



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

(21) PI 1008429-0 A2



(22) Data do Depósito: 16/02/2010

(43) Data da Publicação Nacional: 25/08/2020

(54) Título: MÉTODOS DE CRIAÇÃO E DE TRIAGEM DE BIBLIOTECAS DE DNA CODIFICADO.

(51) Int. Cl.: C40B 40/08; C40B 70/00.

(30) Prioridade Unionista: 13/02/2009 US 61/152,508.

(71) Depositante(es): X-CHEM, INC..

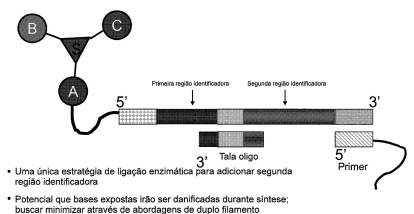
(72) Inventor(es): RICHARD W.WAGNER.

(86) Pedido PCT: PCT US2010024314 de 16/02/2010

(87) Publicação PCT: WO 2010/094036 de 19/08/2010

(85) Data da Fase Nacional: 15/08/2011

(57) Resumo: MÉTODOS DE CRIAÇÃO E DE TRIAGEM DE BIBLIOTECAS DE DNA CODIFICADO. A presente invenção refere-se a uma série de métodos para identificar um ou mais compostos que se ligam a um alvo biológico. Os métodos incluem a síntese de uma biblioteca de compostos, em que os compostos contêm uma porção funcional tendo uma ou mais posições de diversidade. A porção funcional dos compostos é operativamente ligada a um oligonucleotídeo iniciador que identifica a estrutura da molécula funcional.



Relatório Descritivo da Patente de Invenção para **"MÉTODOS DE CRIAÇÃO E DE TRIAGEM DE BIBLIOTECAS DE DNA CODIFICADO"**.  
ANTECEDENTES DA INVENÇÃO

O custo crescente do fármaco descoberto levou à busca permanente por novos métodos de triagem de maior espaço químico tão econômico quanto possível para encontrar moléculas com maior potência e pouca ou nenhuma toxicidade. Abordagens de química combinatória na década de 80 foram originalmente anunciadas como sendo métodos para transcender o paradigma de descoberta de fármacos, mas não tiveram êxito devido a tamanhos da biblioteca insuficientes e métodos inadequados de deconvolução. Recentemente, o uso de bibliotecas combinatórias de DNA exibido de pequenas moléculas criou uma nova mudança de paradigma para a seleção de compostos de chumbo terapêutico.

Morgan et al. (Pedido de Patente norte-americano nº 2007/0224607, incorporado por referência) identifica os principais desafios no uso de abordagens combinatórias de DNA exibido na descoberta do fármaco: (1) a síntese de bibliotecas de complexidade suficiente e (2) a marcação de moléculas que são ativas nas triagens usadas. Além disso, Morgan et al. afirma que quanto maior o grau de complexidade de uma biblioteca, ou seja, o número de estruturas distintas presentes na biblioteca, maior a probabilidade de que a biblioteca contenha moléculas com a atividade de interesse. Assim, a química utilizada na síntese de biblioteca deve ser capaz de produzir um vasto número de compostos dentro de um prazo razoável. Esta abordagem tem sido geralmente bem-sucedida em identificar moléculas com quimiotipos diversificados e de alta afinidade. No entanto, uma série de questões veio à tona com relação à geração de bibliotecas de enorme complexidade e a avaliação da saída de sequenciamento na escala que tem sido descrita. Por exemplo, a purificação de uma biblioteca seguinte a múltiplas transformações químicas (por exemplo, geralmente 3 ou 4 etapas) e transformações biológicas (por exemplo, a ligação enzimática de marcações DNA) é inconveniente e resulta em uma quantidade significativa de "ruído" na biblioteca, seja por síntese incompleta de moléculas ou marcação errada

durante a etapa de ligação. Além disso, a quantidade de sequenciamento que é necessária para interrogar populações selecionadas é impressionante, normalmente exigindo métodos de sequenciamento "nextgeneration" ("próxima geração"). Este último é devido ao fato de que esquemas de marcação genética sofisticada embutidos na parte de DNA da biblioteca, juntamente com algoritmos de bioinformática para a análise da saída de sequenciamento "NextGeneration", são obrigados a filtrar através do ruído e identificar sucessos na biblioteca. Como resultado, mesmo com essas metodologias, o sequenciamento ainda não é avançado o suficiente para capturar totalmente a diversidade de sequências (representando tanto êxito real e "ruído") de uma determinada triagem.

Exibição de DNA de bibliotecas combinatórias de pequena molécula depende de síntese de dividir e agrupar multietapa da biblioteca, acoplada à adição enzimática de marcações de DNA que codificam tanto a etapa sintética quanto o bloco de construção usado. Várias (por exemplo, 3 ou 4) etapas sintéticas são normalmente realizadas e codificadas, e elas incluem posições de diversidade (aqui descritas como A, B e C (figura 1)), tal como aquelas formadas acoplando blocos de construção com, por exemplo, grupos funcionais amina ou carboxilato em um arcabouço químico que mostra os blocos de construção anexados em orientações definidas. Um exemplo de um arcabouço (S), que é frequentemente usado em bibliotecas combinatórias é uma molécula de triazina, que pode ser ortogonalmente derivatizada em três posições sobre a sua estrutura de anel.

O processo de formação da biblioteca pode ser demorado, os produtos são muitas vezes ineficientemente purificados, e o resultado é que as reações não conhecidas podem ocorrer as quais criam moléculas indesejadas e/ou desconhecidas ligadas ao DNA. Além disso, a purificação incompleta da biblioteca pode resultar em contaminação cruzada de marcações durante as etapas de ligação, resultando em marcação errada. O resultado final para a seleção e sequenciamento de hits da biblioteca é que o sequenciamento massivamente paralelo tem que ser empregado devido ao "ruído" inerente de ambos os DNAs que estão ligados às moléculas que são não

intencionais (por exemplo, produtos não reagidos ou secundários) ou que são marcados errado. Assim, a eficiência do sequenciamento é perdida.

Em alguns casos, um oligonucleotídeo iniciador, a partir do qual a biblioteca de pequena molécula é construída, contém uma região de iniciador de ligação para a amplificação da polimerase (por exemplo, PCR) na forma de um oligonucleotídeo de filamento duplo, covalentemente fechado. Esta construção é muito problemática para a realização de reações de polimerase, devido à dificuldade de derreter o oligonucleotídeo dúplice e permitir que um iniciador se ligue e inicie a polimerização, que resulta em uma reação ineficiente, reduzindo o rendimento em 10 - até 1000 vezes ou mais.

Existe uma necessidade de uma abordagem mais faseada para triagem e identificação de moléculas pequenas que têm maior potência e pouca ou nenhuma toxicidade.

#### SUMÁRIO DA INVENÇÃO

A presente invenção apresenta um método para a criação e triagem de bibliotecas simplificadas de DNA codificado, devido ao menor número de etapas sintéticas (por exemplo, nenhuma ligação enzimática ou nenhum oligonucleotídeo iniciador duplo filamentado covalentemente fechado) e, portanto, substancialmente menos "ruído" durante a análise dos oligômeros codificados (aqui chamados de "regiões identificadoras"). Assim, o sequenciamento torna-se muito mais eficiente, ou, alternativamente, a análise de micromatriz torna-se possível, levando em consideração as propensões inerentes que podem confundir a interpretação dos dados que podem ser introduzidos por amplificação da região de codificação. Foram também identificados métodos para a criação de uma maior diversidade de reações químicas em vez de simplesmente limitados a condições aquosas para tornar a biblioteca de DNA codificado mais hidrofóbica e solúvel em solventes orgânicos nas etapas subsequentes da síntese de biblioteca. Desta forma, as reações químicas podem ser realizadas com o rendimento potencialmente maior, uma maior diversidade de blocos de construção, e fidelidade melhorada das reações químicas.

Assim, a presente invenção apresenta um método de marcação

de bibliotecas químicas de DNA codificado ligando um primeiro grupo funcional de um ligante bifuncional a um oligonucleotídeo iniciador na extremidade 5' do oligonucleotídeo iniciador, em que o oligonucleotídeo iniciador forma uma estrutura em formato de grampo, e ligando um segundo grupo funcional do ligante bifuncional a um componente da biblioteca química. O oligonucleotídeo iniciador pode incluir uma primeira região identificadora e uma segunda região identificadora, de tal forma que a segunda região identificadora hibridiza para a primeira região identificadora do oligonucleotídeo iniciador. A segunda região identificadora pode incluir uma marcação fluorescente (por exemplo, um fluoróforo ou GFP) ou o rótulo biotinilado. Além disso, a segunda região identificadora não é amplificada antes da análise após uma etapa de seleção.

Em outra modalidade, a invenção apresenta um método de criação de bibliotecas de DNA codificado pela (a) criação de um primeiro nodo de diversidade, (b) codificação do primeiro nodo de diversidade em vasos separados, (c) agrupamento do primeiro nodo de diversidade, e (d) divisão do primeiro nodo de diversidade agrupado em um segundo conjunto de vasos separados, em que o primeiro nodo de diversidade reage para formar um segundo nodo de diversidade. Em certas modalidades, o segundo nodo de diversidade não é codificado e agrupado.

Em outra modalidade, a presente invenção apresenta um método para a criação de bibliotecas usando reações químicas semi ou não aquosas (por exemplo, orgânicas) com maior rendimento, uma maior diversidade de blocos de construção, e um maior número de reações químicas que podem ser usados para criar mais bibliotecas combinatórias de DNA marcado do que as anteriormente alcançadas.

Em geral, os métodos da presente invenção proveem um conjunto de bibliotecas que contêm, por exemplo, uma ou duas posições de diversidade em um arcabouço químico que pode ser eficientemente gerado em alto rendimento, selecionado para identificar blocos de construção individuais ps ou combinações de blocos de construção que residem em, por exemplo, uma ou duas posições de diversidade, e iterativamente diversificado

em, por exemplo, uma segunda, terceira, e/ou quarta posição de diversidade para criar moléculas com propriedades melhoradas. Além disso, os métodos aqui descritos permitem uma análise ampla e extensa dos compostos selecionados tendo uma propriedade desejada biológica, que, por sua vez, permite que compostos relacionados com relações estruturais familiares sejam identificados (por exemplo, relações estrutura-atividade).

Por "arcabouço" significa uma porção química que exibe nodo (s) de diversidade em uma geometria especial particular. Nodo (s) de diversidade é tipicamente ligado ao arcabouço durante a síntese da biblioteca, mas em alguns casos, um nodo de diversidade pode ser anexado ao arcabouço antes da síntese da biblioteca (por exemplo, em adição das regiões de identificador). Em algumas modalidades, o arcabouço é derivatizado de tal forma que ele pode ser ortogonalmente desprotegido durante a síntese da biblioteca e, posteriormente, reagido com nodos de diversidade diferentes (por exemplo, usando a marcação de identificador em cada etapa).

Por "região identificadora" significa a porção de marcação de DNA da biblioteca que codifica a adição do bloco de construção à biblioteca.

Por "oligonucleotídeo iniciador" significa os oligonucleotídeos de partida para a síntese de biblioteca que contém também um ligante covalentemente ligado e uma porção funcional para adição de um nodo de diversidade ou arcabouço. O oligonucleotídeo pode ser simples ou duplo filamentado. O oligonucleotídeo pode consistir em bases naturais ou modificadas.

Por "porção funcional" significa uma porção química compreendendo um ou mais blocos de construção que podem ser selecionados a partir de qualquer molécula pequena ou projetados e construídos com base em características desejadas, por exemplo, solubilidade, disponibilidade de doadores e receptores de ligação de hidrogênio, graus rotacionais de liberdade das ligações, carga positiva, carga negativa, e assim por diante. A porção funcional deve ser compatível com a modificação química de tal forma que ela reaja com a cabeça. Em certas modalidades, a porção funcional pode ser ainda reagida como uma entidade bifuncional ou trifuncional (ou superior). Porções funcionais também podem incluir blocos de construção que são u-

sados em qualquer um dos nodos de diversidade ou posições. Exemplos de blocos de construção e marcações de DNA de codificação são encontrados nas Tabelas 1 e 2. Ver, por exemplo, publicação do pedido de patente norte-americano nº 2007/0224607, incorporado por referência.

5 Por "bloco de construção" significa uma unidade química estrutural que está ligada a outras unidades químicas estruturais ou pode ser ligada a outras tais unidades. Quando a porção funcional é polimérica ou oligomérica, os blocos de construção são as unidades monoméricas do polímero ou oligômero. Blocos de construção também podem incluir uma estrutura de

10 arcabouço (por exemplo, um bloco de construção de arcabouço) à qual são, ou podem ser anexadas uma ou mais estruturas adicionais (por exemplo, blocos de construção periférica). Os blocos de construção podem ser quaisquer compostos químicos que são complementares (isto é, os blocos de construção devem ser capazes de reagir para formar uma estrutura que in-

15 clui dois ou mais blocos de construção). Normalmente, todos os blocos de construção usados terão pelo menos dois grupos reativos, embora alguns dos blocos de construção usados terão apenas um grupo reativo cada. Grupos reativos em dois blocos de construção diferentes devem ser complementares, ou seja, capazes de reagir para formar uma ligação covalente.

20 Por "ligante" significa uma molécula que liga a parte de ácido nucleico da biblioteca a espécies funcionais exibidas. Tais ligantes são conhecidos na técnica, e aqueles que podem ser usados durante a síntese da biblioteca incluem, mas não estão limitados a, 5'-O-Dimetoxitritil-1',2'-Dideoxirribose-3'-[(2-cianoetil)-(N,N-di-isopropil)]-fosforamidita; 9-O-

25 Dimetoxitritil-trietileno glicol, 1-[(2-cianoetil)-(N,N-di-isopropil)]-fosforamidita; 3-(4,4'-Dimetoxitritilóxi)propil-1-[(2-cianoetil)-(N,N-di-isopropil)]-fosforamidita; e 18-O-Dimetoxitritilhexaetilenoglicol, 1-[(2-cianoetil)-(N,N-di-isopropil)]-fosforamidita. Tais ligantes podem ser adicionados um depois do outro em diferentes combinações para gerar ligantes de diferentes comprimentos de-

30 seados. Por "ligante ramificado" significa uma molécula que se liga à posição de ácido nucleico da biblioteca em 2 ou mais espécies funcionais idênticas da biblioteca. Ligantes ramificados são bem conhecidos na técnica e

exemplos podem consistir em duplicadores simétricos ou assimétricos (1) e (2) ou um triplicador simétrico (3). Ver, por exemplo, Newcome et al., *Dendritic Molecules: Concepts, Synthesis, Perspectives*, VCH Publishers (1996); Boussif et al., *Proc. Natl. Acad. Sci. USA* 92: 7297-7301 (1995); and Jansen et al., *Science* 266: 1226 (1994).

Como usado aqui, o termo "oligonucleotídeo" se refere a um polímero de nucleotídeos. O oligonucleotídeo pode incluir DNA ou qualquer derivado do mesmo conhecido na técnica que pode ser sintetizado e utilizado para reconhecimento de pares de base. O oligonucleotídeo não tem que ter bases contíguas, mas pode ser intercalado com porções de ligante. O polímero de oligonucleotídeo pode incluir nucleosídeos naturais (por exemplo, adenosina, timidina, guanosina, citidina, uridina, desoxiadenosina, desoxitimidina, desoxiguanosina e desoxicitidina), análogos de nucleosídeos (por exemplo, 2-aminoadenosina, 2-tiotimidina, inosina, pirrolo-pirimidina, 3-metil adenosina, C5-propinilcitidina, C5-propiniluridina, C5-bromouridina, C5-fluorouridina, C5-iodouridina, C5-metilcitidina, 7-deazaadenosina, 7-deazaguanosina, 8-oxoadenosina, 8-oxoguanosina, O(6)-metilguanina, e 2-tiocitidina), bases quimicamente modificadas, bases biologicamente modificadas (por exemplo, bases metiladas), bases intercaladas, açúcares modificados (por exemplo, 2'-fluororribose, ribose, 2'-desoxirribose, arabinose e hexose), e/ou grupos de fosfato modificado (por exemplo, fosforotioatos e ligações 5'-N-fosforamidita).

Por "operativamente ligadas" significa que duas estruturas químicas são ligadas entre si de tal forma a manterem-se ligadas através das várias manipulações que espera-se que elas sejam submetidas. Normalmente, a porção funcional e o oligonucleotídeo de codificação estão ligados covalentemente através de um grupo de ligação apropriado. Por exemplo, o grupo de ligação pode ser uma molécula bifuncional com um sítio de fixação para o oligonucleotídeo de codificação e um sítio de fixação para a porção funcional.

Por "pequena molécula" significa uma molécula que tem um peso molecular inferior a cerca de 1000 Dáltons. Pequenas moléculas podem



ser orgânicas ou inorgânicas, e podem ser isoladas a partir de, por exemplo, bibliotecas de compostos ou fontes naturais, ou podem ser obtidas por derivatização de compostos conhecidos.

5 Outras características e vantagens da invenção serão evidentes a partir da seguinte descrição detalhada, desenhos, exemplos, e reivindicações.

#### BREVE DESCRIÇÃO DOS DESENHOS

Figura 1 é um esquema ilustrando as posições de diversidade A, B e C.

10 Figura 2 é um esquema de um membro da biblioteca química de DNA codificado do Modo 1, mostrando, em parte, o oligonucleotídeo iniciador, que inclui uma estrutura em formato de grampo complementar na região identificadora, que foi reagida com nodos de diversidade A e B. A região i-  
dentificadora para B está sendo adicionada. Nesta figura, o nodo de diversi-  
15 dade "C" é a posição potencial para uma posição diversidade adicional a ser adicionada após a adição da região identificadora B.

Figura 3 é um esquema de um membro da biblioteca química de DNA codificado do Modo 1, mostrando, em parte, o oligonucleotídeo iniciador, que inclui uma sequência na região de laço da estrutura em formato de  
20 grampo que pode servir como uma região de ligação de iniciador para amplificação.

Figura 4 é um esquema de um membro da biblioteca química de DNA codificado do Modo 1, mostrando, em parte, o oligonucleotídeo iniciador, que inclui uma sequência não complementar na extremidade 3' da mo-  
25 lécula que pode servir para ligar uma segunda região identificadora à polimerização ou à ligação enzimática.

Figura 5 é um esquema de um membro da biblioteca química de DNA codificado do Modo 1, mostrando, em parte, o oligonucleotídeo iniciador, em que a região de laço do oligonucleotídeo iniciador e pelo menos a  
30 região identificadora no lado 3' da região de laço pode servir para hibridizar com um oligonucleotídeo complementar, que também contém uma segunda região identificadora.

Figura 6 é um esquema de amplificação PCR do modelo em formato de grampo, conforme apresentado na Figura 5.

Figura 7 é um esquema de um membro da biblioteca química de DNA codificado do Modo 2, mostrando um oligonucleotídeo em formato de grampo que é covalentemente fechado (por exemplo, através de uma estrutura em formato de grampo ou quimicamente) na extremidade distal ao ligante.

Figura 8 é um esquema de um membro da biblioteca química de DNA codificado do Modo 2, mostrando a inclusão de nodos de diversidade adicionais.

Figura 9 é um esquema de um membro da biblioteca química de DNA codificado do Modo 2, mostrando as etapas para a triagem de bibliotecas e métodos para deconvolucionar regiões identificadoras.

Figura 10 é um esquema mostrando oligonucleotídeos usados na biblioteca sintética. Cabeça (HP) foi sintetizada por DNA IDT e HPLC purificada. Setas indicam o sítio para restrição de BbvCI (sublinhado) ou digestão denteada de Nb.BbvCI or Nt.BbvCI. Sequências das marcações de DNA A1, B1, e C1 (filamentos superior e inferior), os iniciadores PCR 5' e 3', e a extremidade 3' da HP são também conhecidos.

Figura 11 é um gel eletroforético (TBE-ureia (15%) gel de eletroforese; UV sombreando uma placa de TLC) da cabeça em diferentes etapas desta síntese. Cabeça HP (DNA IDT) foi acilada por Fmoc-amino-PEG2000-NHS (JenKem Technology USA). Rota 1 é a HP (DNA IDT) oligonucleotídeo (42 nts). Rota 2 é HP acilada com Fmoc-amino-PEG2000-NHS. Após adição de Tris-HCl, desproteção de Fmoc é observada. rota 3 é a reação bruta com piperidina, mostrando desproteção completa de Fmoc. rota 4 é a mesma que a rota 3 após dessalgar em uma coluna NAP-5 e liofilização. (XC: xileno cianol (migra como 60 nt DNA); BPB: bromofenol azul (migra como 15 nt DNA))

Figura 12 é um esquema mostrando as etapas em síntese de biblioteca modelo. DTAF foi conjugado com cabeça modificada por amino-PEG (HP-1) na primeira etapa. Após esta etapa, uma porção de HP-1-DTAF foi adicionalmente acilada com pentilamino-biotina.

Figura 13A é um esquema da ligação das marcações de DNA. Fig 13B ilustra 4% de gel de agarose de biblioteca HP-1-DTAF-biotina em diferentes etapas da ligação de marcação de DNA. M: marcador; Rota 1: HP-1-DTAF-biotin; Rota 2: 1+ marcação A apenas; Rota 3: 1 + marcações A, B, e C, bem como extremidade 3' oligoligada. Setas indicam fluorescência verde brilhante (DTAF). Nenhuma separação substancial é observada no gel. Figura 13C ilustra amplificação PCR (24 ciclos) das reações de ligação. M: marcador (banda inferior é 100); Rota 1: amplificação PCR da banda fluorescente verde a partir da rota 1 da figura 14B (HP-1-DTAF-biotina + marcação A); Rota 2: amplificação PCR da banda fluorescente verde a parti da rota 2 da figura 13B (HP-1-DTAF-biotina + todas as 3 marcações e extremidade 3' oligo); Rota 3: amplificação PCR da reação de ligação bruta HP-1-DTAF-biotina + todas as 3 marcações; Rota 4: nenhum controle modelo.

Figura14 é um conjunto de géis eletroforéticos mostrando a purificação do composto modelo XChem e seleção modelo (via uma interação de ligação entre a porção de biotina do composto modelo XChem e estreptavidina). Os géis são géis 4-12% de SDS NuPage com tampão de execução MES. Géis foram escaneados por fluorescência verde usando um laser 450-nm. Figura 14A é um gel mostrando etapas de síntese e purificação. Amostras foram misturadas com tampão de carregamento e fervidas. M: marcador; Rota 1: HP-1 + DTAF; Rotas 2 e 2a: HP-1-DTAF + biotina (duas reações independentes); Rotas 3-6 (etapas de purificação/seleção modelo usando contas de estreptavidina Dynal): Rota 3: fluxo através; Rota 4: última lavagem (lavado com água a 80°C por 10 minutos); Rotas 5 e 5': eluição com 25 mM de EDTA a 90°C (1° e 2°); Rotas 6 e 6': eluição com 25 mM de EDTA e 5 mM de NaOH a 90°C (1° e 2°). Figura 14B é um gel mostrando ligação de HP-1-DTAF-biotina ("biblioteca de 1") a estreptavidina. Amostras foram misturadas com gel carregando tampão e diretamente carregado no gel sem fervura. Amostras, como no gel da figura 14A, foram incubadas com um excesso de estreptavidina em 50 mM de NaCl/10 mM de Tris HCl, pH 7.0, por 10 minutos. "S" indica a adição de estreptavidina. Amostras 5 e 6 foram agrupadas. Rota 1: HP-1-DTAF; Rota 1S: HP-1-DTAF + estreptavidina; Rota

2: HP-1-DTAF-biotina (dessalgada); Rota 2S: HP-1-DTAF-biotina + estreptavidina; Rota 4: última lavagem (lavagem com água a 80°C por 10 minutos); Rota 4S: última lavagem amostra + estreptavidina; Rota 5+6: amostras agrupadas 5, 5', 6 e 6' (frações de eluição a partir de beads de estreptavidina, HP-1-DTAF-biotina purificada e selecionada; Rota 5+6S': HP-1-DTAF-biotina + estreptavidina purificada e selecionada. Note que não existe nenhuma diferença notável na migração entre diferentes etapas de "biblioteca de síntese 1". Figura 14C é 4% de gel de agarose de cabeça (Trilink) HP-T, reagido com DTAF. Rota 1: Marcador; Rota 2: DTAF; Rota 3 HP-T-DTAF. Painel esquerdo: visualização UV do gel (colorido com brometo de etídio); painel direito: mesmo gel escaneado por fluorescência em comprimento de onda de excitação 450 nm (verde, fluoresceína). Figura 14D é 4-12% de SDS de gel NuPage com tampão de execução MES, mostrando ligação de HP-T-DTAF-biotina a estreptavidina. Amostras foram misturadas com tampão carregando em gel e diretamente carregado no gel sem fervura. Amostras, como no gel da figura 14A, foram incubadas com um excesso de estreptavidina em 50 mM de NaCl/10 mM de Tris HCl, pH 7.0, por 10 minutos. Rota 1: DTAF; Rota 2: HP-T-DTAF; Rota 3: HP-T-DTAF + estreptavidina; Rota 4: HP-T-DTAF-biotina (dessalgada); Rota 5: HP-T-DTAF-biotina + estreptavidina; Rota 6: amostras agrupadas 5, 5', 6 e 6' (frações de eluição a partir de beads de estreptavidina, HP-1-DTAF-biotina purificada e selecionada; Rota 7: HP-1-DTAF-biotina + estreptavidina purificada e selecionada.

Figura 15A é um esquema da síntese do construto para experimento de liberação intercelular T7 RNAP. O clone VH dsDNA foi amplificado por PCR para afixar um sítio BsmI na extremidade 5' a montante do promotor T7. Após digestão restrita e purificação, o construto foi ligado a HP-1-DTAF-R7 (cabeça modificada com DTAF e peptídeo (-Arg-εAhx)<sub>6</sub>-Arg). Figura 15B é um gel eletroforético da reação de ligação. Rotas 1 e 2 mostram diferentes amostras HP-1 ligadas a VH; Rota 3 mostra produto PCR VH não ligado; e M é o marcador. Figura 15C é um gel eletroforético mostrando validação pelo promotor T7. O gel mostra uma reação de Megascript T7 (Ambion, Inc.) usando amostras das Rotas 1-3 da figura 15B.

Figura 16 é uma eletroforese em gel de agarose das etapas na síntese 10x10 da biblioteca. Figura 16A é 4% de gel de agarose de cabeça (Trilink) HP-T ligado com marcação A. Rota 1: Marcador; Rota 2: HP-T; Rota 3: marcação A anelada; Rota 4: HP-T ligado com marcação A; Rota 5: HP-T ligado com marcação A e dessalgado em coluna Zeba. Figura 16B é 2% de gel de agarose de ligação HP-T-A com 12 marcações B diferentes. Rota M: Marcador, Rotas 1 e 9: HP-T-A; Rotas 3, 4, 5, 6, 7, 8, 11, 12, 13, 14, 15 e 16: ligação HP-T-A com marcações B1- B12. Figura 16C é 4% de gel de agarose da biblioteca agrupada (biblioteca B), com marcações A e B1-B12 ligadas, após reação com cloreto cianúrico e aminas B1-B12. Rota 1: Marcador; Rota 2: HP-T-A; Rota 3: Biblioteca-B agrupada e dessalgada em coluna Zeba.

### DESCRIÇÃO DA INVENÇÃO

A presente invenção apresenta uma série de métodos para identificar um ou mais compostos que se ligam a um alvo biológico. Os métodos incluem a síntese de uma biblioteca de compostos, em que os compostos contêm uma porção funcional tendo uma ou mais posições de diversidade. A porção funcional dos compostos é operativamente ligada a um oligonucleotídeo iniciador que identifica a estrutura da porção funcional. Em resumo, o Modo 1 provê uma série de métodos para preservar o caráter duplo filamentado do dsDNA durante síntese de biblioteca, o que é importante durante a etapa de reação química, e pode ser usado (como mostrado nas Figuras 2-6) para a geração de até dois nodos de diversidade. O Modo 2 (Figuras 7-9) antecipa um nodo de diversidade e usa um oligonucleotídeo em formato de grampo que é covalentemente fechado (por exemplo, através de uma estrutura em formato de grampo ou quimicamente) na extremidade distal ao ligante. O Modo 3 provê métodos para criar bibliotecas com um, dois, três ou mais nodos de diversidade. Os Modos 1, 2 e 3 são descritos em detalhes abaixo.

#### Modo 1

A presente invenção apresenta um método para identificar um ou mais compostos que se ligam a um alvo biológico. O método inclui a sín-

tese de uma biblioteca de compostos, em que os compostos contêm uma porção funcional tendo não mais do que duas posições de diversidade. A porção funcional dos compostos é operativamente ligada a um oligonucleotídeo iniciador que identifica a estrutura da porção funcional, provendo uma  
5 solução contendo compostos de iniciador A.

O oligonucleotídeo iniciador inclui um ligante L (por exemplo, polietileno glicol) com um número inteiro de um ou mais, em que os oligonucleotídeos iniciadores contêm uma porção funcional que inclui blocos de construção A anexados ao L e separados em vasos de reação A, em que A é um  
10 número inteiro de dois ou maior, que é operativamente ligado a um oligonucleotídeo iniciador que identifica os blocos de construção A.

Em algumas modalidades, os blocos de construção A podem ser adicionalmente derivatizados através de um nodo comum S. Em outras modalidades, A é subsequentemente transformado com S, S sendo uma molécula de arcabouço que permite nodos adicionais da introdução de diversidade. Em algumas modalidades, A-S pode ser triado diretamente, representando um único nodo de diversidade. Em outras modalidades, os vasos de reação A-S (por exemplo, que pode primeiro incluir uma purificação de A-S a partir de matérias-primas) são misturados e aliquotados em vasos de reação  
15 B, em que B é um número inteiro de um ou maior, e reagidos com um dos blocos de construção B. A-S-B, ainda em vasos de reação B, é em alguns casos reagido com um bloco de construção C, em que C é um número inteiro de um, é purificado e submetido a uma reação de polimerização ou ligação usando iniciadores B, em que os iniciadores B diferem em sequência e  
20 identificam os blocos de construção B.

Em certas modalidades, A-S pode ser um número inteiro de um. Em uma modalidade, A-S pode ser ligado diretamente a oligonucleotídeos iniciadores B, e seguinte a reação de blocos de construção B, as reações B são mistas. Em certas modalidades, a mistura A-S-B, em que B representa  
30 somente o nodo de diversidade, é exibida diretamente, representando um único nodo de diversidade. Em outras modalidades, a mistura A-S-B, em que B representa somente o nodo de diversidade, é subsequentemente aliquota-

da em vasos de reação C, reagida com blocos de construção C, e submetida a segunda polimerização de filamento ou reação de ligação usando iniciadores C, na qual os iniciadores C diferem em sequência e identificam os blocos de construção C.

- 5                    Em certas modalidades, B pode ser um número inteiro de um e A-S é maior que um, caso em que A-S, agora derivatizado com B, é aliquotado em vasos de reação C, reagido com blocos de construção C, e submetido a segunda reação de polimerização de filamento usando iniciadores C, em que os iniciadores C diferem em sequência e identificam os blocos de
- 10 construção C. Esta estratégia geral pode ser expandida para incluir nodos de diversidade adicionais (por exemplo, D, E, F, etc.) para que o primeiro nodo de diversidade seja reagido com blocos de construção e/ou S e codificado por um oligonucleotídeo inicial, misturado, realiquotado em vasos e, em seguida, o nodo de diversidade subsequente é derivatizado por blocos de
- 15 construção, que são codificados pelo iniciador utilizado para a polimerização ou reação ligação.

- Em certas modalidades, A pode ser um número inteiro de um, B pode ser um número inteiro de um, e oligonucleotídeos iniciadores C são usados. A-S-B, ligado a oligonucleotídeos iniciadores C, é formado em vasos
- 20 de reação C, reagido com blocos de construção C, e exibido diretamente.

                    Em certas modalidades, S é reagido primeiro com o oligonucleotídeo iniciador, e A, B e/ou C (por exemplo, ou D, E, F, e assim por diante) são posteriormente reagidos.

- Em certas modalidades, A, B ou C (por exemplo, ou D, E, F, e
- 25 assim por diante) podem conter sítios para nodos de diversidade adicionais. Se este for o caso, então S pode ou não pode ser usado ou necessário para introduzir nodos de diversidade adicionais.

- Em uma modalidade, o oligonucleotídeo iniciador inclui uma estrutura em formato de grampo complementar na região identificadora (figura
- 30 2). A região identificadora pode ser, por exemplo, de 2 a 100 pares de bases de comprimento, de preferência 5 a 20 pares de bases de comprimento, e mais preferivelmente 6 a 12 pares de bases de comprimento. O oligonucleo-

tídeo iniciador adicionalmente inclui uma sequência na região de laço da estrutura em formato de grampo que pode servir como uma região de ligação do iniciador para amplificação (figura 3), de tal forma que a região de ligação do iniciador tenha uma temperatura de fusão mais elevada para seu iniciador complementar (por exemplo, que pode incluir acompanhar as regiões identificadoras) do que a região identificadora sozinha.

Em uma modalidade, a região de laço pode incluir bases modificadas que podem formar formações dúplex de mais afinidade do que bases não modificadas, tais bases modificadas sendo conhecidas na técnica (figura 3). O oligonucleotídeo iniciador pode adicionalmente incluir uma sequência não complementar na extremidade 3' da molécula que pode servir para ligar uma segunda região identificadora tanto para a polimerização ou para ligação enzimática (figura 4). Em uma modalidade, os filamentos podem ser posteriormente reticulados, por exemplo, usando psoraleno.

Em outra modalidade, a região de laço e pelo menos a região identificadora no lado 3' da região de laço podem servir para hibridizar com um oligonucleotídeo complementar, que também contém uma segunda região identificadora (figura 5). Nos casos onde muitos blocos de construção e marcações correspondentes são usados (por exemplo, 100 marcações), uma estratégia de misturar e dividir pode ser empregada durante a etapa de síntese de oligonucleotídeos para criar o número necessário de marcações. Tais estratégias de misturar e dividir para a síntese de DNA são conhecidas na técnica. Em uma modalidade, os filamentos podem ser posteriormente reticulados, por exemplo, usando psoraleno. Os membros da biblioteca resultante podem ser amplificados por PCR após seleção para entidades de ligação contra um alvo (s) de interesse (figura 6).

Por exemplo, uma cabeça, que inclui um oligonucleotídeo iniciador, pode ser reagida com um ligante e A, que inclui, por exemplo, 1000 variantes diferentes. Para cada bloco de construção A, uma marcação de DNA A pode ser ligada ou iniciador estendido para a cabeça. Estas reações podem ser realizadas em, por exemplo, uma placa de 1000 poços ou placas de 10 x 100 poços. Todas as reações podem ser agrupadas, opcionalmente



purificadas, e divididas em um segundo conjunto de placas. Em seguida, o mesmo procedimento pode ser realizado com blocos de construção B, que também incluem, por exemplo, 1000 variantes diferentes. Uma marcação de DNA B pode ser ligada à cabeça, e todas as reações podem ser agrupadas.

- 5 Uma biblioteca de 1000 x 1000 combinações de A a B (ou seja, 1.000.000 compostos), marcada por um milhão de diferentes combinações de marcações. A mesma abordagem pode ser estendida para adicionar variantes C, D, E, etc. A biblioteca gerada pode então ser usada para identificar compostos que se ligam ao alvo. A composição dos compostos que se ligam à biblioteca pode ser avaliada por PCR e sequenciamento das marcações de DNA para identificar os compostos que foram enriquecidos.
- 10

### Modo 2

- Em outra modalidade (figura 7), o método inclui sintetizar uma biblioteca de compostos, em que os compostos contêm uma porção funcional tendo não mais do que duas posições de diversidade. A porção funcional dos compostos é operativamente ligada a um oligonucleotídeo iniciador, que contém uma sequência genética única que identifica a estrutura da porção funcional provendo uma solução compreendendo compostos iniciadores A, em que L é um número inteiro de um ou mais, onde os compostos iniciadores incluem uma porção funcional tendo blocos de construção A separados em vasos de reação A, onde, por exemplo, A é um número inteiro de dois ou mais, que é operativamente ligado a um oligonucleotídeo inicial que identifica os blocos de construção A. Em algumas modalidades, os blocos de construção A são pré-derivatizados com um S comum. Em outras modalidades, A é posteriormente transformado com S, sendo S uma molécula de arcabouço que permite nodos adicionais de introdução de diversidade. Em seguida, os vasos de reação A-S (que podem incluir uma primeira purificação de A-S a partir de matérias-primas) são misturados e aliquotados em vasos de reação B, em que B é um número inteiro de um ou maior, e reagido com um dos blocos de construção B. A-S-B, ainda em vasos de reação B é, em algumas modalidades, reagido com um bloco de construção C, onde C é um número inteiro de um, é purificado, e mantido separado em vasos B para triagem.
- 15
- 20
- 25
- 30

Em algumas modalidades, A-S é um número inteiro de um. Em uma modalidade, A-S pode ser ligado diretamente a oligonucleotídeos iniciadores B e, após a reação de blocos de construção B, as reações B são misturadas e ali-

5 mantidas separadas em vasos de reação C, reagidas com blocos de construção C, e mantidas separadas em vasos C para triagem. Em outras modalidades, B pode ser um número inteiro de um e A-S é maior que um, caso em que A-S, agora derivatizado com B, é ali-

10 quotado em vasos de reação C reagidos com blocos de construção C, e mantidos separados em vasos C para triagem. Esta estratégia geral pode ser expandida para incluir nodos de diversidade adicionais (por exemplo, D, E, F, etc.) para que o primeiro nodo de diversidade seja reagido com blocos de construção e/ou S e codificado por um oligonucleotídeo iniciador, misturado, realiquotado em vasos, em seguida, o nodo de diversidade subsequente é derivatizado através de blocos de construção e mantido em seus respectivos vasos para triagem (figura 8).

15 Por exemplo, como descrito no Modo 1, uma cabeça, que inclui um oligonucleotídeo iniciador, pode ser reagida com um ligante e blocos de construção A, que incluem, por exemplo, 1000 variantes diferentes. Bloco de construção A, uma marcação de DNA A pode ser ligada ou iniciador estendido para a cabeça. As reações podem ser agrupadas. Em seguida, o mesmo procedimento pode ser realizado com blocos de construção B, mas uma

20 marcação de DNA não é adicionada para B. Como B não é codificado para, todas as reações de "B" podem ser agrupadas (por exemplo, 1000 reações) e uma etapa de seleção pode ser realizada para identificar todos os blocos de construção A que produzem o efeito desejado de ligação com blocos de construção B desconhecidos. Uma biblioteca de blocos de construção A i-

25 dentificada na etapa de seleção (por exemplo, blocos de construção 10 A) pode ser então reagido com os mesmos blocos de construção B, resultando em uma seleção de 10.000 compostos ou menos. Nesta rodada, marcações DNA para B podem ser adicionadas e blocos B que produzem o efeito de

30 ligação desejado em combinação com os, por exemplo, blocos de construção 10A podem ser identificados, resultando em uma convolução etapa a etapa de uma biblioteca inicial de, por exemplo, 1.000.000 compostos. Um

conjunto destes compostos finais poderá ser testado individualmente para identificar o melhor, por exemplo, aglutinantes, ativadores ou inibidores.

Para evitar o agrupamento de todas as reações após síntese B, um leitor BIND (SRU Biosystems), por exemplo, pode ser usado para monitorar ligação sobre uma superfície de sensor em formato de alta taxa de transferência (por exemplo, placas de 384 poços e placas de 1536 poços). Por exemplo, os blocos de construção A podem ser codificados com as marcações de DNA e os blocos de construção B podem ser codificados em posição. Aglutinantes podem ser identificados usando um sensor BIND, sequenciamento e análise de micromatriz ou análise de restrição de digerir das marcações A. Esta análise permite a identificação de combinações de blocos de construção A e B que produzem as moléculas desejadas. Outros métodos para o monitoramento de ligação conhecidos daqueles versados na técnica podem ser utilizados, incluindo, por exemplo, ELISA.

#### 15 Modos 1 e 2

O oligonucleotídeo iniciador dos Modos 1 e 2 pode conter uma estrutura em formato de grampo, complementares na região identificadora. O oligonucleotídeo iniciador contém ainda uma sequência na região de laço da estrutura em formato de grampo que pode servir como uma região de ligação de iniciador para amplificação, de tal forma que a região de ligação do iniciador tenha uma temperatura de fusão mais elevada por seu iniciador complementar (que pode incluir regiões identificadoras de acompanhamento) do que a região identificadora sozinha.

Em uma modalidade, o oligonucleotídeo iniciador inclui uma molécula ligante capaz de ser funcionalmente reagida com blocos de construção. A molécula ligante pode ser conectada diretamente à extremidade 5' do oligonucleotídeo através de métodos conhecidos na técnica ou pode ser incorporada dentro da molécula, por exemplo, fora de uma base derivatizada (por exemplo, a posição C5 de uridina), ou o ligante pode ser colocado no meio do oligonucleotídeo usando técnicas-padrão conhecida na técnica.

O oligonucleotídeo iniciador pode ser filamento simples ou de filamento duplo. A formação de um oligonucleotídeo de filamento duplo pode

ser alcançada através da formação de estrutura em formato de grampo do oligonucleotídeo ou através de reticulação cruzada usando, por exemplo, uma porção de psoraleno, como são conhecidas na técnica.

O oligonucleotídeo iniciador pode conter duas regiões de ligação primárias (por exemplo, para permitir uma reação de PCR) em ambos os lados da região que codifica o identificador do bloco de construção. Alternativamente, o oligonucleotídeo iniciador pode conter um sítio de ligação de iniciador na extremidade 5'. Em outras modalidades, o oligonucleotídeo iniciador é uma estrutura em formato de grampo e a região do laço forma um sítio de ligação de iniciador ou o sítio de ligação do iniciador é introduzido através de hibridização de um oligonucleotídeo para a região identificadora do lado do 3' do laço. Um oligonucleotídeo iniciador, que contém uma região homóloga à extremidade 3' do oligonucleotídeo iniciador e carregando uma região de ligação do iniciador em sua extremidade 5' (por exemplo, para permitir uma reação de PCR) pode ser hibridizado com o oligonucleotídeo iniciador, e pode conter uma região identificadora que codifica os blocos de construção utilizados em uma das posições de diversidade. O oligonucleotídeo iniciador pode conter informações adicionais, tais como uma região de nucleotídeos aleatorizada, por exemplo, 2 a 16 nucleotídeos de comprimento, que está incluída para análise bioinformática.

Em uma modalidade, o oligonucleotídeo iniciador não contém um sítio de ligação do iniciador de PCR.

Em outra modalidade, a biblioteca de compostos, ou uma parte dela, é contatada com um alvo biológico em condições adequadas por pelo menos um membro da biblioteca de compostos para se ligar ao destino, seguido da remoção de membros da biblioteca que não se ligam ao alvo, e analisam a região ou regiões identificadoras. Alvos biológicos exemplares incluem, por exemplo, enzimas (por exemplo, quinases, fosfatases, metilases, demetilases, proteases e enzimas de reparo do DNA), proteínas envolvidas nas interações de proteínas : proteína (por exemplo, ligantes para receptores), alvos receptores (por exemplo, GPCRs e RTKs ), canais de íons, bactérias, vírus, parasitas, DNA, RNA, príons, ou hidratos de carbono).

Em uma modalidade, a biblioteca de compostos, ou uma parte dela, é contatada com um alvo biológico em condições adequadas a pelo menos um membro da biblioteca de compostos que se liga ao alvo, seguido da remoção dos membros da biblioteca que não se ligam ao alvo, seguido de amplificação da região identificadora por métodos conhecidos na técnica e, posteriormente, analisam a região identificadora ou regiões por métodos conhecidos na técnica.

Em uma modalidade o método de amplificação da região identificadora pode incluir, por exemplo, a reação em cadeia polimerase (PCR), amplificação de cadeia linear (LCR), a amplificação de círculo rolante (RCA), ou qualquer outro método conhecido na técnica para amplificar sequências de ácidos nucleicos.

Em uma modalidade adicional, a biblioteca de compostos não é agrupada após a última etapa de adição de bloco de construção e os grupos são avaliados individualmente para identificar compostos que se ligam a um alvo.

Em outra modalidade, as moléculas que se ligam a um alvo não são submetidas à amplificação, mas são analisadas diretamente. Métodos de análise incluem, por exemplo, análise de micromatriz ou métodos baseados em beads decovolucionar as regiões identificadoras (figura 9). Moléculas que se ligam durante a etapa de triagem também podem ser detectadas por um biossensor de cristal fotônico livre de rótulo.

Em uma modalidade, o oligonucleotídeo iniciador e/ou o oligonucleotídeo iniciador contém uma porção funcional que permite a sua detecção por, por exemplo, marcas fluorescentes, pontos Q, ou biotina.

Em uma modalidade, a análise de micromatriz usa capacidade de detecção avançada, tal como, por exemplo, cristais fotônicos de ressonância evanescente.

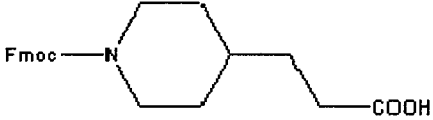
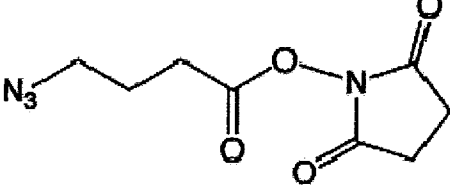
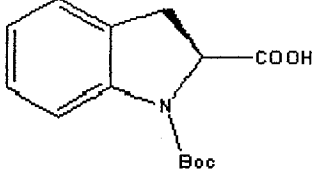
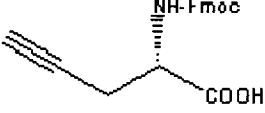
Em uma modalidade, o método de amplificação inclui formar uma emulsão de água em óleo para criar uma pluralidade de microrreatores aquosos, em que pelo menos um dos microrreatores tem pelo menos um membro de uma biblioteca de compostos que se liga ao alvo, uma única be-

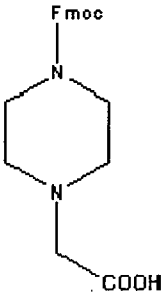
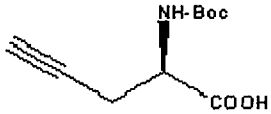
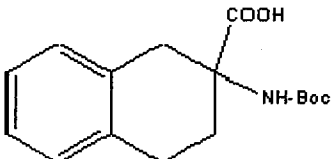
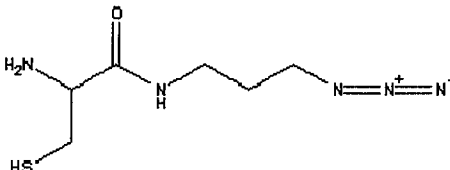

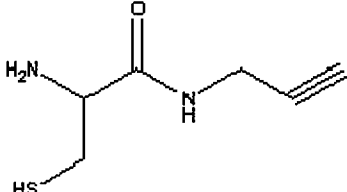
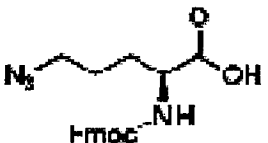
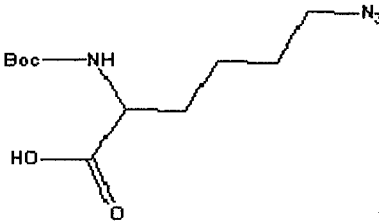
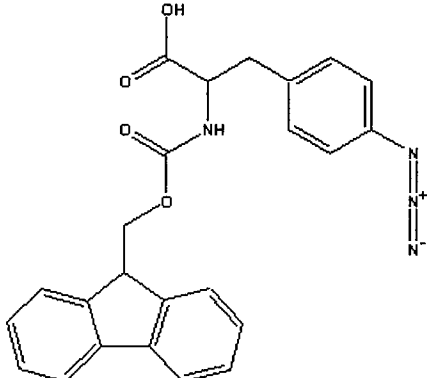
ad capaz de se ligar ao oligonucleotídeo de codificação do pelo menos um membro da biblioteca de compostos que se liga ao alvo, e uma solução da reação de amplificação contendo os reagentes necessários para realizar a amplificação de ácido nucleico, amplificando o oligonucleotídeo de codificação nos microrreatores para formar cópias ampliadas do oligonucleotídeo de codificação, e ligando as cópias amplificadas do oligonucleotídeo de codificação às contas nos microrreatores.

Uma vez que os blocos de construção da primeira biblioteca que se liga ao alvo de interesse foram identificados, uma segunda biblioteca pode ser preparada de forma iterativa, em que um ou dois nodos adicionais de diversidade são adicionados, e a biblioteca é criada e diversidade amostrada como aqui descritos. Este processo pode ser repetido tantas vezes quantas forem necessárias para criar moléculas com propriedades moleculares e farmacêuticas desejadas.

Um bloco de construção exemplar inclui, por exemplo, aminoácidos (não limitados a alfa-aminoácidos), reagentes de clique químico (por exemplo, azida ou cadeias alcalinas) com uma amina ou um reagente tiol. A escolha de um bloco de construção depende, por exemplo, da natureza do grupo reativo utilizado no ligante, a natureza de uma molécula de arcabouço, e o solvente utilizado para a síntese química. Ver, por exemplo, a Tabela 1.

Tabela 1. Blocos de Construção de Posição A Exemplares

 <p>Ácido 3-(1-Fmoc-piperidina-4-il)-propiónico</p>	 <p>Éster de N-Hidroxisuccinimida de ácido 4-Azido-butan-1-oico</p>
 <p>Ácido Boc-L-indolina-2-carboxílico</p>	 <p>Fmoc-L-propargilglicina</p>

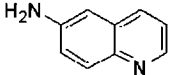
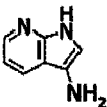
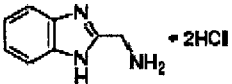
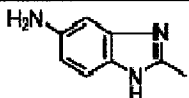

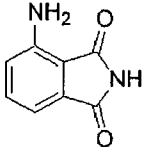
 <p>Fmoc-(4-carboximetil)piperazina</p>	 <p>Boc-D-propargilglicina•DCHA</p>
 <p>Ácido Boc-2-amino-1,2,3,4-tetra- hidronaftaleno-2 -carboxílico</p>	 <p>2-Amino-N-(3-azidopropil)-3-mer- capsuperiorropionamida</p>
 <p>Ácido (S)-(-)-2-azido-6-(boc- amino)hexanoico</p>	 <p>2-Amino-3-mercapto-N-(prop-2 - inil)propionamida</p>
 <p>Ácido (S)-5-azido-2-(Fmoc- amino)pentanoico</p>	 <p>Boc-Lys(N<sub>3</sub>)-OH</p>
 <p>Fmoc-4-azidoFenilalanina</p>	

Blocos de construção B e C exemplares estão descritos nas Tabelas 2 e 3, respectivamente. Um sítio de restrição pode ser introduzido, por

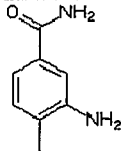
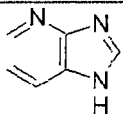
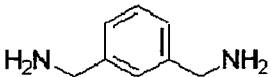
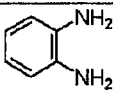
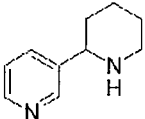
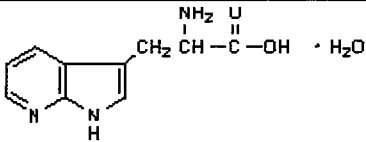
exemplo, na posição B ou C para análise do produto final e seleção através da realização de PCR e restrição de digestão com uma das enzimas de restrição correspondentes.

**Tabela 2. Exemplos de Blocos de Construção de Posição B e marcações de**

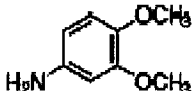
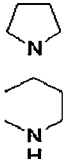
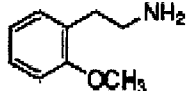
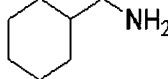

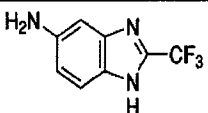
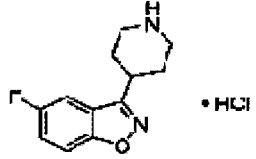
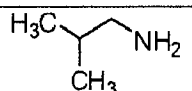
**5 DNA de Codificação**

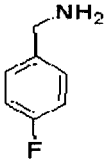
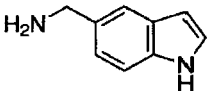
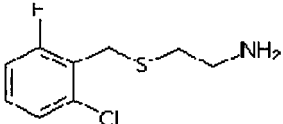
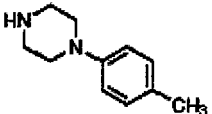
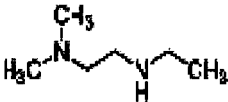
Nome Químico e Estrutura	Sítio de Restrição (Enzima de Restrição)	Filamento Superior Filamento Inferior
 6-Aminoquinolina (B1)	T/CCGGA (BspEI)	5'-Fos-CCTCCGGAGA (SEQ ID NO: 1) 5'-Fos-TCCGGAGGAC (SEQ ID NO: 2)
 3-Amino-7-azaindol, 1H-pirrolo[2,3-b]piridin-3-ilamina (B2)	GGC/GCC (SfoI)	5'-Fos-CCGGCGCCGA (SEQ ID NO: 3) 5'-Fos-GGCGCCGGAC (SEQ ID NO: 4)
 Cloridrato de 2-(aminometil)benzimidazol (B3)	GGTAC/C (KpnI)	5'-Fos-CCGGTACCGA (SEQ ID NO: 5) 5'-Fos-GGTACCGGAC (SEQ ID NO: 6)
 2-Metil-1H-benzimidazol-5-amina (B4)	CAC/GTG (PmlI)	5'-Fos-CCCACGTGGA (SEQ ID NO: 7) 5'-Fos-CACGTGGGAC (SEQ ID NO: 8)
 (Aminometil)ciclopropano (B5)	GAGCT/C (SacI)	5'-Fos-CCGAGCTCGA (SEQ ID NO: 9) 5'-Fos-GAGCTCGGAC (SEQ ID NO: 10)
 3-Aminoftalimida (B6)	G/GATCC (BamHI)	5'-Fos-CCGGATCCGA (SEQ ID NO: 11) 5'-Fos-GGATCCGGAC (SEQ ID NO: 12)



Nome Químico e Estrutura	Sítio de Restrição (Enzima de Restrição)	Filamento Superior Filamento Inferior
 <b>3-Amino-4-metilbenzamida (B7)</b>	AT/CGAT (Bsp-DI)	5'-Fos-CCATCGATGA (SEQ ID NO: 13) 5'-Fos-ATCGATGGAC (SEQ ID NO: 14)
 <b>4-Azabenzimidazol (B8)</b>	A/AGCTT (HindII)	5'-Fos-CCAAGCTTGA (SEQ ID NO: 15) 5'-Fos-AAGCTTGGAC (SEQ ID NO: 16)
 <b>m-Xililendiamina (B9)</b>	A/GATCT (BglII)	5'-Fos-CCAGATCTGA (SEQ ID NO: 17) 5'-Fos-AGATCTGGAC (SEQ ID NO: 18)
 <b>1,2-Fenilenodiamina (B10)</b>	G/AATTC (EcoRI)	5'-Fos-CCGAATTCGA (SEQ ID NO: 19) 5'-Fos-GAATTCGGAC (SEQ ID NO: 20)
 <b>Anabasina (B11)</b>	T/GATCA (BclI)	5'-Fos-CCTGATCAGA (SEQ ID NO: 21) 5'-Fos-TGATCAGGAC (SEQ ID NO: 22)
 <b>Hidrato de DL-7-azatriptofano (B12)</b>	CA/TATG (NdeI)	5'-Fos-CCCATATGGA (SEQ ID NO: 23) 5'-Fos-CATATGGGAC (SEQ ID NO: 24)

**Tabela 3. Exemplos de Blocos de Construção de Posição C de Marcações de DNA de Codificação**

Nome Químico e Estrutura	Filamento Superior Filamento Inferior
 3,4-Dimetoxianilina (C1)	5'-Fos-GAACCTGCTT (SEQ ID NO: 25) 5'-Fos-GCAGGTTCTC (SEQ ID NO: 26)
 4-(1-Pirrolidinil)piperidina (C2)	5'-Fos-GAAGACGCTT (SEQ ID NO: 27) 5'-Fos-GCGTCTTCTC (SEQ ID NO: 28)
 2-Metoxifenetilamina (C3)	5'-Fos-GACCAGACTT (SEQ ID NO: 29) 5'-Fos-GTCTGGTCTC (SEQ ID NO: 30)
 Ciclo-hexanometilamina (C4)	5'-Fos-GACGACTCTT (SEQ ID NO: 31) 5'-Fos-GAGTCGTCTC (SEQ ID NO: 32)
 2-(1-Ciclo-hexenil)etilamina (C5)	5'-Fos-GACGCTTCTT (SEQ ID NO: 33) 5'-Fos-GAAGCGTCTC (SEQ ID NO: 34)
 5-Amino-2-(trifluorometil)benzimidazol (C6)	5'-Fos-GAGCAACCTT (SEQ ID NO: 35) 5'-Fos-GGTTGCTCTC (SEQ ID NO: 36)
 Cloridrato de 5-fluoro-3-(4-piperidinil)- 1,2-benzisoxazol (C7)	5'-Fos-GAGCCATCTT (SEQ ID NO: 37) 5'-Fos-GATGGCTCTC (SEQ ID NO: 38)
 Isobutilamina (C8)	5'-Fos-GCAACCACTT (SEQ ID NO: 39) 5'-Fos-GTGGTTGCTC (SEQ ID NO: 40)

Nome Químico e Estrutura	Filamento Superior Filamento Inferior
 4-Fluorobenzilamina (C9)	5'-Fos-GCACAGACTT (SEQ ID NO: 41) 5'-Fos-GTCTGTGCTC (SEQ ID NO: 42)
 5-(Aminometil)indol (C10)	5'-Fos-GCGATCACTT (SEQ ID NO: 43) 5'-Fos-GTGATCGCTC (SEQ ID NO: 44)
 2-[(2-cloro-6-fluorobenzil)tio]etilamina (C11)	5'-Fos-GCGGTTACTT (SEQ ID NO: 45) 5'-Fos-GTAACCGCTC (SEQ ID NO: 46)
 1-(4-Metilfenil)piperazina (C12)	5'-Fos-GCATGACCTT (SEQ ID NO: 47) 5'-Fos-GGTCATGCTC (SEQ ID NO: 48)
 N,N-Dimetil-N'-etiletilenodiamina (C13)	5'-Fos-GCGTACTCTT (SEQ ID NO: 49) 5'-Fos-GAGTACGCTC (SEQ ID NO: 50)

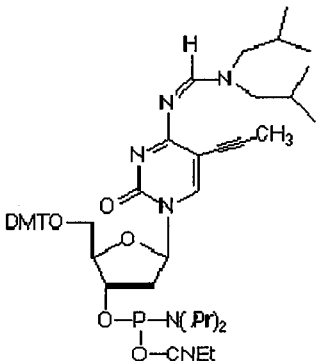
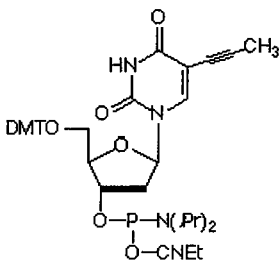
### Modo 3

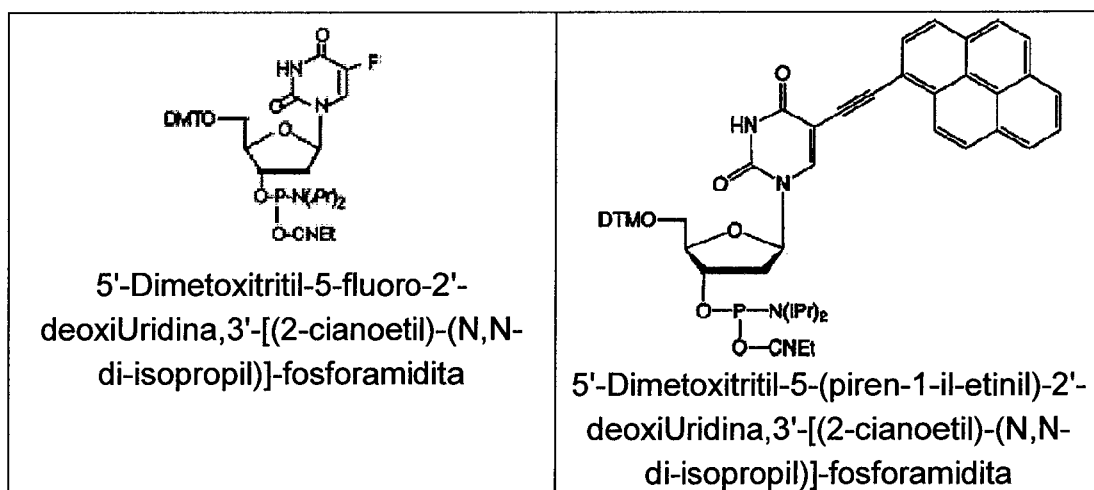
Em qualquer dos modos aqui descritos (por exemplo, modos 1 e 2), o oligonucleotídeo cabeça pode ser modificado para suportar a solubilidade em condições semi ou não aquosas (por exemplo, orgânicas). A cabeça, em certas modalidades, inclui a região identificadora. Em outras modalidades, a cabeça com o ligante pode ser primeiro derivatizado com um bloco de construção (por exemplo, uma porção funcional) ou arcabouço, e a sequência identificadora é então adicionada.

Bases de nucleotídeos do cabeça podem ser tornadas mais hidrofóbicas, modificando, por exemplo, as posições C5 de bases T ou C com cadeias alifáticas sem significativamente prejudicar sua capacidade de ligação de hidrogênio em suas bases complementares. Ver, por exemplo, a Tabela 4 para exemplos de bases modificadas. Além disso, o oligonucleotídeo

cabeça pode ser intercalado com as modificações que promovam a solubilidade em solventes orgânicos. Por exemplo, fosforamidita de azobenzeno pode introduzir uma porção hidrofóbica no projeto cabeça. Tais inserções de amiditas hidrofóbicas na cabeça podem ocorrer em qualquer parte da molécula. No entanto, a inserção não pode interferir com a marcação subsequente usando marcações DNA adicionais durante a síntese de biblioteca ou assegurando reações de PCR, uma vez que uma seleção é finalizada e análise de micromatriz, se usada por deconvolução de marcação. Tais adições ao projeto cabeça aqui descritas tornaria o cabeça solúvel em, por exemplo, 15%, 25%, 30%, 50%, 75%, 90%, 95%, 98%, 99%, ou 100% de solventes orgânicos. Assim, a adição de resíduos hidrofóbicos no projeto cabeça permite a solubilidade melhorada em condições semi ou não aquosas (por exemplo, orgânicas), durante renderização do cabeça competente para a marcação de ácidos nucleicos. Além disso, marcações de DNA que são posteriormente introduzidas na biblioteca também podem ser modificadas na posição C5 de bases T ou C de tal forma que elas também tornam a biblioteca mais hidrofóbica e solúvel em solventes orgânicos para as etapas subsequentes da síntese de biblioteca.

Tabela 4. Exemplar modificado de bases de nucleotídeos

 <p>5'-Dimetoxitritil-N4-di-isobutylaminometilideno-5-(1-Propinil)-2'-deoxiCitidina,3'-[(2-cianoetil)-(N,N-di-isopropil)]-fosforamidita</p>	 <p>5'-Dimetoxitritil-5-(1-Propinil)-2'-deoxyUridina,3'-[(2-cianoetil)-(N,N-di-isopropil)]-fosforamidita</p>
--	--



A molécula ligante entre o cabeça e uma biblioteca de pequena molécula pode ser variada para aumentar a solubilidade do cabeça em solvente orgânico. Uma grande variedade de ligantes está comercialmente disponível os quais podem acoplar a cabeça com a biblioteca de moléculas pequenas. Ligantes são empiricamente selecionados para um projeto de uma determinada biblioteca de pequena molécula (arcabouços e blocos de construção) de tal forma que a biblioteca pode ser sintetizada em solvente orgânico, por exemplo, 15%, 25%, 30%, 50%, 75%, 90%, 95 %, 98%, 99%, ou 100% de solvente orgânico. O ligante pode ser variado utilizando reações modelo anterior à síntese de biblioteca para selecionar o comprimento da cadeia apropriado que solubiliza a cabeça em solvente orgânico. Ligantes podem incluir ligantes com, por exemplo, o comprimento aumentado da cadeia alquila, unidades aumentadas de polietileno glicol, espécies ramificadas com cargas positivas (para neutralizar as cargas de fosfato negativas sobre o cabeça), ou quantidades aumentadas de hidrofobicidade (por exemplo, além de estruturas de anel de benzeno).

A molécula ligante pode fornecer um espaçador adequado entre o DNA de cabeça e o membro de uma biblioteca de química. Por exemplo, ligantes bifuncionais podem ser usados. Em certas modalidades, ligantes bifuncionais podem incluir, por exemplo, três partes. Parte 1 pode ser um grupo reativo, que forma uma ligação covalente com o DNA, tal como, por exemplo, um ácido carboxílico, de preferência ativado por um éster de N-hidróxi succinimida (NHS) para reagir com um grupo amino sobre o DNA

(por exemplo, dT de amino modificado), uma amidita para modificar a extremidade 5' ou 3' de uma cabeça de DNA de filamento único (obtido por meios da química de oligonucleotídeos-padrão), pares de clique química (cicloadição de azida alcino na presença de catalisador Cu (I)) ou grupos reativos tiol. Parte 2 podem ser também um grupo reativo, que forma uma ligação covalente com a biblioteca química, ou um bloco de construção na posição A ou porção arcabouço. Tal um grupo reativo pode ser, por exemplo, uma amina, um tiol, uma azida, ou um alcino para reações à base de água ou vários outros grupos reativos para as reações de base orgânica. Parte 3 pode ser um espaçador quimicamente inerte de comprimento variável, introduzido entre a parte 1 e 2. Como um espaçador pode ser uma cadeia de unidades de etileno glicol (por exemplo, PEGs de comprimentos diferentes), um alcano, um alceno, a cadeia de polieno, ou cadeia peptídica. O ligante pode conter ramificações ou inserções com porções hidrofóbicas (tal como, por exemplo, anéis de benzeno) para melhorar a solubilidade da cabeça em solventes orgânicos, bem como porções fluorescentes (por exemplo, fluoresceína ou Cy-3) utilizadas para fins de detecção de biblioteca.

Exemplos de ligantes comercialmente disponíveis incluem, por exemplo, ligantes amino-carboxílico (por exemplo, peptídeos (por exemplo, Z-Gly-Gly-Gly-Osu or Z-Gly-Gly-Gly-Gly-Gly-Gly-Osu), PEG (por exemplo, Fmoc-aminoPEG2000-NHS ou amino-PEG (12-24)-NHS), ou cadeias de ácido alceno (por exemplo, ácido-Osu Boc-ε-aminocaproico)), ligantes de clique químico (por exemplo, peptídeos (por exemplo, azidohomalanina-Gly-Gly-Gly-OSu ou propargilglicina-Gly-Gly-Gly-OSu), PEG (por exemplo, azido-PEG-NHS), ou cadeias de ácido alceno (por exemplo, ácido 5-azidopentanoico, (S)-2-(azidometil)-1-Boc-pirrolidina, ou éster de N-hidroxissuccinimida de ácido 4-azido-butan-1-óico)), ligantes reativos tiol (por exemplo, PEG (por exemplo, SM(PEG)<sub>n</sub> NHS-PEG-maleimide), cadeias alceno (por exemplo, ácido-Osu 3-(piridin-2-ildissulfanil)-propionico ou 6-(3'-[2-piridilditio]-propionamido)hexanoato))) de sulfossuccinimidila, amiditas para síntese de oligonucleotídeo (por exemplo, modificadores de amino (por exemplo, 6-(trifluoroacetilamino)-hexil-(2-cianoetil)-(N,N-di-isopropil)-

fosforamidita), modificadores de tiol (por exemplo, S-tritil-6-mercaptohexol-1-[(2-cianoetil)-(N,N-di-isopropil)]-fosforamidita, ou modificadores de clique químico (por exemplo, 6-hexin-1-il-(2-cianoetil)-(N,N-di-isopropil)-fosforamidita, 3-Dimetoxitritilóxi-2-(3-(3-propargiloxipropanamido)propanamido)propil-1-O-succinoilal, alquilamino de cadeia longa CPG, ou éster de de N-hidroxissuccinimida de ácido 4-azido-butan-1-oico)).

Resíduos hidrofóbicos no projeto de cabeça podem ser variados com o projeto do ligante para facilitar síntese de biblioteca em solventes orgânicos. Por exemplo, a combinação de cabeça e ligante é projetada para ter resíduos apropriados em que o coeficiente de octanol : água (Poct) seja de, por exemplo, 1.0 a 2.5.

### EXEMPLOS

Os seguintes exemplos pretendem ilustrar a invenção. Eles não pretendem limitar a invenção em nenhuma maneira.

#### Exemplo 1. Preparação da cabeça (Variante 1)

Um oligonucleotídeo cabeça fosforilado (oligo HP) tendo a seguinte sequência foi sintetizado e HPLC purificado por DNA IDT.

5'-(fosfato)TCCTGGCTGAGGCGAGAGTT(dT-C6-NH) TTCTCTCGCCT-CAGCCAGGACC-3' (SEQ ID NO: 51)

O oligonucleotídeo se dobrou em uma estrutura em formato de grampo (figura 10) com uma inclinação e contém um sítio de clivagem (CCTCAGC) para enzima de restrição BbvCI ou versões recortadas desta enzima Nb.BbvCI ou Nt.BbvCI, que podem clivar o filamento superior ou o inferior (New England BioLabs). No meio do laço em formato de grampo, a cadeia lateral dt C5-amino-modificado dT é inserida (dT-C6-NH; "C6" refere-se ao ligante de carbono 6), que foi usado para o acoplamento do ligante amino-PEG (PEG2000, aproximadamente 45 unidades de etileno glicol). Os filamentos superior e inferior das marcações de DNA A, B, e C foram sintetizados e purificados pro DNA IDT e purificados por dessalgação-padrão. Oligonucleotídeos mais longos, tais como a extremidade 3' e os iniciadores PCR, foram sintetizados por DNA IDT e HPLC purificada.

Dez nanomoles do oligo HP foram dissolvidos em 50 µl de água. Um excesso molar de 20 vezes de éster Fmoc-amino-PEG2000-carboxil-NHS (JenKem Technology USA) foi dissolvido em 50 µl de dimetilformamida (DMF) e foi adicionado à solução de oligonucleotídeo em 2 porções durante o curso de 2 horas à temperatura ambiente (composição de solvente final de 50% DMF/50% de água). Subsequentemente, 60 µl de 1 M de Tris HCl, pH 7.0 (concentração final de 200 mM), foram adicionados para resfriar bruscamente o excesso de ésteres NHS, e a solução foi incubada por um adicional de 30 minutos à temperatura ambiente. A mistura de reação resultante foi diluída em 500 µl com água e foi dessalgada passando através de uma coluna NAP-5 (Sephadex-25, GE Healthcare).

O material resultante foi liofilizado e dissolvido em 100 µl de água. 20 µl de piperidina (a uma concentração final de 20%) foram adicionados e incubados por 2 horas à temperatura ambiente. Um precipitado turvo foi formado devido à desproteção da amina e liberação do grupo grupo Fmoc insolúvel em água. A reação foi em seguida filtrada através de filtros amovíveis de 0.2-µm (Millipore) e precipitada usando 300 mM de acetato de sódio pela adição de 3 volumes de etanol. A forma Fmoc protegido do oligonucleotídeo modificado foi encontrada como sendo solúvel em etanol e isopropanol. Devido a alta eficiência de acoplamento, o cabeça resultante (HP-1) foi usado sem purificação adicional (figura 11).

#### Exemplo 2. Preparação do cabeça (Variante 2)

Uma cabeça completa (HP-1) tendo a seguinte sequência foi preparada por Trilink, Inc., seguindo um procedimento similar como descrito acima, e RP-HPLC purificado.

5'-(fosfato)TCCTGGCTGAGGCGAGAGTT(dT-C6-NH)(X) TTCTCTCGCCT-CAGCCAGGACC-3' (SEQ ID NO: 52)

onde X permanece para amino-PEG2000.

#### Exemplo 3. Síntese de um membro da biblioteca modelo

##### Etapa1: Acoplamento de DTAF

A fim de preparar uma "biblioteca de 1," um composto modelo, 5-(4,6-diclorotriazinil-aminofluoresceína) (DTAF; Anaspec) (figura 12), foi aco-



plado ao grupo amino de HP-1. DTAF estruturalmente representa um arcabouço de triclorotriazina com um composto amino acoplado. Para formar uma biblioteca, arcabouços de triclorotriazina podem ser derivatizados com uma diversidade de blocos de construção em cada uma das três posições do cloro. DTAF também provê um rótulo fluorescente para a biblioteca modelo.

5 A reação (10 µl) foi ajustada como segue. A 5 µl de 400 µM de HP-1 dissolvido em água, 2 µl de 750 mM tampão de borato, pH 9.5, e 1 µl de DMF foram adicionados. DTAF foi dissolvido em DMF para 50 mM e 2 µl foi adicionado à reação. Concentrações finais do HP-1 e DTAF foram 200 µM e 10

10 mM, respectivamente, Assim gerando um excesso de 50 vezes do DTAF. A concentração de DMF final foi 30%. Foi notado que HP-1 permaneceu solúvel em até 90% DMF, demonstrando que ele foi solúvel em um solvente orgânico, por exemplo, DMF. A reação foi permitida prosseguir a 4°C por 16-20 horas. A mistura de reação foi em seguida diluída com água a 30-50 µl e

15 dessalgada em uma coluna de rotação Zeba (Pierce). Nenhuma purificação adicional foi finalizada neste ponto.

#### Etapas 2: Acoplamento de amino-biotina

Após DTAF ter sido acoplado a HP-1, um grupo mais reativo na molécula de arcabouço está ainda disponível para modificação. Foi escolhido um análogo de amino-biotina, EZ-Link Pentilamina-Biotina (Pierce), para se acoplar nesta posição a fim de gerar um composto de ligação modelo (figura 12). A reação foi ajustada como segue. 20 µl de uma mistura de reação continha em torno de 200 pmoles de HP-1-DTAF (Etapa 1) dissolvido em 150 mM de tampão de borato, pH 9.5, e 10 nmoles de pentilamina-biotina. A

20 reação foi permitida prosseguir por 4-12 horas a 75°C. A reação foi em seguida purificada por dessalgação em uma coluna de rotação Zeba, como descrito acima.

#### Etapas 3: Ligação das marcações de DNA a HP-1-DTAF-biotina

Marcações de DNA fosforiladas (região iniciador de extremidade 3' e iniciadores PCR 5' e 3') foram sintetizadas por DNA IDT. Sequências de oligonucleotídeo (figura 13) são como seguem.

30

Marcação de DNA A1 (superior): 5'-fos-GGAGGACTGT (SEQ ID NO: 53)

Marcação de DNA A1 (inferior): 5'-fos-AGTCCTCCGG (SEQ ID NO: 54)

Marcação de DNA B1 (superior): 5'-fos-CAGACGACGA (SEQ ID NO: 55)

Marcação de DNA B1 (inferior): 5'-fos-GTCGTCTGAC (SEQ ID NO: 56)

Marcação de DNA C1 (superior): 5'-fos-CGATGCTCTT (SEQ ID NO: 57)

5 Marcação de DNA C1 (inferior): 5'-fos-GAGCATCGTC (SEQ ID NO: 58)

Extremidade 3' (superior): 5'-fos-GCTGTGCAGGTAGAGTGC-3' (SEQ ID NO: 59)

Extremidade 3' (inferior): 5'-AACGACACGTCCATCTCACG (SEQ ID NO: 60)

Iniciador PCR 5': 5'-CTCTCGCCTCAGCCAGGA (SEQ ID NO: 61)

10 Iniciador PCR 3': 5'-GCACTCTACCTGCACAGC (SEQ ID NO: 62)

Quantidades equivalentes de pares superior e inferior de marcações e oligonucleotídeos de extremidade 3' foram dissolvidas em água e aneladas por aquecimento a 85°C e esfriadas para 4°C em 200 mM de NaCl, 50 mM de Tris HCl, pH 7.0, tampão.

15 Primeiro, a marcação A1 de duplo filamento foi ligada à cabeça. A reação de ligação (20 µl) continha 2.5 µM de HP-1-DTAF-biotina e 2.5 µM de a marcação A1 de duplo filamento em tampão 1x T4 DNA ligase e 60 unidades Weiss de T4 DNA ligase (New England BioLabs). A reação foi incubada a 16°C por 16 horas. O produto resultante não foi redissolvida em

20 qualquer um dos geles testados, incluindo diferentes porcentagens de TBE-ureia, NativePage, SDS-PAGE, ou 2% e 4% de E-gel de agarose (Invitrogen, Inc.). Mobilidade do oligonucleotídeo, modificou com ligante PEG e DTAF-biotina, foi a maior parte determinada pela presença desses grupos em vez de pelo DNA propriamente (dados não mostrados). Para testar a eficiência

25 da ligação, todas as marcações foram ligadas e oligonucleotídeos de extremidade 3' e realizados ensaios PCR do construto resultante para confirmar a eficiência da ligação. A reação de ligação (70 µl) continha: 2.5 µM de HP-1-DTAF-biotina; 2.5 µM de de cada uma das Marcações de DNA duplo filamentadas aneladas (A1, B1, e C1), bem como a marcação de extremidade

30 3'; 1x tampão de T4 DNA ligase; e 210 unidades Weiss de T4 DNA ligase. A reação foi incubada a 16°C por 20 horas.

A mistura de reação foi carregada em 4% de gel de agarose e a

banda fluorescente foi extraída do gel. Este material foi usado para amplificação de PCR de 24 ciclos teste usando iniciadores 5' e 3' como descrito acima. Os resultados são resumidos na figura 13.

#### Etapa 4: Purificação de HP-1-DTAF-biotina em contas de estreptavidina

##### 5 tavidinaa e reação com estreptavidina

Purificação de HP-1-DTAF-biotina em estreptavidina (SA) contas magnéticas Dynal M-280 (Invitrogen) servem como um modelo para seleção por afinidade para a biblioteca química de DNA marcado. Beads de SA foram pré-equilibradas em 2x tampão de PBS contendo 0.05% de Triton X-100. 50 pmoles de HP-1-DTAF-biotina foram carregados em 25 µl das beads SA pré-lavadas por 15 minutos à temperatura ambiente com aplicação por rotação em tambor. A corrente foi coletada e as contas foram lavadas 3 vezes por 30 minutos com 1 ml do mesmo tampão. Uma lavagem final foi realizada a 80°C por 10 minutos com 30 µl de água (coletada). As contas foram eluídas com 30 µl de 25 mM de EDTA e 5 mM de NaOH por 10 minutos a 90°C, e o eluente foi imediatamente neutralizado pela adição de 3 µl de 1 M de Tris HCl, pH 7.0.

Para o experimento de ligação de estreptavidina, 5 µl das amostras de eluição foram incubados com um excesso de estreptavidina em 50 mM de NaCl/10 mM Tris HCl, pH 7.0, por 10 minutos. As amostras foram misturadas com tampão de gel de carregamento sem fervura e redissovidas em um gel NuPage SDS a 4-12% (Invitrogen) usando tampão de execução MES. Os resultados são resumidos na figura 14.

#### Exemplo 4. Acoplamento de peptídeo H(-Arg-εAhx)6-Arg-OH a HP-1-DTAF

25 Foi escolhido um peptídeo rico em arginina R7, H(-Arg-εAhx)6-Arg-OH (Bachem), para usar como outra modificação para o último grupo de reação no arcabouço de triazina. Isto é um peptídeo permeável na membrana celular de ácido um arginina-aminohexanoico usado para liberar um composto intracelular. A reação foi ajustada similarmente às condições de reação descritas acima: 20 µl de reação continha em torno de 200 pmoles de HP-1-DTAF (Etapa 1) dissolvido em 150 mM de tampão de borato, pH 9.5, e 10 nmoles de peptídeo R7. Sob essas condições, as cadeias laterais das

argininas não reagiram, e a única amina reativa no peptídeo é a N-terminal. A reação foi permitida prosseguir por 12 horas a 75°C e foi em seguida purificada por dessalgagem em uma coluna de rotação Zeba.

#### Exemplo 5. Construto de DNA para experimento de detecção de liberação de RNAP T7

O construto de DNA para o experimento de liberação intracelular da "biblioteca de 1" química foi preparado a partir de um produto de PCR de um único clone de DNA VH de ~400 pb caracterizando uma região do promotor T7 na extremidade 5' e uma região Cmu constante de pequeno anti-corpo fechada na extremidade 3' da molécula. A fim de ligar o construto de DNA à cabeça modificado da biblioteca química modelo, um sítio de restrição BsmI foi anexado a jusante da região do promotor T7 por amplificação de PCR do clone. Digestão por restrição de BsmI produziu uma saliência em GG 3', que permitiu ligação ao cabeça (saliência CC 3'). O iniciador 5' com o sítio BsmI (sublinhado) foi sintetizado por DNA IDT, Inc.

5'-GGATGCCGAATGCCTAATACGACTCACTATAGGG-  
ACAATTACTATTTACAATTACA (SEQ ID NO: 63)

Seguindo amplificação de PCR, o construto de DNA foi purificado usando um kit de purificação de PCR (Invitrogen), e o DNA resultante foi digerido com 250 U BsmI (New England BioLabs) a 65°C em tampão de NEB 4 por 2 horas. O DNA foi purificado em um gel de agarose a 2%. A reação de ligação (30 µl) continha 2 pmoles de cada construto DNA VH, digerido com BsmI, bem como HP-1-DTAF-R7 (peptídeo de arginina-ácido aminohexanoico) em 1x tampão DNA ligase T4 e 60 unidades Weiss de T4 DNA ligase (New England BioLabs). A reação foi incubada a 16°C por 20 horas. Devido a alta eficiência da ligação, o material foi adicionalmente usado para a liberação intracelular /experimento T7 RNAP sem purificação adicional. Os resultados são resumidos na figura 15.

#### Exemplo 6. Síntese de 10x10 Biblioteca

##### Etapa 1. Ligação da marcação A à cabeça HP-T

Nesta biblioteca exemplar, apenas posições B e C são usadas. Uma marcação A é ligada a HP-T. A marcação tem a seguinte sequência:

Marcação de DNA A1 (superior): 5'-fos-GGAGGACTGT (SEQ ID NO: 64)

Marcação de DNA A1 (inferior): 5'-fos-AGTCCTCCGG (SEQ ID NO: 65)

30 nmoles de HP-T foram misturados com 45 nmols de cada oligos marcação A1 superior e marcação A1 inferior em 1x tampão T4 DNA  
 5 ligase e foram anelados por aquecimento a 95°C por 1 minuto, seguido por resfriamento a 4°C a 0.2°C/segundo. As amostras foram então trazidas para 16°C. 300 Unidades Weiss de T4 DNA ligase foi adicionado e as amostras foram permitidas incubar por 16-20 horas a 16°C. Seguinte à ligação, HP-T-A foi dessalgado usando uma coluna Zeba (Pierce). Ver, por exemplo, figura  
 10 16A.

#### Etapa 2. Ligação de marcações B1-B12 e marcações C

Doze reações de ligação foram ajustas similarmente às reações de ligação descritas acima. Em cada um dos 12 tubos, 5 nmoles de pares de oligos B1-B12 superior e inferior foram adicionados a 1x tampão T4 DNA  
 15 ligase e anelados como descrito acima. HP-T-A foi dissolvido em 1x tampão T4 DNA ligase. 2.5 nmoles de HP-T-A foram aliquotados nestes 12 tubos. 30 Unidades Weiss de T4 DNA ligase foram adicionadas a cada tubo e as reações foram permitidas prosseguir por 20 horas a 16°C. Seguinte à incubação, cada mistura de reação foi individualmente dessalgada em uma coluna  
 20 de rotação Zeba a 0.5 ml, equilibrada com 150 mM de tampão de borato, pH 9.0. A cada tubo, um excesso de 20x de cloreto cianúrico (50 nmoles), dissolvido em acetonitrila, foi adicionado e incubado por 1,5 hora a 4°C. Seguinte a esta incubação, um excesso de 100x (250 nmoles, isto é, 5x excesso relativo a cloreto cianúrico) de aminas B1-B12, dissolvidas em acetonitrila ou  
 25 DMF, foi adicionado em correspondência com as marcações B1-B12 ligadas. A reação com aminas foi permitida prosseguir por 20 horas a 4°C. Seguinte a esta reação a biblioteca foi agrupada, dessalgada duas vezes em 2-ml de coluna Zeba e liofilizada. Ver, por exemplo, Figuras 16B e 16C.

Como as reações acima, as marcações C e aminas são adicionadas usando condições de reação similares àquelas descritas acima.  
 30

#### Outras modalidades

Todas as publicações, patentes e pedidos de patentes mencio-

nados no relatório acima são incorporados por referência. Várias modificações e variações do método descrito e o sistema da invenção serão evidentes para aqueles versados na técnica sem se afastar do escopo e do espírito da invenção. Embora a invenção tenha sido descrita em conexão com modalidades específicas, deve-se compreender que a invenção, tal como reivindicada não deve ser indevidamente limitada a tais modalidades específicas. De fato, várias modificações dos modos descritos para realizar a invenção que são óbvias para aqueles versados na técnica têm a intenção de estar dentro do escopo da invenção.

10                    Outras modalidades estão nas reivindicações.

## REIVINDICAÇÕES

1. Método de marcação de bibliotecas químicas de DNA codificado, o método compreendendo ligar um primeiro grupo funcional de um ligante bifuncional um oligonucleotídeo iniciador na extremidade 5' do oligonucleotídeo iniciador e ligar um segundo grupo funcional do ligante bifuncional a um componente da biblioteca química, em que o ligante bifuncional ou oligonucleotídeo iniciador é modificado para aumentar a solubilidade de um membro da Biblioteca química de DNA codificado em condições orgânicas.

2. Método, de acordo com a reivindicação 1, em que o oligonucleotídeo iniciador ligado ao ligante bifuncional forma uma estrutura em formato de grampo.

3. Método, de acordo com a reivindicação 1, em que o oligonucleotídeo iniciador compreende uma primeira região identificadora.

4. O método, de acordo com a reivindicação 3, em que o oligonucleotídeo iniciador compreende uma segunda região identificadora que hibridiza a primeira região identificadora do oligonucleotídeo iniciador.

5. Método, de acordo com a reivindicação 1, em que o ligante bifuncional é modificado para aumentar a solubilidade de um membro da biblioteca química de DNA codificado em condições orgânicas.

6. Método, de acordo com a reivindicação 5, em que o ligante bifuncional modificado compreende um ou mais dentre uma cadeia alquila, uma unidade de polietileno glicol, uma espécie ramificada com cargas positivas, ou uma estrutura de anel hidrofóbico.

7. Método, de acordo com a reivindicação 1, em que o oligonucleotídeo iniciador é modificado para aumentar a solubilidade de um membro da biblioteca química de DNA codificado em condições orgânicas.

8. Método, de acordo com a reivindicação 7, em que o oligonucleotídeo iniciador compreende uma primeira região identificadora e uma segunda região identificadora, e em que a primeira região identificadora ou a segunda região identificadora é modificada para aumentar a solubilidade de um membro da biblioteca química de DNA codificado em condições orgânicas.

9. Método, de acordo com a reivindicação 7, em que o oligonucleotídeo iniciador modificado compreende um ou mais dentre um nucleotídeo tendo uma porção hidrofóbica ou uma inserção tendo uma porção hidrofóbica.

5                    10. Método, de acordo com a reivindicação 9, em que o oligonucleotídeo iniciador modificado compreende o nucleotídeo tendo uma porção hidrofóbica, e a porção hidrofóbica é uma cadeia alifática na posição C5.

10                   11. Método, de acordo com a reivindicação 9, em que o oligonucleotídeo iniciador modificado compreende a inserção tendo uma porção hidrofóbica, e a porção hidrofóbica é um azobenzeno.

12. Método, de acordo com a reivindicação 1, em que o membro da biblioteca química de DNA codificado tem um coeficiente de octanol: água de 1,0 a 2,5.

15                   13. Método de marcar bibliotecas químicas de DNA codificado, o método compreendendo ligar um primeiro grupo funcional de um ligante bifuncional a um oligonucleotídeo iniciador na extremidade 5' do oligonucleotídeo iniciador, em que o oligonucleotídeo iniciador ligado ao ligante bifuncional forma uma estrutura em formato de grampo, e ligar um segundo grupo funcional do ligante bifuncional a um componente da biblioteca química.

20                   14. Método, de acordo com a reivindicação 13, em que o oligonucleotídeo iniciador compreende uma primeira região identificadora.

15. Método, de acordo com a reivindicação 14, em que o oligonucleotídeo iniciador compreende uma segunda região identificadora que hibridiza a primeira região identificadora do oligonucleotídeo iniciador.

25                   16. Método, de acordo com a reivindicação 15, em que a segunda região identificadora compreende uma marcação fluorescente ou o rótulo de biotina.

30                   17. Método, de acordo com a reivindicação 16, em que a segunda região identificadora não é amplificada antes da análise após uma etapa de seleção.

18. Método, de acordo com qualquer uma das reivindicações 15, em que o ligante bifuncional, oligonucleotídeo iniciador, primeira região iden-



tificadora, ou segunda região identificadora são modificados para aumentar a solubilidade de um membro da biblioteca química de DNA codificado em condições orgânicas.

5 19. Método de criação de bibliotecas de DNA codificado, o método compreendendo:

(a) criar um primeiro nodo de diversidade;

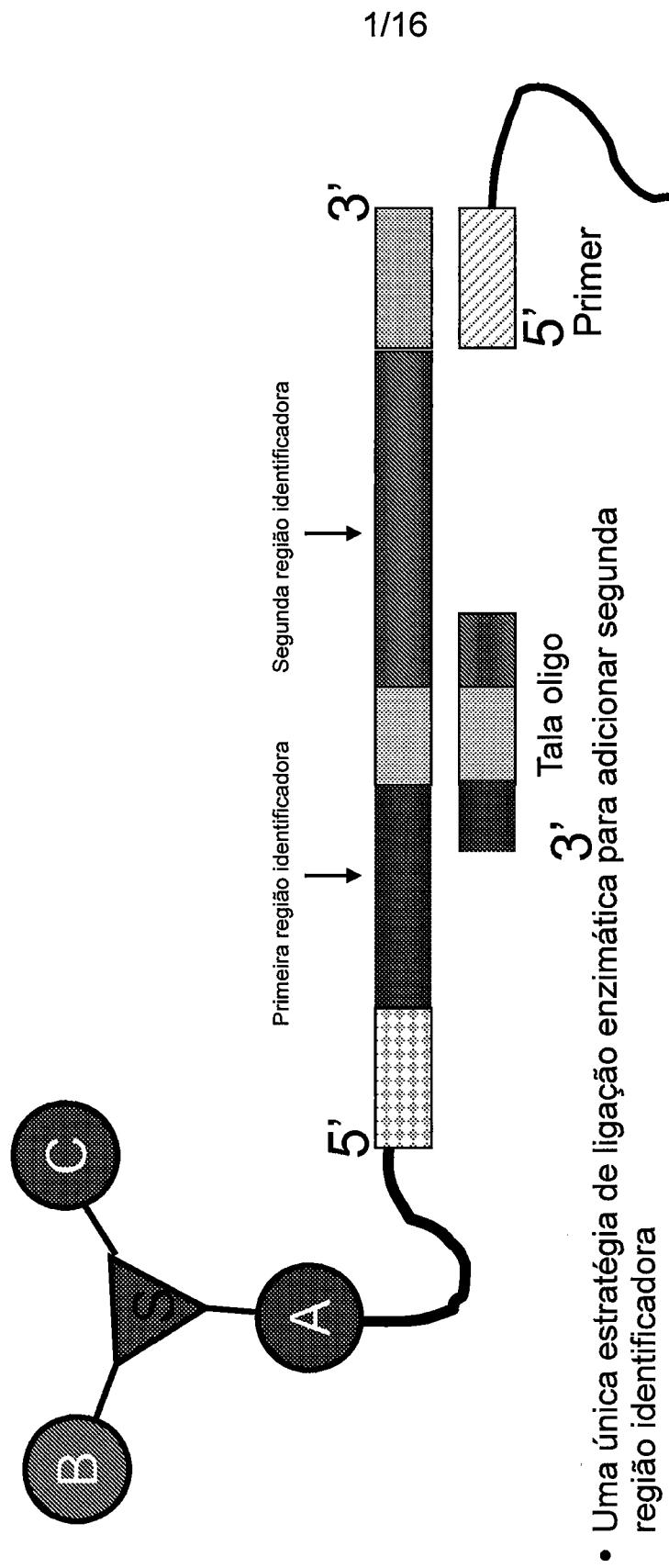
(b) codificar o primeiro nodo de diversidade em vasos separados;

(c) agrupar o primeiro nodo de diversidade, e

10 (d) dividir o primeiro nodo de diversidade agrupado em um segundo conjunto de vasos separados, em que o primeiro nodo de diversidade reage para formar um segundo nodo de diversidade.

20. Método, de acordo com a reivindicação 19, em que o segundo nodo de diversidade não é codificado e agrupado.

FIG. 1



- Uma única estratégia de ligação enzimática para adicionar segunda região identificadora
- Potencial que bases expostas irão ser danificadas durante síntese; buscar minimizar através de abordagens de duplo filamento

FIG. 2

Modo 1

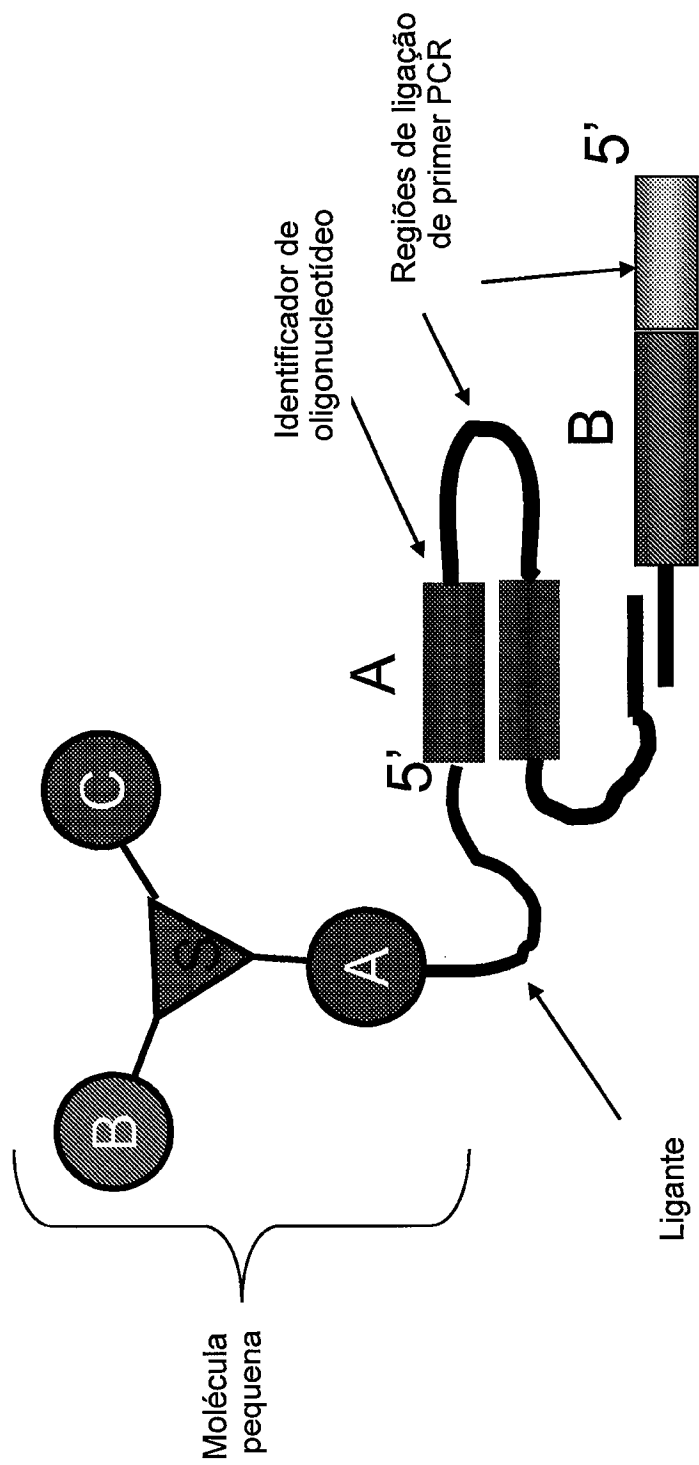


FIG. 3

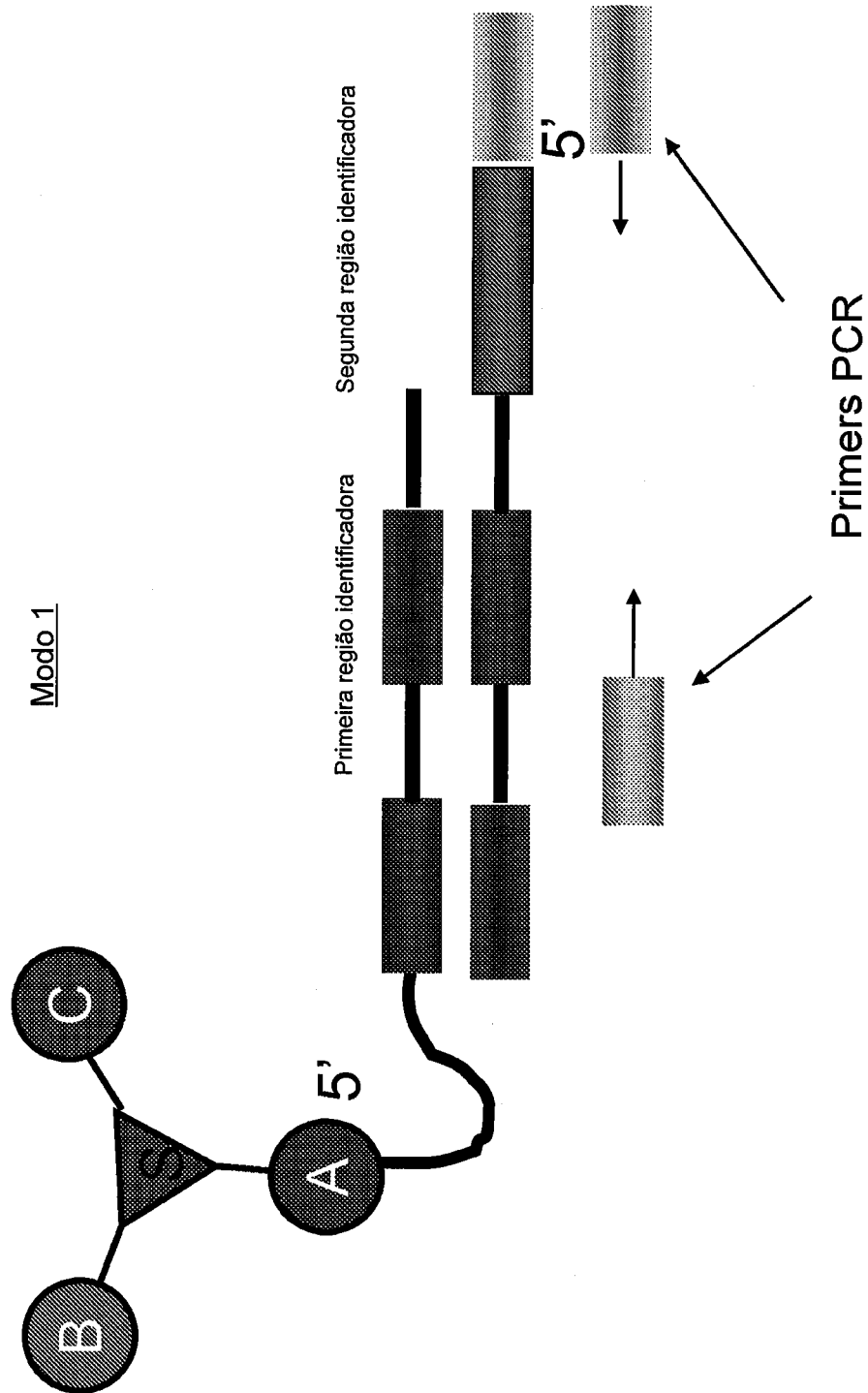
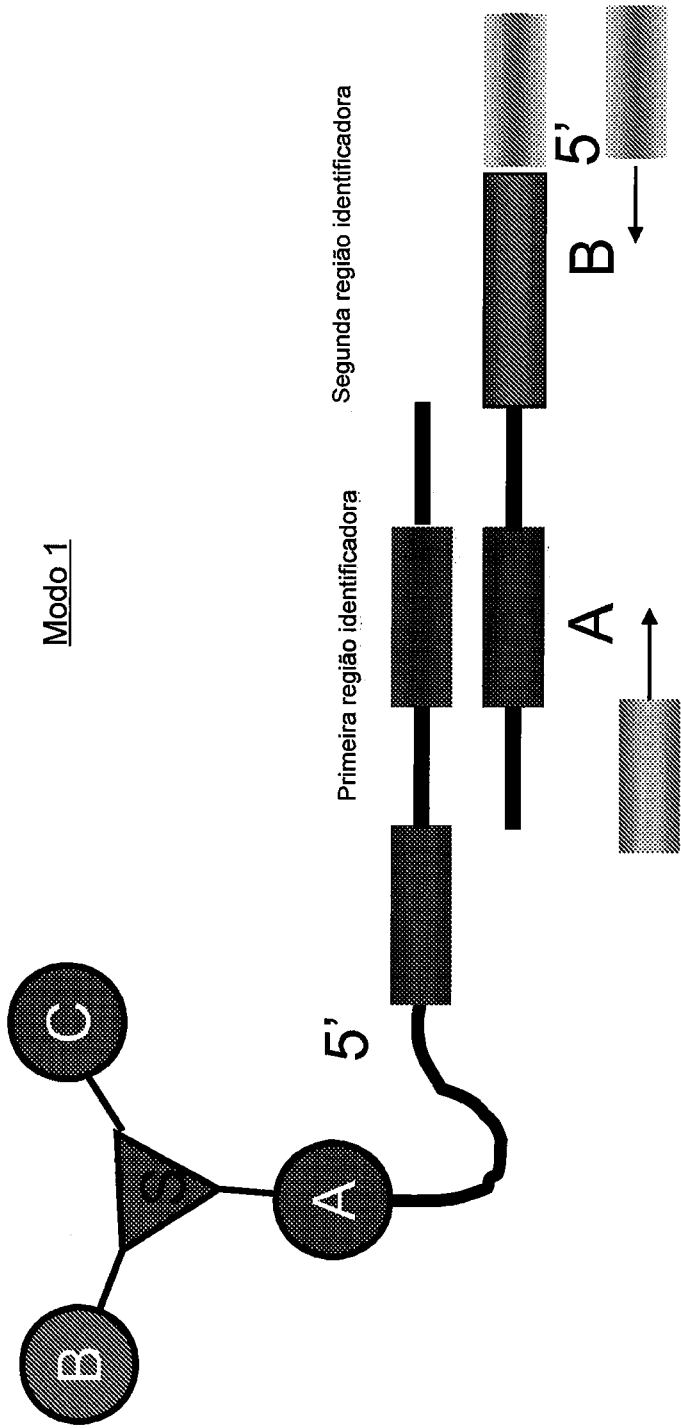


FIG. 4



Alcançado por hibridização de marcações (não enzimáticas)  
Poderia introduzir reticulação cruzada de psoraleno  
Ou polimerização

FIG. 5

Modo 1

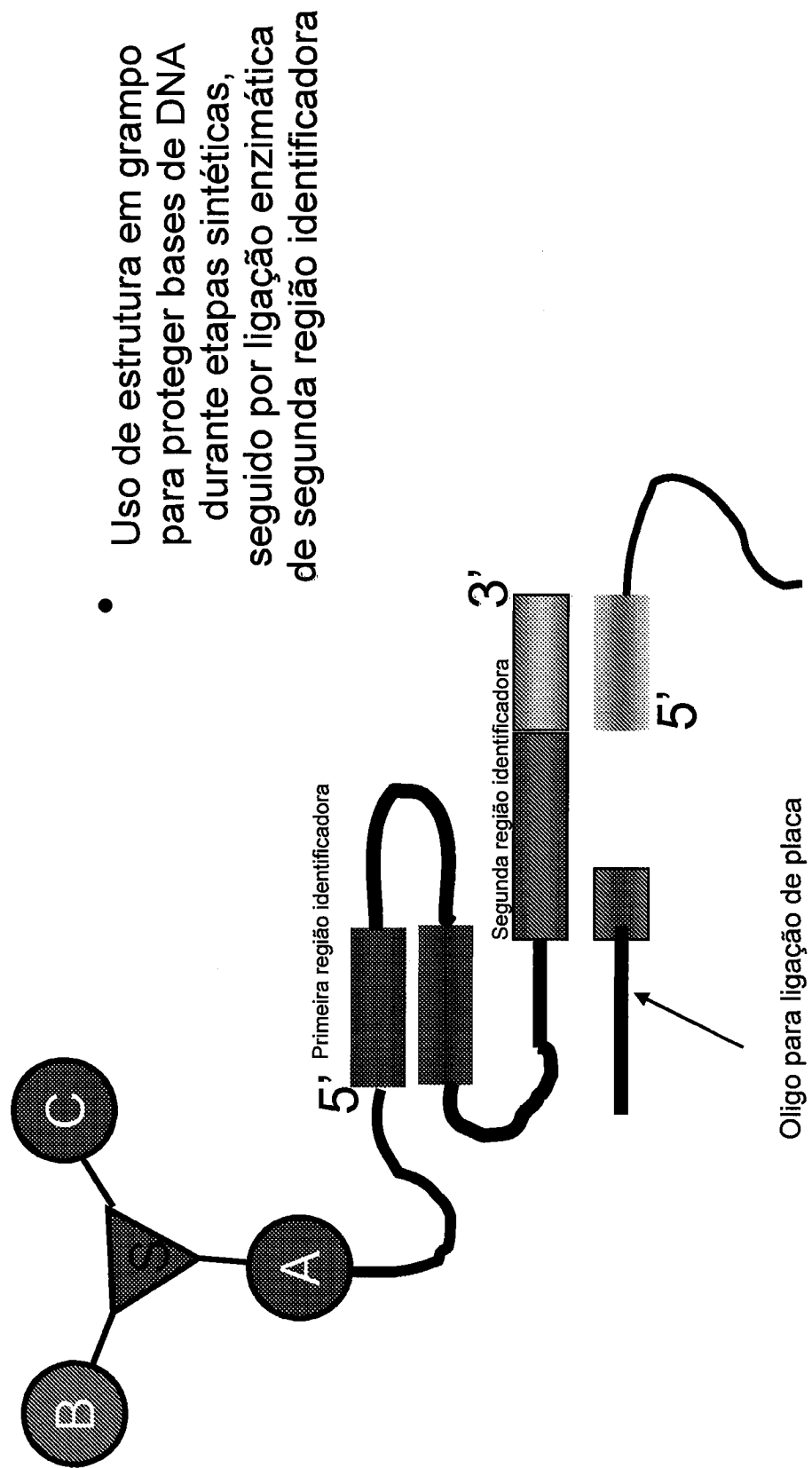
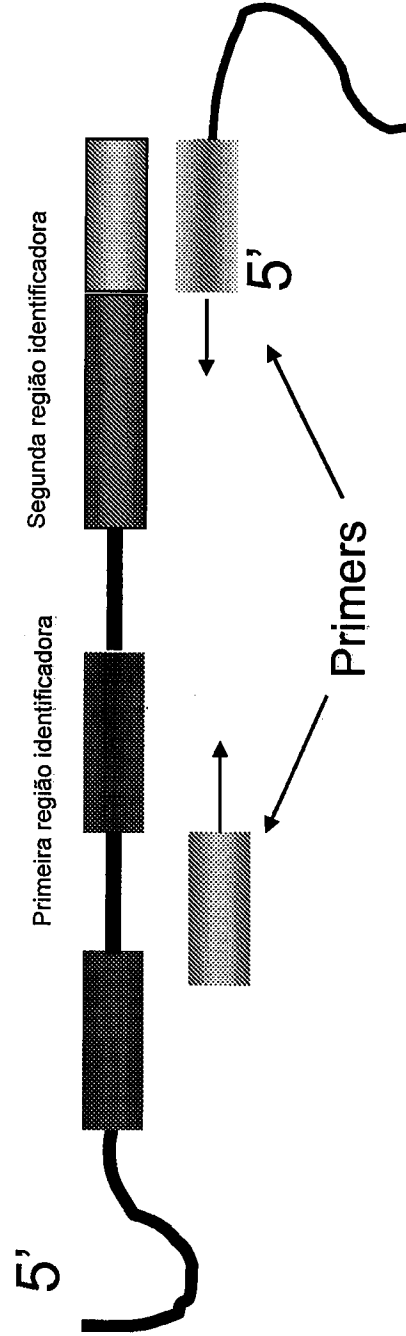


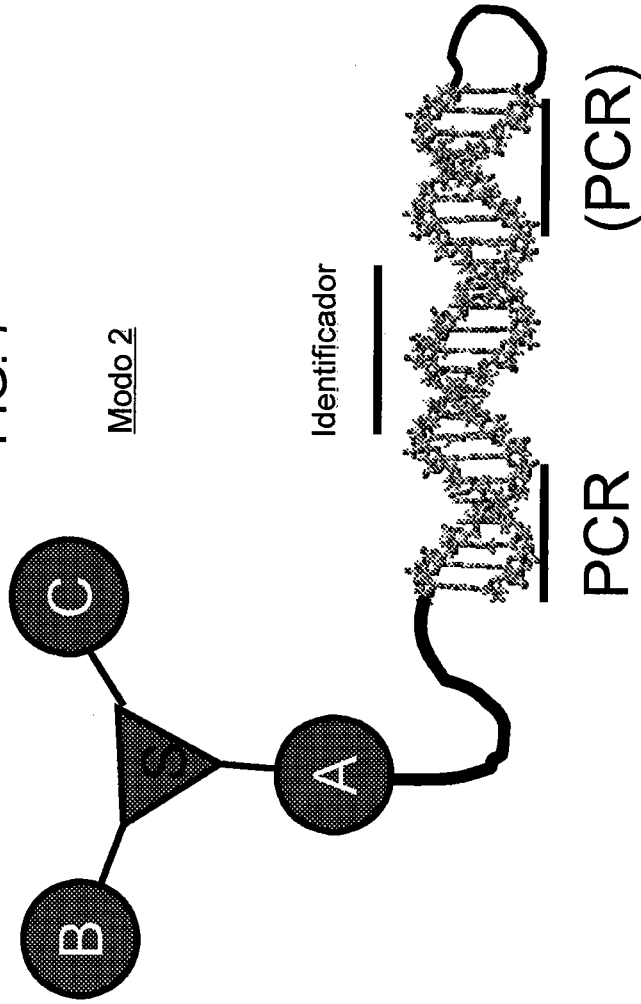
FIG. 6

Modo 1



A partir do modelo em grampo, amplificação de regiões identificadoras

FIG. 7



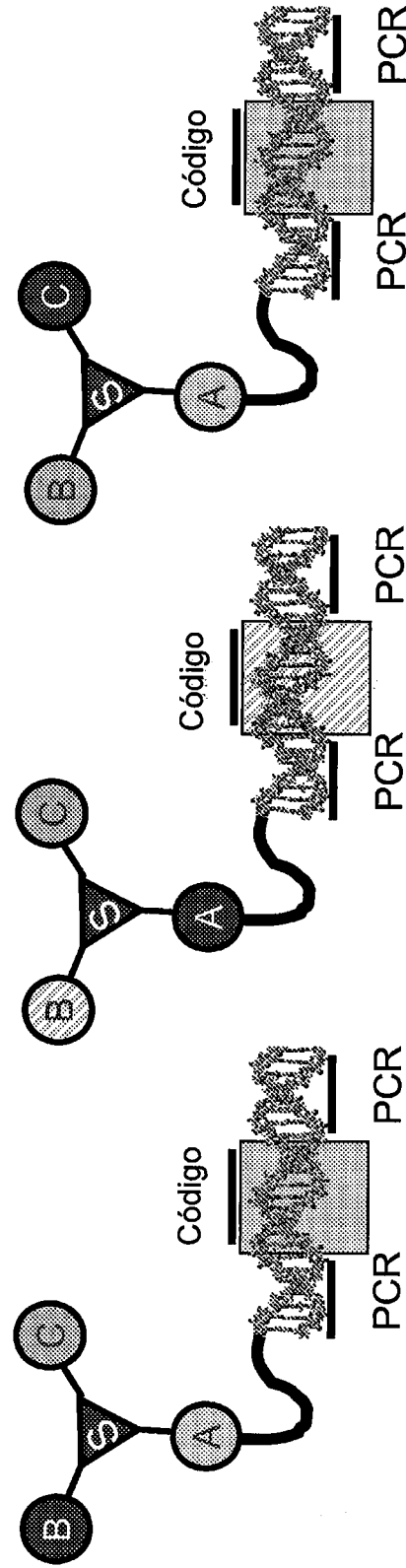
7/16

- Pequena molécula presa a ss ou dsDNA
  - DNA poderia ser em grampo em extremidade oposta
- Ligantes devem variar em tamanho e compor (tipicamente tipo PEG, 20 a 200 átomos)
- Região identificadora pode ser de qualquer tamanho, mas 5 nt – 20 nt preferidos
- Região identificadora apenas codifica A ou B em alguns esquemas, S



FIG. 8

Modo 2

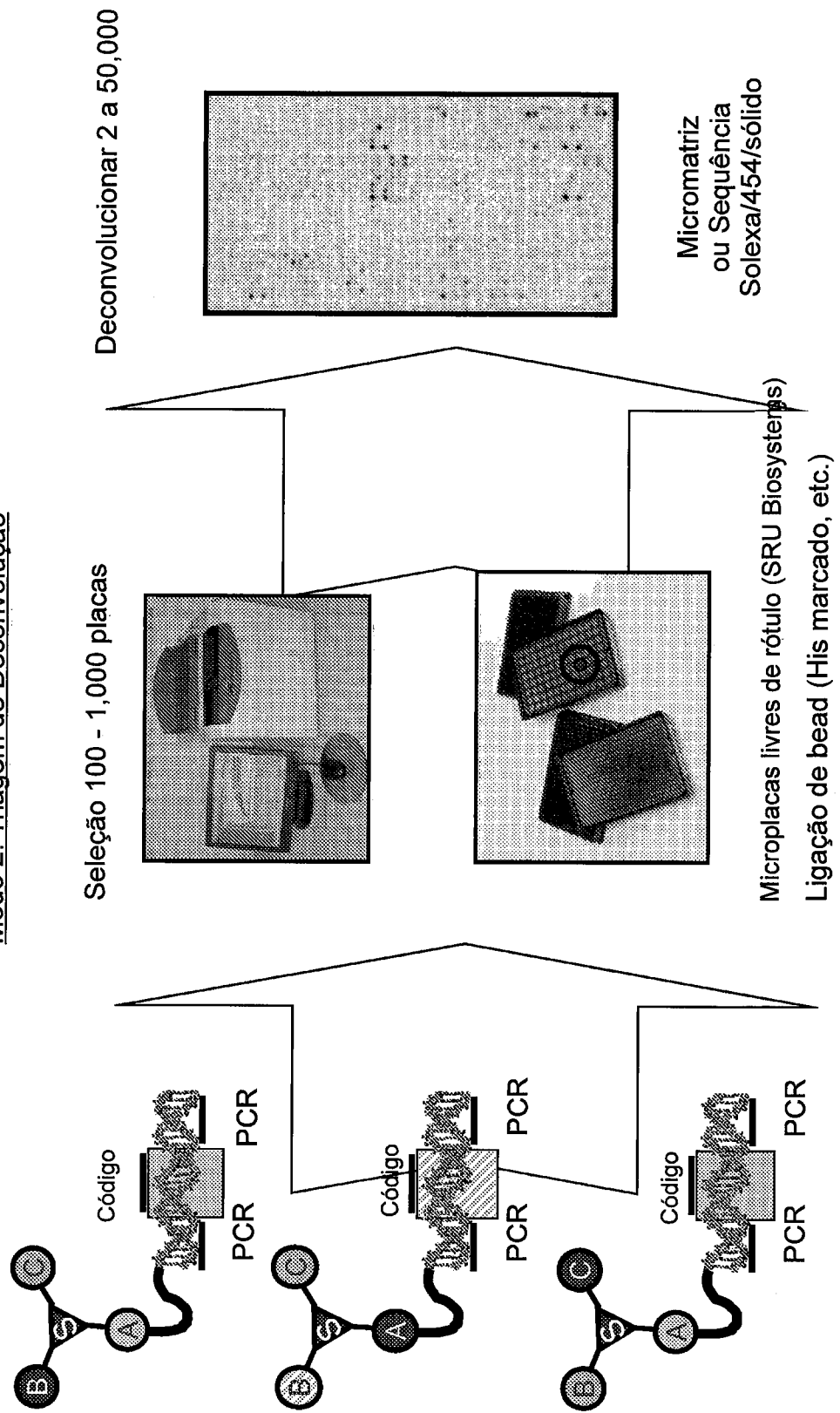


- A é marcado
- B é metila
- C é não agrupado
- Interações A x C (1000 x 1000 = 1,000,000)
- A é marcado
- B é não agrupado
- C é metila
- Interações A X B

- 3.000.000 fragmentos combinatórios examinados

FIG. 9

Modo 2: Triagem de Deconvolução



•



FIG. 11

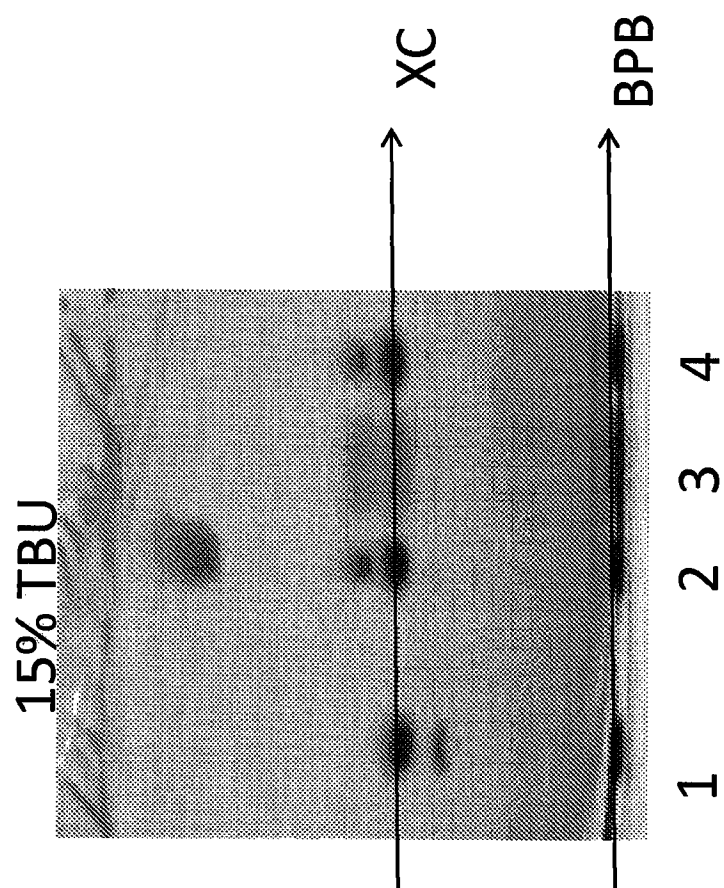


FIG. 12

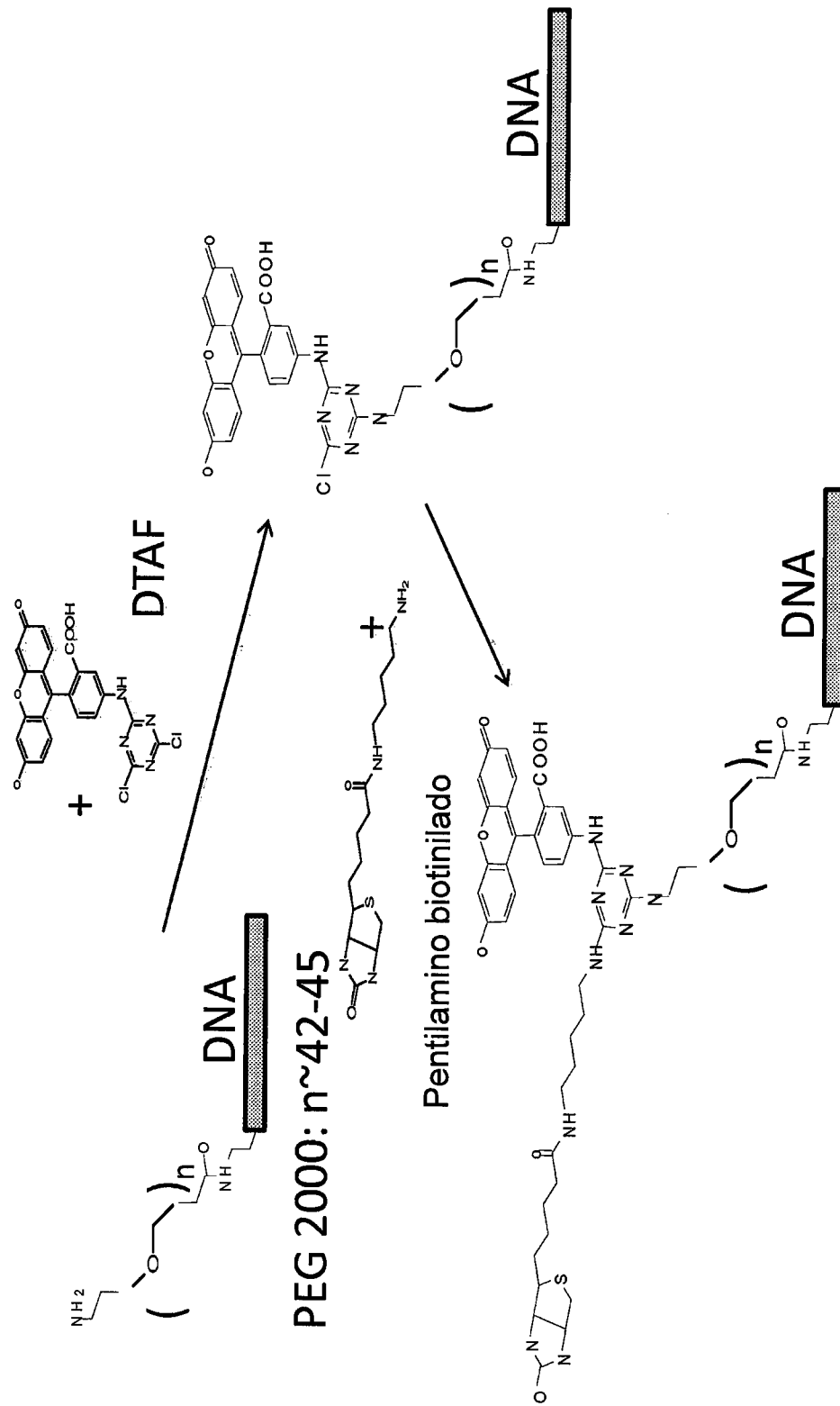
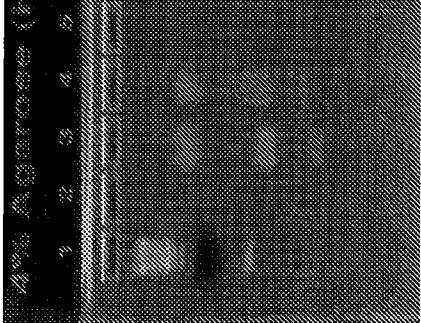


FIG. 13

A. Ligação de Marcações de DNA

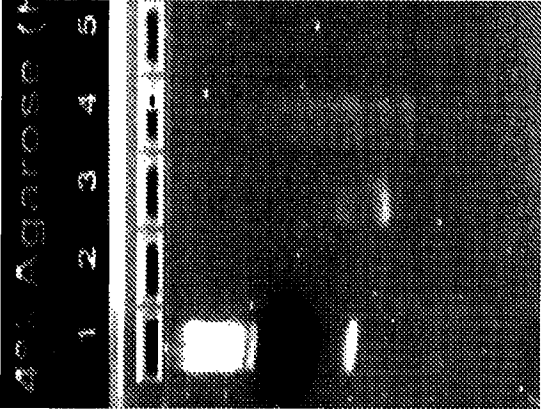
HP		A	B	C	3' end
T T	CTC TCG CCTC AGC CA GGA CC-3'	5' (p)GGAGGACT GT 3'	5' (p)CAGACGAC GA 3'	5' (p)CGATGCTC TT 3'	5'-(p) GCT GTG CAG GTA GAG TGC-3'
H2N-T	GAG AGC GGAG TCG GT CCT (p)-5'	3' GG CCTCCTGA (p) 5'	3' CA GTCTGCTG(p)-5'	3'CT GCTA CGAG (p) 5'	3'-AA CGA CAC GTC CAT CTC ACG-5'

B



M 1 2 3

C



M 1 2 3 4

FIG. 14

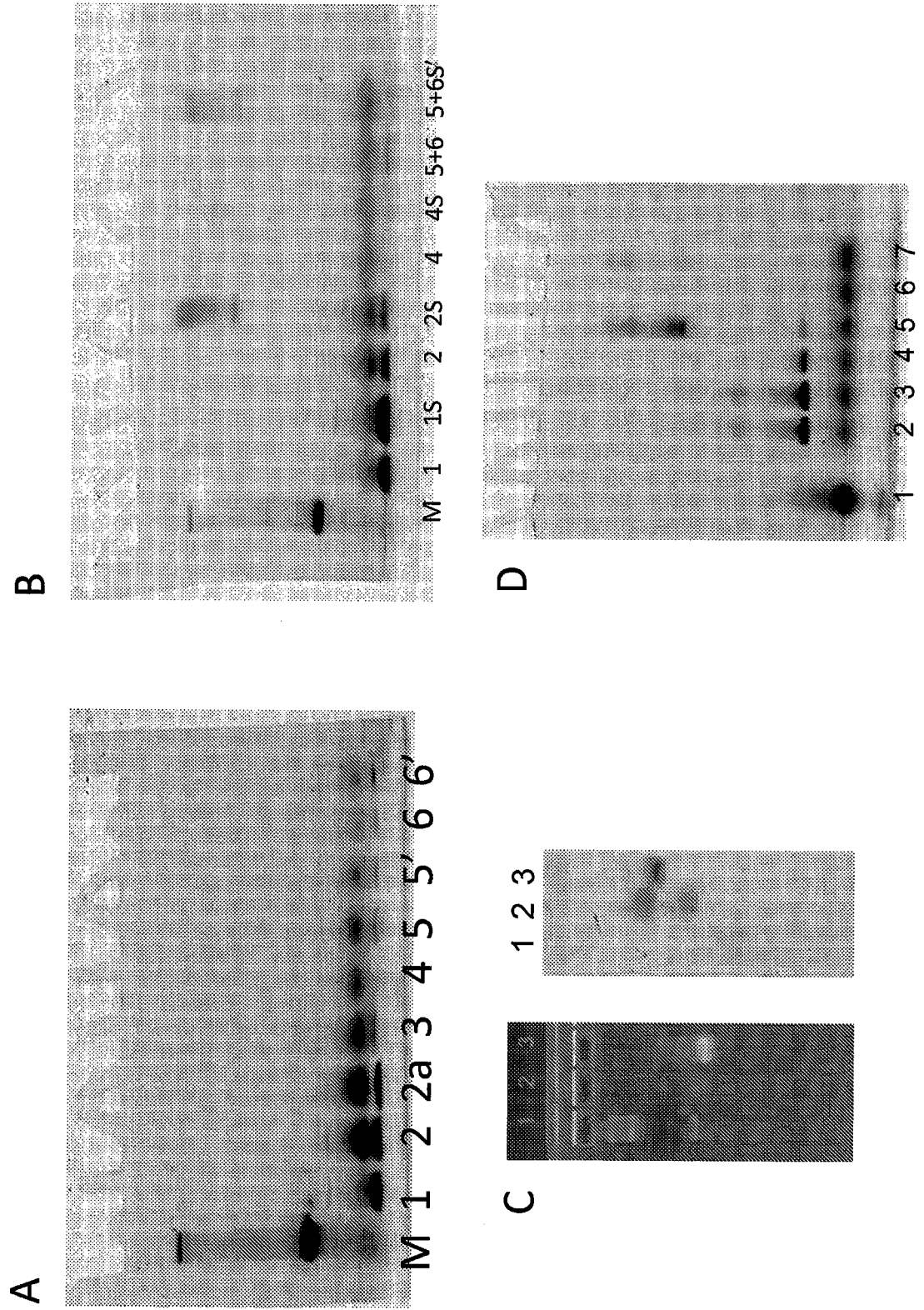


FIG. 15

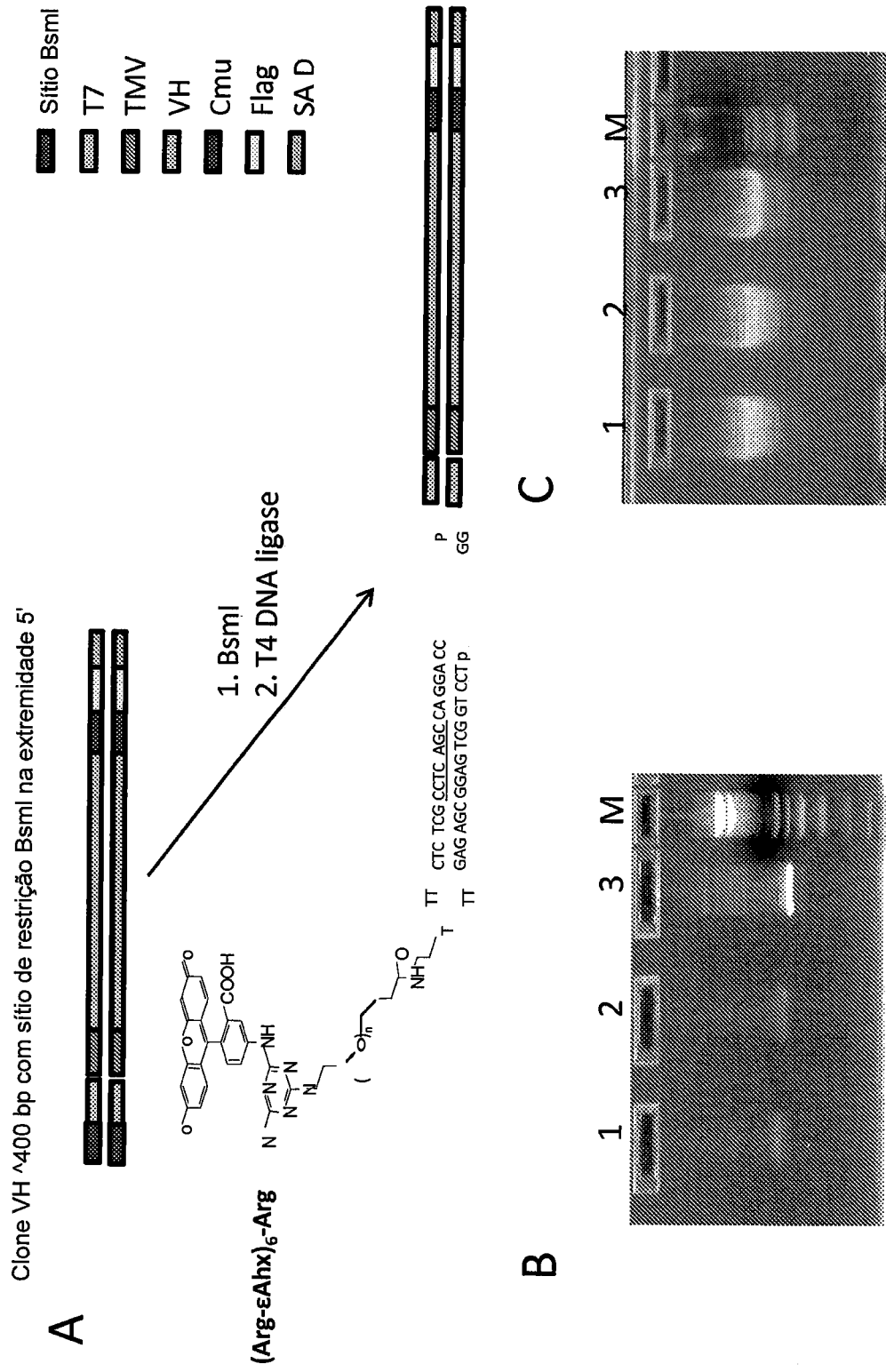
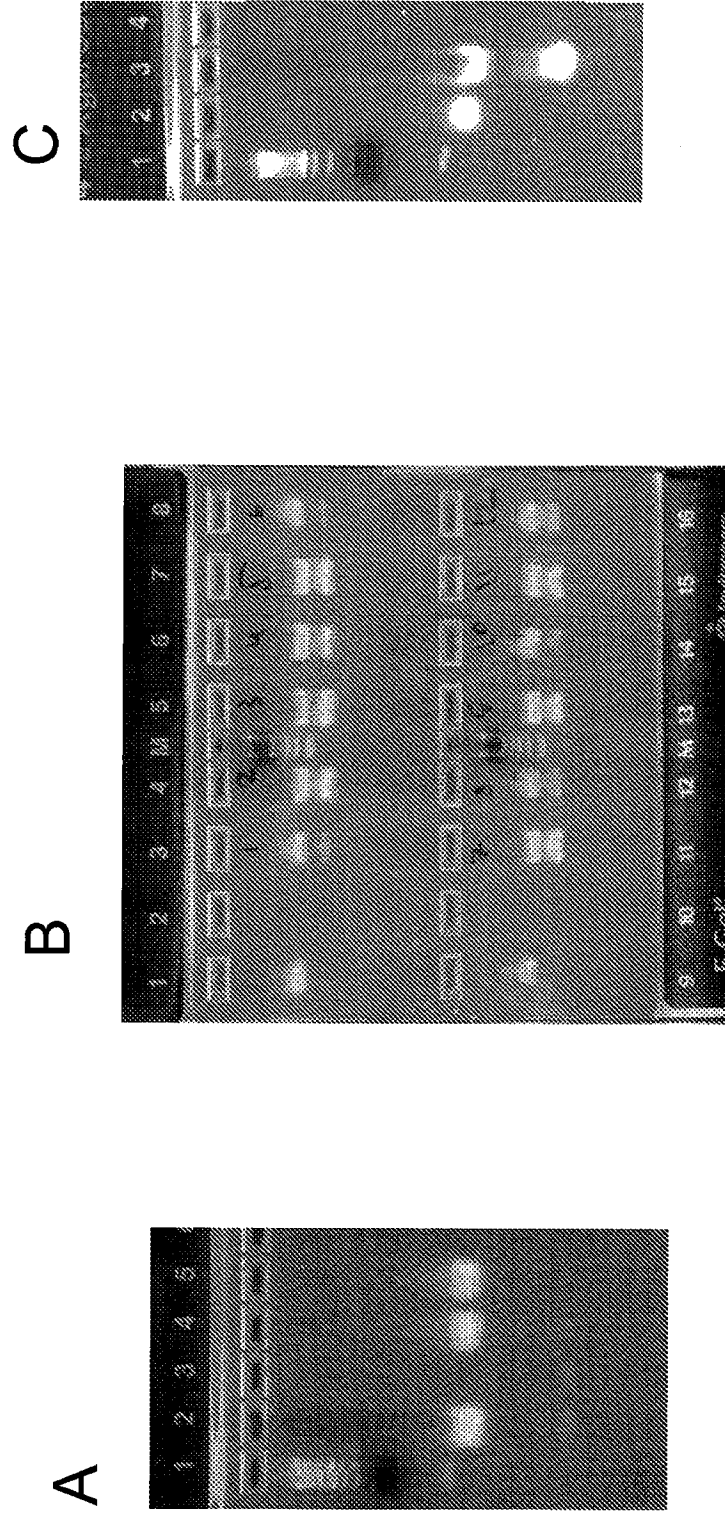




FIG. 16



**RESUMO**

Patente de Invenção: **"MÉTODOS DE CRIAÇÃO E DE TRIAGEM DE BIBLIOTECAS DE DNA CODIFICADO"**.

- 5 A presente invenção refere-se a uma série de métodos para identificar um ou mais compostos que se ligam a um alvo biológico. Os métodos incluem a síntese de uma biblioteca de compostos, em que os compostos contêm uma porção funcional tendo uma ou mais posições de diversidade. A porção funcional dos compostos é operativamente ligada a um oligonucleotídeo iniciador que identifica a estrutura da molécula funcional.

Este anexo apresenta o código de controle da listagem de sequências biológicas de que trata a Resolução INPI 228 de 11/11/2009:

**Código de Controle**

**Campo 1**



C3CDE65A5E0F2DB1

**Campo 2**



BAA16BDFDD8A57B7

**Outras Informações:**

- Nome do Arquivo: p178113-Listagem de sequencia.txt
- Data de Geração do Código: 15-08-2011
- Hora de Geração do Código: 14:06:23
- Código de Controle:
  - Campo 1: C3CDE65A5E0F2DB1
  - Campo 2: BAA16BDFDD8A57B7