República Federativa do Brasil

Ministério do Desenvolvimento, Indústria, Comércio e Serviços

Instituto Nacional da Propriedade Industrial

(11) BR 112017006542-8 B1

(22) Data do Depósito: 29/09/2015

(45) Data de Concessão: 14/11/2023

(54) Título: MUTANTES DE RECOMBINASE

(51) Int.Cl.: C07K 14/01; C12Q 1/68; C12N 15/11; B01J 19/00.

(30) Prioridade Unionista: 29/09/2014 US 62/057,056.

(73) Titular(es): ILLUMINA CAMBRIDGE LIMITED.

(72) Inventor(es): ERIN BOMATI; MATTHEW WILLIAM KELLINGER; JONATHAN MARK BOUTELL.

(86) Pedido PCT: PCT US2015053012 de 29/09/2015

(87) Publicação PCT: WO 2016/054088 de 07/04/2016

(85) Data do Início da Fase Nacional: 29/03/2017

(57) Resumo: MUTANTES DE RECOMBINASE. A presente invenção se refere a recombinases para amplificação de ácidos nucleicos mediada por recombinase aprimoradas, tal como uma biblioteca de PCR que possui regiões de adaptadores de cadeias únicas, em uma superfície de células de fluxo padronizado para amplificação de agrupamento aprimorado, assim como métodos e kits utilizado as mesmas.

MUTANTES DE RECOMBINASE

ANTECEDENTES

[0001] As enzimas recombinases são úteis na amplificação de ácidos nucleicos mediada por recombinase. Por exemplo, as enzimas recombinases podem facilitar o direcionamento de oligonucleotídeos para alvos de DNA para permitir a replicação do DNA por uma polimerase. Permanece uma necessidade de recombinases modificadas com propriedades aprimoradas.

LISTAGEM DE SEQUÊNCIAS

[0002] O presente pedido de patente é depositado junto com uma Listagem de Sequências no formato eletrônico. A Listagem de Sequências é fornecida na forma de um arquivo intitulado IP1264.TXT, criado em 26 de setembro de 2014, que tem o tamanho de 98Kb. A informação no formato eletrônico da Listagem de Sequências é incorporada aqui como referência em sua totalidade.

BREVE SUMÁRIO

[0003] São apresentadas aqui recombinases para amplificação de ácidos nucleicos mediada por recombinase aprimorada. Os presentes inventores identificaram de forma surpreendente certas recombinases alteradas que possuem características substancialmente aprimoradas na distribuição de ácidos nucleicos sobre uma superfície de célula de fluxo (*flow cell*) padronizada. Em certas modalidades, as recombinases alteradas de distribuição aprimorada de uma biblioteca sem a necessidade de PCR, tal como uma biblioteca de PCR que possui regiões adaptadoras de filamento simples, sobre uma superfície de célula de fluxo padronizada para amplificação aprimorada de clusters.

[0004] Em certas modalidades, a recombinase é uma UvsX recombinante e compreende uma mutação de substituição de aminoácido na posição funcionalmente equivalente a Pro256 na sequência de aminoácidos de UvsX de RB49. A sequência de aminoácidos de UvsX de RB49 do tipo selvagem é apresentada na SEQ ID NO: 1. Em certas modalidades, a UvsX recombinante compreende uma sequência de aminoácidos que compreende um aminoácido que é pelo menos 60%, 70%, 80%, 90%, 95%, 99% idêntico à SEQ ID NO: 1 e compreende uma mutação de substituição de aminoácido na posição funcionalmente equivalente a Pro256 na sequência de aminoácidos de UvsX de RB49. Em certas modalidades, a mutação de substituição compreende uma mutação para um resíduo carregado. Em certas modalidades, a mutação de substituição compreende uma mutação para um resíduo básico. Em certas modalidades, a mutação de substituição compreende uma mutação homóloga a Pro256Lys na sequência de aminoácidos de UvsX de RB49.

[0005] Em algumas modalidades, em adição às mutações anteriores, a UvsX recombinante pode compreender ainda mutações de substituição nas posições funcionalmente equivalentes a His63 na sequência de aminoácidos de UvsX de RB49. Por exemplo, em certas modalidades, a UvsX recombinante compreende uma mutação de substituição homóloga a His63Ser na sequência de aminoácidos de UvsX de RB49.

[0006] Em algumas modalidades, em adição a qualquer uma das mutações acima, a UvsX recombinante pode compreender ainda uma mutação selecionada do grupo que consiste de: a adição de um ou mais resíduos de ácido

glutâmico no terminal C; a adição de um ou mais resíduos de ácido aspártico no terminal C; e uma combinação das mesmas.

[0007] Em algumas modalidades, a UvsX recombinante é derivada de um fago de Myoviridae selecionado do grupo que consiste de: T4, T6, Rb69, Aeh1, KVP40, fago 133 de *Acinetobacter*, fago de *Aeromonas* 65, cianofago P-SSM2, cianofago PSSM4, cianofago S-PM2, Rb32, fago nt-1 de *Vibrio*, Rbl6, Rb43 e Rb49.

[0008] Em algumas modalidades, a UvsX recombinante é derivada de um fago de Myoviridae selecionado do grupo que consiste de: T2, Rbl4, fago 25 de *Aeromonas*, phi-1, Fago 31, fago 44RR2.8t, fago Rb3 e fago LZ2.

[0009] Também é apresentada aqui uma UvsX recombinante que compreende a sequência de aminoácidos de qualquer uma das SEQ ID NOs: 2 e 22-35. Em certas modalidades, a UvsX recombinante compreende uma sequência de aminoácidos que compreende um aminoácido que é pelo menos 60%, 70%, 80%, 90%, 95%, 99% idêntico a qualquer uma das SEQ ID NOs: 2 e 22-35 e que compreende uma mutação de substituição de aminoácido na posição funcionalmente equivalente a Pro256 na sequência de aminoácidos de UvsX de RB49.

[0010] Também é apresentada aqui uma UvsX recombinante que compreende uma mutação de substituição no domínio semiconservado que compreende a sequência de aminoácidos de qualquer uma das SEQ ID NOs: 3-5 em que a mutação de substituição compreende uma mutação selecionada de uma substituição na posição 7 para qualquer resíduo diferente de Phe, Pro, Asp, Glu ou Asn. Em certas modalidades, a UvsX recombinante compreende um aminoácido que é pelo menos 60%, 70%, 80%, 90%, 95%, 99% idêntico a

uma recombinase que compreende o domínio semiconservado que compreende a sequência de aminoácidos de qualquer uma das SEQ ID NOs: 3-5 e em que a UvsX recombinante compreende uma mutação de substituição selecionada de uma substituição na posição 7 para qualquer resíduo diferente de Phe, Pro, Asp, Glu ou Asn. Em certas modalidades, a mutação compreende uma mutação para um resíduo carregado. Em certas modalidades, a mutação compreende uma mutação para um resíduo básico. Em certas modalidades, a mutação compreende uma substituição na posição 7 para Lys.

[0011] Também é apresentada aqui uma UvsX recombinante que compreende uma mutação de substituição no domínio semiconservado que compreende a sequência de aminoácidos de qualquer uma das SEQ ID NOs: 6-7 em que a mutação de substituição compreende uma mutação selecionada de uma substituição na posição 12 para qualquer resíduo diferente de Phe, Pro, Asp, Glu ou Asn. Em certas modalidades, a UvsX recombinante compreende um aminoácido que é pelo menos 60%, 70%, 80%, 90%, 95%, 99% idêntico a uma recombinase que compreende o domínio semiconservado que compreende a sequência de aminoácidos de qualquer uma das SEQ ID NOs: 6-7 e em que a UvsX recombinante compreende uma mutação de substituição selecionada de uma substituição na posição 12 para qualquer resíduo diferente de Phe, Pro, Asp, Glu ou Asn. Em certas modalidades, a mutação compreende uma mutação para um resíduo carregado. Em certas modalidades, a mutação compreende uma mutação para um resíduo básico. Em certas modalidades, a mutação compreende uma substituição na posição 12 para Lys.

[0012] Em algumas modalidades, em adição às mutações anteriores, a UvsX recombinante pode compreender

ainda mutações de substituição nas posições funcionalmente equivalentes a His63 na sequência de aminoácidos de UvsX de RB49. Por exemplo, em certas modalidades, a UvsX recombinante compreende uma mutação de substituição homóloga a His63Ser na sequência de aminoácidos de UvsX de RB49.

[0013] Em algumas modalidades, em adição a qualquer uma das mutações anteriores, a UvsX recombinante pode compreender ainda uma mutação selecionada do grupo que consiste de: a adição de um ou mais resíduos de ácido glutâmico no terminal C; a adição de um ou mais resíduos de ácido aspártico no terminal C; e uma combinação das mesmas.

[0014] Em algumas modalidades, a UvsX recombinante é derivada de um fago de Myoviridae selecionado do grupo que consiste de: T4, T6, Rb69, Aeh1, KVP40, fago 133 de *Acinetobacter*, fago de *Aeromonas* 65, cianofago P-SSM2, cianofago PSSM4, cianofago S-PM2, Rb32, fago nt-1 de *Vibrio*, Rbl6, Rb43 e Rb49.

[0015] Em algumas modalidades, a UvsX recombinante é derivada de um fago de Myoviridae selecionado do grupo que consiste de: T2, Rbl4, fago 25 de *Aeromonas*, phi-1, Fago 31, fago 44RR2.8t, fago Rb3 e fago LZ2.

[0016] Também é apresentada aqui uma molécula de ácido nucleico que codifica uma UvsX recombinante que é definida em qualquer uma das modalidades anteriores. Também é apresentado aqui um vetor de expressão que compreende a molécula de ácido nucleico descrita anteriormente. Também é apresentada aqui uma célula hospedeira que compreende o vetor descrito anteriormente.

[0017] Também é apresentado aqui um processo de amplificação por recombinase polimerase de amplificação de

uma molécula de ácido nucleico alvo, que compreende as etapas de: (a) colocar em contato a UvsX recombinante de qualquer uma das modalidades anteriores com um primeiro e um segundo iniciador de ácido nucleico para formar um primeiro e um segundo iniciadores de nucleoproteína, em que o dito iniciador de ácido nucleico compreende uma região de filamento simples em sua extremidade a 3’; (b) colocar em contato o primeiro e o segundo iniciador de nucleoproteína com a dita molécula de ácido nucleico alvo formando assim uma primeira estrutura de filamento duplo em uma primeira porção do dito primeiro filamento e formando uma segunda estrutura de filamento duplo em uma segunda porção do dito segundo filamento de forma que as extremidades a 3' do dito primeiro iniciador de ácido nucleico e do dito segundo iniciador de ácido nucleico fiquem orientadas em direção uma com a outra na mesma molécula de ácido nucleico molde de filamento duplo; (c) estender a extremidade 3’ dos ditos primeiro e segundo iniciadores de ácido nucleico com uma ou mais polimerases e dNTPs para gerar um primeiro e um segundo ácido nucleico de filamento duplo e um primeiro e um segundo filamentos de ácido nucleico deslocados; e (d) continuar da reação através da repetição de (b) e (c) até um grau desejado de amplificação ser atingido.

[0018] Em certas modalidades do processo, a molécula de ácido nucleico alvo compreende ácido nucleico de filamento duplo. Em certas modalidades, a molécula de ácido nucleico alvo compreende ácido nucleico de filamento simples. Por exemplo, em algumas modalidades, o ácido nucleico alvo compreende uma região adaptadora de filamento simples. Em certas modalidades, o processo é realizado na presença de uma proteína de carregamento da recombinase.

Por exemplo, a proteína de carregamento da recombinase pode ser selecionada do grupo que consiste de UvsY de T4, reco de *E. coli*, recR de *E. coli* e uma combinação das mesmas. Em certas modalidades, o processo é realizado na presença de um agente estabilizador de filamento simples selecionado do grupo que consiste de gp32, proteína SSB de *E. coli*, proteína gp32 de T4 e derivados das mesmas. Em certas modalidades, o processo é realizado na presença de um agente de aglomeração selecionado do grupo que compreende polietileno glicol, óxido de polietileno, poliestireno, Ficoll, dextran, PVP e albumina de forma que o agente de aglomeração estimule a amplificação.

[0019] Em certas modalidades, o processo é realizado em um arranjo de sítios de amplificação. Em certas modalidades, cada sítio de amplificação compreende um grande número de iniciadores de amplificação para a amplificação do ácido nucleico alvo. Em certas modalidades, o arranjo de sítios de amplificação compreende um arranjo de características sobre uma superfície. Por exemplo, as características podem ser não contíguas e podem ser separadas por regiões intersticiais da superfície que não possuem os iniciadores de amplificação. Em certas modalidades, o arranjo de sítios de amplificação compreende esferas em solução ou esferas sobre uma superfície. Em certas modalidades, o arranjo de sítios de amplificação compreende uma emulsão. Em certas modalidades, o processo ocorre de forma isotérmica.

[0020] Também é apresentado aqui um kit para a realização de uma reação de recombinase polimerase. Em certas modalidades, o kit pode compreender uma UvsX recombinante que é definida em qualquer uma das modalidades

anteriores e um ou mais dos seguintes: uma proteína de ligação ao DNA de filamento simples; uma DNA polimerase; dNTPs ou uma mistura de dNTPs e ddNTPs; um agente de aglomeração; um tampão; um agente redutor; ATP ou análogo de ATP; uma proteína de carregamento da recombinase; um primeiro iniciador e opcionalmente um segundo iniciador.

[0021] Os detalhes de uma ou mais modalidades são apresentados nos desenhos em anexo e na descrição a seguir. Outras características, objetivos e vantagens serão evidentes partindo da descrição e dos desenhos e das reivindicações.

BREVE DESCRIÇÃO DOS DESENHOS

[0022] A figura 1 é uma representação esquemática que mostra o alinhamento de sequências de aminoácidos de UvsX do fago T4 de *Enterobacteria* (T4) (SEQ ID NO: 8), do fago T6 de *Enterobacteria* (T6) (SEQ ID NO: 9), do fago 133 de *Acinetobacter* (Fagol33) (SEQ ID NO: 10), do fago RB69 de *Enterobacteria* (Rb69) (SEQ ID NO: 11), do fago de *Aeromonas* Aeh1 (Aeh1) (SEQ ID NO: 12), do fago 65 de *Aeromonas* (Ae65) (SEQ ID NO: 13), do fago KVP40 de *Vibrio* (Kvp40) (SEQ ID NO: 14), do fago RB43 de *Enterobacteria* (Rb43) (SEQ ID NO: 15), do fago P-SSM2 de *Prochlorococcus* (PSSM2) (SEQ ID NO: 16) e do fago P-SSM4 de *Prochlorococcus* (PSSM4) (SEQ ID NO: 17), que também são apresentados nos materiais incorporados da US 2009/0029421. Os resíduos que são equivalentes em relação à posição e/ou à função a Pro256 na sequência de aminoácidos de UvsX de RB49 estão destacados e indicados por um símbolo triangular.

[0023] A Figura 2 é uma representação esquemática que mostra o alinhamento de sequências de aminoácidos de UvsX do fago RB49 de *Enterobacteria* (RB49) (SEQ ID NO: 1) e

do fago T4 de *Enterobacteria* (T4) (SEQ ID NO: 8). Os resíduos que são equivalentes em relação à posição e/ou à função a Pro256 na sequência de aminoácidos de UvsX de RB49 estão destacados e indicados por um símbolo triangular.

[0024] A Figura 3 A mostra uma fotografia de uma imagem de cluster de uma biblioteca sem a necessidade de PCR distribuída sobre uma célula de fluxo padronizada utilizando uma formulação de UvsX de T4.

[0025] A Figura 3B mostra uma fotografia de uma imagem de cluster de uma biblioteca sem a necessidade de PCR distribuída sobre uma célula de fluxo padronizada utilizando uma formulação líquida que inclui RB49 P256K recombinase.

[0026] A Figura 4A mostra uma fotografia de uma imagem de cluster de uma biblioteca sem a necessidade de PCR de filamento simples (ssDNA) distribuída sobre uma célula de fluxo padronizada utilizando uma formulação de UvsX de T4.

[0027] A Figura 4B mostra uma fotografia de uma imagem de cluster de uma biblioteca sem a necessidade de PCR de filamento simples distribuída sobre uma célula de fluxo padronizada utilizando uma formulação líquida que inclui RB49 P256K recombinase.

DESCRIÇÃO DETALHADA

[0028] São apresentadas aqui recombinases para a amplificação de ácidos nucleicos mediada por recombinase aprimorada. Os presentes inventores identificaram de forma surpreendente certas recombinases alteradas que possuem características substancialmente aprimoradas na distribuição de ácidos nucleicos sobre uma superfície de célula de fluxo padronizada.

[0029] Como descrito em maiores detalhes posteriormente aqui, os inventores descobriram de forma surpreendente que uma ou mais mutações em um ou mais resíduos na recombinase resultam em melhorias profundas na distribuição de uma biblioteca de DNA, tal como, por exemplo, uma biblioteca de PCR que possui regiões adaptadoras de filamento simples, sobre uma superfície de célula de fluxo padronizada, fornecendo amplificação de cluster aprimorada.

[0030] Em certas modalidades, a mutação de substituição compreende uma mutação em um resíduo que possui uma cadeia lateral carregada. Por exemplo, em algumas modalidades, o aminoácido carregado é um resíduo de aminoácido carregado positivamente. O termo “aminoácido carregado positivamente” se refere a um aminoácido hidrofílico com um valor de pKa de cadeia lateral maior que 7, a saber um aminoácido básico. Os aminoácidos básicos tipicamente possuem cadeias laterais carregadas positivamente em pH fisiológico devido à associação com um íon de hidrônio. Os aminoácidos básicos que ocorrem naturalmente (codificados geneticamente) incluem lisina (Lys, K), arginina (Arg, R) e histidina (His, H), enquanto que os aminoácidos básicos não naturais (não codificados geneticamente ou não-padronizado) incluem, por exemplo, ornitina, ácido 2,3,-diaminopropiônico, ácido 2,4-diaminobutírico, ácido 2,5,6-triaminohexanóico, ácido 2-amino-4-guanidinobutanóico e homoarginina. O termo “aminoácido carregado negativamente” refere-se a um aminoácido natural ou não natural, independentemente da quiralidade, que contém, em adição ao grupo carboxila C-terminal, pelo menos um grupo carregado negativamente

adicional tal como carboxila, fosfato, fosfonato, sulfonato ou similar.

[0031] Também é apresentada aqui uma UvsX recombinante que compreende uma mutação de substituição em um domínio semiconservado da UvsX recombinante. Como é utilizado aqui, o termo “domínio semiconservado” refere-se a uma porção da UvsX recombinante que é completamente conservada ou pelo menos parcialmente conservada entre várias espécies. Foi descoberto de forma surpreendente que uma mutação de um ou mais resíduos no domínio semiconservado afeta a atividade da recombinase especialmente na presença de ácido nucleico molde de filamento simples, resultando no aprimoramento da distribuição e/ou da amplificação em reações de amplificação mediada por recombinase. Estas recombinases submetidas à mutação possuem melhor desempenho na distribuição de bibliotecas sem a necessidade de PCR, tal como uma biblioteca de PCR que possui regiões adaptadoras de filamento simples, sobre uma superfície de célula de fluxo padronizada, resultando na amplificação de cluster aprimorada, que é descrita na seção de Exemplo a seguir.

[0032] Em algumas modalidades, o domínio semiconservado compreende aminoácidos que possuem a sequência apresentada em qualquer uma das SEQ ID NOs: 3-7. As SEQ ID NOs: 3-7 correspondem aos resíduos no domínio semiconservado entre várias espécies. A SEQ ID NO: 3 corresponde aos resíduos 251-258 da sequência de aminoácidos da UvsX de T4, que é apresentada aqui como a SEQ ID NO: 8. Um alinhamento que mostra a conservação entre várias espécies no domínio semiconservado é apresentado nas Figuras 1 e 2. As sequências de UvsX mostradas na Figura 1

foram obtidas nos números de acesso da base de dados do Genbank NP_049656 (T4), YP_004300647 (Fago 133); NP]_861734 (RB69); NP_943894.1 (Aeh1); YP_004300858 (Ae65); NP_899256 (KVP40); YP 239013 (RB43); YP_214417 (P-SSM2); YP_214708 (P-SSM4); e na Publicação US No. 2009/0029421 (T6). A Figura 2 é uma representação esquemática que mostra o alinhamento de sequências de aminoácidos de UvsX do fago RB49 de *Enterobacteria* (RB49) (SEQ ID NO: 1) e do fago T4 de *Enterobacteria* (T4) (SEQ ID NO: 8). Os resíduos que são equivalentes em relação à posição e/ou à função a Pro256 na sequência de aminoácidos de UvsX de RB49 estão destacados e indicados por um triângulo. As sequências de UvsX mostradas na Figura 2 foram obtidas nos números de acesso da base de dados do Genbank NP_891595 (RB49) e NP_049656 (T4).

[0033] Foi descoberto de forma surpreendente que as mutações em um ou mais resíduos no domínio semiconservado aumentam a atividade da recombinase especialmente na presença de ácido nucleico molde de filamento simples, resultando no aprimoramento da distribuição e/ou da amplificação na amplificação mediada por reações da recombinase. Estas recombinases submetidas à mutação possuem melhor desempenho na distribuição de bibliotecas sem a necessidade de PCR, tal como uma biblioteca de PCR que possui regiões adaptadoras de filamento simples, sobre uma superfície de célula de fluxo padronizada, resultando na amplificação de cluster aprimorada, que é descrita na seção de Exemplo a seguir. Por exemplo, em algumas modalidades de uma recombinante UvsX apresentada aqui, a mutação de substituição compreende uma mutação na posição 7 de qualquer uma das SEQ ID NOs: 3-5 para qualquer resíduo diferente de Phe, Pro, Asp, Glu ou Asn. Em certas

modalidades, a UvsX recombinante compreende uma mutação para Lys na posição 7 de qualquer uma das SEQ ID NOs: 3-5. Em algumas modalidades da UvsX recombinante apresentada aqui, a mutação de substituição compreende uma mutação na posição 12 de qualquer uma das SEQ ID NOs: 6-7 para qualquer resíduo diferente de Phe, Pro, Asp, Glu ou Asn. Em certas modalidades, a UvsX recombinante compreende uma mutação para Lys na posição 12 de qualquer uma das SEQ ID NOs: 6-7.

[0034] Em algumas modalidades, a recombinase é uma proteína UvsX. Qualquer recombinase de fago pode ser utilizada nas modalidades apresentadas aqui, incluindo, por exemplo, recombinases de fago tal como UvsX ou recombinase similar à UvsX derivada de um fago de Myoviridae tal como, por exemplo, T4, T6, Rb69, Aeh1, KVP40, fago 133 de *Acinetobacter*, fago de *Aeromonas* 65, cianofago P-SSM2, cianofago PSSM4, cianofago S-PM2, Rb32, fago nt-1 de *Vibrio*, Rbl6, Rb43 e Rb49. Em certas modalidades, a recombinase é uma UvsX ou uma recombinase similar à UvsX derivada de um fago de Myoviridae tal como, por exemplo, T2, Rbl4, fago 25 de *Aeromonas*, phi-1, Fago 31, fago 44RR2.8t, fago Rb3 e fago LZ2. Será facilmente evidente para um perito na arte que outras proteínas recombinases podem ser utilizadas nas modalidades apresentadas aqui. As proteínas recombinases adequadas podem ser identificadas por homologia à UvsX utilizando qualquer número de um número de métodos conhecidos na arte, tal como, por exemplo, alinhamento com BLAST, que é descrito em maiores detalhes a seguir.

[0035] Por “funcionalmente equivalente” entende-se que a recombinase de controle, no caso de estudos

utilizando uma recombinase totalmente diferente, irá conter a substituição de aminoácido que é considerada como ocorrendo na posição de aminoácido na outra recombinase que possui o mesmo papel funcional na enzima. Como um exemplo, a mutação na posição 257 de Fenilalanina para Lisina (F257K) na UvsX de T4 seria funcionalmente equivalente a uma substituição na posição 256 da Prolina para Lisina (P256K) na UvsX de RB49.

[0036] Geralmente as mutações de substituição funcionalmente equivalentes em duas ou mais recombinases diferentes ocorrem em posições de aminoácidos homólogos nas sequências de aminoácidos das recombinases. Assim, o uso aqui do termo “funcionalmente equivalente” abrange também mutações que são “equivalentes em relação à posição” ou “homólogas” a certa mutação, independentemente do fato da função particular do aminoácido submetido à mutação ser conhecida ou não. É possível identificar resíduos de aminoácidos equivalentes em relação à posição ou homólogos nas sequências de aminoácidos de duas ou mais recombinases diferentes com base no alinhamento de sequências e/ou na modelagem molecular. Um exemplo de alinhamento de sequências para identificar resíduos equivalentes em relação à posição e/ou funcionalmente equivalentes é apresentado na Figura 1, que apresenta um alinhamento de sequências de aminoácidos de UvsX do fago T4 de *Enterobacteria* (T4) (SEQ ID NO: 8), fago T6 de *Enterobacteria* (T6) (SEQ ID NO: 9), fago 133 de *Acinetobacter* (Fagol33) (SEQ ID NO: 10), fago RB69 de *Enterobacteria* (Rb69) (SEQ ID NO: 11), fago de *Aeromonas* Aeh1 (Aeh1) (SEQ ID NO: 12), fago de *Aeromonas* 65 (Ae65) (SEQ ID NO: 13), fago de *Vibrio* KVP40 (Kvp40) (SEQ ID NO:

14), fago RB43 de *Enterobacteria* (Rb43) (SEQ ID NO: 15), fago P-SSM2 de *Prochlorococcus* (PSSM2) (SEQ ID NO: 16) e fago P-SSM4 de *Prochlorococcus* (PSSM4) (SEQ ID NO: 17), que também são apresentados nos materiais incorporados da US 2009/0029421. As sequências de UvsX mostradas na Figura 1 foram obtidas nos números de acesso da base de dados do Genbank NP_049656 (T4), YP_004300647 (Fago 133); NP_861734 (RB69); NP_943894.1 (Aeh1); YP_004300858 (Ae65); NP_899256 (KVP40); YP_239013 (RB43); YP_214417 (P-SSM2); YP_214708 (P-SSM4); e na Publicação US No. 2009/0029421 (T6).

[0037] A Figura 2 é uma representação esquemática que mostra o alinhamento de sequências de aminoácidos de UvsX do fago RB49 de *Enterobacteria* (RB49) (SEQ ID NO: 1) e do fago T4 de *Enterobacteria* (T4) (SEQ ID NO: 8). Os resíduos que são equivalentes em relação à posição e/ou à função a Pro256 na sequência de aminoácidos de UvsX de RB49 estão destacados e indicados por um triângulo. As sequências de UvsX mostradas na Figura 2 foram obtidas nos números de acesso da base de dados do Genbank NP_891595 (RB49) e NP_049656 (T4).

[0038] Um resíduo equivalente em relação à posição e/ou funcionalmente equivalente pode ser determinado para uma ou mais de qualquer número de outras sequências de UvsX através do alinhamento de tais sequências com aquela de uma sequência de referência tal como T4 e RB49. Como um exemplo não limitante, as sequências de UvsX do fago S-PM2 de *Synechococcus*, do fago RB32 de *Enterobacteria*, fago nt-1 de *Vibrio*, do fago RB166 de *Enterobacteria* são apresentadas como as SEQ ID NOs: 18-21 e obtidas nos números de acesso da base de dados do Genbank YP_l95169.1; YP_802982.1; YP_008125207.1; YP_003858336.1, podem ser alinhadas com uma

sequência de UvsX de referência tal como, por exemplo, UvsX de T4 (SEQ ID NO: 8) e UvsX de RB49 (SEQ ID NO: 1) e resíduos equivalentes em relação à posição e/ou funcionalmente equivalentes são identificados. Com a finalidade de exemplo, os resíduos mostrados na tabela a seguir são identificados como equivalentes em relação à posição e/ou funcionalmente equivalentes a Pro256 na sequência de aminoácidos de UvsX de RB49. Será facilmente considerado por um técnico no assunto que posições equivalentes em relação à posição e/ou funcionalmente equivalentes para outras proteínas UvsX podem ser determinadas seguindo uma abordagem similar.

Espécie do Fago	SEQ ID NO:	Posição Equivalente em Relação à Posição/Função
T4	8	Phe257
T6	9	Phe259
Fago 133 de *Acinetobacter*	10	Pro257
Rb69	11	Pro258
Aeh1	12	Pro269
Fago 65 de *Aeromonas*	13	Asp266
KVP40	14	Pro267
Rb43	15	Pro259
cianofago P-SSM2	16	Gln261
cianofago PSSM4	17	Glu264
cianofago S-PM2	18	Glu264
Rb32	19	Phe259
fago nt-1 de	20	Pro267

Vibrio		
Rb16	21	Pro259

[0039] As proteínas UvsX recombinantes descritas anteriormente aqui podem compreender mutações de substituição adicionais que são conhecidas por aprimorar um ou mais aspectos da atividade, da estabilidade ou de qualquer outra propriedade desejada da recombinase. Por exemplo, em algumas modalidades, em adição a qualquer uma das mutações anteriores, a UvsX recombinante pode compreender ainda mutações de substituição nas posições funcionalmente equivalentes a His63 na sequência de aminoácidos de UvsX de RB49 como é conhecido na arte e exemplificado pela revelação da US 2009/0029421, que é incorporada como referência em sua totalidade. Por exemplo, em certas modalidades, a UvsX recombinante compreende uma mutação de substituição homóloga a His63Ser na sequência de aminoácidos de UvsX de RB49.

[0040] Em algumas modalidades, em adição a qualquer uma das mutações anteriores, a UvsX recombinante pode compreender mutações de substituição, deleção e/ou adição adicionais quando comparada a uma recombinase do tipo selvagem. Qualquer uma de uma variedade de mutações de substituição em uma ou mais das posições pode ser realizada, como é conhecido na arte e exemplificado pelos materiais incorporados da 2009/0029421. Por exemplo, em algumas modalidades, em adição às mutações anteriores, a UvsX recombinante pode compreender ainda uma mutação selecionada do grupo que consiste de: a adição de um ou mais resíduos de ácido glutâmico no terminal C; a adição de

um ou mais resíduos de ácido aspártico no terminal C; e uma combinação das mesmas.

Mutação das Recombinases

[0041] Vários tipos de mutagênese são opcionalmente utilizados na presente divulgação, por exemplo, para modificar as recombinases para a produção de variantes, por exemplo, de acordo com modelos de recombinase e previsões de modelo ou utilizando abordagens de mutação aleatórias ou semialeatórias. Em geral, qualquer procedimento de mutagênese disponível pode ser utilizado para produzir mutantes de recombinase. Tais procedimentos de mutagênese incluem opcionalmente a seleção de ácidos nucleicos e polipeptídeos mutantes em relação a uma ou mais atividades de interesse (por exemplo, distribuição e/ou amplificação aprimorada sobre um suporte sólido). Os procedimentos que podem ser utilizados incluem, mas não são limitados a: mutagênese pontual direcionada ao sítio, mutagênese pontual aleatória, recombinação homóloga *in vitro* ou *in vivo* (embaralhamento de DNA e PCR de sobreposição combinatória), mutagênese utilizando moldes que contêm uracil, mutagênese direcionada ao oligonucleotídeo, mutagênese de DNA modificado com fosforotioato, mutagênese utilizando DNA em duplex com gaps, reparo de pareamentos errados pontuais, mutagênese utilizando cepas hospedeiras deficientes em relação ao reparo, seleção por restrição e purificação por restrição, mutagênese de deleção, mutagênese através da síntese do gene total, PCR degenerada, reparo de quebras de filamento duplo e muitos outros conhecidos pelos técnicos no assunto. A recombinase inicial para mutação pode ser qualquer uma das citadas aqui, incluindo mutantes de recombinases disponíveis tais como aquelas identificadas,

por exemplo, na US 2009/0029421, que é incorporada como referência em sua totalidade.

[0042] Opcionalmente, a mutagênese pode ser guiada pela informação conhecida de uma molécula de recombinase que ocorre naturalmente ou de uma recombinase alterada ou submetida à mutação conhecida (por exemplo, utilizando uma recombinase mutante existente que é citada nas referências anteriores), por exemplo, sequência, comparações de sequências, propriedades físicas, estrutura de cristal e/ou similar, como é discutido anteriormente. Entretanto, em outra classe de modalidades, a modificação pode ser essencialmente aleatória (por exemplo, como no embaralhamento de DNA clássico ou "familiar", ver, por exemplo, Crameri e outros (1998) "DNA shuffling of a family of genes from diverse species accelerates directed evolution" Nature 391 :288-291).

[0043] Informação adicional sobre os formatos de mutação é encontrada em: Sambrook e outros, Molecular Cloning—A Laboratory Manual (3ª Ed.), Vol. 1-3, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y., 2000 ("Sambrook"); Current Protocols in Molecular Biology, F. M. Ausubel et al., eds., Current Protocols, a joint venture between Greene Publishing Associates, Inc. e John Wiley & Sons, Inc., (suplementado até 2011) ("Ausubel") e PCR Protocols A Guide to Methods and Applications (Innis et al. eds) Academic Press Inc. San Diego, Calif. (1990) ("Innis"). As publicações e as referências a seguir citadas aqui fornecem detalhes adicionais sobre os formatos de mutação: Arnold, Protein engineering for unusual environments, Current Opinion in Biotechnology 4:450-455 (1993); Bass et al., Mutant Trp repressors with new DNA-

binding specificities, Science 242:240-245 (1988); Bordo e Argos (1991) Suggestions for “Safe” Residue Substitutions in Site-directed Mutagenesis 217:721-729; Botstein & Shortle, Strategies and applications of in vitro mutagenesis, Science 229: 1193-1201 (1985); Carter et al., Improved oligonucleotide site-directed mutagenesis using M13 vectors, Nucl. Acids Res. 13: 4431-4443 (1985); Carter, Site-directed mutagenesis, Biochem. J. 237: 1-7 (1986); Carter, Improved oligonucleotide-directed mutagenesis using M13 vectors, Methods in Enzymol. 154: 382-403 (1987); Dale et al., Oligonucleotide-directed random mutagenesis using the phosphorothioate method, Methods Mol. Biol. 57:369-374 (1996); Eghtedarzadeh & Henikoff, Use of oligonucleotides to generate large deletions, Nucl. Acids Res. 14: 5115 (1986); Fritz et al., Oligonucleotide-directed construction of mutations: a gapped duplex DNA procedure without enzymatic reactions in vitro, Nucl. Acids Res. 16: 6987-6999 (1988); Grundstrom et al., Oligonucleotide-directed mutagenesis by microscale “shot-gun” gene synthesis, Nucl. Acids Res. 13: 3305-3316 (1985); Hayes (2002) Combining Computational and Experimental Screening for rapid Optimization of Protein Properties PNAS 99(25) 15926-15931; Kunkel, The efficiency of oligonucleotide directed mutagenesis, em Nucleic Acids & Molecular Biology (Eckstein, F. e Lilley, D. M. J. eds., Springer Verlag, Berlim) (1987); Kunkel, Rapid and efficient site-specific mutagenesis without phenotypic selection, Proc. Natl. Acad. Sci. USA 82:488-492 (1985); Kunkel et al., Rapid and efficient site-specific mutagenesis without phenotypic selection, Methods in Enzymol. 154, 367-382 (1987); Kramer et al., The gapped duplex DNA approach to oligonucleotide-

directed mutation construction, Nucl. Acids Res. 12: 9441-9456 (1984); Kramer & Fritz Oligonucleotide-directed construction of mutations via gapped duplex DNA, Methods in Enzymol. 154:350-367 (1987); Kramer et al., Point Mismatch Repair, Cell 38:879-887 (1984); Kramer et al., Improved enzymatic in vitro reactions in the gapped duplex DNA approach to oligonucleotide-directed construction of mutations, Nucl. Acids Res. 16: 7207 (1988); Ling et al., Approaches to DNA mutagenesis: an overview, Anal Biochem. 254(2): 157-178 (1997); Lorimer e Pastan Nucleic Acids Res. 23, 3067-8 (1995); Mandecki, Oligonucleotide-directed double-strand break repair in plasmids of Escherichia coli: a method for site-specific mutagenesis, Proc. Natl. Acad. Sci. USA, 83:7177-7181(1986); Nakamaye & Eckstein, Inhibition of restriction endonuclease Nci I cleavage by phosphorothioate groups and its application to oligonucleotide-directed mutagenesis, Nucl. Acids Res. 14: 9679-9698 (1986); Nambiar et al., Total synthesis and cloning of a gene coding for the ribonuclease S protein, Science 223: 1299-1301(1984); Sakamar e Khorana, Total synthesis and expression of a gene for the a-subunit of bovine rod outer segment guanine nucleotide-binding protein (transducin), Nucl. Acids Res. 14: 6361-6372 (1988); Sayers et al., Y-T Exonucleases in phosphorothioate-based oligonucleotide-directed mutagenesis, Nucl. Acids Res. 16:791-802 (1988); Sayers et al., Strand specific cleavage of phosphorothioate-containing DNA by reaction with restriction endonucleases in the presence of ethidium bromide, (1988) Nucl. Acids Res. 16: 803-814; Sieber et al., Nature Biotechnology, 19:456-460 (2001); Smith, In vitro mutagenesis, Ann. Rev. Genet. 19:423-462 (1985);

Methods in Enzymol. 100: 468-500 (1983); Methods in Enzymol. 154: 329-350 (1987); Stemmer, Nature 370, 389-91(1994); Taylor et al., The use of phosphorothioate-modified DNA in restriction enzyme reactions to prepare nicked DNA, Nucl. Acids Res. 13: 8749-8764 (1985); Taylor et al., The rapid generation of oligonucleotide-directed mutations at high frequency using phosphorothioate-modified DNA, Nucl. Acids Res. 13: 8765-8787 (1985); Wells et al., Importance of hydrogen-bond formation in stabilizing the transition state of subtilisin, Phil. Trans. R. Soc. Lond. A 317: 415-423 (1986); Wells et al., Cassette mutagenesis: an efficient method for generation of multiple mutations at defined sites, Gene 34:315-323 (1985); Zoller & Smith, Oligonucleotide-directed mutagenesis utilizando M 13 - derived vectors: an efficient e general procedure for the production of point mutations in any DNA fragment, Nucleic Acids Res. 10:6487-6500 (1982); Zoller & Smith, Oligonucleotide-directed mutagenesis of DNA fragments cloned into M13 vectors, Methods in Enzymol. 100:468-500 (1983); Zoller & Smith, Oligonucleotide-directed mutagenesis: a simple method using two oligonucleotide primers and a single-stranded DNA template, Methods in Enzymol. 154:329-350 (1987); Clackson et al. (1991) “Making antibody fragments using phage display libraries” Nature 352:624-628; Gibbs et al. (2001) “Degenerate oligonucleotide gene shuffling (DOGS): a method for enhancing the frequency of recombination with family shuffling” Gene 271:13-20; e Hiraga e Arnold (2003) “General method for sequence-independent site-directed chimeragenesis: J. Mol. Biol. 330:287-296. Detalhes adicionais sobre muitos dos métodos anteriores podem ser

encontrados em Methods in Enzymology Volume 154, que também descreve controles úteis para a resolução de problemas com vários métodos de mutagênese.

Produção e Isolamento de Recombinase Recombinante

[0044] Geralmente, os ácidos nucleicos que codificam uma recombinase como aqui apresentada podem ser produzidos através de clonagem, recombinação, síntese *in vitro*, amplificação *in vitro* e/ou outros métodos disponíveis. Uma variedade de métodos recombinantes pode ser utilizada para expressar um vetor de expressão que codifica uma recombinase que é apresentada aqui. Os métodos para a produção de ácidos nucleicos recombinantes, para a expressão e o isolamento de produtos expressos são bem conhecidos e são descritos na arte. Um número de exemplos de mutações e combinações de mutações, bem como estratégias para o planejamento de mutações desejáveis, é descrito aqui.

[0045] Referências úteis adicionais para métodos de mutação, recombinação e manipulação de ácidos nucleicos *in vitro* (incluindo clonagem, expressão, PCR e similares) incluem Berger e Kimmel, Guide to Molecular Cloning Techniques, Methods in Enzymology volume 152 Academic Press, Inc., San Diego, Calif. (Berger); Kaufman et al. (2003) Handbook of Molecular and Cellular Methods in Biology and Medicine Segunda Edição Ceske (ed) CRC Press (Kaufman); e The Nucleic Acid Protocols Handbook Ralph Rapley (ed) (2000) Cold Spring Harbor, Humana Press Inc (Rapley); Chen et al. (ed) PCR Cloning Protocols, Segunda Edição (Methods in Molecular Biology, volume 192) Humana Press; e em Viljoen et al. (2005) Molecular Diagnostic PCR Handbook Springer, ISBN 1402034032.

[0046] Em adição, uma grande quantidade de kits está disponível comercialmente para a purificação de plasmídeos ou outros ácidos nucleicos relevantes das células, (ver, por exemplo, EasyPrep.TM., FlexiPrep.TM. ambos da Pharmacia Biotech; StrataClean.TM., da Stratagene; e, QIAprep.TM. da Qiagen). Qualquer ácido nucleico isolado e/ou purificado pode ser adicionalmente manipulado para produzir outros ácidos nucleicos, utilizado para transfectar células, incorporado em vetores relacionados para infectar organismos para expressão e/ou similar. Os vetores de clonagem típicos contêm terminadores da transcrição e da tradução, sequências que iniciam a transcrição e a tradução e promotores úteis para a regulação da expressão do ácido nucleico alvo particular. Os vetores opcionalmente compreendem cassetes de expressão genéricos que contêm pelo menos uma sequência terminadora independente, sequências que permitem a replicação do cassete em eucariotos ou procariotos ou ambos, (por exemplo, vetores *shuttle*) e marcadores de seleção tanto para sistemas procarióticos quanto eucarióticos. Os vetores são adequados para replicação e integração em procariotos, eucariotos ou ambos.

[0047] Outras referências úteis, por exemplo, para o isolamento e a cultura de células (por exemplo, para o isolamento subsequente de ácidos nucleicos) incluem Freshney (1994) Culture of Animal Cells, a Manual of Basic Technique, terceira edição, Wiley-Liss, New York e as referências citadas no mesmo; Payne et al. (1992) Plant Cell and Tissue Culture in Liquid Systems John Wiley & Sons, Inc. New York, N.Y.; Gamborg e Phillips (eds) (1995) Plant Cell, Tissue and Organ Culture; Fundamental Methods

Springer Lab Manual, Springer-Verlag (Berlin Heidelberg New York) e Atlas e Parks (eds) The Handbook of Microbiological Media (1993) CRC Press, Boca Raton, Fla.

[0048] Os ácidos nucleicos que codificam as recombinases recombinantes divulgadas aqui também são uma característica das modalidades apresentadas aqui. Um aminoácido particular pode ser codificado por vários códons e certos sistemas de tradução (por exemplo, células procarióticas ou eucarióticas) frequentemente exibem tendências de códons, por exemplo, organismos diferentes frequentemente preferem um dos vários códons sinônimos que codificam o mesmo aminoácido. Dessa maneira, os ácidos nucleicos apresentados aqui são opcionalmente “otimizados em relação ao códon”, significando que os ácidos nucleicos são sintetizados para incluírem os códons que são preferidos pelo sistema de tradução particular que será empregado para expressar a recombinase. Por exemplo, quando for desejável expressar a recombinase em uma célula de bactéria (ou até mesmo uma cepa de bactéria particular), o ácido nucleico pode ser sintetizado para incluir códons mais frequentemente encontrados no genoma de tal célula bacteriana, para a expressão eficiente da recombinase. Uma estratégia similar também pode ser empregada quando for desejável expressar a recombinase em uma célula eucariótica, por exemplo, o ácido nucleico pode incluir códons preferidos por tal célula eucariótica.

[0049] Uma variedade de métodos de isolamento e de detecção de proteína é conhecida na arte e pode ser utilizada para isolar as recombinases, por exemplo, de culturas de células recombinantes que expressam as recombinases recombinantes apresentadas aqui. Uma variedade

de métodos de isolamento e de detecção de proteína é bem conhecida na arte, incluindo, por exemplo, aqueles apresentados em R. Scopes, Protein Purification, Springer-Verlag, N.Y. (1982); Deutscher, Methods in Enzymology Vol. 182: Guide to Protein Purification, Academic Press, Inc. N.Y. (1990); Sandana (1997) Bioseparation of Proteins, Academic Press, Inc.; Bollag et al. (1996) Protein Methods, 2ª Edição Wiley-Liss, NY; Walker (1996) The Protein Protocols Handbook Humana Press, NJ, Harris e Angal (1990) Protein Purification Applications: A Practical Approach IRL Press at Oxford, Oxford, England; Harris e Angal Protein Purification Methods: A Practical Approach IRL Press at Oxford, Oxford, England; Scopes (1993) Protein Purification: Principles and Practice 3ª Edição Springer Verlag, NY; Janson e Ryden (1998) Protein Purification: Principles, High Resolution Methods and Applications, Segunda Edição Wiley-VCH, NY; e Walker (1998) Protein Protocols on CD-ROM Humana Press, NJ; e as referências citadas nos mesmos. Detalhes adicionais em relação aos métodos de purificação e de detecção de proteína podem ser encontrados em Satinder Ahuja ed., Handbook of Bioseparations, Academic Press (2000).

Métodos de uso

[0050] As recombinases alteradas apresentadas aqui podem ser utilizadas em um procedimento de amplificação mediado pela recombinase, tal como uma técnica de amplificação com recombinase polimerase (RPA). Sucintamente, a RPA pode ser iniciada através do contato de um ácido nucleico alvo com uma recombinase e um iniciador de ácido nucleico de filamento simples específico para a molécula de ácido nucleico alvo. O iniciador hibridizado

pode então ser estendido por uma polimerase, tal como uma polimerase capaz de deslocar o filamento na presença de dNTPs para gerar uma molécula de ácido nucleico alvo de filamento duplo e um filamento deslocado da molécula de ácido nucleico. Adicionalmente, a amplificação pode ocorrer através do direcionamento mediado pela recombinase de iniciadores para o filamento deslocado da molécula de ácido nucleico e da extensão do iniciador para gerar uma molécula de ácido nucleico de filamento duplo. O processo de RPA pode ser modulado através da combinação dos componentes descritos anteriormente com, por exemplo, fatores de carregamento da recombinase, polimerases de deslocamento de filamento específico e um sistema de regeneração de energia robusto. Os exemplos de procedimentos, sistemas e componentes de RPA que podem ser facilmente adaptados para uso com as proteínas UvsX recombinantes da presente divulgação são descritos, por exemplo, nas Pat. US Nos. 8,071,308; 7,399,590, 7,485,428, 7,270,981, 8,030,000, 7,666,598, 7,763,427, 8,017,399, 8,062,850 e 7,435,561, das quais cada uma é incorporada aqui como referência.

[0051] Em algumas modalidades, uma amplificação isotérmica pode ser realizada utilizando amplificação de exclusão cinética (KEA), também referida como amplificação de exclusão (ExAmp). Uma biblioteca de ácido nucleico da presente divulgação pode ser produzida utilizando um método que explora a exclusão cinética. A exclusão cinética pode ocorrer quando um processo é realizado a uma taxa suficientemente rápida para impedir eficientemente que outro evento ou processo ocorra. Considerar, por exemplo, a produção de um arranjo de ácido nucleico no qual os sítios do arranjo são distribuídos de forma aleatória com ácidos

nucleicos alvos provenientes de uma solução e cópias do ácido nucleico alvo são geradas em um processo de amplificação para preencher cada um dos sítios distribuídos até sua capacidade. De acordo com os métodos de exclusão cinética da presente divulgação, os processos de distribuição e amplificação podem ocorrer simultaneamente em condições em que a taxa de amplificação excede a taxa de distribuição. Dessa maneira, a taxa relativamente alta na qual as cópias são produzidas em um sítio que foi distribuído por um primeiro ácido nucleico alvo excluirá eficientemente um segundo ácido nucleico da distribuição do sítio para amplificação. Os métodos de amplificação com exclusão cinética podem ser realizados como é descrito em detalhes na divulgação da Publicação de Pedido de Patente U.S. No. 2013/0338042, que é incorporada aqui como referência em sua totalidade.

[0052] Em algumas modalidades, o ácido nucleico alvo que é amplificado é totalmente de filamento duplo. Em algumas modalidades, o ácido nucleico alvo que é amplificado compreende uma região de ácido nucleico de filamento duplo e compreende ainda uma região que possui ácido nucleico de filamento simples. Em certas modalidades, o ácido nucleico alvo compreende um ou mais adaptadores bifurcados com uma região de aproximadamente 5, 10, 15, 20, 25, 30, 35, 40 ou mais que aproximadamente 40 bases de sequência de filamento simples em cada extremidade dos fragmentos da biblioteca. O planejamento e o uso de adaptadores bifurcados são descritos em maiores detalhes nas divulgações das Pat. U.S. Nos. 7,742,463 e 8,563,748, das quais cada uma é incorporada aqui como referência em sua totalidade.

[0053] A exclusão cinética pode explorar uma taxa relativamente lenta para a produção de uma primeira cópia de um ácido nucleico alvo vs. um taxa relativamente alta para a produção de cópias subsequentes do ácido nucleico alvo ou da primeira cópia. No exemplo do parágrafo anterior, a exclusão cinética ocorre devido à taxa relativamente lenta de distribuição do ácido nucleico alvo (por exemplo, difusão ou transporte relativamente lento) vs. a taxa relativamente alta na qual a amplificação ocorre para preencher o sítio com cópias do ácido nucleico distribuído. Em outros exemplos de modalidade, a exclusão cinética pode ocorrer devido a um retardo na formação de uma primeira cópia de um ácido nucleico alvo que foi distribuído em um sítio (por exemplo, ativação retardada ou lenta) vs. a taxa relativamente alta na qual as cópias subsequentes são produzidas para preencher o sítio. Neste exemplo, um sítio individual pode ter sido distribuído com vários ácidos nucleicos alvos diferentes (por exemplo, vários ácidos nucleicos alvos podem estar presentes em cada sítio antes da amplificação). Entretanto, a formação da primeira cópia para qualquer ácido nucleico alvo fornecido pode ser ativada aleatoriamente de forma que a taxa média de formação da primeira cópia seja relativamente lenta comparada com a taxa na qual as cópias subsequentes são geradas. Neste caso, embora um sítio individual possa ter sido distribuído com vários ácidos nucleicos alvos diferentes, a exclusão cinética permitirá que apenas um destes ácidos nucleicos alvos seja amplificado. Mais especificamente, uma vez que um primeiro ácido nucleico alvo foi ativado para amplificação, o sítio irá rapidamente preencher até a capacidade com suas cópias, prevenindo

assim que cópias de um segundo ácido nucleico alvo sejam produzidas no sítio.

[0054] Um reagente de amplificação pode incluir componentes adicionais que facilitam a formação de amplicon e em alguns casos aumentam a taxa de formação de amplicon. Uma recombinase, tal como, por exemplo, UvsX, pode facilitar a formação de amplicon permitindo invasão/extensão repetida. Mais especificamente, uma recombinase pode facilitar a invasão de um ácido nucleico alvo pela polimerase e a extensão de um iniciador pela polimerase utilizando o ácido nucleico alvo como um molde para a formação de amplicon. Este processo pode ser repetido na forma de uma reação em cadeia na qual os amplicons produzidos partindo de cada rodada de invasão/extensão servem como moldes em uma rodada subsequente. O processo pode ocorrer mais rapidamente que a PCR padronizada uma vez que um ciclo de desnaturação (por exemplo, através de aquecimento ou desnaturação química) não é necessário. Dessa maneira, a amplificação facilitada pela recombinase pode ser realizada de forma isotérmica. É geralmente desejável incluir ATP ou outros nucleotídeos (ou em alguns casos análogos não hidrolisáveis dos mesmos) em um reagente de amplificação facilitada pela recombinase para facilitar a amplificação. Uma mistura de recombinase e proteína de ligação de filamento simples (SSB) é particularmente útil uma vez que a SSB pode facilitar adicionalmente a amplificação. Os exemplos de formulações para amplificação facilitada pela recombinase incluem aquelas vendidas comercialmente como TwistAmp kits pela TwistDx (Cambridge, UK). Os componentes úteis de reagente de amplificação facilitada pela recombinase e condições de

reação são apresentados na US 5,223,414 e na US 7,399,590, das quais cada uma é incorporada aqui como referência.

Comparação, Identidade e Homologia de Sequências

[0055] Os termos “idênticas” ou “porcentagem de identidade”, no contexto de duas ou mais sequências de ácidos nucleicos ou de polipeptídeos, referem-se a duas ou mais sequências ou subsequências que são as mesmas ou que possuem uma porcentagem especificada de resíduos de aminoácidos ou nucleotídeos que são os mesmos, quando comparadas e alinhadas para correspondência máxima, que é medida utilizando um dos algoritmos de comparação de sequências descritos a seguir (ou outros algoritmos disponíveis para técnicos no assunto) ou por inspeção visual.

[0056] A expressão “substancialmente idênticos”, no contexto de dois ácidos nucleicos ou polipeptídeos (por exemplo, DNAs que codificam uma recombinase ou a sequência de aminoácidos de uma recombinase) refere-se a duas ou mais sequências ou subsequências que possuem pelo menos aproximadamente 60%, aproximadamente 80%, aproximadamente 90-95%, aproximadamente 98%, aproximadamente 99% ou mais de identidade de nucleotídeos ou de resíduos de aminoácidos, quando comparadas e alinhadas para correspondência máxima, que é medida utilizando um algoritmo de comparação de sequências ou por inspeção visual. Tais sequências “substancialmente idênticas” são tipicamente consideradas como sendo “homólogas”, sem referência à descendência real. Preferencialmente, a “identidade substancial” existe ao longo de uma região das sequências que tem pelo menos aproximadamente 50 resíduos de comprimento, mais preferencialmente ao longo de uma região de pelo menos

aproximadamente 100 resíduos e mais preferencialmente, as sequências são substancialmente idênticas ao longo de pelo menos aproximadamente 150 resíduos ou ao longo do comprimento completo das duas sequências que serão comparadas.

[0057] Proteínas e/ou sequências de proteínas são "homólogas" quando são derivadas, naturalmente ou artificialmente, de uma proteína ou uma sequência de proteína de um ancestral comum. Similarmente, os ácidos nucleicos e/ou as sequências de ácidos nucleicos são homólogas quando são derivadas, naturalmente ou artificialmente, de um ácido nucleico ou uma sequência de ácido nucleico de um ancestral comum. A homologia é geralmente inferida partindo da similaridade de sequência entre dois ou mais ácidos nucleicos ou proteínas (ou sequências dos mesmos). A porcentagem precisa de similaridade entre sequências que é útil no estabelecimento de homologia varia com o ácido nucleico e a proteína em questão, mas tão pouco quanto 25% de similaridade de sequência ao longo de 50, 100, 150 ou mais resíduos é rotineiramente utilizada para estabelecer homologia. Níveis mais altos de similaridade de sequência, por exemplo, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95% ou 99% ou mais, também podem ser utilizados para estabelecer homologia. Os métodos para a determinação das porcentagens de similaridade de sequência (por exemplo, BLASTP e BLASTN utilizando parâmetros pré-estabelecidos) são descritos aqui e estão geralmente disponíveis.

[0058] Para a comparação de sequências e para a determinação de homologia, tipicamente uma sequência age como uma sequência de referência à qual as sequências de

teste são comparadas. Quando se utiliza um algoritmo de comparação de sequências, as sequências de teste e de referência são inseridas em um computador, as coordenadas das subsequências são designadas, se necessário, e os parâmetros do programa do algoritmo de sequência são designados. O algoritmo de comparação de sequências então calcula a porcentagem de identidade de sequência para a(s) sequência(s) de teste em relação à sequência de referência, com base nos parâmetros do programa designados.

[0059] O alinhamento ótimo de sequências para comparação pode ser realizado, por exemplo, através do algoritmo de homologia local de Smith & Waterman, Adv. Appl. Math. 2:482 (1981), através do algoritmo de alinhamento de homologia de Needleman & Wunsch, J. Mol. Biol. 48:443 (1970), através do método de busca por similaridade de Pearson & Lipman, Proc. Nat'l. Acad. Sci. USA 85:2444 (1988), através de implementações computadorizadas destes algoritmos (GAP, BESTFIT, FASTA e TFASTA no Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr., Madison, Wis.) ou através de inspeção visual (ver, de forma geral, Current Protocols in Molecular Biology, Ausubel e outros, eds., Current Protocols, um empreendimento conjunto entre Greene Publishing Associates, Inc. e John Wiley & Sons, Inc., suplementado até 2004).

[0060] Um exemplo de um algoritmo que é adequado para a determinação da porcentagem de identidade de sequência e da similaridade de sequência é o algoritmo BLAST, que é descrito em Altschul et al., J. Mol. Biol. 215:403-410 (1990). O software para a realização das análises de BLAST está disponível publicamente através do

National Center for Biotechnology Information. Este algoritmo envolve uma primeira identificação de pares de sequências de escore alto (HSPs) através da identificação de palavras pequenas de comprimento W na sequência de busca, que se combina ou satisfaz parte do escore de limite de valor positivo T quando alinhada com uma palavra do mesmo comprimento em uma sequência da base de dados. T é referido como o limite de escore de palavra vizinha (Altschul et al., supra). Estes acertos de palavra vizinha iniciais agem como “sementes” para iniciar as pesquisas para encontrar HSPs mais longos que contêm os mesmos. Os acertos de palavras são então estendidos e ambas as direções ao longo de cada sequência o máximo possível à medida que o escore de alinhamento cumulativo pode ser aumentado. Os escores cumulativos são calculados utilizando, para sequências de nucleotídeos, os parâmetros M (escore de recompensa para um par de resíduos que combinam; sempre >0) e N (escore de penalidade para resíduos com pareamentos errados; sempre <0). Para as sequências de aminoácidos, uma matriz de escore é utilizada para calcular o escore cumulativo. A extensão dos acertos de palavras em cada direção é interrompida quando: o escore de alinhamento cumulativo falha pela quantidade X de seu valor máximo atingido; o escore cumulativo vai até zero ou inferior, devido ao acúmulo de um ou mais alinhamentos de resíduos de escore negativo; ou o final de qualquer uma das sequências é atingido. Os parâmetros do algoritmo BLAST W, T e X determinam a sensibilidade e a velocidade do alinhamento. O programa BLASTN (para sequências de nucleotídeos) utiliza como parâmetros pré-estabelecidos um comprimento de palavra (W) de 11, uma expectativa (E) de

10, um limite de 100, M=5, N=-4 e uma comparação de ambos os filamentos. Para as sequências de aminoácidos, o programa BLASTP utiliza como parâmetros pré-estabelecidos um comprimento de palavra (W) de 3, uma expectativa (E) de 10 e a matriz de escore BLOSUM62 (ver Henikoff & Henikoff (1989) Proc. Natl. Acad. Sci. USA 89: 10915).

[0061] Em adição ao cálculo da porcentagem de identidade de sequência, o algoritmo BLAST também realiza uma análise estatística da similaridade entre duas sequências (ver, por exemplo, Karlin & Altschul, Proc. Nafl. Acad. Sci. USA 90:5873-5787 (1993)). Uma medida da similaridade fornecida pelo algoritmo BLAST é a probabilidade da menor soma (P(N)), que fornece uma indicação da probabilidade pela qual uma combinação entre duas sequências de nucleotídeos ou aminoácidos poderia ocorrer ao acaso. Por exemplo, um ácido nucleico é considerado similar a uma sequência de referência se a probabilidade da menor soma em uma comparação do ácido nucleico de teste com o ácido nucleico de referência for menor que aproximadamente 0,1, mais preferencialmente menor que aproximadamente 0,01 e mais preferencialmente menor que aproximadamente 0,001.

Ácidos nucleicos que codificam recombinases alteradas

[0062] São adicionalmente apresentadas aqui as moléculas de ácido nucleico que codificam as enzimas recombinases alteradas apresentadas aqui. Para qualquer recombinase alterada fornecida que é uma submetida à mutação de uma recombinase para a qual a sequência de aminoácidos e preferencialmente também a sequência de nucleotídeos do tipo selvagem que codifica a recombinase são conhecidas, é possível obter uma sequência de

nucleotídeos que codifica o mutante de acordo com os princípios básicos da biologia molecular. Por exemplo, dado que a sequência de nucleotídeos do tipo selvagem que codifica a UvsX recombinase de RB49 é conhecida, é possível deduzir uma sequência de nucleotídeos que codifica qualquer versão mutante fornecida de UvsX de RB49 que possui uma ou mais substituições de aminoácidos utilizando o código genético padronizado. Similarmente, as sequências de nucleotídeos podem ser facilmente derivadas para versões mutantes de outras recombinases tais como, por exemplo, de T4, T6, Rb69, Aeh1, KVP40, fago 133 de *Acinetobacter*, fago de *Aeromonas* 65, cianofago P-SSM2, cianofago PSSM4, cianofago S-PM2, Rb32, fago nt-1 de *Vibrio*, Rbl6, Rb43, T2, Rbl4, fago 25 de *Aeromonas*, phi-1, Fago 31, fago 44RR2.8t, fago Rb3 e fago LZ2, etc. Moléculas de ácido nucleico que possuem a sequência de nucleotídeos requerida podem então ser construídas utilizando técnicas de biologia molecular padronizadas conhecidas na arte.

[0063] De acordo com as modalidades apresentadas aqui, um ácido nucleico definido inclui não somente o ácido nucleico idêntico, mas também quaisquer variações de bases menores incluindo, em particular, substituições em casos que resultam em um códon sinônimo (um códon diferente que especifica o mesmo resíduo de aminoácido) devido ao código degenerado nas substituições conservativas de aminoácidos. O termo “sequência de ácido nucleico” inclui também a sequência complementar a qualquer sequência de filamento simples fornecidas as variações de bases relativas.

[0064] As moléculas de ácido nucleico descritas aqui também podem ser, vantajosamente, incluídas em um vetor de expressão adequado para expressar as proteínas

recombinases codificadas pelas mesmas em um hospedeiro adequado. A incorporação do DNA clonado em um vetor de expressão adequado para transformação subsequente da dita célula e a seleção subsequente das células transformadas é bem conhecida pelos técnicos no assunto como é fornecido em Sambrook et al. (1989), Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, que é incorporado como referência em sua totalidade.

[0065] Tal vetor de expressão inclui um vetor que possui um ácido nucleico de acordo com as modalidades apresentadas aqui ligado de forma operacional às sequências reguladoras, tais como regiões promotoras, que são capazes de realizar a expressão dos ditos fragmentos de DNA. O termo “ligado de forma operacional” se refere a uma justaposição na qual os componentes descritos estão em uma relação que permite que estes funcionem de sua maneira pretendida. Tais vetores podem ser transformados em uma célula hospedeira adequada para possibilitar a expressão de uma proteína de acordo com as modalidades apresentadas aqui.

[0066] A molécula de ácido nucleico pode codificar uma proteína madura ou uma proteína que possui uma pró-sequência, incluindo aquela que codifica uma sequência líder sobre a pré-proteína que é então clivada pela célula hospedeira para formar uma proteína madura. Os vetores podem ser, por exemplo, vetores plasmideais, virais ou de fagos fornecidos com uma origem de replicação e opcionalmente um promotor para a expressão do dito nucleotídeo e opcionalmente um regulador do promotor. Os vetores podem conter um ou mais marcadores selecionáveis,

tal como, por exemplo, um gene de resistência a antibiótico.

[0067] Os elementos reguladores necessários para a expressão incluem sequências promotoras que se ligam à RNA polimerase e que levam a um nível apropriado de sequências de início da transcrição e também de início da tradução para ligação de ribossomos. Por exemplo, um vetor de expressão bacteriano pode incluir um promotor tal como o promotor lac e para o início da tradução a sequência Shine-Dalgarno e o códon de início AUG. Similarmente, um vetor de expressão eucariótico pode incluir um promotor heterólogo ou homólogo para RNA polimerase II, um sinal de poliadenilação a jusante, o códon de início AUG e um códon de término para liberação do ribossomo. Tais vetores podem ser obtidos comercialmente ou ser montados partindo das sequências descritas pelos métodos bem conhecidos na arte.

[0068] A transcrição do DNA que codifica a recombinase por eucariotos superiores pode ser otimizada através da inclusão de uma sequência intensificadora no vetor. Os intensificadores são elementos que atuam em cis do DNA que agem sobre um promotor para aumentar o nível de transcrição. Os vetores também incluirão geralmente origens de replicação em adição aos marcadores selecionáveis.

EXEMPLO 1

[0069] Este exemplo fornece métodos de distribuição de uma biblioteca sem a necessidade de PCR sobre uma superfície de célula de fluxo padronizada para amplificação aprimorada de clusters. Em uma modalidade, o método da invenção utiliza uma formulação de distribuição que inclui uma UvsX que compreende as mutações apresentadas anteriormente aqui, por exemplo, uma UvsX de RB49 mutante

que compreende Pro256Lys (que é apresentada aqui como a SEQ ID NO: 2, referida aqui como “P256K de RB49”). Foi descoberto de forma surpreendente que a amplificação mediada por recombinase utilizando a mesma aprimora substancialmente a distribuição de bibliotecas de PCR com regiões adaptadoras de filamento simples sobre uma superfície de célula de fluxo padronizada. Em outra modalidade, o método da invenção utiliza uma formulação de distribuição que inclui uma concentração relativamente alta de DNA polimerase (por exemplo, eBsu polimerase) em combinação com a P256K recombinase de RB49.

[0070] Para avaliar a eficiência da formulação de P256K de RB49na distribuição de uma biblioteca sem a necessidade de PCR sobre uma superfície de célula de fluxo padronizada, uma biblioteca sem a necessidade de PCR foi gerada utilizando um kit de preparação de amostra sem PCR TruSeq® DNA (Illumina, Inc.). As bibliotecas sem a necessidade de PCR geradas utilizando o kit de preparação de biblioteca TruSeq® possuem adaptadores bifurcados com uma região de aproximadamente 40 bases de sequência de filamento simples em cada extremidade dos fragmentos da biblioteca.

[0071] A Figura 3A mostra uma fotografia 100 de uma imagem de cluster de uma biblioteca sem a necessidade de PCR distribuída sobre uma célula de fluxo padronizada utilizando uma formulação padronizada que compreende UvsX recombinase de T4 (que é apresentada aqui como a SEQ ID NO: 8, referida aqui como “UvsX de T4”). A Figura 3B mostra uma fotografia 150 de uma imagem de cluster de uma biblioteca sem a necessidade de PCR distribuída sobre uma célula de fluxo padronizada utilizando uma formulação líquida que

inclui P256K recombinase de RB49. Neste exemplo, a biblioteca foi misturada com a formulação de UvsX de T4 ou a formulação de P256K de RB49 até a concentração final de 100 pM, descarregada sobre uma célula de fluxo e incubada em um cBot a 38 °C. Após um período de incubação de 1 hora, a temperatura foi reduzida para 20 °C e a célula de fluxo foi lavada com tampão de lavagem HT2 (Illumina). Os clusters foram corados com uma diluição de 1:5.000 de SYBR® Green (Life Technologies) em 0,1 M de Tris/0,1 M de ascorbato de sódio e foram obtidas imagens em um microscópio de fluorescência. Em referência à Figura 3A, a densidade de cluster gerada pela distribuição de uma biblioteca sem a necessidade de PCR sobre uma célula de fluxo padronizada com uma formulação padronizada (por exemplo, UvsX de T4) é relativamente esparsa. Em referência à Figura 3B, a densidade de clusters gerada pela distribuição de uma biblioteca sem a necessidade de PCR sobre uma célula de fluxo padronizada utilizando a formulação que inclui a P256K recombinase de RB49 é substancialmente aumentada.

[0072] A Figura 4A mostra uma fotografia 200 de uma imagem de cluster de uma biblioteca sem a necessidade de PCR de filamento simples (ssDNA) distribuída sobre uma célula de fluxo padronizada utilizando uma formulação padronizada de UvsX de T4. A Figura 4B mostra uma fotografia 250 de uma imagem de cluster de uma biblioteca sem a necessidade de PCR de filamento simples distribuída sobre uma célula de fluxo padronizada utilizando uma formulação líquida que inclui a P256K recombinase de RB49. Neste exemplo, uma biblioteca sem a necessidade de PCR de filamento duplo foi desnaturada utilizando NaOH e

subsequentemente distribuída sobre a célula de fluxo padronizada a uma concentração de 50 pM. Em referência à Figura 4A, a densidade de clusters gerada pela distribuição uma biblioteca de DNA sem PCR de ssDNA sobre uma célula de fluxo padronizada com uma formulação padronizada (por exemplo, UvsX de T4) é relativamente esparsa. Em referência à Figura 4B, a densidade de clusters gerada pela distribuição de uma biblioteca sem a necessidade de PCR de ssDNA sobre uma célula de fluxo padronizada utilizando uma formulação que inclui a P256K recombinase de RB49 é substancialmente aumentada.

EXEMPLO 2

Amplificação aprimorada utilizando mutantes P256K de RB49

[0073] Este exemplo descreve uma comparação do desempenho da amplificação entre as recombinases com e sem a mutação P256K descrita aqui. Para as finalidades deste exemplo, UvsX de RB49 de “controle” (apresentada na SEQ ID NO: 1) compreende ainda uma mutação H63S. Um mutante P256K é gerado através da mutação adicional do controle para carregar um resíduo de Lys na posição 256, como é apresentado aqui pela SEQ ID NO: 2).

[0074] A obtenção de clusters de uma biblioteca sem a necessidade de PCR em uma célula de fluxo padronizada é realizada em um cBot como é descrito anteriormente no Exemplo 1, utilizando controle ou mutante P256K. O sequenciamento é então realizado em um equipamento HiSeq (Illumina, Inc.) e os resultados do sequenciamento são analisados para determinar a resgatabilidade de uma variedade de regiões que são tipicamente insuficientemente representadas nos dados de sequenciamento anteriores.

[0075] Resgatabilidade é uma medida da fração de sítios nos quais um polimorfismo de nucleotídeo único (SNP) é recuperado corretamente. De forma ideal, este valor é 1 (para 100%) significando que em 100% dos sítios dentro de um tipo de região particular (isto é, alto conteúdo de GC etc) os SNPs são recuperados corretamente. A cobertura é uma medida da fração dos sítios que possuem uma cobertura > n, em que n é tipicamente 30x (isto é, a cobertura padronizada para um genoma humano). Os promotores de fosmídeo são um conjunto de 100 promotores gênicos que foram identificados como insuficientemente representados nos dados de sequenciamento anteriores. Os promotores foram clonados em vetores fosmídeos. Uma região de alto conteúdo de GC pode ser definida como uma região com pelo menos 100 pb em que o conteúdo de GC é igual ou superior a 75% (N50 (G + C $\geq$ 0,75) 100 N50). Uma região com conteúdo de GC enorme pode ser definida como uma região com pelo menos 100 pb em que o conteúdo de GC é igual ou superior a 85% (N50 (G + C $\geq$ 0,85) 100 N50). Uma região com conteúdo de GC baixo pode ser definida como uma região com pelo menos 100 pb em que o conteúdo de GC é igual ou menor 40% (N50 (G + C $\geq$ 0,40) 100 N50). Uma região com conteúdo de AT alto pode ser definida como uma região com pelo menos 100 pb em que o conteúdo de AT é igual ou superior a 75% (N50 (A + T $\geq$ 0,75) 100 N50), com taxa de amostragem reduzida para ~ 50k regiões. Uma região com conteúdo de AT enorme pode ser definida como uma região com pelo menos 100 pb em que o conteúdo de AT é igual ou superior a 85% (N50 (A + T $\geq$ 0,85) 100 N50), com taxa de amostragem reduzida para ~ 50K regiões. Uma região de repetição do dinucleotídeo AT pode

ser definida como uma região que inclui extensões longas de repetições ATAT.

[0076] Uma comparação de dados de resgatabilidade para controle vs. mutantes de P256K demonstra que o mutante P256K exibe melhorias inesperadas e significativas na resgatabilidade de uma ou mais das regiões promotoras do fosmídio, Conteúdo de GC Alto, Conteúdo de GC Enorme, Conteúdo de GC Baixo, Conteúdo de AT Alto, Conteúdo de AT Enorme e regiões de repetição do dinucleotídeo AT, comparado com a do controle.

EXEMPLO 3

Amplificação aprimorada utilizando mutantes que possuem mutações homóloga a P256

[0077] A comparação de desempenho descrita anteriormente no Exemplo 2 é repetida para outras recombinases. Neste exemplo, recombinases de “controle” são geradas através da modificação da recombinase do tipo selvagem para compreender uma mutação homóloga a H63S em RB49, como é apresentado na coluna de “controle” na tabela a seguir. Os mutantes de “homólogos de P256K” são gerados através da modificação adicional dos controles para carregarem uma mutação homóloga a P256K em RB49 como é apresentado na coluna de “homólogo de P256K” na tabela a seguir.

[0078] Por exemplo, para UvsX de T6, o controle é gerado através da modificação da UvsX de T6 do tipo selvagem (SEQ ID NO: 9) para carregar uma mutação H66S. O homólogo de P256K é adicionalmente modificado para carregar ambas as mutações H66S e F259K.

Estrutura do Tipo Selvagem	**SEQ ID NO: da Estrutura do**	**Controle**	**Homólogo de P256K**

	Tipo Selvagem		
T4	8	64S	64S F257
T6	9	H66S	H66S F259K
Fago 133 de *Acinetobacter*	10	H64S	H64S P257K
Rb69	11	H64S	H64S P258K
Aeh1	12	H76S	H76S P269K
Fago 65 de *Aeromonas*	13	H73S	H73S D266K
KVP40	14	H64S	H64S P267K
Rb43	15	H66S	H66S P259K
P-SSM2 de cianofago	16	T62S	T62S Q261K
PSSM4 de cianofago	17	T65S	T65S E264K
S-PM2 de cianofago	18	T65S	T65S E264K
Rb32	19	H66S	H66S F259K
Fago nt-1 de *Vibrio*	20	H64S	H64S P267K
Rb16	21	H66S	H66S P259K

[0079] A obtenção de clusters de uma biblioteca sem a necessidade de PCR em uma célula de fluxo padronizada é realizada em um cBot como é descrito anteriormente no Exemplo 1, utilizando controle ou mutante de P256K. O sequenciamento é então realizado em um equipamento HiSeq (Illumina, Inc.) e os resultados do sequenciamento são analisados como descrito anteriormente no Exemplo 2 para determinar a resgatabilidade de uma variedade de regiões que são tipicamente insuficientemente representadas nos dados de sequenciamento anteriores.

[0080] Uma comparação dos dados de resgatabilidade para controle vs. mutantes homólogos de P256K demonstra que os mutantes homólogos de P256K exibem aprimoramentos inesperados e significativos na resgatabilidade de uma ou mais das regiões promotoras do fosmídeo, Conteúdo de GC Alto, Conteúdo de GC Enorme, Conteúdo de GC Baixo, Conteúdo de AT Alto, Conteúdo de AT Enorme e regiões de repetição do dinucleotídeo AT, comparado com a do controle.

[0081] Ao longo de todo este pedido de patente várias publicações, patentes e/ou pedidos de patentes foram referidos. A divulgação destas publicações em suas totalidades é incorporada aqui como referência neste pedido de patente.

[0082] É pretendido que o termo compreendendo seja ilimitado, incluindo não somente os elementos citados, mas abrangendo ainda quaisquer elementos adicionais.

Foi descrito um número de modalidades. Todavia, será entendido que várias modificações podem ser feitas. Consequentemente, outras modalidades estão dentro do âmbito das reivindicações a seguir.

REIVINDICAÇÕES

1. UvsX recombinante **caracterizado pelo** fato de que compreende uma sequência de aminoácidos como estabelecida pela a SEQ ID NO: 1, em que compreende ainda uma mutação de substituição de aminoácido para lisina na posição 256 e para serina na posiçãoo 63;

em que o referido polipeptídeo UvsX recombinante possu atividade recombinase aumentada em comparação com ambos,:

I - um polipeptídeo recombinase UvsX de T4 possuindo a SEQ ID NO: 8; ou

II - um polipeptídeo recombinase UvsX de RB49 possuindo toda a sequência de aminoácidos da SEQ ID NO: 1 exceto para uma substituição His para Ser na posição de aminoácido 63 da SEQ ID NO: 1.

2. UvsX recombinante, de acordo com a reivindicação 1, **caracterizado pelo** fato de que compreende uma mutação selecionada dentre: a adição de um ou mais resíduos de ácido glutâmico no terminal C; a adição de um ou mais resíduos de ácido aspártico no terminal C; ou uma combinação destes.

3. Processo de amplificação da polimerase recombinase de amplificar uma molécula de ácido nucleico alvo **caracterizado pelo** fato de que compreende as etapas de:

(a) colocar a UvsX recombinante conforme definida por qualquer uma das reivindicações 1 a 2 em contato com um primeiro e segundo iniciador de ácido nucleico para formar um primeiro e segundo iniciador de nucleoproteína, em que o referido iniciador de ácido nucleico compreende uma região de cadeia simples na sua extremidade 3';

(b) colocar o primeiro e o segundo iniciador nucleoprotéico em contanto com a referida molécula de ácido nucleico alvo, formando, assim, uma primeira estrutura de cadeia dupla em uma primeira porção da referida primeira cadeia e formando uma segunda estrutura de cadeia dupla em uma segunda porção da referida segunda cadeia das extremidades 3' do referido primeiro iniciador de ácido nucleico e o referido segundo iniciador de ácido nucleico estão orientados um para o outro na mesma molécula de ácido nucleico molde de cadeia dupla;

(c) extender a extremidade 3' do referido primeiro e segundo iniciador de ácido nucleico com uma ou mais polimerases e dNTPs para gerar um primeiro e segundo ácido nucleico de cadeia dupla e uma primeira e segunda cadeias deslocadas de ácido nucleico; e

(d) continuar a reação através da repetição de (b) e (c) até ser atingido um grau desejado de amplificação.

4. Processo, de acordo com a reivindicação 3, **caracterizado pelo** fato de que a referida molécula de ácido nucleico alvo compreende ácido nucleico de cadeia dupla e/ou uma região adaptadora de cadeia simples.

5. Processo, de acordo com a reivindicação 3, **caracterizado pelo** fato de que o referido processo é realizado na presença de uma proteína de carga de recombinase, opcionalmente, em que a proteína de carga da recombinase é selecionada do grupo que consiste em T4 UvsY, *E. coli* recO, *E. coli* recR ou uma combinação destes.

6. Processo, de acordo com a reivindicação 3, **caracterizado pelo** fato de que o referido processo é

realizado na presença de um agente estabilizador de cadeia simples selecionado do grupo que consiste em gp32, proteína SSB de *E. coli*, Proteína gp32 de T4.

7. Processo, de acordo com a reivindicação 3, **caracterizado pelo** fato de que o referido processo é realizado na presença de um agente de aglomeração selccionado do grupo compreendendo polietilenoglicol, óxido de polietileno, poliestireno, Ficoll, dextrano, PVP e albumina, de tal modo que o agente de aglomeração estimule a amplificação.

8. Processo, de acordo com a reivindicação 3, **caracterizado pelo** fato do referido processo ser realizado em um conjunto de locais de amplificação, opcionalmente em que cada local de amplificação compreende uma pluralidade de iniciadores de amplificação para amplificação do ácido nucleico alvo.

9. Processo, de acordo com a reivindicação 3, **caracterizado pelo** fato do conjunto de locais de amplificação compreendeR um conjunto de características em uma superfície, opcionalmente, em que as características são não-contínuas e são separadas por regiões intersticiais da superfície que não possuem os iniciadores de amplificação.

10. Processo, de acordo com a reivindicação 3, **caracterizado pelo** fato de que o processo ocorre isotermicamente.

11. Kit para realizar uma reação de polimerase recombinase **caracterizado pelo** fato de que compreende: um UvsX recombinante conforme definida por qualquer uma das reivindicações 1 a 2 e um ou mais dos seguintes:

uma proteína de ligação ao DNA de cadeia simples;

uma DNA polimerase;

dNTPs ou uma mistura de dNTPs e ddNTPs;

um agente de aglomeração;

um tampão;

um agente redutor;

análogo de ATP ou ATP;

uma proteína de carga de recombinase;

um primeiro iniciador; e

opcionalmente um segundo iniciador em adição ao primeiro iniciador.

ALINHAMENTO DE RELATIVOS DA UvsX

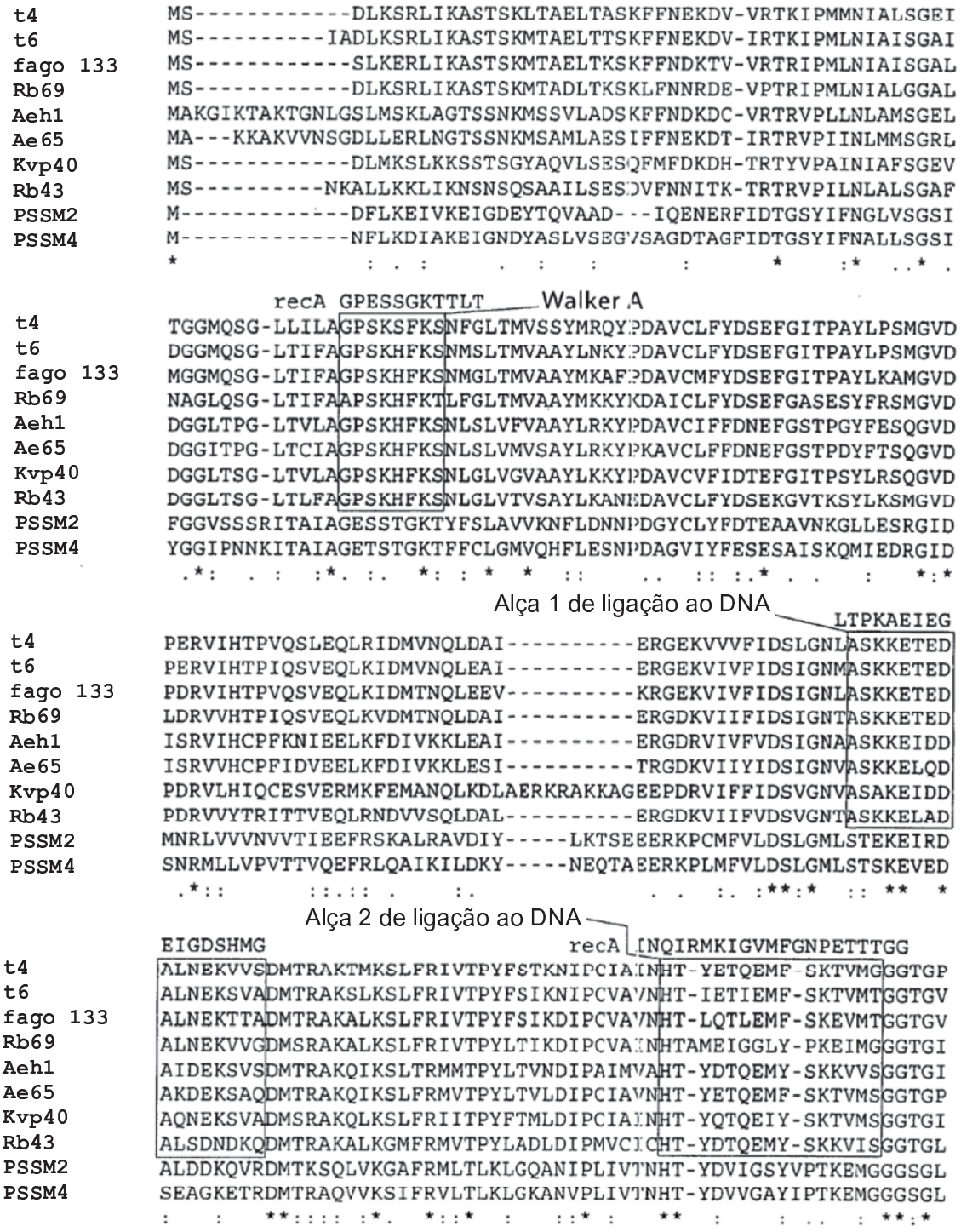


FIGURA 1A

ALINHAMENTO DE RELATIVOS DA UvsX

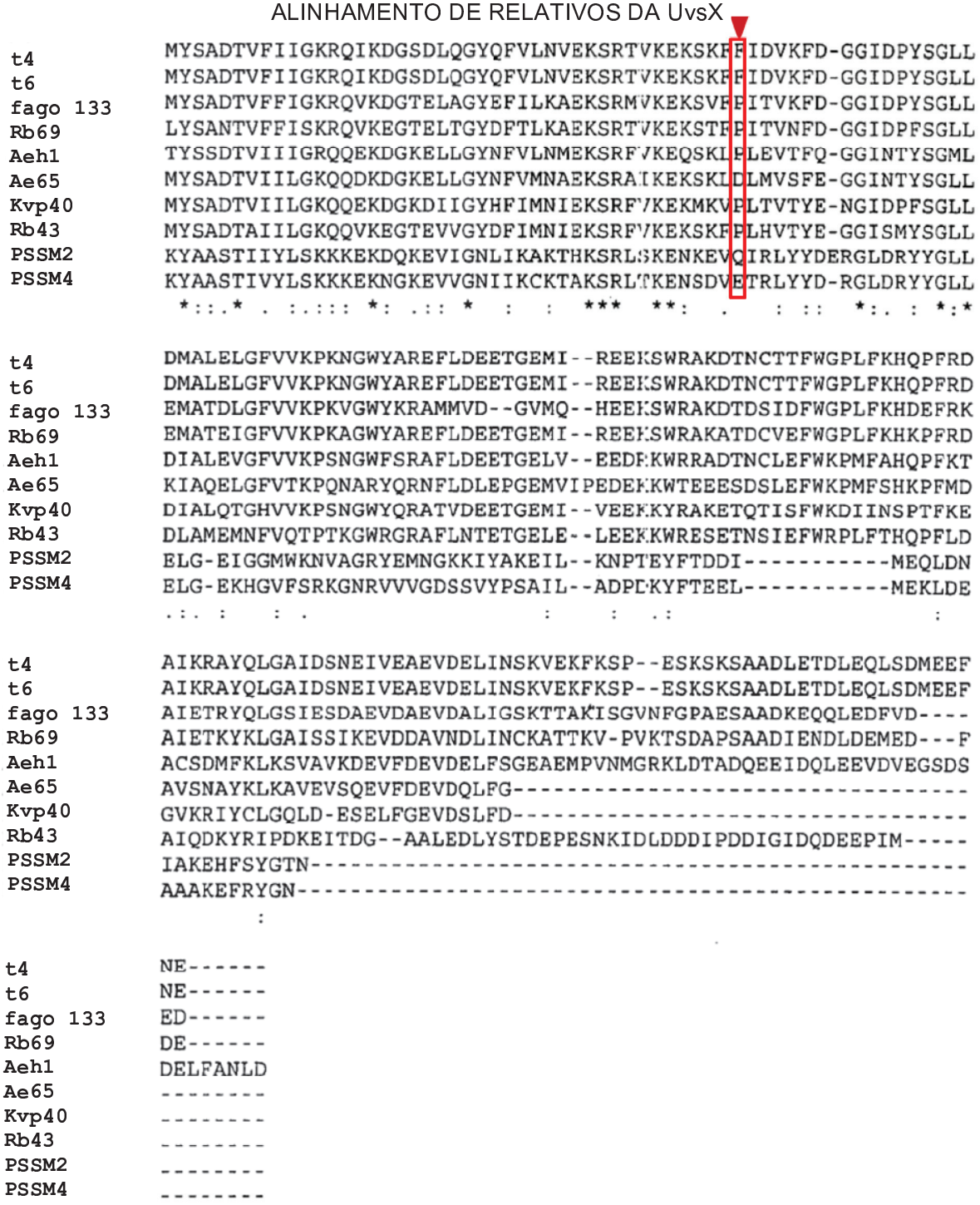


FIGURA 1B

```
RB49    MS-VLEKLKKNSTLKTTAVLSKSSFFNEKTNTRTKIPMLNIAFSGDLKKGFQSGLIFFAGPSKHFK  65
T4      MSDLKSRLIKASTSKLTAELTASKFFNEKDVVRTKIPMMNIALSGEITGGMQSGLLILAGPSKSFK  66

RB49    SNMGLTCVSAYMKQNPDAACLFFDSEFGITSAYLESMGVDPDRVVHVPIKNIEELKFEIM  125
T4      SNFGLTMVSSYMRQYPDAVCLFYDSEFGITPAYLRSMGVDPERVIHTPVQSLEQLRIDMV  126

RB49    NQLEQITREDKVIIFIDSIGNLASKKEVEDAINEKSAQDMTRAKALKGLFRMVTPYLTMN  185
T4      NQLDAIERGEKVVVFIDSLGNLASKKETEDALNEKVVSDMTRAKTMKSLFRIVTPYFSTK  186

RB49    DIPCIAINHTYETQEMFSKTVMSGGTGAMYSANEVFIIGRRQQKEGTEITGYDFILNAEK  245
T4      NIPCIAINHTYETQEMFSKTVMGGGTGPMYSADTVFIIGKRQIKDGSDLQGYQFVLNVEK  246

RB49    SRTVKEKSKFPISVTFSGGIDPYSGLLELAVELGWVVKPSNGWYSRSILNTETGEMETEE  305
T4      SRTVKEKSKFFIDVKFDGGIDPYSGLLDMALELGFVVKPKNGWYAREFLDEETGEMIREE  306

RB49    RKFRAKETNSIEFWKPLLTNDKFNEAINDHYKLGQVISDEAVDKEIEDML  355
T4      KSWRAKDTNCTTFWGPLFKHQPFRDAIKRAYQLGAIDSNEIVEAEVDELI  356
```

FIGURA 2

100

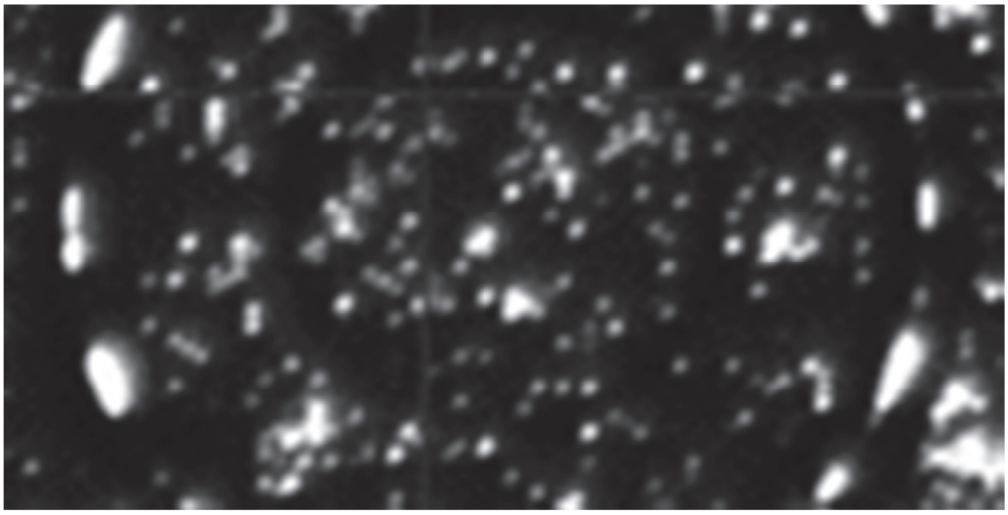


FIGURA 3A

150

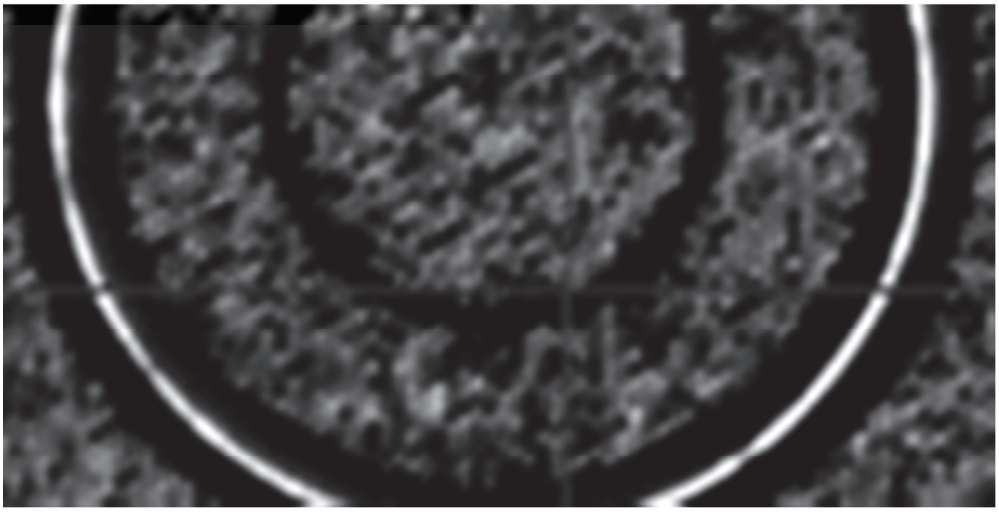


FIGURA 3B

200

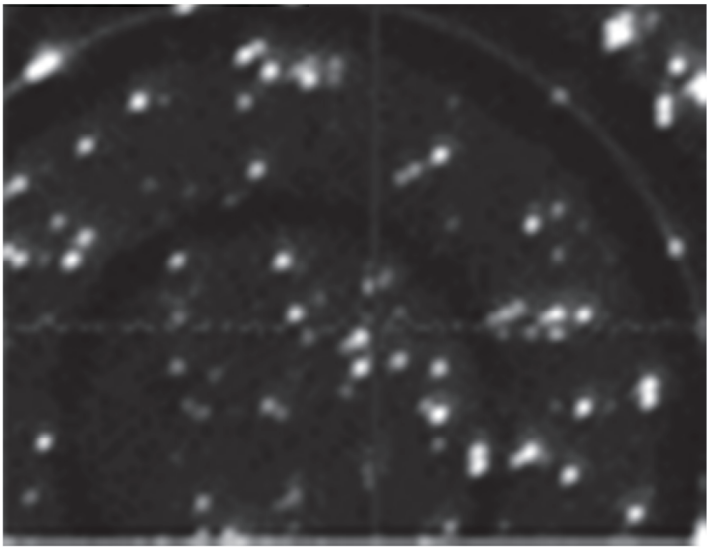


FIGURA 4A

250

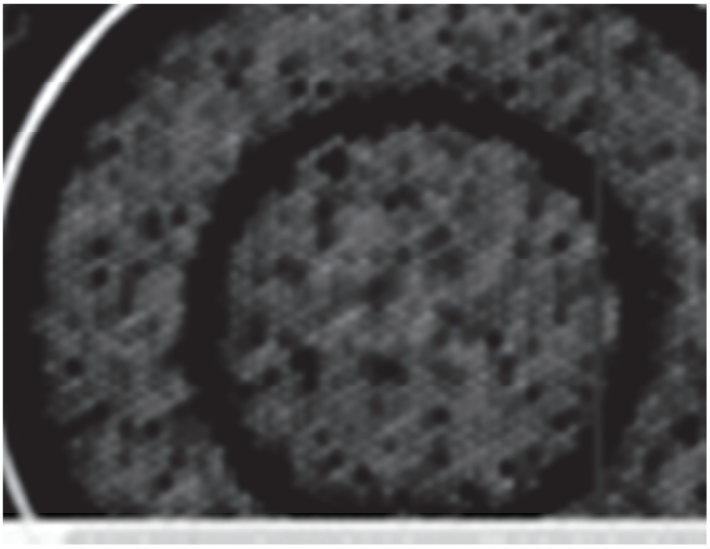


FIGURA 4B