



(19) RU (11) 2 235 773 (13) C2
(51) МПК⁷ С 12 Н 9/12, 15/54, 1/21, С 12
Q 1/68, С 12 Р 21/02

РОССИЙСКОЕ АГЕНТСТВО
ПО ПАТЕНТАМ И ТОВАРНЫМ ЗНАКАМ

(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ПАТЕНТУ РОССИЙСКОЙ ФЕДЕРАЦИИ

(21), (22) Заявка: 97113529/13, 06.08.1997
(24) Дата начала действия патента:
06.08.1997 i.i. 1-20
(30) Приоритет: 06.08.1996 US 60-023.376
(43) Дата публикации заявки: 20.07.1999
(46) Дата публикации: 10.09.2004
(56) Ссылки: EP 0655506 A1, 31.05.1995. EP 0699760 A, 06.03.1996. EP 0701000 A,
13.03.1996. WO 92/09689 A, 11.06.1992. WO 92/03556 A, 05.03.1992.
(98) Адрес для переписки:
101000, Москва, М.Златоустинский пер., 10,
кв.15, "ЕВРОМАРКПАТ", пат.пov.
И.А.Веселицкой

(72) Изобретатель: ГЕЛФЕНД Дэвид Хэрроу (US),
КЭЛМЭН Лиза Вивиан (US), РАЙХЕРТ Фред
Лоренс (US)
(73) Патентообладатель:
Ф. ХОФФМАНН-ЛЯ РОШ АГ (CH)
(74) Патентный поверенный:
Веселицкая Ирина Александровна

(54) МОДИФИЦИРОВАННАЯ ТЕРМОСТАБИЛЬНАЯ ДНК-ПОЛИМЕРАЗА, СПОСОБ ЕЁ ПОЛУЧЕНИЯ И
ПРИМЕНЕНИЕ

(57)
Изобретение относится к области молекулярной биологии и генной инженерии и может быть использовано в различных методах анализа нуклеиновых кислот, связанных с синтезом комплементарных ДНК-последовательностей. Предложена модифицированная форма термостабильной ДНК-полимеразы, получаемая путем замены остатка глутамата в консервативной последовательности SerGlnIleGluLeuArgVal/Ile природного термостабильного фермента другим аминокислотным остатком и отбора мутеинов, проявляющих повышенную способность к включению в реакцию

несвойственных субстратов, предпочтительно рНТФ. Описаны кодирующие модифицированную ДНК-полимеразу нуклеотидные последовательности, содержащие их векторы и штаммы E.coli, трансформированные указанными векторами и обеспечивающие получение рекомбинантных форм мутантных ферментов. Предлагаемое применение новой термостабильной ДНК-полимеразы, в том числе в составе наборов и композиций, в реакции секвенирования нуклеиновой кислоты позволяет использовать новые подходы при анализе результатов исследования. 9 с. и 11 з.п. ф-лы, 4 табл.

R
U
2
2
3
5
7
7
C
2

? 2 3 5 7 7 3 C 2



(19) RU (11) 2 235 773 (13) C2
(51) Int. Cl. ⁷ C 12 N 9/12, 15/54, 1/21, C
12 Q 1/68, C 12 P 21/02

RUSSIAN AGENCY
FOR PATENTS AND TRADEMARKS

(12) ABSTRACT OF INVENTION

(21), (22) Application: 97113529/13, 06.08.1997

(24) Effective date for property rights:
06.08.1997 cl. 1-20

(30) Priority: 06.08.1996 US 60-023.376

(43) Application published: 20.07.1999

(46) Date of publication: 10.09.2004

(98) Mail address:
101000, Moskva, M.Zlatoustinskij per., 10,
kv.15, "EVROMARKPAT", pat.pov. I.A.Veselitskoj

(72) Inventor: GELFEND Devid Khehrrou (US),
KEhLMEhN Liza Vivian (US), RAJKhERT Fred
Lorens (US)

(73) Proprietor:
F. KhOFFMANN-LJa ROSh AG (CH)

(74) Representative:
Veselitskaja Irina Aleksandrovna

(54) MODIFIED THERMOSTABLE DNA POLYMERASE, METHOD FOR IT PREPARING AND APPLICATION

(57) Abstract:

FIELD: molecular biology, genetic engineering, biochemistry.

SUBSTANCE: invention can be used in different methods for analysis of nucleic acids associated with synthesis of complementary DNA sequences. Invention proposes the modified form of thermostable DNA polymerase preparing by replacing glutamate residue in the consensus sequence Ser-Gln-Ile-Glu-Leu-Arg-Val/Ile of the natural thermostable enzyme for other amino acid residue and selection of mutants eliciting enhanced ability for incorporation unusual substrate in reaction, preferably

r-NTP (ribonucleoside triphosphates). Invention describes nucleotide sequences encoding the modified DNA polymerase, vector comprising thereof and strains E. coli transformed with indicated vectors and providing preparing recombinant forms of mutant enzymes. The proposed application of a new thermostable DNA polymerase, among them, as a component of sets and compositions in the sequencing reaction allows using new approaches in analysis of investigation data.

EFFECT: valuable biological and biochemical properties of DNA polymerase.

20 cl, 4 tbl, 6 ex

R
U
2
2
3
5
7
7
3
C
2

R
U
?
2
3
5
7
7
3
C
2

RU 2235773 C2

Область техники, к которой относится изобретение

Настоящее изобретение относится к термостабильным ДНК-полимеразам, способным более эффективно включать рибонуклеозидтрифосфаты. В изобретении описаны методы и способы выделения таких полимераз. Ферменты по изобретению могут применяться для различных целей, в частности для секвенирования нукleinовых кислот. Таким образом, изобретение также относится к улучшенным методам анализа последовательности нукleinовых кислот.

Предпосылки создания изобретения

Секвенирование ДНК обычно включает получение четырех популяций фрагментов одноцепочечной ДНК, имеющих одно определенное окончание и одно вариабельное окончание. Вариабельное окончание обычно заканчивается на конкретных основаниях нуклеотидов [на гуанине (Г), аденине (А), тимидине (Т) либо на цитозине (Ц)]. Каждый из четырех различных наборов фрагментов разделяют на основе их длины. С этой целью используют поликарбамидный гель с высокой разрешающей способностью. Каждая полоса на таких гелях соответствует конкретному нуклеотиду в последовательности ДНК, показывая тем самым его положение в последовательности.

Часто применяемым методом секвенирования ДНК является дидезокси- или цепь-терминирующий метод секвенирования, который включает ферментативный синтез цепи ДНК (Sanger и др., 1977, Proc. Natl. Acad. Sci. 74: 5463). Обычно осуществляют четыре различные реакции синтеза, каждая из которых должна закончиться на определенном основании (Г, А, Т или Ц) путем включения соответствующего терминирующего цепь нуклеотида, такого как дидезоксинуклеотид. Продукты реакции легко определить, поскольку каждая полоса соответствует только Г, А, Т либо Ц.

В дидезокси- или цепь-терминирующем методе отжигают короткий одноцепочечный праймер, соответствующий одноцепочечной матрице. Праймер удлиняют на его 3'-конце путем включения дезоксинуклеотидов (дДНТФ) до включения дидезоксинуклеотида (дДНТФ). После включения дДНТФ удлинение прекращается. Однако для гарантии правильности репликации ДНК ДНК-полимеразы обладают более выраженной способностью включать естественные для них субстраты, т.е. дДНТФ, и не включать аналоги нуклеотидов, называемые несвойственными нуклеотидами. При синтезе ДНК рибонуклеотиды (рНТФ) рассматриваются как несвойственные нуклеотиды, поскольку в отличие от дДНТФ рНТФ обычно не являются естественными субстратами для ДНК-полимеразы *in vivo*. В клетке это свойство способствует снижению включения аномальных оснований, таких как дезоксиинозинтрифосфат (дИТФ) или рНТФ, в растущую цепь ДНК.

Двумя наиболее часто используемыми методами автоматического секвенирования являются секвенирование с использованием окрашенного праймера и окрашенного терминатора. Эти методы пригодны для использования с флуоресцентно меченными фрагментами. Хотя секвенирование также

может быть осуществлено с использованием радиоактивно меченых фрагментов, секвенирование, основанное на флуоресценции, является более предпочтительным. В целом при секвенировании с окрашенным праймером флуоресцентно меченный праймер применяют в комбинации с немеченными дДНТФ. Процесс осуществляется с использованием четырех реакций синтеза и до получения на геле четырех полос для каждой подлежащей секвенированию матрицы (соответствующих продукту терминации, который специфичен для каждого основания). После удлинения праймера подлежащие секвенированию смеси, содержащие продукты терминации с включенными дидезоксинуклеотидами, обычно анализируют с использованием электрофореза на геле для секвенирования ДНК. После электрофоретического разделения флуоресцентно меченные продукты вырезают с помощью лазера из основания геля и флуоресценцию определяют с использованием соответствующего монитора. В автоматических системах при прохождении реакционных смесей через матрикс геля во время электрофореза детектор сканирует основание геля для обнаружения, какой из меченых фрагментов используется (Smith и др., 1986, Nature 321: 674-679). В модификации этого метода каждый из четырех праймеров метят различным флуоресцентным маркером. По завершении четырех различных реакций секвенирования реакционные смеси объединяют и объединенные реакционные смеси подвергают гель-анализу для каждой полосы, посредством чего индивидуально выявляют различные флуоресцентные метки (соответствующие четырем различным продуктам терминации, которые специфичны для каждого основания).

В другом варианте применяют метод секвенирования с использованием окрашенного терминатора. В этом методе для включения дДНТФ применяют ДНК-полимеразу и флуоресцентно меченные дДНТФ на наращиваемом конце ДНК-праймера (Lee и др., 1992, Nucleic Acid Research 20: 2471). Преимуществом этого способа является отсутствие необходимости синтезировать меченные красителем праймеры. Кроме того, реакции с использованием окрашенного терминатора являются более удобными, поскольку при этом все четыре реакции могут быть осуществлены в одной и той же пробирке. Ранее были описаны модифицированные термостабильные ДНК-полимеразы, обладающие пониженной способностью различать дДНТФ (см. публикацию европейской заявки ЕР-A-655506 и заявку на патент США 08/448223). Примером модифицированной термостабильной ДНК-полимеразы является мутантная форма ДНК-полимеразы из *T. aquaticus*, имеющая остаток тирозина в положении 667 (вместо остатка фенилаланина), т.е. так называемая F667Y-мутантная форма ДНК-полимеразы Taq. AmpliTaq® FS, которая производится фирмой Hoffmann-La Roche и продается через фирму Perkin Elmer, снижает количество дДНТФ, необходимое для эффективного

RU ? 2 3 5 7 7 3 C 2

секвенирования нуклеиновой кислоты-мишени, в несколько сотен или несколько тысяч раз. AmpliTaq[®] FS представляет собой мутантную форму ДНК-полимеразы из *T. aquaticus*, имеющую мутацию F667Y и дополнительный остаток аспарагиновой кислоты в положении 46 (вместо остатка глицина; G46D- мутация).

Таким образом, существует необходимость в термостабильных ДНК-полимеразах, применяемых для альтернативных методов синтеза нуклеиновых кислот, пригодных для точного и эффективного с точки зрения стоимости анализа нуклеотидной последовательности в ДНК. Также существует необходимость в разработке методов, основанных на флуоресценции, для которых не требуется применение дидезоксинуклеотидов. Настоящее изобретение направлено на решение этих задач.

Краткое изложение сущности изобретения
Настоящее изобретение относится к зависимым от матрицы термостабильным ДНК-полимеразам, которые включают характерную аминокислотную последовательность-мотив SerGlnIleXaaLeuArgXaa (SEQ ID NO: 1), причем "Хаа" в положении 4 этой последовательности представляет собой остаток любой аминокислоты, кроме остатка глутаминовой кислоты (Glu), а "Хаа" в положении 7 этой последовательности представляет собой остаток валина (Val) или остаток изолейцина (Ile). При выражении в однобуквенном коде аминокислот эта последовательность-мотив может быть обозначена как S Q I X L R V/I, где "Х" в положении 4 этой последовательности представляет собой остаток любой аминокислоты, кроме остатка глутаминовой кислоты. Способность термостабильных ДНК-полимераз, имеющих аминокислотную последовательность, включающую указанную последовательность-мотив, где "Х" в положении 4 этой последовательности-мотива не является остатком глутаминовой кислоты, ограничивать включение рибонуклеотидов по сравнению с ранее описанными термостабильными полимеразами понижена. Для растущей цепи ДНК рибонуклеотиды являются несвойственными нуклеотидами. Таким образом, первым предметом изобретения являются новые ферменты, способные включать несвойственные аналоги оснований, такие как рибонуклеотиды, в растущую цепь ДНК и на несколько порядков более эффективные, чем выявленные ранее термостабильные ДНК-сintéзирующие ферменты. Гены, кодирующие эти ферменты, также относятся к настоящему изобретению, равно как и рекомбинантные векторы экспрессии и клетки-хозяева, включающие эти векторы. С помощью таких трансформированных клеток-хозяев могут быть получены большие количества очищенных термостабильных ферментов полимераз.

В соответствии с настоящим изобретением была выявлена область или последовательность-мотив внутри аминокислотной последовательности термостабильных ДНК-полимераз, увеличивающая способность полимераз

включать рибонуклеотиды, сохраняя при этом способность правильно включать дезоксирибонуклеотиды. Изменения в этой области, например замена одной или нескольких аминокислот (например, введение их путем сайтнаправленного мутагенеза), приводят к получению термостабильного фермента полимеразы, который способен синтезировать РНК или РНК/ДНК-химеры либо гибридную цепь на матрице ДНК.

Другим предметом изобретения являются усовершенствованные методы и композиции, предназначенные для определения последовательности нуклеиновой кислоты-мишени, при этом необходимость в терминирующих цепь дДНТФ отпадает. С помощью представленных в настоящем описании усовершенствованных методов рибонуклеотиды (рНТФ) включаются в продукты удлинения праймеров. Поскольку ферменты, являющиеся предметом изобретения, точно и эффективно включают как рНТФ, так и дНТФ, в реакциях секвенирования могут использоваться смеси обоих нуклеотидов. После удлинения праймера вновь синтезированные олигонуклеотидные продукты могут быть расщеплены по местам включения рНТФ методами, известными в данной области техники, например путем гидролиза, что приводит тем самым к получению популяции фрагментов, пригодных для фракционирования и анализа последовательности стандартными методами, такими как гель-электрофорез. Эти методы, основанные на применении новых термостабильных ферментов полимераз, представлены в данном описании. Таким образом, этот предмет изобретения относится к термостабильным ДНК-полимеразам, которые отличаются тем, что полимераза включает критический мотив SerGlnIleXaaLeuArgXaa (SEQ ID NO: 1), где "Хаа" в положении 4 может представлять собой остаток любой аминокислоты, кроме остатка глутаминовой кислоты (Glu), а "Хаа" в положении 7 представляет собой остаток валина (Val) или остаток изолейцина (Ile).

Другим предметом изобретения являются модифицированные полимеразы, представленные в настоящем описании, которые включают рибонуклеотиды или аналоги, содержащие гидроксильную группу или другой заместитель в положении 2', которые обычно отсутствуют у дезоксирибонуклеотидов. Эти нуклеотиды могут быть по-разному помечены, что создает альтернативные варианты обычному применению дидезоксинуклеотидов для целей секвенирования ДНК.

Мутантные термостабильные полимеразы по изобретению способны более эффективно включать несвойственные нуклеотиды, в частности рибонуклеотиды, по сравнению с соответствующими ферментами дикого типа. В предпочтительном варианте осуществления изобретения подлежащий включению несвойственный нуклеотид может быть аналогом терминирующего цепь основания, таким как 2'-гидрокси-3'-дезоксиАТФ (кордицепнитрифосфат), являющийся "риботерминатор"-аналогом АТФ, или нетерминирующим цепь нуклеотидом, таким как рНТФ.

Другим предметом изобретения являются

RU ? 2 3 5 7 7 3 C 2

мутантные термостабильные полимеразы, которые способны более эффективно включать несвойственные нуклеотиды, в частности рибонуклеотиды, чем соответствующие ферменты дикого типа. Таким образом, этот предмет изобретения относится к рекомбинантным термостабильным ДНК-полимеразам, каждая из которых характеризуется тем, что (а) в своей нативной форме полимераза включает аминокислотную последовательность SerGlnIleGluLeuArgXaa (SEQ ID NO: 2), где "Xaa" в положении 7 этой последовательности обозначает остаток валина (Val) или остаток изолейцина (Ile); (б) в аминокислотной последовательности рекомбинантного фермента, предпочтительно в положении 4 этой последовательности, получают мутацию, при которой остаток глутаминовой кислоты в положении 4 представляет собой остаток другой аминокислоты, предпочтительно остаток глицина; и (в) способность рекомбинантного фермента ограничивать включение рибонуклеотидов и аналогов рибонуклеотидов по сравнению с нативной формой этого фермента понижена.

Другим предметом изобретения являются полимеразы, которые предназначены для фрагментации продуктов амплификации и продуктов удлинения праймера, причем такие фрагментированные продукты могут быть пригодны для использования в методах, основанных на гибридизации, и для различных стратегий определения последовательности.

Ферменты по настоящему изобретению и кодирующие их гены могут применяться для создания композиций, пригодных для осуществления реакций секвенирования ДНК и включающих смесь обычных нуклеотидов и по крайней мере один рибонуклеотид или аналог рибонуклеотида. В предпочтительном варианте осуществления изобретения несвойственный нуклеотид представляет собой рибонуклеотид, а концентрация рибонуклеотида ниже концентрации соответствующего дезоксирибонуклеотида, т.е. отношение рНТФ:дНТФ составляет 1:1 или менее. Ферменты по изобретению также пригодны для продажи в виде наборов, которые также могут включать любой из дополнительных элементов, необходимых для осуществления реакции секвенирования нукleinовой кислоты, такой как, например, дНТФ, рНТФ, буферы и/или праймеры.

Подробное описание изобретения

Настоящее изобретение относится к новым и улучшенным модифицированным термостабильным ДНК-полимеразам, композициям и наборам, охарактеризованным в формуле изобретения. Ферменты по изобретению более эффективно включают несвойственные нуклеозидтрифосфаты в сравнении с ранее известными полимеразами или соответствующими полимеразами дикого типа, из которых получают эти новые полимеразы. Согласно изобретению также предлагаются последовательности ДНК, кодирующие эти модифицированные ферменты, векторы для экспрессии модифицированных ферментов и клетки с такими внедренными в них векторами. Ферменты по изобретению пригодны для практического применения в новых методах секвенирования ДНК, которые обладают

рядом преимуществ по сравнению с методами секвенирования ДНК, известными из уровня техники.

Ниже для облегчения понимания сущности настоящего изобретения определены некоторые понятия.

Понятие "обычный", если оно упоминается в связи с основаниями нукleinовой кислоты, нуклеозидтрифосфатами или нуклеотидами, относится к таковым, которые встречаются в естественных условиях в указанном полинуклеотиде (например, для ДНК это дАТФ, дГТФ, дЦТФ и дТТФ). Кроме того, в реакциях синтеза ДНК *in vitro*, таких как секвенирование, вместо дЦТФ часто используют с7дГТФ и дИТФ (хотя они включаются с меньшей эффективностью). В целом они могут быть обобщены понятием дезоксирибонуклеозидтрифосфаты (дНТФ).

Понятие "система экспрессии" относится к последовательностям ДНК, содержащим требуемую кодирующую последовательность и функционально связанные контролирующие последовательности, так что хозяева, трансформированные этими последовательностями, обладают способностью продуцировать кодируемые протеины. Для осуществления трансформации система экспрессии может быть включена в вектор, хотя пригодная ДНК также может быть включена в хромосому хозяина.

Понятие "ген" относится к последовательности ДНК, которая включает контролирующие и кодирующие последовательности, необходимые для воспроизводимого получения биологически активного полипептида или предшественника. Полипептид может кодироваться полноразмерной последовательностью гена или любой имеющей достаточную длину частью кодирующей последовательности, чтобы сохранить ферментативную активность.

Понятие "клетка(и)-хозяин(ева)" относится как к одноклеточным прокариотическим или эукариотическим организмам, таким как бактерии, дрожжи и актиномицеты, так и к отдельным клеткам из растений или животных высших отрядов, если они способны расти в культуре клеток.

Используемое в настоящем описании понятие "реакционная смесь для секвенирования ДНК" относится к реакционной смеси, которая включает элементы, необходимые для реакции секвенирования ДНК. Таким образом, реакционная смесь для секвенирования ДНК пригодна для использования в методе секвенирования ДНК для определения последовательности нукleinовой кислоты-мишени, хотя вначале реакционная смесь может быть неполной, чтобы пользователь мог контролировать инициацию реакции секвенирования. В этом случае реакция может быть начата после того, как будет добавлен последний элемент, такой как фермент, который вводится для получения полной реакционной смеси для секвенирования ДНК. Обычно реакционная смесь для секвенирования ДНК должна содержать буфер, пригодный для полимеризации, нуклеозидтрифосфаты и по крайней мере один несвойственный нуклеотид. Реакционная смесь также может содержать праймер, который может

удлиняться на мишени полимеразой, полимеразу и нуклеиновую кислоту-мишень. Праймер, равно как и один из нуклеотидов, обычно включает содержащий метку, например, флуоресцентную фрагмент, который может быть обнаружен. Обычно реакционная смесь представляет собой смесь, которая включает четыре обычных нуклеотида и по крайней мере один несвойственный нуклеотид.

В предпочтительном варианте осуществления полимеразы представляет собой термостабильную ДНК-полимеразу, а несвойственный нуклеотид представляет собой рибонуклеотид.

Понятие "олигонуклеотид", используемое в данном описании, относится к молекуле, состоящей из двух или нескольких дезоксирибонуклеотидов или рибонуклеотидов, предпочтительно более чем из трех и обычно более чем из десяти. Точный размер олигонуклеотида может зависеть от многих факторов, включая основную функцию или применение олигонуклеотида.

Олигонуклеотиды могут быть получены любым пригодным методом, включая, например, клонирование и рестрикцию соответствующих последовательностей, и методом прямого химического синтеза, таким как фосфотриэфирный метод Narang и др., 1979, Meth. Enzymol. 68: 90-99; фосфодиэфирный метод Brown и др., 1979, Meth. Enzymol. 68: 109-151; диэтилфосфорамидный метод Beauchage и др., 1981, Tetrahedron Lett. 22: 1859-1862; триэфирный метод Matteucci и др., 1981, J. Am. Chem. Soc. 103: 3185-3191; автоматическими методами синтеза или методом с использованием твердой подложки, описанным в патенте США 4458066.

Понятие "праймер", используемое в данном описании, относится к естественному или синтетическому олигонуклеотиду, который способен служить точкой инициации синтеза при создании условий, при которых начинается удлинение праймера. Праймер предпочтительно представляет собой одноцепочечный олигодезоксирибонуклеотид. Соответствующая длина праймера зависит от предназначения праймера, но обычно составляет от 15 до 35 нуклеотидов. Для коротких молекул праймера обычно необходимы более низкие температуры для образования достаточно стабильных гибридных комплексов с матрицей. Праймер не должен отражать точную последовательность матрицы, но для того, чтобы происходило удлинение праймера, должен обладать достаточной степенью комплементарности для гибридизации с матрицей.

При необходимости праймер может быть помечен путем включения метки, которая может быть обнаружена спектроскопическими, фотохимическими, биохимическими, иммунохимическими или химическими способами. Например, пригодные метки включают ^{32}P , флуоресцентные красители, электронноплотные реагенты, ферменты (такие, как обычно применяемые в методах ELISA), биотин или гаптены и протеины, для которых имеются антисыворотки или

моноклональные антитела.

Понятие "термостабильная полимераза" относится к ферменту, который стабилен при нагревании и который устойчив к нагреванию и сохраняет достаточную активность для последующего осуществления реакции удлинения праймера при воздействии повышенных температур в течение времени, необходимого для осуществления денатурации двухцепочечных нуклеиновых кислот. Согласно данному описанию термостабильная полимераза пригодна для использования в реакции с температурными циклами, такой как полимеразная цепная реакция (ПЦР). Для термостабильной полимеразы под ферментативной активностью понимают способность соответствующим образом катализировать присоединение нуклеотидов для формирования продуктов удлинения праймера, которые комплементарны к цепи нуклеиновой кислоты-матрицы.

Необходимые для денатурации нуклеиновой кислоты условия нагревания могут зависеть, например, от концентрации соли в буфере, а также состава и длины нуклеиновых кислот, подлежащих денатурации, но обычно они находятся в интервале от приблизительно 90°C до приблизительно 105°C, предпочтительно от 90°C до 100°C, а продолжительность нагревания в основном зависит от температуры и длины нуклеиновой кислоты, и обычно она составляет от нескольких секунд до четырех минут.

Понятие "несвойственный" или "модифицированный", когда оно характеризует основание нуклеиновой кислоты, нуклеозидтрифосфат или нуклеотид, включает модификацию, производные или аналоги обычных оснований или нуклеотидов, которые встречаются в ДНК в естественных условиях. В частности, согласно данному описанию несвойственные нуклеотиды модифицированы в положении 2' сахара рибозы по сравнению с обычными дНТФ. Таким образом, хотя рибонуклеотиды (т.е. АТФ, ГТФ, ЦТФ, УТФ, называемые в целом рНТФ) и являются встречающимися в естественных условиях нуклеотидами РНК, в контексте данного описания эти рибонуклеотиды являются несвойственными нуклеотидами, поскольку имеют в положении 2' сахара гидроксильную группу, которая отсутствует в дНТФ. Аналоги рибонуклеотидов, содержащие заместители в положении 2', такие как 2'-фтор-, 2'-аминозамещенные аналоги, подпадают под объем изобретения. Кроме того, аналоги рибонуклеотидов могут быть модифицированы в положении 3', например, путем замещения нормальной гидроксильной группы на водородную группу (3'-дезокси), что приводит к получению аналога-терминатора рибонуклеотида. Такие нуклеотиды также включены в объем понятия "несвойственные нуклеотиды".

Поскольку ДНК обычно состоит из дНТФ, включение рНТФ должно рассматриваться как несвойственное, и, следовательно, рНТФ должен рассматриваться как несвойственное основание. Таким образом, в предпочтительном варианте изобретения при осуществлении методов удлинения ДНК-праймера, включая методы

RU
2
2
3
5
7
7
3
C
2

RU

? 2 3 5 7 7 3 C 2

секвенирования ДНК, в качестве продуктов получают нуклеиновые кислоты, которые содержат как обычные, так и необычные нуклеотиды, но в основном они содержат обычные нуклеотиды, которые представляют собой дНТФ.

Необычные основания могут быть флуоресцентно меченными, например, флуоресцином или родамином; нефлуоресцентно меченными, например, биотином; меченными изотопами, например, ^{32}P , ^{33}P или ^{35}S или немеченными.

Для пояснения сущности изобретения в описании приведены примеры специфических термостабильных ферментов ДНК-полимераз по изобретению, однако эти ссылки не должны рассматриваться как ограничивающие объем изобретения. В предпочтительном варианте термостабильные ферменты по изобретению применяли в различных методах секвенирования нуклеиновых кислот, хотя новые термостабильные полимеразы, приведенные в настоящем описании, могут применяться для любой цели, где такая ферментативная активность является необходимой или желательной. Фермент также может применяться в реакциях амплификации, таких как ПЦР.

Термостабильные полимеразы по изобретению отличаются тем, что каждая из них содержит критический мотив SerGlnIleXaaLeuArgXaa (SEQ ID NO: 1), где "Xaa" в положении 4 этой последовательности представляет собой остаток любой аминокислоты, кроме остатка глутаминовой кислоты (Glu), а "Xaa" в положении 7 этой последовательности представляет собой остаток валина (Val) или остаток изолейцина (Ile). Гены, кодирующие термостабильные полимеразы, которые имеют остаток глутаминовой кислоты в положении 4 этого мотива, могут быть модифицированы согласно настоящему описанию с получением пригодных модифицированных полимераз. Эти модифицированные термостабильные полимеразы отличаются тем, что по сравнению с соответствующими нативными ферментами или ферментами дикого типа они имеют модификацию в аминокислотной последовательности-мотиве SerGlnIleGluLeuArgXaa (SEQ ID NO: 2), где "Xaa" в положении 7 этой последовательности обозначает остаток валина (Val) или остаток изолейцина (Ile); т.е. этот мотив был модифицирован путем замещения остатка глутаминовой кислоты в положении 4 остатком другой аминокислоты. Критический мотив термостабильной ДНК-полимеразы по настоящему изобретению приведен ниже с использованием стандартного однобуквенного кода аминокислот (Lehninger, Biochemistry, New York, Worth Publishers Inc., 1970, стр.67).

SEQ ID NO: 1 SerGlnIleXaaLeuArgXaa, где "Xaa" в положении 4 представляет собой остаток любой аминокислоты, кроме остатка глутаминовой кислоты (Glu), а "Xaa" в положении 7 представляет собой остаток валина (Val) или остаток изолейцина (Ile).

Кодирующие последовательности генов, равно как и протеины, содержащие эту критическую аминокислотную последовательность, где Xaa в положении 4 не является остатком глутаминовой кислоты

(Glu), относятся к полимеразе, способность которой ограничивать включение рНТФ понижена, и включены в объем настоящего изобретения. Внутри критического мотива могут быть осуществлены дополнительные модификации остатков других аминокислот, предпочтительно остатков аминокислот, выбранных из группы, включающей глутамин (Gln или Q), лейцин (Leu или L) или аргинин (Arg или R).

Настоящее изобретение может использоваться для получения термостабильных ДНК-полимераз с улучшенными свойствами путем определенной модификации в последовательности гена, кодирующего термостабильную ДНК-полимеразу. В предпочтительном варианте изобретения последовательность гена и кодируемый фермент происходят из видов рода *Thermus*, хотя эубактерии, не относящиеся к роду *Thermus*, также включены в настоящее изобретение, как это более подробно описано ниже. Аналогично этому, ввиду высококонсервативной природы выявленного в настоящем исследовании критического мотива, новые термостабильные ДНК-полимеразы могут быть обнаружены на основе их гомологии, например, с Таq-полимеразой. Такие термостабильные полимеразы включены в объем настоящего изобретения при условии, что их аминокислотная последовательность включает мотив S Q I X L R V/I, где X обозначает остаток любой аминокислоты, кроме остатка глутаминовой кислоты, и эта аминокислотная последовательность в целом гомологична (идентичность последовательности) по крайней мере приблизительно на 39%, предпочтительно по крайней мере приблизительно на 60%, более предпочтительно по крайней мере приблизительно на 80% аминокислотной последовательности нативной Таq-полимеразы. Вся длина последовательности этой Таq-полимеразы приведена в WO 89/06691 и зарегистрирована под регистрационным номером P90556 в базе данных запатентованных последовательностей GENESEQ или под регистрационным номером M26480 в базе данных последовательностей EMBL и под регистрационным номером A33530 в базе данных последовательностей PIR.

Примером термостабильных ДНК-полимераз по настоящему изобретению являются рекомбинантные производные нативных полимераз из организмов, приведенных в таблице 1. В таблице 1 приведена конкретная последовательность критического мотива и положение остатка "X" для каждой из этих нативных полимераз. Поскольку каждая термостабильная ДНК-полимераза является уникальной, положение аминокислот критического мотива является определенным для каждого фермента. Для перечисленных ниже полимераз аминокислотный остаток в положении "X" критического мотива S Q I X L R V/I представляет собой глутаминовую кислоту. Молекулярная масса предпочтительных полимераз по настоящему изобретению составляет от 85000 до 105000 Да, более предпочтительно от 90000 до 95000 Да. Аминокислотная последовательность этих

R U ? 2 3 5 7 3 C 2

полимераз состоит из приблизительно 750-950 остатков аминокислот, предпочтительно из 800-850 остатков аминокислот. Полимеразы по настоящему изобретению могут состоять из приблизительно 540 или более аминокислот и включают, по крайней мере, полимеразный домен и часть, соответствующую 3'-5'-эксонуклеазному домену (полученная полимераза может обладать или не обладать 3'-5'-эксонуклеазной активностью), и возможно части 5'-3'-эксонуклеазного домена, который расположен на первой трети аминокислотной последовательности многих полноразмерных термостабильных полимераз.

Для термостабильных ДНК-полимераз, не приведенных в таблице 1, выбор соответствующей глутаминовой кислоты, подлежащей модификации, является простым, если определен критический мотив или консенсусный мотив в аминокислотной последовательности.

Независимо от точного положения внутри термостабильной ДНК-полимеразы, замещение остатка глутаминовой кислоты (Glu) на остаток другой аминокислоты внутри последовательности-мотива SerGlnIleGluLeuArgXaa (SEQ ID NO: 2), где "Хaa" в положении 7 этой последовательности обозначает остаток валина (Val) или остаток изолейцина (Ile), в полимеразном домене позволяет получить термостабильные полимеразы, способные к эффективному включению не свойственных нуклеотидов. В предпочтительном варианте глутаминовую кислоту заменяют на аминокислоту, имеющую ненагруженную полярную R-группу, такую как глицин, серин, цистein, треонин, или на аминокислоту, имеющую небольшую неполярную R-группу, такую как, например, аланин. В наиболее предпочтительном варианте остаток глутаминовой кислоты заменяют на остаток глицина (G). Программы, позволяющие установить аминокислотную и нуклеотидную последовательность, могут быть приобретены у фирмы Genetics Computer Group, 575 Science Drive, Медисон, шт. Висконсин. Программы, пригодные для конкретного, определенного в данном описании мотива, включают, например, "GAP", "BESTFIT" и "PILEUP", способствующие выявлению точной последовательности подлежащей модификации области.

Как видно из приведенной таблицы 1, существуют две необходимые формы консервативной последовательности-мотива SerGlnIleGluLeuArgXaa (SEQ ID NO: 2) внутри полимеразного домена термостабильных ДНК-полимераз этих термофильных организмов. Последовательность-мотив SerGlnIleGluLeuArgVal (SEQ ID NO: 3) присутствует в нативных термостабильных полимеразах таких видов рода *Thermus*, как, например, *Thermus aquaticus*, *Thermus caldophilus*, *Thermus thermophilus*, *Thermus flavus* и *Thermus filiformis*, а также видов *Thermus sps17* и *Z05*. Последовательность-мотив SerGlnIleGluLeuArgVal (SEQ ID NO: 3) также присутствует в полимеразном домене других термостабильных ДНК-полимераз, получаемых, например, из *Thermosiphon africanus* и из различных штаммов *Bacillus*, таких как *Bacillus caldotenax* и *Bacillus*

stearothermophilus.

Последовательность-мотив SerGlnIleGluLeuArglle (SEQ ID NO: 4), например, присутствует в нативных термостабильных полимеразах *Thermotoga maritima*, *Thermotoga neapolitana* и *Anaerocellum thermophilum*.

Таблица 1

	Организм	Аминокислотная последовательность консенсусного мотива	Положение глутаминовой кислоты
10		S Q I X L R V/I	
	<i>Thermus aquaticus</i> (Taq)	S Q I E L R V	615
	<i>Thermus caldophilus</i> (Tca)	S Q I E L R V	617
	<i>Thermus thermophilus</i> (Tth)	S Q I E L R V	617
	<i>Thermus flavus</i> (Tfl)	S Q I E L R V	616
	<i>Thermus filiformis</i> (Tfi)	S Q I E L R V	613
15	Вид <i>Thermus sps17</i>	S Q I E L R V	613
	Вид <i>Thermus Z05</i>	S Q I E L R V	617
	<i>Thermotoga maritima</i> (Tma)	S Q I E L R I	678
	<i>Thermotoga neapolitana</i> (Tne)	S Q I E L R I	678
	<i>Thermosiphon africanus</i> (Taf)	S Q I E L R V	677
	<i>Anaerocellum thermophilum</i> (Ath)	S Q I E L R I	632
	<i>Bacillus caldotenax</i> (Bca)	S Q I E L R V	659
20	<i>Bacillus stearothermophilus</i> (Bst)	S Q I E L R V	658, 661 или 736*

* Зависит от выбранной аминокислотной последовательности (см. ниже)

Полная последовательность нуклеиновой кислоты и аминокислотная последовательность для каждой из Taq-, Tth-, Z05-, sps17-, Tma- и Taf-полимераз описана в патенте США 5466591. Последовательность ДНК-полимеразы из Tca, Tfl, Tne, Ath, Bca и Bst описана в следующих публикациях: Tca - в базе данных последовательностей EMBL под регистрационным номером U62584 (см. также у Kwon, 1997, Mol. Cells 7(2): 264-271); Tfl - у Akhmetzjanov и Vakhitov, 1992, Nucleic Acids Research 20(21): 5839; Tne - в WO 97/09451 и в WO 96/41014; Ath - в базе данных последовательностей EMBL под регистрационным номером X98575 (более подробно штамм Ath описан у Rainey и др., 1993, J. Bacteriol. 175(15): 4772-4779); Bst - у Uemori и др., 1993, J. Biochem. 113: 401-410 и в базе данных последовательностей EMBL под регистрационным номером U23149 (см. также у Phang и др., 1995, Gene 163: 65-68). Аминокислотные последовательности Bst-полимеразы, содержащие "E" в критическом мотиве в положении 658, также описаны в опубликованном японском патенте 05/304964 A, опубликованной европейской заявке EP-A-699760 и у Aliotta и др., 1996, Genet. Anal. 12: 185-195; последовательность также может быть получена из базы данных последовательностей EMBL под регистрационным номером U33536. Последовательность, как описано в Gene 163: 65-68 (1995), содержит остаток "E" в положении 661 критического мотива. Bca-полимераза описана у Uemori и др., 1993, J. Biochem. 113: 401-410 и в базе данных последовательностей EMBL под регистрационным номером D12982. Термостабильная ДНК-полимераза из *Thermus filiformis* (см. FEMS Microbiol. Lett. 22: 149-153, 1994; также депонированная в ATCC под номером 43280) может быть восстановлена с использованием методов, описанных в патенте США 4889818, а также на основе информации о последовательности, представленной в таблице 1. Каждая из вышеуказанных последовательностей и публикаций включены в настоящее описание в качестве ссылки.

RU 2 2 3 5 7 7 3 C 2

RU

Гомология (идентичность последовательности) между аминокислотной последовательностью нативной формы Taq-полимеразы, как указано в WO 89/06691, и последовательностью указанной ранее Tfl-полимеразы составляет более 87,4%. Соответствующая гомология с Tth-полимеразой составляет 87,4%, с Tca-полимеразой составляет 86,6%, с Bst-полимеразой (регистрационный номер U23149) составляет 42,0%, с Bca-полимеразой составляет 42,6% и с Ath-полимеразой составляет 39,7%.

Как видно из таблицы 1, критический мотив является заметно консервативным у термостабильных ДНК-полимераз. В случае, когда "X" обозначает остаток глутаминовой кислоты, изменение гена, кодирующего полимеразу, позволяет получить фермент по изобретению, который легко включает рНТФ по сравнению, например, с Taq-полимеразой, у которой критический мотив не модифицирован. Следовательно, изобретение относится к классу ферментов, который также включает, например, термостабильную ДНК-полимеразу, и к соответствующему гену и векторам экспрессии из *Thermus oshimai* (Williams и др., 1996, Int. J. Syst. Bacteriol. 46(2): 403-408); *Thermus silvanus* и *Thermus chliaophilus* (Tenreiro и др., 1995, Int. J. Syst. Bacteriol. 45(4): 633-639); *Thermus scotoductus* (Tenreiro и др., 1995, Res. Microbiol. 146(4): 315-324); *Thermus brockianus* (Munster, 1986, Gen. Microbiol. 132: 1677) и *Thermus ruber* (Loginov и др., 1984, Int. J. Syst. Bacteriol. 34: 498-499; также депонированные в ATCC под номером 35948). Кроме того, изобретение включает, например, модифицированные формы термостабильных ДНК-полимераз и соответствующие гены и векторы экспрессии из *Thermotoga elfii* (Ravot и др., 1995, Int. J. Syst. Bacteriol. 45: 312; также депонированные в DSM под номером 9442) и *Thermotoga thermarum* (Windberger и др., 1992, Int. J. Syst. Bacteriol. 42: 327; также депонированные в DSM под номером 5069). Каждая из вышеуказанных последовательностей и публикаций включены в настоящее описание в качестве ссылки.

В предпочтительном варианте подлежащий модификации критический мотив находится внутри аминокислотной последовательности LeuAspTyrSerGlnIleGluLeuArgValLeuAlaHisLeu Ser (SEQ ID NO: 5). Таким образом, одним из предметов изобретения является получение мутантных термостабильных ДНК-полимераз, проявляющих существенно увеличенную эффективность по включению несвойственных нуклеотидов при использовании матрицы. В особенно предпочтительном примере последовательность полимеразы включает LeuAspTyrSerGlnIleGlyLeuArgValLeuAlaHisLeu Ser (SEQ ID NO: 6). Такие термостабильные ДНК-полимеразы особенно пригодны в таких процессах, как секвенирование ДНК, синтез РНК на основе ДНК и синтез *in vitro* ДНК, имеющих в качестве заместителей рНТФ.

Получение термостабильных ДНК-полимераз с повышенной эффективностью по включению несвойственных оснований может быть

осуществлено с помощью таких способов, как сайт-направленный мутагенез. См., например, публикацию Sambrook и др., Molecular Cloning: A Laboratory Manual, Cold Spring Harbor, 1989, второе издание, глава 15.51, "Oligonucleotide-Mediated Mutagenesis".

Например, мутация, заключающаяся в замене "A" на "G" во втором положении кодона, кодирующего остаток глутаминовой кислоты 615 в последовательности гена ДНК-полимеразы *Thermus aquaticus* (Taq) (см. SEQ ID NO: 7), приводит к более чем 500-кратному увеличению эффективности включения несвойственных нуклеотидов согласно данному описанию, сохраняя при этом способность фермента медиевировать ПЦР в присутствии обычных нуклеотидов, т.е. дНТФ.

Эта конкретная мутация в ДНК-полимеразе Таq приводит к замене аминокислоты Е (глутаминовая кислота) на G (глицин). Хотя эта конкретная замена аминокислоты существенно изменяет способность фермента включать несвойственные нуклеотиды, следует ожидать, что замена остатка глутаминовой кислоты на остатки любых других аминокислот, такие как, например, остатки серина, цистеина, треонина, аланина, валина или лейцина, будет оказывать такое же действие. Следовательно, настоящее изобретение включает и другие аминокислотные заместители, которые замещают E615, хотя мутант E615G является предпочтительным. Таким образом, в основе изобретения лежит тот факт, что четвертый остаток аминокислоты в мотиве, представленном в SEQ ID NO: 1, не является остатком глутаминовой кислоты.

Сайт-направленный мутагенез также может быть осуществлен путем сайт-специфического праймернаправленного мутагенеза. Этот способ в настоящее время является стандартным в данной области техники, и его осуществляют с использованием синтетического олигонуклеотидного праймера, комплементарного одноцепочечной ДНК фага, которая подлежит мутированию, за исключением ограниченного ошибочного спаривания, представляющего собой требуемую мутацию. В целом в качестве праймера для прямого синтеза цепи, комплементарной плазмиде или фагу, применяют синтетический олигонуклеотид, а образовавшуюся двухцепочечную ДНК трансформируют в бактерио-хозяйна, поддерживающую фаг. Образовавшаяся бактерия может быть проанализирована, например, методом секвенирования ДНК или путем гибридизации с зондом с целью выявить те бляшки, которые несут требуемую измененную последовательность гена. В альтернативном варианте могут применяться методы "рекомбинантной ПЦР", описанные в PCR Protocols, San Diego, Academic Press, под ред. Innis и др., 1990, глава 22, озаглавленная "Recombinant PCR", Higuchi, стр.177-183.

Как показано в таблице 1, глутаминовая кислота в критическом мотиве Таq-полимеразы сохраняется в других термостабильных ДНК-полимеразах, однако может быть локализована в другом, но близком положении аминокислотной

RU ? 2 3 5 7 3 C2

RU 2 2 3 5 7 7 3 C2

последовательности. Изменение консервативной глутаминовой кислоты в SEQ ID NO: 2 термостабильных ДНК-полимераз из представителей рода *Thermus* и ДНК-полимераз из родственных родов *Thermotoga*, *Thermosiphon* и *Anaerocellum* должно также усиливать способность полимеразы по эффективному включению несвойственных нуклеотидов по сравнению с Taq-полимеразой, содержащей SEQ ID NO: 2.

Изменение остатка глутаминовой кислоты в критическом мотиве других термостабильных ДНК-полимераз может быть осуществлено с использованием принципов и способов, применяемых в сайт-направленном мутагенезе. В GeneBank или в базе данных SwissProt/PIR существует несколько последовательностей ДНК-полимеразы *Bacillus stearothermophilus*. Эти последовательности имеют высокую степень родства, однако несколько отличаются друг от друга, но каждая из них содержит идентичную критическую последовательность-мотив SerGlnIleGluLeuArgXaa (SEQ ID NO: 2), где "Xaa" в положении 7 этой последовательности обозначает остаток валина (Val) или остаток изолейцина (Ile), хотя и в различных положениях последовательности.

Основываясь на опубликованной информации о доступных аминокислотных и нуклеотидных последовательностях термостабильных ДНК-полимераз, представленных в данном описании, с помощью стандартных методов рекомбинации также можно конструировать химерные полимеразы, которые состоят из доменов, происходящих из различных термостабильных ДНК-полимераз. В патентах США 5466591 и 5374553 описаны методы обмена различными функциональными сегментами термостабильных полимераз, такими как 5'-3'-эксонуклеазный домен, 3'-5'-эксонуклеазный домен и полимеразный домен, с получением новых ферментов. Предпочтительные химерные термостабильные ферменты полимеразы включают 5'-3'-эксонуклеазный домен, 3'-5'-эксонуклеазный домен и полимеразный домен, причем один из доменов происходит от другой полимеразы, и полимеразный домен содержит критический мотив SerGlnIleXaaLeuArgXaa (SEQ ID NO: 1), где "Xaa" в положении 4 этой последовательности представляет собой остаток любой аминокислоты, кроме остатка глутаминовой кислоты (Glu), а "Xaa" в положении 7 этой последовательности представляет собой остаток валина (Val) или изолейцина (Ile). Примерами таких химерных молекул являются химерные ферменты Taq/Tma, строение которых приведено в таблице 2. Как видно из этой таблицы, полимеразный домен этих химерных ферментов Taq/Tma содержит мутацию в указанном выше критическом мотиве.

Таблица 2

	5'-3'- эксонуклеазный домен	3'-5'- эксонуклеазный домен	Полимеразный домен
Taq	а.к. 1-289	а.к. 290-422	а.к. 423-832
Tma	а.к. 1-291	а.к. 292-484	а.к. 485-893
Taq/Tma	а.к. 1-289 (Taq)	а.к. 292-484 (Tma)	а.к. 423-832 (Taq) с мутацией E615G
Taq/Tma	а.к. 1-289 (Taq)	а.к. 292-484 (Tma)	а.к. 485-893 (Tma) с мутацией E678G

Плазмида pC1 в соответствии с

Будапештским договором была депонирована в ATCC 17 июля 1996 г. и получила регистрационный номер 98107. Плазмида pC1 содержит ген, кодирующий термостабильную ДНК-полимеразу с мутацией в кодоне, кодирующем остаток глутаминовой кислоты в положении 615 аминокислотной последовательности нативной

5 Таq-полимеразы, что приводит к получению мутантной формы Таq-полимеразы, которая имеет остаток глицина в положении 615 (E615G-мутантная Таq-полимераза). Этот

10 депозит позволяет применять альтернативные способы для получения термостабильных ДНК-полимераз, обладающих повышенной эффективностью по включению несвойственных аналогов нуклеотидов. В примере 1 показано, что применение flankирующих сайтов рестрикции, пригодных для субклонирования мутации E615G, позволяет создать другие

15 термостабильные ферменты ДНК-полимеразы. Поскольку для

20 многочисленных термостабильных ДНК-полимераз известна полная последовательность гена, основываясь на информации о последовательности

25 критического мотива, приведенной в описании, для введения мутации E615 можно использовать другие известные специалистам в данной области техники способы, такие как расщепление рестриктазами и замещение фрагмента или сайт-специфический мутагенез *in vitro*.

30 Модифицированный ген или фрагмент гена, полученный сайтом-специфическим мутагенезом, может быть получен из плазмида или фага стандартными способами и лигирован с вектором экспрессии для

35 последующего культивирования и выделения образовавшегося ферmenta. Для практического осуществления изобретения пригодны многочисленные клонирующие и экспрессирующие векторы, включая системы,

40 пригодные для млекопитающих и бактерий, и они описаны, например, у Sambrook и др., Molecular Cloning: A Laboratory Manual, Cold Spring Harbor, второе издание, 1989. Для удобства в настоящем изобретении приведен пример использования промотора PL, происходящего из фага лямбда (Shimatake и др., 1981, Nature 292: 128). Применение этого промотора, в частности, описано в патентах США 4711845 и 5079352.

45 Термостабильные ДНК-полимеразы по настоящему изобретению обычно выделяют из микроорганизмов, таких как *E.coli*, которые были трансформированы вектором экспрессии, функционально связанным с геном, кодирующим ДНК-полимеразу дикого типа или модифицированную термостабильную ДНК-полимеразу.

50 Примером пригодного микроорганизма-хозяина является штамм *E. coli* DG116, описанный у Lawyer и др., 1993, PCR Methods and Applications 2: 275-287, и этот штамм также может быть получен из

55 американской коллекции типовых культур (American Type Culture Collection) под регистрационным номером ATCC 53601.

60 Методы очистки термостабильной полимеразы также описаны, например, у Lawyer и др., 1993, PCR Methods and Applications 2: 275-287.

Для специалиста в данной области

R U ? 2 3 5 7 7 3 C 2

техники очевидно, что вышеуказанные термостабильные ДНК-полимеразы, обладающие увеличенной эффективностью по включению несвойственных нуклеотидов, наиболее легко могут быть получены с использованием методов, основанных на технологии рекомбинантной ДНК. Необходимость получения одного из ферментов по настоящему изобретению или производного либо гомолога этих ферментов предполагает создание рекомбинантной формы фермента, которое обычно включает конструирование вектора экспрессии, трансформацию клетки-хозяина этим вектором и культивирование трансформированной клетки-хозяина в условиях, при которых такая экспрессия будет происходить. Способы получения векторов экспрессии, трансформации и культивирования трансформированной клетки-хозяина хорошо известны в данной области техники и подробно описаны, например, у Sambrook и др., 1989, см. выше.

Настоящее изобретение относится к термостабильным ДНК-полимеразам, которые вместе с рибонуклеозидтрифосфатами пригодны для осуществления многочисленных функций, включая амплификацию нуклеиновой кислоты, методы определения и секвенирования ДНК. Применение рибонуклеотидов при секвенировании позволяет исключить использование цепь-терминирующих аналогов, имеющих высокую стоимость, таких как ддНТФ, и, что является важным фактором, облегчает получение новых продуктов амплификации, пригодных не только для анализа последовательности ДНК, но и для других типов анализа, таких как электрофорез или гибридизация, без необходимости последующего проведения реакций секвенирования ДНК.

Ранее уже было подтверждено, что пирофосфатааза улучшает результаты секвенирования при использовании как мезофильных полимераз, так и термостабильных ДНК-полимераз за счет снижения количества накапливающихся продуктов удлинения, полученных в результате пирофосфоролиза. При этом до начала цикла секвенирования необходимо использовать методы, позволяющие вводить дополнительный фермент в реакцию секвенирования. Однако наиболее существенное преимущество настоящего изобретения состоит в том, что для секвенирования ДНК не требуется пирофосфатааза. Таким образом, применение новых, представленных в данном описании ферментов устраниет необходимость в дополнительных расходах по введению второго фермента в смесь для реакции секвенирования.

При применении ферментов по настоящему изобретению реакции амплификации и секвенирования объединяют, что способствует экономии времени и материалов, а также упрощает весь анализ. Эти, а также другие преимущества достигаются, в первую очередь, благодаря тому, что включение как обычных нуклеотидов, так и рибонуклеотидов и их аналогов в продукт удлинения праймера позволяет получить химерную цепь РНК/ДНК, чувствительную к гидролизу РНК. Обработка

5 не оказывает влияния на скелет ДНК и приводит к получению популяции фрагментов нуклеиновой кислоты, каждый из которых заканчивается в том положении, где был встроен рибонуклеотид вместо соответствующего дНТФ. Гидролиз легко осуществляют различными способами, включающими, но не ограничиваясь ими, щелочной гидролиз (например, обработка NaOH в конечной концентрации 0,2 М, как описано ниже в примере VI), нагревание или ферментативная обработка РНКазой (под ред. Vogel и др., *Informational Macromolecular*, New York, Academic Press, 1963, глава, написанная Berg и др., озаглавленная "The Synthesis of Mixed Polynucleotides Containing Ribo- and Deoxyribonucleotide by Purified Preparation of DNA Polymerase from E.coli", стр.467-483).

10 В предпочтительном варианте настоящее изобретение относится к новым и улучшенным композициям, наиболее пригодным для методов секвенирования ДНК. Новые ферменты, приведенные в данном описании, могут использоваться для методов секвенирования нуклеиновых кислот, основанных на применении либо окрашенных терминаторов, либо окрашенных праймеров, а также для других методов секвенирования. Как описано ранее, методы терминации цепи обычно требуют использования матрицы для удлинения праймера в присутствии терминирующих цепь нуклеотидов, что приводит к созданию неполных фрагментов, которые затем разделяют по размеру. В стандартном дидезокси-секвенировании для терминации цепи используются дидезоксинуклеозидтрифосфаты и ДНК-полимераза, такая как фрагмент Кленова из штамма E.coli Pol I (Sanger и др., см. выше).

15 Таким образом, основная методика дидезокси-секвенирования включает (I) отжиг олигонуклеотидного праймера, соответствующего матрице; (II) удлинение праймера с помощью ДНК-полимеразы в четырех различных реакционных смесях, в каждой из которых содержится один меченный нуклеотид или меченный праймер, смесь немеченых дНТФ и один терминирующий цепь ддНТФ; (III) разделение четырех наборов продуктов реакций с помощью, например, гель-электрофореза на денатурирующем полиакриламидном геле с высокой разрешающей способностью/мочевине, капиллярного разделения или других способов разделения; и (IV) получение авторадиографического изображения геля, которое может быть проанализировано для получения заключения о последовательности. В альтернативном варианте для получения информации о последовательности ДНК могут применяться масс-спектрометрические методы или методы, основанные на гибридизации, с использованием флуоресцентно меченых праймеров или нуклеотидов.

20 25 30 35 40 45 50 55 60 Доступность термоустойчивых полимераз, таких как Таq-полимераза, привела к улучшению метода секвенирования (см. патент США 507216) и к его модификациям, называемым "циклическим секвенированием". При циклическом секвенировании повторяют циклы нагревания и охлаждения, что позволяет получить многочисленные продукты удлинения каждой

R U ? 2 3 5 7 3 C 2

R U

молекулы-мишени (см. Murray, 1989, Nucleic Acids Research, 17: 8889). Эта асимметричная амплификация последовательностей-мишеней, комплементарных последовательности матрицы, в присутствии дидезокси-терминаторов цепи приводят к получению семейства продуктов удлинения всех возможных длин.

После денатурации продукта реакции удлинения ДНК-матрицы множественные циклы отжига праймера и удлинения праймера осуществляют в присутствии дидезокси-терминаторов. Термостабильные ДНК-полимеразы имеют несколько преимуществ при циклическом секвенировании; они устойчивы к жестким температурам отжига, которые необходимы для специфической гибридизации праймера с нуклеиновыми кислотами-мишениями, а также устойчивы к множеству циклов высокотемпературной денатурации, которые осуществляются в каждом цикле, т.е. к температуре 90-95°C. По этой причине различные формы ДНК-полимеразы AmpliTaq® были включены в наборы для циклического секвенирования с использованием Taq, которые реализуются через фирму Perkin Elmer, Норвулк, шт. Коннектикут.

Однако способность ДНК-полимеразы Taq ограничивать включение не свойственных нуклеотидов, таких как ддНТФ, связана с проблемой ее использования для циклического секвенирования, поскольку ддНТФ или флуоресцентно меченные ддНТФ должны быть включены в качестве терминаторов цепи. В целом до создания настоящего изобретения секвенирование ДНК с помощью термостабильных ДНК-полимераз требовало наличия смеси терминирующих цепь нуклеотидов, обычно дидезоксинуклеотидов, в высокой концентрации с целью гарантировать получение популяции продуктов удлинения, которая включает фрагменты, имеющие все возможные длины, вплоть до длины в несколько сотен оснований. Часто для снижения стоимости этого процесса применяли протоколы с использованием очень низких концентраций обычных ДНТФ, что делало реакции неэффективными. Эти реакционные смеси, включающие низкую концентрацию ДНТФ и высокую концентрацию ддНТФ, создавали условия, при которых термостабильная полимераза испытывает существенный недостаток нуклеотидных субстратов.

Несмотря на появление модифицированных ферментов, таких как ДНК-полимераза AmpliTaq® FS, позволяющих увеличить концентрацию ДНТФ до более оптимальных уровней, для ферментов-прототипов все еще были необходимы дорогостоящие ддНТФ для секвенирования ДНК. В противоположность этому настоящее изобретение относится к ферментам, которые не только позволяют увеличивать концентрацию ДНТФ, но и исключают применение дорогостоящих ддНТФ, поскольку вместо них в растущую цепь включают РНТФ. Способность новых ферментов эффективно осуществлять частичное замещение рибонуклеотидами облегчает получение ступеней для секвенирования ДНК при отсутствии отдельной реакции по включению

терминирующего нуклеотида.

Выбор не свойственных нуклеотидных аналогов, пригодных для использования в методах секвенирования ДНК, ранее определялся способностью термостабильной ДНК-полимеразы включать эти аналоги. К сожалению, эти нуклеотидные аналоги являются довольно дорогостоящими. Например, цена ддНТФ приблизительно в 25 раз превышает цену как РНТФ, так и ДНТФ. Поскольку ранее термостабильные ДНК-полимеразы не обладали способностью эффективно включать РНТФ в растущую цепь ДНК при использовании матрицы, такие рибонуклеотиды, которые являются легко доступными и недорогими, не могли использоваться для секвенирования ДНК с участием термостабильной ДНК-полимеразы. Настоящее изобретение позволяет исключить необходимость в ддНТФ в реакциях секвенирования ДНК. Таким образом, одним из предметов изобретения является метод секвенирования ДНК, который является существенно менее дорогостоящим по сравнению с ранее применяемыми методами терминации цепи.

Присутствие марганца в реакции удлинения праймера оказывать влияние на способность полимеразы точно включать правильные пары нуклеотидных оснований. Марганец может применяться для усиления некорректного спаривания оснований или для снижения ограничения по инсерции нуклеотидного аналога. Ранее исследователи применяли марганец с той целью, чтобы индуцировать мутагенез при репликации или амплификации ДНК. Таким образом, марганец может влиять на правильность реакции полимеризации, а также на выход продукта реакции. Образовавшаяся последовательность может оказаться неправильной или же - в методе секвенирования ДНК - полученная информация может оказаться неверной. Способы согласно настоящему изобретению не требуют включения двухвалентного катиона марганца в смесь для реакции секвенирования с целью усиления способности полимеразы встраивать не свойственный нуклеотид. В противоположность известным из уровня техники ДНК-полимеразам в настоящем изобретении описан критический мотив, входящий в полимеразный домен, контролирующий способность фермента различать 2'-замещенные и незамещенные нуклеотиды, не используя для этого марганец.

Ферменты по настоящему изобретению не требуют для секвенирования высоких концентраций не свойственных аналогов оснований. До создания настоящего изобретения при осуществлении секвенирования, основанном на терминации цепи ДНК (см. также патент США 5075216 на имя Innis и др.), не свойственные аналоги оснований и соответствующие традиционные основания обычно присутствовали в соотношении (например, дДАТФ:ДАТФ) приблизительно от 1,3:1 до 24:1. Для сравнения термостабильные полимеразы, описанные в настоящем изобретении, позволяют снизить отношение не свойственных аналогов оснований к обычным основаниям в сто или в несколько тысяч раз. Отношение РНТФ:ДНТФ, равное

R U ? 2 3 5 7 3 C 2

R U

1:1 или ниже, в сочетании с новыми ферментами, представленными в данном описании, является достаточным для анализа последовательности ДНК. В предпочтительном варианте отношение рНТФ:дНТФ снижают до менее чем 1:8. Отношение 2'-замещенного нуклеотида к соответствующему естественному дНТФ может составлять даже 1:80 или 1:200, что зависит от конкретного плана эксперимента и требуемой длины фрагментов.

Таким образом, поскольку ферменты по настоящему изобретению легко включают не свойственные нуклеотиды, такие как 2'-замещенные нуклеотиды, нет необходимости усиливать их способность по включению рНТФ с помощью высокой концентрации рНТФ и ограничения концентрации соответствующего дНТФ. Следовательно, методы по настоящему изобретению позволяют применять оптимальные концентрации дНТФ в сочетании с небольшими количествами рНТФ.

Когда в соответствующем методе, таком как секвенирование с окрашенным праймером, применяют модифицированные ферменты полимеразы по настоящему изобретению, хорошие результаты по секвенированию ДНК получают при концентрации каждого дНТФ 50-500 мкМ. Предпочтительно концентрация дНТФ составляет от 100 до 300 мкМ. При использовании этих интервалов соответствующий рНТФ может присутствовать приблизительно в такой же концентрации, что и дНТФ, или в более низкой. Предпочтительно концентрация рНТФ составляет приблизительно от 0,1 мкМ до 100 мкМ, наиболее предпочтительно приблизительно от 2,5 мкМ до 25 мкМ.

Концентрация рНТФ, пригодная для применения модифицированных ферментов по настоящему изобретению, легко может быть определена обычным специалистом в данной области техники путем титрования и оптимизации экспериментов. Необходимое количество рНТФ или аналога определяется типом эксперимента, и на него может влиять размер и чистота мишени, а также выбор буфера и конкретные виды фермента.

Отношение рНТФ:дНТФ будет определять частоту, с которой рНТФ встраиваются в растущий олигонуклеотид. Поскольку в каждом встроенным рНТФ может происходить гидролиз, отношение рНТФ:дНТФ можно регулировать таким образом, чтобы обеспечить пользователю гибкость в возможности увеличения или уменьшения размера образовавшихся фрагментов.

Как хорошо известно, ДНК представляет собой полимер, синтезированный из дНТФ. Каждый дезоксинуклеозидтрифосфат включает сахар рибозу, который имеет гидроксильную группу в 3'-положении и водород в 2'- положении. Рибонуклеотиды также содержат гидроксильную группу в 3'-положении сахара. Однако рНТФ отличаются от дНТФ тем, что в 2'-положении сахара атом водорода замещен на вторую гидроксильную группу. В данном случае рНТФ являются примером, на котором выявляется способность ферментов по изобретению точно включать 2'-замещенные нуклеотиды. Однако соединения по изобретению не ограничены применением не свойственных

нуклеотидов, которые представляют собой рибонуклеотиды. Модификация последовательности термостабильной полимеразы в критическом домене, приведенном в данном описании, позволяет матрице направленно включать альтернативные 2'-замещенные нуклеотиды, такие как 2'-гидроксил, 3'-дезоксинуклеотиды и нуклеотиды, замещенные в 2'-положении фтором или аминогруппой.

Как описано в приведенных ниже примерах, включение 3'-дезокси, 2'-гидроксиАТФ, который в данном описании обозначен также как кордицепинтрифосфат, облегчается наличием второй мутации в термостабильной полимеразе, способность которой ограничивать включение нуклеотида, содержащего дезоксигруппу в 3'-положении рибозы, снижена. Такие ферменты ранее были описаны, например, в EP-A-655506 и в заявке на патент США 08/448223, поданной 23 мая 1995 г. Депозит ATCC 69820, депонированный в соответствии с Будапештским договором 10 мая 1995 г., обеспечивает получение гена, кодирующего модифицированную термостабильную ДНК-полимеразу *Thermus aquaticus*, способность которой ограничивать включение аналогов, таких как ддНТФ, снижена. Дидезоксинуклеотиды имеют заместитель в 3'-положении по сравнению с обычными дНТФ. Таким образом, в сочетании с настоящим изобретением двойная мутация, примером которой в данном описании является E615G, F667Y-мутант Тақ-полимеразы, позволяет использовать нуклеотидные аналоги, которые содержат заместитель в 3'- и 2'-положениях рибозы по сравнению с дНТФ (см. примеры III и V).

Конкретным применением изобретения является метод секвенирования с использованием рНТФ, где секвенирующий праймер является меченым с помощью выявляемой флуоресцентной или радиоактивной метки. В отличие от ддНТФ включение немодифицированного рНТФ не приводит к такому действию, как терминация цепи. Реакционная смесь для секвенирования ДНК включает как рНТФ, так и дНТФ в сочетании с ферментом по изобретению, что приводит к получению смеси беспорядочно замещенных продуктов удлинения праймера, чувствительных к расщеплению по 3'-5'-fosfodiэфирной связи между рибонуклеотидом и соседним дезоксирибонуклеотидом. После удлинения праймера, например, при ПЦР-амплификации или при циклическом секвенировании и до разделения продуктов удлинения праймера, например, с помощью гель-электрофореза реакционную смесь либо обрабатывают щелочью, нагревают, обрабатывают рибонуклеазой, либо другими способами гидролизуют продукты удлинения по каждому встречающемуся рибонуклеотиду. Для каждого меченого продукта удлинения праймера только наиболее крупный 5'-фрагмент, который является промежуточным продуктом удлинения меченого праймера, обнаруживается на секвенирующем геле. Для данной мишени анализ полученного секвенирующего геля дает ступени секвенирования, т.е. серии обнаруживаемых сигналов в Г-, А-, Т- и Ц-линиях, соответствующих нуклеотидной

R U ? 2 3 5 7 7 3 C 2

последовательности мишени. Полученные ступени секвенирования предоставляют такую же информацию, что и обычный метод, предполагающий использование ддНТФ или основанный на использовании рНТФ и новых термостабильных полимераз, представленных в данном описании. Таким образом, при применении настоящего изобретения дорогостоящие ддНТФ для секвенирования ДНК больше не требуются (см. пример VI).

В альтернативном методе секвенирования применяют терминирующие цепь рибонуклеотиды. В этом варианте в качестве терминаторов применяют 2'-гидрокси,3'-дезоксинуклеотиды, такие как кордицепинтрифосфат. Эти аналоги рНТФ могут быть флуоресцентно меченными, и их применяют для секвенирования ДНК. У Lee и др., см. выше, описано применение окрашенных терминаторов ддНТФ. В EP-A-655506 и в заявке на патент США 08/448223, поданной 23 мая 1995 г., описаны модифицированные ферменты для применения с ддНТФ. Термостабильная ДНК-полимераза, содержащая как модификацию, присутствующую в ДНК-полимеразе FS AmpliTaq® (см. выше), так и таковую, представленную в SEQ ID NO: 1, где X не обозначает глутаминовую кислоту (E), как описано ранее, может применяться для эффективного включения меченых аналогов рНТФ при осуществлении реакции секвенирования, которая основана на терминации цепи. Этот процесс может быть автоматизирован и не требует синтеза меченных красителем праймеров. Кроме того, поскольку реакции с окрашенным терминатором позволяют осуществлять в одной и той же пробирке все четыре реакции, они более удобны, чем методы с окрашенным праймером. 2'-гидрокси,3'-дезоксинуклеотиды могут быть синтезированы из коммерчески доступных 3'-нуклеотидов (3'-дА, 3'-дЦ, 3'-дГ и 3'-дТ, например, поставляемых фирмой Sigma Chemical Corporation, Ст. Луис, шт. Миссури) и при добавлении 5'-трифосфата, как описано у Ludwig, Biophosphates and Their Synthesis Structure, Metabolism and Activity, под ред. Bruzik и Stec, Амстердам, Elsevier Science Publishers, 1987, стр.201-204.

Помимо применения в новых методах секвенирования, модифицированные ферменты, представленные в данном описании, могут применяться для различных целей в молекулярной биологии. В одном из примеров осуществления изобретения модифицированные ферменты применяют в реакционных смесях для амплификации, содержащих как обычные, так и не свойственные нуклеотиды, например дНТФ и по крайней мере один обнаруживаемый в результате мечения рНТФ, метки, которые включают, например, флуоресцентные метки или радиоизотопы. Управляемый матрицей синтез комплементарной цепи приводит к получению ДНК-продукта, содержащего рибонуклеозидмонофосфат в различных положениях по его длине. Нагревание и/или обработка щелочью гидролизуют продукт удлинения нуклеиновой кислоты на каждом рибонуклеотиде. Таким образом, получают семейство сегментов ДНК, где каждый фрагмент содержит один меченный остаток на

своем 3'-конце. Размер образовавшихся фрагментов нуклеиновой кислоты может быть изменен путем регулирования отношения и количества рНТФ, включенного в реакцию.

Амплификация мишени с использованием рНТФ и ферментов по настоящему изобретению создает многочисленные преимущества, обусловленные конкретным применением. В описанном выше методе при использовании меченого рНТФ все полученное семейство фрагментов оказывается меченым с равной интенсивностью: одна метка на один олигонуклеотидный фрагмент. Для достижения оптимальных результатов при осуществлении таких методов, как обнаружение нуклеиновой кислоты с использованием олигонуклеотидного зонда, определенным образом фиксированного на тонком слое силикона, обычно необходимо, чтобы амплифицированная мишень была случайным образом фрагментирована в пределах фиксированного и воспроизводимого размера с целью ограничить образование вторичных структур для контроля кинетики гибридизации. Кроме того, для выявления гибридизации при упорядоченном расположении тысяч зондов на тонком слое силикона может оказаться предпочтительным, чтобы фрагменты нуклеиновой кислоты были мечены с равной интенсивностью. Согласно настоящему изобретению были разработаны способы получения семейств фрагментов, удовлетворяющих этому стандарту, облегчая тем самым применение альтернативных форматов выявления, таких, как методы, основанные на использовании тонких слоев, описанные, например, у Cronin и др., 1996, Human Mutation, 7: 244-255.

В другом примере предполагается применение одного меченого праймера и одного немеченого праймера в смеси для реакции амплификации, которая включает термостабильную полимеразу по изобретению, а также рНТФ и дНТФ, что позволяет одновременно улучшить реакции амплификации и секвенирования. Для этого метода требуется, чтобы осуществлялись четыре различные реакции амплификации, по одной для каждого рНТФ. Таким образом, например, поскольку фермент по изобретению пригоден для амплификации мишени, например, с помощью ПЦР или других способов амплификации, полученный продукт, если он присутствует, может быть обнаружен обычными способами, такими как гель-электрофорез или гибридизация с зондом при использовании части продукта реакции. Эти методы определения не приводят к гидролизу включенных рибонуклеотидов, а поведение РНК/ДНК-химерных цепей будет таким, как это ожидается для обычного продукта амплификации нуклеиновой кислоты. Если требуемый продукт обнаружен, оставшаяся часть этой же реакционной смеси может быть обработана щелочью и проанализирована с помощью гель-электрофореза с целью определить последовательность нуклеиновой кислоты. Таким образом, после обнаружения продукта последующая реакция секвенирования не является необходимой. Эта упрощенная методика обеспечивает экономию затрат времени и материалов и

R U ? 2 3 5 7 7 3 C 2

характеризуется повышенной точностью благодаря уменьшению стадий: обнаруженный продукт и является секвенированным продуктом.

Кроме того, также может применяться аналогичная методика с четырьмя мечеными рНТФ и одним биотинилированным (меченным биотином) праймером. После амплификации продукт разлагают с помощью щелочного гидролиза, а праймер, связанный с продуктами, удаляют путем взаимодействия с покрытыми стрептавидином гранулами. Захваченные гранулы продукты затем анализируют на секвенирующем геле. Эта модификация позволяет осуществлять реакцию секвенирования в одной пробирке, исключая тем самым необходимость в четырех различных амплификациях.

В соответствии с другим предметом изобретения ферменты, представленные в данном описании, могут применяться для получения РНК с использованием ДНК-матрицы или для получения замещенной ДНК для медиаируемой щелочью стерилизации без применения обычных стерилизующих агентов, таких как урацил-N-гликозилаза (UNG), как описано в международной заявке WO 92/01814.

В приведенных примерах осуществления изобретения термостабильная полимераза также содержит мутацию в 5'-3'-экзонуклеазном домене, которая приводит к значительному ухудшению экзонуклеазной активности.

Модифицированные формы Тац-полимеразы описаны в патенте США 5466591. В одном из вариантов осуществления изобретения кодон, кодирующий остаток глицина (G) в положении 46 аминокислотной последовательности, заменили кодоном, кодирующим аспарагиновую кислоту (D). Полученный таким образом фермент обладает преимуществом при использовании в реакциях циклического секвенирования благодаря пониженной 5'-3'-экзонуклеазной активности и является предпочтительным для применения по настоящему изобретению. По сравнению с ферментом дикого типа наличие мутации G46D не оказывает влияния на полимеразный домен аминокислотной последовательности и на полимеразную активность.

Согласно изобретению предлагаются также наборы для секвенирования нукleinовых кислот, содержащие термостабильную полимеразу по настоящему изобретению, что является примером коммерческого применения изобретения.

Такие наборы обычно включают дополнительные реагенты для секвенирования ДНК, такие как, например, рНТФ, дНТФ и соответствующие буферы. Если рНТФ являются немечеными, также может быть включен меченный праймер.

Ниже изобретение проиллюстрировано на примерах, не ограничивающих его объем.

Пример I

Экспрессия гена модифицированной Тац-полимеразы с пониженной способностью ограничивать включение не свойственных нуклеотидов

С-концевая часть аминокислотной последовательности ДНК-полимеразы Тац кодирует домен, включающий сайт полимеразной активности (Lawyer и др., 1993,

PCR Methods and Applications 2: 275-287). Фрагмент ДНК, содержащий эту область, выделяли из полноразмерного гена Тац и осуществляли мутагенез с помощью ПЦР-амплификации в присутствии марганца (Leung и др., Technique 1(1): 11-15). В этом примере все рестриктазы были получены от фирмы New England Biolabs, Беверли, шт. Массачусетс. Фрагменты, содержащие мутации, расщепляли рестриктазами PstI и BglII и клонировали в плазмиде, экспрессирующей Тац, в данном случае в плазмиде pLK102, которую предварительно расщепляли PstI и BglII. Плазмода pLK102 представляет собой модифицированную форму плазмиды pSYC1578, которая экспрессирует Тац (Lawyer и др., см. выше). Фрагмент HinclI/EcoRV, локализованный на 3'-области, кодирующй полимеразу, удаляли для создания плазмиды pLK101. PstI-BglII-фрагмент длиной 898 пар оснований затем удаляли из pLK101 и заменили коротким олигонуклеотидным дуплексом PstI-EcoRV-BglII для создания плазмиды pLK102. Таким образом, в результате этой делеции было удалено 900 пар оснований с 3'-конца пол-гена ДНК Тац с их замещением более коротким участком ДНК.

Полученными экспрессирующими плазмидами трансформировали штамм E.coli N1624 (описанный у Gottesman, 1973, J. Mol. Biol. 77: 531; он также может быть получен из E.coli Genetic Stock Center в Йельском Университете, номер штамма CGSC 5066) и полученных трансформантов подвергли скринингу на способность более эффективно по сравнению с ферментом дикого типа включать рНТФ. Используя эту методику, выявляли мутант С1, способный более эффективно включать рНТФ.

Для определения, какая часть гена Тац-полимеразы ответственна за измененный фенотип, содержащую мутацию плазмиду, экспрессирующую Тац, названную pC1 и выделенную из мутанта С1, расщепляли различными рестриктазами и образовавшиеся рестрикционные фрагменты субклонировали в плазмиде pLK101, несущей ген ДНК-полимеразы Тац дикого типа, замещая рестрикционные фрагменты, не содержащие мутаций. Анализ образовавшихся субклонов показал, что мутация, отвечающая за фенотип, находилась в рестрикционном фрагменте NheI-BamHI длиной 265 пар оснований.

Осуществляли анализ последовательности ДНК этой области плазмиды pC1, используя набор ABI PRISM[®] Dye Terminator Cycle Sequencing Core Kit, в который входит ДНК-полимераза FS AmpliTaq[®], поставляемая фирмой Applied Biosystems, Фостер Сити, Калифорния, и система для секвенирования Applied Biosystems, Model 373A DNA Sequencing System. Анализ последовательности выявил две миссенс-мутации в гене Тац-полимеразы между сайтами NheI и BamHI. Мутация аминокислоты в положении 615 привела к тому, что остаток глутаминовой кислоты (E) был замещен остатком глицина (G), а другая мутация в положении 653 привела к замещению остатка аланина (A) на треонин (T). Нумерацию начинают с кодона, кодирующего первый остаток метионина зрелого протеина, как описано в патенте США

R U ? 2 3 5 7 3 C 2

5079352. Мутация E615G привела к замене GAG на GGG в кодоне 615. Мутация A653T привела к замене GCC на ACC в кодоне 653. Плазмида C1 в штамме-хозяине *E.coli* T1624 была депонирована в соответствии с Будапештским договором в ATCC 17 июля 1996 года и получила регистрационный номер 98107.

Две точечные мутации анализировали раздельно путем субклонирования каждой отдельно в гене Taq-полимеразы дикого типа, используя рекомбинантную ПЦР (под. ред Innis и др., PCR Protocols, San Diego, Academic Press, 1990, глава 22, озаглавленная "Recombinant PCR", автор Higuchi, стр.177-183). Образовавшиеся продукты экспрессии анализировали с целью определить, какая из мутаций E615G или A653T либо обе ответственны за фенотип, для которого характерно включение рибонуклеотидов. Результаты эксперимента показали, что за мутантный фенотип ответственна исключительно мутация E615G.

Для дальнейшего анализа и количественной оценки эффективности включения аналогов нуклеотидов ПЦР-фрагмент BamHI-NheI длиной 265 пар оснований, содержащий мутацию E615G, клонировали в векторе экспрессии Taq, обозначенном pRDA3-2. Вектор экспрессии pRDA3-2 содержал полноразмерный ген Taq, функционально связанный с промотором фага-лямбда PL. Экзонуклеазный домен гена Taq в этом векторе содержал точечную мутацию в кодоне, кодирующем глицин, т.е. аминокислотный остаток в положении 46, которая снижает 5'-3'-экзонуклеазную активность. Однако последовательность гена внутри полимеразного домена вектора экспрессии pRDA3-2 идентична последовательности гена Taq дикого типа. Плазмида pRDA3-2 полностью описана в патенте США 5466591, где плазмида обозначена как "клон 3-2". Плазмиду pRDA3-2 расщепляли с помощью BamHI и NheI и ПЦР-фрагмент длиной 265 пар оснований стандартными способами лigationали с вектором.

Полученной плазмидой pLK108 трансформировали штамм *E.coli* DG116 (Lawyer и др., 1993, см. выше, этот штамм также может быть получен из американской коллекции типовых культур под номером ATCC 53606). Плазмида pLK108 кодирует термостабильную ДНК-полимеразу, обозначенную в настоящем описании как G46D, E615G-Taq. С помощью рекомбинантной ПЦР путем комбинации мутаций E615G и P667Y во фрагменте BamHI-NheI получали мутант G46D, E615G, F667Y-Taq. Этот фрагмент клонировали в плазмиде pRDA3-2 для создания плазмиды pLK109. Экспрессируемый плазмидами pLK108 и pLK109 протеин термостабильной ДНК-полимеразы выделяли в соответствии с методом, описанным у Lawyer и др., 1993, см. выше, хотя стадии хроматографии были исключены. Представляющую интерес последовательность подтверждали с помощью анализа последовательности ДНК. Во вставке (инсерции) pLK108 была обнаружена дополнительная мутация в последовательности; однако эта мутация не изменяла аминокислотную последовательность протеина.

После частичной очистки активность модифицированного фермента определяли в соответствии с методом оценки активности, описанным у Lawyer и др., 1989, J. Biol. Chem. 264: 6427-6437. Активность модифицированного фермента оценивали следующим образом: одна единица фермента соответствует 10 нмолям продукта, синтезированного в течение 30 минут. ДНК-полимеразная активность фермента дикого типа прямо пропорциональна концентрации фермента, обеспечивающей включение до 80-100 пмолям ДЦМФ (разбавление фермента составляет до 0,12-0,15 единиц на реакцию). Активность мутантов E615G, G46D и E615G, F667Y, G46D прямо пропорциональна концентрациям, обеспечивающим включение до 0,25-3 пмолям ДЦМФ (разбавление фермента составляет от 6×10^{-4} до 5×10^{-3} единиц на реакцию). Этот препарат фермента использовали для оценки включения и для реакций секвенирования, описанных в примерах III-V. Для примеров II и VI фермент очищали в соответствии с методами, описанными у Lawyer и др. (см. выше).

Пример II

Анализ сравнительной эффективности включения

Относительную способность G46D- и G46D, E615G-Taq включать рНТФ определяли путем измерения количества [α -³²P]рНТФ, которое каждый из ферментов может включить на ДНК-матрице из активированной спермы лосося при ограниченной концентрации этого фермента. Для количественной оценки включения рАТФ реакционную смесь готовили таким образом, чтобы конечные концентрации в 50 мкл реакционной смеси составляли 12,5 мкг ДНК активированной спермы лосося, полученной по описанной ниже методике, 200 мкМ каждого из ДЦТФ, ДГТФ и ДТТФ (фирма Perkin Elmer, Ньюпорт, шт. Коннектикут), 100 мкМ [α -³²P]рАТФ, 1 мМ β -меркаптоэтанола, 25 мМ N-трип[гидроксиметил]метил-3-аминопропансульфоновой кислоты (TAPS), pH 9,5 при 20°C, 50 мМ KCl и 2,25 мМ MgCl₂.

Аналогичную смесь для анализа получали для количественной оценки включения рЦТФ, рГТФ и рУТФ. В каждом случае рНТФ были радиоактивно меченными и присутствовали в концентрации 100 мкМ, а каждый из трех остальных ДНТФ (ДАТФ, ДГТФ и ДТТФ для рЦТФ, ДАТФ, ДЦТФ и ДТТФ для рГТФ и ДАТФ, ДЦТФ и ДГТФ для рУТФ) присутствовал в концентрации 200 мкМ. В качестве стандарта также измеряли включение каждого ферментом соответствующего [α -³²P]ДНТФ. Смесь для анализа в этих опытах была аналогична таковой, применяемой в опыте по включению рНТФ, которая описана выше, за исключением того, что каждый [α -³²P]рНТФ был заменен 100 мкМ соответствующего [α -³²P]ДНТФ. Неочищенную ДНК спермы лосося в количестве 1 г/л, полученную от фирмы Worthington Biochemical (Фрихолд, шт. Нью Йорк), активировали путем инкубации в 10 мМ трис-HCl, pH 7,2, 5 мМ MgCl₂ при температуре 2-8°C в течение 96 часов. Затем добавляли ЭДТК и NaCl до концентрации 12,5 мМ и 0,1 М соответственно. Затем ДНК экстрагировали фенолом/хлороформом и далее осаждали этанолом и

ресуспенсировали в 10 мМ трис, 1 мМ ЭДТК, pH 7,5. Затем препарат активированной ДНК подвергали диализу по отношению к этому же буферу.

45 мкл каждой реакционной смеси разливали в виде аликовотных количеств в пять пробирок объемом 0,5 мл (например, в пробирки Эппendorфа), внося в каждую 5'-меченные предшественники нуклеотидов. Таким образом, каждую полимеразу G46D-Taq и G46D, E615G-Taq оценивали дважды, при этом одну пробирку оставляли в качестве отрицательного контроля. Смесь, полученную в результате реакции полимеризации, из каждого опыта в двух пробирках смешивали сначала с 5 мкл либо полимеразы G46D-Taq (0,02 единицы), либо G46D, E615G-Taq (0,002 единицы). В качестве контроля уровня фона к реакционной смеси, применяемой в качестве отрицательного контроля, вместо фермента добавляли 5 мкл буфера для разбавления фермента.

Каждую реакционную смесь в течение короткого интервала времени центрифугировали и инкубировали в течение 10 минут при температуре 75°C. Реакции прекращали, добавляя 10 мкл 60 мМ ЭДТК, и хранили на льду. Для каждого образца аликовоты объемом 50 мкл из 60 мкл реакционной смеси разбавляли 1 мл 2 мМ ЭДТК, 50 мкг/мл фрагментированной ДНК спермы лосося. ДНК осаждали с помощью трихлоруксусной кислоты (TXU), используя стандартные методы и собирая на фильтровальный диски типа GF/C (фирма Whatman, Кент, Великобритания). Количество включенного [α -³²P]-меченного нуклеотида или рибонуклеотида количественно оценивали с помощью жидкостной сцинтилляционной спектрометрии и затем подсчитывали количество включенных пмолей. Количество пмолей каждого рНТФ, включенное каждым ферментом, соотносили с количеством пмолей соответствующего [α -³²P]дНТФ, включенного каждым ферментом. В табл.3 приведены полученные данные.

Таблица 3

Отношение рНТФ к дНТФ, включенных с использованием полимераз G46D- и G46D, E615G-Taq

Фермент	Количество включенных пмолей (процент)							
	дАТФ	рАТФ	дЦТФ	рЦТФ	дГТФ	рГТФ	дТТФ	рТТФ
G46D	27,74 (100%)	0,052 (0,18%)	34,6 (100%)	0,76 (0,22%)	36,84 (100%)	0,133 (0,36%)	28,79 (100%)	0 (0)
G46D, E615G	0,67 (100%)	1,41 (210%)	2,82 (100%)	5,33 (189%)	3,27 (100%)	5,96 (181%)	0,688 (100%)	0,545 (79%)

Эти результаты показывают, что полимераза G46D, E615G включает рибонуклеотиды более чем в 500 раз эффективнее по сравнению с тем, как это может осуществлять полимераза G46D (например, эффективнее для рГТФ в 181:0,36=502 раза, для рЦТФ в 189:0,22=859 раз и для рАТФ в 210:0,18=1166 раз).

Таким образом, миссенс-мутация в гене полимеразы в кодоне 615 приводит к получению нового фенотипа: к получению термостабильной ДНК-полимеразы, обладающей способностью помимо дезоксирибонуклеотидов эффективно включать рибонуклеотиды.

Пример III

Анализ сравнительной эффективности включения 3'-дезоксиАТФ (кордицепина)

Относительную способность полимераз G46D-; G46D, E615G-; G46D, E615G, F667Y- и G46D, F667Y-Taq включать 3'-дезоксиаденоzin-5'-трифосфат (кордицепинтрифосфат) определяли путем измерения количества [α -³²P]кордицепинтрифосфата, которое каждый фермент может включить на ДНК-матрице из активированной спермы лосося при ограниченной концентрации этого фермента. Для количественной оценки включения [α -³²P] кордицепинтрифосфата опыт проводили таким образом, чтобы конечные концентрации в 50 мкл реакционной смеси составляли: 12,5 мкг ДНК активированной спермы лосося, 200 мкМ каждого из дЦТФ, дГТФ и дТТФ, 50 мкМ дАТФ (фирма Perkin Elmer), 50 мкМ [α -³²P]-3'-дАТФ/3'-дАТФ (фирма New England Nuclear, Sigma), 1 мМ β -меркаптоэтанола, 25 мМ N-трип[гидроксиметил]метил-3-аминопропансульфоновой кислоты (TAPS), pH 9,5 при 20°C, 55 мМ KCl и 2,25 мМ MgCl₂.

45 мкл каждой реакционной смеси разливали в виде аликовотных количеств в девять пробирок объемом 0,5 мл, т.е. каждую реакцию проводили либо с G46D-; G46D, E615G-; G46D, E615G, F667Y-; либо с G46D, F667Y-Taq в двух повторностях, при этом одну пробирку оставляли без фермента в качестве отрицательного контроля. Смесь, полученную в результате реакции полимеризации, для каждого опыта в двух пробирках смешивали сначала с 5 мкл (0,058 единиц) полимеразы G46D-Taq. Такую же операцию проводили с полимеразой G46D, E615G-Taq (0,0025 единиц), полимеразой G46D, E615G, F667Y-Taq (0,0034 единицы) или полимеразой G46D, F667Y-Taq (0,083 единицы). В качестве контроля уровня фона в одну оставшуюся пробирку вместо фермента вначале добавляли буфер для разбавления фермента.

Каждую реакционную смесь в течение короткого интервала времени центрифугировали и инкубировали в течение 10 минут при 75°C. Реакции прекращали, добавляя 10 мкл 60 мМ ЭДТК, и хранили на льду. Каждый образец аликовоты объемом 50 мкл из 60 мкл реакционной смеси разбавляли 1 мл 2 мМ ЭДТК, 50 мкг/мл фрагментированной ДНК спермы лосося. ДНК осаждали с помощью TXU, используя стандартные методы и собирая на фильтровальные диски типа GF/C (фирма Whatman, Кент, Великобритания). Количество включенного [α -³²P]-меченного нуклеотида оценивали с помощью жидкостной сцинтилляционной спектрометрии, после чего подсчитывали количество включенных пмолей. Количество пмолей [α -³²P] кордицепинтрифосфата, включенное каждым ферментом, соотносили с количеством единиц каждого фермента, примененного в опыте, для получения количества пмолей, включенных единицей фермента. В табл.4 приведены полученные данные.

R U ? 2 3 5 7 3 C 2

Таблица 4

Включение [α - 32 P]-кордицепина полимеразами G46D-; G46D, E615G-;

G46D, E615G, F667Y- и G46D, F667Y-Taq

Фермент	Количество пмольей, включенных единицей фермента
G46D	0,221
G46D, E615G	1,56
G46D, E615G, F667Y	893,6
G46D, F667Y	0,74

Эти результаты показывают, что обе мутации E615G и F667Y необходимы для эффективного включения молекулы кордицепина в ДНК.

Пример IV

Секвенирование ДНК на основе щелочного разложения с использованием ДНК-полимеразы G46D, E615G-Taq.

В этом примере показано применение модифицированной полимеразы по изобретению для секвенирования на основе щелочного разложения с использованием ДНК, частично замещенной рНТФ. Отношение рНТФ к дНТФ в реакционных смесях составляло от 1:80 до 1:8. Реакции удлинения праймера осуществляли в буфере, состоящем из 50 мМ Bicine (N,N-бис(2-гидроксигидил) глицин); pH 8,3), 25 мМ KOAc и 2,5 мМ MgCl₂. Проводили четыре отдельные реакции, по одной для каждого из четырех рНТФ. Каждая реакционная смесь (50 мкл) содержала по 200 мкМ каждого из дАТФ, дЦТФ, дГТФ и дТТФ (фирма Perkin Elmer) и по 0,09 пмольей одноцепочечной ДНК-матрицы M13mp18 (фирма Perkin Elmer), которую отжигали с 5'-[³²P]-меченным DG48 (Lawyer и др., 1993, PCR Methods and Applications 2: 275-287). Реакционные смеси также содержали 2,5, 2,5, 2,5 или 25 мкм рАТФ, рЦТФ, рГТФ или рУТФ соответственно.

Каждую из этих четырех реакций начинали путем добавления 7 единиц ДНК-полимеразы G46D, E615GB-Taq и инкубировали в течение 10 минут при температуре 75°C. Реакции прекращали добавлением 10 мкл 60 мМ ЭДТК и помещали на лед. 20 мкл каждой реакционной смеси добавляли к 80 мкл 50 мМ Bicine (pH 8,3), 25 мМ KOAc и 2,5 мМ MgCl₂. Продукты расщепления получали добавлением 7 мкл 1 н. NaOH и инкубированием в течение 15 минут при температуре 98°C. Реакции нейтрализовали добавлением 7 мкл 1 н. HCl. Каждую реакционную смесь осаждали, добавляя 312 мкл 95%-ного этанола и 10 мкл 3 М ацетата натрия (pH 4,8). Реакционные смеси микротитрографировали в течение 15 минут для сбора осадка, супернатант удаляли, дебрис промывали 500 мкл 70%-ного этанола и сушили. Каждый дебрис ре悬浮ировали в 5 мкл 0,5-кратного стоп-буфера (поставляемого фирмой Perkin Elmer, Норвулк, шт. Коннектикут, который содержит 95% формамида, 20 мМ ЭДТК и 0,05% бромфенола синего), выдерживали при температуре 98°C в течение 3 минут и непосредственно загружали на секвенирующий гель для ДНК, состоящий из предварительно подвергнутого электрофорезу 6%-ного полиакриламида/8 М мочевины, и осуществляли электрофорез. Гель сушили и снимали на рентгеновскую пленку. Полученная пленка показала выраженную секвенирующую лестницу, которая дает избыток в 100 оснований правильной последовательности.

Пример V

Секвенирование ДНК с использованием ДНК-полимеразы G46D, E615G, F667Y-Taq и 3'-дезоксинуклеотидтрифосфатов

В этом примере описано применение модифицированной полимеразы G46D, E615G, F667Y-Taq для секвенирования ДНК с использованием

3'-дезоксинуклеотидтрифосфатов. Этот эксперимент осуществляли с использованием 3'-дезоксиАТФ, однако он также может быть осуществлен с использованием других 3'-дезоксинуклеотидов. Реакции удлинения праймера осуществляли в буфере, состоящем из 50 мМ Bicine (pH 8,3), 25 мМ KOAc и 2,5 мМ MgCl₂. Каждая реакционная смесь (50 мкл) содержала по 200 мкМ каждого из дАТФ, дЦТФ, дГТФ и дТТФ (фирма Perkin Elmer) и по 0,09 пмольей одноцепочечной ДНК-матрицы M13mp18 (фирма Perkin Elmer), которую отжигали с 5'-[³²P]-меченным DG48 (Lawyer и др., 1993, PCR Methods and Applications 2: 275-287). Реакционные смеси также содержали 0, 0,1, 0,25, 0,5, 1 или 5 мкм 3'-дезоксиАТФ.

Каждую из реакций начинали путем добавления 7 единиц ДНК-полимеразы G46D, E615G, F667Y-Taq и инкубировали в течение 10 минут при 75°C. Реакции прекращали добавлением 10 мкл 60 мМ ЭДТК и помещали на лед. 30 мкл каждой реакционной смеси осаждали этанолом и ре悬浮ировали в стоп-буфере, выдерживали при температуре 98°C в течение 3 минут и непосредственно загружали на секвенирующий гель для ДНК, состоящий из предварительно подвергнутого электрофорезу 6%-ного полиакриламида/8 М мочевины, и осуществляли электрофорез. Гель сушили и снимали на рентгеновскую пленку. Полосы, которые содержали продукты реакции, проведенной в присутствии кордицепина, свидетельствовали о наличии ясно выраженных лестниц терминации.

Полосы, содержащие наибольшее количество кордицепина, т.е. 5 мкМ, показали лестницу терминации, в которой в среднем полосы были короче по длине по сравнению с полосами, соответствующими более низким уровням кордицепина. Полосы, которые содержали продукты реакции, проведенной в отсутствии кордицепина, свидетельствовали о наличии продукта, имеющего полную длину, и показали отсутствие лестницы терминации. Эти результаты показывают, что мутантный фермент способен включать кордицепин, и включение этой молекулы в продукт удлинения праймера вызывает терминацию. Этот метод также может применяться для создания лестницы секвенирования ДНК с использованием 3'-дезоксиЦТФ, 3'-дезоксиГТФ и 3'-дезоксиУТФ.

Пример VI

ПЦР-секвенирование с окрашенным праймером с использованием ДНК-полимеразы G46D, E615G-Taq

Этот метод иллюстрирует применение модифицированной полимеразы по изобретению для секвенирования с окрашенным праймером с использованием рибонуклеозидтрифосфатов (рНТФ) в ПЦР при соотношении рНТФ:дНТФ не более 1:30. Проводили четыре отдельные реакции, каждую с использованием рНТФ. Реакции секвенирования на основе ПЦР проводили в буфере, состоящем из 25 мМ трис-HCl (pH 9),

R U ? 2 3 5 7 7 3 C 2

5,0 мМ MgCl₂ и 10%-ного глицерина (объем/объем). Каждая реакционная смесь также содержала по 500 мкМ каждого из дАТФ, дЦТФ, дГТФ и дТТФ (фирма Perkin Elmer), матрицу, представляющую собой 5×10⁶ копий/мкл плазиды pBSM13+ (фирма Stratagene), линеаризованной в результате обработки эндонуклеазой рестрикции XmnI, и 0,05 единиц/мкл ДНК-полимеразы G46D, E615G-Taq. Реакционные смеси рибо-АТФ (10 мкл) содержали 2,5 мкМ АТФ (фирма Pharmacia Biotech), 0,1 мкл праймера JOE M13 Reverse Dye Primer (фирма Perkin Elmer) и 0,1 мкМ праймера ASC46 (5'-CGCCATTGCCATTCAAG). Реакционные смеси рибо-ЦТФ (10 мкл) содержали 2,5 мкМ ЦТФ (фирма Pharmacia Biotech), 0,1 мкл праймера FAM M13 Reverse Dye Primer (фирма Perkin Elmer) и 0,1 мкМ праймера ASC46. Реакционные смеси рибо-ГТФ (20 мкл) содержали 2,5 мкМ ГТФ (фирма Pharmacia Biotech), 0,1 мкл праймера TAMRA M13 Reverse Dye Primer (фирма Perkin Elmer) и 0,1 мкМ праймера ASC46. Реакционные смеси рибо-УТФ (20 мкл) содержали 16 мкМ УТФ (фирма Pharmacia Biotech), 0,1 мкл праймера ROX M13 Reverse Dye Primer (фирма Perkin Elmer) и 0,1 мкМ праймера ASC46.

Каждую из четырех реакционных смесей помещали в предварительно нагретую (75°C) термоячейку для проведения ПЦР типа Perkin Elmer GeneAMP® PCR System 9600 и подвергали 30 циклам: выдержка при температуре 95°C в течение 10 секунд, при температуре 55°C в течение 10 секунд, постепенное понижение температуры до 65°C в течение 1 минуты и выдержка при 65°C в течение 5 минут. В результате каждой из реакций с участием рАТФ и рЦТФ получали по 6×10¹¹ копий меченого красителем амплифицированного продукта длиной 300 пар оснований, а в результате реакций с участием рГТФ и УТФ получали по 1,2×10¹² копий меченого красителем амплифицированного продукта длиной 300 пар оснований.

Для определения последовательности ДНК амплифицированных продуктов ПЦР без использования разделения с помощью ферментативных реакций секвенирования ДНК реакционные смеси объединяли, обрабатывали щелочью и нагревали, нейтрализовали и осаждали следующим образом. Объединяли 4 мкл каждой из реакционных смесей с использованием АТФ и ЦТФ и 8 мкл каждой из реакционных смесей с использованием ГТФ и УТФ. К объединенным реакционным смесям добавляли 2 мкл 0,25 М ЭДТК (рН 8,0) (конечная концентрация 10 мМ), 10 мкл 1 М NaOH (конечная концентрация 200 мМ) и 14 мкл H₂O, а затем инкубировали при температуре 95°C в течение 5 минут в термоячейке для проведения ПЦР типа Perkin Elmer GeneAMP® PCR System 9600 и нейтрализовали 10 мкм 1 М HCl. Затем объединенную реакционную смесь осаждали добавлением 150 мкл 95%-ного этанола с последующей инкубацией при температуре 4°C в течение 15 минут. Затем ее микроцентрифугировали в течение 15 минут при температуре 4°C для сбора осадка и супернатант удаляли путем аспирации. Дебрис промывали 300 мкл 70%-ного этанола, микроцентрифугировали в течение 5 минут,

супернатант удаляли путем аспирации и дебрис сушили. Дебрис реусспендирували в 6 мкл формамида, 50 мг/мл голубого декстрана (в 25 мМ ЭДТК, 5:1 (объем/объем)) и выдерживали при температуре 90°C в течение 3 минут. 1,5 мкл реусспендированного дебриса непосредственно загружали на секвенирующий гель, состоящий из предварительно подвергнутого электрофорезу 5%-ного полиакриламида типа Long Ranger (фирма FMC BioProducts) с 6 М мочевиной. Затем проводили электрофорез и анализировали на ДНК-секвенаторе типа Perkin Elmer ABI Prism™ 377 DNA Sequencer в соответствии с инструкцией производителя. Автоматическое определение оснований на основе программного обеспечения для анализа последовательности Perkin Elmer ABI Prism™ 377 Sequencing Analysis показало более чем 99%-ную точность определения последовательности ДНК продукта длиной 300 пар оснований, амплифицированного с использованием ПЦР.

Формула изобретения:

1. Термостабильная ДНК-полимераза, представляющая собой производное природной термостабильной ДНК-полимеразы, включающей аминокислотную последовательность - мотив SerGlnIleGluLeuArgXaa (SEQ ID NO: 2), где Xaa в положении 7 этой последовательности обозначает остаток валина (Val) или остаток изолейцина (Ile), которое получено в результате замены остатка глутаминовой кислоты (Glu) в положении 4 этой последовательности и отбора мутантной (модифицированной) формы, характеризующейся пониженной по сравнению с соответствующим природным ферментом способностью к ограничению включения не свойственных нуклеотидов, предпочтительно рНТФ.

2. Термостабильная ДНК-полимераза по п.1, отличающаяся тем, что она представляет собой рекомбинантное производное природной термостабильной ДНК-полимеразы.

3. Термостабильная ДНК-полимераза по любому из пп.1-2, отличающаяся тем, что ее способность к включению не свойственных нуклеотидов по сравнению со способностью соответствующей природной полимеразы включать не свойственный нуклеотид возрастает по крайней мере в 20 раз.

4. Термостабильная ДНК-полимераза по любому из пп.1-3, отличающаяся тем, что она имеет достаточную активность для применения в реакции секвенирования ДНК, предусматривающей использование не свойственного нуклеотида, предпочтительно рНТФ, и соответствующего обычного нуклеотида при соотношении 1:1 или менее.

5. Термостабильная ДНК-полимераза по любому из пп.1-3, отличающаяся тем, что она имеет достаточную активность для применения в реакции секвенирования ДНК, предусматривающей использование не свойственного нуклеотида, предпочтительно рНТФ, присутствующего в концентрации менее, чем приблизительно 100 мкМ, и соответствующего обычного нуклеотида, присутствующего в концентрации более чем приблизительно 100 мкМ.

R U ? 2 3 5 7 7 3 C 2

R U
2 2
3 3
5 5
7 7
3

C 2

6. Термостабильная ДНК-полимераза по любому из пп.2-5, которая представляет собой рекомбинантное производное природной термостабильной ДНК-полимеразы из организма, выбранного из группы, включающей *Thermus aquaticus*, *Thermus caldophilus*, *Thermus chlariophilus*, *Thermus filiformis*, *Thermus flavus*, *Thermus oshimai*, *Thermus ruber*, *Thermus scotoductus*, *Thermus silvanus*, *Thermus* вид Z05, *Thermus* вид spsI7, *Thermus thermophilus*, *Thermotoga maritima*, *Thermotoga neapolitana*, *Thermosiphon africanus*, *Anaerocellum thermophilum*, *Bacillus caldotenax* и *Bacillus stearothermophilus*.

7. Термостабильная ДНК-полимераза по любому из пп.2-5, которая представляет собой рекомбинантное производное природной термостабильной ДНК-полимеразы из видов рода *Thermus*, предпочтительно термостабильной ДНК-полимеразы, которая включает аминокислотную последовательность
LeuAspTyrSerGlnIleGluLeuArgValLeuAlaHisLeu Ser (SEQ ID NO: 5).

8. Термостабильная ДНК-полимераза по любому из пп.1-5, последовательность которой гомологична аминокислотной последовательности ДНК-полимеразы Таq (SEQ ID NO: 7), по крайней мере, приблизительно на 39%, предпочтительно по крайней мере приблизительно на 60%, более предпочтительно по крайней мере приблизительно на 80%.

9. Нуклеотидная последовательность, кодирующая термостабильную ДНК-полимеразу по любому из пп.1-8.

10. Вектор, обеспечивающий экспрессию термостабильной ДНК-полимеразы и содержащий нуклеотидную последовательность по п.9.

11. Штамм бактерий *E.coli* ATCC № 98107 - продуцент термостабильной ДНК-полимеразы.

12. Способ получения термостабильной ДНК-полимеразы, включающий

(а) культивирование штамма по п.11 в условиях, которые усиливают экспрессию термостабильной ДНК-полимеразы, и (б) выделение термостабильной ДНК-полимеразы.

13. Термостабильная ДНК-полимераза, полученная способом по п.12.

14. Композиция для применения в реакции секвенирования ДНК, которая содержит нуклеиновую кислоту-матрицу; олигонуклеотидный праймер, комплементарный этой матрице; термостабильную ДНК-полимеразу по любому из пп.1-8; смесь дНТФ и по крайней мере один несвойственный нуклеотид, предпочтительно рНТФ, причем соотношение несвойственного

нуклеотида соответствующих обычных нуклеотид составляет 1:1 или менее.

15. Композиция по п.14, где несвойственный нуклеотид представляет собой рибонуклеотид, причем этот рибонуклеотид предпочтительно присутствует в концентрации ниже чем приблизительно 100 мкМ, а соответствующий обычный нуклеотид присутствует в концентрации выше чем приблизительно 100 мкМ.

16. Композиция по п.15, отличающаяся тем, что несвойственный нуклеотид является немеченым.

17. Способ секвенирования нуклеиновой кислоты-мишени, включающий следующие стадии:

(а) подготовку несвойственного нуклеотида, предпочтительно рНТФ, и соответствующего обычного нуклеотида для реакции секвенирования ДНК, причем несвойственный нуклеотид и соответствующий обычный нуклеотид присутствуют в соотношении менее чем приблизительно 1:1; (б) обработку реакционной смеси со стадии (а) в присутствии термостабильной ДНК-полимеразы по любому из пп.1-8 в условиях, пригодных для удлинения праймера, с получением продуктов удлинения праймера, содержащих несвойственный нуклеотид; (в) обработку продуктов удлинения праймера со стадии (б) в условиях, пригодных для гидролиза этих продуктов удлинения праймера; (г) разделение продуктов реакции со стадии (в) и (д) определение последовательности нуклеиновой кислоты-мишени.

18. Способ секвенирования по п.17, где несвойственный нуклеотид представляет собой рибонуклеотид, причем этот рибонуклеотид предпочтительно присутствует в концентрации приблизительно 0,1-100 мкМ.

19. Способ секвенирования по п.17, где соответствующий обычный нуклеотид предпочтительно присутствует в концентрации приблизительно 50-500 мкМ.

20. Набор для секвенирования нуклеиновой кислоты, содержащий термостабильную ДНК-полимеразу по любому из пп.1-8; буферы, используемые в методе секвенирования; один или несколько олигонуклеотидных праймеров; смесь дНТФ и по крайней мере один несвойственный нуклеотид, предпочтительно рНТФ, причем соотношение между несвойственным нуклеотидом и соответствующим обычным нуклеотидом предпочтительно меньше 1.

Конвенционный приоритет установлен от 06.08.1996 по пп.1-20 в соответствии с заявкой 60-023.376, поданной в патентное ведомство США.

R U ? 2 3 5 7 7 3 C 2

ПЕРЕЧЕНЬ ПОСЛЕДОВАТЕЛЬНОСТЕЙ

(1) ОБЩАЯ ИНФОРМАЦИЯ:

(i) ЗАЯВИТЕЛЬ:

- (A) НАИМЕНОВАНИЕ: F.Hoffmann-La Roche Ltd.
- (B) УЛИЦА: Grenzacherstrasse 124
- (C) ГОРОД: Базель
- (D) КАНТОН: BS
- (E) СТРАНА: Швейцария
- (F) ПОЧТОВЫЙ КОД (ZIP): CH-4070
- (G) ТЕЛЕФОН: (0) 61 6882403
- (H) ТЕЛЕФАКС: (0) 61 6881395
- (I) ТЕЛЕКС: 962292/965512 hlr ch

(ii) НАЗВАНИЕ ИЗОБРЕТЕНИЯ: Модифицированная термостабильная ДНК-полимераза

(iii) КОЛИЧЕСТВО ПОСЛЕДОВАТЕЛЬНОСТЕЙ: 8

(iv) ФОРМА ПРЕДСТАВЛЕНИЯ ДЛЯ КОМПЬЮТЕРА:

- (A) ТИП НОСИТЕЛЯ: флоппи-диск
- (B) КОМПЬЮТЕР: совместимый с IBM PC
- (C) ОПЕРАЦИОННАЯ СИСТЕМА: PC-DOS/MS-DOS
- (D) ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ: PatentIn Release #1.0, версия #1.30 (EPO)

(vi) ДАННЫЕ О ПРИОРИТЕТНОЙ ЗАЯВКЕ:

- (A) НОМЕР ЗАЯВКИ: US 60/023376
- (B) ДАТА ПОДАЧИ: 6 августа 1996

(2) ИНФОРМАЦИЯ О ПОСЛЕДОВАТЕЛЬНОСТИ SEQ ID NO: 1:

(i) ХАРАКТЕРИСТИКИ ПОСЛЕДОВАТЕЛЬНОСТИ:

- (A) ДЛИНА: 7 аминокислот
- (B) ТИП: аминокислота
- (D) ТОПОЛОГИЯ: линейная

(ii) ТИП МОЛЕКУЛЫ: пептид

(ix) ПРИЗНАК:

- (A) ИМЯ/КЛЮЧ: пептид
- (B) ПОЛОЖЕНИЕ: 4
- (D) ДРУГАЯ ИНФОРМАЦИЯ: /label = Xaa
/note = "где Xaa представляет собой любую аминокислоту, кроме Glu"

(ix) ПРИЗНАК:

- (A) ИМЯ/КЛЮЧ: пептид

R U ? 2 3 5 7 7 3 C 2

R U
2 2 3 5 7 7 3

C 2

- (B) ПОЛОЖЕНИЕ: 7
(D) ДРУГАЯ ИНФОРМАЦИЯ: /label = Xaa
/note = "где Xaa представляет собой Ile или Val"

(xi) ОПИСАНИЕ ПОСЛЕДОВАТЕЛЬНОСТИ: SEQ ID NO: 1:

Ser Gln Ile Xaa Leu Arg Xaa
1 5

(2) ИНФОРМАЦИЯ О ПОСЛЕДОВАТЕЛЬНОСТИ SEQ ID NO: 2:

(i) ХАРАКТЕРИСТИКИ ПОСЛЕДОВАТЕЛЬНОСТИ:

- (A) ДЛИНА: 7 аминокислот
(B) ТИП: аминокислота
(D) ТОПОЛОГИЯ: линейная

(ii) ТИП МОЛЕКУЛЫ: пептид

(ix) ПРИЗНАК:

- (A) ИМЯ/КЛЮЧ: пептид
(B) ПОЛОЖЕНИЕ: 7
(D) ДРУГАЯ ИНФОРМАЦИЯ: /label = Xaa
/note = "где Xaa представляет собой Ile или Val"

(xi) ОПИСАНИЕ ПОСЛЕДОВАТЕЛЬНОСТИ: SEQ ID NO: 2:

Ser Gln Ile Glu Leu Arg Xaa
1 5

(2) ИНФОРМАЦИЯ О ПОСЛЕДОВАТЕЛЬНОСТИ SEQ ID NO: 3:

(i) ХАРАКТЕРИСТИКИ ПОСЛЕДОВАТЕЛЬНОСТИ:

- (A) ДЛИНА: 7 аминокислот
(B) ТИП: аминокислота
(D) ТОПОЛОГИЯ: линейная

(ii) ТИП МОЛЕКУЛЫ: пептид

(xi) ОПИСАНИЕ ПОСЛЕДОВАТЕЛЬНОСТИ: SEQ ID NO: 3:

Ser Gln Ile Glu Leu Arg Val
1 5

(2) ИНФОРМАЦИЯ О ПОСЛЕДОВАТЕЛЬНОСТИ SEQ ID NO: 4:

(i) ХАРАКТЕРИСТИКИ ПОСЛЕДОВАТЕЛЬНОСТИ:

- (A) ДЛИНА: 7 аминокислот
(B) ТИП: аминокислота

(D) ТОПОЛОГИЯ: линейная

(ii) ТИП МОЛЕКУЛЫ: пептид

(xi) ОПИСАНИЕ ПОСЛЕДОВАТЕЛЬНОСТИ: SEQ ID NO: 4:

Ser Gln Ile Glu Leu Arg Ile
1 5

(2) ИНФОРМАЦИЯ О ПОСЛЕДОВАТЕЛЬНОСТИ SEQ ID NO: 5:

(i) ХАРАКТЕРИСТИКИ ПОСЛЕДОВАТЕЛЬНОСТИ:

- (A) ДЛИНА: 15 аминокислот
- (B) ТИП: аминокислота
- (D) ТОПОЛОГИЯ: линейная

(ii) ТИП МОЛЕКУЛЫ: пептид

(xi) ОПИСАНИЕ ПОСЛЕДОВАТЕЛЬНОСТИ: SEQ ID NO: 5:

Leu Asp Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala His Leu Ser
1 5 10 15

(2) ИНФОРМАЦИЯ О ПОСЛЕДОВАТЕЛЬНОСТИ SEQ ID NO: 6:

(i) ХАРАКТЕРИСТИКИ ПОСЛЕДОВАТЕЛЬНОСТИ:

- (A) ДЛИНА: 15 аминокислот
- (B) ТИП: аминокислота
- (D) ТОПОЛОГИЯ: линейная

(ii) ТИП МОЛЕКУЛЫ: пептид

(xi) ОПИСАНИЕ ПОСЛЕДОВАТЕЛЬНОСТИ: SEQ ID NO: 6:

Leu Asp Tyr Ser Gln Ile Gly Leu Arg Val Leu Ala His Leu Ser
1 5 10 15

(2) ИНФОРМАЦИЯ О ПОСЛЕДОВАТЕЛЬНОСТИ SEQ ID NO: 7:

(i) ХАРАКТЕРИСТИКИ ПОСЛЕДОВАТЕЛЬНОСТИ:

- (A) ДЛИНА: 2626 пар оснований
- (B) ТИП: нуклеиновая кислота
- (C) ХАРАКТЕРИСТИКА ЦЕПИ: двухцепочечная
- (D) ТОПОЛОГИЯ: линейная

(ii) ТИП МОЛЕКУЛЫ: ДНК (геномная)

(iii) ГИПОТЕТИЧНОСТЬ: нет

(iv) АНТИСМЫСЛОВАЯ: нет

(vi) ЕСТЕСТВЕННЫЙ ИСТОЧНИК:

(A) ОРГАНИЗМ: *Thermus aquaticus*

(ix) ПРИЗНАК:

(A) ИМЯ/КЛЮЧ: CDS

(B) ПОЛОЖЕНИЕ: 121...2616

(xi) ОПИСАНИЕ ПОСЛЕДОВАТЕЛЬНОСТИ: SEQ ID NO: 7:

AAGCTCAGAT	CTACCTGCCT	GAGGGCGTCC	GGTTCCAGCT	GGCCCTTCCC	GAGGGGGAGA	60	
GGGAGGCCGT	TCTAAAAGCC	CTTCAGGACG	CTACCCGGGG	GCGGGTGGTG	GAAGGGTAAC	120	
ATG AGG GGG ATG CTG CCC CTC TTT GAG CCC AAG GGC CGG GTC CTC CTG	Met Arg Gly Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu Leu					168	
1	5	10	15				
GTG GAC GGC CAC CAC CTG GCC TAC CGC ACC TTC CAC GCC CTG AAG GGC	Val Asp Gly His His Leu Ala Tyr Arg Thr Phe His Ala Leu Lys Gly					216	
20	25	30					
CTC ACC ACC AGC CGG GGG GAG CCG GTG CAG GCG GTC TAC GGC TTC GCC	Leu Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe Ala					264	
35	40	45					
AAG AGC CTC CTC AAG GCC CTC AAG GAG GAC GGG GAC GCG GTG ATC GTG	Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Asp Ala Val Ile Val					312	
50	55	60					
GTC TTT GAC GCC AAG GCC CCC TCC TTC CGC CAC GAG GCC TAC GGG GGG	Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Gly Gly					360	
65	70	75	80				
TAC AAG GCG GGC CGG GCC CCC ACG CCG GAG GAC TTT CCC CGG CAA CTC	Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln Leu					408	
85	90	95					
GCC CTC ATC AAG GAG CTG GTG GAC CTC CTG GGG CTG GCG CGC CTC GAG	Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Leu Ala Arg Leu Glu					456	
100	105	110					
GTC CCG GGC TAC GAG GCG GAC GAC GTC CTG GCC AGC CTG GCC AAG AAG	Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Ser Leu Ala Lys Lys					504	
115	120	125					
GCG GAA AAG GAG GGC TAC GAG GTC CGC ATC CTC ACC GCC GAC AAA GAC	Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Lys Asp					552	
130	135	140					
CTT TAC CAG CTC CTT TCC GAC CGC ATC CAC GTC CTC CAC CCC GAG GGG	Leu Tyr Gln Leu Leu Ser Asp Arg Ile His Val Leu His Pro Glu Gly					600	
145	150	155	160				

R U
2 2
3 3
5 5
7 7
3 3

? 2 3 5 7 7 3 C 2

C 2

R	U	2	2	3	5	7	7	3	C	2					
TAC	CTC	ATC	ACC	CCG	GCC	TGG	CTT	TGG	GAA	AAG	TAC	GGC	CTG	AGG	CCC
Tyr	Leu	Ile	Thr	Pro	Ala	Trp	Leu	Trp	Glu	Lys	Tyr	Gly	Leu	Arg	Pro
165									170						175
GAC	CAG	TGG	GCC	GAC	TAC	CGG	GCC	CTG	ACC	GGG	GAC	GAG	TCC	GAC	AAC
Asp	Gln	Trp	Ala	Asp	Tyr	Arg	Ala	Leu	Thr	Gly	Asp	Glu	Ser	Asp	Asn
180									185						190
CTT	CCC	GGG	GTC	AAG	GGC	ATC	GGG	GAG	AAG	ACG	GCG	AGG	AAG	CTT	CTG
Leu	Pro	Gly	Val	Lys	Gly	Ile	Gly	Glu	Lys	Thr	Ala	Arg	Lys	Leu	Leu
195									200						205
GAG	GAG	TGG	GGG	AGC	CTG	GAA	GCC	CTC	CTC	AAG	AAC	CTG	GAC	CGG	CTG
Glu	Glu	Trp	Gly	Ser	Leu	Glu	Ala	Leu	Leu	Lys	Asn	Leu	Asp	Arg	Leu
210									215						220
AAG	CCC	GCC	ATC	CGG	GAG	AAG	ATC	CTG	GCC	CAC	ATG	GAC	GAT	CTG	AAG
Lys	Pro	Ala	Ile	Arg	Glu	Lys	Ile	Leu	Ala	His	Met	Asp	Asp	Leu	Lys
225									230						240
CTC	TCC	TGG	GAC	CTG	GCC	AAG	GTG	CGC	ACC	GAC	CTG	CCC	CTG	GAG	GTG
Leu	Ser	Trp	Asp	Leu	Ala	Lys	Val	Arg	Thr	Asp	Leu	Pro	Leu	Glu	Val
245									250						255
GAC	TTC	GCC	AAA	AGG	CGG	GAG	CCC	GAC	CGG	GAG	AGG	CTT	AGG	GCC	TTT
Asp	Phe	Ala	Lys	Arg	Arg	Glu	Pro	Asp	Arg	Glu	Arg	Leu	Arg	Ala	Phe
260									265						270
CTG	GAG	AGG	CTT	GAG	TTT	GGC	AGC	CTC	CTC	CAC	GAG	TTC	GGC	CTT	CTG
Leu	Glu	Arg	Leu	Glu	Phe	Gly	Ser	Leu	Leu	His	Glu	Phe	Gly	Leu	Leu
275									280						285
GAA	AGC	CCC	AAG	GCC	CTG	GAG	GAG	GCC	CCC	TGG	CCC	CCG	CCG	GAA	GGG
Glu	Ser	Pro	Lys	Ala	Leu	Glu	Glu	Ala	Pro	Trp	Pro	Pro	Pro	Glu	Gly
290									295						300
GCC	TTC	GTG	GGC	TTT	GTG	CTT	TCC	CGC	AAG	GAG	CCC	ATG	TGG	GCC	GAT
Ala	Phe	Val	Gly	Phe	Val	Leu	Ser	Arg	Lys	Glu	Pro	Met	Trp	Ala	Asp
305									310						320
CTT	CTG	GCC	CTG	GCC	GCC	AGG	GGG	GGC	CGG	GTC	CAC	CGG	GCC	CCC	
Leu	Leu	Ala	Leu	Ala	Ala	Arg	Gly	Gly	Arg	Val	His	Arg	Ala	Pro	
325									330						335
GAG	CCT	TAT	AAA	GCC	CTC	AGG	GAC	CTG	AAG	GAG	GCG	CGG	GGG	CTT	CTC
Glu	Pro	Tyr	Lys	Ala	Leu	Arg	Asp	Leu	Lys	Glu	Ala	Arg	Gly	Leu	Leu
340									345						350
GCC	AAA	GAC	CTG	AGC	GTT	CTG	GCC	CTG	AGG	GAA	GGC	CTT	GGC	CTC	CCG
Ala	Lys	Asp	Leu	Ser	Val	Leu	Ala	Leu	Arg	Glu	Gly	Leu	Gly	Leu	Pro
355									360						365
CCC	GGC	GAC	GAC	CCC	ATG	CTC	CTC	GCC	TAC	CTC	CTG	GAC	CCT	TCC	AAC
Pro	Gly	Asp	Asp-Pro	Met	Leu	Leu	Ala	Tyr	Leu	Leu	Asp	Pro	Ser	Asn	
370									375						380

R	U	2	3	5	7	7	3	C	2
ACC ACC CCC GAG GGG GTG GCC CGG CGC TAC GGC GGG GAG TGG ACG GAG Thr Thr Pro Glu Gly Val Ala Arg Arg Tyr Gly Gly Glu Trp Thr Glu 385 390 395 400									1320
GAG GCG GGG GAG CGG GCC CTT TCC GAG AGG CTC TTC GCC AAC CTG Glu Ala Gly Glu Arg Ala Ala Leu Ser Glu Arg Leu Phe Ala Asn Leu 405 410 415									1368
TGG GGG AGG CTT GAG GGG GAG GAG AGG CTC CTT TGG CTT TAC CGG GAG Trp Gly Arg Leu Glu Gly Glu Arg Leu Leu Trp Leu Tyr Arg Glu 420 425 430									1416
GTG GAG AGG CCC CTT TCC GCT GTC CTG GCC CAC ATG GAG GCC ACG GGG Val Glu Arg Pro Leu Ser Ala Val Leu Ala His Met Glu Ala Thr Gly 435 440 445									1464
GTG CGC CTG GAC GTG GCC TAT CTC AGG GCC TTG TCC CTG GAG GTG GCC Val Arg Leu Asp Val Ala Tyr Leu Arg Ala Leu Ser Leu Glu Val Ala 450 455 460									1512
GAG GAG ATC GCC CGC CTC GAG GCC GAG GTC TTC CGC CTG GCC GGC CAC Glu Glu Ile Ala Arg Leu Glu Ala Glu Val Phe Arg Leu Ala Gly His 465 470 475 480									1560
CCC TTC AAC CTC AAC TCC CGG GAC CAG CTG GAA AGG GTC CTC TTT GAC Pro Phe Asn Leu Asn Ser Arg Asp Gln Leu Glu Arg Val Leu Phe Asp 485 490 495									1608
GAG CTA GGG CTT CCC GCC ATC GGC AAG ACG GAG AAG ACC GGC AAG CGC Glu Leu Gly Leu Pro Ala Ile Gly Lys Thr Glu Lys Thr Gly Lys Arg 500 505 510									1656
TCC ACC AGC GCC GCC GTC CTG GAG GCC CTC CGC GAG GCC CAC CCC ATC Ser Thr Ser Ala Ala Val Leu Glu Ala Leu Arg Glu Ala His Pro Ile 515 520 525									1704
G TG GAG AAG ATC CTG CAG TAC CGG GAG CTC ACC AAG CTG AAG AGC ACC Val Glu Lys Ile Leu Gln Tyr Arg Glu Leu Thr Lys Leu Lys Ser Thr 530 535 540									1752
TAC ATT GAC CCC TTG CCG GAC CTC ATC CAC CCC AGG ACG GGC CGC CTC Tyr Ile Asp Pro Leu Pro Asp Leu Ile His Pro Arg Thr Gly Arg Leu 545 550 555 560									1800
CAC ACC CGC TTC AAC CAG ACG GGC ACG GGC AGG CTA AGT AGC His Thr Arg Phe Asn Gln Thr Ala Thr Gly Arg Leu Ser Ser 565 570 575									1848
TCC GAT CCC AAC CTC CAG AAC ATC CCC GTC CGC ACC CCG CTT GGG CAG Ser Asp Pro Asn Leu Gln Asn Ile Pro Val Arg Thr Pro Leu Gly Gln 580 585 590									1896
AGG ATC CGC CGG GCC TTC ATC GCC GAG GAG GGG TGG CTA TTG GTG GCC Arg Ile Arg Arg Ala Phe Ile Ala Glu Glu Gly Trp Leu Leu Val Ala 595 600 605									1944

R U 2	2	3	5	7	7	3	C 2
CTG GAC TAT AGC CAG ATA GGG CTC AGG GTG CTG GCC CAC CTC TCC GGC Leu Asp Tyr Ser Gln Ile Gly Leu Arg Val Leu Ala His Leu Ser Gly 610 615 620							1992
GAC GAG AAC CTG ATC CGG GTC TTC CAG GAG GGG CGG GAC ATC CAC ACG Asp Glu Asn Leu Ile Arg Val Phe Gln Glu Gly Arg Asp Ile His Thr 625 630 635 640							2040
GAG ACC GCC AGC TGG ATG TTC GGC GTC CCC CGG GAG GCC GTG GAC CCC Glu Thr Ala Ser Trp Met Phe Gly Val Pro Arg Glu Ala Val Asp Pro 645 650 655							2088
CTG ATG CGC CGG GCG GCC AAG ACC ATC AAC TTC GGG GTC CTC TAC GGC Leu Met Arg Arg Ala Ala Lys Thr Ile Asn Phe Gly Val Leu Tyr Gly 660 665 670							2136
ATG TCG GCC CAC CGC CTC TCC CAG GAG CTA GCC ATC CCT TAC GAG GAG Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile Pro Tyr Glu Glu 675 680 685							2184
GCC CAG GCC TTC ATT GAG CGC TAC TTT CAG AGC TTC CCC AAG GTG CGG Ala Gln Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys Val Arg 690 695 700							2232
GCC TGG ATT GAG AAG ACC CTG GAG GAG GGC AGG AGG CGG GGG TAC GTG Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Arg Arg Gly Tyr Val 705 710 715 720							2280
GAG ACC CTC TTC GGC CGC CGC TAC GTG CCA GAC CTA GAG GCC CGG Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Glu Ala Arg 725 730 735							2328
GTG AAG AGC GTG CGG GAG GCG GCC GAG CGC ATG GCC TTC AAC ATG CCC Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met Pro 740 745 750							2376
GTC CAG GGC ACC GCC GCC CTC ATG AAG CTG GCT ATG GTG AAG CTC Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys Leu 755 760 765							2424
TTC CCC AGG CTG GAG GAA ATG GGG GCC AGG ATG CTC CTT CAG GTC CAC Phe Pro Arg Leu Glu Met Gly Ala Arg Met Leu Leu Gln Val His 770 775 780							2472
GAC GAG CTG GTC CTC GAG GCC CCA AAA GAG AGG GCG GAG GCC GTG GCC Asp Glu Leu Val Leu Glu Ala Pro Lys Glu Arg Ala Glu Ala Val Ala 785 790 795 800							2520
CGG CTG GCC AAG GAG GTC ATG GAG GGG GTG TAT CCC CTG GCC GTG CCC Arg Leu Ala Lys Glu Val Met Glu Gly Val Tyr Pro Leu Ala Val Pro 805 810 815							2568
CTG GAG GTG GAG GTG GGG ATA GGG GAG GAC TGG CTC TCC GCC AAG GAG Leu Glu Val Glu Val Gly Ile Gly Glu Asp Trp Leu Ser Ala Lys Glu 820 825 830							2616
TGATACCACC							2626

(2) ИНФОРМАЦИЯ О ПОСЛЕДОВАТЕЛЬНОСТИ SEQ ID NO: 8:

(i) ХАРАКТЕРИСТИКИ ПОСЛЕДОВАТЕЛЬНОСТИ:

- (A) ДЛИНА: 832 аминокислоты
(B) ТИП: аминокислота
(D) ТОПОЛОГИЯ: линейная

(ii) ТИП МОЛЕКУЛЫ: протеин

(xi) ОПИСАНИЕ ПОСЛЕДОВАТЕЛЬНОСТИ: SEQ ID NO: 8:

Met Arg Gly Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu Leu	1 5 10 15	
Val Asp Gly His His Leu Ala Tyr Arg Thr Phe His Ala Leu Lys Gly	20 25 30	
Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe Ala	35 40 45	
Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Asp Ala Val Ile Val	50 55 60	C 2
Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Gly Gly	65 70 75 80	
Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln Leu	85 90 95	
Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Leu Ala Arg Leu Glu	100 105 110	
Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Ser Leu Ala Lys Lys	115 120 125	
Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Lys Asp	130 135 140	R U
Leu Tyr Gln Leu Leu Ser Asp Arg Ile His Val Leu His Pro Glu Gly	145 150 155 160	
Tyr Leu Ile Thr Pro Ala Trp Leu Trp Glu Lys Tyr Gly Leu Arg Pro	165 170 175	
Asp Gln Trp Ala Asp Tyr Arg Ala Leu Thr Gly Asp Glu Ser Asp Asn	180 185 190	
Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Arg Lys Leu Leu	195 200 205	
Glu Glu Trp Gly Ser Leu Glu Ala Leu Leu Lys Asn Leu Asp Arg Leu	210 215 220	

R U
2 3 5 7 7 3
C 2

? 2 3 5 7 7 3 C 2

R U 2 2 3 5 7 7 C 2	Lys Pro Ala Ile Arg Glu Lys Ile Leu Ala His Met Asp Asp Leu Lys 225 230 235 240 Leu Ser Trp Asp Leu Ala Lys Val Arg Thr Asp Leu Pro Leu Glu Val 245 250 255 Asp Phe Ala Lys Arg Arg Glu Pro Asp Arg Glu Arg Leu Arg Ala Phe 260 265 270 Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly Leu Leu 275 280 285 Glu Ser Pro Lys Ala Leu Glu Glu Ala Pro Trp Pro Pro Pro Glu Gly 290 295 300 Ala Phe Val Gly Phe Val Leu Ser Arg Lys Glu Pro Met Trp Ala Asp 305 310 315 320 Leu Leu Ala Leu Ala Ala Ala Arg Gly Gly Arg Val His Arg Ala Pro 325 330 335 Glu Pro Tyr Lys Ala Leu Arg Asp Leu Lys Glu Ala Arg Gly Leu Leu 340 345 350 Ala Lys Asp Leu Ser Val Leu Ala Leu Arg Glu Gly Leu Gly Leu Pro 355 360 365 Pro Gly Asp Asp Pro Met Leu Leu Ala Tyr Leu Leu Asp Pro Ser Asn 370 375 380 Thr Thr Pro Glu Gly Val Ala Arg Arg Tyr Gly Gly Glu Trp Thr Glu 385 390 395 400 Glu Ala Gly Glu Arg Ala Ala Leu Ser Glu Arg Leu Phe Ala Asn Leu 405 410 415 Trp Gly Arg Leu Glu Gly Glu Glu Arg Leu Leu Trp Leu Tyr Arg Glu 420 425 430 Val Glu Arg Pro Leu Ser Ala Val Leu Ala His Met Glu Ala Thr Gly 435 440 445 Val Arg Leu Asp Val Ala Tyr Leu Arg Ala Leu Ser Leu Glu Val Ala 450 455 460 Glu Glu Ile Ala Arg Leu Glu Ala Glu Val Phe Arg Leu Ala Gly His 465 470 475 480 Pro Phe Asn Leu Asn Ser Arg Asp Gln Leu Glu Arg Val Leu Phe Asp 485 490 495 Glu Leu Gly Leu Pro Ala Ile Gly Lys Thr Glu Lys Thr Gly Lys Arg 500 505 510 Ser Thr Ser Ala Ala Val Leu Glu Ala Leu Arg Glu Ala His Pro Ile 515 520 525	C 2 ? 2 3 5 7 7 3 C 2
------------------------------------------------	------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	--------------------------

R U 2 2 3 5 7 7 C 2	<p>Val Glu Lys Ile Leu Gln Tyr Arg Glu Leu Thr Lys Leu Lys Ser Thr 530 535 540</p> <p>Tyr Ile Asp Pro Leu Pro Asp Leu Ile His Pro Arg Thr Gly Arg Leu 545 550 555 560</p> <p>His Thr Arg Phe Asn Gln Thr Ala Thr Ala Thr Gly Arg Leu Ser Ser 565 570 575</p> <p>Ser Asp Pro Asn Leu Gln Asn Ile Pro Val Arg Thr Pro Leu Gly Gln 580 585 590</p> <p>Arg Ile Arg Arg Ala Phe Ile Ala Glu Glu Gly Trp Leu Leu Val Ala 595 600 605</p> <p>Leu Asp Tyr Ser Gln Ile Gly Leu Arg Val Leu Ala His Leu Ser Gly 610 615 620</p> <p>Asp Glu Asn Leu Ile Arg Val Phe Gln Glu Gly Arg Asp Ile His Thr 625 630 635 640</p> <p>Glu Thr Ala Ser Trp Met Phe Gly Val Pro Arg Glu Ala Val Asp Pro 645 650 655</p> <p>Leu Met Arg Arg Ala Ala Lys Thr Ile Asn Phe Gly Val Leu Tyr Gly 660 665 670</p> <p>Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile Pro Tyr Glu Glu 675 680 685</p> <p>Ala Gln Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys Val Arg 690 695 700</p> <p>Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Arg Arg Gly Tyr Val 705 710 715 720</p> <p>Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Glu Ala Arg 725 730 735</p> <p>Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn Met Pro 740 745 750</p> <p>Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val Lys Leu 755 760 765</p> <p>Phe Pro Arg Leu Glu Glu Met Gly Ala Arg Met Leu Leu Gln Val His 770 775 780</p> <p>Asp Glu Leu Val Leu Glu Ala Pro Lys Glu Arg Ala Glu Ala Val Ala 785 790 795 800</p> <p>Arg Leu Ala Lys Glu Val Met Glu Gly Val Tyr Pro Leu Ala Val Pro 805 810 815</p> <p>Leu Glu Val Glu Val Gly Ile Gly Glu Asp Trp Leu Ser Ala Lys Glu 820 825 830</p>	? 2 3 5 7 7 3 C 2
------------------------------------------------	-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	-------------------