



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

(11) BR 112012032586-8 B1



(22) Data do Depósito: 30/06/2011

(45) Data de Concessão: 17/08/2021

(54) Título: MÉTODOS PARA DETERMINAR A SEQUÊNCIA NUCLEOTÍDICA DE UM ÁCIDO NUCLEÍCO DE INTERESSE E PARA DETERMINAR O GENÓTIPO HLA EM UMA AMOSTRA

(51) Int.Cl.: C12Q 1/68; C12N 15/11.

(30) Prioridade Unionista: 24/12/2010 CN PCT/CN2010/002150; 24/12/2010 CN PCT/CN2010/002149; 30/06/2010 CN 201010213721.2; 30/06/2010 CN 201010213717.6; 30/06/2010 CN 201010213719.5.

(73) Titular(es): BGI GENOMICS CO., LTD.

(72) Inventor(es): JIAN LI; SHIPING CHEN; XIANDONG ZHANG; YING LIU; CAIFEN ZHANG; TAO LIU; MEIRU ZHAO.

(86) Pedido PCT: PCT CN2011076688 de 30/06/2011

(87) Publicação PCT: WO 2012/000445 de 05/01/2012

(85) Data do Início da Fase Nacional: 19/12/2012

(57) Resumo: MÉTODO DE SEQUENCIAMENTO DE PCR E USO DO MESMO NA GENOTIPAGEM HLA A invenção proporciona um método de sequenciamento de PCR, onde a combinação de iniciadores de indexação, uma estratégia de cisalhamento de DNA incompleto e uma técnica de sequenciamento de segunda geração (técnica de sequenciamento de extremidade em pares) pode fazer com que o comprimento dos produtos de PCR que podem ser sequenciados por um sequenciador maiores do que o comprimento do sequenciamento máximo do sequenciador enquanto faz o uso completo das características da técnica de sequenciamento de segunda geração tal como um alto rendimento e baixo custo, desse modo, ampliando muito seu escopo aplicável. Além disso, a presente invenção também proporciona iniciadores de indexação para o método de sequenciamento de PCR e o uso do método em genotipagem, particularmente na análise de HLA, e também proporciona os iniciadores de PCR usados, particularmente os iniciadores de PCR para os genes HLA-A,B, HLA-C e HLA-DQB1.

“MÉTODOS PARA DETERMINAR A SEQUÊNCIA NUCLEOTÍDICA DE UM ÁCIDO NUCLÉICO DE INTERESSE E PARA DETERMINAR O GENÓTIPO HLA EM UMA AMOSTRA”

Pedidos Relevantes

[001] O presente pedido reivindica o direito de prioridade dos Pedidos de Patente Chinesas Nos. 201010213717.6, 201010213719.5, e 201010213721.2, como depositado em 30 de junho de 2010 e o direito de prioridade dos Pedidos Internacionais Nos. PCT/CN2010/002150 e PCT/CN2010/002149 como depositados em 24 de dezembro de 2010, os conteúdos dos quais estão incorporados para referência em sua totalidade.

CAMPO TÉCNICO

[002] A presente invenção refere-se ao campo técnico de sequenciamento de ácido nucléico, em particular, o campo técnico do sequenciamento de PCR. Além disso, a presente invenção também se refere à técnica de código de barras de DNA e uma estratégia de cisalhamento incompleta de DNA. O método da presente invenção é particularmente aplicável a segunda técnica de sequenciamento de geração, especialmente à técnica de sequenciamento de extremidade em pares, e também é aplicável à genotipagem de HLA, em particular, um método para a genotipagem HLA-A, HLA-B, HLA-C e HLA-DQB1, e também proporciona os pares primários para a amplificação do PCR usado no método.

ANTECEDENTES

[003] O método de sequenciamento de PCR refere-se a uma técnica onde os fragmentos de DNA de um gene de interesse são obtidos pelo método de PCR, e os fragmentos de DNA obtidos do gene de interesse são submetidos ao sequenciamento de DNA para obter a formação de sequência de DNA do gene de interesse. Os métodos de sequenciamento de PCR são amplamente aplicados aos campos tais como a detecção de uma mutação de gene e uma genotipagem por muito tempo.

[004] A técnica de sequenciamento de DNA é principalmente classificada na técnica de sequenciamento de DNA de primeira geração representada pelo método de sequenciamento Sanger e a técnica de sequenciamento de DNA de segunda geração representada pela Illumina GA, Roche 454, ABI Solid, e similares. A técnica de sequenciamento DNA Sanger é caracterizada pelas operações experimentais simples, visuais e resultados precisos, e o período experimental curto, e desse modo, é amplamente aplicado nos campos tais como detecção clínica de mutação de gene e genotipagem, onde um tempo de resposta rápida é altamente requerido como para os resultados de detecção. Entretanto, devido às características tais como um baixo rendimento e alto custo, sua aplicação nos campos onde a genotipagem é realizada em uma larga escala é limitada.

[005] Quando comparado com a técnica de sequenciamento de DNA de primeira geração, a técnica de sequenciamento de DNA de segunda geração tem as características tais como um alto rendimento de sequenciamento, baixo custo, alto nível de automação, e sequenciamento de molécula única. Tomado o sequenciamento de Illumina GA de molécula única como um exemplo, uma operação de sequenciamento única gera dados de 50G (cerca de 50 bilhões) bases, 5 bilhões de dados base por dia em média, e a média de custo de sequenciamento para a base é menor do que 1/1000 do custo de sequenciamento no método Sanger. Além disso, a análise dos resultados pode ser diretamente realizada por um computador. Desse modo, a técnica de sequenciamento de DNA de segunda geração é uma técnica quase compatível para os projetos de sequenciamento em larga escala. Entretanto, o comprimento de um sequenciamento contínuo é geralmente pequeno na técnica de sequenciamento de DNA de segunda geração. Atualmente, o comprimento de sequenciamento bidirecional máximo é de 200bp para Illumina GA; embora o comprimento de sequenciamento bidirecional máximo possa ser de até cerca de 500bp para a Roche 454 GS-FLX, o custo de sequenciamento é maior e o rendimento é baixo. Quando

uma ampliação de PCR é de um comprimento maior do que o comprimento de sequenciamento máximo em um sequenciador, o sequenciamento completo de toda a ampliação não pode ser realizado diretamente pelo sequenciamento, e a informação de sequência de DNA toda da ampliação não pode ser obtida. Devido a um curto comprimento de sequenciamento máximo, o aplicativo da técnica de sequenciamento de DNA de segunda geração no método de sequenciamento de PCR é limitado. Além disso, um aperfeiçoamento gradual da técnica de sequenciamento para obter um maior comprimento de sequenciamento máximo, é uma necessidade urgente de desenvolver uma nova técnica para superar a deficiência de um comprimento de sequenciamento máximo do sequenciador de DNA de segunda geração no campo de aplicação de sequenciamento de PCR.

[006] Um antígeno de leucócito humano (HLA) é um dos sistemas de genes encontrados até agora de maior polimorfismo. É um sistema de gene primário para modular uma resposta imune específica em corpos humanos e determinar uma diferença individual na susceptibilidade a doenças, e é estreitamente associada com uma rejeição a um transplante de órgão halogênico. Encontrou-se em estudos que quanto maior o grau de correspondência de genes, tais como HLA-A, B, C, DRB1 e DQB1, assim como a resolução é em um doador e em um receptor, mais tempo um transplante sobrevive. É ainda um item de teste regular submeter um doador em potencial e um receptor a uma genotipagem HLA de alta resolução antes do transplante das células tronco hematopoiéticas.

[007] A técnica atual de genotipagem de alta resolução de HLA de padrão internacional é um método de sequenciamento de PCR com base na técnica de sequenciamento Sanger, que compreende a amplificação de PCR das regiões de gene HLA correspondentes, sequenciamento dos produtos amplificados, submeter o resultado de sequenciamento a uma genotipagem com um software de genotipagem profissional, e finalmente obter a informação de genotipagem HLA da amostra. É

caracterizada pelos resultados visuais, alta resolução e capacidade de detectar novos alelos. Entretanto, devido às características do sequenciamento Sanger, tais como alto custo e baixo rendimento, este aplicativo em institutos como o banco de dados de registro de voluntários de células tronco hematopoiéticas (Banco de Medula Óssea), no qual a detecção da genotipagem HLA em grande escala é requerida, é limitada.

[008] Foi relatado que um método de sequenciamento de PCR com base em GS-FLX Roche 454 foi usado na genotipagem HLA. Entretanto, uma vez que o custo para o sequenciamento foi relativamente alto, não foi significativamente superior sobre a técnica de genotipagem HLA com base no sequenciamento Sanger em termos de rendimento de sequenciamento e custos de sequenciamento. Quando comparado com o Roche 454 GS-FLX, o Illumina GA tem um comprimento de sequenciamento, mas tem vantagens óbvias em termos de rendimento de sequenciamento e custo de sequenciamento. Se o defeito do curto comprimento de sequenciamento máximo de Illumina GA pode ser superado, sua aplicação na genotipagem HLA compensará a escassez do método de genotipagem HLA atual.

CONTEÚDOS DA INVENÇÃO

[009] Quando conduzindo uma análise de sequenciamentos simultaneamente para sequências associadas com um gene específico, a estratégia de sequenciamento de PCR é geralmente empregada, onde a combinação do iniciador de indexação e a técnica de sequenciamento de segunda geração são empregadas diretamente. Quando o comprimento de sequenciamento máximo do sequenciador pode cobrir o comprimento de todo o produto de PCR, a estratégia acima atende os requerimentos. Quando o comprimento de sequenciamento máximo do sequenciador não cobre o comprimento do produto de PCR todo, o Illumina GA precisa ser substituído pelo sequenciador de segunda geração tendo um maior comprimento de sequenciamento máximo (tal como Roche 454 GS-FLX). Se o comprimento de sequenciamento máximo ainda não atende os requerimentos, um sequenciador de primeira geração

tem que ser empregado com escarificação de custos e rendimento.

[010] A situação atual é que o Illumina GA tem um rendimento de sequenciamento super alto, mas seu comprimento de sequenciamento máximo é de 200bp apenas; embora o comprimento de sequenciamento máximo do Roche 454 GS-FLX possa alcançar 500bp, o custo para o sequenciamento é relativamente maior e o rendimento é relativamente abaixo de 1000bp, seu rendimento e custo não são comparáveis a aqueles sequenciadores de segunda geração.

[011] Há uma técnica capaz de aumentar o comprimento dos produtos de PCR que possa ser sequenciada totalmente por um sequenciador sem a escarificação de custo e rendimento? A combinação dos primeiros índices, a estratégia de cisalhamento incompleto de DNA, e a técnica de sequenciamento de segunda geração no presente pedido podem fazer com que o comprimento dos produtos de PCR que possam ser sequenciados pelo sequenciador maiores do que o comprimento de sequenciamento máximo do sequenciador enquanto faz uso completo das características da técnica de sequenciamento de segunda geração tal como um alto rendimento e baixo custo, desse modo, ampliando muito seu escopo aplicável. A técnica de sequenciamento de segunda geração empregada na presente invenção inclui, entre as técnicas de sequenciamento de segunda geração, uma técnica de sequenciamento de extremidade em pares, e uma técnica de sequenciamento de PCR que tem uma sequência de referência de DNA para o modelo PCR.

[012] A presente invenção proporciona métodos para o sequenciamento de PCR, pelo qual a limitação resultada do comprimento de sequenciamento máximo é aliviada e a aplicação de uma segunda técnica de sequenciamento de DNA de segunda geração no campo de aplicação de sequenciamento PRC é amplificada. Por exemplo, quando realizando um sequenciamento com a técnica de sequenciamento de segunda geração, os iniciadores de indexação tendo um iniciador de indexação adicionado à extremidade 5' são usados, os produtos de PCR amplificados são

cisalhados, os produtos cisalhados são terminalmente reparados e tem desoxiadenosina (A) ligada em suas 3' extremidades, e então são ligadas a diferentes adaptadores de PCR livres.

[013] Um método de sequenciamento de PCR, com base na técnica de código de barras de DNA e estratégia de cisalhamento incompleto de DNA, pode aumentar muito o número de amostras legendadas especificamente sem aumentar o número de iniciadores de indexação (figura 5). Na presente invenção, o comprimento sequenciado atualmente de produtos de PCR excede o comprimento de sequenciamento máximo do sequenciador adicionando os iniciadores de indexação aos iniciadores de PCR da frente e reverso, em combinação, em combinação com uma estratégia de cisalhamento incompleto de DNA e aplicando uma técnica de sequenciamento de segunda geração.

[014] A adição de uma sequência de indexação na extremidade frontal de uma amplificação primária é destinada a realizar um sequenciamento simultâneo de uma pluralidade de amostras. Concretamente falando, um único iniciador de indexação é adicionado a cada amostra durante o PCR usando uma técnica de código de barras/índice-PCR em combinação sintetizando um iniciador de indexação adicionando um iniciador de indexação à extremidade 5' de um iniciador de PCR. Como tal, durante o sequenciamento pela técnica de sequenciamento de segunda geração, as amostras têm que ser processadas uma por uma somente na etapa de PCR, e pode ser misturada junto e processada simultaneamente nas etapas experimentais restantes, e o resultado final para cada amostra pode ser traçado devido ao seu único iniciador de indexação.

[015] A técnica de índice "Adaptador" ou "adaptador de biblioteca" refere-se a uma técnica de indexação de biblioteca compreendendo adicionar diferentes adaptadores de biblioteca a bibliotecas de sequenciamento múltiplo (os diferentes adaptadores de biblioteca consistem de sequências diferentes, e a parte diferente entre as

sequências é chamada de índice de adaptador), construindo bibliotecas de sequenciamento indexadas, então realizando o sequenciamento múltiplas bibliotecas de sequenciamento indexadas diferentes em uma piscina, onde o resultado de sequenciamento final para cada biblioteca de sequenciamento indexado é distinguível. O termo “adaptador de biblioteca de PCR livre” refere-se a um segmento desenhado das bases, cujo principal papel reside na fixação auxiliar de uma molécula de DNA no chip de sequenciamento e reside em proporcionar os locais de ligação para os sequenciamentos primários universais, onde o adaptador de PCR livre poder ser diretamente ligado aos dois terminais dos fragmentos de DNA na biblioteca de sequenciamento. Uma vez que nenhum PCR esteja envolvido na introdução do adaptador, o adaptador é chamado de adaptador de biblioteca de PCR livre. Por exemplo, os adaptadores de biblioteca de PCR livre usados nos Exemplos da presente invenção são a partir do ILLUMINA.

[016] Um método de construir uma biblioteca de PCR livre, onde uma técnica de índice de adaptador de biblioteca é usado, refere-se a ligação direta do adaptador de biblioteca aos dois terminais do fragmento de DNA da biblioteca de sequenciamento. Uma vez que nenhum PCR está envolvido na introdução do adaptador de biblioteca, é chamado de construção de biblioteca de PCR livre. Uma ligase de DNA pode ser usada para a ligação no processo de introdução. Uma vez que nenhum PCR esteja envolvido no processo da construção de biblioteca, uma imprecisão dos resultados finais resultados do declive do PCR é envolvido durante a construção de uma biblioteca compreendendo os produtos de PCR de alta similaridade de sequência.

[017] Os métodos de amplificação de DNA, os métodos de extração de DNA, os métodos de purificação de DNA e os métodos de alinhamento de sequência de DNA como envolvidos na presente invenção podem ser quaisquer métodos disponíveis na arte. Tais métodos podem ser selecionados por um perito na arte de acordo com as situações práticas. Como para os métodos de sequenciamento de DNA, um

perito na arte pode realizá-los de acordo com os métodos convencionais ou seguindo as instruções do sequenciador.

[018] O desenho dos iniciadores de indexação varia dependendo da plataforma experimental aplicada. Em vista das características da plataforma de sequenciamento Illumina GA, os seguintes fatores são primeiramente considerados quando desenhando os iniciadores de indexação na presente invenção: 1: uma sequência repetida de mononucleotídeo compreendendo 3 ou mais bases é evitado nas sequências de iniciador de indexação, 2: a quantidade total de base A e base C no mesmo local de todas as contagens de iniciadores de indexação para 30%-70% da quantidade de todas as bases, 3: o conteúdo GC da própria sequência de índice primária é entre 40 e 60%, 4: os iniciadores de indexação diferenciam um do outro por pelo menos 4 bases, 5: as sequências tendo uma alta similaridade de sequência para os primeiros sequenciadores Illumina GA são evitados nas sequências de iniciador de indexação, e 6: a circunstância onde a adição das sequências primárias aos resultados primários de PCR em dímero e grampo sério, são reduzidos.

[019] Na presente invenção, os dois iniciadores de indexação (que são tanto idênticos quanto diferentes) são adicionados a dois terminais de um produto de PCR, respectivamente, de forma que o iniciador de indexação em cada produto de PCR pode especificamente legendar a informação da amostra do produto de PCR. O produto de PCR resultante é submetido a um cisalhamento incompleto. O então chamado “cisalhamento incompleto” refere-se à circunstância onde os produtos compreendem produtos de PCR não cisalhados intactos e produtos de PCR parcialmente cisalhados. Os métodos de cisalhamento incluem, mas não estão limitados a, métodos de cisalhamento químico (tais como digestão enzimática) e métodos de cisalhamento físicos. Os métodos de cisalhamento físicos incluem métodos de cisalhamento ultrassônico ou métodos de cisalhamento mecânicos. O DNA cisalhado é submetido a uma eletroforese de agarose 2%, e todas as faixas de DNA entre o comprimento de

sequenciamento máximo e o comprimento de DNA aplicável máximo do sequenciador são purificados e revestidos cortando o gel (o maior DNA aplicável ao sequenciador Illumina GA é 700bp, e o comprimento refere-se ao comprimento do DNA original, que não compreende o comprimento da sequência de adaptador de biblioteca). Os métodos para a purificação e recuperação incluem, mas não estão limitados a, recuperação por eletroforese e gel fatiado, e recuperação por esferas magnéticas. Os fragmentos de DNA recobertos são submetidos a construção de bibliotecas de sequenciamento de acordo com os procedimentos para a construção das bibliotecas de sequenciamento para o sequenciador de segunda geração, e então são submetidos ao sequenciamento. Preferencialmente, as bibliotecas de sequenciamento são construídas de acordo com os procedimentos de PCR livres para a construção das bibliotecas de sequenciamento, e o método de extremidade em pares é usado como método de sequenciamento. A construção de PCR livre das bibliotecas de sequenciamento é realizada de acordo com os métodos conhecidos por um perito na arte. Nos dados de sequenciamento obtidos, a informação de sequência para todas as amostras de teste podem ser obtidas em virtude das sequências de iniciadores de indexação. As leituras de sequência são alinhadas às sequências de referência de DNA correspondentes dos produtos de PCR por BMA, e a sequência completa é montada por sobreposição w relação de ligação entre as leituras de sequências (figura 1). A ligação aqui se refere a uma relação de ligação de extremidade em pares devido às características de sequenciamento de extremidade em pares.

[020] No sequenciamento Illumina GA (Sequenciador Analisador de Genoma da Illumina Inc., citado como Illumina GA para resumir), a análise de sequência de DNA é realizada com base nos princípios de sequenciamento por síntese. Pode ser aplicado a uma fase aplopica, e os dados obtidos finalmente referem-se a uma série de sequências base e podem ser aplicadas diretamente ao alinhamento com as sequências de referência no banco de dados HLA. Já que não tem o defeito de um

equívoco de picos presentes no software de tipagem tradicional, é vantajoso para a automação do software de tipagem. O Illumina GA tem um rendimento alto de sequenciamento. Atualmente, uma única operação de sequenciamento gera 50G (50 bilhões) de dados base, 5 bilhões de dados base por dia em média. Devido ao alto rendimento de dados, uma alta profundidade de sequenciamento pode ser obtida para cada sequência, assegurando assim a confiança dos resultados de sequenciamento.

[021] Ainda não há estudos na aplicação de Illumina GA no campo de tipagem HLA. A presente invenção aplica um sequenciamento de Illumina GA para o campo de tipagem HLA pela primeira vez, e realiza a tipagem HLA com baixo custo, alto rendimento, alta precisão e alta resolução usando uma técnica de sequenciamento de PCR, com base na técnica de código de barras de DNA, cisalhamento incompleto de DNA e preparação de biblioteca de PCR livre.

[022] Na presente invenção, usando uma técnica de sequenciamento de PCR que é baseada na técnica de código de barras de DNA, cisalhamento incompleto de DNA e preparação de biblioteca de PCR livre, as amostras a serem analisadas são agrupada; as amostras de cada grupo são submetidas à amplificação de um fragmento de interesse de genes HLA com os iniciadores legendados por iniciadores de indexação bidirecionais (comprimento máximo de produtos de PCR dependem do comprimento máximo do DNA que pode ser aplicado em um sequenciador; o comprimento de DNA máximo aplicável é de 700bp no Illumina GA atual, e o comprimento é o comprimento de DNA original, que não compreende o comprimento da sequência de adaptador de biblioteca); os produtos de PCR são agrupados juntos com a mesma quantidade, então submetidos a um cisalhamento incompleto e uma preparação de biblioteca de sequenciamento de DNA de PCR livre indexado. As bibliotecas de sequenciamento indexadas diferentes, como obtidas a partir de vários grupos de amostras, são misturadas em um mol igual, todos os fragmentos de DNA de um comprimento maior do que o comprimento de sequenciamento máximo do sequenciador são

seletivamente recobertos e são sequenciados por um sequenciador Illumina GA. As leituras de sequências de DNA para cada amostra podem ser obtidas rastreando a informação de sequência dos índices de adaptador, os iniciadores de indexação e os Iniciadores de PCR nos dados totais de sequenciamento. As sequências de DNA resultantes depois da montagem são alinhadas com os dados correspondentes em um banco de dados profissional IMGT HLA, determinando assim o genótipo HLA da amostra finalmente.

[023] Nos métodos descritos acima, depois de cisalhar tal DNA, o DNA das amostras de diferentes grupos são ligados a diferentes adaptadores de biblioteca durante a preparação da biblioteca de PCR livre, e, portanto, nas seguintes etapas de tipagem, os dados de sequenciamento resultantes podem ser traçados nas amostras um por um com base nos iniciadores de indexação e índices de adaptadores usados em cada amostra. As sequências de cada amostra são alinhadas à sequência de referência de DNA conhecida correspondente ao produto de PCR por um software. Com base na sequência de sobreposição e na relação de ligação, uma sequência intacta para o produto de PCR é montada a partir das sequências do DNA cisalhado.

[024] A presente invenção proporciona métodos de genotipagem HLA de alta resolução com base em técnicas de sequenciamento Illumina GA, realizando, assim, um sequenciamento haplotípico e um software de automação de tipagem, aumentando o rendimento de genotipagem HLA e reduzindo os custos.

[025] Devido ao requerimento no comprimento do modelo de DNA nas técnicas atuais de sequenciamento e a leitura curta de comprimento nas técnicas de sequenciamento atuais, os Iniciadores de PCR originais para os métodos HLA-SBT não são mais aplicáveis aos novos métodos de tipagem HLA de alta resolução com base na técnica de sequenciamento. A presente invenção projeta novos Iniciadores de PCR com uma boa especificidade e conservação, que amplifica o Éxons 2, 3, 4 do HLA-A, gene B independentemente, e cujos produtos de PCR tem um comprimento não mais

que 700bp e são particularmente aplicáveis a Illumina GA (o comprimento de DNA máximo aplicável ao Illumina GA atual é de 700bp). Um conjunto de Iniciadores de PCR como proporcionado na presente invenção é aplicável a genotipagem HLA para sujeitos (humanos em particular) em uma larga escala, alto rendimento e baixo custo.

[026] Nas soluções técnicas empregadas na presente invenção, todas as últimas sequências de genes HLA-A/B são baixadas a partir de um website da internet IMGT/HLA (<http://www.ebi.ac.uk/imgt/hla/>), e então gravados no disco local como conjunto de dados HLA-A; entretanto, todas as últimas sequências de genes HLA-I outras que não as sequências HLA-A são baixadas como um conjunto de dados de comparação. Tais dois conjuntos de dados são comparados para procurar sequências conservadoras e específicas para cada local de gene nos dois terminais e na parte interna do Éxons 2, 3, 4, e a sequência primária de PCR designada é comparada com a sequência total do genoma humano para homologia. Já que o gene HLA-A/B é altamente similar a outros genes pertencentes as moléculas HLA-I em termos de sequência, quando designam os iniciadores de PCR, o 3' terminal do iniciador deve ser específico tanto quanto possível de forma a assegurar a especificidade da amplificação do gene HLA-A/B com os iniciadores. Entretanto, o comprimento dos produtos de PCR é menor do que 700bp, e a temperatura de anelamento de iniciadores para frente e reversos são substancialmente a mesma.

[027] Os pares múltiplos de iniciadores HLA-A/B candidatos atendendo aos requerimentos de projeto são usados para amplificar os modelos de DNAs de sorotipos HLA-A/B. Entre eles, dois conjuntos de Iniciadores de PCR de HLA-A/B (seis pares para cada conjunto) com o melhor conservatismo e especificidade, para a amplificação de Éxons 2, 3 e 4, respectivamente, são triados.

[028] Os dois conjuntos de Iniciadores de PCR (6 pares para cada conjunto) são usados como iniciadores básicos, com base em que, 95 conjuntos de iniciadores de indexação que são usados para a amplificação de 95 e 950 modelos de DNA de

sorotipos comuns de HLA-A/B (os sorotipos destes modelos incluem todos os sorotipos comuns de HLA-A/B), respectivamente, são designados. Todos os produtos de PCR são sequenciados com Illumina GA de extremidade em pares 100 depois da mistura em uma quantidade igual, e os resultados de sequenciamento depois da montagem são comparados com os resultados de tipagem originais para confirmar o conservatismo e a especificidade dos Iniciadores de PCR.

[029] O HLA-A, os iniciadores B são designados na presente invenção, isto é, dois conjuntos de iniciadores de PCR de HLA-A/B (6 pares para cada conjunto) para a amplificação de Éxons 2, 3 e 4, respectivamente, são mostrados na Tabela 1 e 2.

Tabela 1: Iniciadores de PCR de HLA-A, B

SEQ ID NO:	No. iniciador	sequência de iniciador	o uso do iniciador	comprimento dos produtos
1	A-F2	CCTCTGYGGGGAGAAGCAA	Amplificação	480bp
2	A-R2	ATCTCGGACCCGGAGACTG	Éxon 2 de gene HLA-A	
3	A-F3	CGGGGCCAGGTTCTCACAC	Amplificação	410bp
4	A-R3	GGYGATATTCTAGTGTTGGTCCCA A	Éxon 3 de gene HLA-A	
5	A-F4	GTGTCCCATGACAGATGCAAAA	Amplificação	430bp
6	A-R4	GGCCCTGACCCTGCTAAAGG	Éxon 4 de gene HLA-A	
7	B-F2	AGGAGCGAGGGGACCGCA	Amplificação	400bp
8	B-R2	CGGGCCGGGGTCACTCAC	Éxon 2 de gene HLA-B	
9	B-F3	CGGGGCCAGGGTCTCACA	Amplificação	370bp
10	B-R3	GAGGCCATCCCCGGCGAC	Éxon 3 de	

			gene HLA-B	
11	B-F4	GCTGGTCACATGGGTGGTCCTA	Amplificação	
12	B-R4	CTCCTTACCCCATCTCAGGGTG	Éxon 4 de gene HLA-B	380bp

Tabela 2: Iniciadores de PCR de HLA-A, B

SEQ ID NO:	No. do ini- ciador	sequência de iniciador	O uso do ini- ciador	Compri- mento dos produtos
13	A-F2s	CCTCTGYGGGGAGAAGCAA	Amplificação	481bp
14	A-R2s	GGATCTCGGACCCGGAGACTGT	Éxon 2 de gene HLA-A	
15	A-F3s	TGGGCTGACCGYGGGGTC	Amplificação	
16	A-R3s	GGYGATATTCTAGTGTTGGTCCCA A	Éxon 3 de gene HLA-A	403bp
17	A-F4s	GTGTCCCATKACAGATGCAAAA	Amplificação	
18	A-R4s	GGCCCTGACCCTGCTAAAGG	Éxon 4 de gene HLA-A	405bp
19	B-F2s	AGGAGCGAGGGGACCGCA	Amplificação	
20	B-R2s	CGGGCCGGGGTCACTCAC	Éxon 2 de gene HLA-B	400bp
21	B-F3s	CCAAAATCCCCGCGGGTT	Amplificação	
22	B-R3s	GAGGCCATCCCCGCGAC	Éxon 3 de gene HLA-B	405bp
23	B-F4s	GCTGGTCACATGGGTGGTCCTA	Amplificação	
24	B-R4s	TGACCCCTCATCCCCCTCCT	Éxon 4 de gene HLA-B	374bp

[030] Os iniciadores degenerados referem-se a uma mistura de todas as

sequências diferentes possíveis representando todas as bases diferentes codificando um aminoácido único. A fim de aumentar a especificidade, a degeneração deve ser reduzida de acordo com a tendência do uso base nos diferentes organismos referindo-se a uma tabela códon, onde R=A/G, Y=C/T, M=A/C, K=G/T, S=C/G, W=A/T, H=A/C/T, B=C/G/T, V=A/C/G, D=A/G/T, N=A/C/G/T.

[031] A presente invenção designa 2 conjuntos de iniciadores de PCR (três pares para cada conjunto) para amplificação de Éxons 2, 3 e 4 de HLA-C usando o método de designação de Iniciadores de PCR para a amplificação de Éxons 2, 3 e 4 de gene HLA-A/B.

[032] Nos seguintes exemplos, 95 e 950 amostras de sangue com genótipos HLA conhecidos são submetidos à amplificação de PCR para HLA-C usando os dois conjuntos selecionados de Iniciadores de PCR (3 pares para cada conjunto), respectivamente. Os produtos amplificados são sequenciados pelo método Sanger e pelo método de sequenciamento de segunda geração. Os resultados de sequenciamento são aplicados à tipagem HLA-C, e são comparados com os resultados de tipagem original para confirmar o conservatismo e a especificidade dos Iniciadores de PCR.

[033] A presente invenção proporciona 2 conjuntos de Iniciadores de PCR (três pares para cada conjunto) para amplificação de Éxons 2, 3 e 4 de gene HLA-C, que são SEQ ID NOs: 25 e 26, 27 e 28, e 29 e 30 como mostrado na Tabela 3, e SEQ ID NOs: 31 e 32, 33 e 34, e 35 e 36 como mostrado na Tabela 4. Tais 6 pares de Iniciadores de PCR tem bom conservatismo e especificidade, e podem cobrir as sequências de comprimento completas para Éxons 2, 3 e 4 de HLA-C, onde o comprimento de todos os produtos de PCR é menor do que 700bp, que atende o requerimento de um sequenciamento Illumina Solexa normal. Além disso, os iniciadores da presente invenção também são aplicáveis para o sequenciamento Sanger.

Tabela 3: Iniciadores de PCR de Éxons 2, 3 e 4 de gene HLA-C

SEQ ID NO:	No.	Sequência primária	HLA-C Éxons	Comprimento dos produtos
25	C-F2	GACCCGGGGAGCCGCGCA	2	455bp
26	C-R2	TCGAGGGTCTGGGCGGGTT		
27	C-F3	CCTTTACCCGGTTTCATTTTCRGTTT	3	417bp
28	C-R3	CTACGGGAGATGGGGAAGGCT		
29	C-F4	GTGTCGCAAGAGAGATRCAAAGTGT	4	451bp
30	C-R4	GCTCTGGGAAAGGAGGGAAGG		

Tabela 4: Iniciadores de PCR de Éxons 2, 3 e 4 de gene HLA-C

SEQ ID NO:	No.	Sequência primária	HLA-C Éxons	Comprimento dos produtos
31	C-F2s	GACCCGGGGAGCCGCGCA	2	455bp
32	C-R2s	TCGAGGGTCTGGGCGGGTT		
33	C-F3s	GCCCAGACCCTCGRCCGGA	3	443bp
34	C-R3s	AGATRGGAAGGCTCCCCACT		
35	C-F4s	TCTCAGGATRGTCACATGGGC	4	481bp
36	C-R4s	GCTCTGGGAAARGAGGGAAGG		

[034] De acordo com os métodos como descrito acima, a fim de aplicar a técnica de sequenciamento de segunda geração à genotipagem HLA-DQB1, a presente invenção proporciona os Iniciadores de PCR para a amplificação de Éxons 2 e/ou 3

de HLA-DQB1, que são as SEQ ID NOs: 37-40 como mostrado na Tabela 5. Os Iniciadores de PCR têm um bom conservatismo e especificidade, e podem cobrir as sequências de comprimento completo de Éxons 2, 3 de HLA-DQB1, onde o comprimento de todos os produtos de PCR é menor do que 700bp, que atende ao requerimento de sequenciamento Illumina Solexa normal. Além disso, os iniciadores da presente invenção são também aplicáveis ao sequenciamento Sanger.

Tabela 5 Iniciadores de PCR para a amplificação dos Éxons de HLA-DQB1 correspondentes

SEQ ID NO:	No. iniciador	Sequência de iniciador	Amplificação de alvo	Comprimento dos produtos amplificados
37	Q-F2	GATTCCYCGCAGAGGATTTTCG	Éxon 2 de HLA-DQB1	311bp
38	Q-R2	AGGGGCRACSACGCTCACCTC		
39	Q-F3	CCTGTCTGTTACTGCCCTCAGT	Éxon 3 de HLA-DQB1	339bp
40	Q-R3	GGCCCATAGTAACAGAACTCAATA		

[035] A genotipagem pode ser realizada com base na amplificação de Éxons 2 e/ou 3 de HLA-DQB1 usando pares de iniciadores para a amplificação e métodos de genotipagem como proporcionado na presente invenção. Em relação a arte anterior, os métodos de genotipagem usam uma técnica de sequenciamento Illumina Solexa, que é **caracterizada** pela capacidade de obter uma alta resolução nos resultados de tipagem HLA com alto rendimento e baixo custo.

MODO ESPECÍFICO PARA REALIZAR A INVENÇÃO

[036] Um método para o sequenciamento de ácido nucléico

[037] Em um aspecto, a presente invenção proporciona um método para determinar uma sequência de nucleotídeo de um ácido nucléico de interesse em uma

amostra, compreendendo:

1) proporcionar n amostras, onde n é um número inteiro de ≥ 1 , as amostras são preferencialmente de mamíferos, mais preferencialmente de humanos, particularmente são amostras de sangue humano; opcionalmente as amostras n a serem analisadas são divididas em m grupos, m é um número inteiro e $n \geq m \geq 1$;

2) amplificação: um par ou múltiplos pares de iniciadores de indexação são usados para cada amostra, quando há modelos das amostras, a amplificação de PCR é realizada sob condições compatíveis para a amplificação do ácido nucléico de interesse, onde cada par de iniciadores de indexação consiste de um iniciador de indexação dianteiro e um iniciador de indexação reverso pode ser idêntico ou diferente: os iniciadores de indexação nos pares de iniciadores de indexação usados para amostras diferentes são diferentes;

3) associação: quando $n > 1$, associação de produtos de PCR de cada uma das amostras juntas;

4) cisalhamento: submeter os produtos amplificados a um cisalhamento incompleto, e purificar revestir;

5) sequenciamento: submeter a mistura de DNA recoberta ao sequenciamento usando a técnica de sequenciamento de segunda geração, preferencialmente, uma técnica de extremidade em pares (por exemplo, Illumina GA, Illumina HiSeq 2000), para obter as sequências do DNA cisalhado; e

6) montagem: correspondendo aos dados de sequenciamento obtidos nas amostras uma a uma com base no único iniciador de indexação para cada amostra, alinhando cada sequência lida à sequência de referência de DNA correspondente aos produtos de PCR usando um programa de alinhamento (tal como programa Blast, BWA), montando uma sequência completa de ácidos nucléicos de interesse a partir das sequências de DNA cisalhado em virtude da sobreposição da sequência e relação de ligação.

[038] Em um aspecto da presente invenção, cada par de iniciadores de indexação e um par de iniciadores de PCR a partir de um par de iniciadores de indexação, Iniciadores de PCR dianteiros e reversos tem um iniciador de indexação dianteiro e um iniciador de indexação reverso na 5' extremidade (ou opcionalmente ligada por uma sequência de ligação), respectivamente.

[039] Em uma modalidade da presente invenção, tais Iniciadores de PCR são Iniciadores de PCR para amplificação de um gene HLA, particularmente os Iniciadores de PCR para amplificação do gene HLA-A/B, preferencialmente os Iniciadores de PCR para a amplificação de Éxons 2, 3 e 4 de HLA-A/B e Éxon 2 de HLA-DRB1, preferencialmente de Iniciadores de PCR para amplificação de Éxons 2, 3 e 4 de HLA-A/B como mostrado na Tabela 1 ou Tabela 2, ou preferencialmente de Iniciadores de PCR para a amplificação de Éxon 2 de HLA-DRB1 como mostrado na Tabela 7.

[040] Em uma modalidade da presente invenção, tais Iniciadores de PCR são Iniciadores de PCR para a amplificação de gene HLA, particularmente os Iniciadores de PCR para a amplificação de gene HLA-C, preferencialmente os Iniciadores de PCR para a amplificação de Éxons 2, 3 e/ou 4 de HLA-C; preferencialmente, tais Iniciadores de PCR são mostrados na Tabela 3 ou Tabela 4.

[041] Em uma modalidade da presente invenção, tais Iniciadores de PCR são Iniciadores de PCR para a amplificação de gene HLA, preferencialmente os Iniciadores de PCR para a amplificação de Éxon 2 e/3 de gene HLA-DQB1; preferencialmente, tais Iniciadores de PCR são mostrados na Tabela 5.

[042] Em um aspecto da presente invenção, tais iniciadores de indexação são designados para os Iniciadores de PCR, preferencialmente para os Iniciadores de PCR para a amplificação de gene específico de HLA, mais preferencialmente para os Iniciadores de PCR para a amplificação de Éxons 2, 3 e 4 de HLA-A/B e Éxon 2 para HLA-DRB1, particularmente para os Iniciadores de PCR como mostrado na Tabela 1, Tabela 2 ou Tabela 7; tais iniciadores de indexação particularmente compreendem

pelo menos 10, ou pelo menos 20, ou pelo menos 30, ou pelo menos 40, ou pelo menos 50, ou pelo menos 60, ou pelo menos 70, ou pelo menos 80, ou pelo menos 90, ou 95 pares de 95 pares de iniciadores de indexação como mostrado na Tabela 6 (ou o conjunto de iniciadores de indexação consistindo de 10-95 pares (por exemplo, 10-95 pares, 20-95 pares, 30-95 pares, 40-95 pares, 50-95 pares, 60-95 pares, 70-95 pares, 80-95 pares, 90-95 pares, ou 95 pares) dos 95 pares de iniciadores de indexação como mostrado na Tabela 6); e o conjunto de iniciadores de indexação preferencialmente compreendem pelo menos PI-1 a PI-10, ou PI-11 a PI-20, ou PI-21 a PI-30, ou PI-31 a PI-40, ou PI-41 a PI-50, ou PI-51 a PI-60, ou PI-61 a PI-70, ou PI-71 a PI-80, ou PI-81 a PI-90, ou PI-91 a PI-95 de 95 pares de iniciadores de indexação como mostrado na Tabela 6, ou combinações de quaisquer dois ou mais deles.

[043] Em uma modalidade da presente invenção, tal cisalhamento de DNA inclui métodos de cisalhamento químico e métodos de cisalhamento físicos, onde os métodos de cisalhamento químicos incluem uma digestão enzimática, e os métodos de cisalhamento físico incluem métodos de cisalhamento ultrassônico ou métodos de cisalhamento mecânicos.

[044] Em uma modalidade da presente invenção, depois de tal cisalhamento de DNA, todas as faixas de DNA entre o comprimento lido máximo do sequenciador e o comprimento de DNA máximo aplicável do sequenciador são purificados e revestidos, onde tais métodos de purificação e recuperação incluem, mas não estão limitados a, recuperar por eletroforese e gel fatiado, e recuperação por esferas magnéticas.

[045] Em outra modalidade da presente invenção, um método para o sequenciamento da sequência de nucleotídeo de um ácido nucléico de interesse em uma amostra de teste, compreendendo a etapas de 1) a 4) da reivindicação 1, e as seguintes etapas:

5) construção de uma biblioteca: construir uma biblioteca de sequenciamento de PCR livre usando a biblioteca dos produtos de PCR cisalhados, onde diferentes

adaptadores de biblioteca podem ser adicionados para distinguir as diferentes bibliotecas de sequenciamento de PCR livre, todas as faixas de DNA entre o comprimento lido máximo do sequenciador e o comprimento de DNA máximo aplicável do sequenciador, preferencialmente fragmentos de DNA de 450 a 750 bp, são purificados e recuperados;

6) sequenciamento: submeter a recuperação da mistura de DNA ao sequenciamento usando a técnica de sequenciamento de segunda geração, preferencialmente a técnica de extremidade de pares (por exemplo, Illumina GA, Illumina HiSeq 2000), obtendo as sequências dos DNAs cisalhados;

7) montagem: correspondendo aos dados de sequenciamento obtidos para as amostras uma a uma com base nas diferentes sequências de adaptadores de biblioteca e o único iniciador de indexação para cada amostra, alinhando cada sequência lida à sequência de referência de DNA correspondente aos produtos de PCR usando um programa de alinhamento (tal como programa Blast, BWA), montar uma sequência completa de ácido nucléico de interesse a partir de sequências de DNA cisalhados com base na sobreposição de sequências e relacionamento de ligação.

[046] Em um aspecto, a presente invenção ainda proporciona o uso do método acima mencionado na tipagem de HLA, caracterizada pelo fato de compreender: sequenciar uma amostra (particularmente amostra de sangue) de um paciente portal método, e alinhar os resultados de sequenciamento com os dados de sequência de Éxons de HLA, preferencialmente, Éxons 2, 3, 4 de HLA-A/B, Éxons 2, 3 e/ou 4 de HLA-C, Éxon 2 e/ou 3 de gene HLA-DQB1 e/ou Éxon 2 de HLA-DRB1 em um banco de dados HLA (tal como banco de dados profissional IMGT HLA); onde se o resultado do alinhamento de sequência mostra 100% de equivalência, o genótipo HLA da amostra correspondente é determinado.

[047] Um conjunto de iniciadores de indexação

[048] Em outro aspecto, a presente invenção proporciona um conjunto de

iniciadores de indexação, compreendendo pelo menos 10, ou pelo menos 20, ou pelo menos 30, ou pelo menos 40, ou pelo menos 50, ou pelo menos 60, ou pelo menos 70, ou pelo menos 80, ou pelo menos 90, ou 95 pares de 95 pares de iniciadores de indexação como mostrado na Tabela 6 (ou o conjunto de iniciadores de indexação consistindo de 10-95 pares (por exemplo, 10-95 pares, 20-95 pares, 30-95 pares, 40-95 pares, 50-95 pares, 60-95 pares, 70-95 pares, 80-95 pares, 90-95 pares, ou 95 pares) dos 95 pares de iniciadores de indexação como mostrado na Tabela 6); e

o conjunto de iniciadores de indexação preferencialmente compreendem pelo menos PI-1 a PI-10, ou PI-11 a PI-20, ou PI-21 a PI-30, ou PI-31 a PI-40, ou PI-41 a PI-50, ou PI-51 a PI-60, ou PI-61 a PI-70, ou PI-71 a PI-80, ou PI-81 a PI-90, ou PI-91 a PI-95 de 95 pares de iniciadores de indexação como mostrado na Tabela 6, ou combinações de quaisquer dois ou mais deles.

[049] A presente invenção ainda proporciona o uso de tal conjunto de iniciadores de indexação nos métodos de sequenciamento de PCR, onde em particular, cada par de iniciadores de indexação e um par de Iniciadores de PCR para a amplificação de uma sequência de interesse a ser testada forma um par de iniciadores de indexação, onde os Iniciadores de PCR dianteiros e reversos na 5' extremidade (ou opcionalmente ligado por uma sequência de ligação), respectivamente.

[050] Em um aspecto da presente invenção, tais Iniciadores de PCR são Iniciadores de PCR para a amplificação de um gene específico de HLA, preferencialmente os Iniciadores de PCR para a amplificação de Éxons 2, 3, 4 de gene HLA-A/B e Éxon 2 de HLA-DRB1, preferencialmente os Iniciadores de PCR para a amplificação de Éxons 2, 3 e 4 de HLA-A/B como mostrado na Tabela 1 ou Tabela 2, ou preferencialmente os Iniciadores de PCR para a amplificação de Éxon 2 de HLA-DRB1 como mostrado na Tabela 7; ou preferencialmente os Iniciadores de PCR para a amplificação de Éxons 2, 3 e/ou 4 de HLA-C, preferencialmente tais Iniciadores de PCR são mostrados na Tabela 3 ou Tabela 4; ou preferencialmente os Iniciadores de PCR para

a amplificação de Éxon 2 e/ou 3 de HLA-DQB1, preferencialmente tais Iniciadores de PCR são mostrados na Tabela 5.

[051] Em outro aspecto, a presente invenção proporciona um conjunto de iniciadores de indexação compreendendo tais conjuntos de iniciadores de indexação e um par de Iniciadores de PCR para a ampliação de uma sequência de interesse a ser testada, onde um par de iniciadores de indexação compreende um par de iniciadores de indexação e um par de Iniciadores de PCR, o iniciador de PCR dianteiro e reverso tem um iniciador de indexação dianteiro e um reverso na 5' extremidade (ou opcionalmente ligada por uma sequência de ligador), respectivamente.

[052] Em outra modalidade da presente invenção, tais Iniciadores de PCR são Iniciadores de PCR para a amplificação de um gene específico de HLA, preferencialmente os Iniciadores de PCR para a amplificação de Éxons 2, 3, 4 de gene HLA-A/B e Éxon 2 de HLA-DRB1, preferencialmente os Iniciadores de PCR para a amplificação de Éxons 2, 3 e 4 de HLA-A/B como mostrado na Tabela 1 ou Tabela 2, ou preferencialmente os Iniciadores de PCR para a amplificação de Éxon 2 de HLA-DRB1 como mostrado na Tabela 7; preferencialmente os Iniciadores de PCR para a amplificação de Éxons 2, 3 e/ou 4 de HLA-C, preferencialmente tais Iniciadores de PCR são mostrados na Tabela 3 ou Tabela 4; ou preferencialmente os Iniciadores de PCR para a amplificação de Éxon 2 e/ou 3 de HLA-DQB1, preferencialmente tais Iniciadores de PCR são mostrados na Tabela 5.

[053] Em outro aspecto, a presente invenção ainda proporciona o uso de tais iniciadores de indexação nos métodos de sequenciamento de PCR.

[054] Um método de tipagem HLA

[055] Em um aspecto, a presente invenção proporciona um método de tipagem HLA, compreendendo:

1) proporcionar n amostras, onde n é um número inteiro de ≥ 1 , as amostras são preferencialmente de mamíferos, mais preferencialmente de humanos,

particularmente são amostras de sangue humano;

2) dividir as amostras n a serem analisadas em m grupos, m é um número inteiro e $n \geq m \geq 1$;

3) amplificação: um par ou múltiplos pares de iniciadores de indexação são usados para cada amostra, quando há modelos das amostras, a amplificação de PCR é realizada sob condições compatíveis para a amplificação do ácido nucléico de interesse, onde cada par de iniciadores de indexação consiste de um iniciador de indexação dianteiro e um iniciador de indexação reverso (ambos os quais podem ser iniciadores degenerados) compreendendo iniciadores de indexação, onde os iniciadores de indexação compreendidos no iniciador de indexação dianteiro e no iniciador de indexação reverso pode ser idêntico ou diferente: os iniciadores de indexação nos pares de iniciadores de indexação usados para amostras diferentes são diferentes;

4) associação: associar os produtos de PCR amplificados de cada uma das amostras juntas para obter as bibliotecas de produtos de PCR;

5) cisalhamento: submeter as bibliotecas de produtos de PCR resultantes a um cisalhamento incompleto;

6) construção de uma biblioteca: construir uma biblioteca de sequenciamento de PCR livre da biblioteca dos produtos de PCR cisalhados com uma técnica de índice de adaptador de biblioteca, onde diferentes adaptadores de biblioteca podem ser adicionados para distinguir as diferentes bibliotecas de sequenciamento de PCR livre, todas as faixas de DNA entre o comprimento lido máximo do sequenciador e o comprimento de DNA máximo aplicável do sequenciador, particularmente fragmentos de DNA de 450 a 750 bp, são recuperados;

7) sequenciamento: submeter a recuperação da mistura de DNA ao sequenciamento usando a técnica de sequenciamento de segunda geração, preferencialmente a técnica de extremidade de pares (por exemplo, Illumina GA, Illumina HiSeq 2000), obtendo as sequências dos DNAs cisalhados;

8) montagem: correspondendo aos dados de sequenciamento obtidos para as amostras uma a uma com base nas diferentes sequências de adaptadores de biblioteca das bibliotecas e o único iniciador de indexação para cada amostra, alinhando cada sequência lida à sequência de referência de DNA correspondente aos produtos de PCR usando um programa de alinhamento (tal como programa Blast, BWA), montar uma sequência completa de ácido nucléico de interesse a partir de sequências de DNA cisalhadas com base na sobreposição de sequências e relacionamento de ligação; e

9) tipagem: alinhamento dos resultados de sequenciamento com os dados de sequência de Éxons de HLA, preferencialmente, Éxons 2, 3, 4 de HLA-A/B, Éxons 2, 3 e/ou 4 de HLA-C, Éxon 2 e/ou 3 de gene HLA-DQB1 e/ou Éxon 2 de HLA-DRB1 no banco de dados HLA (tal como o banco de dados profissional IMGT HLA), onde se o resultado do alinhamento de sequência mostra 100% de equivalência, o genótipo HLA da amostra correspondente é determinado.

[056] No método de tipagem HLA da presente invenção, um par de iniciadores de indexação compreende um par de iniciadores de indexação e um par de Iniciadores de PCR, o iniciador de PCR dianteiro e o reverso tem um iniciador de indexação dianteiro e um reverso na 5' extremidade (ou opcionalmente ligada por uma sequência de ligação), respectivamente.

[057] Em uma modalidade da presente invenção, tais Iniciadores de PCR são Iniciadores de PCR para a amplificação de um gene específico de HLA, preferencialmente de Iniciadores de PCR para a amplificação de Éxons 2, 3, 4 de gene HLA-A/B e Éxon 2 de HLA-DRB1, preferencialmente de Iniciadores de PCR para a amplificação de Éxons 2, 3 e 4 de HLA-A/B como mostrado na Tabela 1 ou Tabela 2, ou preferencialmente Iniciadores de PCR para a amplificação de Éxon 2 de HLA-DRB1 como mostrado na Tabela 7; preferencialmente de Iniciadores de PCR para a amplificação de Éxons 2, 3 e/ou 4 de HLA-C, preferencialmente tais Iniciadores de PCR são

mostrados na Tabela 3 ou Tabela 4; ou preferencialmente de Iniciadores de PCR para a amplificação de Éxon 2 e/ou 3 de HLA-DQB1, preferencialmente tais Iniciadores de PCR são mostrados na Tabela 5.

[058] Em uma modalidade da presente invenção, tais iniciadores de indexação são conjuntos de iniciadores de indexação como descrito acima.

[059] Em uma modalidade do método de tipagem HLA da presente invenção, tal cisalhamento de DNA inclui métodos de cisalhamento químico e métodos de cisalhamento físicos, onde os métodos de cisalhamento químicos incluem uma digestão enzimática, e os métodos de cisalhamento físico incluem métodos de cisalhamento ultrassônico ou métodos de cisalhamento mecânicos.

[060] Em uma modalidade do método de tipagem HLA da presente invenção, tais métodos de purificação e recuperação incluem, mas não estão limitados a, recuperação por eletroforese e gel fatiado, e recuperação por esferas magnéticas.

[061] Em uma modalidade do método de tipagem HLA da presente invenção, a construção das bibliotecas de sequenciamento de PCR livre das bibliotecas dos produtos de PCR cisalhados com as técnicas de indexação de adaptadores de biblioteca compreende, adicionar adaptadores de biblioteca m às bibliotecas de produtos de PCR m obtidos em 2), onde cada uma das bibliotecas de produtos de PCR usa um adaptador de biblioteca diferente, construindo assim, bibliotecas de sequenciamento indexado de adaptador m; bibliotecas de sequenciamento indexado de adaptadores m são associados juntos em um mol igual para construir uma mistura de bibliotecas de sequenciamento indexado de adaptadores, onde o método para ligar os adaptadores de biblioteca referem-se à ligação direta usando uma ligase DNA sem um procedimento de PCR.

[062] Iniciadores de PCR para a genotipagem HLA

[063] Em um aspecto, a presente invenção proporciona Iniciadores de PCR para a genotipagem HLA, caracterizados por aqueles tais Iniciadores de PCR são

Iniciadores de PCR para a amplificação de Éxons 2, 3, 4 de gene HLA-A/B e Éxon 2 de HLA-DRB1, preferencialmente de Iniciadores de PCR para a amplificação de Éxons 2, 3 e 4 de HLA-A/B como mostrado na Tabela 1 ou Tabela 2, ou preferencialmente Iniciadores de PCR para a amplificação de Éxon 2 de HLA-DRB1 como mostrado na Tabela 7; preferencialmente de Iniciadores de PCR para a amplificação de Éxons 2, 3 e/ou 4 de HLA-C, preferencialmente tais Iniciadores de PCR são mostrados na Tabela 3 ou Tabela 4; ou preferencialmente de Iniciadores de PCR para a amplificação de Éxon 2 e/ou 3 de HLA-DQB1, preferencialmente tais Iniciadores de PCR são mostrados na Tabela 5.

[064] A presente invenção ainda proporciona um método de sequenciamento usando tais Iniciadores de PCR, compreendendo:

proporcionar uma amostra, particularmente uma amostra de sangue, tal amostra de sangue é preferencialmente de um mamífero, particularmente de um humano;

amplificação: amplificar o DNA das amostras de sangue com os Iniciadores de PCR para obter os produtos de PCR, e purificar os produtos de PCR;

sequenciamento: submeter os produtos de PCR ao sequenciamento, o método de sequenciamento pode ser um método de sequenciamento Sanger, ou um método de sequenciamento de segunda geração (tal como Hiseq 2000, Illumina GA e Roche 454).

[065] Em outro aspecto, a presente invenção ainda proporciona a uso de tais Iniciadores de PCR na genotipagem HLA, caracterizado pelo uso de tais Iniciadores de PCR, realizando a análise de alinhamento e montagem nos resultados obtidos pelo método de sequenciamento acima, e comparando os resultados de sequenciamento com as sequências padrão no banco de dados para obter os resultados de genotipagem HLA.

[066] Em outro aspecto, a presente invenção ainda proporciona um conjunto

para a genotipagem HLA, compreendendo tais Iniciadores de PCR.

[067] Iniciadores de PCR para a genotipagem HLA-A,B

[068] Em um aspecto, a presente invenção proporciona um conjunto de Iniciadores de PCR para a genotipagem HLA-A,B, compreendendo:

proporcionar uma amostra, particularmente uma amostra de sangue, tal amostra de sangue é preferencialmente de um mamífero, particularmente de um humano;

amplificação: amplificar o DNA das amostras de sangue com os Iniciadores de PCR para obter os produtos de PCR, e purificar os produtos de PCR;

sequenciamento: submeter os produtos de PCR ao sequenciamento, o método de sequenciamento pode ser um método de sequenciamento Sanger, ou um método de sequenciamento de segunda geração (tal como Hiseq 2000, Illumina GA e Roche 454).

[069] Em outro aspecto, a presente invenção ainda proporciona o uso de tais Iniciadores de PCR na genotipagem HLA, caracterizada pelo uso de tais Iniciadores de PCR, realizando a análise de alinhamento e montagem nos resultados obtidos pelo método de sequenciamento acima, e comparando os resultados de sequenciamento com as sequências padrão no banco de dados para obter os resultados de genotipagem HLA.

[070] Em outro aspecto, a presente invenção ainda proporciona um conjunto para a genotipagem HLA, compreendendo tais Iniciadores de PCR para a genotipagem HLA-A,B da presente invenção.

[071] Iniciadores de PCR para a genotipagem HLA-C

[072] A presente invenção ainda proporciona um novo método para a amplificação de Éxons 2, 3 e 4 de gene HLA-C, caracterizado por realizar uma amplificação de PCR usando os pares de iniciadores de amplificação da presente invenção, as sequências dos pares de iniciadores de amplificação são mostradas na Tabela 3 ou

Tabela 4.

[073] Desde que os Éxons 2, 3 e 4 de HLA-C possam ser amplificados por uma reação de PCR, o método da presente invenção é particularmente compatível para a genotipagem HLA-C. quando comparado com os métodos de genotipagem HLA-C anteriores, uma vez que os produtos obtidos usando o método e as amplificações primárias da presente invenção são controlados dentro de 700 bp, o HLA-SBT com base na técnica de sequenciamento Illumina Solexa pode ser usado ainda durante a genotipagem.

[074] A presente invenção ainda proporciona um método para o sequenciamento de Éxons 2, 3 e 4 de gene HLA-C nas amostras, compreendendo as seguintes etapas de:

- 1) proporcionar uma amostra e extrair o DNA da amostra;
- 2) amplificar o DNA com o par de iniciadores de PCR para a genotipagem HLA-C da presente invenção para obter os produtos de PCR, preferencialmente purificar os produtos de PCR, tais pares de iniciadores de PCR são preferencialmente selecionados a partir de um grupo que consiste de um par de iniciadores de SEQ ID NO: 25 e SEQ ID NO: 26, SEQ ID NO: 27 e SEQ ID NO: 28, SEQ ID NO: 29 e SEQ ID NO: 30, ou SEQ ID NO: 31 e SEQ ID NO: 32, SEQ ID NO: 33 e SEQ ID NO: 34, SEQ ID NO: 35 e SEQ ID NO: 36;
- 3) submeter os produtos de PCR ao sequenciamento, preferencialmente pelo método de sequenciamento de segunda geração, tais como Illumina Solexa ou Roche454.

[075] A presente invenção ainda proporciona um método de genotipagem HLA-C, compreendendo:

- 1) amplificação de PCR Éxons 2, 3 e/ou 4 de gene HLA-C da amostra a ser testada com o par de iniciadores PCR para a genotipagem HLA-C da presente invenção, tal par de iniciadores de PCR é preferencialmente selecionado a partir de um

grupo que consiste de um par de iniciadores de SEQ ID NO: 25 e SEQ ID NO: 26, SEQ ID NO: 27 e SEQ ID NO: 28, SEQ ID NO: 29 e SEQ ID NO: 30, ou SEQ ID NO: 31 e SEQ ID NO: 32, SEQ ID NO: 33 e SEQ ID NO: 34, SEQ ID NO: 35 e SEQ ID NO: 36;

2) submeter os Éxons amplificados ao sequenciamento, comparando os resultados de sequenciamento com as sequências padrão no banco de dados de forma a determinar os resultados de genotipagem, onde o sequenciamento é realizado por um método de sequenciamento Sanger, ou um método de sequenciamento de segunda geração, tal como Illumina GA e Roche 454.

[076] Em outro aspecto, a presente invenção ainda proporciona um conjunto para a genotipagem HLA-C, compreendendo um par de iniciadores de PCR para a genotipagem HLA- da presente invenção, preferencialmente selecionado a partir de um grupo que consiste de um par de iniciadores de SEQ ID NO: 25 e SEQ ID NO: 26, SEQ ID NO: 27 e SEQ ID NO: 28, SEQ ID NO: 29 e SEQ ID NO: 30, ou SEQ ID NO: 31 e SEQ ID NO: 32, SEQ ID NO: 33 e SEQ ID NO: 34, SEQ ID NO: 35 e SEQ ID NO: 36. Em outra modalidade, tal conjunto ainda compreende agentes adicionais, por exemplo, agentes para a amplificação de DNA, a purificação de DNA, e/ou o sequenciamento de DNA.

[077] A genotipagem pode ser realizada com base na amplificação de Éxons 2, 3 e 4 de HLA-C, usando o par de iniciadores de amplificação e o método de genotipagem como proporcionado na presente invenção. Por isso, quando comparado com a arte anterior, a genotipagem utiliza a técnica de sequenciamento Illumina Solexa, aumenta o rendimento, simplifica o procedimento, e por sua vez economiza tempo e custo.

[078] Iniciadores de PCR para a genotipagem HLA-DQB1

[079] A presente invenção ainda proporciona um novo método para a amplificação de Éxon 2, e/ou 3 de HLA-DQB1, caracterizado por realizar uma amplificação

de PCR com os pares de iniciadores de amplificação da presente invenção, tais pares de iniciadores de amplificação são mostrados na Tabela 5.

[080] Desde que os Éxons 2, e/ou 3 de HLA-DQB1 possam ser amplificados por uma reação de PCR, o método da presente invenção é particularmente compatível para a genotipagem HLA-DQB1. Quando comparado com os métodos de genotipagem HLA-DQB1 anteriores, uma vez que os produtos obtidos usando o método e as amplificações primárias da presente invenção são controlados dentro de 300-400 bp, o HLA-SBT com base na técnica de sequenciamento Illumina Solexa pode ser usado ainda durante a genotipagem.

[081] A presente invenção ainda proporciona um método para o sequenciamento de Éxon 2 e/ou 3 de HLA-DQB1 nas amostras, compreendendo as seguintes etapas de:

- 1) proporcionar uma amostra e extrair o DNA da amostra;
- 2) amplificar o DNA com o par de iniciadores de PCR para a genotipagem HLA-DQB1 da presente invenção, preferencialmente os pares de iniciadores de PCR mostrados na Tabela 5, para obter os produtos de PCR, preferencialmente purificar os produtos de PCR;
- 3) submeter os produtos de PCR ao sequenciamento, preferencialmente pelo método de sequenciamento de segunda geração, tais como Illumina Solexa ou Roche454.

[082] Em outro aspecto da presente invenção, a presente invenção proporciona um método aperfeiçoado para a genotipagem HLA-DQB1, compreendendo:

- 1) amplificar o Éxon 2 e/ou 3 de HLA-DQB1 para ser testado com os pares de iniciadores de PCR para a genotipagem HLA-DQB1 da presente invenção, preferencialmente os pares de iniciadores de PCR como mostrado na Tabela 5;
- 2) submeter os Éxons amplificados ao sequenciamento, comparando os resultados de sequenciamento com as sequências padrão no banco de dados de forma

a determinar os resultados de genotipagem, onde o método de sequenciamento pode ser um método de sequenciamento Sanger, ou o método de sequenciamento de segunda geração, tal como Illumina GA e Roche 454.

[083] Em outro aspecto, a presente invenção ainda proporciona um conjunto para a genotipagem HLA-DQB1, compreendendo um par de iniciadores de PCR para a genotipagem HLA-DQB1 da presente invenção, preferencialmente, os pares primários de amplificação de PCR como mostrados na Tabela 5. Em uma modalidade, tal conjunto ainda compreende agentes adicionais, por exemplo, agentes para a amplificação de DNA, a purificação de DNA, e/ou o sequenciamento de DNA.

DESCRIÇÃO DOS DESENHOS

[084] Figura 1: Um desenho ilustra a montagem da sequência depois de legendada com os iniciadores de indexação, o cisalhamento de DNA e o sequenciamento de DNA. As sequências de iniciador de indexação dianteiro e reverso Index-N-F/R (1) são introduzidos a dois terminais dos produtos de PCR da amostra No. N. os produtos de PCR depois do cisalhamento pelo método de cisalhamento físico, compreendem produtos que carregam as sequências de iniciador de indexação em uma extremidade, produtos carregando nenhuma sequência de iniciador de indexação nos dois terminais, e produtos não cisalhados completamente. Todas as faixas de DNA entre o comprimento lido máximo do sequenciador e o comprimento de DNA máximo aplicável são purificados e recuperados por gel fatiado, e usados para o sequenciamento (2). Os dados de sequenciamento dos produtos de PCR pertencentes a amostra No. N são traçadas usando o Index-N-F/R. As sequências de referência conhecidas dos produtos de PCR são usadas para localizar as posições relativas das sequências lidas, e os resultados de sequenciamento dos produtos de PCR completos são montados com base na sobreposição e na relação de ligação entre as sequências lidas (3, 4).

[085] Figura 2: Um desenho ilustrando os resultados da eletroforese dos

produtos PCR dos Éxons correspondentes de HLA-A/B/DRB1 na amostra No. 1 do Exemplo 2. Pode ser visto do eletroforetograma que os produtos PCR são uma série de faixas únicas de 300bp-500bp, onde a Linha M é um marcador de peso molecular ((DL 2000, Takara Co.), as Linhas 1-7 são os produtos de PCR dos Éxons (A2, A3, A4, B2, B3, B4, DRB1-2) de HLA-A/B/DRB1 da amostra No.1, e não há amplificação de faixa no controle negativo (N). os resultados das outras amostras são similares.

[086] Figura 3: Um desenho ilustrando os resultados de uma eletroforese de DNA depois do cisalhamento HLA-Mix no Exemplo4 (antes e depois do gel fatiado), onde a área de gel fatiado é a área de 450-750bp. A Linha M é um marcador de peso molecular (NEB-50bp DNA escada), e a Linha 1 mostra resultado eletroforético de HLA-Mix antes do gel fatiado, e a Linha 2 é um desenho mostrando o gel de HLA-Mix depois de fatiado.

[087] Figura 4: Uma captura de tela de um programa para a construção de esquema de consenso da Amostra No. 1 no Exemplo 6, ilustrando a montagem da sequência completa dos produtos de PCR com base nos iniciadores de indexação e na relação de sobreposição entre os fragmentos de DNA. Por favor, refira-se ao <http://www.ebi.ac.uk/imgt/hla/align.html> para a nomenclatura de genótipos HLA. Uma pessoa pode encontrar os resultados de todos os códigos de sequências de A*02:03:01 A*11:01:01 no resultado de coluna de saída à esquerda, onde a sequência de Éxon 2 é idêntica ao resultado original conhecido no Modelo 1.

[088] Figura 5: Um desenho ilustrando o produto de PCR depois de legendado com os iniciadores de indexação e o índice de adaptador. Durante a experimentação, os iniciadores de indexação são introduzidos nos dois terminais do produto de PCR de cada amostra por PCR simultaneamente; os produtos de PCR múltiplos carregando diferentes iniciadores de indexação são associados juntos para construir uma biblioteca de sequenciamento múltipla. Durante a construção das bibliotecas de sequenciamento múltiplas, quando a biblioteca de sequenciamento múltipla tem sido

construída, as bibliotecas de sequenciamento podem ser legendadas com os adaptadores de biblioteca carregando diferentes índices de adaptadores. Depois de terminar a construção das bibliotecas, as bibliotecas de sequenciamento múltiplas legendadas com diferentes índices de adaptadores são associadas juntas e são sequenciadas por Illumina GA simultaneamente (os iniciadores de indexação podem ser idênticos entre as bibliotecas de sequenciamento legendadas com diferentes índices de adaptadores). Depois de pegar os resultados de sequenciamento, a informação de sequência de DNA par cada amostra pode ser obtida triando as informações de sequências dos índices de adaptadores e os iniciadores de indexação nos resultados de sequenciamento.

[089] Figura 6: Um desenho ilustrando o resultado eletroforético dos produtos de PCR de Éxons 2, 3, 4 de HLA-C de algumas amostras no Exemplo 8. Pode ser visto pelo do eletroforetograma que os produtos PCR são uma série de faixas únicas de 400bp-500bp, onde a Linha M é um marcador de peso molecular de DNA padrão (DL 2000, Takara Co.).

[090] Figura 7: Um desenho ilustrando os resultados de gel fatiado eletroforético de DNA depois do cisalhamento HLA-Mix no Exemplo 8, onde a área de gel fatiado é a área de 450-750 bp. A Linha M é um marcador de peso molecular (NEB-50bp DNA escada), e a Linha 1 é um desenho mostrando o gel de HLA-Mix antes de fatiado, e a Linha 2 é um desenho mostrando o gel de HLA-Mix depois de fatiado.

[091] Figura 8: Uma captura de tela do programa de construção de sequência de consenso de Éxon 2 do local HLA-C da Amostra No. 2 no Exemplo 8. Primeiramente, as sequências lidas do local C da amostra é alinhado com a sequência de referência pelo software BWA, construindo assim as sequências de consenso de Éxons 2, 3, 4 de local C da amostra; ainda, a sequência de haplótipo de cada Éxon de local C é determinado com base na relação de ligação entre SNPs; e finalmente o tipo da amostra é determinado pela interseção das sequências de haplótipos dos

Éxons. Como mostrado na figura, dois heterozigotos SNP são compreendidos nas áreas 695-764 da sequência de gene C da amostra No. 2, e pode ser determinada pela leitura 1 ou pela leitura 2 que a relação de ligação de SNP seja A-C, G-A (“...” na figura representa as bases idênticas a aquelas da sequência de referência). A sequências correspondem as partes sombreadas das sequências dos tipos C*010201 e C*07020101, respectivamente. O julgamento das relações de ligação de outras áreas é similar.

[092] Figura 9: Um desenho ilustrando os resultados de eletroforese de produtos de PCR de Éxons 2, 3 e 4 de local HLA-C de 26 amostras no Exemplo 9. Como mostrado na figura, todos os produtos de PCR são de um comprimento menor do que 500bp; a banda eletroforética é única; não há banda não específica óbvia; e a eficácia na amplificação do mesmo par de iniciadores é o mesmo em várias amostras.

[093] Figura 10: Um desenho ilustrando os resultados analíticos dos dados de sequenciamento de produtos de amplificação de PCR do Modelo 1 usando um software Tipo-u no Exemplo 9. O resultado da coluna de saída à esquerda mostra o resultado, C*08:01:01 C*15:05:01, que são idênticos ao tipo original conhecido no modelo 1.

[094] Figura 11: Um desenho ilustrando o resultado eletroforético dos produtos PCR de Éxon 2+3 de HLA-DQB1 em 94 amostras do Exemplo 10. Pode ser visto do eletroforetograma que os produtos de PCR são uma série de bandas únicas de 250-500bp, onde a Linha M é referência para os pesos moleculares de DNA padrão (DL 2000, Takara Co.), Linhas PI-1 a PI-94 são produtos de amplificação de PCR de Éxon 2+3 de HLA-DQB1 em 94 amostras, e não há banda de amplificação no controle negativo (N).

[095] Figura 12 mostra os resultados de um gel de fatiado eletroforético de DNA depois do cisalhamento HLA-Q-Mix no Exemplo 10, onde a área de gel fatiado é uma área de 350-550bp. A Linha M é um marcador de pesos moleculares de DNA

padrão (NEB-50bp DNA Escada), e a Linha 1 é um desenho mostrando o gel de HLA-Q-Mix antes de fatiado, e a Linha 2 é um desenho mostrando o gel de HLA-Q-Mix depois de fatiado.

[096] A figura 13 mostra uma captura de tela do programa para a construção da sequência de consenso da Amostra No. 7 no Exemplo 10, ilustrando o procedimento principal da análise de dados. Primeiramente, as sequências lidas do local DQB1 da amostra são alinhados com a sequência de referência, por um software BWA, construindo assim, as sequências de consenso de Éxons 2, 3 de DQB1 da amostra; e as sequências de haplótipo de Éxons 2, 3 de DQB1 são determinadas com base na relação de ligação entre SNPs. Como mostrado na figura, seis heterozigotos SNPs são compreendidos na área 2322-2412 da sequência de gene DQB1 da Amostra No. 7, e pode ser determinada a partir da leitura 1 que o relacionamento de ligação de SNP1-SNP5 é T-G-T-C-C; pode ser determinada pela leitura 2 que o relacionamento de ligação de outro SNP1-SNP5 é C-C-A-G-T; pode ser determinado que a partir da leitura 3 que o relacionamento de ligação de SNP3-SNP6 é A-G-T-G; pode ser determinado que a partir da leitura 4 o relacionamento de ligação de outro SNP3-SNP6 é T-C-C-A; e pode ser determinado que a partir dos relacionamentos de ligação acima de tais SNPs que a leitura 1 é ligada à leitura 4, leitura 2 é ligada ao leitor 3, a combinação de SNP completa nesta área é T-G-T-C-C-A e C-C-A-G-T-G, e as sequências correspondem a partes sombreadas das sequências do tipo DQB1*0303 e DQB1*0602. O julgamento do relacionamento para as outras áreas é similar.

[097] A figura 14 mostra o eletroforetograma dos produtos no Exemplo 11, resultados da amplificação de cada Éxons 2 e 3 de local HLA-DQB1 e amplificação dos Éxons 2 e 3 com dois pares de Iniciadores de PCR, respectivamente. O eletroforetograma mostra três conjuntos de produtos de PCR de sete modelos de DNA, onde todos os produtos de DNA tem um comprimento menor do que 500bp; as faixas eletroforéticas são únicas; e não há uma banda não específica óbvia. Não há

amplificação no controle negativo (N), e a Linha M é referência para os pesos moleculares de DNA padrão (DL2000, Takara Co.).

[098] A figura 15 ilustra os resultados analíticos dos dados de sequenciamento de produtos PCR resultados da amplificação de Éxons 2 e 3 de HLA-DQB1 do Modelo 7, usando um software tipo-u no Exemplo 11. O resultado da coluna de saída à esquerda mostra o resultado, DQB1*03:03 DQB1*06:02, que é idêntico ao resultado original conhecido do Modelo 7.

[099] A figura 16 mostra os resultados eletroforéticos dos produtos PCR dos Éxons correspondentes HLA-A/B/C/DQB1 na amostra No. 1 no Exemplo 12. Pode ser visto a partir do eletroforetograma que os produtos de PCR são uma série de faixas únicas de 300bp-500bp, onde a Linha M é um marcador de pesos moleculares (DL 2000, Takara Co.); as Linhas 1-10 são produtos de PCR amplificados dos Éxons (A2, A3, A4, B2, B3, B4, C2, C3, C4, DQB1) de HLA-A/B/C/DQB1 da Amostra No. 1; nenhuma faixa de amplificação está presente no controle negativo (N). Os resultados de outras amostras são similares.

[0100] A figura 17 ilustra o resultado da recuperação de um gel agarose depois de agrupar HLA-1-Mix, HLA-2-Mix, HLA-3-Mix, HLA-4-Mix, HLA-5-Mix, HLA-6-Mix, HLA-7-Mix, HLA-8-Mix, HLA-9-Mix e HLA-10-Mix em mol iguais no Exemplo 12. A linha M é um marcador de pesos moleculares, e a Linha 1 é um resultado eletroforetograma, e a Linha 2 é uma Linha eletroforetograma depois de fatiado o gel contendo os fragmentos de DNA da média de comprimento de 450 a 750 bp.

[0101] A figura 18 mostra uma captura de tela do programa para a construção de sequência de consenso de Éxon 2 do local HLA-C da Amostra No. 1 no Exemplo 12. Primeiramente, as sequencias lidas do local C da amostra são alinhadas com a sequência de referência por um software BWA, construindo assim as sequências de consenso de Éxons 2, 3, 4 de locais C da amostra; ainda, as sequências de halótipos dos Éxons de local C são determinadas com base no relacionamento de ligação entre

os SNPs; e finalmente o tipo da amostra é determinado pela interseção das sequências de halótipos dos Éxons. Como mostrado na figura, os dois heterozigotos SNPs são compreendidos na área 695-764 da sequência de gene C da Amostra No. 1, e pode ser determinado a partir da leitura 1 e leitura 2 que o relacionamento de ligação de SNPs seja A-C, G-A (“...” na figura representa as bases idênticas àquelas da sequência de referência). As sequências correspondem as partes compartilhadas das sequências do tipo C*010201 e C*07020101, respectivamente. O julgamento do relacionamento de ligação de outras áreas seja similar.

EXEMPLOS

[0102] As modalidades da presente invenção são descritas em detalhes nos seguintes exemplos. Entretanto, um perito na arte entenderia que os seguintes exemplos são usados para ilustrar a presente invenção invés de limitar a restrição do escopo da presente invenção.

[0103] Nos exemplos de 1-6 da presente invenção, os Éxons 2, 3, 4 de HLA-A/B e Éxon 2 de HLA-DRB1 em 95 amostras foram genotipadas usando a combinação dos iniciadores de indexação + estratégia de cisalhamento de DNA incompleto + técnica de sequenciamento 100 de sequenciador de extremidade em pares Illumina GA (produtos de PCR tem um comprimento na faixa de 290bp a 500bp), demonstrando que o método da invenção pode realizar a tipagem dos fragmentos de gene de um comprimento excedendo o máximo comprimento lido do sequenciador enquanto utilizando suficientemente as características do sequenciador de segunda geração, tais como alto rendimento e baixo custo.

[0104] Princípios: para a amostra ser analisada, os iniciadores de indexação foram introduzidos aos dois terminais dos produtos de PCR de Éxons 2, 3, 4 de HLA-A/B e Éxon 2 de HLA-DRB1 por uma reação de PCR de forma a especificamente legendar a informação da amostra dos produtos de PCR. Os produtos de amplificação de PCR de três locais (HLA-A/B/DRB1) em cada grupo de amostras foram associados

juntos para obter uma biblioteca de produtos PCR; depois do cisalhamento ultrassônico incompleto da biblioteca de produtos de PCR, uma biblioteca de sequenciamento de PCR-livre foi construída. A biblioteca de sequenciamento foi submetida a uma eletroforese 2% baixo ponto de derretimento de gel agarose, e todas as faixas de DNA de um comprimento na faixa de 450bp a 750bp foram purificados e recuperados por um gel fatiado (durante a construção da biblioteca de sequenciamento de PCR-livre, uma vez que os adaptadores de biblioteca foram adicionados aos dois terminais dos fragmentos de DNA, o comprimento de uma faixa de DNA como mostrado no eletroforetograma foi de cerca de 250 bp maior do que o comprimento atual dos fragmentos de DNA; portanto, os fragmentos de um comprimento na faixa de 450bp a 700bp como recuperados aqui na realidade correspondem aos fragmentos de DNA de um comprimento original na faixa de 200 bp a 500 bp). O DNA recuperado foi sequenciado por Illumina GA PE-100. A informação de sequência de todas as amostras testadas pode ser traçada pelas sequências de iniciador de indexação, e as sequências de todo o produto de PCR podem ser montadas com base nas sequências de referências conhecidas e na sobreposição e no relacionamento de sobreposição entre as sequências de fragmentos de DNA, a sequência completa dos produtos de PCR originais pode ser montada com o banco de dados padrão dos Éxons correspondentes de HLA-A/B/DRB1, realizando assim a genotipagem HLA-A/B/DRB1.

Exemplo 1

[0105] Extração da amostra

[0106] Os DNAs foram extraídos de 95 amostras de sangue com resultados de tipagem HLA-SBT conhecidos (Programa de Doadores de Medula Óssea da China citado aqui como (CMDP)) usando um Instrumento de Extração Automático KingFisher (US Thermo Co.). As etapas principais foram como segundo: como direcionado no livro, uma certa quantidade de agentes auto contidos foi adicionada a seis placas bem profundas e uma placa bem rasa equipada pelo Instrumento de Extração

Automático KingFisher e todas as placas, as quais os agentes foram adicionados, foram colocadas nas posições correspondentes como requerido. O programa “Bioeasy_200ul Blood DNA_KF.ms2” foi selecionado, e foi implementado para extrair os ácido nucléicos pressionando “começar”. Aproximadamente 100 µl de produtos eluídos (isto é, o DNA extraído) foram coletados da placa de Eluição depois que o programa terminou.

Exemplo 2

[0107] A amplificação de PCR

[0108] Os diferentes iniciadores de indexação de PCR foram feitos sintetizando os Iniciadores de PCR tendo diferentes iniciadores de indexação na 5' extremidade, e tais diferentes iniciadores de indexação de PCR podem ser aplicados a diferentes amostras, onde os Iniciadores de PCR foram Iniciadores de PCR para Éxons 2, 3, 4 de HLA-A/B e Éxon 2 de HLA-DRB1. Portanto, os iniciadores de indexação foram introduzidos aos dois terminais d produtos PCR por uma reação de PCR, assim especificamente legendando os produtos de PCR de diferentes amostras.

[0109] 95 conjuntos de iniciadores de indexação de PCR foram usados para amplificar 95 amostras de DNA, respectivamente, onde cada conjunto de iniciadores de indexação de PCR consistiam de um par de iniciadores de indexação bidirecionais (Tabela 6) e Iniciadores de PCR para a amplificação de Éxons 2, 3, 4 de HLA-A/B (Tabela 1) e de Éxon 2 de HLA-DRB1 (Tabela 7), cada Iniciador de PCR dianteiro tem o iniciador de indexação dianteiro no par de iniciadores de indexação ligado na 5' extremidade, e o iniciador de PCR reverso tem um iniciador de indexação reverso no par de iniciadores de indexação ligado na 5' extremidade. Durante a síntese dos primários, os iniciadores de indexação foram diretamente adicionados a 5' extremidade dos Iniciadores de PCR.

[0110] Os 95 DNAs obtidos das etapas de extração de amostra do Exemplo 1 foram designadas como No. 1-95. A reação de PCR aconteceu em 96 placas boas, 7

placas no total, designadas como HLA-P-A2, HLA-P-A3, HLA-P-A4, HLA-P-B2, HLA-P-B3, HLA-P-B4 e HLA-P-DRB1-2 (A2/A3/A4, B2/B3/B4, DRB1-2 representam os locais amplificados), onde um controle negativo sem a adição de qualquer modelo foi configurada em cada placa, e os iniciadores usados no controle negativo foram os mesmos daqueles usados no Modelo 1. Durante a experimentação, a quantidade de informações das amostras correspondentes a cada par de iniciadores de indexação foi gravada.

Tabela 6: Informações Relevantes dos Iniciadores de indexação

Iniciador de indexação No.	SEQ ID NO:	Iniciador de indexação dianteiro	SEQ ID NO:	Iniciador de indexação reverso	A posição correspondente na placa boa -96	Modelo correspondente (Grupo 1)
PI-1	41	TCGCAGACATCA	42	TGACACGATGCT	A1	1
PI-2	43	TACATCGCACTA	44	TACAGATGCTGA	A2	2
PI-3	45	CTCGATGAGTAC	46	ACGTCTAGACAC	A3	3
PI-4	47	TCTGTATACTCA	48	TGCTGTAGTGAC	A4	4
PI-5	49	TATCTGCTCATA	50	AGATATCGAGCT	A5	5
PI-6	51	TACATGCTGAGC	52	ACGTGTCTATCA	A6	6
PI-7	53	TCATATCGCGAT	54	AGATCGTATAGC	A7	7
PI-8	55	ACAGATGCACGC	56	ATCTCGTGACAG	A8	8
PI-9	57	TAGATCGTACAT	58	ACTAGTACACGC	A9	9
PI-10	59	ACTACACGTCTC	60	ATAGTCACGCGT	A10	10
PI-11	61	AGACTCGCGTAT	62	TACTAGCTGACG	A11	11
PI-12	63	ATACTAGTGCTC	64	TGTATCGTGCTC	A12	12
PI-13	65	CACGATGACATC	66	TAGTGAGCGCAC	B1	13
PI-14	67	TGCTGTCTCGAG	68	CATAGCAGTGTC	B2	14
PI-15	69	TGTGCTCGAGTC	70	TCTGATCGAGCA	B3	15
PI-16	71	CACTCGTACATC	72	AGCGATGCTCAT	B4	16

PI-17	73	CGACGTGCTCGC	74	CGCGTACTGCAG	B5	17
PI-18	75	ACGCATCTATAC	76	CTAGTATCGCAG	B6	18
PI-19	77	CGAGATGACTCT	78	TGTATACACGAT	B7	19
PI-20	79	ACTGTCTCGAGC	80	ACGTAGCGCACA	B8	20
PI-21	81	CATCTGCTATAG	82	TCTAGCTCATGA	B9	21
PI-22	83	ACGCACTCTAGA	84	CTATGCACTGAT	B10	22
PI-23	85	TGAGATACAGTA	86	ATCTGCTATGAC	B11	23
PI-24	87	ACTCATCGTGCT	88	TAGAGCTGTCAC	B12	24
PI-25	89	TAACTGTCTAT	90	CAGCACATAGAT	C1	25
PI-26	91	CACAGTACTCGC	92	CTGCTAGTGTAT	C2	26
PI-27	93	TGTACTATCATA	94	TGTGATAGACAC	C3	27
PI-28	95	CTAGTACTGACG	96	AGCGAGTCTACT	C4	28
PI-29	97	TAGACTGAGCTA	98	ACATACTGAGAC	C5	29
PI-30	99	CAGACGCGTGAG	100	TACATCTCGTAT	C6	30
PI-31	101	CGCGACATCACG	102	TAGCGATGAGAC	C7	31
PI-32	103	AACTCATAGAT	104	CTATCATGACAC	C8	32
PI-33	105	AGCGTATACTAG	106	CATACTCACGTA	C9	33
PI-34	107	TGTCGTGCTATC	108	ACATGACTCACG	C10	34
PI-35	109	CGCTAGACTGTA	110	TACTATAGTCGA	C11	35
PI-36	111	ACAGTGTAGCGC	112	TGATATGCTACA	C12	36
PI-37	113	CACTCTATCGAC	114	TCACGCGATGAG	D1	37
PI-38	115	AACTCTAGTCA	116	ACGTAGATCTAT	D2	38
PI-39	117	CATATGAGATCG	118	AGCAGAGTGCTC	D3	39
PI-40	119	CAGCTATCATAC	120	CACTGCAGACGA	D4	40
PI-41	121	TATACTCTAGAT	122	TGCATAGAGCGC	D5	41
PI-42	123	TGTATGCTCGTC	124	TCGTGACAGATC	D6	42
PI-43	125	TAGTGATGCTCT	126	ACGAGCTGATAT	D7	43

PI-44	127	AGACTCTGAGTC	128	CTGATAGTATCA	D8	44
PI-45	129	CTCATAGACTAC	130	ATCGCGAGTGAC	D9	45
PI-46	131	TCGCTCACTACA	132	TGTCTCGACATC	D10	46
PI-47	133	ATAGAGTCTCAT	134	CGCATAGCGTAT	D11	47
PI-48	135	CGAGACACTCGC	136	TCGTAGTCTACA	D12	48
PI-49	137	CAGCATACTATC	138	TCGTGATACAGA	E1	49
PI-50	139	CAGCTATAGTCT	140	ATGCAGATATCT	E2	50
PI-51	141	TCTATCGATGCA	142	ACACGCAGATCG	E3	51
PI-52	143	CATGAGTATAGC	144	CTAGCTGACGTA	E4	52
PI-53	145	TAGCATATCGAG	146	TACACGTATGAG	E5	53
PI-54	147	ACGACTCGCTAC	148	TCATGACTAGTA	E6	54
PI-55	149	TAGCATACACGC	150	TGACGCGTATAC	E7	55
PI-56	151	CGTCATATGCAG	152	TATAGCGATGAC	E8	56
PI-57	153	TGCAGCGAGTAC	154	TCGACGCTAGCG	E9	57
PI-58	155	CGTGTCGACAGA	156	CAGTCGTGAGCA	E10	58
PI-59	157	ACTCGACGTGAG	158	ACGCGAGTGATA	E11	59
PI-60	159	ACTCGTCTGACG	160	TGCTATCACTGA	E12	60
PI-61	161	CATACTGTATCT	162	TACATAGATGTC	F1	61
PI-62	163	TCTACTCGTGAC	164	CACGTATAGTGA	F2	62
PI-63	165	CTGCACTAGACA	166	ACTCATATCGCA	F3	63
PI-64	167	ACACGAGCTCAT	168	CACTCATATCGA	F4	64
PI-65	169	TACAGATAGTCT	170	TCGTCTGTGATA	F5	65
PI-66	171	TACACTCGTGCT	172	TGACGCTCATCT	F6	66
PI-67	173	TACATGTGACGA	174	TCGTACATGCTC	F7	67
PI-68	175	TGTATGATCTCG	176	CACTGTGCTCAT	F8	68
PI-69	177	CAGTACACTCTA	178	ACTGCATGATCG	F9	69
PI-70	179	CATACTATCACG	180	TCGTGTCACTAC	F10	70

PI-71	181	CACTATACAGAT	182	CGACACGTACTA	F11	71
PI-72	183	ATATCGTAGCAT	184	TCGTGATCACTA	F12	72
PI-73	185	TAGTCTATACAT	186	AGACGCTGTCGA	G1	73
PI-74	187	TGTCACAGTGAC	188	TCATATGATCGA	G2	74
PI-75	189	ATCGACTATGCT	190	CGATCATATGAG	G3	75
PI-76	191	ATACTAGCATCA	192	TCATGCTGACGA	G4	76
PI-77	193	CACTGACGCTCA	194	CACTACATCGCT	G5	77
PI-78	195	TCGCTCATCTAT	196	TAGTACAGAGCT	G6	78
PI-79	197	TGTATCACGAGC	198	ATGATCGTATAC	G7	79
PI-80	199	TACTGCTATCTC	200	CGCTGCATAGCG	G8	80
PI-81	201	CGCGAGCTCGTC	202	ACTCGATGAGCT	G9	81
PI-82	203	TAGAGTCTGTAT	204	TGTCTATCACAT	G10	82
PI-83	205	TACTATCGCTCT	206	TATGTGACATAC	G11	83
PI-84	207	TAGATGACGCTC	208	TACTCGTAGCGC	G12	84
PI-85	209	TCGCGTGACATC	210	ATCTACTGACGT	H1	85
PI-86	211	ACACGCTCTACT	212	ACAGTAGCGCAC	H2	86
PI-87	213	TACATAGTCTCG	214	CTAGTATCATGA	H3	87
PI-88	215	TGAGTAGCACGC	216	TCGATCATGCAG	H4	88
PI-89	217	TAGATGCTATAC	218	TACATGCACTCA	H5	89
PI-90	219	ATCGATGTCACG	220	CAGCTCGACTAC	H6	90
PI-91	221	ATCATATGTAGC	222	CTCTACAGTCAC	H7	91
PI-92	223	TAGCATCGATAT	224	AGATAGCACATC	H8	92
PI-93	225	TGATCGACGCTC	226	CTAGATATCGTC	H9	93
PI-94	227	TGCAGCTCATAG	228	TACAGACTGCAC	H10	94
PI-95	229	CGACGTAGAGTC	230	CAGTAGCACTAC	H11	95

Tabela 7: Iniciadores de PCR para amplificação dos Éxons correspondentes de DRB1 e sem os iniciadores de indexação

SEQ NO:	ID	No. iniciador	Sequência de iniciador	Uso de iniciador	Comprimento dos produtos
231		D2-F1	CACGTTTCTTGGAGTACTCTA	Para amplificação de Éxon 2 de gene HLA-DRB1	300bp
232		D2-F2	GTTTCTTGTGGCAgCTTAAgTT		
233		D2-F3	CCTGTGGCAGGGTAAGTATA		
234		D2-F4	GTTTCTTGAAGCAGGATAAGT T		
235		D2-F5	GCACGTTTCTTGGAGGAGG		
236		D2-F6	TTTCCTGTGGCAGCCTAAGA		
237		D2-F7	GTTTCTTGGAGCAGGTAAAC		
238		D2-R	CCTCACCTCGCCGCTGCAC		

[0111] D2-F1, D2-F2, D2-F3, D2-F4, D2-F5, D2-F6, D2-F7 foram iniciadores dianteiros para a amplificação de Éxon 2 de HLA-DRB1, D2-R foi um iniciador reverso a amplificação de Éxon 2 de HLA-DRB1.

O procedimento PCR para HLA-A/B/DRB1 foi como seguido:

96°C 2min

95°C 30s → 60°C 30s → 72°C 20s (32 ciclos)

15°C ∞

[0112] Sistema de reação de PCR para HLA-A/B foi como seguido, onde todos os agentes foram comprados da Promega (Pequim) Bio-Tech Co.

Promega 5×buffer I (Mg ²⁺ plus)	5,0ul
dNTP Mistura (2,5mM/ µl cada)	2,0ul
Pl _{inf} -A/B-F _{2/3/4} (2pmol/ ul)	1,0ul
Pl _{nr} -A/B-R _{2/3/4} (2pmol/ ul)	1,0ul
Promega Taq (5U/ul)	0,2ul
DNA (cerca de 20 ng/ul)	5,0ul
ddH ₂ O	10,8ul

Total	25,0ul
-------	--------

[0113] O sistema de reação de PCR para HLA-DRB1 foi como segue:

Promega 5×buffer I (Mg ²⁺ plus)	5,0ul
dNTP Mistura (2,5mM/ µl cada)	2,0ul
PI _{nf} -D2-F1 (2pmol/ ul)	1,0ul
PI _{nf} -D2-F2 (2pmol/ ul)	1,0ul
PI _{nf} -D2-F3 (2pmol/ ul)	1,0ul
PI _{nf} -D2-F4 (2pmol/ ul)	1,0ul
PI _{nf} -D2-F5 (2pmol/ ul)	1,0ul
PI _{nf} -D2-F6 (2pmol/ ul)	1,0ul
PI _{nf} -D2-F7 (2pmol/ ul)	1,0ul
PI _{nr} -D2-R (2pmol/ ul)	1,0ul
Promega Taq (5U/ul)	0,2ul
DNA (cerca de 20 ng/ul)	5,0ul
ddH ₂ O	4,8ul
Total	25,0ul

[0114] Onde PI_{nf}-A/B/D2-F_{1/2/3/4/5/6/7} representa o F iniciador de HLA-A/B/DRB1 tendo uma sequência de iniciador de indexação dianteira No. n (Tabela 6) na 5' extremidade, PI_{nr}-A/B/D2-R_{2/3/4} representa o R iniciador de HLA-A/B/DRB1 tendo a sequência de iniciador de indexação reverso No. n na 5' extremidade (aqui n≤95), e o resto pode ser deduzido similarmente. Contudo, cada amostra corresponde a um conjunto específico de Iniciadores de PCR (PI_{nf}-A/B/D2-F_{1/2/3/4/5/6/7}, PI_{nr}-A/B/D2-R_{2/3/4}).

[0115] As reações de PCR foram realizadas em um aparelho PTC-200 PCR da Bio-Rad Co.. Depois da reação de PCR, 2ul de produtos de PCR foram submetidos a 1%de eletroforese de agarose gel. A figura 2 mostrada no resultado eletroforético dos produtos de PCR dos Éxons correspondentes de HLA-A/B/DRB1 da Amostra No.

1, e os marcadores para os pesos moleculares de DNA foram DL 2000 (Takara Co.). Há uma série de bandas únicas de um comprimento na média de 300 bp a 500 bp no eletroforograma, indicando sucesso na amplificação de PCR dos Éxons (A2, A3, A4, B2, B3, B4, DRB1-2) de HLA-A/B/DRB1 da Amostra No.1. Não há amplificação na faixa de controle negativo (N). Os resultados das outras amostras foram similares.

Exemplo 3

[0116] Associação e purificação de produtos de PCR

[0117] 20 µl dos produtos de PCR restantes foram tomadas de cada uma das 96 placas boas de HLA-P-A2 (exceto para o controle negativo), e foi misturada homogeneamente sob agitação em um tubo EP de 3 ml (designado como HLA-A2-Mix). A mesma operação foi aplicada às outras 6 placas boas 96, designadas como HLA-A3-Mix, HLA-A4-Mix, HLA-B2-Mix, HLA-B3-Mix, HLA-B4-Mix e HLA-D2-Mix. 200ul foi tomado de cada HLA-A2-Mix, HLA-A3-Mix, HLA-A4-Mix, HLA-B2-Mix, HLA-B3-Mix, HLA-B4-Mix e HLA-D2-Mix, e foi misturado em um tubo EP de 3 ml, designado como HLA-Mix. 500ul de mistura de DNA de HLA-Mix foi submetida a uma purificação de coluna com um kit de Purificação de DNA Qiagen (QIAGEN Co.) (Para as etapas de purificação específicas, por favor, refira-se as instruções do fabricante). Foi determinado pela Nanodrop 8000 (Thermo Fisher Scientific Co.) que os 200ul de DNA obtidos por purificação tem uma concentração de DNA HLA-Mix de 48 ng/ul.

Exemplo 4

[0118] Cisalhamento dos produtos de PCR, e a construção de bibliotecas de sequenciamento de PCR-livre Illumina GA

1. Cisalhamento de DNA

[0119] Uma quantidade total de 5 ug de DNA, tomado do HLA-Mix purificado, foi contido em um microtubo Covaris com uma fibra AFA e uma tampa de pressão e foi submetida ao cisalhamento no Cisalhador S2DNA Covaris (Covaris Co.). As condições de cisalhamento foram como seguidas:

Frequência de varredura

Ciclo de trabalho	10%
Intensidade	5
Ciclos/Explosão	200
Tempo (segundos)	300

2. Purificação depois do cisalhamento

[0120] Todos os produtos cisalhados de HLA-Mix foram recuperados e purificados por um Kit de purificação PCR QIaquick, e foram dissolvidos em 37,5 µl EB (QIAGEN Elution Buffer), respectivamente.

3. Reação de reparo terminal

[0121] O HLA-Mix purificado depois do cisalhamento foi submetido a uma reação de reparação de DNA, e o sistema de reação foi como seguido (todos os agentes foram comprados da Enzymatics Co.):

DNA	37,5µL
H ₂ O	37,5µL
10x Polinucleotídeo Kinase Buffer (B904)	10 µL
dNTP mistura (Conjunto de Solução (10mM cada))	4 µL
T4 DNA Polimerase	5µL
Klenow Fragmento	1µL
T4 Polinucleotídeo Kinase	5µL
Volume Total	100 µL

[0122] Condições da reação: incubação 20°C por 30 min em um Termomisturador (Termomisturador, Eppendorf Co.).

[0123] Os produtos da reação foram recuperados e purificados por um Conjunto de purificação de PCR QIaquick, e foram dissolvidos em 34 µl EB (QIAGEN Eluição de Buffer).

4. Adição de A na 3' extremidade

[0124] Foi adicionado na 3' extremidade do DNA recuperado na última etapa, e o sistema de reação foi como seguido (todos os agentes foram comprados da Enzymatics Co.):

DNA obtido na última etapa	32 µL
10x buffer azul	5 µL
dATP (1mM, GE Co.)	10 µL
Klenow (3'-5' exo-)	3 µL
Volume Total	50 µL

[0125] Condições da reação: incubação a 37°C por 30 min em um Termomisturador (Termomisturador, Eppendorf Co.).

[0126] Os produtos da reação foram recuperados e purificados por um Conjunto de purificação de PCR MiniElute (QIAGEN Co.), e foram dissolvidos em 13 µl EB (QIAGEN Eluição Buffer).

5. Ligação de adaptador de biblioteca de PCR livre Illumina GA

[0127] O termo “adaptador de biblioteca de PCR livre” refere-se a um segmento das bases designadas, cujo papel principal reside na fixação auxiliar de uma molécula de DNA no chip de sequenciamento e reside em proporcionar os locais de ligação para os sequenciamentos primários universais, onde o adaptador de biblioteca de PCR livre pode ser diretamente ligado a dois terminais dos fragmentos de DNA na biblioteca de sequenciamento; uma vez que nenhum PCR foi envolvido na introdução do adaptador de biblioteca, o adaptador de biblioteca foi chamado de adaptador de biblioteca de PCR livre.

[0128] Os produtos tendo A adicionado foram ligados aos adaptadores de biblioteca de PCR livre Illumina GA, e o sistema de reação foi como seguido (todos os agentes foram comprados da Illumina Co.):

DNA obtido na última etapa	11µL
2x buffer de ligação rápida	15µL

Adaptador oligo mix de PCR (30mM)	1µL
T4 DNA Ligase (Rápido, L603-HC-L)	3µL
Total volume	30 µL

[0129] Condições de reação: incubação a 20°C por 15 min em um Termomisturador (Termomisturador, Eppendorf Co.).

[0130] Os produtos de reação foram purificados por Ampure Beads (Beckman Coulter Genomics), e foram dissolvidos em 50ul de água desionizada, e a concentração de DNA determinada por um PCR quantitativo Fluorescente(QPCR) foi como segue:

	Resultado determinado por qPCR (nM)
HLA-Mix	78,90

6. Recuperação por gel fatiado

[0131] 30µL de HLA-Mix foi submetido a eletroforese 2% baixo ponto de derretimento de gel agarose. As condições eletroforéticas foram 100V, 100min. O marcador de DNA foi o marcador de DNA 50bp da NEB Co.. O gel contendo os fragmentos de DNA em média de 450 a 750bp foram fatiados (Figura 3). Os produtos no gel fatiado foram recuperados e purificados por um conjunto de Purificação de PCR QIAquick (QIAGEN Co.), o volume depois da purificação foi de 32ul, e a concentração de DNA medida pelo PCR quantitativo Fluorescente (QPCR) foi de 10,16nM.

Exemplo 5

[0132] Sequenciamento de Illumina GA

[0133] De acordo com os resultados de QPCR, 10pmol DNA foi tomado e submetido ao sequenciamento por pelo programa Illumina GA PE-100. Para o procedimento de operação específica, por favor, refira-se as instruções de operação de Illumina GA (Illumina GA IIx).

Exemplo 6

[0134] Análise dos resultados

[0135] Os resultados do sequenciamento de Illumina GA foram uma série de sequências de DNA, e procurando as sequências primárias dianteiras e reversas e as sequências primárias nos resultados de sequenciamento, o banco de dados compreendendo os resultados de sequenciamento dos produtos de PCR de vários Éxons de HLA-A/B/DRB1 para cada amostra correspondente ao iniciador de indexação respectivo foram construídas. Os resultados de sequenciamento de cada Éxon foram alinhados à sequência de referência (as sequências de referência foram da <http://www.ebi.ac.uk/imgt/hla/>) do Éxon correspondente por BWA (Burrows-Wheeler Aligner), e portanto, as sequências de consenso de cada base de dados foi construída, e as sequências de DNA no banco de dados foram selecionadas e corrigidas. As sequências de DNA corrigidas foram montadas nas sequências correspondentes de Éxons HLA-A/B/DRB1 com base na sequência de relacionamento de sobreposição e ligação (ligação de extremidade em pares). A sequência de DNA resultante foi alinhada com a sequência de banco de dados dos Éxons correspondentes de HLA-A/B/DRB1 em um banco de dados profissional IMGT HLA. Se o resultado do alinhamento de sequência mostra 100% de ligação, o genótipo HLA-A/B/DRB1 da amostra correspondente foi determinado. Por favor, refira-se a captura de tela do programa para construção da sequência de consenso de Éxon 2 de HLA-A local na Amostra No.1 como ilustrado na figura 4.

[0136] Para todas as 95 amostras, os resultados de tipagem obtidos pelos métodos acima foram completamente consistentes com os resultados de tipagem conhecidos, onde os resultados das Amostras No. 1-32 foram como seguido:

Amostra No.	Tipo HLA-A/B/DRB1 original conhecido					
1	A*02:03	A*11:01	B*38:02	B*48:01	DRB1*14:54	DRB1*15:01
2	A*01:01	A*30:01	B*08:01	B*13:02	DRB1*03:01	DRB1*07:01

3	A*01:01	A*02:01	B*15:11	B*47:01	DRB1*13:02	DRB1*15:01
4	A*24:08	A*26:01	B*40:01	B*51:01	DRB1*04:04	DRB1*09:01
5	A*01:01	A*24:02	B*54:01	B*55:02	DRB1*04:05	DRB1*09:01
6	A*01:01	A*03:02	B*15:11	B*37:01	DRB1*10:01	DRB1*14:54
7	A*11:01	A*30:01	B*13:02	B*15:18	DRB1*04:04	DRB1*07:01
8	A*01:01	A*02:01	B*35:03	B*81:01	DRB1*11:01	DRB1*15:01
9	A*02:06	A*31:01	B*27:07	B*40:02	DRB1*03:01	DRB1*13:02
10	A*01:01	A*66:01	B*37:01	B*49:01	DRB1*10:01	DRB1*13:02
11	A*01:01	A*03:01	B*35:01	B*52:01	DRB1*01:01	DRB1*15:02
12	A*11:01	A*11:01	B*15:01	B*15:05	DRB1*04:06	DRB1*15:01
13	A*01:01	A*11:02	B*07:02	B*15:02	DRB1*09:01	DRB1*15:01
14	A*01:01	A*02:01	B*52:01	B*67:01	DRB1*15:02	DRB1*16:02
15	A*01:01	A*02:05	B*15:17	B*50:01	DRB1*07:01	DRB1*15:01
16	A*01:01	A*11:01	B*37:01	B*40:02	DRB1*10:01	DRB1*12:02
17	A*24:07	A*32:01	B*35:05	B*40:01	DRB1*03:01	DRB1*04:05
18	A*11:01	A*24:02	B*13:01	B*35:01	DRB1*16:02	DRB1*16:02
19	A*11:01	A*11:01	B*40:02	B*55:12	DRB1*04:05	DRB1*15:01
20	A*02:11	A*24:02	B*40:01	B*40:06	DRB1*11:01	DRB1*15:01
21	A*01:01	A*02:06	B*51:01	B*57:01	DRB1*07:01	DRB1*12:01
22	A*01:01	A*29:01	B*07:05	B*15:01	DRB1*04:05	DRB1*07:01
23	A*01:01	A*02:07	B*37:01	B*46:01	DRB1*04:03	DRB1*10:01
24	A*24:85	A*30:01	B*13:02	B*55:02	DRB1*07:01	DRB1*15:01
25	A*11:01	A*31:01	B*07:06	B*51:01	DRB1*12:02	DRB1*14:05
26	A*01:01	A*11:01	B*46:01	B*57:01	DRB1*07:01	DRB1*08:03
27	A*01:01	A*02:01	B*15:18	B*37:01	DRB1*04:01	DRB1*15:01
28	A*01:01	A*24:02	B*37:01	B*46:01	DRB1*09:01	DRB1*10:01
29	A*26:01	A*66:01	B*40:40	B*41:02	DRB1*12:01	DRB1*15:01

30	A*02:01	A*29:02	B*13:02	B*45:01	DRB1*03:01	DRB1*12:02
31	A*01:01	A*11:03	B*15:01	B*57:01	DRB1*07:01	DRB1*15:01
32	A*11:01	A*26:01	B*35:03	B*38:01	DRB1*11:03	DRB1*14:04

Amostra No.	O tipo de HLA-A/B/DRB1 determinado					
1	A*02:03	A*11:01	B*38:02	B*48:01	DRB1*14:54	DRB1*15:01
2	A*01:01	A*30:01	B*08:01	B*13:02	DRB1*03:01	DRB1*07:01
3	A*01:01	A*02:01	B*15:11	B*47:01	DRB1*13:02	DRB1*15:01
4	A*24:08	A*26:01	B*40:01	B*51:01	DRB1*04:04	DRB1*09:01
5	A*01:01	A*24:02	B*54:01	B*55:02	DRB1*04:05	DRB1*09:01
6	A*01:01	A*03:02	B*15:11	B*37:01	DRB1*10:01	DRB1*14:54
7	A*11:01	A*30:01	B*13:02	B*15:18	DRB1*04:04	DRB1*07:01
8	A*01:01	A*02:01	B*35:03	B*81:01	DRB1*11:01	DRB1*15:01
9	A*02:06	A*31:01	B*27:07	B*40:02	DRB1*03:01	DRB1*13:02
10	A*01:01	A*66:01	B*37:01	B*49:01	DRB1*10:01	DRB1*13:02
11	A*01:01	A*03:01	B*35:01	B*52:01	DRB1*01:01	DRB1*15:02
12	A*11:01	A*11:01	B*15:01	B*15:05	DRB1*04:06	DRB1*15:01
13	A*01:01	A*11:02	B*07:02	B*15:02	DRB1*09:01	DRB1*15:01
14	A*01:01	A*02:01	B*52:01	B*67:01	DRB1*15:02	DRB1*16:02
15	A*01:01	A*02:05	B*15:17	B*50:01	DRB1*07:01	DRB1*15:01
16	A*01:01	A*11:01	B*37:01	B*40:02	DRB1*10:01	DRB1*12:02
17	A*24:07	A*32:01	B*35:05	B*40:01	DRB1*03:01	DRB1*04:05
18	A*11:01	A*24:02	B*13:01	B*35:01	DRB1*16:02	DRB1*16:02
19	A*11:01	A*11:01	B*40:02	B*55:12	DRB1*04:05	DRB1*15:01
20	A*02:11	A*24:02	B*40:01	B*40:06	DRB1*11:01	DRB1*15:01
21	A*01:01	A*02:06	B*51:01	B*57:01	DRB1*07:01	DRB1*12:01

22	A*01:01	A*29:01	B*07:05	B*15:01	DRB1*04:05	DRB1*07:01
23	A*01:01	A*02:07	B*37:01	B*46:01	DRB1*04:03	DRB1*10:01
24	A*24:85	A*30:01	B*13:02	B*55:02	DRB1*07:01	DRB1*15:01
25	A*11:01	A*31:01	B*07:06	B*51:01	DRB1*12:02	DRB1*14:05
26	A*01:01	A*11:01	B*46:01	B*57:01	DRB1*07:01	DRB1*08:03
27	A*01:01	A*02:01	B*15:18	B*37:01	DRB1*04:01	DRB1*15:01
28	A*01:01	A*24:02	B*37:01	B*46:01	DRB1*09:01	DRB1*10:01
29	A*26:01	A*66:01	B*40:40	B*41:02	DRB1*12:01	DRB1*15:01
30	A*02:01	A*29:02	B*13:02	B*45:01	DRB1*03:01	DRB1*12:02
31	A*01:01	A*11:03	B*15:01	B*57:01	DRB1*07:01	DRB1*15:01
32	A*11:01	A*26:01	B*35:03	B*38:01	DRB1*11:03	DRB1*14:04

Nota: entre os tipos de HLA-DRB1, o DRB1*1201 não exclui a possibilidade de DRB1*1206/1210/1217, e o DRB1*1454 não exclui a possibilidade de DRB1*1401, porque tais alelos foram completamente identificados na sequência de Éxon 2 de HLA-DRB1.

Exemplo 7

[0137] A genotipagem HLA-A,B e DRB1 usando a técnica de sequenciamento de segunda geração (Illumina GA)

[0138] Extração da Amostra

[0139] Os DNAs foram extraídos de 950 amostras de sangue com resultados de tipagem de HLA-SBT conhecidos (Programa de Doador de Medula Óssea da China citado aqui como (CMDP)) usando um Instrumento de Extração Automático KingFisher (US Thermo Co.). O método foi como descrito no Exemplo 1.

[0140] Amplificação de PCR

[0141] Os 950 DNAs obtidos da etapa de extração de amostra foram designados como No. 1-950, e foram divididos em 10 grupos (95 DNAs para cada), que foram designados como HLA-1, HLA-2, HLA-3, HLA-4, HLA-5, HLA-6, HLA-7, HLA-8, HLA-

9, HLA-10. Para cada grupo de amostras, 95 amostras de DNA foram amplificadas por 95 conjuntos de Iniciadores de PCR (Tabela 1) carregando iniciadores de indexação bidirecionais (Tabela 6) para amplificação de Éxons 2, 3, 4 de HLA-A/B e Iniciadores de PCR (Tabela 7) carregando os iniciadores de indexação bidirecionais (Tabela 6) para a amplificação de Éxon 2 de HLA-DRB1. A reação de PCR aconteceu em 96-placas boas, usando 70 placas no total, designadas como HLA-X-P-A2, HLA-X-P-A3, HLA-X-P-A4, HLA-X-P-B2, HLA-X-P-B3, HLA-X-P-B4 e HLA-X-P-DRB1-2 ("X" representa a informação do número de grupo 1/2/3/4/5/6/7/8/9/10, "A2/3/4", "B2/3/4", "DRB1-2" representa os locais de amplificação), onde um controle negativo sem a adição de qualquer modelo foi configurada em cada placa, e os iniciadores usados para o controle negativo foram iniciadores legendados por PI-1 (Tabela 6). Durante a experimentação, a informação de cada amostra no número de grupo e nos iniciadores de indexação foi gravada. O método foi descrito no Exemplo 2.

[0142] Associação e purificação dos produtos de PCR

[0143] Para as amostras do Grupo X ("X" é 1/2/3/4/5/6/7/8/9/10), 20 µl do resto de produtos PCR foram tomados de cada amostra boa de 96 placas boas HLA-X-P-A2 (exceto para o controle negativo), e foi misturado homogeneamente sob agitação em um tubo EP de 3 ml (designado como HLA-X-A2-Mix). A mesma operação foi aplicada às outras 96 placas boas das amostras do Grupo X, designadas como HLA-X-A3-Mix, HLA-X-A4-Mix, HLA-X-B2-Mix, HLA-X-B3-Mix, HLA-X-B4-Mix e HLA-X-D2-Mix. 200ul foi tomado de cada HLA-X-A2-Mix, HLA-X-A3-Mix, HLA-X-A4-Mix, HLA-X-B2-Mix, HLA-X-B3-Mix, HLA-X-B4-Mix e HLA-X-D2-Mix, e foi misturado em um tubo EP de 3ml, designado como HLA-X-Mix. 500ul de mistura de DNA de HLA-X-Mix foi submetido a uma purificação de coluna com um conjunto de Purificação de DNA Qia-gen DNA (QIAGEN Co.) (Para as etapas de purificação específicas, por favor, refira-se as instruções do fabricante) para obter 200ul de DNA, e sua concentração de DNA foi determinada por Nanodrop 8000 (Thermo Fisher Scientific Co.). A mesma

operação também foi aplicada aos outros grupos. As concentrações de DNA determinadas finalmente foram como seguido.

	HLA- -1- Mix	HLA- -2- Mix	HLA- -3- Mix	HLA- -4- Mix	HLA- -5- Mix	HLA- -6- Mix	HLA- -7- Mix	HLA- 8- Mix	HLA- 9- Mix	HLA- 10- Mix
concentração (ng/ul)	48,0	52,1	49,3	50,2	47,6	48,5	49,1	48,6	51,3	50,8

[0144] O método foi descrito no Exemplo 3.

[0145] A construção das Bibliotecas de Sequenciamento de Illumina GA foi realizada pelo método do Exemplo 4. Os relacionamentos correspondentes entre os grupos de amostra e os adaptadores de biblioteca foram como seguido.

Grupo de Amostra No.	HLA-1	HLA-2	HLA-3	HLA-4	HLA-5	HLA-6	HLA-7	HLA-8	HLA-9	HLA-10
Adaptador de bibliotecar No.	1	2	3	4	5	6	7	8	9	10

[0146] Os produtos de reação foram purificados por Ampure Beads (Beckman Coulter Genomics), e foram dissolvidos em 50ul de água desionizada, e as concentrações molares de DNA determinadas por um PCR de quantitativo Fluorescente (QPCR) foram como segue:

	HLA-1-Mix	HLA-2-	HLA-3-Mix	HLA-4-	HLA-5-	HLA-6-	HLA-7-	HLA-8-	HLA-9-	HLA-10-Mix
--	-----------	--------	-----------	--------	--------	--------	--------	--------	--------	------------

		Mix		Mix	Mix	Mix	Mix	Mix	Mix	
Conc.	78,9	72,1	79,3	80,2	77,6	78,5	89,1	78,6	81,3	80,82
(nM)	0	3	3	1	8	0	2	0	2	

[0147] Recuperação por gel fatiado

[0148] HLA-1-Mix, HLA-2-Mix, HLA-3-Mix, HLA-4-Mix, HLA-5-Mix, HLA-6-Mix, HLA-7-Mix, HLA-8-Mix, HLA-9-Mix e HLA-10-Mix foram misturados em um mol igual (a concentração final foi 72,13 nM/ul), designada como HLA-Mix-10. 30µL HLA-Mix-10 foi submetida a uma eletroforese 2% baixo ponto de derretimento agarose gel. As condições eletroforéticas foram de 100V, 100min. o marcador de DNA foi o 50bp DNA da NEB Co.. O gel contendo os fragmentos de DNA na média de 450 a 750bp foi fatiado. Os produtos no gel fatiado foram recuperados e purificados por um conjunto de Purificação de PCR QIAquick (QIAGEN Co.), o volume depois da purificação foi de 32ul, e a concentração de DNA medida pelo PCR quantitativo Fluorescente (QPCR) foi de 9,96nM.

[0149] O sequenciamento e a análise dos resultados foram realizados como descrito nos Exemplos 5 e 6. Para todas as 950 amostras, os resultados da tipagem obtidos pelo método acima foram completamente consistentes com os resultados de tipagem conhecidos.

Exemplo 8

[0150] Genotipagem de HLA-C usando a técnica de sequenciamento de segunda geração (Illumina GA)

1. Extração de amostra de DNA

As etapas foram descritas no Exemplo 1.

2. Amplificação de PCR

[0151] As etapas foram descritas no Exemplo 2, exceto que os Iniciadores de PCR usados foram Iniciadores de PCR para Éxons 2, 3 e 4 de HLA-C, como mostrado na Tabela 3.

[0152] 95 conjuntos de iniciadores de indexação de PCR foram usados para amplificar 95 amostras de DNA, respectivamente, onde cada conjunto de iniciador de indexação consistia de iniciador de PCR para a amplificação de Éxons 2, 3, 4 de HLA-C (Tabela 3) e um par de iniciadores de indexação bidirecionais (como descrito abaixo), cada Iniciador de PCR dianteiro tem um iniciador de indexação dianteiro de um par de iniciadores de indexação ligados na 5' extremidade, e um Iniciador de PCR reverso tem o iniciador de indexação reverso de um par de iniciadores de indexação reversos ligados na 5' extremidade. Durante a síntese dos iniciadores, os iniciadores de indexação foram adicionados diretamente a 5' extremidade dos Iniciadores de PCR.

[0153] Os 95 DNAs obtidos da etapa de extração de amostra foram designados como No. 1-95. A reação PCR aconteceu em 96-placas boas, 3 placas no total, designadas como HLA-P-C2, HLA-P-C3, HLA-P-A4 (C2/3/4 representam os locais de amplificação), onde um controle negativo sem adicionar qualquer modelo foi configurada em cada placa, e os iniciadores usados no controle negativo foram os mesmos do iniciador PI-96. Durante a experimentação, a quantidade de informação da amostra correspondendo a cada par dos iniciadores de indexação foi recuperada.

[0154] Os iniciadores de indexação usados foram os iniciadores de indexação PI-1 a PI-95 como listado na Tabela 6, e os seguintes controles negativos de iniciador de indexação PI-96 (Tabela 8)

Tabela 8: Informação relevante dos iniciadores de indexação usados para o controle negativo

PI-96	SEQ ID NO:239	CACTGTATAGCT	SEQ ID NO:240	CGACTAGTACTA	H12	Controle Negativo
-------	---------------	--------------	---------------	--------------	-----	-------------------

[0155] Os DNAs, extraídos usando o Instrumento de Extração Automático KingFisher na etapa 1, foram usados como modelo, e a amplificação de PCR foi realizada em tubos únicos usando os iniciadores para Éxons de HLA-C, onde os

iniciadores tem índices na 5' extremidade. O procedimento de PCR foi como segue:

C2: 96°C 2 min

95°C 30s → 62°C 30s → 72°C 20s (35 ciclos)

15°C ∞

C3: 96°C 2 min

95°C 30s → 56°C 30s → 72°C 20s (35 ciclos)

15°C ∞

C4: 96°C 2 min

95°C 30s → 60°C 30s → 72°C 20s (35 ciclos)

15°C ∞

[0156] Sistema de reação PCR de HLA-C foi como segue:

Promega 5×buffer I (Mg ²⁺ plus)	5,0 µL
dNTP mistura (2,5 mM cada)	2,0 µL
PI _{nr} -C-F _{2/3/4} (50 ng/ul)	1,5 µL
PI _{nf} -C-R _{2/3/4} (50 ng/ul)	1,5 µL
Promega Taq (5 U/ul)	0,2 µL
DNA (cerca de 20 ng/ul)	2,0 µL
ddH ₂ O	12,8 µL
Total	25,0 µL

[0157] Onde, PI_{nf}-C-F_{2/3/4} representa o F iniciador de HLA-C tendo uma sequência primária dianteira No. n (Tabela 2) na 5' extremidade, PI_{nr}-C-R_{2/3/4} representa o R iniciador de HLA-C tendo uma sequência de iniciador dianteira No. n na 5' extremidade (aqui n≤96), e o resto pode ser deduzido similarmente. Além disso, cada amostra corresponde a um conjunto de Iniciadores de PCR.

[0158] A reação de PCR foi realizada no aparelho PTC-200 PCR da Bio-Rad Co.. Depois da reação de PCR, 2ul dos produtos PCR foram submetidos a uma eletroforese de 1,5% agarose gel. A figura 6 mostrou o resultado eletroforético dos

produtos PCR dos Éxons correspondentes de HLA-C das primeiras 20 amostras, e o marcador de molécula de DNA foi DL 2000 (Takara Co.). Havia uma série de faixas únicas de um comprimento na média de 400 bp a 500 bp no eletroforograma, indicando sucesso na amplificação de PCR de Éxons (C2, C3, C4) de HLA-C das amostras. Os resultados das outras amostras foram similares.

[0159] Associação e purificação de produtos PCR

[0160] 20 µl do resto dos produtos de PCR foram tomados de cada boa das 96-placas boas HLA-P-C2 (exceto para o controle negativo), e foi associada homogeneamente sob agitação em um tubo EP de 3 ml (designado como HLA-C2-Mix). A mesma operação foi aplicada às outras 2 das 96 placas boas, designadas como HLA-C3-Mix e HLA-C4-Mix. 200ul foi tomado de cada HLA-C2-Mix, HLA-C3-Mix e HLA-C4-Mix, e foi misturado em um tubo EP de 1,5 ml, designado como HLA-Mix. 500ul de mistura de DNA de HLA-Mix foi submetido a uma purificação de coluna com um conjunto de Purificação de DNA Qiagen (QIAGEN Co.) (Para as etapas de purificação específicas, por favor, refira-se as instruções do fabricante). Foi determinado por Nanodrop 8000 (Thermo Fisher Scientific Co.) que os 200ul de DNA obtido por purificação tem uma concentração de DNA HLA-Mix de 50ng/ul.

4. Construção de bibliotecas de sequenciamento de PCR livre Illumina GA

4.1 Cisalhamento de produtos de PCR

[0161] Uma quantidade total de 5 ug de DNA, tomado do HLA-Mix purificado, foi contido em um microtubo Covaris com uma fibra AFA e uma tampa de pressão e foi submetida ao cisalhamento S2 Covaris (Covaris Co.). As condições de cisalhamento foram como segue:

Frequência de varredura

Ciclo de Trabalho	10%
Intensidade	3
Ciclos/Explosão	200

Tempo(s)	180
----------	-----

4.2 Purificação depois do cisalhamento

[0162] Todos os produtos cisalhados de HLA-Mix foram recuperados e purificados por um Kit de purificação PCR QIaquick, e foram dissolvidos em 37,5 µl EB (QIAGEN Elution Buffer), respectivamente.

4.3 Reação de reparo terminal

[0163] Os produtos purificados foram submetidos a uma reação de reparação de DNA, e o sistema de reação foi como segue (todos os agentes foram comprados da Enzymatics Co.):

Produtos purificados na última etapa	37,5 µL
10x Polinucleotídeo Kinase Buffer (B904)	5 µL
dNTP mistura (Configuração de Solução (10mM cada))	2 µL
T4 DNA Polimerase	2,5 µL
Klenow Fragmento	0,5 µL
T4 Polinucleotídeo Kinase	2,5 µL
Volume total	50µL

[0164] Condições da reação: incubação 20°C por 30 min em um Termomisturador (Termomisturador, Eppendorf Co.).

[0165] Os produtos da reação foram recuperados e purificados por um Conjunto de purificação de PCR QIAquick, e foram dissolvidos em 32 µl EB (QIAGEN Eluição de Buffer).

4.4 Adição de A na 3' extremidade

[0166] Foi adicionado na 3' extremidade do DNA recuperado na última etapa, e o sistema de reação foi como segue (todos os agentes foram comprados da

Enzymatics Co.):

DNA obtido na última etapa	32 µL
10x buffer azul	5 µL
dATP (1mM, GE Co.)	10 µL
Klenow (3'-5' exo-)	3 µL
Volume Total	50 µL

[0167] Condições da reação: incubação a 37°C por 30 min em um Termomisturador (Termomisturador, Eppendorf Co.).

[0168] Os produtos da reação foram recuperados e purificados por um Conjunto de purificação de PCR MiniElute (QIAGEN Co.), e foram dissolvidos em 38 µl EB (QIAGEN Eluição Buffer).

4.5 Ligação de adaptador de biblioteca de PCR livre Illumina GA

[0169] Os produtos tendo A adicionado foram ligados aos adaptadores de biblioteca de PCR livre Illumina GA, e o sistema de reação foi como seguido (todos os agentes foram comprados da Illumina Co.):

DNA obtido na última etapa	38 µL
10x buffer de ligação	5 µL
Adaptador de PCR-livre oligo mix (30mM)	2 µL
T4 DNA Ligase (Rápido, L603-HC-L)	5 µL
Volume total	50 µL

[0170] Condições de reação: incubação a 16°C durante a noite em um Termomisturador (Termomisturador, Eppendorf Co.).

[0171] Os produtos de reação foram purificados por Ampure Beads (Beckman Coulter Genomics), e foram dissolvidos em 50ul de água desionizada, e a concentração de DNA determinada por um PCR quantitativo Fluorescente(QPCR) foi como segue:

	resultado determinado por qPCR (nM)
HLA-Mix	122,71

4.6 Recuperação por gel fatiado

[0172] 30µL de HLA-Mix foi submetido a eletroforese 2% baixo ponto de derretimento de gel agarose. As condições eletroforéticas foram 100V, 100min. O marcador de DNA foi o marcador de DNA Ladder 50bp da NEB Co.. O gel contendo os fragmentos de DNA em média de 400 a 750bp foram fatiados (Figura 7). Os produtos no gel fatiado foram recuperados e purificados por um conjunto de Purificação de PCR QIAquick (QIAGEN Co.), o volume depois da purificação foi de 32ul, e a concentração de DNA medida pelo PCR quantitativo Fluorescente (QPCR) foi de 17,16nM.

5. Sequenciamento de Illumina GA

[0173] De acordo com os resultados de QPCR, 10pmol DNA foi tomado e submetido ao sequenciamento por pelo programa Illumina GA PE-100. Para o procedimento de operação específica, por favor, refira-se as instruções de operação de Illumina GA (Illumina GA IIX).

6. Análise dos resultados

[0174] Os resultados do sequenciamento de Illumina GA foram uma série de sequências de DNA, e procurando as sequências primárias dianteiras e reversas e as sequências primárias nos resultados de sequenciamento, o banco de dados compreendendo os resultados de sequenciamento dos produtos de PCR de vários Éxons de HLA-C para cada amostra correspondente ao iniciador de indexação respectivo foram construídas. Os resultados de sequenciamento de cada Éxon foram alinhados à sequência de referência (as sequências de referência foram da <http://www.ebi.ac.uk/imgt/hla/>) do Éxon correspondente por BWA (Burrows-Wheeler Aligner), e as sequências de consenso de cada base de dados foi construída, e as

sequências de DNA lidas foram selecionadas e corrigidas com base no valor qualidade da base de sequenciamento e diferença entre a sequência lida e as sequências de consenso. As sequências de DNA corrigidas foram montadas nas sequências correspondentes de Éxons HLA-C com base na sequência de sobreposição e relacionamento de ligação (ligação de extremidade em pares). A captura de tela da figura 8 ilustra o procedimento para a construção da sequência de consenso de Éxon 2 de HLA-C no local na amostra No.2.

[0175] A sequência de DNA resultante foi alinhada com a sequência do banco de dados profissional do Éxon correspondente HLA-C in IMGT HLA. Se o resultado da sequência de alinhamento mostrou 100% de combinação, o genótipo HLA-C da amostra correspondente foi determinado. Para todas as 95 amostras, os resultados de tipagem obtidos pelos métodos acima foram completamente consistentes com os resultados de tipagem conhecidos, onde os resultados das Amostras No. 1-32 foram como segue: (como mostrado na Tabela 9, todos os resultados obtidos foram idênticos aos resultados originais).

Tabela 9: Comparação dos resultados de tipagem obtidos pelo método acima com os resultados de tipagem originais conhecidos das amostras

Amostr a No.	Genótipo Original conhecido HLA-C		Resultados para HLA-C obtidos neste momento		Idêntico ou não
1	C*08:01	C*15:05	C*08:01	C*15:05	sim
2	C*01:02	C*07:02	C*01:02	C*07:02	sim
3	C*08:01	C*16:02	C*08:01	C*16:02	sim
4	C*01:02	C*03:02	C*01:02	C*03:02	sim
5	C*01:02	C*02:02	C*01:02	C*02:02	sim
6	C*01:02	C*15:02	C*01:02	C*15:02	sim
7	C*01:02	C*03:04	C*01:02	C*03:04	sim

8	C*03:02	C*07:02	C*03:02	C*07:02	sim
9	C*06:02	C*16:02	C*06:02	C*16:02	sim
10	C*01:02	C*03:04	C*01:02	C*03:04	sim
11	C*03:04	C*07:02	C*03:04	C*07:02	sim
12	C*07:02	C*08:01	C*07:02	C*08:01	sim
13	C*01:02	C*15:02	C*01:02	C*15:02	sim
14	C*01:02	C*03:04	C*01:02	C*03:04	sim
15	C*01:02	C*03:04	C*01:02	C*03:04	sim
16	C*07:02	C*12:02	C*07:02	C*12:02	sim
17	C*04:01	C*08:01	C*04:01	C*08:01	sim
18	C*08:01	C*16:02	C*08:01	C*16:02	sim
19	C*14:02	C*15:02	C*14:02	C*15:02	sim
20	C*01:02	C*03:03	C*01:02	C*03:03	sim
21	C*03:03	C*08:01	C*03:03	C*08:01	sim
22	C*03:04	C*07:02	C*03:04	C*07:02	sim
23	C*07:02	C*08:01	C*07:02	C*08:01	sim
24	C*07:02	C*12:02	C*07:02	C*12:02	sim
25	C*07:02	C*12:03	C*07:02	C*12:03	sim
26	C*03:04	C*08:01	C*03:04	C*08:01	sim
27	C*01:02	C*03:04	C*01:02	C*03:04	sim
28	C*07:02	C*12:02	C*07:02	C*12:02	sim
29	C*03:02	C*07:02	C*03:02	C*07:02	sim
30	C*01:02	C*03:03	C*01:02	C*03:03	sim
31	C*01:02	C*07:02	C*01:02	C*07:02	sim
32	C*01:02	C*07:02	C*01:02	C*07:02	sim

Nota: entre os tipos de HLA-C, o C*0303 não exclui a possibilidade de C*0320N, o C*0401 não exclui a possibilidade de C*0409N/0430, o C*0702 não exclui

a possibilidade de C*0750, o C*0801 não exclui a possibilidade de *0822, o C*1505 não exclui a possibilidade de C*1529, porque tais alelos foram completamente identificados na sequência de Éxon 2 de HLA-C.

Exemplo 9: genotipagem HLA-C usando um método de sequenciamento Sanger

1. Extração da amostra de DNA

[0176] Como descrito no Exemplo 1, os DNAs foram extraídos usando um Instrumento de Extração Automática KingFisher de 26 a 95 amostras com os genótipos HLA conhecidos.

2. Amplificação de PCR

[0177] Os DNAs acima, extraídos usando o Instrumento de Extração Automática KingFisher, foram usados como modelo, e a amplificação de PCR foi realizada em tubos únicos usando três pares de Iniciadores de PCR C-F2/C-R2, C-F3/C-R3, C-F4/C-R4 (Tabela 3), respectivamente. O procedimento de PCR para cada par de iniciadores foi como segue :

C2: 96°C 2 min

95°C 30s → 62°C 30s → 72°C 20s (35 ciclos)

15°C ∞

C3: 96°C 2min

95°C 30s → 56°C 30s → 72°C 20s (35 ciclos)

15°C ∞

C4: 96°C 2min

95°C 30s → 60°C 30s → 72°C 20s (35 ciclos)

15°C ∞

[0178] O sistema de reação de PCR de HLA-C foi como segue:

Promega 5×buffer I (Mg ²⁺ plus)	5,0 µL
dNTP Mistura (2,5 mM cada)	2,0 µL

Primer mix (50 ng/μL)	3,0 μL
Promega Taq (5 U/μL)	0,2 μL
DNA (cerca de 20 ng/μL)	2,0 μL
ddH ₂ O	12,8 μL
total	25,0 μL

[0179] Os produtos de PCR foram submetidos a eletroforese agarose gel (Figura 9) antes da purificação.

3. Purificação de produtos PCR

[0180] Os produtos de PCR foram purificados usando placas de purificação Millipore. As etapas principais foram como segue. As fontes a serem usadas foram marcadas com uma caneta marcadora nas 96 fontes de placa purificadora para os produtos de PCR, e 50 μl de água ultrapura foi adicionada a cada uma das fontes a serem usadas. O resto das fontes foram seladas por um filme selador. A placa foi mantida por 15 min ou foi conectada a um sistema de filtração e extração (-10 pa) por 5 min. Quando a placa de purificação foi tomada do sistema de filtração e extração, o líquido na porta de descarga no fundo da placa de purificação foi sugada por um papel absorvente.

[0181] Os produtos de PCR a serem purificados foram centrifugados a 4000 rpm por 1 min; a cobertura ou almofada de sílica gel para os produtos PCR a serem purificados foi removido, e 100 μl de água ultrapura foi adicionada a cada sistema de reação PCR. Então, a placa de purificação, a qual os produtos PCR a serem purificados foram adicionados, foi conectada ao sistema de filtração e extração, e o grau de vácuo foi ajustado para -10 pa como mostra no barômetro. A filtração e a extração foram continuadas até que o líquido foi deixado no filme de celulose regenerável de microporos no fundo da placa de purificação, e nenhuma reflexão de brilho da superfície líquida intacta foi encontrada quando observando sob a luz.

[0182] Nas fontes contendo produtos PCR a serem purificados, 50 μl de água

ultrapure ou TE foram adicionadas ao filme de celulose regenerável de microporo; a placa de purificação foi vibrada na temperatura ambiente em um vibrador de traço por 5 min; e todos os líquidos contidos nas fontes correspondentes foram transferidos para fontes correspondentes de novas 96 fontes de placa PCR.

4. Realização da reação de sequenciamento e purificação de produtos da reação de sequenciamento

[0183] Os produtos de PCR purificados foram usados como modelo para a reação de sequenciamento.

Condições para a reação de sequenciamento

96°C 2 min

96°C 10s → 55°C 5s → 60°C 2min (25 ciclos)

15°C ∞

[0184] O sistema para a reação de sequenciamento foi

Produtos de PCR purificados	1 µL
iniciadores (3,2 pmol/l)	1 µL
2,5 *Grande corante	0,3 µL
5*Buffer	0,85 µL
água	1,85 µL
Volume Total	5 µL

[0185] Os produtos da reação de sequenciamento foram purificados pelas seguintes etapas: a placa de reação de sequenciamento foram balanceadas, e centrifugadas 3000 g por 1 min. Na 96-placa fonte, para cada 5 µL de sistema de reação, 2µL 0,125 mol/L EDTA-Na₂ de solução, 33 µL 85% de etanol foi adicionado, e a placa foi coberta com uma almofada de sílica gel e foi suficientemente vibrada por 3 min. A placa foi então centrifugada a 4°C, 3000 g por 30 min. A placa de sequenciamento foi tomada depois da centrifugação, a almofada de sílica gel foi

removida, e a placa de sequenciamento foi absorvida em papel absorvente para baixo, e foi então submetida a uma centrifugação invertida até que a força de centrifugação alcançasse 185g. Para cada placa fonte 96, 50 µl 70% de etanol foi adicionado. A placa foi coberta com uma almofada de sílica gel, e vibrada por 1,5 min, e centrifugada a 4°C, 3000 g por 15 min. A placa de reação de sequenciamento foi então colocada em um local escuro e ventilado por 30 min de forma a ser secado pelo ar até que o odor de etanol não seja mais sentido. Para cada fonte da placa fonte 96, 10 µL HI-DI formamida foi adicionada (alternativamente, para cada fonte de placa fonte 384, 8 µL foi adicionado), e então a placa foi coberta com um filme selante, e foi centrifugado a 1000 rpm depois de vibrado por 5s.

5. Análise de sequenciamento e resultado

[0186] Os produtos purificados da reação de sequenciamento foram submetidos a um sequenciamento de eletroforese capilar em ABI 3730XL. Os picos do sequenciamento foram analisados por um software Tipo-u (Invitrogen) para obter resultados de tipagem HLA (Figura 10). Todos os resultados obtidos pelo método acima foram idênticos aos resultados originais conhecidos, como mostrado na Tabela 10.

Tabela 10: Comparação dos resultados de tipagem obtidos pelo método acima com os resultados de tipagem original conhecidos

Amostra	Genotipagem Original conhecida HLA-C		Resultados para HLA-C obtidos neste momento		Idêntico ou não
No.					
1	C*08:01	C*15:05	C*08:01	C*15:05	sim
2	C*01:02	C*07:02	C*01:02	C*07:02	sim
3	C*08:01	C*16:02	C*08:01	C*16:02	sim
4	C*01:02	C*03:02	C*01:02	C*03:02	sim
5	C*01:02	C*02:02	C*01:02	C*02:02	sim
6	C*01:02	C*15:02	C*01:02	C*15:02	sim
7	C*01:02	C*03:04	C*01:02	C*03:04	sim

8	C*03:02	C*07:02	C*03:02	C*07:02	sim
9	C*06:02	C*16:02	C*06:02	C*16:02	sim
10	C*01:02	C*03:04	C*01:02	C*03:04	sim
11	C*03:04	C*07:02	C*03:04	C*07:02	sim
12	C*07:02	C*08:01	C*07:02	C*08:01	sim
13	C*01:02	C*15:02	C*01:02	C*15:02	sim
14	C*01:02	C*03:04	C*01:02	C*03:04	sim
15	C*01:02	C*03:04	C*01:02	C*03:04	sim
16	C*07:02	C*12:02	C*07:02	C*12:02	sim
17	C*04:01	C*08:01	C*04:01	C*08:01	sim
18	C*08:01	C*16:02	C*08:01	C*16:02	sim
19	C*14:02	C*15:02	C*14:02	C*15:02	sim
20	C*01:02	C*03:03	C*01:02	C*03:03	sim
21	C*03:03	C*08:01	C*03:03	C*08:01	sim
22	C*03:04	C*07:02	C*03:04	C*07:02	sim
23	C*07:02	C*08:01	C*07:02	C*08:01	sim
24	C*07:02	C*12:02	C*07:02	C*12:02	sim
25	C*07:02	C*12:03	C*07:02	C*12:03	sim
26	C*01:02	C*07:02	C*01:02	C*07:02	sim

Exemplo 10: Genotipagem HLA-DQB1 usando a técnica de sequenciamento de segunda geração (Illumina Solexa)

[0187] 94 amostras de sangue com os resultados de tipagem HLA-SBT conhecidos foram submetidos a genotipagem HLA-DQB1, de acordo com os métodos como descrito no Exemplo 8, exceto os seguintes itens.

[0188] 94 conjuntos de PCR de iniciador de indexação foram usados para amplificar 94 amostras de DNA, respectivamente, onde cada conjunto de PCR de iniciador de indexação consistiu de Iniciadores de PCR par a amplificação de Éxon 2 ou 3 de HLA-DQB1 (Tabela 5) e um par de iniciadores de indexação bidirecionais (como descrito acima), cada Iniciador de PCR dianteiro tem um iniciador de indexação dianteiro de um par de iniciadores de indexação ligados na 5' extremidade, e o

iniciador de PCR tem um iniciador de indexação reverso de um par de iniciadores de indexação ligados 5' extremidade. Durante a síntese dos iniciadores, os iniciadores de indexação foram adicionados diretamente à 5' extremidade dos Iniciadores de PCR, onde os iniciadores foram sintetizados por Shanghai Invitrogen Co.

[0189] Os 94 DNAs obtidos na etapa de extração de amostra foram designados como No. 1-94. A reação de PCR foi realizada em placa fonte 96, Éxons 2, 3 de DQB1 na qual a amostra foi amplificada na mesma fonte. Dois controles negativos sem adicionar qualquer modelo foram configurados em cada placa, e os iniciadores de indexação usados nos controles negativos são PI-95 e PI-96. Durante a experimentação, a informação de quantidade das amostras correspondendo a cada par de iniciadores de indexação foi gravada.

[0190] Os iniciadores de indexação usados foram os iniciadores de indexação PI-1 a PI-94 como listado na Tabela 6, e os seguintes iniciadores de indexação PI-95 e PI-96 (Tabela 11) para controles negativos.

Tabela 11: Informação relevante nos iniciadores de indexação usados para controles negativos

PI-95	SEQ ID NO:229	CGACGTAGAGTC	SEQ ID NO:230	CAGTAGCACTAC	H11	Controle negativo
PI-96	SEQ ID NO:239	CACTGTATAGCT	SEQ ID NO:240	CGACTAGTACTA	H12	Controle negativo

Procedimento PCR para HLA-DQB1 foi como segue:

96°C 2min

95°C 30s→60°C 30s→72°C 20s (32 ciclos)

15°C ∞

Sistema de reação PCR para HLA-DQB1 foi como segue:

Promega 5×buffer I (Mg ²⁺ plus)	5,0 ul
dNTP mistura (2,5 mM cada)	2,0 ul
PIInf-Q-F2 (2 pmol/ul)	1,0 ul
PIInf-Q-R2 (2 pmol/ul)	1,0 ul

Plnf-Q-F3 (2 pmol/ul)	1,0 ul
Plnf-Q-R3 (2 pmol/ul)	1,0 ul
Promega Taq (5U/ul)	0,2 ul
DNA (cerca de 20 ng/ul)	5,0 ul
ddH ₂ O	8,8 ul
Total	25,0 ul

onde, Plnf-Q-F2/3 representa o F iniciador de HLA-DQB1 tendo a sequência de iniciador de indexação dianteiro No. n (Tabela 1) na 5' extremidade; Plnf-Q-R 2/3 representa o R iniciador de HLA-DQB1 tendo uma sequência primária reversa No. n na 5' extremidade (aqui $n \leq 96$); e o resto pode ser deduzido similarmente. Além disso, cada amostra corresponde a um conjunto específico de iniciador de PCR.

[0191] As reações de PCR foram realizadas no aparelho PTC-200 PCR da Bio-Rad Co.. Depois da reação de PCR, 2ul de produto PCR foi submetido a eletroforese 1,5% agarose gel. A figura 11 mostrou o resultado eletroforético dos produtos de PCR de Éxons 2+3 de HLA-DQB1 de 94 amostras, e o marcador molecular de DNA foi DL 2000 (Takara Co.).

[0192] Associação e purificação de produtos PCR

[0193] 20 µl do resto dos produtos de PCR foram tomados de cada fonte da placa fonte 96 HLA-P-DQB1 (exceto para o controle negativo), e foi misturado homogeneamente em um tubo EP de 3 ml (designado como HLA-Q-Mix). 500ul de mistura de DNA de HLA-Q-Mix foi submetida a purificação de coluna com um conjunto de Purificação de DNA Qiagen (QIAGEN Co.) (Para as etapas de purificação específicas, por favor, refira-se às instruções do fabricante). Foi determinado por Nanodrop 8000 (Thermo Fisher Scientific Co.) que os 200ul de DNA obtido por purificação tem uma concentração de DNA HLA-Q-Mix DNA de 48ng/ul.

[0194] Condições para o cisalhamento foi como segue:

Frequência de varredura

Ciclo de trabalho	10%
-------------------	-----

Intensidade	5
Ciclos/Explosão	200
Tempo(s)	300

[0195] Os produtos de reação foram submetidos a uma reação de reparação terminal, e então foram recuperados e purificados por um Conjunto de Purificação de PCR QIAquick, e foram dissolvidos em 34ul EB (QIAGEN Elution Buffer).

[0196] Os produtos de reação foram ainda submetidos a adição de A na 3' extremidade, e então foram recuperados e purificados por um conjunto de purificação de PCR MiniElute (QIAGEN Co.), e foram dissolvidos em 13 µl EB de solução (QIAGEN Elution Buffer).

[0197] Depois da ligação dos adaptadores de biblioteca, os produtos de reação foram purificados por Ampure Beads (Beckman Coulter Genomics), e foram dissolvidos em 50 µl de água desionizada, e a concentração de DNA determinada por um quantitativo Fluorescente de PCR (QPCR) foi como segue:

	Resultado determinado por qPCR (nM)
HLA-Q-Mix	115,3

[0198] O gel contendo os fragmentos de DNA em média de 350 a 550bp foi fatiado (Figura 12). Depois da purificação e recuperação dos produtos do gel, a concentração de DNA, como determinado pelo quantitativo Fluorescente de PCR Fluorescence (QPCR), foi 18,83 nM.

[0199] Análise dos resultados

[0200] Os resultados do sequenciamento de Illumina GA foram uma série de sequências de DNA, e procurando as sequências primárias dianteiras e reversas e as sequências primárias nos resultados de sequenciamento, o banco de dados compreendendo os resultados de sequenciamento dos produtos de PCR de vários Éxons de HLA-DQB1 para cada amostra correspondente ao iniciador de indexação respectivo foram construídas. Os resultados de sequenciamento de cada Éxon foram alinhados à sequência de referência (as sequências de referência foram da

<http://www.ebi.ac.uk/imgt/hla/>) do Éxon correspondente por BWA (Burrows-Wheeler Aligner), e as sequências de consenso foram construídas, as sequências de DNA lidas foram selecionadas e corrigidas com base no valor qualidade da base de sequenciamento e diferença entre a sequência lida e as sequências de consenso. As sequências de DNA corrigidas foram montadas nas sequências correspondentes de Éxons 2, 3 de HLA-DQB1 com base na sequência de sobreposição e relacionamento de ligação (ligação de extremidade em pares). A captura de tela da figura 13 ilustra o procedimento para a construção da sequência de consenso de Éxon 2 de HLA-DQB1 no local na amostra No.7.

[0201] A sequência de DNA resultante para Éxons 2, 3 de HLA-DQB1 foi alinhada com a sequência do banco de dados profissional do Éxon correspondente HLA-DQB1 em IMGT HLA. Se o resultado da sequência de alinhamento mostrou 100% de combinação, o genótipo HLA-DQB1 da amostra correspondente foi determinado.

[0202] Para todas as 94 amostras, os resultados de tipagem obtidos pelos métodos acima foram completamente consistentes com os resultados de tipagem conhecidos, onde os resultados das Amostras No. 1-32 foram mostrados na Tabela 12.

Tabela 12: Os resultados da tipagem das Amostras No. 1-32

Amostra No.	Genótipo Original conhecido DQB1		Resultados para DQB1 obtidos neste momento		Idêntico ou não
1	DQB1*02:02	DQB1*03:01	DQB1*02:02	DQB1*03:01	sim
2	DQB1*02:02	DQB1*04:01	DQB1*02:02	DQB1*04:01	sim
3	DQB1*05:02	DQB1*02:02	DQB1*05:02	DQB1*02:02	sim
4	DQB1*02:02	DQB1*06:03	DQB1*02:02	DQB1*06:03	sim
5	DQB1*03:03	DQB1*04:02	DQB1*03:03	DQB1*04:02	sim
6	DQB1*05:02	DQB1*03:17	DQB1*05:02	DQB1*03:17	sim
7	DQB1*03:03	DQB1*06:02	DQB1*03:03	DQB1*06:02	sim
8	DQB1*05:03	DQB1*04:02	DQB1*05:03	DQB1*04:02	sim
9	DQB1*04:02	DQB1*06:01	DQB1*04:02	DQB1*06:01	sim
10	DQB1*05:01	DQB1*06:10	DQB1*05:01	DQB1*06:10	sim
11	DQB1*03:01	DQB1*03:03	DQB1*03:01	DQB1*03:03	sim
12	DQB1*05:01	DQB1*05:01	DQB1*05:01	DQB1*05:01	sim
13	DQB1*02:02	DQB1*04:02	DQB1*02:02	DQB1*04:02	sim
14	DQB1*05:02	DQB1*02:01	DQB1*05:02	DQB1*02:01	sim
15	DQB1*02:01	DQB1*06:02	DQB1*02:01	DQB1*06:02	sim
16	DQB1*03:03	DQB1*04:01	DQB1*03:03	DQB1*04:01	sim
17	DQB1*05:01	DQB1*03:02	DQB1*05:01	DQB1*03:02	sim
18	DQB1*03:03	DQB1*06:01	DQB1*03:03	DQB1*06:01	sim
19	DQB1*03:03	DQB1*06:10	DQB1*03:03	DQB1*06:10	sim
20	DQB1*05:03	DQB1*04:01	DQB1*05:03	DQB1*04:01	sim
21	DQB1*05:02	DQB1*04:01	DQB1*05:02	DQB1*04:01	sim
22	DQB1*03:01	DQB1*03:03	DQB1*03:01	DQB1*03:03	sim
23	DQB1*05:02	DQB1*05:03	DQB1*05:02	DQB1*05:03	sim
24	DQB1*05:02	DQB1*03:02	DQB1*05:02	DQB1*03:02	sim
25	DQB1*03:03	DQB1*06:01	DQB1*03:03	DQB1*06:01	sim
26	DQB1*05:02	DQB1*06:09	DQB1*05:02	DQB1*06:09	sim
27	DQB1*02:02	DQB1*06:02	DQB1*02:02	DQB1*06:02	sim
28	DQB1*05:02	DQB1*03:01	DQB1*05:02	DQB1*03:01	sim
29	DQB1*02:01	DQB1*03:01	DQB1*02:01	DQB1*03:01	sim
30	DQB1*06:03	DQB1*06:09	DQB1*06:03	DQB1*06:09	sim
31	DQB1*05:02	DQB1*02:02	DQB1*05:02	DQB1*02:02	sim
32	DQB1*05:01	DQB1*06:01	DQB1*05:01	DQB1*06:01	sim

Exemplo 11: Genotipagem HLA-DQB1 usando um método de sequenciamento usando Sanger

1. Extração de amostra de DNA

[0203] Como descrito no exemplo 1, os DNAs foram extraídos usando um Instrumento de Extração Automático KingFisher de 20 para 94 amostras com os genótipos HLA conhecidos.

2. Amplificação de PCR

[0204] Os DNAs acima, extraídos pelo Instrumento de Extração Automática KingFisher, foram usados como modelos, e as amplificações de PCR foram realizadas em tubos únicos usando dois pares de Iniciadores de PCR (Q-F2 e Q-R2, Q-F3 e Q-R3) como listado na Tabela 5, respectivamente. O procedimento PCR para cada par de iniciadores foi como segue:

96°C 2min

95°C 30s→56°C 30s→72°C 20s (35 ciclos)

15°C ∞

Sistema de reação de PCR para HLA-Q foi como segue:

Promega 5×buffer I (Mg ²⁺ plus)	5,0 µL
dNTP Mistura (2,5 mM cada)	2,0 µL
Mistura primária (25 ng/µL)	3,0 µL
Promega Taq (5 U/µL)	0,2 µL
DNA (cerca de 20 ng/µL)	2,0 µL
ddH ₂ O	12,8 µL
total	25,0 µL

[0205] Os produtos PCR foram submetidos a uma eletroforese agarose gel antes da purificação.

3. Purificação de produtos PCR

[0206] Os métodos e as etapas foram as mesmas daquelas descritas no Exemplo 9.

4. Realização da reação de sequenciamento e purificação dos produtos da reação de sequenciamento

[0207] O método e as etapas foram as mesmas daquelas descritas no Exemplo 9

5. Análise de resultados e sequenciamento

[0208] Os produtos purificados da reação de sequenciamento foram submetidos a um sequenciamento de eletroforese capilar em ABI 3730XL. Os picos do sequenciamento foram analisados pelo software Tipo-u (Invitrogen) para obter os resultados de tipagem HLA (Figura 15). Todos os resultados obtidos pelo método acima foram idênticos aos resultados originais conhecidos, como mostrado na Tabela 13.

Tabela 13: Comparação dos resultados de tipagem obtidos pelo método acima com os resultados de tipagem originais conhecidos

Amostra No.	Genótipo original conhecido DQB1		Resultados para DQB1 obtido neste momento		Idêntico ou não
1	DQB1*02:02	DQB1*03:01	DQB1*02:02	DQB1*03:01	sim
2	DQB1*02:02	DQB1*04:01	DQB1*02:02	DQB1*04:01	sim
3	DQB1*05:02	DQB1*02:02	DQB1*05:02	DQB1*02:02	sim
4	DQB1*02:02	DQB1*06:03	DQB1*02:02	DQB1*06:03	sim
5	DQB1*03:03	DQB1*04:02	DQB1*03:03	DQB1*04:02	sim
6	DQB1*05:02	DQB1*03:17	DQB1*05:02	DQB1*03:17	sim
7	DQB1*03:03	DQB1*06:02	DQB1*03:03	DQB1*06:02	sim
8	DQB1*05:03	DQB1*04:02	DQB1*05:03	DQB1*04:02	sim
9	DQB1*04:02	DQB1*06:01	DQB1*04:02	DQB1*06:01	sim
10	DQB1*05:01	DQB1*06:10	DQB1*05:01	DQB1*06:10	sim
11	DQB1*03:01	DQB1*03:03	DQB1*03:01	DQB1*03:03	sim
12	DQB1*05:01	DQB1*05:01	DQB1*05:01	DQB1*05:01	sim
13	DQB1*02:02	DQB1*04:02	DQB1*02:02	DQB1*04:02	sim
14	DQB1*05:02	DQB1*02:01	DQB1*05:02	DQB1*02:01	sim

15	DQB1*02:01	DQB1*06:02	DQB1*02:01	DQB1*06:02	sim
16	DQB1*03:03	DQB1*04:01	DQB1*03:03	DQB1*04:01	sim
17	DQB1*05:01	DQB1*03:02	DQB1*05:01	DQB1*03:02	sim
18	DQB1*03:03	DQB1*06:01	DQB1*03:03	DQB1*06:01	sim
19	DQB1*03:03	DQB1*06:10	DQB1*03:03	DQB1*06:10	sim
20	DQB1*05:03	DQB1*04:01	DQB1*05:03	DQB1*04:01	sim

Exemplo 12: Genotipagem de Éxons 2, 3, 4 de HLA-A/B/C e Éxons 2, 3 de HLA-DQB1 em 950 amostras

[0209] Nos presentes exemplos, os Éxons 2, 3, 4 de HLA-A/B/C e Éxons 2, 3 de HLA-DQB1 em 950 amostras foram genotipados usando uma combinação de iniciadores de indexação, estratégia de cisalhamento de DNA incompleto, índices de biblioteca, preparação de biblioteca de PCR livre, e técnica de sequenciamento 100 de extremidade em pares Illumina GA (produtos PCR tendo um comprimento na média de 300bp a 500bp), demonstrando que o método da presente invenção pode realizar a genotipagem dos fragmentos de gene de um comprimento excedendo o comprimento máximo lido do sequenciador, e também demonstra que a presente invenção pode realizar a genotipagem HLA com baixo custo, alto rendimento, alta precisão, alta segurança e alta resolução.

[0210] Princípio: as amostras a serem analisadas foram divididas em 10 grupos; para amostra de cada grupo, os iniciadores de indexação foram introduzidos aos dois terminais dos produtos de PCR de Éxons 2, 3, 4 de HLA-A/B/C e Éxons 2, 3 de HLA-DQB1 por reação de PCR de forma a legendar especificamente a informação da amostra Dops produtos de PCR. Os produtos de amplificação de PCR de quarto locais (HLA-A/B/C/DQB1) em cada grupo de amostras foram misturados juntos para obter uma biblioteca de produtos de PCR; depois do cisalhamento ultrassônico foram construídas (onde para as bibliotecas de produtos de PCR de cada grupo, um adaptador diferente foi usado, construindo assim 10 bibliotecas de sequenciamento indexadas). As 10 bibliotecas de sequenciamento indexadas foram associadas juntas

em um mol igual para construir uma biblioteca de sequenciamento de índice misturada. A biblioteca de sequenciamento de índice mistura foi submetida a uma 2% baixo ponto de derretimento, e todas as faixas de DNA de um comprimento na média de 450bp a foi sequenciado pelo método Illumina GA PE-100. A informação de sequência de todas as amostras testadas pode ser traçado pelas sequências de iniciador de indexação e as sequências de índice de biblioteca, e a sequência de todo o produto PCR pode ser montada com base nas sequências de referência conhecidas e a relação de sobreposição e ligação entre as sequências de fragmentos de DNA, a sequência completa do produto de PCR original pode ser alinhada com o banco de dados padrão dos Éxons correspondentes HLA-A/B/C/DQB1, realizando assim a genotipagem HLA-A/B/C/DQB1.

1. Extração de amostras

[0211] Os DNAs foram extraídos de 950 amostras de sangue com resultados de tipagem HLA-SBT conhecidos (Programa de Doador de Medula Óssea da China citado aqui como (CMDP)) usando um Instrumento de Extração Automática (US Thermo Co.). o processo foi o mesmo como aquele descrito no Exemplo 1.

2. Amplificação de PCR

[0212] Os 950 DNAs obtidos da etapa de extração de amostra foram designados como No. 1-950, e foram divididos em grupos de 10 (95 DNAs para cada grupo), que foram designados como HLA-1, HLA-2, HLA-3, HLA-4, HLA-5, HLA-6, HLA-7, HLA-8, HLA-9, HLA-10. Para cada grupo de amostras, 95 amostras de DNA foram amplificadas por 95 conjuntos de Iniciadores de PCR carregando iniciadores de indexação bidirecionais (Tabela 6) para a amplificação dos Éxons 2, 3, 4 de HLA-A/B (Tabela 2), para a amplificação dos Éxons 2, 3, 4 de HLA-C (Tabela 4), e para a amplificação de Éxons 2, 3 de HLA-DQB1 (Tabela 5), respectivamente. A reação de PCR aconteceu nas placas de fonte 96, usando placas 100 no total, designadas como HLA-X-P-A2, HLA-X-P-A3, HLA-X-P-A4, HLA-X-P-B2, HLA-X-P-B3, HLA-X-P-B4,

HLA-X-P-C2, HLA-X-P-C3, HLA-X-P-C4 e HLA-X-P-DQB1 (“X” representa a informação do número de grupo 1/2/3/4/5/6/7/8/9/10, “A2/3/4”, “B2/3/4”, “C2/3/4”, “DQB1” representa os locais de), onde um controle negativo sem a adição de qualquer modelo foi configurada em cada placa, e os iniciadores usados para o controle negativo foram iniciadores legendados por PI-1 (Tabela 6). Durante a experimentação, a informação de cada amostra no número de grupo e os iniciadores de indexação foram gravados. Por exemplo, a informação relevante nos iniciadores de indexação PI-1 e PI-2 foi como segue, e o resto pode ser deduzido similarmente.

Iniciador de indexação No.	Iniciador de Indexação dianteiro	Iniciador de Indexação reverso	Posição correspondente e na placa fonte 96	Modelo correspondente (Grupo 1)	Modelo correspondente (Grupo n+1, onde $1 \leq n < 10$, n foi um número inteiro)
PI-1	TCGCAGA CATCA	TGACACG ATGCT	A1	1	$1+95*n$
PI-2	TACATCGC ACTA	TACAGAT GCTGA	A2	2	$2+95*n$

[0213] O procedimento de PCR e o sistema de reação de PCR para HLA-A/B/C foram os mesmos daqueles descritos no Exemplo 2. Os Iniciadores de PCR para a amplificação dos Éxons correspondentes de HLA-A/B foram mostrados na Tabela 2, e os Iniciadores de PCR para a amplificação dos Éxons correspondentes de HLA-C foram mostrados na Tabela 4.

[0214] O procedimento de PCR para HLA-DQB1 foi como segue.

96°C 2min

95°C 30s→55°C 30s→72°C 20s (32ciclos)

15°C ∞

[0215] O sistema de reação múltipla de PCR para HLA-DQB1 (amplificação de Éxons 2, 3 simultaneamente) é o mesmo daquele descrito no Exemplo 10, e os Iniciadores de PCR para a amplificação dos Éxons correspondentes de HLA-DQB1

foram como mostrado na Tabela 5.

[0216] Onde, $PI_{nf-A/B/C-F2/3/4}$ e $PI_{nf-Q-F2/F3}$ representam os F iniciadores de HLA-A/B/C/DQB1 tendo uma sequência de iniciador de indexação dianteiro No. n (Tabela 6) na 5' extremidade, $PI_{nr-A/B/C-R2/3/4}$ e $PI_{nr-Q-R2/R3}$ representa os R iniciadores de HLA-A/B/C/DQB1 tendo uma sequência de índice reverso No. n na 5' extremidade (aqui $n \leq 95$), e o resto pode ser deduzido similarmente. Além disso, cada amostra corresponde a um conjunto específico de Iniciadores de PCR ($PI_{nf-A/B/C-F2/3/4}$, $PI_{nr-A/B/C-R2/3/4}$, $PI_{nf-Q-F2/F3}$, $PI_{nr-Q-R2/R3}$).

[0217] A reação de PCR foi realizada no aparelho PTC-200 PCR da Bio-Rad Co.. Depois da reação de PCR, 3ul de produtos de PCR foram submetidos a uma eletroforese 2% agarose gel. A Figura 16 mostrou o resultado eletroforético dos produtos de PCR dos Éxons correspondentes de HLA-A/B/C/DQB1 da Amostra No. 1, e o marcador molecular de DNA foi DL 2000 (Takara Co.). Havia uma série de faixas de únicas de um comprimento na média de 300 bp a 500 bp no eletroforograma, indicando sucesso na amplificação de PCR de Éxons (A2, A3, A4, B2, B3, B4, C2, C3, C4, DRB1) de HLA-A/B/C/DQB1 da Amostra No.1. Não havia faixa de amplificação no controle negativo (N). os resultados das outras amostras foram similares.

3. Associação e purificação de produtos de PCR

[0218] Para as amostras do Grupo X ("X" é 1/2/3/4/5/6/7/8/9/10), 20 µl do resto de produtos de PCR foram tomados de cada fonte de placa fonte 96 HLA-X-P-A2 (exceto para o controle negativo), e foi misturada homogeneamente sob agitação em um tubo EP de 3 ml (designado como HLA-X-A2-Mix). A mesma operação foi aplicada a outras 9 placas fonte 96 das amostras do Grupo X, designado como HLA-X-A3-Mix, HLA-X-A4-Mix, HLA-X-B2-Mix, HLA-X-B3-Mix, HLA-X-B4-Mix, HLA-X-C2-Mix, HLA-X-C3-Mix, HLA-X-C4-Mix, e HLA-X-DQB1-Mix. 200ul foi tomado de cada HLA-X-A2-Mix, HLA-X-A3-Mix, HLA-X-A4-Mix, HLA-X-B2-Mix, HLA-X-B3-Mix, HLA-X-B4-Mix, HLA-X-C2-Mix, HLA-X-C3-Mix, HLA-X-C4-Mix, e HLA-X-DQB1-Mix, e foi misturado em um

tubo EP de 3ml, designado como HLA-X-Mix. 500ul de mistura de DNA de HLA-X-Mix foi submetido a uma purificação de coluna com um Conjunto de Purificação de DNA Qiagen (QIAGEN Co.) (Para as etapas de purificação específicas, por favor, refira-se as instruções do fabricante) para obter 200ul de DNA purificado, cada concentração de DNA foram determinadas por Nanodrop 8000 (Thermo Fisher Scientific Co.). A mesma operação também foi aplicada a outros 9 grupos de amostras. As concentrações de DNA finalmente determinadas dos 10 grupos de amostras foram como segue.

	HLA-1-Mix	HLA-2-Mix	HLA-3-Mix	HLA-4-Mix	HLA-5-Mix	HLA-6-Mix	HLA-7-Mix	HLA-8-Mix	HLA-9-Mix	HLA-10-Mix
Conc. (ng/ul)	53,1	52,3	56,1	57,2	50,5	55,7	54,2	58,6	53,9	54,8

4. Construção de bibliotecas de sequenciamento de Illumina GA

[0219] Como descrito no Exemplo 4, uma quantidade total de 5ug de DNA, tomado do HLA-X-Mix purificado, foi submetido a um cisalhamento de DNA, purificação depois do cisalhamento, reação de reparação terminal, adição de A na 3' extremidade; e ligação de adaptador de biblioteca de PCR livre Illumina GA.

[0220] O relacionamento correspondente entre os grupos de amostras e os adaptadores de biblioteca foram como segue.

Grupo de amostra No.	HLA-1	HLA-2	HLA-3	HLA-4	HLA-5	HLA-6	HLA-7	HLA-8	HLA-9	HLA-10
Adaptador de biblioteca No.	1	2	3	4	5	6	7	8	9	10

Os produtos de reação obtidos foram purificados por Ampure Beads (Beckman Coulter Genomics), e foram dissolvidos em 50ul de água desionizada, e as concentrações de DNA determinadas por um quantitativo Fluorescente de PCR (QPCR) foram como segue:

	HLA-1-Mix	HLA-2-Mix	HLA-3-Mix	HLA-4-Mix	HLA-5-Mix	HLA-6-Mix	HLA-7-Mix	HLA-8-Mix	HLA-9-Mix	HLA-10-Mix
Conc./nM	86,60	78,21	54,56	87,35	84,37	85,09	96,21	85,81	88,14	88,26

6. Recuperação por gel fatiado

[0221] HLA-1-Mix, HLA-2-Mix, HLA-3-Mix, HLA-4-Mix, HLA-5-Mix, HLA-6-Mix, HLA-7-Mix, HLA-8-Mix, HLA-9-Mix e HLA-10-Mix foram misturados em um mol igual (concentração final foi de 70,86nM/ul), designada como HLA-Mix-10, 30µL HLA-Mix-10 foi submetido a eletroforese 2% baixo ponto de derretimento de agarose gel. A condição eletroforética foi de 100V, 100min. O marcador de DNA foi o 50bp marcador de DNA da NEB Co.. O gel contendo os fragmentos de DNA na média de 450 a 750bp foi fatiado (Figura 17). Os produtos no gel fatiado foram recuperados e purificados por um Conjunto de Purificação de PCR QIAquick (QIAGEN Co.), o volume depois da purificação foi de 32ul, a concentração de DNA medida pelo quantitativo Fluorescente de PCR (QPCR) foi de 10,25nM.

5. Análise de resultados e sequenciamento Illumina GA

[0222] As análises de resultado e sequenciamento foram realizadas de acordo com os métodos como descrito nos Exemplos 5 e 6.

[0223] O banco de dados, compreendendo os resultados das sequenciamentos dos produtos de PCR de vários Éxons de HLA-A/B/C/DQB1 para cada amostra correspondente a um iniciador de indexação respectivo foi construída. A sequência de DNA resultante foi alinhada com a sequência de banco de dados dos Éxons correspondentes do banco de dados profissional HLA-A/B/C/DQB1 em IMGT HLA. Se o resultado do alinhamento de sequência mostrou 100% de combinação, o genótipo HLA-A/B/C/DQB1 da amostra correspondente foi determinado. Por favor, refira-se a captura de tela do programa para a construção da sequência de consenso Éxon 2 de local HLA-C na amostra No.1, como ilustrado na figura 18. Para todas as 950 amostras, os resultados de tipagem obtidos pelo método acima foram

completamente consistentes com os resultados de tipagem originais conhecidos, onde os resultados das Amostras No. 1-32 foram como segue:

No.	Tipo Original conhecido HLA-A/B/C/DQB1
1	A*02:07 A*03:01 B*07:02 B*46:01 C*01:02 C*07:02 DQB1*03:03DQB1*06:02
2	A*11:01 A*31:01 B*15:11 B*38:02 C*03:03 C*07:02 DQB1*03:03DQB1*04:01
3	A*02:07 A*24:02 B*13:01 B*46:01 C*01:02 C*03:04 DQB1*03:02DQB1*06:01
4	A*24:02 A*33:03 B*40:01 B*51:01 C*01:02 C*14:02 DQB1*03:03DQB1*03:03
5	A*31:01 A*31:01 B*15:01 B*35:01 C*04:01 C*04:01 DQB1*03:02DQB1*06:02
6	A*02:07 A*03:01 B*44:02 B*46:01 C*01:02 C*05:01 DQB1*03:01DQB1*06:02
7	A*02:01 A*30:01 B*07:02 B*13:02 C*06:02 C*07:02 DQB1*02:02DQB1*06:01
8	A*02:07 A*02:07 B*46:01 B*46:01 C*01:02 C*01:02 DQB1*05:02DQB1*03:03
9	A*01:01 A*33:03 B*49:01 B*58:01 C*03:02 C*07:01 DQB1*06:04DQB1*06:09
10	A*02:07 A*11:01 B*46:01 B*48:01 C*01:03 C*08:01 DQB1*05:03DQB1*03:02
11	A*02:06 A*30:01 B*13:02 B*15:02 C*06:02 C*08:01 DQB1*03:01DQB1*03:01
12	A*24:02 A*31:01 B*35:01 B*51:01 C*03:03 C*14:02 DQB1*03:03DQB1*06:01
13	A*11:01 A*33:03 B*46:01 B*46:01 C*01:02 C*01:02 DQB1*03:02DQB1*03:03
14	A*01:01 A*02:03 B*38:02 B*57:01 C*06:02 C*07:02 DQB1*05:02DQB1*03:03
15	A*02:06 A*24:02 B*13:01 B*15:25 C*03:04 C*07:02 DQB1*03:01DQB1*06:01
16	A*11:01 A*24:02 B*15:02 B*15:27 C*04:01 C*08:01 DQB1*03:01DQB1*03:03
17	A*24:02 A*24:02 B*40:01 B*46:01 C*01:02 C*03:04 DQB1*03:03DQB1*06:02
18	A*24:02 A*24:02 B*40:01 B*46:01 C*01:02 C*03:04 DQB1*03:03DQB1*06:02
19	A*11:01 A*33:03 B*40:02 B*58:01 C*03:02 C*03:04 DQB1*02:01DQB1*03:02
20	A*24:02 A*30:01 B*13:02 B*40:01 C*03:04 C*06:02 DQB1*06:02DQB1*06:03
21	A*02:01 A*24:02 B*40:01 B*40:01 C*07:02 C*14:02 DQB1*04:02DQB1*06:02
22	A*02:01 A*33:03 B*15:01 B*44:03 C*01:02 C*14:03 DQB1*03:01DQB1*06:04
23	A*26:01 A*33:03 B*15:01 B*58:01 C*03:02 C*08:01 DQB1*02:01DQB1*03:01
24	A*02:01 A*11:01 B*13:01 B*55:02 C*01:06 C*03:04 DQB1*03:01DQB1*03:03
25	A*02:01 A*32:01 B*40:01 B*52:01 C*03:04 C*12:02 DQB1*05:02DQB1*06:01
26	A*02:03 A*02:07 B*40:01 B*46:01 C*01:02 C*07:02 DQB1*03:02DQB1*06:01
27	A*02:07 A*02:07 B*46:01 B*46:01 C*01:02 C*01:02 DQB1*03:03DQB1*06:01
28	A*24:02 A*30:01 B*13:02 B*39:05 C*06:02 C*07:02 DQB1*02:02DQB1*06:01
29	A*31:01 A*33:03 B*15:18 B*58:01 C*03:02 C*07:04 DQB1*04:01DQB1*06:09

30 A*02:06 A*03:01 B*27:05 B*40:02 C*02:02 C*03:04 DQB1*03:01DQB1*03:01
 31 A*02:06 A*33:03 B*15:02 B*58:01 C*03:02 C*08:01 DQB1*05:01DQB1*06:01
 32 A*03:01 A*30:01 B*13:02 B*51:01 C*06:02 C*15:02 DQB1*02:02DQB1*03:01

No.	HLA-A/B/C/DQB1 tipo determinado pelo método da invenção
1	A*02:07 A*03:01 B*07:02 B*46:01 C*01:02 C*07:02 DQB1*03:03DQB1*06:02
2	A*11:01 A*31:01 B*15:11 B*38:02 C*03:03 C*07:02 DQB1*03:03DQB1*04:01
3	A*02:07 A*24:02 B*13:01 B*46:01 C*01:02 C*03:04 DQB1*03:02DQB1*06:01
4	A*24:02 A*33:03 B*40:01 B*51:01 C*01:02 C*14:02 DQB1*03:03DQB1*03:03
5	A*31:01 A*31:01 B*15:01 B*35:01 C*04:01 C*04:01 DQB1*03:02DQB1*06:02
6	A*02:07 A*03:01 B*44:02 B*46:01 C*01:02 C*05:01 DQB1*03:01DQB1*06:02
7	A*02:01 A*30:01 B*07:02 B*13:02 C*06:02 C*07:02 DQB1*02:02DQB1*06:01
8	A*02:07 A*02:07 B*46:01 B*46:01 C*01:02 C*01:02 DQB1*05:02DQB1*03:03
9	A*01:01 A*33:03 B*49:01 B*58:01 C*03:02 C*07:01 DQB1*06:04DQB1*06:09
10	A*02:07 A*11:01 B*46:01 B*48:01 C*01:03 C*08:01 DQB1*05:03DQB1*03:02
11	A*02:06 A*30:01 B*13:02 B*15:02 C*06:02 C*08:01 DQB1*03:01DQB1*03:01
12	A*24:02 A*31:01 B*35:01 B*51:01 C*03:03 C*14:02 DQB1*03:03DQB1*06:01
13	A*11:01 A*33:03 B*46:01 B*46:01 C*01:02 C*01:02 DQB1*03:02DQB1*03:03
14	A*01:01 A*02:03 B*38:02 B*57:01 C*06:02 C*07:02 DQB1*05:02DQB1*03:03
15	A*02:06 A*24:02 B*13:01 B*15:25 C*03:04 C*07:02 DQB1*03:01DQB1*06:01
16	A*11:01 A*24:02 B*15:02 B*15:27 C*04:01 C*08:01 DQB1*03:01DQB1*03:03
17	A*24:02 A*24:02 B*40:01 B*46:01 C*01:02 C*03:04 DQB1*03:03DQB1*06:02
18	A*24:02 A*24:02 B*40:01 B*46:01 C*01:02 C*03:04 DQB1*03:03DQB1*06:02
19	A*11:01 A*33:03 B*40:02 B*58:01 C*03:02 C*03:04 DQB1*02:01DQB1*03:02
20	A*24:02 A*30:01 B*13:02 B*40:01 C*03:04 C*06:02 DQB1*06:02DQB1*06:03
21	A*02:01 A*24:02 B*40:01 B*40:01 C*07:02 C*14:02 DQB1*04:02DQB1*06:02
22	A*02:01 A*33:03 B*15:01 B*44:03 C*01:02 C*14:03 DQB1*03:01DQB1*06:04
23	A*26:01 A*33:03 B*15:01 B*58:01 C*03:02 C*08:01 DQB1*02:01DQB1*03:01
24	A*02:01 A*11:01 B*13:01 B*55:02 C*01:06 C*03:04 DQB1*03:01DQB1*03:03
25	A*02:01 A*32:01 B*40:01 B*52:01 C*03:04 C*12:02 DQB1*05:02DQB1*06:01
26	A*02:03 A*02:07 B*40:01 B*46:01 C*01:02 C*07:02 DQB1*03:02DQB1*06:01
27	A*02:07 A*02:07 B*46:01 B*46:01 C*01:02 C*01:02 DQB1*03:03DQB1*06:01

28 A*24:02 A*30:01 B*13:02 B*39:05 C*06:02 C*07:02 DQB1*02:02DQB1*06:01
 29 A*31:01 A*33:03 B*15:18 B*58:01 C*03:02 C*07:04 DQB1*04:01DQB1*06:09
 30 A*02:06 A*03:01 B*27:05 B*40:02 C*02:02 C*03:04 DQB1*03:01DQB1*03:01
 31 A*02:06 A*33:03 B*15:02 B*58:01 C*03:02 C*08:01 DQB1*05:01DQB1*06:01
 32 A*03:01 A*30:01 B*13:02 B*51:01 C*06:02 C*15:02 DQB1*02:02DQB1*03:01

Note: no caso em que as sequências de Éxons 2, 3, 4 de HLA-A/B/C foram completamente idênticas, um tipo comum foi selecionado.

[0224] 950 amostras com resultados de tipagem HLA-SBT conhecidos foram submetidas a genotipagem de locais HLA-A/B/C/DQB1 pela estratégia técnica da presente invenção foram completamente consistentes com os resultados originais conhecidos.

[0225] Embora as modalidades da presente invenção já tenham sido descritas em detalhes, um perito na arte entenderia aquilo com base em todos os ensinamentos como divulgado, várias modificações e substituições podem ser feitas sem partir do escopo e do espírito da presente invenção. O escopo da presente invenção é definido pelas reivindicações anexadas e qualquer equivalente das mesmas.

Referências

- [1]. <http://www.ebi.ac.uk/imgt/hla/stats.html>
- [2]. Tiercy J M. Base molecular de polimorfismo HLA: implicações no transplante clínico. [J]. Transpl Immunol, 2002, 9: 173-180.
- [3]. C.Antoine, S.Müller, A.Cant, et al. vida longa de sobrevivência e transplante de células tronco hemopoiéticas para imunodeficiências: relato da experiência europeia. 1968-99. [J]. The Lancet, 2003, 361:553-560.
- [4]. H. A. Erlich, G. Opelz, J. Hansen, et al. Tipagem e transplantação de DNA HLA.[J].Immunity, 2001, 14:347-356.
http://www.sciencedirect.com/science?_ob=ArticleURL&_udi=B6WSP-42YW740-2&_user=10&_coverDate=04%2F30%2F2001&_rdoc=1&_fmt=high&_orig=search&

[sort=d& docanchor=&view=c& searchStrId=1194405262& rerunOrigin=scholar.google& acct=C000050221& version=1& urlVersion=0& userid=10&md5=bbe8e010b8bc54730458f34cddd8ec1d - aff2](#)

[5]. Lillo R, Balas A, Vicario JL, et al. Duas novas classes de alelos HLA, DPB1*02014, pela tipagem com base no sequenciamento. [J]. antígenos de tecido, 2002, 59: 47-48.

[6]. A. Dormoy, N. Froelich. Leisenbach, et al. Amplificação mono-alélica de Éxons 2-4 usando grupos específicos de alelos iniciadores para a tipagem com base em sequenciamento (SBT) dos genes HLA-A, -B e -C: Preparação e validação de mini-kits pré-SBT prontos para o uso. [J]. Antígenos de tecido, 2003, 62: 201-216.

[7]. Elaine R. Mardis. O impacto da nova geração de tecnologias de sequenciamento em genética. [J].Tendências em Genética.2008, 24:133-141.

[8]. Christian Hoffmann¹, Nana Minkah¹, Jeremy Leipzig. Código de barras de DNA pirosequenciamento para identificar uma droga rara de mutação resistente de HIV. [J]. Pesquisa de ácidos nucleicos, 2007, 1-8.

[9]. Shannon J.Odelberg, Robert B.Weiss, Akira Hata. Comutação de modelo durante a síntese de DNA por polymerase de DNA *Thermus aquaticus* I. [J]. Pesquisa de ácidos nucleicos.1995, 23:2049-2057.

[10]. Sayer D, Whidborne R, Brestovac B. HLA-DRB1 tipagem com base no sequenciamento de DNA: uma aproximação compatível para uma tipagem de alto rendimento incluindo doadores registrados de medula óssea não. [J]. Antígenos de tecido. 2001, 57(1):46-54.

[11]. Iwanka Kozarewa, Zemin Ning, Michael A Quail. Preparação de biblioteca de sequenciamento Illumina de amplificação livre facilita o mapeamento aperfeiçoado e a montagem de genomas tendenciosos de (G+C). [J]. Métodos naturais. 2009, 6: 291-295.

[12] Marsh, S.G.E., Parham, P. & Barber, L.D. O livro dos fatos HLA 3-91

(Academic Press, London, 2000).

[13]Campbell, K.J. et al. Caracterização de 47 MHC classes I sequências em macacos cynomolgus filipinos. *Imunogenética* 61, 177-187 (2009).

[14]Goulder, P.J.R. & Watkins, D.I. Impacto da classe MHC I diversidade no controle immune de replicação de vírus de imunodeficiência. *Nat. Rev. Immunol.* 8, 619-630 (2008).

[15]O'Leary, C.E. et al. Identificação de nova classe MHC I sequencias em macacos com rado de porco por pirosequenciamento amplificado de sequenciamento e clonagem de cDNA de comprimento completo. *Imunogenética* 61, 689-701 (2009).

[16]Robinson J, Malik A, Parham P, Bodmer JG, Marsh SGE. IMGT/HLA banco de dados- uma sequência de banco de dados para a maioria dos complexos histocompatibilidade de humanos. *Antígenos de tecido.* 55, 80-7 (2000).

[17] Hoffmann C, Minkah N, Leipzig J, Wang G, Arens MQ, Tebas P, Bushman FD. Código de barras e pirosequenciamento de DNA para identificar raras mutações resistentes a droga HIV. *Pesquisa de ácidos nucleicos.* 2007; 35(13):e91.

[18]. WU, D. L. et al. Análise comparativa de tipagem sorológica e tipagem por HLA-II por micro-PCR-SSP. *Di Yi Jun Yi Da Xue Xue Bao*, 2002, 22:247-249.

[19]. Al- Hussein K A, Rama N R, Butt A I, et al. HLA classe II sequência com base em tipagem em indivíduos Saudi normais. *Antígenos de tecido*, 2002, 60: 259-261.

[20]. D. C. Sayer, D. M. Goodridge. Estudo piloto: avaliação de variabilidade interlaboratorial de tipagem com base em qualidade de dados de sequenciamento de DNA. *Antígenos de tecido*, 2007, 69 Supl: 66-68.

[21]. Horton V, Stratton I, Bottazzo G. F. et al. Heterogeneidade genética de diabetes autoimune: idade de apresentação em adultos é influenciada por genótipos HLA DRB1 e DQB1. *Diabetologia*, 1999, 42:608-616.

[22]. C.E.M. Voorter, M.C. Kik1, E.M. van den Berg-Loonen et al. Alta resolução

de tipagem HLA para o gene DQB1 por tipagem com base em sequenciamento. Antígeno de tecido, 2008, 51:80-87.

[23]. G. Bentley, R. Higuchi, B. Hoglund et al. Alta resolução, alto rendimento de genotipagem HLA por sequenciamento de nova geração. Antígenos de tecido, 2009, 74: 393-403.

REIVINDICAÇÕES

1. Método para determinar a sequência nucleotídica de um ácido nucléico de interesse em uma amostra, o método **CARACTERIZADO** pelo fato de que compreende:

- 1) fornecer n amostras, em que n é um número inteiro, e $n \geq 1$;
- 2) dividir as n amostras em m grupos, em que m é um número inteiro, e $n \geq m \geq 1$;
- 3) realizar amplificação por PCR nas amostras sob condições adequadas para a amplificação do ácido nucléico de interesse quando modelos das amostras estão disponíveis, em que um par ou múltiplos pares de iniciadores de indexação são usados para cada amostra, em que cada par de iniciadores de indexação consiste em um iniciador de indexação dianteiro e um iniciador de indexação reverso e os iniciadores de indexação usados para amostras diferentes são diferentes;
- 4) associar produtos de amplificação por PCR de cada uma das amostras em conjunto;
- 5) submeter os produtos amplificados a um cisalhamento incompleto, purificar e recuperar, em que o dito cisalhamento incompleto produz produtos de PCR intactos não-cisalhados e produtos de PCR parcialmente cisalhados;
- 6) construir uma biblioteca de sequenciamento de PCR livre com base nos produtos de PCR intactos não-cisalhados e produtos de PCR parcialmente cisalhados, em que diferentes adaptadores de biblioteca são adicionados para distinguir diferentes bibliotecas de sequenciamento de PCR livre;
- 7) purificar e recuperar faixas de DNA entre o comprimento máximo de leitura e um comprimento de DNA máximo aplicável de um sequenciador;
- 8) submeter a mistura de DNA recuperado ao sequenciamento para obter as sequências do DNA cisalhado; e
- 9) corresponder os dados de sequenciamento obtidos com as amostras correspondentes com base em um único iniciador de indexação para cada amostra,

alinhar as sequências obtidas do DNA cisalhado contra as sequências de DNA de referência correspondentes aos produtos de PCR, e montar uma sequência completa do ácido nucleico de interesse a partir das sequências obtidas do DNA cisalhado com base na sobreposição da sequência e relação de ligação, em que um comprimento da sequência completa do ácido nucleico de interesse excede um comprimento máximo de leitura de um sequenciador usado para a referida sequenciação.

2. Método, de acordo com a reivindicação 1, **CARACTERIZADO** pelo fato de que:

cada par de iniciadores de indexação e um par de iniciadores de PCR formam um par de iniciadores de indexação, e os iniciadores de PCR dianteiros e reversos têm um iniciador de indexação dianteiro e um iniciador de indexação reverso na extremidade 5', respectivamente.

3. Método, de acordo com a reivindicação 1, **CARACTERIZADO** pelo fato de que:

a) os referidos iniciadores de PCR são iniciadores de PCR para amplificação de gene HLA, gene HLA-A/B, Éxons 2, 3 e 4 de HLA-A/B, ou Éxon 2 de HLA-DRB1;

b) os referidos iniciadores de PCR são iniciadores de PCR para amplificação de Éxons 2, 3 e 4 de HLA-A/B como mostrado na SEQ ID NO: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11 e 12;

c) os referidos iniciadores de PCR são iniciadores de PCR para amplificação de Éxons 2, 3 e 4 de HLA-A/B como mostrado na SEQ ID NO: 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23 e 24;

d) os referidos iniciadores de PCR são iniciadores de PCR para amplificação de Éxon 2 de HLA-DRB1 como mostrado na SEQ ID NO: 231, 232, 233, 234, 235, 236, 237 e 238;

e) os referidos iniciadores de PCR são iniciadores de PCR para amplificação de gene HLA, gene HLA-C, ou Éxons 2, 3 e/ou 4 de HLA-C;

f) os referidos iniciadores de PCR são iniciadores de PCR para amplificação de Éxons 2, 3 e/ou 4 de HLA-C como mostrado na SEQ ID NO: 25, 26, 27, 28, 29 e 30;

g) os referidos iniciadores de PCR são iniciadores de PCR para amplificação de Éxons 2, 3 e/ou 4 de HLA-C como mostrado na SEQ ID NO: 31, 32, 33, 34, 35 e 36;

h) os referidos iniciadores de PCR são iniciadores de PCR para amplificação de gene HLA, gene HLA-DQB1, ou Éxon 2 e/ou 3 de gene HLA-DQB1;

i) os referidos iniciadores de PCR são iniciadores de PCR para amplificação de gene HLA, gene HLA-DQB1 ou Éxon 2 e/ou 3 de gene HLA-DQB1 como mostrado na SEQ ID NO: 37, 38, 39 e 40.

4. Método, de acordo com a reivindicação 1, **CARACTERIZADO** pelo fato de que os referidos iniciadores de indexação são designados para:

a) iniciadores de PCR para amplificação de um gene específico de HLA;

b) iniciadores de PCR para a amplificação de Éxons 2, 3 e 4 de HLA-A/B e Éxon 2 de HLA-DRB1;

c) iniciadores de PCR como mostrados na SEQ ID NO: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11 e 12;

d) iniciadores de PCR como mostrados na SEQ ID NO: 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23 e 24; ou

e) iniciadores de PCR como mostrados na in SEQ ID NO: 231, 232, 233, 234, 235, 236, 237 e 238.

5. Método, de acordo com a reivindicação 1, **CARACTERIZADO** pelo fato de que os referidos iniciadores de indexação compreendem pelo menos 10 pares dos seguintes 95 pares de iniciadores de indexação:

Iniciador de
indexação No.
PI-1

Iniciador de indexação dianteiro
SEQ ID NO:41

Iniciador de indexação
reverso
SEQ ID NO: 42

PI-2	SEQ ID NO:43	SEQ ID NO: 44
PI-3	SEQ ID NO:45	SEQ ID NO: 46
PI-4	SEQ ID NO: 47	SEQ ID NO: 48
PI-5	SEQ ID NO: 49	SEQ ID NO: 50
PI-6	SEQ ID NO: 51	SEQ ID NO: 52
PI-7	SEQ ID NO: 53	SEQ ID NO: 54
PI-8	SEQ ID NO: 55	SEQ ID NO: 56
PI-9	SEQ ID NO: 57	SEQ ID NO: 58
PI-10	SEQ ID NO: 59	SEQ ID NO: 60
PI-11	SEQ ID NO: 61	SEQ ID NO: 62
PI-12	SEQ ID NO: 63	SEQ ID NO: 64
PI-13	SEQ ID NO: 65	SEQ ID NO: 66
PI-14	SEQ ID NO: 67	SEQ ID NO: 68
PI-15	SEQ ID NO: 69	SEQ ID NO: 70
PI-16	SEQ ID NO: 71	SEQ ID NO: 72
PI-17	SEQ ID NO: 73	SEQ ID NO: 74
PI-18	SEQ ID NO: 75	SEQ ID NO: 76
PI-19	SEQ ID NO: 77	SEQ ID NO: 78
PI-20	SEQ ID NO: 79	SEQ ID NO: 80
PI-21	SEQ ID NO: 81	SEQ ID NO: 82
PI-22	SEQ ID NO: 83	SEQ ID NO: 84
PI-23	SEQ ID NO: 85	SEQ ID NO: 86
PI-24	SEQ ID NO: 87	SEQ ID NO: 88
PI-25	SEQ ID NO: 89	SEQ ID NO: 90
PI-26	SEQ ID NO: 91	SEQ ID NO: 92
PI-27	SEQ ID NO: 93	SEQ ID NO: 94
PI-28	SEQ ID NO: 95	SEQ ID NO: 96
PI-29	SEQ ID NO: 97	SEQ ID NO: 98
PI-30	SEQ ID NO: 99	SEQ ID NO: 100
PI-31	SEQ ID NO: 101	SEQ ID NO: 102
PI-32	SEQ ID NO: 103	SEQ ID NO: 104
PI-33	SEQ ID NO: 105	SEQ ID NO: 106
PI-34	SEQ ID NO: 107	SEQ ID NO: 108
PI-35	SEQ ID NO: 109	SEQ ID NO: 110
PI-36	SEQ ID NO: 111	SEQ ID NO: 112
PI-37	SEQ ID NO: 113	SEQ ID NO: 114
PI-38	SEQ ID NO: 115	SEQ ID NO: 116
PI-39	SEQ ID NO: 117	SEQ ID NO: 118
PI-40	SEQ ID NO: 119	SEQ ID NO: 120
PI-41	SEQ ID NO: 121	SEQ ID NO: 122
PI-42	SEQ ID NO: 123	SEQ ID NO: 124
PI-43	SEQ ID NO: 125	SEQ ID NO: 126
PI-44	SEQ ID NO: 127	SEQ ID NO: 128
PI-45	SEQ ID NO: 129	SEQ ID NO: 130
PI-46	SEQ ID NO: 131	SEQ ID NO: 132
PI-47	SEQ ID NO: 133	SEQ ID NO: 134
PI-48	SEQ ID NO: 135	SEQ ID NO: 136
PI-49	SEQ ID NO: 137	SEQ ID NO: 138

PI-50	SEQ ID NO: 139	SEQ ID NO: 140
PI-51	SEQ ID NO: 141	SEQ ID NO: 142
PI-52	SEQ ID NO: 143	SEQ ID NO: 144
PI-53	SEQ ID NO: 145	SEQ ID NO: 146
PI-54	SEQ ID NO: 147	SEQ ID NO: 148
PI-55	SEQ ID NO: 149	SEQ ID NO: 150
PI-56	SEQ ID NO: 151	SEQ ID NO: 152
PI-57	SEQ ID NO: 153	SEQ ID NO: 154
PI-58	SEQ ID NO: 155	SEQ ID NO: 156
PI-59	SEQ ID NO: 157	SEQ ID NO: 158
PI-60	SEQ ID NO: 159	SEQ ID NO: 160
PI-61	SEQ ID NO: 161	SEQ ID NO: 162
PI-62	SEQ ID NO: 163	SEQ ID NO: 164
PI-63	SEQ ID NO: 165	SEQ ID NO: 166
PI-64	SEQ ID NO: 167	SEQ ID NO: 168
PI-65	SEQ ID NO: 169	SEQ ID NO: 170
PI-66	SEQ ID NO: 171	SEQ ID NO: 172
PI-67	SEQ ID NO: 173	SEQ ID NO: 174
PI-68	SEQ ID NO: 175	SEQ ID NO: 176
PI-69	SEQ ID NO: 177	SEQ ID NO: 178
PI-70	SEQ ID NO: 179	SEQ ID NO: 180
PI-71	SEQ ID NO: 181	SEQ ID NO: 182
PI-72	SEQ ID NO: 183	SEQ ID NO: 184
PI-73	SEQ ID NO: 185	SEQ ID NO: 186
PI-74	SEQ ID NO: 187	SEQ ID NO: 188
PI-75	SEQ ID NO: 189	SEQ ID NO: 190
PI-76	SEQ ID NO: 191	SEQ ID NO: 192
PI-77	SEQ ID NO: 193	SEQ ID NO: 194
PI-78	SEQ ID NO: 195	SEQ ID NO: 196
PI-79	SEQ ID NO: 197	SEQ ID NO: 198
PI-80	SEQ ID NO: 199	SEQ ID NO: 200
PI-81	SEQ ID NO: 201	SEQ ID NO: 202
PI-82	SEQ ID NO: 203	SEQ ID NO: 204
PI-83	SEQ ID NO: 205	SEQ ID NO: 206
PI-84	SEQ ID NO: 207	SEQ ID NO: 208
PI-85	SEQ ID NO: 209	SEQ ID NO: 210
PI-86	SEQ ID NO: 211	SEQ ID NO: 212
PI-87	SEQ ID NO: 213	SEQ ID NO: 214
PI-88	SEQ ID NO: 215	SEQ ID NO: 216
PI-89	SEQ ID NO: 217	SEQ ID NO: 218
PI-90	SEQ ID NO: 219	SEQ ID NO: 220
PI-91	SEQ ID NO: 221	SEQ ID NO: 222
PI-92	SEQ ID NO: 223	SEQ ID NO: 224
PI-93	SEQ ID NO: 225	SEQ ID NO: 226
PI-94	SEQ ID NO: 227	SEQ ID NO: 228
PI-95	SEQ ID NO: 229	SEQ ID NO: 230.

6. Método, de acordo com a reivindicação 1, **CARACTERIZADO** pelo fato de

que o referido cisalhamento de DNA é realizado por cisalhamento químico e cisalhamento físico, em que o cisalhamento químico inclui digestão enzimática, e o cisalhamento físico inclui cisalhamento ultrassônico ou cisalhamento mecânico.

7. Método, de acordo com a reivindicação 1, **CARACTERIZADO** pelo fato de que os referidos métodos de purificação e recuperação incluem recuperação por eletroforese e corte em gel, e recuperação por esferas magnéticas.

8. Método para determinar o genótipo HLA de uma amostra, o método **CARACTERIZADO** pelo fato de que compreende:

sequenciar uma amostra de um paciente pelo método conforme definido em qualquer uma das reivindicações 1 a 7;

alinhar os resultados do sequenciamento com os dados de sequência de Éxons de HLA, Éxons 2, 3, 4 de HLA-A/B, Éxons 2, 3 e/ou 4 de HLA-C, Éxon 2 e/ou 3 de gene HLA-DQB1 e/ou Éxon 2 de HLA-DRB1 em um banco de dados de HLA; e

determinar o genótipo HLA da amostra se o resultado do alinhamento da sequência mostrar 100% de correspondência.

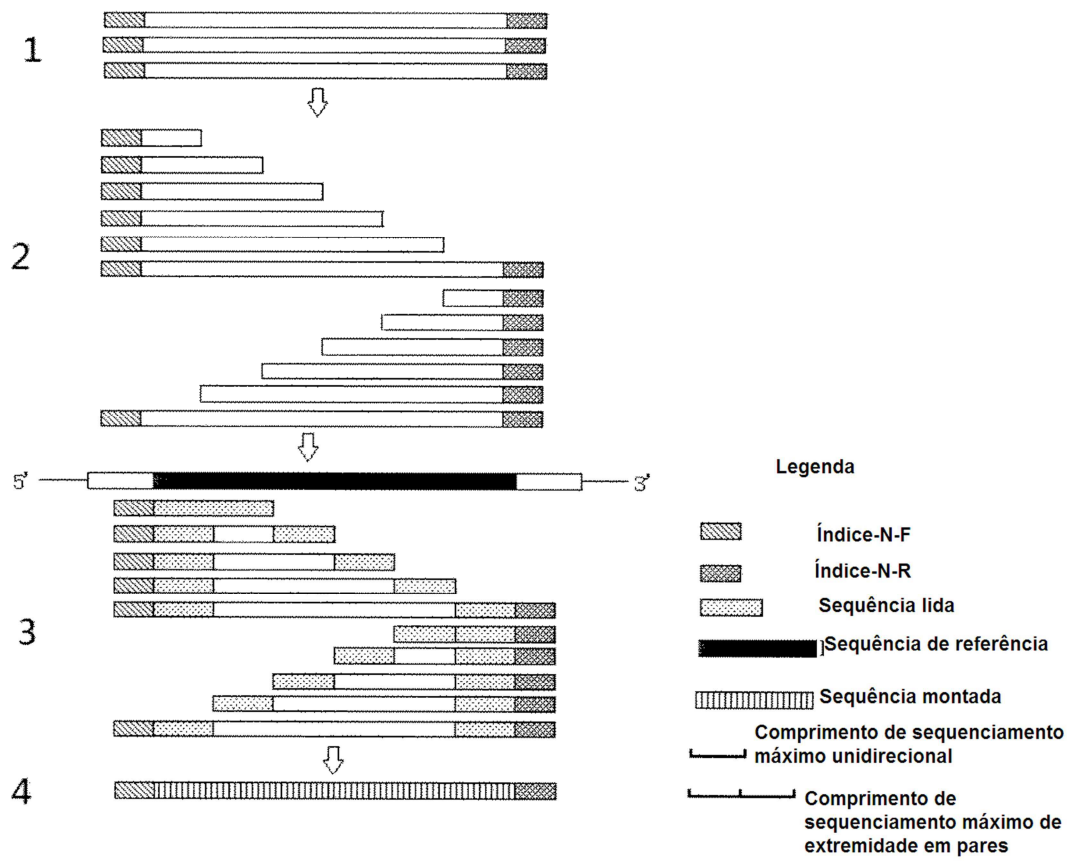


Fig. 1

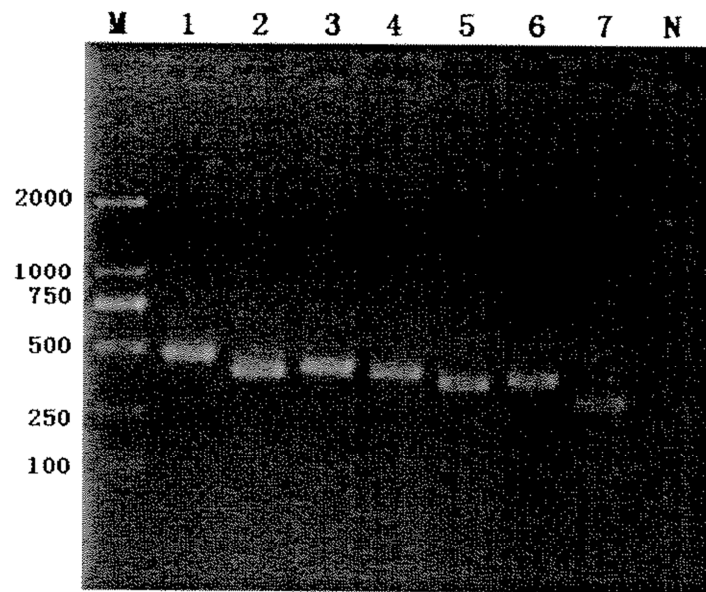


Fig. 2

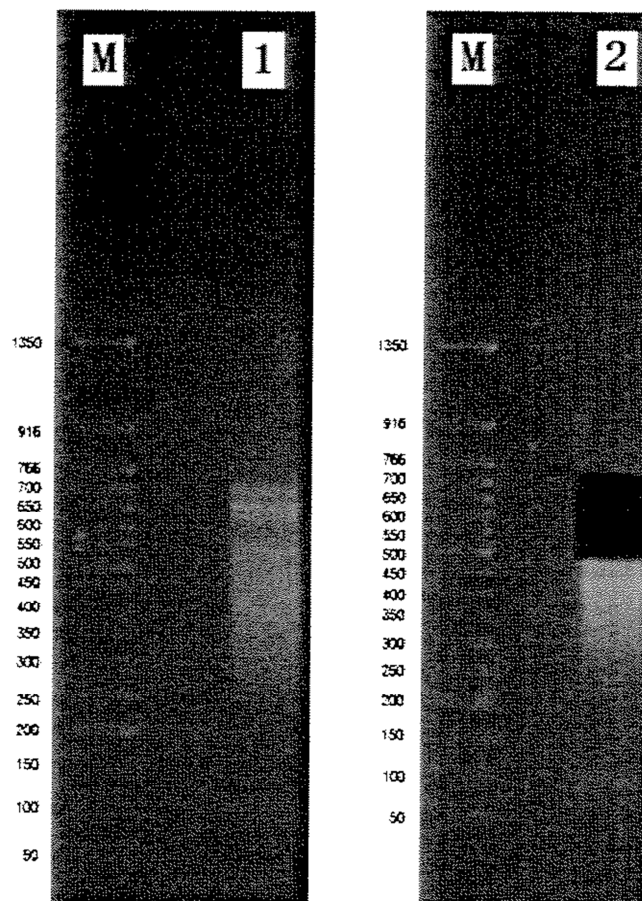


Fig. 3

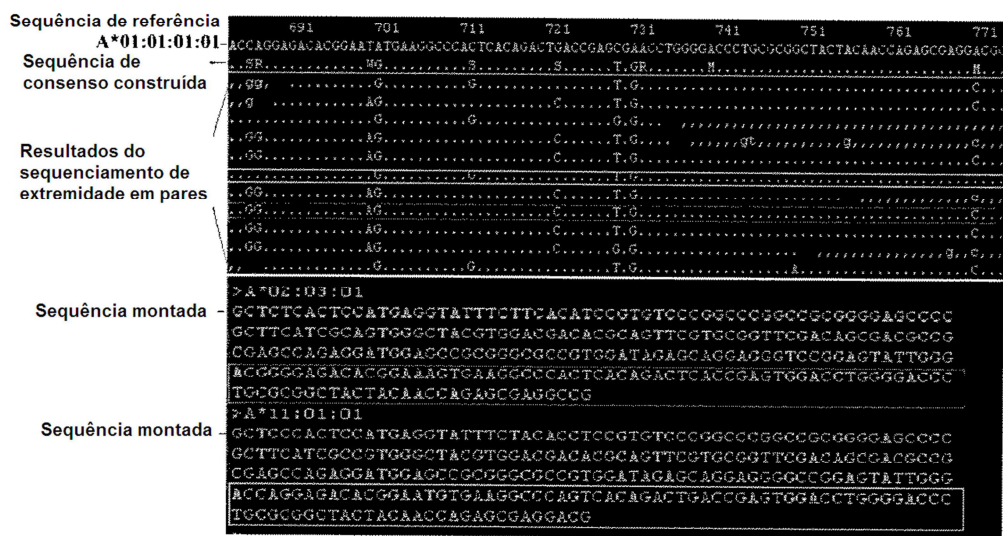


Fig. 4

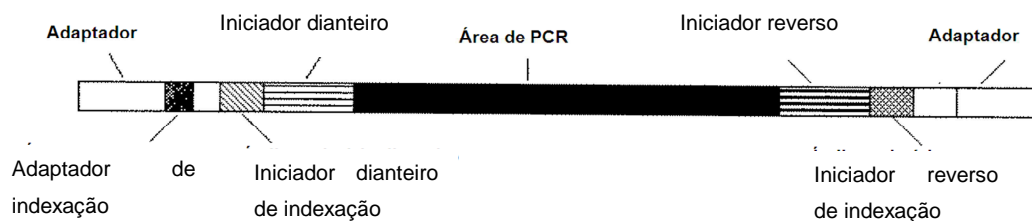


Fig. 5

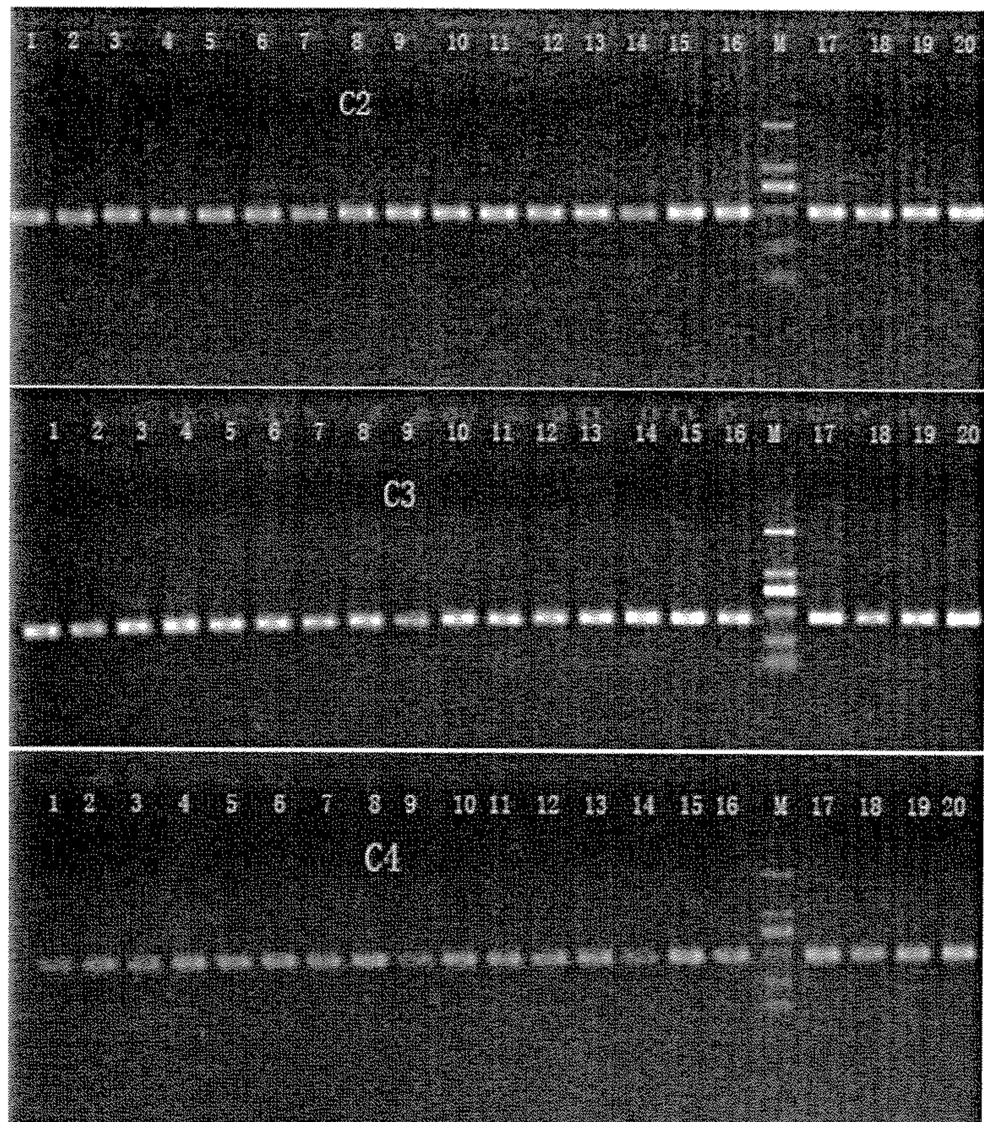


Fig. 6

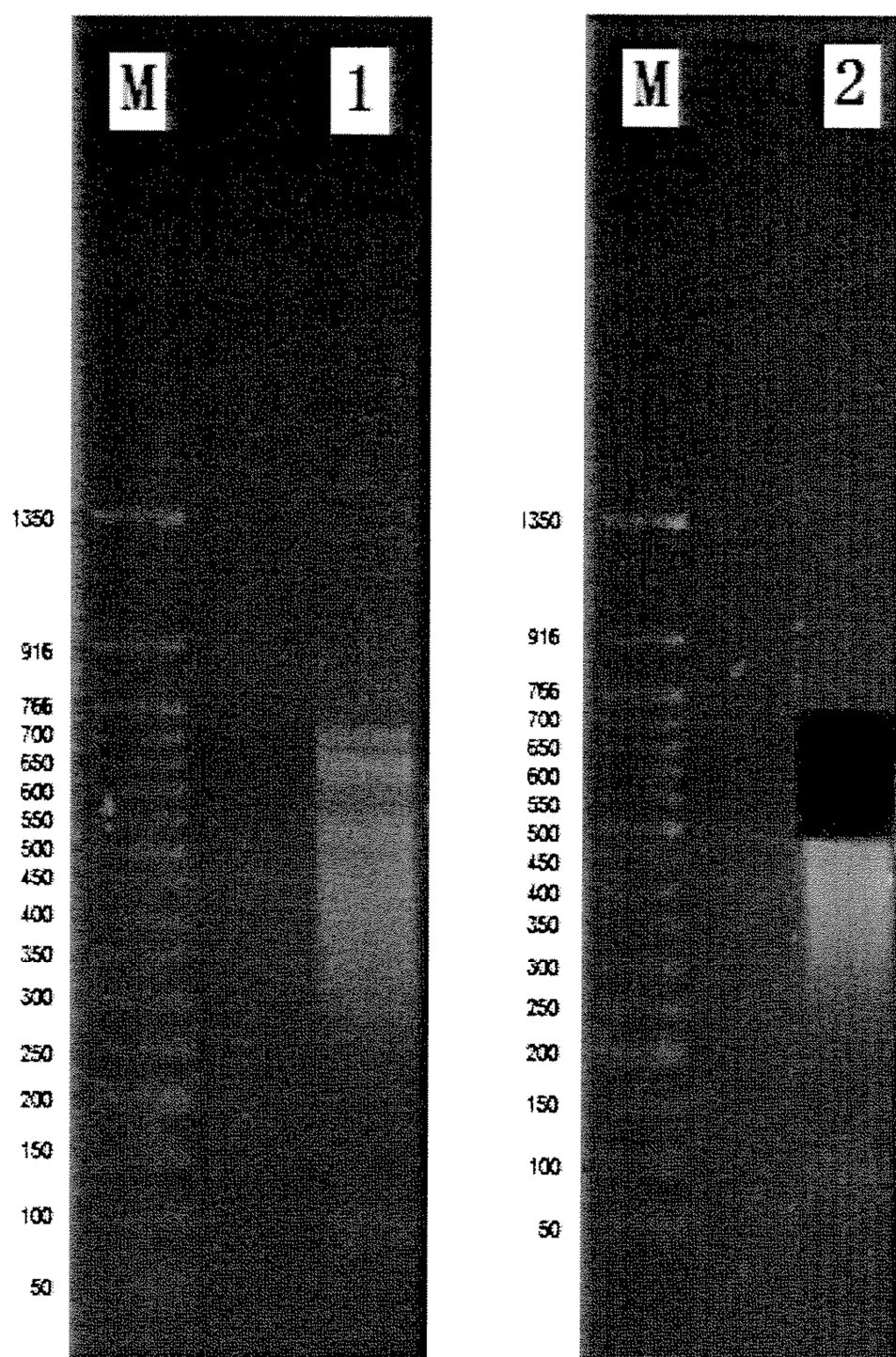


Fig. 7

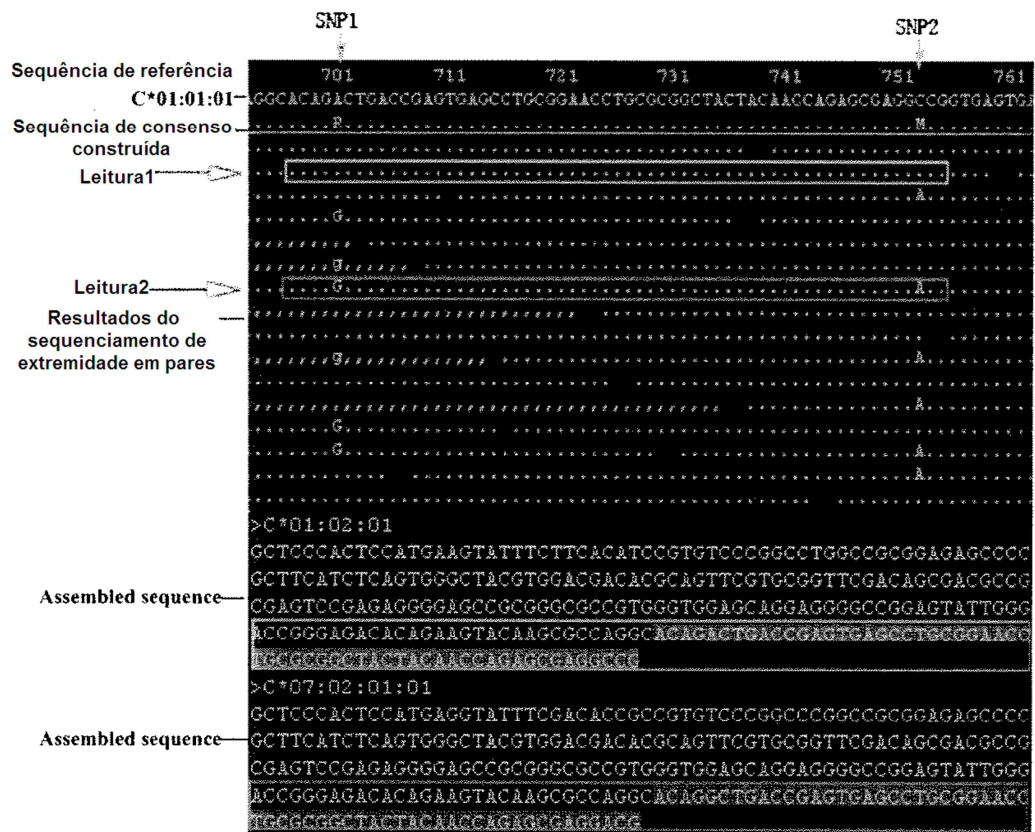


Fig. 8

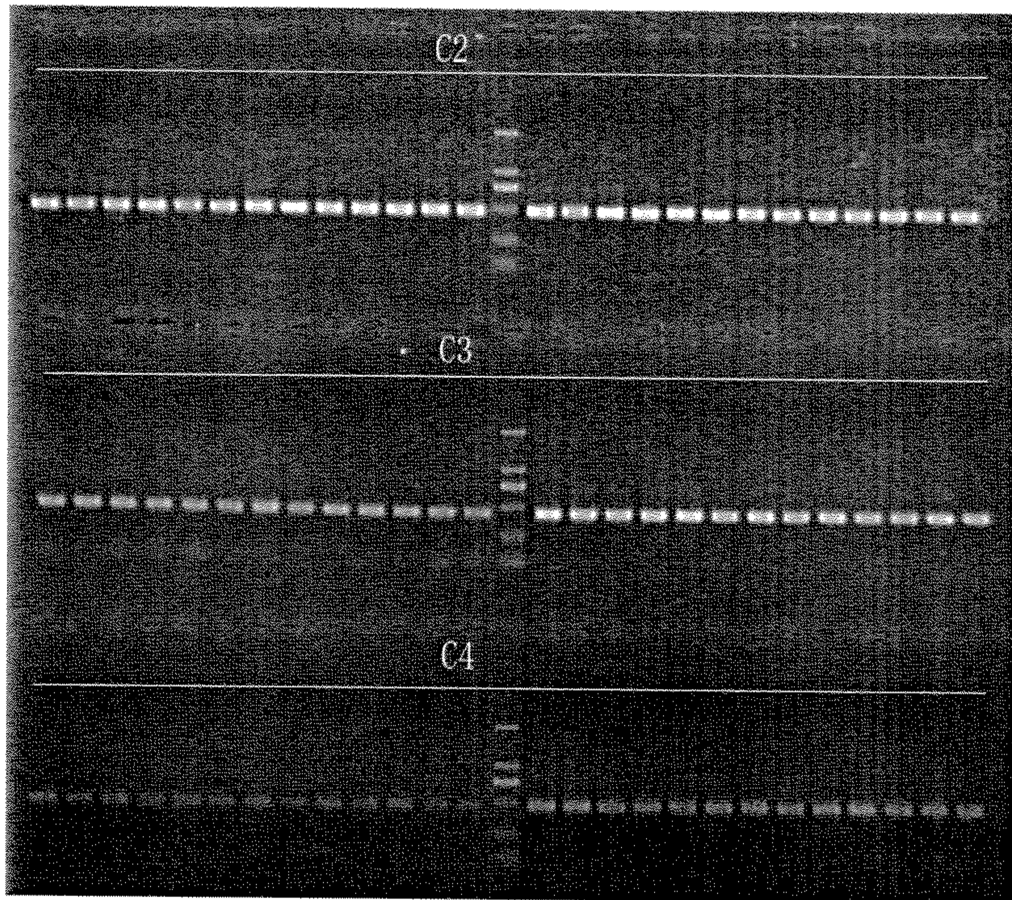


Fig. 9

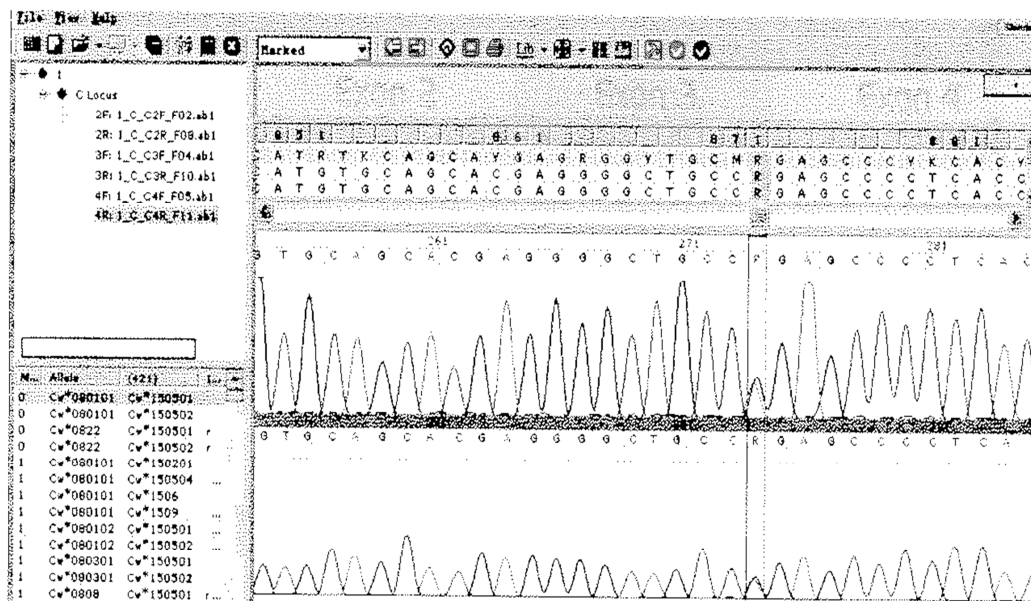


Fig. 10

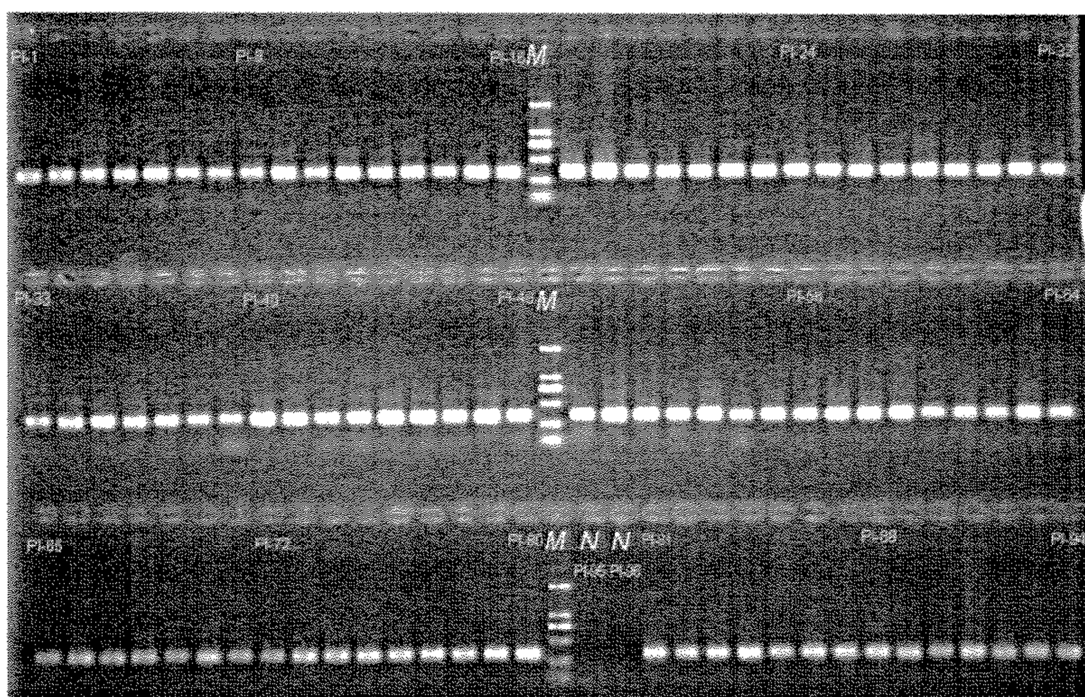


Fig. 11

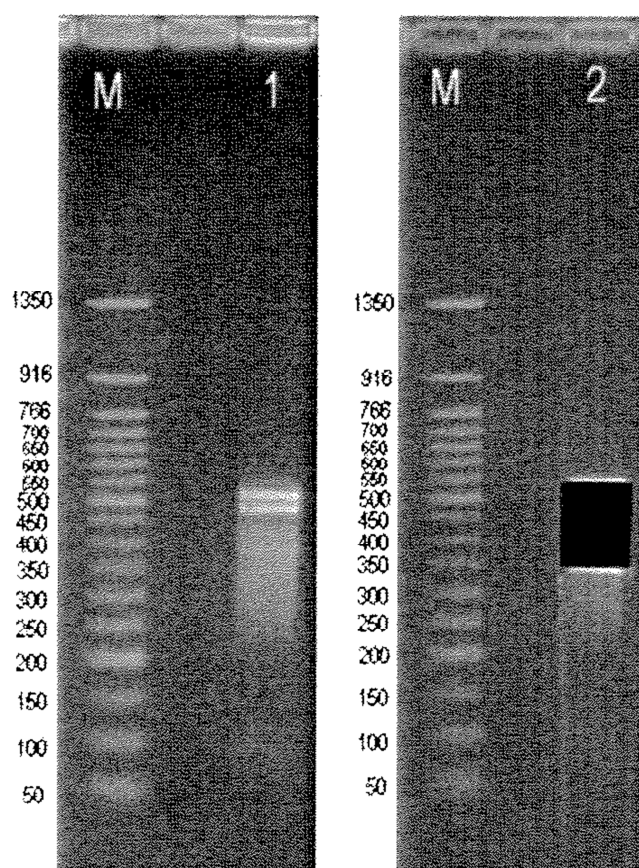


Fig. 12



Fig. 13

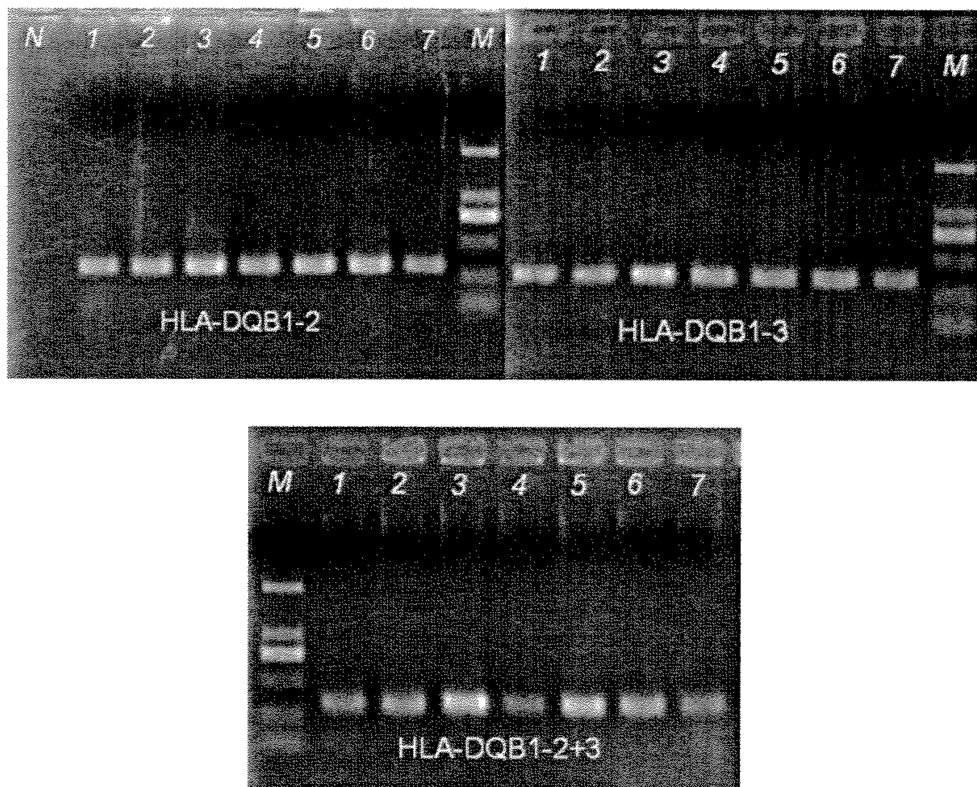


Fig. 14

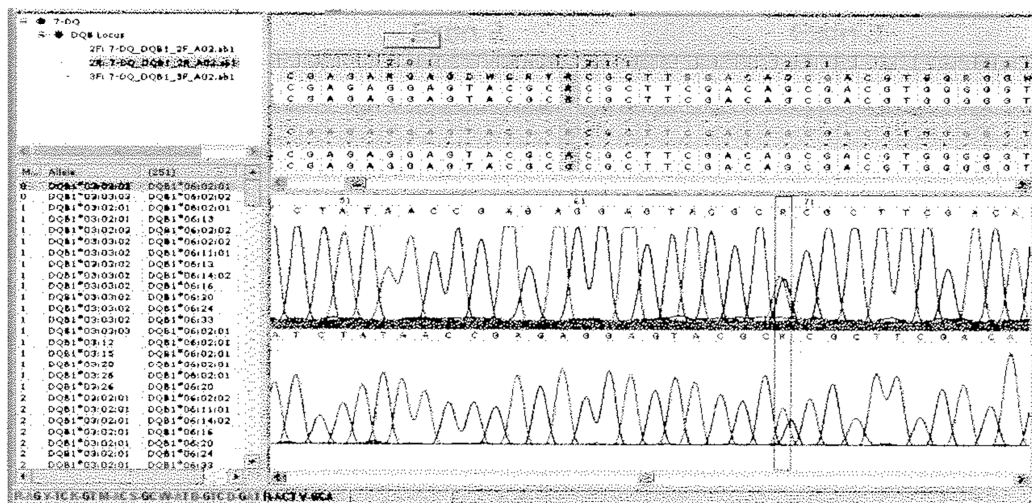


Fig. 15

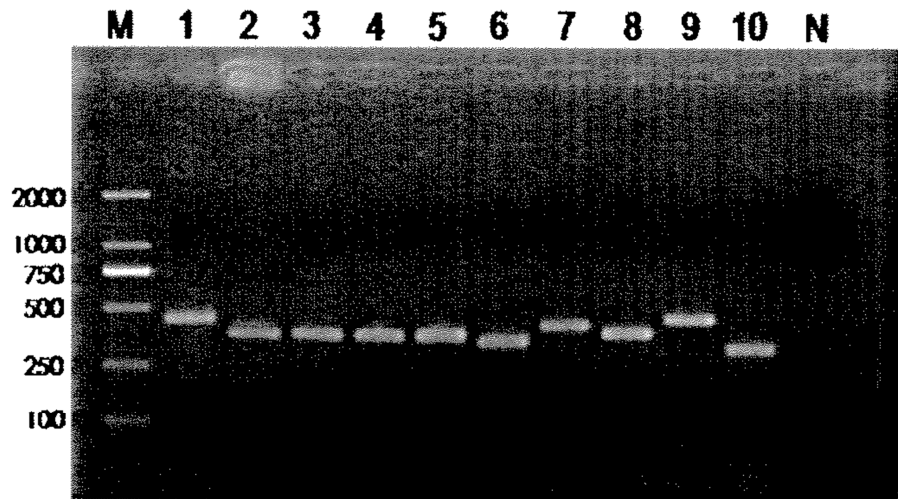


Fig. 16

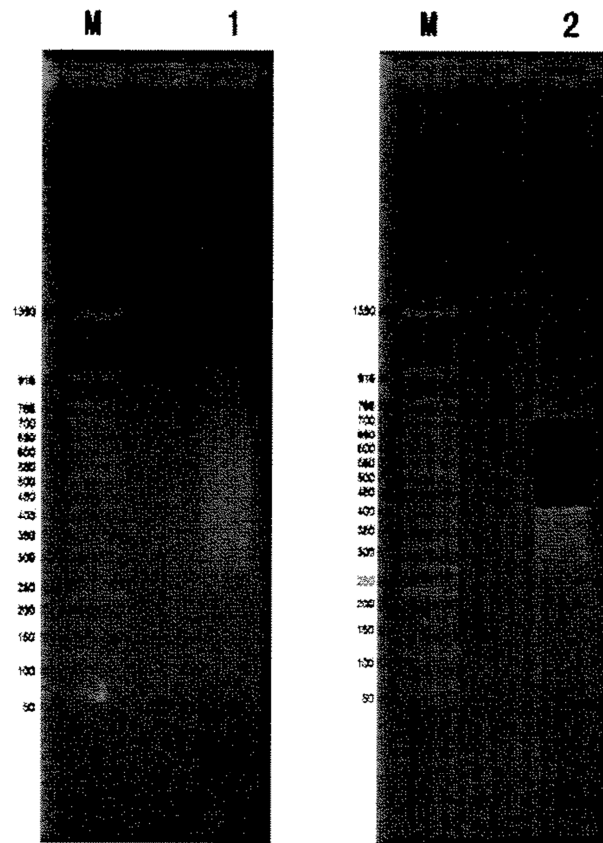


Fig. 17

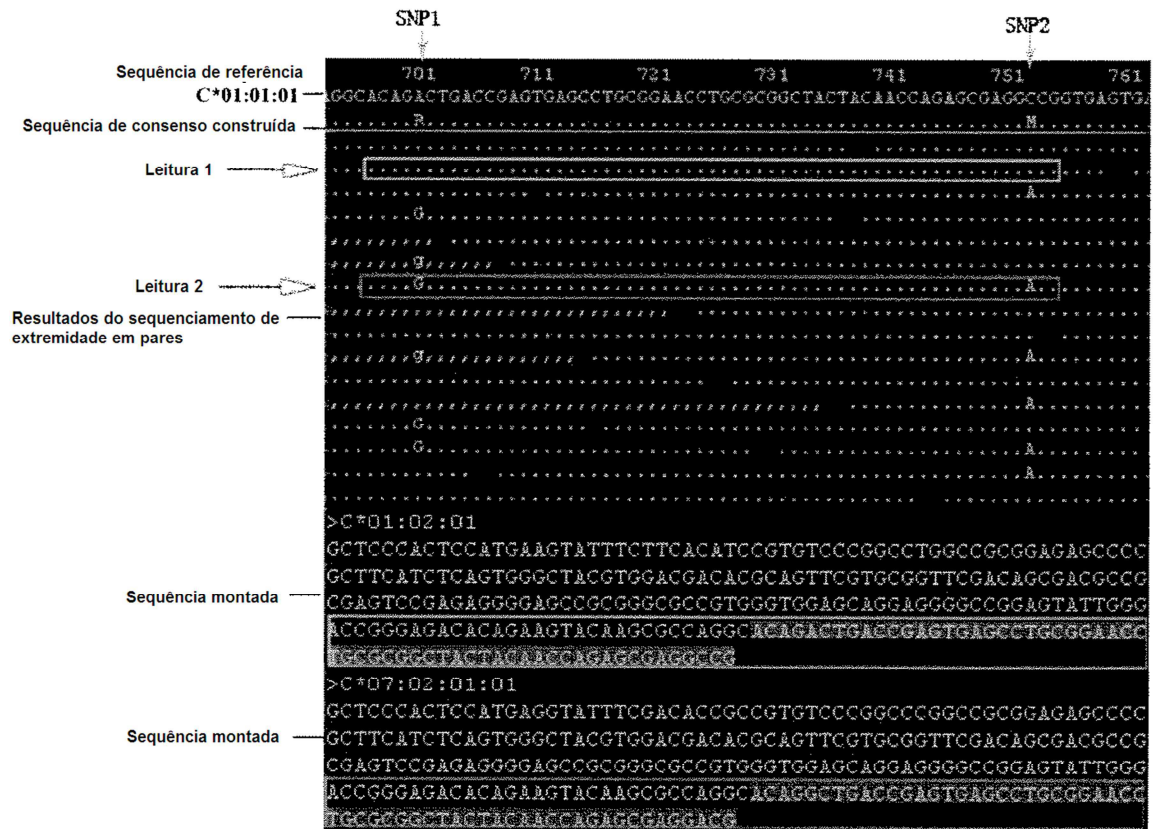


Fig. 18