



ФЕДЕРАЛЬНАЯ СЛУЖБА
ПО ИНТЕЛЛЕКТУАЛЬНОЙ СОБСТВЕННОСТИ

(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ПАТЕНТУ

(21)(22) Заявка: 2013103795/10, 30.06.2011

(24) Дата начала отсчета срока действия патента:
30.06.2011

Приоритет(ы):

(30) Конвенционный приоритет:

30.06.2010 CN 201010213721.2;

30.06.2010 CN 201010213719.5;

30.06.2010 CN 201010213717.6;

24.12.2010 CN PCT/CN2010/002150;

24.12.2010 CN PCT/CN2010/002149

(43) Дата публикации заявки: 20.08.2014 Бюл. № 23

(45) Опубликовано: 20.06.2016 Бюл. № 17

(56) Список документов, цитированных в отчете о поиске: US2009170713 A1, 02.07.2009. WO2007140540 A2, 13.12.2007. WO2005042764 A2, 12.05.2005. US7300755 B1, 27.11.2007. WO9835059 A1, 13.08.1998. JEESUK YU ET AL., Analysis of children with type 1 diabetes in Korea: high prevalence of specific anti-islet autoantibodies, immunogenetic similarities to Western populations with "unique" haplotypes, and lack (см. прод.)

(85) Дата начала рассмотрения заявки РСТ на национальной фазе: 30.01.2013

(86) Заявка РСТ:
CN 2011/076688 (30.06.2011)(87) Публикация заявки РСТ:
WO 2012/000445 (05.01.2012)

Адрес для переписки:

129090, Москва, ул. Б. Спасская, 25, строение 3,
ООО "Юридическая фирма Городисский и
Партнеры"

(72) Автор(ы):

ЛИ Цзянь (CN),

ЧЭНЬ Шипин (CN),

ЧЖАН Сяндун (CN),

ЛЮ Ин (CN),

ЧЖАН Цайфэнь (CN),

ЛЮ Тао (CN),

ЧЖАО Мэйжу (CN)

(73) Патентообладатель(и):

БиДжиАй Дженомикс Ко., Лтд (CN)

(54) **НОВЫЙ СПОСОБ ПЦР-СЕКВЕНИРОВАНИЯ И ЕГО ПРИМЕНЕНИЕ В ГЕНОТИПИРОВАНИИ НЛА**

(57) Реферат:

Изобретение относится к биохимии. Описан способ ПЦР-секвенирования, включающий следующие стадии: получение образца;

амплификация; смешивание; фрагментация; секвенирование и сплайсинг. Настоящее изобретение также относится к праймерным

меткам, используемым в указанном способе, а также к способу генотипирования, в частности, в анализе HLA. Настоящее изобретение также относится к ПЦР-праймам для HLA-A/B, HLA-C и HLA-DQB1. Изобретение позволяет повысить

длину продукта ПЦР, который может быть секвенирован секвенатором, без повышения цены и/или редукции. 3 н. и 16 з.п. ф-лы, 18 ил., 13 табл., 12 пр.

(56) (продолжение):

of discrimination by aspartic acid at position 57 of DQB, Clinical Immunology, 2004, v. 113, p. 318 -325. US5580730 A, 03.12.1996. US5683872 A, 04.11.1997. WO 2009049889 A1, 23.04.2009. STEN LINNARSSON, Recent advances in DNA sequencing methods-general principles of sample preparation, EXPERIMENTAL CELL RESEARCH, 2010, v. 316, no. 8, p. 1339-1343. RU 2393231 C1, 27.06.2010.

RU 2 5 8 7 6 0 6 C 2

RU 2 5 8 7 6 0 6 C 2



FEDERAL SERVICE
FOR INTELLECTUAL PROPERTY

(51) Int. Cl.

C12Q 1/68 (2006.01)*C12N* 15/11 (2006.01)(12) **ABSTRACT OF INVENTION**

(21)(22) Application: 2013103795/10, 30.06.2011

(24) Effective date for property rights:
30.06.2011

Priority:

(30) Convention priority:

30.06.2010 CN 201010213721.2;

30.06.2010 CN 201010213719.5;

30.06.2010 CN 201010213717.6;

24.12.2010 CN PCT/CN2010/002150;

24.12.2010 CN PCT/CN2010/002149

(43) Application published: 20.08.2014 Bull. № 23

(45) Date of publication: 20.06.2016 Bull. № 17

(85) Commencement of national phase: 30.01.2013

(86) PCT application:

CN 2011/076688 (30.06.2011)

(87) PCT publication:

WO 2012/000445 (05.01.2012)

Mail address:

129090, Moskva, ul. B. Spasskaja, 25, stroenie 3,
OOO "JUrIdicheskaja firma Gorodisskij i Partnery"

(72) Inventor(s):

LI TSzyan (CN),

CHEN SHipin (CN),

CHZHAN Syandun (CN),

LYU In (CN),

CHZHAN TSajfen (CN),

LYU Tao (CN),

CHZHAO Mejzhu (CN)

(73) Proprietor(s):

BiDzhiAj Dzhenomiks Ko., Ltd (CN)

(54) **NOVEL METHOD FOR PCR-SEQUENCING AND USE THEREOF FOR HLA GENOTYPING**

(57) Abstract:

FIELD: biochemistry.

SUBSTANCE: invention relates to biochemistry. Described is method of polymerase chain reaction sequencing, involving following steps: obtaining sample; amplification; mixing; fragmentation; sequencing and splicing. Present invention also relates to primer marks used in above method, as well as to

method of genotyping, in particular, in analysis of HLA. Present invention also relates to PCR primers for HLA-A/B, HLA-C and HLA-DQB1.

EFFECT: invention allows to increase length of the PCR product, which can be sequenced by sequencer, without increasing price and/or reduction.

19 cl, 18 dwg, 13 tbl, 12 ex

Родственные заявки

В настоящей заявке испрашивается право на приоритет заявок на патент Китая № 201010213717.6, 201010213719.5 и 201010213721.2, поданных 30 июня 2010 г., и право на приоритет Международных заявок № PCT/CN2010/002150 и PCT/CN2010/002149, поданных 24 декабря 2010 г., содержание которых во всей своей полноте вводится в настоящее описание посредством ссылки.

Область техники, к которой относится изобретения

Изобретение относится к области секвенирования нуклеиновых кислот, в частности к области ПЦР-секвенирования. Настоящее изобретение также относится к технике считывания штрихового кода ДНК и к стратегии неполного фрагментирования ДНК. Способ согласно изобретению может быть применен, в частности, в технологии секвенирования второго поколения, а особенно, в технологии парно-концевого секвенирования второго поколения, а также в генотипировании HLA. Более конкретно, настоящее изобретение относится к способу генотипирования HLA, а в частности, к способу генотипирования HLA-A, HLA-B, HLA-C и HLA-DQB1, а также к парам праймеров, используемым в ПЦР-амплификации, применяемой в этом способе.

Предшествующий уровень техники

Метод ПЦР-секвенирования означает технологию, в которой ДНК-фрагменты представляющего интерес гена получают методом ПЦР, а затем полученные ДНК-фрагменты представляющего интерес гена подвергают ДНК-секвенированию с получением информации о последовательности ДНК представляющего интерес гена. Методы ПЦР-секвенирования широко применяются в таких областях, как детектирование мутации генов и генотипирование в течение длительного периода времени.

Технология ДНК-секвенирования подразделяется, главным образом, на технологию ДНК-секвенирования первого поколения, осуществляемую методом секвенирования по Сэнгеру, и на технологию ДНК-секвенирования второго поколения, осуществляемую на секвенаторе Illumina GA, Roche 454, ABI Solid, и т.п. Метод ДНК-секвенирования по Сэнгеру отличается простотой экспериментальных операций с получением визуальных и точных результатов и непродолжительностью времени проведения эксперимента, а поэтому этот метод широко применяется в таких областях, как клиническое детектирование мутации генов и генотипирование, где необходима быстрая оборачиваемость процесса, особенно, если это касается результатов детектирования. Однако из-за низкой производительности и высокой стоимости этого метода, его применение в области крупномасштабного генотипирования имеет определенные ограничения.

По сравнению с технологией ДНК-секвенирования первого поколения, технология ДНК-секвенирования второго поколения отличается тем, что она имеет высокую производительность секвенирования, низкую стоимость, высокий уровень автоматизации и возможность одномолекулярного секвенирования. Если взять в качестве примера технологию одномолекулярного секвенирования Illumina GA, то путем секвенирования по одной молекуле можно получить данные для 50G (около 50 миллиардов) оснований, то есть в среднем 5 миллиардов оснований в день, при этом средняя стоимость секвенирования основания составляет менее чем 1/1000 от стоимости секвенирования по методу Сэнгера. Кроме того, анализ результатов может быть непосредственно осуществлен на компьютере. Таким образом, технология секвенирования ДНК на секвенаторе второго поколения представляет собой технологию, наиболее подходящую для крупномасштабного секвенирования. Однако в технологии ДНК-секвенирования

второго поколения, длина участка непрерывного секвенирования по всей его длине обычно невелика. В настоящее время максимальная длина двунаправленного секвенирования на секвенаторе Illumina GA составляет 200 п.н., и хотя в методе Roche 454 GS-FLX максимальная длина последовательности для секвенирования может составлять примерно до 500 п.н., однако затраты на такое секвенирование довольно велики, а производительность - низка. Если ПЦР-ампликон имеет длину, превышающую максимальную длину, которая может быть непрерывно секвенирована в данном секвенаторе, то секвенирование всего ампликона не может быть осуществлено путем прямого секвенирования, а поэтому не может быть получена информация о всей длине последовательности ДНК ампликона. Из-за короткой максимальной длины секвенирования применение технологии секвенирования второго поколения в методе ПЦР-секвенирования имеет определенные ограничения. Помимо поэтапного усовершенствования технологии секвенирования в целях увеличения максимальной длины секвенирования, крайне необходимо разработать новый способ, который позволял бы решить существующие в настоящее время проблемы, связанные с максимальной длиной ДНК-секвенирования на секвенаторе второго поколения, и который можно было бы применять в технологии ПЦР-секвенирования.

Главный комплекс гистосовместимости человека (HLA) представляет собой одну из генных систем, которые, как известно, имеют самый высокий полиморфизм. HLA представляет собой главную генную систему для модуляции специфического иммунного ответа в организме человека и определения различий в восприимчивости индивидуумов к развитию заболеваний и непосредственно ассоциируется с отторжением аллогенных органов-трансплантатов. Проведенные исследования показали, что чем выше степень соответствия генов, таких как HLA-A, B, C, DRB1 и DQB1, а также степень их разрешения при генотипировании у донора и реципиента, тем выше выживаемость трансплантата. В настоящее время перед трансплантацией гемопоэтических стволовых клеток уже проводится стандартная процедура тестирования, заключающаяся в высокопроизводительном генотипировании HLA у потенциальных доноров и реципиентов.

Современная международная стандартная технология высокоразрешающего HLA-генотипирования представляет собой метод ПЦР-секвенирования, проводимый на основе технологии секвенирования по Сэнгеру, которая включает ПЦР-амплификацию соответствующих областей гена HLA, секвенирование амплифицированного продукта, генотипирование результатов секвенирования с помощью профессиональной компьютерной программы по генотипированию и, наконец, получение информации о генотипе HLA-образца. Этот метод отличается тем, что он дает визуальные результаты, имеет высокое разрешение и позволяет детектировать новые аллели. Однако из-за высокой стоимости и низкой производительности метода секвенирования по Сэнгеру, его применение в научно-исследовательских институтах, имеющих регистрационную базу данных о добровольцах-донорах гемопоэтических стволовых клеток (банк клеток костного мозга), где требуется проведение крупномасштабного детектирования путем HLA-генотипирования, имеет определенные ограничения.

Сообщалось, что для HLA-генотипирования было применено ПЦР-секвенирование методом Roche 454 GS-FLX. Однако поскольку этот метод является довольно дорогостоящим, то он не имеет явных преимуществ в отношении производительности и стоимости секвенирования по сравнению с методом HLA-генотипирования, проводимым путем секвенирования по Сэнгеру. По сравнению с методом Roche 454 GS-FLX, метод Illumina GA оперирует с более короткой максимальной длиной

непрерывного секвенирования, но имеет явные преимущества, такие как высокая производительность и низкая стоимость секвенирования. Если будет решена проблема, связанная с небольшой максимальной длиной секвенирования в методе Illumina GA, то применение этого метода для HLA-генотипирования позволит компенсировать

недостатки современных методов HLA-генотипирования.

Описание сущности изобретения

При одновременном проведении анализов путем секвенирования последовательностей со специфическим геном в большом числе образцов с применением технологии секвенирования второго поколения обычно применяется стратегия ПЦР-секвенирования, в которой непосредственно используется комбинация индексирования праймера и технологии секвенирования второго поколения. Если максимально возможная длина непрерывного секвенирования в секвенаторе может покрывать длину всего ПЦР-продукта, то вышеуказанная стратегия удовлетворяет предъявляемым требованиям. Если максимальная длина секвенирования в секвенаторе не может покрывать длину всего ПЦР-продукта, то секвенатор Illumina GA необходимо заменить секвенатором второго поколения, имеющим большую максимальную длину секвенирования (таким как Roche 454 GS-FLX). Если максимальная длина секвенирования последовательности все еще не удовлетворяет этим требованиям, то может быть использован секвенатор первого поколения со скарификацией стоимости и производительности.

Фактически, ситуация заключается в том, что секвенатор Illumina GA отличается исключительно высокой производительностью, но его максимальная длина секвенирования составляет лишь 200 п.н.; и хотя максимальная длина секвенирования на секвенаторе Roche 454 GS-FLX может достигать примерно 500 п.н., но стоимость такого секвенирования остается относительно высокой, а производительность - относительно низкой; и кроме того, хотя максимальная длина секвенирования на секвенаторе первого поколения может достигать более чем 1000 п.н., однако его производительность и стоимость не сравнима с производительностью и стоимостью секвенатора второго поколения.

Авторам пока не известен метод, позволяющий увеличивать длину ПЦР-продуктов, которые могут быть секвенированы по всей их длине без скарификации стоимости и производительности. Комбинация индексирования праймеров, стратегии неполного фрагментирования ДНК и применения технологии секвенирования второго поколения, описанных в настоящей заявке, позволяет оперировать с ПЦР-продуктами, которые имеют длину, превышающую максимальную длину непрерывного секвенирования на секвенаторе, и при этом позволяет полностью использовать преимущества технологии секвенирования второго поколения, такие как высокая производительность и низкая стоимость, что значительно расширяет область применения такой технологии. Среди прочих технологий секвенирования второго поколения, технология секвенирования второго поколения, применяемая в настоящем изобретении, включает секвенирование с обоих концов и ПЦР-секвенирование, в котором для ПЦР-матрицы используется эталонная последовательность ДНК.

Настоящее изобретение относится к способам ПЦР-секвенирования, которые позволяют устранить ограничения, связанные с короткой максимальной длиной секвенирования на секвенаторе и расширить область применения технологии второго поколения в ПЦР-секвенировании ДНК. Так, например, при секвенировании с применением технологии второго поколения используют индексированные праймеры, имеющие индекс, добавленный к его 5'-концу, и амплифицированные ПЦР-продукты фрагментируют, после чего фрагментированные продукты репарируют по концам, и

к их 3'-концам присоединяют дезоксиаденозин (А), а затем лигируют с различными адаптерами, полученными без проведения ПЦР.

Метод ПЦР-секвенирования, проводимый на основе штрихового кодирования ДНК и стратегии неполного фрагментирования ДНК, позволяет значительно увеличить
 5 число специфически меченых образцов без увеличения числа индексов праймера (фигура 5). В настоящем изобретении фактически секвенируемая длина ПЦР-продуктов превышает максимальную длину секвенирования на секвенаторе из-за добавления индексов праймера к прямому и обратному ПЦР-праймерам, и секвенирование осуществляют в комбинации со стратегией неполного фрагментирования ДНК и с
 10 технологией секвенирования второго поколения.

Добавление последовательности-индекса к переднему концу праймера для амплификации осуществляют в целях реализации одновременного секвенирования множества образцов. Более конкретно, уникальный индекс праймера добавляют к
 15 каждому образцу в процессе ПЦР с применением технологии ПЦР-индекса/штрих-кода в комбинации с синтезом индексированного праймера путем добавления индекса к 5'-концу ПЦР-праймера. Так, например, во время секвенирования, проводимого с применением технологии секвенирования второго поколения, образцы последовательно процессируют только в стадии ПЦР, а в остальных экспериментальных стадиях эти образцы могут быть смешаны и процессированы одновременно, и конечный результат
 20 для каждого образца может быть идентифицирован по индексу его уникального праймера.

Метод индексирования «адаптера» или «адаптера библиотеки» означает метод индексирования библиотек, включающий добавление к множеству библиотек для секвенирования различных адаптеров библиотек (различных адаптеров библиотек,
 25 состоящих из различных последовательностей и различных частей последовательностей, называемых адаптерными индексами), конструирование индексированных библиотек для секвенирования, и последующее секвенирование множества различных индексированных библиотек для секвенирования в пуле, где конечный результат секвенирования каждой индексированной библиотеки для секвенирования отличается
 30 друг от друга. Термин «адаптер PCR-FREE-библиотеки» означает сегмент, сконструированный из оснований, главная роль которых заключается в дополнительной фиксации молекулы ДНК на чипе для секвенирования и в обеспечении сайтов связывания для универсальных секвенирующих праймеров, где адаптер PCR-FREE-библиотеки может быть непосредственно лигирован с двумя концами ДНК-фрагментов в библиотеке
 35 для секвенирования. Поскольку при введении адаптера не проводят ПЦР, то такой адаптер называется здесь «адаптер PCR-FREE-библиотеки». Так, например, адаптеры PCR-FREE-библиотеки, используемые в примерах настоящего изобретения, были получены от ILLUMINA.

Метод конструирования PCR-FREE-библиотеки, где используется технология
 40 индексирования адаптера библиотеки, заключается в непосредственном лигировании адаптера библиотеки к двум концам ДНК-фрагмента в библиотеке для секвенирования. Поскольку при введении адаптера библиотеки не проводят ПЦР, то такой адаптер называется здесь «адаптер PCR-FREE-библиотеки». Для лигирования в процессе введения может быть использована ДНК-лигаза. Поскольку в процессе конструирования
 45 библиотеки не проводят ПЦР, то неточности конечных результатов ПЦР можно избежать во время конструирования библиотеки, содержащей ПЦР-продукты с большой степенью сходства последовательностей.

Методами амплификации ДНК, методами экстракции ДНК, методами очистки ДНК

и методами выравнивания последовательностей ДНК, применяемыми в настоящем изобретении, могут быть методы, известные специалистам. Указанные методы могут быть выбраны самим специалистом, исходя из практических соображений. Что касается методов секвенирования ДНК, то специалист в данной области может осуществить их в соответствии со стандартными процедурами или в соответствии с инструкцией, прилагаемой к секвенатору.

Конструирование индексов праймера может быть осуществлено в зависимости от применяемой концепции эксперимента. Что касается концепции секвенирования методом Illumina GA, то в настоящем изобретении при конструировании праймерных индексов были предприняты следующие действия: 1) последовательность моонуклеотидного повтора, содержащая 3 или более основания, была удалена в последовательностях праймерных индексов, 2) общее количество оснований А и оснований С в одном и том же сайте всех праймерных индексов составляет 30%-70% от количества всех оснований, 3) содержание GC в последовательности праймерных индексов составляет 40-60%, 4) праймерные индексы отличаются друг от друга по меньшей мере 4 основаниями, 5) последовательности, имеющие высокую степень сходства с праймерами для секвенирования методом Illumina GA, были удалены из последовательностей-индексов праймеров, и 6) случаи, в которых добавление последовательностей праймерных индексов к ПЦР-праймерам приводит к образованию крупных шпилек и димеров, были сведены к минимуму.

В настоящем изобретении два праймерных индекса (которые являются либо идентичными, либо различными) добавляли к двум концам ПЦР-продукта, соответственно, так, чтобы индекс праймера у любого конца ПЦР-продукта мог служить специфической меткой, идентифицирующей образец ПЦР-продукта. Полученный ПЦР-продукт подвергали неполному фрагментированию. Так называемое «неполное фрагментирование» означает процесс получения продуктов, содержащих интактные нефрагментированные ПЦР-продукты и частично фрагментированные ПЦР-продукты. Методами фрагментирования являются, но не ограничиваются ими, методы химического фрагментирования (такие как ферментативное расщепление) и методы физического фрагментирования. Методами физического фрагментирования являются методы ультразвукового фрагментирования или методы механического фрагментирования. Фрагментированную ДНК подвергают электрофорезу в 2% агарозе, и все ДНК-полосы, имеющие длину между максимальной длиной секвенирования и максимально приемлемой длиной ДНК для секвенирования на секвенаторе, очищают и выделяют путем разрезания геля (наибольшая длина ДНК, приемлемая для секвенатора Illumina GA составляет 700 п.н., и эта длина означает исходную длину ДНК, которая не включает длину последовательности адаптера библиотеки). Методами очистки и выделения являются, но не ограничиваются ими, выделение путем электрофореза и разрезания геля, и выделение на магнитных сферах. Из выделенных ДНК-фрагментов конструируют библиотеки для секвенирования в соответствии с процедурами конструирования библиотек для секвенирования на секвенаторе второго поколения, а затем продукты подвергают секвенированию. Библиотеки для секвенирования предпочтительно конструируют в соответствии с PCR-FREE-процедурами, подходящими для конструирования библиотек для секвенирования, а методом секвенирования является метод секвенирования с обоих концов. Конструирование PCR-FREE-библиотек для секвенирования осуществляют методами, известными специалистам. При получении данных секвенирования, информация о последовательностях всех тест-образцов может быть получена исходя из последовательностей праймерных индексов. Считанные

последовательности (последовательности-«риды») сопоставляют путем выравнивания с соответствующими эталонными ДНК-последовательностями ПЦР-продуктов с помощью ВМА, а затем осуществляют сборку полноразмерной последовательности исходя из перекрывания и характера сцепления между последовательностями-ридами (фигура 1). Используемый здесь термин «сцепление» означает присоединение по концам, определяемое параметрами парно-концевого секвенирования.

Секвенирование на секвенаторе Illumina GA (геномный аналитический секвенатор, выпускаемый фирмой Illumina Inc., для краткости называемый здесь Illumina GA), то есть анализ последовательности ДНК, осуществляют, как в принципе и всякое секвенирование, путем синтеза. Этот анализ может быть применен в фазе гаплотипирования, и полученные данные, в конечном счете, представляют собой серию последовательностей оснований, которые могут быть непосредственно подвергнуты выравниванию с эталонными последовательностями, имеющимися в базе данных HLA. Поскольку эта программа не давала каких-либо заметных ошибок, а именно, несоответствий пиков, обычно наблюдаемых у существующих традиционных программ типирования, то желательно, чтобы эта программа типирования была автоматизирована. Секвенатор Illumina GA имеет высокую производительность секвенирования. В настоящее время один единственный раунд секвенирования позволяет получить данные о 50G (50 миллиардов) оснований, то есть, в среднем, 5 миллиардов оснований в день. Благодаря такой высокой производительности, для каждой последовательности может быть достигнута высокая процедурная мощность секвенирования, что дает гарантию надежности результатов секвенирования.

До сих пор каких-либо исследований по применению секвенирования на Illumina GA для HLA-типирования не проводилось. В настоящем изобретении впервые применяется секвенирование на Illumina GA для HLA-типирования, и такой метод HLA-типирования является не дорогостоящим и высокопроизводительным, а также имеет высокую точность и высокое разрешение, если его проводят в комбинации с применением технологии ПЦР-секвенирования и технологии считывания штрих-кода ДНК, с неполным фрагментированием ДНК и созданием PCR-FREE-библиотеки.

В настоящем изобретении с помощью метода ПЦР-секвенирования, основанного на технологии считывании штрих-кода ДНК, неполного фрагментирования ДНК и получения PCR-FREE-библиотеки, анализируемые образцы подразделяют на группы, после чего образцы каждой группы подвергают амплификации фрагментов представляющих интерес генов HLA с использованием праймеров, меченных двунаправленными праймерными индексами (максимальная длина ПЦР-продуктов зависит от максимальной длины ДНК, которая может быть использована в секвенаторе; при этом в современном секвенаторе Illumina GA максимально приемлемая длина считывания ДНК составляет 700 п.н. и такую длину имеет исходная ДНК, которая не включает последовательности адаптера библиотеки), затем ПЦР-продукты объединяют вместе в том же количестве и подвергают неполному фрагментированию с получением индексированной PCR-FREE-библиотеки для секвенирования ДНК. Различные индексированные библиотеки для секвенирования, полученные из образцов различных групп, смешивают в равном молярном объеме, а затем все ДНК-фрагменты, длина которых превышает максимальную длину секвенирования в секвенаторе, селективно выделяют и подвергают секвенированию на секвенаторе Illumina GA.

Последовательности-риды ДНК для каждого образца могут быть получены путем скрининга информации о последовательностях индексов адаптеров, индексов праймеров и ПЦР-праймеров в общих данных по секвенированию. Полученные последовательности

ДНК, после их сборки, сопоставляют с соответствующими данными, имеющимися в профессиональной базе данных HLA IMGT, и наконец, определяют генотип HLA образца.

В вышеописанных методах, после фрагментирования указанной ДНК, ДНК образцов различных групп лигируют с другим адаптером библиотеки во время PCR-FREE-получения индексируемой библиотеки, а поэтому в нижеследующих стадиях типирования полученные данные секвенирования последовательно идентифицируют для каждого образца по индексам праймеров и индексам адаптеров, используемых в каждом образце. Последовательности каждого образца выравнивают с известной эталонной последовательностью ДНК, соответствующей ПЦР-продукту, с помощью компьютерной программы. Исходя из перекрывания последовательностей и характера их сцепления, сборку интактной последовательности для ПЦР-продукта осуществляют из последовательностей фрагментированной ДНК.

Настоящее изобретение относится к методам HLA-генотипирования высокого разрешения, осуществляемым путем секвенирования на Illumina GA, и, тем самым, к секвенированию гаплотипов, а также к автоматизации компьютерной программы типирования, к повышению производительности HLA-генотипирования и к снижению их стоимости.

Из-за требований, предъявляемых к длине ДНК-матрицы в современных методах секвенирования, и из-за небольшой длины ридов, исходные ПЦР-праймеры, используемые в методах HLA-SBT, больше не могут быть применены в HLA-типировании высокого разрешения, основанном на новых методах секвенирования. В настоящем изобретении были сконструированы новые ПЦР-праймеры с хорошей специфичностью и консервативностью, которые позволяют независимо амплифицировать экзоны 2, 3, 4 гена HLA-A, B и дают ПЦР-продукты с длиной не более чем 700 п.н., которые могут быть, в частности, использованы в секвенаторе Illumina GA (максимальная длина ДНК, подходящая для современного секвенатора Illumina GA, составляет 700 п.н.). Настоящее изобретение относится к набору ПЦР-праймеров, подходящих для широкомасштабного, высокопроизводительного и не дорогостоящего HLA-генотипирования у индивидуумов (в частности, у человека).

В техническом решении, используемом в настоящем изобретении, все наиболее поздние последовательности гена HLA-A/B загружают из web-сайта Интернета IMGT/HLA (<http://www.ebi.ac.uk/imgt/hla/>), а затем сохраняют на локальном диске как набор данных HLA-A; при этом все самые последние последовательности гена HLA-I, отличающиеся от последовательностей HLA-A, загружают как набор данных для сравнения. Указанные два набора данных сравнивают для поиска консервативных и специфических последовательностей для каждого генного сайта в двух концевых и внутренних частях экзонов 2, 3, 4, и сконструированную последовательность ПЦР-праймера сравнивают с полноразмерной последовательностью человеческого генома для поиска гомологии. Поскольку последовательность гена HLA-A/B имеет высокую степень сходства с последовательностями других генов, принадлежащих к молекулам HLA-I, то при конструировании ПЦР-праймеров 3'-конец праймера должен иметь, если это возможно, такую специфичность, которая гарантировала бы специфичность амплификации гена HLA-A/B с использованием этих праймеров. При этом длина ПЦР-продуктов составляет менее чем 700 п.н., а температура отжига прямого и обратного праймеров, по существу, является одинаковой.

Множество пар-кандидатов на праймеры HLA-A/B, удовлетворяющие требованиям протокола, используют для амплификации матричных ДНК, принадлежащих к общим серотипам HLA-A/B. Из этих праймеров было скринировано два набора ПЦР-праймеров

HLA-A/B (по 6 пар для каждого набора) с наибольшей консервативностью и специфичностью, для их применения в целях амплификации экзонов 2, 3 и 4, соответственно.

Два набора ПЦР-праймеров (по 6 пар для каждого набора) были использованы в качестве базовых праймеров, на основе которых было сконструировано 95 наборов индексированных праймеров, используемых для амплификации 95 и 950 ДНК-матриц общих серотипов HLA-A/B (серотипами этих матриц являются все общие серотипы HLA-A/B), соответственно. Все ПЦР-продукты секвенировали по обоим концам в секвенаторе Illumina GA Pair-End 100, а затем смешивали в равных количествах, и после сборки, результаты секвенирования сравнивали с первоначальными результатами типирования для подтверждения консервативности и специфичности ПЦР-праймеров.

Праймеры HLA-A, B, сконструированные в соответствии с настоящим изобретением, то есть два набора ПЦР-праймеров HLA-A/B (по 6 пар для каждого набора) для амплификации экзонов 2, 3 и 4, соответственно, представлены в таблицах 1 и 2.

Таблица 1
ПЦР-праймеры HLA-A,B

SEQ ID NO:	Праймер No.	Последовательность праймера	Применение праймера	Длина продукта
1	A-F2	CCTCTGYGGGGAGAAGCAA	Амплификации экзона 2 гена HLA-A	480 п.н.
2	A-R2	ATCTCGGACCCGGAGACTG		
3	A-F3	CGGGGCCAGGTTCTCACAC	Амплификации экзона 3 гена HLA-A	410 п.н.
4	A-R3	GGYGATATTCTAGTGTGGTCCCAA		
5	A-F4	GTGTCCCATGACAGATGCAAAA	Амплификации экзона 4 гена HLA-A	430 п.н.
6	A-R4	GGCCCTGACCTGCTAAAGG		
7	B-F2	AGGAGCGAGGGGACCGCA	Амплификации экзона 2 гена HLA-B	400 п.н.
8	B-R2	CGGGCCGGGGTCACTCAC		
9	B-F3	CGGGGCCAGGGTCTCACA	Амплификации экзона 3 гена HLA-B	370 п.н.
10	B-R3	GAGGCCATCCCGGCGAC		
11	B-F4	GCTGGTCACATGGGTGGTCTTA	Амплификации экзона 4 гена HLA-B	380 п.н.
12	B-R4	CTCCTTACCCCATCTCAGGGTG		

Таблица 2
ПЦР-праймеры HLA-A,B

SEQ ID NO:	Праймер No.	Последовательность праймера	Применение праймера	Длина продукта
13	A-F2s	CCTCTGYGGGGAGAAGCAA	Амплификации экзона 2 гена HLA-A	481 п.н.р
14	A-R2s	GGATCTCGGACCCGGAGACTGT		
15	A-F3s	TGGGCTGACCGYGGGGTTC	Амплификации экзона 3 гена HLA-A	403 п.н.
16	A-R3s	GGYGATATTCTAGTGTGGTCCCAA		
17	A-F4s	GTGTCCCATKACAGATGCAAAA	Амплификации экзона 4 гена HLA-A	405 п.н.
18	A-R4s	GGCCCTGACCTGCTAAAGG		
19	B-F2s	AGGAGCGAGGGGACCGCA	Амплификации экзона 2 гена HLA-B	400 п.н.
20	B-R2s	CGGGCCGGGGTCACTCAC		
21	B-F3s	CCAAAATCCCGCGGGTT	Амплификации экзона 3 гена HLA-B	405 п.н.
22	B-R3s	GAGGCCATCCCGGCGAC		
23	B-F4s	GCTGGTCACATGGGTGGTCTTA	Амплификации экзона 4 гена HLA-B	374 п.н.
24	B-R4s	TGACCCCTCATCCCCCTCCT		

Вырожденность праймеров означает смесь всех возможных различных последовательностей, представляющих собой все различные основания, кодирующие одну аминокислоту. Для повышения специфичности вырожденность может быть снижена путем замены оснований, встречающихся в различных организмах, основаниями в соответствии с таблицей кодонов, где 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.

В настоящем изобретении были сконструированы 2 набора ПЦР-праймеров (по три

пары для каждого набора) для амплификации экзонов 2, 3 и 4 HLA-C с применением метода конструирования ПЦР-праймеров для амплификации экзонов 2, 3 и 4 гена HLA-A/B.

В нижеследующих примерах 95 и 950 проб крови с известными HLA-генотипами подвергали ПЦР-амплификации генов HLA-C с использованием 2 отобранных наборов ПЦР-праймеров (по 3 пары для каждого набора), соответственно. Амплифицированные продукты секвенировали методом Сэнгера и методом секвенирования второго поколения. Результаты секвенирования использовали для HLA-C-типирования и сравнивали с исходными результатами типирования для подтверждения консервативности и специфичности ПЦР-праймеров.

Настоящее изобретение относится к 2 наборам ПЦР-праймеров (по три пары для каждого набора) для амплификации экзонов 2, 3 и 4 гена HLA-C, которые представляют собой SEQ ID NO: 25 и 26, 27 и 28, и 29 и 30, как показано в таблице 3, и SEQ ID NO: 31 и 32, 33 и 34, и 35 и 36, как показано в таблице 4. Указанные 6 пар ПЦР-праймеров имеют хорошую консервативность и специфичность, и могут охватывать полноразмерные последовательности экзонов 2, 3 и 4 HLA-C, где длина всех ПЦР-продуктов составляет менее чем 700 п.н., что удовлетворяет требованиям стандартного метода секвенирования Illumina Solexa. Кроме того, праймеры согласно изобретению могут быть также использованы для секвенирования методом Сэнгера.

Таблица 3 ПЦР-праймеры экзонов 2, 3 и 4 гена HLA-C				
SEQ ID NO:	No.	Последовательность праймера	Экзоны HLA-C	Длина продуктов
25	C-F2	GACCCGGGGAGCCGCGCA	2	455 п.н.
26	C-R2	TCGAGGGTCTGGGCGGGTT		
27	C-F3	CCTTTACCCGGTTTCATTTTCRGTTT	3	417 п.н.
28	C-R3	CTACGGGAGATGGGGAAGGCT		
29	C-F4	GTGTCGCAAGAGAGATRCAAAGTGT	4	451 п.н.
30	C-R4	GCTCTGGGAAAGGAGGRGAAGG		

Таблица 4 ПЦР-праймеры экзонов 2, 3 и 4 гена HLA-C				
SEQ ID NO:	No.	Последовательность праймера	Экзоны HLA-C	Длина продуктов
31	C-F2s	GACCCGGGGAGCCGCGCA	2	455 п.н.
32	C-R2s	TCGAGGGTCTGGGCGGGTT		
33	C-F3s	GCCCAGACCCTCGRCCGGA	3	443 п.н.
34	C-R3s	AGATRGGGAAGGCTCCCCACT		
35	C-F4s	TCTCAGGATRGTCACATGGGC	4	481 п.н.
36	C-R4s	GCTCTGGGAAARGAGGRGAAGG		

В соответствии с описанными выше методами, применяемыми на секвенаторе второго поколения для генотипирования HLA-DQB1, настоящее изобретение относится к ПЦР-праймерам для амплификации экзонов 2 и/или 3 HLA-DQB1, которые имеют последовательности SEQ ID NO: 37-40, представленные в таблице 5. Эти ПЦР-праймеры имеют хорошую консервативность и специфичность, и могут охватывать полноразмерные последовательности экзонов 2, 3 и 4 HLA-DQB1, где длина всех ПЦР-продуктов составляет менее чем 700 п.н., что удовлетворяет требованиям стандартного метода секвенирования Illumina Solexa. Кроме того, праймеры согласно изобретению могут быть также использованы для секвенирования по Сэнгеру.

Таблица 5
ПЦР-праймеры для амплификации соответствующих экзонов HLA-DQB1

SEQ ID NO:	Праймер No.	Последовательность праймера	Мишень амплификации	Длина амплифицированных продуктов
37	Q-F2	GATTCCYCGCAGAGGATTTCG	Экзон 2 HLA-DQB1	311 п.н.
38	Q-R2	AGGGGCRACACGCTCACCTC		
39	Q-F3	CCTGTCTGTTACTGCCCTCAGT	Экзон 3 HLA-DQB1	339 п.н.
40	Q-R3	GGCCCATAGTAACAGAAACTCAATA		

Генотипирование может быть осуществлено путем амплификации экзонов 2 и/или 3 HLA-DQB1 с применением пары праймеров для амплификации, и настоящее изобретение относится к таким методам генотипирования. По сравнению с ранее применяемыми методами, в таком генотипировании используется технология Illumina Solexa, которая позволяет проводить высокоразрешающее HLA-типирование с высокой производительностью и низкими затратами.

Конкретные способы осуществления изобретения

Способ секвенирования нуклеиновой кислоты

В одном из своих аспектов настоящее изобретение относится к способу определения нуклеотидной последовательности представляющей интерес нуклеиновой кислоты в образце, где указанный способ включает:

1) получение n образцов, где n представляет собой целое число ≥ 1 , а указанные образцы предпочтительно берут у млекопитающего, более предпочтительно у человека, в частности, такими образцами являются пробы человеческой крови; причем необязательно указанные n образцов, предназначенные для анализа, разделяют на m групп, где m представляет собой целое число, а $n \geq m \geq 1$;

2) амплификацию, где пару или множество пар индексированных праймеров используют для каждого образца, если от этого образца имеются матрицы; ПЦР-амплификацию осуществляют в условиях, подходящих для амплификации представляющей интерес нуклеиновой кислоты, где каждая пара индексированных праймеров состоит из прямого индексированного праймера и обратного индексированного праймера (оба эти праймера могут быть вырожденными), содержащих праймерные индексы, где указанные праймерные индексы, присутствующие в прямом индексированном праймере и в обратном индексированном праймере, могут быть одинаковыми или различными, то есть индексы праймеров в парах индексированных праймеров, используемых для различных образцов, являются различными;

3) объединение в пул: если $n > 1$, то ПЦР-продукты от каждого образца объединяют в один пул;

4) фрагментирование, где амплифицированные продукты подвергают неполному фрагментированию, очистке и выделению;

5) секвенирование, где выделенную ДНК-смесь подвергают секвенированию с применением технологии секвенирования второго поколения, предпочтительно парно-концевым методом (например, Illumina GA, Illumina HiSeq 2000), с получением последовательностей фрагментированной ДНК; и

6) сборку, где данные секвенирования, полученные для образцов, последовательно сопоставляют по уникальному праймерному индексу для каждого образца; каждую считанную последовательность сопоставляют путем выравнивания с эталонной последовательностью ДНК, соответствующей ПЦР-продуктам, с помощью программы выравнивания (такой как программа Blast, BWA); сборку полноразмерной последовательности представляющей интерес нуклеиновой кислоты осуществляют из последовательностей фрагментированной ДНК по перекрыванию и характеру сцепления последовательностей.

В одном из аспектов настоящего изобретения каждая пара праймерных индексов и

пара ПЦР-праймеров образует пару индексированных праймеров, где прямой и обратный ПЦР-праймеры имеют индекс прямого праймера и индекс обратного праймера у 5'-конца (или, необязательно, присоединенный посредством линкерной последовательности), соответственно.

- 5 В одном из вариантов изобретения указанными ПЦР-праймерами являются ПЦР-праймеры для амплификации гена HLA, в частности ПЦР-праймеры для амплификации гена HLA-A/B, предпочтительно ПЦР-праймеры для амплификации экзонов 2, 3 и 4 HLA-A/B и экзона 2 HLA-DRB1, предпочтительно ПЦР-праймеры для амплификации
10 ПЦР-праймеры для амплификации экзона 2 HLA-DRB1, как показано в таблице 7.

- В одном из вариантов изобретения указанными ПЦР-праймерами являются ПЦР-праймеры для амплификации гена HLA, в частности ПЦР-праймеры для амплификации гена HLA-C, предпочтительно ПЦР-праймеры для амплификации экзонов 2, 3 и 4 HLA-C, где предпочтительные указанные ПЦР-праймеры представлены в таблице 3 или в
15 таблице 4.

- В одном из вариантов изобретения указанными ПЦР-праймерами являются ПЦР-праймеры для амплификации гена HLA, предпочтительно ПЦР-праймеры для амплификации экзона 2 и/или 3 гена HLA-DQB1; где предпочтительные указанные ПЦР-праймеры представлены в таблице 5.

- 20 В одном из аспектов изобретения указанные праймерные индексы были сконструированы для ПЦР-праймеров, предпочтительно для ПЦР-праймеров, используемых для амплификации специфического гена HLA, более предпочтительно для ПЦР-праймеров, используемых для амплификации экзонов 2, 3 и 4 HLA-A/B и экзона 2 для HLA-DRB1, в частности, для ПЦР-праймеров, представленных в таблице
25 1, в таблице 2 или в таблице 7; где указанные праймерные индексы, в частности, содержат по меньшей мере 10, или по меньшей мере 20, или по меньшей мере 30, или по меньшей мере 40, или по меньшей мере 50, или по меньшей мере 60, или по меньшей мере 70, или по меньшей мере 80, или по меньшей мере 90, или все 95 пар из 95 пар праймерных индексов, представленных в таблице 6 (или набор праймерных индексов, состоящий из
30 10-95 пар (например, 10-95 пар, 20-95 пар, 30-95 пар, 40-95 пар, 50-95 пар, 60-95 пар, 70-95 пар, 80-95 пар, 90-95 пар, или 95 пар) из 95 пар праймерных индексов, представленных в таблице 6); и

- набор индексированных праймеров предпочтительно включает по меньшей мере PI-1 - PI-10, или PI-11 - PI-20, или PI-21 - PI-30, или PI-31 - PI-40, или PI-41 - PI-50, или PI-
35 51 - PI-60, или PI-61 - PI-70, или PI-71 - PI-80, или PI-81 - PI-90, или PI-91 - PI-95 из 95 пар праймерных индексов, представленных в таблице 6, или комбинации любых двух или более из указанных пар.

- В одном из вариантов изобретения указанное фрагментирование ДНК включает методы химического фрагментирования и методы физического фрагментирования, где
40 методами химического фрагментирования является ферментативное расщепление, а методами физического фрагментирования являются методы ультразвукового фрагментирования или методы механического фрагментирования.

- В одном из вариантов изобретения, после указанного фрагментирования ДНК, все ДНК-полосы, имеющие длину между максимальной длиной ридов в секвенаторе и
45 максимально приемлемой длиной ДНК в секвенаторе, очищают и выделяют, где указанными методами очистки и выделения являются, но не ограничиваются ими, выделение путем электрофореза и разрезания геля, и выделение на магнитных сферах.

В другом варианте изобретения способ секвенирования нуклеотидной

последовательности представляющей интерес нуклеиновой кислоты в тест-образце включает стадии 1)-4) по п.1 формулы изобретения, и нижеследующие стадии:

5) конструирование библиотеки, а именно, конструирование PCR-FREE-библиотеки для секвенирования с использованием библиотеки фрагментированных ПЦР-продуктов, где различные адаптеры библиотеки могут быть добавлены для распознавания различных PCR-FREE-библиотек для секвенирования, и где все ДНК-полосы, имеющие длину между максимальной длиной ридов в секвенаторе и максимально приемлемой длиной ДНК в секвенаторе, предпочтительно ДНК-фрагменты длиной 450-750 п.н., очищают и выделяют;

6) секвенирование, где выделенную ДНК-смесь подвергают секвенированию с применением технологии секвенирования второго поколения, предпочтительно технологии парно-концевого секвенирования (например, на секвенаторе Illumina GA, Illumina HiSeq 2000), с получением последовательностей фрагментированной ДНК; и

7) сборка, где данные секвенирования, полученные для образцов, последовательно сопоставляют по различным последовательностям адаптеров библиотеки и по уникальному праймерному индексу для каждого образца; каждую считанную последовательность сопоставляют путем выравнивания с эталонной последовательностью ДНК, соответствующей ПЦР-продуктам, с помощью программы выравнивания (такой как программа Blast, BWA); и сборку полноразмерной последовательности представляющей интерес нуклеиновой кислоты осуществляют из последовательностей фрагментированной ДНК по перекрыванию и характеру сцепления последовательностей.

В одном из своих аспектов настоящее изобретение также относится к применению вышеописанного способа для HLA-типирования, где указанный способ отличается тем, что он включает: секвенирование взятого у пациента образца (в частности, пробы крови) указанным методом и сопоставление результатов секвенирования с данными о последовательностях экзонов HLA, предпочтительно экзонов 2, 3, 4 HLA-A/B, экзонов 2, 3 и/или 4 HLA-C, экзона 2 и/или 3 гена HLA-DQB1 и/или экзона 2 HLA-DRB1 в базе данных HLA (такой как профессиональная база данных HLA IMGT); где если результат выравнивания последовательностей дает 100% соответствие, то это означает, что HLA-генотип соответствующего образца был точно определен.

Набор праймерных индексов

В другом своем аспекте настоящее изобретение относится к набору праймерных индексов, где указанный набор содержит по меньшей мере 10, или по меньшей мере 20, или по меньшей мере 30, или по меньшей мере 40, или по меньшей мере 50, или по меньшей мере 60, или по меньшей мере 70, или по меньшей мере 80, или по меньшей мере 90, или все 95 пар из 95 пар праймерных индексов, представленных в таблице 6 (или набор праймерных индексов, состоящий из 10-95 пар (например, 10-95 пар, 20-95 пар, 30-95 пар, 40-95 пар, 50-95 пар, 60-95 пар, 70-95 пар, 80-95 пар, 90-95 пар, или 95 пар) из 95 пар праймерных индексов, представленных в таблице 6); и

указанный набор индексированных праймеров предпочтительно включает по меньшей мере PI-1 - PI-10, или PI-11 - PI-20, или PI-21 - PI-30, или PI-31 - PI-40, или PI-41 - PI-50, или PI-51 - PI-60, или PI-61 - PI-70, или PI-71 - PI-80, или PI-81 - PI-90, или PI-91 - PI-95 из 95 пар праймерных индексов, представленных в таблице 6, или комбинации любых двух или более из указанных пар.

Настоящее изобретение также относится к применению указанного набора праймерных индексов в методах ПЦР-секвенирования, где, в частности, каждая пара праймерных индексов и пара ПЦР-праймеров для амплификации представляющей

интерес тестируемой последовательности образует пару индексированных праймеров, где прямой и обратный ПЦР-праймеры имеют индекс прямого праймера или индекс обратного праймера у 5'-конца (или, необязательно, присоединенный посредством линкерной последовательности), соответственно.

5 В одном из аспектов изобретения указанными ПЦР-праймерами являются ПЦР-праймеры для амплификации специфического гена HLA, в частности, ПЦР-праймеры для амплификации экзонов 2, 3 и 4 HLA-A/B и экзона 2 HLA-DRB1, предпочтительно ПЦР-праймеры для амплификации экзонов 2, 3 и 4 HLA-A/B, как показано в таблице 1 или в таблице 2, или предпочтительно ПЦР-праймеры для амплификации экзона 2 HLA-DRB1, как показано в таблице 7, или предпочтительно ПЦР-праймеры для амплификации экзонов 2, 3 и/или 4 HLA-C, где предпочтительные указанные ПЦР-праймеры представлены в таблице 3 или в таблице 4, или предпочтительно ПЦР-праймеры для амплификации экзона 2 и/или 3 гена HLA-DQB1, где предпочтительные указанные ПЦР-праймеры представлены в таблице 5.

15 В другом своем аспекте настоящее изобретение относится к набору индексированных праймеров, содержащему указанный набор праймерных индексов и пару ПЦР-праймеров для амплификации представляющей интерес тестируемой последовательности, где пара индексированных праймеров содержит пару праймерных индексов и пару ПЦР-праймеров, где прямой и обратный ПЦР-праймеры имеют индекс прямого праймера и индекс обратного праймера у 5'-конца (или, необязательно, присоединенный посредством линкерной последовательности), соответственно.

В одном из вариантов изобретения, указанными ПЦР-праймерами являются ПЦР-праймеры для амплификации специфического гена HLA, в частности, ПЦР-праймеры для амплификации экзонов 2, 3 и 4 гена HLA-A/B и экзона 2 HLA-DRB1, предпочтительно ПЦР-праймеры для амплификации экзонов 2, 3 и 4 HLA-A/B, как показано в таблице 1 или в таблице 2, или предпочтительно ПЦР-праймеры для амплификации экзона 2 HLA-DRB1, как показано в таблице 7, предпочтительно ПЦР-праймеры для амплификации экзонов 2, 3 и/или 4 HLA-C, где предпочтительные указанные ПЦР-праймеры представлены в таблице 3 или в таблице 4, или предпочтительно ПЦР-праймеры для амплификации экзона 2 и/или 3 гена HLA-DQB1, где предпочтительные указанные ПЦР-праймеры представлены в таблице 5.

В другом своем аспекте настоящее изобретение также относится к применению указанных индексированных праймеров в методах ПЦР-секвенирования.

Способ HLA-типирования

35 В одном из своих аспектов настоящее изобретение относится к способу HLA-типирования, включающему:

1) получение n образцов, где n представляет собой целое число ≥ 1 , а указанные образцы предпочтительно, берут у млекопитающего, более предпочтительно у человека, в частности, такими образцами являются пробы человеческой крови;

40 2) разделение анализируемых n образцов на m групп, где m равно целому числу, а $n \geq m \geq 1$;

3) амплификацию, где пару индексированных праймеров используют для каждого образца, если имеются матрицы от этого образца; ПЦР-амплификацию осуществляют в условиях, подходящих для амплификации представляющей интерес нуклеиновой кислоты, где каждая пара индексированных праймеров состоит из прямого индексированного праймера и обратного индексированного праймера (оба эти праймера могут быть вырожденными), содержащих праймерные индексы, где праймерные индексы, присутствующие в прямом индексированном праймере и в обратном

индексированном праймере, могут быть одинаковыми или различными, то есть индексы праймеров в парах индексированных праймеров, используемых для различных образцов, являются различными;

4) объединение в пул, где продукты ПЦР-амплификации от каждого образца объединяют в один пул с получением библиотек ПЦР-продуктов;

5) фрагментирование, где полученные библиотеки ПЦР-продуктов подвергают неполному фрагментированию;

6) конструирование библиотек, а именно, конструирование PCR-FREE-библиотек для секвенирования из библиотеки фрагментированных ПЦР-продуктов с применением технологии индексирования адаптеров библиотек, где различные адаптеры библиотеки могут быть добавлены для распознавания различных PCR-FREE-библиотек для секвенирования, и все ДНК-полосы, имеющие длину между максимальной длиной ридов в секвенаторе и максимально приемлемой длиной ДНК в секвенаторе, предпочтительно ДНК-фрагменты длиной 450-750 п.н., выделяют;

7) секвенирование, где выделенную ДНК-смесь подвергают секвенированию с применением технологии секвенирования второго поколения, предпочтительно технологии парно-концевого секвенирования (например, на Illumina GA, Illumina HiSeq 2000), с получением последовательностей фрагментированных ДНК; и

8) сборку, где данные секвенирования, полученные для образцов, последовательно сопоставляют по различным последовательностям адаптеров библиотеки и по уникальному праймерному индексу для каждого образца; каждый рид сопоставляют путем выравнивания с эталонной последовательностью ДНК, соответствующей ПЦР-продуктам, с помощью программы выравнивания (такой как программа Blast, BWA); сборку полноразмерной последовательности представляющей интерес нуклеиновой кислоты осуществляют из последовательностей фрагментированной ДНК по перекрыванию и характеру сцепления последовательностей; и

9) типирование, где результаты секвенирования сопоставляют с данными о последовательностях экзонов HLA, предпочтительно экзонов 2, 3, 4 HLA-A/B, экзонов 2, 3 и/или 4 HLA-C, экзона 2 и/или 3 гена HLA-DQB1 и/или экзона 2 HLA-DRB1 в базе данных HLA (такой как профессиональная база данных HLA IMGT); при этом если результат выравнивания последовательностей дает 100% соответствие, то это означает, что HLA-генотип соответствующего образца был точно определен.

В способе HLA-типирования согласно изобретению, пара индексированных праймеров содержит пару праймерных индексов и пару ПЦР-праймеров, где прямой и обратный ПЦР-праймеры имеют индекс прямого праймера и индекс обратного праймера у 5'-конца (или, необязательно, присоединены посредством линкерной последовательности), соответственно.

В одном из вариантов изобретения указанными ПЦР-праймерами являются ПЦР-праймеры для амплификации специфического гена HLA, в частности, ПЦР-праймеры для амплификации экзонов 2, 3 и 4 гена HLA-A/B и экзона 2 HLA-DRB1, предпочтительно ПЦР-праймеры для амплификации экзонов 2, 3 и 4 HLA-A/B, как показано в таблице 1 или в таблице 2, или предпочтительно ПЦР-праймеры для амплификации экзона 2 HLA-DRB1, как показано в таблице 7, или предпочтительно ПЦР-праймеры для амплификации экзонов 2, 3 и/или 4 HLA-C, где предпочтительные указанные ПЦР-праймеры представлены в таблице 3 или в таблице 4, или предпочтительно ПЦР-праймеры для амплификации экзона 2 и/или 3 гена HLA-DQB1; где предпочтительные указанные ПЦР-праймеры представлены в таблице 5.

В одном из вариантов изобретения указанные праймерные индексы представляют

собой набор праймерных индексов, описанных выше.

В одном из вариантов способа HLA-типирования согласно изобретению, указанное фрагментирование ДНК включает методы химического фрагментирования и методы физического фрагментирования, где методами химического фрагментирования является ферментативное расщепление, а методами физического фрагментирования являются методы ультразвукового фрагментирования или методы механического фрагментирования.

В одном из вариантов способа HLA-типирования согласно изобретению, указанными методами очистки и выделения являются, но не ограничиваются ими, выделение путем электрофореза и разрезания геля, и выделение на магнитных сферах.

В одном из вариантов способа HLA-типирования согласно изобретению, конструирование библиотек для PCR-FREE-секвенирования из библиотек фрагментированных ПЦР-продуктов методом индексирования адаптера библиотеки включает: добавление m адаптеров библиотеки к m библиотекам ПЦР-продуктов, полученных в стадии 2), где в каждой библиотеке ПЦР-продуктов используется другой адаптер библиотеки; и конструирование m библиотек индексированных адаптеров для секвенирования; объединение в пул m библиотек индексированных адаптеров для секвенирования в равном молярном соотношении с получением смеси библиотек индексированных адаптеров для секвенирования, где указанный способ присоединения адаптеров библиотек означает прямое связывание посредством ДНК-лигазы без проведения ПЦР.

ПЦР-праймеры для HLA-генотипирования

В одном из своих аспектов настоящее изобретение относится к ПЦР-праймерам для HLA-генотипирования, где указанные праймеры отличаются тем, что они представляют собой ПЦР-праймеры для амплификации экзонов 2, 3 и 4 гена HLA-A/B и экзона 2 HLA-DRB1, предпочтительно ПЦР-праймеры для амплификации экзонов 2, 3 и 4 HLA-A/B, как показано в таблице 1 или в таблице 2, или предпочтительно ПЦР-праймеры для амплификации экзона 2 HLA-DRB1, как показано в таблице 7, предпочтительно ПЦР-праймеры для амплификации экзонов 2, 3 и/или 4 HLA-C, где предпочтительные указанные ПЦР-праймеры представлены в таблице 3 или в таблице 4, или предпочтительно ПЦР-праймеры для амплификации экзона 2 и/или 3 гена HLA-DQB1, где предпочтительные указанные ПЦР-праймеры представлены в таблице 5.

Настоящее изобретение также относится к способу секвенирования с использованием указанных ПЦР-праймеров, включающему:

получение образца, в частности, пробы крови, где указанную пробу крови берут предпочтительно у млекопитающего, в частности, у человека;

амплификацию, где амплификацию ДНК из пробы крови осуществляют с использованием ПЦР-праймеров с получением ПЦР-продуктов и очистку этих ПЦР-продуктов;

секвенирование, где ПЦР-продукты подвергают секвенированию, и где методом секвенирования может быть метод секвенирования по Сэнгеру или метод секвенирования второго поколения (такой как метод HiSeq 2000, Illumina GA и Roche454).

В другом своем аспекте настоящее изобретение также относится к применению указанных ПЦР-праймеров для HLA-генотипирования, отличающемуся тем, что оно включает: использование указанных ПЦР-праймеров, сборку и анализ путем сопоставления результатов, полученных вышеуказанным методом секвенирования; и сравнение результатов секвенирования со стандартными последовательностями, имеющимися в базе данных с получением результатов HLA-генотипирования.

В другом своем аспекте настоящее изобретение также относится к набору для HLA-генотипирования, содержащему указанные ПЦР-праймеры.

ПЦР-праймеры для HLA-A,B-генотипирования

В одном из своих аспектов настоящее изобретение относится к набору ПЦР-праймеров для HLA-A,B-генотипирования, где указанный набор отличается тем, что он содержит указанные ПЦР-праймеры, представленные в таблице 1 или в таблице 2.

В другом своем аспекте настоящее изобретение относится к способу секвенирования с использованием указанных ПЦР-праймеров для HLA-A,B-генотипирования, где указанный способ включает:

- 10 получение образца, в частности, пробы крови, где указанную пробу крови берут предпочтительно от млекопитающего, в частности, у человека;
- амплификацию, где амплификацию ДНК из пробы крови осуществляют с использованием ПЦР-праймеров с получением ПЦР-продуктов и очистку этих ПЦР-продуктов;
- 15 секвенирование, где ПЦР-продукты подвергают секвенированию, и где методом секвенирования может быть метод секвенирования по Сэнгеру или метод секвенирования второго поколения (такой как метод HiSeq 2000, Illumina GA и Roche454).

В другом своем аспекте, настоящее изобретение также относится к применению указанных ПЦР-праймеров для HLA-генотипирования, отличающемуся тем, что оно включает использование указанных ПЦР-праймеров, сборку и анализ путем сопоставления результатов, полученных вышеуказанным методом секвенирования, и сравнение результатов секвенирования со стандартными последовательностями, имеющимися в базе данных, с получением результатов HLA-генотипирования.

В другом своем аспекте настоящее изобретение также относится к набору для HLA-генотипирования, содержащему указанные ПЦР-праймеры для HLA-A,B-генотипирования согласно изобретению.

ПЦР-праймеры для HLA-C-генотипирования

Настоящее изобретение также относится к новому способу амплификации экзонов 2, 3 и 4 гена HLA-C, отличающемуся тем, что он включает ПЦР-амплификацию с использованием пар праймеров для амплификации согласно изобретению, последовательности которых представлены в таблице 3 или в таблице 4.

Поскольку экзоны 2, 3 и 4 HLA-C могут быть амплифицированы с помощью ПЦР, то способ согласно изобретению является особенно подходящим для HLA-C-генотипирования. Поскольку указанные продукты были получены с применением такого способа, а праймеры для амплификации согласно изобретению имеют размеры в пределах 700 п.н., то в процессе последующего генотипирования может быть использован не известный уже метод HLA-C-генотипирования, а метод HLA-SBT, разработанный на основе технологии секвенирования Illumina Solexa.

Настоящее изобретение также относится к способу секвенирования экзонов 2, 3 и 4 гена HLA-C в образцах, где указанный способ включает следующие стадии:

- 1) получение образца и экстракция ДНК из этого образца;
- 2) амплификация ДНК с использованием пары ПЦР-праймеров для HLA-C-генотипирования согласно изобретению с получением ПЦР-продуктов, и предпочтительно очистка ПЦР-продуктов, где указанная пара ПЦР-праймеров
- 45 предпочтительно выбрана из группы, состоящей из пары праймеров SEQ ID NO: 25 и SEQ ID NO: 26, SEQ ID NO: 27 и SEQ ID NO: 28, SEQ ID NO: 29 и SEQ ID NO: 30, или SEQ ID NO: 31 и SEQ ID NO: 32, SEQ ID NO: 33 и SEQ ID NO: 34, SEQ ID NO: 35 и SEQ ID NO: 36;

3) секвенирование ПЦР-продуктов предпочтительно методом секвенирования второго поколения, таким как Illumina Solexa или Roche454.

Настоящее изобретение также относится к способу HLA-C-генотипирования, включающему:

5 1) ПЦР-амплификацию экзонов 2, 3 и/или 4 гена HLA-C тестируемого образца с использованием пары ПЦР-праймеров для HLA-C-генотипирования согласно изобретению, где указанная пара ПЦР-праймеров предпочтительно выбрана из группы, состоящей из пар праймеров SEQ ID NO: 25 и SEQ ID NO: 26, SEQ ID NO: 27 и SEQ ID NO: 28, SEQ ID NO: 29 и SEQ ID NO: 30, или SEQ ID NO: 31 и SEQ ID NO: 32, SEQ ID NO: 33 и SEQ ID NO: 34, SEQ ID NO: 35 и SEQ ID NO: 36;

10 2) секвенирование амплифицированных экзонов; сравнение результатов секвенирования с результатами для стандартных последовательностей, имеющихся в базе данных, и получение результатов генотипирования, где указанное секвенирование осуществляют методом секвенирования по Сэнгеру или методом секвенирования второго поколения, таким как Illumina Solexa или Roche454.

В другом своем аспекте настоящее изобретение также относится к набору для HLA-C-генотипирования, включающему пару ПЦР-праймеров для HLA-C-генотипирования согласно изобретению, предпочтительно выбранную из группы, состоящей из пар праймеров: SEQ ID NO: 25 и SEQ ID NO: 26, SEQ ID NO: 27 и SEQ ID NO: 28, SEQ ID NO: 29 и SEQ ID NO: 30, или SEQ ID NO: 31 и SEQ ID NO: 32, SEQ ID NO: 33 и SEQ ID NO: 34, SEQ ID NO: 35 и SEQ ID NO: 36. В одном из вариантов изобретения указанный набор также включает дополнительные агенты, например, агенты для амплификации ДНК, очистки ДНК и/или секвенирования ДНК.

Генотипирование может быть осуществлено посредством амплификации экзонов 2, 3 и 4 HLA-C с использованием пары праймеров для амплификации и с применением способа генотипирования согласно изобретению. Способ генотипирования согласно изобретению, по сравнению с известным методом, может быть проведен с применением метода секвенирования Illumina Solexa, способствует повышению производительности, упрощает процедуру и при этом позволяет экономить время и средства.

30 ПЦР-праймеры для HLA-DQB1-генотипирования

Настоящее изобретение также относится к новому способу амплификации экзона 2 и/или 3 HLA-DQB1, отличающемуся тем, что он включает ПЦР-амплификацию с использованием пар праймеров для амплификации согласно изобретению, где указанные пары праймеров для амплификации представлены в таблице 5.

35 Поскольку экзоны 2 и/или 3 HLA-DQB1 могут быть амплифицированы с помощью ПЦР, то способ согласно изобретению является особенно подходящим для HLA-DQB1-генотипирования. Поскольку указанные продукты были получены с применением такого способа, и праймеры для амплификации согласно изобретению имеют размеры в пределах 300-400 п.н., то в процессе последующего типирования может быть применен 40 не известный метод HLA-DQB1-генотипирования, а метод HLA-SBT на основе технологии секвенирования Illumina Solexa.

Настоящее изобретение также относится к способу секвенирования экзонов 2 и/или 3 гена HLA-DQB1 в образцах, включающему нижеследующие стадии:

1) получение образца и экстракция ДНК из этого образца;

45 2) амплификация ДНК с использованием пары ПЦР-праймеров для HLA-DQB1-генотипирования согласно изобретению, предпочтительно с использованием пар ПЦР-праймеров, представленных в таблице 5, с получением ПЦР-продуктов, и предпочтительно очистка этих ПЦР-продуктов;

3) секвенирование ПЦР-продуктов, предпочтительно методом секвенирования второго поколения, таким как Illumina Solexa или Roche454.

В другом своем аспекте настоящее изобретение относится к усовершенствованному способу HLA-DQB1-генотипирования, включающему:

- 5 1) амплификацию тестируемых экзонов 2 и/или 3 гена HLA-DQB1 с использованием пары ПЦР-праймеров для HLA-DQB1-генотипирования согласно изобретению, предпочтительно с использованием пар ПЦР-праймеров, представленных в таблице 5;
- 2) секвенирование амплифицированных экзонов; сравнение результатов
- 10 секвенирования со стандартными последовательностями, имеющимися в базе данных; и получение результатов генотипирования, где указанным способом секвенирования может быть метод секвенирования Сэнгера или метод секвенирования второго поколения, такой как Illumina Solexa или Roche454.

В другом своем аспекте настоящее изобретение также относится к набору для HLA-DQB1-генотипирования, включающему пару ПЦР-праймеров для HLA-DQB1-генотипирования согласно изобретению, предпочтительно пары ПЦР-праймеров для амплификации, представленные в таблице 5. В одном из вариантов изобретения указанный набор также включает дополнительные агенты, например, агенты для амплификации ДНК, очистки ДНК и/или секвенирования ДНК.

20 Описание графического материала

Фигура 1. На этой фигуре проиллюстрирована сборка последовательностей после мечения праймерными индексами, фрагментирования ДНК и секвенирования ДНК. Последовательности индексов прямого и обратного праймеров, Индекс-N-F/R (1), вводят в оба конца ПЦР-продуктов образца No. N. После фрагментирования физическим

25 методом, ПЦР-продукты включают продукты, несущие последовательности праймерных индексов у одного конца, продукты, не содержащие последовательности праймерных индексов у своих двух концов, и полностью нефрагментированные продукты. Все ДНК-полосы, имеющие длину между максимальной длиной ридов секвенатора и максимально приемлемой длиной ДНК в секвенаторе, очищают и выделяют путем разрезания геля,

30 а затем используют для секвенирования (2). Данные секвенирования ПЦР-продуктов, принадлежащих к образцу No. N, метят с использованием последовательности Индекс-N-F/R. Известные эталонные последовательности ПЦР-продуктов используют для определения относительной локализации положений считываемых последовательностей, и результаты секвенирования полноразмерных ПЦР-продуктов собирают с учетом

35 перекрывания и характера сцепления считываемых последовательностей (3, 4).

Фигура 2. На этой фигуре проиллюстрированы результаты электрофореза ПЦР-продуктов соответствующих экзонов HLA-A/B/DRB1 в образце No.1 примера 2. На электрофореграмме видно, что ПЦР-продукты представляют собой серию отдельных полос размером 300 п.н. - 500 п.н., где дорожка М представляет собой маркер

40 молекулярной массы (DL 2000, Takara Co.), дорожки 1-7 представляют собой ПЦР-продукты экзонов (A2, A3, A4, B2, B3, B4, DRB1-2) HLA-A/B/DRB1 образца No.1, а в негативном контроле (N) полоса амплификации отсутствует. Другие образцы показывают аналогичные результаты.

Фигура 3. На этой фигуре проиллюстрированы результаты электрофореза ДНК после фрагментирования HLA-Mix, как описано в примере 4 (до и после разрезания геля), где площадь гелевых срезов составляет 450-750 п.н. Дорожка М представляет собой маркер молекулярной массы (ДНК-лэддер NEB-50 п.н.), дорожка 1 представляет собой результаты электрофореза HLA-Mix до разрезания геля, а дорожка 2 представляет

собой гель с HLA-Mix после разрезания.

Фигура 4. Программа с захватом изображения на экране для конструирования консенсусной последовательности образца No.1 примера 6, иллюстрирующая сборку полноразмерной последовательности ПЦР-продуктов на основе праймерных индексов и перекрывания ДНК-фрагментов. Номенклатуру HLA-генотипов можно найти на сайте <http://www.ebi.ac.uk/imgt/hla/align.html>. Результаты для всей кодирующей последовательности A*02:03:01 A*11:01:01 приводятся в столбце выходных данных слева, где последовательность экзона 2 идентична исходному известному результату для Матрицы 1.

Фигура 5. На этой фигуре проиллюстрирован ПЦР-продукт после мечения праймерными индексами и адаптерным индексом. Во время эксперимента праймерные индексы вводят в оба конца ПЦР-продукта каждого образца одновременно с помощью ПЦР, а затем множество ПЦР-продуктов, несущих различные праймерные индексы, объединяют в пул и конструируют библиотеку для секвенирования. В процессе конструирования библиотек для секвенирования, в том случае, если необходимо сконструировать множество таких библиотек, эти библиотеки могут быть помечены соответствующими адаптерами, несущими различные адаптерные индексы. После завершения конструирования библиотек, множество библиотек для секвенирования, меченных различными адаптерными индексами, объединяют в пул и одновременно секвенируют в секвенаторе Illumina GA (праймерные индексы могут быть идентичными в различных библиотеках для секвенирования, помеченных различными адаптерными индексами). После получения результатов секвенирования, информация о последовательностях ДНК для каждого образца может быть получена путем скрининга результатов секвенирования в целях поиска информации о последовательностях адаптерных индексов и праймерных индексов.

Фигура 6. На этой фигуре проиллюстрированы результаты электрофореза ПЦР-продуктов для экзонов 2, 3, 4 гена HLA-C некоторых образцов примера 8. На электрофореграмме видно, что ПЦР-продукты представляют собой серию отдельных полос размером 400-500 п.н., где дорожка М представляет собой эталон молекулярных масс стандартных ДНК (DL 2000, Takara Co.).

Фигура 7. На этой фигуре проиллюстрированы результаты электрофореза ДНК с разрезанием геля после фрагментирования HLA-Mix, как описано в примере 8, где площадь гелевых срезов составляет 450-750 п.н. Дорожка М представляет собой маркер молекулярной массы (ДНК-лэддер NEB-50 п.н.), дорожка 1 представляет собой гель с HLA-Mix до разрезания, а дорожка 2 представляет собой гель с HLA-Mix после разрезания.

Фигура 8. Программа с захватом изображения на экране для конструирования консенсусной последовательности экзона 2 HLA-C-сайта образца No.2 в примере 8. Сначала последовательности-риды С-сайта образца сопоставляют путем выравнивания с эталонной последовательностью с помощью компьютерной программы BWA, затем конструируют консенсусные последовательности экзонов 2, 3, 4 С-сайта образца, после чего определяют последовательность гаплотипа каждого экзона С-сайта исходя из характера сцепления между SNP; и наконец, определяют тип образца по перекрыванию последовательностей экзонов этого гаплотипа. Как показано на этой фигуре, два гетерозиготных SNP находились в области 695-764 п.н. С-генной последовательности образца No.2, и исходя из считываемых последовательностей «рид 1» и «рид 2» можно определить, что характер сцепления SNP представляет собой А-С, G-А («...» на этой фигуре означает, что эти основания идентичны основаниям эталонной

последовательности). Эти последовательности соответствуют заштрихованным частям последовательностей типа C*010201 и C*07020101, соответственно. Оценка характера сцепления в других областях давала аналогичные результаты.

Фигура 9. На этой фигуре проиллюстрированы результаты электрофореза ПЦР-продуктов для экзонов 2, 3, 4 HLA-C-сайта для 26 образцов примера 9. На этой фигуре все ПЦР-продукты имеют длину менее чем 500 п.н.; наблюдается одна электрофоретическая полоса, неспецифическая полоса отсутствует, а эффективности амплификации одной той же пары праймеров в различных образцах являются аналогичными.

Фигура 10. На этой фигуре проиллюстрированы результаты анализа данных секвенирования продуктов ПЦР-амплификации матрицы 1 с помощью компьютерной программы uTure в примере 9. В столбце выходных данных слева приведены результаты *08:01:01 C*15:05:01, которые идентичны исходному известному результату для матрицы 1.

Фигура 11. На этой фигуре проиллюстрированы результаты электрофореза ПЦР-продуктов для экзонов 2+3 гена HLA-DQB1 в 94 образцах примера 10. На электрофореграмме видно, что ПЦР-продукты представляют собой серию отдельных полос размером 250 п.н. - 500 п.н., где дорожка М представляет собой эталон молекулярных масс стандартных ДНК (DL 2000, Takara Co.). Дорожки PI-1 - PI-94 представляют собой продукты ПЦР-амплификации экзонов 2+3 гена HLA-DQB1 в 94 образцах, а в негативном контроле (N) полоса амплификации отсутствует.

На фигуре 12 проиллюстрированы результаты электрофореза ДНК с разрезанием геля после фрагментирования HLA-Q-Mix, как описано в примере 10, где площадь гелевых срезов составляет 350-550 п.н. Дорожка М представляет собой маркер молекулярной массы стандартных ДНК (ДНК-лэддер NEB-50 п.н.), дорожка 1 представляет собой гель с HLA-Q-Mix до разрезания, а дорожка 2 представляет собой гель с HLA-Q-Mix после разрезания.

На фигуре 13 представлена программа с захватом изображения на экране для конструирования консенсусной последовательности образца No.7 в примере 10, иллюстрирующая главную процедуру анализа данных. Сначала считываемую последовательность DQB1-сайта образца выравнивают с эталонной последовательностью с помощью компьютерной программы BWA, затем конструируют консенсусные последовательности экзонов 2, 3 DQB1-образца, после чего определяют последовательность гаплотипов экзонов 2, 3 DQB1 исходя из характера сцепления между SNP. Как показано на этой фигуре, шесть гетерозиготных SNP находятся в области 2322-2412 последовательности гена DQB1 образца No.7; и исходя из считываемой последовательности «рид 1» можно определить, что характер сцепления SNP1-SNP5 составляет T-G-T-C-C; исходя из последовательности «рид 2» можно определить, что характер сцепления других SNP1-SNP5 составляет C-C-A-G-T; исходя из считываемой последовательности «рид 3» можно определить, что характер сцепления SNP3-SNP6 составляет A-G-T-G; исходя из считываемой последовательности «рид 4» можно определить, что характер сцепления других SNP3-SNP6 составляет T-C-C-A; а исходя из вышеуказанных характер сцепления для вышеуказанных SNP было определено, что считываемая последовательность «рид 1» сцеплена с последовательностью «рид 4», а считываемая последовательность «рид 2» сцеплена с последовательностью «рид 3», при этом полная комбинация SNP в этой области составляет T-G-T-C-C-A и C-C-A-G-T-G, и эти последовательности соответствуют заштрихованным частям последовательностей типа DQB1*0303 и DQB1*0602. Оценка характера сцепления в

других областях давала аналогичные результаты.

На фигуре 14 представлена электрофореграмма продуктов примера 11, полученных после амплификации экзонов 2 и 3 HLA-DQB1-сайта и амплификации экзонов 2 и 3 с использованием двух пар ПЦР-праймеров, соответственно. На этой электрофореграмме показаны три серии ПЦР-продуктов от семи ДНК-матриц, где все ПЦР-продукты имеют длину менее чем 500 п.н., наблюдается одна электрофоретическая полоса, а неспецифическая полоса отсутствует. В негативной контроле (N) полоса амплификации отсутствует, а дорожка М представляет собой эталон молекулярных масс стандартных ДНК (DL 2000, Takara Co.).

На фигуре 15 проиллюстрированы результаты анализа данных секвенирования продуктов ПЦР-амплификации экзонов 2 и 3 HLA-DQB1 матрицы 7 с помощью компьютерной программы uType в примере 11. В столбце выходных данных слева приведены результаты DQB1*03:03 DQB1*06:02, которые идентичны исходному известному результату для матрицы 7.

На фигуре 16 представлены результаты электрофореза ПЦР-продуктов для соответствующих экзонов HLA-A/B/C/DQB1 в образце No.1 примера 12. На электрофореграмме видно, что ПЦР-продукты представляют собой серию отдельных полос размером 300 п.н. - 500 п.н., где дорожка М представляет собой маркер молекулярных масс (DL 2000, Takara Co.), дорожки 1-10 представляют собой продукты ПЦР-амплификации экзонов (A2, A3, A4, B2, B3, B4, C2, C3, C4, DQB1) HLA-A/B/C/DQB1 образца No.1, а в негативном контроле (N) полоса амплификации отсутствует. Результаты для других образцов являются аналогичными.

На фигуре 17 представлен результат выделения из агарозного геля после объединения в пул 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 и HLA-10-Mix в равном молярном соотношении, как описано в примере 12. Дорожка М представляет собой маркер молекулярных масс, дорожка 1 представляет собой результаты электрофоретического анализа пула, а дорожка 2 представляет собой электрофореграмму гелевых срезов, содержащих ДНК-фрагменты длиной 450-750 п.н.

На фигуре 18 представлена программа с захватом изображения на экране для конструирования консенсусной последовательности экзона 2 HLA-C-сайта образца No.1 в примере 12. Сначала считываемую последовательность С-сайта образца сопоставляют путем выравнивания с эталонной последовательностью с помощью компьютерной программы BWA, затем конструируют консенсусные последовательности экзонов 2, 3, 4 С-сайта образца, после чего определяют последовательности гаплотипа каждого экзона С-сайта исходя из характера сцепления между SNP; и наконец, определяют тип образца по перекрыванию последовательностей экзонов этого гаплотипа. Как показано на этой фигуре, два гетерозиготных SNP находились в области 695-764 С-генной последовательности образца No.1, и исходя из считываемых последовательностей «рид 1» и «рид 2» можно определить, что характер сцепления SNP представляет собой А-С, G-A («...» на этой фигуре означает, что эти основания идентичны основаниям эталонной последовательности). Эти последовательности соответствуют заштрихованным частям последовательностей типа С*010201 и С*07020101, соответственно. Оценка характера сцепления в других областях давала аналогичные результаты.

Примеры

Варианты настоящего изобретения подробно описаны в нижеследующих примерах. Однако для специалиста в данной области очевидно, что нижеследующие примеры

приводятся лишь для иллюстрации настоящего изобретения и не должны рассматриваться как ограничение его объема.

В примерах 1-6 настоящего изобретения экзоны 2, 3, 4 HLA-A/B и экзон 2 HLA-DRB1 в 95 образцах были генотипированы с использованием комбинации праймерных индексов+стратегии неполного фрагментирования ДНК+парно-концевого метода секвенирования на секвенаторе Illumina GA 100 (ПЦР-продукты имели длину в пределах от 290 п.н. до 500 п.н.), что указывает на то, что такой способ согласно изобретению может быть применен для типирования генных фрагментов, длина которых превышает максимальную длину считываемых последовательностей в секвенаторе, который в достаточной степени удовлетворяет требованиям, предъявляемым к секвенатору второго поколения, таким как высокая производительность и низкая стоимость.

Принцип: для анализируемого образца, в два конца ПЦР-продуктов экзонов 2, 3, 4 HLA-A/B и экзона 2 HLA-DRB1 с помощью ПЦР-реакции вводили праймерные индексы в качестве специфических меток для получения информации о ПЦР-продуктах данного образца. Продукты ПЦР-амплификации трех сайтов (HLA-A/B/DRB1) в каждой группе образцов объединяли в пул с получением библиотеки ПЦР-продуктов, и после неполного ультразвукового фрагментирования библиотеки ПЦР-продуктов конструировали библиотеку для PCR-FREE-секвенирования. Полученную библиотеку для секвенирования подвергали электрофорезу в 2% агарозном геле с низкой температурой плавления, и все ДНК-полосы, длина которых составляла в пределах от 450 п.н. до 750 п.н., очищали и выделяли путем разрезания геля (поскольку адаптеры библиотеки добавляли к двум концам ДНК-фрагментов, то, во время конструирования библиотеки для PCR-FREE-секвенирования, длина ДНК-полосы, как было показано на электрофореграмме, примерно на 250 п.н. превышала фактическую длину ДНК-фрагментов, а поэтому выделенные здесь фрагменты, длина которых составляла в пределах от 450 п.н. до 700 п.н., фактически соответствовали ДНК-фрагментам, исходная длина которых составляла в пределах от 200 п.н. до 500 п.н.). Выделенную ДНК секвенировали на секвенаторе Illumina GA PE-100. Информация о последовательностях всех тестируемых образцов может быть получена благодаря последовательностям праймерных индексов, а последовательность полноразмерного ПЦР-продукта может быть получена путем сборки исходя из известных эталонных последовательностей и с учетом перекрытия и характера сцепления между последовательностями ДНК-фрагментов. Полноразмерная последовательность исходного ПЦР-продукта может быть получена путем сборки с использованием стандартной базы данных соответствующих экзонов HLA-A/B/DRB1, и последующего HLA-A/B/DRB1-генотипирования.

Пример 1

Экстракция образцов

ДНК экстрагировали из 95 проб крови с использованием известных результатов HLA-SBT-типирования (по Программе взятия костного мозга у доноров в Китае, далее называемой CMDP) на автоматическом устройстве для экстракции KingFisher (US Thermo Co.). При этом проводили следующие основные стадии: как описано в руководстве, определенное количество отдельно взятых агентов добавляли в шесть планшетов с глубокими лунками и в один планшет с неглубокими лунками с подсоединенным автоматическим оборудованием для экстракции KingFisher, и все планшеты, в которые были добавлены указанные агенты, помещали, если это необходимо, в соответствующие позиции. Была выбрана программа «Bioeasy_200ul Blood DNA_KF.msz», после чего запуск экстракции нуклеиновых кислот осуществляли путем нажатия кнопки «start». После завершения программы с планшета для элюирования было собрано

приблизительно 100 мкл элюированных продуктов (то есть экстрагированной ДНК).

Пример 2

ПЦР-амплификация

Различные индексированные ПЦР-праймеры получали путем синтеза ПЦР-праймеров, имеющих различные праймерные индексы у 5'-конца, и такие различные индексированные ПЦР-праймеры могут быть нанесены на различные образцы, где ПЦР-праймерами являются ПЦР-праймеры для экзонов 2, 3, 4 HLA-A/B и экзона 2 HLA-DRB1. Затем праймерные индексы вводили в оба конца ПЦР-продуктов с помощью ПЦР-реакции для специфического мечения ПЦР-продуктов от различных образцов.

Для амплификации 95 образцов ДНК использовали 95 серий индексированных ПЦР-праймеров, соответственно, где каждая серия индексированных ПЦР-праймеров состояла из пары двунаправленных праймерных индексов (таблица 6) и ПЦР-праймеров для амплификации экзонов 2, 3, 4 HLA-A/B (таблица 1) и экзона 2 HLA-DRB1 (таблица 7), причем каждый прямой ПЦР-праймер имел индекс прямого праймера в паре праймерных индексов, присоединенных у 5'-конца, а обратный ПЦР-праймер имел индекс обратного праймера в паре праймерных индексов, присоединенных у 5'-конца. Во время синтеза праймеров эти праймерные индексы были непосредственно добавлены к 5'-концу ПЦР-праймеров.

95 ДНК, полученных после проведения стадии экстракции образцов, описанной в примере 1, обозначали как No. 1-95. ПЦР проводили в 96-луночных планшетах, всего в 7 планшетах, обозначенных HLA-P-A2, HLA-P-A3, HLA-P-A4, HLA-P-B2, HLA-P-B3, HLA-P-B4 и HLA-P-DRB1-2 (A2/A3/A4, B2/B3/B4, DRB1-2 представляют собой амплифицированные сайты), где негативный контроль в каждом планшете был добавлен без какой-либо матрицы, а праймеры, используемые в негативном контроле, были идентичны праймерам, используемым для матрицы 1. Во время проведения эксперимента регистрировали информацию о нумерации образцов, соответствующих каждой паре праймерных индексов.

Необходимая информация о праймерных индексах

№ праймерного индекса	SEQ ID NO:	Индекс прямого праймера	SEQ ID NO:	Индекс обратного праймера	Соответствующее положение в 96-луночном планшете	Соответствующая матрица (группа-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

53

5	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
10	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
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
20	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
25	PI-24	87	ACTCATCGTGCT	88	TAGAGCTGTCAC	B12	24
	PI-25	89	TACACTGTCTAT	90	CAGCACATAGAT	C1	25
	PI-26	91	CACAGTACTCGC	92	CTGCTAGTGTAT	C2	26
	PI-27	93	TGTACTATCATA	94	TGTGATAGACAC	C3	27
30	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
35	PI-32	103	ACACTCATAGAT	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
40	PI-36	111	ACAGTGTAGCGC	112	TGATATGCTACA	C12	36
	PI-37	113	CACTCTATCGAC	114	TCACGCGATGAG	D1	37

5	PI-38	115	ACACTCTAGTCA	116	ACGTAGATCTAT	D2	38
	PI-39	117	CATATGAGATCG	118	AGCAGAGTGCTC	D3	39
	PI-40	119	CAGCTATCATAAC	120	CACTGCAGACGA	D4	40
	PI-41	121	TATACTCTAGAT	122	TGCATAGAGCGC	D5	41
	PI-42	123	TGTATGCTCGTC	124	TCGTGACAGATC	D6	42
10	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
15	PI-47	133	ATAGAGTCTCAT	134	CGCATAGCGTAT	D11	47
	PI-48	135	CGAGACACTCGC	136	TCGTAGTCTACA	D12	48
	PI-49	137	CAGCATACTATC	138	TCGTGATACAGA	E1	49
20	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
25	PI-54	147	ACGACTCGCTAC	148	TCATGACTAGTA	E6	54
	PI-55	149	TAGCATACACGC	150	TGACGCGTATAC	E7	55
	PI-56	151	CGTCATATGCAG	152	TATAGCGATGAC	E8	56
30	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
35	PI-61	161	CATACTGTATCT	162	TACATAGATGTC	F1	61
	PI-62	163	TCTACTCGTGAC	164	CACGTATAGTGA	F2	62
	PI-63	165	CTGCACTAGACA	166	ACTCATATCGCA	F3	63
40	PI-64	167	ACACGAGCTCAT	168	CACTCATATCGA	F4	64
	PI-65	169	TACAGATAGTCT	170	TCGTCTGTGATA	F5	65
	PI-66	171	TACACTCGTGCT	172	TGACGCTCATCT	F6	66
45	PI-67	173	TACATGTGACGA	174	TCGTACATGCTC	F7	67

5

10

15

20

25

30

35

40

PI-68	175	TGTATGATCTCG	176	CACTGTGCTCAT	F8	68
PI-69	177	CAGTACACTCTA	178	ACTGCATGATCG	F9	69
PI-70	179	CATACTATCACG	180	TCGTGTCACCTAC	F10	70
PI-71	181	CACTATACAGAT	182	CGACACGTACTA	F11	71
PI-72	183	ATATCGTAGCAT	184	TCGTGATCACTA	F12	72
PI-73	185	TAGTCTATACAT	186	AGACGCTGTCTGA	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

45

ПЦР-праймеры для амплификации соответствующих экзонов DRB1 и без
праймерных индексов

SEQ ID NO:	№ праймера	Последовательность праймера	Применение праймера	Длина продуктов
231	D2-F1	CACGTTTCTTGGAGTACTCTA	Для амплификации и экзона 2 гена HLA- DRB1	300bp
232	D2-F2	GTTTCTTGTGGCAGCTTAAgTT		
233	D2-F3	CCTGTGGCAGGGTAAGTATA		
234	D2-F4	GTTTCTTGAAGCAGGATAAGTT		
235	D2-F5	GCACGTTTCTTGGAGGAGG		
236	D2-F6	TTTCCTGTGGCAGCCTAAGA		
237	D2-F7	GTTTCTTGGAGCAGGTAAAC		
238	D2-R	CCTCACCTCGCCGCTGCAC		

D2-F1, D2-F2, D2-F3, D2-F4, D2-F5, D2-F6, D2-F7 представляют собой прямые праймеры для амплификации экзона 2 HLA-DRB1, D2-R представляет собой обратный праймер для амплификации экзона 2 HLA-DRB1.

ПЦР для HLA-A/B/DRB1 проводили в следующих условиях:

96°C 2 мин

95°C 30 сек → 60°C 30 сек → 72°C 20 сек (32 цикла)

15°C ∞

ПЦР-реакцию для HLA-A/B проводили на нижеследующей системе, где все реагенты были закуплены у Promega (Beijing) Bio-Tech Co.

Promega 5×буфер I (плюс Mg ²⁺)	5,0 мкл
Смесь dNTP (2,5 мМ/мкл каждого)	2,0 мкл
PI _{inf} -A/B-F _{2/3/4} (2 пмоль/мкл)	1,0 мкл

PI _{inf} -A/B-R _{2/3/4} (2 пмоль/мкл)	1,0 мкл
Promega Taq (5 ед./мкл)	0,2 мкл
ДНК (примерно 20 нг/мкл)	5,0 мкл
ddH ₂ O	10,8 мкл
Всего	25,0 мкл

ПЦР-реакцию для HLA-DRB1 проводили по следующей системе:

Promega 5×буфер I (плюс Mg ²⁺)	5,0 мкл
Смесь dNTP (2,5 мМ/мкл каждого)	2,0 мкл
PI _{inf} -D2-F1 (2 пмоль/мкл)	1,0 мкл
PI _{inf} -D2-F2 (2 пмоль/мкл)	1,0 мкл

PI _{nf} -D2-F3 (2 пмоль/мкл)	1,0 мкл
PI _{nf} -D2-F4 (2 пмоль/мкл)	1,0 мкл
PI _{nf} -D2-F5 (2 пмоль/мкл)	1,0 мкл
PI _{nf} -D2-F6 (2 пмоль/мкл)	1,0 мкл
PI _{nf} -D2-F7 (2 пмоль/мкл)	1,0 мкл
PI _{nf} -D2-R (2 пмоль/мкл)	1,0 мкл
Promega Taq (5 ед./мкл)	0,2 мкл
ДНК (примерно 20 нг/мкл)	5,0 мкл
ddH ₂ O	4,8 мкл
Всего	25,0 мкл

Если PI_{nf}-A/B/D2-F_{1/2/3/4/5/6/7} представляет собой F-праймер HLA-A/B/DRB1, имеющий последовательность индекса прямого праймера No. n (таблица 6) у 5'-конца, то PI_{nf}-A/B/D2-R_{2/3/4} представляет собой R-праймер HLA-A/B/DRB1, имеющий последовательность индекса обратного праймера No. n у 5'-конца (в данном случае n≤95), а остальные праймеры могут быть получены аналогичным образом. Кроме того, каждый образец соответствует конкретной серии ПЦР-праймеров (PI_{nf}-A/B/D2-F_{1/2/3/4/5/6/7}, PI_{nf}-A/B/D2-R_{2/3/4}).

ПЦР-реакцию осуществляли на оборудовании для ПЦР PTC-200, поставляемом Bio-Rad Co. После проведения ПЦР 2 мкл ПЦР-продуктов подвергали электрофорезу в 1% агарозном геле. На фигуре 2 представлены результаты электрофореза ПЦР-продуктов соответствующих экзонов HLA-A/B/DRB1 образца No.1, а маркером молекулярных масс ДНК является DL 2000 (Takara Co.). На электрофореграмме наблюдалась серия одиночных полос длиной в пределах от 300 п.н. до 500 п.н., что указывало на успешную ПЦР-амплификацию экзонов (A2, A3, A4, B2, B3, B4, DRB1-2) HLA-A/B/DRB1 образца No.1. При этом в негативном контроле (N) полоса амплификации не наблюдалась. Результаты для других образцов были аналогичными.

Пример 3

Объединение в пул и очистка ПЦР-продуктов

20 мкл остальных ПЦР-продуктов, взятых из каждой лунки 96-луночного планшета с HLA-P-A2 (за исключением негативного контроля), смешивали со встряхиванием в 3-миллилитровой EP-пробирке до получения гомогенной смеси (обозначенной HLA-A2-Mix). Аналогичную процедуру проводили в других шести 96-луночных планшетах, обозначенных HLA-A3-Mix, HLA-A4-Mix, HLA-B2-Mix, HLA-B3-Mix, HLA-B4-Mix и HLA-D2-Mix. Из каждого планшета брали 200 мкл смеси HLA-A2-Mix, HLA-A3-Mix, HLA-A4-Mix, HLA-B2-Mix, HLA-B3-Mix, HLA-B4-Mix и HLA-D2-Mix и смешивали в 3-миллилитровой EP-пробирке, обозначенной HLA-Mix. 500 мкл ДНК-смеси от HLA-Mix подвергали очистке на колонке с использованием набора для очистки ДНК Qiagen (QIAGEN Co.) (Для подробного описания стадий очистки см. инструкции производителей). На оборудовании Nanodrop 8000 (Thermo Fisher Scientific Co.) было определено, что 200 мкл ДНК, полученной путем очистки, имеет концентрацию ДНК HLA-Mix, равную 48 нг/мкл.

Пример 4

Фрагментирование ПЦР-продуктов и конструирование библиотек для PCR-FREE-секвенирования на Illumina GA

1. Фрагментирование ДНК

Общее количество ДНК, составляющее 5 мкг и взятое из очищенной HLA-Mix, помещали в микропробирку Covaris с AFA-волоком, снабженную защелкивающейся

крышкой, и подвергали фрагментированию на фрагментаторе Covaris S2DNA Shearer (Covaris Co.). Фрагментирование проводили в следующих условиях:

Качание частоты

5	Рабочий цикл	10%
	Интенсивность	5
	Циклы/«Вспышка»	200
	Время (секунды)	300

2. Очистка после фрагментирования

10 Все фрагментированные продукты HLA-Mix выделяли и очищали с помощью набора для ПЦР-очистки QIAquick и растворяли в 37,5 мкл EB (элюирующего буфера QIAGEN), соответственно.

3. Реакция репарации концов

15 После фрагментации очищенную смесь HLA-Mix подвергали реакции репарации концов ДНК в следующей реакционной системе (все реагенты были закуплены у Enzymatics Co.):

	ДНК	37,5 мкл
	H ₂ O	37,5 мкл
	10x буфер с полинуклеотид-киназой (B904)	10 мкл
20	dNTP-смесь (набор растворов (10 мМ каждого))	4 мкл
	ДНК-полимераза T4	5 мкл
	Фрагмент Кленова	1 мкл
	Полинуклеотид-киназа T4	5 мкл
	Общий объем	100 мкл

25 Условия реакции: инкубирование при 20°C в течение 30 минут на термомиксере (Thermomixer, Eppendorf Co.).

Продукты реакции выделяли и очищали с помощью набора для ПЦР-очистки QIAquick и растворяли в 34 мкл EB (элюирующего буфера QIAGEN).

4. Добавление А у 3'-конца

30 А добавляли к 3'-концу ДНК, выделенной в последней стадии, с использованием следующей реакционной системы (все реагенты были закуплены у Enzymatics Co.):

	ДНК, полученная в последней стадии	32 мкл
	10xбуфер Blue	5 мкл
	dATP (1 мМ, GE Co.)	10 мкл
35	Фрагмент Кленова (3'-5' экзо-)	3 мкл
	Общий объем	50 мкл

Условия реакции: инкубирование при 37°C в течение 30 минут на термомиксере (Thermomixer, Eppendorf Co.).

40 Продукты реакции выделяли и очищали с помощью набора для ПЦР-очистки MiniElute (QIAGEN Co.) и растворяли в 13 мкл EB (элюирующего буфера QIAGEN).

5. Лигирование адаптера PCR-FREE-библиотеки на секвенаторе Illumina GA

45 Термин «адаптер PCR-FREE-библиотеки» означает сегмент, сконструированный из оснований, главная роль которых заключается в дополнительной фиксации молекулы ДНК на чипе для секвенирования и в обеспечении сайтов связывания для универсальных секвенирующих праймеров, где адаптер PCR-FREE-библиотеки может быть непосредственно лигирован с двумя концами ДНК-фрагментов в библиотеке для секвенирования, а поскольку при введении адаптера библиотеки ПЦР не проводят, то такой адаптер библиотеки называется здесь «адаптером PCR-FREE-библиотеки».

Продукты, имеющие добавленный А, лигировали с адаптерами PCR-FREE-библиотеки на Illumina GA с использованием следующей реакционной системы (все реагенты были закуплены у Illumina Co.):

ДНК, полученная в последней стадии	11 мкл
2×буфер для быстрого лигирования	15 мкл
Олиго-смесь PCR-FREE-адаптеров (30 мМ)	1 мкл
ДНК-лигаза T4 (Rapid, L603-NC-L)	3 мкл
Общий объем	30 мкл

Условия реакции: инкубирование при 20°C в течение 15 минут на термомиксере (Thermomixer, Eppendorf Co.).

Продукты реакции очищали на сферах Ampure (Beckman Coulter Genomics) и растворяли в 50 мкл деионизованной воды, а затем определяли концентрацию ДНК с помощью количественной ПЦР с активацией флуоресценции (QPCR) следующим образом:

15

	Результат, полученный с помощью кол. ПЦР (нМ)
HLA-Mix	78,90

6. Выделение путем разрезания геля

30 мкл HLA-Mix подвергали электрофорезу в 2% агарозном геле с низкой температурой плавления. Электрофорез проводили при 100 В, 100 минут. ДНК-маркер представлял собой 50 п.н. ДНК-маркера от NEB Co. Гель, содержащий ДНК-фрагменты размером от 450 до 750 п.н., разрезали (фигура 3). Продукты, выделенные из разрезанного геля, очищали с помощью набора для ПЦР-очистки QIAquick (QIAGEN Co.), и после такой очистки объем составлял 32 мкл, а концентрация ДНК, измеренная с помощью количественной ПЦР с активацией флуоресценции (QPCR), составляла 10,16 нМ.

Пример 5

Секвенирование на Illumina GA

В соответствии с результатами QPCR брали 10 пмоль ДНК и подвергали секвенированию по программе Illumina GA PE-100. Более подробное описание процедуры секвенирования см. в инструкции по секвенированию на Illumina GA (Illumina GA IIx).

Пример 6

Анализ результатов

Результаты секвенирования на Illumina GA представляют собой серию последовательностей ДНК, и после поиска последовательностей индексов прямого и обратного праймеров и последовательностей праймеров в результатах секвенирования создавали базы данных, содержащие результаты секвенирования ПЦР-продуктов различных экзонов HLA-A/B/DRB1 для каждого образца, соответствующего данному праймерному индексу. Результаты секвенирования каждого экзона сопоставляли путем выравнивания с эталонной последовательностью (эталонные последовательности можно найти на сайте <http://www.ebi.ac.uk/imgt/hla/>) соответствующего экзона с помощью программы BWA (Burrows-Wheeler Aligner), и таким образом конструировали консенсусные последовательности для каждой базы данных, а затем в этой базе данных отбирали последовательности ДНК и корректировали. Скорректированные последовательности ДНК получали путем их сборки в соответствующие последовательности экзонов HLA-A/B/DRB1 исходя из перекрывания последовательностей и характера сцепления (парно-концевого сцепления). Полученную

последовательность ДНК сопоставляли путем выравнивания с последовательностью базы данных, соответствующей экзону HLA-A/B/DRB1 в профессиональной базе данных HLA IMGT. Если результат выравнивания последовательностей дает 100% соответствие, то это означает, что HLA-A/B/DRB1-генотип соответствующего образца был точно
5 определен. См. программу с захватом изображения на экране для конструирования консенсусной последовательности экзона 2 HLA-A-сайта образца No.1, как показано на фигуре 4.

Для всех 95 образцов результаты типирования, полученные вышеописанным методом, полностью соответствовали известным результатам типирования, где результаты для
10 образцов No.1-32 представлены ниже:

15

20

25

30

35

40

45

Образец No.

Исходный HLA-A/B/DRB1 известного типа

5	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
10	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
15	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
20	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
25	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
30	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
35	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
40	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

45

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
5 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
10 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

Образец No.

Исходный HLA-A/B/DRB1 определенного типа

15 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
20 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
25 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
30 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
35 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
40 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

45

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
5 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
10 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
15 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
20 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

Примечание: в случае HLA типа HLA-DRB1, не исключено, что DRB1*1201 может представлять собой DRB1*1206/1210/1217, а DRB1*1454 может представлять собой DRB1*1401, поскольку указанные аллели полностью идентичны в последовательности экзона 2 гена HLA-DRB1.

Пример 7

Генотипирование HLA-A,B и DRB1 путем секвенирования на секвенаторе второго поколения (Illumina GA)

Экстракция образцов

ДНК экстрагировали из 950 проб крови с использованием известных результатов HLA-SBT-типирования (по Программе взятия костного мозга у доноров в Китае, далее называемой CMDP) на автоматическом оборудовании для экстракции KingFisher (US Thermo Co.). Этот метод описан в примере 1.

ПЦР-амплификация

950 ДНК, полученных после проведения каждой стадии экстракции образцов, были обозначены No.1-950, и были подразделены на 10 групп (95 ДНК для каждой группы), которые были обозначены как HLA-1, HLA-2, HLA-3, HLA-4, HLA-5, HLA-6, HLA-7, HLA-8, HLA-9, HLA-10. Для каждой группы образцов 95 образцов ДНК амплифицировали с использованием 95 наборов ПЦР-праймеров (таблица 1), несущих двунаправленные праймерные индексы (таблица 6), для амплификации экзонов 2, 3, 4 HLA-A/B и ПЦР-праймеров (таблица 7), несущих двунаправленные праймерные индексы (таблица 6), для амплификации экзона 2 HLA-DRB1. ПЦР-реакцию проводили в 96-луночных планшетах с использованием всего 70 планшетов, обозначенных HLA-X-P-A2, HLA-X-P-A3, HLA-X-P-A4, HLA-X-P-B2, HLA-X-P-B3, HLA-X-P-B4 и HLA-X-P-DRB1-2 («X» указывает на номер группы 1/2/3/4/5/6/7/8/9/10, а «A2/3/4», «B2/3/4», «DRB1-2» представляют собой сайты амплификации), где в каждом планшете имеется негативный контроль без какой-либо матрицы, а праймеры, используемые для негативного контроля, представляют собой праймеры, помеченные PI-1 (таблица 6). Во время эксперимента регистрировали номер каждого образца в группе и информацию о праймерных индексах.

Этот способ описан в примере 2.

Объединение в пул и очистка ПЦР-продуктов

Для образцов группы X («X» представляет собой 1/2/3/4/5/6/7/8/9/10), 20 мкл остальных ПЦР-продуктов, взятых из каждой лунки 96-луночного планшета с HLA-X-P-A2 (за исключением негативного контроля), смешивали со встряхиванием в 3-миллилитровой ЕР-пробирке до получения гомогенной смеси (обозначенной HLA-X-A2-Mix).

Аналогичную процедуру проводили в других шести 96-луночных планшетах с образцами группы X, обозначенных HLA-X-A3-Mix, HLA-X-A4-Mix, HLA-X-B2-Mix, HLA-X-B3-Mix, HLA-X-B4-Mix и HLA-X-D2-Mix. Из каждого планшета брали 200 мкл смеси HLA-X-A2-Mix, HLA-X-A3-Mix, HLA-X-A4-Mix, HLA-X-B2-Mix, HLA-X-B3-Mix, HLA-X-B4-Mix и HLA-X-D2-Mix, и смешивали в 3-миллилитровой ЕР-пробирке, обозначенной HLA-X-Mix. 500 мкл ДНК-смеси от HLA-X-Mix подвергали очистке на колонке с использованием набора для очистки ДНК Qiagen (QIAGEN Co.) (Для подробного описания стадий очистки см. инструкции производителей) с получением 200 мкл ДНК, концентрацию которой определяли на оборудовании Nanodrop 8000 (Thermo Fisher Scientific Co.). Ту же самую процедуру проводили и для других групп. Конечные определенные концентрации ДНК указаны ниже.

	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
Концентр. (нг/мкл)	48,0	52,1	49,3	50,2	47,6	48,5	49,1	48,6	51,3	50,8

Этот метод описан в примере 3.

Конструирование библиотек для секвенирования осуществляли на Illumina GA методом, описанным в примере 4. Ниже показана соответствующая взаимосвязь между группами образцов и адаптерами библиотеки.

Образец группы No.	HLA-1	HLA-2	HLA-3	HLA-4	HLA-5	HLA-6	HLA-7	HLA-8	HLA-9	HLA-10
Адаптер библиотечки No.	1	2	3	4	5	6	7	8	9	10

Продукты реакции очищали на сферах Ampure (Beckman Coulter Genomics) и растворяли в 50 мкл деионизованной воды, а затем определяли молярные концентрации ДНК с помощью количественной ПЦР с активацией флуоресценции (QPCR) следующим образом:

	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
Конц. (нМ)	78.90	72.13	79.33	80.21	77.68	78.50	89.12	78.60	81.32	80.82

Выделение путем разрезания геля

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 и HLA-10-Mix смешивали в равных молярных концентрациях (конечная концентрация составляла 72,13 нМ/мкл) и обозначали HLA-Mix-10. 30 мкл HLA-Mix-10 подвергали электрофорезу в 2% агарозном геле с низкой температурой плавления. Электрофорез проводили при 100 В, 100 минут. ДНК-маркер представлял собой 50 п.н. ДНК-маркера от NEB Co. Гель, содержащий ДНК-фрагменты размером от 450 до 750 п.н., разрезали. Продукты, выделенные из разрезанного геля, очищали с помощью набора для ПЦР-очистки QIAquick (QIAGEN Co.), и после такой очистки

объем составлял 32 мкл, а концентрация ДНК, измеренная с помощью количественной ПЦР с активацией флуоресценции (QPCR) составляла 9,96 нМ.

Секвенирование и анализ результатов осуществляли, как описано в примерах 5 и 6. Для всех 950 образцов результаты типирования, полученные вышеописанным методом, полностью соответствовали известным результатам типирования.

Пример 8

HLA-C-генотипирование с применением метода секвенирования на секвенаторе второго поколения (Illumina GA)

1. Экстракция ДНК-образцов

Стадии проводили, как описано в примере 1.

2. ПЦР-амплификация

Стадии проводили, как описано в примере 2, за исключением того, что используемыми ПЦР-праймерами являются ПЦР-праймеры для экзонов 2, 3 и 4 HLA-C, как показано в таблице 3.

Для амплификации 95 образцов ДНК использовали 95 серий индексированных ПЦР-праймеров, соответственно, где каждая серия индексированных ПЦР-праймеров состояла из ПЦР-праймеров для амплификации экзонов 2, 3, 4 HLA-C (таблица 3) и пары двунаправленных праймерных индексов (как описано ниже), где каждый прямой ПЦР-праймер имел индекс прямого праймера в паре праймерных индексов, присоединенных у 5'-конца, а обратный ПЦР-праймер имел индекс обратного праймера в паре праймерных индексов, присоединенных у 5'-конца. Во время синтеза праймеров, эти праймерные индексы были непосредственно добавлены к 5'-концу ПЦР-праймеров.

95 ДНК, полученных после проведения стадии экстракции образцов, обозначали No. 1-95. ПЦР проводили в 96-луночных планшетах, всего в 3 планшетах, обозначенных HLA-P-C2, HLA-P-C3, HLA-P-A4 (C2/3/4 представляют собой сайты амплификации), где негативный контроль в каждом планшете был добавлен без какой-либо матрицы, а праймеры, используемые в негативном контроле, были идентичны праймерам PI-96. Во время проведения эксперимента регистрировали информацию о нумерации образцов, соответствующих каждой паре праймерных индексов.

Используемыми праймерными индексами являются праймерные индексы PI-1 - PI-95, перечисленные в таблице 6, и указанный ниже праймерный индекс PI-96, используемый в качестве негативного контроля (таблица 8)

Таблица 8

Необходимая информация о праймерных индексах, используемых для

негативного контроля

PI	SEQ		SEQ			Негатив-
- 96	ID	CACTGTATAGCT	ID	CGACTAGTACTA	H12	ный
	NO: 239		NO: 240			контроль

ДНК, экстрагированные на автоматическом оборудовании для экстракции KingFisher в стадии 1, использовали в качестве матриц и ПЦР-амплификацию осуществляли в отдельных пробирках с использованием праймеров для экзонов HLA-C, где праймеры были индексированы у 5'-конца. ПЦР проводили следующим образом:

C2: 96°C, 2 мин

95°C 30 сек → 62°C 30 сек → 72°C 20 сек (35 циклов)

15°C ∞

C3: 96°C 2 мин

95°C 30 сек → 56°C 30 сек → 72°C 20 сек (35 циклов)

15°C ∞

C4: 96 °C 2 мин

95°C 30 сек → 60°C 30сек → 72°C 20 сек (35 циклов)

15°C ∞

ПЦР-реакцию для HLA-C проводили на нижеследующей системе:

Promega 5xбуфер I (плюс Mg^{2+})	5,0 мкл
Смесь dNTP (2,5 мМ/мкл каждого)	2,0 мкл
PI _{inf} -C-F2/3/4 (50 нг/мкл)	1,5 мкл
PI _{inf} -C-R2/3/4 (50 нг/мкл)	1,5 мкл
Promega Taq (5 ед./мкл)	0,2 мкл
ДНК (примерно 20 нг/мкл)	2,0 мкл
ddH ₂ O	12,8 мкл
Всего	25,0 мкл

Если PI_{inf}-C-F_{2/3/4} представляет собой F-праймер HLA-C, имеющий последовательность индекса прямого праймера No. n (таблица 2) у 5'-конца, то PI_{inf}-C-R_{2/3/4} представляет собой R-праймер HLA-C, имеющий последовательность индекса обратного праймера No. n у 5'-конца (в данном случае $n \leq 96$), а остальные праймеры могут быть получены аналогичным образом. Кроме того, каждый образец соответствует конкретной серии ПЦР-праймеров.

ПЦР-реакцию осуществляли на оборудовании для ПЦР PTC-200, поставляемом Bio-Rad Co. После проведения ПЦР 2 мкл ПЦР-продуктов подвергали электрофорезу в 1,5% агарозном геле. На фигуре 6 представлены результаты электрофореза ПЦР-продуктов соответствующих экзонов HLA-C первых 20 образцов, а маркером молекулярных масс ДНК является DL 2000 (Takara Co.). На электрофореграмме наблюдалась серия одиночных полос длиной в пределах от 400 п.н. до 500 п.н., что указывало на успешную ПЦР-амплификацию экзонов (C2, C3, C4) HLA-C-образцов. Результаты для других образцов были аналогичными.

Объединение в пул и очистка ПЦР-продуктов

20 мкл остальных ПЦР-продуктов, взятых из каждой лунки 96-луночного планшета с HLA-P-C2 (за исключением негативного контроля), объединяли в гомогенный пул со встряхиванием в 3-миллилитровой EP-пробирке (обозначенный HLA-C2-Mix).

Аналогичную процедуру проводили в других двух 96-луночных планшетах, обозначенных HLA-C3-Mix и HLA-C4-Mix. Из каждого планшета брали 200 мкл каждого из HLA-C2-Mix, HLA-C3-Mix и HLA-C4-Mix, и смешивали в 1,5-миллилитровой EP-пробирке, обозначенной HLA-Mix. 500 мкл ДНК-смеси от HLA-Mix подвергали очистке на колонке с использованием набора для очистки ДНК Qiagen (QIAGEN Co.) (Для подробного описания стадий очистки см. инструкции производителей). На оборудовании Nanodrop 8000 (Thermo Fisher Scientific Co.) было определено, что 200 мкл ДНК, полученной путем очистки, имеет концентрацию ДНК HLA-Mix, равную 50 нг/мкл.

4. Конструирование библиотек для PCR-FREE-секвенирования на Illumina GA

4.1. Фрагментирование ПЦР-продуктов

Общее количество ДНК, составляющее 5 мкг и взятое из очищенной HLA-Mix, помещали в микропробирку Covaris с AFA-волоком, снабженную защелкивающейся крышкой, и подвергали фрагментированию на фрагментаторе Covaris S2 (Covaris Co.). Фрагментирование проводили в следующих условиях:

Качание частоты

Рабочий цикл	10%
Интенсивность	3
Циклы/«Вспышка»	200
Время (секунды)	180

4.2. Очистка фрагментированных ПЦР-продуктов

Все фрагментированные продукты HLA-Mix выделяли и очищали с помощью набора для ПЦР-очистки QIAquick и растворяли в 37,5 мкл EB (элюирующего буфера QIAGEN), соответственно.

4.3. Реакция репарации концов

Очищенные продукты подвергали реакции репарации концов ДНК в следующей реакционной системе (все реагенты были закуплены у Enzymatics Co.):

Продукты, очищенные в последней стадии	37,5 мкл
10× буфер с полинуклеотид-киназой (B904)	5 мкл
dNTP-смесь (Набор растворов (10 мМ каждого))	2 мкл
ДНК-полимераза T4	2,5 мкл
Фрагмент Кленова	0,5 мкл
Полинуклеотид-киназа T4	2,5 мкл
Общий объем	50 мкл

Условия реакции: инкубирование при 20°C в течение 30 минут на термомиксере (Thermomixer, Eppendorf Co.).

Продукты реакции выделяли и очищали с помощью набора для ПЦР-очистки QIAquick и растворяли в 32 мкл EB (элюирующего буфера QIAGEN).

4.4. Добавление А у 3'-конца

А добавляли к 3'-концу ДНК, выделенной в последней стадии, с использованием следующей реакционной системы (все реагенты были закуплены у Enzymatics Co.):

ДНК, полученная в последней стадии	32 мкл
10× буфер Blue	5 мкл
dATP (1 мМ GE Co.)	10 мкл
Фрагмент Кленова (3'-5'-экзо)	3 мкл
Общий объем	50 мкл

Условия реакции: инкубирование при 37°C в течение 30 минут в термомиксере (Thermomixer, Eppendorf Co.).

Продукты реакции выделяли и очищали с помощью набора для ПЦР-очистки MiniElute (QIAGEN Co.) и растворяли в 38 мкл EB (элюирующего буфера QIAGEN).

4.5. Лигирование адаптера PCR-FREE-библиотеки на секвенаторе Illumina GA

Продукты, имеющие добавленный А, лигировали с адаптерами PCR-FREE-библиотеки на Illumina GA с использованием следующей реакционной системы (все реагенты были закуплены у Illumina Co.):

ДНК, полученная в последней стадии	38 мкл
10× буфер для лигирования	5 мкл
Олиго-смесь PCR-FREE-адаптеров (30 мМ)	2 мкл
ДНК-лигаза T4 (Rapid, L603-NC-L)	5 мкл
Общий объем	50 мкл

Условия реакции: инкубирование при 16°C в течение ночи в термомиксере (Thermomixer, Eppendorf Co.).

Продукты реакции очищали на сферах Ampure (Beckman Coulter Genomics) и растворяли в 50 мкл деионизированной воды, а затем определяли концентрацию ДНК с помощью количественной ПЦР с активацией флуоресценции (QPCR) следующим

образом:

	Результат, полученный с помощью кол. ПЦР (нМ)
HLA-Mix	122,71

4.6. Выделение путем разрезания геля

30 мкл HLA-Mix подвергали электрофорезу в 2% агарозном геле с низкой температурой плавления. Электрофорез проводили при 100 В, 100 минут. ДНК-маркер представлял собой 50 п.н. ДНК-лэддера от NEB Co. Гель, содержащий ДНК-фрагменты размером от 450 до 750 п.н., разрезали (фигура 7). Продукты, выделенные из разрезанного геля, очищали с помощью набора для ПЦР-очистки QIAquick (QIAGEN Co.), и после такой очистки объем составлял 32 мкл, а концентрация ДНК, измеренная с помощью количественной ПЦР с активацией флуоресценции (QPCR), составляла 17,16 нМ.

5. Секвенирование на Illumina GA

В соответствии с результатами QPCR-детектирования брали 10 пмоль ДНК и подвергали секвенированию по программе Illumina GA PE-100. Более подробное описание процедуры секвенирования см. в инструкции по секвенированию на Illumina GA (Illumina GA IIx).

6. Анализ результатов

Результаты секвенирования, полученные на Illumina GA, представляют собой серию последовательностей ДНК, и после поиска последовательностей индексов прямого и обратного праймеров и последовательностей праймеров в этих результатах секвенирования создавали базы данных, содержащие результаты секвенирования ПЦР-продуктов различных экзонов HLA-C для каждого образца, соответствующего данному праймерному индексу. Результаты секвенирования каждого экзона сопоставляли путем выравнивания с эталонной последовательностью (эталонные последовательности можно найти на сайте <http://www.ebi.ac.uk/imgt/hla/>) соответствующего экзона с помощью программы BWA (Burrows-Wheeler Aligner), и таким образом конструировали консенсусные последовательности для каждой базы данных, а затем в этой базе данных отбирали последовательности «риды» и корректировали исходя из качественной оценки секвенирования оснований и различия между последовательностями-«ридами» и консенсусными последовательностями. Скорректированные последовательности ДНК получали путем сборки в соответствующие последовательности экзонов HLA-C исходя из перекрывания последовательностей и характера сцепления (парно-концевого сцепления). Процедура конструирования консенсусных последовательностей экзона 2 HLA-C-сайта в образце № 2, проводимая с помощью программы с захватом изображения на экране, проиллюстрирована на фигуре 8.

Полученную ДНК-последовательность сопоставляли с последовательностями базы данных для соответствующего экзона HLA-C, имеющегося в профессиональной базе данных HLA IMGT. Если результат выравнивания последовательностей давал 100% соответствие, то это означает, что HLA-C-генотип соответствующего образца был точно определен. Для всех 95 образцов результаты типирования, полученные вышеописанным методом, полностью соответствовали известным результатам типирования, где результаты типирования для образцов No.1-32 представлены ниже (как видно из таблицы 9, все полученные результаты были идентичны исходным известным результатам):

Таблица 9

Сравнение результатов типирования образцов, полученных вышеописанным методом,
с результатами типирования образцов, полученными известным методом

5

Образец №		Исходный известный HLA-C-генотип		Результаты для HLA-C, полученные на данный момент времени		Идентичный или нет
10	1	C*08:01	C*15:05	C*08:01	C*15:05	да
	2	C*01:02	C*07:02	C*01:02	C*07:02	да
	3	C*08:01	C*16:02	C*08:01	C*16:02	да
	4	C*01:02	C*03:02	C*01:02	C*03:02	да
15	5	C*01:02	C*02:02	C*01:02	C*02:02	да
	6	C*01:02	C*15:02	C*01:02	C*15:02	да
	7	C*01:02	C*03:04	C*01:02	C*03:04	да
	8	C*03:02	C*07:02	C*03:02	C*07:02	да
20	9	C*06:02	C*16:02	C*06:02	C*16:02	да
	10	C*01:02	C*03:04	C*01:02	C*03:04	да
	11	C*03:04	C*07:02	C*03:04	C*07:02	да
	12	C*07:02	C*08:01	C*07:02	C*08:01	да
25	13	C*01:02	C*15:02	C*01:02	C*15:02	да
	14	C*01:02	C*03:04	C*01:02	C*03:04	да
	15	C*01:02	C*03:04	C*01:02	C*03:04	да
	16	C*07:02	C*12:02	C*07:02	C*12:02	да
30	17	C*04:01	C*08:01	C*04:01	C*08:01	да
	18	C*08:01	C*16:02	C*08:01	C*16:02	да
	19	C*14:02	C*15:02	C*14:02	C*15:02	да
	20	C*01:02	C*03:03	C*01:02	C*03:03	да
35	21	C*03:03	C*08:01	C*03:03	C*08:01	да
	22	C*03:04	C*07:02	C*03:04	C*07:02	да
	23	C*07:02	C*08:01	C*07:02	C*08:01	да
	24	C*07:02	C*12:02	C*07:02	C*12:02	да
40	25	C*07:02	C*12:03	C*07:02	C*12:03	да
	26	C*03:04	C*08:01	C*03:04	C*08:01	да
	27	C*01:02	C*03:04	C*01:02	C*03:04	да
	27	C*01:02	C*03:04	C*01:02	C*03:04	да

28	C*07:02	C*12:02	C*07:02	C*12:02	да
29	C*03:02	C*07:02	C*03:02	C*07:02	да
30	C*01:02	C*03:03	C*01:02	C*03:03	да
31	C*01:02	C*07:02	C*01:02	C*07:02	да
32	C*01:02	C*07:02	C*01:02	C*07:02	да

10 Примечание: в случае HLA типа HLA-C, не исключено, что C*0303 может представлять собой C*0320N, C*0401 может представлять собой C*0409N/0430, C*0702 может представлять собой C*0750, C*0801 может представлять собой C*0822, а C*1505 может представлять собой C*1529, поскольку указанные аллели полностью идентичны в последовательностях экзонов 2, 3, 4 HLA-C.

Пример 9

HLA-C-генотипирование методом секвенирования по Сэнгеру

15 1. Экстракция образцов ДНК

Как описано в примере 1, на автоматическом оборудовании для экстракции KingFisher экстрагировали 26 из 95 образцов ДНК с известными HLA-генотипами.

2. ПЦР-амплификация

20 Вышеуказанные ДНК, экстрагированные на автоматическом оборудовании для экстракции KingFisher, использовали в качестве матриц и ПЦР-амплификацию осуществляли в отдельных пробирках с использованием трех пар ПЦР-праймеров C-F2/C-R2, C-F3/C-R3, C-F4/C-R4 (таблица 3), соответственно. Для каждой пары праймеров проводили ПЦР в следующих условиях:

C2: 96°C, 2 мин

25 95°C 30 сек → 62°C 30 сек → 72°C 20 сек (35 циклов)

15°C ∞

C3: 96°C 2 мин

95°C 30 сек → 56°C 30 сек → 72°C 20 сек (35 циклов)

15°C ∞

30 C4: 96°C 2 мин

95°C 30 сек → 60°C 30 сек → 72°C 20 сек (35 циклов)

15°C ∞

ПЦР-реакцию для HLA-C проводили на нижеследующей системе:

35	Promega 5xбуфер I (плюс Mg ²⁺)	5,0 мкл
	Смесь dNTP (2,5 мМ каждого)	2,0 мкл
	Смесь праймеров (50 нг/мкл)	3,0 мкл
	Promega Taq (5 ед./мкл)	0,2 мкл
	ДНК (примерно 20 нг/мкл)	2,0 мкл
	ddH ₂ O	12,8 мкл
40	Всего	25,0 мкл

Перед очисткой ПЦР-продукты подвергали электрофорезу в агарозном геле (фигура 9).

3. Очистка ПЦР-продуктов

45 ПЦР-продукты очищали на планшетах для очистки Millipore. Были проведены следующие основные стадии. Используемые лунки 96-луночного планшета для очистки ПЦР-продуктов помечали маркером и в каждую используемую лунку добавляли 50 мкл сверхчистой воды. Остальные лунки герметично заклеивали пленкой-герметиком. Планшет оставляли на 15 минут или подсоединяли на 5 минут к дренажной и

фильтрующей системе (~10 Па). При отсоединении планшета для очистки от дренажной и фильтрующей системы, жидкость, находящуюся в съемном порте на дне планшета для очистки, удаляли с помощью впитывающей бумаги.

Предназначенные для очистки ПЦР-продукты центрифугировали в течение 1 минуты при 4000 об/мин, покрытие или слой силикагеля для очищаемых ПЦР-продуктов удаляли и в каждую систему для ПЦР-реакции добавляли 100 мкл сверхчистой воды. Затем планшеты для очистки, в который были добавлены очищаемые ПЦР-продукты, подсоединяли к дренажной и фильтрующей системе и уровень вакуума доводили до ~10 Па, на что указывал барометр. Дренаж и фильтрование продолжали до тех пор, пока с микропористой регенерируемой целлюлозной пленки, находящейся на дне планшета для очистки, не исчезала жидкость и не пропадал блеск, наблюдаемый в результате отражения света от интактной поверхности жидкости.

В лунках, содержащих очищаемые ПЦР-продукты, в микропористую регенерируемую целлюлозную пленку добавляли 50 мкл сверхчистой воды или ТЕ, а затем планшеты для очистки подвергали встряхиванию при комнатной температуре в шейкере-самописце в течение 5 минут и всю жидкость, содержащуюся в соответствующих лунках, переносили в соответствующие лунки нового 96-луночного ПЦР-планшета.

4. Осуществление реакции секвенирования и очистка продуктов реакции секвенирования

ПЦР-продукты, очищенные, как описано выше, использовали в качестве матриц для реакции секвенирования.

Условия проведения реакции секвенирования:

96°C 2 мин

96°C 10 сек → 55°C 5 сек → 60°C 2 мин (25 циклов)

15°C ∞

Система для реакции секвенирования

Очищенные ПЦР-продукты	1 мкл
Праймеры (3,2 пмоль/л)	1 мкл
2,5 *BigDye	0,3 мкл
5* буфер	0,85 мкл
Вода	1,85 мкл
Общий объем	5 мкл

Продукты реакции секвенирования очищали путем проведения следующих стадий: планшеты для реакции секвенирования уравнивали и центрифугировали при 3000×g в течение 1 минуты. В 96-луночном планшете к каждой реакционной 5-мкл системе добавляли 2 мкл 0,125 моль/л EDTA-Na₂-раствора, 33 мкл 85% этанола и планшеты покрывали слоем силикагеля, а затем достаточно интенсивно встряхивали в течение 3 минут. Затем планшеты центрифугировали при 4°C, 3000×g в течение 30 минут. После центрифугирования планшеты для секвенирования вынимали, слой силикагеля удаляли и планшеты для секвенирования помещали верхом вниз на впитывающую бумагу, после чего подвергали инверсионному центрифугированию до тех пор, пока центрифужная сила не достигала 185×g. Затем в каждую лунку 96-луночного планшета добавляли 50 мкл 70% этанола. Планшеты покрывали слоем силикагеля, встряхивали в течение 1,5 минут и центрифугировали при 4°C, 3000×g в течение 15 минут. Затем планшеты для реакции секвенирования помещали на 30 минут в темное и проветриваемое помещение и сушили воздухом до тех пор, пока не исчезал запах этанола. В каждую лунку 96-луночного планшета добавляли 10 мкл HI-DI-формамида (альтернативно, в каждую лунку 384-луночного планшета добавляли 8 мкл HI-DI-формамида), а затем

этот планшет закрывали пленкой-герметиком и после встряхивания в течение 5 секунд центрифугировали при 1000 об/мин.

5. Секвенирование и анализ результатов

Очищенные продукты реакции секвенирования подвергали капиллярному электрофорезу в секвенирующем геле ABI 3730XL. Пики секвенирования анализировали с помощью компьютерной программы uType (Invitrogen) и получали результаты HLA-типирования (фигура 10). Все результаты, полученные вышеописанным методом, были идентичны первоначальным известным результатам, как показано в таблице 10.

Таблица 10

Сравнение результатов типирования, полученных вышеописанным методом, с результатами типирования, полученными известным методом

Образец №	Исходный известный HLA-C-генотип	Результаты для HLA-C, полученные на данный момент времени	Идентичный или нет
1	C*08:01 C*15:05	C*08:01 C*15:05	да
2	C*01:02 C*07:02	C*01:02 C*07:02	да
3	C*08:01 C*16:02	C*08:01 C*16:02	да
4	C*01:02 C*03:02	C*01:02 C*03:02	да
5	C*01:02 C*02:02	C*01:02 C*02:02	да
6	C*01:02 C*15:02	C*01:02 C*15:02	да
7	C*01:02 C*03:04	C*01:02 C*03:04	да
8	C*03:02 C*07:02	C*03:02 C*07:02	да
9	C*06:02 C*16:02	C*06:02 C*16:02	да
10	C*01:02 C*03:04	C*01:02 C*03:04	да
11	C*03:04 C*07:02	C*03:04 C*07:02	да
12	C*07:02 C*08:01	C*07:02 C*08:01	да
13	C*01:02 C*15:02	C*01:02 C*15:02	да
14	C*01:02 C*03:04	C*01:02 C*03:04	да
15	C*01:02 C*03:04	C*01:02 C*03:04	да
16	C*07:02 C*12:02	C*07:02 C*12:02	да
17	C*04:01 C*08:01	C*04:01 C*08:01	да
18	C*08:01 C*16:02	C*08:01 C*16:02	да
19	C*14:02 C*15:02	C*14:02 C*15:02	да
20	C*01:02 C*03:03	C*01:02 C*03:03	да
21	C*03:03 C*08:01	C*03:03 C*08:01	да
22	C*03:04 C*07:02	C*03:04 C*07:02	да

23	C*07:02	C*08:01	C*07:02	C*08:01	да
24	C*07:02	C*12:02	C*07:02	C*12:02	да
25	C*07:02	C*12:03	C*07:02	C*12:03	да
26	C*01:02	C*07:02	C*01:02	C*07:02	да

Пример 10: HLA-DQB1-генотипирование методом секвенирования второго поколения (Illumina Solexa)

94 пробы крови с известными результатами HLA-SBT-типирования подвергали HLA-DQB1-генотипированию методами, описанными в примере 8, за исключением нижеследующих изменений, представленных ниже.

Для амплификации 94 образцов ДНК использовали 94 серии индексированных ПЦР-праймеров, соответственно, где каждая серия индексированных ПЦР-праймеров состояла из ПЦР-праймеров для амплификации экзонов 2 или 3 HLA-DQB1 (таблица 5) и пары двунаправленных праймерных индексов (как описано выше), где каждый прямой ПЦР-праймер имел индекс прямого праймера в паре праймерных индексов, присоединенных у 5'-конца, а обратный ПЦР-праймер имел индекс обратного праймера в паре праймерных индексов, присоединенных у 5'-конца. Во время синтеза праймеров, эти праймерные индексы были непосредственно добавлены к 5'-концу ПЦР-праймеров, где указанные праймеры были синтезированы фирмой Shanghai Invitrogen Co.

94 ДНК, полученных после проведения стадии экстракции образцов, обозначали No.1-94. ПЦР проводили в 96-луночных планшетах, где в одной и той же лунке были амплифицированы экзоны 2, 3 DQB1 в каждом образце. Два негативных контроля в каждом планшете были добавлены без какой-либо матрицы, а праймерными индексами, используемыми в качестве негативного контроля, являются PI-95 и PI-96. Во время проведения эксперимента регистрировали данные о нумерации образцов, соответствующих каждой паре праймерных индексов.

В качестве праймерных индексов использовали праймерные индексы PI-1 - PI-94, перечисленные в таблице 6, а указанные ниже праймерные индексы PI-95 и PI-96 (таблица 11) были использованы для негативного контроля.

Таблица 11

Необходимая информация о праймерных индексах, используемых для негативного контроля

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

ПЦР-процедура для HLA-DQB1:

96°C 2 мин

95°C 30 сек → 60°C 30 сек → 72°C 20 сек (32 цикла)

15°C ∞

ПЦР-реакцию для HLA-DQB1 проводили на нижеследующей системе:

	Promega 5×буфер I (плюс Mg^{2+})	5,0 мкл
5	Смесь dNTP (2,5 мМ каждого)	2,0 мкл
	PInf-Q-F2 (2 пмоль/мкл)	1,0 мкл
	PInf-Q-R2 (2 пмоль/мкл)	1,0 мкл
10	PInf-Q-F3 (2 пмоль/мкл)	1,0 мкл
	PInf-Q-R3 (2 пмоль/мкл)	1,0 мкл
	Promega Taq (5 ед./мкл)	0,2 мкл
15	ДНК (примерно 20 нг/мкл)	5,0 мкл
	ddH ₂ O	8,8 мкл
20	Всего	25,0 мкл

Если PInf-Q-F2/3 представляет собой F-праймер HLA-DQB1, имеющий последовательность индекса прямого праймера No. n

(таблица 1) у 5'-конца, то PInf-Q-R2/3 представляет собой R-праймер HLA-DQB1, имеющий последовательность индекса обратного праймера No. n у 5'-конца (в данном случае $n \leq 96$), а остальные праймеры могут быть получены аналогичным образом. Кроме того, каждый образец соответствует конкретной серии ПЦР-праймеров.

ПЦР-реакцию осуществляли на оборудовании для ПЦР PTC-200, поставляемом Bio-Rad Co. После проведения ПЦР 2 мкл ПЦР-продуктов подвергали электрофорезу в 1,5% агарозном геле. На фигуре 11 представлены результаты электрофореза ПЦР-продуктов экзонов 2+3 HLA-DQB1 для 94 образцов, а маркером молекулярных масс ДНК является DL 2000 (Takara Co.).

Объединение в пул и очистка ПЦР-продуктов

20 мкл остальных ПЦР-продуктов, взятых из каждой лунки 96-луночного планшета с HLA-P-DQB1 (за исключением негативного контроля), смешивали до гомогенности в пул в 3-миллилитровой EP-пробирке (обозначенной HLA-Q-Mix). 500 мкл ДНК-смеси от HLA-Q-Mix подвергали очистке на колонке с использованием набора для очистки ДНК Qiagen (QIAGEN Co.) (Для подробного описания стадий очистки см. инструкции производителей). На оборудовании Nanodrop 8000 (Thermo Fisher Scientific Co.) было определено, что 200 мкл ДНК, полученной путем очистки, имеет концентрацию ДНК HLA-Q-Mix, равную 48 нг/мкл.

Фрагментирование проводили в следующих условиях:

Качание частоты

Рабочий цикл	10%
Интенсивность	5
Циклы/«Вспышка»	200
Время (секунды)	300

Продукты реакции подвергали реакции репарации концов, а затем выделяли и

очищали с помощью набора для ПЦР-очистки QIAquick, и растворяли в 34 мкл EB (элюирующего буфера QIAGEN).

Затем к продуктам реакции присоединяли А у 3'-конца, после чего эти продукты выделяли и очищали с помощью набора для ПЦР-очистки MiniElute (QIAGEN Co.) и растворяли в 13 мкл раствора EB (элюирующего буфера QIAGEN).

После лигирования адаптеров библиотеки, продукты реакции очищали на сферах Ampure (Beckman Coulter Genomics) и растворяли в 50 мкл деионизованной воды, а затем определяли концентрацию ДНК с помощью количественной ПЦР с активацией флуоресценции (QPCR) следующим образом:

	Результат, полученный с помощью кол. ПЦР (нМ)
HLA-Q-Mix	115,3

Гель, содержащий ДНК-фрагменты размером от 350 до 550 п.н., разрезали (фигура 12). После очистки и выделения продуктов из геля, концентрация ДНК, измеренная с помощью количественной ПЦР с активацией флуоресценции (QPCR), составляла 18,83 нМ.

Анализ результатов

Результаты секвенирования, полученные на Illumina GA, представляют собой серию последовательностей ДНК, и после поиска последовательностей индексов прямого и обратного праймеров и последовательностей праймеров в результатах секвенирования создавали базы данных, содержащие результаты секвенирования ПЦР-продуктов различных экзонов HLA-DQB1 для каждого образца, соответствующего данному праймерному индексу. Результаты секвенирования каждого экзона сопоставляли путем выравнивания с эталонной последовательностью (эталонные последовательности можно найти на сайте <http://www.ebi.ac.uk/imgt/hla/>) соответствующего экзона с помощью программы BWA (Burrows-Wheeler Aligner), и таким образом конструировали консенсусные последовательности для каждой базы данных, а затем в этой базе данных отбирали последовательности «ридов» и корректировали исходя из качественной оценки секвенирования оснований и различий между последовательностями «ридов» и консенсусными последовательностями. Скорректированные последовательности ДНК получали путем сборки в соответствующие последовательности экзонов 2,3 HLA-DQB1 исходя из перекрытия и характера сцепления (парно-концевого сцепления) последовательностей. Процедура конструирования консенсусной последовательности экзона 2 HLA-DQB1-сайта в образце по.7, проводимая с помощью программы с захватом изображения на экране, проиллюстрирована на фигуре 13.

Полученную ДНК-последовательность для экзонов 2,3 HLA-DQB1 сопоставляли с последовательностями базы данных для соответствующего экзона HLA-DQB1, имеющегося в профессиональной базе данных HLA IMGT. Если результат выравнивания последовательностей дает 100% соответствие, то это означает, что HLA-DQB1-генотип соответствующего образца был точно определен.

Для всех 94 образцов результаты типирования, полученные вышеописанным методом, полностью соответствовали известным результатам типирования, где результаты для образцов No.1-32 представлены в таблице 12.

Таблица 12

Результаты типирования образцов No. 1-32

Образец №	Исходный известный DQB1-генотип		Результаты для DQB1, полученные на данный момент времени		Идентичный или нет
1	DQB1*02:02	DQB1*03:01	DQB1*02:02	DQB1*03:01	да
2	DQB1*02:02	DQB1*04:01	DQB1*02:02	DQB1*04:01	да
3	DQB1*05:02	DQB1*02:02	DQB1*05:02	DQB1*02:02	да
4	DQB1*02:02	DQB1*06:03	DQB1*02:02	DQB1*06:03	да
5	DQB1*03:03	DQB1*04:02	DQB1*03:03	DQB1*04:02	да
6	DQB1*05:02	DQB1*03:17	DQB1*05:02	DQB1*03:17	да
7	DQB1*03:03	DQB1*06:02	DQB1*03:03	DQB1*06:02	да
8	DQB1*05:03	DQB1*04:02	DQB1*05:03	DQB1*04:02	да
9	DQB1*04:02	DQB1*06:01	DQB1*04:02	DQB1*06:01	да
10	DQB1*05:01	DQB1*06:10	DQB1*05:01	DQB1*06:10	да
11	DQB1*03:01	DQB1*03:03	DQB1*03:01	DQB1*03:03	да
12	DQB1*05:01	DQB1*05:01	DQB1*05:01	DQB1*05:01	да
13	DQB1*02:02	DQB1*04:02	DQB1*02:02	DQB1*04:02	да
14	DQB1*05:02	DQB1*02:01	DQB1*05:02	DQB1*02:01	да
15	DQB1*02:01	DQB1*06:02	DQB1*02:01	DQB1*06:02	да
16	DQB1*03:03	DQB1*04:01	DQB1*03:03	DQB1*04:01	да
17	DQB1*05:01	DQB1*03:02	DQB1*05:01	DQB1*03:02	да
18	DQB1*03:03	DQB1*06:01	DQB1*03:03	DQB1*06:01	да
19	DQB1*03:03	DQB1*06:10	DQB1*03:03	DQB1*06:10	да
20	DQB1*05:03	DQB1*04:01	DQB1*05:03	DQB1*04:01	да
21	DQB1*05:02	DQB1*04:01	DQB1*05:02	DQB1*04:01	да
22	DQB1*03:01	DQB1*03:03	DQB1*03:01	DQB1*03:03	да
23	DQB1*05:02	DQB1*05:03	DQB1*05:02	DQB1*05:03	да
24	DQB1*05:02	DQB1*03:02	DQB1*05:02	DQB1*03:02	да
25	DQB1*03:03	DQB1*06:01	DQB1*03:03	DQB1*06:01	да
26	DQB1*05:02	DQB1*06:09	DQB1*05:02	DQB1*06:09	да

27	DQB1*02:02	DQB1*06:02	DQB1*02:02	DQB1*06:02	да
28	DQB1*05:02	DQB1*03:01	DQB1*05:02	DQB1*03:01	да
29	DQB1*02:01	DQB1*03:01	DQB1*02:01	DQB1*03:01	да
30	DQB1*06:03	DQB1*06:09	DQB1*06:03	DQB1*06:09	да
31	DQB1*05:02	DQB1*02:02	DQB1*05:02	DQB1*02:02	да
32	DQB1*05:01	DQB1*06:01	DQB1*05:01	DQB1*06:01	да

Пример 11: HLA-DQB1-генотипирование методом секвенирования по Сэнгеру

1. Экстракция образцов ДНК

Как описано в примере 1, на автоматическом оборудовании для экстракции KingFisher экстрагировали 20 из 94 образцов ДНК с известными HLA-генотипами.

2. ПЦР-амплификация

Вышеуказанные ДНК, экстрагированные на автоматическом оборудовании для экстракции KingFisher, использовали в качестве матриц и ПЦР-амплификацию осуществляли в отдельных пробирках с использованием двух пар ПЦР-праймеров (Q-F2 и Q-R2, Q-F3 и Q-R3), перечисленных в таблице 5, соответственно. Для каждой пары праймеров проводили ПЦР в следующих условиях:

96°C 2 мин

95°C 30 сек → 56°C 30 сек → 72°C 20 сек (35 циклов)

15°C ∞

ПЦР-реакцию для HLA-Q проводили на нижеследующей системе:

Promega 5×буфер I (плюс Mg ²⁺)	5,0 мкл
Смесь dNTP (2,5 мМ каждого)	2,0 мкл
Смесь праймеров (25 нг/мкл)	3,0 мкл
Promega Taq (5 ед./мкл)	0,2 мкл
ДНК (примерно 20 нг/мкл)	2,0 мкл
ddH ₂ O	12,8 мкл
Всего	25,0 мкл

Перед очисткой ПЦР-продукты подвергали электрофорезу в агарозном геле.

3. Очистка ПЦР-продуктов

Очистку проводили с применением метода и стадий, описанных в примере 9.

4. Осуществление реакции секвенирования и очистки продуктов реакции секвенирования

Эту реакцию и очистку проводили с применением метода и стадий, описанных в примере 9.

5. Секвенирование и анализ результатов

Очищенные продукты реакции секвенирования подвергали капиллярному электрофорезу в секвенирующем геле ABI 37 30XL. Пики секвенирования анализировали с помощью компьютерной программы uType (Invitrogen) и получали результаты HLA-типирования (фигура 15). Все результаты, полученные вышеописанным методом, были идентичны первоначальным известным результатам, как показано в таблице 13.

Таблица 13

Сравнение результатов типирования, полученных вышеописанным методом, с результатами типирования, полученными известным методом

Образец №	Исходный известный DQB1-генотип		Результаты для DQB1, полученные на данный момент времени		Идентичный или нет
1	DQB1*02:02	DQB1*03:01	DQB1*02:02	DQB1*03:01	да
2	DQB1*02:02	DQB1*04:01	DQB1*02:02	DQB1*04:01	да
3	DQB1*05:02	DQB1*02:02	DQB1*05:02	DQB1*02:02	да
4	DQB1*02:02	DQB1*06:03	DQB1*02:02	DQB1*06:03	да
5	DQB1*03:03	DQB1*04:02	DQB1*03:03	DQB1*04:02	да
6	DQB1*05:02	DQB1*03:17	DQB1*05:02	DQB1*03:17	да
7	DQB1*03:03	DQB1*06:02	DQB1*03:03	DQB1*06:02	да
8	DQB1*05:03	DQB1*04:02	DQB1*05:03	DQB1*04:02	да
9	DQB1*04:02	DQB1*06:01	DQB1*04:02	DQB1*06:01	да
10	DQB1*05:01	DQB1*06:10	DQB1*05:01	DQB1*06:10	да
11	DQB1*03:01	DQB1*03:03	DQB1*03:01	DQB1*03:03	да
12	DQB1*05:01	DQB1*05:01	DQB1*05:01	DQB1*05:01	да
13	DQB1*02:02	DQB1*04:02	DQB1*02:02	DQB1*04:02	да
14	DQB1*05:02	DQB1*02:01	DQB1*05:02	DQB1*02:01	да
15	DQB1*02:01	DQB1*06:02	DQB1*02:01	DQB1*06:02	да
16	DQB1*03:03	DQB1*04:01	DQB1*03:03	DQB1*04:01	да

17	DQB1*05:01	DQB1*03:02	DQB1*05:01	DQB1*03:02	да
18	DQB1*03:03	DQB1*06:01	DQB1*03:03	DQB1*06:01	да
19	DQB1*03:03	DQB1*06:10	DQB1*03:03	DQB1*06:10	да
20	DQB1*05:03	DQB1*04:01	DQB1*05:03	DQB1*04:01	да

Пример 12: Генотипирование экзонов 2, 3, 4 HLA-A/B/C и экзонов 2, 3 HLA-DQB1 в 950 образцах

В примерах настоящего изобретения экзоны 2, 3, 4 HLA-A/B/C и экзоны 2,3 HLA-DQB1 в 950 образцах были генотипированы с помощью комбинации праймерных индексов, стратегии неполного фрагментирования ДНК, индексов библиотеки, получения PCR-FREE-библиотек и парно-концевого метода секвенирования на секвенаторе Illumina GA 100 (ПЦР-продукты имели длину в пределах от 300 п.н. до 500 п.н.), что указывает на то, что такой способ согласно изобретению может быть применен для генотипирования генных фрагментов, длина которых превышает максимальную длину «ридов» в секвенаторе, и что настоящее изобретение позволяет осуществлять недорогостоящее HLA-генотипирование с высокой производительностью, высокой точностью и высоким разрешением.

Принцип: анализируемые образцы разделяли на 10 групп, и в образцах каждой группы в два конца ПЦР-продуктов экзонов 2, 3, 4 HLA-A/B/C и экзонов 2,3 HLA-DQB1 с помощью ПЦР-реакции вводили праймерные индексы в качестве специфических меток для получения информации о ПЦР-продуктах данного образца. Продукты ПЦР-амплификации четырех сайтов (HLA-A/B/C/DQB1) в каждой группе образцов смешивали с получением библиотеки ПЦР-продуктов, и после неполного ультразвукового фрагментирования библиотек ПЦР-продуктов конструировали индексированные PCR-FREE-библиотеки для секвенирования (если для библиотеки ПЦР-продуктов в образце каждой группы использовали различные адаптеры, то было сконструировано 10 индексированных библиотек для секвенирования). 10 индексированных библиотек для секвенирования объединяли в пул в равной молярной концентрации и получали смешанную индексированную библиотеку для секвенирования. Эту смешанную индексированную библиотеку для секвенирования подвергали электрофорезу в 2% агарозном геле с низкой температурой плавления, и все ДНК-полосы, длина которых составляла в пределах от 450 п.н. до 750 п.н., очищали и выделяли путем разрезания геля. Выделенную ДНК секвенировали методом Illumina GA PE-100. Информация о последовательностях всех тестируемых образцов может быть получена исходя из последовательностей праймерных индексов и последовательностей индексов библиотеки, а последовательность полноразмерного ПЦР-продукта может быть получена путем сборки исходя из известных эталонных последовательностей и с учетом перекрытия и характера сцепления между последовательностями ДНК-фрагментов. Полноразмерная последовательность исходного ПЦР-продукта может быть получена путем сопоставления с соответствующими экзонами HLA-A/B/C/DQB1 в стандартной базе данных и последующего HLA-A/B/C/DQB1-генотипирования.

1. Экстракция образцов

ДНК экстрагировали из 950 проб крови с использованием известных результатов HLA-SBT-типирования (по Программе забора костного мозга у доноров в Китае, далее называемой CMDP) на автоматическом оборудовании для экстракции KingFisher (US Thermo Co.). Этот метод описан в примере 1.

2. ПЦР-амплификация

950 ДНК, полученных после проведения каждой стадии экстракции образцов, были обозначены No.1-950 и были подразделены на 10 групп (95 ДНК в каждой группе), которые были обозначены HLA-1, HLA-2, HLA-3, HLA-4, HLA-5, HLA-6, HLA-7, HLA-8, HLA-9, HLA-10. Для каждой группы образцов 95 образцов ДНК амплифицировали с использованием 95 наборов ПЦР-праймеров, несущих двунаправленные праймерные индексы (таблица 6) для амплификации экзонов 2, 3, 4 HLA-A/B (таблица 2), для амплификации экзонов 2, 3, 4 HLA-C (таблица 4) и для амплификации экзонов 2, 3 HLA-DQB1 (таблица 5), соответственно. ПЦР-реакцию проводили в 96-луночных планшетах всего в 100 планшетах, обозначенных 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 и HLA-X-P-DQB1 («X» указывает на номер группы 1/2/3/4/5/6/7/8/9/10, а «A2/3/4», «B2/3/4», «C2/3/4», «DQB1» представляют собой сайты амплификации), где в каждом планшете имеется негативный контроль без какой-либо матрицы, а праймеры, используемые для негативного контроля, представляют собой праймеры, помеченные PI-1 (таблица 6). Во время эксперимента регистрировали информацию о каждом образце по номеру группы и праймерным индексам. Так, например, ниже представлена соответствующая информация о праймерных индексах PI-1 и PI-2, а остальная информация может быть получена аналогичным способом.

№ Прай- мерного индекса	SEQ ID NO:	Индекс прямого праймера	SEQ ID NO:	Индекс обратного праймера	Соответ- ствующее положение в 96-луночном планшете	Соответ- ствующая матрица (группа 1)	Соответ- ствующая матрица (группа n+1, где $1 \leq n < 10$, n представляет собой целое число)
PI-1	41	TCGCAGACATCA	42	TGACACGATGCT	A1	1	1+95*n
PI-2	43	TACATCGCACTA	44	TACAGATGCTGA	A2	2	2+95*n

Процедуру ПЦР и систему ПЦР-реакции для HLA-A/B/C осуществляли методом, описанным в примере 2. ПЦР-праймеры для амплификации соответствующих экзонов HLA-A/B представлены в таблице 2, а ПЦР-праймеры для амплификации соответствующих экзонов HLA-C представлены в таблице 4.

ПЦР для HLA-DQB1 проводили в следующих условиях:

96°C 2 мин

95°C 30 сек → 55°C 30 сек → 72°C 20 сек (32 цикла)

15°C ∞

Множество ПЦР-реакций для HLA-DQB1 (одновременной амплификации экзонов 2, 3) проводили с использованием системы, описанной в примере 10, а ПЦР-праймеры для амплификации соответствующих экзонов HLA-DQB1 представлены в таблице 5.

Если $PI_{nF}\text{-A/B/C-F}_{2/3/4}$ и $PI_{nF}\text{-Q-F2/F3}$ представляют собой F-праймеры HLA-A/B/C/DQB1, имеющие последовательности индекса прямого праймера No. n (таблица 6) у 5'-конца, то $PI_{nF}\text{-A/B/C-R}_{2/3/4}$ и $PI_{nF}\text{-Q-R2/R3}$ представляют собой R-праймеры HLA-A/B/C/DQB1, имеющие последовательность индекса обратного праймера No. n у 5'-конца (в данном случае $n \leq 95$), а остальные праймеры могут быть получены аналогичным образом. Кроме того, каждый образец соответствует конкретной серии ПЦР-праймеров ($PI_{nF}\text{-A/B/C-F}_{2/3/4}$, $PI_{nF}\text{-A/B/C-R}_{2/3/4}$, $PI_{nF}\text{-Q-F2/F3}$, $PI_{nF}\text{-Q-R2/R3}$).

ПЦР-реакцию осуществляли на оборудовании для ПЦР PTC-200, поставляемом Bio-Rad Co. После проведения ПЦР 3 мкл ПЦР-продуктов подвергали электрофорезу в 2% агарозном геле. На фигуре 16 представлены результаты электрофореза ПЦР-продуктов соответствующих экзонов HLA-A/B/C/DQB1 образца No.1, а маркером молекулярных масс ДНК является DL 2000 (Takara Co.). На электрофореграмме наблюдалась серия одиночных полос длиной в пределах от 300 п.н. до 500 п.н., что указывало на успешную ПЦР-амплификацию экзонов (A2, A3, A4, B2, B3, B4, C2, C3, C4, DRB1) HLA-A/B/C/DQB1 образца No.1. При этом в негативном контроле (N) полоса амплификации не наблюдалась. Результаты для других образцов были аналогичными.

3. Объединение в пул и очистка ПЦР-продуктов

Для образцов группы X («X» представляет собой 1/2/3/4/5/6/7/8/9/10), 20 мкл остальных ПЦР-продуктов, взятых из каждой лунки 96-луночного планшета с HLA-X-P-A2 (за

исключением негативного контроля), смешивали со встряхиванием в 3-миллилитровой ЕР-пробирке до получения гомогенной смеси (обозначенной HLA-X-A2-Mix).

Аналогичную процедуру проводили в других девяти 96-луночных планшетах с образцами группы X, обозначенных 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 и HLA-X-DQB1-Mix. Из каждого планшета брали 200 мкл смеси 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 и HLA-X-DQB1-Mix и смешивали в 3-миллилитровой ЕР-пробирке, обозначенной HLA-X-Mix. 500 мкл ДНК-смеси от HLA-X-Mix подвергали очистке на колонке с использованием набора для очистки ДНК Qiagen (QIAGEN Co.) (Для подробного описания стадий очистки см. инструкции производителей) с получением 200 мкл очищенной ДНК, концентрацию которой определяли на оборудовании Nanodrop 8000 (Thermo Fisher Scientific Co.). Ту же самую процедуру проводили и для других 9 групп образцов. Конечные определенные концентрации ДНК для 10 групп образцов указаны ниже.

	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
Концентр. (нг/мкл)	53,1	52,3	56,1	57,2	50,5	55,7	54,2	58,6	53,9	54,8

4. Конструирование библиотек для секвенирования Illumina GA

Как описано в примере 4, ДНК в общем количестве 5 мкг, взятых из очищенной HLA-X-Mix, подвергали фрагментации, а после фрагментации очищали, подвергали реакции репарации концов, присоединяли А к 3'-концу и лигировали с адаптером PCR-FREE-библиотеки Illumina GA.

Ниже показана соответствующая взаимосвязь между группами образцов и адаптерами библиотеки.

№ Образца группы	HLA-1	HLA-2	HLA-3	HLA-4	HLA-5	HLA-6	HLA-7	HLA-8	HLA-9	HLA-10
№ адаптера библиотеки	1	2	3	4	5	6	7	8	9	10

Полученные продукты реакции очищали на сферах Ampure (Beckman Coulter Genomics) и растворяли в 50 мкл деионизованной воды, а затем определяли концентрации ДНК с помощью количественной ПЦР с активацией флуоресценции (QPCR) следующим образом:

	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
Конц. (нМ)	86,60	78,21	54,56	87,35	84,37	85,09	96,21	85,81	88,14	88,26

5. Выделение путем разрезания геля

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 и HLA-10-Mix смешивали в равных молярных концентрациях (конечная концентрация составляла 70,86 нМ/мкл) и обозначали HLA-Mix-10. 30 мкл HLA-Mix-10 подвергали электрофорезу в 2% агарозном геле с низкой температурой плавления. Электрофорез проводили при 100 В, в течение 100 минут. ДНК-маркер представлял собой 50 п.н. ДНК-маркера от NEB Co. Гель, содержащий ДНК-фрагменты размером в пределах от 450 до 750 п.н., разрезали (фигура 17). Продукты, выделенные из разрезанного геля, очищали с помощью набора для ПЦР-очистки QIAquick (QIAGEN Co.), и после такой очистки объем составлял 32 мкл, а концентрация ДНК, измеренная с помощью количественной ПЦР с активацией флуоресценции (QPCR), составляла 10,25

нМ.

6. Секвенирование методом Illumina GA и анализ результатов

Секвенирование и анализ результатов осуществляли методами, описанными в примерах 5 и 6.

- 5 Были созданы базы данных, содержащих результаты секвенирования ПЦР-продуктов различных экзонов HLA-A/B/C/DQB1 для каждого образца, соответствующего определенному праймерному индексу. Полученную последовательность ДНК сопоставляли путем выравнивания с последовательностями базы данных для соответствующего экзона HLA-A/B/C/DQB1, имеющегося в профессиональной базе
- 10 данных HLA IMGT. Если результат выравнивания последовательностей дает 100% соответствие, то это означает, что HLA-A/B/C/DQB1-генотип соответствующего образца был точно определен. См. программу с захватом изображения на экране для конструирования консенсусной последовательности экзона 2 HLA-A-сайта образца No.1, как показано на фигуре 18. Для всех 950 образцов результаты типирования,
- 15 полученные вышеописанным методом, полностью соответствовали первоначальным известным результатам типирования, где результаты для образцов No.1-32 представлены ниже:

No.		Исходный известный тип HLA-A/B/C/DQB1							
20	1	A*02:	A*03:	B*07:	B*46:	C*01:	C*07:	DQB1*03	DQB1*06
		07	01	02	01	02	02	:03	:02
25	2	A*11:	A*31:	B*15:	B*38:	C*03:	C*07:	DQB1*03	DQB1*04
		01	01	11	02	03	02	:03	:01
30	3	A*02:	A*24:	B*13:	B*46:	C*01:	C*03:	DQB1*03	DQB1*06
		07	02	01	01	02	04	:02	:01
35	4	A*24:	A*33:	B*40:	B*51:	C*01:	C*14:	DQB1*03	DQB1*03
		02	03	01	01	02	02	:03	:03
40	5	A*31:	A*31:	B*15:	B*35:	C*04:	C*04:	DQB1*03	DQB1*06
		01	01	01	01	01	01	:02	:02
45	6	A*02:	A*03:	B*44:	B*46:	C*01:	C*05:	DQB1*03	DQB1*06
		07	01	02	01	02	01	:01	:02
	7	A*02:	A*30:	B*07:	B*13:	C*06:	C*07:	DQB1*02	DQB1*06
		01	01	02	02	02	02	:02	:01
	8	A*02:	A*02:	B*46:	B*46:	C*01:	C*01:	DQB1*05	DQB1*03
		07	07	01	01	02	02	:02	:03

5	9	A*01:	A*33:	B*49:	B*58:	C*03:	C*07:	DQB1*06	DQB1*06
		01	03	01	01	02	01	:04	:09
10	10	A*02:	A*11:	B*46:	B*48:	C*01:	C*08:	DQB1*05	DQB1*03
		07	01	01	01	03	01	:03	:02
10	11	A*02:	A*30:	B*13:	B*15:	C*06:	C*08:	DQB1*03	DQB1*03
		06	01	02	02	02	01	:01	:01
15	12	A*24:	A*31:	B*35:	B*51:	C*03:	C*14:	DQB1*03	DQB1*06
		02	01	01	01	03	02	:03	:01
15	13	A*11:	A*33:	B*46:	B*46:	C*01:	C*01:	DQB1*03	DQB1*03
		01	03	01	01	02	02	:02	:03
20	14	A*01:	A*02:	B*38:	B*57:	C*06:	C*07:	DQB1*05	DQB1*03
		01	03	02	01	02	02	:02	:03
20	15	A*02:	A*24:	B*13:	B*15:	C*03:	C*07:	DQB1*03	DQB1*06
		06	02	01	25	04	02	:01	:01
25	16	A*11:	A*24:	B*15:	B*15:	C*04:	C*08:	DQB1*03	DQB1*03
		01	02	02	27	01	01	:01	:03
30	17	A*24:	A*24:	B*40:	B*46:	C*01:	C*03:	DQB1*03	DQB1*06
		02	02	01	01	02	04	:03	:02
30	18	A*24:	A*24:	B*40:	B*46:	C*01:	C*03:	DQB1*03	DQB1*06
		02	02	01	01	02	04	:03	:02
35	19	A*11:	A*33:	B*40:	B*58:	C*03:	C*03:	DQB1*02	DQB1*03
		01	03	02	01	02	04	:01	:02
40	20	A*24:	A*30:	B*13:	B*40:	C*03:	C*06:	DQB1*06	DQB1*06
		02	01	02	01	04	02	:02	:03
40	21	A*02:	A*24:	B*40:	B*40:	C*07:	C*14:	DQB1*04	DQB1*06
		01	02	01	01	02	02	:02	:02
45	22	A*02:	A*33:	B*15:	B*44:	C*01:	C*14:	DQB1*03	DQB1*06

		01	03	01	03	02	03	:01	:04
5	23	A*26:	A*33:	B*15:	B*58:	C*03:	C*08:	DQB1*02	DQB1*03
		01	03	01	01	02	01	:01	:01
	24	A*02:	A*11:	B*13:	B*55:	C*01:	C*03:	DQB1*03	DQB1*03
		01	01	01	02	06	04	:01	:03
10	25	A*02:	A*32:	B*40:	B*52:	C*03:	C*12:	DQB1*05	DQB1*06
		01	01	01	01	04	02	:02	:01
	26	A*02:	A*02:	B*40:	B*46:	C*01:	C*07:	DQB1*03	DQB1*06
15		03	07	01	01	02	02	:02	:01
	27	A*02:	A*02:	B*46:	B*46:	C*01:	C*01:	DQB1*03	DQB1*06
		07	07	01	01	02	02	:03	:01
20	28	A*24:	A*30:	B*13:	B*39:	C*06:	C*07:	DQB1*02	DQB1*06
		02	01	02	05	02	02	:02	:01
	29	A*31:	A*33:	B*15:	B*58:	C*03:	C*07:	DQB1*04	DQB1*06
25		01	03	18	01	02	04	:01	:09
	30	A*02:	A*03:	B*27:	B*40:	C*02:	C*03:	DQB1*03	DQB1*03
		06	01	05	02	02	04	:01	:01
	31	A*02:	A*33:	B*15:	B*58:	C*03:	C*08:	DQB1*05	DQB1*06
30		06	03	02	01	02	01	:01	:01
	32	A*03:	A*30:	B*13:	B*51:	C*06:	C*15:	DQB1*02	DQB1*03
		01	01	02	01	02	02	:02	:01
35	No.	тип HLA-A/B/C/DQB1, определенный способом согласно изобретению							
	1	A*02:	A*03:	B*07:	B*46:	C*01:	C*07:	DQB1*03	DQB1*06
		07	01	02	01	02	02	:03	:02
40	2	A*11:	A*31:	B*15:	B*38:	C*03:	C*07:	DQB1*03	DQB1*04
		01	01	11	02	03	02	:03	:01
45									

	3	A*02:	A*24:	B*13:	B*46:	C*01:	C*03:	DQB1*03	DQB1*06
		07	02	01	01	02	04	:02	:01
5	4	A*24:	A*33:	B*40:	B*51:	C*01:	C*14:	DQB1*03	DQB1*03
		02	03	01	01	02	02	:03	:03
10	5	A*31:	A*31:	B*15:	B*35:	C*04:	C*04:	DQB1*03	DQB1*06
		01	01	01	01	01	01	:02	:02
	6	A*02:	A*03:	B*44:	B*46:	C*01:	C*05:	DQB1*03	DQB1*06
		07	01	02	01	02	01	:01	:02
15	7	A*02:	A*30:	B*07:	B*13:	C*06:	C*07:	DQB1*02	DQB1*06
		01	01	02	02	02	02	:02	:01
20	8	A*02:	A*02:	B*46:	B*46:	C*01:	C*01:	DQB1*05	DQB1*03
		07	07	01	01	02	02	:02	:03
	9	A*01:	A*33:	B*49:	B*58:	C*03:	C*07:	DQB1*06	DQB1*06
		01	03	01	01	02	01	:04	:09
25	10	A*02:	A*11:	B*46:	B*48:	C*01:	C*08:	DQB1*05	DQB1*03
		07	01	01	01	03	01	:03	:02
30	11	A*02:	A*30:	B*13:	B*15:	C*06:	C*08:	DQB1*03	DQB1*03
		06	01	02	02	02	01	:01	:01
	12	A*24:	A*31:	B*35:	B*51:	C*03:	C*14:	DQB1*03	DQB1*06
		02	01	01	01	03	02	:03	:01
35	13	A*11:	A*33:	B*46:	B*46:	C*01:	C*01:	DQB1*03	DQB1*03
		01	03	01	01	02	02	:02	:03
	14	A*01:	A*02:	B*38:	B*57:	C*06:	C*07:	DQB1*05	DQB1*03
		01	03	02	01	02	02	:02	:03
40	15	A*02:	A*24:	B*13:	B*15:	C*03:	C*07:	DQB1*03	DQB1*06
		06	02	01	25	04	02	:01	:01
45	16	A*11:	A*24:	B*15:	B*15:	C*04:	C*08:	DQB1*03	DQB1*03

		01	02	02	27	01	01	:01	:03
5	17	A*24:	A*24:	B*40:	B*46:	C*01:	C*03:	DQB1*03	DQB1*06
		02	02	01	01	02	04	:03	:02
	18	A*24:	A*24:	B*40:	B*46:	C*01:	C*03:	DQB1*03	DQB1*06
		02	02	01	01	02	04	:03	:02
10	19	A*11:	A*33:	B*40:	B*58:	C*03:	C*03:	DQB1*02	DQB1*03
		01	03	02	01	02	04	:01	:02
	20	A*24:	A*30:	B*13:	B*40:	C*03:	C*06:	DQB1*06	DQB1*06
15		02	01	02	01	04	02	:02	:03
	21	A*02:	A*24:	B*40:	B*40:	C*07:	C*14:	DQB1*04	DQB1*06
		01	02	01	01	02	02	:02	:02
20	22	A*02:	A*33:	B*15:	B*44:	C*01:	C*14:	DQB1*03	DQB1*06
		01	03	01	03	02	03	:01	:04
	23	A*26:	A*33:	B*15:	B*58:	C*03:	C*08:	DQB1*02	DQB1*03
		01	03	01	01	02	01	:01	:01
25	24	A*02:	A*11:	B*13:	B*55:	C*01:	C*03:	DQB1*03	DQB1*03
		01	01	01	02	06	04	:01	:03
	25	A*02:	A*32:	B*40:	B*52:	C*03:	C*12:	DQB1*05	DQB1*06
30		01	01	01	01	04	02	:02	:01
	26	A*02:	A*02:	B*40:	B*46:	C*01:	C*07:	DQB1*03	DQB1*06
		03	07	01	01	02	02	:02	:01
35	27	A*02:	A*02:	B*46:	B*46:	C*01:	C*01:	DQB1*03	DQB1*06
		07	07	01	01	02	02	:03	:01
	28	A*24:	A*30:	B*13:	B*39:	C*06:	C*07:	DQB1*02	DQB1*06
40		02	01	02	05	02	02	:02	:01
	29	A*31:	A*33:	B*15:	B*58:	C*03:	C*07:	DQB1*04	DQB1*06
		01	03	18	01	02	04	:01	:09
45									

	30	A*02:	A*03:	B*27:	B*40:	C*02:	C*03:	DQB1*03	DQB1*03
		06	01	05	02	02	04	:01	:01
5	31	A*02:	A*33:	B*15:	B*58:	C*03:	C*08:	DQB1*05	DQB1*06
		06	03	02	01	02	01	:01	:01
	32	A*03:	A*30:	B*13:	B*51:	C*06:	C*15:	DQB1*02	DQB1*03
10		01	01	02	01	02	02	:02	:01

Примечание: В случае если последовательности экзонов 2, 3, 4 HLA-A/B/C были полностью идентичны, то был выбран общий тип.

950 образцов с известными результатами HLA-SBT-типирования подвергали генотипированию в HLA-A/B/C/DQB1-сайтах с применением технической стратегии согласно изобретению, и полученные результаты показали, что данные, полученные с применением технической стратегии согласно изобретению, полностью соответствуют ранее полученным известным данным.

Хотя в настоящей заявке подробно описаны варианты осуществления изобретения, однако для специалиста в данной области очевидно, что исходя из описания изобретения, в него могут быть внесены различные модификации и изменения, не выходящие за рамки существа и объема изобретения. Объем настоящего изобретения определен прилагаемой формулой изобретения и любым ее эквивалентом.

Источники информации

- [1]. <http://www.ebi.ac.uk/imgt/hla/stats.html>.
- [2]. Tiercy J M. Molecular basis of HLA polymorphism: implications in clinical transplantation. [J]. Transpl Immunol, 2002, 9:173-180.
- [3]. C.Antoine, S.Müller, A.Cant, et al. Long-term survival and transplantation of haemopoietic stem cells for immunodeficiencies: report of the European experience. 1968-99. [J]. The Lancet, 2003, 9357:553-560.
- [4]. H. A. Erlich, G. Opelz, J. Hansen, et al. HLA DNA Typing and Transplantation.[J]. Immunity, 2001, 14:347-356.
- [5]. Lillo R, Balas A, Vicario JL, et al. Two new HLA class allele, DPB1*02014, by sequence-based typing. [J]. Tissue Antigens, 2002, 59:47-48.
- [6]. A. Dormoy, N. Froelich. Leisenbach, et al. Mono-allelic amplification of exons 2-4 using allele group-specific primers for sequence-based typing (SBT) of the HLA-A, -B and -C genes: Preparation and validation of ready-to-use pre-SBT mini-kits. [J]. Tissue Antigens, 2003, 62:201-216.
- [7]. Elaine R. Mardis. The impact of next-generation sequencing technology on genetics. [J]. Trends in Genetics. 2008, 24:133-141.
- [8]. Christian Hoffmann1, Nana Minkah1, Jeremy Leipzig. DNA barcoding and pyrosequencing to identify rare HIV drug resistance mutations. [J]. Nucleic Acids Research, 2007, 1-8.
- [9]. Shannon J. Odelberg, Robert B. Weiss, Akira Hata. Template-switching during DNA synthesis by Thermus aquaticus DNA polymerase I. [J]. Nucleic Acids Research. 1995, 23:2049-2057.
- [10]. Sayer D, Whidborne R, Brestovac B. HLA-DRB1 DNA sequencing based typing: an approach suitable for high throughput typing including unrelated bone marrow registry donors. [J]. Tissue Antigens. 2001, 57(1):46-54.
- [11]. Iwanka Kozarewa, Zemin Ning, Michael A Quail. Amplification-free Illumina sequencing-library preparation facilitates improved mapping and assembly of (G+C)-biased genomes. [J].

Nature Methods. 2009, 6:291-295.

[12]. Marsh, S.G.E., Parham, P. & Barber, L.D. The HLA Facts Book 3-91 (Academic Press, London, 2000).

[13]. Campbell, K.J. et al. Characterization of 47 MHC class I sequences in Filipino cynomolgus macaques. Immunogenetics 61, 177-187 (2009).

[14]. Goulder, P.J.R. & Watkins, D.I. Impact of MHC class I diversity on immune control of immunodeficiency virus replication. Nat. Rev. Immunol. 8, 619-630 (2008).

[15]. O'Leary, C.E. et al. Identification of novel MHC class I sequences in pig-tailed macaques by amplicon pyrosequencing and full-length cDNA cloning and sequencing. Immunogenetics 61, 689-701 (2009).

[16]. Robinson J, Malik A, Parham P, Bodmer JG, Marsh SGE. IMGT/HLA database-a sequence database for the human major histocompatibility complex. Tissue Antigens 55, 80-7 (2000).

[17]. Hoffmann C, Minkah N, Leipzig J, Wang G, Arens MQ, Tebas P, Bushman FD. DNA bar coding and pyrosequencing to identify rare HIV drug resistance mutations. Nucleic Acids Res. 2007; 35(13):e91.

[18]. WU, D. L. et al. Comparative analysis of serologic typing and HLA-II typing by 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 class II sequence based typing in normal Saudi individuals. Tissue Antigens, 2002, 60: 259-261.

[20]. D. C. Sayer, D. M. Goodridge. Pilot study: assessment of interlaboratory variability of sequencing-based typing DNA sequence data quality. Tissue Antigens, 2007, 69 Suppl: 66-68.

[21]. Horton V, Stratton I, Bottazzo G. F. et al. Genetic heterogeneity of autoimmune diabetes: age of presentation in adults is influenced by HLA DRB1 and DQB1 genotypes. Diabetologia, 1999, 42:608-616.

[22]. C.E.M. Voorter, M.C. Kik1, E.M. van den Berg-Loonen et al. High-resolution HLA typing for the DQB1 gene by sequence-based typing. Tissue Antigens, 2008, 51:80-87.

[23]. G. Bentley, R. Higuchi, B. Hoglund et al. High-resolution, high-throughput HLA genotyping by next-generation sequencing. Tissue Antigens, 2009, 74:393-403.

Формула изобретения

1. Способ определения нуклеотидной последовательности представляющей интерес нуклеиновой кислоты в образце, где указанный способ включает:

1) получение n образцов, где n представляет собой целое число ≥ 1 ; причем необязательно указанные n образцов, предназначенные для анализа, разделяют на m групп, где m представляет собой целое число и $n \geq m \geq 1$;

2) амплификацию, где пару или множество пар индексированных праймеров используют для каждого образца, если от этого образца имеются матрицы; ПЦР-амплификацию осуществляют в условиях, подходящих для амплификации представляющей интерес нуклеиновой кислоты, где каждая пара индексированных праймеров состоит из прямого индексированного праймера и обратного индексированного праймера, содержащих праймерные индексы, где праймерные индексы, присутствующие в прямом индексированном праймере и в обратном индексированном праймере каждой пары индексированных праймеров, могут быть одинаковыми или различными и где индексы праймеров в парах индексированных праймеров, используемых для различных образцов, являются различными;

где каждая пара индексированных праймеров образована парой праймерных индексов и парой ПЦР-праймеров, где прямой и обратный ПЦР-праймеры имеют прямой праймерный индекс и обратный праймерный индекс, необязательно связанные

линкерной последовательностью на 5'-конце, соответственно;

3) объединение в пул, где если $n > 1$, то ПЦР-продукты от каждого образца объединяют в один пул;

4) фрагментирование, где амплифицированные продукты подвергают неполному
5 фрагментированию, очистке и выделению;

5) секвенирование, где выделенную ДНК-смесь подвергают секвенированию методом секвенирования второго поколения с получением последовательностей фрагментированной ДНК; и

6) сборку, где устанавливают соответствие между полученными данными
10 секвенирования и образцами один за другим по уникальному праймерному индексу для каждого образца; риды каждой последовательности сопоставляют путем выравнивания с эталонной последовательностью ДНК, соответствующей ПЦР-продуктам, с помощью программы выравнивания; сборку полноразмерной последовательности представляющей интерес нуклеиновой кислоты осуществляют из
15 последовательностей фрагментированной ДНК по перекрыванию и характеру сцепления последовательностей.

2. Способ по п. 1, где указанная пара праймерных индексов выбрана из 95 пар праймерных индексов, показанных в Таблице 6; и указанные ПЦР-праймеры выбраны из SEQ ID NO: 1-40 и 231-238.

20 3. Способ по п. 1 или 2, где указанные ПЦР-праймеры выбраны из:

(1) ПЦР-праймеров для амплификации экзонов 2, 3 и 4 HLA-A/B, как показано в таблице 1 или в таблице 2;

(2) ПЦР-праймеров для амплификации экзона 2 HLA-DRB1, как показано в таблице
7;

25 (3) ПЦР-праймеров для амплификации экзонов 2, 3 и/или 4 HLA-C, представленных в таблице 3 или в таблице 4; и

(4) ПЦР-праймеров для амплификации экзона 2 и/или 3 гена HLA-DQB1, представленных в таблице 5.

4. Способ по п. 1 или 2, где указанные праймерные индексы содержат по меньшей
30 мере 10, или по меньшей мере 20, или по меньшей мере 30, или по меньшей мере 40, или по меньшей мере 50, или по меньшей мере 60, или по меньшей мере 70, или по меньшей мере 80, или по меньшей мере 90, или все 95 пар из 95 пар праймерных индексов, представленных в таблице 6; и/или

где указанные праймерные индексы состоят из 10-95 пар, 20-95 пар, 30-95 пар, 40-
35 95 пар, 50-95 пар, 60-95 пар, 70-95 пар, 80-95 пар, 90-95 пар или 95 пар из 95 пар праймерных индексов, представленных в таблице 6; и/или

где указанные праймерные индексы включают по меньшей мере PI-1-PI-10, или PI-11-PI-20, или PI-21-PI-30, или PI-31-PI-40, или PI-41-PI-50, или PI-51-PI-60, или PI-61-PI-70, или PI-71-PI-80, или PI-81-PI-90, или PI-91-PI-95 из 95 пар праймерных индексов,
40 представленных в таблице 6, или комбинации любых двух или более из указанных пар.

5. Способ по п. 1 или 2, где указанное фрагментирование ДНК включает методы химического фрагментирования и методы физического фрагментирования, где методы химического фрагментирования включают ферментативное расщепление, а методы физического фрагментирования включают методы ультразвукового фрагментирования
45 или методы механического фрагментирования.

6. Способ по п. 1 или 2, где после указанного фрагментирования ДНК все ДНК-полосы, имеющие длину между максимальной длиной ридов в секвенаторе и максимально приемлемой длиной ДНК в секвенаторе, очищают и выделяют, где

указанные методы очистки и выделения включают выделение путем электрофореза и разрезания геля и выделение на магнитных сферах.

7. Способ по п. 1, где указанный способ, помимо стадий 1)-4) п. 1, включает следующие стадии:

5 5) конструирование библиотеки, где конструирование PCR-FREE-библиотеки для секвенирования осуществляют с использованием библиотеки фрагментированных ПЦР-продуктов, где различные адаптеры библиотеки могут быть добавлены для различения различных PCR-FREE-библиотек для секвенирования и где все ДНК-полосы, имеющие длину между максимальной длиной ридов в секвенаторе и максимально приемлемой длиной ДНК в секвенаторе, очищают и выделяют;

10 6) секвенирование, где выделенную ДНК-смесь подвергают секвенированию методом секвенирования второго поколения с получением последовательностей фрагментированных ДНК; и

15 7) сборку, где устанавливают соответствие между полученными данными секвенирования и образцами один за другим по различным последовательностям адаптеров библиотеки для библиотек и по уникальному праймерному индексу для каждого образца; каждую последовательность рида сопоставляют путем выравнивания с эталонной последовательностью ДНК, соответствующей ПЦР-продуктам, с помощью программы выравнивания; а сборку полноразмерной последовательности представляющей интерес нуклеиновой кислоты осуществляют из последовательностей фрагментированной ДНК по перекрыванию и характеру сцепления последовательностей.

8. Способ по п. 7, где на стадии 5) очищают и выделяют фрагменты ДНК длиной 450-750 п.о.

25 9. Способ по п. 1 или 7, где прямой индексированный праймер и/или обратный индексированный праймер представляют собой вырожденные праймеры и/или где метод секвенирования второго поколения представляет собой парно-концевой метод.

10. Способ по п. 1 или 7, где метод секвенирования второго поколения представляет собой метод секвенирования Illumina GA или Illumina HiSeq 2000.

30 11. Способ по п. 1 или 7, где программа выравнивания представляет собой Blast или BWA.

12. Способ по п. 1 или 2, где указанные образцы берут у млекопитающего.

13. Способ по п. 12, где указанные образцы берут у человека.

14. Способ по п. 12, где указанные образцы представляет собой пробы человеческой крови.

35 15. Применение способа по любому из пп. 1-14 в HLA-типировании, где указанный способ отличается тем, что он включает секвенирование взятого у пациента образца указанным способом по любому из пп. 1-13 и сопоставление результатов секвенирования путём выравнивания с данными о последовательностях экзонов 2, 3, 4 HLA-A/B, экзонов 2, 3 и/или 4 HLA-C, экзона 2 и/или 3 гена HLA-DQB1 и/или экзона 2 HLA-DRB1 в базе данных HLA; где если результат выравнивания последовательностей дает 100% соответствие, то это означает, что HLA-генотип соответствующего образца был точно определен.

45 16. Применение по п. 15, где указанный образец представляет собой пробу крови и/или где указанная база данных HLA представляет собой профессиональную базу данных HLA IMGT.

17. Способ по п. 1 или 2, где указанный способ используется для способа HLA-типирования и дополнительно включает следующую стадию:

типирование, где результаты секвенирования сопоставляют с данными о

последовательностях экзонов 2, 3, 4 HLA-A/B, экзонов 2, 3 и/или 4 HLA-C, экзона 2 и/или 3 гена HLA-DQB1 и/или экзона 2 HLA-DRB1 в базе данных HLA; где если результат выравнивания последовательностей дает 100% соответствие, то это означает, что HLA-генотип соответствующего образца был точно определен.

5 18. Способ по п. 17, где указанная база данных HLA представляет собой профессиональную базу данных HLA IMGT.

19. Применение набора в способе по любому из пп. 1-14, где указанный набор включает пару или совокупность пар индексированных праймеров и, необязательно, средства для амплификации ДНК, очистки ДНК и/или секвенирования ДНК, где каждая
10 пара индексированных праймеров включает пару праймерных индексов и пару ПЦР-праймеров, где прямой и обратный ПЦР-праймеры имеют прямой и обратный праймерные индексы, необязательно связанные линкерной последовательностью на 5'-конце, соответственно; где указанная пара праймерных индексов выбрана из 95 пар праймерных индексов, показанных в Таблице 6; и указанные ПЦР-праймеры выбраны
15 из SEQ ID NO: 1-40 и 231-238.

20

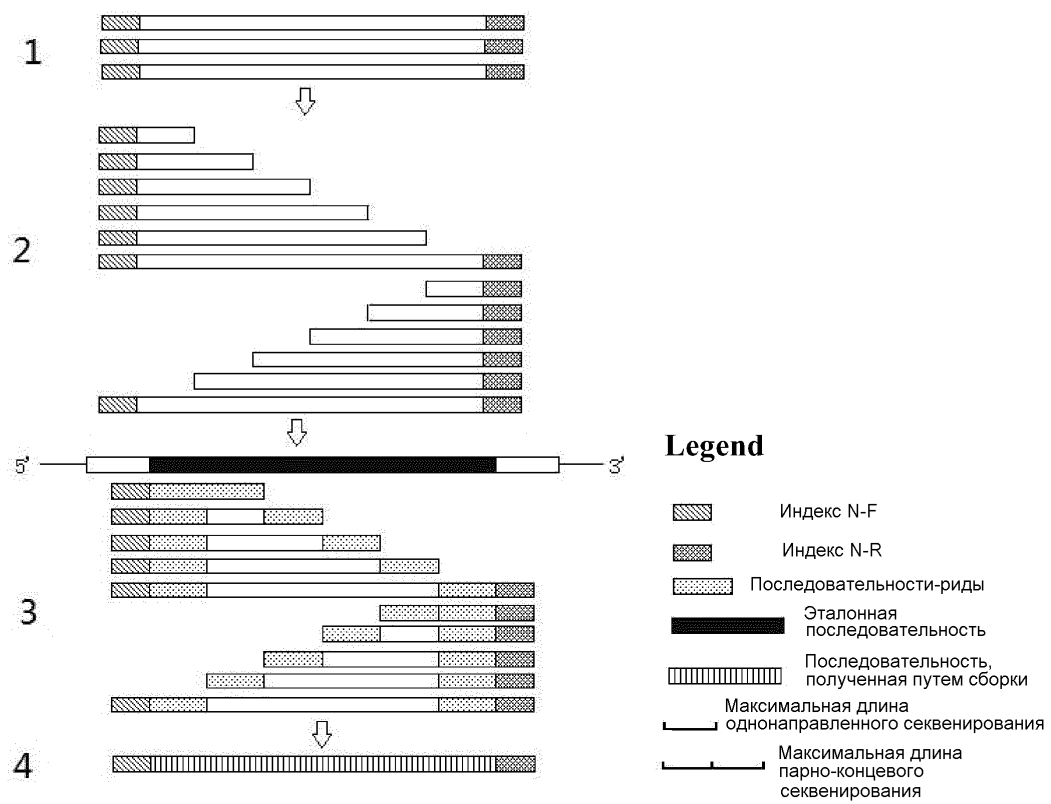
25

30

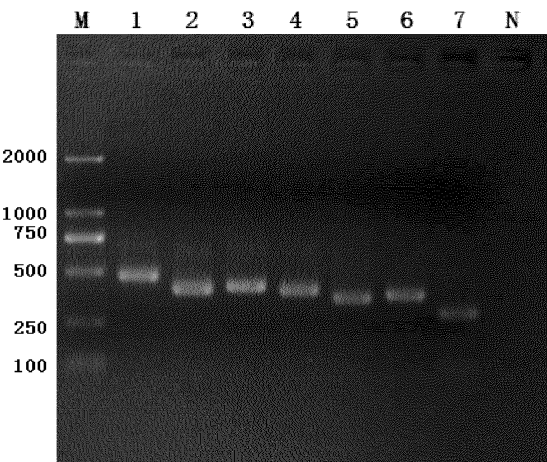
35

40

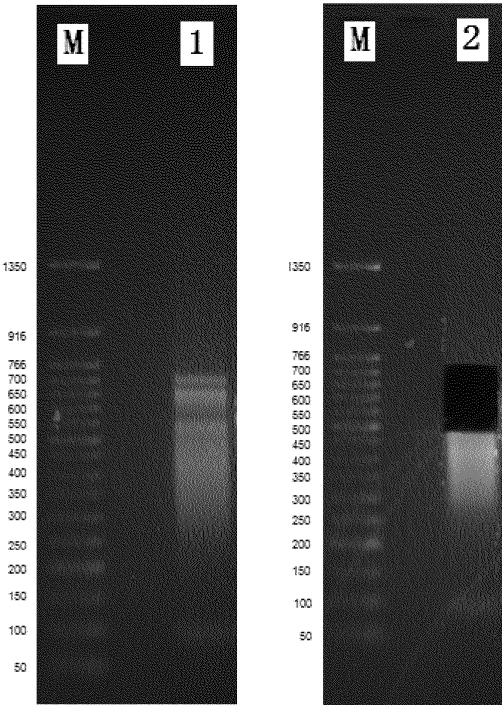
45



Фиг.1

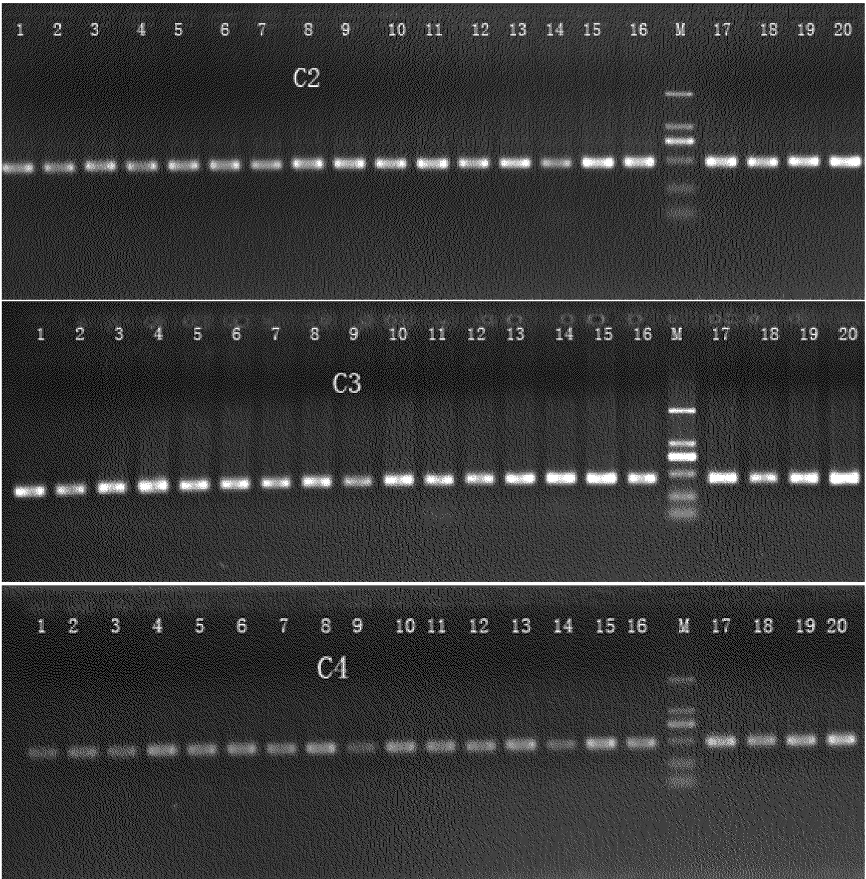


Фиг.2

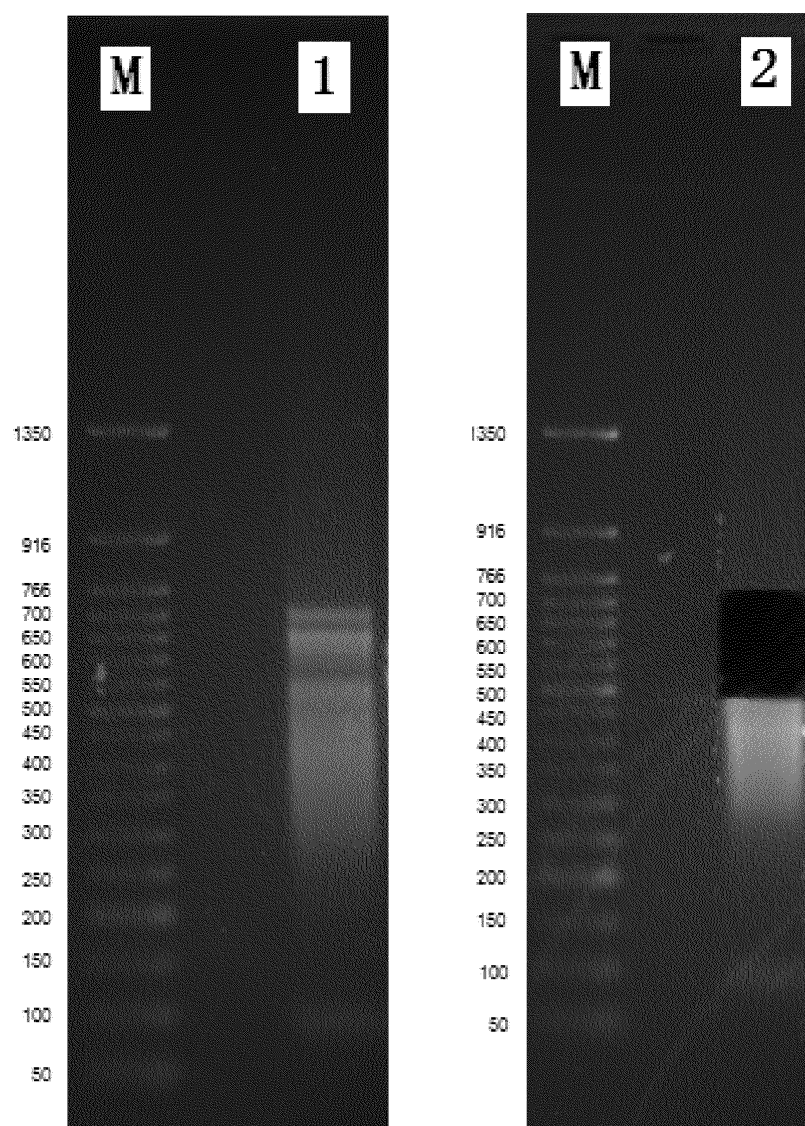


Фиг.3



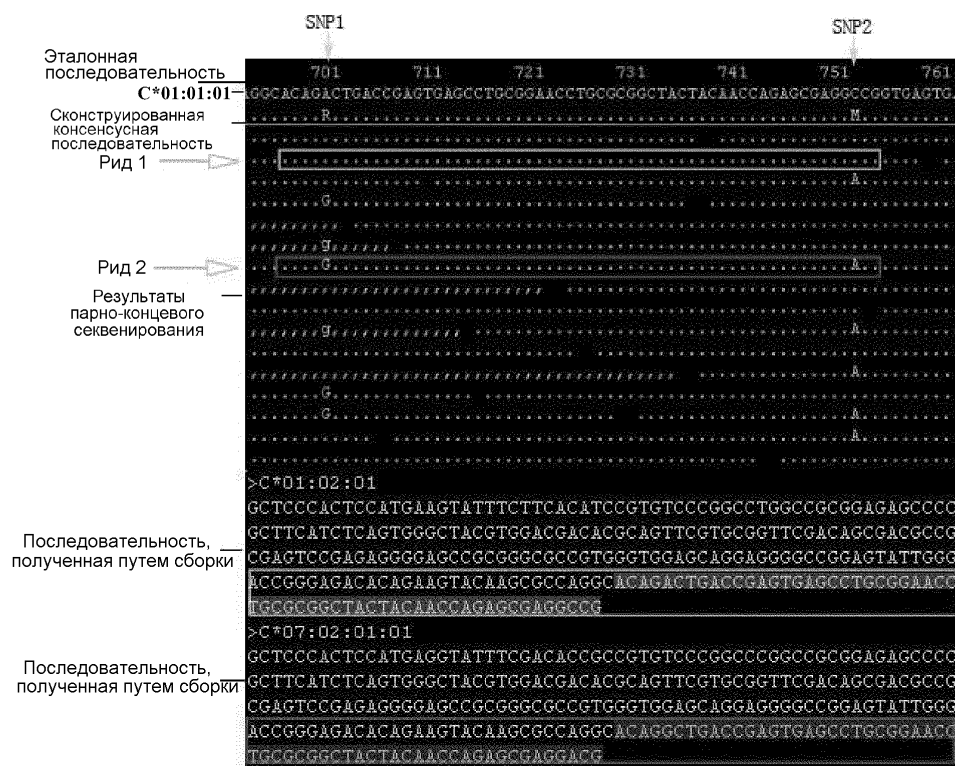


Фиг.6



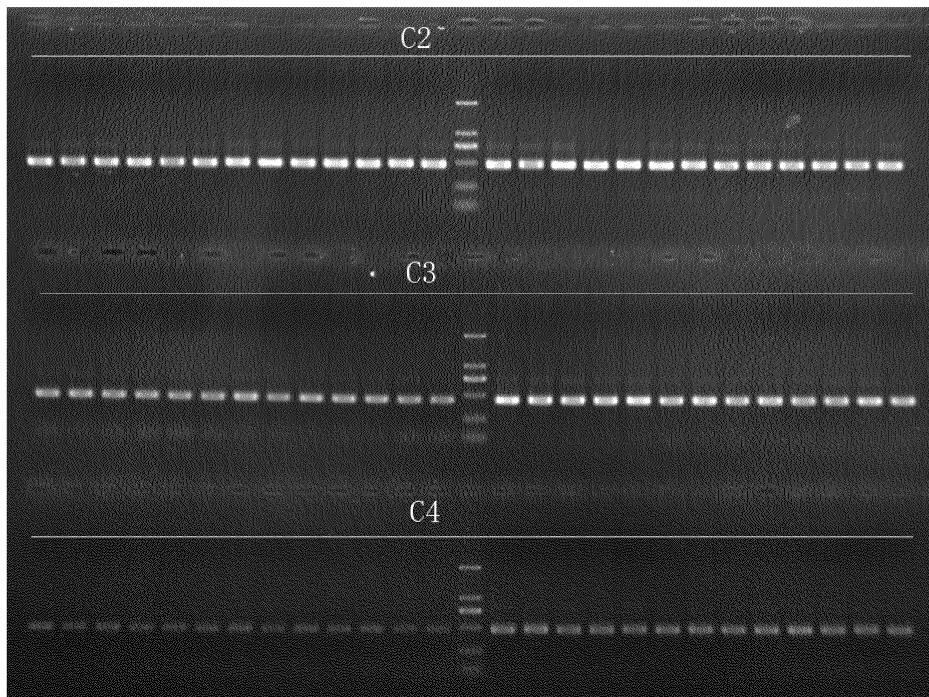
Фиг.7

6/12

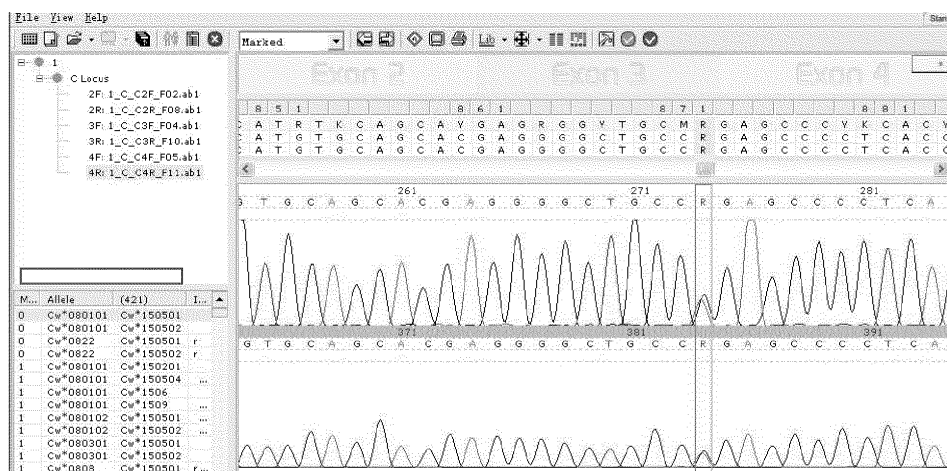


Фиг.8

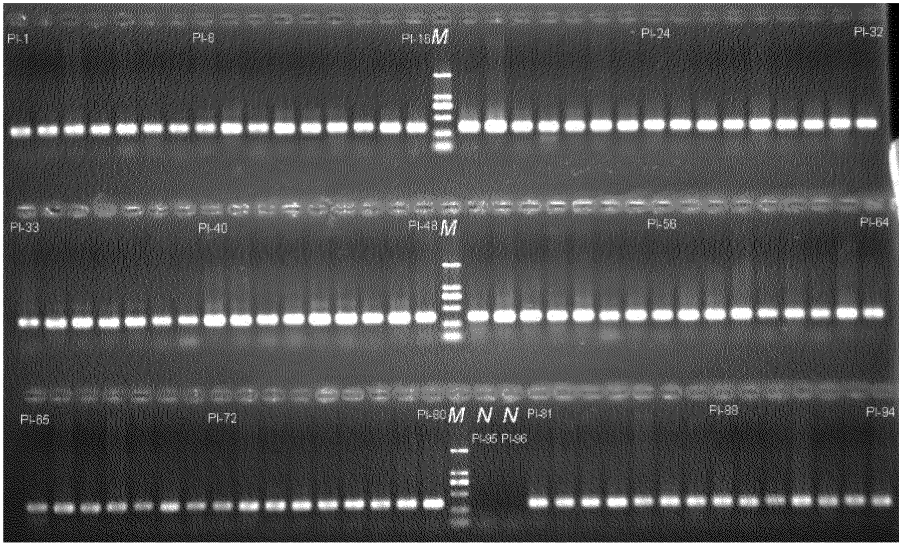
7/12



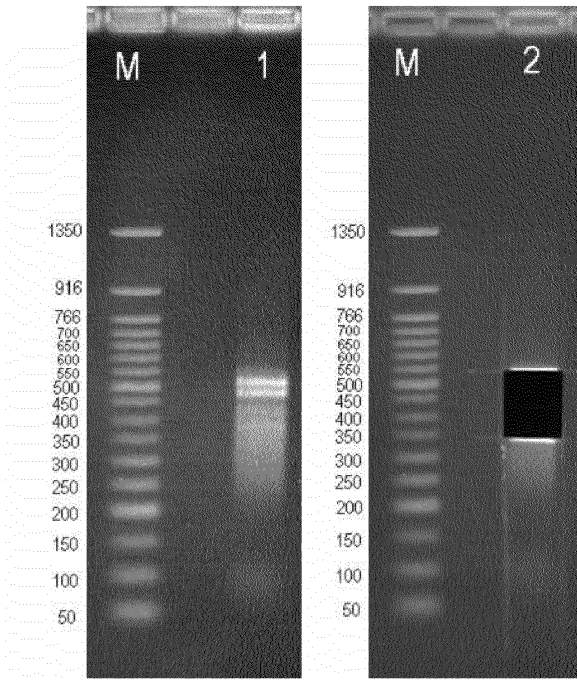
Фиг.9



Фиг.10

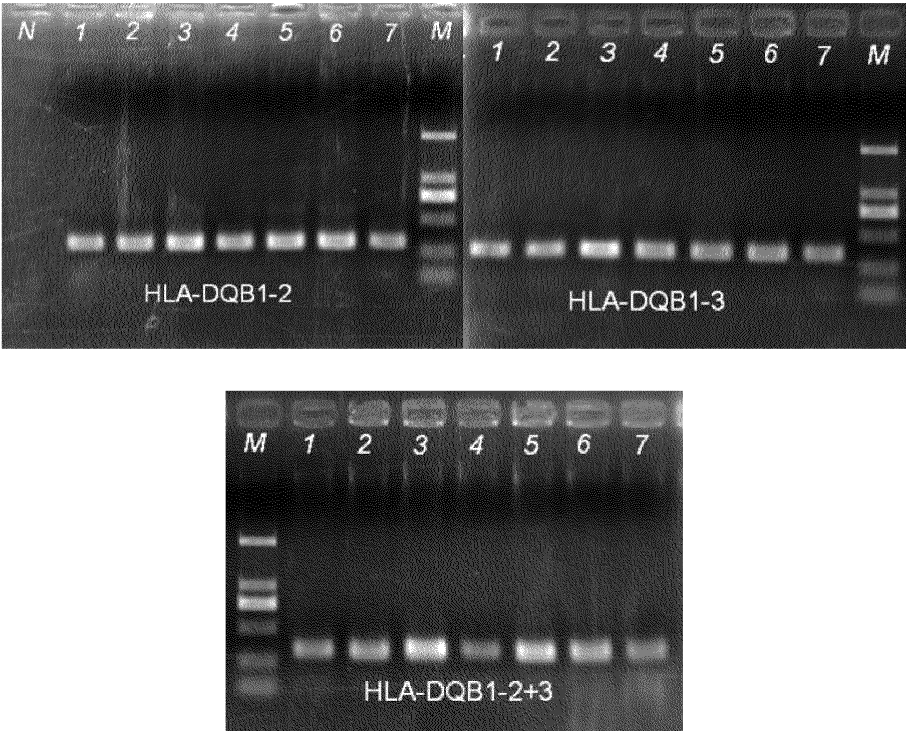


Фиг.11

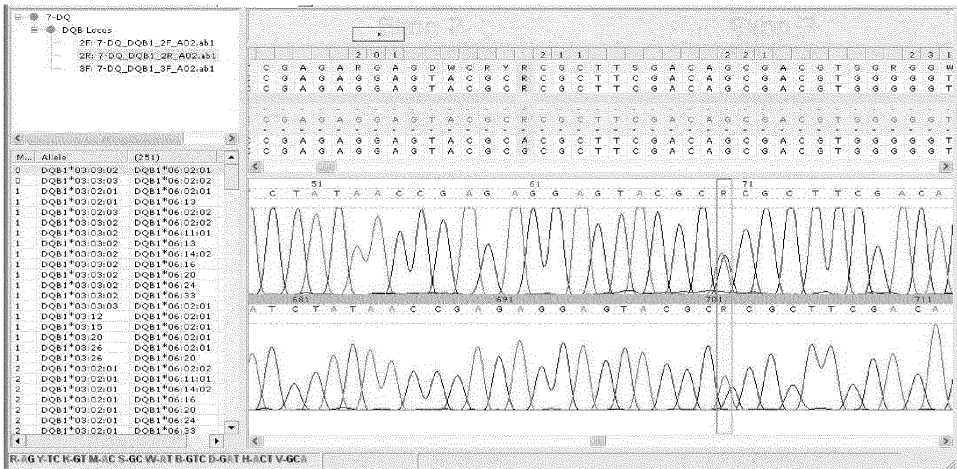


Фиг.12

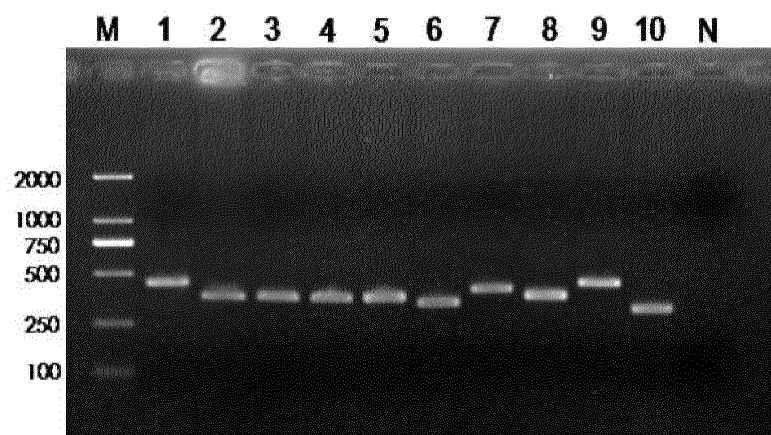




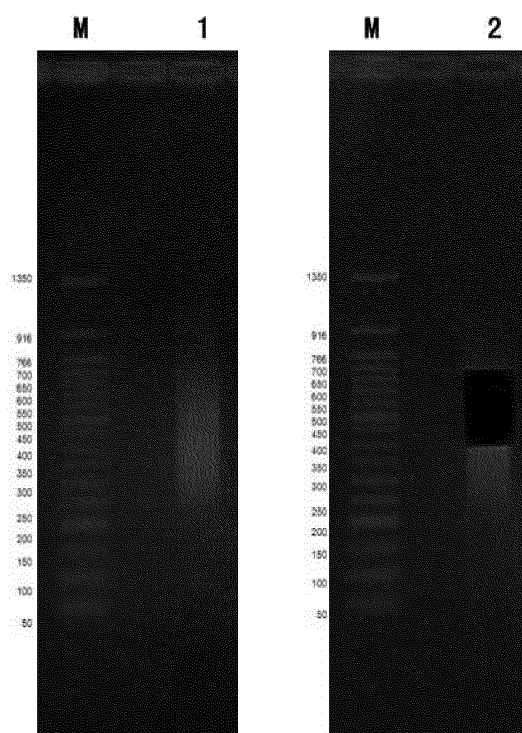
Фиг.14



Фиг.15



Фиг.16



Фиг.17



СПИСОК ПОСЛЕДОВАТЕЛЬНОСТЕЙ

<110> БиджиАй ШЭНЬЧЖЭНЬ КО., ЛИМИТЕД
 <120> Новый способ ПЦР-секвенирования
 и его применение в генотипировании HLA
 <130> IEC110063
 <150> 201010213721.2
 <151> 2010-06-30
 <150> 201010213719.5
 <151> 2010-06-30
 <150> 201010213717.6
 <151> 2010-06-30
 <150> PCT/CN2010/002150
 <151> 2010-12-24
 <150> PCT/CN2010/002149
 <151> 2010-12-24
 <160> 240
 <170> PatentIn version 3.5
 <210> 1
 <211> 19
 <212> ДНК
 <213> искусственная
 <220>
 <223> последовательность праймера
 <400> 1
 cctctgyggg gagaagcaa 19
 <210> 2
 <211> 19
 <212> ДНК
 <213> искусственная
 <220>
 <223> последовательность праймера
 <400> 2
 atctcggacc cggagactg 19
 <210> 3
 <211> 19
 <212> ДНК
 <213> искусственная
 <220>
 <223> последовательность праймера
 <400> 3
 cggggccagg ttctcacac 19

<210> 4
 <211> 25
 <212> ДНК
 <213> искусственная

<220>
 <223> последовательность праймера

<400> 4
 ggygatattc tagtggttgt cccaa 25

<210> 5
 <211> 22
 <212> ДНК
 <213> искусственная

<220>
 <223> последовательность праймера

<400> 5
 gtgtcccatg acagatgcaa aa 22

<210> 6
 <211> 20
 <212> ДНК
 <213> искусственная

<220>
 <223> последовательность праймера

<400> 6
 ggccctgacc ctgctaaagg 20

<210> 7
 <211> 18
 <212> ДНК
 <213> искусственная

<220>
 <223> последовательность праймера

<400> 7
 aggagcgagg ggaccgca 18

<210> 8
 <211> 18
 <212> ДНК
 <213> искусственная

<220>
 <223> последовательность праймера

<400> 8
 cgggccgggg tcactcac 18

<210> 9
 <211> 18

<212> ДНК
 <213> искусственная

<220>
 <223> последовательность праймера

<400> 9
 cggggccagg gtctcaca 18

<210> 10
 <211> 18
 <212> ДНК
 <213> искусственная

<220>
 <223> последовательность праймера

<400> 10
 gagggccatcc ccggcgac 18

<210> 11
 <211> 22
 <212> ДНК
 <213> искусственная

<220>
 <223> последовательность праймера

<400> 11
 gctggtcaca tgggtggtcc ta 22

<210> 12
 <211> 22
 <212> ДНК
 <213> искусственная

<220>
 <223> последовательность праймера

<400> 12
 ctcccttacc catctcaggg tg 22

<210> 13
 <211> 19
 <212> ДНК
 <213> искусственная

<220>
 <223> последовательность праймера

<400> 13
 cctctgyggg gagaagcaa 19

<210> 14
 <211> 22
 <212> ДНК
 <213> искусственная

<220>
 <223> последовательность праймера

 <400> 14
 ggatctcgga cccggagact gt 22

 <210> 15
 <211> 18
 <212> ДНК
 <213> искусственная

 <220>
 <223> последовательность праймера

 <400> 15
 tgggctgacc gyggggtc 18

 <210> 16
 <211> 25
 <212> ДНК
 <213> искусственная

 <220>
 <223> последовательность праймера

 <400> 16
 ggygatattc tagtggtggt cccaa 25

 <210> 17
 <211> 22
 <212> ДНК
 <213> искусственная

 <220>
 <223> последовательность праймера

 <400> 17
 gtgtcccatk acagatgcaa aa 22

 <210> 18
 <211> 20
 <212> ДНК
 <213> искусственная

 <220>
 <223> последовательность праймера

 <400> 18
 ggccctgacc ctgctaaagg 20

 <210> 19
 <211> 18
 <212> ДНК
 <213> искусственная

 <220>
 <223> последовательность праймера

<400> 19 aggagcgagg ggaccgca	18
<210> 20 <211> 18 <212> ДНК <213> искусственная	
<220> <223> последовательность праймера	
<400> 20 cgggccgggg tcactcac	18
<210> 21 <211> 18 <212> ДНК <213> искусственная	
<220> <223> последовательность праймера	
<400> 21 сcaaaatccc cgggggtt	18
<210> 22 <211> 18 <212> ДНК <213> искусственная	
<220> <223> последовательность праймера	
<400> 22 gaggccatcc ccggcgac	18
<210> 23 <211> 22 <212> ДНК <213> искусственная	
<220> <223> последовательность праймера	
<400> 23 gctggtcaca tgggtggtcc ta	22
<210> 24 <211> 20 <212> ДНК <213> искусственная	
<220> <223> последовательность праймера	
<400> 24 tgaccctca tccccctct	20

<210> 25
 <211> 18
 <212> ДНК
 <213> искусственная

 <220>
 <223> последовательность праймера

 <400> 25
 gacccgggga gccgcgca 18

 <210> 26
 <211> 19
 <212> ДНК
 <213> искусственная

 <220>
 <223> последовательность праймера

 <400> 26
 tcgaggggtct gggcgggtt 19

 <210> 27
 <211> 26
 <212> ДНК
 <213> искусственная

 <220>
 <223> последовательность праймера

 <400> 27
 cctttacccg gtttcatttt crgttt 26

 <210> 28
 <211> 21
 <212> ДНК
 <213> искусственная

 <220>
 <223> последовательность праймера

 <400> 28
 ctacgggaga tggggaaggc t 21

 <210> 29
 <211> 25
 <212> ДНК
 <213> искусственная

 <220>
 <223> последовательность праймера

 <400> 29
 gtgtcgcaag agagatrcaa agtgt 25

 <210> 30
 <211> 22

<212> ДНК
 <213> искусственная

<220>
 <223> последовательность праймера

<400> 30
 gctctgggaa aggaggrgaa gg 22

<210> 31
 <211> 18
 <212> ДНК
 <213> искусственная

<220>
 <223> последовательность праймера

<400> 31
 gacccgggga gccgcgca 18

<210> 32
 <211> 19
 <212> ДНК
 <213> искусственная

<220>
 <223> последовательность праймера

<400> 32
 tcgaggggtct gggcgggtt 19

<210> 33
 <211> 19
 <212> ДНК
 <213> искусственная

<220>
 <223> последовательность праймера

<400> 33
 gccsagaccc tcgrccgga 19

<210> 34
 <211> 21
 <212> ДНК
 <213> искусственная

<220>
 <223> последовательность праймера

<400> 34
 agatrgggaa ggctccccac t 21

<210> 35
 <211> 21
 <212> ДНК
 <213> искусственная

<220>
 <223> последовательность праймера

 <400> 35
 tctcaggatr gtcacatggg c 21

<210> 36
 <211> 22
 <212> ДНК
 <213> искусственная

<220>
 <223> последовательность праймера

 <400> 36
 gctctgggaa argaggrgaa gg 22

<210> 37
 <211> 21
 <212> ДНК
 <213> искусственная

<220>
 <223> последовательность праймера

 <400> 37
 gattccycgc agaggatttc g 21

<210> 38
 <211> 21
 <212> ДНК
 <213> искусственная

<220>
 <223> последовательность праймера

 <400> 38
 aggggcracs acgctcacct c 21

<210> 39
 <211> 22
 <212> ДНК
 <213> искусственная

<220>
 <223> последовательность праймера

 <400> 39
 cctgtctgtt actgccctca gt 22

<210> 40
 <211> 25
 <212> ДНК
 <213> искусственная

<220>
 <223> последовательность праймера

<400> 40 ggcccatagt aacagaaact caata	25
<210> 41 <211> 12 <212> ДНК <213> искусственная	
<220> <223> праймерный индекс	
<400> 41 tcgcagacat ca	12
<210> 42 <211> 12 <212> ДНК <213> искусственная	
<220> <223> праймерный индекс	
<400> 42 tgacacgatg ct	12
<210> 43 <211> 12 <212> ДНК <213> искусственная	
<220> <223> праймерный индекс	
<400> 43 tacatcgcac ta	12
<210> 44 <211> 12 <212> ДНК <213> искусственная	
<220> <223> праймерный индекс	
<400> 44 tacagatgct ga	12
<210> 45 <211> 12 <212> ДНК <213> искусственная	
<220> <223> праймерный индекс	
<400> 45 ctcgatgagt ac	12

<210> 46
 <211> 12
 <212> ДНК
 <213> искусственная

<220>
 <223> праймерный индекс

<400> 46
 acgtctagac ac 12

<210> 47
 <211> 12
 <212> ДНК
 <213> искусственная

<220>
 <223> праймерный индекс

<400> 47
 tctgtatact ca 12

<210> 48
 <211> 12
 <212> ДНК
 <213> искусственная

<220>
 <223> праймерный индекс

<400> 48
 tgctgtagtg ac 12

<210> 49
 <211> 12
 <212> ДНК
 <213> искусственная

<220>
 <223> праймерный индекс

<400> 49
 tatctgctca ta 12

<210> 50
 <211> 12
 <212> ДНК
 <213> искусственная

<220>
 <223> праймерный индекс

<400> 50
 agatatcgag ct 12

<210> 51
 <211> 12

<212>	ДНК	
<213>	искусственная	
<220>		
<223>	праймерный индекс	
<400>	51	
tacatgctga	gc	12
<210>	52	
<211>	12	
<212>	ДНК	
<213>	искусственная	
<220>		
<223>	праймерный индекс	
<400>	52	
acgtgtctat	ca	12
<210>	53	
<211>	12	
<212>	ДНК	
<213>	искусственная	
<220>		
<223>	праймерный индекс	
<400>	53	
tcatatcgcg	at	12
<210>	54	
<211>	12	
<212>	ДНК	
<213>	искусственная	
<220>		
<223>	праймерный индекс	
<400>	54	
agatcgata	gc	12
<210>	55	
<211>	12	
<212>	ДНК	
<213>	искусственная	
<220>		
<223>	праймерный индекс	
<400>	55	
acagatgcac	gc	12
<210>	56	
<211>	12	
<212>	ДНК	
<213>	искусственная	

12

<220>
 <223> праймерный индекс

 <400> 56
 atctcgtgac ag 12

<210> 57
 <211> 12
 <212> ДНК
 <213> искусственная

<220>
 <223> праймерный индекс

 <400> 57
 tagatcgtac at 12

<210> 58
 <211> 12
 <212> ДНК
 <213> искусственная

<220>
 <223> праймерный индекс

 <400> 58
 actagtagac gc 12

<210> 59
 <211> 12
 <212> ДНК
 <213> искусственная

<220>
 <223> праймерный индекс

 <400> 59
 actacacgtc tc 12

<210> 60
 <211> 12
 <212> ДНК
 <213> искусственная

<220>
 <223> праймерный индекс

 <400> 60
 atagtcacgc gt 12

<210> 61
 <211> 12
 <212> ДНК
 <213> искусственная

<220>
 <223> праймерный индекс

<400> 61 agactcgcgt at	12
<210> 62 <211> 12 <212> ДНК <213> искусственная	
<220> <223> праймерный индекс	
<400> 62 tactagctga cg	12
<210> 63 <211> 12 <212> ДНК <213> искусственная	
<220> <223> праймерный индекс	
<400> 63 atactagtgc tc	12
<210> 64 <211> 12 <212> ДНК <213> искусственная	
<220> <223> праймерный индекс	
<400> 64 tgtatcgtgc tc	12
<210> 65 <211> 12 <212> ДНК <213> искусственная	
<220> <223> праймерный индекс	
<400> 65 cacgatgaca tc	12
<210> 66 <211> 12 <212> ДНК <213> искусственная	
<220> <223> праймерный индекс	
<400> 66 tagtgagcgc ac	12

<210> 67
 <211> 12
 <212> ДНК
 <213> искусственная

<220>
 <223> праймерный индекс

<400> 67
 tgctgtctcg ag 12

<210> 68
 <211> 12
 <212> ДНК
 <213> искусственная

<220>
 <223> праймерный индекс

<400> 68
 catagcagtg tc 12

<210> 69
 <211> 12
 <212> ДНК
 <213> искусственная

<220>
 <223> праймерный индекс

<400> 69
 tgtgctcgag tc 12

<210> 70
 <211> 12
 <212> ДНК
 <213> искусственная

<220>
 <223> праймерный индекс

<400> 70
 tctgatcgag ca 12

<210> 71
 <211> 12
 <212> ДНК
 <213> искусственная

<220>
 <223> праймерный индекс

<400> 71
 cactcgtaca tc 12

<210> 72
 <211> 12

<212> ДНК
 <213> искусственная

 <220>
 <223> праймерный индекс

 <400> 72
 agcgatgctc at 12

 <210> 73
 <211> 12
 <212> ДНК
 <213> искусственная

 <220>
 <223> праймерный индекс

 <400> 73
 cgacgtgctc gc 12

 <210> 74
 <211> 12
 <212> ДНК
 <213> искусственная

 <220>
 <223> праймерный индекс

 <400> 74
 cgcgtactgc ag 12

 <210> 75
 <211> 12
 <212> ДНК
 <213> искусственная

 <220>
 <223> праймерный индекс

 <400> 75
 acgcatctat ac 12

 <210> 76
 <211> 12
 <212> ДНК
 <213> искусственная

 <220>
 <223> праймерный индекс

 <400> 76
 ctagtatcgc ag 12

 <210> 77
 <211> 12
 <212> ДНК
 <213> искусственная

<220>
 <223> праймерный индекс

 <400> 77
 cgagatgact ct 12

<210> 78
 <211> 12
 <212> ДНК
 <213> искусственная

<220>
 <223> праймерный индекс

 <400> 78
 tgtatacacg at 12

<210> 79
 <211> 12
 <212> ДНК
 <213> искусственная

<220>
 <223> праймерный индекс

 <400> 79
 actgtctcga gc 12

<210> 80
 <211> 12
 <212> ДНК
 <213> искусственная

<220>
 <223> праймерный индекс

 <400> 80
 acgtagcgca ca 12

<210> 81
 <211> 12
 <212> ДНК
 <213> искусственная

<220>
 <223> праймерный индекс

 <400> 81
 catctgctat ag 12

<210> 82
 <211> 12
 <212> ДНК
 <213> искусственная

<220>
 <223> праймерный индекс

<400> 82 tctagctcat ga	12
<210> 83 <211> 12 <212> ДНК <213> искусственная	
<220> <223> праймерный индекс	
<400> 83 acgcactcta ga	12
<210> 84 <211> 12 <212> ДНК <213> искусственная	
<220> <223> праймерный индекс	
<400> 84 ctatgcactg at	12
<210> 85 <211> 12 <212> ДНК <213> искусственная	
<220> <223> праймерный индекс	
<400> 85 tgagatacag ta	12
<210> 86 <211> 12 <212> ДНК <213> искусственная	
<220> <223> праймерный индекс	
<400> 86 atctgctatg ac	12
<210> 87 <211> 12 <212> ДНК <213> искусственная	
<220> <223> праймерный индекс	
<400> 87 actcatcgtg ct	12

<210> 88
 <211> 12
 <212> ДНК
 <213> искусственная

<220>
 <223> праймерный индекс

<400> 88
 tagagctgtc ac 12

<210> 89
 <211> 12
 <212> ДНК
 <213> искусственная

<220>
 <223> праймерный индекс

<400> 89
 tacactgtct at 12

<210> 90
 <211> 12
 <212> ДНК
 <213> искусственная

<220>
 <223> праймерный индекс

<400> 90
 cagcacatag at 12

<210> 91
 <211> 12
 <212> ДНК
 <213> искусственная

<220>
 <223> праймерный индекс

<400> 91
 cacagtactc gc 12

<210> 92
 <211> 12
 <212> ДНК
 <213> искусственная

<220>
 <223> праймерный индекс

<400> 92
 ctgctagtgt at 12

<210> 93
 <211> 12

<212>	ДНК	
<213>	искусственная	
<220>		
<223>	праймерный индекс	
<400>	93	
	tgtactatca ta	12
<210>	94	
<211>	12	
<212>	ДНК	
<213>	искусственная	
<220>		
<223>	праймерный индекс	
<400>	94	
	tgtgatagac ac	12
<210>	95	
<211>	12	
<212>	ДНК	
<213>	искусственная	
<220>		
<223>	праймерный индекс	
<400>	95	
	ctagtactga cg	12
<210>	96	
<211>	12	
<212>	ДНК	
<213>	искусственная	
<220>		
<223>	праймерный индекс	
<400>	96	
	agcgagtcta ct	12
<210>	97	
<211>	12	
<212>	ДНК	
<213>	искусственная	
<220>		
<223>	праймерный индекс	
<400>	97	
	tagactgagc ta	12
<210>	98	
<211>	12	
<212>	ДНК	
<213>	искусственная	

<220>
 <223> праймерный индекс

 <400> 98
 acatactgag ac 12

 <210> 99
 <211> 12
 <212> ДНК
 <213> искусственная

 <220>
 <223> праймерный индекс

 <400> 99
 cagacgcgtg ag 12

 <210> 100
 <211> 12
 <212> ДНК
 <213> искусственная

 <220>
 <223> праймерный индекс

 <400> 100
 tacatctcgt at 12

 <210> 101
 <211> 12
 <212> ДНК
 <213> искусственная

 <220>
 <223> праймерный индекс

 <400> 101
 cgcgacatca cg 12

 <210> 102
 <211> 12
 <212> ДНК
 <213> искусственная

 <220>
 <223> праймерный индекс

 <400> 102
 tagcgatgag ac 12

 <210> 103
 <211> 12
 <212> ДНК
 <213> искусственная

 <220>
 <223> праймерный индекс

<400> 103 acactcatag at	12
<210> 104 <211> 12 <212> ДНК <213> искусственная	
<220> <223> праймерный индекс	
<400> 104 ctatcatgac ac	12
<210> 105 <211> 12 <212> ДНК <213> искусственная	
<220> <223> праймерный индекс	
<400> 105 agcgtatact ag	12
<210> 106 <211> 12 <212> ДНК <213> искусственная	
<220> <223> праймерный индекс	
<400> 106 catactcacg ta	12
<210> 107 <211> 12 <212> ДНК <213> искусственная	
<220> <223> праймерный индекс	
<400> 107 tgtcgtgcta tc	12
<210> 108 <211> 12 <212> ДНК <213> искусственная	
<220> <223> праймерный индекс	
<400> 108 acatgactca cg	12

<210> 109
 <211> 12
 <212> ДНК
 <213> искусственная

<220>
 <223> праймерный индекс

<400> 109
 cgctagactg ta 12

<210> 110
 <211> 12
 <212> ДНК
 <213> искусственная

<220>
 <223> праймерный индекс

<400> 110
 tactatagtc ga 12

<210> 111
 <211> 12
 <212> ДНК
 <213> искусственная

<220>
 <223> праймерный индекс

<400> 111
 acagtgtagc gc 12

<210> 112
 <211> 12
 <212> ДНК
 <213> искусственная

<220>
 <223> праймерный индекс

<400> 112
 tgatatgcta ca 12

<210> 113
 <211> 12
 <212> ДНК
 <213> искусственная

<220>
 <223> праймерный индекс

<400> 113
 cactctatcg ac 12

<210> 114
 <211> 12

<212> ДНК
 <213> искусственная

 <220>
 <223> праймерный индекс

 <400> 114
 tcacgcgatg ag 12

<210> 115
 <211> 12
 <212> ДНК
 <213> искусственная

 <220>
 <223> праймерный индекс

 <400> 115
 acactctagt ca 12

<210> 116
 <211> 12
 <212> ДНК
 <213> искусственная

 <220>
 <223> праймерный индекс

 <400> 116
 acgtagatct at 12

<210> 117
 <211> 12
 <212> ДНК
 <213> искусственная

 <220>
 <223> праймерный индекс

 <400> 117
 catatgagat cg 12

<210> 118
 <211> 12
 <212> ДНК
 <213> искусственная

 <220>
 <223> праймерный индекс

 <400> 118
 agcagagtgc tc 12

<210> 119
 <211> 12
 <212> ДНК
 <213> искусственная

<220>
 <223> праймерный индекс

 <400> 119
 cagctatcat ac 12

<210> 120
 <211> 12
 <212> ДНК
 <213> искусственная

<220>
 <223> праймерный индекс

 <400> 120
 cactgcagac ga 12

<210> 121
 <211> 12
 <212> ДНК
 <213> искусственная

<220>
 <223> праймерный индекс

 <400> 121
 tatactctag at 12

<210> 122
 <211> 12
 <212> ДНК
 <213> искусственная

<220>
 <223> праймерный индекс

 <400> 122
 tgcatagagc gc 12

<210> 123
 <211> 12
 <212> ДНК
 <213> искусственная

<220>
 <223> праймерный индекс

 <400> 123
 tgtatgctcg tc 12

<210> 124
 <211> 12
 <212> ДНК
 <213> искусственная

<220>
 <223> праймерный индекс

<400> 124 tcgtgacaga tc	12
<210> 125 <211> 12 <212> ДНК <213> искусственная	
<220> <223> праймерный индекс	
<400> 125 tagtgatgct ct	12
<210> 126 <211> 12 <212> ДНК <213> искусственная	
<220> <223> праймерный индекс	
<400> 126 acgagctgat at	12
<210> 127 <211> 12 <212> ДНК <213> искусственная	
<220> <223> праймерный индекс	
<400> 127 agactctgag tc	12
<210> 128 <211> 12 <212> ДНК <213> искусственная	
<220> <223> праймерный индекс	
<400> 128 ctgatagtat ca	12
<210> 129 <211> 12 <212> ДНК <213> искусственная	
<220> <223> праймерный индекс	
<400> 129 ctcatagact ac	12

<210> 130
<211> 12
<212> ДНК
<213> искусственная

<220>
<223> праймерный индекс

<400> 130
atcgcgagtg ac 12

<210> 131
<211> 12
<212> ДНК
<213> искусственная

<220>
<223> праймерный индекс

<400> 131
tcgctcacta ca 12

<210> 132
<211> 12
<212> ДНК
<213> искусственная

<220>
<223> праймерный индекс

<400> 132
tgtctcgaca tc 12

<210> 133
<211> 12
<212> ДНК
<213> искусственная

<220>
<223> праймерный индекс

<400> 133
atagagtctc at 12

<210> 134
<211> 12
<212> ДНК
<213> искусственная

<220>
<223> праймерный индекс

<400> 134
cgcatagcgt at 12

<210> 135
<211> 12

<212>	ДНК	
<213>	искусственная	
<220>		
<223>	праймерный индекс	
<400>	135	
	cgagacactc gc	12
<210>	136	
<211>	12	
<212>	ДНК	
<213>	искусственная	
<220>		
<223>	праймерный индекс	
<400>	136	
	tcgtagtcta ca	12
<210>	137	
<211>	12	
<212>	ДНК	
<213>	искусственная	
<220>		
<223>	праймерный индекс	
<400>	137	
	cagcatacta tc	12
<210>	138	
<211>	12	
<212>	ДНК	
<213>	искусственная	
<220>		
<223>	праймерный индекс	
<400>	138	
	tcgtgataca ga	12
<210>	139	
<211>	12	
<212>	ДНК	
<213>	искусственная	
<220>		
<223>	праймерный индекс	
<400>	139	
	cagctatagt ct	12
<210>	140	
<211>	12	
<212>	ДНК	
<213>	искусственная	

<220>
 <223> праймерный индекс

 <400> 140
 atgcagatat ct 12

<210> 141
 <211> 12
 <212> ДНК
 <213> искусственная

<220>
 <223> праймерный индекс

 <400> 141
 tctatcgatg ca 12

<210> 142
 <211> 12
 <212> ДНК
 <213> искусственная

<220>
 <223> праймерный индекс

 <400> 142
 acacgcagat cg 12

<210> 143
 <211> 12
 <212> ДНК
 <213> искусственная

<220>
 <223> праймерный индекс

 <400> 143
 catgagtata gc 12

<210> 144
 <211> 12
 <212> ДНК
 <213> искусственная

<220>
 <223> праймерный индекс

 <400> 144
 ctagctgacg ta 12

<210> 145
 <211> 12
 <212> ДНК
 <213> искусственная

<220>
 <223> праймерный индекс

<400> 145 tagcatatcg ag	12
<210> 146 <211> 12 <212> ДНК <213> искусственная	
<220> <223> праймерный индекс	
<400> 146 tacacgtatg ag	12
<210> 147 <211> 12 <212> ДНК <213> искусственная	
<220> <223> праймерный индекс	
<400> 147 acgactcgct ac	12
<210> 148 <211> 12 <212> ДНК <213> искусственная	
<220> <223> праймерный индекс	
<400> 148 tcatgactag ta	12
<210> 149 <211> 12 <212> ДНК <213> искусственная	
<220> <223> праймерный индекс	
<400> 149 tagcatacac gc	12
<210> 150 <211> 12 <212> ДНК <213> искусственная	
<220> <223> праймерный индекс	
<400> 150 tgacgcgtat ac	12

<210> 151
 <211> 12
 <212> ДНК
 <213> искусственная

<220>
 <223> праймерный индекс

<400> 151
 cgtcatatgc ag 12

<210> 152
 <211> 12
 <212> ДНК
 <213> искусственная

<220>
 <223> праймерный индекс

<400> 152
 tatagcgatg ac 12

<210> 153
 <211> 12
 <212> ДНК
 <213> искусственная

<220>
 <223> праймерный индекс

<400> 153
 tgcagcgagt ac 12

<210> 154
 <211> 12
 <212> ДНК
 <213> искусственная

<220>
 <223> праймерный индекс

<400> 154
 tcgacgctag cg 12

<210> 155
 <211> 12
 <212> ДНК
 <213> искусственная

<220>
 <223> праймерный индекс

<400> 155
 cgtgtcgaca ga 12

<210> 156
 <211> 12

<212>	ДНК	
<213>	искусственная	
<220>		
<223>	праймерный индекс	
<400>	156	
	cagtcgtgag ca	12
<210>	157	
<211>	12	
<212>	ДНК	
<213>	искусственная	
<220>		
<223>	праймерный индекс	
<400>	157	
	actcgacgtg ag	12
<210>	158	
<211>	12	
<212>	ДНК	
<213>	искусственная	
<220>		
<223>	праймерный индекс	
<400>	158	
	acgcgagtga ta	12
<210>	159	
<211>	12	
<212>	ДНК	
<213>	искусственная	
<220>		
<223>	праймерный индекс	
<400>	159	
	actcgtctga cg	12
<210>	160	
<211>	12	
<212>	ДНК	
<213>	искусственная	
<220>		
<223>	праймерный индекс	
<400>	160	
	tgctatcact ga	12
<210>	161	
<211>	12	
<212>	ДНК	
<213>	искусственная	

<220>		
<223>	праймерный индекс	
<400>	161	
catactgtat	ct	12
<210>	162	
<211>	12	
<212>	ДНК	
<213>	искусственная	
<220>		
<223>	праймерный индекс	
<400>	162	
tacatagatg	tc	12
<210>	163	
<211>	12	
<212>	ДНК	
<213>	искусственная	
<220>		
<223>	праймерный индекс	
<400>	163	
tctactcgtg	ac	12
<210>	164	
<211>	12	
<212>	ДНК	
<213>	искусственная	
<220>		
<223>	праймерный индекс	
<400>	164	
cacgtatagt	ga	12
<210>	165	
<211>	12	
<212>	ДНК	
<213>	искусственная	
<220>		
<223>	праймерный индекс	
<400>	165	
ctgcactaga	ca	12
<210>	166	
<211>	12	
<212>	ДНК	
<213>	искусственная	
<220>		
<223>	праймерный индекс	

<400> 166 actcatatcg ca	12
<210> 167 <211> 12 <212> ДНК <213> искусственная	
<220> <223> праймерный индекс	
<400> 167 acacgagctc at	12
<210> 168 <211> 12 <212> ДНК <213> искусственная	
<220> <223> праймерный индекс	
<400> 168 cactcatatc ga	12
<210> 169 <211> 12 <212> ДНК <213> искусственная	
<220> <223> праймерный индекс	
<400> 169 tacagatagt ct	12
<210> 170 <211> 12 <212> ДНК <213> искусственная	
<220> <223> праймерный индекс	
<400> 170 tcgtctgtga ta	12
<210> 171 <211> 12 <212> ДНК <213> искусственная	
<220> <223> праймерный индекс	
<400> 171 tacactcgtg ct	12

<210> 172
 <211> 12
 <212> ДНК
 <213> искусственная

<220>
 <223> праймерный индекс

<400> 172
 tgacgctcat ct 12

<210> 173
 <211> 12
 <212> ДНК
 <213> искусственная

<220>
 <223> праймерный индекс

<400> 173
 tacatgtgac ga 12

<210> 174
 <211> 12
 <212> ДНК
 <213> искусственная

<220>
 <223> праймерный индекс

<400> 174
 tcgtacatgc tc 12

<210> 175
 <211> 12
 <212> ДНК
 <213> искусственная

<220>
 <223> праймерный индекс

<400> 175
 tgtatgatct cg 12

<210> 176
 <211> 12
 <212> ДНК
 <213> искусственная

<220>
 <223> праймерный индекс

<400> 176
 cactgtgctc at 12

<210> 177
 <211> 12

<212> ДНК
 <213> искусственная

 <220>
 <223> праймерный индекс

 <400> 177
 cagtacactc ta 12

 <210> 178
 <211> 12
 <212> ДНК
 <213> искусственная

 <220>
 <223> праймерный индекс

 <400> 178
 actgcatgat cg 12

 <210> 179
 <211> 12
 <212> ДНК
 <213> искусственная

 <220>
 <223> праймерный индекс

 <400> 179
 catactatca cg 12

 <210> 180
 <211> 12
 <212> ДНК
 <213> искусственная

 <220>
 <223> праймерный индекс

 <400> 180
 tcgtgtcact ac 12

 <210> 181
 <211> 12
 <212> ДНК
 <213> искусственная

 <220>
 <223> праймерный индекс

 <400> 181
 cactatacag at 12

 <210> 182
 <211> 12
 <212> ДНК
 <213> искусственная

<220>
 <223> праймерный индекс

 <400> 182
 cgacacgtac ta 12

<210> 183
 <211> 12
 <212> ДНК
 <213> искусственная

<220>
 <223> праймерный индекс

 <400> 183
 atatcgtagc at 12

<210> 184
 <211> 12
 <212> ДНК
 <213> искусственная

<220>
 <223> праймерный индекс

 <400> 184
 tcgtgatcac ta 12

<210> 185
 <211> 12
 <212> ДНК
 <213> искусственная

<220>
 <223> праймерный индекс

 <400> 185
 tagtctatac at 12

<210> 186
 <211> 12
 <212> ДНК
 <213> искусственная

<220>
 <223> праймерный индекс

 <400> 186
 agacgctgtc ga 12

<210> 187
 <211> 12
 <212> ДНК
 <213> искусственная

<220>
 <223> праймерный индекс

<400> 187 tgtcacagtg ac	12
<210> 188 <211> 12 <212> ДНК <213> искусственная	
<220> <223> праймерный индекс	
<400> 188 tcatatgatc ga	12
<210> 189 <211> 12 <212> ДНК <213> искусственная	
<220> <223> праймерный индекс	
<400> 189 atcgactatg ct	12
<210> 190 <211> 12 <212> ДНК <213> искусственная	
<220> <223> праймерный индекс	
<400> 190 cgatcatatg ag	12
<210> 191 <211> 12 <212> ДНК <213> искусственная	
<220> <223> праймерный индекс	
<400> 191 atactagcat ca	12
<210> 192 <211> 12 <212> ДНК <213> искусственная	
<220> <223> праймерный индекс	
<400> 192 tcatgctgac ga	12

<210> 193
 <211> 12
 <212> ДНК
 <213> искусственная

<220>
 <223> праймерный индекс

<400> 193
 cactgacgct ca 12

<210> 194
 <211> 12
 <212> ДНК
 <213> искусственная

<220>
 <223> праймерный индекс

<400> 194
 cactacatcg ct 12

<210> 195
 <211> 12
 <212> ДНК
 <213> искусственная

<220>
 <223> праймерный индекс

<400> 195
 tcgctcatct at 12

<210> 196
 <211> 12
 <212> ДНК
 <213> искусственная

<220>
 <223> праймерный индекс

<400> 196
 tagtacagag ct 12

<210> 197
 <211> 12
 <212> ДНК
 <213> искусственная

<220>
 <223> праймерный индекс

<400> 197
 tgtatcacga gc 12

<210> 198
 <211> 12

<212> ДНК	
<213> искусственная	
<220>	
<223> праймерный индекс	
<400> 198	
atgatcgtat ac	12
<210> 199	
<211> 12	
<212> ДНК	
<213> искусственная	
<220>	
<223> праймерный индекс	
<400> 199	
tactgctatc tc	12
<210> 200	
<211> 12	
<212> ДНК	
<213> искусственная	
<220>	
<223> праймерный индекс	
<400> 200	
cgctgcatag cg	12
<210> 201	
<211> 12	
<212> ДНК	
<213> искусственная	
<220>	
<223> праймерный индекс	
<400> 201	
cgcgagctcg tc	12
<210> 202	
<211> 12	
<212> ДНК	
<213> искусственная	
<220>	
<223> праймерный индекс	
<400> 202	
actcgatgag ct	12
<210> 203	
<211> 12	
<212> ДНК	
<213> искусственная	

<220>
 <223> праймерный индекс

 <400> 203
 tagagtctgt at 12

<210> 204
 <211> 12
 <212> ДНК
 <213> искусственная

<220>
 <223> праймерный индекс

 <400> 204
 tgtctatcac at 12

<210> 205
 <211> 12
 <212> ДНК
 <213> искусственная

<220>
 <223> праймерный индекс

 <400> 205
 tactatcgct ct 12

<210> 206
 <211> 12
 <212> ДНК
 <213> искусственная

<220>
 <223> праймерный индекс

 <400> 206
 tatgtgacat ac 12

<210> 207
 <211> 12
 <212> ДНК
 <213> искусственная

<220>
 <223> праймерный индекс

 <400> 207
 tagatgacgc tc 12

<210> 208
 <211> 12
 <212> ДНК
 <213> искусственная

<220>
 <223> праймерный индекс

<400> 208 tactcgtagc gc	12
<210> 209 <211> 12 <212> ДНК <213> искусственная	
<220> <223> праймерный индекс	
<400> 209 tcgctgaca tc	12
<210> 210 <211> 12 <212> ДНК <213> искусственная	
<220> <223> праймерный индекс	
<400> 210 atctactgac gt	12
<210> 211 <211> 12 <212> ДНК <213> искусственная	
<220> <223> праймерный индекс	
<400> 211 acacgctcta ct	12
<210> 212 <211> 12 <212> ДНК <213> искусственная	
<220> <223> праймерный индекс	
<400> 212 acagtagcgc ac	12
<210> 213 <211> 12 <212> ДНК <213> искусственная	
<220> <223> праймерный индекс	
<400> 213 tacatagtct cg	12

<210> 214
 <211> 12
 <212> ДНК
 <213> искусственная

<220>
 <223> праймерный индекс

<400> 214
 ctagtatcat ga 12

<210> 215
 <211> 12
 <212> ДНК
 <213> искусственная

<220>
 <223> праймерный индекс

<400> 215
 tgagtagcac gc 12

<210> 216
 <211> 12
 <212> ДНК
 <213> искусственная

<220>
 <223> праймерный индекс

<400> 216
 tcgatcatgc ag 12

<210> 217
 <211> 12
 <212> ДНК
 <213> искусственная

<220>
 <223> праймерный индекс

<400> 217
 tagatgctat ac 12

<210> 218
 <211> 12
 <212> ДНК
 <213> искусственная

<220>
 <223> праймерный индекс

<400> 218
 tacatgcact ca 12

<210> 219
 <211> 12

<212>	ДНК	
<213>	искусственная	
<220>		
<223>	праймерный индекс	
<400>	219	
	atcgatgtca cg	12
<210>	220	
<211>	12	
<212>	ДНК	
<213>	искусственная	
<220>		
<223>	праймерный индекс	
<400>	220	
	cagctcgact ac	12
<210>	221	
<211>	12	
<212>	ДНК	
<213>	искусственная	
<220>		
<223>	праймерный индекс	
<400>	221	
	atcatatgta gc	12
<210>	222	
<211>	12	
<212>	ДНК	
<213>	искусственная	
<220>		
<223>	праймерный индекс	
<400>	222	
	ctctacagtc ac	12
<210>	223	
<211>	12	
<212>	ДНК	
<213>	искусственная	
<220>		
<223>	праймерный индекс	
<400>	223	
	tagcatcgat at	12
<210>	224	
<211>	12	
<212>	ДНК	
<213>	искусственная	

<220>
 <223> праймерный индекс

 <400> 224
 agatagcaca tc 12

<210> 225
 <211> 12
 <212> ДНК
 <213> искусственная

<220>
 <223> праймерный индекс

 <400> 225
 tgatcgacgc tc 12

<210> 226
 <211> 12
 <212> ДНК
 <213> искусственная

<220>
 <223> праймерный индекс

 <400> 226
 ctagatatcg tc 12

<210> 227
 <211> 12
 <212> ДНК
 <213> искусственная

<220>
 <223> праймерный индекс

 <400> 227
 tgcagctcat ag 12

<210> 228
 <211> 12
 <212> ДНК
 <213> искусственная

<220>
 <223> праймерный индекс

 <400> 228
 tacagactgc ac 12

<210> 229
 <211> 12
 <212> ДНК
 <213> искусственная

<220>
 <223> праймерный индекс

<400> 229 cgacgtagag tc	12
<210> 230 <211> 12 <212> ДНК <213> искусственная	
<220> <223> праймерный индекс	
<400> 230 cagtagcact ac	12
<210> 231 <211> 21 <212> ДНК <213> искусственная	
<220> <223> последовательность праймера	
<400> 231 cacgtttctt ggagtactct a	21
<210> 232 <211> 22 <212> ДНК <213> искусственная	
<220> <223> последовательность праймера	
<400> 232 gtttcttgtg gcagcttaag tt	22
<210> 233 <211> 20 <212> ДНК <213> искусственная	
<220> <223> последовательность праймера	
<400> 233 cctgtggcag ggtaagtata	20
<210> 234 <211> 22 <212> ДНК <213> искусственная	
<220> <223> последовательность праймера	
<400> 234 gtttcttgaa gcaggataag tt	22

<210> 235
 <211> 19
 <212> ДНК
 <213> искусственная

<220>
 <223> последовательность праймера

<400> 235
 gcacgtttct tggaggagg 19

<210> 236
 <211> 20
 <212> ДНК
 <213> искусственная

<220>
 <223> последовательность праймера

<400> 236
 tttcctgtgg cagcctaaga 20

<210> 237
 <211> 21
 <212> ДНК
 <213> искусственная

<220>
 <223> последовательность праймера

<400> 237
 gtttcttgga gcaggtaaa c 21

<210> 238
 <211> 19
 <212> ДНК
 <213> искусственная

<220>
 <223> последовательность праймера

<400> 238
 cctcacctcg ccgctgcac 19

<210> 239
 <211> 12
 <212> ДНК
 <213> искусственная

<220>
 <223> праймерный индекс

<400> 239
 cactgtatag ct 12

<210> 240
 <211> 12

47

<212> ДНК
<213> искусственная

<220>
<223> праймерный индекс

<400> 240
cgactagtac ta

12