



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

(21) PI 0808171-9 A2



* B R P I O 8 0 8 1 7 1 A 2 *

(22) Data de Depósito: 27/02/2008
(43) Data da Publicação: 05/08/2014
(RPI 2274)

(51) Int.Cl.:
C12N 15/10
C07K 16/00

(54) Título: MÉTODO PARA CLONAGEM DE ANTICORPOS COGNATOS

(57) Resumo:

(30) Prioridade Unionista: 05/03/2007 US 60/904,772, 01/03/2007 DK PA 2007 00316, 01/03/2007 DK PA 2007 00316, 05/03/2007 US 60/904,772

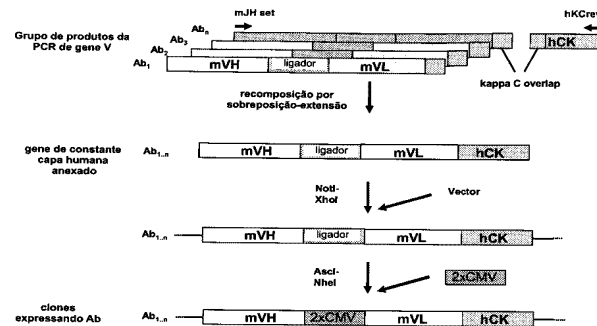
(73) Titular(es): Symphogen A/S

(72) Inventor(es): Jesper Kastrup, Lars Soegaard Nielsen, Per-Johan Meijer

(74) Procurador(es): Dannemann, Siemsen, Bigler & Ipanema Moreira

(86) Pedido Internacional: PCT DK2008050048 de 27/02/2008

(87) Publicação Internacional: WO 2008/104184 de 04/09/2008



Relatório Descritivo da Patente de Invenção para "**MÉTODO PARA CLONAGEM DE ANTICORPOS COGNATOS**".

Campo Técnico da Invenção

A presente invenção refere-se a um procedimento para ligação
5 de pares cognatos de sequências de codificação de V_H e V_L de uma população de células enriquecidas nos marcadores de antígeno de superfície específicos. O procedimento de ligação envolve um procedimento de ampliação molecular multiplexa capaz de ligar sequências de nucleotídeo de interesse, em conexão com a ampliação, especificamente, reação em cadeia de
10 polimerase (PCR multiplexa). O método é especificamente vantajoso para a geração de bibliotecas de pares cognatos, bem como, bibliotecas combinatórias de sequências de codificação de região variável de imunoglobulina. A invenção também refere-se aos métodos para geração de anticorpos humanos/não-humanos quiméricos e bibliotecas de expressão geradas por tais
15 métodos.

Antecedentes da Invenção

O WO 2005/042774 (Symphogen) descreve um método para ligação de sequências de nucleotídeo de interesse especificamente pares cognatos de sequências de codificação de V_H e V_L usando um procedimento
20 molecular multiplexo. As sequências de interesse são preferivelmente ampliadas e ligadas de células simples isoladas seguindo diluição limitada ou outras técnicas de separação de célula. A referência descreve vários modos de enriquecer um linfócito contendo população de células de modo a obter uma população de células plasmáticas que são especificamente apropriadas para
25 o procedimento de ampliação molecular multiplexa.

Sumário da Invenção

A presente invenção foca os métodos para geração de bibliotecas de sequências de codificação de imunoglobulina de animais não-humanos e métodos para geração em algumas etapas adaptadas para clonagem e classificação de alto rendimento, bibliotecas de vetores codificando
30 anticorpos quiméricos compreendendo regiões constantes e regiões variáveis não-humanas.

Em um primeiro aspecto, a invenção refere-se a um método para produzir uma biblioteca de pares cognatos compreendendo sequências de codificação de região variável ligada, dito método compreendendo:

5 a) provisão de uma fração celular compreendendo linfócito de um doador;

b) obtenção de uma população de células simples isoladas, compreendendo distribuição de células da dita fração celular individualmente em vários recipientes, onde pelo menos uma subpopulação das células expressa antígenos CD43 e CD138 ou antígenos MHCII e B220; e

10 c) ampliação e efeito da ligação das sequências de codificação de região variável contidas na dita população de célula simples isoladas por ampliação, em um procedimento de ampliação molecular multiplexa, sequências de nucleotídeo de interesse usando um gabarito derivado de uma célula simples isolada ou uma população de células isogênicas; e efetuando
15 ligação das sequências de nucleotídeo de interesse ampliadas.

Este método provê uma biblioteca de anticorpos de par cognato ou fragmentos de anticorpo.

Em outro aspecto a invenção refere-se a um método de ligação aleatória de várias sequências de nucleotídeo descontínuas de interesse, o
20 dito método compreendendo:

a) ampliação, em um procedimento de ampliação molecular multiplexa, das sequências de nucleotídeo de interesse empregando um gabarito derivado de uma população de células geneticamente diversas;

25 b) onde as células geneticamente diversas são derivadas de uma fração celular compreendendo linfócito, de um doador;

c) onde pelo menos uma subpopulação das células expressa antígenos CD43 e CD138 ou antígenos MHCII e B220 antígeno; e

d) efetuando a ligação das sequências de nucleotídeo de interesse ampliadas na etapa a).

30 Este método provê uma biblioteca combinatória de domínios variáveis de cadeia pesada e leve.

Evidência experimental provida no presente pedido estabelece

que as populações de células isoladas de esplenócitos de camundongo e sendo positivas para os antígenos de superfície listados fornecem um bom material de partida para clonagem das sequências de codificação de anticorpo usando um método de ampliação molecular multiplexa. Os métodos da invenção podem ser facilmente aplicados a outras espécies expressando os ortólogos de CD43 e CD138 ou MHCII e B220, especificamente os métodos podem ser aplicados a outro roedor, tal como rato.

O método provê várias vantagens com relação à alternativa, isto é, geração de hibridoma. Quando da preparação do hibridoma de um camundongo imunizado, as linhagens de célula estabelecidas codificarão um repertório de diferentes isotipos de anticorpo. As linhagens de células de hibridoma subsequentemente precisarão ser classificadas para ambas as funções (por exemplo, ligação a um antígeno específico, eficácia na neutralização de um agente patogênico) e para o isotipo de anticorpo. Assim, isto requer um procedimento de classificação de duas etapas para selecionar um hibridoma com um isotipo de anticorpo específico e com um efeito específico em um ensaio funcional. Finalmente, a fim de gerar uma linhagem de célula produtora, o anticorpo secretado pelo hibridoma selecionado precisará ser clonado e sequenciado antes de ser transferido para uma linhagem de célula produtora.

Utilizando o método da presente invenção, o isotipo de anticorpo é determinado pelos iniciadores empregados na ampliação molecular multiplexa e, portanto, isto não precisa ser determinado. O isotipo de anticorpo pode também ser determinado (e pode ser alterado) pela ligação subsequente ou recomposição do(s) domínio(s) constante(s) para sequências variáveis clonadas. Adicionalmente, o método da invenção provê uma biblioteca de polinucleotídeos que podem ser facilmente sequenciados e/ou inseridos nos vetores, tais como vetores de expressão, transferência, exibição ou de transporte bidirecional, de modo que uma vez que um anticorpo específico tenha sido selecionado, ele já foi clonado, sua sequência já é conhecida e ele pode ser transferido facilmente para um vetor de expressão apropriado para produção do anticorpo.

Espera-se que as células classificadas de acordo com o protocolo forneçam uma fonte de anticorpos de alta afinidade, potencialmente com afinidades na faixa picomolar. Anticorpos monoclonais do hibridoma não possuirão afinidades na faixa picomolar e precisarão ser de afinidade sinteticamente maturada para alcançar estas afinidades.

De acordo com uma concretização, estes métodos compreendem, adicionalmente, avaliação antes da amplificação molecular multiplexa de que a população de células compreendendo linfócito consista em células expressando antígenos CD43 e CD138 ou antígenos MHCII e B220, preferivelmente CD43 e CD138. Também, os métodos podem incluir enriquecimento da dita fração celular compreendendo linfócito para uma população de linfócito expressando antígenos CD43 e CD138 ou MHCII e B220 antes da ampliação molecular multiplexa.

Preferivelmente os métodos compreendem, adicionalmente, isolamento da dita população de células compreendendo linfócito expressando antígenos CD43 e CD138 ou antígenos MHCII e B220 antes da ampliação molecular multiplexa. Em uma concretização preferida as células isoladas ou subpopulação de células são CD138 alto/CD43 alto ou CD138 intermediário/CD43 alto em relação à fração celular compreendendo linfócito. Mais preferivelmente as células isoladas ou subpopulação de células são CD138 alto/CD43 alta em relação à fração celular compreendendo linfócito.

Enriquecimento ou isolamento compreende, preferivelmente, um procedimento de classificação automatizado, tal como, MACS ou FACS.

Em um aspecto adicional, a invenção refere-se a um método para geração de um vetor codificando um anticorpo quimérico com regiões constantes humanas e regiões variáveis não-humanas, dito método compreendendo:

a) provisão de uma fração celular compreendendo linfócito de um animal diferente do ser humano;

b) obtenção de uma população de células simples isoladas, compreendendo distribuição de células da dita fração celular individualmente em vários recipientes;

c) ampliação e realização da ligação da região variável codifican-
do ácidos nucleicos contidos na dita população de células simples isoladas
por ampliação, em um procedimento de ampliação molecular multiplexa, di-
tos ácidos nucleicos empregando um gabarito derivado de uma célula sim-
ples isolada ou uma população de células isogênicas; e efetuando ligação
5 dos ácidos nucleicos ampliados codificando regiões variáveis de cadeias
pesadas e leves;

d) efetuando a ligação das regiões variáveis ampliadas às regi-
ões constantes humanas; e

10 e) inserção do ácido nucleico obtido em um vetor.

Preferivelmente o animal diferente do ser humano é um camun-
dongo. À medida que os métodos da invenção são aplicados às células do
camundongo, os métodos são denominados: Symplex[®] ou mSymplex[®] de
camundongo.

15 De acordo com este aspecto da invenção é provido um novo mé-
todo para geração de bibliotecas de anticorpos humanos/diferentes dos hu-
manos quiméricos. Isto é possível por combinação da ampliação molecular
multiplexa e clonagem subsequente em uma estrutura de vetor com ligação
e/ou recomposição dos domínios constantes de cadeia pesada e leve huma-
20 na. Tradicionalmente, em um método para geração de anticorpos huma-
nos/diferentes dos humanos quiméricos, a quimerização tem sido a última
etapa após o hibridoma ter sido estabelecido e classificado e o anticorpo
codificado ter sido clonado. A quimerização pode afetar a especificidade da
ligação e/ou afinidade de um anticorpo e assim existe um risco de que um
25 bom anticorpo monoclonal de camundongo perca sua eficácia quando for
quimerizado em um anticorpo humano/camundongo.

Quando da provisão de um método que gera diretamente um an-
ticorpo de anticorpos quiméricos, a classificação pode ser realizada nos pro-
dutos que não necessitam ser modificados adicionalmente antes do desen-
volvimento pré-clínico e clínico.
30

As regiões constantes humanas podem ser providas em uma e-
tapa de ampliação molecular ou elas podem ser providas como parte de uma

estrutura de vetor, onde as regiões variáveis são clonadas seguindo-se ampliação molecular. Em uma concretização preferida, o método compreende uma etapa de ampliação adicional, onde um polinucleotídeo codificando uma cadeia leve constante humana ou seu fragmento com uma sobreposição capaz de prover ligação à cadeia leve variável, é adicionado à mistura da PCR em conjunto com um conjunto de iniciador capaz de ampliação de um construto compreendendo em ordem: uma cadeia VH de murino, um ligante, uma cadeia VL de murino e uma cadeia leve constante humana.

Em outra concretização o método compreende uma etapa de ampliação adicional, onde um polinucleotídeo codificando cadeia pesada constante humana ou seu fragmento com uma sobreposição capaz de provisão de ligação à cadeia pesada variável é adicionado à mistura da PCR em conjunto com um conjunto de iniciador capaz de ampliação de um construto compreendendo em ordem: a cadeia pesada constante humana, uma cadeia VH de murino, um ligante, e uma cadeia VL de murino. Consequentemente, também é provida uma biblioteca de vetores codificando anticorpos quiméricos, cada elemento do anticorpo consistindo em sequências de imunoglobulina diferentes daquelas dos seres humanos codificando região variável e regiões constantes de cadeia leve e pesada de imunoglobulina humana. Preferivelmente, os vetores são vetores de expressão que permitam a expressão de elementos de anticorpo da biblioteca para classificação posterior. Mais preferivelmente o vetor de expressão se destina à expressão em mamíferos.

Os vetores da biblioteca podem ser obtidos por um método da invenção.

Em uma concretização a região constante de cadeia leve é uma região constante capa. As sequências diferentes das humanas podem ser de rato, carneiro, cabra, coelho, porquinho-da-Índia ou outros animais apropriados, para os quais os protocolos de imunização foram descritos, para os quais as informações de sequência estão disponíveis para permitir o projeto de iniciadores apropriados e para os quais técnicas de classificação de célula apropriadas permitem classificação de células plasmáticas para ampliação

molecular multiplexa de célula única, de modo a ligar pares cognatos de sequências de região variável. Em uma concretização, as sequências diferentes das humanas são de origem primata não-humana, tais como, macaco caranguejeiro, macaco Resus, chimpanzé ou símios. Preferivelmente as sequências diferentes das humanas são de roedores, tais como, murino ou rato. Em outra concretização as sequências diferentes das humanas são sequências de coelhos.

Preferivelmente as regiões variáveis dos anticorpos são pares cognatos.

Em outro aspecto a invenção refere-se à sub-biblioteca que codifica anticorpos exibindo especificidades de ligação desejadas, dirigidas contra um alvo específico, selecionadas de uma biblioteca de acordo com a invenção.

Descrição das Figuras

Figura 1: Murino - PCR com mSymplex[®]: Sobreposição-extensão multiplexa RT-PCR para a ampliação e ligação cognata de genes de anticorpo de cadeia pesada e leve de uma célula única. Misturas de iniciadores exemplares usadas para RT-PCR e PCR aninhada são descritas em detalhes nas Tabelas 2 e 3 (ou Tabela 5).

Figura 2: Clonagem de repertório de murino: um grupo de produtos da PCR mSymplex[®] codificando pares de gene VH/VL de células plasmáticas únicas foi recomposto no gene codificando a cadeia leve constante capa humana por recomposição por sobreposição-extensão. O grupo de genes, codificando anticorpos quiméricos de humanos-camundongos completos, foi inserido em um vetor de expressão seguido por uma inserção de um cassete promotor bidirecional (2xCMV). As misturas de iniciador usadas para a recomposição de capa humana são descritas em detalhes na Tabela 6.

Figura 3A: Classificação de esplenócitos de camundongo. (A) Para isolamento das células plasmáticas definidas como CD43 alto, CD138 alto, PI (iodeto de propídio) positivo ou células mortas foram excluídos no painel esquerdo inferior (não P1). Então as células plasmáticas retidas como CD43 alto, CD138 alto no painel direito inferior (P2). Finalmente, dupletos

foram excluídos no gráfico do painel direito superior SSC-H, SSC-W (P3). As células positivas para todas as três retenções foram classificadas em placas ELISPOT. (B) Para isolamento dos blastos plasmáticos definidos como intermediário MHCII, intermediário B220, PI (iodeto de propídio) positivo ou células mortas foram excluídos no painel esquerdo inferior (não P1). Então, os blastos plasmáticos foram excluídos como MHCII intermediário, B220 intermediário, painel direito inferior (P2). Dupletos foram excluídos no gráfico do painel direito superior SSC-H, SSC-V (P3) e finalmente as células foram retidas por tamanho no painel esquerdo superior (P4). Células positivas para todas as quatro retenções foram classificadas nas placas ELISPOT.

Figura 4: Classificação dos esplenócitos de camundongo. Primeiramente, células PI positivo ou mortas foram excluídas no painel esquerdo inferior (P1). No gráfico de pontos à direita, parte superior, são ilustrados CD138 PE e CD43 FITC. Quatro retenções foram ajustadas nas diferentes populações de célula fenotípicas: P2 é CD138 intermediário, CD43 alto, P3 é CD138 alto, CD43 alto, P4 é CD138 alto, CD43 negativo, P5 é CD138 intermediário, CD43 baixo. 10.000 células positivas para P1 e cada uma das quatro retenções foi classificada nos tubos de teste e congelada para avaliação por symplex.

Figura 5: Eletroforese de gel para produtos da PCR a partir de titulação da PCR Symplex de lisado celular a partir das 4 frações classificadas, P2, P3, P4 e P5 são as retenções classificadas de acordo com a figura 4. M representa marcadores de peso molecular.

Figura 6: representação esquemática de um vetor de expressão de anticorpo e comprimento pleno de mamífero 00-VP-002. Gene de resistência à ampicilina, Amp e Amp pro e seu promotor; origem pUC, origem pUC de replicação; CMV, promotor de mamífero acionando a expressão da cadeia leve e da cadeia pesada; Líder IGHV, líder da cadeia pesada humana genômica; enchimento H, inserção que é trocada pela sequência de codificação de região variável de cadeia pesada; IGHG1, sequência de codificação de região constante de cadeia pesada G1 e isotipo de imunoglobulina genômica (sequência mostrada no Anexo 1); B-globina A de coelho; se-

quência beta-globina poliA de coelho; líder IGKV, líder capa murino; enchimento L, inserção que é trocada pela sequência de codificação de cadeia leve; termo SV40, sequência terminadora 40 do vírus símio; FRT, sítio alvo de reconhecimento Flp; gene resistente à neomicina, Neo, SV40 poliA, sequência de sinal 40 poli A do vírus de símio.

Figura 7: Análise de um repertório de anticorpos quiméricos anti-hEGFR. Análise do grupo de diferença de absorvência em 450-620 nm. Os sobrenadantes são agrupados por reatividade conforme indicado pelo número (1 a 4) seguindo o número do clone. Cinza escuro indica uma diminuição no número de células metabolicamente ativas, considerando-se que cinza claro indica um aumento no número de células metabolicamente ativas. O preto indica sobrenadantes sem efeito no número de células metabolicamente ativas.

Descrição da Invenção

A presente invenção provê possibilidades adicionais para emprego do método de ampliação e ligação descrito no WO2005/042774 para provisão de coleções de vetores de anticorpo de animais não-humanos. Estes aperfeiçoamentos permitem a clonagem de sequências de codificação de anticorpo quimérico humano/não-humano com pares cognatos de regiões variáveis a serem ajustados a um formato de alto rendimento. Isto é basicamente obtido por aperfeiçoamento de um novo material de partida para processos de ampliação e ligação e por provisão de métodos para geração de bibliotecas de anticorpos humanos/não-humanos quiméricos com pares cognatos de regiões variáveis.

Um aspecto da invenção é um método de ligação de sequências variáveis de cadeia pesada e leve compreendendo ampliação, em um procedimento de ampliação molecular multiplexa, as sequências de nucleotídeo relevantes empregando um gabarito derivado de uma célula simples isolada, uma população de células isogênicas ou uma população de células geneticamente diversas e realização de uma ligação subsequente das sequências ampliadas.

Definições

O termo "par cognato" descreve um par original de ácidos nucleicos descontínuos de interesse que estão contidos dentro ou são derivados de uma célula simples. Nas concretizações preferidas, um par cognato compreende duas sequências de codificação de região variável que em conjunto
5 codificam um domínio variável de proteína de ligação, as sequências gênicas destas sendo derivadas da mesma célula. Assim, quando expressas tanto como uma proteína de ligação completa quanto com um seu fragmento estável, elas preservam a afinidade de ligação e especificidade da proteína de ligação originalmente expressa desta célula. Um par cognato pode ser
10 compreendido, por exemplo, de uma sequência de codificação de cadeia pesada variável de anticorpo associada a uma sequência de codificação de cadeia leve variável da mesma célula, ou uma sequência de codificação de cadeia α de receptor de célula α associada a uma sequência de codificação de cadeia β da mesma célula. Uma biblioteca de pares cognatos é uma coleção de tais pares cognatos.
15

O termo "polimerase de partida aquecida" descreve polimerases que são inativas ou possuem atividade muito lenta em temperaturas usadas para transcrição reversa. Tais polimerases precisam ser ativadas por temperaturas altas (90 a 95°C) para se tornarem funcionais. Isto é, por exemplo,
20 uma vantagem em procedimentos de RT-PCR de etapa única, uma vez que proíbe a interferência da polimerase com a reação de transcriptase reversa.

O termo "população de células isogênica" descreve uma população de células geneticamente idêntica. Especificamente, uma população de células isogênica derivada por expansão clonal de uma célula simples isolada é de interesse na presente invenção.
25

O termo "célula simples isolada" descreve uma célula que foi fisicamente separada de uma população de células correspondendo a uma "célula simples em um recipiente único". Quando da distribuição de uma população de células individualmente entre vários recipientes, uma população de células simples isoladas é obtida. Conforme especificado na seção intitulada
30 "*Fontes de Gabarito*" a proporção dos recipientes com uma célula simples não é necessariamente 100%, a fim de denominar a mesma uma população

de células simples.

Os termos derivados de "ligar" ou "ligação" em relação à ampliação descrevem a associação das sequências de ácido nucleico ampliadas codificando as sequências de ácido nucleico de interesse em um segmento
5 simples. Em relação aos pares cognatos um segmento compreende sequências de ácido nucleico codificando um domínio variável, por exemplo, uma região variável de cadeia pesada de anticorpo associada a uma sequência de codificação de região variável de cadeia leve de anticorpo derivada da mesma célula. A ligação pode ser tanto ativada simultaneamente com a ampliação ou como uma etapa imediata seguindo-se a ampliação. Não existem
10 requisitos para a forma ou funcionalidade do segmento, podendo ser linear, circular, de filamento simples ou filamento duplo. A ligação não é necessariamente permanente, uma das sequências de ácido nucleico de interesse pode ser isolada do segmento, caso desejado, uma da sequência de codificação da região variável pode ser, por exemplo, isolada de um segmento de
15 par cognato. Contudo, à medida que as regiões variáveis originais constituindo o par cognato não são misturadas desordenadamente com outras regiões variáveis, elas ainda são consideradas um par cognato, embora não ligadas em conjunto a um segmento simples. A ligação é considerada um par cognato, embora não ligada em conjunto a um segmento simples. A ligação é preferivelmente uma ligação de fosfodiéster de nucleotídeo. Contudo, a
20 ligação pode também ser obtida por procedimentos de ligação cruzada química diferentes.

O termo "ampliação molecular multiplexa" descreve a ampliação
25 simultânea de duas ou mais seqüências-alvo na mesma reação. Métodos de ampliação apropriados incluem a reação em cadeia de polimerase (PCR) (U.S. 4.683.202), reação da cadeia de ligase (LCR), (Wu and Wallace, 1989, Genomics 4, 560-9), técnica de ampliação de deslocamento do filamento (SDA) (Walker e outros,, 1992, Nucl. Acids Res. 20, 1691-6), replicação de
30 sequência autossustentada (Guatelli e outros, 1990, Proc. Nat. Acad. Sci. U.S.A., 87, 1874-8) e ampliação de sequência com base em ácido nucleico (NASBA) (Compton J., 1991, Nature 350, 91-2). Os últimos dois métodos de

ampliação envolvem reações isotérmicas com base na transcrição isotérmica, os quais produzem RNA de filamento simples (RNAss) e DNA de filamento duplo (DNAds).

5 O termo "PCR multiplexa" descreve uma variante da PCR onde duas ou mais sequências-alvo são ampliadas simultaneamente, por inclusão de mais de um conjunto de iniciador na mesma reação, por exemplo, um conjunto de iniciador adaptado para ampliação da região variável de cadeia pesada e um conjunto de iniciador adaptado para ampliação da região variável de cadeia leve na mesma reação da PCR. Adicional ou alternativamente, um conjunto de iniciador adaptado para ampliação da região variável de cadeia lambda pode ser combinado com estes conjuntos de iniciadores.

O termo "RT-PCR multiplexa" descreve uma reação da PCR multiplexa, que é precedida por uma etapa de transcrição reversa (RT). A RT-PCR multiplexa pode tanto ser realizada como um processo de duas etapas com uma etapa RT separada antes da PCR multiplexa, como um processo de etapa simples onde todos os componentes para ambas RT e PCR multiplexa são combinados em um tubo simples.

Os termos "PCR de extensão-sobreposição multiplexa" e "RT-PCR de extensão-sobreposição multiplexa" implicam no fato de que a PCR multiplexa ou RT-PCR multiplexa é realizada utilizando uma mistura de iniciador de extensão-sobreposição multiplexa para ampliar as sequências-alvo, pelo que, permitindo ampliação simultânea e ligação das sequências-alvo.

O termo "vários recipientes" descreve qualquer objeto (ou coleção de objetos) que permitem a separação física de uma célula simples de uma população de células. Estes podem ser tubos, placas de múltiplos poços (por exemplo, placas de 96 poços, 384 poços, placas de microtitulação ou outras placas de múltiplos poços), fileiras, microfileiras, microchips, géis ou uma matriz de gel. Preferivelmente o objeto é aplicável à ampliação da PCR.

30 O termo "proteína policlonal" ou "policlonalidade" conforme usado neste documento refere-se a uma composição de proteína compreendendo moléculas de proteína diferentes, porém homólogas, preferivelmente sele-

cionadas da superfamília da imunoglobulina. Assim, cada molécula de proteína é homóloga às outras moléculas da composição, porém também contém um ou mais estiramentos de sequência de polipeptídeo variável, que é/são caracterizado(s) por diferenças na sequência de aminoácido entre os elementos individuais da proteína policlonal. Exemplos conhecidos de tais proteínas policlonais incluem moléculas de anticorpo ou imunoglobulina, receptores de célula T e receptores de célula B. Uma proteína policlonal pode consistir em um subconjunto definido de moléculas de proteína, que foi definido por um aspecto comum, tal como, atividade de ligação compartilhada na direção de um alvo desejado, por exemplo, um anticorpo monoclonal exibindo especificidade de ligação na direção de um antígeno-alvo desejado.

O termo "uma população de células geneticamente diversas" conforme usado no presente documento refere-se a uma população de células onde as células individuais na população diferem uma da outra no nível genômico. Tal população de célula geneticamente diversa é, por exemplo, uma população de células derivadas de um doador ou uma fração de tais células, por exemplo, um linfócito B ou um linfócito T contendo uma fração de célula.

O termo "conjunto de iniciador" é empregado intercambiavelmente com o termo "par iniciador" e descreve dois ou mais iniciadores que em conjunto são capazes de iniciar a ampliação de uma sequência de nucleotídeo de interesse (isto é, um elemento de um par cognato). Um conjunto de iniciador da presente invenção pode ser projetado para iniciar uma família de sequências de nucleotídeo contendo sequências de codificação de região variável. Exemplos de famílias diferentes são cadeias leve capa de anticorpo, cadeias leve lambda, regiões variáveis de cadeia pesada, e regiões variáveis de receptor de célula T α , β , γ ou δ . Um conjunto de iniciador para a ampliação de uma família de sequências de nucleotídeo contendo sequências de codificação de região variável frequentemente constitui vários iniciadores onde os vários iniciadores podem ser iniciadores degenerados.

O termo "identidade de sequência" é expresso como uma porcentagem que indica o grau de identidade entre as sequências de ácido nucleico

com relação ao comprimento da mais curta das duas sequências. Pode ser calculada como $(N_{ref} - N_{dif}) \square 100\% / N_{ref}$, onde N_{ref} é o número de resíduos na mais curta das sequências e onde N_{dif} é o número total de resíduos não idênticos em um emparelhamento alinhado otimamente longo N_{ref} entre as duas sequências. Conseqüentemente, a sequência de DNA AGTCAGTC (Seq. NO. 32) terá uma identidade de sequência de 75% com a sequência TAATCAATCGG (Seq. NO. 33) ($N_{dif}=2$ e $N_{ref}=8$) (o sublinhamento mostra o alinhamento ótimo, e o negrito indica os dois resíduos não-idênticos dos 8).

Os termos "aleatoriamente" ou "aleatória" com relação à ligação refere-se à ligação de sequências de nucleotídeo que não são derivadas da mesma célula, porém são ligadas transversalmente entre uma população de células geneticamente diversas. Se as sequências de nucleotídeo de interesse forem sequências de codificação de região variável, isto resultará em uma biblioteca combinatória de sequências ligadas. Se, por outro lado, as sequências de nucleotídeo de interesse codificarem uma proteína heteromérica não diversa, as sequências ligadas aleatoriamente parecerão similares às sequências ligadas de uma célula simples.

O termo "gabarito derivado de uma célula simples isolada" com relação à transcrição reversa, refere-se aos ácidos nucleicos dentro de tal célula isolada. Os ácidos nucleicos podem estar, por exemplo, na forma de RNA, RNAm, DNA ou DNA genômico. Os ácidos nucleicos podem tanto ser isolados da célula como ainda estarem com o conteúdo remanescente da célula, onde a célula está em uma forma intacta ou uma forma lisada.

O termo "CD43" refere-se a um antígeno de superfície de camundongo conhecido sob inúmeros sinônimos incluindo antígeno 3E8, A630014B01Rik, antígeno LP-3 de diferenciação de célula B, Cd43, antígeno CD43, Galgp, sialoglicoproteína de leucócito, precursor de leucosialina, Ly48, Ly-48, Sialoforina, bem como marcadores de superfície ortólogos de outros animais.

O termo "CD138" refere-se a um antígeno de superfície de camundongo conhecido por inúmeros sinônimos incluindo Syndecan-1, A-A408134, AA409076, CD138, syn-1, Synd, Synd1, SYND1, Synd-1, precur-

sor de Syndecan-1, bem como marcadores de superfície ortólogos de outros animais.

O termo "MHCII" refere-se a um antígeno de superfície de camundongo conhecido por inúmeros sinônimos incluindo antígeno CD74, CLIP, DHLAG, cadeia gama de antígeno de histocompatibilidade de H-2 classe II, HLADG, HLA-DR-GAMMA, cadeia não variante associada ao antígeno Ia, Ia-GAMMA, Ii, cadeia não variante associada à classe II MHC, bem como marcadores de superfície ortólogos de outros animais.

O termo "B220" refere-se a um antígeno de superfície de camundongo conhecido por inúmeros sinônimos incluindo B220, Cd45, CD45, antígeno CD45, CD45R, L-CA, precursor de antígeno comum ao leucócito, loc, Ly-5, antígeno Ly-5 comum ao linfócito, Lyt-4, T200, bem como aos marcadores de superfície ortólogos de outros animais.

Descrição Detalhada da Invenção

15 Processo de ampliação e ligação

Um aspecto da presente invenção reduz o número de tubos necessários para ampliar as sequências de nucleotídeo de interesse, empregando uma variante de PCR onde duas ou mais sequências-alvo são ampliadas simultaneamente no mesmo tubo por inclusão de um ou mais conjuntos de iniciadores, por exemplo, todos os iniciadores necessários para ampliar sequências de codificação de região variável, na mesma reação. De modo geral, esta abordagem é conhecida como reação da cadeia de polimerase multiplexa (PCR multiplexa).

Um aspecto adicional da presente invenção é que duas ou mais sequências-alvo ampliadas por PCR multiplexa são ligadas em proximidade ao processo de ampliação. Especificamente, pares cognatos de sequências de codificação de região variável são ligados por este processo.

Uma concretização da presente invenção descreve que uma mistura de iniciador multiplexa pode ser projetada para trabalhar em um procedimento de PCR de sobreposição-extensão, resultando em uma ampliação simultânea e ligação das sequências de nucleotídeo de interesse. Esta técnica de PCR de sobreposição-extensão multiplexa serve para reduzir o nú-

mero de reações necessárias para isolar e ligar sequências de nucleotídeo de interesse, especificamente pares cognatos de regiões variáveis ligadas.

Outras concretizações da presente invenção aplicam ligação por união ou por recombinação como uma alternativa à ligação por PCR de sobreposição-extensão multiplexa. Nestes procedimentos, a ligação não é realizada simultaneamente com a ampliação da PCR multiplexa, porém como uma etapa imediata seguindo-se a ampliação. Contudo, a ligação ainda pode ser realizada no mesmo tubo como onde a PCR multiplexa foi realizada.

A PCR de sobreposição-extensão multiplexa requer a presença de dois ou mais conjuntos de iniciadores (uma mistura de iniciador multiplexa), onde pelo menos um iniciador de cada conjunto é equipado com uma cauda de sobreposição-extensão. As caudas de sobreposição-extensão permitem a ligação dos produtos gerados em cada um dos conjuntos de iniciadores durante a ampliação. Tal mistura de iniciador é denominada uma mistura de iniciador de sobreposição-extensão multiplexa. A PCR de sobreposição-extensão multiplexa, difere da PCR de sobreposição-extensão convencional pelo que as sequências a serem ligadas são geradas simultaneamente no mesmo tubo, pelo que, provendo ligação imediata das sequências-alvo durante ampliação, sem qualquer purificação intermediária.

Um aspecto adicional da presente invenção é uma etapa de transcrição reversa (RT) precedendo a PCR multiplexa ou ampliação da PCR de sobreposição-extensão multiplexa, utilizando um gabarito derivado de uma célula simples isolada ou uma população de células isogênicas.

Um aspecto adicional da presente invenção é o emprego de sequências de nucleotídeo derivadas de uma célula simples isolada ou uma população de células isogênicas como gabarito para a ampliação da PCR multiplexa. Preferivelmente, RNA de uma célula simples sofre transcrição reversa no DNAC antes da PCR multiplexa. Para a ampliação de algumas sequências de ácido nucleico de interesse, DNA genômico pode ser usado como alternativa ao RNAm. Utilizando-se células simples isoladas ou uma população de células isogênicas derivadas da expansão clonal de uma célula simples isolada como fonte de gabarito, é possível evitar mistura desorde-

nada das sequências de nucleotídeo codificando uma proteína heteromérica de interesse, com sequências de nucleotídeo derivadas de diferentes células dentro de uma população de células. Isto é importante caso seja desejada obtenção da composição original de sequências de interesse. Especialmente para a geração de um par cognato de sequências de codificação de região variável, o emprego de uma célula simples isolada ou uma população de células isogênicas como fonte de gabarito é um aspecto importante.

Adicionalmente, a presente invenção facilita a geração de bibliotecas de sequências de ácido nucleico de interesse ligadas, especificamente bibliotecas combinatórias e bibliotecas de pares cognatos de regiões variáveis. Adicionalmente, a presente invenção utiliza ácidos nucleicos derivados de células simples, preferivelmente na forma de RNA que não precisa ser isolado do teor da célula remanescente antes de ser utilizado como gabarito.

Uma concretização da presente invenção engloba a ligação de várias sequências de nucleotídeo de interesse descontínuas. O método compreende ampliação, em um procedimento de PCR multiplexa ou ampliação de RT-PCR multiplexa, sequências de nucleotídeo de interesse empregando um gabarito derivado de uma célula simples isolada ou uma população de células isogênicas e efetuando ligação das sequências de nucleotídeo ampliadas de interesse. Adicionalmente, o método compreende uma etapa opcional de realização de ampliação adicional dos produtos ligados.

Uma concretização adicional da presente invenção engloba um método de produção de uma biblioteca de pares cognatos compreendendo sequências ligadas de codificação de região variável. O método compreende provisão de uma fração celular contendo linfócito de um doador, que é opcionalmente enriquecido por uma população de linfócito específica da dita fração celular, ou onde uma população de linfócito específica foi isolada da dita fração celular. Adicionalmente, uma população de células simples isoladas é obtida por distribuição das células da fração celular contendo linfócito, ou da fração celular enriquecida, individualmente entre vários recipientes. A ampliação molecular multiplexa (ampliação de RT-PCR multiplexa) das sequências de codificação de região variável contidas na população de células

simples isoladas é realizada e a ligação dos pares de sequências de codificação de região variável é efetuada, onde um par individual de sequências de região variável é derivado de uma célula simples, dentro da população da célula simples isolada. Adicionalmente, a técnica compreende duas etapas
5 opcionais: na primeira etapa a célula simples isolada, individual na população das células simples é expandida para uma população de células isogênicas antes da realização da ampliação de RT-PCR multiplexa. Desta forma, são obtidos vários recipientes com população diversa de células isogênicas (uma população das células isogênicas em um recipiente). A segunda etapa
10 opcional engloba realização de uma ampliação adicional das sequências ligadas de codificação de região variável.

Nas concretizações preferidas da presente invenção, um elemento individual da dita biblioteca de pares cognatos compreendido de uma sequência de codificação de região variável de cadeia leve de imunoglobulina
15 é associado a uma sequência de codificação de região variável de cadeia pesada de imunoglobulina, originando-se da mesma célula ou das sequências de codificação de um domínio de ligação do receptor de célula T, constituído de uma região variável de cadeia alfa associada à região variável de cadeia-beta ou uma região variável de cadeia gama associada à região variável de cadeia delta, onde as regiões variáveis associadas se originam da
20 mesma célula.

A ampliação de RT-PCR multiplexa da presente invenção pode ser realizada tanto como um processo de duas etapas, onde a transcrição reversa (RT) é realizada separada da ampliação da PCR multiplexa (ou ampliação molecular multiplexa alternativa), ou como um processo de etapa
25 única, onde as etapas de RT e ampliação da PCR multiplexa são realizadas com os mesmos iniciadores em um tubo.

A transcrição reversa (RT) é realizada com uma enzima contendo atividade de transcriptase reversa resultando na geração de DNAc do RNA total, RNAm ou RNA específico-alvo de uma célula simples isolada. Iniciadores
30 que podem ser utilizados para a transcrição reversa são, por exemplo, iniciadores oligo-dT, hexâmeros aleatórios, decâmeros aleatórios ou outros

iniciadores aleatórios ou iniciadores que são específicos para as sequências de nucleotídeo de interesse.

O procedimento de ampliação de RT-PCR multiplexa de duas etapas permite que o DNAc gerado na etapa de RT, seja distribuído a mais de um recipiente, permitindo armazenamento de uma fração do gabarito antes do procedimento com a ampliação. Adicionalmente, a distribuição do DNAc a mais de um tubo permite o desempenho de mais de uma ampliação de PCR multiplexa de ácido nucleico derivada do mesmo gabarito. Embora isto resulte em um número aumentado de reações separadas, abre também a possibilidade de aumentar a complexidade da mistura de iniciador multiplexa se isto for desejado. Esta abordagem de duas etapas pode ser aplicada, por exemplo, para ampliar e ligar região variável de cadeia pesada e sequências de codificação de região variável de cadeia leve capa em um tubo e região variável de cadeia pesada e sequências de codificação de região variável de cadeia leve lambda em um tubo diferente empregando o mesmo gabarito. Uma célula simples geralmente expressa apenas uma das cadeias leves. Contudo, será frequentemente mais fácil realizar as reações simultaneamente ao invés de esperar o resultado de uma das reações antes de realizar a outra. Adicionalmente, a ampliação de ambas a capa e lambda serve como um controle interno negativo, uma vez que seria esperado que apenas capa e lambda ampliem de uma célula simples.

O procedimento de única etapa de RT-PCR multiplexa, transcrição reversa e ampliação da PCR multiplexa é realizado no mesmo recipiente. Todos os componentes necessários para realizar ambas a transcrição reversa e a PCR multiplexa em uma única etapa são inicialmente adicionados aos recipientes e a reação é realizada. De modo geral, não é necessário acrescentar componentes uma vez que a reação tenha começado. A vantagem de uma ampliação de RT-PCR multiplexa de etapa única é que ela reduz o número de etapas necessárias para gerar as sequências de nucleotídeo ligadas da presente invenção. Isto é especificamente útil quando da realização da RT-PCR multiplexa em uma fileira de células simples, quando a mesma reação precisa ser realizada em vários recipientes. A RT-PCR multi-

plexa de etapa única é realizada por emprego de iniciadores reversos presentes na mistura de iniciador multiplexa necessária para a ampliação da PCR multiplexa como iniciadores para a transcrição reversa. De modo geral, a composição necessária para a RT-PCR multiplexa de etapa única compreende um gabarito de ácido nucleico, uma enzima com atividade de transcriptase reversa, uma enzima com atividade de DNA polimerase, mistura de trifosfato de desoxinucleosídeo (mistura de dNTP compreendendo dATP, dCTP, dGTP e dTTP) e mistura de um iniciador multiplexo. O gabarito de ácido nucleico é preferivelmente RNA total ou RNAm derivado de uma célula simples isolada tanto na forma purificada como um lisado da célula ou ainda dentro da célula intacta. De modo geral, a composição exata da mistura de reação requer alguma otimização de cada mistura de iniciador multiplexa a ser usada com a presente invenção. Isto se aplica para os procedimentos de RT-PCT multiplexa de etapa única e duas etapas.

Para algumas reações de RT-PCR multiplexa de etapa única pode ser vantajoso acrescentar componentes adicionais durante a reação. Por exemplo, a adição de polimerase seguindo-se a etapa de RT. Outros componentes seriam, por exemplo, uma mistura de dNTP ou uma mistura de iniciador multiplexa possivelmente com uma composição de iniciador diferente. Isto pode então ser considerado como uma RT-PCR multiplexa de um tubo, que geralmente possui as mesmas vantagens que uma RT-PCR multiplexa de uma etapa, uma vez que também limita o número de tubos necessários para obtenção dos produtos ligados desejados.

As sequências de nucleotídeo de interesse, ampliadas pela RT-PCR multiplexa podem ser ligadas uma à outra por vários métodos, tais como, RT-PCR de sobreposição-extensão multiplexa, ligação ou recombinação, empregando diferentes misturas de iniciador multiplexas. Preferivelmente a ampliação de RT-PCR multiplexa e processo de ligação é um processo de etapa única ou de duas etapas. Contudo, um processo de ligação também pode ser realizado como um processo de múltiplas etapas, empregando, por exemplo, um fragmento de enchimento para unir as sequências de ácido nucleico de interesse, tanto com PCR, ligação ou recombinação. Tal

fragmento de enchimento pode conter elementos cis, elementos promotores ou uma sequência de codificação relevante ou sequência de reconhecimento. Em uma concretização preferida um processo de ligação é realizado no mesmo recipiente que o da ampliação de RT-PCR multiplexa.

5 Em uma concretização da presente invenção a ligação de várias sequências de nucleotídeo de interesse descontínuas é realizada em associação com a ampliação da PCR multiplexa, utilizando a mistura de iniciador de sobreposição-extensão multiplexa. Isto resulta na ampliação e ligação combinadas das sequências-alvo. De modo geral, a composição necessária
10 para a PCR de sobreposição-extensão multiplexa compreende, um gabarito de ácido nucleico, uma enzima com atividade de DNA polimerase, mistura de trifosfato desoxinucleosídeo (mistura de dNTP compreendendo dATP, dCTP, dGTP e dTTP) e uma mistura de iniciador de sobreposição-extensão multiplexa.

15 Em uma concretização específica da presente invenção, a ligação das várias sequências de nucleotídeo de interesse descontínuas é realizada por RT-PCR de sobreposição-extensão multiplexa empregando um gabarito derivado de uma célula simples isolada ou uma população de células isogênicas. Adicionalmente, o método compreende uma etapa opcional
20 de realização de uma ampliação molecular adicional dos produtos ligados. Preferivelmente, a RT-PCR de sobreposição-extensão multiplexa é realizada como uma reação de etapa única/um tubo.

 A mistura de iniciador de sobreposição-extensão multiplexa da presente invenção compreende pelo menos dois conjuntos de iniciadores
25 capazes de iniciar a ampliação e ligação de pelo menos duas sequências de codificação de região variável, por exemplo, ampliação e ligação das sequências de famílias de região variável de cadeia pesada de imunoglobulina com famílias de região variável de cadeia leve capa ou lambda, ou ampliação e ligação das sequências de famílias de receptores de células T α , β , γ ,
30 ou δ .

 Em outra concretização da presente invenção as várias sequências de nucleotídeo de interesse, ampliadas por RT-PCR multiplexa, são li-

gadas por união. De modo a obter isto, a mistura de iniciador multiplexa usada para a RT-PCR multiplexa é projetada, tal que, as sequências-alvo ampliadas podem ser clivadas com enzimas de restrição apropriadas e ligação covalente por união de DNA pode ser realizada (o projeto do iniciador é descrito na seção "*Misturas de Iniciador e Projeto*"). Seguindo-se a RT-PCR multiplexa com tal mistura de iniciador multiplexa, as enzimas de restrição necessárias para formar terminações compatíveis das sequências-alvo são adicionadas à mistura em conjunto com a ligase. Nenhuma purificação dos produtos da PCR é necessária antes desta etapa, embora a purificação possa ser realizada. A temperatura de reação para a clivagem de restrição combinada e ligação está aproximadamente entre 0 e 40°C. Contudo, se a polimerase da reação PCR multiplexa ainda estiver presente na mistura, uma temperatura de incubação abaixo da temperatura ambiente é preferida, mais preferivelmente temperaturas entre 4 e 16°C.

Ainda em outra concretização da presente invenção, as várias sequências de nucleotídeo de interesse, ampliadas por RT-PCR multiplexa, são unidas por recombinação. Nesta abordagem, as sequências-alvo ampliadas podem ser unidas empregando sítios de recombinação idênticos. A ligação é então realizada por adição de recombinases que facilitam a recombinação. Alguns sistemas de recombinase apropriados são Flp recombinase com vários sítios FRT, Cre recombinase com vários sítios lox, integrase Φ C31 que realiza a recombinação entre o sítio attB, o sistema β -recombinase-six, bem como o sistema Gin-gix. A ligação por recombinação foi exemplificada para duas sequências de nucleotídeo (V_H ligado com V_L) (Chapal, N. e outros, 1997 *BioTechniques* 23, 518-524).

Em uma concretização preferida da presente invenção, as sequências de nucleotídeo de interesse compreendem sequências de codificação de região variável e a ligação gera um par cognato de sequências de codificação de região variável. Tal par cognato pode compreender uma ou mais sequências de codificação de região constante além das regiões variáveis. Preferivelmente as regiões constantes são de origem humana e o par cognato da região variável é de origem diferente, tal como, camundongo,

rato ou coelho.

Em uma concretização mesmo mais preferida da presente invenção, as sequências de nucleotídeo de interesse compreendem região variável codificando sequências de imunoglobulina e a ligação gera um par cognato da região variável de cadeia leve e sequências de codificação de região variável de cadeia pesada. Tal par cognato pode compreender uma ou mais sequências de codificação de região constante além das regiões variáveis. Adicionalmente, tal par cognato pode ser isolado de gabarito derivado de células da linhagem de linfócito B enriquecidas da fração celular contendo linfócito, tais como sangue integral, células mononucleares ou leucócitos.

Em outra concretização da presente invenção, as sequências de nucleotídeo de interesse compreendem sequências TcR codificando região variável e a ligação gera um par cognato de sequências de codificação de região variável de cadeia α e região variável de cadeia β ou sequências de codificação de região variável de cadeia γ e região variável de cadeia δ . Tal par cognato pode compreender uma ou mais sequências de codificação de região constante além das regiões variáveis. Adicionalmente, tal par cognato pode ser isolado de gabarito derivado de células da linhagem de linfócito T enriquecidas da fração celular contendo linfócito, tais como sangue integral, células mononucleares ou leucócitos.

Outro aspecto da presente invenção se constitui na utilização da RT-PCR multiplexa com uma população de células geneticamente diversas como fonte de gabarito. A maior parte das sequências de codificação da proteína heteromérica não varia de célula para célula como é o caso com as sequências de codificação de região variável das proteínas de ligação. Assim, quando utilizando a presente invenção para a clonagem de sequências de codificação de proteína heteromérica não-variável, não há necessidade de se realizar um isolamento inicial das células simples.

Nesta concretização da presente invenção, várias sequências de nucleotídeo de interesse descontínuas são ligadas aleatoriamente por um método compreendendo, realização de ampliação de RT-PCR multiplexa de sequências de nucleotídeo de interesse empregando um gabarito derivado

de uma população de células geneticamente diversas e efetuando ligação das sequências de nucleotídeo ampliadas de interesse. Adicionalmente, o método compreende uma etapa opcional de realização de uma ampliação adicional dos produtos ligados. Como com a abordagem da célula simples, a
5 ligação pode ser realizada tanto utilizando a mistura de iniciador de sobreposição-extensão multiplexa para a ampliação ou alternativamente por ligação ou recombinação. Preferivelmente o gabarito derivado da população de células não está estritamente contido dentro das células. A população de células pode ser, por exemplo, lisada.

10 A aplicação do processo de ligação aleatória em uma população de células expressando proteínas de ligação variantes, possibilita a geração simplificada de bibliotecas combinatórias de sequências de codificação de região variável. Preferivelmente, a população de células constitui células que expressam proteínas de ligação de região variáveis, tais como, linfócitos B,
15 linfócitos T, células de hibridoma, células plasmáticas, plasmablastos, ou uma mistura destas células.

A população de células na concretização mencionada acima pode, por exemplo, ser permeabilizada ou lisada, sem purificação adicional ou os ácidos nucleicos de gabarito podem ser isolados das células por procedimentos-padrão. O procedimento de RT-PCR multiplexa de etapa simples é
20 preferido. Contudo, o procedimento de duas etapas também pode ser usado na concretização.

Um modo eficiente de aumentar a especificidade, sensibilidade e rendimento do processo de ligação de RT-PCR multiplexa é por realização
25 de uma ampliação molecular adicional das sequências de nucleotídeo ligadas obtidas da RT-PCR multiplexa seguida por ligação por união ou recombinação ou ligação empregando a RT-PCR de sobreposição-extensão multiplexa. Esta ampliação adicional é realizada preferivelmente com ampliação da PCR, utilizando uma mistura de iniciador adaptada para ampliação das
30 sequências de ácido nucleico de interesse ligadas. A mistura de iniciador utilizada pode se constituir em iniciadores externos da mistura de iniciador multiplexa ou mistura de iniciador de sobreposição-extensão multiplexa, sig-

nificando os iniciadores que recombinaem com a extremidade 5' e 3' mais extremas do filamento de senso das sequências ligadas de codificação de região variável, pelo que, permitindo a ampliação de todo produto ligado. Os iniciadores externos também podem ser descritos como os iniciadores da
5 mistura de iniciador de sobreposição-extensão multiplexa que não contém caudas de sobreposição-extensão. Alternativamente, um conjunto de iniciador aninhado ou semianinhado pode ser usado para a ampliação adicional das sequências de nucleotídeo ligadas. Tal PCR aninhada serve especialmente para aumentar a especificidade do método, bem como para aumentar
10 a quantidade do produto ligado. Para a presente invenção, PCR semianinhada (conforme descrita na seção intitulada *Misturas de Iniciador e Projeto*) é considerada funcionando bem como a PCR aninhada. Assim, é desejado, embora não necessário que a presente invenção realize uma ampliação adicional da PCR dos produtos ligados a partir da RT-PCR de sobreposição-
15 extensão multiplexa ou dos produtos ligados por união ou recombinação, preferivelmente empregando PCR aninhada ou PCR semianinhada.

A ampliação adicional pode tanto ser realizada diretamente usando uma fração quanto todo o produto de reação da RT-PCR de sobreposição-extensão multiplexa ou produto de ligação ou produto de recombinação ou uma fração de qualquer um destes produtos ou empregando produtos ligados parcialmente purificados de qualquer uma destas reações, por exemplo, por realização de uma eletroforese de gel agarose dos produtos ligados, e excisando o fragmento correspondendo ao tamanho esperado das sequências de codificação da região variável ligada. Para os produtos ligados por RT-PCR de sobreposição-extensão multiplexa, a ampliação adicional é preferivelmente realizada diretamente em uma fração a partir da reação de RT-PCR de sobreposição-extensão multiplexa, uma vez que isto ajudaria na ligação das sequências-alvo individuais que não foram ligadas na primeira reação.

30 **Sequências de Interesse**

As sequências de nucleotídeo de interesse da presente invenção podem ser selecionadas de sequências que codificam diferentes subunida-

des ou domínios, as quais, quando expressas, formam uma proteína ou parte de uma proteína. Tais proteínas que são compostas de pelo menos duas subunidades não idênticas são conhecidas como proteínas heteroméricas. As proteínas heteroméricas são comuns em todos os tipos de espécies. Algumas das classes as quais tais proteínas pertencem são, por exemplo, enzimas, inibidores, proteínas estruturais, toxinas, proteínas de canal, proteínas G, proteínas de receptor, proteínas da superfamília das imunoglobulinas, proteínas de transporte etc. As sequências de nucleotídeo codificando tais proteínas heteroméricas são descontínuas, significando, por exemplo, que elas se originam de genes diferentes, ou moléculas de RNA diferentes. Contudo, descontínuas conforme empregado na presente invenção também pode significar sequências de nucleotídeo codificando domínios da mesma proteína, onde os domínios são separados por sequências de nucleotídeo que não são de interesse.

Em uma concretização da presente invenção as sequências de nucleotídeos de interesse contêm sequências de codificação de região variável da superfamília da imunoglobulina, tais como, imunoglobulinas (anticorpos), receptores de célula B e receptores de célula T (TcR's). Especialmente, as sequências de codificação de região variável das imunoglobulinas são de interesse. Tais sequências de codificação de região variável compreendem anticorpos de comprimento pleno, bem como, Fab's, Fv's, scFv's e combinações de fragmentos das sequências de codificação de região variável, por exemplo, regiões determinantes de complementaridade (CDR's), genes de união ou genes V ou suas combinações. De modo geral a presente invenção pode ser aplicada a quaisquer combinações de sequências de codificação de região variável e seus fragmentos. A presente aplicação permite a ligação apenas dos domínios variáveis das cadeias pesada e leve gerando sequências de codificação de Fv ou scFv. Ou a ligação de toda cadeia leve com a região variável de cadeia pesada + domínio de região constante C_{H1} + partes da região de articulação, gerando Fab, Fab' ou $F(ab)_2$. Adicionalmente é possível acrescentar qualquer região dos domínios de região constante de cadeia pesada à cadeia pesada variável, pelo que, gerando sequências de

codificação de anticorpo de comprimento pleno ou sequências de codificação de anticorpo truncado. Em um aspecto da invenção, as sequências variáveis não-humanas são ligadas às regiões constantes humanas para gerar anticorpos humanos/não-humanos completamente quiméricos, preferivelmente anticorpos quiméricos com regiões humanas constantes.

Em uma concretização adicional da presente invenção, sequências de codificação de região variável compreendem um tipo de sequência de codificação de cadeia leve de imunoglobulina (capa ou lambda) e uma sequência de codificação de região variável de cadeia pesada de imunoglobulina. Isto é obtido por seleção de iniciadores ampliando apenas um isotipo de cadeia leve e pesada. O isotipo pode também ser determinado por ligação ou recomposição de regiões constantes humanas de um ou mais isotipos de cadeia pesada e leve.

Sequências de codificação de região variável derivadas de receptores de célula T (TcR's) são também de interesse. Tais sequências de codificação de TcR compreendem sequências de codificação de cadeias alfa e beta ou gama e delta de comprimento pleno, bem como TcR's solúveis ou apenas domínios variáveis destas cadeias ou proteínas de fusão de cadeia simples das mesmas (por exemplo, cadeia $\alpha\beta$ simples ou cadeia simples $\gamma\delta$).

Fontes de Gabarito

Um aspecto da presente invenção é a capacidade de ligar sequências de nucleotídeo derivadas de uma célula simples isolada, uma população de células isogênicas, ou uma população de células geneticamente diversa que não foi separada em recipientes simples.

Um aspecto preferido da presente invenção é o emprego de células simples isoladas ou uma população de células isogênicas como fonte de gabarito, uma vez que a mistura desordenada das sequências de ácido nucleico de interesse, especificamente sequências de codificação de região variável é evitada. Isto é importante, se é desejado um par original, por exemplo, de sequências de codificação de região variável.

Outro aspecto preferido da presente invenção é a obtenção de

uma célula simples ou população das células simples a partir da fração celular compreendendo linfócitos, tais como, linfócitos B, linfócitos T, as células plasmáticas e/ou vários estágios de desenvolvimento destas linhagens de células. Outras populações de células que expressam proteínas de ligação da superfamília das imunoglobulinas podem também ser usadas, de modo a obter células simples. Linhagens de células, tais como, células de hibridoma, linhagens de células de linfócito B ou linfócito T ou linhagens de célula imortalizadas de vírus ou células derivadas de doador participando na resposta imune são também aplicáveis na presente invenção. Frações celulares contendo linfócito derivadas de doador podem também ser obtidas de tecido natural ou fluido que é rico em tais células, por exemplo, sangue, medula óssea, nodos linfáticos, tecido do baço, tecido das amídalas ou de infiltrações em e ao redor de tumores ou infiltrações de tecido inflamatório. Preferivelmente, no caso de animais não-humanos, o tecido do baço ou medula óssea é usado. Doadores podem tanto ser não desafiados ou hiperimunes com relação ao alvo desejado. Para o isolamento de proteínas de ligação de antígeno com especificidades de ligação na direção de um alvo desejado, os doadores hiperimunes são preferidos. Tais doadores hiperimunes podem tanto ser doadores imunizados com o alvo quanto fragmentos do alvo ou podem ser pacientes convalescentes, ou indivíduos não-saudáveis que estão operando uma resposta imune natural na direção do alvo, por exemplo, pacientes autoimunes, pacientes com câncer, pacientes com doenças infecciosas, por exemplo, pacientes com HIV, pacientes com hepatite A, B ou C, pacientes SARS, etc. ou pacientes com doenças crônicas. Contudo, em uma concretização especificamente preferida, o doador é um animal não-humano que foi imunizado com um antígeno próprio humano, tal como, proteína humana implicada em câncer, tal como, EGFR.

Para uso na presente invenção, doadores de células podem ser da mesma espécie a ser tratada com os produtos obtidos das sequências de nucleotídeo ligadas da presente invenção. Preferivelmente, um doador de célula é um animal doméstico, de estimação, o homem. Um aspecto especial da presente invenção é que ela permite a geração de bibliotecas de anticor-

po humano/não-humano quiméricas para uso em terapia humana. Tal abordagem é preferida quando os anticorpos são direcionados contra os assim chamados antígenos próprios, isto é, antígenos humanos.

O doador pode também ser animal transgênico, especificamente
5 um camundongo transgênico. Animais transgênicos transportando locais de imunoglobulina humana são descritos na Patente US número 6.111.166 e Kuroiwa, Y. e outros, *Nature Biotechnology*, 2002; 20: 889-893. Tais animais transgênicos são capazes de produzir imunoglobulinas humanas. Assim, anticorpos humanos completos contra um alvo específico podem surgir por
10 técnicas de imunização comuns de tais animais transgênicos. Isto permite a geração de bibliotecas codificando proteínas de ligação com especificidades na direção de alvos mais difíceis, tais como, antígenos humanos para os quais não existe ou existe resposta limitada de anticorpo humano natural. Tais animais transgênicos podem da mesma forma ser desenvolvidos para
15 produzir receptores de célula T humana.

Em uma concretização adicional da presente invenção, a fração celular contendo linfócito é constituída de sangue integral, medula óssea, células mononucleares ou leucócitos obtidos de um doador. Células mononucleares podem ser isoladas de sangue, medula óssea, nodos linfáticos, baço, infiltrações ao redor das células cancerígenas e infiltrações inflamatórias. Células mononucleares podem ser isoladas por técnicas de centrifugação de densidade, por exemplo, gradientes de Ficoll. Se as células mononucleares forem isoladas das amostras compostas de tecido, o tecido será desintegrado antes da centrifugação gradiente ser realizada. A desintegração
20 pode ser realizada, por exemplo, por métodos mecânicos, tais como, moagem, eletroporação e/ou por métodos químicos, tais como, tratamentos enzimáticos. O isolamento dos leucócitos pode ser realizado diretamente de doadores empregando leucoforese. Preparações brutas, por exemplo, de medula óssea ou tecido, que contém linfócitos, também podem ser usadas
25 na presente invenção. Tais preparações precisarão ser desintegradas, por exemplo, conforme descrito acima, a fim de facilitar a distribuição da célula simples.

Um aspecto adicional da presente invenção é o enriquecimento da fração celular contendo linfócito, por exemplo, sangue integral, células mononucleares, leucócitos ou medula óssea, com relação a uma população de linfócito específica, tais como células da linhagem de linfócito B ou linfócito T. O enriquecimento dos linfócitos B pode ser realizado, por exemplo, empregando classificação de célula em microesfera magnética (MACS) ou classificação de célula ativada em fluorescência (FACS) tirando vantagem das proteínas marcadoras de superfície de célula específica de linhagem, tais como, CD19 ou outros marcadores específicos de linhagem de célula B, tais como, B220. O enriquecimento de linfócitos T pode ser realizado, por exemplo, utilizando um marcador de superfície celular, tal como, CD3 ou marcadores específicos de linhagem de célula T.

Um aspecto preferido da presente invenção é a classificação dos linfócitos B enriquecidos adicionalmente, a fim de adquirir as células plasmáticas, antes da distribuição das células individualmente entre vários recipientes. O isolamento das células plasmáticas é realizado, de modo geral, por classificação MACS ou FACS, utilizando marcadores de superfície, tais como CD19. Outros marcadores de superfície específicos de célula plasmática ou suas combinações podem ser utilizados, por exemplo, CD138, CD43, CD19, MHC-II, a escolha exata do marcador dependendo da fonte de célula plasmática, por exemplo, braço, amídalas, sangue ou medula óssea. Naturalmente, a escolha exata dos marcadores de superfície também depende das espécies, a partir das quais as células são isoladas.

Em um aspecto da invenção, os marcadores usados para classificação e/ou seleção de células são CD43 e CD138 ou MHCII e B220 ou ortólogos. Preferivelmente, a combinação dos marcadores é CD43 e CD138 e preferivelmente as células selecionadas possuem uma expressão intermediária ou alta destes marcadores em relação à população de células compreendendo linfócitos da qual eles são selecionados ou isolados. Mais preferivelmente, o nível de expressão de CD43 e CD138 é relativamente alto em relação à população de células compreendendo linfócitos da qual eles são selecionados ou isolados.

As células plasmáticas também podem ser obtidas de uma população de células compreendendo linfócitos obtidos de qualquer uma destas fontes. As células plasmáticas isoladas do sangue são algumas vezes denominadas como células plasmáticas ou plasmablastos. Na presente invenção, estas células são também denominadas células plasmáticas. As células plasmáticas são desejadas para isolamento dos pares cognatos de sequências de codificação de imunoglobulina, uma vez que uma frequência mais alta destas células produz anticorpos específicos de antígeno que refletem a imunidade adquirida na direção do antígeno desejado e a maior parte das células sofreu hipermutação somática e, portanto, codificam anticorpos de alta afinidade. Adicionalmente, os níveis de RNAm nas células plasmáticas são elevados em comparação à população de linfócito B remanescente, assim o procedimento de transcrição reversa é mais eficiente quando usando células plasmáticas simples. Como uma alternativa ao isolamento das células plasmáticas, as células de memória B podem ser isoladas da fração de célula contendo linfócito utilizando um marcador de superfície de célula, tal como, CD27 e IgG.

Um aspecto alternativo da presente invenção é a seleção de linfócitos B enriquecidos para especificidade do antígeno antes da distribuição das células entre os vários recipientes. Isolamento dos linfócitos B específicos de antígeno é realizado por contato dos linfócitos B enriquecidos com o antígeno desejado ou antígenos permitindo ligação do antígeno à superfície exposta à imunoglobulina, seguido por isolamento dos ligantes. Isto pode ser feito, por exemplo, por biotinação do antígeno desejado ou antígenos seguido por técnicas de classificação de célula apropriadas. As células plasmáticas bem como linfócitos B, células mononucleares não enriquecidas, leucócitos, sangue integral, medula óssea ou preparações de tecido podem ser submetidas ao isolamento com relação à especificidade de antígeno se isto for desejado.

Outro aspecto da presente invenção é classificar linfócitos T enriquecidos (por exemplo, células CD3 positivas) empregando marcadores de superfície, tais como, por exemplo, CD27, de modo a obter uma fração de

células T de memória. Os linfócitos T podem também ser selecionados de especificidade de antígeno-MHC usando complexos de peptídeo-MHC (por exemplo, Callan, M.F. e outros, 1998. *J. Exp. Med.* 187, 1395-1402; Novak, E.J. e outros, 1999. *J. Clin. Invest* 104, R63-R67).

5 Como uma alternativa à classificação das células expressando determinados marcadores de superfície, isto é, uma seleção positiva, é concebível que células NÃO expressando os marcadores sejam esgotadas da composição de células, deixando as células atrás daquelas que realmente expressam os marcadores.

10 Um aspecto adicional da presente invenção é a imortalização de qualquer uma das frações de célula isolada descritas acima (por exemplo, linfócitos B, as células plasmáticas, células de memória ou linfócitos T). A imortalização pode ser realizada, por exemplo, com vírus Epstein-Barr (Traggiai, E., e outros, 2004. *Nat Med* 10, 871-875) antes da distribuição ce-
15 lular. Alternativamente, células simples isoladas podem ser imortalizadas e expandidas antes da transcrição reversa. Traggiai e outros, *Nat Med.* Agosto de 2004; 10(8):871-5.

 Um aspecto adicional da presente invenção é a distribuição de uma população de células desejadas (por exemplo, células de hibridoma,
20 linhagens de células de linfócito B ou linhagem de linfócito T, células de sangue integral, células de medula óssea, células mononucleares, leucócitos, linfócitos B, células plasmáticas, linfócitos B específicos de antígeno, células B de memória, linfócitos T, linfócitos T específicos de peptídeo/MHC ou células T de memória) individualmente, em vários recipientes, a fim de se obter
25 uma população de células simples isoladas. Este isolamento de células simples refere-se à separação física das células de uma população de células, de tal modo que um recipiente simples contenha uma célula simples ou uma microfileira, chip ou matriz de gel sendo carregada em um modo que produ-
30 za células simples. As células podem ser distribuídas diretamente aos vários recipientes, tais como, fileiras de recipientes simples por limitação da diluição. Os recipientes simples utilizados na presente invenção são preferivelmente aqueles aplicáveis na PCR (por exemplo, tubos de PCR e placas de

PCR com 96 poços ou 384 poços ou fileiras maiores de recipientes). Contudo, outros recipientes também podem ser usados. Quando da distribuição das células simples em um grande número de recipientes simples (por exemplo, placas com 384 poços), uma população de células simples é obtida.

5 Tal distribuição pode ser realizada, por exemplo, por dispensa de um volume em um recipiente simples que na média engloba uma concentração de célula de um, 0,5 ou 0,3 célula, desta forma obtendo recipientes que em média contém uma célula simples ou menos. Uma vez que a distribuição das células por limitação da diluição é um evento estatístico, uma fração de recipientes será cheia, uma fração maior conterá uma célula simples e uma fração

10 menor conterá duas ou mais células. Quando duas ou mais células estiverem presentes em um recipiente, alguma mistura desordenada das sequências de codificação de a região variável poderá ocorrer entre as células presentes no recipiente. Contudo, uma vez que este é um evento menor, ele

15 não afetará a utilidade total da presente invenção. Adicionalmente, combinações de sequências de codificação de região variável que não possuem a afinidade e especificidade de ligação desejada mais provavelmente não serão selecionadas e conseqüentemente serão eliminadas durante o processo de classificação. Portanto, eventos menores de mistura desordenada não afeta-

20 ram significativamente a biblioteca final da presente invenção.

Existem alternativas para distribuição celular por limitação da diluição usando, por exemplo, classificadores de células, tais como, máquinas de FACS ou robôs, que podem ser programados para dispensar acuradamente células simples dentro de recipientes simples. Estas alternativas são

25 preferíveis, uma vez que elas são menos trabalhosas e são mais eficientes na obtenção uniforme de distribuição das células simples dentro dos recipientes simples. Os procedimentos de enriquecimento, classificação e isolamento descritos acima são realizados, tal que, a maior parte das células é mantida intacta. A ruptura das células durante o enriquecimento e classifica-

30 ção pode resultar na mistura desordenada das sequências de codificação de região variável. Contudo, não se espera que isto constitua um problema, uma vez que a ruptura é esperada como sendo baixa. A lavagem é possível

tratamento das células com RNase antes da distribuição nos recipientes simples removerá qualquer RNA que tenha vazado durante o processo.

Adicionalmente, quando se considera as descrições acima de como distribuir as células de modo a se obter uma população de células simples em uma população de recipientes simples, não deve ser interpretado como um aspecto absolutamente necessário que cada recipiente deva conter uma célula simples. Ao invés disto, isto indica que a maioria dos recipientes contém células simples, por exemplo, o número de recipientes com duas ou mais células está abaixo de 25% da quantidade total de células distribuídas, ou mesmo melhor está abaixo de 10%.

Um aspecto adicional da presente invenção é o desempenho de uma transcrição reversa empregando gabarito derivado de células distribuídas individualmente entre vários recipientes. Para a finalidade da transcrição reversa (RT), de acordo com a presente invenção, os ácidos nucleicos dentro de uma célula simples que devem servir como fonte de gabarito para RT, são considerados como sendo derivados de uma célula simples, embora eles não tenham sido necessariamente separados do teor remanescente daquela célula simples.

Quando a distribuição final das células simples em seus recipientes simples tiver sido realizada, as células simples podem ser expandidas a fim de se obter uma população de células isogênicas antes da transcrição reversa. Este processo rende mais RNAm a ser usado como gabarito, que pode ser importante se um alvo raro precisar se ampliado e ligado. Contudo, as células permaneceriam geneticamente idênticas com relação ao gene-alvo durante a expansão. As células isoladas ou a população das células isogênicas pode tanto ser mantida intacta quanto lisada, à medida que o gabarito para a transcrição reversa não for degradado. Preferivelmente, as células são lisadas a fim de facilitar a transcrição reversa e ampliação da PCR que se seguem.

Em uma concretização diferente da presente invenção, o método de RT-PCR de sobreposição-extensão multiplexa ou RT-PCR multiplexa seguida por ligação por união ou recombinação descrevidos podem também

ser utilizados no gabarito derivado de uma população de células geneticamente diversa que não foi separada em recipientes simples, porém permanecem juntas como um grupo de células. Este método pode ser usado para a geração de bibliotecas combinatorias. Tal abordagem não necessitará de distribuição das células simples. Contudo, as células que podem ser usadas nesta abordagem são as mesmas que as descritas para a abordagem da célula simples, por exemplo, a população (grupo) de linfócitos B ou linfócitos T classificada. Quando da realização da RT-PCR de sobreposição-extensão multiplexa de etapa única ou RT-PCR multiplexa de etapa única seguida por ligação por união ou recombinação em tal população de células, é preferível lisar as células antes da reação e, caso desejado, RNA total ou RNAm pode ser isolado do lisado. A sensibilidade da RT-PCR de sobreposição-extensão multiplexa de etapa única da presente invenção permite o emprego de uma quantidade muito pequena de gabarito, por exemplo, uma quantidade de gabarito correspondendo ao lisado de uma célula simples.

Misturas de Iniciador e Projeto

As misturas de iniciador da presente invenção compreendem pelo menos quatro iniciadores que formam conjuntos de iniciadores dois a dois, que são capazes de ampliação pelo menos de duas sequências-alvo de interesse diferentes. Misturas de dois ou mais de tais conjuntos de iniciadores constituem uma mistura de iniciador multiplexa. Preferivelmente, uma mistura multiplexa compreende pelo menos 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 ou 20, 30, 40, 50, 60, 70, 80, 90 100, 110, 120, 130, 140 ou 150 conjuntos de iniciadores (pares de iniciadores). Especificamente para a ampliação de sequências de codificação de região variável, um conjunto de iniciador individual dentro da mistura de iniciador multiplexa pode constituir muito mais de dois iniciadores. Preferivelmente, um conjunto de iniciador individual compreende pelo menos 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 20, 25, 30, 35, 45, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 220, 240, 260, 280 ou 300 iniciadores. Preferivelmente o número total de iniciadores em uma mistura de iniciador multiplexa é de pelo menos 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 20, 25, 30, 35, 45, 50, 60, 70, 80,

90, 100, 125, 150 ou 200 e no máximo de 225, 250, 275, 300, 325, 350, 375 ou 400 iniciadores.

Todos os iniciadores da presente invenção compreendem uma região específica gênica e preferivelmente, todos os iniciadores são adicionalmente equipados com uma cauda de iniciador na extremidade 5' do iniciador, isto é, sequências não codificando 5' que são fundidas à extremidade 3' da parte do iniciador específico gênico. Tal cauda de iniciador possui aproximadamente 6 a 50 nucleotídeos de comprimento, porém ela pode também ser mais longa, caso desejado. Quando da ampliação, as caudas de iniciadores são adicionadas às sequências-alvo. As caudas de iniciadores da presente invenção são, por exemplo, caudas de clonagem e caudas de ligação, tais como, caudas adaptadas para ligação por união, caudas adaptadas para ligação por recombinação ou caudas de sobreposição-extensão.

As caudas de clonagem podem possuir 6 a 20 nucleotídeos de comprimento ou ser mais longas e compreendem sítios de restrição e/ou sítios de recombinação, que são úteis para a inserção do produto ligado a um vetor apropriado.

De modo a habilitar ligação por união, os conjuntos de iniciadores da mistura de iniciador multiplexa são designados tal que uma parte (iniciador(es) a frente e reverso) do primeiro conjunto de iniciador é equipada com uma cauda de ligação contendo um sítio de restrição que, mediante clivagem será compatível com um sítio de restrição localizado em uma cauda de ligação de uma parte do segundo conjunto de iniciador. Para ligação de mais de duas sequências-alvo, a segunda parte do segundo conjunto de iniciador é equipadas com um sítio de restrição que mediante clivagem será compatível com um sítio de restrição localizado em uma parte do terceiro conjunto de iniciador. Este segundo sítio de restrição localizado no segundo conjunto de iniciador seria não-compatível com aquele do primeiro conjunto de iniciador. Um número considerável de sequências-alvo podem ser ligadas por projetar-se o conjunto de iniciador desta forma. Os sítios de restrição com uma frequência baixa ou sem ocorrência nas sequências-alvo seriam escolhidos. Adicionalmente, é preferível que sítios de restrição compatíveis não sejam

idênticos, tal que, o sítio de ligação se torna resistente à clivagem para as enzimas de restrição específicas usadas. Isto acionará a reação na direção da ligação da sequência-alvo um com a sequência-alvo dois, uma vez que a ligação entre sequências-alvo idênticas será clivável pelas enzimas de restrição. Pares apropriados de sítios de restrição são, por exemplo, SpeI com XbaI (alternativamente NheI ou AvrII podem substituir um ou ambos), NcoI com BspHI, EcoRI com MfeI ou PstI com NsiI. Para ligação, SpeI pode estar localizado, por exemplo, na sequência-alvo um, XbaI pode estar localizado na sequência-alvo dois, NcoI pode estar localizado na outra extremidade da sequência-alvo dois e BspHI na sequência-alvo três e assim por diante. Adicionalmente para simplificar o processo, será uma vantagem se a função das enzimas de restrição estiver no mesmo tampão.

De modo a habilitar a ligação por recombinação, os conjuntos de iniciadores da mistura de iniciador multiplexa podem, por exemplo, ser projetados conforme exemplificado no artigo por Chapal (1997 BioTechniques 23, 518-524), que é incorporado no presente documento como referência.

De modo a habilitar a ligação das sequências de nucleotídeo de interesse na mesma etapa que a ampliação da PCR multiplexa, as caudas adaptadas para PCR de sobreposição-extensão são adicionadas a pelo menos um iniciador de cada conjunto de iniciador da mistura de iniciador multiplexa, pelo que, gerando uma mistura de iniciador de sobreposição-extensão multiplexa.

As caudas de sobreposição-extensão são tipicamente mais longas, variando de 8 a 75 nucleotídeos de comprimento e podem conter sítios de restrição ou sítios de recombinação que permitem a inserção subsequente de elementos reguladores, tais como, promotores, sítios de ligação ribossômica, sequências de terminação, ou sequências ligantes, tais como, em um scFv. A cauda de sobreposição-extensão também pode conter um códon de para se isto for desejado. De modo geral, existem três tipos de caudas de sobreposição-extensão, conforme ilustrado na figura 1 do WO 2005/042774. No tipo I, as caudas de sobreposição-extensão dos dois conjuntos de iniciadores se sobrepõem unicamente uma com a outra. Não necessariamente

todos os nucleotídeos das duas caudas de sobreposição-extensão são complementares um ao outro. Em um aspecto da presente invenção, os nucleotídeos complementares representam entre 60 a 85% da cauda de sobreposição-extensão. No tipo II, as caudas de sobreposição-extensão, 4 a 6 dos nucleotídeos 5' são complementares à região específica gênica da sequência-alvo adjacente. No tipo III das caudas de sobreposição-extensão, toda a sobreposição é complementar à sequência-alvo adjacente. Os tipos I e II das caudas de sobreposição-extensão são preferidos quando elementos reguladores e semelhantes são mais tarde inseridos entre as sequências-alvo ligadas. O tipo II das caudas de sobreposição-extensão é o preferido se as sequências-alvo forem ligadas por um ligador definido conforme visto com scFv. O tipo III de caudas de sobreposição-extensão é o preferido se as sequências-alvo forem ligadas em estrutura uma a outra.

O projeto das caudas de sobreposição-extensão depende dos aspectos da sequência, tais como, comprimento, teor de GC relativo (% de GC), presença dos sítios de restrição, palíndromos, temperatura de fusão, a parte específica gênica a qual eles são acoplados, etc. O comprimento das caudas de sobreposição-extensão estaria entre 8 e 75 nucleotídeos de comprimento, preferivelmente de 15 a 40 nucleotídeos de comprimento. Mesmo mais preferivelmente, elas teriam de 22 a 28 nucleotídeos de comprimento. O emprego de caudas de sobreposição-extensão muito longas (50 a 75 nucleotídeos) favoreceria a ligação dos produtos fabricados pelo conjunto iniciador independente. Contudo, a proporção entre o comprimento da cauda de sobreposição-extensão e a região específica gênica provavelmente precisaria ser ajustado quando se emprega caudas de sobreposição-extensão muito longas. A preferência de % de GC depende do comprimento da cauda de sobreposição-extensão. Uma vez que caudas mais curtas possuem área mais curta onde elas são complementares, elas precisam de uma %GC maior para resistirem a interação que as caudas mais longas. Outros princípios do projeto de iniciador seriam da mesma forma observados, por exemplo, dimerização do iniciador e formação de cabeças diminutas seriam minimizadas. Adicionalmente, sabe-se que Taq DNA polimerase frequentemente adi-

ciona uma adenosina (A) na extremidade 3' do filamento de DNA recentemente sintetizado, e isto pode ser acomodado no projeto de cauda de sobreposição-extensão por permitir que as caudas de sobreposição-extensão acomodem a adição de não-gabarito A 3'.

5 A escolha dos iniciadores que portam uma cauda de ligação, por exemplo, a cauda de sobreposição-extensão, cauda adaptada para ligação ou união ou cauda adaptada para ligação por recombinação define a ordem e a direção da ligação das sequências-alvo. Não é essencial para a invenção se for(em) o(s) iniciador(es) a frente ou iniciador(es) inverso(s) de um conjunto de iniciador ou possivelmente ambos iniciadores a frente e reversos
10 que são equipados com uma cauda de ligação. Contudo, alguma consideração seria feita com relação a isto, uma vez que a ordem e direção das sequências-alvo no produto final pode ser de relevância, por exemplo, para a inserção de elementos reguladores, tais como, promotores e sequências
15 terminadoras ou para a ligação na estrutura das sequências-alvo individuais. Para a ligação das duas sequências de nucleotídeo de interesse, a cauda de ligação pode ser adicionada tanto no(s) iniciador(es) reverso(s) quanto iniciador(es) a frente de cada conjunto de iniciador usado para a ampliação da PCR de cada sequência-alvo.

20 A presente invenção exemplifica a adição das caudas de sobreposição-extensão e caudas adaptadas para ligação por união, aos iniciadores de mVH e mVK de cada conjunto. Isto resulta em uma direção de ligação dos produtos que é 5' para 5' (cabeça-a-cabeça e bidirecional). Contudo, as caudas de ligação podem também ser adaptadas ao(s) iniciador(es) reverso(s) de cada conjunto. Isto resulta em uma direção de ligação do produto
25 que é 3' para 3' (cauda-a-cauda e bidirecional). Uma terceira opção é a adição de caudas de ligação ao(s) iniciador(es) reverso(s) do primeiro conjunto de iniciador e o(s) iniciador(es) a frente do segundo conjunto de iniciador ou vice-versa. Isto resulta em uma orientação 3' para 5' (cabeça-a-cauda e unidirecional).
30

Quando se liga mais de duas sequências de nucleotídeo de interesse, alguns dos conjuntos de iniciadores precisam ter caudas de ligação

em ambos iniciadores a frente e reverso, tal que, uma cauda seja complementar a cauda do conjunto de iniciador precedente e a outra cauda seja complementar a um dos iniciadores do conjunto de iniciador subsequente. Este princípio mantém todos os conjuntos de iniciadores que ampliam sequências-alvo que devem estar ligados entre duas outras sequências-alvo. O projeto da parte de iniciador específica gênica geralmente observaria regras de projeto de iniciador conhecidas, tais como, minimização da dimerização do iniciador, formação de cabeça diminuta e recombinação não-específica. Adicionalmente, múltiplos nucleotídeos G ou C como as bases 3' devem ser evitados, quando possível. A temperatura de fusão (T_m) das regiões específicas gênicas em um conjunto de iniciador preferivelmente seria igual uma a outra mais/menos 5°C . Na presente invenção, valores de T_m entre 45°C e 75°C são desejáveis e valores de T_m de cerca de 60°C são ótimos para a maioria das aplicações. Vantajosamente, o projeto de iniciador primário pode ser auxiliado por programas de computador desenvolvidos para esta tarefa. Contudo, projetos de iniciador geralmente precisam de testes laboratoriais e otimização de rotina. Isto pode ser realizado, por exemplo, por análise de tamanho, polimorfismo de comprimento de fragmento (RFLP) e sequenciamento dos produtos de ampliação obtidos usando os conjuntos de iniciador. O uso de posições de degeneração dentro dos iniciadores é uma abordagem útil quando se ampliam sequências com regiões variáveis ou quando se pesquisa novos membros da família pertencendo a uma classe específica de proteínas. Os números de posições de degeneração podem também precisar de otimização.

Um aspecto da presente invenção são misturas de iniciadores compostas de pelo menos dois conjuntos de iniciadores que são capazes de iniciar a ampliação e promover a ligação de pelo menos duas sequências de nucleotídeo de interesse. As misturas de iniciadores da presente invenção são capazes de iniciar a ampliação de pelo menos duas subunidades ou domínios das proteínas heteroméricas, por exemplo, pertencendo a classe das enzimas, inibidores, proteínas estruturais, toxinas, proteínas de canal, proteínas-G, proteínas de receptor, proteínas da superfamília das imunoglo-

bulinas, proteínas da transporte etc., preferivelmente imunoglobulinas.

Um aspecto adicional da presente invenção é uma mistura de iniciador de sobreposição-extensão multiplexa compreendendo conjuntos de iniciadores onde pelo menos um elemento do conjunto de iniciador de cada conjunto de iniciador compreende uma cauda de sobreposição-extensão capaz de hibridizar com a cauda de sobreposição-extensão de um elemento do conjunto de iniciador de um segundo conjunto de iniciador.

As caudas de sobreposição-extensão permitem a ligação imediata dos nucleotídeos de interesse durante a ampliação da PCR de sobreposição-extensão multiplexa por equipar cada produto individual surgindo dos conjuntos de iniciadores com uma cauda que é complementar ao produto de anexação. Isto contudo, não significa que a ligação ocorre necessariamente durante esta primeira ampliação da PCR. Dependendo do estabelecimento da reação, a maior parte da ligação real pode ser realizada durante uma ampliação adicional com os iniciadores externos da primeira ampliação da PCR (ampliação da PCR multiplexa).

Um aspecto adicional da presente invenção é um conjunto de iniciador projetado para ampliar uma família de sequências de nucleotídeo contendo sequências de codificação de região variável. Exemplos de tais famílias são cadeias leve capa (por exemplo, VK1-19 em camundongos), cadeias leve lambda (por exemplo, VL1-8 em camundongos) e cadeias pesadas variáveis (por exemplo, VH1-15 em camundongos) a partir de imunoglobulinas e regiões variáveis α , β , γ ou δ TcR. Um conjunto de iniciador para a ampliação de uma família de sequências de nucleotídeo contendo sequências de codificação de região variável frequentemente compreende vários iniciadores onde inúmeros iniciadores podem ser iniciadores degenerados. A ampliação das famílias das sequências de codificação da região variável da cadeia leve de imunoglobulina é, por exemplo, realizada usando um conjunto iniciador compreendido de vários iniciadores complementares à extremidade 5' da região variável da cadeia capa (iniciador(es) mVK)) ou a sequência líder capa e/ou cadeia lambda ou a sequência líder lambda (iniciadores a frente) em conjunto com a região constante capa (iniciador(es) mKappar1)) e/ou

iniciadores lambda (iniciadores reversos) ou vários de tais iniciadores.

Alternativamente, cadeia leve unindo iniciadores de região pode ser usada como iniciadores reversos ao invés dos iniciadores de região constante. Alternativamente, os iniciadores a frente recombina-
5 UTR precedendo a sequência líder da cadeia leve variável podem ser usa-
dos. Igualmente, famílias de sequências de codificação de região variável de
cadeia pesada de imunoglobulina podem ser ampliadas com um conjunto de
iniciador empregando várias combinações de iniciador. Por exemplo, uma
pluralidade de iniciadores complementares à extremidade 5' da região variá-
10 vel de cadeia pesada (iniciador(es) mVH) ou a sequência líder desta região
(iniciadores a frente) em conjunto com vários iniciadores de região ligados a
cadeia pesada ou iniciador(es) de região constante de cadeia pesada (inici-
adores reversos). O iniciador mCH pode ser específico para isotipo e em
princípio, qualquer iniciador mCH pode ser utilizado, também um que resul-
15 tasse em uma cadeia pesada de comprimento pleno. Preferivelmente, um
iniciador mCH é empregado, o qual não amplia a cadeia pesada de compri-
mento pleno, de modo a habilitar a adição de uma região constante de ca-
deia pesada humana. Alternativamente, os iniciadores a frente recombina-
do na região UTR precedendo a sequência líder da cadeia pesada variável
20 podem ser usados.

O emprego de iniciadores a frente recombina-
do na sequência líder ao invés da extremidade 5' da região variável é especificamente útil se
hibridização cruzada, devido ao alto grau de similaridade da sequência, for
observada para os iniciadores de região variável. Mutações devido à hibridi-
25 zação cruzada empregando iniciadores líder serão eliminadas da proteína
final porque as sequências líder são clivadas durante o processamento da
proteína dentro da célula.

Um aspecto da presente invenção se constitui nos iniciadores
que recombina-
do na extremidade 3' da sequência de codificação líder prece-
30 dendo uma sequência de codificação de região variável e seu uso para am-
pliação das sequências de codificação de região variável.

Em uma concretização da presente invenção, a mistura de inicia-

dor de sobreposição-extensão multiplexa utilizada para a PCR de sobreposição-extensão multiplexa e possivelmente para a etapa de transcrição reversa compreende:

5 a) pelo menos um iniciador mKappar1 ou hmJK complementar ao filamento de sentido de uma sequência de codificação de região de cadeia leve de imunoglobulina;

b) pelo menos um iniciador mVK complementar ao filamento de antissenso de uma sequência de codificação de região variável de cadeia leve de imunoglobulina ou sequência líder de região variável de cadeia leve, e capaz de formar um conjunto de iniciador com o iniciador(es) na etapa a);

10 a) pelo menos um iniciador mCHrev1, mHCrev1-ext ou mJH complementar ao filamento de sentido de uma sequência de codificação de domínio de cadeia pesada de imunoglobulina; e

b) pelo menos um iniciador mVH complementar ao filamento de antissenso de uma sequência de codificação de região variável de cadeia pesada de imunoglobulina ou sequência líder de região variável de cadeia pesada e capaz de formar um conjunto de iniciador com o iniciador(es) na etapa c).

20 Em uma concretização da presente invenção o iniciador de cadeia leve é adaptado para ampliação de ambas as sequências de codificação de região variável de cadeia leve κ e λ .

25 Em uma concretização adicional da presente invenção, os iniciadores mVK de imunoglobulina portam caudas de ligação, preferivelmente na forma de caudas de sobreposição-extensão complementares. Isto gera sequências de codificação de região variável que são ligadas em um modo cabeça-a-cabeça. Para a ligação de sequências de codificação de região variável em um modo cabeça-cauda, tanto os iniciadores mKappar1 quanto mVH contém caudas de ligação quanto os iniciadores mVK e mCHrev1 contém caudas de ligação, preferivelmente na forma de caudas de sobreposição-extensão complementares. Para a ligação das sequências de codificação de região variável e em um modo cauda-a-cauda, os iniciadores mCH e mKappar1 contém caudas de ligação, preferivelmente na forma de caudas

de sobreposição-extensão complementares. Preferivelmente, as misturas de iniciador multiplexas, incluindo misturas de iniciador de sobreposição-extensão multiplexas da presente invenção compreendem dois conjuntos de iniciadores. Assim, uma mistura de iniciador multiplexa compreende pelo
5 menos quatro iniciadores diferentes. Em um aspecto adicional da presente invenção uma mistura de iniciador multiplexa compreende mais de quatro iniciadores diferentes. Uma mistura de iniciador multiplexa da presente invenção é empregada para a ampliação de sequências-alvo em um recipiente simples. Por exemplo, regiões variáveis de cadeia pesada, capa e lambda
10 podem ser ampliadas no mesmo recipiente.

A presente invenção também engloba iniciadores para uma ampliação adicional da PCR dos produtos ligados obtidos por RT-PCR multiplexa seguida por ligação por união ou recombinação ou por RT-PCR de sobreposição-extensão multiplexa. Esta ampliação adicional da PCR pode ser
15 realizada empregando uma mistura de iniciador adaptada para ampliação das sequências-alvo ligadas. Tal mistura de iniciador pode compreender os iniciadores externos da mistura de iniciador multiplexa ou mistura de iniciador de sobreposição-extensão multiplexa, significando os iniciadores que recombina-
20 m em relação a extremidade mais externa 5' e extremidade 3' do filamento de sentido das sequências de nucleotídeo ligadas, pelo que permitindo seletivamente a ampliação de todo o produto ligado. Este processo de modo geral serve para aumentar a quantidade de produto ligado obtida da RT-PCR multiplexa seguida por ligação por união ou recombinação ou da RT-PCR de sobreposição-extensão multiplexa.

Alternativamente, um conjunto de iniciador que é aninhado em
25 comparação aos iniciadores externos empregados na RT-PCR multiplexa primária ou reação de RT-PCR de sobreposição-extensão multiplexa pode ser usado para a ampliação adicional das sequências de nucleotídeo ligadas. Na presente invenção, tal conjunto de iniciador é denominado é denomi-
30 nado um conjunto de iniciador aninhado. O projeto de iniciadores aninhados de modo geral observa as mesmas regras de projeto dos iniciadores específicos gênicos descritos previamente, exceto que eles iniciam parcial

ou totalmente 3' com relação à posição de recomposição dos iniciadores externos usados na RT-PCR multiplexa ou RT-PCR de sobreposição-extensão multiplexa. O produto resultante da PCR aninhada, portanto, pode ser mais curto que o produto ligado obtido pela RT-PCR multiplexa seguida por ligação por união ou recombinação ou por RT-PCR de sobreposição-extensão multiplexa. Além do aumento da quantidade de produto ligado, a PCR aninhada serve para aumentar a especificidade total, especialmente da tecnologia de RT-PCR de sobreposição-extensão multiplexa. Contudo, deve ser observado que nem todas as misturas de iniciador multiplexas/misturas de iniciador de sobreposição-extensão multiplexas que foram descritas previamente são apropriadas para combinação com um conjunto de iniciador aninhado quando da realização de uma ampliação adicional. Em tais casos, os iniciadores externos da mistura de iniciador multiplexa/mistura de iniciador de sobreposição-extensão multiplexa podem ser usados para a ampliação adicional ou uma PCR semianinhada pode ser aplicada conforme descrito mais adiante.

Em uma concretização da presente invenção, uma mistura de iniciadores J_L e J_H iniciadores é empregada como iniciadores aninhados para a ampliação adicional das sequências de codificação de região variável de imunoglobulina ligadas.

Conjunto de iniciador aninhados da presente invenção também podem ser compreendidos de iniciador(es) externo reverso(s) (ou a frente) a partir da primeira mistura de iniciador multiplexa/mistura de iniciador de sobreposição-extensão multiplexa e um segundo iniciador aninhado que inicia 3' em relação à posição de recomposição do(s) iniciador(es) externo(s) a frente (ou reverso(s)) da primeira mistura de iniciador multiplexa/mistura de iniciador de sobreposição-extensão multiplexa. O emprego de tal um conjunto de iniciador para uma ampliação adicional da PCR é de modo geral conhecido como uma PCR semianinhada. A PCR semianinhada pode ser aplicada, se for difícil projetar um iniciador aninhado em uma região específica, por exemplo, para as sequências de região variável, uma vez que tal iniciador teria de se recompor nas regiões determinantes de complementaridade

(CDRs). Adicionalmente, PCR semianinhada pode ser empregada quando é desejável manter uma extremidade das sequências ligadas intacta, por exemplo, para fins de clonagem.

Otimização da PCR de sobreposição-extensão multiplexa

5 Os parâmetros da etapa da PCR de sobreposição-extensão multiplexa de ambos o procedimento de etapa única e duas etapas podem ser otimizados de várias formas (vide, por exemplo, Henegariu, O. e outros, 1997. *BioTechniques* 23, 504-511; Markoulatos, P. e outros, 2002. *J. Clin. Lab. Anal.* 16, 47-51). De modo geral, os mesmos parâmetros de otimização
10 se aplicam à RT-PCR multiplexa, embora a razão entre os iniciadores externo e interno seja menos importante para tal reação.

a. Concentração do Iniciador

A concentração dos iniciadores transportando a cauda de sobreposição-extensão (por exemplo os iniciadores V_H e V_L) é preferivelmente
15 inferior à concentração dos iniciadores externos sem cauda de sobreposição-extensão (por exemplo iniciadores J_H e J_L).

Se uma das sequências-alvo ampliar com uma eficiência menor que as outras, por exemplo, como resultado de %GC maior, será possível equalizar a eficácia da ampliação. Isso pode ser realizado por emprego de
20 uma concentração maior do conjunto de iniciador que media ampliação com baixa eficiência, ou diminuição da concentração de outros iniciadores. Por exemplo, sequências codificando regiões variáveis de cadeia pesada tendem a apresentar %GC maior e conseqüentemente eficiência de ampliação mais baixa que as regiões variáveis de cadeia leve. Isto aponta na direção
25 do emprego de iniciadores V_L em uma concentração inferior que os iniciadores V_H .

Adicionalmente, quando se emprega um número maior de iniciadores, a concentração total de iniciador pode ser uma questão a ser observada. O limite superior é determinado experimentalmente por experimentos
30 de titulação. Para o sistema PCR "AmpliTaq Gold" da PCR Applied Biosystems, o limite superior foi encontrado como sendo de 1,1 μM de concentração total de oligonucleotídeo, para outros sistemas, isto contudo, pode ser

tão alto quanto de 2,4 μ M. Tal limite superior de concentração de oligonucleotídeo total influencia a concentração máxima de iniciadores individuais. Se a concentração de iniciador individual for muito baixa, é provável que cause uma baixa sensibilidade de PCR.

- 5 Foi verificado que a qualidade dos iniciadores de oligonucleotídeo é importante para a PCR de sobreposição-extensão multiplexa. Oligonucleotídeos purificados por HPLC produziram os melhores resultados.

b. Condições de Ciclagem da PCR:

Preferivelmente as condições de ciclagem são como se segue:

Desnaturação	10-30 segundos	94°C	
Recombinação	30-60 segundos	50-70°C	Aproxunmadamente 5°C abaixo dos iniciadores
Extensão: Tm	1 minuto x EPL	65-72°C	ELP é o Comprimento do produto esperado em kb
Número de ciclos	30-80		
Extensão final	10 minutos	65-72°C	

- 10 Para a RT-PCR de sobreposição-extensão multiplexa de etapa única as etapas que se seguem foram constituídas no programa de ciclagem antes da ciclagem de ampliação ressaltada acima:

Transcrição reversa	30 minutos	42-60°C	Estas condições são também usadas quando a transcrição reversa separada é realizada
Ativação da polimerase	10-15 minutos	95°C	Polimerases iniciadas a quente são favoráveis na RT-PCR de etapa única. Ativação de acordo com as instruções do fabricante

- 15 É possível otimizar todos estes parâmetros. A temperatura de recombinação é especialmente importante. Assim, inicialmente todos os conjuntos de iniciadores individual que constituem a mistura de iniciador final devem ser testados separadamente, a fim de se identificar temperatura e tempo de recombinação ótimos, bem como tempos de alongamento de desnaturação. Isto fornecerá uma boa idéia a respeito da janela dentro da qual estes parâmetros podem ser otimizados para a mistura de iniciador de so-

breposição-extensão multiplexa.

Problemas com baixa sensibilidade de PCR, por exemplo, devido à baixa concentração do iniciador ou de concentração de gabarito podem ser superados empregando um número maior de ciclos térmicos. Um número maior de ciclos térmicos representa entre 35 e 80 ciclos, preferivelmente em torno de 40 ciclos. Adicionalmente, tempos de extensão mais longos podem aperfeiçoar os processos da PCR de sobreposição-extensão multiplexa. Tempos de extensão mais longos constituem 1,5 - 5 minutos x EPL em comparação à extensão de 1 minuto normal.

10 **c. Uso de Adjuvantes**

Reações de PCR multiplexa podem ser significativamente aperfeiçoadas por emprego de um aditivo da PCR, tal como, DMSO, glicerol, formamida ou betaína, com relaxamento do DNA, assim tornando a desnaturação do gabarito mais fácil.

15 **d. dNTP e MgCl₂**

A qualidade e a concentração do trifosfato de desoxinucleosídeo (dNTP) são importantes para a PCR de sobreposição-extensão multiplexa. A melhor concentração de dNTP está entre 200 e 400 μM de cada dNTP (dATP, dCTP, dGTP and dTTP), acima do que a ampliação é rapidamente inibida. Concentrações de dNTP mais baixas (100 μM de cada dNTP) são suficientes para obter ampliação da PCR. Soluções mestras de dNTP são sensíveis aos ciclos de descongelamento/congelamento. Após três a cinco de tais ciclos, a PCR multiplexa frequentemente não trabalha bem. Para evitar tais problemas, pequenas alíquotas de dNTP podem ser fabricadas e mantidas congeladas a -20°C .

A otimização da concentração de Mg^{2+} é crítica uma vez que a maioria das DNA polimerases são enzimas magnésio dependentes. Além da DNA polimerase, os iniciadores de gabarito de DNA e de dNT se ligam ao Mg^{2+} . Portanto, a ótima concentração de Mg^{2+} dependerá da concentração de dNTP, gabarito de DNA e composição de amostra de tampão. Se os iniciadores e/ou tampões de gabarito de DNA contiverem quelantes, tais como, EDTA ou EGTA, o aparente Mg^{2+} ótimo pode ser alterado. Concentração de

Mg²⁺ excessiva estabiliza o filamento duplo de DNA e impede desnaturação completa do DNA, o que reduz o rendimento. Concentração em excesso de Mg²⁺ pode também estabilizar recombinação espúria do iniciador para sítios de gabarito incorretos, pelo que, diminuindo a especificidade. Por outro lado, uma concentração de Mg²⁺ inadequada reduz a quantidade do produto. Um bom equilíbrio entre dNTP e MgCl₂ é de aproximadamente 200 a 400 μM dNTP (de cada) a 1,5 a 3 mM de MgCl₂.

e. Concentração do Tampão da PCR

De modo geral os tampões com base em KCl são suficientes para PCR de sobreposição-extensão multiplexa; contudo, tampões com base em outros componentes, tais como, (NH₄)₂SO₄, MgSO₄, Tris-HCl, ou combinações dos mesmos podem também ser otimizados para funcionar com a PCR de sobreposição-extensão multiplexa. Pares de iniciadores envolvidos na ampliação de produtos mais longos trabalham melhor em concentrações de sal mais baixas (por exemplo, 20 a 50 mM KCl), considerando-se que pares de iniciadores envolvidos na ampliação de produtos curtos trabalham melhor em concentrações de sal maiores (por exemplo, 80 a 100 mM KCl). A elevação da concentração de tampão para 2 vezes ao invés de 1 vez pode aperfeiçoar a eficiência da reação multiplexa.

f. DNA Polimerase

A presente invenção é exemplificada com Taq polimerase. Alternativamente, outros tipos de DNA polimerases resistentes ao aquecimento incluindo, por exemplo, Pfu, Phusion, Pwo, Tgo, Tth, Vent, Deep-vent podem ser empregadas. Polimerases sem ou com atividade de 3' a 5' exonuclease podem tanto ser usadas sozinhas quanto em combinação uma com a outra.

Vetores e Bibliotecas

A ligação das sequências de nucleotídeo de interesse de acordo com a presente invenção produz um segmento de nucleotídeo compreendendo as sequências de nucleotídeo ligadas codificando regiões variáveis das imunoglobulinas. Adicionalmente, bibliotecas de tais sequências de ácido nucleico ligadas são produzidas pelos métodos da presente invenção, especificamente bibliotecas de sequências de codificação de região variável

não-humanas ligadas ou recompostas com relação às sequências de região constante humana (cadeia pesada e leve).

Um aspecto da presente invenção é a inserção de um segmento contendo sequências de nucleotídeo ligadas de interesse ou uma biblioteca de sequências de nucleotídeo ligadas de interesse, geradas por um método 5 da presente invenção, em vetores apropriados. As bibliotecas podem ser bibliotecas combinatórias ou mais preferivelmente bibliotecas de pares cognatos de sequências de codificação de região variável. Os sítios de restrição gerados pelos iniciadores externos, iniciadores aninhados ou iniciadores semianinhados são preferivelmente projetados para emparelhar sítios de 10 restrição apropriados do vetor de escolha. As sequências de ácido nucleico de interesse ligadas podem também ser inseridas nos vetores por recombinação, se um dos iniciadores aninhados, semianinhados ou outros iniciadores forem equipados com um sítio de recombinação apropriado e o vetor de 15 escolha contiver um. Basicamente, não existem limitações aos vetores que podem ser usados como veículos dos produtos gerados por um dos métodos de ligação RT-PCR multiplexa da presente invenção. Os vetores de escolha podem ser aqueles apropriados para ampliação e expressão nas células incluindo, por exemplo, bactérias, levedura, outros fungos, células de insetos, 20 células de plantas ou células de mamíferos. Tais vetores podem ser usados para facilitar etapas de clonagem adicionais, transporte bidirecional entre sistemas de vetores, exibição do produto inserido no vetor, expressão do produto inserido e/ou integração no genoma de uma célula hospedeira.

Vetores de clonagem e transporte bidirecional são preferivelmente 25 vetores bacterianos. Contudo, os outros tipos de vetores podem também ser aplicados nos procedimentos de clonagem e transporte bidirecional.

Vetores de exibição podem ser, por exemplo, vetores de fago ou vetores de fagemido originários da classe de fd, M13 ou de bacteriófagos filamentosos f1. Tais vetores são capazes de facilitar a exibição de uma pro- 30 teína incluindo, por exemplo, uma proteína de ligação ou um fragmento da mesma, na superfície de um bacteriófago filamentoso. Os vetores de exibição apropriados para exibição nos ribossomas, DNA, células de levedura ou

células de mamíferos são também conhecidos na técnica. Estes compreendem, por exemplo, vetores viróticos ou vetores codificando proteínas quiméricas.

Vetores de expressão existem para todas as espécies mencionadas e o escolhido dependerá completamente da proteína a ser expressa. Alguns vetores de expressão são adicionalmente capazes de integração no genoma de uma célula hospedeira tanto por integração aleatória quanto por integração específica de sítio empregando sítios de recombinação apropriados. Os vetores de expressão podem ser designados para prover sequências de codificação adicionais que, quando o produto ligado é inserido na estrutura para estas sequências, permitem a expressão de uma proteína maior, por exemplo, um anticorpo monoclonal de comprimento pleno, quando introduzidos em uma célula hospedeira apropriada. Esta inserção na estrutura pode também facilitar a expressão das proteínas quiméricas que facilitam a exibição na superfície de um bacteriófago filamentosos ou célula. Em um sistema de exibição de bacteriófago, as sequências de nucleotídeo ligadas de interesse podem ser inseridas na estrutura em relação a sequência de codificação de uma proteína de revestimento, tal como, pIII ou pVIII (Barbas, C.F. e outros, 1991. Proc. Natl. Acad. Sci. USA 88, 7978-7982; Kang, A.S. e outros, 1991. Proc. Natl. Acad. Sci. USA 88, 4363-4366).

Em uma concretização da presente invenção, os segmentos individuais das sequências de nucleotídeo ligadas de interesse são compreendidos de uma sequência de codificação de região variável de cadeia pesada de imunoglobulina associada a uma sequência de codificação de região variável de cadeia leve de uma espécie, inserida em um vetor que contém sequência(s) de codificação de um ou mais domínios constantes de imunoglobulina, preferivelmente ambas regiões constantes de cadeia leve e pesada humanas. A inserção é construída, tal que, as sequências de codificação de sequência variável de cadeia leve e/ou região variável de cadeia pesada ligadas são inseridas na estrutura em relação às sequências de codificação de região constante. Tal inserção pode gerar, por exemplo, um vetor de expressão Fab ou F(ab')₂, um vetor de expressão de anticorpo de comprimen-

to pleno ou um vetor de expressão codificando um anticorpo de comprimento pleno. Preferivelmente, tal vetor é um vetor de expressão apropriado para expressão (por exemplo, *E. coli*, fagemido, ou vetores de mamíferos) e as sequências de codificação de cadeia pesada de região constante são escolhidas de classes de imunoglobulina humana IgG1, IgG2, IgG3, IgG4, IgM, IgA1, IgA2, IgD, ou IgE, pelo que permitindo a expressão de um anticorpo recombinante de comprimento pleno ou Fab. Além das sequências de codificação de cadeia pesada constantes, o vetor pode também conter uma sequência de codificação de cadeia leve constante escolhida de cadeias lambda ou capa humanas. Isto é preferido na geração de anticorpos quiméricos como as sequências de nucleotídeo ligadas nestes casos apenas codificam a região variável codificando sequências de imunoglobulina (FV's) das espécies não-humanas. Em uma concretização alternativa, a(s) região(ões) constante(s) humana(s) é/são recompostas ou ligadas às regiões variáveis não-humanas em uma etapa do procedimento de ampliação molecular, por adição aos recipiente de uma região constante humana provendo uma sobreposição com a sequência não-humana e iniciadores apropriados assegurando a ampliação de ambas regiões variável e constante na estrutura. Desta forma, a cadeia capa ou lambda constante humana pode ser adicionada e/ou uma cadeia pesada constante humana pode ser adicionada. Empregando-se este procedimento, não há necessidade de provisão de um sítio de restrição dentro da sequência de codificação, o que é uma vantagem.

Em outra concretização da presente invenção, os segmentos individuais das sequências de nucleotídeo ligadas são compreendidos de uma sequência de codificação de região variável de cadeia TcR α associada a uma sequência de codificação de região variável de cadeia β ou uma sequência de codificação de região variável de cadeia γ associada a uma sequência de codificação de região variável de cadeia δ . Preferivelmente, estas sequências ligadas são inseridas em um vetor que contém sequências que codificam um ou mais domínios constantes de TcR. A inserção é construída, tal que, as sequências ligadas de codificação de região variável estão na estrutura em relação às TcR sequências de codificação de região cons-

tante correspondentes TcR. Em uma concretização adicional, tal vetor é um vetor de expressão quimérico compreendendo sequências que codificam uma leucina zipper em relação às regiões constantes. Foi mostrado que tais construções aumentam a estabilidade das TcRs solúveis (Willcox, B.E. e outros, 1999. Protein Sci 8, 2418-2423).

5 Bibliotecas de pares cognatos da presente invenção podem ser introduzidas nos vetores por duas abordagens diferentes. Na primeira abordagem, os pares cognatos simples são inseridos individualmente em um vetor apropriado. A biblioteca de vetores pode então ser tanto mantida separa-
10 da ou agrupada. Na segunda abordagem, todos os pares cognatos são agrupados antes da inserção do vetor, seguido por inserção em massa nos vetores apropriados gerando uma biblioteca agrupada de vetores. Tal biblioteca de vetores compreende uma grande diversidade de pares de sequências de codificação de região variável.

15 Um aspecto da presente invenção é uma biblioteca de anticorpos com pares cognatos de sequências ligadas de codificação de região variável. Preferivelmente, os anticorpos individuais da biblioteca compreendem uma sequência de codificação de região variável de cadeia leve de imunoglobulina associada a uma sequência de codificação de região variável de
20 cadeia pesada de uma espécie e regiões constantes humanas.

Outra biblioteca preferida de pares cognatos compreende sequências de codificação de região de TcR ligada, onde cada sequência de codificação de região TcR individual compreende uma sequência de codificação de região variável de cadeia alfa associada a uma sequência de codi-
25 ficação de região variável de cadeia-beta e/ou uma sequência de codificação de região variável de cadeia gama TcR associada a uma sequência de codificação de região variável de cadeia delta.

Uma concretização da presente invenção é uma sub-biblioteca de pares cognatos de sequências ligadas de codificação de região variável
30 que codificam especificidades de ligação desejadas direcionadas contra um alvo específico. Preferivelmente estes pares cognatos compreendem região variável de cadeia leve de imunoglobulina ligada e sequências de codifica-

ção de região variável de cadeia pesada, região variável de cadeia alfa-TcR e sequências de codificação de região variável de cadeia-beta região variável de cadeia gama TcR e sequências de codificação de região variável delta.

5 Uma concretização adicional é uma sub-biblioteca de uma biblioteca de origem de pares cognatos de sequências de codificação de região variável conforme descritas através de toda a invenção.

Uma concretização preferida da presente invenção é uma biblioteca ou sub-biblioteca de codificação de pares cognatos de imunoglobulinas quiméricas de comprimento pleno, selecionadas das classes de imunoglobulina humanas IgA1, IgA2, IgD, IgE, IgG1, IgG2, IgG3, IgG4, ou IgM.

10 Outro aspecto preferido da presente invenção é uma biblioteca ou sub-biblioteca de codificação de pares cognatos de TcRs solúveis e estáveis.

15 Um aspecto da presente invenção é a diversidade das ditas bibliotecas, que são compreendidas de pelo menos 5, 10, 20, 50, 100, 1000, 10^4 , 10^5 ou 10^6 diferentes anticorpos de par cognato.

Em uma concretização adicional da presente invenção, as ditas bibliotecas de pares cognatos das sequências ligadas de codificação de região variável são obtidas por um método compreendendo as etapas descritas no presente documento. Esta biblioteca é também denominada biblioteca de origem.

Classificação e Seleção

25 Espera-se que a biblioteca de origem de pares de sequências ligadas de codificação de região variável isolada de um doador, utilizando um dos métodos da presente invenção, represente uma diversidade de proteínas de ligação das quais algumas serão irrelevantes, isto é, não ligando a um alvo desejado, especificamente para bibliotecas combinatórias. Portanto, a presente invenção engloba enriquecimento e classificação de uma sub-

30 biblioteca codificando um subconjunto de diversidades de especificidades de ligação direcionadas contra um alvo específico.

Para bibliotecas de pares cognatos, espera-se que a diversidade

da biblioteca represente a diversidade presente no material do doador, com apenas um pequeno número de regiões aleatoriamente variáveis ligadas. Assim, uma etapa de enriquecimento pode não ser necessária antes da classificação para afinidades de ligação específicas do alvo em uma biblioteca composta de pares cognatos.

Em uma concretização adicional da presente invenção, o método de geração de uma biblioteca pares de sequências ligadas de codificação de região variável, compreende, adicionalmente, a criação de uma sub-biblioteca por seleção de subconjuntos de pares de sequências de região variável ligadas que codificam proteínas de ligação com uma especificidade de alvo desejada. Tal seleção de sequências de codificação de região variável ligada é também denominada uma biblioteca de pares cognatos específicos para alvo.

Em uma concretização preferida da presente invenção, a biblioteca de pares cognatos específicos de alvo das sequências de codificação de região variável é transferida para um vetor de expressão. O vetor de expressão pode ser um vetor de expressão de mamífero, um vetor de expressão de levedura, um vetor de expressão de fungo, um vetor de expressão de planta, um vetor de expressão bacteriana dependendo do tipo de célula usada para classificação. Preferivelmente, o vetor de expressão é mamífero.

Ensaio imunológico são, de modo geral, apropriados para a seleção de sequências de codificação de região variável de imunoglobulina específicas de alvo. Tais ensaios são bem conhecidos na técnica e constituem, por exemplo, FMAT, FLISA, ELISPOT, ELISA, ensaios de membrana (por exemplo, Western blots), fileiras nos filtros ou FACS. Os ensaios podem tanto ser realizados de modo direito, utilizando os polipeptídeos produzidos das sequências de codificação de região variável de imunoglobulina. Alternativamente, os imunoenaios podem ser realizados em combinação com ou seguindo-se os métodos de enriquecimento, tais como, exibição de fago, exibição de ribossomo, exibição de superfície bacteriana, exibição de levedura, exibição de vírus eucariótico, exibição de RNA ou exibição covalente (revisto em FitzGerald, K., 2000. Drug Discov. Today 5, 253-258). Ambas

bibliotecas de expressão Fab cognato e bibliotecas de expressão de anticorpo de comprimento pleno cognato podem ser submetidas à classificação, pelo que, gerando uma sub-biblioteca de clones positivos. Tais ensaios de classificação e procedimentos de enriquecimento são também apropriados para os fragmentos Fv ou scFv ou bibliotecas combinatórias de regiões variáveis ligadas.

Além da classificação imunológica, um aspecto especial da invenção é que ela permite o emprego de vários tipos de classificação funcional para selecionar clones de secreção anticorpos com propriedades desejadas. Tais ensaios de classificação incluem, porém não estão limitados aos ensaios de proliferação, ensaios de inativação de vírus, ensaios de mortalidade de célula. Preferivelmente os ensaios funcionais podem ser realizados em formato de alto rendimento usando sobrenadantes de células transfectadas com vetores de expressão da invenção.

Em uma concretização preferida da presente invenção, a seleção de uma sub-biblioteca de pares cognatos específicos de alvo ou pares combinatórios de sequências de codificação de região variável é realizada por emprego de um ensaio de classificação de alto rendimento. Ensaios de classificação de alto rendimento poderiam ser, porém não estariam restritos aos ensaios ELISA realizados com equipamento semiautomatizado ou completamente automatizado. Eles também seriam um ensaio de membrana no qual as bactérias são coletadas roboticamente e passadas por grade em uma membrana apropriada na parte superior das placas de ágar gerando ensaios de colônias expressando moléculas de ligação de antígeno. As moléculas são secretadas através da membrana em uma segunda membrana revestida com antígeno subjacente que pode ser desenvolvido separadamente e usado para identificar clones que secretam moléculas de ligação de antígeno na direção do alvo desejado (de Wildt, R.M., e outros, 2000. *Nat. Biotechnol.* 18, 989-994).

Quando uma sub-biblioteca de pares cognatos ou pares combinatórios de clones de ligação de antígenos tiver sido selecionada por uma tecnologia apropriada é possível realizar uma análise adicional por sequên-

ciamento de DNA da

região variável de cadeia leve de imunoglobulina ligada e sequências de codificação de região variável de cadeia pesada. Primeiramente, tal sequenciamento do DNA fornecerá informações a respeito da diversidade da biblioteca, tal como, origem germinal, distribuição da família e maturação dentro das regiões CDR. Tal análise permitirá a seleção de clones que representam uma ampla diversidade e abandono de clones repetidos. Em segundo lugar, o sequenciamento do DNA descreverá mutações introduzidas durante o processo de isolamento.

Quando da análise das sequências de codificação de região variável existem três tipos de mutações a considerar quando da avaliação se uma mutação for aceitável: i) O tempo mais frequente das mutações resulta da iniciação cruzada, quando um iniciador do gene V devido às similaridades da sequência inicia uma sequência germinal a qual ele não é totalmente homólogo. As alterações introduzidas são principalmente substituições de códons ocorrendo naturalmente em uma posição específica. Devido ao alto grau de homologia da sequência entre sequências gênicas V. Algumas destas alterações podem ser significativas, algumas sem contraparte natural. Tais alterações poderiam afetar potencialmente a imunogenicidade da região variável por criação de novos epítopos. Tais alterações podem ser facilmente identificadas e subsequentemente reparadas usando técnicas biológicas moleculares padrão ou os clones podem ser excluídos da biblioteca; ii) Erros criados pela Taq DNA polimerase são mais facilmente identificados nas sequências de codificação de região constante e podem ser facilmente eliminados. Contudo, mutações induzidas por Taq naturalmente também estão presentes nas sequências de codificação de região variável onde elas não são distinguíveis das mutações somáticas ocorrendo naturalmente, que são também o resultado das mutações aleatórias nas sequências de codificação de região variável. Considerando-se que as mutações não são sistemáticas e apenas afetam pares específicos de modos distintos, parece razoável desconsiderar tais alterações. Em uma concretização adicional da presente invenção, a sub-biblioteca de pares específicos alvo e possivelmente analisa-

dos de sequência de região variável de cadeia leve de imunoglobulina ligada e sequências de codificação de região variável de cadeia pesada são transferidos para um vetor de expressão de mamífero. Tal transferência pode ser realizada em qualquer um dos vetores descritos na seção anterior, permitindo a expressão de um anticorpo recombinante de comprimento pleno. Se a classificação for realizada com uma biblioteca de expressão de anticorpo de comprimento pleno cognato de mamífero, tal transferência pode não ser necessária.

Em outra concretização da presente invenção, a biblioteca de origem é gerada de uma fração celular contendo linfócito que é enriquecida para linfócitos T. Os pares de sequências ligadas de codificação de região variável constituindo a biblioteca de origem podem ser selecionados para codificação de um subconjunto de pares de sequências de região variável ligadas, compostas de cadeias alfa e beta e/u gama e delta que codificam proteínas de ligação com uma especificidade alvo desejada, gerando uma sub-biblioteca de pares cognatos de pares combinatórios. Os receptores de célula T específicos de antígeno podem ser subseqüentemente identificados de um grupo de células transfectadas usando metodologia padrão, tal como, coloração com complexos de peptídeo MHC tetramérico (por exemplo, Callan, M.F. e outros, 1998. J. Exp. Med. 187, 1395-1402; Novak, E.J. e outros, 1999. J. Clin. Invest 104, R63-R67), medição de respostas celulares na forma de liberação de IL-2 ou por meios mais sofisticados, tais como, técnicas de exibição de levedura ou retrovirótica.

Células hospedeiras e Expressão

As bibliotecas da presente invenção podem ser transferidas para os vetores apropriados para expressão e produção de proteínas codificadas das sequências de ácido nucleico de interesse ligadas, especificamente região variável contendo proteínas de ligação ou seus fragmentos. Tais vetores são descritos na seção *Vetores e Bibliotecas* e fornecem expressão, por exemplo, de anticorpos de comprimento pleno, fragmentos Fab, fragmentos Fv, scFv, membrana ligada ou fragmentos de TcRs ou TcR de uma espécie de escolha.

Um aspecto da presente invenção é a introdução em uma célula hospedeira de uma biblioteca ou uma sub-biblioteca de vetores de pares cognatos de sequências ligadas de codificação de região variável ou um clone simples codificando um par cognato de sequências ligadas de codificação de região variável, para ampliação e/ou expressão. Células hospedeiras podem ser escolhidas de bactérias, levedura, outros fungos, células de inseto, células de plantas ou células de mamíferos. Para fins de expressão, células de mamíferos, tais como, células de ovário de hamster chinês (CHO, células COS, células BHK, células de mieloma (por exemplo, células Sp2/0, NS0), NIH 3T3, fibroblasto ou células humanas imortalizadas, tais como, células HeLa, células HEK 293 ou PER.C6 são preferidas.

A introdução de vetores nas células hospedeiras pode ser realizada por vários métodos de transformação ou transfecção conhecidos dos versados na técnica, incluindo precipitação de fosfato de cálcio, eletroporação, microinjeção, fusão de lipossoma, fusão de espectro RBC, fusão de protoplasma, infecção virótica e similares. A produção de anticorpos monoclonais de comprimento pleno, fragmentos Fab, fragmentos Fv e fragmentos scFv é bem conhecida.

A produção de anticorpos policlonais recombinantes a serem usados para tratamento é uma área bem nova. Uma tecnologia de fabricação policlonal recombinante foi descrita no Pedido PCT WO2004/061104. Em resumo, esta tecnologia envolve a geração de várias células, apropriadas como uma linha de células de fabricação. A descrição que se segue da técnica é realizada para uma biblioteca de pares cognatos, contudo é apenas aplicável a uma biblioteca combinatória. As células individuais na coleção de células são capazes de expressar um elemento distinto da proteína de ligação policlonal recombinante, por exemplo, de uma biblioteca de pares cognatos. A fim de garantir que as células individuais expressem um par cognato simples e não vários pares cognatos da proteína de ligação policlonal, as sequências de ácido nucleico codificando os pares cognatos são introduzidas em um sítio específico de sítio simples no genoma de cada célula individual. Isto é um aspecto importante da coleção de células, uma vez que isto

impede a mistura desordenada das cadeias pesada e leve expressas de cada célula, porém também, porque gera células que são virtualmente idênticas uma a outra, exceto pelas pequenas diferenças nas regiões variáveis dos pares cognatos individuais. Este traço permitirá um crescimento não inclinado da coleção de células por um período de tempo necessário para a produção. De modo a garantir integração específica de sítio simples, uma linhagem de célula hospedeira com apenas um sítio de integração seria empregada, havendo células CHO Flip-In da Invitrogen disponíveis comercialmente contendo um sítio FRT simples. Vetores apropriados para esta linhagem de célula contém um sítio FRT correspondente e são introduzidos no genoma usando a recombinase Fip. Existem várias outras recombinases conhecidas, por exemplo, Cre, beta-recombinase, Gin, Pin, PinB, PinD, R/RS, lambda integrase, ou fago Φ C31 integrase que podem ser usadas em combinação com seus sítios de recombinação correspondentes. Adicionalmente, vetores apropriados contêm um marcador de seleção que permite a seleção de integrantes específicos de sítio.

A geração de uma linhagem de célula de fabricação policlonal e a produção de uma proteína policlonal recombinante de tal linhagem de célula pode ser obtida por várias estratégias de transfecção e fabricação diferentes.

Um meio é utilizar uma biblioteca de vetores misturada em conjunto em uma composição simples, para a transfecção de uma linhagem de célula hospedeira com um sítio de integração simples por célula. Este método é denominado transfecção por volume ou transfecção em volume. De modo geral, o projeto do vetor e da célula hospedeira descrito anteriormente garantirá que uma linhagem de célula policlonal capaz de crescimento não inclinado será obtida quando da seleção apropriada. Uma solução mestra congelada da linhagem de célula policlonal será gerada antes do início da fabricação da proteína policlonal recombinante.

Outro modo é usar uma biblioteca de vetores dividida em frações, contendo 5 a 50 vetores individuais da biblioteca em uma composição, para transfecção. Preferivelmente, uma fração da biblioteca constitui 10 a 20 vetores individuais. Cada composição é então transfectada em uma alíquota de

células hospedeiras. Este método é denominado transfecção em semivolu-
me. O número de alíquotas transfectadas dependerá do tamanho da biblio-
teca e do número de vetores individuais em cada fração. Se a biblioteca
constituir, por exemplo, 100 pares cognatos distintos, que são divididos em
5 frações contendo 20 elementos distintos em uma composição, 5 alíquotas de
células hospedeiras precisariam ser transfectadas com uma composição de
biblioteca constituindo uma fração distinta da biblioteca original. As alíquotas
das células hospedeiras são selecionadas para integração específica do sí-
tio. Preferivelmente, as alíquotas distintas são selecionadas separadamente.
10 Contudo, elas podem também ser agrupadas antes da seleção. As alíquotas
podem ser analisadas quanto a sua diversidade clonal e apenas aquelas
com diversidade suficiente serão empregadas para gerar uma solução mes-
tra de biblioteca de par cognato policlonal. De modo a obter a linhagem de
célula policlonal desejada para fabricação, as alíquotas podem ser mistura-
15 das antes da geração da solução mestra de congelamento, imediatamente
após terem sido recuperados da solução mestra ou após um curto tempo de
proliferação e adaptação. Opcionalmente, as alíquotas das células são man-
tidas separadas através de toda a produção e a composição da proteína po-
liclonal é montada para combinar os produtos de cada alíquota ao invés das
20 alíquotas das células antes da produção.

Um terceiro modo é um método de alto rendimento no qual as cé-
lulas hospedeiras são transfectadas separadamente usando os vetores indi-
viduais constituindo a biblioteca de pares cognatos. Este método é denomi-
nado transfecção individual. As células hospedeiras transfectadas individu-
25 almente são preferivelmente selecionadas para integração específica do sítio
separadamente. Os clones de células individuais gerados da seleção podem
ser analisados com relação ao tempo de proliferação e, preferivelmente, a-
queles com taxas de crescimento semelhantes são usados para gerar uma
solução mestra de biblioteca de par cognato policlonal. Os clones de célula
30 individuais podem ser misturados para obter a linhagem de célula policlonal
desejada antes da geração da solução mestra, imediatamente após terem
sido recuperados da solução mestra ou após um tempo de proliferação e

5 adaptação curto. Esta abordagem pode eliminar qualquer inclinação de sequência residual possível durante a transfecção, integração e seleção. Alternativamente, as células hospedeiras individualmente transfectadas são misturadas antes da seleção ser realizada, isto permitindo controle da inclinação da sequência devido à transfecção.

Um aspecto compartilhado nas estratégias de fabricação ressaltadas acima é que todos os pares cognatos individuais constituindo a proteína policlonal recombinante podem ser produzidos em um ou um número limitado de biorreatores. A única diferença é o estágio no qual se escolhe gerar a coleção de células que constituem a linhagem de célula de fabricação policlonal.

15 Uma concretização da presente invenção é uma população de células hospedeiras compreendendo uma biblioteca de pares cognatos ou sub-biblioteca de pares ligados das sequências de codificação de região variável.

Em uma concretização adicional, uma população de células hospedeiras compreende uma biblioteca obtida de uma população de células simples isoladas constituindo linfócitos, utilizando a ampliação de RT-PCR multiplexa seguida por ligação por união ou recombinação ou tecnologia de RT-PCR de sobreposição-extensão multiplexa da presente invenção, de ligar os pares cognatos.

25 Outra concretização da presente invenção é uma população de células hospedeiras compreendendo uma biblioteca combinatória ou sub-biblioteca de pares ligados de sequências de codificação de região variável.

Uma população de células hospedeiras de acordo com a presente invenção englobará uma população de células diversa correspondendo à diversidade da biblioteca, as células tendo sido transformadas/transfectadas. Preferivelmente, cada célula da população de células apenas constitui um par cognato de toda biblioteca de pares cognatos e nenhum elemento individual da biblioteca de pares cognatos excede mais de 50%, mais preferivelmente 25%, ou o mais preferido 10% do número total de elementos individuais da população das células hospedeiras.

Em uma concretização preferida da presente invenção, a população das células hospedeiras é de células de mamíferos.

Uma população de células hospedeiras conforme descritas acima pode ser utilizada para a expressão de uma proteína de ligação policlonal recombinante, uma vez que células individuais da população constituem se-
5 quências de codificação de região variável de diversidade diferente.

Uma concretização da presente invenção é uma proteína policlonal recombinante expressa de uma população de células hospedeiras compreendendo uma biblioteca de vetores codificando pares cognatos diversos
10 de sequências ligadas de codificação de região variável, onde tal biblioteca é obtida pelo método da presente invenção. Tipicamente, uma proteína policlonal recombinante da presente invenção é compreendida de pelo menos 2, 5, 10, 20, 50, 100, 1000, 10^4 , 10^5 ou 10^6 proteínas compostas de diferentes pares cognatos.

Uma concretização preferida da presente invenção é uma imunoglobulina policlonal recombinante expressa de uma população de células hospedeiras compreendendo uma biblioteca de vetores codificando pares cognatos diversos da região variável de cadeia pesada e sequências de co-
15 dificação de região variável de cadeia leve.

Outra concretização preferida da presente invenção é uma TcR policlonal recombinante expressa de uma população de células hospedeiras compreendendo uma biblioteca de vetores codificando pares cognatos diversos da região variável de cadeia alfa-TcR ligados com as sequências de codificação de região variável e/ou região variável de cadeia gama TcR liga-
20 da às sequências de codificação de região variável da cadeia delta.

Outra concretização da presente invenção é uma célula hospedeira apropriada para produção de uma proteína monoclonal. Especificamente, um anticorpo monoclonal compreendido de um par cognato de uma região variável de cadeia leve com uma região variável de cadeia pesada ou
30 uma TcR monoclonal compreendida de um par cognato de uma região variável alfa com uma região variável beta ou uma região variável delta com uma região variável gama. Preferivelmente, tal linhagem de célula de produ-

ção monoclonal não é uma linhagem de célula de hibridoma.

Tal anticorpo monoclonal ou TcR pode ser gerado por adição das etapas que se seguem ao método de ligação de várias sequências de nucleotídeo de interesse descontínuas a) inserção das sequências de ácido nucleico ligadas em um vetor; b) introdução do dito vetor em uma célula hospedeira; c) cultivo das ditas células hospedeiras sob condições apropriadas para expressão; e d) obtenção do produto da proteína expresso do vetor inserido na dita célula hospedeira. Preferivelmente, o vetor introduzido na célula hospedeira codifica um par cognato individual de sequências de codificação de região variável.

Aplicações da Invenção

Uma das principais aplicações da presente invenção é a ligação de pares cognatos de sequências de codificação de região variável, especialmente sequências de codificação de região variável de cadeia pesada e leve de imunoglobulina ou sequências de codificação de região variável de cadeia alfa e beta TcR ou gama e delta, por um método de alto rendimento para a geração de bibliotecas de pares cognatos. Além da geração de bibliotecas de pares cognatos, a RT-PCR multiplexa seguida por ligação por união ou recombinação ou técnicas de RT-PCR de sobreposição-extensão multiplexa da presente invenção pode ser utilizada na geração de bibliotecas combinatórias de anticorpos humanos/não-humanos quiméricos por reação da técnica em uma população de células geneticamente diversas, lisados de células de tal população de células, ou em RNA purificado de tal população de células. As bibliotecas, sub-bibliotecas, ou clones simples de uma destas bibliotecas facilitam a expressão das proteínas policlonais ou monoclonais. Anticorpos especialmente monoclonais ou policlonais podem ser obtidos das bibliotecas da presente invenção.

É bem-conhecido o emprego de anticorpos monoclonais recombinantes no diagnóstico, tratamento e profilaxia. Os anticorpos monoclonais e policlonais recombinantes gerados pela presente invenção terão as mesmas aplicações que os produtos de anticorpo gerados pelas tecnologias existentes. Especificamente, uma composição farmacêutica compreendendo

uma imunoglobulina recombinante policlonal como ingrediente ativo, combinada a pelo menos um excipiente farmacologicamente aceitável, pode ser produzida por meio da presente invenção. São mais preferidas as composições farmacêuticas onde a imunoglobulina recombinante policlonal é compreendida de pares cognatos de sequências de codificação de região variável. Tais composições farmacêuticas de imunoglobulinas recombinantes policlonais podem ser usadas como medicamentos. A imunoglobulina recombinante policlonal da composição pode ser específica para ou reativa contra um alvo de doença predeterminado e a composição pode assim ser usada para o tratamento, melhora ou prevenção de doenças, tais como, câncer, infecções, doenças inflamatórias, alergia, asma e outras doenças respiratórias, doenças autoimunes, má função imunológica, doenças cardiovasculares, doenças no sistema nervoso central, doenças metabólicas e endócrinas, rejeição ao transplante ou gravidez indesejada em um mamífero, tal como, ser humano, animal doméstico ou animal de estimação.

Exemplos

Exemplo 1:

Camundongos Balb/c foram imunizados subcutaneamente com 50 μ g de toxoide tetânico (TT) em adjuvante de Freund completo. Os camundongos receberam reforço no dia 14 com 50 μ g TT em adjuvante incompleto de Freund. Após mais 30 dias, os camundongos receberam reforço com 50 μ g TT em adjuvante incompleto de Freund. Três dias após o último reforço, os camundongos foram sacrificados e o baço retirado e transferido para um tubo contendo 30 mL de RPMI 1640 FCS 10% em peso a 4°C. O tecido foi transferido para um crivo de célula de 74 μ m (Corning, 136350-3479) em um prato de 10 cm. Com a parte posterior de um êmbolo de seringa, os baços foram macerados através do filtro. O filtro foi enxaguado com 10 mL de RPMI 1640, solução de 10% de FCS. O filtro foi removido e o prato cheio com 20 mL de RPMI 1640 frio, 10% de FCS. As células foram transferidas para um tubo de 50 mL e centrifuga a 300 x g a 2 a 8°C por 5 minutos. As células foram ressuspensas em 5 a 10 mL a 4°C RPMI 1640 FCS 1% em peso e filtradas através de um filtro de seringa de 50 μ m (Becton Dickinson, 340603).

As células foram micronizadas e ressuspensas em FCS; 10% DMSO de modo a obter uma densidade de célula em 2×10^7 células/ampola e congeladas.

Ampolas congeladas com esplenócitos em suspensão de célula simples de camundongos Balb/c imunizados com toxoide tetânico foram des-
5 congeladas a 37°C e transferidas para tubo de 15 mL com gelo ainda presente. 10 mL de RPMI resfriado em gelo, 10% de FCS foram acrescentados gota a gota ao tubo enquanto rotacionando. Após uma lavagem em 10 mL FACS PBS, 4 mL PBS, 2% FCS foi adicionado antes da filtração das células através de Filcon 50 µm (Becton Dickinson, número CAT 340603). As células foram micronizadas e ressuspensas em 1 mL de PBS, 2% de FCS (final
10 volume) e subsequentemente coloridas tanto com anti CD43 FITC diluído a 1:100 (Número BD CAT 553270) e anti CD138 PE diluído a 1:40 (Número BD CAT 553714) em 1 mL de PBS, 2% de FCS ou com anti B220 APC (Número BD CAT 553092) diluído a 1:40 e anti MHCII FITC (Número BD CAT
15 553547) diluído a 1:200. As células foram incubadas a 4°C por 20 minutos no escuro. Finalmente, as células foram lavadas 2 vezes com 2 mL de PBS, 2% de FCS e adicionadas até 15 mL de PBS, 2% de FCS (soro bovino fetal). Imediatamente antes da classificação, PI foi adicionado a 1:100, e as células foram classificadas em uma contagem de aproximadamente 1.000 a 2.000
20 células/segundo. Exclusões para ambas colorações são ilustradas na figura 3. Figura 3A: Células PI positivo (mortas) as células foram excluídas no painel esquerdo inferior (P1). Então, as células plasmáticas foram excluídas com CD43 alto, CD138 alto no painel direito inferior (P2). Finalmente, os dupletos foram excluídos no painel direito superior do gráfico SSC-H, SSC-W
25 (P3). As células positivas para todas as três portas foram classificadas em placas ELISPOT de acordo com a tabela 1.

ELISPOT:

Placas de 96 poços com fundo de nitrocelulose (placas HA, Millipore, Bedford, MA) foram pré-umectadas com PBS (também os poços de
30 bloqueio) antes de serem revestidas com 100 µL, 25 µg/mL de toxoide tetânico (TT) ou IgG anti-murino de carneiro (Jackson Immuno Research, Número CAT 515-005-062) em PBS. O mesmo volume de PBS estava presente

nos poços de controle. A placa foi deixada a 4°C. No dia seguinte, os poços foram lavados três vezes com PBS antes de serem bloqueados com 200 µL de RPMI + leite em pó desnatado a 2% e deixados a 4°C. A placa foi movida para o incubador (37°C, 5% CO₂, 100% de umidade) 1 hora antes da adição das células. 100 µL de células em RPMI completo foram adicionados ao TT, poços revestidos com anti-IgG e apenas poços bloqueados. Células sem meio foram incluídas como controle. A placa foi movida para um incubador (37°C, 5% CO₂, 100% de umidade). No dia seguinte, a placa foi lavada seis vezes para remover as células, 3 vezes no tampão: PBS+0,01% Tween20 e 3x em PBS. Subsequentemente, os poços receberam IgG anti-camundongo de cabra conjugado com HRP (Caltag M30007) diluído 1:3000 em RPMI + leite em pó desnatado a 2% (100 µL/poço). Após 2 horas de incubação a 37°C, os poços foram lavados três vezes em PBS+0,01% Tween 20 seguido três vezes apenas com PBS. As manchas foram então desenvolvidas com 100 µL de substrato cromogênico consistindo em 0,015% de H₂O₂ e 0,3 mg/mL de 3-amino-9-etilcarbazol em 0,1M de acetato de sódio, pH 5,1. O desenvolvimento da cor foi parado por lavagem com água de torneira após 5 minutos.

Julgamento dos resultados da tabela 1. Células plasmáticas (PC; CD43 alto, CD138 alto, porta P3 na Figura 3A) específica para TT estão presentes em aproximadamente 2%, considerando-se plasma blastos específicos para TT (PB; B220 e MHCII positivo, Porta P4 na Figura 3B) estão presentes em aproximadamente 4%. Isto ilustra a superioridade de PC para PB na produção de anticorpos específicos após imunização com TT.

Tabela 1

		Número de manchas (número de células classificadas)		
Revestimento: PBS	PC	0 (1.001)	ND	ND
	PB	0 (13.976)	0 (1500)	0 (150)
Revestimento: TT	PC	22 (1.002)	2 (250)	0 (50)
	PB	54 (13.257)	0 (1500)	0 (150)
Revestimento: α-IgG	PC	55 (1.000)	6 (250)	1 (50)
	PB	48 (5000)	8 (1500)	0 (150)

Exemplo 2

A ampola congelada de esplenócitos foi colorida conforme descrito no exemplo 1. Quatro diferentes fenótipos foram classificados de quatro modos. As exclusões de classificação de classificação são ilustradas na figura 4. Primeiramente, células mortas ou PI positivas foram excluídas no painel esquerdo inferior (P1). P2 é CD138 intermediário, CD43 alto. P3 é CD138 alto, CD43 alto. P4 é CD138 alto, Cd43 negativo P5 é CD138 intermediário, CD43 baixo. 10.000 células positivas para P1 e cada uma das quatro exclusões foram classificadas em tubos de teste e congeladas para avaliação por Symplex de camundongo.

As frações P2, P3, P4 e P5 foram classificadas por volume nos tubos, cada um contendo 10.000 células. Os tubos foram centrifugados e ressuspensos em meio de Eagle modificado da Dulbecco contendo 2 U/ μ L de inibidor de RNase (RNasin, Promoga, CAT número N2511) a uma concentração de 250 células/ μ L, 10 μ L por tubo e congeladas a -80°C . Uma ampliação Symplex de camundongo foi realizada em uma diluição em série de cada um dos quatro conjuntos de linfócitos classificados para comparar o teor de RNAm de anticorpo IgG-capa. Os princípios básicos das reações foram:

- Primeiro é realizada uma reação RT, onde as cadeias pesada e leve são iniciadas por iniciadores de região constante;
- Em segundo lugar é realizada uma reação multiplexa empregando iniciadores de região VH e VK 5' cobrindo todas regiões variáveis e equipados com projeções complementares facilitando a formação das faixas de sobreposição. Os iniciadores 3' estão localizados na região constante das cadeias pesada e leve;
- Por ultimo é realizada uma reação aninhada ampliando apenas VH e VK conjugados usando iniciadores JH e JK;
- O produto de reação final consiste em extremidade 5' conjugada em VH e VK para extremidade 5' e conectado com um ligador. O tamanho seria de aproximadamente 700 pares de base.

Para a reação de RT-PCR multiplexa combinada foi empregado o

conjunto de iniciadores mostrado na tabela 2 e empregando o kit Qiagen OneStep RT-PCR essencialmente de acordo com as instruções do fabricante. As células congeladas foram descongeladas sobre gelo, ressuspensas e centrifugadas. Em cada série de diluições foi usado lisado de célula correspondendo a 100, 32, 10, 3,2, 1, 0,32, 0,1 e 0 células. Os volumes de reação

5 totais foram de 20 μ L. As condições de ciclagem foram:

- 55°C, 30 minutos
- 95°C, 15 minutos
- 94°C, 30 segundos
- 10 • 60°C, 30 segundos 35 ciclos
- 72°C, 5 minutos
- 72°C, 10 minutos

A reação aninhada foi realizada com o conjunto de iniciador mostrado na tabela 3 empregando FastStart polimerase (Roche) e reagentes fornecidos essencialmente de acordo com as instruções dos fabricantes. 1

15 μ L do produto de reação de RT-PCR foi empregado por reação aninhada em um volume total de 20 μ L. As condições de reação foram:

- 95°C, 30 segundos
- 60°C, 30 segundos 35 ciclos
- 20 • 72°C, 90 segundos
- 72°C, 10 minutos

10 μ L de cada produto de reação final foram finalmente analisados em um gel agarose a 1%.

A partir do resultado do Symplex nos lisados de célula titulada (figura 5), fica claro que nós queremos ligar regiões variáveis de

25 cadeias pesada e leve com células de P3 a jusante para aproximadamente 0,1 célula e de P2 a partir de aproximadamente 3,2 células. As outras exclusões foram perda eficiente com ligação apenas sendo possível de aproximadamente 32 células e mais. Na conclusão, o CD43

30 alto, CD138 alto (P3) é mais útil para SymplexTM em nível de célula simples, enquanto CD43 alto, CD138 intermediário seriam usados porém sendo menos eficientes.

Tabela 2: Conjunto de iniciador Symplex de camundongo empregado para reação de RT e multiplexa combinada

Iniciador	Conc. (nM)	Sequência	Número da Sequência
mH-Crev1	0,2	GACSGATGGGCCCTTGGTGG	1
mKap-par1	0,2	GCTGTAGGTGCTGTCTTTGC	2
mVH set			
mVH A	0,04	TATTCCCATGGCGCGCCSAGGTCCARCTGCAR-CAGYCTG	3
mVH B	0,04	TATTCCCATGGCGCGCCGARGTG-MAGCTKGTKGAGTC	4
mVH C	0,04	TATTCCCATGGCGCGCCSAGGTGCAGCTK-MAGGAGTC	5
mVH 8	0,04	TATTCCCATGGCGCGCCCAGGTTACTCTGAAA-GAGTC	6
mVH 9	0,04	TATTCCCATGGCGCGCCCAGATCCAGTTGGTG-CAGTCTG	7
Conjunto mVK			
mVK D	0,04	GGCGCGCCATGGGAATAGCTAGCCGAYATCCA GATGACHCARWCT	8
mVK E	0,04	GGCGCGCCATGGGAATAGCTAGCCRACATTGT GMTGACHCAGTC	9
mVK F	0,04	GGCGCGCCATGGGAATAGCTAGCCSAMATTGTK CTSACCCARTCTC	10
mVK 1-2	0,04	GGCGCGCCATGGGAATAGCTAGCCGATRRTGT GATGACBCARRCT	11

W=A/T, R=A/G, S=G/C, Y=C/T, K=G/T, M=A/C, H=ACT, B=GCT;

Conc. - concentração final.

Tabela 3: Iniciadores para PCR aninhada de Symplex de camundongo

Iniciador	Concentração final nM	Sequência	Número da Sequência
Conjunto mJH			
mJH1	200	GGAGGCGCTCGAGACGGT-GACCGTGGTCCC	12
mJH2	200	GGAGGCGCTCGAGACTGTGA-GAGTGGTGCC	13
mJH3	200	GGAGGCGCTCGAGACAGTGAC-CAGAGTCCC	14
mJH4	200	GGAGGCGCTCGAGACGGTGACT-GAGGTTCC	15
Conjunto mJK			
mJK1	200	GATGGTGCAGCCACAGTTCGTTT-GATTTCCAGCTTGGTG	16
mJK2	200	GATGGTGCAGCCACAGTTCGTTT-TATTTCCAGCTTGGTC	30
mJK4	200	GATGGTGCAGCCACAGTTCGTTT-TATTTCCAACCTTGTG	31
mJK5	200	GATGGTGCAGCCACAGTTCGTTT-CAGCTCCAGCTTGGTC	17

EXEMPLO 3. Clonagem de Anticorpos Anti-EGFR

Imunizações

- 5 Fêmeas de BALB/c, cepa A, ou camundongo C57B16 (8 a 10 semanas de vida) foram usados para imunizações por injeções com diferentes proteínas purificadas além das células superexpressando EGFR. Proteínas de EGFR disponíveis comercialmente (R&D systems, CAT número 1095-ER ou Sigma número E3641) foram usadas para algumas das imunizações. Para outras imunizações foram usados EGFR e EGFRvIII humano
- 10 recombinante produzidos como proteínas de fusão consistindo em ECD de

EGFR ou EGFRvIII e hormônio do crescimento humano (hGH), também incluindo um sítio de clivagem do Vírus de Gravação de Tabaco (TEV) além do marcador His. Em alguns casos, o ECD de EGFR foi isolado por clivagem de TEV-protease e subsequente purificação em uma coluna de níquel.

5 A linhagem de célula cancerígena de cabeça e pescoço humanos, HN5 (Easty DM, Easty GC, Carter RL, Monaghan P, Butler LJ. Br J Cancer., junho de 1981; 43(6):772-85. Dez linhagens de células de carcinoma humano derivadas de carcinomas escamosos de cabeça pescoço) expressando aproximadamente 10^7 receptores/célula foram usadas para imunizações com base em célula. As células foram cultivadas em meio DMEM suplementado com FBS a 10% (Soro Bovino Fetal), Glicerol a 3 mM, Piruvato de Sódio 5 mM e 1% de Penicilina Estreptomicina. Antes de cada imunização, as células foram lavadas em PBS, tripsinizadas com TrypLE e res-suspensas em meio de crescimento. Subsequentemente, as suspensões
10 celulares foram lavadas duas vezes em PBS por centrifugação em 250Xg, por 5 minutos, com desalojamento e ressuspensão em 15 mL de PBS estéril.

 Células ou antígeno foram diluídos em PBS e então misturados 1:1 com adjuvante de Freund. O adjuvante é usado para melhorar e modular a resposta imune. Para as primeiras imunizações, Adjuvante de Freund
20 Completo (CFA) foi usado, considerando-se que Adjuvante de Freund Incompleto (IFA) foi empregado para as imunizações subsequentes. IFA é uma emulsão de óleo em água composta de óleos minerais e CFA é IFA ao qual espécies de *Mycobacterium* secas, mortas por aquecimento foram adicionadas. Ambos os adjuvantes possuem um efeito de depósito. CFA fornece persistência de longo prazo à resposta imune e é usado para as primeiras
25 imunizações para reforçar a resposta imune e IFA é usado para imunizações subsequentes. As emulsões foram testadas por adição de uma gota na superfície de um vidro com água. Se a gota permanecesse como uma gota, a emulsão era estável e as injeções poderiam ser realizadas. Apenas emul-sões estáveis foram administradas aos camundongos. Dependendo da pro-gramação (vide Tabela 4), 25 a 100 μ g de antígeno ou 10^7 células foram u-sadas para cada injeção. No total, os camundongos receberam 4 injeções.

Todos os camundongos receberam injeções de 300 μ L ou 200 μ L de emulsão. Dependendo da programação, as injeções foram realizadas subcutaneamente (s.c), intraperitonealmente (i.p.) ou intravenosamente (i.v.).

5 No término, os camundongos foram sacrificados por deslocamento cervical e os baços foram removidos e transferidos para um coador de célula de 74 cm (Corning número 136350-3479). As células foram maceradas através do filtro, ressuspensas em RPMI frio 1640 com FBS a 10% e centrifugadas a 300Xg por 5 minutos. A microesfera celular foi ressuspensa em RPMI 1640 com FBS a 1%, filtrada através de um filtro de seringa de 50
10 μ m (BD número 340603) e coletada por filtração. O pélete celular foi criopreservado após ressuspensão em FCS com DMSO a 10% e as células congeladas armazenadas a 80% até classificação FACS.

Classificação FACS de células plasmáticas de murino

Ampolas com esplenócitos congelados foram descongeladas a
15 37°C e transferidas para tubo de 15 mL com gelo ainda presente. 10 mL de RPMI resfriado em gelo, FBS a 10% (soro bovino fetal) foram adicionados, gota a gota ao tubo enquanto agitando. Após uma lavagem em 10 mL de FACS PBS, 5 mL de FCS PBS são adicionados antes da filtração das células através de Filcon 50 μ m. As células foram então micronizadas e ressus-
20 pensas em 1 mL de PBS com FBS a 2% (volume final) e coloridas com anti-CD43-FITC e anti-D138-PE de acordo com a diluição específica a uma concentração final de aproximadamente 5 μ g/mL. As células foram incubadas a 4°C por 20 minutos no escuro. Subsequentemente, as células foram lavadas duas vezes com 2 mL de tampão FACS. Até 15 mL de FACS PBS foram
25 adicionados. Iodeto de propídio (PI) foi adicionado a 1:100 e as células foram subsequentemente classificadas em placas PCR de 96 poços contendo tampão de reação da PCR (vide abaixo) e rotacionadas a jusante por 2 minutos 400Xg antes das placas serem congeladas a -80°C. As células plasmáticas foram excluídas como CD43 positivo/CD-138 positivo.

30 *Ligação de pares cognatos V_H e V_L*

A ligação das sequências de codificação de V_H e V_L foi realizada nas células simples excluídas como células plasmáticas, facilitando o pare-

lhamento cognato das sequências de codificação de V_H e V_L . O procedimento utilizou um procedimento de PCR de duas etapas com base na

5 RT-PCR de sobreposição-extensão multiplexa de etapa única seguida por uma PCR aninhada. As misturas de iniciadores usadas no presente exemplo apenas ampliam as cadeias leve capa. Os iniciadores capazes de ampliação das cadeias leve lambda poderiam, contudo, se adicionados à mistura de iniciador multiplexa e mistura de iniciador da PCR aninhada, caso desejado. Se os iniciadores de lambda fossem adicionados, o procedimento de classificação seria adaptado, tal que, as células positivas
10 lambda não seriam excluídas. O princípio para ligação de sequências V_H e V_L cognatas é ilustrado na figura 1. As placas de PCR de 96 poços produzidas foram descongeladas e as células classificadas serviram como gabarito para a RT-PCR de sobreposição-extensão multiplexa. O tampão de classificação adicionado a cada poço antes da classificação da célula simples continha tampão de reação (Tampão RT-PCR de etapa única; Qiagen), iniciadores
15 para RT-PCR (vide tabela 2 acima) e inibidor de RNase (RNasina, Promega). Isto foi suplementado com Mistura de Enzima de RT-PCR de Etapa única (diluição de 25 vezes da Qiagen) e mistura dNTP (200 μ M cada) para obter a concentração final fornecida em um volume de reação de 20 μ L. As
20 placas foram incubadas por 30 minutos a 55°C para permitir transcrição reversa de RNA de cada célula. Seguindo-se a RT, as placas foram submetidas ao seguinte ciclo da PCR: 10 minutos a 94°C, 35 x (40 segundos a 94°C, 40 segundos a 60°C, 5 minutos a 72°C), 10 minutos a 72°C. As reações da PCR foram realizadas em ciclador térmico H20BIT com Cesta de
25 Vedação de Casca para placas de 24 a 96 poços (ABgene) para facilitar um alto rendimento. As placas PCR foram armazenadas a -20°C após ciclagem.

Para a etapa PCR aninhada, as placas PCR de 96 poços foram preparadas com a mistura que se segue em cada poço (reações de 20 μ L) para obter a concentração final fornecida: 1 x tampão FastStart (Roche),
30 mistura dNTP (200 μ M cada), mistura de iniciador aninhado (vide Tabela 5), DNA Polimerase Phusion (0,08 U; Finnzymes) e Combinação de Enzima FastStart High Fidelity (0,8 U; Roche). Como o gabarito para a PCR aninha-

da, foi transferido 1 μ L para as reações de PCR de sobreposição-extensão multiplexa. As placas de PCR aninhada foram submetidas à ciclagem térmica que se segue: 35 x (30 segundos a 95°C, 30 segundos a 60°C, 90 segundos a 72°C), 10 minutos a 72°C.

5 As reações selecionadas aleatoriamente foram analisadas em gel agarose a 1% para verificar a presença de fragmento de sobreposição-extensão de aproximadamente 890 bp (pares de base). As placas foram armazenadas a -20°C até processamento adicional dos fragmentos da PCR.

10 Os repertórios de pares de codificação de V_H e V_L ligados da PCR aninhada foram agrupados, sem misturar os pares de doadores diferentes e foram purificados por eletroforese de gel agarose a 1% de preparação. A sequência de codificação de cadeia leve constante capa humana foi recomposta por sobreposição-extensão para a região de codificação de V_L dos produtos da PCR agrupados dos pares de codificação de V_H e V_L (Figura 2). A sequência de codificação de cadeia leve constante capa humana foi ampliada de um plasmídeo contendo a sequência de codificação de um anticorpo humano com uma cadeia leve capa em uma reação contendo: Enzima Phusion (2 U; Finnzymes), 1 x tampão Phusion, mistura de dNTP (200 μ M cada), iniciador hKCforw-v2 e iniciador capa3' (Tabela 6), e gabarito de plasmídeo pLL138 (10 ng/ μ L) em um volume total de 50 μ L. A reação foi submetida à termociclagem que se segue: 25 x(30 segundos a 95°C, 30 segundos a 55°C, 45 segundos a 72°C), 10 minutos a 72°C. O fragmento de PCR resultante foi purificado por eletroforese de gel agarose a 1% de preparação.

25 Os fragmentos de PCR agrupados e purificados de cada repertório foram recompostos para o fragmento de PCR ampliado e purificado da região de codificação de constante capa humana (Apêndice 1) pela recomposição que se segue por PCR de sobreposição-extensão (50 μ L, volume total) contendo: fragmento de região de codificação de constante capa humana (1,4 ng/ μ L), fragmento de PCR agrupado e purificado (1,4 ng/ μ L), DNA Polimerase Phusion (0,5 U; Finnzymes) e Combinação de Enzima FastStart High Fidelity (0,2 U; Roche), 1x tampão FastStart (Roche), mistura de dNTP

30

(200 μ M cada), iniciador de mhKCreV e iniciadores de conjunto mJH (vide Tabela 6). A reação foi submetida à termociclagem que se segue: 2 minutos a 95°C, 25 x (30 segundos a 95°C, 30 segundos a 55°C, 1 minuto a 72°C), 10 minutos a 72°C. O fragmento de PCR resultante (aproximadamente 1.070 bp) foi purificado por eletroforese de gel agarose a 1% de preparação.

Inserção de pares de codificação V_H e V_L cognatos em um vetor de classificação

A fim de identificar os anticorpos com especificidade de ligação ao EGFR, as sequências de codificação de V_H e V_L obtidas foram expressas como anticorpos de comprimento pleno. Isto envolveu a inserção do repertório de pares de codificação V_H e V_L em um vetor de expressão e transfecção em uma célula hospedeira.

O procedimento de clonagem de duas etapas foi empregado para geração de um repertório de vetores de expressão contendo os pares de codificação V_H e V_L ligados. Estatisticamente, se o repertório dos vetores de expressão contiver dez vezes tantos plasmídeos recombinantes quanto o número de produtos de PCR de V_H e V_L emparelhados e cognatos usados para a geração do repertório de classificação, existirá uma probabilidade de 99% de que todos os pares gênicos únicos sejam representados. Assim, se 400 fragmentos gênicos de V de sobreposição-extensão forem obtidos, um repertório de pelo menos 4.000 clones será gerado para classificação.

Em resumo, o produto de PCR purificado dos repertórios de pares de codificação V_H e V_L ligados, recompostos na região de codificação constante capa humana foi clivado com DNA endonucleases de *XhoI* e *NotI* nos sítios de reconhecimento introduzidos nos terminos dos produtos de PCR. Os fragmentos clivados e purificados foram ligados em um vetor de expressão de IgG de mamífero digerido de *XhoI/NotI*, o OO-VP-002 (figura 6), por procedimentos de união-padrão. A mistura de ligação sofreu eletroforese em *E.coli* e foi adicionada às placas 2xYT contendo o antibiótico apropriado e incubada a 37°C por toda noite. O repertório ampliado de vetores foi purificado das células recuperadas das placas usando métodos de purificação de DNA-padrão (Qiagen). Os plasmídeos foram preparados por inserção

dos fragmentos promotor-líder por clivagem usando

Ascl e *NheI* endonucleases. Os sítios de restrição para estas enzimas estavam localizados entre os pares de genes de codificação V_H e V_L . Seguindo-se a purificação do vetor, um fragmento líder-promotor de mamífero bidirecional digerido *Ascl-NheI* foi inserido nos sítios de restrição *Ascl* e *NheI* por procedimentos de ligação-padrão. O vetor ligado foi ampliado em *E.coli* e o plasmídeo foi purificado usando métodos-padrão. O repertório gerado de vetores de classificação foi transformado no *E.coli* por procedimentos convencionais. As colônias obtidas foram consolidadas em placas principais de 384 poços e armazenadas. O número de colônias enfileiradas excedeu o número de produtos PCR que deram entrada em pelo menos três vezes, assim fornecendo uma probabilidade de 95% para presença de todos pares gênicos V únicos obtidos.

O plasmídeo do DNA foi preparado de clones selecionados acima e células CHO-S FreeStyle (Invitrogen) foram transfectadas em escala de 2 mL para expressão de anticorpos (de acordo com as instruções do fabricante). Os sobrenadantes foram colhidos 96 horas após transfecção.

Classificação para ligação ao domínio extracelular de EGFR

Em geral, a classificação foi realizada como um procedimento de duas etapas. As bibliotecas de anticorpo foram classificadas quanto à reatividade da proteína de EGFR recombinante em ELISA após o que FMAT (FLISA) foi empregado como uma abordagem com base em célula, com a linhagem de célula NR6wtEGFR (Batra e outros, 1995, Cell Growth Differ, 6(10):1251-9), para detecção de ligação dos anticorpos de EGFR ao EGFR expresso na superfície da célula. Para as bibliotecas 101 e 108/109 (Tabela 4) o ELISA foi realizado com EGFR recombinante representando o domínio extracelular de EGFR.

Em resumo, para o ELISA, placas Nunc maxisorb (número CAT 464718) foram revestidas com 1 $\mu\text{g/mL}$ de proteína (produzida domesticamente), diluída em PBS a 4°C por toda noite. Antes do bloqueio em 50 μL de 2%-Leite-PBS-T as placas foram lavadas uma vez com PBS + 0,05% de tween 20% (PBS-T). As placas foram lavadas uma vez com PBS-T, 20 μL de

2%-leite-PBS-T e 5 µL de sobrenadantes dos transfectantes CHO-S FreeStyle (vide acima) foram adicionados e incubados por 1 ½ hora TR após o que as placas foram lavadas uma vez com PBS-T a 20 µL por poço. Anticorpo secundário (HRP-IgG anti-humano de cabra, Jackson, número CAT 109-035-097) diluído 1:10.000 em 2% leite-PBS-T foi adicionado para detectar os anticorpos ligados aos poços e incubado por 1 hora a temperatura ambiente. As placas foram lavadas uma vez em PBS-T antes da adição de 25 µL de substrato (Kem-em-tec Diagnostics, número CAT 4390) que foi incubado por 5 minutos. 25 µL de ácido sulfúrico a 1M foram adicionados após a incubação para parar a reação. O sinal específico foi detectado em um leitor ELISA a 450 nm.

Para a detecção de FMAT com base na célula de anticorpos anti-EGFR, SKBR-3 (ATCC número HTB-30) ou NR6wtEGFR as células foram mantidas em meio de crescimento. As células foram contadas e diluídas para 125.000 células/mL com o anticorpo IgG anti-humano de cabra conjugado Alexa-647 (H-L) (Molecular probes número A21445, lote número 34686A) diluídas 1:40.000. Um total de 20 µL desta suspensão foram transferidos para placas Nunc de fundo limpo de 384 poços. subsequentemente, 10 µL de sobrenadante de transfecção foram adicionados às células. O sinal FMAT da reação foi medido após 6 a 10 horas de incubação.

Os dados da classificação indicam que 221 (4,8%) dos clones totais foram positivos no ELISA. 93 (2,0%) destes clones foram também positivos em FMAT. No total 220 (4,8%) dos clones foram positivos no FMAT e dentre aqueles 127 (220-93) unicamente positivos para o antígeno de superfície de célula. A biblioteca 111 foi classificada de um modo semelhante, porém uma vez que o procedimento de imunização foi realizado para gerar anticorpos específicos para a detecção de um mutante EGFR receptor EGFRvIII, as classificações de ELISA incluíram ensaios para detectar ambos EGFR do tipo selvagem e EGFRvIII. Sete clones foram identificados para serem específicos para o EGFRvIII no ELISA e de forma interessante aqueles clones foram negativos para coloração de células expressando wtEGFR no FMAT. Treze clones foram identificados como sendo positivos para wEGFR

em FMAT e ELISA, porém não para o EGFRvIII, que foram únicos para esta biblioteca, em comparação às bibliotecas 101 e 108/109. Todos os clones ELISA positivos foram selecionados para análises adicionais.

Análise da Sequência e Seleção do Clone

5 Os clones identificados como específicos de EGFR no ELISA foram recuperados das placas principais originais (formato de 384 poços) e consolidados em novas placas. DNA foi isolado dos clones e submetido ao sequenciamento de DNA dos genes V. As sequências foram alinhadas e todos os clones únicos foram selecionados. Alinhamentos múltiplos de se-
10 quências obtidas descreveram a singularidade de cada clone específico e permitiram a identificação de anticorpos únicos. Seguindo-se a análise da sequência de 220 clones, 70 grupos de sequência de anticorpo geneticamente distintos foram identificados. Cada grupo de sequências correlatas foi provavelmente derivado de hipermutações somáticas de um clone precursor
15 comum. No total, um a dois clones de cada grupo foi escolhido para validação da sequência e especificidade.

Sequência e Validação da Especificidade

A fim de validar os clones de codificação do anticorpo, o plasmídeo de DNA foi preparado e a transfecção de células CHO-S FreeStyle (Invitrogen) em escala de 2 mL foi realizada para expressão. O sobrenadante foi
20 colhido 96 horas após a transfecção. Os níveis de expressão foram estimados com ELISA anti-IgG padrão e a especificidade foi determinada por ELISA específico para EGFR-EGFRvIII, 85% dos clones mostraram possuir a especificidade e sequência corretas.

Classificação para Efeitos Antiproliferativos

25 A lesão celular pode resultar inevitavelmente na perda da capacidade da célula de manter e prover energia para função e crescimento da célula metabólica. Os ensaios de atividade metabólica se baseiam nesta premissa. Geralmente eles medem a atividade mitocondrial. O Reagente de Proliferação Celular WST-1 (Roche, número CAT 11 644 807 001) é um
30 substrato pronto para uso que mede a atividade metabólica das células viáveis. Então se presume que a atividade metabólica se correlacione ao núme-

ro de células viáveis. Neste exemplo, o ensaio WST-1 foi usado para medir o número de células metabolicamente ativas após tratamento com os sobrenadantes de cultura de célula contendo diferentes anticorpos anti-EGFR.

Antes da realização do ensaio WST-1, volumes diferentes de 2 mL de sobrenadantes (0, 10, 25, 50 e 150 μ L) foram transferidos para poços apropriados em uma placa de 96 poços.

Células HN5 foram então lavadas com 1xPBS e destacadas por tripsinação com 3 mL de solução de tripsina. 17 mL de meios completos foram então adicionados e as células rotacionadas a jusante a 300 x g (1.200 rcf) por 5 minutos. O sobrenadante foi removido e as células ressuspensas em DMEM + 0,5% FBS. As células foram contadas e sua concentração ajustada e 1.500 células foram adicionadas aos poços com sobrenadante, de modo que cada poço contivesse 200 μ L de meios no total. As placas foram incubadas por 4 dias em um incubador umidificado a 37°C. Então 20 μ L de reagente WST-1 foram adicionados por poço e as placas incubadas por uma hora a 37°C. As placas foram então transferidas para um agitador de placa orbital e deixadas mais uma hora. A absorvência foi medida a 450 e 620 nm (comprimento de onda de referência) em um leitor ELISA. A diferença nos níveis de células metabolicamente ativas (MAC) foi calculada como percentual dos sobrenadantes de controle, como se segue:

$$\%MAC = \left(1 - \frac{(OD_{exp.} - OD_{media})}{(OD_{untreat.} - OD_{media})} \right) \times 100$$

N.T.: Na formula acima: OD exp. - meios OD; OD não-tratado.

Estes valores foram então empregados como base para a análise de grupo hierárquica supervisionada (agrupados com base na reatividade no ELISA) realizada usando os softwares livres Cluster e TreeView.

É preferível ser capaz de se classificar anticorpos funcionais em um estágio anterior no processo de seleção de anticorpos. Os sobrenadantes de cultura de 83 transfeções de 2 mL foram usados para classificar as funções inibidores de crescimento em um ensaio de proliferação realizado usando células HN5 em FBS a 0,5%. Os resultados foram visualizados por análise de grupo hierárquica simples. Conforme pode ser visto na análise de

grupo (figura 7), foi verificado que vários sobrenadantes diminuem o número de células HN5 metabolicamente ativas (cinza-escuro) em um modo dependente de concentração (Grupo 2). De modo semelhante, alguns sobrenadantes aumentaram o número de células HN5 metabolicamente ativas (cinza claro) em um modo dependente da concentração (grupos 1, 3 e 4). Uma observação interessante foi a de que os sobrenadantes, que diminuíram o número de células HN5 metabolicamente ativas, possuíram reatividade 2 (setas pretas) considerando-se que os sobrenadantes que aumentaram o número de células HN5 metabolicamente ativas apresentaram reatividade 1 (setas cinzas). Os sobrenadantes com reatividade 2 foram positivos em ambos ELISAs wtEGFR e EGFRvIII, enquanto os sobrenadantes com reatividade 1 apresentaram reatividade apenas na direção de wtEGFR. Assim, tais análises podem prover relações entre reatividade do anticorpo em ELISA e funcionalidade nos ensaios celulares.

15 **Reparo do Clone**

Quando se emprega a abordagem de PCR multiplexa, se espera um determinado grau de iniciação cruzada da família gênica intra e inter-V devido à degeneração do iniciador e o alto grau de homologia. A iniciação cruzada introduz aminoácidos que não ocorrem naturalmente na estrutura da imunoglobulina com várias consequências em potencial, por exemplo, alterações estruturais e imunogenicidade aumentada, tudo resultando em uma atividade terapêutica diminuída.

A fim de eliminar estas desvantagens e assegurar que clones selecionados se igualem à resposta imune humoral natural, tais mutações de iniciação cruzada foram corrigidas em um processo denominado reparo de clone.

Na primeira etapa do procedimento de reparo de clone, a sequência V_H foi ampliada em PCR com um conjunto iniciador contendo a sequência correspondendo ao gene de V_H , o clone de interesse originado do mesmo, pelo que, corrigindo quaisquer mutações introduzidas por iniciação cruzada. O fragmento da PCR foi digerido com *XhoI* e *Ascl* e ligado de volta no vetor de expressão de mamífero digerido *XhoI/Ascl* (Figura 6) empregando procedimentos de ligação convencionais. O vetor ligado foi ampliado em

E.coli e o plasmídeo foi purificado por métodos-padrão. A sequência de V_H foi sequenciada para verificar a correção e o vetor foi digerido com *NheI/NotI* para preparar o mesmo para inserção na cadeia leve.

Na segunda etapa, a cadeia leve completa era PCR ampliada com um conjunto de iniciador contendo a sequência correspondendo ao gene de V_L, o clone de interesse originário do mesmo, pelo que, corrigindo quaisquer mutações introduzidas por iniciação cruzada. O fragmento de PCR foi digerido com *NheI/NotI* e ligado no vetor contendo V_H preparado acima. O produto de ligação foi ampliado em *E.coli* e o plasmídeo foi purificado por métodos-padrão. Subsequentemente, a cadeia leve foi sequenciada para verificar a correção.

No caso onde a região constante capa de um clone selecionado contém mutações introduzidas durante a ampliação dos genes, ela é substituída por uma região constante não-mutada. Isto é realizado em uma PCR de sobreposição onde o gene de V_L não-reparado (ampliado sem a região constante) foi fundido a uma região constante com sequência correta (obtida em uma PCR separada). A sequência integral é ampliada e clonada no vetor contendo V_H, conforme descrito acima e a cadeia leve reparada é sequenciada para verificar a correção.

Tabela 4 - Programações de imunização empregadas para gerar material de partida para clonagem anti-EGFR

Programação, Grupo de camundongos	Ce-pa	Injeção 1	Injeção 2	Injeção 3	Injeção 4	Término
101	Balb/ c	Dia 1 25 µg rhEGFR (R&D systems 1095- ER) CFA s.c.	Dia 35 25 µg rh- GH-EGFR (Sympho- gen) IFA s.c	Dia 56 25 µg rhEGFR* (Sympho- gen) IFA s.c	Dia 70 25 µg rhEGFR* (Sympho- gen) IFA s.c	Dia 73

Programação, Grupo de camundongos	Ce-pa	Injeção 1	Injeção 2	Injeção 3	Injeção 4	Término
108	Balb/ c	Dia 1 1x10 ⁷ células de HN5 CFA i.p.	Dia 28 25 µg rhEGFR* (Sympho- gen) IFA s.c.	Dia 42 1x10 ⁷ célu- las de HN5 IFA i.p.	Dia 56 25 µg rhEGFR*, (Sympho- gen) IFA s.c.	Dia 59
109	Balb/ c	Dia 1 1x10 ⁷ células de HN5 CFA i.p.	Dia 28 25 µg rhEGFR* (Sympho- gen) IFA s.c.	Dia 42 1x10 ⁷ célu- las de HN5 IFA i.p.	Dia 56 25 µg rhEGFR* (Sympho- gen) PBS i.v.	Dia 59
111	Balb/ c	Dia 1 25 µg rhEG- FR* (Sym- phogen) CFA s.c.	Dia 28 25 µg rhEGFR+ rhEGFR- vIII** (Sympho- gen) IFA s.c.	Dia 42 25 µg rhEGFR+ rhEGFRvII- I** (Sym- phogen) IFA s.c.	Dia 56 25 µg rhEGFR+ rhEGFRvII- I** (Sym- phogen) IFA s.c.	Dia 59
118	Balb/ c	Dia 1 1x10 ⁷ células de HN5 CFA i.p.	Dia 29 100 µg rhGH- EGFR (Sympho- gen) IFA s.c.	Dia 44 1x10 ⁷ célu- las de HN5 IFA i.p.	Dia 58 25 µg rhEGFR, (Sigma E3641) IFA s.c.	Dia 61
119	C57 B	Dia 1 1x10 ⁷ células de HN5 CFA i.p.	Dia 29 100 µg rhGH- EGFR (Sympho- gen) IFA s.c.	Dia 44 1x10 ⁷ célu- las HN5 IFA i.p.	Dia 58 25 µg rhEGFR, (Sigma E3641) IFA s.c.	Dia 61

Tabela 5 - Conjunto de iniciador aninhado

Denominação do iniciador	Conc. (nM)	Sequência	Número da Sequência
mHCrev1-	0.2	GGACAGGGMTCCAAGTTCCADKT	18
Conjunto hmJK			
hmJK1-v2	0.2	GACAGATGGTGCAGCCACAGTTCGTTT-	19
hmJK2-v2	0.2	GACAGATGGTGCAGCCACAGTTCGTTT-	20
hmJK4-v2	0.2	GACAGATGGTGCAGCCACAGTTCGTTT-	21
hmJK5-v2	0.2	GACAGATGGTGCAGCCACAGTTCGTTT-	22

K=G/T, M=A/C, D=AGT; Conc. - concentração final.

Tabela 6 - Conjunto de iniciador de recomposição de constante capa

Iniciador	Conc. (nM)	Sequência	Número sequência
Ampliação da constante capa humana			
hKC-	0.2	GAAGTGTGGCTGCACCATCTGTC	23
Capa3'	0.2	ACCGCCTCCACCGGCGGCCGCTTATTAA-	24
Recomposição por sobreposição-extensão			
mhK-	0.2	ACCGCCTCCACCGGCGGCCGCTTATTAACACTCT	25
mJH			
mJH1	0.2	GGAGGCGCTCGAGACGGTGACCGTGGTCCC	12
mJH2	0.2	GGAGGCGCTCGAGACTGTGAGAGTGGTGCC	13
mJH3	0.2	GGAGGCGCTCGAGACAGTGACCAGAGTCCC	14
mJH4	0.2	GGAGGCGCTCGAGACGGTGACTGAGGTTCC	15

5 Apêndice 1, Sequências de região constante de anticorpo

>Região IGKC humana (Sequência número 26)

ttcatctcccgccatctgatgagcagttgaaatctggaactgcctctgtgtgtgcctgctgaataact
 tctatcccagagaggccaaagtacagtggaaggtggataacgccctccaatcgggtaactcccaggagag
 tgtcacagagcaggacagcaaggacagcacctacagcctcagcagcacctgacgctgagcaaagcagac
 10 tacgagaaacacaaagtctacgcctgcgaagtcacccatcagggcctgagctgcccgtcacaaagagct
 tcaacaggggagagtgttaataagcggccgctggaggcgggt

>Região IGKC humana (Sequência número 27)

TVAAPSVFIFPPSDEQLKSGTASVCLLNNFYPREAKVQWKVDNALQSGNS

QESVTEQDS

KDSTYSLSSTLTLSKADYKHKVYACEVTHQGLSSPVTKSFNRGEC

Exon1 1..298

Intron 299..689

5 Exon2 690..734

Intron 735..852

Exon3 853..1182

Intron 1183..1279

Exon4 1280..1602

10 >Sequência genômica de domínio constante de IGHG1 humana (Sequência número 28)

agtcctccaccaagggcccatcggtctccccctggcaccctctccaagagcacctctgggggcacag

cggccctgggctgctggtcaaggactactccccgaacoggtgacggtgctggaactcaggcgcct

gaccagcggcgtgcacacctcccggctgtctacagtctcaggactctacccctcagcagcgtggg

15 accgtgccctccagcagctgggcacccagacctacatctgcaacgtgaatcacaagcccagcaacacca

aggtggacaagagagttggtgagaggccagcacagggagggagggtgctgctggaagccaggctcagcg

ctctgctgagcgcacatcccggctatgcagctccagtcaggcagcaaggcaggccccgtctgccttt

cacccggaggcctctgcccgccccactcatgctcaggagaggggtctctggctttccccaggctctg

ggcaggcacaggctaggtgcccctaaccaggccctgcacacaaaggggcaggtgctgggctcagacctg

20 ccaagagccatatccgggaggaccctgcccctgacctaagcccaccccaaaggccaaactctccactccc

tcagctcggacacctctctcctcccagattccagtaactcccaatcttctctctgagagcccaaatct

tgtgacaaaactcacacatgcccaccgtgcccaggtaagccagcccaggcctcgccctccagctcaaggc

gggacagggtccctagagtagcctgcatccaggacaggccccagccgggtgctgacacgtccacctcca

tctctctcagcacctgaaactctggggggaccgtcagctctctcttcccccaaaacccaaggacac

25 cctcatgatctccggaccctgaggtcacatgcgtggtgggacgtgagccacgaagaccctgaggtc

aagtcaactggtacgtggacggcgtggaggtgcataatgccaagacaaagccgcccggaggagcagtaca

acagcacgtaccgtggtcagcgtcctcaccgtcctgcaccaggactggctgaatggcaaggagtacaa

gtgcaaggtctcaacaaagccctcccagccccatcgagaaaacctctcaaagccaaaggtgggacc

cgtgggggtgcgagggccacatggacagaggccggctcggcccaccctctgcctgagagtgaccgtgta

30 ccaacctctgtccctacagggcagccccgagaaccacaggtgtacacctgccccatcccgggaggaga

tgaccaagaaccaggctcagcctgacctgctggtcaaaggcttctatcccagcgacatcgccgtggagtg

ggagagcaatgggcagccggagaacaactacaagaccacgcctcccgtgctggactccgacggctcctc

ttcctctatagcaagctcacogtggacaagagcaggtggcagcaggggaacgtcttctcatgctccgtga

tgcatgaggctctgcacaaccactacacgcagaagagcctctccctgtccccgggtaaata

>IGHG1 (Sequência número 29)

SASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALTSG

5 VHTFPAVLQSSGLYSLSSVVT

VPSSSLGTQTYICNVNHKPSNTKVDKRVEPKSCDKTHTCPPCPAPELLGGP

SVFLFPPKPKDTLMISRTP

EVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLT

VLHQDWLNGKEYKCKVSNKA

10 LPAPIEKTISKAKGQPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAV

EWESNGQPENNYKTTTPV

LDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK

LISTAGEM DE SEQUÊNCIAS P 160899

Listagem de Sequências

15 <110> Symphogen AS

Meijer, Per Johan

Kastrup, Jesper

Nielsen, Lars

<120> Método para Clonagem de Anticorpos Cognatos

20 <130> P112PC00

<160> 33

<170> PatentIn versão 3.4

<210> 1

<211> 20

25 <212> DNA

<213> artificial

<220>

<223> iniciador da PCR

<400> 1

30 gacsgatggg cccttggtgg

<210> 2

<211> 20

<212> DNA
 <213> artificial
 <220>
 <223> iniciador da PCR
 5 <400> 2
 gctgtaggtg ctgtctttgc 20
 <210> 3
 <211> 39
 <212> DNA
 10 <213> artificial
 <220>
 <223> iniciador da PCR
 <400> 3
 tattcccatg gcgcgccsag gtccarctgc arcagycytg 39
 15 <210> 4
 <211> 37
 <212> DNA
 <213> artificial
 <220>
 20 <223> iniciador da PCR
 <400> 4
 tattcccatg gcgcgcccgar gtgmagctkg tkgagtc 37
 <210> 5
 <211> 37
 25 <212> DNA
 <213> artificial
 <220>
 <223> iniciador da PCR
 <400> 5
 30 tattcccatg gcgcgccsag gtgcagctkm aggagtc 37
 <210> 6
 <211> 37

<212> DNA
 <213> artificial
 <220>
 <223> iniciador da PCR
 5 <400> 6
 tattcccatg gcgcgccag gttactctga aagagtc 37
 <210> 7
 <211> 39
 <212> DNA
 10 <213> artificial
 <220>
 <223> iniciador da PCR
 <400> 7
 tattcccatg gcgcgccag atccagttgg tgcagtctg 39
 15 <210> 8
 <211> 45
 <212> DNA
 <213> artificial
 <220>
 20 <223> iniciador da PCR
 <400> 8
 ggcgcccat ggaatagct agccgayatc cagatgachc arwct 45
 <210> 9
 <211> 44
 25 <212> DNA
 <213> artificial
 <220>
 <223> iniciador da PCR
 <400> 9
 30 ggcgcccat ggaatagct agccracatt gtgmtgachc agtc 44
 <210> 10
 <211> 46

<212> DNA
 <213> artificial
 <220>
 <223> iniciador da PCR
 5 <400> 10
 ggcgcgccat ggaatagct agccsamatt gtkotsacce artctc 46
 <210> 11
 <211> 45
 <212> DNA
 10 <213> artificial
 <220>
 <223> iniciador da PCR
 <400> 11
 ggcgcgccat ggaatagct agccgatrtrt gtgatgacbc arrct 45
 15 <210> 12
 <211> 30
 <212> DNA
 <213> artificial
 <220>
 20 <223> iniciador da PCT
 <400> 12
 ggagggcgctc gagacgggtga ccgtgggtccc 30
 <210> 13
 <211> 30
 25 <212> DNA
 <213> artificial
 <220>
 <223> iniciador da PCR
 <400> 13
 30 ggagggcgctc gagactgtga gagtgggtgcc 30
 <210> 14
 <211> 30

<212> DNA
 <213> artificial
 <220>
 <223> iniciador da PCR
 5 <400> 14
 ggagggcgtc gagacagtga ccagagtccc 30
 <210> 15
 <211> 30
 <212> DNA
 10 <213> artificial
 <220>
 <223> iniciador da PCT
 <400> 15
 ggagggcgtc gagacggtga ctgaggttcc 30
 15 <210> 16
 <211> 39
 <212> DNA
 <213> artificial
 <220>
 20 <223> iniciador da PCR
 <400> 16
 gatggtgcag ccacagttcg tttgatttcc agcttgggtg 39
 <210> 17
 <211> 39
 25 <212> DNA
 <213> artificial
 <220>
 <223> iniciador da PCT
 <400> 17
 30 gatggtgcag ccacagttcg tttcagctcc agcttgggtc 39
 <210> 18
 <211> 24

<212> DNA
 <213> artificial
 <220>
 <223> iniciador da PCT
 5 <400> 18
 ggacagggmt ccakagttcc adkt 24
 <210> 19
 <211> 43
 <212> DNA
 10 <213> artificial
 <220>
 <223> iniciador da PCR
 <400> 19
 gacagatggt gcagccacag ttcgtttgat ttccagcttg gtg 43
 15 <210> 20
 <211> 43
 <212> DNA
 <213> artificial
 <220>
 20 <223> iniciador da PCR
 <400> 20
 gacagatggt gcagccacag ttcgttttat ttccagcttg gtc 43
 <210> 21
 <211> 43
 25 <212> DNA
 <213> artificial
 <220>
 <223> iniciador da PCR
 <400> 21
 30 gacagatggt gcagccacag ttcgttttat ttccaacttt gtc 43
 <210> 22
 <211> 43

<212> DNA
 <213> artificial
 <220>
 <223> iniciador da PCR
 5 <400> 22
 gacagatggt gcagccacag ttcgtttcag ctccagcttg gtc 43
 <210> 23
 <211> 23
 <212> DNA
 10 <213> artificial
 <220>
 <223> iniciador da PCR
 <400> 23
 gaactgtggc tgcaccatct gtc 23
 15 <210> 24
 <211> 43
 <212> DNA
 <213> artificial
 <220>
 20 <223> iniciador da PCR
 <400> 24
 accgcctcca ccggcgcccg cttattaaca ctctcccctg ttg 43
 <210> 25
 <211> 51
 25 <212> DNA
 <213> artificial
 <220>
 <223> iniciador da PCR
 <400> 25
 30 accgcctcca ccggcgcccg cttattaaca ctctcccctg ttgaagctct t 51
 <210> 26
 <211> 324

<212> DNA

<213> artificial

<220>

<223> iniciador da PCR

5 <400> 26

```

ttcatcttcc cgccatctga tgagcagttg aaatctggaa ctgcctctgt tgtgtgcctg      60
ctgaataact tctatcccag agaggccaaa gtacagtgga aggtggataa cgcctcccaa      120
tcgggtaact cccaggagag tgtcacagag caggacagca aggacagcac ctacagcctc      180
agcagcaccg tgacgctgag caaagcagac tacgagaaac acaaagtcta cgcctgcgaa      240
10 gtcacccatc agggcctgag ctogcccgtc acaaagagct tcaacagggg agagtgttaa      300
taagcggccg ccggtggagg cggt                                             324

```

<210> 27

<211> 106

<212> PRT

15 <213> Homo sapiens

<400> 27

```

Thr Val Ala Ala Pro Ser Val Phe Ile Phe Pro Pro Ser Asp Glu Gln
1           5           10           15
Leu Lys Ser Gly Thr Ala Ser Val Val Cys Leu Leu Asn Asn Phe Tyr
20           20           25           30
Pro Arg Glu Ala Lys Val Gln Trp Lys Val Asp Asn Ala Leu Gln Ser
           35           40           45
Gly Asn Ser Gln Glu Ser Val Thr Glu Gln Asp Ser Lys Asp Ser Thr
           50           55           60
25 Tyr Ser Leu Ser Ser Thr Leu Thr Leu Ser Lys Ala Asp Tyr Glu Lys
           65           70           75           80
His Lys Val Tyr Ala Cys Glu Val Thr His Gln Gly Leu Ser Ser Pro
           85           90           95
Val Thr Lys Ser Phe Asn Arg Gly Glu Cys
30           100           105

```

<210> 28

<211> 1602

<212> DNA

<213> artificial

<220>

<223> iniciador da PCR

5 <400> 28

```

agtgccctcca ccaagggccc atcgggtcttc cccttgccac cctcctccaa gagcacctct    60
gggggcacag cggccctggg ctgcctggtc aaggactact tccccgaacc ggtgacggtg    120
tcgtggaact caggcgcctt gaccagcggc gtgcacacct tcccggctgt cctacagtcc    180
tcaggactct actccctcag cagcgtggtg accgtgcctt ccagcagctt gggcacccag    240
10 acctacatct gcaacgtgaa tcacaagccc agcaacacca aggtggacaa gagagtgggt    300
gagaggccag cacagggagg gaggggtgtct gctggaagcc aggctcagcg ctctgcctg    360
gacgcatccc ggctatgcag tccagtgcca gggcagcaag gcaggccccg tctgctctt    420
caccgggagg cctctgcccc cccactcat gctcagggag agggctctct ggctttttcc    480
ccaggctctg ggcaggcaca ggctaggtgc ccctaaccga ggccctgcac acaaaggggc    540
15 aggtgctggg ctgagacctg ccaagagcca tatccgggag gaccctgccc ctgacctaaag    600
cccaccccaa aggccaaact ctccactccc tcagctcggg caccttctct cctcccagat    660
tccagtaact cccaatcttc tctctgcaga gcccaaatct tgtgacaaaa ctcacacatg    720
cccaccgtgc ccaggtaagc cagcccaggc ctgcctctcc agctcaaggc gggacaggtg    780
ccctagagta gcctgcatcc aggacagggc ccagccggg tgctgacacg tccacctcca    840
20 tctcttctc agcacctgaa ctctggggg gaccgtcagt ctctctctc cccccaaaac    900
ccaaggacac cctcatgatc tcccggacct ctgaggtcac atgcgtggtg gtggacgtga    960
gccacgaaga cctgaggtc aagttcaact ggtacgtgga cggcgtggag gtgcataatg    1020
ccaagacaaa gccgcgggag gagcagtaca acagcacgta ccgtgtggtc agcgtcctca    1080
ccgtcctgca ccaggactgg ctgaatggca aggagtacaa gtgcaaggtc tccaacaaag    1140
25 ccctcccagc ccccatcgag aaaacatct ccaaagccaa aggtgggacc cgtgggggtgc    1200
gagggccaca tggacagagg ccggctcggc ccaccctctg ccctgagagt gaccgctgta    1260
ccaacctctg tcctacagc gcagccccga gaaccacagg tgtacacct gcccccatcc    1320
cgggaggaga tgaccaagaa ccaggtcagc ctgacctgcc tggtaaagg cttctatccc    1380
agcgacatcg ccgtggagtg ggagagcaat gggcagccgg agaacaacta caagaccacg    1440
30 cctcccgtgc tggactccga cggctcctc ttctctata gcaagctcac cgtggacaag    1500
agcaggtggc agcaggggaa cgtcttctca tgctccgtga tgcatgaggc tctgcacaac    1560
cactacacgc agaagagcct ctccctgtcc ccgggtaaat ga    1602

```

<210> 29

<211> 331

<212> PRT

<213> Homo sapiens

5 <400> 29

Ser Ala Ser Thr Lys Gly Pro Ser Val Phe Pro Leu Ala Pro Ser Ser

1 5 10 15

Lys Ser Thr Ser Gly Gly Thr Ala Ala Leu Gly Cys Leu Val Lys Asp

20 25 30

10 Tyr Phe Pro Glu Pro Val Thr Val Ser Trp Asn Ser Gly Ala Leu Thr

35 40 45

Ser Gly Val His Thr Phe Pro Ala Val Leu Gln Ser Ser Gly Leu Tyr

50 55 60

Ser Leu Ser Ser Val Val Thr Val Pro Ser Ser Ser Leu Gly Thr Gln

15 65 70 75 80

Thr Tyr Ile Cys Asn Val Asn His Lys Pro Ser Asn Thr Lys Val Asp

85 90 95

Lys Arg Val Glu Pro Lys Ser Cys Asp Lys Thr His Thr Cys Pro Pro

100 105 110

20 Cys Pro Ala Pro Glu Leu Leu Gly Gly Pro Ser Val Phe Leu Phe Pro

115 120 125

Pro Lys Pro Lys Asp Thr Leu Met Ile Ser Arg Thr Pro Glu Val Thr

130 135 140

Cys Val Val Val Asp Val Ser His Glu Asp Pro Glu Val Lys Phe Asn

25 145 150 155 160

Trp Tyr Val Asp Gly Val Glu Val His Asn Ala Lys Thr Lys Pro Arg

165 170 175

Glu Glu Gln Tyr Asn Ser Thr Tyr Arg Val Val Ser Val Leu Thr Val

180 185 190

30 Leu His Gln Asp Trp Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser

195 200 205

Asn Lys Ala Leu Pro Ala Pro Ile Glu Lys Thr Ile Ser Lys Ala Lys

<211> 8
<212> DNA
<213> artificial
<220>
5 <223> Sequência de Referência
<400> 32
agtcagtc 8
<210> 33
<211> 11
10 <212> DNA
<213> artificial
<220>
<223> Sequencia de referência
<400> 33
15 taatcaatcg g 11

REIVINDICAÇÕES

1. Método de produção de uma biblioteca de pares cognatos compreendendo sequências ligadas de codificação de região variável, dito método compreendendo:

5 a) provisão de uma fração celular compreendendo linfócito de um doador;

b) obtenção de uma população de células simples isoladas, compreendendo distribuição de células da dita fração celular individualmente em vários recipientes, em que pelo menos uma subpopulação das células expressa antígenos CD43 e CD138 ou antígenos MHCII e B220; e

10 c) ampliação e realização da ligação das sequências de codificação de região variável contidas na dita população das células simples isoladas por ampliação, em um procedimento de ampliação molecular multiplexo, sequências de nucleotídeo de interesse empregando um gabarito derivado de uma célula simples isolada ou uma população de células isogênicas; e

15 efetuando ligação das sequências de nucleotídeo de interesse ampliadas.

2. Método de acordo com a reivindicação 1, em que a célula simples isolada, individual na população das células simples é expandida para uma população de células isogênicas, antes da realização da ampliação e

20 ligação (etapa c)).

3. Método de acordo com qualquer uma das reivindicações precedentes, em que a fração celular contendo linfócito compreende esplenócitos, sangue integral, medula óssea, células mononucleares ou leucócitos.

4. Método de acordo com a reivindicação 1, em que a fração celular contendo linfócito compreende esplenócitos ou medula óssea, preferivelmente esplenócitos.

25

5. Método de acordo com qualquer uma das reivindicações precedentes, em que a fração celular contendo linfócito ou linhagem de linfócito B é enriquecida para as células plasmáticas ou plasma blastos.

30 6. Método de acordo com qualquer uma das reivindicações precedentes, em que as sequências de nucleotídeo de interesse compreendem região variável codificando sequências de imunoglobulina e a ligação gera

um par cognato de uma sequência de codificação de região variável de cadeia leve associada à sequência de codificação de região variável de cadeia pesada.

5 7. Método de ligação aleatória de várias sequências de nucleotídeo de interesse descontínuas, dito método compreendendo:

a) ampliação, em um procedimento de ampliação molecular multiplexo, de sequências de nucleotídeo de interesse empregando um gabarito derivado de uma população de células geneticamente diversas;

10 b) em que as células geneticamente diversas são derivadas de uma fração celular compreendendo linfócito de um doador;

c) em que pelo menos uma subpopulação das células expressa antígenos CD43 e CD138 ou antígenos MHCII e B220; e

d) efetuando a ligação das sequências de nucleotídeo de interesse ampliadas na etapa a).

15 8. Método de acordo com a reivindicação 7, em que a dita população de células é lisada.

9. Método de acordo com a reivindicação 7 ou 8, em que as ditas sequências de nucleotídeo de interesse compreendem sequências de codificação de região variável e a ligação gera uma biblioteca combinatória de pares de sequências de codificação de região variável.

20 10. Método de acordo com a reivindicação 9, em que as ditas sequências de nucleotídeo de interesse compreendem região variável codificando sequências de imunoglobulina e a ligação gera uma biblioteca combinatória de pares de região variável de cadeia leve e sequências de codificação de região variável de cadeia pesada.

25 11. Método de acordo com qualquer reivindicação precedente, compreendendo, adicionalmente, avaliação antes da ampliação molecular multiplexa se a população de células compreendendo linfócito consiste em células expressa antígenos CD43 e CD138 ou antígenos MHCII e B220, preferivelmente CD43 e CD138.

30 12. Método de acordo com qualquer uma das reivindicações precedentes, compreendendo, adicionalmente, enriquecimento da dita fração

celular compreendendo linfócito para uma população de linfócito expressando antígenos CD43 e CD138 ou MHCII e B220 antes da ampliação molecular multiplexa.

5 13. Método de acordo com qualquer uma das reivindicações precedentes, compreendendo, adicionalmente, isolamento da dita população de células compreendendo linfócito expressando antígenos CD43 e CD138 ou antígenos MHCII e B220 antes da ampliação molecular multiplexa.

10 14. Método de acordo com qualquer uma das reivindicações precedentes, em que as células isoladas ou subpopulação de células se constitui em CD138 alto/CD43 alto ou CD138 intermediário/CD43 alto em relação à fração celular compreendendo linfócito.

15 15. Método de acordo com a reivindicação 14, em que as células isoladas ou subpopulação de células se constitui em CD138 alto/CD43 alto em relação à fração celular compreendendo linfócito.

16 16. Método de acordo com qualquer uma das reivindicações precedentes, em que o enriquecimento ou isolamento compreende um procedimento de classificação automatizado.

17 17. Método de acordo com a reivindicação 16, em que o procedimento de classificação automatizado é MACS ou FACS.

20 18. Método de acordo com qualquer uma das reivindicações precedentes, em que as células são derivadas de um mamífero.

19. Método de acordo com a reivindicação 18, em que o mamífero é não-humano.

25 20. Método de acordo com a reivindicação 19, em que o mamífero não-humano é transgênico e expressa imunoglobulinas humanas.

21. Método de acordo com a reivindicação 18, em que o doador é um roedor.

22. Método de acordo com a reivindicação 21, em que o doador é um camundongo.

30 23. Método de acordo com a reivindicação 18, em que o doador é um primata.

24. Método de acordo com a reivindicação 18, em que o doador é

um coelho.

25. Método de acordo com qualquer uma das reivindicações precedentes, em que o dito procedimento de ampliação molecular multiplexo é uma ampliação de RT-PCR multiplexa.

5 26. Método de acordo com a reivindicação 25, em que a dita ampliação de RT-PCR multiplexa é um processo de duas etapas compreendendo uma etapa separada de transcrição reversa (RT) antes da ampliação da PCR multiplexa.

10 27. Método de acordo com a reivindicação 25, em que a dita ampliação de RT-PCR multiplexa é realizada em uma etapa única compreendendo, inicialmente, adição de todos os componentes necessários para realizar ambas a transcrição reversa (RT) e ampliação da PCR multiplexa em um recipiente simples.

15 28. Método de acordo com qualquer uma das reivindicações precedentes, em que a dita ligação das sequências de nucleotídeo de interesse é realizada no mesmo recipiente como a ampliação molecular multiplexa.

20 29. Método de acordo com qualquer uma das reivindicações 25 a 28, em que a dita ligação das sequências de nucleotídeo de interesse é efetuada em associação com a ampliação da PCR multiplexa, utilizando a mistura de iniciador de sobreposição-extensão multiplexa.

30. Método de acordo com qualquer uma das reivindicações precedentes, em que a dita ligação das sequências de nucleotídeo de interesse é efetuada por união.

25 31. Método de acordo com qualquer uma das reivindicações precedentes, em que é realizada uma ampliação molecular adicional utilizando uma mistura de iniciador adaptada para ampliação das sequências de ácido nucleico de interesse ligadas.

30 32. Método de acordo com a reivindicação 29, em que a mistura de iniciador de sobreposição-extensão multiplexa compreende conjunto de iniciador, em que pelo menos um elemento do conjunto de iniciador de cada conjunto de iniciador compreende uma cauda de sobreposição-extensão capaz de hibridizar com relação à cauda de sobreposição-extensão de um

elemento do conjunto de iniciador de um segundo conjunto de iniciador.

33. Método de acordo com a reivindicação 29 ou 32, em que a mistura de iniciador de sobreposição-extensão multiplexa compreende:

5 a) pelo menos um iniciador mKappar1 ou hmJK complementar ao filamento de sentido de uma sequência de codificação de região de cadeia leve de imunoglobulina;

b) pelo menos um iniciador mVK complementar ao filamento de antissenso de uma sequência de codificação de região variável de cadeia leve de imunoglobulina ou sequência líder de região variável de cadeia leve e capaz de formar um conjunto de iniciador com o(s) iniciador(es) na etapa a);

c) pelo menos um iniciador mCHrev1, mHCrev1-ext, ou mJH complementar ao filamento de sentido de uma sequência de codificação de domínio de cadeia pesada de imunoglobulina; e

15 d) pelo menos um iniciador mVH complementar ao fita de antissenso de uma sequência de codificação de região variável de cadeia pesada de imunoglobulina ou sequência líder de região variável de cadeia pesada, e capaz de formar um conjunto de iniciador com o(s) iniciador(es) na etapa c).

20 34. Método de acordo com qualquer uma das reivindicações precedentes, compreendendo, adicionalmente, inserção das sequências de nucleotídeo ligadas ou uma biblioteca de pares cognatos em um vetor.

35. Método de acordo com a reivindicação 34, em que o dito vetor é selecionado dentre vetores de clonagem, vetores de transporte bidirecional, vetores de exibição ou vetores de expressão.

25 36. Método de acordo com a reivindicação 34 ou 35, em que as sequências de nucleotídeo ligadas ou os elementos individuais da biblioteca de pares cognatos compreendem uma sequência de codificação de região variável de cadeia pesada de imunoglobulina associada a uma sequência de codificação de região variável de cadeia leve e ditas sequências são inseridas na estrutura em um vetor já contendo sequências codificando um ou

30 mais domínios constantes de imunoglobulina ou fragmentos dos mesmos.

37. Método de acordo com qualquer uma das reivindicações 34 a

36, compreendendo, adicionalmente, criação de uma sub-biblioteca por seleção de um subconjunto de pares cognatos de sequências ligadas de região variável que codificam proteínas de ligação com especificidade-alvo desejada, gerando uma biblioteca de pares cognatos específicos alvo de sequências de codificação de região variável.

38. Método de acordo com reivindicação 36 a 37, compreendendo, adicionalmente, transferência do dito par cognato ou biblioteca de pares cognatos específicos alvo de sequências de codificação de região variável para um vetor de expressão de mamífero.

39. Método de acordo com a reivindicação 38, em que o vetor de expressão de mamífero codifica um ou mais domínios de região constante selecionados das classes de imunoglobulina humana IgA1, IgA2, IgD, IgE, IgG1, IgG2, IgG3, IgG4, IgM, cadeia leve capa ou cadeia leve lambda.

40. Método de acordo com qualquer uma das reivindicações 34 a 39, compreendendo, adicionalmente, as etapas:

a) introdução de um vetor codificando um segmento de sequências de nucleotídeo ligadas em uma célula hospedeira;

b) cultivo das ditas células hospedeiras sob condições adaptadas para expressão; e

c) obtenção do produto da proteína expresso do vetor inserido na dita célula hospedeira.

41. Método de acordo com a reivindicação 40, em que dito produto de proteína é um anticorpo compreendendo um par cognato de uma região variável de cadeia leve associada com à região variável de cadeia pesada.

42. Placa de múltiplos poços compreendendo na maioria dos poços,

- uma célula derivada de uma fração celular compreendendo linfócito, dita célula expressando antígenos CD43 e CD138 ou antígenos MHCII e B220, e

- tampões e reagentes necessários para realizar transcrição reversa do RNAm e para ampliação das regiões de codificação variável de

cadeia pesada e leve.

43. Método para geração de um vetor codificando um anticorpo quimérico com regiões constantes humanas e regiões variáveis não-humanas, dito método compreendendo:

5 a) provisão de uma fração celular compreendendo linfócito de um animal diferente do ser humano;

b) obtenção de uma população de células simples isoladas, compreendendo distribuição das células da dita fração celular individualmente em vários recipientes;

10 c) ampliação e realização da ligação da região variável codificando ácidos nucleicos contidos na dita população de células simples isoladas por ampliação, em um procedimento de ampliação molecular multiplexo, ditos ácidos nucleicos empregando um gabarito derivado de uma célula simples isolada ou uma população de células isogênicas; e efetuando ligação
15 dos ácidos nucleicos ampliados codificando regiões variáveis de cadeias pesadas e leves;

d) efetuando a ligação das regiões variáveis ampliadas às regiões constantes humanas; e

e) inserção do ácido nucleico obtido em um vetor.

20 44. Método de acordo com a reivindicação 43, em que o dito procedimento de ampliação molecular multiplexo é uma ampliação de RT-PCR multiplexa.

45. Método de acordo com a reivindicação 44, em que a dita ampliação de RT-PCR multiplexa é um processo de duas etapas compreendendo uma etapa separada de transcrição reversa (RT) antes da ampliação da PCR multiplexa.
25

46. Método de acordo com a reivindicação 44, em que a dita ampliação de RT-PCR multiplexa é realizada em uma etapa única compreendendo inicialmente adição de todos os componentes necessários para realizar ambas a transcrição reversa (RT) e ampliação da PCR multiplexa em um
30 recipiente simples.

47. Método de acordo com qualquer uma das reivindicações pre-

cedentes 43 a 46, em que a dita ligação das sequências de nucleotídeo de interesse é realizada no mesmo recipiente como a ampliação molecular multiplexa.

5 48. Método de acordo com qualquer uma das reivindicações 44 a 47, em que a dita ligação das sequências de nucleotídeo de interesse é efetuada em associação à ampliação da PCR multiplexa, utilizando a mistura de iniciador de sobreposição-extensão multiplexa.

10 49. Método de acordo com qualquer uma das reivindicações precedentes 43 a 48, em que a dita ligação das sequências de nucleotídeo de interesse é efetuada por união.

50. Método de acordo com qualquer uma das reivindicações precedentes 43 a 49, em que é realizada uma ampliação molecular adicional, utilizando uma mistura de iniciador adaptada para ampliação das sequências de ácido nucleico de interesse ligadas.

15 51. Método de acordo com a reivindicação 48, em que a mistura de iniciador de sobreposição-extensão multiplexa compreende conjunto de iniciador, em que pelo menos um elemento do conjunto de iniciador de cada conjunto de iniciador compreende uma cauda de sobreposição-extensão capaz de hibridizar com relação à cauda de sobreposição-extensão de um elemento do conjunto de iniciador de um segundo conjunto de iniciador.

20 52. Método de acordo com a reivindicação 48 ou 51, em que a mistura de iniciador de sobreposição-extensão multiplexa compreende:

25 a) pelo menos um iniciador mKappar1 ou hmJK complementar à fita de senso de uma sequência de codificação de região de cadeia leve de imunoglobulina;

b) pelo menos um iniciador mVK complementar à fita de antisenso de uma sequência de codificação de região variável de cadeia leve de imunoglobulina ou sequência líder de região variável de cadeia leve e capaz de formar um conjunto de iniciador com o(s) iniciador(es) na etapa a);

30 c) pelo menos um iniciador mCHrev1, mHCrev1-ext, ou mJH complementar à fita de senso de uma sequência de codificação de domínio de cadeia pesada de imunoglobulina; e

d) pelo menos um iniciador mVH complementar à fita de antissenso de uma sequência de codificação de região variável de cadeia pesada de imunoglobulina ou sequência líder de região variável de cadeia pesada e capaz de formar um conjunto de iniciador com o(s) iniciador(es) na etapa c).

5 53. Método de acordo com a reivindicação 44, em que o produto da PCR é inserido em um vetor de expressão.

 54. Método de acordo com a reivindicação 53, em que um cassete de promotor duplo é inserido na construção de expressão, o cassete promotor duplo sendo capaz de direcionar a expressão simultânea de cadeias pesadas e leves, preferivelmente em que o cassete promotor duplo for bidirecional.

 55. Método de acordo com a reivindicação 54, em que o cassete promotor duplo inclui adicionalmente uma sequência de ácido nucleico codificando peptídeos de sinal duplo.

15 56. Método de acordo com a reivindicação 53, em que a estrutura do vetor de expressão compreende uma sequência de codificação de cadeia leve constante humana ou seu fragmento e/ou uma sequência de codificação de cadeia pesada constante humana ou seu fragmento.

 57. Método de acordo com qualquer uma das reivindicações 43 a 20 56, compreendendo uma etapa de ampliação adicional, em que um polinucleotídeo codificando uma cadeia leve constante humana ou seu fragmento com uma sobreposição capaz de provisão de ligação à cadeia leve variável é adicionado à mistura da PCR em conjunto com um conjunto de iniciador capaz de ampliação de uma construção compreendendo em ordem: uma 25 cadeia VH de murino, um ligador, uma cadeia VL de murino e uma cadeia leve constante humana.

 58. Método de acordo com qualquer uma das reivindicações 43 a 56, compreendendo uma etapa de ampliação adicional, em que um polinucleotídeo codificando cadeia pesada constante humana ou seu fragmento 30 com uma sobreposição capaz de provisão de ligação a cadeia pesada variável, é adicionado à mistura da PCR em conjunto com um conjunto de iniciador capaz de ampliação de uma construção compreendendo em ordem: a

cadeia pesada constante humana, uma cadeia VH de murino, um ligador e uma cadeia VL de murino.

5 59. Biblioteca de vetores codificando anticorpos quiméricos, cada elemento de anticorpo consistindo em sequências de codificação de região variável de imunoglobulina não-humana e regiões constantes de cadeia pesada e leve de imunoglobulina humana.

60. Biblioteca de acordo com a reivindicação 59, em que os ditos vetores são obtidos pelo método como definido nas reivindicações 1 a 41 ou método como definido nas reivindicações 43 a 58.

10 61. Biblioteca de acordo com a reivindicação 59, em que a região constante de cadeia leve é uma região constante capa.

62. Biblioteca de acordo com a reivindicação 59, em que as sequências diferentes das humanas são de roedor, preferivelmente de murino.

15 63. Biblioteca de acordo com a reivindicação 59, em que as sequências diferentes das humanas são sequências de coelho.

64. Biblioteca de acordo com a reivindicação 59, em que os vetores são vetores de expressão.

65. Biblioteca de acordo com a reivindicação 59, em que as regiões variáveis são pares cognatos.

20 66. Biblioteca de acordo com a reivindicação 59, em que a região constante de imunoglobulina humana é selecionada dentre as classes de imunoglobulina humana IgA1, IgA2, IgD, IgE, IgG1, IgG2, IgG3, IgG4 ou IgM.

67. Biblioteca de acordo com a reivindicação 66, em que a região constante é selecionada dentre IgG1 e IgG2.

25 68. Sub-biblioteca que codifica anticorpos exibindo especificidades de ligação desejadas direcionadas contra um alvo específico, selecionada de uma biblioteca como definida em uma das reivindicações 59 a 67.

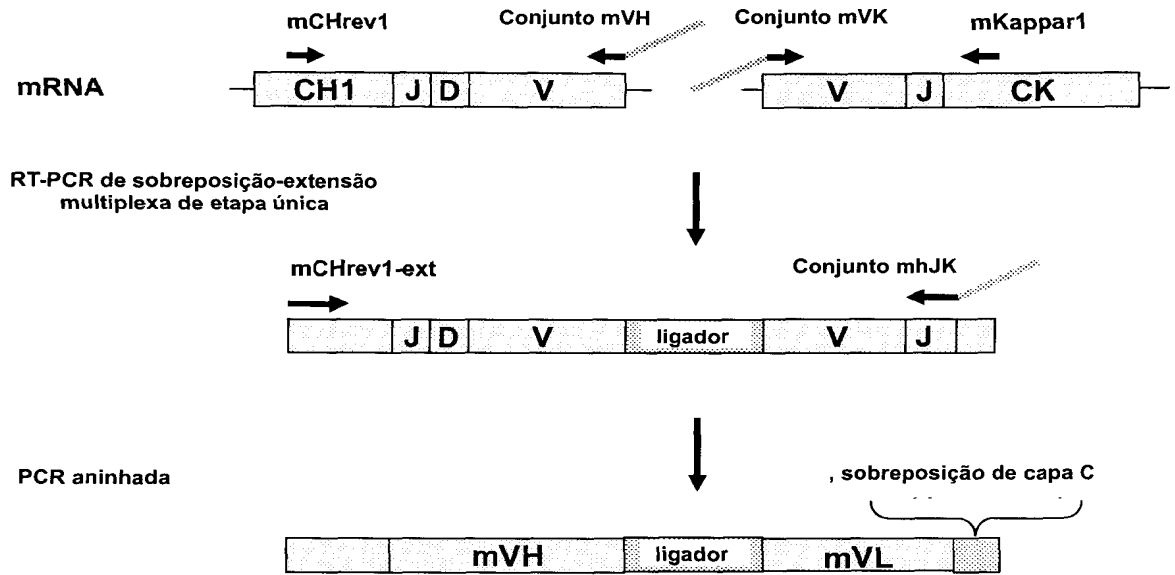


Fig. 1

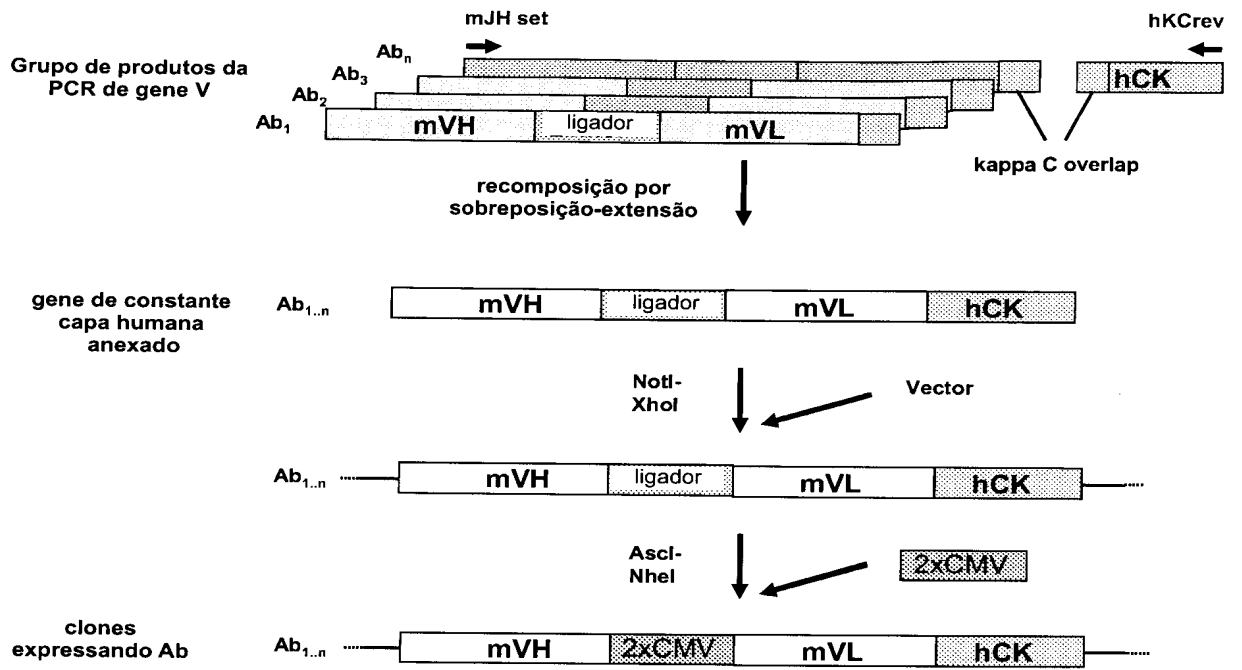


Fig. 2

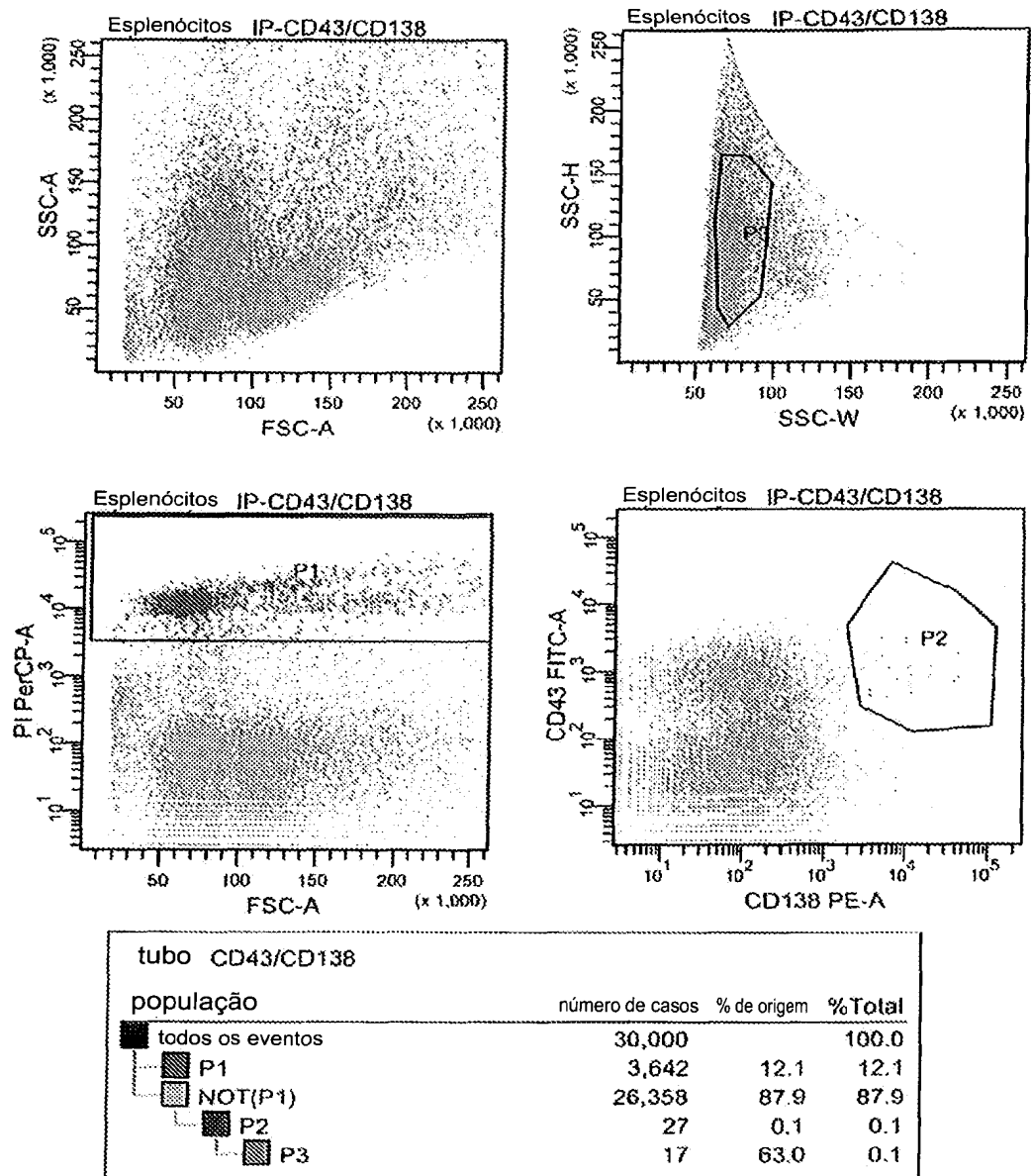


Fig. 3 A

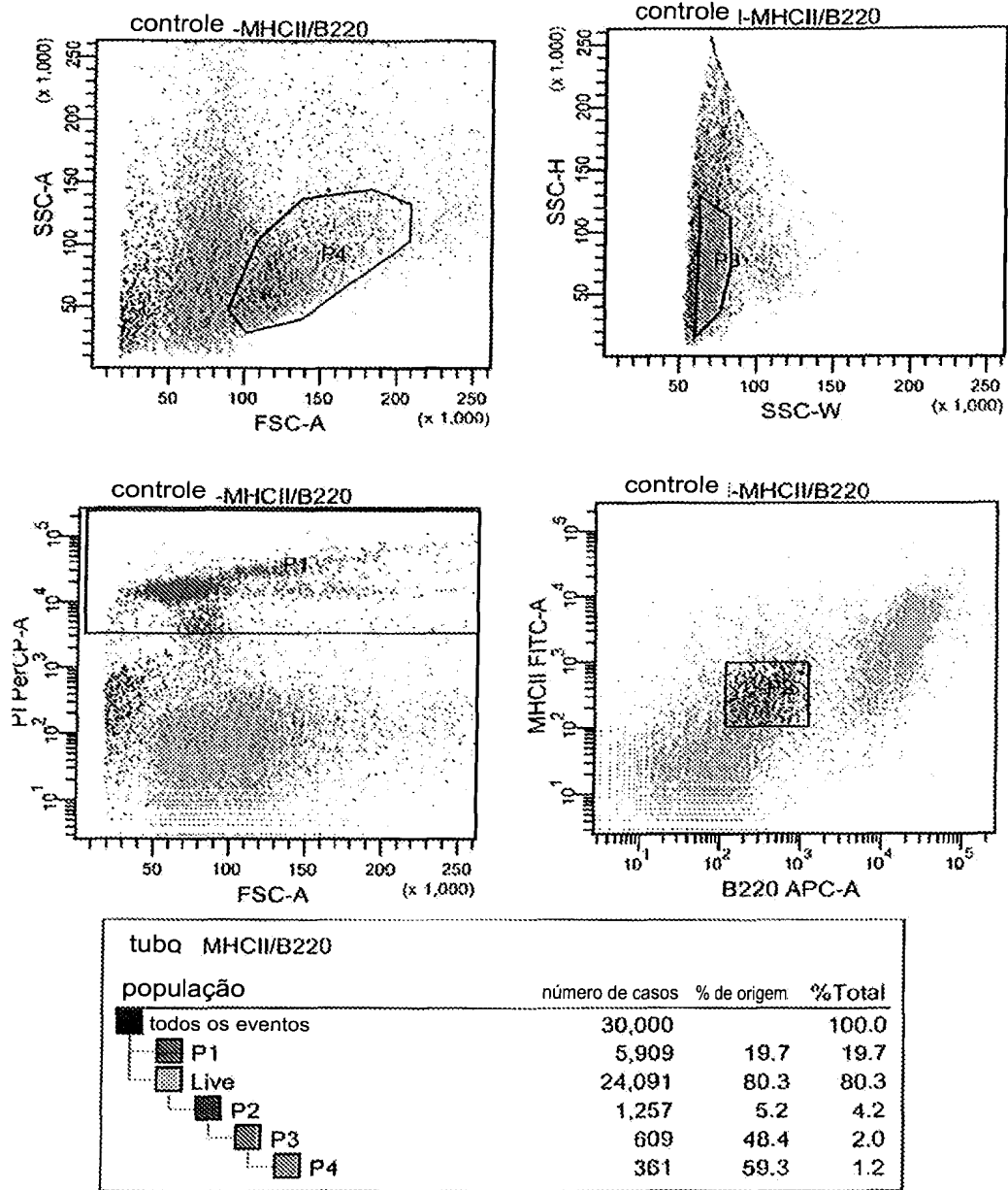


Fig.3 B

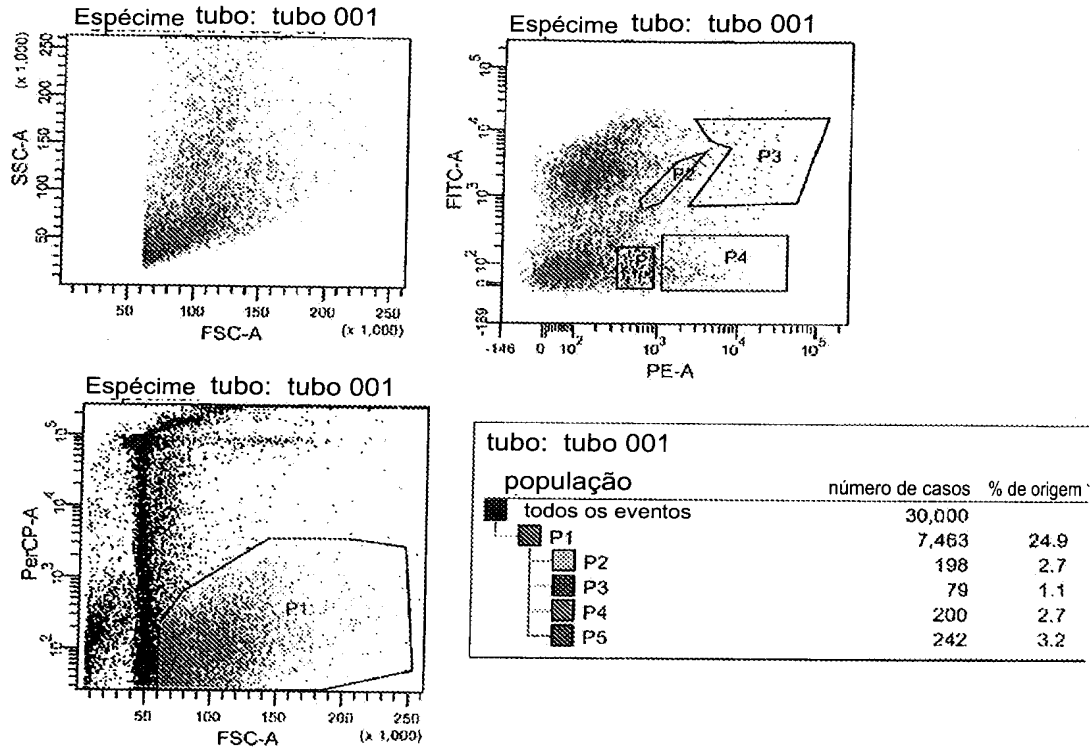


Fig. 4

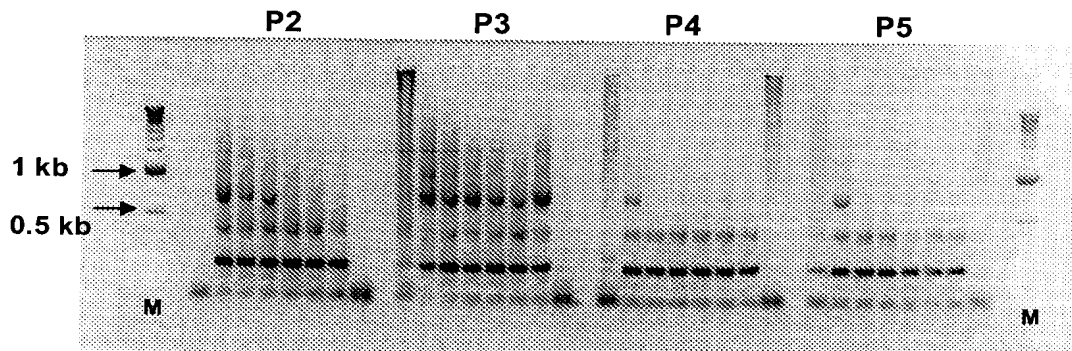


Fig. 5

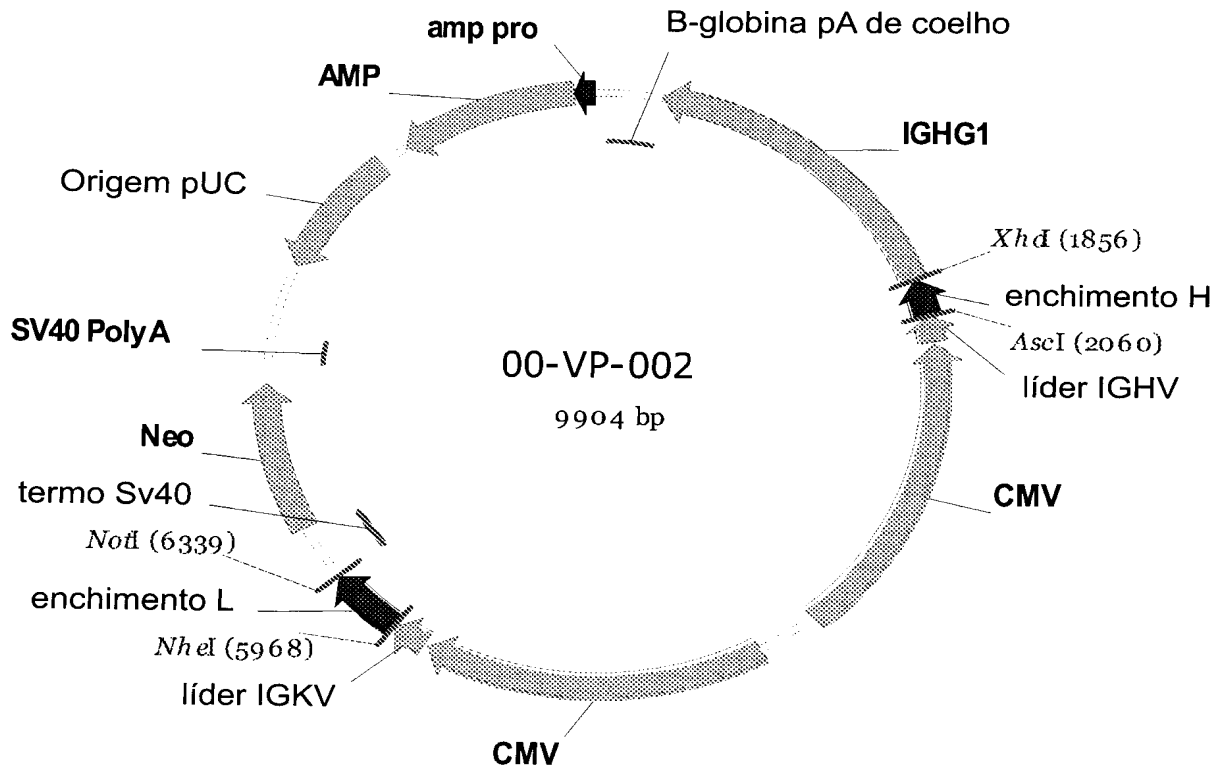


Fig. 6

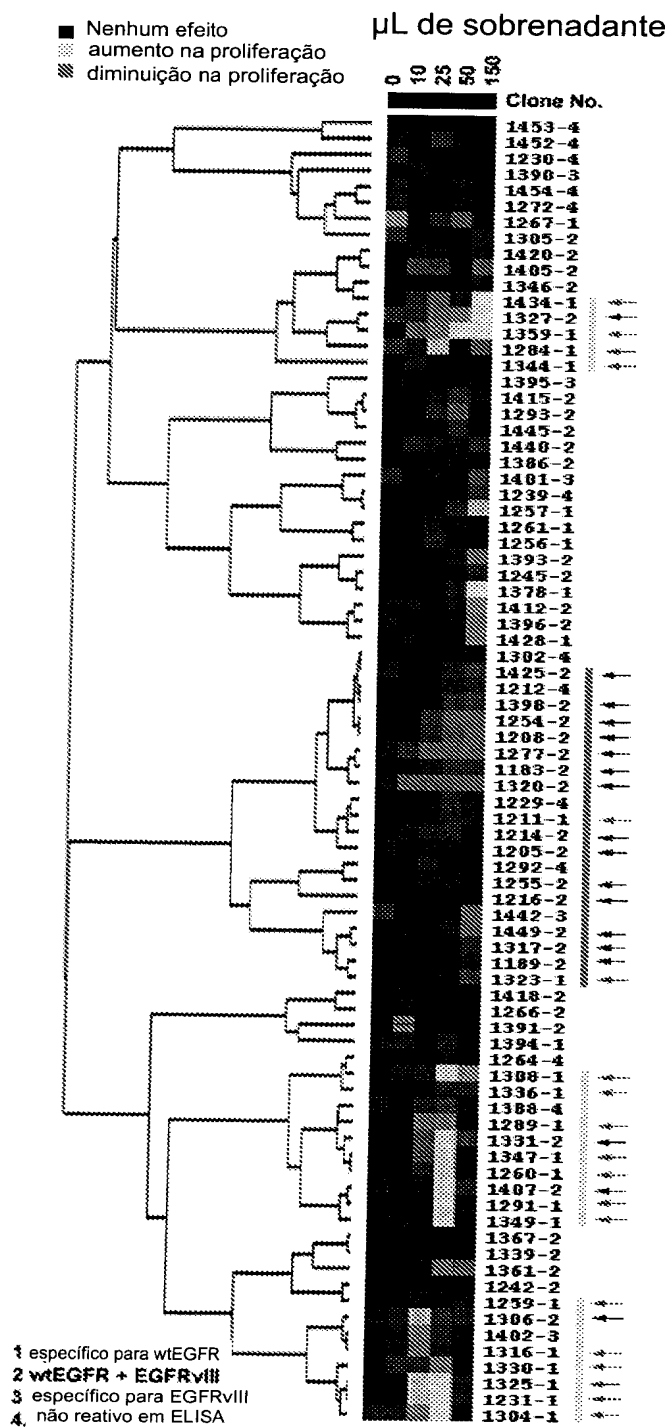


Fig. 7

RESUMO

Patente de Invenção: **"MÉTODO PARA CLONAGEM DE ANTICORPOS COGNATOS"**.

5 A presente invenção refere-se a um procedimento para ligação de pares cognatos de sequências de codificação de V_H e V_L de uma população de células enriquecidas especificamente nos marcadores de antígeno de superfície. O procedimento de ligação envolve um procedimento de ampliação molecular multiplexa capaz de ligar sequências de nucleotídeo de interesse, em conexão com a ampliação, especificamente reação em cadeia da
10 polimerase (PCR multiplexa). O método é especificamente vantajoso para a geração de bibliotecas de pares cognatos, bem como bibliotecas combinatórias de sequências de codificação de região variável de imunoglobulinas. A invenção também refere-se aos métodos para geração de anticorpos humanos/não-humanos quiméricos e bibliotecas de expressão geradas por tais
15 métodos.