



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

**(11) PI 0922187-5 B1**



**(22) Data do Depósito:** 19/11/2009

**(45) Data de Concessão:** 04/05/2021

**(54) Título:** MÉTODOS PARA GERAR UM POLINUCLEOTÍDEO E PARA GERAR UMA CÉLULA HOSPEDEIRA

**(51) Int.Cl.:** C12N 15/06; C12N 15/10; C40B 40/06.

**(30) Prioridade Unionista:** 19/11/2008 US 61/116,109; 20/03/2009 US 61/162,230.

**(73) Titular(es):** AMYRIS, INC..

**(72) Inventor(es):** ZACH SERBER; RAYMOND LOWE; JEFFREY A. UBERSAX; SUNIL S. CHANDRAN; ERICK JEDEDIAH DEAN; DARREN M. PLATT; KENNETH TOSHIKI TAKEOKA.

**(86) Pedido PCT:** PCT US2009065048 de 19/11/2009

**(87) Publicação PCT:** WO 2010/059763 de 27/05/2010

**(85) Data do Início da Fase Nacional:** 19/05/2011

**(57) Resumo:** COMPOSIÇÃO, VETOR, KIT PARA A MONTAGEM DE UM POLINUCLEOTÍDEO, BIBLIOTECA DE MOLÉCULAS DE ÁCIDO NUCLEICO, MÉTODOS PARA GERAR UM POLINUCLEOTÍDEO, E UMA CÉLULA HOSPEDEIRA, CÉLULA HOSPEDEIRA, POLINUCLEOTÍDEO, E, MÉTODO PARA FABRICAR UM VETOR. A presente invenção fornece composições e métodos para a montagem rápida de um ou mais polinucleotídeos montados a partir de uma pluralidade de polinucleotídeos componentes. Os métodos da invenção utilizam vetores de ácido nucléico circular que compreendem um segmento de DNA D flanqueado por uma sequência ligadora recozível, pares de sequência de ligador recozíveis LA e LB ou sequência de ligador recozível / pares de segmento de ligação de iniciador LA e PC ou PA e LB. A Digestão de endonuclease de restrição de uma pluralidade de vetores contendo os segmentos de DNA ao montado gera uma pluralidade de fragmentos de DNA que compreendem os elementos PA-D-LB, LA-D-LB e LA-DPB ou D-LB, LA-D-LB e LA-D. As sequências de sequências de ligador recozível LA e LB fornecem terminais complementares aos fragmentos de DNA, que são utilizados na recombinação homóloga mediada por célula hospedeira ou junto com os segmentos de ligação de promotor PA e PB em uma reação de montagem (...).

## “MÉTODOS PARA GERAR UM POLINUCLEOTÍDEO E PARA GERAR UMA CÉLULA HOSPEDEIRA”

[001] Este pedido reivindica o benefício da prioridade do Pedido Provisório U. S. N° 61/116.109, depositado em 19 de novembro de 2008 e Pedido Provisório U. S. N° 61/162.230, depositado em 23 de março de 2009, cujos conteúdos são incorporados aqui por referência em suas totalidades.

### 1. CAMPO DA INVENÇÃO

[002] A presente invenção, em geral, diz respeito ao campo da tecnologia de DNA recombinante e, mais particularmente, a métodos melhorados para a montagem ordenada de uma pluralidade de segmentos de DNA em um polinucleotídeo montado.

### 2. FUNDAMENTOS DA INVENÇÃO

[003] A recombinação de polinucleotídeos pode ser realizada usando-se muitos métodos conhecidos na técnica. As técnicas tradicionais para a recombinação de ácidos nucleicos utilizaram enzimas de restrição e que ligam enzimas para a criação de novas moléculas de ácido nucleico. As moléculas recombinantes, tais como os vetores de clonagem e de expressão podem ser utilizados para integrar uma sequência de ácido nucleico de interesse no genoma de uma célula hospedeira e/ou conduz a expressão de um ou mais genes de interesse. A utilização de um vetor para conduzir a expressão de um gene de interesse na célula, por exemplo, uma célula de levedura, requer que o vetor contém os elementos genéticos requeridos que permitem a replicação e a expressão do gene de interesse. Estes elementos podem incluir, por exemplo, o gene ou genes de interesse, uma sequência promotora, uma sequência terminadora, marcadores selecionáveis, locais de integração e outros.

[004] A montagem dos elementos em um vetor simples usando-se métodos com base em enzima de restrição e ligação tradicionais pode ser consumidora de tempo e trabalhosa. Cada etapa de sub-clonagem, isto é, a

introdução de um novo fragmento de ácido nucleico em um polinucleotídeo existente, pode requerer que o clone resultante seja avaliado e caracterizado antes da introdução de fragmentos adicionais. Os clones produzidos pela ligação de extremidade abrupta requerem confirmação que o fragmento foi introduzido na orientação apropriada. Por outro lado, a ligação de extremidade pegajosa requer que os locais de restrição utilizados produzem as extremidades pegajosas do fragmento receptor também estão presentes no fragmento doador, mas não em um local que deve interromper a sequência de interesse dentro do fragmento doador. Desta maneira, a seleção de locais de restrição trabalháveis depende totalmente das composições dos pedaços sendo montados e devem ser cuidadosamente considerados em cada caso. Além disso, estes métodos frequentemente introduzem sequências de ácido nucleico estranhas ao clone resultante que pode interferir com a estrutura e função dos produtos genéticos desejados. A limitação adicional da eficiência dos métodos de clonagem com base na enzima de restrição é a limitação intrínseca de diversas moléculas de ácido nucleico que podem ser ligadas juntas em uma reação simples.

[005] A reação da cadeia de polimerase (PCR) é uma técnica poderosa pela qual as sequências de nucleotídeo específicas, incluindo DNA, cDNA e mRNA genômicos, são amplificados *in vitro*. O PCR tipicamente compreende contactar os filamentos complementares separados de um ácido nucleico alvo com dois iniciadores de oligonucleotídeo sob condições que permitem a formação dos produtos de extensão de iniciador complementar em ambos os filamentos. Estes filamentos atuam como padrões para a síntese de cópias das sequências de ácido nucleico desejadas. Repetindo-se as etapas de separação e síntese em um sistema automatizado, a duplicação exponencial das sequências alvo podem ser atingidas.

[006] Um método de PCR, denominado "união por extensão de sobreposição" ("SOE"; ver, por exemplo Patente U. S. N° 5,023,171), facilita

a montagem de moléculas de DNA em junções precisas sem o uso de enzimas de restrição ou ligase. Os fragmentos de componente a serem recombinados são gerados em reações de cadeia de polimerase separadas usando-se iniciadores unicamente projetados que produzem amplicons tendo terminais de complementaridade um com o outro. Na mistura e na desnaturação destes amplicons, os filamentos tendo sequências complementares em sua sobreposição de extremidades 3' e atuam como iniciadores um com o outro. A extensão desta sobreposição pela DNA polimerase produz a molécula de ácido nucleico em que as sequências originais são “unidas” juntas. As séries subsequentes de PCR amplifica o polinucleotídeo montado resultante.

[007] SOE, enquanto mais eficiente do que os métodos com base em enzima de ligação tradicional para a combinação de uma pluralidade de fragmentos de ácido nucleico, não requer tempo para otimizar as sequências iniciadoras e condições de amplificação para a produção dos produtos desejados. Cada junção entre os fragmentos a serem montados juntos deve ser individualmente considerada e um par de iniciadores deve ser projetado para cada fragmento a fim de tornar as extremidades compatíveis. As considerações tradicionais para o projeto de iniciadores de PCR, por exemplo, temperatura de fusão, teor de G-C, evitar a formação de grampo de cabelo e dímero e estringência para os locais de imprimação falsa, ainda deve ser considerada mais cuidadosamente quando o número de fragmentos a serem montados na reação SOE aumenta.

[008] Desta maneira, a despeito dos avanços na tecnologia de DNA recombinante, existe uma necessidade quanto a métodos melhorados que fornecem a montagem rápida e ordenada de polinucleotídeos. São particularmente necessários métodos que podem facilitar a montagem de diversos polinucleotídeos com a manipulação e a caracterização mínimas de produtos intermediários e sem a necessidade das etapas de otimização de iniciador. Estas e outras necessidades podem ser satisfeitas por composições e

métodos da presente invenção.

### 3. SUMÁRIO DA INVENÇÃO

[009] As composições e métodos fornecidos aqui permitem a montagem rápida e ordenada ou "costura" de polinucleotídeos componentes em polinucleotídeo montados. Em algumas formas de realização, os métodos fornecidos aqui utilizam vetores de montagem de ácido nucleico circulares. Em certas formas de realização, um vetor de montagem compreende um polinucleotídeo componente em que o polinucleotídeo componente compreende um segmento de DNA flanqueado por: (i) um ligante anelável na extremidade 3'; (ii) um segmento de ligação de iniciador na extremidade 5' e um ligante anelável na extremidade 3'; (iii) um ligante anelável tanto na extremidade 3' quanto na extremidade 5'; (vi) um ligante anelável na extremidade 5' e segmento de ligação de iniciador na extremidade 3' ou (v) um ligante anelável na extremidade 5'.

[0010] Em algumas formas de realização, uma pluralidade de polinucleotídeos componentes podem ser montados fornecendo uma pluralidade de vetores de montagem em um recipiente de reação simples. Em certas formas de realização, os polinucleotídeos componentes pode ser cortados de seus vetores de montagem dentro do recipiente de reação. Em algumas formas de realização, os polinucleotídeos componentes podem ser desnaturados, as sequências de ligante anelável podem ser aneladas a filamentos complementares em um polinucleotídeo componente adjacente e os polinucleotídeos componentes podem ser montados em um polinucleotídeo montado unindo-se ela extensão de sobreposição (SOE) seguido por PCR. Em outras formas de realização, os polinucleotídeos componentes cortados dos vetores de montagem podem ser montados em um polinucleotídeo montado *in vivo* pela recombinação homóloga dentro de uma célula hospedeira transformada com os polinucleotídeos componentes. Os polinucleotídeos montados ainda podem ser combinados *in vivo* por uma recombinação

homóloga mediada por célula hospedeira.

[0011] A eficiência da montagem de polinucleotídeo pode ser intensificada pelo fornecimento de uma série padrão de sequências de ligante anelável e são usados dentro do vetor de montagem, por exemplo, aqueles descritos como SEQ ID N°: 1 a 23. As sequências de ligante anelável fornecem sobreposição de sequência entre polinucleotídeos componentes adjacentes na reação de montagem. Idealmente, as sequências de ligante anelável precisam de estrutura secundária estimável tanto no nível de DNA quanto de RNA, não reagem de uma maneira indesejada um com o outro e têm temperaturas de fusão relativamente altas ( $T_m$ ). Consequentemente, diversos polinucleotídeos componentes podem ser montados sem a necessidade de projetar iniciadores únicos para cada polinucleotídeo componente, desse modo economizando tempo e trabalho. As composições e métodos fornecidos aqui também podem ser usados para montar muitos tipos de polinucleotídeos, incluindo genes sintéticos, construções, vetores de clonagem, vetores de expressão, cromossomos, genomas, bibliotecas de peptídeo e outros.

[0012] Em um aspecto, é fornecido um vetor, isto é, um vetor de montagem, que pode ser usado na montagem de um ou mais polinucleotídeos montados a partir de uma pluralidade de polinucleotídeos componentes.

[0013] Em algumas formas de realização, o vetor de montagem é um polinucleotídeo circular que compreende, em uma orientação de 5' a 3', um local de restrição RA, uma sequência de ligante anelável LA, um segmento de DNA D, uma sequência de ligante anelável LB e um local de restrição RB (isto é, 5'-RALA-D-LB-RB-3'). Em algumas formas de realização, o vetor de montagem é um polinucleotídeo circular que compreende, em uma orientação de 5' a 3', um local de restrição RA, um segmento de DNA D, uma sequência de ligante anelável LB e um local de restrição RB (isto é, 5'-RA-D-LB-RB-3'). Em algumas formas de realização, o vetor de montagem é um

polinucleotídeo circular que compreende, em uma orientação de 5' a 3', um local de restrição RA, uma sequência de ligante anelável LA, um segmento de DNA D e um local de restrição RB (isto é, 5'- RA-LA-D-RB-3'). Em algumas formas de realização, o vetor de montagem é um polinucleotídeo circular que compreende, em uma orientação de 5' a 3', um local de restrição RA, um segmento de ligação de iniciador PA, um segmento de DNA D, uma sequência de ligante anelável LB e um local de restrição RB (isto é, 5'-RA-PA-D-LB-RB-3'). Em algumas formas de realização, o vetor de montagem é um polinucleotídeo circular que compreende, em uma orientação de 5' a 3', um local de restrição RA, uma sequência de ligante anelável LA, um segmento de DNA D, um segmento de ligação de iniciador PB e um local de restrição RB (isto é, 5'- RA-LA-D-PB-RB-3'). Em algumas formas de realização, o vetor de montagem é um polinucleotídeo circular que compreende, em uma orientação de 5' a 3', um local de restrição RA, uma sequência de ligante anelável LA, um segmento de DNA D, uma sequência de ligante anelável LB e um local de restrição RB (isto é, 5'-RA-LA-D-LB-RB-3'). Os vetores de montagem exemplares são fornecidos em FIG. 1B e FIG. 2.

[0014] Em algumas formas de realização, um segmento de ligação de iniciador (isto é, PA ou PB) pode ser qualquer sequência de nucleotídeo que não é complementar com qualquer uma das sequências de ligante anelável que são usadas para preparar um polinucleotídeo montado. Em algumas formas de realização, um segmento de ligação de iniciador inclui uma endonuclease de local de reconhecimento de restrição e/ou local de clivagem. Em algumas formas de realização, um segmento de ligação de iniciador compreende uma sequência de ácido nucleico de uma das sequências de ligante disponíveis (por exemplo, um da SEQ ID N°: 1 a 23) ou complementos do mesmo, não sendo usados na reação de montagem particular. Em algumas formas de realização, a sequência de ácido nucleico do segmento de ligação de iniciador PA é selecionado do grupo que consiste

de SEQ ID N°: 24, 25 e complementos do mesmo. Em algumas formas de realização, a sequência de ácido nucleico do segmento de ligação de iniciador PB é selecionado do grupo que consiste de SEQ ID N°: 24, 25 e complementos do mesmo. Nas formas de realização preferíveis, segmento de ligação de iniciador PA e segmento de ligação de iniciador PB não são idênticas na sequência.

[0015] Em algumas formas de realização, as duas ou mais sequências de ligante anelável são pelo menos 24 nucleotídeos de comprimento e têm um  $T_m$  de pelo menos 60°C.

[0016] Em algumas formas de realização, duas ou mais sequências de ligante anelável têm um teor de G-C de pelo menos 70% e um  $T_m$  de pelo menos 70°C e não forma estruturas de DNA secundária estimáveis. Em algumas formas de realização, a sequência de nucleotídeo da sequência de ligante anelável LA é selecionado do grupo que consiste de SEQ ID N°: 1 a 8 e complementos do mesmo. Em algumas formas de realização, a sequência de nucleotídeo da sequência de ligante anelável LB é selecionado do grupo que consiste de SEQ ID N°: 1 a 8 e complementos do mesmo. Em algumas formas de realização, as sequências nucleicas de sequência de ligante anelável LA e sequência de ligante anelável LB são selecionados do grupo que consiste de SEQ ID N°: 1 a 8 e complementos do mesmo.

[0017] Em algumas formas de realização, duas ou mais sequências de ligante anelável têm um teor de A-T de pelo menos 30% e um  $T_m$  de pelo menos 65°C e não forma estruturas de DNA ou RNA secundárias estimáveis. Em algumas formas de realização, duas ou mais sequências de ligante anelável têm um teor de G-C baixo e um  $T_m$  de pelo menos 65°C e compreende o motivo de sequência 5'- ANNNAANTANNTTNANA-3', em que A significa adenina, N para qualquer nucleotídeo e T para timina. Em algumas formas de realização, a sequência de nucleotídeo da sequência de ligante anelável LA é selecionado do grupo que consiste de SEQ ID N°: 9 a



23 e complementos do mesmo. Em algumas formas de realização, a sequência de nucleotídeo da sequência de ligante anelável LB é selecionado do grupo que consiste de SEQ ID N°: 9 a 23 e complementos do mesmo. Em algumas formas de realização, as sequências nucleicas de sequência de ligante anelável LA e sequência de ligante anelável LB são selecionados do grupo que consiste de SEQ ID N°: 9 a 23 e complementos do mesmo.

[0018] A montagem ordenada de uma pluralidade de polinucleotídeos componentes pode ser controlada pela seleção de sequências de ligante anelável que flanqueiam um segmento de DNA dentro do vetor de montagem. Consequentemente, em algumas formas de realização, para garantir que os polinucleotídeos componentes podem ser montados de uma maneira ordenada, as sequências de um par de sequência de ligante anelável/sequência de ligante anelável dentro de um vetor de montagem particular não são complementares. Similarmente, em algumas formas de realização, as sequências de um segmento de ligação de iniciador/sequência de ligante anelável dentro de um vetor de montagem particular não são complementares.

[0019] Em uma forma de realização particular, os locais de restrição RA e RB são clivados pela mesma endonuclease de restrição a fim de facilitar o corte do polinucleotídeo componente do vetor de montagem. Em algumas formas de realização, o local de restrição RA ou RB é clivável por uma endonuclease de restrição que cliva uma projeção 5' ou 3'. Em outras formas de realização, o local de restrição RA ou RB é clivável por uma endonuclease de restrição que cliva uma extremidade abrupta. Em algumas formas de realização, os locais de restrição RA e RB são clivados pela mesma endonuclease de restrição. Ainda, em outras formas de realização, os locais de restrição RA e RB são clivados por uma endonuclease de restrição Tipo IIS. Em algumas formas de realização, os locais de restrição RA e RB são clivados pela mesma endonuclease de restrição Tipo IIS. Em uma forma de realização particular, os locais de restrição RA e RB são cliváveis por

endonucleases de restrição SapI ou LguI.

[0020] Em um outro aspecto, a invenção fornece um vetor de entrada útil na preparação de um vetor de montagem que compreende um segmento de DNA.

[0021] Em algumas formas de realização, o vetor de entrada é um polinucleotídeo circular que compreende, em uma orientação de 5' a 3', um local de restrição RA, um local de restrição RY, um segmento de DNA D, um local de restrição RZ, uma sequência de ligante anelável LB e um local de restrição RB (isto é, 5'-RA-RY-D-RZ-LB-RB-3'). Em algumas formas de realização, o vetor de entrada é um polinucleotídeo circular que compreende, em uma orientação de 5' a 3', um local de restrição RA, uma sequência de ligante anelável LA, um local de restrição RY, um segmento de DNA D, um local de restrição RZ e um local de restrição RB (isto é, 5'-RA-LA-RY-D-RZ-RB-3'). Em algumas formas de realização, o vetor de entrada é um polinucleotídeo circular que compreende, em uma orientação de 5' a 3', um local de restrição RA, uma sequência de ligante anelável LA, um local de restrição RY, um segmento de DNA D, um local de restrição RZ, uma sequência de ligante anelável LB e um local de restrição RB (isto é, 5'-RA-LA-RY-D-RZ-LB-RB-3'). Em algumas formas de realização, o vetor de entrada é um polinucleotídeo circular que compreende, em uma orientação de 5' a 3', um local de restrição RA, um segmento de ligação de iniciador PA, um local de restrição RY, um segmento de DNA D, um local de restrição RZ, uma sequência de ligante anelável LB e um local de restrição RB (isto é, 5'-RA-PA-RY-D-RZ-LB-RB-3'). Em algumas formas de realização, o vetor de entrada é um polinucleotídeo circular que compreende, em uma orientação de 5' a 3', um local de restrição RA, uma sequência de ligante anelável LA, um local de restrição RY, um segmento de DNA D, um local de restrição RZ, um segmento de ligação de iniciador PB e um local de restrição RB (isto é, 5'-RA-LA-RY-D-RZ-PB-RB-3'). Um vetor de entrada exemplar é fornecido em

FIG. 1A.

[0022] A digestão de um vetor de entrada com uma ou mais endonucleases de restrição capazes de clivar RY e RZ pode criar um vetor linearizado capaz de aceitar um segmento de DNA. O segmento de DNA pode ser ligado em locais RY e RZ usando-se técnicas de clonagem padrão para a geração de um vetor de montagem da invenção. Em algumas formas de realização, os locais de restrição RY e RZ do vetor de entrada são clivados pela mesma endonuclease de restrição. Em algumas formas de realização, os locais de restrição RY e RZ do vetor de entrada são clivados por uma endonuclease de restrição Tipo IIS. Em algumas formas de realização, os locais de restrição RY e RZ do vetor de entrada são clivados pela mesma endonuclease de restrição Tipo IIS. Nas formas de realização particulares, a endonuclease de restrição Tipo IIS é SchI ou MlyI.

[0023] Em algumas formas de realização, os locais de restrição RA e RB do vetor de entrada são clivados pela mesma endonuclease de restrição. Em algumas formas de realização, os locais de restrição RA e RB do vetor de entrada são clivados por uma endonuclease de restrição Tipo IIS. Em algumas formas de realização, os locais de restrição RA e RB do vetor de entrada são clivados pela mesma endonuclease de restrição Tipo IIS. Nas formas de realização particulares, a endonuclease de restrição Tipo IIS é SapI ou LguI.

[0024] Em um outro aspecto, a invenção fornece uma composição de montagem que compreende uma pluralidade de vetores de montagem para o uso na montagem de um ou mais polinucleotídeos montados a partir de uma pluralidade de polinucleotídeos componentes. Em algumas formas de realização, a composição de montagem compreende:

(a) uma ou mais primeiras moléculas de ácidos nucleicos, em que cada primeira molécula de ácido nucleico é circular e compreende, em uma orientação de 5' a 3', um primeiro local de restrição RA<sub>0</sub>, qualquer segmento de DNA selecionado do grupo D<sub>0</sub>, uma sequência de ligante

anelável  $LB_0$ , e um segundo local de restrição  $RB_0$ ;

(b) uma ou mais moléculas de ácido nucleico intermediárias em que cada molécula de ácido nucleico intermediária  $n$  é circular e compreende, em uma orientação de 5' a 3', um primeiro local de restrição  $RA_n$ , uma primeira sequência de ligante anelável  $LA_n$ , qualquer segmento de DNA selecionado do grupo  $D_n$ , uma segunda sequência de ligante anelável  $LB_n$ , e um segundo local de restrição  $RB_n$  e em que  $n$  representa um número inteiro de uma das diversas moléculas de ácido nucleico intermediárias e

(c) uma ou mais das últimas moléculas de ácido nucleico, em que cada última molécula de aminoácido é circular e compreende, em uma orientação de 5' a 3', um primeiro local de restrição  $RA_m$ , uma sequência de ligante anelável  $LA_m$ , qualquer segmento de DNA selecionado do grupo  $D_m$ , um segundo local de restrição  $RB_m$  em que  $m$  representa um número inteiro, um maior do que o número de moléculas de ácido nucleico intermediárias;

após o que a clivagem dos locais de restrição  $RA_0$  até  $RB_m$  e a desnaturação das moléculas de ácido nucleico lineares resultantes, cada sequência de ligante anelável  $LB_{(p-1)}$  é capaz de hibridizar ao complemento de sequência de ligante anelável  $LA_p$ , em que  $n$  é um número inteiro que varia de 1 a  $(m-1)$ , em que  $p$  representa um número inteiro de 1 a  $m$  e em que cada grupo  $D_0, \dots D_n, \dots$  e  $D_m$  consiste de um ou mais segmentos de DNA.

[0025] Em certas formas de realização, uma ou mais primeiras moléculas de ácidos nucleicos compreende adicionalmente um segmento de ligação de iniciador PA posicionado 5' com relação ao segmento de DNA selecionado do grupo  $D_0$ . Em certas formas de realização, uma ou mais das últimas moléculas de ácido nucleico compreende adicionalmente um segmento de ligação de iniciador PB posicionado 3' com relação ao segmento de DNA selecionado do grupo  $D_m$ .

[0026] Em certas formas de realização, a composição de montagem compreende duas ou mais moléculas de ácido nucleico intermediárias. Em

certas formas de realização, a composição de montagem compreende três ou mais moléculas de ácido nucleico intermediárias. Em certas formas de realização, a composição de montagem compreende quatro ou mais moléculas de ácido nucleico intermediárias. Em certas formas de realização, a composição de montagem compreende cinco ou mais moléculas de ácido nucleico intermediárias. Em certas formas de realização da montagem, a composição compreende seis ou mais moléculas de ácido nucleico intermediárias. Em certas formas de realização, a composição de montagem compreende sete ou mais moléculas de ácido nucleico intermediárias. Em certas formas de realização, a composição de montagem compreende oito ou mais moléculas de ácido nucleico intermediárias. Em certas formas de realização, a composição de montagem compreende nove ou mais moléculas de ácido nucleico intermediárias. Em certas formas de realização, a composição de montagem compreende dez ou mais moléculas de ácido nucleico intermediárias. Em certas formas de realização, a composição de montagem compreende quinze ou mais moléculas de ácido nucleico intermediárias. Em certas formas de realização, a composição de montagem compreende vinte ou mais moléculas de ácido nucleico intermediárias.

[0027] Em certas formas de realização,  $m$  é igual a 1. Em certas formas de realização,  $m$  é igual a 2. Em certas formas de realização,  $m$  é igual a 3. Em certas formas de realização,  $m$  é igual a 4. Em certas formas de realização,  $m$  é igual a 5. Em certas formas de realização,  $m$  é igual a 6. Em certas formas de realização,  $m$  é igual a 7. Em certas formas de realização,  $m$  é igual a 8. Em certas formas de realização,  $m$  é igual a 9. Em certas formas de realização,  $m$  é igual a 10. Em certas formas de realização,  $m$  é igual a ou maior do que 10.

[0028] Em algumas formas de realização, na clivagem dos locais de restrição  $RA_0$  até  $RB_m$  e a desnaturação das moléculas de ácido nucleico lineares resultantes, cada sequência de ligante anelável  $LB_{(p-1)}$  é capaz de

hibridizar seletivamente ao complemento da sequência de ligante anelável  $LA_p$  em comparação com outras sequências de ligante anelável ou seus componentes, em uma composição de montagem. Em algumas formas de realização, cada sequência de ligante anelável  $L_{(p-1)}$  é idêntico na sequência à sequência de ligante anelável  $LA_p$ .

[0029] Em uma forma de realização particular, os locais de restrição  $RA_0$  até  $RB_m$ , são clivados pela mesma endonuclease de restrição a fim de facilitar o corte dos polinucleotídeos componentes dos vetores de montagem. Em algumas formas de realização, os locais de restrição  $RA_0$  até  $RB_m$  são cliváveis pelas endonucleases de restrição SspI e/ou LguI.

[0030] Em um outro aspecto, a invenção fornece uma composição de componentes que compreende uma pluralidade de moléculas de ácido nucleico lineares em que as moléculas de ácido nucleico lineares pode ser formadas pela digestão de uma composição de montagem com uma ou mais endonucleases de restrição capazes de clivar os locais de restrição  $RA_0$  até  $RB_m$ , em que a composição de montagem compreende:

(a) uma ou mais primeiras moléculas de ácidos nucleicos, em que cada primeira molécula de ácido nucleico é circular e compreende, em uma orientação de 5' a 3', um primeiro local de restrição  $RA_0$ , qualquer segmento de DNA selecionado do grupo  $D_0$ , uma sequência de ligante anelável  $LB_0$  e um segundo local de restrição  $RB_0$ ;

(b) uma ou mais moléculas de ácido nucleico intermediárias em que cada molécula de ácido nucleico intermediária  $n$  é circular e compreende, em uma orientação de 5' a 3', um primeiro local de restrição  $RA_n$ , uma primeira sequência de ligante anelável  $LA_n$ , qualquer segmento de DNA selecionado do grupo  $D_n$ , uma segunda sequência de ligante anelável  $LB_n$ , e um segundo local de restrição  $RB_n$  e em que  $n$  representa um número inteiro de uma das diversas moléculas de ácido nucleico intermediárias e

(c) uma ou mais das últimas moléculas de ácido nucleico, em

que cada última molécula de aminoácido é circular e compreende, em uma orientação de 5' a 3', um primeiro local de restrição  $RA_m$ , uma sequência de ligante anelável  $LA_m$ , qualquer segmento de DNA selecionado do grupo  $D_m$ , um segundo local de restrição  $RB_m$ , em que  $m$  representa um número inteiro, um maior do que o número de moléculas de ácido nucleico intermediárias;

após o que a clivagem dos locais de restrição  $RA_0$  até  $RB_n$  e a desnaturação das moléculas de ácido nucleico lineares resultantes, cada sequência de ligante anelável  $LB_{(p-1)}$  é capaz de hibridizar ao complemento de sequência de ligante anelável  $LA_p$ , em que  $n$  é um número inteiro que varia de 1 a  $(m-1)$ , em que  $p$  representa um número inteiro de 1 a  $m$  e em que cada grupo  $D_0, \dots, D_n, \dots, D_m$  consiste de um ou mais segmentos de DNA.

[0031] Em certas formas de realização, uma ou mais primeiras moléculas de ácidos nucleicos compreende adicionalmente um segmento de ligação de iniciador PA posicionado 5' com relação ao segmento de DNA selecionado do grupo  $D_0$ . Em certas formas de realização, uma ou mais das últimas moléculas de ácido nucleico compreende adicionalmente um segmento de ligação de iniciador PB posicionado 3' com relação ao segmento de DNA selecionado do grupo  $D_m$ .

[0032] Em certas formas de realização, a composição de componentes compreende duas ou mais moléculas de ácido nucleico intermediárias. Em certas formas de realização, a composição de componentes compreende três ou mais moléculas de ácido nucleico intermediárias. Em certas formas de realização, a composição de componentes compreende quatro ou mais moléculas de ácido nucleico intermediárias. Em certas formas de realização, a composição de componentes compreende cinco ou mais moléculas de ácido nucleico intermediárias. Em certas formas de realização, a composição de componentes compreende seis ou mais moléculas de ácido nucleico intermediárias. Em certas formas de realização, a composição de componentes compreende sete ou mais moléculas de ácido nucleico

intermediárias. Em certas formas de realização, a composição de montagem compreende oito ou mais moléculas de ácido nucleico intermediárias. Em certas formas de realização, a composição de montagem compreende nove ou mais moléculas de ácido nucleico intermediárias. Em certas formas de realização, a composição de montagem compreende dez ou mais moléculas de ácido nucleico intermediárias. Em certas formas de realização, a composição de montagem compreende quinze ou mais moléculas de ácido nucleico intermediárias. Em certas formas de realização, a composição de montagem compreende vinte ou mais moléculas de ácido nucleico intermediárias.

[0033] Em certas formas de realização,  $m$  é igual a 1. Em certas formas de realização,  $m$  é igual a 2. Em certas formas de realização,  $m$  é igual a 3. Em certas formas de realização,  $m$  é igual a 4. Em certas formas de realização,  $m$  é igual a 5. Em certas formas de realização,  $m$  é igual a 6. Em certas formas de realização,  $m$  é igual a 7. Em certas formas de realização,  $m$  é igual a 8. Em certas formas de realização,  $m$  é igual a 9. Em certas formas de realização,  $m$  é igual a 10. Em certas formas de realização,  $m$  é igual a ou maior do que 10.

[0034] Em um outro aspecto, é fornecido aqui, um kit útil para montagem de uma pluralidade de polinucleotídeo de acordo com os métodos fornecidos aqui. Em algumas formas de realização, o kit compreende: (a) um ou mais vetores de entrada descrito neste; (b) uma ou mais endonucleases de restrição capazes de clivar os locais de restrição RA e RB do vetores de entrada e (c) uma ou mais endonucleases de restrição capazes de clivar os locais de restrição RY e RZ do vetores de entrada.

[0035] Em um outro aspecto, a invenção fornece uma biblioteca de moléculas de ácido nucleico.

[0036] Em algumas formas de realização, a molécula de ácido nucleico de uma biblioteca compreende um primeiro local de restrição RA, um segmento de DNA D, uma sequência de ligante anelável LB, e um



segundo local de restrição RB. Em algumas formas de realização, a molécula de ácido nucleico de uma biblioteca compreende um primeiro local de restrição RA, um segmento de ligação de iniciador PA, um segmento de DNA D, uma sequência de ligante anelável LB, e um segundo local de restrição RB. Em algumas formas de realização, a molécula de ácido nucleico de uma biblioteca compreende um primeiro local de restrição RA, uma sequência de ligante anelável LA, um segmento de DNA D, uma sequência de ligante anelável LB, e um segundo local de restrição RB. Em algumas formas de realização, a molécula de ácido nucleico de uma biblioteca compreende um primeiro local de restrição RA, uma sequência de ligante anelável LA, um segmento de DNA D, e um segundo local de restrição RB. Em algumas formas de realização, uma molécula de ácido nucleico de uma biblioteca compreende um primeiro local de restrição RA, uma sequência de ligante anelável LA, um segmento de DNA D, um segmento de ligação de iniciador PB, e um segundo local de restrição RB.

[0037] Em algumas formas de realização, a biblioteca compreende pelo menos um de cada um dos seguintes vetores:

(a) um vetor que consiste de um polinucleotídeo circular que compreende, em uma orientação de 5' a 3', um local de restrição RA, um segmento de DNA D, uma sequência de ligante anelável LB e um local de restrição RB;

(b) um vetor que consiste de um polinucleotídeo circular que compreende, em uma orientação de 5' a 3', um local de restrição RA, uma sequência de ligante anelável LA, um segmento de DNA D, uma sequência de ligante anelável LB e um local de restrição RB e

(c) um vetor que consiste de um polinucleotídeo circular que compreende, em uma orientação de 5' a 3', um local de restrição RA, uma sequência de ligante anelável LA, um segmento de DNA D e um local de restrição RB<sub>0</sub>.

[0038] Em algumas formas de realização, a biblioteca compreende pelo menos um de cada um dos seguintes vetores:

(a) um vetor que consiste de um polinucleotídeo circular que compreende, em uma orientação de 5' a 3', um local de restrição RA, um segmento de ligação de iniciador PA, um segmento de DNA D, uma sequência de ligante anelável LB e um local de restrição RB;

(b) um vetor que consiste de um polinucleotídeo circular que compreende, em uma orientação de 5' a 3', um local de restrição RA, uma sequência de ligante anelável LA, um segmento de DNA D, uma sequência de ligante anelável LB e um local de restrição RB e

(c) um vetor que consiste de um polinucleotídeo circular que compreende, em uma orientação de 5' a 3', um local de restrição RA, uma sequência de ligante anelável LA, um segmento de DNA D, um segmento de ligação de iniciador PB e um local de restrição RB<sub>0</sub>.

[0039] Em algumas formas de realização, o segmento de DNA D compreende uma sequência nucleíca selecionado do grupo que consiste de um marcador selecionável, um promotor, sequência alvejadora genômica, uma sequência de ácido nucleico que codifica um etiqueta de epítipo, e uma sequência de ácido nucleico que codifica um gene de interesse, uma sequência de ácido nucleico que codifica um códon de terminação e lacZ.

[0040] Em algumas formas de realização, a biblioteca compreende pelo menos uma de cada uma das seguintes moléculas de ácido nucleico:

(a) uma primeira molécula de ácido nucleico em que a primeira molécula de ácido nucleico é circular e compreende, em uma orientação de 5' a 3', um primeiro local de restrição RA<sub>0</sub>, um segmento de DNA D<sub>0</sub>, uma sequência de ligante anelável LB<sub>0</sub> e um segundo local de restrição RB<sub>0</sub>;

(b) uma molécula de ácido nucleico intermediária em que a molécula de ácido nucleico intermediária n é circular e compreende, em uma

orientação de 5' a 3', um primeiro local de restrição  $RA_n$ , uma primeira sequência de ligante anelável  $LA_n$ , um segmento de DNA  $D_n$ , uma segunda sequência de ligante anelável  $LB_n$  e um segundo local de restrição  $RB_n$  e em que  $n$  representa um número inteiro de uma das diversas moléculas de ácido nucleico intermediárias e

(c) a última molécula de aminoácido em que a última molécula de aminoácido é circular e compreende, em uma orientação de 5' a 3', um primeiro local de restrição  $RA_m$ , uma sequência de ligante anelável  $LA_m$ , um segmento de DNA  $D_m$ , um segundo local de restrição  $RB_m$  em que  $m$  representa um número inteiro, um maior do que o número de moléculas de ácido nucleico intermediárias;

após o que a clivagem dos locais de restrição  $RA_0$  até  $RB_m$  e a desnaturação das moléculas de ácido nucleico lineares resultantes, cada sequência de ligante anelável  $LB_{(p-1)}$  é capaz de hibridizar ao complemento de sequência de ligante anelável  $LA_p$  em que  $p$  representa os inteiros de 1 a  $m$ .

[0041] Em algumas formas de realização, uma primeira molécula de ácido nucleico compreende adicionalmente um segmento de ligação de iniciador PA posicionado 5' com relação ao segmento de DNA selecionado do grupo  $D_0$ . Em algumas formas de realização, a última molécula de aminoácidos compreende adicionalmente um segmento de ligação de iniciador PB posicionado 3' com relação ao segmento de DNA selecionado do grupo  $D_m$ .

[0042] Em certas formas de realização, a biblioteca compreende duas ou mais moléculas de ácido nucleico intermediárias. Em certas formas de realização, a biblioteca compreende três ou mais moléculas de ácido nucleico intermediárias. Em certas formas de realização, a biblioteca compreende quatro ou mais moléculas de ácido nucleico intermediárias. Em certas formas de realização, a biblioteca compreende cinco ou mais moléculas de ácido nucleico intermediárias. Em certas formas de realização, a biblioteca

compreende seis ou mais moléculas de ácido nucleico intermediárias. Em certas formas de realização, a biblioteca compreende sete ou mais moléculas de ácido nucleico intermediárias. Em certas formas de realização, a composição de montagem compreende oito ou mais moléculas de ácido nucleico intermediárias. Em certas formas de realização, a composição de montagem compreende nove ou mais moléculas de ácido nucleico intermediárias. Em certas formas de realização, a composição de montagem compreende dez ou mais moléculas de ácido nucleico intermediárias. Em certas formas de realização, a composição de montagem compreende quinze ou mais moléculas de ácido nucleico intermediárias. Em certas formas de realização, a composição de montagem compreende vinte ou mais moléculas de ácido nucleico intermediárias.

[0043] Em certas formas de realização,  $m$  é igual a 1. Em certas formas de realização,  $m$  é igual a 2. Em certas formas de realização,  $m$  é igual a 3. Em certas formas de realização,  $m$  é igual a 4. Em certas formas de realização,  $m$  é igual a 5. Em certas formas de realização,  $m$  é igual a 6. Em certas formas de realização,  $m$  é igual a 7. Em certas formas de realização,  $m$  é igual a 8. Em certas formas de realização,  $m$  é igual a 9. Em certas formas de realização,  $m$  é igual a 10. Em certas formas de realização,  $m$  é igual a ou maior do que 10.

[0044] Em um outro aspecto, são fornecidos aqui métodos de montar um ou mais polinucleotídeo montado a partir de uma pluralidade de polinucleotídeos componentes, que compreende as etapas de:

(a) digerir uma composição de montagem com uma ou mais endonucleases de restrição para gerar uma composição de componentes, a composição de montagem compreendendo:

(i) uma ou mais primeiras moléculas de ácidos nucleicos, em que cada primeira molécula de ácido nucleico é circular e compreende, em uma orientação de 5' a 3', um primeiro local de restrição  $RA_0$ , qualquer

segmento de ligação de iniciador selecionado do grupo PA, qualquer segmento de DNA selecionado do grupo  $D_0$ , uma sequência de ligante anelável  $LB_0$  e um segundo local de restrição  $RB_0$ ;

(ii) uma ou mais moléculas de ácido nucleico intermediárias em que cada molécula de ácido nucleico intermediária  $n$  é circular e compreende, em uma orientação de 5' a 3', um primeiro local de restrição  $RA_n$ , uma primeira sequência de ligante anelável  $LA_n$ , qualquer segmento de DNA selecionado do grupo  $D_n$ , uma segunda sequência de ligante anelável  $LB_n$ , e um segundo local de restrição  $RB_n$  e em que  $n$  representa um número inteiro de uma das diversas moléculas de ácido nucleico intermediárias e

(iii) uma ou mais das últimas moléculas de ácido nucleico, em que cada última molécula de aminoácido é circular e compreende, em uma orientação de 5' a 3', um primeiro local de restrição  $RA_m$ , uma sequência de ligante anelável  $LA_m$ , um segmento de DNA selecionado do grupo  $D_m$ , qualquer segmento de ligação de iniciador selecionado do grupo PB, um segundo local de restrição  $RB_m$  em que  $m$  representa um número inteiro, um maior do que o número de moléculas de ácido nucleico intermediárias; após o que a clivagem dos locais de restrição  $RA_0$  até  $RB_m$  e a desnaturação das moléculas de ácido nucleico lineares resultantes, cada sequência de ligante anelável  $LB_{(p-1)}$  é capaz de hibridizar ao complemento de sequência de ligante anelável  $LA_p$ , em que  $n$  é um número inteiro que varia de 1 a  $(m-1)$ , em que  $p$  representa um número inteiro de 1 a  $m$  e em que cada grupo  $D_0, \dots D_n, \dots$  e  $D_m$  consiste de um ou mais segmentos de DNA;

em que a uma ou mais endonucleases de restrição são capazes de clivar os locais de restrição  $RA_0$  até  $RB_m$  e

(b) contactar a composição de componentes com DNA polimerase, trifosfatos de desoxirribonucleosídeo e um ou mais primeiros iniciadores e um ou mais segundos iniciadores, sob condições adequadas para a desnaturação da moléculas de ácido nucleico, anelamento da sequência de

ligante anelável  $LB_{(p-1)}$  à sequência de ligante anelável  $LA_p$  e extensão deste; em que cada dito primeiro iniciador é capaz de hibridizar a um dos ditos segmentos de ligação de iniciador selecionado do grupo PA e cada dito segundo iniciador é capaz de hibridizar a um dos ditos segmentos de ligação de iniciador selecionado do grupo PB e submeter a composição de componentes à reação de cadeia de polimerase,

em que um polinucleotídeo é montado compreendendo, em uma orientação de 5' a 3', um segmento de DNA selecionado de cada um dos grupos  $D_0, \dots, D_n, \dots$  e  $D_m$ . No método,  $p$  representa os inteiros de 1 a  $m$ .

[0045] Em certas formas de realização, a composição de montagem compreende duas ou mais moléculas de ácido nucleico intermediárias. Em certas formas de realização, a composição de montagem compreende três ou mais moléculas de ácido nucleico intermediárias. Em certas formas de realização, a composição de montagem compreende quatro ou mais moléculas de ácido nucleico intermediárias. Em certas formas de realização, a composição de montagem compreende cinco ou mais moléculas de ácido nucleico intermediárias. Em certas formas de realização, a composição de montagem compreende seis ou mais moléculas de ácido nucleico intermediárias. Em certas formas de realização, a composição de montagem compreende sete ou mais moléculas de ácido nucleico intermediárias. Em certas formas de realização, a composição de montagem compreende oito ou mais moléculas de ácido nucleico intermediárias. Em certas formas de realização, a composição de montagem compreende nove ou mais moléculas de ácido nucleico intermediárias. Em certas formas de realização, a composição de montagem compreende dez ou mais moléculas de ácido nucleico intermediárias. Em certas formas de realização, a composição de montagem compreende quinze ou mais moléculas de ácido nucleico intermediárias. Em certas formas de realização, a composição de montagem compreende vinte ou mais moléculas de ácido nucleico intermediárias.

[0046] Em certas formas de realização, m é igual a 1. Em certas formas de realização, m é igual a 2. Em certas formas de realização, m é igual a 3. Em certas formas de realização, m é igual a 4. Em certas formas de realização, m é igual a 5. Em certas formas de realização, m é igual a 6. Em certas formas de realização, m é igual a 7. Em certas formas de realização, m é igual a 8. Em certas formas de realização, m é igual a 9. Em certas formas de realização, m é igual a 10. Em certas formas de realização, m é igual a ou maior do que 10.

[0047] Em algumas formas de realização, a composição de montagem compreende uma primeira molécula de ácido nucleico e uma última molécula de aminoácido. Em outras formas de realização, a composição de montagem compreende mais do que uma primeira molécula de ácido nucleico e mais do que uma última molécula de aminoácido e os métodos de montagem fornecem a montagem ordenada de polinucleotídeos componentes múltiplos em uma pluralidade de polinucleotídeo montado de uma maneira combinatória. Em certas formas de realização, a composição de montagem compreende pelo menos duas moléculas de ácido nucleico que compreendem a mesma sequência de ligante anelável LA ou LB ou o mesmo segmento de ligação de iniciador PA ou PB ou o mesmo par de sequências de ligante anelável LA e LB ou o mesmo par de sequência de ligante anelável/segmento de ligação de iniciador LA e PB ou LB e PA.

[0048] Em um outro aspecto, aqui, são fornecidos métodos para a geração de células hospedeiras que compreende polinucleotídeo montados. Em algumas formas de realização, os métodos compreendem transformar uma célula hospedeira com um polinucleotídeo montado gerado pelos métodos de montagem de polinucleotídeo descrito neste. Em outras formas de realização, os métodos compreendem transformar uma célula hospedeira com uma pluralidade de polinucleotídeo montado gerado pelos métodos de montagem de polinucleotídeo descrito neste. Em uma forma de realização particular, a

célula hospedeira dois ou mais polinucleotídeos montados em um ou mais polinucleotídeos combinados pela recombinação homóloga. Ainda, em outras formas de realização, os métodos compreendem transformar uma célula hospedeira com uma pluralidade de polinucleotídeos componentes e permitindo que a célula hospedeira gere um ou mais polinucleotídeos montados ou combinados pela recombinação homóloga.

[0049] Em um outro aspecto, a presente invenção fornece métodos para a geração de uma pluralidade de células hospedeiras que compreende uma pluralidade de polinucleotídeos montados. Em algumas formas de realização, a pluralidade de células hospedeiras é gerada pela transformação de células hospedeiras com uma composição que compreende uma pluralidade de polinucleotídeo montado gerada pela montagem combinatória de polinucleotídeos componentes. Em outras formas de realização, a pluralidade de células hospedeiras é gerada pela transformação das células hospedeiras com uma composição que compreende uma pluralidade de polinucleotídeo montado dos quais pelo menos dois polinucleotídeos montado compreendem segmentos não funcionais de um marcador selecionável que na recombinação homóloga mediada por célula hospedeira gera um marcador selecionável funcional e selecionando-se as células hospedeiras que compreendem um polinucleotídeo combinado. Ainda, em outras formas de realização, a pluralidade de células hospedeiras é gerada por métodos combinatórios pela transformação de células hospedeiras com uma composição de componente que compreende componentes de polinucleotídeo múltiplos dos quais pelo menos dois polinucleotídeos componentes compreendem a mesma sequência de ligante anelável LA ou LB ou o mesmo par de sequências de ligante anelável LA e LB e pela seleção de células hospedeiras que compreende um polinucleotídeo montado.

[0050] Em um outro aspecto, é fornecido um polinucleotídeo tendo uma sequência selecionada do grupo que consiste de SEQ ID N°: 1 a 25.



[0051] Em um outro aspecto, é fornecido um polinucleotídeo que compreende uma ou mais sequências selecionadas do grupo que consiste de SEQ ID N°: 1 a 25.

#### 4. BREVE DESCRIÇÃO DAS FIGURAS

[0052] A FIG. 1A fornece um esquema de um vetor de entrada útil para a preparação de um vetor de montagem da invenção. O vetor contém um local de restrição  $RA_0$ , um segmento de ligação de iniciador PA ou uma sequência de ligante anelável LA, um local de restrição RY, um segmento de DNA D, um local de restrição RZ, um segmento de ligação de iniciador PB ou uma sequência de ligante anelável LB e um local de restrição RB.

[0053] A FIG. 1B fornece um método exemplar de preparar um vetor de entrada para a aceitação de um segmento de DNA para formar um vetor de montagem. No Exemplo,  $RY = RZ = \text{SchI}$ . A digestão com SchI, a endonuclease de restrição Tipo IIS que é capaz de produzir extremidades abruptas permite o isolamento do vetor com os locais ligantes abertos para serem fundidos com relação ao segmento de DNA (D). A ligação de extremidade abrupta de D no vetor de entrada pode ser realizada por métodos tradicionais usando-se, por exemplo, T4 DNA ligase.

[0054] A FIG. 2 apresenta um esquema de uma composição de uma composição de montagem que compreende uma pluralidade de vetores de montagem (primeiro, intermediário e último), cada um dos quais compreendendo um segmento de DNA de interesse ( $D_0, D_m, D_n$ ). A primeira molécula de ácido nucleico compreende um primeiro local de restrição  $RA_0$ , um segmento de ligação de iniciador PA, um segmento de DNA  $D_0$ , uma sequência de ligante anelável  $LB_0$ , e um segundo local de restrição  $RB_0$ . A uma ou mais moléculas de ácido nucleico intermediárias compreendem um primeiro local de restrição  $RA_n$ , uma primeira sequência de ligante anelável  $LA_n$ , um segmento de DNA  $D_n$ , uma segunda sequência de ligante anelável  $LB_n$ , e um segundo local de restrição  $RB_n$  em que n representa um número

inteiro de uma das diversas moléculas de ácido nucleico intermediárias e a última molécula de aminoácido compreende um primeiro local de restrição  $RA_m$ , uma sequência de ligante anelável  $LA_m$ , um segmento de DNA  $D_m$ , um segmento de ligação de iniciador PB, um segundo local de restrição  $RB_m$  em que  $m$  representa um número inteiro, um maior do que o número de moléculas de ácido nucleico intermediárias.

[0055] A FIG. 3 apresenta um método exemplar de montagem, isto é, "costura" de um polinucleotídeo montado a partir de quatro (4) polinucleotídeos componentes. Os vetores de montagem que compreendem segmentos de DNA a serem montados são montados em um tubo simples e digeridos com SapI para a liberação dos fragmentos de polinucleotídeo componente das estruturas de vetor de montagem. Seguindo a inativação por calor de SapI, os fragmentos de polinucleotídeo componente estão sujeitos às condições de desnaturação, seguido por condições de anelamento suficientes para a hibridização dos pares aneláveis de ligante complementar. Seguindo a extensão de iniciador na presença de DNA polimerase e dNTPs, iniciadores complementares a PA e PB são adicionados, seguido pela amplificação de PCR tradicional. Um polinucleotídeo montado que compreende polinucleotídeos componentes  $D_0$ ,  $D_1$ ,  $D_2$  e  $D_3$  montado em uma direção 5' a 3' é produzido como um resultado da reação de montagem.

[0056] A FIG. 4 mostra um mapa do vetor de pRYSE.

[0057] A FIG. 5 mostra polinucleotídeo montado obtido pela montagem de 2 a 4 polinucleotídeos componentes (Montagens 1 até 6 na Tabela 7) usando-se métodos diferentes para a remoção da SapI endonuclease de restrição (purificação de coluna ou inativação por calor), concentrações de DNA de vetor de montagem diferente (5 ng (concentração de DNA baixa) ou 50 ng (concentração de DNA alta) de fragmento menor com a concentração molar igual de todos os outros fragmentos e temperaturas de anelamento diferentes para a amplificação de PCR (54°C e 72°C).

[0058] FIG. 6 mostra o polinucleotídeo montado obtido pela montagem de 6 ou 9 polinucleotídeos componentes (Montagens 7 e 13 até 16 na Tabela 7) usando-se DNA polimerases diferentes (Phusion (New England Biolabs, Ipswich, MA) e PfuUltraII (Stratagene/Agilent, La Jolla, CA)).

[0059] A FIG. 7 mostra um mapa do vetor pMULE. O vetor de entrada pMULE difere do vetor de entrada pRYSE em que este precisa de um segmento de ligação de iniciadores ou sequências de ligante anelável.

[0060] A FIG. 8 apresenta um método exemplar de combinar polinucleotídeo montado em um polinucleotídeo combinado pela recombinação homóloga mediada por célula hospedeira e interação do polinucleotídeo combinado em um cromossomo da célula hospedeira. O polinucleotídeo montado A compreende um segmento de DNA  $D_m$ , que codifica a primeiro segmento não funcional de um marcador selecionável e um segmento de DNA  $D_0$ , que codifica uma sequência alvejadora genômica a montante. O polinucleotídeo montado B compreende um segmento de DNA  $D_{m2}$  que codifica um segundo segmento não funcional do marcador selecionável e um segmento de DNA  $D_{02}$  que codifica a sequência alvejadora genômica a jusante. A célula hospedeira recombina o polinucleotídeo montado A e o polinucleotídeo montado B na região de homologia em segmentos de DNA  $D_{m1}$  e  $D_{m2}$  para formar um polinucleotídeo combinado que compreende um marcador selecionável funcional e usa as sequências alvejadoras genômicas codificadas por segmentos de DNA  $D_{01}$  e  $D_{02}$  para inserir o polinucleotídeo combinado pela recombinação homóloga em seu cromossomo.

[0061] A FIG. 9 apresenta um método exemplar de gerar um polinucleotídeo montado pela recombinação homóloga em uma célula hospedeira e integração do polinucleotídeo montado no cromossomo da célula hospedeira. Na primeira etapa, uma composição de montagem que compreende vetores de montagem é digerida com uma endonuclease de

restrição, resultando no corte de polinucleotídeos componentes do vetor de estruturas de montagem. Em uma segunda etapa, os polinucleotídeos componentes são introduzidos em uma célula hospedeira onde estas são recombinadas nas regiões de homologia nas sequências de ligante anelável para formar um polinucleotídeo montado e o polinucleotídeo montado é integrado no cromossomo da célula hospedeira.

[0062] A FIG. 10 apresenta um método exemplar de montagem uma pluralidade de polinucleotídeo montado de sete (7) polinucleotídeos componentes na mesma reação. Os vetores de montagem que compreendem segmentos de DNA a serem montados são únicos em um tubo simples e digerido com SapI para a liberação de polinucleotídeos componentes do vetor de estruturas de montagem. Seguindo a inativação por calor de SapI os fragmentos de polinucleotídeo componente são submetidos às condições de desnaturação, seguido por condições de anelamento suficientes para a hibridização dos pares de ligante anelável complementares. Seguindo a extensão de iniciador na presença de DNA polimerase e dNTPs, iniciadores complementares a PA e PB são adicionados, seguido por amplificação de PCR tradicional. A reação de montagem resulta na produção de um polinucleotídeo montado que compreende polinucleotídeos componentes D<sub>01/02</sub>, D<sub>1/2</sub>, D<sub>3</sub> e D<sub>41/42</sub> montado em uma direção 5' a 3'.

[0063] A FIG. 11 apresenta um método exemplar de gerar uma pluralidade de células hospedeiras que compreendem polinucleotídeos combinatoriamente combinados. Os polinucleotídeo montados A1 e A2, cada um compreendendo a mesma sequência alvejadora genômica a montante na mesma primeira porção não funcional de um marcador selecionável e polinucleotídeo montado B1 e B2, cada um compreendendo a mesma sequência alvejadora genômica a jusante e a mesma segunda função não funcional de um marcador selecionável, são combinatoriamente combinadas pela recombinação homóloga mediada por célula hospedeira para a geração

de quatro polinucleotídeos combinados diferentes, A1/B1, A1/B2, A2/B1 e A2/B2, cada um dos quais compreendendo um marcador selecionável funcional, que pode ser inserido em um cromossomo para a geração de quatro células hospedeiras diferentes.

[0064] FIG. 12A shows os polinucleotídeos componentes usados no Exemplo 10 para a geração de rendimento alto de polinucleotídeo combinatoriamente montado e células de levedura que compreendem polinucleotídeos combinatoriamente montados e combinados e os polinucleotídeos montados e combinados esperados. US = sequência alvejadora genômica a montante, DS = sequência alvejadora genômica a jusante, P = várias sequências promotoras, G = várias sequências codificadoras de proteína, URA = 5' segmento de marcador selecionável, RA3 = 3' segmento de marcador selecionável, PA = segmento de ligação de iniciador PmeI-5', PB = segmento de ligação de iniciador PmeI-3', LB<sub>0</sub> = sequência de ligante anelável RYSE 2, LA<sub>n1</sub> = sequência de ligante anelável

[0065] RYSE 2, LB<sub>n1</sub> = sequência de ligante anelável RYSE 15, LA<sub>n2</sub> = sequência de ligante anelável RYSE 3, LB<sub>n2</sub> = sequência de ligante anelável RYSE16, LA<sub>n3</sub> = sequência de ligante anelável RYSE 15, LB<sub>,,3</sub> = sequência de ligante anelável RYSE 3, LA<sub>n4</sub> = sequência de ligante anelável RYSE 16, LB<sub>n4</sub> = sequência de ligante anelável RYSE 4, LA<sub>m1</sub> = sequência de ligante anelável RYSE 3, LA<sub>m2</sub> = sequência de ligante anelável RYSE 4, LA<sub>m3</sub> = sequência de ligante anelável RYSE 3.

[0066] A FIG. 12B mostra polinucleotídeo montado exemplar (em caixa) gerado como descrito no Exemplo 10 e redissolvido em um gel de agarose a 1%.

[0067] A FIG. 12C mostra a análise de restrição para colônias celulares exemplares obtidas como descrito no Exemplo 10.

[0068] A FIG. 13A mostra o polinucleotídeo montado e os polinucleotídeos componentes usados no Exemplo 11 e o local cromossômico

esperado obtido na montagem e a integração cromossômica por uma célula hospedeira.

[0069] A FIG. 13B mostra os resultados da análise de cPCR obtidos por transformantes de célula de levedura gerados no Exemplo 11 que compreende polinucleotídeos cromossomicamente integrados montados.

[0070] A FIG. 14 mostra os polinucleotídeos componentes usados no Exemplo 12 para a geração de rendimento alto de células de levedura que compreendem polinucleotídeos cromossomicamente integrados, combinatoriamente montados e combinatoriamente combinados e os polinucleotídeos combinados esperados obtidos na montagem e na combinação por uma recombinação homóloga mediada por célula hospedeira. US = sequência alvejadora genômica a montante, DS = sequência alvejadora genômica a jusante, P = várias sequências promotoras, G = várias sequências codificadoras de proteína, URA = 5' segmento de marcador selecionável, RA3 = 3' segmento de marcador selecionável, PA = segmento de ligação de iniciador PmeI-5', PB = segmento de ligação de iniciador PmeI-3', LB<sub>0</sub> = sequência de ligante anelável RYSE 2, LA<sub>n1</sub> = sequência de ligante anelável RYSE 2, LB<sub>n1</sub> = sequência de ligante anelável RYSE 15, LA<sub>n2</sub> = sequência de ligante anelável RYSE 3, LB<sub>n2</sub> = sequência de ligante anelável RYSE16, LA<sub>n3</sub> = sequência de ligante anelável RYSE 15, LB<sub>n3</sub> = sequência de ligante anelável RYSE 3, LA<sub>n4</sub> = sequência de ligante anelável RYSE 16, LB<sub>n4</sub> = sequência de ligante anelável RYSE 4, LA<sub>m1</sub> = sequência de ligante anelável RYSE 3, LA<sub>m2</sub> = sequência de ligante anelável RYSE 4, LA<sub>m3</sub> = sequência de ligante anelável RYSE 3.

## 5. DESCRIÇÃO DETALHADA DAS FORMAS DE REALIZAÇÃO

### 5.1. Definições

[0071] Como usado neste, o termo "polinucleotídeo" refere-se a um polímero composto de unidades de nucleotídeo como deve ser entendido por uma pessoa habilitada na técnica. As unidades de nucleotídeo preferidas

incluem, mas não se limitam àquelas que compreendem adenina (A), guanina (G), citosina (C), timina (T) e uracila (U). as unidades de nucleotídeo modificadas úteis incluem mas não limitam-se àquelas que compreende 4-acetilcitidina, 5-(carboxiidroxilmetil)uridina, 2-O-metilcitidina, 5-carboximetilaminometil-2-tiouridina, 5-carboximetilamino-metiluridina, diidrouridina, 2-O-metilpseudouridina, 2-O-metilguanosina, inosina, N6-isopentiladenosina, 1-metiladenosina, 1-metilpseudouridina, 1-metilguanosina, 1-metilinosina, 2,2-dimetilguanosina, 2-metiladenosina, 2-metilguanosina, 3-metilcitidina, 5-metilcitidina, N6-metiladenosina, 7-metilguanosina, 5-metilaminometiluridina, 5-metoxiaminometil-2-tiouridina, 5-metoxiuridina, 5-metoxycarbonilmetil-2-tiouridina, 5-metoxycarbonilmetiluridina, 2-metiltio-N6-isopentiladenosina, éster metílico do ácido uridino-5-oxiacético, ácido uridino-5-oxiacético, wybutoxosina, wybutosina, pseudouridina, queuosina, 2-tiocitidina, 5-metil-2-tiouridina, 2-tiouridina, 4-tiouridina, 5-metiluridina, 2-O-metil-5-metiluridina, 2-O-metiluridina e outros. Os polinucleotídeos incluem ácidos nucleicos de ocorrência natural, tais como ácido desoxirribonucleico ("DNA") e ácido ribonucleico ("RNA"), bem como análogos de ácido nucleico. Os análogos de ácido nucleico incluem aqueles que incluem bases de ocorrência não natural, nucleotídeos que se dedicam em ligações com outros nucleotídeos que não a ligação de fosfodiéster de ocorrência natural ou que inclui bases ligadas através de outras ligações que não ligações de fosfodiéster. Desta maneira, os análogos de nucleotídeo incluem, por exemplo, e sem limitação, fosforotioates, fosforoditioatos, fosforotriésteres, fosforamidatos, boranofosfatos, metilfosfonates, fosfonatos de quirais-metila, 2-O-metil ribonucleotídeos, peptídeo-ácidos nucleicos (PNAs) e outros.

[0072] Como usado neste, um "componente de polinucleotídeo" refere-se a uma sequência de polinucleotídeo que pode ser unida para formar um "polinucleotídeo montado" usando-se os métodos de montagem de

polinucleotídeo descritos neste. Quando uma pluralidade de vetores de montagem é digerida com uma ou mais endonucleases de restrição capazes de cortar os polinucleotídeos componentes dos vetores de montagem, a população resultante de polinucleotídeos componentes pode compreender a totalidade dos segmentos de DNA a serem montados em um polinucleotídeo montado.

[0073] Como usado neste, um “polinucleotídeo montado” refere-se a um polinucleotídeo produzido pelos métodos de montagem de polinucleotídeo descrito neste. O polinucleotídeo montado pode ser compreendido dos dois ou mais polinucleotídeos componentes. Em algumas formas de realização, o polinucleotídeo montado compreende 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15 ou mais polinucleotídeos componentes. O comprimento do polinucleotídeo montado pode variar de cerca de 100 a cerca de 20.000 nucleotídeos, ou mais. Em algumas formas de realização, o comprimento do polinucleotídeo montado varia de cerca de 200 a cerca de 10.000, cerca de 200 a cerca de 8000, cerca de 200 a cerca de 5000, cerca de 200 a cerca de 3000 ou cerca de 200 a cerca de 1000 nucleotídeos. Em outras formas de realização, o comprimento do polinucleotídeo montado pode variar de cerca de 200 a cerca de 2000, cerca de 2000 a cerca de 5000, cerca de 5000 a cerca de 10.000, cerca de 10.000 a cerca de 20.000, ou maior do que 20.000 nucleotídeos.

[0074] A notação convencional é usada neste para descrever sequências de polinucleotídeo: a extremidade direita de uma sequência de polinucleotídeo de filamento simples é a extremidade 5'; a direção esquerda de uma sequência de polinucleotídeo de filamento duplo é referido como a direção 5'.

[0075] Como usado neste, o termo “segmento de DNA,” alternativamente referido como “Pedaços” nos exemplos abaixo, refere-se a qualquer molécula isolada ou isolável de DNA. Os exemplos incluem mas não limitam-se a uma sequência codificadora de proteína, gene repórter,



sequência codificador de marcador fluorescente, promotor, intensificador, terminador, íntron, éxon, cauda poly-A, local de clonagem múltiplo, sinal de localização nuclear, sinal de estabilização de mRNA, marcador selecionável, locais de integração, sequência codificadora de etiqueta de epítipo, sinal de degradação ou qualquer outra molécula de DNA de ocorrência natural ou sintética. Em algumas formas de realização, o segmento de DNA pode ser de origem natural. Alternativamente, um segmento de DNA pode ser completamente de origem sintética, produzido *in vitro*. Além disso, um segmento de DNA pode compreender qualquer combinação de moléculas de DNA de ocorrência natural isoladas ou qualquer combinação de uma molécula de DNA de ocorrência natural isolada e uma molécula de DNA sintética. Por Exemplo, um segmento de DNA pode compreender um promotor heterólogo operacionalmente ligado a uma sequência codificadora de proteína, uma sequência codificadora de proteína ligada a uma cauda poly-A, uma sequência codificadora de proteína ligada na estrutura com uma sequência codificadora de etiqueta de epítipo e outros.

[0076] "Complementar" refere-se à compatibilidade ou comparação topológica junto das superfícies de interação de dois polinucleotídeos como entendido por aqueles habilitados na técnica. Desta maneira, duas sequências são "complementares" umas às outras se estas forem capazes de hibridizar uma à outra para formar uma estrutura de ácido nucleico de filamento duplo anti-paralelo estável. Um primeiro polinucleotídeo é complementar a um segundo polinucleotídeo se a sequência de nucleotídeo do primeiro polinucleotídeo for substancialmente idêntica à sequência de nucleotídeo do parceiro de ligação de polinucleotídeo do segundo polinucleotídeo, ou se o primeiro polinucleotídeo pode hibridizar-se ao segundo polinucleotídeo sob condições de hibridização estringentes. Desta maneira, o polinucleotídeo cuja sequência 5'-TATAC-3' é complementar a um polinucleotídeo cuja sequência é 5'-GTATA-3'.

[0077] "Iniciador" refere-se a uma sequência de polinucleotídeo que é capaz de hibridizar especificamente a uma sequência padrão de polinucleotídeo, por exemplo, um segmento de ligação de iniciador e é capaz de fornecer um ponto de início para a síntese de um polinucleotídeo complementar sob condições adequadas para a síntese, isto é, na presença de nucleotídeos e um agente que catalisa a reação de síntese (por exemplo, uma DNA polimerase). O iniciador é complementar à sequência padrão de polinucleotídeo, mas não necessita ser um complemento exato da sequência de padrão de polinucleotídeo. Por exemplo, um iniciador pode ser de pelo menos cerca de 80, 85, 90, 95, 96, 97, 98 ou 99% idêntico ao complemento da sequência padrão de polinucleotídeo. Um iniciador pode ser de comprimento variável mas, em geral, é de pelo menos 15 bases. Em algumas formas de realização, o iniciador está entre 15 e 35 bases de comprimento. Em algumas formas de realização, o iniciador é mais do que 35 bases de comprimento. Em outras formas de realização, o iniciador tem uma temperatura de fusão ( $T_m$ ), isto é, a temperatura em que metade do duplex de DNA dissociará para tornar-se de filamento múltiplo, de pelo menos 50°C. Em outras formas de realização, o iniciador tem uma  $T_m$  entre cerca de 50°C e 70°C. Ainda, em outras formas de realização, o iniciador não forma estruturas secundárias de DNA ou RNA estimáveis de modo que não impactem a eficiência de hibridização à sequência padrão de polinucleotídeo.

[0078] Como usado neste, o termo "segmento de ligação de iniciador" é uma sequência de polinucleotídeo que liga-se a um iniciador a fim de fornecer um ponto de iniciação para a síntese de uma polinucleotídeo complementar sob condições adequadas para a síntese. Em algumas formas de realização, a sequência de ligação de iniciador é uma dos ligantes aneláveis da presente invenção. Uma sequência é uma sequência de ligante de iniciador em vez de um ligante anelável pela ausência de um ligante complementar dentro de uma dada série de vetores de montagem ou polinucleotídeos componentes

dentro de uma composição de montagem. Em algumas formas de realização, o segmento de ligação de iniciador pode funcionar como uma sequência alvejadora genômica, por exemplo, uma sequência alvejadora genômica a montante ou a jusante.

[0079] Como usado neste, o termo "sequência ligante" e "sequência de ligante anelável" são usados de maneira intercambiável e referem-se a uma sequência de polinucleotídeo contido dentro de um vetor de entrada e vetor de montagem descrito neste. Em particular, uma sequência de ligante anelável flanqueia um segmento de DNA dentro de um vetor de entrada ou vetor de montagem. No corte de um polinucleotídeo componente de um vetor de montagem e a desnaturação das polinucleotídeo componente, um ligante anelável é capaz de hibridizar especificamente a uma sequência de ligante anelável complementar de um polinucleotídeo componente adjacente em uma reação de montagem de polinucleotídeo, as descrito neste. Um ligante anelável, no anelamento com um filamento de ligante complementar, pode fornecer um ponto de início para a síntese de um polinucleotídeo complementar.

[0080] Como usado neste, o termo "vetor" é usado em referência a moléculas de ácido nucleico extracromossômicas capazes de replicação em uma célula e ao qual uma sequência de inserção pode ser operacionalmente ligada a fim de realizar a replicação da sequência de inserção. Os exemplos úteis incluem mas não são limitados a moléculas de DNA circulares tais como construções de plasmídeo, construções de fago, vetores de cosmídeo, *etc.*, bem como construções de ácido nucleico lineares (por exemplo, construções e fago lambda, cromossomos artificiais bacterianos (BACs), cromossomos artificiais de levedura (YACs), *etc.*). um vetor pode incluir sinais de expressão, tais como um promotor e/ou um terminador, um marcador selecionável tal como um gene que confere resistência a um antibiótico e um ou mais locais de restrição em que as sequências de inserção podem ser

clonadas. Os vetores podem ter outras características únicas (tais como o tamanho da inserção de DNA que estes podem acomodar).

[0081] Como usado neste, o termo "vetor de entrada" refere-se a um plasmídeo de vetor de clonagem que pode servir como um vetor parental para a preparação de um vetor de montagem a ser usado nos métodos de montagem de polinucleotídeo fornecidos neste. Um vetor de entrada compreende duas sequências de ligante anelável ou uma sequência de ligante anelável e um segmento de ligação de iniciador, que flanqueia os locais de restrição que podem ser utilizados para a introdução de um segmento de DNA para formar um vetor de montagem. Como usado neste, um "vetor de montagem" refere-se a um vetor de entrada ao qual um segmento de DNA foi introduzido. Um vetor de montagem pode ser usado nos métodos de montagem de polinucleotídeo descritos neste para fornecer um polinucleotídeo componente a ser montado em um polinucleotídeo montado.

[0082] Como usado neste, o termo "vetor de montagem" refere-se a um vetor que compreende uma sequência de ligante anelável, duas sequências de ligante anelável ou uma sequência de ligante anelável e um segmento de ligação de iniciador, e um segmento de DNA.

[0083] Como usado neste, o termo "enzima de restrição" ou "endonuclease de restrição" refere-se a um membro ou membros de uma classificação de moléculas catalíticas que ligam uma sequência cognada de DNA e cliva a molécula de DNA em um local preciso dentro daquela sequência. As endonucleases de restrição incluem endonucleases de restrição Tipo IIS. Esta classe de enzimas difere de outras endonucleases de restrição naquela sequência de reconhecimento é separada do local de clivagem. Alguns exemplos de enzimas de restrição Tipo IIS incluem AlwI, BsaI, BbsI, BbuI, BsmAI, BsrI, BsmI, BspMI, EarI, Esp3I, FokI, HgaI, HphI, IguI, MboII, MnII, PfiI, SapI, SchI, SfaNI e outros. Muitas destas endonucleases de restrição estão comercialmente disponíveis e são bem conhecidas por aqueles

habilitados na técnica.

[0084] Como usado neste, o termo "sequência de ligante anelável duplex" refere-se a um filamento de sequência de ligante anelável alinhada com um filamento de sequência de ligante anelável substancialmente complementar em associação antiparalela. A complementaridade não necessita ser perfeita; os duplexes de sequência de ligante anelável podem conter pares de base mal comparados ou bases não comparadas, embora nas formas de realização particulares, o duplex de sequência de ligante anelável compreende dois filamentos de sequências de ligante aneláveis tendo complementaridade perfeita.

[0085] Como usado neste, o termo "sequência alvejadora genômica" refere-se a uma sequência de nucleotídeo que está presente no genoma de uma célula hospedeira em um local em que um polinucleotídeo da invenção deve ser inserido por uma recombinação homóloga mediada por célula hospedeira. Os termos "sequência alvejadora genômica a montante" e "sequência alvejadora genômica a jusante" refere-se às sequências alvejadoras genômicas que estão localizadas a montante e à jusante um do outro no genoma de uma célula hospedeira.

[0086] Como usado neste, o termo "sequência de alvejamento cromossômico" refere-se a uma sequência de nucleotídeo que está presente em um cromossomo de uma célula hospedeira em um local em que um polinucleotídeo da invenção deve ser inserido por uma recombinação homóloga mediada por célula hospedeira. Os termos "sequência de alvejamento cromossômico a montante" e "sequência de alvejamento cromossômico a jusante" refere-se às sequências de alvejamento cromossômico que estão localizados a montante e a jusante um do outro em um cromossomo de uma célula hospedeira.

## 5.2 Métodos de montagem de polinucleotídeo

[0087] Em um aspecto, a presente invenção fornece métodos rápidos,

robustos e de rendimento alto para a montagem ordenada de uma pluralidade de polinucleotídeos componentes em um ou mais polinucleotídeos montados. Os métodos da invenção utilizam vetores de ácido nucleico circular, denominados vetores de montagem, cada um compreendendo um segmento de DNA, D, flanqueado por uma sequência ligante anelável (isto é, LA ou LB), um par de sequências de ligante anelável (isto é, LA e LB) ou uma sequência de ligante anelável e um segmento de ligação de iniciador (isto é, LA e PB ou LB e PA) e um par dos locais de restrição, RA e RB (FIG. 1B). A digestão de endonuclease de restrição de uma pluralidade de vetores de montagem nos locais de restrição RA e RB gera uma pluralidade de polinucleotídeos componentes que compreendem os elementos 5'-LAD-3', 5'-D-LB-3', 5'-LA-D-LB-3', 5'-LA-D-PB-3' ou 5'-LB-D-PA-3' (FIG. 3). Nos métodos da invenção, as sequências de ligante anelável LA e LB fornecem os polinucleotídeos componentes com terminais complementares que são utilizados em uma reação de montagem de extensão de sobreposição de união seguido por reação de cadeia de polimerase (SOE/PCR) para montar os polinucleotídeos componentes em um polinucleotídeo montado com uma sequência ordenada.

[0088] Em particular, os métodos podem fornecer a montagem em um polinucleotídeo montado simples de diversos elementos de DNA funcionais, que incluem mas não limitam-se a sequências codificadoras de proteína, genes repórteres, sequências codificadoras de marcador fluorescente, promotores, intensificadores, terminadores, íntrons, éxons, caudas poly-A, locais de clonagem múltiplos, sinais de localização nucleares, sinais de estabilização de mRNA, marcadores selecionáveis, locais de integração, sequências codificadoras de etiqueta de epítipo e sinais de degradação. Os métodos podem ser usados para a montagem e qualquer tipo de polinucleotídeo montado, que inclui mas não limita-se a genes sintéticos, construções, vetores de clonagem, vetores de expressão, cromossomos,

construções de integração genômica, genomas e bibliotecas de DNA. Além disso, os métodos podem ser usados para a montagem de segmentos de DNA em uma reação simples sem a necessidade de manipulação e caracterização de produtos intermediários.

[0089] Em algumas formas de realização, os métodos também podem fornecer a montagem de um polinucleotídeo montado a partir de uma pluralidade de polinucleotídeos componentes que não se originam de um vetor de montagem (isto é, segmentos de DNA obtidos por procedimentos padrão conhecidos na técnica, tais como, por exemplo, amplificação de PCR, síntese química e outros, que são flanqueados por uma ou duas sequências de ligante anelável, LA e/ou LB ou por uma sequência de ligante anelável e um segmento de ligação de iniciador (isto é, LA e PB ou LB e PA). Os polinucleotídeos componentes que não originam-se de um vetor de montagem podem ser adicionados a uma reação de montagem em qualquer estágio anterior à reação de SOE/PCR ou uma recombinação homóloga mediada por célula hospedeira para a montagem no polinucleotídeo montado. Desta maneira, em algumas formas de realização, os métodos de montagem podem ser usados para montar: (1) polinucleotídeos componentes derivados de vetores de montagem que compreendem uma ou duas sequências de ligante anelável ou uma sequência de ligante anelável e um segmento de ligação de iniciador e gerado pela digestão dos vetores de montagem; (2) fragmentos de DNA flanqueados sem vetor por uma ou duas sequências de ligante anelável ou por uma sequência de ligante anelável e um segmento de ligação de iniciador e (3) combinações destes.

[0090] Em algumas formas de realização, são fornecidos aqui métodos de montar uma pluralidade de polinucleotídeos componentes em um ou mais polinucleotídeos montados, que compreende as etapas de:

(a) digerir uma composição de montagem com uma ou mais endonucleases de restrição para gerar uma composição de componentes, a

composição de montagem compreendendo:

(i) uma ou mais primeiras moléculas de ácidos nucleicos, em que cada primeira molécula de ácido nucleico é circular e compreende, em uma orientação de 5' a 3', um primeiro local de restrição  $RA_0$ , qualquer segmento de ligação de iniciador selecionado do grupo PA, qualquer segmento de DNA selecionado do grupo  $D_0$ , uma sequência de ligante anelável  $LB_0$ , e um segundo local de restrição  $RB_0$ ;

(ii) uma ou mais moléculas de ácido nucleico intermediárias em que cada molécula de ácido nucleico intermediária  $n$  é circular e compreende, em uma orientação de 5' a 3', um primeiro local de restrição  $RA_n$  uma primeira sequência de ligante anelável  $LA_n$  qualquer segmento de DNA selecionado do grupo  $D_n$ , uma segunda sequência de ligante anelável  $LB_n$  e um segundo local de restrição  $RB_n$  e em que  $n$  representa um número inteiro de uma das diversas moléculas de ácido nucleico intermediárias e

(iii) uma ou mais das últimas moléculas de ácido nucleico, em que cada última molécula de aminoácido é circular e compreende, em uma orientação de 5' a 3', um primeiro local de restrição  $RA_m$ , uma sequência de ligante anelável  $LA_m$ , um segmento de DNA selecionado do grupo  $D_m$ , qualquer segmento de ligação de iniciador selecionado do grupo PB, um segundo local de restrição  $RB_m$  em que  $m$  representa um número inteiro, um maior do que o número de moléculas de ácido nucleico intermediárias; após o que a clivagem dos locais de restrição  $RA_0$  até  $RB_m$  e a desnaturação das moléculas de ácido nucleico lineares resultantes, cada sequência de ligante anelável  $LB_{(p-1)}$  é capaz de hibridizar ao complemento de sequência de ligante anelável  $LAp$ , em que  $n$  é um número inteiro que varia de 1 a  $(m-1)$ , em que  $p$  representa um número inteiro de 1 a  $m$  e em que cada grupo  $D_0, \dots D_n, \dots$  e  $D_m$  consiste de um ou mais segmentos de DNA;

em que a uma ou mais endonucleases de restrição são capazes de clivar os locais de restrição  $RA_0$  até  $RB_m$  e



(b) contactar a composição de componentes com DNA polimerase, trifosfatos de desoxirribonucleosídeo e um ou mais primeiro iniciadores e um ou mais segundos iniciadores, sob condições adequadas para a desnaturação da moléculas de ácido nucleico, anelamento da sequência de ligante anelável  $LB_{(p-1)}$  à sequência de ligante anelável LAp e extensão deste; em que cada dito primeiro iniciador é capaz de hibridizar a um dos ditos segmentos de ligação de iniciador selecionado do grupo PA e cada dito segundo iniciador é capaz de hibridizar a um dos ditos segmentos de ligação de iniciador selecionado do grupo PB e submeter a composição de componentes à reação de cadeia de polimerase,

em que um polinucleotídeo é montado compreendendo, em uma orientação de 5' a 3', um segmento de DNA selecionado de cada um dos grupos  $D_0, \dots, D_n, \dots$  e  $D_m$ . No método, p representa os inteiros de 1 a m.

[0091] A FIG. 3 descreve uma forma de realização dos métodos de montagem da invenção para os propósitos ilustrativos. Neste exemplo, um total de quatro polinucleotídeos componentes são mostrados para produzir um polinucleotídeo montado. Entretanto, os métodos de montagem fornecidos neste podem ser usados para montar qualquer número de polinucleotídeos componentes em um ou mais polinucleotídeos montados. Em algumas formas de realização, os métodos fornecidos aqui resultam na montagem de 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, ou mais polinucleotídeos componentes em um ou mais polinucleotídeos montados.

[0092] No exemplo ilustrado em FIG. 3, a composição de montagem a partir da qual o polinucleotídeo montado é gerado compreende quatro vetores de montagem de entrada, indicado "primeiro", "intermediário 1 ( $int_1$ )," "intermediário 2 ( $int_2$ )," e "último". Cada vetor de montagem compreende um segmento de DNA flanqueado por uma sequência de ligante anelável e um segmento de ligação de iniciador ou por duas sequências de ligante anelável. Especificamente, o segmento de DNA  $D_0$  é flanqueado por 5'

segmento de ligação de iniciador PA e 3' sequência de ligante anelável LB<sub>0</sub>. O segmento de DNA D<sub>1</sub> é flanqueado por 5' e 3' sequências de ligante anelável LA<sub>1</sub> e LB<sub>1</sub> e segmento de DNA D<sub>2</sub> é flanqueado por 5' e 3' sequências de ligante anelável LA<sub>2</sub> e LB<sub>2</sub>. O segmento de DNA D<sub>3</sub> é flanqueado por 3' segmento de ligação de iniciador PB e 5' sequência de ligante anelável LA<sub>3</sub>. Os elementos 5'-PA-D-LB-3', 5'-LA-D-LB-3' ou 5'-LA-D-PB-3' nos vetores de montagem ainda são flanqueados por locais de endonuclease de restrição SapI.

[0093] Na primeira etapa da reação de montagem mostrada na FIG. 3, a composição de montagem é digerida com SapI, resultando no corte de polinucleotídeos componentes, que compreendem os elementos 5'-PA-D-LB-3', 5'-LA-D-LB-3' ou 5'-LA-D-PB-3', do vetor de estruturas de montagem em uma composição de componente. Por causa do Sap I ser uma endonuclease de restrição Tipo IIS, seu local de reconhecimento é distal ao seu local de clivagem e a clivagem ocorre fora de sua sequência de reconhecimento. Esta propriedade torna as endonucleases de restrição Tipo IIS particularmente úteis no montagem de um polinucleotídeo de acordo com os métodos fornecidos aqui, visto que os polinucleotídeos podem ser montados não compreendendo uma cicatriz no local de restrição, que pode, de outra maneira resultar da clivagem dos locais de restrição RA e RB com uma endonuclease de restrição que não Tipo IIS. Referindo-se à Figura 2, o local de reconhecimento Tipo IIS é 5' do local de clivagem correspondente para cada um de RA<sub>0</sub>, RA<sub>n</sub> e RA<sub>m</sub> e 3' de seu local de clivagem RB<sub>0</sub>, RA<sub>n</sub> e RA<sub>m</sub>. desta maneira, os locais de restrição RA<sub>0</sub> até RB<sub>m</sub> são orientados de modo que a clivagem por uma ou mais endonucleases de restrição Tipo IIS capazes de clivar RA<sub>0</sub> até RB<sub>m</sub> resultam na separação de RA<sub>0</sub> de D<sub>0</sub>, LB<sub>0</sub> de RB<sub>0</sub>, RA<sub>n</sub>, de LA<sub>n</sub>, LB<sub>n</sub> de RB<sub>n</sub>, de LA<sub>n</sub>, e D<sub>m</sub> de RB<sub>m</sub>, em que as moléculas de ácido nucleico linearizadas resultantes compreendem D<sub>0</sub>, LB<sub>0</sub>, RA<sub>n</sub>, LB<sub>n</sub>, LA<sub>m</sub> ou D<sub>m</sub> não compreende qualquer um de RA<sub>0</sub> até RB<sub>m</sub> como uma consequência, os polinucleotídeos

componentes não incluem qualquer traço de locais de reconhecimento ou clivagem de enzimas de restrição. Como um resultado, os métodos inventivos de montagem de polinucleotídeo podem ser usados para a transformação de células hospedeiras múltiplas vezes sem a introdução das repetições de sequência que possam causar a instabilidade genética.

[0094] Subsequentemente, a endonuclease de restrição é opcionalmente inativada. Se a inativação for desejada, qualquer método conhecido na técnica para a inativação da atividade da enzima endonuclease pode ser utilizado, incluindo métodos de purificação com base em coluna ou gel. Um método conveniente é a inativação por calor, por exemplo, a 65°C por 20 minutos, que requer pouca ou nenhuma manipulação de uma composição de componentes fora do tubo de reação.

[0095] A montagem dos polinucleotídeos componentes em um polinucleotídeo montado é permitida pelos duplexes de sequência formados pela sobreposição de filamentos de terminais complementares entre os polinucleotídeos componentes. Especificamente, as sequências de ligante anelável são projetados tal que a sequência de ligante anelável LB<sub>0</sub> pode hibridizar ao complemento da sequência de ligante anelável LA<sub>1</sub>, sequência de ligante anelável LB<sub>1</sub> pode hibridizar ao complemento da sequência de ligante anelável LA<sub>2</sub> e sequência de ligante anelável LB<sub>2</sub> pode hibridizar ao complemento da sequência de ligante anelável LA<sub>3</sub>. Desta maneira, em uma segunda etapa da reação de montagem, os polinucleotídeos componentes são submetidos às condições de desnaturação (por exemplo, calor) para gerar os componentes de polinucleotídeos de filamento simples, que concomitante com ou subsequente à etapa de desnaturação da reação de montagem são contactados com uma DNA polimerase termoestável e trifosfatos de desoxirribonucleosídeo.

[0096] A DNA polimerase termoestável pode ser qualquer DNA polimerase termoestável considerada adequada por aqueles habilitados na

técnica. As DNA polimerases termoestáveis adequadas para o uso nos presente métodos incluem mas não limitam-se a DNA polimerase de *Thermus thermophilus* (Tth), DNA polimerase de *Thermus aquaticus* (Taq) DNA polimerase, DNA polimerase de *Thermotoga neopolitana* (Tne), DNA polimerase de *Thermotoga maritima* (Tma), DNA polimerase de *Thermococcus litoralis* (Tli or VENT<sup>TM</sup>) DNA polimerase, *Pyrococcus furiosus* (Pfu or DEEPVENT<sup>TM</sup>) DNA polimerase, *Pyrococcus woosii* (Pwo) DNA polimerase, *Bacillus sterothermophilus* (Bst) DNA polimerase, *Sulfolobus acidocaldarius* (SAC) DNA polimerase, *Thermoplasma acidophilum* (Tac) DNA polimerase, *Thermus flavus* (Tfl/Tub) DNA polimerase, *Thermus Tuber* (Tru) DNA polimerase, *Thermus brockianus* (DYNAZYME<sup>TM</sup>) DNA polimerase, *Methanobacterium thermoautotrophicum* (Mth) DNA polimerase, and mutants, variants, and derivatives thereof. Thermostable DNA polimerases having high fidelity (isto é, proofreading properties) and low error rates are preferred. Em certas formas de realização, a DNA polimerase é Phusion<sup>TM</sup> DNA Polimerase (New England Biolabs, Ipswich, MA). Em outras formas de realização, a DNA Polimerase é PfuUltra<sup>TM</sup> II Fusion DNA Polimerase (Stratagene/Agilent, La Jolla, CA).

[0097] A reação de montagem é então submetida às condições que permitem o alongamento de filamento das porções de 3'-hidroxila adas sequências de ligante anelável de sobreposição, durante o que a DNA polimerase termoestável acumula-se na porção entre as sequências de ligante anelável de sobreposição. A reação de montagem é submetida a um número limitado de ciclos de repetição de desnaturação/anelamento/extensão (por exemplo, por 5 a 15 ciclos) durante os quais uma quantidade substancial de polinucleotídeo de filamento duplo montado é formada. Durante este ciclo, os polinucleotídeos componentes atuam tanto como iniciadores quanto padrão para gerar um padrão de comprimento total para o polinucleotídeo montado.

Em certas formas de realização, as etapas de anelamento e extensão do PCR podem ser realizadas a 72°C.

[0098] Em contraste com as sequências de ligante anelável LA e LB, o segmento de ligação de iniciadores PA e PB são projetados para não sobrepor um com o outro ou qualquer uma das sequências de ligante anelável ou segmentos de DNA, mas em vez disso servem como locais de ligação para os iniciadores usados para amplificar o polinucleotídeo montado de comprimento total. Desta maneira, nas etapas 4 e 5 da reação de montagem, iniciadores complementares a segmento de ligação de iniciadores PA e PB são adicionados e a composição é submetida à amplificação de condições tradicionais de PCR. As condições de amplificação de PCR podem ser quaisquer condições de amplificação de PCR consideradas adequadas por aqueles habilitados na técnica, incluindo aqueles descritos em *PCR Technology: Principles and Applications for DNA Amplification*, ed. HA Erlich, Stockton Press, New York, N.Y. (1989); *PCR Protocols: A Guide to Methods and Applications*, eds. Innis, Gelfland, Snisky, and White, Academic Press, San Diego, Calif. (1990); Mattila *et al.* (1991) *Nucleic Acids Res.* 19: 4967; Eckert, K. A. and Kunkel, T. A. (1991) *PCR Methods and Applications* 1: 17; e Patente U.S. N° 4,683,202 e 4,965,188, cada um dos quais são incorporados neste pode referência. Em certas formas de realização, a etapa de PCR da reação de montagem compreende cerca de 35 ciclos de desnaturação, anelamento e extensão na presença de iniciadores complementares ao segmento de ligação de iniciadores PA e PB. Em certas formas de realização, as etapas de anelamento e de extensão do PCR podem ser ambos realizados a 72°C. Entretanto, uma pessoa habilitada na técnica entenderá que as condições ótimas para a amplificação bem sucedida dependerá da DNA polimerase termoestável e as sequências de ligante anelável utilizadas e estas condições podem ser ajustadas consequentemente.

[0099] Opcionalmente, o polinucleotídeo montado pode ser purificado

por qualquer técnica evidente para uma pessoa habilitada na técnica, por exemplo, métodos de purificação de eletroforese em gel e usados para uma variedade de propósitos. Por exemplo, o polinucleotídeo montado pode ser inserido em uma estrutura de vetor de expressão em uma estrutura de vetor de expressão para a verificação de sequência.

### 5.3 Métodos de Gerar Células Hospedeiras Compreendendo Polinucleotídeos Montados

[00100] Em um outro aspecto, a presente invenção fornece métodos para a geração de células hospedeiras que compreende polinucleotídeo montados. Em algumas formas de realização, o polinucleotídeo montado tem pelo menos 3 kb de tamanho. Em outras formas de realização, o polinucleotídeo montado tem pelo menos 5 kb de tamanho. Ainda, em outras formas de realização, o polinucleotídeo montado tem pelo menos 6, 7, 8, 9 ou 10 kb de tamanho. Ainda, em outras formas de realização, o polinucleotídeo montado é maior do que 10 kb de tamanho. Ainda, em outras formas de realização, o polinucleotídeo montado é maior do que 15 kb de tamanho. Ainda, em outras formas de realização, o polinucleotídeo montado é maior do que 20 kb de tamanho.

[00101] Em algumas formas de realização, os métodos são fornecidos compreendendo transformar uma célula hospedeira com um polinucleotídeo montado gerado pelos métodos de montagem de polinucleotídeo descrito neste. O polinucleotídeo montado pode ser circularizado antes da transformação ou pode ser transformado como uma molécula linear. O polinucleotídeo montado pode ser mantido em uma célula hospedeira como um polinucleotídeo extracromossomal. Alternativamente, o polinucleotídeo montado pode ser integrado no genoma de uma célula hospedeira, por exemplo, por uma recombinação homóloga mediada por célula hospedeira. Para a integração de um polinucleotídeo montado no genoma pela recombinação homóloga, o polinucleotídeo montado deve compreender, em

um terminal uma sequência de ácido nucleico que compreende uma sequência alvejadora genômica a montante e no outro terminal uma sequência de ácido nucleico que compreende a sequência alvejadora genômica a jusante. Consequentemente, um polinucleotídeo montado que deve ser integrado em um cromossomo de uma célula hospedeira é gerada a partir de uma composição de montagem que compreende uma primeira molécula de ácido nucleico que compreende uma sequência de alvejamento cromossômico a montante e a última molécula de aminoácido que compreende a sequência de alvejamento cromossômico a jusante, cada sequência de alvejamento cromossômico sendo de comprimento suficiente para iniciar a recombinação homóloga por uma célula hospedeira com seu cromossomo.

[00102] Em outras formas de realização, os métodos compreendem transformar uma célula hospedeira com uma pluralidade de polinucleotídeo montado gerado pelos métodos de montagem de polinucleotídeo descrito neste. Em uma forma de realização particular, a célula hospedeira dois ou mais polinucleotídeos montados em um polinucleotídeo combinado simples pela recombinação homóloga. Os transformantes de célula hospedeira que compreendem os polinucleotídeos combinados são selecionados e virtude da expressão de um marcador selecionável que é gerado no processo de combinar os polinucleotídeos montados. O método é particularmente útil para inserir pedaços relativamente grandes de polinucleotídeo em um polinucleotídeo alvo pela recombinação homóloga. Para a integração cromossômica ocorrer, o polinucleotídeo combinado deve compreender uma sequência alvejadora genômica a montante localizada 5' ou 3' da sequência codificadora do marcador selecionável e a sequência alvejadora genômica a jusante localizada 3' ou 5' da sequência codificadora do marcador selecionável, respectivamente. A integração genômica como usado neste inclui a integração cromossomal, isto é, integração de um polinucleotídeo em um cromossomo de uma célula hospedeira. Os locais de integração

cromossomal em *Saccharomyces cerevisiae* incluem mas não limitam-se ao local *NDT80*, *HO*, *GAL2* e *GAL1-GAL10-GAL7*. O método também pode ser útil para a geração de células hospedeiras que compreende um polinucleotídeo extracromossomicamente mantido, por exemplo, os vetores e plasmídeos de expressão. A estabilidade de um polinucleotídeo cromossomicamente integrado ou um extracromossomicamente mantido é aumentada quando o polinucleotídeo combinado não compreende sequências de ligante anelável idênticas ou segmentos de DNA dispostos como repetições diretas que podem, de outra maneira, iniciar os eventos de recombinação homóloga adicionais resultando no corte de segmentos do polinucleotídeo componente. Portanto, em algumas formas de realização, o polinucleotídeo montado compreende sequências de ligante anelável e segmentos de DNA únicos. Em outras formas de realização, o polinucleotídeo montado contém um ou mais sequências de ligante anelável ou segmentos de DNA idênticos que, na combinação do polinucleotídeo montado, são dispostos como repetições invertidas no polinucleotídeo combinado.

[00103] A geração de um polinucleotídeo combinado exemplar e integração do polinucleotídeo combinado em um cromossomo da célula hospedeira pela recombinação homóloga é ilustrada na FIG. 8. Dois polinucleotídeos montados A e B são absorvidos por uma célula hospedeira que é capaz de recombinação homóloga. Cada polinucleotídeo montado compreende um segmento de DNA  $D_{m1}$  que codifica um segmento de um marcador selecionável, em que o segmento de DNA do polinucleotídeo montado A codifica um primeiro segmento de um marcador selecionável e o segmento de DNA  $D_{m2}$  de polinucleotídeo montado B codifica um segundo segmento do marcador selecionável, em que o segmento de DNA e o segmento de DNA  $D_{m2}$  compreende uma região de homologia suficiente para iniciar a recombinação homóloga mediada por célula hospedeira e em que nem o segmento de DNA  $D_{m1}$  nem o segmento de DNA  $D_{m2}$  produz um



marcador selecionável funcional mas, visto que a recombinação homóloga por uma célula hospedeira a marcador selecionável funcional é gerada. Cada polinucleotídeo montado compreende adicionalmente um segmento de DNA  $D_0$  que codifica a sequência de alvejamento cromossômico de comprimento suficiente para iniciar a recombinação homóloga mediada por hospedeiro, em que segmento de DNA  $D_{01}$  de polinucleotídeo montado A codifica uma sequência de alvejamento cromossômico a montante e o segmento de DNA  $D_{02}$  de polinucleotídeo montado B codifica uma sequência de alvejamento cromossômico a jusante. Uma vez dentro da célula, a célula hospedeira recombina o polinucleotídeo montado A e o polinucleotídeo montado B na região de homologia em segmentos de DNA  $D_{m1}$  e  $D_{m2}$  para formar um polinucleotídeo combinado. Além disso, a célula hospedeira usa as sequências de alvejamento cromossômico codificados por segmentos de DNA  $D_{01}$  e  $D_{02}$  para a inserção do polinucleotídeo combinado pela recombinação homóloga em seu cromossomo. As células hospedeiras que compreendem o polinucleotídeo combinado podem ser facilmente identificadas com base no marcador selecionável funcional gerado.

[00104] Ainda, em outras formas de realização, os métodos compreendem transformar uma célula hospedeira com uma pluralidade de polinucleotídeos componentes e permitindo que a célula hospedeira gere um ou mais polinucleotídeo montado pela recombinação homóloga. O polinucleotídeo montado pode ser extracromossomicamente mantido em uma célula hospedeira ou integrado no cromossomo da célula hospedeira. A geração de um polinucleotídeo exemplar montado pela recombinação homóloga em uma célula hospedeira e integração do polinucleotídeo montado no cromossomo da célula é ilustrada na FIG. 9. Na primeira etapa, uma composição de montagem que compreende vetores de montagem é digerida com a endonuclease de restrição Tipo IIS tal como SapI ou LguI, resultando na excisão do vetor de estruturas de montagem de polinucleotídeos

componentes. Nesta forma de realização, D<sub>0</sub> e D<sub>3</sub> podem ser a sequência de alveamento cromossômico amontante e a jusante, caso no qual a presença de um segmento de ligação de iniciador no primeiro e no último vetor de montagem é opcional. Alternativamente, os dois segmentos de ligação de iniciador podem funcionar como as sequências alvejadoras genômicas a montante e a jusante.

[00105] Uma vez cortado, cada polinucleotídeo componente cortado compreende uma sequência de ligante anelável LB que é homólogo a uma sequência de ligante anelável LA de um outro polinucleotídeo componente e que é de comprimento suficiente para iniciar a recombinação homóloga mediada por hospedeiro. O polinucleotídeo componente cortado do primeiro vetor de montagem compreende adicionalmente umn sequência de alveamento cromossômico a montante e o polinucleotídeo componente cortado do último vetor de montagem compreende adicionalmente uma sequência de alveamento cromossômico a jusante, em que tanto a sequência de alveamento cromossômica são de comprimento suficiente para iniciar a recombinação homóloga mediada por hospedeiro com um cromossomo de uma célula hospedeira. A endonuclease de restrição pode ser subsequentemente inativada. Em uma segunda etapa do método, as composições componentes são introduzidas em uma célula hospedeira capaz de recombinação homóloga. Uma vez dentro da célula, a célula hospedeira recombina os polinucleotídeos componentes nas regiões de homologia entre as sequências de ligante anelável para a formação de um polinucleotídeo montado e o polinucleotídeo montado é integrado no cromossomo. As células hospedeiras que compreendem o polinucleotídeo montado podem ser facilmente identificadas com base em um marcador selecionável codificado por um segmento de DNA do polinucleotídeo montado.

[00106] Qualquer célula hospedeira pode ser usada nos métodos descritos neste. Nas formas de realização particulares, as células hospedeiras

adequadas são células hospedeiras que são capazes de recombinar polinucleotídeos com base nas tensões de sequência complementar tal como fornecido pelos segmentos de marcador selecionáveis, sequências alvejadoras genômicas e sequências de ligante anelável fornecidos neste. Os exemplos ilustrativos de tais células hospedeiras incluem mas não limitam-se a *Saccharomyces cerevisiae*. As condições adequadas para a entrada de DNA por tais células hospedeiras são bem conhecidas na técnica.

[00107] Os transformantes de célula que compreendem um polinucleotídeo montado ou combinado podem ser facilmente identificados em virtude de expressar um marcador selecionável codificado pelo polinucleotídeo montado ou pelo polinucleotídeo combinado que permite a seleção para ou contra o desenvolvimento das células. O marcador selecionável pode ser codificado por um segmento de DNA simples presente em um vetor de montagem de uma composição de montagem. Alternativamente, segmentos não funcionais do marcador selecionável podem ser codificados pelos segmentos de DNA presentes em vetores de montagem múltiplos de uma composição de montagem ou em polinucleotídeos múltiplos montados tal que um marcador selecionável funcional é gerado apenas na geração de um polinucleotídeo montado ou na geração de um polinucleotídeo combinado, respectivamente.

[00108] Uma ampla variedade de marcadores selecionáveis é conhecida na técnica (ver, por exemplo, Kaufman, *Meth. Enzymol.*, 185:487 (1990); Kaufman, *Meth. Enzymol.*, 185:537 (1990); Srivastava and Schlessinger, *Gene*, 103:53 (1991); Romanos *et al.*, in *DNA Cloning 2: Expression Systems*, 2ª Edição, páginas 123 a 167 (IRL Press 1995); Markie, *Methods Mol. Biol.*, 54:359 (1996); Pfeifer *et al.*, *Gene*, 188:183 (1997); Tucker and Burke, *Gene*, 199:25 (1997); Hashida-Okado *et al.*, *FEBS Letters*, 425:117 (1998)). Em algumas formas de realização, o marcador selecionável é um marcador resistente a medicamento. Um marcador resistente a

medicamento permite que as células realizem a desintoxicação de um medicamento exógeno que, de outra maneira, deve matar a célula. Os exemplos ilustrativos de marcadores resistentes a medicamento incluem mas não limitam-se àquelas que conferem resistência a antibióticos, tais como ampicilina, tetraciclina, canamicina, bleomicina, estreptomicina, higromicina, neomicina, Zeocin<sup>TM</sup>, e outros. Em outras formas de realização, o marcador selecionável é um marcador auxotrófico. Um marcador auxotrófico permite que as células sintetizem um componente essencial (usualmente um aminoácido) enquanto desenvolvem-se no meio que precisa daquele componente essencial. As sequências de gene auxotrófico selecionáveis incluem, por exemplo, hisD, que permite o desenvolvimento em meio isento de histidina na presença de histidinol. Outros marcadores selecionáveis incluem um gene resistente à bleomicina, um gene de metalotioneína, um gene de higromicina B-fosfotransferase, o gene AURI, um gene de adenosina deaminase, um gene de aminoglicosídeo fosfotransferase, um gene de diidrofolato redutase, um gene de timidina cinase, um gene de xantina-guanina fosforibosiltransferase e outros.

[00109] A auxotrofia também pode ser usada para identificar os transformantes de célula hospedeira que compreendem um polinucleotídeo montado ou combinado cromossomicamente integrados quando a integração do polinucleotídeo montado ou combinado resulta na interrupção de um gene que a célula hospedeira requer para sintetizar um componente essencial para o desenvolvimento celular, desta maneira, tornando a célula auxotrófica.

[00110] Os transformantes de célula hospedeira que compreendem um polinucleotídeo montado ou combinado cromossomicamente integrado também pode ser identificado pela seleção dos transformantes de célula hospedeira que apresentam outras características codificadas por segmentos individuais de DNA ou por combinações dos segmentos de DNA, por exemplo, expressão de peptídeos em emitem luz ou pela análise

molecular de colônias de célula hospedeira individuais, por exemplo, pelo mapeamento de enzima de restrição, amplificação de PCR ou análise de sequência de polinucleotídeo montado isolado ou locais e integração cromossômica.

#### 5.4. Métodos combinatórios do montagem de polinucleotídeo e célula hospedeira

##### Geração

[00111] Em outro aspecto, a presente invenção fornece métodos rápidos, robustos e rendimento alto para a montagem ordenada dos polinucleotídeos de componentes múltiplos em uma pluralidade de polinucleotídeos montados. Os métodos confiam no uso de uma composição de montagem que compreende vetores de montagem que cada um compreende um segmento de DNA D, flanqueado por uma sequência LA ou LB ligante anelada, um par das sequências LA ou LB ligantes aneladas ou por uma sequência ligante anelada e um segmento de ligação iniciador, isto é, LA e PB ou LB e PA, flanqueado por um par dos locais de restrição RA e RB (FIG. 1 B). Entretanto, para gerar uma diversidade de polinucleotídeos montados usando os métodos divulgados neste, as sequências ligantes aneladas e segmentos de ligação iniciadores são escolhidos tal que mais do que uma combinação dos polinucleotídeos componentes podem ser montados em um polinucleotídeo montado na reação. Deste modo, em algumas formas de realização, a composição de montagem compreende pelo menos dois vetores de montagem que tem a mesma sequência LA ou LB ligante anelada ou o mesmo segmento de ligação iniciador PA ou PB, mas difere-se com relação ao segmento DNA. Em outras formas de realização, a composição de montagem compreende pelo menos dois vetores de montagem que tem o mesmo par das sequências LA ou LB ligantes aneladas ou a mesma sequência ligante anelada e par de segmento de ligação iniciador, isto é, LA e PB ou LB e PA. mas difere-se com relação ao segmento DNA.

[00112] FIG. 10 apresenta um método exemplar de geração de uma pluralidade de polinucleotídeos montados a partir de sete (7) polinucleotídeos componentes na mesma reação. Os vetores de montagem que compreendem os segmentos de DNA a serem montados são montados em um tubo simples e digeridos com SapI para liberar os fragmentos de polinucleotídeos componentes a partir das estruturas dos vetores de montagem. Seguindo a inativação de calor de SapI, os polinucleotídeos componentes são submetidos as condições de desnaturação, seguido pelas condições de anelamento suficientes para a hibridização dos pares ligantes anelados complementares. Seguindo a extensão iniciadora na presença de polimerase de DNA e dNTPs, iniciadores complementares aos segmentos de ligação iniciadores PA e PB são adicionados ao polinucleotídeo de comprimento total diferente oito (8) de amplificação de PCR montados que compreendem os segmentos de DNA  $D_{01/02}$ ,  $D_{1/2}$ ,  $D_3$  e  $D_{41/42}$  montados em várias combinações possíveis. Polinucleotídeos montados individuais podem ser isolados a partir da composição dos polinucleotídeos montados misturados, por exemplo, por outro ciclo de amplificação de PCR usando iniciadores complementares às regiões de segmentos de DNA  $D_{01}$ ,  $D_{02}$ ,  $D_{41}$  e  $D_{42}$ . Alternativamente, uma série de polinucleotídeos montados podem ser isolados pelo primeiro e últimos vetores de montagem que compreendem um de um grupo de segmentos de ligação iniciadores PA e/ou PB e usando iniciadores para amplificação de PCR que hibridizam para selecionar apenas um grupo de segmentos de ligação iniciadores PA e PB. Os polinucleotídeos montados isolados podem ser usados, por exemplo, para transformar as células hospedeiras para gerar uma pluralidade de células hospedeiras que compreendem polinucleotídeos montados. Alternativamente, as células hospedeiras podem ser diretamente transformada com a composição dos polinucleotídeos montados misturados e transformantes da célula hospedeira que compreendem cada

polinucleotídeo montado podem ser isolados, por exemplo, pela análise molecular das colônias de célula hospedeira individual ou pela seleção de transformantes da célula hospedeira que compreendem marcadores selecionáveis ou que exibem outros traços codificados pelos segmentos de DNA individuais ou pelas combinações de segmentos de DNA.

[00113] Em outras formas de realização, uma pluralidade de células hospedeiras que compreendem uma pluralidade de polinucleotídeos montados pelos métodos combinatoriais são gerados pela transformação das células hospedeiras com uma composição que compreende polinucleotídeos montados múltiplos de que pelo menos dois polinucleotídeos montados compreendem segmentos não funcionais de um marcador selecionável que na recombinação homóloga gera um marcador selecionável funcional e pela seleção das células hospedeiras que compreendem um polinucleotídeo combinado. A FIG. 11 ilustra um método combinatorial para gerar uma pluralidade das células hospedeiras que compreendem os polinucleotídeos combinados. No exemplo, polinucleotídeos montados A1 e A2, cada um que compreende a mesma sequência de alvejamento cromossomal a montante e a mesma primeira porção de um marcador selecionável e polinucleotídeos montados B1 e B2, cada um que compreende a mesma sequência de alvejamento cromossomal a jusante e a mesma segunda porção de um marcador selecionável, são combinatoriamente combinadas pela recombinação homóloga mediada pela célula hospedeira para gerar quatro polinucleotídeos combinados diferentes, A1/B1, A1/B2, A2/B1 e A2/B2, que podem ser inseridos em um cromossomo para gerar quatro células hospedeiras diferentes.

[00114] Já em outras formas de realização, uma pluralidade de células hospedeiras que compreendem uma pluralidade de polinucleotídeos montados e combinados pelos métodos combinatoriais são gerados pela transformação das células hospedeiras com uma composição componente

que compreende polinucleotídeos de componentes múltiplos de que pelo menos dois polinucleotídeos componentes compreendem segmentos não funcionais de um marcador selecionável que na recombinação homóloga mediada pela célula hospedeira gera um marcador selecionável funcional e pela seleção de células hospedeiras que compreendem um polinucleotídeo combinado ou montado.

### 5.5. Vetores de entrada

[00115] Em outro aspecto, fornecido neste é um vetor, isto é, um vetor de entrada, que pode ser usado para preparar um vetor de montagem. Em algumas formas de realização, um vetor de entrada é um polinucleotídeo circular que compreende um marcador selecionável, uma origem de replicação e um segmento DNA imediatamente flanqueado por dois locais de restrição que facilitam a subclonagem de segmentos de DNA diferentes a serem montados nos métodos de montagem fornecido neste. O vetor de entrada compreende adicionalmente uma ou duas sequências ligantes aneladas ou uma sequência ligante anelada e um segmento de ligação iniciador, flanqueando os locais de restrição. O vetor de entrada compreende adicionalmente um par adicional de locais de restrição posicionados nos flancos externos do segmento DNA, por exemplo, que flanqueiam um ou duas sequências ligantes aneladas ou a sequência ligante anelada e segmento de ligação iniciador. Deste modo, em algumas formas de realização, o vetor de entrada é um polinucleotídeo circular que compreende, em uma orientação da extremidade 5' a 3', um local de restrição RA, uma sequência ligante anelada LA, um local de restrição RY, um segmento de DNA D, um local de restrição RZ e um local de restrição RB. Em outras formas de realização, o vetor de entrada é um polinucleotídeo circular que compreende, em uma orientação a extremidade de 5' a 3', um local de restrição RA, um local de restrição RY, um segmento de DNA D, um local de restrição RZ, uma sequência ligante anelada LB e um local de restrição RB. Em outras formas



de realização, o vetor de entrada é um polinucleotídeo circular que compreende, em uma orientação da extremidade 5' a 3', um local de restrição RA, um segmento de ligação iniciador PA ou uma sequência ligante anelada LA, um local de restrição RY, um segmento de DNA D, um local de restrição RZ, um segmento de ligação iniciador PB ou uma sequência ligante anelada LB e um local de restrição RB.

[00116] Em algumas formas de realização, a sequência do segmento de DNA D do vetor de entrada é o gene repórter lac Z. O gene repórter lac Z é útil para facilitar a seleção branca/azul das colônias transformadas com vetores que compreendem os segmentos de DNA outros do que lac Z, por exemplo, durante a preparação de um vetor de montagem descrito neste.

[00117] Em algumas formas de realização, o vetor de entrada é um polinucleotídeo circular que compreende, em uma orientação da extremidade 5' a 3', um local de restrição RA, uma sequência ligante anelada LA, um local de restrição RY, um segmento de DNA D, um local de restrição RZ e um local de restrição RB (isto é, 5'-RA-LA-RY-D-RZ-RB-3'). Em algumas formas de realização, o vetor de entrada é um polinucleotídeo circular que compreende, em uma orientação da extremidade 5' a 3', um local de restrição RA, um local de restrição RY, um segmento de DNA D, um local de restrição RZ, uma sequência ligante anelada LB e um local de restrição RB (isto é, 5'-RA-RY-D-RZ-LB-RB-3'). Em algumas formas de realização, o vetor de entrada é um polinucleotídeo circular que compreende, em uma orientação da extremidade 5' a 3', um local de restrição RA, uma sequência ligante anelada LA, um local de restrição RY, um segmento de DNA D, um local de restrição RZ, uma sequência ligante anelada LB e um local de restrição RB (isto é, 5'-RA-LA-RY-DRZ-LB-RB-3'). Em algumas formas de realização, o vetor de entrada é um polinucleotídeo circular que compreende, em uma orientação da extremidade 5' a 3', um local de restrição RA, um segmento de ligação iniciador PA, um local de restrição RY, um segmento de DNA D, um local de

restrição RZ, uma sequência ligante anelada LB e um local de restrição RB (isto é, 5'-RA-PA-RY-D-RZ-LB-RB-3'). Em algumas formas de realização, o vetor de entrada é um polinucleotídeo circular que compreende, em uma orientação da extremidade 5' a 3', um local de restrição RA, uma sequência ligante anelada LA, um local de restrição RY, um segmento de DNA D, um local de restrição RZ, um segmento de ligação iniciador PB e um local de restrição RB (isto é, 5'-RALA-RY-D-RZ-PB-RB-3'). Um vetor de entrada exemplar é fornecido na FIG. 1A.

[00118] O segmento de ligação iniciador pode ser qualquer sequência de nucleotídeo que não é complementar com qualquer uma das sequências ligantes aneladas que são usadas para preparar um polinucleotídeo montado. Em algumas formas de realização, os dois segmentos de ligação iniciador incluem um reconhecimento de endonuclease de restrição e local de clivagem. Em algumas formas de realização, o segmento de ligação iniciador é simplesmente uma das sequências ligantes disponíveis que não estão sendo usadas em uma reação de montagem particular. Em algumas formas de realização, a sequência de ácido nucleico do segmento de ligação iniciador PA é selecionada a partir do grupo que consiste de SEQ ID N°: 24 e 25. Em algumas formas de realização, a sequência de ácido nucleico do segmento de ligação iniciador PB é selecionada a partir do grupo que consiste de SEQ ID N°: 24 e 25. Em algumas formas de realização, a sequência de ácidos nucleicos do segmento de ligação iniciador PA e segmento de ligação iniciador PB são selecionados a partir do grupo que consiste de SEQ ID N°: 24 e 25. Nas formas de realização preferidas, PA e PB não são idênticos na sequência.

[00119] Em algumas formas de realização, a sequência de ácido nucleico da sequência LA ou LB ligante anelada é pelo menos 24 nucleotídeos e tem um  $T_m$  de pelo menos 60°C. Em algumas formas de realização, a sequência de ácido nucleico da sequência ligante anelada LA é selecionada a partir do grupo que consiste de SEQ ID N°: 1 a 23. Em

algumas formas de realização, a sequência nucléica da sequência ligante anelada LB é selecionada a partir do grupo que consiste de SEQ ID N°: 1 a 23. Em algumas formas de realização, as sequências nucléicas da sequência ligante anelada LA e sequência ligante anelada LB são selecionadas a partir do grupo que consiste de SEQ ID N°: 1 a 23.

[00120] Os locais de restrição RY e RZ podem ser utilizados como locais de clonagem para introduzir vários segmentos de DNA para a geração de um vetor de montagem. Em algumas formas de realização, RY e RZ não são idênticos na sequência. Em algumas formas de realização, RY e RZ são clivados pela mesma endonuclease de restrição. Em algumas formas de realização, RY e RZ são idênticos na sequência. Em algumas formas de realização, locais de restrição RY e RZ são clivados pela endonuclease de restrição que geram extremidades escalonadas, isto é terminais tendo uma projeção de extremidade 5' ou 3'. Em outras formas de realização, locais de restrição RY e RZ são clivados pela endonuclease de restrição que geram extremidades abruptas.

[00121] Embora os locais de restrição RY e RZ podem ser qualquer local de restrição conhecido na técnica, locais de restrição reconhecidos pela endonucleases de restrição tipo IIS são particularmente úteis. As endonucleases de restrição tipo IIS tem domínios de ligação de DNA que são distintas a partir dos seus domínios de clivagem. Portanto, estes reconhecem uma sequência específica mas clivam uma distância definida. Por exemplo, a Endonuclease de restrição Tipo IIS SchI (que também é conhecido como MlyI) liga-se a um local de reconhecimento contendo a sequência GAGTC e cliva-se quatro (4) pares de base distantes do local de reconhecimento, criando uma molécula de DNA de extremidade abrupta. Os locais de restrição Tipo IIS são particularmente úteis para a preparação de um vetor de montagem a partir de um vetor de entrada. Por exemplo, em um procedimento de subclonagem em que o segmento DNA de um vetor de entrada, por

exemplo lacZ, é substituído com um segmento DNA de interesse, excisão de lacZ com um endonuclease de restrição tipo IIS pode resultar na remoção completa da sequência de reconhecimento do local de restrição. Como um resultado, na ligação do segmento DNA de interesse ao vetor de entrada linearizado, a sequência estranha entre a sequência ligante anelada ou o segmento de ligação iniciador e o segmento DNA novamente introduzido é minimizado.

[00122] Deste modo, em algumas formas de realização, locais de restrição RY e RZ são locais de restrição reconhecidos e clivados por qualquer endonuclease de restrição tipo IIS conhecido na técnica. A endonucleases de restrição tipo IIS adequada inclui mas não são limitadas a seguinte endonuclease e seus isosquizômeros, que são indicados nos parênteses: Alw26I (BsmAI), AlwI (AclWI, BinI), AsuHPI (HphI), BbvI (Bst71I), BceFI, BstF5I (BseGI, FokI), FauI, HgaI, SapI (LguI), MboII, PleI, SapI, SchI (MlyI), SfaNI e TspRI, AceIII, BbsI (BbvII, BpiI, BpuAI), Bce83I, BciVI, BfiI (BmrI), BpmI (GsuI), BsaI (Eco31I), BseRI, BsgI, BsmBI (Esp3I), BsmFI, BspMI, BsrDI (Bse3DI), Bsu6I (Eam1104I, Earl, Ksp632I), Eco57I, Fail, Mmel, RleAI, TagII e Tth111II. Em formas de realização particulares, locais de restrição RY e RZ são reconhecidos e clivados pela endonuclease de restrição SchI .

[00123] Em algumas formas de realização, RA e RB não são idênticos na sequência. Em algumas formas de realização, RA e RB são clivados pela mesma endonuclease de restrição. Em algumas formas de realização, RA e RB são idênticos na sequência. Em algumas formas de realização, locais de restrição RA e RB são clivados pela endonuclease de restrição que geram extremidades escalonadas, isto é terminais tendo uma projeção de extremidade 5' ou 3'. Em outras formas de realização, locais de restrição RA e RB são clivados pela endonuclease de restrição que geram extremidades abruptas.

[00124] Embora os locais de restrição RA e RB podem ser quaisquer

locais de restrição conhecidos na técnica, locais de restrição que são relativamente infrequentes em DNA (por exemplo, cDNA) de um ou mais organismos (isto é, um cortador infrequente) são particularmente úteis. Em algumas formas de realização, locais de restrição RA e RB são reconhecidos e clivados pela endonuclease de restrição que tem locais de restrição relativamente infrequentes no DNA humano. Em algumas formas de realização, locais de restrição RA e RB são reconhecidos e clivados pela endonuclease de restrição que tem os locais de restrição relativamente infrequentes no DNA de camundongo. Em algumas formas de realização, locais de restrição RA e RB são reconhecidos e clivados pela endonuclease de restrição que tem os locais de restrição relativamente infrequentes no DNA de levedura, por exemplo, no DNA de *Saccharomyces cerevisiae*, *Pichia pastoris*, *Kluyveromyces lactis*, *Arxula adeninivorans* ou *Hansenula polymorpha*. Em algumas formas de realização, locais de restrição RA e RB são reconhecidos e clivados pela endonuclease de restrição que tem relativamente poucos locais de restrição no DNA de bactéria, por exemplo, no DNA de *Escherichia coli* ou *Bacillus subtilis*.

[00125] Em algumas formas de realização, locais de restrição RA e RB são reconhecidos e clivados pela endonuclease de restrição tipo IIS em que o local de reconhecimento é distal à polissequência de nucleotídeo que compreende, por exemplo, PA/LA-D-PB/LB. Em algumas formas de realização, cada local de restrição RA e RB é independentemente reconhecido e clivado pela endonuclease de restrição selecionado a partir do grupo que consiste de MssI, NruI (Bsp68I, MluB2I, Sbo13I, SpoI), SnaBI (BstSNI, Eco1051), SrfI e SwaI (BstRZ2461, BstSWI, MspSWI, SmlI), HpaI, HincII, PshAI, OliI, AluI, Alw261, Ball, DraI, DpnI, EcoR47III, EcoRCRI, EcoRV, FokI, HaeIII, HincII, MboI, MspA1I, NaeI, RsaI, PvuII, ScaI, SmaI, SspI, StuI, Xmnt, EcaBC3I, SciI, HincII, DraI, BsaBI, Cac8I, Hpy8I, MlyI, PshAI, SspD51, BfrBI, BsaAI, BfrBI, BtrI, CdiI, CviJI, CviRI, Eco47III, Eco78I,

EcoICRI, FnuDII, FspAI, HaeI, LpnI, MlyI, MslI, MstI, NaeI, NlaIV, NruI, NspBII, OsiI, PmaCI, PshAI, PsiI, SrfI, StuI, XcaI, XmnI, ZraI e isosquizômeros destes. Em uma forma de realização particular, locais de restrição RA e RB são reconhecidos e clivados pela endonuclease de restrição. SapI ou LguI. LguI é um isosquizômero de SapI tendo uma mesma especificidade de clivagem e reconhecimento.

[00126] Em algumas formas de realização, o vetor de entrada fornecido neste também compreende uma ou mais da sequência de ácido nucleicos que geralmente tem a mesma função na replicação, na manutenção ou integridade do vetor (por exemplo origem das replicações) bem como um ou mais marcadores selecionáveis. Origens de replicação são polinucleotídeos únicos que compreendem sequências repetidas curtas múltiplas que são reconhecidas pelas proteínas de ligação de origem multimérica e que desempenham um papel chave na montagem das enzimas de replicação de DNA no local de origem. Origens adequadas de replicação para uso nos vetores de entrada e montagem fornecidos neste incluem mas não são limitados a *E. coli* oriC, origem de plasmídeo colE1, 2  $\mu$  e ARS (ambos úteis no sistema de levedura), ski, SV40 EBV oriP (úteis no sistema de mamíferos) ou aqueles observados em pSC101. Os marcadores selecionáveis podem ser elementos úteis nos vetores como estes fornecem um significado para selecionar para ou contra o desenvolvimento das células que foram sucessivamente transformadas com um vetor contendo um marcador selecionável e expressam o marcador.

[00127] Em algumas formas de realização, qualquer vetor pode ser usado para construir o vetor de entrada como fornecido neste. Em particular, vetores conhecidos na técnica e aqueles comercialmente disponíveis (e variantes ou derivados destes) podem ser projetados para incluir um local de restrição RA, opcionalmente um segmento de ligação iniciador PA ou uma sequência ligante anelada LA, um local de restrição RY, um segmento de DNA D, um local de restrição RZ, opcionalmente um segmento de ligação

iniciador PB ou uma sequência ligante anelada LB e um local de restrição RB, para uso nos métodos fornecido neste. Tais vetores podem ser obtidos de, por exemplo, Vector Laboratories Inc., InVitrogen, Promega, Novagen, NEB, Clontech, Boehringer Mannheim, Pharmacia, EpiCenter ou iGenes Technologies Inc., Stratagene, Perkin Elmer, Pharmingen, Life Technologies, Inc. e Research Genetics. Classes gerais de vetores de interesse particular incluem vetores de clonagem procarióticos e/ou eucarióticos, vetores de expressão, vetores de fusão, vetores de dois híbridos reversos ou dois híbridos, vetores oscilantes para uso nos hospedeiros diferentes, vetores de mutagênese, vetores de transcrição, vetores para receber inserções grandes e outros. Outros vetores de interesse incluem vetores de origem viral (vetores M13, vetores  $\gamma$  de fago bacteriano, vetores de adenovírus e vetores de retrovírus), vetores do número de cópias ajustáveis, alta e baixa, vetores que tem replicons compatíveis para uso na combinação em um hospedeiro simples (PACYC184 e pBR322) e vetores de replicação espissomal eucariótico (pCDM8).

[00128] Em formas de realização particulares, vetores de entrada para o uso de acordo com os métodos fornecidos neste são os vetores pRYSE, tendo uma sequência de nucleotídeos da SEQ ID N°: 207 até 221. Um esquemático dos vetores pRYSE é fornecido na FIG. 4 e a preparação dos vetores pRYSE é descrito no Exemplo 1 abaixo.

### 5.6 Vetores de montagem

[00129] Em outro aspecto, fornecido neste é um vetor, isto é, um vetor de montagem, que pode ser usado na montagem de uma pluralidade de polinucleotídeos componentes em um ou mais polinucleotídeos montados. Em algumas formas de realização, um vetor de montagem é um polinucleotídeo circular que compreende um marcador selecionável, uma origem de replicação e um segmento DNA flanqueado por uma sequência ligante anelada, um par da sequência ligante anelada ou por uma sequência

ligante anelada/par de segmento de ligação iniciador, flanqueado por um par de locais de restrição. Os locais de restrição podem servir para facilitar a excisão do polinucleotídeo componente a partir da estrutura de vetor de montagem durante a reação de montagem. Deste modo, em algumas formas de realização, o vetor de montagem é um polinucleotídeo circular que compreende, em uma orientação da extremidade 5' a 3', um local de restrição RA, um segmento de ligação iniciador PA ou uma sequência ligante anelada LA, um segmento de DNA D e um local de restrição RB. Em algumas formas de realização, o vetor de montagem é um polinucleotídeo circular que compreende, em uma orientação da extremidade 5' a 3', um local de restrição RA, um segmento de DNA D, um segmento de ligação iniciador PB ou uma sequência ligante anelada LB e um local de restrição RB. Em certas formas de realização, o vetor de montagem é um polinucleotídeo circular que compreende, em uma orientação da extremidade 5' a 3', um local de restrição RA, um segmento de ligação iniciador PA ou uma sequência ligante anelada LA, um segmento de DNA D, um segmento de ligação iniciador PB ou uma sequência ligante anelada LB e um local de restrição RB.

[00130] Em algumas formas de realização, o vetor de montagem é um polinucleotídeo circular que compreende, em uma orientação da extremidade 5' a 3', um local de restrição RA, uma sequência ligante anelada LA, um segmento de DNA D e um local de restrição RB (isto é, 5'-RA-LA-D-RB-3'). Em algumas formas de realização, o vetor de montagem é um polinucleotídeo circular que compreende, em uma orientação da extremidade 5' a 3', um local de restrição RA, um segmento de DNA D, uma sequência ligante anelada LB e um local de restrição RB (isto é, 5'-RA-D-LB-RB-3'). Em algumas formas de realização, o vetor de montagem é um polinucleotídeo circular que compreende, em uma orientação da extremidade 5' a 3', um local de restrição RA, uma sequência ligante anelada LA, um segmento de DNA D, uma sequência ligante anelada LB e um local de restrição RB (isto é, 5'-RA-LA-D-



LB-RB-3'). Em algumas formas de realização, o vetor de montagem é um polinucleotídeo circular que compreende, em uma orientação da extremidade 5' a 3', um local de restrição RA, um segmento de ligação iniciador PA, um segmento de DNA D, uma sequência ligante anelada LB e um local de restrição RB (isto é, 5'-RA-PA-D-LB-RB-3'). Em algumas formas de realização, o vetor de montagem é um polinucleotídeo circular que compreende, em uma orientação da extremidade 5' a 3', um local de restrição RA, uma sequência ligante anelada LA, um segmento de DNA D, um segmento de ligação iniciador PB e um local de restrição RB (isto é, 5'-RA-LA-D-PB-RB-3'). Vetores de montagem exemplares são fornecidos na FIG. 1B e FIG. 2.

[00131] Em algumas formas de realização, a sequência de ácido nucleico do segmento de ligação iniciador PA é selecionada a partir do grupo que consiste de SEQ ID N°: 24 e 25. Em algumas formas de realização, a sequência de ácido nucleico do segmento de ligação iniciador PB é selecionada a partir do grupo que consiste de SEQ ID N°: 24 e 25. Em algumas formas de realização, a sequência de ácidos nucleicos do segmento de ligação iniciador PA e segmento de ligação iniciador PB são selecionados a partir do grupo que consiste de SEQ ID N°: 24 e 25. Nas formas de realização preferidas, a sequência de ácidos nucleicos do segmento de ligação iniciador PA e segmento de ligação iniciador PB não são idênticos.

[00132] Em algumas formas de realização, a sequência de ácido nucleico da sequência LA ou LB ligante anelada é pelo menos 24 nucleotídeos e tem um  $T_m$  de pelo menos 60°C. Em algumas formas de realização, a sequência de ácido nucleico da sequência ligante anelada LA é selecionada a partir do grupo que consiste de SEQ ID N°: 1 a 23. Em algumas formas de realização, a sequência de ácido nucleico da sequência ligante anelada LB é selecionado a partir do grupo que consiste de SEQ ID N°: 1 a

23. Em algumas formas de realização, a sequência de ácido nucleicos da sequência ligante anelada LA e a sequência ligante anelada LB são selecionados a partir do grupo que consiste de SEQ ID N°: 1 a 23.

[00133] Em algumas formas de realização, RA e RB não são idênticos na sequência. Em algumas formas de realização, RA e RB são clivados pela mesma endonuclease de restrição. Em algumas formas de realização, RA e RB são idênticos na sequência. Em algumas formas de realização, locais de restrição RA e RB são clivados pela endonuclease de restrição que geram extremidades escalonadas, isto é terminais tendo uma projeção de extremidade 5' ou 3'. Em outras formas de realização, locais de restrição RA e RB são clivados pela endonuclease de restrição que geram extremidades abruptas.

[00134] Embora os locais de restrição RA e RB podem ser quaisquer locais de restrição conhecido na técnica, locais de restrição que são relativamente infrequentes em DNA (por exemplo, cDNA) de um ou mais organismos (isto é, um cortador infrequente) são particularmente úteis. Em algumas formas de realização, locais de restrição RA e RB são reconhecidos e clivados pela endonuclease de restrição que tem os locais de restrição relativamente infrequentes no DNA humano. Em algumas formas de realização, locais de restrição RA e RB são reconhecidos e clivados pela endonuclease de restrição que tem os locais de restrição relativamente infrequentes no DNA de camundongo. Em algumas formas de realização, locais de restrição RA e RB são reconhecidos e clivados pela endonuclease de restrição que tem os locais de restrição relativamente infrequentes no DNA de levedura, por exemplo, no DNA de *Saccharomyces cerevisiae*, *Pichia pastoris*, *Kluyveromyces lactis*, *Arxula adeninivorans* ou *Hansenula polymorpha*. Em algumas formas de realização, locais de restrição RA e RB são reconhecidos e clivados pela endonuclease de restrição que tem relativamente poucos locais de restrição no DNA de bactéria, por exemplo, no DNA de *Escherichia coli* ou *Bacillus subtilis*.

[00135] Em algumas formas de realização, locais de restrição RA e RB são reconhecidos e clivados pela endonuclease de restrição tipo IIS. Exemplos ilustrativos são endonucleases de restrição tipo IIS adequadas incluem mas não são limitadas a: MssI, NruI (Bsp68I, MluB2I, Sbo13I, SpoI), SnaBI (BstSNI, Eco105I), SrfI e SwaI (BstRZ246I, BstSWI, MspSWI, SmiI), HpaI, HincII, PshAI, OliI, AluI, Alw26I, BalI, DraI, DpnI, EcoR471II, EcoRCRI, EcoRV, FokI, HaeIII, HincII, MboI, MspA1I, NaeI, RsaI, PvuII, ScaI, SmaI, SspI, StuI, XmnI, EcaBC3I, SciI, HincII, DraI, BsaBI, Cac8I, Hpy8I, MlyI, PshAI, SspD51, BfrBI, BsaAI, BsrBI, BtrI, CdiI, CviII, CviRI, Eco47III, Eco78I, EcoICRI, FnuDII, FspAI, HaeI, LpnI, MlyI, MslI, MstI, NaeI, NlaIV, NruI, NspBII, OliI, PmaCI, PshAI, PsiI, SrfI, StuI, XcaI, XmnI, ZraI ou isosquizômeros destes. Em uma forma de realização particular, locais de restrição RA e RB são reconhecidos e clivados pela endonuclease de restrição SapI ou LguI.

[00136] Preferivelmente, o segmento DNA de um vetor de montagem não compreende uma sequência de ácido nucleico que pode ser reconhecida e clivada pela endonuclease de restrição que pode clivar qualquer um dos locais de restrição RA e RB dentro do vetor de montagem. Este garante que o segmento DNA permanece intacto durante um primeiro estágio da reação de montagem, durante o qual, o polinucleotídeo componente é taxado a partir da estrutura de vetor de montagem. Em formas de realização particulares, o segmento de DNA D não compreende um local SapI/LguI e RA e RB são clivados por SapI ou LguI. Mutagênese direcionada ao local (ver Carter, BioChem. 1 237:1-7 (1986); Zoller and Smith, Methods Enzymol. 154:329-50 (1987)), mutagênese de cassete, mutagênese de seleção de restrição (Wells et al., Gene 34:315-323 (1985)), mutagênese mediada pelo oligonucleotídeo (direcionada ao local), mutagênese de PCR ou outras técnicas conhecidas podem ser realizadas para modificar tais quaisquer sequências dentro do segmento DNA antes ou depois da ligação do segmento DNA ao vetor de

entrada.

[00137] Em algumas formas de realização, o vetor de montagem fornecido neste também compreende uma ou mais sequência de ácido nucleicos que geralmente tem a mesma função na replicação, na manutenção ou integridade do vetor (por exemplo origem das replicações) bem como um ou mais marcadores selecionáveis. As origens de replicação são polinucleotídeos únicos que compreendem sequências repetidas curtas múltiplas que são reconhecidas pelas proteínas de ligação de origem multimérica e que desempenham um papel chave na montagem das enzimas de replicação de DNA no local de origem. As origens adequadas de replicação para uso nos vetores de entrada e montagem fornecidos neste incluem mas não são limitadas a *E. coli* oriC, origem de plasmídeo colE 1, 2  $\mu$  e ARS (ambos úteis no sistema de levedura), sfl, SV40 EBV oriP (úteis no sistema de mamíferos) ou aqueles observados em pSC101. Marcadores selecionáveis podem ser elementos úteis nos vetores como estes fornecem para selecionar para ou contra o desenvolvimento das células que foram sucessivamente transformadas com um vetor contendo um marcador selecionável e expressam o marcador.

[00138] Em algumas formas de realização, qualquer vetor pode ser usado para construir o vetor de montagem como fornecido neste. Em particular, vetores conhecidos na técnica e aqueles comercialmente disponíveis (e variantes ou derivados destes) podem ser projetados para incluir um local de restrição RA, um segmento de ligação iniciador PA ou uma sequência ligante anelada LA, um segmento de DNA D, um segmento de ligação iniciador PB ou uma sequência ligante anelada LB e um local de restrição RB, para uso nos métodos fornecido neste. Tais vetores podem ser obtidos de, por exemplo, Vector Laboratories Inc., InVitrogen, Promega, Novagen, NEB, Clontech, Boehringer Mannheim, Pharmacia, EpiCenter ou iGenes Technologies Inc., Stratagene, Perkin Elmer, Pharmingen, Life

Technologies, Inc. e Research Genetics. Classes gerais de vetores de interesse particular incluem vetores de clonagem procarióticos e/ou eucarióticos, vetores de expressão, vetores de fusão, vetores de dois híbridos reversos ou dois híbridos, vetores oscilantes para uso nos hospedeiros diferentes, vetores de mutagênese, vetores de transcrição, vetores para receber inserções grandes e outros. Outros vetores de interesse incluem vetores de origem viral (vetores M13, vetores  $\gamma$  de fago bacteriano, vetores de adenovírus e vetores de retrovírus), vetores do número de cópias ajustáveis, alta e baixa, vetores que tem replicons compatíveis para uso na combinação em um hospedeiro simples (PACYC 184 e pBR322) e vetores de replicação espissomal eucariótico (pCDM8).

[00139] Um vetor de montagem pode ser preparado a partir de um vetor de entrada. Para preparar um vetor de montagem a partir de um vetor de entrada, o vetor de entrada pode ser digerido com um ou mais endonucleases de restrição capazes da clivagem RY e RZ por meio deste linearizando o vetor tal que este pode aceitar um segmento DNA. O segmento DNA pode ser ligado em locais RY e RZ usando técnicas de clonagem padrão para gerar um vetor de montagem da invenção. Por exemplo, o segmento DNA pode ser obtido pelos procedimentos padrão conhecidos na técnica do DNA clonado (por exemplo, uma "biblioteca" de DNA), pela síntese química, pela clonagem de cDNA ou pela clonagem do DNA genômico ou fragmentos destes, purificados a partir da célula desejada ou pela amplificação de PCR e clonagem. Ver, por exemplo, Sambrook *et al.*, *Molecular Cloning, A Laboratory Manual*, 3d. ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York (2001); Glover, D.M. (ed.), *DNA Cloning: A Practical Approach*, 2d. ed., MRL Press, Ltd., Oxford, U.K. (1995).

[00140] Um vetor de montagem também pode ser preparado a partir de outro vetor que não compreende uma sequência ligante anelada, um par da sequência ligante anelada ou uma sequência ligante anelada/par de segmento

de ligação iniciador flanqueando o local da inserção do segmento DNA. Para preparar um vetor de montagem a partir de um tal vetor, o vetor pode ser digerido com um ou mais endonucleases de restrição capazes da clivagem do vetor em um local adequado para a inserção de um fragmento de DNA, por exemplo, em um local de clonagem múltipla, por meio deste linearizando o vetor tal que este pode aceitar um fragmento de DNA. O fragmento de DNA a ser inserido pode ser obtido pelos procedimentos padrão conhecido na técnica tal como, por exemplo, clonagem, síntese química ou amplificação de PCR. O fragmento de DNA compreende um segmento de DNA flanqueado por uma sequência ligante anelada, um par da sequência ligante anelada ou uma sequência ligante anelada/par de segmento de ligação iniciador. Deste modo, em algumas formas de realização, o fragmento de DNA compreende, em uma orientação da extremidade 5' a 3', uma sequência ligante anelada LA ou um segmento de ligação iniciador PA, um segmento de DNA D e uma sequência ligante anelada LB ou um segmento de ligação iniciador PB (isto é, 5'-LA-D-LB-3' or 5'-PA-D-LB-3' or 5'-LA-D-PB-3'). Em algumas formas de realização, o fragmento de DNA compreende, em uma orientação da extremidade 5' a 3', um segmento de DNA D e uma sequência ligante anelada LB ou um segmento de ligação iniciador PB (isto é, 5'-D-LB-3' ou 5'-D-PB-3'). Em algumas formas de realização, o fragmento de DNA compreende, em uma orientação da extremidade 5' a 3', uma sequência ligante anelada LA ou um segmento de ligação iniciador PA e um segmento de DNA D, (isto é, 5'-LA-D-3' ou 5'-PA-D-3'). O fragmento de DNA ainda pode compreender um par da locais de restrição que flanqueiam uma sequência ligante anelada, um par da sequência ligante anelada ou uma sequência ligante anelada/par de segmento de ligação iniciador e que na clivagem pela endonuclease de restrição produz terminais que são compatíveis com os terminais produzidos pela linearização do vetor em que o fragmento de DNA será inserido. Alternativamente, o fragmento de DNA gerado tal que este contém tais

terminais compatíveis e não requerem a digestão adicional com uma endonuclease de restrição para produzir os terminais compatíveis. Na ligação do fragmento de DNA com o vetor linearizado para gerar um vetor de montagem, os locais de restrição usados para gerar os terminais compatíveis podem ser preservados para servir como locais de restrição RA e RB do vetor de montagem. Alternativamente, a ligação pode remover os locais de restrição originais mas locais de restrição adicionais podem estar presentes no vetor linearizado que podem servir como os locais de restrição RA e RB do vetor de montagem.

[00141] Métodos exemplares para a geração de um vetor de montagem a partir de um vetor de entrada (isto é, um vetor pRYSE) ou de outro vetor (isto é, um vetor pMULE) são fornecidos no Exemplo 6 abaixo.

#### 5.7 Sequências ligantes aneladas

[00142] Em outro aspecto, fornecido neste são as sequências ligantes aneladas que flanqueiam o segmento DNA localizado dentro dos vetores de entrada e vetores de montagem. As sequências ligantes aneladas fornecem a sobreposição da sequência entre os polinucleotídeos componentes adjacentes em uma reação de montagem e deste modo servem para preparar um polinucleotídeo componente para a montagem em um polinucleotídeo montado. Deste modo, nas formas de realização preferidas, as sequências LA ou LB ligantes aneladas dos vetores de montagem e entrada são otimizadas para fornecer a imprimação eficiente e precisa as sequências ligantes aneladas complementares durante uma reação de montagem.

[00143] Em algumas formas de realização, o comprimento de uma sequência ligante anelada é longo o bastante para fornecer especificidade adequada com sua sequência de complemento ligante anelada de, já curta o bastante para anelar prontamente a sua sequência de complemento ligante anelada na temperatura de anelamento da reação de montagem. Em algumas formas de realização o comprimento de uma sequência ligante anelada é

longo o bastante permitindo a recombinação homóloga mediada pela célula hospedeira com sua sequência complemento ligante anelada.

[00144] Em algumas formas de realização, a sequência ligante anelada é cerca de 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75 ou 80 nucleotídeos em comprimento. Em algumas formas de realização, a sequência ligante anelada é pelo menos 10, 12, 14, 16, 18, 20, 22, 24, 26, 28 ou 30 nucleotídeos em comprimento. Em algumas formas de realização, a sequência ligante anelada é maior do que 30, 40, 50, 60, 70, 80, 90, 100, 500, 1000, 5000 ou 10.000 nucleotídeos em comprimento. Em algumas formas de realização, o ligante anelado é pelo menos 18 nucleotídeos em comprimento e é um número dividido por três, deste modo para facilitar leitura através da transcrição do ligado quando ligado a um segmento DNA codificado. Em formas de realização particulares, o ligante anelado é 18, 21, 24, 27, 30, 33, 36, 39, 42, 45, 48, 51, 54, 57 ou 60 nucleotídeos em comprimento.

[00145] Em algumas formas de realização, uma sequência ligante anelada tem uma temperatura de fusão relativamente alta ( $T_m$ ), isto é, a temperatura em que uma metade de um duplex da sequência ligante anelada dissociará para tornar-se filamentado simples. A  $T_m$  de um ligante anelado pode ser calculado de acordo com SantaLucia, PNAS, 95:-1460-1465 (1998) usando um algoritmo vizinho mais próximo. Uma  $T_m$  relativamente alta pode fornecer para mais imprimação específica durante uma reação de montagem. Uma  $T_m$  relativamente alta também pode permitir a combinação das etapas de anelamento e extensão de PCR ou reduz a quantidade do período necessário para ajustar as temperaturas entre as etapas de anelamento e extensão de PCR e deste modo capacita a maior eficiência no uso dos métodos de montagem da invenção. Deste modo, em algumas formas de realização, um duplex da sequência ligante anelada tem uma  $T_m$  de cerca de 60°C a 80°C. Em algumas formas de realização, um duplex sequência ligante anelada tem uma  $T_m$  de cerca de 65°C a 75°C. Em algumas formas de



realização, um duplex sequência ligante anelada tem uma  $T_m$  maior do que 50°C, 55°C, 60°C, 65°C, 70°C, 75°C, 80°C, 85°C ou 90°C.

[00146] Em algumas formas de realização, sequências ligantes aneladas não formam as estruturas secundárias apreciáveis (por exemplo, grampo de cabelo, auto-dímeros) produzidos por intermédio das interações intramoleculares (isto é, dentro de uma mesma molécula) sob as condições dos métodos descrito neste, no nível DNA ou no nível RNA ou tanto no nível de DNA quanto RNA. A presença da estrutura secundária em DNA pode levar ao rendimento deficiente ou nulo de polinucleotídeo montado da reação de montagem. A presença da estruturas secundárias em RNA pode levar a eficiência de tradução diminuída, que são de interesse particular quando uma sequência ligante anelada é usada para montar os polinucleotídeos componentes que compreendem um promotor e uma sequência codificadora de proteína em um polinucleotídeo montado em que uma sequência ligante anelada é posicionada entre o promotor e a sequência codificadora de proteína. Consequentemente, as sequências ligantes aneladas úteis nos métodos de montagem da invenção são projetados para não formar as estruturas de RNA e/ou DNA. A capacidade de uma sequência ligante anelada para formar estruturas de RNA ou DNA secundários podem ser determinados usando ferramentas de software tal como, por exemplo, IDT Oligo Analyzer (Integrated DNA Technologies, Coralville, IA), mFold (Zuker 2003 *Nucleic Acids Res.* 31 (13), 3406-15) ou RNAfold (Hofacker & Stadler (2006) *Bioinformatics* 22 (10): 1172-6). No geral, estas ferramentas calcula a energia livre de Gibbs ( $\Delta G$ ) para a transição de uma sequência a partir do estado linear ao dobrado. O  $\Delta G$  maior, o mesmo semelhante aquele que a sequência formará uma estrutura secundária. Consequentemente, em algumas formas de realização, sequências ligantes aneladas são projetadas para ter os valores  $\Delta G$  maiores para a transição a partir dos estados dobrados a lineares. Em algumas formas de realização, sequências ligantes aneladas são projetadas para ter os

valores  $\Delta G$  para a transição a partir dos estados dobrados a lineares são iguais ou maiores do que os valores  $\Delta G$  para a transição a partir dos estados dobrados a lineares de n-bases que sustentam imediatamente a montante das sequências codificadoras dos genes altamente expressados no genoma *Saccharomyces cerevisiae*, em que n representa um inteiro que corresponde ao número de bases em uma sequência ligante anelada. Em algumas formas de realização, sequências ligantes aneladas são 36 bases longas e tem um valor  $\Delta G$  para a transição a partir dos estados dobrados a lineares de -1 ou maior.

[00147] Em algumas formas de realização, sequências ligantes aneladas também são projetadas para evitar as interações intermoleculares não pretendidas (isto é, entre moléculas diferentes). Deste modo, em algumas formas de realização, uma sequência ligante anelada não anela substancialmente com quaisquer outras sequências dentro do vetor de montagem que contém uma sequência ligante anelada (por exemplo, sequências de estrutura de vetor) e/ou com quaisquer outras sequências dentro de outros vetores de montagem de uma composição de montagem aparte das sequências ligantes aneladas complementares requeridas para a montagem de polinucleotídeo pelos métodos fornecidos no mesmo. Em algumas formas de realização, uma sequência ligante anelada não anela com outras sequências ligantes aneladas dentro dos vetores de montagem de uma composição de montagem fornecidos neste.

[00148] Em algumas formas de realização, uma sequência ligante anelada tem um conteúdo G-C alto, isto é, o número de nucleotídeos de citosina e guanina em uma sequência ligante anelada como uma porcentagem do número total de bases em uma sequência ligante anelada. As sequências ligantes aneladas que tem um conteúdo G-C alto são geralmente úteis nos métodos da invenção por causa de um conteúdo G-C alto geralmente fornecido para uma  $T_m$  alta, que volta para fornecer para imprimação mais específica durante uma reação de montagem e para economia de processo e

tempo permitindo a combinação das etapas de anelamento e extensão de SOE/PCR. Em algumas formas de realização, o conteúdo G-C da sequência ligante anelada está entre cerca de 20-80%. Em algumas formas de realização, o conteúdo G-C da sequência ligante anelada está entre cerca de 40 a 60%. Em algumas formas de realização, o conteúdo G-C da sequência ligante anelada é cerca de 40, 45, 50, 55, 60 ou 70%. Em formas de realização particulares, uma sequência ligante anelada tem um conteúdo G-C maior do que 70%. Exemplos ilustrativos das sequências ligantes aneladas que tem um conteúdo G-C alto, não formam estruturas de DNA secundárias apreciáveis e tem uma T<sub>m</sub> de 70°C ou maior são SEQ ID N°: 1 a 8.

[00149] Em algumas formas de realização, uma sequência ligante anelada tem um conteúdo A-T alto, isto é, o número de nucleotídeos de adenina e timina em uma sequência ligante anelada como uma porcentagem do número total de bases em uma sequência ligante anelada. Um conteúdo A-T alta pode fornecer para a propensão reduzida da sequência ligante anelada para formar estruturas secundárias substanciais, que podem ser de interesse particular quando uma sequência ligante anelada é usada para montar os polinucleotídeos componentes que compreendem um promotor e uma sequência codificadora de proteína em um polinucleotídeo montado no qual uma sequência ligante anelada é posicionado entre o promotor e a sequência codificadora de proteína. Em algumas formas de realização, o conteúdo A-T da sequência ligante anelada está entre cerca de 20 a 80%. Em algumas formas de realização, o conteúdo A-T da sequência ligante anelada está entre cerca de 40 a 60%. Em algumas formas de realização, o conteúdo A-T da sequência ligante anelada é cerca de 30, 35, 40, 45, 50, 55 ou 60%. Em algumas formas de realização, a sequência ligante anelada tem um conteúdo A-T maior do que 30%. Em algumas formas de realização, a sequência de 3'-mais 26 bases de uma sequência ligante anelada satisfaz o seguinte motivo consenso: 5'-A ANNNAANTANNTTNANA-3', em que A significa adenina,

N para qualquer nucleotídeo e T para timina. Este motivo consenso é frequentemente observado nas 26 bases que estão a montante dos códons de partida dos genes altamente expressados no genoma de *Saccharomyces cerevisiae*. Exemplos ilustrativos das sequências ligantes aneladas que compreendem este motivo consenso, tem um conteúdo A-T relativamente alto, não formam RNA secundário apreciável ou estruturas de DNA e tem um T<sub>m</sub> de 65°C ou maior são SEQ ID N°: 9 a 23.

[00150] Em algumas formas de realização, uma sequência ligante anelada compreende um ou mais locais de restrição. A incorporação dos locais de restrição em uma sequência ligante anelada permite a excisão de um segmento DNA a partir de um vetor de montagem ou entrada enquanto mantém os locais de restrição RA e RB dentro do vetor de entrada ou vetor de montagem. Locais de restrição dentro de uma sequência ligante anelada também facilita a subclonagem direcional dos segmentos de DNA em outros vetores de montagem ou de entrada. Esta característica facilita a construção eficiente dos vetores de montagem que compreendem o mesmo segmento DNA mas tendo pares da sequência ligante anelada diferente ou pares de segmento de ligação iniciador/sequência ligante anelada, por exemplo, para gerar uma biblioteca de vetores de montagem que compreendem pares da sequência ligante anelada diferente como descrito abaixo. Esta característica também pode evitar a necessidade de re-amplificar e sequenciar um segmento DNA para criar vetores de montagem adicionais que compreendem o segmento DNA. Deste modo, em algumas formas de realização, a sequência ligante anelada compreende um local de restrição único. Em algumas formas de realização, o local de restrição é um local de restrição de par de 7-base, isto é, é clivado pela endonuclease de restrição que reconhece uma sequência de nucleotídeo de par de 7-base. Em algumas formas de realização, o local de restrição é um local de restrição de par de 8-base. Em formas de realização particulares, o local de restrição dentro de uma sequência ligante anelada é

reconhecido e clivado por MreI, FseI, SbfI, AsiSI, NotI, AscI ou BbvCI.

[00151] Em algumas formas de realização, a sequência ligante anelada compreende uma sequência que permite para a transcrição através da leitura uma vez que o ligante é ligado e um DNA de segmento codificado. Em algumas formas de realização, uma sequência ligante anelada permite a transcrição através da leitura tanto na extremidade 5' quanto 3' e orientação da extremidade 3' a 5'. Nestas formas de realização, o comprimento da sequência ligante anelada, preferivelmente, é um número de nucleotídeos divididos por três (3).

[00152] Em formas de realização particulares, uma sequência ligante anelada não compreende códons que são raramente usados em *Escherichia coli* (*E. coli*) ou *Saccharomyces cerevisiae* (*S. cerevisiae*). A expressão eficiente dos genes heterólogos em *E. coli* ou *S. cerevisiae* podem ser adversamente afetados pela presença de códons infreqüentemente usados e níveis de expressão da proteína heteróloga freqüentemente aumentada quanto os códons raros são substituídos por um mais comum. Ver, por exemplo, Williams *et al.*, *Nucleic Acids Res.* 16: 10453-10467, 1988 and Hoog *et al.*, *Gene* 43: 13-21, 1986. Consequentemente, uma sequência ligante anelada que compreende uma sequência através da leitura preferivelmente não compreende códons raros usados em *E. coli* ou *S. cerevisiae*, deste modo capaz da expressão eficiente das proteínas codificadas por um polinucleotídeo montado que compreende a sequência ligante anelada.

[00153] Em algumas formas de realização, a série das sequências ligantes aneladas são sequências únicas que não são observadas em um organismo hospedeiro pretendido. Em algumas formas de realização, a série das sequências ligantes aneladas são sequências únicas que não são observadas em *E. coli*. Em outras formas de realização, a série das sequências ligantes aneladas são sequências únicas que não são observadas em *S. cerevisiae*.

[00154] Em algumas formas de realização, sequências ligantes aneladas adequadas são identificadas em um polinucleotídeo montado de teste. Um polinucleotídeo montado de teste compreende a sequência ligante anelada a ser testada e elementos adicionais que permitem o teste de uma sequência ligante anelada. Por exemplo, para testar se um ligante anelado é adequado para a montagem um primeiro polinucleotídeo componente que compreende uma sequência promotora e um segundo polinucleotídeo componente que compreende a sequência codificadora de proteína a ser colocada sob o controle do promotor no polinucleotídeo montado, um polinucleotídeo montado testado pode ser montado a partir do primeiro polinucleotídeo componente que compreende, em uma orientação da extremidade 5' a 3', um segmento de ligação iniciador ou uma sequência ligante anelada, um segmento DNA que compreende o promotor e a sequência ligante anelada a ser testada e o segundo polinucleotídeo componente que compreende, em uma orientação da extremidade 5' a 3', a sequência ligante anelada a ser testada, um segmento DNA que codifica um gene repórter (por exemplo, proteína fluorescente verde (GFP)), e um segmento de ligação iniciador ou sequência ligante anelada. O polinucleotídeo montado de teste pode ser testado in vivo ou in vitro para a eficiência da expressão do gene repórter. Os polinucleotídeos montados de teste similares podem ser montados para testar a conveniência das sequências ligantes aneladas para a montagem dos polinucleotídeos componentes que compreendem os segmentos de DNA que compreendem outros elementos, tal como um intensificador, terminador, cauda poli-A, sinal de localização nuclear, sinal de estabilização de mRNA, marcador selecionável, sequência codificadora de etiqueta de epítipo, sinal de degradação e outros. O polinucleotídeo montado de teste pode compreender polinucleotídeos componentes adicionais que capacitam o teste, tal como, por exemplo, sequências de alvejamento genômico e marcadores selecionáveis que

capacitam a introdução do polinucleotídeo montado de teste nas células hospedeiras e seleção dos transformantes positivos para o teste *in vivo*.

[00155] A tabela 1 apresenta a T<sub>m</sub>, locais de restrição e aminoácidos através da leitura das sequências ligantes aneladas exemplares correspondentes a SEQ ID N°: 1-23.

Tabela 1 – Sequência e características das sequências ligantes aneladas								
Sequência ligante anelável	Nome da sequência	Comprimento (bases)	G-C	% A-T	Temperatura de fusão (T <sub>m</sub> )	Local de enzima de restrição	Aminoácidos através da leitura	
							Fwd	Rev
SEQ ID N°: 1	RYSE 1	24	79,2	20,8	72,4			
SEQ ID N°: 2	RYSE 2	24	75,0	25,0	71,4	MreI		
SEQ ID N° 3 <sup>D</sup>	RYSE 3	24	75,0	25,0	73,7	FseI		TAGQA RGD
SEQ ID N°:4	RYSE 4	24	70,8	29,2	71,5	SbfI	NLQA ASAD	IGARG LQV
SEQ ID N°:5	RYSE 5	24	70,8	29,2	71,2	AsiSI	NAIAD AAD	IGGVG DRV
SEQ ID N°: 6	RYSE 6	24	70,8	29,2	70,9	NotI	KAAA GEGD	ISLASG RL
SEQ ID N°: 7	RYSE 7	24	70,8	29,2	71,5	AscI	KARH GRRD	
SEQ ID N°: 8	RYSE 8	24	75,0	25,0	70,7	BbvCI		
SEQ ID N°: 9	RYSE 9	36	50,0	50,0	67,4			
SEQ ID N°: 10	RYSE 10	36	52,8	47,2	67,7			
SEQ ID N°: 11	RYSE 11	36	58,3	41,7	69,2			
SEQ ID N°: 12	RYSE 12	36	50,0	50,0	67,4			
SEQ ID N°: 13	RYSE 13	36	58,3	41,7	69,4			
SEQ ID N°: 14	RYSE 14	36	52,8	47,2	67,4			
SEQ ID N°: 15	RYSE 15	36	52,8	47,2	67,8			
SEQ ID N°: 16	RYSE 16	36	52,8	47,2	67,8			
SEQ ID N°: 17	RYSE 17	36	52,8	47,2	68,4			
SEQ ID N°: 18	RYSE 18	36	50,0	50,0	67,8			
SEQ ID N°: 19	RYSE 19	36	52,8	47,2	68,1			
SEQ ID N°: 20	RYSE 20	36	55,6	44,4	68,3			
SEQ ID N°: 21	RYSE 21	36	55,6	44,4	67,9			
SEQ ID N°: 22	RYSE 22	36	52,8	47,2	67,4			
SEQ ID N°:23	RYSE 23	36	55,6	44,4	68,8			

## 5.8 Bibliotecas

[00156] Em outro aspecto, fornecido neste é uma biblioteca que compreende uma pluralidade de vetores de montagem. A biblioteca podem servir para facilitar a montagem eficiente de uma pluralidade de polinucleotídeos componentes em um ou mais polinucleotídeos montados que são funcionais em procariotas e eucariotas e deste modo facilitando a geração de organismos únicos, por exemplo, cepas recombinantes de bactéria e levedura, sem a necessidade para o período de consumo da endonuclease de restrição e enzima de ligase com base nas técnicas de clonagem. Os métodos de montagem e composições fornecidas neste podem facilitar a substituição eficiente ou introdução das unidades de DNA funcionais, por exemplo, promotores, intensificadores de origem das replicações, etc., dentro de uma construção da expressão e deste modo podem fornecer para a otimização eficiente da replicação de e/ou expressão de, a construção de expressão dentro de um organismo hospedeiro.

[00157] A biblioteca pode compreender uma pluralidade de vetores de montagem montados dentro de uma composição ou recipiente simples, por exemplo, uma composição ou recipiente adequado para realizar os métodos de montagem fornecidos no mesmo. Alternativamente, a biblioteca pode compreender uma pluralidade de vetores de montagem que não são montados dentro de uma mesma composição ou recipiente. Em algumas formas de realização, a biblioteca compreende pelo menos 3, pelo menos 6, pelo menos 10, pelo menos 20, pelo menos 50 ou mais do que 50 vetores de montagem, cada um que compreende a segmento DNA.

[00158] Em algumas formas de realização, a biblioteca compreende uma pluralidade de vetores de montagem em que cada um dos vetores de montagem compreendem, em uma orientação da extremidade 5' a 3', um primeiro local de restrição RA, um segmento de DNA D, uma sequência ligante anelada LB e um segundo local de restrição RB. Em algumas formas



de realização, a biblioteca compreende uma pluralidade de vetores de montagem em que cada um dos vetores de montagem compreendem, em uma orientação da extremidade 5' a 3', um primeiro local de restrição RA, um segmento de ligação iniciador PA ou uma primeira sequência ligante anelada LA, um segmento de DNA D e um segundo local de restrição RB. Em algumas formas de realização, a biblioteca compreende uma pluralidade de vetores de montagem em que cada um dos vetores de montagem compreende, em uma orientação da extremidade 5' a 3', um primeiro local de restrição RA, uma primeira sequência ligante anelada LA, um segmento de DNA D, uma sequência ligante anelada LB ou um segmento de ligação iniciador PB e um segundo local de restrição RB. Em algumas formas de realização, um par da sequência ligante anelada ou sequência ligante anelada/par de segmento de ligação primária dentro de cada vetor de montagem de uma biblioteca não compreende uma mesma sequência. Em algumas formas de realização, a sequência de ácido nucleico de uma sequência ligante anelada LA e/ou LB dentro de cada vetor de montagem é selecionado a partir do grupo que consiste de SEQ ID N°: 1 a 23. Em algumas formas de realização, a sequência de ácido nucleico do segmento de ligação iniciadora PA ou PB dentro de cada vetor de montagem é selecionado a partir do grupo que consiste de SEQ ID N°: 24 e 25.

[00159] Em algumas formas de realização, a biblioteca compreende pelo menos um de cada um dos seguintes vetores:

(a) um vetor que consiste de um polinucleotídeo circular que compreende, em uma orientação da extremidade 5' a 3', um local de restrição RA, um segmento de DNA D, uma sequência ligante anelada LB e um local de restrição RB;

(b) um vetor que consiste de um polinucleotídeo circular que compreende, em uma orientação da extremidade 5' a 3', um local de restrição RA, uma sequência ligante anelada LA, um segmento de DNA D,

uma sequência ligante anelada LB e um local de restrição RB; e

(c) um vetor que consiste de um polinucleotídeo circular que compreende, em uma orientação da extremidade 5' a 3', um local de restrição RA, uma sequência ligante anelada LA, um segmento de DNA D e um local de restrição RB<sub>0</sub>.

[00160] Em algumas formas de realização, a biblioteca compreende pelo menos um de cada um dos seguintes vetores:

(a) um vetor que consiste de um polinucleotídeo circular que compreende, em uma orientação da extremidade 5' a 3', um local de restrição RA, um segmento de ligação iniciador PA, um segmento de DNA D, uma sequência ligante anelada LB e um local de restrição RB;

(b) um vetor que consiste de um polinucleotídeo circular que compreende, em uma orientação da extremidade 5' a 3', um local de restrição RA, uma sequência ligante anelada LA, um segmento de DNA D, uma sequência ligante anelada LB e um local de restrição RB; e

(c) um vetor que consiste de um polinucleotídeo circular que compreende, em uma orientação da extremidade 5' a 3', um local de restrição RA, uma sequência ligante anelada LA, um segmento de DNA D, um segmento de ligação iniciador PB e um local de restrição RB<sub>0</sub>.

[00161] Em algumas formas de realização, a sequência de ácido nucleico do segmento de ligação iniciador PA é selecionada a partir do grupo que consiste de SEQ ID N°: 24 e 25. Em algumas formas de realização, a sequência de ácido nucleico do segmento de ligação iniciador PB é selecionada a partir do grupo que consiste de SEQ ID N°: 24 e 25. Em algumas formas de realização, a sequência de ácidos nucleicos do segmento de ligação iniciador PA e segmento de ligação iniciador PB são selecionados a partir do grupo que consiste de SEQ ID N°: 24 e 25.

[00162] Em algumas formas de realização, a sequência de ácido nucleico de quaisquer uma das ligantes aneladas LA e sequências ligantes

aneladas LB em uma biblioteca são selecionados a partir do grupo que consiste de SEQ ID N°: 1 a 23. Em algumas formas de realização, a sequência de ácido nucleicos de pelo menos uma das sequências ligantes aneladas LA e pelo menos uma das sequências ligantes aneladas LB em uma biblioteca são selecionados a partir do grupo que consiste de SEQ ID N°: 1 a 23. Em algumas formas de realização, a sequência de ácido nucleico de cada uma das sequências ligantes aneladas LA e as sequências ligantes aneladas LB em uma biblioteca é selecionada a partir do grupo que consiste de SEQ ID N°: 1 a 23.

[00163] Em algumas formas de realização, o segmento de DNA D compreende uma sequência nucléica selecionada a partir do grupo que consiste de um marcador selecionável, um promotor, uma sequência de alvejamento genômico, uma sequência de ácido nucleico que codifica um etiqueta de epítipo, a sequência de ácido nucleico que codifica um gene de interesse, a sequência de ácido nucleico que codifica um códon de terminação e lacZ.

[00164] Em algumas formas de realização, a biblioteca compreende pelo menos uma de cada uma das seguintes moléculas de ácido nucleico:

(a) uma primeira molécula de ácido nucleico em que a primeira molécula de ácido nucleico é circular e compreende, em uma orientação da extremidade 5' a 3', um primeiro local de restrição  $RA_0$ , qualquer segmento DNA selecionado a partir do grupo  $D_0$ , uma sequência ligante anelada  $LB_0$  e um segundo local de restrição  $RB_0$ ;

(b) uma molécula de ácido nucleico intermediária em que a molécula de ácido nucleico intermediária n é circular e compreende, em uma orientação da extremidade 5' a 3', um primeiro local de restrição  $RA_n$ , uma primeira sequência ligante anelada  $LA_n$ , qualquer segmento DNA selecionado a partir do grupo  $D_n$ , um segundo sequência ligante anelada  $LB_n$ , e um segundo local de restrição  $RB_n$ , e em que n representa um inteiro

de um ao número de moléculas de ácido nucleicos intermediárias; e

(c) uma última molécula de ácido nucleico em que a última molécula de ácido nucleico é circular e compreende, em uma orientação da extremidade 5' a 3', um primeiro local de restrição  $RA_m$ , uma sequência ligante anelada  $LA_m$ , qualquer segmento DNA selecionado a partir do grupo  $D_m$ , um segundo local de restrição  $RB_m$  em que  $m$  representa um inteiro um maior do que o número de moléculas de ácido nucleicos intermediárias;

[00165] Considerando a clivagem dos locais de restrição  $RA_0$  através de  $RB_m$  e desnaturação das moléculas de ácido nucleico lineares resultantes, cada sequência ligante anelada  $LB_{(p-1)}$  é capaz de hibridizar ao complemento da sequência ligante anelada  $LA_p$  em que  $p$  representa os inteiros de 1 a  $m$  e em que cada grupo  $D_0, \dots D_n, \dots$  e  $D_m$  consistem de um ou mais segmentos de DNA. Em algumas formas de realização, uma primeira molécula de ácido nucleico compreende adicionalmente um segmento de ligação iniciador PA posicionado na extremidade 5' ao segmento DNA selecionado a partir do grupo  $D_0$ . Em algumas formas de realização, uma última molécula de ácido nucleico compreende adicionalmente um segmento de ligação iniciador PB posicionado na extremidade 3' ao segmento DNA selecionado a partir do grupo  $D_m$ .

[00166] Em algumas formas de realização, na clivagem dos locais de restrição  $RA_0$  através de  $RB_m$ , e desnaturação das moléculas de ácido nucleico lineares resultantes, cada sequência ligante anelada  $LB_{(p-1)}$  é capaz de hibridizar seletivamente ao complemento da sequência ligante anelada  $LA_p$  comparado a outras sequências ligantes aneladas ou complementos do mesmo, na composição dos componentes. Em algumas formas de realização, cada sequência ligante anelada  $LB_{(p-1)}$  é idêntica na sequência à sequência ligante anelada  $LA_p$ .

[00167] Em uma forma de realização particular, os locais de restrição  $RA_0$  através de  $RB_m$  são clivados pela mesma endonuclease de restrição deste

modo para facilitar a excisão do polinucleotídeos componentes a partir dos vetores de montagem. Em algumas formas de realização, os locais de restrição  $RA_0$  através de  $RB_m$ , são clivados por endonucleases de restrição SapI e LguI.

[00168] Em algumas formas de realização, a sequência de ácido nucleico do segmento de ligação iniciador PA é selecionada a partir do grupo que consiste de SEQ ID N°: 24 e 25. Em algumas formas de realização, a sequência de ácido nucleico do segmento de ligação iniciador PB é selecionada a partir do grupo que consiste de SEQ ID N°: 24 e 25. Em algumas formas de realização, a sequência de ácido nucleicos do segmento de ligação iniciador PA e segmento de ligação iniciador PB são selecionados a partir do grupo que consiste de SEQ ID N°: 24 e 25. Nas formas de realização preferidas, a sequência de ácido nucleicos do segmento de ligação iniciador PA e segmento de ligação iniciador PB não são idênticos.

[00169] Em algumas formas de realização, a sequência de ácido nucleico de quaisquer uma das ligantes aneladas LA e sequências ligantes aneladas LB em uma biblioteca é selecionado a partir do grupo que consiste de SEQ ID N°: 1 a 23. Em algumas formas de realização, a sequência de ácido nucleicos de pelo menos uma das sequências ligantes aneladas LA e pelo menos uma das sequências ligantes aneladas LB em uma biblioteca são selecionados a partir do grupo que consiste de SEQ ID N°: 1 a 23. Em algumas formas de realização, a sequência de ácido nucleico de cada uma das sequências ligantes aneladas LA e as sequências ligantes aneladas LB em uma biblioteca é selecionada a partir do grupo que consiste de SEQ ID N°: 1 a 23. Em algumas formas de realização, a sequência de ácido nucleico de cada uma das sequências ligantes aneladas LA em uma composição não são idênticas a uma outra. Em algumas formas de realização, a sequência de ácido nucleico de cada uma das sequências ligantes aneladas LB em uma

composição não são idênticas a uma outra.

[00170] Em uma forma de realização particular, a biblioteca compreende as seguintes moléculas de ácido nucleico:

(a) duas primeiras moléculas de ácido nucleico, em que uma primeira molécula de ácido nucleico compreende, em uma orientação da extremidade 5' a 3', um primeiro local de restrição  $RA_0$ , um segmento de ligação iniciador PA, um segmento de DNA  $D_{01}$ , uma sequência ligante anelada  $LB_0$  e um segundo local de restrição  $RB_0$ , em que outra primeira molécula de ácido nucleico compreende, em uma orientação da extremidade 5' a 3', um primeiro local de restrição  $RA_0$ , um segmento de ligação iniciador PA, um segmento de DNA  $D_{02}$ , uma sequência ligante anelada  $LB_0$  e um segundo local de restrição  $RB_0$ , em que o segmento de DNA  $D_{01}$  codifica uma primeira sequência de alvejamento genômico, em que o segmento de DNA  $D_{02}$  codifica um segundo sequência de alvejamento genômico localizado a jusante da primeira sequência de alvejamento genômico em um genoma alvo e em que o segmento de DNA  $D_{02}$  é posicionado na orientação oposta como o segmento de DNA  $D_{01}$  relativo ao segmento de ligação iniciador PA e sequência ligante anelada  $LB_0$ ;

(b) pelo menos uma molécula de ácido nucleico intermediária que compreende, em uma orientação da extremidade 5' a 3', um primeiro local de restrição  $RA_n$ , uma primeira sequência ligante anelada  $LA_n$ , um segmento de DNA  $D_n$ , um segundo sequência ligante anelada  $LB_n$  e um segundo local de restrição  $RB_n$ , em que  $n$  representa um inteiro de um ao número de moléculas de ácido nucleicos intermediárias; e

(c) duas últimas moléculas de ácido nucleico, em que uma última molécula de ácido nucleico compreende, em uma orientação da extremidade 5' a 3', um primeiro local de restrição  $RA_m$ , uma sequência ligante anelada  $LA_m$ , um segmento de DNA  $D_{m1}$ , um segmento de ligação iniciador PB e um segundo local de restrição  $RB_m$ , em que outras última

molécula de ácido nucleico que compreende, em uma orientação da extremidade 5' a 3', um primeiro local de restrição  $RA_m$ , uma sequência ligante anelada  $LA_m$ , um segmento de DNA  $D_{m2}$ , um segmento de ligação iniciador PB e um segundo local de restrição  $RB_m$ , em que  $m$  representa um inteiro um maior do que o número de moléculas de ácido nucleicos intermediárias, em que o segmento de DNA  $D_{m1}$  codifica um primeiro segmento de um marcador selecionável, em que o segmento de DNA  $D_{m2}$  codifica um segundo segmento do marcador selecionável, em que o segmento de DNA  $D_{m2}$  é posicionado na orientação oposta ao segmento de DNA  $D_{m1}$  relativo ao sequência ligante anelada  $LA_m$  e segmento de ligação iniciador PB, em que nenhum segmento de DNA  $D_{m1}$  ou segmento de DNA  $D_{m2}$  produz um marcador selecionável funcional mas considerando recombinação homóloga dos segmentos de DNA  $D_{m1}$  e  $D_{m2}$  um marcador selecionável funcional é gerado, em que cada sequência ligante anelada  $LB_{(p-1)}$  é idêntica à sequência ligante anelada  $LA_p$ , em que  $p$  representa os inteiros de 1 a  $m$ .

[00171] Em algumas formas de realização, a biblioteca compreende uma pluralidade de vetores de montagem em que cada vetor de montagem compreende a mesma sequência ligante anelada, par da sequência ligante anelada ou sequência ligante anelada/par de segmento de ligação primária mas difere-se na sequência de seu fragmento de DNA  $D$  respectivo.

[00172] Em outras formas de realização, a biblioteca compreende uma pluralidade de vetores de montagem em que cada vetor de montagem compreende o mesmo segmento de DNA  $D$  flanqueado por uma sequência ligante anelada única, par da sequência ligante anelada ou sequência ligante anelada/par de segmento de ligação iniciador. Uma tal biblioteca pode servir para facilitar a montagem rápida do segmento de DNA  $D$  em uma posição ou orientação relativo a outros segmentos de DNA sendo montados no polinucleotídeo montado.

Em algumas formas de realização, os membros de uma

biblioteca compreendem os segmentos de DNA que tem formado as características estruturais ou funcionais. Por exemplo, uma biblioteca pode compreender uma pluralidade de vetores de montagem que compreende a mesma unidade de DNA funcional. As unidades de DNA funcionais exemplares incluem mas não são limitadas as sequências codificadoras de proteínas, gene repórter, marcadores fluorescentes, promotores, intensificadores, terminadores, íntrons, éxons, caudas poli-A, locais de clonagem múltiplos, sinal de localização nuclear, sinais de exportação nuclear, estabilização de mRNA, marcadores selecionáveis, locais de integração, etiquetas de epítipo e sinais de degradação. Em algumas formas de realização, a biblioteca compreende uma pluralidade de vetores de montagem em que cada vetor de montagem compreende a mesmo promotor. Os vetores de montagem podem compreender qualquer sequência promotora procariótica ou eucariótica conhecida na técnica. Os promotores eucarióticos exemplares incluem mas não são limitadas a um promotor de metalotioneína, um promotor tardio principal de adenovírus constitutivo, a promotor MMTV indutível por dexametasona, um promotor SV40, um promotor MRP pol III, um promotor MPSV constitutivo, um promotor RSV, um promotor CMV indutível por tetraciclina (tal como o promotor CMV precoce-imediato humano) e um promotor CMV constitutivo. Em formas de realização particulares, os vetores de montagem compreendem uma sequência promotora de levedura. Os promotores de levedura exemplares incluem mas não são limitados a PGAL3, PGAL7, PCTR3, PMET3, PPGK1, PTDH1, PTDH3, PFBA1, PTEF1, PENO1, PENO2, PCYC1, PTDH2, PCUP1, PGAL80, PGAL2, PBNA6, PTMA29, PSBP1, PPUP3, PACS2, PTPO1, PRPT1, PAAT2, PAHP1, PSSE1, PTEF2, PNPL3, PPET9, PTUB2, POLE1, PCPR1, PIPPP1 e PSOD1.

[00173] Em algumas formas de realização, a biblioteca compreende uma pluralidade de vetores de montagem em que cada vetor de montagem



compreende a mesma sequência terminadora. Os vetores de montagem podem compreender qualquer sequência terminadora procariótica ou eucariótica conhecida na técnica. Em formas de realização particulares, os vetores de montagem compreendem uma sequência terminadora de levedura. terminadores de levedura exemplares incluem mas não são limitados a TADH1, TENO1, TENO2, TCYC1, TNDT80, TTDH3, TTDH1 e TPGK1.

[00174] Em algumas formas de realização, a biblioteca compreende uma pluralidade de vetores de montagem em que cada vetor de montagem compreende o mesmo marcador selecionável. Os vetores de montagem podem compreender qualquer marcador selecionável procariótico ou eucariótico conhecido na técnica. Exemplos de marcadores selecionáveis incluem mas não são limitados a marcadores de resistência ao antibiótico (por exemplo, genes que codificam a resistência a canamicina, ampicilina, cloramfenicol, gentamicina ou trimetoprim) e marcadores metabólicos (por exemplo, genes de síntese de aminoácido ou genes de transferência de RNA).

### 5.9Kits

[00175] Em outro aspecto, fornecido neste é um kit para a montagem de um polinucleotídeo, o dito kit que compreende dois ou mais dos seguintes: (a) um ou mais vetores de entrada descritos neste; (b) um ou mais endonucleases de restrição capazes da clivagem dos locais de restrição RA e RB do dito um ou mais vetores de entrada; (c) um ou mais endonucleases de restrição capazes da clivagem dos locais de restrição RY e RZ dos ditos vetores de entrada; e (d) iniciadores de oligonucleotídeos capazes do anelamento aos segmentos de ligação iniciadores PA e PB do dito um ou mais vetores de entrada.

[00176] Em algumas formas de realização, locais de restrição RA e RB de cada vetor de entrada do kit são reconhecidos e clivados pela endonuclease de restrição SapI e o kit compreende endonuclease de restrição SapI. Em algumas formas de realização, locais de restrição RY e RZ de cada vetor de

entrada do kit são reconhecidos e clivados pela endonuclease de restrição SchI (ou MlyI) e o kit compreende endonuclease de restrição SchI (ou MlyI).

[00177] Em algumas formas de realização, a sequência de ácido nucleico do segmento de ligação iniciador PA de um ou mais vetores de entrada no kit é selecionado a partir do grupo que consiste de SEQ ID N°: 24 e 25. Em algumas formas de realização, a sequência de ácido nucleico do segmento de ligação iniciador PB um ou mais vetores de entrada no kit é selecionado a partir do grupo que consiste de SEQ ID N°: 24 e 25. Nas formas de realização preferidas, a sequência de ácidos nucleicos do segmento de ligação iniciador PA e segmento de ligação iniciador PB não são idênticos.

[00178] Em algumas formas de realização, a sequência nucleica da sequência ligante anelada LA de um ou mais vetores de entrada no kit é selecionado a partir do grupo que consiste de SEQ ID N°: 1 a 23. Em algumas formas de realização, a sequência nucleica da sequência ligante anelada LB um ou mais vetores de entrada no kit é selecionado a partir do grupo que consiste de SEQ ID N°: 1 a 23. Em algumas formas de realização, as sequências nucleicas da sequência ligante anelada LA e sequência ligante anelada LB de todos os vetores de entrada no kit são selecionados a partir do grupo que consiste de SEQ ID N°: 1 a 23.

[00179] Em algumas formas de realização, o kit compreende o vetor pRYSE #1, a sequência de que é fornecido neste como a SEQ ID N°: 221. Em algumas formas de realização, o kit compreende o vetor pRYSE #2, a sequência de que é fornecido neste como a SEQ ID N°: 207. Em algumas formas de realização, o kit compreende o vetor pRYSE #3, a sequência de que é fornecido neste como a SEQ ID N°: 208. Em algumas formas de realização, o kit compreende o vetor pRYSE #4, a sequência de que é fornecido neste como a SEQ ID N°: 209. Em algumas formas de realização, o kit compreende o vetor pRYSE #5, a sequência de que é fornecido neste como a SEQ ID N°: 210. Em algumas formas de realização, o kit compreende o vetor pRYSE #6,

a sequência de que é fornecido neste como a SEQ ID N°: 211. Em algumas formas de realização, o kit compreende o vetor pRYSE #7, a sequência de que é fornecido neste como a SEQ ID N°:212. Em algumas formas de realização, o kit compreende o vetor pRYSE #8, a sequência de que é fornecido neste como a SEQ ID N°: 213. Em algumas formas de realização, o kit compreende o vetor pRYSE #9, a sequência de que é fornecido neste como a SEQ ID N°: 214. Em algumas formas de realização, o kit compreende o vetor pRYSE #10, a sequência de que é fornecido neste como a SEQ ID N°: 215. Em algumas formas de realização, o kit compreende o vetor pRYSE #11, a sequência de que é fornecido neste como a SEQ ID N°: 216. Em algumas formas de realização, o kit compreende o vetor pRYSE #12, a sequência de que é fornecido neste como a SEQ ID N°: 217. Em algumas formas de realização, o kit compreende o vetor pRYSE #13, a sequência de que é fornecido neste como a SEQ ID N°: 218. Em algumas formas de realização, o kit compreende o vetor pRYSE #14, a sequência de que é fornecido neste como a SEQ ID N°:219. Em algumas formas de realização, o kit compreende o vetor pRYSE #15, a sequência de que é fornecido neste como a SEQ ID N°: 220.

[00180] Em algumas formas de realização, o kit compreende adicionalmente as instruções para o uso que descreve o método de montagem de polinucleotídeo divulgado neste. Em algumas formas de realização, a polimerase de polinucleotídeo, tal como a polimerase de DNA termoestável (por exemplo, polimerase Pfu de DNA) e trifosfatos de desoxiribonucleosídeo (dNTPs) também estão presentes no kit. Em algumas formas de realização, dois ou mais vetores de montagem cada um que compreende um polinucleotídeo componente a ser montado em um polinucleotídeo montado pode ser fornecido no kit. Por exemplo, vetores de montagem podem ser fornecidos que compreendem um polinucleotídeo componente útil para a calibração e/ou para uso como um controle positivo para verificar o

desempenho correto do kit. Outros exemplos incluem mas não são limitados a vetores de montagem que compreendem como um polinucleotídeo componente uma sequência codificadora de proteína, gene repórter, sequência codificadora marcadora fluorescente, promotor, intensificador, terminador, íntron, éxon, cauda poli-A, local de clonagem múltipla, sinal de localização nuclear, sinal de estabilização de mRNA, marcador selecionável, local de integração, sequência codificadora de etiqueta de epítipo e sinal de degradação.

## 6.EXEMPLOS

[00181] A invenção é ilustrada pelos seguintes exemplos, que não são pretendidos ser limitantes em qualquer maneira. As construções *Saccharyomices cerevisiae* descritas nos exemplos foram derivados da cepa *Saccharyomices cerevisiae* CEN.PK2. A cepa *Saccharyomices cerevisiae* S288c não semelhante, a sequência genômica da cepa CEN.PK2 não é publicamente disponível. Algumas das construções foram descritas foram verificadas pelas sequências e de modo que as sequências fornecidas são aquelas das construções derivadas CEN.PK2 atuais. Para as construções que não são verificadas pela sequência, as sequências fornecidas são baseadas na sequência genômica publicada da cepa S288c e deste modo podem incluir diferenças polimórficas às sequências das construções derivadas por CEN.PK2 atuais.

### Exemplo 1

[00182] Este exemplo descreve os métodos para a preparação dos vetores pRYSE. Os vetores pRYSE compreendem, em uma orientação da extremidade 5' a 3', um primeiro do local de reconhecimento da enzima de restrição SapI, uma primeira sequência ligante anelada ou segmento de ligação iniciadora, um primeiro local de reconhecimento da enzima de restrição SchI, uma proteína fluorescente verde (GFP) ou gene marcador lacZ, um segundo local de reconhecimento de enzima de restrição SchI, uma

segunda sequência ligante anelada ou segmento de ligação iniciador e um segundo local de reconhecimento da enzima de restrição SapI.

[00183] Um fragmento de DNA que codifica  $\beta$ -lactamase foi amplificado por PCR a partir do vetor pUC19 (Acessão Genbank L09137) usando iniciadores JCB158-17C (SEQ ID N°: 229) e JCB158-17D (SEQ ID N°: 230) após o local de reconhecimento da enzima de restrição SchI no gene bla de pUC 19 foram removidos pela mutagênese direcionada ao local de pUC 19 usando os iniciadores PCR JCB158-17A (SEQ ID N°: 227) e JCB158-17B (SEQ ID N°: 228). O produto PCR foi purificado por gel e então ligado no vetor TOPO (Invitrogen, Carlsbad, CA), de que este foi liberado novamente pela digestão da construção a finalização usando as enzimas de restrição SphI e MfeI, produzindo o "fragmento de DNA bla".

[00184] Os fragmentos de DNA 1040 (SEQ ID N°: 224), 1041 (SEQ ID N°: 225) e 1042 (SEQ ID N°: 226) foram gerados sinteticamente (Biosearch Technologies, Novato, CA). Os fragmentos de DNA 1040 e 1041 foram digeridos até a finalização usando enzima de restrição *BstXI* e cada fragmento digerido foi ligado com a cadeia principal do vetor 2,65 kb que foi gerado por corte até a finalização pAM1466 (SEQ ID N°: 223; gerado sinteticamente por Biosearch Technologies, Novato, CA) usando enzimas de restrição *SacI* e *KpnI*. A construção de DNA 1040\_pAM1466 foi digerido até a finalização usando enzimas de restrição *BsmBI* e *BstXI*, a mistura de reação foi determinada pela eletroforese em gel e um fragmento de DNA de aproximadamente 3,5 kb que compreende o fragmento de DNA 1040 foi purificado por gel. A construção de DNA 1041 pAM1466 foi digerida até a finalização usando as enzimas de restrição *BsaI* e *BstXI*, a mistura de reação foi determinada pela eletroforese em gel e um 1041 fragmento de DNA de aproximadamente 0,9 kb que compreende o fragmento de DNA 1041 que foi purificado por gel. Os fragmentos de DNA purificados foram ligados, produzindo construção de DNA 1040\_1041 pAM1466. O fragmento de DNA

1042 foi montado a construção de DNA 1040\_1041 por uma reação de "costura" de PCR usando os iniciadores J036 (SEQ ID N°: 69) e J037 (SEQ ID N°: 70) para gerar o fragmento de DNA 1040\_1041, iniciadores J038 (SEQ ID N°: 71) e J039 (SEQ ID N°: 72) para gerar o fragmento de DNA 1042 com uma sequência terminal que sobrepõe uma sequência terminal do fragmento de DNA 1040\_1041 e iniciadores J039 (contendo um local de reconhecimento da enzima de restrição *SphI*) (SEQ ID N°: 72) e J036 (contendo um local de reconhecimento da enzima de restrição *MfeI*) (SEQ ID N°: 69) para unir os dois produtos PCR. O produto 1040\_1041\_1042 PCR foi digerido até a finalização usando as enzimas de restrição *SphI* e *MfeI*, a mistura de reação foi determinada pela eletroforese em gel, o fragmento de DNA de aproximadamente 2,4 kb 1040\_1041\_1042 foi purificado por gel e o fragmento de DNA purificado foi ligado ao fragmento bla purificado por gel, produzindo a construção de DNA "1040 1041 1042 bla".

[00185] O segmento da construção de DNA 1040\_1041\_1042\_bla que codifica o gene GFP foi amplificado por PCR usando iniciadores de PCR 1 e 2 (ver Tabela 2). Ao fragmento terminal GFP amplificado dos locais de reconhecimento das enzimas de restrição *Sad* e *XhoI* foram adicionados pela amplificação de PCR usando como modelos os fragmentos GFP extraídos por gel gerados no primeiro ciclo de reações PCR e iniciadores PCR 3 e 4 (ver Tabela 2). Os produtos PCR amplificados foram extraídos em gel, então digeridos até a finalização usando as enzimas de restrição *XhoI* e *SacI*, as enzimas de restrição foram inativadas por calor por 20 minutos a 65°C e os produtos PCR digeridos foram purificados por coluna e então ligados com o fragmento de DNA de aproximadamente 2,2 kb purificado por gel que resulta a partir da digestão da construção de DNA 1040\_1041\_1042\_bla até a finalização usando as enzimas de restrição *XhoI* e *SacI*. Os vetores resultantes foram amplificados por PCR usando iniciadores PCR 5 e 6 (ver Tabela 3), as misturas de reação foram determinadas pela eletroforese em gel

e as "estruturas dos vetores pRYSE de aproximadamente 2,2 kb" foram purificadas por gel.

Tabela 2 – Iniciadores de PCR usados para gerar inserções de GFP flanqueadas por pares ligantes anelados ou ligantes anelados/pares de segmento de ligação iniciadora e locais de enzima de restrição SacI e XhoI

Fragmento GFP	Ligante anelado ou Segmento 2 de ligação iniciadora	Ligante anelado ou Segmento 2 de ligação iniciadora	Iniciador 1	Iniciador 2	Iniciador 3	Iniciador 4
1	PmeI-5'	RYSE 1	J018 (SEQ ID N°: 73)	J073 (SEQ ID N°:106)	J055 (SEQ ID N°: 88)	J064 (SEQ ID N°: 97)
2	RYSE 1	RYSE 2	J019 (SEQ ID N°: 74)	J074 (SEQ ID N°:107)	J056 (SEQ ID N°: 89)	J065 (SEQ ID N°: 98)
3	RYSE 2	RYSE 3	J020 (SEQ ID N°: 75)	J029 (SEQ ID N°:82)	J057 (SEQ ID N°: 90)	J066 (SEQ ID N°: 99)
4	RYSE 3	RYSE 4	J021 (SEQ ID N°: 76)	J030 (SEQ ID N°:83)	J058 (SEQ ID N°: 91)	J067 (SEQ ID N°: 100)
5	RYSE 4	RYSE 5	J022 (SEQ ID N°: 77)	J031 (SEQ ID N°:84)	J059 (SEQ ID N°: 92)	J068 (SEQ ID N°: 101)
6	RYSE 5	RYSE 6	J023 (SEQ ID N°: 78)	J032 (SEQ ID N°:85)	J060 (SEQ ID N°: 93)	J069 (SEQ ID N°: 102)
7	RYSE 6	RYSE 7	J024 (SEQ ID N°: 79)	J033 (SEQ ID N°:86)	J061 (SEQ ID N°: 94)	J070 (SEQ ID N°: 103)
8	RYSE 7	RYSE 8	J025 (SEQ ID N°: 80)	J034 (SEQ ID N°:87)	J062 (SEQ ID N°: 95)	J071 (SEQ ID N°: 104)
9	RYSE 2	PmeI-3'	J020 (SEQ ID N°: 75)	J075 (SEQ ID N°:108)	J057 (SEQ ID N°: 90)	J072 (SEQ ID N°: 105)
10	RYSE 3	PmeI-3'	J021 (SEQ ID N°: 76)	J075 (SEQ ID N°:108)	J058 (SEQ ID N°: 91)	J072 (SEQ ID N°: 105)
11	RYSE 4	PmeI-3'	J022 (SEQ ID N°: 77)	J075 (SEQ ID N°:108)	J059 (SEQ ID N°: 92)	J072 (SEQ ID N°: 105)
12	RYSE 5	PmeI-3'	J023 (SEQ ID N°: 78)	J075 (SEQ ID N°:108)	J060 (SEQ ID N°: 93)	J072 (SEQ ID N°: 105)
13	RYSE 6	PmeI-3'	J024 (SEQ ID N°: 79)	J075 (SEQ ID N°:108)	J061 (SEQ ID N°: 94)	J072 (SEQ ID N°: 105)

14	RYSE 7	Pmel-3'	J025 (SEQ ID Nº: 80)	J075 (SEQ ID Nº:108)	J062 (SEQ ID Nº: 95)	J072 (SEQ ID Nº: 105)
15	RYSE 8	Pmel-3'	J026 (SEQ ID Nº: 81)	J075 (SEQ ID Nº: 108)	J063 (SEQ ID Nº: 96)	J072 (SEQ ID Nº: 105)

Tabela 3 – Pares da sequência ligante anelada ou sequência ligante anelada/pares de segmento da ligação iniciadora presente nos vetores pRYSE e iniciadores de PCR usados para gerar as cadeias principais do vetor pRYSE

Vetor pRYSE	Ligante anelado ou Segmento 1 de ligação iniciadora (ver Tabela1)	Ligante anelado ou Segmento 2 de ligação iniciadora (ver Tabela1)	Iniciador 5	Iniciador 6
1	Pmel-5'	RYSE 1	S001 (SEQ ID Nº:46)	S002 (SEQ ID Nº:47)
2	RYSE 1	RYSE 2	S003 (SEQ ID Nº:48)	S004 (SEQ ID Nº:49)
3	RYSE 2	RYSE 3	S005 (SEQ ID Nº:50)	S006 (SEQ ID Nº:51)
4	RYSE 3	RYSE 4	S007 (SEQ ID Nº:52)	S008 (SEQ ID Nº:53)
5	RYSE 4	RYSE 5	S009 (SEQ ID Nº:54)	S010 (SEQ ID Nº:55)
6	RYSE 5	RYSE 6	S011 (SEQ ID Nº:56)	S012 (SEQ ID Nº:57)
7	RYSE 6	RYSE 7	S013 (SEQ ID Nº: 58)	S014 (SEQ ID Nº: 59)
8	RYSE 7	RYSE 8	S015 (SEQ ID Nº:60)	S016 (SEQ ID Nº:61)
9	RYSE 2	Pmel-3'	S005 (SEQ ID Nº:50)	S018 (SEQ ID Nº:63)
10	RYSE 3	Pmel-3'	S007 (SEQ ID Nº:52)	S018 (SEQ ID Nº:63)
11	RYSE 4	Pmel-3'	S009 (SEQ ID Nº:54)	S018 (SEQ ID Nº: 63)
12	RYSE 5	Pmel-3'	S011 (SEQ ID Nº:56)	S018 (SEQ ID Nº:63)
13	RYSE 6	Pmel-3'	S013 (SEQ ID Nº:58)	S018 (SEQ ID Nº:63)
14	RYSE 7	Pmel-3'	S015 (SEQ ID Nº:60)	S018 (SEQ ID Nº:63)
15	RYSE 8	Pmel-3'	S017 (SEQ ID Nº:62)	S018 (SEQ ID Nº:63)

[00186] O gene lacZ foi amplificado por PCR a partir do vetor pUC19 usando iniciadores S027 (SEQ ID Nº: 65) e S028 (SEQ ID Nº: 66), que cada um compreende um local de reconhecimento da enzima de restrição SchI. A mistura de reação foi determinada pela eletroforese em gel, do produto PCR de aproximadamente 0,5 kb PCR foi purificado por gel e o produto PCR



purificado foi ligado com cada uma das cadeias principais do vetor pRYSE. A mutagênese direcionada ao local foi realizada nos vetores resultantes usando os iniciadores PCR L012 (SEQ ID N°: 231) e L013 (SEQ ID N°: 232) para remover um local de reconhecimento da enzima de restrição SchI a partir da origem da replicação. Finalmente, uma segunda mutagênese direcionada ao local foi realizada usando iniciadores PCR S036 (SEQ ID N°: 67) e S037 (SEQ ID N°: 68) para remover o local de reconhecimento da enzima de restrição SchI a partir do fragmento *lacZ*, deste modo produzindo os vetores pRYSE 1 através de 15 (ver FIG. 4 para um mapa de plasmídeo dos vetores pRYSE e SEQ ID N°: 207 através de 221 para uma sequência de nucleotídeo dos vetores pRYSE 1 através de 15).

### Exemplo 2

[00187] Este exemplo descreve métodos alternativos para a preparação dos vetores pRYSE.

[00188] Os vetores pRYSE 1 através de 15 podem ser gerados sinteticamente usando como modelo a SEQ ID N°: 207 através de 221 (por exemplo, por Biosearch Technologies, Novato, CA). os vetores pRYSE adicionais que compreendem as sequências ligantes aneladas diferentes podem ser geradas sinteticamente usando como modelo a SEQ ID N°: 221 em que o segmento de ligação iniciador Pmel-5' e/ou a sequência ligante anelada RYSE 1 são mudados a outra sequência ligante anelada adequada ou segmento de ligação iniciador (ver Tabela 1).

### Exemplo 3

[00189] Este exemplo descreve os métodos para a preparação de um vetor pMULE, que compreende, em uma orientação da extremidade 5' a 3', um primeiro local de reconhecimento da enzima de restrição SapI, um primeiro local de reconhecimento da enzima de restrição SchI, um gene marcador *lacZ*, um segundo local de reconhecimento da enzima de restrição SchI e um segundo local de reconhecimento da enzima de restrição SapI. O

vetor pMULE pode ser usado para clonar Mules.

[00190] A estrutura dos vetores pRYSE 8 foi amplificada por PCR usando os iniciadores K162 (SEQ ID N°: 109) e K163 (SEQ ID N°: 110). A mistura de reação foi determinada pela eletroforese em gel e a cadeia principal do vetor de aproximadamente 2,2 kb foi purificado por gel. Um fragmento de DNA que compreende o gene *lacZ* foi gerado pela digestão até a finalização do vetor pRYSE 8 usando a enzima de restrição *SchI*, inativando por calor a enzima a 65°C por 20 minutos, determinado uma mistura de reação pela eletroforese em gel e purificação em gel do fragmento de DNA de aproximadamente 0,5 kb. O fragmento de DNA purificado que compreende o gene *lacZ* foi ligado com a cadeia principal do vetor purificado, produzindo o vetor pMULE (ver FIG. 7 para um mapa de plasmídeo).

#### Exemplo 4

[00191] Este exemplo descreve os métodos para a preparação de "Pedaços". Os pedaços são fragmentos de DNA que podem ser inseridos nos vetores pRYSE para gerar os vetores de montagem que compreendem os polinucleotídeos componentes que podem ser montados nos polinucleotídeos montados usando os métodos divulgados neste. Os pedaços podem codificar os genes ou elementos de interesse genético (por exemplo, promotores, terminadores, marcadores selecionáveis, local de integração, etiqueta de epítipo, sinais de localização, sinais de degradação, marcadores fluorescentes, locais de clonagem múltiplos). Os pedaços foram amplificados por PCR a partir de um modelo usando os iniciadores como descrito na Tabela 4.

Tabela 4 – Pedaços amplificados				
Pedaço	Tipo	Iniciadores	Tamanho	Modelo
atoB	Gs	L229 (SEQ ID N°: 40) L230 (SEQ ID N°: 41)	1185	DNA de plasmídeo que compreende o gene <i>atoB</i> de <i>Escherichia coli</i> (número de acesso Genbank NC 000913 REGIÃO: 2324131..2325315)
mvaS	Gs	L235 (SEQ ID N°: 42) L236 (SEQ ID N°: 43)	1152	fragmento de DNA sintético que compreende o gene <i>mvaS</i> de <i>Enterococcus faecalis</i> (Número de acesso Genbank AF290092 REGIÃO: 142..1293)

				códon otimizado pela expressão em <i>Saccharomyces cerevisiae</i> e que compreende na posição 110 uma modificação de alanina a glicina para aumentar a atividade de enzima (ver Steussy <i>et al.</i> (2006) <i>Biochemistry</i> 45(48):14407-14414)
ERG13-1	GsT	L109 (SEQ ID N°: 235) L110 (SEQ ID N°: 26)	1726	Cepa <i>Saccharomyces cerevisiae</i> CEN.PK2 DNA genômico
3' NDT80	D	L221 (SEQ ID N°: 34) L222 (SEQ ID N°: 35)	516	Cepa <i>Saccharomyces cerevisiae</i> CEN.PK2 DNA genômico
5' NDT80	U	L219 (SEQ ID N°: 32) L220 (SEQ ID N°: 33)	495	Cepa <i>Saccharomyces cerevisiae</i> CEN.PK2 DNA genômico
tP <sub>FBA1</sub>	P	L225 (SEQ ID N°: 37) L057 (SEQ ID N°: 234)	526	Cepa <i>Saccharomyces cerevisiae</i> CEN.PK2 DNA genômico
tP <sub>TDH3</sub>	P	L224 (SEQ ID N°: 36) L054 (SEQ ID N°: 233)	559	Cepa <i>Saccharomyces cerevisiae</i> CEN.PK2 DNA genômico
ERG10-1	Gs	L226 (SEQ ID N°: 38) L227 (SEQ ID N°: 39)	1182	fragmento sintetizado que codifica o acetil-CoA acetiltransferase de <i>Ralstonia eutropha</i> (acessão GenBank NC 008313 REGIÃO: 183291..184469) códon otimizado pela expressão em <i>Saccharomyces cerevisiae</i> e seguido por um códon de interrupção adicional
tEN01	T	L248 (SEQ ID N°: 44) L176 (SEQ ID N°: 27)	265	Cepa <i>Saccharomyces cerevisiae</i> CEN.PK2 DNA genômico
tTDH3	T	L185 (SEQ ID N°: 28) L186 (SEQ ID N°: 29)	260	Cepa <i>Saccharomyces cerevisiae</i> CEN.PK2 DNA genômico
HphA	M	TRIX_L_193 (SEQ ID N°: 184) TRIX_L_194 (SEQ ID N°: 185)	1912	DNA de plasmídeo que compreende o promotor TEFL e terminador de <i>Kluyveromyces lactis</i> (acessão GenBank CR382122 REGIÕES:788874..789380 e 787141..787496, respectivamente) e o gene hph de <i>Klebsiella pneumonia</i>
tHMG1	GsT	TRIX_L_232 (SEQ ID N°: 186)	1742	Cepa <i>Saccharomyces cerevisiae</i> CEN.PK2 DNA genômico
		TRIX_L_233 (SEQ ID N°: 187)		
tP <sub>GAL1-10</sub>	P	TRIX_L266 (SEQ ID N°: 190) TRIX_L_267 (SEQ ID N°: 191)	620	Cepa <i>Saccharomyces cerevisiae</i> CEN.PK2 DNA genômico
ERG10-2	GsT	TRIX_L_106 (SEQ ID N°: 170) TRIX_L_107 (SEQ ID N°: 171)	1467	Cepa <i>Saccharomyces cerevisiae</i> CEN.PK2 DNA genômico
ERG13-2	GsT	TRIX_L_109 (SEQ ID N°: 172)	1726	Cepa <i>Saccharomyces cerevisiae</i> CEN.PK2 DNA genômico

		TRIX_L_110 (SEQ ID N°: 173)		
GAL80US	U	JU-218-168-130-GAL80US-F (SEQ ID N°: 134) JU-219-168-130-GAL80US-R (SEQ ID N°: 135)	500	Cepa <i>Saccharomyces cerevisiae</i> CEN.PK2 DNA genômico
GAL80DS	D	JU-220-168-130-GAL80DS-F (SEQ ID N°: 136) JU-221-168-130-GAL80DS-R (SEQ ID N°: 137)	500	Cepa <i>Saccharomyces cerevisiae</i> CEN.PK2 DNA genômico
P <sub>TDH3</sub>	P	L224 (SEQ ID N°: 36) TRIX_L_053 (SEQ ID N°: 169)	583	Cepa <i>Saccharomyces cerevisiae</i> CEN.PK2 DNA genômico
NatA	M	TRIX_L_193 (SEQ ID N°: 184) TRIX_L_194 (SEQ ID N°: 185)	1456	DNA de plasmídeo que compreende o promotor TEF 1 e terminador de <i>Kluyveromyces lactis</i> (acessão GenBank CR382122 REGIÕES:788874..789380 e 787141..787496, respectivamente) e o gene natl. de <i>S. noursei</i>
ERG12	GsT	TRIX_L_112 (SEQ ID N°: 174) TRIX_L_113 (SEQ ID N°: 175)	1582	Cepa <i>Saccharomyces cerevisiae</i> CEN.PK2 DNA genômico
ERG8	GsT	TRIX_L_118 (SEQ ID N°: 178) TRIX_L_119 (SEQ ID N°: 179)	1616	Cepa <i>Saccharomyces cerevisiae</i> CEN.PK2 DNA genômico
P <sub>GAL4OC</sub>	P	TRIX_K_131 (SEQ ID N°: 165) PW-91-093-CPK422-G (SEQ ID N°: 162)	270	DNA de plasmídeo que compreende uma versão "constitutiva operativa" do promotor do gene GAL4 da cepa <i>Saccharomyces cerevisiae</i> CEN.PK2 (Griggs & Johnston (1991) <i>PNAS</i> 88(19):8597-8601)
GAL4-1	G	JU-286-275-31-GAL4-F (SEQ ID N°: 140) JU-285-275-31-GAL4-FIX-R2 (SEQ ID N°: 139)	526	Cepa <i>Saccharomyces cerevisiae</i> CEN.PK2 DNA genômico
GAL4-2	G	JU-284-275-31-GAL4-FIX-F2 (SEQ ID N°: 138) JU-287-275-31-GAL4-R (SEQ ID N°: 141)	2414	Cepa <i>Saccharomyces cerevisiae</i> CEN.PK2 DNA genômico
KanA	M	TRIX_L_193 (SEQ ID N°: 184) TRIX_L_194 (SEQ ID N°: 185)	1696	DNA de plasmídeo que compreende o promotor TEF1 e terminador de <i>Kluyveromyces lactis</i> (acessão GenBank CR382122)

				REGIÕES:788874..789380 e 787141..787496, respectivamente) e o gene kanR do transposon Tn903
ERG 19	GsT	TRIX_L_115 (SEQ ID N°: 176) TRIX_L_116 (SEQ ID N°: 177)	1441	Cepa <i>Saccharomyces cerevisiae</i> CEN.PK2 DNA genômico
ERG20	GsT	TRIX_L_124 (SEQ ID N°: 182) TRIX_L_125 (SEQ ID N°: 183)	1319	Cepa <i>Saccharomyces cerevisiae</i> CEN.PK2 DNA genômico
PCAL7	P	TRIX_L_34 (SEQ ID N°: 166) TRIXL35 (SEQ ID N°: 167)	500	Cepa <i>Saccharomyces cerevisiae</i> CEN.PK2 DNA genômico
tPcAL7	P	TRIX_L_34 (SEQ ID N°: 166) TRIX_L_36 (SEQ ID N°: 168)	476	Cepa <i>Saccharomyces cerevisiae</i> CEN.PK2 DNA genômico
IDI1	GsT	TRIX_L_121 (SEQ ID N°: 180) TRIX_L_122 (SEQ ID N°: 181)	1127	Cepa <i>Saccharomyces cerevisiae</i> CEN.PK2 DNA genômico
tP <sub>CTR3</sub>	P	TRIX_K_0142 (SEQ ID N°: 163) TRIX_K_0143 (SEQ ID N°: 164)	710	DNA de plasmídeo que compreende o promotor do gene CTR3 <i>Saccharomyces da cepa cerevisiae</i> CEN.PK2
LEU2US	U	JU-164-168-110-LEU2 US-f (SEQ ID N°: 129) JU-165-168-110-LEU2 US-r (SEQ ID N°: 130)	500	Cepa <i>Saccharomyces cerevisiae</i> CEN.PK2 DNA genômico
LEU2DS	D	JU-162-168-110-LEU2 DS-f (SEQ ID N°: 127) JU-163-168-110-LEU2 DS-r (SEQ ID N°: 128)	500	Cepa <i>Saccharomyces cerevisiae</i> CEN.PK2 DNA genômico
ERG9US	U	JU-108-168-110-ERG9 US-f (SEQ ID N°: 126) JU-172-168-110-ERG9 US-rl (SEQ ID N°: 133)	499	Cepa <i>Saccharomyces cerevisiae</i> CEN.PK2 DNA genômico
ERG9CDS	G	JU-106-168-110-ERG9 CDS-f (SEQ ID N°: 124) JU-107-168-110-ERG9 CDS-r (SEQ ID N°: 125)	501	Cepa <i>Saccharomyces cerevisiae</i> CEN.PK2 DNA genômico
STE5US	U	TRIXRN017 (SEQ ID N°: 192)	600	Cepa <i>Saccharomyces cerevisiae</i> CEN.PK2 DNA genômico

		TRIX_RN018 (SEQ ID N°: 193)		
STE5DS	D	TRIX_RN019 (SEQ ID N°: 194) TRIX_RN020 (SEQ ID N°: 195)	600	Cepa <i>Saccharomyces cerevisiae</i> CEN.PK2 DNA genômico
URA3	M	JU-169-168-110-URA3-f (SEQ ID N°: 131) JU-170-168-110-URA3-r (SEQ ID N°: 132)	1554	Cepa <i>Saccharomyces cerevisiae</i> CEN.PK2 DNA genômico
* G=gene; s=códon de interrupção; T=terminador; M=marcador; D=região de integração a jusante; U= região de integração a montante; P=promotor.				

[00192] As amplificações de PCR foram feitas usando a polimerase de DNA Phusion (New England Biolabs, Ipswich, MA) como pelos protocolos sugeridos pelos fabricantes. As reações de PCR foram determinadas pela eletroforese em gel, os pedaços foram purificados em gel e os pedaços purificados foram tratados com T4 polinucleotídeo cinase (PNK) (New England Biolabs, Ipswich, MA) como pelo protocolo sugerido pelos fabricantes. O PNK foi inativado por calor a 65°C por 20 minutos e as amostras foram armazenadas a -20°C.

#### Exemplo 5

[00193] Este modelo descreve os métodos para a preparação de "MULEs." Os MULEs são fragmentos de DNA que podem ser inseridos nos vetores pMULE para gerar os vetores de montagem que compreendem os polinucleotídeos componentes que podem ser montados nos polinucleotídeos montados usando os métodos divulgados neste. Os MULEs podem codificar os genes ou elementos genéticos de interesse (por exemplo, promotores, terminadores, marcadores selecionáveis, local de integração, etiqueta de epítopos, sinais de localização, sinais de degradação, marcadores fluorescentes, locais múltiplos de clonagem) flanqueados pelos pares da sequência ligante anelada ou sequência ligante anelada/iniciadores de par de segmento de ligação. Os MULEs foram amplificados por PCR a partir do modelo usando iniciadores no qual a extremidade 3' anela a sequência alvo e a extremidade 5' compreende uma sequência ligante anelada ou um segmento de ligação iniciador (ver Tabela 1 para as sequências ligantes aneladas

adequadas), como descrito na Tabela 5.

Tabela 5 – MULEs amplificados				
MULE	Tipo	Iniciadores	Tamanho (bp)	Modelo
tHMG1-a	G	KMH8-276-1-ligante4.tHMG1.fwd (SEQ ID N°: 157) KMH9-276-1-ligante9.tHMG 1.rev (SEQ ID N°: 160)	1794	RABit 254 DNA de plasmídeo
ERG12	G	KMH46-276-43-ERG 12ligante4.fwd (SEQ ID N°: 151) KMH14-276-4-ligante9.ERG 12.rev (SEQ ID N°: 145)	1634	RABit 250 DNA de plasmídeo
ERG19	G	KMH47-276-43-ERG 19ligante4.fwd (SEQ ID N°: 152) KMH15-276-4-ligante9. ERG 19.rev (SEQ ID N°: 146)	1493	RABit 241 DNA de plasmídeo
P <sub>TDH3</sub> -a	P	KMH81-276-116-TDH3.rev.tHMG1 (SEQ ID N°: 155) S004 (SEQ ID N°: 49)	626	RABit 54 DNA de plasmídeo
P <sub>TDH3</sub> -b	P	KMH9 1 -276-1 16-TDH3.rev.FS (SEQ ID N°: 158) S004 (SEQ ID N°: 49)	546	RABit 54 DNA de plasmídeo
tHMG1-b	G	KMH82-276-116-tHMG1.fwd.TDH3 (SEQ ID N°: 156) S009 (SEQ ID N°: 54)	1801	RABit 20 DNA de plasmídeo
IME 1 US	U	KB454-266-53 (SEQ ID N°: 142) KB455-266-53 (SEQ ID N°: 143)	578	Cepa <i>Saccharomyces cerevisiae</i> CEN.PK2 DNA genômico
IME 1 DS	D	KMH93-276-130-3' IME.ligante4.fwd (SEQ ID N°: 161) KB457-266-53 (SEQ ID N°: 144)	554	Cepa <i>Saccharomyces cerevisiae</i> CEN.PK2 DNA genômico
LEU2	M	VH296-235-55-Leu2 12-1 F (SEQ ID N°: 30) VH296-235-55-Leu2 12-1 R (SEQ ID N°: 31)	1795	DNA de plasmídeo que compreende LEU2 Local da cepa <i>Saccharomyces cerevisiae</i> CEN.PK2 (Sikorski RS, Hieter (1989) Genetics 122(1):19-27)
FS-a	G	KMH5-276-1-Ligante3.FS(Kozak).fwd (SEQ ID N°: 153) KMH7-276-1-linker4.TCYC1.rev (SEQ ID N°: 154)	1981	DNA de plasmídeo que compreende a sequência codificadora da sintase farneseno de <i>Artemisia annua</i> (número de acesso GenBank AY835398) códon otimizado pela expressão em <i>Saccharomyces cerevisiae</i> e terminador da cepa do gene CYC1 <i>Saccharomyces cerevisiae</i> CEN.PK2
FS-b	G	KMH92-276-116-FS.fwd.TDH3 (SEQ ID N°: 159) KMH7-276-1-linker4.TCYC1.rev (SEQ ID N°: 154)	1976	DNA de plasmídeo que compreende a sequência codificadora da sintase de farneseno de <i>Artemisia annua</i> (número de acesso GenBank AY835398) códon otimizado pela expressão em <i>Saccharomyces cerevisiae</i> e terminador da cepa do gene

				CYC1 <i>Saccharomyces cerevisiae</i> CEN.PK2
URA3b Último	M	VH228-235-7- URA3LOF3RYSE 12-1F (SEQ ID N°: 204) VH229-235-7- URA3LOF3RYSE12-1R (SEQ ID N°: 205)	1565	Modelo URA-3 blaster **

\* G=gene; s=códon de interrupção; T=terminador; M=marcador; D= região de integração a jusante; U= região de integração a montante; P = promotor.

\*\* O modelo URA-3 blaster foi feito pela primeira geração dos fragmentos de DNA que flanqueia a sequência A (gerada a partir do fragmento de DNA sintético que compreende SEQ ID N°: 206 usando iniciadores PCR TRI\_X\_Z025 (SEQ ID N°: 196) e TRI\_X\_Z026 (SEQ ID N°: 197)), flanqueando a sequência B (gerada a partir do fragmento de DNA sintético que compreende SEQ ID N°: 206 usando iniciadores PCR TRI\_X\_Z027 (SEQ ID N°: 198) e TRI\_X\_Z028 (SEQ ID N°: 199)), URA3-c (gerado a partir da cepa *Saccharomyces cerevisiae* CEN.PK2 DNA genômico usando iniciadores PCR TRI\_X\_Z033 (SEQ ID N°: 200) e TRI\_X\_Z036 (SEQ ID N°: 203)) e URA3-d (gerado da cepa *Saccharomyces cerevisiae* CEN.PK2 DNA genômico usando iniciadores PCR TRI\_X\_Z034 (SEQ ID N°: 201) e TRI\_X\_Z035 (SEQ ID N°: 202)). Os fragmentos de DNA que flanqueiam a sequência A, URA3-c e URA3-d foram então costurados juntos no fragmento de DNA A usando os iniciadores de PCR TRI\_X\_Z025 e TRI\_X\_Z034 e fragmentos de DNA URA3-c, URA3-d e flanqueiam a sequência B foram costurados juntos no fragmento de DNA B usando iniciadores de PCR TRI\_X\_Z028 e TRI\_X\_Z033. Finalmente, os fragmentos de DNA A e B foram costurados juntos usando iniciadores de PCR TRI\_X\_Z025 e TRI\_X\_Z028, produzindo o modelo URA-3 blaster.

[00194] As amplificações de PCR foram feitas usando a polimerase de DNA Phusion (New England Biolabs, Ipswich, MA) como pelo protocolo sugeridos pelos fabricantes. As reações de PCR foram determinadas pela eletroforese em gel, os MULEs foram purificados em gel e os MULEs purificados foram tratados com polinucleotídeo cinase T4 (PNK) (New England Biolabs, Ipswich, MA) como pelo protocolo sugeridos pelos fabricantes. O PNK foi inativado por calor a 65°C por 20 minutos e as amostras foram armazenadas a -20°C.

#### Exemplo 6

[00195] Este modelo descreve os métodos para inserção dos pedaços nos vetores pRYSE ou MULEs no vetor pMULE para gerar os vetores de montagem.

[00196] Os vetores pRYSE 1 através do vetor 8 e pRYSE 15 foram digeridos até a finalização usando enzima de restrição SchI e os fragmentos de DNA digeridos foram tratados com Fosfatase Antarctic (New England Biolabs, Ipswich, MA). A fosfatase foi inativada por calor a 65°C por 20 minutos, as misturas de reação foram determinadas pela eletroforese em gel e as cadeias principais do vetor de aproximadamente 2,2 kb pRYSE



(desprovidos de *lacZ*) foram purificados em gel. As cadeias principais do vetor pRYSE purificadas foram ligadas com pedaços como detalhados na Tabela 6, deste modo produzindo os vetores de montagem.

[00197] O vetor pMULE é digerido até a finalização usando enzima de restrição *SchI*, a mistura de reação é determinada pela eletroforese em gel e cadeia principal do vetor de aproximadamente 2,2 kb pMULE (desprovidos de *lacZ*) é purificada por gel. A cadeia principal do vetor pMULE purificada é tratada com uma fosfatase (por exemplo, Fosfatase Antarctic (New England Biolabs, Ipswich, MA), CLAP (New England Biolabs, Ipswich, MA), SAP (New England Biolabs, Ipswich, MA; Fermentas, Glen Burnie, MD) ou FastAP (Fermentas, Glen Burnie, MD)), a fosfatase é inativada por calor (por exemplo, 20 minutos a 65°C) e a cadeia principal do vetor pMULE é ligado com MULEs, deste modo produzindo os vetores de montagem.

Tabela 6 - Vetores de montagem gerados		
Pedaco (ver Tabela 4)	Vetor pRYSE (ver Tabela 3)	Vetor de montagem
atoB	4	2
mvaS	7	5
ERG13-1	7	12
3' NDT80	15	29
	10	24
5' NDT80	1	30
	1	97
tP <sub>FBA1</sub>	6	35
tP <sub>TDH3</sub>	3	53
ERG 10-1	4	60
tENO 1	8	62
tTDH3	5	64
GAL80US	1	270
HphA	2	22
tHMG1	3	254
tP <sub>GAL1,10</sub>	4	229
ERG 10-2	5	244
ERG13-2	6	253
tPGAL 1,10	7	228
tHMG1	8	255
GAL80DS	15	271
LEU2US	1	187

NatA	2	262
ERG12	3	250
ERG8	5	252
P <sub>GAL4OC</sub>	6	268
GAL4 *	7	265
LEU2DS	14	263
ERG9US	1	186
KanA	2	261
ERG19	3	241
ERG20	5	251
tP <sub>GAL7</sub>	6	249
ID11	7	237
tP <sub>CTR3</sub>	8	269
ERG9CDS	15	185
P <sub>GAL7</sub>	3	44
STE5US	1	567
URA3	2	556 (orientação 1) 555 (orientação 2)
P <sub>TDH3</sub>	3	54
tHMG1	4	20
STE5DS	11	563

As ligações foram realizadas usando 50 ng da cadeia principal do vetor, 3 excessos molares de Bit e a ligase (por exemplo, Ligase rápida (New England Biolabs, Ipswich, MA), T4 DNA ligase (concentrações regulares e altas; vendor, Fermentas, Glen Burnie, MD ), Ligase rápida (Fermentas, Glen Burnie, MD)) como pelo protocolo sugeridos pelos fabricantes.

\* Peça GAL4 foi gerado pelo corte junto aos pedaços GAL4-1 e GAL4-2 (ver Tabela 4) usando iniciadores JU-286-275-31-GAL4-F (SEQ ID N°: 140) e JU-287-275-31-GAL4-R (SEQ ID N°: 141).

[00198] Os vetores de montagem foram transformados em células precursoras TOP 10 *Escherichia coli* quimicamente competentes (Invitrogen, Carlsbad, CA). Os transformantes da célula hospedeira foram selecionadas em ágar Luria Bertoni (LB) contendo 100 ug/mL de carbenicilina e 40 ug/mL X-gal. As colônias brancas simples foram transferidas a partir de ágar LB para cultivar os tubos contendo 5 mL do meio líquido LB e carbenicilina e as culturas foram incubadas durante a noite a 37°C em um agitador rotativo a 250 rpm. Os DNAs de plasmídeo foram extraídos e sequenciados para identificar os clones contendo uma sequência correta na orientação correta. As células foram armazenadas a -80°C em criofrascos em de alíquotas de choque de 1 mL feitas até 400 uL de glicerol 50%

estéril e 600 uL de cultura líquida.

### Exemplo 7

[00199] Este modelo descreve os métodos para os polinucleotídeos componentes de montagem em um polinucleotídeo montado usando os vetores de montagem e/ou MULEs.

[00200] Os vetores de montagem (ver Tabela 7) foram colocados juntos em um tubo (333 fmol de cada RABit) e digerido usando enzima de restrição LguI (Fermentas, Glen Burnie, MD). A enzima de restrição foi removida pela centrifugação de coluna ou inativada por calor por 20 minutos a 65°C. Para a reação de montagem que envolve os MULEs ou polinucleotídeos montados, 333 fmol de cada MULE ou polinucleotídeo montado (ver Tabela 7) foram colocados juntos em um tubo ou foram adicionados aos vetores de montagem digeridos. As amostras foram unidas nas três reações de 30 uL; água, tampão, dNTPs e polimerase de DNA foram adicionados a cada mistura de reação e um primeiro ciclo de amplificação de PCR foi iniciado. As amostras foram colocadas em gel, 0,5 uM de cada iniciador terminal (Tabela 7) foram adicionados as misturas de reação e um segundo ciclo de amplificação de PCR foi realizada. As três misturas de reação de PCR foram combinadas em um tubo, as misturas de reação foram determinadas pela eletroforese em gel e os produtos PCRs foram purificados em gel.

Tabela 7 - Iniciadores terminais para a montagem de polinucleotídeos montados				
Montagem	Vetores de montagem (ver Tabela ou MULEs (ver Tabela 5) a ser combinada*	Polinucleotídeo montado Tamanho (kb) (sequência)	Iniciador terminal 1	Iniciador terminal 2
1	3 0 2 2 5 3 6 0	4,3	S000 (SEQ ID N°: 45)	S009 (SEQ ID N°: 54)
2	30_22_53	3,1	5000 (SEQ ID N°: 45)	S007 (SEQ ID N°: 52)
3	22_53_60	3,7	S002 (SEQ ID N°: 47)	S009 (SEQ ID N°: 54)
4	3022	2,5	S000 (SEQ ID N°: 45)	S005 (SEQ ID N°: 50)
5	22_53	2,5	S002	S007

			(SEQ ID N°: 47)	(SEQ ID N°: 52)
6	53_60	1,8	S004 (SEQ ID N°: 49)	S009 (SEQ ID N°: 54)
7	302253_60_6435_12_62_29	7,7 (SEQ ID N°: 222)	S000 (SEQ ID N°: 45)	S019 (SEQ ID N°: 64)
8	30_22_53_60_64_35_5_62_29	7,1	S000 (SEQ ID N°: 45)	S019 (SEQ ID N°: 64)
9	30_22_53_2_64_35_562_29	7,1	S000 (SEQ ID N°: 45)	S019 (SEQ ID N°: 64)
10	60_64_35_5_62_29	4,1	S006 (SEQ ID N°: 51)	S019 (SEQ ID N°: 64)
11	2_64_35_5_62_29	4,1	S006 (SEQ ID N°: 51)	S019 (SEQ ID N°: 64)
Fase I-A	270_22_254_229_244_253	8,1 (SEQ ID N°: 111)	S000 (SEQ ID N°: 45)	S013 (SEQ ID N°: 58)
Fase I-B	228_255_271	3,0 (SEQ ID N°: 112)	S013 (SEQ ID N°: 58)	S019 (SEQ ID N°: 64)
Fase II Completa	187_262_250_229_252_268_265_263	9,7 (SEQ ID N°: 113)	S000 (SEQ ID N°: 45)	S019 (SEQ ID N°: 64)
Fase III-A	186_261_241_229	4,4 (SEQ ID N°: 114)	S000 (SEQ ID N°: 45)	S008 (SEQ ID N°: 53)
Fase III-B	251_249_237_269_185	4,3 (SEQ ID N°: 115)	S009 (SEQ ID N°: 54)	S018 (SEQ ID N°: 63)
Reciclina marcadora de Fase I	270_URA3blaster_44_FS-a_tHMG1-a	6,3 (SEQ ID N°: 116)	S000 (SEQ ID N°: 45)	S019 (SEQ ID N°: 64)
Reciclina marcadora de Fase II	187_URA3blaster_44_FS-a_ERG12	6,2 (SEQ ID N°: 117)	S000 (SEQ ID N°: 45)	S019 (SEQ ID N°: 64)
Reciclina marcadora de Fase III	186URA3blaster44FS-a_ERG19	6,0 (SEQ ID N°: 118)	S000 (SEQ ID N°: 45)	S019 (SEQ ID N°: 64)
Nocaute STE5	567556 PTDH3-a_tH <sup>MG1</sup> -b 5 6 3	5,2 (SEQ ID N°: 119)	S000 (SEQ ID N°: 45)	S019 (SEQ ID N°: 64)
Nocaute IME1	IME 1 US_LEU2_PTDH3-bFS-bIME1DS	5,4 (SEQ ID N°: 120)	S000 (SEQ ID N°: 45)	S019 (SEQ ID N°: 64)

O primeiro ciclo de amplificação de PCR foi realizado como seguem: um ciclo de desnaturação a 98°C por 2 minutos; 5 ciclos de desnaturação a 98°C por 30 segundos e recozer/estender a 72°C por 30 segundos por quilobase do produto PCR. O segundo ciclo de amplificação de PCR foi realizado como seguem: um ciclo de desnaturação a 98°C por 2 minutos; 35 ciclos de desnaturação a 98°C por 12 segundos e anelar /estender a 72°C por 20-25 segundos por quilobase de produto PCR; um ciclo de extensão final a 72°C por 7 minutos; e um final mantendo a 4°C. Quando a temperatura de anelamento não foi 72°C (isto é, quando foi 54°C ou 65°C), no primeiro ciclo de amplificação de PCR a 1 minuto da etapa de anelamento por 30 segundos por quilobase do produto de PCR da etapa de extensão a 72°C foi usada e para o segundo ciclo de amplificação de PCR a 15 segundos da etapa de anelamento seguido por 20 segundos por quilobase do produto

PCR da etapa de extensão a 72°C foi usado.

\* Vetores de montagem são projetados com os números e MULEs com nomes.

Como mostrado nas FIGS. 5 e 6, 2 a 9 polinucleotídeos componentes foram corretamente montados nos polinucleotídeos montados longos de até 7,7 kb.

#### Exemplo 8

[00201] Este modelo descreve os métodos para geneticamente gerar os microorganismos hospedeiros alterados usando polinucleotídeos montados, montados pelos métodos divulgados neste.

[00202] Os polinucleotídeos montados de Fase I-A e Fase I-B (ver Tabela 7) foram clonados no vetor de clonagem TOPO Zero Blunt II (Invitrogen, Carlsbad, CA), que produzem os plasmídeos TOPO-Fase I-A e TOPO-Fase I-B, respectivamente. As construções foram propagadas nas células TOP 10 (Invitrogen, Carlsbad, CA) desenvolvido em ágar LB contendo 50 µg/ml de canamicina. Cada plasmídeo foi digerido até a finalização usando endonuclease de restrição NotI, as inserções Fase I-A e Fase I-B foram extraídas em gel usando um kit de purificação em gel (Qiagen, Valencia, CA) e razões molares iguais dos fragmentos de DNA purificados foram ligados usando T4 DNA ligase (New England Biolabs, Ipswich, MA), produzindo o polinucleotídeo montado de Fase I completa. O polinucleotídeo montado de Fase I completa foi clonado no vetor de clonagem TOPO Zero Blunt II (Invitrogen, Carlsbad, CA), produzindo o plasmídeo TOPO-Fase I. A construção foi propagada em células TOP10 (Invitrogen, Carlsbad, CA) desenvolvido em ágar LB contendo 50 µg/ml de canamicina.

[00203] O polinucleotídeo montado de Fase II completa (ver Tabela 7) foi clonado no vetor de clonagem TOPO Zero Blunt II (Invitrogen, Carlsbad, CA), produzindo o plasmídeo TOPOFase II. A construção foi propagada em células TOP10 (Invitrogen, Carlsbad, CA) desenvolvido em ágar LB contendo 50 µg/ml de canamicina.

[00204] Os polinucleotídeos montados de Fase III-A e Fase III-B (ver

Tabela 7) foram clonados no vetor de clonagem TOPO Zero Blunt II (Invitrogen, Carlsbad, CA), que produzem os plasmídeos TOPO-Fase III-A e TOPO-Fase III-B, respectivamente. As construções foram propagadas nas células TOP10 (Invitrogen, Carlsbad, CA) desenvolvido em ágar LB contendo 50 µg/ml de canamicina. Cada plasmídeo foi digerido até a finalização usando endonuclease de restrição BamHI e SbfI, as inserções de Fase III-A e Fase III-B foram extraídas em gel usando um kit de purificação em gel (Qiagen, Valencia, CA) e razões molares iguais dos fragmentos de DNA purificados foram ligados usando T4 DNA ligase (New England Biolabs, Ipswich, MA), produzindo o polinucleotídeo montado de Fase III completa. O polinucleotídeo montado de Fase III completa foi clonado no vetor de clonagem TOPO Zero Blunt II (Invitrogen, Carlsbad, CA), produzindo o plasmídeo TOPO-Fase III. A construção foi propagada em células TOP10 (Invitrogen, Carlsbad, CA) desenvolvido em ágar LB contendo 50 µg/ml de canamicina.

[00205] Para as transformações de célula de levedura, 25 ml do meio de dextrose de peptona de extrato de levedura (YPD) foi inoculado com uma colônia simples de uma cepa hospedeira de partida. A cultura foi desenvolvida durante a noite a 30°C em um agitador rotativo a 200 rpm. O OD600 de uma cultura foi medida e a cultura foi então usada para inocular 50 ml do meio YPD a um OD600 de 0,15. A cultura novamente inoculada foi desenvolvida a 30°C em um agitador rotativo a 200 rpm até um OD600 de 0,7 a 0,9, em que os pontos das células foram transformados com 1 µg de DNA. As células foram deixadas recuperar no meio YPD por 4 horas antes destes serem colocados no ágar contendo um agente seletivo para identificar os transformantes da célula hospedeira.

[00206] A cepa hospedeira iniciadora Y1198 foi gerada pela levedura PE-2 seca ativa recolocada em suspensão (isolado em 1994 a Santelisa Vale, Sertãozinho, Brasil) em 5 mL do meio YPD contendo 100 µg/mL de

carbamicilina e 50 ug/mL de canamicina. A cultura foi incubada durante a noite a 30°C em um agitador rotativo a 200 rpm. Uma alíquota de 10 uL de uma cultura foi então colocada em uma placa de YPD e deixado secar. As células foram seriamente movidas pelas colônias simples e incubadas por 2 dias a 30°C. Doze colônias simples foram selecionadas, consertada em uma placa de YPD e deixada desenvolver durante a noite a 30°C. As identidades da cepa das colônias foram verificadas pela análise dos seus tamanhos cromossomais em um sistema Bio-Rad CHEF DR II (Bio-Rad, Hercules, CA) usando o kit de tampão de DNA genômico Bio-Rad CHEF (Bio-Rad, Hercules, CA) de acordo com as especificações dos fabricantes. Uma colônia foi selecionada e estocada como cepa Y1198.

[00207] As cepas Y1661, Y1662, Y1663 e Y1664 foram geradas a partir da cepa Y1198 para render a cepa haplóide. A cepa Y1198 foi desenvolvida durante a noite em 5 mL do meio YPD a 30°C em um tubo de vidro em um tambor. O OD600 foi medido e as células foram diluídas em um OD600 de 0,2 em 5 mL do meio YP contendo 2% de acetato de potássio. A cultura foi desenvolvida durante a noite a 30°C em um tubo de vidro em um tambor. O OD600 foi medido novamente e 4 OD600\*mL das células foi coletado pela centrifugação a 5.000g por 2 minutos. O grânulo celular foi lavado uma vez com água estéril e então recolocado em suspensão em 3 mL de 2% de acetato de potássio contendo 0,02% de rafinose. As células foram desenvolvidas por 3 dias a 30°C em um tubo de vidro em um tambor. A esporulação foi confirmada pela microscopia. Uma alíquota de 33 uL de uma cultura foi transferida em um tubo de microfuge 1,5 mL e foi centrifugado a 14.000rpm por 2 minutos. O grânulo celular foi recolocado em suspensão em 50 uL de água estéril contendo 2 uL de 10 mg/mL de Zymolyase 100T (MP Biomedicals, Solon, OH) e as células foram incubadas por 10 minutos em um banho de água 30°C. O tubo foi transferido ao gelo e 150 uL de água gelada foi adicionada. Uma alíquota de 10 uL desta mistura foi adicionada em uma

placa de YPD de 12 mL e tetrads foram dissecados em um microscópio de dissecação Singer MSM 300 (Singer, Somerset, UK). A placa de YPD foi incubada a 30°C por 3 dias, após o qual os esporos foram consertados em uma placa fresca de YPD e desenvolvida durante a noite a 30°C. Os tipos mating de cada esporo de 8 tetrads de quatro esporos foram analisados pela colônia de PCR. Um tetrad de 4 esporos simples com 2 esporos MATA e 2 MATalfa foi selecionado e estocado como cepas Y1661 (MATA), Y1662 (MATA), Y1663 (MATalfa) e Y1664 (MATalfa).

[00208] A cepa hospedeira 1515 foi gerada pela transformação da cepa Y1664 com plasmídeo TOPO-Fase I digerido até a finalização usando endonuclease de restrição PmeI. Os transformantes da célula hospedeira foram selecionados no meio YPD contendo 300 ug/mL de higromicina B.

[00209] A cepa hospedeira 1762 foi gerada pela transformação da cepa Y1515 com plasmídeo TOPO-Fase II digerido até a finalização usando endonuclease de restrição PmeI. Os transformantes da célula hospedeira foram selecionados no meio YPD contendo 100 ug/mL de nourseotricina.

[00210] A cepa hospedeira 1770 foi gerada pela transformação da cepa Y1762 em duas etapas com plasmídeo de expressão pAM404 e plasmídeo TOPO-Fase III digerido até a finalização usando endonuclease de restrição PmeI. O plasmídeo de expressão pAM404 foi derivado do plasmídeo pAM353, que foi gerado pela inserção de uma sequência de nucleotídeo que codifica uma sintase  $\beta$ -farneseno no vetor pRS425-Gall (Mumberg et. al. (1994) *Nucl. Acids. Res.* 22(25): 5767-5768). A inserção da sequência de nucleotídeo foi gerada sinteticamente, usando como um modelo da sequência codificadora do gene de sintase  $\beta$ -farneseno de *Artemisia annua* (Número de acesso Genbank AY835398) códon otimizado pela expressão em *Saccharomyces cerevisiae* (SEQ ID N°: 121). A sequência de nucleotídeo sinteticamente gerada foi flanqueada por locais de restrição 5' BamHI e 3' XhoI e deve ser deste modo clonado nos locais de restrição compatíveis de



um vetor de clonagem tal como um pUC padrão ou vetor de origem pACYC. A sequência de nucleotídeo sinteticamente gerada foi isolada pela digestão até a finalização da construção da síntese de DNA usando enzimas de restrição *BamHI* e *XhoI*. A mistura de reação foi determinada pela eletroforese em gel, o fragmento de DNA de aproximadamente 1,7 kb que compreende a sequência codificadora de sintase  $\beta$ -farneseno foi extraído em gel e o fragmento de DNA isolado foi ligado no local de restrição *BamHI XhoI* do vetor pRS425-Gall, produzindo o plasmídeo de expressão pAM353. A sequência de nucleotídeo que codifica a sintase  $\beta$ -farneseno foi amplificado por PCR de pAM353 usando iniciadores GW-52-84 pAM326 *BamHI* (SEQ ID N°: 188) e GW-52-84 pAM326 *NheI* (SEQ ID N°: 189). O produto PCR resultante foi digerido até a finalização usando enzimas de restrição *BamHI* e *NheI*, a mistura de reação foi determinada pela eletroforese em gel, o fragmento de DNA de aproximadamente 1,7 kb que compreende a sequência codificadora de sintase  $\beta$ -farneseno foi extraído em gel e o fragmento de DNA isolado foi ligado em o local de restrição *BamHI NheI* do vector pAM178 (SEQ ID N°: 122), produzindo o plasmídeo de expressão pAM404. Os transformantes da célula hospedeira com pAM404 foram selecionados no meio sintético completo (CSM) desprovido de metionina e leucina. Os transformantes da célula hospedeira com pAM404 e polinucleotídeo montado de Fase III completa foram selecionados no CSM desprovidos de metionina e leucina e contendo 200 ug/mL G418.

[00211] A cepa hospedeira 1793 foi gerada pela transformação da cepa Y1770 com uma construção nocaute URA3 (SEQ ID N°: 123). A construção nocaute foi gerada pela geração dos fragmentos de DNAs URA3US (gerado da cepa *Saccharomyces cerevisiae* CEN.PK2 DNA genômico usando PCR iniciadores KMH33-276-21-URA3 5'.fwd (SEQ ID N°: 147) e KMH34-276-21-URA3 5'.rev (SEQ ID N°: 148)) e URA3DS (gerado da cepa *Saccharomyces cerevisiae* CEN.PK2 DNA genômico usando PCR iniciadores

KMH35-276-21-URA3 3'.fwd (SEQ ID N°: 149) e KMH36-276-21-URA3 3'.rev (SEQ ID N°: 150); seguido pela costura de dois fragmentos de DNA junto usando PCR iniciadores KMH33-276-21-URA3 5'.fwd e KMH36-276-21-URA3 3'.rev. Os transformantes da célula hospedeira foram selecionados no meio YPD contendo 5-FOA.

[00212] A cepa hospedeira YAAA foi gerada pela transformação da cepa Y1793 com um marcador de Fase I reciclando o polinucleotídeo montado (ver Tabela 7). Os transformantes da célula hospedeira foram selecionados no CSM desprovidos de metionina e uracila. O marcador URA3 foi taxado pelo desenvolvimento das células durante a noite no meio YPD a 30°C em um agitador rotativo a 200 rpm e então banhando as células em ágar contendo 5-FOA. O corte marcador foi confirmado pela colônia de PCR.

[00213] A cepa hospedeira YBBB foi gerada pela transformação da cepa YAAA com o marcador de Fase II reciclando o polinucleotídeo montado (ver Tabela 7). Os transformantes da célula hospedeira foram selecionados no CSM desprovidos de metionina e uracila. O marcador URA3 foi taxado pelo desenvolvimento das células durante a noite no meio YPD a 30°C em um agitador rotativo a 200 rpm e então colocando as células em ágar contendo 5-FOA. O corte marcador foi confirmado pela colônia de PCR.

[00214] A cepa hospedeira Y1912 foi gerada pela transformação da cepa YBBB com o marcador de Fase III reciclando o polinucleotídeo montado (ver Tabela 7). Os transformantes da célula hospedeira foram selecionados no CSM desprovidos de metionina e uracila. O marcador URA3 foi taxado pelo desenvolvimento das células durante a noite no meio YPD a 30°C em um agitador rotativo a 200 rpm e então colocando as células em ágar contendo 5-FOA. o corte marcador foi confirmado pela colônia de PCR.

[00215] A cepa hospedeira Y1913 foi gerada pela transformação da cepa Y1912 com o polinucleotídeo montado nocaute STE5 (ver Tabela 7). Os transformantes da célula hospedeira foram selecionados no CSM desprovidos

de metionina e uracila.

[00216] A cepa hospedeira Y1915 foi gerada a partir da cepa Y1913 pela cura da cepa de pAM404 e transformando a cepa resultante com o polinucleotídeo montado nocaute IME 1 (ver Tabela 7). A cepa Y1913 foi propagada no meio YPD não selecionado a 30°C em um agitador rotativo a 200 rpm. Aproximadamente 100 células foram colocadas no meio YPD sólido e deixado desenvolver por 3 dias a 30°C antes destas terem suas réplicas colocadas nas placas CSM desprovidas de metionina e leucina onde estes foram desenvolvidos por outros 3 dias a 30°C. As células curadas foram identificadas por sua capacidade de desenvolver no meio mínimo contendo leucina e sua incapacidade de desenvolver no meio desprovido de leucina. Uma tal simples colônia foi selecionada e transformada com o polinucleotídeo montado nocaute IME1. Os transformantes da célula hospedeira foram selecionados no CSM desprovido de metionina e uracila.

#### Exemplo 9

[00217] Este modelo descreve os métodos para seleção das sequências ligantes aneladas a serem usadas para montar os polinucleotídeos componentes que codificam um promotor e uma sequência codificadora de proteína em um polinucleotídeo montado pelos métodos inventivos divulgados neste.

[00218] Os promotores que codificam os MULEs seguido pelas duas sequências ligantes aneladas candidatas diferentes, a sequência ligante anelada RYSE 15 (R15; SEQ ID N°: 15) e sequência ligante anelada RYSE 7 (R7; SEQ ID N°: 7), bem como MULEs que codificam GFP precedidos pelas duas sequências ligantes aneladas, foram amplificadas por PCR como descrito na Tabela 8.

Tabela 8 – Promotores que codificam os MULEs amplificados e GFP com sequências ligantes aneladas RYSE 15 (R15) ou Sequência ligante anelada RYSE 7 (R7)				
MULE	Tipo *	Iniciadores	Tamanho	Modelo
pGAL1 -R15	P	Plan X19 (SEQ ID N°: 236) Plan X20 (SEQ ID N°: 237)	698	Cepa <i>S. cerevisiae</i> CEN.PK2 DNA genômico

pTDH3 R15	P	Plan X47(SEQ ID N°: 238) Plan X48(SEQ ID N°: 239)	613	Cepa <i>S. cerevisiae</i> CEN.PK2 DNA genômico
p CYC 1 -R15	P	Plan X 1 i (SEQ ID N°: 240) Plan X12(SEQ ID N°: 241)	645	Cepa <i>S. cerevisiae</i> CEN.PK2 DNA genômico
pGALI-R7	p	Plan X19 (SEQ ID N°: 236) Plan X64 (SEQ ID N°: 242)	692	Cepa <i>S. cerevisiae</i> CEN.PK2 DNA genômico
pTDH3-R7	p	Plan X47(SEQ ID N°: 238) Plan X71(SEQ ID N°: 243)	607	Cepa <i>S. cerevisiae</i> CEN.PK2 DNA genômico
p CYC 1-R7	p	Plan X 1 I (SEQ ID N°: 240) Plan X78(SEQ ID N°: 244)	639	Cepa <i>S. cerevisiae</i> CEN.PK2 DNA genômico
R7-GFP	GsT	Plan X96(SEQ ID N°: 247) Plan X88(SEQ ID N°: 245)	1378	DNA de plasmídeo **RABit 634
A-GFP	GsT	Plan X89(SEQ ID N°: 246) Plan X88(SEQ ID N°: 245)	1385	DNA de plasmídeo **RABit 634

Reações PCR contidas: 67 uL de ddH<sub>2</sub>O, 20 uL 5x HF de tampão, 2 uL de cada iniciador (10 uM), 1 uL de mistura dNTP (200 uM), 1 uL de polimerase de DNA Phusion (New England Biolabs, Ipswich, MA) e 9 uL de DNA genômico Y002 ou DNA de plasmídeo RABit 634.

Amplificação de PCR foi realizada como seguem: 1 ciclo de desnaturação a 98°C por 2 minutos; 9 ciclos de desnaturação a 98°C por 15 segundos, anelar a 61°C por 30 segundos diminuindo por 1°C cada ciclo entendido a 72°C por 1 minuto; 26 ciclos de desnaturação a 98°C por 15 segundos, anelar a 52°C por 30 segundos e entendido a 72°C por 1 minuto; 1 ciclo de extensão final a 72°C por 7 minutos; e um final mantido a 4°C.

\* G=gene; s=códon de interrupção; T=terminador; P=promotor.

\*\* RABit 634 compreende a sequência codificadora da proteína fluorescente verde (GFP) seguido pelo terminador do gene ADH1 de *Saccharomyces cerevisiae*.

[00219] As reações de PCR foram determinadas pela eletroforese em gel, os MULEs foram purificados em gel e os MULEs purificados foram usados para montar os polinucleotídeos montados testados. Neste final, os MULEs e vetores de montagem (ver Tabela 6) a serem montados (ver Tabela 9) foram colocados juntos em um tubo (333 fmol de cada vetor de montagem, 667 fmol de cada MULE) e digerido usando enzima de restrição LguI (Fermentas, Glen Burnie, MD). A enzima de restrição foi inativada por calor por 20 minutos a 65°C. As amostras foram unidas em três reações de 30 uL; água, tampão, dNTPs e polimerase de DNA foram adicionados a cada mistura de reação e um primeiro ciclo de amplificação de PCR foi iniciado. Os iniciadores terminais foram então adicionados nas misturas de reação e um segundo ciclo de amplificação de PCR foi realizado (ver Tabela 9).

[00220] As três misturas de reação de PCR foram combinadas em um tubo, as misturas de reação foram determinadas pela eletroforese em gel e os

produtos PCRs foram purificados em gel.

Tabela 9 - Iniciadores terminais para a montagem dos polinucleotídeos montados testados				
Montagem	MULEs (ver Tabela 8) e Vetores de montagem (ver Tabela 6) a ser combinado	Tamanho do polinucleotídeo montado (kb)	Iniciador de terminal 1	Iniciador de terminal 2
1	97_555_pGAL1-A_A-GFP_24	4,7	S000 (SEQ ID N°: 45)	S019 (SEQ ID N°: 64)
2	97_555_pTDH3-A_A-GFP_24	4,6		
3	97_555_pCYC 1 -A_A-GFP_24	4,7		
7	97_555_pGAL1-R7_R7-GFP_24	4,7		
8	97555_pTDH3-R7_R7-GFP_24	4,6		
9	97_555_pCYC1-R7_R7-GFP_24	4,6		
<p>As reações de PCR continham: 41 uL de ddH<sub>2</sub>O, 20 uL de 5x tampão HF, 5 uL de cada iniciador terminal (1 uM), 2 uL de mistura de dNTP (200 uM), 1,8 uL de Phusion DNA Polimerase e 30 uL MULE ou vetor de montagem digerido por LguI.</p> <p>O primeiro ciclo de amplificação de PCR foi realizado como segue: 1 ciclo de desnaturação a 98°C por 2 minutos; 5 ciclos de desnaturação a 98°C por 30 segundos, anelamento a 60°C por 30 segundos e estendido a 72°C por 2,5 minutos; seguido mantendo-se a 4°C pela adição dos dois terminais iniciadores terminais. O segundo ciclo de amplificação de PCR foi realizado como segue: 1 ciclo de desnaturação a 98°C por 2 minutos; 35 séries de desnaturação a 98°C por 12 segundos, anelamento a 60°C por 30 segundos e extensão a 72°C por 2,5 minutos; 1 ciclo de extensão final a 72°C por 7 minutos e uma manutenção final a 4°C.</p> <p>* Os vetores de montagem são projetados com números e MULEs com nomes.</p>				

[00221] Os polinucleotídeos montados de teste foram usados para transformar uma cepa hospedeira de *Saccharomyces cerevisiae* que foi deficiente em URA3 e teve uma anulação do local GAL80. Os transformantes da célula hospedeira foram selecionados no CSM desprovidos de uracila e integração genômica correta do polinucleotídeo montado foi confirmada pela colônia PCR. Duas colônias verificadas de cada transformação foram coletadas em 360 uL de Meio Bird Seed (BSM) contendo 2% de sacarose e as culturas foram incubadas por 48 horas a 30°C em um agitador rotativo a 999 rpm. Uma alíquota de 14,4 uL foi retirada de cada reservatório e transferido a 1,1 mL de BSM contendo 4% de sacarose em uma placa de bloco de 96 reservatórios e cultivado por outras 6 horas a 30°C em um agitador rotativo a 999 rpm, ponto no qual 100 uL de cada cultura foi transferida a um reservatório de uma placa de 96 reservatórios de fundo claro para a análise da expressão de GFP. A expressão de GFP em cada reservatório foi analisada pela medição de 515 nm de emissão após 485 nm de excitação em um espectrômetro leitor de placa M5 (Molecular Devices, Sunnyvale, CA). As concentrações de GFP medidas foram normalizadas para o desenvolvimento de cultura celular por divisão pela leitura de OD600 para cada cultura.

[00222] Como mostrado nas Tabela 10, a sequência ligante anelada RYSE 15 permitiu a expressão conduzida pelo promotor de GAL1, TDH3 e CYC1 aumentada do gene repórter GFP nos polinucleotídeos montados de teste em comparação com a sequência ligante anelada RYSE 7.

Tabela 10 – Expressão de GFP em células hospedeiras que abrigam Polinucleotídeos montados de teste que compreende Sequência ligante anelada RYSE 15 (R15) ou Sequência de Ligante Anelável RYSE 7 (R7) Entre Promotor e Repórter GFP				
Sequência de ligante anelável posicionado entre o promotor e o gene repórter GFP no polinucleotídeo montado de teste	% média de expressão de GFP (em comparação com a% média de expressão de GFP obtido com células hospedeiras que abrigam um de 3 construções de controle sem costura *; Média para 2 células hospedeiras independentes isoladas)			
	Promotor GAL1	Promotor TDH3	Promotor CYC1	Média através de todos os três promotores
R15	79,34	91,42	81,92	84,22
R7	27,43	54,68	46,31	42,81
* As construções de controle sem costura têm uma estrutura idêntica como os polinucleotídeos montados de teste exceto que as sequências promotoras foram ligadas sem costura ao gene repórter GFP (isto é, sem uma sequência ligante de intervenção anelada).				

### Exemplo 10

[00223] Este modelo descreve os métodos para a montagem combinatorial de rendimento alto de polinucleotídeos e métodos para a geração de rendimento alto de células hospedeiras que compreende polinucleotídeos combinatoriamente combinados.

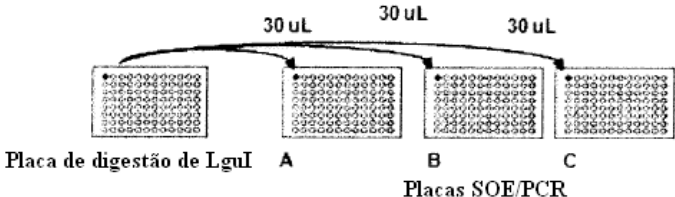
[00224] Os polinucleotídeos componentes usados neste modelo e os polinucleotídeos montados e combinados esperados gerados a partir destes polinucleotídeos componentes, são esquematicamente ilustrados na FIG. 12A. Os polinucleotídeos componentes compreenderam segmentos de DNA que codificam uma sequência alvejadora cromossomal a montante e a jusante (US e DS), 6 promotores diferentes (P), 35 proteínas diferentes (G) e um segmento 5' e um 3' do marcador selecionável URA3 (URA e RA3, respectivamente), flanqueado por pares de sequência de ligante anelável ou pares de segmento de ligação iniciador/sequência ligante anelada.

[00225] Os polinucleotídeos componentes foram liberados a partir de vetores de montagem pela digestão de RABits ou MULES usando endonuclease de restrição LguI. Para este fim, as placas de 96 reservatórios

("placas de digestão de LguI") foram ajustadas como mostrado na tabela abaixo e as placas foram incubadas a 37°C por 75 minutos, após o que a endonuclease de restrição LguI foi inativada por calor a 65°C por 20 minutos em uma máquina de PCR.

Placas de digestão de LguI	
Componente (por reservatório)	Volume (uL)
667 fMoles de RABit ou MULE	Variável
10x Tampão Tango (Fermentas, Glen Burnie, MD)	10
LguI (Fermentas, Glen Burnie, MD)	2,5
ddH2O	a 100

Os polinucleotídeos componentes foram montados por SOE. Para cada placa de digestão de LguI, placas de 96 reservatórios em triplicata ("placas SOE/PCR") foram ajustadas e submetidas ao termociclo em uma máquina de PCR como mostrado na tabela abaixo.

Placas SOE/PCR			
Componente (por reservatório)		Volume (uL)	
ddH2O		41	
5x tampão de Phusion HF (New England Biolabs, Ipswich, MA)		20	
10 mM de mistura de dNTP		2	
Phusion DNA polimerase (New England Biolabs, Ipswich, MA)		1,8	
RABits ou MULEs digeridos por a serem montados		30	
			
		Total: 95	
Condições de termociclo			
Desnaturação inicial		98°C	2 minutos
7 ciclos	Desnaturação	98°C	30 segundos
	Anelamento	67°C	30 segundos
	Extensão	72°C	5 minutos
Manutenção		4°C	∞

[00226] Os polinucleotídeo montados foram amplificados por PCR. Cada placa de SOE/PCR recebeu reagente adicional e foi submetido ao termociclo em uma máquina de PCR como mostrado na tabela abaixo. Os reservatórios correspondentes em placas de SOE/PCR foram unidas em

blocos de 96 reservatórios profundos e os polinucleotídeos montados foram purificados usando-se o Kit Omega Biotek E-Z 96® Cycle-Pure (Omega Bio-Tek Inc., Norcross, GA) como pelo protocolo sugerido pelos fabricantes (volumes finais aproximados de 45 uL).

Placas de SOE/PCR			
Componente adicional (por reservatório)		Volume (uL)	
Estoque de 10 mM de iniciadores terminais S000 (SEQ ID N° 45 e S019 (SEQ ID N°: 64)		10	
Condições de termociclo			
Desnaturação inicial		98°C	2 minutos
35 ciclos	Desnaturação	98°C	12 segundos
	Recozimento	67°C	30 segundos
	Extensão	72°C	4,5 minutos
Extensão Final		72°C	7 minutos
Manutenção		4°C	∞

[00227] A FIG. 12B mostra polinucleotídeos montados exemplares (em caixa) dissolvido em um gel de agarose a 1%.

[00228] Os polinucleotídeos montados purificados foram digeridos com endonuclease de restrição LguI para gerar extremidades pegajosas para clonagem. Para este fim, as placas de 96 reservatórios ("Placas de Digestão Unidas em Polinucleotídeo LguI") foram ajustadas como mostrado na tabela abaixo e as placas foram incubadas a 37°C por 60 minutos, após o que a endonuclease de restrição LguI foi inativado por calor a 65°C por 20 minutos em uma máquina de PCR. Os polinucleotídeos digeridos por LguI montados foram purificados em gel usando o Kit de Recuperação de DNA em gel ZR-96 Zymoclean™ (Zymo Research Corporation orange, CA) como pelo protocolo recomendado pelo fabricante.

LguI Polinucleotídeo montado Placas de digestão de LguI	
Componente (por reservatório)	Volume (uL)
polinucleotídeo montado purificado	43
10x tampão Tango	5
LguI	2



[00229] Os polinucleotídeo montados foram ligados em uma cadeia principal do vetor com base em pUC-19. Quando nenhuma inserção é ligada neste vetor, um promotor pTRC (isto é, o promotor do gene TRC de *Saccharomyces cerevisiae*) conduz a expressão de GluRS e mata a célula hospedeira. As placas de 96 reservatórios ("Placas de Ligação") foram ajustadas como mostrado na tabela abaixo e as placas foram incubadas a 24°C por 15 minutos e então a 16°C durante a noite. Os produtos de ligação foram purificados usando-se ZR-96 DNA Clean & Concentrator™-5 (Zymo Research Corporation orange, CA) como pelo protocolo sugeridos pelos fabricantes.

Placas de ligação	
Componente (por reservatório)	Volume (uL)
ddH <sub>2</sub> O	5
10x tampão de T4 DNA Ligase	2
Cadeia principal de vetor	2
polinucleotídeo montado purificado	10
T4 DNA ligase (NEB, Ipswich, MA)	1

[00230] Os produtos de ligação foram submetidos à eletroporação em células competentes de *E. coli*. Placas de eletroporação de 96 reservatórios pré-esfriados foram ajustadas e as eletroporações foram realizadas como mostrado na tabela abaixo.

Placas de Eletroporação		
Componente (por reservatório)		Volume (uL)
Produtos de ligação purificados		10
Células competentes Lucigen 10G (Lucigen Corporation, Middleton. WI)		25
Ajustes de eletroporação		
2400V	750 $\Omega$	25 uF

[00231] 1,1 mL de placas de cultura de 96 reservatórios ("placas de cultura") contendo 250 uL de SOC pré-aquecido foram apresentados e 100 uL de SOC foi tirado de cada reservatório e adicionado as células eletroporadas imediatamente após a eletroporação. O SOC e as células foram misturadas e 100 uL de cada mistura foi transferido novamente a uma placa de cultura. As placas de cultura foram incubadas a 37°C por 1 hora em um agitador

incubador Multitron II (ATR Biotech, Laurel, MD). Duas diluições das células (3 ul e 240 ul) foram colocadas em ágar LB que compreende 50 ug/mL de canamicina e incubadas durante a noite a 37°C. As colônias foram selecionadas e desenvolvidas nas placas de 96 reservatórios profundos que compreendem 1 mL de meio LB com canamicina por reservatório e DNA foi extraído para a análise de restrição usando endonuclease de restrição LguI. Os resultados de tal análise de restrição por 22 de 24 colônias exemplares que compreendem um polinucleotídeo combinado aproximadamente de 8 kb são mostrados na FIG. 12C.

[00232] As células de levedura que compreendem polinucleotídeos combinados cromossomicamente integrados foram gerados pela recombinação homóloga mediada pela célula hospedeira entre sequências de alvejamento cromossomal terminal e segmentos marcadores selecionáveis dos polinucleotídeos montados. Para este fim, as placas de PCR de 96 reservatórios ("placas de transformação de levedura") foram apresentados e o choque por calor até as transformações serem realizadas em uma máquina de PCR como mostrado na tabela abaixo.

Placas de transformação de levedura	
Componente (por reservatório)	Volume (uL)
DNA Miniprep (20 ng/uL)	10
Células de levedura competentes*	40
Mistura PEG/SS/LiAc master **	200
Choque por calor	
30°C	30 min
42°C	45 min
24°C (opcional)	30 min
*preparado pelas células de desenvolvimento em 100 mL de YPD durante a noite, diluindo a cultura e desenvolvimento a um OD600 de cerca de 0,8 durante a noite, rotacionando as culturas a 3.000g por 5 minutos, lavando o grânulo celular com 1 L de ddH2O, lavando o grânulo celular com 1 L de acetato de lítio a 100 mM (LiAc) e recolocando em suspensão o grânulo celular a um volume total de 18 mL em 100 mM de LiAc.	
** mistura Master suficiente por 4 placas de PCR contendo 100 mL de 50% PEG, 4 mL submetido a ebulição (95°C por pelo menos 10 minutos) DNA filamentado simples, 15 mL 1M de LiAc.	

[00233] As placas de transformação de levedura foram rotacionadas a

2.000g por 2 minutos, os sobrenadantes foram removidos e os grânulos celulares foram lavados três vezes com 200 uL de ddH<sub>2</sub>O. Os grânulos celulares foram recolocados em suspensão com 100 uL de meio Bird Seed frio (BSM) tirado de previamente preparado pelas placas de cultura de 96 reservatórios pré-resfriadas ("placas de semente") contendo 360 uL de BSM frio por reservatório. As células recolocadas em suspensão foram transferidas as placas de semente e foram desenvolvidas durante a noite a 30°C em um agitador incubador Multitron II. As placas de semente foram rotacionadas a 3.000g por 5 minutos, todos mais 60 uL do líquido foi removido e placas de semente recuperadas foram agitadas a 1.000 rpm aos grânulos celulares recolocados em suspensão.

#### Exemplo 11

[00234] Este modelo descreve os métodos para a geração as células de levedura que compreendem os polinucleotídeos montados gerados pela recombinação homóloga mediada pela célula hospedeira.

[00235] O polinucleotídeo montado e polinucleotídeos componentes usados neste modelo e o local cromossomal esperado obtido na montagem e integração cromossomal, são esquematicamente ilustrados na FIG. 13A.

[00236] As transformações da célula de levedura foram realizadas como descrito na tabela abaixo. Seguindo o choque por calor, as células foram rotacionadas, o sobrenadante foi removido, as células foram recolocadas em suspensão em 400 uL de ddH<sub>2</sub>O, e os transformantes da célula hospedeira foram selecionados colocando-se 100-200 uL da suspensão celular em ágar desprovido de uracila.

Transformação de levedura	
Componente	Volume (uL)
Componente e polinucleotídeos montados (300-500 ng cada)	20
Células de levedura competentes *	Grânulo celular *
50% de solução PEG	240
1 M de LiAc pH 8,4-8,9	36
DNA filamentado simples submetido a ebulição (95°C por 5 minutos) (10 mg/mL) (Invitrogen, Carlsbad, CA)	10
Choque por calor ddH <sub>2</sub> O	54

42°C	40 minutos
*Preparado pelo desenvolvimento das células a partir da colônia em 25 mL de YPD durante a noite a 30°C a um OD600 de 0,7-0,9, rotacionado as células, lavando o grânulo celular com 5-10 mL de ddH <sub>2</sub> O, lavando o grânulo celular com 1 mL de ddH <sub>2</sub> O, lavando o grânulo celular com 1 mL de acetato de lítio a 100 mM (LiAc), rotacionado na microcentrífuga por 30 segundos para granular as células e descartar o sobrenadante.	

[00237] A integração bem sucedida de polinucleotídeos montados foi determinada por cPCR usando iniciadores cPCR A, B, E e F (junção da extremidade 5' do local de integração cromossomal) ou iniciadores cPCR C, D, G e H (junção da extremidade 3' do local de integração cromossomal) (FIG. 13A). Como mostrado nas FIG. 13B, todas as 8 colônias analisadas produzidas pelo grupo de PCR a 700 bp indicativas de um evento de integração cromossomal positivo do polinucleotídeo montado esperado e perdeu o grupo PCR a 950 bp que o local natural teria produzido.

#### Exemplo 12

[00238] Este modelo descreve os métodos para a geração em rendimento alto das células de levedura que compreende polinucleotídeos combinatoriamente combinados ou combinatoriamente montados pela recombinação homóloga mediada pela célula hospedeira.

[00239] Os polinucleotídeos componentes usados neste modelo e os polinucleotídeos combinados esperados obtidos na montagem e combinação pela recombinação homóloga mediada pela célula hospedeira, são esquematicamente ilustrados na FIG. 14A. Os polinucleotídeos componentes compreendem os segmentos de DNA que codificam uma sequência de alvejamento cromossomal a jusante e a montante (US e DS), 6 promotores diferentes (P), 35 proteínas diferentes (G) e um segmento da extremidade 5' e 3' do marcador selecionável URA3 (URA e RA3, respectivamente), flanqueado pelos pares das sequências ligantes aneladas ou segmento de ligação iniciador/pares da sequência ligante anelada.

[00240] Os componentes dos polinucleotídeos foram liberados a partir dos vetores de montagem pela digestão de RABits ou MULES usando endonuclease de restrição LguI. Para este fim, as placas de 96 reservatórios

("placas de digestão LguI") foram apresentados como mostrado na tabela abaixo e as placas foram incubadas a 37°C por 75 minutos, após o qual a endonuclease de restrição LguI foi inativada por calor a 65°C por 20 minutos em uma máquina PCR.

Placas de digestão LguI	
Componente (por reservatório)	Volume (uL)
667 fMoles RABit ou MULE	Variável
Tampão 10x Tango (Fermentas, Glen Burnie, MD)	5
LguI (Fermentas, Glen Burnie, MD)	2,5
ddH2O	a 50

[00241] Para gerar as células de levedura que compreendem as placas de PCR de 96 reservatórios de polinucleotídeos combinatoriamente combinados e combinatoriamente montados cromossomicamente integrados ("placas de transformação de levedura") foram apresentados e transformações de choque por calor foram realizados em uma máquina de PCR como mostrado na tabela abaixo.

Placas de transformação de levedura	
Componente (por reservatório)	Volume (uL)
Polinucleotídeos componentes	10
Células de levedura competentes*	40
Mistura PEG/SS/LiAc master**	200
Choque por calor	
30°C	30 minutos
42°C	45 minutos
24°C (opcional)	30 minutos
*Preparada pelo desenvolvimento das células em 100 mL de YPD durante a noite, diluindo a cultura e desenvolvimento a um OD600 de cerca de 0,8 durante a noite, rotacionando as culturas a 3.000g por 5 minutos, lavando o grânulo celular com 1 L de ddH2O, lavando o grânulo celular com 1 L de acetato de lítio a 100mM (LiAc) e recolocando em suspensão o grânulo celular a um volume total de 18 mL em 100 mM de LiAc.	
** Mistura Master suficiente por 4 placas de PCR contém 100 mL de 50% de PEG, 4 mL submetido a ebulição (95°C por pelo menos 10 minutos) de DNA filamentado simples, 15 mL de 1M de LiAc.	

[00242] As placas de transformação de levedura foram rotacionados a 2.000 g por 2 minutos, os sobrenadantes foram removidos e os grânulos celulares foram lavados três vezes com 200 uL de ddH2O. As placas celulares

foram recolocadas em suspensão com 100 uL do meio Bird Seed frio (BSM) tirados previamente preparados pelas placas de cultura de 96 reservatórios pré-resfriado ("placas de semente") contendo 360 uL de BSM frio por reservatório. As células recolocadas em suspensão foram transfectadas as placas de semente e foram desenvolvidas durante a noite a 30°C em um agitador incubador Multitron II. As placas de semente foram rotacionadas a 3.000g por 5 minutos, todos mas 60 uL do líquido foi removido e placas de semente recuperadas foram agitadas a 1.000 rpm para recolocar em suspensão os grânulos celulares. Várias diluições das células foram colocadas em ágar desprovido de uracila e incubados durante a noite a 37°C. As colônias dos transformantes da célula de levedura que abrigam um marcador selecionável URA3 funcional foram selecionados e analisados.

[00243] Todas as Publicações, Patentes e Pedidos de Patentes citados nesta especificação são neste incorporados por referência como se cada publicação individual ou Pedido de Patente foram especificamente e individualmente indicados ser incorporados por referência. Embora a invenção precedente foi descrita em alguns detalhes por meio da ilustração e exemplo para os propósitos da clareza do entendimento, será prontamente aparente aquela pessoa com habilidade comum na técnica na luz dos ensinamentos desta invenção que certas mudanças e modificações podem ser feitas neste sem divergir do espírito ou escopo das reivindicações anexas. As formas de realização da presente invenção descritas acima são pretendidas ser meramente exemplares e aquela pessoa habilitada na técnica reconhecerá ou será capaz para determinar o uso não mais do que a experimentação de rotina, equivalentes numerosas aqueles procedimentos específicos descritos neste. Todos os tais equivalentes são considerados estar dentro do escopo da presente invenção e são convertidos pelas seguintes reivindicações. Além disso, como usado nesta especificação e reivindicações, as formas singulares "um," "uma" e "o" incluem formas plurais a não ser do conteúdo claramente

de outra maneira.

## REIVINDICAÇÕES

1. Método para gerar um polinucleotídeo montado a partir de uma pluralidade de polinucleotídeos de componente, caracterizado pelo fato de que compreende as etapas de:

(a) digerir uma composição de montagem com uma ou mais endonucleases de restrição Tipo IIS para gerar uma composição de componentes, a composição de montagem compreendendo:

(i) uma ou mais primeiras moléculas de ácidos nucleicos, em que cada primeira molécula de ácido nucleico é circular e compreende, em uma orientação de 5' a 3', um primeiro local de restrição  $RA_0$ , qualquer segmento de ligação de iniciador selecionado do grupo PA, qualquer segmento de DNA selecionado do grupo  $D_0$ , uma sequência de ligante anelável  $LB_0$  e um segundo local de restrição  $RB_0$ ;

(ii) uma ou mais moléculas de ácido nucleico intermediárias em que cada molécula de ácido nucleico intermediária  $n$  é circular e compreende, em uma orientação de 5' a 3', um primeiro local de restrição  $RA_n$ , uma primeira sequência de ligante anelável  $LA_D$ , qualquer segmento de DNA selecionado do grupo  $D_n$ , uma segunda sequência de ligante anelável  $LB_n$  e um segundo local de restrição  $RB_n$  e em que  $n$  representa um número inteiro de uma das diversas moléculas de ácido nucleico intermediárias; e

(iii) uma ou mais das últimas moléculas de ácido nucleico, em que cada última molécula de ácido nucleico é circular e compreende, em uma orientação de 5' a 3', um primeiro local de restrição  $RA_m$ , uma sequência de ligante anelável  $LA_m$ , um segmento de DNA selecionado do grupo  $D_m$ , qualquer segmento de ligação de iniciador selecionado do grupo PB, um segundo local de restrição  $RB_m$  em que  $m$  representa um número inteiro, um maior do que o número de moléculas de ácido nucleico intermediárias; após o que a clivagem dos locais de restrição  $RA_0$  até  $RB_m$  e a desnaturação das moléculas de ácido nucleico lineares resultantes, cada sequência de ligante



anelável  $LB_{(p-1)}$  é capaz de hibridizar ao complemento de sequência de ligante anelável  $LA_p$ , e servindo como um ponto de iniciação para síntese de um polinucleotídeo complementar, em que  $n$  é um número inteiro que varia de 1 a  $(m-1)$ , em que  $p$  representa um número inteiro de 1 a  $m$  e em que cada grupo  $D_0, \dots D_n, \dots$  e  $D_m$  consiste de um ou mais segmentos de DNA; e

(b) contactar a composição de componentes com DNA polimerase, trifosfatos de desoxirribonucleosídeo e um ou mais primeiro iniciadores e um ou mais segundos iniciadores, sob condições adequadas para a desnaturação da moléculas de ácido nucleico, anelamento da sequência de ligante anelável  $LB_{(p-1)}$  à sequência de ligante anelável  $LA_p$  e extensão deste; em que cada dito primeiro iniciador é capaz de hibridizar a um dos ditos segmentos de ligação de iniciador selecionado do grupo PA e cada dito segundo iniciador é capaz de hibridizar a um dos ditos segmentos de ligação de iniciador selecionado do grupo PB e submeter a composição de componentes à reação de cadeia de polimerase,

em que um polinucleotídeo é montado compreendendo, em uma orientação de 5' a 3', um segmento de DNA selecionado de cada um dos grupos  $D_0, \dots D_n, \dots$  e  $D_m$ .

2. Método para gerar uma célula hospedeira que compreende um polinucleotídeo, caracterizado pelo fato de que compreende as etapas de:

(a) transformar uma célula hospedeira com uma composição que compreende:

(i) uma ou mais primeiras moléculas de ácido nucleico, em que cada molécula compreende, em uma orientação de 5' a 3', qualquer segmento de DNA selecionado do grupo  $D_0$  e uma sequência de ligante anelável  $LB_0$ ;

(ii) opcionalmente, uma ou mais moléculas de ácido nucleico lineares intermediárias, em que cada molécula de ácido nucleico linear intermediária compreende, em uma orientação de 5' a 3', uma primeira sequência de ligante anelável  $LA_n$ , qualquer segmento de DNA selecionado

do grupo  $D_n$  e uma segunda sequência de ligante anelável  $LB_n$ , em que  $n$  representa um número inteiro de uma das diversas moléculas de ácido nucleico intermediárias; e

(iii) uma ou mais das últimas moléculas de ácido nucleico lineares, em que cada última molécula de ácido nucleico linear compreende, em uma orientação de 5' a 3', uma sequência de ligante anelável  $LA_m$  e qualquer segmento de DNA selecionado do grupo  $D_m$ , em que  $m$  representa um número inteiro, um maior do que o número de moléculas de ácido nucleico intermediárias;

em que  $n$  é um número inteiro que varia de 1 a  $(m-1)$ ,

em que cada grupo  $D_0, \dots, D_n, \dots, D_m$  consiste de um ou mais segmentos de DNA,

em que cada sequência de ligante anelável  $LB_{(p-1)}$  compreende uma região de homologia com sequência de ligante anelável  $LA_p$  de comprimento suficiente para iniciar a recombinação homóloga mediada por célula hospedeira entre  $LB_{(p-1)}$  e  $LA_p$  em que  $p$  representa um número inteiro de 1 a  $m$ , em que a dita recombinação homóloga resulta na montagem de um polinucleotídeo; e

(b) selecionar uma célula hospedeira que compreende um polinucleotídeo montado, em que o polinucleotídeo montado compreende em uma orientação de 5' a 3', um segmento de DNA selecionado de cada um dos grupos  $D_0, \dots, D_n, \dots, D_m$ .

3. Método de acordo com a reivindicação 1, caracterizado pelo fato de que cada um dos locais de restrição RA e RB é clivável por uma endonuclease de restrição Tipo IIS.

4. Método de acordo com a reivindicação 1 ou 2, caracterizado pelo fato de que dito um ou mais endonucleases de restrição são uma ou mais endonucleases de restrição do Tipo IIS.

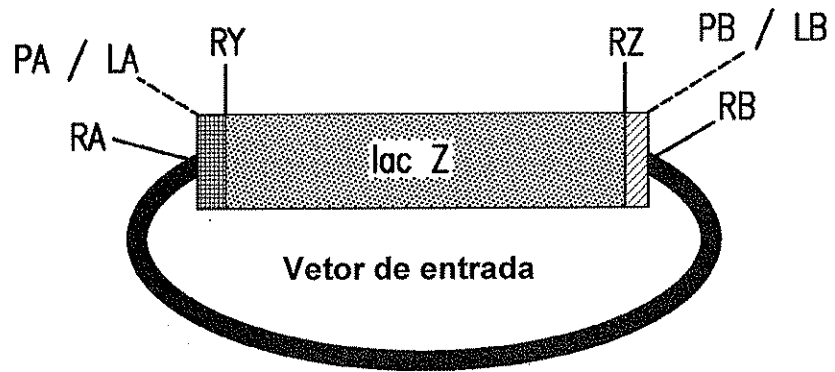


FIG.1A

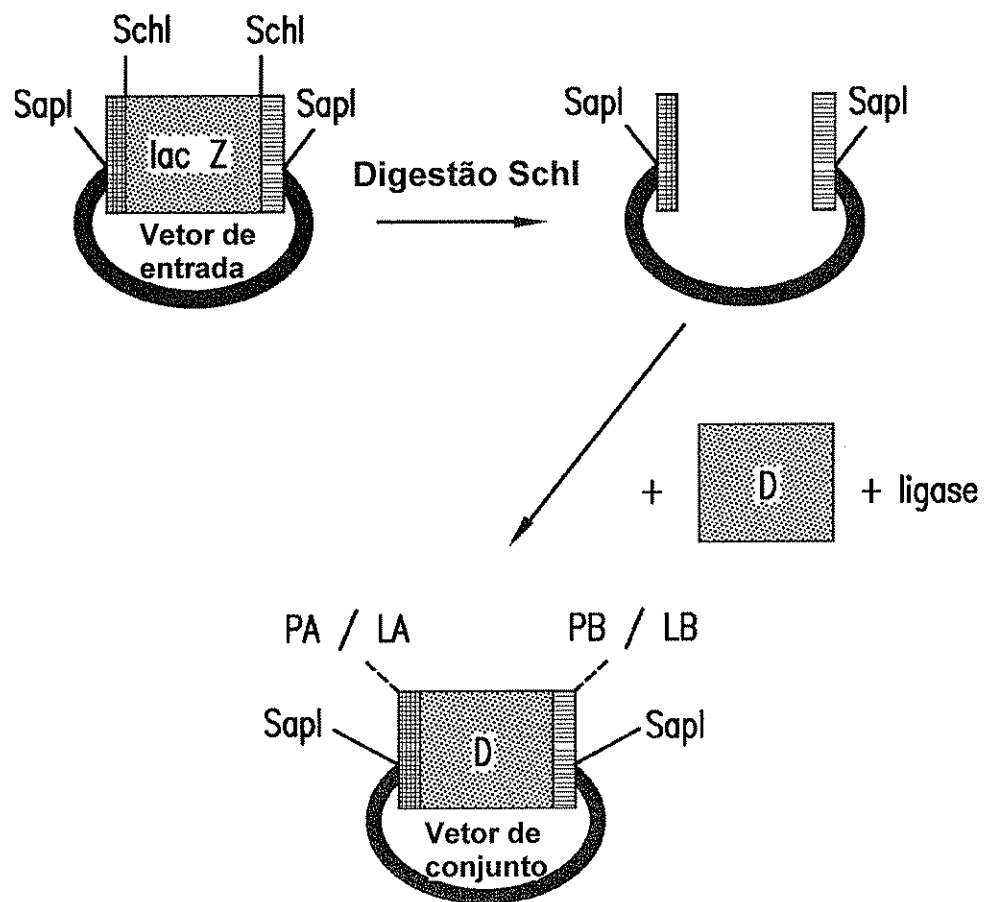


FIG.1B

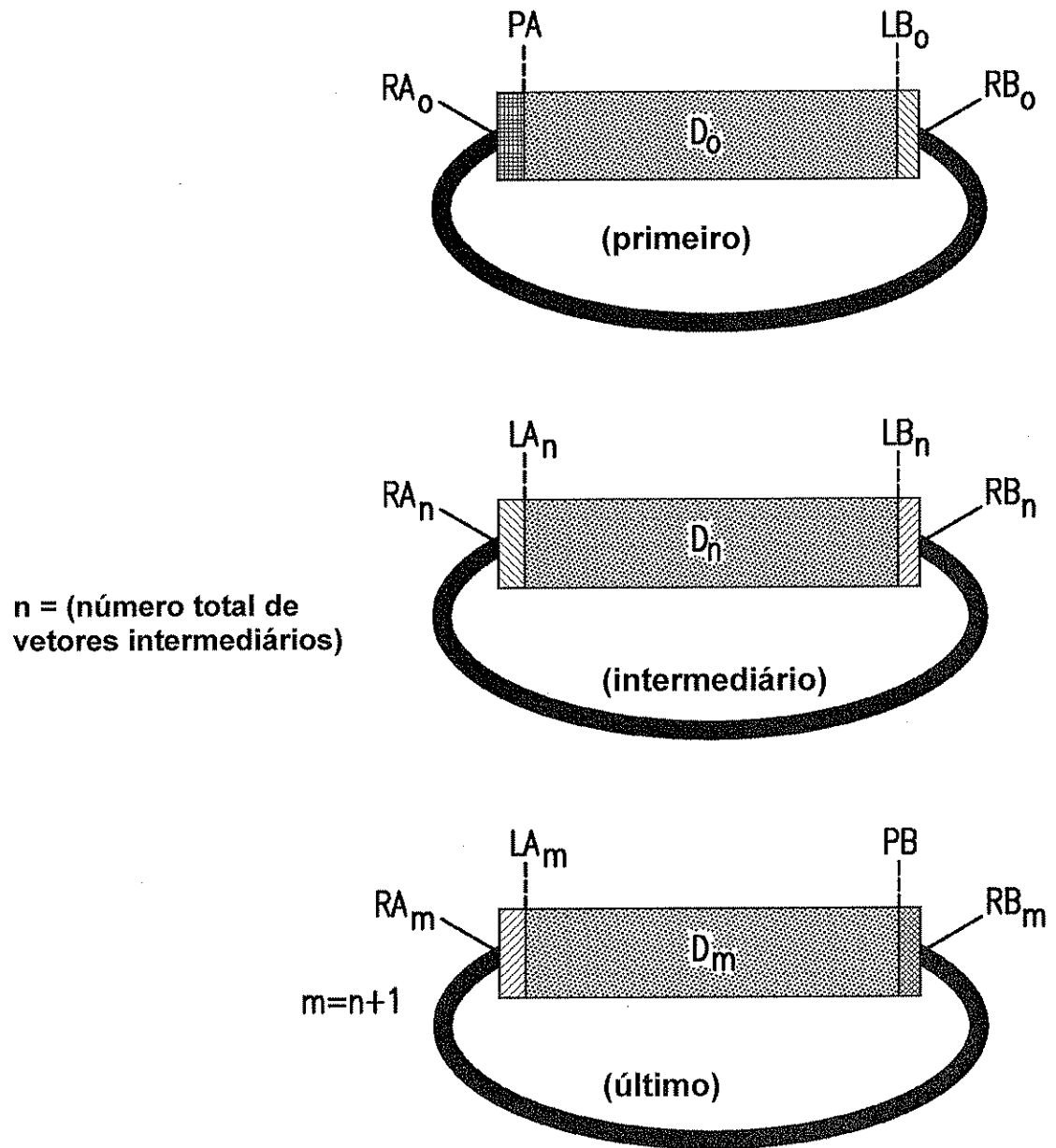


FIG.2

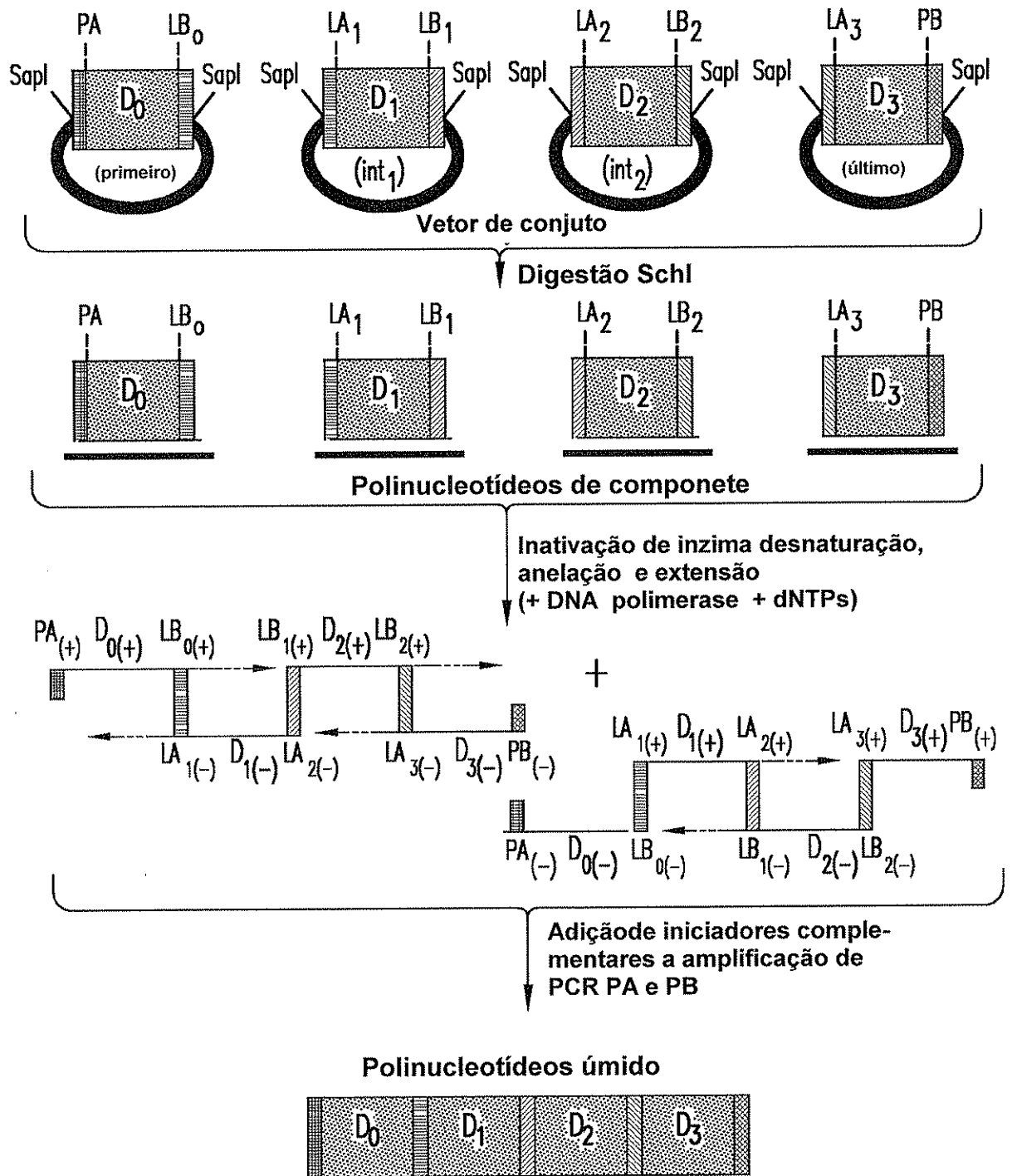


FIG.3

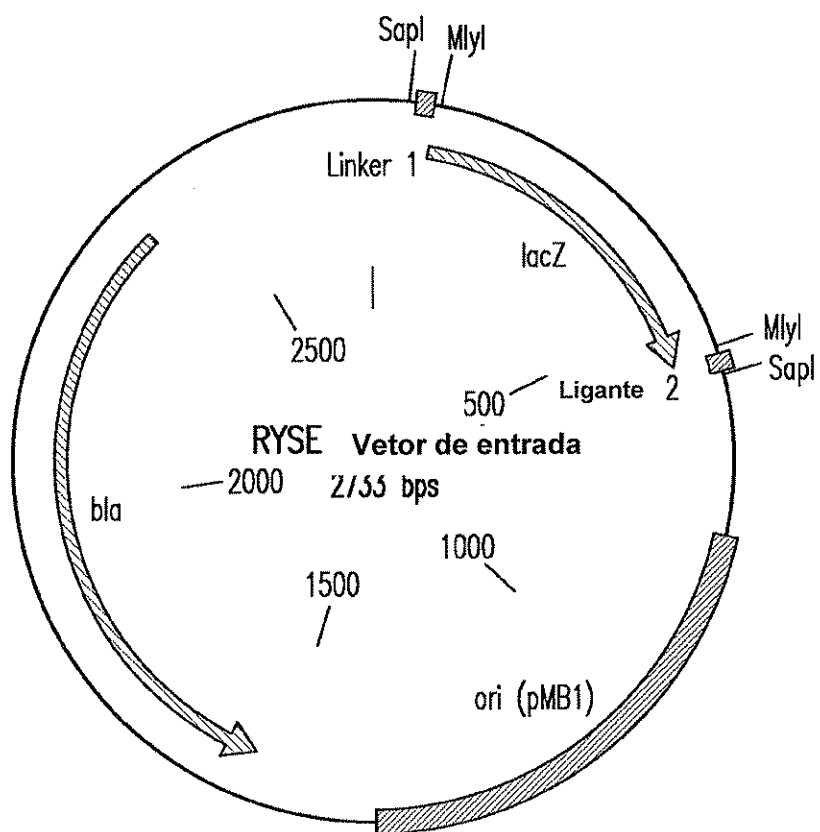


FIG.4

Purificação de  
coluna

Inativação por  
calor

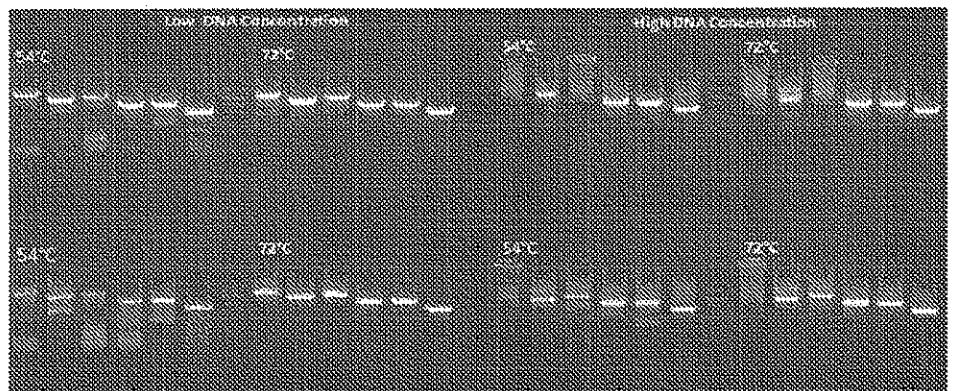


FIG.5

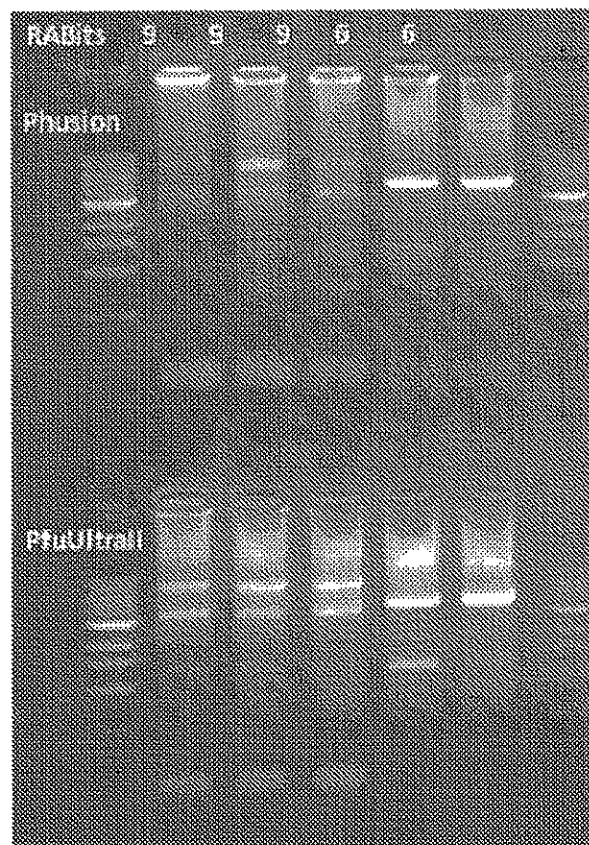


FIG.6



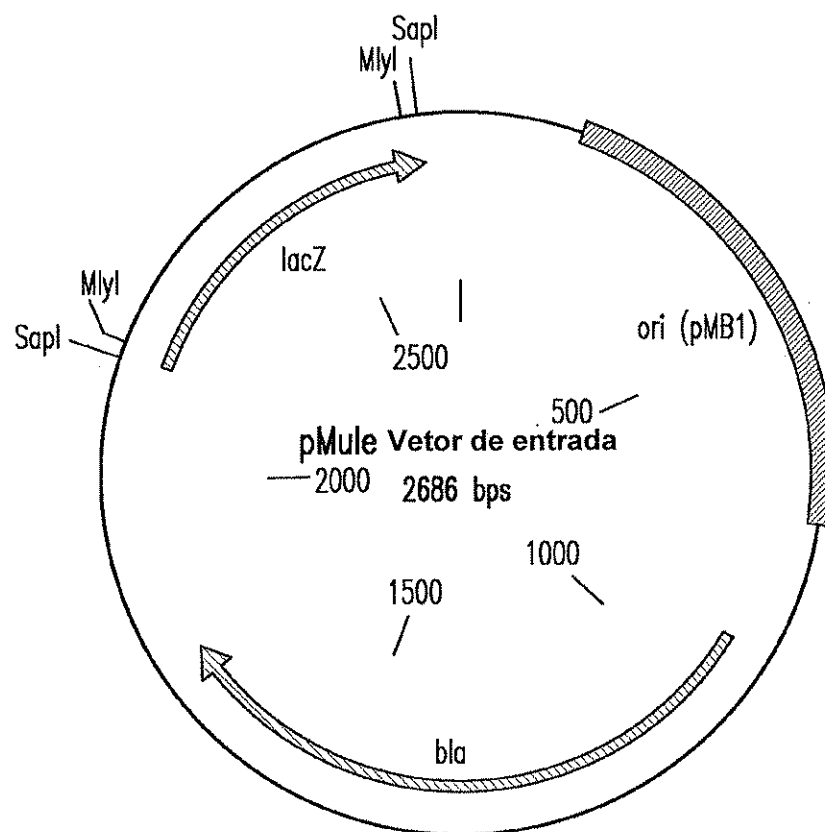


FIG. 7

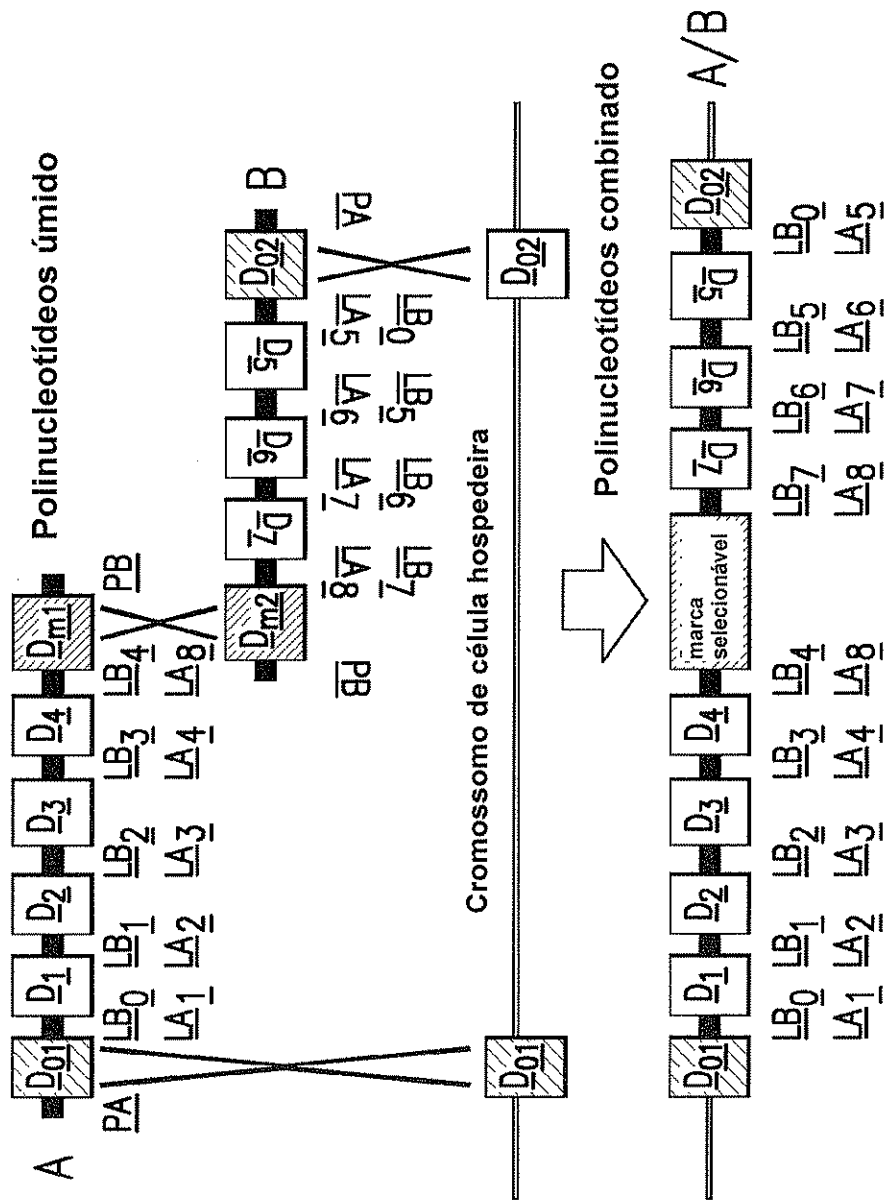


FIG.8

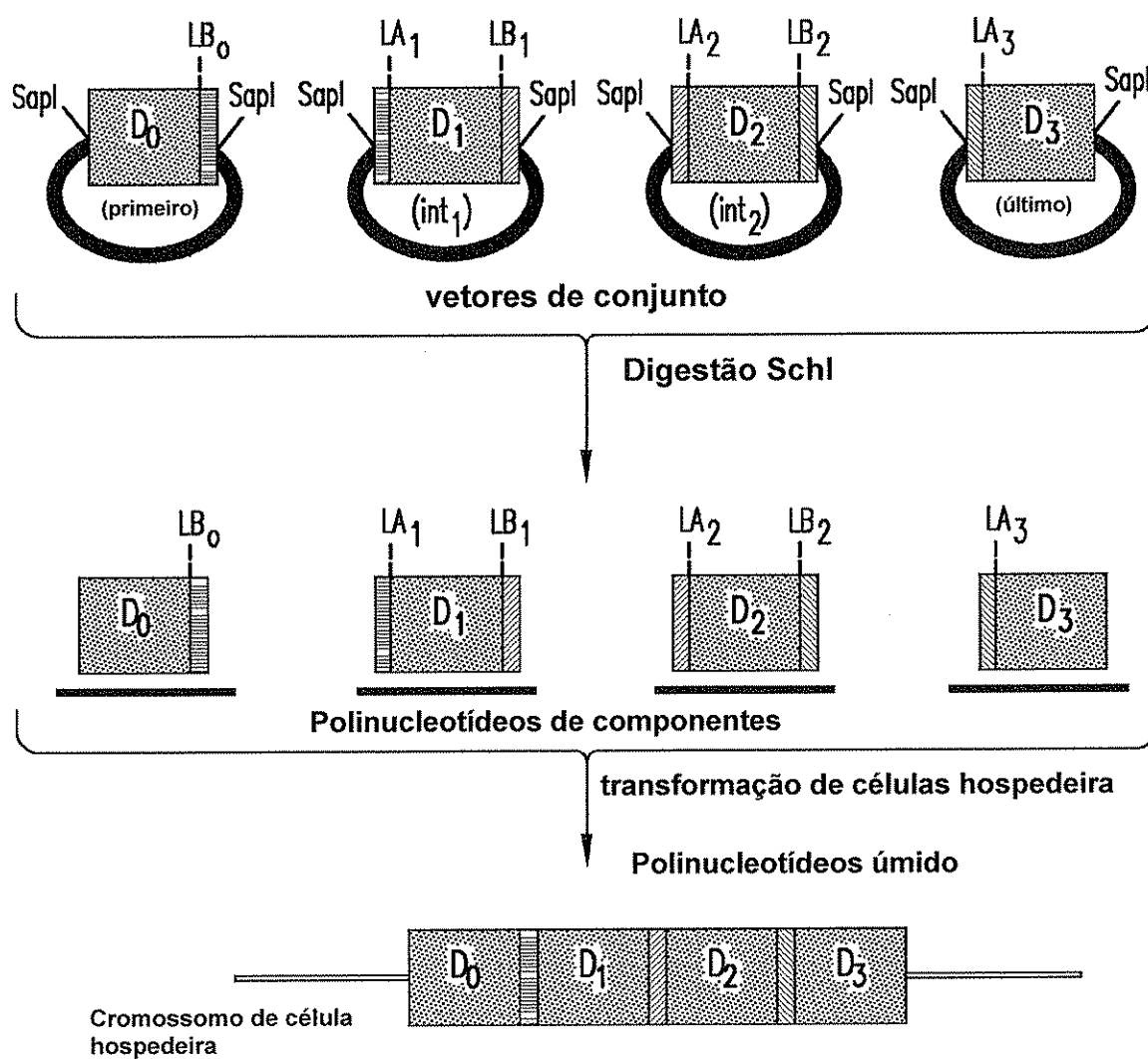


FIG.9

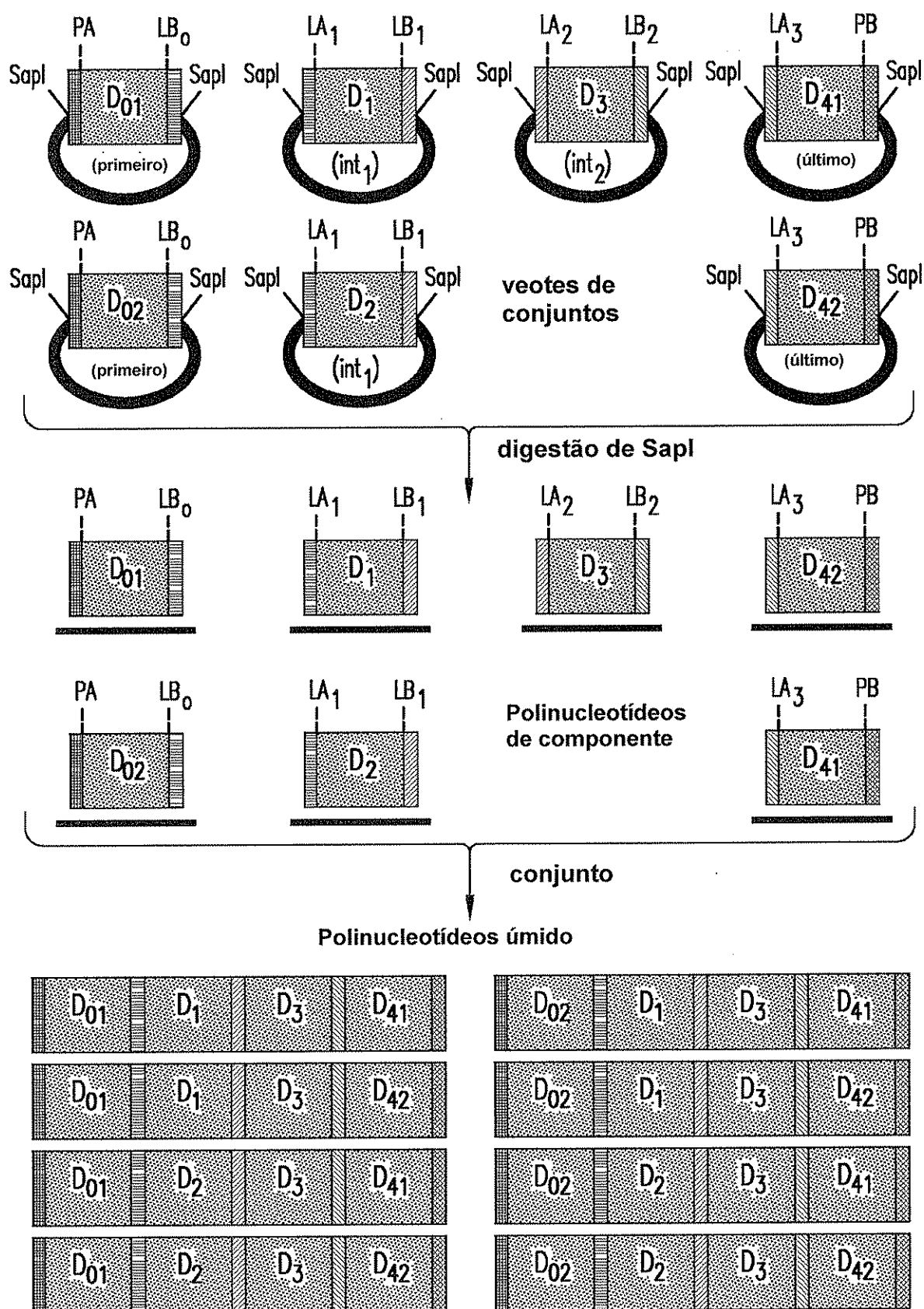


FIG.10

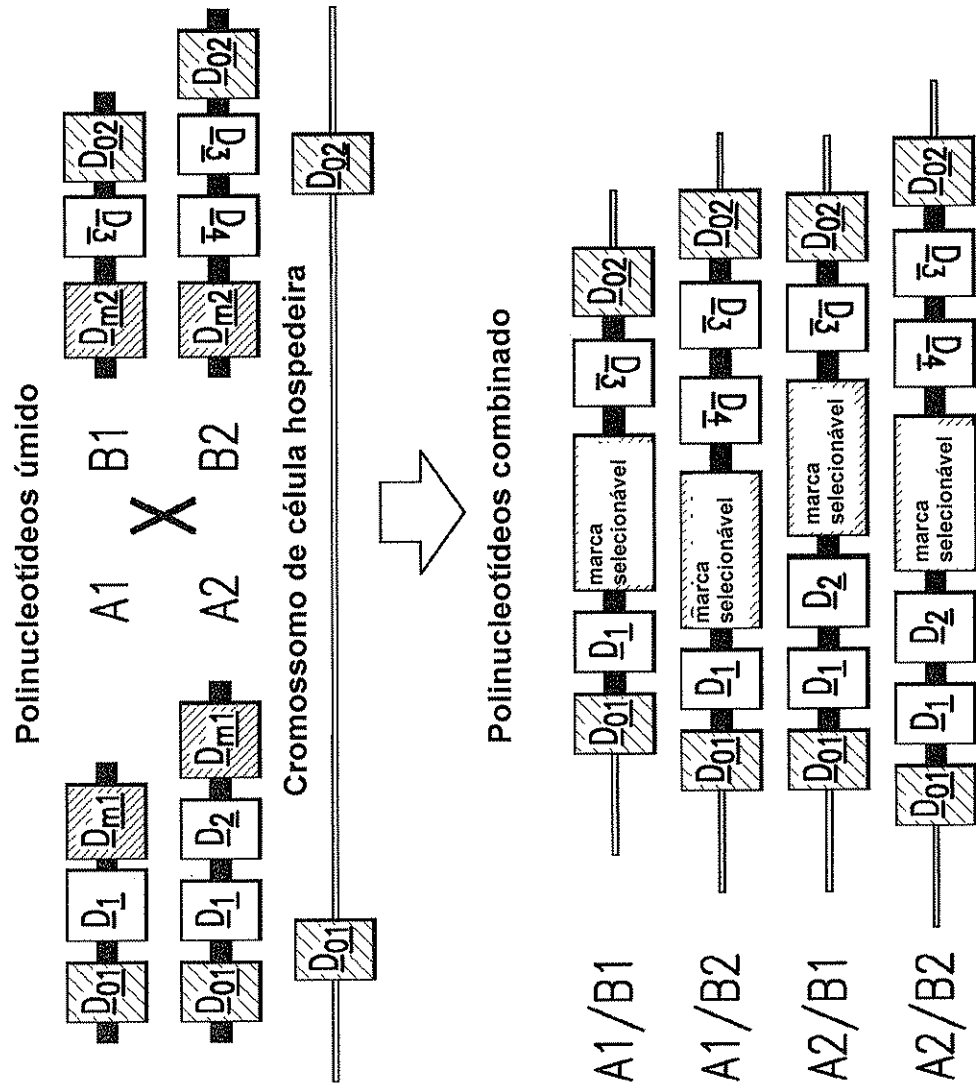


FIG.11

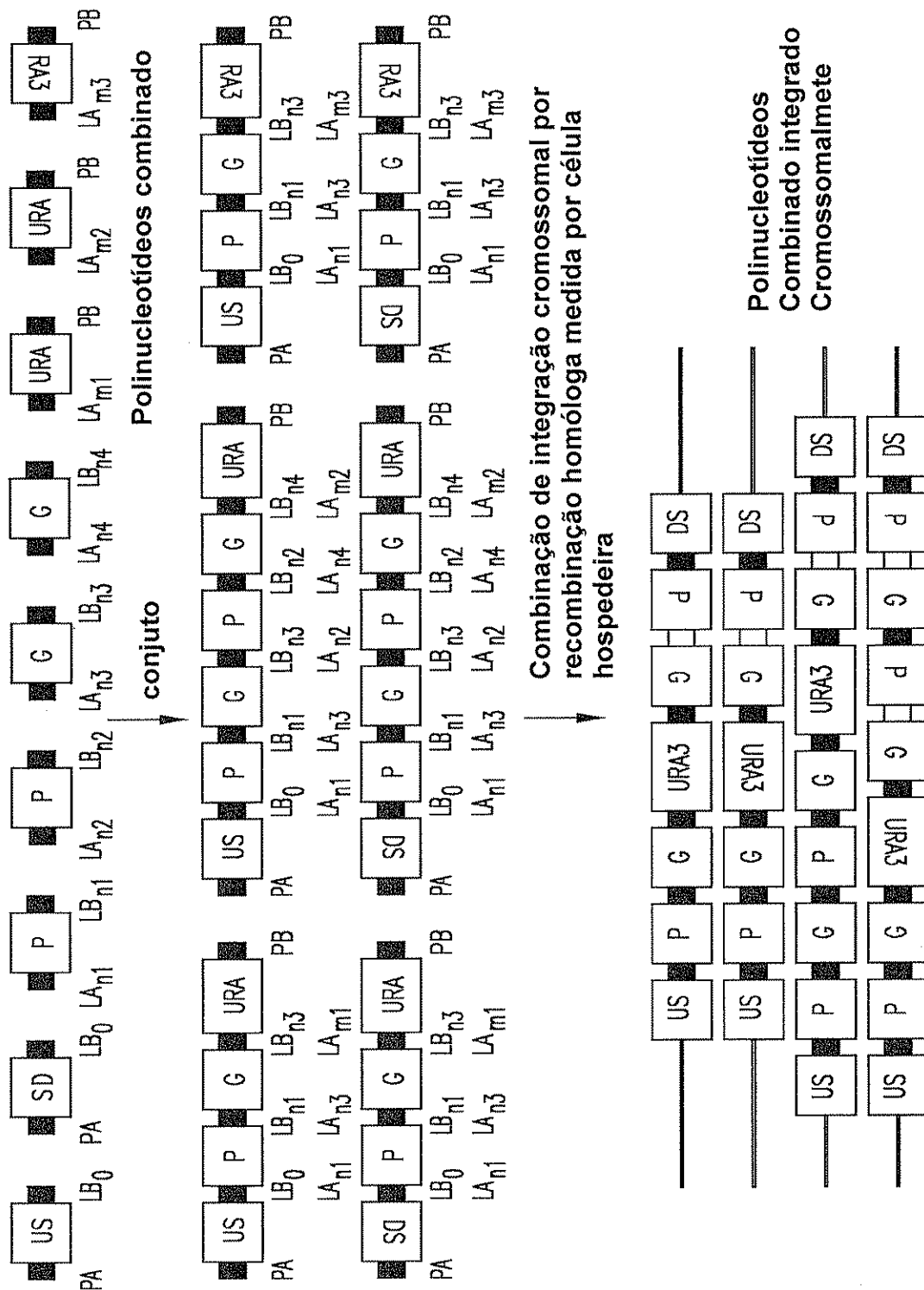
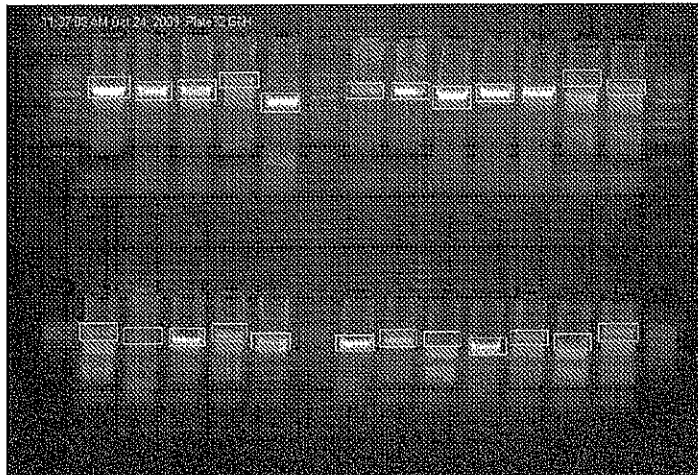


FIG.12A

Top panel									
Escala		1621	3909	<div> <div>Tamanho de Polinucleotídeo úmido (bp)</div> <div>Número de Polinucleotídeos de componentes úmido</div> <div>Número de conjunto</div> </div>					
		1622	3896						
		1623	4027						
		1624	6425						
		1625	2709						
		1626	5310						
Escala		1627	5415						
		1628	3518						
		1629	3874						
		1630	3933						
		1631	8459						
		1632	5370						
Escala									



Bottom panel	<b>Escala</b>		5877	6	1633	<b>Número de conjunto</b>		
			3848	4	1634		<b>Número de Polinucleotídeos de componentes úmido</b>	
			4035	4	1635			<b>Tamanho de Polinucleotídeo úmido (bp)</b>
			5135	6	1636			
			3603	4	1637			
	<b>Escala</b>		3516	4	1638			
			6457	6	1639			
			5492	6	1640			
			3458	4	1641			
			5453	6	1642			
		4758	6	1643				
<b>Escala</b>		6818	6	1644				

FIG. 12B

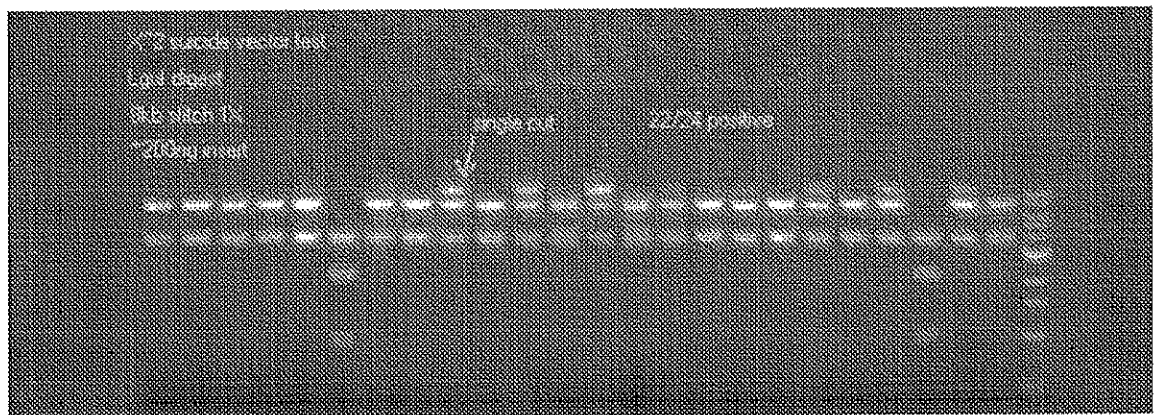


FIG.12C



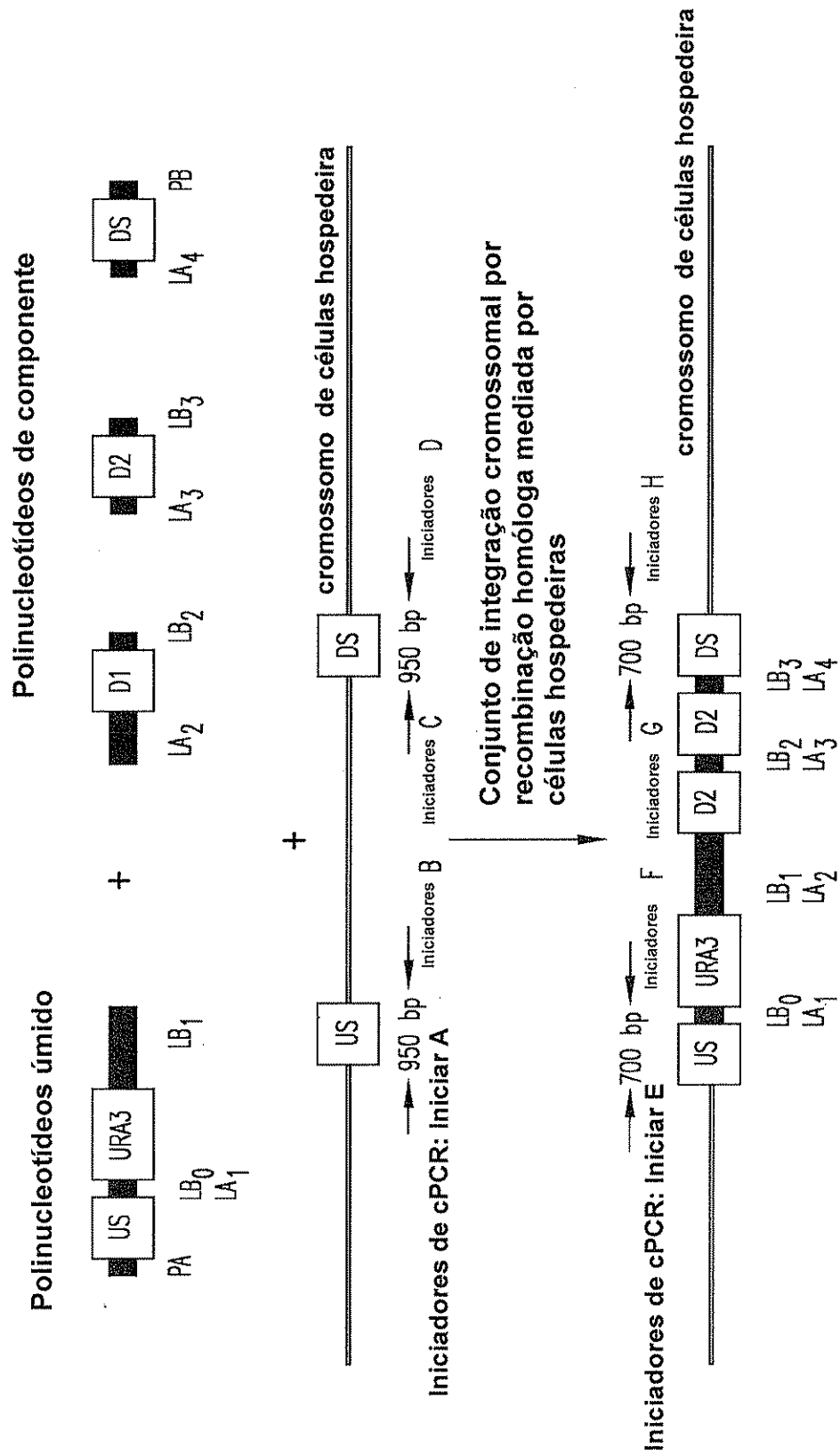


FIG.13A

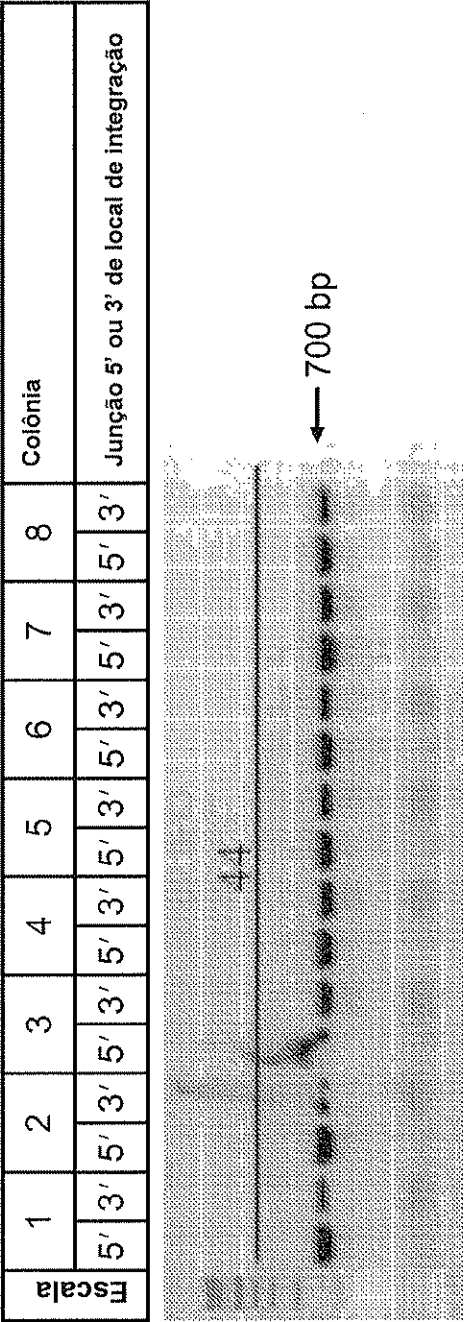
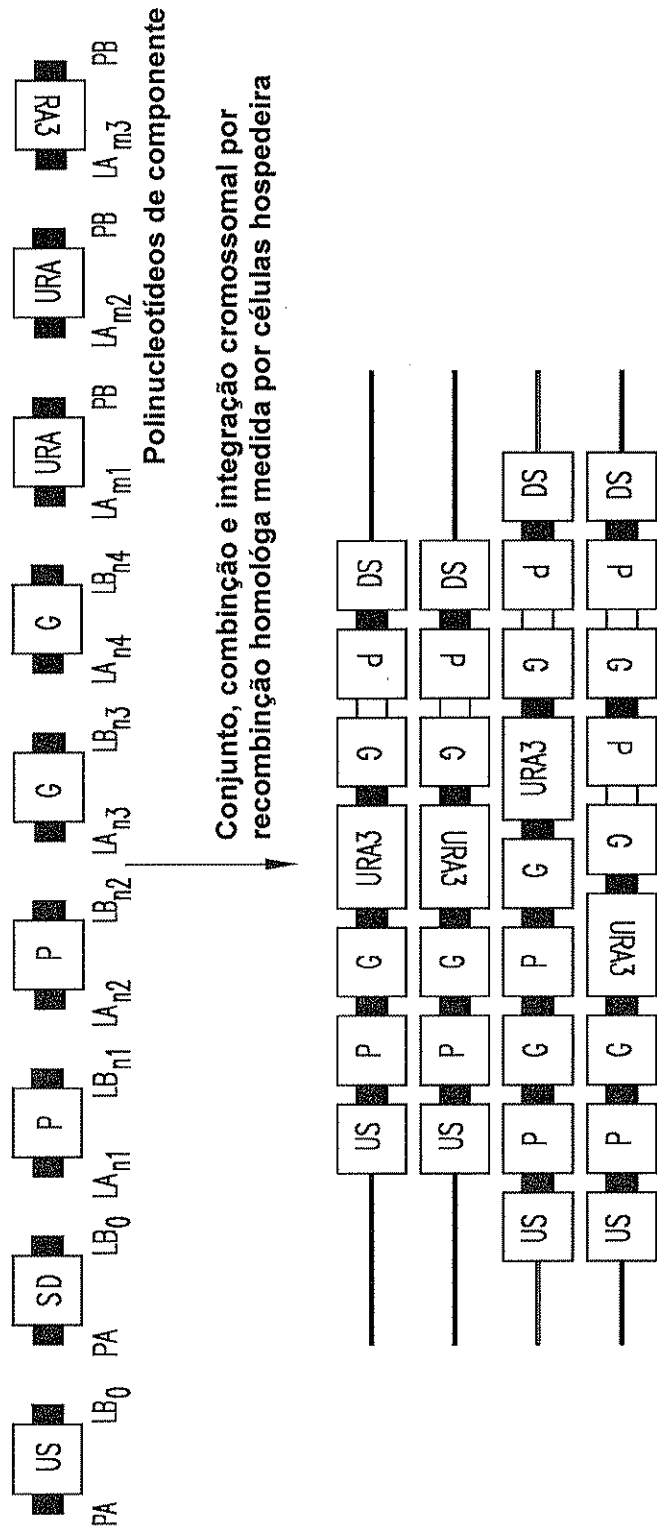


FIG.13B



Polinucleotídeos combinados cromossomalmente in tegrados

FIG.14