



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

**(11) BR 112016003480-5 B1**



**(22) Data do Depósito:** 19/08/2014

**(45) Data de Concessão:** 16/08/2022

---

**(54) Título:** ENSAIOS PARA DETECÇÃO DE MOLÉCULA ÚNICA E SEU USO

**(51) Int.Cl.:** G01N 33/53.

**(30) Prioridade Unionista:** 19/08/2013 US 61/867.559; 19/08/2013 US 61/867.554.

**(73) Titular(es):** SINGULAR BIO, INC..

**(72) Inventor(es):** ADRIAN NIELSEN FEHR; PATRICK JAMES COLLINS; JILL LYNDON HERSCHLEB; HYWEL BOWDEN JONES.

**(86) Pedido PCT:** PCT US2014051763 de 19/08/2014

**(87) Publicação PCT:** WO 2015/026873 de 26/02/2015

**(85) Data do Início da Fase Nacional:** 19/02/2016

**(57) Resumo:** ENSAIOS PARA DETECÇÃO DE MOLÉCULA ÚNICA E SEU USO. A invenção está relacionada a métodos de detecção de uma variação genética em uma amostra genética de um paciente que utiliza sondagens rotuladas e à contagem do número de rótulos nas sondagens.

**ENSAIOS PARA DETECÇÃO DE MOLÉCULA ÚNICA E SEU USO**

## Antecedente da Invenção

[0001] A invenção está relacionada aos métodos de detecção de uma variação genética em uma amostra genética de um paciente. A detecção de uma variação genética é importante em muitos aspectos da biologia humana.

## Sumário

[0002] A invenção está relacionada aos métodos de detecção de uma variação genética em uma amostra genética de um paciente. A invenção ainda está relacionada aos métodos de detecção de uma variação genética em uma amostra genética de um paciente que usa sondagens rotuladas e contagem do número de rótulos nas sondagens.

Descrição breve dos desenhos

[0003] A figura 1 representa membros de grupos que consistem na ligação de parceiros, marcas, marcas de afinidade, sondagens de marcação, conjuntos de sondagem e/ou conjuntos de sondagem litigados descritos nestes documento ou em um documento base.

[0004] A figura 2 representa um histograma normalizado de intensidade de sinal medida com base nas amostras de rótulo únicas e anticorpos de vários rótulos.

[0005] A figura 3 representa perfis de branqueamento médio de vários rótulos.

[0006] As figuras 4-13 mostram os gráficos de intensidade de rótulo integrado no decorrer do tempo para vários rótulos Alexa 488.

[0007] A figura 14 representa o espectro de excitação e o espectro de emissão por meio de uma operação padrão quando a excitação de um fluoróforo é alcançada iluminando-se com uma faixa espectral estreita alinhada à absorção máxima dessas espécies.

[0008] A figura 15 representa o espectro da excitação e o espectro da emissão através da interrogação com várias cores de excitação e faixas de emissão coletadas diferentes (ou além) do caso para a operação padrão.

[0009] A figura 16 mostra os resultados quando a luz dessas várias configurações de imagens, por ex., vários filtros de emissão, é coletada e comparada aos valores de calibração para os fluoróforos de interesse.

[0010] A figura 17 mostra os resultados coletados com várias referências, inclusive aquelas com um perfil de emissão plana (Contaminante 1; triângulos) ou um perfil de peso azul (Contaminante 2; estrelas).

[0011] A figura 18 representa faixas de excitação de dois fluoróforos diferentes de modo significativo.

[0012] A figura 19 representa um exemplo do fluxograma do sistema.

[0013] A figura 20 representa um exemplo do fluxograma do sistema que inclui vários métodos para analisar dados.

[0014] As figuras 21-46 representam exemplos de conjuntos de sondagem descritos aqui.

[0015] As figuras 47 e 48 mostram os padrões de fluorescência resultantes quando os produtos contêm sequências de marca de afinidade exclusivas e o substrato subjacente contém complementos para cada uma das marcas de afinidade exclusiva na mesma localidade (por ex., como o mesmo membro) em um substrato.

[0016] As figuras 49 e 51 mostram os padrões de fluorescência resultantes quando diferentes produtos contêm sequências de marca de afinidade idênticas e o substrato subjacente contém o complemento da marca de afinidade.

[0017] As figuras 50 e 52 mostram as localidades com zoom nas figuras 49 e 51, respectivamente.

[0018] As figuras 53 e 54 mostram os padrões de fluorescência resultantes quando os produtos contêm sequências de marca de afinidade exclusiva e o substrato subjacente tem uma localidade (por ex., como um membro) contendo o complemento a um complemento de marca de afinidade e outro local separado (por ex., como outro membro) contendo o complemento a outra marca de afinidade.

[0019] A figura 55 representa dois conjuntos de sondagem; um conjunto de sondagem para o Locus 1 e um conjunto de sondagem para o Locus 2 - embora, conforme citado anteriormente, vários conjuntos de sondagem possam ser designados para cada locus genômico.

[0020] A figura 56 representa um fluxo de trabalho prático que seria aplicado ao grupo de conjuntos de sondagem.

[0021] A figura 57 representa uma versão modificada do fluxo de trabalho prático ilustrado na figura 56.

[0022] A figura 58 fornece um exemplo de como os produtos de sondagem para Locus 1 e Locus 2 podem ser rotulados com diferentes moléculas de rótulo.

[0023] A figura 59 fornece evidência de que os produtos que representam vários locais genômicos para um locus podem ser gerados em uma maneira específica de enzima ligase que usa o processo de hibridização-ligação.

[0024] A figura 60 fornece dados que indicam que os conjuntos de sondagem podem ser usados para detectar mudanças relativas no estado de número de cópia.

[0025] A figura 61 fornece evidência de que misturas de produtos de sondagem podem ser usados para gerar dados de microconjunto quantitativo.

[0026] As figuras 62-64 ilustram modificações do procedimento geral descrito nas figuras 55 a 58.

[0027] A figura 65 representa uma outra representação do procedimento modificado descrito na figura 62.

[0028] A figura 66 representa ainda outra representação do procedimento representado na figura 65.

[0029] A figura 67 representa exemplos de conjuntos de sondagem usados nos métodos descritos aqui.

[0030] A figura 68 representa exemplos de conjuntos de sondagem usados nos métodos descritos aqui com translocações que têm pontos de interrupção que são analisados.

[0031] A figura 69 representa exemplos de conjuntos de sondagem usados nos métodos descritos aqui quando as mutações em SNPs são pretendidas.

#### Descrição detalhada da invenção

[0032] Os métodos descritos neste documento, exceto se indicado de outra forma, podem empregar técnicas e descrições convencionais de biologia molecular (incluindo técnicas recombinantes), biologia celular, bioquímica e microarranjo e tecnologia de sequenciamento, conhecidos pelos especialistas na arte. Essas técnicas convencionais incluem sínteses de arranjo de polímero, hibridização e ligação de oligonucleotídeos, sequenciamento de oligonucleotídeos e detecção de hibridização usando um rótulo. Ilustrações específicas de técnicas adequadas podem ser usadas como referência para os exemplos deste documento. Entretanto, os procedimentos convencionais equivalentes podem, é claro, também ser usados. Essas técnicas e descrições convencionais podem ser encontradas, por exemplo, em Kimmel and Oliver, DNA Microarrays (2006) Elsevier; Campbell, DNA Microarray, Synthesis and Synthetic DNA (2012) Nova Science; Bowtell and Sambrook, DNA Microarrays: Molecular Cloning Manual (2003) Cold Spring Harbor Laboratory Press. Antes de as composições presentes, ferramentas e métodos de pesquisa serem descritos, é necessário compreender que esta invenção não está

limitada aos métodos, composições, alvos e usos específicos descritos, como tal, eles podem variar, é claro. Também deve ser compreendido que a terminologia usada neste documento tem como finalidade apenas descrever os aspectos particulares e não tem intenção de limitar o escopo da presente invenção, que será limitado apenas pelas reivindicações anexas.

[0033] A invenção está relacionada aos métodos de detecção de uma variação genética em uma amostra genética de um paciente. A variação genética aqui pode incluir, mas não está limitada a, uma ou mais substituição, inversão, inserção, exclusão ou mutação nas sequências de nucleotídeo (por ex., DNA e RNA) e proteínas (por ex., peptídeo e proteína) um ou mais alelo, polimorfismo, polimorfismo de nucleotídeo único (SNP), polimorfismo genético de larga escala, como inversões e translocações, diferenças na abundância e/ou número de cópia (por ex., variantes do número de cópia, CNVs) de uma ou mais moléculas de nucleotídeo (por ex., DNA), trissomia, monossomia e rearranjos genômicos. Em algumas representações, a variação genética pode estar relacionada às metástases, presença, ausência e/ou risco de uma doença, como o câncer, variabilidade farmacocinética, toxicidade da droga, eventos adversos, recorrência e/ou presença, ausência ou risco de rejeição de transplante de órgão no paciente. Por exemplo, as mudanças do número de cópia no gene *HER2* podem afetar na resposta ou não de um paciente com câncer de mama ao tratamento com Herceptina. Da mesma maneira, a detecção no aumento no número de cópia de cromossomo 21 (ou 18 ou 13 ou cromossomos sexuais) no sangue de uma mulher grávida pode ser usada para um diagnóstico não invasivo de Síndrome de Down em uma criança ainda não nascida. Um exemplo adicional é a detecção de alelos de um órgão transplantado que não estão presentes no genoma receptor - monitoramento da frequência ou número de cópia desses alelos

podem identificar sinais de possível rejeição do órgão. Vários métodos podem ser usados para detectar essas alterações (por ex., rtPCR, sequenciamento e microarranjos). Um dos métodos é contar moléculas individuais rotuladas para detectar a presença de uma mutação (por ex., EGFR, mutação no câncer) ou um excesso de sequência ou região genômica específica (por ex., Cromossomo 21 na Síndrome de Down). A contagem de moléculas únicas pode ser feita de várias maneiras, com uma leitura comum para depositar as moléculas em uma superfície e imagem.

[0034] Além disso, a variação genética pode ser novamente mutações genéticas, como mutações de base única ou de várias bases, translocações, ampliações subcromossômicas, exclusões e aneuploidia. Em algumas representações, a variação genética pode indicar uma sequência de nucleotídeo alternativa em um locus genético que pode estar presente em uma população de indivíduos e que inclui substituições, inserções e exclusões de nucleotídeo em relação aos outros membros da população. Em representações adicionais, a variação genética pode ser aneuploidia. Ainda nas representações, a variação genética pode ser trissomia 13, trissomia 18, trissomia 21, aneuploidia de X (por ex., trissomia XXX e trissomia XXY) ou aneuploidia de Y (por ex., trissomia XYY). Em outras representações, a variação genética pode ser na região 22q11.2, 1q21.1, 9q34, 1p36, 4p, 5p, 7q11.23, 11q24.1, 17p, 11p15, 18q ou 22q13. Em outras representações, a variação genética pode ser microexclusão ou microamplificação.

[0035] Em algumas representações, a detecção, descoberta, determinação, medição, avaliação, contagem e avaliação da variação genética são usadas de modo intercambiável e incluem determinações quantitativas e/ou qualitativas, incluindo, por exemplo, identificação da variação genética, determinação de presença e/ou ausência da variação genética e quantificação da

variação genética. Em outras representações, os métodos da presente divulgação podem detectar várias variações genéticas. O termo "e/ou" usado neste documento é definido para indicar qualquer combinação dos componentes. Além disso, as formas singulares "um", "uma" e "o/a" pode ainda incluir referentes plurais, a menos que o contexto defina claramente de outra forma. Dessa forma, por exemplo, a referência a uma "região de nucleotídeo" indica uma, mais de uma ou misturas dessas regiões e a referência a "um ensaio" pode incluir a referência a etapas e métodos equivalentes conhecidos pelos especialistas na arte e assim por diante.

[0036] "Amostra" significa uma quantidade de material de origem biológica, ambiental, médica ou de paciente na qual é observada a detecção, a medição ou a rotulação dos ácidos nucleicos, peptídeos e/ou proteínas almejados. Por um lado, é indicada a inclusão de um espécime ou cultura (por ex., culturas microbiológicas). Por outro lado, é indicada a inclusão de amostras biológicas e ambientais. Uma amostra pode incluir um espécime de origem sintética. As amostras ambientais incluem material ambiental, como matéria da superfície, óleo, água e amostras industriais, bem como amostras obtidas de instrumentos de processamento de alimentos e laticínios, aparelho, equipamento, utensílios, itens descartáveis e não descartáveis. "Amostra genética" pode ser qualquer amostra líquida ou sólida com informações biológicas hereditárias e/ou não hereditárias, codificadas nas sequências de nucleotídeo dos ácidos nucleicos. A mostra pode ser obtida de uma origem, incluindo, mas não limitado a, todo o sangue, soro, plasma, urina, saliva, suor, matéria fecal, lágrimas, fluido intestinal, amostras de membrana da mucosa, tecido do pulmão, tumores, órgãos transplantados, feto e/ou outras origens. As amostras genéticas podem ser de fluido, sólido, (por exemplo, fezes) ou tecido animal, inclusive

humano. As amostras genéticas podem incluir os materiais obtidos de um paciente, incluindo, mas não limitado a, culturas, sangue, saliva, fluido cerebrospinal, leite, linfa, escarro, sêmen, aspirados de agulha e similares. Além disso, a amostra genética pode ser um material genético fetal de uma amostra de sangue materno. O material genético fetal pode ser isolado e separado da amostra de sangue materno. A amostra genética pode ser uma mistura de material genético fetal e materno. Além disso, a amostra genética pode incluir sequências genéticas anômalas que surgem da formação de tumores ou metástase e/ou assinaturas de DNA do doador presentes em um recipiente transparente. Em representações adicionais, quando a amostra genética é plasma, o método pode consistir no isolamento do plasma a partir de uma amostra sanguínea do paciente. Em outras representações, quando a amostra genética é soro, o método pode consistir no isolamento do soro a partir de uma amostra sanguínea do paciente. Ainda em representações adicionais, quando a amostra genética é uma amostra de DNA livre de células (cfDNA), o método também consiste no isolamento da amostra de DNA livre de células a partir de uma amostra obtida da origem descrita neste documento. A amostra de DNA livre de células descrita neste documento significa uma população de moléculas de DNA que circula livremente no fluxo sanguíneo, fora de qualquer célula ou organela. No caso de gravidez, o DNA livre de células da mãe carrega uma mistura do DNA materno e também do DNA fetal. Esses exemplos não devem ser interpretados como limitadores dos tipos de amostra aplicáveis a presente invenção.

[0037] Em algumas representações, o método da presente divulgação pode compreender a seleção e/ou o isolamento do locus genético ou locus de interesse e a quantificação da quantidade de cada locus presente (por exemplo, para a determinação do número de cópia) e/ou as quantidades relativas de diferentes

variantes de locus (por exemplo, dois alelos de uma determinada sequência de DNA). Região, região de interesse, locus ou locus de interesse em referência a um genoma ou polinucleotídeo alvo usado neste documento indica uma sub-região ou segmento do genoma ou polinucleotídeo alvo. Conforme usado neste documento, região, regiões de interesse, locus ou locus de interesse em uma molécula de nucleotídeo pode se referir a posição de um nucleotídeo, um gene ou uma porção de um gene em um genoma, incluindo DNA mitocondrial ou outro DNA não cromossômico ou pode se referir a qualquer porção contígua de sequência genômica esteja ela dentro ou não de, ou associada a, um gene. Uma região, região de interesse, locus ou locus de interesse em uma molécula de nucleotídeo pode ser de um único nucleotídeo para um segmento de algumas centenas ou alguns milhares de nucleotídeos em extensão ou mais. Em algumas representações, uma região ou locus de interesse pode ter uma sequência de referência associada a ela. O termo "sequência de referência" usado neste documento representa uma sequência à qual um locus de interesse em um ácido nucleico está sendo comparada. Em certas representações, uma sequência de referência é considerada uma sequência de "tipo selvagem" para um locus de interesse. Um ácido nucleico que contém um locus de interesse tendo uma sequência que varia de uma sequência de referência para o locus de interesse é ocasionalmente referida como "polimórfica", ou "mutante", ou de "variação genética". Um ácido nucleico que contém um locus de interesse tendo uma sequência que não varia de uma sequência de referência para o locus de interesse é ocasionalmente referida como "não polimórfica", ou "tipo selvagem", ou de "variação não genética". Em determinadas representações, um locus de interesse pode ter mais de uma sequência de referência distinta associada a ele (por ex., onde um locus de interesse é conhecido por ter um polimorfismo que

deve ser considerado um tipo normal ou selvagem). Em algumas representações, o método da presente divulgação também pode compreender a escolha e/ou o isolamento de peptídeo ou peptídeos de interesse e qualificação da quantidade de cada peptídeo presente e/ou quantidades relativas de diferentes peptídeos.

[0038] Em representações adicionais, a região de interesse descrita neste documento pode incluir "sequência de variante genética de consenso" que se refere à sequência de ácido nucleico ou proteína, aos ácidos nucleicos ou aminoácidos os quais podem ocorrer com alta frequência em uma população de indivíduos que carregam o gene cujos códigos para uma proteína que não funciona normalmente ou na qual o próprio ácido nucleico não funcional normalmente. Além disso, a região de interesse descrita neste documento pode incluir a "sequência de gene normal de consenso" que se refere a uma sequência de ácido nucleico, ao ácido nucleico os quais podem ocorrer em suas respectivas posições com alta frequência em uma população de indivíduos que carregam o gene cujos códigos para uma proteína não funcionam normalmente ou que ela própria não funciona normalmente. Em outras representações, a região de controle que não é a região de interesse ou a sequência de referência descrita neste documento pode incluir a "sequência normal de consenso" que se refere à sequência de ácido nucleico ou proteína, aos ácidos nucleicos ou aminoácidos os quais podem ocorrer com alta frequência em uma população de indivíduos que carregam o gene cujos códigos para uma proteína que funciona normalmente ou na qual o próprio ácido nucleico tem função normal.

[0039] Os métodos descritos aqui podem produzir medições altamente precisas de variação genética. Um tipo de variação descrita neste documento inclui abundância relativa de dois ou mais locais genômicos distintos. Neste caso, os locais podem ser

pequenos (por ex., 300, 250, 200, 150, 100 ou 50 nucleotídeos ou menos) de tamanho moderado (por ex., 1.000, 10.000, 100.000 ou um milhão de nucleotídeos) e tão grandes como uma porção de um braço de cromossomo ou cromossomo inteiro ou conjuntos de cromossomos. Os resultados deste método podem determinar a abundância de um locus para outro. A precisão e a exatidão dos métodos da presente divulgação podem permitir a detecção de alterações muito pequenas no número de cópia (aproximadamente 25, 10, 5, 4, 3, 2, 1, 0,5, 0,1, 0,05, 0,02 ou 0,01 % ou menos), o que permite a identificação de uma assinatura muito fraca de variação genética. Por exemplo, uma assinatura de aneuploidia fetal pode ser encontrada em uma amostra de sangue materno onde a anomalia genética fetal é diluída pelo sangue materno e um número de cópia observável de mudança de aproximadamente 2% é indicativo de trissomia fetal.

[0040] Conforme usado neste documento, o termo "aproximadamente" indica a modificação, por exemplo, de extensões de sequências de nucleotídeo, graus de erros, dimensões, quantidade de um ingrediente em uma composição, concentrações, volumes, temperaturas do processo, tempo do processo, rendimento, taxas de vazão, pressões e valores similares, além de suas variações, refere-se à variação na quantidade numérica que pode ocorrer, por exemplo, por meio de medição típica e procedimentos de manipulação usados para fazer compostos, composições, concentrados ou usar formulações; através de erro involuntário nesses procedimentos; através de diferenças na fabricação, origem ou pureza de materiais iniciais ou ingredientes usados para executar os métodos e considerações semelhantes. O termo "aproximadamente" também inclui quantidades que diferem devido à maturação de, por exemplo, uma composição, formulação ou cultura de célula com uma concentração ou mistura inicial e quantidades que diferem devido à mistura ou

processamento de uma composição ou formulação com uma concentração ou mistura inicial em particular. Se forem modificadas pelo termo "aproximadamente", as reivindicações anexas incluirão equivalentes a essas quantidades. O termo "aproximadamente" pode se referir ainda a um intervalo de valores que sejam similares ao valor de referência determinado. Em determinadas representações, o termo "aproximadamente" se refere a um intervalo de valores cuja variação está entre 50, 25, 10, 9, 8, 7, 6, 5, 4, 3, 2, 1 por cento ou menos do valor de referência determinado.

[0041] Em algumas representações, a paciente pode ser grávida, uma pessoa, um paciente com risco elevado de doença genética (por ex., câncer), todas as outras várias famílias de animais domésticos, bem como animais silvestres ou selvagens. Em algumas representações, a variação genética pode ser uma variação genética no feto da paciente grávida (por ex., variantes de número de cópia e aneuploidia no feto). Em algumas representações, a paciente é uma paciente grávida, e a variação genética é uma variação no feto da paciente grávida em uma região selecionada no grupo que consiste em 22q11.2, 1q21.1, 9q34, 1p36, 4p, 5p, 7q11.23, 11q24.1, 17p, 11p15, 18q e 22q13, (por ex., uma mutação e/ou mudança de número de cópia em qualquer uma das regiões 22q11.2, 1q21.1, 9q34, 1p36, 4p, 5p, 7q11.23, 11q24.1, 17p, 11p15, 18q e 22q13). O feto descrito neste documento pode indicar uma prole que ainda não nasceu de um ser humano ou outro animal. Em algumas representações, o feto pode ser a prole mais de 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 semanas após a concepção. Em representações adicionais, o feto pode ser uma prole concebida por implantes, fertilização in vitro, várias gravidezes ou nascimento de gêmeos. Em representações adicionais, o feto pode

ser parte de um par de gêmeos (idênticos ou não idênticos) ou um trio ou tripletos (idênticos ou não idênticos).

[0042] As invenções, de acordo com algumas representações, incluem pelo menos dois componentes importantes: um ensaio para a identificação seletiva dos locais genômicos e uma tecnologia para quantificação desses locais com alta precisão. O ensaio pode incluir métodos de rotulação e/ou isolamento seletivo de uma ou mais sequências de ácido nucleico, de maneira que a etapa de rotulação seja suficiente para produzir moléculas (definidas como "produtos de sondagem", "conjunto de sondagem ligada", "conjunto de sondagem conjugada", "sondagens ligadas", "sondagens conjugadas" ou "moléculas rotuladas" nesta invenção) contendo todas as informações necessárias para identificação de uma sequência particular no contexto de um ensaio particular. Por exemplo, o ensaio pode compreender contato, ligação e/ou hibridização de sondagens de uma amostra, ligação e/ou conjugação das sondagens, amplificando opcionalmente as sondagens ligadas/conjugadas e imobilização das sondagens em um substrato. Em algumas representações, os ensaios e métodos descritos neste documento podem ser realizados em uma única amostra de entrada em paralelo como um ensaio multiplex conforme descrito neste documento.

[0043] O produto da sondagem, conjunto de sondagem ligada, conjunto de sondagem conjugada, sondagens ligadas, sondagens conjugadas e moléculas rotuladas podem ser moléculas únicas contíguas resultantes do desempenho da ação enzimática em um conjunto de sondagem, como um ensaio. Em um produto de sondagem ou uma molécula rotulada uma ou mais sondagens individuais de um conjunto de sondagem podem ser modificadas de modo que formem uma espécie molecular singular distinta quando comparada às sondagens ou conjuntos de sondagens. Como resultado, os produtos de sondagem ou uma molécula rotulada podem ser quimicamente

distintos e, portanto, podem ser identificados, contados, isolados ou manipulados sem considerar as sondagens ou conjuntos de sondagem.

[0044] Por exemplo, os produtos de sondagem podem conter um ou mais rótulos de identificação e uma ou mais marcas de afinidade para isolamento e/ou imobilização. Em algumas representações, nenhuma modificação adicional de produtos de sondagem (por ex., determinação de sequência de DNA) precisa ser realizada. Em algumas representações, não é exigida nenhuma investigação adicional da sequência de DNA. Os produtos de sondagem que contêm os rótulos podem ser diretamente contados, geralmente após uma etapa de imobilização em um substrato sólido. Por exemplo, os rótulos de fluoróforo orgânico são usados para rotular produtos de sondagem, e os produtos de sondagem são contados diretamente pela imobilização dos produtos de sondagem para um substrato de vidro e geração de imagem subsequente via microscópio fluorescente e uma câmera digital. Em outras representações, o rótulo pode ser seletivamente eliminado ou removido dependendo de a molécula rotulada ter interagido com seu locus genômico complementar. Em representações adicionais, dois rótulos em porções opostas do produto de sondagem podem funcionar de acordo com um sinal de distribuição de uma transferência de energia de ressonância fluorescente (FRET) dependendo de a molécula rotulada ter interagido com seu locus genômico complementar. Para um determinado locus genômico, as sondagens de rotulação contendo os rótulos são designadas para qualquer região de sequência nesse locus. Um conjunto de várias sondagens de rotulação com rótulos iguais ou diferentes também pode ser designado para um único locus genômico. Nesse caso, uma sondagem pode seletivamente isolar e rotular uma região diferente em um locus particular ou regiões sobrepostas em um locus. Em algumas

representações, os produtos de sondagem contendo as marcas de afinidade são imobilizados no substrato via marcas de afinidade. Por exemplo, as marcas de afinidade são usadas para imobilizar produtos de sondagem no substrato, e os produtos de sondagem contendo as marcas de afinidade são contados diretamente. Para um determinado locus genômico, as sondagens de marcação contendo as marcas de afinidade são designadas para qualquer região de sequência nesse locus. Um conjunto de várias sondagens de marcação com marcas iguais ou diferentes também pode ser designado para um único locus genômico. Nesse caso, uma sondagem pode seletivamente isolar e marcar uma região diferente em um locus particular ou regiões sobrepostas em um locus.

[0045] Sob um aspecto, os métodos da presente divulgação podem consistir nos conjuntos de sondagem de contato descritos com a amostra genética apresentada neste documento. Em algumas representações, os métodos da presente divulgação podem consistir no contato de vários conjuntos de sondagem, como um primeiro e segundo conjuntos de sondagem, para a amostra genética. Nas representações adicionais, cada um dos conjuntos de sondagem compreende uma sondagem de rotulação e uma sondagem de marcação. Por exemplo, o primeiro conjunto de sondagem compreende uma primeira sondagem de rotulação e uma primeira sondagem de marcação, e o segundo conjunto de sondagem compreende uma segunda sondagem de rotulação e uma segunda sondagem de marcação.

[0046] O contato dos conjuntos de sondagem com a amostra genética pode ser realizado simultaneamente ou após a hibridização, ligação, amplificação e/ou imobilização das sondagens. Além disso, o contato dos conjuntos de sondagem com a amostra genética pode ser realizado simultaneamente ou antes da hibridização, ligação, amplificação e/ou imobilização das sondagens.

[0047] Para um determinado locus genômico ou região de uma molécula de nucleotídeo na amostra genética, uma única sequência de ácido nucleico nesse locus ou várias sequências de ácido nucleico nesse locus podem ser investigadas e/ou quantificadas via criação de produtos de sondagem. As sequências investigadas em um locus genômico podem ser distintas e/ou sobrepostas e podem conter ou não os polimorfismos genéticos. Um produto de sondagem é formado pelo esquema de um ou mais nucleotídeos chamado "conjunto de sondagem". Por exemplo, o produto de sondagem pode ser formado ligando-se o conjunto de sondagem, ligando-se as sondagens no conjunto de sondagem. Um conjunto de sondagem compreende pelo menos uma sondagem que hibridiza, conjuga, liga ou imobiliza uma molécula alvo, incluindo ácidos nucleicos (por ex., DNA e RNA), peptídeos e proteínas. Em algumas representações, uma sondagem pode compreender um material isolado, purificado, que ocorre naturalmente, que ocorre não naturalmente e/ou artificial, por exemplo, incluindo oligonucleotídeos de qualquer comprimento (por ex., 5, 10, 20, 30, 40, 50, 100 ou 150 nucleotídeos ou menos), no qual pelo menos uma porção (por ex., 50, 60, 70, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 ou 100%) das sequências de oligonucleotídeos é complementar a um motivo de sequência e/ou domínio de hibridização presente em uma ou mais moléculas alvo, de modo que a sondagem seja configurada para hibridizar (ou interagir de maneira semelhante) em parte ou totalmente com uma ou mais moléculas alvo ou região de ácido nucleico de interesse. A parte da molécula alvo ou a região de interesse do ácido nucleico na qual uma sondagem é hibridizada é chamada de "domínio de hibridização" da sondagem, que pode ser, em parte ou totalmente, da molécula alvo ou região de interesse do ácido nucleico conforme descrito neste documento.

[0048] Uma sondagem pode ser de filamento único ou de filamento duplo. Em algumas representações, a sondagem pode ser preparada com base em uma digestão de restrição purificada ou produzida sinteticamente, de forma recombinante ou por amplificação de PCR. Nas representações adicionais, a sondagem pode consistir em um material que se liga a uma sequência de peptídeo em particular. Um conjunto de sondagem descrito neste documento pode consistir em um conjunto de uma ou mais sondagens designadas para corresponder a um locus genômico único ou um peptídeo em uma sequência de proteína.

[0049] O termo "nucleotídeo" usado neste documento indica um desoxirribonucleotídeo ou um ribonucleotídeo ou qualquer análogo de nucleotídeo (por ex., DNA e RNA). Os análogos de nucleotídeo incluem nucleotídeos com modificações na estrutura química da base, açúcar e/ou fosfato, incluindo, mas não limitado a, modificações de pirimidina na posição 5'-, modificações de purina na posição 8-, modificações em aminas exocíclicas de citocina, substituição de 5-bromo-uracil e similares e modificações de açúcar na posição 2'-, incluindo, mas não limitado a, ribonucleotídeos modificados de açúcar nos quais o 2'-OH é substituído por um grupo selecionado de H, OR, R, halo, SH, SR, NH<sub>2</sub>, NHR, NR<sub>2</sub> ou CN. shRNAs também podem consistir em elementos não naturais, como nucleotídeos não naturais, por ex., ionosina e xantina, açúcares não naturais, por ex., 2'-metóxi ribose ou ligações de fosfodiéster não natural, por ex., metilfosfonatos, fosforotioatos e peptídeos. Em uma representação, o shRNA também consiste em um elemento ou uma modificação que produz o shRNA resistente à digestão de nuclease. O termo "polinucleotídeo" ou "oligonucleotídeo" é usado de modo intercambiável e cada um significa um polímero linear de monômeros de nucleotídeo. Os monômeros que constituem polinucleotídeos e oligonucleotídeos são capazes de se ligarem

especificamente a um polinucleotídeo natural e/ou artificial por meio de um padrão regular de interações de monômero com monômero, como o tipo Watson-Crick de pareamento de bases, empilhamento de bases, tipos Hoogsteen ou Hoogsteen reverso de pareamento de bases ou similares. Esses monômeros e suas ligações internucleosídicas podem ocorrer naturalmente ou podem ser seus análogos, por ex., análogos que ocorrem naturalmente ou não naturalmente. Os análogos que ocorrem não naturalmente podem incluir PNAs, LNAs, ligações internucleosídicas de fosforotioato, nucleotídeos que contêm grupos de ligação que permitem a anexação de rótulos, como fluoróforos ou haptenos e similares. Sempre que o uso de um oligonucleotídeo ou polinucleotídeo requer o processamento enzimático, como extensão por uma polimerase, ligação por uma ligase ou similar, uma pessoa de habilidade comum compreenderia que os oligonucleotídeos ou polinucleotídeos nessas instâncias não conteriam determinados análogos de ligações internucleosídicas, como metades de açúcar ou nucleotídeos em qualquer ou algumas posições. Os polinucleotídeos geralmente variam em tamanho de algumas unidades monoméricas quando são referidos como "oligonucleotídeos" a vários milhares de unidades monoméricas. Sempre que um polinucleotídeo ou oligonucleotídeo é representado por uma sequência de letras (letra maiúscula ou letra minúscula) como "ATGCCTG", será compreendido que os nucleotídeos estão na ordem 5'→3' da esquerda para a direita. Geralmente os polinucleotídeos consistem em quatro nucleotídeos naturais (por ex., desoxiadenosina, desoxicitidina, desoxiguanosina, desoxitimidina para DNA ou suas contrapartes de ribose para RNA) ligados por vínculos de fosfodiéster; entretanto, eles também podem compreender análogos de nucleotídeo não natural, por ex., incluindo nucleotídeos modificados, açúcares ou ligações internucleosídicas. Fica claro para os especialistas na arte que

onde uma enzima tem exigências específicas de substrato de oligonucleotídeo ou polinucleotídeo para atividade, por ex., DNA, RNA, RNA/DNA dúplice de filamento único, ou similares, a seleção da composição apropriada para os substratos de oligonucleotídeo ou polinucleotídeo está bem dentro do conhecimento de uma pessoa de conhecimento comum.

[0050] Em outro aspecto, os métodos da presente divulgação podem consistir na hibridização de pelo menos partes do primeiro e segundo conjuntos de sondagem nas primeira e segunda regiões de ácido nucleico de interesse nas moléculas de nucleotídeo da amostra genética, respectivamente. A hibridização das sondagens no ácido nucleico de interesse pode ser realizada simultaneamente ou após o contato das sondagens com a amostra genética, ligando, amplificando e/ou imobilizando as sondagens. Além disso, a hibridização das sondagens no ácido nucleico de interesse pode ser realizada simultaneamente ou antes da ligação amplificação e/ou imobilização das sondagens. Uma parte ou parte completa da sondagem pode hibridizar em uma parte ou parte completa da região de interesse em moléculas de nucleotídeo de filamento único ou duplo, proteína ou anticorpo em uma amostra. A região de interesse hibridizada na sondagem pode ser de 1 a 50 nucleotídeos, 50 a 1000 nucleotídeos, 100 a 500 nucleotídeos, 5, 10, 50, 100, 200 nucleotídeos ou menos ou 2, 5, 10, 50, 100, 200, 500, 1000 nucleotídeos ou mais. As sondagens podem ser desenvolvidas ou configuradas para hibridizar perfeitamente com uma região ou molécula alvo ou podem ser desenvolvidas de modo que uma incompatibilidade de base única (por ex., um local de polimorfismo de nucleotídeo único ou SNP), ou um número pequeno de incompatibilidades, falhe ao produzir um híbrido de sondagem e molécula alvo.

[0051] Em representações adicionais, a primeira sondagem de rotulação e/ou a primeira sondagem de marcação são hibridizadas

na primeira região de ácido nucleico de interesse e a segunda sondagem de rotulação e/ou a segunda sondagem de marcação são hibridizadas na segunda região de ácido nucleico de interesse. Em representações adicionais, várias ou todas as sondagens e/ou outros componentes (por ex., sondagens de rotulação, sondagens de marcação e sondagens de lacuna) de um conjunto de sondagem que são hibridizados em uma região de ácido nucleico de interesse são adjacentes entre si. Quando duas das sondagens e/ou componentes hibridizados na região de ácido nucleico de interesse são "adjacentes" ou "imediatamente adjacentes", não há nucleotídeos entre os domínios de hibridização das duas sondagens na região do ácido nucleico de interesse. Nesta representação, as diferentes sondagens em um conjunto de sondagem podem ser ligadas de maneira covalente juntas para formar uma molécula de oligonucleotídeo maior. Em outra representação, um conjunto de sondagem pode ser designado para hibridizar em uma porção não contígua, mas proximal, da região do ácido nucleico de interesse, de modo que haja uma "lacuna" de um ou mais nucleotídeos na região do ácido nucleico de interesse, entre as sondagens hibridizadas com base em um conjunto de sondagem, que não é ocupada por uma sondagem. Nesta representação, uma polimerase de DNA ou outra enzima pode ser usada para sintetizar uma nova sequência de polinucleotídeo, em alguns casos unindo de maneira covalente duas sondagens de um único conjunto de sondagem. Em um conjunto de sondagem, qualquer sondagem pode carregar um ou mais rótulos ou marcas de afinidade usados para identificação ou isolamento do locus. Sob um aspecto, a primeira e segunda sondagens de rotulação são hibridizadas na primeira e segunda regiões de ácido nucleico de interesse nas moléculas de nucleotídeo da amostra genética, respectivamente; a primeira e a segunda sondagens de marcação são hibridizadas na primeira e segunda regiões de ácido nucleico

de interesse nas moléculas de nucleotídeo da amostra genética, respectivamente; a primeira sondagem de rotulação é hibridizada em uma região adjacente na qual a primeira sondagem de marcação é hibridizada; e a segunda sondagem de rotulação é hibridizada em uma região adjacente na qual a segunda sondagem de marcação é hibridizada.

[0052] A hibridização ocorre de maneira que as sondagens em um conjunto de sondagem possam ser modificadas para formar uma nova entidade molecular maior (por ex., um produto de sondagem). As sondagens podem hibridizar nas regiões de ácido nucleico de interesse sob condições rigorosas. Conforme usado neste documento, o termo "rigor" é usado em referência às condições de temperatura, resistência iônica e presença de outros compostos, como solventes orgânicos, sob os quais as hibridizações de ácido nucleico são conduzidas. O "rigor" geralmente ocorre em um intervalo de aproximadamente  $T_m$ ° C a aproximadamente 20° C a 25° C abaixo de  $T_m$ . Uma hibridização rigorosa pode ser usada para isolar e detectar sequências idênticas de polinucleotídeo ou para isolar e detectar sequências semelhantes ou relacionadas de polinucleotídeo. Sob as "condições rigorosas", a sequência de nucleotídeo, em sua totalidade ou suas porções, hibridizarão em seu complemento exato e sequências próximas relacionadas. Condições de baixo rigor compreendem condições equivalentes à ligação ou hibridização a 68° C em uma solução que consiste em 5×SSPE (43,8 g/l NaCl, 6,9 g/l  $\text{NaH}_2\text{PO}_4 \cdot \text{H}_2\text{O}$  e 1,85 g/l EDTA, pH ajustado em 7,4 com NaOH), 0,1% SDS, 5× reagente de Denhardt (50× conteúdo de Denhardt por 500 ml: 5 g de Ficoll (Tipo 400), 5 g de BSA) e 100 µg/ml de DNA de esperma de salmão desnaturado seguido por uma lavagem em um solução que consiste em 2,0×SSPE, 0,1% SDS em temperatura ambiente quando uma sondagem de aproximadamente 100 a aproximadamente 1000 nucleotídeos de comprimento é usada. É bem conhecido na arte que várias

condições equivalentes podem ser empregadas para compreender condições de baixo rigor, fatores como o comprimento e a natureza (DNA, RNA, composição de base) da sondagem e natureza do alvo (DNA, RNA, composição de base, presente na solução ou imobilizada etc.) e a concentração dos sais e outros componentes (por ex., a presença ou ausência de formamida, sulfato de dextrano, polietileno glicol), bem como componentes da solução de hibridização podem ser variados para gerar condições de hibridização de baixo rigor diferentes das, mas equivalentes às, condições listadas anteriormente. Além disso, as condições que promovem a hibridização sob condições de rigor elevado (por ex., aumento na temperatura da hibridização e/ou etapas de lavagem, o uso de formamida na solução de hibridização etc.) também são bem conhecidas na arte. As condições de rigor elevado, quando usadas em referência à hibridização do ácido nucleico, compreendem condições equivalentes à ligação ou hibridização a 68° C em uma solução que consiste em 5×SSPE, 1% SDS, 5× reagente de Denhardt e 100 µg/ml de DNA de esperma de salmão desnaturado seguido por lavagem em uma solução que consiste em 0,1×SSPE e 0,1% SDS a 68° C quando uma sondagem de aproximadamente 100 a aproximadamente 1000 nucleotídeos de comprimento é empregada.

[0053] Em algumas representações, o produto de sondagem pode ser formado somente se as sondagens em um conjunto de sondagem forem corretamente hibridizadas. Portanto, os produtos de sondagem podem ser formados com rigor elevado e alta precisão. Novamente, os produtos de sondagem podem conter informações suficientes para identificar a sequência genômica para a qual o produto de sondagem foi designado para investigação. Portanto, a geração e a quantificação direta de um produto de sondagem particular (neste caso, por contagem molecular) pode refletir a abundância de uma sequência genética em particular na amostra de origem.

[0054] Nas representações adicionais, as regiões de ácido nucleico de interesse, nas quais as sondagens são configuradas para hibridização, estão localizadas em diferentes cromossomos. Por exemplo, uma primeira região de ácido nucleico de interesse está localizada no cromossomo 21 e a segunda região de ácido nucleico de interesse não está localizada no cromossomo 21 (por ex., localizada no cromossomo 18).

[0055] Em outro aspecto, os métodos da presente divulgação podem consistir na ligação da primeira sondagem de rotulação e na primeira sondagem de marcação e ligação da segunda sondagem de rotulação e segunda sondagem de marcação. A ligação das sondagens pode ser realizada simultaneamente ou após o contato das sondagens com a amostra genética, amplificando e/ou imobilizando as sondagens. Além disso, a ligação das sondagens pode ser realizada simultaneamente ou antes do contato das sondagens com a amostra genética, amplificando e/ou imobilizando as sondagens. A ligação aqui significa o processo de união de duas sondagens (por ex., união de duas moléculas de nucleotídeo). Por exemplo, a ligação aqui pode envolver a formação de uma união de 3', 5'-fosfodiéster que vincula dois nucleotídeos, e um agente de união que é um agente capaz de provocar a ligação pode ser uma enzima ou um elemento químico.

[0056] Em outro aspecto, os métodos da presente divulgação podem consistir na amplificação das sondagens ligadas e/ou conjuntos de sondagens ligadas. A amplificação das sondagens ligadas pode ser realizada simultaneamente ou após o contato das sondagens com a amostra genética, ligando, hibridizando e/ou imobilizando as sondagens. Além disso, a amplificação das sondagens ligadas podem ser realizadas simultaneamente ou antes da imobilização das sondagens. A amplificação aqui é definida como produção de cópias adicionais da sondagem e/ou produto de sondagem e pode ser executada com o uso de tecnologias de reação

em cadeia de polimerase bem conhecidas na arte. Conforme usado neste documento, o termo "reação em cadeia de polimerase" ("PCR") se refere a um método para aumentar a concentração de um segmento de uma sequência alvo (por ex., em uma mistura de DNA genômico) sem clonagem ou purificação. A extensão do segmento amplificado da sequência de destino desejada é determinada pelas posições relativas de dois primers de oligonucleotídeo com relação entre si e, portanto, essa extensão é um parâmetro controlável. Devido ao aspecto de repetição do processo, o método é referido como "reação em cadeia de polimerase" (de agora em diante "PCR"). Devido aos segmentos amplificados da sequência alvo se tornarem as sequências predominantes (em termos de concentração) na mistura, eles são chamados de "PCR amplificado". Com o PCR é possível amplificar uma única cópia de uma sequência alvo específica no DNA genômico para um nível detectável por diversas metodologias diferentes (por ex., hibridização com uma sondagem rotulada). Além do DNA genômico, qualquer sequência de oligonucleotídeo pode ser amplificada com o conjunto apropriado de moléculas de primer. Em particular, os segmentos amplificados criados pelo processo de PCR são modelos eficientes para amplificações subsequentes de PCR. Uma amplificação poderá ser uma amplificação em "tempo real" se estiver disponível uma química de detecção que permita que um produto da reação seja medido à medida que a reação de amplificação progride, por ex., "PCR em tempo real" ou "NASBA em tempo real" conforme descrito em Leone et al, *Nucleic Acids Research*, 26: 2150-2155 (1998).

[0057] Os primers geralmente são de filamento único para a máxima eficiência na amplificação, mas como alternativa podem ser de filamento duplo. Se for de filamento duplo, o primer geralmente será tratado primeiro para separar seus filamentos antes de ser usado para preparar os produtos de extensão. Essa

etapa de desnaturação geralmente é influenciada pelo calor, mas como alternativa pode ser executada com o uso de álcali seguido pela neutralização. Dessa forma, um "primer" é complementar a um modelo e se mistura pela ligação de hidrogênio ou hibridização com o modelo para produzir um complexo de primer/modelo para iniciação de síntese por uma polimerase que é estendida pela adição de nucleotídeos ligados de maneira covalente vinculados ao seu final 3' complementar ao modelo no processo de síntese de DNA.

[0058] Um "par de primer" conforme usado neste documento se refere a um primer avançado e um primer reverso correspondente com sequências de ácido nucleico adequadas para a amplificação de um ácido nucleico alvo baseada em ácido nucleico. Esses pares de primer geralmente incluem um primeiro primer com uma sequência que é igual ou similar à de uma primeira porção de ácido nucleico e um segundo primer com uma sequência que é complementar a uma segunda porção de um ácido nucleico para fornecer a amplificação do ácido nucleico alvo ou seu fragmento. A referência ao "primeiro" e "segundo" primers aqui é arbitrária, exceto se indicado especificamente de outra forma. Por exemplo, o primeiro primer pode ser designado como um "primer avançado" (que inicia a síntese de ácido nucleico de um final 5' do ácido nucleico alvo) ou como um "primer reverso" (que inicia a síntese de ácido nucleico de um final 5' do produto de extensão produzido a partir da síntese iniciada no primer avançado). Da mesma maneira, o segundo primer pode ser designado como um primer avançado ou um primer reverso.

[0059] Em algumas representações, a região do ácido nucleico de interesse na molécula de nucleotídeo pode ser amplificada pelos métodos de amplificação descritos aqui. Os ácidos nucleicos em uma amostra podem ser ou não amplificados antes da análise usando um método de amplificação universal (por ex.,

amplificação de todo o genoma ou todo o PCR do genoma). A amplificação da região do ácido nucleico de interesse pode ser realizada simultaneamente ou após o contato das sondagens com a amostra genética, ligando, amplificando e/ou imobilizando as sondagens. Além disso, a amplificação das sondagens ligadas pode ser realizada simultaneamente ou antes do contato das sondagens com a amostra genética, ligando as sondagens, imobilizando as sondagens e/ou contando os rótulos.

[0060] Em representações adicionais, o método exclui a amplificação das moléculas de nucleotídeo da amostra genética após a hibridização ou ligação. Em outras representações, o método exclui a amplificação das moléculas de nucleotídeo da amostra genética após a hibridização e ligação.

[0061] Em outro aspecto, os métodos da presente divulgação podem consistir na imobilização das sondagens de marcação em um local predeterminado em um substrato. A imobilização da sondagem em um substrato pode ser realizada simultaneamente ou após o contato das sondagens com a amostra genética, hibridização das sondagens na região do ácido nucleico de interesse, ligação e/ou amplificação das sondagens. Além disso, a imobilização da sondagem em um substrato pode ser realizada simultaneamente ou antes do contato das sondagens com a amostra genética, hibridização das sondagens na região do ácido nucleico de interesse, ligação, amplificação e/ou contagem das sondagens. A imobilização neste documento significa a ligação direta ou indireta das sondagens de marcação ao local predeterminado no substrato por uma união física ou química. Em algumas representações, o substrato pode consistir em um parceiro de ligação que é configurado para entrar em contato e se ligar a uma marca parcial ou completa na sondagem de marcação descrita aqui e imobilizar a marca e dessa forma a sondagem de marcação que consiste na marca. A marca da sondagem de marcação pode

consistir em um parceiro de ligação correspondente do parceiro de ligação no substrato conforme descrito neste documento.

[0062] A imobilização pode ser realizada pela hibridização de uma parte ou toda a sondagem de marcação em uma parte ou toda a parceria de ligação no substrato. Por exemplo, a etapa de imobilização consiste na hibridização de pelo menos uma parte da marca ou sequência do nucleotídeo de marcação em uma molécula de nucleotídeo correspondente imobilizada no substrato. Aqui, a molécula de nucleotídeo correspondente é um parceiro de ligação da marca ou sequência de nucleotídeo de marcação que é configurada para hibridizar parcialmente ou totalmente na marca ou sequência de nucleotídeo de marcação. Em algumas representações, os parceiros de ligação de oligonucleotídeo ou polinucleotídeo podem ser de filamento único e podem ser anexados de maneira covalente ao substrato, por exemplo, pelo final 5' ou pelo final 3'. A imobilização também pode ser executada pelos seguintes parceiros de ligação de exemplo e meio de ligação: Biotina-oligonucleotídeo misturada com Avidin, Streptavidin ou Neutravidin; SH-oligonucleotídeo ligado de maneira covalente via ligação de bissulfeto a um SH-superfície; Amina-oligonucleotídeo ligada de maneira covalente a um carboxilato ativado ou um grupo de aldeído; ácido fenilborônico (PBA)-oligonucleotídeo complexado com ácido salicílico-hidroxâmico (SHA); Acrydite-oligonucleotídeo reagido com tiol ou superfície de silano ou copolimerizado com monômero de acrilamida para formar poli(acrilamida) ou por outros métodos conhecidos na arte. Para algumas aplicações onde é preferível ter uma superfície carregada, as camadas de superfície podem ser compostas por uma estrutura de multicamada de polieletrólito (PEM) conforme mostrado na Publicação do requerimento de patente norte-americana N° 2002/025529. Em algumas representações, a imobilização pode ser realizada pelos procedimentos bem

conhecidos, por exemplo, consistindo num contato das sondagens com o suporte tendo parceiros de ligação anexados por um determinado período e após as sondagens serem esgotadas para a extensão, o suporte com os produtos de extensão imobilizada é opcionalmente enxaguado em um líquido adequado. Nas representações adicionais, a imobilização dos produtos de sondagem em um substrato podem permitir lavagem rigorosa para remover componentes da amostra biológica e o ensaio, reduzindo dessa forma o ruído de fundo e melhorando a precisão.

[0063] "Suporte sólido", "suporte", "substrato" e "suporte de fase sólida" são usados de maneira intercambiável e se referem a um material ou grupo de materiais com uma superfície ou superfícies rígidas ou semirrígidas. Em algumas representações, pelo menos uma superfície do substrato será consideravelmente plana, embora em algumas representações possa ser desejável separar fisicamente as regiões de síntese para diferentes compostos com, por exemplo, poços, regiões elevadas, pinos, valetas entalhadas ou similares. Nas representações adicionais, o substrato pode compreender pelo menos um suporte de fase sólida planar (por ex., slide de microscópio de vidro). De acordo com outras representações, os substratos assumirão a forma de esferas, resinas, géis, microesferas ou outras configurações geométricas. Sob um aspecto, de acordo com algumas representações da presente divulgação, o substrato exclui esferas, resinas, géis e/ou microesferas.

[0064] Em algumas representações, conforme mostrado na figura 1, os parceiros de ligação, as marcas, as marcas de afinidade, os rótulos, as sondagens (por ex., sondagens de marcação e sondagens de rotulação) e/ou conjuntos de sondagem descritos aqui podem ser imobilizadas em um substrato (1) como um arranjo (2). O arranjo aqui tem vários membros (3-10) que podem ter ou não uma sobreposição (6) entre os membros. Cada membro pode ter

pelo menos uma área sem sobreposição com outro membro (3-5 e 7-10). Nas representações adicionais, cada membro pode ter diferentes formatos (por ex., pontos circulares (3-8), triângulos (9) e quadrados (10)) e dimensões. Um membro de um arranjo pode ter uma área de aproximadamente 1 a  $10^7$   $\mu\text{m}^2$ , de 100 a  $10^7$   $\mu\text{m}^2$ , de  $10^3$  a  $10^8$   $\mu\text{m}^2$ , de  $10^4$  a  $10^7$   $\mu\text{m}^2$ ; de  $10^5$  a  $10^7$   $\mu\text{m}^2$ ; aproximadamente 0,0001, 0,001, 0,01, 0,1, 1, 10, 100,  $10^3$ ,  $10^4$ ,  $10^5$ ,  $10^6$ ,  $10^7$ ,  $10^8$  ou mais  $\mu\text{m}^2$ ; e/ou aproximadamente 0,001, 0,01, 0,1, 1, 10, 100,  $10^3$ ,  $10^4$ ,  $10^5$ ,  $10^6$ ,  $10^7$ ,  $10^8$  ou menos  $\mu\text{m}^2$ . Uma imagem de um membro de exemplo (8) de acordo com algumas representações da presente invenção é mostrada conforme o item 12. Além disso, dois ou mais membros que compreendem os parceiros de ligação, as marcas, as marcas de afinidade, os rótulos, as sondagens (por ex., sondagens de marcação e sondagens de rotulação) e/ou os conjuntos de sondagem do mesmo tipo podem ter a mesma forma e dimensão. Especificamente, os membros de um arranjo consistem nos parceiros de ligação, marcas, marcas de afinidade, rótulos, sondagens de marcação e/ou conjuntos de sondagem configurados ou usados para detectar a mesma variação genética ou um controle de acordo com os métodos descritos neste documento pode ter as mesmas formas e dimensões. Além disso, todo e cada membro dos arranjos no substrato podem ter as mesmas formas e dimensões. Em outras representações, os membros de um arranjo consistem nos parceiros de ligação, marcas, marcas de afinidade, rótulos, sondagens e/ou conjuntos de sondagem configurados ou usados para detectar diferentes variações genéticas e/ou controles de acordo com os métodos descritos neste documento podem ter as mesmas formas e dimensões. Além disso, cada membro do arranjo pode compreender diferentes parceiros de ligação, marcas, marcas de afinidade, rótulos, sondagens e/ou conjuntos de sondagem.

[0065] Em algumas representações, dois membros do arranjo podem ser separados por (i) uma distância na qual não haver ou haver apenas pouquíssimos parceiros de ligação, marcas, marcas de afinidade, rótulos, sondagens (por ex., sondagens de marcação e sondagens de rotulação) e/ou conjuntos de sondagem imobilizados e/ou (ii) qualquer separador que distingue um membro do outro (por ex., substrato intensificado, qualquer material que impeça a ligação dos parceiros de ligação, marcas, marcas de afinidade, sondagens (por ex., sondagens de marcação) e/ou conjuntos de sondagem no substrato e qualquer material que não seja de sondagem entre os membros). Nas representações adicionais, os membros do arranjo podem ser distinguidos entre si pelo menos por seus locais. Os membros do arranjo podem ser separados por uma distância de aproximadamente 0 a  $10^4$  microns, de 0 a  $10^3$  microns, de  $10^2$  a  $10^4$  microns ou de  $10^2$  a  $10^3$  microns; aproximadamente 0, 0,001, 0,1, 1, 2, 3, 4, 5, 10, 50, 100,  $10^3$ ,  $10^4$ ,  $10^5$ ,  $10^6$ ,  $10^7$  ou  $10^8$  microns ou mais; e/ou aproximadamente 0, 0,001, 0,1, 1, 2, 3, 4, 5, 10, 50, 100,  $10^3$ ,  $10^4$ ,  $10^5$ ,  $10^6$ ,  $10^7$  ou  $10^8$  microns ou menos. Aqui, a distância pela qual dois membros do arranjo são separados pode ser determinada pela distância mais curta entre as bordas dos membros. Por exemplo, na figura 1, a distância pela qual dois membros, itens 3 e 4, de um arranjo (2) são separados e a distância indicada pelo item n. Além disso, por exemplo, a distância mais curta pela qual os membros do arranjo (2) em um substrato (1) são separados é 0, como a distância pela qual dois membros, itens 10 e 11, do arranjo são separados. Em outras representações, dois membros do arranjo podem não ser separados e podem ser sobrepostos (6). Nessas representações, cada membro pode ter pelo menos uma área sem sobreposição com outro membro (7).

[0066] Em representações adicionais, um arranjo e os membros do arranjo dos parceiros de ligação, as marcas, as marcas de

afinidade, os rótulos, as sondagens e/ou os conjuntos de sondagem descritos aqui podem ser localizados em locais predeterminados no substrato e as formas e dimensões de cada membro do arranjo e a distância entre os membros podem ser predeterminadas antes da imobilização. O local predeterminado aqui significa um local que é determinado ou identificado antes da imobilização. Por exemplo, a forma e a dimensão de cada membro de um arranjo são determinadas ou identificadas antes da imobilização.

[0067] Em representações adicionais, o substrato pode consistir em um arranjo de parceiros de ligação, cada membro do arranjo compreendendo os parceiros de ligação, como oligonucleotídeos e polinucleotídeos, que são imobilizados (por ex., por uma ligação química que não seria quebrada durante a hibridização de sondagens nos parceiros de ligação do substrato descrito aqui) para uma região ou local definido espacialmente; ou seja, as regiões ou os locais são distintos ou separados espacialmente por uma região ou local definido no substrato. Em representações adicionais, o substrato pode compreender um arranjo, cada membro do qual compreende os parceiros de ligação que se ligam a uma região ou local definido espacialmente. Cada um dos locais definidos espacialmente configurados para compreender os parceiros de ligação podem ser adicionalmente "endereçáveis" porque seu local e a identidade de seus parceiros de ligação imobilizados são conhecidos ou predeterminados, por exemplo, antes de seu uso, análise ou anexação a seus parceiros de ligação nas sondagens de marcação e/ou conjuntos de sondagem. O termo "endereçável" com relação aos conjuntos de sondagem imobilizados no substrato significa que a sequência de nucleotídeo ou outras características físicas e/ou químicas de uma parte anexada ao final (por ex., um parceiro de ligação do parceiro de ligação do substrato, marca, marca de afinidade e

sondagem de marcação) de um conjunto de sondagem descrito aqui pode ser determinada a partir de seu endereço, isto é, uma correspondência um para um entre a sequência ou outra propriedade da parte anexada ao final do conjunto de sondagem e um local espacial em, ou característica de, substrato no qual o conjunto de sondagem é imobilizado. Por exemplo, um endereço de uma parte anexada ao final é um local espacial, por ex., as coordenadas planares de uma região particular que imobiliza cópias da parte anexada ao final do conjunto de sondagem. Entretanto, as partes anexadas ao final dos conjuntos de sondagem podem ser endereçadas de outras maneiras também, por ex., por cor, frequência de microtransponder ou similares, por ex., Chandler et al, publicação de PCT WO 97/14028, que está incorporada neste documento como referência integralmente para todos os fins. Em outras representações, os métodos descritos neste documento excluem o "microarranjo aleatório", que se refere a um microarranjo cujas regiões espacialmente distintas de parceiros de ligação (por ex., oligonucleotídeos ou polinucleotídeos) do substrato e/ou partes anexadas ao final dos conjuntos de sondagem não são endereçadas espacialmente. Ou seja, a identidade dos parceiros de ligação anexados, marca, marca de afinidade, sondagem de marcação e/ou conjuntos de sondagem não é discernível, pelo menos inicialmente, a partir de seu local. Sob um aspecto, os métodos descritos aqui excluem microarranjos aleatórios que são arranjos planares de microesferas.

[0068] Um arranjo de ácido nucleico de acordo com algumas representações da presente divulgação pode ser produzido por qualquer método bem conhecido na arte, incluindo, mas não limitado àqueles descritos na Publicação do requerimento de patente norte-americana N° 2013/0172216, que está incorporada como referência em sua totalidade para todos os fins; Schena,

Microarrays: A Practical Approach (IRL Press, Oxford, 2000). Por exemplo, um arranjo de captura de DNA pode ser usado. O arranjo de captura de DNA é um substrato sólido (por ex., um slide de vidro) com oligonucleotídeos localizados anexados de maneira covalente à superfície. Esses oligonucleotídeos podem ter um ou mais tipos na superfície e podem ser segregados geograficamente no substrato. Sob condições de hibridização, os arranjos da captura de DNA preferencialmente ligarão os alvos complementares comparados a outras metades não específicas, agindo dessa forma para localizar alvos na superfície e separá-los de espécie indesejada.

[0069] Em algumas representações, a primeira e a segunda sondagens de rotulação e/ou suas sondagens de rotulação amplificadas ligadas às sondagens de marcação imobilizada compreendem o primeiro e o segundo rótulos, respectivamente.

[0070] A sondagem de rotulação neste documento significa uma sondagem que compreende ou é configurada para se ligar a um rótulo. A própria sondagem de rotulação pode consistir em um rótulo ou pode ser modificada para compreender ou se ligar a um rótulo. A sondagem amplificada é definida neste documento como cópias adicionais de uma sondagem inicial produzida após a amplificação da sondagem inicial conforme descrito neste documento. Assim sendo, as sondagens amplificadas podem ter uma sequência que seja a sequência de nucleotídeo das sondagens iniciais e/ou sequência complementar das sequências de nucleotídeo das sondagens iniciais. As sondagens amplificadas podem conter uma sequência que é uma correspondência parcial ou completa às sequências de nucleotídeo das sondagens iniciais. Os termos "complementar" ou "complementaridade" são usados em referência a uma sequência de nucleotídeos relacionada pelas regras de pareamento de bases. Por exemplo, a sequência "5'-CAGT-3'", é complementar à sequência "5'-ACTG-3' ". A

complementaridade pode ser "parcial" ou "total". A complementaridade "parcial" é onde um ou mais nucleotídeos de ácido nucleico em uma sondagem não são correspondidos de acordo com as regras de pareamento de bases enquanto outros são correspondidos. A complementaridade "total" ou "completa" entre os ácidos nucleicos é onde toda e cada base de ácido nucleico na sondagem se corresponde com outra base sob as regras de pareamento de bases.

[0071] A sondagem imobilizada é definida neste documento como uma sondagem que se liga direta ou indiretamente ao substrato por uma ligação física ou química. Em algumas representações, uma sondagem de rotulação pode ser imobilizada em um substrato indiretamente via ligação a uma sondagem de marcação imobilizada ao substrato descrito aqui.

[0072] Um rótulo neste documento significa uma molécula, tintura, ou metade sintética, artificial que ocorre naturalmente ou não naturalmente tendo uma propriedade ou característica que é capaz de detecção e, opcionalmente, quantificação. Um rótulo pode ser diretamente detectável (por ex., radioisótopos, fluoróforos, quemiluminóforos, enzimas, partículas coloidais, substâncias fluorescentes, pontos quânticos ou outras nanopartículas, nanoestruturas, compostos de metal, rótulos organometálicos e aptâmeros de peptídeo); ou um rótulo pode ser detectável indiretamente usando parceiros de ligação específicos. Exemplos das substâncias fluorescentes incluem tinturas fluorescentes, como fluoresceína, fósforo, rodamina, derivados de tintura de polimetina e similares. Exemplos de uma substância fluorescente comercialmente disponível incluem tinturas fluorescentes, como BODYPY FL (marca registrada, produzida pela Molecular Probes, Inc.), FluorePrime (nome de produto, produzido pela Amersham Pharmacia Biotech, Inc.), Fluoredite (nome de produto, produzido pela Millipore

Corporation), FAM (produzido pela ABI Inc.), Cy 3 e Cy 5 (produzido pela Amersham Pharmacia), TAMRA (produzido pela Molecular Probes, Inc.), Pacific Blue, TAMRA, Alexa 488, Alexa 594, Alexa 647, Atto 488, Atto 590, Atto 647N e similares. "Ponto quântico" (QD) significa uma estrutura cristalina semicondutora em escala nano geralmente feita de seleneto de cádmio e absorve luz e depois a retransmite alguns nanossegundos depois em uma cor específica. Os QDs com uma variedade de superfícies conjugadas ou reativas, por ex., amino, carboxil, estreptavidina, proteína A, biotina e imunoglobulinas, também estão incluídos na presente divulgação.

[0073] Em representações adicionais, o primeiro e o segundo rótulo são diferentes de modo que os rótulos possam ser distintos entre si. Em representações adicionais, o primeiro e o segundo rótulo são diferentes em suas propriedades físicas, óticas e/ou químicas.

[0074] Em algumas representações, os rótulos imobilizados são oticamente determináveis. O termo "rótulo oticamente determinável" ou "rótulo oticamente e individualmente determinável" significa, neste documento, um grupo de rótulos que podem ser distintos entre si por sua emissão fotônica ou outras propriedades óticas, por exemplo, após imobilização, conforme descrito neste documento. Nas representações adicionais, muito embora os rótulos possam ter as mesmas propriedades de emissão ótica e/ou espectral, os rótulos imobilizados podem ser distintos entre si espacialmente. Em algumas representações, os rótulos do mesmo tipo, que são definidos como rótulos que têm as mesmas propriedades óticas, são imobilizados no substrato, por exemplo, como um membro de um arranjo descrito neste documento a uma densidade e/ou espaçamento, de modo que os produtos de sondagem individual sejam determináveis conforme mostrado no item 12 da figura 1.

Nesta divulgação, "rótulos iguais" são definidos como rótulos com composições químicas e físicas idênticas. Neste documento, "rótulos diferentes" indicam rótulos que têm composições químicas e/ou físicas diferentes, incluindo "rótulos de tipos diferentes" com propriedades óticas diferentes. Neste documento, "rótulos diferentes do mesmo tipo" significam rótulos com composições químicas e/ou físicas diferentes, mas com as mesmas propriedades óticas.

[0075] O item 12 da figura 1 representa uma imagem de um membro de exemplo de um arranjo que consiste nos rótulos imobilizados. Nestas representações, os rótulos são espacialmente endereçáveis como o local em que uma molécula especifica sua identidade (e em síntese combinatória espacial, a identidade é uma consequência de local). Nas representações adicionais, um membro do arranjo no substrato pode ter uma ou várias sondagens rotuladas imobilizadas no membro. Quando várias sondagens rotuladas são imobilizadas em um membro do arranjo, os rótulos do mesmo tipo nas sondagens rotuladas imobilizadas em um membro de um arranjo no substrato podem ser distintos entre si espacialmente conforme mostrado no item 12 da figura 1. Em algumas representações, os rótulos imobilizados do mesmo tipo são separados por uma distância de aproximadamente 1 a 1000 nm, de 5 a 100 nm ou de 10 a 100 nm; aproximadamente 100, 150, 200, 250, 300, 350 ou 400 nm ou mais; e/ou aproximadamente 50, 100, 150, 200, 250, 300, 350 ou 400 nm ou menos em todas as dimensões. A densidade dos produtos de sondagem e seus rótulos nos substratos pode ser de até muitos milhões (e até um bilhão ou mais) de produtos de sondagem a serem contados por substrato. A capacidade de contar números grandes de produtos de sondagem contendo os rótulos permite a quantificação precisa de sequências de ácido nucleico.

[0076] Em algumas representações, a primeira e a segunda sondagens de marcação imobilizadas e/ou suas sondagens de marcação amplificadas consistem na primeira e segunda marcas, respectivamente. Neste documento, sondagem de marcação significa uma sondagem que é configurada para se ligar diretamente ou indiretamente ao substrato. A própria sondagem de marcação pode se ligar ao substrato ou pode ser modificada para se ligar ao substrato. Uma marca ou marca de afinidade neste documento significa um motivo para isolamento específico, enriquecimento ou imobilização dos produtos de sondagem. Exemplos da marca ou marca de afinidade incluem um parceiro de ligação descrito neste documento, sequências de DNA exclusivas que permitem a captura específica da sequência incluindo sequência genômica natural e/ou não genômica artificial, biotina-estreptavidina, His-tags, octopeptídeo FLAG, química click (por ex., pares de grupos funcionais que reagem rápida e seletivamente entre si sob condições aquosas moderadas) e anticorpos (por ex., azida-ciclina). Por exemplo, a etapa de imobilização consiste na hibridização de pelo menos uma parte da marca, marca de afinidade ou sequência do nucleotídeo de marcação em uma molécula de nucleotídeo correspondente imobilizada no substrato. A marca ou marca de afinidade é configurada para se ligar a entidades, incluindo, mas não limitado a, uma esfera, uma esfera magnética, um slide microscópico, uma lamela, um microarranjo ou uma molécula. Em algumas representações, a etapa de imobilização é realizada imobilizando-se as marcas no local predeterminado do substrato.

[0077] Em outro aspecto, os números de diferentes rótulos imobilizados no substrato e, dessa forma, os números de diferentes produtos de sondagem imobilizada que compreendem os rótulos são contados. Por exemplo, os produtos de sondagem de cada locus genético são agrupados juntos, e os rótulos nos

produtos de sondagem imobilizada são contados. Em algumas representações, várias sequências em um locus genômico podem ser investigadas via criação de vários tipos de produto de sondagem. Para este exemplo, diferentes produtos de sondagem para o mesmo locus genômico podem ser combinados (possivelmente via imobilização em um local comum de um substrato, por ex., como um membro de um arranjo descrito neste documento), e os rótulos nestes produtos de sondagem podem ser diretamente contados. Diferentes produtos de sondagem para o mesmo locus genômico também podem ser separados (possivelmente via imobilização em localidades diferentes de um substrato, por ex., como diferentes membros de um arranjo descrito neste documento), e os rótulos nestes produtos de sondagem podem ser diretamente contados. Nas representações adicionais, o substrato pode ter uma ou mais marca de afinidade específica em cada localidade em um substrato, por ex., como um membro de um arranjo no substrato. Portanto, outro método para quantificar sequências de ácido nucleico ocorre via imobilização de produtos de sondagem para um único locus genômico (esse pode ser um tipo de produto de sondagem ou pode ser um conjunto de mais de um produto de sondagem para um locus genômico em particular) na mesma localidade de um substrato (por ex., como o mesmo membro de um arranjo descrito neste documento) como produtos de sondagem correspondentes a um segundo locus genômico que pode servir ou não como uma referência ou locus de controle. Neste caso, os produtos de sondagem do primeiro locus genômico será distinguível dos produtos de sondagem do segundo locus genômico, com base na presença de diferentes rótulos usados na geração dos produtos de sondagem.

[0078] Em um exemplo, para detectar trissomia 21 (aneuploidia) de um feto por meio do exame de uma amostra sanguínea materna, um conjunto de produtos de sondagem

correspondente ao cromossomo 21 seria gerado, por exemplo, com um rótulo de fluoróforo vermelho e contado. Um segundo conjunto de produtos também seria gerado a partir de uma referência ou locus de controle, por exemplo, cromossomo 18 e contado. Esse segundo conjunto de produtos de sondagem pode ser gerado, por exemplo, com um rótulo de fluoróforo verde.

[0079] Em algumas representações, esses produtos de sondagem podem ser preparados de modo que sejam agrupados por locus (neste caso o cromossomo 21 ou o cromossomo 18) e contados separadamente em um substrato. Ou seja, os produtos de sondagem correspondentes ao cromossomo 21 podem ser isolados e contados separadamente e os produtos de sondagem correspondentes ao cromossomo 18 podem ser isolados e contados separadamente. Em representações adicionais, esses produtos de sondagem também podem ser preparados de maneira que sejam agrupados na mesma localidade de um substrato (por ex., como o mesmo membro de um arranjo descrito neste documento). Neste caso, na mesma região de um substrato, os produtos de sondagem que carregam um fluoróforo vermelho corresponderão ao cromossomo 21, e os produtos de sondagem com um fluoróforo verde corresponderão ao cromossomo 18. Por exemplo, como todos esses produtos de sondagem podem ser determinados individualmente e, portanto, podem ser contados de maneira muito precisa, uma frequência maior dos produtos de sondagem de cromossomo 21 relativa aos produtos de sondagem de cromossomo 18 (mesmo de 0,01, 0,1, um ou mais por cento ou menos) significarão a presença de trissomia 21 em um feto. Neste caso, os produtos de sondagem para o cromossomo 18 podem servir como um controle.

[0080] Em outro aspecto, os métodos da presente divulgação podem consistir na contagem de rótulos dos conjuntos de sondagem imobilizados no substrato. Em algumas representações, os métodos podem compreender a contagem (i) um primeiro número do primeiro

rótulo imobilizado no substrato e (ii) um segundo número do segundo rótulo imobilizado no substrato. A etapa de contagem pode ser realizada após a imobilização do conjunto de sondagem ligado a um substrato, e o substrato com os conjuntos de sondagem ligados imobilizados podem ser armazenados em uma condição para evitar a degradação dos conjuntos de sondagem ligados (por ex., em temperatura ambiente ou em temperatura abaixo da temperatura ambiente) antes da realização da etapa de contagem.

[0081] Para quantificar de modo preciso a abundância relativa de diferentes sequências genômicas, por exemplo, para quantificação de número de cópia de DNA ou para quantificação de frequência de alelo, um grande número de produtos de sondagem pode ser contado. Por exemplo, um rótulo pode ser detectado e contado com base na medição, por exemplo, propriedades físico-químicas, eletromagnéticas, elétricas, optoeletrônicas ou eletroquímicas ou características do rótulo imobilizado.

[0082] Em algumas representações, o rótulo pode ser detectado por microscopia de sondagem por escaneamento (SPM), microscopia de tunelamento por escaneamento (STM) e microscopia de força atômica (AFM), microscopia de elétrons, técnicas de investigação/detecção ótica incluindo, mas não limitado a, microscopia ótica de escaneamento de campo próximo (NSOM), microscopia confocal e excitação de onda evanescente. Versões mais específicas dessas técnicas incluem microscopia confocal de campo próximo, microscopia de dois fótons, epi-iluminação de campo selvagem e microscopia de reflexão interna total (TIR). Muitas das técnicas acima também podem ser usadas em um modo espectroscópico. A detecção real é feita por câmeras de dispositivo de carga acoplada (CCD) e CCDs intensificados, fotodiodos e/ou tubos fotomultiplicadores. Em algumas representações, a etapa de contagem consiste em uma análise

ótica que detecta uma propriedade ótica de um rótulo. Nas representações adicionais, a análise ótica consiste em uma análise de imagem conforme descrito neste documento.

[0083] Em outro aspecto, a etapa de contagem consiste na leitura do substrato no primeiro e segundo canais de imagem que correspondem ao primeiro e segundo rótulos, respectivamente, e produção de uma ou mais imagens do substrato em que a primeira e segunda sondagem de rotulação podem ser determinadas em uma ou mais imagens. Em algumas representações, a etapa de contagem consiste na filtragem espacial para segmentação da imagem. Nas representações adicionais, a etapa de contagem consiste na análise de derramamento de água ou em um método híbrido para segmentação da imagem.

[0084] Os métodos descritos neste documento também podem examinar a frequência de alelos diferentes no mesmo locus genético (por ex., dois alelos de polimorfismos de nucleotídeo único). A precisão desses métodos pode detectar mudanças muito pequenas na frequência (por ex.; aproximadamente 10, 5, 4, 3, 2, 1, 0,5, 0,1 ou 0,01 % ou menos). Como um exemplo, no caso de transplante de órgão, uma amostra sanguínea conterá uma assinatura genética muito fraca do órgão doado. Essa assinatura pode ser a presença de um alelo que não está no recipiente do genoma do órgão doado. Os métodos descritos neste documento podem detectar desvios muito pequenos na frequência de alelo (por ex., aproximadamente 10, 5, 4, 3, 2, 1, 0,5, 0,1 ou 0,01 % ou menos) e pode identificar a presença do DNA do doador em uma amostra do hospedeiro (por ex., amostra sanguínea). Um órgão não saudável transplantado pode resultar em níveis elevados de DNA do doador no sangue do hospedeiro - uma elevação de apenas uma pequena porcentagem (por ex., aproximadamente 10, 5, 4, 3, 2, 1, 0,5, 0,1 ou 0,01 % ou menos). Os métodos descritos neste documento podem ser suscetíveis o suficiente para identificar

alterações na frequência de alelo com a sensibilidade necessária e, portanto, podem determinar de maneira precisa a presença e as quantidades variáveis de DNA do doador no sangue do hospedeiro.

[0085] Em outro aspecto, os métodos da presente divulgação podem consistir na comparação do primeiro e segundo números para determinar a variação genética na amostra genética. Em algumas representações, a etapa de comparação consiste na obtenção da estimativa de um número relativo das moléculas de nucleotídeo tendo a primeira e a segunda regiões de ácido nucleico de interesse.

[0086] Em outro aspecto, os métodos da presente divulgação podem compreender a rotulação da primeira e segunda sondagens de rotulação com o primeiro e o segundo rótulo, respectivamente, antes da etapa de contato (por ex., durante a fabricação das sondas). A rotulação da sondagem pode ser realizada simultaneamente ou após o contato das sondagens com a amostra genética, hibridizando, ligando, amplificando e/ou imobilizando as sondagens. Além disso, a rotulação da sondagem pode ser realizada simultaneamente ou antes do contato das sondagens com a amostra genética, hibridizando, ligando, amplificando e/ou imobilizando as sondagens. A rotulação da sondagem pode consistir na adição, imobilização ou ligação de um rótulo na sondagem por uma ligação física ou química. Os rótulos podem ser posicionados em qualquer lugar na sequência de uma sondagem, incluindo o final 5' ou 3'.

[0087] Em outro aspecto, os métodos da presente divulgação podem compreender a marcação da primeira e segunda sondagens de marcação com a primeira e a segunda marca, respectivamente, antes da etapa de contato. (por ex., durante a fabricação das sondas). A marcação da sondagem pode ser realizada simultaneamente ou após o contato das sondagens com a amostra genética, hibridizando, ligando, amplificando e/ou rotulando as

sondagens. Além disso, a marcação da sondagem pode ser realizada simultaneamente ou antes do contato das sondagens com a amostra genética, hibridizando, ligando, amplificando, imobilizando e/ou rotulando as sondagens. A marcação de uma sondagem pode consistir na adição, imobilização ou ligação de uma marca na sondagem por uma ligação física ou química. As marcas rótulos podem ser posicionados em qualquer lugar na sequência de uma sondagem, incluindo o final 5' ou 3'.

[0088] Em outro aspecto, os conjuntos de sondagem neste documento podem ser desenvolvidos para ter as marcas de acordo com a localização predeterminada na qual as marcas devem ser imobilizadas. Em algumas representações, as marcas em todos os conjuntos de sondagem configurados para detectar uma variação genética são iguais e são configuradas para serem imobilizadas nas mesmas localidades no substrato direta ou indiretamente. Em representações adicionais, a primeira e segunda marcas são iguais e cada uma do restante das marcas é diferente da primeira ou da segunda marca. Em outras representações, cada membro ou grupo de membros do arranjo de várias localidades predeterminadas em um substrato podem ter uma marca exclusiva para ser imobilizada.

[0089] Em outro aspecto, os conjuntos de sondagem de acordo com algumas representações podem ser amplificados e os conjuntos de sondagem rotulados podem ser produzidos durante o processo de amplificação. Em outro aspecto, cada uma das sondagens de rotulação pode consistir em uma sequência de preparação de avanço ou reversa e cada uma das sondagens de marcação pode consistir em uma sequência de preparação reversa ou de avanço correspondente e uma sequência de nucleotídeo de marcação como uma marca. As sequências de preparação de avanço e reversa são as sequências configuradas para hibridizar nos primers avançados e reverso correspondentes, respectivamente. Em algumas

representações, a etapa de amplificação consiste na amplificação (i) das primeiras sondagens de rotulação e marcação ligadas com os primeiros primers avançados e reverso que são hibridizados nas sequências de preparação de avanço e reversa, respectivamente, em que o primeiro primer avançado ou reverso que é hibridizado na primeira sondagem de rotulação consiste no primeiro rótulo e (ii) as segundas sondagens de rotulação e marcação ligadas com o segundos primers avançados e reverso que são hibridizados nas sequências de preparação de avanço e reversa, respectivamente, em que o segundo primer avançado ou reverso que é hibridizado na segunda sondagem de rotulação compreende o segundo rótulo. Nas representações adicionais, as sequências de nucleotídeos de marcação amplificadas das sondagens de marcação são immobilizadas em uma localidade predeterminada em um substrato, no qual as sequências de nucleotídeos de marcação amplificada da primeira e segunda sondagem de marcação são a primeira e segunda marcas. Em algumas representações, a primeira e segunda marcas são iguais e/ou são configuradas para se ligarem à mesma localidade no substrato. Em outra representação, a primeira e a segunda marca são diferentes e/ou são configuradas para se ligarem a diferentes localidades no substrato. Em outras representações, quando as sondagens são amplificadas, o método consiste nos números de contagem dos rótulos nas sondagens amplificadas e/ou conjuntos de sondagem immobilizados no substrato. Por exemplo, o primeiro número é o número do primeiro rótulo no primeiro conjunto de sondagem amplificada immobilizado no substrato e o segundo número é o número do segundo rótulo no conjunto de sondagem amplificada immobilizado no substrato.

[0090] Em outro aspecto, os conjuntos de sondagem, de acordo com algumas representações, podem ser amplificados, e os conjuntos de sondagem rotulados podem ser produzidos com o uso

de primers reversos rotulados sem o uso de um primer avançado. Em outro aspecto, cada uma das sondagens de rotulação pode consistir em uma sequência de preparação reversa e cada uma das sondagens de marcação pode consistir em uma sequência de nucleotídeo de marcação como uma marca. Em algumas representações, a etapa de amplificação pode compreender a amplificação (i) das primeiras sondagens de rotulação e marcação ligadas com um primeiro primer reverso que é hibridizado em uma primeira sequência de preparação reversa da primeira sondagem de rotulação, em que o primeiro primer reverso compreende o primeiro rótulo e (ii) as segundas sondagens de rotulação e marcação ligadas com um segundo primer reverso que é hibridizado em uma segunda sequência de preparação reversa da segunda sondagem de rotulação, em que o segundo primer reverso consiste no segundo rótulo. Nas representações adicionais, as sequências de nucleotídeos de marcação amplificadas das sondagens de marcação são imobilizadas em uma localidade predeterminada em um substrato, no qual as sequências de nucleotídeos de marcação amplificada da primeira e segunda sondagem de marcação são a primeira e segunda marcas. Em outras representações, o primeiro número é o número do primeiro rótulo no primeiro conjunto de sondagem amplificada imobilizado no substrato e o segundo número é o número do segundo rótulo no conjunto de sondagem amplificada imobilizado no substrato.

[0091] Em outro aspecto, os conjuntos de sondagem ligados de acordo com algumas representações podem ser produzidos com o uso de uma reação em cadeia de ligase. Em outro aspecto, o método descrito neste documento consiste no contato de terceiro e quarto conjuntos de sondagem com a amostra genética, na qual o terceiro conjunto de sondagem consiste em uma terceira sondagem de rotulação e uma terceira sondagem de marcação, e o quarto conjunto de sondagem consiste em uma quarta sondagem de

rotulação e uma quarta sondagem de marcação. O método pode ainda consistir na hibridização do primeiro e do segundo conjunto de sondagem no primeiro e segundo filamentos de ácido nucleico de sentido de interesse em moléculas de nucleotídeo de filamento único a partir de moléculas de nucleotídeo de filamento duplo da amostra genética, respectivamente; e hibridização do terceiro e quarto conjuntos de sondagem nos filamentos de ácido nucleico antissentido do primeiro e segundo filamentos de ácido nucleico de sentido de interesse, respectivamente. O método pode ainda consistir na produção de primeiro, segundo, terceiro e quarto conjuntos de sondagens ligados pelo menos ligando (i) a primeira sondagem de rotulação e a primeira sondagem de marcação, (ii) a segunda sondagem de rotulação e a segunda sondagem de marcação, (iii) a terceira sondagem de rotulação e a terceira sondagem de marcação e (iv) a quarta sondagem de rotulação e a quarta sondagem de marcação. O método pode ainda consistir na realização de uma reação em cadeia de ligase conhecida na arte para amplificar a sondagem ligada e/ou os conjuntos de sondagem ligados. Em algumas representações, a reação em cadeia da ligase pode compreender a hibridização do primeiro, segundo, terceiro e quarto conjuntos de sondagem não ligada no terceiro, quarto, primeiro e segundo conjuntos de sondagem ligada, respectivamente, e ligando pelo menos (i) a primeira sondagem de rotulação e a primeira sondagem de marcação, (ii) a segunda sondagem de rotulação e a segunda sondagem de marcação, (iii) a terceira sondagem de rotulação e a terceira sondagem de marcação e (iv) a quarta sondagem de rotulação e a quarta sondagem de marcação dos conjuntos de sondagem não ligados. O método pode ainda consistir na imobilização das sondagens de marcação na localidade predeterminada em um substrato, no qual a primeira, segunda, terceira e quarta sondagens de rotulação ligadas à primeira, segunda, terceira e quarta sondagens de marcação

imobilizadas, respectivamente, consistem no primeiro, segundo, terceiro e quarto rótulos, respectivamente; os rótulos imobilizados são oticamente determináveis; a primeira, segunda, terceira e quarta sondagens de marcação imobilizadas consistem na primeira, segunda, terceira e quarta marcas, respectivamente, e a etapa de imobilização é realizada pela imobilização das marcas na localidade predeterminada. O método pode também consistir na contagem (i) da primeira soma do primeiro e terceiro rótulos imobilizados no substrato e (ii) da segunda soma do segundo e quarto rótulos imobilizados no substrato e comparação da primeira e segunda somas para determinar a variação genética na amostra genética. Ainda em representações adicionais, o método também consiste na rotulação da primeira, segunda, terceira e quarta sondagens de rotulação com o primeiro, segundo, terceiro e quarto rótulos, respectivamente, antes da etapa de contato. Ainda em outras representações, o primeiro e o terceiro rótulo são iguais, e o segundo e quarto rótulo também são iguais.

[0092] Em outro aspecto, o método descrito neste documento consiste no contato do terceiro e quarto conjuntos de sondagem com a amostra genética, onde o terceiro conjunto de sondagem consiste em uma terceira sondagem de rotulação e uma terceira sondagem de marcação, e o quarto conjunto de sondagem consiste em uma quarta sondagem de rotulação e uma quarta sondagem de marcação, a primeira e a terceira sondagens de rotulação consistem em uma primeira sequência de preparação reversa, a segunda e a quarta sondagens de rotulação consistem em uma segunda sequência de preparação reversa e cada uma das sondagens de marcação consiste em uma sequência de nucleotídeo de marcação como uma marca. O método pode compreender a hibridização do primeiro e segundo conjuntos de sondagem em um primeiro e segundo filamentos de ácido nucleico de interesse,

respectivamente, nas moléculas de nucleotídeo de filamento único a partir das moléculas de nucleotídeo de filamento duplo da amostra genética; e a hibridização de pelo menos partes do terceiro e quarto conjuntos de sondagem nos filamentos de ácido nucleico antissentido do primeiro e segundo filamentos de ácido nucleico de sentido de interesse, respectivamente; produzindo o primeiro, o segundo, o terceiro e o quarto conjuntos de sondagem ligados pela ligação (i) da segunda sondagem de rotulação e a segunda sondagem de marcação, (iii) a terceira sondagem de rotulação e a terceira sondagem de marcação e (iv) a quarta sondagem de rotulação e a quarta sondagem de marcação. O método pode ainda consistir na execução de uma reação em cadeia de ligase. Em algumas representações, a reação em cadeia da ligase consiste na hibridização de pelo menos partes do primeiro, segundo, terceiro e quarto conjuntos de sondagem não ligada no terceiro, quarto, primeiro e segundo conjuntos de sondagem ligada, respectivamente, e ligando (i) a primeira sondagem de rotulação e a primeira sondagem de marcação, (ii) a segunda sondagem de rotulação e a segunda sondagem de marcação, (iii) a terceira sondagem de rotulação e a terceira sondagem de marcação e (iv) a quarta sondagem de rotulação e a quarta sondagem de marcação do conjunto de sondagem não ligado. O método pode ainda consistir na amplificação (i) do primeiro e terceiro conjuntos de sondagem ligados com um primeiro primer reverso que é hibridizado na primeira sequência de preparação reversa, na qual o primeiro primer reverso consiste no primeiro rótulo e (ii) no segundo e quarto conjuntos de sondagem ligados com um segundo primer reverso que é hibridizado na segunda sequência de preparação reversa, na qual o segundo primer reverso consiste no segundo rótulo, as sequências de nucleotídeo de marcação amplificadas das sondagens de marcação são imobilizadas em uma localidade predeterminada em um substrato, no qual as sequências

de nucleotídeo de marcação amplificadas da primeira, segunda, terceira e quarta sondagens de marcação são a primeira, segunda, terceira e quarta marcas, o primeiro número é o número do primeiro rótulo no primeiro e terceiro conjuntos de sondagem immobilizados no substrato e o segundo número é o número do segundo rótulo no segundo e quarto conjuntos de sondagem amplificados immobilizados no substrato.

[0093] Em outro aspecto, a primeira e a segunda sondagem de rotulação ligadas estão no final 3' do primeiro e segundo conjunto de sondagem ligados e consistem na primeira e segunda sequências de preparação reversas hibridizadas no primeiro e segundo primers reversos, respectivamente. Em algumas representações, o primeiro e o segundo primer reverso consistem no primeiro e segundo rótulos. Nas representações adicionais, a primeira e segunda sondagens de marcação ligadas estão no final 5' do primeiro e segundo conjunto de sondagem ligados. Em outras representações, a primeira e a segunda sondagens de marcação ligadas estão no final 5' do primeiro e segundo conjunto de sondagem ligados e consistem na primeira e segunda sequências de preparação de avanço correspondentes hibridizadas no primeiro e segundo primers avançados, respectivamente.

[0094] Em outro aspecto, o método deste documento consiste na digestão de moléculas de filamento duplo na amostra para produzir moléculas de filamento único. Em algumas representações, a etapa de amplificação consiste no contato de uma exonuclease com a sondagem e/ou conjunto de sondagem amplificada e digestão da sondagem e/ou conjunto de sondagem amplificada do final 5' da sondagem e/ou conjunto de sondagem amplificada de um filamento ou de filamento duplo. Por exemplo, a etapa de amplificação consiste no contato de uma exonuclease com a sondagem amplificada em um conjunto de sondagem e digestão do conjunto de sondagem amplificada do final 5' do conjunto de

sondagem amplificada de um filamento ou de filamento duplo. Em representações adicionais, o filamento da sondagem e/ou conjunto de sondagem amplificada em contato com a exonuclease não tem nenhum rótulo no final 5'. O contato da exonuclease com as sondagens de filamento duplo não rotulados podem digerir o filamento não rotulado do final 5' produzindo sondagens de filamento único. Em outro aspecto, o final 5' do conjunto de sondagem amplificada que consiste no rótulo no final 5' pode ser protegido contra a digestão de exonuclease.

[0095] Em outro aspecto, o método pode detectar de 1 a 100, de 1 a 50, de 2 a 40 ou de 5 a 10 variações genéticas; 2, 3, 4, 5, 6, 7, 8, 9, 10 ou mais variações genéticas; e 100, 50, 30, 20, 10 ou menos variações genéticas. Em algumas representações, o método descrito neste documento pode detectar um número  $x$  de variações genéticas usando pelo menos o número  $(x+1)$  de diferentes conjuntos de sondagem. Nessas representações, vários rótulos de um tipo de conjunto de sondagem podem ser comparados com um ou mais números de rótulos do restante dos diferentes tipos de conjuntos de sondagem. Em algumas representações, o método descrito neste documento pode detectar a variação genética de maneira contínua em todo o genoma em várias resoluções, por exemplo, em 300.000 resoluções de base de modo que 100 variações distribuídas em todos os cromossomos sejam separadamente investigadas e quantificadas. Nas representações adicionais, a resolução de base está no intervalo de um ou dez a 100 mil nucleotídeos até um milhão, dez milhões ou 100 milhões ou mais nucleotídeos.

[0096] Em outro aspecto, o método, de acordo com algumas representações, pode detectar pelo menos duas variações genéticas. Em algumas representações, o método descrito neste documento pode ainda consistir no contato de um quinto conjunto de sondagem com a amostra genética em que o quinto conjunto de

sondagem consiste em uma quinta sondagem de rotulação e uma quinta sondagem de marcação. O método pode ainda consistir na hibridização de pelo menos uma parte do quinto conjunto de sondagem na terceira região de ácido nucleico de interesse nas moléculas de nucleotídeo da amostra genética, em que a terceira região de ácido nucleico de interesse é diferente da primeira e segunda regiões de ácido nucleico de interesse. O método pode também consistir na ligação do quinto conjunto de sondagem pelo menos ligando a quinta sondagem de rotulação e quinta sondagem de marcação. O método pode consistir na amplificação dos conjuntos de sondagem ligados. O método pode ainda consistir na imobilização de cada uma das sondagens de marcação em uma localidade predeterminada em um substrato, no qual a quinta sondagem de rotulação e/ou sua sondagem de rotulação amplificada ligada à sondagem de marcação imobilizada consiste em um quinto rótulo, o quinto rótulo é diferente do primeiro e segundo rótulos, os rótulos imobilizados são óticamente determináveis, a quinta sondagem de marcação imobilizada e/ou sua sondagem de marcação amplificada consiste em uma quinta marca, e a etapa de imobilização é executada pela imobilização das marcas na localidade predeterminada. O método pode consistir na contagem de um terceiro número do quinto rótulo imobilizado para o substrato e na comparação do terceiro número para o primeiro e/ou segundo número(s) para determinar a segunda variação genética na amostra genética. Em algumas representações, a paciente pode ser uma paciente grávida, a primeira variação genética é trissomia 21 no feto da paciente grávida e a segunda variação genética é selecionada no grupo que consiste em trissomia 13, trissomia 18, aneuploidia de X e aneuploidia de Y no feto da paciente grávida.

[0097] Em outro aspecto, o método, de acordo com algumas representações, pode detectar pelo menos três variações

genéticas. Em algumas representações, o método descrito neste documento consiste ainda no contato de um sexto conjunto de sondagem com a amostra genética em que o sexto conjunto de sondagem consiste em uma sexta sondagem de rotulação e uma sexta sondagem de marcação. O método pode ainda consistir na hibridização de pelo menos uma parte do sexto conjunto de sondagem na quarta região de ácido nucleico de interesse nas moléculas de nucleotídeo da amostra genética, em que a quarta região de ácido nucleico de interesse é diferente da primeira, segunda e terceira regiões de ácido nucleico de interesse. O método pode também consistir na ligação do sexto conjunto de sondagem pelo menos ligando a sexta sondagem de rotulação e sexta sondagem de marcação. O método pode consistir na amplificação dos conjuntos de sondagem ligados. O método pode ainda consistir na imobilização de cada uma das sondagens de marcação em uma localidade predeterminada em um substrato, no qual a sexta sondagem de rotulação e/ou sua sondagem de rotulação amplificada ligada à sondagem de marcação imobilizada consiste em um sexto rótulo, o sexto rótulo é diferente do primeiro e segundo rótulos, os rótulos imobilizados são oticamente determináveis, a sexta sondagem de marcação imobilizada e/ou sua sondagem de marcação amplificada consiste em uma sexta marca, e a etapa de imobilização é executada pela imobilização das marcas na localidade predeterminada. O método pode ainda consistir na contagem de um quarto número do sexto rótulo imobilizado para o substrato e na comparação do quarto número com o primeiro e/ou terceiro número para determinar a terceira variação genética na amostra genética.

[0098] Em outro aspecto, o método, de acordo com algumas representações, pode detectar pelo menos quatro variações genéticas. Em algumas representações, o método descrito neste documento consiste ainda no contato de um sétimo conjunto de

sondagem com a amostra genética em que o sétimo conjunto de sondagem consiste em uma sétima sondagem de rotulação e sétima sexta sondagem de marcação. O método pode ainda consistir na hibridização de pelo menos uma parte do sétimo conjunto de sondagem na quinta região de ácido nucleico de interesse nas moléculas de nucleotídeo da amostra genética, em que a quinta região de ácido nucleico de interesse é diferente da primeira, segunda, terceira e quarta regiões de ácido nucleico de interesse. O método pode também consistir na ligação do sétimo conjunto de sondagem pelo menos ligando a sétima sondagem de rotulação e sétima sondagem de marcação. O método pode consistir opcionalmente na amplificação dos conjuntos de sondagem ligados. O método pode ainda consistir na imobilização de cada uma das sondagens de marcação em uma localidade predeterminada em um substrato, no qual a sétima sondagem de rotulação e/ou sua sondagem de rotulação amplificada ligada à sondagem de marcação imobilizada consiste em um sétimo rótulo, o sétimo rótulo é diferente do primeiro e segundo rótulos, os rótulos imobilizados são ópticamente determináveis, a sétima sondagem de marcação imobilizada e/ou sua sondagem de marcação amplificada consiste em uma sétima marca, e a etapa de imobilização é executada pela imobilização das marcas na localidade predeterminada. O método pode ainda consistir na contagem de um quinto número do sétimo rótulo imobilizado para o substrato e na comparação do quinto número com o primeiro, segundo, terceiro e/ou quarto número(s) para determinar a quarta variação genética na amostra genética.

[0099] Em outro aspecto, o método, de acordo com algumas representações, pode detectar pelo menos cinco variações genéticas. Em algumas representações, o método descrito neste documento consiste ainda no contato de um oitavo conjunto de sondagem com a amostra genética em que o oitavo conjunto de sondagem consiste em uma oitava sondagem de rotulação e oitava

sondagem de marcação. O método pode ainda consistir na hibridização de pelo menos uma parte do oitavo conjunto de sondagem na sexta região de ácido nucleico de interesse nas moléculas de nucleotídeo da amostra genética, em que a sexta região de ácido nucleico de interesse é diferente da primeira, segunda, terceira, quarta e quinta regiões de ácido nucleico de interesse. O método pode também consistir na ligação do oitavo conjunto de sondagem pelo menos ligando a oitava sondagem de rotulação e oitava sondagem de marcação. O método pode consistir na amplificação dos conjuntos de sondagem ligados. O método pode ainda consistir na imobilização de cada uma das sondagens de marcação em uma localidade predeterminada em um substrato, no qual a oitava sondagem de rotulação e/ou sua sondagem de rotulação amplificada ligada à sondagem de marcação imobilizada consiste em um oitavo rótulo, o oitavo rótulo é diferente do primeiro e segundo rótulos, os rótulos imobilizados são ópticamente determináveis, a oitava sondagem de marcação imobilizada e/ou sua sondagem de marcação amplificada consiste em uma oitava marca, e a etapa de imobilização é executada pela imobilização das marcas na localidade predeterminada. O método pode ainda consistir na contagem de um sexto número do oitavo rótulo imobilizado para o substrato e na comparação do sexto número com o primeiro, segundo, terceiro, quarto e/ou quinto número(s) para determinar a quinta variação genética na amostra genética. Em algumas representações, o paciente é uma paciente grávida, e a primeira, a segunda, a terceira, a quarta e a quinta variações genéticas são trissomia 13, trissomia 18, trissomia 21, aneuploidia X e aneuploidia Y no feto da paciente grávida.

[00100] Em outro aspecto, o paciente é uma paciente grávida, a variação genética ser trissomia 21 no feto da paciente grávida, a primeira região de ácido nucleico de interesse está localizada

no cromossomo 21 e a segunda região de ácido nucleico de interesse não está localizada no cromossomo 21.

[00101] Em outro aspecto, o paciente é uma paciente grávida, a variação genética é trissomia 21 no feto da paciente grávida, a primeira região de ácido nucleico de interesse está localizada no cromossomo 21 e a segunda região de ácido nucleico de interesse está localizada no cromossomo 18.

[00102] Em um aspecto, o conjunto de sondagem deste documento pode consistir em dois, três, quatro, cinco ou mais sondagens de rotulação e/ou dois, três, quatro, cinco ou mais rótulos. Em algumas representações, o método descrito neste documento pode ainda compreender o primeiro e o segundo conjuntos de sondagem e consistir no terceiro e quarto conjuntos de rotulação, respectivamente, o primeiro conjunto de sondagem imobilizada e/ou primeiro conjunto de sondagem amplificada consistir ainda em um nono rótulo na terceira sondagem de rotulação e/ou seu produto amplificado e o segundo conjunto de sondagem imobilizada e/ou segundo conjunto de sondagem amplificada consistir ainda em um décimo rótulo na quarta sondagem de rotulação e/ou seu produto amplificado. Nessas representações, se o nono e décimo rótulos forem diferentes do primeiro e segundo rótulos, esse método poderá ser usado para confirmar o número contado para o primeiro e segundo rótulos. Se o nono e o décimo rótulos forem iguais ao primeiro e segundo rótulos, respectivamente, este método poderá ser usado para melhorar a precisão dos rótulos de detecção imobilizados em cada uma das regiões do ácido nucleico de interesse. Por exemplo, o uso de vários rótulos seria mais claro do que o uso de um rótulo e, portanto, vários rótulos podem ser mais facilmente detectados que um rótulo.

[00103] Em representações adicionais, (i) o primeiro conjunto de sondagem imobilizada e/ou primeiro conjunto de sondagem amplificada consistem ainda em um décimo primeiro rótulo na

sondagem de rotulação e (ii) pelo segundo conjunto de sondagem imobilizada e/ou segundo conjunto de sondagem amplificada ainda consistir em uma décima segunda rotulação que é diferente do décimo primeiro rótulo na sondagem de rotulação. Em outras representações, em que o primeiro, segundo, décimo primeiro e décimo segundo rótulos são diferentes entre si e a etapa de contagem ainda consiste nos números de contagem do décimo primeiro e décimo segundo rótulos imobilizados no substrato.

[00104] Em outro aspecto, o método descrito neste documento pode ser executado com uma amostra de controle. Em algumas representações, o método pode ainda consistir na repetição das etapas com uma amostra de controle diferente da amostra genética do paciente. O método pode ainda consistir na contagem de números de controle dos rótulos imobilizados no substrato e na comparação dos números de controle com o primeiro, segundo, terceiro, quarto, quinto e/ou sexto número para confirmar a variação genética na amostra genética.

[00105] Em outro aspecto, a paciente pode ser uma paciente grávida e a variação genética é uma variação genética no feto da paciente grávida. Nessas representações, o método pode usar um local de polimorfismo de nucleotídeo único (SNP) para determinar se a proporção (por ex., concentração e porcentagem de número com base no número de moléculas de nucleotídeo na amostra) do material fetal (por ex., fração fetal) é suficiente para que a variação genética do feto possa ser detectada com base em uma amostra da paciente grávida com um significância estatística razoável. Nas representações adicionais, o método pode consistir ainda no contato de conjuntos de sondagem materna e paterna com a amostra genética, em que o conjunto de sondagem materna compreende uma sondagem de rotulação materna, e o conjunto de sondagem paterna consiste em uma sondagem de rotulação paterna e uma sondagem de marcação paterna. O método pode ainda consistir

na hibridização de pelo menos uma parte de cada um dos conjuntos de sondagem materna e paterna em uma região de ácido nucleico de interesse nas moléculas de nucleotídeo da amostra genética, na região de ácido nucleico de interesse consistindo em um local de SNP predeterminado no qual pelo menos uma parte de um conjunto de sondagem materna é hibridizado em um primeiro alelo no local de SNP, pelo menos uma parte do conjunto de sondagem paterna é hibridizada em um segundo alelo no local de SNP, e o primeiro e segundo alelos são diferentes entre si. O método pode ainda consistir na ligação dos conjuntos de sondagem materna e paterna pelo menos por ligação (i) das sondagens de rotulação e marcação maternas e (ii) das sondagens de rotulação e marcação paterna. O método pode também consistir na amplificação das sondagem ligadas. O método pode ainda consistir na imobilização das sondagens de marcação em uma localidade predeterminada em um substrato, em que as sondagens de rotulação materna e paterna e/ou suas sondagens de rotulação amplificadas ligadas às sondagens de marcação imobilizadas consistem nos rótulos materno e paterno, respectivamente; os rótulos materno e paterno são diferentes, e os rótulos imobilizados são ópticamente determináveis. O método pode ainda consistir na contagem de números de rótulos materno e paterno e determinar se uma proporção de um material fetal na amostra genética é suficiente para detectar a variação genética no feto com base nos números dos rótulos materno e paterno. O método pode ainda consistir na determinação da proporção do material fetal na amostra genética.

[00106] Em algumas representações, quando a paciente é uma paciente grávida, e a variação genética é uma variação no feto da paciente grávida, o método pode ainda consistir no contato dos conjuntos de sondagem do alelo A e do alelo B que são específicos do alelo na amostra genética, na qual o conjunto de sondagem do alelo A consiste em uma sondagem de rotulação do

alelo A e uma sondagem de marcação do alelo A, e o conjunto de sondagem do alelo B consiste em uma sondagem de rotulação do alelo B e em uma sondagem de marcação do alelo B. O método pode ainda consistir na hibridização de pelo menos uma parte de cada um dos conjuntos de sondagem do alelo A e do alelo B em uma região de ácido nucleico de interesse nas moléculas de nucleotídeo da amostra genética, região de ácido nucleico de interesse que consiste em um local de polimorfismo de nucleotídeo único (SNP) predeterminado para o qual um perfil alélico materno (isto é, genótipo) difere de um perfil alélico fetal no local de SNP (por exemplo, a composição alélica materna pode ser AA e a composição alélica fetal pode ser AB ou BB. Em outro exemplo, a composição alélica materna pode ser AB, e a composição alélica fetal pode ser AA ou BB. ), em que pelo menos uma parte do conjunto de sondagem ao alelo A é hibridizado em um primeiro alelo no local do SNP, pelo menos uma parte do conjunto de sondagem do alelo B é hibridizado em um segundo alelo no local do SNP e o primeiro e o segundo alelos são diferentes entre si. O método pode ainda consistir na ligação dos conjuntos de sondagem do alelo A e alelo B pelo menos por ligação (i) das sondagens de rotulação e marcação do alelo A e (ii) das sondagens de rotulação e marcação do alelo B. O método pode consistir na amplificação dos conjuntos de sondagem ligados. O método pode ainda consistir na imobilização das sondagens de marcação em uma localidade predeterminada em um substrato, em que as sondagens de rotulação do alelo A e alelo B e/ou suas sondagens de rotulação amplificadas ligadas às sondagens de marcação imobilizadas consistem nos rótulos do alelo A e alelo b, respectivamente; os rótulos do alelo A e alelo B são diferentes, e os rótulos imobilizados são ópticamente determináveis. O método pode ainda consistir na contagem de números de rótulos do alelo A e alelo B e determinar se uma

proporção de um material fetal na amostra genética é suficiente para detectar a variação genética no feto com base nos números dos rótulos do alelo A e alelo B. O método pode ainda consistir na determinação da proporção do material fetal na amostra genética.

[00107] Em algumas representações, quando a paciente é uma paciente grávida, a variação genética é uma variação no feto da paciente grávida, e a amostra genética consiste em um cromossomo Y, o método pode ainda consistir no contato dos conjuntos de sondagem materna e paterna na amostra genética, na qual o conjunto de sondagem materna consiste em uma sondagem de rotulação materna e uma sondagem de marcação materna, e o conjunto de sondagem paterna consiste em uma sondagem de rotulação paterna e uma sondagem de marcação paterna. O método pode ainda consistir na hibridização de pelo menos partes dos conjuntos de sondagem materna e paterna em regiões de ácido nucleico materno e paterno de interesse nas moléculas de nucleotídeo da amostra genética, respectivamente, em que a região do ácido nucleico paterno de interesse está localizada no cromossomo Y, e a região do ácido nucleico materno de interesse não está localizada no cromossomo Y. O método pode ainda consistir na ligação dos conjuntos de sondagem materna e paterna pelo menos por ligação (i) das sondagens de rotulação e marcação maternas e (ii) das sondagens de rotulação e marcação paterna. O método pode também consistir na amplificação das sondagem ligadas. O método pode ainda consistir na região de ácido nucleico de interesse que compreende um local de polimorfismo de nucleotídeo único (SNP) que contém mais de um SNP, por exemplo, dois ou três SNPs. Além disso, o local de SNP pode conter SNPs com alto desequilíbrio de vinculação de modo que as sondagens de rotulação e marcação sejam configuradas para se obter vantagem da energética melhorada de várias compatibilidades ou

incompatibilidades de SNP versus apenas uma. O método pode ainda consistir na imobilização das sondagens de marcação em uma localidade predeterminada em um substrato, em que as sondagens de rotulação materna e paterna e/ou suas sondagens de rotulação amplificadas ligadas às sondagens de marcação imobilizadas consistem nos rótulos materno e paterno, respectivamente, os rótulos materno e paterno são diferentes, e os rótulos imobilizados são óticamente determináveis. O método pode ainda consistir na contagem de números de rótulos materno e paterno e determinar se uma proporção de um material fetal na amostra genética é suficiente para detectar a variação genética no feto com base nos números dos rótulos materno e paterno. O método pode ainda consistir na determinação da proporção do material fetal na amostra genética.

[00108] Nas representações adicionais, outras variações genéticas (por ex., exclusão de base única, microssatélite e pequenas inserções) podem ser usados no lugar da variação genética no local de SNP descrito neste documento.

[00109] Em um aspecto, o conjunto de sondagem descrito neste documento pode compreender três ou mais sondagens, incluindo pelo menos uma sondagem entre as sondagens de rotulação e de marcação. Em algumas representações, o primeiro e o segundo conjuntos de sondagem consistem ainda na primeira e segunda sondagens de lacuna, respectivamente; a primeira sondagem de lacuna é hibridizada em uma região entre as regiões em que a primeira sondagem de rotulação e a primeira sondagem de marcação são hibridizadas; a segunda sondagem de lacuna é hibridizada em uma região entre as regiões em que a segunda sondagem de rotulação e a segunda sondagem de marcação são hibridizadas. O método pode ainda consistir na etapa de ligação que liga pelo menos (i) a primeira sondagem de rotulação, a primeira sondagem de marcação e a primeira sondagem de lacuna e (ii) a segunda

sondagem de rotulação, a segunda sondagem de marcação e a segunda sondagem de lacuna. Em representações adicionais, a sondagem de lacuna pode consistir em um rótulo. Por exemplo, a primeira e a segunda sondagens de lacuna e/ou seus produtos amplificados são rotulados com rótulos (por ex., décimo terceiro e décimo quarto rótulos, respectivamente) e cada um dos rótulos pode ser diferente do restante dos rótulos (por ex., primeiro e segundo rótulos). Os rótulos nas sondagens de lacuna (por ex., décimo terceiro e décimo quarto rótulos) podem ser iguais ou diferentes uns dos outros. Em outro aspecto, a primeira e segunda sondagens de rotulação são hibridizadas na primeira e segunda regiões de ácido nucleico de interesse nas moléculas de nucleotídeo da amostra genética, respectivamente; a primeira e a segunda sondagens de marcação são hibridizadas na primeira e segunda regiões de ácido nucleico de interesse nas moléculas de nucleotídeo da amostra genética, respectivamente; a primeira e a segunda sondagens de lacuna são hibridizadas na primeira e segunda regiões de ácido nucleico de interesse nas moléculas de nucleotídeo da amostra genética, respectivamente. Em algumas representações, há de 0 a 100 nucleotídeos, 1 a 100 nucleotídeos, 2 a 50 nucleotídeos; 3 a 30 nucleotídeos, 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 100, 150 ou 200 ou mais; ou 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 25, 35, 45, 55, 110, 160 ou 300 ou menos entre as regiões nas quais a primeira sondagem de rotulação e as sondagens de marcação são hibridizadas; e há de 0 a 100 nucleotídeos, 1 a 100 nucleotídeos, 2 a 50 nucleotídeos; 3 a 30 nucleotídeos, 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 100, 150 ou 200 nucleotídeos ou mais; ou 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 25, 35, 45, 55, 110, 160 ou 300 nucleotídeos ou menos entre as regiões nas quais a segunda sondagem de rotulação e as sondagens de marcação são hibridizadas. Nas representações adicionais, a

sondagem de lacuna entre uma sondagem de rotulação e uma sondagem de marcação pode ter um comprimento de 0 a 100 nucleotídeos, 1 a 100 nucleotídeos, 2 a 50 nucleotídeos; 3 a 30 nucleotídeos, 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 100, 150 ou 200 ou mais; ou 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 25, 35, 45, 55, 110, 160 ou 300 ou menos.

[00110] Em outro aspecto, o conjunto de sondagem descrito neste documento pode compreender um espaçador ligado e/ou conjugado à sondagem de rotulação e à sondagem de marcação. O espaçador pode ou não consistir em oligonucleotídeos. O espaçador pode compreender um material isolado, purificado que ocorre naturalmente e não naturalmente, incluindo nucleotídeo de qualquer comprimento (por ex., 5, 10, 20, 30, 40, 50, 100 ou 150 nucleotídeos ou menos). Em algumas representações, a sondagem pode estar em uma digestão de restrição purificada ou produzida sinteticamente, de forma recombinante ou por amplificação de PCR. Por exemplo, as primeiras sondagens de rotulação e marcação são conjugadas por um primeiro espaçador, as segundas sondagens de rotulação e marcação serem conjugadas por um segundo espaçador e o primeiro e o segundo espaçadores não são hibridizados nas moléculas de nucleotídeo da amostra genética. Em algumas representações, o método ainda consiste na digestão da amostra genética hibridizada com uma enzima e quebra de uma ligação no primeiro e segundo espaçadores após a digestão.

[00111] Em outro aspecto, o método descrito neste documento exclui a identificação de uma sequência nas moléculas de nucleotídeo da amostra genética e/ou sequenciamento das regiões de ácido nucleico de interesse e/ou as sondagens. Em algumas representações, o método que exclui o sequenciamento das sondagens inclui a exclusão de sequenciamento de um código de barra e/ou marca de afinidade em uma sondagem de marcação. Nas representações adicionais, os conjuntos de sondagem imobilizados

detectam diferentes variações genéticas, regiões de nucleotídeo de interesse e/ou os peptídeos de interesse não precisam ser detectados ou escaneados separadamente porque o sequenciamento não é exigido nos métodos descritos aqui. Nas representações adicionais, os números de diferentes rótulos imobilizados no substrato foram contados simultaneamente (por ex., por um único escaneamento e/ou geração de imagem) e dessa forma os números de diferentes rótulos não foram contados separadamente. Em outro aspecto, o método descrito neste documento exclui a leitura ou a quantificação análoga do arranjo em grande volume. A leitura do arranjo em grande volume aqui significa uma única medição que mede o sinal cumulativo combinado de vários rótulos de um único tipo, opcionalmente combinado com uma segunda medição do sinal cumulativo combinado de vários rótulos de um segundo tipo, sem resolver um sinal de cada rótulo. Um resultado é representado a partir da combinação de uma ou mais dessas medições no qual os rótulos individuais não são determinados. Em outro aspecto, o método descrito neste documento pode incluir uma única medição que mede os mesmos rótulos, diferentes rótulos do mesmo tipo e/ou rótulos do mesmo tipo nos quais os rótulos individuais são determinados. O método descrito neste documento pode excluir a quantificação análoga e pode empregar quantificação digital, na qual apenas o número de rótulos é determinado (confirmado por meio de medições de intensidade e formato de rótulo individual), e não a intensidade ótica cumulativa ou combinada dos rótulos.

[00112] Em outro aspecto, o conjunto de sondagem descrito neste documento pode consistir em um ligador. Um ligador é o mesmo material da marca ou marca de afinidade descrita aqui. Em algumas representações, o método consiste ainda na imobilização do ligador em uma fase sólida após as etapas de ligação. O método pode consistir no isolamento dos conjuntos de sondagem ligados a partir de sondagens não ligadas. Nas representações

adicionais, o ligador consiste em biotina, e a fase sólida consiste em uma esfera magnética.

[00113] Em outro aspecto, a etapa de contagem descrita neste documento pode também compreender a calibração, a verificação e/ou a confirmação dos números contados. A calibração aqui significa a verificação e/ou o ajuste da exatidão do número contado. A verificação e a confirmação significam determinar se o número contado é exato ou não e/ou o nível de erro, se houver.

[00114] Em outro aspecto, a intensidade e/ou o sinal-ruído é usado como método de identificação de rótulos únicos. Quando as moléculas de tintura ou outros rótulos óticos que estão próximos, muitas vezes eles são impossíveis de discriminar com imagem baseada em fluorescência devido ao limite intrínseco da difração da luz. Ou seja, dois rótulos que estão próximos serão indistinguíveis sem lacuna visível entre eles. Um método de exemplo para determinar o número de rótulos em uma determinada localidade é examinar o sinal relativo e/ou sinal-ruído comparado às localidades conhecidas por ter um flúor único. Dois ou mais rótulos geralmente emitirão um sinal mais claro (e um que possa ser mais claramente diferenciado do fundo) e será um flúor único. A figura 2 mostra o histograma normalizado da intensidade de sinal medida com base nas amostras de rótulo únicas e anticorpos de vários rótulos (ambos Alexa 546; verificados por meio de perfis de descoloração). As duas populações eram claramente separáveis e vários rótulos podem ser claramente distintos de rótulos únicos.

[00115] Em algumas representações, a etapa de contagem pode consistir na medição de sinais óticos dos rótulos imobilizados e na calibração dos números contados distinguindo-se um sinal ótico de um único rótulo do restante dos sinais óticos de fundo e/ou vários rótulos. Em algumas representações, a distinção consiste no cálculo de um sinal relativo e/ou intensidade de

sinal-ruído do sinal ótico comparado a uma intensidade de um sinal ótico de um rótulo único. A distinção pode ainda consistir em determinar se o sinal ótico provém de um rótulo único. Nas representações adicionais, o sinal ótico proverá de um rótulo único se o sinal relativo e/ou intensidade de sinal-ruído de um sinal ótico diferir de uma intensidade de um sinal ótico de um rótulo único por uma quantidade predeterminada ou menos. Em outras representações, a quantidade predeterminada é de 0% a 100%, de 0% a 150%, 10% a 200%, 0, 1, 2, 3, 4, 5, 10, 20, 30 ou 40% ou mais e/ou 300, 200, 100, 50, 30, 10 ou 5% ou menos da intensidade do sinal ótico de um rótulo único.

[00116] Em outro aspecto, os rótulos diferentes podem ter diferentes níveis de cintilação e propriedades de descoloração. Eles também podem ter diferentes propriedades de excitação. Para comparar o número de moléculas de tintura para dois rótulos diferentes, é necessário garantir que as duas tinturas se comportem de maneira semelhante e tenham características de emissão similares. Por exemplo, se uma tintura for mais turva que a outra, o número de moléculas poderá ser contado a menor nesse canal. Vários fatores podem ser titulados para produzir a equivalência ideal entre as tinturas. Por exemplo, a etapa de contagem e/ou a etapa de calibração pode consistir na otimização de (i) capacidade de fonte de luz para excitar os rótulos, (ii) tipos de fontes de luz, (iii) tempos de exposição para os rótulos e/ou (iv) conjuntos de filtros para os rótulos corresponderem aos sinais óticos dos rótulos e medição do sinal ótico dos rótulos. Esses fatores podem variar isoladamente ou em combinação. Além disso, a métrica otimizada pode variar. Por exemplo, pode ser intensidade geral, sinal-ruído, menos fundo, variação mais baixa em intensidade ou qualquer outra característica.

[00117] Os perfis de descoloração são específicos do rótulo e podem ser usados para adicionar informações para distinguir tipos de rótulo. A figura 3 mostra perfis de branqueamento médio de vários rótulos. O gráfico mostra as contagens normalizadas por tipo de rótulo como uma função de imagem sucessivas que foram coletadas em um intervalo de 60 segundos. O item c1 é o flúor Cy3, o item c2 é o flúor Atto647 e o item c3 é o flúor Alexa488.

[00118] Em outro aspecto, o comportamento de cintilação pode ser usado como um método de identificação de rótulos únicos. Muitas moléculas de tintura são conhecidas por temporariamente passarem a um estado escuro (por ex., Burnette et al., Proc. Natl. Acad. Sci. USA (2011) 108: 21081-21086). Isso produz um efeito de cintilação onde um rótulo passará por uma ou mais etapas de brilho-escuridão-brilho. O comprimento e o número desses períodos de escuridão podem variar. A invenção atual usa esse comportamento de cintilação para discriminar um rótulo de dois ou mais rótulos que podem parecer similares nas imagens limitadas de difração. Se houver vários rótulos presentes, será improvável que o sinal desapareça completamente durante a cintilação. É mais provável que a intensidade diminua quando um dos rótulos escurece, mas os outros não. A probabilidade de todos os rótulos cintilarem simultaneamente (e assim parecerem-se com um flúor único) pode ser calculada com base nas características de cintilação específicas de uma tintura.

[00119] Em algumas representações, os sinais óticos provenientes dos rótulos são medidos em pelo menos dois momentos e um sinal ótico será proveniente de um único rótulo se a intensidade do sinal ótico for reduzida por uma única função de etapa. Em algumas representações, os dois momentos no tempo podem ser separados por 0,1 a 30 minutos, 1 segundo a 20 minutos, 10 segundos a 10 minutos; 0,01, 0,1, 1, 2, 3, 4, 5, 10,

20, 30, 40, 50, 60 segundos ou mais; e/ou 1, 2, 3, 4, 5, 10, 20, 30, 40, 50, 60 segundos ou menos. Nas representações adicionais, uma intensidade do sinal ótico proveniente de um único rótulo tem uma única diminuição de etapa no decorrer do tempo e uma intensidade do sinal ótico de dois ou mais rótulos tem várias diminuições de etapa no decorrer do tempo. Em representações adicionais, os sinais óticos provenientes dos rótulos são medidos em pelo menos dois momentos e são normalizados nos perfis de branqueamento dos rótulos. Em outro aspecto, o método descrito neste documento e/ou a etapa de contagem pode ainda consistir na medição de um sinal ótico de um rótulo de controle em pelo menos dois momentos e comparação do sinal ótico do rótulo de controle com os sinais óticos dos rótulos para determinar um aumento ou diminuição do sinal ótico proveniente dos rótulos.

[00120] Em outro aspecto, a etapa de contagem consiste também na confirmação da contagem com o uso de uma molécula de controle. Uma molécula de controle pode ser usada para determinar a alteração na frequência de um tipo de molécula. Com frequência, o objetivo experimental é determinar a abundância de dois ou mais tipos de moléculas de maneira absoluta ou um em relação ao outro. Considere o exemplo de duas moléculas rotuladas com duas tinturas diferentes. Se a hipótese nula for a de estarem em frequência igual, elas poderão ser enumeradas em um arranjo de molécula única e a proporção das contagens comparadas à hipótese nula. A expressão "arranjo de molécula única" neste documento é definida como um arranjo configurado para detectar uma molécula única, incluindo, por exemplo, os arranjos descritos na Publicação do requerimento de patente norte-americana N° 2013/0172216. Se a proporção variar de 1:1, isso implicará que as duas moléculas estão em frequências diferentes. Entretanto, pode não ficar claro a priori se uma tem

a abundância ampliada ou se a outra abundância diminuída. Se uma terceira tintura for usada como uma molécula de controle que também deva estar em frequência igual, ela deverá ter uma proporção 1:1 com as duas outras tinturas. Considere o exemplo de duas moléculas rotuladas com as tinturas A e B, o objetivo é ver se a molécula rotulada com a tintura B está em frequência ampliada ou diminuída comparada à molécula rotulada com a tintura A. Uma terceira molécula rotulada com a tintura C é incluída no experimento de modo que deva ter a mesma abundância das outras duas moléculas. Se a proporção de moléculas rotuladas A e B for respectivamente 1:2, a primeira molécula terá a frequência diminuída ou a segunda terá a frequência aumentada. Se a proporção das moléculas rotuladas A e C for 1:1 e a proporção das moléculas rotuladas B e C for 1:2, é provável que a molécula rotulada com a tintura B tenha a frequência aumentada em relação à molécula rotulada com a tintura A. Um exemplo disso seria determinar as mudanças no número de cópia de DNA no genoma diploide. É importante saber se a sequência é ampliada ou excluída e se o uso de uma molécula de controle permite essa determinação. Note que o controle pode ser outra região do genoma ou uma sequência de controle artificial.

[00121] Em algumas representações, os resultados do método descrito neste documento (por ex., números contados de rótulos) podem ser confirmados com o uso de diferentes rótulos, mas com as mesmas marcas usadas no método inicial. Essa confirmação pode ser executada simultaneamente com o método inicial ou após a realização do método inicial. Além das representações adicionais, a confirmação descrita neste documento consiste no contato do primeiro e segundo conjuntos de sondagem com a amostra genética, em que o primeiro conjunto de sondagem de controle consiste em uma primeira sondagem de rotulação e primeira sondagem de marcação, que é a mesma marca do primeiro

conjunto de sondagem descrito neste documento e o segundo conjunto de sondagem de controle consiste em uma sondagem de rotulação de controle e na segunda sondagem de marcação, que é a mesma marca do segundo conjunto de sondagem descrito neste documento. A confirmação pode ainda consistir na hibridização de pelo menos parte do primeiro e segundo conjuntos de sondagem na primeira e segunda regiões de ácido nucleico de interesse nas moléculas de nucleotídeo da amostra genética, respectivamente. A confirmação pode ainda consistir na ligação do primeiro conjunto de sondagem de controle pelo menos ligando a primeira sondagem de rotulação de controle e primeira sondagem de marcação. A confirmação pode ainda consistir na ligação do segundo conjunto de sondagem de controle pelo menos ligando a segunda sondagem de rotulação de controle e segunda sondagem de marcação. A confirmação pode consistir na amplificação dos conjuntos de sondagem ligados. A confirmação pode ainda consistir na imobilização de cada uma das sondagens de marcação em uma localidade predeterminada em um substrato, em que a primeira e a segunda sondagens de rotulação de controle e/ou suas sondagens de rotulação amplificadas ligadas às sondagens de marcação imobilizadas consistem no primeiro e segundo rótulos de controle, respectivamente, o primeiro e o segundo rótulos de controle são diferentes, e os rótulos imobilizados são ópticamente determináveis. A confirmação pode ainda consistir na medição dos sinais óticos dos rótulos de controle imobilizados para o substrato. A confirmação pode consistir também na comparação dos sinais óticos do primeiro e segundo rótulos de controle imobilizados para os sinais óticos do primeiro e segundo rótulos imobilizados para determinar se existe um erro com base nos rótulos. A expressão "erro baseado em um rótulo" usada neste documento significa qualquer erro causado pelo rótulo que pode não ter ocorrido se um rótulo diferencial for

usado no método. Em algumas representações, o primeiro e o segundo rótulos de controle são iguais e o segundo rótulo e o primeiro rótulo de controle são iguais.

[00122] A descoloração também pode ser usada como um método de identificação de rótulos únicos. Um elemento importante da leitura é que os rótulos individuais devem ser "determináveis", isto é, distintos. Isso é trivial em baixas densidades em uma superfície quando a probabilidade de rótulos na proximidade é muito baixa. Para densidades mais elevadas, supondo-se que os rótulos estejam em localidades aleatórias (isto é, distribuição de Poisson), as chances de vizinhanças próximas aumenta até o ponto no qual números significativos de rótulos tenham vizinhos cuja emissão fluorescente parcialmente (ou totalmente) se sobreponha com sua própria emissão. Nesse ponto, os rótulos não são mais "determináveis", e existe um regime de transição entre a detecção de rótulo único (isto é, detecção digital) e a detecção de tipo de arranjo multirrótulo clássica (por ex., detecção análoga) na qual o sinal médio de muitas moléculas é medido. De outra forma, um regime de contagem digital de moléculas individuais é alternado para um regime análogo de intensidade fluorescente média de muitas moléculas.

[00123] Uma solução para aumentar o intervalo de carregamento durante a manutenção da capacidade de determinação individual é obter vantagem da descoloração de fluoróforo. A exposição ampliada à luz pode fazer com que os rótulos sejam descolorados, ou seja, percam sua propriedade de fluorescência. Ou seja, com o passar do tempo, um rótulo pode ser extinto. Isso geralmente ocorre como uma função de etapa, com o rótulo parecendo "desativado". A invenção atual pode usar esse comportamento de descoloração para discriminar um rótulo de dois ou mais rótulos que podem parecer similares nas imagens limitadas de difração. Para vários rótulos, a extinção seria esperada por meio de uma

série de diminuições em etapas na intensidade do sinal. Por exemplo, as figuras 4-13 mostram a intensidade de rótulo integrada vs gráficos de tempo (mostrando os eventos de descoloração como mudanças na intensidade) que foram obtidos para vários rótulos Alexa 488. Espécies de rótulo único versus múltiplo podem ser facilmente diferenciadas (por ex., dependendo de a intensidade do sinal ótico ser reduzida por etapas únicas versus múltiplas conforme mostrado nos gráficos).

[00124] Em outro aspecto, o método descrito neste documento pode consistir na calibração e/ou confirmação dos números contados por uma troca de rótulo ou troca de tintura. Em algumas representações, onde o produto de sondagem 1 e 2 são rotulados com os rótulos 1 e 2, respectivamente, vários modos de erro podem imitar a frequência diferencial dos produtos de sondagem. Por exemplo, se uma proporção de 1:2 for observada entre o rótulo 1 e o rótulo 2, isso pode ocorrer devido a diferenças genuínas na frequência (o produto de sondagem 2 é duas vezes tão comum quanto o produto de sondagem 1), diferenças na eficiência de hibridização (os produtos de sondagem têm abundância igual, mas o produto de sondagem 2 é hibridizado de maneira mais eficiente que o produto de sondagem 1) ou diferenças nas propriedades dos rótulos (por exemplo, se os rótulos forem tinturas fluorescentes, o rótulo 1 poderá descolorar primeiro, cintilar com mais frequência, produzir sinal menor ou sinal-ruído menor que o rótulo 2). Se o mesmo experimento for repetido com os rótulos alternados, a proporção deverá ser revertida, se for uma observação genuína de diferentes frequências das moléculas, com o rótulo 1 agora duas vezes tão comum quanto o rótulo 2. Entretanto, se for devido à eficiência de hibridização diferencial, a proporção será  $\leq 2:1$ . Se a proporção 1:2 foi devido às propriedades dos rótulos, a proporção mudará para 2:1 do rótulo 1 para o rótulo 2 se estiverem de fato em frequências

iguais. Essa abordagem pode ser estendida a qualquer número de conjuntos de sondagem rotulada.

[00125] Em algumas representações, a primeira região do ácido nucleico de interesse está localizada em um primeiro cromossomo, e a segunda região de ácido nucleico de interesse está localizada em um segundo cromossomo, diferente do primeiro cromossomo. A etapa de contagem pode ainda compreender a confirmação da contagem, na qual a etapa de confirmação consiste em manter o contato do primeiro e segundo conjuntos de sondagem de controle com a amostra genética, em que o primeiro conjunto de sondagem de controle consiste em uma primeira sondagem de rotulação de controle e uma primeira sondagem de marcação de controle, e o segundo conjunto de sondagem de controle consiste em uma segunda sondagem de rotulação de controle e segunda sondagem de marcação de controle. A etapa de confirmação pode ainda compreender a hibridização de pelo menos parte do primeiro e segundo conjuntos de sondagem de controle na primeira e segunda região de controle localizadas no primeiro e segundo cromossomos, respectivamente, em que a primeira e a segunda regiões de controle são diferentes da primeira e segunda regiões de ácido nucleico de interesse. A etapa de confirmação pode ainda consistir na ligação do primeiro e segundo conjuntos de sondagem de controle pelo menos por ligação (i) das primeiras sondagens de rotulação de controle e marcação e (ii) das segundas sondagens de rotulação de controle e marcação. A etapa de confirmação pode consistir na amplificação dos conjuntos de sondagem ligados. A etapa de confirmação pode ainda consistir na imobilização (i) do primeiro conjunto de sondagem e do segundo conjunto de sondagem de controle para uma primeira localidade predeterminada e (ii) do segundo conjunto de sondagem e do primeiro conjunto de sondagem de controle em uma segunda localidade predeterminada. Em algumas representações, a primeira

e a segunda sondagens de rotulação de controle e/ou suas sondagens de rotulação amplificadas ligadas às sondagens de marcação imobilizadas consistem em um primeiro e segundo rótulos de controle, respectivamente, o primeiro e o segundo rótulos de controle são diferentes, o segundo rótulo e o primeiro rótulo de controle são diferentes, os rótulos imobilizados são óticamente determináveis, a primeira e a segunda sondagens imobilizadas de marcação de controle e/ou suas sondagens de marcação amplificadas consistem na primeira e segunda marcas de controle, respectivamente, e a etapa de imobilização é realizada pela imobilização das marcas nas localidades predeterminadas. A etapa de confirmação pode ainda consistir na medição dos sinais óticos dos rótulos de controle imobilizados para o substrato. A etapa de confirmação pode consistir também na comparação dos sinais óticos dos rótulos de controle imobilizados para os sinais óticos do primeiro e segundo rótulos imobilizados para determinar se existe um erro com base na região de ácido nucleico de interesse. Em outras representações, a primeira e a segunda marca de controle são iguais e a segunda marca e a primeira marca de controle são iguais.

[00126] Em outro aspecto, a etapa de contagem do método descrito neste documento pode ainda compreender a calibração e/ou confirmação dos números contados (i) pela repetição de algumas ou de todas as etapas dos métodos (por ex., etapas que incluem o contato, a ligação, a hibridização, a ligação, a amplificação e/ou a imobilização) descritos neste documento com um conjunto de sondagem diferente configurado para se ligar e/ou hibridizar na mesma região de nucleotídeo e/ou peptídeo de interesse ou uma região diferente no mesmo cromossomo de interesse e (ii) cálculo da média dos números contados de rótulos nos conjuntos de sondagem ligados e/ou hibridizados na mesma região de nucleotídeo e/ou peptídeo de interesse ou mesmo

cromossomo de interesse. Em algumas representações, a etapa de cálculo de média pode ser executada antes da etapa de comparação para que os números contados médios de rótulos em um grupo de diferentes conjuntos de sondagem que se ligam e/ou são hibridizados na mesma região de nucleotídeo e/ou peptídeo de interesse sejam comparados, em vez dos números contados dos rótulos nos conjuntos de sondagem individual. Em outro aspecto, o método descrito neste documento pode ainda compreender a calibração e/ou confirmação da detecção da variação genética (i) pela repetição de algumas ou de todas as etapas dos métodos (por ex., etapas que incluem o contato, a ligação, a hibridização, a ligação, a amplificação, a imobilização e/ou a contagem) descritos neste documento com conjuntos de sondagem diferentes configurados para se ligarem e/ou ser hibridizados nas regiões que controle que não têm qualquer variação genética conhecida e (ii) cálculo da média dos números contados de rótulos nos conjuntos de sondagem e/ou hibridizados nas regiões de controle. Em algumas representações, os números médios dos rótulos nos conjuntos de sondagem que se ligam e/ou são hibridizados nas regiões de controle são comparados aos números dos rótulos nos conjuntos de sondagem que se ligam e/ou são hibridizados nas regiões de interesse descritas neste documento para confirmar a variação genética na amostra genética. Em outro aspecto, as etapas de calibração e/ou confirmação podem ser repetidas simultaneamente com as etapas iniciais ou após a execução das etapas iniciais.

[00127] Em outro aspecto, os rótulos (por ex., tinturas fluorescentes) de uma ou mais populações podem ser medidas e/ou identificadas com base em suas características espectrais subjacentes. A maioria dos sistemas de geração de imagens fluorescentes inclui a opção de coletar imagens em vários canais espectrais controlados pela combinação de fonte de luz e

excitação/emissão espectral/filtros dicróicos. Isso permite que a mesma espécie fluorescente em uma determinada amostra seja investigada com várias faixas de cores de luz diferentes bem como com a captura de faixas de cores de luz de saída desejadas. Em operação normal, a excitação de um fluoróforo é obtida pela iluminação com uma faixa espectral estreita alinhada à absorção máxima dessa espécie (por ex., com um LED de banda larga ou lâmpada a arco e filtro de excitação para modelar de forma espectral a saída ou um laser homogêneo de forma espectral) e a maior parte da emissão do fluoróforo é coletada com um filtro de emissão correspondente e um dicróico de passagem longa para diferenciar a excitação e a emissão (Figura 14). Em operações alternativas, a identidade exclusiva de uma parte fluorescente pode ser confirmada por meio de investigação com várias cores de excitação e faixas de emissão coletadas diferentes (ou além) do caso para a operação padrão (Figura 15.) A luz dessas várias configurações de imagens, por ex., vários filtros de emissão, é coletada e comparada aos valores de calibração para os fluoróforos de interesse (Figura 16). No caso de exemplo, a medição experimental (pontos) corresponde aos dados de calibração/referência esperados para esse fluoróforo (triângulos), mas não corresponde bem à hipótese alternativa (quadrados). Com os dados de teste e calibração apresentados para um ou mais canais, uma bondade de ajuste ou chi-quadrado pode ser calculado para cada espectro de calibração de hipótese, e o melhor ajuste selecionado, de maneira automática e robusta. Várias referências podem ser de interesse, incluindo fluoróforos usados no sistema, bem como contaminantes fluorescentes comuns, por ex., aqueles com um perfil de emissão plana (Contaminante 1; triângulo) ou um perfil ponderado azul (Contaminante 2; estrelas) (Figura 17).

[00128] As restrições do projeto para a seleção de filtro podem ser diferentes dos projetos padrão para os quais o objetivo é simplesmente maximizar a luz coletada em um único canal enquanto evita contribuições significativas de outros canais. Em nossa invenção, o objetivo é a seletividade espectral, e não unicamente a coleta de luz. Por exemplo, considere dois fluoróforos com faixas de excitação significativamente diferentes, mostrado na figura 18 (observação: somente as regiões de excitação são mostradas e nenhum espectro de excitação). Um projeto padrão maximizaria a captura da emissão de Flúor 1 (com filtro Em1, linha sólida) e minimizaria a obtenção da vantagem do Flúor 2, e o Flúor 2 seria capturado de maneira ideal pelo Em2 (que muda levemente para o vermelho a fim de evitar a coleta significativa da luz do Flúor 1). Em nosso projeto, a verificação da presença de Flúor 2 com o filtro Em1 é desejada para levar à amplitude da faixa a ser capturada ("Em1+", linha tracejada fina). Isso cria informações adicionais para verificar a identidade do Flúor 2. Da mesma maneira, o Em2 pode ser expandido ou deslocado para o Flúor 1 a fim de capturar mais dessa luz de flúor (Em2+, linha tracejada fina). Esse aumento nas informações espectrais também devem ser balanceados com a luz disponível total de um determinado fluoróforo para manter a capacidade de detecção. De outra forma, a contribuição de um determinado fluoróforo em um dado canal só será significativa se o sinal correspondente estiver acima do ruído de fundo e, portanto, informativo, exceto se um controle negativo for pretendido. Dessa maneira, a assinatura espectral de uma entidade fluorescente pode ser usada para identificação robusta e a captura de mais luz poderá ser uma segunda prioridade se recursos exclusivos da espécie puderem ser quantificados de maneira mais efetiva.

[00129] Os produtos de sondagem fornecidos podem ser rotulados com mais de um tipo de fluoróforo de modo que a assinatura espectral seja mais complexa. Por exemplo, os produtos de sondagem podem sempre carregar um flúor universal, por ex., Alexa647, e um fluoróforo específico do locus, por ex., Alexa 555 para o locus 1 e Alexa 594 para o locus 2. Como os contaminantes raramente portarão a assinatura de dois flúor, isso pode ainda aumentar a segurança da rejeição de contaminação. A implementação envolveria a geração de imagens em três ou mais canais neste exemplo de modo que a presença ou ausência de cada flúor possa ser confirmada pelo método de bondade de ajuste citado anteriormente comparando o teste à referência, produzindo chamadas de locus 1, locus 2 ou nenhum produto de locus. A adição de flúor extra auxilia na identificação, pois há mais disponibilidade de luz para a coleta, mas à custa da produção de produtos de ensaio adequadamente criados e tempo total de geração de imagem (canais extras podem ser necessários). Outros modificadores espectrais também podem ser usados para aumentar as informações espectrais e a singularidade, incluindo pares FRET que deslocam a cor quando bem próximo ou em outras metades.

[00130] Em outro aspecto, conforme descrito neste documento, o método da presente divulgação pode ser usado para detectar uma variação genética no peptídeo ou proteínas. Nesse caso, os métodos podem consistir no contato do primeiro e segundo conjuntos de sondagem com a amostra genética, na qual o primeiro conjunto de sondagem compreende uma primeira sondagem de rotulação e uma primeira sondagem de marcação, e o segundo conjunto de sondagem compreende uma segunda sondagem de rotulação e uma segunda sondagem de marcação. Os métodos podem ainda consistir na ligação dos conjuntos de sondagem às regiões do peptídeo de interesse por uma ligação física ou química, no

lugar da etapa de hibridização descrita neste documento no caso da detecção da variação genética nas moléculas de ácido nucleico. Especificamente, os métodos podem consistir na ligação pelo menos de partes do primeiro e segundo conjuntos de sondagem à primeira e segunda regiões de peptídeo de interesse em um peptídeo de proteína da amostra genética, respectivamente. Por exemplo, a ligação pode ser realizada tendo um ligador em pelo menos uma sondagem no conjunto de sondagem que se liga especificamente à região do peptídeo de interesse.

[00131] Em algumas representações, os métodos para detectar uma variação genética no peptídeo ou proteínas pode ainda consistir na conjugação do primeiro conjunto de sondagem por uma ligação química pelo menos conjugando a primeira sondagem de rotulação e a primeira sondagem de marcação e conjugando o segundo conjunto de sondagem pelo menos pela conjugação da segunda sondagem de rotulação e a segunda sondagem de marcação, no lugar da etapa de ligação descrita neste documento no caso de detectar a variação genética nas moléculas de ácido nucleico. O método pode ainda consistir na imobilização das sondagens de marcação em um local predeterminado em um substrato conforme descrito neste documento. Nas representações adicionais, a primeira e a segunda sondagem de rotulação conjugadas nas sondagens de marcação imobilizadas consistem no primeiro e segundo rótulos, respectivamente; o primeiro e o segundo rótulos são diferentes; os rótulos imobilizados são ópticamente determináveis; a primeira e a segunda sondagens de marcação imobilizadas e/ou suas sondagens de marcação compreendem na primeira e segunda marcas, respectivamente; e a etapa de imobilização é realizada pela imobilização das marcas na localidade predeterminada. Os métodos podem ainda consistir, conforme descrito neste documento, na contagem (i) de um primeiro número do primeiro rótulo imobilizado no substrato e

(ii) um segundo número do segundo rótulo imobilizado no substrato; e comparação do primeiro e segundo números para determinar a variação genética na amostra genética.

[00132] Um sistema para detectar uma variação genética de acordo com os métodos descritos neste documento inclui vários elementos. Alguns elementos incluem a transformação de uma amostra biológica bruta em um analito útil. Esse analito é detectado, gerando os dados que são processados em um relatório. Vários módulos que podem ser incluídos no sistema são mostrados na figura 19. Mais detalhes de vários métodos para analisar dados, incluindo, por ex., processamento de imagem, são mostrados na Figura 20. A análise pode ser executada em um computador e envolve uma rede conectada ao dispositivo que gera os dados e um servidor de dados para armazenamento de dados e relatório. Como opção, as informações adicionais além dos dados do analito podem ser incorporadas no relatório final, por ex., idade materna ou riscos conhecidos anteriores. Em algumas representações, o sistema de teste inclui uma série de módulos, alguns dos quais são opcionais ou podem ser repetidos, dependendo dos resultados dos módulos anteriores. O teste pode consistir: (1) no recebimento de uma requisição, por ex., de um clínico ou médico solicitante, (2) no recebimento de uma amostra de paciente, (3) na realização de um ensaio que inclui os controles de qualidade nessa amostra resultando em um produto de ensaio em um substrato de geração de imagem apropriado (por ex., contato, ligação e/ou hibridização de sondagens em uma amostra, ligando as sondagens, opcionalmente amplificando as sondagens ligadas, e imobilizando as sondagens para um substrato conforme descrito neste documento), (4) na geração de imagens do substrato em um ou mais canais espectrais, (5) na análise de dados da imagem, (6) na realização de cálculos estatísticos (por ex., comparação do primeiro e segundo números para determinar a

variação genética na amostra genética), (7) na criação e aprovação do relatório clínico e (8) no retorno do relatório para o clínico ou médico solicitante. O sistema de teste pode compreender um módulo configurado para receber uma requisição, por ex., de um clínico ou médico solicitante, um módulo configurado para receber a amostra de um paciente, (3) um módulo configurado para executar um ensaio que inclua controles de qualidade nessa amostra resultando em um produto de ensaio em um substrato de geração de imagem apropriado, (4) um módulo configurado para gerar a imagem do substrato em um ou mais canais espectrais, (5) um módulo configurado para analisar os dados de imagem, (6) um módulo configurado para executar os cálculos estatísticos, (7) um módulo configurado para criar e confirmar o relatório clínico e/ou (8) um módulo configurado para retornar o relatório ao clínico ou médico solicitante.

[00133] Em um aspecto, os ensaios e os métodos descritos neste documento podem ser realizados em uma única amostra de entrada simultaneamente. Por exemplo, o método pode consistir na verificação da presença de moléculas genômicas letais em um limite mínimo ou acima, conforme descrito neste documento, seguido por uma etapa de estimativa do estado do número de cópia alvo somente se esse limite mínimo for alcançado. Portanto, é possível executar separadamente um ensaio específico de alelo na amostra de entrada para realizar o cálculo de fração fetal e um ensaio alvo genômico para calcular o estado do número de cópia. Em outras representações, os ensaios e os métodos descritos neste documento podem ser executados em paralelo na mesma amostra ao mesmo tempo no mesmo volume fluídico. Além disso, os ensaios do controle de qualidade também podem ser executados em paralelo com as mesmas etapas de processamento de ensaio universal. Como as marcas, marcas de afinidade e/ou sondagens de marcação nos produtos de sondagem, conjunto de sondagem ligado

ou molécula rotulada para ser imobilizada no substrato podem ser exclusivamente desenvolvidos para cada ensaio e cada produto do ensaio, todos os produtos de ensaio paralelo podem ser localizados, ter sua imagem gerada e quantificados em diferentes localidades físicas nas imagens do substrato. Em outro aspecto, o mesmo ensaio ou método (ou alguma ou suas etapas) descrito neste documento usando as mesmas sondagens e/ou detectando a mesma variação genética ou controle podem ser executados em várias amostras simultaneamente no mesmo ou em diferentes módulos (por ex., tubo de teste) descrito aqui. Em outro aspecto, os ensaios e os métodos (ou algumas de suas etapas) descritos neste documento usando diferentes sondagens e/ou detectando diferentes variações genéticas ou controles podem ser executados em uma ou várias amostras simultaneamente no mesmo ou em diferentes módulos (por ex., tubo de teste).

[00134] Em outro aspecto, a análise da imagem pode incluir o pré-processamento da imagem, a segmentação da imagem para identificar os rótulos, a caracterização da qualidade do rótulo, a filtragem da população de rótulos detectados com base na qualidade e a realização de cálculos estatísticos, dependendo da natureza dos dados da imagem. Em algumas instâncias, como quando um ensaio específico alelo é realizado e sua imagem é gerada, a fração fetal pode ser calculada. Em outras, como o ensaio de alvo genômico e geração de imagem, o estado do número de cópia relativo entre as duas regiões genômicas alvo é calculado. A análise dos dados da imagem pode ocorrer em tempo real no mesmo computador que está controlando a aquisição da imagem ou em um computador em rede, de modo que resulte da análise, pode ser incorporada na árvore de decisões do fluxo de trabalho de teste quase em tempo real.

[00135] Em outro aspecto, as etapas (4) e (5) do teste acima podem ser repetidas várias vezes para diferentes partes da

geração de imagens do substrato de modo que os resultados determinem as próximas etapas. Por exemplo, os testes e os métodos descritos neste documento consistem na confirmação da presença e do nível preciso de uma amostra fetal em uma amostra genética obtida de um paciente antes do teste do estado de número de cópia relativo de alvos genômicos. Conforme descrito neste documento, um ensaio sensível a alelo pode ser usado para quantificar os níveis de DNA fetal relativo ao DNA maternal. Os produtos de sondagem resultantes podem ser baixados para uma região de fração fetal 1 no substrato e ter suas imagens geradas. Em algumas representações, somente se a fração fetal calculada estiver acima do requisito de sistema mínimo, o teste poderá continuar e produzir um resultado válido. Dessa forma, o teste de amostras que falham em confirmar pelo menos a fração fetal de entrada mínima pode ser finalizado antes de ocorrer a geração de imagens adicionais e a análise. De modo oposto, se a fração fetal estiver acima do limite mínimo, a geração de imagens (etapa 4 do teste) dos alvos genômicos (por ex., cromossomo 21, 18 ou 13) pode continuar seguida pela análise adicional (etapa 5 do teste). Outros critérios também podem ser usados e testados.

[00136] Em outro aspecto, cada SNP sondado no ensaio específico do alelo pode resultar em informações úteis. Por exemplo, o material genômico materno pode ter alelos heterozigotos para um determinado SNP (por ex., par de alelos AB) e o material fetal também pode ser heterozigoto nesse local (por ex., AB), conseqüentemente o material fetal é indistinguível, e o cálculo da fração fetal falha. Outro local de SNP para a mesma amostra de entrada, entretanto, pode novamente mostrar o material materno a ser heterozigoto (por ex., AB) enquanto o material fetal é homozigoto (por ex., AA). Neste exemplo, o ensaio específico alelo pode produzir levemente

mais contagens A do que contagens B devido à presença do DNA fetal, como base no qual a fração fetal pode ser calculada. Como o perfil SNP (ou seja, genótipo) não pode ser conhecido a priori para uma determinada amostra, vários ou numerosos locais de SNP devem ser desenvolvidos de modo que quase cada amostra possível produza um local de SNP informativo. Cada local de SNP pode ser localizado em uma diferente localidade física na geração de imagem do substrato, por exemplo, usando uma marca diferente para cada SNP. Entretanto, para um determinado teste, a fração fetal só pode ser calculada com êxito uma vez. Portanto, uma ou várias localidades no substrato usado para investigar os SNPs podem ter a imagem gerada podem ser analisadas (por ex., em grupos de um, dois, três, quatro, cinco, dez, vinte, cinquenta ou menos e/ou um, dois, três, quatro, cinco, dez, vinte, cinquenta ou mais) até um SNP informativo ser detectado. Ao alternar as imagens e a análise, é possível desviar a geração de imagem de todos os pontos de SNP possíveis e reduzir de maneira significativa a duração média do teste enquanto mantém a exatidão e a robustez.

[00137] Em outro aspecto, a determinação da fração fetal de uma amostra pode ajudar outros aspectos do sistema além dos testes de terminação para os quais a porção da fração fetal em uma amostra é inadequada. Por exemplo, se a fração fetal for elevada (por ex., 20%), para uma determinada capacidade estatística, o número de contagens necessário por alvo genético (por ex., chr21) será inferior, se a fração fetal for baixa (por ex., 1%) então para a mesma capacidade estatística, um número muito elevado de contagens é necessário por alvo genômico para alcançar o mesmo significado estatístico. Portanto, após a geração de imagens (4-1) da região de fração fetal 1, (5-1) a análise desses dados resultantes em um rendimento de contagem necessário por alvo genômico, (4-2) a geração de imagens da

região de alvo 2 começa no rendimento necessário, seguida pela (5-2) análise desses dados de imagem e resultado de teste para variação genômica dos alvos de entrada.

[00138] Em outro aspecto, as etapas (4) e (5) do teste acima podem ser repetidas também para fins de controle de qualidade, incluindo avaliação de níveis de flúor de fundo na geração de imagem do substrato, metades contaminadas, controles positivos ou outras causas de variação de número de cópia além do teste imediato (por ex., câncer na mãe ou no feto, quimerismo, geração de gêmeos). Como a análise da imagem pode ser em tempo real e não requer conclusão de toda a execução da imagem antes de gerar os resultados (diferentemente dos métodos de sequenciamento de DNA), os resultados intermediários podem determinar as etapas seguintes de uma árvore de decisões e adequar o teste para o desempenho ideal de uma amostra individual. O controle de qualidade também pode incluir verificar se a amostra é de qualidade aceitável e presente, se a geração de imagem do substrato está configurada adequadamente, se o produto do ensaio está presente e/ou na concentração ou densidade correta, se há níveis aceitáveis de contaminação, se a geração de imagem do instrumento é funcional e se a análise está produzindo resultados adequados, todos alimentando um relatório de teste final para revisão pela equipe clínica.

[00139] Em outro aspecto, o teste acima consiste em uma ou mais das seguintes etapas: (1) recebimento de uma requisição (por exemplo, de um clínico ou médico solicitante), (2) recebimento de uma amostra de paciente, (3) realização de um ensaio (incluindo uma porção específica do alelo, porção de alvo genômico e controles de qualidade) nessa amostra resultando em um produto de ensaio contendo geração de imagem de substrato, (4-1) geração de imagem da região específica do alelo do substrato em um ou mais canais espectrais, (5-1) análise dos

dados de imagem específicos do alelo para calcular a fração fetal, (fração fetal suficiente pendente) (4-2) geração de imagem da região do alvo genômico do substrato em um ou mais canais espectrais, (5-2) análise de dados de imagem de região de alvo genômico, (4-3) geração de imagens da região de controle de qualidade do substrato em um ou mais canais espectrais, (5-3) análise dos dados de imagem de controle de qualidade para computar, validar e verificar o teste, (6) realização de cálculos estatísticos, (7) criação e aprovação do relatório clínico e (8) envio do relatório de volta para o clínico ou médico solicitante.

[00140] Na descrição a seguir, várias representações de exemplo são demonstradas nas figuras.

[00141] A figura 21 é uma implementação de um ensaio para quantificar o número de cópia genômica em dois locus genômicos. Nesta representação do ensaio, 105 e 106 são as moléculas alvo. 105 contém a sequência correspondente ao primeiro locus genômico "Locus 1" investigado quanto ao número de cópia (por exemplo, cromossomo 21) e 106 contém a sequência correspondente ao segundo locus genômico "Locus 2" investigado quanto ao número de cópia (por exemplo, cromossomo 18). A figura 21 contém um exemplo de um conjunto de sondagem por locus genômico, mas em algumas representações deste ensaio, vários conjuntos de sondagem serão desenvolvidos para investigar várias regiões em um locus genômico. Por exemplo, mais de 10 ou mais de 100 ou mais de 500 conjuntos de sondagem podem ser desenvolvidos para corresponder ao cromossomo 21. A figura 21 ilustra apenas um único conjunto de sondagem para cada locus genômico, mas com importância o escopo desta invenção permite vários conjuntos de sondagem para cada locus genômico. A figura 21 também ilustra um único evento de hibridização entre uma molécula alvo e um conjunto de sondagem. Na prática, haverá várias moléculas alvo

presentes em um ensaio de amostra. Muitas moléculas alvo conterão as sequências necessárias para hibridização em um conjunto de sondagem e formação de um produto de sondagem. Diferentes moléculas de sondagem podem ser hibridizadas nos conjuntos de sondagem, pois determinadas moléculas carregarão polimorfismos genéticos. Além disso, as moléculas alvo que surgem do DNA genômico podem ter uma classificação aleatória de tamanhos de moléculas, bem como várias sequências iniciais e finais. Em essência, há várias moléculas alvo que podem ser hibridizadas em um determinado conjunto de sondagem. Em um único ensaio, várias cópias de uma determinada sondagem são adicionadas. Portanto, em um único ensaio até milhares ou centenas de milhares ou milhões de produtos de sondagem específicos podem ser criados.

[00142] A figura 21 representa dois conjuntos de sondagem; um conjunto de sondagem para Locus 1 e um conjunto de sondagem para o Locus 2, embora, conforme citado anteriormente, vários conjuntos de sondagem possam ser designados para cada locus genômico. Um primeiro conjunto de sondagem contém sondagens de membro 101, 102, 103. O item 101 contém um rótulo (100) do tipo "A". O item 103 contém uma marca de afinidade (104) que pode ser usada para isolamento e identificação do produto de sondagem. 102 pode não conter modificações, como um rótulo ou código de barras. Um segundo conjunto de sondagem com as sondagens de membro 108, 109, 110 apresenta as respectivas características como no primeiro conjunto de sondagem. Entretanto, 108 contém um rótulo (107) do tipo "B", distinguível do tipo "A". O item 110 contém uma marca de afinidade (111) que pode ser idêntica a ou exclusiva de 104. Muitos conjuntos de sondagem podem ser designados de modo que o "Locus 1" alvo contenha sequências de sondagem exclusivas, mas do mesmo tipo de rótulo "A". De modo similar, muitos conjuntos de sondagem podem ser designados de

modo que o "Locus 2" alvo contenha sequências de sondagem exclusivas, mas o mesmo tipo de rótulo "B". Nesta representação, as marcas de afinidade para os muitos conjuntos de sondagem para o Locus 1 podem ser idênticas ou exclusivas e as marcas de afinidade para os muitos conjuntos de sondagem para o Locus 2 podem ser idênticas ou exclusivas.

[00143] Um ou mais conjuntos de sondagem são adicionados a moléculas alvo em um único recipiente e expostos nas condições de hibridização específicas da sequência.

[00144] Para cada conjunto de sondagem, as três sondagens (por ex., 101, 102, 103) são hibridizadas (ou anexadas via uma interação alvo de sondagem similar) na molécula alvo (105) de modo que não haja lacunas entre as sondagens na molécula alvo. Ou seja, as sondagens do conjunto de sondagem são adjacentes entre si e com ligação capacitada.

[00145] A ligase é adicionada às sondagens hibridizadas e expostas nas condições de ligase padrão. As sondagens ligadas formam um produto de sondagem. Todos (ou uma maioria) de produtos de sondagem do Locus 1 têm um rótulo do tipo "A". Todos os produtos de sondagem do Locus 2 têm um rótulo do tipo "B". A quantificação dos produtos de sondagem correspondente aos locus genômicos 1 e 2 ocorre com o uso dos rótulos "A" e "B".

[00146] Em algumas representações, os produtos de sondagem são imobilizados em um substrato usando suas marcas de afinidade. Por exemplo, se a marca de afinidade for uma sequência de DNA, os produtos de sondagem poderão ser hibridizados em regiões de um arranjo de captura de DNA na densidade apropriada para a geração de imagens subsequentes.

[00147] Em algumas representações, as marcas de afinidade 104 e 111 contêm sequências exclusivas e ortogonais que permitem o posicionamento, baseado na superfície, para um ou mais locais, o que pode ser compartilhado entre os produtos de hibridização ou

não. As figuras 47 e 48 mostram os padrões de fluorescência resultantes quando os produtos contêm sequências de marca de afinidade exclusivas e o substrato subjacente contém complementos para cada uma das marcas de afinidade exclusiva na mesma região (por ex., como o mesmo membro de um grupo) em um substrato. As imagens são da mesma região de um substrato, mas a figura 47 mostra os rótulos Cy3 (ligados de maneira covalente ao produto do cromossomo 18), e a figura 48 mostra os rótulos Alexa Fluor 647 (ligados de maneira covalente ao produto do cromossomo 21). Padrões similares podem ser gerados para outras representações de ensaio que seguem.

[00148] Em outra representação, as marcas de afinidade 104 e 111 contêm sequências idênticas que permitem o posicionamento baseado na superfície na mesma região (por ex., como o mesmo membro de um grupo) em um substrato. Ou seja, produtos diferentes competem pelos mesmos locais de ligação. As figuras 49 e 51 mostram os padrões de fluorescência resultantes quando diferentes produtos contêm sequências de marca de afinidade idênticas e o substrato subjacente contém o complemento da marca de afinidade. As imagens são do mesmo local em um substrato, mas a figura 49 mostra os rótulos Cy3 (ligados de maneira covalente ao produto do cromossomo 18), e a figura 51 mostra os rótulos Alexa Fluor 647 (ligados de maneira covalente ao produto do cromossomo 21). As figuras 50 e 52 mostram regiões em zoom nas figuras 49 e 51 respectivamente, demonstrando claramente a resolução de molécula única e os rótulos distinguíveis individualmente. Padrões similares podem ser gerados para outras representações de ensaio que seguem.

[00149] Em outra representação, as marcas de afinidade 104 e 111 contêm sequências exclusivas e ortogonais que permitem o posicionamento baseado em superfície em mais de um local em um substrato. As figuras 53 e 54 mostram os padrões de

fluorescência resultantes quando os produtos contêm sequências de marca de afinidade exclusiva e o substrato subjacente tem uma região contendo o complemento a um complemento de marca de afinidade e outra região separada contendo o complemento da outra marca de afinidade. As imagens são de duas regiões separadas de um substrato, com cada região contendo um único complemento de marca de afinidade conforme descrito previamente. A figura 53 mostra os rótulos Cy3 (ligados de maneira covalente ao produto do cromossomo 21), e a figura 54 mostra os rótulos Alexa Fluor 647 (ligados de maneira covalente ao produto do cromossomo 18). Padrões similares podem ser gerados para outras representações de ensaio que seguem.

[00150] Uma característica desta invenção, de acordo com algumas representações, é que a especificidade é alcançada por meio da combinação de várias sondagens adjacentes que devem ser ligadas com êxito juntas para que o produto de sondagem seja criado, capturado e detectado com êxito. Se um produto de sondagem não for formado com êxito por qualquer motivo, ele não poderá ser isolado ou enriquecido para usar uma marca de afinidade e detectado. Por exemplo, se a sondagem 101 não for ligada com êxito à sondagem 102, o produto resultante não poderá ser detectado. De modo similar, se a sondagem 103 não for ligada com êxito à sondagem 102, o produto resultante não poderá ser isolado ou enriquecido com o uso de uma marca de afinidade.

[00151] A exigência de todas as sondagens do conjunto de sondagem para hibridização com êxito na molécula alvo e ligação com êxito fornece especificidade elevada e reduz muito os problemas de hibridização cruzada e, portanto, sinais falso-positivos.

[00152] Neste ensaio, a especificidade é alcançada por meio da hibridização e ligação específica da sequência. Em uma representação preferencial, a especificidade de formar produtos

de sondagem ocorre no recipiente de reação, antes do isolamento ou enriquecimento para produtos de sondagem, por exemplo, imobilização em uma superfície ou outro substrato sólido. Isso evita o desafio da hibridização baseada na superfície padrão (por ex., microarranjo genômico) na qual a especificidade deve ser inteiramente alcançada por meio da hibridização somente com sequências de oligonucleotídeo longas (>40bp) (por ex., grupos Agilent e Affymetrix).

[00153] O uso das marcas de afinidade permite que os produtos de sondagem sejam imobilizados em um substrato e, portanto, as sondagens não ligadas em demasia sejam lavadas com o uso de métodos padrão ou removidas com o uso dos métodos padrão. Portanto, todos ou a maioria dos rótulos na superfície são parte de um produto de sondagem criado especificamente que seja imobilizado na superfície.

[00154] Uma característica desta invenção, de acordo com algumas representações, é que a captura da superfície não afeta a exatidão. Ou seja, não introduz qualquer tendência. Em um exemplo, se a mesma marca de afinidade for usada para conjuntos de sondagem de diferentes locus genômicos, com conjuntos de sondagem buscando cada locus com um rótulo diferente. Os produtos de sondagem dos dois locus genômicos podem ser imobilizados no mesmo local no substrato usando a mesma marca de afinidade. Isso significa que os produtos de sondagem do Locus 1 e do Locus 2 serão capturados com a mesma eficiência, não introduzindo, portanto, qualquer tendência específica do locus.

[00155] Em algumas representações, algumas ou todas as sondagens não ligadas e/ou moléculas alvo são removidas antes da captura de superfície com o uso dos métodos padrão. Isso aumenta interferência entre as sondagens não ligadas e/ou moléculas alvo e os produtos de sondagem durante a captura de superfície.

[00156] Um recurso desta invenção, de acordo com algumas representações, é que vários tipos de marca de afinidade podem ser colocados na mesma região do substrato (por exemplo, o mesmo ponto do grupo ou membro do grupo). Isso tem muitas vantagens, incluindo o posicionamento do controle ou marcadores de calibração. As figuras 22-46 descrevem representações de exemplo adicionais desta invenção. Essas figuras não representam todas as representações possíveis, e todas as outras variações deste ensaio estão incluídas como parte desta invenção. Além disso, todos os recursos da representação descritos na Figura 21 são aplicáveis a todas as outras representações adicionais do ensaio descrito nesta aplicação.

[00157] A figura 22 representa uma modificação do procedimento geral descrito na figura 21. A figura 22 representa dois conjuntos de sondagem; um conjunto de sondagem para Locus 1 e um conjunto de sondagem para o Locus 2, embora, conforme citado anteriormente, vários conjuntos de sondagem possam ser designados para cada locus genômico. 207 e 214 são moléculas alvo correspondentes ao Locus 1 e Locus 2, respectivamente. Um primeiro conjunto de sondagem contém sondagens de membro 202, 204, 206. 202 contém um rótulo (201) do tipo "A". 206 contém uma marca de afinidade (205) que pode ser usada para isolamento e identificação do produto de sondagem. Um segundo conjunto de sondagem com as sondagens de membro 209, 211, 231 apresenta as respectivas características como no primeiro conjunto de sondagem. Entretanto, 209 contém um rótulo (208) do tipo "B", distinguível do tipo "A". 213 contém uma marca de afinidade (212) que pode ser idêntica a ou exclusiva de 205. Muitos conjuntos de sondagem podem ser designados de modo que o "Locus 1" alvo contenha sequências de sondagem exclusivas, mas do mesmo tipo de rótulo "A". De modo similar, muitos conjuntos de sondagem podem ser designados de modo que o "Locus 2" alvo

contenha sequências de sondagem exclusivas, mas o mesmo tipo de rótulo "B". Nesta representação, as marcas de afinidade para os muitos conjuntos de sondagem para o Locus 1 podem ser idênticas ou exclusivas em uma mistura de idênticas e exclusivas, e as marcas de afinidade para os muitos conjuntos de sondagem para o Locus 2 podem ser idênticas ou exclusivas ou uma mistura de idêntica e exclusiva. Nesta representação, as sondagens 204 e 211 podem conter um ou mais rótulos (203, 210) do tipo "C". Portanto, os produtos de sondagem conterão uma combinação de rótulos. Para o Locus 1, os produtos de sondagem conterão rótulos do tipo "A" e tipo "C", enquanto os produtos de sondagem do Locus 2 conterão rótulos do tipo "B" e do tipo "C".

[00158] A figura 23 representa uma modificação do procedimento geral descrito na figura 21. A figura 23 representa dois conjuntos de sondagem; um conjunto de sondagem para Locus 1 e um conjunto de sondagem para o Locus 2, embora, conforme citado anteriormente, vários conjuntos de sondagem possam ser designados para cada locus genômico. 307 e 314 são moléculas alvo correspondentes ao Locus 1 e Locus 2, respectivamente. Um primeiro conjunto de sondagem contém sondagens de membro 302, 303, 305. 302 contém um rótulo (301) do tipo "A". 305 contém uma marca de afinidade (306) que pode ser usada para isolamento e identificação do produto de sondagem. Um segundo conjunto de sondagem com as sondagens de membro 309, 310, 312 apresenta as respectivas características como no primeiro conjunto de sondagem. Entretanto, 309 contém um rótulo (308) do tipo "B", distinguível do tipo "A". 312 contém uma marca de afinidade (313) que pode ser idêntica a ou exclusiva de 306. Muitos conjuntos de sondagem podem ser designados de modo que o "Locus 1" alvo contenha sequências de sondagem exclusivas, mas do mesmo tipo de rótulo "A". De modo similar, muitos conjuntos de sondagem podem ser designados de modo que o "Locus 2" alvo

contenha sequências de sondagem exclusivas, mas o mesmo tipo de rótulo "B". Nesta representação, as marcas de afinidade para os muitos conjuntos de sondagem para o Locus 1 podem ser idênticas ou exclusivas e as marcas de afinidade para os muitos conjuntos de sondagem para o Locus 2 podem ser idênticas ou exclusivas. Nesta representação, as sondagens 305 e 312 contêm um ou mais rótulos (304, 311) do tipo "C". Portanto, os produtos de sondagem conterão uma combinação de rótulos. Para o Locus 1, os produtos de sondagem conterão rótulos do tipo "A" e tipo "C", enquanto os produtos de sondagem do Locus 2 conterão rótulos do tipo "B" e do tipo "C".

[00159] A figura 24 representa uma modificação do procedimento geral descrito na figura 21. A figura 24 representa dois conjuntos de sondagem; um conjunto de sondagem para Locus 1 e um conjunto de sondagem para o Locus 2, embora, conforme citado anteriormente, vários conjuntos de sondagem possam ser designados para cada locus genômico. 407 e 414 são moléculas alvo correspondentes ao Locus 1 e Locus 2, respectivamente.

[00160] Um primeiro conjunto de sondagem contém os membros 402, 405. 402 contém um rótulo (401) do tipo "A". 405 contém uma marca de afinidade (406) que pode ser usada para isolamento e identificação do produto de sondagem.

[00161] Um segundo conjunto de sondagem com as sondagens de membro 409, 412 apresenta as respectivas características como no primeiro conjunto de sondagem. Entretanto, 409 contém um rótulo (408) do tipo "B", distinguível do tipo "A". 412 contém uma marca de afinidade (413) que pode ser idêntica a ou exclusiva de 406. Muitos conjuntos de sondagem podem ser designados de modo que o "Locus 1" alvo contenha sequências de sondagem exclusivas, mas do mesmo tipo de rótulo "A". De modo similar, muitos conjuntos de sondagem podem ser designados de modo que o "Locus 2" alvo contenha sequências de sondagem exclusivas, mas o mesmo

tipo de rótulo "B". Nesta representação, as marcas de afinidade para os muitos conjuntos de sondagem para o Locus 1 podem ser idênticas ou exclusivas e as marcas de afinidade para os muitos conjuntos de sondagem para o Locus 2 podem ser idênticas ou exclusivas.

[00162] Nesta representação, as sondagens 402 e 405 são hibridizadas nas sequências correspondentes ao Locus 1, mas há uma "lacuna" na molécula de destino que consiste em um ou mais nucleotídeos entre as sondagens hibridizadas 402 e 405. Nesta representação, uma polimerase de DNA ou outra enzima pode ser usada para sintetizar uma nova espécie de polinucleotídeo (404) que une de forma covalente 402 e 405. Ou seja, o produto de sondagem formado neste exemplo é uma molécula de ácido nucleico contígua única com uma sequência correspondente ao Locus 1 e que contém os rótulos e/ou marcas de afinidade acima. Além disso, a 404 pode conter um ou mais rótulos do tipo "C", possivelmente como um resultado de incorporação de um de mais nucleotídeos portadores de um rótulo do tipo "C". Esse exemplo também pode consistir no produto de sondagem criado para o Locus 2, contendo as sondagens 409 e 412. Portanto, os produtos de sondagem conterão uma combinação de rótulos. Para o Locus 1, os produtos de sondagem conterão rótulos do tipo "A" e tipo "C", enquanto os produtos de sondagem do Locus 2 conterão rótulos do tipo "B" e do tipo "C".

[00163] A figura 25 representa uma modificação do procedimento geral descrito na figura 21. A figura 25 representa dois conjuntos de sondagem; um conjunto de sondagem para Locus 1 e um conjunto de sondagem para o Locus 2, embora, conforme citado anteriormente, vários conjuntos de sondagem possam ser designados para cada locus genômico. 505 e 510 são moléculas alvo correspondentes ao Locus 1 e Locus 2, respectivamente. Um primeiro conjunto de sondagem contém os membros 502, 503. 502

contém um rótulo (501) do tipo "A". 503 contém uma marca de afinidade (504) que pode ser usada para isolamento e identificação do produto de sondagem. Um segundo conjunto de sondagem com as sondagens de membro 507, 508 apresenta as respectivas características como no primeiro conjunto de sondagem. Entretanto, 507 contém um rótulo (506) do tipo "B", distinguível do tipo "A". 508 contém uma marca de afinidade (509) que pode ser idêntica a ou exclusiva de 504. Muitos conjuntos de sondagem podem ser designados de modo que o "Locus 1" alvo contenha sequências de sondagem exclusivas, mas do mesmo tipo de rótulo "A". De modo similar, muitos conjuntos de sondagem podem ser designados de modo que o "Locus 2" alvo contenha sequências de sondagem exclusivas, mas o mesmo tipo de rótulo "B". Nesta representação, as marcas de afinidade para os muitos conjuntos de sondagem para o Locus 1 podem ser idênticas ou exclusivas e as marcas de afinidade para os muitos conjuntos de sondagem para o Locus 2 podem ser idênticas ou exclusivas.

[00164] A figura 26 representa uma modificação do procedimento geral descrito na figura 21. A figura 26 representa dois conjuntos de sondagem; um conjunto de sondagem para Locus 1 e um conjunto de sondagem para Locus 2, embora, conforme citado anteriormente, vários conjuntos de sondagem possam ser designados para cada locus genômico. 606 e 612 são moléculas alvo correspondentes ao Locus 1 e Locus 2, respectivamente. Um primeiro conjunto de sondagem contém os membros 602, 603. 602 contém um rótulo (601) do tipo "A". 603 contém uma marca de afinidade (605) que pode ser usada para isolamento e identificação do produto de sondagem. Um segundo conjunto de sondagem com as sondagens de membro 608, 609 apresenta as respectivas características como no primeiro conjunto de sondagem. Entretanto, 608 contém um rótulo (607) do tipo "B", distinguível do tipo "A". 609 contém uma marca de afinidade

(611) que pode ser idêntica a ou exclusiva de 605. Muitos conjuntos de sondagem podem ser designados de modo que o "Locus 1" alvo contenha sequências de sondagem exclusivas, mas do mesmo tipo de rótulo "A". De modo similar, muitos conjuntos de sondagem podem ser designados de modo que o "Locus 2" alvo contenha sequências de sondagem exclusivas, mas o mesmo tipo de rótulo "B". Nesta representação, as marcas de afinidade para os muitos conjuntos de sondagem para o Locus 1 podem ser idênticas ou exclusivas e as marcas de afinidade para os muitos conjuntos de sondagem para o Locus 2 podem ser idênticas ou exclusivas.

[00165] Nesta representação, as sondagens 603 e 609 contêm um ou mais rótulos (604, 610) do tipo "C". Portanto, os produtos de sondagem conterão uma combinação de rótulos. Para o Locus 1, os produtos de sondagem conterão rótulos do tipo "A" e tipo "C", enquanto os produtos de sondagem do Locus 2 conterão rótulos do tipo "B" e do tipo "C".

[00166] A figura 27 representa uma modificação do procedimento geral descrito na figura 21. A figura 27 representa dois conjuntos de sondagem para a identificação de vários alelos do mesmo locus genômico. Por exemplo, para distinguir os alelos maternos e fetais, no caso de um DNA livre de célula isolado de uma mulher grávida ou para distinguir os alelos do hospedeiro e do doador, no caso de DNA livre de célula de um receptor de um transplante de órgão. A figura 27 representa dois conjuntos de sondagem - um conjunto de sondagem para o Alelo 1 e um conjunto de sondagem para o Alelo 2. 706 e 707 são moléculas alvo correspondentes ao Alelo 1 e Alelo 2, respectivamente. Um primeiro conjunto de sondagem contém sondagens de membro 702, 703, 704. 702 contém um rótulo (701) do tipo "A". 704 contém uma marca de afinidade (705) que pode ser usada para isolamento e identificação do produto de sondagem. Um segundo conjunto de sondagem com as sondagens de membro 709, 703, 704 apresenta as

respectivas características como no primeiro conjunto de sondagem. Nesta representação, 703 e 704 são idênticos para os dois conjuntos de sondagem. Entretanto, o item 709 contém um rótulo (708) de tipo "B", distinguível do tipo "A". Nesta representação, 702 e 709 contêm sequências que são quase idênticas e diferem em apenas um nucleotídeo na sequência. Portanto, as sequências de hibridização dessas duas sondagens, que estão configuradas para hibridização nas regiões para o Alelo 1 e o Alelo 2, contêm regiões complementares para o Alelo 1 (702) e Alelo 2 (709). Além disso, o comprimento de cada domínio de hibridização em 702 e 709, bem como as condições de hibridização experimentais são desenvolvidas de modo que a sondagem 702 será hibridizada apenas no Alelo 1 e a sondagem 709 só será hibridizada no Alelo 2. A finalidade desse tipo de ensaio é quantificar precisamente a frequência do Alelo 1 e do Alelo 2 em uma amostra.

[00167] A figura 28 representa uma modificação do procedimento geral descrito na figura 21. A figura 28 representa dois conjuntos de sondagem para a identificação de vários alelos do mesmo locus genômico. Por exemplo, para distinguir os alelos maternos e fetais, no caso de um DNA livre de célula isolado de uma mulher grávida ou para distinguir os alelos do hospedeiro e do doador, no caso de DNA livre de célula de um receptor de um transplante de órgão. A figura 28 representa dois conjuntos de sondagem - um conjunto de sondagem para o Alelo 1 e um conjunto de sondagem para o Alelo 2. 807 e 810 são moléculas alvo correspondentes ao Alelo 1 e Alelo 2, respectivamente. Um primeiro conjunto de sondagem contém sondagens de membro 802, 804, 805. 802 contém um rótulo (801) do tipo "A". 805 contém uma marca de afinidade (806) que pode ser usada para isolamento e identificação do produto de sondagem. Um segundo conjunto de sondagem com as sondagens de membro 809, 804, 805 apresenta as

respectivas características como no primeiro conjunto de sondagem. Nesta representação, 804 e 805 são idênticos para os dois conjuntos de sondagem. Entretanto, o item 809 contém um rótulo (808) de tipo "B", distinguível do tipo "A". Nesta representação, 802 e 809 contêm sequências que são quase idênticas e diferem em apenas um nucleotídeo na sequência. Portanto, as sequências de hibridização dessas duas sondagens contêm regiões complementares para o Alelo 1 (802) e Alelo 2 (809). Além disso, o comprimento de cada domínio de hibridização em 802 e 809, bem como as condições de hibridização experimentais são desenvolvidas de modo que a sondagem 802 será hibridizada apenas no Alelo 1 e a sondagem 809 só será hibridizada no Alelo 2. A finalidade desse tipo de ensaio é ser capaz de quantificar precisamente a frequência do Alelo 1 e do Alelo 2 em uma amostra. Nesta representação, a sondagem 804 contém um ou mais rótulos (803) do tipo "C". Portanto, os produtos de sondagem conterão uma combinação de rótulos. Para o Alelo 1, os produtos de sondagem conterão rótulos do tipo "A" e tipo "C", enquanto os produtos de sondagem do Alelo 2 conterão rótulos do tipo "B" e do tipo "C".

[00168] A figura 29 representa uma modificação do procedimento geral descrito na figura 21. A figura 29 representa dois conjuntos de sondagem para a identificação de vários alelos do mesmo locus genômico. Por exemplo, para distinguir os alelos maternos e fetais, no caso de um DNA livre de célula isolado de uma mulher grávida ou para distinguir os alelos do hospedeiro e do doador, no caso de DNA livre de célula de um receptor de um transplante de órgão. A figura 29 representa dois conjuntos de sondagem - um conjunto de sondagem para o Alelo 1 e um conjunto de sondagem para o Alelo 2.

[00169] 907 e 910 são moléculas alvo correspondentes ao Alelo 1 e Alelo 2, respectivamente. Um primeiro conjunto de sondagem

contém os membros 902, 905. 902 contém um rótulo (901) do tipo "A". O item 905 contém uma marca de afinidade (906) que pode ser usada para isolamento e identificação do produto de sondagem. Um segundo conjunto de sondagem com as sondagens de membro 909, 905 apresenta as respectivas características como no primeiro conjunto de sondagem. Nesta representação, 905 é idêntico para os dois conjuntos de sondagem. Entretanto, o item 909 contém um rótulo (908) de tipo "B", distinguível do tipo "A". Nesta representação, 902 e 909 contêm sequências que são quase idênticas e diferem em apenas um nucleotídeo na sequência. Portanto, as sequências de hibridização dessas duas sondagens contêm regiões complementares para o Alelo 1 (902) e Alelo 2 (909). Além disso, o comprimento de cada domínio de hibridização em 902 e 909, bem como as condições de hibridização experimentais são desenvolvidas de modo que a sondagem 902 será hibridizada apenas no Alelo 1 e a sondagem 909 só será hibridizada no Alelo 2. A finalidade desse tipo de ensaio é ser capaz de quantificar precisamente a frequência do Alelo 1 e do Alelo 2 em uma amostra.

[00170] Nesta representação, as sondagens 902 e 905 são hibridizadas nas sequências correspondentes ao Alelo 1, de modo que há uma "lacuna" na molécula de destino que consiste em um ou mais nucleotídeos entre as sondagens hibridizadas 902 e 905. Nesta representação, uma polimerase de DNA ou outra enzima pode ser usada para sintetizar uma nova espécie de polinucleotídeo (904) que une de forma covalente 902 e 905. Ou seja, o produto de sondagem formado neste exemplo é uma molécula de ácido nucleico contígua única com uma sequência correspondente ao Alelo 1 e que contém os rótulos e/ou marcas de afinidade acima. Além disso, a 904 pode conter um ou mais rótulos do tipo "C", possivelmente como um resultado de incorporação de um nucleotídeo portador de um rótulo do tipo "C". Esse exemplo

também pode consistir no produto de sondagem criado para o Alelo 2, contendo as sondagens 909 e 905.

[00171] A figura 30 representa uma modificação do procedimento geral descrito na figura 21. A figura 30 representa dois conjuntos de sondagem para a identificação de vários alelos do mesmo locus genômico. Por exemplo, para distinguir os alelos maternos e fetais, no caso de um DNA livre de célula isolado de uma mulher grávida ou para distinguir os alelos do hospedeiro e do doador, no caso de DNA livre de célula de um receptor de um transplante de órgão. A figura 30 representa dois conjuntos de sondagem - um conjunto de sondagem para o Alelo 1 e um conjunto de sondagem para o Alelo 2.

[00172] 1006 e 1007 são moléculas alvo correspondentes ao Alelo 1 e Alelo 2, respectivamente. Um primeiro conjunto de sondagem contém sondagens de membro 1001, 1003, 1004. 1003 contém um rótulo (1002) do tipo "A". 1004 contém uma marca de afinidade (1005) que pode ser usada para isolamento e identificação do produto de sondagem.

[00173] Um segundo conjunto de sondagem com as sondagens de membro 1001, 1009, 1004 apresenta as respectivas características como no primeiro conjunto de sondagem. Nesta representação, 1001 é idêntico para ambos os conjuntos de sondagem e 1004 é idêntico para ambos os conjuntos de sondagem. Entretanto, 1009 contém um rótulo (1008) do tipo "B", distinguível do tipo "A".

[00174] Nesta representação, 1003 e 1009 contêm sequências que são quase idênticas e diferem em apenas um nucleotídeo na sequência. Portanto, as sequências de hibridização dessas duas sondagens contêm regiões complementares para o Alelo 1 (1003) e Alelo 2 (1009), respectivamente. Além disso, o comprimento de cada domínio de hibridização em 1003 e 1009, bem como as condições de hibridização experimentais são desenvolvidas de modo que a sondagem 1003 será hibridizada apenas no Alelo 1 e a

sondagem 1009 só será hibridizada no Alelo 2. A finalidade desse tipo de ensaio é ser capaz de quantificar precisamente a frequência do Alelo 1 e do Alelo 2 em uma amostra. Nesta representação, a sondagem 1001 contém um ou mais rótulos (1000) do tipo "C". Portanto, os produtos de sondagem conterão uma combinação de rótulos. Para o Alelo 1, os produtos de sondagem conterão rótulos do tipo "A" e tipo "C", enquanto os produtos de sondagem do Alelo 2 conterão rótulos do tipo "B" e do tipo "C".

[00175] A figura 31 representa uma modificação do procedimento geral descrito na figura 21. A figura 31 representa dois conjuntos de sondagem para a identificação de vários alelos do mesmo locus genômico. Por exemplo, para distinguir os alelos maternos e fetais, no caso de um DNA livre de célula isolado de uma mulher grávida ou para distinguir os alelos do hospedeiro e do doador, no caso de DNA livre de célula de um receptor de um transplante de órgão. A figura 31 representa dois conjuntos de sondagem - um conjunto de sondagem para o Alelo 1 e um conjunto de sondagem para o Alelo 2. 1104 e 1105 são moléculas alvo correspondentes ao Alelo 1 e Alelo 2, respectivamente. Um primeiro conjunto de sondagem contém os membros 1101, 1102. 1101 contém um rótulo (1100) do tipo "A". 1102 contém uma marca de afinidade (1103) que pode ser usada para isolamento e identificação do produto de sondagem. Um segundo conjunto de sondagem com as sondagens de membro 1107, 1102 apresenta as respectivas características como no primeiro conjunto de sondagem. Nesta representação, 1102 é idêntico para os dois conjuntos de sondagem. Entretanto, o item 1107 contém um rótulo (1106) de tipo "B", distinguível do tipo "A". Nesta representação, 1101 e 1107 contêm sequências que são quase idênticas e diferem em apenas um nucleotídeo na sequência. Portanto, as sequências de hibridização dessas duas sondagens contêm regiões complementares para o Alelo 1 (1101) e Alelo 2

(1107). Além disso, o comprimento de cada domínio de hibridização em 1101 e 1107, bem como as condições de hibridização experimentais são desenvolvidas de modo que a sondagem 1101 será hibridizada apenas no Alelo 1 e a sondagem 1107 só será hibridizada no Alelo 2. A finalidade desse tipo de ensaio é ser capaz de quantificar precisamente a frequência do Alelo 1 e do Alelo 2 em uma amostra.

[00176] A figura 32 representa uma modificação do procedimento geral descrito na figura 21. A figura 32 representa dois conjuntos de sondagem para a identificação de vários alelos do mesmo locus genômico. Por exemplo, para distinguir os alelos maternos e fetais, no caso de um DNA livre de célula isolado de uma mulher grávida ou para distinguir os alelos do hospedeiro e do doador, no caso de DNA livre de célula de um receptor de um transplante de órgão. A figura 32 representa dois conjuntos de sondagem - um conjunto de sondagem para o Alelo 1 e um conjunto de sondagem para o Alelo 2. 1206 e 1207 são moléculas alvo correspondentes ao Alelo 1 e Alelo 2, respectivamente. Um primeiro conjunto de sondagem contém os membros 1202, 1203. 1202 contém um rótulo (1201) do tipo "A". 1203 contém uma marca de afinidade (1205) que pode ser usada para isolamento e identificação do produto de sondagem. Um segundo conjunto de sondagem com as sondagens de membro 1209, 1203 apresenta as respectivas características como no primeiro conjunto de sondagem. Nesta representação, 1203 é idêntico para os dois conjuntos de sondagem. Entretanto, o item 1209 contém um rótulo (1208) de tipo "B", distinguível do tipo "A". Nesta representação, 1202 e 1209 contêm sequências que são quase idênticas e diferem em apenas um nucleotídeo na sequência. Portanto, as sequências de hibridização dessas duas sondagens contêm regiões complementares para o Alelo 1 (1202) e Alelo 2 (1209). Além disso, o comprimento de cada domínio de

hibridização em 1202 e 1209, bem como as condições de hibridização experimentais são desenvolvidas de modo que a sondagem 1202 será hibridizada apenas no Alelo 1 e a sondagem 1209 só será hibridizada no Alelo 2. A finalidade desse tipo de ensaio é ser capaz de quantificar precisamente a frequência do Alelo 1 e do Alelo 2 em uma amostra. Nesta representação, a sondagem 1203 contém um ou mais rótulos (1204) do tipo "C". Portanto, o produto de sondagem conterão uma combinação de rótulos. Para o Alelo 1, os produtos de sondagem conterão rótulos do tipo "A" e tipo "C", enquanto os produtos de sondagem do Alelo 2 conterão rótulos do tipo "B" e do tipo "C".

[00177] A figura 33 representa uma modificação do procedimento geral descrito na figura 21. A figura 33 representa dois conjuntos de sondagem; um conjunto de sondagem para Locus 1 e um conjunto de sondagem para o Locus 2, embora, conforme citado anteriormente, vários conjuntos de sondagem possam ser designados para cada locus genômico. 1304 e 1305 são moléculas alvo correspondentes ao Locus 1 e Locus 2, respectivamente. Um primeiro conjunto de sondagem contém os membros 1301, 1302. 1301 contém um rótulo (1300) do tipo "A". 1301 contém uma marca de afinidade (1303) que pode ser usada para isolamento e identificação do produto de sondagem. Um segundo conjunto de sondagem com as sondagens de membro 1307, 1308 apresenta as respectivas características como no primeiro conjunto de sondagem. Entretanto, 1307 contém um rótulo (1306) do tipo "B", distinguível do tipo "A". 1307 contém uma marca de afinidade (1309) que pode ser idêntica a ou exclusiva de 1303. Muitos conjuntos de sondagem podem ser designados de modo que o "Locus 1" alvo contenha sequências de sondagem exclusivas, mas do mesmo tipo de rótulo "A". De modo similar, muitos conjuntos de sondagem podem ser designados de modo que o "Locus 2" alvo contenha sequências de sondagem exclusivas, mas o mesmo tipo de

rótulo "B". Nesta representação, as marcas de afinidade para os muitos conjuntos de sondagem para o Locus 1 podem ser idênticas ou exclusivas e as marcas de afinidade para os muitos conjuntos de sondagem para o Locus 2 podem ser idênticas ou exclusivas. Nesta representação, as sondagens 1301 e 1307 têm estruturas similares. Por exemplo, na sondagem 1301, há dois domínios de hibridização distintos, de modo que a sondagem 1302 pode ser ligada a cada final de 1301, formando um produto de sondagem que consiste em uma molécula contígua topologicamente fechada de DNA (por ex., uma molécula circular). A sequência de não hibridização na sondagem 1301 pode conter características adicionais, possivelmente locais de enzima de restrição ou locais de ligação do primer para amplificação universal.

[00178] Uma característica dessa representação é que todos os produtos de sondagem são moléculas circulares contíguas. Dessa maneira, os produtos de sondagem podem ser isolados de todos os outros ácidos nucleicos via degradação enzimática de todas as moléculas de ácido nucleico lineares, por exemplo, usando uma exonuclease.

[00179] A figura 34 representa uma modificação do procedimento geral descrito na figura 21. A figura 34 representa dois conjuntos de sondagem; um conjunto de sondagem para Locus 1 e um conjunto de sondagem para o Locus 2, embora, conforme citado anteriormente, vários conjuntos de sondagem possam ser designados para cada locus genômico. 1405 e 1406 são moléculas alvo correspondentes ao Locus 1 e Locus 2, respectivamente. Um primeiro conjunto de sondagem contém os membros 1401, 1403. 1401 contém um rótulo (1400) do tipo "A". 1401 contém uma marca de afinidade (1404) que pode ser usada para isolamento e identificação do produto de sondagem. Um segundo conjunto de sondagem com as sondagens de membro 1408, 1410 apresenta as respectivas características como no primeiro conjunto de

sondagem. Entretanto, 1408 contém um rótulo (1407) do tipo "B", distinguível do tipo "A". 1408 contém uma marca de afinidade (1411) que pode ser idêntica a ou exclusiva de 1404. Muitos conjuntos de sondagem podem ser designados de modo que o "Locus 1" alvo contenha sequências de sondagem exclusivas, mas do mesmo tipo de rótulo "A". De modo similar, muitos conjuntos de sondagem podem ser designados de modo que o "Locus 2" alvo contenha sequências de sondagem exclusivas, mas o mesmo tipo de rótulo "B". Nesta representação, as marcas de afinidade para os muitos conjuntos de sondagem para o Locus 1 podem ser idênticas ou exclusivas e as marcas de afinidade para os muitos conjuntos de sondagem para o Locus 2 podem ser idênticas ou exclusivas. Nesta representação, as sondagens 1401 e 1408 têm estruturas similares. Por exemplo, na sondagem 1401, há dois domínios de hibridização distintos, de modo que a sondagem 1403 pode ser ligada a cada final de 1401, formando um produto de sondagem que consiste em uma molécula contígua topologicamente fechada de DNA (por ex., uma molécula circular). A sequência de não hibridização na sondagem 1401 pode conter características adicionais, possivelmente locais de enzima de restrição ou locais de ligação do primer para amplificação universal.

[00180] Uma característica dessa representação é que todos os produtos de sondagem são moléculas circulares contíguas. Dessa maneira, os produtos de sondagem podem ser isolados de todos os outros ácidos nucleicos via degradação enzimática de todas as moléculas de ácido nucleico lineares, por exemplo, usando uma exonuclease. Nesta representação, as sondagens 1403 e 1410 contêm um ou mais rótulos (1402, 1409) do tipo "C". Portanto, os produtos de sondagem conterão uma combinação de rótulos. Para o Locus 1, os produtos de sondagem conterão rótulos do tipo "A" e tipo "C", enquanto os produtos de sondagem do Locus 2 conterão rótulos do tipo "B" e do tipo "C".

[00181] A figura 35 representa uma modificação do procedimento geral descrito na figura 21. A figura 35 representa dois conjuntos de sondagem; um conjunto de sondagem para Locus 1 e um conjunto de sondagem para o Locus 2, embora, conforme citado anteriormente, vários conjuntos de sondagem possam ser designados para cada locus genômico. 1505 e 1506 são moléculas alvo correspondentes ao Locus 1 e Locus 2, respectivamente. Um primeiro conjunto de sondagem contém a sondagem de membro 1501. 1501 contém um rótulo (1500) do tipo "A". 1501 contém uma marca de afinidade (1504) que pode ser usada para isolamento e identificação do produto de sondagem. Um segundo conjunto de sondagem com a sondagem de membro 1508 apresenta as respectivas características como no primeiro conjunto de sondagem. Entretanto, 1508 contém um rótulo (1507) do tipo "B", distinguível do tipo "A". 1508 contém uma marca de afinidade (1511) que pode ser idêntica a ou exclusiva de 1504. Muitos conjuntos de sondagem podem ser designados de modo que o "Locus 1" alvo contenha sequências de sondagem exclusivas, mas do mesmo tipo de rótulo "A". De modo similar, muitos conjuntos de sondagem podem ser designados de modo que o "Locus 2" alvo contenha sequências de sondagem exclusivas, mas o mesmo tipo de rótulo "B". Nesta representação, as marcas de afinidade para os muitos conjuntos de sondagem para o Locus 1 podem ser idênticas ou exclusivas e as marcas de afinidade para os muitos conjuntos de sondagem para o Locus 2 podem ser idênticas ou exclusivas. Nesta representação, as sondagens 1501 e 1508 têm estruturas similares.

[00182] Por exemplo, na sondagem 1501, há dois domínios de hibridização distintos, de modo que quando hibridizados em uma molécula alvo, há uma lacuna entre os dois domínios de hibridização. Nesta representação, uma polimerase de DNA ou outra enzima pode ser usada para sintetizar uma nova espécie de

polinucleotídeo (1503) que preenche a lacuna de forma covalente entre os domínios de 1501. Ou seja, o produto de sondagem formado neste exemplo é uma molécula de DNA contígua única, topologicamente fechada (por ex., uma molécula circular) com uma sequência correspondente ao Locus 1 e que contém os rótulos e/ou marcas de afinidade acima. Além disso, a 1503 pode conter um ou mais rótulos do tipo "C", possivelmente como um resultado de incorporação de um nucleotídeo portador de um rótulo do tipo "C". Esse exemplo também pode consistir no produto de sondagem criado para o Locus 2, contendo a sondagem 1508. A sequência de não hibridização na sondagem 1501 e sondagem 1508 pode conter características adicionais, possivelmente locais da enzima de restrição. Uma característica dessa representação é que todos os produtos de sondagem são moléculas circulares contíguas. Dessa maneira, os produtos de sondagem podem ser isolados de todos os outros ácidos nucleicos via degradação enzimática de todas as moléculas de ácido nucleico lineares, por exemplo, usando uma exonuclease. Os produtos de sondagem conterão uma combinação de rótulos. Para o Locus 1, os produtos de sondagem conterão rótulos do tipo "A" e tipo "C", enquanto os produtos de sondagem do Locus 2 conterão rótulos do tipo "B" e do tipo "C".

[00183] A figura 36 representa uma modificação do procedimento geral descrito na figura 21. A figura 36 representa dois conjuntos de sondagem; um conjunto de sondagem para Locus 1 e um conjunto de sondagem para o Locus 2, embora, conforme citado anteriormente, vários conjuntos de sondagem possam ser designados para cada locus genômico. 1605 e 1606 são moléculas alvo correspondentes ao Locus 1 e Locus 2, respectivamente.

[00184] Um primeiro conjunto de sondagem contém a sondagem de membro 1602. 1602 contém um rótulo (1600) do tipo "A". 1602 contém uma marca de afinidade (1601) que pode ser usada para isolamento e identificação do produto de sondagem.

[00185] Um segundo conjunto de sondagem com a sondagem de membro 1609 apresenta as respectivas características como no primeiro conjunto de sondagem. Entretanto, 1609 contém um rótulo (1608) do tipo "B", distinguível do tipo "A". 1609 contém uma marca de afinidade (1607) que pode ser idêntica a ou exclusiva de 1601. Muitos conjuntos de sondagem podem ser designados de modo que o "Locus 1" alvo contenha sequências de sondagem exclusivas, mas do mesmo tipo de rótulo "A". De modo similar, muitos conjuntos de sondagem podem ser designados de modo que o "Locus 2" alvo contenha sequências de sondagem exclusivas, mas o mesmo tipo de rótulo "B". Nesta representação, as marcas de afinidade para os muitos conjuntos de sondagem para o Locus 1 podem ser idênticas ou exclusivas e as marcas de afinidade para os muitos conjuntos de sondagem para o Locus 2 podem ser idênticas ou exclusivas.

[00186] Nesta representação, as sondagens 1602 e 1609 são hibridizadas nas sequências correspondentes ao Locus 1 e Locus 2, respectivamente, e uma polimerase de DNA ou outra enzima pode ser usada para sintetizar uma nova sequência de polinucleotídeo, por exemplo, 1603 no caso de Locus 1 ou 1611 no caso de Locus 2. Nesta representação, 1603 e 1611 podem conter um ou mais rótulos (1604) do tipo "C", possivelmente como um resultado de incorporação de um de mais nucleotídeos portadores de um rótulo do tipo "C". Esse exemplo também leva ao produto de sondagem criado para o Locus 2. Portanto, os produtos de sondagem conterão uma combinação de rótulos. Para o Locus 1, os produtos de sondagem conterão rótulos do tipo "A" e tipo "C", enquanto os produtos de sondagem do Locus 2 conterão rótulos do tipo "B" e do tipo "C". Esta representação resulta nos produtos de sondagem com especificidade elevada para sequências no Locus 1 ou Locus 2, respectivamente.

[00187] A figura 37 representa uma modificação do procedimento geral descrito na figura 21. A figura 37 representa dois conjuntos de sondagem; um conjunto de sondagem para Locus 1 e um conjunto de sondagem para o Locus 2, embora, conforme citado anteriormente, vários conjuntos de sondagem possam ser designados para cada locus genômico. 1704 e 1705 são moléculas alvo correspondentes ao Locus 1 e Locus 2, respectivamente.

[00188] Um primeiro conjunto de sondagem contém a sondagem de membro 1702. 1702 contém uma marca de afinidade (1700) que pode ser usada para isolamento e identificação do produto de sondagem.

[00189] Um segundo conjunto de sondagem com a sondagem de membro 1708 apresenta as respectivas características como no primeiro conjunto de sondagem. 1708 contém a marca de afinidade (1706) que pode ser idêntica ou exclusiva de 1700. Muitos conjuntos de sondagem podem ser desenvolvidos de modo que o "Locus 1" alvo contenha sequências de sondagem exclusivas. Da mesma maneira, muitos conjuntos de sondagem podem ser desenvolvidos de modo que o "Locus 2" alvo contenha sequências de sondagem exclusivas. Nesta representação, as marcas de afinidade para os muitos conjuntos de sondagem do Locus 1 podem ser idênticas ou exclusivas, e as marcas de afinidade para os muitos conjuntos de sondagem do Locus 2 podem ser idênticas ou exclusivas.

[00190] Nesta representação, as sondagens 1702 e 1708 são hibridizadas nas sequências correspondentes ao Locus 1 e ao Locus 2, respectivamente. Os projetos de cada sondagem para Locus 1 e Locus 2 são aqueles em que o primeiro nucleotídeo adjacente próximo aos domínios de hibridização contém um nucleotídeo diferente do Locus 2 para Locus 1. Neste exemplo, o primeiro nucleotídeo adjacente próximo ao domínio de hibridização de 1702 é "A", enquanto o primeiro nucleotídeo

adjacente próximo ao domínio de hibridização de 1708 é "T". Nesta representação, todas as sondagens para Locus 1 devem ser desenvolvidas de modo que o primeiro nucleotídeo imediatamente adjacente ao domínio de hibridização consista em diferentes nucleotídeos do primeiro nucleotídeo imediatamente adjacente ao domínio de hibridização das sondagens para o Locus 2. Ou seja, pelo projeto, os conjuntos de sondagem de Locus 1 e Locus 2 podem ser distinguidos uns dos outros com base na identidade do primeiro nucleotídeo imediatamente adjacente ao domínio de hibridização.

[00191] Nesta representação, uma polimerase de DNA ou outra enzima será usada para adicionar pelo menos um nucleotídeo adicional a cada uma das sequências de sondagem. Neste exemplo, os substratos de nucleotídeo para a polimerase de DNA são competentes para uma adição única, por exemplo, os nucleotídeos podem ser terminadores da cadeia de dideoxi. Ou seja, somente um novo nucleotídeo deve ser adicionado a cada sequência de sondagem. Neste exemplo, o nucleotídeo adicionado à sondagem 1702 conterá um ou mais rótulos (1703) do tipo "A". O nucleotídeo adicionado à sondagem 1708 conterá um ou mais rótulos (1709) do tipo "B", de modo que os produtos de sondagem do Locus 1 possam ser distinguidos dos produtos de sondagem do Locus 2.

[00192] A figura 38 representa uma modificação do procedimento geral descrito na figura 21. A figura 38 representa dois conjuntos de sondagem; um conjunto de sondagem para Locus 1 e um conjunto de sondagem para o Locus 2, embora, conforme citado anteriormente, vários conjuntos de sondagem possam ser designados para cada locus genômico. 1804 e 1805 são moléculas alvo correspondentes ao Locus 1 e Locus 2, respectivamente.

[00193] Um primeiro conjunto de sondagem contém a sondagem de membro 1802. 1802 contém uma marca de afinidade (1800) que pode

ser usada para isolamento e identificação do produto de sondagem.

[00194] Um segundo conjunto de sondagem com a sondagem de membro 1808 apresenta as respectivas características como no primeiro conjunto de sondagem. 1808 contém a marca de afinidade (1806) que pode ser idêntica ou exclusiva de 1800. Muitos conjuntos de sondagem podem ser desenvolvidos de modo que o "Locus 1" alvo contenha sequências de sondagem exclusivas. Da mesma maneira, muitos conjuntos de sondagem podem ser desenvolvidos de modo que o "Locus 2" alvo contenha sequências de sondagem exclusivas. Nesta representação, as marcas de afinidade para os muitos conjuntos de sondagem do Locus 1 podem ser idênticas ou exclusivas, e as marcas de afinidade para os muitos conjuntos de sondagem do Locus 2 podem ser idênticas ou exclusivas.

[00195] Nesta representação, as sondagens 1802 e 1808 são hibridizadas nas sequências correspondentes ao Locus 1 e ao Locus 2, respectivamente. Os projetos de cada sondagem para Locus 1 e Locus 2 são aqueles em que o primeiro nucleotídeo adjacente próximo aos domínios de hibridização contém um nucleotídeo diferente do Locus 2 para Locus 1. Neste exemplo, o primeiro nucleotídeo adjacente próximo ao domínio de hibridização de 1802 é "A", enquanto o primeiro nucleotídeo adjacente próximo ao domínio de hibridização de 1808 é "T". Nesta representação, todas as sondagens para Locus 1 devem ser desenvolvidas de modo que o primeiro nucleotídeo imediatamente adjacente ao domínio de hibridização consista em diferentes nucleotídeos do primeiro nucleotídeo imediatamente adjacente ao domínio de hibridização das sondagens para o Locus 2. Ou seja, pelo projeto, os conjuntos de sondagem de Locus 1 e Locus 2 podem ser distinguidos uns dos outros com base na identidade do

primeiro nucleotídeo imediatamente adjacente ao domínio de hibridização.

[00196] Nesta representação, uma polimerase de DNA ou outra enzima será usada para adicionar pelo menos um nucleotídeo adicional a cada uma das sequências de sondagem. Neste exemplo, os substratos de nucleotídeo para a polimerase de DNA são competentes para uma adição única, talvez porque os nucleotídeos adicionados à mistura da reação são nucleotídeos de dideoxi. Ou seja, somente um novo nucleotídeo deve ser adicionado a cada sequência de sondagem. Neste exemplo, o nucleotídeo adicionado à sondagem 1802 conterá um ou mais rótulos (1803) do tipo "A". O nucleotídeo adicionado à sondagem 1808 conterá um ou mais rótulos (1809) do tipo "B", de modo que os produtos de sondagem do Locus 1 possam ser distinguidos dos produtos de sondagem do Locus 2.

[00197] Nesta representação, as sondagens 1802 e 1808 contêm um ou mais rótulos (1801, 1806) do tipo "C". Portanto, os produtos de sondagem conterão uma combinação de rótulos. Para o Locus 1, os produtos de sondagem conterão rótulos do tipo "A" e tipo "C", enquanto os produtos de sondagem do Locus 2 conterão rótulos do tipo "B" e do tipo "C".

[00198] A figura 39 representa uma modificação do procedimento geral descrito na figura 21. A figura 39 representa dois conjuntos de sondagem; um conjunto de sondagem para Locus 1 e um conjunto de sondagem para o Locus 2, embora, conforme citado anteriormente, vários conjuntos de sondagem possam ser designados para cada locus genômico. 1906 e 1907 são moléculas alvo correspondentes ao Locus 1 e Locus 2, respectivamente.

[00199] Um primeiro conjunto de sondagem contém a sondagem de membro 1902. 1902 contém uma marca de afinidade (1901) que pode ser usada para isolamento e identificação do produto de sondagem.

[00200] Um segundo conjunto de sondagem com a sondagem de membro 1910 apresenta as respectivas características como no primeiro conjunto de sondagem. 1910 contém a marca de afinidade (1908) que pode ser idêntica ou exclusiva de 1901. Muitos conjuntos de sondagem podem ser desenvolvidos de modo que o "Locus 1" alvo contenha sequências de sondagem exclusivas. Da mesma maneira, muitos conjuntos de sondagem podem ser desenvolvidos de modo que o "Locus 2" alvo contenha sequências de sondagem exclusivas. Nesta representação, as marcas de afinidade para os muitos conjuntos de sondagem do Locus 1 podem ser idênticas ou exclusivas, e as marcas de afinidade para os muitos conjuntos de sondagem do Locus 2 podem ser idênticas ou exclusivas.

[00201] Nesta representação, as sondagens 1902 e 1910 são hibridizadas nas sequências correspondentes ao Locus 1 e ao Locus 2, respectivamente. Os projetos de cada sondagem para Locus 1 e Locus 2 são aqueles em que o primeiro nucleotídeo adjacente próximo aos domínios de hibridização contém um nucleotídeo diferente do Locus 2 para Locus 1. Neste exemplo, o primeiro nucleotídeo adjacente próximo ao domínio de hibridização de 1902 é "A", enquanto o primeiro nucleotídeo adjacente próximo ao domínio de hibridização de 1910 é "T". Nesta representação, todas as sondagens para Locus 1 devem ser desenvolvidas de modo que o primeiro nucleotídeo imediatamente adjacente ao domínio de hibridização consista em diferentes nucleotídeos do primeiro nucleotídeo imediatamente adjacente ao domínio de hibridização das sondagens para o Locus 2. Ou seja, pelo projeto, os conjuntos de sondagem de Locus 1 e Locus 2 podem ser distinguidos de outro nucleotídeo na identidade do primeiro nucleotídeo imediatamente adjacente ao domínio de hibridização. Um nucleotídeo diferente, não usado para distinguir sondagens de Locus 1 ou Locus 2 deve servir como um

terminador de cadeia. Neste exemplo em particular, um nucleotídeo "A" em uma molécula alvo é usado para distinguir sondagens do Locus 1, e um nucleotídeo "T" é usado para distinguir sondagens do Locus 2. Neste exemplo, um nucleotídeo "C" pode servir como um terminador de cadeia. Neste caso, um nucleotídeo "C" será adicionado ao ensaio que não será capaz do alongamento da cadeia (por exemplo, um dideoxi C). Uma restrição adicional é que as sequências de sondagem são desenvolvidas de modo que não haja casos de uma identificação de nucleotídeo do Locus 2 presente em 1906 entre o nucleotídeo de distinção do Locus 1 e o nucleotídeo de terminação da cadeia. Neste exemplo, não haverá nucleotídeos "T" presentes em 1906 após o domínio de hibridização de 1902 e antes de G, que será emparelhado com o terminador C da cadeia.

[00202] Nesta representação, a polimerase de DNA ou uma enzima similar será usada para sintetizar nova sequências de nucleotídeo, e o nucleotídeo adicionado no local do nucleotídeo distinto para o Locus 1 conterá um ou mais rótulos (1903) do tipo "A". O nucleotídeo adicionado no local do nucleotídeo distinto para o Locus 2 conterá 1 ou mais rótulos (1911) do tipo "B", de modo que os produtos de sondagem para o Locus 1 possam ser distintos dos produtos de sondagem do Locus 2. Nesta representação, o nucleotídeo adicionado na posição de terminação da cadeia conterá um ou mais rótulos (1912) do tipo "C". Portanto, os produtos de sondagem conterão uma combinação de rótulos. Para o Locus 1, os produtos de sondagem conterão rótulos do tipo "A" e tipo "C", enquanto os produtos de sondagem do Locus 2 conterão rótulos do tipo "B" e do tipo "C".

[00203] Em outra representação, o terminador da cadeia pode não conter rótulo. Nesta representação, um quarto nucleotídeo pode ser adicionado ao ensaio que contém um ou mais rótulos do tipo "C". Esse quarto nucleotídeo não é emparelhado com o

nucleotídeo de identificação para o Alelo 1 (neste exemplo, A), não é emparelhado com o nucleotídeo de identificação para o Alelo 2 (neste exemplo, T), não é emparelhado com o nucleotídeo de terminação da cadeia (neste exemplo, G). Neste exemplo, o quarto nucleotídeo que teria um ou mais rótulos do tipo "C" é G e será emparelhado com os locais C em 1906 e 1907. Portanto, os produtos de sondagem conterão uma combinação de rótulos. Para o Locus 1, os produtos de sondagem conterão rótulos do tipo "A" e tipo "C", enquanto os produtos de sondagem do Locus 2 conterão rótulos do tipo "B" e do tipo "C".

[00204] A figura 40 representa uma modificação do procedimento geral descrito na figura 21. A figura 40 representa dois conjuntos de sondagem; um conjunto de sondagem para Locus 1 e um conjunto de sondagem para o Locus 2, embora, conforme citado anteriormente, vários conjuntos de sondagem possam ser designados para cada locus genômico. 2005 e 2006 são moléculas alvo correspondentes ao Locus 1 e Locus 2, respectivamente.

[00205] Um primeiro conjunto de sondagem contém a sondagem de membro 2001. 2001 contém uma marca de afinidade (2000) que pode ser usada para isolamento e identificação do produto de sondagem.

[00206] Um segundo conjunto de sondagem com a sondagem de membro 2008 apresenta as respectivas características como no primeiro conjunto de sondagem. 2008 contém a marca de afinidade (2007) que pode ser idêntica ou exclusiva de 2000. Muitos conjuntos de sondagem podem ser desenvolvidos de modo que o "Locus 1" alvo contenha sequências de sondagem exclusivas. Da mesma maneira, muitos conjuntos de sondagem podem ser desenvolvidos de modo que o "Locus 2" alvo contenha sequências de sondagem exclusivas. Nesta representação, as marcas de afinidade para os muitos conjuntos de sondagem do Locus 1 podem ser idênticas ou exclusivas, e as marcas de afinidade para os

muitos conjuntos de sondagem do Locus 2 podem ser idênticas ou exclusivas.

[00207] Nesta representação, as sondagens 2001 e 2008 são hibridizadas nas sequências correspondentes ao Locus 1 e ao Locus 2, respectivamente. Os projetos de cada sondagem para o Locus 1 e o Locus 2 são aqueles em que há um ou mais casos de um nucleotídeo de distinção (neste exemplo, "A" é um nucleotídeo de distinção do Locus 1 e "T" é nucleotídeo de distinção do Locus 2) seguido por um nucleotídeo de terminação de cadeia (neste exemplo, "G") adjacente ao domínio de hibridização das sondagens. É importante notar que não haverá casos do nucleotídeo de distinção para o Locus 2 (neste exemplo, "T") presente entre o domínio de hibridização de 2001 em 2005 e o nucleotídeo de terminação de cadeia em 2005. De modo similar, não haverá ocorrência do nucleotídeo de distinção para o Locus 1 (neste exemplo, "T") presente entre o domínio de hibridização de 2008 em 2006 e o nucleotídeo de terminação de cadeia em 2006.

[00208] Nesta representação, a polimerase de DNA ou uma enzima similar será usada para sintetizar nova sequências de nucleotídeo (2004, 2011) até a adição de um nucleotídeo de terminação de cadeia, um exemplo possível seria um dideóxi C. Nesta representação, os nucleotídeos adicionados no local do nucleotídeo distinto para o Locus 1 conterão um ou mais rótulos (2003) do tipo "A". Os nucleotídeos adicionados nos locais do nucleotídeo distinto para o Locus 2 conterão 1 ou mais rótulos (2010) do tipo "B", de modo que os produtos de sondagem para o Locus 1 possam ser claramente distintos dos produtos de sondagem do Locus 2.

[00209] A figura 41 representa uma modificação do procedimento geral descrito na figura 21. A figura 41 representa dois conjuntos de sondagem; um conjunto de sondagem para Locus 1 e um conjunto de sondagem para o Locus 2, embora, conforme citado

anteriormente, vários conjuntos de sondagem possam ser designados para cada locus genômico. 2105 e 2106 são moléculas alvo correspondentes ao Locus 1 e Locus 2, respectivamente.

[00210] Um primeiro conjunto de sondagem contém a sondagem de membro 2102. 2102 contém uma marca de afinidade (2100) que pode ser usada para isolamento e identificação do produto de sondagem.

[00211] Um segundo conjunto de sondagem com a sondagem de membro 2109 apresenta as respectivas características como no primeiro conjunto de sondagem. 2109 contém a marca de afinidade (2107) que pode ser idêntica ou exclusiva de 2100. Muitos conjuntos de sondagem podem ser desenvolvidos de modo que o "Locus 1" alvo contenha sequências de sondagem exclusivas. Da mesma maneira, muitos conjuntos de sondagem podem ser desenvolvidos de modo que o "Locus 2" alvo contenha sequências de sondagem exclusivas. Nesta representação, as marcas de afinidade para os muitos conjuntos de sondagem do Locus 1 podem ser idênticas ou exclusivas, e as marcas de afinidade para os muitos conjuntos de sondagem do Locus 2 podem ser idênticas ou exclusivas.

[00212] Nesta representação, as sondagens 2102 e 2109 são hibridizadas nas sequências correspondentes ao Locus 1 e ao Locus 2, respectivamente. Os projetos de cada sondagem para o Locus 1 e o Locus 2 são aqueles em que há um ou mais casos de um nucleotídeo de distinção (neste exemplo, "A" é um nucleotídeo de distinção do Locus 1 e "T" é nucleotídeo de distinção do Locus 2) seguido por um nucleotídeo de terminação de cadeia (neste exemplo, "G") adjacente ao domínio de hibridização das sondagens. É importante notar que não haverá casos do nucleotídeo de distinção para o Locus 2 (neste exemplo, "T") presente entre o domínio de hibridização de 2102 em 2105 e o nucleotídeo de terminação de cadeia em 2105. De modo similar,

não haverá ocorrência do nucleotídeo de distinção para o Locus 1 (neste exemplo, "T") presente entre o domínio de hibridização de 2109 em 2106 e o nucleotídeo de terminação de cadeia em 2106.

[00213] Nesta representação, a polimerase de DNA ou uma enzima similar será usada para sintetizar nova sequências de nucleotídeo (2104, 2110) até a adição de um nucleotídeo de terminação de cadeia, um exemplo possível seria um dideóxi C. Nesta representação, os nucleotídeos adicionados no local do nucleotídeo distinto para o Locus 1 conterão um ou mais rótulos (2103) do tipo "A". Os nucleotídeos adicionados nos locais do nucleotídeo distinto para o Locus 2 conterão 1 ou mais rótulos (2110) do tipo "B", de modo que os produtos de sondagem para o Locus 1 possam ser claramente distintos dos produtos de sondagem do Locus 2.

[00214] Nesta representação, as sondagens 2102 e 2109 contêm um ou mais rótulos (2101, 2108) do tipo "C". Portanto, os produtos de sondagem conterão uma combinação de rótulos. Para o Locus 1, os produtos de sondagem conterão rótulos do tipo "A" e tipo "C", enquanto os produtos de sondagem do Locus 2 conterão rótulos do tipo "B" e do tipo "C".

[00215] A figura 42 representa uma modificação do procedimento geral descrito na figura 21. A figura 42 representa dois conjuntos de sondagem para a identificação de vários alelos do mesmo locus genômico. Por exemplo, para distinguir os alelos maternos e fetais, no caso de um DNA livre de célula isolado de uma mulher grávida ou para distinguir os alelos do hospedeiro e do doador, no caso de DNA livre de célula de um receptor de um transplante de órgão. A figura 42 representa dois conjuntos de sondagem - um conjunto de sondagem para o Alelo 1 e um conjunto de sondagem para o Alelo 2. 2203 e 2204 são moléculas alvo correspondentes ao Alelo 1 e Alelo 2, respectivamente.

[00216] Um primeiro conjunto de sondagem contém a sondagem de membro 2201. 2201 contém uma marca de afinidade (2200) que pode ser usada para isolamento e identificação do produto de sondagem. Nesta representação, os conjuntos de sondagem usados para identificação de dois alelos diferentes são iguais. Ou seja, o conjunto de sondagem para o Alelo 2 consiste na sondagem do membro 2201. Nesta representação, a sondagem 2201 é hibridizada em uma sequência correspondente no Alelo 1 e no Alelo 2, respectivamente na figura 42. O projeto da sondagem 2201 é aquele em que o primeiro nucleotídeo adjacente próximo ao domínio de hibridização contém um nucleotídeo diferente do Alelo 2 para o Alelo 1. Em outras palavras, o primeiro nucleotídeo adjacente ao domínio de hibridização pode ser um polimorfismo de nucleotídeo único ou SNP. Neste exemplo, o primeiro nucleotídeo adjacente em 2203 próximo ao domínio de hibridização de 2201 é um "A", enquanto o primeiro nucleotídeo adjacente em 2204 próximo ao domínio de hibridização de 2201 é um "T". Ou seja, os produtos de sondagem do Alelo 1 e Alelo 2 podem ser distinguidos entre si com base na identidade do primeiro nucleotídeo imediatamente adjacente ao domínio de hibridização.

[00217] Nesta representação, uma polimerase de DNA ou outra enzima será usada para adicionar pelo menos um nucleotídeo adicional a cada uma das sequências de sondagem. Neste exemplo, os substratos de nucleotídeo para a polimerase de DNA são competentes para uma adição única, talvez porque os nucleotídeos adicionados à mistura da reação são nucleotídeos de dideoxi. Ou seja, somente um novo nucleotídeo deve ser adicionado a cada sequência de sondagem. Neste exemplo, o nucleotídeo adicionado à sondagem 2201 para o Alelo 1 conterá um ou mais rótulos (2202) do tipo "A". O nucleotídeo adicionado à sondagem 2201 conterá um ou mais rótulos (2205) do tipo "B", de modo que os produtos de sondagem do Alelo 1 possam ser claramente distinguidos dos

produtos de sondagem do Alelo 2. Ou seja, o produto de sondagem para o Alelo 1 consiste na sondagem 2201 mais um nucleotídeo adicional que possui um ou mais rótulos do tipo "A", e os produtos de sondagem para o Alelo 2 consistem na sondagem 2201 mais um nucleotídeo adicional portador de um ou mais rótulos do tipo "B".

[00218] A figura 43 representa uma modificação do procedimento geral descrito na figura 21. A figura 43 representa dois conjuntos de sondagem para a identificação de vários alelos do mesmo locus genômico. Por exemplo, para distinguir os alelos maternos e fetais, no caso de um DNA livre de célula isolado de uma mulher grávida ou para distinguir os alelos do hospedeiro e do doador, no caso de DNA livre de célula de um receptor de um transplante de órgão. A figura 43 representa dois conjuntos de sondagem - um conjunto de sondagem para o Alelo 1 e um conjunto de sondagem para o Alelo 2. 2304 e 2305 são moléculas alvo correspondentes ao Alelo 1 e Alelo 2, respectivamente.

[00219] Um primeiro conjunto de sondagem contém a sondagem de membro 2302. 2302 contém uma marca de afinidade (2300) que pode ser usada para isolamento e identificação do produto de sondagem. Nesta representação, os conjuntos de sondagem usados para identificação de dois alelos diferentes são iguais. Ou seja, o conjunto de sondagem para o Alelo 2 consiste na sondagem do membro 2302. Nesta representação, a sondagem 2302 é hibridizada em uma sequência correspondente no Alelo 1 e no Alelo 2, respectivamente na figura 43. O projeto da sondagem 2302 é aquele em que o primeiro nucleotídeo adjacente próximo aos domínios de hibridização contém um nucleotídeo diferente do Alelo 2 para o Alelo 1. Em outras palavras, o primeiro nucleotídeo adjacente ao domínio de hibridização pode ser um polimorfismo de nucleotídeo único ou SNP. Neste exemplo, o primeiro nucleotídeo adjacente em 2304 próximo ao domínio de

hibridização de 2302 é um "A", enquanto o primeiro nucleotídeo adjacente em 2305 próximo ao domínio de hibridização de 2302 é um "T". Ou seja, os produtos de sondagem do Alelo 1 e Alelo 2 podem ser distinguidos entre si com base na identidade do primeiro nucleotídeo imediatamente adjacente ao domínio de hibridização.

[00220] Nesta representação, uma polimerase de DNA ou outra enzima será usada para adicionar pelo menos um nucleotídeo adicional a cada uma das sequências de sondagem. Neste exemplo, os substratos de nucleotídeo para a polimerase de DNA são competentes para uma adição única, talvez porque os nucleotídeos adicionados à mistura da reação são nucleotídeos de dideoxi. Ou seja, somente um novo nucleotídeo deve ser adicionado a cada sequência de sondagem. Neste exemplo, o nucleotídeo adicionado à sondagem 2302 para o Alelo 1 conterá um ou mais rótulos (2303) do tipo "A". O nucleotídeo adicionado à sondagem 2302 conterá um ou mais rótulos (2306) do tipo "B", de modo que os produtos de sondagem do Alelo 1 possam ser claramente distinguidos dos produtos de sondagem do Alelo 2. Ou seja, o produto de sondagem para o Alelo 1 consiste na sondagem 2302 mais um nucleotídeo adicional que possui um ou mais rótulos do tipo "A", e os produtos de sondagem para o Alelo 2 consistem na sondagem 2302 mais um nucleotídeo adicional portador de um ou mais rótulos do tipo "B".

[00221] Nesta representação, as sondagens 2302 contêm um ou mais rótulos (2301) do tipo "C". Portanto, os produtos de sondagem conterão uma combinação de rótulos. Para o Alelo 1, os produtos de sondagem conterão rótulos do tipo "A" e tipo "C", enquanto os produtos de sondagem do Alelo 2 conterão rótulos do tipo "B" e do tipo "C".

[00222] A figura 44 representa uma modificação do procedimento geral descrito na figura 21. A figura 44 representa dois

conjuntos de sondagem para a identificação de vários alelos do mesmo locus genômico. Por exemplo, para distinguir os alelos maternos e fetais, no caso de um DNA livre de célula isolado de uma mulher grávida ou para distinguir os alelos do hospedeiro e do doador, no caso de DNA livre de célula de um receptor de um transplante de órgão. A figura 44 representa dois conjuntos de sondagem - um conjunto de sondagem para o Alelo 1 e um conjunto de sondagem para o Alelo 2. 2405 e 2406 são moléculas alvo correspondentes ao Alelo 1 e Alelo 2, respectivamente.

[00223] Um primeiro conjunto de sondagem contém a sondagem de membro 2401. 2401 contém uma marca de afinidade (2400) que pode ser usada para isolamento e identificação do produto de sondagem. Nesta representação, os conjuntos de sondagem usados para identificação de dois alelos diferentes são iguais. Ou seja, o conjunto de sondagem para o Alelo 2 consiste na sondagem do membro 2401. Nesta representação, a sondagem 2401 é hibridizada em uma sequência correspondente no Alelo 1 e no Alelo 2, respectivamente na figura 44. O projeto da sondagem para 2401 é aquele em que o primeiro nucleotídeo adjacente próximo aos domínios de hibridização contém um nucleotídeo diferente do Alelo 2 para o Alelo 1. Em outras palavras, o primeiro nucleotídeo adjacente ao domínio de hibridização pode ser um polimorfismo de nucleotídeo único ou SNP. Neste exemplo, o primeiro nucleotídeo adjacente em 2405 próximo ao domínio de hibridização de 2401 é um "A", enquanto o primeiro nucleotídeo adjacente em 2406 próximo ao domínio de hibridização de 2401 é um "T". Ou seja, os produtos de sondagem do Alelo 1 e Alelo 2 podem ser distinguidos entre si com base na identidade do primeiro nucleotídeo imediatamente adjacente ao domínio de hibridização.

[00224] Nesta representação, uma polimerase de DNA ou outra enzima será usada para adicionar pelo menos um nucleotídeo

adicional a cada uma das sequências de sondagem. Neste exemplo, o nucleotídeo adicionado à sondagem 2401 para o Alelo 1 conterá um ou mais rótulos (2402) do tipo "A". O nucleotídeo adicionado à sondagem 2401 conterá um ou mais rótulos (2407) do tipo "B", de modo que os produtos de sondagem do Locus 1 possam ser claramente distinguidos dos produtos de sondagem do Locus 2. Ou seja, o produto de sondagem para o Alelo 1 contém a sondagem 2401 mais um nucleotídeo adicional que possui um ou mais rótulos do tipo "A" e o produto de sondagem para o Alelo 2 contém a sondagem 2401 mais um nucleotídeo portador de um ou mais rótulos do tipo "B". Um nucleotídeo diferente, e não aquele usado para distinguir o Alelo 1 do Alelo 2, deverá servir como um terminador de cadeia. Neste exemplo em particular, um nucleotídeo "A" em uma molécula alvo é usado para identificar o Alelo 1 e um nucleotídeo "T" é usado para identificar o Alelo 2. Neste exemplo, um nucleotídeo "C" pode servir como um terminador de cadeia. Neste caso, um nucleotídeo "C" será adicionado ao ensaio que não será capaz do alongamento da cadeia (por exemplo, um dideóxi C). Uma restrição adicional é que as sequências de sondagem são desenvolvidas de modo que não haja casos de uma identificação de nucleotídeo do Alelo 2 presentes em 2405 entre o nucleotídeo de distinção do Alelo 1 e o nucleotídeo de terminação da cadeia. Neste exemplo, não haverá nucleotídeos "T" presentes em 2405 após o domínio de hibridização de 2401 e antes de G, que será emparelhado com o terminador C da cadeia.

[00225] Nesta representação, a polimerase de DNA ou uma enzima similar será usada para sintetizar nova sequências de nucleotídeo, e o nucleotídeo adicionado no local do nucleotídeo distinto para o Alelo 1 conterá um ou mais rótulos (2402) do tipo "A". O nucleotídeo adicionado no local do nucleotídeo distinto para o Alelo 2 conterá 1 ou mais rótulos (2407) do tipo "B", de modo que os produtos de sondagem para o Alelo 1 possam

ser claramente distintos dos produtos de sondagem do Alelo 2. Nesta representação, o nucleotídeo adicionado na posição de terminação da cadeia conterá um ou mais rótulos (2403) do tipo "C". Portanto, os produtos de sondagem conterão uma combinação de rótulos. Para o Alelo 1, os produtos de sondagem conterão rótulos do tipo "A" e tipo "C", enquanto os produtos de sondagem do Alelo 2 conterão rótulos do tipo "B" e do tipo "C".

[00226] A figura 45 representa uma modificação do procedimento geral descrito na figura 21. A figura 45 representa dois conjuntos de sondagem para a identificação de vários alelos do mesmo locus genômico. Por exemplo, para distinguir os alelos maternos e fetais, no caso de um DNA livre de célula isolado de uma mulher grávida ou para distinguir os alelos do hospedeiro e do doador, no caso de DNA livre de célula de um receptor de um transplante de órgão. A figura 45 representa dois conjuntos de sondagem - um conjunto de sondagem para o Alelo 1 e um conjunto de sondagem para o Alelo 2. 2505 e 2506 são moléculas alvo correspondentes ao Alelo 1 e Alelo 2, respectivamente.

[00227] Um primeiro conjunto de sondagem contém a sondagem de membro 2501. 2501 contém uma marca de afinidade (2500) que pode ser usada para isolamento e identificação do produto de sondagem. Nesta representação, os conjuntos de sondagem usados para identificação de dois alelos diferentes são iguais. Ou seja, o conjunto de sondagem para o Alelo 2 consiste na sondagem do membro 2501. Nesta representação, a sondagem 2501 é hibridizada em uma sequência correspondente no Alelo 1 e no Alelo 2, respectivamente na figura 45. O projeto da sondagem para 2501 é aquele em que o primeiro nucleotídeo adjacente próximo aos domínios de hibridização contém um nucleotídeo diferente do Alelo 2 para o Alelo 1. Em outras palavras, o primeiro nucleotídeo adjacente ao domínio de hibridização pode ser um polimorfismo de nucleotídeo único ou SNP. Neste exemplo,

o primeiro nucleotídeo adjacente em 2505 próximo ao domínio de hibridização de 2501 é um "A", enquanto o primeiro nucleotídeo adjacente em 2506 próximo ao domínio de hibridização de 2501 é um "T". Ou seja, os produtos de sondagem do Alelo 1 e Alelo 2 podem ser distinguidos entre si com base na identidade da primeira base imediatamente adjacente ao domínio de hibridização.

[00228] Nesta representação, uma polimerase de DNA ou outra enzima será usada para adicionar pelo menos um nucleotídeo adicional a cada uma das sequências de sondagem. Neste exemplo, o nucleotídeo adicionado à sondagem 2501 para o Alelo 1 conterà um ou mais rótulos (2502) do tipo "A". O nucleotídeo adicionado à sondagem 2501 conterà um ou mais rótulos (2507) do tipo "B", de modo que os produtos de sondagem do Locus 1 possam ser claramente distinguidos dos produtos de sondagem do Locus 2. Ou seja, o produto de sondagem para o Alelo 1 contém a sondagem 2501 mais um nucleotídeo adicional que possui um ou mais rótulos do tipo "A" e o produto de sondagem para o Alelo 2 contém a sondagem 2501 mais um nucleotídeo portador de um ou mais rótulos do tipo "B". Um nucleotídeo diferente, e não aquele usado para distinguir o Alelo 1 do Alelo 2, deverá servir como um terminador de cadeia. Neste exemplo em particular, um nucleotídeo "A" em uma molécula alvo é usado para identificar o Alelo 1 e um nucleotídeo "T" é usado para identificar o Alelo 2. Neste exemplo, um nucleotídeo "C" pode servir como um terminador de cadeia. Neste caso, um nucleotídeo "C" será adicionado ao ensaio que não será capaz do alongamento da cadeia (por exemplo, um dideóxi C). Uma restrição adicional é que as sequências de sondagem são desenvolvidas de modo que nenhum caso de uma identificação de nucleotídeo do Alelo 2 esteja presente em 2505 entre o nucleotídeo de distinção do Alelo 1 e o nucleotídeo de terminação da cadeia. Neste exemplo, não haverá nucleotídeos "T"

presentes em 2505 após o domínio de hibridização de 2501 e antes de G, que será emparelhado com o terminador C da cadeia.

[00229] Nesta representação, a polimerase de DNA ou uma enzima similar será usada para sintetizar novas sequências de nucleotídeo, e o nucleotídeo adicionado no local do nucleotídeo distinto para o Alelo 1 conterá um ou mais rótulos (2502) do tipo "A". O nucleotídeo adicionado no local do nucleotídeo distinto para o Alelo 2 conterá 1 ou mais rótulos (2507) do tipo "B", de modo que os produtos de sondagem para o Alelo 1 possam ser claramente distintos dos produtos de sondagem do Alelo 2. Nesta representação, um quarto nucleotídeo pode ser adicionado ao ensaio que contém um ou mais rótulos (2508, 2503) do tipo "C". Esse quarto nucleotídeo não é emparelhado com o nucleotídeo de identificação para o Alelo 1 (neste exemplo, A), não é emparelhado com o nucleotídeo de identificação para o Alelo 2 (neste exemplo, T), não é emparelhado com o nucleotídeo de terminação da cadeia (neste exemplo, G). Neste exemplo, o quarto nucleotídeo que teria um ou mais rótulos do tipo "C" é G e será emparelhado com os locais C em 2505 e 2506. Portanto, os produtos de sondagem conterão uma combinação de rótulos. Para o Alelo 1, os produtos de sondagem conterão rótulos do tipo "A" e tipo "C", enquanto os produtos de sondagem do Alelo 2 conterão rótulos do tipo "B" e do tipo "C".

[00230] A figura 46 representa uma modificação do procedimento geral descrito na figura 21. A figura 46 representa dois conjuntos de sondagem para a identificação de vários alelos do mesmo locus genômico. Por exemplo, para distinguir os alelos maternos e fetais, no caso de um DNA livre de célula isolado de uma mulher grávida ou para distinguir os alelos do hospedeiro e do doador, no caso de DNA livre de célula de um receptor de um transplante de órgão. A figura 46 representa dois conjuntos de sondagem - um conjunto de sondagem para o Alelo 1 e um conjunto

de sondagem para o Alelo 2. 2605 e 2606 são moléculas alvo correspondentes ao Alelo 1 e Alelo 2, respectivamente. Um primeiro conjunto de sondagem contém a sondagem de membro 2602. 2602 contém um rótulo (2601) do tipo "A". 2602 contém uma marca de afinidade (2600) que pode ser usada para isolamento e identificação do produto de sondagem.

[00231] Um segundo conjunto de sondagem com a sondagem de membro 2609 apresenta as respectivas características como no primeiro conjunto de sondagem. Entretanto, 2609 contém um rótulo (2608) do tipo "B", distinguível do tipo "A". 2609 contém uma marca de afinidade (2607) que pode ser idêntica a ou exclusiva de 2600.

[00232] Nesta representação, 2602 e 2609 contêm sequências que são quase idênticas e diferem em apenas um nucleotídeo na sequência. Portanto, as sequências de hibridização dessas duas sondagens são complementares para o Alelo 1 (2605) ou Alelo 2 (2606). Além disso, o comprimento de cada domínio de hibridização em 2602 e 2609, bem como as condições de hibridização experimentais são desenvolvidas de modo que a sondagem 2602 será hibridizada apenas no Alelo 1 e a sondagem 2609 só será hibridizada no Alelo 2. A finalidade desse tipo de ensaio é ser capaz de quantificar precisamente a frequência do Alelo 1 e do Alelo 2 em uma amostra.

[00233] Nesta representação, a polimerase de DNA ou outra enzima pode ser usada para sintetizar uma nova sequência de polinucleotídeo, por exemplo, 2604 no caso de Alelo 1 ou 2611 no caso de Alelo 2. Nesta representação, 2604 e 2611 podem conter um ou mais rótulos (2603, 2610) do tipo "C", possivelmente como um resultado de incorporação de um de mais nucleotídeos portadores de um rótulo do tipo "C". Portanto, os produtos de sondagem conterão uma combinação de rótulos. Para o Alelo 1, os produtos de sondagem conterão rótulos do tipo "A" e tipo "C",

enquanto os produtos de sondagem do Alelo 2 conterão rótulos do tipo "B" e do tipo "C". Esta representação resulta nos produtos de sondagem com especificidade elevada para sequências no Alelo 1 ou Alelo 2, respectivamente.

[00234] As figuras 55-58 ilustram uma modificação do procedimento geral descrito em referência às 21-46. A figura 55 representa dois conjuntos de sondagem; um conjunto de sondagem para o Locus 1 e um conjunto de sondagem para o Locus 2 - embora, conforme citado anteriormente, vários conjuntos de sondagem possam ser designados para cada locus genômico. O braço esquerdo do conjunto de sondagem do Locus 1 consiste em uma sequência de preparação de avanço, uma sequência de marca de afinidade e um homólogo à sequência de Locus 1. O braço direito do conjunto de sondagem do Locus 1 consiste em um homólogo da sequência do Locus 1 e uma sequência de preparação reversa para rotular o conjunto de sondagem do Locus 1 com um rótulo "A". O braço esquerdo do conjunto de sondagem do Locus 2 consiste em uma sequência de preparação de avanço, uma sequência de marca de afinidade e um homólogo à sequência do Locus 2. O braço direito do conjunto de sondagem do Locus 2 consiste em um homólogo da sequência do Locus 2 e uma sequência de preparação reversa para rotular o conjunto de sondagem do Locus 2 com um rótulo "B". A sequência de preparação de avanço e a sequência de marca de afinidade são idênticas para os conjuntos de sondagem do Locus 1 e do Locus 2. As sequências homólogas são específicas para um único locus genômico. As sequências homólogas do locus para cada conjunto de sondagem são imediatamente adjacentes entre si de modo que quando são hibridizadas em seus locus alvo, elas imediatamente se tocam e, dessa forma, podem ser ligadas para formar uma molécula contígua. A sequência de preparação reversa é específica para o rótulo (por ex., rótulo A ou rótulo B) a ser

usado na rotulação dos produtos de sondagem para um locus em particular de uma sequência de marca de afinidade específica.

[00235] A figura 56 representa um fluxo de trabalho prático que seria aplicado ao grupo de conjuntos de sondagem, de modo como os conjuntos de sondagem ilustrados na Figura 55. Esta representação é baseada em um conjunto de sondagem para um locus genômico (por ex., o conjunto de sondagem para o Locus 1 mostrado na figura 55). Na etapa 1, o grupo de conjuntos de sondagem é misturado com o DNA livre de célula purificado. Na etapa 2, as sequências específicas do locus em cada conjunto de sondagem são hibridizadas em suas sequências homólogas correspondentes na amostra de DNA livre de célula. Na Etapa 3, uma enzima ligase é adicionada para catalisar a formação de uma ligação de fosfodiéster entre a base 3' no homólogo do braço esquerdo e o braço 5' do homólogo direito, fechando a incisão entre os dois braços e, assim, formando uma molécula contínua que é o produto de sondagem. Na etapa 4, os primers modificados e os componentes de reação PCR (polimerase Taq, dNTPs e tampão de reação) são adicionados para amplificar o produto de sondagem ligado. O Primer Avançado é modificado, desde que ele tenha um grupo de fosfato 5' que o torne um modelo preferencial para a exonuclease Lambda usada na etapa 6 e o Primer Reverso é modificado desde que contenha o rótulo (círculo azul) que é específico para produtos de sondagem de um locus particular para uma marca de afinidade em particular. Na etapa 5, o produto de sondagem é PCR amplificado para produzir um produto de PCR de filamento duplo no qual o filamento avançado contém um grupo de fosfato 5' e o filamento reverso contenha um rótulo 5'. Na etapa 6, a exonuclease Lambda é adicionada para digerir o filamento avançado em uma direção 5' para 3' - o grupo de fosfato 5' no filamento avançado o torna um modelo preferido para a digestão de exonuclease Lambda. O material resultante possui filamento

único (somente filamento reverso) com um rótulo 5'. Isso representa o material alvo rotulado para hibridização em um microarranjo ou monocamada.

[00236] A figura 57 representa uma versão modificada do fluxo de trabalho prático ilustrado na figura 56. Nesta representação, o braço esquerdo de cada conjunto de sondagem contém uma molécula de biotina de terminal conforme indicado por um "B" nas etapas 1 a 6 da figura. Essa biotinilação permite a purificação do grupo de produtos de sondagem após a conclusão da reação de hibridização-ligação e antes da amplificação de PCR. O fluxo de trabalho desta representação é idêntico ao descrito na Figura 57 para as Etapas 1 a 3. Na Etapa 4, as esferas revestidas com estreptavidina são adicionadas à reação da hibridização-ligação. A molécula da biotina contida nos produtos de sondagem ligará os produtos à estreptavidina. Na Etapa 5, as esferas magnéticas são lavadas para remover o DNA não biotinilado (DNA genômico livre de célula e oligonucleotídeo do braço direito), resultando em um produto de sondagem purificado. As etapas 6 a 9 são realizadas da mesma maneira descrita para as etapas 4 a 7 na figura 56.

[00237] A figura 58 fornece um exemplo de como os produtos de sondagem para Locus 1 e Locus 2 podem ser rotulados com diferentes moléculas de rótulo. Na Figura 58A, os produtos de sondagem do Locus 1 são rotulados com o rótulo A (verde), e os produtos de sondagem do Locus 2 são rotulados com o rótulo B (vermelho) em uma reação de amplificação de PCR. Os produtos de sondagem para os dois locus contêm a sequência A de marca de afinidade. Na figura 58B, a mistura de produtos de sondagem rotulada de maneira diferente é hibridizada em um local de microarranjo no qual a sequência da sondagem de captura é complementar à sequência A da marca de afinidade. Nas figuras 58C, a imagem do local do microarranjo é gerada, e o número de moléculas do rótulo A e rótulo B é contado para fornecer uma

medida relativa dos níveis de Locus 1 e Locus 2 presentes na amostra.

[00238] A figura 59 fornece evidência de que os produtos que representam vários locais genômicos para um locus podem ser gerados em uma maneira específica de enzima ligase que usa o processo de hibridização-ligação. Oito conjuntos de sondagem, cada um consistindo em um componente de braço esquerdo e braço direito conforme descrito na figura 55 e contendo homólogos a oito locais de cromossomo 18 foram hibridizados em modelos de oligonucleotídeo sintético (cerca de 48 nucleotídeos) e ligados usando uma enzima ligase para unir os braços esquerdo e direito para cada. Os produtos da reação foram analisados com o uso de desnaturação de eletroforese de gel de poliacrilamida. A pista gel 1 contém uma escada de peso molecular para indicar tamanhos de faixa de DNA. As pistas 2 a 9 contêm produtos de reação de hibridização-ligação para os oito conjuntos de sondagem de cromossomo 18. Uma faixa de DNA de aproximadamente 100 nucleotídeos, representando o produto de sondagem de aproximadamente 60 nucleotídeos no braço esquerdo e aproximadamente 40 nucleotídeos no braço direito está presente em cada uma das faixas 2 a 9. As pistas 10 e 11 contêm reações de controle negativas às quais nenhuma enzima ligase foi adicionada. Nenhuma faixa de DNA de aproximadamente 100 nucleotídeos está presente nas pistas 10 e 11.

[00239] A figura 60 fornece dados que indicam que os conjuntos de sondagem podem ser usados para detectar mudanças relativas no estado de número de cópia. Uma mistura de oito conjuntos de sondagem contendo homólogos a oito locais distintos do cromossomo X foi usada para analisar as linhagens de célula contendo diferentes números de cromossomo X indicados na Tabela 1.

**Tabela 1:** Linhagens de célula contendo diferentes números de cópia do cromossomo X

<b>ID de linhagem de célula Coriell</b>	<b>Número de cópias do cromossomo X</b>
NA12138	1
NA13783	2
NA00254	3
NA01416	4
NA06061	5

[00240] O PCR quantitativo foi usado para determinar a quantidade do produto de sondagem presente para cada linhagem de célula após os processos de hibridização-ligação e purificação descritos na figura 57 (etapas 1 a 5). Conforme ilustrado pela figura 60A, o estado do número da cópia medido para várias linhas de célula seguido pela tendência esperada indicada na Tabela 1. Por exemplo, qPCR indicou um estado de número de cópia de menos de dois para NA12138, que tem uma cópia do cromossomo X. O estado de número de cópia medido para NA00254 (três cópias de X) era maior que dois, para NA01416 (quatro cópias de X) era maior que três e para NA06061 (cinco cópias de X) era maior que quatro. A receptividade do processo na detecção de diferenças no estado do número de cópia também é ilustrada pela Figura 60B na qual o estado do número de cópia é representado novamente no estado do número de cópia teórico.

[00241] A figura 61 fornece evidência de que misturas de produtos de sondagem podem ser usados para gerar dados de microconjunto quantitativo conforme descrito nas figuras 56 e 57.

[00242] A figura 61A apresenta imagens de fluorescência representativas de dois pontos de arranjo em dois canais de

imagens ortogonais (Alexa 488: verde, Alexa 594; vermelho). Uma região de interesse (ROI) é automaticamente selecionada (círculo grande), com qualquer contaminante brilhante indesejado sendo mascarado na imagem (regiões esboçadas menores na ROI). Fluoróforos únicos em produtos de ensaio hibridizados únicos são visualizados como características pontuadas pequenas dentro do ponto do arranjo. (i) Um ponto "balanceado" (representando entradas de alvos genômicos em uma proporção de concentração 1:1 no ensaio) com imagem no canal verde e (ii) o mesmo ponto com imagem no canal vermelho. (i) Um ponto "aumentado" (representando entradas de alvos genômicos em uma proporção de concentração > 1:1 no ensaio) com imagem no canal verde e (iv) o mesmo ponto com imagem no canal vermelho.

[00243] A figura 61B apresenta contagens brutas dos fluoróforos detectados em dois canais para cinco pontos cada das condições "balanceadas" e "aumentadas". Apesar de alguma variação no número absoluto de flúor, os números nos dois canais estão muito próximos no caso "balanceado", mas demonstram separação clara no caso "aumentado".

[00244] A figura 61C apresenta os valores de proporção calculados para vários flúor no canal verde dividido pelo número de flúor no canal vermelho, para os cinco pontos de cada uma das condições "balanceadas" e "aumentadas". O caso "balanceado" é centralizado em uma proporção aproximada de 1.0 e o caso "aumentado" está em uma proporção elevada. Considerando o caso "balanceado" ao comparar dois locus genômicos balanceados e o caso "aumentado" como um onde um locus é aumentado em relação ao outro, podemos calcular a confiança das duas condições usando um teste T independente de 2 grupos, produzindo um valor p de  $8 \times 10^{-14}$ .

[00245] A figura 62 ilustra uma modificação do procedimento geral descrito na figura 55 a 58. Nesta representação, um

segundo conjunto de sondagem, Conjunto de sondagem B é projetado para cada local genômico de modo que as sequências homólogas do genoma no Conjunto de sondagem B sejam um complemento reverso das sequências homólogas do genoma no Conjunto de sondagem A. O Conjunto de sondagem A será hibridizado no filamento reverso do DNA genômico e o Conjunto de sondagem B será hibridizado no filamento avançado do DNA genômico. Esta representação fornecerá sensibilidade aumentada em relação à representação descrita nas figuras 55 a 58 pois produzirá aproximadamente o dobro do número de produtos de sondagem por locus.

[00246] A figura 63 ilustra uma modificação no procedimento geral descrito na figura 57. Nesta representação, o Primer Reverso usado na etapa 6 é modificado adicionalmente porque as quatro ligações que vinculam os primeiro cinco nucleotídeos na sequência de oligonucleotídeo são ligações de fosforotioato. Essa modificação resultará em todos os produtos de PCR gerados durante a amplificação de PCR (Etapa 7) tendo a modificação de fosforotioato no final 5'. Essa modificação protegerá o filamento reverso de qualquer digestão que possa ocorrer durante o tratamento com a exonuclease Lambda na etapa 8.

[00247] Embora o grupo de fosfato 5' no filamento avançado torne-o um modelo preferencial para a digestão de exonuclease Lambda, o filamento reverso pode ainda ter alguma vulnerabilidade à digestão. A modificação de fosforotioato do final 5' do filamento reverso reduzirá sua vulnerabilidade à digestão de exonuclease Lambda.

[00248] A figura 64 ilustra uma modificação do procedimento geral descrito na figura 55 a 58. Nesta representação, a amplificação de PCR do produto de sondagem é substituída pela amplificação linear adicionando o Primer Reverso, mas não o Primer Avançado à reação da amplificação na etapa 6. Se apenas o Primer Reverso estiver presente, o produto de amplificação será

de filamento único - o filamento reverso com um rótulo do final 5'. Como o produto de amplificação já é de filamento único, ele não requer processamento adicional antes da hibridização em um microarranjo, ou seja, a digestão de exonuclease Lambda pode ser omitida. Como um primer avançado não é usado nesta representação, é desnecessário para o braço esquerdo do conjunto de sondagem conter uma sequência de preparação de avanço. O braço esquerdo consistiria em uma sequência de marca de afinidade e uma sequência homóloga de locus apenas como ilustrado na figura 64.

[00249] Uma representação adicional do procedimento geral descrito nas figuras 55 a 58 é aquela na qual o processo de reação de ligação única na etapa 3 é substituído pelo processo de reação de ligação em ciclo. Isso é realizado pela substituição da enzima ligase termolábil (por ex., ligase T4) usada para catalisar a reação de ligação com uma ligase termoestável (por ex., ligase *Taq*). Quando a ligase termoestável é usada, a reação de hibridização-ligação pode ser aquecida a uma temperatura que fundirá todos os duplexes de DNA (por ex., 95 °C) após o ciclo inicial de hibridização e ligação ter ocorrido. Isso tornará o DNA de modelo genômico totalmente disponível para hibridização e ligação de outro conjunto de sondagem. A redução subsequente da temperatura (por ex., a 45 °C) permitirá que esse próximo evento de hibridização e ligação ocorra. Cada termociclo da reação de hibridização e ligação entre uma temperatura que fundirá os duplexes de DNA e uma que permitirá que a hibridização e a ligação ocorra aumentará linearmente a quantidade do produto de sondagem produzido a partir da reação. Se a reação for exposta a 30 ciclos, até 30 vezes, a quantidade do produto de sondagem será produzida com base em um processo no qual uma reação de ligação única é usada.

[00250] A figura 65 representa uma outra representação do procedimento modificado descrito na figura 62. Essa representação aproveita a vantagem da reação em cadeia da ligase (LCR) na combinação da presença do complemento reverso para cada conjunto de sondagem com o uso da ligase termoestável para permitir uma reação de ligação em ciclo na qual o produto é exponencialmente amplificado. A figura 65 representa dois conjuntos de sondagem, um Conjunto de sondagem A e Conjunto de sondagem B para um locus; onde as sequências homólogas do genoma no Conjunto de sondagem B são o complemento reverso das sequências homólogas do genoma no Conjunto de sondagem A. O braço 5' de cada Conjunto de sondagem consiste em uma sequência de marcação de afinidade e um homólogo enquanto o braço 3' de cada Conjunto de sondagem consiste em uma sequência homóloga com um rótulo anexo. No primeiro ciclo de uma reação em termociclo, o DNA genômico será o único modelo disponível para permitir que a hibridização e a ligação ocorram para gerar um produto de sondagem como ilustrado na figura 65A. Entretanto, no segundo ciclo, o Produto de sondagem B gerado no primeiro ciclo atuará como um modelo adicional para o Conjunto de sondagem A e, do mesmo modo, o Produto de sondagem A gerado no primeiro ciclo atuará como um modelo adicional para o Conjunto de sondagem B como ilustrado na Figura 65B. Dessa mesma maneira, os produtos de sondagem de cada ciclo sucessivo atuarão como modelo para a hibridização e ligação do conjunto de sondagem no próximo ciclo. Esse processo eliminaria a necessidade da amplificação de PCR do produto de sondagem que pode ser diretamente usada como alvo de microarranjo.

[00251] Outra representação do procedimento representado na Figura 65 é aquela que emprega o LCR, mas usa os conjuntos de sondagem cuja estrutura é descrita na figura 55, isto é, os braços esquerdo e direito são flanqueados pelas sequências de

preparação, o braço esquerdo contém uma molécula de biotina e o braço direito não contém um rótulo. Após a conclusão de LCR, os produtos de sondagem são purificados com o uso de esferas magnéticas (opcional) e depois o PCR amplificado e o alvo de microarranjo são preparados conforme ilustrado nas figuras 56 e 57.

[00252] A figura 66 representa ainda outra representação do procedimento representado na figura 65. O braço 5' de cada Conjunto de sondagem consiste em uma sequência de marca de afinidade e um homólogo enquanto o braço 3' de cada Conjunto de sondagem consiste em uma sequência homóloga e em uma sequência de preparação sem um rótulo anexado conforme ilustrado na figura 66A. Após a conclusão do LCR, o produto de sondagem pode ser purificado. O produto do LCR seria amplificado de maneira linear pela adição de um único primer que tem um rótulo anexado, juntamente com os componentes da reação (polimerase Taq, dNTPs e tampão da reação) conforme ilustrado na figura 66B. O produto desta amplificação seria de filamento único (apenas filamento reverso) com um rótulo 5' conforme ilustrado na figura 66C. Conseqüentemente, não seria necessário tratá-lo com a exonuclease Lambda, mas em vez disso poderia ser diretamente usado como alvo de microarranjo.

[00253] Em outro aspecto, a variação genética determinada pelos métodos descritos neste documento indica a presença ou ausência de câncer, variabilidade farmacocinética, toxicidade da droga, rejeição do transplante ou aneuploidia no paciente. Em outro aspecto, a variação genética determinada indica a presença ou ausência de câncer. Como consequência, os métodos descritos neste documento podem ser executados para diagnosticar câncer.

[00254] Um desafio significativo em oncologia é a descoberta precoce do câncer. Isso é particularmente verdadeiro em cânceres cuja geração de imagens ou biópsia é difícil (por ex., câncer

pancreático, câncer do pulmão). O DNA do tumor livre de célula (cfDNA de tumor) no sangue de um paciente oferece um método não invasivo para detectar um tumor. Podem ser tumores sólidos, tumores benignos, tumores líquidos, metástase ou outros crescimentos somáticos. A detecção pode estar em qualquer estágio no desenvolvimento do tumor, embora precoce de modo ideal (Estágio I ou Estágio II). A descoberta precoce permite intervenção (por ex., cirurgia, quimioterapia, tratamento farmacêutico) que pode estender a vida ou levar à remissão. Outros problemas na oncologia incluem o monitoramento da eficácia do tratamento, a titulação da dose de um agente terapêutico, a recorrência de um tumor no mesmo órgão como o tumor primário ou em locais distais e detecção de metástase. A invenção atual pode ser usada para todas essas aplicações.

[00255] Em algumas representações, os conjuntos de sondagem da presente divulgação podem ser configurados para ter como objetivo as variações genéticas conhecidas associadas aos tumores. Isso pode incluir mutações, SNPs, variantes de número de cópia (por ex., ampliações, exclusões), variantes neutras de cópia (por ex., inversões, translocações) e/ou combinações de complexo dessas variantes. Por exemplo, as variações de genética conhecidas associadas aos tumores incluem aquelas listadas em [cancer.sanger.ac.uk/cancergenome/projects/cosmic](http://cancer.sanger.ac.uk/cancergenome/projects/cosmic); [nature.com/ng/journal/v45/n10/full/ng.2760.html#supplementary-information](http://nature.com/ng/journal/v45/n10/full/ng.2760.html#supplementary-information); e nas Tabelas 2 e 3 abaixo: <sup>B</sup>GENE = valor p do corrigido para FDR no pico; oncogene amplificado frequentemente <sup>K</sup>Conhecido ou TSG excluído; gene de câncer <sup>P</sup>Putativo, regulador <sup>E</sup>Epigenético; gene associado ao <sup>M</sup>Mitocôndrias; \*\*Imediatamente adjacente à região de pico; <sup>T</sup>Adjacente ao telômero ou centrômero de cromossomo acrocêntrico.

Tabela 2: Variações genéticas de exemplo associada a tumores (Amplificação do gene)

Nome de pico	Classificação	Localização genômica	Região de pico	GISTIC valor q	Contagem de genes	Alvo(s)	Genes que sofreram mutação com frequência <sup>B</sup>
CCND1	1	11q13.3	chr11:69464719-69502928	2.05E-278	2	<i>CCND1</i> <sup>K</sup>	<i>CCND1</i> = 6.6e-08
EGFR	2	7p11.2	chr7:55075808-55093954	2.30E-240	1	<i>EGFR</i> <sup>K</sup>	<i>EGFR</i> = 2.2e-15
MYC	3	8q24.21	chr8:128739772-128762863	6.50E-180	1	<i>MYC</i> <sup>K</sup>	
TERC	4	3q26.2	chr3:169389459-169490555	5.40E-117	2	<i>TERC</i> <sup>P</sup>	
ERBB2	5	17q12	chr17:37848534-37877201	1.59E-107	1	<i>ERBB2</i> <sup>K</sup>	<i>ERBB2</i> = 1.3e-06
CCNE1	6	19q12	chr19:30306758-30316875	4.77E-90	1	<i>CCNE1</i> <sup>K</sup>	
MCL1	7	1q21.3	chr1:150496857-150678056	1.25E-80	6	<i>MCL1</i> <sup>K</sup>	
MDM2	8	12q15	chr12:69183279-69260755	2.59E-62	2	<i>MDM2</i> <sup>K</sup>	
INTS4	9	11q1	chr11:77610	1.01E-	1	<i>INTS4</i>	

4		4.1	143- 77641464	54			
WHSC 1L1	10	8p11 .23	chr8:381918 04-38260814	3.43E- 46	2	<i>WHSC1L1</i> <sup>E</sup> , <i>LETM2</i> <sup>M</sup>	
CDK4	11	12q1 4.1	chr12:58135 797- 58156509	5.14E- 41	5	<i>CDK4</i> <sup>K</sup>	<i>CDK4</i> = 0.0048
KAT6 A	12	8p11 .21	chr8:417513 00-41897859	2.97E- 39	2	<i>KAT6A</i> <sup>P/E</sup> , <i>IKBKB</i> <sup>**</sup>	
SOX2	13	3q26 .33	chr3:181151 312- 181928394	1.21E- 38	2	<i>SOX2</i> <sup>K</sup>	
PDGF RA	14	4q12	chr4:549247 94-55218386	1.08E- 37	3	<i>PDGFRA</i> <sup>K</sup>	
BDH1	15	3q29	chr3:197212 101- 197335320	1.21E- 31	1	<i>BDH1</i> <sup>M</sup>	
1q44	16	1q44 <sup>T</sup>	chr1:242979 907- 249250621	4.48E- 31	83	<i>SMYD3</i> <sup>E</sup>	
MDM4	17	1q32 .1	chr1:204367 383- 204548517	1.98E- 29	3	<i>MDM4</i> <sup>K</sup>	
TERT	18	5p15 .33	chr5:128770 4-1300024	9.34E- 27	1	<i>TERT</i> <sup>K</sup>	
KDM5 A	19	12p1 3.33 <sup>T</sup>	chr12:1- 980639	1.59E- 25	11	<i>KDM5A</i> <sup>E</sup>	
MYCL 1	20	1p34 .2	chr1:403179 71-40417342	3.99E- 25	2	<i>MYCL1</i> <sup>K</sup>	
IGF1	21	15q2	chr15:98667	8.62E-	9	<i>IGF1R</i> <sup>K</sup>	

R		6.3	475- 100292401	25			
PARP 10	22	8q24 .3	chr8:144925 436- 145219779	5.44E- 20	15	<i>PARP10<sup>P,E</sup></i> <i>, CYC1<sup>M</sup></i>	
G6PD	23	Xq28	chrX:153760 870- 153767853	3.66E- 19	1	<i>G6PD</i>	
PHF1 2	24	17q1 1.2	chr17:27032 828- 27327946	1.75E- 16	21	<i>PHF12<sup>E</sup></i> <i>ERAL1<sup>M</sup></i>	
20q1 3.33	25	20q1 3.33	chr20:62187 847- 62214354	2.96E- 16	2		
PAF1	26	19q1 3.2	chr19:39699 366- 39945515	1.66E- 15	13	<i>PAF1<sup>P,E</sup></i>	<i>IL28A=0</i> <i>,021,</i> <i>SUPT5H=</i> <i>0,084</i>
BCL2 L1	27	20q1 1.21	chr20:30179 028- 30320705	2.85E- 15	4	<i>BCL2L1<sup>K</sup></i>	
TUBD 1	28	17q2 3.1	chr17:57922 443- 57946458	7.19E- 15	1	<i>TUBD1</i>	<i>TUBD1 =</i> <i>0.009</i>
[ZNF 703]	29	8p11 .23	chr8:374926 69-37527108	2.44E- 14	0		
1q23 .3	30	1q23 .3	chr1:160949 115- 161115281	7.73E- 13	9		
8q22 .2	31	8q22 .2	chr8:101324 079-	4.22E- 11	3		<i>SNX31 =</i> <i>0,015</i>

			101652657				
BRD4	32	19p1 3.12	chr19:15310 246- 15428182	5.04E- 10	3	<i>NOTCH3<sup>P</sup></i> , <i>BRD4<sup>P,E</sup></i>	
KRAS	33	12p1 2.1	chr12:24880 663- 25722878	9.47E- 10	7	<i>KRAS<sup>K</sup></i>	<i>KRAS</i> = 1.5e-14
NKX2 -1	34	14q1 3.2	chr14:35587 755- 37523513	1.33E- 09	14	<i>NKX2-1<sup>K</sup></i>	<i>NFKBIA</i> = 0,0098, <i>RALGAPA</i> 1=0,027
NFE2 L2	35	2q31 .2	chr2:178072 322- 178171101	5.48E- 09	5	<i>NFE2L2</i>	<i>NFE2L2</i> = 3.9e- 14
ZNF2 17	36	20q1 3.2	chr20:52148 496- 52442225	5.83E- 08	1	<i>ZNF217<sup>K</sup></i>	<i>ZNF217</i> = 0,0082
13q3 4	37	13q3 4 <sup>T</sup>	chr13:10881 8892- 115169878	6.28E- 08	45	<i>ING1<sup>E</sup></i>	<i>ING1</i> = 0,00026
KAT6 B	38	10q2 2.2	chr10:76497 097- 77194071	1.41E- 07	9	<i>KAT6B<sup>E</sup></i> , <i>VDAC2<sup>M</sup></i>	
NSD1	39	5q35 .3	chr5:176337 344- 177040112	1.75E- 06	22	<i>NSD1<sup>E</sup></i> , <i>PRELID1<sup>M</sup></i>	<i>NSD1</i> = 4.9e-10
FGFR 3	40	4p16 .3	chr4:177879 7-1817427	2.14E- 06	2	<i>FGFR3<sup>E</sup></i> , <i>LETM1<sup>M</sup></i>	<i>FGFR3</i> = 0,00018
9p13 .3	41	9p13 .3	chr9:356523 85-35739486	2.55E- 06	8		
COX1	42	4q13	chr4:735302	2.68E-	7	<i>COX18<sup>M</sup></i>	

8		.3	10-74658151	06			
7q36 .3	43	7q36 .3 <sup>T</sup>	chr7:153768 037- 159138663	3.19E- 06	30	<i>PTPRN2</i> <sup>L</sup> , <i>DPP6</i> <sup>L</sup>	
18q1 1.2	44	18q1 1.2	chr18:23857 484- 24119078	3.83E- 06	2		
SOX1 7	45	8q11 .23	chr8:550697 81-55384342	2.02E- 05	1	<i>SOX17</i>	<i>SOX17</i> = 0,00092
11q2 2.2	46	11q2 2.2	chr11:10229 5593- 102512085	0,00015 337	3		
CBX8	47	17q2 5.3	chr17:77770 110- 77795534	0,00023 029	1	<i>CBX8</i> <sup>E</sup>	
AKT1	48	14q3 2.33	chr14:10518 2581- 105333748	0,00028 451	7	<i>AKT1</i> <sup>K</sup>	<i>AKT1</i> = 1.1e-14
CDK6	49	7q21 .2	chr7:921960 92-92530348	0,00069 831	3	<i>CDK6</i> <sup>K</sup>	
6p21 .1	50	6p21 .1	chr6:415199 30-44297771	0,00104 59	70		
EHF	51	11p1 3	chr11:34574 296- 34857324	0,00110 02	1	<i>EHF</i>	
6q21	52	6q21	chr6:107098 934- 107359899	0,00118 06	4		
19q1 3.42	53	19q1 3.42 <sup>T</sup>	chr19:55524 376- 59128983	0,00133 19	13 8	<i>TRIM28</i> <sup>E</sup> , <i>SUV420H2</i> <sup>E</sup>	<i>ZNF471</i> = 5.4e- 05

17q2 1.33	54	17q2 1.33	chr17:47346 425- 47509605	0,00257 75	2		
BPTF	55	17q2 4.2	chr17:65678 858- 66288612	0,00283 75	11	<i>BPTF</i> <sup>E</sup>	
E2F3	56	6p22 .3	chr6:196107 94-22191922	0,00336 58	7	<i>E2F3</i> <sup>K</sup>	
19p1 3.2	57	19p1 3.2	chr19:10260 457- 10467501	0,00380 41	12	<i>MRPL4</i> <sup>M</sup>	<i>DNMT1</i> = 0,099
17q2 5.1	58	17q2 5.1	chr17:73568 926- 73594884	0,01233 7	2		
KDM2 A	59	11q1 3.2	chr11:67025 375- 67059633	0,01244 5	3	<i>KDM2A</i> <sup>E</sup>	
8q21 .13	60	8q21 .13	chr8:804325 52-81861219	0,02054 8	6	<i>MRPS28</i> <sup>M</sup>	
2p15	61	2p15	chr2:591432 37-63355557	0,02105 6	25		<i>XPO1</i> = 1.1e-05
14q1 1.2	62	14q1 1.2 <sup>T</sup>	chr14:1- 21645085	0,02780 3	57		
NEDD 9	63	6p24 .2	chr6:111804 26-11620845	0,08260 6	2	<i>NEDD9</i> <sup>K</sup>	
5p13 .1	64	5p13 .1	chr5:354596 50-50133375	0,09465 7	61		<i>SLC1A3</i> = 0,0021, <i>IL7R</i> =0, 0021
LINC 0053	65	8q23 .3	chr8:116891 361-	0,09529 4	1	<i>LINC0053</i> 6	

6			117360815				
10p1		10p1	chr10:41900				
5.1	66	5.1	59-6130004	0,10391	21		
22q1		22q1	chr22:18613				
1.21	67	1.21	558- 23816427	0,13213	10 5		
PHF3	68	6q12	chr6:638831 56-64483307	0,17851	4	<i>PHF3</i> <sup>E</sup> , <i>EYS</i> <sup>L</sup>	<i>PHF3</i> = 0,051
PAX8	69	2q13	chr2:113990 138- 114122826	0,19717	2	<i>PAX8</i> <sup>K</sup>	
9p24		9p24	chr9:1-			<i>SMARCA2</i> <sup>E</sup> , <i>KDM4C</i> <sup>E</sup> , <i>UHRF2</i> <sup>E</sup> , <i>KIAA2026</i> <sup>E</sup>	
.2	70	.2 <sup>T</sup>	7379570	0,20405	45		

Tabela 3: Variações genéticas de exemplo associada a tumores (Exclusão do gene)

Nome de pico	Classificação	Localização genômica	Região de pico	GISTIC valor q	Contagem de genes	Alvo(s)	Genes que sofreram mutação com frequência <sup>B</sup>
CDKN2A	1	9p21.3	chr9:21865498-22448737	0	4	<i>CDKN2A</i> <sup>K</sup>	<i>CDKN2A</i> = 4.4e-15
STK11	2	19p13.3	chr19:1103715-1272039	1.46E-238	7	<i>STK11</i> <sup>K</sup>	<i>STK11</i> = 2.5e-13
PDE4	3	5q11	chr5:58260298	2.02E-	3	<i>PDE4D</i> <sup>L</sup>	

D		.2	-59787985	143			
PARK 2	4	6q26	chr6:16169309 9-163153207	5.85E- 137	1	<i>PARK2</i> <sup>L</sup> , <sup>K</sup>	
LRP1 B	5	2q22 .1	chr2:13965561 7-143637838	4.25E- 107	1	<i>LRP1B</i> <sup>L</sup>	
CSMD 1	6	8p23 .2	chr8:2079140- 6262191	2.39E- 96	1	<i>CSMD1</i> <sup>L</sup>	
1p36 .23	7	1p36 .23	chr1:7829287- 8925111	1.23E- 93	8		
ARID 1A	8	1p36 .11	chr1:26900639 -27155421	5.74E- 87	2	<i>ARID1A</i> <sup>K</sup>	<i>ARID1A</i> = 1.5e- 14
PTEN	9	10q2 3.31	chr10:8961513 8-90034038	1.12E- 79	2	<i>PTEN</i> <sup>K</sup>	<i>PTEN</i> = 2.2e-15
WWOX	10	16q2 3.1	chr16:7812905 8-79627770	8.14E- 76	1	<i>WWOX</i> <sup>L</sup>	<i>WWOX</i> = 0,092
RB1	11	13q1 4.2	chr13:4883376 7-49064807	3.88E- 75	2	<i>RB1</i> <sup>K</sup>	<i>RB1</i> = 1.7e-13
FAM1 90A	12	4q22 .1	chr4:90844993 -93240505	9.26E- 75	1	<i>FAM190</i> <i>A</i> <sup>L</sup>	
2q37 .3	13	2q37 .3 <sup>T</sup>	chr2:24154452 7-243199373	1.77E- 70	29	<i>ING5</i> <sup>E</sup>	
22q1 3.32	14	22q1 3.32 <sup>T</sup>	chr22:4802691 0-51304566	8.20E- 65	45	<i>BRD1</i> <sup>E</sup> , <i>HDAC10</i> <sup>E</sup>	
11p1 5.5	15	11p1 5.5 <sup>T</sup>	chr11:1- 709860	1.02E- 62	34	<i>SIRT3</i> <sup>E</sup> , <i>PHRF1</i> <sup>E</sup>	<i>HRAS</i> = 7.8e-13
LINC 0029 0	16	4q34 .3	chr4:17891187 4-183060693	1.21E- 55	1	<i>LINC00</i> <i>290</i>	

FHIT	17	3p14 .2	chr3:59034763 -61547330	3.01E- 55	1	<i>FHIT</i> <sup>L</sup>	
RBFOX1	18	16p1 3.3	chr16:5144019 -7771745	1.00E- 45	1	<i>RBFOX1</i> <sup>L</sup>	
PTPRD	19	9p24 .1	chr9:8310705- 12693402	3.24E- 38	1	<i>PTPRD</i> <sup>L</sup>	
18q23	20	18q2 3 <sup>T</sup>	chr18:7497970 6-78077248	1.69E- 37	12		
FAT1	21	4q35 .2	chr4:18747587 5-188227950	6.81E- 36	1	<i>FAT1</i> <sup>K</sup>	<i>FAT1</i> = 2.4e-15
MPHOSPH8	22	13q1 2.11 <sup>T</sup>	chr13:1- 20535070	2.57E- 31	10	<i>MPHOSP</i> <i>H8</i> <sup>E</sup>	
15q15.1	23	15q1 5.1	chr15:4179590 1-42068054	2.71E- 29	4		<i>MGA</i> = 0,0083, <i>RPAP1</i> =0 ,035
11q25	24	11q2 5 <sup>T</sup>	chr11:1334002 80-135006516	4.93E- 26	14		
1p13.2	25	1p13 .2	chr1:11004852 8-117687124	1.69E- 25	10 0	<i>TRIM33</i> <sup>E</sup>	<i>NRAS</i> =1. 8e-13, <i>CD58</i> =0, 079
NF1	26	17q1 1.2	chr17:2932673 6-29722618	6.59E- 23	5	<i>NF1</i> <sup>K</sup>	<i>NF1</i> = 3.3e-13
MACROD2	27	20p1 2.1	chr20:1430287 6-16036135	9.00E- 19	3	<i>MACROD</i> <i>2</i> <sup>L</sup>	
7p22.3	28	7p22 .3 <sup>T</sup>	chr7:1- 1496620	1.04E- 17	18		
6p25.3	29	6p25 .3	chr6:1608837- 2252425	3.01E- 17	2		

21q1 1.2	30	21q1 1.2 <sup>T</sup>	chr21:1- 15482604	2.34E- 14	14		
9p13 .1	31	9p13 .1	chr9:38619152 -71152237	9.75E- 14	48		
ZNF1 32	32	19q1 3.43 <sup>T</sup>	chr19:5866158 2-59128983	3.77E- 13	24	<i>TRIM28</i> <sup>E</sup> , <i>ZNF132</i>	
5q15	33	5q15	chr5:73236070 -114508587	8.15E- 13	15 6	<i>APC</i> <sup>K</sup> , <i>CHD1</i> <sup>E</sup>	<i>APC</i> =2.6 e-13, <i>RASA1</i> =0 ,0029
MLL3	34	7q36 .1	chr7:15181741 5-152136074	9.26E- 13	1	<i>MLL3</i> <sup>K,E</sup>	<i>MLL3</i> = 1.1e-05
19q1 3.32	35	19q1 3.32	chr19:4733268 6-47763284	2.38E- 12	10		
15q1 2	36	15q1 2 <sup>T</sup>	chr15:1- 32929863	3.40E- 11	15 5		<i>OTUD7A</i> = 0,027
12q2 4.33	37	12q2 4.33 <sup>T</sup>	chr12:1316929 56-133851895	1.24E- 10	27		<i>POLE</i> =3. 9e-05, <i>PGAM5</i> =0 ,038
10q2 6.3	38	10q2 6.3 <sup>T</sup>	chr10:1351902 63-135534747	2.09E- 10	14		
6q21	39	6q21	chr6:86319089 -117076132	4.56E- 10	14 1	<i>PRDM1</i> <sup>E</sup> , <i>HDAC2</i> <sup>E</sup> , <i>PRDM13</i> <sup>E</sup>	<i>PRDM1</i> = 0,00054
PPP2 R2A	40	8p21 .2	chr8:25896447 -26250295	1.78E- 09	1	<i>PPP2R2</i> <i>A</i>	

IKZF 2	41	2q34	chr2:21154263 7-214143899	3.24E- 09	4	<i>IKZF2</i> <sup>K</sup> , <i>ERBB4</i> <sup>L</sup>	<i>ERBB4</i> = 0,00058
CNTN 4	42	3p26 .3 <sup>T</sup>	chr3:1- 3100786	6.44E- 09	3	<i>CNTN4</i> <sup>L</sup>	
3p12 .2	43	3p12 .2	chr3:75363575 -86988125	1.22E- 07	12	<i>ROBO1</i> <sup>L</sup> , <i>CADM2</i> <sup>L</sup>	
RAD5 1B	44	14q2 4.1	chr14:6827537 5-69288431	1.38E- 07	2	<i>RAD51B</i> <sup>L</sup>	<i>ZFP36L1</i> = 0,0016
11q2 3.1	45	11q2 3.1	chr11:1058491 58-117024891	5.31E- 07	84	<i>ATM</i> <sup>K</sup>	<i>ATM</i> =1.4 e-06, <i>POU2AF1</i> =0,082
IMMP 2L	46	7q31 .1	chr7:10959946 8-111366370	5.74E- 07	2	<i>IMMP2L</i> <sup>L</sup>	
NEGR 1	47	1p31 .1	chr1:71699756 -74522473	7.25E- 07	2	<i>NEGR1</i> <sup>L</sup>	
BRCA 1	48	17q2 1.31	chr17:4117876 5-41336147	7.25E- 07	2	<i>BRCA1</i> <sup>K</sup>	<i>BRCA1</i> = 3.5e-08
9q34 .3	49	9q34 .3	chr9:13544181 0-139646221	8.73E- 06	94	<i>NOTCH1</i> <sup>K</sup> , <i>BRD3</i> <sup>E</sup> , <i>GTF3C4</i> <sup>E</sup>	<i>NOTCH1</i> = 1e-08, <i>RXRA</i> =2. 1e-05, <i>COL5A1</i> = 0,0022, <i>TSC1</i> =0, 012
ANKS 1B	50	12q2 3.1	chr12:9912400 1-100431272	8.73E- 06	2	<i>ANKS1B</i> <sup>L</sup>	

DMD	51	Xp21 .2	chrX:30865118 -34644819	5.15E- 05	4	<i>DMD</i> <sup>L</sup>	
ZMYN D11	52	10p1 5.3 <sup>T</sup>	chr10:1- 857150	7.12E- 05	4	<i>ZMYND1</i> <i>1</i> <sup>E</sup>	
PRKG 1	53	10q1 1.23	chr10:5264408 5-54061437	9.79E- 05	3	<i>PRKG1</i> <sup>L</sup>	
FOXK 2	54	17q2 5.3	chr17:8044343 2-80574531	0,00019 271	1	<i>FOXK2</i>	
AGBL 4	55	1p33	chr1:48935280 -50514967	0,00021 9	2	<i>AGBL4</i> <sup>L</sup>	
CDKN 1B	56	12p1 3.1	chr12:1271099 0-12966966	0,00035 777	5	<i>CDKN1B</i> <sup>K</sup>	<i>CDKN1B</i> = 2.2e- 06
14q3 2.33	57	14q3 2.33 <sup>T</sup>	chr14:9438142 9-107349540	0,00074 358	22 7	<i>SETD3</i> <sup>E</sup> , <i>TDRD9</i> <sup>E</sup>	<i>AKT1</i> =2. 1e-13, <i>TRAF3</i> =9 .7e-05
14q1 1.2	58	14q1 1.2 <sup>T</sup>	chr14:1- 30047530	0,00101 81	16 2	<i>PRMT5</i> <sup>E</sup> , <i>CHD8</i> <sup>E</sup>	<i>CHD8</i> = 0,034
2p25 .3	59	2p25 .3 <sup>T</sup>	chr2:1- 20072169	0,00111 37	86	<i>MYCN</i> <sup>K</sup>	<i>MYCN</i> =0, 068
5q35 .3	60	5q35 .3 <sup>T</sup>	chr5:15384047 3-180915260	0,00285 15	21 2	<i>NSD1</i> <sup>E</sup> , <i>ODZ2</i> <sup>L</sup>	<i>NPM1</i> =3. 5e-13, <i>NSD1</i> =1. 9e-09, <i>ZNF454</i> = 0.0019, <i>UBLCP1</i> = 0,03, <i>GABRB2</i> =

							0,07
PTTG 1IP	61	21q2 2.3	chr21:4623068 7-46306160	0,01222 7	1	<i>PTTG1I</i> <i>P</i>	
22q1 1.1	62	22q1 1.1 <sup>T</sup>	chr22:1- 17960585	0,02033 2	15		
SMAD 4	63	18q2 1.2	chr18:4847208 3-48920689	0,03686 6	3	<i>SMAD4<sup>K</sup></i>	<i>SMAD4</i> = 6.6e-15
17p1 3.3	64	17p1 3.3 <sup>T</sup>	chr17:1- 1180022	0,04081 4	16		
4p16 .3	65	4p16 .3 <sup>T</sup>	chr4:1- 1243876	0,05634 5	27		
9p21 .2	66	9p21 .2	chr9:27572512 -28982153	0,09174 2	3		
10q2 5.1	67	10q2 5.1	chr10:9934008 4-113910615	0,11879	13 7	<i>HPSE2<sup>L</sup></i> <i>SMNDC1</i> <sup>E</sup>	<i>SMC3</i> =0, 00031, <i>GSTO2</i> =0 ,086
SMYD 3	68	1q44	chr1:24528226 7-247110824	0,15417	8	<i>SMYD3<sup>E</sup></i>	
8p11 .21	69	8p11 .21	chr8:42883855 -47753079	0,17382	4		
Xp22 .33	70	Xp22 .33 <sup>T</sup>	chrX:1- 11137490	0,21462	52		<i>MXRA5</i> = 0,031

[00256] No método de diagnóstico do câncer de acordo com algumas representações, as inversões que ocorrem em locais conhecidos (Figura 67A) podem ser facilmente buscadas desenvolvendo-se as sondagens que pelo menos parcialmente se sobrepõem ao ponto de interrupção de um braço da sondagem. Uma primeira sondagem que liga a sequência "normal" tem como alvo o

material genômico não invertido (Figura 67B) e possui um primeiro tipo de rótulo. Uma segunda sondagem que liga o alvo "invertido" carrega um segundo tipo de rótulo (Figura 67C). Um braço de sondagem de direito comum liga a sequência nativa que não está suscetível à inversão, imediatamente adjacente às primeiras sondagens. O braço de sondagem direito ainda carrega uma marca de ativação comum que localiza os produtos de sondagem na mesma região de um substrato de imagens. Dessa maneira, os pares de sondagem podem hibridizar nos alvos genômicos, ligar e ter suas imagens geradas para produzir contagens relativas das duas espécies subjacentes.

[00257] Da mesma maneira, as translocações que têm pontos de interrupção também podem ser analisadas. A Figura 68A mostra dois elementos genéticos que estão em sua ordem nativa ou translocados. Os braços de sondagem que pelo menos parcialmente sobrepõem esses pontos de interrupção de translocação permitem a diferenciação entre os pedidos normais e transpostos do material genético. Conforme mostrado nas figuras 68B e 68C, com a escolha de rótulos exclusivos nos dois braços esquerdos, os produtos de sondagem ligados resultantes podem ser distinguidos e contados durante a geração de imagens.

[00258] Esses métodos para detectar mudanças neutras de cópia (por ex., inversões, translocação) também podem ser usados para detectar variantes de linha germinativa no câncer ou em outra doença ou condições.

[00259] As mutações ou SNPs também são implicadas em numerosos cânceres e são buscadas de maneira similar às que são investigadas na determinação da fração fetal na aplicação de diagnósticos de pré-natal. Em algumas representações mostradas nas figuras 69A e 69B, os braços de sondagem esquerdos são desenvolvidos para aproveitar a vantagem de um desequilíbrio energético causado por um ou mais SNPs não correspondidos. Isso

faz com que um braço de sondagem (1101, contendo um rótulo) se ligue mais favoravelmente do que um segundo braço de sondagem (1107, contendo um segundo tipo de rótulo). Ambos os projetos ligam-se ao mesmo braço de sondagem direito (1102) que carrega a marca de ativação universal.

[00260] O sangue de um determinado paciente pode ser sondado por um método ou um híbrido de mais de um método. Além disso, em alguns casos, pode ser valiosa a personalização de sondagens específicas para um paciente. Isso envolveria a caracterização das características do tumor (SNPs, translocações, inversões etc.) em uma amostra do tumor primário (por ex., uma biópsia) e a criação de um ou mais conjuntos de sondagem personalizados que são otimizados para detectar essas variações genéticas específicas do paciente no sangue do paciente, fornecendo um método de monitoramento de baixo custo não invasivo. Isso teria valor significativo no caso de reincidência, onde é ideal a detecção de recorrência de baixo nível de um tipo de tumor (idêntico ou relacionado ao tumor original) o mais cedo possível.

[00261] Para as vias comuns de progressão da doença, os painéis adicionais podem ser designados para antecipar e monitorar o avanço da doença. Por exemplo, se as mutações tendem a se acumular em uma determinada ordem, as sondagens podem ser designadas a monitorar o status atual e os "pontos de verificação" de progressão e orientar opções de terapia.

[00262] Descoberta precoce do câncer: Por exemplo, a translocação ALK foi associada ao câncer de pulmão. Uma sondagem designada a investigar a translocação ALK pode ser usada para detectar tumores desse tipo via amostra sanguínea. Isso seria extremamente vantajoso, pois o método padrão de descoberta de tumores no pulmão é feito via raio x do tórax, um procedimento

caro que pode ser prejudicial à saúde do paciente e, portanto, não é realizado como padrão.

[00263] Detecção de recorrência do tipo de tumor primário: Por exemplo, um tumor de mama HER2+ é removido por cirurgia, e o paciente está em remissão. Uma sondagem que tem como objetivo o gene HER2 pode ser usada para monitorar as amplificações do gene HER2 em um ou mais momentos. Se for detectado, o paciente poderá ter um segundo tumor HER2+ no local primário ou em outro lugar.

[00264] Detecção dos tipos de tumor não primário: Por exemplo, um tumor de mama HER2+ é removido por cirurgia, e o paciente está em remissão. Uma sondagem que tem como objetivo o gene EGFR pode ser usada para monitorar os tumores EGFR+. Se for detectado, o paciente poderá ter um segundo tumor EGFR+ no local primário ou em outro lugar.

[00265] Detecção de metástase: Por exemplo, o paciente tem um tumor de mama HER2+. Uma sondagem designada a investigar a translocação ALK pode ser usada para detectar tumores desse tipo via amostra sanguínea. Esse tumor pode não estar na mama e é mais provável que esteja no pulmão. Se for detectado, o paciente poderá ter um tumor metastático distal do órgão primário.

[00266] Determinação da heterogeneidade do tumor: Muitos tumores têm várias populações clonais caracterizados por diferentes variantes genéticas. Por exemplo, um tumor de mama pode ter uma população de células que possam ser HER2+ e outra população de células EGRF+. O uso das sondagens designadas a essas variantes permitiria a identificação dessa heterogeneidade genética subjacente.

[00267] Medição de carga do tumor: Em todos os exemplos acima, a quantidade de cfDNA de tumor pode ser medida e pode ser usada para determinar o tamanho, a taxa de crescimento, a agressividade, o estágio, o prognóstico, o diagnóstico e outros atributos do tumor e do paciente. De modo ideal, as medições são

feitas em mais de um momento para mostrar as alterações na quantidade de cfDNA de tumor.

[00268] Tratamento de monitoramento: Por exemplo, um tumor de mama HER2+ é tratado com Herceptin. Uma sondagem que tem como objetivo o gene HER2 pode ser usada para monitorar a quantidade de cfDNA de tumor, que pode ser um substituto para o tamanho do tumor. Isso pode ser usado para determinar se o tumor está mudando em tamanho e o tratamento pode ser modificado para otimizar o resultado do paciente. Isso pode incluir alteração na dose, interrupção do tratamento, alteração para outra terapia, combinação de várias terapias.

[00269] Triagem de DNA de tumor: Atualmente, não há nenhuma triagem universal para o câncer. A presente invenção oferece uma maneira de detectar tumores em alguns ou em todos os locais do corpo. Por exemplo, um painel de sondagens é desenvolvido em um espaçamento de 100 kb no genoma. Esse painel pode ser usado como uma maneira de detectar a variação genética no genoma. Em um exemplo, o painel detecta mudanças no número de cópia de um determinado tamanho no genoma. Essas mudanças de número de cópia são associadas às células de tumor e, portanto, o teste detecta a presença de células tumorais. Diferentes tipos de tumor podem produzir diferentes quantidades de cfDNA de tumor ou podem ter variação em diferentes partes do genoma. Dessa forma, o teste pode ser capaz de identificar qual órgão é afetado. Além disso, a quantidade de cfDNA de tumor medido pode indicar o estágio ou o tamanho do tumor ou o local do tumor. Dessa maneira, o teste é uma triagem totalmente de genoma para muitos ou todos os tipos de tumor.

[00270] Para todos os testes acima, para mitigar os falsos positivos, um limite pode ser estabelecido para determinar a presença ou certeza de um tumor. Além disso, o teste pode ser repetido em várias amostras ou em vários momentos para aumentar

a certeza dos resultados. Os resultados também podem ser combinados com outras informações ou sintomas para fornecer mais informações ou determinadas informações sobre o tumor.

[00271] Os conjuntos de sondagem e os primers de exemplo que podem ser usados no método descrito neste documento para medir o número de cópia de regiões de ácido nucleico de interesse estão listados na Tabela 4 abaixo. Cada um dos conjuntos de sondagem de exemplo na Tabela 4 consiste em duas sondagens. A primeira sondagem (marcação) tem uma estrutura que inclui um local de preparação de avanço, marca e homologia 1. A segunda sondagem (rotulação) tem estrutura, incluindo a homologia 2 e o local do primer reverso, que é usado na rotulação. As sequências de componente das sondagens (sequência de marca, homologia etc.) também são mostradas.

Tabela 4: Sondagens de primers de exemplo.

Cromossomo	ID do Locus	Sondagem de marcação (Primer avançado + marca + 5pHom)	Sondagem de rotulação (3' - Salto + Primer reverso)	Primer avançado	Marca	Hom 5p	Hom 3p	Primer reverso
18	18-1	GCCCTCATCTTC TTCCCTGCGTTC TCACCACCCTCA CCAAGGAAGAAG TGAGGGCTTCTC (N° ID SEQ.: 1)	CGTGCTAATAGTC TCAGGGCTTCCTC CACCGAACGTGTC T (N° ID SEQ.: 17)	GCCCTCAT CTTCTTCC CTGC (N° ID SEQ.: 33)	GTTCTCA CCACCCT CACCAA (N° ID SEQ.: 34)	GGAAGAA GTGAGGG CTTCTC (N° ID SEQ.: 35)	CGTGCTA ATAGTCT CAGGGC (N° ID SEQ.: 51)	TTCCTCCA CCGAACGT GTCT (N° ID SEQ.: 67)
18	18-2	GCCCTCATCTTC TTCCCTGCGTTC TCACCACCCTCA CCAAAATCAAG GTGACCAGCTCC (N° ID SEQ.: 2)	CGACGCTTCATTG CTTCATTTTCCTC CACCGAACGTGTC T (N° ID SEQ.: 18)	GCCCTCAT CTTCTTCC CTGC (N° ID SEQ.: 33)	GTTCTCA CCACCCT CACCAA (N° ID SEQ.: 34)	AAATCAA GGTGACC AGCTCC (N° ID SEQ.: 36)	CGACGCT TCATTGC TTCATT (N° ID SEQ.: 52)	TTCCTCCA CCGAACGT GTCT (N° ID SEQ.: 67)
18	18-3	GCCCTCATCTTC TTCCCTGCGTTC TCACCACCCTCA	CTTGCGCCAAACA ATTGTCCTTCCTC CACCGAACGTGTC	GCCCTCAT CTTCTTCC CTGC (N°	GTTCTCA CCACCCT CACCAA	TCATCTG CCAAGAC AGAAGTT	CTTGCGC CAAACAA TTGTCC	TTCCTCCA CCGAACGT GTCT (N°

		CCAATCATCTGC CAAGACAGAAGT TC (N° ID SEQ.: 3)	T (N° ID SEQ.: 19)	ID SEQ.: 33)	(N° ID SEQ.: 34)	C (N° ID SEQ.: 37)	(N° ID SEQ.: 53)	ID SEQ.: 67)
18	18-4	GCCCTCATCTTC TTCCCTGCGTTC TCACCACCCTCA CCAAGCAGGAGA GTCAAAGGTCTG (N° ID SEQ.: 4)	GCTGCAGAGTTTG CATTCATTTTCCTC CACCGAACGTGTC T (N° ID SEQ.: 20)	GCCCTCAT CTTCTTCC CTGC (N° ID SEQ.: 33)	GTTCTCA CCACCCT CACCAA (N° ID SEQ.: 34)	GCAGGAG AGTCAAA GGTCTG (N° ID SEQ.: 38)	GCTGCAG AGTTTGC ATTCAT (N° ID SEQ.: 54)	TTCCTCCA CCGAACGT GTCT (N° ID SEQ.: 67)
18	18-5	GCCCTCATCTTC TTCCCTGCGTTC TCACCACCCTCA CCAAGTTGCCAT GGAGATTGTTGC (N° ID SEQ.: 5)	CATACACACAGAC CGAGAGTCTTCCT CCACCGAACGTGT CT (N° ID SEQ.: 21)	GCCCTCAT CTTCTTCC CTGC (N° ID SEQ.: 33)	GTTCTCA CCACCCT CACCAA (N° ID SEQ.: 34)	GTTGCCA TGGAGAT TGTTGC (N° ID SEQ.: 39)	CATACAC ACAGACC GAGAGTC (N° ID SEQ.: 55)	TTCCTCCA CCGAACGT GTCT (N° ID SEQ.: 67)
18	18-6	GCCCTCATCTTC TTCCCTGCGTTC TCACCACCCTCA CCAACAGCTCAG	GGATGTCAGCCAG CATAAGTTTCCTC CACCGAACGTGTC T (N° ID	GCCCTCAT CTTCTTCC CTGC (N° ID SEQ.: (N° ID	GTTCTCA CCACCCT CACCAA (N° ID	CAGCTCA GTGATGT CATTGC (N° ID	GGATGTC AGCCAGC ATAAGT (N° ID	TTCCTCCA CCGAACGT GTCT (N° ID SEQ.: (N° ID

		TGATGTCATTGC (N° ID SEQ.: 6)	SEQ.: 22)	33)	SEQ.: 34)	SEQ.: 40)	SEQ.: 56)	67)
18	18-7	GCCCTCATCTTC TTCCCTGCGTTC TCACCACCCTCA CCAACCTTGACC TCTGCTAATGTG G (N° ID SEQ.: 7)	GCAAGTGCCAAAC AGTTCTCTTCCTC CACCGAACGTGTC T (N° ID SEQ.: 23)	GCCCTCAT CTTCTTCC CTGC (N° ID SEQ.: 33)	GTTCTCA CCACCCT CACCAA (N° ID SEQ.: 34)	CCTTGAC CTCTGCT AATGTGG (N° ID SEQ.: 41)	GCAAGTG CCAAACA GTTCTC (N° ID SEQ.: 57)	TTCCTCCA CCGAACGT GTCT (N° ID SEQ.: 67)
18	18-8	GCCCTCATCTTC TTCCCTGCGTTC TCACCACCCTCA CCAACACCTGTC CAACAGCTACAG (N° ID SEQ.: 8)	GATTCCAGCACAC TTGAGTCTTTCCT CCACCGAACGTGT CT (N° ID SEQ.: 24)	GCCCTCAT CTTCTTCC CTGC (N° ID SEQ.: 33)	GTTCTCA CCACCCT CACCAA (N° ID SEQ.: 34)	CACCTGT CCAACAG CTACAG (N° ID SEQ.: 42)	GATTCCA GCACACT TGAGTCT (N° ID SEQ.: 58)	TTCCTCCA CCGAACGT GTCT (N° ID SEQ.: 67)
X	X-1	GCCCTCATCTTC TTCCCTGCGTTC TCACCACCCTCA CCAAAGAATGTA TCTTCAGGCCTG	CCGTTGCAGGTTT AAATGGCGCCCTA TTGCAAGCCCTCT T (N° ID SEQ.: 25)	GCCCTCAT CTTCTTCC CTGC (N° ID SEQ.: 33)	GTTCTCA CCACCCT CACCAA (N° ID SEQ.: 34)	AGAATGT ATCTTCA GGCCTGC (N° ID SEQ.: 41)	CCGTTGC AGGTTTA AATGGC (N° ID SEQ.: 57)	GCCCTATT GCAAGCCC TCTT (N° ID SEQ.: 68)

		C (N° ID SEQ.: 9)			34)	43)	59)	
X	X-2	GCCCTCATCTTC TTCCCTGCGTTC TCACCACCCTCA CCAAAAGTAATC ACTCTGGGTGGC (N° ID SEQ.: 10)	CAAGAGTGCTTTA TGGGCCTGCCCTA TTGCAAGCCCTCT T (N° ID SEQ.: 26)	GCCCTCAT CTTCTTCC CTGC (N° ID SEQ.: 33)	GTTCTCA CCACCCT CACCAA (N° ID SEQ.: 34)	AAGTAAT CACTCTG GGTGGC (N° ID SEQ.: 44)	CAAGAGT GCTTTAT GGGCCT (N° ID SEQ.: 60)	GCCCTATT GCAAGCCC TCTT (N° ID SEQ.: 68)
X	X-3	GCCCTCATCTTC TTCCCTGCGTTC TCACCACCCTCA CCAAAGCTCACA GACAACCTTGTG (N° ID SEQ.: 11)	GCACTCAAGGAGA TCAGACTGGCCCT ATTGCAAGCCCTC TT (N° ID SEQ.: 27)	GCCCTCAT CTTCTTCC CTGC (N° ID SEQ.: 33)	GTTCTCA CCACCCT CACCAA (N° ID SEQ.: 34)	AGCTCAC AGACAAC CTTGTG (N° ID SEQ.: 45)	GCACTCA AGGAGAT CAGACTG (N° ID SEQ.: 61)	GCCCTATT GCAAGCCC TCTT (N° ID SEQ.: 68)
X	X-4	GCCCTCATCTTC TTCCCTGCGTTC TCACCACCCTCA CCAAGCAATAGA CACCTACAGGCG (N° ID SEQ.: 10)	GGCTATCGAACTA CAACCACAGCCCT ATTGCAAGCCCTC TT (N° ID SEQ.: 28)	GCCCTCAT CTTCTTCC CTGC (N° ID SEQ.: 33)	GTTCTCA CCACCCT CACCAA (N° ID SEQ.: 34)	GCAATAG ACACCTA CAGGCG (N° ID SEQ.: 46)	GGCTATC GAACTAC AACCACA (N° ID SEQ.: 62)	GCCCTATT GCAAGCCC TCTT (N° ID SEQ.: 68)

		12)						
X	X-5	GCCCTCATCTTC TTCCCTGCGTTC TCACCACCCTCA CCAAGCACATTA TCAAAGGCCACG (N° ID SEQ.: 13)	GTAGCTGTCTGTG GTGTGATCGCCCT ATTGCAAGCCCTC TT (N° ID SEQ.: 29)	GCCCTCAT CTTCTTCC CTGC (N° ID SEQ.: 33)	GTTCTCA CCACCCT CACCAA (N° ID SEQ.: 34)	GCACATT ATCAAAG GCCACG (N° ID SEQ.: 47)	GTAGCTG TCTGTGG TGTGATC (N° ID SEQ.: 63)	GCCCTATT GCAAGCCC TCTT (N° ID SEQ.: 68)
X	X-6	GCCCTCATCTTC TTCCCTGCGTTC TCACCACCCTCA CCAACAACGACC TAAAGCATGTGC (N° ID SEQ.: 14)	CAAGAACTTCGA GCCTTAGCAGCCC TATTGCAAGCCCT CTT (N° ID SEQ.: 30)	GCCCTCAT CTTCTTCC CTGC (N° ID SEQ.: 33)	GTTCTCA CCACCCT CACCAA (N° ID SEQ.: 34)	CAACGAC CTAAAGC ATGTGC (N° ID SEQ.: 48)	CAAGAAA CTTCGAG CCTTAGC A (N° ID SEQ.: 64)	GCCCTATT GCAAGCCC TCTT (N° ID SEQ.: 68)
X	X-7	GCCCTCATCTTC TTCCCTGCGTTC TCACCACCCTCA CCAAGACATACA TGGCTTTGGCAG (N° ID SEQ.: 15)	GTGAACCAGTCCG AGTGAAAGCCCTA TTGCAAGCCCTCT T (N° ID SEQ.: 31)	GCCCTCAT CTTCTTCC CTGC (N° ID SEQ.: 33)	GTTCTCA CCACCCT CACCAA (N° ID SEQ.: 34)	GACATAC ATGGCTT TGGCAG (N° ID SEQ.: 49)	GTGAACC AGTCCGA GTGAAA (N° ID SEQ.: 65)	GCCCTATT GCAAGCCC TCTT (N° ID SEQ.: 68)

X	X-8	GCCCTCATCTTC TTCCTGCGTTC TCACCACCCTCA CCAAGAGATACT GCCACTTATGCA CG (N° ID SEQ.: 16)	GCAAATGATGTTC AGCACCACGCCCT ATTGCAAGCCCTC TT (N° ID SEQ.: 32)	GCCCTCAT CTTCTTCC CTGC (N° ID SEQ.: 33)	GTTCTCA CCACCCT CACCAA (N° ID SEQ.: 34)	GAGATAC TGCCACT TATGCAC G (N° ID SEQ.: 50)	GCAAATG ATGTTCA GCACCAC (N° ID SEQ.: 66)	GCCCTATT GCAAGCCC TCTT (N° ID SEQ.: 68)
---	-----	---	---	---	--	--	---	---

[00272] Os conjuntos de sondagem e os primers de exemplo que podem ser usados no método descrito neste documento para detectar um polimorfismo em um local SNP são listados na Tabela 5 abaixo. Cada um dos conjuntos de sondagem de exemplo na Tabela 5 consiste em três sondagens, duas sondagens específicas de alelo (que são usadas para rotulação) e uma sondagem de marcação. Nesses exemplos, as duas sondagens específicas do alelo têm sequências de homologia que são diferentes em um ou mais nucleotídeos. A estrutura na primeira sondagem alélica inclui um Alelo 1 do local do primer avançado e o Alelo 1 de homologia; e a estrutura da segunda sondagem alélica inclui um Alelo 2 do local do primer avançado e Alelo 2 de homologia. Na prática, os primers rotulados podem ser usados com diferentes rótulos nos dois primers (portanto, os rótulos são específicos do alelo). Nesses exemplos, também há uma sondagem de 3' universal que inclui uma região de homologia (sem qualquer SNP), a sequência de marcação e um local de primer reverso. As sequências de componente das sondagens (sequência de marca, homologia etc.) também são mostradas.

Tabela 5: Sondagens de primers de exemplo.

Cromossomo	Sondagem de rotulação - Alelo 1 (Alelo 1 do primer avançado + Hom 5p Alelo 1)	Sondagem de rotulação - Alelo 2 (Alelo 2 do primer avançado + Hom 5p Alelo 2)	Sondagem de marcação (Hom 3p + Marca + Primer reverso)	Primer avançado - Alelo 1	Primer avançado - Alelo 2	Hom 5p - Alelo 1	Hom 5p - Alelo 2	Hom 3p	Marca	Primer reverso
chr2 1	TTCCTCCACC GAACGTGTCT AGACCAGCAC AACTTACTcg (N° ID SEQ.: 69)	GCCCTATTGCAA GCCCTCTTAGAC CAGCACAACCTTA CTta (N° ID SEQ.: 112)	CACTTGACAA AGTTCTCACG CGCCGAAGTT CTCCGAAGGA TGCCCTCATC TTCTTCCCTG C (N° ID SEQ.: 155)	TTCCTC CACCGA ACGTGT CT (N° ID SEQ.: 67)	GCCCTAT TGCAAGC CCTCTT (N° ID SEQ.: 68)	AGACCA GCACAA CTTACT cg (N° ID SEQ.: 198)	AGACC AGCAC AACTT ACTta (N° ID SEQ.: 241)	CACTT GACAA AGTTC TCACG C (N° ID SEQ.: 284)	GCCGA AGTTC TCCGA AGGAT ID SEQ.: 327)	GCCCT CATCT TCTTC CCTGC ID SEQ.: 33)
chr3	TTCCTCCACC GAACGTGTCT CCAAATgCAC CTGCctg	GCCCTATTGCAA GCCCTCTTCCAA ATtCACCTGCCc a (N° ID	CATTAGGGAT TAACGGCTTG GGACAGACTG ACGGAGCTTC	TTCCTC CACCGA ACGTGT CT (N°	GCCCTAT TGCAAGC CCTCTT (N° ID	CCAAAT gCACCT GCctg (N° ID	CCAAA TtCAC CTGCC ca	CATTA GGGAT TAACG GCTTG	GACAG ACTGA CGGAG CTTCA	GCCCT CATCT TCTTC CCTGC

	(N° ID SEQ.: 70)	SEQ.: 113)	AGCCCTCATC TTCTTCCCTG C (N° ID SEQ.: 156)	ID SEQ.: 67)	SEQ.: 68)	SEQ.: 199)	(N° ID SEQ.: 242)	G (N° ID SEQ.: 285)	(N° ID SEQ.: 328)	(N° ID SEQ.: 33)
chr1 3	TTCCTCCACC GAACGTGTCT AGTTTGGACA AAGGCaATTc g (N° ID SEQ.: 71)	GCCCTATTGCAA GCCCTCTTAGTT TGGACAAAGGCg ATTta (N° ID SEQ.: 114)	CACACGTTAA GAAGACTTTC TGCTGACTCT GCCGCACATG ATCGCCCTCA TCTTCTTCCC TGC (N° ID SEQ.: 157)	TTCCTC CACCGA ACGTGT CT (N° ID SEQ.: 67)	GCCCTAT TGCAAGC CCTCTT (N° ID SEQ.: 68)	AGTTTG GACAAA GGCaAT Tcg (N° ID SEQ.: 200)	AGTTT GGACA AAGGC gATTt a (N° ID SEQ.: 243)	CACAC GTTAA GAAGA CTTTC TGC (N° ID SEQ.: 286)	TGACT CTGCC GCACA TGATC (N° ID SEQ.: 329)	GCCCT CATCT TCTTC CCTGC (N° ID SEQ.: 33)
chr3	TTCCTCCACC GAACGTGTCT TGAGCTTAGC CAATATCAAg AAGg (N° ID SEQ.: 72)	GCCCTATTGCAA GCCCTCTTTGAG CTTAGCCAATAT CAAcAAGa (N° ID SEQ.: 115)	CTAAGTGCCC TCCATGAGAA AGGATCCGAT AGCCCTCTGC AGGCCCTCAT CTTCTTCCCT GC (N° ID SEQ.: 158)	TTCCTC CACCGA ACGTGT CT (N° ID SEQ.: 67)	GCCCTAT TGCAAGC CCTCTT (N° ID SEQ.: 68)	TGAGCT TAGCCA ATATCA AgAAGg (N° ID SEQ.: 201)	TGAGC TTAGC CAATA TCAAc AAGa (N° ID SEQ.: 244)	CTAAG TGCCC TCCAT GAGAA AG (N° ID SEQ.: 287)	GATCC GATAG CCCTC TGCAG (N° ID SEQ.: 330)	GCCCT CATCT TCTTC CCTGC (N° ID SEQ.: 33)

chr9	TTCCTCCACC GAACGTGTCT ACGTGAACTT TCCTTGGTAc Ac (N° ID SEQ.: 73)	GCCCTATTGCAA GCCCTCTTACGT GAACTTTCCTTG GTAAAt (N° ID SEQ.: 116)	GCACAGATTT CCCACACTCT CAACAGGCCT GCTAAACACC GCCCTCATCT TCTTCCCTGC (N° ID SEQ.: 159)	TTCCTC CACCGA ACGTGT CT (N° ID SEQ.: 67)	GCCCTAT TGCAAGC CCTCTT (N° ID SEQ.: 68)	ACGTGA ACTTTC CTTGGT AcAc (N° ID SEQ.: 202)	ACGTG AACTT TCCTT GGTAA At (N° ID SEQ.: 245)	GCACA GATTT CCCAC ACTCT (N° ID SEQ.: 288)	CAACA GGCCT GCTAA ACACC (N° ID SEQ.: 331)	GCCCT CATCT TCTTC CCTGC (N° ID SEQ.: 33)
chr3	TTCCTCCACC GAACGTGTCT TGAAGATGTT CTAATACCTT GCcg (N° ID SEQ.: 74)	GCCCTATTGCAA GCCCTCTTTGAA GATGTTCTAATA CCTTGcta (N° ID SEQ.: 117)	CTTACAGGAG GTCTGGCATC AGGTCAACAA CCGAGGGACT CGCCCTCATC TTCTTCCCTG C (N° ID SEQ.: 160)	TTCCTC CACCGA ACGTGT CT (N° ID SEQ.: 67)	GCCCTAT TGCAAGC CCTCTT (N° ID SEQ.: 68)	TGAAGA TGTTCT AATACC TTGCcg (N° ID SEQ.: 203)	TGAAG ATGTT CTAAT ACCTT GCta (N° ID SEQ.: 246)	CTTAC AGGAG GTCTG GCATC A (N° ID SEQ.: 289)	GGTCA ACAAC CGAGG GACTC (N° ID SEQ.: 332)	GCCCT CATCT TCTTC CCTGC (N° ID SEQ.: 33)
chr1 7	TTCCTCCACC GAACGTGTCT CAGTGTGGAG ActGAACg	GCCCTATTGCAA GCCCTCTTCAGT GTGGAGACcGAA Ca (N° ID	CCACAATGAG AAGGCAGAGT TGTCATTAAT GCTGGCGGCG	TTCCTC CACCGA ACGTGT CT (N°	GCCCTAT TGCAAGC CCTCTT (N° ID	CAGTGT GGAGAC tGAACg (N° ID	CAGTG TGGAG ACcGA ACa	CCACA ATGAG AAGGC AGAG	TTGTC ATTAA TGCTG GCGGC	GCCCT CATCT TCTTC CCTGC

	(N° ID SEQ.: 75)	SEQ.: 118)	CCCTCATCTT CTTCCCTGC (N° ID SEQ.: 161)	ID SEQ.: 67)	SEQ.: 68)	SEQ.: 204)	(N° ID SEQ.: 247)	(N° ID SEQ.: 290)	(N° ID SEQ.: 333)	(N° ID SEQ.: 33)
chr1 6	TTCCTCCACC GAACGTGTCT AGGCAGGGTA ATGTCATGAA aTg (N° ID SEQ.: 76)	GCCCTATTGCAA GCCCTCTTAGGC AGGGTAATGTCA TGAAgTt (N° ID SEQ.: 119)	GCTGTGGCAT AGCTACACTC CGGTGACGGT TTGCAACTTT GCCCTCATCT TCTTCCCTGC (N° ID SEQ.: 162)	TTCCTC CACCGA ACGTGT CT (N° ID SEQ.: 67)	GCCCTAT TGCAAGC CCTCTT (N° ID SEQ.: 68)	AGGCAG GGTAAT GTCATG AAaTg (N° ID SEQ.: 205)	AGGCA GGGTA ATGTC ATGAA gTt (N° ID SEQ.: 248)	GCTGT GGCAT AGCTA CACTC (N° ID SEQ.: 291)	CGGTG ACGGT TTGCA ACTTT (N° ID SEQ.: 334)	GCCCT CATCT TCTTC CCTGC (N° ID SEQ.: 33)
chr2 1	TTCCTCCACC GAACGTGTCT GATTGTCTGG AGcGCTg (N° ID SEQ.: 77)	GCCCTATTGCAA GCCCTCTTGATT GTCTGGAGgGCT c (N° ID SEQ.: 120)	CAGGGTAATT TGTGGGTCTG GTCCGGCAGT TAAGGGTCTC GCCCTCATCT TCTTCCCTGC (N° ID SEQ.: 163)	TTCCTC CACCGA ACGTGT CT (N° ID SEQ.: 67)	GCCCTAT TGCAAGC CCTCTT (N° ID SEQ.: 68)	GATTGT CTGGAG cGCTg (N° ID SEQ.: 206)	GATTG TCTGG AGgGC Tc (N° ID SEQ.: 249)	CAGGG TAATT TGTGG GTCTG (N° ID SEQ.: 292)	GTCCG GCAGT TAAGG GTCTC (N° ID SEQ.: 335)	GCCCT CATCT TCTTC CCTGC (N° ID SEQ.: 33)
chr2	TTCCTCCACC	GCCCTATTGCAA	GGGCTATCCA	TTCCTC	GCCCTAT	AGGGAG	AGGGA	GGGCT	TACTC	GCCCT

	GAACGTGTCT AGGGAGCAAT AGGCcg (N° ID SEQ.: 78)	GCCCTCTTAGGG AGCAATAGGcta (N° ID SEQ.: 121)	GAAAGATAAG AATACTCACa AACGACTGCG CAGCCCTCAT CTTCTTCCCT GC (N° ID SEQ.: 164)	CACCGA ACGTGT CT (N° ID SEQ.: 67)	TGCAAGC CCTCTT (N° ID SEQ.: 68)	CAATAG GCcg (N° ID SEQ.: 207)	GCAAT AGGCt a (N° ID SEQ.: 250)	ATCCA GAAAG ATAAG AA (N° ID SEQ.: 293)	ACAAA CGACT GCGCA (N° ID SEQ.: 336)	CATCT TCTTC CCTGC (N° ID SEQ.: 33)
chr2	TTCCTCCACC GAACGTGTCT CTGCAGGGTA CAAcACg (N° ID SEQ.: 79)	GCCCTATTGCAA GCCCTCTTCTGC AGGGTACAAgAC a (N° ID SEQ.: 122)	CATAACTGGT GGAGTATTTc ACTCGTATAT GGCCGACTGG AGGGCCCTCA TCTTCTTCCC TGC (N° ID SEQ.: 165)	TTCCTC CACCGA ACGTGT CT (N° ID SEQ.: 67)	GCCCTAT TGCAAGC CCTCTT (N° ID SEQ.: 68)	CTGCAG GGTACA AcACg (N° ID SEQ.: 208)	CTGCA GGGTA CAAgA Ca (N° ID SEQ.: 251)	CATAA CTGGT GGAGT ATTTC ACT (N° ID SEQ.: 294)	CGTAT ATGGC CGACT GGAGG (N° ID SEQ.: 337)	GCCCT CATCT TCTTC CCTGC (N° ID SEQ.: 33)
chr1 9	TTCCTCCACC GAACGTGTCT CGTATCTGGG AAGAcGGc (N° ID	GCCCTATTGCAA GCCCTCTTCGTA TCTGGGAAGAtG Gg (N° ID SEQ.: 123)	CTTCAAGGAA GAAATTCAAC AGGGTAGGGT TTGCGGCGAT AAGGGCCCTC	TTCCTC CACCGA ACGTGT CT (N° ID SEQ.: ID	GCCCTAT TGCAAGC CCTCTT (N° ID SEQ.: ID	CGTATC TGGGAA GAcGGc (N° ID SEQ.: ID	CGTAT CTGGG AAGAt GGg (N° ID	CTTCA AGGAA GAAAT TCAAC AGGG	TAGGG TTTGC GGCGA TAAGG (N°	GCCCT CATCT TCTTC CCTGC (N°

	SEQ.: 80)		ATCTTCTTCC CTGC (N° ID SEQ.: 166)	SEQ.: 67)	68)	209)	ID SEQ.: 252)	(N° ID SEQ.: 295)	ID SEQ.: 338)	ID SEQ.: 33)
chr9	TTCCTCCACC GAACGTGTCT CCTGTAATCC CTTGCAATgc (N° ID SEQ.: 81)	GCCCTATTGCAA GCCCTCTTCCCTG TAATCCCTTGCA ATaa (N° ID SEQ.: 124)	CATGGATTCA ACACAGCAAA CACCAAGTCA ACCACCCGAG ACGCCCTCAT CTTCTTCCCT GC (N° ID SEQ.: 167)	TTCCTC CACCGA ACGTGT CT (N° ID SEQ.: 67)	68)	GCCCTAT TGCAAGC CCTCTT (N° ID SEQ.: 210)	CCTGT AATCC CTTGC AATaa (N° ID SEQ.: 253)	CATGG ATTCA ACACA GCAAA CA (N° ID SEQ.: 296)	CCAAG TCAAC CACCC GAGAC (N° ID SEQ.: 339)	GCCCT CATCT TCTTC CCTGC (N° ID SEQ.: 33)
chr1 6	TTCCTCCACC GAACGTGTCT GGTCTCAGCA CGGTtCTg (N° ID SEQ.: 82)	GCCCTATTGCAA GCCCTCTTGGTC TCAGCACGGTcC Tt (N° ID SEQ.: 125)	CTCTGACCTC CTTCACTCTT ACACTTCCCT GGCCTTCCCTT CTGCCCTCAT CTTCTTCCCT GC (N° ID SEQ.: 168)	TTCCTC CACCGA ACGTGT CT (N° ID SEQ.: 67)	68)	GCCCTAT TGCAAGC CCTCTT (N° ID SEQ.: 211)	GGTCT CAGCA GTtCTg (N° ID SEQ.: 254)	CTCTG ACCTC CTTCA CTt (N° ID SEQ.: 297)	ACTTC CCTGG CCTTC CTTCT (N° ID SEQ.: 340)	GCCCT CATCT TCTTC CCTGC (N° ID SEQ.: 33)

chr9	TTCCTCCACC GAACGTGTCT GCACCTCCCT AcCACAc (N° ID SEQ.: 83)	GCCCTATTGCAA GCCCTCTTGcAC CTCCCTAtCACA t (N° ID SEQ.: 126)	GCTTTCATTT GTGCTAAACC TCGCTTGGGT CCTCTCCTGA ACGCCCTCAT CTTCTTCCCT GC (N° ID SEQ.: 169)	TTCCTC CACCGA ACGTGT CT (N° ID SEQ.: 67)	GCCCTAT TGCAAGC CCTCTT (N° ID SEQ.: 68)	GCACCT CCCTAc CACAc (N° ID SEQ.: 212)	GCACC TCCCT AtCAC At (N° ID SEQ.: 255)	GCTTT CATTT GTGCT AAACC TC (N° ID SEQ.: 298)	GCTTG GGTCC TCTCC TGAAC (N° ID SEQ.: 341)	GCCCT CATCT TCTTC CCTGC (N° ID SEQ.: 33)
chr3	TTCCTCCACC GAACGTGTCT GCCTCTAGCT AGAGAGAAGt c (N° ID SEQ.: 84)	GCCCTATTGCAA GCCCTCTTGcCT CTAGCTAGAGAG AAGcg (N° ID SEQ.: 127)	CATCCCAGAT GCCCTCATAA CGTCCGAACC ACAATGCTGC CCTCATCTTC TTCCCTGC (N° ID SEQ.: 170)	TTCCTC CACCGA ACGTGT CT (N° ID SEQ.: 67)	GCCCTAT TGCAAGC CCTCTT (N° ID SEQ.: 68)	GCCTCT AGCTAG AGAGAA Gtc (N° ID SEQ.: 213)	GCCTC TAGCT AGAGA GAAGc g (N° ID SEQ.: 256)	CATCC CAGAT GCCCT CAT (N° ID SEQ.: 299)	AACGT CCGAA CCACA ATGCT (N° ID SEQ.: 342)	GCCCT CATCT TCTTC CCTGC (N° ID SEQ.: 33)
chr2 0	TTCCTCCACC GAACGTGTCT CTGGCAGTCT AGCCgTTAc (N° ID	GCCCTATTGCAA GCCCTCTTCTGG CAGTCTAGCCaT TAt (N° ID SEQ.: 128)	GTAGAAATCC CAAGGCAATC AGCTCCTCGC ATCCAACAGT CGGCCCTCAT	TTCCTC CACCGA ACGTGT CT (N° ID	GCCCTAT TGCAAGC CCTCTT (N° ID SEQ.: ID	CTGGCA GTCTAG CCgTTA c (N° ID	CTGGC AGTCT AGCCa TTAt (N° AG	GTAGA AATCC CAAGG CAATC AG	CTCCT CGCAT CCAAC AGTCG (N°	GCCCT CATCT TCTTC CCTGC (N°

	SEQ.: 85)		CTTCTTCCCT GC (N° ID SEQ.: 171)	SEQ.: 67)	68)	SEQ.: 214)	ID SEQ.: 257)	(N° ID SEQ.: 300)	ID SEQ.: 343)	ID SEQ.: 33)
chrX	TTCCTCCACC GAACGTGTCT TGTCTTAGAA TTTGGA ACT gGc (N° ID SEQ.: 86)	GCCCTATTGCAA GCCCTCTTTGTC TTAGAATTTGGC AACTaGt (N° ID SEQ.: 129)	GAACA ACTAA CTCCACAGAA CCCCACCGT AGCACTCCTT CTTGGCCCTCA TCTTCTTCCC TGC (N° ID SEQ.: 172)	TTCCTC CACCGA ACGTGT CT (N° ID SEQ.: 67)	(N° ID SEQ.: 68)	GCCCTAT TGCAAGC CCTCTT TGGCAA CTgGc (N° ID SEQ.: 215)	TGTCTT TGTCT TAGAA TTTGG CAACT aGt (N° ID SEQ.: 258)	GAACA ACTAA CTCCA CAGAA CCC (N° ID SEQ.: 301)	CCACC GTAGC ACTCC TTCTT (N° ID SEQ.: 344)	GCCCT CATCT TCTTC CCTGC (N° ID SEQ.: 33)
chr7	TTCCTCCACC GAACGTGTCT GCAGGAAAGC CTAcTGAac (N° ID SEQ.: 87)	GCCCTATTGCAA GCCCTCTTGcAG GAAAGCCTAtTG AAAt (N° ID SEQ.: 130)	GTGCAGAGGA CAGGAAGAAC GGAGCGTcGG TAGTGTAAG CCCTCATCTT CTTCCCTGC (N° ID SEQ.: 173)	TTCCTC CACCGA ACGTGT CT (N° ID SEQ.: 67)	(N° ID SEQ.: 68)	GCCCTAT TGCAAGC CCTCTT AcTGAA c (N° ID SEQ.: 216)	GCAGG AAAGC CTAtT GAAt (N° ID SEQ.: 259)	GTGCA GAGGA CAGGA AGAA (N° ID SEQ.: 302)	CGGAG CGTCG GTAGT GTAAA (N° ID SEQ.: 345)	GCCCT CATCT TCTTC CCTGC (N° ID SEQ.: 33)
chr3	TTCCTCCACC	GCCCTATTGCAA	GGTGCTTCAA	TTCCTC	GCCCTAT	GGGAGC	GGGAG	GGTGC	ACAAC	GCCCT

	GAACGTGTCT GGGAGCCAGA GAAATgTc (N° ID SEQ.: 88)	GCCCTCTTGGGA GCCAGAGAAAt Tct (N° ID SEQ.: 131)	GACATACACC TTAACAACTC GACGAACCTA CCGGCCCTCA TCTTCTTCCC TGC (N° ID SEQ.: 174)	CACCGA ACGTGT CT (N° ID SEQ.: 67)	TGCAAGC CCTCTT (N° ID SEQ.: 68)	CAGAGA AATgTC c (N° ID SEQ.: 217)	CCAGA GAAAT tTct (N° ID SEQ.: 260)	TTCAA GACAT ACACC TTA (N° ID SEQ.: 303)	TCGAC GAACC TACCG (N° ID SEQ.: 346)	CATCT TCTTC CCTGC (N° ID SEQ.: 33)
chr2	TTCCTCCACC GAACGTGTCT TGTCTCCAGT TCCAATTcAT tTAg (N° ID SEQ.: 89)	GCCCTATTGCAA GCCCTCTTTGTC TCCAGTTCCACT TCATgTAa (N° ID SEQ.: 132)	GGAACCTCTG TGACCTTGGA TGGCCCATCC TTATGTGCTG GCCCTCATCT TCTTCCCTGC (N° ID SEQ.: 175)	TTCCTC CACCGA ACGTGT CT (N° ID SEQ.: 67)	GCCCTAT TGCAAGC CCTCTT (N° ID SEQ.: 68)	TGTCTC CAGTTC CACTTC ATtTAg (N° ID SEQ.: 218)	TGTCT CCAGT TCCAC TTCAT gTAa (N° ID SEQ.: 261)	GGAAC CTCTG TGACC TTGGA (N° ID SEQ.: 304)	TGGCC CATCC TTATG TGCTG (N° ID SEQ.: 347)	GCCCT CATCT TCTTC CCTGC (N° ID SEQ.: 33)
chr1 5	TTCCTCCACC GAACGTGTCT CCCGTTAATT GCCTAcTcg (N° ID	GCCCTATTGCAA GCCCTCTTCCCG TTAATTGCCTAt Tta (N° ID SEQ.: 133)	CCCAGTGGTA CCTTCTGAAG GTCGTTATTG CTCAAGCCCG CCCTCATCTT	TTCCTC CACCGA ACGTGT CT (N° ID SEQ.: ID	GCCCTAT TGCAAGC CCTCTT (N° ID SEQ.: ID	CCCGTT AATTGC CTAcTc g (N° ID	CCCGT TAATT GCCTA tTta (N° ID	CCCAG TGGTA CCTTC TGAA (N° ID	GGTCG TTATT GCTCA AGCCC (N° ID	GCCCT CATCT TCTTC CCTGC (N° ID

	SEQ.: 90)		CTTCCCTGC (N° ID SEQ.: 176)	SEQ.: 67)	68)	SEQ.: 219)	ID SEQ.: 262)	ID SEQ.: 305)	ID SEQ.: 348)	ID SEQ.: 33)
chr1 5	TTCCTCCACC GAACGTGTCT CTCGGTCCCA CTGGaAAg (N° ID SEQ.: 91)	GCCCTATTGCAA GCCCTCTTCTCG GTCCCACTGGgA Aa (N° ID SEQ.: 134)	CTTCTGTTGC TTATTTGGGT AACTTGATTC TGGCCCTCCC ATCGCCCTCA TCTTCTTCCC TGC (N° ID SEQ.: 177)	TTCCTC CACCGA ACGTGT CT (N° ID SEQ.: 67)	GCCCTAT TGCAAGC CCTCTT (N° ID SEQ.: 68)	CTCGGT CCCCT GGaAAg (N° ID SEQ.: 220)	CTCGG TCCCA CTGGg AAa (N° ID SEQ.: 263)	CTTCT GTTGC TTATT TGGGT AAC (N° ID SEQ.: 306)	TTGAT TCTGG CCCTC CCATC (N° ID SEQ.: 349)	GCCCT CATCT TCTTC CCTGC (N° ID SEQ.: 33)
chr2	TTCCTCCACC GAACGTGTCT ACACCCATGA TTCAGTTACT g (N° ID SEQ.: 92)	GCCCTATTGCAA GCCCTCTTACAC CCATGATTCAGT TACca (N° ID SEQ.: 135)	CCCCTGGAT GCCTCCCTCA CGCCGGCTAT TTAGGTGCC TCATCTTCTT CCCTGC (N° ID SEQ.: 178)	TTCCTC CACCGA ACGTGT CT (N° ID SEQ.: 67)	GCCCTAT TGCAAGC CCTCTT (N° ID SEQ.: 68)	ACACCC ATGATT CAGTTA Ctg (N° ID SEQ.: 221)	ACACC CATGA TTCAG TTACC a (N° ID SEQ.: 264)	CCCAC TGGAT GCCTC C (N° ID SEQ.: 307)	CTCAC GCCGG CTATT TAGGT (N° ID SEQ.: 350)	GCCCT CATCT TCTTC CCTGC (N° ID SEQ.: 33)
chr9	TTCCTCCACC GAACGTGTCT	GCCCTATTGCAA GCCCTCTTGCTA	CGGAGAGACG CATCTGAAAG	TTCCTC CACCGA	GCCCTAT TGCAAGC	GCTAGT ATGAAC	GCTAG TATGA	CGGAG AGACG	AGTCT GGGTA	GCCCT CATCT

	GCTAGTATGA ACATCACAgG c (N° ID SEQ.: 93)	GTATGAACATCA CAaGt (N° ID SEQ.: 136)	TCTGGGTAGG TGGAGGACGC CCTCATCTTC TTCCCTGC (N° ID SEQ.: 179)	ACGTGT CT (N° ID SEQ.: 67)	CCTCTT (N° ID SEQ.: 68)	ATCACa gGc (N° ID SEQ.: 222)	ACATC ACaAG t (N° ID SEQ.: 265)	CATCT GAA (N° ID SEQ.: 308)	GGTGG AGGAC (N° ID SEQ.: 351)	TCTTC CCTGC (N° ID SEQ.: 33)
chr7	TTCCTCCACC GAACGTGTCT ACAAATGAGT AAGAAGCGAG Tcg (N° ID SEQ.: 94)	GCCCTATTGCAA GCCCTCTTACAA ATGAGTAAGAAG CGAGTta (N° ID SEQ.: 137)	CAGGATTTCC AGCTTACAGG GCGACTGAGC CACATCCAAC TGCCCTCATC TTCTTCCCTG C (N° ID SEQ.: 180)	TTCCTC CACCGA ACGTGT CT (N° ID SEQ.: 67)	GCCCTAT TGCAAGC CCTCTT (N° ID SEQ.: 68)	ACAAAT GAGTAA GAAGCG AGTcg (N° ID SEQ.: 223)	ACAAA TGAGT AAGAA GCGAG Tta (N° ID SEQ.: 266)	CAGGA TTTCC AGCTT ACAGG G (N° ID SEQ.: 309)	CGACT GAGCC ACATC CAACT (N° ID SEQ.: 352)	GCCCT CATCT TCTTC CCTGC (N° ID SEQ.: 33)
chr2 0	TTCCTCCACC GAACGTGTCT GATAAGGGTT GCTCTgCg (N° ID SEQ.: 95)	GCCCTATTGCAA GCCCTCTTGATA AGGGTTGCTCTa Ca (N° ID SEQ.: 138)	CTTGCAAGAT GTGCCTCTTA GAGCCTCAGC CGGAATTGAA GCCCTCATCT TCTTCCCTGC (N° ID SEQ.: 67)	TTCCTC CACCGA ACGTGT CT (N° ID SEQ.: 67)	GCCCTAT TGCAAGC CCTCTT (N° ID SEQ.: 68)	GATAAG GGTTGC TCTgCg (N° ID SEQ.: 224)	GATAA GGGTT GCTCT aCa (N° ID SEQ.: 266)	CTTGC AAGAT GTGCC TCTTA (N° ID SEQ.: 309)	GAGCC TCAGC CGGAA TTGAA (N° ID SEQ.: 352)	GCCCT CATCT TCTTC CCTGC (N° ID SEQ.: 33)

			SEQ.: 181)				267)	310)	353)	33)
chr2 0	TTCCTCCACC GAACGTGTCT CCATGCACCA GCTACcc (N° ID SEQ.: 96)	GCCCTATTGCAA GCCCTCTTCCAT GCACCAGCTACT a (N° ID SEQ.: 139)	GGGTGGTTTC TCTAAACACA AATTGCCATT CTGCACCAAT GCGCCCTCAT CTTCTTCCCT GC (N° ID SEQ.: 182)	TTCCTC CACCGA ACGTGT CT (N° ID SEQ.: 67)	GCCCTAT TGCAAGC CCTCTT (N° ID SEQ.: 68)	CCATGC ACCAGC TACcc (N° ID SEQ.: 225)	CCATG CACCA GCTAC ta (N° ID SEQ.: 268)	GGGTG GTTTC TCTAA ACACA AA (N° ID SEQ.: 311)	TTGCC ATTCT GCACC AATGC (N° ID SEQ.: 354)	GCCCT CATCT TCTTC CCTGC (N° ID SEQ.: 33)
chr1	TTCCTCCACC GAACGTGTCT AACTGTACCC TACTCCCAGc (N° ID SEQ.: 97)	GCCCTATTGCAA GCCCTCTTAACT GTACCCTACTCC CAat (N° ID SEQ.: 140)	GCAGGGTATT GAGAGAAGGA TCTATTGGTG TTCGCGGCTG ATGCCCTCAT CTTCTTCCCT GC (N° ID SEQ.: 183)	TTCCTC CACCGA ACGTGT CT (N° ID SEQ.: 67)	GCCCTAT TGCAAGC CCTCTT (N° ID SEQ.: 68)	AACTGT ACCCTA CTCCCA gc (N° ID SEQ.: 226)	AACTG TACCC TACTC CCAat (N° ID SEQ.: 269)	GCAGG GTATT GAGAG AAGGA TC (N° ID SEQ.: 312)	TATTG GTGTT CGCGG CTGAT (N° ID SEQ.: 355)	GCCCT CATCT TCTTC CCTGC (N° ID SEQ.: 33)
chr2	TTCCTCCACC GAACGTGTCT AGGACCAAGG	GCCCTATTGCAA GCCCTCTTAGGA CCAAGGGACCAG	GTGCACATTT CTTGATGAAG GGATGGGCGT	TTCCTC CACCGA ACGTGT	GCCCTAT TGCAAGC CCTCTT	AGGACC AAGGGA CCAGTT	AGGAC CAAGG GACCA	GTGCA CATTT CTTGA	ATGGG CGTAA CAGGA	GCCCT CATCT TCTTC

	GACCAGTTtA g (N° ID SEQ.: 98)	TTcAc (N° ID SEQ.: 141)	AACAGGAGGA CTGCCCTCAT CTTCTTCCCT GC (N° ID SEQ.: 184)	CT (N° ID SEQ.: 67)	(N° ID SEQ.: 68)	tAg (N° ID SEQ.: 227)	GTTcA c (N° ID SEQ.: 270)	TGAAG GG (N° ID SEQ.: 313)	GGACT (N° ID SEQ.: 356)	CCTGC (N° ID SEQ.: 33)
chr7	TTCCTCCACC GAACGTGTCT AGAGTTCCTC CAAGAAATTG cg (N° ID SEQ.: 99)	GCCCTATTGCAA GCCCTCTTAGAG TTCCTCCAAGAA ATTGta (N° ID SEQ.: 142)	GAGCAATGCC TGTTTCATGA GAGGAATGGC CTACCTGCAT CAGCCCTCAT CTTCTTCCCT GC (N° ID SEQ.: 185)	TTCCTC CACCGA ACGTGT CT (N° ID SEQ.: 67)	GCCCTAT TGCAAGC CCTCTT (N° ID SEQ.: 68)	AGAGTT CCTCCA AGAAAT TGcg (N° ID SEQ.: 228)	AGAGT TCCTC CAAGA AATTG ta (N° ID SEQ.: 271)	GAGCA ATGCC TGTTT CATGA GA (N° ID SEQ.: 314)	GGAAT GGCCT ACCTG CATCA (N° ID SEQ.: 357)	GCCCT CATCT TCTTC CCTGC (N° ID SEQ.: 33)
chr5	TTCCTCCACC GAACGTGTCT ACATTATACA GCATGCTGGc TAtc (N° ID SEQ.: 100)	GCCCTATTGCAA GCCCTCTTACAT TATACAGCATGC TGGtTAga (N° ID SEQ.: 143)	GTTAACATTA TACAGCATGG TGGCCCCGTT GTTGTCATCG CATCGCCCTC ATCTTCTTCC CTGC (N°	TTCCTC CACCGA ACGTGT CT (N° ID SEQ.: 67)	GCCCTAT TGCAAGC CCTCTT (N° ID SEQ.: 68)	ACATTA TACAGC ATGCTG GcTAtc (N° ID SEQ.: 229)	ACATT ATACA GCATG CTGGt TAga (N° ID SEQ.: ID	GTTAA CATT TACAG CATGG TGGC (N° ID SEQ.: ID	CCCGT TGTTG TCATC GCATC (N° ID SEQ.: SEQ.: ID	GCCCT CATCT TCTTC CCTGC (N° ID SEQ.: SEQ.: ID

			ID SEQ.: 186)				SEQ.: 272)	SEQ.: 315)	358)	33)
chr2	TTCCTCCACC GAACGTGTCT GAGGAAGAAA GTGAGgTTTG c (N° ID SEQ.: 101)	GCCCTATTGCAA GCCCTCTTGAGG AAGAAAGTGAGa TTTGt (N° ID SEQ.: 144)	GCAGAACATG TCCTGAAGCG TTCGATGCGT CCCATGAGTG CCCTCATCTT CTTCCCTGC (N° ID SEQ.: 187)	TTCCTC CACCGA ACGTGT CT (N° ID SEQ.: 67)	GCCCTAT TGCAAGC CCTCTT (N° ID SEQ.: 68)	GAGGAA GAAAGT GAGgTT TGc (N° ID SEQ.: 230)	GAGGA AGAAA GTGAG aTTTG t (N° ID SEQ.: 273)	GCAGA ACATG TCCTG AAGC (N° ID SEQ.: 316)	GTTCG ATGCG TCCCA TGAGT (N° ID SEQ.: 359)	GCCCT CATCT TCTTC CCTGC (N° ID SEQ.: 33)
chr1 5	TTCCTCCACC GAACGTGTCT CTGAATTATG TGCTTACCAa GAGc (N° ID SEQ.: 102)	GCCCTATTGCAA GCCCTCTTCTGA ATTATGTGCTTA CCAgGAGt (N° ID SEQ.: 145)	CAGCTTGTTT CCAAACCCAT CAACCCGCGT AGATGTTCCCT GCCCTCATCT TCTTCCCTGC (N° ID SEQ.: 188)	TTCCTC CACCGA ACGTGT CT (N° ID SEQ.: 67)	GCCCTAT TGCAAGC CCTCTT (N° ID SEQ.: 68)	CTGAAT TATGTG CCTTACC AaGAGc (N° ID SEQ.: 231)	CTGAA TTATG TGCTT ACCAg GAGt (N° ID SEQ.: 274)	CAGCT TGTTT CCAAA CCCAT (N° ID SEQ.: 317)	CAACC CGCGT AGATG TTCCT (N° ID SEQ.: 360)	GCCCT CATCT TCTTC CCTGC (N° ID SEQ.: 33)
chr9	TTCCTCCACC GAACGTGTCT TGGGTTCTGA	GCCCTATTGCAA GCCCTCTTTGGG TTCTGATAACCT	CAAAGTGTGG AAGTTGCTTC CGCCAGCTCA	TTCCTC CACCGA ACGTGT	GCCCTAT TGCAAGC CCTCTT	TGGGTT CTGATA ACCTTA	TGGGT TCTGA TAACC	CAAAG TGTGG AAGTT	GCCAG CTCAA GAGTG	GCCCT CATCT TCTTC

	TAACCTTATC AAgc (N° ID SEQ.: 103)	TATCAAAct (N° ID SEQ.: 146)	AGAGTGTAGC CGCCCTCATC TTCTTCCCTG C (N° ID SEQ.: 189)	CT (N° ID SEQ.: 67)	(N° ID SEQ.: 68)	TCAAgc (N° ID SEQ.: 232)	TTATC AAAct (N° ID SEQ.: 275)	GCTTC C (N° ID SEQ.: 318)	TAGCC (N° ID SEQ.: 361)	CCTGC (N° ID SEQ.: 33)
chr2	TTCCTCCACC GAACGTGTCT GGTTAGTCAA ACATGcTGc (N° ID SEQ.: 104)	GCCCTATTGCAA GCCCTCTTGGTT AGTCAAACATGt TGt (N° ID SEQ.: 147)	GGTCGACTTT GTCCATCCTT CTTGATCCTG CGCGATGTGC CCTCATCTTC TTCCCTGC (N° ID SEQ.: 190)	TTCCTC CACCGA ACGTGT CT (N° ID SEQ.: 67)	GCCCTAT TGCAAGC CCTCTT (N° ID SEQ.: 68)	GGTTAG TCAAAC ATGcTG c (N° ID SEQ.: 233)	GGTTA GTCAA ACATG tTGt (N° ID SEQ.: 276)	GGTCG ACTTT GTCCA TCC (N° ID SEQ.: 319)	TTCTT GATCC TGCGC GATGT (N° ID SEQ.: 362)	GCCCT CATCT TCTTC CCTGC (N° ID SEQ.: 33)
chr1 7	TTCCTCCACC GAACGTGTCT GACACTGGCA GAATCAAAtC Ac (N° ID SEQ.: 105)	GCCCTATTGCAA GCCCTCTTGACA CTGGCAGAATCA AAcCAa (N° ID SEQ.: 148)	CTCTGTTGCC TGTGGACTCA TCGCAGGCGT TCCCTATACG CCCTCATCTT CTTCCCTGC (N° ID SEQ.: 191)	TTCCTC CACCGA ACGTGT CT (N° ID SEQ.: 67)	GCCCTAT TGCAAGC CCTCTT (N° ID SEQ.: 68)	GACACT GGCAGA ATCAAA tCAc (N° ID SEQ.: 234)	GACAC TGGCA GAATC AAAcC Aa (N° ID SEQ.: 320)	CTCTG TTGCC TGTGG ACTC (N° ID SEQ.: 320)	ATCGC AGGCG TTCCC TATAC (N° ID SEQ.: 363)	GCCCT CATCT TCTTC CCTGC (N° ID SEQ.: 33)

							277)			
chr6	TTCCTCCACC GAACGTGTCT AGAGTTACAC CTTTAGCTAA CcAc (N° ID SEQ.: 106)	GCCCTATTGCAA GCCCTCTTAGAG TTACACCTTTAG CTAAcAg (N° ID SEQ.: 149)	CTAACTAGAA TTAGTCTGCC TGCCTATTGG ACCTCCGACC ACGAGCCCTC ATCTTCTTCC CTGC (N° ID SEQ.: 192)	TTCCTC CACCGA ACGTGT CT (N° ID SEQ.: 67)	GCCCTAT TGCAAGC CCTCTT (N° ID SEQ.: 68)	AGAGTT ACACCT TTAGCT AACcAc (N° ID SEQ.: 235)	AGAGT TACAC CTTTA GCTAA CtAg (N° ID SEQ.: 278)	CTAAC TAGAA TTAGT CTGCC TGCC (N° ID SEQ.: 321)	TATTG GACCT CCGAC CACGA (N° ID SEQ.: 364)	GCCCT CATCT TCTTC CCTGC (N° ID SEQ.: 33)
chr7	TTCCTCCACC GAACGTGTCT CCAGGAGTTC AAGaAGCg (N° ID SEQ.: 107)	GCCCTATTGCAA GCCCTCTTCCAG GAGTTCAAGgAG Ca (N° ID SEQ.: 150)	GTGAGCCATA ATCGTGTCAA GCCACCATTT AGATCCGCGG CCCTCATCTT CTTCCCTGC (N° ID SEQ.: 193)	TTCCTC CACCGA ACGTGT CT (N° ID SEQ.: 67)	GCCCTAT TGCAAGC CCTCTT (N° ID SEQ.: 68)	CCAGGA GTTCAA GaAGCg (N° ID SEQ.: 236)	CCAGG AGTTC AAGgA GCa (N° ID SEQ.: 279)	GTGAG CCATA ATCGT GTCA (N° ID SEQ.: 322)	AGCCA CCATT TAGAT CCGCG (N° ID SEQ.: 365)	GCCCT CATCT TCTTC CCTGC (N° ID SEQ.: 33)
chr4	TTCCTCCACC GAACGTGTCT ACCACTCCTT TCTCCCaTCT	GCCCTATTGCAA GCCCTCTTACCA CTCCTTTCTCCC gTCTt (N° ID	GAGAATTAAT GCTCCCTCTC CTGGACCAGT AGAAGTCTGC	TTCCTC CACCGA ACGTGT CT (N°	GCCCTAT TGCAAGC CCTCTT (N° ID	ACCACT CCTTTC TCCCAt CTc	ACCAC TCCTT TCTCC CgTCT	GAGAA TTAAT GCTCC CTCTC	GACCA GTAGA AGTCT GCCCCG	GCCCT CATCT TCTTC CCTGC

	c (N° ID SEQ.: 108)	SEQ.: 151)	CCGGCCCTCA TCTTCTTCCC TGC (N° ID SEQ.: 194)	ID SEQ.: 67)	SEQ.: 68)	(N° ID SEQ.: 237)	t (N° ID SEQ.: 280)	CTG (N° ID SEQ.: 323)	(N° ID SEQ.: 366)	(N° ID SEQ.: 33)
chr2	TTCCTCCACC GAACGTGTCT GTCTTATGGG ACAATGGTtG ATAg (N° ID SEQ.: 109)	GCCCTATTGCAA GCCCTCTTGTCT TATGGGACAATG GTcGATAt (N° ID SEQ.: 152)	GTGGTCTGCT GTTGACCAAT TTCAGAATGG CCGAGCTGTG CCCTCATCTT CTTCCCTGC (N° ID SEQ.: 195)	TTCCTC CACCGA ACGTGT CT (N° ID SEQ.: 67)	GCCCTAT TGCAAGC CCTCTT (N° ID SEQ.: 68)	GTCTTA TGGGAC AATGGT tGATAg (N° ID SEQ.: 238)	GTCTT ATGGG ACAAT GGTcG ATAt (N° ID SEQ.: 281)	GTGGT CTGCT GTTGA CCAA (N° ID SEQ.: 324)	TTTCA GAATG GCCGA GCTGT (N° ID SEQ.: 367)	GCCCT CATCT TCTTC CCTGC (N° ID SEQ.: 33)
chr1 7	TTCCTCCACC GAACGTGTCT CTACCCTCAA CCCTCgTc (N° ID SEQ.: 110)	GCCCTATTGCAA GCCCTCTTCTAC CCTCAACCCTCa Tt (N° ID SEQ.: 153)	GGTTGCAACT GCTGATCTAT AGGTGACCTT CTTGACGCC GCCCTCATCT TCTTCCCTGC (N° ID SEQ.: 196)	TTCCTC CACCGA ACGTGT CT (N° ID SEQ.: 67)	GCCCTAT TGCAAGC CCTCTT (N° ID SEQ.: 68)	CTACCC TCAACC CCTCTT CTCgTc (N° ID SEQ.: 239)	CTACC CTCAA CCCTC aTt (N° ID SEQ.: 282)	GGTTG CAACT GCTGA TCTAT (N° ID SEQ.: 325)	AGGTG ACCTT CTTGT ACGCC (N° ID SEQ.: 368)	GCCCT CATCT TCTTC CCTGC (N° ID SEQ.: 33)

chr7	TTCCTCCACC GAACGTGTCT CCAAGACTGA TCATGCcg (N° ID SEQ.: 111)	GCCCTATTGCAA GCCCTCTTCCAA GACTGATCATGC ta (N° ID SEQ.: 154)	CTTTCCCAGT CAAGGCAGGG CGCGTCCTTA TTTCCATCGC CCTCATCTTC TTCCCTGC (N° ID SEQ.: 197)	TTCCTC CACCGA ACGTGT CT (N° ID SEQ.: 67)	GCCCTAT TGCAAGC CCTCTT (N° ID SEQ.: 68)	CCAAGA CTGATC ATGCcg (N° ID SEQ.: 240)	CCAAG ACTGA TCATG Cta (N° ID SEQ.: 283)	CTTTC CCAGT CAAGG CAG (N° ID SEQ.: 326)	GGCGC GTCCT TATTT CCATC (N° ID SEQ.: 369)	GCCCT CATCT TCTTC CCTGC (N° ID SEQ.: 33)
------	--	---	--	--	--	---	--	--	--	---

## EXEMPLOS

[00273] O seguinte protocolo descreve o processamento de até 24 amostras de DNA livre de célula por meio da hibridização-ligação, purificação, amplificação, preparação de alvo de microarranjo, hibridização de microarranjo e lavagem de microarranjo.

[00274] Os materiais a seguir foram preparados ou obtidos: DNA livre de célula (cfDNA) em um volume de 20 µL de água; mistura de sondagem: mistura de todos os oligonucleotídeos de sondagem de marcação e rotulação em uma concentração de 2 nM cada; Ligase *Taq* (40 U/µL); Esferas magnéticas: MyOne Streptavidin C1 Dynabeads; Ligação de esferas e Tampão de lavagem, concentrações 1X e 2X; Primer de amplificação de avanço, fosfato 5' modificado; primer de amplificação reversa, rotulado; Enzima AmpliTaq Gold (5 U/µL); Mistura dNTP; Exonuclease Lambda (5 U/µL); Tampão de hibridização, 1.25X; Oligonucleotídeos de controle de hibridização; Tampão de Lavagem de Microarranjo A; Tampão de Lavagem de Microarranjo B; Tampão de Lavagem de Microarranjo

[00275] **Reação de hibridização-ligação:** As amostras de cfDNA (20 µL) foram adicionadas aos poços A3-H3 de uma placa de reação de 96 poços. Os seguintes reagentes foram adicionados a cada amostra de cfDNA para um volume de reação total de 50 µL e misturado por pipetagem 5-8 vezes para cima e para baixo.

Componente	Volume
H <sub>2</sub> O	19,33 µL
Mistura de sondagem	5 µL
Tampão Ligase	5 µL

10X <i>Taq</i>	
Ligase <i>Taq</i>	0,67 µL

[00276] A placa foi colocada em um ciclador térmico e ligada usando o seguinte perfil cíclico: (i) 95 °C durante 5 minutos; (ii) 95 °C durante 30 segundos; (iii) 45 °C durante 25 minutos; (iv) Repita as etapas b a c 4 vezes; suspenso a (v) 4 °C.

**Purificação do produto de hibridização-ligação:**

[00277] Dynabeads de lavagem: um frasco de Dynabeads foi vortexado em configuração mais elevada durante 30 segundos. 260 µL de esferas foram transferidas para um tubo de 1,5 mL. 900 µL de Ligação de esferas 2X e Tampão de Lavagem e esferas de mistura foram combinados por pipetagem 5-8 vezes para cima e para baixo. O tubo foi posicionado em um suporte magnético durante 1 minuto e o sobrenadante foi descartado. O tubo do suporte magnético foi removido e suspendeu novamente as esferas magnéticas lavadas em 900 µL de Ligação de esferas 2X e Tampão de Lavagem por pipetagem 5-8 vezes para cima e para baixo. O tubo foi posicionado no suporte magnético durante 1 minuto e o sobrenadante foi descartado. O tubo foi removido do suporte magnético e foi adicionado 1,230 µL de Ligação de esferas 2X e Tampão de Lavagem. As esferas foram suspensas novamente por pipetagem 5-8 vezes para cima e para baixo.

[00278] Imobilizar produtos HL: 50 µL de esferas lavadas foram transferidos para cada produto de reação de hibridização-ligação na placa de reação de 96 poços e misturado por pipetagem 8 vezes para cima e para baixo, foram incubados durante 15 minutos em temperatura ambiente, misturados em um ímã de placa duas vezes durante o tempo de incubação. As esferas foram separadas em um ímã de placa durante 3 minutos e depois removidas e o sobrenadante foi descartado. A placa foi removida do ímã de placa, foram adicionados 200 µL de Ligação de esferas 1X e

Tampão de Lavagem, e as esferas foram suspensas novamente por pipetagem 5-8 vezes para cima e para baixo. A placa foi posicionada no ímã de placa durante 1 minuto, e o sobrenadante foi descartado. A placa foi removida do ímã de placa, foram adicionados 180 µL de 1X SSC, e as esferas foram suspensas novamente por pipetagem 5-8 vezes para cima e para baixo. A placa foi posicionada no ímã de placa durante 1 minuto, e o sobrenadante foi descartado.

[00279] Purificação de produtos de hibr.-ligação: Foram adicionados 50 µL de 0,15 M NaOH preparado recentemente em cada poço, e as esferas foram suspensas novamente por pipetagem 5-8 vezes para cima e para baixo e incubadas em temperatura ambiente durante 10 minutos. A placa foi posicionada no ímã de placa durante 2 minutos e depois foi removido, e o sobrenadante foi descartado. A placa foi removida do ímã de placa, foram adicionados 200 µL de 0,1 M NaOH, e as esferas foram suspensas novamente por pipetagem 5-8 vezes para cima e para baixo. A placa foi posicionada no ímã de placa durante 1 minuto, e o sobrenadante foi descartado. A placa foi removida do ímã de placa e foram adicionados 180 µL de 0,1 M NaOH, e as esferas foram suspensas novamente por pipetagem 5-8 vezes para cima e para baixo. A placa foi posicionada no ímã de placa durante 1 minuto, e o sobrenadante foi descartado. A placa foi removida do ímã de placa, foram adicionados 200 µL de Ligação de esferas 1X e Tampão de Lavagem, e as esferas foram suspensas novamente por pipetagem 5-8 vezes para cima e para baixo. Posicione a placa no ímã de placa durante 1 minuto e descarte o sobrenadante. Remova a placa do ímã de placa, adicione 180 µL TE, e as esferas foram suspensas novamente por pipetagem 5-8 vezes para cima e para baixo. A placa foi posicionada no ímã de placa durante 1 minuto, e o sobrenadante foi descartado. 20 µL de água foram adicionados a cada poço, e as esferas foram suspensas novamente por

pipetagem 5-8 vezes para cima e para baixo. A placa foi selada e armazenada a 4 °C até que fosse usada nas etapas subsequentes.

[00280] **Amplificação:** Os reagentes a seguir foram adicionados a cada produto de reação de hibridização-ligação na placa de reação de 96 poços para um volume de reação total de 50 µL.

Componente	Volume
H2O	17,25 µL
Primer avançado, 10 µM	2,5 µL
Primer reverso, 10 µM	2,5 µL
Mistura de 4 mM dNTP (L/N 052114)	2,5 µL
Tampão 10X AmpliTaq Gold	5 µL
Enzima AmpliTaq Gold	0,25 µL

[00281] A placa foi colocada em um ciclador térmico, e as sondagens foram ligadas usando o seguinte perfil cíclico: (i) 95 °C durante 5 minutos; (ii) 95 °C durante 30 segundos; (iii) 45 °C durante 25 minutos; (iv) Repita as etapas b a c 4 vezes; suspenso a (v) 4 °C.

[00282] Purificação do produto de hibridização-ligação: os reagentes foram misturados por pipetagem 5-8 vezes para cima e para baixo. A placa foi colocada em um ciclador térmico, e as sondagens foram amplificadas usando o seguinte perfil cíclico: (i) 95 °C durante 5 minutos; (ii) 95 °C durante 30 segundos; (iii) 54 °C durante 30 segundos; (iv) 72 °C durante 60 segundos, (v) Repita as etapas b a d 29 vezes; (vi) 72 °C durante 5 minutos; (vii) Repita as etapas b a c 4 vezes; e mantenha (v) 4 °C.

[00283] **Preparação de alvo de microarranjo (digestão de filamento único):** os reagentes a seguir foram adicionados a cada produto de reação na placa de reação de 96 poços para um volume de reação total de 60 µL.

Componente	Volume
H <sub>2</sub> O	3 µL
Tampão 10X Exonuclease Lambda	6 µL
Enzima Exonuclease Lambda	1 µL

[00284] Os reagentes foram misturados por pipetagem 5-8 vezes para cima e para baixo. A placa foi colocada em um ciclador térmico, e as sondagens foram digeridas usando o seguinte perfil cíclico: (i) 37 °C durante 60 minutos; (ii) 80 °C durante 30 minutos; suspenso a (iii) 4 °C. A placa foi posicionada em um Speed-vac e as amostras foram secadas em ambiente aquecido por aproximadamente 60 minutos ou até que todo o líquido fosse evaporado. As amostras foram lavadas a 4 °C no escuro até que fossem usadas em etapas subsequentes.

[00285] Hibridização de microarranjo: os seguintes reagentes foram adicionados a cada Alvo de Microarranjo seco na placa de reação de 96 poços para um volume de reação total de 20 µL.

Componente	Volume
H <sub>2</sub> O	3 µL
Tampão de hibridização 1.25X	16 µL
Oligonucleotídeos de controle de hibridização	1 µL

[00286] Os reagentes foram misturados pela pipetagem 10-20 vezes para cima e para baixo para serem suspensos novamente e

foram agitados em movimento giratório brevemente para empurrar o conteúdo para o fundo dos poços da placa. A placa foi colocada em um ciclador térmico, e as sondagens foram desnaturadas usando o seguinte perfil cíclico: (i) 70 °C durante 30 minutos; suspenso a (ii) 42 °C. O código de barras do microarranjo a ser usado foi registrado para cada amostra na folha de acompanhamento. É preparada uma câmara de hibridização contendo um Lifter Slip para cada microarranjo a ser processado. Para cada amostra, 15 µL de alvo de Microarray foram adicionados ao centro de um Lifter Slip em uma câmara de hibridização e o microarranjo apropriado foi imediatamente colocado no fluido alvo posicionando-se a borda no lifter slip e permitindo que ela repousasse lentamente em posição horizontal. As câmaras de hibridização foram fechadas e incubadas a 42 °C durante 60 minutos. As câmaras de hibridização foram abertas e cada microarranjo foi removido dos Lifter Slips e colocados em um rack imerso no Tampão de Lavagem de Microarranjo A. Assim que todos os microarranjos estavam no rack, o rack foi agitado a 650 rpm durante 5 minutos. O rack de microarranjos foi removido do Tampão de Lavagem de Microarranjo A, o líquido em excesso foi drenado em uma área limpa, e o rack foi rapidamente colocado em um Tampão de Lavagem de Microarranjo B. O rack foi agitado a 650 rpm durante 5 minutos. O rack de microarranjos foi removido do Tampão de Lavagem de Microarranjo B, o líquido em excesso foi drenado em uma área limpa, e o rack foi rapidamente colocado em um Tampão de Lavagem de Microarranjo C. O rack foi agitado a 650 rpm durante 5 minutos. Imediatamente após a conclusão da lavagem de 5 minutos no Tampão de Lavagem de Microarranjo C, o rack dos microarranjos foi lentamente removido do tampão. Isso leva 5-10 segundos para maximizar a laminação do tampão de lavagem da superfície da lamínula. O líquido em excesso foi drenado em uma área limpa. Um aspirador a vácuo foi usado para remover

quaisquer gotículas de tampão remanescentes presentes na superfície de cada microarranjo. Os microarranjos foram armazenados em um rack deslizando sob nitrogênio no escuro até que os microarranjos fossem analisados.

REIVINDICAÇÕES

1. Um método de detecção de uma variação genética em uma amostra genética de um objeto, caracterizado por compreender o contato do primeiro e segundo conjuntos de sondagem com uma amostra genética na qual o primeiro conjunto de sondagem compreende uma primeira sondagem de rotulação e uma primeira sondagem de marcação e um segundo conjunto de sondagem compreende uma segunda sondagem de rotulação e uma segunda sondagem de marcação, hibridização de pelo menos partes do primeiro e segundo conjuntos de sondagem na primeira e segunda região de ácido nucleico de interesse nas moléculas de nucleotídeo da amostra genética, respectivamente, ligando o primeiro conjunto de sondagem pelo menos ligando a primeira sondagem de rotulação e a primeira sondagem de marcação, ligando o segundo conjunto de sondagem pelo menos ligando a segunda sondagem de rotulação e a segunda sondagem de marcação, amplificando opcionalmente os conjuntos de sondagem ligados, imobilizando as sondagens de marcação e/ou as sondas de marcação amplificadas das mesmas em um local predeterminado em um substrato no qual a primeira e a segunda sondagem de marcação e/ou as sondagens de marcação amplificadas ligadas às sondagens de marcação imobilizadas e/ou as sondas de marcação amplificadas das mesmas que compreendem um primeiro e segundo rótulos, respectivamente, o primeiro e o segundo rótulos são diferentes, os rótulos imobilizados são opticamente determináveis, a primeira e a segunda sondagem de marcação imobilizada compreendem a primeira e a segunda marca, respectivamente e a etapa de imobilização é realizada pela imobilização de marcas no local predeterminado, contagem (i) de um primeiro número do primeiro rótulo imobilizado no substrato e (ii) um segundo número do segundo rótulo imobilizado no substrato e comparação dos primeiro e segundo números para determinar a variação genética na amostra genética.

2. O método, de acordo com a reivindicação 1, caracterizado por compreender a rotulação da primeira e da segunda sondagem de rotulação com o primeiro e o segundo rótulo, respectivamente,

antes da etapa de contato e/ou compreender a marcação da primeira e da segunda sondagem de marcação com a primeira e a segunda marca, respectivamente, antes da etapa de contato.

3. O método, de acordo com qualquer uma das reivindicações 1 e 2, caracterizado pelo método consistir na amplificação dos conjuntos de sondagem ligados com ou sem rotulação das sondagens durante a amplificação.

4. O método, de acordo com a reivindicação 1, caracterizado pela primeira e segunda sondagem de rotulação compreenderem uma sequência de preparação de avanço ou retrocesso e cada uma das primeira e segunda sondagens de marcação consistir em uma sequência de preparação reversa ou de avanço correspondente e uma sequência de nucleotídeo de marcação como uma marca, o método consiste na amplificação dos conjuntos de sondagem ligados, a etapa de amplificação consiste na amplificação (i) da primeira rotulação ligada e das sondagens de marcação com o primeiro e o segundo primer de avanço e reversa que são hibridizadas nas primeiras sequências de preparação de avanço e retrocesso, respectivamente, em que o primeiro primer de avanço ou retrocesso que é hibridizado na primeira sondagem de rotulação consiste no primeiro rótulo e (ii) das segundas sondagens de rotulação e marcação ligadas com as segundas preparações de avanço e retrocesso que são hibridizadas nas sequências de preparação de avanço e retrocesso, respectivamente, em que o segundo primer de avanço ou retrocesso que é hibridizado na segunda sondagem de rotulação consiste no segundo rótulo, as sequências de nucleotídeo de marcação amplificada das sondagens de marcação são imobilizadas em um local predeterminado em um substrato, em que as sequências de nucleotídeo de marcação amplificada da primeira e segunda sondagem de marcação são a primeira e a segunda marca, o primeiro número é o número do primeiro rótulo no primeiro conjunto de sondagem amplificada imobilizada para o substrato e o segundo número é o número do segundo rótulo no segundo conjunto de sondagem amplificada imobilizada para o substrato.

5. O método, de acordo com qualquer uma das reivindicações 1-4, caracterizado pela etapa de amplificação compreender o contato de uma exonuclease com a sondagem amplificada, digerindo o final 5' do conjunto da sondagem amplificada que não tem nenhum rótulo no final 5' em que o final 5' do conjunto de sondagem amplificado que consiste no rótulo no final 5' ser protegido contra a digestão de exonuclease.
6. O método, de acordo com qualquer uma das reivindicações 1-5, caracterizado pela variação genética determinada indicar a presença ou ausência de câncer, variabilidade farmacocinética, toxicidade da droga, rejeição do transplante ou aneuploidia no paciente.
7. O método, de acordo com qualquer uma das reivindicações 1-5, caracterizado pelo paciente ser uma paciente grávida, e a variação genética ser uma variação genética no feto da paciente grávida.
8. O método, de acordo com qualquer uma das reivindicações 1-5, caracterizado pelo paciente ser uma paciente grávida, e a variação genética ser selecionada no grupo que consiste em trissomia 13, trissomia 18, trissomia 21, aneuploidia de X e aneuploidia de Y no feto da paciente grávida.
9. O método, de acordo com qualquer uma das reivindicações 1-8, caracterizado pela amostra genética ser selecionada no grupo que consiste de todo o sangue, soro, plasma, urina, saliva, suor, matéria fecal e lágrimas.
10. O método, de acordo com qualquer uma das reivindicações 1-9, caracterizado pela primeira sondagem de rotulação e a primeira sondagem de marcação serem hibridizadas na primeira região de ácido nucleico de interesse e a segunda sondagem de rotulação e a segunda sondagem de marcação serem hibridizadas na segunda região de ácido nucleico de interesse.
11. O método, de acordo com qualquer uma das reivindicações 1-10, caracterizado pela etapa de contagem compreender filtragem espacial ou  
pela etapa de contagem compreender a análise de derramamento de água.

12. O método, de acordo com qualquer uma das reivindicações 1-11, caracterizado pela etapa de comparação compreender a obtenção da estimativa de um número relativo das moléculas de nucleotídeo tendo a primeira e a segunda regiões de ácido nucleico de interesse.

13. O método, de acordo com qualquer uma das reivindicações 1-42, caracterizado pelos rótulos imobilizados do mesmo tipo serem separados por uma distância de pelo menos 250 nm nas duas dimensões.

14. O método, de acordo com qualquer uma das reivindicações 1-13, caracterizado pelos rótulos serem corantes fluorescentes.

15. O método, de acordo com qualquer uma das reivindicações 1-14, caracterizado pelo substrato compreender um parceiro de ligação que entra em contato e imobiliza as marcas.

16. O método, de acordo com qualquer uma das reivindicações 1-15, caracterizado pela etapa de imobilização compreender a hibridização de pelo menos uma parte da marca ou sequência do nucleotídeo de marcação em uma molécula de nucleotídeo correspondente imobilizada no substrato.

17. O método, de acordo com qualquer uma das reivindicações 1-16 caracterizado pelas sondagens de marcação serem imobilizadas em um grupo de vários locais predeterminados em um substrato.

18. O método, de acordo com qualquer uma das reivindicações 1-17, caracterizado pelo paciente ser uma paciente grávida, a variação genética ser uma variação genética no feto da paciente grávida e o método ainda consistir no contato dos conjuntos de sondagem maternal e paternal com a amostra genética em que o conjunto de sondagem maternal consiste em uma sondagem de rotulação maternal e uma sondagem de marcação maternal e o conjunto de sondagem paternal consiste em uma sondagem de rotulação paternal e uma sondagem de marcação paternal, hibridização de pelo menos uma parte de cada conjunto de sondagem maternal e paternal em uma região de ácido nucleico de interesse nas moléculas de nucleotídeo da amostra genética, a região de ácido nucleico de interesse consiste em um local de Polimorfismo de Nucleotídeo Único (SNP, Single Nucleotide

Polymorphism) predeterminado em que pelo menos uma parte do conjunto de sondagem maternal é hibridizada em um primeiro alelo no local SNP, pelo menos uma parte do conjunto de sondagem paternal é hibridizada em um segundo alelo no local SNP e o primeiro e o segundo alelos são diferentes uns dos outros, ligação dos conjuntos de sondagem maternal e paternal pelo menos pela ligação (i) das sondagens de rotulação e marcação maternal e (ii) das sondagens de rotulação e marcação paternal, amplificação opcionalmente dos conjuntos de sondagem ligados, imobilização das sondagens de marcação e/ou as sondas de marcação amplificadas das mesmas em um local predeterminado em um substrato em que as sondagens de rotulação maternal e paternal e/ou suas sondagens de rotulação amplificada ligadas às sondagens de marcação imobilizadas e/ou as sondas de marcação amplificadas das mesmas que compreendem nos rótulos maternal e paternal, respectivamente, os rótulos maternal e paternal são diferentes, e os rótulos imobilizados podem ser opticamente determinados, contando-se os números dos rótulos maternal e paternal e determinando-se se uma proporção de um material fetal na amostra genética é suficiente para detectar a variação genética no feto com base nos números de rótulos maternal e paternal.

19. O método, de acordo com qualquer uma das reivindicações 1-18, caracterizado pelo paciente ser uma paciente grávida, a variação genética ser uma variação genética no feto da paciente grávida e o método ainda consistir no contato dos conjuntos de sondagem de alelo A e B que são específicos de alelo com a amostra genética em que o conjunto de sondagem de alelo A consiste em uma sondagem de rotulação de alelo A e uma sondagem de marcação de alelo A e o conjunto de sondagem de alelo B consiste em uma sondagem de rotulação de alelo B e uma sondagem de marcação de alelo B, hibridização de pelo menos uma parte de cada conjunto de sondagem de alelo A e alelo B em uma região de ácido nucleico de interesse nas moléculas de nucleotídeo da amostra genética, a região de ácido nucleico de interesse consiste em um local de Polimorfismo de Nucleotídeo Único (SNP,

Single Nucleotide Polymorphism) para o qual um perfil alélico maternal difere de um perfil alélico fetal no local SNP e o primeiro e o segundo alelos são diferentes uns dos outros, ligando-se os conjuntos de sondagem do alelo A e B pelo menos pela ligação (i) das sondagens de rotulação e marcação de alelo A e (ii) das sondagens de rotulação e marcação de alelo B, amplificação opcionalmente dos conjuntos de sondagem ligados, imobilização das sondagens de marcação e/ou as sondas de marcação amplificadas das mesmas em um local predeterminado em um substrato em que as sondagens de rotulação de alelo A e alelo B e/ou suas sondagens de rotulação amplificada ligadas às sondagens de marcação imobilizadas e/ou as sondas de marcação amplificadas das mesmas que compreendem os rótulos de alelo A e alelo B, respectivamente, os rótulos de alelo A e alelo B são diferentes, e os rótulos imobilizados podem ser opticamente determinados, contando-se os números dos rótulos de alelo A e alelo B e determinando-se se uma proporção de um material fetal na amostra genética é suficiente para detectar a variação genética no feto com base nos números de rótulos de alelo A e alelo B.

20. O método, de acordo com qualquer uma das reivindicações 1-19, caracterizado por excluir o sequenciamento das regiões de ácido nucleico de interesse e/ou sondagens.

21. O método, de acordo com qualquer uma das reivindicações 1-20, caracterizado por excluir a leitura do grupo de volume ou a quantificação análoga.

22. O método, de acordo com qualquer uma das reivindicações 1-21, caracterizado pela etapa de contagem compreender a medição de sinais ópticos dos rótulos imobilizados e na calibração dos números contados distinguindo-se um sinal óptico de um único rótulo do restante dos sinais ópticos de fundo e/ou vários rótulos.

23. O método, de acordo com qualquer uma das reivindicações 1-22, caracterizado pela etapa de contagem ainda consistir na confirmação da contagem na qual a confirmação consiste no contato do primeiro e segundo conjuntos de sondagem de controle

com a amostra genética, em que o primeiro conjunto de sondagem de controle consiste em uma primeira sondagem de rotulação e na primeira sondagem de marcação e o segundo conjunto de sondagem de controle consiste em uma segunda sondagem de rotulação de controle e na segunda sondagem de marcação, hibridização de pelo menos uma parte do primeiro e segundo conjuntos de sondagem de controle na primeira e segunda regiões de ácido nucleico de interesse nas moléculas de nucleotídeo da amostra genética, respectivamente, ligação do primeiro conjunto de sondagem de controle pelo menos pela ligação da primeira sondagem de rotulação de controle e primeira sondagem de marcação, ligação do segundo conjunto de sondagem de controle pelo menos por ligação da segunda sondagem de rotulação de controle e segunda sondagem de marcação, amplificação opcionalmente dos conjuntos de sondagem ligados, imobilização de cada uma das sondagens de marcação e/ou as sondas de marcação amplificadas das mesmas a um local predeterminado em um substrato em que a primeira e a segunda sondagens de rotulação de controle e/ou suas sondagens de rotulação amplificadas ligadas às sondagens de marcação imobilizadas e/ou as sondas de marcação amplificadas das mesmas que compreendem o primeiro e segundo rótulos de controle, respectivamente, o primeiro e segundo rótulos de controle são diferentes e os rótulos imobilizados podem ser determinados opticamente, medindo-se os sinais ópticos dos rótulos de controle imobilizados para o substrato e comparando-se os sinais ópticos do primeiro e segundo rótulos de controle imobilizado com os sinais ópticos do primeiro e segundo rótulos imobilizados para determinar se existe um erro baseado nos rótulos.

24. O método, de acordo com a reivindicação 23, caracterizado pelo primeiro rótulo e o segundo rótulo de controle serem iguais e o segundo rótulo e o primeiro rótulo de controle serem iguais.

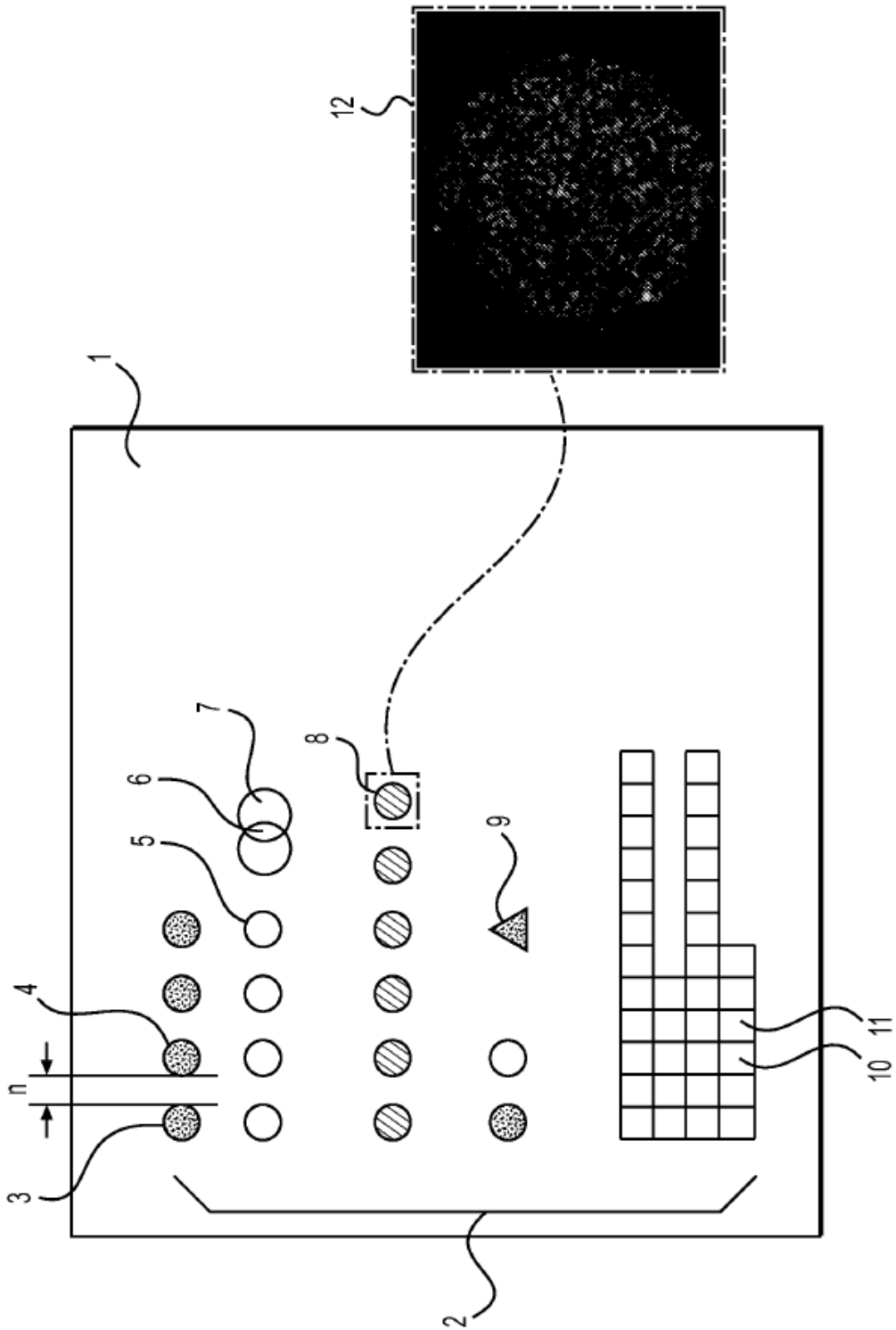


Figura 1

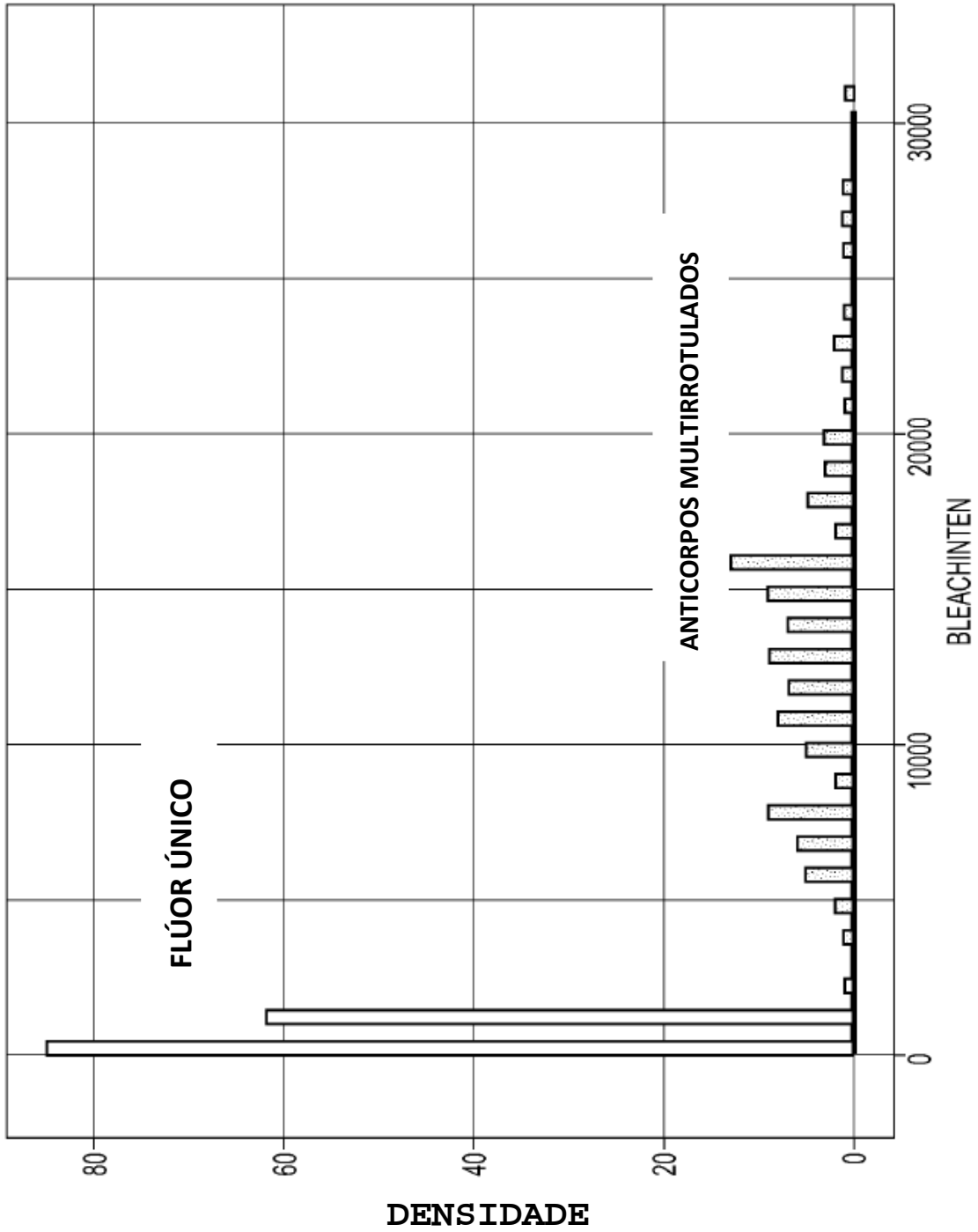


Figura 2

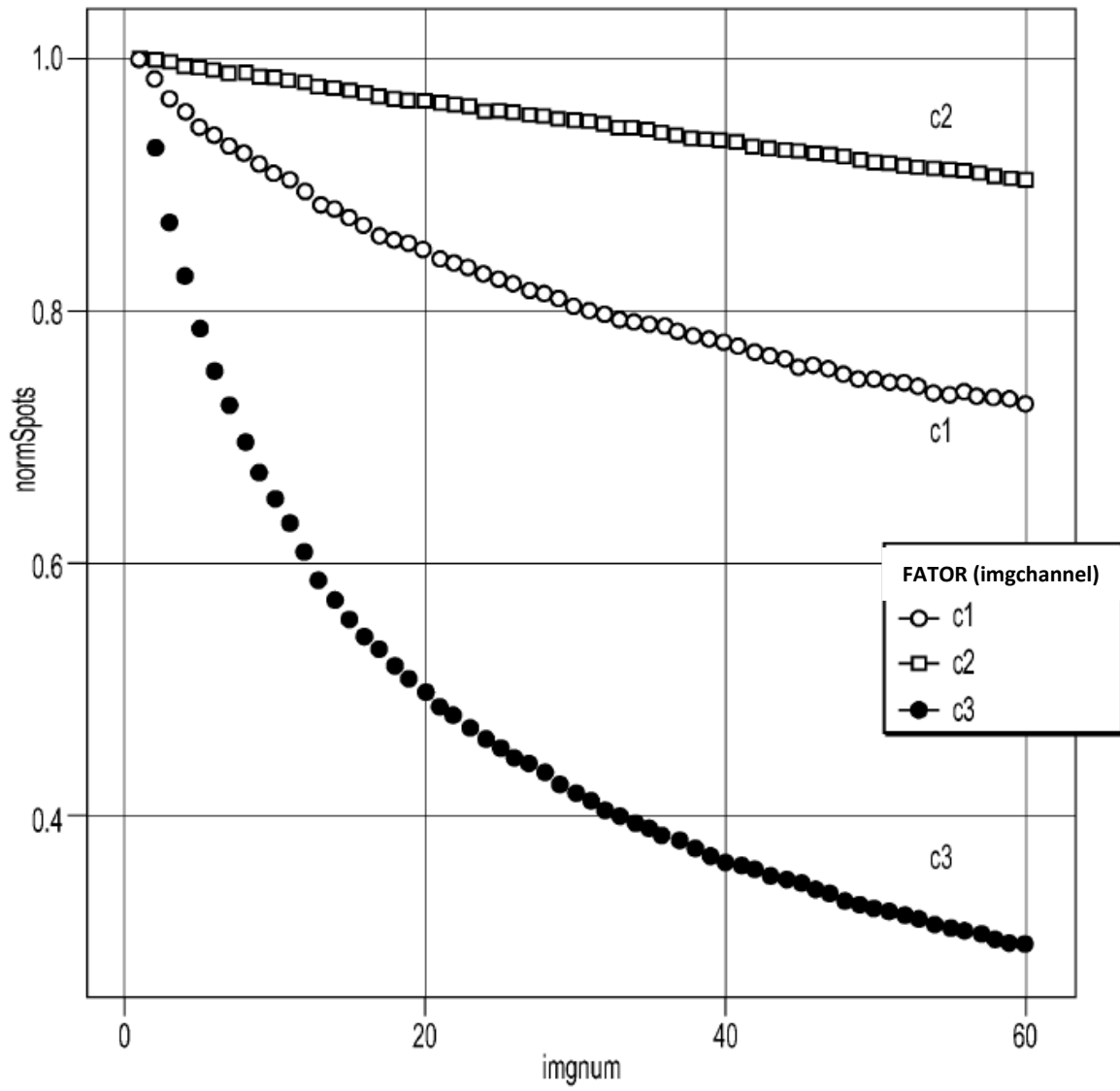


Figura 3

Figura 4

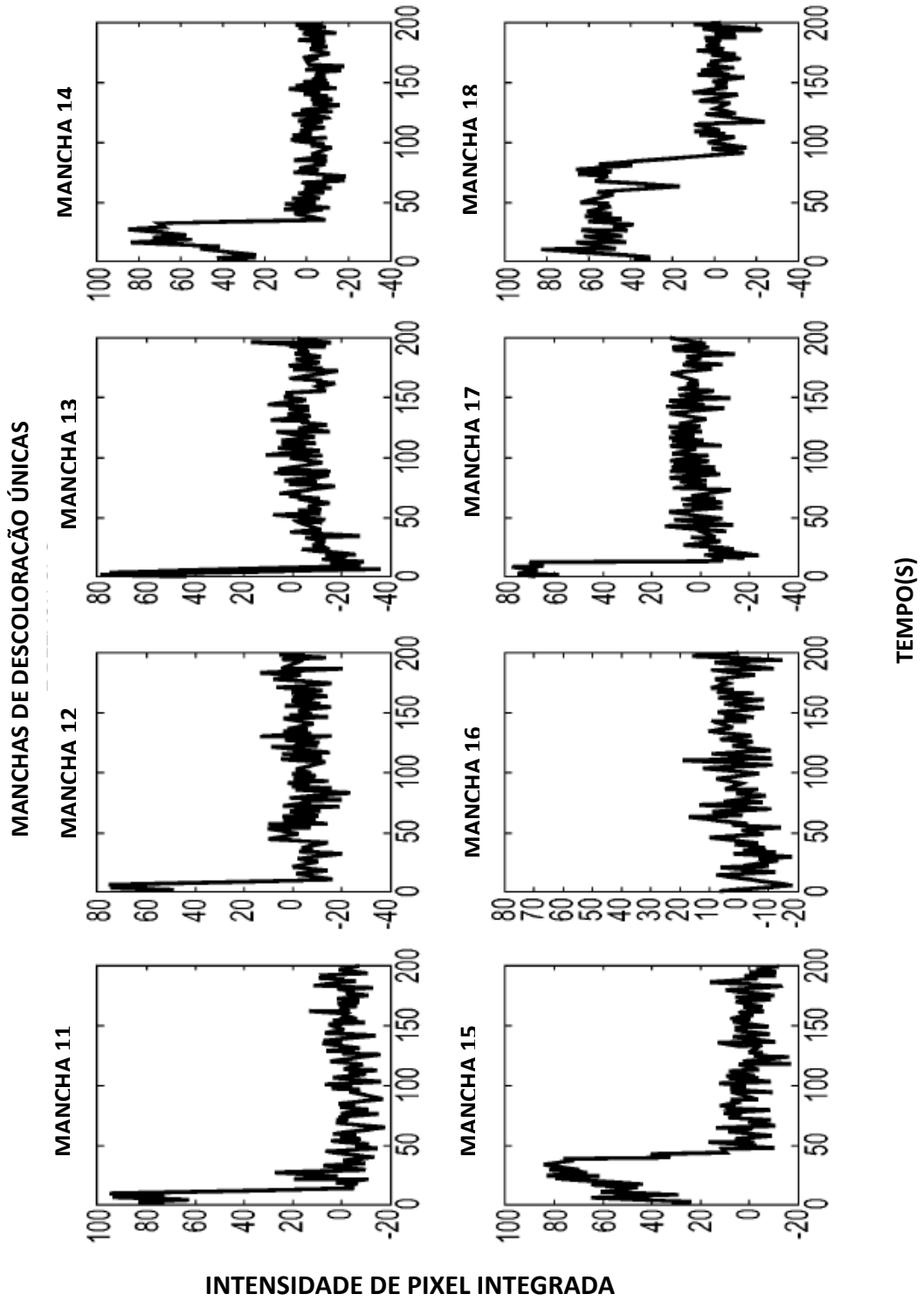


Figura 5

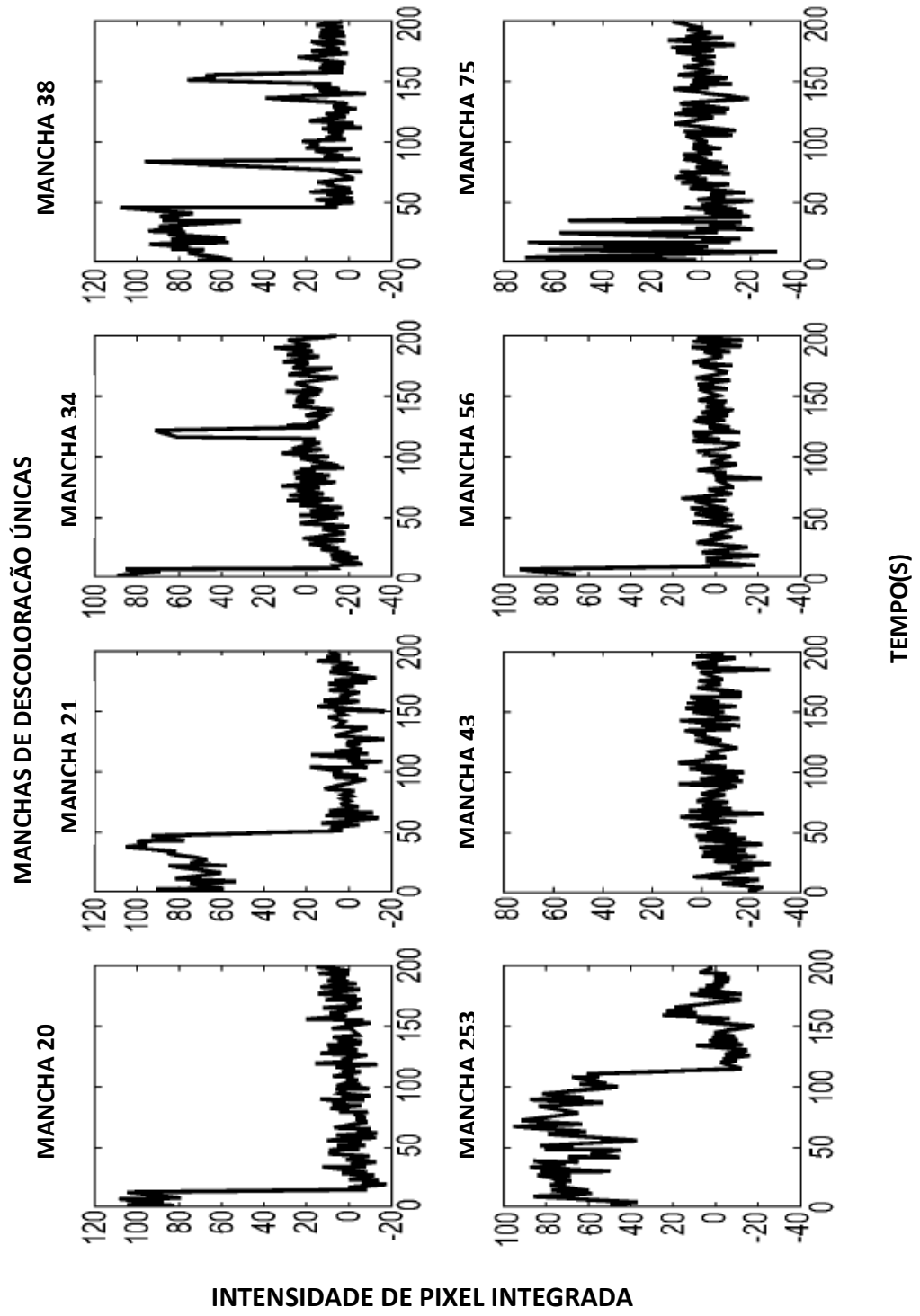


Figura 6

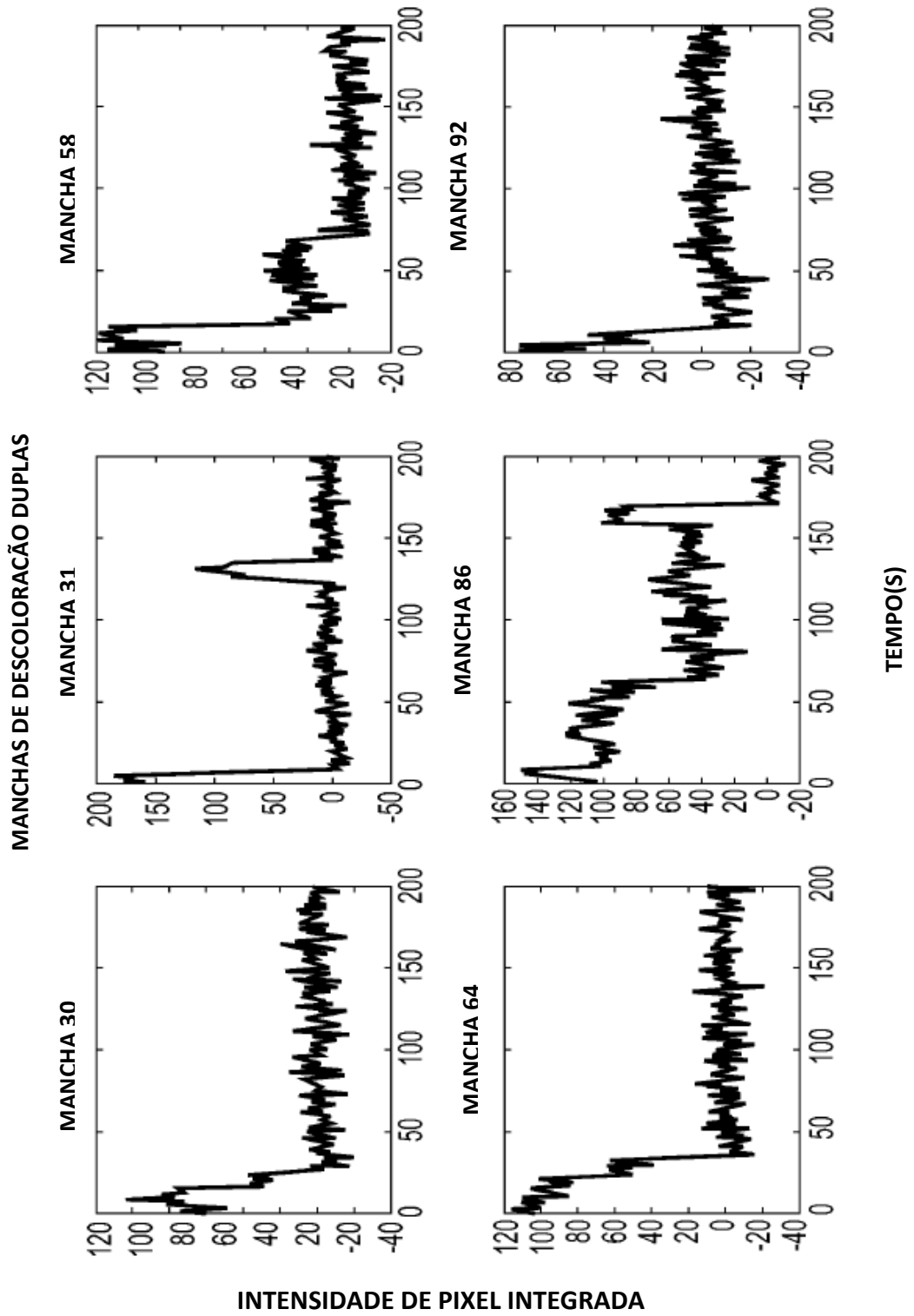


Figura 7

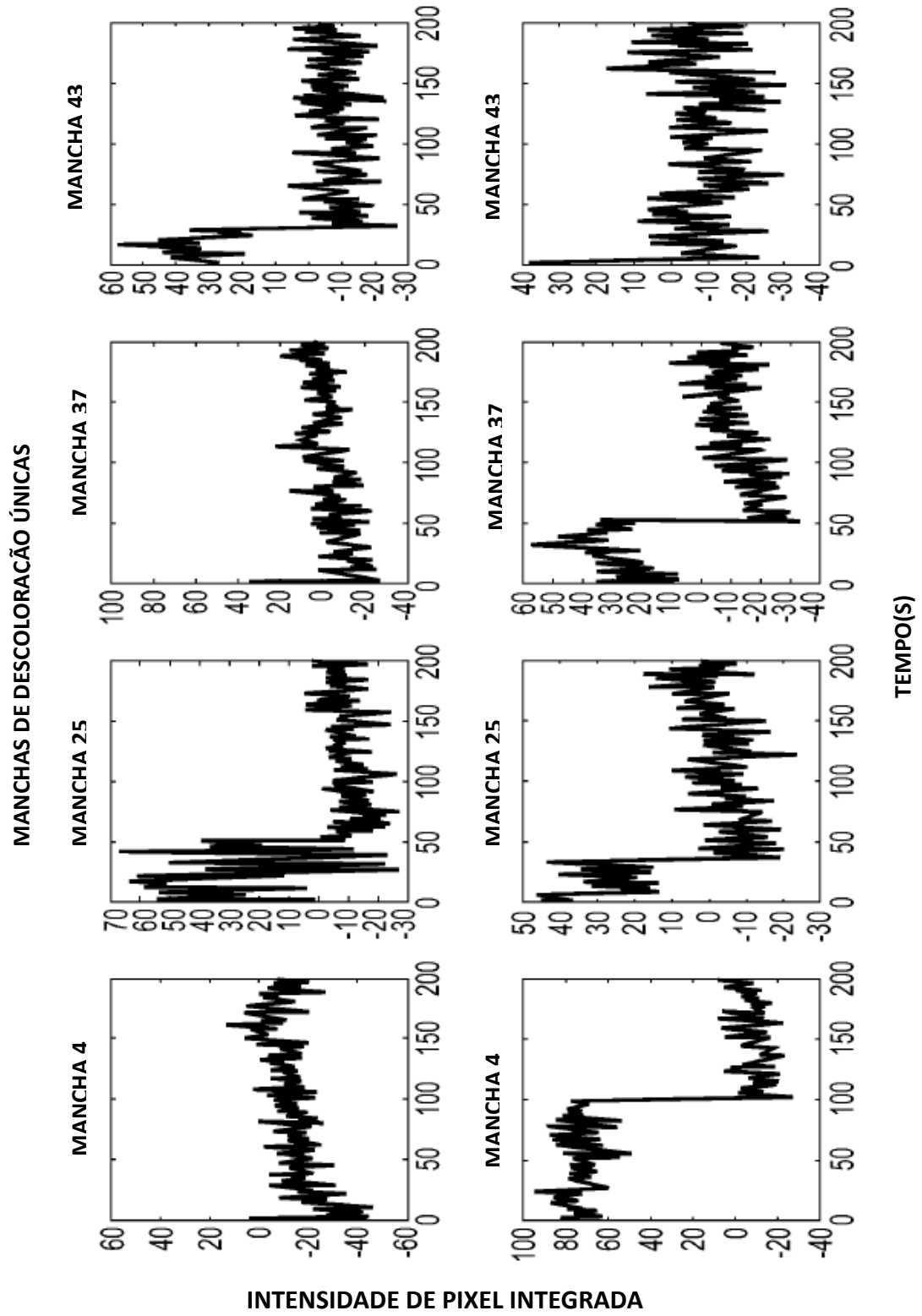


Figura 8

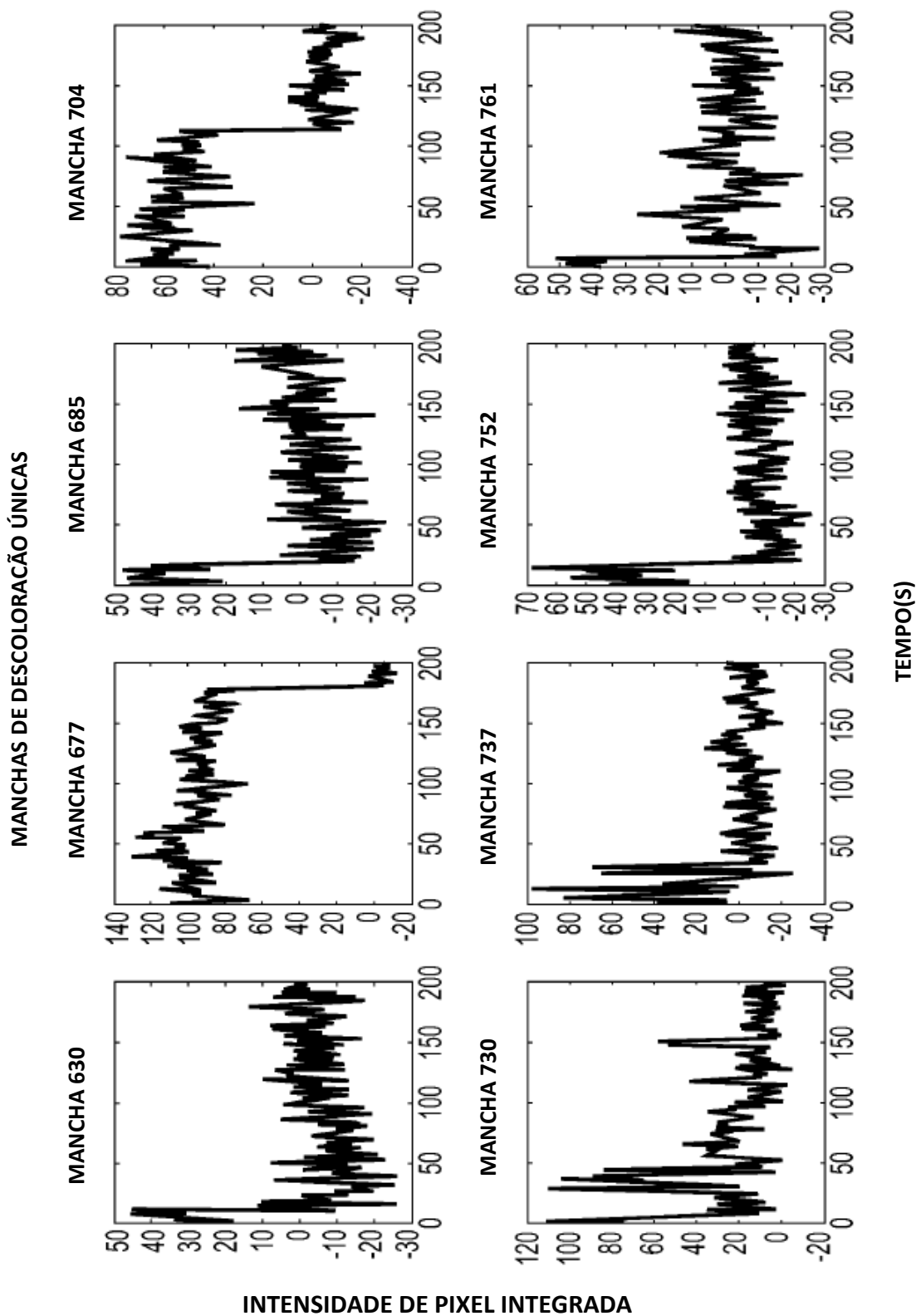


Figura 9

MANCHAS DE DESCOLORAÇÃO DUPLAS

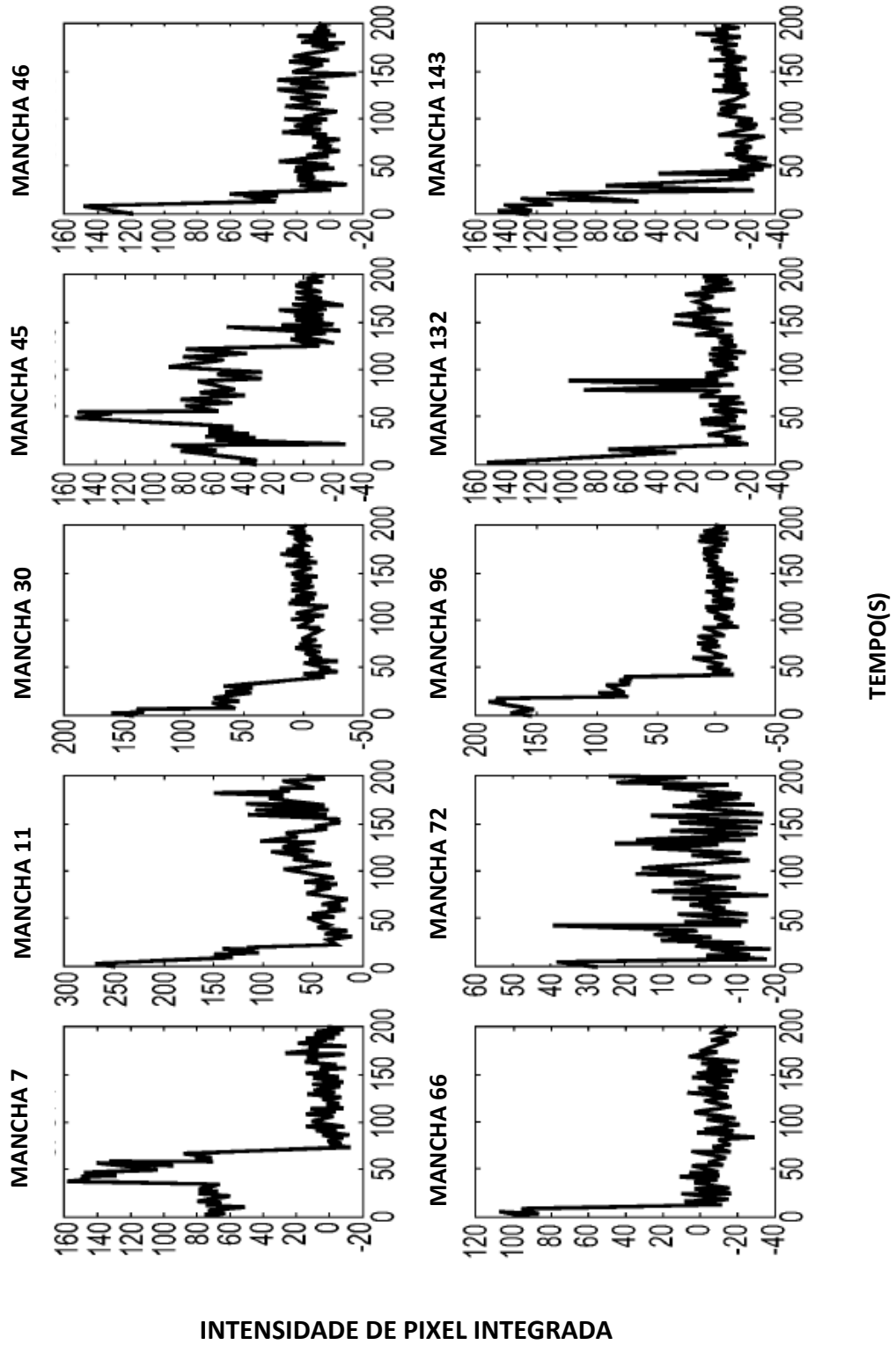


Figura 10

MANCHAS DE DESCOLORAÇÃO DUPLAS

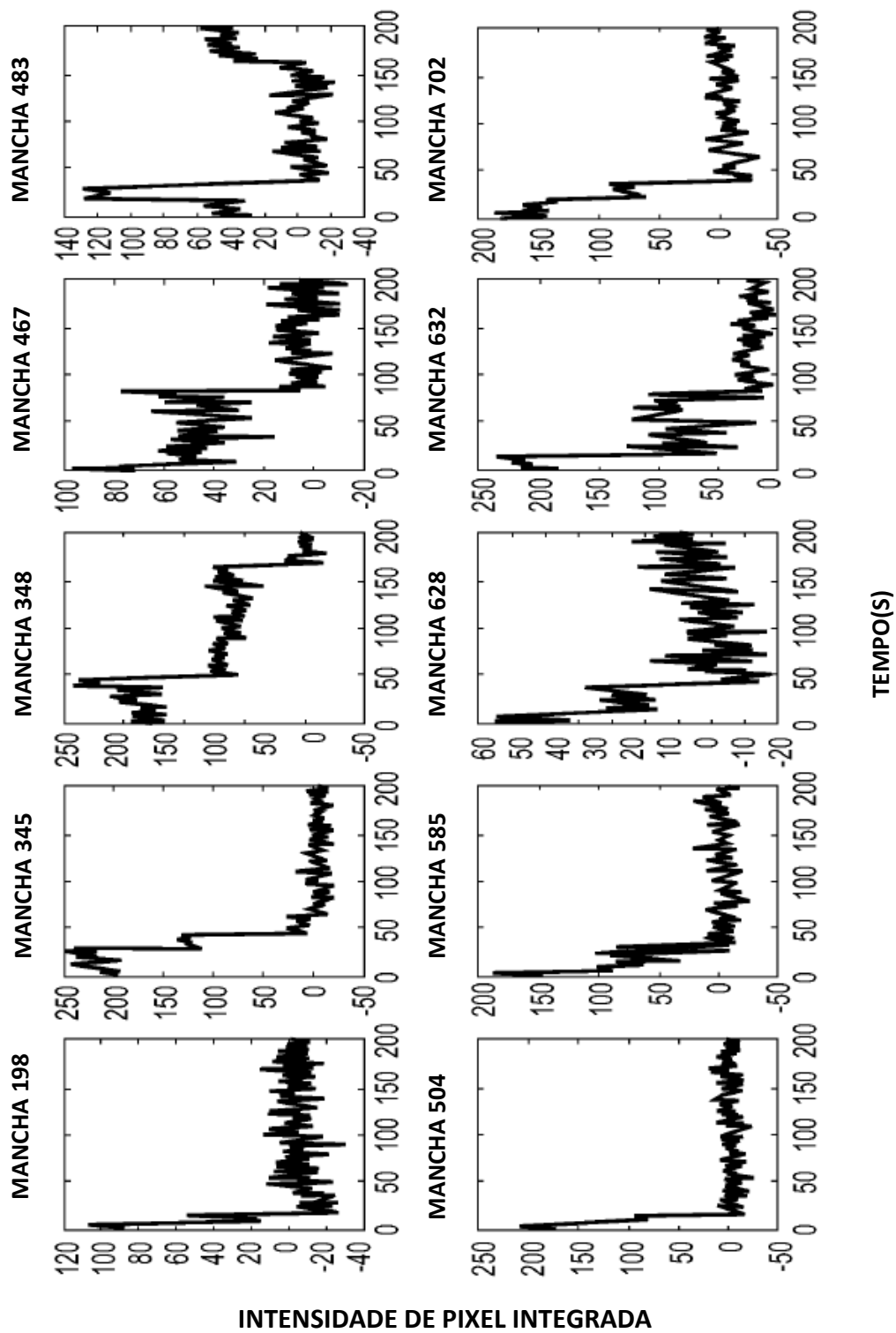


Figura 11

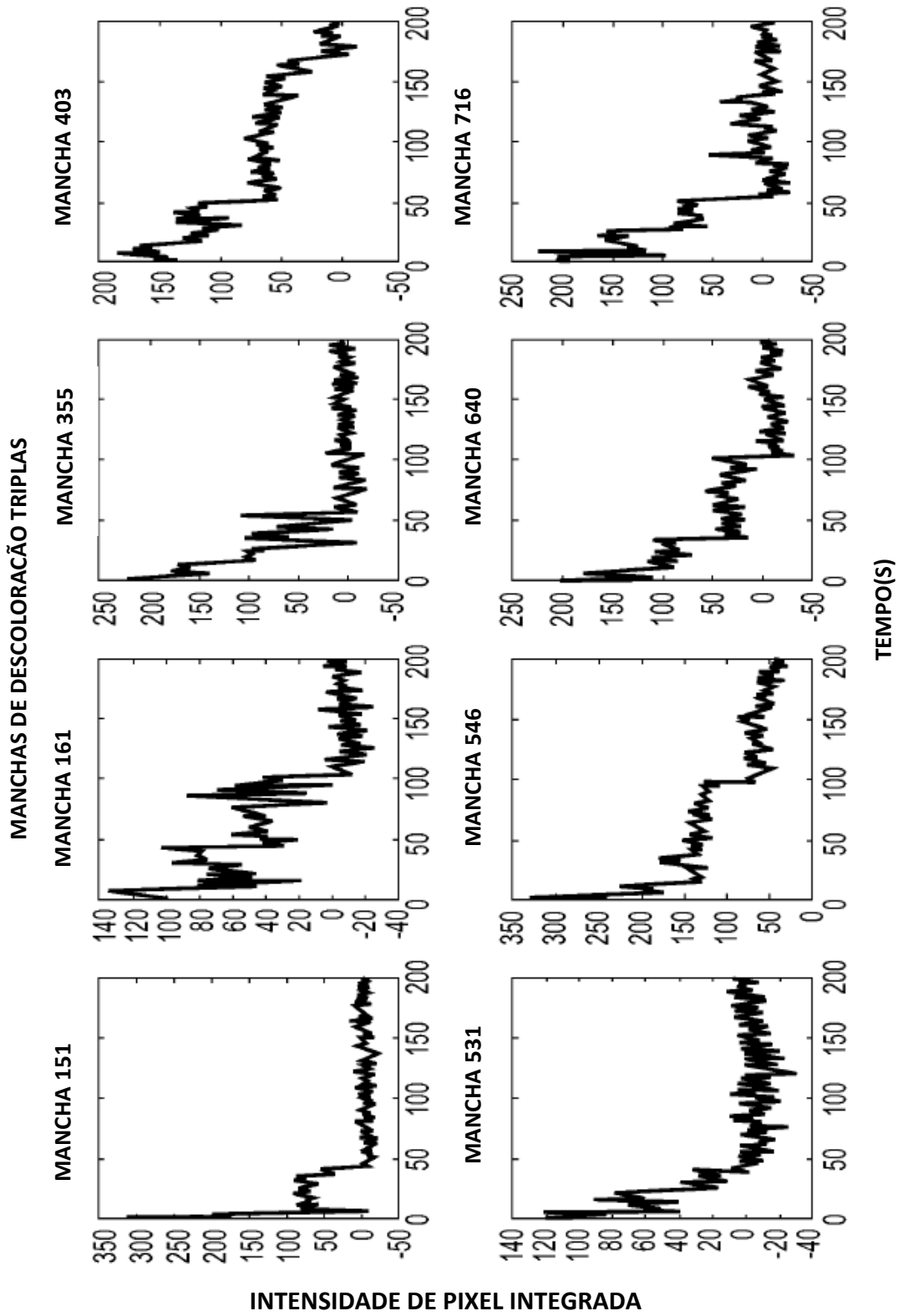


Figura 12

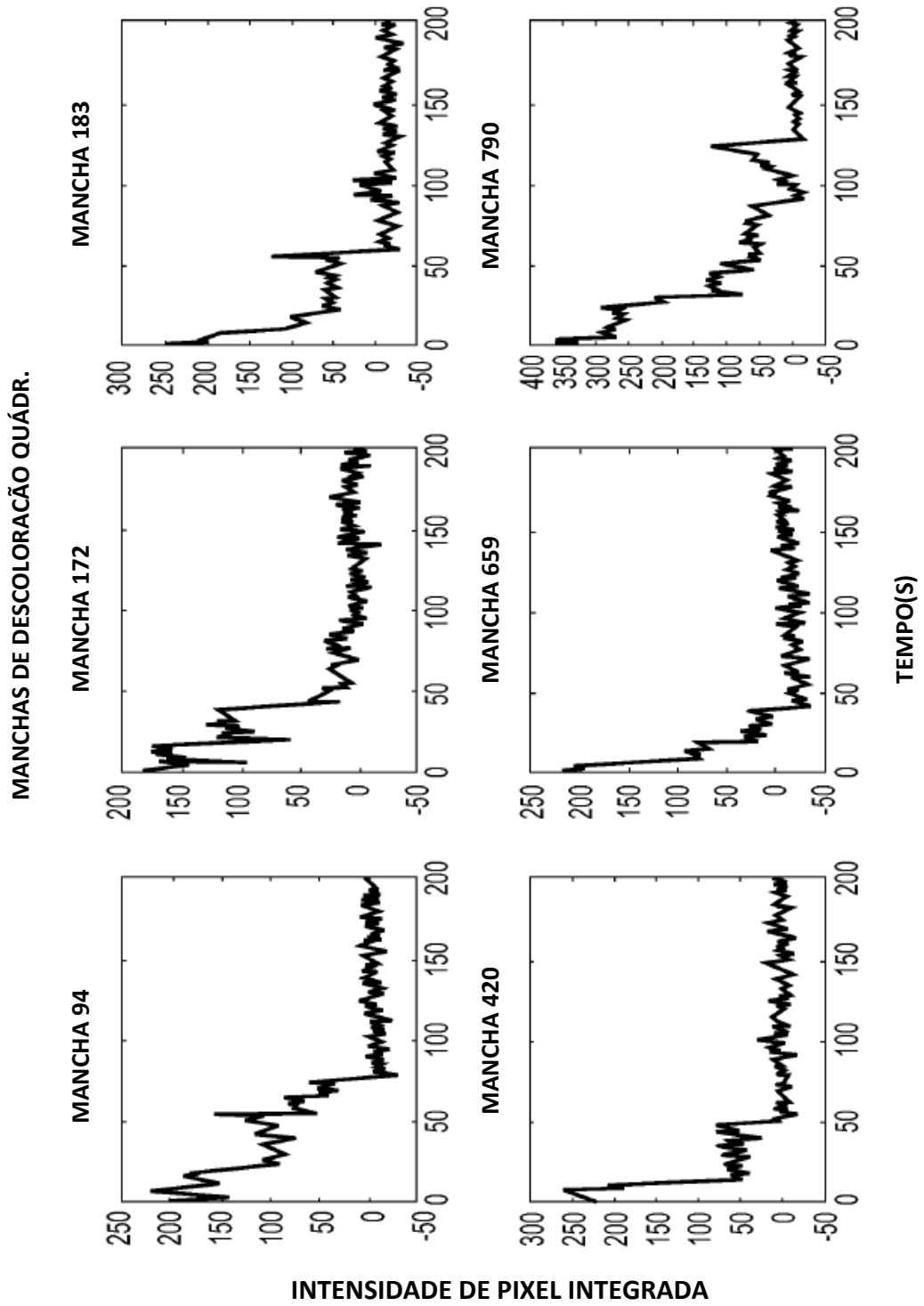
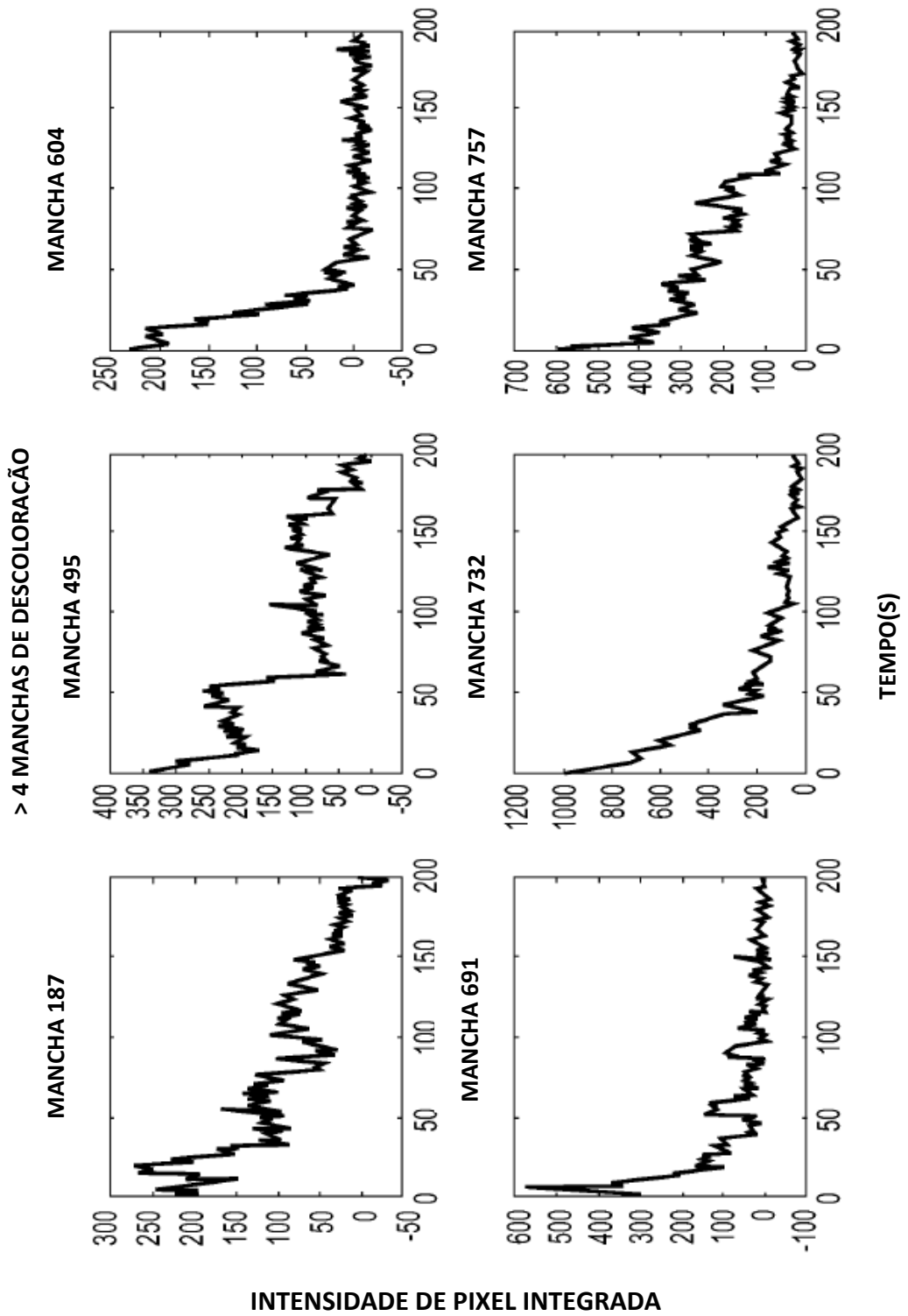


Figura 13



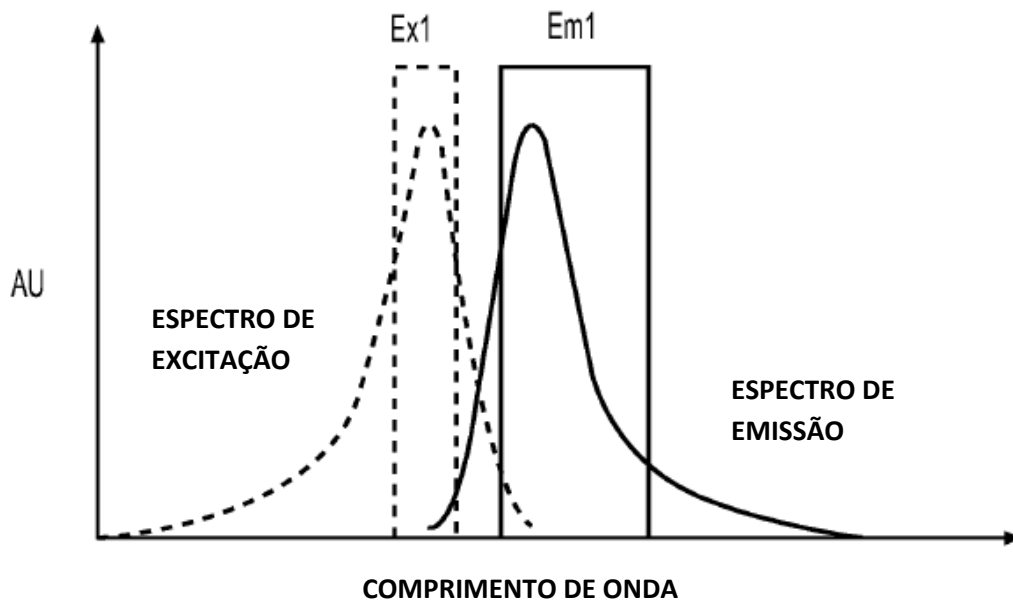


Figura 14

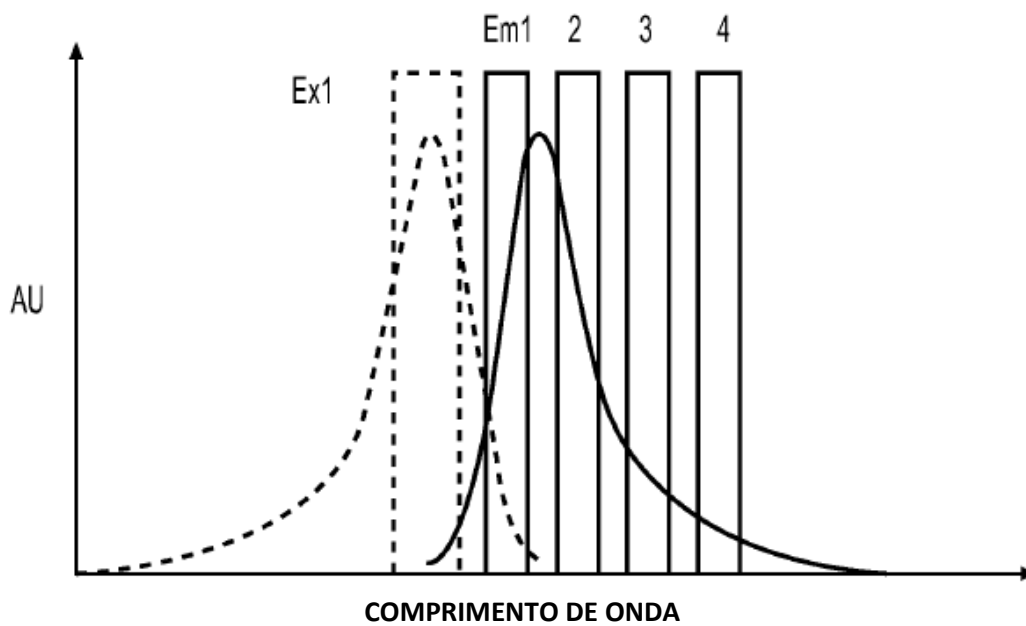


Figura 15

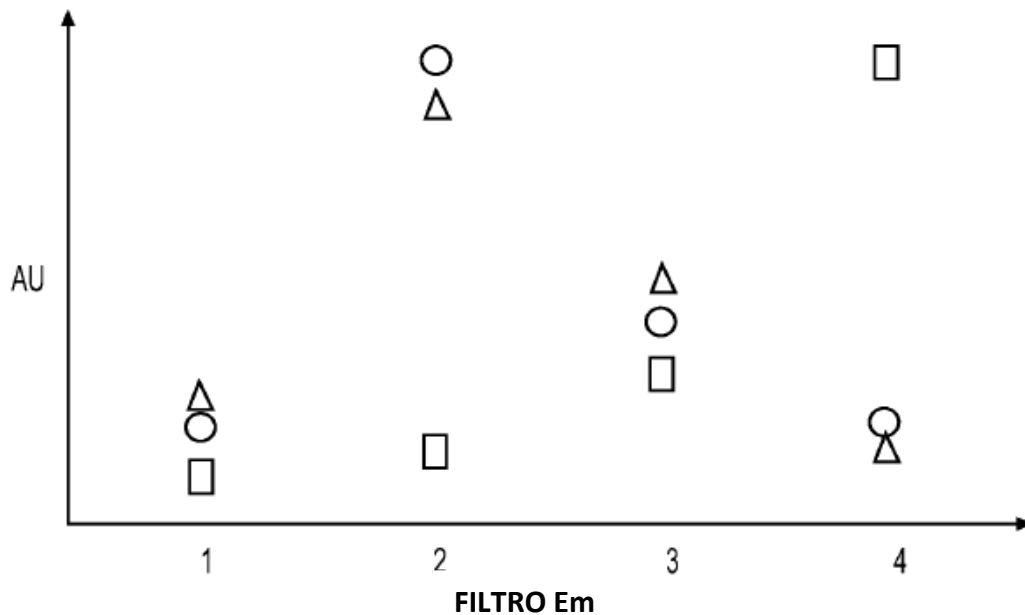


Figura 16

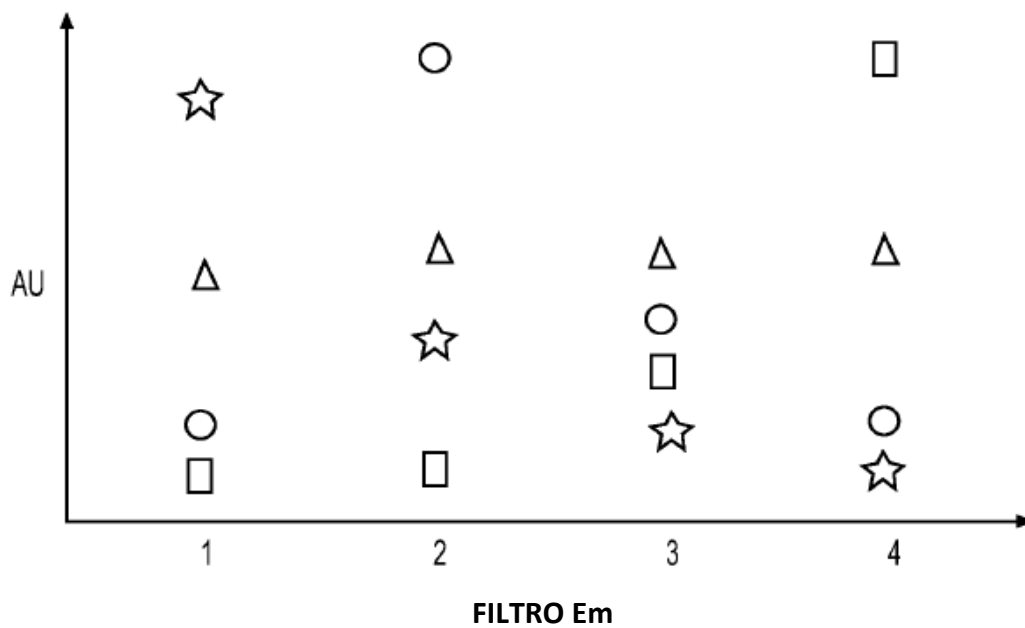


Figura 17

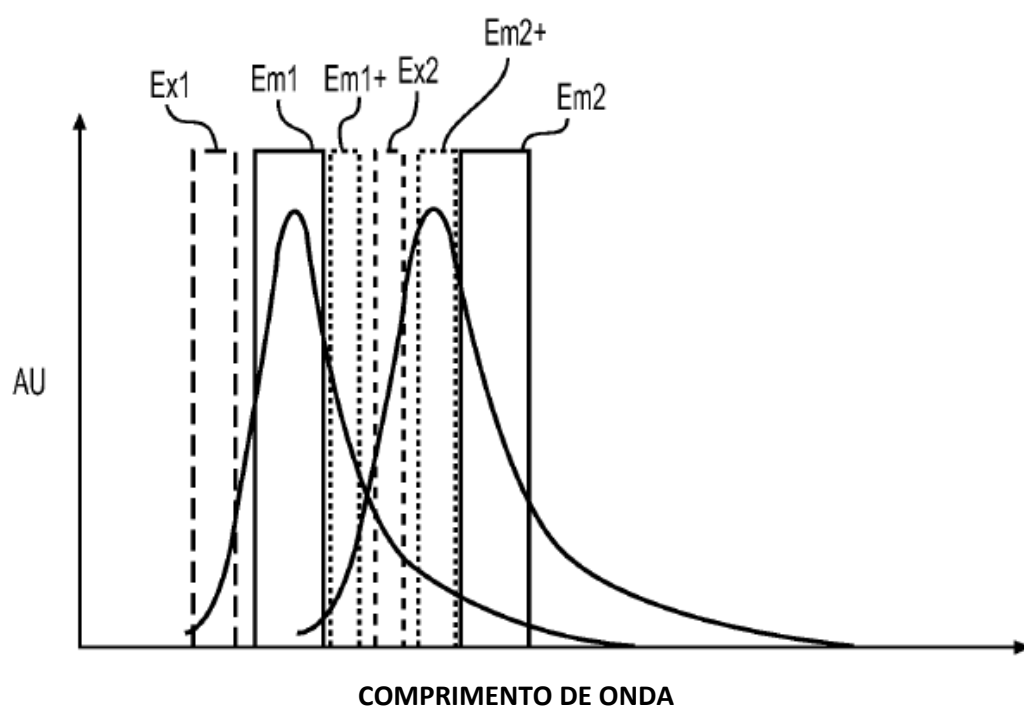


Figura 18

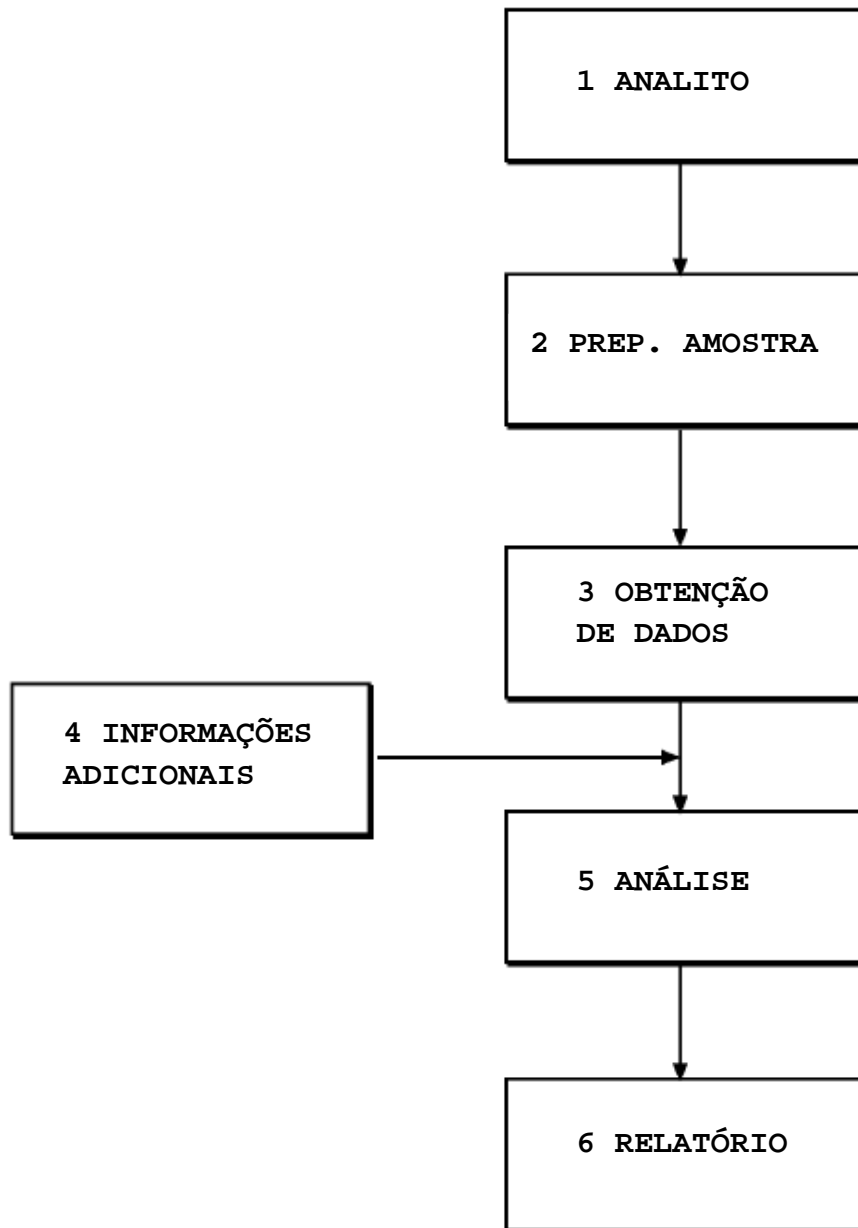


Figura 19

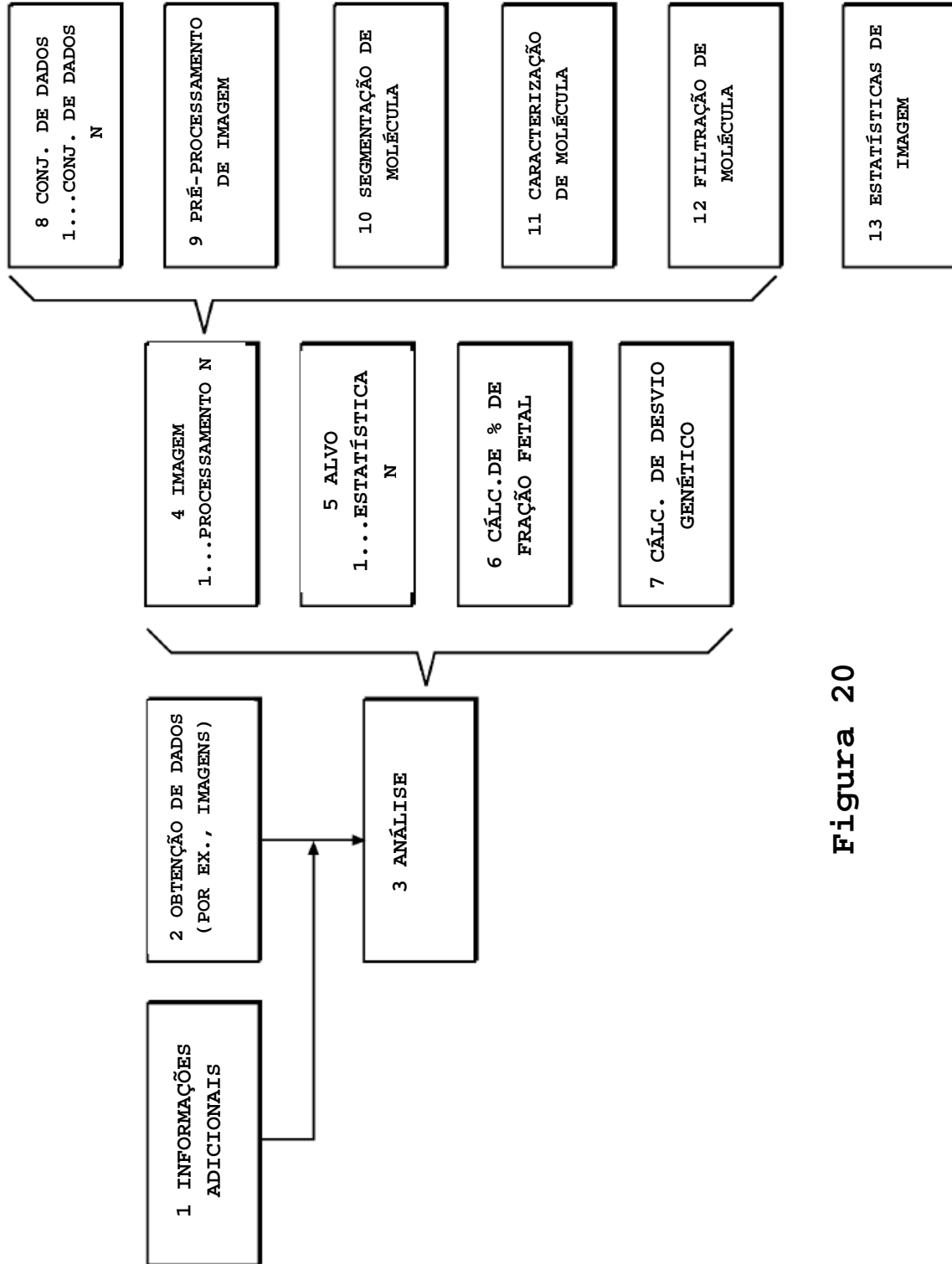


Figura 20

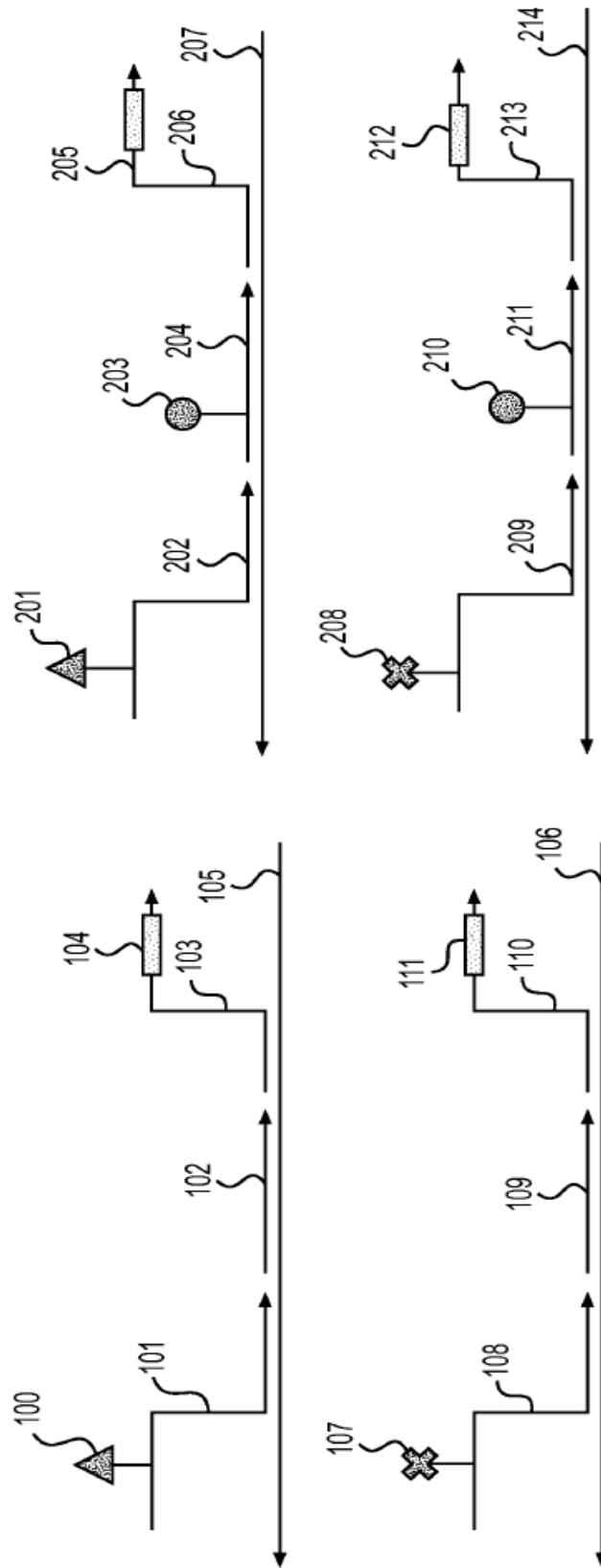


Figura 21

Figura 22

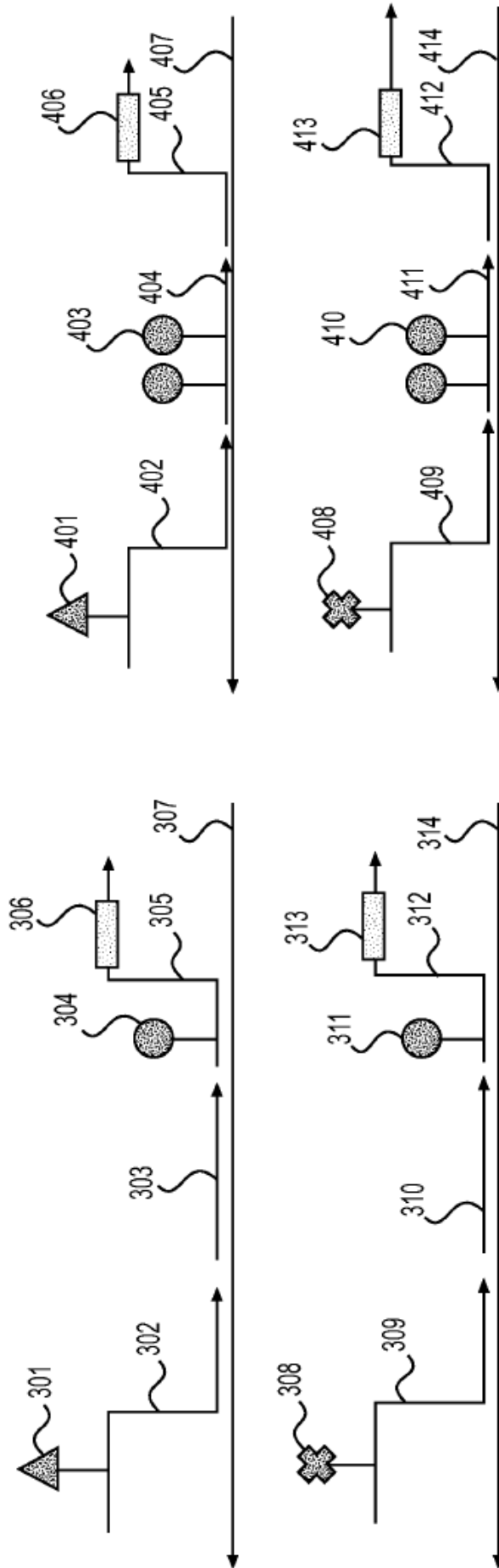


Figura 24

Figura 23

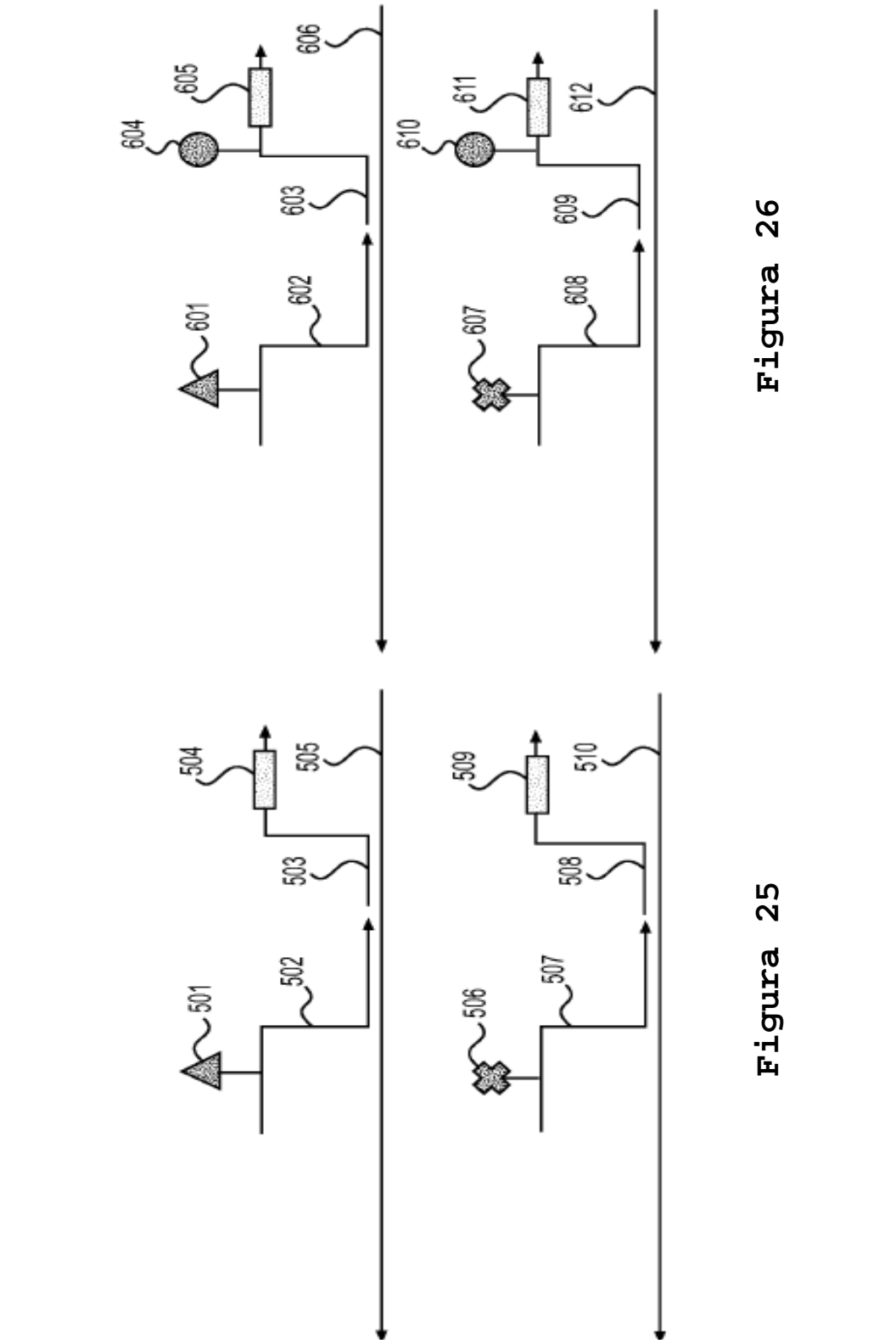


Figura 26

Figura 25

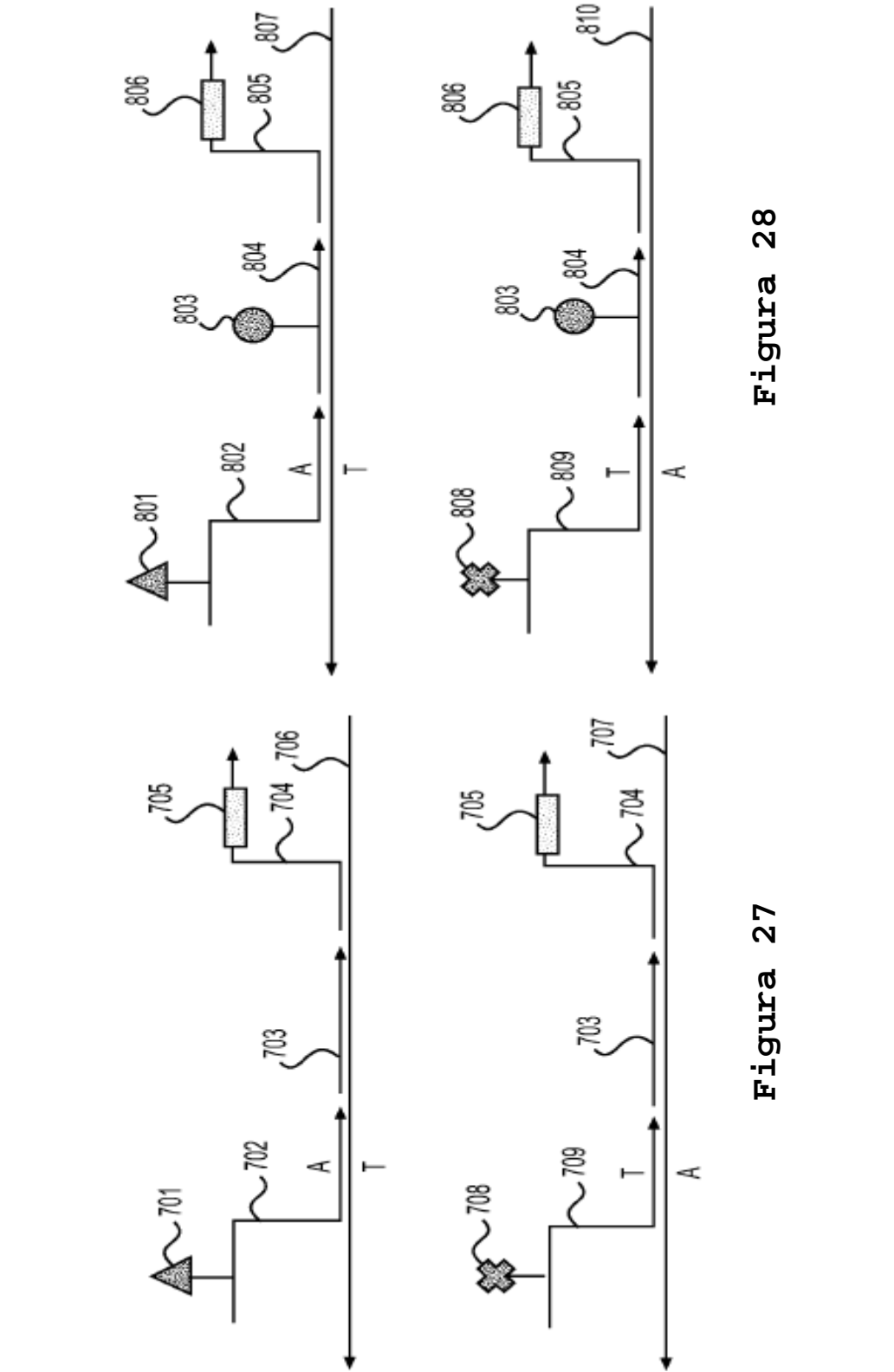


Figura 28

Figura 27

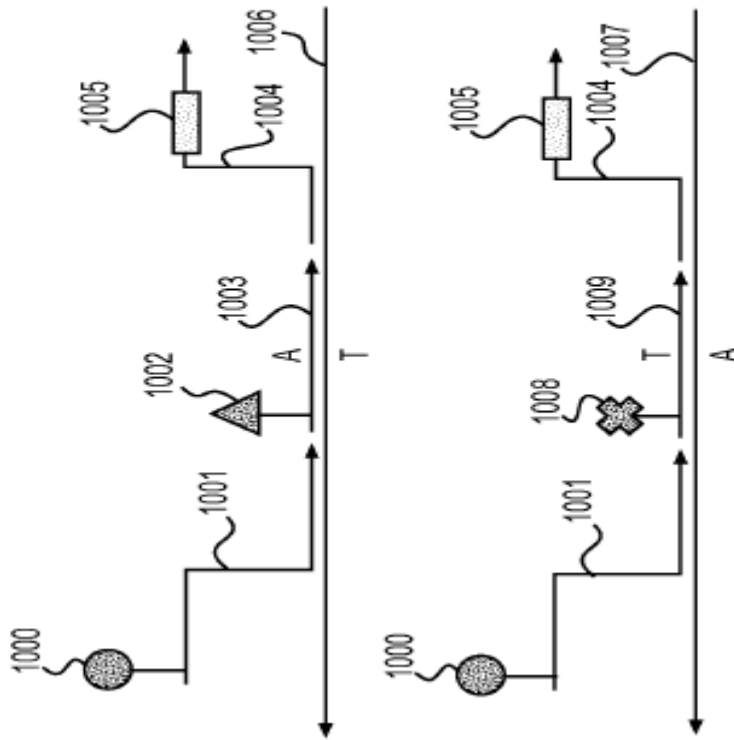


Figura 29

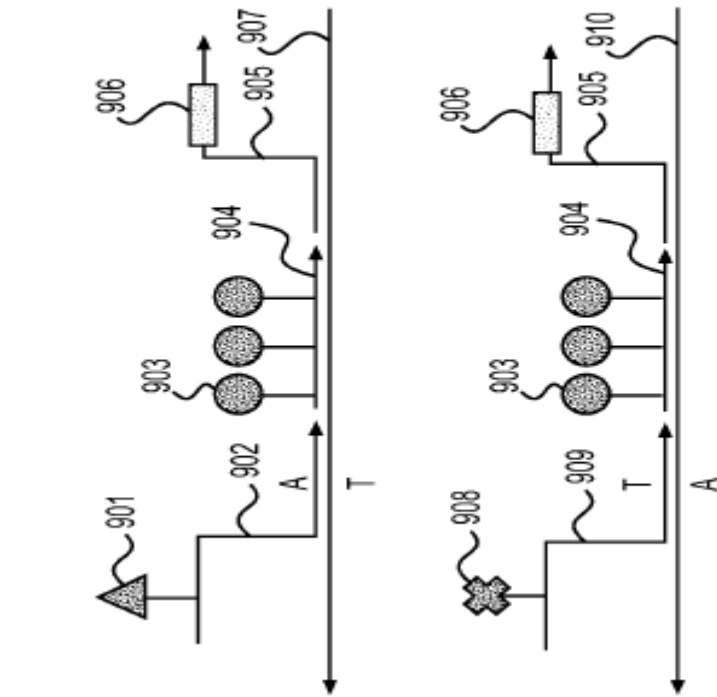


Figura 30

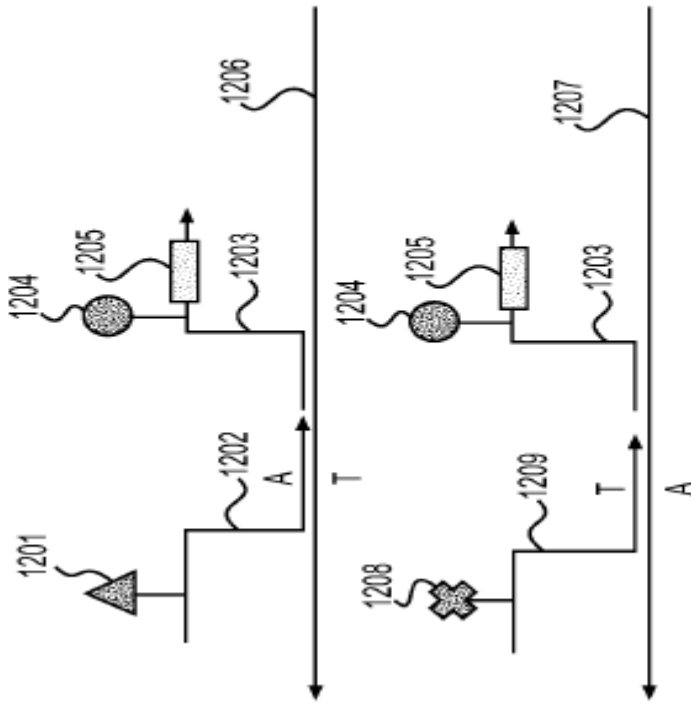


Figura 32

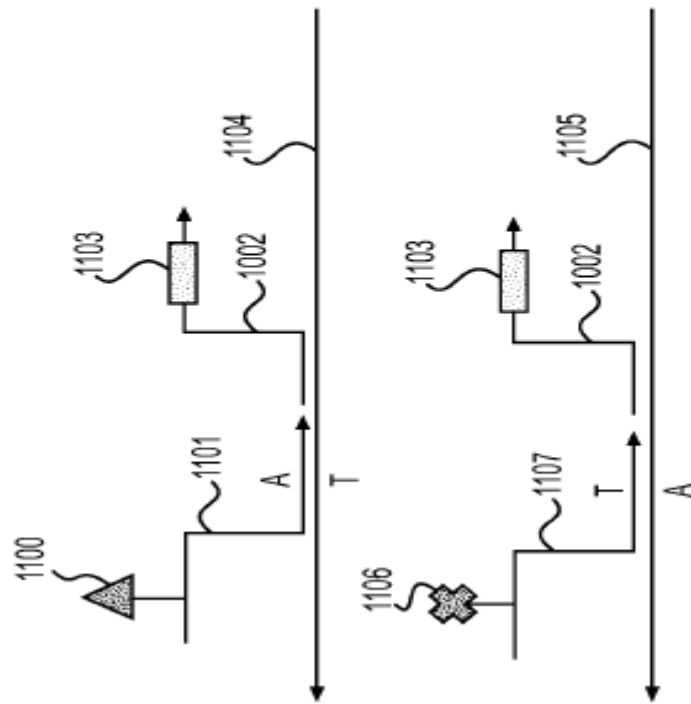


Figura 31

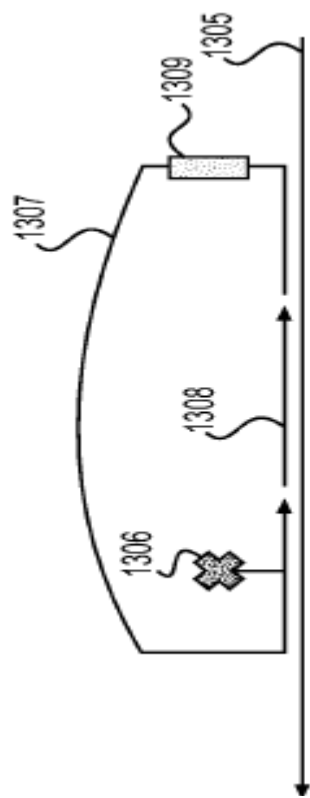
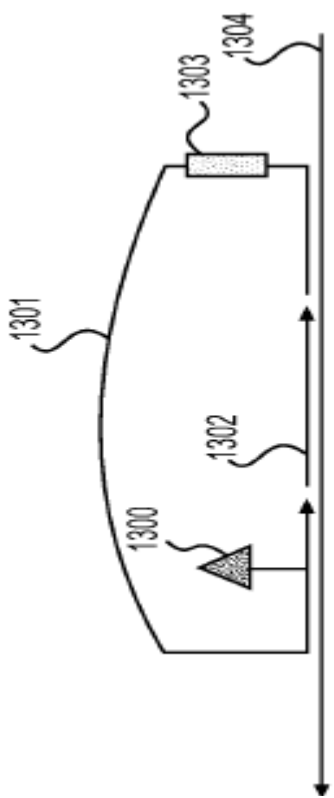


Figura 33

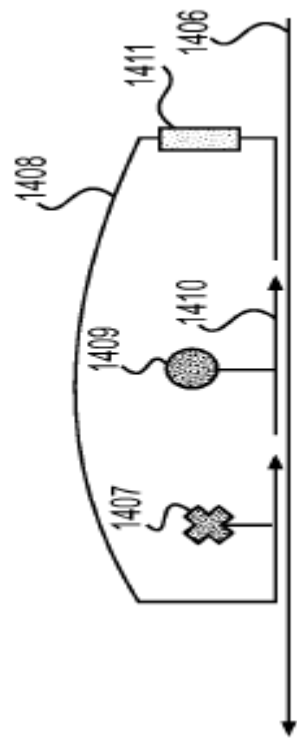
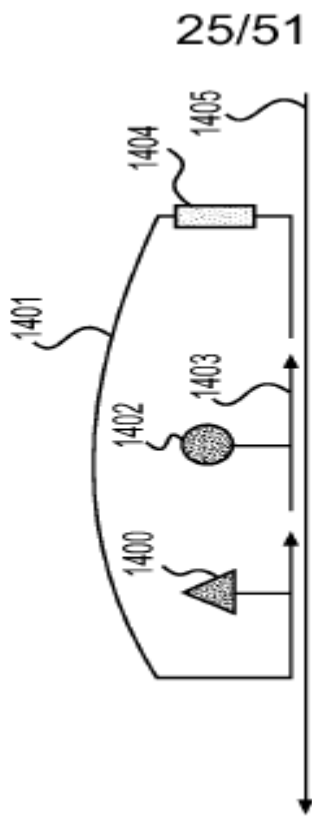


Figura 34

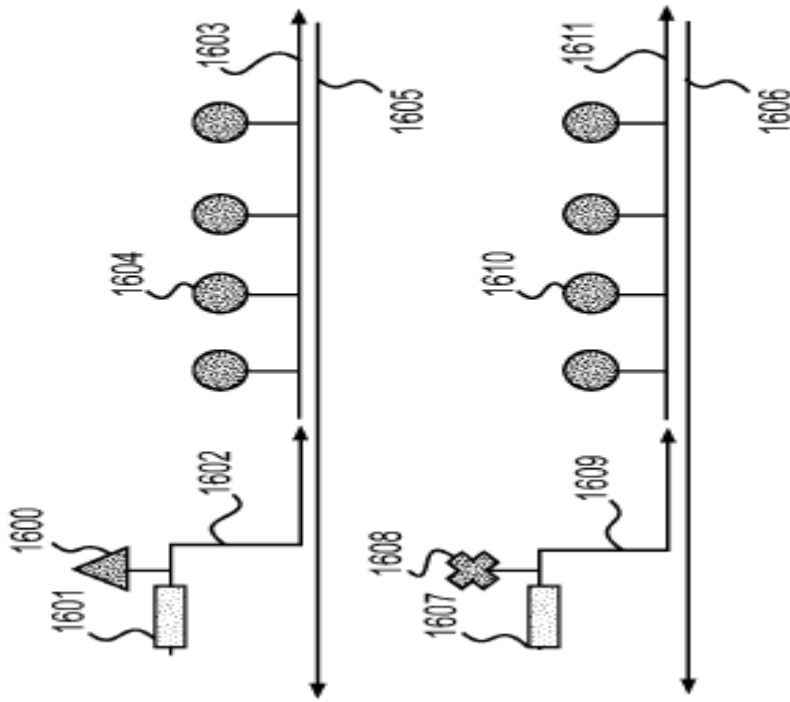


Figura 36

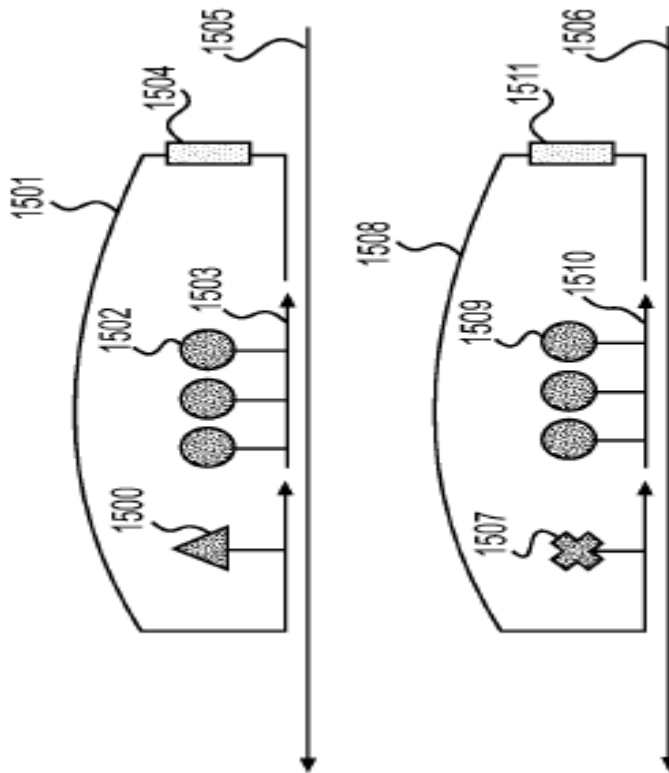


Figura 35

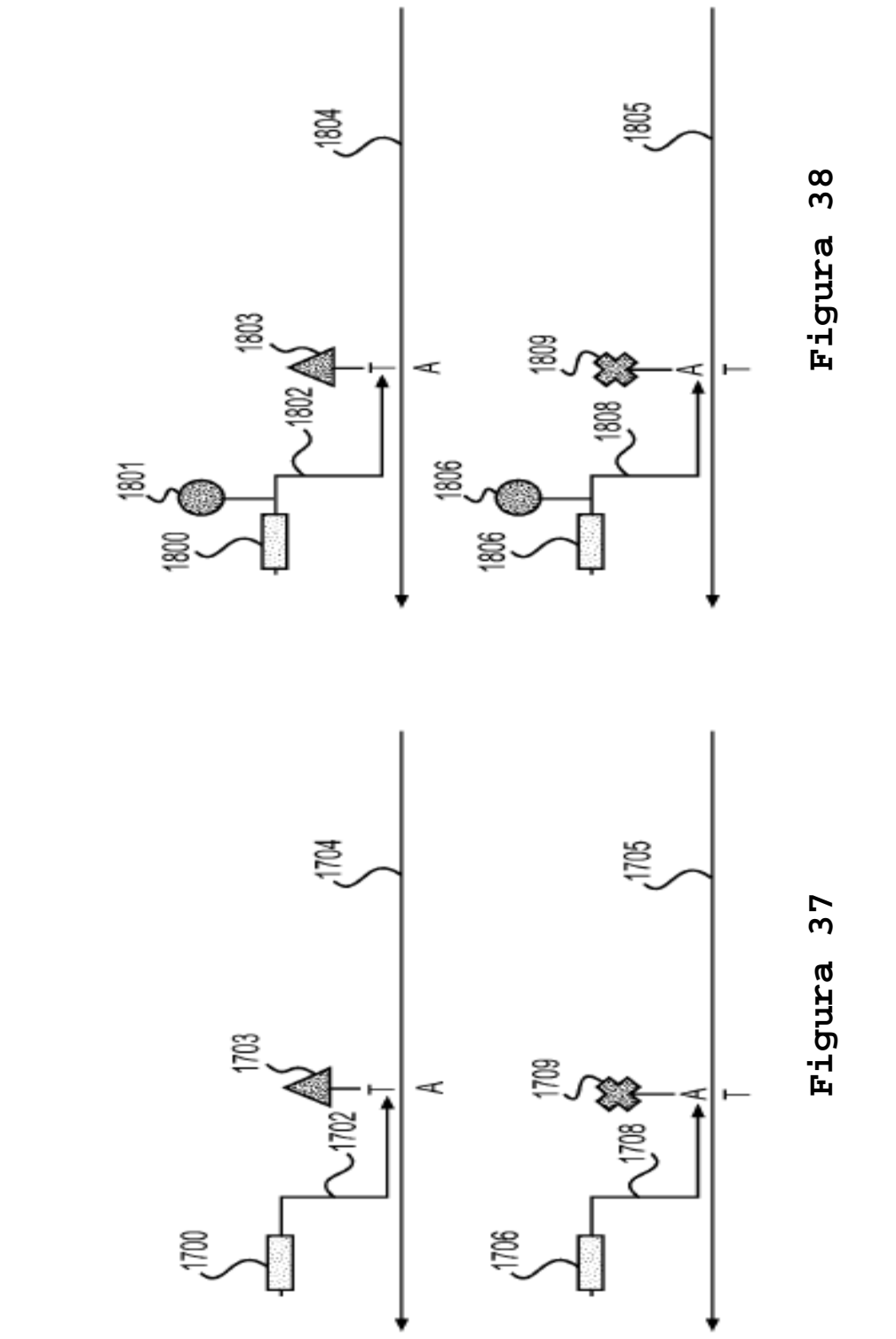


Figura 38

Figura 37

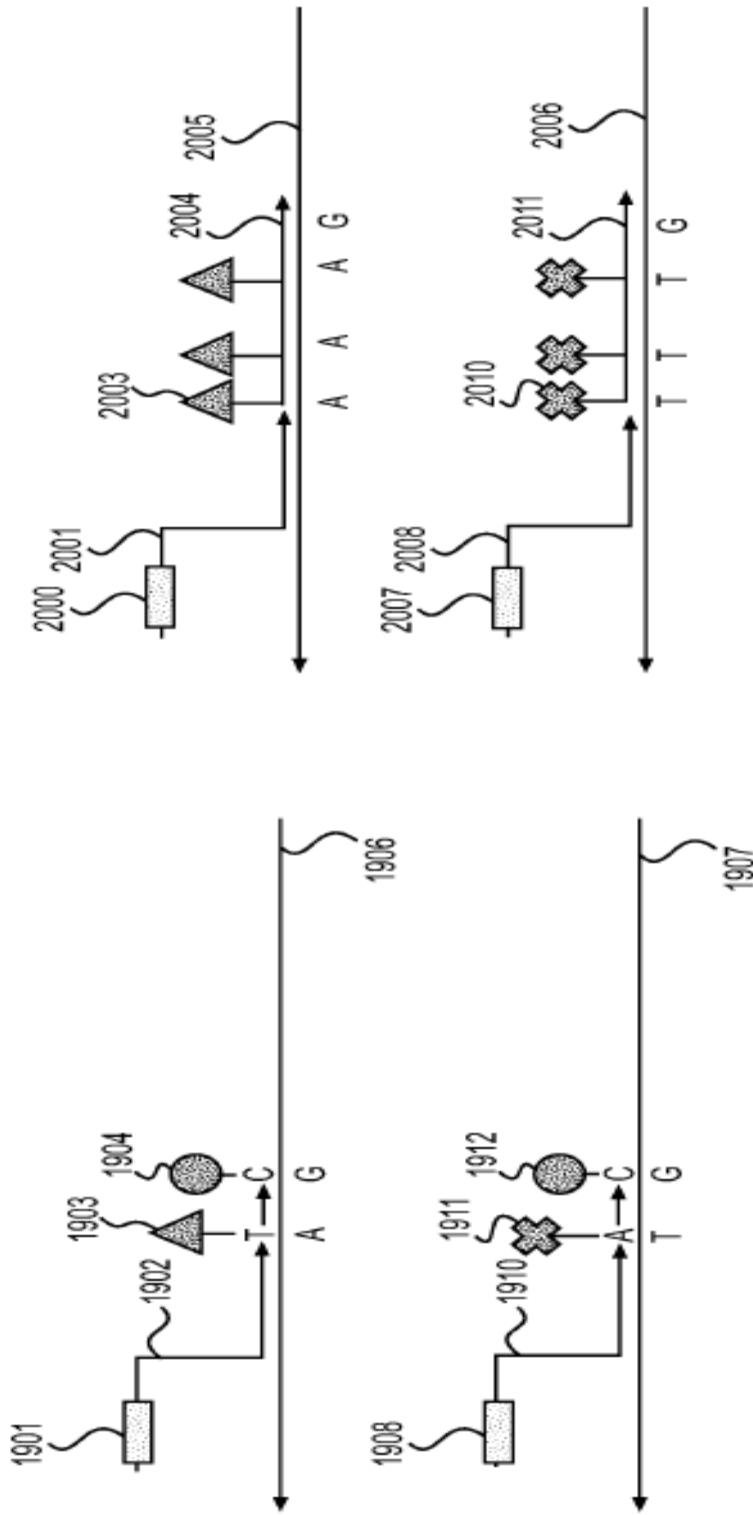


Figura 39

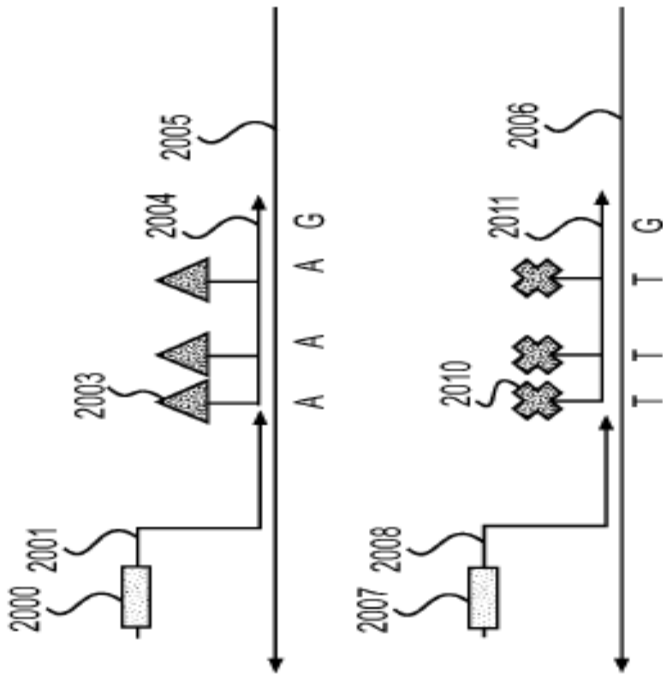


Figura 40

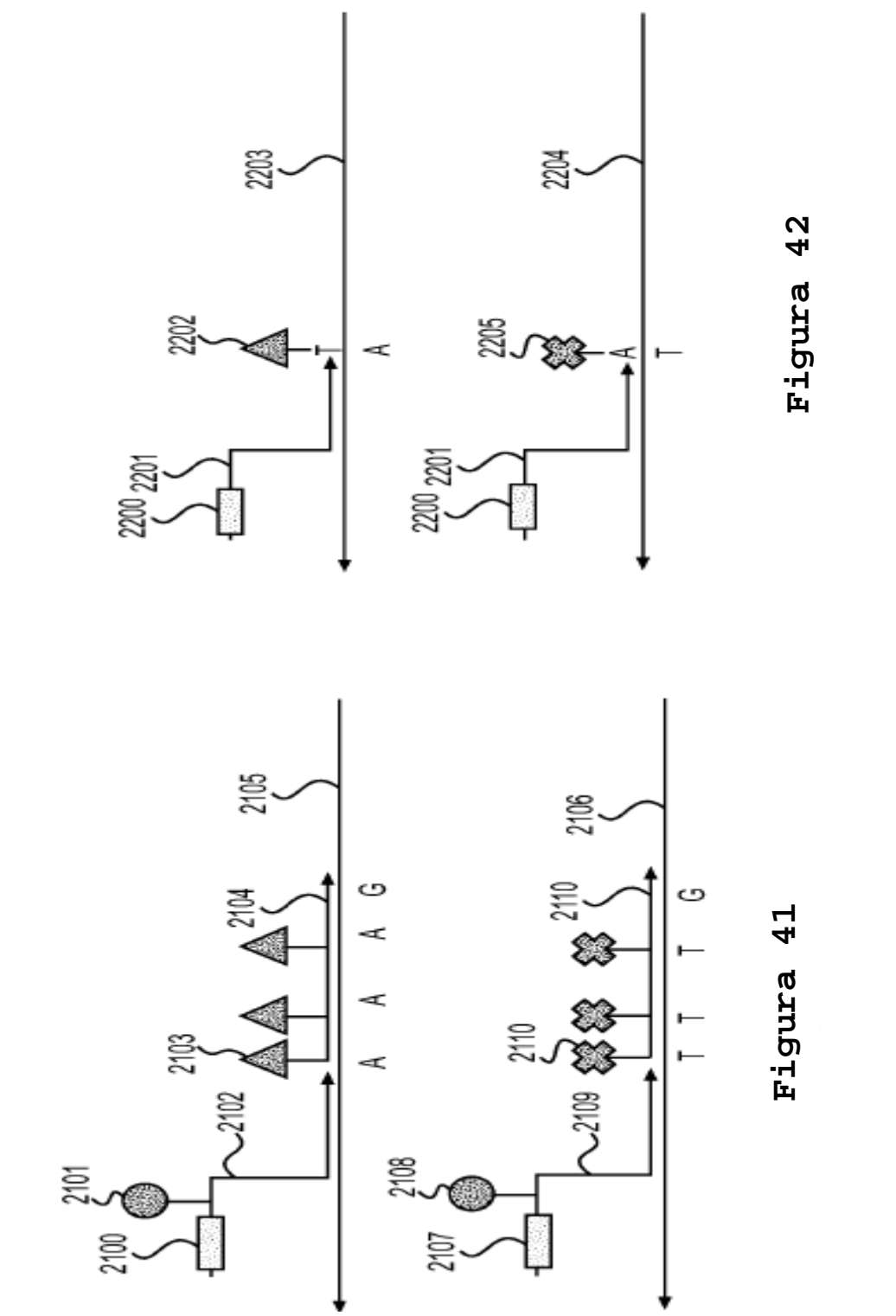


Figura 42

Figura 41

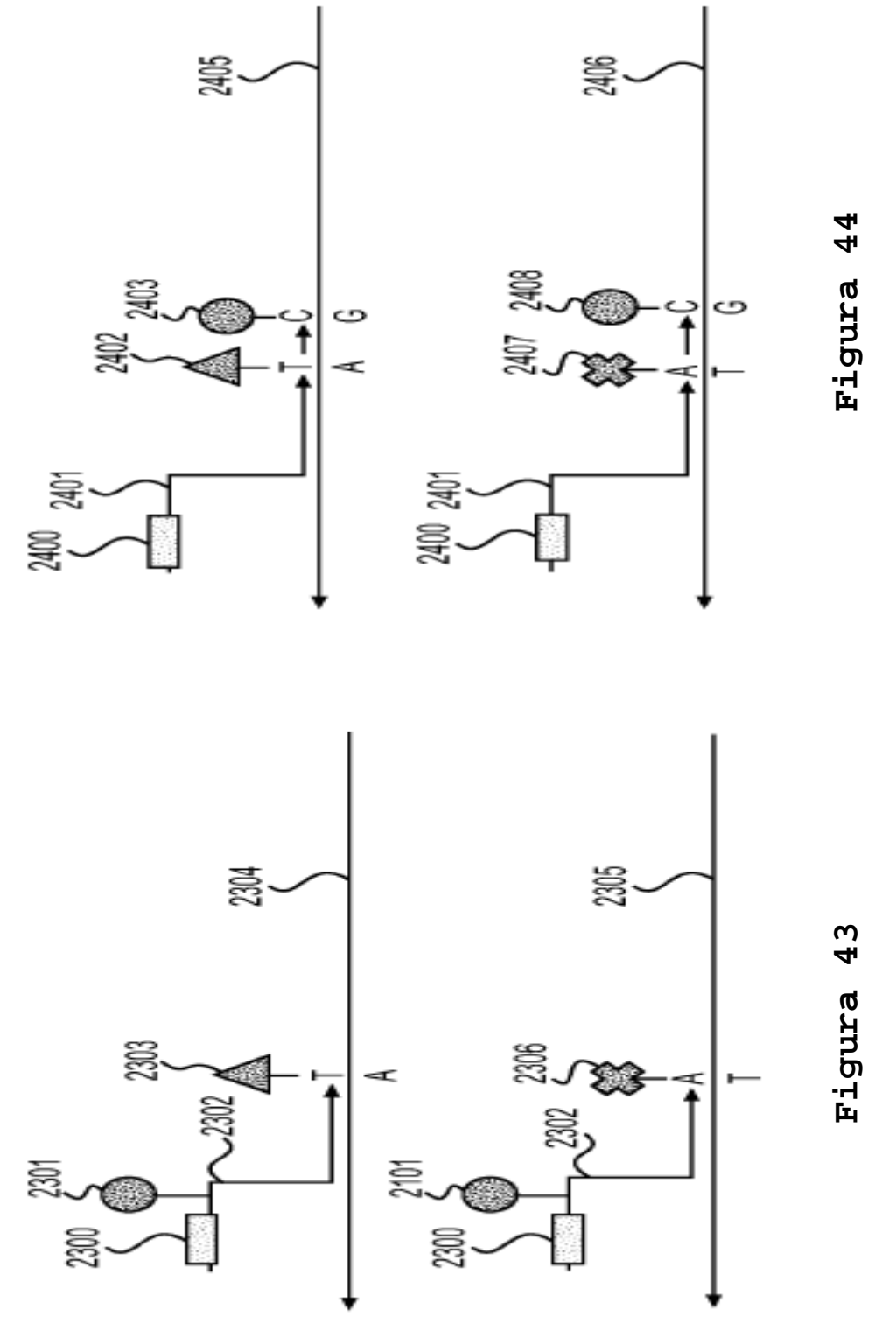


Figura 44

Figura 43

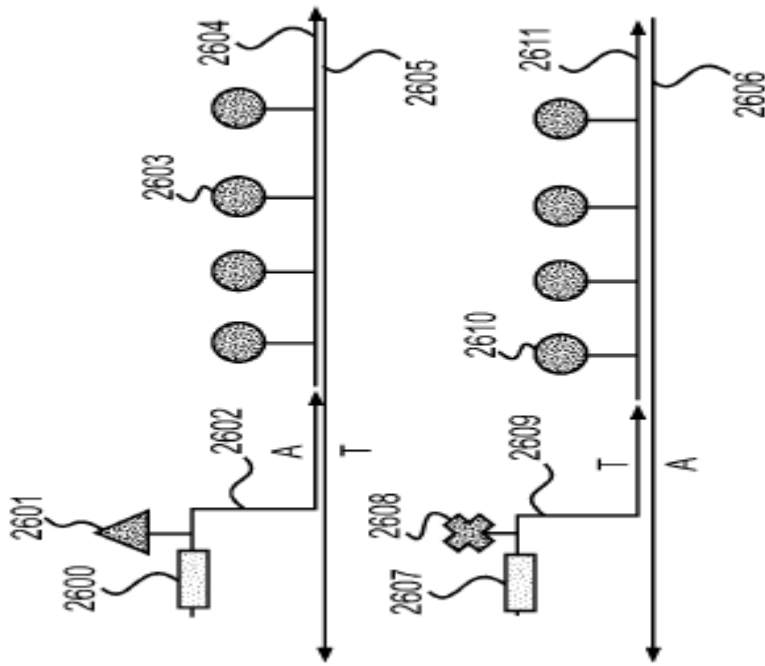


Figura 46

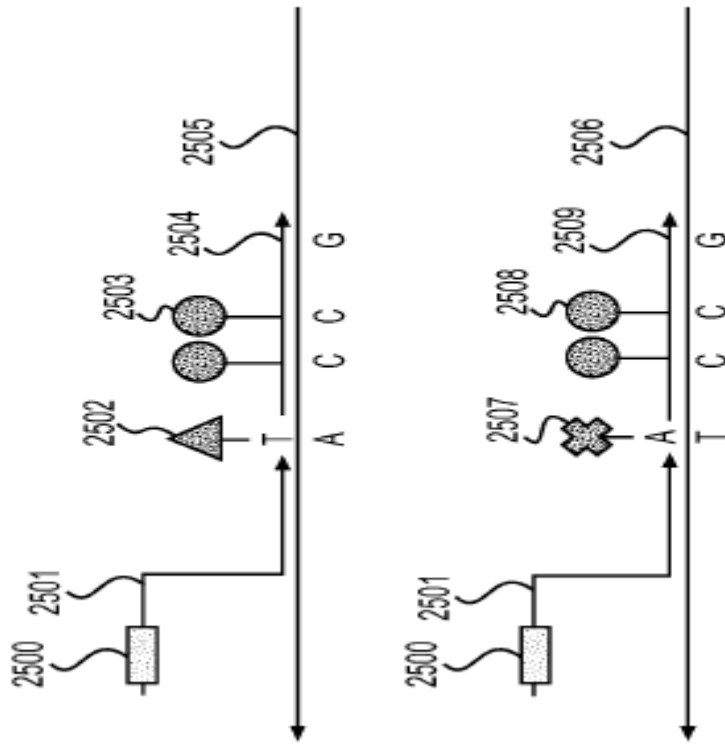


Figura 45

32/51

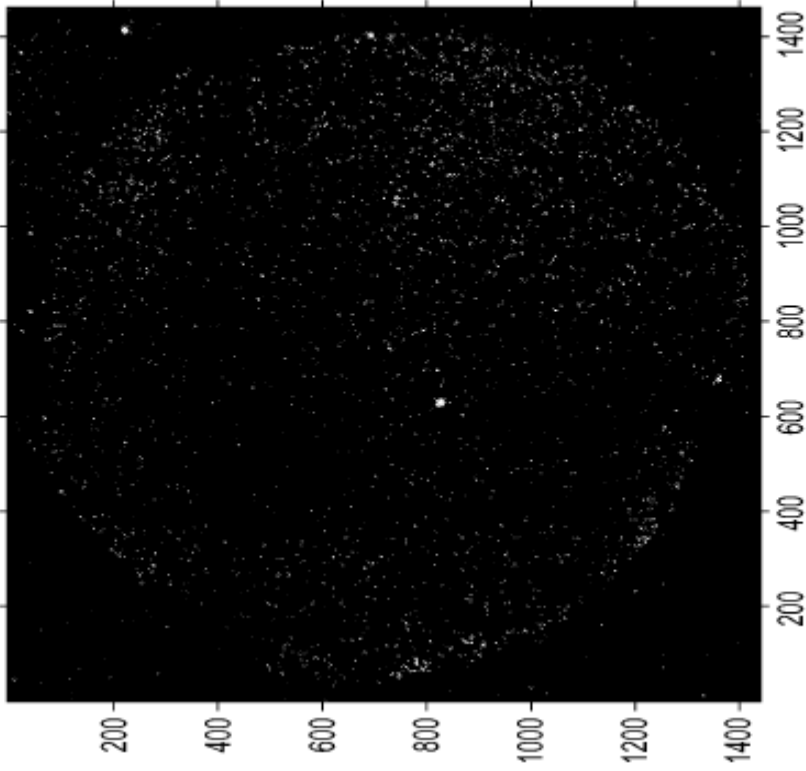


Figura 48

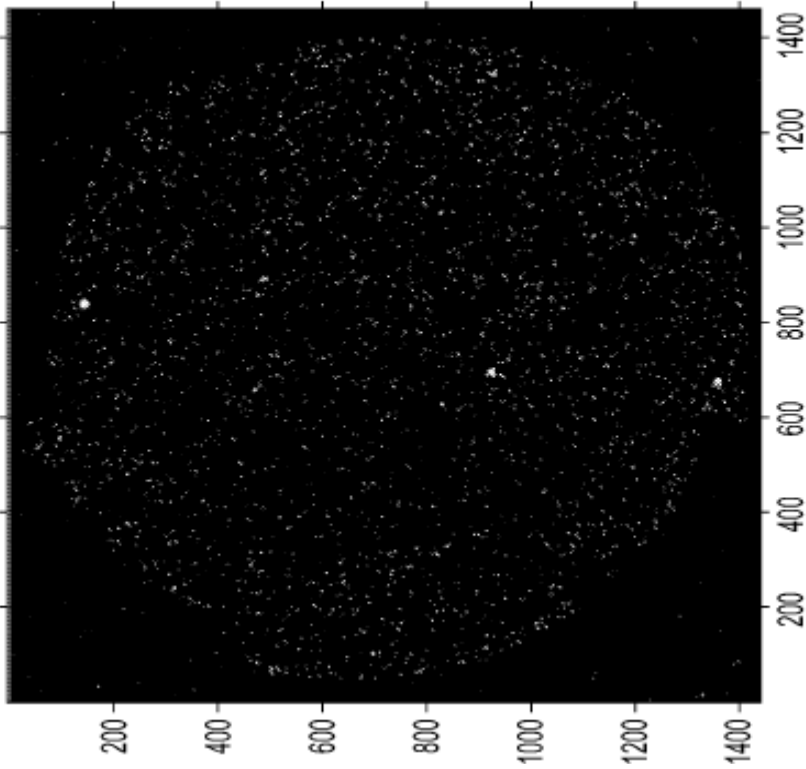


Figura 47

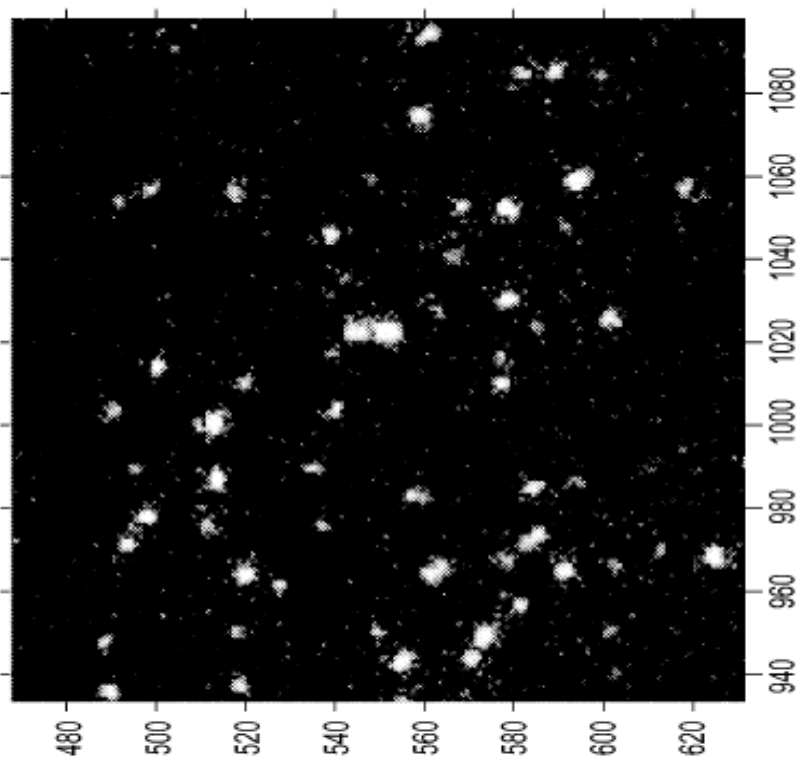


Figura 50

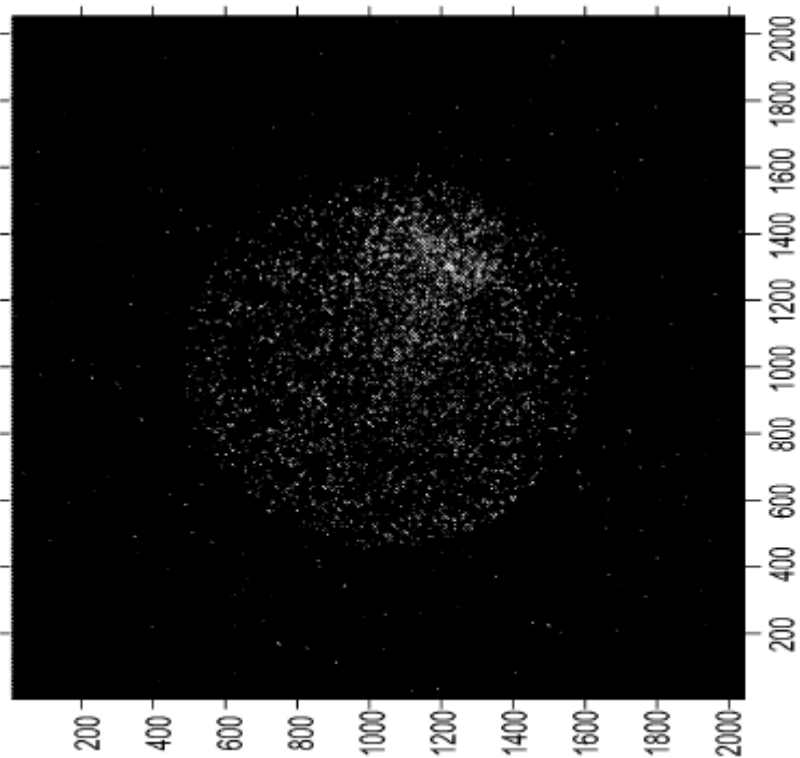


Figura 49

34/51

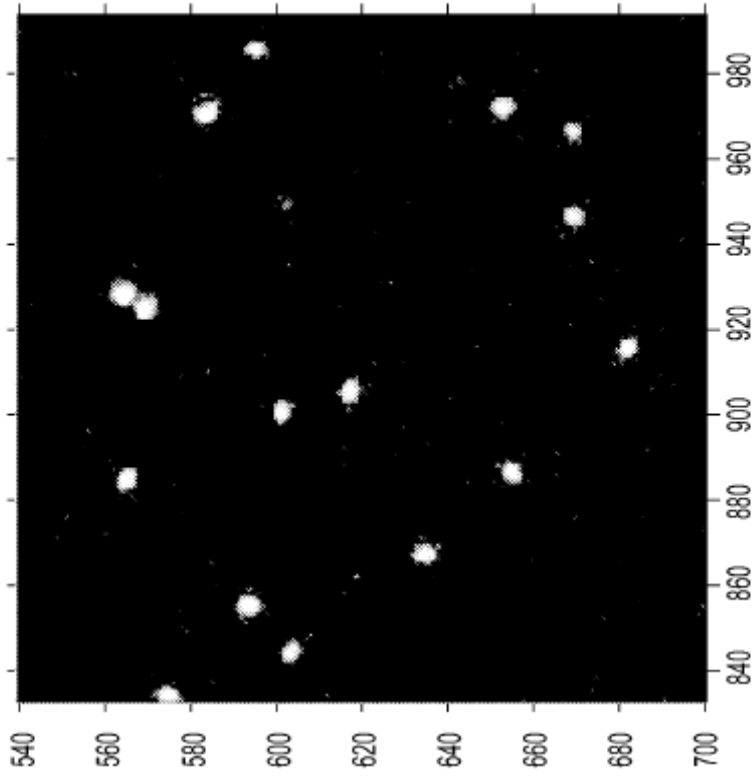


Figura 52

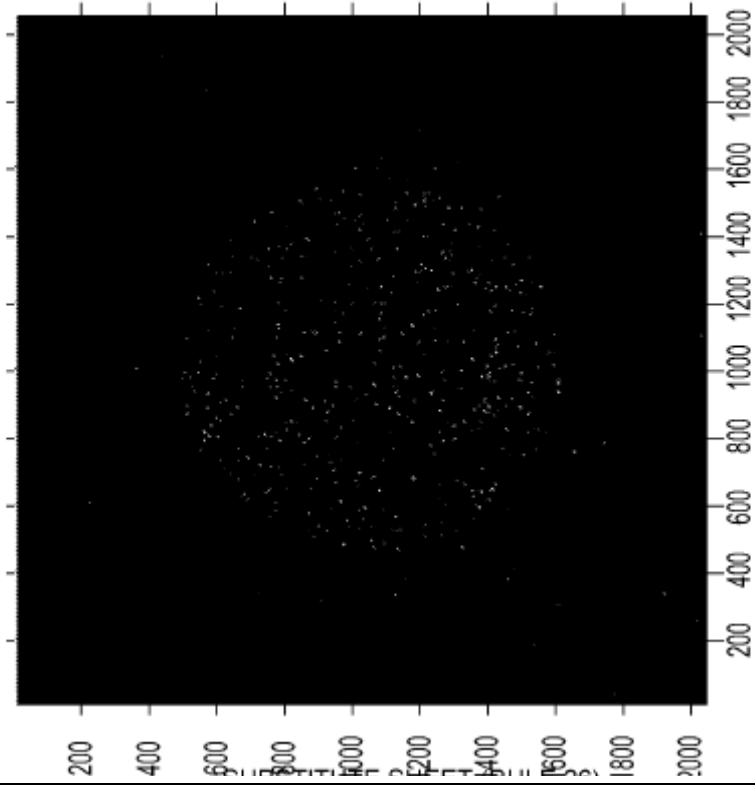


Figura 51

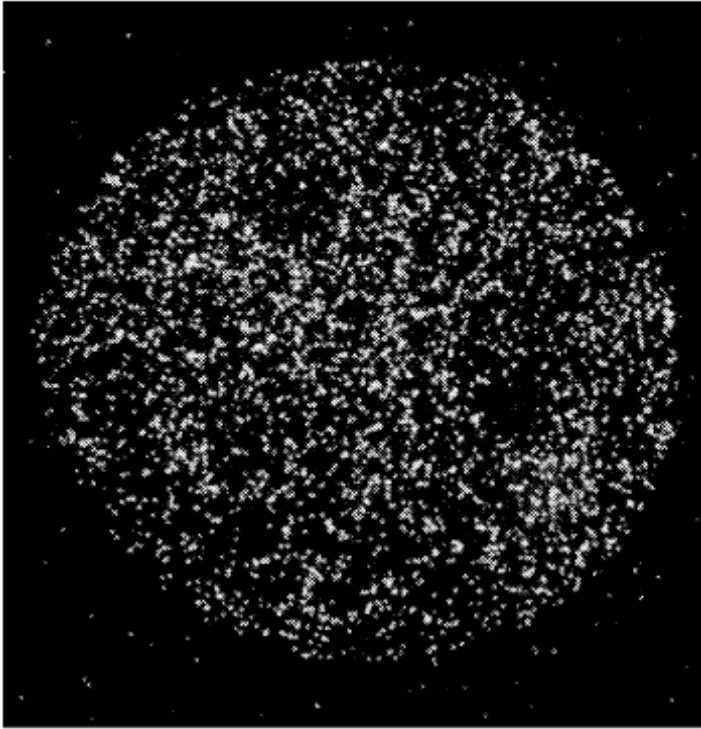


Figura 54

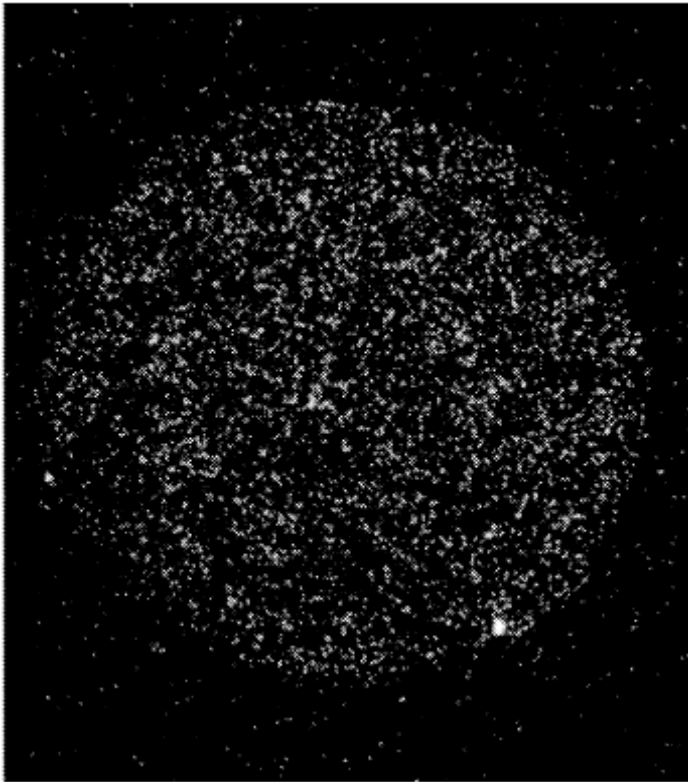


Figura 53

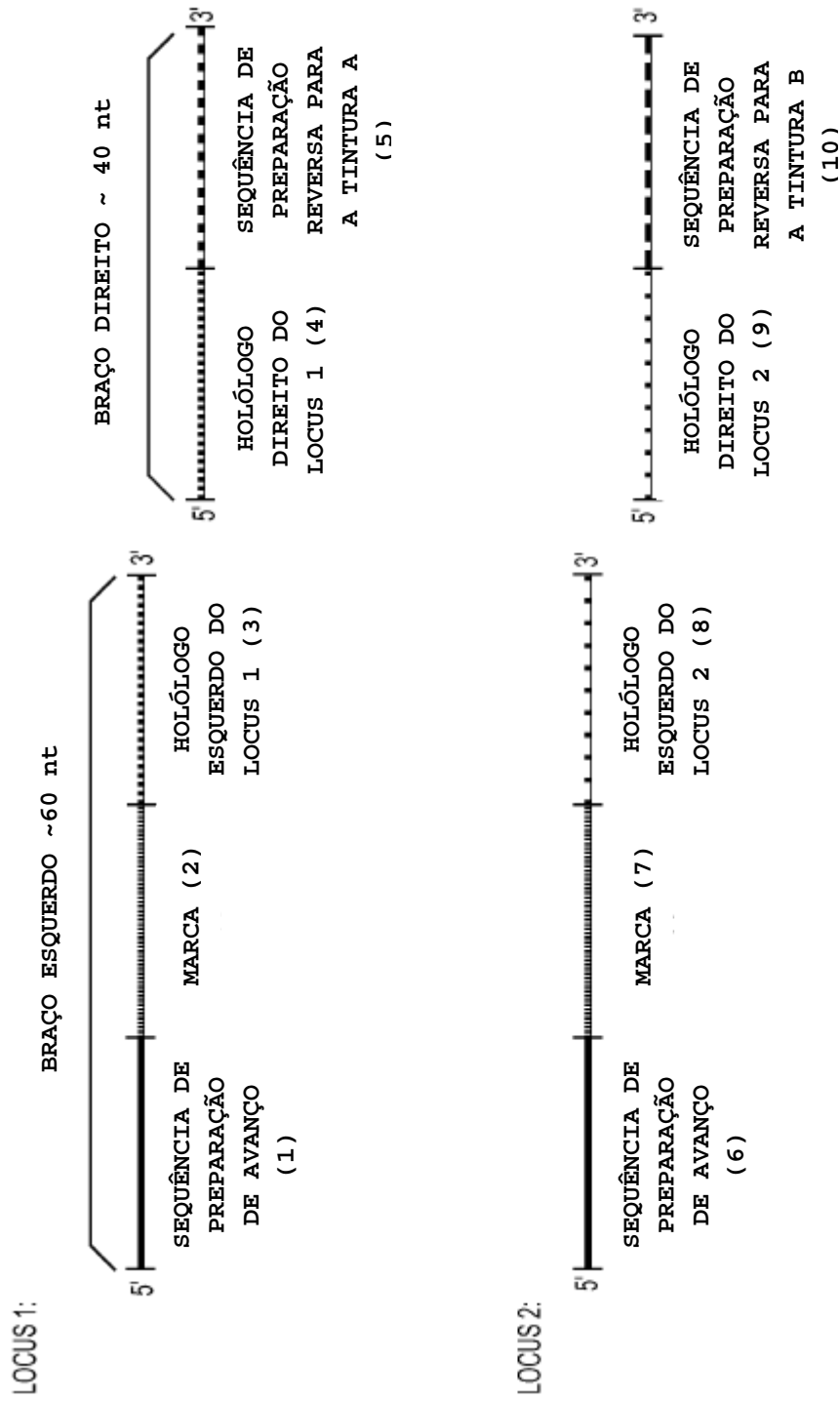


Figura 55

37/51

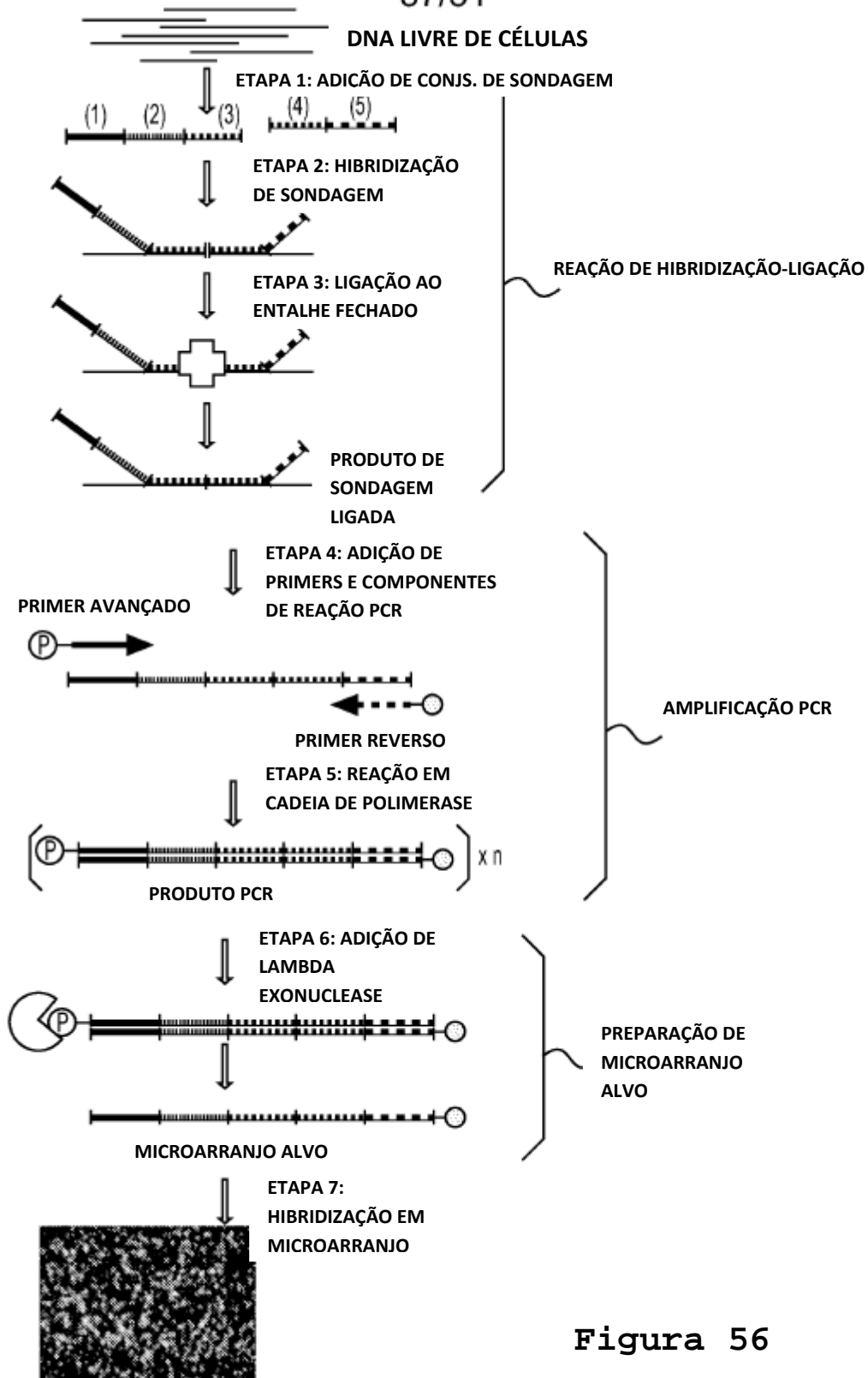


Figura 56

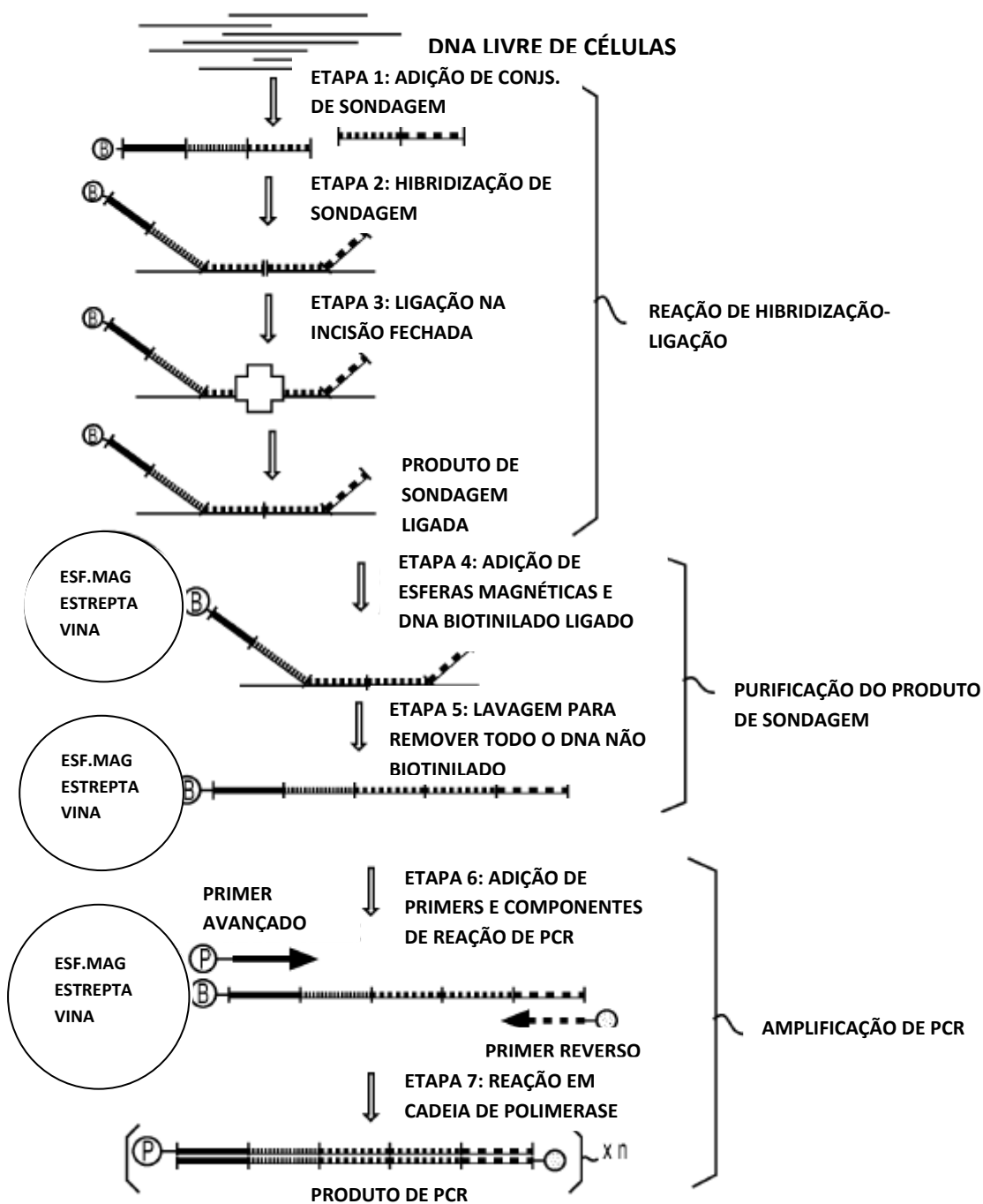


Figura 57

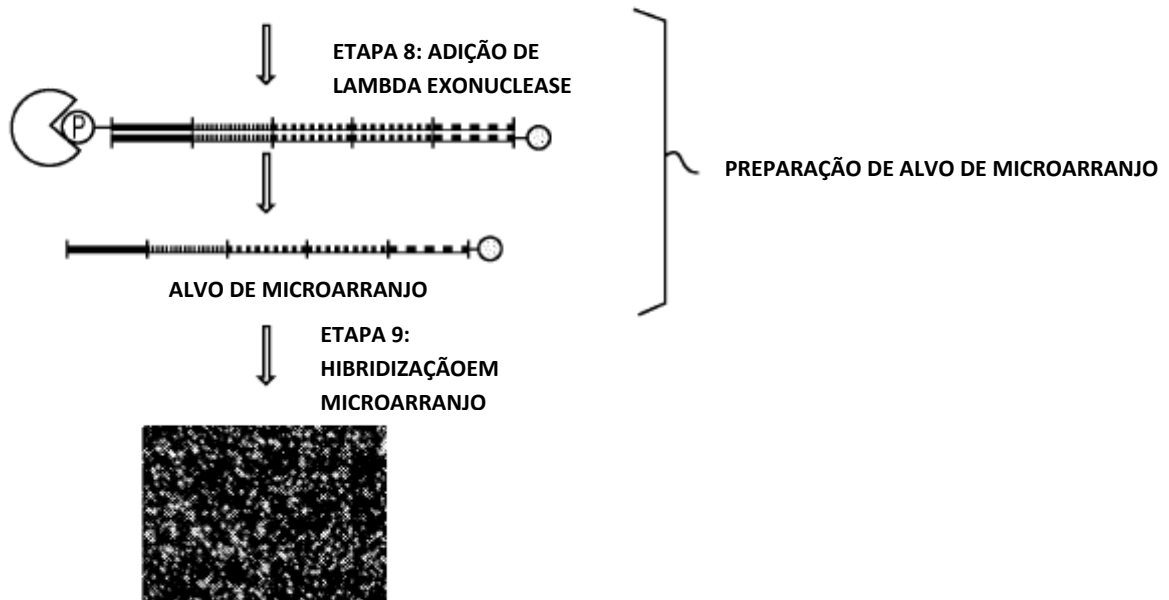


Figura 57 Continuação

40/51

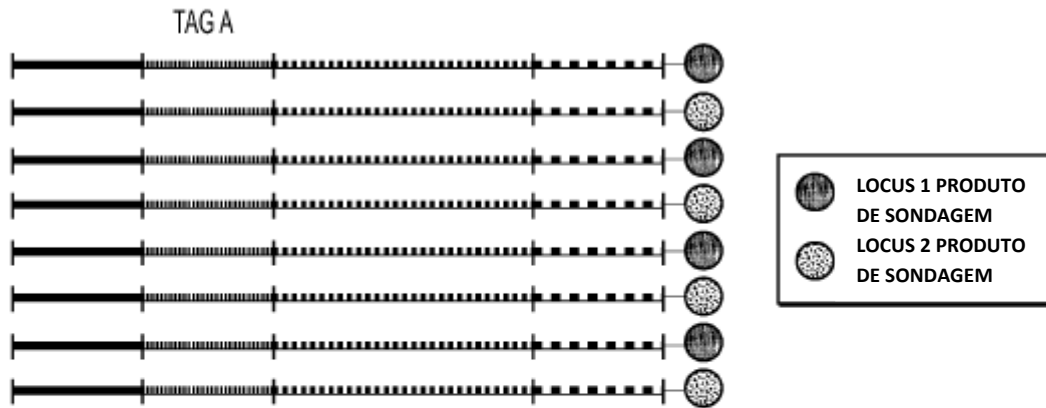


Figura 58A

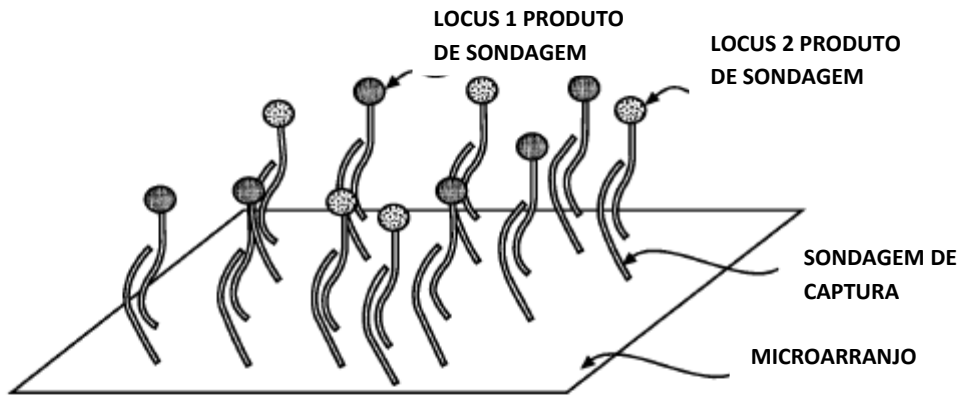


Figura 58B

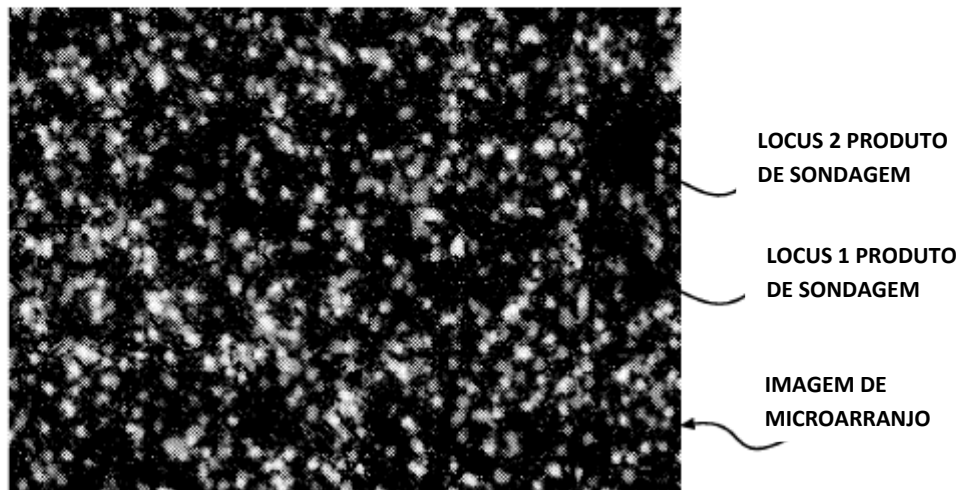


Figura 58C

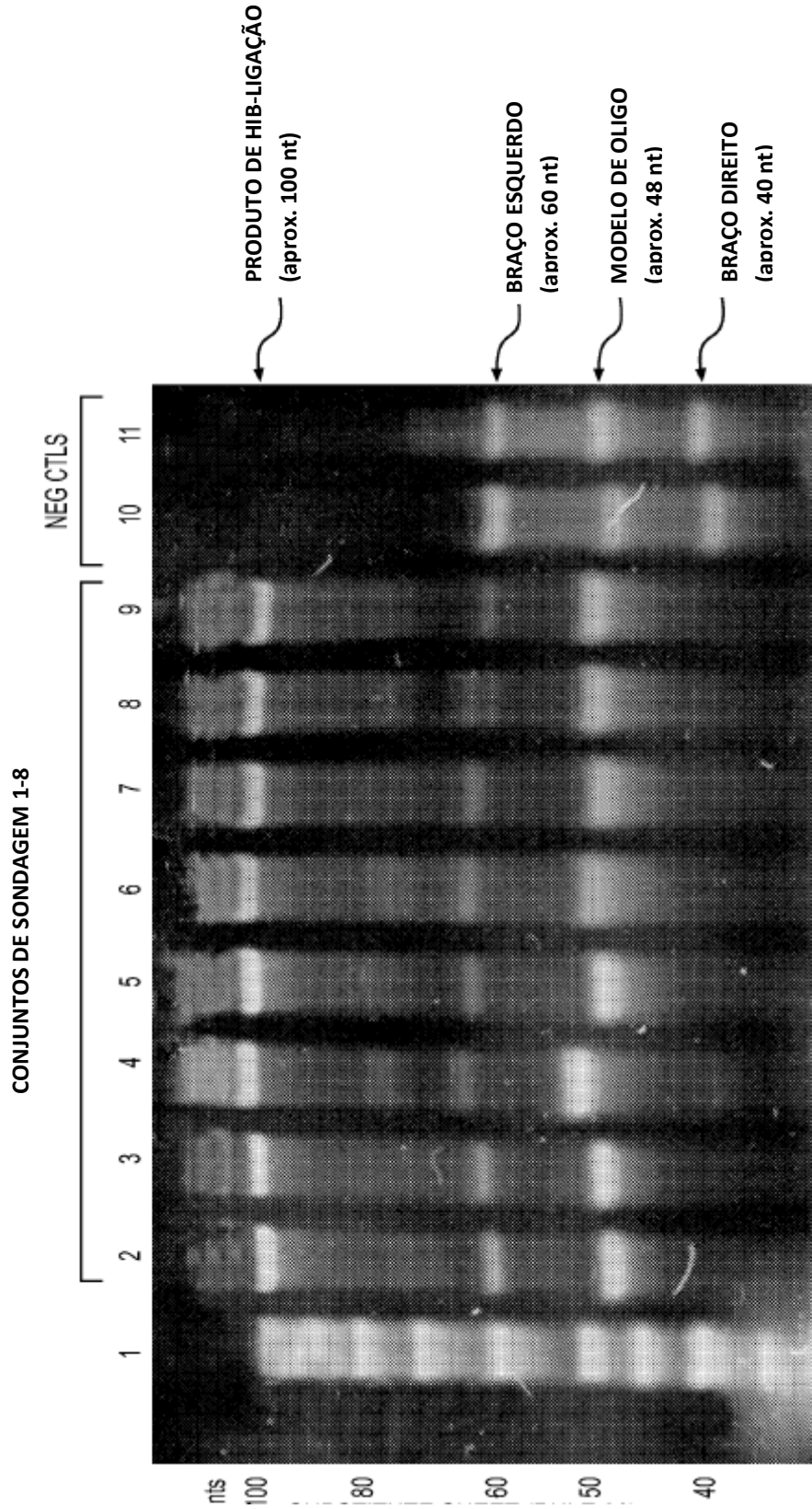


Figura 59

NÚMERO DE CÓPIA MEDIDO

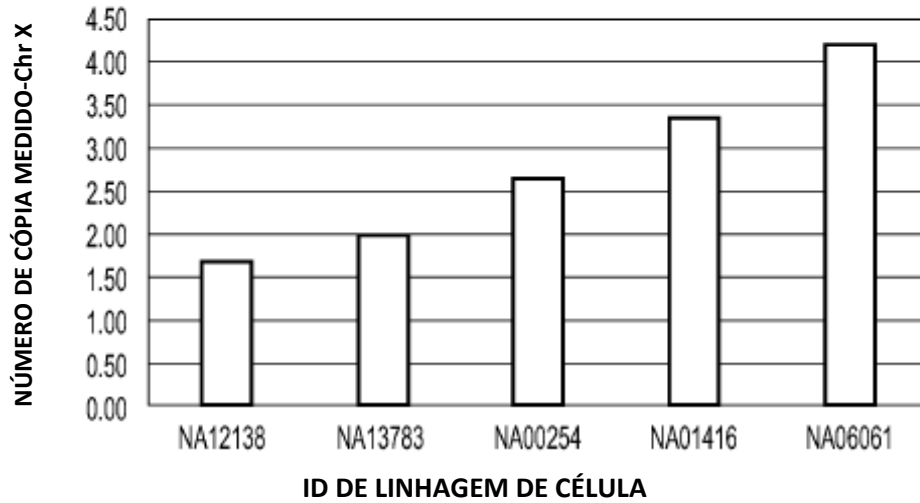


Figura 60A

NÚMERO DE CÓPIA MEDIDO VS TEÓRICO

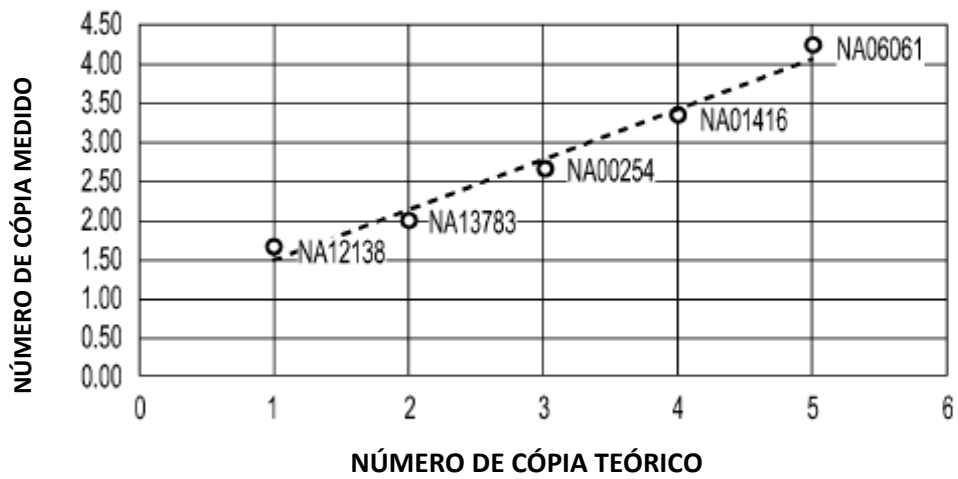
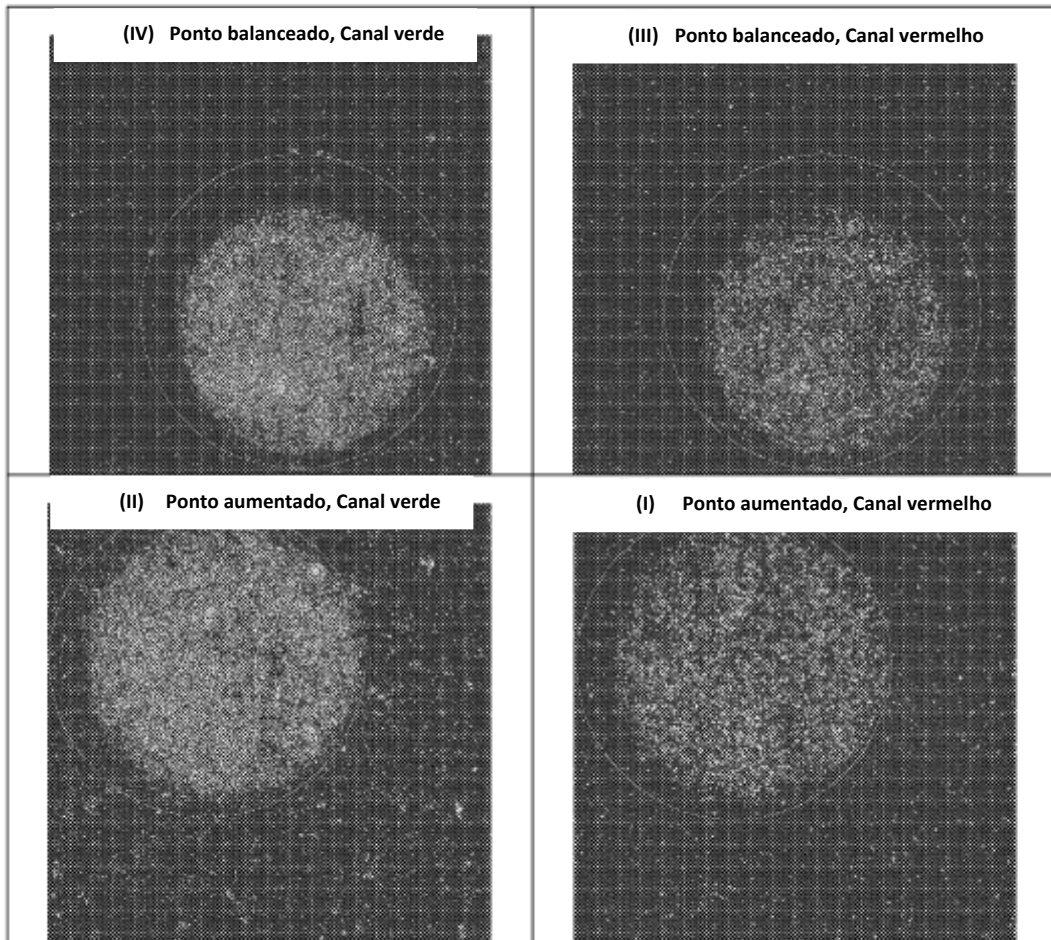


Figura 60B



**Figura 61A**

---

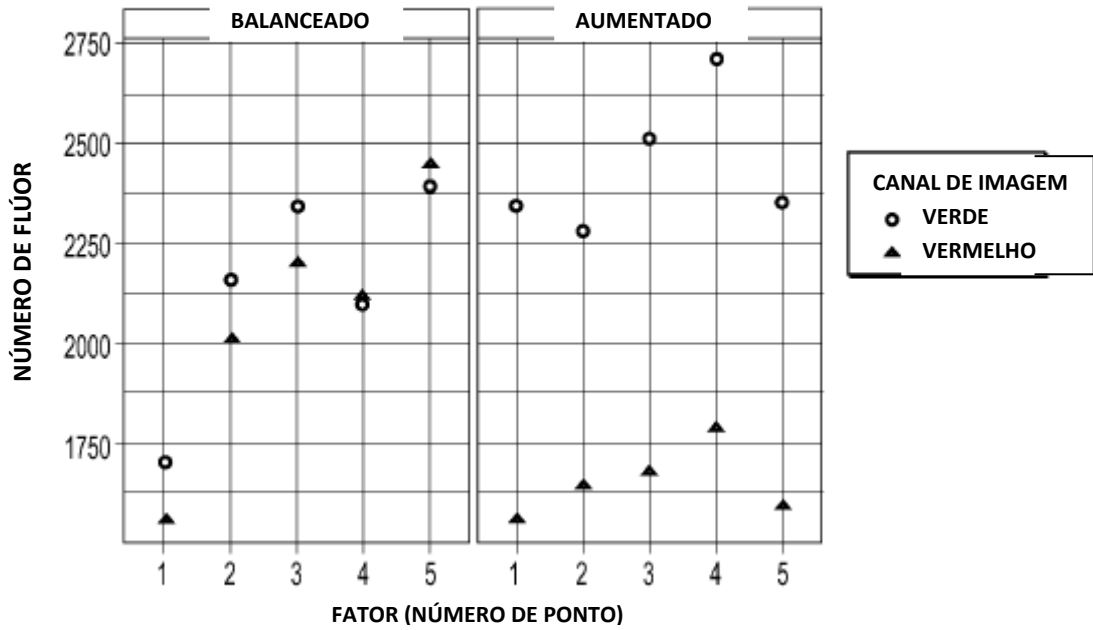


Figura 61B

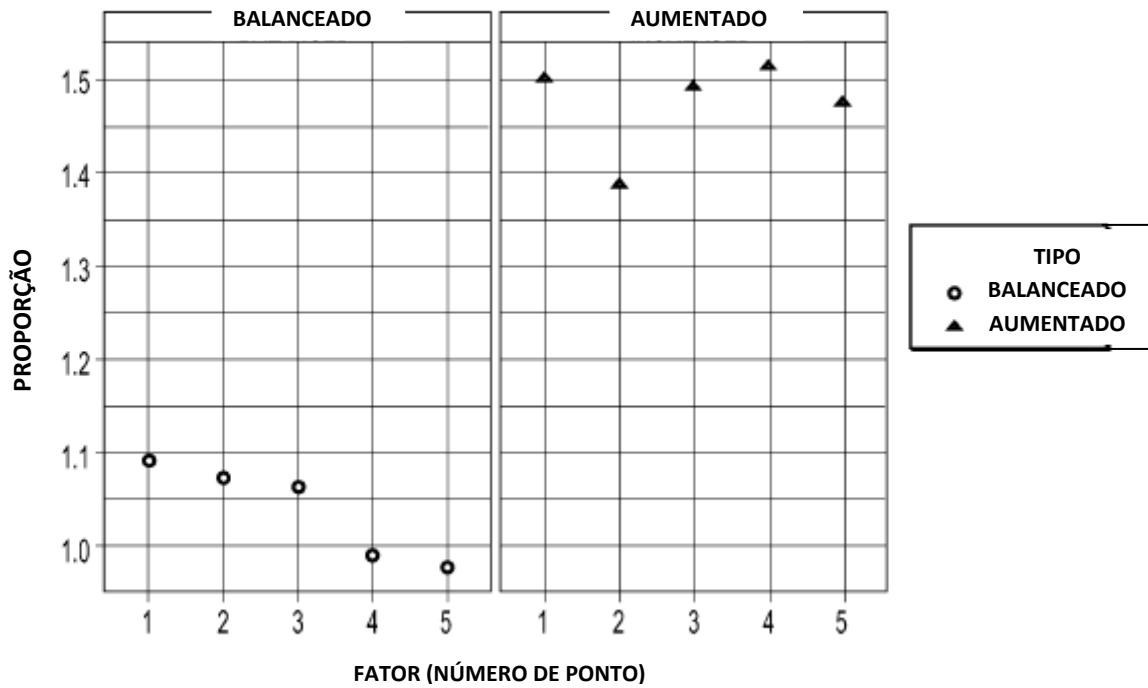


Figura 61C

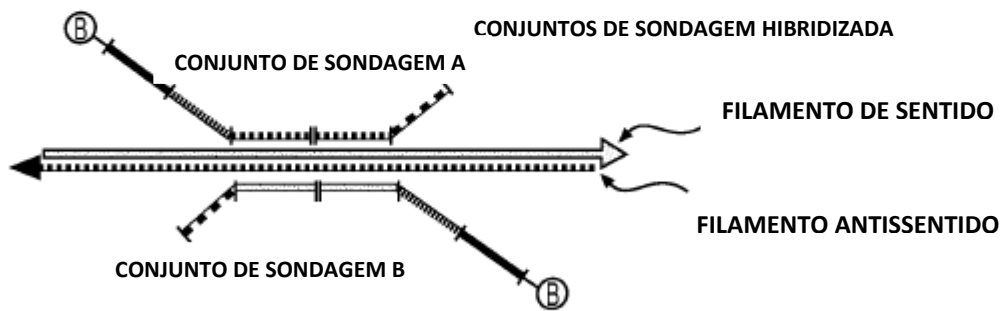


Figura 62

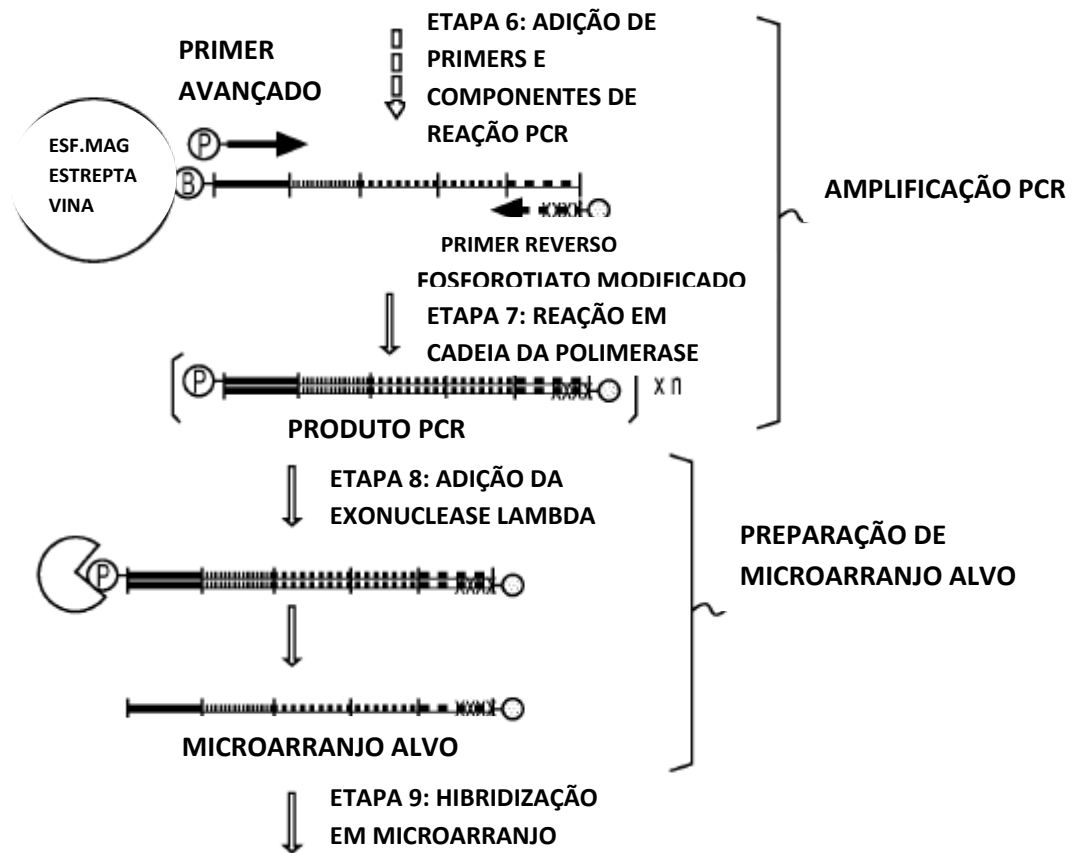


Figura 63

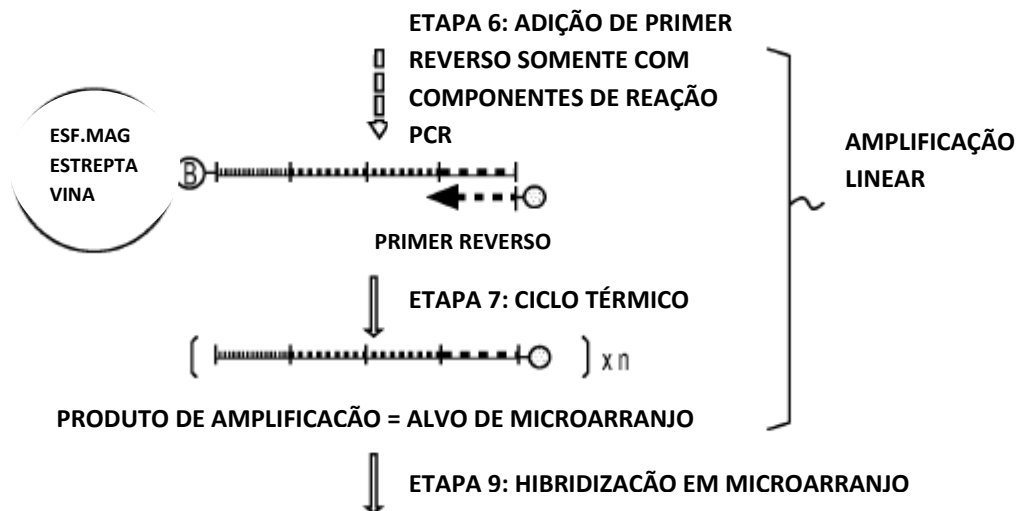
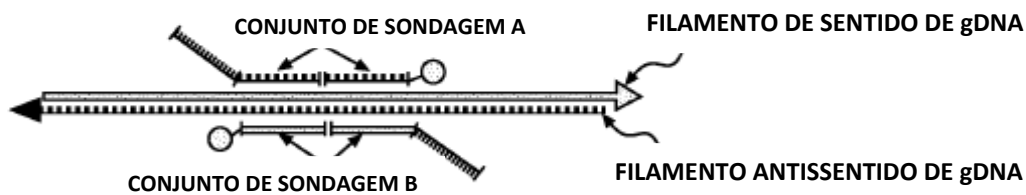


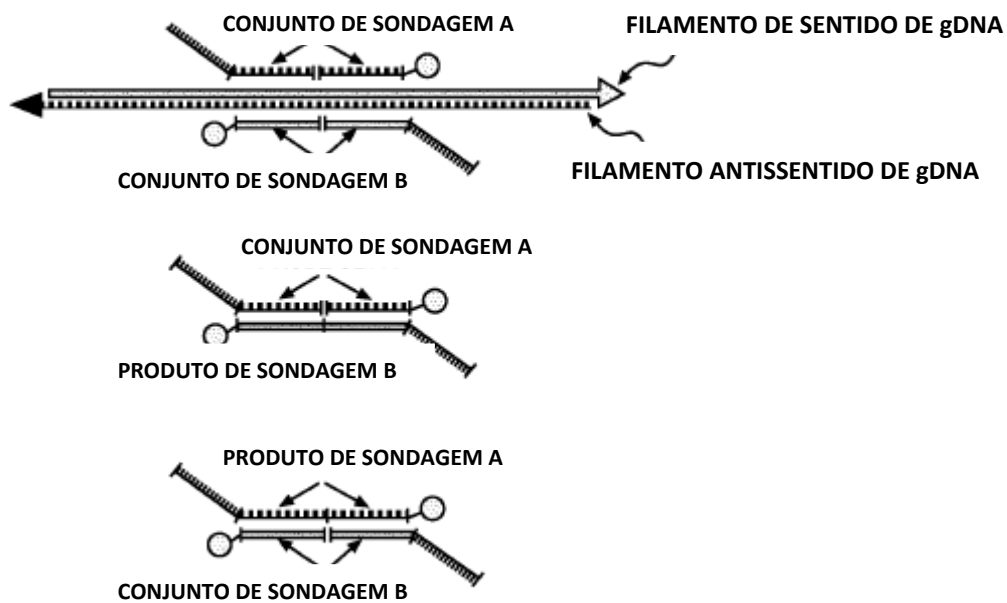
Figura 64

**CICLO 1 EVENTOS DE HIBRIDIZAÇÃO-LIGAÇÃO**



**Figura 65A**

**CICLO 2 EVENTOS DE HIBRIDIZAÇÃO-LIGAÇÃO**



**Figura 65B**

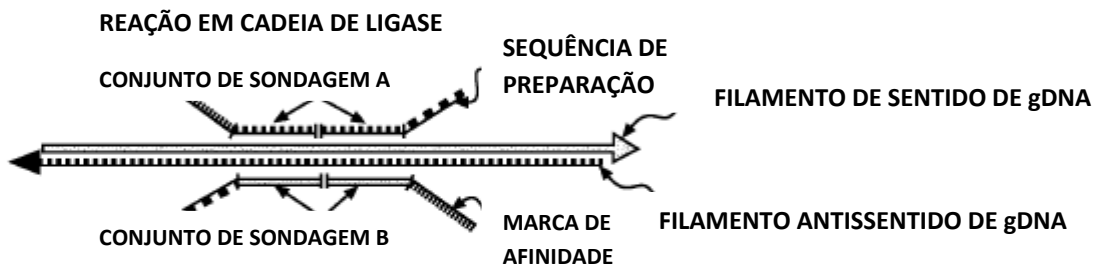


Figura 66A

AMPLIFICAÇÃO LINEAR E ROTULAÇÃO COM UM PRIMER

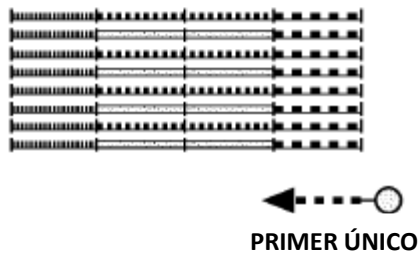


Figura 66B

ALVO DE MICROARRANJO

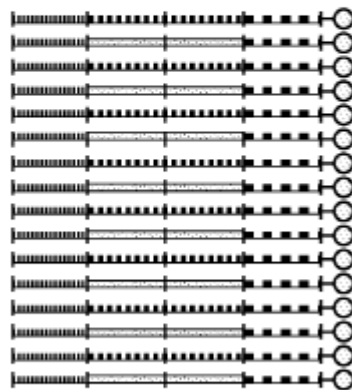
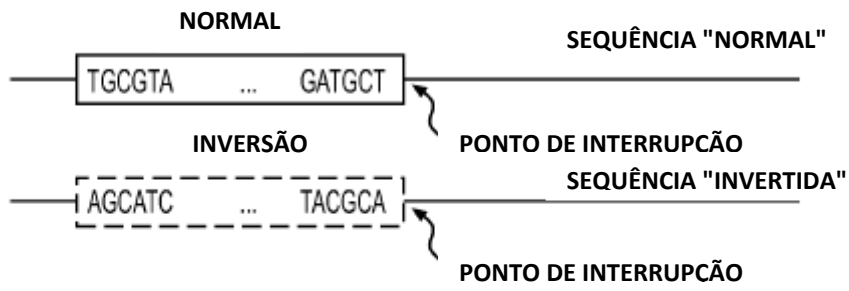
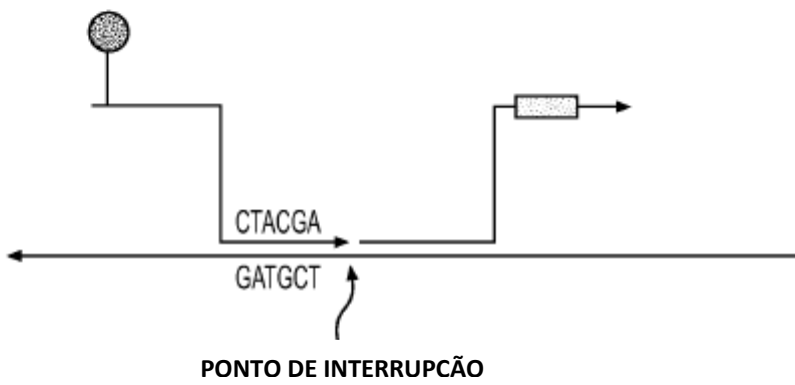


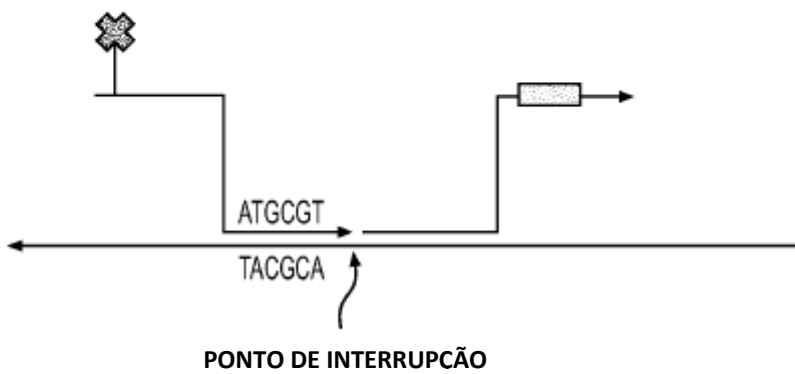
Figura 66C



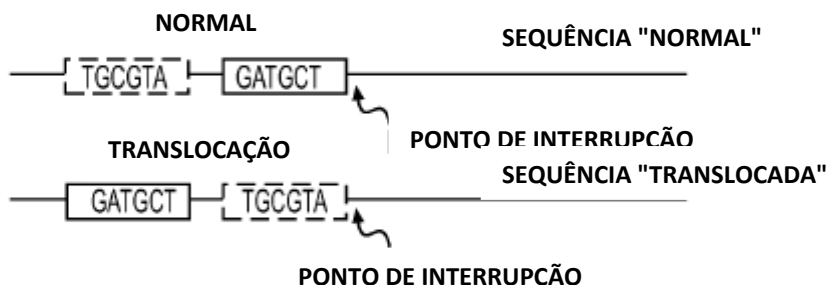
**Figura 67A**



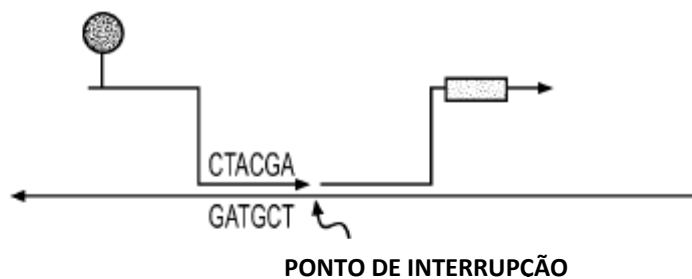
**Figura 67B**



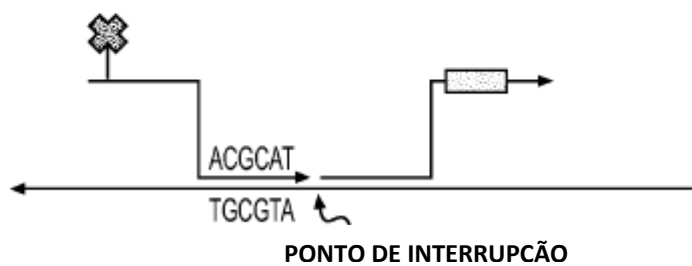
**Figura 67C**



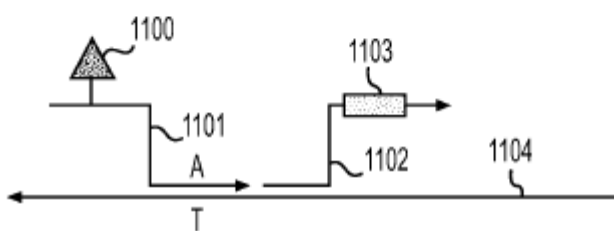
PONTO DE INTERRUPTÃO  
**Figura 68A**



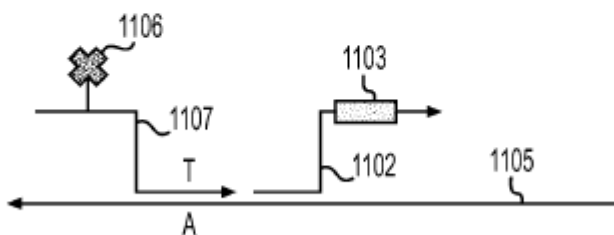
PONTO DE INTERRUPTÃO  
**Figura 68B**



PONTO DE INTERRUPTÃO  
**Figura 68C**



**Figura 69A**



**Figura 69B**