais muitos dos elementos CR se inserem dentro de um clado filogenético altamente conservado de elementos Ty3/gypsy (Miller et al. (1998) Theor Appl Genet 96:832-839; Presting et al. (1998) Plant J 16: 721-728; Langdon et al. (2000) Genetics 156:313-325). A 15 homologia de DNA é suficiente para que sondas CR de sorgo ou *Brachypodium sylvaticum* identifiquem centrômeros na maioria ou todos os cromossomos em cereais significantes agronomicamente tais como arroz, milho, trigo, sorgo, cevada e centeio (Aragon-Alcaide et al. (1996) Chromosoma 20 105:261-268; Jiang et al. (1996) Proc Natl Acad Sci USA 93:14210-14213; Miller et al. (1998) Theor Appl Genet 96:832-839).

Retrotransposons, também conhecidos como elementos 25 transponíveis de classe I, consistem de dois subtipos, a repetição terminal longa (LTR) e os retrotransposons não-LTR. Os subtipos de repetição terminal longa tem LTRs diretas que variam de ~100 bp a acima de 5 kb em tamanho. Retrotransposons LTR são ainda classificados dentro dos

grupos Ty1-copia-like (*Pseudoviridae*) e Ty3-gypsy-like (*Metaviridae*) baseados em ambos os graus de similaridade de seqüência e ordem de produtos gênicos codificados. Os grupos Ty1-copia e Ty3-gypsy de retrotransposons são comumente encontrados em alto número de cópias (até poucos milhões de cópias por núcleo haplóide) em plantas com grandes genomas. O retrotransposons Ty1-copia são abundantes em espécies variando de algas de célula simples a briófitas, gimnospermas, e angiospermas. Retrotransposons Ty3-gypsy são também amplamente distribuídos, incluindo ambas gimnospermas e angiospermas. Retrotransposons LTR perfazem aproximadamente 8% do genoma humano. Retrotransposons não-LTR consistem de dois subtipos, elementos nucleares intercalados longos (LINEs) e elementos nucleares intercalados curtos (SINEs). Eles também podem ser encontrados em altos números de cópias (até 250,000) em espécies vegetais. Transposons de planta, incluindo retrotransposons, são revisados por Feschotte *et al.* (2002) *Nat Rev Genet* 3:329-341. Retrotransposons de planta são revisados por Kumar & Bennetzen (1999) *Ann Rev Genet* 33:479-532.

Retrotransposons centroméricos são identificáveis baseados na classificação unificada de elementos transcrição-reversa usada para estudos de filogenia e taxonomia. Retroelementos completos e retrovírus incluem duas ou mais fases abertas de leitura (ORFs) que codificam proteínas simples ou poliproteínas. A ordem dos genes nos elementos varia, mas são classificadas com base no

alinhamento de aminoácidos e resíduos chaves conservados ou domínios dentro da transcriptase reversa (RT), RNase H 15 (RH), integrase (INT) e genes da protease aspártica (PR) e em um domínio tipo zinc-finger da cisteína-histidina conservada (CH). Os retroelementos também compreendem 5 seqüências de repetições terminais longas (LTR) que flanqueiam a região interna do retroelemento. Cada família de retrotransposons tem diferentes, LTRs hibridizando não-cruzados, e componentes dentro de uma família que pode 10 variar (0-50%) nas suas seqüências LTR. No processo de transposição, os dois LTRs são usualmente idênticos no momento da inserção, mas com o passar do tempo substituições podem causar divergência de seqüência. Muitos retroelementos são conhecidos, incluindo retrotransposons 15 específicos de centrômeros (ver, por exemplo, SanMiguel *et al.* (1998) *Nat Genet* 20:43-45; Turcotte *et al.* (2001) *Plant J* 25:169-179; Feng *et al.* (2002) *Nature* 420:316; Nagaki *et al.* (2004) *Nat Genet* 36:138; Nagaki *et al.* (2003) *Genetics* 163:750-770; Wu *et al.* (2004) *Plant Cell* 16:967-976; Hansen & Haslop-Harrison (2004) *Adv Bot Res* 41:165-193). 20

Existe variação significativa entre centrômeros de diferentes cromossomos de milho com relação ao seu tamanho relativo e a composição da repetição. Clusters CentC de 25 milho podem ser tão pequenos quanto cerca de 100 kb, ou mais do que cerca de 2000 kb em diferentes cromossomos, mas comumente na variação de cerca de 200 kb a cerca de 300 kb. Dado o intervalo de tamanho pequeno, é possível que uma porção central inteira da região centromérica de milho

possam ser encontrada dentro de um simples clone BAC. O polimorfismo estrutural observado sugere que um centrômero de milho é composto de blocos funcionais redundantes, cada um dos quais pode ser capaz de suportar função de centrômero. Uma variação significativa (pelo menos 10 vezes) no tamanho de centrômeros como definido pelo comprimento e/ou número de cópia das repetições em tandem centroméricas CentC é observada entre os diferentes cromossomos de milho. Existe também uma variação significativa no tamanho do centrômero entre cromossomos homólogos de diferentes inbreds.

Em outro aspecto, o minicromossomo artificial de planta da invenção pode compreender pelo menos um telômero funcional.

Telômeros são caps de nucleoproteína na extremidade dos cromossomos eucarióticos lineares essenciais para a manutenção da extremidade cromossômico. A síntese de DNA do telômero é feita pela telomerase, uma ribonucleoproteína com atividade transcriptase reversa (McKnight *et al.* (2002) *Plant Mol Biol* 48:331-337). A telomerase adiciona DNA telomérico colocado na extremidade 3' dos cromossomos através da cópia de uma seqüência molde curta dentro de sua subunidade de RNA. Os telômeros da maioria dos organismos consistem de seqüências curtas altamente conservadas repetidas assimetricamente.

Muitas seqüências repetidas teloméricas são

conhecidas, incluindo CCCCAA ( $C_4A_2$ , *Tetrahymena* & *Paramecium*);  $C_4A_4$  (*Oxytricha* & *Euplotes*);  $C_3TA$  (*Trypanosoma*, *Leishmania*, & *Physarum*);  $C_{1-3}A$  (*Saccharomyces*);  $C_{1-8}T$  (*Dictyostelium*); e  $C_3TA_3$  (*Arabidopsis*, humano, rato, *Caenorhabditis*). O número de repetições observado em cromossomos nativos varia amplamente entre os organismos, pex. Alguns ciliados têm cerca de 50 repetições, menos que 350 repetições tem sido observado em *Arabidopsis*, e repetições totalizando cerca de 300-500 bp observadas em *Saccharomyces*.

O comprimento do telômero em plantas, que tipicamente varia de cerca de 2-75 kb, é controlado por fatores genéticos e de desenvolvimento. Regiões teloméricas têm sido isoladas de *Arabidopsis*, e mostram repetições em tandem heterogêneas em tamanho (Richards & Ausubel (1988) Cell 53:127-136). Uma diferença de 25 vezes nos comprimentos de telômeros entre linhagens puras de milho foi encontrada, variando menos do que 2 kb para a linhagem WF9 para cerca de 40 kb para a linhagem CM37 (Burr et al. (1992) Plant Cell 4:953-960). Mais próximo em direção ao centrômero, a repetição telomérica canônica é freqüentemente encontrada misturada com outros elementos repetitivos do genoma vegetal. Em contraste, *Drosophila* usa transposons nas extremidades de seus cromossomos. Os transposons, elementos HeT-A e TART, são encontrados em múltiplas cópias na extremidade de cada cromossomo. Gradual encurtamento dos telômeros pode ser invertido através de transposição de novas repetições de transposons nas

extremidades. Similar à manutenção do telômero pela telomerase, o modelo de transposição em *Drosophila* invoca um mecanismo usando uma transposição intermediária de RNA que é convertida em extremidade de DNA pela transcriptase reversa.

Replicação de DNA é o processo pelo qual células fazem uma cópia completa de sua informação genética antes da divisão celular. Em *E. coli*, viroses de mamíferos, e *S. cerevisiae*, a iniciação de replicação de DNA é controlada pelas conduções de proteínas de iniciação que interagem com seqüências replicadoras de DNA agindo em cis. Para *S. cerevisiae*, replicadores englobam 100-200 bp e incluem os grandes sítios de origem de replicação onde a síntese de DNA inicia. Esses replicadores contêm uma seqüência replicadora autônoma conservada de 11bp (ARS) que liga o complexo de reconhecimento de origem (ORC) à formação nucleada de complexos de pré-replicação (Gilbert (2001) Science 294:96-100).

20

Em eucariotos desenvolvidos a replicação de DNA pode ser iniciada simultaneamente em centenas ou milhares de sítios cromossômicos. Seqüências de origem definidas não são requeridas, muitas origens de replicação potenciais existem consistindo de amplas zonas de sítios de iniciação espaçados próximos, alguns dos quais podem ser usados mais freqüentemente.

No entanto, diversas origens de replicação

eucarióticas específicas são conhecidas tais como a origem de replicação para 18S-26S rDNA que está localizada em um espaçador não transcrito (Ivessa & Zakian (2002) Genes Dev 16:2459-2464). Esta região é capaz de promover amplificação de construções transgênicos (Hemann et al. (1994) DNA Cell Biol 13:437-445). Outra origem específica é encontrada na região abaixo do gene dihidrofolato redutase (DHFR) em células de ovário do hamster chinês (CHO) (Altman & Fanning (2001) Mol Cell Biol 21:1098-1110). Sítios preferenciais de iniciação de replicação foram também encontrados no segmento do cromossomo de *Drosophila* contendo genes córion (Levine & Spradling (1985) Chromosoma 92:136-142).

A maquinaria de replicação de células vegetais e animais é provavelmente capaz de replicar qualquer tipo de DNA introgrado, incluindo construções integradas, episômas, cromossomos inteiros, ou seus fragmentos (Gilbert (2001) Science 294:96-100).

Minicromossomos artificiais são moléculas de DNA lineares ou circulares construídas de elementos de seqüências de DNA agindo em cis responsáveis pela correta replicação e divisão dos cromossomos para as células filha. Os elementos agindo em cis incluem: origens de replicação (ori), os sítios para iniciação de replicação de DNA, também conhecidos como seqüências de replicação autônomas (ARS); centrômeros, os sítios de cinetócoros montados para a correta segregação dos cromossomos replicados na mitose e meiose; e telômeros, estruturas repetidas de DNA



especializadas que estabilizam a extremidade cromossomos lineares e facilitam a completa replicação das extremidades do cromossomo.

- 5           Diversas estratégias para produzir minicromossomos eucarióticos estão disponíveis, incluindo, mas não estando limitado uma auto-montagem *in vivo* de um minicromossomo de elementos componentes pela maquinaria de manutenção cromossômica celular endógena em célula eucariótica, 10 montagem de um minicromossomo eucariótico de elementos componentes em uma célula procariótica, e montagem *in vitro* de um minicromossomo eucariótico de elementos componentes.

Minicromossomos artificiais foram primeiramente 15 construídos em *Saccharomyces cerevisiae* (Murray *et al.* (1986) *Mol Cell Biol* 6:3166-3172; Blackburn & Szostak (1984) *Ann Rev Biochem* 53:163-194). Um plasmídeo circular compreendendo o centrômero de levedura de 125 bp, uma origem de replicação, um marcador de seleção, e um arranjo 20 palindrômico de dois trechos de DNA telomérico foi montado através de técnicas de DNA recombinante convencionais e introduzido dentro de leveduras por transformação de esferoplasto onde ele resultou em uma molécula linear simples. Construções lineares de 50 kb em comprimento 25 contendo um centrômero, uma origem de replicação, e dois telômeros replicados e segregados na mitose com ~99% de exatidão, e retido em culturas em divisão por pelo menos 20 gerações. A geração de YACs indicaram o potencial para montar cromossomos artificiais em outros eucariotos tais

como plantas e animais. Experimentos em YACs indicaram que três seqüências de DNA agindo em cis são necessárias para construir um cromossomo artificial: telômeros; origens(s) de replicação; e um centrômero.

5

Cromossomos artificiais de animais têm sido gerados por duas abordagens diferentes: geração de cromossomos *de novo* de segmentos de DNA clonados; ou através de fragmentação e rearranjos de um cromossomo natural (Brown

10

*et al.* (2000) *Trends Biotechnol* 18:402-403; Cooke (2001) *Cloning Stem Cells* 3:243-249; Lipps *et al.* (2003) *Gene* 304:23-33). A abordagem *de novo*, refere-se à montagem ou abordagem *bottom-up*, que gera cromossomos artificiais através da combinação de componentes clonados essenciais.

15

Co-transfecção de uma mistura de DNA alfoide humano, telômeros, DNA genômico humano, e um marcador de seleção em células HT1080 resultaram na formação de minicromossomos (Harrington *et al.* (1997) *Nat Genet* 15:345-355).

20

Caracterização dos minicromossomos revelaram que eles todos têm estruturas citogenéticas complexas, e foram estavelmente mantidas na ausência de qualquer seleção. Foi concluído que os minicromossomos e seus centrômeros foram formados *de novo* de entradas de DNA através de rearranjos

25

complexos. Subseqüentemente, outros grupos também usaram células HT1080 para introduzir construções de DNA circulares ou lineares contendo DNA alfoide humano e telômeros clonados em YACs, PACs, ou BACs (Compton *et al.* (1999) *Nucleic Acids Res* 27:1762-1765; Grimes *et al.* (2001)

EMBO Rep 2:910-914). Minicromossomos foram observados com diferentes frequências e mostraram diferentes estabilidades mitóticas. Todos os minicromossomos produzidos foram significativamente maiores do que os construções originais, variando de 5 a 10 Mb. Portanto, um cromossomo de mamífero totalmente funcional poderia ser gerado começando com DNA clonado servindo como uma espinha dorsal para a abordagem *de novo*.

10 Fragmentação e rearranjo de cromossomos naturais retendo regiões centroméricas e teloméricas é uma outra estratégia para a produção de minicromossomo. Pequenos fragmentos cromossômicos podem ser isolados através de campo de pulso de gel de eletroforese, readaptado com genes  
15 desejáveis, e reintroduzidos dentro de célula hospedeira. Minicromossomos fragmentados foram observados em células cancerígenas, e outros tipos celulares depois da irradiação, no entanto os fragmentos foram muito grandes para isolamento e não existia nenhuma forma de controlar a  
20 composição gênica.

Uma abordagem para controlar a redução do tamanho do cromossomo foi baseada em telômero associado à fragmentação cromossômica (TACF) ou truncação direcionado à telômero  
25 (TDT) (Heller *et al.* (1996) Proc Natl Acad Sci USA 93:7125-7130; Shen *et al.* (1997) Hum Mol Genet 6:1375-1382). Ela envolve fragmentação sucessiva dos cromossomos hospedeiros humanos específicos em minicromossomos menores usando um vetor alvo englobando um segmento de telômero terminal, um

marcador de seleção, e às vezes uma região de homologia ao cromossomo alvo. Os 'minicromossomos engenheirados' resultantes permanecem autônomos e segregam normalmente. Minicromossomos tão pequenos quanto 0.5 Mb têm sido gerados  
5 contendo DNA alóide como a sequência centromérica funcional em humanos, linhagens híbridas celulares somáticas hamster-humanos, ou células de galinhas.

Recentemente, cromossomos artificiais humanos foram  
10 usados para criar bezerros clonados transcromossômicos produzindo imunoglobulina humana. Um vetor de minicromossomo humano (HAC) construído por Cre/loxP translocações cromossômicas mediados e trunicações cromossômicas telômero-direcionadas em células DT40 de  
15 galinha proficientes em recombinação homóloga foi introduzido para fibroblastos fetais primários bovinos através de transferência cromossômica microcélula mediada (MMCT). Núcleos isolados de fibroblastos fetais com HAC foram transferidos para oócitos maduros enucleados para  
20 produzir bezerros clonados (Kuroiwa et al. (2002) Nat Biotechnol 20:889-894). Uma abordagem *in vivo* para geração de cromossomos artificiais tem sido desenvolvida, baseada na indução de mecanismos de amplificação em larga escala intrínsecos de células de mamíferos. Integração direcionada  
25 de DNA satélite centromérico e o espaçador não transcrito do rDNA em um cromossomo específico resultou em uma amplificação em larga escala de regiões centroméricas. Esses cromossomos amplificados se tornaram instáveis e passaram por rearranjos significantes produzindo

minicromossomos estáveis preferencialmente compostos de DNA satélite (Kereso *et al.* (1996) *Chromosome Res* 4:226-239; Hadlaczky (2001) *Curr Opin Mol Ther* 3:125-132).

5 Um cromossomo artificial contendo múltiplos sítios acceptor de recombinação específico de seqüência foi desenvolvido (plataforma ACE). Seqüências de interesse são providas em um vetor alvo, e a enzima lambda integrase usada para catalisar a recombinação entre a plataforma ACE  
10 e o vetor alvo.

Processos similares têm sido observados em plantas. Fragmentação espontânea de cromossomos nativos em plantas tem sido observada. Minicromossomos foram descobertos em  
15 *Arabidopsis* (Murata *et al.* (2006) *Chromosoma*, publicado online em 11 de abril de 2006), e milho (Brock & Pryor (1996) *Chromosoma* 104:575-584; Kato *et al.* (2005) *Cytogenet Genome Res* 109:156-165). Em alguns casos, minicromossomos foram induzidos por radiação ionizante (Riera-Lizarazu *et*  
20 *al.* (2000) *Genetics* 156:327-339).

Um mapa físico do centrômero de arroz 5 tem sido construído, e poderia ser usado para criar um cromossomo de arroz artificial (Nonomura & Kurata (2001) *Chromosoma*  
25 110:284-291). Uma abordagem similar foi proposta para a construção de um cromossomo artificial para beterraba, *Beta procumbens* (Gindullis *et al.* (2001) *Genome* 44:846-855). Concatemerização de construções transgênicos, ligações, e rearranjos podem ser encontrados em eventos de

transformação de planta. Transformação de planta em geral com construções padrões pode produzir rearranjos complexos, concatemerização, e amplificação da construção (Svitashev & Somers (2001) *Genome* 44:691-697; Svitashev *et al.* (2002) *Plant J* 32:443-445). Co-transformação de plantas com plasmídeos múltiplos pode produzir loci transgênico contendo combinações dos diferentes transgenes (Wu *et al.* (2002) *Transgenic Res* 11:533-541). Similar aos estudos em células animais, montagem *de novo* de minicromossomos artificiais via concatemerização espontânea e ligação de componentes pode ocorrer em células vegetais (ver Figuras 1-10, e 14-15).

A presente invenção diz respeito a um minicromossomo artificial de planta compreendendo um centrômero funcional, onde o centrômero se liga especificamente à proteína C centromérica.

Cinetócoros ligam o DNA centromérico ao aparato das fibras do fuso. Autoanticorpos humanos que ligam especificamente próximos facilitaram clonagem de proteínas associadas a centrômeros (CENPs, Rattner (1991) *Bioassays* 13:51-56). Pelo menos uma dessas proteínas pertence à superfamília cinesina de motores de microtúbulos (Yen (1991) *EMBO J* 10:1245-1254). Proteínas de leveduras que se ligam a centrômeros têm sido identificadas através de estudos genéticos e bioquímicos (Bloom (1993) *Cell* 73:621-624; Lechner *et al.* (1991) *Cell* 64:717-725). CENH3 é uma proteína altamente conservada que substitui a histona H3 em

centrômeros, é considerada para recrutar outras proteínas requeridas para o movimento dos cromossomos. CENH3 está presente inteiramente no ciclo celular e co-localiza com a proteína C centromérica do cinetócoro (CENPC) em células  
5 meióticas.

Anticorpos específicos às proteínas associadas a centrômero podem ser usados para confirmar montagem do centrômero em uma construção de DNA e/ou minicromossomo.

10 Imunolocalização de um CENP, tal como CENH3 e/ou CENPC, para o centrômero de um minicromossomo indica a formação de um centrômero funcional compreendida de elementos de DNA centromérico e as proteínas de ligação associadas. Antisoro para histona centromérica de milho H3 (CENH3, 17kD) foi  
15 feito e testado em cromossomos de milho nativos (Zhong et al. (2002) Plant Cell 14:2825-2836). Imunoprecipitação de cromatina demonstrou que a CentC e CRM2 interagem especificamente com CENH3. Aproximadamente 38 e 33% de CentC e CRM2 foram precipitados no ensaio de  
20 imunoprecipitação da cromatina, confirmando que muito de CENH3 co-localiza com CentC. Um homólogo de milho de CENPC de mamífero foi isolado por Dawe et al. ((1999) Plant Cell 11:1227-1238) e mostrado para ser um componente do cinetócoro em milho. Um peptídeo conservado de 20  
25 aminoácidos do domínio amino terminal foi usado para produzir antisoro específico para CENPC de milho, que foi marcado diretamente e usado para demonstrar que CENPC é especificamente localizado no centrômero de minicromossomos nativos e artificiais em milho (ver, pex., Figuras 3, 4, 8,

e 10)

Os elementos repetidos centroméricos CentA, CentC, CRM1, e CRM2 incluem seqüências que são substancialmente idênticas às seqüências de milho para CentA, CentC, CRM1, e CRM2 das SEQ ID NOs:1-4. Seqüências substancialmente idênticas incluem seqüências que têm uma alta homologia uma com a outra como exemplificado por ter uma significativa porcentagem de identidade de seqüência, e/ou através de hibridizar seletivamente sob condições estringentes para um CentA, um CentC, um CRM1, ou um CRM2 (SEQ ID NOs: 1-4), ou um complemento dos mesmos. Seqüências que hibridizam seletivamente sob condições de hibridização estringentes incluem seqüências que hibridizam à seqüência alvo pelo menos 2 vezes sobre o background e a exclusão substancial de ácidos nucléicos não alvo. Seqüências hibridizando seletivamente têm tipicamente cerca de pelo menos 80, 85, 90, 95, 96, 97, 98, 99, ou 100% de identidade de seqüência à seqüência alvo. Quaisquer condições de hibridização adequadas e tampões conhecidos no estado da técnica podem ser usados, exemplos dos quais têm sido descritos aqui. Identidade de seqüência pode ser usada para comparar a estrutura primária de duas seqüências de polinucleotídeos ou polipeptídeos. Identidade de seqüência mede os resíduos nas duas seqüências que são as mesmas quando alinhadas para a correspondência máxima. Relações de seqüências podem ser analisadas usando algoritmos implementados por computador. A relação de seqüência entre dois ou mais polinucleotídeos, ou dois ou mais polipeptídeos pode ser determinada pela determinação do melhor alinhamento das seqüências, e



pontuar os compartilhamentos e os gaps no alinhamento, que produz a porcentagem de identidade de seqüência e a porcentagem de similaridade de seqüência. Relações de polinucleotídeo podem ser também descritas baseadas em uma

5 comparação dos polipeptídeos que cada um codifica. Muitos programas e algoritmos para comparação e análise de seqüências são conhecidos. A menos que indicado de outra forma, valores de identidade/similaridade de seqüência

10 providos aqui referem-se a valores obtidos usando GAP Versão 10 (GCG, Accelrys, San Diego, CA) usando os seguintes parâmetros: % de identidade e % de similaridade para uma seqüência de nucleotídeo usando grau de ruptura (GAP Weight) de 50 e Length Weight e 3, e a matriz de pontuação nwsgapdna.cmp; % de identidade e % de

15 similaridade para uma seqüência de aminoácido usando GAP Weight de 8 e Length Weight de 2, e a matriz de pontuação BLOSUM62 (Henikoff & Henikoff (1992) Proc Natl Acad Sci USA 89:10915-10919). GAP usa o algoritmo de Needleman & Wunsch (1970) J Mol Biol 48:443-453, para encontrar o alinhamento

20 de duas seqüências completas que maximiza o número de compartilhamentos e minimiza o número de gaps. Substancialmente idêntico inclui seqüências tendo pelo menos 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% ou mais de identidade de seqüência, onde as

25 seqüências são esperadas reter a função nativa baseada na porcentagem de identidade de seqüência total, a similaridade de seqüência, o alinhamento total da seqüência primária, a presença de blocos conservados de resíduos, a presença de elementos conservados e/ou domínios, a presença

de domínios funcionais conservados, a presença de regiões de ligação, a presença de resíduos catalíticos, a(s) estrutura(s) predita secundária(s) e/ou terciária(s), a viabilidade de estruturas tridimensionais conhecidas, e  
5 outros critérios usados por um especialista na área para identificar e predizer um homólogo funcional de qualquer seqüência particular.

Polinucleotídeos variantes incluem polinucleotídeos  
10 tendo pelo menos uma deleção, adição, e/ou substituição em pelo menos uma das extremidades 5', 3', e/ou sítios internos incluindo íntrons ou éxons, quando comparados com o polinucleotídeo nativo. Polinucleotídeos variantes incluem variantes ocorrendo naturalmente bem como  
15 polinucleotídeos derivados sinteticamente, por exemplo, aqueles gerados usando mutagênese sítio dirigida. Variantes conservadas incluem seqüências que mantêm sua função, codificam o mesmo polipeptídeo, ou codificam um polipeptídeo variante com identidade, função e/ou atividade  
20 substancialmente similar com o polinucleotídeo nativo. Variantes podem ser identificadas com técnicas conhecidas, por exemplo, reação de polimerase em cadeia (PCR), e/ou técnicas de hibridização. Geralmente, variantes de um polinucleotídeo particular terá pelo menos cerca de 40%,  
25 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% ou mais identidade de seqüência para aquele polinucleotídeo particular. Polinucleotídeos variantes podem também ser avaliados pela comparação da porcentagem de seqüência entre os

polipeptídeos codificados usando programas e parâmetros de alinhamento padrões. Quando avaliada pela comparação da porcentagem de identidade de seqüência compartilhada por dois polipeptídeos codificados pelos polinucleotídeos, a  
5 porcentagem de identidade de seqüência entre os dois polipeptídeos codificados é tipicamente pelo menos cerca de 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% ou mais de identidade de seqüência.

10

Proteínas variantes incluem proteínas tendo pelo menos uma deleção, adição, e/ou substituição em pelo menos uma das extremidades N-terminal, C-terminal, e/ou um sítio interno, quando comparadas com o polipeptídeo nativo.

15 Proteínas variantes possuem a atividade biológica desejada da proteína. Variantes incluem polipeptídeos ocorrendo naturalmente, bem como aqueles gerados por manipulação humana. Variantes ativas biologicamente de uma proteína tipicamente tem pelo menos cerca de 40%, 45%, 50%, 55%,  
20 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% ou mais de identidade de seqüência com uma seqüência de aminoácido para a proteína nativa como determinado pelos programas de alinhamento de seqüência. Uma variante biologicamente ativa de uma proteína pode  
25 diferir daquela proteína por tão pouco quanto 1-15 resíduos de aminoácidos. Substituições conservadas geralmente referem-se a troca de um aminoácido com outro tendo propriedades similares. Por exemplo, o modelo de Dayhoff et al. (1978) *Atlas of Protein Sequence and Structure* (Natl

Biomed Res Found, Washington, D.C.) provê orientação em substituições de aminoácidos que não são esperadas afetar a atividade biológica da proteína.

- 5 Polinucleotídeos variantes e proteínas englobam seqüências derivadas de procedimentos mutagênicos e/ou recombino-gênicos, tais como mutagênese e/ou embaralhamento de DNA. Métodos para mutagênese e alterações na seqüência de nucleotídeo são conhecidos (ver, pex., Kunkel (1985)
- 10 Proc Natl Acad Sci USA 82:488-492; Kunkel et al. (1987) Methods Enzymol 154:367-382; U.S. Patent 4.873.192; Walker & Gastra, eds. (1983) Techniques in Molecular Biology (MacMillan Publ. Co., NY) e as referências citadas aqui). Por exemplo, uma ou mais seqüências diferentes codificando
- 15 a recombinase podem ser manipuladas para criar e selecionar uma nova proteína recombinase possuindo as propriedades desejadas. Tipicamente, bibliotecas de polinucleotídeos recombinantes são geradas de uma população de seqüências relacionadas e podem ser homologamente recombinadas *in*
- 20 *vitro* ou *in vivo* (ver, pex., Stemmer (1994) Proc Natl Acad Sci USA 91:10747-10751; Stemmer (1994) Nature 370:389-391; Crameri et al. (1997) Nat Biotechnol 15:436-438; Moore et al. (1997) J Mol Biol 272:336-347; Zhang et al. (1997) Proc Natl Acad Sci USA 94:4504-4509; Crameri et al. (1998)
- 25 Nature 391:288-291; e patente americanas US 5.605.793, e US 5.837.458). Geralmente, modificações em um polinucleotídeo codificando um polipeptídeo não deve alterar a fase de leitura, ou criar e/ou alterar a estrutura secundária do DNA ou mRNA. Ver, publicação de pedido de patente EP No.

75,444.

Oligonucleotídeos sobrepostos, denominados overgos, são os primeiros pares que estendem cerca de 40 bp em comprimento e são usualmente constituídos de dois  
5 oligonucleotídeos de 24-bp que têm uma região de sobreposição de 8-bp na extremidade 3'. Essa característica permite ao primeiro par de overgo prover um com o outro e sintetizar suas cadeias complementares com nucleotídeos marcados pelo método de preenchimento de Klenow (McPherson  
10 (1999) *Genome Analysis: A Laboratory Manual*, Vol. 4, pp. 207-213, ed. Birren et al., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY). Uma variedade de nucleotídeos marcados pode ser usada, incluindo mas não limitada a nucleotídeos marcados com radioatividade ou  
15 nucleotídeos marcados com fluorescência. Isso é útil para a geração de sondas por diferentes métodos de hibridização, incluindo mas não limitado a hibridização de colônia, dot blots, Southern blots, e hibridização *in situ*, tal como FISH. A maior vantagem de sondas overgo sobre as sondas  
20 convencionais para hibridização de bibliotecas é que as seqüências para o desenho de overgos podem ser selecionadas, e então seqüências repetidas presentes em um fragmento de DNA de sonda convencional podem ser evitadas; por essa razão, o problema de hibridização cruzada que está  
25 freqüentemente associado com rastreamento de biblioteca de DNA de grandes genomas pode ser minimizado. Por causa dessa vantagem, hibridização overgo combinada com a estratégia de agrupamento de sondas (Cai et al. (1998) *Genomics* 54:387-397; Chang et al. (2001) *Genetics* 159:1231-1242; Tao et al.

(2001) *Genetics* 158:1711-1724; Romanov et al. (2003) *Cytogenet Genome Res* 101:277-281) tem emergido como um método para rastreamento de biblioteca BAC de alta-vazão para identificação de clone e mapeamento genético físico.

5

Em alguns exemplos genes ou polipeptídeos codificados que possam fortalecer ou estimular o crescimento celular são providos com ou dentro das construções de DNA. Genes que fortalecem ou estimulam o crescimento celular incluem

10 genes envolvidos na regulação da transcrição, regulação de gene homeótico, manutenção e proliferação de célula tronco, divisão celular, e/ou diferenciação celular tais como homólogos WUS (Mayer et al. (1998) *Cell* 95:805-815; W001/0023575; US2004/0166563); *aintegumenta* (ANT) (Klucher et al. (1996) *Plant Cell* 8:137-153; Elliott et al. (1996) *Plant Cell* 8:155-168; números de acesso ao GenBank U40256, U41339, Z47554) ; *clavata* (pex., CLV1, CVL2, CLV3) (W003/093450; Clark et al. (1997) *Cell* 89:575-585; Jeong et al. (1999) *Plant Cell* 11:1925-1934; Fletcher et al. (1999) *Science* 283:1911-1914); *Clavata* e genes da região ao redor do embrião (pex., CLE) (Sharma et al. (2003) *Plant Mol Biol* 51:415-425; Hobe et al. (2003) *Dev Genes Evol* 213:371-381; Cock & McCormick (2001) *Plant Physiol* 126:939-942; Casamitjana-Martinez et al. (2003) *Curr Biol* 13:1435-1441);

25 *baby boom* (pex., BNM3, BBM, ODP1, ODP2) (W000/75530; Boutileir et al. (2002) *Plant Cell* 14:1737-1749); *Zwille* (Lynn et al. (1999) *Dev* 126:469-481); *cotilédone coposo* (pex., Lec1, Lec2) (Lotan et al. (1998) *Cell* 93:1195-1205; W000/28058; Stone et al. (2001) *Proc Natl Acad Sci USA*

98:11806-11811; patente americana 6,492,577); Shoot Meristem-less (STM) (Long et al. (1996) Nature 379:66-69); ultrapetala (ULT) (Fletcher (2001) Dev 128:1323-1333); proteína quinase ativadoras mitogênicas (MAPK) (Jonak et al. (2002) Curr Opin Plant Biol 5:415); proteína fosfatase associada à quinase (KAPP) (Williams et al. (1997) Proc Natl Acad Sci USA 94:10467-10472; Trotochaud et al. (1999) Plant Cell 11:393-406); ROP GTPase (Wu et al. (2001) Plant Cell 13:2841-2856; Trotochaud et al. (1999) Plant Cell 11:393-406); fasciata (pex. FAS1, FAS2) (Kaya et al. (2001) Cell 104:131-142); genes do ciclo celular (patente americana US 6,518,487; WO99/61619; WO02/074909), Shepherd (SHD) (Ishiguro et al. (2002) EMBO J. 21:898-908); Poltergeist (Yu et al. (2000) Dev 127:1661-1670; Yu et al. (2003) Curr Biol 13:179-188); Pickle (PKL) (Ogas et al. (1999) Proc Natl Acad Sci USA 96:13839-13844); genes knox (pex., KN1, KNAT1) (Jackson et al. (1994) Dev 120:405-413; Lincoln et al. (1994) Plant Cell 6:1859-1876; Venglat et al. (2002) Proc Natl Acad Sci USA 99:4730-4735); fertilização independente de endosperma (FIE) (Ohad et al. (1999) Plant Cell 11:407-415), e semelhantes. As combinações dos polinucleotídeos incluem múltiplas cópias de qualquer um dos polinucleotídeos de interesse, e as combinações podem ter qualquer combinação de expressão acima-regulada e abaixo-regulada dos polinucleotídeos combinados. As combinações podem ou não podem ser combinadas em uma construção para transformação da célula hospedeira, e portanto pode ser provida seqüencialmente ou simultaneamente. A célula hospedeira pode ser uma célula

tipo selvagem ou mutante, em um estado normal ou aneuplóide.

Sistemas recombinase sítio-específicos podem ser  
5 usados com qualquer sistema de minicromossomo. Ambas  
integrases e recombinases capazes de catalisar ambas as  
reações para frente ("forward") e reversa, são úteis para  
introduzir modificações após as construções de DNA ou  
minicromossomos terem sido estabelecidos na célula vegetal.  
10 Várias modificações intramoleculares, tais como deleção ou  
inversão de seqüências definidas podem ser feitas. Além  
disso, inserções intermoleculares e trocas podem ser  
feitas, incluindo translocações com cromossomos endógenos  
compreendendo sítios de recombinação sítio-específicos  
15 compatíveis. Os sistemas recombinase podem também ser  
usados para estabelecer sítios alvo (sítios "docking")  
dentro do minicromossomo para integração sítio-específica  
tardia dos polinucleotídeos de interesse providos por  
qualquer método, incluindo cruzamento ou liberação direta.

20

Elementos dos sistemas de recombinação, tais como  
recombinases, e sítios de recombinação podem ser usados,  
por exemplo em uma construção de DNA, um sítio alvo, e/ou  
um cassete de transferência. Um sítio alvo compreende um  
25 polinucleotídeo integrado dentro do genoma, o  
polinucleotídeo compreendendo um promotor operacionalmente  
ligado a pelo menos um sítio de recombinação. Um cassete de  
transferência compreende pelo menos um primeiro sítio de  
recombinação operacionalmente ligado a um polinucleotídeo



de interesse e/ou um polinucleotídeo codificando um marcador de seleção, onde o primeiro sítio de recombinação é recombinogênico com um sítio de recombinação no sítio alvo. Uma semente ou planta alvo tem incorporado

5 estavelmente dentro do seu genoma uma construção de DNA que tem sido gerado e/ou manipulado através do uso de um sistema de recombinação. Métodos de recombinação sítio-específico que resultam em vários eventos de integração, alteração, e/ou excisão para gerar a construção de DNA

10 relatado podem ser empregados para gerar uma semente alvo. Ver, pex., WO99/25821, WO99/25854, WO99/25840, WO99/25855, WO99/25853, WO99/23202, WO99/55851, WO01/07572, WO02/08409, e WO03/08045.

15 Uma recombinase é um polipeptídeo que catalisa uma recombinação sítio-específica entre seus sítios de recombinação compatíveis, e inclui seqüências de recombinase ocorrendo naturalmente, variantes, e/ou fragmentos que retenham a atividade. Um sítio de

20 recombinação é uma seqüência de nucleotídeo que é especificamente reconhecida por uma enzima recombinase, e engloba seqüências de sítio de recombinação ocorrendo naturalmente, variantes, e/ou fragmentos que retenham a atividade. Para revisão de recombinases sítio-específicas,

25 ver Sauer (1994) Curr Op Biotech 5:521-527; Sadowski (1993) FASEB 7:760-767; Groth & Calos (2004) J Mol Biol 335:667-678; and Smith & Thorpe (2002) Mol Microbiol 44:299-307. Qualquer sistema de recombinação, ou combinação de sistemas, pode ser usado incluindo mas não limitado a

recombinases e sítios de recombinação das famílias integrase e/ou resolvase, variantes e fragmentos biologicamente ativos dos mesmos, e/ou quaisquer outros ocorrendo naturalmente ou recombinantemente enzima  
5 produzida ou variante da mesma que catalisa recombinação conservada sítio-específica entre sítios de recombinação especificados, e ocorrendo naturalmente ou sítios de recombinação modificados ou variantes dos mesmos que sejam especificamente reconhecidos por uma recombinase para gerar  
10 um evento de recombinação.

Os sítios de recombinação empregados podem ser sítios correspondentes ou sítios desiguais. Sítios de recombinação correspondentes, ou um grupo de sítios de recombinação  
15 correspondentes, são sítios tendo uma seqüência de nucleotídeo idêntica. Um grupo de sítios de recombinação correspondente, na presença da recombinase apropriada, recombinará eficientemente com um outro. Sítios de recombinação desiguais têm uma seqüência diferente,  
20 compreendendo pelo menos um nucleotídeo de diferença quando comparados um com o outro. Os sítios de recombinação dentro de um grupo de sítios de recombinação desiguais podem ser recombinogênicos ou não recombinogênicos com relação um ao outro. Cada sítio de recombinação dentro do grupo de sítios  
25 desiguais é biologicamente ativo e pode recombinar com um sítio idêntico. Sítios recombinogênicos são capazes de recombinar com um outro na presença da recombinase apropriada. Sítios recombinogênicos incluem aqueles sítios onde a eficiência de excisão relativa de recombinação entre

os sítios recombinogênicos está acima do limite detectável sob condições padrão em um ensaio de excisão quando comparado com o controle tipo selvagem, tipicamente, maior do que 2%, 5%, 10%, 20%, 50%, 100%, ou maior. Sítios não recombinaogênicos não recombinarão com um outro na presença da recombinase apropriada, ou recombinação entre os sítios não é detectável. Sítios de recombinação não recombinaogênicos incluem aqueles sítios que recombinam com um outro em uma frequência menor do que o limite detectável sob condições padrões em um ensaio de excisão quando comparados com controle selvagem, tipicamente, menores do que 2%, 1.5%, 1%, 0.75%, 0.5%, 0.25%, 0.1%, 0.075, 0.005%, 0.001%. Quaisquer sítios de recombinação não recombinaogênicos adequados podem ser utilizados, incluindo um sítio FRT ou variante ativo do mesmo, um sítio lox ou variante ativo do mesmo, um sítio att ou variante ativo do mesmo, qualquer combinação dos mesmos, ou qualquer outra combinação de sítios de recombinação não recombinaogênicos. Sítios de recombinação repetidos diretamente em um grupo de sítios de recombinação recombinogênicos são arranjados na mesma orientação, recombinação entre esses sítios resultam na excisão da seqüência de DNA interveniente. Sítios de recombinação invertidos em um grupo de sítios de recombinação recombinogênicos são arranjados na orientação oposta, recombinação entre esses sítios resultam na inversão da seqüência de DNA interveniente.

A família Integrase das recombinases tem mais de cem membros e inclui, por exemplo, FLP, Cre, Dre, Int, e R.

Para outros membros da família Integrase, ver por exemplo, Esposito *et al.* (1997) *Nucleic Acids Res* 25:3605-3614; Nunes-Duby *et al.* (1998) *Nucleic Acids Res* 26:391-406; Abremski *et al.* (1992) *Protein Eng* 5:87-91; Groth & Calos (2004) *J Mol Biol* 335:667-678; and Smith & Thorpe (2002) *Mol Microbiol* 44:299-307. Outros sistemas de recombinação incluem, por exemplo, bacteriófago estreptomiceto phiC31 (Kuhstoss *et al.* (1991) *J Mol Biol* 20:897-908); bacteriófago  $\lambda$  (Landy (1989) *Ann Rev Biochem* 58:913-949, e Landy (1993) *Curr Op Genet Dev* 3:699-707); sistema de recombinação sítio-específico SSV1 de *Sulfolobus shibatae* (Maskhelishvili *et al.* (1993) *Mol Gen Genet* 237:334-342); e um sistema de integração baseado na integrase retroviral (Tanaka *et al.* (1998) *Gene* 17:67-76). Em alguns exemplos, a recombinase é uma que não requer co-fatores ou um substrato super enrolado. Tais recombinases incluem Cre, FLP, phiC31 Int, mutante  $\lambda$  Int, R, SSV1, Dre, ou variantes ativas ou fragmentos dos mesmos. FLP recombinase catalisa uma reação sítio específica entre dois sítios FRT, e está envolvida em amplificar o número de cópias do plasmídeo 2 microns de *S. cerevisiae* durante a replicação de DNA. A proteína FLP tem sido clonada e expressa. Ver, por exemplo, Cox (1993) *Proc Natl Acad Sci USA* 80:4223-4227. A recombinase FLP usada pode ser derivada do gênero *Saccharomyces*. Em alguns exemplos um polinucleotídeo sintetizado usando códons preferidos de planta codificando a recombinase é utilizado. A enzima FLP codificada por uma seqüência de nucleotídeo compreendendo os códons preferidos de milho (FLPm) que catalisa eventos de recombinação sítio específicos é

conhecida (Patente US 5.929.301). Variantes funcionais adicionais e fragmentos de FLP são conhecidas. Ver, por exemplo, Buchholz et al. (1998) Nat Biotechnol 16:617-618, Hartung et al. (1998) J Biol Chem 273:22884-22891, Saxena et al. (1997) Biochim Biophys Acta 1340:187-204, Hartley et al. (1980) Nature 286:860-864, Shaikh & Sadowski (2000) J Mol Biol 302:27-48, Voziyanov et al. (2002) Nucleic Acids Res 30:1656-1663, e Voziyanov et al. (2003) J Mol Biol 326:65-76. A recombinase Cre do bacteriófago P1 catalisa a recombinação sítio específica entre dois sítios lox. Ver, por exemplo, Guo et al. (1997) Nature 389:40-46; Abremski et al. (1984) J Biol Chem 259:1509-1514; Chen et al. (1996) Somat Cell Mol Genet 22:477- 488; Shaikh et al. (1977) J Biol Chem 272:5695- 5702; e, Buchholz et al. (1998) Nat Biotechnol 16:617-618. Sequências do polinucleotídeo Cre podem também ser sintetizadas usando códons preferidos de planta, por exemplo, moCre (ver, pex., WO 99/25840), e outras variantes são conhecidas, ver por exemplo Vergunst et al. (2000) Science 290:979-982, Santoro & Schulz (2002) Proc Natl Acad Sci USA 99:4185-4190, Shaikh & Sadowski (2000) J Mol Biol 302:27-48, Rufer & Sauer (2002) Nucleic Acids Res 30:2764-2771, Wierzbicki et al. (1987) Mol Biol 195:785-794, Petyuk et al. (2004) J Biol Chem 279:37040-37048, Hartung & Kisters-Wolke (1998) J Biol Chem 273:22884-22891, Koresawa et al. (2000) J Biochem (Tokyo) 127:367-372, Patente americana US 6.890.726, e Buchholz & Stewart (2001) Nat Biotechnol 19:1047-1052. Um homólogo Cre tem sido identificado em fagos P1-relacionados, a recombinase isolada do fago D6 é conhecida como Dre que é

uma tirosina recombinase intimamente relacionada ao Cre, mas que reconhece diferentes sítios rox 32 bp (Sauer & McDermott (2004) *Nucleic Acids Res* 32:1-10). A phiC31 integrase e variantes são conhecidas (Kushtoss *et al.* (1991) *J Mol Biol* 222:897-908, WO03/066867, WO05/017170, US2005/0003540, e Scilimenti *et al.* (2001) *Nucleic Acids Res* 29:5044-5051. A  $\lambda$  integrase e co-fatores (Hoess *et al.* (1980) *Proc Natl Acad Sci USA* 77:2482-2486, Blattner *et al.* (1997) *Science* 277:1453-1474), e variantes dos mesmos são conhecidos, incluindo variantes Int cofator-independente (Miller *et al.* (1980) *Cell* 20:721-729, Lange-Gustafson e Nash (1984) *J Biol Chem* 259:12724-12732, Christ *et al.* (1998) *J Mol Biol* 288:825-836, e Lorbach *et al.* (2000) *J Mol Biol* 296:1175-1181), variantes de reconhecimento de sítio att (Dorgai *et al.* (1995) *J Mol Biol* 252:178-188, Yagu *et al.* (1995) *J Mol Biol* 252:163-167, and Dorgai *et al.* (1998) *J Mol Biol* 277:1059-1070), bem como códon de milho otimizado Int, variante, e seqüências co-fatoras (WO03/08045). Outras integrases e variantes são conhecidas, tais como HK022 integrase (Kolot *et al.* (1999) *Mol Biol Rep* 26:207-213) e variantes tais como variantes de reconhecimento de sítio att (Dorgai *et al.* (1995) *J Mol Biol* 252:178-188, Yagu *et al.* (1995) *J Mol Biol* 252:163-167, and Dorgai *et al.* (1998) *J Mol Biol* 277:1059-1070).

25

Sítios de recombinação tipo selvagem, mutante, ou qualquer combinação do tipo selvagem e/ou sítios mutantes podem ser usados. Tais sítios de recombinação incluem, por exemplo, tipo selvagem lox, FRT, e sítios att, e mutante

lox, FRT, e sítios att. Uma análise da atividade de recombinação dos sítios mutantes lox é apresentada em Lee *et al.* (1998) *Gene* 216:55-65. Outros sítios de recombinação e variantes são conhecidos, ver por exemplo, Hoess *et al.* (1982) *Proc Natl Acad Sci USA* 79:3398-3402; Hoess *et al.* (1986) *Nucleic Acids Res* 14:2287-2300; Thomson *et al.* (2003) *Genesis* 36:162-167; Schlake & Bode (1994) *Biochemistry* 33:12746-12751; Siebler & Bode (1997) *Biochemistry* 36:1740-1747; Huang *et al.* (1991) *Nucleic Acids Res* 19:443-448; Sadowski (1995) in *Progress in Nucleic Acid Research and Molecular Biology* Vol. 51, pp. 53-91; Cox (1989) in *Mobile DNA*, Berg & Howe (eds) American Society of Microbiology, Washington D.C., pp. 116-670; Dixon *et al.* (1995) *Mol Microbiol* 18:449-458; Umlauf & Cox (1988) *EMBO J* 7:1845-1852; Buchholz *et al.* (1996) *Nucleic Acids Res* 24:3118-3119; Kilby *et al.* (1993) *Trends Genet* 9:413-421; Rossant & Geagy (1995) *Nat Med* 1:592-594; Bayley *et al.* (1992) *Plant Mol Biol* 18:353-361; Odell *et al.* (1990) *Mol Gen Genet* 223:369-378; Dale & Ow (1991) *Proc Natl Acad Sci USA* 88:10558-10562; Qui *et al.* (1994) *Proc Natl Acad Sci USA* 91:1706-1710; Stuurman *et al.* (1996) *Plant Mol Biol* 32:901-913; Dale *et al.* (1990) *Gene* 91:79-85; Albert *et al.* (1995) *Plant J* 7:649-659, Patente americana US 6.465.254, WO01/23545, WO99/55851, e WO01/11058. Em alguns exemplos, grupos de sítios de recombinação desiguais e correspondentes podem ser usados, por exemplo sítios de diferentes sistemas de recombinação. Assim sendo, qualquer sítio de recombinação adequado ou grupo de sítios de recombinação pode ser usado, incluindo

um sítio FRT, uma variante biologicamente ativa de um sítio FRT, um sítio lox, uma variante biologicamente ativa de um sítio lox, um sítio att, uma variante biologicamente ativa de um sítio att, qualquer combinação dos mesmos, ou  
5 qualquer outra combinação dos sítios de recombinação. Exemplos de sítios FRT incluem, por exemplo, o sítio FRT tipo selvagem mínimo (FRT1), e vários sítios FRT mutantes, incluindo mas não limitado a FRT5, FRT6, e FRT7 (ver patente americana US 6.187.994). Sítios FRT variantes  
10 adicionais são conhecidos, (ver, pex., WO01/23545, e publicação US 2007/0015195, aqui incorporadas por referência). Outros sítios de recombinação que podem ser usados incluem sítios att, tais como aqueles revelados em Landy (1989) Ann Rev Biochem 58:913-949, Landy (1993) Curr  
15 Op Genet Dev 3:699-707, patente americana US 5.888.732, WO01/07572, e Thygarajan et al. (2001) Mol Cell Biol 21:3926-3934. A recombinase sítio-específica usada depende dos sítios de recombinação no sítio alvo e cassete de transferência. Se sítios FRT são utilizados, recombinase  
20 FLP é provida, quando sítios lox são utilizados, recombinase Cre é provida, quando sítios  $\lambda$  att são usados,  $\lambda$  Int é provida, quando sítios phiC31 att são usados, phiC31 Int é provida. Se os sítios de recombinação usados compreendem sítios de diferentes sistemas, por exemplo um  
25 sítio FRT e um sítio lox, ambas as atividades recombinase podem ser providas, cada uma como entidades separadas, ou como uma recombinase quimérica, por exemplo FLP/Cre (ver, pex., WO 99/25840).



Um marcador provê para a identificação e/ou seleção de uma célula, planta, e/ou semente expressando o marcador. Marcadores incluem, pex. marcador rastreável visual, e/ou de seleção. Um marcador de seleção é qualquer marcador, que  
5 quando expresso em um nível suficiente, confere resistência a um agente seletivo. Por exemplo marcadores visuais podem ser usados para identificar células transformadas compreendendo os construções de DNA introduzidos. Em outro exemplo o marcador visual é uma proteína fluorescente. Tais  
10 proteínas fluorescentes incluem mas não estão limitadas a proteína fluorescente amarela (YFP), proteína fluorescente verde (GFP), proteína fluorescente ciano (CFP), e proteína fluorescente vermelha (RFP). Em ainda em outros exemplos, o marcador visual é codificado por um polinucleotídeo tendo  
15 códons preferidos de milho. Em exemplos adicionais, o marcador visual compreende GFPm, AmCyan, ZsYellow, ou DsRed. Ver, Wenck et al. (2003) Plant Cell Rep. 22:244-251.

Marcadores de seleção e seus agentes seletivos  
20 correspondentes incluem, mas não estão limitados a, genes de resistência a herbicida e herbicidas; genes de resistência a antibiótico e antibióticos; e outros genes de resistência química com seus correspondentes agentes químicos. Genes de resistência a drogas bacterianas  
25 incluem, mas não estão limitados a, neomicina fosfotransferase II (nptII) que confere resistência a canamicina, paromicina, neomicina, e G418, e higromicina fosfotransferase (hph) que confere resistência a higromicina B. Ver também, Bowen (1993) Markers for Plant

Gene Transfer, Transgenic Plants, Vol. 1, Engineering and Utilization; Everett et al. (1987) Bio/Technology 5:1201-1204; Bidney et al. (1992) Plant Mol Biol 18:301-313; e W097/05829.

5

Resistência pode também ser conferida para herbicidas de diversos grupos, incluindo inibidores da síntese de aminoácidos, inibidores de fotossíntese, inibidores de lipídeos, reguladores de crescimento, desregulador de  
10 membrana celular, inibidores de pigmento, inibidores de crescimento de plântulas, incluindo mas não limitado a imidazolinonas, sulfonilureas, triazolopirimidinas, glifosato, setoxidim, fenoxaprop, glufosinato, fosfinotricina, triazinas, bromoxinil, e semelhantes. Ver,  
15 por exemplo, Holt (1993) Ann Rev Plant Physiol Plant Mol Biol 44:203-229; and Miki et al. (2004) J Biotechnol 107:193-232. Marcadores de seleção incluem seqüências que conferem resistência a herbicidas, incluindo mas não limitado a, gene bar, que codifica fosfinotricina acetil  
20 transferase (PAT) que confere resistência ao glufosinato (Thompson et al. (1987) EMBO J 6:2519-2523); glifosato oxidoreductase (GOX), glifosato N-acetiltransferase (GAT), e 5-enol piruvilshikimato-3-fosfato sintase (EPSPS) que confere resistência ao glifosato (Barry et al. (1992) in  
25 Biosynthesis and Molecular Regulation of Amino Acids in Plants, B.K. Singh et al. (Eds) pp.139-145; Kishore et al. (1992) Weed Tech 6:626-634; Castle (2004) Science 304:1151-1154; Zhou et al. (1995) Plant Cell Rep 15:159-163; W097/04103; W002/36782; e W003/092360). Outros marcadores

de seleção incluem dihidrofolato redutase (DHFR), que confere resistência ao metotrexato (ver, pex., Dhir et al. (1994) *Improvements of Cereal Quality by Genetic Engineering*, R.J. Henry (ed), Plenum Press, New York; and  
5 Hauptmann et al. (1988) *Plant Physiol* 86:602- 606). Acetohidroxi ácido sintase (AHAS ou ALS) seqüências mutantes levam à resistência a imidazolinonas e/ou sulfonilureas tais como imazetapir e/ou clorsulfuron (ver, pex., Zu et al. (2000) *Nat Biotechnol* 18:555-558; Patentes  
10 americanas US 6,444,875, e US 6,660,910; Sathasivan et al. (1991) *Plant Physiol* 97:1044-1050; Ott et al. (1996) *J Mol Biol* 263:359-368; e Fang et al. (1992) *Plant Mol Biol* 18:1185-1187).

15 Em adição, genes de resistência química incluem ainda triptofano decarboxilase que confere resistência ao 4-metil triptofano (4-mT) (Goodijn et al. (1993) *Plant Mol Biol* 22:907-912); e bromoxinil nitrilase que confere resistência ao bromoxinil. O marcador de seleção pode compreender  
20 cianamida hidratase (Cah), ver, por exemplo, Greiner et al. (1991) *Proc Natl Acad Sci USA* 88:4260-4264; e Weeks et al. (2000) *Crop Sci* 40:1749-1754. A enzima cianamida hidratase converte cianamida em uréia, por meio disso conferindo  
25 resistência à cianamida. Qualquer forma ou derivado de cianamida pode ser usado como a um agente de seleção incluindo, mas não limitado a, cianamida de cálcio (Perlka® (SKW, Trotberg Germany) e cianamida hidrogenada (Dormex® (SKW)). Ver também, Patentes americanas US 6.096.947, e US 6.268.547. Variantes dos polinucleotídeos e/ou

polipeptídeos da hidratase cianamida reterá atividade hidratase cianamida. Uma variante ativa biologicamente da hidratase cianamida reterá a habilidade para converter cianamida em uréia. Métodos para ensaios para tais  
5 atividades incluem análise para resistência de plantas expressando a hidratase cianamida para cianamida. Ensaios adicionais incluem o ensaio colorimétrico da hidratase cianamida (ver, *pex.*, Weeks et al. (2000) Crop Sci 40:1749-1754; e Patente americana US 6.268.547).

10

A presente invenção também engloba um polinucleotídeo isolado compreendendo: (a) pelo menos dois arranjos de repetições em tandem de CentC em uma orientação invertida onde o primeiro arranjo compreende pelo menos dez cópias de  
15 CentC e o segundo arranjo compreende pelo menos dez cópias de CentC; e, (b) pelo menos uma cópia de um elemento retrotransponível, onde o elemento retrotransponível está situado entre o primeiro e o segundo arranjo. Elementos retrotransponíveis adequados são discutidos acima.

20

Também dentro do escopo da invenção está um polinucleotídeo isolado compreendendo: (a) pelo menos um arranjo de repetições em tandem de CentC, o arranjo compreendendo pelo menos dez cópias de CentC; e, (b) pelo  
25 menos uma cópia de um elemento retrotransponível selecionado do grupo consistindo de CentA, CRM1, e CRM2.

Em ainda outro aspecto, a presente invenção engloba um polinucleotídeo isolado compreendendo: (a) pelo menos um

arranjo de repetições em tandem de CentC, o arranjo compreendendo pelo menos dez cópias de CentC; e, (b) pelo menos uma cópia cada de CentA, CRM1, e CRM2.

- 5 Os polinucleotídeos isolados compreendem pelo menos um arranjo de repetições em tandem de CentC. Cada arranjo de repetições CentC pode compreender pelo menos 5, 10, 15, 20, 25, 30, 40, 50, 60, 70, 80, 90, 100, 120, 140, 150, 160, 180, 200, 220, 240, 250, 260, 280, ou 300 cópias de CentC.
- 10 Além disso, cada arranjo de repetições em tandem de CentC pode ser interrompida por outro elemento de seqüência, incluindo mas não limitado a um retrotransposon, que é inserido entre cópias de CentC, ou dentro de um elemento CentC, ou dentro de um retrotransposon, ou qualquer outro
- 15 elemento de seqüência no arranjo. Retrotransposons incluem, mas não estão limitados a, CentA, CRM1, e CRM2.

Um polinucleotídeo inclui qualquer molécula de ácido nucléico, e compreende ribonucleotídeos

20 deoxiribonucleotídeos ocorrendo naturalmente, sintéticos e/ou modificados, e combinações dos ribonucleotídeos e deoxiribonucleotídeos. Polinucleotídeos englobam todas as formas de seqüências incluindo, mas não limitado a, fita simples, fita dupla, linear, circular, ramificada, grampos,

25 estruturas haste-volta, e semelhantes.

Também dentro do escopo da invenção está uma construção recombinante compreendendo qualquer dos polinucleotídeos isolados da invenção.

Uma construção de DNA recombinante compreende um polinucleotídeo que quando presente no genoma de uma planta é heterólogo ou exógeno para aquela localização cromossômica no genoma vegetal. No preparo da construção de DNA, vários fragmentos podem ser manipulados para prover as seqüências em uma orientação própria e/ou na fase de leitura correta. Adaptadores ou ligantes podem ser empregados para unir os fragmentos. Outras manipulações podem ser usadas para prover sítios de restrição convenientes, remoção de DNA supérfluos, ou remoção de sítios de restrição. Por exemplo, mutagênese *in vitro*, reparo de iniciadores, restrição, anelamento, resubstituições, transições, transversões, ou sistemas de recombinação podem ser usados. Polinucleotídeos de interesse referem-se a qualquer molécula de ácido nucléico incluída nos construções de DNA para qualquer propósito, incluindo mas não limitado a regiões não traduzidas, regiões regulatórias, regiões de iniciação de transcrição, regiões de iniciação da tradução, íntrons, éxons, polinucleotídeos codificando um RNA, marcadores de seleção, marcadores rastreáveis, marcadores fenotípicos, polinucleotídeos codificando uma recombinase, sítios de recombinação, sítios alvo, cassetes de transferência, sítios de restrição, sítios de reconhecimento, isoladores, fortalecedores, seqüências espaçadoras, origens de replicação, seqüência telomérica, operadores, e semelhantes, podem ser providos em uma construção (s) de DNA. A construção pode incluir seqüências regulatórias 5' e

3' operacionalmente ligadas às seqüências apropriadas. As construções de DNA podem incluir na direção 5' a 3' de transcrição pelo menos um dos seguintes, uma região de iniciação transcricional e traducional, o polinucleotídeo, e uma região de terminação transcricional e traducional funcional em plantas. Alternativamente, as construções de DNA podem necessitar de pelo menos um elemento regulatório 5' e/ou 3'. Por exemplo, construções de DNA podem ser designados de tal forma que na introdução dentro de uma célula e na presença da recombinase apropriada um evento de recombinação no sítio alvo operacionalmente ligada as regiões regulatórias 5' e/ou 3' com as seqüências apropriadas das construções de DNA.

Elementos regulatórios podem ser usados em uma variedade de maneiras dependendo do elemento polinucleotídeo, sítio de recombinação, cassete de transferência e/ou sítio alvo empregado. Em alguns exemplos seqüências intervenientes podem estar presentes entre os elementos operacionalmente ligados e não desorganizam a ligação funcional. Por exemplo, uma ligação operacional entre um promotor e um polinucleotídeo de interesse permite ao promotor iniciar e mediar a transcrição do polinucleotídeo de interesse. Em alguns exemplos um sítio de iniciação transcricional é operacionalmente ligado a um sítio de recombinação. Em alguns exemplos, um sítio de recombinação está dentro de um íntron.

Um cassete pode adicionalmente compreender pelo menos

uma a seqüência adicional para ser introduzida dentro da planta. Alternativamente, seqüências adicionais podem ser providas separadamente. Construções de DNA podem ser providos com uma pluraridade de sítios de restrição ou  
5 sítios de recombinação para manipulação dos vários componentes e elementos. Construções de DNA podem adicionalmente conter genes marcadores de seleção.

Uma região de iniciação transcricional pode ser  
10 nativa, análoga, exógena, ou heteróloga à planta hospedeira ou ao polinucleotídeo de interesse, e pode ser uma seqüência natural, uma seqüência modificada, ou uma seqüência sintética. Um número de promotores pode ser usado para expressar uma seqüência codificadora.

15

Uma variedade de promotores úteis em plantas é revisto em Potenza *et al.* (2004) *In Vitro Cell Dev Biol Plant* 40:1-22. Em alguns exemplos, o promotor expressando o marcador de seleção é ativo na semente. Promotores ativos na semente  
20 incluem promotores constitutivos, por exemplo, a parte principal do promotor Rsyn7 e outros promotores constitutivos revelados em WO99/43838 e a patente americana US 6.072.050; a parte principal do promotor CaMV 35S (Odell  
25 *et al.* (1985) *Nature* 313:810-812); o promotor MVV (mirabilis mosaic virus) (Dey & Maiti (1999) *Plant Mol Biol* 40:771-782); actina de arroz (McElroy *et al.* (1990) *Plant Cell* 2:163-171); ubiquitina (Christensen *et al.* (1989) *Plant Mol Biol*.12:619-632, and Christensen *et al.* (1992) *Plant Mol Biol* 18:675-689); pEMU (Last *et al.* (1991) *Theor*



Appl Genet 81:581-588); MAS (Velten et al. (1984) EMBO J 3:2723-2730); promotor ALS (patente americana US 5.659.026), e semelhantes. Outros promotores constitutivos incluem aqueles revelados em, pex., patentes americanas US  
5 5.608.149; 5.608.144; 5.604.121; 5.569.597; 5.466.785; 5.399.680; 5.268.463; 5.608.142; e 6.177.611.

O promotor pode ser um promotor tecido-específico, para direcionar a expressão reforçada dentro de um tecido  
10 particular vegetal. Em alguns exemplos, um promotor específico de semente é usado para expressar o marcador de seleção. Promotores específicos de semente incluem ambos os promotores específicos de semente, ativos durante o desenvolvimento da semente, bem como os promotores de  
15 germinação da semente, ativos durante a germinação da semente. Ver Thompson et al. (1989) BioEssays 10:108. Promotores específicos da semente incluem, mas não estão limitados a, Cim1 (mensagem induzida por citocinina); cZ19B1 (zeína 19 kDa de milho); milps (mio-inositol-1-  
20 fosfato sintase) (ver WO00/11177, e patente americana US 6.225.529),  $\beta$ -faseolina de feijão, napina,  $\beta$ -conglucina, lectina de soja, cruciferina, zeína 15 kDa de milho, zeína 22 kDa, zeína 27 kDa, cera, shrunken 1, shrunken 2, globulina 1, end1, e end2 (WO00/12733), e semelhantes.

25

Um promotor químico-regulado pode ser usado para modular a expressão na semente através da aplicação de um regulador químico exógeno. O promotor pode ser um promotor químico-induzível, onde a aplicação de químico induz a

expressão gênica, ou um promotor químico-repressível, onde a aplicação do químico reprime a expressão gênica. Promotores químico induzíveis incluem, mas não estão limitados a, promotor In2-2 de milho, ativado pelo herbicida protetor benzenosulfonamida; o promotor GST de milho, ativado por compostos hidrofóbicos eletrofílicos (pex., alguns herbicidas pré-emergentes); e o promotor PR-la de tabaco, ativado pelo ácido salicílico. Outros promotores químicos regulados de interesse incluem promotores esteróides-responsivos (ver, por exemplo, o promotor glucocorticóide-induzível em Schena et al. (1991) Proc Natl Acad Sci USA 88:10421-10425 and McNellis et al. (1998) Plant J 14:247-257) e promotores tetraciclina-induzível e tetraciclina-repressível (ver, pex., Gatz et al. (1991) Mol Gen Genet 227:229-237, patentes americanas US 5.814.618, e US 5.789.156).

As construções de DNA podem compreender unidades de expressão. Unidades de expressão podem ter elementos incluindo, mas não limitado a, íntrons, fortalecedores, isoladores líderes, espaçadores, regiões codificando um RNA, genes marcadores, sítios de recombinação, regiões de terminação, seqüências codificando recombinases, fortalecedores, ligadores, sítios de reconhecimento, etc. Em adição, as construções de DNA podem compreender cassetes de transferência, sítios alvo, ou quaisquer porções ou combinações das mesmas. Os construções de DNA podem ser modificados em uma variedade de formas incluindo mas não limitado a métodos de integração/recombinação sítio

específico ou transposições baseadas em transposon, para prover um número de variações nos construções de DNA. Seqüências de polinucleotídeos podem ser modificadas para expressão na planta. Ver, *pex.*, Campbell & Gowri (1990) Plant Physiol 92:1-11. Métodos para sintetizar genes preferidos de planta incluem, *pex.*, Patentes americanas US 5.380.831, 5.436.391, e Murray et al. (1989) Nucleic Acids Res 17:477-498.

10           Modificações de seqüências adicionais são conhecidas para fortalecer expressão gênica em um hospedeiro celular. Isso inclui a eliminação de seqüências codificando sinais de poliadenilação espúrios, sinais de sítios de união éxon-intron, repetições como transposons, e outras seqüências bem caracterizadas que possam ser deletérias à expressão

15           gênica. O conteúdo G-C da seqüência pode ser ajustado para níveis médios para um dado hospedeiro, como calculado por referência a genes endógenos expressos no hospedeiro. A seqüência pode também ser modificada para evitar estruturas

20           de RNAm secundárias. Cassetes podem adicionalmente conter seqüências líderes 5' no cassete de DNA que podem agir para fortalecer a tradução. Líderes de tradução incluem, *pex.*, líderes pimaizeavirus tais como líder EMCV (Elroy-Stein et al. (1989) Proc Natl Acad Sci USA 86:6126-6130); líderes

25           potivirus tais como líder TEV (Gallie et al. (1995) Gene 165:233-238), líder MDMV (Kong et al. (1988) Arch Virol 143:1791-1799), e a proteína de ligação à cadeia pesada da imunoglobulina humana (BiP) (Macejak et al. (1991) Nature 353:9094); líder não traduzida do RNAm da proteína da capa

do vírus do mosaico de alfafa (AMV RNA 4) (Jobling *et al.* (1987) *Nature* 325:622-625); líder do vírus do mosaico de tabaco (TMV) (Gallie *et al.* (1989) in *Molecular Biology of RNA*, ed. Cech (Liss, New York), pp. 237-256); e líder do  
5 vírus mosqueado clorótico de milho (MCMV) (Lommel *et al.* (1991) *Virology* 81:382-385). Ver também, Della-Cioppa *et al.* (1987) *Plant Physiol* 84:965-968. Outros métodos ou seqüências conhecidas para fortalecer a tradução pode também ser utilizada, tais como íntrons, e semelhantes.

10

Seqüências de interesse incluem, pex., zinc fingers, quinases, proteínas de choque térmico, fatores de transcrição, reparo de DNA, tratos agronômicos, resistência a insetos, resistência à doença, resistência a herbicida,  
15 esterilidade, óleo, proteína, amido, digestibilidade, tamanho do grão de milho, maturidade, composição de nutriente, níveis ou metabolismo, e semelhantes. Genes de resistência a insetos podem codificar resistência a pragas tais como besouros, moscas, European Maize Borer, e  
20 semelhante. Tais genes incluem, pex., genes de proteínas tóxicas de *B. thuringiensis* (patentes americanas US 5.366.892; 5.747.450; 5.736.514; 5.723.756; 5.593.881; Geiser *et al.* (1986) *Gene* 48:109) e semelhantes. Tratos de resistência à doenças incluem genes de detoxificação, tais  
25 como contra fumonosina (patente americana US 5.792.931); genes de avirulência (avr) e de resistência a doenças (R) (Jones *et al.* (1994) *Science* 266:789; Martin *et al.* (1993) *Science* 262:1432; Mindrinos *et al.* (1994) *Cell* 78:1089); e semelhantes. Tratos de resistência a herbicida incluem

genes codificando para resistência a herbicidas incluindo herbicidas tipo sulfoniluréia (pex., mutações S4 e/ou Hra em ALS), herbicidas que agem para inibir ação da glutamina sintase, tais como fosfinotricina ou basta (pex., o gene bar), EPSPS (patentes americanas US 6.,867.293; 5.188.642; 5 e 5.627.061), GOX (Zhou et al. (1995) Plant Cell Rep 15:159-163), e GAT (patente americana US 6.395.485). Genes de resistência a antibióticos podem também ser usados, tais como o gene nptII que codifica resistência aos antibióticos canamicina e geneticina. Genes de esterilidade podem também 10 ser usados, por exemplo como uma alternativa para a retirada da borla, incluindo genes específicos do tecido masculino e genes com fenótipos de esterilidade masculina tais como QM (pex., patente americana US 5.583.210), 15 quinases, e aqueles codificando compostos tóxicos para qualquer desenvolvimento gametofítico feminino ou masculino.

Redução da atividade de genes específicos, 20 silenciamento e/ou supressão pode ser desejada. Muitas técnicas para silenciamento gênico são conhecidas, incluindo mas não limitado a tecnologia antisense (ver, pex., Sheehy et al. (1988) Proc Natl Acad Sci USA 85:8805-8809; e patentes americanas US 5.107.065; 5.453.566; e 25 5.759.829); co-supressão (pex., Taylor (1997) Plant Cell 9:1245; Jorgensen (1990) Trends Biotech 8:340-344; Flavell (1994) Proc Natl Acad Sci USA 91:3490-3496; Finnegan et al. (1994) Bio/Technology 12:883-888; e Neuhuber et al. (1994) Mol Gen Genet 244:230-241); RNA interference (Napoli et al.

(1990) *Plant Cell* 2:279-289; patente americana US 5.034.323; Sharp (1999) *Genes Dev* 13:139-141; Zamore et al. (2000) *Cell* 101:25-33; Javier (2003) *Nature* 425:257-263; e, Montgomery et al. (1998) *Proc Natl Acad Sci USA* 95:15502-5  
5 15507), silenciamento gênico induzido por vírus (Burton et al. (2000) *Plant Cell* 12:691-705; e Baulcombe (1999) *Curr Op Plant Bio* 2:109-113); ribozimas específicas de RNA alvo (Haseloff et al. (1988) *Nature* 334: 585-591); estruturas em grampo (Smith et al. (2000) *Nature* 407:319-320; WO99/53050;  
10 WO02/00904; e WO98/53083); ribozimas (Steinecke et al. (1992) *EMBO J* 11:1525; patente americana US 4.987.071; e, Perriman et al. (1993) *Antisense Res Dev* 3:253); modificação direcionada mediada por oligonucleotídeo (pex, WO03/076574: e WO99/25853); moléculas direcionadas Zn-  
15 finger (pex., WO01/52620; WO03/048345; e WO00/42219); e outros métodos, ou combinações dos métodos acima.

A região de terminação pode ser nativa com a região de iniciação transcricional, pode ser nativa com a seqüência  
20 de DNA de interesse operacionalmente ligada, ou pode ser derivada de outra origem. Regiões de terminação convenientes são acessíveis do plasmídeo Ti de *A. tumefaciens*, tais como as regiões de terminação da octopina sintase e nopalina sintase. Ver também Guerineau et al.  
25 (1991) *Mol Gen Genet* 262:141-144; Proudfoot (1991) *Cell* 64:671-674; Sanfacon et al. (1991) *Genes Dev* 5:141-149; Mogen et al. (1990) *Plant Cell* 2:1261-1272; Munroe et al. (1990) *Gene* 91:151-158; Ballas et al. (1989) *Nucleic Acids Res* 17:7891-7903; e Joshi et al. (1987) *Nucleic Acids Res*

15:9627-9639.

Em ainda outro aspecto, a presente invenção diz respeito a um método para fazer uma planta transgênica de milho compreendendo um minicromossomo de planta artificial tendo um centrômero funcional, o método compreende:

(a) fazer contato de pelo menos uma célula vegetal de milho com uma mistura compreendendo uma construção recombinante da invenção;

(b) identificar pelo menos uma célula vegetal de milho da etapa (a) compreendendo um minicromossomo de planta artificial tendo um centrômero funcional; e

(c) regenerar uma planta de milho fértil da célula vegetal de milho da etapa (b) onde a dita planta de milho compreende um minicromossomo artificial de planta tendo um centrômero funcional.

A mistura pode ainda compreender um polinucleotídeo codificando um polipeptídeo que estimula o crescimento celular. Exemplos de polipeptídeos que estimulam crescimento celular incluem, mas não estão limitados a, um wuschel, um baby boom, um RepA, ou um Lecl.

Qualquer método para introduzir uma seqüência dentro de uma planta pode ser usado, enquanto o polinucleotídeo ou polipeptídeo ganha acesso ao interior de pelo menos uma célula. Métodos para introduzir seqüências dentro de plantas são conhecidos e incluem, mas não estão limitados

a, transformação estável, transformação transiente, métodos mediados por vírus, e reprodução sexual. Incorporado estavelmente indica que o polinucleotídeo introduzido está integrado dentro de um genoma e é capaz de ser herdado pela progênie. Transformação transiente indica que uma seqüência introduzida não integra dentro do genoma tal que ela é herdável pela progênie do hospedeiro. As plantas e sementes empregadas podem ter uma construção de DNA estavelmente incorporado dentro do seu genoma. Qualquer protocolo pode ser usado para introduzir a construção de DNA, qualquer componente de sistemas de recombinação sítio específico, um polipeptídeo, ou qualquer outro polinucleotídeo de interesse. Prover compreende qualquer método que reunir qualquer polipeptídeo e/ou polinucleotídeo com quaisquer outros componentes relatados. Qualquer meio pode ser usado para reunir um sítio alvo, cassete de transferência, e recombinase apropriada, incluindo, por exemplo, transformação estável, liberação transiente, e cruzamento sexual (ver, *pex.*, WO99/25884). Em alguns exemplos, a recombinase pode ser provida na forma do polipeptídeo ou mRNA. Uma série de protocolos pode ser usada com a finalidade de reunir os vários componentes. Por exemplo, uma célula pode ser provida com pelo menos um desses componentes via uma variedade de métodos incluindo métodos de transformação transiente e estável; co-introdução de uma DNA recombinase, mRNA ou proteína diretamente dentro da célula; empregando um organismo (*pex.*, uma cepa ou linhagem) que expressa a recombinase; ou crescer/cultivar a célula ou organismo carregando um sítio alvo, cruzar com um



organismo expressando uma proteína recombinase ativa, e selecionar eventos na progênie. Um simples padrão de integração é produzido quando o cassete de transferência integra predominantemente no sítio alvo. Qualquer promotor, incluindo promotor constitutivo, induzível, regulado desenvolvimentalmente, temporal, e/ou espacialmente, etc., que seja capaz de regular expressão no organismo possa ser usado.

10           Protocolos de transformação bem como protocolos para introdução de seqüências de polipeptídeos ou polinucleotídeos dentro de plantas podem variar dependendo do tipo de planta ou célula vegetal direcionada para transformação. Métodos adequados de introduzir polipeptídeos e polinucleotídeos dentro das células vegetais incluem microinjeção (Crossway et al. (1986) Biotechniques 4:320-334, patente americana US 6.300.543; e pedidos americanos 11/427.947 e 11/427.371 todos os quais são aqui incorporadas por referência), eletroporação (Riggs et al. (1986) Proc Natl Acad Sci USA 83:5602-5606, transformação mediada por *Agrobacterium* (patentes americanas US 5.563.055; e US 5.981.840), transferência gênica direta (Paszkowski et al. (1984) EMBO J 3:2717-2722), e aceleração de partícula balística (patentes americanas US 4.945.050; US 5.879.918; US 5.886.244; e US 5.932.782; Tomes et al. (1995) in Plant Cell, Tissue, and Organ Culture: Fundamental Methods, ed. Gamborg & Phillips (Springer-Verlag, Berlin); McCabe et al. (1988) Biotechnology 6:923-926); e transformação de Lecl

(W000/28058). Também ver Weissinger *et al.* (1988) *Ann Rev Genet* 22:421-477; Sanford *et al.* (1987) *Particulate Science and Technology* 5:27-37 (cebola); Christou *et al.* (1988) *Plant Physiol* 87:671-674 (soja); Finer & McMullen (1991) In  
5 *Vitro Cell Dev Biol* 27P:175-182 (soja); Singh *et al.* (1998) *Theor Appl Genet* 96:319-324 (soja); Datta *et al.* (1990) *Biotechnology* 8:736-740 (arroz); Klein *et al.* (1988) *Proc Natl Acad Sci USA* 85:4305-4309 (milho); Klein *et al.* (1988) *Biotechnology* 6:559-563 (milho); Patentes americanas US  
10 5.240.855; US 5.322.783; e, US 5.324.646; Klein *et al.* (1988) *Plant Physiol* 91:440- 444 (milho); Fromm *et al.* (1990) *Biotechnology* 8:833-839 (milho); Hooykaas-Van Slogteren *et al.* (1984) *Nature* 311:763-764; Patente americana US 5.736.369 (cereais); Bytebier *et al.* (1987)  
15 *Proc Natl Acad Sci USA* 84:5345-5349 (Liliaceae); De Wet *et al.* (1985) in *The Experimental Manipulation of Ovule Tissues*, ed. Chapman *et al.* (Longman, New York), pp. 197-209 (pólen); Kaeppler *et al.* (1990) *Plant Cell Rep* 9:415-418; e Kaeppler *et al.* (1992) *Theor Appl Genet* 84:560-566  
20 (transformação mediada por fibra); D'Halluin *et al.* (1992) *Plant Cell* 4:1495-1505 (eletroporação); Li *et al.* (1993) *Plant Cell Rep* 12:250-255; Christou & Ford (1995) *Ann Bot* 75:407-413 (arroz); Osjoda *et al.* (1996) *Nat Biotechnol* 14:745-750 (milho via *A. tumefaciens*); and Ch. 8, pp. 189-  
25 253 in *Advances in Cellular and Molecular Biology of Plants*, Vol. 5, Ed. Vasil, Kluwer Acad Publ (Dordrecht, The Netherlands) 1999.

Vários compostos podem ser usados na união com

quaisquer métodos de liberação direta para introduzir dentro de células vegetais qualquer polinucleotídeo, polipeptídeo, ou combinação dos mesmos, opcionalmente contendo outros componentes. Por exemplo, microprojéteis para um método de arma de partículas podem ser preparados através da associação de construções de DNA com microprojéteis na presença de uma solução lipídica catiônica, solução lipossômica, polímero catiônico, proteína de ligação ao DNA, proteína catiônica, peptídeo catiônico, ácido poliamino catiônico, ou combinação dos mesmos. Em alguns exemplos, microprojéteis para um método de arma de partícula são preparados através da associação de construções de DNA com os microprojéteis na presença de Tfx-10, Tfx-20, Tfx-50, Lipofectina, Lipofectamina, Celfectina, Efecteno, Citofectina GSV, Lipídeos Perfect, DOTAP, DMRIE-C, FuGENE-6, Superfect, Polyfect, polietileneimina, quitosana, protamina Cl, proteínas de ligação ao DNA, histona H1, histona CENH3, pola-L lisina, DMSA, e semelhantes.

20

O polinucleotídeo pode ser introduzido em plantas através do contato de plantas com um vírus, ou ácidos nucléicos virais. Geralmente, tais métodos envolvem incorporação de um polinucleotídeo desejado dentro de uma molécula de DNA viral ou RNA viral. A seqüência pode inicialmente ser sintetizada em uma poliproteína viral e depois processada *in vivo* ou *in vitro* para produzir uma proteína desejada. Promotores úteis englobam promotores utilizados para transcrição através de RNA polimerases

virais. Métodos para introduzir polinucleotídeos dentro de plantas e expressar em plantas uma proteína codificada, envolvendo moléculas de DNA ou RNA viral, são conhecidas, ver, *pex.*, Patentes americanas US 5.889.191; US 5.889.190; 5 US 5.866.785; US 5.589.367; US 5.316.931; e Porta *et al.* (1996) *Mol Biotech* 5:209-221.

Vários componentes, incluindo aqueles de um sistema de recombinação sítio-específico, podem ser providos para uma 10 planta usando uma variedade de métodos transientes. Tais métodos de transformação transiente incluem, mas não estão limitados a, introdução da recombinase ou fragmento ativo ou variante dos mesmos diretamente, introdução do mRNA da recombinase, ou usando um método não integrativo, ou 15 introduzir baixos níveis de DNA dentro da planta. Tais métodos incluem, por exemplo, microinjeção, bombardeamento de partículas, sistemas de vetores virais, e/ou precipitação do polinucleotídeo onde a transcrição ocorre do DNA ligado à partícula sem substantiva liberação da 20 partícula ou integração dentro do genoma, tais métodos geralmente usam partículas cobertas com polietilimina, (ver, *pex.*, Crossway *et al.* (1986) *Mol Gen Genet* 202:179-185; Nomura *et al.* (1986) *Plant Sci* 44:53-58; Hepler *et al.* (1994) *Proc Natl Acad Sci USA* 91:2176-2180; e Hush *et al.* 25 (1994) *J Cell Sci* 107:775-784).

As células transformadas podem ser regeneradas em plantas usando meios e protocolos padrões, ver *pex.*, McCormick *et al.* (1986) *Plant Cell Rep* 5:81-84. Essas

plantas podem então ser crescidas e auto-polinizadas, retrocruzadas, e/ou cruzadas, e a progênie resultante tendo a característica desejada identificada. Duas ou mais gerações podem ser crescidas para assegurar que a característica é estavelmente mantida e herdada e então as sementes colhidas. Desta forma sementes transformadas/transgênicas tendo a construção de DNA relatado estavelmente incorporado dentro do seu genoma são providas. Uma planta e/ou uma semente tendo estavelmente incorporado a construção de DNA pode ser ainda caracterizada para expressão, potencial de integração sítio-específica, agronômica, e número de cópias (ver, pex., patente americana US 6.187.994).

Fragmentos e variantes de sítios de recombinação, recombinases, marcadores de seleção, e seqüências de nucleotídeo de interesse podem ser usados, e ao menos que estipulado ao contrário, indicar que a variante ou fragmento retenha pelo menos algumas das atividades/função da composição original. Em exemplos onde o polinucleotídeo codifica uma proteína, um fragmento de um polinucleotídeo pode codificar fragmentos de proteínas que retenham a atividade biológica da proteína de comprimento total. Fragmentos de um polinucleotídeo podem variar de pelo menos cerca de 20 nucleotídeos, cerca de 50 nucleotídeos, cerca de 100 nucleotídeos, e até o polinucleotídeo de comprimento total. Um fragmento de um polinucleotídeo que codifica uma porção ativa biologicamente de uma proteína tipicamente codifica pelo menos 15, 25, 30, 50, 100, 150, 200, 250,

300, 325, 350, 375, 400, 420, ou 450 de aminoácidos contíguos, ou qualquer inteiro nesta faixa até e incluindo o número total de aminoácidos presentes em uma proteína de comprimento total. Um fragmento ativo biologicamente de um polipeptídeo pode ser preparado através de isolamento de uma porção de um dos polinucleotídeos codificando a porção do polipeptídeo de interesse, expressando o fragmento de proteína, e avaliando a atividade.

10 Alternativamente, um fragmento biologicamente ativo de um polipeptídeo pode ser produzido através de clivagem seletivamente química ou proteolítica do polipeptídeo de comprimento total, e a atividade medida. Por exemplo, polinucleotídeos que codificam fragmentos de um polipeptídeo recombinase podem compreender seqüências de nucleotídeo compreendendo pelo menos 16, 20, 50, 75, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 800, 900, 1,000, 1,100, 1,200, 1,300, ou 1,400 nucleotídeos, ou qualquer inteiro nesta faixa até e incluindo o número total de um polinucleotídeo de comprimento total. Em adição, fragmentos de um sítio de recombinação retêm a atividade biológica do sítio de recombinação, passando por um evento de recombinação na presença da recombinase apropriada. Fragmentos de um sítio de recombinação podem variar de pelo menos cerca de 5, 10, 15, 20, 25, 30, 35, 40 nucleotídeos, até o comprimento total de um sítio de recombinação. Por exemplo, sítios de FRT de comprimento total, lox, attB, e attP são conhecidos e variam de cerca de 50 nucleotídeos a cerca de 250

nucleotídeos, e totalmente mínimo ativo são conhecidos e variam de cerca de 20, 25, 30, 35, 40, 45, e 50 nucleotídeos.

5           Ensaio para medir a atividade biológica dos sítios de recombinação e recombinases são conhecidos (ver, *pex.*, Senecoll *et al.* (1988) *J Mol Biol* 201:406-421; Voziyanov *et al.* (2002) *Nucleic Acids Res* 30:7; patente Americana US 6.187.994; WO01/00158; Albert *et al.* (1995) *Plant J* 7:649-10 659; Hartang *et al.* (1998) *J Biol Chem* 273:22884-22891; Saxena *et al.* (1997) *Biochim Biophys Acta* 1340:187-204; e Hartley *et al.* (1980) *Nature* 280-860-864). Ensaio para atividade da recombinase geralmente medem a atividade global da enzima em substratos de DNA contendo sítios de 15 recombinação. Por exemplo, para analisar para atividade FLP, uma inversão de uma seqüência de DNA em um plasmídeo circular contendo dois sítios invertidos de FRT pode ser detectada como uma mudança na posição dos sítios de enzima de restrição (ver, *pex.*, Vetter *et al.* (1983) *Proc Natl 20 Acad Sci USA* 80:7284). Alternativamente, excisão de DNA de uma molécula linear ou freqüência de recombinação intermolecular induzida pela enzima pode ser analisada (ver, *pex.*, Babineau *et al.* (1985) *J Biol Chem* 260:12313; Meyer-Leon *et al.* (1987) *Nucleic Acids Res* 15:6469; and 25 Gronostajski *et al.* (1985) *J Biol Chem* 260:12328). Atividade recombinase pode também ser medida por excisão de uma seqüência flanqueada por sítios recombinogênicos FRT para ativar um gene marcador analisável.

**EXEMPLOS**

A presente invenção é ainda definida nos seguintes Exemplos, onde partes e porcentagens são por peso e graus são Celsius, a menos que estipulado de outra forma. Deve ser entendido que esses Exemplos, enquanto indicar concretizações preferidas da invenção, são dados de forma ilustrativa apenas. Da discussão acima e desses Exemplos, um especialista no assunto pode apurar as características essenciais desta invenção, e sem sair do espírito e escopo da mesma, pode fazer várias mudanças e modificações da invenção para adaptá-la a vários usos e condições. Então, várias modificações da invenção em adição aquelas mostradas e descritas aqui serão aparentes para aqueles especialistas no estado da técnica da descrição precedente. Tais modificações são também destinadas a cair dentro do escopo das reivindicações anexadas.

O significado das abreviações é o seguinte: "seg" significa segundo(s), "min" significa minuto(s), "h" significa hora(s), "d" significa dia(s), "µl" significa microlitro (s), "mL" significa mililitro(s), "L" significa litro(s), "µM" significa micromolar, "mM" significa milimolar, "M" significa molar, "mmol" significa milimol(s), "µmole" significa micromol(s), "g" significa grama(s), "µg" significa micrograma(s), "ng" significa nanograma(s), "U" significa unidade(s), "bp" significa pares de base(s) e "kB" significa kilobase(s).



**EXEMPLO 1. Identificação e isolamento de centrômeros de milho**

Para avaliar o tamanho, composição, e organização estrutural de centrômeros individuais, sondas específicas marcadas para CentC, CentA, CRM1, e/ou CRM2, foram usadas individualmente e/ou em um coquetel para hibridização fluorescente *in situ* (FISH) em cromossomos no paquíteno, metáfase e anáfase I da meiose de milho e para moléculas de DNA entendidas (fibra-FISH). Essas quatro sondas foram também usadas para rastrear bibliotecas BAC genômicas de milho.

A. Hibridização *In situ*

Multi-color FISH para cromossomos na metáfase de milho revelam que essas quatro repetições centroméricas são específicas de centrômeros e estão co-localizadas em regiões centroméricas em todos os cromossomos em células somáticas. Análises FISH mostraram que os retrotransposons CRM1, CRM2, e CentA, ocupam aproximadamente a mesma região em centrômeros de milho. Existe significativa variação na composição da repetição e tamanho relativo das regiões de repetição entre centrômeros de diferentes cromossomos de milho.

Resultados FISH mostraram que a sonda CentA teve o sinal de hibridização mais fraco; a sonda CRM1 mostrou um padrão de hibridização como gradiente com o sinal mais forte ao redor da constrição primária do cromossomo em metáfase, com o sinal desvanecendo gradualmente na

periferia das regiões do centrômero, e a sonda CRM2 mostrou o sinal de hibridização mais claro e compacto. A força do sinal FISH das repetições CentC foi altamente dependente do número de cópias CentC, que é variável entre os centrômeros de diferentes cromossomos de milho. Em alguns centrômeros CentC está firmemente aglomerada, mostrando ligeira sobreposição com outras repetições centroméricas, em outros cromossomos a distribuição da repetição CentC mostra mais sobreposição com todas as outras repetições. FISH de cromossomos meióticos na anáfase I em microsporócitos com todas as quatro repetições centroméricas revelou que a região centromérica neste estágio é altamente estendida e apenas um pequeno segmento da região centromérica inteira está realmente ligado ao cinetócoro. Todas as quatro repetições co-localizadas no segmento ligado ao microtúbulo, sugerem que uma região do centrômero funcional nativo compreende todas as quatro repetições centroméricas. Fibra-FISH em moléculas de DNA estendidas foi usada para caracterizar melhor a distribuição e arranjo das repetições centroméricas em uma resolução mais alta.

Aveia pelo cruzamento de milho gerou embriões F1 que retinham um ou mais cromossomos de milho (ver, *pex.*, Riera-Lizarazu *et al.* (1996) *Theor Appl Genet* 93:123-135; Ananiev *et al.* (1997) *Proc Natl Acad Sci USA* 94:3524-3529). Essas linhagens fornecem um meio para estudar cromossomos de milho individuais sem as complexidades de fundo dos outros nove cromossomos de milho. Um número de linhagens adicionais de aveia-milho está disponível da Ron Phillips

at University of Minnesota (St. Paul, MN, USA), incluindo as linhagens adicionais de aveia-milho Seneca 60, A188, e B73 usadas aqui.

5 DNA do cromossomo das linhagens adicionais aveia-milho foi usado para análise das regiões centroméricas de cromossomos de milho individuais. Multicolor fibra-FISH em cromossomo de linhagens de aveia-milho adicionais revelou trechos de hibridização ao longo de megabase de repetições  
10 centroméricas únicas para cada cromossomo (Figura 11). Nos cromossomos 1, 7, e 8 todas as quarto repetições foram intercaladas ao longo da região centromérica inteira. Em outros cromossomos, CentC estava presente como trechos relativamente curtos (cerca de 300 kb) flanqueados por  
15 arranjos "avulsos" de outras três repetições centroméricas. O comprimento total das regiões centroméricas variou grandemente entre diferentes cromossomos de milho como observado através de FISH. CentC revelou significante polimorfismo entre centrômeros de cromossomos individuais  
20 na abundância dessa repetição, com uma diferença de 10 vezes o observado dentro de qualquer dado genótipo. O cromossomo 7 teve o maior bloco das repetições em tandem CentC nos cromossomos em metáfase e paquíteno. Similarmente a linhagem adicional aveia-milho com cromossomo 7 teve o  
25 trecho mais longo de fibras de DNA que hibridizou com a sonda CentC. Inversamente, o centrômero do cromossomo 4 de milho teve o menor bloco das repetições CentC nos cromossomos na metáfase e a menor extensão de CentC no cromossomo 4 das linhagens adicionais de aveia-milho,

especialmente no cromossomo 4 da linhagem de milho B73. Quando analisados por fibra-FISH os retrotransposons centroméricos CentA, CRM1, e CRM2 mostraram um padrão semelhante a pontilhados com grandes lacunas entre os sinais de hibridização positivos. Quando sondas para esses 5 três retrotransposons foram misturadas juntas e usadas como uma sonda coquetel elas revelaram mais fibras de DNA marcadas contiguamente intercaladas com blocos de repetições CentC. Os flaqueios de retrotransposons centroméricos marcados contiguamente mostrando um padrão 10 tipo pontilhado ao longo das moléculas de DNA indicou que retrotransposons centroméricos foram intercalados com outros tipos de seqüências de DNA, incluindo elementos específicos não centroméricos. Os retrotransposons centroméricos podem formar arranjos avulsos de até 1 Mb em 15 centrômeros de cromossomos com pequenos blocos de repetições CentC, tais como cromossomo 4. O híbrido de milho Zapalote chico tem um cromossomo-B super numérico. FISH de cromossomos meióticos Zapalote chico indicou que o 20 centrômero funcional do cromossomo B de milho contém todas as quatro repetições centroméricas, similar àquelas observadas em todos os cromossomos-A. No entanto, grupos de repetições CentC podem ser encontrados também em vários sítios não centroméricos no braço longo do cromossomo B. 25 Aqueles sítios são aparentemente livres de outras repetições centroméricas.

Os resultados de FISH nos cromossomos mitóticos e meióticos, e fibra FISH sugeriram que o segmento

centromérico nativo funcional responsável pela formação do cinetócoro no cromossomo de milho geralmente compreende arranjos de repetições em tandem CentC intermisturada com três outras repetições centroméricas, CRM1, CRM2 e CentA  
5 (Figura 12).

### B. Bibliotecas BAC

Vetores BAC permitem a clonagem de grandes fragmentos de DNA genômico, até cerca de 300 kb em tamanho, que pode  
10 ser mantido em um hospedeiro bacteriano, tipicamente *E. coli*. Uma ampla variedade de bibliotecas BAC têm sido geradas de espécies de plantas e animais e estão disponíveis ao público, ver, por exemplo a informação na Cleidson University Genome Institute (CUGI; ver website na  
15 genome.clemson.edu) e Children's Hospital Oakland Research Institute (CHORI; ver website na chori.org). Bibliotecas BAC genômicas de milho representando mais do que 13X a cobertura usando múltiplas enzimas para construção da biblioteca de dois genótipos de milho diversos, B73 e Mo17,  
20 representando os grupos heteróticos Dent e Lancaster respectivamente, foram rastreadas para seqüências centroméricas de milho.

#### *i. Biblioteca BAC genômica de milho Mo17*

25 Os vetores de clonagem BAC pIndigoBac536 (Shizuya, unpublished) e pBeloBAC11 (Kim et al. (1996) Genomics 34:213-218) foram desenvolvidos de pBAC108L (Shizuya et al. (1992) Proc Natl Acad Sci USA 89:8794-8797). O pBAC108L é um mini fator F baseado em plasmídeo. O fator F codifica

para os genes que regulam sua própria replicação e número de cópias na célula. O vetor pBeloBAC11 foi gerado pela introdução do gene LacZ para facilitar a identificação do clone recombinante pelos fenótipos azul ou ausência de cor (branco). O pBeloBAC11 tem três sítios de clonagem únicos: *Bam*HI, *Sph*I, e *Hind*III, que são flanqueados pelos promotores T7 e SP6. Os sítios de restrição de corte raro *Not*I, *Eag*I, *Xma*I, *Sma*I, *Bgl*II, e *Sfi*I podem ser usados para excitar o inserto de pBeloBAC11. No vetor pIndigoBac536, um sítio *Eco*RI tem sido modificado no gene cloranfenicol ( $CM^R$ ) para que o sítio *Eco*RI no sítio de clonagem possa ser usado para construção da biclioteca. Os vetores pBeloBAC11 e pIndigoBac536 têm dois marcadores de seleção, LacZ e  $CM^R$  para seleção dos transformantes.

15

Uma biblioteca BAC genômica proprietária de milho da linhagem pura pública de milho Mo17 foi construída em pBeloBAC11 ou pIndigoBac536 essencialmente como descrita em Kim *et al.* ((1996) *Genomics* 34:213-218) sobre contrato com o laboratório Shizuya no Instituto de Tecnologia da Califórnia. Brevemente, DNA genômico Mo17 foi parcialmente digerido com as enzimas de restrição *Hind*III ou *Eco*RI. Os fragmentos de DNA foram fracionados por tamanho em gel de agarose e clonados em sítios *Hind*III no pBeloBAC11 ou nos sítios *Eco*RI no pIndigoBac536. O tamanho médio do inserto foi cerca de 150 kb. A biblioteca BAC genômica inteira de Mo17 consiste de 433 placas de 384 poços ou 166,272 clones BAC totais. A primeira metade da biblioteca compreende de 214 placas contendo clones BAC com insertos *Hind*III,

25

enquanto que a segunda metade da biblioteca compreende 219 placas, contém clones BAC com insertos *EcoRI*. Os clones BAC são mantidos em *E. coli* DH10B (BRL Life Technologies).

5 *ii. Bibliotecas BAC genômicas de milho B73*

Duas bibliotecas públicas BAC genômicas de milho B73 foram obtidas. A biblioteca ZMMBBb está disponível do Clemson University Genome Institute (CUGI, University of Georgia, Athens, GA, USA). A biblioteca BAC ZMMBBb foi criada no CUGI através da clonagem parcialmente digerida com *HindIII* do DNA genômico de milho B73 no vetor pIndigoBac536 compreendendo um gene de resistência ao cloranfenicol ( $CM^R$ ). A biblioteca BAC ZMMBBb compreende 10 247,680 de clones BAC totais com uma media do tamanho do inserto de cerca de 137 kb, representando um cobertura 15 genômica de 14X. A segunda biblioteca BAC B73 BAC, CHORI-201 (ZMMBBc) criada por Pieter de Jong's laboratory no Children's Hospital Oakland Research Institute (CHORI), está disponível do BACPAC Resource Center at CHORI. Para 20 construir esta biblioteca, DNA genômico foi isolado do núcleo de milho B73. O primeiro segmento da biblioteca foi construído usando DNA parcialmente digerido com uma combinação de *EcoRI* e *EcoRI* metilase, o segundo segmento foi construído usando DNA parcialmente digerido com *MboI*. 25 DNA selecionado por tamanho foi clonado dentro do vetor pTARBAC2.1 (segmento 1, placas 1-288) entre os sítios *EcoRI* e dentro do vetor pTARBAC1.3 (segmento 2, placas 289-576) entre os sítios *BamHI*. Os produtos de ligação foram transformados dentro de células de *E. coli* DH10B

eletrocompetentes (BRL Life Technologies). Os clones BAC para cada segmento de biblioteca em cada vetor têm sido arranjados em 288 placas de 384-poços microtitulados. O segmento 1 compreende 106,637 clones BAC individuais com  
5 uma média de tamanho de inserto de 163 kb, representando uma cobertura genômica de 6.9X. O segmento 2 compreende 105,579 clones BAC individuais com uma média de tamanho de inserto de 167 kb, representando uma cobertura genômica de 7.0X. A biblioteca total ZMMBBc compreende 212,216 clones  
10 BAC individuais com uma média de tamanho de inserto de 165 kb, representando uma cobertura genômica de 13.9X.

### C. Rastreamento de biblioteca BAC

Bibliotecas de milho B73 e Mo17 BAC foram rastreadas  
15 com quatro sondas separadas para seqüências centroméricas CentA, CentC, CRM1, e CRM2. As sondas foram desenhadas como oligonucleotídeos OVERGO de 40 bp de comprimento e foram únicas para cada elemento de centrômero. Através do uso de marcações apropriadas, essas sondas podem ser usadas para  
20 colônia, e hibridização blot, e FISH e fiber-FISH.

#### *i. Sondas Overgo*

Sondas Overgo são tipicamente desenhadas como dois oligonucleotídeos curtos que têm uma região de sobreposição  
25 de 8 bp de complementariedade. Os oligonucleotídeos curtos estão tipicamente na variação de 23-28 bp, com 24 bp sendo mais comumente usado. Após anelamento, os oligonucleotídeos formam dímeros com DNA de fita simples de 16 bp em ambos os lados. A sonda parcialmente de fita dupla é marcada pelo



preenchimento do recesso da extremidade 3' usando atividade de polimerização da enzima Klenow na presença de nucleotídeos marcados. A sonda overgo final compreende uma sonda marcada de fita dupla de 40 bp. TABELA 1 lista iniciadores e sondas usados para geração, rastreamento, e caracterização dos clones BAC, construções de DNA, e eventos de minicromossomos de milho.

**TABLE 1**

SEQ ID	Biocódigo	Nome do Oligo	Seqüência
5		PCR-Telomere-F	AGGGTTTAGGGTTTAGGGTTTAGGGTTTAGGGTTTAGGGTTT AGGG
6		PCR-Telomere-R	CCCTAAACCCTAAACCCTAAACCCTAAACCCTAAACC C
7	65644	CentC-OVG-1-40f	GGTCCGGTGGCAAAACTCGTGC
8	65645	CentC-OVG-1-40r	TGTCGGTGCATACAAAGCACGAGT
9	65646	CentC-OVG-51-90f	GAATGGGTGACGTGCGACAACGAA
10	65647	CentC-OVG-51-90r	GGTGGTTTCTCGCAATTTTCGTGT
11	65648	CentC-OVG-101-140f	GTTTTGGACCTAAAGTAGTGGATT
12	104790	CentC-OVG-101-140r	CACAACGAACATGCCCAATCCACT
13	69509	CRM1-LTR-OVG1f	CTTGGTCTTGGACAGTACCTCACT
14	69510	CRM1-LTR-OVG2f	CCCTTGCGATCCGACTACGACGAG
15	69511	CRM1-LTR-OVG3f	TCACGAAGATCGTTTTCTGTGCGC
16	69512	CRM1-LTR-OVG4f	CAGCGCAGATTAGCGCGTGTTCGA
17	69513	CRM1-LTR-OVG5f	CCAACCCTAGGTCGTCCATTATGG
18	69514	CRM1-LTR-OVG6f	TTCAATTCTCTTGACGGGCCCGA
19	69515	CRM1-LTR-OVG1r	TCAGGTCTACTTCATCAGTGAGGT
20	69516	CRM1-LTR-OVG2r	TGGCGCCTCGGGCTTGCTCGTCGT
21	69517	CRM1-LTR-OVG3r	TGTTTCGTTCTTCGATTGCGCACAG
22	69518	CRM1-LTR-OVG4r	TTAGCCTTAGCTACTCTCGAACAC
23	69519	CRM1-LTR-OVG5r	CCAGCCCAATTGCGGCCATAATG
24	69520	CRM1-LTR-OVG6r	CACCTGGGCCAGTGACTCGGGCCC
25	69521	CRM2-LTR-OVG1f	TGATGAAGACATCCACTACTGA
26	69522	CRM2-LTR-OVG2f	TTGAACATGCTGGATTCCGACTGC
27	69523	CRM2-LTR-OVG3f	CTGCCCATGGTGCTGCGTCACCCT
28	69524	CRM2-LTR-OVG4f	GCGCGTGCTAGTTCAGCCGCCGCT
29	69525	CRM2-LTR-OVG5f	GTATCGGTTGCTAAGGCAGCGT
30	69526	CRM2-LTR-OVG1r	TATTGGTATAGATGCATCAGTAGT
31	69527	CRM2-LTR-OVG2r	AAGTTGGTGTTCTTCTGCAGTCCG
32	69528	CRM2-LTR-OVG3r	CCCATTGGGCCAAAATAGGGTGACG
33	69529	CRM2-LTR-OVG4r	TTCCGAAGACAAGAAGACGGGCGG

34	69530	CRM2-LTR-OVG5r	CTACAGCCTTCCAAAGACGCTGCG
35	69531	CentA-LTR-OVG1f	TGATGAGAACATAAACCCGCACAGA
36	69532	CentA-LTR-OVG2f	AGGATGATGAGGACATCACTGCCA
37	69533	CentA-LTR-OVG3f	AACCATCTAGAATTTGAGAAGGCA
38	69534	CentA-LTR-OVG4f	GTCCAGAAACTGCCGAGTGAAGTC
39	65535	CentA-LTR-OVG5f	GAGAGAGTTTCGTTCTCCATTAGA
40	69536	CentA-LTR-OVG6f	GTTCTTGCTTGTTCGATTGCTT
41	69537	CentA-LTR-OVG7f	TTGGTTGTGGTAGTCGGGCAGCCA
42	69538	CentA-LTR-OVG1r	CATTAACATGGTCATATCTGTGCG
43	69539	CentA-LTR-OVG2r	TGGTGTGGTGTATTGATGGCAGTG
44	69540	CentA-LTR-OVG3r	CTTTTATTGCCTTGTTGCCTTCT
45	69541	CentA-LTR-OVG4r	GACTTGGGTAGAGCAGGAGTTCAC
46	69542	CentA-LTR-OVG5r	AGGAATAGAAAGGAGTTCTAATGG
47	69543	CentA-LTR-OVG6r	ACAGCCTTGAACCTGCAAGCAATC
48	69544	CentA-LTR-OVG7r	TGTTGGAGAACGACGTTGGCTGCC
49	69555	Cent4-250-OVG1f	TAAGTGCAAACCATTGTTAAATTT
50	69556	Cent4-250-OVG2f	CACAAACCCTTAACCTCGAAACTAT
51	69557	Cent4-250-OVG3f	ATCGAAAGATAACTCATATGGCTT
52	69558	Cent4-250-OVG4f	TCCACTAAAGAACCAAGATTGTGA
53	69559	Cent4-250-OVG1r	AATTGTACTATCTCTAAAATTTAA
54	69560	Cent4-250-OVG2r	TTTAGGGTTTGGGGTTATAGTTTC
55	69561	Cent4-250-OVG3r	GACCATAATGGTCAAAAAGCCATA
56	69562	Cent4-250-OVG4r	ATATGTTGGACACAAATCACAAATC
57	69634	18-26SrDNANTS-OvG1f	CCGGAATAAGCAAAGTCCAAGCG
58	69635	18-26SrDNANTS-OvG2f	TATGTCTTGGGTGAAGGGCATGGC
59	69636	18-26SrDNANTS-OvG3f	CGCAAGGCGACGGGCGGCATGGCT
60	69637	18-26SrDNANTS-OvG4f	CGAGGGGTTCCCATGGCGCACGG
61	69638	18-26SrDNANTS-OvG1r	TCGGTGTCTTTCCACACGCTTGGA
62	69639	18-26SrDNANTS-OvG2r	GTTTTCCCTCCGTTCCGCCATGCC
63	69640	18-26SrDNANTS-OvG3r	AGACGCAAGGCCGAACAGCCATGC
64	69641	18-26SrDNANTS-OvG4r	GGCCTCAGTTTTTCGGCCCCGTGCGC
65	74794	subtelo-TR430-OvG2f	GACACATGTTTTTGTCGTCGAACA
66	74795	subtelo-TR430-OvG2r	GGAGGCACGAAATCGCTGTTTCGAC
67	74796	subtelo-TR430-OvG3f	CGACCGCCACCCATGATTTGACCA
68	74797	subtelo-TR430-OvG3r	ACCTTACCAGTCTCTATGGTCAAA
69	74799	subtelo-TR430-	TCCCGTGAGCTATAGCACACGTTT

		OvG4f	
70	74800	subtelo-TR430-OvG4r	GGTCGCTCGGCCATGAAAACGTGT
71	74801	subtelo-TR430-OvG5f	CCGTGTTCCCTCCACACGTGTTTT
72	74802	subtelo-TR430-OvG5r	AAGGTGCTCCGGGGACAAAAACAC
73	74803	subtelo-TR430-OvG6f	TTGGCCTCCCGCGAGCTATATCAC
74	74804	subtelo-TR430-OvG6r	TTGGCCACGGAAATGTGTGATATA
75	74805	subtelo-TR430-OvG7f	TTATGTATCCGACCTGCCACCTTC
76	74806	subtelo-TR430-OvG7r	CTCCCCGGTCTAAAACGAAGGTGG
77	74807	subtelo-TR430-OvG8f	GCCACCCGTGAGCTATAGCACACG
78	74808	subtelo-TR430-OvG8r	TAGGTTTCCATAAAAATCGTGTGCT
79	65650	180knobOvG21-60f	TGTCGAAAATAGCCATGAACGACC
80	65651	180knobOvG21-60r	CGGTATTATTGGAAATGGTCGTTTC
81	65652	180knobOvG71-110f	CCTACGGATTTTTGACCAAGAAAT
82	65653	180knobOvG71-110r	ATTTCTAGTGGAGACCATTTCTTG
83	65654	180knobOvG141-180f	ATGTGGGGTGAGGTGTATGAGCCT
84	65655	180knobOvG141-180r	ATGAGCCTCTGGTCGATGATCAAT
85	65656	5SrDNAOvG1-40f	GGATGCGATCATAACCAGCACTAAA
86	65657	5SrDNAOvG1-40r	TGATGGGATCCGGTGCTTTAGTGC
87	65658	5SrDNAOvG61-100f	CTTGGGCGAGAGTAGTACTAGGAT
88	65659	5SrDNAOvG61-100r	TCCCAGGAGGTCACCCATCCTAGT
89	65660	5SrDNAOvG161-200f	ACCATAGTAAAAATGGGTGACCGT
90	65661	5SrDNAOvG161-200r	TAATTTAACACGAGAACGGTCAC
91	65662	5SrDNAOvG261-230f	CCGTGGGCGAGCCGAGCACGGAGG
92	65663	5SrDNAOvG261-230r	TCCTCTTATGCCACACCTCCGTG
93	65664	350knobOvG31-70f	CTCAAATGACGTTTCTATGATATT
94	65665	350knobOvG31-70r	TGAATACAATGCCCTCAATATCAT
95	65666	350knobOvG121-160f	CTAGGTTTCTATAATCCCCTCTA
96	65667	350knobOvG121-160r	CTAGGTATGCCTTGAATAGAGGG
97	65668	350knobOvG161-200f	ATGTTGTTTATGTCCACTCAAGTA
98	65669	350knobOvG161-200r	ATGGTGTACGGTGTTTTACTTGAG
99	65670	350knobOvG261-300f	GTGAGATCTGTCCAAACATAGGTT
100	65671	350knobOvG261-300r	GGTGCCTTACAACCGTAACCTATG
101		b010.m7 fis31	GCAAACCTTTATGTGATCCCTTCCTCGCTG AACGAGATGAG
102		b108.h15 fis47	GGGACGGCAAGTCACGGTAAGACCAGTCC AACCGAATGAT
103		Cen3n.pk0001.g11	CCAAACTTGCTGAGATTACTGGGCAATCT GTTTCGCTCGCA
104	103022	23715-3101-3200f	CCAGGTAGTTTGAAACAGTATTCT

105	103023	23715-3501-3600f	ATAAAGGAAAAGGGCAAACCAAAC
106	103024	23715-1401-1500f	GATGCCACATTATAGTGATTAGC
107	103025	23715-2901-3000f	CCACATATAGCTGCTGCATATGCC
108	103026	23715-3701-3800f	CGGATCTAACACAAACATGAACAG
109	103027	23715-1-100f	CGATGAATTTTCTCGGGTGTTC
110	103028	23715-101-200f	CCTGCAGCCCTAATAATTCAGAAG
111	103029	23715-301-400f	CACAGTCGATGAATCCAGAAAAGC
112	103030	23715-901-1000f	GCGTGCAATCCATCTTGTTCAATC
113	103031	23715-3201-3300f	CAACCACACCACATCATCACAACC
114	103032	23715-3601-3700f	ACTGGCAAGTTAGCAATCAGAACG
115	103033	23715-4901-5000f	CATGAACGTGTCTTCAACTAGAGG
116	103034	23715-4201-4300f	GACGGCGTTTAAACAGGCTGGCATT
117	103035	23715-201-300f	CCAAGCTCTTCAGCAATATCACGG
118	103036	23715-601-700f	ATACTTTCTCGGCAGGAGCAAGGT
119	103037	23715-1001-1100f	ATCCTTGCGCGCAAGAAAGCCATC
120	103038	23715-1101-1200f	GCAAGCTACCTGCTTTCTCTTTC
121	103039	23715-1601-1700f	GCTTCTTGCCATGTAGATGGACT
122	103040	23715-1801-1900f	TTCACGCCGATGAACTTCACCTTG
123	103041	23715-5001-5087f	AAGCTTGCCAACGACTACGCACTA
124	103042	23715-401-500f	CCCTGATGCTCTTCGTCCAGATCA
125	103043	23715-801-900f	AGAGCAGCCGATTGTCTGTTGTGC
126	103044	23715-1301-1400f	CAGGATCCCGTAACTATAACGGTC
127	103045	23715-2801-2900f	CGACCTGCAGAAGTAACACCAAAC
128	103046	23715-3401-3500f	ATCTAGAACGACCGCCCAACCAGA
129	103047	23715-3801-3900f	ATTTGGGGGAGATCTGGTTGTGTG
130	103048	23715-3901-4000f	GAGGGGGTGTCTATTTATTACGGC
131	103049	23715-4801-4900f	CATGCAAGCTGATCTGAGCTTGGC
132	103050	23715-2101-2200f	TCCATGCGCACCTTGAAGCGCATG
133	103051	23715-501-600f	TTCCATCCGAGTACGTGCTCGCTC
134	103052	23715-1201-1300f	ATCCACTAGTAACGGCCGCCAGTG
135	103053	23715-4001-4100f	GCCACGCAATTTCTGGATGCCGAC
136	103054	23715-701-800f	CGATAGCCGCGCTGCCTCGTCTTG
137	103055	23715-1901-2000f	CACTTGAAGCCCTCGGGGAAGGAC
138	103056	23715-1701-1800f	TCCTTCAGCTTCAGGGCCTTGTGG
139	103057	23715-2001-2100f	CACCTTGGAGCCGTACTGGAAGTG
140	103058	23715-2601-2700f	TGCGGCTCGGTGCGGAAGTTCACG
141	103059	23715-4101-4200f	ACGCGACGCTGCTGGTTCGCTGGT
142	103060	23715-3101-3200r	CGTTCTAGATCGGAGTAGAATACT
143	103061	23715-3501-3600r	TGTTTCGTTGCATAGGGTTTGGTT
144	33332	23715-1401-1500r	GCACACATAGTGACATGCTAATCA
145	103062	23715-2901-3000r	GATATACTTGGATGATGGCATATG
146	103063	23715-3701-3800r	CCCGGTAGTTCTACTTCTGTTTCAT
147	103064	23715-1-100r	ATTCGAGCCAATATGCGAGAACAC
148	103065	23715-101-200r	GCCTTCTTGACGAGTTCTTCTGAA
149	103066	23715-301-400r	ATGGTGGAAAATGGCCGCTTTTCT
150	103067	23715-901-1000r	GAGGATCGTTTCGCATGATTGAAC
151	103068	23715-3201-3300r	TGCTTTTTGTTTCGCTTGGTTGTGA
152	103069	23715-3601-3700r	ACCTGTACGTCAGACACGTTCTGA

153	103070	23715-4901-5000r	AATTAAGTCAGGCGCGCCTCTAGT
154	103071	23715-4201-4300r	CTTGTTTCGAGTAGATAATGCCAG
155	103072	23715-201-300r	ACATAGCGTTGGCTACCCGTGATA
156	103073	23715-601-700r	GATCTCCTGTCATCTCACCTTGCT
157	103074	23715-1001-1100r	CCTGCAAAGTAAACTGGATGGCTT
158	103075	23715-1101-1200r	AAGGGAAAACGCAAGCGCAAAGAG
159	103076	23715-1601-1700r	TACCTGGTGGAGTTCAAGTCCATC
160	103077	23715-1801-1900r	ACGGCTGCTTCATCTACAAGGTGA
161	103078	23715-5001-5087r	TGAAGCTCTTGTTGGCTAGTGCCT
162	103079	23715-401-500r	GTCTTGTCGATCAGGATGATCTGG
163	103080	23715-801-900r	ATTCGGCTATGACTGGGCACAACA
164	103081	23715-1301-1400r	CGCTTCGCTACCTTAGGACCGTTA
165	103082	23715-2801-2900r	CGATGCTCACCTGTTGTTTGGTG
166	88245	23715-3401-3500r	GGTTGTGATGATGTGGTCTGGTTG
167	103083	23715-3801-3900r	GTTCCGGAGCGCACACACACAAC
168	103084	23715-3901-4000r	TTCCCTTCCTCGCCCCCGCTAAT
169	103085	23715-4801-4900r	TAAAACGACGGCCAGTGCCAAGCT
170	103086	23715-2101-2200r	ACGTCATCACCGAGTTCATGCGCT
171	103087	23715-501-600r	AGCGAAACATCGCATCGAGCGAGC
172	103088	23715-1201-1300r	AAGCCGAATCCAGCACACTGGCG
173	103089	23715-4001-4100r	TTGGACTTGCTCCGCTGTCCGGCAT
174	103090	23715-701-800r	TGCCCTGAATGAACTGCAAGACGA
175	103091	23715-1901-2000r	CCGACTACAAGAAGCTGTCCTTCC
176	103092	23715-1701-1800r	TGCTGAAGGGCGAGACCCACAAGG
177	103093	23715-2001-2100r	GGACATCCTGTCCCCCAGTTCCA
178	103094	23715-2601-2700r	ACATCGAGACCTCCACCGTGAACCT
179	103095	23715-4101-4200r	AGTCTAACGGACACCAACCAGCGA
180		PCRbacmpk108h15f	GATCGTCGAATGGGAATCCATGGG
181		PCRbacmpk108h15r	CCCTGAGTGAACCATTTAGGAAGATCAG
182		PCRbacmpk108h15- 2.fis47f	TGCAACATCCAAGACCCAACATG
183		PCRbacmpk108h15- 2.fis47r	TTCCAACATGGTTGGTGGTCAG
184		PCRbacmpk010m07fis 31f	TGTCATGACATCTTGTTGCTACCCTG
185		PCRbacmpk010m07fis 31r	AAACCCGGAGTTTCTATGCAGG
192	75319	Telo-31overgo primer1	AGGGTTTAGGGTTTAGGGTTTAGGGTTTA GGG
193	39612	Telo-31overgo primer2	CCCTAAACCCTAAACCCTAAACCCTAAAC CC

ii. Resultados do rastreamento da biblioteca BAC

O rastreamento da hibridização da colônia identificou um grupo de aproximadamente 8000 clones BAC que

hibridizaram com pelo menos uma das quatro sondas específicas de centrômero. Os 8000 clones BAC foram classificados em 4 grupos baseados em seu perfil de hibridização (Tabela 2).

5 **TABELA 2**

<b>Grupo</b>	<b>Total</b>
Todos BACs contendo CentA	842
Todos BACs contendo CentC	2479
Todos BACs contendo CRM2	2968
Todos BACs contendo CRM1	6012

Baseados na composição da repetição centromérica os clones BAC foram ainda classificados em 5 grupos baseados na composição das sondas que hibridizaram com cada clone BAC particular (Tabela 3).

10

**TABELA 3**

<b>Grupo</b>	<b># de BACs</b>
BACs contendo CentA & CentC & CRM1 & CRM2	247
BACs CentA & CentC & CRM2; não CRM1	6
BACs contendo CentA & CentC & CRM1; não CRM2	45
BACs contendo CentA & CRM1 & CRM2; não CentC	116
BACs contendo CentC & CRM1 & CRM2; não CentA	730
BACs CentA & CentC; não CRM1 não CRM2	4
BACs contendo CentA & CRM1; não CentC não CRM2	131
BACs contendo CentA & CRM2; não CentC não CRM1	27
BACs contendo CentC & CRM2; não CentA não CRM1	97
BACs contendo CentC & CRM1; não CentA não CRM2	829
BACs contendo CRM1 & CRM2; não CentA não CentC	749
BACs CentC; não CentA não CRM1 não CRM2	521
BACs contendo CRM2; não CentA não CentC não CRM1	966
BACs contendo CRM1; não CentA não CentC não CRM2	3165
BACs contendo CentA; não CentC não CRM1 não CRM2	266

Os clones BAC foram ainda classificados baseados no

somatório dos clones BAC que hibridizaram para cada sonda particular (Tabela 4).

5 **TABELA 4**

Todos BACs contendo CentA	842
Todos BACs contendo CentA & CentC	302
Todos BACs contendo CentA, CentC, & CRM1	292
Todos BACs contendo CentA, CentC, & CRM2	253
Todos BACs contendo CentA & CRM1	539
Todos BACs contendo CentA, CRM1, & CRM2	363
Todos BACs contendo CentA & CRM2	396
Todos BACs contendo CentA, CentC, CRM1, & CRM2	247
Todos BACs contendo CentC	2479
Todos BACs contendo CentC & CRM1	1851
Todos BACs contendo CentC & CRM2	1080
Todos BACs contendo CentC, CRM1, & CRM2	977
Todos BACs contendo CRM1	6012
Todos BACs contendo CRM1 & CRM2	1842
Todos BACs contendo CRM2	2968

Um grupo de 247 clones BAC contém todas as quatro repetições centroméricas. Eles compreendem 0.15% do genoma do milho ou podem estar presentes em um segmento de DNA de  
 10 cerca de 300 kb por centrômero em média. Este grupo de clones BAC foi identificado como o grupo principal a ser usado primeiro em experimentos para construir um minicromossomo de milho. O DNA foi purificado de todos os  
 247 BACs no grupo principal, digeridos com *XmnI* ou *RsaI*,  
 15 transferencia de DNA (blotted) e hibridizado com cada uma das quatro repetições centroméricas. Hibridização southern blot, confirmou que clones neste grupo principal continham todas as quatro repetições centroméricas. Os BACs mostraram diferenças gerais na composição do fragmento de restrição e  
 20 padrões de hibridização, e foram ainda classificados dentro

de 87 grupos com base na similaridade dos fragmentos de restrição. Um representante de cada um dos 87 grupos (Tabela 5) foi tomado para gerar um grupo principal de construções de DNA e/ou grupos de grupos principais de construções BAC para transformação e montagem de minicromossomo.

TABELA 5

No.	Nome	Inserto (kb)	No.	Nome	Inserto (kb)
1	bacm.pk101.n23	50	45	bacm2.pk002.g7	125
2	bacm2.pk064.e15	50	46	bacm.pk135.l7	125
3	bacm.pk036.e13	60	47	bacm.pk090.o5	125
4	bacm2.pk179.e1	70	48	bacm2.pk100.j24	130
5	bacm.pk030.a6	70	49	bacm2.pk013.c9	130
6	bacm2.pk179.b18	75	50	bacm.pk166.n7	130
7	bacm.pk133.b11	75	51	bacm.pk043.o23	130
8	bacm2.pk066.m12	80	52	bacm.pk001.n1	130
9	bacm.pk119.a23	80	53	bacm.pk106.j20	135
10	bacm.pk098.h2	85	54	bacm.pk015.d19	135
11	bacm2.pk174.e4	90	55	bacm.pk007.a2	140
12	bacm2.pk116.g16	90	56	bacm.pk148.e2	140
13	bacm2.pk023.e24	90	57	bacm.pk141.j4	140
14	bacm.pk178.c10	90	58	bacm.pk138.e14	140
15	bacm.pk135.l6	90	59	bacm.pk135.j2	140
16	bacm.pk098.f3	90	60	bacm.pk134.f15	140
17	bacm.pk075.l6	90	61	bacm.pk085.k5	140
18	bacm.pk066.j14	95	62	bacm.pk077.b21	140
19	bacm2.pk099.m24	100	63	bacm.pk124.j24	145
20	bacm2.pk093.h11	100	64	bacm.pk023.i5	145
21	bacm2.pk083.a2	100	65	bacm.pk039.m16	150
22	bacm.pk179.d4	100	66	bacm2.pk169.a21	150
23	bacm.pk076.m3	100	67	bacm2.pk130.e20	150
24	bacm.pk070.h17	100	68	bacm.pk156.i17	150
25	bacm.pk064.n1	100	69	bacm.pk143.m18	150
26	bacm.pk011.l8	100	70	bacm.pk112.p1	150
27	bacm.pk068.p16	105	71	bacm.pk102.i4	150
28	bacm.pk012.n20	105	72	bacm.pk087.m4	150
29	bacm.pk077.k5	110	73	bacm.pk079.m11	150
30	bacm2.pk053.g23	110	74	bacm.pk041.e16	150
31	bacm2.pk034.j8	110	75	bacm.pk129.a4	150
32	bacm.pk164.b11	110	76	bacm.pk164.e18	155
33	bacm.pk062.c14	110	77	bacm.pk161.h1	155



34	bacm.pk013.m8	110	78	bacm.pk089.l8	155
35	bacm.pk056.j19	110	79	bacm.pk076.o15	160
36	bacm.pk051.g11	115	80	bacm.pk039.a3	160
37	bacm2.pk179.o14	120	81	bacm.pk019.h24	160
38	bacm2.pk096.d23	120	82	bacm2.pk158.f12	160
39	bacm2.pk070.g7	120	83	bacm2.pk075.n6	170
40	bacm2.pk034.g20	120	84	bacm2.pk137.f2	175
41	bacm2.pk012.g19	120	85	bacm.pk093.d8	175
42	bacm2.pk115.o22	125	86	bacm.pk133.b10	180
43	bacm2.pk094.f14	125	87	bacm.pk178.o20	180
44	bacm2.pk003.g6	125			

#### D. Identificação de arranjos invertidos de repetições CentC

Bibliotecas BAC de linhagens de milho Mo17 e B73 foram buscadas para arranjos em tandem invertidos de CentC. Uma busca no BLAST de um banco de dados de seqüências da extremidade BAC Mo17 revelaram 591 extremidades BAC contendo repetições. Dessas, apenas 45 clones BAC continham repetições CentC em ambas as extremidades, e 44 BACs tinham repetições CentC na mesma orientação, com apenas um BAC tendo repetições CentC em uma orientação invertida (bacm.pk128.j21). Um segundo clone BAC, bacm.pk008.d20 tendo repetições CentC em uma orientação invertida foi encontrado por análises de hibridização em Southern. As análises de Southern deste clone mostraram um padrão de hibridização muito similar ao padrão observado para bacm.pk128.j21. Uma busca no BLAST do banco de dados de seqüências públicas de extremidades BAC B73 revelou 136 extremidades BAC contendo repetições CentC. Dessas, apenas 5 clones BAC continham repetições CentC e ambas as extremidades, e 4 BACs tinham repetições CentC na mesma orientação, com apenas um BAC tendo repetições CentC na orientação invertida (ZMMBBb0243L15). O DNA de

5 bacm.pk128.j21 e bacm.pk008.d20 foram digeridos com a enzima de restrição *XmnI*, que cliva repetições CentC em curtos fragmentos monoméricos ou diméricos. Um fragmento *XmnI* de 10 kb foi isolado, subclonado e seqüenciado. A análise da seqüência mostrou que o elemento CRM1 (SEQ ID NO: 191) está localizado entre duas repetições invertidas CentC.

10 E. Isolamento de clones BAC centroméricos do cromossomo 4 de milho

O cromossomo 4 de milho contém o arranjo de repetição mais curto de repetições CentC. Esses arranjos estão presentes em um simples trecho de DNA de aproximadamente 300 kb como estimado por fibra-FISH. Este segmento pode conter as seqüências de DNA centromérico funcional principais, e poderia potencialmente ser representado por 2-4 clones BAC sobrepostos. Os clones BAC centroméricos específicos do cromossomo 4 podem ser identificados através seqüências de DNA únicas encontradas localizadas na região centromérica do cromossomo 4.

25 A biblioteca BAC genômica do milho Mo17, compreendendo 10,965 extremidades de seqüências BAC foi analisada para identificar extremidades únicas de seqüências BAC representadas apenas uma vez na biblioteca. Oitenta e uma extremidades de seqüências BAC únicas foram identificadas e selecionadas para caracterizações posteriores. Um par de iniciadores de PCR foi desenhado para cada uma das 81 extremidades de seqüências BAC únicas para mapeamento no

painel de linhagens adicionais de cromossomos aveia-milho e cada seqüência única designada para um cromossomo de milho individual.

5           A extremidade de seqüência BAC de bacm.pk108.h15 (170 kb) de Mol7 foi mapeada para o cromossomo 4. Este BAC foi sequenciado e 6 seqüências únicas, bem como todas as quatro repetições centroméricas CentA, CentC, CRM1 e CRM2 foram encontradas. Usando PCR, este BAC foi designada para um  
10           contig contendo diversos BACs que também hibridizam com CentC. Seqüenciamento confirmou que dois ou mais clones BAC deste contig, bacm.pk010.m7 (170 kb), e bacm.pk184.c21 (150 kb) sobrepunham parcialmente com bacm.pk108.h15 e compartilham alguns marcadores únicos. Três seqüências de  
15           DNA únicas foram identificadas dentro desses três clones BAC e sua localização no cromossomo 4 foi confirmada por PCR em DNA da linhagem adicional de aveia-milho. Sondas overgo correspondentes (SEQ ID NOs: 102-104 na Tabela 1) foram desenvolvidas e usadas para rastreamento de uma  
20           biblioteca BAC pública B73.

          Sete clones BAC da biblioteca BAC B73 foram selecionados baseados na hibridização para todas as três sondas específicas do cromossomo 4. O DNA desses clones BAC  
25           foi digerido com *XmnI*, transferido para uma membrana e hibridizado com todas as quatro sondas da repetição centromérica. Quatro dos clones BAC B73 selecionados contém os elementos repetidos centroméricos CentC, CRM1, e CRM2: bacb.0424.d20 (150 kb); bacb.0155.h15 (175 kb);

bacc.0048.g5 (170 kb); e bacc.0237.m8 (125 kb). Outros três clones BAC B73 contém apenas os elementos repetidos centroméricos CRM1 e CRM2: bacc.0143.i9 (205 kb); bacc.0237.j16 (175 kb); e bacc.0270.c1 (180 kb).

5 Sequenciamento do clone BAC bacb.0155.h15 confirmou que ele contém regiões significantes de homologia para os clones BAC Mol7 específicos do cromossomo 4 bacm.pk010.m7, e bacm.pk108.h15.

10 Dois grupos de clones BAC representando a região centromérica do cromossomo 4 das linhagens puras Mol7 e B73 foram usados para a produção de construções de DNA para montagem de minicromossomo.

15 F. Isolamento e purificação dos fragmentos de DNA centroméricos cromossômicos

Essencialmente todo DNA genômico de milho é fortemente metilado, e este padrão de metilação pode desempenhar um papel na montagem, função, e/ou manutenção dos centrômeros de milho. DNA genômico de milho isolado mantendo a metilação e/ou outras características genômicas nativas, tais como tamanho, organização dos elementos, e outras modificações de nucleotídeo nativas, pode ser usado para gerar construções de DNA para montagem do minicromossomo de milho.

25

*i. Seleção das enzimas de restrição*

Análises de seqüência de repetições centroméricas de milho identificaram um grande número de enzimas de

restrição (seis enzimas) com nenhum sítio de reconhecimento dentro de qualquer das repetições centroméricas CentA, CentC, CRM1, ou CRM2 (Tabela 6). Essas enzimas de restrição deveriam digerir o volume do DNA genômico em pequenos  
5 fragmentos de DNA, a maioria das quais sendo cerca de 1-20 kb em tamanho, enquanto o DNA centromérico é esperado ser significativamente mais longo. Regiões centroméricas cromossômicas de milho podem ser isoladas por digestão parcial ou completa do DNA genômico de milho de alto peso  
10 molecular (HMW) com pelo menos uma dessas enzimas de restrição. A fração do DNA genômico HMW digerido compreendendo grandes fragmentos de aproximadamente 50 kb - cerca de 1000 kb podem ser purificados após o campo de gel de eletroforese pulsado (PFGE) de núcleos de milho  
15 embebidos em blocos de agarose.

*ii. Preparação e caracterização de DNA HMW genômico de milho*

DNA HMW genômico de milho de Mol7 foi preparado essencialmente como descrito por Liu & Whittier ((1994)  
20 Nucleic Acids Res 22:2168-2169) do DNA embebido em blocos de agarose através da digestão com várias enzimas de restrição da TABELA 6 e fracionização por PFGE. Cinco enzimas de restrição, *BspTI*, *AatII*, *Cfr9I*, *MbiI*, *MluI*, foram selecionadas para análises iniciais. Dessas, *BspTI*  
25 foi selecionada para todas as outras preparações. Hibridação Blot com sonda centromérica CentC marcada revelaram que a enzima de restrição *BspTI* produziu um grupo de fragmentos de DNA centroméricos genômicos variando de cerca de 50 kb a cerca de 600 kb que foram bem separados do

resto do DNA genômico. Hibridização dos mesmos fragmentos de DNA com três outras sondas centroméricas (CentA, CRM1, e CRM2) confirmaram que esses longos fragmentos de DNA compreendendo todas as quatro repetições centroméricas não tiveram essencialmente nenhum sítio de restrição *BspTI*. As bandas de hibridização podem representar fragmentos de DNA centroméricos individuais que podem ser isolados e usados para gerar construções de DNA para montagem de minicromossomo.

10

TABELA 6

<b>Enzima</b>	<b>Sítio restrição</b>	<b>Enzima</b>	<b>Sítio restrição</b>
<i>AatI</i>	AGGCCT	<i>NgoPIII</i>	CCGCGG
<i>AatII</i>	GACGTC	<i>Pac25I</i>	CCCGGG
<i>AccBSI</i>	CCGCTC	<i>Pae14kI</i>	CCGCGG
<i>AflIII</i>	CTTAAG	<i>Pae5kI</i>	CCGCGG
<i>AhyI</i>	CCCGGG	<i>PaeAI</i>	CCGCGG
<i>AspMI</i>	AGGCCT	<i>PaeBI</i>	CCCGGG
<i>Bbi24I</i>	ACGCGT	<i>PaeQI</i>	CCGCGG
<i>BfrI</i>	CTTAAG	<i>PceI</i>	AGGCCT
<i>BpuB5I</i>	CGTACG	<i>Pfl123II</i>	CGTACG
<i>BsiWI</i>	CGTACG	<i>Pme55I</i>	AGGCCT
<i>BspTI</i>	CTTAAG	<i>PpuAI</i>	CGTACG
<i>BsrBI</i>	GAGCGG	<i>PspAI</i>	CCCGGG
<i>Bst31NI</i>	CCGCTC	<i>PspALI</i>	CCCGGG
<i>Bst98I</i>	CTTAAG	<i>PspLI</i>	CGTACG
<i>BstD102I</i>	CCGCTC	<i>SacII</i>	CCGCGG
<i>BstPZ740I</i>	CTTAAG	<i>SarI</i>	AGGCCT
<i>BvuBI</i>	CGTACG	<i>SchZI</i>	CCGCGG
<i>Cfr42I</i>	CCGCGG	<i>SenPT14bI</i>	CCGCGG
<i>Cfr9I</i>	CCCGGG	<i>SexBI</i>	CCGCGG
<i>CfrJ4I</i>	CCCGGG	<i>SexCI</i>	CCGCGG
<i>CscI</i>	CCGCGG	<i>Sfr303I</i>	CCGCGG
<i>Eae46I</i>	CCGCGG	<i>SgrBI</i>	CCGCGG
<i>EaeAI</i>	CCCGGG	<i>SmaI</i>	CCCGGG
<i>EclRI</i>	CCCGGG	<i>SplI</i>	CGTACG
<i>Eco147I</i>	AGGCCT	<i>SpuI</i>	CCGCGG
<i>Eco29kI</i>	CCGCGG	<i>Sru30DI</i>	AGGCCT
<i>Esp4I</i>	CTTAAG	<i>SseBI</i>	AGGCCT
<i>GalI</i>	CCGCGG	<i>Ssp5230I</i>	GACGTC

<i>Gce</i> GLI	CCGCGG	<i>Sst</i> II	CCGCGG
<i>Gce</i> I	CCGCGG	<i>Ste</i> I	AGGCCT
<i>Gdi</i> I	AGGCCT	<i>Stu</i> I	AGGCCT
<i>Kpn</i> 378I	CCGCGG	<i>Sun</i> I	CGTACG
<i>Ksp</i> I	CCGCGG	<i>Vha</i> 464I	CTTAAG
<i>Mae</i> K81I	CGTACG	<i>Xcy</i> I	CCCGGG
<i>Mbi</i> I	CCGCTC	<i>Xma</i> CI	CCCGGG
<i>Mlu</i> I	ACGCGT	<i>Xma</i> I	CCCGGG
<i>Msp</i> CI	CTTAAG	<i>Zra</i> I	GACGTC
<i>Ngo</i> AIII	CCGCGG		

## EXEMPLO 2. Identificação e isolamento de seqüências teloméricas

Qualquer região telomérica funcional, nativa, clonada,  
 5 ou sintética, compreendendo uma repetição telomérica pode  
 ser usada para fazer construções de DNA. Diversas  
 repetições teloméricas são conhecidas, incluindo aquelas de  
*Tetrahymena*, *Paramecium*, *Oxytricha*, *Euplotes*,  
*Dictyostelium*, *Saccharomyces*, *Caenorhabditis*, *Trypanosoma*,  
 10 *Leishmania*, *Physarum*, *Arabidopsis*, humanos, e ratos.

### Repetição telomérica

### Organismos exemplares

CCCCAA (C <sub>4</sub> A <sub>2</sub> )	<i>Tetrahymena</i> , <i>Paramecium</i>
CCCCAAA (C <sub>4</sub> A <sub>4</sub> )	<i>Oxytricha</i> , <i>Euplotes</i>
15 CCCTA (C <sub>3</sub> TA)	<i>Trypanosoma</i> , <i>Leishmania</i> , <i>Physarum</i>
C <sub>1-3</sub> A	<i>Saccharomyces</i>
C <sub>1-8</sub> T	<i>Dictyostelium</i>
CCCTAAA (C <sub>3</sub> TA <sub>3</sub> )	<i>Arabidopsis</i> , humanos, ratos, <i>Caenorhabditis</i>

20

### A. Seqüências teloméricas sintéticas

A natureza altamente conservada, repetitiva das  
 seqüências teloméricas permite a síntese química e/ou

amplificação de PCR de longas regiões teloméricas disponíveis para construção do vetor. Longos tratos de repetições teloméricas, pex., (CCCTAAA)<sub>n</sub> para flanquear extremidades de minicromossomos podem ser geradas.

5

Longos trecho de repetições teloméricas em tandem podem ser produzidos por diversos ciclos de amplificação PCR usando pares iniciadores das SEQ ID NOs: 5 e 6 por mútuo preparo de dois oligonucleotídeos teloméricos complementares e seus produtos. Uma reação PCR usando uma baixa concentração de iniciadores (< 0.1µM) pode produzir segmentos de DNA de cerca de 100-10000 bp. Preferencialmente, repetições teloméricas sintéticas podem ser produzidas por ligação dos oligos fosforilados. Segmentos de DNA teloméricos foram clonados e usados para produzir construções de DNA.

## B. Identificação e isolamento de seqüências subteloméricas

### *i. Clones BAC contendo repetições teloméricas*

Clones BAC contendo regiões subteloméricas compreendendo repetições teloméricas podem ser usados para estabilizar extremidades cromossômicas de uma construção de minicromossomo. Um número de seqüências foram previamente identificadas como repetições subteloméricas (Burr et al. (1992) J Plant Cell 4:953-60). O banco de dados Genbank sequence pesquisado por palavra chave para seqüências teloméricas e subteloméricas. Seqüências selecionadas foram alinhadas e um elemento repetitivo comum identificado (Telo266, SEQ ID NO:189). Usando SEQ ID NO:189, diversos



oligonucleotídeos foram desenhados e usados como sonda para rastrear a biblioteca Mo17 BAC. Um número de BACs foi recuperado, um foi selecionado (bacm.pk107.g1), marcado, e hibridizado para cromossomos paquítenos. Sequências do clone BAC foram encontradas em grupos de 6 dos 20 subtelômeros em cromossomos de milho. O inserto do BAC bacm.pk107.g1 foi subclonado e sequenciado. Análises de seqüência revelaram um elemento repetitivo comum (TR430, SEQ ID NO:190) que foi usado para desenhar sondas overgo (Tabela 1). Localização subtelomérica daquelas repetições foi confirmada por FISH para cromossomos paquítenos de milho Mo17 e B73. Usando as mesmas sondas, bibliotecas BAC genômicas de milho Mo17 foram rastreadas por hibridização de colônia.

15

Aproximadamente 71 clones BAC contendo blocos de repetições subteloméricas de milho foram confirmados como tendo a repetição subtelomérica TR430 (Tabela 7).

20 **TABELA 7**

bacm.pk155.e24	bacm.pk166.a12	bacm.pk173.m16	bacm.pk203.j15
bacm2.pk022.m14	bacm2.pk092.a9	bacm2.pk114.i4	bacm2.pk169.b21
bacm2.pk177.j18	bacm2.pk190.m10	bacm2.pk220.h7	bacm.pk001.k4
bacm.pk009.c19	bacm.pk024.j15	bacm.pk024.k8	bacm.pk036.g23
bacm.pk038.g6	bacm.pk061.i6	bacm.pk062.g4	bacm.pk064.f6
bacm.pk070.j17	bacm.pk071.c12	bacm.pk073.m7	bacm.pk082.m9
bacm.pk101.h5	bacm.pk107.g1	bacm.pk110.h10	bacm.pk112.b18
bacm.pk123.e21	bacm.pk125.n6	bacm.pk132.h6	bacm.pk141.p12
bacm.pk142.b15	bacm.pk146.l14	bacm.pk148.j17	bacm.pk154.a21
bacm.pk155.p12	bacm.pk157.d2	bacm.pk164.n4	bacm.pk165.n1
bacm.pk169.n16	bacm.pk171.d3	bacm.pk172.m20	bacm.pk172.n19
bacm.pk172.n16	bacm.pk173.e9	bacm.pk173.i12	bacm.pk174.g4
bacm.pk176.g2	bacm.pk184.e5	bacm.pk185.o19	bacm.pk189.a10
bacm.pk197.m23	bacm.pk198.f9	bacm.pk198.k3	bacm.pk200.c20
bacm.pk208.j1	bacm.pk213.f2	bacm.pk214.i17	bacm.pk214.k16

bacm.pk214.l11	bacm.pk214.m20	bacm2.pk007.d1	bacm2.pk034.k22
bacm2.pk043.g14	bacm2.pk043.j16	bacm2.pk073.o7	bacm2.pk102.o18
bacm2.pk108.a3	bacm2.pk117.h13	bacm2.pk160.l2	bacm.pk203.j15
bacm.pk155.e24	bacm.pk166.a12	baacm.pk173.m16	bacm2.pk169.b21
bacm2.pk022.m14	bacm2.pk092.a9	bacm2.pk114.i4	bacm.pk001.k4
bacm2.pk177.j18	bacm2.pk190.m10	bacm2.pk220.h7	bacm.pk036.g23
bacm.pk009.c19	bacm.pk024.j15	bacm.pk024.k8	

Impressão digital de restrição com *DpnI* e hibridização blot com sondas TR430, sonda (CCCTAAA)<sub>n</sub>, e sondas para repetições knob 180 bp apresentaram pelo menos 3 tipos de clones BAC subteloméricos. O primeiro tipo tem longos tratos de repetições relacionadas de TR430 mais longas do que 10-20 kb. O segundo dos clones BAC tem repetições TR430 relacionadas que têm um sítio de restrição dentro da unidade, onde a unidade pode ser 800 bp ou 900 bp. Alguns clones BAC continham ambas dessas duas repetições. O terceiro tipo de clones BAC tem unidade relatada TR430 bp ao redor de 500 bp. Alguns desses clones BAC também têm repetições teloméricas relacionadas (CCCTAAA)<sub>n</sub>. Repetições knob de 180 bp estão também presentes em 37 clones BAC subteloméricos sugerindo que as repetições de 180 bp podem ser uma parte de algumas regiões subteloméricas. Clones BAC representativos de cada tipo foram pegos para análises posteriores, experimentos readaptados, e experimentos transgênicos:

bacm.pk038.g6;                      bacm2.pk063.g24;  
 bacm.pk071.c12;                      bacm.pk112.b18;                      bacm.pk142.b15;  
 bacm.pk173.e9. Insetos BAC com fragmentos subteloméricos podem ser usados em construções de DNA para montagem de minicromossomos *in vitro*, ou montagem de uma célula vegetal.

ii. *Isolamento dos fragmentos de DNA teloméricos cromossômicos nativos*

Fragmentos teloméricos cromossômicos que retenham pelo  
5 menos uma característica genômica nativa, tais como padrão  
de metilação, foram purificados de DNA genômico de milho  
através de fracionamento de tamanho do DNA genômico de  
milho digerido com enzimas de restrição que tenham um curto  
sítio de reconhecimento de 4 bp ou menor. Sequência  
10 telomérica de milho nativo compreende centenas ou milhares  
de repetições em tandem de CCCTAAA em cada telômero, essas  
curtas repetições em tandem no telômero não têm nenhum  
sítio de reconhecimento para qualquer enzima de restrição  
conhecida. Quaisquer enzimas de restrição de pequeno corte  
15 que reconhece uma sequência de 2-4 bp pode ser usada,  
enquanto elas não tiverem especificidade para repetição em  
tandem telomérica canônica (CCCTAAA)<sub>n</sub>. Pequenos cortadores  
digerem a maioria do DNA genômico em pequenos fragmentos  
que podem ser separados do grande DNA telomérico. O uso de  
20 uma combinação de duas ou mais enzimas de restrição de  
corte pequeno pode eliminar outros fragmentos de DNA não  
teloméricos não fragmentados pela primeira enzima de  
restrição. Não existem enzimas de restrição conhecidas  
tendo um sítio de reconhecimento dentro da repetição  
25 telomérica em tandem canônica.

DNA genômico de milho de Mo17 foi digerido com a  
enzima de restrição *Sau3A*, a maioria do DNA genômico é  
reduzido a pequenos fragmentos bem abaixo de 1 kb, enquanto

a maioria dos fragmentos de DNA telomérico são maiores do que cerca de 15 kb como determinado pela hibridização blot. O comprimento total dos segmentos de DNA telomérico *Sau3A* por genoma haplóide é cerca de 400 kb, ou 0.02% do genoma haplóide de milho total. Aproximadamente 1 mg do DNA genômico de milho total produz aproximadamente 200 ng de fragmentos de DNA telomérico na fração restante não digerida. A fração de DNA telomérica genômica pode ser purificada do gel e usada para gerar construções de DNA para montagem de minicromossomos.

### **EXEMPLO 3. Origem de replicação**

As construções de DNA são readaptadas com segmentos de DNA carregando origens de replicação para permitir replicação própria da construção e/ou minicromossomo no núcleo de células vegetais transgênicas. Qualquer origem de replicação que funciona em uma célula vegetal pode ser usada. Origens de replicação disponíveis são conhecidas e incluem origens de replicação vegetais, e origens de replicação virais. Opcionalmente, se uma construção for mantido em uma célula hospedeira não vegetal pelo menos uma origem de replicação apropriada pode ser incluída na construção, por exemplo origens de replicação bacteriana e/ou de levedura.

25

#### A. Espaçador não transcrito do rDNA 18-26S

Uma origem de replicação eucariótica bem estabelecida é um espaçador não transcrito do rDNA 18-26S (Ivessa & Zakian (2002) Genes Dev 16:2459-2464) que é provavelmente

funcional em plantas (Hernandez et al. (1993) EMBO J 12:1475-85). As seqüências de DNA 18-26S rDNA NTS podem ser isoladas de uma variedade de diferentes espécies de plantas, tais como *Zea mays*, *Triticum aestivum*, *Avena*  
5 *sativa*, *Hordeum vulgare*, *Arabidopsis thaliana*, e/ou *Glycine max*. Essas seqüências são clonadas em construções como cópias dispersas únicas ou múltiplas. Cromossomos eucarióticos tipicamente têm múltiplas origens de replicação, por essa razão a inclusão de múltiplas origens  
10 de replicação nos construções de DNA pode ser útil. A menos que disposto de outra forma, a seqüência 18-26S rDNA NTS de milho é usada nos construções de DNA (Toloczyki & Feix (1986) Nucleic Acids Res 14:4969-86).

15 B. Iniciador de proteína (Rep) do "Wheat dwarf virus" (WDV)

Iniciador de proteína (Rep) do "Wheat dwarf virus" (WDV) e sua origem de replicação cognata podem ser usados para gerar construções de DNA para montagem de minicromossomos. O iniciador de proteína (Rep) do "Wheat  
20 dwarf virus" (WDV) e sua origem de replicação cognata podem ser usados para suportar replicação de construções de minicromossomos em células de milho. A origem de replicação WDV pode ser provida na construção de DNA (em cis), enquanto genes necessários para o iniciador da proteína Rep  
25 e proteína RepA estimuladora do ciclo celular podem ser providos através da co-transformação em construções de plasmídeo independentes (in trans) (Sanz-Burgos & Gutierrez (1998) Virology 243:119-129.).

**EXEMPLO 4. Polinucleotídeos e polipeptídeos que estimulam crescimento**

Polinucleotídeos e/ou polipeptídeos que realçam o  
5 crescimento celular através da promoção da divisão celular,  
entrada na fase S, estimulação da divisão celular e/ou  
crescimento em cultura, ou melhora da transformação podem  
ser providos antes, durante, ou depois da introdução de  
construções de DNA compreendendo seqüência centromérica de  
10 milho e/ou fragmento subtelomérico. Qualquer tal  
composição, ou combinação dos mesmos pode ser usada  
incluindo polinucleotídeos, polipeptídeos, e/ou outros  
fatores usando qualquer método de liberação adequado.

**15 A. Proteína A associada à replicação.**

Proteína A de replicação do "wheat dwarf virus" (WDV)  
pode ser provida para fortalecer crescimento celular e/ou  
recuperação de eventos transgênicos. Ambas as RepA que  
retém atividade de replicação e uma RepA modificada que não  
20 auxilia na replicação viral podem ser usadas. Por exemplo,  
um plasmídeo carregando promotor nos::RepA pode ser co-  
liberado em células vegetais com construções de DNA.  
Expressão transiente de RepA durante os primeiros três dias  
é esperada ser suficiente para estimular divisão celular e  
25 aumentar evento de recuperação (ver, por exemplo,  
W000/50614, aqui incorporada por referência).

**B. Ciclinas**

Proteínas ciclinas, envolvidas na modulação do ciclo

celular podem fortalecer crescimento celular e recuperação de eventos transgênicos. Por exemplo, ciclina D de milho pode estimular divisão celular e crescimento de calo na cultura e melhorar a transformação de milho. Expressão  
5 ectópica de E2F induziu proliferação celular em *Arabidopsis*, este efeito foi aumentado pela co-expressão de DPa (de Veylder et al. (2002) EMBO J 21:1360-1368). Muitos homólogos de ciclo celular, incluindo ciclina D, ciclina E, weel, Rb, Rbr3, E2F/DP, e semelhantes têm sido isolados de  
10 plantas (Patente americana US 6.518.487; WO99/61619; WO00/37645; WO02/074909; Xie et al. (1996) EMBO J 15:4900-4908; todas as quais são aqui incorporadas por referência), e podem ser introduzidos dentro de vetores para liberação em células vegetais.

15

### C. Wuschel

Genes que causam vias de desenvolvimento específico são úteis no fortalecimento do crescimento celular. Por exemplo, membros da família WOX, tais como wuschel (WUS)  
20 parecem estimular divisão celular em ambas as células expressando WUS e células adjacentes. Uma construção compreendendo um polinucleotídeo codificando um polipeptídeo WUS pode ser usado para estimular divisão celular através de co-transformação com os construções de  
25 DNA. Diversos homólogos WUS são conhecidos em plantas, tais como *Arabidopsis* e milho (pex., Mayer et al. (1998) Cell 95:805-815; WO01/0023575; e US2004/0166563, todas as quais são aqui incorporadas por referência), e podem ser usadas para melhorar o crescimento de células transformadas. Por

exemplo, uma construção compreendendo um gene WUS de milho foi construída:

PHP21139 ubi pro::ubi 5' UTR::ubi intron::WUS::pinII

5

#### D. Protéina 2 do desenvolvimento de óvulo

Outros genes de interesse incluem aqueles revelados à família AP2/ERF dos fatores de transcrição que são preferencialmente expressos em embriões em desenvolvimento e sementes, incluindo Proteína 2 do desenvolvimento de óvulo (ZmODP2) que é expresso precocemente em embriogênese de milho. Quando ectopicamente expresso, ODP2 pode estimular crescimento celular em uma variedade de tecidos, incluindo tecidos não embriônicos, que podem facilitar a recuperação de eventos transgênicos. Essa família de genes inclui baby boom (BBM, BNM3, ODP2) que tem sido mostrado induzir embriões somáticos ectópicos em plantas (Boutilier *et al.* (2002) *Plant Cell* 14:1737-1749). Homólogos BBM/ODP2 são conhecidos, incluindo homólogos de milho (WO00/75530, aqui incorporada por referência) e podem ser liberadas para células vegetais para melhorar crescimento celular. Por exemplo, uma construção compreendendo um gene de milho ODP2 foi construída:

25 PHP21875 ubi pro::ubi 5' UTR::ubi intron::ODP2::pinII

#### E. Knotted-1

Genes Homeobox, incluindo membros da família gênica knox, tais como KN1, KNAT1, e STM atuam na iniciação do



meristema e/ou manutenção em plantas (Jackson *et al.* (1994) Dev 120:405-413; Lincoln *et al.* (1994) Plant Cell 6:1859-1876; Venglat *et al.* (2002) Proc Natl Acad Sci USA 99:4730-4735). Muitos membros da família knox são conhecidos em  
5 plantas, incluindo homólogos de milho (Vollbrecht & Hake (1991) Nature 350:241-243; Kerstetter *et al.* (1994) Plant Cell 6:1877-1887; Serikawa *et al.* (1996) Plant Mol Biol 32:673-683) e podem ser usados para construir vetores para liberação em células vegetais.

10

#### F. Lec1

Genes de cotilédones frondoso, tais como Lec1 e Lec2, estão envolvidos na regulação da embriogênese e atividade transcricional em plantas (Meinke *et al.* (1994) Plant Cell  
15 6:1049-1064; Lotan *et al.* (1998) Cell 93:1195-1205; W000/28058; Stone *et al.* (2001) Proc Natl Acad Sci USA 98:11806-11811; Patente americana US 6.492.577, aqui incorporadas por referência). Muitos homólogos são conhecidos que podem ser usados para construir vetores para  
20 serem liberados dentro das células vegetais.

#### G. Combinação dos polinucleotídeos estimulando crescimento

Uma combinação de polinucleotídeos e/ou polipeptídeos que aumentam crescimento celular através da promoção da  
25 divisão celular, entrada na fase S, estímulo da divisão celular e/ou crescimento na cultura, ou melhora na transformação podem ser providos antes, durante, ou depois da introdução de construções de DNA compreendendo seqüências centroméricas de milho e/ou fragmentos

subteloméricos. Por exemplo, polinucleotídeos codificando um ODP2 (PHP21875) de milho e um WUS (PHI21139) de milho podem ser usados em experimentos de transformação com construções de DNA compreendendo regiões centroméricas e/ou subteloméricas de milho. Em geral ODP2 e/ou WUS em co-transformação de bombardeamento de partículas de embriões de milho imaturos, como descrito no Exemplo 6D, mostrou um aumento significativo na frequência de eventos transgênicos como determinado pelo fenótipo BAR<sup>R</sup> e expressão da proteína marcadora fluorescente (DsRed). Em média 1008 eventos/4800 embriões primários (21%) foram observados quando foram providos na mistura de transformação. Sem PHP21139 ou PHP21875, 8 eventos/706 embriões primários (~ 1%) foram observados. Análises posteriores de eventos transgênicos indicaram que as construções ODP2 e/ou WUS co-bombardeados não foram integradas dentro do genoma ou minicromossomos montados.

#### **EXEMPLO 5. Construção de vetor**

Vetores, circulares ou lineares, para liberação dentro de células vegetais usando qualquer protocolo de transformação padrão são construídos usando protocolos de biologia molecular padrões, ver e.g. Sambrook et al. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd Ed., Cold Spring Harbor Laboratory Vols. 1-3. Vetores para transformação de células vegetais são construídos através da combinação de elementos cromossômicos isolados, opcionalmente com outros polinucleotídeos de interesse, usando técnicas padrões. Os vetores incluem aqueles

desenhados para serem mantidos em um sistema hospedeiro conveniente tal como *E. coli*, *Agrobacterium*, ou levedura, bem como em células vegetais. Tipicamente, a construção compreende ainda um marcador de seleção e/ou rastreável que  
5 funciona em células vegetais para auxiliar na manutenção, identificação, e/ou seleção de células vegetais compreendendo a construção de minicromossomo. Além disso, a construção tipicamente compreende diversos sítios de restrição únicos onde os polinucleotídeos adicionais de  
10 interesse podem ser clonados. A construção pode também compreender sítios de recombinação sítio-específicos úteis para clonagem recombicional, e/ou para direcionamento tardio e/ou modificação do minicromossomo. Construções de DNA derivados dos clones BAC de milho compreendendo  
15 seqüências centroméricas para liberação direta ou transformação de planta mediada por *Agrobacterium* são descritos abaixo. Vários componentes podem ser fornecidos na construção do clone BAC, e/ou em construções de DNA separados em trans.

20

#### A. Marcadores

Uma variedade de marcadores podem ser usados para identificar células transformadas compreendendo as construções de DNA introduzidas. Marcadores visuais incluem  
25 proteínas fluorescentes, tais como AmCyan, ZsYellow, ou DsRed (ClonTech Laboratories, Inc., Mountain View, CA, USA). Marcadores selecionáveis incluem PAT, BAR, GAT, e semelhantes.

Um cassete de expressão, PHP 23715, para liberação em células vegetais compreendendo uma proteína fluorescente vermelha (DsRed2) e um marcador de seleção PAT foi construído compreendendo os seguintes componentes operacionalmente ligados:

ubi pro::ubi 5'UTR::ubi intron::DsRed2::moPAT::pinII

Uma construção de DNA, PHP 23714, compreendendo uma proteína fluorescente ciano (AmCyan) para liberar em células vegetais é construído compreendendo os seguintes componentes operacionalmente ligados:

ubi pro::ubi 5'UTR::ubi intron::AmCyan1::moPAT::pinII

#### B. Vetores *Agrobacterium*

Plasmídeos binários de *Agrobacterium* são feitos usando o sistema híbrido descrito por Komari et al. ((1996) Plant J 10:165-174). Derivados de pSB11 são construídos como construções de T-DNA intermediários contendo a configuração desejada entre as seqüências de borda de T-DNA. O plasmídeo pSB11 é obtido da Japan Tobacco Inc. (Tokyo, Japan). Construção de pSB11 de pSB21, e construção de pSB21 dos vetores iniciadores, é descrita por Komari et al. ((1996) Plant J 10:165-174). Descrição da integração do plasmídeo T-DNA dentro do plasmídeo superbinário pSB1 por recombinação homóloga pode ser encontrada em EP672752 A1. O plasmídeo pSB1 é também obtido da Japan Tobacco Inc. Esses plasmídeos são usados para transformação mediada por

*Agrobacterium* depois de fazer a co-integração em LBA4404. Células eletro-competentes da cepa LBA4404 de *Agrobacterium* abrigando o pSB1 são criadas usando o protocolo como descrito por Lin (1995) em Methods in Molecular Biology,  
5 ed. Nickoloff, J.A. (Humana Press, Totowa, NJ). Células e DNA são preparados por eletroporação através da mistura de 1 µl de DNA plasmidial (~100ng) com 20 ml de células competentes em uma cubeta de eletrodo de 0.15cm da Life Technologies (agora Whatman Biometra) (Whatman Biometra  
10 #11608-031). Eletroporação é realizada em um aparelho de eletroporação porador de células usando a unidade de controle de pulso (Whatman Biometra #11604-014) no ambiente de 330 µF com um impulsor de voltagem (Whatman Biometra #11612-017) fixado em 4kW. O sistema libera  
15 aproximadamente 1.8 kV para as células de *Agrobacterium*. Recombinação com sucesso é verificada através de análise de restrição do plasmídeo co-integrante seguindo isolamento e transformação de volta em células *E. coli* DH5α para amplificação.

20

### C. Montagem *In vitro* de construções de DNA através de ligação

Vetores de minicromossomo da construção de DNA linear são produzidos através de preparação dos componentes de  
25 fragmentos de DNA tais como seqüências centroméricas de milho, marcadores de seleção (DsRed2 e AmCyan), origens de replicação eucarióticas (ori), seqüências teloméricas (TEL), e gene conferindo resistência a bialophos (PAT) sob controle do promotor de ubiquitina (ubi). Em um exemplo,

vetores de minicromossomo linear foram feitos do clone BAC centromérico bacm.pk128.j21 que compreende repetições invertidas de arranjos em tandem de CentC flanqueando um elemento de repetição centromérica CRM1. Fragmentos de DNA foram gerados de bacm.pk128.j21 através de digestão com NotI e purificação em gel de agarose. O fragmento purificado compreendendo a região centromérica foi combinado com fragmentos digeridos com restrição específica compreendendo marcadores de seleção e origem de replicação:

ubi pro::ubi 5' UTR::ubi intron::DsRed::moPAT-18S-26S rDNA NTS (NotI/SpeI), e um segundo cassete de marcador de seleção: ubi pro::ubi 5' UTR::ubi intron::AmCyan::moPAT (NotI/SmaI), e seqüências teloméricas (SpeI/XhoI ou SmaI/KpnI) de seus construções. Fragmentos de DNA foram preparados de tal modo que cada fragmento compreendeu sítios de reconhecimento únicos para montagem *in vitro* de uma estrutura linear única durante a ligação. O vetor linearizado montado compreende:

TEL-(SpeI)-ubi pro::ubi 5' UTR:ubi intron::AmCyan::moPAT-(NotI)-bacm.pk128.j21-(NotI)-ori-ubi pro::ubi 5' UTR::ubi intron::DsRed::moPAT-(SmaI)-TEL

#### D. Construções de DNA - Vetores de clone BAC readaptado circular

Qualquer clone BAC centromérico e/ou subtelomérico ou clone de fragmento cromossômico pode ser readaptado com componentes adicionais para transformação de planta.

O sistema de construção de vetor transposon EPICENTRE

EZ::TN™ pMOD™-2 MCS (EpiCentre Madison, WI, USA) é usado para reconverter polinucleotídeos de interesse em clones BAC existentes. O pMOD-2 é um plasmídeo baseado em pUC com uma origem de replicação colE1 e um sítio múltiplo de clonagem (MCS) entre a extremidade em mosaico hiperativa de 5 19 bp (ME) reconhecida pela transposase EZ-Tn5. O transposon Tn5-2 integra randomicamente em cada DNA alvo, portanto cada reação de transposição gera uma pequena biblioteca de construções representando diferentes sítios de integração. 10 Preparações de DNA de clones readaptados individuais ou um grupo de clones podem ser usados para transformação de células vegetais.

*i. BACs Centroméricos*

15 Dois BACs representativos de apenas CentC, bacm.pk018.113 e bacm2.pk174.o21, foram selecionados baseados em sua digestão com enzimas de restrição e padrões de hibridização de Southern. Esses clones BAC foram readaptados usando o sistema de construção EPICENTRE 20 EZ::TN™ pMOD™-2 MCS para gerar construções de DNA circular para transformação de plantas e montagem de minicromossomo.

O MCS foi usado para inserir um fragmento de DNA compreendendo marcadores de seleção: ubi pro::ubi 5' 25 UTR::ubi intron::DsRed::moPAT com ou sem um 18-26S rDNA NTS ori de milho para produzir uma primeira versão de uma construção de transposon prático, Tn5-1s. Depois de clonar as seqüências de DNA de interesse, o transposon é gerado através da digestão com a enzima de restrição PshAI. Após a

integração os construções BAC são transformados em *E. coli*, clones positivos selecionados através da hibridização da colônia com as sondas de transposon, e o DNA isolado de clones positivos selecionados.

5

*ii. BACs subteloméricos*

Seis clones BAC representativos foram selecionados do BAC do grupo BAC subtelomérico: bacm.pk038.g06, bacm2.pk063.g24, bacm.pk071.c12, bacmpk112.b18, 10 bacm.pk142.b15, e bacm.pk173.e09. Novos construções de transposon personalizados Tn5-2 compreendendo 18-26S rDNA NTS ori-ubi pro::ubi 5' UTR::ubi intron::DsRed::moPAT, um gene Kan<sup>r</sup>, e sítios para três diferentes enzimas de restrição da origem: I-PpoI, I-CeuI, and PI-SceI, foram 15 construídos e usados para reconverter os clones BAC subteloméricos. As construções BAC readaptadas são transformadas em *E. coli* e selecionados em canamicina e cloranfenicol, DNA é isolado dos clones positivos selecionados.

20

E. Construções de DNA - Clones BAC readaptados linearizados

Adicionais construções de transposon customizados Tn5-3 foram gerados. Esses vetores Tn5-3 compreendem 18-26S rDNA NTS ori-ubi pro::ubi 5' UTR::ubi intron::DsRed::moPAT. 25 As construções também compreendem um gene Kan<sup>r</sup> flanqueados por dois segmentos de DNA em orientação invertida cada uma composta de dois sítios de reconhecimento para as enzimas de restrição de origem I-CeuI e PI-SceI, e seqüência telomérica compreendendo arranjos de repetições



teloméricas. Depois da clonagem das seqüências de DNA de interesse, o transposon é gerado pela digestão com a enzima de restrição *PshAI*. Após a integração as construções BAC são transformadas em *E. coli* e selecionados canamicina e cloranfenicol, DNA é isolado dos clones positivos selecionados. DNA BAC readaptado recombinante é digerido in vitro com enzimas de restrição de origem (I-CeuI ou PI-SceI) convertendo o DNA circular em uma construção de DNA linear flanqueado com seqüências teloméricas na orientação correta, e removendo o fragmento compreendendo Kan<sup>r</sup> (Figura 13).

Três tipos de clones BAC centroméricos foram readaptados com este vetor Tn5-3:

15

1. Clone BAC centromérico com blocos invertidos de repetições centroméricas CentC flanqueando um elemento centromérico CRM1 bacm.pk128.j21, não seqüências CentA ou CRM2;

20

2. Clones BAC centroméricos pertencendo ao grupo principal dos clones BAC centroméricos contendo todas as quatro repetições específicas de centrômero CentA, CentC, CRM1, e CRM2 (Tabela 8); e,

25

3. Clones BAC centroméricos do cromossomo 4 de milho (Tabela 9).

Amostras de DNA de cada clone BAC são fracionadas em um gel de agarose e a banda contendo a construção BAC readaptado linear foi excisada. DNA é eletroeluído da

agarose e usado para transformação biolística de embriões Hi-II imaturos 8-11 DAP (dias depois da polinização). Opcionalmente, essas construções podem ser usadas para microinjeção do DNA, ou convertido em vetores para transformação mediada por *Agrobacterium*.

TABELA 8

Grupo 1	Grupo 2	Grupo 3	grupo 4
bacm.pk007.a2	bacm.pk011.l8	bacm.pk001.n1	bacm.pk109.h24
bacm.pk036.e13	bacm.pk012.n20	bacm.pk023.i5	bacm.pk039.a3
bacm.pk066.j14	bacm.pk013.m8	bacm.pk043.o23	bacm.pk039.m16
bacm.pk075.l6	bacm.pk062.c14	bacm.pk051.g11	bacm.pk041.e16
bacm.pk076.m3	bacm.pk064.n1	bacm.pk056.j19	bacm.pk077.b21
bacm.pk119.a23	bacm.pk068.p16	bacm.pk076.o15	bacm.pk079.m11
bacm.pk133.b10a	bacm.pk070.h17	bacm.pk087.m4	bacm.pk085.k5
bacm.pk133.b10b	bacm.pk090.o5	bacm.pk089.l8	bacm.pk098.h2
bacm.pk133.b11	bacm.pk098.f3	bacm.pk093.d8	bacm.pk102.i4
bacm.pk135.i6	bacm.pk135.l7	bacm.pk106.j20	bacm.pk112.p1
bacm.pk178.c10	bacm2.pk002.g7	bacm.pk129.a4	bacm.pk124.j24
bacm2.pk023.e24	bacm2.pk003.g6	bacm.pk134.f15	bacm.pk143.m18
bacm2.pk064.e15	bacm2.pk012.g19	bacm.pk135.j2	bacm.pk148.e2
bacm2.pk066.m12	bacm2.pk013.c9	bacm.pk138.e14	bacm.pk156.i17
bacm2.pk083.a2	bacm2.pk034.g20	bacm.pk141.j4	bacm.pk164.b11
bacm2.pk093.h11	bacm2.pk053.g23	bacm.pk161.h1	bacm.pk166.n7
bacm2.pk099.m24	bacm2.pk070.g7	bacm.pk164.e18	bacm.pk178.o20
bacm2.pk116.g16	bacm2.pk094.f14	bacm.pk179.d4	bacm2.pk034.j8
bacm2.pk174.e4	bacm2.pk096.d23	bacm2.pk130.e20	bacm2.pk075.n6
bacm2.pk179.b18	bacm2.pk100.j24	bacm2.pk137.f2	bacm2.pk115.o22
bacm2.pk179.e1	bacm2.pk179.o14	bacm2.pk158.f12	bacm2.pk169.a21

TABELA 9

Grupo B73 específico do cromossomo 4	Grupo Mol7 específico do cromossomo 4
Baccpk0143i9	bacm.pk010m7
Bacbpk0155h15	bacm.pk108h15
Bacbpk0424d20	bacm.pk184c21

10

F. Construções de DNA - Vetores de combinação BAC múltiplo recombinados

Clones BAC centroméricos pertencendo ao grupo principal dos clones BAC centroméricos contendo todas as quatro repetições específicas de centrômero CentA, CentC, CRM1, e CRM2 (Tabela 8) foram também recombinadas com o  
5 vetor Tn5-2. Construções Tn5-2 compreendem ori-ubi pro::ubi  
5' UTR::ubi intron::DsRed2::moPAT, um gene Kan<sup>r</sup>, e sítios  
para três enzimas de restrição de origem: I-PpoI, I-CeuI,  
and PI-SceI. Os BACs readaptados foram cortados com enzimas  
de restrição de origem I-CeuI e PI-SceI, separados por gel  
10 de eletroforese de campo pulsado (PFGE) sob condições  
padrões: 1% agarose, 1X TAE, pulso inicial de 5 seg, pulso  
final de 10 seg, tempo total de corrida 12 hs à 12°C.  
Grandes fragmentos foram purificados, e sujeitos a ligação  
para formar construções de DNA multiméricos de até 1 Mb  
15 longos.

#### **EXEMPLO 6: Transformação de planta**

Qualquer método de transformação de planta disponível  
pode ser usado. Similarmente qualquer célula vegetal e/ou  
20 tecido que possa ser transformado, cultivado, e/ou  
regenerado em uma planta pode ser usado. Essas células e  
tecidos vegetais, bem como meio de cultura e condições,  
métodos de transformação disponíveis, e meio de regeneração  
e condições são bem conhecidos.

25

##### A. Tipos celulares

Uma variedade de tipos celulares foram avaliados para  
o seu potencial como alvo para geração de minicromosomo e  
liberação da construção, incluindo células em suspensão

Black Mexican Sweet (BMS), células meristemáticas, o zigoto, células escutelares no embrião imaturo, células nos embriões somáticos cultivados, célula central e células do endosperma precoce. Métodos estão disponíveis para produzir  
5 embriões haplóides através do cruzamento de um dado genótipo com a linhagem RWS, ou outra linhagem indutora. Embriões imaturos haplóides podem ser um bom alvo para a liberação do minicromossomo, tanto em células escutelares de 10-12 dias depois da polinização (DAP) ou dentro de  
10 meristemas apicais expostos dos embriões de estágio coleóptilo (7-8 DAP). Comparações importantes no comportamento de minicromossomos introduzidos em um ambiente diplóide ou haplóide podem ser executadas, além disso, se a introdução do minicromossomo é seguida por  
15 duplicação cromossômica quimicamente induzida (pex. colchicina, ou óxido nítrico), esses embriões duplo-haplóides podem ser rapidamente regenerados para produzir uma linhagem pura contendo minicromossomo. Todos os tipos celulares haplóides e diplóides acima mencionados podem ser  
20 convertidos em cultura de suspensão e/ou protoplastos, ou culturas em suspensão estabelecidas, tais como BMS são adequados e podem ser usados para transformação. Células em suspensão e/ou protoplastos podem prover fácil acessibilidade e claridade ótica para monitoramento  
25 microscópico após a liberação da construção de DNA. Qualquer método adequado para liberação da construção para a cultura de protoplasto de planta pode ser usado, incluindo métodos padrões de liberação direta de eletroporação e mediado por PEG, ver pex., C. 8, pp. 189-

253 in *Advances in Cellular and Molecular Biology of Plants*, Vol. 5, Ed. Vasil, Kluwer Acad Publ (Dordrecht, The Netherlands) 1999.

#### 5 B. Microinjeção de milho

Qualquer método adequado para microinjeção de células vegetais, tecidos, e/ou embriões pode ser usado. Além disso, qualquer composição ou combinação/mistura de composições pode ser injetada, incluindo polinucleotídeos, polipeptídeos, cofatores, químicos, adjuvantes, e semelhantes. Liberação direta em um zigoto provê uma oportunidade para produzir uma planta transgênica sem os passos intermediários da cultura de tecido e regeneração. Por exemplo, microinjeção de milho pode ser feita essencialmente como descrita na patente americana US 6.300.543. Brevemente, óvulos de milho imaturos são seccionados para produzir placas nucelares compreendendo o saco embrionário, que é direcionado para liberação de microinjeção da composição de transformação. Seguindo a microinjeção, os sacos embrionários são cultivados em meio apropriado para propagação e regeneração da planta.

#### C. Transformação mediada por *Agrobacterium*

Transformação de milho mediada por *Agrobacterium* é realizada essencialmente como descrita por Zhao (WO98/32326). Brevemente, embriões imaturos são isolados de óvulos de milho e os embriões contactados com uma suspensão de *Agrobacterium* contendo um T-DNA, onde as bactérias são capazes de transferir a construção de DNA para pelo menos

uma célula de pelo menos um dos embriões imaturos. Opcionalmente, o tecido alvo pode ser co-transformado com múltiplas linhagens de *Agrobacterium* compreendendo T-DNAs com diferentes construções de DNA e/ou polinucleotídeos de interesse.

5  
10  
Passo 1: Passo de infecção. Embriões imaturos são imersos em uma suspensão de *Agrobacterium* para iniciação da inoculação.

Passo 2: Passo de co-cultivo. Os embriões são co-cultivados por um tempo com o *Agrobacterium*.

15  
Passo 3: Passo de descanso. Opcionalmente, seguindo o co-cultivo, um passo de descanso pode ser executado. Os embriões imaturos são cultivados em meio sólido com antibiótico, mas sem um agente de seleção, para eliminação do *Agrobacterium* e para uma fase de descanso para as células infectadas.

20  
25  
Passo 4: Passo de seleção. Embriões inoculados são cultivados em meio contendo um agente de seleção e calos transformados desenvolvidos são recuperados. Os embriões imaturos são cultivados em meio sólido com um agente de seleção resultando no crescimento seletivo das células transformadas.

Passo 5: Passo de regeneração. Calos crescidos em meio seletivo são cultivados em meio sólido para regenerar

plantas.

D. Bombardeamento de partículas de milho

Embriões imaturos de milho são bombardeados com uma construção de DNA circular ou linear compreendendo uma  
5 seqüência centromérica de milho isolada, e opcionalmente  
regiões subteloméricas, origens de replicação, sítios de  
recombinação docking, polipeptídeos, e/ou marcadores, por  
exemplo um gene marcador de seleção tal como PAT (Wohlleben  
*et al.* (1988) Gene 70:25-37) que confere resistência ao  
10 herbicida Bialaphos, ou outro marcador de seleção adequado  
ou marcadores rastreáveis, tais como RFP e/ou CFP. A  
construção pode também compreender outros genes marcadores,  
ou ser co-transformados com construções de polinucleotídeos  
adicionais compreendendo marcadores. Transformação é  
15 executada essencialmente como segue.

Espigas de milho imaturas de 8-11 DAP são esterilizadas na superfície em uma solução de alvejante 30%  
mais 0.5% de micro detergente por 20 minutos, e lavadas  
20 duas vezes com água estéril. Os embriões imaturos são  
excisados, colocados com o lado do eixo embrionário para  
baixo (lado do escutelo para cima), 50 embriões por placa,  
em 560L de meio por 1-3 dias a 26°C no escuro. Antes da  
transformação os embriões imaturos são transferidos para um  
25 meio 560Y por 4 horas, e então alinhados dentro de uma zona  
alvo de 2.5-cm na preparação para o bombardeamento.

O DNA é precipitado em péletes de ouro de 0.6 µm  
(diâmetro médio) usando um lipídeo catiônico solúvel em

água Tfx™-50 (Cat# E1811, Promega, Madison, WI, USA) como segue: preparar uma solução de DNA em gelo usando 1 µg da construção de DNA centromérico de milho (10 µl); opcionalmente outras construções para co-bombardeamento tais como 50 ng (0.5 µl) de PHP21875 (BBM), e 50 ng de (0.5 µl) PHP21139 (WUS); solução de DNA misturado. Para o DNA pré-misturado adicionar 20 µl de partículas de ouro preparadas (15 mg/ml) em água; 10 µl de Tfx-50 em água; misturar cuidadosamente. Isso pode ser estocado em gelo durante a preparação de macrocarregadores, tipicamente por cerca de 10 min. Partículas de ouro são peletizadas em uma microcentrífuga à 10,000rpm por 1 min, remover o sobrenadante. Cuidadosamente lavar o pélete com 100 ml de EtOH 100% sem ressuspender o pélete, cuidadosamente remover o EtOH por lavagem. Adicionar 20 µl de EtOH 100% e cuidadosamente ressuspender as partículas por sonicação breve, 10 µl gotejados no centro de cada macrocarregador e permitido secar por cerca de 2 minutos antes do bombardeamento.

20

As placas de amostras de embriões alvos de milho são bombardeadas duas vezes por placa usando aproximadamente 0.5 µg de DNA por tiro usando o aparelho Bio-Rad PDS-1000/He (Bio-Rad Laboratories, Hercules, CA) com uma pressão de ruptura de 450 PSI, uma pressão de vácuo de 27-28 polegadas de Hg, e uma distância de voo (flight) da partícula de 8.5 cm.

Após o bombardeamento, os embriões são transferidos



para o meio sólido 560P mantido no escuro à 26°C por 4-6 dias, então transferidos para o meio de seleção 560R contendo 3 mg/L de Bialaphos, e subcultivado a cada 2 semanas. Depois de aproximadamente 10 semanas de seleção, clones de calos resistentes à seleção são transferidos para o meio 288J para iniciar a regeneração da planta. Após a maturação do embrião somático (2-4 semanas), embriões somáticos bem desenvolvidos são transferidos para o meio 272V para germinação e transferidos para um quarto de cultivo iluminado. Aproximadamente 7-10 dias mais tarde, plântulas em desenvolvimento são transferidas para um meio 272V livre de hormônio em tubos por 7-10 dias até as plântulas estarem bem estabelecidas. Plantas são então transferidas para inserções em frascos (equivalentes a potes de 2.5") contendo solo envasado e deixadas crescer por 1 semana em uma câmara de crescimento, subsequentemente crescidas em um adicional de 1-2 semanas na casa de vegetação, então transferidas para 600 potes clássicos (1.6 galão) e crescidas até a maturidade.

20

#### E. Bombardeamento de partícula de soja

Um polinucleotídeo, uma mistura de polinucleotídeos, e opcionalmente, polipeptídeos, podem ser introduzidos em culturas de suspensão embriogênica de soja através de bombardeamento de partícula usando essencialmente os métodos descritos em Parrot et al. (1989) Plant Cell Rep 7:615-617. Este método, com modificações, é descrito abaixo.

Semente é removida de vagens imaturas e cotilédones menores do que 4mm em comprimento são selecionados. As sementes são esterilizadas por 15 minutos em uma solução alvejante 0.5% v/v e então lavadas com água destilada  
5 estéril. Os cotilédones imaturos são excisados através de primeiro suprimir a porção da semente que contenha o eixo do embrião. Os cotilédones são então removidos do tegumento empurrando gentilmente a extremidade distal da semente com a extremidade brusca da lâmina de bisturi. Os cotilédones  
10 são então colocados em placas de petri (lado para cima) com meio de iniciação SB1. As placas de petri são incubadas na luz (16 hr dia; 75-80  $\mu$ E) à 26°C. Depois de 4 semanas de incubação os cotilédones são transferidos para meio SB1 fresco. Depois de um adicional de duas semanas, embriões  
15 somáticos no estágio globular que exibem áreas proliferativas são excisados e transferidos para meio líquido FN Lite (Samoylov et al. (1998) In Vitro Cell Dev Biol Plant 34:8-13). Cerca de 10 a 12 pequenos grupos de embriões somáticos são colocados em frascos de 250 ml  
20 contendo 35 ml de meio SB172. As culturas em suspensão embriogênicas de soja são mantidas em meio líquido de 35 mL em um agitador rotativo, 150 rpm, à 26°C com luz fluorescente (20  $\mu$ E) em um cronograma de 16:8 hora dia/noite. Culturas são sub-cultivadas a cada duas semanas  
25 através da inoculação de aproximadamente 35 mg de tecido em 35 mL de meio líquido.

Culturas em suspensão embriogênicas de soja são então transformadas usando bombardeamento de partícula (Klein et

al. (1987) Nature 327:70; patente americana US 4.945.050). Um instrumento de biolística da BioRad PDS1000/HE pode ser usado para essas transformações. Um gene marcador de seleção pode ser usado para facilitar a transformação de soja, por exemplo, um cassete de expressão pode ser usado 5 compreendendo o promotor 35S do vírus mosaico da couve-flor (Odell et al. (1985) Nature 313:810-812), o gene higromicina fosfotransferase do plasmídeo pJR225 (de *E. coli*; Gritz et al. (1983) Gene 25:179-188) e a região 3' do 10 gene da nopalina sintetase do T-DNA do plasmídeo Ti de *Agrobacterium tumefaciens*.

Para 50 µL de uma suspensão de particular de ouro de 60 mg/mL 1 µm é adicionado (em ordem): 5 µL de DNA (1 15 mg/mL), 20 µl de espermidina (0.1 M), e 50µmL de CaCl<sub>2</sub> (2.5 M). A preparação de partícula é agitada por três minutos, centrifugada em uma microcentrífuga por 10 segundos e o sobrenadante removido. As partículas cobertas com DNA são lavadas uma vez em 400 µL de etanol 70% e então 20 ressuspendidas em 40 µL de etanol anidro. A suspensão DNA/partícula é sonicada três vezes por um segundo cada. Cinco µL das partículas de ouro cobertas com DNA são então carregadas em cada disco macrocarregador.

25 Aproximadamente 300-400 mg de uma cultura em suspensão de duas semanas de idade são colocados em uma placa de petri vazia de 60x15 mm e o líquido residual removido do tecido com uma pipeta. Pressão de ruptura de membrana é fixada em 1100 psi e a câmara é evacuada para um vácuo de

28 polegadas de mercúrio. O tecido é colocado  
aproximadamente 8 cm distante da tela de retenção, e é  
bombardeado três vezes. Após o bombardeamento, o tecido é  
dividido na metade e colocado em dois frascos separados com  
5 35 ml de meio FN Lite por frasco.

Cinco a sete dias depois do bombardeamento, o meio  
líquido é trocado com meio fresco. Onze dias após o  
bombardeamento o meio é trocado com meio fresco contendo  
10 higromicina 50 mg/mL. Esse meio seletivo é substituído  
semanalmente. Sete a oito semanas após o bombardeamento,  
tecido verde transformado será observado crescendo dos  
grupos embriogênicos necróticos não transformados. Tecido  
verde isolado é removido e inoculado em frascos individuais  
15 para gerar novas, clonalmente propagadas, transformadas  
culturas de suspensão embriogênicas. Cada nova linhagem é  
tratada como um evento de transformação independente. Essas  
suspensões são então subcultivadas e mantidas como grupos  
de embriões imaturos, ou tecido é regenerado em plantas  
20 inteiras através de maturação e germinação dos embriões  
individuais.

Para regeneração, eventos são removidos da cultura  
líquida e o processo de maturação é iniciado em meio  
25 sólido. Grupos embriogênicos são removidos do líquido  
SB196, colocados em papel de filtro estéril, e colocado em  
meio Agar sólido SB166 por 1-2 semanas. Grupos de tecidos  
são quebrados ou gentilmente esmagados com colher. Cerca de  
10-20 grupos de tecido de cerca de 4-5 mm de diâmetro são

subcultivados por três semanas em meio SB103 ou SB148, para gerar embriões. Embriões são cultivados por 4-6 semanas a 26°C sob luz fluorescente branca fria e bulbos Agro (40 watt) em um fotoperíodo de 16:8 hr com intensidade luminosa de 90-120  $\mu\text{E}/\text{m}^2\text{s}$ . Após 4-6 semanas de maturação, embriões individuais são desidratados através da colocação em uma grande placa de petri estéril (60 x 25 mm) selada com fita de fibra, ou colocados em uma caixa plástica (com nenhuma fita de fibra) por 4-7 dias. Embriões desidratados são plantados em meio sólido SB71-4 em cada vaso de cultura circular ventilado (RCV) ou em placa de petri de 100x25 mm, e germinados a 26°C sob luz fluorescente branca fria e bulbo Agro (40 watt) em um fotoperíodo de 16:8 hr com intensidade luminosa de 90-120  $\mu\text{E}/\text{m}^2\text{s}$  para produzir plântulas. Plântulas são colocadas em bandejas celulares e colocadas em um incubador em condições de 16 hr de fotoperíodo, 26°C/24°C de temperatura dia/noite por cerca de 2 semanas antes do transplante para solo para produção de semente.

20

#### F. Meio de cultura de célula vegetal

Meio 560L compreende 4.0 g/L de sais basais N6 (Sigma C-1416), 1.0 ml/L de mistura de vitaminas Eriksson's (1000X Sigma 1511), 0.5 mg/L de tiamina HCl, 20 g/L de sacarose, 1.0 mg/L de 2,4-D, e 2.88 g/L de L-prolina (trazida para o volume com D-I H<sub>2</sub>O seguindo ajuste para pH 5.8 com KOH); 2.0 g/L Gelrite® (adicionado após trazer para o volume com D-I H<sub>2</sub>O); e 8.5 mg/L de nitrato de prata (adicionado após esterilização do meio e esfriamento para temperatura

ambiente).

Meio 560P compreende 4.0 g/L de sais basais N6 (Sigma C-1416), 1.0 ml/L de mistura de vitaminas Eriksson's (1000X Sigma 1511), 0.5 mg/L de tiamina HCl, 30 g/L de sacarose,  
5 2.0 mg/L de 2,4-D, e 0.69 g/L de L-prolina (trazida para o volume com D-I H<sub>2</sub>O seguindo ajuste para pH 5.8 com KOH);  
3.0 g/L Gelrite® (adicionado após trazer para o volume com D-I H<sub>2</sub>O); e 0.85 mg/L de nitrato de prata (adicionado após  
esterilização do meio e esfriamento para temperatura  
10 ambiente).

Meio 560Y compreende 4.0 g/L de sais basais N6 (Sigma C-1416), 1.0 ml/L de mistura de vitaminas Eriksson's (1000X Sigma 1511), 0.5 mg/L de tiamina HCl, 120 g/L de sacarose,  
15 1.0 mg/L de 2,4-D, e 2.88 g/L de L-prolina (trazida para o volume com D-I H<sub>2</sub>O seguindo ajuste para pH 5.8 com KOH);  
2.0 g/L Gelrite® (adicionado após trazer para o volume com D-I H<sub>2</sub>O); e 8.5 mg/L de nitrato de prata (adicionado após  
esterilização do meio e esfriamento para temperatura  
20 ambiente).

Meio 560R compreende 4.0 g/L de sais basais N6 (Sigma C-1416), 1.0 ml/L de mistura de vitaminas Eriksson's (1000X Sigma 1511), 0.5 mg/L de tiamina HCl, 30.0 g/L de sacarose,  
25 e 2.0 mg/L de 2,4-D (trazido para o volume com D-I H<sub>2</sub>O seguindo ajuste para pH 5.8 com KOH); 3.0 g/L Gelrite (adicionado após trazer para o volume com D-I H<sub>2</sub>O); e 0.85 mg/L de nitrato de prata e 3.0 mg/L de bialaphos (ambos adicionados após esterilização do meio e esfriamento para

temperatura ambiente).

Meio 288J compreende: 4.3 g/L de sais MS (Gibco 11117-074), 5.0 ml/L de solução estoque de vitaminas MS (0.100 g/L de ácido nicotínico, 0.02 g/L de tiamina HCl, 0.10 g/L de piridoxina HCl, e 0.40 g/L de glicina trazida para o volume com D-I H<sub>2</sub>O) (Murashige & Skoog (1962) *Physiol Plant* 15:473), 100 mg/L de mio-inositol, 0.5 mg/L de zeatina, 60 g/L de sacarose, e 1.0 ml/L de ácido abscísico 0.1 mM (trazido para o volume com D-I H<sub>2</sub>O após ajustar para pH 5.6); 3.0 g/L Gelrite (adicionado após trazer para o volume com D-I H<sub>2</sub>O); e 1.0 mg/L de ácido indol acético e 3.0 mg/L de bialaphos (adicionado após esterilizar o meio e esfriar para 60°C).

Meio 272V compreende: 4.3 g/L de sais MS (Gibco 11117-074), 5.0 ml/L de solução estoque de vitaminas MS (0.100 g/L de ácido nicotínico, 0.02 g/L de tiamina HCl, 0.10 g/L de piridoxina HCl, e 0.40 g/L de glicina trazida para o volume com D-I H<sub>2</sub>O), 0.1 g/L de mio-inositol, e 40.0 g/L de sacarose (trazida para o volume com D-I H<sub>2</sub>O após ajustar pH para 5.6); e 6 g/L de bacto-agar (adicionado após ajustar volume com D-I H<sub>2</sub>O), esterilizar e esfriar para 60°C.

Meio SB1 compreende sais MS (Gibco/ BRL - Cat# 11117-066, 1 pk/L), estoque de vitaminas B5 1ml/L, 20 mg/L de 2,4-D, 31.5 g/L de sacarose, 8 g/L de TC Agar, pH 5.8

Estoque de vitaminas B5 1000X compreende 10 g de mio-inositol, 100 mg de ácido nicotínico, 100 mg de piridoxina

HCl, 1 g de tiamina, D-I H<sub>2</sub>O para 100 ml, alicotar e estocar a -20°C.

G. Isolamento de DNA de calos e tecido foliar

5 Putativos eventos de transformação podem ser rastreados para a presença do transgene. DNA genômico pode ser extraído de caules, folhas, ou outro tecido usando separação de núcleo vegetal, lise, e purificação HMW, ou  
10 alternativamente usando uma modificação do CTAB (*brometo de cetiltriethylamônio*, Sigma H5882) método descrito por Stacey & Isaac (1994 In Methods in Molecular Biology Vol. 28, pp. 9-15, Ed. P.G. Isaac, Humana Press, Totowa, NJ). Aproximadamente 100-200 mg de tecido congelado é macerado em pó em nitrogênio líquido e homogeneizado em 1 ml de  
15 tampão de extração CTAB (2% CTAB, 0.02 M EDTA, 0.1 M TrisHCl pH 8, 1.4 M NaCl, 25 mM DTT) por 30 min a 65°C. Amostras homogeneizadas são permitidas esfriar em temperatura ambiente por 15 min antes de uma simples extração de proteína com aproximadamente 1 ml 24:1 v/v  
20 clorofórmio:octanol é feita. Amostras são centrifugadas por 7 min a 13,000 rpm e a camada superior do sobrenadante coletada usando ponteiras de pipeta de grande largura. DNA é precipitado do sobrenadante através de incubação em etanol 95% em gelo por 1 h. Fibras de DNA são enroladas em  
25 um gancho de vidro, lavadas em etanol 75% contendo acetato de sódio 0.2 M por 10 min, secas ao ar por 5 min e ressuspensas em tampão TE. Cinco µl de RNase A é adicionada para as amostras e incubadas a 37°C por 1 h. Para quantificação do DNA genômico, gel de eletroforese é



realizado usando um gel de agarose 0.8% em tampão 1x TBE. Um microlitro de cada uma das amostras é fracionado ao lado de 200, 400, 600 e 800 ng  $\mu\text{l}^{-1}$   $\lambda$  marcadores de DNA intactos.

5

#### **EXEMPLO 7. Resultados da Transformação**

Embriões de milho imaturos Hi-II em 8-11 DAP foram transformados por bombardeamento de partículas essencialmente como descrito no Exemplo 6D. Junto com 10 construções de DNA BAC readaptados, embriões foram co-transformados com ODP2, WUS, e/ou vetores ODP2+WUS. Nas duas semanas pós-bombardeamento, células transformadas tinham proliferado para formar um calo embriogênico com múltiplos embriões somáticos. Alguns desses embriões 15 somáticos expressaram o gene marcador fluorescente DsRed2 indicativo de herança estável. Embriões somáticos individuais foram excisados e propagados como eventos independentes no meio de seleção Bialophos. Cultura de calos clonalmente propagados foi estabelecida de cada 20 evento.

Um primeiro rastreamento de cada evento de transformação foi feito usando FISH. Embriões somáticos foram usados para fazer distribuição cromossômica para FISH 25 como descrito no Exemplo 8. Cada evento foi caracterizado usando sondas FISH separadas para o marcador mo-PAT/DsRed2 (PHP23715), e as repetições centroméricas CentC em tandem para detectar do marcador transgênico e herança de seqüências de DNA centroméricas, respectivamente.

Seguindo o primeiro rastreamento, eventos de transformação selecionados de interesse foram transferidos para um meio de regeneração para produzir plântulas, que foram eventualmente transferidas em solos para recuperar plantas. Após um período de crescimento, as plantas selecionadas foram rastreadas uma segunda vez através de análises FISH da ponta da raiz comprimida para reafirmar a herança.

10

A. Experimentos de co-transformação de grupos de BACs

Os embriões foram co-transformados com grupos de construções de DNA. Esses grupos podem compreender combinações de construções de DNA derivados de clones BAC compreendendo repetições centroméricas de milho, construções de DNA derivados de clones BAC compreendendo segmentos teloméricos e/ou subteloméricos, plasmídeo marcador visual PHP23715, e polinucleotídeos codificando proteínas fortalecedoras do crescimento Ovule Development Protein-2, plasmídeos ODP-2 (PHP21875) e Wushel (PHP21139). Construções de DNA derivados dos clones BAC centroméricos incluem clones BAC tendo CentC apenas, CRM2 apenas, CentC e CRM2 apenas, e BACs principais tendo todas as quatro repetições centroméricas CentA, CentC, CRM1, e CRM2.

25

Análises FISH dos 80 calos transformados com grupos de BACs compreendendo DNA centromérico revelaram 42 eventos citogeneticamente detectáveis de novos grupos de CentC em adição aos sítios de centrômeros normais. Em alguns

exemplos, os elementos centroméricos de milho usados para transformação inseridos dentro dos cromossomos nativos, resultaram nas estruturas dicêntricas. Essas inserções de seqüências de DNA centroméricas variam em tamanho (número de repetições), número de inserções por cromossomo (até 3 detectáveis em um único cromossomo), e número de cromossomos com inserções (até 4 cromossomos com pelo menos uma inserção) e todas as inserções co-localizadas com a sonda plasmidial marcadora RFP. Isso indica que fragmentos de DNA exógeno podem ser montados em grandes blocos e integrados dentro de um cromossomo de milho.

B. Transformação com construções de DNA protótipo de minicromossomo linear montados através de ligação *in vitro* de um clone BAC centromérico, seqüências teloméricas e seqüências marcadoras.

O clone Mo17 BAC, bacm.pk128.j21, com orientação invertida das repetições em tandem CentC foi identificado como descrito no Exemplo 1D. Seqüências teloméricas foram geradas através de amplificação de PCR de oligonucleotídeos teloméricos e clonadas em um vetor plasmidial. Uma construção de DNA linear foi gerado deste clone BAC através de montagem *in vitro* com marcadores de seleção (moPAT, AmCyan1, DsRed2), uma origem de replicação (ori) 18-26S rRNA NTS, e seqüências teloméricas (TEL). Cada fragmento de DNA tinha sítios de reconhecimento que permitiam a montagem de uma estrutura única mediante ligação. A construção linearizada montada compreende:

TEL-(*SpeI*)-ubi pro::ubi 5' UTR::ubi intron::AmCyan::moPAT-(*NotI*)-bacm.pk.128J21-(*NotI*)-ori-ubi pro::ubi 5' UTR::ubi intron::DsRed::moPAT-(*SmaI*)-TEL

5           A mistura de ligação inteira, contenda construção montado bem como os sub-produtos da ligação, foram liberados dentro de embriões imaturos Hi-II através da transformação biolística. Mais do que duzentos de eventos foram propagados como clones de calos individuais baseados  
10 no marcador de seleção selecionável e fluorescente (PAT). Três grupos de clones foram recuperados: aqueles que mostraram apenas marcadores fluorescentes vermelhos (72), apenas marcadores fluorescentes azuis (83), ou ambos apenas marcadores fluorescentes (137). Eventos expressando ambos  
15 os marcadores foram selecionados de análises posteriores através de FISH.

Em adição aos eventos de integração simples, um número de eventos de integração múltiplos foram observados cada um  
20 no mesmo cromossomo, ou em cromossomos diferentes. Em dois eventos nós observamos rearranjos cromossômicos. Os sítios de inserção adicionais da repetição centromérica CentC co-localizada com a sonda marcadora PHP23715 sugere possível formação cromossômica dicêntrica. Análises de células  
25 dividindo na anáfase mostraram pontes cromossômicas consistentes com a presença de cromossomos dicêntricos com dois centrômeros funcionais devido à integração das seqüências de DNA centroméricas exógenas CentC. A função centromérica é indicada pela formação dos cromossomos

dicêntricos, aparecimento das pontes cromossômicas na anáfase, e a indução de quebras cromossômicas. Esses resultados indicam que os elementos cromossômicos podem se auto-arranjar dentro da célula vegetal em blocos de multicópias, associados com proteínas cromatina, e em alguns casos podem adquirir função centromérica.

Um evento mostrou um cromossomo 6 rearranjado tendo dois sítios de inserção da construção de DNA centromérico perto da região de organização nucleolar (NOR), bem como uma estrutura adicional semelhante à estrutura do minicromossomo com uma grande região centromérica e um pequeno sítio de inserção adicional da construção de DNA centromérico. A citologia deste evento pode ser uma indicação de ruptura cromossômica devido à formação de um cromossomo dicêntrico.

#### C. Transformação com grupos de clones BAC readaptados linearizados

Diversos clones BAC foram readaptados com Tn5-3 transposons feitos usando o sistema transposase (EPICENTRE EZ::TN™ pMOD™-2 MCS Transposon Construction Vector system (EpiCentre, Madison, WI, USA)) essencialmente como descrito no Exemplo 5E. Eles foram linearizados e usados por transformação biolística de embriões imaturos de milho Hi-II:

1. Sete diferentes variantes do clone readaptado bacm.pk128.j21 com blocos invertidos de repetições

centroméricas CentC representando diferentes inserções geradas por transposase dentro do mesmo clone BAC foram unidas em um grupo;

- 5 2. 84 clones BAC grupos principais centroméricos readaptados foram combinadas para gerar 4 grupos com 21 variantes individuais cada (Tabela 8). Cada um dos 4 grupos foi usado individualmente por transformação biolística;
- 10 3. Clones BAC centroméricos readaptados do cromossomo 4 foram divididos em 2 grupos contendo três clones BAC de B73 e três clones BAC de Mo17 (Tabela 9); e,
- 15 4. Grupo 1 da Tabela 8, foi dividido em 4 subgrupos de 5 ou 6 clones BAC grupos principais centroméricos readaptados (Tabela 10). Cada um dos subgrupos foi usado individualmente para transformação biolística.

**TABELA 10**

Subgrupo 1.1	Subgrupo 1.2	Subgrupo 1.3	Subgrupo 1.4
bacm.pk007.a2	bacm.pk133.b10	bacm.pk119.a23	bacm.pk075.16
bacm.pk036.e13	bacm.pk077.k5	bacm2.pk174.e4	bacm.pk0066.j14
bacm.pk178.c10	bacm2.pk179.b18	bacm2.pk116.g16	bacm2.pk099.m24
bacm2.pk179.e1	bacm.pk0133.b11	bacm2.pk023.e24	bacm2.pk093.h11
bacm2.pk064.e15	bacm2.pk066.m12	bacm.pk135.16	bacm2.pk083.a2
			bacm.pk076.m3

20

Para cada exemplo acima, os embriões imaturos Hi-II foram co-transformados com ODP2, WUS, e/ou vetores de expressão ODP2+WUS e grupos BAC readaptados.

Diversas diferentes classes de eventos de integração foram encontradas quando construções readaptados linearizados dos clones BAC foram usados para  
5 transformação. Por exemplo, quando os construções BAC contendo blocos invertidos das repetições centroméricas CentC (bacm.pk128.j21), ou grupos ou subgrupos readaptados do grupo principal dos clones BAC foram usados para transformação:

10

1. Integrações únicas dentro de regiões eucromáticas dos cromossomos hospedeiros;
2. Integrações múltiplas dentro de regiões eucromáticas de cromossomos hospedeiros;
- 15 3. Integrações únicas dentro da região centromérica de cromossomos hospedeiros;
4. Integrações múltiplas dentro de regiões centroméricas de cromossomos hospedeiros;
5. Integrações que resultaram em quebras cromossômicas,  
20 tais como novas variantes não usuais de cromossomos de milho com braços cromossômicos reduzidos, ou duplicação de certas regiões cromossômicas, por exemplo um cromossomo 6 com dois NORs, ou formação cromossômica dicêntrica;
6. Amplificação local de marcadores e construções  
25 centroméricos na integração;
7. Amplificação de marcador e construções centroméricas dentro de segmentos de cromatina extra-cromossômica em algumas células;
8. Criação de novos minicromossomos tendo um centrômero

funcional similar aos cromossomos nativos, por exemplo segregação autônoma na mitose.

Essas observações indicam que clone BAC centromérico readaptado bacm.pk128.j21 e o grupo central dos clones BAC reunidos são capazes de induzir uma variedade de efeitos citogenéticos tais como formação de cromossomos dicêntricos, quebras cromossômicas, amplificação local de construções transgênicos e formação de elementos extra cromossômicos, i.e. minicromossomos.

Eventos de minicromossomos de sucesso que resultaram da reconversão (montagem) de um único BAC ou grupo de BACs específicos de centrômero com a construção Tn5-3 e sua subsequente linearização dentro de uma construção de transformação linear são descritos abaixo:

- 1) Grupo 1 região principal dos BACs específicos de centrômero (Tabela 8), ou subgrupos do grupo 1 (Tabela 10);
- 2) Grupo 3 região principal dos BACs específicos de centrômero (Tabela 8);
- 3) um único clone BAC, bacm.pk128.j21, com repetições CentC invertidas; e,
- 4) três clones BAC específicos de centrômero do cromossomo 4 B73 (Tabela 9).

O primeiro evento de minicromossomo de milho (CMC3 grupo 1 evento #14) foi encontrado entre eventos gerados por transformação biolística com grupo 1 da região



principal reconvertida do BAC Tn5-3 linearizado (Tabela 8). No meio seletivo calos embriogênicos crescendo ativamente expressaram o marcador visual DsRed2. Análises FISH no estágio de metáfase mostraram 0, 1, 2, ou 3 minicromossomos adicionais tendo várias formas e tamanhos (Figuras 1-4). Neste evento, 60 dos 80 núcleos pesquisados tiveram 1, 2, ou 3 minicromossomos juntamente com o complemento normal de 20 cromossomos nativos. Esses cromossomos artificiais variaram em tamanho de cerca de 20% a cerca de 50% da média do cromossomo nativo de milho como medido na metáfase. Preliminarmente medidas na prometáfase mostrou os minicromossomos relativamente inalterados no tamanho, enquanto os cromossomos nativos são cerca de 4-5 vezes maiores, portanto os minicromossomos medidos neste estágio são cerca de 5% para cerca de 15% do comprimento de um cromossomo médio nativo na prometáfase de milho. Como determinado por FISH os minicromossomos são predominantemente compostos de repetições centroméricas e componentes Tn5-3. Diversos exemplos de formação de anel cromossômico que tem uma organização mais complexa foram também observados. Esses minicromossomos recentemente formados são aparentemente capazes de replicação e segregação durante a mitose (Figura 4), no entanto a segregação não é perfeita e algumas não disjunções foram observadas, resultando em células com uma mudança no número de cromossomos. Calos do grupo 1 CMC3 evento #14 foram mantidos crescendo ativamente sob seleção por pelo menos 10 meses, amostrados em vários momentos, e analisados por FISH para demonstrar manutenção estável do minicromossomo

através de muitas rodadas de divisão celular mitótica. Este evento, grupo 1 CMC3 evento #14, foi também analisado por FISH para a presença de telômeros usando as sondas overgo Telo-31 (SEQ ID NOS: 192 e 193) usando núcleos de calos na metáfase. Dois dos quatro loci positivos telo-31 foram observados em cada minicromossomo, onde os dois loci observados podem representar 4 loci separados que não podem ser distinguidos nesta resolução. A intensidade do sinal telo-31 foi geralmente mais fraca no minicromossomo quando comparada com o sinal observado para os cromossomos nativos em cada amostra. Plantas foram regeneradas deste evento e suas pontas das raízes foram analisadas com FISH para determinar se os minicromossomos foram herdáveis através de sucessivas divisões mitóticas em um ambiente de casa de vegetação. Cinco das 19 plantas que regeneraram deste evento de transformação mostraram a presença de um minicromossomo. Quatro plantas tiveram uma alta incidência de núcleos com um único minicromossomo mais o complemento normal de 20 cromossomos nativos. A quinta planta teve uma maioria deste núcleo com 1, 2, ou 3 minicromossomos mais o complemento normal de 20 cromossomos nativos. Todos os minicromossomos descritos acima foram compreendidos predominantemente de repetições centroméricas e componentes Tn5-3.

25

Subsequentemente, grupo 1 da região principal do BAC readaptado foi ainda dividido em quatro subgrupos tendo 5-6 dos clones BAC da região principal reconvertida (Tabela 10). Análises FISH demonstraram a presença de

minicromossomos nos calos embriônicos gerados por subgrupos 1.1 e 1.3. Dois eventos de minicromossomos foram produzidos do subgrupo 1.1: o primeiro evento teve o complemento normal dos 20 cromossomos, mais 1 minicromossomo que não hibridiza com o marcador PHP23715 ou CentC em um nível detectável; o segundo evento mostrou 24-28 cromossomos, 3 cópias do cromossomo 6, e 1 minicromossomo. Baseado em observações FISH, este minicromossomo foi positivo para CentC, mas não foi consistentemente positivo para a sonda PHP23715. Este evento pode ter sido produzido por integração e ruptura de um cromossomo nativo, e/ou condições produzidas por ou resultante de aneuploidia. Subgrupo 1.3 produziu 5 eventos. Três dos cinco eventos parecem ser de novo formação de minicromossomo e tinham complemento normal de 20 cromossomos mais 1 minicromossomo, e os minicromossomos foram positivos para o marcador PHP23715 e CentC por análises FISH dos eventos de calos primários na metáfase. Um desses eventos, CMC3 subgrupo 1.3 evento #27, foi ainda analisado por FISH para a presença de telômeros usando as sondas overgo Telo-31 (SEQ ID NOS: 192 e 193) usando núcleos de calos na metáfase. Este evento tem um minicromossomo muito pequeno com dois fortes loci CentC e dois loci telo-31 em cada minicromossomo. Os dois loci telo-31 observados podem representar 4 loci separados que não podem ser distinguidos nesta resolução. Este evento tem um minicromossomo menor do que o observado em eventos prévios. Quando medido na metáfase, o minicromossomo é aproximadamente 0.5 a 1 micron em comprimento, que é cerca de um terço para cerca de um meio do tamanho dos

minicromossomos em outros eventos independentes. O sinal FISH para telo-31 foi geralmente mais fraco no minicromossomo do que em cromossomos nativos. Calos do subgrupo 1.3 CMC3 evento #27 têm sido crescidos ativamente sob seleção por aproximadamente 10 meses, amostrados em vários momentos, e analisados por FISH para demonstrar manutenção estável do minicromossomo através de muitas rodadas de divisão celular mitótica. Os dois outros eventos do subgrupo 1.3 tiveram 19 cromossomos normais mais um minicromossomo, possivelmente como um resultado da integração e quebra cromossômica. Usando análises FISH um dos dois minicromossomos foi positivo para CentC apenas, e o segundo foi positivo para ambos os marcadores PHP23715 e CentC. Usando FISH no núcleo na metáfase dos calos, todos dos eventos do subgrupo parecem essencialmente similares quando comparados com mostras de outros eventos de minicromossomos gerados, como mostrado nas Figuras 1, 2, 5, 6, e 9.

Outro evento de minicromossomo foi observado (Figuras 5-8) de um evento de transformação biolística de um embrião imaturo Hi-II com região principal reconvertida linearizada dos clones BAC grupo 3 (Tabela 8). O evento de calos embriogênicos resultante foi positivo no meio de seleção Bialophos e expressou a proteína marcadora fluorescente DsRed. Análises FISH mostraram que este evento foi tetra-aneuplóide, com apenas 39 cromossomos observados porque um cromossomo 6 estava ausente. Cada núcleo teve 0, 1, ou 2 minicromossomos neste evento. Como descrito com o primeiro

evento de minicromossomo do grupo 1, os minicromossomos neste evento foram compreendidos predominantemente de repetições centroméricas e componentes Tn5-3. Na anáfase, as cromátides irmãs dos minicromossomos foram capazes de  
5 segregar (Figura 7) indicando a presença de centrômeros funcionais. O quadro acima indica que os minicromossomos estão replicando autonomamente e mostram estabilidade através de sucessivas divisões mitóticas.

10 Outro evento de minicromossomo foi observado (Figuras 9-10) de um evento de transformação biolística de um embrião imaturo Hi-II com clone BAC Tn5-3 readaptado linearizado bacm.pk128.j21. Novamente, os calos embriogênicos deste terceiro evento foram positivos no meio  
15 de seleção Bialophos e expressaram a proteína fluorescente DsRed. Uma planta foi regenerada deste evento e as pontas das raízes rastreadas via FISH. Cada núcleo teve 0 ou 1 minicromossomo. Naqueles núcleos com um minicromossomo, apenas 19 dos 20 cromossomos nativos foram observados.  
20 Análises FISH na distribuição metafásica mostraram que o minicromossomo único foi composto principalmente de repetições centroméricas e componentes Tn5-3.

25 Outro evento de minicromossomo, bCMC4 evento #73, foi observado de um evento de transformação biolística de um embrião imaturo Hi-II com linearizados, Tn5-3 readaptado três clones BAC específicos de centrômero do cromossomo 4 (Tabela 9). O evento de calos embriogênicos resultantes foi positivo no meio de seleção Bialophos e expressou a

proteína marcadora fluorescente DsRed. Análises FISH mostraram que este evento foi aneuplóide, com apenas 19 cromossomos e 1 ou 2 minicromossomos. Similar aos minicromossomos descritos acima, os minicromossomos neste evento foram compreendidos predominantemente de repetições centroméricas e componentes Tn5-3.

Observações em todos os eventos de minicromossomos indicaram que minicromossomos recém formados predominantemente resultaram de concatenação e/ou amplificação das construções de DNA linear liberados para células vegetais para produzir um minicromossomo *de novo*.

Três dos eventos de minicromossomos foram posteriormente analisados utilizando imunofluorescência com um anticorpo marcado fluorescentemente levantada contra a proteína específica para centrômero/cinetócoro, proteína C centromérica (CENPC). Imunocoloração de distribuição nuclear revelou que CENPC liga especificamente à região centromérica dos cromossomos nativos. Em adição, a CENPC localizada em distintas posições em todos os minicromossomos em todos os três eventos de minicromossomos estudados (Figuras 3, 4, 8, e 10). Ligação FISH com imunolocalização mostrou que a repetição CentC e a localização da sonda marcadora DsRed2 sobrepôs com CENPC nos minicromossomos. Como visto para os cromossomos nativos, na metáfase os minicromossomos têm dois diferentes loci de CENPC (Figuras 3, 8, e 10), e na anáfase as cromátides irmãs dos minicromossomos separadas e cada

cromátide irmã tem um único loci de CENPC (Figura 4). Os resultados acima indicam que os minicromossomos podem recrutar as proteínas necessárias, tais como CENPC, para formação de cinetócoro, e, portanto agem autonomamente dos cromossomos nativos durante a replicação e segregação dentro das células filha durante a mitose e meiose.

Vários milhares bialophos-resistente, eventos transgênicos de milho DsRed positivo tem sido gerados e pelo menos diversas centenas foram citologicamente caracterizados. Os eventos mostram uma alta incidência de integração dentro dos cromossomos hospedeiros, com cerca de 60% dos eventos mostrando integração detectável por FISH. Ambos os marcadores visual e de seleção estão presentes em quase 39% dos eventos, mas não detectáveis por análises FISH. Até o momento a maioria das combinações das construções recombinantes produziram minicromossomos contendo ambos marcadores e repetições CentC detectáveis por FISH em apenas cerca de 1% dos eventos (4 eventos, Figuras 1-10). A exceção é o subgrupo 1.3 que gerou minicromossomos contendo ambos marcadores e repetições CentC em cerca de 12% dos eventos analisados (4 dos 34).

#### D. Medidas do tamanho dos minicromossomos artificiais

Três dos eventos com minicromossomos de milho autônomos foram posteriormente caracterizados pela medida do tamanho dos minicromossomos montados e o cromossomo 6, que é facilmente identificado pela sonda 18-26S rDNA FISH. Todas as medidas foram tomadas no núcleo metafásico, que

deram medidas mais consistentes. Outros estágios são menos definidos e altamente variáveis no tamanho do cromossomo, por exemplo, medidas preliminarmente na prometáfase mostram os minicromossomos relativamente inalterados no tamanho em  
 5 relação às medidas na metáfase, enquanto os cromossomos nativos são cerca de 4-5 vezes maiores, portanto os minicromossomos medidos neste estágio são cerca de 5% a cerca de 15% do comprimento de um cromossomo na prometáfase de milho nativo médio. Portanto, minicromossomos medidos na  
 10 metáfase provavelmente parecem maior com relação aos cromossomos nativos do que se medidos em um estágio diferente. Cromossomos foram medidos usando um microscópio fluorescente Leica DMRXA, imagens capturadas com uma câmara Photometrics CoolSnap CCD e medidas capturadas com software  
 15 de análise de imagem Metamorph® (Molecular Devices, Sunnyvale, CA, USA). Todas as medidas estão em microns.

Cromossomo 6 nativo (n=29):

Média = 4.62 (l) e 2.38 (w)

20 Faixa = 3.16 - 5.78 (l) e 2.06 - 2.70 (w)

Minicromossomo (n=37):

Média = 1.29 (l) e 1.67 (w)

Faixa = 0.75 - 3.07 (l) e 1.12 - 3.17 (w)

25

Os minicromossomos de milho estão em média cerca de 28% do cromossomo 6 em comprimento, mas pode variar de cerca de 13-97% do comprimento total do cromossomo 6 na metáfase.



O tamanho dos minicromossomos de milho observados podem também ser estimados em Mb. Por exemplo, o genoma de milho compreende cerca de 2500 Mb do DNA total, com cromossomos variando em tamanho de cerca de 150-350 Mb, cromossomo 6 é aproximadamente 200 Mb (Seneca 60).

### **EXEMPLO 8: Métodos**

Isolamento de DNA de espigas imaturas ou folhas verdes das plantas de milho foi realizado essencialmente como descrito em Ananiev et al. (1997) Proc Natl Acad Sci USA 94:3524-3529. DNA de clone BAC foi isolado usando o kit de plasmídeo Nucleobond (BD Biosciences Clontech, California) de acordo com as recomendações do fabricante. Preparação de DNA de alto peso molecular em blocos de agarose foi realizada essencialmente como descrito em Liu & Whittier Nucleic Acids Res (1994) 22:2168-2169. Digestão de restrição de DNA, gel de eletroforese, Southern blotting, e filtro de hibridização foram realizados usando técnicas padrões como descrito em Sambrook et al. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd Ed., Cold Spring Harbor Laboratory Vols. 1-3. As referências acima são todas aqui incorporadas por referência.

#### **A. Marcação de sondas overgo para colônia e hibridização southern**

Overgos agrupadas para cada sonda (5 pmol de cada oligo) foram combinadas com 2  $\mu$ l de tampão Klenow 10x, 1  $\mu$ l da enzima Klenow (5 U/ $\mu$ l), 1  $\mu$ l de 1mM dGTP, 1  $\mu$ l de 1mM

dTTP, [ $\alpha^{32}\text{P}$ ]dCTP and [ $\alpha^{32}\text{P}$ ]dATP - 5  $\mu\text{l}$  cada, e água estéril para um volume final de 20  $\mu\text{l}$ . A mistura de reação foi incubada à 14°C por 2 horas. Porcentagem de incorporação foi calculada e foi considerada aceitável a 50% ou mais.

5

#### B. Preparação de membrana e hibridização

Membranas foram preparadas usando 432 placas de 384 poços uniformemente distribuídos entre as bibliotecas BAC Mol7 *EcoRI* e *HindIII*. Um padrão de grade 4 x 4 que permitiu 10 96 placas com 384 poços serem espotadas em uma única membrana de nylon Millipore Imobilon N+ (Bedford, MA) foi usado. As 96 placas gradeadas compreendiam 90 placas de clone BAC e 6 placas de clone plasmidial usadas como marcadores de grade. Depois de gradear, membranas foram 15 cuidadosamente colocadas com as bactérias viradas para cima em placas ágar Luria-Bertani com 17  $\mu\text{g/mL}$  de cloranfenicol, as placas foram cobertas, invertidas, e crescidas à 37°C durante à noite. Depois do crescimento da colônia as membranas foram removidas das placas e denaturadas em 1.5 M 20 NaCl e 0.5 M NaOH por 5 min cada, seguida por neutralização em 1.5 M NaCl, 1 M Tris-HCl duas vezes por 5 min cada. Membranas foram secas e tratadas com proteinase K (100 mls em 1 mg/mL; Sigma, St. Louis, MO) por 50 min à 37°C.

25 Cada membrana foi embebida em 6x SSC, 0.5% de solução SDS em caixas plástica. Filtros foram pré-hibridizados à 56°C em 6x SSC, 0.5% SDS com constante agitação por pelo menos 20 minutos. Sondas overgo agrupadas foram desnaturadas à 100°C por 5 min e adicionadas à solução de

hibridização que tinha sido usada para pré-hibridização. Hibridização foi por 12-16 h à 56°C. Membranas foram lavadas progressivamente por 1 h cada à 56°C em 2x SSC e 0.1% SDS (lavagem 1), 1.5x SSC e 0.1% SDS (lavagem 2), e 5 0.1x SSC e 0.1% SDS (lavagem 3). Membranas foram seladas em plástico moldado e expostas ao filme raio X por 3 h durante à noite. Após a hibridização, os filtros foram despojados em 100 ml de 0.1x SSC e 0.1% SDS à 90°C por 10 min e estocados à -20°C. Membranas foram usadas múltiplas vezes.

10

### C. Métodos citológicos

Quaisquer métodos adequados, e composições, incluindo muitos métodos citológicos padrões, preparações, e semelhantes são conhecidos no estado da técnica e podem ser 15 usados para analisar tecidos vegetais.

#### *i. Preparação do núcleo de tecido de calos de milho*

1. Calos usados para fazer preparações nucleares foram primeiro intoxicados com óxido nitroso à 150psi por 3 horas e então imediatamente fixados. Óxido nitroso detém o núcleo 20 na metáfase que permite uma distribuição cromossômica melhorada para análises FISH.

2. Fixar as amostras de tecidos dos calos em ácido acético 50% por pelo menos 1 hora. Tecidos podem ser 25 estocados indefinidamente em ácido acético 50% à -20°C.

3. Separar embriões somáticos dos calos e colocar em 50 µl de gota de tampão PIM (50mM CaCl<sub>2</sub>, 10mM de acetato de

sódio, pH 5.8) em uma placa de petri pequena.

4. Dissecar embriões somáticos em pedaços menores do que 0.5 mm.
- 5 5. Lavar o tecido em tampão PIM 3-5 vezes sobre 1 hora para remover o fixador. Pipetar lentamente diversas vezes para lavar e substituir com tampão PIM fresco.
6. Remover cuidadosamente tampão PIM. Adicionar 50 µl de  
10 solução de enzima de digestão (2% p/v celulase (Cat# CEL, Worthington Biochemical Corp. (Lakewood, NJ, USA)), 0.2% p/v pectinase (Cat# PASE, Worthington Biochemical Corp. (Lakewood, NJ, USA)); 0.5% p/v albumina sérica bovina).
- 15 7. Digerir o tecido à temperatura ambiente, no escuro, em uma câmara úmida por 1-2 horas. Como o tecido começa a ficar mole, pipetar muito gentilmente e/ou esmagar com sonda para desmembrar grandes pedaços e liberar as células.
- 20 8. Remover cuidadosamente a solução de digestão da enzima e substituir com cerca de 50 µl de tampão PIM.
9. Transferir células/núcleos livres para um tubo de microcentrífuga. Adicionar mais tampão PIM para o tecido  
25 digerido remanescente e gentilmente pipetar para liberar células, transferir essas células para um tubo de microcentrífuga, repetir quando necessário.
10. Peletizar células em uma microcentrífuga à 500 rpm por

3 minutos, remover sobrenadante. Adicionar tampão PIM fresco e gentilmente ressuspender as células. Repetir esse passo de lavagem mais 3 vezes.

- 5 11. Remover tampão PIM e substituir com ácido acético 50%. Ressuspender gentilmente as células, peletizar à 500 rpm por 10 min., remover o sobrenadante e adicionar ácido acético 50%. Repetir.
- 10 12. Estocar núcleo isolado em ácido acético 50% à -20°C. O volume final de ácido acético 50% deve ser 2X o volume do pélete nuclear.
- 15 13. Transferir 5 µl do núcleo ressuspendido para uma lâmina de vidro, adicionar uma lamela de 18 mm<sup>2</sup>.
14. Aquecer lâmina em uma placa quente à 70°C por 15 segundos.
- 20 15. Remover lâmina do calor e gentilmente pressionar para baixo na lamela lamínula para esmagar o núcleo.
16. Permitir a lâmina esfriar brevemente, então imergir lâmina no nitrogênio líquido por 10-15 segundos.
- 25 17. Remover lâmina do nitrogênio líquido e aquecer lamínula com seu hálito.
18. Remover rapidamente lamínula com a ponta de uma lâmina

de navalha.

19. Colocar a lâmina em 2 mudanças de etanol 100% por 2 minutos cada.

5 20. Permitir a secagem das lâminas ao ar. Estocar lâminas à -20°C até o necessário.

*ii. FISH seguido por imunolocalização direta do núcleo*

*a. Preparação de sonda overgo para FISH*

10

Sondas overgo são descritas na Tabela 1.

1. Adicionar 10 µl da mistura overgo 100 µM, compreendendo concentrações iguais de cada overgo, para 5 µl de água deionizada.

15

2. Aquecer à 95°C por 1 min, então transferir para o gelo.

3. Adicionar a mistura acima:

20

- 2 µl de tampão DNA polimerase 10X (100mM Tris-HCl, pH 7.5, 100 mM MgCl<sub>2</sub>, 7.5 mM DTT)

- 0.5 µl de fluoróforo dUTP

a) dUTP-Cy3 (Amersham)

b) dUTP-FITC (Roche)

c) dUTP-Texas Red (Molecular Probes)

25

- 2 µl dNTPs (200 µM A-, G-, CTP; 40 µM TTP)

- 0.5 µl Klenow

4. Incubar à 37°C por 20 min.

5. Lavar sonda usando o kit de extração de nucleotídeo da Quigen. Eluir em 50 ml de formamida 50% em um tampão de eluição do kit.

5            *b. Hibridização fluorescente in situ (FISH)*

FISH dos núcleos de milho em lâminas foi feita essencialmente como segue:

- 10    1. Fixar lâmina 10 min. em paraformaldeído 1% v/v m solução salina tamponada com fosfato (PBS) pH 7.2
2.    2. Lavar 2X 5 min. em PBS
- 15    4. Lavar 2 min. em água destilada/deionizada
5.    5. Secar lâmina
- 20    6. Hibridizar 2 min à 80°C em sonda fluorescente titulada em formamida 50% em uma concentração final de 50 mM MgCl<sub>2</sub>
7.    7. Hibridizar 30 min - durante à noite na câmara úmida à 37°C
- 25    8. Lavar 5 min. em 2X SSC
9.    9. Lavar 5 min. em 0.2X SSC
10.    10. Secar lâmina ao ar

11. Adicionar Vectashield<sup>®</sup> com DAPI (Cat# H-1200, Vector Laboratories, Burlingame, CA, USA) e lamínula (5 ml meio montagem/22mm lamínula)

5

12. Examinar sobre microscópio usando conjunto de filtros apropriados e/ou óleo de imersão quando necessário.

*c. Imunolocalização*

10 Após exame e caracterização da localização da sonda FISH, essas mesmas amostras podem ser processadas e usadas para imunolocalização usando uma sonda anticorpo identificada diretamente. Imunolocalização do anticorpo anti-CENPC de Coelho policlonal fluorescente-marcada foi  
15 feita essencialmente como segue:

1. Remover lamínula

20 2. Lavar 5 min. em etanol 70% v/v EtOH para remover meio montagem e óleo de imersão

3. Lavar 3X 5 min. em PBS

25 4. Bloquear 1 hora à 37°C me uma câmara úmida em soro de Coelho normal 5% v/v (Jackson Immunoresearch, West Grove, PA, USA) em PBS-BT (PBS com BSA 3% p/v, 0.02% p/v Na azida, 0.5% v/v Triton X-100)

5. Enxaguar em PBS



6. Incubar durante à noite à 37°C em uma câmara úmida com 1° anticorpo em soro de coelho normal 5% v/v em PBS-BT. Anti-CENPC-Cy3 (ou -FITC) de Coelho foi usado em uma diluição 1:200, concentração final 2.5 µg/mL de anticorpo marcado

7. Lavar 3X em PBS sobre um período de 1 hora

10 8. Secar lâmina

9. Adicionar Vectashield® com DAPI (Cat# H-1200, Vector Laboratories, Burlingame, CA, USA) e lamínula (5 ml de meio montagem/22mm lamínula)

15

10. Selar lamínula com unha polida

11. Examinar sobre microscópio usando filtros apropriados e/ou óleo de imersão quando necessário.

20

*d. Produção de anticorpo CENPC e marcação*

Um CENPC de milho homólogo de mamífero foi isolado por Dawe et al. (1999 Plant Cell 11:1227-1238) e mostrou ser um componente do cinetócoro em milho. Um peptídeo conservado de 20 aminoácidos do domínio amino terminal foi sintetizado e usado para produção de anticorpo policlonal em coelhos (Openbiosystems, Huntsville, AL, USA). Os anticorpos resultantes foram diretamente marcados com fluoróforos processando o kit de marcação Fluorolink-AbCy3 (GE

Healthcare, UK) ou kit de marcação de proteína fluorescente (Roche Diagnostics Corp., Indianapolis, IN, USA).

*iii. Fibra-FISH*

5           Fibras de DNA estendidas nas lâminas citológicas foram preparadas como descrito em Jackson *et al.* (1998) Genome 41:566-572. Sondas para fibra-FISH foram marcadas com biotina-11-dUTP (Roche, Germany) ou DIG-dUTP (Roche, Germany) usando o kit Nik Translation Labeling (Roche, Germany) de acordo com recomendações do fabricante. Após a precipitação, as sondas foram re-dissolvidas em tampão TE e estocadas à -20°C. Para fibra-FISH, as sondas foram hibridizadas para fibras de DNA em uma mistura de formamida 50% (v/v), SDS 10% (v/v), e SSC 2x em um volume final de 10 µL. As lâminas foram cobertas com lamínulas, selados com cimento de borracha e incubadas à 80°C por 2 min para desnaturar ambas as sondas e o DNA alvo, seguido pela incubação à 37°C. As lavagens pós-hibridização e detecção do sinal foram executados como descrito por Zhong *et al.* (1996) Plant Mol Biol Rep 14: 232-242. As sondas marcadas com biotina foram detectadas com fluoresceína-avidina DN (Vector Laboratories, Burlingame, CA, USA), biotinilada anti-avidina D (Vector Laboratories, Burlingame, CA, USA) e novamente com fluoresceína-avidina DN (Vector Laboratories, Burlingame, CA, USA). As sondas DIG-marcadas foram detectadas através de anticorpos monoclonais anti-DIG de rato (Jackson Jackson ImmunoResearch, West Grove, PA, USA) e anticorpos anti-rato Cy3-conjugada em ovelha (Jackson ImmunoResearch, West Grove, PA, USA). As lâminas foram

então montadas em um meio de montagem Vectashield (Vector Laboratories, Burlingame, CA, USA). Preparações foram examinadas usando um microscópio fluorescente Leica DMRXA, imagens capturadas com uma câmara Photometrics CoolSnap  
5 CCD. Imagens foram capturadas usando software de análise de imagem Metamorph<sup>®</sup> (Molecular Devices, Sunnyvale, CA, USA). Fibra-FISH foi realizada em 3 para 5 preparações de cada linhagem.

## LISTAGEM DE SEQUÊNCIAS

<110> Depositante: PIONEER HI-BRED INTERNATIONAL, INC.

<120> Título da Invenção: Microsomo Artificial de Planta

<130> Referência do documento: 2083-PCT

<150> Número do Documento de Prioridade: 60/801,004

<151> Data de Depósito do Documento de Prioridade: 2006-05-17

<160> Quantidade de SEQ ID NOs.: 193

<170> Software: FastSEQ for Windows Version 4.0

<210> SEQ ID NO: 1

<211> Comprimento: 4635

<212> Tipo: DNA

<213> Organismo: Zea mays

<220>

<221> Nome: Origem

<222> Localização: (1)...(4635)

<223> Outras Informações: CentA

<400> SEQ ID No. 1

```

tgatgagaac ataaccgca cagatatgac catgttaatg gctcctgcta caaagacatt 60
gaggaacaaa gaagttgatt ggggaccaag taatgatatt tccaacattt ccaacaaagc 120
aagcacatca tcaaatttaa agatatactt gggtgaggag catacactag agtcgaggac 180
gactctatta caagaagggg aggatgatga ggacatcact gccatcaata caccacacca 240
gcgacctcct tcaccattta ataatggacc agtaaacgag tccgtgcacg taaatTTTTat 300
tatcagggtga actcgttccct tattgttgaa gctaatacatt ccttaaataga ggtactaata 360
ccttgtgatt actttattaa tctaaggtgt ttgggaggtg aaccatctag aatttgagaa 420
ggcaacaagg caataaaaagc tgctccactt gaggggattt cgaaactaca acaagtgcaa 480
gtttaagagg gcatatcttt cagctcctaa ggttgtttaa tgcaataag cacttgttgg 540
aaaggctctct ttgtctactt tctagtggat caagaatcaa cgagagatca gacactaagt 600
gtccagaaac tgccgagtga actcctgctc tacccaagtc aatttcgtaa ctgcagcatg 660
caccaaatta aatggagcat aactttccac tcccaaggtt gtttagtgca aataactact 720
tgttggaaag ctctcttcgt ctactttcat gtgcatcaat aatcaatgac agaaaccaa 780
cgagggcgtcc agaaactgcc gagagagttt cgttctccat tagaactcct ttctattcct 840
ctatttaagc aactagcagc caccaaagaa cttgggtttt tgtttgatgt aagtttagcc 900
tttgctactt cttgtaaac gcatgtgtcg gctagaccac cgggatactt gaaacagaa 960
cccaactcta tcagatccgt gagtgtctgc tttttatctt gttcttgctt gttctcgatt 1020
gcttgacagt tcaaggctgt tcttggcagc gcaagggcag caacaacagg agccgatgta 1080
actatcgcta aggcgcagca cccttgggtg tgttgtagt ggatagcaca acgtcgacct 1140
ccaccccaaa tcgtagttat caggagacgg tgtacctgtc gctcaaggcg ccacaccatc 1200
ttggttggtg tagtcgggca gccaacgtcg ttctccaaca agtttccacc tccatcatct 1260
ctcatcgaaa gatcgggcac ctttctacct gttgggttta tcaagtggta tcaaatttca 1320
ggttgctcgg tgagagatct caatcttctt tgttttgttt acctacagtc cacttttgcc 1380
caaagatata tttagagcag aatttcacct aaaaacagtt tgagcctttg ctttactact 1440
tagttttcga cttggtgaat tccggtagct gcatttgggt cgagttgctg gtctaaagtt 1500
ttcttaccgc tagagtttcg agttcgcgcc acctgtttt aatcaccagt ttagacctct 1560
tgatgcaatt caacaaaaaa gaagagaaa caaaaggcga gtgcacaaaa aaagccgcac 1620
aacacagcaa acaaaaaaaa gacacgtgca aaaaacaaa gagagaaaaa aaccagttct 1680
gaattttggg agataaaatt tgtaagtgca acaaaacaaa aggcagtttg tgtgccttct 1740
ttttatagtt tcagaaatca gattgttgtt ctgagctttt ggtgatacta tttgtgtaac 1800
ggctcgcgct tctattacgg tttggactag gaccagcaca acaccttggt gaacgtttat 1860
tcaacttggt gtggctaacg tggtagtagc tattccttgg aactattggt taacagcca 1920

```

```

cctataaatc cacaaaattt tctacaacac caccaggttg tgctagcagc cactgttgtt 1980
gttggtcgtg ctgtttgcca ggcctcctg ctttgcgtgg tgagaacttg taagaacttg 2040
ttaaccagt ttgagagtga gagattacaa caatgattcc tagtagttta tagaatcaaa 2100
gatatttttt attgtttctt gtctttacta aacatggcag gtgatatgga catttttgac 2160
ccaaccgaac gttatattgg aggcattcatt caacacttgc ctttatatgc cggtaaattc 2220
gatcctcatg catacattga ttgggagcta aagctagata aggaatttga taagcatgat 2280
ctatctcaaa aacaaaagat ttatattgcc tctaatttgt taactgagca cgcattgatg 2340
gaatggaaat acatttgtag gcacaacaaa gttccacaat cttgggaaga cttcaaactt 2400
cattttagag atgcattcat tccctgcatc tatgctgac atttgcttct taaattagac 2460
accttaaagc aggggtgctag gactgtgaaa gattattatt atgattttaa aatttttacc 2520
atgtttgctc gtttagatga atgcatggaa gatgcatga ctaggttcat gaaaggactc 2580
aattctgaaa ttcagactat agtcatgcat gaagcataca aacacatttc tcacttgttt 2640
ttgcttgcac gtaaagctga aatgagatt ctattataca attatacaag cactgaacat 2700
gtgagccata attcctcttt tgcattctct ctacatgctg atcaagaaca caaaataatg 2760
aaaccagctg ttgtttttcc atcatcacia gaagaattga ttgctgacac ttgtgatagt 2820
gaagatttgt gggataatga ttcacatgta ctaagacaac aactagtaaa tgaacatggt 2880
acatctatta ttgaacacaaa cattttggct aaaaaggaac atgtaatttg tattgcaaac 2940
gaaactgaag aaataaattt gctctctctt ttaaatactt ggggctatat tgaatttgat 3000
gatctttttg agctcggtaa tttggaaaa attttatttg ctagattcaa ctataccatg 3060
tccttctcat gatataattt atattgctgg caagtacaac aacataggac aatttcttgt 3120
gcatagaatt tctatttcat ctagatagt tgttcttca ctttgtgcaa ataagattt 3180
ggtatgttct caagaagaaa agaattctct gtttccatgt actttagttg aagtttcagg 3240
tttatatttg aaagacatta ataaaagctt agtcatcaac atcaatcatg atgcaaaacc 3300
gaggacggtt tgctatcaag aaggggagaa tgatgagaac ataaccgca cagatatgac 3360
catgttaatg gctcctgcta caaagacatt aaggaacaaa gaagttgatt ggggaccaag 3420
taatgatatt ttcaacattt ccaacaaagc aagcacatca tcaaatttaa agatatactt 3480
gggtgaggag catacactag agtcgaggac gactctatta caagaagggg aggacgatga 3540
ggacatcact gccatcaata caccacacca gcgacctct tcaccattta ataatggacc 3600
agtaaagcag tccgtgcacg taaatttaat tatcagggtga actcgttctt tgttgttga 3660
gtaatcatt ccttaaatag ggtactaata ccttgtgatt actttattat tctaagggtg 3720
ttgggaggtg aaccatctag aatttgagaa ggcaacaagg caataaaagt tgctccactt 3780
gaggggattt cgaaactaca acaagtgcaa gtttaagagg gcatactctt cagctcctaa 3840
ggttgtttaa tgcaataaag cacttgttgg aaaggtctct ttgtctactt tctagtggat 3900
caagaatcaa cgagagatca gacactaagt gtccagaaac tgccgagtga actcctgctc 3960
taccgaagtc aatttcgtaa ctgcagcatg caccaaatta aatggagcat aactttccac 4020
tcccaaggtt gtttagtgca aataactact tgttggaaag ctctcttctg ctactttcat 4080
gtgcatcaat aatcaatgac agaaaccaa cgaggcgtcc agaaactgcc gagagagttt 4140
cgttctccat tagaactcct ttctattcct ctatttaagc aactagcagc caccaaagaa 4200
cttgggtttt tgtttgatgt aagtttagcc tttgtactt ccttgtaaac tcatgtgtcg 4260
gctagaccac cgggatactt gaaacaaaac cccaactcta tcagatccgt gagtgtctgc 4320
tttttatctt gttcttgctt gttctcgatt gcttgcaggt tcaaggctgt tcttggcacg 4380
gcaagagcag caacaacagg agccggtgta actatcgcta aggcgcagca cccttgtggt 4440
tgttgtagtc ggatagcaca acgtcgacct ccacccaaa tcgtagttat caggagacgg 4500
tgtacctgtc gctcaaggca ccacaccatc ttggttgtgg tagtcgggca gccaacgtcg 4560
ttctccaaca agttttccac ctccatcatc tctcatcgaa agatcgggca cccttctacc 4620
cgttgcgctt atcaa 4635

```

```

<210> SEQ ID NO: 2
<211> Comprimento: 156
<212> Tipo: DNA
<213> Organismo: Zea mays

```

```

<220>
<221> Nome: Origem
<222> Localização: (1)...(156)
<223> Outras Informações: CentC

```

```

<400> SEQ ID NO. 2
ggttccggtg gcaaaaactc gtgctttgta tgcacccgca caccggtttt cggaatgggt 60
gacgtgcgac aacgaaattg cgcgaaacca ccccaaacat gagttttgga cctaaagtag 120

```

tggattgggc atgttcggtg cgaaaaacga agaaat

156

<210> SEQ ID NO: 3  
 <211> Comprimento: 6915  
 <212> Tipo: DNA  
 <213> Organismo: Zea mays

<220>  
 <221> Nome: Origem  
 <222> Localização: (1)...(6915)  
 <223> Outras Informações: CRM1

<400> SEQ ID No. 3

cttgggtcttg	gacagtacct	cactgatgaa	gtagacctga	tgaagctgag	gtgcccgatc	60
tttcggcgag	tagagataat	tccgatttgg	cggaagatga	cccttgcgat	ccgactacga	120
cgagcaagcc	cgagggcgcca	atgcaatcgc	tgaaccaact	ccctgtgggt	accgaccttg	180
ctgatgagag	atcggcctga	tcacgaagat	cgtttcctgt	gcgcaatcga	agaacgaaca	240
agaataagat	gagagcaatc	taatctatta	ctcgaggggtg	gagttctgaa	tacacgaaga	300
cagcagagat	tagcgcgtgt	tcgagagtag	ctaaggctaa	cgtaaaacaa	aactcaggaa	360
ataaaggagg	cgcagctcct	gaataaatag	agagggggcg	cagcccctag	gggcccggcaa	420
ccctaggtcg	tccattatgg	gccgcaattg	ggctggctcg	ctattcttcc	gggccttcgt	480
tctttaacaa	catgatgtag	ttcaattctc	ttgcacgggc	ccgagtcact	ggcccagggtg	540
gaaggggtgg	cgctgggct	ggagaagggtg	ctgctgggtg	agatgtgctc	gtgttggtgt	600
tgatgtccgc	atcatcctcc	ccttcttgaa	ctgaagtcgt	cctcgacggc	aactcctcat	660
cttctcccgc	atacggtttc	aaatctgcaa	cattaaaact	agtggaaaca	ccaaactccg	720
caggtaggtc	gagggataaa	gcattatcat	taatcttggg	tagtatctta	aaaggaccag	780
cagcagagag	catcaattta	gaacggcgca	aagtaggaaa	acgatccttt	ctcaaatacga	840
accaaacat	atcaccaggt	tcaaaagtaa	catgtttctc	tcctttacta	ccagcaacct	900
gattttttgt	attagtagca	gcaatgttct	gttctgtttg	ttcatgtatg	gtaatcattt	960
gttcaacatg	tgcagaagca	tctacatgty	ggcgcttcgc	agcattaagt	gaaatcaaat	1020
caataggtgc	cctaggaatg	taaccataaa	caatctggaa	agggcacatc	tttgtagaag	1080
aatgcgtggc	atgattgtaa	gcaaattcaa	catgaggcaa	gcaatcctcc	caacgtcgca	1140
aattcttgtc	taaaacagcc	ctaagcatgg	tagataaagt	tctatttact	acctcagttt	1200
gtccatcagt	ttgaggggtga	caagtagtgc	taaacaacaa	tttagttccc	aattttattcc	1260
aaagagatct	ccaaaaatgg	cttagaaaact	tagcatcgcg	atcagagact	attgtttttg	1320
gaataccatg	taaacgaata	atcttctctaa	agaacaattc	agcaacaatg	ctagcatcat	1380
cagttttatg	acaaggtata	aagttagcca	ttttagagaa	tctatcaaca	accacaaaaa	1440
tactatccct	ccccttctta	gttctagcca	agccccaaac	aaagtccata	gagatatcaa	1500
gccaaggaga	agaagggaca	ggcaaaggca	tatacaaac	atggttgttc	aaccgtgact	1560
tagctttctg	acaagtagtg	cagcgtgcaa	caaggcgtc	cacatcagcg	cgcatccgag	1620
gccaaaagaa	gtgggcagcc	aacacctcat	gtgtcttgta	gacgccaaaa	tgccccatga	1680
gaccgcctcc	atgcgcttcc	tgtaacaaca	aaagacgaac	cgagctagct	ggaacacaca	1740
gcttgttagc	gcgaaacagg	aacccatcct	gtatgtgaaa	tttgccccat	ggttttccat	1800
taatacaatg	gccgaaagca	tctttaaaat	cagcatcgtc	aacatattga	tccttcacag	1860
tgtccaaacc	aaagatttta	aaatctaact	gtgacagcat	ggtatagcga	cgagacaaaag	1920
catcagcaat	aacattgtcc	ttcccgttct	tgtgtttaat	aatgtaagga	aaggactcaa	1980
tgaattctac	ccatttagca	tgacgacggg	tcagatttgt	ttgggtacga	atatgtttta	2040
aagcctcatg	atcagaatgg	attatgaact	cacgatgcca	aagatagtgc	tgccatgtat	2100
gcaaagtgcg	cactaaagcg	taaagctcct	tatcataagt	agaatatttc	agactagcac	2160
cgcttaatth	ttcactaaaa	taagcaactg	gttttccttc	ttgtaacaaa	acagcaccca	2220
gcccaatacc	actagcatcg	cattcaagct	cgaaaacttt	attaaaatca	ggcaattgca	2280
agaggggagc	ttgggttaac	ttatctttca	aagtgtgtaa	cgctacctcc	tgcgaaatcac	2340
tccaggcaaa	cggcacatct	ttctttgtaa	gctcatgtag	agggcgtgca	atggagctaa	2400
aatcacgaac	aaatctgcgg	tagaaaccgg	caagtccaag	aaagctccga	atttgtgtga	2460
ccgtcgtcgg	tgttagccac	tcccgaatgg	cagcaatctt	gctgctatcc	acctcaatgc	2520
cctgctgagt	aaccacataa	ccaagaaacg	agacacgttg	cgtgcaaaat	gtgcactttt	2580
ccatggttacc	aaacaagcga	gcagcacgca	aagcgtcaaa	aacagcacgc	aaatgttcta	2640
aatgctcctc	tatagacttg	ctgtaaataa	ggatatcatc	gaaataaaca	acaacaaaca	2700
aacctatgaa	ggcccgtaga	acttcgttca	ttaaaccgat	aaaggtgctg	ggagcattag	2760
tcaatccaaa	tggcataacc	aaccattcat	ataaaccaaa	tttctgtttta	aaagccgttt	2820

tccattcattc	accgagtttc	attctaattct	gatggtaacc	gctacgcaa	tcaaccttag	2880
agaaaataac	ggcaccacta	agctcatcta	gcataatc	aaggcgtgga	ataggatatac	2940
gataacgaac	tgtgatgtta	ttaatagcac	gacaatctac	gcacatacgc	catgacccat	3000
ctttctttgg	aacgagtaaa	acgggaaccg	agcaagggct	aagagactca	cgaatgtaac	3060
ccttatcaag	cagcgtctgc	acctggcgct	ggatctcctt	cgtctcatct	ggatttgtag	3120
ggtacggagc	gcggttcggg	agagaagtgc	cggggatgag	atcgatctga	tgctcaatgc	3180
cacggagggg	aggaagacc	ggtggttaag	ccgtaggata	aacatcagca	tactcctgca	3240
aaaggttaac	aacagcaggg	ggtatatcca	aagacggtgc	atcatcaagc	ggaacaagca	3300
ttcgcgagca	tacaagtgca	taacagggca	tatgagcttc	atggagatca	tcaaaatcag	3360
cacgtgtagc	aagtaaaaca	ggagagtgc	acttgatttc	agatttaata	ggtgcggtg	3420
atgtcagttt	gacttgttgt	gcagttttag	cagccctagt	aagatcatct	ttcaaaattt	3480
ggtcaggggt	cattggatgt	ataattattt	tctggccttt	aaacatgaaa	gaataatgat	3540
ttgaacgacc	atgatgtaaa	ctatcagtat	catattgcca	aggtcgacc	aataacaag	3600
agcatgcttc	cataggaata	acatcgcaat	caacataatc	agaataagaa	cccagcgaaa	3660
aggggacacg	taccgaacgt	gttaccttta	ttttaccacc	atcattaagc	cattgaatgt	3720
gatacggatg	tggatgtgtg	cgagtgggca	aggataattt	ctctaccaac	gctgtacttg	3780
ccaaattggt	gcagctgcc	ctatcgatga	tgatgcgaat	cgaccgttcg	tgacgacgc	3840
ccttggtag	gaatagagtg	tgctcgctgat	ttttttcggc	ctgggcaacc	tgtgtgctga	3900
gaacacgctg	cacaacaaga	ctctcatacc	tatcagcgtc	gatgggatca	acgtggactt	3960
ctcatcttc	tgcatggta	gtggcaatca	tagcatgact	agtttctca	gaatcactgg	4020
ctgaagagta	ctcaccattg	tcacgtataa	gcaaggtacg	cttgtttggg	cagtccecgaa	4080
tcatgtgccc	aaaccctctg	caacgatggc	actgaatatac	ccgtgtacgt	cctgtggaag	4140
aagcggcgcc	tttggcaggg	ggcgccactg	gcttggccgg	ccctgtgctg	gatgtagtgc	4200
taggcgtagg	aggtgcaggg	ctggaaggag	ttgagctgtg	tggtggacc	cggcctgcaa	4260
aagagttagt	atatgtcttt	gatcgtcgtc	cctgcacttc	acgttcagct	ttgcaagcat	4320
attcaaaaca	tgtggttata	tcaaaataat	ccttataatc	aagtatatcc	tgaatttccc	4380
tgttcaaacc	accacgaaaa	cgcgccatag	cagcgtcatc	tgactcaacc	aaaccacaac	4440
gaagcatacc	cttttgcaac	tcctggtaat	attcctcaac	agatttgtgaa	ccttgttgaa	4500
aacgtctgcat	tttgttaagc	aaatcacgag	cataatagga	aggaacaaat	ctgtggcgcga	4560
tggcagtttt	taattgggtc	caagtaatga	cactgttaat	gggaagtttt	tgttatact	4620
cacgccacca	aattaaagca	aaatcagtaa	attcactaat	ggcagccttc	acttggctat	4680
tagcaggaat	atcatggcat	gaaaatttct	gttctacctc	taattcccaa	tcaagatatg	4740
cagcaggatc	atattttacca	ttaaagatg	gaattttaaa	tttaattctta	gaaaataagt	4800
cattaggggg	atgacgaacc	acacgacgtg	cacgaccacg	gcgatctcca	tcgtcctgct	4860
cagtgtcacc	gccgtattcc	tgctccatct	ttgtgggcaa	tgcatcaagg	cgtgccagga	4920
tggtgtcgag	tgtggtgcga	gtcggcgttt	gagcaaggtc	aagttggttg	aaacgctcgg	4980
tcgtcgaagt	gatcgttgaa	tcaagccgtt	catgcatcgt	cctaattgtca	gcagcaagtc	5040
catcaacttg	gcccttact	tccctgaact	gggcatccac	catgtcgtgt	gctcctgcca	5100
tagttagcgc	aaacaccaa	aggagaaaa	ccaacgacaa	aaacaggggt	gtactgctca	5160
caaggcgctc	acactagtgc	tgttatcaag	ttcttatccg	ttcttaccaa	gccacagtgg	5220
tgaactgcaa	ccaacaggtg	gaaccggtga	aagattggat	gagcgtattgc	ttggagaaac	5280
agaaacctgc	tcgtcgtaga	aatatgtgga	gttgtgggta	ggctgcactc	aagtcaagga	5340
ttagcacgat	caacaataa	tgcaaagttg	aattatagtg	caaaacacga	aactatattg	5400
ctggccacag	gtgcaaagga	tggatggatg	gaaatagcag	aatggcagta	acgtaaatat	5460
tgtagtagtg	atgccaaaa	ggcactagta	caaatcacag	gtgattttgt	ttttctttt	5520
tgtatgattt	ttttgatatt	tttctcagca	caagaagcaa	caagatagga	gctacacgaa	5580
gtttcaccta	aaacagatat	cagatgtggt	ctacagaaaa	tcaggaagtt	ctctgaaaag	5640
cgtgcgagaa	ctttgacgga	ttttttctt	tattttctg	aatttttttg	acaattttgt	5700
cgaaccccaa	acagaccgta	ggtgagtttg	gccgggctca	gaatgggtgc	aactagctcc	5760
tgtaaaaatt	tcagattttt	tggacaccgg	agcgaaaagt	tatgcccggt	ttaaggaagg	5820
tacctcaaat	tatgttttca	aacgaccgga	atgaaacaac	cgtatccttt	ctccttcggt	5880
gttttttggt	tctgtttttt	tttttacgta	accgaaggag	aaaaacaagg	aaacgatggt	5940
gactcgggtt	gttttttttt	ctgttttttt	ttctgttttt	tttcgtaacc	gaaggagaaa	6000
atcaaggaaa	acagccgttg	actcggtttg	ttttttctgt	ttttttttga	cgtaaccgaa	6060
ggagaaaaaac	aaggaacaa	tgttgactcg	gtttgtgggtg	tgatcaaacg	agagatggtg	6120
gcgcgcctag	ggtttgaatg	gtggaagaac	acaattgcaac	cagcaacaaa	tgacgcgaaa	6180
gacacacaa	tcaacaatgc	agattattga	aagaaagtgc	gaggctcaaa	aggggtgctg	6240
gataagatct	aacctgaatt	tttatgtggt	tttgtggact	gtaggaaaaa	aaaacgctcg	6300
ataaactcac	cgatcaacct	agaaatctga	taccaattga	tgaagctgag	gtgccgatc	6360
tttcggcgag	tagagataat	tccgatttgg	cggaagatga	cccttgcat	ccgactacga	6420
cgagcaagcc	cgaggcgcca	atgcaatcgc	tgaaccaact	ccctgtggtt	accgacctg	6480

```

ctgatgag atcggcctga tcacgaagat cgtttcctgt gcgcaatcga agaacgaaca 6540
agaataagat gcgagcaatc taatctatta ctcgagggtg gagttctgaa tacacgaaga 6600
cagcgcagat tagcgcgtgt tgcgagtag ctaaggctaa cgtaaaacaa aactcaggaa 6660
ataaaggagg cgcagctcct gaataaatag agagggggcg cagcccctag gggcggccaa 6720
ccctaggtcg tccattatgg gccgcaattg ggctggctgt ctattcttcc gggccttcgt 6780
tctttaacaa catgatgtag ttcaattctc ttgcacgggc cggagtcact ggcccagggtg 6840
gaaggggtgg cgctgggct ggagaaggtg ctgctgggtg agatgtgctc gtgttggtgt 6900
tgatgtccgc atcaa 6915

```

<210> SEQ ID NO: 4

<211> Comprimento: 7572

<212> Tipo: DNA

<213> Organismo: Zea mays

<220>

<221> Nome: Origem

<222> Localização: (1)...(7572)

<223> Outras Informações: CRM2

<400> SEQ ID NO. 4

```

tgatgaag atccacacta ctgatgcatc tataccaata caagtaccaa tttctggctc 60
cattactcgc gctcgtgctc gtcaactcaa ccatcagggtg attacactct tgagttcatg 120
tccatcatat ttagaccatg gagaccctgt cactcttggt ttgcttagga atcagggaga 180
agaccgaaag ggaaaaggat ttgaacatgc tggattcggga ctgcagaaga acaccaactt 240
gtgacggtea ccacggtcag atgcccgtc ggattggaat gttcaagcac aacatggaaa 300
gcttatcaag tctactttca tatggatccg gaattatagt catatctggt ctgaggccgc 360
cgtaatcatt gttttcttac cgagacattt cctgcctttt ctgcccattg tgctgcgtca 420
ccctattttg gcccaatggg tegtgtatca agttagggtcc attagggagc catcctaggg 480
ttgcagcaga accccaatac ccttgtggtc gtccctccat gtttataaac ccctagccg 540
ccaccaagaa cagcgggttt tgtttagatc aagttagct ctcgctactt gcttgcaagc 600
gcgcgtgcta gttcagccgc cegtcttctt gtcttcggaa cccaccata ttggagtttg 660
atctttaaac ctacatttag atctggtaat tcagtaactg tctacttgt tcttgctag 720
tcttcgattg cttgcaggac gagtgcccta gtggccaggg tgtcacgctc cacaagatcg 780
tgacagccat aggaggtggt gtatcgggtg ctaaggcgca gcgtctttgg aaggctgtag 840
tcgggccgtg aacgtcgtct cctcccccaa tcgagttatt ccacaccctc tcatcgaag 900
atcgggcaat cacccaacgg gtgcacatca gttggtaatc agagcaaggt ttatcgggtg 960
gagatttact tttcttctgt gttttcttat ctctatagt ccagaaaaag ccaaaaaaat 1020
agtagattag tttaccgca atcctataaa ccattgagca tttactagta ctacttagtt 1080
agggttgtt gaggtttttg ttgcatcgtt tgtctcgagt tgctggtctt agtttattcc 1140
tttagagttt tgagttctac caggttttgg tcaccacgag atccaccatc accaaaaaca 1200
tctctggttc gtttttgcca ccacggatac atatcatatc cgatttgga gtttaaatc 1260
aatctggaaa gcttatctta tcttcttcc aacggatctg accttatctc aaaattcgtt 1320
ctgagcgtc cgcaatcctc gtagagattt ctggacttct tatattaaca agatttgta 1380
aatctgattt aaaggggttg ttagcaatat ctttattggt tgggttgta tagtgaaaa 1440
aagggtttag gccctgcaa aaaaaaacag aagaagaaga aaaaaaaga atagaaaaga 1500
aaaaaaagga aaagaagaag aagggggctg aatctctaaa tctgttggtt ctctttgtgc 1560
tgtgctagtt gttcttttct agtgactacc tttgtgctca ggctcacgctc tctagcctgg 1620
tttagcctag gaccagcaca gtaccaccgt tgaacgatta ttcagcttgc ttttgtaact 1680
aacgtggtac tagtgtattc cttgcttcag cccacctaca actctacata tttcgactac 1740
agtttgacag gctgtgttgc tgcggcaccg atacacttat tccacggttg cagacttgtt 1800
ggttgctgac cctcctggtt gtgcaaggta agaattggta agagcttggt tggcaggttg 1860
agagtgagcg ccttgcaagta gctacatcct aatagttgta gagtttttat tcttcacat 1920
tttttttct tgttgctctt gttcgtctaa ccatggcagg attggaggtt gatgatgctt 1980
ctcgaatat gccacactct cctcgcacca agggatcat acaacacttt gtaaggctgg 2040
tgaaaacgca cacggaaggt cttgataatg acatgcaggt gacgaatgaa aagatggggc 2100
aattggaggc cacacagatc gacacaaaca ccaacttgc aatgtggaa atgacagttg 2160
ctcatattga caagagcctt gtcgactct tgaggcgatt tgatgagatg catgctaata 2220
ccaatgggtg gcgtgatgag ggcgcccgaag gtaactggga tgactatggt gctgatactg 2280
aacaagatga ccaagaagca cctaactgcc ggcgactacg tactaacctg agaggatgg 2340
gtggttttca ccgacgtgag gtacatggta atgatgatgc ttttagtaag gttaaattta 2400

```



aaatacctcc	ttttgatggt	aaatatgacc	ctgatgctta	cattacttgg	gagattgceg	2460
ttgatcaaaa	gtttgcatgc	catgaatttc	ctgagaatgc	gcggggttaga	gctgctacta	2520
gtgagtttac	tgaatttgct	tctgtttggt	ggatagaaca	tggtaagaag	aatcctaata	2580
acatgccaca	aacttgggat	gcgttgaaac	gggtcatgcg	ggctagattt	gttccttctt	2640
attatgcacg	tgatatgta	aacaagttgc	aacaattgag	acaggggtact	aaaagtgtag	2700
aagaatatta	tcaggaatta	caaatgggta	tgctgcggtg	taacatagag	gaggggtgagg	2760
aatctgctat	ggctagattt	ttgggcgggt	taaataggga	aattcaggac	atccttgctt	2820
ataaagatta	tgctaagtga	acccgattgt	ttcatcttgc	ttgcaaagct	gaaaggggaag	2880
tgcaggggacg	acgtgctagt	gcaaggtcta	atgtttctgc	aggaaaatct	acaccatggc	2940
aacagcgcac	gactacgtcc	atgaccggcc	gtacactagc	accaactccc	togccaagtc	3000
gaccagcacc	cccgccttcc	tcagcgcaca	aaccacgtgc	atcttccaca	aattcagcaa	3060
ccaaactctgc	ccagaaacca	gcaggtagtg	cctcttcagt	agcctccacg	ggtagaacia	3120
gagatgttct	gtgttatcga	tgcaagggct	atggacacgt	gcagcgtgat	tgctcctaata	3180
agcgtgtttt	ggtggtaaaa	gacgatgggtg	ggtattctct	tgctagtgat	ttggatgaag	3240
ctacacttgc	tttgcttgcg	gctgatgatg	caggcactaa	ggaaccaccc	gaagaacaga	3300
ttggtgcaga	tgatgcagag	cattatgaga	gcctcattgt	acagcgtgtg	cttagtgcac	3360
aaatggagaa	ggcagagcag	aatcagcgcac	atacgttgtt	tcaaacaaag	tgtgtcatta	3420
aggagcgttc	atgtcgtttg	atcattgatg	gaggtagctg	caacaacttg	gctagcagcg	3480
acatggttga	gaagcttgca	cttacgacca	aaccgcacc	gcatccatat	cacattcaat	3540
ggctcaacaa	tagtggtaag	gtcaaggtaa	ccaagttggt	acgaattaat	tttgcatttg	3600
gttcatatcg	tgatgttgtt	gactgtgatg	ttgtgcctat	ggatgcttgt	aatattctgc	3660
taggtagacc	atggcaattt	gattcagatt	gtatgcatca	tggtagatca	aatcaatatt	3720
ctctcataca	ccatgataag	aaaattatct	tgcttcccat	gtcccctgag	gctattgtgc	3780
gtgatgatgt	tgctaaagct	accaaagcta	aaactgagaa	caacaagaat	attaaagttg	3840
ttggttaata	caaagatggg	ataaaaattga	aaggacattg	cttgcttgca	acaaaaactg	3900
atgtaaatga	attatttgct	tccactactg	ttgcctacgc	cttggtatgc	aaggatgctt	3960
tgatttcaat	tcaagatatg	cagcattctt	tgctctctgt	tattactaac	atthttgcagg	4020
agtattctga	tgtatttcca	agttagatac	cagaggggct	gccacctata	cgagggattg	4080
agcaccaaat	tgatcttatt	cctgggtgat	ctttgcgaa	tcgtgcgcca	tataggacia	4140
atccagagga	aacaaaagaa	atcagcgcac	aagtgcaga	actactcgac	aaaggttacg	4200
tgcgtgagtc	tcttagtccg	tgtgctgttc	cggttattht	agtgcctaaa	aaagatggaa	4260
catggcgtat	gtgtgttgat	tgtagggcta	ttaataatat	cacgatacgt	tatcgacacc	4320
ctattccacg	tttagatgat	atgcttgatg	aattgagtgg	tgccattgtc	ttttctaaag	4380
ttgatttgcg	tagtgggtac	caccagattc	gtatgaaatt	gggagatgaa	tggaaaactg	4440
ctttcaaaac	taagttcgga	ttgtatgagt	ggttagtcat	gccttttggg	ttaactaatg	4500
cacctagcac	tttcatgaga	ttaatgaacg	aggttttgcg	tgcttctatt	ggaaaatttg	4560
tgtagtata	ctttgatgac	atattaatct	acagcaaact	tatggatgaa	catggtgatc	4620
acatgcgtgc	tgtttttaat	gctttacgag	atgcacgttt	atthggtaac	cttgagaagt	4680
gcacattttg	caccgatcga	gtttcgtttc	ttggttatgt	tgtgactcca	caggggaattg	4740
aggttgatca	agccaaggta	gaagcgatac	atggatggcc	tatgccaag	actatcacac	4800
agggtcggag	tttctagga	cttgctggct	tctatcgccg	ttttgtgaag	gactttagca	4860
ccattgctgc	acctttgaat	gagcttacga	agaagggagt	gcattttagt	tggggcaaag	4920
tacaagagca	cgctttcaac	gtgctgaaa	ataagttgac	acatgcacct	ctcctccaac	4980
ttcttgattt	taataagact	tttgagcttg	aatgtgatgc	tagtgggaatt	ggattgggtg	5040
gtgttttggt	acaagaaggc	aaacctggtg	catattttag	tgaaaaattg	agtgggtctg	5100
ttctaaatta	ttctacttat	gataaggaat	tatatgctct	tgtgcaaca	ttagaacat	5160
ggcagcatta	tttgtggccc	aaagagtttg	ttattcattc	tgatcatgaa	tctttgaaac	5220
atattcgtag	tcaaggaaaa	ctgaaccgta	gacatgctaa	gtgggttgaa	tttatcgaa	5280
cgtttcttca	tgttattaag	cacaagaaa	gaaaagagaa	tatcattgct	gacgctttgt	5340
ctaggagata	tactttgctg	aatcaacttg	actacaaaat	ctttggatta	gagacgatta	5400
aagaccaata	tgttcatgat	gctgatttta	aagatgtggt	gctgcattgt	aaagatggga	5460
aaggatggaa	caaatatata	gtttagtgatg	ggtttgtggt	tagagctaac	aagctatgca	5520
ttccagctag	ctccgttcgt	ttgttggtgt	tacaggaagc	acatggaggt	ggcttaatgg	5580
gacattttgg	agcaaagaaa	acggaggaca	tacttgctgg	tcatttcttt	tggcccaaga	5640
tgagaagaga	tgtggtgaga	ttggttgctc	gttgacgac	atgccaaaag	gcgaagtcac	5700
ggtaaatcc	acacggtttg	tatttgctc	taccggttcc	tagtgctcct	tgggaagata	5760
tttctatgga	ttttgtgctg	ggattgccta	ggactaggaa	gggacgtgat	agtgtgtttg	5820
tggttggtga	tagattttct	aagatggcac	atthcatacc	atgtcataaa	actgacgatg	5880
ctactcatat	tgctgatttg	ttctttctgtg	aaattggtcg	cttgcatggg	tgcccaaca	5940
caatcgtttc	tgatcgtgat	gctaaatttc	ttagtcatth	ttggaggact	ttgtgggcaa	6000
aatgggggac	taagctthta	ttthctacta	catgtcatcc	tcaaactgat	ggtcaaac	6060

```

aagttgtgaa tagaactttg tctactatgt taagggcagt tctaaagaag aatattaaga 6120
tgtgggagga ctgtttgcct catattgaat ttgcttataa tcgatcattg cattctacta 6180
caaagatgtg cccatttcag attgtatag gtttgttacc tcgtgctcct attgatttaa 6240
tgcctttgcc atcttctgaa aaactaaatt ttgatgctac taggcgtgct gaattgatgt 6300
taaaactgca cgaaactact aaagaaaaca tagagcgtat gaatgctaga tataagtttg 6360
ctagtgataa aggtagaaaag gaaataaatt ttgaacctgg agatttagtt tggttgcatt 6420
tgagaaaagga aaggtttcct gaattacgaa aatctaaatt gttgcctcga gccgatggac 6480
cgtttaaagt gctagagaaa attaacgaca atgcatatag gctagatctg cctgcagact 6540
ttgggggtag cccacattt aacattgcag atttaaagcc ctacttggga gaggaagttg 6600
agcttgagtc gaggacgact caaatgcaag aaggggagaa tgatgaagac atccacacta 6660
ctgatgcatac tataccaata caagtacaa tttctggtcc cattactcgc gctcgtgctc 6720
gtcaactcaa ccatcaggtg attacactct tgagttcatg tccatcatat ttagagccat 6780
ggagaccogt gcactcttgt tttgcttagg aatcagggag aagaccgaaa gggaaaagga 6840
tttgaacatg ctggattcgg actgcagaag aacaccaact tgtgacggtc accacgggtca 6900
gatgcgggct cggattggaa tgttcaagca caacatggaa agcttatcaa gtctactttc 6960
atatggatcc ggaattatag tcatatctgt tctgaggccg ccgtaatcat tgttttctta 7020
ccgagacatt tctgccttt tctgcccag gtgctgcgtc accctatfff ggcccaatgg 7080
gtcgtgtatc aagttaggtc cattagggag gcatectagg gttgcagcac gaccccaata 7140
cccttggtg cgtcctccca tgtttataaa ccccctagcc gccaccaaga acagcggggt 7200
ttgtttgat caagtttagc tctcgtact tgcttgtaag cgcgcgtgct agttcagccg 7260
cccgtcttct tgtcttcgga accccacat attggagttt gattttgaaa cctacattta 7320
gatctggtaa ttcagtactt gttctacttg ttcttgctag ttcttogatt gcttgcagga 7380
cgagtgcctt agtggccagg gtgtcacgct ccacaagatc gtgacagcca taggaggtgg 7440
tgtatcgggt gctaaggcgc agcgtctttg gaaggctgta gtcgggcccgt gaacgtcgtc 7500
tctccccca atcgagttat tccacacct ctcatcgaag gatcgggcaa tcacccaacg 7560
ggtgcacatc ag 7572

```

<210> SEQ ID NO: 5  
<211> Comprimento: 32  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Overgo telomere primer1

<400> SEQ ID NO. 5  
agggtttagg gtttaggggt tagggtttag gg 32

<210> SEQ ID NO: 6  
<211> Comprimento: 30  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Overgo telomere primer2

<400> SEQ ID NO. 6  
ccctaaacc taaaccctaa accctaaacc 30

<210> SEQ ID NO: 7  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Primer CentC-OVG-1-40-F Biocode 65644

<400> SEQ ID NO. 7  
ggttccggtg gcaaaaactc gtgc 24

<210> SEQ ID NO: 8  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Primer CentC-OVG-1-40-R Biocode 65645

<400> SEQ ID NO. 8  
tgtcggtgca tacaaagcac gagt 24

<210> SEQ ID NO: 9  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Primer CentC-OVG-51-90-F Biocode 65646

<400> SEQ ID NO. 9  
gaatgggtga cgtgcgacaa cgaa 24

<210> SEQ ID NO: 10  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Primer CentC-OVG-51-90-R Biocode 65647

<400> SEQ ID NO. 10  
ggtggtttct cgcaatttcg ttgt 24

<210> SEQ ID NO: 11  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Primer CentC-OVG-101-140-F Biocode 65648

<400> SEQ ID NO. 11  
gttttgacc taaagtagtg gatt 24

<210> SEQ ID NO: 12  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Primer CentC-OVG-101-140-R Biocode 104790

<400> SEQ ID NO. 12  
cacaacgaac atgcccaatc cact 24

<210> SEQ ID NO: 13  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Primer CRM1-LTR-OVG1-F Biocode 69509

<400> SEQ ID NO. 13  
cttggctcttg gacagtacct cact 24

<210> SEQ ID NO: 14  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Primer CRM1-LTR-OVG2-F Biocode 69510

<400> SEQ ID NO. 14  
cccttgcgat cggactacga cgag 24

<210> SEQ ID NO: 15  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Primer CRM1-LTR-OVG3-F Biocode 69511

<400> SEQ ID NO. 15  
tcacgaagat cgtttcctgt gcgc 24

<210> SEQ ID NO: 16  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Primer CRM1-LTR-OVG4-F Biocode 69512

<400> SEQ ID NO. 16  
cagcgcagat tagcgcgtgt tcga 24

<210> SEQ ID NO: 17  
<211> Comprimento: 24

<212> Tipo: DNA  
<213> Organismo: Seqüência Artificial

<220> Características:  
<223> Outras Informações: Primer CRM1-LTR-OVG5-F Biocode 69513

<400> SEQ ID NO. 17  
ccaaccctag gtcgtccatt atgg 24

<210> SEQ ID NO: 18  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Seqüência Artificial

<220> Características:  
<223> Outras Informações: Primer CRM1-LTR-OVG6-F Biocode 69514

<400> SEQ ID NO. 18  
ttcaattctc ttgcacgggc ccga 24

<210> SEQ ID NO: 19  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Seqüência Artificial

<220> Características:  
<223> Outras Informações: Primer CRM1-LTR-OVG1-R Biocode 69515

<400> SEQ ID NO. 19  
tcaggtctac ttcacagtg aggt 24

<210> SEQ ID NO: 20  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Seqüência Artificial

<220> Características:  
<223> Outras Informações: Primer CRM1-LTR-OVG2-R Biocode 69516

<400> SEQ ID NO. 20  
tggcgctcg ggcttgctcg tcgt 24

<210> SEQ ID NO: 21  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Seqüência Artificial

<220> Características:  
<223> Outras Informações: Primer CRM1-LTR-OVG3-R Biocode 69517

<400> SEQ ID NO. 21  
tggtcgttct tcgattgcgc acag 24

<210> SEQ ID NO: 22  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Seqüência Artificial

<220> Características:  
<223> Outras Informações: Primer CRM1-LTR-OVG4-R Biocode 69518

<400> SEQ ID NO. 22  
ttagccttag ctactctcga acac 24

<210> SEQ ID NO: 23  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Primer CRM1-LTR-OVG5-R Biocode 69519

<400> SEQ ID NO. 23  
ccagcccaat tgcggcccat aatg 24

<210> SEQ ID NO: 24  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Primer CRM1-LTR-OVG6-R Biocode 69520

<400> SEQ ID NO. 24  
cacctgggcc agtgactcgg gcc 24

<210> SEQ ID NO: 25  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Primer CRM2-LTR-OVG1-F Biocode 69521

<400> SEQ ID NO. 25  
tgatgaagac atccacacta ctga 24

<210> SEQ ID NO: 26  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Primer CRM2-LTR-OVG2-F Biocode 69522

<400> SEQ ID NO. 26  
ttgaacatgc tggattcgga ctgc 24

<210> SEQ ID NO: 27  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:

<223> Outras Informações: Primer CRM2-LTR-OVG3-F Biocode 69523

<400> SEQ ID NO. 27

ctgcccattgg tgctgcgtca ccct

24

<210> SEQ ID NO: 28

<211> Comprimento: 24

<212> Tipo: DNA

<213> Organismo: Sequência Artificial

<220> Características:

<223> Outras Informações: Primer CRM2-LTR-OVG4-F Biocode 69524

<400> SEQ ID NO. 28

gcgcgtgcta gttcagccgc ccgt

24

<210> SEQ ID NO: 29

<211> Comprimento: 24

<212> Tipo: DNA

<213> Organismo: Sequência Artificial

<220> Características:

<223> Outras Informações: Primer CRM2-LTR-OVG5-F Biocode 69525

<400> SEQ ID NO. 29

gtatcggttg ctaaggcgc gcgt

24

<210> SEQ ID NO: 30

<211> Comprimento: 24

<212> Tipo: DNA

<213> Organismo: Sequência Artificial

<220> Características:

<223> Outras Informações: Primer CRM2-LTR-OVG1-R Biocode 69526

<400> SEQ ID NO. 30

tattggtata gatgcatcag tagt

24

<210> SEQ ID NO: 31

<211> Comprimento: 24

<212> Tipo: DNA

<213> Organismo: Sequência Artificial

<220> Características:

<223> Outras Informações: Primer CRM2-LTR-OVG2-R Biocode 69527

<400> SEQ ID NO. 31

aagttggtgt tcttctgcag tccg

24

<210> SEQ ID NO: 32

<211> Comprimento: 25

<212> Tipo: DNA

<213> Organismo: Sequência Artificial

<220> Características:

<223> Outras Informações: Primer CRM2-LTR-OVG3-R Biocode 69528

<400> SEQ ID NO. 32  
cccattgggc caaaataggg tgacg 25

<210> SEQ ID NO: 33  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Primer CRM2-LTR-OVG4-R Biocode 69529

<400> SEQ ID NO. 33  
ttccgaagac aagaagacgg gcgg 24

<210> SEQ ID NO: 34  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Primer CRM2-LTR-OVG5-R Biocode 69530

<400> SEQ ID NO. 34  
ctacagcctt ccaagacgc tgcg 24

<210> SEQ ID NO: 35  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Primer CentA-LTR-OVG1-F Biocode 69531

<400> SEQ ID NO. 35  
tgatgagaac ataaccgca caga 24

<210> SEQ ID NO: 36  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Primer CentA-LTR-OVG2-F Biocode 69532

<400> SEQ ID NO. 36  
aggatgatga ggacatcact gccca 24

<210> SEQ ID NO: 37  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Primer CentA-LTR-OVG3-F Biocode 69533



<400> SEQ ID NO. 37  
aaccatctag aatttgagaa ggca 24

<210> SEQ ID NO: 38  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Primer CentA-LTR-OVG4-F Biocode 69534

<400> SEQ ID NO. 38  
gtccagaaac tgccgagtga actc 24

<210> SEQ ID NO: 39  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Primer CentA-LTR-OVG5-F Biocode 69535

<400> SEQ ID NO. 39  
gagagagttt cgttctccat taga 24

<210> SEQ ID NO: 40  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Primer CentA-LTR-OVG6-F Biocode 69536

<400> SEQ ID NO. 40  
gttcttgctt gttctcgatt gctt 24

<210> SEQ ID NO: 41  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Primer CentA-LTR-OVG7-F Biocode 69537

<400> SEQ ID NO. 41  
ttggttgagg tagtcgggca gccca 24

<210> SEQ ID NO: 42  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Primer CentA-LTR-OVG1-R Biocode 69538

<400> SEQ ID NO. 42  
cattaacatg gtcatatctg tgcg 24

<210> SEQ ID NO: 43  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Primer CentA-LTR-OVG2-R Biocode 69539

<400> SEQ ID NO. 43  
tggtgtggtg tattgatggc agtg 24

<210> SEQ ID NO: 44  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Primer CentA-LTR-OVG3-F Biocode 69540

<400> SEQ ID NO. 44  
cttttattgc cttggtgcct tct 23

<210> SEQ ID NO: 45  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Primer CentA-LTR-OVG4-R Biocode 69541

<400> SEQ ID NO. 45  
gacttgggta gagcaggagt tcac 24

<210> SEQ ID NO: 46  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Primer CentA-LTR-OVG5-R Biocode 69542

<400> SEQ ID NO. 46  
aggaatagaa aggagttcta atgg 24

<210> SEQ ID NO: 47  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Primer CentA-LTR-OVG6-R Biocode 69543

<400> SEQ ID NO. 47  
acagccttga acctgcaagc aatc 24

<210> SEQ ID NO: 48  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Primer CentA-LTR-OVG7-R Biocode 69544

<400> SEQ ID NO. 48  
tgttggagaa cgacggttggc tgcc 24

<210> SEQ ID NO: 49  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Primer Cent4-250-OVG1-F Biocode 69555

<400> SEQ ID NO. 49  
taagtgcaaa ccattgttaa attt 24

<210> SEQ ID NO: 50  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Primer Cent4-250-OVG2-F Biocode 69556

<400> SEQ ID NO. 50  
caciaaccct taactcgaaa ctat 24

<210> SEQ ID NO: 51  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Primer Cent4-250-OVG3-F Biocode 69557

<400> SEQ ID NO. 51  
atcgaaagat aactcatatg gctt 24

<210> SEQ ID NO: 52  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Primer Cent4-250-OVG4-F Biocode 69558

<400> SEQ ID NO. 52  
tccactaaag aaccaagatt gtga 24

<210> SEQ ID NO: 53  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Primer Cent4-250-OVG1-F Biocode 69559

<400> SEQ ID NO. 53  
aattgtacta tctctaaaat ttaa 24

<210> SEQ ID NO: 54  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Primer Cent4-250-OVG2-R Biocode 69560

<400> SEQ ID NO. 54  
tttagggttt ggggttatag tttc 24

<210> SEQ ID NO: 55  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Primer Cent4-250-OVG3-R Biocode 69561

<400> SEQ ID NO. 55  
gaccataatg gtcaaaaagc cata 24

<210> SEQ ID NO: 56  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Primer Cent4-250-OVG4-R Biocode 69562

<400> SEQ ID NO. 56  
atatgttgga cacaaatcac aatc 24

<210> SEQ ID NO: 57  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Primer 18-26SrDNANTS-OVG1-F Biocode 69634

<400> SEQ ID NO. 57  
ccggaataa gcaaagtcca agcg 24

<210> SEQ ID NO: 58  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Primer 18-26SrDNANTS-OVG2-F Biocode 69635

<400> SEQ ID NO. 58  
tatgtcttgg gtgaaggcca tggc 24

<210> SEQ ID NO: 59  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Primer 18-26SrDNANTS-OVG3-F Biocode 69636

<400> SEQ ID NO. 59  
cgcaaggcga cgggcggcat ggct 24

<210> SEQ ID NO: 60  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Primer 18-26SrDNANTS-OVG4-F Biocode 69637

<400> SEQ ID NO. 60  
cgaggggttc cccatggcgc acgg 24

<210> SEQ ID NO: 61  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Primer 18-26SrDNANTS-OVG1-R Biocode 69638

<400> SEQ ID NO. 61  
tcggtgtctt tccacagct tgga 24

<210> SEQ ID NO: 62  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Primer 18-26SrDNANTS-OVG2-R Biocode 69639

<400> SEQ ID NO. 62  
gttttcctc cgttccgcca tgcc 24

<210> SEQ ID NO: 63  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Primer 18-26SrDNANTS-OVG3-R Biocode 69640

<400> SEQ ID NO. 63  
agacgcaagg ccgaacagcc atgc 24

<210> SEQ ID NO: 64  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Primer 18-26SrDNANTS-OVG4-R Biocode 69641

<400> SEQ ID NO. 64  
ggcctcagtt ttcggcccgt gcgc 24

<210> SEQ ID NO: 65  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Primer Subtelomere-266 Biocode 74794

<400> SEQ ID NO. 65  
gacacatggt tttgtcgtcg aaca 24

<210> SEQ ID NO: 66  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Primer Subtelomere-266 Biocode 74795

<400> SEQ ID NO. 66  
ggaggcacga aatcgctggt cgac 24

<210> SEQ ID NO: 67  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Primer Subtelomere-266 Biocode 74796

<400> SEQ ID NO. 67  
cgaccgccac ccatgatttg acca 24

<210> SEQ ID NO: 68  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Primer Subtelomere-266 Biocode 74797

<400> SEQ ID NO. 68  
accttaccag tctctatggt caaa 24

<210> SEQ ID NO: 69  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Primer Subtelomere-266 Biocode 74799

<400> SEQ ID NO. 69  
tcccgtgagc tatagcacac gttt 24

<210> SEQ ID NO: 70  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Primer Subtelomere-266 Biocode 74800

<400> SEQ ID NO. 70  
acacgttttc atggccgagc gacc 24

<210> SEQ ID NO: 71  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Primer Subtelomere-266 Biocode 74801

<400> SEQ ID NO. 71  
ccgtgttctt ccacacgtgt tttt 24

<210> SEQ ID NO: 72  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Primer Subtelomere-266 Biocode 74802

<400> SEQ ID NO. 72  
aaggtgctcc ggggacaaaa acac 24

<210> SEQ ID NO: 73  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Primer Subtelomere-266 Biocode 74803

<400> SEQ ID NO. 73  
ttggcctccc gcgagctata tcac 24

<210> SEQ ID NO: 74  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Primer Subtelomere-266 Biocode 74803

<400> SEQ ID NO. 74  
ttggccacgg aaatgtgtga tata 24

<210> SEQ ID NO: 75  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Primer Subtelomere-266 Biocode 74805

<400> SEQ ID NO. 75  
ttatgtatcc gacctgccac cttc 24

<210> SEQ ID NO: 76  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Primer Subtelomere-266 Biocode 74806

<400> SEQ ID NO. 76  
ctccccggtc taaaacgaag gtgg 24

<210> SEQ ID NO: 77  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Primer Subtelomere-266 Biocode 74807



<400> SEQ ID NO. 77  
gccacccgtg agctatagca cacg 24

<210> SEQ ID NO: 78  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Seqüência Artificial

<220> Características:  
<223> Outras Informações: Primer Subtelomere-266 Biocode 74808

<400> SEQ ID NO. 78  
taggtttcca taaaatcgtg tgct 24

<210> SEQ ID NO: 79  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Seqüência Artificial

<220> Características:  
<223> Outras Informações: Primer 180-knob-OVG-21-60-F Biocode 65650

<400> SEQ ID NO. 79  
tgtcgaaaat agccatgaac gacc 24

<210> SEQ ID NO: 80  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Seqüência Artificial

<220> Características:  
<223> Outras Informações: Primer 180-knob-OVG-21-60-r Biocode 65651

<400> SEQ ID NO. 80  
cggtattatt ggaaatggtc gttc 24

<210> SEQ ID NO: 81  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Seqüência Artificial

<220> Características:  
<223> Outras Informações: Primer 180-knob-OVG-71-110-F Biocode 65652

<400> SEQ ID NO. 81  
cctacggatt tttgaccaag aaat 24

<210> SEQ ID NO: 82  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Seqüência Artificial

<220> Características:  
<223> Outras Informações: Primer 180-knob-OVG-71-110-r Biocode 65653

<400> SEQ ID NO. 82  
atttctagtg gagaccattt cttg 24

<210> SEQ ID NO: 83  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Primer 180-knob-OVG-141-180-F Biocode 65654

<400> SEQ ID NO. 83  
atgtgggggtg aggtgtatga gcct 24

<210> SEQ ID NO: 84  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Primer 180-knob-OVG-141-180-r Biocode 65655

<400> SEQ ID NO. 84  
atgagcctct ggtcgatgat caat 24

<210> SEQ ID NO: 85  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Primer 5S-rDNA-OVG-1-40-F Biocode 65656

<400> SEQ ID NO. 85  
ggatgcatc ataccagcac taaa 24

<210> SEQ ID NO: 86  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Primer 5S-rDNA-OVG-1-40-r Biocode 65657

<400> SEQ ID NO. 86  
tgatgggatc cgggtgcttta gtgc 24

<210> SEQ ID NO: 87  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Primer 5S-rDNA-OVG-61-100-F Biocode 65658

<400> SEQ ID NO. 87  
cttgggcgag agtagtacta ggat 24

<210> SEQ ID NO: 88  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Primer 5S-rDNA-OVG-61-100-R Biocode 65659

<400> SEQ ID NO. 88  
tcccaggagg tcacccatcc tagt 24

<210> SEQ ID NO: 89  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Primer 5S-rDNA-OVG-161-200-F Biocode 65660

<400> SEQ ID NO. 89  
accatagtaa aaatgggtga ccgt 24

<210> SEQ ID NO: 90  
<211> Comprimento: 23  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Primer 5S-rDNA-OVG-161-200-R Biocode 65661

<400> SEQ ID NO. 90  
taatttaaca cgagaacggt cac 23

<210> SEQ ID NO: 91  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Primer 5S-rDNA-OVG-261-230-F Biocode 65662

<400> SEQ ID NO. 91  
ccgtgggcca gccgagcacg gagg 24

<210> SEQ ID NO: 92  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Primer 5S-rDNA-OVG-261-230-F Biocode 65662

<400> SEQ ID NO. 92  
tcctcttatg cccacacctc cgtg 24

<210> SEQ ID NO: 93  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Primer 350-knob-OVG-31-70-F Biocode 65664

<400> SEQ ID NO. 93  
ctcaaagac gtttctatga tatt 24

<210> SEQ ID NO: 94  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Primer 350-knob-OVG-31-70-R Biocode 65665

<400> SEQ ID NO. 94  
tgaatacaat gccctcaata tcat 24

<210> SEQ ID NO: 95  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Primer 350-knob-OVG-121-160-F Biocode 65666

<400> SEQ ID NO. 95  
ctaggtttcc tataatcccc tcta 24

<210> SEQ ID NO: 96  
<211> Comprimento: 23  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Primer 350-knob-OVG-121-160-R Biocode 65667

<400> SEQ ID NO. 96  
ctaggtatgc cttgaataga ggg 23

<210> SEQ ID NO: 97  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Primer 350-knob-OVG-161-200-F Biocode 65668

<400> SEQ ID NO. 97  
atggttggtta tgtccactca agta 24

<210> SEQ ID NO: 98  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Primer 350-knob-OVG-161-200-R Biocode 65669

<400> SEQ ID NO. 98  
atggtgtacg gtgttttact tgag 24

<210> SEQ ID NO: 99  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Primer 350-knob-OVG-261-3000-F Biocode 65670

<400> SEQ ID NO. 99  
gtgagatctg tccaaacata gggt 24

<210> SEQ ID NO: 100  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Primer 350-knob-OVG-261-300-R Biocode 65671

<400> SEQ ID NO. 100  
ggtgccttac aaccgtaacc tatg 24

<210> SEQ ID NO: 101  
<211> Comprimento: 40  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Primer to b010.m7 fis31

<400> SEQ ID NO. 101  
gcaaaacttta tgtgatccct tctcgtgta acgagatgag 40

<210> SEQ ID NO: 102  
<211> Comprimento: 40  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Primer to b108.h15 fis47

<400> SEQ ID NO. 102  
gggacggcaa gtcacggtaa gaccagtcca accgaatgat 40

<210> SEQ ID NO: 103  
<211> Comprimento: 40  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Primer to Cen3n.pk0001.g11

<400> SEQ ID NO. 103  
ccaaacttgc tgagattact gggcaatctg ttcgctcgca 40

<210> SEQ ID NO: 104  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Overgo probe primer 23715-3101-3200f Biocode 103022

<400> SEQ ID NO. 104  
ccaggtagtt tgaaacagta ttct 24

<210> SEQ ID NO: 105  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Overgo probe primer 23715-3101-3600f Biocode 103023

<400> SEQ ID NO. 105  
ataaaggaaa agggcaaacc aaac 24

<210> SEQ ID NO: 106  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Overgo probe primer 23715-1401-1500f Biocode 103024

<400> SEQ ID NO. 106  
gatgccaca ttatagtgat tagc 24

<210> SEQ ID NO: 107  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Overgo probe primer 23715-2901-3000f Biocode 103025

<400> SEQ ID NO. 107  
ccacatatag ctgctgcata tgcc

24

<210> SEQ ID NO: 108  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Overgo probe primer 23715-3701-3800f Biocode 103026

<400> SEQ ID NO. 108  
cggatctaac acaaacatga acag

24

<210> SEQ ID NO: 109  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Overgo probe primer 23715-1-100f Biocode 103027

<400> SEQ ID NO. 109  
cgatgaattt tctcgggtgt tctc

24

<210> SEQ ID NO: 110  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Overgo probe primer 23715-3101-3200f Biocode 103022

<400> SEQ ID NO. 110  
cctgcagccc taataattca gaag

24

<210> SEQ ID NO: 111  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Overgo probe primer 23715-301-400f Biocode 103029

<400> SEQ ID NO. 111  
cacagtcgat gaatccagaa aagc

24

<210> SEQ ID NO: 112  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Overgo probe primer 23715-901-1000f Biocode 103030

<400> SEQ ID NO. 112  
gcgtgcaatc catcttggtc aatc

24

<210> SEQ ID NO: 113  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Overgo probe primer 23715-3201-3300f Biocode 103031

<400> SEQ ID NO. 113  
caaccacacc acatcatcac aacc

24

<210> SEQ ID NO: 114  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Overgo probe primer 23715-3601-3700f Biocode 103032

<400> SEQ ID NO. 114  
actggcaagt tagcaatcag aacg

24

<210> SEQ ID NO: 115  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Overgo probe primer 23715-4901-5000f Biocode 103033

<400> SEQ ID NO. 115  
catgaacgtg tcttcaacta gagg

24

<210> SEQ ID NO: 116  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Overgo probe primer 23715-4201-4300f Biocode 103034

<400> SEQ ID NO. 116  
gacggcgttt aacaggctgg catt

24

<210> SEQ ID NO: 117  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Overgo probe primer 23715-201-300f Biocode 103035



<400> SEQ ID NO. 117  
ccaagctctt cagcaatatac acgg

24

<210> SEQ ID NO: 118  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Overgo probe primer 23715-601-700f Biocode 103036

<400> SEQ ID NO. 118  
atactttctc ggcaggagca aggt

24

<210> SEQ ID NO: 119  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Overgo probe primer 23715-1001-1100f Biocode 103037

<400> SEQ ID NO. 119  
atccttggcg gcaagaaagc catc

24

<210> SEQ ID NO: 120  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Overgo probe primer 23715-1101-1200f Biocode 103038

<400> SEQ ID NO. 120  
gcaagctacc tgctttctct ttgc

24

<210> SEQ ID NO: 121  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Overgo probe primer 23715-1601-1700f Biocode 103039

<400> SEQ ID NO. 121  
gcttcttggc catgtagatg gact

24

<210> SEQ ID NO: 122  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Overgo probe primer 23715-1801-1900f Biocode 103040

<400> SEQ ID NO. 122  
ttcacgccga tgaacttcac cttg 24

<210> SEQ ID NO: 123  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Overgo probe primer 23715-5001-5087f Biocode 103041

<400> SEQ ID NO. 123  
aagcttgcca acgactacgc acta 24

<210> SEQ ID NO: 124  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Overgo probe primer 23715-401-500f Biocode 103042

<400> SEQ ID NO. 124  
ccctgatgct cttcgtccag atca 24

<210> SEQ ID NO: 125  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Overgo probe primer 23715-801-900f Biocode 103043

<400> SEQ ID NO. 125  
agagcagccg attgtctggt gtgc 24

<210> SEQ ID NO: 126  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Overgo probe primer 23715-1301-1400f Biocode 103044

<400> SEQ ID NO. 126  
caggatcccg taactataac ggtc 24

<210> SEQ ID NO: 127  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Overgo probe primer 23715-2801-2900f Biocode 103045

<400> SEQ ID NO. 127  
cgacctgcag aagtaacacc aaac

24

<210> SEQ ID NO: 128  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Overgo probe primer 23715-3401-3500f Biocode 103046

<400> SEQ ID NO. 128  
atctagaacg accgccaac caga

24

<210> SEQ ID NO: 129  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Overgo probe primer 23715-3801-3900f Biocode 103047

<400> SEQ ID NO. 129  
atttggggga gatctggttg tgtg

24

<210> SEQ ID NO: 130  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Overgo probe primer 23715-3901-4000f Biocode 103048

<400> SEQ ID NO. 130  
gagggggtgt ctatttatta cggc

24

<210> SEQ ID NO: 131  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Overgo probe primer 23715-4801-4900f Biocode 103049

<400> SEQ ID NO. 131  
catgcaagct gatctgagct tggc

24

<210> SEQ ID NO: 132  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Overgo probe primer 23715-2101-2200f Biocode 103050

<400> SEQ ID NO. 132  
tccatgcgca ccttgaagcg catg

24

<210> SEQ ID NO: 133  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Overgo probe primer 23715-501-600f Biocode 103051

<400> SEQ ID NO. 133  
ttccatccga gtacgtgctc gctc

24

<210> SEQ ID NO: 134  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Overgo probe primer 23715-1201-1300f Biocode 103052

<400> SEQ ID NO. 134  
atccactagt aacggccgcc agtg

24

<210> SEQ ID NO: 135  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Overgo probe primer 23715-4001-4100f Biocode 103053

<400> SEQ ID NO. 135  
gccacgcaat ttctggatgc cgac

24

<210> SEQ ID NO: 136  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Overgo probe primer 23715-701-800f Biocode 103054

<400> SEQ ID NO. 136  
cgatagccgc gctgcctcgt cttg

24

<210> SEQ ID NO: 137  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Overgo probe primer 23715-1901-2000f Biocode 103055

<400> SEQ ID NO. 137  
cacttgaagc cctcggggaa ggac

24

<210> SEQ ID NO: 138  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Overgo probe primer 23715-1701-1800f Biocode 103056

<400> SEQ ID NO. 138  
tccttcagct tcagggcctt gtgg

24

<210> SEQ ID NO: 139  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Overgo probe primer 23715-2001-2100f Biocode 103057

<400> SEQ ID NO. 139  
caccttggag ccgtactgga actg

24

<210> SEQ ID NO: 140  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Overgo probe primer 23715-2601-2700f Biocode 103058

<400> SEQ ID NO. 140  
tgcggctcgg tgcggaagtt cacg

24

<210> SEQ ID NO: 141  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Overgo probe primer 23715-4101-4200f Biocode 103059

<400> SEQ ID NO. 141  
acgcgacgct gctggttcgc tggc

24

<210> SEQ ID NO: 142  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Overgo probe primer 23715-3101-3200r Biocode 103060

<400> SEQ ID NO. 142  
cgttctagat cggagtagaa tact 24

<210> SEQ ID NO: 143  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Overgo probe primer 23715-3501-3600r Biocode 103061

<400> SEQ ID NO. 143  
tgtttcggtg catagggttt ggtt 24

<210> SEQ ID NO: 144  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Overgo probe primer 23715-1401-1500r Biocode 33332

<400> SEQ ID NO. 144  
gcacacatag tgacatgcta atca 24

<210> SEQ ID NO: 145  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Overgo probe primer 23715-2901-3000r Biocode 103062

<400> SEQ ID NO. 145  
gatatacttg gatgatggca tatg 24

<210> SEQ ID NO: 146  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Overgo probe primer 23715-3701-3800r Biocode 103063

<400> SEQ ID NO. 146  
cccggtagtt ctacttctgt tcat 24

<210> SEQ ID NO: 147  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Overgo probe primer 23715-1-100r Biocode 103064

<400> SEQ ID NO. 147  
attcgagcca atatgcgaga acac

24

<210> SEQ ID NO: 148  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:

<223> Outras Informações: Overgo probe primer 23715-101-200r Biocode 103065

<400> SEQ ID NO. 148  
gccttcttga cgagttcttc tgaa

24

<210> SEQ ID NO: 149  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:

<223> Outras Informações: Overgo probe primer 23715-301-400r Biocode 103066

<400> SEQ ID NO. 149  
atggtggaaa atggccgctt ttct

24

<210> SEQ ID NO: 150  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:

<223> Outras Informações: Overgo probe primer 23715-901-1000r Biocode 103067

<400> SEQ ID NO. 150  
gaggatcggt tcgcatgatt gaac

24

<210> SEQ ID NO: 151  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:

<223> Outras Informações: Overgo probe primer 23715-3201-3300r Biocode 103068

<400> SEQ ID NO. 151  
tgctttttgt tcgcttggtt gtga

24

<210> SEQ ID NO: 152  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:

<223> Outras Informações: Overgo probe primer 23715-3601-3700r Biocode 103069

<400> SEQ ID NO. 152  
acctgtacgt cagacacgtt ctga

24

<210> SEQ ID NO: 153  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Overgo probe primer 23715-4901-5000r Biocode 103070

<400> SEQ ID NO. 153  
aattaagtca ggcgcgcctc tagt

24

<210> SEQ ID NO: 154  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Overgo probe primer 23715-4201-4300r Biocode 103071

<400> SEQ ID NO. 154  
cttgtttcga gtagataatg ccag

24

<210> SEQ ID NO: 155  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Overgo probe primer 23715-201-300r Biocode 103072

<400> SEQ ID NO. 155  
acatagcgtt ggctaccctg gata

24

<210> SEQ ID NO: 156  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Overgo probe primer 23715-601-700r Biocode 103073

<400> SEQ ID NO. 156  
gatctcctgt catctcacct tgct

24

<210> SEQ ID NO: 157  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Overgo probe primer 23715-1001-1100r Biocode 103074



<400> SEQ ID NO. 157  
cctgcaaagt aaactggatg gctt

24

<210> SEQ ID NO: 158  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:

<223> Outras Informações: Overgo probe primer 23715-1101-1200r Biocode 103075

<400> SEQ ID NO. 158  
aagggaaaac gcaagcgcaa agag

24

<210> SEQ ID NO: 159  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:

<223> Outras Informações: Overgo probe primer 23715-1601-1700r Biocode 103076

<400> SEQ ID NO. 159  
tacctggtgg agttcaagtc catc

24

<210> SEQ ID NO: 160  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:

<223> Outras Informações: Overgo probe primer 23715-1801-1900r Biocode 103077

<400> SEQ ID NO. 160  
acggctgctt catctacaag gtga

24

<210> SEQ ID NO: 161  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:

<223> Outras Informações: Overgo probe primer 23715-5001-5087r Biocode 103078

<400> SEQ ID NO. 161  
tgaagctctt gttggctagt gcgt

24

<210> SEQ ID NO: 162  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:

<223> Outras Informações: Overgo probe primer 23715-401-500r Biocode 103079

<400> SEQ ID NO. 162  
gtcttgcga tcaggatgat ctgg 24

<210> SEQ ID NO: 163  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Overgo probe primer 23715-801-900r Biocode 103080

<400> SEQ ID NO. 163  
attcggctat gactgggcac aaca 24

<210> SEQ ID NO: 164  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Overgo probe primer 23715-1301-1400r Biocode 103081

<400> SEQ ID NO. 164  
cgcttcgcta ccttaggacc gtta 24

<210> SEQ ID NO: 165  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Overgo probe primer 23715-2801-2900r Biocode 103082

<400> SEQ ID NO. 164  
cgatgctcac cctggtggtt ggtg 24

<210> SEQ ID NO: 166  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Overgo probe primer 23715-3401-3500r Biocode 88145

<400> SEQ ID NO. 166  
ggttgtgatg atgtggtctg gttg 24

<210> SEQ ID NO: 167  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Overgo probe primer 23715-3801-3900r Biocode 103083

<400> SEQ ID NO. 167  
gttcggagcg cacacacaca caac

24

<210> SEQ ID NO: 168  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:

<223> Outras Informações: Overgo probe primer 23715-3901-4000r Biocode 103084

<400> SEQ ID NO. 168  
tttcccttcc tcgcccgccg taat

24

<210> SEQ ID NO: 169  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:

<223> Outras Informações: Overgo probe primer 23715-4801-4900r Biocode 103085

<400> SEQ ID NO. 169  
taaaacgacg gccagtgcca agct

24

<210> SEQ ID NO: 170  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:

<223> Outras Informações: Overgo probe primer 23715-2101-2200r Biocode 103086

<400> SEQ ID NO. 170  
acgtcatcac cgagttcatg cgct

24

<210> SEQ ID NO: 171  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:

<223> Outras Informações: Overgo probe primer 23715-501-600r Biocode 103087

<400> SEQ ID NO. 171  
agcgaaacat cgcacgcgac gacg

24

<210> SEQ ID NO: 172  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:

<223> Outras Informações: Overgo probe primer 23715-1201-1300r Biocode 103088

<400> SEQ ID NO. 172  
aagccgaatt ccagcacact ggcg

24

<210> SEQ ID NO: 173  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Overgo probe primer 23715-4001-4100r Biocode 103089

<400> SEQ ID NO. 173  
ttggacttgc tccgctgtcg gcat

24

<210> SEQ ID NO: 174  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Overgo probe primer 23715-701-800r Biocode 103090

<400> SEQ ID NO. 174  
tgcctgaat gaactgcaag acga

24

<210> SEQ ID NO: 175  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Overgo probe primer 23715-1901-2000r Biocode 103091

<400> SEQ ID NO. 175  
ccgactacaa gaagctgtcc ttcc

24

<210> SEQ ID NO: 176  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Overgo probe primer 23715-1701-1800r Biocode 103092

<400> SEQ ID NO. 176  
tgctgaaggg cgagaccac aagg

24

<210> SEQ ID NO: 177  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Overgo probe primer 23715-2001-2100r Biocode 103093

<400> SEQ ID NO. 177  
ggacatcctg tccccccagt tcca

24

<210> SEQ ID NO: 178  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Overgo probe primer 23715-2601-2700r Biocode 103094

<400> SEQ ID NO. 178  
acatcgagac ctccaccgtg aact

24

<210> SEQ ID NO: 179  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: Overgo probe primer 23715-4101-4200r Biocode 103095

<400> SEQ ID NO. 179  
agtctaacgg acaccaacca gcga

24

<210> SEQ ID NO: 180  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: PCR primer para bacm.pk108.h15.f sequência única  
sequence

<400> SEQ ID NO. 180  
gatcgtcgaa tgggaatcca tggg

24

<210> SEQ ID NO: 181  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: PCR primer para bacm.pk108.h15.r sequência única  
sequence

<400> SEQ ID NO. 181  
ccctgagtga accatttagg aagatcag

28

<210> SEQ ID NO: 182  
<211> Comprimento: 24  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: PCR primer para bacm.pk108.h15.2 FIS47.f sequência única

<400> SEQ ID NO. 182  
tgcaacatcc aaagacccaa catg 24

<210> SEQ ID NO: 183  
<211> Comprimento: 22  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: PCR primer para bacm.pk108.h15.2 FIS47.r sequência única

<400> SEQ ID NO. 183  
ttccaacatg gttggtggtc ag 22

<210> SEQ ID NO: 184  
<211> Comprimento: 26  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: PCR primer para bacm.pk010.m07. fis31.f sequência única

<400> SEQ ID NO. 184  
tgtcatgaca tcttgttgc accctg 26

<210> SEQ ID NO: 185  
<211> Comprimento: 222  
<212> Tipo: DNA  
<213> Organismo: Sequência Artificial

<220> Características:  
<223> Outras Informações: PCR primer para bacm.pk010.m07 fis31.r sequência única

<400> SEQ ID NO. 185  
aaaccggag tttctatgca gg 22

████████████████████

<211> Comprimento: 591  
<212> Tipo: DNA  
<213> Organismo: Zea mays

<220> Características:  
<221> Nome: misc\_features  
<222> Localização (1)...(591)  
<223> Outras Informações: n = a, t, c, ou g

<221> Nome: Origem

<222> Localização (1)...(591)

<223> Outras Informações: bacm.pk108.h15 sequência única

<400> SEQ ID NO. 186

```

ttgctcgttaa cagatgggttc angnnngatt gatcgtcgaa tgggaatcca tgggcaccca 60
cttgaatttc aggttttctt tttgctacac ctatgtatat tttctgtttc atacgggtct 120
tttttcccaa gttgattttt tgtgattggt tttgaggcac cttttaaaag aataaaatac 180
acaaacattc ttcaaattgt ctgggaatgt catctagggt cccaaacgat tagtttggat 240
tcaaaacatc cctgatcttc ctaaattggt cactcagggt tcgatccttc aaaatcagct 300
agtccacgac catcctactt ggcagccctt acatctcttt ctccccctc tcgttcacac 360
cttggttaatg tccatcagca tagagtcttg ttagtgtcca cgctgccagc caaaggattt 420
accatggggt ttgcacctg agtgaaccac ataacaagt gagggacatg aaattgcaaa 480
ttaatagctc agggatctcg ataacatgct tggacaagt ttagcnaactg ctgatgcac 540
ttagtcttat taaaagnntn nnnnacagtg cnacgncccc tatttttacac g 591

```

<210> SEQ ID NO: 187

<211> Comprimento: 2000

<212> Tipo: DNA

<213> Organismo: Zea mays

<220> Características:

<221> Nome: Origem

<222> Localização (1)...(2000)

<223> Outras informações: bacm.pk108.h15-2.fis47 sequência única

<400> SEQ ID NO. 187

```

tgagaggctg atatgtcttc tttttttct tattttttct aaaatcctga ttcctttcaa 60
tgtcagtttt gaattccaaa ttcaaattta atttgattct caagcttcaa ttttaatgca 120
acatccaaag acccaacatg aatgcataa ttccatttta tttatcttat tttatctatt 180
aaacaaagag tttcttaata tgaacttat atacacaaaa gacactattc taagaaaaca 240
attcctcata tatcatctta aatcttttgc tgaatattct ttaaacatta attctaaata 300
gttttatttg tacaagaatt tgtgattatt tcctacgaag agatggttcc taggcaactat 360
aatgaatagt tttgctaatt aaacaattga aaatgtttct atgttctctg ttttaacctt 420
agtttagagt ttttaacttc aagtttgaac taccaaagt tgaacctctt tttttatatt 480
ctttataact ttaaaataca tttaaactca aatcttttgg agaaccttta taaatcccaa 540
aatagggttt tgaggtgta taaacgagga gtctgtgcat gggacggcaa gtcacggtaa 600
gaccagtcca accgaatgat ggggcagctc ggtgtcgtct caattaatgc atctcttaag 660
ttgctgacca ccaaccatgt tggaacctga tcagttggta cccattgtct gtttaccatg 720
ggtagcttct aattggttcg cgtgatcttg aattttaaaa aattgaaata acatttttat 780
aaaattagaa tctagcttgt ggtaacattg ctactccat tgttgaatat atcaaatgg 840
agaaagcacg acgttaaaat gcttgcaaca tttatgtagg agtggtattt tatgtttttg 900
cgaggagtat aatcgtagtg tcaactgttg caatctttgg cgacttttag ctaaaggaga 960
ggagagacaa tttcttgcta atgataggaa ttatagattg catatattga aagtgataga 1020
gctagagtgc ccgatcttc ggtgagtgga gataattcog atttggtgga agtagaccct 1080
cagatccga ctacgacgag cgaaccgaa gcgccaatgc aatcgtgaa ccaactccca 1140
atggttaccg accttgctta tgcgagatcg gctgatcac gaagatcgtt tcctgtgctc 1200
aatcgaagaa cgaacaagaa aaagatgcca gcaatcttaa tcaactcga ggtggagttc 1260
tgaatcacag aggacaacac gtatttgtgt gtgttctggg gtagctaaag ctagatgtaa 1320
aacaaaactc aagttctaaa tgaacagga ctctgactaa atagaggaga ggcgtgaaca 1380
gtaggtcgac gctacagtac cacgtttact gttcacgact tacctaggcg cccatcccg 1440
gccttctat tggaccatct tctaactctc tgggccttcg ttctttaaca acatgatgta 1500
gttcaattct cttgcacgag cctgagtcac tggcccaagt ggaaggggtg cgctggact 1560
agggaaagtg gtgcctgggc tggagaaggt gctgctggtg tagatgtact catgttgggt 1620
ttgatgtct cctcatctc ccttcttga actgaagtgc tctcgcacgg caactcctca 1680
tcttctcccg caacggttt caaatctgca acattaaat tagtggaaac accaaactc 1740
gcaggtaggt caaggtata agcattagca ttaatcttgg ttagtatctt aaaaggacca 1800
gcaacacgag gcatcaattt agaacggcgc aaagtaggaa aacgatcctt tctcaaatgc 1860
aaccacaaca tatcaccagg ttcaaaagta actaggttct tgggctttac taccaataat 1920
ctgatattta gcattagcag caacaatgtt ttgttgagtt tgttctggag gttatcatt 1980
gttcacatgt gcacatcatc 2000

```

<210> SEQ ID NO: 188  
 <211> Comprimento: 1541  
 <212> Tipo: DNA  
 <213> Organismo: Zea mays

<220> Características:  
 <221> Nome: Origem  
 <222> Localização (1)...(1541)  
 <223> Outras Informações: bacm.pk010.m07-fis31 sequência única

<400> SEQ ID NO. 188  
 atcgatacc ttaattggga gataactggt atattaatag tgaaaaatcc atcctattat 60  
 aatttttgggt atatctttta catggccatc agtaatcaca gttaagagtt ggacaaaggc 120  
 actatgggagc gtggcccttc taattgtggt gctgaagttc ttattaaggt ttgccaattt 180  
 tttggtaggt gttccaaagc acagaaggct cttgagctaa gccactttt aggagtcaac 240  
 atttcttcaa cacaggtact tctgaagtag tcatgagtca ttgcttatct gaagtaagga 300  
 gagaaagtta ttgcttttct tttctgatgc gtgaggttgt gcatgtacat tctatagagt 360  
 gtgcctttca agtactacat caataatttc ttattttctc atagacctgt agcacctgaa 420  
 tgaaacctat ttttgaactc tgcatttttt gaagatgaca ttatgtatca atttagttct 480  
 gtattgggtt ttgccaaaaa atttgtaaga tattgagata gataatata gtgatattt 540  
 tttatgtagt tgttttgcaa taacattttg atgaatattt ttgtatatta ttgtaacaca 600  
 taggcaccta actagtgtat atataggatg gaaacacaaa caccatcgac gtcgatgatc 660  
 tcttggatag cattatctag tgtaggtgta catgggtatt atgattatca tagtgtggtt 720  
 atggttgtca tgacatcttg ttgctaccct gacaggtaac ctaaaaatcc aaaacacaat 780  
 gcaagtgcag gaatgtatga ccaatgcaaa agctaaagta aagtgccttc acattcgttg 840  
 ccaatgaaaag tttgaaaatg cttgaccaga ttttatcaac taataaatac tgtatataat 900  
 aggtctaaat ttgtgataac aaaacaagaa aaatagaaaa gtgggggaaac aaaaaatagt 960  
 agaaagagc gtggggagag gcgcgagaga ggtggttggg tgggggtgga ggaccgtcgg 1020  
 ccttccccgt ttcgtcttcc ggctctcggt ttatcccctc gcccttgcaac tctcccgcc 1080  
 caccctttc ctcttccgc cctccgcgc gcctggcttt ctcccttccc gagaggccc 1140  
 gctccatccc cgcccgcgt gctcccctg gcgatcgcg gcgaccgcg cccaacgaaa 1200  
 gcgtcgtggt cggccggtaa gagttccttg ctttctctc cgcgttgct ctgccgttac 1260  
 gcaaacctta tgtgatccct tctctgctga acgagatgag aagaaaagtt gctgactttc 1320  
 gttcacacgt gcacctgcat agaaactccg ggttttggct tccgttgga aattttgata 1380  
 aatgacttgc cacgcttctg cttacttcaa attcgttcaa attaattact tgctatgctt 1440  
 catcttactt caaatctgtt cgattcttgc tgggttctc tgattatata tttttttgta 1500  
 tgatcaagcg taaagatacc gtcgacctcc cgctttttga a 1541

<210> SEQ ID NO: 189  
 <211> Comprimento: 355  
 <212> Tipo: DNA  
 <213> Organismo: Sequência Artificial

<220> Características:  
 <223> Outras Informações: TEL0-266 consenso 355 bp repetida

<400> SEQ ID NO. 189  
 attttagtgt cgaaaccatg gtaaaatgaa atttgagtct cccaaccata tcaactttgc 60  
 tgacatagta gaatttttagt gtccaaacca tagtacacat tttgggtccc ggaggccggt 120  
 aaggctatth ttggcctccc gtvacacatg ttttcatcgt caaacaacga tttcatgcct 180  
 cccgctcgc acccatgatt tggccacwga gactgctaag gctrtttatg gctcccgtta 240  
 agctatagca camgttttca tggctgagcg actatthtta agtacgtgtt ccaccaccg 300  
 cgttttggct cccagagcac cttaaagttg ttcttggcct cccacgagct gtagg 355

<210> SEQ ID NO: 190  
 <211> Comprimento: 430  
 <212> Tipo: DNA  
 <213> Organismo: Sequência Artificial



<220> Características:

<223> Outras Informações: TR430 subtelomeric repetido

<400> SEQ ID NO. 190

```

gacatggttag aatttttagtg tccaaaccaa agtacacggt ttagtccccg gaggcccgta 60
aggctatttt ggccctcccg gacacatggt tttgtcgtcg aacagcgatt tcgtgcctcc 120
cgaccgccac ccatgatttg accatagaga ctggtaaggt tgttttggcc tcccgtgagc 180
tatagcacac gttttcatgg cggagcgacc atttttatgt ccgtgttcc tccacacgtgt 240
ttttgtcccc ggagcacctt aaagcggttc ttggcctccc gcgagctata tcacacattt 300
ccgtggccaa acaaccaatt ttatgtatcc gacctgccac cttcgtttta gaccggggag 360
gccgttatgg caattttttt gccaccctg agctatagca cacgatttta tggaaacctt 420
gaccctaat 430

```

<210> SEQ ID NO: 191

<211> Comprimento: 10368

<212> Tipo: DNA

<213> Organismo: Zea mays

<400> SEQ ID NO. 191

```

ggttccggtg gcaaaaactc gtgctttgta tgcacccgac acccgttttc ggaaagggtg 60
acgtgcgaca acgaaattgc gcgaaaccac cccaacatg agttttggac ctaaagtagt 120
ggattgggca tgtttgttgc tgatgtagct gaggtgccc atctttcggc gagtagagat 180
aattccgatt tggcgggaaga tgacccttgc gatccgacta cgacgagcaa gcccgaggcg 240
ccaatgcaat cgctgaacca actccctgtg gttaccgacc ttgctgatgc gagatcggcc 300
tgatcacgaa gatcgtttcc tgtgcgcaat cgaagaacga acaagaacaa gatgcgagca 360
atctaacta ttactogagg gtggagtctt gaatacacga ggacagcgca gatttgcgcg 420
tgttcggaa tagtaactag taacgtaaaa caaaactccc caaaaataaa ggaggcgag 480
ctcctgtata aatagagagg gggcgcagcc cctaggggcg gccaaccta ggtcgtccat 540
tatgggccc aattgggctg gtcgtctatc cttccgggcc ttcgttcttt aacaacatga 600
tgtagttcaa ttctcttgca cgggcccag tcaactggccc aggtggaagg ggtggcgcc 660
gggctggaga aggtgctgct ggtgtagatg tgctcgtgtt ggtgttgatg tcctcatcat 720
cctcccctc ttgaaactgaa gtcgtcctcg acggcaactc ctcatcttct cccgcatacg 780
gtttcaaatc tgcaacatta aaactagtg aaacaccaa ctccgcaggc aggtcaagga 840
tataagcatt atcattaatc ttggttagta ccttaaagg accagcagca cgaggcatca 900
atttagaacg gcgcaaagta ggaaaacgat cttttctcaa atgcaacaa accatatcac 960
caggttcaaa agtaactagt ttccggcctt tactaccagt aatctgatat ttagcattag 1020
cagcagcaat gttttgttga gtttgttcat ggagattaat catttgttca acatgtgcaa 1080
gagcatctat gtgtggggcg tccgtagcat caagcgaaaa caaagcaata ggccgacctag 1140
gaatgtaacc ataaacaatt tgaaaagggc acatctttgt agaagaatgt gttgcatgat 1200
tataagcaaa ctcaacatga ggtaagcaat cctcccaacg tttcaaattt ttgtctaaaa 1260
cagccctaag catggtagac aaagtccgat taactacctc agtttgacca tcagtctgag 1320
ggtgacaagt ggtgctaaac agcaatttag ttccataattt attccacaga gatctccaaa 1380
aatgactcag aaacttgga tcacgatccg agactattgt attgggaata ccgtgcaaac 1440
gaataatctc tctaaaaaac aattcagcaa cattgctagc atcatcagtc ttatgacaag 1500
gtatgaaatg agccattttg gagaatcgat caacaaccac aaaaatgta tccctcccct 1560
tcttagttct aggcaatccc aaacaaaaat ccatcgaaat atcaatccaa gggaaagtag 1620
gaacaggcaa aggcataatc aaacctaggt tgttcaaccg tgacttagct ttctgacaag 1680
tagtgcagcg tgcaacaagg cgctcaacat cagcgcgcat ccgaggccaa aagaagtggg 1740
cagccaacac ctcatgtgtc ttgtagacgc caaagtgcc catgagaccg cctccatgtg 1800
cttctgttaa caacaaaaga cgaaccgagc tagctggaac acacagcttg ttagcgcgaa 1860
acaggaaccc atcctgtatg tgaaatttgc cccatggttt cccattaata caatggccga 1920
aagcatcttt aaaatcagca tcgtaacat attgatcttt tacagtgtgc aaaccaaaga 1980
ttttaaatac taactgtgac agcatggat agcgacgaga caaagcatca gcaataacat 2040
tgtccttccc gttcttgtgt ttaataatgt aaggaaaaga ctcaatgaat tctaccatt 2100
tagcatgacg acggttcaga tttgttggg tagcaatag ttttaaagcc tcatgatcag 2160
aatgaattat gaactcacga tgccaaagat agtgcgtcca tgtatgtaaa gtgcgacta 2220
acgcgtaaag ctcttatca taagtagaat atttcagact agcaccgctt aatttttcac 2280
taaaataagc aactggtttt cttcttggta ataaaacagc acctagccca ataccgctag 2340
catcgcattc aagctcaaat actttattaa aatcaggcaa ttgcaatagg ggagcttggg 2400
ttaacttatc tttcaaagtg ctgaacgctt cctcctgcga atcactccaa gcaaatggca 2460

```

catctttctt	tgtaagctca	tgtagaggcg	ctgcaatgga	gctaaaatca	cgaacaaatc	2520
tgcggtagaa	accggcaagt	ccaagaaagc	tccgaatttg	tgtgaccgtc	gtcgggtgtag	2580
gccactcccg	aatggcagca	atcttgctgc	tatccacctc	aatgccctgt	ggagtaacaa	2640
cataaccaag	aaacgagaca	cgctgtgtgc	aaaagatgca	tttttccatg	ttagcgaata	2700
actggggcggc	acgcaatgca	tcaaaaacag	cactttaaag	ttccaaatgc	tctttcttag	2760
at ttgctgta	aataaggata	tcatcgaaat	aaacaaccac	aaacaatcct	atgaagggcc	2820
tcagaacttc	attcatcact	cgcataaaag	tgctgggagc	attagtcact	ccaaacggca	2880
taaccaacca	ttcatataaa	ccaaatttcg	ttttgaaggc	tgttttccat	tcatcaccta	2940
gtttcattct	aatctggtgg	taaccactac	gcaaatcaat	cttagtgaaa	ataatggcac	3000
cactaagctc	atctagcata	tcatcaaggc	gtggatagga	atagcgataa	cgaatagtga	3060
tattattaat	agcagcagag	tctacacaca	tacgccatga	cccatecttc	ttgggaacaa	3120
gtaacacagg	aacagagcaa	gggctaagag	actcacgaat	gtatcctttg	tcaagcaacg	3180
cctttacctg	gcgctgaatc	tccttcgtct	catccggatt	tgtacgggat	gggctgcggg	3240
ttggaagctg	tgcaccggga	atgaggtcga	tctgggtgctc	aatgccacga	agcgggtggga	3300
gacccggtgg	taagtctttg	ggaaagacat	cagcgtactc	ctgcaaaagg	ttagcaacca	3360
tagggggaat	agccaaagat	ggtgcatcat	caagtgaat	gaggacacta	gagcatacaa	3420
gtgcatagca	tggcaaatga	gcaccgtgta	gatcatcaaa	atcagcacgt	gtagcaagta	3480
aaacaggagc	cttcaactta	at ttcagaag	gaacagagtg	cgatggatca	agttgtttag	3540
cagttatagc	agctcgggca	agatcatctt	taacaatttg	ttcaggtgctc	attggatgta	3600
taattat ttt	ctgaccctta	aaaaatgaa	aaaaatgatt	cgaacgacca	tgatgcaagc	3660
tatcagtatc	atattgccaa	ggtcagccaa	gtaacaaaga	gcatgcttcc	atggggaataa	3720
catcacaatc	aaaaaaatca	gaataagcac	ccatggagaa	aggaactcgc	acggaacgcg	3780
tgatttttat	tttaccacca	tcattaagcc	attgaatgtg	atatgggttc	ggatgcttac	3840
gagtgggtaa	cgacaacttt	tcgaccagca	tggtactcgc	caaattgttg	caactgccgc	3900
tgtcgataat	gatgcgaatt	gaccgctcct	gcacaacacc	ctttgtatgg	aacagagtgt	3960
gtcgtctgatt	cttctcgggc	aaagcgacct	gtgtactgag	aacacgctgc	acaacaagac	4020
tttcatactc	atcagcgtcg	cccgggttga	catgtacctc	cgcttagct	gcatggtcag	4080
tggcaagtga	tgcatgttca	at t tcttcag	aatcactagc	ggaagagtac	tcaccatctg	4140
cacgaatgag	taaagtgcgc	ttgtttggac	agtcccgaat	cacatggcca	aatcctctac	4200
agcgatgaca	ctgaatatcc	cgtgtacgac	ctgtgggtgg	ggcagtgctg	gctggtttgg	4260
tcgtcttctc	gcgtggagta	gtggtggacg	tggatggcgc	agggggaact	gggtgctgagg	4320
tggatgtcga	gcttcgtcct	gcaaaagggt	tagaatattg	cttcgagcgg	cgctcctgca	4380
cttcacgttc	agctttgcaa	gcatattcaa	ataaagtagt	catatcataa	tattccttat	4440
aatcaagtat	atcttggatt	tcgctgttta	aaccaccacg	aaaacgtgcc	atagcagcgt	4500
cgttgtcctc	caatatccca	caacgaatca	tacctttttg	taactcctgg	aaatagtcct	4560
caacagattg	agaaccttgc	tgaaaacgct	gcattttatt	aagcaaatca	cgagcataat	4620
acgaaggcac	aaatcgattg	cgcacgcagc	ctttcaactg	atcccaagtg	gttggaaatgg	4680
tagtgggatg	ttttatttta	aactcacgcc	accaaat taa	agcaaaatca	gtaaat tca	4740
taatagcagc	tttaacctga	gaatgtgcag	caatatcatg	gcatgaaaat	ttttgttcaa	4800
cctctaactc	ccaatcaaga	tatgcagcag	gatcatattt	gccattaaaa	gggtggaattt	4860
taaatttaat	cttagaaaaat	aagtcattag	ggggattacg	aaccacacgg	cgtgcacgac	4920
cacggcgatc	tccatcgtcc	cactcagtg	caccgccgta	gtcctgctcc	atctttgtgg	4980
tcaacgcate	aaggcgtgct	aggatgggtg	cgagtgtggt	gcgagtcgcc	gtttgagcaa	5040
ggtcaagctg	gttgaaaacgc	tcggtcgtcg	aagtgaccgt	tgaatcaagc	cgttcatgca	5100
tcgtccta	gtcatcagca	agtccatcaa	cttgtccctt	tacttccagc	aactgggcat	5160
ccaccgtgtc	atgtgctcct	gccatagtta	gcgcaaacac	caaaacacca	aaaaaaacga	5220
caaaaacagg	ggtgtactgc	tcacaaggcg	ctcacactag	tgctgttatc	aaattcttat	5280
ccgttcttac	caagccacag	tggtgaaactg	caaccaacag	gtggaaccgg	tgaagatttg	5340
gatgagcgat	tgcttgagga	aacagaaacc	tactcgttgt	agaaatatgt	ggagttctgg	5400
gtaggctgca	ctcaagtcaa	ggattagcac	gaccaaacaa	taatgcaaag	ttgaattata	5460
gtgcaaaaaca	cgaaactata	ttgctggcca	caagtgcaaa	ggacggatgg	aactagcaga	5520
atggcagtac	cgtaaatatt	gtactagcga	ggccactagt	aggaatcaca	agtgattttg	5580
tttttctttt	ttgtatgatt	tttttgggat	ttttctcagc	acaagaagca	acaaaatagg	5640
agctacacga	agtttcacct	aaaacagagt	tcaaatgtgg	tctacagaaa	atcagaaaat	5700
tctctaaaaa	gcgtgcgaga	actttgaag	at t t t t t t t	tattttctg	aatttttttt	5760
gacaattttg	tcgaacccea	acagaccgaa	ggtcagtttg	gccggcctca	gaatgggtgc	5820
aactagctcc	tgtaaaaatt	tcagattttt	cggacaccgg	agcgaaaagt	tatgccgggt	5880
ttaaggaagg	tacctcaaat	tatgttttca	aacgaccgga	atgaaacaac	cgtatccttt	5940
ctccttcggt	gttttttgggt	tctgtttttt	tttttacgta	accgaaggag	aaaaacaagg	6000
aaacgatggt	gactcgggtt	gttttttttt	ctgttttttt	ttctgttttt	tttcgtaacc	6060
gaaggagaaa	atcaaggaaa	acagccggtg	actcgggttg	ttttttctgt	tttttttga	6120

cgtaaccgaa	ggagaaaaac	aaggaaacaa	tgttgactcg	gtttgtggtg	tgatcaaacg	6180
agagatggtg	gcggcgctag	ggtttgaatg	gtggaagaac	acaatgcaac	cagcaacaaa	6240
tgacgcgaaa	gcacacaaat	tcaacaatgc	agattattga	aagaaagtgc	gaggctcaaa	6300
aggggtgctgg	gataagatct	aacctgaatt	tttatgtggt	tttgtggact	gtaggaaaaa	6360
aaaacgctcg	ataaactcac	cgatcaacct	agaaatctga	taccaattga	tgaagctgag	6420
gtgcccgatc	tttcggcgag	tagagataat	tccgatttgg	cggaagatga	cccttgcgat	6480
ccgactacga	cgagcaagcc	cgaggcgcca	atgcaatcgc	tgaaccaact	ccctgtggtt	6540
accgaccttg	ctgatgcgag	atcggcctga	tacagaaat	cgtttctctg	gcgcaatcga	6600
agaacgaaca	agaataagat	gcgagcaatc	taatctatta	ctcgagggtg	gagttctgaa	6660
tacacgaaga	cagcgagat	tagcgcgtg	tcgagagtga	ctaaggctaa	cgtaaaacaa	6720
aactcaggaa	ataaaggagg	cgcagctcct	gaataaatag	agaggggcg	cagcccctag	6780
gggcgcccaa	ccctaggctg	tccattatgg	gccgcaattg	ggctggtcgt	ctattcttcc	6840
gggccttcgt	tctttaacaa	catgatgtag	ttcaattctc	ttgcacgggc	ccgagtcact	6900
ggcccaggtg	gaaggggtgg	cgctgggct	ggagaaggtg	ctgctggtgt	agatgtgctc	6960
gtgttggtgt	tgatgtccgc	atcaacatcg	caatcaacat	aatcagaata	agaaccagc	7020
gaaaagggga	cacgtaccga	acgtgttacc	tttattttac	caccatcatt	aagccattga	7080
atgtgatagc	gatgtggatg	tgtgcgagtg	ggcaaggata	atctctctac	caacgctgta	7140
cttgccaaat	tgttgcagct	gccactatcg	atgatgatgc	gaatcgaccg	ttcgtgcacg	7200
acgccccttg	tatggaatag	agtgtgtcgc	tgattttttt	cggcctgggc	aacctgtgtg	7260
ctgagaacac	gctgcacaac	aagactctca	tacctctcag	cgctgatggg	atcaacgtg	7320
acttctcat	tttctgcatg	gttagtggca	atcatagcat	gactagtctc	ctcagaatca	7380
ctggctgaag	agtactcacc	attgtcacgt	ataagcaagg	tacgcttgtt	tgggcagctc	7440
cgaatcatgt	gcccacaccc	tctgcaacga	tggcactgaa	tatcccgtgt	acgtcctgtg	7500
gaagaagcgg	cgcccttggc	agggggcgcc	actggcttgg	ccggccctgt	gcgcgatgta	7560
gtgctaggcg	taggaggtgc	agggctggaa	ggagttgagc	tgtgtgttgg	accccgccct	7620
gcaaaagagt	tagtatatgt	ctttgatcgt	cgccccgca	cttcacgttc	agctttgcaa	7680
gcatattcaa	acaatgtggt	tatatcaaaa	taatccttat	aatcaagtat	atcctgaatt	7740
tcctgtttca	aaccaccag	aaaacgcgcc	atagcagcgt	catctgactc	aaccaaacca	7800
caacgaagca	tacccttttg	caactcctgg	taatactcct	caacagattg	tgaacctgtg	7860
tgaaaacgct	gcattttgtt	aagcaaatca	cgagcataat	aggaaggaac	aaatctgtgg	7920
cgcattggcag	tttttaattg	ggtccaagta	atgacactgt	taatgggaag	tttttgttta	7980
tactcacgcc	accaaattaa	agcaaaatca	gtaaatcac	taatggcagc	cttcacttgg	8040
ctattagcag	gaatatcatg	gcatgaaaat	ttctgttcta	cctctaattc	ccaatcaaga	8100
tatgcagcag	gatcatattt	accattaaaa	gatggaattt	taaatttaat	cttagaaaat	8160
aagtcattag	ggggatgacg	aaccacacga	cgtgcacgac	cacggcgatc	tccatcgctc	8220
tgctcagtgt	caccgcccga	ttcctgctcc	atctttgtgg	tcaatgcatc	aaggcgtgcc	8280
aggatggtgt	cgagtggtgt	cgagtcgcc	gtttgagcaa	ggtcaagttg	gttgaaacgc	8340
tcggctgctg	aagtgatcgt	tgaatcaagc	cgttcatgca	tcgtcctaata	gtcagcagca	8400
agtccatcaa	cttggccctc	tacttctctg	aactgggcat	ccaccatgtc	gtgtgctcct	8460
gccatagtta	gcgcaaacac	caaaaggaga	aaaaccaacg	acaaaaacag	gggtgtactg	8520
ctcacaaggc	gctcacacta	gtgctgttat	caagttctta	tccgttctta	ccaagccaca	8580
gtggtgaaact	gcaaccaaca	ggtggaaccg	gtgaaagatt	ggatgagcga	ttgcttggag	8640
aaacagaaaac	ctgctcgtcg	tagaaatatg	tggagtgtg	ggtaggctgc	actcaagtca	8700
aggattagca	cgatcaaaac	ataatgcaaa	gttgaattat	agtgcaaaaac	acgaaactat	8760
attgctggcc	acaggtgcaa	aggatggatg	gatgaaaata	gcagaatggc	agtaacgtaa	8820
aattgtact	agtgatgcca	aaaaggcact	agcacaaatc	acaggtgatt	ttgtttttct	8880
tttttgtatg	atttttttga	tatttttctc	agcaacaaga	gcaacaagat	aggagctaca	8940
cgaagtttca	cctaaaacag	atatcagatg	tggcttacag	aaaatcagga	agttctctga	9000
aaagcgtgcg	agaactttga	cggatttttt	tctttatttt	cctgaatttt	tttgacaatt	9060
ttgtcgaacc	ccaaacagac	cgtaggtgag	tttggccggg	ctcagaatgg	tgtcaactag	9120
ctcctgtaaa	aatttcagat	tttttggaca	cccgagcgaa	aagttatgcc	cggtttaagg	9180
aaggtagccct	caagttatgt	tttcaaacga	ccgggatgaa	acaaccgcat	cctttctcct	9240
tcgttgtttt	tttttgtttc	tgtttttttt	ttgacgtaac	cgaaggagaa	aaacaaggaa	9300
acgatgttgc	ctcggttttt	ttttttctgt	ttttttctgt	aaccgaagga	gaaaaacaag	9360
gaaacggccg	ttgactcggg	ttgttttttt	tgtttttttt	tacgtaaccg	aaggagaaaa	9420
acaaggaaac	aatgttgact	cggtttggca	cgtgatcaaa	ggggagatgg	tggcggcgct	9480
aggatatgaa	tgggtggaaga	acacaatgca	accagcaaca	aggaaacgcg	aaagcacaca	9540
aattcaacaa	tgcagattat	tgaagaaag	tgcgaggctc	aaaagggtgc	tgggataaga	9600
actaacctga	atttttatgt	ggttttgtgg	actgtaggaa	aaaaaacgct	cgataaaactc	9660
accgatcaac	ctggaaatct	gataccaatt	gatgtagctg	aggtgcccg	tctttcggcg	9720
agtagagata	attccgattt	ggcggaagat	gacccttgcg	atccgactac	gacgagcaag	9780

```

cccgaggtgc caatgcaatc gctgaaccaa ctccctgtgg ttaccgacct tgctgatgcg 9840
agatcggcct gatcacgaag atcgtttctt gtgcgcaatc gaagaacgaa caagaacaag 9900
atgcgagcaa tctaattctat tactcgaggg tggagttctg aatacacgag gacagcgcag 9960
atttgcgctg gttcgggaagt agctaaggct aacgtaaaac aaaactccca aaaataaagg 10020
aggcgcagct cctgtataaa tagagagggg gcgagcccc taggggcggc caaccctagg 10080
tcgtccatta tgggccgcaa ttgggctggt cgtctatcct tccgggcctt cgttctttaa 10140
caacatgatg tagttcaatt ctcttgacag ggccccgagtc actggcccag gtggaagggg 10200
tggcgccctgg gctgggagaag gtgctgctgg tgtagatgtg ctcggtgttg tgttgatgtc 10260
ctcatcacat gtttgggggtg gtttcgcgca atttcgttgt cgcacgtcac acattccgaa 10320
aacggttgtc ggggtgcata caaagcacga gtttttgaca ccggaacc 10368

```

```

<210> SEQ ID NO: 192
<211> Comprimento: 32
<212> Tipo: DNA
<213> Organismo: Sequência Artificial

```

```

<220> Características:
<223> Outras Informações: Telo-31 overgo primer1 Biocode 75319

```

```

<400> SEQ ID NO. 192
agggtttagg gtttagggtt tagggtttag gg 32

```

```

<210> SEQ ID NO: 193
<211> Comprimento: 32
<212> Tipo: DNA
<213> Organismo: Sequência Artificial

```

```

<220> Características:
<223> Outras Informações: Telo-31overgo primer2 Biocode 39612

```

```

<400> SEQ ID NO. 193
ccctaaacc taaaccctaa accctaaacc c 31

```

App. Ref.: 2083-PCT

**REIVINDICAÇÕES**

1. Minicromossomo artificial de planta caracterizado por compreender um centrômero funcional contendo: (a) pelo menos dois arranjos de repetição em tandem de CentC em uma  
5 orientação invertida em que o primeiro arranjo compreende pelo menos cinquenta cópias de CentC e o segundo arranjo compreende pelo menos cinquenta cópias de CentC; e (b) pelo menos uma cópia de um elemento retrotransponível, em que o elemento retrotransponível está localizado entre o primeiro  
10 e o segundo arranjo.
2. Minicromossomo artificial de planta de acordo com a reivindicação 1, caracterizado pelo fato de que o elemento retrotransponível é selecionado do grupo que consiste em CentA, CRM1 e CRM2.
- 15 3. Minicromossomo artificial de planta de acordo com a reivindicação 1, caracterizado pelo fato de que o referido minicromossomo compreende, adicionalmente, pelo menos um telômero funcional.
4. Minicromossomo artificial de planta de acordo com  
20 qualquer uma das reivindicações de 1 a 3, caracterizado pelo fato de que o centrômero funcional se liga especificamente à proteína centromérica C (CENPC).

5. Planta de milho caracterizada por compreender o minicromossomo artificial de acordo com qualquer uma das reivindicações de 1 a 3.

6. Planta de milho caracterizada por compreender o  
5 minicromossomo artificial de acordo com a reivindicação 4.

7. Minicromossomo artificial de planta caracterizado por compreender um centrômero funcional em que o centrômero se liga especificamente à proteína centromérica C (CENPC).

8. Planta de milho caracterizada por compreender o  
10 minicromossomo artificial de acordo com a reivindicação 7.

9. Polinucleotídeo isolado caracterizado por compreender:  
(a) pelo menos dois arranjos de repetições em tandem de CentC em uma orientação invertida em que o primeiro arranjo compreende pelo menos dez cópias de CentC e o segundo  
15 arranjo compreende pelo menos dez cópias de CentC; e (b) pelo menos uma cópia de um elemento retrotransponível, em que o elemento retrotransponível está localizado entre o primeiro e o segundo arranjo.

10. Polinucleotídeo isolado de acordo com a reivindicação  
20 9, caracterizado pelo fato de que o elemento retrotransponível é selecionado do grupo que consiste em CentA, CRM1 e CRM2.

11. Polinucleotídeo isolado caracterizado por compreender:  
(a) pelo menos um arranjo de repetições em tandem de CentC, o arranjo compreendendo pelo menos 10 cópias de CentC; e  
(b) pelo menos uma cópia de um elemento retrotransponível selecionada do grupo que consiste em CentA, CRM1 e CRM2.
- 5
12. Polinucleotídeo isolado caracterizado por compreender:  
(a) pelo menos um arranjo de repetições em tandem de CentC, o arranjo compreendendo pelo menos 10 cópias de CentC; e  
(b) pelo menos uma cópia de cada de CentA, CRM1 e CRM2.
- 10
13. Construção recombinante caracterizada por compreender o polinucleotídeo isolado de acordo com qualquer uma das reivindicações de 9 a 12.
14. Construção recombinante de acordo com a reivindicação 13, caracterizada por compreender adicionalmente um
- 15
- fragmento de DNA compreendendo um arranjo de pelo menos 30 cópias de repetições teloméricas.
15. Planta transgênica de milho caracterizada por compreender a construção recombinante de acordo com a reivindicação 13.
- 20
16. Planta transgênica de milho caracterizada por compreender a construção recombinante de acordo com a reivindicação 14.

17. Método para produzir uma planta transgênica de milho compreendendo um minicromossomo artificial de planta tendo um centrômero funcional caracterizado por compreender:

5 (a) contatar pelo menos uma célula vegetal de milho com uma mistura compreendendo a construção recombinante de acordo com a reivindicação 14;

(b) identificar pelo menos uma célula vegetal de milho da etapa (a) compreendendo um minicromossomo artificial de planta tendo um centrômero funcional; e

10 (c) regenerar uma planta de milho fértil a partir da célula vegetal de milho da etapa (b) em que a planta de milho citada compreende um minicromossomo artificial de planta tendo um centrômero funcional.

18. Método de acordo com a reivindicação 17, caracterizado  
15 pelo fato de que a mistura compreende adicionalmente um polipeptídeo que estimula crescimento de célula.

19. Método de acordo com a reivindicação 18, caracterizado pelo fato de que o polipeptídeo é selecionado do grupo que consiste em um *wuschel*, um *baby boom*, um RepA ou um Lec1.



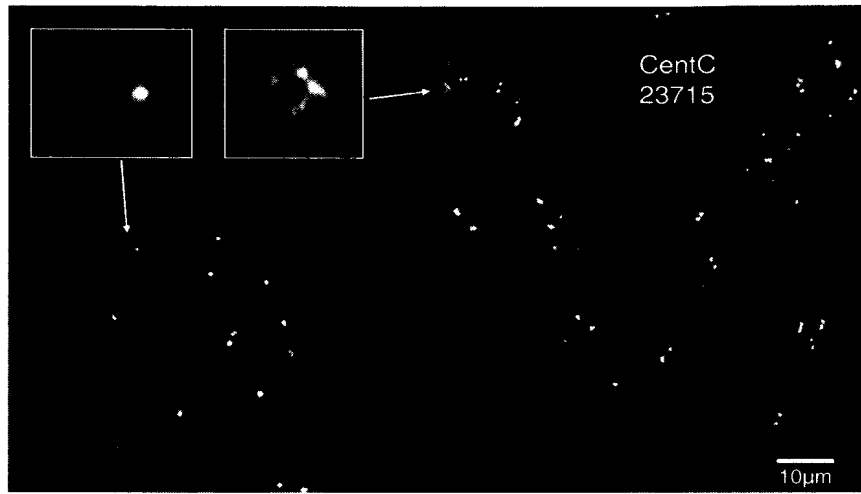


FIGURA 1

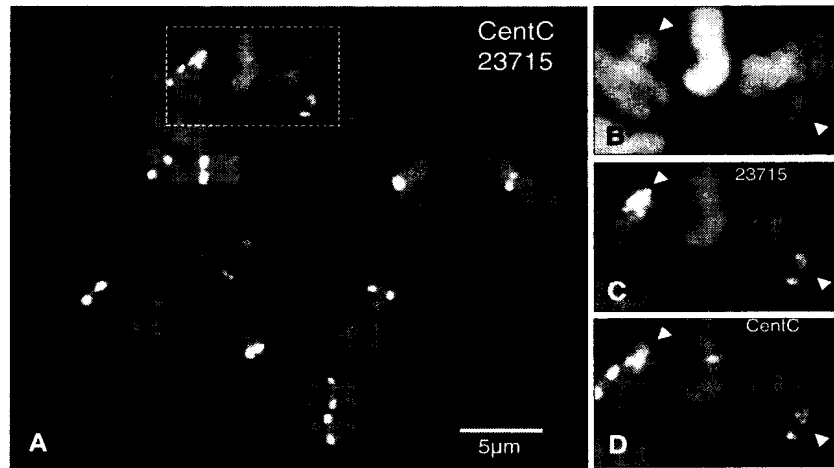


FIGURA 2

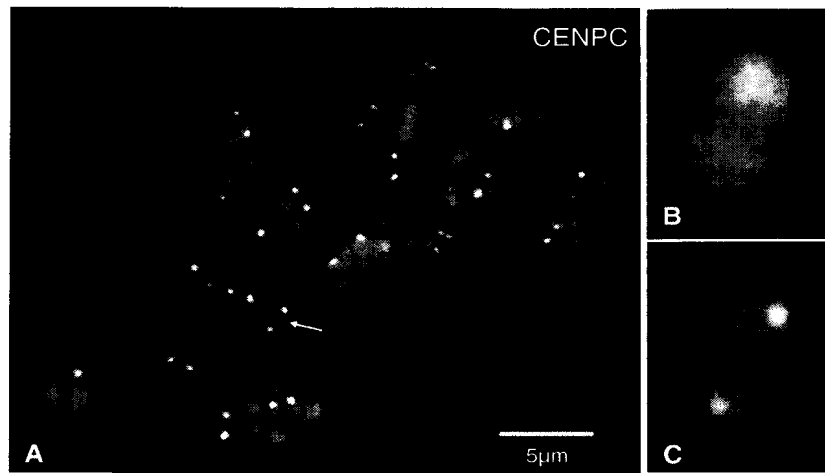


FIGURA 3

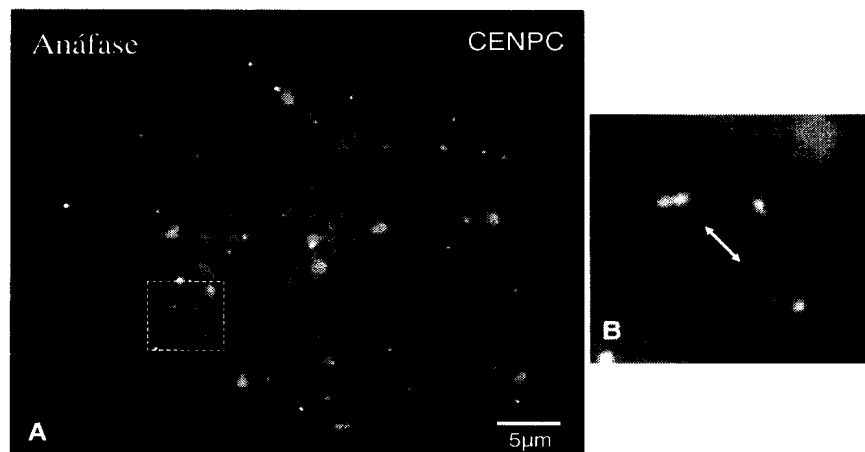


FIGURA 4

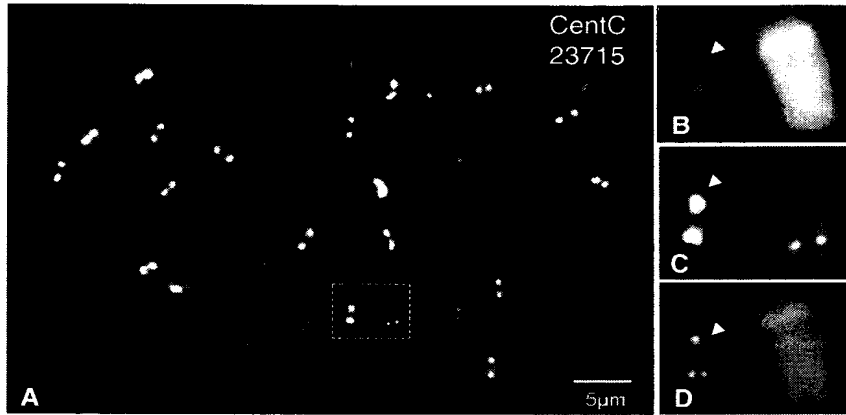


FIGURA 5

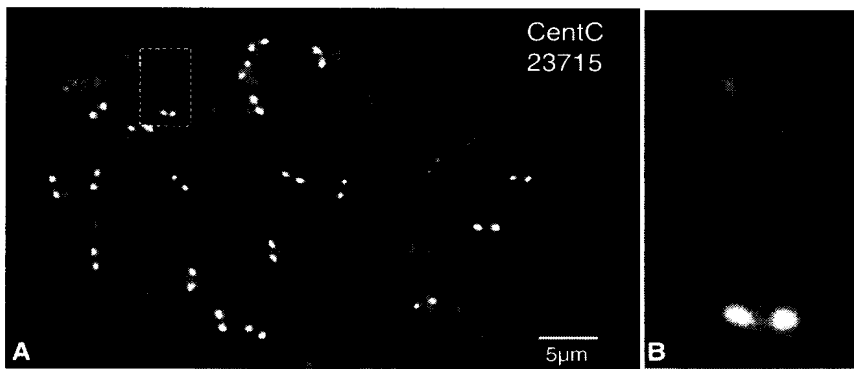


FIGURA 6

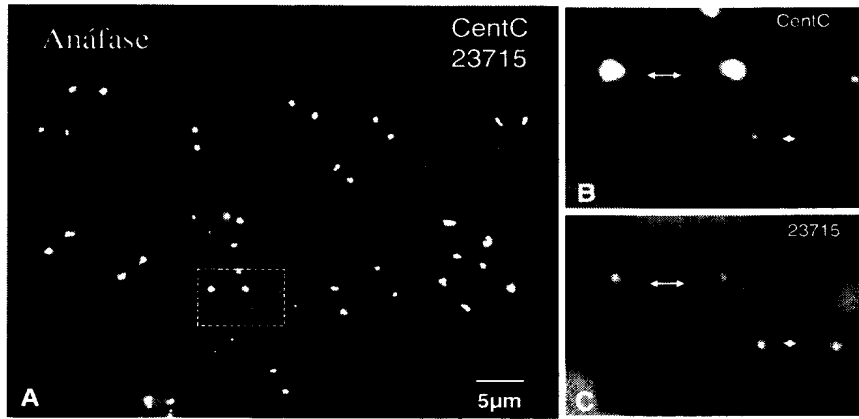


FIGURA 7



FIGURA 8

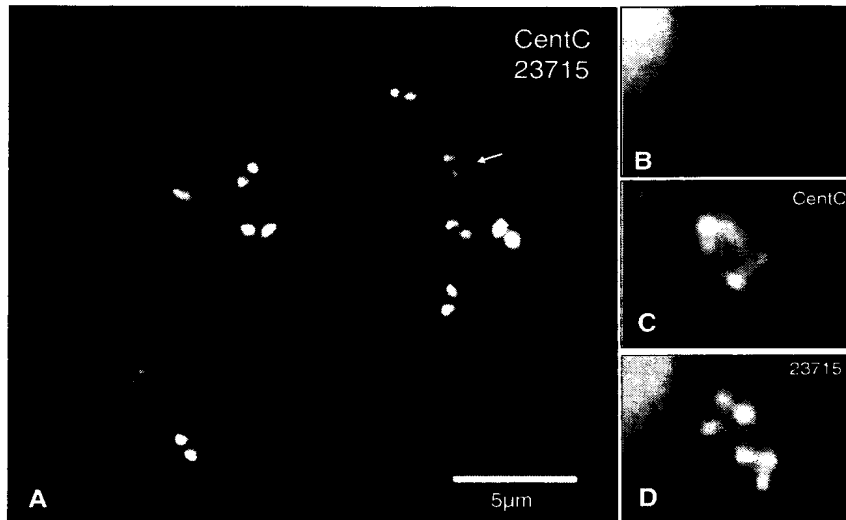


FIGURA 9

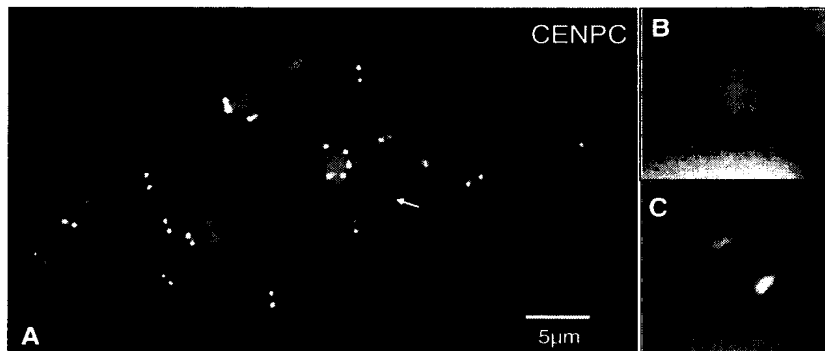


FIGURA 10

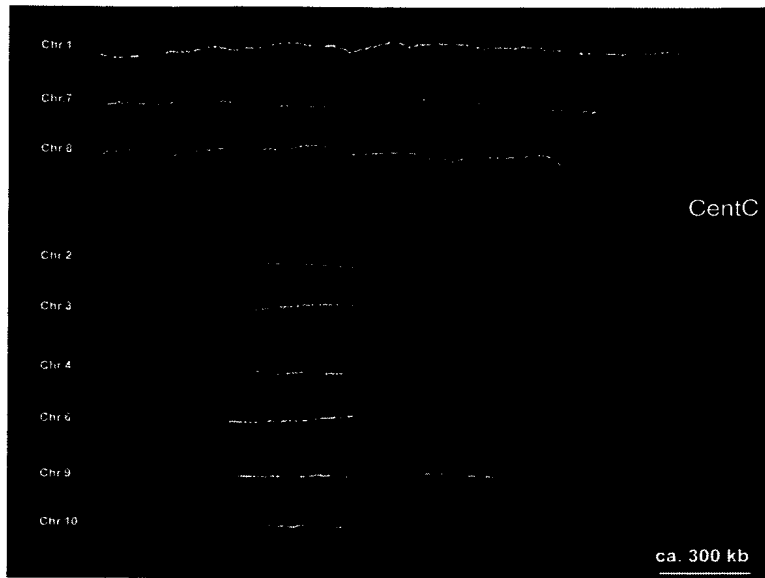


FIGURA 11

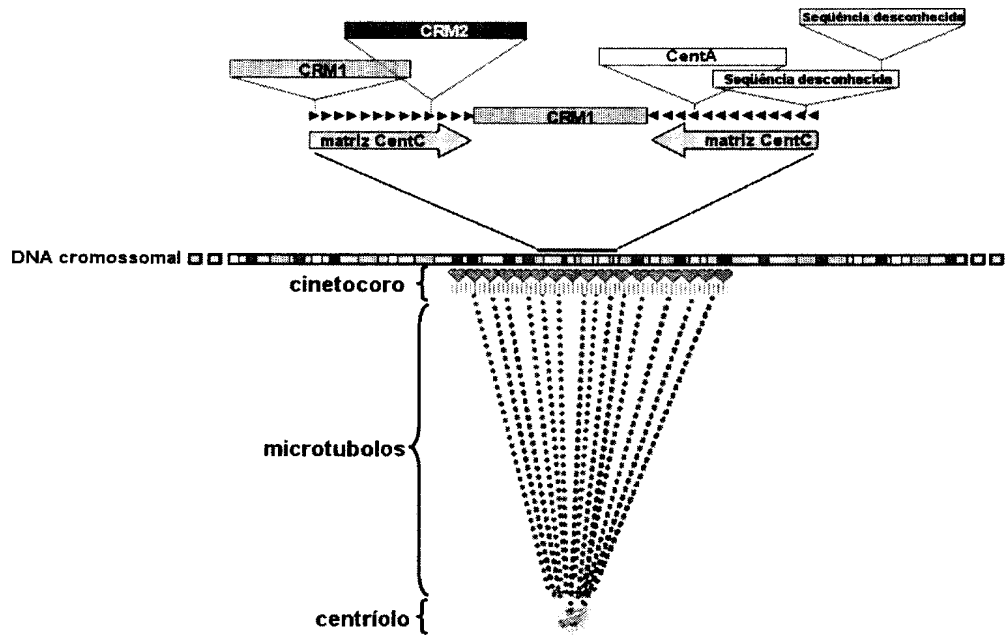


FIGURA 12

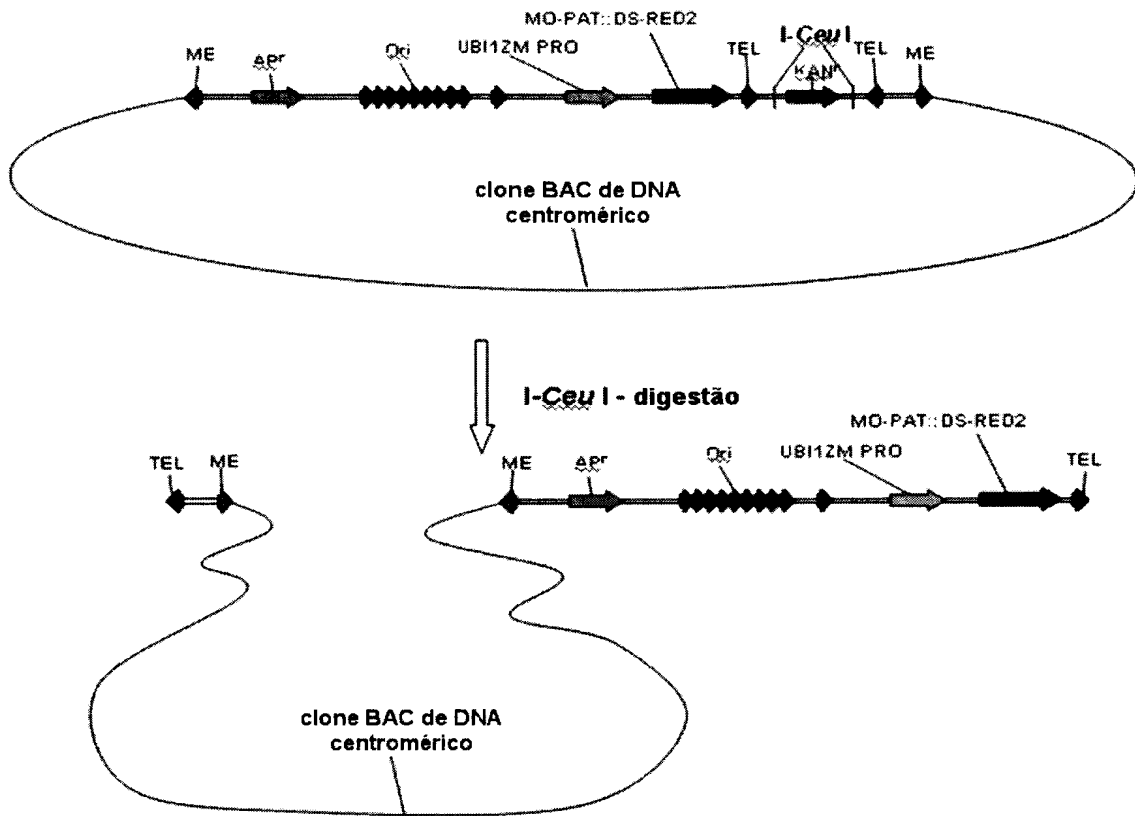


FIGURA 13

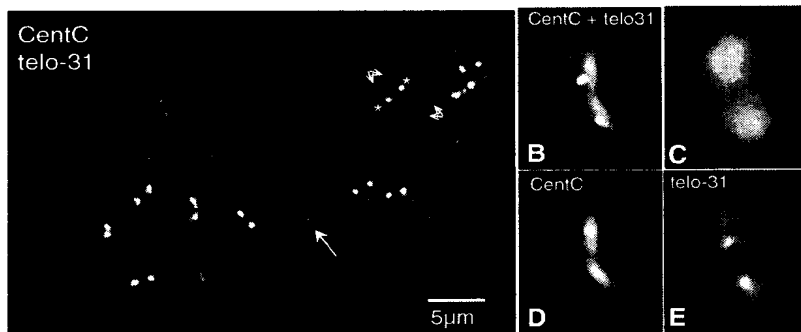


FIGURA 14

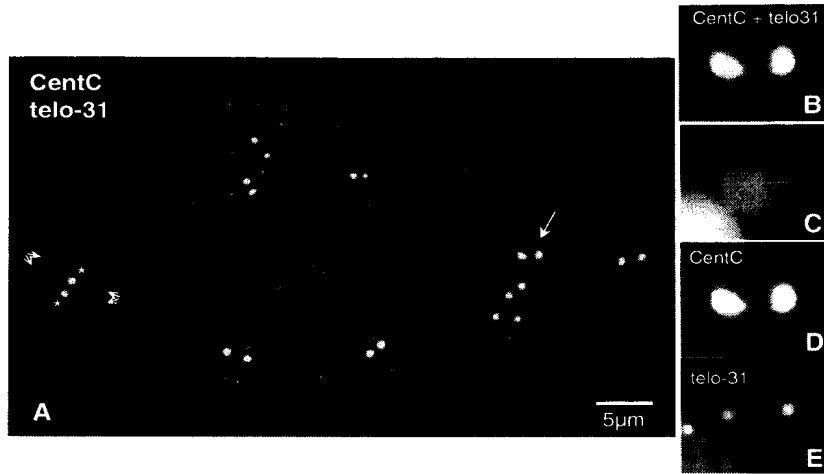


FIGURA 15



**RESUMO**

**MINICROMOSSOMO ARTIFICIAL DE PLANTA, PLANTA DE MILHO,  
POLINUCLEOTÍDEO ISOLADO, CONSTRUÇÃO RECOMBINANTE, PLANTA  
TRANSGÊNICA DE MILHO E MÉTODO PARA PRODUZIR UMA PLANTA  
5 TRANSGÊNICA DE MILHO COMPREENDENDO UM MINICROMOSSOMO  
ARTIFICIAL DE PLANTA TENDO UM CENTRÔMERO FUNCIONAL**

São descritos minicromossomos artificiais de planta  
compreendendo um centrômero funcional o qual se liga  
10 especificamente à proteína centromérica C (CENPC) e métodos  
para produzir tais minicromossomos.