

(19)



Евразийское
патентное
ведомство

(21) 201991116 (13) A1

(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ЕВРАЗИЙСКОЙ ЗАЯВКЕ

(43) Дата публикации заявки
2019.12.30

(22) Дата подачи заявки
2017.12.07

(51) Int. Cl. *A61K 39/12* (2006.01)
A61K 39/235 (2006.01)
C07K 14/075 (2006.01)
C12N 15/861 (2006.01)

(54) ПОЛИНУКЛЕОТИДЫ И ПОЛИПЕПТИДЫ АДЕНОВИРУСА

(31) 1620968.6

(32) 2016.12.09

(33) GB

(86) PCT/IB2017/057738

(87) WO 2018/104911 2018.06.14

(71) Заявитель:
ГЛАКСОСМИТКЛАЙН
БАЙОЛОДЖИКАЛС СА (BE)

(72) Изобретатель:
Аммендола Вирджиния, Капоне
Стефаниа, Коллока Стефано,
Фольгори Антонелла, Мероне Розелла
(IT)

(74) Представитель:
Поликарпов А.В., Соколова М.В.,
Путинцев А.И., Черкас Д.А., Игнатьев
А.В., Билык А.В., Дмитриев А.В. (RU)

(57) Данное изобретение относится к последовательностям выделенных полинуклеотидов и полипептидов, имеющих происхождение из нового аденовируса шимпанзе ChAd157, а также к рекомбинантным полинуклеотидам, векторам, аденовирусам, клеткам и композициям, содержащим указанные последовательности полинуклеотидов и полипептидов.

A1

201991116

201991116

A1

РСТ/ВВ2017/057738

МПК: А61К 39/12 (2006.01)

С07К 14/075 (2006.01)

А61К 39/235 (2006.01)

С12N 15/861 (2006.01)

ПОЛИНУКЛЕОТИДЫ И ПОЛИПЕПТИДЫ АДЕНОВИРУСА**ОБЛАСТЬ ИЗОБРЕТЕНИЯ**

Данное изобретение относится к последовательностям выделенных полинуклеотидов и полипептидов, имеющих происхождение из нового аденовируса шимпанзе ChAd157, а также к рекомбинантным полинуклеотидам, векторам, аденовирусам, клеткам и композициям, содержащим указанные последовательности полинуклеотидов и полипептидов.

ПРЕДШЕСТВУЮЩИЙ УРОВЕНЬ ТЕХНИКИ

Аденовирусы широко применяли для переноса генов благодаря их способности обеспечивать высокоэффективный перенос генов в различных тканях-мишенях и большой трансгенной емкости. Обычно осуществляют делецию генов E1 аденовируса и заменяют их трансгенной кассетой, состоящей из выбранного промотора, последовательности кДНК (комплементарная ДНК) гена, представляющего интерес, и сигнала поли А (полиаденилирования), в результате чего получается дефектный по репликации рекомбинантный аденовирус.

Рекомбинантные аденовирусы полезны в генной терапии и в качестве вакцин. При разработке генетических вакцин вирусные векторы на основе аденовируса шимпанзе являются альтернативой применению векторов на основе аденовирусов, имеющих происхождение из человека. Аденовирусы, выделенные у шимпанзе, близкородственны аденовирусам, выделенным у человека, о чем свидетельствует их эффективное размножение в клетках человеческого происхождения. Однако, поскольку аденовирусы человека и шимпанзе являются близкородственными, между двумя видами вирусов возможна серологическая перекрестная реактивность.

Существует потребность в векторах, которые будут эффективно доставлять молекулы к мишени и минимизировать влияние предсуществующего в популяции иммунитета к отдельным серотипам аденовирусов. Одним аспектом предсуществующего иммунитета, который наблюдается у человека, является гуморальный иммунитет, который может привести к продуцированию и персистенции антител, специфических к аденовирусным белкам. Гуморальный ответ, вызываемый аденовирусом, преимущественно направлен против трех основных структурных белков капсида: фибриллы, пентона и гексона.

СУЩНОСТЬ ИЗОБРЕТЕНИЯ

Предложен выделенный полинуклеотид, который кодирует полипептид, выбранный из группы, состоящей из:

- (а) полипептида, имеющего аминокислотную последовательность SEQ ID NO: 1, и
- (б) функционального производного полипептида, имеющего аминокислотную последовательность SEQ ID NO: 1, где функциональное производное имеет аминокислотную последовательность, которая по меньшей мере на 99,8% идентична на всем своем протяжении аминокислотной последовательности SEQ ID NO: 1.

Также предложен рекомбинантный полинуклеотид, содержащий полинуклеотид, выбранный из группы, состоящей из:

- (а) полинуклеотида, кодирующего полипептид, имеющий аминокислотную последовательность SEQ ID NO: 1, и

(б) полинуклеотида, кодирующего функциональное производное полипептида, имеющего аминокислотную последовательность SEQ ID NO: 1, где функциональное производное имеет аминокислотную последовательность, которая по меньшей мере на 99,8% идентична на всем своем протяжении аминокислотной последовательности SEQ ID NO: 1.

Также предложен рекомбинантный вектор, содержащий полинуклеотид, выбранный из группы, состоящей из:

(а) полинуклеотида, кодирующего полипептид, имеющий аминокислотную последовательность SEQ ID NO: 1, и

(б) полинуклеотида, кодирующего функциональное производное полипептида, имеющего аминокислотную последовательность SEQ ID NO: 1, где функциональное производное имеет аминокислотную последовательность, которая по меньшей мере на 99,8% идентична на всем своем протяжении аминокислотной последовательности SEQ ID NO: 1.

Также предложен рекомбинантный аденовирус, содержащий по меньшей мере один полинуклеотид или полипептид, выбранный из группы, состоящей из:

(а) полинуклеотида, кодирующего полипептид, имеющий аминокислотную последовательность SEQ ID NO: 1;

(б) полинуклеотида, кодирующего функциональное производное полипептида, имеющего аминокислотную последовательность SEQ ID NO: 1, где функциональное производное имеет аминокислотную последовательность, которая по меньшей мере на 99,8% идентична на всем своем протяжении аминокислотной последовательности SEQ ID NO: 1;

(в) полипептида, имеющего аминокислотную последовательность SEQ ID NO: 1, и

(г) функционального производного полипептида, имеющего аминокислотную последовательность SEQ ID NO: 1, где функциональное производное имеет аминокислотную последовательность, которая по меньшей мере на 99,8% идентична на всем своем протяжении аминокислотной последовательности SEQ ID NO: 1.

Также предложена композиция, содержащая по меньшей мере одно из следующего:

(а) выделенный полинуклеотид, кодирующий полипептид, имеющий аминокислотную последовательность SEQ ID NO: 1;

(б) выделенный полинуклеотид, кодирующий функциональное производное полипептида, имеющего аминокислотную последовательность SEQ ID NO: 1, где функциональное производное имеет аминокислотную последовательность, которая по меньшей мере на 99,8% идентична на всем своем протяжении аминокислотной последовательности SEQ ID NO: 1;

(в) выделенный полипептид, имеющий аминокислотную последовательность SEQ ID NO: 1;

(г) выделенное функциональное производное полипептида, имеющего аминокислотную последовательность SEQ ID NO: 1, где функциональное производное имеет аминокислотную последовательность, которая по меньшей мере на 99,8% идентична на всем своем протяжении аминокислотной последовательности SEQ ID NO: 1;

(д) вектор, содержащий полинуклеотид, описанный в (а) или (б) выше, и

(е) рекомбинантный аденовирус, содержащий полинуклеотид, описанный в (а) или (б) выше, и фармацевтически приемлемый эксципиент.

Также предложена клетка, содержащая по меньшей мере одно из следующего:

(а) выделенный полинуклеотид, кодирующий полипептид, имеющий аминокислотную последовательность SEQ ID NO: 1,

(б) выделенный полинуклеотид, кодирующий функциональное производное полипептида, имеющего аминокислотную последовательность SEQ ID NO: 1, где функциональное производное имеет аминокислотную последовательность, которая по меньшей мере на 99,8% идентична на всем своем протяжении аминокислотной последовательности SEQ ID NO: 1;

(в) вектор, содержащий полинуклеотид, описанный в (а) или (б) выше, и

(г) рекомбинантный аденовирус, содержащий полинуклеотид, описанный в (а) или (б) выше.

Также предложен выделенный аденовирусный полипептид, выбранный из группы, состоящей из:

(а) полипептида, имеющего аминокислотную последовательность SEQ ID NO: 1, и

(б) функционального производного полипептида, имеющего аминокислотную последовательность SEQ ID NO: 1, где функциональное производное имеет аминокислотную последовательность, которая по меньшей мере на 99,8% идентична на всем своем протяжении аминокислотной последовательности SEQ ID NO: 1.

КРАТКОЕ ОПИСАНИЕ ГРАФИЧЕСКИХ МАТЕРИАЛОВ

Фиг. 1А-Г - выравнивание последовательностей белка фибриллы указанных аденовирусов обезьян.

ChAd157 (SEQ ID NO:1)

ChAd3 (SEQ ID NO:27)

PanAd3 (SEQ ID NO:28)

ChAd17 (SEQ ID NO:29)

ChAd19 (SEQ ID NO:30)

ChAd24 (SEQ ID NO:31)

ChAd155 (SEQ ID NO:7)

ChAd11 (SEQ ID NO:32)

ChAd20 (SEQ ID NO:33)

ChAd31 (SEQ ID NO:34)

PanAd1 (SEQ ID NO:35)

PanAd2 (SEQ ID NO:36)

Фиг. 2 - схематическое изображение челнока ВАС (искусственной бактериальной хромосомы) подгруппы С.

Фиг. 3 - схематическое изображение челночной плазмиды подгруппы С.

Фиг. 4 - схематическое изображение вектора pChAd157 ΔE1/TetO hCMV GAG

Фиг. 5 - схематическое изображение челнока pARS SpeciesC Ad5orf6-2

Фиг. 6 - схематическое изображение плазмиды, несущей ChAd157 RG

Фиг. 7 - экспрессия трансгена ChAd157/GAG, ChAd19/GAG и ChAd155/GAG

Фиг. 8 - анализ методом вестерн-блоттинг лизатов клеток Hela, инфицированных ChAd155/RG и ChAd157/RG

Фиг. 9 - иммуногенность ChAd157/GAG, ChAd155/GAG и ChAd19 GAG у мышей BALB/c

Фиг. 10 - иммуногенность ChAd157/RG и ChAd155/RG у мышей BALB/c

Фиг. 11 - нейтрализующие титры после предварительной иммунизации мышей различными векторами ChAd

Фиг. 12 - IFN-γ ELISpot после вакцинации мышей ChAd157/GAG после различных схем

предварительной иммунизации

ОПИСАНИЕ ПОСЛЕДОВАТЕЛЬНОСТЕЙ

- SEQ ID NO: 1 - Полипептидная последовательность фибриллы ChAd157
- SEQ ID NO: 2 - Полинуклеотидная последовательность, кодирующая фибриллу ChAd157
- SEQ ID NO: 3 - Полипептидная последовательность пентона ChAd157
- SEQ ID NO: 4 - Полинуклеотидная последовательность, кодирующая пентон ChAd157
- SEQ ID NO: 5 - Полипептидная последовательность гексона ChAd157
- SEQ ID NO: 6 - Полинуклеотидная последовательность, кодирующая гексон ChAd157
- SEQ ID NO: 7 - Полипептидная последовательность фибриллы ChAd155
- SEQ ID NO: 8 - Полинуклеотидная последовательность, кодирующая фибриллу ChAd155
- SEQ ID NO: 9 - Полипептидная последовательность пентона ChAd155
- SEQ ID NO: 10 - Полинуклеотидная последовательность, кодирующая пентон ChAd155
- SEQ ID NO: 11 - Полипептидная последовательность гексона ChAd155
- SEQ ID NO: 12 - Полинуклеотидная последовательность, кодирующая гексон ChAd155
- SEQ ID NO: 13 - Полинуклеотидная последовательность, кодирующая ChAd155 дикого типа
- SEQ ID NO: 14 - Полинуклеотидная последовательность челночной ВАС подгруппы C (#1365)
- SEQ ID NO: 15 - Полинуклеотидная последовательность pChAd157ΔE1 TetO hCMV RpsL-Kana#1551
- SEQ ID NO: 16 - Полинуклеотидная последовательность HIV Gag
- SEQ ID NO: 17 - Полинуклеотидная последовательность pChAd157 ΔE1/TetO hCMV GAG#1557
- SEQ ID NO: 18 - Полинуклеотидная последовательность 1 праймера Ad5orf6
- SEQ ID NO: 19 - Полинуклеотидная последовательность 2 праймера Ad5orf6
- SEQ ID NO: 20 - Полинуклеотидная последовательность 1 праймера Fiber-E4 polyA
- SEQ ID NO: 21 - Полинуклеотидная последовательность 2 праймера Fiber-E4 polyA
- SEQ ID NO: 22 - Полинуклеотидная последовательность ChAd157 ΔE1E4_Ad5E4orf6/TetO hCMV RpsL-Kana#1594
- SEQ ID NO: 23 - Полинуклеотидная последовательность гликопротеина вируса бешенства
- SEQ ID NO: 24 - Полинуклеотидная последовательность pChAd157 ΔE1E4_Ad5E4orf6/TetO hCMV RG#1559
- SEQ ID NO: 25 - Полинуклеотидная последовательность праймера CMVfor
- SEQ ID NO: 26 - Полинуклеотидная последовательность праймера CMVrev
- SEQ ID NO: 27 - Аминокислотная последовательность белка фибриллы ChAd3
- SEQ ID NO: 28 - Аминокислотная последовательность белка фибриллы PanAd3
- SEQ ID NO: 29 - Аминокислотная последовательность белка фибриллы ChAd17
- SEQ ID NO: 30 - Аминокислотная последовательность белка фибриллы ChAd19
- SEQ ID NO: 31 - Аминокислотная последовательность белка фибриллы ChAd24
- SEQ ID NO: 32 - Аминокислотная последовательность белка фибриллы ChAd11
- SEQ ID NO: 33 - Аминокислотная последовательность белка фибриллы ChAd20
- SEQ ID NO: 34 - Аминокислотная последовательность белка фибриллы ChAd31
- SEQ ID NO: 35 - Аминокислотная последовательность белка фибриллы PanAd1
- SEQ ID NO: 36 - Аминокислотная последовательность белка фибриллы PanAd2

SEQ ID NO: 37 - Полинуклеотидная последовательность hCMV(tetO)

SEQ ID NO: 38 - Полинуклеотидная последовательность челночной плазмиды подгруппы С #1376

SEQ ID NO: 39 - Полинуклеотидная последовательность BGH polyA

SEQ ID NO: 40 - Полинуклеотидная последовательность pARS SpeciesC Ad5orf6-2

SEQ ID NO: 41 - Полинуклеотидная последовательность зонда CMVFAM-TAMRA

SEQ ID NO: 42 - Полинуклеотидная последовательность, кодирующая усиленный промотор hCMV

ПОДРОБНОЕ ОПИСАНИЕ ИЗОБРЕТЕНИЯ

Векторы, композиции и способы по данному изобретению могут обладать одной или более улучшенными характеристиками относительно существующего уровня техники, включая более высокую продуктивность, улучшенную иммуногенность, повышенную экспрессию трансгена или определенный профиль серологической перекрестной реактивности, не ограничиваясь ими.

Векторы, композиции и способы по данному изобретению могут демонстрировать комбинацию свойств, таких как продуктивность, иммуногенность, экспрессия трансгена и/или серологическая перекрестная реактивность, что означает, что они представляют важную альтернативу известным подходам.

Аденовирус

Аденовирусы отличаются характерной морфологией с икосаэдрическим капсидом, содержащим три основных белка: гексон (II), основание пентона (III) и булавовидная фибрилла (IV), а также несколько других минорных белков: VI, VIII, IX, IIIa и IVa2. Вирусный геном представлен линейной двуцепочечной ДНК. Вирусная ДНК находится в тесной связи с высокоосновным белком VII и малым пептидом рХ (ранее обозначавшимся мю). Другой белок, V, упакован внутри данного комплекса ДНК-белок и обеспечивает структурную связь с капсидом через белок VI. Вирус также содержит кодируемую вирусом протеазу, которая необходима для процессинга некоторых структурных белков для продуцирования зрелого инфекционного вируса.

Геном аденовируса хорошо охарактеризован. Общая организация генома аденовирусов обычно консервативна в том, что определенные открытые рамки считывания имеют одинаковое расположение, например, расположение генов E1A, E1B, E2A, E2B, E3, E4, L1, L2, L3, L4 и L5 каждого вируса. С каждой стороны на концах аденовирусного генома содержится последовательность, известная как инвертированный концевой повтор (ITR), необходимая для вирусной репликации. Вирус также содержит кодируемую вирусом протеазу, которая необходима для процессинга некоторых структурных белков, требующихся для продуцирования патогенных вирионов. Структуру аденовирусного генома описывают в соответствии с порядком, в котором вирусные гены экспрессируются после трансдукции клеток-хозяев. Более конкретно, вирусные гены обозначают как ранние (E) или поздние (L) гены в зависимости от того, происходит ли их транскрипция до или после начала репликации ДНК. На ранней фазе трансдукции экспрессируются гены аденовируса E1A, E1B, E2A, E2B, E3 и E4, чтобы подготовить клетку-хозяина к репликации вируса. На поздней фазе инфекции активируется экспрессия поздних генов L1-L5, которые кодируют структурные компоненты вирусных частиц.

Аденовирусы являются видоспецифичными, и у разных видов млекопитающих были выделены различные серотипы, то есть типы вируса, которые не подвергаются перекрестной нейтрализации антителами. Например, у человека было выделено более 50 серотипов, которые делят на шесть подгрупп (A–F; B подразделяют на B1 и B2) на основе гомологии последовательностей и их способности вызывать агглютинацию эритроцитов (Tatsis and Ertl Molecular Therapy (2004) 10:616–629). У обезьян, таких как

шимпанзе, бонобо, макаки резус и гориллы, выделили множество аденовирусов, которые классифицировали на такие же группы, как у человека, на основании филогенетических связей, установленных на основании последовательностей гексона или фибриллы (Colloca et al. (2012) *ScienceTranslational Medicine* 4:1-9; Roy et al. (2004) *Virology* 324: 361-372; Roy et al. (2010) *Journal of Gene Medicine* 13:17-25).

В WO2005071093 описаны аденовирусы шимпанзе, включая ChAd19. В WO2016198621 (PCT/EP2016/063329) описаны аденовирусы шимпанзе ChAd155, данный документ включен путем ссылки для определения векторов, имеющих происхождение от ChAd155.

Белки капсида аденовируса, включая белок фибриллу, и полинуклеотиды, кодирующие указанные белки

Как упоминалось выше, капсид аденовируса содержит три основных белка, гексон, пентон и фибриллу. Гексон составляет большинство структурных компонентов капсида, который состоит из 240 тримерных капсомеров типа гексон и 12 оснований пентона. Гексон имеет три консервативных двойных цилиндра, а на его вершине находятся три бугорка, каждый из которых содержит петлю из каждой субъединицы. Гексоны формируют большую часть капсида. Основание гексона у разных серотипов аденовируса является в высокой степени консервативным, тогда как поверхностные петли переменны (Tatsis and Ertl *Molecular Therapy* (2004) 10:616–629).

Пентон представляет собой другой белок капсида аденовируса, который формирует пентамерное основание, к которому присоединяется фибрилла. Тримерный белок фибриллы отходит от основания пентона в каждой из 12 вершин капсида, и имеет булабовидную структуру. Наличие длинного тонкого белка фибриллы является важным отличием поверхности аденовирусных капсидов от большинства других икосаэдрических вирусов. Основная роль белка фибриллы заключается в прикреплении вирусного капсида к клеточной поверхности путем его взаимодействия с клеточным рецептором.

Белки фибриллы аденовирусов многих серотипов имеют схожую архитектуру: N-концевой хвост, центральный стебель, построенный из повторяющихся последовательностей, и C-концевой глобулярный домен, образующий шаровидное утолщение (или «головку»). Центральный стеблевой домен состоит из переменного числа бета-повторов. Бета-повторы соединяются, образуя продолговатую структуру из трех скрученных спиральных нитей, которые являются очень жесткими и стабильными. Стебель соединяет N-концевой хвост с глобулярной шаровидной структурой, которая отвечает за взаимодействие с клеточным рецептором-мишенью. Глобулярная природа шаровидного домена аденовируса обеспечивает большую поверхность для латерального и апикального связывания с рецептором. Благодаря такой архитектуре рецептор-связывающий сайт отстоит далеко от вирусного капсида, таким образом вирус освобождается от стерических препятствий, представленных относительно плоской поверхностью капсида.

Несмотря на сходство общей архитектуры фибрилл аденовирусов многих серотипов, их аминокислотные последовательности переменны, что влияет на их функцию, а также на структуру. Например, ряд областей, экспонированных на поверхности шаровидного утолщения фибриллы, представляют легко адаптирующиеся сайты связывания рецепторов. Глобулярная форма шаровидного утолщения фибриллы позволяет рецепторам связываться по бокам или на вершине шаровидного утолщения фибриллы. Эти сайты связывания обычно располагаются в экспонированных на поверхности петлях, соединяющих бета-нити, консервативность которых среди аденовирусов человека мала. Экспонированные боковые цепи этих петель обеспечивают разнообразие характеристик поверхности шаровидного утолщения при сохранении третичной и четвертичной структуры. Например, электростатический потенциал и

распределения заряда на поверхности шаровидного утолщения могут варьировать благодаря широкому диапазону изоэлектрических точек у последовательностей шаровидного утолщения, от pI приблизительно 9 для Ad 8, Ad 19 и Ad 37 до приблизительно 5 для аденовирусов подгруппы В. Как структурно сложный вирусный лиганд, белок фибриллы позволяет экспонировать разнообразные связывающие поверхности (шаровидное утолщение) в различных ориентациях и на разном удалении (стебель) от вирусного капсида.

Одним из наиболее очевидных различий между некоторыми серотипами является длина фибриллы. Исследования показали, что длина стебля фибриллы сильно влияет на взаимодействие шаровидного утолщения и вируса с его рецепторами-мишенями. Кроме того, серотипы могут различаться по способности белков фибриллы сгибаться. Несмотря на то, что бета-повторы в стебле образуют высоко стабильную и регулярную структуру, исследования с использованием электронной микроскопии (ЕМ) продемонстрировали существование в фибрилле отдельных шарниров. Анализ белковой последовательности нескольких серотипов аденовирусов указывает на нарушение повторяющихся последовательностей стебля в третьем от N-конца бета-повторе, что сильно коррелирует с одним из шарниров в стебле, как обнаружено посредством ЕМ. Шарниры в фибрилле позволяют шаровидному утолщению принимать различные ориентации относительно вирусного капсида, что позволяет преодолевать стерические препятствия для связывания с рецептором, требующего надлежащей презентации рецептор-связывающего сайта на шаровидном утолщении. Например, жестким фибриллам Ad подгруппы D требуется гибкий рецептор или рецептор, заранее позиционированный для присоединения вируса, поскольку они сами не способны сгибаться. (Nicklin et al *Molecular Therapy* 2005 12:384–393).

Применение технологии псевдотипирования фибрилл позволило идентифицировать конкретные клеточные рецепторы для разных серотипов Ad и узнать, как они вносят вклад в тропизм ткани. Несмотря на то, что Ad некоторых подгрупп используют CAR (химерный антигенный рецептор) в качестве первичного рецептора, становится понятным, что многие Ad используют альтернативные первичные рецепторы, что ведет к существенно различному тропизму *in vitro* и *in vivo*. Фибриллы этих серотипов демонстрируют четкие различия в своей первичной и третичной структуре, такие как жесткость стебля фибриллы, длина стебля фибриллы и отсутствие сайта связывания CAR и/или предполагаемого связывающего HSPG (гепаринсульфатпротеогликан) мотива, а также различия в общем заряде в пределах шаровидного утолщения фибриллы. Таким образом, псевдотипирование частиц Ad 5 с заменой стебля фибриллы и шаровидного утолщения позволяет удалить важные связывающиеся с клеткой домены и, помимо этого, может обеспечить более эффективную (и потенциально более селективную в отношении клеток) доставку трансгенов в определенные типы клеток по сравнению с доставкой, обеспечиваемой Ad 5. Нейтрализация частиц Ad с псевдотипированными фибриллами может также быть снижена при использовании фибрилл от Ad с более низким доминированием серотипа у человека или в экспериментальных моделях, что будет способствовать успешному внедрению вектора (Nicklin et al *Molecular Therapy* (2005) 12:384–393). Кроме того, полноразмерная фибрилла, а также выделенные области шаровидного утолщения фибриллы, но не гексон или пентон в отдельности, способны индуцировать созревание дендритных клеток и связаны с индуцированием выраженного ответа CD8⁺ Т-клеток (Molinier-Frenkel et al. *J. Biol. Chem.* (2003) 278:37175–37182). Суммируя вышеизложенное, аденовирусные фибриллы играют важную роль по меньшей мере в связывании с рецептором и иммуногенности аденовирусных векторов.

Выравнивание, приведенное на Фиг. 1, демонстрирует различия между белками фибрилл

аденовирусов обезьян группы С. Поразительно, что последовательности фибрилл этих аденовирусов можно упрощенно разделить на группы с короткой фибриллой, как ChAd157, или с длинной фибриллой, как ChAd155. Это различие в длине обусловлено делецией 36 аминокислот в короткой фибрилле приблизительно в положении 321 по сравнению с длинной фибриллой. Кроме того, существует ряд аминокислотных замен, которые различаются между подгруппами с короткой и длинной фибриллами, но согласуются внутри каждой подгруппы. Хотя точная функция этих различий до сих пор не выяснена, с учетом функции и иммуногенности фибрилл, они могут быть существенными. Было показано, что одной из детерминант вирусного тропизма является длина стебля фибриллы. Было показано, что вектор Ad5 с более коротким стеблем менее эффективно связывается с рецептором CAR и имеет более низкую инфицирующую способность (Ambriović-Ristov A. et al.: Virology. (2003) 312(2):425-33); Полагают, что данное нарушение является результатом повышенной жесткости более короткой фибриллы, ведущей к менее эффективному присоединению к клеточному рецептору (Wu, E et al.: J Virol. (2003) 77(13): 7225–7235).

В одном аспекте изобретения предложен выделенный полипептид фибриллы аденовируса шимпанзе ChAd157 и выделенные полинуклеотиды, кодирующие полипептид фибриллы аденовируса шимпанзе ChAd157.

Ожидают, что белок фибриллы способствует низкому доминированию серотипа и, следовательно, его можно применять независимо от полипептидов гексона и пентона ChAd157 или в комбинации (с гексоном или с пентоном или с обоими) для подавления аффинности аденовируса к предсуществующим нейтрализующим антителам, например, для получения рекомбинантного аденовируса с пониженным доминированием серотипа. Такой рекомбинантный аденовирус может быть химерным аденовирусом, включающим белки капсида от различных серотипов, и по меньшей мере белок фибриллы ChAd157.

Последовательность полипептида фибриллы ChAd157 приведена в SEQ ID NO: 1.

Последовательность полипептида пентона ChAd157 приведена в SEQ ID NO: 3.

Последовательность полипептида гексона ChAd157 приведена в SEQ ID NO: 5.

Полипептиды, рекомбинантные аденовирусы, композиции или клетки, содержащие полипептидные последовательности фибриллы ChAd157 или ее функционального производного

Предпочтительно, выделенный полипептид, рекомбинантный аденовирус, композиция или клетка по изобретению содержит полипептид, имеющий аминокислотную последовательность SEQ ID NO: 1.

Полипептид, рекомбинантный аденовирус, композиция или клетка по изобретению может содержать полипептид, являющийся функциональным производным полипептида, имеющего аминокислотную последовательность SEQ ID NO: 1, где функциональное производное имеет аминокислотную последовательность, которая по меньшей мере на 99,8% идентична на всем своем протяжении аминокислотной последовательности SEQ ID NO: 1.

В альтернативном варианте функциональное производное имеет не более одной вставки, делеции или замены по сравнению с SEQ ID NO: 1, например одну замену по сравнению с SEQ ID NO: 1.

Предпочтительно, полипептид, рекомбинантный аденовирус, композиция или клетка по изобретению дополнительно содержит:

(а) полипептид, имеющий аминокислотную последовательность SEQ ID NO: 3; или

(б) функциональное производное полипептида, имеющего аминокислотную последовательность SEQ ID NO: 3, где функциональное производное имеет аминокислотную последовательность, которая по меньшей мере на 60% идентична на всем своем протяжении аминокислотной последовательности SEQ ID

NO: 3,

и/или

(а) полипептид, имеющий аминокислотную последовательность SEQ ID NO: 5; или

(б) функциональное производное полипептида, имеющего аминокислотную последовательность SEQ ID NO: 5, где функциональное производное имеет аминокислотную последовательность, которая по меньшей мере на 60% идентична на всем своем протяжении аминокислотной последовательности SEQ ID NO: 5.

Предпочтительно, функциональное производное полипептида, имеющего аминокислотную последовательность SEQ ID NO: 3, имеет аминокислотную последовательность, которая по меньшей мере на 70% идентична на всем своем протяжении аминокислотной последовательности SEQ ID NO: 3, например по меньшей мере на 80%, в частности по меньшей мере на 90%, например по меньшей мере на 95% или по меньшей мере на 98%.

Предпочтительно, функциональное производное полипептида, имеющего аминокислотную последовательность SEQ ID NO: 5, имеет аминокислотную последовательность, которая по меньшей мере на 70% идентична на всем своем протяжении аминокислотной последовательности SEQ ID NO: 5, например по меньшей мере на 80%, в частности по меньшей мере на 90%, например по меньшей мере на 95% или по меньшей мере на 98%.

В частности, полипептид, рекомбинантный аденовирус, композиция или клетка по изобретению дополнительно содержит:

(а) полипептид, имеющий аминокислотную последовательность SEQ ID NO: 3;

и/или

(б) полипептид, имеющий аминокислотную последовательность SEQ ID NO: 5.

Выделенные полинуклеотиды, векторы, рекомбинантные аденовирусы, композиции или клетки, содержащие полинуклеотиды, кодирующие фибриллу ChAd157 или ее функциональное производное

Предпочтительно, выделенный полинуклеотид, вектор, рекомбинантный аденовирус, композиция или клетка по изобретению содержит полинуклеотид, который кодирует полипептид, имеющий аминокислотную последовательность SEQ ID NO: 1. Предпочтительно, полинуклеотид имеет последовательность SEQ ID NO: 2.

Когда выделенный полинуклеотид, вектор, рекомбинантный аденовирус, композиция или клетка по изобретению содержит полинуклеотид, который кодирует функциональное производное полипептида, имеющего аминокислотную последовательность SEQ ID NO: 1, где функциональное производное имеет аминокислотную последовательность, которая по меньшей мере на 99,8% идентична на всем своем протяжении аминокислотной последовательности SEQ ID NO: 1, предпочтительно, полинуклеотид имеет последовательность SEQ ID NO: 2, где один кодон был добавлен, удален или изменен для кодирования отличающейся аминокислоты.

Предпочтительно, полинуклеотид, вектор, рекомбинантный аденовирус, композиция или клетка по изобретению дополнительно содержит полинуклеотид, который кодирует:

(а) полипептид, имеющий аминокислотную последовательность SEQ ID NO: 3; или

(б) функциональное производное полипептида, имеющего аминокислотную последовательность SEQ ID NO: 3, где функциональное производное имеет аминокислотную последовательность, которая по меньшей мере на 60% идентична на всем своем протяжении аминокислотной последовательности SEQ ID

NO: 3,

и/или

(а) полипептид, имеющий аминокислотную последовательность SEQ ID NO: 5; или

(б) функциональное производное полипептида, имеющего аминокислотную последовательность SEQ ID NO: 5, где функциональное производное имеет аминокислотную последовательность, которая по меньшей мере на 60% идентична на всем своем протяжении аминокислотной последовательности SEQ ID NO: 5.

Предпочтительно, функциональное производное полипептида, имеющего аминокислотную последовательность SEQ ID NO: 3, имеет аминокислотную последовательность, которая по меньшей мере на 70% идентична на всем своем протяжении аминокислотной последовательности SEQ ID NO: 3, например по меньшей мере на 80%, в частности по меньшей мере на 90%, например по меньшей мере на 95% или по меньшей мере на 98%.

Предпочтительно, функциональное производное полипептида, имеющего аминокислотную последовательность SEQ ID NO: 5, имеет аминокислотную последовательность, которая по меньшей мере на 70% идентична на всем своем протяжении аминокислотной последовательности SEQ ID NO: 5, например по меньшей мере на 80%, в частности по меньшей мере на 90%, например по меньшей мере на 95% или по меньшей мере на 98%.

В частности, полинуклеотид, вектор, рекомбинантный аденовирус, композиция или клетка по изобретению дополнительно содержит полинуклеотид, который кодирует:

(а) полипептид, имеющий аминокислотную последовательность SEQ ID NO: 3;

и/или

(б) полипептид, имеющий аминокислотную последовательность SEQ ID NO: 5.

Полинуклеотид, вектор, рекомбинантный аденовирус, композиция или клетка по изобретению может дополнительно содержать:

(а) полинуклеотид согласно SEQ ID NO: 4;

и/или

(б) полинуклеотид согласно SEQ ID NO: 6.

Остовы ChAd157

В изобретении предложены последовательности выделенных полинуклеотидов аденовируса шимпанзе ChAd157, включая таковые дикого типа, немодифицированного ChAd157 и конструкций ChAd157 с модифицированным остовом. Указанные конструкции с модифицированным остовом включают приведенные в данном документе в качестве примера, такие как pChAd157ΔE1 TetO hCMV RpsL-Kana#1551 (SEQ ID NO: 15) и ChAd157 ΔE1E4_Ad5E4orf6/TetO hCMV RpsL-Kana#1594 (SEQ ID NO: 22). Остовы ChAd157 можно применять в конструировании рекомбинантных репликативно-компетентных или репликативно-некомпетентных аденовирусов, например, для доставки трансгенов.

Аннотирование последовательности pChAd157 ΔE1/TetO hCMV GAG (SEQ ID NO: 17) приведено ниже.

Annotations	ChAd157DE1_TetOhCMV_GAG
IX	3187..3651
IVa2	Complement (3710..5045,5325..5337)
Pol	Complement(4816..8397, 13762..13770)
VA RNAI	10230..10391
pTP	Complement(8196..10199,13762..13770)

48K	10652..11914
pIIIa	11938..13714
III	13807..15588
pVII	15603..16199
V	16275..17390
pX	17415..17660
pVI	17750..18508
Hexon	18623..21499
Protease	21529..22158
DBP	Complement(22274..23926)
92K	23976..26447
22K	26164..26739
33K	Join(26164..26473,26679..27061)
E2e promoter	Complement(27027..27274)
pVIII	27136..27819
E3 12K	27820..28137
E3 CR1-alpha p0	28635..28835
E3 gp18K	28838..29329
E3A 11K	30776..31072
E3 RID alpha	31084..31356
E3 RID beta	31359..31757
E3 15K	31750..32136
U exon	Complement(32167..32331)
fibre	32342..33973
E4 ORF6/7	Complement(34181..34456,35168..35341)
E4 ORF6	Complement(34457..35341)
E4 ORF4	Complement(35241..35606)
E4 ORF3	Complement(35622..35969)
E4 ORF2	Complement(35966..36358)
E4 ORF1	Complement(36411..36797)

В одном воплощении предложены фрагменты последовательностей SEQ ID NO: 15, 22 и их комплементарных цепей, кДНК и РНК, комплементарных им. Предпочтительно, фрагменты имеют длину по меньшей мере 15 нуклеотидов, более предпочтительно длину 30 нуклеотидов, более предпочтительно длину 60 нуклеотидов, более предпочтительно длину 120 нуклеотидов, более предпочтительно длину 240 нуклеотидов, более предпочтительно длину 480 нуклеотидов и охватывают функциональные фрагменты, то есть фрагменты, которые представляют биологический интерес. Например, функциональный фрагмент может экспрессировать желаемый аденовирусный продукт или может быть полезным в производстве рекомбинантных вирусных векторов. Такие фрагменты включают последовательности генов, приведенные выше. В некоторых воплощениях предложены выделенные последовательности SEQ ID NO: 15, 22 и их комплементарных цепей, кДНК и РНК, комплементарных им.

Предложены генные продукты аденовируса ChAd157, такие как белки, ферменты и их фрагменты, которые кодируются аденовирусными нуклеиновыми кислотами, описанными в данном документе, и их вышеупомянутые фрагменты. Такие белки включают белки, кодируемые открытыми рамками считывания, указанными выше, и белки, кодируемые полинуклеотидами, приведенными в перечне последовательностей.

Дополнительные полинуклеотиды и полипептиды ChAd157

В некоторых воплощениях полинуклеотид по изобретению содержит полинуклеотид, кодирующий полипептид фибриллу; полипептид гексон и полипептид фибриллу; полипептид пентон и полипептид фибриллу или полипептид гексон, полипептид пентон и полипептид фибриллу по изобретению; и может дополнительно содержать дополнительные аденовирусные полинуклеотиды, предпочтительно полинуклеотиды ChAd157. Таким образом, предпочтительно, полинуклеотиды по изобретению содержат

одно или более из следующего:

- (а) аденовирусный 5'-инвертированный концевой повтор (ITR);
- (б) аденовирусная область E1A или ее фрагмент, выбранный из областей E1A_280R и E1A_243R;
- (в) аденовирусная область E1B или IX или ее фрагмент, выбранный из группы, состоящей из областей E1B_19K, E1B_55K и IX;
- (г) аденовирусная область E2B или ее фрагмент, выбранный из группы, состоящей из областей E2B_pTP, E2B_полимераза и E2B_IVa2;
- (д) аденовирусная область L1 или ее фрагмент, где указанный фрагмент кодирует аденовирусный белок, выбранный из группы, состоящей из белков L1_13.6K, L1_52K и L1_pIIIa;
- (е) аденовирусная область L2 или область L2, содержащая полинуклеотид, кодирующий белок пентон по изобретению, или ее фрагмент, где указанный фрагмент кодирует аденовирусный белок, выбранный из группы, состоящей из белка L2_пентон, белка L2_pVII, белка L2_V и белка L2_pX;
- (ж) аденовирусная область L3 или область L3, содержащая полинуклеотид, кодирующий белок гексон по изобретению, или ее фрагмент, где указанный фрагмент кодирует аденовирусный белок, выбранный из группы, состоящей из белка L3_pVI, белка L3_гексон и белка L3_протеаза;
- (з) аденовирусная область E2A;
- (и) аденовирусная область L4 или ее фрагмент, где указанный фрагмент кодирует аденовирусный белок, выбранный из группы, состоящей из белка L4_100k, белка L4_33K, белка L4_22K и белка L4_VIII;
- (к) аденовирусная область E3 или ее фрагмент, выбранный из группы, состоящей из: E3 ORF1, E3 ORF2, E3 ORF3, E3 ORF4, E3 ORF5, E3 ORF6, E3 ORF7, E3 ORF8 и E3 ORF9;
- (л) аденовирусная область L5 или область L5, содержащая полинуклеотид, кодирующий полипептид L5_фибрилла по изобретению;
- (м) аденовирусная область E4 (такого как Ad5) или ее фрагмент, выбранный из группы, состоящей из E4 ORF7, E4 ORF6, E4 ORF4, E4 ORF3, E4 ORF2 и E4 ORF1; в частности, ORF6 указанной области E4;
- (н) аденовирусный 3'-ITR и/или
- (о) аденовирусная область VAI или VAII PHK, предпочтительно аденовирусная область VAI или VAII PHK из аденовируса, отличного от ChAd157, более предпочтительно из Ad5.

Определения

Предпочтительно, полинуклеотиды или полипептиды по изобретению являются выделенными. «Выделенный» полинуклеотид это полинуклеотид, который отделен от его исходного окружения. Например, полинуклеотид естественного происхождения является выделенным, если он отделен от всего или части вещества, существующего наряду с ним в естественной системе. Полинуклеотид считают выделенным, например, если он клонирован в вектор, который не является частью его естественного окружения или если он входит в состав кДНК.

Предпочтительно, полинуклеотиды по изобретению являются рекомбинантными. Рекомбинантный означает, что полинуклеотид представляет собой продукт по меньшей мере одной стадии клонирования, рестрикции или лигирования, или других процедур, которые приводят к получению полинуклеотида, отличного от полинуклеотида, встречающегося в природе. Рекомбинантный аденовирус представляет собой аденовирус, содержащий рекомбинантный полинуклеотид. Рекомбинантный вектор представляет собой вектор, содержащий рекомбинантный полинуклеотид. «Рекомбинантный вирус» включает потомков исходного рекомбинантного вируса. «Рекомбинантный вектор» включает репликаты исходного

рекомбинантного вектора. «Рекомбинантный полинуклеотид» включает репликаты исходного рекомбинантного полинуклеотида.

Предпочтительно, полипептидная последовательность по данному изобретению содержит по меньшей мере одну модификацию относительно нативной последовательности. Предпочтительно, последовательности полинуклеотидов по данному изобретению содержат по меньшей мере одну модификацию относительно нативной последовательности. Например, полинуклеотид, внедренный генноинженерными способами в плазмиду или вектор, имеющий происхождение из другого вида (и зачастую, другого рода, подсемейства или семейства), представляет собой гетерологичный полинуклеотид. Промотор, отделенный от его нативной кодирующей последовательности и функционально связанный с кодирующей последовательностью, с которой он не связан в природе, представляет собой гетерологичный промотор. Специфический сайт рекомбинации, клонированный в геном вируса или в вирусный вектор, где геном вируса не содержит его в естественных условиях, представляет собой гетерологичный сайт рекомбинации. Гетерологичная нуклеиновокислотная последовательность также включает последовательность, встречающуюся в естественных условиях в аденовирусном геноме, но расположенную в положении аденовирусного вектора, не являющемся нативным.

Обычно «гетерологичный» означает полученный из субъекта, генотипически отличного от того субъекта, с которым проводят сравнение. Гетерологичная нуклеотидная последовательность относится к любой нуклеотидной последовательности, которая не является выделенной из, имеющей происхождение из, или основанной на нуклеотидной последовательности аденовирусного вектора, имеющей естественное происхождение. Гетерологичная белковая последовательность относится к любой последовательности белка, которая не является выделенной из, имеющей происхождение из, или основанной на последовательности белка аденовирусного вектора естественного происхождения. «Естественного происхождения» означает, что последовательность встречается в природе и не получена путем синтеза и не модифицирована. Последовательность «имеет происхождение из» источника, когда она выделена из источника, но модифицирована (например, посредством делеции, замены (мутации), вставки или другой модификации), предпочтительно, таким образом, чтобы не нарушать нормальной функции гена-источника.

«Функциональное производное» или полипептид предпочтительно относится к модифицированному варианту полипептида, например, в котором осуществили делецию, вставку, модификацию и/или замену одной или более аминокислот полипептида. Производное немодифицированного белка капсида аденовируса считают функциональным, если, например:

(а) аденовирус, содержащий в составе своего капсида производное белка капсида, сохраняет по существу такое же или имеет более низкое доминирование серотипа по сравнению с аденовирусом, содержащим немодифицированный белок капсида, и/или

(б) аденовирус, содержащий в составе своего капсида производное белка капсида, сохраняет по существу такую же или обладает более высокой способностью инфицировать клетку-хозяина по сравнению с аденовирусом, содержащим немодифицированный белок капсида, и/или

(в) аденовирус, содержащий в составе своего капсида производное белка капсида, сохраняет по существу такую же или имеет более высокую иммуногенность по сравнению с аденовирусом, содержащим немодифицированный белок капсида, и/или

(г) аденовирус, содержащий в составе своего капсида производное белка капсида, сохраняет по существу такую же или обеспечивает более высокую продуктивность в отношении трансгена по сравнению

с аденовирусом, содержащим немодифицированный белок капсида.

Свойства (а)-(г), перечисленные выше, предпочтительно можно оценивать способами, описанными в разделе Примеры ниже.

Предпочтительно, полипептид, вектор или рекомбинантный аденовирус имеет низкое доминирование серотипа в человеческой популяции. «Низкое доминирование серотипа» может означать пониженный уровень предсуществующих нейтрализующих антител по сравнению с человеческим аденовирусом 5 (Ad5). Аналогично или в качестве альтернативы, «низкое доминирование серотипа» может означать доминирование серотипа менее чем приблизительно 20%, доминирование серотипа менее чем приблизительно 15%, доминирование серотипа менее чем приблизительно 10%, доминирование серотипа менее чем приблизительно 5%, доминирование серотипа менее чем приблизительно 4%, доминирование серотипа менее чем приблизительно 3%, доминирование серотипа менее чем приблизительно 2%, доминирование серотипа менее чем приблизительно 1% или отсутствие детектируемого доминирования серотипа. Доминирование серотипа можно определять, как выраженную в процентах долю индивидуумов, имеющих клинически значимый нейтрализующий титр (который можно охарактеризовать как 50% нейтрализующий титр более 200), используя способы, описанные Aste-Amézaga et al., Hum. Gene Ther. (2004) 15(3):293-304.

Термины полипептид, пептид и белок в данном документе используют взаимозаменяемо.

Термин «обезьяний» обычно охватывает приматов, не являющихся человеком, например обезьян Старого Света, обезьян Нового Света, больших человекообразных обезьян и гиббонов. В частности, обезьяний может относиться к большим человекообразным обезьянам, таким как шимпанзе (*Pan troglodyte*), бонобо (*Pan paniscus*) и гориллы (род *Gorilla*). Нечеловекообразные обезьяны могут включать макака резус (*Macaca mulatta*).

Сравнение последовательностей

Для сравнения двух близкородственных полинуклеотидных или полипептидных последовательностей можно рассчитывать «% идентичности» между первой последовательностью и второй последовательностью с помощью программ выравнивания, таких как BLAST® (доступна по адресу blast.ncbi.nlm.nih.gov, последний доступ 09 марта 2015) с использованием стандартных настроек. % идентичности представляет собой количество идентичных остатков, поделенное на количество остатков эталонной последовательности, умноженное на 100. Значения % идентичности, указанные выше и в формуле изобретения, представляют собой процент, рассчитанный данным способом. Согласно альтернативному определению, % идентичности представляет собой количество идентичных остатков, поделенное на количество выровненных остатков, умноженное на 100. Альтернативные способы включают применение способа с пробелами, в котором пробелы в выравнивании, например делеции в одной последовательности относительно другой последовательности, при расчете веса выравнивания учитывают, присваивая им определенный вес или цену. Для более подробной информации см. обзор по BLAST® доступный по адресу ftp.ncbi.nlm.nih.gov/pub/factsheets/HowTo_BLASTGuide.pdf, последний доступ 09 марта 2015.

Полагают, что последовательности, сохраняющие функциональность полинуклеотида или кодируемого им полипептида, имеют большую идентичность. Полипептидные или полинуклеотидные последовательности называют одинаковыми или идентичными другим полипептидным или полинуклеотидным последовательностям, если они демонстрируют 100% идентичность

последовательностей на всем своем протяжении.

«Различие» между последовательностями относится к вставке, делеции или замене одного аминокислотного остатка в положении второй последовательности относительно первой последовательности. Между двумя полипептидными последовательностями может существовать одно, два или более таких аминокислотных различий. Вставки, делеции или замены во второй последовательности, которая в остальном идентична (100% идентичность последовательностей) первой последовательности, приводят к снижению процента идентичности последовательностей. Например, если идентичные последовательности имеют длину 9 аминокислотных остатков, одна замена во второй последовательности приводит к 88,9% идентичности последовательности. Если идентичные последовательности имеют длину 17 аминокислотных остатков, две замены во второй последовательности приводят к 88,2% идентичности последовательности. Если идентичные последовательности имеют длину 7 аминокислотных остатков, три замены во второй последовательности приводят к 57,1% идентичности последовательности. Если первая и вторая полипептидные последовательности имеют длину 9 аминокислотных остатков и 6 идентичных остатков, первая и вторая полипептидные последовательности имеют идентичность более 66% (первая и вторая полипептидные последовательности имеют идентичность 66,7%). Если первая и вторая полипептидные последовательности имеют длину 17 аминокислотных остатков и 16 идентичных остатков, первая и вторая полипептидные последовательности имеют идентичность более 94% (первая и вторая полипептидные последовательности имеют идентичность 94,1%). Если первая и вторая полипептидные последовательности имеют длину 7 аминокислотных остатков и 3 идентичных остатка, первая и вторая полипептидные последовательности имеют идентичность более 42% (первая и вторая полипептидные последовательности имеют идентичность 42,9%).

Альтернативно, для сравнения первой, эталонной полипептидной последовательности со второй, сравниваемой полипептидной последовательностью, может быть установлено количество добавлений, замен и/или делеций, сделанных в первой последовательности для получения второй последовательности. Добавление представляет собой добавление одного аминокислотного остатка в последовательность первого полипептида (включая добавления на любом конце первого полипептида). Замена представляет собой замену одного аминокислотного остатка в последовательности первого полипептида одним отличным аминокислотным остатком. Делеция представляет собой удаление одного аминокислотного остатка из последовательности первого полипептида (включая делецию на любом конце первого полипептида).

Для сравнения первой, эталонной полинуклеотидной последовательности со второй, сравниваемой полинуклеотидной последовательностью, может быть установлено количество добавлений, замен и/или делеций, сделанных в первой последовательности для получения второй последовательности. Добавление представляет собой добавление одного нуклеотидного остатка в последовательность первого полинуклеотида (включая добавление на любом конце первого полинуклеотида). Замена представляет собой замену одного нуклеотидного остатка в последовательности первого полинуклеотида одним отличным нуклеотидным остатком. Делеция представляет собой удаление одного нуклеотидного остатка из последовательности первого полинуклеотида (включая делецию на любом конце первого полинуклеотида).

Предпочтительно, замены в последовательностях по настоящему изобретению могут представлять собой консервативные замены. Консервативная замена включает замену аминокислоты другой аминокислотой, имеющей химические свойства, схожие с замещаемой аминокислотой (см., например, Stryer et al, Biochemistry, 5th Edition 2002, стр. 44-49). Предпочтительно, консервативная замена представляет

собой замену, выбранную из группы, состоящей из: (1) замены основной аминокислоты другой, отличной основной аминокислотой; (2) замены кислой аминокислоты другой, отличной кислой аминокислотой; (3) замены ароматической аминокислоты другой, отличной ароматической аминокислотой; (4) замены неполярной алифатической аминокислоты другой, отличной неполярной алифатической аминокислотой и (5) замены полярной незаряженной аминокислоты другой, отличной полярной незаряженной аминокислотой. Основная аминокислота предпочтительно выбрана из группы, состоящей из аргинина, гистидина и лизина. Кислая аминокислота предпочтительно представляет собой аспартат или глутамат. Ароматическая аминокислота предпочтительно выбрана из группы, состоящей из фенилаланина, тирозина и триптофана. Неполярная алифатическая аминокислота предпочтительно выбрана из группы, состоящей из глицина, аланина, валина, лейцина, метионина и изолейцина. Полярная незаряженная аминокислота предпочтительно выбрана из группы, состоящей из серина, треонина, цистеина, пролина, аспарагина и глутамина. В отличие от консервативной аминокислотной замены, неконсервативная аминокислотная замена представляет собой замену одной аминокислоты любой аминокислотой, которая не подпадает под перечисленные выше консервативные замены (1) - (5).

Векторы и рекомбинантный аденовирус

Последовательности ChAd155 по изобретению полезны в качестве терапевтических агентов и в конструировании различных систем векторов, рекомбинантного аденовируса и клеток-хозяев. Предпочтительно, термин «вектор» относится к нуклеиновой кислоте, которая была существенно изменена (например, произведена делеция и/или инактивация гена или функциональной области) относительно последовательности дикого типа и/или включает гетерологичную последовательность, то есть нуклеиновую кислоту, полученную из другого источника (также называемую «вставкой»), и реплицирующая и/или экспрессирующая вставленную полинуклеотидную последовательность при введении в клетку (например, клетку-хозяина). Например, вставка может представлять собой всю или часть последовательностей ChAd157, описанных в данном документе. Дополнительно или альтернативно, вектор ChAd157 может представлять собой аденовирус ChAd157, включающий одну или более делеций или инактиваций вирусных генов, таких как E1 или другой вирусный ген или функциональная область, описанные в данном документе. Такой ChAd157, который может содержать или не содержать гетерологичную последовательность, часто обозначается «остовом» и может быть использован в существующем виде или в виде отправной точки для дополнительных модификаций вектора.

Вектор может представлять собой любую подходящую молекулу нуклеиновой кислоты, включая голую ДНК, плазмиду, вирус, космиду, фаговый вектор, такой как вектор лямбда, искусственную хромосому, такую как BAC (искусственная бактериальная хромосома) или эписому. Альтернативно, вектор может представлять собой транскрипционную и/или экспрессионную единицу для бесклеточной транскрипции или экспрессии *in vitro*, такую как T7-совместимая система. Векторы можно применять отдельно или в комбинации с другими аденовирусными последовательностями или фрагментами, или в комбинации с элементами из неаденовирусных последовательностей. Последовательности ChAd157 также полезны в бессмысловых векторах для доставки, векторах для генной терапии или векторных вакцинах. Таким образом, дополнительно предложены векторы для доставки генов и клетки-хозяева, которые содержат последовательности ChAd157.

Термин «репликативно-компетентный» аденовирус относится к аденовирусу, который может реплицироваться в клетке-хозяине в отсутствие каких-либо рекомбинантных хелперных белков,

содержащихся в клетке. Предпочтительно, «репликативно-компетентный» аденовирус содержит следующие интактные или функциональные существенные ранние гены: E1A, E1B, E2A, E2B, E3 и E4. Аденовирусы дикого типа, выделенные у конкретного животного, будут репликативно-компетентными у этого животного.

Термин «репликативно-некомпетентный» или «дефектный по репликации» аденовирус относится к аденовирусу, который не способен к репликации, поскольку он был сконструирован так, чтобы иметь по меньшей мере функциональную делецию (или мутацию с «потерей функции»), то есть делецию или мутацию, которая нарушает функцию гена без полного его удаления, например введение искусственных стоп-кодонов, делецию или мутацию активных сайтов или доменов взаимодействия, мутацию или делецию регуляторной последовательности гена и т.д., или полное удаление гена, кодирующего генный продукт, необходимый для вирусной репликации, такого как один или более из аденовирусных генов, выбранных из E1A, E1B, E2A, E2B, E3 и E4 (таких как E3 ORF1, E3 ORF2, E3 ORF3, E3 ORF4, E3 ORF5, E3 ORF6, E3 ORF7, E3 ORF8, E3 ORF9, E4 ORF7, E4 ORF6, E4 ORF4, E4 ORF3, E4 ORF2 и/или E4 ORF1). Наиболее предпочтительно, осуществлена делеция E1 и, возможно E3 и/или E4. При делеции, указанная удаленная область гена предпочтительно не будет учитываться при выравнивании при определении % идентичности с другой последовательностью.

В данном изобретении предложены векторы, такие как рекомбинантный аденовирус, которые доставляют белок, предпочтительно гетерологичный белок, в клетки, либо в целях терапии, либо в целях вакцинации. Вектор может включать любой генетический элемент, включая голую ДНК, фаг, транспозон, космиду, эписому, плазмиду или вирус. Такие векторы содержат ДНК ChAd157, согласно данному описанию, и миниген. Под «минигеном» (или «экспрессионной кассетой») понимают комбинацию выбранного гетерологичного гена (трансгена) и другие регуляторные элементы, необходимые для осуществления трансляции, транскрипции и/или экспрессии продукта гена в клетке-хозяине.

Обычно аденовирусный вектор на основе ChAd157 создают таким образом, чтобы миниген располагался в молекуле нуклеиновой кислоты, содержащей другие аденовирусные последовательности в нативной для выбранного гена аденовируса области. При необходимости миниген может быть встроен в область существующего гена для нарушения функции данной области. В альтернативном варианте миниген может быть встроен в сайт аденовирусного гена, подвергшегося частичной или полной делеции. Например, миниген может быть расположен в сайте мутации, вставки или делеции, которая делает нефункциональным по меньшей мере один ген в области генома, выбранной из группы, состоящей из E1A, E1B, E2A, E2B, E3 и E4. Термин «делает нефункциональным» означает, что удален или иным образом разрушен достаточный участок гена, так что ген больше не способен продуцировать функциональные продукты экспрессии гена. При желании можно удалять весь участок гена целиком (и предпочтительно заменять минигеном).

Например, для продуцирующего вектора, полезного в создании рекомбинантного вируса, вектор может содержать миниген и либо 5' конец аденовирусного генома, либо 3' конец аденовирусного генома, либо как 5', так и 3' концы аденовирусного генома. 5' конец аденовирусного генома содержит 5' цис-элементы, необходимые для упаковки и репликации; то есть, 5' ITR последовательности (которые функционируют как точки начала репликации) и нативные 5' домены упаковки и энхансера (которые содержат последовательности, необходимые для упаковки линейных геномов Ad и энхансерные элементы для промотора E1). 3' конец аденовирусного генома содержит 3' цис-элементы (включая ITR), необходимые для упаковки и заключения в капсид. Предпочтительно, рекомбинантный аденовирус содержит и 5', и 3' аденовирусные цис-элементы, и миниген (предпочтительно содержащий трансген) располагается между 5' и

3' аденовирусными последовательностями. Аденовирусный вектор на основе ChAd157 может также содержать дополнительные аденовирусные последовательности.

Предпочтительно, векторы на основе ChAd157 содержат один или более аденовирусных элементов, имеющих происхождение из аденовирусного генома ChAd157 по изобретению. В одном воплощении векторы содержат аденовирусные ITR из ChAd157 и дополнительные аденовирусные последовательности из того же аденовирусного серотипа. В другом воплощении векторы содержат аденовирусные последовательности, которые имеют происхождение из другого аденовирусного серотипа, чем тот, который обеспечивает ITR.

Согласно данному описанию, псевдотипированный аденовирус относится к аденовирусу, у которого белки капсида аденовируса имеют происхождение из другого аденовируса, чем аденовирус, который обеспечивает ITR.

Кроме того, можно сконструировать химерные или гибридные аденовирусы, используя аденовирусы, описанные в данном документе, с применением способов, известных специалистам в области техники (например, US 7291498).

ITR и любые другие аденовирусные последовательности, присутствующие в векторе по настоящему изобретению, могут быть получены из многих источников. Разнообразные штаммы аденовирусов предоставляются Американской коллекцией типовых культур, Манассас, Виргиния, или, по запросу, различными коммерческими и академическими организациями. Кроме того, доступ к последовательностям многих таких штаммов предоставляется различными базами данных, включая, например, PubMed и GenBank. В опубликованной литературе описаны гомологичные аденовирусные векторы, полученные из других аденовирусов шимпанзе или человека (например, US 5240846). Доступ к последовательностям ДНК аденовирусов нескольких типов предоставляется GenBank, включая тип Ad5 (номер доступа GenBank M73370). Могут быть получены последовательности аденовирусов любых известных серотипов аденовирусов, таких как 2, 3, 4, 7, 12 и 40, также включая любые из идентифицированных к настоящему времени типов аденовирусов человека. Аналогично, в конструкциях векторов по данному изобретению можно также применять аденовирусы, которые, как известно, инфицируют животных, не являющихся человеком (например, обезьян) (например, US 6083716). Вирусные последовательности, вирусы-хелперы (при необходимости) и рекомбинантные вирусные частицы, а также другие компоненты векторов и последовательности, использованные при конструировании векторов, описанных в данном документе, могут быть получены, как описано ниже.

Получение последовательности, вектора и аденовируса

Последовательности по изобретению могут быть получены любым подходящим способом, включая рекомбинантное получение, химический синтез или другие способы синтеза. Подходящие методики получения хорошо известны специалистам в области техники. Альтернативно, пептиды также можно синтезировать хорошо известными способами твердофазного синтеза пептидов.

Для получения аденовирусных векторов можно использовать аденовирусные плазмиды (или другие векторы). В одном воплощении аденовирусные векторы представляют собой частицы аденовируса, которые являются репликативно-некомпетентными. В одном воплощении частицы аденовируса становятся репликативно-некомпетентными в результате делеций в генах E1A и/или E1B, в частности, E1A и E1B. В альтернативном варианте аденовирусы делают репликативно-некомпетентными другими способами, возможно при сохранении генов E1A и/или E1B. Аналогично, в некоторых воплощениях снижение

иммунного ответа на вектор может достигаться в результате делеций в генах E2B и/или ДНК-полимеразы. Аденовирусные векторы также могут содержать другие мутации аденовирусного генома, например чувствительные к температуре мутации или делеции в других генах. В других воплощениях желательно сохранять интактные области E1A и/или E1B в аденовирусных векторах. Так, интактная область E1 может располагаться в аденовирусном геноме в нативном положении или помещаться в сайт, где произведена делеция нативного аденовирусного генома (например, в область E3).

При конструировании аденовирусных векторов для доставки гена в клетку млекопитающего (такого как человек) в векторах можно использовать ряд модифицированных нуклеотидных последовательностей аденовируса. Например, из аденовирусной последовательности, входящей в состав рекомбинантного вируса, можно удалить весь задержанный ранний ген E3 аденовируса или его часть. Полагают, что функция E3 не важна для функции и продукции частицы рекомбинантного вируса. Аденовирусные векторы могут также быть сконструированы с делециями по меньшей мере области ORF6 гена E4, и более предпочтительно всей области E4 из-за избыточности функции указанной области. Еще один вектор по изобретению содержит делецию в задержанном раннем гене E2A. Можно также осуществлять делеции в любом из поздних генов L1 - L5 генома аденовируса. Аналогично, для некоторых задач могут быть полезны делеции в промежуточных генах IX и IVa2. Другие делеции могут быть осуществлены в других структурных или неструктурных генах аденовируса. Указанные выше делеции можно использовать по отдельности, то есть аденовирусная последовательность для применения, как описано здесь, может содержать делеции только в одной области. Альтернативно, делеции целых генов или их частей, эффективно разрушающие биологическую активность, можно использовать в любой комбинации. Например, в одном приведенном в качестве примера векторе последовательность аденовируса может иметь делеции генов E1 и гена E4, или генов E1, E2A и E3, или генов E1 и E3, или генов E1, E2A и E4, с делецией или без делеции E3 и т.д. Любой один или более E генов предпочтительно могут быть заменены E геном (или одной или более открытыми рамками считывания E гена), источником которых является аденовирус другого штамма. Наиболее предпочтительно, осуществляют делецию генов ChAd157 E1 и E3 и ген ChAd157E4 заменяют на E4Ad5orf6. Как обсуждалось выше, для достижения желаемого результата такие делеции и/или замены можно использовать в комбинации с другими мутациями, такими как чувствительные к температуре мутации.

Аденовирусный вектор, лишенный одной или более важных аденовирусных последовательностей (например, E1A, E1B, E2A, E2B, E4 ORF6, L1, L2, L3, L4 и L5), можно культивировать в присутствии продуктов отсутствующих аденовирусных генов, необходимых для инфицирующей способности вируса и размножения частиц аденовируса. Указанные хелперные функции могут обеспечиваться культивированием аденовирусного вектора в присутствии одной или более хелперных конструкций (например, плазмиды или вируса) или пакующей клетки-хозяина.

Комплементация репликативно-некомпетентных векторов

Для создания рекомбинантных аденовирусов с делециями в любом из генов согласно описанию выше, функцию области гена рекомбинантного вируса, делецию которой произвели, если она существенна для репликации и инфицирующей способности вируса, необходимо восполнить с использованием хелперного вируса или клеточной линии, то есть комплементирующей или пакующей клеточной линии.

Хелперные вирусы

В зависимости от генов аденовируса, содержащихся в вирусных векторах, использованных как носителей минигена, можно применять хелперный аденовирус или нереплицирующийся фрагмент вируса,

чтобы обеспечить достаточно последовательностей гена аденовируса, необходимых для продукции инфекционной частицы рекомбинантного вируса, содержащего миниген. Пригодные хелперные вирусы содержат выбранные последовательности генов аденовируса, отсутствующие в конструкции аденовирусного вектора и/или не экспрессирующиеся пакующей клеточной линией, трансфицированной вектором. В одном воплощении хелперный вирус является дефектным по репликации и содержит гены аденовируса в дополнение, предпочтительно, к одной или более последовательностям по данному описанию. Такой хелперный вирус предпочтительно применяют в комбинации с клеточной линией, экспрессирующей E1 (и возможно дополнительно экспрессирующей E3).

Хелперный вирус может возможно содержать репортерный ген. Ряд таких репортерных генов известен в области техники, а также описан в данном документе. Присутствие репортерного гена в хелперном вирусе, отличающегося от трансгена в аденовирусном векторе, позволяет независимо отслеживать как аденовирусный вектор, так и хелперный вирус. Этот репортер применяют для возможности разделения полученного рекомбинантного вируса и хелперного вируса при очистке.

Линии комплементирующих клеток

Во многих обстоятельствах для транскомплементации вектора на основе аденовируса шимпанзе можно использовать линию клеток, экспрессирующих один или более отсутствующих генов, существенных для репликации и инфекционной способности вируса, таких как человеческий E1. Это особенно выгодно, поскольку из-за различий между последовательностями аденовируса шимпанзе по изобретению и последовательностями аденовируса человека, обнаруживаемыми в существующих на данный момент пакующих клетках, применение существующих клеток человека, содержащих E1, предупреждает появление репликативно-компетентных аденовирусов в ходе процессов репликации и продукции.

В альтернативном варианте, при желании можно использовать последовательности, приведенные в данном документе, для создания пакующей клетки или линии клеток, экспрессирующей, как минимум, ген E1 из ChAd157 или из другого аденовируса (например, из аденовируса человека, например, E1 из hAd5, или E1 из другого ChAd (аденовируса шимпанзе)) под транскрипционным контролем промотора для экспрессии в выбранной родительской клеточной линии. Для этой задачи можно использовать индуцибельные или конститутивные промоторы. Примеры таких промоторов подробно описаны в других разделах данного документа. Родительскую клетку выбирают для создания новой клеточной линии, экспрессирующей любой желаемый ген ChAd157. Без ограничения, такая родительская клеточная линия может представлять собой клетки HeLa [номер доступа ATCC CCL 2], A549 [номер доступа ATCC CCL 185], HEK 293, KB [CCL 17], Detroit [например, Detroit 510, CCL 72] и WI-38 [CCL 75], помимо прочих. Данные клеточные линии можно получить из Американской коллекции типовых культур, 10801 University Boulevard, Manassas, Virginia 20110-2209.

Такие экспрессирующие E1 клеточные являются полезными в создании рекомбинантных аденовирусных векторов с делецией E1. Дополнительно или альтернативно, клеточные линии, которые экспрессируют продукты одного или более генов аденовируса, например E1A, E1B, E2A, E3 и/или E4, могут быть сконструированы с применением по существу тех же процедур, которые использовали при создании рекомбинантных вирусных векторов. Такие клеточные линии можно применять для транскомплементации аденовирусных векторов с делецией принципиально значимых генов, кодирующих указанные продукты, или для обеспечения хелперных функций, необходимых для упаковки хелпер-зависимого вируса (например, аденоассоциированного вируса). Получение клетки-хозяина включает такие методики, как сборка

отобранных последовательностей ДНК.

В другом альтернативном варианте источниками продуктов принципиально значимых генов аденовируса *in trans* являются аденовирусный вектор и/или вирус-хелпер. В этом случае подходящая клетка-хозяин может быть выбрана из любого биологического организма, включая прокариотические (например, бактериальные) клетки и эукариотические клетки, включая клетки насекомых, клетки дрожжей и клетки млекопитающих.

Клетки-хозяева могут быть выбраны из клеток любых видов млекопитающих, включая, без ограничения, такие клетки, как клетки A549, WENI, 3T3, 10T1/2, HEK 293 или Per.C6 (обе из которых экспрессируют функциональный аденовирусный E1) [Fallaux, FJ et al, (1998), Hum Gene Ther, 9:1909-1917], Saos, C2C12, L клетки, HT1080, HerG2 и первичные фибробласты, гепатоциты и миобласты, полученные у млекопитающих, включая человека, обезьяну, мышь, крысу, кролика и хомяка.

Особенно подходящей линией комплементирующих клеток является линия клеток Procell92. Линия клеток Procell92 получена на основе клеток HEK 293, экспрессирующих гены E1 аденовируса, трансфицированных Tet репрессором под контролем промотора фосфоглицераткиназы-1 (PGK) человека и гена резистентности к G418 (Vitelli et al. PLOS One (2013) 8(e55435):1-9). Procell92.S адаптирована для роста в суспензии и является полезной в получении аденовирусных векторов, экспрессирующих токсические белки (www.okairos.com/e/innerns.php?m=00084, последний доступ 13 апреля 2015).

Сборка вирусной частицы и трансфекция клеточной линии

Как правило, при доставке вектора, содержащего миниген, путем трансфекции, вектор доставляют в количестве от приблизительно 5 мкг до приблизительно 100 мкг ДНК, и предпочтительно от приблизительно 10 до приблизительно 50 мкг ДНК в от приблизительно 1×10^4 клеток до приблизительно 1×10^{13} клеток, и предпочтительно в приблизительно 10^5 клеток. Однако относительные количества ДНК вектора в клетках-хозяевах можно корректировать с учетом таких факторов, как выбранный вектор, способ доставки и выбранные клетки-хозяева.

Внедрение вектора в клетку-хозяина можно осуществлять любым способом, известным в области техники, включая трансфекцию и инфицирование. Один или более аденовирусных генов могут быть стабильно интегрированы в геном клетки-хозяина, экспрессироваться стабильно в виде эписом или экспрессироваться транзитивно. Все продукты генов могут экспрессироваться транзитивно, в эписоме или стабильно интегрированном виде, или некоторые из продуктов генов могут экспрессироваться стабильно, тогда как другие экспрессируются транзитивно.

Внедрение векторов в клетку-хозяина может также достигаться с использованием методик, известных специалистам. Предпочтительно, применяют стандартные способы трансфекции, например трансфекцию CaPC или электропорацию.

Сборку отобранных последовательностей ДНК аденовируса (а также трансгена и других элементов вектора) в различные промежуточные плазмиды и использование плазмид и векторов для получения рекомбинантной вирусной частицы осуществляют с использованием стандартных методик. Такие методики включают стандартные методики клонирования кДНК, применение перекрывающихся олигонуклеотидных последовательностей геномов аденовирусов, полимеразную цепную реакцию и любые подходящие способы, которые позволяют получить желаемую нуклеотидную последовательность. Используют стандартные методики трансфекции и ко-трансфекции, например методики преципитации CaPC. Другие применяемые стандартные способы включают гомологичную рекомбинацию вирусных геномов, бляшкообразование

вирусов в слое агара, способы измерения формируемого сигнала и т.п.

Например, после конструирования и сборки желаемого вирусного вектора, содержащего миниген, вектором трансфицируют *in vitro* пакующую клеточную линию в присутствии хелперного вируса. Происходит гомологичная рекомбинация между хелперными и векторными последовательностями, что позволяет последовательностям аденовируса-трансгена в векторе реплицироваться и упаковываться в капсиды вириона, в результате чего образуются частицы рекомбинантного вирусного вектора. Полученные рекомбинантные аденовирусы являются полезными в переносе выбранного трансгена в выбранную клетку. В экспериментах *in vivo* с рекомбинантным вирусом, выращиваемым в пакующих клеточных линиях, рекомбинантные аденовирусные векторы с делецией E1 по изобретению демонстрируют пригодность для переноса трансгена в клетку млекопитающего, не являющегося обезьяной, предпочтительно, человека.

Трансгены

Трансген представляет собой нуклеиновокислотную последовательность, гетерологичную по отношению к последовательностям вектора, фланкирующим трансген, кодирующую белок, представляющий интерес. Кодирующая нуклеиновокислотная последовательность функционально связана с регуляторными компонентами таким образом, который позволяет трансгену транскрибироваться, транслироваться и/или экспрессироваться в клетке-хозяине.

Состав последовательности трансгена будет зависеть от применения полученного вектора. Например, трансген может быть терапевтическим трансгеном или иммуногенным трансгеном. Альтернативно, последовательность трансгена может включать репортерную последовательность, которая при экспрессии производит выявляемый сигнал. Такие репортерные последовательности включают, без ограничения, последовательности ДНК, кодирующие β -лактамазу, β -галактозидазу (LacZ), щелочную фосфатазу, тимидин-киназу, зеленый флуоресцентный белок (GFP), хлорамфениколацетилтрансферазу (CAT), люциферазу, мембраносвязанные белки, включая, например, CD2, CD4, CD8, белок гемагглютинин вируса гриппа и другие хорошо известные в области техники, к которым существуют или могут быть получены стандартными способами высокоаффинные антитела, и слитые белки, содержащие мембрано-связанный белок, надлежащим образом слитый с доменом антигенной метки, помимо прочих, из гемагглютинина или Мус. Эти кодирующие последовательности, будучи ассоциированными с регуляторными элементами, которые направляют их экспрессию, обеспечивают сигналы, выявляемые стандартными способами, включая ферментативный, радиографический, колориметрический, флуоресцентный или иной спектральный анализ, проточную цитометрию и иммунологические анализы, включая иммуноферментный анализ (ELISA), радиоиммунный анализ (RIA) и иммуногистохимию.

В одном воплощении трансген представляет собой немаркерную последовательность, кодирующую продукт, который является полезным в биологии и медицине, такой как терапевтический трансген или иммуногенный трансген, такой как белки, РНК, ферменты или каталитические РНК. Желательные молекулы РНК включают тРНК (транспортная РНК), днРНК (двунитевая РНК), рибосомальную РНК, каталитические РНК и антисмысловые РНК. Одним примером полезной последовательности РНК является последовательность, которая подавляет экспрессию нуклеиновокислотной последовательности-мишени в животном, получающем лечение.

Трансген можно применять для лечения, например, генетических дефектов, в качестве противоопухолевого средства или вакцины, для индуцирования иммунного ответа и/или для профилактической вакцинации. В данном документе индуцирование иммунного ответа относится к

способности белка индуцировать Т-клеточный и/или гуморальный иммунный ответ на белок.

Термин «профилактика» означает получение лекарственного средства заранее, это может быть перед контактом с патогеном (доконтактная профилактика) или перед развитием симптомов заболевания (постконтактная профилактика). Термины «лечение» и «терапия» в данном описании используются взаимозаменяемо и означают введение лекарственного средства во время заболевания.

Под термином «заболевание» понимают нарушение структуры или функции субъекта, в частности, вызывающее специфические симптомы или оказывающее воздействие на определенный участок, и не являющееся просто прямым результатом физического повреждения.

Регуляторные элементы

В дополнение к трансгену вектор также включает стандартные контрольные элементы, которые функционально связаны с трансгеном таким образом, что обеспечивают его транскрипцию, трансляцию и/или экспрессию в клетке, трансфицированной плазмидным вектором или инфицированной вирусом, полученным по изобретению. В данном документе выражение «функционально связанный» охватывает как контролирующие экспрессию последовательности, которые являются смежными с интересующим геном, так и контролирующие экспрессию последовательности, которые контролируют интересующий ген, действуя *in trans* или на расстоянии.

Контролирующие экспрессию последовательности включают подходящие последовательности инициации и терминации транскрипции, последовательности промотора и энхансера; сигналы эффективного процессинга РНК, такие как сигналы сплайсинга и полиаденилирования (поли А), включая полиА бета-глобина кролика; последовательности, стабилизирующие цитоплазматическую мРНК; последовательности, повышающие эффективность трансляции (например, консенсусная последовательность Козак); последовательности, усиливающие стабильность белка и, при необходимости, последовательности, усиливающие секрецию кодируемого продукта. Среди прочих последовательностей можно использовать химерные интроны.

В некоторых воплощениях с трансгеном может быть функционально связан пост-транскрипционный регуляторный элемент вируса гепатита сурков (WPRE) (Zuffrey et al. (1999) J Virol; 73(4):2886-9). Пример WPRE приведен в SEQ ID NO: 26.

«Промотор» представляет собой нуклеотидную последовательность, которая позволяет связываться РНК полимеразе и направляет транскрипцию гена. Как правило, промотор расположен в 5' некодирующей области гена, вблизи сайта начала транскрипции гена. Элементы последовательности в составе промоторов, которые функционируют при инициации транскрипции, часто характеризуются консенсусными нуклеотидными последовательностями. Примеры промоторов включают промоторы бактерий, дрожжей, растений, вирусов и млекопитающих (включая человека), но не ограничиваются ими. В области техники известно и может применяться большое количество последовательностей, контролирующих экспрессию, включая промоторы, которые являются внутренними, нативными, конститутивными, индуцибельными и/или тканеспецифичными.

Примеры конститутивных промоторов включают промотор TBG (тироксинсвязывающий глобулин), промотор LTR (длинные концевые повторы) вируса саркомы Рауса (возможно с энхансером), промотор цитомегаловируса (CMV) (возможно с энхансером CMV, см., например, Boshart et al, Cell, 41:521-530 (1985)), промотор CASI, промотор SV40, промотор дигидрофолатредуктазы, промотор β -актина, промотор фосфоглицеролкиназы (PGK) и промотор EF1a (фактор элонгации 1a) (Invitrogen), но не ограничиваются

ими.

В некоторых воплощениях промотор представляет собой промотор CASI (см., например, WO2012/115980). Промотор CASI представляет собой синтетический промотор, содержащий часть энхансера CMV, часть промотора бета-актина цыпленка и часть энхансера UBC (убиквитин C). В некоторых воплощениях промотор CASI может включать нуклеиновокислотную последовательность, идентичную по меньшей мере приблизительно на 90%, по меньшей мере приблизительно на 95%, по меньшей мере приблизительно на 96%, по меньшей мере приблизительно на 97%, по меньшей мере приблизительно на 98%, по меньшей мере приблизительно на 99% или более SEQ ID NO: 12. В некоторых воплощениях промотор содержит нуклеотидную последовательность SEQ ID NO: 12 или состоит из нее.

Индукцибельные промоторы позволяют регулировать экспрессию гена и могут регулироваться посредством экзогенно вносимых соединений, факторов окружающей среды, таких как температура, или наличия определенного физиологического состояния, например острой фазы, конкретного состояния дифференцировки клетки или только в реплицирующихся клетках. Индукцибельные промоторы и индукцибельные системы доступны из различных коммерческих источников, включая Invitrogen, Clontech и Ariad, но не ограничиваясь ими. Описаны многие другие системы, которые могут быть легко выбраны специалистом в области техники. Например, индукцибельные промоторы включают цинк-индуцируемый промотор металлотионеина овцы (MT) и дексаметазон (Dex)-индуцируемый промотор вируса опухоли молочной железы мыши (MMTV). Другие индукцибельные системы включают систему промотора T7-полимеразы (WO 98/10088); промотор экдизона насекомых (No et al, Proc. Natl. Acad. Sci. USA, 93:3346-3351 (1996)), тетрациклин-репрессируемую систему (Gossen et al, Proc. Natl. Acad. Sci. USA, 89:5547-5551 (1992)), тетрациклин-индуцируемую систему (Gossen et al, Science, 378:1766-1769 (1995), см. также Harvey et al, Curr. Opin. Chem. Biol., 2:512-518 (1998)). Другие системы включают димер FK506, VP16 или p65, с использованием кастрадиола, дифенол-муристерона, RU486-индуцируемую систему (Wang et al, Nat. Biotech., 15:239-243 (1997) и Wang et al, Gene Ther., 4:432-441 (1997)) и рапамицин-индуцируемую систему (Magari et al, J. Clin. Invest., 100:2865-2872 (1997)). Эффективность некоторых индукцибельных промоторов увеличивается со временем. В таких случаях можно повысить эффективность таких систем путем встраивания нескольких tandemно расположенных репрессоров, например TetR, связанный с TetR посредством IRES (участок внутренней посадки рибосомы).

В некоторых воплощениях промотор представляет собой усиленный промотор hCMV, такой как представленный в SEQ ID NO: 42.

В другом воплощении применяют нативный для трансгена промотор. Нативный промотор может быть предпочтителен, когда желательно, чтобы экспрессия трансгена имитировала нативную экспрессию. Нативный промотор можно применять, когда экспрессия трансгена должна регулироваться временно или в процессе развития, или тканеспецифичным образом, или в ответ на определенные транскрипционные стимулы. В следующем воплощении для имитации нативной экспрессии можно также применять другие нативные контролирующие экспрессию элементы, такие как энхансерные элементы, сайты полиаденилирования или консенсусные последовательности Козак.

Трансген может быть функционально связан с тканеспецифичным промотором. Например, если желательна экспрессия в скелетной мышце, следует использовать промотор, активный в мышце. Они включают промоторы генов, кодирующих скелетный β -актин, легкую цепь 2A миозина, дистрофин, креатинкиназу мышц, а также синтетические промоторы мышц с активностью, превышающей таковую

естественных промоторов (см. Li et al, Nat. Biotech., 17:241-245 (1999)). Известны примеры тканеспецифичных промоторов для печени (альбумин Miyatake et al, J. Virol, 71:5124-32 (1997); основной промотор вируса гепатита В, Sandig et al, Gene Ther., 3:1002-9 (1996); альфа-фетопrotein (AFP), Arbuthnot et al., Hum. Gene Ther., 7: 1503-14 (1996)), остеокальцин кости (Stein et al, Mol. Biol. Rep., 24:185-96 (1997)); сиалопротеин кости (Chen et al., J. Bone Miner. Res., 11:654-64 (1996)), лимфоциты (CD2, Hansal et al, J. Immunol, 161:1063-8 (1998); тяжелая цепь иммуноглобулина; цепь Т клеточного рецептора), нейрональные, такие как промотор нейрон-специфической енолазы (NSE) (Andersen et al, Cell. Mol. Neurobiol, 13:503-15 (1993)), ген легкой цепи нейрофиламента (Piccioli et al, Proc. Natl. Acad. Sci. USA, 88:5611-5 (1991)) и ген нейрон-специфического *vgr* (Piccioli et al, Neuron, 15:373-84 (1995)), среди прочих.

Возможно, векторы, несущие трансгены, кодирующие терапевтически полезные или иммуногенные продукты, могут также включать селективные маркеры или репортерные гены, которые могут включать последовательности, обеспечивающие резистентность к генетицину, гиромидину или пуромицину, помимо прочих. Такие селективные репортерные или маркерные гены (предпочтительно расположенные вне вирусного генома, упаковываемого в вирусную частицу) можно применять, чтобы сигнализировать о наличии плазмид в бактериальных клетках, например как в случае резистентности к ампициллину. Другие компоненты вектора могут включать точку начала репликации.

Указанные векторы создают с применением методик и последовательностей, приведенных в данном документе, в сочетании со способами, известными специалистам в области техники. Такие методики включают стандартные методики клонирования кДНК, описанные в публикациях, применение перекрывающихся олигонуклеотидных последовательностей геномов аденовирусов, полимеразную цепную реакцию и любые подходящие способы, которые обеспечивают нужную нуклеотидную последовательность.

Терапия и профилактика

Рекомбинантные векторы на основе ChAd157 являются полезными для переноса генов человеку или млекопитающему, не являющемуся обезьяной, *in vitro*, *ex vivo* и *in vivo*.

Рекомбинантные аденовирусные векторы, описанные в данном документе, можно применять в качестве экспрессирующих векторов для получения продуктов, кодируемых гетерологичными трансгенами *in vitro*. Например, рекомбинантным репликативно-некомпетентным аденовирусом, содержащим трансген, может быть трансфицирована линия комплементирующих клеток, как описано выше.

Рекомбинантный аденовирусный вектор на основе ChAd157 обеспечивает эффективное средство переноса генов, которое может доставлять выбранный трансген в выбранную клетку-хозяина *in vivo* или *ex vivo*, даже когда организм имеет нейтрализующие антитела к одному или более серотипам аденовируса. В одном воплощении вектор и клетки смешивают *ex vivo*, инфицированные клетки культивируют с использованием стандартных методик и осуществляют реинфузию трансдуцированных клеток пациенту. Данные методики являются наиболее подходящими для доставки генов в терапевтических целях и для иммунизации, включая индуцирование защитного иммунитета.

Иммуногенные трансгены

Рекомбинантные векторы на основе ChAd157 также можно вводить в иммуногенных композициях. Иммуногенная композиция, описанная в данном документе, представляет собой композицию, содержащую один и более рекомбинантных векторов ChAd157, способных, после доставки млекопитающему, предпочтительно человеку, индуцировать иммунный ответ, например гуморальный (например, антительный) и/или клеточно-опосредованный (например, цитотоксическими Т-клетками) ответ против

трансгенного продукта, доставляемого вектором. Рекомбинантный аденовирус может содержать (предпочтительно, в любой из его делеций гена) ген, кодирующий желательный иммуноген и может, таким образом, применяться в вакцине. Рекомбинантные аденовирусы можно использовать в качестве профилактических или терапевтических вакцин против любого патогена, у которого идентифицирован(ы) антиген(ы), играющий(ие) важнейшую роль для индуцирования иммунного ответа и способные ограничивать распространение патогена, и для которых доступна кДНК.

Термин «иммуноген» означает полипептид, способный индуцировать иммунный ответ. Предпочтительно, иммуноген представляет собой антиген, который содержит по меньшей мере один В- или Т-клеточный эпитоп. Индуцированный иммунный ответ может представлять собой антиген-специфичный В-клеточный ответ, который приводит к образованию нейтрализующих антител. Индуцированный иммунный ответ может представлять собой антиген-специфичный Т-клеточный ответ, который может быть системным и/или местным ответом. Антиген-специфичный Т-клеточный ответ может включать CD4+ Т-клеточный ответ, такой как ответ, включающий CD4+ Т-клетки, экспрессирующие множество цитокинов, например, IFN (интерферон) гамма, TNF (фактор некроза опухоли) альфа и/или IL 2 (интерлейкин 2). В качестве альтернативы или дополнительно, антиген-специфичный Т-клеточный ответ включает CD8+ Т-клеточный ответ, такой как ответ, включающий CD8+ Т-клетки, экспрессирующие множество цитокинов, например IFN гамма, TNF альфа и/или IL 2.

Следовательно, термин «иммунизировать» означает введение иммуногена (или полинуклеотида, кодирующего иммуноген, в соответствии с контекстом) для индуцирования иммунного ответа.

Такие вакцины или иные иммуногенные композиции могут быть изготовлены в подходящем для доставки носителе. Как правило, дозы иммуногенных композиций находятся в диапазоне, который указан далее в разделе «Способы доставки и дозы». Для определения необходимости в бустерных введениях можно оценивать уровень иммунитета к выбранному гену. После определения титров антител в сыворотке могут быть желательными возможные бустерные иммунизации.

Возможно, вакцина или иммуногенная композиция по изобретению может быть изготовлена так, чтобы содержать другие компоненты, включая, например, адъюванты, стабилизаторы, регуляторы pH, консерванты и т.п. Примеры подходящих адъювантов приведены ниже в разделе «Адъюванты». Такой адъювант можно вводить с примирующей ДНК-вакциной, кодирующей антиген, для усиления антиген-специфичного иммунного ответа по сравнению с иммунным ответом, развивающимся после первичной вакцинации только ДНК-вакциной, кодирующей антиген. Альтернативно, такой адъювант можно вводить с полипептидным антигеном, который вводят в схеме введения, включающей векторы ChAd157 по изобретению (как описано ниже в разделе «Схемы введения»).

Рекомбинантные аденовирусы вводят в иммуногенном количестве, то есть в количестве рекомбинантного аденовируса, которое является эффективным при способе введения для трансфекции желаемых целевых клеток и обеспечения достаточных уровней экспрессии выбранного гена для индуцирования иммунного ответа. При обеспечении защитного иммунитета рекомбинантные аденовирусы считают вакцинными композициями, полезными для предупреждения инфекции и/или рецидивирующего заболевания.

Предполагают, что рекомбинантные векторы, описанные в данном документе, будут в высокой степени эффективными для индуцирования цитолитических Т-клеток и антител, направленных на вставленный гетерологичный антигенный белок, экспрессируемый вектором.

Иммуногены, экспрессируемые векторами по изобретению, которые полезны для иммунизации человека или животного, не являющегося человеком, против других патогенов, включают, например, бактерии, грибы, паразитические микроорганизмы или многоклеточных паразитов, инфицирующих человека и позвоночных, не являющихся человеком, или иммуногены из опухолевой клетки или раковой клетки. Например, иммуногены могут быть выбраны из ряда вирусных семейств. Примеры вирусных семейств, против которых желателен иммунный ответ, включают лиссавирусы, такие как вирус бешенства, респираторные вирусы, такие как респираторный синцитиальный вирус (RSV) и другие парамиксовирусы, такие как метапневмовирус человека, hMPV, и вирусы парагриппа (PIV).

Подходящие антигены вируса бешенства, которые полезны в качестве иммуногенов для иммунизации человека или животного, не являющегося человеком, могут быть выбраны из: гликопротеина вируса бешенства (G), РНК-полимеразы (L), белка матрикса (M), нуклеопротеина (N) и фосфопротеина (P). Термин «белок G» или «гликопротеин» или «полипептид белка G» или «полипептид гликопротеина» относится к полипептиду или белку, имеющему всю или часть аминокислотной последовательности полипептида гликопротеина вируса бешенства. Термин «белок L» или «белок РНК полимеразы» или «полипептид белка L» или «полипептид белка РНК полимеразы» относится к полипептиду или белку, имеющему всю или часть аминокислотной последовательности полипептида белка РНК полимеразы вируса бешенства. Термин «белок M» или «белок матрикса» или «полипептид белка M» или «полипептид белка матрикса» относится к полипептиду или белку, имеющему всю или часть аминокислотной последовательности полипептида белка матрикса вируса бешенства. Аналогично, термин «белок N» или «нуклеопротеин» или «полипептид белка N» или «полипептид нуклеопротеина» относится к полипептиду или белку, имеющему всю или часть аминокислотной последовательности полипептида нуклеопротеина вируса бешенства. Аналогично, термин «белок P» или «фосфопротеин» или «полипептид белка P» или «полипептид фосфопротеина» относится к полипептиду или белку, имеющему всю или часть аминокислотной последовательности полипептида фосфопротеина вируса бешенства.

Подходящие антигены RSV, которые полезны в качестве иммуногенов для иммунизации человека или животного, не являющегося человеком, могут быть выбраны из: белка слияния (F), белка присоединения (G), белка матрикса (M2) и нуклеопротеина (N). Термин «белок F» или «белок слияния», или «полипептид белка F», или «полипептид белка слияния» относится к полипептиду или белку, имеющему всю или часть аминокислотной последовательности полипептида белка слияния RSV. Аналогично, термин «белок G» или «полипептид белка G» относится к полипептиду или белку, имеющему всю или часть аминокислотной последовательности полипептида белка присоединения RSV. Термин «белок M» или «белок матрикса», или «полипептид белка M» относится к полипептиду или белку, имеющему всю или часть аминокислотной последовательности белка матрикса RSV, и может включать либо один, либо оба из генных продуктов M2-1 (который в данном документе может указываться как M2.1) и M2-2. Аналогично, термин «белок N» или «белок нуклеокапсида», или «полипептид белка N» относится к полипептиду или белку, имеющему всю или часть аминокислотной последовательности нуклеопротеина RSV.

Описаны две группы штаммов RSV человека, группы A и B, основанные, главным образом, на различиях антигенности гликопротеина G. В настоящее время выделены различные штаммы RSV, любые из которых являются подходящими, когда речь идет об антигенах иммуногенных комбинаций, описанных в данном документе. Примеры штаммов, обозначенных номерами доступа GenBank и/или EMBL, можно найти в опубликованной заявке на патент US 2010/0203071 (WO2008114149), включенной в данное

описание путем ссылки для изложения нуклеотидных и полипептидных последовательностей белков F и G RSV, подходящих для применения в данном изобретении. В воплощении белок F RSV может представлять собой эктодомен белка F RSV (FATM).

Примеры нуклеиновокислотных и белковых последовательностей белков M и N можно найти, например, в опубликованной заявке на патент US 2014/0141042 (WO2012/089833), включенной в данное описание для изложения нуклеотидных и полипептидных последовательностей белков M и N RSV, подходящих для применения в данном изобретении.

Подходящим образом, для применения в данном изобретении нуклеиновая кислота кодирует антиген F RSV и антигены M и N RSV. Более конкретно, нуклеиновая кислота кодирует FATM антиген RSV и M2-1 и N антигены RSV, где сайт саморасщепления находится между FATM антигеном RSV и M2-1 RSV, а гибкий линкер расположен между антигенами M2-1 и N RSV. В одном воплощении подходящая нуклеиновая кислота кодирует полипептид, представленный SEQ ID NO:37.

В одном воплощении иммуноген может иметь происхождение из ретровируса, например лентивируса, такого как вирус иммунодефицита человека (HIV). В таком воплощении иммуногены могут иметь происхождение из HIV-1 или HIV-2.

Геном HIV кодирует ряд различных белков, каждый из которых может быть иммуногенным в целом виде или в виде фрагмента, при экспрессии векторами по данному изобретению. Белки оболочки включают, например, gp120, gp41 и предшественник Env gp160. Белки HIV, не являющиеся белками оболочки, включают, например, внутренние структурные белки, такие как продукты генов gag и pol и другие неструктурные белки, такие как Rev, Nef, Vif и Tat. В воплощении вектор по изобретению кодирует один или более полипептидов, содержащих HIV Gag.

Ген Gag транслируется в виде полипротеина-предшественника, который расщепляется протеазой с образованием продуктов, включающих белок матрикса (p17), капсид (p24), нуклеокапсид (p9), p6 и два соединительных пептида, p2 и p1, все из которых являются примерами фрагментов Gag.

Ген Gag дает начало белку-предшественнику Gag размером 55 килодальтон (кДа), также обозначаемому как p55, который экспрессируется с не прошедшей сплайсинг вирусной мРНК. В ходе трансляции N конец p55 подвергается миристоилированию, что запускает его ассоциацию с цитоплазматической стороной клеточных мембран. Ассоциированный с мембраной полипротеин Gag захватывает две копии вирусной геномной РНК наряду с другими вирусными и клеточными белками, тем самым запуская процесс отпочковывания вирусной частицы от поверхности инфицированной клетки. После отпочковывания p55 расщепляется протеазой, кодируемой вирусом (продукт гена pol) в ходе процесса созревания вируса, на четыре более мелких белка, обозначаемых как MA (матрикс [p17]), CA (капсид [p24]), NC (нуклеокапсид [p9]) и p6, все из которых являются примерами фрагментов Gag. В одном воплощении векторы по данному изобретению содержат полипептид Gag SEQ ID NO: 16.

Адьюванты

Термин «адьювант» в данном документе относится к композиции, которая усиливает иммунный ответ на иммуноген. Примеры таких адьювантов включают неорганические адьюванты (например, неорганические соли металлов, такие как фосфат алюминия или гидрат окиси алюминия), органические адьюванты (например, сапонины, такие как QS21 или сквален), адьюванты на основе масел (например, полный адьювант Фрейнда и неполный адьювант Фрейнда), цитокины (например, IL-1 β (интерлейкин 1 β), IL-2, IL-7, IL-12, IL-18, GM-CSF и INF- γ (интерферон γ)), адьюванты в виде частиц (например,

иммуностимулирующие комплексы (ISCOMS), липосомы или биodeградируемые микросферы), виросомы, бактериальные адъюванты (например, монофосфориллипид А, такой как 3-де-О-ацилированный монофосфориллипид А (3D-MPL) или мурамилпептиды), синтетические адъюванты (например, неионные блок-сополимеры, аналоги мурамилпептидов или синтетический липид А), синтетические полинуклеотидные адъюванты (например, полиаргинин или полилизин) и иммуностимулирующие олигонуклеотиды, содержащие неметилированные динуклеотиды CpG («CpG»), но не ограничиваются ими.

Одним из подходящих адъювантов является монофосфориллипид А (MPL), в частности, 3-де-О-ацилированный монофосфориллипид А (3D-MPL). Химически он часто поставляется в виде смеси 3-де-О-ацилированного монофосфориллипид А с 4, 5 или 6 ацилированными цепями. Его можно очистить и получить способами, изложенными в GB 2122204B, где также изложено получение дифосфориллипид А и его 3-О-деацилированных вариантов. Описаны другие очищенные и синтетические липополисахариды (патенты US 6005099 и EP 0 729 473 B1; Hilgers et al., 1986, *Int. Arch. Allergy. Immunol.*, 79(4):392-6; Hilgers et al., 1987, *Immunology*, 60(1):141-6; и EP 0 549 074 B1).

Сапонины также представляют собой подходящие адъюванты (см. Lacaille-Dubois, M and Wagner H, A review of the biological and pharmacological activities of saponins. *Phytomedicine* vol 2 pp 363-386 (1996)). Например, сапонин Quil А (полученный из коры южноамериканского дерева *Quillaja Saponaria* Molina) и его фракции, описанные в патенте US 5057540 и Kensil, *Crit. Rev. Ther. Drug Carrier Syst.*, 1996, 12:1-55; и EP 0 362 279 B1. Очищенные фракции Quil А также известны как иммуностимуляторы, такие как QS21 и QS17; способы их получения описаны в патентах US 5057540 и EP 0 362 279 B1. В указанных источниках также описан QS7 (негемолитическая фракция Quil-A). Применение QS21 также описано Kensil et al. (1991, *J. Immunology*, 146: 431-437). Также известны комбинации QS21 с полисорбатом или циклодекстрином (WO 99/10008). Системы адъювантов в виде частиц, содержащие фракции QuilA, такие как QS21 и QS7, описаны в WO 96/33739 и WO 96/11711.

Другие адъюванты представляют собой иммуностимулирующие олигонуклеотиды, содержащие неметилированные динуклеотиды CpG ("CpG") (Krieg, *Nature* 374:546 (1995)). CpG это аббревиатура, обозначающая динуклеотидный мотив цитозин-гуанозин, присутствующий в ДНК. CpG известен как адъювант, вводимый как системно, так и через слизистую (WO 96/02555, EP 468520, Davis et al, *J. Immunol*, 1998, 160:870-876; McCluskie and Davis, *J. Immunol.*, 1998, 161:4463-6). CpG, включенный в состав вакцин, можно вводить в виде свободного раствора вместе со свободным антигеном (WO 96/02555) или ковалентно конъюгированный с антигеном (WO 98/16247), или приготовленный с носителем, таким как гидрат окиси алюминия (Brazolot-Millan et al., *Proc. Natl. Acad. Sci., USA*, 1998, 95:15553-8).

Адъюванты, такие как описанные выше, могут быть приготовлены вместе с носителями, такими как липосомы, эмульсии типа «масло в воде» и/или солями металлов (включая соли алюминия, такие как гидрат окиси алюминия). Например, 3D-MPL может быть приготовлен вместе с гидратом окиси алюминия (EP 0 689 454) или эмульсиями масло в воде (WO 95/17210); QS21 может быть приготовлен вместе с липосомами, содержащими холестерин (WO 96/33739), эмульсиями типа «масло в воде» (WO 95/17210) или квасцами (WO 98/15287); CpG может быть приготовлен вместе с квасцами (Brazolot-Millan, см. выше) или другими катионными носителями.

В данном изобретении можно использовать комбинации адъювантов, в частности комбинацию монофосфориллипид А и производного сапонины (см., например, WO 94/00153; WO 95/17210; WO 96/33739; WO 98/56414; WO 99/12565; WO 99/11241), более конкретно комбинацию QS21 и 3D-MPL, как

описано в WO 94/00153, или композицию, где QS21 содержится в липосомах, содержащих холестерин (DQ), как описано в WO 96/33739. Альтернативно, комбинация CpG и сапонины, такого как QS21, представляет собой адъювант, подходящий для применения в данном изобретении. Эффективная адъювантная композиция, включающая QS21, 3D-MPL и токоферол в эмульсии «масло в воде» описана в WO 95/17210 и представляет собой другую композицию для применения в данном изобретении. Сапониновые адъюванты можно готовить в липосомах и комбинировать с иммуностимулирующим олигонуклеотидом. Так, подходящие адъювантные системы включают, например, комбинацию монофосфориллипида А, предпочтительно 3D-MPL, вместе с солью алюминия (например, как описано в WO00/23105). Дополнительные примеры адъювантов включают QS21 и/или MPL и/или CpG. QS21 может содержаться в липосомах, содержащих холестерин, как описано в WO 96/33739.

Другие подходящие адъюванты включают алкил-глюкозаминид-фосфаты (AGP), такие, как описанные в WO9850399 или в US 6303347 (также описывает способы получения AGP), или фармацевтически приемлемые соли AGP, как описано в патенте US 6764840. Некоторые AGP представляют собой агонисты TLR4, а некоторые являются антагонистами TLR4. И те, и другие можно применять в качестве адъювантов.

Обнаружили, что слияние инвариантной цепи с антигеном, входящим в состав системы экспрессии, используемой для вакцинации, повышает иммунный ответ на указанный антиген, если его вводят с аденовирусом (WO 2007/062656, опубликованная как US 2011/0293704 и включенная путем ссылки для изложения инвариантных последовательностей цепей). Соответственно, в одном воплощении изобретения иммуногенный трансген может экспрессироваться совместно с инвариантной цепью в составе рекомбинантного вирусного вектора ChAd157.

В другом воплощении изобретения предложено применение капсида ChAd157 (возможно применяют интактную или рекомбинантную вирусную частицу или пустой капсид) для индуцирования ответа - иммуномодулирующего эффекта или для усиления ответа цитотоксических Т-клеток на другой активный агент или выполнения функции адъюванта посредством доставки субъекту капсида ChAd157. Капсид ChAd157 можно доставлять в отдельности или в комбинации с активным агентом для усиления иммунного ответа на него. Предпочтительно, желаемый эффект может достигаться без инфицирования хозяина аденовирусом.

Схемы введения

Главным образом, рекомбинантные векторы на основе аденовируса ChAd157 будут использовать для доставки терапевтических или иммуногенных молекул (таких как белки). Нетрудно понять, что при обоих применениях рекомбинантные аденовирусные векторы по изобретению особенно хорошо подходят для применения в схемах, в которые вовлечено повторное введение рекомбинантных аденовирусных векторов. Такие схемы, как правило, включают доставку серии вирусных векторов, в которых вирусные капсиды сменяют друг друга. Вирусные капсиды можно менять при каждом последующем введении или после определенного числа введений капсида конкретного серотипа (например, одного, двух, трех, четырех или более). Таким образом, схема может включать доставку рекомбинантного аденовируса с первым капсидом, доставку рекомбинантного аденовируса со вторым капсидом и доставку рекомбинантного аденовируса с третьим капсидом. Специалистам в области техники будет очевидно множество других схем, в которых используют аденовирусные капсиды по изобретению в отдельности, в комбинации друг с другом или в комбинации с другими аденовирусами (которые предпочтительно не демонстрируют

иммунологической перекрестной реактивности). Возможно такая схема может включать введение рекомбинантного аденовируса с капсидами других аденовирусов приматов, не являющихся человеком, аденовирусов человека или искусственными последовательностями, как описано в данном документе.

Аденовирусные векторы по изобретению особенно хорошо подходят для схем лечения, в которых необходима многократная доставка трансгенов, опосредованная аденовирусами, например в схемах, включающих повторяющуюся доставку того же самого трансгена или в комбинированных схемах, включающих доставку других трансгенов. Такие схемы могут включать введение аденовирусного вектора ChAd157 с последующим повторным введением вектора на основе аденовируса того же самого серотипа. Наиболее предпочтительные схемы включают введение аденовирусного вектора ChAd157, в котором источник аденовирусных капсидных последовательностей у вектора, доставляемого при первом введении, отличается от источника аденовирусных капсидных последовательностей у вирусного вектора, используемого в одном или более последующих введениях. Например, схема лечения включает введение вектора ChAd157 и повторное введение одного или более аденовирусных векторов того же самого или других серотипов.

В другом примере схема лечения включает введение аденовирусного вектора с последующим повторным введением вектора ChAd157, который имеет капсид, который отличается от источника капсида в аденовирусном векторе, введенном первым, и возможно дополнительное введение другого вектора, который является таким же или, предпочтительно, отличается от источника аденовирусного капсида вектора на предшествующих стадиях введения. Указанные схемы не ограничиваются доставкой аденовирусных векторов, сконструированных с применением последовательностей ChAd157. Скорее, в этих схемах можно легко применять другие аденовирусные последовательности, включая, без ограничения, другие аденовирусные последовательности, включая другие аденовирусные последовательности приматов, не являющихся человеком, или аденовирусные последовательности человека, в комбинации с векторами ChAd157.

В дополнительном примере схема лечения может включать либо одновременную (такую как совместное введение), либо последовательную (такую как прайм-буст) доставку (1) одного или более аденовирусных векторов ChAd157 и (2) дополнительного компонента, такого как векторы, не являющиеся аденовирусными, невирусные векторы и/или ряд других терапевтически полезных соединений или молекул, таких как антигенные белки, возможно вводимые одновременно с адъювантом. Примеры совместного введения включают гомолатеральное совместное введение и контралатеральное совместное введение (более подробно описанные ниже в разделе «Способы доставки и дозы»).

Подходящие векторы, не являющиеся аденовирусными, для применения в одновременной или особенно в последовательной доставке (такой как прайм-буст) с одним или более аденовирусными векторами на основе ChAd157 включают один или более поксвирусных векторов. Предпочтительно, поксвирусные векторы принадлежат к подсемейству *chordopoxvirinae*, более предпочтительно к роду в указанном подсемействе, выбранному из группы, состоящей из *orthopox*, *parapox*, *yatapox*, *avipox* (предпочтительно, *canarypox* (ALVAC) или *fowlpox* (FPV)) и *molluscipox*. Еще более предпочтительно, поксвирусный вектор относится к роду *orthopox* и выбран из группы, состоящей из вируса *vaccinia*, NYVAC (производного штамма *vaccinia* Copenhagen), модифицированного штамма *Vaccinia Ankara* (MVA), вируса коровьей оспы и вируса обезьяньей оспы. Более предпочтительно, поксвирусный вектор представляет собой MVA.

«Одновременное» введение предпочтительно относится к тому же существующему иммунному ответу. Предпочтительно, оба компонента вводят в одно и то же время (например, совместное введение как ДНК, так и белка), однако, один компонент можно вводить в течение нескольких минут (например, в ходе того же посещения медицинского учреждения или визите врача), в течение нескольких часов. Такое введение также обозначают как совместное введение. В некоторых воплощениях совместное введение может относиться к введению аденовирусного вектора, адъюванта и белкового компонента. В других воплощениях совместное введение относится к введению аденовирусного вектора и другого вирусного вектора, например второго аденовирусного вектора или поксвируса, такого как MVA. В других воплощениях совместное введение относится к введению аденовирусного вектора и белкового компонента, к которому возможно добавлен адъювант.

Можно применять схему прайм-буст. Прайм-буст относится к двум отдельным иммунным ответам: (1) первичная стимуляция иммунной системы с последующей (2) вторичной, или бустерной, стимуляцией иммунной системы через несколько недель или месяцев после развития первичного иммунного ответа.

Такая схема может включать введение рекомбинантного вектора ChAd157 для первичной стимуляции иммунной системы и второе бустерное введение обычного антигена, такого как белок (возможно вводимого совместно с адъювантом) или рекомбинантного вируса, несущего последовательности, кодирующие такой антиген (например, WO 00/11140). Альтернативно, схема иммунизации может включать введение рекомбинантного вектора ChAd157 для усиления иммунного ответа на вектор (или вирусный, или на основе ДНК), кодирующий антиген. В другом альтернативном варианте схема иммунизации включает введение белка с последующим бустерным введением рекомбинантного вектора ChAd157, кодирующего антиген. В одном примере схема прайм-буст может обеспечивать защитный иммунный ответ на вирус, бактерию или другой организм, из которого имеет происхождение антиген. В другом воплощении схема прайм-буст обеспечивает терапевтический эффект, который можно измерить с использованием стандартных анализов для выявления наличия состояния, для которого вводят терапию.

Предпочтительно, бустерную композицию вводят через от приблизительно 2 до приблизительно 27 недель после введения субъекту примиряющей композиции. Введение бустерной композиции осуществляют с использованием эффективного количества бустерной композиции, содержащей или способной доставлять тот же антиген или другой антиген по сравнению с введенным в составе примиряющей вакцины. Бустерная композиция может состоять из рекомбинантного вирусного вектора, имеющего происхождение из того же самого вирусного источника или из другого источника. Альтернативно, бустерная композиция может представлять собой композицию, индуцирующую иммунный ответ у хозяина, содержащую тот же антиген, что и кодируемый примиряющей вакциной, но в форме белка. Основные требования к бустерной композиции состоят в том, чтобы антиген композиции представлял собой тот же антиген или антиген, перекрестно-реагирующий с антигеном, кодируемым примиряющей композицией.

Низкая перекрестная реактивность между нейтрализующими антителами к ChAd157 и некоторым другим аденовирусным векторам, таким как ChAd155, обладает преимуществами, когда требуются многократные введения векторов. Многократные введения могут осуществляться для отдельной доставки различных транстенов (например, кодирующих иммуногены, ассоциированные с различными медицинскими показаниями) или доставки одинаковых или похожих транстенов (например, согласно схеме прайм-буст для усиления иммунного ответа при конкретном медицинском показании).

Следовательно, предложен рекомбинантный аденовирусный вектор по изобретению, кодирующий

трансген, для введения субъекту, который ранее был в контакте с рекомбинантным аденовирусным вектором, который не содержит фибриллу ChAd157 или ее функциональное производное, как описано в данном документе (например, не содержит фибриллу, гексон или пентон ChAd157, как описано в данном документе, таким как рекомбинантный аденовирусный вектор, содержащий фибриллу, гексон и/или пентон ChAd155, в частности рекомбинантный аденовирусный вектор, содержащий фибриллу, гексон и пентон ChAd155). В частности, предложен рекомбинантный аденовирусный вектор по изобретению, кодирующий трансген, для введения субъекту, которому ранее вводили рекомбинантный аденовирусный вектор, который не содержит фибриллу ChAd157 или ее функциональное производное, как описано в данном документе (например, не содержит фибриллу, гексон или пентон ChAd157, как описано в данном документе, таким как рекомбинантный аденовирусный вектор, содержащий фибриллу, гексон и/или пентон ChAd155, в частности рекомбинантный аденовирусный вектор, содержащий фибриллу, гексон и пентон ChAd155). Предпочтительно, рекомбинантный аденовирусный вектор, который не содержит фибриллу ChAd157, обладает низкой перекрестной реактивностью с ChAd157. В одном воплощении рекомбинантный аденовирусный вектор, который не содержит фибриллу ChAd157, кодирует трансген, направленный на иное медицинское показание или показания, нежели трансген рекомбинантного аденовирусного вектора по изобретению. В другом воплощении рекомбинантный аденовирусный вектор, который не содержит фибриллу ChAd157, кодирует трансген, направленный на то же медицинское показание или показания, что и трансген рекомбинантного аденовирусного вектора по изобретению (например, тот же самый трансген).

Также предложен рекомбинантный аденовирусный вектор по изобретению, кодирующий трансген, для введения субъекту, который может (то есть который намерен или ожидаемо будет) впоследствии оказаться в контакте с рекомбинантным аденовирусным вектором, который не содержит фибриллу ChAd157 или ее функциональное производное, как описано в данном документе (например, не содержит фибриллу, гексон или пентон ChAd157, как описано в данном документе, таким как рекомбинантный аденовирусный вектор, содержащий фибриллу, гексон и/или пентон ChAd155, в частности, с рекомбинантным аденовирусным вектором, содержащим фибриллу, гексон и пентон ChAd155). В частности, предложен рекомбинантный аденовирусный вектор по изобретению, кодирующий трансген, для введения субъекту, которому впоследствии могут вводить рекомбинантный аденовирусный вектор, который не содержит фибриллу ChAd157 или ее функциональное производное, как описано в данном документе (например, не содержит фибриллу, гексон или пентон ChAd157, как описано в данном документе, таким как рекомбинантный аденовирусный вектор, содержащий фибриллу, гексон и/или пентон ChAd155, в частности рекомбинантный аденовирусный вектор, содержащий фибриллу, гексон и пентон ChAd155). Предпочтительно, рекомбинантный аденовирусный вектор, который не содержит фибриллу ChAd157, обладающий низкой перекрестной реактивностью с ChAd157. В одном воплощении рекомбинантный аденовирусный вектор, который не содержит фибриллу ChAd157, кодирует трансген, направленный на иное медицинское показание или показания, нежели трансген рекомбинантного аденовирусного вектора по изобретению. В другом воплощении рекомбинантный аденовирусный вектор, который не содержит фибриллу ChAd157, кодирует трансген, направленный на то же медицинское показание или показания, что и трансген рекомбинантного аденовирусного вектора по изобретению (например, тот же самый трансген).

Таким образом, в данном изобретении предложен способ индуцирования иммунного ответа у субъекта, включающий:

- (а) введение субъекту рекомбинантного аденовирусного вектора по изобретению, кодирующего

первый трансген, и

(б) введение субъекту рекомбинантного аденовирусного вектора, который не содержит фибриллу ChAd157 или ее функциональное производное, как описано в данном документе, где вектор кодирует второй трансген;

где стадии (а) и (б) можно осуществлять в любом порядке и первый и второй трансгены могут быть одинаковыми или различными.

Первый и второй трансгены обычно кодируют иммуногены, которые являются полезными для иммунизации человека или животного, не являющегося человеком, против патогена, такого как бактерии, грибы, паразитические микроорганизмы или многоклеточные паразиты, инфицирующие человека и позвоночных, не являющихся человеком, или против раковой клетки или опухолевой клетки. Первый и второй трансгены могут кодировать одинаковые или различные иммуногены. Когда кодируемые иммуногены различаются, они могут быть направлены на одинаковые или разные патогены или раковые клетки или опухолевые клетки.

Следовательно, также предложен способ профилактики или лечения субъекта, включающий:

(а) введение субъекту рекомбинантного аденовирусного вектора по изобретению, кодирующего первый трансген, кодирующий иммуноген, который полезен для иммунизации человека или животного, не являющегося человеком, против патогена, такого как бактерии, грибы, паразитические микроорганизмы или многоклеточные паразиты, инфицирующие человека и позвоночных, не являющихся человеком, или против раковой клетки или опухолевой клетки; и

(б) введение субъекту рекомбинантного аденовирусного вектора, не содержащего фибриллу ChAd157 или ее функциональное производное, как описано в данном документе, где вектор кодирует второй трансген, кодирующий иммуноген, который полезен для иммунизации человека или животного, не являющегося человеком, против иного патогена, такого как бактерии, грибы, паразитические микроорганизмы или многоклеточные паразиты, инфицирующие человека и позвоночных, не являющихся человеком, или против раковой клетки или опухолевой клетки;

где стадии (а) и (б) можно осуществлять в любом порядке.

Рекомбинантный аденовирусный вектор, не содержащий фибриллу ChAd157 или ее функциональное производное, как описано в данном документе, предпочтительно не содержащий фибриллу ChAd157, гексон ChAd157 или фибриллу ChAd157, такой как не содержащий фибриллу ChAd157, гексон ChAd157 или фибриллу ChAd157 или ее функциональные производные, имеющие по меньшей мере 98% идентичности с ней.

Рекомбинантный аденовирусный вектор, который не содержит фибриллу ChAd157 или ее функциональное производное, как описано в данном документе, может представлять собой рекомбинантный аденовирусный вектор, содержащий фибриллу, гексон и/или пентон ChAd155, в частности рекомбинантный аденовирусный вектор, содержащий фибриллу, гексон и пентон ChAd155.

Как упоминалось, рекомбинантный аденовирусный вектор по изобретению можно применять для доставки терапевтических или иммуногенных молекул вместе с рекомбинантным аденовирусным вектором, содержащим фибриллу, гексон и/или пентон ChAd155. Рекомбинантный аденовирусный вектор, содержащий фибриллу, гексон и/или пентон ChAd155, будет содержать фибриллу, пентон и/или гексон в соответствии с SEQ ID NO: 7, 9 и 11, в частности фибриллу, пентон и гексон в соответствии с SEQ ID NO: 7, 9 и 11.

Термин «низкая перекрестная реактивность» означает, что иммунизация первым вектором не индуцирует значимого ответа, представляющего собой выработку нейтрализующих антител на второй вектор, то есть существенно не влияют на иммунологическую эффективность второго вектора. Ответ, представляющий собой выработку нейтрализующих антител, можно определять способами, аналогичными приведенным в Примере 7 в настоящем документе. Желательно, двукратная иммунизация первым вектором вызывает нейтрализующий титр, который в среднем составляет менее 50% от уровня, развивающегося после иммунизации вторым вектором, например менее 75%, предпочтительно менее 90%.

Термин «субъект» означает любое животное, предпочтительно млекопитающее, и в частности человека.

Способы доставки и дозы

Вектор может быть подготовлен для введения посредством суспендирования или растворения в фармацевтически или физиологически приемлемом носителе, таком как изотонический физиологический раствор; изотонический солевой раствор или другие композиции будут очевидны специалистам в области техники. Подходящий носитель будет очевиден специалистам в области техники и будет зависеть по большей части от способа введения. Композиции, описанные в данном документе, можно вводить млекопитающему в композиции с замедленным высвобождением, с применением биodeградируемого биосовместимого полимера или посредством местной доставки с применением мицелл, гелей и липосом.

В некоторых воплощениях рекомбинантный аденовирус по изобретению вводят субъекту путем внутримышечной инъекции, интравагинальной инъекции, внутривенной инъекции, внутрибрюшинной инъекции, подкожной инъекции, нанесения на кожу, внутрикожного введения, назального введения или перорального введения. Сублингвальное введение также может представлять интерес.

Если схема терапии включает совместное введение одного или более векторов на основе аденовируса ChAd157 и дополнительного компонента, каждый из которых входит в разные композиции, их предпочтительно вводить близко в один и тот же участок или вблизи одного и того же участка. Например, компоненты можно вводить (например, способом введения, выбранным из внутримышечного, чрескожного, внутрикожного, подкожного) на той же стороне или в ту же конечность («ко-латеральное» введение) или на противоположной стороне или в противоположную конечность («контра-латеральное» введение).

Дозы вирусных векторов будут зависеть преимущественно от таких факторов, как состояние, подлежащее лечению, возраст, масса тела и здоровье пациента, и, следовательно, могут варьировать у пациентов. Например, терапевтически эффективная доза вирусного вектора для взрослого человека или терапевтически эффективная ветеринарная доза обычно содержит от 1×10^5 до 1×10^{15} вирусных частиц, например от 1×10^8 до 1×10^{12} (например 1×10^8 , $2,5 \times 10^8$, 5×10^8 , 1×10^9 , $1,5 \times 10^9$, $2,5 \times 10^9$, 5×10^9 , 1×10^{10} , $1,5 \times 10^{10}$, $2,5 \times 10^{10}$, 5×10^{10} , 1×10^{11} , $1,5 \times 10^{11}$, $2,5 \times 10^{11}$, 5×10^{11} , 1×10^{12} частиц). Альтернативно, вирусный вектор можно вводить в дозе, которая обычно составляет от 1×10^5 до 1×10^{10} бляшкообразующих единиц (PFU), например 1×10^5 PFU, $2,5 \times 10^5$ PFU, 5×10^5 PFU, 1×10^6 PFU, $2,5 \times 10^6$ PFU, 5×10^6 PFU, 1×10^7 PFU, $2,5 \times 10^7$ PFU, 5×10^7 PFU, 1×10^8 PFU, $2,5 \times 10^8$ PFU, 5×10^8 PFU, 1×10^9 PFU, $2,5 \times 10^9$ PFU, 5×10^9 PFU или 1×10^{10} PFU. Дозы будут варьировать в зависимости от размера животного и способа введения. Например, подходящая доза для человека или ветеринарная доза (для животного массой приблизительно 80 кг) для внутримышечной инъекции находится в диапазоне от приблизительно 1×10^9 до приблизительно 5×10^{12} на мл для одного участка. Возможно, можно использовать несколько участков введения. В другом примере подходящая доза для человека или ветеринарная доза может находиться в диапазоне от приблизительно 1×10^{11} до

приблизительно 1×10^{15} частиц для пероральной композиции.

Количественное определение вирусного вектора можно осуществлять анализом посредством количественной PCR (полимеразной цепной реакции) (Q-PCR), например, с праймерами и зондом, сконструированными для области промотора CMV, используя в качестве стандартной кривой серию разведений плазмидной ДНК, содержащей геном вектора с экспрессионной кассетой, содержащей промотор HCMV. Число копий в исследуемом образце определяют методом анализа параллельных линий. Альтернативными способами количественного определения частиц вектора могут быть аналитическая HPLC (высокоэффективная жидкостная хроматография) или спектрофотометрический метод, основанный на поглощении при 260 нм (A_{260}).

Иммунологически эффективное количество нуклеиновой кислоты может предпочтительно составлять от 1 нг до 100 мкг. Например, подходящее количество может составлять от 1 мкг до 100 мкг. Специалисты в области техники могут легко установить подходящее количество конкретной нуклеиновой кислоты (например, вектора). Примеры эффективных количеств нуклеиновокислотного компонента могут составлять от 1 нг до 100 мкг, например от 1 нг до 1 мкг (например, 100 нг - 1 мкг) или от 1 мкг до 100 мкг, например 10 нг, 50 нг, 100 нг, 150 нг, 200 нг, 250 нг, 500 нг, 750 нг или 1 мкг. Эффективные количества нуклеиновой кислоты могут также включать от 1 мкг до 500 мкг, например от 1 мкг до 200 мкг, например от 10 до 100 мкг, например 1 мкг, 2 мкг, 5 мкг, 10 мкг, 20 мкг, 50 мкг, 75 мкг, 100 мкг, 150 мкг или 200 мкг. Альтернативно, приведенное в качестве примера эффективное количество нуклеиновой кислоты может составлять от 100 мкг до 1 мг, например от 100 мкг до 500 мкг, например 100 мкг, 150 мкг, 200 мкг, 250 мкг, 300 мкг, 400 мкг, 500 мкг, 600 мкг, 700 мкг, 800 мкг, 900 мкг или 1 мг.

Как правило, доза для человека будет находиться в объеме от 0,1 мл до 2 мл. Таким образом, описанная в данном документе композиция может быть приготовлена в объеме, например, 0,1, 0,15, 0,2, 0,5, 1,0, 1,5 или 2,0 мл человеческой дозы для каждого из иммуногенных компонентов в отдельности или для их комбинации.

Специалист в области техники может скорректировать указанные дозы в зависимости от способа введения и терапевтического или вакцинного применения, при котором используют рекомбинантный вектор. Для определения частоты введения доз можно отслеживать уровни экспрессии трансгена или, для адъюванта, уровень циркулирующего антитела.

Если применяют одну или более стадий первичной и/или бустерной иммунизации, данная стадия может включать однократную дозу, которую вводят ежечасно, ежедневно, еженедельно, ежемесячно или ежегодно. Например, млекопитающие могут получать одну или две дозы, содержащие от приблизительно 10 мкг до приблизительно 50 мкг плазмиды в носителе. Количество или участок доставки желательно выбирать, основываясь на особенностях и состоянии млекопитающего.

Для определения необходимости в бустерных введениях, если таковая возникнет, можно оценивать терапевтические уровни или уровень иммунного ответа против белка, кодируемого заданным трансгеном. После оценки ответа CD8⁺ Т клеток или возможно титра антител в сыворотке, могут быть желательны возможные бустерные иммунизации. Возможно, рекомбинантные векторы ChAd155 могут быть доставлены в виде однократного введения или в различных комбинированных схемах, например в комбинации со схемой или курсом лечения, включающей(им) другие активные ингредиенты или в составе схемы прайм-буст.

Далее изобретение будет описано более подробно посредством примеров, которые не являются

исчерпывающими.

ОПИСАНИЕ ПРИМЕРОВ ОСУЩЕСТВЛЕНИЯ ИЗОБРЕТЕНИЯ

ПРИМЕР 1. Выделение ChAd157 и конструирование вектора

29 различных аденовирусов шимпанзе дикого типа выделяли у здоровых молодых шимпанзе, которых содержали в различных вивариях Европы, используя стандартные процедуры, как описано Colloca et al. Sci Transl Med. 2012 Jan 4;4(115):115ra2 и в WO2010/086189, содержание которых включено путем ссылки в целях описания выделения аденовируса и способов характеристики.

Затем 29 вирусов дикого типа пулировали; вирусный геном пула клонировали посредством гомологичной рекомбинации в клетки E. coli BJ5183 с использованием челнока ВАС для создания минибиблиотеки векторов, несущих делецию области E1. Минибиблиотекой векторов ΔE1 трансфицировали клетки линии Procell 92; регенерированные векторы последовательно пассировали в течение 16 пассажей. На 16 пассаже из амплифицированного вектора получали ДНК вируса и клонировали посредством гомологичной рекомбинации в клетки E. coli BJ5183 с использованием челночной плазмиды. Преобладающий вид вектора идентифицировали как вектор ChAd157ΔE1 и затем модифицировали для включения в остоу вектора следующих дополнительных модификаций:

- а) делеция области E4 вируса ΔE1 (от 34413 п.н. до 37127 п.н.) ;
- б) вставка E4orf6, имеющей происхождение из человеческого Ad5.

1.1: Создание минибиблиотеки ΔE1

Для получения пулированного вирусного генома использовали пул 29 вирусов дикого типа. Пулированный вирусный геном клонировали в вектор ВАС посредством гомологичной рекомбинации в клетках E. coli штамма BJ5183, ко-трансформированных пулированной вирусной ДНК и челночным вектором ВАС Subgroup C (#1365) (SEQ ID NO: 14). Как показано на схеме на Фиг. 2, челночный вектор ВАС Subgroup C представляет собой вектор ВАС, предназначенный для клонирования ChAd, относящихся к виду C, и, следовательно, содержит ген pIX и фрагменты ДНК, имеющие происхождение из правых и левых концов (включая правые и левые ITR) вирусов ChAd вида C.

Челночная ВАС вида C также содержит кассету RpsL-Kana, встроенную между левым концом и геном pIX. Кроме того, между геном pIX и правым концом вирусного генома находится селекционная кассета Amp-LacZ-SacB, фланкированная сайтами рестрикции IScel. В частности, челночный вектор ВАС имеет следующие характеристики: Левый ITR: от 27 до 139 п.н., кассета hCMV(tetO) RpsL-Kana: от 493 до 3396 п.н., ген pIX: от 3508 до 3972 п.н., сайты рестрикции IScel: 3990 и 7481 п.н., селекционная кассета Amp-LacZ-SacB: от 4000 до 7471 п.н., правый ITR: от 7805 до 7917 п.н.. hCMV(tetO) представлен в SEQ ID NO: 37.

Клетки BJ5183 ко-трансформировали путем электропорации пулом очищенных вирусных ДНК и челночным вектором Subgroup C ВАС, разрезанным рестрикционным ферментом IScel и затем выделенным из геля. Гомологичная рекомбинация, происходящая между геном pIX и последовательностями правого ITR (находящимися на концах линеаризованной ДНК челночного вектора Species C ВАС) и гомологичными последовательностями, присутствующими в пулированной вирусной ДНК, приводила к встраиванию различных вирусных геномных ДНК в челночный вектор ВАС. В то же время вирусные области E1 были делетированы и замещены кассетой RpsL-Kana, с образованием ВАС/минибиблиотека ΔE1/TetO hCMV RpsL-Kana.

1.2: Амплификация минибиблиотеки ΔE1 в клетках линии Procell 92 и клонирование вектора ChAd157ΔE1.

Минибibliothek ΔE1 разрезали посредством PmeI и использовали для трансфекции пакующей клеточной линии Procell 92, чтобы собрать библиотеку различных вирусов в совокупности. Через 10 суток после трансфекции клетки собирали и клеточный лизат подвергали трем циклам замораживания (-70°C) и оттаивания (+37°C), просветляли центрифугированием при 2000 об/мин и использовали для инфицирования свежих клеток. Для отбора видов вируса по эффективности размножения в клетках Procell92 проводили 16 последовательных пассажей амплификации вируса. Вирус(ы) на 16 пассаже очищали посредством двух центрифугирований в градиенте CsCl и выделяли вирусную ДНК и клонировали посредством гомологичной рекомбинации в клетки E. coli BJ5183 с применением челночной плазмиды. Подробнее, клетки BJ5183 ко-трансформировали очищенной вирусной ДНК и челночной плазмидой Subgroup C (SEQ ID NO: 38). Как показано на схеме на Фиг. 3, челночная плазида Subgroup C представляет собой плазмидный вектор, предназначенный для клонирования ChAd, относящихся к виду C, и, следовательно, содержит фрагменты ДНК, имеющие происхождение из правых и левых концов (включая правые и левые ITR) вирусов ChAd вида C.

Гомологичная рекомбинация между последовательностями правых и левых ITR, присутствующих на концах линейаризованной челночной плазмиды Subgroup C (разрезанной PshAI/NdeI/XbaI), и вирусными геномными ДНК позволяла встроить их в плазмидный вектор. Амплифицировали 30 различных клонов, проанализировали с помощью рестрикционного анализа и идентифицировали 9 различных видов. 19/30 клонов демонстрировали одинаковые паттерны рестрикции и представляли преобладающие виды; отобрали один из этих клонов и обозначили его pChAd157ΔE1 TetO hCMV RpsL-Kana#1551 (SEQ ID NO: 15).

1.3: Конструирование ChAd157 ΔE1/TetO hCMV GAG#1557

Кассету GAG (полинуклеотидная последовательность GAG SEQ ID NO: 16) клонировали в линейаризованный пре-адено акцепторный вектор между последовательностями промотора HCMV и BGH polyA (SEQ ID NO: 39) посредством гомологичной рекомбинации в E.coli, основанной на существующей гомологии.

Плазмиду pARS CV32TetOhCMV GAG расщепляли посредством SpeI и SphI чтобы вырезать фрагмент 2,44 кб, содержащий промотор HCMV с tetO, HIV-GAG и последовательностью BGH polyA.

Фрагмент HIV-GAG 2.44 кб клонировали посредством гомологичной рекомбинации в акцепторный вектор pChAd157 ΔE1 /TetO hCMV RpsL-Kana (#1551) (разрезанный посредством SnaBI), несущий селекционную кассету RpsL-Kana под контролем HCMV и BGHPA. Полученную конструкцию обозначали как вектор pChAd157 ΔE1/TetO hCMV GAG#1557 (SEQ ID NO: 17).

Структура плазмиды, несущей ChAd157 GAG, показана на Фиг. 4.

1.4: Конструирование ChAd157 ΔE1E4 Ad5E4orf6/TetO hCMV RpsL-Kana#1594

Затем вектор ChAd157ΔE1 модифицировали, чтобы его остов имел следующие модификации:

- а) делеция области E4 вируса ΔE1 (от 34413 п.н. до 37127 п.н.) ;
- б) вставка E4orf6, имеющей происхождение из человеческого Ad5.

Осуществляли делецию области E4, расположенную между нуклеотидами 34413-37127 (координаты последовательности вектора ΔE1) путем замещения нативной области E4 последовательностью, кодирующей Ad5 E4orf6, применяя стратегию, включающую несколько стадий клонирования и гомологичной рекомбинации в E.coli. Осуществили делецию всей кодирующей области E4, при этом нативный промотор E4 и сигнал полиаденилирования сохранили. Для этого сконструировали челночный вектор, позволяющий осуществить вставку Ad5orf6 путем замены нативной области E4 ChAd157 в

результате гомологичной рекомбинации в E.coli BJ5183, как подробно описано ниже.

- Конструирование pARS SpeciesC Ad5E4orf6-1:

Фрагмент ДНК, содержащий Ad5orf6, получили с помощью PCR, используя ДНК Ad5 в качестве матрицы и олигонуклеотиды: 5'-ATACGGACTAGTGGAGAAGTACTCGCCTACATG-3' (SEQ ID NO: 18) и 5'-ATACGGAAGATCTAAGACTTCAGGAAATATGACTAC-3' (SEQ ID NO: 19). Фрагмент PCR разрезали посредством BglII и SpeI и клонировали в челночный вектор pARS Species C RLD-EGFP, разрезанный посредством BglII и SpeI, с получением плазмиды pARS Species C Ad5orf6-1.

- Конструирование pARS SpeciesC Ad5E4orf6-2:

Фрагмент ДНК 144 п.н., содержащий Fiber-E4 polyA (от 34269 п.н. до 34412 п.н. вектора ChAd157ΔE1) амплифицировали при помощи PCR с использованием плазмиды pChAd157 ΔE1 /TetO hCMV RpsL-Kana (#1551) в качестве матрицы со следующими олигонуклеотидами: 5'-ATTCACTGTACAGGCGCGCCAAAGCATGACACTGATGTTTCATTTTC-3' (SEQ ID NO: 20) и 5'-ACTAGGACTAGTTATAAGCTAGAATGGGGCTTTGC-3' (SEQ ID NO: 21). PCR-фрагмент разрезали посредством BsrGI и SpeI и клонировали в pARS SubGroupC Ad5orf6-1, разрезанный посредством BsrGI и SpeI, с получением плазмиды pARS SpeciesC Ad5orf6-2 (SEQ ID NO: 40).

Полученную плазмиду pARS SpeciesC Ad5orf6-2 затем использовали для замещения области E4 в составе остова ChAd157 на Ad5orf6. Для этого плазмиду pChAd157ΔE1 TetO hCMV RpsL-Kana#1551 разрезали посредством PacI и ко-трансформировали клетки BJ5183 плазмидой pARS SpeciesC Ad5orf6-2, разрезанной посредством BamHI/AscI, для получения преадено плазмиды pChAd157 ΔE1E4_Ad5E4orf6/TetO hCMV RpsL-Kana (#1594) (SEQ ID NO: 22).

1.5: Конструирование ChAd157 ΔE1E4_Ad5E4orf6/TetO hCMV RG#1559

Кассету, экспрессирующую гликопротеин вируса бешенства (RG) (полинуклеотидная последовательность гликопротеина вируса бешенства SEQ ID NO: 23) клонировали в линейаризованный преадено акцепторный вектор между последовательностями промотора HCMV и BGH polyA посредством гомологичной рекомбинации в E.coli, основанной на существующей гомологии.

Плазмиду pvjTetOhCMV-bghpolyA_RG расщепляли посредством SpeI и AsiSI, чтобы вырезать фрагмент 2,59 кб, содержащий промотор HCMV с tetO, RG и последовательностью BGH polyA.

Полученный фрагмент RG 2,59 т.п.н. клонировали посредством гомологичной рекомбинации в акцепторный вектор pChAd157 ΔE1E4_Ad5E4orf6/TetO hCMV RpsL-Kana (#1594), несущий селекционную кассету RpsL-Kana под контролем HCMV и BGHpA. Акцепторную преадено плазмиду линейаризовали с помощью рестрикционной эндонуклеазы SnaBI. Полученную конструкцию обозначали как вектор pChAd157 ΔE1E4_Ad5E4orf6/TetO hCMV RG#1559 (SEQ ID NO: 24).

Структура плазмиды, несущей ChAd157 RG, показана на Фиг. 6.

ПРИМЕР 2. Получение вектора

Продуктивность в отношении ChAd157 оценивали в сравнении с ChAd19 и ChAd155 в клетках линии Procell 92.

2.1. Получение векторов, содержащих трансген HIV Gag

ChAd157/GAG, ChAd19/GAG, ChAd155/GAG (векторы ChAd157, ChAd19 и ChAd155, экспрессирующие трансген HIV Gag) регенерировали и амплифицировали в Procell 92; лизаты использовали для инфицирования каждым вектором клеток Procell 92, культивируемых в виде монослоя в 1 флаконе T25. Использовали множественность заражения (MOI) 300 вч/клетку (вирусных частицах на клетку) и

осуществляли заражение в присутствии тетрациклина, поскольку у ChAd19/GAG отсутствовал транскрипционный контроль, опосредованный вставкой оператора TetO в промотор hCMV. Инфицированные клетки собирали, когда становился очевидным полный цитопатогенный эффект (через 48 часов после инфицирования ChAd157/GAG и ChAd155/GAG и через 5 суток после инфицирования ChAd19/GAG); для высвобождения вирусов из инфицированных клеток осуществляли 3 цикла замораживания/оттаивания (от -70° до 37°C), после чего лизат просветляли путем центрифугирования. Проводили количественный анализ просветленных лизатов при помощи количественной PCR с праймерами и зондом, комплементарными области промотора CMV. Последовательности олигонуклеотидов были следующими: CMVfor 5'-CATCTACGTATTAGTCATCGCTATTACCA-3' (SEQ ID NO: 25), CMVrev 5'-GACTTGGAATCCCCGTGAGT-3' (SEQ ID NO: 26), зонд CMVFAM-TAMRA 5'-ACATCAATGGGCGTGGATAGCGGTT-3' (SEQ ID NO: 41) (QPCR выполняли с использованием прибора ABI Prism 7900 Sequence detector – Applied Biosystem).

Итоговые титры, выраженные как количество вирусных частиц на единицу объема (вч/мл), которые определяли в просветленных лизатах, и специфическая продуктивность, выраженная как количество вирусных частиц на клетку (вч/клетку), представлены в Таблице 1 ниже.

Таблица 1. Продуктивность вектора GAG

Вектор	Продуктивность на единицу объема (вч/мл)	Всего вирусных частиц	Специфическая продуктивность (вч/клетку)
ChAd157/GAG	4,61E+09	2,30E+10	7,68E+03
ChAd155/GAG	5,42E+09	2,71E+10	9,04E+03
ChAd19/GAG	4,80E+08	2,40E+09	8,00E+02

2.2. Получение векторов, содержащих трансген RG

Для оценки продуктивности вакцинных векторов RG в клетках Procell 92, культивируемых в суспензии, провели другой ряд экспериментов. В эксперименте проводили параллельное сравнение ChAd157/RG и ChAd155/RG путем инфицирования Procell 92 при плотности клеток 5×10^5 клеток/мл. Использовали множественность заражения (MOI) 300 вч/клетку. Инфицированные клетки собирали через 4 дня после инфицирования; для высвобождения вируса из инфицированных клеток осуществляли 3 цикла замораживания/оттаивания и просветляли лизат путем центрифугирования. Затем проводили количественный анализ просветленных лизатов при помощи количественной PCR, как описано выше.

В Таблице 2 ниже представлена продуктивность на единицу объема и специфическая продуктивность на клетку.

Таблица 2. Продуктивность вектора RG

Вектор	Продуктивность на единицу объема (вч/мл)	Всего вирусных частиц	Специфическая продуктивность (вч/клетку)
ChAd157/RG	9,39E+09	4,69E+11	1,88E+04

ChAd155/RG	1,41E+10	7,04E+11	2,81E+04
-------------------	----------	----------	----------

Пример 3. Уровни экспрессии трансгенов

3.1. Уровень экспрессии трансгена HIV Gag

Уровни экспрессии сравнивали в параллельных экспериментах путем инфицирования клеток HeLa векторами ChAd19, ChAd155 и ChAd157, содержащими трансген HIV Gag.

Клетки HeLa высевали в 35 мм чашки и инфицировали очищенными вирусами ChAd19/GAG, ChAd157/GAG и ChAd155/GAG, используя MOI, составляющую 250 вч/клетку. Супернатанты инфицированных клеток HeLa собирали через 48 часов после инфицирования и определяли количество секретируемого белка HIV GAG с использованием коммерческого набора для ELISA (HIV-1 p24 ELISA Kit, PerkinElmer Life Science). Количественное определение проводили в соответствии с инструкциями производителя, используя стандартную кривую антигена HIV-1 p24.

Результаты, выраженные в пг/мл белка GAG, представлены на Фиг. 7.

3.2. Уровень экспрессии трансгена RG

Также проводили анализ методом вестерн-блоттинг для оценки экспрессии гликопротеина вируса бешенства, обеспечиваемой вектором ChAd157/RG, в сравнении с вектором ChAd155/RG. Для этого клетки HeLa высевали в 35 мм чашки и инфицировали очищенными вирусами ChAd157/RG и ChAd155/RG, используя MOI, составляющую 250 вч/клетку. Клеточные лизаты собирали через 48 часов после инфицирования и оценивали уровень экспрессии трансгена при помощи SDS-PAGE (электрофорез в полиакриламидном геле в присутствии додецилсульфата натрия) в восстанавливающих условиях, после чего проводили анализ вестерн-блот.

В гель вносили эквивалентное количество белковых экстрактов; после электрофоретического разделения в присутствии SDS в восстанавливающих условиях белки переносили на нитроцеллюлозную мембрану для инкубации с поликлональным антителом кролика к GP (кат. номер: RBVGP11-S α Diagnostic, в разведении 1:1000). После инкубации с первичным антителом мембрану отмывали и затем инкубировали со вторичным антителом к иммуноглобулинам кролика, конъюгированным с пероксидазой хрена (HRP). В конце анализа применяли стандартный электрохемилюминесцентный метод (реагент для выявления ECL (W3252282 PIERCE)). Результаты вестерн-блота показаны на Фиг. 8.

Полоса приблизительно 57 кДа, указанная стрелкой, была выявлена с помощью поликлонального антитела к GP, что соответствовало ожидаемой массе гликопротеина вируса бешенства.

Результат свидетельствует, что уровень экспрессии, обеспечиваемый вектором ChAd157, сопоставим с таковым для вектора ChAd155.

Пример 4. Оценка иммуногенности в экспериментах с иммунизацией мышей

4.1. Иммуногенность векторов, содержащих трансген HIV Gag

Иммуногенность вектора ChAd157/GAG оценивали параллельно с ChAd155/GAG и ChAd19/GAG на мышах BALB/c (по 6 в группе). Для проведения эксперимента осуществляли внутримышечные инъекции 10^7 вирусных частиц. Т-клеточный ответ оценивали через 3 недели после иммунизации *ex vivo*, используя иммуноферментный спот-анализ интерферон- γ (IFN- γ) (ELISpot) с использованием эпитопа GAG для CD8+ Т клеток, картированного у мышей BALB/c. Полученные результаты, выраженные как количество IFN- γ -секретирующих клеток, образующих пятна (SFC), на миллион спленоцитов, показаны на Фиг. 9.

Каждый кружок соответствует ответу одной мыши, а линия соответствует среднему геометрическому значению для каждой дозовой группы. Доля мышей, положительных в отношении CD8-иммунодоминантного пептида, показана на оси х.

4.2 Иммуногенность векторов, содержащих трансген RG

Иммуногенность векторов ChAd157/RG и ChAd155/RG оценивали на мышах BALB/c. Осуществляли внутримышечные инъекции обоих векторов в дозах 10^7 и 10^6 вч. Через семь недель после вакцинации выделяли спленоциты иммунизированных мышей и анализировали методом IFN γ ELISpot (Фиг. 10), используя в качестве антигена пулы пептидов RG.

Как и ожидалось, с уменьшением дозы уровни иммунного ответа снижались. Более того, вектор ChAd155RG индуцировал более выраженный Т-клеточный ответ по сравнению с ChAd157RG, хотя различие не было статистически достоверным (Фиг. 10).

Пример 5. Оценка инфекционности

5.1. Инфекционность векторов, содержащих трансген HIV Gag

Инфекционность вирусов оценивали в адгезированных клетках Procell 92, используя антитело против белка гексон аденовируса для визуализации инфицированных клеток при помощи иммуноцитохимического окрашивания. Антитело к белку гексону распознает все серотипы аденовируса. Для этого клетки Procell92 высевали в 24-луночный планшет при плотности 2×10^5 жизнеспособных клеток/мл и инфицировали в двойной повторности векторами ChAd157/GAG и ChAd155/GAG и ChAd19/GAG, используя MOI, составляющую 1 вч/клетку, 0,5 вч/клетку и 0,25 вч/клетку. Через 48 часов после инфицирования инфицированные клетки фиксировали холодным метанолом и затем метили антителом к гексону. Излишек антитела удаляли. Затем меченные клетки инкубировали со вторичным антителом, конъюгированным с пероксидазой хрена и выявляли с использованием в качестве субстрата коммерческого набора VECTOR NOVARED (SK-4800). Выявление основано на реакции ферментной метки, пероксидазы хрена, с субстратом DAB, в результате чего образуется темно-коричневый продукт. Затем проводили количественное определение меченных темно-коричневых клеток при помощи световой микроскопии и рассчитывали инфекционный титр. Результаты представлены в таблице, приведенной ниже.

Вirus	Вirusных частиц/мл	Инфекционных единиц/мл	R (вч/инф. ед.)
ChAd155 GAG	1,32E+11	1,58E+09	84
ChAd157 GAG	1,17E+11	1,23E+09	95
ChAd19 GAG	4,46E+10	3,86E+08	116

Результаты свидетельствуют, что инфекционность вирусов ChAd155 и ChAd157 сопоставима и превышает таковую для ChAd19.

5.2 Инфекционность векторов, содержащих трансген RG

Инфекционность очищенных вирусов ChAd157/RG и ChAd155/RG оценивали в адгезированных клетках Procell 92 при помощи иммуноцитохимического окрашивания, как описано выше. Результаты представлены в таблице, приведенной ниже.

Вirus	Вirusных частиц/мл	Инфекционных единиц/мл	R (вч/инф. ед.)
-------	-----------------------	---------------------------	--------------------

ChAd155/RG	4,23E+11	4,06E+09	104
ChAd157/RG	1,97E+11	1,46E+09	133

Результаты свидетельствуют, что инфекционность вирусов ChAd155 и ChAd157 сопоставима.

Пример 6. Оценка перекрестной нейтрализации между векторами ChAd155 и ChAd157

6.1. Исследование in vivo того, имеют ли векторы ChAd155 и ChAd157 разный серотип

Перекрестную нейтрализацию между векторами ChAd155 и ChAd157 оценивали на мышах BALB/c (6 на группу). Мышей дважды предварительно иммунизировали на 0 и 3 неделе 10^9 вч ChAd155 или ChAd157, экспрессирующими RG, или ложно вакцинировали забуференным физиологическим раствором. Три недели спустя всех мышей иммунизировали однократно 10^9 вч ChAd157, кодирующим HIV gag.

Группы	n	Двукратная преиммунизация на 0 неделе и 3 неделе	доза (вч)	Иммунизация 6 неделя	доза (вч)
1	6	PBS (забуференный фосфатом физиологический раствор)	-	ChAd157-GAG	10^9
2	6	ChAd155-RG	10^9	ChAd157-GAG	10^9
3	6	ChAd157-RG	10^9	ChAd157-GAG	10^9

На 5 неделе (через 2 недели после второй инъекции) в сыворотках определяли нейтрализующие титры к векторам, использованным для предварительной иммунизации, при помощи теста нейтрализации in vitro (Фиг. 11). Наконец, через 3 недели после иммунизации оценивали Т-клеточный ответ против gag на спленоцитах при помощи IFN- γ ELISpot с использованием эпитопа GAG для CD8+ Т клеток, картированного у мышей BALB/c (Фиг. 12). Дозы векторов, использованные для предварительной иммунизации, были способны вызывать хорошую нейтрализующую активность против двух Ad(аденовирусных)-векторов, несмотря на наличие некоторой вариабельности. Нейтрализующие антитела к ChAd155 не реагируют перекрестно с ChAd157 и наоборот (Фиг. 11). Более того, предсуществующий иммунитет к ChAd155 не влиял на ChAd157-Gag Т-клеточный ответ, подтверждая отсутствие перекрестной нейтрализации (Фиг. 12).

Взятые вместе, эти данные позволяют предположить, что ChAd155 и ChAd157 представляют собой аденовирусы разных серотипов.

SEQUENCE LISTING

<110> GlaxoSmithKline Biologicals S.A.

<120> ADENOVIRUS POLYNUCLEOTIDES AND POLYPEPTIDES

<130> VU66143

<150> GB1620968.6

<151> 2016-12-09

<160> 42

<170> PatentIn version 3.5

<210> 1

<211> 543

<212> PRT

<213> Pan troglodytes

<400> 1

Met	Lys	Arg	Thr	Lys	Thr	Ser	Asp	Glu	Ser	Phe	Asn	Pro	Val	Tyr	Pro
1				5					10					15	

Tyr	Asp	Thr	Glu	Ser	Gly	Pro	Pro	Ser	Val	Pro	Phe	Leu	Thr	Pro	Pro
			20					25					30		

Phe	Val	Ser	Pro	Asp	Gly	Phe	Gln	Glu	Ser	Pro	Pro	Gly	Val	Leu	Ser
		35					40					45			

Leu	Asn	Leu	Ala	Glu	Pro	Leu	Val	Thr	Ser	His	Gly	Met	Leu	Ala	Leu
	50						55				60				

Lys	Met	Gly	Ser	Gly	Leu	Ser	Leu	Asp	Asp	Ala	Gly	Asn	Leu	Thr	Ser
65					70					75					80

Gln	Asp	Ile	Thr	Thr	Ala	Ser	Pro	Pro	Leu	Lys	Lys	Thr	Lys	Thr	Asn
				85					90						95

Leu	Ser	Leu	Glu	Thr	Ser	Ser	Pro	Leu	Thr	Val	Ser	Thr	Ser	Gly	Ala
			100					105					110		

Leu	Thr	Val	Ala	Ala	Ala	Ala	Pro	Leu	Ala	Val	Ala	Gly	Thr	Ser	Leu
		115					120					125			

Thr	Met	Gln	Ser	Glu	Ala	Pro	Leu	Thr	Val	Gln	Asp	Ala	Lys	Leu	Thr
	130					135					140				

Leu	Ala	Thr	Lys	Gly	Pro	Leu	Thr	Val	Ser	Glu	Gly	Lys	Leu	Ala	Leu
145					150					155					160

Gln	Thr	Ser	Ala	Pro	Leu	Thr	Ala	Ala	Asp	Ser	Ser	Thr	Leu	Thr	Val
				165					170					175	

Ser	Ala	Thr	Pro	Pro	Ile	Asn	Val	Ser	Ser	Gly	Ser	Leu	Gly	Leu	Asp	180	185	190
Met	Glu	Asp	Pro	Met	Tyr	Thr	His	Asn	Gly	Lys	Leu	Gly	Ile	Arg	Ile	195	200	205
Gly	Gly	Pro	Leu	Arg	Val	Val	Asp	Ser	Leu	His	Thr	Leu	Thr	Val	Val	210	215	220
Thr	Gly	Asn	Gly	Leu	Thr	Val	Asp	Asn	Asn	Ala	Leu	Gln	Thr	Lys	Val	225	230	235
Thr	Gly	Ala	Leu	Gly	Tyr	Asp	Thr	Ser	Gly	Asn	Leu	Gln	Leu	Arg	Ala	245	250	255
Ala	Gly	Gly	Met	Arg	Ile	Asp	Ala	Asn	Gly	Gln	Leu	Ile	Leu	Asn	Val	260	265	270
Ala	Tyr	Pro	Phe	Asp	Ala	Gln	Asn	Asn	Leu	Ser	Leu	Arg	Leu	Gly	Gln	275	280	285
Gly	Pro	Leu	Tyr	Ile	Asn	Thr	Asp	His	Asn	Leu	Asp	Leu	Asn	Cys	Asn	290	295	300
Arg	Gly	Leu	Thr	Thr	Thr	Thr	Thr	Asn	Asn	Thr	Lys	Lys	Leu	Glu	Thr	305	310	315
Lys	Ile	Ser	Ser	Gly	Leu	Asp	Tyr	Asp	Thr	Asn	Gly	Ala	Val	Ile	Ile	325	330	335
Lys	Leu	Gly	Thr	Gly	Leu	Ser	Phe	Asp	Asn	Thr	Gly	Ala	Leu	Thr	Val	340	345	350
Gly	Asn	Thr	Gly	Asp	Asp	Lys	Leu	Thr	Leu	Trp	Thr	Thr	Pro	Asp	Pro	355	360	365
Ser	Pro	Asn	Cys	Arg	Ile	His	Ser	Asp	Lys	Asp	Cys	Lys	Phe	Thr	Leu	370	375	380
Val	Leu	Thr	Lys	Cys	Gly	Ser	Gln	Ile	Leu	Ala	Ser	Val	Ala	Ala	Leu	385	390	395
Ala	Val	Ser	Gly	Asn	Leu	Ala	Ser	Ile	Thr	Gly	Thr	Val	Ala	Ser	Val	405	410	415
Thr	Ile	Phe	Leu	Arg	Phe	Asp	Gln	Asn	Gly	Val	Leu	Met	Glu	Asn	Ser	420	425	430

Ser Leu Asp Arg Gln Tyr Trp Asn Phe Arg Asn Gly Asn Ser Thr Asn
435 440 445

Ala Ala Pro Tyr Thr Asn Ala Val Gly Phe Met Pro Asn Leu Ala Ala
450 455 460

Tyr Pro Lys Thr Gln Ser Gln Thr Ala Lys Asn Asn Ile Val Ser Gln
465 470 475 480

Val Tyr Leu Asn Gly Asp Lys Ser Lys Pro Met Thr Leu Thr Ile Thr
485 490 495

Leu Asn Gly Thr Asn Glu Ser Ser Glu Thr Ser Gln Val Ser His Tyr
500 505 510

Ser Met Ser Phe Thr Trp Ala Trp Glu Ser Gly Gln Tyr Ala Thr Glu
515 520 525

Thr Phe Ala Thr Asn Ser Phe Thr Phe Ser Tyr Ile Ala Glu Gln
530 535 540

<210> 2
<211> 1629
<212> DNA
<213> Pan troglodytes

<400> 2
atgaagcgca ccaaaacgtc tgacgagagc ttcaaccccg tgtacccta tgacacggaa 60
agcggccctc cctccgtccc tttcctcacc cctcccttcg tgtctcccga tggattccaa 120
gaaagtcccc ccgggggtcct gtctctgaac ctggccgagc ccctgggtcac ttcccacggc 180
atgctcgccc tgaaaatggg aagtggcctc tccctggacg acgctggcaa cctcacctct 240
caagatatca ccaccgctag ccctccctc aaaaaaacca agaccaacct cagcctagaa 300
acctcatccc ccctaactgt gagcacctca ggcgcctca ccgtagcagc cgccgctccc 360
ctggcggtgg ccggcacctc cctcaccatg caatcagagg cccccctgac agtacaggat 420
gcaaaactca ccctggccac caaaggcccc ctgaccgtgt ctgaaggcaa actggccttg 480
caaacatcgg ccccgctgac ggccgctgac agcagcacc tcaccgtag cgccacacca 540
ccaattaatg taagcagtgg aagtttaggc ttagacatgg aagaccctat gtatactcac 600
aatggaaaac tgggaataag aattgggggt ccactaagag tagtagacag cttgcataca 660
ctcactgtag ttaccgaaa tggactaact gtagataaca atgccctcca aactaaagtt 720
acgggcgccc taggttatga cacatcagga aatctacaat taagagctgc aggaggtatg 780
cgaattgacg caaatggcca acttatcctt aatgtggcat acccatttga tgctcagaac 840

```

aatctcagcc ttagacttgg tcagggaccc ctgtatataa acacagacca caacctggat      900
ttgaattgca acagaggtct aaccacaact accaccaaca acacaaaaaa acttgagact      960
aaaattagct caggcttaga ctatgacacc aatggtgctg tcattattaa acttggcact     1020
ggctctaagct tcgacaacac aggcgcacct actgtgggaa aacttggtga tgataaactg     1080
actctgtgga cgaccccgaga cccatctcca aattgcagaa ttcactcaga caaagactgc     1140
aagtttactc tagtcctaac taagtgtgga agccaaatcc tggcctctgt cgccgcacct     1200
gcggtatcag gaaatctggc ttcgataaca ggcaccgttg ccagcgttac catctttctc     1260
agatttgatc agaatggagt gcttatggaa aactcctcgc tagacaggca gtactggaac     1320
ttcagaaatg gcaactcaac taacgctgcc ccctacacca atgcagttgg gttcatgcca     1380
aacctcgcag cataccccaa aacgcaaagc cagactgcta aaaacaacat tgtaagtcag     1440
gtttacttga atggagacaa atccaaaccc atgaccctta ccatcaccct caatggaact     1500
aatgaatcca gtgaaactag ccaggtgagt cactactcca tgtcatttac atgggcttgg     1560
gaaagtgggc aatatgccac tgaaaccttt gccaccaact ccttcacctt ttcttacatt     1620
gctgaacaa                                     1629

```

```

<210> 3
<211> 593
<212> PRT
<213> Pan troglodytes

<400> 3

```

```

Met Arg Arg Ala Ala Met Tyr Gln Glu Gly Pro Pro Pro Ser Tyr Glu
1           5           10           15

```

```

Ser Val Val Gly Ala Ala Ala Ala Ala Pro Ser Ser Pro Phe Ala Ser
          20           25           30

```

```

Gln Leu Leu Glu Pro Pro Tyr Val Pro Pro Arg Tyr Leu Arg Pro Thr
          35           40           45

```

```

Gly Gly Arg Asn Ser Ile Arg Tyr Ser Glu Leu Ala Pro Leu Phe Asp
50           55           60

```

```

Thr Thr Arg Val Tyr Leu Val Asp Asn Lys Ser Ala Asp Val Ala Ser
65           70           75           80

```

```

Leu Asn Tyr Gln Asn Asp His Ser Asn Phe Leu Thr Thr Val Ile Gln
          85           90           95

```

```

Asn Asn Asp Tyr Ser Pro Ser Glu Ala Ser Thr Gln Thr Ile Asn Leu
100           105           110

```

Asp	Asp	Arg	Ser	His	Trp	Gly	Gly	Asp	Leu	Lys	Thr	Ile	Leu	His	Thr	115	120	125
Asn	Met	Pro	Asn	Val	Asn	Glu	Phe	Met	Phe	Thr	Asn	Lys	Phe	Lys	Ala	130	135	140
Arg	Val	Met	Val	Ser	Arg	Ser	His	Thr	Lys	Glu	Asp	Arg	Val	Glu	Leu	145	150	155
Lys	Tyr	Glu	Trp	Val	Glu	Phe	Glu	Leu	Pro	Glu	Gly	Asn	Tyr	Ser	Glu	165	170	175
Thr	Met	Thr	Ile	Asp	Leu	Met	Asn	Asn	Ala	Ile	Val	Glu	His	Tyr	Leu	180	185	190
Lys	Val	Gly	Arg	Gln	Asn	Gly	Val	Leu	Glu	Ser	Asp	Ile	Gly	Val	Lys	195	200	205
Phe	Asp	Thr	Arg	Asn	Phe	Arg	Leu	Gly	Leu	Asp	Pro	Val	Thr	Gly	Leu	210	215	220
Val	Met	Pro	Gly	Val	Tyr	Thr	Asn	Glu	Ala	Phe	His	Pro	Asp	Ile	Ile	225	230	235
Leu	Leu	Pro	Gly	Cys	Gly	Val	Asp	Phe	Thr	Tyr	Ser	Arg	Leu	Ser	Asn	245	250	255
Leu	Leu	Gly	Ile	Arg	Lys	Arg	Gln	Pro	Phe	Gln	Glu	Gly	Phe	Arg	Ile	260	265	270
Thr	Tyr	Glu	Asp	Leu	Glu	Gly	Gly	Asn	Ile	Pro	Ala	Leu	Leu	Asp	Val	275	280	285
Glu	Ala	Tyr	Gln	Asp	Ser	Leu	Lys	Glu	Asn	Glu	Ala	Gly	Gln	Glu	Asp	290	295	300
Thr	Ala	Pro	Ala	Ala	Ser	Ala	Ala	Ala	Glu	Gln	Gly	Glu	Asp	Ala	Ala	305	310	315
Asp	Thr	Ala	Ala	Ala	Asp	Gly	Ala	Glu	Ala	Asp	Pro	Ala	Met	Val	Val	325	330	335
Glu	Ala	Ala	Glu	Gln	Glu	Glu	Asp	Met	Asn	Asp	Ser	Ala	Val	Arg	Gly	340	345	350
Asp	Thr	Phe	Val	Thr	Arg	Gly	Glu	Glu	Lys	Gln	Ala	Glu	Ala	Glu	Ala	355	360	365

Ala Ala Glu Glu Lys Gln Leu Ala Ala Ala Ala Ala Ala Ala Leu
370 375 380

Ala Ala Ala Glu Ala Glu Ser Glu Gly Thr Lys Pro Ala Lys Glu Pro
385 390 395 400

Val Ile Lys Pro Leu Thr Glu Asp Ser Lys Lys Arg Ser Tyr Asn Leu
405 410 415

Leu Lys Asp Ser Thr Asn Thr Ala Tyr Arg Ser Trp Tyr Leu Ala Tyr
420 425 430

Asn Tyr Gly Asp Pro Ser Thr Gly Val Arg Ser Trp Thr Leu Leu Cys
435 440 445

Thr Pro Asp Val Thr Cys Gly Ser Glu Gln Val Tyr Trp Ser Leu Pro
450 455 460

Asp Met Met Gln Asp Pro Val Thr Phe Arg Ser Thr Arg Gln Val Ser
465 470 475 480

Asn Phe Pro Val Val Gly Ala Glu Leu Leu Pro Val His Ser Lys Ser
485 490 495

Phe Tyr Asn Asp Gln Ala Val Tyr Ser Gln Leu Ile Arg Gln Phe Thr
500 505 510

Ser Leu Thr His Val Phe Asn Arg Phe Pro Glu Asn Gln Ile Leu Ala
515 520 525

Arg Pro Pro Ala Pro Thr Ile Thr Thr Val Ser Glu Asn Val Pro Ala
530 535 540

Leu Thr Asp His Gly Thr Leu Pro Leu Arg Asn Ser Ile Gly Gly Val
545 550 555 560

Gln Arg Val Thr Val Thr Asp Ala Arg Arg Arg Thr Cys Pro Tyr Val
565 570 575

Tyr Lys Ala Leu Gly Ile Val Ser Pro Arg Val Leu Ser Ser Arg Thr
580 585 590

Phe

<210> 4
<211> 1779
<212> DNA
<213> Pan troglodytes

<400> 4

atgcgggcgcg	cggcgatgta	ccaggagggga	cctcctccct	cttacgagag	cgtgggtgggc	60
gcggcgggcgg	cggcgccttc	ttctcccttt	gcgtcgcagc	tgctggagcc	gccgtacgtg	120
cctccgcgct	acctgcggcc	tacggggggg	agaaacagca	tccgttactc	ggagctggcg	180
cccctgttcg	acaccacccg	ggtgtacctg	gtggacaaca	agtcggcgga	cgtggcctcc	240
ctgaactacc	agaacgacca	cagcaatttt	ttgaccacgg	tcattccagaa	caatgactac	300
agcccgagcg	aggccagcac	ccagaccatc	aatctggatg	accggtcgca	ctggggcggc	360
gacctgaaaa	ccatcctgca	caccaacatg	cccaacgtga	acgagttcat	gttcaccaat	420
aagttcaagg	cgcggggtgat	ggtgtcgcgc	tcgcacacca	aggaagaccg	ggtggagctg	480
aagtacgagt	gggtggagtt	cgagctgcca	gagggcaact	actccgagac	catgaccatt	540
gacctgatga	acaacgcgat	cgtggagcac	tatctgaaag	tgggcaggca	gaacggggtc	600
ctggagagcg	acatcggggg	caagtctgac	accaggaact	tccgcctggg	gctggacccc	660
gtgaccgggc	tggttatgcc	cgggggtgtac	accaacgagg	ccttccatcc	cgacatcatc	720
ctgctgcccg	gctgcggggg	ggacttcact	tacagccgcc	tgagcaacct	cctgggcatc	780
cgcaagcggc	agcccttcca	ggagggcctc	aggatcacct	acgaggacct	ggaggggggc	840
aacatccccg	cgctcctcga	tgtggaggcc	taccaggata	gcttgaagga	aaatgaggcg	900
ggacaggagg	ataccgcccc	cgccgcctcc	gccgccgccg	agcagggcga	ggatgctgct	960
gacaccgcgg	ccgcggacgg	ggcggaggcc	gaccccgcta	tggtggtgga	ggctgccgag	1020
caggaggagg	acatgaatga	cagtgcggtg	cgcgagagaca	ccttcgtcac	ccgggggggag	1080
gaaaagcaag	cggaggccga	ggccgcggcc	gaggaaaagc	aactggcggc	agcagcggcg	1140
gcggcggcgt	tggccgcggc	ggaggctgag	tctgagggga	ccaagcccgc	caaggagccc	1200
gtgattaagc	ccctgaccga	agatagcaag	aagcgcagtt	acaacctgct	caaggacagc	1260
accaacaccg	cgtaccgcag	ctggtacctg	gcctacaact	acggcgaccc	gtcgacgggg	1320
gtgcgctcct	ggaccctgct	gtgcacgccg	gacgtgacct	gcggtctgga	gcaggtgtac	1380
tggtcgctgc	ccgacatgat	gcaagacccc	gtgaccttcc	gctccacgcg	gcaggtcagc	1440
aacttccccg	tggtgggcgc	cgagctgctg	cccgtgcact	ccaagagctt	ctacaacgac	1500
caggccgtct	actcccagct	catccgccag	ttcacctctc	tgaccacagt	gttcaatcgc	1560
tttctgaga	accagattct	ggcgcgcccc	cccgccccca	ccatcaccac	cgtcagtgaa	1620
aacgttcctg	ctctcacaga	tcacgggacg	ctaccgctgc	gcaacagcat	cggaggagtc	1680
cagcgagtga	ccgttactga	cgccagacgc	cgcacctgcc	cctacgttta	caaggccttg	1740
ggcatagtct	cgccgcgcgt	cctttccagc	cgcactttt			1779

<210> 5

<211> 958
<212> PRT
<213> Pan troglodytes

<400> 5

Met Ala Thr Pro Ser Met Met Pro Gln Trp Ser Tyr Met His Ile Ser
1 5 10 15

Gly Gln Asp Ala Ser Glu Tyr Leu Ser Pro Gly Leu Val Gln Phe Ala
20 25 30

Arg Ala Thr Glu Ser Tyr Phe Ser Leu Ser Asn Lys Phe Arg Asn Pro
35 40 45

Thr Val Ala Pro Thr His Asp Val Thr Thr Asp Arg Ser Gln Arg Leu
50 55 60

Thr Leu Arg Phe Ile Pro Val Asp Arg Glu Asp Thr Ala Tyr Ser Tyr
65 70 75 80

Lys Ala Arg Phe Thr Leu Ala Val Gly Asp Asn Arg Val Leu Asp Met
85 90 95

Ala Ser Thr Tyr Phe Asp Ile Arg Gly Val Leu Asp Arg Gly Pro Thr
100 105 110

Phe Lys Pro Tyr Ser Gly Thr Ala Tyr Asn Ser Leu Ala Pro Lys Gly
115 120 125

Ala Pro Asn Ser Cys Glu Trp Glu Gln Leu Glu Glu Ala Gln Ala Ala
130 135 140

Val Glu Asp Glu Glu Leu Glu Asp Glu Asp Glu Glu Pro Gln Asp Glu
145 150 155 160

Ala Pro Val Lys Lys Thr His Val Tyr Ala Gln Ala Pro Leu Ser Gly
165 170 175

Glu Glu Ile Thr Lys Asn Gly Leu Gln Ile Gly Ser Asp Asn Thr Glu
180 185 190

Ala Gln Ser Lys Pro Ile Tyr Ala Asp Pro Thr Phe Gln Pro Glu Pro
195 200 205

Gln Ile Gly Glu Ser Gln Trp Asn Glu Ala Asp Ala Thr Val Ala Gly
210 215 220

Gly Arg Val Leu Lys Lys Ser Thr Pro Met Lys Pro Cys Tyr Gly Ser
225 230 235 240

Tyr	Ala	Arg	Pro	Thr	Asn	Ser	Asn	Gly	Gly	Gln	Gly	Val	Leu	Val	Ala	245	250	255
Asp	Asp	Lys	Gly	Val	Leu	Gln	Ser	Lys	Val	Glu	Leu	Gln	Phe	Phe	Ser	260	265	270
Asn	Thr	Thr	Thr	Leu	Asn	Gln	Arg	Glu	Gly	Asn	Asp	Thr	Lys	Pro	Lys	275	280	285
Val	Val	Leu	Tyr	Ser	Glu	Asp	Val	His	Met	Glu	Thr	Pro	Asp	Thr	His	290	295	300
Ile	Ser	Tyr	Lys	Pro	Thr	Lys	Ser	Asp	Asp	Asn	Ser	Lys	Ile	Met	Leu	305	310	315
Gly	Gln	Gln	Ser	Met	Pro	Asn	Arg	Pro	Asn	Tyr	Ile	Gly	Phe	Arg	Asp	325	330	335
Asn	Phe	Ile	Gly	Leu	Met	Tyr	Tyr	Asn	Ser	Thr	Gly	Asn	Met	Gly	Val	340	345	350
Leu	Ala	Gly	Gln	Ala	Ser	Gln	Leu	Asn	Ala	Val	Val	Asp	Leu	Gln	Asp	355	360	365
Arg	Asn	Thr	Glu	Leu	Ser	Tyr	Gln	Leu	Leu	Leu	Asp	Ser	Met	Gly	Asp	370	375	380
Arg	Thr	Arg	Tyr	Phe	Ser	Met	Trp	Asn	Gln	Ala	Val	Asp	Ser	Tyr	Asp	385	390	395
Pro	Asp	Val	Arg	Ile	Ile	Glu	Asn	His	Gly	Thr	Glu	Asp	Glu	Leu	Pro	405	410	415
Asn	Tyr	Cys	Phe	Pro	Leu	Gly	Gly	Ile	Gly	Val	Thr	Asp	Thr	Tyr	Gln	420	425	430
Ala	Ile	Lys	Thr	Asn	Gly	Asn	Gly	Gln	Glu	Asn	Pro	Thr	Trp	Glu	Lys	435	440	445
Asp	Thr	Glu	Phe	Ala	Asp	Arg	Asn	Glu	Ile	Gly	Val	Gly	Asn	Asn	Phe	450	455	460
Ala	Met	Glu	Ile	Asn	Leu	Ser	Ala	Asn	Leu	Trp	Arg	Asn	Phe	Leu	Tyr	465	470	475
Ser	Asn	Val	Ala	Leu	Tyr	Leu	Pro	Asp	Lys	Leu	Lys	Tyr	Asn	Pro	Ser	485	490	495

Asn Val Asp Ile Ser Asp Asn Pro Asn Thr Tyr Asp Tyr Met Asn Lys
500 505 510

Arg Val Val Ala Pro Gly Leu Val Asp Cys Tyr Ile Asn Leu Gly Ala
515 520 525

Arg Trp Ser Leu Asp Tyr Met Asp Asn Val Asn Pro Phe Asn His His
530 535 540

Arg Asn Ala Gly Leu Arg Tyr Arg Ser Met Leu Leu Gly Asn Gly Arg
545 550 555 560

Tyr Val Pro Phe His Ile Gln Val Pro Gln Lys Phe Phe Ala Ile Lys
565 570 575

Asn Leu Leu Leu Leu Pro Gly Ser Tyr Thr Tyr Glu Trp Asn Phe Arg
580 585 590

Lys Asp Val Asn Met Val Leu Gln Ser Ser Leu Gly Asn Asp Leu Arg
595 600 605

Val Asp Gly Ala Ser Ile Lys Phe Glu Ser Ile Cys Leu Tyr Ala Thr
610 615 620

Phe Phe Pro Met Ala His Asn Thr Ala Ser Thr Leu Glu Ala Met Leu
625 630 635 640

Arg Asn Asp Thr Asn Asp Gln Ser Phe Asn Asp Tyr Leu Ser Ala Ala
645 650 655

Asn Met Leu Tyr Pro Ile Pro Ala Asn Ala Thr Asn Val Pro Ile Ser
660 665 670

Ile Pro Ser Arg Asn Trp Ala Ala Phe Arg Gly Trp Ala Phe Thr Arg
675 680 685

Leu Lys Thr Lys Glu Thr Pro Ser Leu Gly Ser Gly Phe Asp Pro Tyr
690 695 700

Tyr Thr Tyr Ser Gly Ser Ile Pro Tyr Leu Asp Gly Thr Phe Tyr Leu
705 710 715 720

Asn His Thr Phe Lys Lys Val Ser Val Thr Phe Asp Ser Ser Val Ser
725 730 735

Trp Pro Gly Asn Asp Arg Leu Leu Thr Pro Asn Glu Phe Glu Ile Lys
740 745 750

Arg Ser Val Asp Gly Glu Gly Tyr Asn Val Ala Gln Cys Asn Met Thr
755 760 765

Lys Asp Trp Phe Leu Val Gln Met Leu Ala Asn Tyr Asn Ile Gly Tyr
770 775 780

Gln Gly Phe Tyr Ile Pro Glu Ser Tyr Lys Asp Arg Met Tyr Ser Phe
785 790 795 800

Phe Arg Asn Phe Gln Pro Met Ser Arg Gln Val Val Asp Gln Thr Lys
805 810 815

Tyr Lys Asp Tyr Gln Glu Val Gly Ile Ile His Gln His Asn Asn Ser
820 825 830

Gly Phe Val Gly Tyr Leu Ala Pro Thr Met Arg Glu Gly Gln Ala Tyr
835 840 845

Pro Ala Asn Phe Pro Tyr Pro Leu Ile Gly Lys Thr Ala Val Asp Ser
850 855 860

Ile Thr Gln Lys Lys Phe Leu Cys Asp Arg Thr Leu Trp Arg Ile Pro
865 870 875 880

Phe Ser Ser Asn Phe Met Ser Met Gly Ala Leu Ser Asp Leu Gly Gln
885 890 895

Asn Leu Leu Tyr Ala Asn Ser Ala His Ala Leu Asp Met Thr Phe Glu
900 905 910

Val Asp Pro Met Asp Glu Pro Thr Leu Leu Tyr Val Leu Phe Glu Val
915 920 925

Phe Asp Val Val Arg Val His Gln Pro His Arg Gly Val Ile Glu Thr
930 935 940

Val Tyr Leu Arg Thr Pro Phe Ser Ala Gly Asn Ala Thr Thr
945 950 955

<210> 6
<211> 2874
<212> DNA
<213> Pan troglodytes

<400> 6
atggcgaccc catcgatgat gccgcagtg tctgacatgc acatctcggg ccaggacgcc 60
tcggagtacc tgagccccgg gctggtgcag ttcgccgcg ccaccgagag ctacttcagc 120

ctgagtaaca agtttaggaa cccacggtg gcgcccacgc acgatgtgac caccgaccgg	180
tctcagcgcc tgacgctgcg gttcattccc gtggaccgcg aggacaccgc gtactcgtac	240
aaggcgcggt tcacctggc cgtgggcgac aaccgcgtgc tggacatggc ctccacctac	300
tttgacatcc gcggggtgct ggaccggggc cccactttca agccttactc tggcaccgcc	360
tacaactccc tggcccccac gggcgctccc aactcctgcg agtgggagca attagaagaa	420
gccagggccg ctgtggaaga cgaagaatta gaagatgaag acgaggaacc acaggatgag	480
gcacctgtga aaaaaaccca tgtatacgct caggctcccc tttctggaga agaaattact	540
aaaaacggtt tgcaaatagg gtcagataac acagaagccc agtctaagcc catatatgca	600
gacacctacat tccagcctga accccaaatc ggggaatccc agtggaatga ggcagatgct	660
acagttgccg gcggtagagt gctaaagaaa tccactccca tgaagccatg ctatggttcc	720
tatgcaagac ccacaaactc caatggaggt caaggtgtgc tgggtggctga tgataagggg	780
gttcttcaat ctaaagttga attgcaattt ttttcaaata ctactactct taatcagcgg	840
gagggtaacg atacaaaacc aaaagtgggt ctgtatagcg aagatgtgca catggaaact	900
ccagacacc acattttctta caagcccaca aaaagcgatg acaattcaaa aatcatgctg	960
ggtcagcagt ccatgcccaa cagacctaat tacatcgggt tcagagacaa ctttatcggc	1020
ctcatgtatt acaatagcac tggcaacatg ggagtgcctg caggtcaggc ctctcagttg	1080
aatgcagtgg tggacttgca agacagaaac acagaactgt cctaccagct cttgcttgat	1140
tccatgggtg acagaaccag atacttttcc atgtggaatc aggcagtgga cagttatgac	1200
ccagatgtca gaattattga aaatcatgga actgaagacg agctcccaa ctattgtttc	1260
cctctgggcg gcataggggt aactgacact taccaggcca ttaaaaccaa tggcaatggt	1320
caagaaaacc caacctggga aaaagataca gagtttgcag accgcaatga aataggggtg	1380
ggaaacaatt tcgctatgga gatcaacctc agtgccaacc tgtggagaaa cttcctgtac	1440
tccaacgtgg cgctgtacct gccagacaag cttaagtaca accctccaa tgtggacatc	1500
tctgacaacc ccaacaccta cgattacatg aacaagcgag tgggtggccc ggggctgggtg	1560
gactgctaca tcaacctggg gcgcgctgg tgcgtggact acatggacaa cgtcaacccc	1620
ttcaaccacc accgcaatgc gggcctgcgc taccgctcca tgctcctggg caacgggcgc	1680
tacgtgccct tccacatcca ggtgccccag aagttctttg ccatcaagaa ctcctcctc	1740
ctgccgggct cctacaccta cgagtggaac ttcaggaagg atgtcaacat ggtcctccag	1800
agctctctgg gtaacgatct cagggtggac ggggccagca tcaagttcga gagcatctgc	1860
ctctacgcca ctttcttccc catggcccac aacacggcct ccacgctcga ggccatgctc	1920
aggaacgaca ccaacgacca gtccttcaat gactacctct ccgccgcaa catgctctac	1980
cccatacccg ccaacgccac caacgtcccc atctccatcc cctcgcgcaa ctggggcgcc	2040

ttccgcggtt	gggccttcac	ccgcctcaag	accaaggaga	ccccctccct	gggctcggga	2100
ttcgaccctt	actacaccta	ctcggggtcc	attccctacc	tggacggcac	cttctacctc	2160
aaccacactt	tcaagaaggt	ctcgggtcacc	ttcgactcct	cggtcagctg	gccgggcaac	2220
gaccgtctgc	tcacccccaa	cgagttcgaa	atcaagcgct	cggtcgacgg	ggaggggtac	2280
aacgtggccc	agtgaacat	gaccaaggac	tggttcctgg	tccagatgct	ggccaactac	2340
aacatcgggt	accagggctt	ctacatccca	gagagctaca	aggacaggat	gtactccttc	2400
ttcaggaact	tccagcccat	gagccggcag	gtgggtggacc	agaccaagta	caaggactac	2460
caggaggtgg	gcacatcca	ccagcacaac	aactcggggt	tcgtgggcta	cctcgcccc	2520
accatgcgcg	agggacaggc	ctaccccgcc	aacttcccct	accgctcat	aggcaagacc	2580
gcggtcgaca	gcacaccca	gaaaaagttc	ctctgcgac	gcaccctctg	gcgcaccccc	2640
ttctccagca	acttcatgtc	catgggtgcg	ctctcggacc	tgggccagaa	cttgctctac	2700
gccaactccg	cccacgcct	cgacatgacc	ttcgaggtcg	accccatgga	cgagcccacc	2760
cttctctatg	ttctgttcga	agtctttgac	gtgggtccggg	tccaccagcc	gcaccgcggc	2820
gtcatcgaga	ccgtgtacct	gcgtacgccc	ttctcggccg	gcaacgccac	cacc	2874

<210> 7
 <211> 578
 <212> PRT
 <213> Pan troglodytes

<400> 7

Met	Lys	Arg	Thr	Lys	Thr	Ser	Asp	Glu	Ser	Phe	Asn	Pro	Val	Tyr	Pro
1				5					10					15	

Tyr	Asp	Thr	Glu	Ser	Gly	Pro	Pro	Ser	Val	Pro	Phe	Leu	Thr	Pro	Pro
			20					25					30		

Phe	Val	Ser	Pro	Asp	Gly	Phe	Gln	Glu	Ser	Pro	Pro	Gly	Val	Leu	Ser
		35					40					45			

Leu	Asn	Leu	Ala	Glu	Pro	Leu	Val	Thr	Ser	His	Gly	Met	Leu	Ala	Leu
	50						55				60				

Lys	Met	Gly	Ser	Gly	Leu	Ser	Leu	Asp	Asp	Ala	Gly	Asn	Leu	Thr	Ser
65					70					75				80	

Gln	Asp	Ile	Thr	Thr	Ala	Ser	Pro	Pro	Leu	Lys	Lys	Thr	Lys	Thr	Asn
				85					90					95	

Leu	Ser	Leu	Glu	Thr	Ser	Ser	Pro	Leu	Thr	Val	Ser	Thr	Ser	Gly	Ala
			100					105						110	

Leu	Thr	Val	Ala	Ala	Ala	Ala	Pro	Leu	Ala	Val	Ala	Gly	Thr	Ser	Leu	115	120	125	
Thr	Met	Gln	Ser	Glu	Ala	Pro	Leu	Thr	Val	Gln	Asp	Ala	Lys	Leu	Thr	130	135	140	
Leu	Ala	Thr	Lys	Gly	Pro	Leu	Thr	Val	Ser	Glu	Gly	Lys	Leu	Ala	Leu	145	150	155	160
Gln	Thr	Ser	Ala	Pro	Leu	Thr	Ala	Ala	Asp	Ser	Ser	Thr	Leu	Thr	Val	165	170	175	
Ser	Ala	Thr	Pro	Pro	Leu	Ser	Thr	Ser	Asn	Gly	Ser	Leu	Gly	Ile	Asp	180	185	190	
Met	Gln	Ala	Pro	Ile	Tyr	Thr	Thr	Asn	Gly	Lys	Leu	Gly	Leu	Asn	Phe	195	200	205	
Gly	Ala	Pro	Leu	His	Val	Val	Asp	Ser	Leu	Asn	Ala	Leu	Thr	Val	Val	210	215	220	
Thr	Gly	Gln	Gly	Leu	Thr	Ile	Asn	Gly	Thr	Ala	Leu	Gln	Thr	Arg	Val	225	230	235	240
Ser	Gly	Ala	Leu	Asn	Tyr	Asp	Thr	Ser	Gly	Asn	Leu	Glu	Leu	Arg	Ala	245	250	255	
Ala	Gly	Gly	Met	Arg	Val	Asp	Ala	Asn	Gly	Gln	Leu	Ile	Leu	Asp	Val	260	265	270	
Ala	Tyr	Pro	Phe	Asp	Ala	Gln	Asn	Asn	Leu	Ser	Leu	Arg	Leu	Gly	Gln	275	280	285	
Gly	Pro	Leu	Phe	Val	Asn	Ser	Ala	His	Asn	Leu	Asp	Val	Asn	Tyr	Asn	290	295	300	
Arg	Gly	Leu	Tyr	Leu	Phe	Thr	Ser	Gly	Asn	Thr	Lys	Lys	Leu	Glu	Val	305	310	315	320
Asn	Ile	Lys	Thr	Ala	Lys	Gly	Leu	Ile	Tyr	Asp	Asp	Thr	Ala	Ile	Ala	325	330	335	
Ile	Asn	Ala	Gly	Asp	Gly	Leu	Gln	Phe	Asp	Ser	Gly	Ser	Asp	Thr	Asn	340	345	350	
Pro	Leu	Lys	Thr	Lys	Leu	Gly	Leu	Gly	Leu	Asp	Tyr	Asp	Ser	Ser	Arg	355	360	365	

Ala Ile Ile Ala Lys Leu Gly Thr Gly Leu Ser Phe Asp Asn Thr Gly
 370 375 380

Ala Ile Thr Val Gly Asn Lys Asn Asp Asp Lys Leu Thr Leu Trp Thr
 385 390 395 400

Thr Pro Asp Pro Ser Pro Asn Cys Arg Ile Tyr Ser Glu Lys Asp Ala
 405 410 415

Lys Phe Thr Leu Val Leu Thr Lys Cys Gly Ser Gln Val Leu Ala Ser
 420 425 430

Val Ser Val Leu Ser Val Lys Gly Ser Leu Ala Pro Ile Ser Gly Thr
 435 440 445

Val Thr Ser Ala Gln Ile Val Leu Arg Phe Asp Glu Asn Gly Val Leu
 450 455 460

Leu Ser Asn Ser Ser Leu Asp Pro Gln Tyr Trp Asn Tyr Arg Lys Gly
 465 470 475 480

Asp Leu Thr Glu Gly Thr Ala Tyr Thr Asn Ala Val Gly Phe Met Pro
 485 490 495

Asn Leu Thr Ala Tyr Pro Lys Thr Gln Ser Gln Thr Ala Lys Ser Asn
 500 505 510

Ile Val Ser Gln Val Tyr Leu Asn Gly Asp Lys Ser Lys Pro Met Thr
 515 520 525

Leu Thr Ile Thr Leu Asn Gly Thr Asn Glu Thr Gly Asp Ala Thr Val
 530 535 540

Ser Thr Tyr Ser Met Ser Phe Ser Trp Asn Trp Asn Gly Ser Asn Tyr
 545 550 555 560

Ile Asn Glu Thr Phe Gln Thr Asn Ser Phe Thr Phe Ser Tyr Ile Ala
 565 570 575

Gln Glu

<210> 8

<211> 1734

<212> DNA

<213> Pan troglodytes

<400> 8

atgaagcgca ccaaaacgtc tgacgagagc ttcaaccccg tgtacccta tgacacggaa

agcggccctc cctccgtccc tttcctcacc cctcccttcg tgtctcccga tggattccaa	120
gaaagtcccc ccgggggtcct gtctctgaac ctggccgagc ccctgggtcac ttcccacggc	180
atgctcgcgc tgaaaatggg aagtggcctc tccctggacg acgctggcaa cctcacctct	240
caagatatca ccaccgctag ccctccccc ctaaaaaacca agaccaacct cagcctagaa	300
acctcatccc ccctaactgt gagcacctca ggcgcctca ccgtagcagc cgccgctccc	360
ctggcggtgg ccggcacctc cctcaccatg caatcagagg cccccctgac agtacaggat	420
gcaaaactca ccctggccac caaaggcccc ctgaccgtgt ctgaaggcaa actggccttg	480
caaacatcgg ccccgctgac ggccgctgac agcagcacc tcacagtcag tgccacacca	540
ccccttagca caagcaatgg cagcttgggt attgacatgc aagcccccat ttacaccacc	600
aatggaaaac taggacttaa ctttggcgct cccctgcatg tggtagacag cctaaatgca	660
ctgactgtag ttactggcca aggtcttacg ataaacggaa cagccctaca aactagagtc	720
tcaggtgccc tcaactatga cacatcagga aacctagaat tgagagctgc aggggggatg	780
cgagttgatg caaatggtca acttatcctt gatgtagctt acccatttga tgcacaaaac	840
aatctcagcc ttaggcttgg acagggaccc ctgtttgtta actctgcca caacttggat	900
gttaactaca acagaggcct ctacctgttc acatctggaa ataccaaaaa gctagaagtt	960
aatatcaaaa cagccaaggg tctcatttat gatgacactg ctatagcaat caatgcgggt	1020
gatgggctac agtttgactc aggtcagat acaaatccat taaaaactaa acttggatta	1080
ggactggatt atgactccag cagagccata attgctaaac tgggaactgg cctaagcttt	1140
gacaacacag gtgccatcac agtaggcaac aaaaatgatg acaagcttac cttgtggacc	1200
acaccagacc catcccctaa ctgtagaatc tattcagaga aagatgctaa attcacactt	1260
gttttgacta aatgcggcag tcaggtgttg gccagcgttt ctgttttatc tgtaaaaggt	1320
agccttgcg ccatcagtgg cacagtaact agtgctcaga ttgtcctcag atttgatgaa	1380
aatggagttc tactaagcaa ttcttccctt gaccctcaat actggaacta cagaaaaggt	1440
gaccttacag agggcactgc atataccaac gcagtgggat ttatgcccaa cctcacagca	1500
tacccaaaaa cacagagcca aactgctaaa agcaacattg taagtcaggt ttacttgaat	1560
ggggacaaat ccaaaccat gaccctcacc attaccctca atggaactaa tgaaacagga	1620
gatgccacag taagcactta ctccatgtca ttctcatgga actggaatgg aagtaattac	1680
attaatgaaa cgttccaaac caactccttc accttctcct acatgcccc agaa	1734

<210> 9
 <211> 593
 <212> PRT
 <213> Pan troglodytes
 <400> 9

Met	Arg	Arg	Ala	Ala	Met	Tyr	Gln	Glu	Gly	Pro	Pro	Pro	Ser	Tyr	Glu	1	5	10	15
Ser	Val	Val	Gly	Ala	Ala	Ala	Ala	Ala	Pro	Ser	Ser	Pro	Phe	Ala	Ser	20	25	30	
Gln	Leu	Leu	Glu	Pro	Pro	Tyr	Val	Pro	Pro	Arg	Tyr	Leu	Arg	Pro	Thr	35	40	45	
Gly	Gly	Arg	Asn	Ser	Ile	Arg	Tyr	Ser	Glu	Leu	Ala	Pro	Leu	Phe	Asp	50	55	60	
Thr	Thr	Arg	Val	Tyr	Leu	Val	Asp	Asn	Lys	Ser	Ala	Asp	Val	Ala	Ser	65	70	75	80
Leu	Asn	Tyr	Gln	Asn	Asp	His	Ser	Asn	Phe	Leu	Thr	Thr	Val	Ile	Gln	85	90	95	
Asn	Asn	Asp	Tyr	Ser	Pro	Ser	Glu	Ala	Ser	Thr	Gln	Thr	Ile	Asn	Leu	100	105	110	
Asp	Asp	Arg	Ser	His	Trp	Gly	Gly	Asp	Leu	Lys	Thr	Ile	Leu	His	Thr	115	120	125	
Asn	Met	Pro	Asn	Val	Asn	Glu	Phe	Met	Phe	Thr	Asn	Lys	Phe	Lys	Ala	130	135	140	
Arg	Val	Met	Val	Ser	Arg	Ser	His	Thr	Lys	Glu	Asp	Arg	Val	Glu	Leu	145	150	155	160
Lys	Tyr	Glu	Trp	Val	Glu	Phe	Glu	Leu	Pro	Glu	Gly	Asn	Tyr	Ser	Glu	165	170	175	
Thr	Met	Thr	Ile	Asp	Leu	Met	Asn	Asn	Ala	Ile	Val	Glu	His	Tyr	Leu	180	185	190	
Lys	Val	Gly	Arg	Gln	Asn	Gly	Val	Leu	Glu	Ser	Asp	Ile	Gly	Val	Lys	195	200	205	
Phe	Asp	Thr	Arg	Asn	Phe	Arg	Leu	Gly	Leu	Asp	Pro	Val	Thr	Gly	Leu	210	215	220	
Val	Met	Pro	Gly	Val	Tyr	Thr	Asn	Glu	Ala	Phe	His	Pro	Asp	Ile	Ile	225	230	235	240
Leu	Leu	Pro	Gly	Cys	Gly	Val	Asp	Phe	Thr	Tyr	Ser	Arg	Leu	Ser	Asn	245	250	255	

Leu Leu Gly Ile Arg Lys Arg Gln Pro Phe Gln Glu Gly Phe Arg Ile
 260 265 270

Thr Tyr Glu Asp Leu Glu Gly Gly Asn Ile Pro Ala Leu Leu Asp Val
 275 280 285

Glu Ala Tyr Gln Asp Ser Leu Lys Glu Asn Glu Ala Gly Gln Glu Asp
 290 295 300

Thr Ala Pro Ala Ala Ser Ala Ala Ala Glu Gln Gly Glu Asp Ala Ala
 305 310 315 320

Asp Thr Ala Ala Ala Asp Gly Ala Glu Ala Asp Pro Ala Met Val Val
 325 330 335

Glu Ala Pro Glu Gln Glu Glu Asp Met Asn Asp Ser Ala Val Arg Gly
 340 345 350

Asp Thr Phe Val Thr Arg Gly Glu Glu Lys Gln Ala Glu Ala Glu Ala
 355 360 365

Ala Ala Glu Glu Lys Gln Leu Ala Ala Ala Ala Ala Ala Ala Ala Leu
 370 375 380

Ala Ala Ala Glu Ala Glu Ser Glu Gly Thr Lys Pro Ala Lys Glu Pro
 385 390 395 400

Val Ile Lys Pro Leu Thr Glu Asp Ser Lys Lys Arg Ser Tyr Asn Leu
 405 410 415

Leu Lys Asp Ser Thr Asn Thr Ala Tyr Arg Ser Trp Tyr Leu Ala Tyr
 420 425 430

Asn Tyr Gly Asp Pro Ser Thr Gly Val Arg Ser Trp Thr Leu Leu Cys
 435 440 445

Thr Pro Asp Val Thr Cys Gly Ser Glu Gln Val Tyr Trp Ser Leu Pro
 450 455 460

Asp Met Met Gln Asp Pro Val Thr Phe Arg Ser Thr Arg Gln Val Ser
 465 470 475 480

Asn Phe Pro Val Val Gly Ala Glu Leu Leu Pro Val His Ser Lys Ser
 485 490 495

Phe Tyr Asn Asp Gln Ala Val Tyr Ser Gln Leu Ile Arg Gln Phe Thr
 500 505 510

Ser Leu Thr His Val Phe Asn Arg Phe Pro Glu Asn Gln Ile Leu Ala
515 520 525

Arg Pro Pro Ala Pro Thr Ile Thr Thr Val Ser Glu Asn Val Pro Ala
530 535 540

Leu Thr Asp His Gly Thr Leu Pro Leu Arg Asn Ser Ile Gly Gly Val
545 550 555 560

Gln Arg Val Thr Val Thr Asp Ala Arg Arg Arg Thr Cys Pro Tyr Val
565 570 575

Tyr Lys Ala Leu Gly Ile Val Ser Pro Arg Val Leu Ser Ser Arg Thr
580 585 590

Phe

<210> 10
<211> 1779
<212> DNA
<213> Pan troglodytes

<400> 10
atgcggcgcg cggcgatgta ccaggaggga cctcctccct cttacgagag cgtggtgggc 60
gcggcgggcg cggcgccctc ttctcccttt gcgtcgcagc tgctggagcc gccgtacgtg 120
cctccgcgct acctgcggcc tacggggggg agaaacagca tccgttactc ggagctggcg 180
ccctgttcg acaccacccg ggtgtacctg gtggacaaca agtcggcgga cgtggcctcc 240
ctgaactacc agaacgacca cagcaatttt ttgaccacgg tcatccagaa caatgactac 300
agcccgagcg aggccagcac ccagaccatc aatctggatg accggtcgca ctggggcggc 360
gacctgaaaa ccatcctgca caccaacatg cccaacgtga acgagttcat gttcaccaat 420
aagttcaagg cgcggtgat ggtgtcgcgc tcgcacacca aggaagaccg ggtggagctg 480
aagtacgagt ggggtggagtt cgagctgcca gagggcaact actccgagac catgaccatt 540
gacctgatga acaacgcgat cgtggagcac tatctgaaag tgggcaggca gaacggggtc 600
ctggagagcg acatcggggg caagtctgac accaggaact tccgcctggg gctggacccc 660
gtgaccgggc tggttatgcc cggggtgtac accaacgagg ccttccatcc cgacatcatc 720
ctgctgcccc gctgcggggg ggacttcact tacagccgcc tgagcaacct cctgggcatc 780
cgcaagcggc agcccttcca ggagggttc aggatcacct acgaggacct ggaggggggc 840
aacatccccg cgctcctcga tgtggaggcc taccaggata gcttgaagga aaatgaggcg 900
ggacaggagg ataccgcccc cgccgcctcc gccgccgccg agcagggcga ggatgctgct 960
gacaccgcgg ccgcggacgg ggcagaggcc gaccccgcta tgggtggtgga ggctcccag 1020

caggaggagg acatgaatga cagtgcggtg cgcgagagaca ccttcgtcac ccgggggggag 1080
gaaaagcaag cggaggccga ggccgcggcc gaggaaaagc aactggcggc agcagcggcg 1140
gcggcggcgt tggccgcggc ggaggctgag tctgagggga ccaagcccgc caaggagccc 1200
gtgattaagc ccctgaccga agatagcaag aagcgcagtt acaacctgct caaggacagc 1260
accaacaccg cgtaccgcag ctggtacctg gcctacaact acggcgaccc gtcgacgggg 1320
gtgcgctcct ggaccctgct gtgcacgccg gacgtgacct gcggtcggga gcaggtgtac 1380
tggtcgctgc ccgacatgat gcaagacccc gtgaccttcc gctccacgcg gcaggtcagc 1440
aacttcccgg tggtggggcg cgagctgctg cccgtgact ccaagagctt ctacaacgac 1500
caggccgtct actcccagct catccgccag ttcacctctc tgaccacagt gttcaatcgc 1560
tttctgaga accagattct ggcgcgccc cccgccccca ccatcaccac cgtcagtga 1620
aacgttcttg ctctcacaga tcacgggacg ctaccgctgc gcaacagcat cggaggagtc 1680
cagcgagtga ccgttactga cgccagacgc cgcacctgcc cctacgttta caaggccttg 1740
ggcatagtct cgccgcgcgt cctttccagc cgcactttt 1779

<210> 11
<211> 960
<212> PRT
<213> Pan troglodytes

<400> 11

Met Ala Thr Pro Ser Met Met Pro Gln Trp Ser Tyr Met His Ile Ser
1 5 10 15

Gly Gln Asp Ala Ser Glu Tyr Leu Ser Pro Gly Leu Val Gln Phe Ala
20 25 30

Arg Ala Thr Glu Ser Tyr Phe Ser Leu Ser Asn Lys Phe Arg Asn Pro
35 40 45

Thr Val Ala Pro Thr His Asp Val Thr Thr Asp Arg Ser Gln Arg Leu
50 55 60

Thr Leu Arg Phe Ile Pro Val Asp Arg Glu Asp Thr Ala Tyr Ser Tyr
65 70 75 80

Lys Ala Arg Phe Thr Leu Ala Val Gly Asp Asn Arg Val Leu Asp Met
85 90 95

Ala Ser Thr Tyr Phe Asp Ile Arg Gly Val Leu Asp Arg Gly Pro Thr
100 105 110

Phe Lys Pro Tyr Ser Gly Thr Ala Tyr Asn Ser Leu Ala Pro Lys Gly
115 120 125

Ala	Pro	Asn	Ser	Cys	Glu	Trp	Glu	Gln	Glu	Glu	Thr	Gln	Ala	Val	Glu	130	135	140	
Glu	Ala	Ala	Glu	Glu	Glu	Glu	Glu	Asp	Ala	Asp	Gly	Gln	Ala	Glu	Glu	145	150	155	160
Glu	Gln	Ala	Ala	Thr	Lys	Lys	Thr	His	Val	Tyr	Ala	Gln	Ala	Pro	Leu	165	170	175	
Ser	Gly	Glu	Lys	Ile	Ser	Lys	Asp	Gly	Leu	Gln	Ile	Gly	Thr	Asp	Ala	180	185	190	
Thr	Ala	Thr	Glu	Gln	Lys	Pro	Ile	Tyr	Ala	Asp	Pro	Thr	Phe	Gln	Pro	195	200	205	
Glu	Pro	Gln	Ile	Gly	Glu	Ser	Gln	Trp	Asn	Glu	Ala	Asp	Ala	Thr	Val	210	215	220	
Ala	Gly	Gly	Arg	Val	Leu	Lys	Lys	Ser	Thr	Pro	Met	Lys	Pro	Cys	Tyr	225	230	235	240
Gly	Ser	Tyr	Ala	Arg	Pro	Thr	Asn	Ala	Asn	Gly	Gly	Gln	Gly	Val	Leu	245	250	255	
Thr	Ala	Asn	Ala	Gln	Gly	Gln	Leu	Glu	Ser	Gln	Val	Glu	Met	Gln	Phe	260	265	270	
Phe	Ser	Thr	Ser	Glu	Asn	Ala	Arg	Asn	Glu	Ala	Asn	Asn	Ile	Gln	Pro	275	280	285	
Lys	Leu	Val	Leu	Tyr	Ser	Glu	Asp	Val	His	Met	Glu	Thr	Pro	Asp	Thr	290	295	300	
His	Leu	Ser	Tyr	Lys	Pro	Ala	Lys	Ser	Asp	Asp	Asn	Ser	Lys	Ile	Met	305	310	315	320
Leu	Gly	Gln	Gln	Ser	Met	Pro	Asn	Arg	Pro	Asn	Tyr	Ile	Gly	Phe	Arg	325	330	335	
Asp	Asn	Phe	Ile	Gly	Leu	Met	Tyr	Tyr	Asn	Ser	Thr	Gly	Asn	Met	Gly	340	345	350	
Val	Leu	Ala	Gly	Gln	Ala	Ser	Gln	Leu	Asn	Ala	Val	Val	Asp	Leu	Gln	355	360	365	
Asp	Arg	Asn	Thr	Glu	Leu	Ser	Tyr	Gln	Leu	Leu	Leu	Asp	Ser	Met	Gly	370	375	380	

Asp Arg Thr Arg Tyr Phe Ser Met Trp Asn Gln Ala Val Asp Ser Tyr
385 390 395 400

Asp Pro Asp Val Arg Ile Ile Glu Asn His Gly Thr Glu Asp Glu Leu
405 410 415

Pro Asn Tyr Cys Phe Pro Leu Gly Gly Ile Gly Val Thr Asp Thr Tyr
420 425 430

Gln Ala Val Lys Thr Asn Asn Gly Asn Asn Gly Gly Gln Val Thr Trp
435 440 445

Thr Lys Asp Glu Thr Phe Ala Asp Arg Asn Glu Ile Gly Val Gly Asn
450 455 460

Asn Phe Ala Met Glu Ile Asn Leu Ser Ala Asn Leu Trp Arg Asn Phe
465 470 475 480

Leu Tyr Ser Asn Val Ala Leu Tyr Leu Pro Asp Lys Leu Lys Tyr Asn
485 490 495

Pro Ser Asn Val Asp Ile Ser Asp Asn Pro Asn Thr Tyr Asp Tyr Met
500 505 510

Asn Lys Arg Val Val Ala Pro Gly Leu Val Asp Cys Tyr Ile Asn Leu
515 520 525

Gly Ala Arg Trp Ser Leu Asp Tyr Met Asp Asn Val Asn Pro Phe Asn
530 535 540

His His Arg Asn Ala Gly Leu Arg Tyr Arg Ser Met Leu Leu Gly Asn
545 550 555 560

Gly Arg Tyr Val Pro Phe His Ile Gln Val Pro Gln Lys Phe Phe Ala
565 570 575

Ile Lys Asn Leu Leu Leu Leu Pro Gly Ser Tyr Thr Tyr Glu Trp Asn
580 585 590

Phe Arg Lys Asp Val Asn Met Val Leu Gln Ser Ser Leu Gly Asn Asp
595 600 605

Leu Arg Val Asp Gly Ala Ser Ile Lys Phe Glu Ser Ile Cys Leu Tyr
610 615 620

Ala Thr Phe Phe Pro Met Ala His Asn Thr Ala Ser Thr Leu Glu Ala
625 630 635 640

Met	Leu	Arg	Asn	Asp	Thr	Asn	Asp	Gln	Ser	Phe	Asn	Asp	Tyr	Leu	Ser	645	650	655	
Ala	Ala	Asn	Met	Leu	Tyr	Pro	Ile	Pro	Ala	Asn	Ala	Thr	Asn	Val	Pro	660	665	670	
Ile	Ser	Ile	Pro	Ser	Arg	Asn	Trp	Ala	Ala	Phe	Arg	Gly	Trp	Ala	Phe	675	680	685	
Thr	Arg	Leu	Lys	Thr	Lys	Glu	Thr	Pro	Ser	Leu	Gly	Ser	Gly	Phe	Asp	690	695	700	
Pro	Tyr	Tyr	Thr	Tyr	Ser	Gly	Ser	Ile	Pro	Tyr	Leu	Asp	Gly	Thr	Phe	705	710	715	720
Tyr	Leu	Asn	His	Thr	Phe	Lys	Lys	Val	Ser	Val	Thr	Phe	Asp	Ser	Ser	725	730	735	
Val	Ser	Trp	Pro	Gly	Asn	Asp	Arg	Leu	Leu	Thr	Pro	Asn	Glu	Phe	Glu	740	745	750	
Ile	Lys	Arg	Ser	Val	Asp	Gly	Glu	Gly	Tyr	Asn	Val	Ala	Gln	Cys	Asn	755	760	765	
Met	Thr	Lys	Asp	Trp	Phe	Leu	Val	Gln	Met	Leu	Ala	Asn	Tyr	Asn	Ile	770	775	780	
Gly	Tyr	Gln	Gly	Phe	Tyr	Ile	Pro	Glu	Ser	Tyr	Lys	Asp	Arg	Met	Tyr	785	790	795	800
Ser	Phe	Phe	Arg	Asn	Phe	Gln	Pro	Met	Ser	Arg	Gln	Val	Val	Asp	Gln	805	810	815	
Thr	Lys	Tyr	Lys	Asp	Tyr	Gln	Glu	Val	Gly	Ile	Ile	His	Gln	His	Asn	820	825	830	
Asn	Ser	Gly	Phe	Val	Gly	Tyr	Leu	Ala	Pro	Thr	Met	Arg	Glu	Gly	Gln	835	840	845	
Ala	Tyr	Pro	Ala	Asn	Phe	Pro	Tyr	Pro	Leu	Ile	Gly	Lys	Thr	Ala	Val	850	855	860	
Asp	Ser	Ile	Thr	Gln	Lys	Lys	Phe	Leu	Cys	Asp	Arg	Thr	Leu	Trp	Arg	865	870	875	880
Ile	Pro	Phe	Ser	Ser	Asn	Phe	Met	Ser	Met	Gly	Ala	Leu	Ser	Asp	Leu	885	890	895	

Gly Gln Asn Leu Leu Tyr Ala Asn Ser Ala His Ala Leu Asp Met Thr
900 905 910

Phe Glu Val Asp Pro Met Asp Glu Pro Thr Leu Leu Tyr Val Leu Phe
915 920 925

Glu Val Phe Asp Val Val Arg Val His Gln Pro His Arg Gly Val Ile
930 935 940

Glu Thr Val Tyr Leu Arg Thr Pro Phe Ser Ala Gly Asn Ala Thr Thr
945 950 955 960

<210> 12
<211> 2880
<212> DNA
<213> Pan troglodytes

<400> 12
atggcgaccc catcgatgat gccgcagtgg tcgtacatgc acatctcggg ccaggacgcc 60
tcggagtacc tgagccccgg gctggtgcag ttgccccg cgccaccgagag ctacttcagc 120
ctgagtaaca agtttaggaa cccacgggtg gcgcccacgc acgatgtgac caccgaccgg 180
tctcagcgcc tgacgctgcg gttcattccc gtggaccg cgaggacaccgc gtactcgtac 240
aaggcgcggt tcaccctggc cgtggggcgac aaccgcgtgc tggacatggc ctccacctac 300
tttgacatcc gcgggggtgct ggaccggggg cccactttca agccctactc tggcaccgcc 360
tacaactccc tggcccccaa gggcgctccc aactcctgcg agtgggagca agaggaaact 420
caggcagttg aagaagcagc agaagaggaa gaagaagatg ctgacgggtca agctgaggaa 480
gagcaagcag ctacaaaaaa gactcatgta tatgctcagg ctcccccttc tggcgaaaaa 540
attagtaaag atggtctgca aataggaacg gacgctacag ctacagaaca aaaacctatt 600
tatgcagacc ctacattcca gcccgacccc caaatcgggg agtcccagtg gaatgaggca 660
gatgctacag tcgccggcgg tagagtgtca aagaaatcta ctcccatgaa accatgctat 720
ggttcctatg caagaccac aaatgcta atggaggtcagg gtgtactaac ggcaaagtc 780
cagggacagc tagaatctca ggttgaaatg caattctttt caacttctga aaacgcccgt 840
aacgaggcta acaacattca gcccaaattg gtgctgtata gtgaggatgt gcacatggag 900
accccgata cgcacctttc ttacaagccc gcaaaaagcg atgacaattc aaaaatcatg 960
ctgggtcagc agtccatgcc caacagacct aattacatcg gcttcagaga caactttatc 1020
ggcctcatgt attacaatag cactggcaac atgggagtg ttgcagggtca ggcctctcag 1080
ttgaatgcag tgggtggactt gcaagacaga aacacagaac tgtcctacca gctcttgctt 1140
gattccatgg gtgacagaac cagatacttt tccatgtgga atcaggcagt ggacagttat 1200

gaccagatg ttagaattat tgaaaatcat ggaactgaag acgagctccc caactattgt	1260
ttccctctgg gtggcatagg ggtaactgac acttaccagg ctgttaaaac caacaatggc	1320
aataacgggg gccaggtgac ttggacaaaa gatgaaactt ttgcagatcg caatgaaata	1380
ggggtgggaa acaatttcgc tatggagatc aacctcagt ccaacctgtg gagaaacttc	1440
ctgtactcca acgtggcgct gtacctacca gacaagctta agtacaaccc ctccaatgtg	1500
gacatctctg acaaccccaa cacctacgat tacatgaaca agcgagtggg ggccccgggg	1560
ctggtggact gctacatcaa cctggggcgc cgctggctgc tggactacat ggacaacgtc	1620
aacccttca accaccaccg caatgcgggc ctgcgctacc gctccatgct cctgggcaac	1680
gggcgctacg tgcccttcca catccagggtg cccagaagt tctttgccat caagaacctc	1740
ctcctctgc cgggctccta cacctacgag tggaacttca ggaaggatgt caacatggtc	1800
ctccagagct ctctgggtaa cgatctcagg gtggacgggg ccagcatcaa gttcgagagc	1860
atctgcctct acgccacctt cttcccatg gccacaaca cggcctccac gctcgaggcc	1920
atgctcagga acgacaccaa cgaccagtcc ttcaatgact acctctccgc cgccaacatg	1980
ctctaccca taccgcgcaa cgccaccaac gtcccatct ccatccctc gcgcaactgg	2040
gcggccttcc gcggctgggc cttcaccgc ctcaagacca aggagacccc ctccctgggc	2100
tcgggattcg acccctacta cacctactcg ggctccattc cctacctgga cggcaccttc	2160
tacctcaacc acactttcaa gaaggctctg gtcaccttcg actcctcggg cagctggcgc	2220
ggcaacgacc gtctgctcac cccaacgag ttcgagatca agcgctcggg cgacggggag	2280
ggctacaacg tggcccagtg caacatgacc aaggactggg tcctgggtcca gatgctggcc	2340
aactacaaca tcggctacca gggcttctac atcccagaga gctacaagga caggatgtac	2400
tccttcttca ggaacttcca gcccatgagc cggcagggtg tggaccagac caagtacaag	2460
gactaccagg aggtgggcat catccaccag cacaacaact cgggcttcgt gggctacctc	2520
gccccacca tgcgcgaggg acaggcctac cccgccaact tccctatcc gctcatagge	2580
aagaccgagg tcgacagcat ccccagaaa aagtctctct gcgaccgcac cctctggcgc	2640
atcccttct ccagcaactt catgtccatg ggtgcgctct cggacctggg ccagaacttg	2700
ctctacgcca actccgcca cgccctcgac atgaccttcg aggtcgacct catggacgag	2760
cccacccttc tctatgttct gttcgaagtc tttgacgtgg tccgggtcca ccagccgcac	2820
cgcgcggtca tcgagaccgt gtacctgcgt acgcccttct cggccggcaa cgccaccacc	2880

<210> 13

<211> 37830

<212> DNA

<213> Pan troglodytes

<400> 13

catcatcaat aatatacctt attttggatt gaagccaata tgataatgag atgggcgggcg	60
--	----

cggggcggga ggcgggtccg ggggcgggcc ggcgggcggg gcggtgtggc ggaagtggac	120
tttgtaagtg tggcggatgt gacttgctag tgccgggcgc ggtaaaagtg acgttttccg	180
tgcgcgacaa cgcccacggg aagtgcatt tttccgcgg tttttaccgg atgtttagt	240
gaatttgggc gtaaccaagt aagatttggc cattttcgcg ggaaaactga aacggggaag	300
tgaaatctga ttaatttcgc gttagtcata ccgcgtaata tttgtcgagg gccgagggac	360
tttgccgat tacgtggagg actcgcccag gtgttttttg aggtgaattt ccgcgttccg	420
ggtcaaagtc tccgttttat tattatagtc agctgacgcg gagtgtattt ataccctctg	480
atctcgtcaa gtggccactc ttgagtcca gcgagtagag ttttctctc tgccgctctc	540
cgctccgctc cgctcggctc tgacaccggg gaaaaaatga gacatttcac ctacgatggc	600
ggtgtgctca ccggccagct ggctgctgaa gtcttgaca cctgatcga ggaggtattg	660
gccgataatt atctccctc gactcctttt gagccaccta cacttcacga actctacgat	720
ctggatgtgg tggggcccag cgatccgaac gagcaggcgg tttccagttt tttccagag	780
tccatgttgt tggccagcca ggagggggtc gaacttgaga cccctcctc gatcgtggat	840
tccccgatc cgccgcagct gactaggcag ccgagcgcgt gtgcgggacc tgagactatg	900
ccccagctgc tacctgaggt gatcgatctc acctgtaatg agtctggttt tccaccagc	960
gaggatgagg acgaagagg tgagcagttt gtgttagatt ctgtggaaca acccgggcga	1020
ggatgcaggt cttgtcaata tcaccgaaa aacacaggag actcccagat tatgtgttct	1080
ctgtgttata tgaagatgac ctgtatgtt atttacagta agtttatcat ctgtgggcag	1140
gtgggctata gtgtgggtgg tggctcttgg ggggtttttt aatatatgtc aggggttatg	1200
ctgaagactt ttttattgtg atttttaag gtccagtgtc tgagcccag caagaacctg	1260
aaccggagcc tgagccttct cgcccagga gaaagcctgt aatcttaact agaccagcg	1320
caccggtagc gagaggcctc agcagcgcgg agaccaccga ctccggtgct tcctcatcac	1380
ccccggagat tcacccctg gtgccctgt gtcccgtaa gcccgttgcc gtgagagtca	1440
gtgggcggcg gtctgctgtg gagtgcattg aggacttgct ttttgattca caggaaacctt	1500
tggacttgag cttgaaacgc ccaggcatt aaacctggtc acctggactg aatgagttga	1560
cgcctatgtt tgcttttgaa tgacttaatg tgtatagata ataaagagtg agataatgtt	1620
ttaattgcat ggtgtgttta acttgggcgg agtctgctgg gtatataagc ttccctgggc	1680
taaacttggg tacacttgac ctcatggagg cctgggagtg tttggagaac tttgccggag	1740
ttcgtgcctt gctggacgag agctctaaca atacctcttg gtggtggagg tatttgtggg	1800
gctctcccca gggcaagtta gttttagaa tcaaggagga ttacaagtgg gaatttgaag	1860
agcttttgaa atcctgtggg gagctattgg attctttgaa tctaggccac caggctctct	1920
tccaggagaa ggtcatcagg actttggatt tttccacacc ggggcgcatt gcagccgcgg	1980

ttgcttttct agcttttttg aaggatagat ggagcgaaga gaccacttg agttcgggct	2040
acgtcctgga ttttctggcc atgcaactgt ggagagcatg gatcagacac aagaacaggc	2100
tgcaactgtt gtcttccgtc cgtccgttgc tgattccggc ggaggagcaa caggccgggt	2160
cagaggaccg ggcccgtcgg gatccggagg agagggcacc gaggccgggc gagaggagcg	2220
cgctgaacct gggaaccggg ctgagcggcc atccacatcg ggagtgaatg tcgggcaggt	2280
ggtggatctt tttccagaac tgcggcggat tttgactatt agggaggatg ggcaatttgt	2340
taagggtctt aagagggaga ggggggcttc tgagcataac gaggaggcca gtaatttagc	2400
tttagcttg atgaccagac accgtccaga gtgcatcact tttcagcaga ttaaggacaa	2460
ttgtgccaat gagttggatc tgttgggtca gaagtatagc atagagcagc tgaccactta	2520
ctggctgcag ccgggtgatg atctggagga agctattagg gtgtatgcta aggtggccct	2580
gcggcccgat tgcaagtaca agctcaaggg gctggtgaat atcaggaatt gttgctacat	2640
ttctggcaac ggggcggagg tggagataga gaccgaagac aggggtggctt tcagatgcag	2700
catgatgaat atgtggccgg ggggtgctgg catggacggg gtggtgatta tgaatgtgag	2760
gttcacgggg cccaacttta acggcacggg gtttttgggg aacaccaacc tggtcctgca	2820
cggggtgagc ttctatgggt ttaacaacac ctgtgtggag gcctggaccg atgtgaaggt	2880
ccgcggttgc gccttttatg gatgttggaa ggccatagtg agccgcccta agagcaggag	2940
ttccattaag aaatgcttgt ttgagaggtg caccttgggg atcctggccg agggcaactg	3000
cagggtgctc cacaatgtgg cctccgagtg cggttgcttc atgctagtca agagcgtggc	3060
ggtaatcaag cataatatgg tgtgcggcaa cagcgaggac aaggcctcac agatgctgac	3120
ctgcacggat ggcaactgcc acttgctgaa gaccatccat gtaaccagcc acagccggaa	3180
ggcctggccc gtgttcgagc acaacttgct gaccgcgtgc tccttgcatc tgggcaacag	3240
gcgggggggtg ttcttgccct atcaatgcaa ctttagtcac accaagatct tgctagagcc	3300
cgagagcatg tccaaggtga acttgaacgg ggtgtttgac atgaccatga agatctggaa	3360
ggtgctgagg tacgacgaga ccaggtcccg gtgcagaccc tgcgagtgcg ggggcaagca	3420
tatgaggaac cagcccgtga tgctggatgt gaccgaggag ctgaggacag accacttggt	3480
tctggcctgc accagggccg agtttggttc tagcgatgaa gacacagatt gaggtgggtg	3540
agtgggcgtg gcctgggggtg gtcatgaaaa tatataagtt gggggctctta gggctctctt	3600
atgtgtgttg cagagaccgc cggagccatg agcgggagca gcagcagcag cagtagcagc	3660
agcgccttgg atggcagcat cgtgagccct tatttgacga cgcggatgcc cactgggcc	3720
ggggtgcgtc agaatgtgat gggctccagc atcgacggcc gaccgcctc gcccgcaaat	3780
tccgccacgc tgacctatgc gaccgtcgcg gggacgccgt tggacgccac cgccgccgc	3840
gccgccaccg cagccgcctc ggccgtgcgc agcctggcca cggactttgc attcctggga	3900

ccactggcga	caggggctac	ttctcgggcc	gctgctgccg	cgttcgcga	tgacaagctg	3960
accgccctgc	tggcgcagtt	ggatgcgctt	actcggaac	tgggtgacct	ttctcagcag	4020
gtcatggccc	tgccccagca	ggtctcctcc	ctgcaagctg	gcgggaatgc	ttctcccaca	4080
aatgccgttt	aagataaata	aaaccagact	ctgtttggat	taaagaaaag	tagcaagtgc	4140
attgctctct	ttatttcata	attttccgcg	cgcgataggc	cctagaccag	cgttctcggt	4200
cgttgagggg	gcgggtgtatc	ttctccagga	cgtggtagag	gtggctctgg	acgttgagat	4260
acatgggcat	gagcccgtcc	cgggggtgga	ggtagcacca	ctgcagagct	tcatgctccg	4320
gggtggtggt	gtagatgatc	cagtcgtagc	aggagcgctg	ggcatggtgc	ctaaaaatgt	4380
ccttcagcag	caggccgatg	gccaggggga	ggcccttggg	gtaagtgttt	acaaaacggg	4440
taagttggga	aggggtgcatt	cggggagaga	tgatgtgcat	cttggaactgt	atttttagat	4500
tggcgatggt	tccgcccaga	tcccttctgg	gattcatggt	gtgcaggacc	accagtacag	4560
tgtatccggg	gcacttgggg	aatttgtcat	gcagcttaga	gggaaaagcg	tggaagaact	4620
tggagacgcc	tttgtggcct	cccagatttt	ccatgcattc	gtccatgatg	atggcaatgg	4680
gcccgcggga	ggcagcttgg	gcaaagatat	ttctgggggtc	gctgacgtcg	tagttgtggt	4740
ccagggtgag	gtcgtcatag	gccattttta	caaagcgcg	gcggagggtg	cccgaactgg	4800
ggatgatggg	cccctctggc	cctggggcgt	agttgccctc	gcagatctgc	atttcccagg	4860
ccttaatctc	ggagggggga	atcatatcca	cctgcggggc	gatgaagaaa	acggtttccg	4920
gagccgggga	gattaactgg	gatgagagca	ggtttctaag	cagctgtgat	tttccacaac	4980
cgggtgggcc	ataaataaca	cctataaccg	gttgcagctg	gtagtttaga	gagctgcagc	5040
tgccgtcgtc	ccggaggagg	ggggccacct	cgttgagcat	gtccctgacg	cgcatgttct	5100
ccccgaccag	atccgccaga	aggcgctcgc	cgcccaggga	cagcagctct	tgcaaggaag	5160
caaagttttt	cagcggcttg	aggccgtccg	ccgtgggcat	gtttttcagg	gtctgggtca	5220
gcagctccag	gcgggtccag	agctcgggtg	cgtgctctac	ggcatctcta	tccagcatat	5280
ctcctcggtt	cgcggttgg	ggcgactttc	gctgtagggc	accaagcggg	ggctcgtccag	5340
cggggccaga	gtcatgtcct	tccatgggcg	cagggtcctc	gtcaggggtg	tctgggtcac	5400
ggtgaagggg	tgcgctccgg	gctgagcgct	tgccaagggtg	cgcttgaggc	tggttctgct	5460
ggtgctgaag	cgctgccggg	cttcgccctg	cgcgctggcc	aggtagcatt	tgaccatggg	5520
gtcatagtcc	agccccctcc	cggcgtgtcc	cttggcgcg	agcttgccct	tggaggtggc	5580
gccgcacgag	gggcagagca	ggctcttgag	cgcgtagagc	ttgggggcca	ggaagaccga	5640
ttcgggggag	taggcgtccg	cgccgcagac	cccgcacacg	gtctcgcact	ccaccagcca	5700
ggtgagctcg	gggcgcgccg	ggtcaaaaac	caggtttccc	ccatgctttt	tgatgcgttt	5760
cttacctcgg	gtctccatga	ggtggtgtcc	ccgctcgggtg	acgaagaggc	tgtccgtgtc	5820

tccgtagacc	gacttgaggg	gtcttttctc	cagggggggtc	cctcgggtctt	cctcgtagag	5880
gaactcggac	cactctgaga	cgaaggcccg	cgtccaggcc	aggacgaagg	aggctatgtg	5940
ggaggggtag	cggtcgttgt	ccactagggg	gtccaccttc	tccaagggtgt	gaagacacat	6000
gtcgccttcc	tcggcgtcca	ggaaggatgat	tggctttag	gtgtaggcca	cgtgaccggg	6060
ggttcctgac	gggggggtat	aaaaggggggt	gggggcgcgc	tcgtcgtcac	tctcttccgc	6120
atcgctgtct	gcgagggcca	gctgctgggg	tgagtattcc	ctctcgaagg	ggggcatgac	6180
ctccgcgctg	aggttgtcag	tttccaaaaa	cgaggaggat	ttgatgttca	cctgtcccga	6240
ggtgatacct	ttgagggtac	ccgcgtccat	ctggtcagaa	aacacgatct	ttttattgtc	6300
cagcttggtg	gcgaacgacc	cgtagagggc	gttggagagc	agcttggcga	tggagcgcag	6360
ggtctggttc	ttgtccctgt	cggcgcgctc	cttggccgcg	atgttgagct	gcacgtactc	6420
gcgcgcgacg	cagcgccact	cggggaagac	ggtggtgcgc	tcgtcgggca	ccaggcgcac	6480
gcgccagccg	cggttgtgca	gggtgaccag	gtccacgctg	gtggcgacct	cgccgcgcag	6540
gcgctcgttg	gtccagcaga	gacggccgcc	cttgcgcgag	cagaaggggg	gcaggggggtc	6600
gagctgggtc	tcgtccgggg	ggtccgcgtc	cacggtgaaa	accccggggc	gcaggcgcgc	6660
gtcgaagtag	tctatcttgc	aaccttgcac	gtccagcgcc	tgctgccagt	cgcgggcgcc	6720
gagcgcgcgc	tcgtaggggt	tgagcggcgg	gccccagggc	atgggggtggg	tgagtgcgga	6780
ggcgtacatg	ccgcagatgt	catagacgta	gaggggctcc	cgcaggacct	cgatgtaggt	6840
ggggtagcag	cggccgcgcg	ggatgctggc	gcgcacgtag	tcatacagct	cgtgcgaggg	6900
ggcgaggagg	tcggggccca	ggttggtgcg	ggcggggcgc	tccgcgcgga	agacgatctg	6960
cctgaagatg	gcatgcgagt	tggaagagat	ggtggggcgc	tggaagacgt	tgaagctggc	7020
gtcctgcagg	ccgacggcgt	cgcgcacgaa	ggaggcgtag	gagtcgcgca	gcttgtgtac	7080
cagctcggcg	gtgacctgca	cgtcgaagcg	gcagtagtcg	agggctctcg	ggatgatgtc	7140
atatttagcc	tgcccccttct	ttttccacag	ctcgcgggtg	aggacaaact	cttcgcgggtc	7200
tttccagtac	tcttgatcg	ggaaaccgtc	cggttccgaa	cggtaagagc	ctagcatgta	7260
gaactggttg	acggcctggt	aggcgcagca	gcccttctcc	acggggaggg	cgtaggcctg	7320
cgcggccttg	cggagcagag	tgtgggtcag	ggcgaagggt	tccctgacca	tgactttgag	7380
gtactggtgc	ttgaagtcgg	agtcgtcgca	gccgccccgc	tcccagagcg	agaagtcggt	7440
gcgcttcttg	gagcgggggt	tgggcagagc	gaaggtagaca	tcgttgaaga	ggattttgcc	7500
cgcgcggggc	atgaagttgc	gggtgatgcg	gaagggcccc	ggcacttcag	agcggttggt	7560
gatgacctgg	gcggcgagca	cgatctcgtc	gaagccggtg	atgttgtggc	ccacgatgta	7620
gagttccagg	aagcggggcc	ggccctttac	ggtgggcagc	ttcttttagct	cttcgtaggt	7680
gagctcctcg	ggcgaggcga	ggccgtgctc	ggccagggcc	cagtcgcgca	ggtgcggggt	7740

gtctctgagg	aaggacttcc	agaggtcgcg	ggccaggagg	gtctgcaggc	ggtctctgaa	7800
ggtcctgaac	tggcgggcca	cggccatttt	ttcgggggtg	atgcagtaga	aggtgagggg	7860
gtcttgctgc	cagcgggtccc	agtcgagctg	cagggcgagg	tcgcgcgcgg	cggtgaccag	7920
gcgctcgtcg	ccccgaatt	tcatgaccag	catgaagggc	acgagctgct	ttccgaaggc	7980
ccccatccaa	gtgtaggtct	ctacatcgta	ggtgacaaag	aggcgctccg	tgcgaggatg	8040
cgagccgatc	gggaagaact	ggatctcccc	ccaccagttg	gaggagtggc	tgttgatgtg	8100
gtggaagtag	aagtcccgtc	gccggggccga	acactcgtgc	tggcttttgt	aaaagcgagc	8160
gcagtactgg	cagcgctgca	cgggctgtac	ctcatgcacg	agatgcacct	ttcgcccgcg	8220
cacgaggaag	ccgaggggaa	atctgagccc	ccgcctggc	tcgcggcatg	gctggttctc	8280
ttctactttg	gatgcgtgtc	cgtctccgtc	tggctcctcg	aggggtgtta	cggtgagcgc	8340
gaccaccacg	ccgcgcgagc	cgcaggtcca	gatatcggcg	cgcggcggtc	ggagtttgat	8400
gacgacatcg	cgcagctggg	agctgtccat	ggtctggagc	tcccgcggcg	gcggcaggtc	8460
agccgggagt	tcttgcaagg	tcacctcgca	gagtcggggc	agggcgcggg	gcaggtctag	8520
gtggtacctg	atctctaggg	gcgtgttggg	ggcggcgtcg	atggcttgca	ggagcccgcg	8580
gccccggggg	gcgacgacgg	tgccccgcgg	ggtggtgggtg	gtggtggcgg	tgcagctcag	8640
aagcggtgcc	gcgggcgggc	ccccggaggt	agggggggct	ccggtcccgc	gggcaggggc	8700
ggcagcggca	cgtcggcgtg	gagcgcgggc	aggagtgggt	gctgtgcccg	gaggttgctg	8760
gcgaaggcga	cgacgcggcg	gttgatctcc	tggatctggc	gcctctgcgt	gaagacgacg	8820
ggccccggtga	gcttgaacct	gaaagagagt	tcgacagaat	caatctcggt	gtcattgacc	8880
gcggcctggc	gcaggatctc	ctgcacgtct	cccagattgt	cttggtaggc	gatctcggcc	8940
atgaactgct	cgatctcttc	ctcctggagg	tctccgcgtc	cggcgcgttc	cacggtggcc	9000
gccaggtcgt	tggagatgcg	ccccatgagc	tgcgagaagg	cgttgagtcc	gccctcgttc	9060
cagactcggc	tgtagaccac	gccccctgg	tcatcgcggg	cgcgcatgac	cacctgcgcg	9120
aggttgagct	ccacgtgccg	cgcgaagacg	gcgtagtgtc	gcagacgctg	gaagaggtag	9180
ttgaggggtg	tggcggtgtg	ctcggccacg	aagaagttca	tgaccagcg	gcgcaacgtg	9240
gattcgttga	tgtcccccaa	ggcctccagc	cgttccatgg	cctcgtagaa	gtccacggcg	9300
aagttgaaaa	actgggagtt	gcgcgccgac	acggtcaact	cctcctccag	aagacggatg	9360
agctcggcga	cgggtgtcgcg	cacctcgcgc	tcgaaggcta	tggggatctc	ttcctccgct	9420
agcatcacca	cctcctctc	ttcctctct	tctggcactt	ccatgatggc	ttcctctct	9480
tcgggggggtg	gcggcggcgg	cgggtggggga	gggggcgctc	tcgcgccggc	gcggcgccac	9540
gggaggcggg	ccacgaagcg	cgcgatcatc	tccccgcggc	ggcggcgcac	ggtctcgggtg	9600
acggcgcggc	cgttctcccc	ggggcgcgag	tggaagacgc	cgcgggacac	ctggtgctgg	9660

ggcgggtggc	cgtgaggcag	cgagacggcg	ctgacgatgc	atctcaacaa	ttgctgcgta	9720
ggtacgccgc	cgagggacct	gagggagtcc	atatccaccg	gatccgaaaa	cctttcgagg	9780
aaggcgtcta	accagtcgca	gtcgcaaggt	aggctgagca	ccgtggcggg	cggcgggggg	9840
tggggggagt	gtctggcgga	ggtgctgctg	atgatgtaat	tgaagtaggc	ggacttgaca	9900
cggcggatgg	tcgacaggag	caccatgtcc	ttgggtccgg	cctgctggat	gcggaggcgg	9960
tcggctatgc	cccaggcttc	gttctggcat	cggcgcaggt	ccttgtagta	gtcttgcatg	10020
agcctttcca	ccggcacctc	ttctccttcc	tcttctgctt	cttccatgtc	tgcttcggcc	10080
ctggggcggc	gccgcgcccc	cctgcccccc	atgcgcgtga	ccccgaaccc	cctgagcggg	10140
tggagcaggg	ccaggtcggc	gacgacgcgc	tcggccagga	tggcctgctg	cacctgcgtg	10200
aggggtggttt	ggaagtcatc	caagtccacg	aagcggtggt	aggcgcccgt	gttgatgggtg	10260
taggtgcagt	tggccatgac	ggaccagttg	acggctctgg	ggcccggttg	cgacatctcg	10320
gtgtacctga	gtcgcgagta	ggcgcgggag	tcgaagacgt	agtcgttgca	agtccgcacc	10380
aggtactggg	agcccaccag	gaagtgcggc	ggcggctggc	ggtagagggg	ccagcgcagg	10440
gtggcggggg	ctccgggggc	caggtcttcc	agcatgaggc	ggtggtaggc	gtagatgtac	10500
ctggacatcc	aggtgatacc	cgcggcgggtg	gtggaggcgc	gcgggaagtc	gcgcacccgg	10560
ttccagatgt	tgcgcgaggg	cagaaagtgc	tccatggtag	gcgtgctctg	tccagtcaga	10620
cgcgcgcagt	cgttgatact	ctagaccagg	gaaaacgaaa	gccggtcagc	gggcactctt	10680
ccgtgggtctg	gtgaatagat	cgcaagggta	tcatggcgga	gggcctcggt	tcgagccccg	10740
ggtccggggc	ggacgggtccg	ccatgatcca	cgcggttacc	gcccgcgtgt	cgaacccagg	10800
tgtgcgacgt	cagacaacgg	tggagtgttc	cttttggcgt	ttttctggcc	gggcgcgggc	10860
gccgcgtaag	agactaagcc	gcgaaagcga	aagcagtaag	tggctcgctc	cccgtagccg	10920
gagggatcct	tgctaagggg	tgcgttgcgg	cgaaccccg	ttcgaatccc	gtactcgggc	10980
cggccggacc	cgcggtctag	gtgttggtat	ggcctcccc	tcgtataaag	accccgcttg	11040
cggattgact	ccggacacgg	ggacgagccc	cttttatttt	tgctttcccc	agatgcatcc	11100
ggtgctgcgg	cagatgcgcc	ccccgcccc	gcagcagcaa	caacaccagc	aagagcggca	11160
gcaacagcag	cgggagtcac	gcagggcccc	ctcaccacc	ctcggcgggc	cggccacctc	11220
ggcgtccgcg	gccgtgtctg	gcgcctgcgg	cggcggcggg	gggcgggctg	acgacccga	11280
ggagcccccg	cggcgcaggg	ccagacacta	cctggacctg	gaggagggcg	agggcctggc	11340
gcggctgggg	gcgccgtctc	ccgagcgcca	cccgcgggtg	cagctgaagc	gcgactcgcg	11400
cgaggcgtag	gtgcctcggc	agaacctgtt	cagggaccgc	gcgggcgagg	agcccagagg	11460
gatgcgggac	aggaggttca	gcgcagggcg	ggagctgcgg	caggggctga	accgcgagcg	11520
gctgctgcgc	gaggaggact	ttgagcccga	cgcgcgggac	gggatcagcc	ccgcgcgcgc	11580

gcacgtggcg gccgccgacc tgggtgacggc gtacgagcag acggtgaacc aggagatcaa	11640
cttccaaaag agtttcaaca accacgtgcg cacgctgggtg gcgcgcgagg aggtgaccat	11700
cgggctgatg cacctgtggg actttgtaag cgcgctgggtg cagaacccca acagcaagcc	11760
tctgacggcg cagctgttcc tgatagtgca gcacagcagg gacaacgagg cgtttagggg	11820
cgcgctgctg aacatcaccc agcccgaggg tcggtggctg ctggacctga ttaacatcct	11880
gcagagcata gtggtgcagg agcgcagcct gagcctggcc gacaaggtgg cggccatcaa	11940
ctactcgatg ctgagcctgg gcaagtttta cgcgcgcaag atctaccaga cgccgtacgt	12000
gcccatagac aaggaggtga agatcgacgg tttttacatg cgcatggcgc tgaaggtgct	12060
caccctgagc gacgacctgg gcgtgtaccg caacgagcgc atccacaagg ccgtgagcgt	12120
gagccggcgg cgcgagctga gcgaccgca gctgatgcac agcctgcagc gggcgctggc	12180
gggcgccggc agcggcgaca gggaggcgga gtcctacttc gatgcggggg cggacctgcg	12240
ctgggcgccc agccggcggg ccctggaggc cgcgggggtc cgcgaggact atgacgagga	12300
cggcgaggag gatgaggagt acgagctaga ggaggcgag tacctggact aaaccgcggg	12360
tgggtgtttcc ggtagatgca agaccgaac gtggtggacc cggcgctgcg ggcggctctg	12420
cagagccagc cgtccggcct taactcctca gacgactggc gacaggatcat ggaccgcac	12480
atgtcgctga cggcgcgtaa cccggacgcg ttccggcagc agccgcaggc caacaggctc	12540
tccgccatcc tggaggcggg ggtgcctgcg cgctcgaacc ccacgcacga gaaggtgctg	12600
gccatagtga acgcgctggc cgagaacagg gccatccgcc cggacgaggc cgggctggtg	12660
tacgacgcgc tgctgcagcg cgtggccccg tacaacagcg gcaacgtgca gaccaacctg	12720
gaccggctgg tgggggacgt gcgcgaggcg gtggcgagc gcgagcgcg ggatcggcag	12780
ggcaacctgg gctccatggt ggcgctgaat gccttcctga gcacgcagcc ggccaacctg	12840
ccgcgggggc aggaagacta caccaacttt gtgagcgcg tcgcgctgat ggtgaccgag	12900
acccccaga gcgaggtgta ccagtcgggc ccggactact tcttcagac cagcagacag	12960
ggcctgcaga cggatgaacct gagccaggct ttcaagaacc tgcgggggct gtggggcgctg	13020
aaggcgccca ccggcgaccg ggcgacgggtg tccagcctgc tgacgcccga ctcgcgcctg	13080
ctgctgctgc tgatcgcgcc gttcacggac agcggcagcg tgtcccggga cacctacctg	13140
gggcacctgc tgacctgta ccgcgaggcc atcgggcagg cgcaggtgga cgagcacacc	13200
ttccaggaga tcaccagcgt gagccgcgcg ctggggcagg aggacacgag cagcctggag	13260
gcgactctga actacctgct gaccaaccgg cggcagaaga ttccctcgct gcacagcctg	13320
acctccgagg aggagcgcat cttgcgctac gtgcagcaga gcgtgagcct gaacctgatg	13380
cgcgacgggg tgacgcccag cgtggcgctg gacatgaccg cgcgcaacat ggaaccgggc	13440
atgtacgccg cgcaccggcc ttacatcaac cgctgatgg actacctgca tcgcgcggcg	13500

gccgtgaacc	ccgagtactt	taccaacgcc	atcctgaacc	cgcactggct	cccgcgcgcc	13560
gggttctaca	gcgggggctt	cgaggtcccg	gagaccaacg	atggcttcct	gtgggacgac	13620
atggacgaca	gcgtgttctc	cccgcggccg	caggcgctgg	cggaagcgtc	cctgctgcgt	13680
ccaagaagg	aggaggagga	ggaggcgagt	cgccgccgcg	gcagcagcgg	cgtggcttct	13740
ctgtccgagc	tggggggcgg	agccgccgcg	cgccccgggt	ccctgggcgg	cagccccctt	13800
ccgagcctgg	tggggtctct	gcacagcgag	cgcaccaccc	gccctcggct	gctgggcgag	13860
gacgagtacc	tgaataactc	cctgctgcag	ccggtgcggg	agaaaaacct	gcctcccgcc	13920
ttccccaaca	acgggataga	gagcctgggt	gacaagatga	gcagatggaa	gacctatgcg	13980
caggagcaca	gggacgcgcc	tgcgctccgg	ccgcccacgc	ggcgccagcg	ccacgaccgg	14040
cagcgggggc	tgggtgtgga	tgacgaggac	tccgcggacg	atagcagcgt	gctggacctg	14100
ggagggagcg	gcaacccggt	cgcgcacctg	cgcccccgcc	tggggaggat	gttttaaaaa	14160
aaaaaaaaaa	aagcaagaag	catgatgcaa	aaattaaata	aaactcacca	aggccatggc	14220
gaccgagcgt	tggtttcttg	tgttcccttc	agtatgcggc	gcgcggcgat	gtaccaggag	14280
ggacctctc	cctcttacga	gagcgtgggt	ggcgggcgcg	cgggggcgcc	ctcttctccc	14340
tttgcgtcgc	agctgctgga	gccgccgtac	gtgcctccgc	gctacctgcg	gcctacgggg	14400
gggagaaaca	gcatccgtta	ctcgagctg	gcgcccctgt	tcgacaccac	ccgggtgtac	14460
ctggtggaca	acaagtcggc	ggacgtggcc	tccctgaact	accagaacga	ccacagcaat	14520
tttttgacca	cggatcatcca	gaacaatgac	tacagcccg	gcgaggccag	caccagacc	14580
atcaatctgg	atgaccggtc	gcactggggc	ggcgacctga	aaaccatcct	gcacaccaac	14640
atgccaacg	tgaacgagtt	catgttcacc	aataagttca	aggcgcgggg	gatggtgtcg	14700
cgctcgaca	ccaaggaaga	ccgggtggag	ctgaagtacg	agtgggtgga	gttcgagctg	14760
ccagagggca	actactccga	gaccatgacc	attgacctga	tgaacaacgc	gatcgtggag	14820
cactatctga	aagtgggcag	gcagaacggg	gtcctggaga	gcgacatcgg	ggtcaagttc	14880
gacaccagga	acttccgcct	ggggctggac	ccgctgaccg	ggctggttat	gcccgggggt	14940
tacaccaacg	aggccttcca	tcccgcacatc	atcctgctgc	ccggctgcgg	ggtggacttc	15000
acttacagcc	gcctgagcaa	cctcctgggc	atccgcaagc	ggcagccctt	ccaggagggc	15060
ttcaggatca	cctacgagga	cctggagggg	ggcaacatcc	ccgcgctcct	cgatgtggag	15120
gcctaccagg	atagcttgaa	ggaaaatgag	gcgggacagg	aggataccgc	ccccgcgcgc	15180
tccgccgccg	ccgagcaggg	cgaggatgct	gctgacaccg	cgcccgcgga	cggggcagag	15240
gccgaccccg	ctatggtggt	ggaggctccc	gagcaggagg	aggacatgaa	tgacagtgcg	15300
gtgcgcggag	acaccttcgt	caccgggggg	gaggaaaagc	aagcggaggc	cgaggccgcg	15360
gccgaggaaa	agcaactggc	ggcagcagcg	gcggcgggcg	cgttggccgc	ggcggaggct	15420

gagtctgagg	ggaccaagcc	cgccaaggag	cccgtgatta	agccccctgac	cgaagatagc	15480
aagaagcgca	gttacaacct	gctcaaggac	agcaccaaca	ccgcgtaccg	cagctgggtac	15540
ctggcctaca	actacggcga	cccgtcgcgc	ggggtgcgct	cctggaccct	gctgtgcacg	15600
ccggacgtga	cctgcggctc	ggagcaggtg	tactggtcgc	tgcccgacat	gatgcaagac	15660
cccgtgacct	tccgctccac	gcggcaggtc	agcaacttcc	cggtggtggg	cgccgagctg	15720
ctgccccgtgc	actccaagag	cttctacaac	gaccaggccg	tctactccca	gctcatccgc	15780
cagttcacct	ctctgaccca	cgtgttcaat	cgctttcctg	agaaccagat	tctggcgcg	15840
ccgccccgcc	ccaccatcac	caccgtcagt	gaaaacgttc	ctgctctcac	agatcacggg	15900
acgctaccgc	tgcgcaacag	catcggagga	gtccagcgag	tgaccgttac	tgacgccaga	15960
cgccgcacct	gcccctacgt	ttacaaggcc	ttgggcatag	tctcgccgcg	cgctccttcc	16020
agccgcactt	tttgagcaac	accaccatca	tgtccatcct	gatctcacc	agcaataact	16080
ccggctgggg	actgctgcgc	gcgcccagca	agatgttcgg	aggggcgagg	aagcgttccg	16140
agcagcaccc	cgtgcgcgtg	cgcgggcact	tccgcgcccc	ctggggagcg	cacaaacgcg	16200
gccgcgcggg	gcgcaccacc	gtggacgacg	ccatcgactc	ggtggtggag	caggcgcgca	16260
actacaggcc	cgcggtctct	accgtggacg	cggccatcca	gaccgtggtg	cggggcgcgc	16320
ggcggtacgc	caagctgaag	agccgcggga	agcgcggtgg	ccgccgccac	cgccgcgcgac	16380
ccggggccgc	cgccaaacgc	gccgccgcgg	ccctgcttcg	ccggggccaag	cgcacggggc	16440
gccgcgccgc	catgaggggc	gcgcgcgcgt	tggccgcggg	catcacgcgc	gccaccatgg	16500
ccccccgtac	ccgaagacgc	gcggccgcgg	ccgccgcgcg	cgccatcagt	gacatggcca	16560
gcaggcgccg	gggcaacgtg	tactgggtgc	gcgactcggt	gaccggcacg	cgcggtgccg	16620
tgcgcttccg	ccccccgcgg	acttgagatg	atgtgaaaaa	acaacactga	gtctcctgct	16680
gttgtgtgta	tcccagcggc	ggcggcgcgc	gcagcgtcac	gtccaagcgc	aaaatcaaag	16740
aagagatgct	ccaggtcgtc	gcgccggaga	tctatggggc	cccgaagaag	gaagagcagg	16800
attcgaagcc	ccgaagata	aagcgggtca	aaaagaaaaa	gaaagatgat	gacgatgccg	16860
atggggaggt	ggagttcctg	cgcgccacgg	cgcccaggcg	cccgggtgcag	tggaagggcc	16920
ggcgcgtaaa	gcgcgtcctg	cgccccggca	ccgcggtggt	cttcacgccc	ggcgagcgct	16980
ccaccgggac	tttcaagcgc	gtctatgacg	aggtgtacgg	cgacgaagac	ctgctggagc	17040
aggccaacga	gcgcttcgga	gagtttgctt	acgggaagcg	tcagcggggc	ctggggaagg	17100
aggacctgct	ggcgctgccg	ctggaccagg	gcaacccac	ccccagtctg	aagcccgtga	17160
ccctgcagca	ggtgctgccg	agcagcgcac	cctccgaggc	gaagcggggg	ctgaagcgcg	17220
agggcgggca	cctggcgccc	accgtgcagc	tcatgggtgc	caagcggcag	aggctggagg	17280
atgtgctgga	gaaaatgaaa	gtagaccccg	gtctgcagcc	ggacatcagg	gtccgcccc	17340

tcaagcaggt	ggcgccgggc	ctcggcgtgc	agaccgtgga	cgtgggtcatc	cccaccggca	17400
actccccgc	cgccgccacc	actaccgctg	cctccacgga	catggagaca	cagaccgatc	17460
ccgccgcagc	cgcagccgca	gccgccgccc	cgacctctc	ggcggaggtg	cagacggacc	17520
cctggctgcc	gccggcgatg	tcagctcccc	gcgcgcgtcg	cgggcgcagg	aagtacggcg	17580
ccgccaacgc	gctcctgccc	gagtacgcct	tgcatccttc	catcgcgccc	acccccggct	17640
accgaggcta	tacctaccgc	ccgcgaagag	ccaagggttc	caccgcgcgt	ccccgccgac	17700
gcgccgccgc	caccaccgcg	cgccgccgcc	gcagacgcca	gcccgcactg	gctccagtct	17760
ccgtgaggaa	agtggcgcg	gacggacaca	ccctgggtgt	gcccagggcg	cgctaccacc	17820
ccagcatcgt	ttaaaagcct	gttgtggttc	ttgcagatat	ggccctcact	tgccgcctcc	17880
gtttccccgt	gccgggatac	cgaggaggaa	gatcgcgccg	caggaggggt	ctggccggcc	17940
gcggcctgag	cggaggcagc	cgccgcgcgc	accggcgggc	acgcgccacc	agccgacgca	18000
tgcgcgggcg	ggtgctgccc	ctgttaatcc	ccctgatcgc	cgcggcgatc	ggcgccgtgc	18060
ccgggatcgc	ctccgtggcc	ttgcaagcgt	cccagaggca	ttgacagact	tgcaaacttg	18120
caaatatgga	aaaaaaaaacc	ccaataaaaa	agtctagact	ctcacgctcg	cttgggtcctg	18180
tgactatttt	gtagaatgga	agacatcaac	tttgcgctcg	tggccccgcg	tcacggctcg	18240
cgcccgttcc	tgggacactg	gaacgatatc	ggcaccagca	acatgagcgg	tggcgcttcc	18300
agttggggct	ctctgtggag	cggcattaaa	agtatcgggt	ctgccgttaa	aaattacggc	18360
tccccggcct	ggaacagcag	cacgggccag	atgttgagag	acaagttgaa	agagcagaac	18420
ttccagcaga	aggtggtgga	gggcctggcc	tccggcatca	acggggtggt	ggacctggcc	18480
aaccaggccg	tgcagaataa	gatcaacagc	agactggacc	cccggccgcc	ggtggaggag	18540
gtgccgccgg	cgctggagac	ggtgtcccc	gatgggcgtg	gcgagaagcg	cccgcggccc	18600
gataggggaag	agaccactct	ggtcacgcag	accgatgagc	cgcccccgta	tgaggaggcc	18660
ctgaagcaag	gtctgcccac	cacgcggccc	atcgcgccca	tggccaccgg	ggtggtgggc	18720
cgccacaccc	ccgccacgct	ggacttgctc	ccgcccgccg	atgtgccgca	gcagcagaag	18780
gcggcacagc	cgggcccgcg	cgcgaccgcc	tcccgttcct	ccgccgggtc	tctgcgcgc	18840
gcggccagcg	gcccccgcg	gggggtcgcg	aggcacggca	actggcagag	cacgctgaac	18900
agcatcgtgg	gtctgggggt	gcggtccgtg	aagcgccgcc	gatgctactg	aatagcttag	18960
ctaacgtgtt	gtatgtgtgt	atgcgcccta	tgtcgccgcc	agaggagctg	ctgagtcgcc	19020
gccgttcgcg	cgcccaccac	caccgccact	ccgcccctca	agatggcgac	cccatcgatg	19080
atgccgcagt	ggtcgtacat	gcacatctcg	ggccaggacg	cctcggagta	cctgagcccc	19140
gggctggtgc	agttcgcccc	cgccaccgag	agctacttca	gcctgagtaa	caagtttagg	19200
aacccccacg	tggcgcccac	gcacgatgtg	accaccgacc	ggtctcagcg	cctgacgctg	19260

cggttcattc	ccgtggaccg	cgaggacacc	gcgtactcgt	acaaggcgcg	gttcaccctg	19320
gccgtgggcg	acaaccgcgt	gctggacatg	gcctccacct	actttgacat	ccgcgggggtg	19380
ctggaccggg	gtcccacttt	caagccctac	tctggcaccg	cctacaactc	cctggccccc	19440
aaggggcgctc	ccaactcctg	cgagtgggag	caagaggaaa	ctcaggcagt	tgaagaagca	19500
gcagaagagg	aagaagaaga	tgctgacggg	caagctgagg	aagagcaagc	agctaccaaa	19560
aagactcatg	tatatgctca	ggctccccct	tctggcgaaa	aaattagtaa	agatgggtctg	19620
caaataggaa	cggacgctac	agctacagaa	caaaaaccta	tttatgcaga	ccctacattc	19680
cagcccgaac	cccaaactcg	ggagtcccag	tggaatgagg	cagatgctac	agtcgcgggc	19740
ggtagagtgc	taaagaaatc	tactcccatg	aaaccatgct	atggttccta	tgcaagaccc	19800
acaaatgcta	atggagggtca	gggtgtacta	acggcaaattg	cccagggaca	gctagaatct	19860
caggttgaaa	tgcaattctt	ttcaacttct	gaaaacgccc	gtaacgaggc	taacaacatt	19920
cagcccaaatt	tggtgctgta	tagtgaggat	gtgcacatgg	agaccccgga	tacgcacctt	19980
tcttacaagc	ccgcaaaaag	cgatgacaat	tcaaaaatca	tgctgggtca	gcagtccatg	20040
cccaacagac	ctaattacat	cggcttcaga	gacaacttta	tcggcctcat	gtattacaat	20100
agcactggca	acatgggagt	gcttgacagg	caggcctctc	agttgaatgc	agtgggtggac	20160
ttgcaagaca	gaaacacaga	actgtcctac	cagctcttgc	ttgattccat	gggtgacaga	20220
accagatact	tttccatgtg	gaatcaggca	gtggacagtt	atgaccacaga	tgttagaatt	20280
attgaaaatc	atggaactga	agacgagctc	cccaactatt	gtttccctct	gggtggcata	20340
ggggtaactg	acacttacca	ggctgttaaa	accaacaatg	gcaataacgg	gggccagggtg	20400
acttggacaa	aagatgaaac	ttttgcagat	cgcaatgaaa	taggggtggg	aaacaatttc	20460
gctatggaga	tcaacctcag	tgccaacctg	tgagaaaact	tcctgtactc	caacgtggcg	20520
ctgtacctac	cagacaagct	taagtacaac	ccctccaatg	tggacatctc	tgacaacccc	20580
aacacctacg	attacatgaa	caagcgagtg	gtggccccgg	ggctgggtgga	ctgctacatc	20640
aacctgggcg	cgcgctggtc	gctggactac	atggacaacg	tcaaccctt	caaccaccac	20700
cgcaatgcgg	gcctgcgcta	ccgctccatg	ctcctgggca	acgggcgcta	cgtgcccttc	20760
cacatccagg	tgccccagaa	gttctttgcc	atcaagaacc	tcctcctcct	gccgggctcc	20820
tacacctacg	agtggaactt	caggaaggat	gtcaacatgg	tcctccagag	ctctctgggt	20880
aacgatctca	gggtggacgg	ggccagcatc	aagtctgaga	gcattctgct	ctacgccacc	20940
ttcttcccca	tggcccacaa	cacggcctcc	acgctcgagg	ccatgctcag	gaacgacacc	21000
aacgaccagt	ccttcaatga	ctacctctcc	gccgccaaca	tgctctaccc	catacccgcc	21060
aacgccacca	acgtccccat	ctccatcccc	tcgcgcaact	ggggggcctt	ccgcgggtgg	21120
gccttcaccc	gcctcaagac	caaggagacc	ccctccctgg	gctcgggatt	cgaccctac	21180

tacacctact	cggggtccat	tccctacctg	gacggcacct	tctacctcaa	ccacactttc	21240
aagaaggtct	cggtcacctt	cgactcctcg	gtcagctggc	cgggcaacga	ccgtctgctc	21300
acccccaacg	agttcgagat	caagcgctcg	gtcgacgggg	agggctacaa	cgtggcccag	21360
tgcaacatga	ccaaggactg	gttcctggtc	cagatgctgg	ccaactacaa	catcggttac	21420
cagggcttct	acatcccaga	gagctacaag	gacaggatgt	actccttctt	caggaacttc	21480
cagcccatga	gccggcaggt	ggtggaccag	accaagtaca	aggactacca	ggaggtgggc	21540
atcatccacc	agcacaacaa	ctcgggcttc	gtgggctacc	tcgccccac	catgcgcgag	21600
ggacaggcct	accccgccaa	cttcccctat	ccgctcatag	gcaagaccgc	ggtcgacagc	21660
atcaccaga	aaaagtctct	ctgcgaccgc	accctctggc	gcatccccct	ctccagcaac	21720
ttcatgtcca	tgggtgcgct	ctcggacctg	ggccagaact	tgtctacgc	caactccgcc	21780
cacgcctcg	acatgacctt	cgaggctcgac	cccatggacg	agcccaccct	tctctatggt	21840
ctgttcgaag	tctttgacgt	ggtccgggtc	caccagccgc	accgcggcgt	catcgagacc	21900
gtgtacctgc	gtacgccctt	ctcggccggc	aacgccacca	cctaaagaag	caagccgcag	21960
tcatcgccgc	ctgcatgccg	tcgggttcca	ccgagcaaga	gtcaggggc	atcgtcagag	22020
acctgggatg	cgggccctat	tttttgggca	ccttcgacaa	gcgcttcctt	ggctttgtct	22080
ccccacacaa	gctggcctgc	gccatcgctc	acacggccgg	ccgcgagacc	gggggcgtgc	22140
actggctggc	cttcgcctgg	aaccgcgct	ccaaaacatg	cttcctcttt	gacccttcg	22200
gcttttcgga	ccagcggctc	aagcaaactt	acgagttcga	gtacgagggc	ttgctgcgtc	22260
gcagcgccat	cgctcctcg	cccgaccgct	gcgtcacctt	cgaaaagtcc	accagaccg	22320
tcagggggcc	cgactcggcc	gcctgcggtc	tcttctgctg	catgtttctg	cacgcctttg	22380
tgactggcc	tcagagtccc	atggaccgca	acccaccat	gaacttgctg	acgggggtgc	22440
ccaactccat	gctccagagc	ccccaggctg	agcccaccct	gcgccgcaac	caggagcagc	22500
tctacagctt	cctggagcgc	cactgcctt	acttcgcgcg	ccacagcgca	cagatcagga	22560
gggccacctc	cttctgccac	ttgcaagaga	tgcaagaagg	gtaataacga	tgtacacact	22620
ttttttctca	ataaatggca	tctttttatt	tatacaagct	ctctggggta	ttcatttccc	22680
accaccacc	gccgttgctg	ccatctggct	ctatttagaa	atcgaaaggg	ttctgccggg	22740
agtcgccgtg	cgccacgggc	agggacacgt	tgcgatactg	gtagcgggtg	ccccacttga	22800
actcgggcac	caccaggcga	ggcagctcgg	ggaagttttc	gctccacagg	ctgcgggtca	22860
gcaccagcgc	gttcatcagg	tcgggcgcgcg	agatcttgaa	gtcgcagttg	gggcgcgcgc	22920
cctgcgcgcg	cgagttgcgg	tacaccgggt	tgcagcactg	gaacaccaac	agcgccgggt	22980
gcttcacgct	ggccagcacg	ctgcggtcgg	agatcagctc	ggcgccagg	tcctccgcgt	23040
tgctcagcgc	gaacggggtc	atcttgggca	cttgccgcgc	caggaagggc	gcgtgccccg	23100

gtttcgagtt	gcagtcgcag	cgcagcggga	tcagcaggtg	cccgtagccg	gactcggcgt	23160
tggggtacag	cgcgcgcatg	aaggcctgca	tctggcgga	ggccatctgg	gccttggcgc	23220
cctccgagaa	gaacatgccg	caggacttgc	ccgagaactg	gtttgcgggg	cagctggcgt	23280
cgtgcaggca	gcagcgcgcg	tcggtgttgg	cgatctgcac	cacgttgccg	ccccaccggt	23340
tcttcacgat	cttggccttg	gacgattgct	ccttcagcgc	gcgctgcccg	ttctcgctgg	23400
tcacatccat	ctcgatcaca	tgttccttgt	tcaccatgct	gctgccgtgc	agacacttca	23460
gctcgccctc	cgtctcgggtg	cagcgggtgct	gccacagcgc	gcagcccgtg	ggctcgaaag	23520
acttgtaggt	cacctccgcg	aaggactgca	ggtacccttg	caaaaagcgg	cccatcatgg	23580
tcacgaagg	cttggtgctg	ctgaagggtca	gctgcagccc	gcggtgctcc	tcgttcagcc	23640
aggtcttgca	cacggccgcc	agcgcctcca	cctggtcggg	cagcatcttg	aagttcacct	23700
tcagctcatt	ctccacgtgg	tacttggtcca	tcagcgtgcg	cgcgcctcc	atgcccttct	23760
cccaggccga	caccagcggc	aggctcacgg	ggttcttcac	catcacctg	gccgcgcct	23820
cgcgcgcgt	ttcgctttcc	gccccgctgt	tctcttcctc	ttcctcctct	tcctcgccgc	23880
cgcctactcg	cagccccgc	accacgggggt	cgtcttcctg	caggcgctgc	accttgcgct	23940
tgccgttgcg	ccctgcttg	atgcgcacgg	gcgggttgct	gaagcccacc	atcaccagcg	24000
cggcctcttc	ttgctcgctc	tcgctgtcca	gaatgacctc	cggggagggg	gggttggtca	24060
tcctcagtac	cgaggcacgc	ttctttttct	tcctgggggc	gttcgccagc	tcgcgggctg	24120
cggccgctgc	cgaggtcgaa	ggccgagggc	tgggcgtgcg	cggcaccagc	gcgtcctgcg	24180
agccgtcctc	gtcctcctcg	gactcgagac	ggaggcgggc	ccgcttcttc	gggggcgcgc	24240
ggggcggcgg	aggcggcggc	ggcgacggag	acggggacga	gacatcgctc	agggtgggtg	24300
gacggcgggc	cgcgccgcgt	ccgcgctcgg	gggtggtctc	gcgctggtcc	tcttcccagc	24360
tggccatctc	ccactgctcc	ttctcctata	ggcagaaaga	gatcatggag	tctctcatgc	24420
gagtcgagaa	ggaggaggac	agcctaaccg	ccccctctga	gccctccacc	accgcgcgca	24480
ccaccgccaa	tgccgccgcg	gacgacgcgc	ccaccgagac	caccgccagt	accaccctcc	24540
ccagcgacgc	acccccgctc	gagaatgaag	tgctgatcga	gcaggaccgc	ggttttgtga	24600
gcggagagga	ggatgaggtg	gatgagaagg	agaaggagga	ggtcgccgcc	tcagtgccaa	24660
aagaggataa	aaagcaagac	caggacgacg	cagataagga	tgagacagca	gtcgggcggg	24720
ggaacggaag	ccatgatgct	gatgacggct	acctagacgt	gggagacgac	gtgctgctta	24780
agcacctgca	ccgccagtgc	gtcatcgtct	gcgacgcgct	gcaggagcgc	tgccaagtgc	24840
ccctggacgt	ggcggaggtc	agccgcgcct	acgagcggca	cctcttcgcg	ccgcacgtgc	24900
cccccaagcg	ccgggagaa	ggcacctgcg	agcccaacct	gcgtctcaac	ttctaccgcg	24960
tcttcgcgggt	acccgaggtg	ctggccacct	accacatctt	tttccaaaac	tgcaagatcc	25020

ccctctcctg	ccgcgcacaac	cgcacccgcg	ccgacaaaac	cctgaccctg	cggcagggcg	25080
cccacatacc	tgatatcgcc	tctctggagg	aagtgcccaa	gatcttcgag	ggtctcggtc	25140
gcgacgagaa	acgggcggcg	aacgctctgc	acggagacag	cgaaaacgag	agtcactcgg	25200
gggtgctggt	ggagctcgag	ggcgacaacg	cgcgccctggc	cgtactcaag	cgcagcatag	25260
aggtcaccca	ctttgcctac	ccggcgctca	acctgcccc	caaggatcatg	agtgtggtca	25320
tgggcgagct	catcatgcgc	cgcgcccagc	cctggccgcg	ggatgcaaac	ttgcaagagt	25380
cctccgagga	aggcctgccc	gcggtcagcg	acgagcagct	ggcgcgctgg	ctggagaccc	25440
gcgaccccg	gcagctggag	gagcggcgca	agctcatgat	ggccgcgggtg	ctggtcaccg	25500
tggagctcga	gtgtctgcag	cgcttcttcg	cggaccccg	gatgcagcgc	aagctcgagg	25560
agaccctgca	ctacaccttc	cgcaggggct	acgtgcgcca	ggcctgcaag	atctccaacg	25620
tggagctctg	caacctggtc	tcctacctgg	gcctcctgca	cgagaaccgc	ctcgggcaga	25680
acgtcctgca	ctccaccctc	aaaggggagg	cgcgccgcga	ctacatccgc	gactgcgcct	25740
acctcttcct	ctgctacacc	tggcagacgg	ccatgggggt	ctggcagcag	tgcttgagg	25800
agcgcaacct	caaggagctg	gaaaagctcc	tcaagcgcac	cctcagggac	ctctggacgg	25860
gcttcaacga	gcgctcgggtg	gccgccgcgc	tggcggacat	catctttccc	gagcgctgc	25920
tcaagaccct	gcagcagggc	ctgcccagct	tcaccagcca	gagcatgctg	cagaacttca	25980
ggactttcat	cctggagcgc	tcgggcatcc	tgccggccac	ttgctgcgcg	ctgcccagcg	26040
acttcgtgcc	catcaagtac	agggagtgcc	cgccgccgct	ctggggccac	tgctacctct	26100
tccagctggc	caactacctc	gcctaccact	cggacctcat	ggaagacgtg	agcggcgagg	26160
gcctgctcga	gtgccactgc	cgctgcaacc	tctgcacgcc	ccaccgctct	ctagtctgca	26220
accgcagct	gctcagcgag	agtcagatta	tcggtacctt	cgagctgcag	ggtccctcgc	26280
ctgacgagaa	gtccgcggct	ccagggctga	aactcactcc	ggggctgtgg	acttccgcct	26340
acctacgcaa	atgtgtacct	gaggactacc	acgcccacga	gatcagggttc	tacgaagacc	26400
aatcccgc	gcccaaggcg	gagctcaccg	cctgcgtcat	caccagggg	cacatcctgg	26460
gccaattgca	agccatcaac	aaagcccgc	gagagttctt	gctgaaaaag	ggtcgggggg	26520
tgtacctgga	ccccagtc	ggcgaggagc	taaaccgcct	accccgccg	ccgccccagc	26580
agcgggacct	tgcttcccag	gatggcaccc	agaaagaagc	agcagccgc	gccgcgcgcg	26640
cagccataca	tgcttctgga	ggaagaggag	gaggactggg	acagtcaggc	agaggaggtt	26700
tcggacgagg	agcaggagga	gatgatggaa	gactgggagg	aggacagcag	cctagacgag	26760
gaagcttcag	aggccgaaga	ggtggcagac	gcaacaccat	cgccctcggt	cgcagcccc	26820
tcgccggggc	ccctgaaatc	ctccgaacct	agcaccagcg	ctataacctc	cgctcctccg	26880
gcgccgggcg	caccgcgccg	cagacccaac	cgtagatggg	acaccacagg	aaccgggggtc	26940

ggtaagtcca	agtgcccgcc	gccgccaccg	cagcagcagc	agcagcagcg	ccagggctac	27000
cgctcgtggc	gcggggcaca	gaacgccata	gtcgcttgct	tgcaagactg	cgggggcaac	27060
atctctttcg	cccgccgctt	cctgctattc	caccacgggg	tcgcctttcc	ccgcaatgtc	27120
ctgcattact	accgtcatct	ctacagcccc	tactgcagcg	gcgaccaga	ggcggcagcg	27180
gcagccacag	cggcgaccac	cacctaggaa	gatatacctc	gcggggcaaga	cagcggcagc	27240
agcggccagg	agaccgcggg	cagcagcggc	gggagcgggtg	ggcgactgc	gcctctcgcc	27300
caacgaaccc	ctctcgaccc	gggagctcag	acacaggatc	ttccccactt	tgtatgccat	27360
cttccaacag	agcagaggcc	aggagcagga	gctgaaaata	aaaaacagat	ctctgcgctc	27420
cctcacccgc	agctgtctgt	atcacaaaag	cgaagatcag	cttcggcgca	cgctggagga	27480
cgcgagggca	ctcttcagca	aatactgcgc	gctcactctt	aaagactagc	tccgcgccct	27540
tctcgaattt	aggcgggaga	aaactacgtc	atcgccggcc	gccgcccagc	ccgcccagcc	27600
gagatgagca	aagagattcc	cacgccatac	atgtggagct	accagccgca	gatgggactc	27660
gcggcgggag	cggcccagga	ctactccacc	cgcatgaact	acatgagcgc	gggacccac	27720
atgatctcac	aggtcaacgg	gatccgcgcc	cagcgaaacc	aaatactgct	ggaacaggcg	27780
gccatcaccg	ccacgccccg	ccataatctc	aacccccgaa	attggccccg	cgccctcgtg	27840
taccaggaaa	ccccctccgc	caccaccgta	ctacttccgc	gtgacgcca	ggccgaagtc	27900
cagatgacta	actcaggggc	gcagctcgcg	ggcggctttc	gtcacggggc	gcggccgctc	27960
cgaccaggta	taagacacct	gatgatcaga	ggccgaggta	tccagctcaa	cgacgagtcg	28020
gtgagctctt	cgctcggctc	ccgtccggac	ggaactttcc	agctcgccgg	atccggccgc	28080
tcttcgttca	cgccccgcca	ggcgtacctg	actctgcaga	cctcgtcctc	ggagccccgc	28140
tccggcggca	tcggaaccct	ccagtctctg	gaggagtctg	tgccctcggg	ctacttcaac	28200
cccttctcgg	gacctcccgg	acgtacccc	gaccagttca	ttccgaactt	tgacgcgggtg	28260
aaggactcgg	cggacggcta	cgactgaatg	tcagggtgtc	aggcagagca	gcttcgcctg	28320
agacacctcg	agcactgccg	ccgccacaag	tgcttcgccc	gcggttcttg	tgagttctgc	28380
tactttcagc	tacccgagga	gcataccgag	gggccggcgc	acggcgctcc	cctgaccacc	28440
cagggcgagg	ttacctgttc	cctcatccgg	gagtttacc	tccgtcccct	gctagtggag	28500
cgggagcggg	gtccctgtgt	cctaactatc	gcctgcaact	gccctaacc	tggattacat	28560
caagatcttt	gctgtcatct	ctgtgctgag	tttaataaac	gctgagatca	gaatctactg	28620
gggctcctgt	cgccatcctg	tgaacgccac	cgtcttcacc	caccccgacc	aggcccaggc	28680
gaacctcacc	tgcggtctgc	atcggagggc	caagaagtac	ctcacctggg	acttcaacgg	28740
cacccctttt	gtggtttaca	acagcttcga	cggggacgga	gtctccctga	aagaccagct	28800
ctccgggtctc	agctactcca	tccacaagaa	caccaccctc	caactcttcc	ctccctacct	28860

gccgggaacc	tacgagtgcg	tcaccggccg	ctgcacccac	ctcaccgcgc	tgatcgtaaa	28920
ccagagcttt	ccgggaacag	ataactccct	cttccccaga	acaggaggtg	agctcaggaa	28980
actccccggg	gaccagggcg	gagacgtacc	ttcgaccctt	gtgggggttag	gattttttat	29040
taccggggtg	ctggctcttt	taatcaaagt	ttccttgaga	tttgttcttt	ccttctacgt	29100
gtatgaacac	ctcaacctcc	aataactcta	ccctttcttc	ggaatcaggt	gacttctctg	29160
aaatcgggct	tggtgtgctg	cttactctgt	tgattttttt	ccttatcata	ctcagccttc	29220
tgtgcctcag	gctcgccgc	tgctgcgcac	acatctatat	ctactgctgg	ttgctcaagt	29280
gcaggggtcg	ccaccaaga	tgaacaggta	catggtccta	tcgatcctag	gcctgctggc	29340
cctggcgggc	tgacgcgcg	ccaaaaaaga	gattaccttt	gaggagcccg	cttgcaatgt	29400
aactttcaag	cccgaggggtg	accaatgcac	cacctcgtc	aatgcgtta	ccaatcatga	29460
gaggctgcgc	atcgactaca	aaaacaaaac	tggccagttt	gcggtctata	gtgtgtttac	29520
gcccggagac	ccctctaact	actctgtcac	cgtcttccag	ggcggacagt	ctaagatatt	29580
caattacact	ttcccttttt	atgagttatg	cgatgcggtc	atgtacatgt	caaaacagta	29640
caacctgtgg	cctccctctc	cccaggcgtg	tgtggaaaat	actgggtctt	actgctgtat	29700
ggctttcgca	atcactacgc	tcgctcta	ctgcacgggtg	ctatacataa	aattcaggca	29760
gaggcgaatc	tttatcgatg	aaaagaaaat	gccttgatcg	ctaacaccgg	ctttctatct	29820
gcagaatgaa	tgcaatcacc	tccctactaa	tcaccaccac	cctccttgcg	attgcccattg	29880
ggttgacacg	aatcgaagtg	ccagtggggt	ccaatgtcac	catggtgggc	cccgcgggca	29940
attccaccct	catgtgggaa	aaatttgtcc	gcaatcaatg	ggttcatttc	tgctctaacc	30000
gaatcagtat	caagcccaga	gccatctgcg	atgggcaaaa	tctaactctg	atcaatgtgc	30060
aaatgatgga	tgctgggtac	tattacgggc	agcggggaga	aatcattaat	tactggcgac	30120
cccacaagga	ctacatgctg	catgtagtcg	aggcacttcc	cactaccacc	cccactacca	30180
ccttctccac	caccaccacc	actactacta	ctactactac	tactactact	actaccacta	30240
ccgctgcccc	ccataccgc	aaaagcacca	tgattagcac	aaagccccct	cgtgctcact	30300
cccacgcccg	cgggcccac	ggtgcgacct	cagaaaccac	cgagctttgc	ttctgccaat	30360
gcactaacgc	cagcgctcat	gaactgttcg	acctggagaa	tgaggatgtc	cagcagagct	30420
ccgcttgcc	gaccaggag	gctgtggagc	ccgttgccct	gaagcagatc	ggtgattcaa	30480
taattgactc	ttcttctttt	gccactcccc	aataccctcc	cgattctact	ttccacatca	30540
cgggtaccaa	agaccctaac	ctctctttct	acctgatgct	gctgctctgt	atctctgtgg	30600
tctcttccgc	gctgatgtta	ctggggatgt	tctgctgcct	gatctgccgc	agaaagagaa	30660
aagctcgctc	tcagggccaa	ccactgatgc	ccttccccta	ccccccgat	tttgagata	30720
acaagatatg	agctcgctgc	tgacactaac	cgctttacta	gcctgcgctc	taacccttgt	30780

cgcttgcgac	tcgagattcc	acaatgtcac	agctgtggca	ggagaaaatg	ttactttcaa	30840
ctccacggcc	gatacccagt	ggtcgtggag	tggtctcaggt	agctacttaa	ctatctgcaa	30900
tagctccact	tcccccgga	tatccccaac	caagtaccaa	tgcaatgcca	gcctgttcac	30960
cctcatcaac	gcttcacccc	tggaacaatgg	actctatgta	ggctatgtac	cctttggtgg	31020
gcaaggaaaag	accacgcgtt	acaacctgga	agttcgccag	cccagaaacca	ctaccaagc	31080
ttctcccacc	accaccacca	ccaccaccat	caccagcagc	agcagcagca	gcagccacag	31140
cagcagcagc	agattattga	ctttggtttt	ggccagctca	tctgccgcta	cccaggccat	31200
ctacagctct	gtgcccga	ccactcagat	ccaccgcca	gaaacgacca	ccgccaccac	31260
cctacacacc	tccagcgatc	agatgccgac	caacatcacc	cccttggttc	ttcaaattggg	31320
acttacaagc	cccactccaa	aaccagtgga	tgcgcccgag	gtctccgccc	tcgtcaatga	31380
ctgggcgggg	ctgggaatgt	ggtggttcgc	cataggcatg	atggcgctct	gcctgcttct	31440
gctctggctc	atctgctgcc	tccaccgcag	gcgagccaga	ccccccatct	atagacccat	31500
cattgtcctg	aaccccgata	atgatgggat	ccatagattg	gatggcctga	aaaacctact	31560
tttttctttt	acagtatgat	aaattgagac	atgcctcgca	ttttcttgta	catgttcctt	31620
ctcccacctt	ttctgggggtg	ttctacgctg	gccgctgtgt	ctcacctgga	ggtagactgc	31680
ctctcaccct	tcactgtcta	cctgctttac	ggattgggtca	ccctcactct	catctgcagc	31740
ctaatacacag	taatcatcgc	cttcatccag	tgcatgtgatt	acatctgtgt	gcgcctcgca	31800
tacttcagac	accaccgcga	gtaccgagac	aggaacattg	cccaacttct	aagactgctc	31860
taatcatgca	taagactgtg	atctgccttc	tgatcctctg	catcctgccc	accctcacct	31920
cctgccagta	caccacaaaa	tctccgcgca	aaagacatgc	ctcctgccgc	ttcacccaac	31980
tgtggaatat	acccaaatgc	tacaacgaaa	agagcgagct	ctccgaagct	tggtctgtatg	32040
gggtcatctg	tgtcttagtt	ttctgcagca	ctgtctttgc	cctcataatc	taccctact	32100
ttgatttggg	atggaacgcg	atcgatgcca	tgaattaccc	cacctttccc	gcacccgaga	32160
taattccact	gcgacaagtt	gtaccggttg	tcgttaatca	acgcccccca	tcccctacgc	32220
ccactgaaat	cagctacttt	aacctaacag	gcggagatga	ctgacgccct	agatctagaa	32280
atggacggca	tcagtaccga	gcagcgtctc	ctagagaggc	gcaggcaggc	ggctgagcaa	32340
gagcgctca	atcaggagct	ccgagatctc	gttaacctgc	accagtgcaa	aagaggcatc	32400
ttttgtctgg	taaagcaggc	caaagtcacc	tacgagaaga	ccggcaacag	ccaccgctc	32460
agttacaaat	tgcccaccca	gcgccagaag	ctggtgctca	tggtgggtga	gaatcccatc	32520
accgtcacc	agcactcgg	agagaccgag	gggtgtctgc	actccccctg	tcgggggtcca	32580
gaagacctct	gcacctgggt	aaagacctg	tgcggtctca	gagatttagt	cccctttaac	32640
taatcaaaca	ctggaatcaa	taaaaagaat	cacttactta	aaatcagaca	gcagggtctct	32700

gtccagttta	ttcagcagca	cctccttccc	ctcctcccaa	ctctgggtact	ccaaacgcct	32760
tctggcggca	aacttctctc	acaccctgaa	gggaatgtca	gattctttgct	cctgtccctc	32820
cgcaccact	atcttcatgt	tgttgagat	gaagcgcacc	aaaacgtctg	acgagagctt	32880
caaccccg	tacccctatg	acacggaaag	cggccctccc	tccgtccctt	tcctcacccc	32940
tcccttcgtg	tctcccgatg	gattccaaga	aagtcccccc	ggggtcctgt	ctctgaacct	33000
ggccgagccc	ctgggtcactt	cccacggcat	gctcgccctg	aaaatgggaa	gtggcctctc	33060
cctggacgac	gctggcaacc	tcacctctca	agatatcacc	accgctagcc	ctccccctca	33120
aaaaaccaag	accaacctca	gcctagaaac	ctcatcccc	ctaactgtga	gcacctcagg	33180
cgcctcacc	gtagcagccg	ccgctcccct	ggcgggtggc	ggcacctccc	tcaccatgca	33240
atcagaggcc	cccctgacag	tacaggatgc	aaaactcacc	ctggccacca	aaggccccct	33300
gaccgtgtct	gaaggcaaac	tggccttgca	aacatcggcc	ccgctgacgg	ccgctgacag	33360
cagcacctc	acagtcagtg	ccacaccacc	ccttagcaca	agcaatggca	gcttgggtat	33420
tgacatgcaa	gccccattt	acaccaccaa	tggaaaacta	ggacttaact	ttggcgctcc	33480
cctgcatgtg	gtagacagcc	taaatgcact	gactgtagtt	actggccaag	gtcttacgat	33540
aaacggaaca	gccctacaaa	ctagagtctc	aggtgccctc	aactatgaca	catcaggaaa	33600
cctagaattg	agagctgcag	ggggtatg	agttgatgca	aatgggtcaac	ttatccttga	33660
tgtagcttac	ccatttgatg	cacaaaacaa	tctcagcctt	aggcttggac	agggaccctt	33720
gtttgttaac	tctgcccaca	acttgatgt	taactacaac	agaggcctct	acctgttcac	33780
atctggaaat	acaaaaagc	tagaagttaa	tatcaaaaca	gccaagggtc	tcatttatga	33840
tgacactgct	atagcaatca	atgcgggtga	tgggctacag	tttgactcag	gctcagatac	33900
aatccatta	aaaactaaac	ttggattagg	actggattat	gactccagca	gagccataat	33960
tgctaaactg	ggaactggcc	taagctttga	caacacaggt	gccatcacag	taggcaacaa	34020
aatgatgac	aagcttacct	tgtggaccac	accagaccca	tcccctaact	gtagaatcta	34080
ttcagagaaa	gatgctaaat	tcacacttgt	tttgactaaa	tgcggcagtc	aggtgttggc	34140
cagcgtttct	gttttatctg	taaaaggtag	ccttgcgccc	atcagtggca	cagtaactag	34200
tgctcagatt	gtcctcagat	ttgatgaaaa	tggagttcta	ctaagcaatt	cttcccttga	34260
ccctcaatac	tggaaactaca	gaaaagggtga	ccttacagag	ggcactgcat	ataccaacgc	34320
agtgggattt	atgccaacc	tcacagcata	cccaaaaaca	cagagccaaa	ctgctaaaag	34380
caacattgta	agtcaggttt	acttgaatgg	ggacaaatcc	aaacccatga	ccctcaccat	34440
tacctcaat	ggaactaatg	aaacaggaga	tgccacagta	agcacttact	ccatgtcatt	34500
ctcatggaac	tggaatggaa	gtaattacat	taatgaaacg	ttccaaacca	actccttcac	34560
cttctcctac	atcgcccaag	aataaaaagc	atgacgctgt	tgatttgatt	caatgtgttt	34620

ctgttttatt	ttcaagcaca	acaaaatcat	tcaagtcatt	cttccatctt	agcttaatag	34680
acacagtagc	ttaatagacc	cagtagtgca	aagccccatt	ctagcttata	gatcagacag	34740
tgataattaa	ccaccaccac	caccatacct	tttgattcag	gaaatcatga	tcatcacagg	34800
atcctagtct	tcaggccgcc	ccctccctcc	caagacacag	aatacacagt	cctctcccc	34860
cgactggctt	taaataacac	catctgggtg	gtcacagaca	tgttcttagg	ggttatattc	34920
cacacggctc	cctgccgcgc	caggcgctcg	tcgggtgatgt	tgataaaactc	tcccggcagc	34980
tcgctcaagt	tcacgtcgct	gtccagcggc	tgaacctccg	gctgacgcga	taactgtgcg	35040
accggctgct	ggacgaacgg	aggccgcgcc	tacaaggggg	tagagtcata	atcctcggtc	35100
aggatagggc	ggtgatgcag	cagcagcgag	cgaaacatct	gctgccgcgc	ccgctccgtc	35160
cggcaggaaa	acaacacgcc	ggtggtctcc	tccgcgataa	tccgcaccgc	ccgcagcatc	35220
agcttcctcg	ttctccgcgc	gcagcacctc	acccttatct	cgctcaaatc	ggcgcagtag	35280
gtacagcaca	gcaccacgat	gttattcatg	atcccacagt	gcagggcgct	gtatccaaag	35340
ctcatgccgg	gaaccaccgc	ccccacgtgg	ccatcgtacc	acaagcgcac	gtaaatcaag	35400
tgtcgacccc	tcatgaacgc	gctggacaca	aacattactt	ccttgggcat	gttgtaattc	35460
accacctccc	ggtaccagat	aaacctctgg	ttgaacaggg	caccttccac	caccatcctg	35520
aaccaagagg	ccagaacctg	cccaccggct	atgcactgca	gggaaccgcg	gttggaaaca	35580
tgacaatgca	gactccaagg	ctcgtaaccg	tggtatcatcc	ggctgctgaa	ggcatcgatg	35640
ttggcacaac	acagacacac	gtgcatgcac	tttctcatga	ttagcagctc	ttccctcgtc	35700
aggatcatat	ccaaggaat	aaccattctt	tgaatcaacg	taaaaccac	acagcagggg	35760
aggcctcgca	cataactcac	gttgtgcatg	gtcagcgtgt	tgcattccgg	aaacagcgga	35820
tgatcctcca	gtatcgaggc	gcgggtctcc	ttctcacagg	gaggtaaagg	gtccctgctg	35880
tacggactgc	gccgggacga	ccgagatcgt	gttgagcgta	gtgtcatgga	aaagggaaacg	35940
ccggacgtgg	tcatacttct	tgaagcagaa	ccaggttcgc	gcgtggcagg	cctccttgcg	36000
tctgcggtct	cgccgtctag	ctcgctccgt	gtgatagtgt	tagtacagcc	actccgcgag	36060
agcgtcgagg	cgcaccctgg	cttccggatc	tatgtagact	ccgtcttgca	ccgcggccct	36120
gataatatcc	accaccgtag	aataagcaac	accagccaa	gcaatacact	cgctctgcga	36180
gcggcagaca	ggaggagcgg	gcagagatgg	gagaaccatg	ataaaaaact	ttttttaag	36240
aatatcttcc	aattcttcga	aagtaagatc	tatcaagtgg	cagcgtccc	ctccactggc	36300
gcggtcaaac	tctacggcca	aagcacagac	aacggcattt	ctaagatggt	ccttaatggc	36360
gtccaaaaga	cacaccgctc	tcaagttgca	gtaaactatg	aatgaaaacc	catccggctg	36420
atcttccaat	atagacgcgc	cggcagcgtc	caccaaacc	agataatctt	cttctctcca	36480
gcgggtttacg	atctgtctaa	gcaaaccct	tatatcaagt	ccgacctgac	caaaaatctg	36540

ctcaagagcg	ccctccacct	tcatgtacaa	gcagcgcac	atgattgcaa	aaattcaggt	36600
tcttcagaga	cctgtataag	attcaaaatg	ggaacattaa	caaaaattcc	tctgtcgcgc	36660
agatcccttc	gcagggcaag	ctgaacataa	tcagacaggt	ccgaacggac	cagtgaggcc	36720
aaatccccac	caggaaccag	atccagagac	cctatactga	ttatgacgcg	catactcggg	36780
gctatgctga	ccagcgtagc	gccgatgtag	gcgtgctgca	tgggcggcga	gataaaatgc	36840
aaagtgtctg	ttaaaaaatc	aggcaaagcc	tcgcgcacaaa	aagctaacac	atcataatca	36900
tgctcatgca	ggtagttgca	ggtaagctca	ggaaccacaaa	cggaataaca	cacgattttc	36960
ctctcaaa	tgacttcgcg	gatactgcgt	aaaacacaaa	attataaata	aaaaattaat	37020
taaataactt	aaacattgga	agcctgtctc	acaacaggaa	aaaccacttt	aatcaacata	37080
agacggggcca	cgggcatgcc	ggcatagccg	taaaaaaatt	ggccccgtg	attaacaagt	37140
accacagaca	gctccccggt	catgtcgggg	gtcatcatgt	gagactctgt	atacacgtct	37200
ggattgtgaa	catcagacaa	acaaagaaat	cgagccacgt	agccccggagg	tataatcacc	37260
cgcaggcgga	ggtacagcaa	aacgaccccc	ataggaggaa	tcacaaaatt	agtaggagaa	37320
aaaaatacat	aaacaccaga	aaaaccctgt	tgctgaggca	aaatagcgcc	ctcccgatcc	37380
aaaacaacat	aaagcgcttc	cacaggagca	gccataacaa	agacccgagt	cttaccagta	37440
aaagaaaaaa	gatctctcaa	cgcagcacca	gcaccaacac	ttcgcagtgt	aaaaggccaa	37500
gtgccgagag	agtatatata	ggaataaaaa	gtgacgtaaa	cgggcacaaagt	ccaaaaaacg	37560
cccagaaaaa	ccgcacgcga	acctacgccc	cgaaacgaaa	gccaaaaaac	actagacact	37620
cccttcgggc	gtcaacttcc	gctttccccc	gctacgtcac	ttcccccggt	caaacaaact	37680
acatatcccg	aacttccaag	tcgccacgcc	caaaacaccg	cctacacctc	cccggccgcc	37740
ggcccgcccc	cggacccgcc	tcccgcctcc	cgccgcccac	ctcattatca	tattggcttc	37800
aatccaaaat	aaggtatatt	attgatgatg				37830

<210> 14
 <211> 15422
 <212> DNA
 <213> Pan troglodytes

<400> 14		
tgcccgcttt	ccagtcggga	aacctgtcgt gccagctgca ttaatgaatc ggccaacgcg 60
cggggagagg	cggtttgcgt	attgggcgct agatctccca tcacatatac ctgccgttca 120
ctattattta	gtgaaatgag	atattatgat attttctgaa ttgtgattaa aaaggcaact 180
ttatgcccat	gcaacagaaa	ctataaaaaa tacagagaat gaaaagaaac agatagattt 240
tttagttctt	tagggccgta	gtctgcaaat ctttttatga ttttctatca aacaaaagag 300
gaaaatagac	cagttgcaat	ccaaacgaga gtctaataga atgaggtcga aaagtaaact 360

gcgcggggttt gttactgata aagcaggcaa gacctaaaat gtgtaaaggg caaagtgtat	420
actttggcgt cacccttac atatttttagg tcttttttta ttgtgcgtaa ctaacttgcc	480
atcttcaaac aggagggtg gaagaagcag accgctaaca cagtacataa aaaaggagac	540
atgaacgatg aacatcaaaa agtttgcaaa acaagcaaca gtattaacct ttactaccgc	600
actgctggca ggaggcgcaa ctcaagcggt tgcgaaagaa acgaaccaa agccatataa	660
ggaaacatac ggcatttccc atattacacg ccatgatatg ctgcaaatcc ctgaacagca	720
aaaaaatgaa aaatatcaag ttcctgaatt cgattcgtcc acaattaaaa atatctcttc	780
tgcaaaaggc ctggacgttt gggacagctg gccattacaa aacgctgacg gcactgtcgc	840
aaactatcac ggctaccaca tcgtctttgc attagccgga gatcctaaaa atgcggatga	900
cacatcgatt tacatgttct atcaaaaagt cggcgaaact tctattgaca gctggaaaaa	960
cgctggccgc gtctttaag acagcgacaa attcgatgca aatgattcta tcctaaaaga	1020
ccaaacacaa gaatggtcag gttcagccac atttacatct gacggaaaaa tccgtttatt	1080
ctacactgat ttctccggta aacattacgg caaacaacaa ctgacaactg cacaagttaa	1140
cgtatcagca tcagacagct ctttgaacat caacgggtgta gaggattata aatcaatctt	1200
tgacggtgac ggaaaaacgt atcaaaatgt acagcagttc atcgatgaag gcaactacag	1260
ctcaggcgac aaccatacgc tgagagatcc tctactacgta gaagataaag gccacaaata	1320
cttagtattt gaagcaaaca ctggaactga agatggctac caaggcgaag aatctttatt	1380
taacaaagca tactatggca aaagcacatc attcttccgt caagaaagtc aaaaacttct	1440
gcaaagcgat aaaaaacgca cggctgagtt agcaaacggc gctctcggta tgattgagct	1500
aaacgatgat tacacactga aaaaagtgat gaaaccgctg attgcatcta acacagtaac	1560
agatgaaatt gaacgcgcga acgtctttaa aatgaacggc aaatgggtacc tgttcactga	1620
ctcccgcgga tcaaaaatga cgattgacgg cattacgtct aacgatattt acatgcttgg	1680
ttatgtttct aattctttta ctggcccata caagccgctg aacaaaactg gccttggtgt	1740
aaaaatggat cttgatccta acgatgtaac ctttacttac tcacacttcg ctgtacctca	1800
agcgaaagga aacaatgtcg tgattacaag ctatatgaca aacagaggat tctacgcaga	1860
caaacaatca acgtttgcgc caagcttcct gctgaacatc aaaggcaaga aaacatctgt	1920
tgtcaaagac agcatccttg aacaaggaca attaacagtt aacaaataat agggataaca	1980
gggtaatgct agaagacccg agtcttacca gtaaaagaaa aaagatctct caacgcagca	2040
ccagcaccaa cacttcgcag tgtaaaaggc caagtgccga gagagtatat ataggaataa	2100
aaagtgacgt aaacgggcaa agtccaaaaa acgcccagaa aaaccgcacg cgaacctacg	2160
ccccgaaacg aaagccaaaa aacactagac actcccttcc ggcgtcaact tccgctttcc	2220
cacgctacgt cacttgcccc agtcaaacaa actacatatc ccgaacttcc aagtcgccac	2280

gccccaaaaca	ccgcctacac	ctccccgccc	gccggcccg	ccccaaaccc	gcctcccgc	2340
ccgcgccccg	ccccgcgccc	cccatctcat	tatcatattg	gcttcaatcc	aaaataaggt	2400
atattattga	tgatggttta	aacggatcct	ctagagtcga	cctgcaggca	tgcaagcttg	2460
agtattctat	agtgtcacct	aaatagcttg	gcgtaatcat	ggtcatagct	gtttcctgtg	2520
tgaaattgtt	atccgctcac	aattccacac	aacatacgag	ccggaagcat	aaagtgtaaa	2580
gcctgggggtg	cctaattgagt	gagctaactc	acattaattg	cgttgcgctc	actgcccgt	2640
ttccagtcgg	gaaacctgtc	gtgccagctg	cattaatgaa	tcggccaacg	cgaacccctt	2700
gcggccgccc	gggcccgtcg	ccaattctca	tgtttgacag	cttatcatcg	aattttctgcc	2760
attcatccgc	ttattatcac	ttattcaggc	gtagcaacca	ggcgtttaag	ggcaccaata	2820
actgccttaa	aaaaattacg	ccccgccttg	ccactcatcg	cagtactggt	gtaattcatt	2880
aagcattctg	ccgacatgga	agccatcaca	aacggcatga	tgaacctgaa	tcgccagcgg	2940
catcagcacc	ttgtcgcctt	gcgtataata	tttgcccatg	gtgaaaacgg	gggcgaagaa	3000
gttgtccata	ttggccacgt	ttaaatcaaa	actggtgaaa	ctcaccaggg	gattggctga	3060
gacgaaaaac	atatttctcaa	taaacccttt	agggaaatag	gccagggtttt	caccgtaaca	3120
cgccacatct	tgcgaaatata	tgtgtagaaa	ctgccggaaa	tcgtcgtggg	attcaactcca	3180
gagcgatgaa	aacgtttcag	tttgctcatg	gaaaacgggtg	taacaagggt	gaacactatc	3240
ccatatcacc	agctcacctg	ctttcattgc	catacggaat	tccggatgag	cattcatcag	3300
gcgggcaaga	atgtgaataa	aggccggata	aaacttgtgc	ttatTTTTTct	ttacgggtctt	3360
taaaaaggcc	gtaatatcca	gctgaacggg	ctggttatag	gtacattgag	caactgactg	3420
aatgcctca	aatgtttctt	tacgatgcca	ttgggatata	tcaacgggtg	tatatccagt	3480
gatttttttc	tccatttttag	cttccttagc	tcctgaaaat	ctcgataact	caaaaaatac	3540
gcccggtagt	gatcttattt	cattatgggtg	aaagttggaa	cctcttacgt	gccgatcaac	3600
gtctcatttt	cgccaaaagt	tggcccagg	cttcccggta	tcaacaggga	caccaggatt	3660
tattttattct	gcgaagtgat	cttccgtcac	aggtatttat	tcgcgataag	ctcatggagc	3720
ggcgtaaccg	tcgcacagga	aggacagaga	aagcgcggat	ctgggaagtg	acggacagaa	3780
cggtcaggac	ctggattggg	gaggcggttg	ccgccgctgc	tgctgacggg	gtgacgttct	3840
ctgttccggg	cacaccacat	acgttccgcc	attcctatgc	gatgcacatg	ctgtatgccg	3900
gtataccgct	gaaagtcttg	caaagcctga	tgggacataa	gtccatcagt	tcaacggaag	3960
tctacacgaa	ggtttttg	ctggatgtgg	ctgcccggca	ccgggtgcag	tttgcgatgc	4020
cggagtctga	tgcggttg	atgctgaaac	aattatcctg	agaataaatg	ccttggcctt	4080
tatatggaaa	tgtggaactg	agtggatatg	ctgtttttgt	ctgttaaaca	gagaagctgg	4140
ctgttatcca	ctgagaagcg	aacgaaacag	tcgggaaaat	ctccattat	cgtagagatc	4200

cgcattatta atctcaggag cctgtgtagc gtttatagga agtagtggtc tgtcatgatg	4260
cctgcaagcg gtaacgaaaa cgatttgaat atgccttcag gaacaataga aatcttcgtg	4320
cgggtgttacg ttgaagtgga gcggattatg tcagcaatgg acagaacaac ctaatgaaca	4380
cagaaccatg atgtggtctg tcctttttaca gccagtagtg ctgcgccgag tcgagcgaca	4440
gggcgaagcc ctcgagttag cgaggaagca ccagggaaca gcacttatat attctgctta	4500
cacacgatgc ctgaaaaaac ttcccttggg gttatccact tatccacggg gatattttta	4560
taattatttt ttttatagtt ttttagatctt cttttttaga gcgccttgta ggcctttatc	4620
catgctgggt ctagagaagg tgttgtagaca aattgccctt tcagtgtgac aaatcacctt	4680
caaatgacag tcctgtctgt gacaaattgc ccttaaccct gtgacaaatt gccctcagaa	4740
gaagctgttt tttcacaag ttatccctgc ttattgactc ttttttattt agtgtgacaa	4800
tctaaaaact tgtcacactt cacatggatc tgtcatggcg gaaacagcgg ttatcaatca	4860
caagaaacgt aaaaatagcc cgcgaatcgt ccagtcaaac gacctcactg aggcggcata	4920
tagtctctcc cgggatcaaa aacgtatgct gtatctgttc gttgaccaga tcagaaaatc	4980
tgatggcacc ctacaggaac atgacgggat ctgcgagatc catgttgcta aatatgctga	5040
aatattcggg ttgacctctg cggaagccag taaggatata cggcaggcat tgaagagttt	5100
cgcggggaag gaagtgggtt tttatcgccc tgaagaggat gccggcgatg aaaaaggcta	5160
tgaatctttt ccttggttta tcaaacgtgc gcacagtcca tccagagggc tttacagtgt	5220
acatatcaac ccatatctca ttcccttctt tatcggttta cagaaccggt ttacgcagtt	5280
tcggcttagt gaaacaaaag aaatcaccaa tccgtatgcc atgcgtttat acgaatccct	5340
gtgtcagtat cgtaagccgg atggctcagg catcgtctct ctgaaaatcg actggatcat	5400
agagcgttac cagctgcctc aaagttacca gcgtatgcct gacttccgcc gccgcttctt	5460
gcaggctctgt gttaatgaga tcaacagcag aactccaatg cgcctctcat acattgagaa	5520
aaagaaaggc cgccagacga ctcatatcgt attttccctc cgcgatatca cttccatgac	5580
gacaggatag tctgaggggt atctgtcaca gatttgaggg tggttcgtca catttgttct	5640
gacctactga gggtaatttg tcacagtttt gctgtttcct tcagcctgca tggattttct	5700
catacttttt gaactgtaat ttttaaggaa gccaaatttg agggcagttt gtcacagttg	5760
atttccctct ctttcccttc gtcattgtgac ctgatatcgg ggggttagttc gtcattcattg	5820
atgaggggtt attatcacag tttattactc tgaattggct atccgcgtgt gtacctctac	5880
ctggagtttt tcccacggtg gatatttctt cttgcgctga gcgtaagagc tatctgacag	5940
aacagttctt ctttgcttcc tcgccagttc gctcgctatg ctcggttaca cggctgcggc	6000
gagcgctagt gataataagt gactgaggta tgtgctcttc ttatctcctt ttgtagtggt	6060
gctcttattt taaacaactt tgcgggtttt tgatgacttt gcgattttgt tgttgctttg	6120

cagtaaattg	caagatttaa	taaaaaaacg	caaagcaatg	attaaaggat	gttcagaatg	6180
aaactcatgg	aaacacttaa	ccagtgcata	aacgctggtc	atgaaatgac	gaaggctatc	6240
gccattgcac	agtttaatga	tgacagcccg	gaagcgagga	aaataacccg	gcgctggaga	6300
ataggtgaag	cagcggattt	agttgggggt	tcttctcagg	ctatcagaga	tgccgagaaa	6360
gcagggcgac	taccgcaccc	ggatatggaa	attcgaggac	gggttgagca	acgtgttggt	6420
tatacaattg	aacaaattaa	tcatatgcgt	gatgtgtttg	gtacgcgatt	gcgacgtgct	6480
gaagacgtat	ttccaccggg	gatcgggggt	gctgcccata	aagggtggcgt	ttacaaaacc	6540
tcagtttctg	ttcatcttgc	tcaggatctg	gctctgaagg	ggctacgtgt	tttgctcgtg	6600
gaaggtaacg	acccccaggg	aacagcctca	atgtatcacg	gatgggtacc	agatcttcat	6660
attcatgcag	aagacactct	cctgcctttc	tatcttgggg	aaaaggacga	tgtcacttat	6720
gcaataaagc	ccacttgctg	gccgggggct	gacattattc	cttcctgtct	ggctctgcac	6780
cgtattgaaa	ctgagttaat	gggcaaattt	gatgaaggta	aactgcccac	cgatccacac	6840
ctgatgctcc	gactggccat	tgaaactgtt	gctcatgact	atgatgtcat	agttattgac	6900
agcgcgccta	acctgggtat	cggcacgatt	aatgtcgtat	gtgctgctga	tgtgctgatt	6960
gttcccacgc	ctgctgagtt	gtttgactac	acctccgcac	tgcagttttt	cgatatgctt	7020
cgtgatctgc	tcaagaacgt	tgatcttaaa	gggttcgagc	ctgatgtacg	tattttgctt	7080
accaaataca	gcaatagtaa	tggtctctcag	tccccgtgga	tggaggagca	aattcgggat	7140
gcctggggaa	gcatggttct	aaaaaatgtt	gtacgtgaaa	cggatgaagt	tggtaaagggt	7200
cagatccgga	tgagaactgt	ttttgaacag	gccattgata	aacgctcttc	aactgggtgcc	7260
tggagaaatg	ctctttctat	ttgggaacct	gtctgcaatg	aaattttcga	tcgtctgatt	7320
aaaccacgct	gggagattag	ataatgaagc	gtgcgcctgt	tattccaaaa	catacgtcca	7380
atactcaacc	ggttgaagat	acttcgttat	cgacaccagc	tgccccgatg	gtggattcgt	7440
taattgcgcg	cgtaggagta	atggctcgcg	gtaatgcat	tactttgcct	gtatgtggtc	7500
gggatgtgaa	gtttactctt	gaagtgctcc	ggggtgatag	tgttgagaag	acctctcggg	7560
tatggtcagg	taatgaacgt	gaccaggagc	tgcttactga	ggacgcactg	gatgatctca	7620
tcccttcttt	tctactgact	ggtcaacaga	caccggcggt	cggtcgaaga	gtatctgggtg	7680
tcatagaaat	tgccgatggg	agtcgccgtc	gtaaagctgc	tgcacttacc	gaaagtgatt	7740
atcgtgttct	ggttggcgag	ctggatgatg	agcagatggc	tgcattatcc	agattgggta	7800
acgattatcg	ccaacaagt	gcttatgaac	gtggtcagcg	ttatgcaagc	cgattgcaga	7860
atgaatttgc	tggaaatatt	tctgcgctgg	ctgatgcgga	aaatatttca	cgtaagatta	7920
ttacccgctg	tatcaacacc	gccaaattgc	ctaaatcagt	tgttgctctt	ttttctcacc	7980
ccggtgaact	atctgcccgg	tcagggtgatg	cacttcaaaa	agcctttaca	gataaagagg	8040

aattacttaa	gcagcaggca	tctaaccttc	atgagcagaa	aaaagctggg	gtgatatttg	8100
aagctgaaga	agttatcact	cttttaactt	ctgtgcttaa	aacgtcatct	gcatcaagaa	8160
ctagtttaag	ctcacgacat	cagtttgctc	ctggagcgac	agtattgtat	aagggcgata	8220
aaatggtgct	taacctggac	aggtctcgtg	ttccaactga	gtgtatagag	aaaattgagg	8280
ccattcttaa	ggaacttgaa	aagccagcac	cctgatgcga	ccacgtttta	gtctacgttt	8340
atctgtcttt	acttaatgtc	ctttgttaca	ggccagaaag	cataactggc	ctgaatatct	8400
tctctggggc	cactgttcca	cttgatcgt	cggctctgata	atcagactgg	gaccacggct	8460
ccactcgat	cgctcggctg	attattagtc	tgggaccacg	gtccactcgt	tatcgctcgt	8520
ctgattatta	gtctgggacc	acgggtccac	tcgtatcgtc	ggctctgata	tcagactggg	8580
accacggctc	cactcgtatc	gtcggctctga	ttattagtc	gggaccatgg	tccactcgt	8640
atcgctcggc	tgattattag	tctgggacca	cgggtccact	cgtatcgtcg	gtctgattat	8700
tagtctggaa	ccacgggtcc	actcgtatcg	tcggctctgat	tattagtcgt	ggaccacgg	8760
cccactcgta	tcgtcggctc	gattattagt	ctgggaccac	gatccactc	gtgttgctcg	8820
tctgattatc	ggctctgggac	cacgggtcca	cttgatttgt	cgatcagact	atcagcgtga	8880
gactacgatt	ccatcaatgc	ctgtcaaggg	caagtattga	catgtcgtcg	taacctgtag	8940
aacggagtaa	cctcgggtgtg	cggttgtatg	cctgctgtgg	attgctgctg	tgtcctgctt	9000
atccacaaca	ttttgcgcac	ggttatgtgg	acaaaatacc	tggttacca	ggcgtgccc	9060
gcacgttaac	cgggctgcat	ccgatgcaag	tgtgtcgtcg	tcgacgagct	cgcgagctcg	9120
gacatgaggt	tgccccgtat	tcagtgtcgc	tgatttgtat	tgtctgaagt	tgtttttacg	9180
ttaagttgat	gcagatcaat	taatacgata	cctgcgtcat	aattgattat	ttgacgtggt	9240
ttgatggcct	ccacgcacgt	tgtgatatgt	agatgataat	cattatcact	ttacgggtcc	9300
tttccgggtga	tccgacaggt	tacggggcgg	cgacctcgcg	ggttttcgct	atttatgaaa	9360
attttccgg	ttaaggcgtt	tccgttcttc	ttcgtcataa	cttaatgttt	ttatttaaaa	9420
taccctctga	aaagaaagga	aacgacaggt	gctgaaagcg	agcttttttg	cctctgtcgt	9480
ttcctttctc	tgtttttgct	cgtggaatga	acaatggaag	tccgagctca	tcgctaataa	9540
cttcgtatag	catacattat	acgaagtatt	attcgatgcg	gccgcaaggg	gttcgcgtca	9600
gcgggtgttg	gcgggtgtcg	gggctggctt	aactatgcgg	catcagagca	gattgtactg	9660
agagtgcacc	atatgcgggtg	tgaaataaccg	cacagatgcg	taaggagaaa	ataccgcac	9720
aggcgccatt	cgccattcag	gctgcgcaac	tgttgggaag	ggcgatcgg	gcgggcctct	9780
tcgctattac	gccagctggc	gaaaggggga	tgtgctgcaa	ggcgattaag	ttgggtaacg	9840
ccagggtttt	cccagtcacg	acgttgtaaa	acgacggcca	gtgaattgta	atacgactca	9900
ctatagggcg	aattcgagct	cgggtaccgg	ggatcctcgt	ttaaaccatc	atcaataata	9960

taccttattt	tggattgaag	ccaatatgat	aatgagatgg	gcggcgcggg	gcggggcgcg	10020
gggcgggagg	cgggtttggg	ggcggggccg	cgggcggggc	ggtgtggcgg	aagtggactt	10080
tgtaagtgtg	gcggatgtga	cttgctagt	cggggcgcg	taaaagtgac	gttttccgtg	10140
cgcgacaacg	cccccgga	gtgacatttt	tcccgcggtt	tttaccggat	gttgtagtga	10200
atttgggcgt	aaccaagtaa	gatttggcca	ttttcgcg	aaaactgaaa	cggggaagt	10260
aaatctgatt	aattttgcgt	tagtcatacc	gcgtaatat	tgtctagggc	cgagggactt	10320
tggccgatta	cgtggaggac	tgcgccaggt	gttttttgag	gtgaatttcc	gcgttccggg	10380
tcaaagtctg	cgttttatta	ttataggata	tccattgca	tacgttgtat	ccatatcata	10440
atatgtacat	ttatattggc	tcatgtccaa	cattaccgcc	atgttgacat	tgattattga	10500
ctagttatta	atagtaatca	attacgggg	cattagttca	tagcccatat	atggagttcc	10560
gcgttacata	acttacggta	aatggcccg	ctggctgacc	gccaacgac	ccccgcccat	10620
tgacgtcaat	aatgacgtat	gttcccatag	taacgcgaat	agggactttc	cattgacgtc	10680
aatgggtgga	gtatttacgg	taaactgccc	acttggcagt	acatcaagt	tatcatatgc	10740
caagtacgcc	ccctattgac	gtcaatgacg	gtaaatggcc	cgctggcat	tatgccag	10800
acatgacctt	atgggacttt	cctacttggc	agtacatcta	cgtattagtc	atcgctatta	10860
ccatggtgat	gcggttttgg	cagtacatca	atgggcgtgg	atagcggttt	gactcacggg	10920
gatttccaag	tctccacccc	attgacgtca	atgggagttt	gttttggcac	caaaatcaac	10980
gggactttcc	aaaatgtcgt	aacaactccg	ccccattgac	gcaaatgggc	ggtaggcgtg	11040
tacggtggga	ggtctatata	agcagagctc	tccctatcag	tgatagagat	ctccctatca	11100
gtgatagaga	tcgtcgacga	gctcgtttag	tgaaccgtca	gatcgctgg	agacgccatc	11160
cacgctgttt	tgacctccat	agaagacacc	gggaccgatc	cagcctccgc	ggccgggaac	11220
ggtgcattgg	aacgcggatt	ccccgtgcca	agagtgagat	cttccgttta	tctaggtacc	11280
gggccccccc	tcgaggtcga	cggtatcgat	aagcttcacg	ctgcgcgaag	cactcagggc	11340
gcaagggctg	ctaaaggaag	cggaacacgt	agaaagccag	tccgcagaaa	cgggtgctgac	11400
cccggatgaa	tgtcagctac	tgggctatct	ggacaaggg	aaacgcaagc	gcaaagagaa	11460
agcaggtagc	ttgcagtggg	cttacatggc	gatagctaga	ctgggcggtt	ttatggacag	11520
caagcgaacc	ggaattgcca	gctggggcgc	cctctggtaa	ggttgggaag	ccctgcaaag	11580
taaactggat	ggctttcttg	ccgccaagga	tctgatggcg	caggggatca	agatctaacc	11640
aggagctatt	taatggcaac	agttaaccag	ctggtacgca	aaccacgtgc	tcgcaaagtt	11700
gcgaaaagca	acgtgcctgc	gctggaagca	tgcccgcaaa	aacgtggcgt	atgtactcgt	11760
gtatatacta	ccactcctaa	aaaaccgaac	tccgcgtgc	gtaaagtatg	ccgtgttcgt	11820
ctgactaacg	gtttcgaagt	gacttcctac	atcggtggtg	aaggtcacaa	cctgcaggag	11880

cactccgtga	tcctgatccg	tggcggtcgt	gttaaagacc	tcccgggtgt	tcgttaccac	11940
accgtacgtg	gtgcgcttga	ctgctccggc	gttaaagacc	gtaagcaggc	tcgttccaag	12000
tatggcgtga	agcgtcctaa	ggcttaatgg	tagatctgat	caagagacag	gatgacggtc	12060
gtttcgcgatg	cttgaacaag	atggattgca	cgcaggttct	cgggccgctt	gggtggagag	12120
gctattcggc	tatgactggg	cacaacagac	aatcggctgc	tctgatgccg	ccgtgttccg	12180
gctgtcagcg	cagggggcgcc	cggttctttt	tgtcaagacc	gacctgtccg	gtgccctgaa	12240
tgaactgcag	gacgaggcag	cgcggctatc	gtggctggcc	acgacggggc	ttccttgccg	12300
agctgtgctc	gacgttgtca	ctgaagcggg	aagggactgg	ctgctattgg	gcgaagtgcc	12360
ggggcaggat	ctcctgtcat	ctcaccttgc	tcctgccgag	aaagtatcca	tcattggctga	12420
tgcaatgcgg	cggctgcata	cgcttgatcc	ggctacctgc	ccattcgacc	accaagcgaa	12480
acatcgcatc	gagcgagcac	gtactcggat	ggaagccggt	cttgtcgatc	aggatgatct	12540
ggacgaagag	catcaggggc	tcgcgccagc	cgaactgttc	gccaggctca	aggcgcgcat	12600
gcccgcggc	gaggatctcg	tcgtgacca	tggcgatgcc	tgcttgccga	atatcatggt	12660
ggaaaatggc	cgcttttctg	gattcatcga	ctgtggccgg	ctgggtgtgg	cggaccgcta	12720
tcaggacata	gcgttggcta	cccgtgatat	tgctgaagag	cttggcgggc	aatgggctga	12780
ccgcttcctc	gtgctttacg	gtatcgccgc	tcccgattcg	cagcgcacgc	ccttctatcg	12840
ccttcttgac	gagttcttct	gagcgggact	ctggggttcg	aaatgaccga	ccaagcgacg	12900
cccaacctgc	catcacgaga	tttcgattcc	accgccgcct	tctatgaaag	gttgggcttc	12960
ggaatcgttt	tccgggacgc	cggctggatg	atcctccagc	gcggggatct	catgctggag	13020
ttcttcgccc	accccgggct	cgatcccctc	ggggggaatc	agaattcagt	cgacagcggc	13080
cgcgatctgc	tgtgccttct	agttgccagc	catctgttgt	ttgccccctc	cccgtgcctt	13140
ccttgaccct	ggaagggtgc	actcccactg	tcctttccta	ataaaatgag	gaaattgcat	13200
cgcatgtct	gagtaggtgt	cattctattc	tggggggtgg	ggtggggcag	gacagcaagg	13260
gggaggattg	ggaagacaat	agcaggcatg	ctggggatgc	ggtgggctct	atggccgacg	13320
agcgatcgct	gaggtgggtg	agtgggcgtg	gcctgggggtg	gtcatgaaaa	tatataagtt	13380
gggggtctta	gggtctcttt	atgtgtgttg	cagagaccgc	cggagccatg	agcgggagca	13440
gcagcagcag	cagtagcagc	agcgccttgg	atggcagcat	cgtgagccct	tatttgacga	13500
cgcggatgcc	ccactggggc	gggggtgcgtc	agaatgtgat	gggctccagc	atcgacggcc	13560
gaccgcctct	gcccgcaaat	tccgccacgc	tgacctatgc	gaccgtcgcg	gggacgccgt	13620
tggacgccac	cgccgccgcc	gccgccaccg	cagccgcctc	ggccgtgcgc	agcctggcca	13680
cggactttgc	attcctggga	ccactggcga	caggggctac	ttctcggggc	gctgctgccg	13740
ccgttcgcga	tgacaagctg	accgccctgc	tggcgagttt	ggatgcgctt	actcgggaac	13800

tgggtgacct	ttctcagcag	gtcatggccc	tgcgccagca	ggctctctcc	ctgcaagctg	13860
gcgggaatgc	ttctcccaca	aatgccgttt	aagggcgcg	ctagggataa	cagggtaata	13920
cccctatttg	tttatttttc	taaatacatt	caaatatgta	tccgctcatg	agacaataac	13980
cctgataaat	gcttcaataa	tattgaaaaa	ggaagagtat	gagtattcaa	catttccgtg	14040
tcgcccttat	tccctttttt	gcggcatttt	gccttcctgt	ttttgctcac	ccagaaacgc	14100
tgggtgaaagt	aaaagatgct	gaagatcagt	tgggtgcacg	agtgggttac	atcgaactgg	14160
atctcaacag	cggtaaagtc	cttgagagtt	ttcgccccga	agaacgtttt	ccaatgatga	14220
gcacttttaa	agttctgcta	tgtggcgcg	tattatcccg	tattgacgcc	gggcaagagc	14280
aactcggtcg	ccgcatacac	tattctcaga	atgacttggg	tgagtactca	ccagtcacag	14340
aaaagcatct	tacggatggc	atgacagtaa	gagaattatg	cagtgtgcc	ataaccatga	14400
gtgataacac	tgcggccaac	ttacttctga	caacgatcgg	aggaccgaag	gagctaaccg	14460
cttttttgca	caacatgggg	gatcatgtaa	ctcgccttga	tcgttgggaa	ccggagctga	14520
atgaagccat	accaaacgac	gagcgtgaca	ccacgatgcc	tgtagcaatg	gcaacaacgt	14580
tgcgcaaact	attaactggc	gaactactta	ctctagcttc	ccggcaacaa	ttaatagact	14640
ggatggaggc	ggataaagtt	gcaggaccac	ttctgcgctc	ggcccttcg	gctggctggg	14700
ttattgctga	taaatctgga	gccggtgagc	gtgggtctcg	cggtatcatt	gcagcactgg	14760
ggccagatgg	taagccctcc	cgtatcgtag	ttatctacac	gacggggagt	caggcaacta	14820
tggatgaacg	aaatagacag	atcgctgaga	taggtgcctc	actgattaag	cattggtaac	14880
tgtcagacca	agtttactca	tatatacttt	agattgattt	aaaatacgta	tatatgtatt	14940
agtcatcgct	attaccatgg	ttaatgcgcc	gctacagggc	gcgtccattc	gccattcagg	15000
ctgcgcaact	gttgggaagg	gcgatcgggtg	cgggcctctt	cgctattacg	ccagctggcg	15060
aaagggggat	gtgctgcaag	gcgattaagt	tgggtaacgc	cagggttttc	ccagtcacga	15120
cgttgtaaaa	cgacggccag	tgaattgtaa	tacgactcac	tatagggcga	attgggccct	15180
ctagatgcat	gctcgagcgg	ccgccagtgt	gatggatatc	tgcagaattc	cagcacactg	15240
gcggccgtta	ctagtggatc	cgagctcggt	accaagcttg	gcgtaatcat	ggtcatagct	15300
gtttcctgtg	tgaaattggt	atccgctcac	aattccacac	aacatacgag	ccggaagcat	15360
aaagtgtaaa	gcctgggggtg	cctaattgagt	gagctaactc	acattaattg	cgttgcgctc	15420
ac						15422

<210> 15
 <211> 39535
 <212> DNA
 <213> Pan troglodytes

<220>

<221> misc_feature
<222> (24391)..(24393)
<223> n iðääñòàâëÿåò ñîáîé a, c, g, èèè t

<400> 15
catcatcaat aatatacctt attttggatt gaagccaata tgataatgag atgggcggcg 60
cggggcgggg cgcggggcgg gaggcgggtt tgggggcggg ccggcgggcg gggcggtgtg 120
gcggaagtgg actttgtaag tgtggcggat gtgacttgct agtgccgggc gcggtaaaag 180
tgacgttttc cgtgcgcgac aacgcccccg ggaagtgaca tttttccgc ggtttttacc 240
ggatgttgta gtgaatttgg gcgtaaccaa gtaagatttg gccatttttcg cgggaaaact 300
gaaacgggga agtgaaatct gattaatttt gcgttagtca taccgcgtaa tatttgtcta 360
gggccgaggg actttggccg attacgtgga ggactcgccc aggtgttttt tgaggtgaat 420
ttccgcgttc cgggtcaaag tctgcgtttt attattatag gatatcccat tgcatacgtt 480
gtatccatat cataatatgt acatttatat tggctcatgt ccaacattac cgccatgttg 540
acattgatta ttgactagtt attaatagta atcaattacg gggtcattag ttcatagccc 600
atatatggag ttccgcgtta cataacttac ggtaaattggc ccgcctggct gaccgcccaa 660
cgacccccgc ccattgacgt caataatgac gtatgttccc atagtaacgc caatagggac 720
tttcattga cgtcaatggg tggagtattt acggtaaact gccacttgg cagtacatca 780
agtgtatcat atgccaagta cgccttat tgacgtcaat gacggtaaatt ggccgcctg 840
gcattatgcc cagtacatga ccttatggga ctttcctact tggcagtaca tctacgtatt 900
agtcacgct attaccatgg tgatgcggtt ttggcagtac atcaatgggc gtggatagcg 960
gtttgactca cggggatttc caagtctcca cccattgac gtcaatggga gtttgttttg 1020
gcacaaaaat caacgggact ttccaaaatg tcgtaacaac tccgccccat tgacgcaaatt 1080
gggcggtagg cgtgtacggt gggaggtcta tataagcaga gctctcccta tcagtgatag 1140
agatctccct atcagtgata gagatcgtcg acgagctcgt ttagtgaacc gtcagatcgc 1200
ctggagacgc catccacgt gttttgacct ccatagaaga caccgggacc gatccagcct 1260
ccgcggcccg gaacggtgca ttggaacgcg gattccccgt gccaaagagt agatcttccg 1320
tttatctagg taccgggccc cccctcgagg tcgacggtat cgataagctt cacgctgccg 1380
caagcactca gggcgcaagg gctgctaaag gaagcggaac acgtagaaag ccagtccgca 1440
gaaacggtgc tgaccccgga tgaatgtcag ctactgggct atctggacaa gggaaaacgc 1500
aagcgcaaag agaaagcagg tagcttgcag tgggcttaca tggcgatagc tagactgggc 1560
ggttttatgg acagcaagcg aaccggaatt gccagctggg gcgccctctg gtaaggttgg 1620
gaagccctgc aaagtaaact ggatggcttt cttgccgcca aggatctgat ggcgcagggg 1680
atcaagatct aaccaggagc tatttaatgg caacagttaa ccagctggta cgcaaaccac 1740
gtgctcgcaa agttgcgaaa agcaacgtgc ctgcgctgga agcatgcccg caaaaacgtg 1800

gcgtatgtac	tcgtgtatat	actaccactc	ctaaaaaacc	gaactccgcg	ctgcgtaaag	1860
tatgccgtgt	tcgtctgact	aacggtttcg	aagtgacttc	ctacatcggg	ggtgaaggtc	1920
acaacctgca	ggagcactcc	gtgatcctga	tccgtggcgg	tcgtgttaaa	gacctcccgg	1980
gtgttcgtta	ccacaccgta	cgtgggtgcg	ttgactgctc	cggcgttaaa	gaccgtaagc	2040
aggctcgttc	caagtatggc	gtgaagcgtc	ctaaggctta	atggtagatc	tgatcaagag	2100
acaggatgac	ggtcgttttcg	catgcttgaa	caagatggat	tgcacgcagg	ttctccggcc	2160
gcttgggtgg	agaggctatt	cggctatgac	tgggcacaa	agacaatcgg	ctgctctgat	2220
gccgccgtgt	tccggctgtc	agcgcagggg	cgcccggttc	tttttgtcaa	gaccgacctg	2280
tccggtgccc	tgaatgaact	gcaggacgag	gcagcgcggc	tatcgtggct	ggccacgacg	2340
ggcgttcctt	gcgcagctgt	gctcgacgtt	gtcactgaag	cgggaaggga	ctggctgcta	2400
ttgggcgaag	tgccggggca	ggatctcctg	tcattctacc	ttgctcctgc	cgagaaagta	2460
tccatcatgg	ctgatgcaat	gcggcggctg	catacgcttg	atccggctac	ctgcccattc	2520
gaccaccaag	cgaaacatcg	catcgagcga	gcacgtactc	ggatggaagc	cggtcttgtc	2580
gatcaggatg	atctggacga	agagcatcag	gggctcgcgc	cagccgaact	gttcgccagg	2640
ctcaaggcgc	gcatgcccga	cggcgaggat	ctcgtcgtga	cccatggcga	tgcttgcctg	2700
ccgaatatca	tggtggaaaa	tgcccgcttt	tctggattca	tcgactgtgg	cgggctgggt	2760
gtggcggacc	gctatcagga	catagcgttg	gctaccctg	atattgctga	agagcttggc	2820
ggcgaatggg	ctgaccgctt	cctcgtgctt	tacggtatcg	ccgctcccga	ttcgcagcgc	2880
atcgctttct	atcgctttct	tgacgagttc	ttctgagcgg	gactctgggg	ttcgaaatga	2940
ccgaccaagc	gacgcccac	ctgccatcac	gagatttcga	ttccaccgcc	gccttctatg	3000
aaagggttggg	cttcggaatc	gttttcgggg	acgccggctg	gatgatcctc	cagcgcgggg	3060
atctcatgct	ggagtcttct	gcccaccccg	ggctcgatcc	cctcgggggg	aatcagaatt	3120
cagtcgacag	cggccgcgat	ctgctgtgcc	ttctagttgc	cagccatctg	ttgtttgccc	3180
ctccccctg	ccttccttga	ccctggaagg	tgccactccc	actgtccttt	cctaataaaa	3240
tgaggaaatt	gcatcgcatt	gtctgagtag	gtgtcattct	attctggggg	gtgggggtggg	3300
gcaggacagc	aagggggagg	attgggaaga	caatagcagg	catgctgggg	atgcggtggg	3360
ctctatggcc	gatcagcgat	cgctgaggtg	ggtgagtggt	cgtggcctgg	ggtgggtcatg	3420
aaaatatata	agttgggggt	cttagggctc	ctttatttgt	gttgcagaga	ccgccggagc	3480
catgagcggg	agcagcagca	gcagcagtag	cagcagcgcc	ttggatggca	gcatcgtgag	3540
cccttatttg	acgacgcgga	tgccccactg	ggccgggggtg	cgtcagaatg	tgatgggctc	3600
cagcatcgac	ggccgaccgg	tcctgcccgc	aaattccgcc	acgctgacct	atgcgaccgt	3660
cgccggggacg	ccgttggacg	ccaccgccgc	cgccgccgcc	accgcagccg	cctcggccgt	3720

gcgagcctg gccacggact ttgcattcct gggaccactg ggcacagggg ctactttctcg	3780
ggccgctgct gccgccgttc gcgatgacaa gctgaccgcc ctgctggcgc agttggatgc	3840
gcttactcgg gaactgggtg acctttctca gcaggtcatg gccctgcgcc agcaggctctc	3900
ctccctgcaa gctggcgagg atgcttctcc cacaaatgcc gtttaagata aataaaacca	3960
gactctgttt ggattaaaga aaagtagcaa gtgcattgct ctctttattt cataattttc	4020
cgcgcgcgat aggccctaga ccagcgttct cggtcgttga gggcgcggtg tatcttctcc	4080
aggacgtggt agaggtggct ctggacgttg agatacatgg gcatgagccc gtccccgggg	4140
tggaggtagc accactgcag agcttcatgc tccgggggtg tgttgtagat gatccagtcg	4200
tagcaggagc gctgggcatg gtgcctaaaa atgtccttca gcagcaggcc gatggccagg	4260
gggaggccct tgggtgaagt gtttacaaaa cggttaagtt ggggaagggtg cattcgggga	4320
gagatgatgt gcatcttgga ctgtatTTTT agattggcga tgtttccgcc cagatccctt	4380
ctgggattca tgttgtagcag gaccaccagt acagtgtatc cgggtgcactt ggggaatttg	4440
tcatgcagct tagagggaaa agcgtggaag aacttggaaga cggccttgtag gcctcccaga	4500
ttttccatgc attcgtccat gatgatggca atggggccgc gggaggcagc ttggggcaaag	4560
atatttctgg ggtcgctgac gtcgtagttg tgttccaggg tgaggtcgtc ataggccatt	4620
tttacaaagc gcgggcggag ggtgcccga c tgggggatga tgggtccctc tggccccggg	4680
gcgtagttgc cctcgcatg ctgcatttcc caggccttaa tctcggaggg gggaaatcata	4740
tccacctgcg gggcgatgaa gaaaacggtt tccggagccg gggagattaa ctgggatgag	4800
agcaggtttc taagcagctg tgattttcca caaccggtgg gcccataaat aacacctata	4860
accggttgca gctggtagtt tagagagctg cagctgccgt cgtcccggag gaggggggccc	4920
acctcgttga gcatgtccct gacgcgcgatg ttctccccga ccagatccgc cagaaggcgc	4980
tcgcccacca gggacagcag ctcttgcaag gaagcaaagt ttttcagcgg cttgaggccg	5040
tccgccgtgg gcatgttttt cagggctctg ctcagcagct ccaggcggtc ccagagctcg	5100
gtgacgtgct ctacggcatc tctatccagc atatctcctc gtttcgcggg ttggggcgac	5160
tttcgctgta gggcaccaag cgggtggtcgt ccagcggggc cagagtcatg tccttccatg	5220
ggcgaggggt cctcgtcagg gtggtctggg tcacggtgaa ggggtgcgct ccgggctgag	5280
cgcttgccaa ggtgcgcttg aggctgggtc tgctgggtgct gaagcgctgc cggctctcgc	5340
cctgcgcgtc ggccaggtag catTTTgacca tgggtgtcata gtccagcccc tccgcggcgt	5400
gtcccttggc gcgcagcttg cccttgaggg tggcgccgca cgaggggcag agcaggctct	5460
tgagcgcgta gagcttgggg gcgaggaaga ccgattcggg ggagtaggcg tccgcgcgc	5520
agaccccgca cacggtctcg cactccacca gccagggtgag ctcggggcgc gccgggtcaa	5580
aaaccagggt tcccccatgc tttttgatgc gtttcttacc tggggtctcc atgagggtgt	5640

gtccccgctc	ggtgacgaag	aggctgtccg	tgtctccgta	gaccgacttg	aggggtcttt	5700
tctccagggg	ggtccctcgg	tcttctctgt	agaggaactc	ggaccactct	gagacgaagg	5760
cccgcgtcca	ggccaggacg	aaggaggcta	tgtgggaggg	gtagcggtcg	ttgtccacta	5820
gggggtccac	cttctccaag	gtgtgaagac	acatgtcgcc	ttcctcggcg	tccaggaagg	5880
tgattggctt	gtaggtgtag	gccacgtgac	cgggggttcc	tgacgggggg	gtataaaagg	5940
gggtgggggc	gcgctcgtcg	tcactctctt	ccgcatcgct	gtctgcgagg	gccagctgct	6000
ggggtgagta	ttccctctcg	aaggcgggca	tgacctccgc	gctgaggttg	tcagtttcca	6060
aaaacgagga	ggatttgatg	ttcacctgtc	ccgaggtgat	acctttgagg	gtaccgcgct	6120
ccatctggtc	agaaaacacg	atctttttat	tgtccagctt	ggtggcgaac	gacccgtaga	6180
gggcgttggg	gagcagcttg	gcgatggagc	gcagggtctg	gttcttgtcc	ctgtcggcgc	6240
gctccttggc	cgcgatgttg	agctgcacgt	actcgcgcgc	gacgcagcgc	cactcgggga	6300
agacggtggg	gcgctcgtcg	ggcaccaggc	gcacgcgcca	gccgcggttg	tgcaggggtga	6360
ccagggtccac	gctggtggcg	acctcgccgc	gcaggcgctc	gttgggtccag	cagagacggc	6420
cgcccttgcg	cgagcagaag	gggggcaggg	ggtcgagctg	ggtctcgtcc	gggggggtccg	6480
cgtccacggg	gaaaaccccc	gggcgcaggc	gcgcgtcgaa	gtagtctatc	ttgcaacctt	6540
gcatgtccag	cgcttctgtc	cagtcgcggg	cggcgagcgc	gcgctcgtag	gggttgagcg	6600
gcgggcccc	gggcatgggg	tgggtgagtg	cggaggcgta	catgccgcag	atgtcataga	6660
cgtagagggg	ctccgcagg	accccgatgt	aggtggggta	gcagcgggcg	ccgcggatgc	6720
tggcgcgcac	gtagtcatac	agctcgtgcg	agggggcgag	gaggtcgggg	cccaggttgg	6780
tgcgggcggg	gcgctccgtg	cggaagacga	tctgcctgaa	gatggcatgc	gagttggaag	6840
agatggtggg	gcgctggaag	acgttgaagc	tggcgtcctg	caggccgacg	gcgtcgcgca	6900
cgaaggaggc	gtaggagtgc	cgcagcttgt	gtaccagctc	ggcgggtgacc	tgcacgtcga	6960
gcgcgcagta	gtcgaggggtc	tcgcggatga	tgtcatattt	agcctgcccc	ttcttttttcc	7020
acagctcgcg	gttgaggaca	aactcttcgc	ggtctttcca	gtactcttgg	atcgggaaac	7080
cgtccggttc	cgaacggtaa	gagcctagca	tgtagaactg	gttgacggcc	tggtaggcgc	7140
agcagccctt	ctccacgggg	agggcgtagg	cctgcgcggc	cttgcggagc	gaggtgtggg	7200
tcagggcgaa	ggtgtccctg	accatgactt	tgaggtactg	gtgcttgaag	tcggagtctg	7260
cgcagccgcc	ccgctcccag	agcgagaagt	cgggtgcgctt	cttgagagcg	gggttgggca	7320
gagcgaagg	gacatcgttg	aagaggatth	tgcccgcgcg	gggcatgaag	ttgcgggtga	7380
tgcggaagg	ccccggcact	tcagagcggg	tgttgatgac	ctgggcggcg	agcacgatct	7440
cgtcgaagcc	gttgatgttg	tggcccacga	tgtagagtth	caggaagcgg	ggccggccct	7500
ttacggtggg	cagcttcttt	agctcttcgt	aggtgagctc	ctcgggcgag	gcgaggccgt	7560

gctcggccag	ggcccagtc	gcgaggtgcg	ggttgctct	gaggaaggac	tcccagaggt	7620
cgcgggccag	gagggctctgc	aggcggtccc	tgaaggtcct	gaactggcgg	cccacggcca	7680
ttttttcggg	ggtgatgcag	tagaaggtga	gggggtcttg	ctgccagcgg	tcccagtcga	7740
gctgcagggc	gaggtcgcgc	gcggcggtga	ccaggcgctc	gtcgcccccg	aatttcatga	7800
ccagcatgaa	gggcacgagc	tgctttccga	aggcccccat	ccaagtgtag	gtctctacat	7860
cgtaggtgac	aaagaggcgc	tccgtgcgag	gatgcgagcc	gatcggggaag	aactggatct	7920
cccgccacca	gttgaggag	tggctgttga	tgtggtggaa	gtagaagtcc	cgtcgccggg	7980
ccgaacactc	gtgctggctt	ttgtaaaagc	gagcgcagta	ctggcagcgc	tgcacgggct	8040
gtacctcatg	cacgagatgc	acctttcgcc	cgcgcacgag	gaagccgagg	ggaaatctga	8100
gcccccgcc	tggctcgcgg	catggctgg	gctcttctac	tttgatgcg	tgtccgtctc	8160
cgtctggctc	ctcgaggggt	gttacgggtg	agcggaccac	cacgccgcgc	gagccgcagg	8220
tccagatata	ggcgcgcggc	ggtcggagtt	tgatgacgac	atcgcgcagc	tgggagctgt	8280
ccatggtctg	gagctccgc	ggcggcggca	ggtcagccgg	gagttcttgc	aggttcacct	8340
cgcagagtcg	ggccagggcg	cggggcaggt	ctaggtggta	cctgatctct	aggggcgtgt	8400
tgggtggcgg	gtcgatggct	tgcaggagcc	cgcagccccg	gggggcgacg	acggtgcccc	8460
gcggggtgg	ggtggtgg	gcggtgcagc	tcagaagcgg	tgccgcgggc	gggcccccg	8520
aggtaggggg	ggctccggct	ccgcgggcag	gggcggcagc	ggcacgtcgg	cgtggagcgc	8580
gggcaggagt	tggtgctgtg	cccggaggtt	gctggcgaag	gcgacgacgc	ggcggttgat	8640
ctcctggatc	tggcgctct	gcgtgaagac	gacgggcccc	gtgagcttga	acctgaaaga	8700
gagttcgaca	gaatcaatct	cgggtgtcatt	gaccgcggcc	tggcgcagga	tctcctgcac	8760
gtctcccgag	ttgtcttgg	aggcgatctc	ggccatgaac	tgctcgatct	cttctctctg	8820
gaggtctccg	cgtccggcgc	gttccacgg	ggccgccagg	tcgttggaga	tgcgccccat	8880
gagctgcgag	aaggcgttga	gtccgcctc	gttccagact	cggctgtaga	ccacgcccc	8940
ctggtcatcg	cgggcgcgca	tgaccacctg	cgcgaggttg	agctccacgt	gccgcgcgaa	9000
gacggcgtag	ttgcgcagac	gctggaagag	gtagttgagg	gtggtggcgg	tgtgctcggc	9060
cacgaagaag	ttcatgacct	agcggcgcaa	cgtggattcg	ttgatgtccc	ccaaggcctc	9120
cagccgttcc	atggcctcgt	agaagtccac	ggcgaagtgg	aaaaactggg	agttgcgcgc	9180
cgacacggtc	aactctcct	ccagaagacg	gatgagctcg	gcgacggtgt	cgcgcacctc	9240
gcgctcgaag	gctatgggga	tctcttctc	cgctagcatc	accacctcct	cctcttctc	9300
ctcttctggc	acttccatga	tggcttctc	ctcttcgggg	ggtggcgggc	gcggcggtgg	9360
gggagggggc	gctctgcgcc	ggcgggcg	caccgggagg	cgggtccacga	agcgcgcgat	9420
catctccccg	cggcgggcg	gcatggtctc	ggtgacggcg	cggccgttct	ccggggggcg	9480

cagttggaag	acgccgccgg	acatctggtg	ctggggcggg	tggccgtgag	gcagcgagac	9540
ggcgctgacg	atgcatctca	acaattgctg	cgtaggtacg	cgcgcgaggg	acctgagggg	9600
gtccatatcc	accggatccg	aaaacctttc	gaggaaggcg	tctaaccagt	cgcagtcgca	9660
aggtaggctg	agcaccgtgg	cgggcggcgg	ggggtggggg	gagtgtcttg	cggaggtgct	9720
gctgatgatg	taattgaagt	aggcggactt	gacacggcgg	atggtcgaca	ggagcaccat	9780
gtccttgggt	ccggcctgct	ggatgcggag	gcggtcggct	atgcccagg	cttcgttctg	9840
gcatcggcgc	aggtccttgt	agtagtcttg	catgagcctt	tccaccggca	cctcttctcc	9900
ttcctcttct	gcttcttcca	tgtctgcttc	ggccctgggg	cggcgccgcg	ccccctgcc	9960
ccccatgcmc	gtgaccccg	acccctgag	cggttggagc	agggccagg	cggcgacgac	10020
gcgctcggcc	aggatggcct	gctgcacctg	cgtgagggcg	gtttggaagt	catccaagtc	10080
cacgaagcgg	tggtaggcgc	ccgtgttgat	ggtgtaggtg	cagttggcca	tgacggacca	10140
gttgacggtc	tggtggcccg	gttgcgacat	ctcgggtgac	ctgagtcgcg	agtaggcgcg	10200
ggagtccaag	acgtagtcgt	tgcaagtccg	caccaggtag	tggtagccca	ccaggaagtg	10260
cggcgccggc	tgccggtaga	ggggccagcg	caggggtggc	ggggctccgg	gggccaggtc	10320
ttccagcatg	aggcgggtgt	aggcgtagat	gtacctggac	atccagggtg	taccgcgggc	10380
ggtggtggag	gcgcgcggga	agtcgcgcac	ccggttccag	atgttgcgca	ggggcagaaa	10440
gtgctccatg	gtaggcgtgc	tctgtccagt	cagacgcgcg	cagtcgttga	tactctagac	10500
cagggaaaac	gaaagccggt	cagcgggcac	tcttccgtgg	tctggtgaat	agatcgcaag	10560
ggtatcatgg	cggagggcct	cggttcgagc	ccgggtccg	ggccggacgg	tccgccatga	10620
tccacgcggt	taccgcccgc	gtgtcgaacc	caggtgtgcg	acgtcagaca	acggtggagt	10680
gttccttttg	gcgtttttct	ggccgggcgc	cggcgccgcg	taagagacta	agccgcgaaa	10740
gcgaaagcag	taagtggctc	gctccccgta	gccggaggga	tccttgctaa	gggttgcggt	10800
gcggcgaaac	ccggttcgaa	tcccgtactc	gggcgggcgc	gaccgcgggc	taagggtgtg	10860
gattggcctc	cccctcgat	aaagaccccg	cttgccgatt	gactccggac	acggggacga	10920
gcccctttta	tttttgcttt	ccccagatgc	atccgggtgt	gcgacagatg	cgcctccgcg	10980
cccagcagca	gcaacaacac	cagcaagagc	ggcagcaaca	gcagcgggag	tcatgcaggg	11040
ccccctcacc	caccctcggc	ggcccgccca	cctcggcgct	cgcggccgtg	tctggcgcc	11100
gcggcgccgc	cggcgggggg	ccggctgacg	accccgagga	gccccgcgg	cgcagggcca	11160
gacactacct	ggacctggag	gagggcgagg	gcctggcgcg	gctggggggc	ccgtctcccg	11220
agcgccaccc	gcgggtgcag	ctaaagcgcg	actcgcgcca	ggcgtagctg	cctcggcaga	11280
acctgttcag	ggaccgcgcg	ggcgaggagc	ccgaggagat	gcgggacagg	aggttcagcg	11340
cggggcgggg	gctgcggcag	gggctgaacc	gcgagcggct	gctgcgcgag	gaggactttg	11400

agcccgacgc	gcggacgggg	atcagccccg	cgcgcgcgca	cgtggcgggc	gccgacctgg	11460
tgacggcgta	cgagcagacg	gtgaaccagg	agatcaactt	ccaaaagagt	ttcaacaacc	11520
acgtgcgcac	gctggtggcg	cgcgaggagg	tgaccatcgg	gctgatgcac	ctgtgggact	11580
ttgtgagcgc	gctggtgcag	aaccccaata	gcaagcctct	gacggcgcgag	ctgttcctga	11640
tagtgcagca	cagcagggac	aacgaggcgt	ttagggacgc	gctgctgaac	atcaccgagc	11700
ccgagggccg	gtggctgctg	gacctgatta	acatcctgca	gagcatagtg	gtgcaggagc	11760
gcagcctgag	cctggccgac	aaggtggcgg	ccatcaacta	ctcgatgctg	agcctgggca	11820
agttttacgc	gcgcaagatc	taccagacgc	cgtacgtgcc	catagacaag	gaggtgaaga	11880
tcgacggttt	ttacatgcgc	atggcgctga	aggtgctcac	cctaagcgac	gacctgggcg	11940
tgtaccgcaa	cgagcgcac	cacaaggccg	tgagcgtgag	ccggcgggcg	gagctgagcg	12000
accgcgagct	gatgcatagc	ctgcagcggg	cgctggcggg	cgccggcagc	ggcgacaggg	12060
aggcggagtc	ctacttcgat	gcggggggcg	acctgcgctg	ggcgcccagc	cggcggggcc	12120
tggaggccgc	gggggtccgc	gaggactatg	acgaggacgg	cgaggaggat	gaggagtacg	12180
agctagagga	gggcgagtac	ctggactaaa	ccgcgggtgg	tgtttccggt	agatgcaaga	12240
cccgaacgtg	gtggaccgcc	cgctgcgggc	ggctctgcag	agccagccgt	ccggccttaa	12300
ctcctcagac	gactggcgac	aggtcatgga	ccgcatcatg	tcgctgacgg	cgcgtaaccc	12360
ggacgcgttc	cggcagcagc	cgcaggccaa	caggctctcc	gccatcctgg	aggcgggtgt	12420
gcctgcgcgc	tcgaacccca	cgcacgagaa	ggtgctggcc	atagtgaacg	cgctggccga	12480
gaacagggcc	atccgcccgg	acgaggccgg	gctggtgtac	gacgcgctgc	tgcagcgcgt	12540
ggcccgtac	aacagcggca	acgtgcagac	caacctggac	cggctggtgg	gggacgtgcg	12600
cgaggcggtg	gcgcagcgcg	agcgcgcgga	tcggcagggc	aacctgggct	ccatggtggc	12660
gctgaatgcc	ttcctgagca	cgcagccggc	caacgtgccg	cgggggcagg	aagactacac	12720
caactttgtg	agcgcgctgc	ggctgatggt	gaccgagacc	ccccagagcg	aggtgtacca	12780
gtcgggtccg	gactacttct	tccagaccag	cagacagggc	ctgcagacgg	tgaacctgag	12840
ccaggctttc	aagaacctgc	gggggctgtg	gggcgtgaag	gcgcccaccg	gcgaccgggc	12900
gacggtgtcc	agcctgctga	cgcccaactc	gcgcctgctg	ctgctgctga	tcgcgccttt	12960
cacggacagc	ggcagcgtgt	cccgggacac	ctacctgggg	cacctgctga	ccctgtaccg	13020
cgaggccatc	gggcaggcgc	aggtggacga	gcacaccttc	caggagatca	ccagcgttag	13080
ccgcgcgctg	gggcaggagg	acacgagcag	cctggaggcg	actctgaact	acctgctgac	13140
caaccggcgg	cagaagattc	cctcgtgca	cagcctgacc	tccgaggagg	agcgcactct	13200
gcgctacgtg	cagcagagcg	tgagcctgaa	cctgatgcgc	gacgggggtga	cgcccagtgt	13260
ggcgctggac	atgaccgcgc	gcaacatgga	accgggcatg	tacgccgcgc	accggcctta	13320

catcaaccgc	ctgatggact	acctgcatcg	cgcgggcgcc	gtgaaccccg	agtactttac	13380
caacgccatc	ctgaacccgc	actggctccc	gccgcccggg	ttctacagcg	ggggcttcga	13440
ggtcccggag	gccaacgatg	gcttcctgtg	ggacgacatg	gacgacagcg	tgtttctccc	13500
gcggccgcag	gcgctggcgg	aagcgtccct	gctgctccc	aagaaggagg	aggaggcgag	13560
tcgccgccgc	ggcagcagcg	gcgtggcttc	tctgtccgag	ctggggggcg	cagccgccgc	13620
gcgccccggg	tccttggggc	gcagccccct	tccgagcctg	gtgggggtctc	tgcacagcga	13680
gcgcaccacc	cgccctcggc	tgctggggcg	ggacgagtag	ctgaataact	ccctgctgca	13740
gccggtgcgg	gagaaaaacc	tgccctccgc	cttccccaac	aacgggatag	agagcctggt	13800
ggacaagatg	agcagatgga	agacctatgc	gcaggagcac	agggacgcgc	ccgcgctccg	13860
gccgcccacg	cggcgccagc	gccacgaccg	gcagcggggg	ctggtgtggg	atgacgagga	13920
ctccgcggac	gatagcagcg	tgctggacct	gggagggagc	ggcaaccctg	tcgcgcacct	13980
gcgcccccg	ctgggggagga	tgttttaaaa	aaaaaaaaaa	gcaagaagca	tgatgcaaaa	14040
attaaataaa	actcaccaag	gccatggcga	ccgagcgctg	gtttcttgtg	ttcccttcag	14100
tatgcggcgc	gcggcgatgt	accaggaggg	acctcctccc	tcttacgaga	gcgtgggtggg	14160
cgcgggcgcg	gcggcgccct	cttctccctt	tgcgctcgag	ctgctggagc	cgcgctacgt	14220
gcctccgcgc	tacctgcggc	ctacgggggg	gagaaacagc	atccgttact	cggagctggc	14280
gcccctgttc	gacaccaccc	gggtgtacct	ggtggacaac	aagtcggcgg	acgtggcctc	14340
cctgaactac	cagaacgacc	acagcaatth	tttgaccacg	gtcatccaga	acaatgacta	14400
cagcccagag	gaggccagca	cccagaccat	caatctggat	gaccggctcg	actggggcgg	14460
cgacctgaaa	accatcctgc	acaccaacat	gcccacgtg	aacgagttca	tgttcaccaa	14520
taagttcaag	gcgcgggtga	tgggtgtcgc	ctcgcacacc	aaggaagacc	gggtggagct	14580
gaagtacgag	tgggtggagt	tcgagctgcc	agagggcaac	tactccgaga	ccatgaccat	14640
tgacctgatg	aacaacgcga	tcgtggagca	ctatctgaaa	gtgggcaggc	agaacggggg	14700
cctggagagc	gacatcgggg	tcaagttcga	caccaggaac	ttccgcctgg	ggctggaccc	14760
cgtgaccggg	ctggttatgc	ccgggggtga	caccaacgag	gccttccatc	ccgacatcat	14820
cctgctgccc	ggctgcgggg	tggacttcac	ttacagccgc	ctgagcaacc	tcctgggcat	14880
ccgcaagcgg	cagcccttcc	aggagggttt	caggatcacc	tacgaggacc	tggagggggg	14940
caacatcccc	gcgctcctcg	atgtggaggc	ctaccaggat	agcttgaagg	aaaatgaggc	15000
gggacaggag	gataccgccc	ccgccgcctc	cgccgccgcc	gagcagggcg	aggatgctgc	15060
tgacaccgcg	gccgcggacg	gggcggaggc	cgaccccgct	atggtggtgg	aggctgccga	15120
gcaggaggag	gacatgaatg	acagtgcggt	gcgcggagac	accttcgtca	ccggggggga	15180
ggaaaagcaa	gcggaggccg	aggccgcggc	cgaggaaaag	caactggcgg	cagcagcggc	15240

ggcggcgggcg	ttggccgcgg	cggaggctga	gtctgagggg	accaagcccg	ccaaggagcc	15300
cgtgattaag	cccctgaccg	aagatagcaa	gaagcgcagt	tacaacctgc	tcaaggacag	15360
caccaacacc	gcgtaccgca	gctggtacct	ggcctacaac	tacggcgacc	cgtcgacggg	15420
ggtgcgctcc	tggaccctgc	tgtgcacgcc	ggacgtgacc	tgcggtcgg	agcaggtgta	15480
ctggtcgctg	cccgacatga	tgcaagaccc	cgtgaccttc	cgctccacgc	ggcaggtcag	15540
caacttcccg	gtgggtggggc	ccgagctgct	gcccgtagac	tccaagagct	tctacaacga	15600
ccaggccgtc	tactcccagc	tcatccgcca	gttcacctct	ctgacctcacg	tgttcaatcg	15660
ctttcctgag	aaccagattc	tggcgcgccc	gcccgcctcc	accatcacca	ccgtcagtga	15720
aaacgttcct	gctctcacag	atcacgggac	gctaccgctg	cgcaacagca	tcggaggagt	15780
ccagcgagtg	accgttactg	acgccagacg	ccgcacctgc	ccctacgttt	acaaggcctt	15840
gggcatagtc	tcgccgcgcg	tcctttccag	ccgcactttt	tgagcaacac	caccatcatg	15900
tccatcctga	tctcaccag	caataactcc	ggctggggac	tgctgcgcgc	gcccagcaag	15960
atgttcggag	gggcgaggaa	gcgttccgag	cagcaccccg	tgcgcgtagc	cgggcacttc	16020
cgcgccccct	ggggagcgca	caaacgcggc	cgcggggggc	gcaccaccgt	ggacgacgcc	16080
atcgactcgg	tggtggagca	ggcgcgcaac	tacaggcccg	cggtctctac	cgtggacgcg	16140
gccatccaga	ccgtggtgcg	gggcgcgcg	cggtacgcca	agctgaagag	ccgccggaag	16200
cgcgtagggc	gccgccaccg	ccgccgaccc	ggggccgccc	ccaaacgcgc	cgccgcggcc	16260
ctgcttcgcc	gggccaagcg	cacgggcccgc	cgcgccgcca	tgagggccgc	gcgccgcttg	16320
gccgccggca	tcaccgccgc	caccatggcc	ccccgtaccc	gaagacgcgc	ggccgccgcc	16380
gccgccgccc	ccatcagtga	catggccagc	aggcgccggg	gcaacgtgta	ctgggtgcgc	16440
gactcggtga	ccggcacgcg	cgtgcccgtg	cgcttccgcc	ccccgcggac	ttgagatgat	16500
gtgaaaaaac	aacactgagt	ctcctgctgt	tgtgtgtatc	ccagcggcgg	cggcggcgcg	16560
cgcagcgtca	tgtccaagcg	caaaatcaaa	gaagagatgc	tccaggtcgt	cgcgccggag	16620
atctatgggc	ccccgaagaa	ggaagagcag	gattcgaagc	cccgaagat	aaagcgggtc	16680
aaaaagaaaa	agaaagatga	tggcgatgcc	gatggggagg	tggagttcct	gcgcgccacg	16740
gcgcccaggc	gcccgggtga	gtggaagggc	cggcgcgtaa	agcgcgctct	gcgccccggc	16800
accgcggtgg	tcttcacgcc	cggcgagcgc	tccaccggga	ctttcaagcg	cgtctatgac	16860
gaggtgtacg	gcgacgaaga	cctgctggag	caggccaacg	agcgcttcgg	agagtttgct	16920
tacgggaagc	gtcagcggcc	gctggggaag	gaggacctgc	tggcgctgcc	gctggaccag	16980
ggcaacccca	ccccagttct	gaagcccgtg	accctgcagc	aggtgctgcc	gagcagcgca	17040
ccctccgagg	cgaagcgggg	tctgaagcgc	gagggcggcg	acctggcgcc	caccgtgcag	17100
ctcatggtgc	ccaagcggca	gaggctggag	gatgtgctgg	agaaaatgaa	agtagacccc	17160

ggtctgcagc	cggacatcag	ggtccgtccc	atcaagcagg	tggcgccggg	cctcggcgtg	17220
cagaccgtgg	acgtgggtcat	ccccaccggc	aactcccccg	cggccaccac	cactaccgct	17280
gcctccacgg	acatggagac	acagaccgat	cccgcgcgag	cggccgccac	cggcgccgcc	17340
gcgacctcct	cggcgagggt	gcagacggac	ccctgggtgc	cggcggcgat	gtcagctccc	17400
cgcgcgcgtc	gcggggcgag	gaagtacggc	gccgccaacg	cgctcctgcc	cgagtacgcc	17460
ttgcatcctt	ccatcgcgcc	cacccccggc	taccgagggt	atacctaccg	cccgcgaaga	17520
gccaaggggt	ccaccgcgcg	tccccgccga	cgcgcgcgcc	ccaccacccg	cggccgcgcg	17580
cgcagacgcc	agcccgcact	ggctccagtc	tccgtgagga	gagtggcgcg	cgacggacac	17640
accctgggtg	tgcccagggc	gcgctaccac	cccagcatcg	tttaaaagcc	tgttgtgggt	17700
cttgagata	tggccctcac	ttgccgcctc	cgtttcccg	tgccgggata	cggaggagga	17760
agatcgcgcc	gcaggagggg	tctggccggc	cgcggcctga	gcggaggcag	cggccgcgcg	17820
caccggcggc	gacgcgccac	cagccgacgc	atgcgcggcg	gggtgctgcc	cctgttaatc	17880
cccctgatcg	cgcggcgat	cggcgccgtg	cccgggatcg	cctccgtggc	cttgaggcg	17940
tcccagaggc	attgacagac	ttgcaaactt	gcaaatatgg	aaaaaaacc	caataaaaaa	18000
gtctagactc	tcacgctcgc	ttggtcctgt	gactattttg	tagaatggaa	gacatcaact	18060
ttgcgtcgt	ggccccgcgt	cacggctcgc	gcccgttcct	gggacactgg	aacgatatcg	18120
gcaccagcaa	catgagcgg	ggcgccttca	gttggggctc	tctgtggagc	ggcattaaaa	18180
gtatcgggtc	tgccgttaaa	aattacggct	cccgggcctg	gaacagcagc	acgggccaga	18240
tgttgagaga	caagttgaaa	gagcagaact	tccagcagaa	ggtggtggag	ggcctggcct	18300
ccggcatcaa	cggggtggtg	gacctggcca	accaggccgt	gcagaataag	atcaacagca	18360
gactggaccc	cggccgcgcg	gtggaggagg	tgccgcgggc	gctggagacg	gtgtcccccg	18420
atgggcgtgg	cgagaagcgc	cgcgggcccc	ataggggaaga	gaccactctg	gtcacgcaga	18480
ccgatgagcc	gccccgtat	gaggaggccc	tgaagcaagg	tctgccacc	acgcggccca	18540
tcgcgcccat	ggccaccggg	gtggtggggc	gccacacccc	cgccacgctg	gacttgccct	18600
cgcgcgccga	tgtgccgcag	cagcagcaga	aggcggcaca	gccgggcccc	cccgtgaccg	18660
cctcccgttc	ctccgcgggt	cctctgcgcc	gcgcggccag	cggccccccg	gggggggtcg	18720
cgaggcacgg	caactggcag	agcacgctga	acagcatcgt	gggtctgggg	gtgcgggtccg	18780
tgaagcgccg	ccgatgctac	tgaatagctt	agctaacgtg	ttgtatgtgt	gtatgcgccc	18840
tatgtcgccg	ccagaggagc	tgctgagtcg	cgcgcgttcg	cgcgcccacc	accaccaccg	18900
ccactccgcc	cctcaagatg	gcgaccccat	cgatgatgcc	gcagtgggtcg	tacatgcaca	18960
tctcgggcca	ggacgcctcg	gagtacctga	gccccgggct	ggtgcagttc	gccgcgcgca	19020
ccgagagcta	cttcagcctg	agtaacaagt	ttaggaaccc	cacggtggcg	cccacgcacg	19080

atgtgaccac	cgaccggtct	cagcgccctga	cgctgcgggtt	cattcccgtg	gaccgcgagg	19140
acaccgcgta	ctcgtacaag	gcgcggttca	ccctggccgt	gggcgacaac	cgcggtgctgg	19200
acatggcctc	cacctacttt	gacatccgcg	gggtgctgga	ccggggcccc	actttcaagc	19260
cttactctgg	caccgcctac	aactccctgg	cccccaaggg	cgctcccaac	tcctgcgagt	19320
gggagcaatt	agaagaagcc	caggccgctg	tggaagacga	agaattagaa	gatgaagacg	19380
aggaaccaca	ggatgaggca	cctgtgaaaa	aaacccatgt	atacgctcag	gctccccctt	19440
ctggagaaga	aattactaaa	aacggtttgc	aaatagggtc	agataacaca	gaagcccagt	19500
ctaagcccat	atatgcagat	cctacattcc	agcctgaacc	ccaaatcggg	gaatcccagt	19560
ggaatgaggc	agatgctaca	gttgccggcg	gtagagtgt	aaagaaatcc	actcccatga	19620
agccatgcta	tggttcctat	gcaagacca	caaactccaa	tggagggtcaa	ggtgtgctgg	19680
tggctgatga	taaggggggtt	cttcaatcta	aagttgaatt	gcaatTTTTT	tcaaatacta	19740
ctactcttaa	tcagcgggag	ggtaacgata	caaaacaaaa	agtgggtgctg	tatagcgaag	19800
atgtgcacat	ggaaactcca	gacaccaca	tttcttacia	gccacaaaa	agcgatgaca	19860
attcaaaaat	catgctgggt	cagcagtcca	tgcccaacag	acctaattac	atcggtctca	19920
gagacaactt	tatcggcctc	atgtattaca	atagcactgg	caacatggga	gtgcttgacg	19980
gtcaggcctc	tcagttgaat	gcagtgggtg	acttgcaaga	cagaaacaca	gaactgtcct	20040
accagctctt	gcttgattcc	atgggtgaca	gaaccagata	cttttccatg	tggaatcagg	20100
cagtggacag	ttatgacca	gatgtcagaa	ttattgaaaa	tcatggaact	gaagacgagc	20160
tccccaacta	ttgtttccct	ctgggcgga	taggggtaac	tgacacttac	caggccatta	20220
aaaccaatgg	caatggtcaa	gaaaacccaa	cctgggaaaa	agatacagag	tttgcagacc	20280
gcaatgaaat	aggggtggga	aacaatttcg	ctatggagat	caacctcagt	gccaacctgt	20340
ggagaaaactt	cctgtactcc	aacgtggcgc	tgtacctgcc	agacaagctt	aagtacaacc	20400
cctccaatgt	ggacatctct	gacaaccca	acacctacga	ttacatgaac	aagcgagtgg	20460
tggccccggg	gctggtggac	tgctacatca	acctgggcgc	gcgctggctg	ctggactaca	20520
tggacaacgt	caacccttc	aaccaccacc	gcaatgcggg	cctgcgctac	cgctccatgc	20580
tcctgggcaa	cgggcgctac	gtgcccttcc	acatccaggt	gccccagaag	ttctttgcca	20640
tcaagaacct	cctcctcctg	ccgggctcct	acacctacga	gtggaacttc	aggaaggatg	20700
tcaacatggt	cctccagagc	tctctgggta	acgatctcag	ggtggacggg	gccagcatca	20760
agttcgagag	catctgcctc	tacgccacct	tcttccccat	ggccccacaac	acggcctcca	20820
cgctcgaggc	catgctcagg	aacgacacca	acgaccagtc	cttcaatgac	tacctctccg	20880
ccgccaacat	gctctacccc	ataccgcga	acgccaccaa	cgccccatc	tccatcccc	20940
cgcgcaactg	ggcggccttc	cgcggtggg	ccttcacccg	cctcaagacc	aaggagaccc	21000

cctccctggg	ctcgggattc	gacccctact	acacctactc	ggggtccatt	ccctacctgg	21060
acggcacctt	ctacctcaac	cacactttca	agaagggtctc	ggtcaccttc	gactcctcgg	21120
tcagctggcc	gggcaacgac	cgtctgctca	cccccaacga	gttcgaaatc	aagcgctcgg	21180
tcgacgggga	gggctacaac	gtggcccagt	gcaacatgac	caaggactgg	ttcctgggtcc	21240
agatgctggc	caactacaac	atcggtctacc	agggcttcta	catcccagag	agctacaagg	21300
acaggatgta	ctccttcttc	aggaacttcc	agcccatgag	cgggcaggtg	gtggaccaga	21360
ccaagtacaa	ggactaccag	gaggtgggca	tcattccacca	gcacaacaac	tcggggcttcg	21420
tgggctacct	cgccccacc	atgcgcgagg	gacaggccta	ccccgccaac	ttcccctacc	21480
cgctcatagg	caagaccgcg	gtcgacagca	tcaccagaa	aaagtccctc	tgcgatcgca	21540
ccctctggcg	catccccctc	tccagcaact	tcattgtccat	gggtgcgctc	tcggacctgg	21600
gccagaactt	gctctacgcc	aactccgccc	acgccctcga	catgaccttc	gaggtcgacc	21660
ccatggacga	gcccaccctt	ctctatgttc	tgttcgaagt	ctttgacgtg	gtccgggtcc	21720
accagccgca	ccgcggcgtc	atcgagaccg	tgtacctgcg	tacgcccttc	tcggccggca	21780
acgccaccac	ctaaagaagc	aagccgcagt	catcgccgcc	tgcattgccgt	cgggttccac	21840
cgagcaagag	ctcagggcca	tcgtcagaga	cctgggatgc	gggccctatt	ttttgggcac	21900
tttcgacaag	cgcttccctg	gctttgtctc	cccacacaag	ctggcctgcg	ccatcgtcaa	21960
cacggccggc	cgcgagaccg	ggggcggtca	ctggctggcc	ttcgccctgga	acccgcgctc	22020
caaaacatgc	ttcctctttg	accccttcgg	cttttcggac	cagcgggtca	agcaaattcta	22080
cgagttcgag	tacgagggct	tgctgcgtcg	cagcgccatc	gcctcctcgc	ccgaccgctg	22140
cgtcaccctc	gaaaagtcca	cccagaccgt	gcagggggccc	gactcggccg	cctgcgggtct	22200
cttctgctgc	atgtttctgc	acgcctttgt	gcactggcct	cagagtccca	tggaccgcaa	22260
ccccaccatg	aacttgctga	cgggggtgcc	caactccatg	ctccagagcc	cccaggctga	22320
gcccaccctg	cgccgcaacc	aggagcagct	ctacagcttc	ctggagcgcc	actcgcccta	22380
cttccgccgc	cacagcgac	agatcaggag	ggccacctcc	ttctgccact	tgcaagagat	22440
gcaagaaggg	taataacgat	gtacacactt	ttttctcaat	aaatggcatt	tttttattta	22500
tacaagctct	ctgggggtatt	catttcccac	caccaccacc	acccgccgtt	gtcgccatct	22560
ggctctattt	agaaatcgaa	agggttctgc	cgggagtcgc	cgtgcgccac	gggcagggac	22620
acgttgcgat	actggtagcg	ggtgccccac	ttgaactcgg	gcaccaccag	gcgaggcagc	22680
tcggggaagt	tttcgctcca	caggctgcgg	gtcagcacca	gcgcgttcat	caggctgggc	22740
gccgagatct	tgaagtcgca	gttggggccg	ccgccctgcg	cgcgcgagtt	gcggtacacc	22800
gggttgcgac	actggaacac	caacagcgcc	gggtgcttca	cgctagccag	cacgctgcgg	22860
tcgggagatca	gctcggcgtc	caggctcctcc	gcgttgctca	gcgcgaacgg	ggtcatcttg	22920

ggcacttgcc	tccccaggaa	ggg'gcgctgc	cccggtttctg	agttgcagtc	gcagcgcagc	22980
gggatcagca	ggtgcccattg	cccggactcg	gcgttgggggt	acagcgcgcg	catgaaggcc	23040
tgcattctggc	ggaaggccat	ctgggccttg	gcgccctccg	agaagaacat	gccgcaggac	23100
ttgcccgaga	actggtttgc	ggggcagctg	gcgtcgtgca	ggcagcagcg	cgcgctcggtg	23160
ttggcgatct	gcaccacggt	gcgccccac	cggttcttca	cgatcttggc	cttggacgat	23220
tgctccttca	gcgcgcgctg	cccgttctcg	ctggtcacat	ccatctcgat	cacatgttcc	23280
ttgttcacca	tgctgctgcc	gtgcaggcac	ttcagctcgc	cctccgtctc	ggtgcagcgg	23340
tgctgccaca	gcgcgcagcc	cgtgggctcg	aaagacttgt	aggtcacctc	cgcgaaggac	23400
tgcaggtacc	cctgcaaaaa	gcggcccatc	atggtcacga	aggtcttgtt	gctgctgaag	23460
gtcagctgca	gcccgcgggtg	ctcctcgttc	agccaggctc	tgcacacggc	cgccagcgcc	23520
tccacctggg	cgggcagcat	cttgaagttc	accttcagct	cattctccac	gtggtacttg	23580
tccatcagcg	tgcgcgccgc	ctccatgcc	ttctcccagg	ccgacaccag	cggcaggctc	23640
acgggggttct	tcaccatcac	cgtggccgcc	gcctccgccg	cgctttcgct	ttccgccccg	23700
ctgttctctt	cctcttcttc	ctcttctctg	ccgcgcgcca	ctcgcagccc	ccgcaccacg	23760
gggtcgtctt	cctgcaggcg	ctgcaccttg	cgcttgccgt	tgcgcccctg	cttgatgcgc	23820
acgggccccg	tgctgaagcc	caccatcacc	agcgcggcct	cttcttgctc	gtcctcgctg	23880
tccagaatga	cctccgggga	ggggggggtg	gtcatcctca	gtaccgaggc	acgcttcttt	23940
ttcttctctg	gggcgttcgc	cagctccgcg	gctgcggccg	ctgccgaggt	cgaaggccga	24000
gggctgggcg	tgcgcggcac	cagcgcgtcc	tgcgagccgt	cctcgtcctc	ctcggactcg	24060
agacggaggc	gggcccgtt	cttcgggggc	gcgcggggcg	gcggaggcgg	cggcggcgac	24120
ggagacgggg	acgagacatc	gtccagggtg	ggtggacggc	gggccgcgcc	gcgtccgcgc	24180
tcggggggtg	tctcgcgctg	gtcctcttcc	cgactggcca	tctcccactg	ctccttctcc	24240
tataggcaga	aagagatcat	ggagtctctc	atgcgagtcg	agaaggagga	ggacagccta	24300
accgccccct	ctgagccctc	caccaccgcc	gccaccaccg	ccaatgccgc	cgcggacgac	24360
gcgcccaccg	agaccaccgc	cagtaccacc	nnctcccca	gcgacgcacc	cccgtcgcag	24420
aatgaagtgc	tgatcgagca	ggaccggggt	tttgtgagcg	gagaggagga	tgaggtggat	24480
gagaaggaga	aggaggaggt	cgccgcctca	gtgccaaaag	aggataaaaa	gcaagaccag	24540
gacgacgcag	ataaggatga	gacagcagtc	gggcggggga	acggaagcca	tgatgctgat	24600
gacggctacc	tagacgtggg	agacgacgtg	ctgcttaagc	acctgcaccg	ccagtgcgtc	24660
atcgtctgcg	acgcgctgca	ggagcgggtg	gaagtgcgcc	tggacgtggc	ggaggtcagc	24720
cgcgctacg	agcggcacct	cttcgcgccg	cacgtgcgcc	ccaagcgccg	ggagaacggc	24780
acctgcgagc	ccaaccgcgc	tctcaacttc	taccgggtct	tcgcggtacc	cgaggtgctg	24840

gccacctacc	acatcttttt	cctaaactgc	aagatcccc	tctcctgccg	cgctaaccgc	24900
acccgcgcgc	acaaaaccct	gaccctgcgc	cagggcgccc	acataacctga	tattgcctct	24960
ctggaggaag	tgcccaagat	cttcgaggg	ctcggtcgcg	acgagaaacg	ggcggcgaa	25020
gctctgcacg	gagacagcga	aaacgagagt	cactcggggg	tgctggtgga	gctcgagggc	25080
gacaacgcgc	gcctggccgt	actcaagcgc	agcatagagg	tcaccactt	tgctacctgc	25140
gcgctcaacc	tgcccccaa	ggatcatgagt	gtggtcatgg	gcgagctcat	catgcgcgc	25200
gccagcccc	tgcccgcgga	tgcaaaactg	caagagtcct	cagaggaagg	cctgccccgc	25260
gtcagcgacg	agcagctggc	gcgctggctg	gagaccgcgc	accccgcgca	gctggaggag	25320
cggcgcaagc	tcatgatggc	cgcggtgctg	gtcacctggg	agctcgagtg	tctgcagcgc	25380
ttcttcgcgc	accccgagat	gcagcgcaag	ctcaggagga	cctgcacta	caccttcgc	25440
cagggctacg	tgcgccaggc	ctgcaagatc	tccaacgtgg	agctctgcaa	cctggtctcc	25500
tacctgggca	tcctgcacga	gaaccgcctc	gggcagaacg	tcctgcactc	cacctcaaa	25560
ggggaggcgc	gccgcgacta	catccgcgac	tgccctacc	ttctcctctg	ctacacctgg	25620
cagacggcca	tggggggtctg	gcagcagtg	ctggaggagc	gcaacctcaa	ggagctggaa	25680
aagctcctca	agcgaccct	cagggacctc	tggaacgggt	tcaacgagcg	ctcggtggcc	25740
gccgcgctgg	cggacatcat	cttccccgag	cgctgtctca	agacctgca	gcagggcctg	25800
cccgacttca	ccagccagag	catgctgcag	aacttcagga	ctttcatcct	ggagcgctcg	25860
ggcatcctgc	cggccacttg	ctgcgcgctg	cccagcgact	tcgtgcccat	caagtacagg	25920
gagtgcgcgc	cgccgctctg	gggccactgc	tacctcttcc	agctggccaa	ctacctcgcc	25980
tacctctcg	acctcatgga	agacgtgagc	ggcgagggcc	tgctcgagtg	ccactgcgc	26040
tgcaacctct	gcacgcccc	ccgctctcta	gtctgcaacc	cgagctgct	cagcgagagt	26100
cagattatcg	gtaccttcga	gctgcaggg	ccctgcctg	acgagaagtc	cgcggtccg	26160
gggctgaaac	tactccggg	gctgtggact	tccgcctacc	tacgcaaatt	tgtacctgag	26220
gactaccacg	cccacgagat	caggttctac	gaagaccaat	cccgcccgcc	caaggcggag	26280
ctcaccgcct	gcgtcatcac	ccaggggcac	atcctggggc	aattgcaagc	catcaacaaa	26340
gcccgcgcgag	agttcttgct	gaaaaaggg	cggggggtgt	acctggacct	ccagtccggc	26400
gaggagctaa	acccgctacc	cccgccgcgc	ccccagcagc	gggaccttgc	ttcccaggat	26460
ggcaccaga	aagaagcagc	agccgcgcgc	gcagccatac	atgcttctgg	aggaagagga	26520
ggaggactgg	gacagtcagg	cagaggagat	gatggaagac	tgaggaggag	acagcagcct	26580
agacgaggaa	gcttcagagg	ccgaagaggt	ggcagacgca	acaccatcac	cctcggtcgc	26640
agccccctcg	ccggggcccc	tgaaatcctc	cgaaccagc	accagcgcta	taacctccgc	26700
tcctccggcg	ccggcgccac	ccgcccgcag	acccaaccgt	agatgggaca	ccacaggaac	26760

cggggtcggt	aagtccaagt	gcccgcgcgc	gccaccgcag	cagcagcagc	agcagcgcca	26820
gggctaccgc	tcgtggcgcg	ggcacaagaa	cgccatagtc	gcctgcttgc	aagactgcgg	26880
gggcaacatc	tctttcgccc	gccgcttcct	gctattccac	cacggggtcg	cctttccccg	26940
caatgtcctg	cattactacc	gtcatctcta	cagcccctac	tgcagcgggc	accagagggc	27000
ggcagcgcca	gccacagcgg	cgaccaccac	ctaggaagat	atcctccgcg	ggcaagacag	27060
cggcagcagc	ggccaggaga	cccgcggcag	cagcggcggg	agcgggtggc	gcaactgcgc	27120
tctcgcccaa	cgaacccttc	tcgaccgggg	agctcagaca	caggatcttc	cccactttgt	27180
atgccatctt	ccaacagagc	agaggccagg	agcaggagct	gaaaataaaa	aacagatctc	27240
tgcgctccct	caccgcagc	tgtctgtatc	acaaaagcga	agatcagctt	cggcgcacgc	27300
tggaggacgc	ggaggcactc	ttcagcaa	actgcgcgct	cactcttaaa	gactagctcc	27360
gcgccttct	cgaatttagg	cgggagaaaa	ctacgtcatc	gccggccgcg	gccagccccg	27420
cccagccgag	atgagcaaag	agattcccac	gccatacatg	tggagctacc	agccgcagat	27480
gggactcgcg	gcgggagcgg	cccaggacta	ctccaccgcg	atgaactaca	tgagcgcggg	27540
accccatatg	atctcacagg	tcaacgggat	ccgcgcccag	cgaaaccaaa	tactgctgga	27600
acaggcgggc	atcaccgcca	cgccccgcca	taatctcaac	ccccgaaatt	ggcccccgcg	27660
cctagtgtac	caggaaacct	cctccgccac	caccgtacta	cttccgcgtg	acgcccaggc	27720
cgaagtccag	atgactaact	caggggcgca	gctcgcgggc	ggctttcgtc	acggggcgcg	27780
gccgctccga	ccaggataaa	gacacctgat	gatcagaggc	cgaggatatc	agctcaacga	27840
cgagtcggtg	agctcttcgc	tcggtctccg	tcgggacgga	actttccagc	tcgccggatc	27900
cggtcgctct	tcgttcacgc	cccgccaggc	gtacctgact	ctgcagacct	cgctcctcga	27960
gccccgctcc	ggcggcatcg	gaaccctcca	gttcgtggag	gagttcgtgc	cctcgggtcta	28020
cttcaacccc	ttctcgggac	ctccccggag	ctaccccgac	cagttcatte	cgaactttga	28080
cgcggtgaag	gactcggcgg	acggctacga	ctgaatgtca	ggtgccgagg	cagagcagct	28140
tcgcctgaga	cacctcgagc	actgccgccg	ccacaagtgc	ttcgcccgcg	gttcgggtga	28200
gttctgctac	tttcagctac	ccgaggagca	taccgagggg	ccggcgcacg	gcgtccgcct	28260
gaccaccag	ggcgaggtta	cctgttcctc	catccgggag	ttcacctcc	gtccccctgt	28320
agtggagcgg	gagcggggtc	cctgtgtcct	aactatcgcc	tgcaactgcc	ctaaccctgg	28380
attacatcaa	gatctttgct	gtcatctctg	tgctgagttt	aataaacgct	gagatcagaa	28440
tctactgggg	ctcctgtcgc	catcctgtga	acgccaccgt	cttcaccac	cccgaccagg	28500
cccaggcgaa	cctcacctgc	ggtctgcac	ggaggcca	gaagtacctc	acctgggtact	28560
tcaacggcac	cccctttgtg	gtttacaaca	gcttcgacgg	ggacggagtc	tcctgaaag	28620
accagctctc	cgggtctcagc	tactccatcc	acaagaacac	caccctccaa	ctcttccttc	28680

cctacctgcc	gggaacctac	gagtgcgtca	ccggccgctg	caccacacctc	acccgcctga	28740
tcgtaaacca	gagctttccg	ggaacagata	actccctctt	ccccagaaca	ggaggtgagc	28800
tcaggaaact	ccccggggac	cagggcggag	acgtaccttc	gacccttggtg	gggttaggat	28860
tttttattac	cgggttgctg	gctcttttaa	tcaaagcttc	cttgagattt	gttctttcct	28920
tctacgtgta	tgaacacctc	agcctccaat	aactctaccc	tttcttcggg	atcaggtgac	28980
ttttctgaaa	tcgggcttgg	tgtgctgctt	actctgttga	tttttttctt	tatcatactc	29040
agccttctgt	gcctcaggct	cgccgcctgc	tgcgcacaca	tctatatcta	ctgctgggtg	29100
ctcaagtgca	ggggtcgcca	ccaagatga	acaggtacat	ggtcctatcg	atcctaggcc	29160
tgctggccct	ggcggcctgc	agcgccgcca	aaaaagagat	tacctttgag	gagcccgctt	29220
gcaatgtaac	tttcaagccc	gagggtgacc	aatgcaccac	cctcgtcaaa	tgcgttacca	29280
atcatgagaa	gctgcgcata	gactacaaaa	acaaaactgg	cgggtttgcg	gtctatagtg	29340
tgtttacgcc	cggagacccc	tctaactact	ctgtcacctg	cttcaggggc	ggacagtcta	29400
agatattcaa	ttacactttc	cctttttatg	agttgtgcga	tgcggtcatg	tacatgtcaa	29460
aacagtacaa	cctgtggcct	ccctctcccc	aggcgtgtgt	ggaaaatact	gggtcttact	29520
gctgtatggc	tttggcaatc	actacgctcg	ctctaattctg	cacgggtgcta	tatataaaat	29580
tcaggcagag	gcgaatcttt	atcgatgaaa	agaaaatgcc	ttgatcgcta	acaccggctt	29640
tctatctgca	gaatgaatgc	aatcacctcc	ctactaatca	ccaccaccct	ccttgcgatt	29700
gcccattgggt	tgacacgaat	cgaagtgcc	gtgggggtcca	atgtcaccat	ggtgggcccc	29760
gccggcaatt	ccaccctcat	gtgggaaaaa	tttgtccgca	atcaatgggt	tcattttctgc	29820
tctaaccgaa	tcagtatcaa	gccagagcc	atctgcgatg	ggcaaaatct	aactctgatc	29880
aatgtgcaaa	tgatggatgc	tgggtactat	tacgggcagc	ggggagaaat	cattaattac	29940
tggcgacccc	acaaggacta	catgctgcat	gtagtcgagg	cacttcccac	taccaccccc	30000
actaccacct	ctcccaccac	cactaccact	actactacta	ctactaccac	taccgctgcc	30060
cgccataccc	gcaaaagcac	catgattagc	acaaagcccc	ctcgtgctca	ctcccacgcc	30120
ggcggggcca	tcggtgcgac	ctcagaaacc	accgagcttt	gcttctgcc	atgcactaac	30180
gccagcgctc	atgaactggt	cgacctggag	aatgaggatg	cccagcagag	ctccgcttgc	30240
ctgaccaggg	aggctgtgga	gcccgttgcc	ctgaagcaga	tcggtgattc	aataattgac	30300
tcttcttctt	ttgccactcc	cgaataccct	cccgattcta	ctttccacat	cacgggtacc	30360
aaagacccta	acctctcttt	ctacctgatg	ctgctgctct	gtatctctgt	ggtctcttcc	30420
gcgctgatgt	tactggggat	gttctgctgc	ctgatctgcc	gcagaaagag	aaaagctcgc	30480
tctcagggcc	aaccactgat	gcccttcccc	tacccccggg	attttgcaga	taacaagata	30540
tgagctcgct	gctgacacta	accgctttac	tagcctgcgc	tctaaccctt	gtcgcttgcg	30600

actcgagatt	ccacaatgtc	acagctgtgg	caggagaaaa	tgttactttc	aactccacgg	30660
ccgataccca	gtggtcgtgg	agtggctcag	gtagctactt	aactatctgc	aatagctcca	30720
cttccccag	catatcccca	accaagtacc	aatgcaatgc	cagcctgttc	accctcatca	30780
acgcttccac	cctggacaat	ggactctatg	taggctatgt	accctttggt	gggcaaggaa	30840
agaccacgc	ttacaacctg	gaagttcgcc	agcccagaac	cactacccaa	gctwcymcca	30900
ycaccagcac	cagcagcagc	agccacagca	gcagcagcag	attattgact	ttgggttttg	30960
ccagctcatc	tgccgctacc	caggccatct	acagctctgt	gcccgaacc	actcagaccc	31020
accgcccaga	aacgaccacc	gccaccaccc	tacacacctc	cagcgatcag	atgccgacca	31080
acatcacccc	cttggctctt	caaatgggac	ttacaagccc	cactccaaaa	ccagtggatg	31140
cggccgaggt	ctccgccctc	gtcaatgact	gggcggggct	gggaatgtgg	tggttcgcca	31200
taggcatgat	ggcgctctgc	ctgcttctgc	tctggctcat	ctgctgcctc	caccgcaggc	31260
gagccagacc	ccccatctat	agacccatca	ttgtcctgaa	ccccgataat	gatgggatcc	31320
atagattgga	tggcctgaaa	aacctacttt	tttcttttac	agtatgataa	attgagacat	31380
gcctcgcatt	ttcttgtaca	tgttccttct	cccacctttt	ctggggtggt	ctacgctggc	31440
cgctgtgtct	cacctggagg	tagactgcct	ctcacccttc	actgtctacc	tgctttacgg	31500
attggtcacc	ctcactctca	tctgcagcct	aatcacagta	atcatcgcct	tcatccagtg	31560
cattgattac	atctgtgtgc	gcctcgcata	cttcagacac	caccgcagct	accgagacag	31620
gaacattgcc	caacttctaa	gactgctcta	atcatgcata	agactgtgat	ctgccttctg	31680
atcctctgca	tcttgcccac	cctcacctcc	tgccagtaca	ccacaaaatc	tccgcgcaaa	31740
agacatgcct	cctgccgctt	cacccaactg	tggaatatac	ccaaatgcta	caacgaaaag	31800
agcgagctct	ccgaagcttg	gctgtatggg	gtcatctgtg	tcttagtttt	ctgcagcact	31860
gtctttgccc	tcatgatcta	cccctacttt	gatttgggat	ggaacgcgat	cgatgccatg	31920
aattacccca	cctttcccg	acccgagata	attccactgc	gacaagttgt	accgttgtc	31980
gttaatcaac	gcccccatc	ccctacgcc	actgaaatca	gctactttta	cctaacaggc	32040
ggagatgact	gacgccctag	atctagaaat	ggacggcatc	agtaccgagc	agcgtctcct	32100
agagaggcgc	aggcaggcgg	ctgagcaaga	gcgcctcaat	caggagctcc	gagatctcgt	32160
taacctgcac	cagtgcaaaa	gaggcatctt	ttgtctggta	aagcaggcca	aagtcaccta	32220
cgagaagacc	ggcaacagcc	accgcctcag	ttacaaattg	cccaccagc	gccagaagct	32280
ggtgctcatg	gtgggtgaga	atcccatcac	cgtcaccag	cactcggtag	agaccgaggg	32340
gtgtctgcac	tccccctgtc	ggggtccaga	agacctctgc	accctggtaa	agaccctgtg	32400
cggctctcaga	gatttagtcc	cctttaacta	atcaaacact	ggaatcaata	aaaagaatca	32460
cttacttaaa	atcagacagc	aggtctctgt	ccagtttatt	cagcagcacc	tccttcccct	32520

cctcccaact	ctggtactcc	aaacgccttc	tggcggcaaa	cttcctccac	accctgaagg	32580
gaatgtcaga	ttcttgctcc	tgtccctccg	caccactat	cttcattgtt	ttgcagatga	32640
agcgcaccaa	aacgtctgac	gagagcttca	accccggtga	cccctatgac	acggaaagcg	32700
gccctccctc	cgtccctttc	ctcacccttc	ccttcgtgtc	tcccgatgga	ttccaagaaa	32760
gtcccccccg	ggtcctgtct	ctgaacctgg	ccgagccctt	ggtcacttcc	cacggcatgc	32820
tcgcccctgaa	aatgggaagt	ggcctctccc	tggacgacgc	tggcaacctc	acctctcaag	32880
atatcaccac	cgctagccct	cccctcaaaa	aaaccaagac	caacctcagc	ctagaaacct	32940
catccccctt	aactgtgagc	acctcaggcg	ccctcacctg	agcagccgcc	gctccccctg	33000
cgggtggccg	cacctccctc	accatgcaat	cagaggcccc	cctgacagta	caggatgcaa	33060
aactcacctt	ggccaccaa	ggccccctga	ccgtgtctga	aggcaaactg	gccttgcaaa	33120
catcggcccc	gctgacggcc	gctgacagca	gcacctcac	cgttagcgcc	acaccaccaa	33180
ttaatgtaag	cagtggaagt	ttaggcttag	acatggaaga	ccctatgtat	actcacaatg	33240
gaaaactggg	aataagaatt	gggggtccac	taagagtagt	agacagcttg	catacactca	33300
ctgtagttac	cggaaatgga	ctaactgtag	ataacaatgc	cctccaaact	aaagttacgg	33360
gcgccctagg	ttatgacaca	tcaggaaatc	tacaattaag	agctgcagga	ggtatgcgaa	33420
ttgacgcaaa	tggccaactt	atccttaatg	tggcataccc	atttgatgct	cagaacaatc	33480
tcagccttag	acttggtcag	ggaccctgtt	atataaacac	agaccacaac	ctggatttga	33540
attgcaacag	aggtctaacc	acaactacca	ccaacaacac	aaaaaaactt	gagactaaaa	33600
ttagctcagg	cttagactat	gacaccaatg	gtgctgtcat	tattaaactt	ggcactggtc	33660
taagcttcga	caacacaggc	gccctaactg	tgggaaacac	tggtgatgat	aaactgactc	33720
tgtggacgac	cccagaccca	tctccaaatt	gcagaattca	ctcagacaaa	gactgcaagt	33780
ttactctagt	cctaactaag	tgtggaagcc	aaatcctggc	ctctgtcgcc	gccctagcgg	33840
tatcaggaaa	tctggcttcg	ataacaggca	ccgttgccag	cgttaccatc	tttctcagat	33900
ttgatcagaa	tggagtgcct	atggaaaact	cctcgctaga	caggcagtac	tggaaacttca	33960
gaaatggcaa	ctcaactaac	gctgccccct	acaccaatgc	agttgggttc	atgccaaacc	34020
tcgcagcata	ccccaaaacg	caaagccaga	ctgctaataa	caacattgta	agtcaggttt	34080
acttgaatgg	agacaaatcc	aaacccatga	cccttaccat	cacctcaat	ggaactaatg	34140
aatccagtga	aactagccag	gtgagtcact	actccatgtc	atttacctgg	gcttgggaaa	34200
gtgggcaata	tgccactgaa	acctttgcc	ccaactcctt	caccttttct	tacattgctg	34260
aacaataaaa	agcatgacac	tgatgttcat	ttctgattct	tattttatta	ttttcaaaca	34320
caacaaaatc	attcaagtca	ttcttccatc	ttagcttaat	agacacagta	gcttaataga	34380
cccagtagtg	caaagcccca	ttctagctta	tagatcagac	agtgataatt	aaccaccacc	34440

accaccatac	cttttgattc	aggaaatcat	gatcatcaca	ggatcctagt	cttcaggccg	34500
ccccctccct	cccaagacac	agaatacaca	gtcctctccc	cccgactggc	tttaaataac	34560
accatctggc	tggtcacaga	catgttctta	ggggttatat	tccacacggc	ctcctgccgc	34620
gccaggcgct	cgtcgggat	gttgataaac	tctcccggca	gctcgctcaa	gttcacgtcg	34680
ctgtccagcg	gctgaacctc	cggctgacgc	gataactgtg	cgaccggctg	ctggacaaac	34740
ggaggcccg	cctacaaggg	ggtagagtca	taatcctcgg	tcaggatagg	gcgggtgatgc	34800
agcagcagcg	agcgaaacat	ctgctgccgc	cgccgctccg	tccggcagga	aaacaacacg	34860
ccggtggctc	cctccgcgat	aatccgcacc	gcccgcagca	tcagcttcct	cgttctccgc	34920
gcgagcacc	tcacctgat	ctcgctcaag	tcggcgagct	aggtacagca	cagcaccacg	34980
atgttattca	tgatcccaca	gtgcagggcg	ctgtatccaa	agctcatgcc	gggaaccacc	35040
gccccacgt	ggccatcgta	ccacaagcgc	acgtaaatta	agtgtcgacc	cctcatgaac	35100
gtgctggaca	caaacattac	tctcttgggc	atgttgtaat	tcaccacctc	ccggtaccag	35160
ataaacctct	ggttaaacag	ggcaccttcc	accaccatcc	tgaaccaaga	ggccagaacc	35220
tgcccaccgg	ctatgcactg	caggggaacc	gggttggaac	aatgacaatg	cagactccaa	35280
ggctcgtaac	cgtaggatcat	ccggctgctg	aaggcatcga	tgttggcaca	acacagacac	35340
acgtgcatgc	actttctcat	gattagcagc	tcttccctcg	tcaggatcat	atcccaagga	35400
ataaccatt	cttgaatcaa	cgtaaaacc	acacagcagg	gaaggcctcg	cacataactc	35460
acgttgtagc	tggtcagcgt	gttgcatctc	ggaaacagcg	gatgatcctc	cagtatcgag	35520
gcgcgggtct	ccttctcaca	gggaggtaaa	gggtccctgc	tgtacggact	gcgcggggac	35580
gaccgagatc	gtgttgagcg	tagtgtcatg	gaaaagggaa	cgccggacgt	ggtcatactt	35640
cttgaagcag	aaccaggttc	gcgcgtggca	ggcctccttg	cgtctgcggc	ctcgccgtct	35700
agctcgctcc	gtgtgatagt	tgtagtacag	ccactcccgc	agagcgtcga	ggcgcacctc	35760
ggcttccgga	tctatgtaga	ctccgtcttg	caccgcggcc	ctgataatat	ccaccaccgt	35820
agaataagca	acaccagcc	aagcaatata	ctcgctctgc	gagcggcaga	caggaggagc	35880
gggcagagat	gggagaacca	tgataaaaaa	ctttttttta	agaatatttt	ccaattcttc	35940
gaaagtaaga	tctatcaagt	ggcagcgctc	ccctccactg	gcgcgggtcaa	actctacggc	36000
caaagcacag	acaacggcat	ttctaagatg	ttccttaatg	gcgtccaaaa	gacacaccgc	36060
tctcaagtcg	cagtaaaacta	tgaatgaaaa	cccatccggc	tgattttcca	atatagacgc	36120
gccggcggcg	tccaccaaac	ccagataatt	ttcttctctc	cagcggttta	gaatctgtct	36180
aagcaaatcc	cttatatcaa	gtccggccat	gccaaaaatc	tgctcaagag	cgccctccac	36240
cttcatgacc	aagcagcgca	tcatgattgc	aaaaattcag	gttcttcaga	gacctgtata	36300
agattcaaaa	tgggaacatt	aacaaaaatt	cctctgtcgc	gcagatccct	tcgcagggca	36360

agctgaacat	aatcagacag	gtctgaacgg	accagtgagg	ccaaatcccc	accaggaacc	36420
agatccagag	accctatact	gattatgacg	cgcatactcg	gggctatgct	gaccagcgta	36480
gcgccgatgt	aggcgtgctg	catggggcggc	gagataaaat	gcaaagtgct	ggttaaaaaa	36540
tcaggcaaag	cctcgcgcaa	aaaagctaac	acatcataat	catgctcatg	caggtagttg	36600
caggtaagct	caggaaccaa	aacggaataa	cacacgattt	tcctctcaaa	catgacttgc	36660
cggatactgc	gtaaaacaaa	aattataaat	aaaaaattaa	ttaacttaaa	cattggaagc	36720
ctgtctcaca	acaggaaaaa	ccactttaat	caacataaga	cggggccacgg	gcatgccggc	36780
atagccgtaa	aaaaattggg	ccccgtgatt	aacaagtacc	acagacagct	ccccgggcat	36840
gtcgggggtc	atcatgtgag	actctgtata	cacgtctgga	ttgtgaacat	cagacaaaca	36900
aagaaatcga	gccacgtagc	ccggaggtat	aatcacccgc	aggcggagggt	acagcaaaac	36960
gacccccata	ggaggaatca	caaaattagt	aggagaaaaa	aatacataaa	caccagaaaa	37020
accctgttgc	tgaggcaaaa	tagcgccctc	ccgatccaaa	acaacataaa	gcgcttccac	37080
aggagcagcc	ataacaaaaga	cccgagtctt	accagtaaaa	agaaaaaaga	tctctcaacg	37140
cagcaccagc	accaacactt	cgcagtgtaa	aaggccaagt	gccgagagag	tatatatagg	37200
aataaaaagt	gacgtaaacg	ggcaaagtcc	aaaaaacgcc	cagaaaaacc	gcacgcgaac	37260
ctacgccccg	aaacgaaagc	caaaaaacac	tagacactcc	cttcgggcgt	caacttccgc	37320
tttcccacgc	tacgtcactt	gccccagtca	aacaaactac	atatcccgaa	cttccaagtc	37380
gccacgcca	aaacaccgcc	tacacctccc	cgcccgcggg	cccgccccca	aaccgcctc	37440
ccgccccgcg	ccccgccccg	cgccgccccat	ctcattatca	tattggcttc	aatccaaaat	37500
aaggtatatt	attgatgatg	gtttaaacgg	atccaattct	tgaagacgaa	agggcctcgt	37560
gatacgcta	tttttatagg	ttaatgtcat	gataataatg	gtttcttaga	cgtcagggtg	37620
cacttttcgg	ggaaatgtgc	gcggaacccc	tatttgttta	tttttctaaa	tacattcaaa	37680
tatgtatccg	ctcatgagac	aataaccctg	ataaatgctt	caataatatt	gaaaaaggaa	37740
gagtatgagt	attcaacatt	tccgtgtcgc	ccttattccc	ttttttgcgg	cattttgcct	37800
tcctgttttt	gctcacccag	aaacgctggg	gaaagtaaaa	gatgctgaag	atcagttggg	37860
tgcacgagtg	ggttacatcg	aactggatct	caacagcggg	aagatccttg	agagttttcg	37920
ccccgaagaa	cgttttccaa	tgatgagcac	ttttaaagtt	ctgctatgtg	gcgcgggtatt	37980
atccccgtgtt	gacgccgggc	aagagcaact	cggtcgcgcg	atacactatt	ctcagaatga	38040
cttggttgag	tactcaccag	tcacagaaaa	gcatcttacg	gatggcatga	cagtaagaga	38100
attatgcagt	gctgccataa	ccatgagtga	taacactgcg	gccaacttac	ttctgacaac	38160
gatcggagga	ccgaaggagc	taaccgcttt	tttgacaaac	atgggggatc	atgtaactcg	38220
ccttgatcgt	tgggaaccgg	agctgaatga	agccatacca	aacgacgagc	gtgacaccac	38280

gatgcctgta gcaatggcaa caacgttgcg caaactatta actggcgaac tacttactct	38340
agcttccccg caacaattaa tagactggat ggaggcggat aaagttgcag gaccacttct	38400
gcgctcggcc cttccggctg gctggtttat tgctgataaa tctggagccg gtgagcgtgg	38460
gtctcgcggt atcattgcag cactggggcc agatggtaag ccctcccgta tcgtagttat	38520
ctacacgacg gggagtcagg caactatgga tgaacgaaat agacagatcg ctgagatagg	38580
tgcctcactg attaagcatt ggtaactgtc agaccaagtt tactcatata tactttagat	38640
tgatttaaaa ggatctaggt gaagatcctt ttgataatc tcatgaccaa aatcccttaa	38700
cgtgagtttt cgttccactg agcgtcagac cccgtagaaa agatcaaagg atcttcttga	38760
gacccctttt ttctgcgct aatctgctgc ttgcaaacia aaaaaccacc gctaccagcg	38820
gtggtttgtt tgccggatca agagctacca actctttttc cgaaggtaac tggcttcagc	38880
agagcgcaga taccaaatac tgtccttcta gtgtagccgt agttaggcca ccacttcaag	38940
aactctgtag caccgcctac atacctcgct ctgctaattc tgttaccagt ggctgctgcc	39000
agtggcgata agtcgtgtct taccgggttg gactcaagac gatagttacc ggataaggcg	39060
cagcggtcgg gctgaacggg gggttcgtgc acacagccca gcttggagcg aacgacctac	39120
accgaactga gatacctaca gcgtgagcta tgagaaagcg ccacgcttcc cgaagggaga	39180
aaggcggaca ggtatccggt aagcggcagg gtcggaacag gagagcgcac gagggagctt	39240
ccagggggaa acgcctggta tctttatagt cctgtcgggt ttcgccacct ctgacttgag	39300
cgtcgatttt tgtgatgctc gtcagggggg cggagcctat ggaaaaacgc cagcaacgcg	39360
gcctttttac ggttcctggc cttttgctgg ccttgaagct gtccttgatg gtcgtcatct	39420
acctgcctgg acagcatggc ctgcaacgcg ggcattccga tgccgccgga agcgagaaga	39480
atcataatgg ggaaggccat ccagcctcgc gtcgcagatc cgaattcgtt taaac	39535

<210> 16
 <211> 1500
 <212> DNA
 <213> Human immunodeficiency virus

<400> 16	
atgggtgcta gggcttctgt gctgtctggt ggtgagctgg acaagtggga gaagatcagg	60
ctgaggcctg gtggcaagaa gaagtacaag ctaaagcaca ttgtgtgggc ctccagggag	120
ctggagaggt ttgctgtgaa ccctggcctg ctggagacct ctgaggggtg caggcagatc	180
ctgggccagc tccagccctc cctgcaaaca ggctctgagg agctgaggtc cctgtacaac	240
acagtggcta ccctgtactg tgtgcaccag aagattgatg tgaaggacac caaggaggcc	300
ctggagaaga ttgaggagga gcagaacaag tccaagaaga aggcccagca ggctgctgct	360
ggcacaggca actccagcca ggtgtcccag aactacccca ttgtgcagaa cctccagggc	420

cagatggtgc accagggccat ctcccccccg accctgaatg cctgggtgaa ggtggtggag	480
gagaaggcct tctcccctga ggtgatcccc atgttctctg ccctgtctga gggtgccacc	540
ccccaggacc tgaacaccat gctgaacaca gtggggggcc atcaggctgc catgcagatg	600
ctgaaggaga ccatcaatga ggaggctgct gagtgggaca ggctgcatcc tgtgcacgct	660
ggccccattg cccccggcca gatgaggag cccaggggct ctgacattgc tggcaccacc	720
tccaccctcc aggagcagat tggctggatg accaacaacc ccccatccc tgtgggggaa	780
atctacaaga ggtggatcat cctgggcctg aacaagattg tgaggatgta ctccccacc	840
tccatcctgg acatcaggca gggccccaag gagcccttca gggactatgt ggacaggttc	900
tacaagacc tgagggctga gcaggcctcc caggaggtga agaactggat gacagagacc	960
ctgctggtgc agaatgcaa ccctgactgc aagaccatcc tgaaggccct gggccctgct	1020
gccaccctgg aggagatgat gacagcctgc caggggggtgg ggggccctgg tcacaaggcc	1080
agggtgctgg ctgaggccat gtcccagggtg accaactccg ccaccatcat gatgcagagg	1140
ggcaacttca ggaaccagag gaagacagtg aagtgccttca actgtggcaa ggtggggccac	1200
attgccaaga actgtagggc ccccaggaag aagggtgctg ggaagtgtgg caaggagggc	1260
caccagatga aggactgcaa tgagaggcag gccaaacttc tgggcaaaat ctggccctcc	1320
cacaagggca ggcctggcaa cttcctccag tccaggcctg agcccacagc ccctcccag	1380
gagtccttca ggtttgggga ggagaagacc acccccagcc agaagcagga gccattgac	1440
aaggagctgt accccctggc ctccctgagg tccctgtttg gcaacgaccc ctccctccag	1500

<210> 17
 <211> 39240
 <212> DNA
 <213> Pan troglodytes

<220>
 <221> misc_feature
 <222> (24096)..(24098)
 <223> n iöääñòàâëÿåò ñíáíé a, c, g, èèè t

<400> 17	
catcatcaat aatatacctt attttggatt gaagccaata tgataatgag atgggcggcg	60
cggggcgggg cgcggggcgg gaggcgggtt tgggggcggg ccggcgggcg gggcggtgtg	120
gcggaagtgg actttgtaag tgtggcggat gtgacttgc agtgccgggc gcggtaaaag	180
tgacgttttc cgtgcgcgac aacgcccccg ggaagtgaca tttttccgc ggtttttacc	240
ggatgttgta gtgaatttgg gcgtaaccaa gtaagatttg gccattttcg cgggaaaact	300
gaaacgggga agtgaaatct gattaatttt gcgttagtca taccgcgtaa tatttgtcta	360
gggccgaggg actttggccg attacgtgga ggactcgccc aggtgttttt tgaggatgaat	420
ttccgcgttc cgggtcaaag tctgcgtttt attattatag gatatcccat tgcatacgtt	480

gtatccatat cataatatgt acattttatat tggctcatgt ccaacattac cgccatgttg	540
acattgatta ttgactagtt attaatagta atcaattacg gggtcattag ttcatagccc	600
atatatggag ttccgcgtta cataacttac ggtaaattggc ccgcctggct gaccgcccac	660
cgacccccgc ccattgacgt caataatgac gtatgttccc atagtaacgc caatagggac	720
tttccattga cgtcaatggg tggagtattt acggtaaact gccacttg cagtacatca	780
agtgtatcat atgccaaagta cgtcccttat tgacgtcaat gacggtaaat ggccgcctg	840
gcattatgcc cagtacatga ctttatggga ctttctact tggcagtaca tctacgtatt	900
agtcacgcgt attaccatgg tgatgcggtt ttggcagtac atcaatgggc gtggatagcg	960
gtttgactca cggggatttc caagtctcca cccattgac gtcaatggga gtttgttttg	1020
gcacaaaat caacgggact ttccaaaatg tcgtaacaac tccgccccat tgacgcaaat	1080
gggcggtagg cgtgtacggt gggaggtcta tataagcaga gctctcccta tcagtgatag	1140
agatctccct atcagtgata gagatcgtcg acgagctcgt ttagtgaacc gtcagatcgc	1200
ctggagacgc catccacgct gttttgacct ccatagaaga caccgggacc gatccagcct	1260
ccgcggccgg gaacgggtgca ttggaacgcg gattccccgt gccaaagagt agatctacca	1320
tgggtgctag ggcttctgtg ctgtctggtg gtgagctgga caagtgggag aagatcaggc	1380
tgaggcctgg tggcaagaag aagtacaagc taaagcacat tgtgtggggc tccagggagc	1440
tggagaggtt tgctgtgaac cctggcctgc tggagacctc tgaggggtgc aggcagatcc	1500
tggggccagct ccagccctcc ctgcaaacag gctctgagga gctgagggtc ctgtacaaca	1560
cagtggctac cctgtactgt gtgcaccaga agattgatgt gaaggacacc aaggaggccc	1620
tggagaagat tgaggaggag cagaacaagt ccaagaagaa ggcccagcag gctgctgctg	1680
gcacaggcaa ctccagccag gtgtcccaga actaccccat tgtgcagaac ctccagggcc	1740
agatggtgca ccaggccatc tccccccgga ccctgaatgc ctgggtgaag gtggtggagg	1800
agaaggcctt ctcccctgag gtgatcccca tgttctctgc cctgtctgag ggtgccacc	1860
cccaggacct gaacaccatg ctgaacacag tggggggcca tcaggctgcc atgcagatgc	1920
tgaaggagac catcaatgag gaggctgctg agtgggacag gctgcacccct gtgcacgctg	1980
gccccattgc ccccgccag atgagggagc ccaggggctc tgacattgct ggcaccacct	2040
ccaccctcca ggagcagatt ggctggatga ccaacaaccc ccccatccct gtgggggaaa	2100
tctacaagag gtggatcatc ctgggcctga acaagattgt gaggatgtac tccccacct	2160
ccatcctgga catcaggcag ggccccaaag agcccttcag ggactatgtg gacaggttct	2220
acaagaccct gagggctgag caggcctccc aggaggtgaa gaactggatg acagagaccc	2280
tgctggtgca gaatgccaac cctgactgca agaccatcct gaaggccctg ggccctgctg	2340
ccaccctgga ggagatgatg acagcctgcc aggggggtggg gggccctggt cacaaggcca	2400

gggtgctggc	tgaggccatg	tcccaggtga	ccaactccgc	caccatcatg	atgcagaggg	2460
gcaacttcag	gaaccagagg	aagacagtga	agtgttcaa	ctgtggcaag	gtggggcaca	2520
ttgccaagaa	ctgtagggcc	cccaggaaga	agggtgctg	gaagtgtggc	aaggagggcc	2580
accagatgaa	ggactgcaat	gagaggcagg	ccaacttcct	gggcaaaatc	tggccctccc	2640
acaagggcag	gcctggcaac	ttcctccagt	ccaggcctga	gcccacagcc	cctcccaggg	2700
agtccttcag	gtttggggag	gagaagacca	ccccagcca	gaagcaggag	cccattgaca	2760
aggagctgta	ccccctggcc	tccctgaggt	ccctgtttgg	caacgacccc	tcctcccagt	2820
aaaataaagc	ccgggcagat	ctgatctgct	gtgccttcta	gttgccagcc	atctgttgtt	2880
tgccctccc	ccgtgccttc	cttgaccctg	gaaggtgcca	ctccactgt	cctttcctaa	2940
taaaatgagg	aaattgcatc	gcattgtctg	agtaggtgtc	attctattct	ggggggtggg	3000
gtggggcagc	acagcaaggg	ggaggattgg	gaagacaata	gcaggcatgc	tggggatgcg	3060
gtgggctcta	tggccgatca	gcgatcgctg	aggtgggtga	gtgggcgtgg	cctgggggtg	3120
tcatgaaaat	atataagttg	ggggtcttag	ggtctcttta	tttgtgttgc	agagaccgcc	3180
ggagccatga	gcgggagcag	cagcagcagc	agtagcagca	gcgccttgga	tggcagcatc	3240
gtgagccctt	atttgacgac	gcggatgccc	cactggggccg	gggtgcgta	gaatgtgatg	3300
ggctccagca	tcgacggccg	accgtcctg	ccgcgaaatt	ccgccacgct	gacctatgcg	3360
accgtcgcg	ggacgccgtt	ggacgccacc	gccgccgccg	ccgccaccgc	agccgcctcg	3420
gccgtgcgca	gcctggccac	ggactttgca	ttcctgggac	cactggcgac	aggggctact	3480
tctcgggccg	ctgctgccgc	cgttcgcat	gacaagctga	ccgccctgct	ggcgcagttg	3540
gatgcgtta	ctcggaact	gggtgacctt	tctcagcagg	tcatggccct	gcgccagcag	3600
gtctcctccc	tgcaagctgg	cgggaatgct	tctcccacaa	atgccgttta	agataaataa	3660
aaccagactc	tgtttggatt	aaagaaaagt	agcaagtgca	ttgtctcttt	tatttcataa	3720
ttttccgcgc	gcgataggcc	ctagaccagc	gttctcggtc	gttgaggggtg	cgggtgtatct	3780
tctccaggac	gtggtagagg	tggctctgga	cgttgagata	catgggcatg	agcccgctcc	3840
gggggtggag	gtagcaccac	tgcagagctt	catgctccgg	ggtggtgttg	tagatgatcc	3900
agtcgtagca	ggagcgctgg	gcatggtgcc	taaaaatgtc	cttcagcagc	aggccgatgg	3960
ccagggggag	gcccttggtg	taagtgttta	caaacgggtt	aagttgggaa	gggtgcattc	4020
ggggagagat	gatgtgcatc	ttggactgta	tttttagatt	ggcgatgttt	ccgccagat	4080
cccttctggg	attcatgttg	tgcaggacca	ccagtacagt	gtatccggtg	cacttggggga	4140
atttgtcatg	cagcttagag	ggaaaagcgt	ggaagaactt	ggagacgccc	ttgtggcctc	4200
ccagattttc	catgcattcg	tccatgatga	tggcaatggg	ccgcggggag	gcagcttggg	4260
caaagatatt	tctgggggtcg	ctgacgtcgt	agttgtgttc	caggggtgagg	tcgtcatagg	4320

ccattttttac	aaagcgcggg	cggaggggtgc	ccgactgggg	gatgatgggc	ccctctggcc	4380
ccggggcgta	gttgccctcg	cagatctgca	tttcccaggc	cttaatctcg	gaggggggaa	4440
tcatatccac	ctgcggggcg	atgaagaaaa	cggttttccg	agccggggag	attaactggg	4500
atgagagcag	gtttctaagc	agctgtgatt	ttccacaacc	ggtgggcca	taaataacac	4560
ctataaccgg	ttgcagctgg	tagtttagag	agctgcagct	gccgtcgtcc	cggaggaggg	4620
ggggccacctc	gttgagcatg	tccctgacgc	gcatgttctc	cccgaccaga	tccgccagaa	4680
ggcgctcgcc	gcccagggac	agcagctctt	gcaaggaagc	aaagtttttc	agcggcctga	4740
ggccgtccgc	cgtgggcatg	tttttcaggg	tctggctcag	cagctccagg	cggctcccaga	4800
gctcggtgac	gtgctctacg	gcatctctat	ccagcatatc	tcctcgtttc	gcgggttggg	4860
gcgactttcg	ctgtagggca	ccaagcggtg	gtcgtccagc	ggggccagag	tcattgtcctt	4920
ccatgggcg	agggtcctcg	tcagggtggg	ctgggtcacg	gtgaaggggt	gcgctccggg	4980
ctgagcgctt	gccaaggtgc	gcttgaggct	ggttctgctg	gtgctgaagc	gctgccggtc	5040
ttcgccctgc	gcgtcggcca	ggtagcattt	gaccatgggt	tcatagtcca	gcccctccgc	5100
ggcgtgtccc	ttggcgcgca	gcttgccctt	ggaggtggcg	ccgcacgagg	ggcagagcag	5160
gctcttgagc	gcgtagagct	tggggggcgag	gaagaccgat	tcgggggagt	aggcgtccgc	5220
gccgcagacc	ccgcacacgg	tctcgctcct	caccagccag	gtgagctcgg	ggcgcgccgg	5280
gtcaaaaacc	aggtttcccc	catgcttttt	gatgcgtttc	ttacctcggg	tctccatgag	5340
gtggtgtccc	cgctcgggtga	cgaagaggct	gtccgtgtct	ccgtagaccg	acttgagggg	5400
tcttttctcc	aggggggtcc	ctcggctctt	ctcgtagagg	aactcggacc	actctgagac	5460
gaaggcccg	gtccaggcca	ggacgaagga	ggctatgtgg	gaggggtagc	ggtcgttgtc	5520
cactaggggg	tccaccttct	ccaaggtgtg	aagacacatg	tcgccttcct	cggcgtccag	5580
gaaggtgatt	ggcttgtagg	tgtaggccac	gtgaccgggg	gttcctgacg	gggggggtata	5640
aaaggggggtg	ggggcgcgct	cgctcgtcact	ctcttccgca	tcgctgtctg	cgagggccag	5700
ctgctggggg	gagtattccc	tctcgaaggc	gggcatgacc	tccgcgctga	ggttgctcagt	5760
ttccaaaaac	gaggaggatt	tgatgttcac	ctgtcccag	gtgatacctt	tgagggtacc	5820
cgcgcccatc	tggtcagaaa	acacgatctt	tttattgtcc	agcttggtgg	cgaacgaccc	5880
gtagagggcg	ttggagagca	gcttggcgat	ggagcgcagg	gtctggttct	tgtccctgtc	5940
ggcgcgctcc	ttggcccgca	tgttgagctg	cacgtactcg	cgcgcgacgc	agcgccactc	6000
ggggaagacg	gtggtgcgct	cgctcgggcac	caggcgcacg	cgccagccgc	ggttggtgcag	6060
ggtgaccagg	tccacgctgg	tggcgacctc	gccgcgcagg	cgctcgttgg	tccagcagag	6120
acggccgccc	ttgcgcgagc	agaagggggg	caggggggtcg	agctgggtct	cgtccggggg	6180
gtccgcgtcc	acggtgaaaa	ccccggggcg	caggcgcgcg	tcgaagtagt	ctatcttgca	6240

accttgcattg	tccagcgcct	gctgccagtc	gcgggcggcg	agcgcgcgct	cgtaggggtt	6300
gagcggcggg	ccccagggca	tgggggtgggt	gagtgcggag	gcgtacatgc	cgcagatgtc	6360
atagacgtag	aggggctccc	gcaggacccc	gatgtaggtg	gggtagcagc	ggccgccgcg	6420
gatgctggcg	cgcacgtagt	catacagctc	gtgcgagggg	gcgaggaggt	cggggcccag	6480
gttggtgcgg	gcggggcgct	ccgtgcggaa	gacgatctgc	ctgaagatgg	catgcgagtt	6540
ggaagagatg	gtggggcgct	ggaagacgtt	gaagctggcg	tcctgcaggc	cgacggcgctc	6600
gcgcacgaag	gaggcgtagg	agtcgcgcag	cttgtgtacc	agctcggcgg	tgacctgcac	6660
gtcgagcgcg	cagtagtcga	gggtctcgcg	gatgatgtca	tatttagcct	gccccctctt	6720
tttccacagc	tcgcggttga	ggacaaactc	ttcgcggtct	ttccagtact	cttggtatcg	6780
gaaaccgtcc	ggttccgaac	ggtaagagcc	tagcatgtag	aactggttga	cggcctggta	6840
ggcgcagcag	cccttctcca	cggggagggc	gtaggcctgc	gcggccttgc	ggagcgaggt	6900
gtgggtcagg	gcgaaggtgt	ccctgaccat	gactttgagg	tactggtgct	tgaagtcgga	6960
gtcgtcgcag	ccgccccgct	cccagagcga	gaagtcggtg	cgcttcttgg	agcgggggtt	7020
gggcagagcg	aaggtgacat	cgttgaagag	gattttgccc	gcgcggggca	tgaagttgcg	7080
ggtgatgcgg	aagggccccg	gcacttcaga	gcggttggtg	atgacctggg	cggcgagcac	7140
gatctcgtcg	aagccgttga	tggtgtggcc	cacgatgtag	agttccagga	agcggggccg	7200
gccctttacg	gtgggcagct	tctttagctc	ttcgtaggtg	agctcctcgg	gcgaggcgag	7260
gccgtgctcg	gccagggccc	agtcgcgcag	gtgcgggttg	tctctgagga	aggactccca	7320
gaggtcgcgg	gccaggaggg	tctgcaggcg	gtccctgaag	gtcctgaact	ggcggcccac	7380
ggccattttt	tcgggggtga	tgcaagtaga	ggtgaggggg	tcttgctgcc	agcgggtccca	7440
gtcgagctgc	agggcgaggt	cgcgcgcggc	ggtgaccagg	cgctcgtcgc	ccccgaattt	7500
catgaccagc	atgaagggca	cgagctgctt	tccgaaggcc	cccatccaag	tgtaggtctc	7560
tacatcgtag	gtgacaaaga	ggcgctccgt	gcgaggatgc	gagccgatcg	ggaagaactg	7620
gatctcccg	caccagttgg	aggagtggct	gttgatgtgg	tggaagtaga	agtcccgctg	7680
ccgggcccga	cactcgtgct	ggcttttgta	aaagcgagcg	cagtactggc	agcgtgcac	7740
gggctgtacc	tcatgcacga	gatgcacctt	tcgcccgcgc	acgaggaagc	cgaggggaaa	7800
tctgagcccc	ccgcttggct	cgcgcatgg	ctggtgctct	tctactttgg	atgcgtgtcc	7860
gtctccgtct	ggctcctcga	ggggtgttac	ggtggagcgg	accaccacgc	cgcgcgagcc	7920
gcaggtccag	atatcggcgc	gcggcggtcg	gagtttgatg	acgacatcgc	gcagctggga	7980
gctgtccatg	gtctggagct	cccgcggcgg	cggcaggtca	gccgggagtt	cttgcaggtt	8040
cacctcgcag	agtcggggca	gggcgcgggg	caggtctagg	tggtacctga	tctctagggg	8100
cgtgttggtg	gcggcgctga	tggcttgcag	gagcccgcag	ccccgggggg	cgacgacggt	8160

gccccgcggg	gtggtggtg	tgggtggcggt	gcagctcaga	agcgggtgccg	cgggccccgc	8220
cccggaggta	gggggggctc	cgggtcccgcg	ggcagggggcg	gcagcggcac	gtcggcggtg	8280
agcgcgggca	ggagttggtg	ctgtgcccgcg	aggttgctg	cgaaggcgac	gacgcggcg	8340
ttgatctcct	ggatctggcg	cctctgcgtg	aagacgacg	gcccggtgag	cttgaacctg	8400
aaagagagtt	cgacagaatc	aatctcggtg	tcattgaccg	cggcctggcg	caggatctcc	8460
tgcacgtctc	ccgagttgtc	ttggtagggc	atctcggcca	tgaactgctc	gatctcttcc	8520
tcctggaggt	ctccgcgtcc	ggcgcgttcc	acgggtggccg	ccaggtcggt	ggagatgcgc	8580
cccatgagct	gcgagaaggc	gttgagtccg	ccctcgttcc	agactcggct	gtagaccacg	8640
ccccctggt	catcgcgggc	gcgcatgacc	acctgcgcga	ggttgagctc	cacgtgccgc	8700
gcgaagacgg	cgtagttgcg	cagacgctgg	aagaggtagt	tgagggtggt	ggcgggtgtc	8760
tcggccacga	agaagttcat	gacccagcgg	cgcaacgtgg	attcgttgat	gtcccccaag	8820
gcctccagcc	gttccatggc	ctcgtagaag	tccacggcga	agttgaaaaa	ctgggagttg	8880
cgcgccgaca	cgggtcaactc	ctcctccaga	agacggatga	gctcggcgac	ggtgtcgcgc	8940
acctcgcgt	cgaaggctat	ggggatctct	tcctccgcta	gcataccac	ctcctctct	9000
tcctctctct	ctggcacttc	catgatggct	tcctctctct	cgggggggtg	cggcgggcgc	9060
ggtgggggag	ggggcgctct	gcgccggcgg	cggcgccaccg	ggaggcggtc	cacgaagcgc	9120
gcgatcatct	ccccgcggcg	gcggcgcatg	gtctcggtag	cggcgcgggc	gttctcccgg	9180
gggcgcagtt	ggaagacgcc	gccggacatc	tggtgctggg	gcgggtggcc	gtgaggcagc	9240
gagacggcgc	tgacgatgca	tctcaacaat	tgctgcgtag	gtacgccgcc	gagggacctg	9300
agggagtcca	tatccaccgg	atccgaaaac	ctttcgagga	aggcgtctaa	ccagtcgcag	9360
tcgcaaggta	ggctgagcac	cgtggcgggc	ggcggggggt	ggggggagtg	tctggcgag	9420
gtgctgctga	tgatgtaatt	gaagtaggcg	gacttgacac	ggcggatggt	cgacaggagc	9480
accatgtcct	tgggtccggc	ctgctggatg	cggaggcggt	cggctatgcc	ccaggcttcg	9540
ttctggcatc	ggcgcaggtc	cttgtagtag	tcttgcatga	gcctttccac	cggcacctct	9600
tctccttct	cttctgcttc	ttccatgtct	gcttcggccc	tggggcgggc	ccgcgcccc	9660
ctgcccccca	tgcgcgtagc	cccgaacccc	ctgagcgggt	ggagcagggc	caggtcggcg	9720
acgacgcgt	cggccaggat	ggcctgctgc	acctgcgtga	gggtggtttg	gaagtcatcc	9780
aagtccacga	agcggtggtg	ggcgcccggt	ttgatggtgt	aggtgcagtt	ggccatgacg	9840
gaccagttga	cggctctggtg	gcccggttgc	gacatctcgg	tgtacctgag	tcgcgagtag	9900
gcgcgggagt	cgaagacgta	gtcgttgcaa	gtccgcacca	ggtactggta	gcccaccagg	9960
aagtgcggcg	gcggctggcg	gtagaggggc	cagcgcaggg	tggcgggggc	tccggggggc	10020
aggtcttcca	gcatgaggcg	gtggtagggc	tagatgtacc	tggacatcca	ggtgataccc	10080

gcggcggtgg	tggaggcgcg	cggaagtcg	cgcacccggt	tccagatggt	gcgaggggc	10140
agaaagtgct	ccatggtagg	cgtgctctgt	ccagtcagac	gcgcgagtc	gttgatactc	10200
tagaccagg	aaaacgaaag	ccggtcagcg	ggcactcttc	cgtggtctgg	tgaatagatc	10260
gcaagggtat	catggcgag	ggcctcggtt	cgagccccgg	gtccgggccc	gacggtccgc	10320
catgatccac	gcggttaccg	cccgcgtgtc	gaaccaggt	gtgcgacgtc	agacaacggt	10380
ggagtgttcc	ttttggcggt	tttctggccg	ggcgccggcg	ccgcgtaaga	gactaagccg	10440
cgaaagcgaa	agcagtaagt	ggctcgctcc	ccgtagccgg	agggatcctt	gctaaggggt	10500
gcgttgcggc	gaaccccggt	tcgaatcccg	tactcggggc	ggccggaccc	gcggctaagg	10560
tgttggattg	gcctccccct	cgtataaaga	ccccgcttgc	ggattgactc	cggacacggg	10620
gacgagcccc	ttttatTTTT	gctttcccca	gatgcatccg	gtgttgcgac	agatgcgccc	10680
cccgccccag	cagcagcaac	aacaccagca	agagcggcag	caacagcagc	gggagtcatg	10740
cagggcccc	tcaccacccc	tcggcgggcc	ggccacctcg	gcgtccgcgg	ccgtgtctgg	10800
cgcctgcggc	ggcgcgggcg	gggggcccgc	tgacgacccc	gaggagcccc	cgcggcgcag	10860
ggccagacac	tacctggacc	tggaggaggg	cgagggcctg	gcgcggctgg	gggcgcgctc	10920
tcccgagcgc	caccgcgggg	tgcagctaaa	gcgcgactcg	cgcgaggcgt	acgtgcctcg	10980
gcagaacctg	ttcagggacc	gcgcgggcga	ggagcccag	gagatgcggg	acaggagggt	11040
cagcgcgggg	cgggagctgc	ggcaggggct	gaaccgcgag	cggctgctgc	gcgaggagga	11100
ctttgagccc	gacgcgcgga	cggggatcag	ccccgcgcgc	gcgcacgtgg	cggccgcgca	11160
cctggtgacg	gcgtacgagc	agacggtgaa	ccaggagatc	aacttccaaa	agagtttcaa	11220
caaccacgtg	cgcacgctgg	tggcgcgca	ggaggtgacc	atcgggctga	tgcacctgtg	11280
ggactttgtg	agcgcgctgg	tgcagaaccc	caatagcaag	cctctgacgg	cgcagctggt	11340
cctgatagtg	cagcacagca	gggacaacga	ggcgtttagg	gacgcgctgc	tgaacatcac	11400
cgagccccgag	ggccggtggc	tgctggacct	gattaacatc	ctgcagagca	tagtggtgca	11460
ggagcgcagc	ctgagcctgg	ccgacaaggt	ggcgccatc	aactactcga	tgctgagcct	11520
gggcaagttt	tacgcgcgca	agatctacca	gacgccgtac	gtgcccatag	acaaggagggt	11580
gaagatcgac	ggtttttaca	tgcgcatggc	gctgaagggtg	ctcaccctaa	gcgacgacct	11640
gggcgtgtac	cgcaacgagc	gcatccacaa	ggccgtgagc	gtgagccggc	ggcgcgagct	11700
gagcgaccgc	gagctgatgc	atagcctgca	gcgggcgctg	gcgggcgccc	gcagcggcga	11760
cagggaggcg	gagtcctact	tcgatgcggg	ggcggacctg	cgtggggcgc	ccagccggcg	11820
ggccctggag	gccgcggggg	tccgcgagga	ctatgacgag	gacggcgagg	aggatgagga	11880
gtacgagcta	gaggagggcg	agtacctgga	ctaaaccgcg	ggtggtgttt	ccggtagatg	11940
caagacccga	acgtggtgga	cccggcgctg	cgggcggtc	tgagagcca	gccgtccggc	12000

cttaactcct	cagacgactg	gcgacaggtc	atggaccgca	tcatgtcgct	gacggcgct	12060
aaccgagcg	cgttcggca	gcagccgag	gccaacaggc	tctccgcat	cctggaggcg	12120
gtggtgcctg	cgcgctcgaa	ccccacgcac	gagaaggtgc	tggccatagt	gaacgcgctg	12180
gccgagaaca	gggccatccg	cccggacgag	gccgggctgg	tgtacgacgc	gctgctgcag	12240
cgcgaggccc	gctacaacag	cggcaacgtg	cagaccaacc	tggaccggct	ggtgggggac	12300
gtgcgagagg	cgggtggcgca	gcgcgagcgc	gcggatcggc	agggcaacct	gggctccatg	12360
gtggcgctga	atgccttcct	gagcacgcag	ccggccaacg	tgcccggggg	gcaggaagac	12420
tacaccaact	ttgtgagcgc	gctgcggctg	atggtgaccg	agacccccca	gagcgagggtg	12480
taccagtcgg	gtccggacta	cttcttccag	accagcagac	agggcctgca	gacggtgaac	12540
ctgagccagg	ctttcaagaa	cctgcggggg	ctgtggggcg	tgaaggcgcc	caccggcgac	12600
cgggcgacgg	tgtccagcct	gctgacgccc	aactcgcgcc	tgctgctgct	gctgatcgcg	12660
ccgttcacgg	acagcggcag	cgtgtcccgg	gacacctacc	tggggcacct	gctgaccctg	12720
taccgcgagg	ccatcgggca	ggcgaggtg	gacgagcaca	ccttccagga	gatcaccagc	12780
gttagccgcg	cgctggggca	ggaggacacg	agcagcctgg	agggcactct	gaactacctg	12840
ctgaccaacc	ggcggcgaaa	gattccctcg	ctgcacagcc	tgacctccga	ggaggagcgc	12900
atcttgcgct	acgtgcagca	gagcgtgagc	ctgaacctga	tgcgcgacgg	ggtgacgccc	12960
agtgtggcgc	tggacatgac	cgcgcgcaac	atggaaccgg	gcatgtacgc	cgcgcacccg	13020
ccttacatca	accgcctgat	ggactacctg	catcgcgcg	cggccgtgaa	ccccgagtac	13080
tttaccaacg	ccatcctgaa	cccgactgg	ctccgcgcgc	cggggttcta	cagcgggggc	13140
ttcgagggtc	cggaggccaa	cgatggcttc	ctgtgggacg	acatggacga	cagcgtgttc	13200
tccccgcggc	cgcaggcgct	ggcggaagcg	tccctgctgc	gtcccaagaa	ggaggaggag	13260
gcgagtcgcc	gccgcggcag	cagcggcgctg	gcttctctgt	ccgagctggg	ggcggcagcc	13320
gccgcgcgcc	ccgggtccct	gggcggcagc	ccctttccga	gcctggtggg	gtctctgcac	13380
agcgagcgca	ccaccgccc	tcggctgctg	ggcgaggacg	agtacctgaa	taactccctg	13440
ctgcagccgg	tgcgggagaa	aaacctgcct	ccgccttcc	ccaacaacgg	gatagagagc	13500
ctggtggaca	agatgagcag	atggaagacc	tatgcgcagg	agcacaggga	cgcgcccgcg	13560
ctccggccgc	ccacgcggcg	ccagcgccac	gaccggcagc	gggggctggg	gtgggatgac	13620
gaggactccg	cggacgatag	cagcgtgctg	gacctgggag	ggagcggcaa	cccgttcgcg	13680
cacctgcgcc	ccgcctggg	gaggatgttt	taaaaaaaaa	aaaaagcaag	aagcatgatg	13740
caaaaattaa	ataaaactca	ccaaggccat	ggcgaccgag	cgttggtttc	ttgtgttccc	13800
ttcagtatgc	ggcgcgggc	gatgtaccag	gagggacctc	ctccctctta	cgagagcgctg	13860
gtgggcgcgg	cggcgggggc	gccctcttct	ccctttgcgt	cgcagctgct	ggagccgccc	13920

tacgtgcctc	cgcgctacct	gcggcctacg	ggggggagaa	acagcatccg	ttactcggag	13980
ctggcgcccc	tgttcgacac	cacccgggtg	tacctggtgg	acaacaagtc	ggcggacgtg	14040
gcctccctga	actaccagaa	cgaccacagc	aattttttga	ccacggtcac	ccagaacaat	14100
gactacagcc	cgagcgaggc	cagcacccag	accatcaatc	tggatgaccg	gtcgcactgg	14160
ggcggcgacc	tgaaaaccaa	cctgcacacc	aacatgcccc	acgtgaacga	gttcatgttc	14220
accaataagt	tcaaggcgcg	ggtgatggtg	tcgcgctcgc	acaccaagga	agaccgggtg	14280
gagctgaagt	acgagtgggt	ggagttcgag	ctgccagagg	gcaactactc	cgagaccatg	14340
accattgacc	tgatgaacaa	cgcgatcgtg	gagcactatc	tgaaagtggg	caggcagaa	14400
ggggtcctgg	agagcgacac	cgggggtcaag	ttcgacacca	ggaacttccg	cctggggctg	14460
gaccccgatg	ccgggctggt	tatgcccggg	gtgtacacca	acgaggcctt	ccatcccagc	14520
atcatcctgc	tgcccggctg	cgggggtggac	ttcacttaca	gccgcctgag	caacctcctg	14580
ggcatccgca	agcggcgacc	cttccaggag	ggcttcagga	tcacctacga	ggacctggag	14640
gggggcaaca	tccccgcgct	cctcgatgtg	gaggcctacc	aggatagctt	gaaggaaaat	14700
gaggcgggac	aggaggatac	cgcccccgcc	gcctccgccc	ccgccgagca	gggcgaggat	14760
gctgctgaca	ccgcggccgc	ggacggggcg	gaggccgacc	ccgctatggt	ggtggaggct	14820
gccgagcagg	aggaggacac	gaatgacagt	gcggtgcgcg	gagacacctt	cgtcaccgag	14880
ggggaggaaa	agcaagcgga	ggccgaggcc	gcggccgagg	aaaagcaact	ggcggcgaca	14940
gcggcgggcg	cggcgttggc	cgcgggcgag	gctgagtctg	aggggaccaa	gcccgcgaag	15000
gagcccgatg	ttaagcccct	gaccgaagat	agcaagaagc	gcagttacaa	cctgctcaag	15060
gacagcacca	acaccgcgta	ccgcagctgg	tacctggcct	acaactacgg	cgacccgctc	15120
acgggggtgc	gctcctggac	cctgctgtgc	acgccggacg	tgacctgcgg	ctcggagcag	15180
gtgtactggt	cgctgcccga	catgatgcaa	gaccccgatg	ccttccgctc	cacgcggcag	15240
gtcagcaact	tcccggtggt	gggcgccgag	ctgctgcccg	tgcactccaa	gagcttctac	15300
aacgaccagg	ccgtctactc	ccagctcatc	cgccagttca	cctctctgac	ccacgtgttc	15360
aatcgctttc	ctgagaacca	gattctggcg	cgcccggccc	ccccaccat	caccaccgtc	15420
agtgaaaacg	ttcctgctct	cacagatcac	gggacgctac	cgctgcgcaa	cagcatcgga	15480
ggagtccagc	gagtgaccgt	tactgacgcc	agacgccgca	cctgccccta	cgttttacaag	15540
gccttgggca	tagtctcgcc	gcgcgtcctt	tccagccgca	ctttttgagc	aacaccacca	15600
tcatgtccat	cctgatctca	cccagcaata	actccggctg	gggactgctg	cgcgcgccca	15660
gcaagatgtt	cggagggggc	aggaagcggt	ccgagcagca	ccccgtgcgc	gtgcgcgggc	15720
acttccgcgc	cccctgggga	gcgcacaaac	gcggccgcgc	ggggcgacc	accgtggagc	15780
acgccatcga	ctcgggtggt	gagcaggcgc	gcaactacag	gcccggggtc	tctaccgtgg	15840

acgcggccat	ccagaccgtg	gtgcggggcg	cgcggcggta	cgccaagctg	aagagccgcc	15900
ggaagcgcgt	ggcccgcgc	caccgccgcc	gacccggggc	cgccgccaaa	cgcgccgccg	15960
cggccctgct	tcgccgggccc	aagcgcacgg	gccgccgcgc	cgccatgagg	gccgcgcgcc	16020
gcttgggcgc	cggcatcacc	gccgccacca	tggccccccg	taccgaaga	cgcgcgccg	16080
ccgccgcgc	cgccgccatc	agtgacatgg	ccagcaggcg	ccggggcaac	gtgtactggg	16140
tgcgcgactc	ggtgaccggc	acgcgcgtgc	ccgtgcgctt	cgcccccccg	cggacttgag	16200
atgatgtgaa	aaaacaacac	tgagtctcct	gctgttgtgt	gtatcccagc	ggcggcgggc	16260
gcgcgcgcag	cgtcatgtcc	aagcgcaaaa	tcaaagaaga	gatgctccag	gtcgtcgcgc	16320
cggagatcta	tgggcccccg	aagaaggaag	agcaggattc	gaagccccgc	aagataaagc	16380
gggtcaaaaa	gaaaaagaaa	gatgatggcg	atgccgatgg	ggaggtggag	ttcctgcgcg	16440
ccacggcgcc	caggcgcccc	gtgcagtgga	agggccggcg	cgtaaagcgc	gtcctgcgcc	16500
ccggcaccgc	ggtggtcttc	acgcccggcg	agcgtccac	ccggactttc	aagcgcgtct	16560
atgacgaggt	gtacggcgac	gaagacctgc	tggagcaggc	caacgagcgc	ttcggagagt	16620
ttgcttacgg	gaagcgtcag	cggccgctgg	ggaaggagga	cctgctggcg	ctgccgctgg	16680
accagggcaa	ccccaccccc	agtctgaagc	ccgtgaccct	gcagcagggtg	ctgccgagca	16740
gcgcaccctc	cgaggcgaag	cggggctctga	agcgcgaggg	cggcgacctg	gcgcccaccg	16800
tgcagctcat	ggtgcccagg	cggcagaggc	tggaggatgt	gctggagaaa	atgaaagtag	16860
accccggtct	gcagccggac	atcaggggtcc	gtcccatcaa	gcaggtggcg	ccgggcctcg	16920
gcgtgcagac	cgtggacgtg	gtcatcccca	ccggcaactc	ccccgcgcgc	accaccacta	16980
ccgtgcctc	cacggacatg	gagacacaga	ccgatcccg	cgcagccgcc	gccaccgccg	17040
ccgccgcgac	ctcctcggcg	gaggtgcaga	cggaccctg	gctgccgcgc	gcgatgtcag	17100
ctccccgcgc	gcgtcgcggg	cgcaggaagt	acggcgccgc	caacgcgctc	ctgcccaggt	17160
acgccttgca	tccttccatc	gcgcccaccc	ccggctaccg	aggctatacc	taccgcccgc	17220
gaagagccaa	gggttccacc	cgccgtcccc	gccgacgcgc	cgccgccacc	acccgcgcgc	17280
gccgccgcag	acgccagccc	gactggctc	cagtctccgt	gaggagagtg	gcgcgcgacg	17340
gacacaccct	ggtgctgccc	agggcgcgct	accaccccag	catcgtttaa	aagcctgttg	17400
tggttcttgc	agatatggcc	ctcacttgcc	gcctccgttt	cccggtgccg	ggataccgag	17460
gaggaagatc	gcgccgcagg	aggggtcttg	ccggccgcgg	cctgagcggg	ggcagccgcc	17520
gcgcgcaccg	gcggcgacgc	gccaccagcc	gacgcatgcg	cgggggggtg	ctgcccctgt	17580
taatccccct	gatcgccgcg	gcgatcggcg	ccgtgcccgg	gatcgccctc	gtggccttgc	17640
aggcgtccca	gaggcattga	cagacttgca	aacttgcaaa	tatggaaaaa	aacccaata	17700
aaaaagtcta	gactctcacg	ctcgcttggt	cctgtgacta	ttttgtagaa	tggaagacat	17760

caacttttgcg	tcgctggccc	cgcgtcacgg	ctcgcgcccc	ttcctgggac	actggaacga	17820
tatcggcacc	agcaacatga	gcggtggcgc	cttcagttgg	ggctctctgt	ggagcggcat	17880
taaaagtatc	gggtctgccg	ttaaaaatta	cggctccccg	gcctggaaca	gcagcacggg	17940
ccagatgttg	agagacaagt	tgaaagagca	gaacttccag	cagaaggtgg	tggagggcct	18000
ggcctccggc	atcaacgggg	tggtggacct	ggccaaccag	gccgtgcaga	ataagatcaa	18060
cagcagactg	gacccccggc	cgccggtgga	ggaggtgccg	cgggcgctgg	agacggtgtc	18120
ccccgatggg	cgtggcgaga	agcgcccgcg	gcccgatagg	gaagagacca	ctctgggtcac	18180
gcagaccgat	gagccgcccc	cgtatgagga	ggccctgaag	caaggtctgc	ccaccacgcg	18240
gcccacgcg	cccatggcca	cgggggtggt	gggccgccac	acccccgcca	cgctggactt	18300
gcctccgccc	gccgatgtgc	cgcagcagca	gcagaaggcg	gcacagccgg	gcccgcgccg	18360
gaccgcctcc	cgttctctcg	cgggtcctct	gcgccgcgcg	gccagcggcc	cccgcggggg	18420
ggtcgcgagg	cacggcaact	ggcagagcac	gctgaacagc	atcgtgggtc	tgggggtgcg	18480
gtccgtgaag	cgccgccgat	gctactgaat	agcttagcta	acgtgttgta	tgtgtgtatg	18540
cgcctatgt	cgccgccaga	ggagctgctg	agtcgccgcc	gttcgcgcgc	ccaccaccac	18600
caccgccact	ccgcccctca	agatggcgac	cccacgatg	atgccgcagt	ggtcgtacat	18660
gcacatctcg	ggccaggacg	cctcggagta	cctgagcccc	gggctggtgc	agttcgcccc	18720
cgccaccgag	agctacttca	gcctgagtaa	caagtttagg	aaccccacgg	tggcgcccac	18780
gcacgatgtg	accaccgacc	ggtctcagcg	cctgacgctg	cggttcattc	ccgtggaccg	18840
cgaggacacc	gcgtactcgt	acaaggcgcg	gttcaccctg	gccgtgggcg	acaaccgcgt	18900
gctggacatg	gcctccacct	actttgacat	ccgcgggggtg	ctggaccggg	gccccacttt	18960
caagccttac	tctggcaccg	cctacaactc	cctggccccc	aaggcgctc	ccaactcctg	19020
cgagtgggag	caattagaag	aagcccaggc	cgctgtggaa	gacgaagaat	tagaagatga	19080
agacaggaa	ccacaggatg	aggcacctgt	gaaaaaaacc	catgtatacg	ctcaggctcc	19140
cctttctgga	gaagaaatta	ctaaaaacgg	tttgcaaata	gggtcagata	acacagaagc	19200
ccagtctaag	cccatatatg	cagatcctac	attccagcct	gaaccccaaa	tcgggggaatc	19260
ccagtggaat	gaggcagatg	ctacagttgc	cggcggtaga	gtgctaaaga	aatccactcc	19320
catgaagcca	tgctatggtt	cctatgcaag	accacaaaac	tccaatggag	gtcaagggtg	19380
gctggtggct	gatgataagg	gggttcttca	atctaaagtt	gaattgcaat	ttttttcaaa	19440
tactactact	cttaatcagc	gggagggtaa	cgatacaaaa	ccaaaagtgg	tgctgtatag	19500
cgaagatgtg	cacatggaaa	ctccagacac	ccacatttct	tacaagccca	caaaaagcga	19560
tgacaattca	aaaatcatgc	tgggtcagca	gtccatgccc	aacagacctc	attacatcgg	19620
cttcagagac	aactttatcg	gcctcatgta	ttacaatagc	actggcaaca	tgggagtgct	19680

tgcaggtcag	gcctctcagt	tgaatgcagt	ggcggacttg	caagacagaa	acacagaact	19740
gtcctaccag	ctcttgcttg	attccatggg	tgacagaacc	agatactttt	ccatgtggaa	19800
tcaggcagtg	gacagttatg	accagatgt	cagaattatt	gaaaatcatg	gaactgaaga	19860
cgagctcccc	aactattgtt	tccctctggg	cggcataggg	gtaactgaca	cttaccaggc	19920
cattaaaacc	aatggcaatg	gtcaagaaaa	cccaacctgg	gaaaaagata	cagagtttgc	19980
agaccgcaat	gaaatagggg	tgggaaacaa	tttcgctatg	gagatcaacc	tcagtgccaa	20040
cctgtggaga	aacttcctgt	actccaacgt	ggcgtgttac	ctgccagaca	agcttaagta	20100
caaccctcc	aatgtggaca	tctctgacaa	cccaacacc	tacgattaca	tgaacaagcg	20160
agtgggtggc	ccggggctgg	tggactgcta	catcaacctg	ggcgcgcgct	ggcgtgtgga	20220
ctacatggac	aacgtcaacc	ccttcaacca	ccaccgcaat	gcgggcctgc	gctaccgctc	20280
catgctcctg	ggcaacgggc	gctacgtgcc	cttccacatc	caggtgcccc	agaagttctt	20340
tgccatcaag	aacctcctcc	tcctgccggg	ctctacacc	tacgagtgga	acttcaggaa	20400
ggatgtcaac	atggctcctcc	agagctctct	gggtaacgat	ctcaggggtg	acggggccag	20460
catcaagttc	gagagcatct	gcctctacgc	caccttcttc	cccatggccc	acaacacggc	20520
ctccacgctc	gaggccatgc	tcaggaacga	caccaacgac	cagtccttca	atgactacct	20580
ctccgccgcc	aacatgctct	acccataacc	cgccaacgcc	accaacgtcc	ccatctccat	20640
cccctcgcg	aactgggcgg	ccttcgcggg	ctgggccttc	acccgcctca	agaccaagga	20700
gacccctcc	ctgggctcgg	gattcgaccc	ctactacacc	tactcgggct	ccattcccta	20760
cctggacggc	accttctacc	tcaaccacac	tttcaagaag	gtctcgggtc	ccttcgactc	20820
ctcggtcagc	tggccgggca	acgaccgtct	gtcaccccc	aacgagttcg	aatcaagcg	20880
ctcggtcgac	ggggagggct	acaacgtggc	ccagtgaac	atgaccaagg	actggttctt	20940
ggtccagatg	ctggccaact	acaacatcgg	ctaccagggc	ttctacatcc	cagagagcta	21000
caaggacagg	atgtactcct	tcttcaggaa	cttcagccc	atgagccggc	aggtgggtgga	21060
ccagaccaag	tacaaggact	accaggaggt	gggcatcatc	caccagcaca	acaactcggg	21120
cttcgtgggc	tacctcgccc	ccaccatgcg	cgagggacag	gcctaccccg	ccaacttccc	21180
ctacccgctc	ataggcaaga	ccgcggtcga	cagcatcacc	cagaaaaagt	tcctctgcga	21240
tcgcacctc	tggcgcatcc	ccttctccag	caacttcatg	tccatgggtg	cgctctcgga	21300
cctgggccag	aacttgctct	acgccaactc	cgcccacgcc	ctcgacatga	ccttcgaggt	21360
cgaccccatg	gacgagccca	cccttctcta	tgttctgttc	gaagtctttg	acgtgggtccg	21420
ggtccaccag	ccgcaccgcg	gcgtcatcga	gaccgtgtac	ctgcgtacgc	ccttctcggc	21480
cggcaacgcc	accacctaaa	gaagcaagcc	gcagtcatcg	ccgcctgcat	gccgtcgggt	21540
tccaccgagc	aagagctcag	ggccatcgtc	agagacctgg	gatgcggggc	ctatTTTTTg	21600

ggcacttttcg	acaagcgctt	ccctggcctt	gtctccccac	acaagctggc	ctgcgccatc	21660
gtcaacacgg	ccggccgcga	gaccgggggc	gtgactggc	tggccttcgc	ctggaacccg	21720
cgctccaaaa	catgcttcct	ctttgacccc	ttcggctttt	cggaccagcg	gctcaagcaa	21780
atctacgagt	tcgagtacga	gggcttgctg	cgtcgcagcg	ccatcgccctc	ctcgcccgcac	21840
cgctgcgtca	ccctcgaaaa	gtccacccag	accgtgcagg	ggcccgcactc	ggccgcctgc	21900
ggtctctttct	gctgcatggt	tctgcacgcc	tttgtgact	ggcctcagag	tcccatggac	21960
cgcaacccca	ccatgaactt	gctgacgggg	gtgcccaact	ccatgctcca	gagccccag	22020
gtcgagccca	ccctgcgccg	caaccaggag	cagctctaca	gcttcctgga	gcgccactcg	22080
ccctacttcc	gccgccacag	cgcacagatc	aggagggcca	cctccttctg	ccacttgcaa	22140
gagatgcaag	aagggtaata	acgatgtaca	cacttttttc	tcaataaatg	gcattttttt	22200
atttatacaa	gctctctggg	gtattcattt	cccaccacca	ccaccaccgc	ccgttgctgc	22260
catctggctc	tatttagaaa	tcgaaaggg	tctgccggga	gtcgccgtgc	gccacgggca	22320
gggacacgtt	gcgatactgg	tagcgggtgc	cccacttgaa	ctcgggcacc	accaggcgag	22380
gcagctcggg	gaagttttcg	ctccacaggc	tgcgggtcag	caccagcgcg	ttcatcaggt	22440
cgggcgccga	gatcttgaag	tcgcagttgg	ggccgcgcgc	ctgcgcgcgc	gagttgcggt	22500
acaccgggtt	gcagcactgg	aacaccaaca	gcgcggggtg	cttcacgcta	gccagcacgc	22560
tgcggtcgga	gatcagctcg	gcgtccaggt	cctccgcgtt	gtcagcgcg	aacggggcca	22620
tcttgggcac	ttgcctcccc	aggaagggcg	cgtgccccgg	tttcgagttg	cagtcgcagc	22680
gcagcgggat	cagcaggtgc	ccatgcccgg	actcggcggt	ggggtacagc	gcgcgcatga	22740
aggcctgcat	ctggcggaag	gccatctggg	ccttggcgcc	ctccgagaag	aacatgccgc	22800
aggacttgcc	cgagaactgg	tttgcggggc	agctggcgctc	gtgcaggcag	cagcgcgcg	22860
cgggtgttgg	gatctgcacc	acgttgcgcc	cccaccgggt	cttcacgata	ttggccttgg	22920
acgattgctc	cttcagcgcg	cgctgcccgt	tctcgctggt	cacatccatc	tcgatcacat	22980
gttccttggt	caccatgctg	ctgccgtgca	ggcacttcag	ctcgccctcc	gtctcggtgc	23040
agcgggtgctg	ccacagcgcg	cagcccgtgg	gctcgaaaga	cttgtaggtc	acctccgcga	23100
aggactgcag	gtacccctgc	aaaaagcggc	ccatcatggt	cacgaaggtc	ttgttgctgc	23160
tgaaggtcag	ctgcagcccc	cgggtgctcct	cgttcagcca	ggtcttgcac	acggccgcca	23220
gcgcctccac	ctggtcgggc	agcatcttga	agttcacctt	cagctcattc	tccacgtggt	23280
acttgctcat	cagcgtgcgc	gccgcctcca	tgcccttctc	ccaggccgac	accagcggca	23340
ggctcacggg	gttcttcacc	atcacctggg	ccgccgcctc	cgccgcgctt	tcgctttccg	23400
ccccgctggt	ctcttctctt	tcctcctctt	cctcgccgcc	gcccactcgc	agccccgcga	23460
ccacgggggtc	gtcttctctg	aggcgctgca	ccttgcgctt	gccgttgcg	ccctgcttga	23520

tgcgcacggg	cgggttgctg	aagcccacca	tcaccagcgc	ggcctcttct	tgctcgtcct	23580
cgctgtccag	aatgacctcc	ggggaggggg	ggttggtcat	cctcagtacc	gaggcacgct	23640
tctttttctt	cctgggggcg	ttcgccagct	ccgcggctgc	ggccgctgcc	gaggtcgaag	23700
gccgagggct	gggcgtgcgc	ggcaccagcg	cgtcctgcga	gccgtcctcg	tcctcctcgg	23760
actcgagacg	gaggcggggc	cgcttcttct	ggggcgcgcg	gggcggcgga	ggcggcggcg	23820
gcgacggaga	cggggacgag	acatcggtcca	gggtgggtgg	acggcggggc	gcgccgcgtc	23880
cgcgctcggg	ggtggtctcg	cgctggtcct	cttcccgact	ggccatctcc	cactgctcct	23940
tctcctatag	gcagaaagag	atcatggagt	ctctcatgcg	agtcgagaag	gaggaggaca	24000
gcctaaccgc	cccctctgag	ccctccacca	ccgccgccac	caccgccaat	gccgccgcgg	24060
acgacgcgcc	caccgagacc	accgccagta	ccaccnnnct	ccccagcgac	gcacccccgc	24120
tcgagaatga	agtgtgatc	gagcaggacc	cgggttttgt	gagcggagag	gaggatgagg	24180
tggtatgagaa	ggagaaggag	gaggtcgccg	cctcagtgcc	aaaagaggat	aaaaagcaag	24240
accaggacga	cgcagataag	gatgagacag	cagtcggggc	ggggaacgga	agccatgatg	24300
ctgatgacgg	ctacctagac	gtgggagacg	acgtgctgct	taagcacctg	caccgccagt	24360
gcgtcatcgt	ctgcgacgcg	ctgcaggagc	ggtgcgaagt	gcccctggac	gtggcgaggg	24420
tcagccgcgc	ctacgagcgg	cacctcttct	cgccgcacgt	gccccccaag	cgccgggaga	24480
acggcacctg	cgagcccaac	ccgcgtctca	acttctaccc	ggtcttcgcg	gtacccgagg	24540
tgctggccac	ctaccacatc	tttttccaaa	actgcaagat	ccccctctcc	tgccgcgcta	24600
accgcacccg	cgccgacaaa	accctgaccc	tgccggcagg	cgccacata	cctgatattg	24660
cctctctgga	ggaagtgcc	aagatcttct	agggctctcg	tcgcgacgag	aaacgggcgg	24720
cgaacgctct	gcacggagac	agcgaaaacg	agagtcactc	gggggtgctg	gtggagctcg	24780
agggcgacaa	cgcgcgcctg	gccgtactca	agcgcagcat	agaggtcacc	cactttgcct	24840
acccggcgct	caacctgccc	ccaagggtca	tgagtgtggt	catgggcgag	ctcatcatgc	24900
gccgcgcca	gcccctggcc	gcggatgcaa	acttgcaaga	gtcctcagag	gaaggcctgc	24960
ccgcggtcag	cgacgagcag	ctggcgcgct	ggctggagac	ccgcgacccc	gcgcagctgg	25020
aggagcggcg	caagctcatg	atggccgcgg	tgctggtcac	cgtggagctc	gagtgtctgc	25080
agcgcttctt	cgcggacccc	gagatgcagc	gcaagctcga	ggagaccctg	cactacacct	25140
tccgccaggg	ctacgtgcgc	caggcctgca	agatctccaa	cgtggagctc	tgcaacctgg	25200
tctcctacct	gggcatcctg	cacgagaacc	gcctcgggca	gaacgtcctg	cactccaccc	25260
tcaaagggga	ggcgcgcgcg	gactacatcc	gcgactgcgc	ctacctcttc	ctctgctaca	25320
cctggcagac	ggccatgggg	gtctggcagc	agtgccctgga	ggagcgcaac	ctcaaggagc	25380
tggaaaagct	cctcaagcgc	accctcaggg	acctctggac	gggcttcaac	gagcgctcgg	25440

tggccgcccgc	gctggcgggac	atcatctttcc	ccgagcgcct	gctcaagacc	ctgcagcagg	25500
gcctgccccga	cttcaccagc	cagagcatgc	tgcagaactt	caggactttc	atcctggagc	25560
gctcgggcat	cctgccggcc	acttgctgcg	cgctgcccag	cgacttcgtg	cccatcaagt	25620
acagggagtg	cccgccggcc	ctctggggcc	actgctacct	cttcagctg	gccaaactacc	25680
tcgcctacca	ctcggacctc	atggaagacg	tgagcggcga	gggcctgctc	gagtgccact	25740
gccgctgcaa	cctctgcacg	ccccaccgct	ctctagtctg	caaccgcgag	ctgctcagcg	25800
agagtcatat	tatcgggtacc	ttcgagctgc	aggggtccctc	gcctgacgag	aagtccgcgg	25860
ctccggggct	gaaactcact	ccggggctgt	ggacttccgc	ctacctacgc	aaatttgtac	25920
ctgaggacta	ccacgcccac	gagatcaggt	tctacgaaga	ccaatcccgc	ccgcccgaagg	25980
cggagctcac	cgctgcgtc	atcaccagc	ggcacatcct	gggccaattg	caagccatca	26040
acaaagcccc	ccgagagttc	ttgctgaaaa	agggtcgggg	ggtgtacctg	gacccccagt	26100
ccggcgagga	gctaaacccg	ctacccccgc	cgccgcccga	gcagcgggac	cttgcttccc	26160
aggatggcac	ccagaaagaa	gcagcagccg	ccgccgcagc	catacatgct	tctggaggaa	26220
gaggaggagg	actgggacag	tcaggcagag	gagatgatgg	aagactggga	ggaggacagc	26280
agcctagacg	aggaagcttc	agaggccgaa	gaggtggcag	acgcaacacc	atcacctcgc	26340
gtcgcagccc	cctcgccggg	gcccctgaaa	tcttcggaac	ccagcaccag	cgctataacc	26400
tccgctcctc	cggcgcgggc	gccacccgcc	cgcagaccca	accgtagatg	ggacaccaca	26460
ggaaccgggg	tcggtaagtc	caagtgcccg	ccgccgccac	cgcagcagca	gcagcagcag	26520
cgccagggct	accgctcgtg	gcgcggggcac	aagaacgcca	tagtcgcctg	cttgcaagac	26580
tgccgggggca	acatctcttt	cgcccgccgc	ttctgctat	tccaccacgg	ggtcgccttt	26640
ccccgcaatg	tcttgcatca	ctaccgtcat	ctctacagcc	cctactgcag	cggcgaccca	26700
gaggcggcag	cggcagccac	agcggcgacc	accacctagg	aagatatact	ccgcggggcaa	26760
gacagcggca	gcagcggcca	ggagaccgcg	ggcagcagcg	gcgggagcgg	tgggcgcact	26820
gcgcctctcg	cccaacgaac	ccctctcgac	ccgggagctc	agacacagga	tcttccccac	26880
tttgatgccc	atcttccaac	agagcagagg	ccaggagcag	gagctgaaaa	taaaaaacag	26940
atctctgcgc	tccctcacc	gcagctgtct	gtatcacaaa	agcgaagatc	agcttcggcg	27000
cacgctggag	gacgcggagg	cactcttcag	caaatactgc	gcgctcactc	ttaaagacta	27060
gctccgcgcc	cttctcgaat	ttaggcggga	gaaaactacg	tcctcgccgg	ccgccgcccga	27120
gcccgccag	ccgagatgag	caaagagatt	cccacgcat	acatgtggag	ctaccagccg	27180
cagatgggac	tcgcggcggg	agcggcccag	gactactcca	cccgcataga	ctacatgagc	27240
gcgggacccc	acatgatctc	acaggtcaac	gggatccgcg	cccagcgaaa	ccaaatactg	27300
ctggaacagg	cggccatcac	cgccacgccc	cgccataatc	tcaacccccg	aaattggccc	27360

gccgccctag	tgtaccagga	aacccccctcc	gccaccaccg	tactacttcc	gcgtgacgcc	27420
caggccgaag	tccagatgac	taactcaggg	gcgcagctcg	cgggcgggctt	tcgtcacggg	27480
gcgcggccgc	tccgaccagg	tataagacac	ctgatgatca	gaggccgagg	tatccagctc	27540
aacgacgagt	cggtgagctc	ttcgctcggg	ctccgtccgg	acggaacttt	ccagctcgcc	27600
ggatccggtc	gctcttcggt	cacgccccgc	caggcgtacc	tgactctgca	gacctcgtec	27660
tcggagcccc	gctccggcgg	catcggaacc	ctccagttcg	tggaggagtt	cgtgccctcg	27720
gtctacttca	accctttctc	gggacctccc	ggacgctacc	ccgaccagtt	cattccgaac	27780
tttgacgcgg	tgaaggactc	ggcggacggc	tacgactgaa	tgtcaggtgc	cgaggcagag	27840
cagcttcgcc	tgagacacct	cgagcactgc	cgccgccaca	agtgcctcgc	ccgcgggtcc	27900
ggtgagttct	gctactttca	gctacccgag	gagcataccg	aggggccggc	gcacggcgtc	27960
cgctgacca	cccagggcga	ggttacctgt	tcctctatcc	gggagttcac	cctccgtccc	28020
ctgctagtgg	agcgggagcg	gggtccctgt	gtcctaacta	tcgcctgcaa	ctgccctaac	28080
cctggattac	atcaagatct	ttgctgtcat	ctctgtgctg	agtttaataa	acgctgagat	28140
cagaatctac	tggggctcct	gtcgccatcc	tgtgaacgcc	accgtcttca	cccacccga	28200
ccaggcccag	gcgaacctca	cctgcgggtct	gcatcggagg	tccaagaagt	acctcacctg	28260
gtacttcaac	ggcacccccct	ttgtggttta	caacagcttc	gacggggacg	gagtctccct	28320
gaaagaccag	ctctccggtc	tcagctactc	catccacaag	aacaccaccc	tccaactctt	28380
ccctccctac	ctgccgggaa	cctacgagtg	cgtcaccggc	cgctgcaccc	acctcacccg	28440
cctgatcgta	aaccagagct	ttccgggaac	agataactcc	ctcttcccca	gaacaggagg	28500
tgagctcagg	aaactccccg	gggaccaggg	cggagacgta	ccttcgaccc	ttgtgggggt	28560
aggatttttt	attaccgggt	tgctggctct	tttaatcaaa	gcttccttga	gatttggtct	28620
ttccttctac	gtgtatgaac	acctcagcct	ccaataactc	taccctttct	tcgggatcag	28680
gtgacttttc	tgaaatcggg	cttggtgtgc	tgcttactct	gttgattttt	ttccttatca	28740
tactcagcct	tctgtgcctc	aggctcgccg	cctgctgcgc	acacatctat	atctactgct	28800
ggttgctcaa	gtgcaggggt	cgccacccaa	gatgaacagg	tacatgggtcc	tatcgatcct	28860
aggcctgctg	gccctggcgg	cctgcagcgc	cgccaaaaaa	gagattacct	ttgaggagcc	28920
cgcttgcaat	gtaactttca	agcccgaggg	tgaccaatgc	accaccctcg	tcaaatgcgt	28980
taccaatcat	gagaagctgc	gcatcgacta	caaaaacaaa	actggccggg	ttgcgggtcta	29040
tagtgtgttt	acgcccggag	accctcttaa	ctactctgtc	accgtcttcc	agggcggaca	29100
gtctaagata	ttcaattaca	ctttcccttt	ttatgagttg	tgcgatgcgg	tcatgtacat	29160
gtcaaaacag	tacaacctgt	ggcctccctc	tcccagggcg	tgtgtggaaa	atactgggtc	29220
ttactgctgt	atggcttttg	caatcactac	gctcgctcta	atctgcacgg	tgctatatat	29280

aaaattcagg	cagaggcgaa	tctttatcga	tgaaaagaaa	atgccttgat	cgctaacacc	29340
ggcttttctat	ctgcagaatg	aatgcaatca	cctccctact	aatcaccacc	accctccttg	29400
cgattgccc	tgggttgaca	cgaatcgaag	tgccagtggg	gtccaatgtc	accatggtgg	29460
gccccgccg	caattccacc	ctcatgtggg	aaaaatttgt	ccgcaatcaa	tgggttcatt	29520
tctgctctaa	ccgaatcagt	atcaagccca	gagccatctg	cgatgggcaa	aatctaactc	29580
tgatcaatgt	gcaaatgatg	gatgctgggt	actattacgg	gcagcgggga	gaaatcatta	29640
attactggcg	acccacaag	gactacatgc	tgcatgtagt	cgaggcactt	cccactacca	29700
ccccactac	cacctctccc	accaccacta	ccactactac	tactactact	accactaccg	29760
ctgcccccca	taccgcgaaa	agcaccatga	ttagcacaaa	gccccctcgt	gctcactccc	29820
acgccggcg	gcccacgggt	gcgacctcag	aaaccaccga	gctttgcttc	tgccaatgca	29880
ctaacgccag	cgctcatgaa	ctgttcgacc	tggagaatga	ggatgccag	cagagctccg	29940
cttgccctgac	ccaggaggct	gtggagcccg	ttgccctgaa	gcagatcggg	gattcaataa	30000
ttgactcttc	ttcttttgcc	actcccgaat	accctccga	ttctactttc	cacatcacgg	30060
gtaccaaaga	ccctaacctc	tctttctacc	tgatgctgct	gctctgtatc	tctgtggtct	30120
cttccgcgct	gatgttactg	gggatgttct	gctgcctgat	ctgccgcaga	aagagaaaag	30180
ctcgctctca	gggccaacca	ctgatgccct	tcccctaccc	cccggatttt	gcagataaca	30240
agatatgagc	tcgctgctga	cactaaccgc	tttactagcc	tgcgctctaa	cccttgctgc	30300
ttgcgactcg	agattccaca	atgtcacagc	tgtggcagga	gaaaatgtta	ctttcaactc	30360
cacggccgat	accagtggt	cgtggagtgg	ctcaggtagc	tacttaacta	tctgcaatag	30420
ctccacttcc	cccagcatat	ccccaaccaa	gtaccaatgc	aatgccagcc	tgttcaccct	30480
catcaacgct	tccaccctgg	acaatggact	ctatgtaggc	tatgtaccct	ttggtgggca	30540
aggaaagacc	cacgcttaca	acctggaagt	tcgccagccc	agaaccacta	ccaagctwc	30600
ymccaycacc	agcaccagca	gcagcagcca	cagcagcagc	agcagattat	tgacttttgg	30660
tttggccagc	tcattctgcc	ctaccaggcc	catctacagc	tctgtgcccg	aaaccactca	30720
gaccaccgc	ccagaaacga	ccaccgccac	caccctacac	acctccagcg	atcagatgcc	30780
gaccaacatc	acccccttgg	ctcttcaa	gggacttaca	agccccactc	caaaaccagt	30840
ggatgcggcc	gaggtctccg	ccctcgtaa	tgactgggcg	gggctgggaa	tgtggtggtt	30900
cgccataggc	atgatggcgc	tctgcctgct	tctgctctgg	ctcatctgct	gcctccaccg	30960
caggcgagcc	agacccccca	tctatagacc	catcattgtc	ctgaaccccg	ataatgatgg	31020
gatccataga	ttggatggcc	tgaaaaacct	acttttttct	tttacagtat	gataaattga	31080
gacatgcctc	gcattttctt	gtacatgttc	cttctcccac	cttttctggg	gtgttctacg	31140
ctggccgctg	tgtctcacct	ggaggtagac	tgccctctcac	ccttcaactgt	ctacctgctt	31200

tacggattgg	tcaccctcac	tctcatctgc	agcctaataca	cagtaatcat	cgccttcate	31260
cagtgcattg	attacatctg	tgtgcgcctc	gcataacttca	gacaccaccc	gcagtaccga	31320
gacaggaaca	ttgcccaact	tctaagactg	ctctaatacat	gcataagact	gtgatctgcc	31380
ttctgatcct	ctgcatcctg	cccaccctca	cctcctgcc	gtacaccaca	aaatctccgc	31440
gcaaaagaca	tgccctcctgc	cgccttcaccc	aactgtggaa	tatacccaaa	tgctacaacg	31500
aaaagagcga	gctctccgaa	gcttgggtgt	atgggggtcat	ctgtgtctta	gttttctgca	31560
gcactgtctt	tgccctcatg	atctaccct	actttgat	gggatggaac	gcgatcgatg	31620
ccatgaatta	ccccaccttt	cccgcacccg	agataattcc	actgcgacaa	gttgtaccgc	31680
ttgtcgtaaa	tcaacgcccc	ccatccccta	cgcacctga	aatcagctac	tttaacctaa	31740
caggcggaga	tgactgacgc	cctagatcta	gaaatggacg	gcacagctac	cgagcagcgt	31800
ctcctagaga	ggcgcaggca	ggcggctgag	caagagcgcc	tcaatcagga	gctccgagat	31860
ctcgtaaac	tgaccagtg	caaaagaggc	atcttttgtc	tggtaaagca	ggccaaagtc	31920
acctacgaga	agaccggcaa	cagccaccgc	ctcagttaca	aattgcccac	ccagcgccag	31980
aagctggtgc	tcatggtggg	tgagaatccc	atcacgtca	cccagcactc	ggtagagacc	32040
gaggggtgtc	tgactcccc	ctgtcggggg	ccagaagacc	tctgcaccct	ggtaaagacc	32100
ctgtgcggtc	tcagagattt	agccccctt	aactaatcaa	acactggaat	caataaaaag	32160
aatcacttac	ttaaaatcag	acagcaggtc	tctgtccagt	ttattcagca	gcacctcctt	32220
ccccctctcc	caactctggt	actccaaacg	ccttctggcg	gcaaacttcc	tccacaccct	32280
gaagggaatg	tcagattctt	gctcctgtcc	ctccgcaccc	actatcttca	tggtgttgca	32340
gatgaagcgc	acaaaaacgt	ctgacgagag	cttcaacccc	gtgtaccctt	atgacacgga	32400
aagcggccct	ccctccgtcc	ctttcctcac	ccctcccttc	gtgtctcccg	atggattcca	32460
agaaagtccc	cccgggggtcc	tgtctctgaa	cctggccgag	cccctgggtca	cttcccacgg	32520
catgctcgcc	ctgaaaatgg	gaagtggcct	ctccctggac	gacgctggca	acctcacctc	32580
tcaagatatc	accaccgcta	gccctcccct	caaaaaaac	aagaccaacc	tcagcctaga	32640
aacctcatcc	cccctaactg	tgagcacctc	aggcgccctc	accgtagcag	ccgccgctcc	32700
cctggcggtg	gccggcacct	ccctcaccat	gcaatcagag	gccccctga	cagtacagga	32760
tgcaaaactc	acctggcca	ccaaaggccc	cctgaccgtg	tctgaaggca	aactggcctt	32820
gcaaacatcg	gccccgctga	cggccgctga	cagcagcacc	ctcacgtta	gcgccacacc	32880
accaattaat	gtaagcagtg	gaagttagg	cttagacatg	gaagacccta	tgtataactca	32940
caatggaaaa	ctgggaataa	gaattggggg	tccactaaga	gtagtagaca	gcttgcatat	33000
actcactgta	gttaccggaa	atggactaac	tgtagataac	aatgccctcc	aaactaaagt	33060
tacgggcgcc	ctaggttatg	acacatcagg	aaatctacaa	ttaagagctg	caggaggtat	33120

gcgaattgac	gcaaattggcc	aacttatacct	taatgtggca	taccattttg	atgctcagaa	33180
caatctcagc	cttagacttg	gtcagggacc	cctgtatata	aacacagacc	acaacctgga	33240
tttgaattgc	aacagaggtc	taaccacaac	taccaccaac	aacacaaaaa	aacttgagac	33300
taaaattagc	tcaggcttag	actatgacac	caatgggtgct	gtcattatta	aacttggcac	33360
tgggtctaagc	ttcgacaaca	caggcgccct	aactgtggga	aacactgggtg	atgataaaact	33420
gactctgtgg	acgaccccg	acccatctcc	aaattgcaga	attcactcag	aaaagactg	33480
caagtttact	ctagtcttaa	ctaagtgtgg	aagccaaatc	ctggcctctg	tcgccgcctt	33540
agcggtatca	ggaaatctgg	cttcgataac	aggcacggtt	gccagcggtta	ccatctttct	33600
cagatttgat	cagaatggag	tgcttatgga	aaactcctcg	ctagacaggc	agtactggaa	33660
cttcagaaat	ggcaactcaa	ctaacgctgc	cccctacacc	aatgcagttg	ggttcatgcc	33720
aaacctcgca	gcatacccca	aaacgcaaag	ccagactgct	aaaaacaaca	ttgtaagtca	33780
ggtttacttg	aatggagaca	aatccaaacc	catgaccctt	accatcacc	tcaatggaac	33840
taatgaatcc	agtgaaacta	gccagggtgag	tcactactcc	atgtcattta	catgggcttg	33900
ggaaagtggg	caatatgcc	ctgaaacctt	tgccaccaac	tccttcacct	tttcttacat	33960
tgctgaacaa	taaaaagcat	gacactgatg	ttcatttctg	attcttattt	tattattttc	34020
aaacacaaca	aatcattca	agtcattctt	ccatcttagc	ttaatagaca	cagtagctta	34080
atagaccag	tagtgcaaag	ccccattcta	gcttatagat	cagacagtga	taattaacca	34140
ccaccaccac	catacctttt	gattcaggaa	atcatgatca	tcacaggatc	ctagtcttca	34200
ggccgcccc	tccttcccaa	gacacagaat	acacagtcct	ctccccccga	ctggctttta	34260
ataacaccat	ctggttggtc	acagacatgt	tcttaggggt	tatattccac	acggtctcct	34320
gccgcgccag	gcgctcgtcg	gtgatgttga	taaactctcc	cggcagctcg	ctcaagttca	34380
cgtcgctgtc	cagcggctga	acctccggct	gacgcgataa	ctgtgcgacc	ggctgctgga	34440
caaacggagg	ccgcgcctac	aagggggtag	agtcataatc	ctcggtcagg	atagggcggt	34500
gatgcagcag	cagcgagcga	aacatctgct	gccgcgccg	ctccgtccgg	caggaaaaca	34560
acacgccggt	ggtctcctcc	gcgataatcc	gcaccgcccg	cagcatcagc	ttcctcgttc	34620
tccgcgcgca	gcacctcacc	ctgatctcgc	tcaagtcggc	gcagtaggta	cagcacagca	34680
ccacgatgtt	attcatgata	ccacagtgca	gggcgctgta	tccaaagctc	atgccgggaa	34740
ccaccgcccc	cacgtggcca	tcgtaccaca	agcgcacgta	aattaagtgt	cgacccctca	34800
tgaacgtgct	ggacacaaac	attacttcct	tgggcatgtt	gtaattcacc	acctcccgtt	34860
accagataaa	cctctggtta	aacagggcac	cttccaccac	catcctgaac	caagaggcca	34920
gaacctgccc	accggctatg	cactgcaggg	aaccggggtt	ggaacaatga	caatgcagac	34980
tccaaggctc	gtaaccgtgg	atcatccggc	tgctgaaggc	atcgatgttg	gcacaacaca	35040

gacacacgtg	catgcacttt	ctcatgatta	gcagctcttc	cctcgtcagg	atcatatccc	35100
aaggaataac	ccattcttga	atcaacgtaa	aacccacaca	gcagggaagg	cctcgcacat	35160
aactcacgtt	gtgcatggtc	agcgtgttgc	attccggaaa	cagcggatga	tcctccagta	35220
tcgaggcgcg	ggtctccttc	tcacaggag	gtaaagggtc	cctgctgtac	ggactgcgcc	35280
gggacgaccg	agatcgtgtt	gagcgtagt	tcattggaaaa	gggaacgccg	gacgtggtca	35340
tacttcttga	agcagaacca	ggttcgcgcg	tggcaggcct	ccttgcgctc	gcggtctcgc	35400
cgtctagctc	gctccgtgtg	atagttgtag	tacagccact	cccgcagagc	gtcgaggcg	35460
accctggctt	ccggatctat	gtagactccg	tcttgccaccg	cggccctgat	aatatccacc	35520
accgtagaat	aagcaacacc	cagccaagca	atacactcgc	tctgcgagcg	gcagacagga	35580
ggagcgggca	gagatgggag	aaccatgata	aaaaactttt	tttaaagaat	attttccaat	35640
tcttcgaaag	taagatctat	caagtggcag	cgctcccctc	cactggcgcg	gtcaaactct	35700
acggccaaag	cacagacaac	ggcatttcta	agatgttcct	taatggcgtc	caaaagacac	35760
accgctctca	agtcgcagta	aactatgaat	gaaaacccat	cgggctgatt	ttccaatata	35820
gacgcgccgg	cggcgtccac	caaaccaga	taattttctt	ctctccagcg	gtttagaatc	35880
tgtctaagca	aatcccttat	atcaagtccg	gccatgccaa	aaatctgctc	aagagcgccc	35940
tccaccttca	tgaccaagca	gcgcatcatg	attgcaaaaa	ttcaggttct	tcagagacct	36000
gtataagatt	caaaatggga	acattaacaa	aaattcctct	gtcgcgcaga	tccttctgca	36060
gggcaagctg	aacataatca	gacaggctctg	aacggaccag	tgaggccaaa	tccccaccag	36120
gaaccagatc	cagagaccct	atactgatta	tgacgcgcac	actcggggct	atgctgacca	36180
gcgtagcgcc	gatgtaggcg	tgctgcatgg	gcggcgagat	aaaatgcaaa	gtgctggtta	36240
aaaaatcagg	caaagcctcg	cgcaaaaaag	ctaacacatc	ataatcatgc	tcatgcaggt	36300
agttgcaggt	aagctcagga	acaaaaacgg	aataacacac	gattttctct	tcaaacaatga	36360
cttcgcggat	actgcgtaaa	acaaaaatta	taaataaaaa	attaattaac	ttaaaccattg	36420
gaagcctgtc	tcacaacagg	aaaaaccact	ttaatcaaca	taagacgggc	cacgggcag	36480
ccggcatagc	cgtaaaaaaa	ttggtccccg	tgattaacaa	gtaccacaga	cagctccccg	36540
gtcatgtcgg	gggtcatcat	gtgagactct	gtatacacgt	ctggattgtg	aacatcagac	36600
aaacaaagaa	atcgagccac	gtagcccgga	ggtataatca	cccgcaggcg	gaggtacagc	36660
aaaacgaccc	ccataggagg	aatcacaaaa	ttagtaggag	aaaaaaatac	ataaacacca	36720
gaaaaaccct	gttgctgagg	caaaatagcg	ccctcccgat	ccaaaacaac	ataaagcgct	36780
tccacaggag	cagccataac	aaagaccgga	gtcttaccag	taaaaagaaa	aaagatctct	36840
caacgcagca	ccagcaccaa	cacttcgcag	tgtaaaaggc	caagtgccga	gagagtatat	36900
ataggaataa	aaagtgcagt	aaacgggcaa	agtcacaaaa	acgccagaaa	aaaccgcacg	36960

cgaacctacg	ccccgaaacg	aaagccaaaa	aacactagac	actcccttcc	ggcgtcaact	37020
tccgctttcc	cacgctacgt	cacttgcccc	agtcaaacaa	actacatatc	ccgaacttcc	37080
aagtcgccac	gccccaaaaca	ccgcctacac	ctccccgccc	gccggccccgc	ccccaaaccc	37140
gcctccccgc	ccgcgccccg	ccccgcgccc	cccattctcat	tatcatattg	gcttcaatcc	37200
aaaataaggt	atattattga	tgatggttta	aacggatcca	attcttgaag	acgaaagggc	37260
ctcgtgatac	gcctatTTTT	ataggTTaat	gtcatgataa	taatggTTtc	ttagacgtca	37320
ggtggcactt	ttcggggaaa	tgtgcgcgga	accctatTTt	gTTtattTTt	ctaaatacat	37380
tcaaatatgt	atccgctcat	gagacaataa	ccctgataaa	tgcttcaata	atattgaaaa	37440
aggaagagta	tgagtattca	acatttccgt	gtcgccctta	ttccctTTTT	tgcggcattt	37500
tgcttctctg	TTTTtgctca	cccagaaacg	ctggtgaaag	taaaagatgc	tgaagatcag	37560
ttgggtgcac	gagtgggtta	catcgaactg	gatctcaaca	gcggtaatga	ccttgagagt	37620
tttcgccccg	aagaacgttt	tccaatgatg	agcactTTta	aagTTctgct	atgtggcgcg	37680
gtattatccc	gtgttgacgc	cgggcaagag	caactcggtc	gccgcataca	ctattctcag	37740
aatgacttgg	ttgagtactc	accagtcaca	gaaaagcatc	ttacggatgg	catgacagta	37800
agagaattat	gcagtgtctg	cataaccatg	agtgataaca	ctgcggccaa	cttacttctg	37860
acaacgatcg	gaggaccgaa	ggagctaacc	gctTTTTtgc	acaacatggg	ggatcatgta	37920
actcgcttgg	atcgttggga	accggagctg	aatgaagcca	taccaaacga	cgagcgtgac	37980
accacgatgc	ctgtagcaat	ggcaacaacg	ttgcgcaaac	tattaactgg	cgaactactt	38040
actctagctt	cccggcaaca	attaatagac	tggatggagg	cggataaagt	tgcaggacca	38100
cttctgcgct	cggcccttcc	ggctggctgg	tttattgctg	ataaatctgg	agccggtgag	38160
cgtgggtctc	gcggtatcat	tgacgactg	gggccagatg	gtaagccctc	ccgtatcgta	38220
gttatctaca	cgacggggag	tcaggcaact	atggatgaac	gaaatagaca	gatcgctgag	38280
ataggtgcct	cactgattaa	gcattggtaa	ctgtcagacc	aagTTtactc	atatatactt	38340
tagattgatt	taaaaggatc	taggtgaaga	tcTTTTtga	taatctcatg	acaaaaatcc	38400
cttaacgtga	gttttcgttc	cactgagcgt	cagaccccg	agaaaagatc	aaaggatctt	38460
cttgagatcc	TTTTtttctg	cgcgtaatct	gctgcttgca	aacaaaaaaaa	ccaccgctac	38520
cagcggtggt	ttgtttgccg	gatcaagagc	taccaactct	TTTTccgaag	gtaactggct	38580
tcagcagagc	gcagatacca	aatactgtcc	ttctagtgtg	gccgtagtta	ggccaccact	38640
tcaagaactc	tgtagcaccg	cctacatacc	tcgctctgct	aatcctgtta	ccagtggctg	38700
ctgccagtgg	cgataagtcg	tgtcttaccg	ggttggactc	aagacgatag	ttaccggata	38760
aggcgcagcg	gtcgggctga	acgggggggt	cgtgcacaca	gccagcttg	gagcgaacga	38820
cctacaccga	actgagatac	ctacagcgtg	agctatgaga	aagcgccacg	cttcccgaag	38880

ggagaaaggc ggacaggtat ccggtaaagcg gcagggtcgg aacaggagag cgcacgaggg	38940
agcttccagg gggaaacgcc tggatatcttt atagtcctgt cgggtttcgc cacctctgac	39000
ttgagcgtcg atttttgtga tgctcgtcag gggggcggag cctatggaaa aacgccagca	39060
acgcggcctt tttacggttc ctggcctttt gctggccttg aagctgtccc tgatggtcgt	39120
catctacctg cctggacagc atggcctgca acgcgggcat cccgatgccg ccggaagcga	39180
gaagaatcat aatggggaag gccatccagc ctcgcgctcg agatccgaat tcgtttaaac	39240

<210> 18
 <211> 33
 <212> DNA
 <213> Pan troglodytes

<400> 18	
atacggacta gtggagaagt actcgcctac atg	33

<210> 19
 <211> 36
 <212> DNA
 <213> Pan troglodytes

<400> 19	
atacgaaga tctaagactt caggaaatat gactac	36

<210> 20
 <211> 45
 <212> DNA
 <213> Pan troglodytes

<400> 20	
attcagtgt caggcgcgcc aaagcatgac actgatgttc atttc	45

<210> 21
 <211> 35
 <212> DNA
 <213> Pan troglodytes

<400> 21	
actaggacta gttataagct agaatggggc tttgc	35

<210> 22
 <211> 37740
 <212> DNA
 <213> Pan troglodytes

<220>
 <221> misc_feature
 <222> (24391)..(24393)
 <223> n iðååñòàâäëÿàò ñîáîé a, c, g, èèè t

<400> 22	
catcatcaat aatatacctt attttggatt gaagccaata tgataatgag atgggcggcg	60

cggggcgggg	cgcggggcgg	gaggcggggt	tgggggcggg	ccggcgggcg	gggcggtgtg	120
gcggaagtgg	actttgtaag	tgtggcggat	gtgacttgct	agtgcggggc	gcggtaaaag	180
tgacgttttc	cgtgcgcgac	aacgcccccg	ggaagtgaca	tttttcccgc	ggtttttacc	240
ggatgttgta	gtgaatttgg	gcgtaaccaa	gtaagatttg	gccattttcg	cgggaaaact	300
gaaacgggga	agtgaaatct	gattaatttt	gcgttagtca	taccgcgtaa	tatttgtcta	360
gggccgaggg	actttggccg	attacgtgga	ggactcgccc	aggtgttttt	tgagggtgaat	420
ttccgcgttc	cgggtcaaag	tctgcgtttt	attattatag	gatatcccat	tgcatacgtt	480
gtatccatat	cataatatgt	acatttatat	tggctcatgt	ccaacattac	cgccatgttg	540
acattgatta	ttgactagtt	attaatagta	atcaattacg	gggtcattag	ttcatagccc	600
atatatggag	ttccgcgtta	cataacttac	ggtaaatggc	ccgcctggct	gaccgcccaa	660
cgacccccgc	ccattgacgt	caataatgac	gtatgttccc	atagtaacgc	caatagggac	720
ttccattga	cgtcaatggg	tggagtattt	acggtaaact	gccacttgg	cagtacatca	780
agtgtatcat	atgccaagta	cgccccctat	tgacgtcaat	gacggtaaata	ggcccgcctg	840
gcattatgcc	cagtacatga	ccttatggga	ctttcctact	tggcagtaca	tctacgtatt	900
agtcacgcgt	attaccatgg	tgatgcgggt	ttggcagtag	atcaatgggc	gtggatagcg	960
gtttgactca	cggggatttc	caagtctcca	ccccattgac	gtcaatggga	gtttgttttg	1020
gcacaaaaat	caacgggact	ttccaaaatg	tcgtaacaac	tccgccccat	tgacgcaaata	1080
gggcggtagg	cgtgtacggt	gggaggtcta	tataagcaga	gctctcccta	tcagtgatag	1140
agatctccct	atcagtgata	gagatcgtcg	acgagctcgt	ttagtgaacc	gtcagatcgc	1200
ctggagacgc	catccacgct	gttttgacct	ccatagaaga	caccgggacc	gatccagcct	1260
ccgcggccgg	gaacggtgca	ttggaacgcg	gattccccgt	gccaagagtg	agatcttccg	1320
tttatctagg	taccggggcc	cccctcgagg	tcgacggtat	cgataagctt	cacgctgccg	1380
caagcactca	gggcgcaagg	gctgctaaag	gaagcggaac	acgtagaaag	ccagtccgca	1440
gaaacggtgc	tgaccccgga	tgaatgtcag	ctactgggct	atctggacaa	gggaaaacgc	1500
aagcgcaaag	agaaagcagg	tagcttgcag	tgggcttaca	tggcgatagc	tagactgggc	1560
ggttttatgg	acagcaagcg	aaccggaatt	gccagctggg	gcgccctctg	gtaagggttg	1620
gaagccctgc	aaagtaaact	ggatggcttt	cttgccgcca	aggatctgat	ggcgcagggg	1680
atcaagatct	aaccaggagc	tatttaatgg	caacagttaa	ccagctggta	cgcaaaccac	1740
gtgctcgcaa	agttgcgaaa	agcaacgtgc	ctgcgctgga	agcatgcccg	caaaaacgtg	1800
gcgtatgtac	tcgtgtatat	actaccactc	ctaaaaaacc	gaactccgcg	ctgcgtaaag	1860
tatgccgtgt	tcgtctgact	aacggtttcg	aagtgacttc	ctacatcggt	ggtgaaggtc	1920
acaacctgca	ggagcactcc	gtgatcctga	tccgtggcgg	tcgtgttaaa	gacctcccgg	1980

gtgttcgtta ccacaccgta cgtggtgctc ttgactgctc cggcggttaa gaccgtaagc	2040
aggctcgttc caagtatggc gtgaagcgtc ctaaggctta atggtagatc tgatcaagag	2100
acaggatgac ggtcgtttcg catgcttgaa caagatggat tgcacgcagg ttctccggcc	2160
gcttgggtgg agaggctatt cggctatgac tgggcacaa agacaatcgg ctgctctgat	2220
gccgccgtgt tccggctgtc agcgcagggg cgcccggttc tttttgtcaa gaccgacctg	2280
tccgggtgcc tgaatgaact gcaggacgag gcagcgcggc tatcgtggct ggccacgacg	2340
ggcgttcctt gcgcagctgt gctcgacgtt gtcactgaag cgggaaggga ctggctgcta	2400
ttgggcgaag tgccggggca ggatctcctg tcatctcacc ttgctcctgc cgagaaagta	2460
tccatcatgg ctgatgcaat gcggcggtg catacgcttg atccggctac ctgcccattc	2520
gaccaccaag cgaaacatcg catcgagcga gcacgtactc ggatggaagc cggctctgtc	2580
gatcaggatg atctggacga agagcatcag gggctcgcgc cagccgaact gttcgccagg	2640
ctcaaggcgc gcatgcccga cggcgaggat ctcgctgtga cccatggcga tgctgttg	2700
ccgaatatca tgggtgaaaa tggccgcttt tctggattca tgcactgtgg ccggctgggt	2760
gtggcggacc gctatcagga catagcgttg gctaccctg atattgctga agagcttggc	2820
ggcgaatggg ctgaccgctt cctcgtgctt tacggtatcg ccgctcccga ttgcagcgc	2880
atcgcttct atcgcttct tgacgagttc ttctgagcgg gactctgggg ttcgaaatga	2940
ccgaccaagc gacgcccac ctgccatcac gagatttcga ttccaccgcc gccttctatg	3000
aaaggttggg cttcggaatc gttttccggg acgccggctg gatgatcctc cagcgcgggg	3060
atctcatgct ggagttcttc gccaccccg ggctcgatcc cctcgggggg aatcagaatt	3120
cagtcgacag cggccgcgat ctgctgtgcc ttctagttgc cagccatctg ttgtttgcc	3180
ctccccctg ccttccttga ccctggaagg tgccactccc actgtccttt cctaataaaa	3240
tgaggaaatt gcatcgcatt gtctgagtag gtgtcattct attctggggg gtgggggtgg	3300
gcaggacagc aagggggagg atttggaaga caatagcagg catgctgggg atgcgggtgg	3360
ctctatggcc gatcagcgat cgctgaggtg ggtgagtggt cgtggcctgg ggtggtcatg	3420
aaaatatata agttgggggt cttaggggtc ttttatttgt gttgcagaga ccgccggagc	3480
catgagcggg agcagcagca gcagcagtag cagcagcgcc ttggatggca gcatcgtgag	3540
cccttatttg acgacgcgga tgccccactg ggccgggggt cgtcagaatg tgatgggctc	3600
cagcatcgac ggccgacctg tcctgccccg aaattccgcc acgctgacct atgcgacct	3660
cgccgggacg ccgttgacg ccaccgccgc cgccgccgcc accgcagccg cctcggccgt	3720
gcgcagcctg gccacggact ttgcattcct gggaccactg gcgacagggg ctacttctcg	3780
ggccgctgct gccgccgttc gcgatgacaa gctgaccgcc ctgctggcgc agttggatgc	3840
gcttactcgg gaactgggtg acctttctca gcaggatcat gccctgcgcc agcagggtctc	3900

ctccctgcaa gctggcgggga atgcttctcc cacaaatgcc gtttaagata aataaaacca	3960
gactctgttt ggattaaaga aaagtagcaa gtgcattgct ctctttattt cataattttc	4020
cgcgcgcat aggcctaga ccagcgttct cggtcgttga gggtgcggtg tatcttctcc	4080
aggacgtggt agaggtggct ctggacgttg agatacatgg gcatgagccc gtcccggggg	4140
tggaggtagc accactgcag agcttcatgc tccgggggtg tgttgtagat gatccagtcg	4200
tagcaggagc gctgggcatg gtgcctaaaa atgtccttca gcagcaggcc gatggccagg	4260
gggaggccct tgggtgtaagt gtttacaaaa cggttaagtt gggaagggtg cattcgggga	4320
gagatgatgt gcatcttgga ctgtatTTTT agattggcga tgtttccgcc cagatccctt	4380
ctgggattca tgttgTgcag gaccaccagt acagtgtatc cgggtgcactt ggggaatttg	4440
tcatgcagct tagagggaaa agcgtggaag aacttggaaga cgcccttgTg gcctcccaga	4500
ttttccatgc attcgtccat gatgatggca atgggcccgc gggaggcagc ttgggcaaag	4560
atatttctgg ggtcgctgac gtcgtagttg tgttccaggg tgaggTcgTc ataggccatt	4620
tttaciaaagc gcgggcggag ggtgcccga tgggggatga tggTcccctc tggccccggg	4680
gcgtagttgc cctcgcatat ctgcatttcc caggccttaa tctcgagggg gggaatcata	4740
tccacctgcg gggcgatgaa gaaaacggtt tccggagccg gggagattaa ctgggatgag	4800
agcaggtttc taagcagctg tgattttcca caaccggtgg gccataaat aacacctata	4860
accggttgca gctggtagtt tagagagctg cagctgccgt cgtcccggag gaggggggcc	4920
acctcgttga gcatgtccct gacgcgcatg ttctccccga ccagatccgc cagaaggcgc	4980
tgcgcgcca gggacagcag ctcttgcaag gaagcaaagt ttttcagcgg cttgaggccg	5040
tccgccgtgg gcatgttttt cagggTctgg ctcagcagct ccaggcggtc ccagagctcg	5100
gtgacgtgct ctacggcatc tctatccagc atatctctc gtttcgcggg ttggggcgac	5160
tttcgctgta gggcaccaag cggTggTcgt ccagcggggc cagagTcatg tccttccatg	5220
ggcgcagggT cctcgTcagg gtggtctggg tcacggTgaa ggggtgcgct ccgggctgag	5280
cgcttgccaa ggtgcgcttg aggctggTtc tgctggTgct gaagcgctgc cggTcttcgc	5340
cctgcgcgTc ggccaggtag catttgacca tggTgtcata gtccagcccc tccgcggcgt	5400
gtcccttggc gcgcagcttg cccttgagg tggcgccgca cgaggggcag agcaggctct	5460
tgagcgcgta gagcttgggg gcgaggaaga ccgattcggg ggagtaggcg tccgcgcgc	5520
agaccccgca cacggtctcg cactccacca gccaggTgag ctcgggggcg gccgggtcaa	5580
aaaccaggtt tccccatgc tttttgatgc gtttcttacc tcgggtctcc atgaggTggt	5640
gtccccgctc ggtgacgaag aggctgtccg tgtctccgta gaccgacttg aggggtcttt	5700
tctccagggg ggtccctcgg tcttctcgt agaggaactc ggaccactct gagacgaagg	5760
cccgcgTcca ggccaggacg aaggaggcta tgtgggaggg gtagcggtcg ttgtccacta	5820

gggggtccac	cttctccaag	gtgtgaagac	acatgtcgcc	ttcctcggcg	tccaggaagg	5880
tgattggctt	gtaggtgtag	gccacgtgac	cgggggttcc	tgacgggggg	gtataaaagg	5940
gggtgggggc	gcgctcgtcg	tcactctctt	ccgcatcgct	gtctgcgagg	gccagctgct	6000
ggggtgagta	ttccctctcg	aaggcgggca	tgacctccgc	gctgagggtt	tcagtttcca	6060
aaaacgagga	ggatttgatg	ttcacctgtc	ccgagggtgat	acctttgagg	gtacccgcgt	6120
ccatctggtc	agaaaacacg	atctttttat	tgtccagctt	ggtggcgaa	gacccgtaga	6180
gggctgttga	gagcagcttg	gcatgggagc	gcaggggtctg	gttcttgtcc	ctgtcggcgc	6240
gctccttggc	cgcatgttg	agctgcacgt	actcgcgcgc	gacgcagcgc	cactcgggga	6300
agacggtggt	gcgctcgtcg	ggcaccaggc	gcacgcgcca	gccgcggtt	tgacgggtga	6360
ccagggtccac	gctggtggcg	acctcgccgc	gcaggcgctc	gttgggtccag	cagagacggc	6420
cgcccttgcg	cgagcagaag	gggggcaggg	ggtcgagctg	ggtctcgtcc	ggggggtccg	6480
cgtccacggt	gaaaaccccg	gggcgcaggc	gcgcgtcgaa	gtagtctatc	ttgcaacctt	6540
gcatgtccag	cgctcgtgc	cagtcgcggg	cggcgagcgc	gcgctcgtag	gggttgagcg	6600
gcgggcccc	gggcatgggg	tgggtgagtg	cggaggcgta	catgccgcag	atgtcataga	6660
cgtagagggg	ctccgcagg	acccgatgt	aggtggggta	gcagcggccg	ccgcggatgc	6720
tggcgcgcac	gtagtcatac	agctcgtcg	agggggcgag	gaggtcgggg	cccaggttgg	6780
tgcgggcggg	gcgctccgtg	cggaagacga	tctgcctgaa	gatggcatgc	gagttggaag	6840
agatggtggg	gcgctggaag	acgttgaagc	tggcgtcctg	caggccgacg	gcgtcgcgca	6900
cgaaggaggc	gtaggagtcg	cgcagcttgt	gtaccagctc	ggcggtgacc	tgacagtcga	6960
gcgcgcagta	gtcaggggtc	tcgcggatga	tgtcataatt	agcctgcccc	ttctttttcc	7020
acagctcgcg	gttgaggaca	aactcttcgc	ggtctttcca	gtactcttgg	atcgggaaac	7080
cgtccgggtc	cgaacggtaa	gagcctagca	tgtagaactg	gttgacggcc	tggtaggcgc	7140
agcagccctt	ctccacgggg	agggcgtagg	cctgcgcggc	cttgcgaggc	gaggtgtggg	7200
tcagggcgaa	ggtgtccctg	accatgactt	tgaggtactg	gtgcttgaag	tcggagtcgt	7260
cgcagccgcc	ccgctcccag	agcgagaagt	cggtgcgctt	cttgagcgcg	gggttgggca	7320
gagcgaagg	gacatcgttg	aagaggattt	tgcccgcgcg	gggcatgaag	ttgcgggtga	7380
tgcggaagg	ccccggcact	tcagagcgg	tgttgatgac	ctgggcggcg	agcacgatct	7440
cgtcgaagcc	gttgatgttg	tggcccacga	tgtagagtcc	caggaagcgg	ggccggccct	7500
ttacggtggg	cagcttcttt	agctcttcgt	aggtgagctc	ctcgggcgag	gcgaggccgt	7560
gctcggccag	ggcccagtc	gcgagggtcg	ggttgtctct	gaggaaggac	tcccagaggt	7620
cgcgggccag	gaggggtctg	aggcgggtcc	tgaaggctct	gaactggcgg	cccacggcca	7680
ttttttcggg	ggtgatgcag	tagaagggtga	gggggtcttg	ctgccagcgg	tcccagtcga	7740

gctgcagggc	gaggtcgcgc	gcggcggtga	ccaggcgctc	gtcgcccccg	aatttcatga	7800
ccagcatgaa	gggcacgagc	tgctttccga	aggcccccat	ccaagtgtag	gtctctacat	7860
cgtaggtgac	aaagaggcgc	tccgtgcgag	gatgcgagcc	gatcgggaag	aactggatct	7920
cccgccacca	gttggaggag	tggtctgtga	tgtggtggaa	gtagaagtcc	cgtcgcgggg	7980
ccgaacactc	gtgctggctt	ttgtaaaagc	gagcgagta	ctggcagcgc	tgcacgggct	8040
gtacctcatg	cacgagatgc	accttttcgc	cgcgacgag	gaagccgagg	ggaaatctga	8100
gccccccg	tggtctcgcg	catggctggg	gctcttctac	tttggatgcg	tgtccgtctc	8160
cgtctggctc	ctcgaggggt	gttacgggtg	agcggaccac	cacgccgcgc	gagccgcagg	8220
tccagatata	ggcgcgcggc	ggtcggagtt	tgatgacgac	atcgcgagc	tgggagctgt	8280
ccatggtctg	gagctccgc	ggcggcggca	ggtcagccgg	gagttcttgc	aggttcacct	8340
cgcagagtcg	ggccagggcg	cggggcagg	ctaggtggta	cctgatctct	aggggcgtgt	8400
tggtggcggc	gtcgatggct	tgaggagcc	cgcagccccg	gggggcgacg	acggtgcccc	8460
gcggggtggt	ggtggtggtg	gcggtgcagc	tcagaagcgg	tgccgcgggc	gggcccccg	8520
aggtaggggg	ggctccggct	ccgcgggcag	gggcggcagc	ggcacgtcgg	cgtggagcgc	8580
gggcaggagt	tggtgctgtg	cccggagggt	gctggcgaag	gcgacgacgc	ggcgggtgat	8640
ctcctggatc	tggcgcctct	gcgtgaagac	gacgggcccc	gtgagcttga	acctgaaaga	8700
gagttcgaca	gaatcaatct	cggtgtcatt	gaccgcggcc	tggcgcagga	tctcctgcac	8760
gtctcccag	ttgtcttgg	aggcgatctc	ggccatgaac	tgctcgatct	cttcctcctg	8820
gaggtctccg	cgtccggcgc	gttccacgg	ggccgccagg	tcgttggaga	tgcgccccat	8880
gagctgcgag	aaggcgttga	gtccgcctc	gttccagact	cggctgtaga	ccacgcccc	8940
ctggtcatcg	cgggcgcgca	tgaccacctg	cgcgaggttg	agctccacgt	gccgcgcgaa	9000
gacggcgtag	ttgcgcagac	gctggaagag	gtagttgagg	gtggtggcgg	tgtgctcggc	9060
cacgaagaag	ttcatgacct	agcggcgcaa	cgtggattcg	ttgatgtccc	ccaaggcctc	9120
cagccgttcc	atggcctcgt	agaagtccac	ggcgaagtgt	aaaaactggg	agttgcgcgc	9180
cgacacggtc	aactcctcct	ccagaagacg	gatgagctcg	gcgacggtgt	cgcgcacctc	9240
gcgctcgaag	gctatgggga	tctcttcctc	cgctagcatc	accacctcct	cctcttcctc	9300
ctcttctggc	acttccatga	tggttctcct	ctcttcgggg	ggtggcgggc	gcggcggtgg	9360
gggagggggc	gctctgcgcc	ggcggcggcg	caccgggagg	cggtcacga	agcgcgcgat	9420
catctccccg	cggcggcggc	gcatggtctc	ggtgacggcg	cggccgttct	cccgggggcg	9480
cagttggaag	acgccgcggg	acatctggtg	ctggggcggg	tggccgtgag	gcagcgagac	9540
ggcgctgacg	atgcatctca	acaattgctg	cgtaggtacg	ccgccgaggg	acctgaggga	9600
gtccatatcc	accggatccg	aaaacctttc	gaggaaggcg	tctaaccagt	cgcagtcgca	9660

aggtaggctg	agcacctggt	cgggcggcgg	ggggtggggg	gagtgtcttg	cggaggtgct	9720
gctgatgatg	taattgaagt	aggcggactt	gacacggcgg	atggtcgaca	ggagcaccat	9780
gtccttgggt	ccggcctgct	ggatgcggag	gcggtcggct	atgccccagg	cttcgttctg	9840
gcatcggcgc	aggtccttgt	agtagtcttg	catgagcctt	tccaccggca	cctctttctcc	9900
ttcctcttct	gcttcttcca	tgtctgcttc	ggccctgggg	cggcgccgcg	ccccctgcc	9960
ccccatgcgc	gtgaccccg	acccctgag	cggttggagc	agggccaggt	cggcgacgac	10020
gcgctcggcc	aggatggcct	gctgcacctg	cgtgaggggtg	gtttggaagt	catccaagtc	10080
cacgaagcgg	tggtaggcgc	ccgtgttgat	ggtgtaggtg	cagttggcca	tgacggacca	10140
gttgacggtc	tggtggcccc	gttgcgacat	ctcgggtgtac	ctgagtcgcg	agtaggcgcg	10200
ggagtcgaag	acgtagtcgt	tgcaagtccg	caccaggtac	tggtagccca	ccaggaagtg	10260
cggcggcggc	tggcggtaga	ggggccagcg	caggggtggcg	ggggctccgg	gggccaggtc	10320
ttccagcatg	aggcgggtgt	aggcgtagat	gtacctggac	atccaggtga	taccgcgggc	10380
ggtggtggag	gcgcgcggga	agtcgcgcac	ccggttccag	atgttgcgca	ggggcagaaa	10440
gtgctccatg	gtaggcgtgc	tctgtccagt	cagacgcgcg	cagtcggtga	tactctagac	10500
cagggaaaac	gaaagccggt	cagcgggcac	tcttcctgtg	tctggtgaat	agatcgcaag	10560
ggtatcatgg	cggagggcct	cggttcgagc	ccgggtccg	ggccggacgg	tccgccatga	10620
tccacgcggt	taccgcccgc	gtgtcgaacc	caggtgtgcg	acgtcagaca	acggtggagt	10680
gttccttttg	gcgtttttct	ggccggggcg	cggcgccgcg	taagagacta	agccgcgaaa	10740
gcgaaagcag	taagtggctc	gctccccgta	gccggaggga	tccttgctaa	gggttgcggt	10800
gcggcgaacc	ccggttcgaa	tcccgtactc	gggccggccg	gaccgcgggc	taaggtgttg	10860
gattggcctc	cccctcgat	aaagaccccg	cttgcggtat	gactccggac	acggggacga	10920
gcccctttta	tttttgcttt	ccccagatgc	atccggtgtt	gcgacagatg	cgccccccgc	10980
cccagcagca	gcaacaacac	cagcaagagc	ggcagcaaca	gcagcgggag	tcatgcaggg	11040
ccccctcacc	caccctcggc	ggccccggcca	cctcggcgctc	cgcggccgtg	tctggcgctt	11100
gcggcggcgg	cggcgggggg	ccggctgacg	accccagga	gcccccgcg	cgcagggcca	11160
gacactacct	ggacctggag	gagggcgagg	gcctggcgcg	gctggggggcg	ccgtctcccg	11220
agcgccaccc	gcgggtgcag	ctaaagcgcg	actcgcgcg	ggcgtagctg	cctcggcaga	11280
acctgttcag	ggaccgcgcg	ggcgaggagc	ccgaggagat	gcgggacagg	aggttcagcg	11340
cggggcggga	gctgcggcag	gggctgaacc	gcgagcggct	gctgcgcgag	gaggactttg	11400
agcccgcgcg	gcggacgggg	atcagccccg	cgcgcgcgca	cgtggcggcc	gccgacctgg	11460
tgacggcgta	cgagcagacg	gtgaaccagg	agatcaactt	ccaaaagagt	ttcaacaacc	11520
acgtgcgcac	gctggtggcg	cgcgaggagg	tgaccatcgg	gctgatgcac	ctgtgggact	11580

ttgtgagcgc	gctggtgcag	aaccccaata	gcaagcctct	gacggcgcag	ctgttcctga	11640
tagtgcagca	cagcagggac	aacgaggcgt	ttagggacgc	gctgctgaac	atcaccgagc	11700
ccgagggccg	gtggctgctg	gacctgatta	acatcctgca	gagcatagtg	gtgcaggagc	11760
gcagcctgag	cctggccgac	aaggtggcgg	ccatcaacta	ctcgatgctg	agcctgggca	11820
agttttacgc	gcgcaagatc	taccagacgc	cgtacgtgcc	catagacaag	gaggtgaaga	11880
tcgacggttt	ttacatgcgc	atggcgctga	aggtgctcac	cctaagcgac	gacctgggcg	11940
tgtaccgcaa	cgagcgcac	cacaaggccg	tgagcgtgag	ccggcgggcg	gagctgagcg	12000
accgcgagct	gatgcatagc	ctgcagcggg	cgctggcggg	cgccggcgagc	ggcgacaggg	12060
aggcggagtc	ctacttcgat	gcggggggcg	acctgcgctg	ggcgcccagc	cggcggggccc	12120
tggaggccgc	gggggtccgc	gaggactatg	acgaggacgg	cgaggaggat	gaggagtacg	12180
agctagagga	gggcgagtac	ctggactaaa	ccgcgggtgg	tgtttccggt	agatgcaaga	12240
cccgaacgtg	gtggaccg	cgctgcgggc	ggctctgcag	agccagccgt	ccggccttaa	12300
ctcctcagac	gactggcgac	aggtcatgga	ccgcatcatg	tcgctgacgg	cgcgtaaccc	12360
ggacgcgttc	cggcagcagc	cgcaggccaa	caggctctcc	gccatcctgg	aggcgggtggt	12420
gcctgcgcgc	tcgaacccca	cgcacgagaa	ggtgctggcc	atagtgaacg	cgctggccga	12480
gaacagggcc	atccgcccgg	acgaggccgg	gctggtgtac	gacgcgctgc	tgcagcgctg	12540
ggcccgtac	aacagcggca	acgtgcagac	caacctggac	cggctggtgg	gggacgtgcg	12600
cgaggcggtg	gcgcagcgcg	agcgcgcgga	tcggcagggc	aacctgggct	ccatggtggc	12660
gctgaatgcc	ttcctgagca	cgcagccggc	caacgtgccg	cgggggcagg	aagactacac	12720
caactttgtg	agcgcgctgc	ggctgatggt	gaccgagacc	ccccagagcg	aggtgtacca	12780
gtcgggtccg	gactacttct	tccagaccag	cagacagggc	ctgcagacgg	tgaacctgag	12840
ccaggctttc	aagaacctgc	gggggctgtg	gggcgtgaag	gcgcccaccg	gcgaccgggc	12900
gacggtgtcc	agcctgctga	cgcccaactc	gcgcctgctg	ctgctgctga	tcgcgcggtt	12960
cacggacagc	ggcagcgtgt	cccgggacac	ctacctgggg	cacctgctga	cctgtaccg	13020
cgaggccatc	gggcaggcgc	aggtggacga	gcacaccttc	caggagatca	ccagcgttag	13080
ccgcgcgctg	gggcaggagg	acacgagcag	cctggaggcg	actctgaact	acctgctgac	13140
caaccggcgg	cagaagattc	cctcgctgca	cagcctgacc	tccgaggagg	agcgcatctt	13200
gcgctacgtg	cagcagagcg	tgagcctgaa	cctgatgcgc	gacggggtga	cgcccagtgt	13260
ggcgctggac	atgaccgcgc	gcaacatgga	accgggcatg	tacgccgcgc	accggcctta	13320
catcaaccgc	ctgatggact	acctgcatcg	cgcggcggcc	gtgaaccccg	agtactttac	13380
caacgccatc	ctgaaccgcg	actggctccc	gccgcccggg	ttctacagcg	ggggcttcga	13440
ggtcccggag	gccaacgatg	gcttcctgtg	ggacgacatg	gacgacagcg	tgtttctccc	13500

gcggccgcag	gcgctggcgg	aagcgtccct	gctgcgtccc	aagaaggagg	aggaggcgag	13560
tcgccgccgc	ggcagcagcg	gcgtggcttc	tctgtccgag	ctggggggcg	cagccgccgc	13620
gcgccccggg	tccttgggcg	gcagcccctt	tccgagcctg	gtgggggtctc	tgcacagcga	13680
gcgcaccacc	cgccctcggc	tgctgggcga	ggacgagtac	ctgaataact	ccctgctgca	13740
gccggtgcgg	gagaaaaacc	tgccctccgc	cttccccaac	aacgggatag	agagcctggt	13800
ggacaagatg	agcagatgga	agacctatgc	gcaggagcac	agggacgcgc	ccgcgctccg	13860
gccgcccacg	cgccgcccagc	gccacgaccg	gcagcggggg	ctggtgtggg	atgacgagga	13920
ctccgcggac	gatagcagcg	tgctggacct	gggagggagc	ggcaaccctg	tcgcgcacct	13980
gcgccccgc	ctggggagga	tgttttaaaa	aaaaaaaaaa	gcaagaagca	tgatgcaaaa	14040
attaaataaa	actcaccaag	gccatggcga	ccgagcgttg	gtttcttgtg	ttcccttcag	14100
tatgcggcgc	gcggcgatgt	accaggaggg	acctcctccc	tcttacgaga	gcgtggtggg	14160
cgcgggcggc	gcggcgccct	cttctccctt	tgcgctgcag	ctgctggagc	cgccgtacgt	14220
gcctccgcgc	tacctgcggc	ctacgggggg	gagaaacagc	atccgttact	cggagctggc	14280
gccctgttc	gacaccaccc	gggtgtacct	ggtggacaac	aagtcggcgg	acgtggcctc	14340
cctgaactac	cagaacgacc	acagcaatth	tttgaccacg	gtcatccaga	acaatgacta	14400
cagcccagac	gaggccagca	cccagaccat	caatctggat	gaccggtcgc	actggggcgg	14460
cgacctgaaa	accatcctgc	acaccaacat	gccaacgtg	aacgagttca	tgttcaccaa	14520
taagttcaag	gcgcgggtga	tggtgtcgcg	ctcgcacacc	aaggaagacc	gggtggagct	14580
gaagtacgag	tgggtggagt	tcgagctgcc	agagggcaac	tactccgaga	ccatgaccat	14640
tgacctgatg	aacaacgcga	tcgtggagca	ctatctgaaa	gtgggcaggc	agaacggggg	14700
cctggagagc	gacatcgggg	tcaagttcga	caccaggaac	ttccgcctgg	ggctggaccc	14760
cgtgaccggg	ctggttatgc	ccggggtgta	caccaacgag	gccttccatc	ccgacatcat	14820
cctgctgccc	ggctgcgggg	tggacttcac	ttacagccgc	ctgagcaacc	tcctgggcat	14880
ccgcaagcgg	cagcccttcc	aggagggctt	caggatcacc	tacgaggacc	tggagggggg	14940
caacatcccc	gcgctcctcg	atgtggaggc	ctaccaggat	agcttgaagg	aaaatgaggc	15000
gggacaggag	gataaccgcc	ccgccgcctc	cgccgccgcc	gagcagggcg	aggatgctgc	15060
tgacaccgcg	gccgcggacg	gggaggaggc	cgaccccgct	atggtggtgg	aggctgccga	15120
gcaggaggag	gacatgaatg	acagtgcggg	gcgcggagac	accttcgtca	ccggggggga	15180
ggaaaagcaa	gcggaggccg	aggccgcggc	cgaggaaaag	caactggcgg	cagcagcggc	15240
ggcggcggcg	ttggccgcgg	cggaggctga	gtctgagggg	accaagcccg	ccaaggagcc	15300
cgtgattaag	cccctgaccg	aagatagcaa	gaagcgcagt	tacaacctgc	tcaaggacag	15360
caccaacacc	gcgtaccgca	gctggtacct	ggcctacaac	tacggcgacc	cgtcgacggg	15420

ggtgcgctcc	tggaccctgc	tgtgcacgcc	ggacgtgacc	tgcggctcgg	agcaggtgta	15480
ctggtcgcctg	cccgacatga	tgcaagaccc	cgtgaccttc	cgctccacgc	ggcaggtcag	15540
caacttcccc	gtggtgggcg	ccgagctgct	gcccgatgcac	tccaagagct	tctacaacga	15600
ccaggccgctc	tactcccagc	tcataccgcca	gttcacctct	ctgaccacgc	tgttcaatcg	15660
ctttcctgag	aaccagattc	tggcgcgccc	gcccgcctcc	accatcacca	ccgtcagtga	15720
aaacgttcct	gctctcacag	atcacgggac	gctaccgctg	cgcaacagca	tcggaggagt	15780
ccagcgagtg	accgttactg	acgccagacg	ccgcacctgc	ccctacgttt	acaaggcctt	15840
gggcatagtc	tcgccgcgcg	tcctttccag	ccgcactttt	tgagcaacac	caccatcatg	15900
tccatcctga	tctcacccag	caataactcc	ggctggggac	tgctgcgcgc	gcccagcaag	15960
atgttcggag	gggagaggaa	gcgttccgag	cagcaccccc	tgcgcgctgc	cgggcacttc	16020
cgcgccccct	ggggagcgca	caaacgcggc	cgcgcggggc	gcaccaccgt	ggacgacgcc	16080
atcgactcgg	tggtggagca	ggcgcgcaac	tacaggcccc	cggtctctac	cgtggacgcg	16140
gcatccaga	ccgtggtgcg	gggagcgcg	cggtacgcca	agctgaagag	ccgcccgaag	16200
cgcgaggccc	gcccgcaccg	ccgcccagcc	ggggccgccc	ccaaacgcgc	cgccgcccgc	16260
ctgcttcgcc	gggccaagcg	cacggggcgc	cgcgccgcca	tgagggcgcg	gcgcccgttg	16320
gcccgcggca	tcaccgccc	caccatggcc	ccccgtaccc	gaagacgcgc	ggccgcccgc	16380
gcccgcggcg	ccatcagtga	catggccagc	aggcgccggg	gcaacgtgta	ctgggtgcgc	16440
gactcgggtg	ccggcacgcg	cgtgcccgtg	cgcttcgcc	ccccgcggac	ttgagatgat	16500
gtgaaaaaac	aacactgagt	ctcctgctgt	tgtgtgtatc	ccagcggcgg	cgggcgcgcg	16560
cgagcgctca	tgtccaagcg	caaaatcaaa	gaagagatgc	tccaggctcg	cgcgccggag	16620
atctatgggc	ccccgaagaa	ggaagagcag	gattcgaagc	cccgcaagat	aaagcgggtc	16680
aaaaagaaaa	agaaagatga	tggcgatgcc	gatggggagg	tgaggttcct	gcgcgccacg	16740
gcgcccaggc	gcccgggtga	gtggaagggc	cggcgcgtaa	agcgcgctct	gcgccccggc	16800
accgcggtgg	tcttcacgcc	cggcgagcgc	tccaccggga	ctttcaagcg	cgtctatgac	16860
gaggtgtacg	gcgacgaaga	cctgctggag	caggccaacg	agcgcttcgg	agagtttgct	16920
tacgggaagc	gtcagcggcc	gctggggaag	gaggacctgc	tggcgctgcc	gctggaccag	16980
ggcaaccca	ccccagttct	gaagcccgtg	accctgcagc	aggtgctgcc	gagcagcgca	17040
ccctccgagg	cgaagcgggg	tctgaagcgc	gagggcgggc	acctggcgcc	caccgtgcag	17100
ctcatggtgc	ccaagcggca	gaggctggag	gatgtgctgg	agaaaatgaa	agtagacccc	17160
ggtctgcagc	cggacatcag	ggtccgtccc	atcaagcagg	tggcgccggg	cctcggcgtg	17220
cagaccgtgg	acgtgggtcat	ccccaccggc	aactcccccg	ccgccaccac	cactaccgct	17280
gcctccacgg	acatggagac	acagaccgat	ccgcgcgag	ccgcgcgcc	cgccgcgcgc	17340

gcgacctcct	cggcggaggt	gcagacggac	ccctggctgc	cgccggcgat	gtcagctccc	17400
cgcgcgctc	gcgggcgag	gaagtacggc	gccgccaacg	cgctcctgcc	cgagtacgcc	17460
ttgcatcctt	ccatcgcgcc	cacccccggc	taccgaggct	atacctaccg	cccgcgaaga	17520
gccaaggggt	ccaccgcgcg	tccccgccga	cgcgccgcgcg	ccaccaccgcg	ccgcgcgcgc	17580
cgcagacgcc	agcccgcaact	ggctccagtc	tccgtgagga	gagtggcgcg	cgacggacac	17640
accctggctgc	tgcccagggc	gcgctaccac	cccagcatcg	tttaaaagcc	tgttgtgggt	17700
cttgagata	tgccctcac	ttgccgcctc	cgtttcccg	tgccgggata	ccgaggagga	17760
agatcgcgcc	gcaggagggg	tctggccggc	cgcggcctga	gcggaggcag	ccgcgcgcgcg	17820
caccggcggc	gacgcgccac	cagccgacgc	atgcgcggcg	gggtgctgcc	cctgttaatc	17880
ccctgatcg	ccgcggcgat	cggcgccgtg	cccgggatcg	cctccgtggc	cttgaggcg	17940
tcccagaggc	attgacagac	ttgcaaactt	gcaaatatgg	aaaaaaacc	caataaaaaa	18000
gtctagactc	tcacgctcgc	ttggtcctgt	gactattttg	tagaatggaa	gacatcaact	18060
ttgcgtcgct	ggccccgcgt	cacggctcgc	gcccgctcct	gggacactgg	aacgatatcg	18120
gcaccagcaa	catgagcggg	ggcgccctca	gttggggctc	tctgtggagc	ggcattaaaa	18180
gtatcgggtc	tgccgttaaa	aattacggct	cccgggcctg	gaacagcagc	acggggccaga	18240
tgttgagaga	caagttgaaa	gagcagaact	tccagcagaa	ggtggtggag	ggcctggcct	18300
ccggcatcaa	cgggggtggtg	gacctggcca	accaggccgt	gcagaataag	atcaacagca	18360
gactggaccc	ccggccgcgcg	gtggaggagg	tgccgcgggc	gctggagacg	gtgtcccccg	18420
atgggcgtgg	cgagaagcgc	ccgcggccccg	ataggggaaga	gaccactctg	gtcacgcaga	18480
ccgatgagcc	gccccgtat	gaggaggccc	tgaagcaagg	tctgcccacc	acgcggccca	18540
tcgcgcccac	ggccaccggg	gtggtggggc	gccacacccc	cgccacgctg	gacttgccctc	18600
cgcccgccga	tgtgccgcag	cagcagcaga	aggcggcaca	gccggggcccg	cccgtgaccg	18660
cctcccgttc	ctccgcgggt	cctctgcgcc	gcgcggccag	cgccccccgc	gggggggtcg	18720
cgaggcacgg	caactggcag	agcacgctga	acagcatcgt	gggtctgggg	gtgcgggtccg	18780
tgaagcgccg	ccgatgctac	tgaatagctt	agctaacgtg	ttgtatgtgt	gtatgcgccc	18840
tatgtcgccg	ccagaggagc	tgctgagtcg	ccgccgttcg	cgcgcccacc	accaccaccg	18900
ccactccgcc	cctcaagatg	gcgaccccat	cgatgatgcc	gcagtggctg	tacatgcaca	18960
tctcggggcca	ggacgcctcg	gagtacctga	gccccgggct	ggtgcagttc	gcccgcgcca	19020
ccgagagcta	cttcagcctg	agtaacaagt	ttaggaaccc	cacggtggcg	cccacgcacg	19080
atgtgaccac	cgaccggtct	cagcgccctga	cgctgcgggt	cattcccgtg	gaccgcgagg	19140
acaccgcgta	ctcgtacaag	gcgcgggttca	ccctggccgt	gggcgacaac	cgcgtgctgg	19200
acatggcctc	cacctacttt	gacatccgcg	gggtgctgga	ccggggcccc	actttcaagc	19260

cttactctgg	caccgcctac	aactccctgg	cccccaaggg	cgctcccaac	tcctgcgagt	19320
gggagcaatt	agaagaagcc	caggccgctg	tggaagacga	agaattagaa	gatgaagacg	19380
aggaaccaca	ggatgaggca	cctgtgaaaa	aaacccatgt	atacgctcag	gctccccctt	19440
ctggagaaga	aattactaaa	aacggtttgc	aaataggggtc	agataacaca	gaagcccagt	19500
ctaagcccat	atatgcagat	cctacattcc	agcctgaacc	ccaaatcggg	gaatcccagt	19560
ggaatgaggc	agatgctaca	gttgccggcg	gtagagtgtc	aaagaaatcc	actcccatga	19620
agccatgcta	tggttcctat	gcaagacca	caaactccaa	tggagggtcaa	ggtgtgctgg	19680
tggtgatga	taaggggggt	cttcaatcta	aagttgaatt	gcaatTTTTT	tcaaatacta	19740
ctactcttaa	tcagcgggag	ggtaacgata	caaaacaaaa	agtgggtgctg	tatagcgaag	19800
atgtgcacat	ggaaactcca	gacaccaca	tttcttacia	gccacaaaa	agcgatgaca	19860
attcaaaaat	catgctgggt	cagcagtcca	tgcccaacag	acctaattac	atcggttca	19920
gagacaactt	tatcggcctc	atgtattaca	atagcactgg	caacatggga	gtgcttgcag	19980
gtcaggcctc	tcagttgaat	gcagtgggtg	acttgcaaga	cagaaacaca	gaactgtcct	20040
accagctctt	gcttgattcc	atgggtgaca	gaaccagata	cttttccatg	tggaatcagg	20100
cagtggacag	ttatgacca	gatgtcagaa	ttattgaaaa	tcatggaact	gaagacgagc	20160
tccccaacta	ttgtttccct	ctgggcgga	taggggtaac	tgacacttac	caggccatta	20220
aaaccaatgg	caatgggtcaa	gaaaacccaa	cctgggaaaa	agatacagag	tttgcagacc	20280
gcaatgaaat	aggggtggga	aacaatttcg	ctatggagat	caacctcagt	gccaacctgt	20340
ggagaaactt	cctgtactcc	aacgtggcgc	tgtacctgcc	agacaagctt	aagtacaacc	20400
cctccaatgt	ggacatctct	gacaacccca	acacctacga	ttacatgaac	aagcgagtgg	20460
tggccccggg	gctggtggac	tgctacatca	acctgggcgc	gcgctggctg	ctggactaca	20520
tggacaacgt	caacccttc	aaccaccacc	gcaatgcggg	cctgcgctac	cgctccatgc	20580
tcctgggcaa	cgggcgctac	gtgcccttcc	acatccaggt	gccccagaag	ttctttgcc	20640
tcaagaacct	cctcctcctg	ccgggctcct	acacctacga	gtggaacttc	aggaaggatg	20700
tcaacatggt	cctccagagc	tctctgggta	acgatctcag	ggtggacggg	gccagcatca	20760
agttcgagag	catctgcctc	tacgccacct	tcttccccat	ggcccacaac	acggcctcca	20820
cgctcgaggc	catgctcagg	aacgacacca	acgaccagtc	cttcaatgac	tacctctccg	20880
ccgccaacat	gctctacccc	ataccgcga	acgccaccaa	cgccccatc	tccatccct	20940
cgcgcaactg	ggcggccttc	cgcggtggg	ccttcacccg	cctcaagacc	aaggagaccc	21000
cctccctggg	ctcgggattc	gaccctact	acacctactc	gggctccatt	ccctacctgg	21060
acggcacctt	ctacctcaac	cacactttca	agaagggtctc	ggtcaccttc	gactcctcgg	21120
tcagctggcc	gggcaacgac	cgtctgctca	cccccaacga	gttcgaaatc	aagcgctcgg	21180

tcgacgggga	gggctacaac	gtggcccagt	gcaacatgac	caaggactgg	ttcctgggtcc	21240
agatgctggc	caactacaac	atcggtctacc	agggctttcta	catcccagag	agctacaagg	21300
acaggatgta	ctcctttcttc	aggaacttcc	agcccatgag	cgggcaggtg	gtggaccaga	21360
ccaagtacaa	ggactaccag	gaggtgggca	tcatccacca	gcacaacaac	tcgggcttcg	21420
tgggctacct	cgccccacc	atgcgcgagg	gacaggccta	ccccgccaac	ttcccctacc	21480
cgctcatagg	caagaccgcg	gtcgacagca	tcaccagaa	aaagtctctc	tgcgatcgca	21540
ccctctggcg	catcccccttc	tccagcaact	tcatgtccat	gggtgcgctc	tcggacctgg	21600
gccagaactt	gctctacgcc	aactccgccc	acgccctcga	catgaccttc	gaggtcgacc	21660
ccatggacga	gcccaccctt	ctctatgttc	tgttcgaagt	ctttgacgtg	gtccgggtcc	21720
accagccgca	ccgcggcgctc	atcgagaccg	tgtacctgcg	tacgcccttc	tcggccggca	21780
acgccaccac	ctaaagaagc	aagccgcagt	catcgccgcc	tgcatgccgt	cgggttccac	21840
cgagcaagag	ctcagggcca	tcgtcagaga	cctgggatgc	gggccctatt	ttttgggcac	21900
tttcgacaag	cgcttccctg	gctttgtctc	cccacacaag	ctggcctgcg	ccatcgtaaa	21960
cacggccggc	cgcgagaccg	ggggcggtgca	ctggctggcc	ttcgccctgga	accgcgctc	22020
caaaacatgc	ttcctctttg	accccttcgg	cttttcggac	cagcgggtca	agcaaatact	22080
cgagttcgag	tacgagggct	tgctgcgtcg	cagcgccatc	gcctcctcgc	ccgaccgctg	22140
cgtcaccttc	gaaaagtcca	cccagaccgt	gcagggggccc	gactcggccg	cctgcgggtct	22200
cttctgctgc	atgtttctgc	acgcctttgt	gcactggcct	cagagtccca	tggaaccgaa	22260
ccccaccatg	aacttgctga	cgggggtgcc	caactccatg	ctccagagcc	cccaggctga	22320
gcccaccctg	cgccgcaacc	aggagcagct	ctacagcttc	ctggagcgcc	actcgcccta	22380
cttccgccgc	cacagcgcac	agatcaggag	ggccacctcc	ttctgccact	tgcaagagat	22440
gcaagaaggg	taataacgat	gtacacactt	ttttctcaat	aaatggcatt	tttttattta	22500
tacaagctct	ctgggggtatt	catttccac	caccaccacc	acccgcggtt	gtcgccatct	22560
ggctctattt	agaaatcgaa	agggttctgc	cgggagtcgc	cgtgcgccac	gggcagggac	22620
acgttgcgat	actggtagcg	ggtgccccac	ttgaactcgg	gcaccaccag	gcgaggcagc	22680
tcggggaagt	tttcgctcca	caggctgcgg	gtcagcacca	gcgcgttcat	caggctgggc	22740
gccgagatct	tgaagtcgca	gttggggccg	ccgccctgcg	cgcgcgagtt	gcggtacacc	22800
gggttgcgac	actggaacac	caacagcgcc	gggtgcttca	cgctagccag	cacgctgcgg	22860
tcggagatca	gctcggcgctc	caggtcctcc	gcgttgctca	gcgcgaacgg	ggtcatcttg	22920
ggcacttgcc	tccccaggaa	gggcgcgtgc	cccggtttcg	agttgcagtc	gcagcgcagc	22980
gggatcagca	ggtgccccatg	cccggactcg	gcgttgggggt	acagcgcgcg	catgaaggcc	23040
tgcatctggc	ggaaggccat	ctgggccttg	gcgccctccg	agaagaacat	gccgcaggac	23100

ttgccccgaga	actgggtttgc	ggggcagctg	gcgtcgtgca	ggcagcagcg	cgcgctcggtg	23160
ttggcgatct	gcaccacgtt	gcgccccac	cggttcttca	cgatcttggc	cttggacgat	23220
tgctccttca	gcgcgcgctg	cccgttctcg	ctggtcacat	ccatctcgat	cacatgttcc	23280
ttgttcacca	tgctgctgcc	gtgcaggcac	ttcagctcgc	cctccgtctc	ggtgcagcgg	23340
tgctgccaca	gcgcgcagcc	cgtgggctcg	aaagacttgt	aggtcacctc	cgcgaaggac	23400
tgcaggtacc	cctgcaaaaa	gcggcccatc	atggtcacga	aggctcttgt	gctgctgaag	23460
gtcagctgca	gcccgcggtg	ctcctcgttc	agccaggctc	tgcacacggc	cgccagcgcc	23520
tccacctggg	cgggcagcat	cttgaagttc	accttcagct	cattctccac	gtgggtactg	23580
tccatcagcg	tgcgcgccgc	ctccatgccc	ttctcccagg	ccgacaccag	cggcaggctc	23640
acgggggttct	tcaccatcac	cgtggccgcc	gcctccgccg	cgctttcgct	ttccgccccg	23700
ctgtttctct	cctcttctc	ctcttctctg	ccgccgccca	ctcgcagccc	ccgcaccacg	23760
gggtcgtctt	cctgcaggcg	ctgcaccttg	cgcttgccgt	tgcgcccctg	cttgatgcgc	23820
acgggcgggt	tgctgaagcc	caccatcacc	agcgcgccct	cttcttgctc	gtcctcgctg	23880
tccagaatga	cctccgggga	ggggggggtg	gtcatcctca	gtaccgaggc	acgcttcttt	23940
ttcttctctg	gggcgttcgc	cagctccgcg	gctgcgggcg	ctgccgagg	cgaaggccga	24000
gggctgggcg	tgcgcgccac	cagcgcgctc	tgcgagccgt	cctcgtctc	ctcggactcg	24060
agacggaggc	gggcccgtt	cttcgggggc	gcgcggggcg	gcggaggcgg	cggcggcgac	24120
ggagacgggg	acgagacatc	gtccagggtg	ggtggacggc	gggccgcgcc	gcgtccgcgc	24180
tcgggggtgg	tctcgcgctg	gtcctcttcc	cgactggcca	tctcccactg	ctccttctcc	24240
tataggcaga	aagagatcat	ggagtctctc	atgcgagtcg	agaaggagga	ggacagccta	24300
accgccccct	ctgagccctc	caccaccgcc	gccaccaccg	ccaatgccgc	cgcggaacgac	24360
gcgcccaccg	agaccaccgc	cagtaccacc	nnnctcccca	gcgacgcacc	cccgtcgcag	24420
aatgaagtgc	tgatcgagca	ggacccgggt	tttgtgagcg	gagaggagga	tgagggtggat	24480
gagaaggaga	aggaggagg	cgccgcctca	gtgccaaaag	aggataaaaa	gcaagaccag	24540
gacgacgcag	ataaggatga	gacagcagtc	gggcggggga	acggaagcca	tgatgctgat	24600
gacggctacc	tagacgtggg	agacgacgtg	ctgcttaagc	acctgcaccg	ccagtgcgtc	24660
atcgtctgcg	acgcgctgca	ggagcgggtg	gaagtgcccc	tggacgtggc	ggagggtcagc	24720
cgcgctacg	agcggcacct	cttcgcgcgc	cacgtgcccc	ccaagcgccg	ggagaacggc	24780
acctgcgagc	ccaaccgcg	tctcaacttc	taccgggtct	tcgcggtacc	cgagggtgctg	24840
gccacctacc	acatcttttt	ccaaaactgc	aagatcccc	tctcctgccg	cgctaaccgc	24900
acccgcgcgc	acaaaaccct	gaccctgcgg	cagggcgccc	acataacctga	tattgcctct	24960
ctggaggaag	tgcccaagat	cttcgagggt	ctcggtcgcg	acgagaaacg	ggcggcggaac	25020

gctctgcacg	gagacagcga	aaacgagagt	cactcggggg	tgctggtgga	gctcgagggc	25080
gacaacgcgc	gcctggccgt	actcaagcgc	agcatagagg	tcacccactt	tgcctacccg	25140
gcgctcaacc	tgcccccaa	ggtcattgagt	gtgggtcatgg	gcgagctcat	catgcgccgc	25200
gccagcccc	tggccgcgga	tgcaaacttg	caagagtcct	cagaggaagg	cctgcccgcg	25260
gtcagcgacg	agcagctggc	gcgctggctg	gagaccgcgc	accccgcgca	gctggaggag	25320
cggcgcaagc	tcatgatggc	cgcggtgctg	gtcacctggg	agctcgagtg	tctgcagcgc	25380
ttcttcgcgg	accccagat	gcagcgcaag	ctcgaggaga	ccctgcacta	caccttcgcg	25440
cagggctacg	tgcgccaggc	ctgcaagatc	tccaacgtgg	agctctgcaa	cctggtctcc	25500
tacctgggca	tcctgcacga	gaaccgcctc	gggcagaacg	tcctgcactc	caccctcaaa	25560
ggggaggcgc	gccgcgacta	catccgcgac	tgcgccctacc	tcttcctctg	ctacacctgg	25620
cagacggcca	tgggggtctg	gcagcagtgc	ctggaggagc	gcaacctcaa	ggagctggaa	25680
aagctcctca	agcgcaccct	cagggacctc	tggacgggct	tcaacgagcg	ctcgggtggc	25740
gccgcgctgg	cggacatcat	cttccccgag	cgcttgctca	agaccctgca	gcagggcctg	25800
cccgaattca	ccagccagag	catgctgcag	aacttcagga	ctttcatcct	ggagcgctcg	25860
ggcatcctgc	cggccacttg	ctgcgcgctg	cccagcgact	tcgtgcccac	caagtacagg	25920
gagtggccgc	cgccgctctg	gggccactgc	tacctcttcc	agctggccaa	ctacctcgcc	25980
taccactcgg	acctcatgga	agacgtgagc	ggcgagggcc	tgctcgagtg	ccactgcgcg	26040
tgcaacctct	gcacgcccc	ccgctctcta	gtctgcaacc	cgcagctgct	cagcgagagt	26100
cagattatcg	gtaccttcga	gctgcagggg	ccctcgcttg	acgagaagtc	cgcggtcccg	26160
gggctgaaac	tactccggg	gctgtggact	tccgcctacc	tacgcaaatt	tgtacctgag	26220
gactaccacg	cccacgagat	caggttctac	gaagaccaat	cccggccgcc	caaggcggag	26280
ctcaccgcct	gcgtcatcac	ccaggggcac	atcctggggc	aattgcaagc	catcaacaaa	26340
gcccgcgag	agttcttgct	gaaaaagggg	cggggggtgt	acctggacce	ccagtcgggc	26400
gaggagctaa	acccgctacc	cccgcgcgcg	cccagcagc	gggaccttgc	ttcccaggat	26460
ggcaccaga	aagaagcagc	agccgcgcgc	gcagccatac	atgcttcttg	aggaagagga	26520
ggaggactgg	gacagtcagg	cagaggagat	gatggaagac	tgggaggagg	acagcagcct	26580
agacgaggaa	gcttcagagg	ccgaagaggt	ggcagacgca	acaccatcac	cctcggtcgc	26640
agccccctcg	ccggggcccc	tgaaatcctc	cgaaccagc	accagcgcta	taacctccgc	26700
tcctccggcg	ccggcgccac	ccgcccgcag	acccaaccgt	agatgggaca	ccacaggaac	26760
cggggtcggg	aagtccaagt	gcccgcgcgc	gccaccgcag	cagcagcagc	agcagcgcca	26820
gggctaccgc	tcgtggcgcg	ggcacaagaa	cgccatagtc	gcctgcttgc	aagactgcgg	26880
gggcaacatc	tctttcgccc	gccgcttcct	gctattccac	cacgggggtcg	cctttccccc	26940

caatgtcctg	cattactacc	gtcatctcta	cagccccctac	tgcagcggcg	accagagggc	27000
ggcagcggca	gccacagcgg	cgaccaccac	ctaggaagat	atcctccgcg	ggcaagacag	27060
cggcagcagc	ggccaggaga	cccgcggcag	cagcggcggg	agcgggtgggc	gcaactgcgc	27120
tctcgcccaa	cgaaccctc	tcgacccg	agctcagaca	caggatcttc	cccactttgt	27180
atgccatctt	ccaacagagc	agaggccagg	agcaggagct	gaaaataaaa	aacagatctc	27240
tgcgctccct	caccgcagc	tgtctgtatc	acaaaagcga	agatcagctt	ggcgcgacgc	27300
tggaggacgc	ggaggcactc	ttcagcaa	actgcgcgct	cactcttaaa	gactagctcc	27360
gcgcccttct	cgaatttagg	cgggagaaaa	ctacgtcatc	gccggccgc	gccagccccg	27420
cccagccgag	atgagcaaag	agattccac	gccatacatg	tggagctacc	agccgcagat	27480
gggactcgcg	gcgggagcgg	cccaggacta	ctccaccgc	atgaactaca	tgagcgcggg	27540
accccatatg	atctcacagg	tcaacgggat	ccgcgcccag	cgaaaccaaa	tactgctgga	27600
acaggcggcc	atcaccgcca	cgccccgcca	taatctcaac	ccccgaaatt	ggcccgcgc	27660
cctagtgtac	caggaaacc	cctccgccac	caccgtacta	cttccgcgtg	acgcccaggc	27720
cgaagtccag	atgactaact	caggggcgca	gctcgcgggc	ggctttcgtc	acggggcgcg	27780
gccgctccga	ccaggtataa	gacacctgat	gatcagaggc	cgaggtatcc	agctcaacga	27840
cgagtcggtg	agctcttcgc	tcggtctccg	tccggacgga	actttccagc	tcgccggatc	27900
cggtcgctct	tcgttcacgc	cccgccaggc	gtacctgact	ctgcagacct	cgtectcgga	27960
gccccgctcc	ggcggcatcg	gaaccctcca	gttcgtggag	gagttcgtgc	cctcgggtcta	28020
cttcaacccc	ttctcgggac	ctcccggacg	ctaccccgac	cagttcatte	cgaactttga	28080
cgcggtgaag	gactcggcgg	acggctacga	ctgaatgtca	ggtgccgagg	cagagcagct	28140
tcgcctgaga	cacctcgagc	actgccgcg	ccacaagtgc	ttcgcccgcg	gttcgggtga	28200
gttctgctac	tttcagctac	ccgaggagca	taccgagggg	ccggcgcacg	gcgtccgcct	28260
gaccaccag	ggcgaggtta	cctgttcct	catccgggag	ttcacctcc	gtcccctgct	28320
agtggagcgg	gagcggggtc	cctgtgtcct	aactatcgcc	tgcaactgcc	ctaaccctgg	28380
attacatcaa	gatctttgct	gtcatctctg	tgctgagttt	aataaacgct	gagatcagaa	28440
tctactgggg	ctcctgtcgc	catcctgtga	acgccaccgt	cttcaccac	cccgaccagg	28500
cccaggcgaa	cctcacctgc	ggtctgcata	ggagggtcaa	gaagtacctc	acctggtact	28560
tcaacggcac	cccctttgtg	gtttacaaca	gcttcgacgg	ggacggagtc	tcctgaaag	28620
accagctctc	cggctctcagc	tactccatcc	acaagaacac	cacctccaa	ctcttcctc	28680
cctacctgcc	gggaacctac	gagtgcgtca	ccggccgctg	caccacctc	accgcctga	28740
tcgtaaacca	gagctttccg	ggaacagata	actccctctt	ccccagaaca	ggaggtgagc	28800
tcaggaaact	ccccggggac	cagggcggag	acgtaccttc	gacccttg	gggttaggat	28860

tttttattac	cggggttgctg	gctcttttta	tcaaagcttc	cttgagattt	gttcttttct	28920
tctacgtgta	tgaacacctc	agcctccaat	aactctaccc	tttcttcggg	atcaggtgac	28980
ttttctgaaa	tcgggcttgg	tgtgctgctt	actctgttga	tttttttctt	tatcatactc	29040
agccttctgt	gcctcaggct	cgccgcctgc	tgcgcacaca	tctatatcta	ctgctgggtg	29100
ctcaagtgca	ggggtcgcca	cccaagatga	acaggtacat	ggtcctatcg	atcctaggcc	29160
tgctggccct	ggcggcctgc	agcgccgcca	aaaaagagat	tacctttgag	gagcccgtt	29220
gcaatgtaac	tttcaagccc	gaggggtgacc	aatgcaccac	cctcgtcaaa	tgcgttacca	29280
atcatgagaa	gctgcgcatc	gactacaaaa	acaaaactgg	cgggtttgcg	gtctatagt	29340
tgtttacgcc	cggagacccc	tctaactact	ctgtcacctg	cttcaggggc	ggacagtcta	29400
agatattcaa	ttacactttc	cctttttatg	agttgtgcga	tgcggtcatg	tacatgtcaa	29460
aacagtacaa	cctgtggcct	ccctctcccc	aggcgtgtgt	ggaaaatact	gggtcttact	29520
gctgtatggc	tttggcaatc	actacgctcg	ctctaactct	cacgggtgcta	tatataaaat	29580
tcaggcagag	gcgaatcttt	atcgatgaaa	agaaaatgcc	ttgatcgcta	acaccggctt	29640
tctatctgca	gaatgaatgc	aatcacctcc	ctactaatca	ccaccaccct	ccttgcgatt	29700
gcccattgggt	tgacacgaat	cgaagtgcc	gtgggggtcca	atgtcaccat	ggtgggcccc	29760
gccggcaatt	ccaccctcat	gtgggaaaaa	tttgtccgca	atcaatgggt	tcattttctgc	29820
tctaaccgaa	tcagtatcaa	gcccagagcc	atctgcgatg	ggcaaaatct	aactctgatc	29880
aatgtgcaaa	tgatggatgc	tgggtactat	tacgggcagc	ggggagaaaat	cattaattac	29940
tggcgacccc	acaaggacta	catgctgcat	gtagtcgagg	cacttcccac	taccaccccc	30000
actaccacct	ctcccaccac	cactaccact	actactacta	ctactaccac	taccgctgcc	30060
cgccataccc	gcaaaagcac	catgattagc	acaaagcccc	ctcgtgctca	ctcccacgcc	30120
ggcgggcccc	tcggtgcgac	ctcagaaacc	accgagcttt	gcttctgcca	atgcactaac	30180
gccagcgctc	atgaactggt	cgacctggag	aatgaggatg	cccagcagag	ctccgcttgc	30240
ctgaccaggg	aggctgtgga	gcccgttgcc	ctgaagcaga	tcggtgattc	aataattgac	30300
tcttcttctt	ttgccactcc	cgaataccct	cccgattcta	ctttccacat	cacgggtacc	30360
aaagacccta	acctctcttt	ctacctgatg	ctgctgctct	gtatctctgt	ggtctcttcc	30420
gcgctgatgt	tactggggat	gttctgctgc	ctgatctgcc	gcagaaagag	aaaagctcgc	30480
tctcagggcc	aaccactgat	gcccttcccc	tacccccggg	attttgcaga	taacaagata	30540
tgagctcgct	gctgacacta	accgctttac	tagcctgcgc	tctaaccctt	gtcgcttgcg	30600
actcgagatt	ccacaatgtc	acagctgtgg	caggagaaaa	tgttactttc	aactccacgg	30660
ccgataccca	gtggtcgtgg	agtggctcag	gtagctactt	aactatctgc	aatagctcca	30720
cttccccccag	catatcccca	accaagtacc	aatgcaatgc	cagcctgttc	accctcatca	30780

acgcttccac	cctggacaat	ggactctatg	taggctatgt	accctttggt	gggcaaggaa	30840
agaccacgc	ttacaacctg	gaagttcgcc	agcccagaac	cactacccaa	gctwcymcca	30900
ycaccagcac	cagcagcagc	agccacagca	gcagcagcag	attattgact	ttgggttttgg	30960
ccagctcatc	tgccgctacc	caggccatct	acagctctgt	gcccgaacc	actcagaccc	31020
accgcccaga	aacgaccacc	gccaccaccc	tacacacctc	cagcgatcag	atgccgacca	31080
acatcacccc	cttggctctt	caaattgggac	ttacaagccc	cactccaaaa	ccagtggatg	31140
cggccgaggt	ctccgccctc	gtcaatgact	gggcggggct	gggaatgtgg	tggttcgccca	31200
taggcatgat	ggcgctctgc	ctgcttctgc	tctggctcat	ctgctgcctc	caccgcaggc	31260
gagccagacc	ccccatctat	agacccatca	ttgtcctgaa	ccccgataat	gatgggatcc	31320
atagattgga	tggcctgaaa	aacctacttt	tttcttttac	agtatgataa	attgagacat	31380
gcctcgcatt	ttcttgtaca	tgttccttct	cccacctttt	ctgggggtgtt	ctacgctggc	31440
cgctgtgtct	cacctggagg	tagactgcct	ctcacccttc	actgtctacc	tgctttacgg	31500
attggtcacc	ctcactctca	tctgcagcct	aatcacagta	atcatcgcct	tcattccagt	31560
cattgattac	atctgtgtgc	gcctcgcata	cttcagacac	caccgcagct	accgagacag	31620
gaacattgcc	caacttctaa	gactgctcta	atcatgcata	agactgtgat	ctgccttctg	31680
atcctctgca	tctgcccac	cctcacctcc	tgccagtaca	ccacaaaatc	tccgcgcaaa	31740
agacatgcct	cctgccgctt	cacccaactg	tggaatatac	ccaaatgcta	caacgaaaag	31800
agcgagctct	ccgaagcttg	gctgtatggg	gtcatctgtg	tcttagtttt	ctgcagcact	31860
gtctttgccc	tcattgatcta	cccctacttt	gatttgggat	ggaacgcgat	cgatgccatg	31920
aattacccca	cctttcccg	acccgagata	attccactgc	gacaagttgt	accggttgtc	31980
gttaatcaac	gcccccatc	ccctacgccc	actgaaatca	gctactttta	cctaacaggc	32040
ggagatgact	gacgccctag	atctagaaat	ggacggcatc	agtaccgagc	agcgtctcct	32100
agagaggcgc	aggcaggcgg	ctgagcaaga	gcgcctcaat	caggagctcc	gagatctcgt	32160
taacctgcac	cagtgcacaaa	gaggcatctt	ttgtctggta	aagcaggcca	aagtcaccta	32220
cgagaagacc	ggcaacagcc	accgcctcag	ttacaaattg	cccaccacgc	gccagaagct	32280
ggtgctcatg	gtgggtgaga	atcccatcac	cgtcaccacg	cactcggtag	agaccgaggg	32340
gtgtctgcac	tccccctgtc	ggggtccaga	agacctctgc	accctggtaa	agaccctgtg	32400
cggctctcaga	gatttagtcc	cctttaacta	atcaaact	ggaatcaata	aaaagaatca	32460
cttacttaaa	atcagacagc	aggtctctgt	ccagtttatt	cagcagcacc	tccttccct	32520
cctcccaact	ctgggtactcc	aaacgccttc	tggcggcaaa	cttcctccac	accctgaagg	32580
gaatgtcaga	ttcttgcctc	tgtccctccg	caccactat	cttcattgtg	ttgcagatga	32640
agcgcaccaa	aacgtctgac	gagagcttca	acccgtgtga	cccctatgac	acggaaagcg	32700

gccctccctc	cgcccccttc	ctcacccctc	ccttcgtgtc	tcccgatgga	ttccaagaaa	32760
gtcccccccg	ggtcctgtct	ctgaacctgg	ccgagccctt	ggtcacttcc	cacggcatgc	32820
tcgccctgaa	aatgggaagt	ggcctctccc	tggacgacgc	tggcaacctc	acctctcaag	32880
atatcaccac	cgctagccct	cccccaaaa	aaaccaagac	caacctcagc	ctagaaacct	32940
catccccctt	aactgtgagc	acctcaggcg	ccctcacctg	agcagccgcc	gctccccctg	33000
cgggtggccg	cacctccctc	accatgcaat	cagaggcccc	cctgacagta	caggatgcaa	33060
aactcacctt	ggccacaaaa	ggccccctga	ccgtgtctga	aggcaaactg	gccttgcaaa	33120
catcggtccc	gctgacggcc	gctgacagca	gcacctcac	cgttagcgcc	acaccaccaa	33180
ttaatgtaag	cagtgggaagt	ttaggcttag	acatggaaga	ccctatgtat	actcacaatg	33240
gaaaactggg	aataagaatt	gggggtccac	taagagtagt	agacagcttg	catacactca	33300
ctgtagttac	cggaaatgga	ctaactgtag	ataacaatgc	cctccaaact	aaagttacgg	33360
gcgccctagg	ttatgacaca	tcaggaaatc	tacaattaag	agctgcagga	ggtatgcgaa	33420
ttgacgcaaa	tggccaactt	atccttaatg	tggcataccc	atttgatgct	cagaacaatc	33480
tcagcccttag	acttggtcag	ggacccctgt	atataaacac	agaccacaac	ctggatttga	33540
attgcaacag	aggtctaacc	acaactacca	ccaacaacac	aaaaaaactt	gagactaaaa	33600
ttagctcagg	cttagactat	gacaccaatg	gtgctgtcat	tattaaactt	ggcactggtc	33660
taagcttcga	caacacaggc	gccctaactg	tgggaaacac	tggatgatgat	aaactgactc	33720
tgtggacgac	cccagaccca	tctccaaatt	gcagaattca	ctcagacaaa	gactgcaagt	33780
ttactctagt	cctaactaag	tgtggaagcc	aaatcctggc	ctctgtcgcc	gccctagcgg	33840
tatcaggaaa	tctggcttcg	ataacaggca	ccgttgccag	cgttaccatc	tttctcagat	33900
ttgatcagaa	tggagtgctt	atggaaaact	cctcgctaga	caggcagtac	tggaaacttca	33960
gaaatggcaa	ctcaactaac	gctgccccct	acaccaatgc	agttggggtc	atgccaaacc	34020
tcgcagcata	ccccaaaacg	caaagccaga	ctgctaaaaa	caacattgta	agtcagggtt	34080
acttgaatgg	agacaaatcc	aaacccatga	cccttaccat	cacctcaat	ggaactaatg	34140
aatccagtga	aactagccag	gtgagtcact	actccatgtc	atttacatgg	gcttgggaaa	34200
gtgggcaata	tgccactgaa	acctttgcca	ccaactcctt	caccttttct	tacattgctg	34260
aacaataaaa	agcatgacac	tgatgttcat	ttctgattct	tattttatta	ttttcaaaca	34320
caacaaaatc	attcaagtca	ttcttccatc	ttagcttaat	agacacagta	gcttaataga	34380
cccagtagtg	caaagcccca	ttctagctta	taactagtgg	agaagtactc	gcctacatgg	34440
gggtagagtc	ataatcgtgc	atcaggatag	ggcggtggtg	ctgcagcagc	gcgcgaataa	34500
actgctgccg	ccgccgctcc	gtcctgcagg	aataacaacat	ggcagtggtc	tcctcagcga	34560
tgattcgcac	cgcccgagc	ataaggcgcc	ttgtcctccg	ggcacagcag	cgcacctga	34620

tctcacttaa	atcagcacag	taactgcagc	acagcaccac	aatattgttc	aaaatcccac	34680
agtgcaaggc	gctgtatcca	aagctcatgg	cggggaccac	agaaccacg	tggccatcat	34740
accacaagcg	caggtagatt	aagtggcgac	ccctcataaa	cacgctggac	ataaacatta	34800
cctcttttgg	catgttgtaa	ttcaccacct	cccgtacca	tataaacctc	tgattaaaca	34860
tggcgccatc	caccaccatc	ctaaaccagc	tggccaaaac	ctgcccgcg	gctatacact	34920
gcagggaacc	gggactggaa	caatgacagt	ggagagccca	ggactcgtaa	ccatggatca	34980
tcatgctcgt	catgatatca	atgttggcac	aacacaggca	cacgtgcata	cacttctctca	35040
ggattacaag	ctcctcccg	gttagaacca	tatcccaggg	aacaacccat	tcctgaatca	35100
gcgtaaatcc	cacactgcag	ggaagacctc	gcacgtaact	cacgttgtgc	attgtcaaag	35160
tgttacattc	gggcagcagc	ggatgatcct	ccagtatggg	agcgcggggt	tctgtctcaa	35220
aaggaggtag	acgatcccta	ctgtacggag	tgcgccgaga	caaccgagat	cgtgttggtc	35280
gtagtgtcat	gccaaatgga	acgccggacg	tagtcatatt	tcctgaagtc	ttagatctct	35340
caacgcagca	ccagcaccaa	cacttcgcag	tgtaaaaggc	caagtgccga	gagagtatat	35400
ataggaataa	aaagtgcagt	aaacgggcaa	agtccaaaaa	acgccagaa	aaaccgcacg	35460
cgaacctacg	ccccgaaacg	aaagccaaaa	aacactagac	actcccttcc	ggcgtcaact	35520
tccgctttcc	cacgctacgt	cacttgcccc	agtcaaacaa	actacatatc	ccgaacttcc	35580
aagtcgccac	gccccaaaaca	ccgcctacac	ctccccgccc	gccggccccg	ccccaaaccc	35640
gcctccccgc	ccgcgccccg	ccccgcgcgc	cccatctcat	tatcatattg	gcttcaatcc	35700
aaaataaggt	atattattga	tgatggttta	aacggatcca	attcttgaag	acgaaagggc	35760
ctcgtgatac	gcctattttt	ataggttaat	gtcatgataa	taatggtttc	ttagacgtca	35820
ggtggcactt	ttcggggaaa	tgtgcgcgga	accctatttt	gtttattttt	ctaaatacat	35880
tcaaatatgt	atccgctcat	gagacaataa	ccctgataaa	tgcttcaata	atattgaaaa	35940
aggaagagta	tgagtattca	acatttccgt	gtcgccctta	ttcccttttt	tgcggcattt	36000
tgccttctctg	tttttgtctca	cccagaaacg	ctggtgaaag	taaaagatgc	tgaagatcag	36060
ttgggtgcac	gagtgggtta	catcgaactg	gatctcaaca	gcggtaagat	ccttgagagt	36120
tttcgccccg	aagaacgttt	tccaatgatg	agcactttta	aagttctgct	atgtggcgcg	36180
gtattatccc	gtgttgacgc	cgggcaagag	caactcggtc	gccgcataca	ctattctcag	36240
aatgacttgg	ttgagtactc	accagtcaca	gaaaagcatc	ttacggatgg	catgacagta	36300
agagaattat	gcagtgtgc	cataaccatg	agtgataaca	ctgcggccaa	cttacttctg	36360
acaacgatcg	gaggaccgaa	ggagctaacc	gcttttttgc	acaacatggg	ggatcatgta	36420
actcgccttg	atcgttggga	accggagctg	aatgaagcca	taccaaacga	cgagcgtgac	36480
accacgatgc	ctgtagcaat	ggcaacaacg	ttgcgcaaac	tattaactgg	cgaactactt	36540

actctagctt	cccggcaaca	attaatagac	tggatggagg	cggataaagt	tgcaggacca	36600
cttctgcgct	cggcccttcc	ggctggctgg	tttattgctg	ataaatctgg	agccggtgag	36660
cgtgggtctc	gcggtatcat	tgcagcactg	gggccagatg	gtaagccctc	ccgtatcgta	36720
gttatctaca	cgacggggag	tcaggcaact	atggatgaac	gaaatagaca	gatcgctgag	36780
ataggtgcct	cactgattaa	gcattggtaa	ctgtcagacc	aagtttactc	atatatactt	36840
tagattgatt	taaaaggatc	taggtgaaga	tcctttttga	taatctcatg	acccaaatcc	36900
cttaacgtga	gttttcgttc	cactgagcgt	cagaccccg	agaaaagatc	aaaggatctt	36960
cttgagatcc	tttttttctg	cgcgtaatct	gctgcttgca	aacaaaaaaaa	ccaccgctac	37020
cagcggtggt	ttgtttgccg	gatcaagagc	taccaactct	ttttccgaag	gtaactggct	37080
tcagcagagc	gcagatacca	aatactgtcc	ttctagtgtg	gccgtagtta	ggccaccact	37140
tcaagaactc	tgtagcaccg	cctacatacc	tcgctctgct	aatcctgtta	ccagtggctg	37200
ctgccagtgg	cgataagtcg	tgtcttaccg	ggttggactc	aagacgatag	ttaccggata	37260
aggcgcagcg	gtcgggctga	acgggggggt	cgtgcacaca	gccagcttg	gagcgaacga	37320
cctacaccga	actgagatac	ctacagcgtg	agctatgaga	aagcgccacg	cttcccgaag	37380
ggagaaaggc	ggacaggtat	ccggtaagcg	gcagggtcgg	aacaggagag	cgcacgaggg	37440
agcttccagg	gggaaacgcc	tggtatcttt	atagtcctgt	cgggtttcgc	cacctctgac	37500
ttgagcgtcg	atTTTTgtga	tgctcgtcag	gggggcgag	cctatggaaa	aacgccagca	37560
acgcggcctt	tttacggttc	ctggcctttt	gctggccttg	aagctgtccc	tgatggctgt	37620
catctacctg	cctggacagc	atggcctgca	acgcgggcat	cccgatgccg	ccggaagcga	37680
gaagaatcat	aatggggaag	gccatccagc	ctcgcgtcgc	agatccgaat	tcgtttaaac	37740

<210> 23
 <211> 1572
 <212> DNA
 <213> Rabies virus

<400> 23		
atggtgcctc	aagccctgct	gttcgtgccc
ctgctggtct	tctccctctg	ctttggcaag
ttccccatct	acaccatccc	tgacaagctc
ggccccctgg	ccccattga	catacatcac
ctcagctgcc	ccaacaacct	ggtggtggag
gatgagggct	gcacaaacct	gagcggcttc
tcctatatgg	aactcaaggt	gggctatata
tccgccatca	aggtcaatgg	attcacatgc
accggcgtcg	tgacagaggg	tgaaacatac
accaactttg	tgggctacgt	caccaccaca
ttcaagagga	agcacttcag	gccacccct
gacgcttgca	gggctgccta	caattggaag
atggctggcg	acccaggtg	tgaggagtcc
ctgcacaatc	cctaccccg	ctaccattgg
ctcaggacag	tcaagaccac	caaggagtcc
ctggctatta	tctcccctag	cgtggccgac
ctagacccgt	atgacaaaag	cctgcactcc
agggctcttc	ctagcggcaa	atgctccggc

60
 120
 180
 240
 300
 360
 420
 480
 540

attacagtga gctccaccta ctgcagcaca aaccacgact acaccatctg gatgcctgag	600
aatcctaggc tcggcacctc ctgtgacata tttaaaata gcaggggcaa gagggcttcc	660
aaaggcagca aaacctgcgg ctttgtcgac gaaagaggcc tgtacaagtc cctcaagggc	720
gcttgtaaac tcaagctgtg cggagtgtg ggactcagac tcatggacgg cacatgggtg	780
gccatgcaga ccagcgatga gaccaagtgg tgcccccccg atcagctggt gaatctgcac	840
gacttcaggt ccgacgaaat tgagcacctc gtgggtcgagg agctggtgaa gaagagagaa	900
gagtgcctgg atgctctgga gtccatcatg accaccaaat ccgtgtcctt cagaaggctg	960
agccacctca ggaagctggt ccccggttt ggcaaggcct acacaatttt caataagaca	1020
ctgatggagg ccgatgtca ctacaaatcc gtgaggacct ggaacgagat catccccctc	1080
aaaggctgcc tgagggtggg aggaagatgc cccccccag tcaacggcgt cttcttcaac	1140
ggcattatcc tcggaccgga tggccatgtc ctgatccctg aaatgcaaag ctccctgctg	1200
cagcagcaca tggaactcct ggagagctcc gtcatcccc tgatgcaccc tctcgctgac	1260
cccagcaccg tgtttaaaga cggcgatgag gccgaggact tcgtggaagt gcatctgcct	1320
gatgtgcata agcaagtcag cggcgctgat ctgggcctgc ctaattgggg caagtatgtc	1380
ctgctctccg ccggagctct gattgccctg atgctgatca tcttctgat gacctgctgc	1440
agaagagtca acagacctga gagcacccaa agatccctcg gcggaaccgg aaggaaggtc	1500
agcgtgacca gccagtccgg caaagtgatt tcctcctggg agagctataa aagcggcgga	1560
gagaccaggc tg	1572

<210> 24
 <211> 37536
 <212> DNA
 <213> Pan troglodytes

<220>
 <221> misc_feature
 <222> (24187)..(24189)
 <223> n iöääñòàâëÿåð ñíáíé a, c, g, èèè t

<400> 24	
catcatcaat aatatacctt attttggatt gaagccaata tgataatgag atgggcggcg	60
cggggcgggg cgcggggcgg gaggcgggtt tgggggcggg ccggcgggcg gggcggtgtg	120
gcggaagtgg actttgtaag tgtggcggat gtgacttgct agtgccgggc gcggtaaaag	180
tgacgttttc cgtgcgcgac aacgccccg ggaagtgaca tttttccgc ggtttttacc	240
ggatgttgta gtgaatttgg gcgtaaccaa gtaagatttg gccattttcg cgggaaaact	300
gaaacgggga agtgaaatct gattaatttt gcgttagtca taccgcgtaa tatttgtcta	360
gggccgaggg actttggccg attacgtgga ggactcgccc aggtgttttt tgagggtgaat	420

ttccgcgttc	cgggtcaaag	tctgcgtttt	attattatag	gatatcccat	tgcatacgtt	480
gtatccatat	cataatatgt	acattttatat	tggctcatgt	ccaacattac	cgccatgttg	540
acattgatta	ttgactagtt	attaatagta	atcaattacg	gggtcattag	ttcatagccc	600
atatatggag	ttccgcgtta	cataacttac	ggtaaattggc	ccgcctggct	gaccgcccac	660
cgacccccgc	ccattgacgt	caataatgac	gtatgttccc	atagtaacgc	caatagggac	720
tttccattga	cgtcaatggg	tggagtattt	acggtaaact	gcccacttgg	cagtacatca	780
agtgtatcat	atgccaaagta	cgccccctat	tgacgtcaat	gacggtaaata	ggccccgctg	840
gcattatgcc	cagtacatga	ccttatggga	ctttcctact	tggcagtaca	tctacgtatt	900
agtcacgcct	attaccatgg	tgatgcgggt	ttggcagtac	atcaatgggc	gtggatagcg	960
gtttgactca	cggggatttc	caagtctcca	ccccattgac	gtcaatggga	gtttgttttg	1020
gaacccaaat	caacgggact	ttccaaaatg	tcgtaacaac	tccgccccat	tgacgcaaata	1080
gggcggtagg	cgtgtacggg	gggaggtcta	tataagcaga	gctctcccta	tcagtgatag	1140
agatctccct	atcagtgata	gagatcgtcg	acgagctcgt	ttagtgaacc	gtcagatcgc	1200
ctggagacgc	catccacgct	gttttgacct	ccatagaaga	caccgggacc	gatccagcct	1260
ccgcgggccg	gaacgggtga	ttggaacgcg	gattccccgt	gccaagagtg	agatcttccg	1320
tttatctagg	taccagatat	cgccaccatg	gtgcctcaag	ccctgctgtt	cgtgccccctg	1380
ctgggtcttct	ccctctgctt	tggcaagttc	cccatctaca	ccatccctga	caagctcggc	1440
ccctgggtccc	ccattgacat	acatcacctc	agctgcccc	acaacctggg	gggtggaggat	1500
gagggctgca	caaacctgag	cggcttctcc	tatatggaac	tcaaggtggg	ctatatctcc	1560
gccatcaagg	tcaatggatt	cacatgcacc	ggcgtcgtga	cagaggctga	aacatacacc	1620
aactttgtgg	gctacgtcac	caccacattc	aagaggaagc	acttcaggcc	caccctgac	1680
gcttgacagg	ctgcctacaa	ttggaagatg	gctggcgacc	ccaggatatga	ggagtccctg	1740
cacaatccct	accccgacta	ccattggctc	aggacagtca	agaccaccaa	ggagtccctg	1800
gtcattatct	cccctagcgt	ggccgacct	gaccgctatg	acaaaagcct	gactccagg	1860
gtcttcccta	gcggcaaagt	ctccggcatt	acagtgaagt	ccacctactg	cagcacaac	1920
cacgactaca	ccatctggat	gcctgagaat	cctaggctcg	gcacctcctg	tgacatat	1980
acaaatagca	ggggcaagag	ggcttccaaa	ggcagcaaaa	cctgcggctt	tgctgacgaa	2040
agaggcctgt	acaagtccct	caagggcgct	tgtaaactca	agctgtgcgg	agtgtggga	2100
ctcagactca	tggacggcac	atgggtggcc	atgcagacca	gcgatgagac	caagtgggtgc	2160
cccccgatc	agctggtgaa	tctgcacgac	ttcaggctccg	acgaaattga	gcacctcgtg	2220
gtcgaggagc	tgggtgaagaa	gagagaagag	tgcttggtatg	ctctggagtc	catcatgacc	2280
accaaattccg	tgctcttcag	aaggctgagc	cacctcagga	agctgggtccc	cggctttggc	2340

aaggcctaca	caatthttcaa	taagacactg	atggaggccg	atgctcacta	caaatccgtg	2400
aggacctgga	acgagatcat	ccccccaaa	ggctgcctga	gggtgggagg	aagatgccac	2460
ccccacgtca	acggcgtctt	cttcaacggc	attatcctcg	gacccgatgg	ccatgtcctg	2520
atccctgaaa	tgcaaagctc	cctgctgcag	cagcacatgg	aactcctgga	gagctccgtc	2580
atccccctga	tgacacctct	cgctgacccc	agcaccgtgt	ttaaagacgg	cgatgaggcc	2640
gaggacttcg	tggaagtgca	tctgcctgat	gtgcataagc	aagtcagcgg	cgtcgatctg	2700
ggcctgccta	attggggcaa	gtatgtcctg	ctctccgccg	gagctctgat	tgccctgatg	2760
ctgatcatct	tcctgatgac	ctgctgcaga	agagtcaaca	gacctgagag	cacccaaaga	2820
tcctcggcgg	gaaccggaag	gaaggctcagc	gtgaccagcc	agtcgggcaa	agtgatttcc	2880
tcctgggaga	gctataaaaag	cgggcgagag	accaggctgt	gatgagcggc	cgcgatctgc	2940
tgtgccttct	agttgccagc	catctgttgt	ttgccccctc	cccgtgcctt	ccttgacctt	3000
ggaaggtgcc	actccactg	tcctttccta	ataaaatgag	gaaattgcat	cgcattgtct	3060
gagtaggtgt	cattctattc	tgggggggtg	gggtggggcag	gacagcaagg	gggaggattg	3120
ggaagacaat	agcaggcatg	ctgggggatgc	gggtgggctct	atggccgatc	agcgatcgct	3180
gaggtgggtg	agtgggcgtg	gcctgggggtg	gtcatgaaaa	tatataagtt	gggggtctta	3240
gggtctcttt	atttgtgttg	cagagaccgc	cggagccatg	agcgggagca	gcagcagcag	3300
cagtagcagc	agcgccttgg	atggcagcat	cgtgagccct	tatttgacga	cgcggatgcc	3360
ccactgggcc	ggggtgcgtc	agaatgtgat	gggctccagc	atcgacggcc	gacccgtcct	3420
gcccgcaa	atccgccagc	tgacctatgc	gaccgtcgcg	gggacgccgt	tggaagccac	3480
cgccgccgcc	gccgccaccg	cagccgcctc	ggccgtgcgc	agcctggcca	cggactttgc	3540
attcctggga	ccactggcga	caggggctac	ttctcgggcc	gctgctgccg	ccgttcgcga	3600
tgacaagctg	accgccctgc	tggcgcagtt	ggatgcgctt	actcgggaac	tgggtgacct	3660
ttctcagcag	gtcatggccc	tgcgccagca	ggtctcctcc	ctgcaagctg	gcgggaatgc	3720
ttctcccaca	aatgccgttt	aagataaata	aaaccagact	ctgtttggat	taaagaaaag	3780
tagcaagtgc	attgctctct	ttatttcata	atthttccgcg	cgcgataggc	cctagaccag	3840
cgthctcgg	cgthgagggt	gcgggtgtatc	ttctccagga	cgtggtagag	gtggctctgg	3900
acgttgagat	acatgggcat	gagcccgctc	cgggggtgga	ggtagcacca	ctgcagagct	3960
tcatgctccg	gggtgggtgt	gtagatgatc	cagtcgtagc	aggagcgtg	ggcatgggtc	4020
ctaaaaatgt	ccttcagcag	caggccgatg	gccaggggga	ggcccttggt	gtaagtgttt	4080
acaaaacgg	taagttggga	aggggtgcatt	cggggagaga	tgatgtgcat	cttggactgt	4140
atthtttagat	tggcgatgtt	tccgccca	tcctthctgg	gattcatgtt	gtgcaggacc	4200
accagtacag	tgtatccggt	gcacttgggg	aattthgtcat	gcagcttaga	gggaaaagcg	4260

tggaagaact	tggagacgcc	cttgtggcct	cccagatddd	ccatgcattc	gtccatgatg	4320
atggcaatgg	gcccgcggga	ggcagcttgg	gcaaagatat	ttctggggtc	gctgacgtcg	4380
tagttgtgtt	ccaggggtgag	gtcgtcatag	gccatdddta	caaagcgcg	gcggaggggtg	4440
cccgactggg	ggatgatgg	cccctctggc	cccggggcgt	agttgccctc	gcagatctgc	4500
atttcccagg	ccttaatctc	ggagggggga	atcatatcca	cctgcggggc	gatgaagaaa	4560
acggtttccg	gagccgggga	gattaactgg	gatgagagca	ggtttctaag	cagctgtgat	4620
tttccacaac	cgggtggggcc	ataaataaca	cctataaccg	gttgacagctg	gtagttttaga	4680
gagctgcagc	tgccgtcgtc	ccggaggagg	ggggccacct	cgttgagcat	gtccctgacg	4740
cgcatgttct	ccccgaccag	atccgccaga	aggcgctcgc	cgccaggga	cagcagctct	4800
tgcaaggaag	caaagttttt	cagcggcctt	aggccgtccg	cgtgggcat	gtttttcagg	4860
gtctggctca	gcagctccag	gcgggtcccag	agctcgggtga	cgtgctctac	ggcatctcta	4920
tccagcatat	ctctcgttt	cgcggttgg	ggcgactttc	gctgtagggc	accaagcgg	4980
ggtcgtccag	cggggccaga	gtcatgtcct	tccatgggcg	cagggtcctc	gtcaggggtg	5040
tctgggtcac	ggtgaagggg	tgcgctccgg	gctgagcgct	tgccaagggtg	cgcttgaggc	5100
tggttctgct	ggtgctgaag	cgctgccgg	cttcgccctg	cgcgtcggcc	aggtagcatt	5160
tgaccatgg	gtcatagtcc	agccctccg	cgcggtgtcc	cttggcgcg	agcttgccct	5220
tggaggtggc	gccgcacgag	gggcagagca	ggctcttgag	cgcgtagagc	ttgggggcga	5280
ggaagaccga	ttcgggggag	taggcgtccg	cgccgcagac	cccgcacacg	gtctcgcaact	5340
ccaccagcca	ggtgagctcg	gggcgcgccg	ggtcaaaaac	caggtttccc	ccatgctttt	5400
tgatgcgttt	cttacctcgg	gtctccatga	ggtggtgtcc	ccgctcgggtg	acgaagaggc	5460
tgtccgtgtc	tccgtagacc	gacttgagg	gtcttttctc	caggggggtc	cctcgggtctt	5520
cctcgtagag	gaactcggac	cactctgaga	cgaaggcccg	cgtccaggcc	aggacgaagg	5580
aggctatgtg	ggaggggtag	cggtcgttgt	ccactagggg	gtccaccttc	tccaagggtg	5640
gaagacacat	gtcgccttcc	tggcggtcca	ggaagggtgat	tggcttgtag	gtgtaggcca	5700
cgtgaccggg	ggttctctgac	gggggggtat	aaaaggggg	gggggcgcgc	tcgtcgtcac	5760
tctcttccgc	atcgctgtct	gcgagggcc	gctgctgggg	tgagtattcc	ctctcgaagg	5820
cgggcatgac	ctccgcgctg	aggttgtcag	tttccaaaa	cgaggaggat	ttgatgttca	5880
cctgtcccga	ggtgatacct	ttgagggtac	ccgcgtccat	ctggtcagaa	aacacgatct	5940
ttttattgtc	cagcttggtg	gcgaacgacc	cgtagagggc	gttgagagc	agcttggcga	6000
tggagcgcag	ggtctggttc	ttgtccctgt	cggcgcgctc	cttggccgcg	atgttgagct	6060
gcacgtactc	gcgcgcgacg	cagcgccact	cggggaagac	ggtggtgcgc	tcgtcgggca	6120
ccaggcgcac	gcgccagccg	cggttgtgca	gggtgaccag	gtccacgctg	gtggcgacct	6180

cgccgcgcag	gcgctcgttg	gtccagcaga	gacggccgcc	cttgcgcgag	cagaaggggg	6240
gcaggggggtc	gagctgggtc	tcgtccgggg	ggtccgcgtc	cacggtgaaa	accccggggc	6300
gcaggcgcg	gtcgaagtag	tctatcttgc	aaccttgcat	gtccagcgcc	tgctgccagt	6360
cgcgggcggc	gagcgcgcg	tcgtaggggt	tgagcggcgg	gccccagggc	atgggggtggg	6420
tgagtgcgga	ggcgtacatg	ccgcagatgt	catagacgta	gaggggctcc	cgcaggaccc	6480
cgatgtaggt	ggggtagcag	cggccgccgc	ggatgctggc	gcgcacgtag	tcatacagct	6540
cgtgcgaggg	ggcgaggagg	tcggggccca	ggttgggtgcg	ggcggggcgc	tccgtgcgga	6600
agacgatctg	cctgaagatg	gcatgcgagt	tggaagagat	ggtggggcgc	tggaagacgt	6660
tgaagctggc	gtcctgcagg	ccgacggcgt	cgcgcacgaa	ggaggcgtag	gagtcgcgca	6720
gcttgtgtac	cagctcggcg	gtgacctgca	cgtcgagcgc	gcagtagtcg	agggctctcg	6780
ggatgatgtc	atatttagcc	tgcccccttct	ttttccacag	ctcgcggttg	aggacaaact	6840
cttcgcggtc	tttccagtag	tcttggtatcg	ggaaaccgtc	cggttccgaa	cggtaagagc	6900
ctagcatgta	gaactgggtg	acggcctggg	aggcgcagca	gcccttctcc	acggggaggg	6960
cgtaggcctg	cgcggccttg	cggagcgagg	tgtgggtcag	ggcgaagggtg	tccctgacca	7020
tgactttgag	gtactggtgc	ttgaagtcgg	agtcgtcgca	gccgccccgc	tcccagagcg	7080
agaagtcggg	gcgcttcttg	gagcgggggt	tgggcagagc	gaagggtgaca	tcgttgaaga	7140
ggattttgcc	cgcgcggggc	atgaagttgc	gggtgatgcg	gaagggcccc	ggcacttcag	7200
agcggttgtt	gatgacctgg	gcggcgagca	cgatctcgtc	gaagccggtg	atgttgtggc	7260
ccacgatgta	gagttccagg	aagcggggcc	ggccctttac	ggtgggcagc	ttcttttagct	7320
cttcgtaggt	gagctcctcg	ggcgaggcga	ggccgtgctc	ggccagggcc	cagtccgcga	7380
ggtgcgggtt	gtctctgagg	aaggactccc	agaggtcgcg	ggccaggagg	gtctgcaggc	7440
ggtccctgaa	ggtcctgaac	tggcggccca	cggccatttt	ttcgggggtg	atgcagtaga	7500
aggtgagggg	gtcttgctgc	cagcgggtccc	agtcgagctg	cagggcgagg	tcgcgcgcgg	7560
cggtgaccag	gcgctcgtcg	ccccgaatt	tcatgaccag	catgaagggc	acgagctgct	7620
ttccgaaggc	ccccatccaa	gtgtaggtct	ctacatcgta	ggtgacaaag	aggcgctccg	7680
tgcgaggatg	cgagccgatc	gggaagaact	ggatctcccg	ccaccagttg	gaggagtggc	7740
tgttgatgtg	gtggaagtag	aagtcccgtc	gccgggcca	acactcgtgc	tggcttttgt	7800
aaaagcgagc	gcagtactgg	cagcgtgca	cgggctgtac	ctcatgcacg	agatgcacct	7860
ttcgcccgcg	cacgaggaag	ccgaggggaa	atctgagccc	cccgcctggc	tcgcggcatg	7920
gctgggtgctc	ttctactttg	gatgcgtgtc	cgtctccgtc	tggctcctcg	aggggtgtta	7980
cggtgagcgc	gaccaccacg	ccgcgcgagc	cgcagggtcca	gatatcggcg	cgcggcggtc	8040
ggagtttgat	gacgacatcg	cgcagctggg	agctgtccat	ggtctggagc	tcccgcggcg	8100

gcggcaggtc agccgggagt tcttgcaggt tcacctcgca gagtcggggc agggcgcggg	8160
gcaggtctag gtggtacctg atctctaggg gcgtgttggt ggcggcgctg atggcttgca	8220
ggagcccgca gcccggggg gcgacgacgg tgccccgcgg ggtggtggtg gtggtggcgg	8280
tgcagctcag aagcgggtgcc gcggggcgggc ccccgaggt agggggggct ccgggtccgc	8340
gggcaggggc ggcagcgga cgtcggcgtg gagcgcgggc aggagttggt gctgtgccg	8400
gaggttgctg gcgaaggcga cgacgcggcg gttgatctcc tggatctggc gcctctgcgt	8460
gaagacgacg ggcccggtag gcttgaacct gaaagagagt tcgacagaat caatctcggt	8520
gtcattgacc gcggcctggc gcaggatctc ctgcacgtct cccgagttgt cttggtaggc	8580
gatctcggcc atgaactgct cgatctcttc ctctggagg tctccgcgtc cggcgcgttc	8640
cacggtggcc gccaggtcgt tggagatgcg ccccatgagc tgcgagaagg cgttgagtcc	8700
gccctcgttc cagactcggc tgtagaccac gccccctgg tcatcgcggg cgcgcatgac	8760
cacctgcgcg aggttgagct ccacgtgccg cgcaagacg gcgtagttgc gcagacgctg	8820
gaagaggtag ttgagggtag tggcggtagt ctcgccacg aagaagttca tgaccagcg	8880
gcgcaacgtg gattcgttga tgtcccccac ggcctccagc cgttccatgg cctcgtagaa	8940
gtccacggcg aagttgaaaa actgggagtt gcgcgccgac acggtcaact cctcctccag	9000
aagacggatg agctcggcga cgggtgcgcg cacctcgcgc tcgaaggcta tggggatctc	9060
ttcctccgct agcatcacca cctcctctc ttcctcctct tctggcactt ccatgatggc	9120
ttcctcctct tcggggggtag gcggcgggcg cggtagggga gggggcgctc tgcgcggcg	9180
gcggcgacc gggaggcggc ccacgaagcg gcgatcatc tccccgcggc ggcggcgcat	9240
ggtctcggtag acggcgcggc cgttctcccg ggggcgcagt tggaagacgc cgccggacat	9300
ctggtgctgg ggcgggtggc cgtgaggcag cgagacggcg ctgacgatgc atctcaacaa	9360
ttgctgcgta ggtacgccgc cgagggacct gagggagtcc atatccaccg gatccgaaaa	9420
cctttcgagg aaggcgtcta accagtcgca gtcgcaagg aggtgagca ccgtggcggg	9480
cggcgggggg tggggggagt gtctggcgga ggtgctgctg atgatgtaat tgaagtaggc	9540
ggacttgaca cggcggtagg tcgacaggag caccatgtcc ttgggtccgg cctgctggat	9600
gcggaggcgg tcggctatgc cccaggcttc gttctggcat cggcgcaggc ccttgtagta	9660
gtcttgcatg agcctttcca ccggcacctc ttctccttc tcttctgctt cttccatgtc	9720
tgcttcggcc ctggggcggc gccgcgcccc cctgcccccc atgcgcgtga cccgaaccc	9780
cctgagcggg tggagcaggg ccaggtcggc gacgacgcgc tcggccagga tggcctgctg	9840
cacctgcgtg aggggtggtt ggaagtcac caagtccacg aagcggtggt aggcgcccgt	9900
gttgatggtg taggtgcagt tggccatgac ggaccagttg acggtctggt ggcccggttg	9960
cgacatctcg gtgtacctga gtcgcgagta ggcgcgggag tcgaagacgt agtcgttgca	10020

agtccgcacc	aggtactggt	agcccaccag	gaagtgcggc	ggcggctggc	ggtagagggg	10080
ccagcgcagg	gtggcggggg	ctccgggggc	caggtcttcc	agcatgaggc	ggtggtaggc	10140
gtagatgtac	ctggacatcc	aggtgatacc	cgcggcggtg	gtggaggcgc	gcgggaagtc	10200
gcgcaccccg	ttccagatgt	tgcgcagggg	cagaaagtgc	tccatggtag	gcgtgctctg	10260
tccagtcaga	cgcgcgcagt	cgttgatact	ctagaccagg	gaaaacgaaa	gccggtcagc	10320
gggcactctt	ccgtggtctg	gtgaatagat	cgcaagggtg	tcattggcgg	gggcctcggg	10380
tcgagccccg	ggtccggggc	ggacggtccg	ccatgatcca	cgcggttacc	gcccgcgtgt	10440
cgaaccagg	tgtgcgacgt	cagacaacgg	tggagtgttc	cttttggcgt	ttttctggcc	10500
gggcgcgggc	gccgcgtaag	agactaagcc	gcgaaagcga	aagcagtaag	tggctcgctc	10560
cccgtagccg	gagggatcct	tgctaagggt	tgcgttgccg	cgaaccccg	ttcgaatccc	10620
gtactcgggc	cggccggacc	cgcggtctag	gtgttggtg	ggcctcccc	tcgtataaag	10680
accccgcttg	cggattgact	ccggacacgg	ggacgagccc	cttttatctt	tgctttcccc	10740
agatgcatcc	ggtgttgcca	cagatgcgcc	ccccgcccc	gcagcagcaa	caacaccagc	10800
aagagcggca	gcaacagcag	cgggagtcag	gcagggcccc	ctcaccacc	ctcggcggcc	10860
cggccacctc	ggcgtccgcg	gccgtgtctg	gcgcctgcgg	cggcggcgcc	ggggggcgcc	10920
ctgacgacct	cgaggagccc	ccgcggcgca	gggccagaca	ctacctggac	ctggaggagg	10980
gcgagggcct	ggcgcggcgt	ggggcgccgt	ctcccagcgc	ccaccgcggg	gtgcagctaa	11040
agcgcgactc	gcgcgaggcg	tacgtgcctc	ggcagaacct	gttcaggggc	cgcgcgggcg	11100
aggagcccga	ggagatgcgg	gacaggaggt	tcagcgcggg	gcgggagctg	cggcaggggc	11160
tgaaccgcga	gcggctgctg	cgcgaggagg	actttgagcc	cgacgcgcgg	acggggatca	11220
gccccgcgcg	cgcgcacgtg	gcggccgcgg	acctggtgac	ggcgtacgag	cagacggtga	11280
accaggagat	caacttccaa	aagagtttca	acaaccacgt	gcgcacgctg	gtggcgcgcg	11340
aggaggtgac	catcgggctg	atgcacctgt	gggactttgt	gagcgcgctg	gtgcagaacc	11400
ccaatagcaa	gcctctgacg	gcgcagctgt	tcctgatagt	gcagcacagc	agggacaacg	11460
aggcgtttag	ggacgcgctg	ctgaacatca	ccgagcccga	gggccgggtg	ctgctggacc	11520
tgattaacat	cctgcagagc	atagtgggtg	aggagcgcag	cctgagcctg	gccgacaagg	11580
tggcggccat	caactactcg	atgctgagcc	tgggcaagtt	ttacgcgcgc	aagatctacc	11640
agacgccgta	cgtgcccata	gacaaggagg	tgaagatcga	cggttttttac	atgcgcattg	11700
cgctgaaggt	gctcacccta	agcgacgacc	tgggcgtgta	ccgcaacgag	cgcatccaca	11760
aggccgtgag	cgtgagccgg	cggcgcgagc	tgagcgaccg	cgagctgatg	catagcctgc	11820
agcgggcgct	ggcgggcgcc	ggcagcgggc	acagggaggc	ggagtcctac	ttcgatgcgg	11880
gggcggacct	gcgctggggc	cccagccggc	gggccctgga	ggccgcgggg	gtccgcgagg	11940

actatgacga	ggacggcgag	gaggatgagg	agtacgagct	agaggagggc	gagtacctgg	12000
actaaaccgc	gggtggtggt	tccggtagat	gcaagacccg	aacgtggtgg	acccggcgct	12060
gcgggcggt	ctgcagagcc	agccgtccgg	ccttaactcc	tcagacgact	ggcgacaggt	12120
catggaccgc	atcatgtcgc	tgacggcgcg	taaccgggac	gcgttccggc	agcagccgca	12180
ggccaacagg	ctctccgcc	tcctggaggc	ggtggtgcct	gcgcgctcga	acccacgcga	12240
cgagaaggtg	ctggccatag	tgaacgcgct	ggccgagaac	agggccatcc	gcccggacga	12300
ggccggggtg	gtgtacgacg	cgctgctgca	gcgcgtggcc	cgctacaaca	gcggcaacgt	12360
gcagaccaac	ctggaccggc	tggtggggga	cgtgcgcgag	gcggtggcgc	agcgcgagcg	12420
cgcggatcgg	cagggcaacc	tgggctccat	ggtggcgctg	aatgccttcc	tgagcacgca	12480
gccggccaac	gtgccgcggg	ggcaggaaga	ctacaccaac	tttgtgagcg	cgctgcggct	12540
gatggtgacc	gagaccccc	agagcgaggt	gtaccagtcg	ggtccggact	acttcttcca	12600
gaccagcaga	cagggcctgc	agacggtgaa	cctgagccag	gctttcaaga	acctgcgggg	12660
gctgtggggc	gtgaaggcgc	ccaccggcga	ccgggcgacg	gtgtccagcc	tgctgacgcc	12720
caactcgcg	ctgctgctgc	tgctgatcgc	gccgttcacg	gacagcggca	gcgtgtcccg	12780
ggacacctac	ctggggcacc	tgctgaccct	gtaccgcgag	gccatcgggc	aggcgcaggt	12840
ggacgagcac	accttccagg	agatcaccag	cgttagccgc	gcgctggggc	aggaggacac	12900
gagcagcctg	gaggcgactc	tgaactacct	gctgaccaac	cggcggcaga	agattccctc	12960
gctgcacagc	ctgacctccg	aggaggagcg	catcttgcg	tacgtgcagc	agagcgtgag	13020
cctgaacctg	atgcgcgacg	gggtgacgcc	cagtgtggcg	ctggacatga	ccgcgcgcaa	13080
catggaaccg	ggcatgtacg	ccgcgcaccg	gccttacatc	aaccgcctga	tggactacct	13140
gcatcgcgcg	gcggccgtga	accccgagta	ctttaccaac	gccatcctga	acccgcactg	13200
gctcccgccg	cccgggttct	acagcggggg	cttcgaggtc	ccggaggcca	acgatggctt	13260
cctgtgggac	gacatggacg	acagcgtggt	ctccccgcgg	ccgcaggcgc	tggcggaagc	13320
gtccctgctg	cgtcccaaga	aggaggagga	ggcgagtcgc	cgccgcggca	gcagcggcgt	13380
ggcttctctg	tccgagctgg	gggcggcagc	cgccgcgcgc	cccgggtccc	tgggcggcag	13440
cccctttccg	agcctggtgg	ggtctctgca	cagcgagcgc	accacccgcc	ctcggctgct	13500
gggcgaggac	gagtacctga	ataactccct	gctgcagccg	gtgcgggaga	aaaacctgcc	13560
tcccgcttcc	cccaacaacg	ggatagagag	cctggtggac	aagatgagca	gatggaagac	13620
ctatgcgcag	gagcacaggg	acgcgcccgc	gctccggccg	cccacgcggc	gccagcgcca	13680
cgaccggcag	cgggggctgg	tgtgggatga	cgaggactcc	gcggacgata	gcagcgtgct	13740
ggacctggga	gggagcggca	acccgttcgc	gcacctgcgc	ccccgcctgg	ggaggatggt	13800
ttaaaaaaaaa	aaaaaagcaa	gaagcatgat	gcaaaaatta	aataaaaactc	accaaggcca	13860

tggcgaccga	gcgttggttt	cttgtgttcc	cttcagtatg	cggcgcgcg	cgatgtacca	13920
ggagggacct	cctccctctt	acgagagcgt	ggtgggcgcg	gggcgggcg	cgccctcttc	13980
tccctttgcg	tcgcagctgc	tggagccgcc	gtacgtgcct	ccgcgctacc	tgcggcctac	14040
gggggggaga	aacagcatcc	gttactcgga	gctggcgccc	ctgttcgaca	ccaccgggt	14100
gtacctggtg	gacaacaagt	cggcggacgt	ggcctccctg	aactaccaga	acgaccacag	14160
caattttttg	accacggtca	tccagaacaa	tgactacagc	ccgagcgagg	ccagcaccca	14220
gaccatcaat	ctggatgacc	ggtcgcactg	ggcgggcgac	ctgaaaacca	tcctgcacac	14280
caacatgccc	aacgtgaacg	agttcatgtt	caccaataag	ttcaaggcgc	gggtgatggt	14340
gtcgcgctcg	cacaccaagg	aagaccgggt	ggagctgaag	tacgagtggg	tggagttcga	14400
gctgccagag	ggcaactact	ccgagaccat	gaccattgac	ctgatgaaca	acgcgatcgt	14460
ggagcactat	ctgaaagtgg	gcaggcagaa	cggggtcctg	gagagcgaca	tcgggggtcaa	14520
gttcgacacc	aggaacttcc	gcctggggct	ggaccccgctg	accgggctgg	ttatgcccg	14580
ggtgtacacc	aacgaggcct	tccatcccga	catcatcctg	ctgcccggct	gcgggggtgga	14640
cttcacttac	agccgcctga	gcaacctcct	gggcatccgc	aagcggcagc	ccttccagga	14700
gggcttcagg	atcacctacg	aggacctgga	ggggggcaac	atccccgcgc	tcctcgatgt	14760
ggaggcctac	caggatagct	tgaaggaaaa	tgaggcgga	caggaggata	ccgccccgc	14820
cgcctccgcc	gccgccgagc	agggcgagga	tgctgctgac	accgcggccg	cggacggggc	14880
ggaggccgac	cccgtatgg	tggtggaggc	tgccgagcag	gaggaggaca	tgaatgacag	14940
tgcggtgcgc	ggagacacct	tcgtcacccg	gggggaggaa	aagcaagcgg	aggccgaggc	15000
cgcggccgag	gaaaagcaac	tggcggcagc	agcggcggcg	gcggcgttgg	ccgcggcgga	15060
ggctgagtct	gaggggacca	agcccgccaa	ggagcccgtg	attaagcccc	tgaccgaaga	15120
tagcaagaag	cgcagttaca	acctgctcaa	ggacagcacc	aacaccgcgt	accgcagctg	15180
gtacctggcc	tacaactacg	gcgaccgctc	gacgggggtg	cgtccttgga	ccctgctgtg	15240
cacgccggac	gtgacctgcg	gctcggagca	ggtgtactgg	tcgctgcccg	acatgatgca	15300
agaccccgctg	accttccgct	ccacgcggca	ggtcagcaac	ttcccgggtg	tgggcgccga	15360
gctgctgccc	gtgcactcca	agagcttcta	caacgaccag	gccgtctact	cccagctcat	15420
ccgccagttc	acctctctga	cccacgtgtt	caatcgcttt	cctgagaacc	agattctggc	15480
gcgccccgcc	gccccacca	tcaccaccgt	cagtgaaaac	gttcctgctc	tcacagatca	15540
cgggacgcta	ccgctgcgca	acagcatcgg	aggagtccag	cgagtgaccg	ttactgacgc	15600
cagacgccgc	acctgcccct	acgtttacaa	ggccttgggc	atagtctcgc	cgcgcgtcct	15660
ttccagccgc	actttttgag	caacaccacc	atcatgtcca	tcctgatctc	accagcaat	15720
aactccggct	ggggactgct	gcgcgcgccc	agcaagatgt	tcggaggggc	gaggaagcgt	15780

tccgagcagc	accccgtgcg	cgtgcgcggg	cacttccgcg	ccccctgggg	agcgcacaaa	15840
cgcgcccgcg	cggggcgcac	caccgtggac	gacgccatcg	actcggtggt	ggagcaggcg	15900
cgcaactaca	ggcccgcggt	ctctaccgtg	gacgcggcca	tccagaccgt	ggtgcggggc	15960
gcgcggcggt	acgccaagct	gaagagccgc	cggaagcgcg	tggcccgcgc	ccaccgcgc	16020
cgacccgggg	ccgccgcaa	acgcgccgcc	gcggccctgc	ttcgccgggc	caagcgcacg	16080
ggccgcccgc	ccgccatgag	ggccgcgcgc	cgcttgccgc	ccggcatcac	cgccgccacc	16140
atggccccc	gtacccgaag	acgcgcggcc	gccgccgcgc	ccgccgccat	cagtgcacatg	16200
gccagcaggc	gccggggcaa	cgtgtactgg	gtgcgcgact	cggtgaccgg	cacgcgcgtg	16260
cccgtgcgct	tccgcccccc	gcggacttga	gatgatgtga	aaaaacaaca	ctgagtctcc	16320
tgctgtttgt	tgtatcccag	cgggggcggc	ggcgcgcgca	gcgtcatgtc	caagcgcaaa	16380
atcaaagaag	agatgctcca	ggtcgtcgcg	ccggagatct	atgggcccc	gaagaaggaa	16440
gagcaggatt	cgaagccccg	caagataaag	cggttcaaaa	agaaaaagaa	agatgatggc	16500
gatgccgatg	gggaggtgga	gttcctgcgc	gccacggcgc	ccaggcgccc	ggtgcagtgg	16560
aagggccggc	gcgtaaagcg	cgtcctgcgc	cccggcaccg	cggtggtctt	cacgcccggc	16620
gagcgctcca	cccggacttt	caagcgcgtc	tatgacgagg	tgtacggcga	cgaagacctg	16680
ctggagcagg	ccaacgagcg	cttcggagag	tttgcttacg	ggaagcgtca	gcggccgctg	16740
gggaaggagg	acctgctggc	gctgccgctg	gaccagggca	acccaccccc	cagtctgaag	16800
cccgtgacct	tgcagcaggt	gctgccgagc	agcgcaccct	ccgaggcgaa	gcggggctctg	16860
aagcgcgagg	gcggcgacct	ggcgcccacc	gtgcagctca	tggtgcccaa	gcggcagagg	16920
ctggaggatg	tgctggagaa	aatgaaagta	gaccccggtc	tgcagccgga	catcagggtc	16980
cgtcccatca	agcaggtggc	gccgggcctc	ggcgtgcaga	ccgtggacgt	ggtcatcccc	17040
accggcaact	ccccgcgcgc	caccaccact	accgctgcct	ccacggacat	ggagacacag	17100
accgatcccc	ccgcagccgc	cgccaccgcc	gccgcgcgca	cctcctcggc	ggaggtgcag	17160
acggacccct	ggctgccgcc	ggcgatgtca	gctccccgcg	cgcgtcgcgg	gcgcaggaag	17220
tacggcgccg	ccaacgcgct	cctgcccagag	tacgccttgc	atccttccat	cgcgcccacc	17280
cccggctacc	gaggctatac	ctaccgcccc	cgaagagcca	agggttccac	ccgcggtccc	17340
cgccgacgcg	ccgccgccac	cacccgccgc	cgccgccgca	gacgccagcc	cgcactggct	17400
ccagtctccg	tgaggagagt	ggcgcgcgac	ggacacaccc	tggtgctgcc	cagggcgcgc	17460
taccaccca	gcatcgttta	aaagcctgtt	gtggttcttg	cagatatggc	cctcacttgc	17520
cgctccggtt	tcccggtgcc	gggataccga	ggaggaagat	cgcgccgcag	gaggggtctg	17580
gccggccgcg	gcctgagcgg	aggcagccgc	cgcgcgccacc	ggcggcgacg	cgccaccagc	17640
cgacgcatgc	gcggcggggt	gctgcccctg	ttaatcccc	tgatcgccgc	ggcgatcggc	17700

gccgtgcccc	ggatcgccctc	cgtggccttg	caggcgctccc	agaggcattg	acagacttgc	17760
aaacttgcaa	atatggaaaa	aaacccaat	aaaaaagtct	agactctcac	gctcgcttgg	17820
tcctgtgact	atlttgtaga	atggaagaca	tcaactttgc	gtcgctggcc	ccgcgtcacg	17880
gctcgcgccc	gttcctggga	cactggaacg	atatcggcac	cagcaacatg	agcgggtggcg	17940
ccttcagttg	gggctctctg	tggagcggca	ttaaaagtat	cgggtctgcc	gttaaaaatt	18000
acggctcccc	ggcctggaac	agcagcacgg	gccagatggt	gagagacaag	ttgaaagagc	18060
agaacttcca	gcagaaggtg	gtggagggcc	tggcctccgg	catcaacggg	gtggtggacc	18120
tggccaacca	ggccgtgcag	aataagatca	acagcagact	ggacccccgg	ccgccggtgg	18180
aggaggtgcc	gccggcgctg	gagacggtgt	ccccgatgg	gcgtggcgag	aagcgcccg	18240
ggcccgatag	ggaagagacc	actctggtca	cgcagaccga	tgagccgccc	ccgtatgagg	18300
aggccctgaa	gcaaggtctg	cccaccacgc	ggcccatcgc	gcccatggcc	accgggggtg	18360
tgggccgcca	cacccccgcc	acgttgact	tgcctccgcc	cgccgatgtg	ccgcagcagc	18420
agcagaaggc	ggcacagccg	ggcccgcccc	tgaccgcctc	ccgttcctcc	gccggtcctc	18480
tgcgccgcgc	ggccagcggc	ccccgcgggg	gggtcgcgag	gcacggcaac	tggcagagca	18540
cgctgaacag	catcgtgggt	ctgggggtgc	ggtccgtgaa	gcgcgcgcca	tgctactgaa	18600
tagcttagct	aacgtgttgt	atgtgtgtat	gcgccctatg	tcgccgccag	aggagctgct	18660
gagtcgccc	cgttcgcgcg	cccaccacca	ccaccgccac	tccgccccctc	aagatggcga	18720
ccccatcgat	gatgccgcag	tggtcgtaca	tgcacatctc	gggccaggac	gcctcggagt	18780
acctgagccc	cggtctggtg	cagttcgccc	gcgccaccga	gagctacttc	agcctgagta	18840
acaagtttag	gaacccacg	gtggcgccca	cgcacgatgt	gaccaccgac	cggtctcagc	18900
gcctgacgct	gcggttcatt	cccgtggacc	gcgaggacac	cgcgtactcg	tacaaggcgc	18960
ggttcaccct	ggcgtgggc	gacaaccgcg	tgctggacat	ggcctccacc	tactttgaca	19020
tccgcgggg	gctggaccgg	ggccccactt	tcaagcetta	ctctggcacc	gcctacaact	19080
ccctggcccc	caagggcgct	cccaactcct	gcgagtggga	gcaattagaa	gaagcccagg	19140
ccgctgtgga	agacgaagaa	ttagaagatg	aagacgagga	accacaggat	gaggcacctg	19200
tgaaaaaaac	ccatgtatac	gctcaggctc	ccctttctgg	agaagaaatt	actaaaaacg	19260
gtttgcaa	agggtcagat	aacacagaag	cccagtctaa	gcccatatat	gcagatccta	19320
cattccagcc	tgaaccccaa	atcggggaat	cccagtggaa	tgaggcagat	gctacagttg	19380
ccggcggtag	agtgtctaa	aatccactc	ccatgaagcc	atgctatggt	tcctatgcaa	19440
gaccacaaa	ctccaatgga	ggtcaagggt	tgctgggtgg	tgatgataag	ggggttcttc	19500
aatctaaagt	tgaattgcaa	tttttttcaa	atactactac	tcttaatcag	cgggagggta	19560
acgatacaaa	accaaagtgt	gtgctgtata	gcgaagatgt	gcacatggaa	actccagaca	19620

cccacatttc	ttacaagccc	acaaaaagcg	atgacaattc	aaaaatcatg	ctggggtcagc	19680
agtccatgcc	caacagacct	aattacatcg	gcttcagaga	caactttatc	ggcctcatgt	19740
attacaatag	cactggcaac	atgggagtg	ttgcagggtca	ggcctctcag	ttgaatgcag	19800
tgggtggactt	gcaagacaga	aacacagaac	tgctctacca	gctcttgctt	gattccatgg	19860
gtgacagaac	cagatacttt	tccatgtgga	atcaggcagt	ggacagtatt	gaccagatg	19920
tcagaattat	tgaaaatcat	ggaactgaag	acgagctccc	caactattgt	ttccctctgg	19980
gcgggcatagg	ggtaactgac	acttaccagg	ccattaaaaac	caatggcaat	ggtcaagaaa	20040
acccaacctg	ggaaaaagat	acagagtttg	cagaccgcaa	tgaaataggg	gtgggaaaca	20100
atttcgctat	ggagatcaac	ctcagtgcc	acctgtggag	aaacttcctg	tactccaacg	20160
tggcgctgta	cctgccagac	aagcttaagt	acaaccctc	caatgtggac	atctctgaca	20220
acccaacac	ctacgattac	atgaacaagc	gagtgggtggc	cccggggctg	gtggactgct	20280
acatcaacct	gggcgcgcgc	tggtcgctgg	actacatgga	caacgtcaac	cccttcaacc	20340
accaccgcaa	tgcgggcctg	cgctaccgct	ccatgctcct	gggcaacggg	cgctacgtgc	20400
ccttccacat	ccagggtgcc	cagaagttct	ttgccatcaa	gaacctctc	ctcctgccgg	20460
gctctacac	ctacgagtgg	aacttcagga	aggatgtcaa	catgggtctc	cagagctctc	20520
tgggtaacga	tctcaggggtg	gacggggcca	gcatcaagtt	cgagagcatc	tgctcttacg	20580
ccaccttctt	ccccatggcc	cacaacacgg	cctccacgct	cgaggccatg	ctcaggaacg	20640
acaccaacga	ccagtccttc	aatgactacc	tctccgcgc	caacatgctc	tacccatac	20700
ccgccaacgc	caccaacgtc	cccatctcca	tccctcgcg	caactgggcg	gccttcgcg	20760
gctgggcctt	caccgcctc	aagaccaagg	agacccctc	cctgggctcg	ggattcgacc	20820
cctactacac	ctactcgggc	tccattccct	acctggacgg	caccttctac	ctcaaccaca	20880
ctttcaagaa	ggtctcggtc	accttcgact	cctcggtcag	ctggccgggc	aacgaccgtc	20940
tgctcacccc	caacgagttc	gaaatcaagc	gctcggtcga	cggggaggggc	tacaacgtgg	21000
cccagtgcaa	catgaccaag	gactgggttc	tgggtccagat	gctggccaac	tacaacatcg	21060
gctaccaggg	cttctacatc	ccagagagct	acaaggacag	gatgtactcc	ttcttcagga	21120
acttcagcc	catgagccgg	caggtgggtg	accagaccaa	gtacaaggac	taccaggagg	21180
tgggcatcat	ccaccagcac	aacaactcgg	gcttcgtggg	ctacctcgcc	cccaccatgc	21240
gcgaggggaca	ggcctacccc	gccaacttcc	cctaccgct	cataggcaag	accgcggtcg	21300
acagcatcac	ccagaaaaag	ttcctctgcg	atcgcaccct	ctggcgcatac	cccttctcca	21360
gcaacttcat	gtccatgggt	gcgctctcgg	acctgggcca	gaacttgctc	tacgccaact	21420
ccgcccacgc	cctcgacatg	accttcgagg	tcgaccccat	ggacgagccc	acccttctct	21480
atgttctgtt	cgaagtcttt	gacgtgggtcc	gggtccacca	gccgcaccgc	ggcgtcatcg	21540

agaccgtgta	cctgcgtacg	cccttctcgg	ccggcaacgc	caccacctaa	agaagcaagc	21600
cgcagtcac	gccgcctgca	tgccgtcggg	ttccaccgag	caagagctca	gggccatcgt	21660
cagagacctg	ggatgcgggc	cctatTTTTT	gggcactttc	gacaagcgct	tccctggctt	21720
tgtctcccca	cacaagctgg	cctgcgccat	cgtcaacacg	gccggccgcg	agaccggggg	21780
cgtgcactgg	ctggccttcg	cctggaaccc	gcgctccaaa	acatgcttcc	tctttgaccc	21840
cttcggcttt	tcggaccagc	ggctcaagca	aatctacgag	ttcgagtacg	agggtttgct	21900
gcgtcgcagc	gccatcgcct	cctcgcccga	ccgctgcgtc	accctcgaaa	agtccaccca	21960
gaccgtgcag	gggcccgaact	cgcccgccctg	cggtctcttc	tgctgcatgt	ttctgcacgc	22020
ctttgtgcac	tggcctcaga	gtcccatgga	ccgcaacccc	accatgaact	tgctgacggg	22080
ggtgcccac	tccatgctcc	agagccccc	ggtcgagccc	accctgcgcc	gcaaccagga	22140
gcagctctac	agcttcctgg	agcgccactc	gccctacttc	cgcgcgcaca	gcgcacagat	22200
caggagggcc	acctccttct	gccacttgca	agagatgcaa	gaagggtaat	aacgatgtac	22260
acactTTTTT	ctcaataaat	ggcattTTTT	tatttataca	agctctctgg	ggtattcatt	22320
tcccaccacc	accaccaccc	gccgttgctg	ccatctggct	ctatttagaa	atcgaaaggg	22380
ttctgccggg	agtcgccgtg	cgccacgggc	agggacacgt	tgcgatactg	gtagcgggtg	22440
ccccacttga	actcgggcac	caccaggcga	ggcagctcgg	ggaagttttc	gctccacagg	22500
ctgcgggtca	gcaccagcgc	gttcacaggg	tcggggcgccg	agatcttgaa	gtcgcagttg	22560
gggccgccgc	cctgcgcgcg	cgagttgcgg	tacaccgggt	tgagcactg	gaacaccaac	22620
agcgccgggt	gcttcacgct	agccagcacg	ctgcggtcgg	agatcagctc	ggcgtccagg	22680
tcctccgcgt	tgctcagcgc	gaacggggtc	atcttgggca	cttgccctcc	caggaagggc	22740
gcgtgccccg	gtttcgagtt	gcagtcgcag	cgcagcggga	tcagcaggtg	cccatgcccg	22800
gactcggcgt	tggggtagag	cgcgcgcatg	aaggcctgca	tctggcgga	ggccatctgg	22860
gccttggcgc	cctccgagaa	gaacatgccg	caggacttgc	cgcgagaactg	gtttgcgggg	22920
cagctggcgt	cgtgcaggca	gcagcgcgcg	tcgggtgttg	cgatctgcac	cacgttgcgc	22980
ccccaccggt	tcttcacgat	cttggccttg	gacgattgct	ccttcagcgc	gcgctgcccg	23040
ttctcgctgg	tcacatccat	ctcgatcaca	tgttccttgt	tcacatgct	gctgccgtgc	23100
aggcacttca	gctcgccttc	cgtctcgggtg	cagcgggtgct	gccacagcgc	gcagcccgtg	23160
ggctcgaaag	acttgtaggt	cacctccgcg	aaggactgca	ggtacccttg	caaaaagcgg	23220
cccatcatgg	tcacgaaggt	cttggtgctg	ctgaaggtca	gctgcagccc	gcggtgctcc	23280
tcgttcagcc	aggtcttgca	cacggccgcc	agcgccctca	cctggtcggg	cagcatcttg	23340
aagttcacct	tcagctcatt	ctccacgtgg	tacttgctca	tcagcgtgcg	cgcgcctcc	23400
atgcccttct	cccaggccga	caccagcggc	aggctcacgg	ggttcttcac	catcaccggtg	23460

gccgccgcct	ccgccgcgct	ttcgctttcc	gccccgctgt	tctcttcctc	ttcctcctct	23520
tcctcgccgc	cgcccactcg	cagccccgcg	accacggggg	cgtcttcctg	cagggcgctgc	23580
accttgcgct	tgccgttgcg	cccctgcttg	atgcgcacgg	gcggggttgct	gaagcccacc	23640
atcaccagcg	cggcctcttc	ttgctcgctc	tcgctgtcca	gaatgacctc	cggggagggg	23700
gggttggtca	tcctcagtac	cgaggcacgc	ttctttttct	tcctgggggc	gttcgcccagc	23760
tccgcggctg	cggccgctgc	cgaggtcgaa	ggccgagggc	tgggcgtgcg	cggcaccagc	23820
gcgtcctgcg	agccgtcctc	gtcctcctcg	gactcgagac	ggaggcgggc	ccgcttcttc	23880
gggggcgcgc	ggggcgggcg	aggcgggcgc	ggcgacggag	acggggacga	gacatcgctc	23940
agggtgggtg	gacggcgggc	cgcgccgcgt	ccgcgctcgg	gggtgggtctc	gcgctgggtc	24000
tcttcccgac	tggccatctc	ccactgctcc	ttctcctata	ggcagaaaga	gatcatggag	24060
tctctcatgc	gagtcgagaa	ggaggaggac	agcctaaccg	ccccctctga	gccctccacc	24120
accgccgcca	ccaccgcaa	tgccgccgcg	gacgacgcgc	ccaccgagac	caccgccagt	24180
accaccnnnc	tccccagcga	cgcacccccg	ctcgagaatg	aagtgctgat	cgagcaggac	24240
ccgggttttg	tgagcggaga	ggaggatgag	gtggatgaga	aggagaagga	ggaggtcgcc	24300
gcctcagtgc	caaaagagga	taaaaagcaa	gaccaggacg	acgcagataa	ggatgagaca	24360
gcagtcgggc	gggggaacgg	aagccatgat	gctgatgacg	gctacctaga	cgtgggagac	24420
gacgtgctgc	ttaagcacct	gcaccgccag	tgcgctcatcg	tctgcgacgc	gctgcaggag	24480
cggtgcgaa	tgccccctga	cgtggcgagg	gtcagccgcg	cctacgagcg	gcacctcttc	24540
gcgccgcacg	tgccccccaa	gcgccgggag	aacggcacct	gcgagcccaa	cccgcgtctc	24600
aacttctacc	cggctcttcg	ggtacccgag	gtgctggcca	cctaccacat	ctttttccaa	24660
aactgcaaga	tccccctctc	ctgccgcgct	aaccgcaccc	gcgccgacaa	aaccctgacc	24720
ctgcggcagg	gcgcccacat	acctgatatt	gcctctctgg	aggaagtgcc	caagatcttc	24780
gagggctctg	gtcgcgacga	gaaacggggc	gcgaacgctc	tgcacggaga	cagcgaaaac	24840
gagagtcact	cggggggtgct	ggtggagctc	gagggcgaca	acgcgcgcct	ggccgtactc	24900
aagcgcagca	tagaggtcac	ccactttgcc	taccggcgcg	tcaacctgcc	ccccaaaggtc	24960
atgagtgtgg	tcattgggca	gctcatcatg	cgccgcgccc	agcccctggc	cgcggatgca	25020
aacttgcaag	agtcctcaga	ggaaggcctg	cccgcggtca	gcgacgagca	gctggcgcgcg	25080
tggctggaga	cccgcgaccc	cgcgcagctg	gaggagcggc	gcaagctcat	gatggccgcg	25140
gtgctgggtca	ccgtggagct	cgagtgtctg	cagcgcttct	tcgcggaccc	cgagatgcag	25200
cgcaagctcg	aggagaccct	gcactacacc	ttccgccagg	gctacgtgcg	ccaggcctgc	25260
aagatctcca	acgtggagct	ctgcaacctg	gtctcctacc	tgggcatacct	gcacgagaac	25320
cgctcggggc	agaacgtcct	gcactccacc	ctcaaagggg	aggcgcgccg	cgactacatc	25380

cgcgactgcg	cctacctctt	cctctgctac	acctggcaga	cggccatggg	ggtctggcag	25440
cagtgcctgg	aggagcgcaa	cctcaaggag	ctggaaaagc	tcctcaagcg	cacctcagg	25500
gacctctgga	cgggcttcaa	cgagcgctcg	gtggccgccc	cgctggcgga	catcatcttc	25560
cccgagcgcc	tgctcaagac	cctgcagcag	ggcctgcccg	acttcaccag	ccagagcatg	25620
ctgcagaact	tcaggacttt	catcctggag	cgctcgggca	tcctgccggc	cacttgctgc	25680
gcgctgccc	gcgacttcgt	gcccataaag	tacagggagt	gcccgcgcgc	gctctggggc	25740
cactgctacc	tcttcagct	ggccaactac	ctcgcctacc	actcggacct	catggaagac	25800
gtgagcggcg	agggcctgct	cgagtgccac	tgccgctgca	acctctgcac	gccccaccgc	25860
tctctagtct	gcaaccgcga	gctgctcagc	gagagtcaga	ttatcggtac	cttcgagctg	25920
cagggtcctt	cgctgacga	gaagtccgcg	gctccggggc	tgaaactcac	tccggggctg	25980
tggacttccg	cctacctacg	caaatttgta	cctgaggact	accacgccc	cgagatcagg	26040
ttctacgaag	accaatccc	cccgcccaag	gcggagctca	ccgctgcgt	catcacccag	26100
gggcacatcc	tgggccaatt	gcaagccatc	aacaaagccc	gccgagagtt	cttgctgaaa	26160
aagggtcggg	gggtgtacct	ggacccccag	tccggcgagg	agctaaacct	gctacccccg	26220
ccgcccgcgc	agcagcggga	ccttgcttcc	caggatggca	cccagaaaga	agcagcagcc	26280
gccgcccag	ccatacatgc	ttctggagga	agaggaggag	gactgggaca	gtcaggcaga	26340
ggagatgatg	gaagactggg	aggaggacag	cagcctagac	gaggaagctt	cagaggccga	26400
agaggtggca	gacgcaacac	catcaccttc	ggtcgcagcc	ccctcgccgg	ggccccctgaa	26460
atcctccgaa	cccagcacca	gcgctataac	ctccgctcct	ccggcgccgg	cgccaccgcg	26520
ccgcagacct	aaccgtagat	gggacaccac	aggaaccggg	gtcggtaagt	ccaagtgccc	26580
gccgcccga	ccgcagcagc	agcagcagca	gcgccagggc	taccgctcgt	ggcgcgggca	26640
caagaacgcc	atagtcgcct	gcttgcaaga	ctgcgggggc	aacatctctt	tcgcccgcgc	26700
cttctgcta	ttccaccacg	gggtcgcctt	tccccgcaat	gtcctgcatt	actaccgtca	26760
tctctacagc	ccctactgca	gcggcgaccc	agaggcgggc	gcggcgagcca	cagcgggcag	26820
caccacctag	gaagatatcc	tccgcgggca	agacagcggc	agcagcggcc	aggagacctg	26880
cggcagcagc	ggcgggagcg	gtgggcgcac	tgcgctcttc	gcccacgaa	cccctctcga	26940
cccgggagct	cagacacagg	atcttcccc	ctttgtatgc	catcttccaa	cagagcagag	27000
gccaggagca	ggagctgaaa	ataaaaaaca	gatctctgcg	ctccctcacc	cgcagctgtc	27060
tgtatcacia	aagcgaagat	cagcttcggc	gcacgctgga	ggacgcggag	gcactcttca	27120
gcaaatactg	cgcgctcact	cttaaagact	agctccgcgc	ccttctcgaa	tttagggggg	27180
agaaaactac	gtcatcgccg	gccgcccgc	agcccgcaca	gccgagatga	gcaaagagat	27240
tcccacgcca	tacatgtgga	gctaccagcc	gcagatggga	ctcgcgggcg	gagcgggcca	27300

ggactactcc	acccgcatga	actacatgag	cgcggggaccc	cacatgatct	cacaggtcaa	27360
cgggatccgc	gcccagcgaa	accaaatact	gctggaacag	gcgggccatca	ccgccacgcc	27420
ccgccataat	ctcaaccccc	gaaattggcc	cgccgccccta	gtgtaccagg	aaacccccctc	27480
cgccaccacc	gtactacttc	cgcgtgacgc	ccaggccgaa	gtccagatga	ctaactcagg	27540
ggcgcagctc	gcggggcggct	ttcgtcacgg	ggcgcgggccg	ctccgaccag	gtataagaca	27600
cctgatgatc	agaggccgag	gtatccagct	caacgacgag	tcggtgagct	cttcgctcgg	27660
tctccgtccg	gacggaactt	tccagctcgc	cggatccggt	cgctcttcgt	tcacgccccg	27720
ccaggcgtag	ctgactctgc	agacctcgtc	ctcggagccc	cgctccggcg	gcatcggaac	27780
cctccagttc	gtggaggagt	tcgtgccctc	ggtctacttc	aaccccttct	cgggacctcc	27840
cggacgctac	cccgaccagt	tcattccgaa	ctttgacgcg	gtgaaggact	cggcggacgg	27900
ctacgactga	atgtcagggt	ccgaggcaga	gcagcttcgc	ctgagacacc	tcgagcactg	27960
ccgccgccac	aagtgtttcg	cccgcggttc	cggtaggttc	tgctactttc	agctacccga	28020
ggagcatacc	gagggggccgg	cgcacggcgt	ccgcctgacc	accagggcg	aggttacctg	28080
ttccctcatc	cgggagttca	ccctccgtcc	cctgctagt	gagcgggagc	gggggtccctg	28140
tgtcctaact	atgcctgca	actgccctaa	ccctggatta	catcaagatc	tttgctgtca	28200
tctctgtgct	gagtttaata	aacgctgaga	tcagaatcta	ctggggctcc	tgctgccatc	28260
ctgtgaacgc	caccgtcttc	accacccccg	accaggccca	ggcgaacctc	acctgcggtc	28320
tgcatcggag	gtccaagaag	tacctcacct	ggtacttcaa	cggcaccccc	tttgtggttt	28380
acaacagctt	cgacggggac	ggagtctccc	tgaaagacca	gctctccggt	ctcagctact	28440
ccatccacaa	gaacaccacc	ctccaactct	tcctcccta	cctgccggga	acctacgagt	28500
gcgtcaccgg	ccgtgcacc	cacctaccc	gcctgatcgt	aaaccagagc	tttccgggaa	28560
cagataactc	cctcttcccc	agaacaggag	gtgagctcag	gaaactcccc	ggggaccagg	28620
gcggagacgt	accttcgacc	cttgtgggg	taggattttt	tattaccggg	ttgctggctc	28680
ttttaatcaa	agcttccttg	agatttggtc	tttccttcta	cgtgtatgaa	cacctcagcc	28740
tccaataact	ctaccctttc	ttcgggatca	ggtgactttt	ctgaaatcgg	gcttggtgtg	28800
ctgcttactc	tgttgatttt	tttccttata	atactcagcc	ttctgtgcct	caggctcgcc	28860
gcctgctgcg	cacacatcta	tatctactgc	tggttgctca	agtgacgggg	tcgccaccca	28920
agatgaacag	gtacatggtc	ctatcgatcc	taggcctgct	ggccctggcg	gcctgcagcg	28980
ccgcaaaaaa	agagattacc	tttgaggagc	ccgcttgcaa	tgtaactttc	aagcccaggg	29040
gtgaccaatg	caccaccctc	gtcaaatgcg	ttaccaatca	tgagaagctg	cgcacgcact	29100
acaaaaacaa	aactggccgg	tttgcggtct	atagtgtgtt	tacgcccgga	gacccctcta	29160
actactctgt	caccgtcttc	cagggcggac	agtctaagat	attcaattac	actttccctt	29220

tttatgagtt gtgcgatgcg gtcattgtaca tgtcaaaaca gtacaacctg tggcctccct	29280
ctccccaggc gtgtgtggaa aatactgggt cttactgctg tatggctttg gcaatcacta	29340
cgctcgctct aatctgcacg gtgctatata taaaattcag gcagaggcga atctttatcg	29400
atgaaaagaa aatgccttga tcgctaacac cggctttcta tctgcagaat gaatgcaatc	29460
acctccctac taatcaccac caccctcctt gcgattgccc atggggttgac acgaatcgaa	29520
gtgccagtgg ggtccaatgt caccatgggtg ggccccgcgc gcaattccac cctcatgtgg	29580
gaaaaatttg tccgcaatca atggggttcat ttctgctcta accgaatcag tatcaagccc	29640
agagccatct gcgatgggca aaatctaact ctgatcaatg tgcaaatgat ggatgctggg	29700
tactattacg ggcagcgggg agaaatcatt aattactggc gacccacaa ggactacatg	29760
ctgcatgtag tcgaggcact tcccactacc accccacta ccacctctcc caccaccact	29820
accactacta ctactactac taccactacc gctgccccgc ataccgcga aagcaccatg	29880
attagcacia agccccctcg tgctcactcc cagcgcggcg ggcccatcgg tgcgacctca	29940
gaaaccaccg agctttgctt ctgccaatgc actaacgcca gcgctcatga actgttcgac	30000
ctggagaatg aggatgcccc gcagagctcc gcttgccctga cccaggaggc tgtggagccc	30060
gttgccctga agcagatcgg tgattcaata attgactctt cttcttttgc cactcccga	30120
taccctcccg attctacttt ccacatcacg ggtaccaaag accctaacct ctctttctac	30180
ctgatgctgc tgctctgtat ctctgtggtc tcttccgcgc tgatgttact ggggatgttc	30240
tgctgcctga tctgccgcag aaagagaaaa gctcgctctc agggccaacc actgatgccc	30300
ttccccctacc ccccgattt tgcagataac aagatatgag ctcgctgctg aactaaccg	30360
ctttactagc ctgcgctcta acccttgctg cttgcgactc gagattccac aatgtcacag	30420
ctgtggcagg agaaaatgtt actttcaact ccacggccga taccagtggt tcgtggagtg	30480
gctcaggtag ctacttaact atctgcaata gctccacttc cccagcata tccccacca	30540
agtaccaatg caatgccagc ctgttcaccc tcatcaacgc ttccacctg gacaatggac	30600
tctatgtagg ctatgtaccc tttgggtgggc aaggaaagac ccacgcttac aacctggaag	30660
ttcgccagcc cagaaccact acccaagctw cymccaycac cagcaccagc agcagcagcc	30720
acagcagcag cagcagatta ttgactttgg ttttggccag ctcatctgcc gctaccagg	30780
ccatctacag ctctgtgccc gaaaccactc agaccaccg cccagaaacg accaccgcca	30840
ccaccctaca cacctccagc gatcagatgc cgaccaacat caccctcttg gctcttcaaa	30900
tgggacttac aagccccact ccaaaaccag tggatgcggc cgaggtctcc gccctcgtca	30960
atgactgggc ggggctggga atgtggtggt tcgccatagg catgatggcg ctctgcctgc	31020
ttctgctctg gctcatctgc tgccctcacc gcaggcgagc cagaccccc atctatagac	31080
ccatcattgt cctgaacccc gataatgatg ggatccatag attggatggc ctgaaaaacc	31140

tacttttttc	ttttacagta	tgataaattg	agacatgcct	cgcattttct	tgtacatggt	31200
ccttctccca	ccttttctgg	ggtgttctac	gctggccgct	gtgtctcacc	tggaggtaga	31260
ctgcctctca	cccttctactg	tctacctgct	ttacggattg	gtcaccctca	ctctcatctg	31320
cagcctaatc	acagtaatca	tcgccttcat	ccagtgcatt	gattacatct	gtgtgcgcct	31380
cgcatacttc	agacaccacc	cgcagtaccg	agacaggaac	attgcccaac	ttctaagact	31440
gctctaatac	tgcataagac	tgtgatctgc	cttctgatcc	tctgcatect	gccaccctc	31500
acctcctgcc	agtacaccac	aaaatctccg	cgcaaaagac	atgcctcctg	ccgcttcacc	31560
caactgtgga	atatacccaa	atgctacaac	gaaaagagcg	agctctccga	agcttggctg	31620
tatgggggtc	tctgtgtctt	agttttctgc	agcactgtct	ttgccctcat	gatctacccc	31680
tactttgatt	tgggatggaa	cgcgatcgat	gccatgaatt	acccacctt	tcccgacccc	31740
gagataattc	cactgcgaca	agttgtacct	gttgtcggtt	atcaacgccc	cccatccct	31800
acgcccactg	aaatcagcta	ctttaacct	acaggcggag	atgactgacg	ccctagatct	31860
agaaatggac	ggcatcagta	ccgagcagcg	tctcctagag	aggcgcaggc	aggcggctga	31920
gcaagagcgc	ctcaatcagg	agctccgaga	tctcgttaac	ctgcaccagt	gcaaaagagg	31980
catcttttgt	ctggtaaagc	aggccaaagt	cacctacgag	aagaccggca	acagccaccg	32040
cctcagttac	aaattgccca	cccagcgcca	gaagctgggt	ctcatgggtg	gtgagaatcc	32100
catcacgctc	accagcact	cggtagagac	cgaggggtgt	ctgcactccc	cctgtcgggg	32160
tccagaagac	ctctgcaccc	tggtaaagac	cctgtgcggt	ctcagagatt	tagtcccctt	32220
taactaatca	aacactggaa	tcaataaaaa	gaatcactta	cttaaaatca	gacagcaggt	32280
ctctgtccag	tttattcagc	agcacctcct	tcccctcctc	ccaactctgg	tactccaaac	32340
gccttctggc	ggcaaacttc	ctccacaccc	tgaagggaat	gtcagattct	tgctcctgtc	32400
cctccgcacc	cactatcttc	atgttggtgc	agatgaagcg	caccaaaccg	tctgacgaga	32460
gcttcaaccc	cgtgtacccc	tatgacacgg	aaagcggccc	tccctccgtc	cctttcctca	32520
cccctccctt	cgtgtctccc	gatggattcc	aagaaagtcc	ccccggggtc	ctgtctctga	32580
acctggccga	gcccctgggtc	acttcccacg	gcatgctcgc	cctgaaaatg	ggaagtggcc	32640
tctccctgga	cgacgctggc	aacctcacct	ctcaagatat	caccaccgct	agccctcccc	32700
tcaaaaaaac	caagaccaac	ctcagcctag	aaacctcatc	ccccctaact	gtgagcacct	32760
caggcgcctt	caccgtagca	gccgccgctc	ccttggcggt	ggccggcacc	tccctcacca	32820
tgcaatcaga	ggccccctg	acagtacagg	atgcaaaact	caccttggcc	accaaaggcc	32880
ccctgaccgt	gtctgaaggc	aaactggcct	tgcaaacatc	ggccccgctg	acggccgctg	32940
acagcagcac	cctcacggtt	agcgccacac	caccaattaa	tgtaagcagt	ggaagttag	33000
gcttagacat	ggaagaccct	atgtatactc	acaatggaaa	actgggaata	agaattgggg	33060

gtccactaag	agtagtagac	agcttgcata	cactcactgt	agttaccgga	aatggactaa	33120
ctgtagataa	caatgccctc	caaactaaag	ttacggggcg	cctaggttat	gacacatcag	33180
gaaatctaca	attaagagct	gcaggaggta	tgcggaattga	cgcaaatggc	caacttatcc	33240
ttaatgtggc	ataccatttt	gatgctcaga	acaatctcag	ccttagactt	ggtcagggac	33300
ccctgtatat	aaacacagac	cacaacctgg	atttgaattg	caacagaggt	ctaaccacaa	33360
ctaccaccaa	caacacaaaa	aaacttgaga	ctaaaattag	ctcaggctta	gactatgaca	33420
ccaatgggtg	tgctattatt	aaacttggca	ctgggtctaag	cttcgacaac	acagggcgcc	33480
taactgtggg	aaacactggg	gatgataaac	tgactctgtg	gacgaccca	gacccatctc	33540
caaattgcag	aattcactca	gacaaagact	gcaagtttac	tctagtccta	actaagtgtg	33600
gaagccaaat	cctggcctct	gtcgccgccc	tagcgggtatc	aggaaatctg	gcttcgataa	33660
caggcacctg	tgccagcggt	accatctttc	tcagatttga	tcagaatgga	gtgcttatgg	33720
aaaactcctc	gctagacagg	cagtactgga	acttcagaaa	tggcaactca	actaacgctg	33780
ccccctacac	caatgcagtt	gggttcatgc	caaacctcgc	agcatacccc	aaaacgcaaa	33840
gccagactgc	taaaaacaac	attgtaagtc	aggtttactt	gaatggagac	aaatccaaac	33900
ccatgaccct	taccatcacc	ctcaatggaa	ctaatgaatc	cagtgaaact	agccagggtga	33960
gtcactactc	catgtcattt	acatgggctt	gggaaagtgg	gcaatatgcc	actgaaacct	34020
ttgccaccaa	ctccttcacc	ttttcttaca	ttgctgaaca	ataaaaagca	tgacactgat	34080
gttcatttct	gattcttatt	ttattatttt	caaacacaac	aaaatcattc	aagtcattct	34140
tccatcttag	cttaatagac	acagtagctt	aatagaccca	gtagtgcaaa	gccccattct	34200
agcttataac	tagtggagaa	gtactcgcct	acatgggggt	agagtcataa	tcgtgcatca	34260
ggatagggcg	gtggtgctgc	agcagcgcg	gaataaactg	ctgccgccc	cgctccgtcc	34320
tgcaggaata	caacatggca	gtggtctcct	cagcgatgat	tcgcaccgcc	cgcagcataa	34380
ggcgcttgt	cctccgggca	cagcagcgca	ccctgatctc	acttaaataca	gcacagtaac	34440
tgcagcacag	caccacaata	ttgttcaaaa	ttccacagtg	caaggcgctg	tatccaaagc	34500
tcatggcggg	gaccacagaa	cccacgtggc	catcatacca	caagcgcagg	tagattaagt	34560
ggcgaccct	cataaacacg	ctggacataa	acattacctc	ttttggcatg	ttgtaattca	34620
ccacctcccg	gtaccatata	aacctctgat	taaacatggc	gccatccacc	accatcctaa	34680
accagctggc	caaaacctgc	ccgccggcta	tacactgcag	ggaaccggga	ctggaacaat	34740
gacagtggag	agcccaggac	tcgtaaccat	ggatcatcat	gctcgtcatg	atatcaatgt	34800
tggcacaaca	caggcacacg	tgcatacact	tcctcaggat	tacaagctcc	tcccgcgtta	34860
gaaccatatc	ccagggaaca	acccattcct	gaatcagcgt	aaatcccaca	ctgcagggaa	34920
gacctcgcac	gtaactcacg	ttgtgcattg	tcaaagtgtt	acattcgggc	agcagcggat	34980

gatcctccag	tatggttagcg	cgggttttctg	tctcaaaagg	aggtagacga	tcctactgt	35040
acggagtgcg	ccgagacaac	cgagatcgtg	ttggtcgtag	tgtcatgcc	aatggaacgc	35100
cggacgtagt	catatttctt	gaagtcttag	atctctcaac	gcagcaccag	caccaacact	35160
tcgcagtgt	aaaggccaag	tgccgagaga	gtatatatag	gaataaaaag	tgacgtaaac	35220
gggcaaagtc	caaaaaacgc	ccagaaaaac	cgcacgcgaa	cctacgcccc	gaaacgaaag	35280
ccaaaaaaca	ctagacactc	ccttcggcg	tcaacttccg	ctttccacg	ctacgtcact	35340
tgccccagtc	aaacaaacta	catatcccga	acttccaagt	cgccacgccc	aaaacaccgc	35400
ctacacctcc	ccgcccgcgc	gcccccccc	aaaccgcct	cccgccccgc	gccccgcccc	35460
gcgcgcgcca	tctcattatc	atattggctt	caatccaaaa	taaggatat	tattgatgat	35520
ggtttaaagc	gatccaattc	ttgaagacga	aagggcctcg	tgatacgct	atttttatag	35580
gttaatgtca	tgataataat	ggtttcttag	acgtcaggtg	gcacttttgc	gggaaatgtg	35640
cgcggaaccc	ctatttgttt	atttttctaa	atacattcaa	atatgtatcc	gctcatgaga	35700
caataaccct	gataaatgct	tcaataatat	tgaaaaagga	agagtatgag	tattcaacat	35760
ttccgtgtcg	cccttattcc	cttttttgcg	gcattttgcc	ttcctgtttt	tgctcaccca	35820
gaaacgctgg	tgaaagtaaa	agatgctgaa	gatcagttgg	gtgcacgagt	gggttacatc	35880
gaactggatc	tcaacagcgg	taagatcctt	gagagttttc	gccccgaaga	acgtttttcca	35940
atgatgagca	cttttaaagt	tctgctatgt	ggcgcggtat	tatcccgtgt	tgacgcgggg	36000
caagagcaac	tcggtcgccg	catacactat	tctcagaatg	acttggttga	gtactcacca	36060
gtcacagaaa	agcatcttac	ggatggcatg	acagtaagag	aattatgcag	tgctgccata	36120
accatgagtg	ataaactgct	ggccaactta	cttctgacaa	cgatcggagg	accgaaggag	36180
ctaaccgctt	ttttgcacaa	catgggggat	catgtaactc	gccttgatcg	ttgggaaccg	36240
gagctgaatg	aagccatacc	aaacgacgag	cgtgacacca	cgatgcctgt	agcaatggca	36300
acaacgttgc	gcaaactatt	aactggcgaa	ctacttactc	tagcttcccg	gcaacaatta	36360
atagactgga	tggaggcgga	taaagttgca	ggaccacttc	tgcgctcggc	ccttcgggct	36420
ggctggttta	ttgctgataa	atctggagcc	ggtgagcgtg	ggtctcgcg	tatcattgca	36480
gcactggggc	cagatggtaa	gccctcccgt	atcgtagtta	tctacacgac	ggggagtcag	36540
gcaactatgg	atgaacgaaa	tagacagatc	gctgagatag	gtgcctcact	gattaagcat	36600
tggtaaactgt	cagaccaagt	ttactcatat	atactttaga	ttgatttaaa	aggatctagg	36660
tgaagatcct	ttttgataat	ctcatgacca	aaatccctta	acgtgagttt	tcgttccact	36720
gagcgtcaga	ccccgtagaa	aagatcaaag	gatcttcttg	agatcctttt	tttctgcgcg	36780
taatctgctg	cttgcaaaca	aaaaaaccac	cgctaccagc	ggtggtttgt	ttgccggatc	36840
aagagctacc	aactcttttt	ccgaaggtaa	ctggcttcag	cagagcgcag	ataccaaata	36900

```

ctgtccttct agtgtagccg tagttaggcc accacttcaa gaactctgta gcaccgccta 36960
catacctcgc tctgctaata ctgttaccag tggctgctgc cagtggcgat aagtcgtgtc 37020
ttaccggggtt ggactcaaga cgatagttac cggataaggc gcagcggtcg ggctgaacgg 37080
gggggttcgtg cacacagccc agcttggagc gaacgaccta caccgaactg agatacctac 37140
agcgtgagct atgagaaaagc gccacgcttc ccgaaggag aaaggcggac aggtatccgg 37200
taagcggcag ggtcggaaca ggagagcgca cgaggagct tccaggggga aacgcctggt 37260
atctttatag tcctgtcggg tttcgccacc tctgacttga gcgtcgattt ttgtgatgct 37320
cgtcaggggg gcggagccta tggaaaaacg ccagcaacgc ggccttttta cggttcctgg 37380
ccttttgctg gccttgaagc tgtccctgat ggctgctatc tacctgcctg gacagcatgg 37440
cctgcaacgc gggcatcccg atgccgccg aagcgagaag aatcataatg gggaaggcca 37500
tccagcctcg cgtcgcagat ccgaattcgt ttaaac 37536

```

```

<210> 25
<211> 29
<212> DNA
<213> Human cytomegalovirus

```

```

<400> 25
catctacgta ttagtcatcg ctattacca 29

```

```

<210> 26
<211> 21
<212> DNA
<213> Human cytomegalovirus

```

```

<400> 26
gacttggaat tccccgtgag t 21

```

```

<210> 27
<211> 543
<212> PRT
<213> Pan troglodytes

```

```

<400> 27

```

```

Met Lys Arg Thr Lys Thr Ser Asp Glu Ser Phe Asn Pro Val Tyr Pro
1          5          10          15

```

```

Tyr Asp Thr Glu Ser Gly Pro Pro Ser Val Pro Phe Leu Thr Pro Pro
          20          25          30

```

```

Phe Val Ser Pro Asp Gly Phe Gln Glu Ser Pro Pro Gly Val Leu Ser
          35          40          45

```

```

Leu Asn Leu Ala Glu Pro Leu Val Thr Ser His Gly Met Leu Ala Leu
          50          55          60

```

Lys	Met	Gly	Ser	Gly	Leu	Ser	Leu	Asp	Asp	Ala	Gly	Asn	Leu	Thr	Ser	65	70	75	80
Gln	Asp	Ile	Thr	Thr	Ala	Ser	Pro	Pro	Leu	Lys	Lys	Thr	Lys	Thr	Asn	85	90	95	
Leu	Ser	Leu	Glu	Thr	Ser	Ser	Pro	Leu	Thr	Val	Ser	Thr	Ser	Gly	Ala	100	105	110	
Leu	Thr	Val	Ala	Ala	Ala	Ala	Pro	Leu	Ala	Val	Ala	Gly	Thr	Ser	Leu	115	120	125	
Thr	Met	Gln	Ser	Glu	Ala	Pro	Leu	Thr	Val	Gln	Asp	Ala	Lys	Leu	Thr	130	135	140	
Leu	Ala	Thr	Lys	Gly	Pro	Leu	Thr	Val	Ser	Glu	Gly	Lys	Leu	Ala	Leu	145	150	155	160
Gln	Thr	Ser	Ala	Pro	Leu	Thr	Ala	Ala	Asp	Ser	Ser	Thr	Leu	Thr	Val	165	170	175	
Ser	Ala	Thr	Pro	Pro	Ile	Asn	Val	Ser	Ser	Gly	Ser	Leu	Gly	Leu	Asp	180	185	190	
Met	Glu	Asp	Pro	Met	Tyr	Thr	His	Asp	Gly	Lys	Leu	Gly	Ile	Arg	Ile	195	200	205	
Gly	Gly	Pro	Leu	Arg	Val	Val	Asp	Ser	Leu	His	Thr	Leu	Thr	Val	Val	210	215	220	
Thr	Gly	Asn	Gly	Leu	Thr	Val	Asp	Asn	Asn	Ala	Leu	Gln	Thr	Arg	Val	225	230	235	240
Thr	Gly	Ala	Leu	Gly	Tyr	Asp	Thr	Ser	Gly	Asn	Leu	Gln	Leu	Arg	Ala	245	250	255	
Ala	Gly	Gly	Met	Arg	Ile	Asp	Ala	Asn	Gly	Gln	Leu	Ile	Leu	Asn	Val	260	265	270	
Ala	Tyr	Pro	Phe	Asp	Ala	Gln	Asn	Asn	Leu	Ser	Leu	Arg	Leu	Gly	Gln	275	280	285	
Gly	Pro	Leu	Tyr	Ile	Asn	Thr	Asp	His	Asn	Leu	Asp	Leu	Asn	Cys	Asn	290	295	300	
Arg	Gly	Leu	Thr	Thr	Thr	Thr	Thr	Asn	Asn	Thr	Lys	Lys	Leu	Glu	Thr	305	310	315	320

Lys Ile Ser Ser Gly Leu Asp Tyr Asp Thr Asn Gly Ala Val Ile Ile
325 330 335

Lys Leu Gly Thr Gly Leu Ser Phe Asp Asn Thr Gly Ala Leu Thr Val
340 345 350

Gly Asn Thr Gly Asp Asp Lys Leu Thr Leu Trp Thr Thr Pro Asp Pro
355 360 365

Ser Pro Asn Cys Arg Ile His Ser Asp Lys Asp Cys Lys Phe Thr Leu
370 375 380

Val Leu Thr Lys Cys Gly Ser Gln Ile Leu Ala Ser Val Ala Ala Leu
385 390 395 400

Ala Val Ser Gly Asn Leu Ala Ser Ile Thr Gly Thr Val Ala Ser Val
405 410 415

Thr Ile Phe Leu Arg Phe Asp Gln Asn Gly Val Leu Met Glu Asn Ser
420 425 430

Ser Leu Asp Arg Gln Tyr Trp Asn Phe Arg Asn Gly Asn Ser Thr Asn
435 440 445

Ala Ala Pro Tyr Thr Asn Ala Val Gly Phe Met Pro Asn Leu Ala Ala
450 455 460

Tyr Pro Lys Thr Gln Ser Gln Thr Ala Lys Asn Asn Ile Val Ser Gln
465 470 475 480

Val Tyr Leu Asn Gly Asp Lys Ser Lys Pro Met Thr Leu Thr Ile Thr
485 490 495

Leu Asn Gly Thr Asn Glu Ser Ser Glu Thr Ser Gln Val Ser His Tyr
500 505 510

Ser Met Ser Phe Thr Trp Ala Trp Glu Ser Gly Gln Tyr Ala Thr Glu
515 520 525

Thr Phe Ala Thr Asn Ser Phe Thr Phe Ser Tyr Ile Ala Glu Gln
530 535 540

<210> 28
<211> 542
<212> PRT
<213> Pan paniscus

<400> 28

Met Lys Arg Ala Lys Thr Ser Asp Glu Thr Phe Asn Pro Val Tyr Pro

1		5		10		15													
Tyr	Asp	Thr	Glu	Asn	Gly	Pro	Pro	Ser	Val	Pro	Phe	Leu	Thr	Pro	Pro				
			20					25					30						
Phe	Val	Ser	Pro	Asp	Gly	Phe	Gln	Glu	Ser	Pro	Pro	Gly	Val	Leu	Ser				
		35					40					45							
Leu	Arg	Leu	Ser	Glu	Pro	Leu	Val	Thr	Ser	His	Gly	Met	Leu	Ala	Leu				
	50					55					60								
Lys	Met	Gly	Asn	Gly	Leu	Ser	Leu	Asp	Asp	Ala	Gly	Asn	Leu	Thr	Ser				
65					70					75					80				
Gln	Asp	Val	Thr	Thr	Val	Thr	Pro	Pro	Leu	Lys	Lys	Thr	Lys	Thr	Asn				
				85					90					95					
Leu	Ser	Leu	Gln	Thr	Ser	Ala	Pro	Leu	Thr	Val	Ser	Ser	Gly	Ser	Leu				
			100					105					110						
Thr	Val	Ala	Ala	Ala	Ala	Pro	Leu	Ala	Val	Ala	Gly	Thr	Ser	Leu	Thr				
		115					120					125							
Met	Gln	Ser	Gln	Ala	Pro	Leu	Thr	Val	Gln	Asp	Ala	Lys	Leu	Gly	Leu				
	130					135					140								
Ala	Thr	Gln	Gly	Pro	Leu	Thr	Val	Ser	Glu	Gly	Lys	Leu	Thr	Leu	Gln				
145					150					155					160				
Thr	Ser	Ala	Pro	Leu	Thr	Ala	Ala	Asp	Ser	Ser	Thr	Leu	Thr	Val	Gly				
				165				170						175					
Thr	Thr	Pro	Pro	Ile	Ser	Val	Ser	Ser	Gly	Ser	Leu	Gly	Leu	Asp	Met				
			180					185					190						
Glu	Asp	Pro	Met	Tyr	Thr	His	Asp	Gly	Lys	Leu	Gly	Ile	Arg	Ile	Gly				
		195					200					205							
Gly	Pro	Leu	Gln	Val	Val	Asp	Ser	Leu	His	Thr	Leu	Thr	Val	Val	Thr				
	210					215					220								
Gly	Asn	Gly	Ile	Thr	Val	Ala	Asn	Asn	Ala	Leu	Gln	Thr	Lys	Val	Ala				
225					230					235					240				
Gly	Ala	Leu	Gly	Tyr	Asp	Ser	Ser	Gly	Asn	Leu	Glu	Leu	Arg	Ala	Ala				
				245				250						255					
Gly	Gly	Met	Arg	Ile	Asn	Thr	Gly	Gly	Gln	Leu	Ile	Leu	Asp	Val	Ala				

			260				265				270					
Tyr	Pro	Phe 275	Asp	Ala	Gln	Asn	Asn 280	Leu	Ser	Leu	Arg	Leu 285	Gly	Gln	Gly	
Pro	Leu 290	Tyr	Val	Asn	Thr	Asn 295	His	Asn	Leu	Asp	Leu 300	Asn	Cys	Asn	Arg	
Gly 305	Leu	Thr	Thr	Thr	Thr 310	Ser	Ser	Asn	Thr	Thr 315	Lys	Leu	Glu	Thr	Lys 320	
Ile	Asp	Ser	Gly	Leu 325	Asp	Tyr	Asn	Ala	Asn 330	Gly	Ala	Ile	Ile	Ala 335	Lys	
Leu	Gly	Thr	Gly 340	Leu	Thr	Phe	Asp	Asn 345	Thr	Gly	Ala	Ile	Thr 350	Val	Gly	
Asn	Thr	Gly 355	Asp	Asp	Lys	Leu	Thr 360	Leu	Trp	Thr	Thr	Pro 365	Asp	Pro	Ser	
Pro	Asn 370	Cys	Arg	Ile	His	Ala 375	Asp	Lys	Asp	Cys	Lys 380	Phe	Thr	Leu	Val	
Leu 385	Thr	Lys	Cys	Gly	Ser 390	Gln	Ile	Leu	Ala	Ser 395	Val	Ala	Ala	Leu	Ala 400	
Val	Ser	Gly	Asn	Leu 405	Ser	Ser	Met	Thr	Gly 410	Thr	Val	Ser	Ser	Val 415	Thr	
Ile	Phe	Leu	Arg 420	Phe	Asp	Gln	Asn	Gly 425	Val	Leu	Met	Glu	Asn 430	Ser	Ser	
Leu	Asp	Lys 435	Glu	Tyr	Trp	Asn	Phe 440	Arg	Asn	Gly	Asn	Ser 445	Thr	Asn	Ala	
Thr	Pro 450	Tyr	Thr	Asn	Ala	Val 455	Gly	Phe	Met	Pro	Asn 460	Leu	Ser	Ala	Tyr	
Pro 465	Lys	Thr	Gln	Ser	Gln 470	Thr	Ala	Lys	Asn	Asn 475	Ile	Val	Ser	Glu	Val 480	
Tyr	Leu	His	Gly	Asp 485	Lys	Ser	Lys	Pro	Met 490	Ile	Leu	Thr	Ile	Thr 495	Leu	
Asn	Gly	Thr	Asn 500	Glu	Ser	Ser	Glu	Thr 505	Ser	Gln	Val	Ser	His 510	Tyr	Ser	
Met	Ser	Phe	Thr	Trp	Ser	Trp	Asp	Ser	Gly	Lys	Tyr	Ala	Thr	Glu	Thr	

515

520

525

Phe Ala Thr Asn Ser Phe Thr Phe Ser Tyr Ile Ala Glu Gln
 530 535 540

<210> 29

<211> 543

<212> PRT

<213> Pan troglodytes

<400> 29

Met Lys Arg Thr Lys Thr Ser Asp Glu Ser Phe Asn Pro Val Tyr Pro
 1 5 10 15

Tyr Asp Thr Glu Ser Gly Pro Pro Ser Val Pro Phe Leu Thr Pro Pro
 20 25 30

Phe Val Ser Pro Asp Gly Phe Gln Glu Ser Pro Pro Gly Val Leu Ser
 35 40 45

Leu Asn Leu Ala Glu Pro Leu Val Thr Ser His Gly Met Leu Ala Leu
 50 55 60

Lys Met Gly Ser Gly Leu Ser Leu Asp Asp Ala Gly Asn Leu Thr Ser
 65 70 75 80

Gln Asp Ile Thr Ser Thr Thr Pro Pro Leu Lys Lys Thr Lys Thr Asn
 85 90 95

Leu Ser Leu Glu Thr Ser Ser Pro Leu Thr Val Ser Thr Ser Gly Ala
 100 105 110

Leu Thr Val Ala Ala Ala Ala Pro Leu Ala Val Ala Gly Thr Ser Leu
 115 120 125

Thr Met Gln Ser Glu Ala Pro Leu Ala Val Gln Asp Ala Lys Leu Thr
 130 135 140

Leu Ala Thr Lys Gly Pro Leu Thr Val Ser Glu Gly Lys Leu Ala Leu
 145 150 155 160

Gln Thr Ser Ala Pro Leu Thr Ala Ala Asp Ser Ser Thr Leu Thr Val
 165 170 175

Ser Ser Thr Pro Pro Ile Ser Val Ser Ser Gly Ser Leu Gly Leu Asp
 180 185 190

Met Glu Asp Pro Met Tyr Thr His Asp Gly Lys Leu Gly Ile Arg Ile
 195 200 205

Gly	Gly	Pro	Leu	Arg	Val	Val	Asp	Ser	Leu	His	Thr	Leu	Thr	Val	Val	210	215	220
Thr	Gly	Asn	Gly	Leu	Thr	Val	Asp	Asn	Asn	Ala	Leu	Gln	Thr	Arg	Val	225	230	235
Thr	Gly	Ala	Leu	Gly	Tyr	Asp	Thr	Ser	Gly	Asn	Leu	Gln	Leu	Arg	Ala	245	250	255
Ala	Gly	Gly	Met	Arg	Ile	Asp	Ala	Asn	Gly	Gln	Leu	Ile	Leu	Asp	Val	260	265	270
Ala	Tyr	Pro	Phe	Asp	Ala	Gln	Asn	Asn	Leu	Ser	Leu	Arg	Leu	Gly	Gln	275	280	285
Gly	Pro	Leu	Tyr	Val	Asn	Thr	Asp	His	Asn	Leu	Asp	Leu	Asn	Cys	Asn	290	295	300
Arg	Gly	Leu	Thr	Thr	Thr	Thr	Thr	Asn	Asn	Thr	Lys	Lys	Leu	Glu	Thr	305	310	315
Lys	Ile	Ser	Ser	Gly	Leu	Asp	Tyr	Asp	Thr	Asn	Gly	Ala	Val	Ile	Ile	325	330	335
Lys	Leu	Gly	Thr	Gly	Leu	Ser	Phe	Asp	Asn	Thr	Gly	Ala	Leu	Thr	Val	340	345	350
Gly	Asn	Thr	Gly	Asp	Asp	Lys	Leu	Thr	Leu	Trp	Thr	Thr	Pro	Asp	Pro	355	360	365
Ser	Pro	Asn	Cys	Arg	Ile	His	Ser	Asp	Lys	Asp	Cys	Lys	Phe	Thr	Leu	370	375	380
Val	Leu	Thr	Lys	Cys	Gly	Ser	Gln	Ile	Leu	Ala	Ser	Val	Ala	Ala	Leu	385	390	395
Ala	Val	Ser	Gly	Asn	Leu	Ala	Ser	Ile	Thr	Gly	Thr	Val	Ala	Ser	Val	405	410	415
Thr	Ile	Phe	Leu	Arg	Phe	Asp	Gln	Asn	Gly	Val	Leu	Met	Glu	Asn	Ser	420	425	430
Ser	Leu	Asp	Lys	Gln	Tyr	Trp	Asn	Phe	Arg	Asn	Gly	Asn	Ser	Thr	Asn	435	440	445
Ala	Ala	Pro	Tyr	Thr	Asn	Ala	Val	Gly	Phe	Met	Pro	Asn	Leu	Ala	Ala	450	455	460

Tyr Pro Lys Thr Gln Ser Gln Thr Ala Lys Asn Asn Ile Val Ser Gln
465 470 475 480

Val Tyr Leu Asn Gly Asp Lys Ser Lys Pro Met Thr Leu Thr Ile Thr
485 490 495

Leu Asn Gly Thr Asn Glu Ser Ser Glu Thr Ser Gln Val Ser His Tyr
500 505 510

Ser Met Ser Phe Thr Trp Ala Trp Glu Ser Gly Gln Tyr Ala Thr Glu
515 520 525

Thr Phe Ala Thr Asn Ser Phe Thr Phe Ser Tyr Ile Ala Glu Gln
530 535 540

<210> 30

<211> 543

<212> PRT

<213> Pan troglodytes

<400> 30

Met Lys Arg Thr Lys Thr Ser Asp Lys Ser Phe Asn Pro Val Tyr Pro
1 5 10 15

Tyr Asp Thr Glu Asn Gly Pro Pro Ser Val Pro Phe Leu Thr Pro Pro
20 25 30

Phe Val Ser Pro Asp Gly Phe Gln Glu Ser Pro Pro Gly Val Leu Ser
35 40 45

Leu Asn Leu Ala Glu Pro Leu Val Thr Ser His Gly Met Leu Ala Leu
50 55 60

Lys Met Gly Ser Gly Leu Ser Leu Asp Asp Ala Gly Asn Leu Thr Ser
65 70 75 80

Gln Asp Val Thr Thr Thr Thr Pro Pro Leu Lys Lys Thr Lys Thr Asn
85 90 95

Leu Ser Leu Glu Thr Ser Ala Pro Leu Thr Val Ser Thr Ser Gly Ala
100 105 110

Leu Thr Leu Ala Ala Ala Ala Pro Leu Ala Val Ala Gly Thr Ser Leu
115 120 125

Thr Met Gln Ser Glu Ala Pro Leu Thr Val Gln Asp Ala Lys Leu Thr
130 135 140

Leu Ala Thr Lys Gly Pro Leu Thr Val Ser Glu Gly Lys Leu Ala Leu
145 150 155 160

Gln Thr Ser Ala Pro Leu Thr Ala Ala Asp Ser Ser Thr Leu Thr Val
165 170 175

Ser Ala Thr Pro Pro Ile Ser Val Ser Ser Gly Ser Leu Gly Leu Asp
180 185 190

Met Glu Asp Pro Met Tyr Thr His Asp Gly Lys Leu Gly Ile Arg Ile
195 200 205

Gly Gly Pro Leu Arg Val Val Asp Ser Leu His Thr Leu Thr Val Val
210 215 220

Thr Gly Asn Gly Ile Ala Val Asp Asn Asn Ala Leu Gln Thr Arg Val
225 230 235 240

Thr Gly Ala Leu Gly Tyr Asp Thr Ser Gly Asn Leu Gln Leu Arg Ala
245 250 255

Ala Gly Gly Met Arg Ile Asp Ala Asn Gly Gln Leu Ile Leu Asp Val
260 265 270

Ala Tyr Pro Phe Asp Ala Gln Asn Asn Leu Ser Leu Arg Leu Gly Gln
275 280 285

Gly Pro Leu Tyr Val Asn Thr Asp His Asn Leu Asp Leu Asn Cys Asn
290 295 300

Arg Gly Leu Thr Thr Thr Thr Thr Asn Asn Thr Lys Lys Leu Glu Thr
305 310 315 320

Lys Ile Gly Ser Gly Leu Asp Tyr Asp Thr Asn Gly Ala Val Ile Ile
325 330 335

Lys Leu Gly Thr Gly Val Ser Phe Asp Ser Thr Gly Ala Leu Ser Val
340 345 350

Gly Asn Thr Gly Asp Asp Lys Leu Thr Leu Trp Thr Thr Pro Asp Pro
355 360 365

Ser Pro Asn Cys Arg Ile His Ser Asp Lys Asp Cys Lys Phe Thr Leu
370 375 380

Val Leu Thr Lys Cys Gly Ser Gln Ile Leu Ala Ser Val Ala Ala Leu
385 390 395 400

Ala Val Ser Gly Asn Leu Ala Ser Ile Thr Gly Thr Val Ser Ser Val
405 410 415

Thr Ile Phe Leu Arg Phe Asp Gln Asn Gly Val Leu Met Glu Asn Ser
420 425 430

Ser Leu Asp Lys Gln Tyr Trp Asn Phe Arg Asn Gly Asn Ser Thr Asn
435 440 445

Ala Thr Pro Tyr Thr Asn Ala Val Gly Phe Met Pro Asn Leu Ala Ala
450 455 460

Tyr Pro Lys Thr Gln Ser Gln Thr Ala Lys Asn Asn Ile Val Ser Gln
465 470 475 480

Val Tyr Leu Asn Gly Asp Lys Ser Lys Pro Met Thr Leu Thr Ile Thr
485 490 495

Leu Asn Gly Thr Asn Glu Ser Ser Glu Thr Ser Gln Val Ser His Tyr
500 505 510

Ser Met Ser Phe Thr Trp Ala Trp Glu Ser Gly Gln Tyr Ala Thr Glu
515 520 525

Thr Phe Ala Thr Asn Ser Phe Thr Phe Ser Tyr Ile Ala Glu Gln
530 535 540

<210> 31
<211> 543
<212> PRT
<213> Pan troglodytes

<400> 31

Met Lys Arg Thr Lys Thr Ser Asp Glu Ser Phe Asn Pro Val Tyr Pro
1 5 10 15

Tyr Asp Thr Glu Asn Gly Pro Pro Ser Val Pro Phe Leu Thr Pro Pro
20 25 30

Phe Val Ser Pro Asp Gly Phe Gln Glu Ser Pro Pro Gly Val Leu Ser
35 40 45

Leu Asn Leu Ala Glu Pro Leu Val Thr Ser His Gly Met Leu Ala Leu
50 55 60

Lys Met Gly Ser Gly Leu Ser Leu Asp Asp Ala Gly Asn Leu Thr Ser
65 70 75 80

Gln	Asp	Val	Thr	Thr	Thr	Thr	Pro	Pro	Leu	Lys	Lys	Thr	Lys	Thr	Asn	
				85					90					95		
Leu	Ser	Leu	Glu	Thr	Ser	Ala	Pro	Leu	Thr	Val	Ser	Thr	Ser	Gly	Ala	
			100					105					110			
Leu	Thr	Leu	Ala	Ala	Ala	Ala	Pro	Leu	Ala	Val	Ala	Gly	Thr	Ser	Leu	
		115					120					125				
Thr	Met	Gln	Ser	Glu	Ala	Pro	Leu	Thr	Val	Gln	Asp	Ala	Lys	Leu	Thr	
	130					135					140					
Leu	Ala	Thr	Lys	Gly	Pro	Leu	Thr	Val	Ser	Glu	Gly	Lys	Leu	Ala	Leu	
145					150					155					160	
Gln	Thr	Ser	Ala	Pro	Leu	Thr	Ala	Ala	Asp	Ser	Ser	Thr	Leu	Thr	Val	
				165					170					175		
Ser	Ala	Thr	Pro	Pro	Ile	Asn	Val	Ser	Ser	Gly	Ser	Leu	Gly	Leu	Asp	
			180					185					190			
Met	Glu	Asn	Pro	Met	Tyr	Thr	His	Asp	Gly	Lys	Leu	Gly	Ile	Arg	Ile	
		195					200					205				
Gly	Gly	Pro	Leu	Arg	Val	Val	Asp	Ser	Leu	His	Thr	Leu	Thr	Val	Val	
	210					215					220					
Thr	Gly	Asn	Gly	Ile	Ala	Val	Asp	Asn	Asn	Ala	Leu	Gln	Thr	Arg	Val	
225					230					235					240	
Thr	Gly	Ala	Leu	Gly	Tyr	Asp	Thr	Ser	Gly	Asn	Leu	Gln	Leu	Arg	Ala	
				245					250					255		
Ala	Gly	Gly	Met	Arg	Ile	Asp	Ala	Asn	Gly	Gln	Leu	Ile	Leu	Asp	Val	
			260					265					270			
Ala	Tyr	Pro	Phe	Asp	Ala	Gln	Asn	Asn	Leu	Ser	Leu	Arg	Leu	Gly	Gln	
		275					280					285				
Gly	Pro	Leu	Tyr	Val	Asn	Thr	Asp	His	Asn	Leu	Asp	Leu	Asn	Cys	Asn	
	290					295					300					
Arg	Gly	Leu	Thr	Thr	Thr	Thr	Thr	Asn	Asn	Thr	Lys	Lys	Leu	Glu	Thr	
305					310					315					320	
Lys	Ile	Gly	Ser	Gly	Leu	Asp	Tyr	Asp	Thr	Asn	Gly	Ala	Val	Ile	Ile	
				325					330					335		

Lys Leu Gly Thr Gly Val Ser Phe Asp Ser Thr Gly Ala Leu Ser Val
340 345 350

Gly Asn Thr Gly Asp Asp Lys Leu Thr Leu Trp Thr Thr Pro Asp Pro
355 360 365

Ser Pro Asn Cys Arg Ile His Ser Asp Lys Asp Cys Lys Phe Thr Leu
370 375 380

Val Leu Thr Lys Cys Gly Ser Gln Ile Leu Ala Ser Val Ala Ala Leu
385 390 395 400

Ala Val Ser Gly Asn Leu Ala Ser Ile Thr Gly Thr Val Ser Ser Val
405 410 415

Thr Ile Phe Leu Arg Phe Asp Gln Asn Gly Val Leu Met Glu Asn Ser
420 425 430

Ser Leu Asp Lys Gln Tyr Trp Asn Phe Arg Asn Gly Asn Ser Thr Asn
435 440 445

Ala Thr Pro Tyr Thr Asn Ala Val Gly Phe Met Pro Asn Leu Ala Ala
450 455 460

Tyr Pro Lys Thr Gln Ser Gln Thr Ala Lys Asn Asn Ile Val Ser Gln
465 470 475 480

Val Tyr Leu Asn Gly Asp Lys Ser Lys Pro Met Ile Leu Thr Ile Thr
485 490 495

Leu Asn Gly Thr Asn Glu Ser Ser Glu Thr Ser Gln Val Ser His Tyr
500 505 510

Ser Met Ser Phe Thr Trp Ala Trp Glu Ser Gly Gln Tyr Ala Thr Glu
515 520 525

Thr Phe Ala Thr Asn Ser Phe Thr Phe Ser Tyr Ile Ala Glu Gln
530 535 540

<210> 32

<211> 578

<212> PRT

<213> Pan troglodytes

<400> 32

Met Lys Arg Thr Lys Thr Ser Asp Glu Ser Phe Asn Pro Val Tyr Pro
1 5 10 15

Tyr Asp Thr Glu Asn Gly Pro Pro Ser Val Pro Phe Leu Thr Pro Pro

20

25

30

Phe Val Ser Pro Asp Gly Phe Gln Glu Ser Pro Pro Gly Val Leu Ser
 35 40 45

Leu Asn Leu Ala Glu Pro Leu Val Thr Ser His Gly Met Leu Ala Leu
 50 55 60

Lys Met Gly Ser Gly Leu Ser Leu Asp Asp Ala Gly Asn Leu Thr Ser
 65 70 75 80

Gln Asp Val Thr Thr Thr Thr Pro Pro Leu Lys Lys Thr Lys Thr Asn
 85 90 95

Leu Ser Leu Glu Thr Ser Ala Pro Leu Thr Val Ser Thr Ser Gly Ala
 100 105 110

Leu Thr Leu Ala Ala Ala Val Pro Leu Ala Val Ala Gly Thr Ser Leu
 115 120 125

Thr Met Gln Ser Glu Ala Pro Leu Thr Val Gln Asp Ala Lys Leu Thr
 130 135 140

Leu Ala Thr Lys Gly Pro Leu Thr Val Ser Glu Gly Lys Leu Ala Leu
 145 150 155 160

Gln Thr Ser Ala Pro Leu Thr Ala Ala Asp Ser Ser Thr Leu Thr Ile
 165 170 175

Ser Ala Thr Pro Pro Leu Ser Thr Ser Asn Gly Ser Leu Gly Ile Asp
 180 185 190

Met Gln Ala Pro Ile Tyr Thr Thr Asn Gly Lys Leu Gly Leu Asn Phe
 195 200 205

Gly Ala Pro Leu His Val Val Asp Ser Leu Asn Ala Leu Thr Val Val
 210 215 220

Thr Gly Gln Gly Leu Thr Ile Asn Gly Thr Ala Leu Gln Thr Arg Val
 225 230 235 240

Ser Gly Ala Leu Asn Tyr Asp Ser Ser Gly Asn Leu Glu Leu Arg Ala
 245 250 255

Ala Gly Gly Met Arg Val Asp Ala Asn Gly Lys Leu Ile Leu Asp Val
 260 265 270

Ala Tyr Pro Phe Asp Ala Gln Asn Asn Leu Ser Leu Arg Leu Gly Gln

275

280

285

Gly Pro Leu Phe Val Asn Ser Ala His Asn Leu Asp Val Asn Tyr Asn
290 295 300

Arg Gly Leu Tyr Leu Phe Thr Ser Gly Asn Thr Lys Lys Leu Glu Val
305 310 315 320

Asn Ile Lys Thr Ala Lys Gly Leu Ile Tyr Asp Asp Thr Ala Ile Ala
325 330 335

Ile Asn Pro Gly Asp Gly Leu Glu Phe Gly Ser Gly Ser Asp Thr Asn
340 345 350

Pro Leu Lys Thr Lys Leu Gly Leu Gly Leu Glu Tyr Asp Ser Ser Arg
355 360 365

Ala Ile Ile Ala Lys Leu Gly Thr Gly Leu Ser Phe Asp Asn Thr Gly
370 375 380

Ala Ile Thr Val Gly Asn Lys Asn Asp Asp Lys Leu Thr Leu Trp Thr
385 390 395 400

Thr Pro Asp Pro Ser Pro Asn Cys Arg Ile Tyr Ser Glu Lys Asp Ala
405 410 415

Lys Phe Thr Leu Val Leu Thr Lys Cys Gly Ser Gln Val Leu Ala Ser
420 425 430

Val Ser Val Leu Ser Val Lys Gly Ser Leu Ala Pro Ile Ser Gly Thr
435 440 445

Val Thr Ser Ala Gln Ile Ile Leu Arg Phe Asp Glu Asn Gly Val Leu
450 455 460

Leu Ser Asn Ser Ser Leu Asp Pro Gln Tyr Trp Asn Tyr Arg Lys Gly
465 470 475 480

Asp Leu Thr Glu Gly Thr Ala Tyr Thr Asn Ala Val Gly Phe Met Pro
485 490 495

Asn Leu Thr Ala Tyr Pro Lys Thr Gln Ser Gln Thr Ala Lys Ser Asn
500 505 510

Ile Val Ser Gln Val Tyr Leu Asn Gly Asp Lys Ser Lys Pro Met Ile
515 520 525

Leu Thr Ile Thr Leu Asn Gly Thr Asn Glu Thr Gly Asp Ala Thr Val

530

535

540

Ser Thr Tyr Ser Met Ser Phe Ser Trp Asn Trp Asn Gly Ser Asn Tyr
 545 550 555 560

Ile Asn Glu Thr Phe Gln Thr Asn Ser Phe Thr Phe Ser Tyr Ile Ala
 565 570 575

Gln Glu

<210> 33
 <211> 578
 <212> PRT
 <213> Pan troglodytes

<400> 33

Met Lys Arg Thr Lys Thr Ser Asp Glu Ser Phe Asn Pro Val Tyr Pro
 1 5 10 15

Tyr Asp Thr Glu Ser Gly Pro Pro Ser Val Pro Phe Leu Thr Pro Pro
 20 25 30

Phe Val Ser Pro Asp Gly Phe Gln Glu Ser Pro Pro Gly Val Leu Ser
 35 40 45

Leu Asn Leu Ala Glu Pro Leu Val Thr Ser His Gly Met Leu Ala Leu
 50 55 60

Lys Met Gly Ser Gly Leu Ser Leu Asp Asp Ala Gly Asn Leu Thr Ser
 65 70 75 80

Gln Asp Ile Thr Thr Ala Ser Pro Pro Leu Lys Lys Thr Lys Thr Asn
 85 90 95

Leu Ser Leu Glu Thr Ser Ser Pro Leu Thr Val Ser Thr Ser Gly Ala
 100 105 110

Leu Thr Val Ala Ala Ala Ala Pro Leu Ala Val Ala Gly Thr Ser Leu
 115 120 125

Thr Met Gln Ser Glu Ala Pro Leu Thr Val Gln Asp Ala Lys Leu Thr
 130 135 140

Leu Ala Thr Lys Gly Pro Leu Thr Val Ser Glu Gly Lys Leu Ala Leu
 145 150 155 160

Gln Thr Ser Ala Pro Leu Thr Ala Ala Asp Ser Ser Thr Leu Thr Val
 165 170 175

Ser	Ala	Thr	Pro	Pro	Leu	Ser	Thr	Ser	Asn	Gly	Ser	Leu	Gly	Ile	Asp	180	185	190
Met	Gln	Ala	Pro	Ile	Tyr	Thr	Thr	Asn	Gly	Lys	Leu	Gly	Leu	Asn	Phe	195	200	205
Gly	Ala	Pro	Leu	His	Val	Val	Asp	Ser	Leu	Asn	Ala	Leu	Thr	Val	Val	210	215	220
Thr	Gly	Gln	Gly	Leu	Thr	Ile	Asn	Gly	Thr	Ala	Leu	Gln	Thr	Arg	Val	225	230	235
Ser	Gly	Ala	Leu	Asn	Tyr	Asp	Thr	Ser	Gly	Asn	Leu	Glu	Leu	Arg	Ala	245	250	255
Ala	Gly	Gly	Met	Arg	Val	Asp	Ala	Asn	Gly	Gln	Leu	Ile	Leu	Asp	Val	260	265	270
Ala	Tyr	Pro	Phe	Asp	Ala	Gln	Asn	Asn	Leu	Ser	Leu	Arg	Leu	Gly	Gln	275	280	285
Gly	Pro	Leu	Phe	Val	Asn	Ser	Ala	His	Asn	Leu	Asp	Val	Asn	Tyr	Asn	290	295	300
Arg	Gly	Leu	Tyr	Leu	Phe	Thr	Ser	Gly	Asn	Thr	Lys	Lys	Leu	Glu	Val	305	310	315
Asn	Ile	Lys	Thr	Ala	Lys	Gly	Leu	Ile	Tyr	Asp	Asp	Thr	Ala	Ile	Ala	325	330	335
Ile	Asn	Ala	Gly	Asp	Gly	Leu	Gln	Phe	Asp	Ser	Gly	Ser	Asp	Thr	Asn	340	345	350
Pro	Leu	Lys	Thr	Lys	Leu	Gly	Leu	Gly	Leu	Asp	Tyr	Asp	Ser	Ser	Arg	355	360	365
Ala	Ile	Ile	Ala	Lys	Leu	Gly	Thr	Gly	Leu	Ser	Phe	Asp	Asn	Thr	Gly	370	375	380
Ala	Ile	Thr	Val	Gly	Asn	Lys	Asn	Asp	Asp	Lys	Leu	Thr	Leu	Trp	Thr	385	390	395
Thr	Pro	Asp	Pro	Ser	Pro	Asn	Cys	Arg	Ile	Tyr	Ser	Glu	Lys	Asp	Ala	405	410	415
Lys	Phe	Thr	Leu	Val	Leu	Thr	Lys	Cys	Gly	Ser	Gln	Val	Leu	Ala	Ser	420	425	430

Val Ser Val Leu Ser Val Lys Gly Ser Leu Ala Pro Ile Ser Gly Thr
435 440 445

Val Thr Ser Ala Gln Ile Val Leu Arg Phe Asp Glu Asn Gly Val Leu
450 455 460

Leu Ser Asn Ser Ser Leu Asp Pro Gln Tyr Trp Asn Tyr Arg Lys Gly
465 470 475 480

Asp Leu Thr Glu Gly Thr Ala Tyr Thr Asn Ala Val Gly Phe Met Pro
485 490 495

Asn Leu Thr Ala Tyr Pro Lys Thr Gln Ser Gln Thr Ala Lys Ser Asn
500 505 510

Ile Val Ser Gln Val Tyr Leu Asn Gly Asp Lys Ser Lys Pro Met Thr
515 520 525

Leu Thr Ile Thr Leu Asn Gly Thr Asn Glu Thr Gly Asp Ala Thr Val
530 535 540

Ser Thr Tyr Ser Met Ser Phe Ser Trp Asn Trp Asn Gly Ser Asn Tyr
545 550 555 560

Ile Asn Glu Thr Phe Gln Thr Asn Ser Phe Thr Phe Ser Tyr Ile Ala
565 570 575

Gln Glu

<210> 34
<211> 578
<212> PRT
<213> Pan troglodytes

<400> 34

Met Lys Arg Thr Lys Thr Ser Asp Glu Ser Phe Asn Pro Val Tyr Pro
1 5 10 15

Tyr Asp Thr Glu Ser Gly Pro Pro Ser Val Pro Phe Leu Thr Pro Pro
20 25 30

Phe Val Ser Pro Asp Gly Phe Gln Glu Ser Pro Pro Gly Val Leu Ser
35 40 45

Leu Asn Leu Ala Glu Pro Leu Val Thr Ser His Gly Met Leu Ala Leu
50 55 60

Lys	Met	Gly	Ser	Gly	Leu	Ser	Leu	Asp	Asp	Ala	Gly	Asn	Leu	Thr	Ser	
65					70					75					80	
Gln	Asp	Ile	Thr	Thr	Ala	Ser	Pro	Pro	Leu	Lys	Lys	Thr	Lys	Thr	Asn	
			85						90					95		
Leu	Ser	Leu	Glu	Thr	Ser	Ser	Pro	Leu	Thr	Val	Ser	Thr	Ser	Gly	Ala	
			100					105					110			
Leu	Thr	Val	Ala	Ala	Ala	Ala	Pro	Leu	Ala	Val	Ala	Gly	Thr	Ser	Leu	
		115					120					125				
Thr	Met	Gln	Ser	Glu	Ala	Pro	Leu	Thr	Val	Gln	Asp	Ala	Lys	Leu	Thr	
	130					135					140					
Leu	Ala	Thr	Lys	Gly	Pro	Leu	Thr	Val	Ser	Glu	Gly	Lys	Leu	Ala	Leu	
145					150					155					160	
Gln	Thr	Ser	Ala	Pro	Leu	Thr	Ala	Ala	Asp	Ser	Ser	Thr	Leu	Thr	Val	
				165					170					175		
Ser	Ala	Thr	Pro	Pro	Leu	Ser	Thr	Ser	Asn	Gly	Ser	Leu	Gly	Ile	Asp	
			180					185					190			
Met	Gln	Ala	Pro	Ile	Tyr	Thr	Thr	Asn	Gly	Lys	Leu	Gly	Leu	Asn	Phe	
		195					200					205				
Gly	Ala	Pro	Leu	His	Val	Val	Asp	Ser	Leu	Asn	Ala	Leu	Thr	Val	Val	
	210					215					220					
Thr	Gly	Gln	Gly	Leu	Thr	Ile	Asn	Gly	Thr	Ala	Leu	Gln	Thr	Arg	Val	
225					230					235					240	
Ser	Gly	Ala	Leu	Asn	Tyr	Asp	Thr	Ser	Gly	Asn	Leu	Glu	Leu	Arg	Ala	
				245					250					255		
Ala	Gly	Gly	Met	Arg	Val	Asp	Ala	Asn	Gly	Gln	Leu	Ile	Leu	Asp	Val	
			260					265					270			
Ala	Tyr	Pro	Phe	Asp	Ala	Gln	Asn	Asn	Leu	Ser	Leu	Arg	Leu	Gly	Gln	
		275					280					285				
Gly	Pro	Leu	Phe	Val	Asn	Ser	Ala	His	Asn	Leu	Asp	Val	Asn	Tyr	Asn	
	290					295					300					
Arg	Gly	Leu	Tyr	Leu	Phe	Thr	Ser	Gly	Asn	Thr	Lys	Lys	Leu	Glu	Val	
305					310					315					320	

Asn Ile Lys Thr Ala Lys Gly Leu Ile Tyr Asp Asp Thr Ala Ile Ala
325 330 335

Ile Asn Ala Gly Asp Gly Leu Gln Phe Asp Ser Gly Ser Asp Thr Asn
340 345 350

Pro Leu Lys Thr Lys Leu Gly Leu Gly Leu Asp Tyr Asp Ser Ser Arg
355 360 365

Ala Ile Ile Ala Lys Leu Gly Thr Gly Leu Ser Phe Asp Asn Thr Gly
370 375 380

Ala Ile Thr Val Gly Asn Lys Asn Asp Asp Lys Leu Thr Leu Trp Thr
385 390 395 400

Thr Pro Asp Pro Ser Pro Asn Cys Arg Ile Tyr Ser Glu Lys Asp Ala
405 410 415

Lys Phe Thr Leu Val Leu Thr Lys Cys Gly Ser Gln Val Leu Ala Ser
420 425 430

Val Ser Val Leu Ser Val Lys Gly Ser Leu Ala Pro Ile Ser Gly Thr
435 440 445

Val Thr Ser Ala Gln Ile Val Leu Arg Phe Asp Glu Asn Gly Val Leu
450 455 460

Leu Ser Asn Ser Ser Leu Asp Pro Gln Tyr Trp Asn Tyr Arg Lys Gly
465 470 475 480

Asp Leu Thr Glu Gly Thr Ala Tyr Thr Asn Ala Val Gly Phe Met Pro
485 490 495

Asn Leu Thr Ala Tyr Pro Lys Thr Gln Ser Gln Thr Ala Lys Ser Asn
500 505 510

Ile Val Ser Gln Val Tyr Leu Asn Gly Asp Lys Ser Lys Pro Met Thr
515 520 525

Leu Thr Ile Thr Leu Asn Gly Thr Asn Glu Thr Gly Asp Ala Thr Val
530 535 540

Ser Thr Tyr Ser Met Ser Phe Ser Trp Asn Trp Asn Gly Ser Asn Tyr
545 550 555 560

Ile Asn Glu Thr Phe Gln Thr Asn Ser Phe Thr Phe Ser Tyr Ile Ala
565 570 575

Gln Glu

<210> 35
<211> 577
<212> PRT
<213> Pan paniscus

<400> 35

Met	Lys	Arg	Ala	Lys	Thr	Ser	Asp	Glu	Thr	Phe	Asn	Pro	Val	Tyr	Pro
1				5					10					15	

Tyr	Asp	Thr	Glu	Asn	Gly	Pro	Pro	Ser	Val	Pro	Phe	Leu	Thr	Pro	Pro
			20					25					30		

Phe	Val	Ser	Pro	Asp	Gly	Phe	Gln	Glu	Ser	Pro	Pro	Gly	Val	Leu	Ser
		35					40					45			

Leu	Arg	Leu	Ser	Glu	Pro	Leu	Val	Thr	Ser	His	Gly	Met	Leu	Ala	Leu
	50						55				60				

Lys	Met	Gly	Asn	Gly	Leu	Ser	Leu	Asp	Asp	Ala	Gly	Asn	Leu	Thr	Ser
65					70					75					80

Gln	Asp	Val	Thr	Thr	Val	Thr	Pro	Pro	Leu	Lys	Lys	Thr	Lys	Thr	Asn
				85					90					95	

Leu	Ser	Leu	Gln	Thr	Ser	Ala	Pro	Leu	Thr	Val	Ser	Ser	Gly	Ser	Leu
			100					105					110		

Thr	Val	Ala	Ala	Ala	Ala	Pro	Leu	Ala	Val	Ala	Gly	Thr	Ser	Leu	Thr
		115					120					125			

Met	Gln	Ser	Gln	Ala	Pro	Leu	Thr	Val	Gln	Asp	Ala	Lys	Leu	Gly	Leu
	130					135					140				

Ala	Thr	Gln	Gly	Pro	Leu	Thr	Val	Ser	Glu	Gly	Lys	Leu	Thr	Leu	Gln
145					150					155					160

Thr	Ser	Ala	Pro	Leu	Thr	Ala	Ala	Asp	Ser	Ser	Thr	Leu	Thr	Val	Ser
				165					170					175	

Ala	Thr	Pro	Pro	Leu	Ser	Thr	Ser	Asn	Gly	Ser	Leu	Ser	Ile	Asp	Met
				180				185					190		

Gln	Ala	Pro	Ile	Tyr	Thr	Thr	Asn	Gly	Lys	Leu	Ala	Leu	Asn	Ile	Gly
		195					200					205			

Ala	Pro	Leu	His	Val	Val	Asp	Thr	Leu	Asn	Ala	Leu	Thr	Val	Val	Thr
	210					215					220				
Gly	Gln	Gly	Leu	Thr	Ile	Asn	Gly	Arg	Ala	Leu	Gln	Thr	Arg	Val	Thr
	225				230					235					240
Gly	Ala	Leu	Ser	Tyr	Asp	Thr	Glu	Gly	Asn	Ile	Gln	Leu	Gln	Ala	Gly
				245					250					255	
Gly	Gly	Met	Arg	Ile	Asp	Asn	Asn	Gly	Gln	Leu	Ile	Leu	Asn	Val	Ala
			260					265					270		
Tyr	Pro	Phe	Asp	Ala	Gln	Asn	Asn	Leu	Ser	Leu	Arg	Leu	Gly	Gln	Gly
		275						280				285			
Pro	Leu	Ile	Val	Asn	Ser	Ala	His	Asn	Leu	Asp	Leu	Asn	Leu	Asn	Arg
	290					295					300				
Gly	Leu	Tyr	Leu	Phe	Thr	Ser	Gly	Asn	Thr	Lys	Lys	Leu	Glu	Val	Asn
	305				310					315					320
Ile	Lys	Thr	Ala	Lys	Gly	Leu	Phe	Tyr	Asp	Gly	Thr	Ala	Ile	Ala	Ile
				325					330					335	
Asn	Ala	Gly	Asp	Gly	Leu	Gln	Phe	Gly	Ser	Gly	Ser	Asp	Thr	Asn	Pro
			340					345					350		
Leu	Gln	Thr	Lys	Leu	Gly	Leu	Gly	Leu	Glu	Tyr	Asp	Ser	Asn	Lys	Ala
		355					360					365			
Ile	Ile	Thr	Lys	Leu	Gly	Thr	Gly	Leu	Ser	Phe	Asp	Asn	Thr	Gly	Ala
	370					375					380				
Ile	Thr	Val	Gly	Asn	Lys	Asn	Asp	Asp	Lys	Leu	Thr	Leu	Trp	Thr	Thr
	385				390					395					400
Pro	Asp	Pro	Ser	Pro	Asn	Cys	Arg	Ile	Asn	Ser	Glu	Lys	Asp	Ala	Lys
				405					410					415	
Leu	Thr	Leu	Val	Leu	Thr	Lys	Cys	Gly	Ser	Gln	Val	Leu	Ala	Ser	Val
			420					425					430		
Ser	Val	Leu	Ser	Val	Lys	Gly	Ser	Leu	Ala	Pro	Ile	Ser	Gly	Thr	Val
		435					440					445			
Thr	Ser	Ala	Gln	Ile	Val	Leu	Arg	Phe	Asp	Glu	Asn	Gly	Val	Leu	Leu
						450					460				

Ser Asn Ser Ser Leu Asp Pro Gln Tyr Trp Asn Tyr Arg Lys Gly Asp
465 470 475 480

Ser Thr Glu Gly Thr Ala Tyr Thr Asn Ala Val Gly Phe Met Pro Asn
485 490 495

Leu Thr Ala Tyr Pro Lys Thr Gln Ser Gln Thr Ala Lys Ser Asn Ile
500 505 510

Val Ser Gln Val Tyr Leu Asn Gly Asp Lys Thr Lys Pro Met Thr Leu
515 520 525

Thr Ile Thr Leu Asn Gly Thr Asn Glu Thr Gly Asp Ala Thr Val Ser
530 535 540

Thr Tyr Ser Met Ser Phe Ser Trp Asn Trp Asn Gly Ser Asn Tyr Ile
545 550 555 560

Asn Asp Thr Phe Gln Thr Asn Ser Phe Thr Phe Ser Tyr Ile Ala Gln
565 570 575

Glu

<210> 36
<211> 577
<212> PRT
<213> Pan paniscus

<400> 36

Met Lys Arg Ala Lys Thr Ser Asp Glu Thr Phe Asn Pro Val Tyr Pro
1 5 10 15

Tyr Asp Thr Glu Asn Gly Pro Pro Ser Val Pro Phe Leu Thr Pro Pro
20 25 30

Phe Val Ser Pro Asp Gly Phe Gln Glu Ser Pro Pro Gly Val Leu Ser
35 40 45

Leu Arg Leu Ser Glu Pro Leu Val Thr Ser His Gly Met Leu Ala Leu
50 55 60

Lys Met Gly Asn Gly Leu Ser Leu Asp Asp Ala Gly Asn Leu Thr Ser
65 70 75 80

Gln Asp Val Thr Thr Val Thr Pro Pro Leu Lys Lys Thr Lys Thr Asn
85 90 95

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

		100						105				110			
Thr	Val	Ala	Ala	Ala	Ala	Pro	Leu	Ala	Val	Ala	Gly	Thr	Ser	Leu	Thr
		115					120					125			
Met	Gln	Ser	Gln	Ala	Pro	Leu	Thr	Val	Gln	Asp	Ala	Lys	Leu	Gly	Leu
	130					135					140				
Ala	Thr	Gln	Gly	Pro	Leu	Thr	Val	Ser	Glu	Gly	Lys	Leu	Thr	Leu	Gln
145					150					155					160
Thr	Ser	Ala	Pro	Leu	Thr	Ala	Ala	Asp	Ser	Ser	Thr	Leu	Thr	Val	Ser
				165					170					175	
Ala	Thr	Pro	Pro	Leu	Ser	Thr	Ser	Asn	Gly	Ser	Leu	Ser	Ile	Asp	Met
			180					185					190		
Gln	Ala	Pro	Ile	Tyr	Thr	Thr	Asn	Gly	Lys	Leu	Ala	Leu	Asn	Ile	Gly
		195					200					205			
Ala	Pro	Leu	His	Val	Val	Asp	Thr	Leu	Asn	Ala	Leu	Thr	Val	Val	Thr
	210					215					220				
Gly	Gln	Gly	Leu	Thr	Ile	Asn	Gly	Arg	Ala	Leu	Gln	Thr	Arg	Val	Thr
225					230					235					240
Gly	Ala	Leu	Ser	Tyr	Asp	Thr	Glu	Gly	Asn	Ile	Gln	Leu	Gln	Ala	Gly
				245					250					255	
Gly	Gly	Met	Arg	Ile	Asp	Asn	Asn	Gly	Gln	Leu	Ile	Leu	Asn	Val	Ala
			260					265					270		
Tyr	Pro	Phe	Asp	Ala	Gln	Asn	Asn	Leu	Ser	Leu	Arg	Leu	Gly	Gln	Gly
		275					280					285			
Pro	Leu	Ile	Val	Asn	Ser	Ala	His	Asn	Leu	Asp	Leu	Asn	Leu	Asn	Arg
	290					295					300				
Gly	Leu	Tyr	Leu	Phe	Thr	Ser	Gly	Asn	Thr	Lys	Lys	Leu	Glu	Val	Asn
305					310					315					320
Ile	Lys	Thr	Ala	Lys	Gly	Leu	Phe	Tyr	Asp	Gly	Thr	Ala	Ile	Ala	Ile
				325					330					335	
Asn	Ala	Gly	Asp	Gly	Leu	Gln	Phe	Gly	Ser	Gly	Ser	Asp	Thr	Asn	Pro
			340					345					350		
Leu	Gln	Thr	Lys	Leu	Gly	Leu	Gly	Leu	Glu	Tyr	Asp	Ser	Asn	Lys	Ala

355

360

365

Ile Ile Thr Lys Leu Gly Thr Gly Leu Ser Phe Asp Asn Thr Gly Ala
 370 375 380

Ile Thr Val Gly Asn Lys Asn Asp Asp Lys Leu Thr Leu Trp Thr Thr
 385 390 395 400

Pro Asp Pro Ser Pro Asn Cys Arg Ile Asn Ser Glu Lys Asp Ala Lys
 405 410 415

Leu Thr Leu Val Leu Thr Lys Cys Gly Ser Gln Val Leu Ala Ser Val
 420 425 430

Ser Val Leu Ser Val Lys Gly Ser Leu Ala Pro Ile Ser Gly Thr Val
 435 440 445

Thr Ser Ala Gln Ile Val Leu Arg Phe Asp Glu Asn Gly Val Leu Leu
 450 455 460

Ser Asn Ser Ser Leu Asp Pro Gln Tyr Trp Asn Tyr Arg Lys Gly Asp
 465 470 475 480

Ser Thr Glu Gly Thr Ala Tyr Thr Asn Ala Val Gly Phe Met Pro Asn
 485 490 495

Leu Thr Ala Tyr Pro Lys Thr Gln Ser Gln Thr Ala Lys Ser Asn Ile
 500 505 510

Val Ser Gln Val Tyr Leu Asn Gly Asp Lys Thr Lys Pro Met Thr Leu
 515 520 525

Thr Ile Thr Leu Asn Gly Thr Asn Glu Thr Gly Asp Ala Thr Val Ser
 530 535 540

Thr Tyr Ser Met Ser Phe Ser Trp Asn Trp Asn Gly Ser Asn Tyr Ile
 545 550 555 560

Asn Asp Thr Phe Gln Thr Asn Ser Phe Thr Phe Ser Tyr Ile Ala Gln
 565 570 575

Glu

<210> 37

<211> 845

<212> DNA

<213> Human cytomegalovirus

<400> 37

ccattgcata cgttgtatcc atatcataat atgtacat	ttt atattggctc atgtccaaca	60
ttaccgccat gttgacattg attattgact agttattaat	agtaatcaat tacgggggtca	120
ttagttcata gcccatatat ggagttccgc gttacataac	ttacggtaaa tggcccgct	180
ggctgaccgc ccaacgaccc ccgcccattg acgtcaataa	tgacgtatgt tcccatagta	240
acgccaatag ggactttcca ttgacgtcaa tgggtggagt	atttacggta aactgccac	300
ttggcagtac atcaagtgt tcatatgcc agtacgccc	ctattgacgt caatgacggt	360
aaatggcccg cctggcatta tgcccagtac atgaccttat	gggactttcc tacttggcag	420
tacatctacg tattagtcac cgctattacc atgggtgatgc	ggttttggca gtacatcaat	480
gggcgtggat agcggtttga ctcacgggga tttccaagtc	tccaccccat tgacgtcaat	540
gggagtttgt tttggcacca aaatcaacgg gactttccaa	aatgtcgtaa caactccgcc	600
ccattgacgc aaatgggcgg taggcgtgta cgggtgggagg	tctatataag cagagctctc	660
cctatcagtg atagagatct ccctatcagt gatagagatc	gtcgacgagc tcgttttagtg	720
aaccgtcaga tcgcctggag acgccatcca cgctgttttg	acctccatag aagacaccgg	780
gaccgatcca gcctccgcg cggggaacgg tgcattggaa	cgcggtattcc ccgtgccaa	840
agtga		845

<210> 38
 <211> 15422
 <212> DNA
 <213> Pan troglodytes

<400> 38

tgcccgcttt ccagtcggga aacctgtcgt gccagctgca	ttaatgaatc ggccaacgcg	60
cggggagagg cggtttgcgt attgggcgct agatctccca	tcacatatac ctgccgttca	120
ctattattta gtgaaatgag atattatgat attttctgaa	ttgtgattaa aaaggcaact	180
ttatgcccat gcaacagaaa ctataaaaaa tacagagaat	gaaaagaaac agatagattt	240
tttagttctt taggcccgta gtctgcaaat ctttttatga	ttttctatca aacaaaagag	300
gaaaatagac cagttgcaat ccaaacgaga gtctaataga	atgaggtcga aaagtaaadc	360
gcgcggggtt gttactgata aagcaggcaa gacctaaat	gtgtaaaggg caaagtgtat	420
actttggcgt cacccttac atatttttagg tcttttttta	ttgtgcgtaa ctaacttgcc	480
atcttcaaac aggagggctg gaagaagcag accgctaaca	cagtacataa aaaaggagac	540
atgaacgatg aacatcaaaa agtttgcaaa acaagcaaca	gtattaacct ttactaccgc	600
actgctggca ggaggcgcaa ctcaagcggt tgcgaaagaa	acgaaccaa agccatataa	660
ggaaacatac ggcatttccc atattacacg ccatgatatg	ctgcaaattcc ctgaacagca	720
aaaaaatgaa aaatatcaag ttcctgaatt cgattcgtcc	acaattaaaa atatctcttc	780

tgcaaaaggc	ctggacgttt	gggacagctg	gccattacaa	aacgctgacg	gcactgtcgc	840
aaactatcac	ggctaccaca	tcgtcttttc	attagccgga	gacctaataa	atgcggatga	900
cacatcgatt	tacatgttct	atcaaaaagt	cggcgaaact	tctattgaca	gctggaaaaa	960
cgctggccgc	gtcttttaaag	acagcgacaa	attcgatgca	aatgattcta	tcctaataaga	1020
ccaaacacaa	gaatgggtcag	gttcagccac	atttacatct	gacggaaaaa	tccgtttatt	1080
ctacactgat	ttctccggta	aacattacgg	caaacaaaca	ctgacaactg	cacaagttaa	1140
cgtatcagca	tcagacagct	ctttgaacat	caacgggtgta	gaggattata	aatcaatctt	1200
tgacggtgac	ggaaaaacgt	atcaaaatgt	acagcagttc	atcgatgaag	gcaactacag	1260
ctcaggcgac	aaccatacgc	tgagagatcc	tcactacgta	gaagataaag	gccacaaata	1320
cttagtattt	gaagcaaaca	ctggaactga	agatggctac	caaggcgaag	aatctttatt	1380
taacaaagca	tactatggca	aaagcacatc	attcttccgt	caagaaagtc	aaaaacttct	1440
gcaaagcgat	aaaaaacgca	cggctgagtt	agcaaacggc	gctctcggta	tgattgagct	1500
aaacgatgat	tacacactga	aaaaagtgat	gaaaccgctg	attgcatcta	acacagtaac	1560
agatgaaatt	gaacgcgcga	acgtctttta	aatgaacggc	aaatgggtacc	tgttcactga	1620
ctcccgcgga	tcaaaaatga	cgattgacgg	cattacgtct	aacgatattt	acatgcttgg	1680
ttatgtttct	aattctttta	ctggcccata	caagccgctg	aacaaaactg	gccttgtgtt	1740
aaaaatggat	cttgatccta	acgatgtaac	ctttacttac	tcacacttcg	ctgtacctca	1800
agcgaaagga	aacaatgtcg	tgattacaag	ctatatgaca	aacagaggat	tctacgcaga	1860
caaacaatca	acgtttgcgc	caagcttcct	gctgaacatc	aaaggcaaga	aaacatctgt	1920
tgtcaaagac	agcatccttg	aacaaggaca	attaacagtt	aacaaataat	agggataaca	1980
gggtaatgct	agaagacccg	agtcttacca	gtaaaagaaa	aaagatctct	caacgcagca	2040
ccagcaccaa	cacttcgcag	tgtaaaaggc	caagtgccga	gagagtatat	ataggaataa	2100
aaagtgacgt	aaacgggcaa	agtccaaaaa	acgccagaaa	aaaccgcacg	cgaacctacg	2160
ccccgaaacg	aaagccaaaa	aacactagac	actcccttcc	ggcgtcaact	tccgctttcc	2220
cacgctacgt	cacttgcccc	agtcaaacaa	actacatata	ccgaacttcc	aagtcgccac	2280
gccccaaaaca	ccgcctacac	ctccccgccc	gccggcccg	ccccaaacct	gcctcccgc	2340
ccgcgccccg	ccccgcgcgc	cccatctcat	tatcatattg	gcttcaatcc	aaaataagg	2400
atattattga	tgatggttta	aacggatcct	ctagagtcga	cctgcaggca	tgcaagcttg	2460
agtattctat	agtgtcacct	aaatagcttg	gcgtaatcat	ggtcatactg	gtttcctgtg	2520
tgaaattgtt	atccgctcac	aattccacac	aacatacgag	ccggaagcat	aaagtgtaaa	2580
gcctgggggtg	cctaattgagt	gagctaactc	acattaattg	cgttgcgctc	actgcccgc	2640
ttccagtcgg	gaaacctgtc	gtgccagctg	cattaatgaa	tcggccaacg	cgaacccctt	2700

gcgggccgccc	gggcccgtcga	ccaattctca	tgtttgacag	cttatcatcg	aattttctgcc	2760
attcatccgc	ttattatcac	ttattcaggc	gtagcaacca	ggcgtttaag	ggcaccaata	2820
actgccttaa	aaaaattacg	ccccgcctg	ccactcatcg	cagtactggt	gtaattcatt	2880
aagcattctg	ccgacatgga	agccatcaca	aacggcatga	tgaacctgaa	tcgccagcgg	2940
catcagcacc	ttgtcgcctt	gcgtataata	tttgcccatg	gtgaaaacgg	gggcgaagaa	3000
gttgtccata	ttggccacgt	ttaaatcaaa	actggtgaaa	ctcaccacgg	gattggctga	3060
gacgaaaaac	atattctcaa	taaacccttt	agggaaatag	gccagggtttt	caccgtaaca	3120
cgccacatct	tgcgaaatata	tgtgtagaaa	ctgccggaaa	tcgtcgtggt	attcactcca	3180
gagcgatgaa	aacgtttcag	tttgctcatg	gaaaacgggt	taacaagggt	gaacactatc	3240
ccatatcacc	agctcacctg	ctttcattgc	catacggaat	tccggatgag	cattcatcag	3300
gcgggcaaga	atgtgaataa	aggccggata	aaacttgtgc	ttatTTTTTct	ttacgggtctt	3360
taaaaaggcc	gtaatatcca	gctgaacggg	ctgggttatag	gtacattgag	caactgactg	3420
aaatgcctca	aaatgttctt	tacgatgcca	ttgggatata	tcaacgggtgg	tatatccagt	3480
gatttttttc	tccatttttag	cttccttagc	tcctgaaaat	ctcgataact	caaaaaatac	3540
gcccggtagt	gatcttattt	cattatgggtg	aaagttggaa	cctcttacgt	gccgatcaac	3600
gtctcatttt	cgccaaaagt	tggcccaggg	cttcccggta	tcaacaggga	caccaggatt	3660
tatttattct	gcgaagtgat	cttccgtcac	aggtatttat	tcgcgataag	ctcatggagc	3720
ggcgtaaccg	tcgcacagga	aggacagaga	aagcgcggat	ctgggaagtg	acggacagaa	3780
cggtcaggac	ctggattggg	gaggcggttg	ccgccgctgc	tgctgacggg	gtgacgttct	3840
ctgttccggg	cacaccacat	acgttccgcc	attcctatgc	gatgcacatg	ctgtatgccg	3900
gtataccgct	gaaagtcttg	caaagcctga	tgggacataa	gtccatcagt	tcaacggaag	3960
tctacacgaa	ggtttttgcg	ctggatgtgg	ctgcccggca	ccgggtgcag	tttgcgatgc	4020
cggagtctga	tgcggttgcg	atgctgaaac	aattatcctg	agaataaatg	ccttggcctt	4080
tatatggaaa	tgtggaactg	agtggatatg	ctgtttttgt	ctgttaaaca	gagaagctgg	4140
ctgttatcca	ctgagaagcg	aacgaaacag	tcgggaaaat	ctcccattat	cgtagagatc	4200
cgcattatta	atctcaggag	cctgtgtagc	gtttatagga	agtagtggtc	tgtcatgatg	4260
cctgcaagcg	gtaacgaaaa	cgatttgaat	atgccttcag	gaacaataga	aatcttcgtg	4320
cgggtgttac	ttgaagtgga	gcggattatg	tcagcaatgg	acagaacaac	ctaatgaaca	4380
cagaaccatg	atgtggtctg	tcctttttaca	gccagtagtg	ctcgccgcag	tcgagcgaca	4440
gggcgaagcc	ctcgagtgag	cgaggaagca	ccagggaaca	gcacttatat	attctgctta	4500
cacacgatgc	ctgaaaaaac	ttcccttggg	gttatccact	tatccacggg	gatattttta	4560
taattatttt	ttttatagtt	tttagatctt	ctttttttaga	gcgccttgta	ggcctttatc	4620

catgctgggt	ctagagaagg	tgttgtgaca	aattgccctt	tcagtgtgac	aaatcacccct	4680
caaatgacag	tcctgtctgt	gacaaattgc	ccttaaccct	gtgacaaatt	gccctcagaa	4740
gaagctgttt	tttcacaaag	ttatccctgc	ttattgactc	ttttttat	agtgtgacaa	4800
tctaaaaact	tgtcacactt	cacatggatc	tgtcatggcg	gaaacagcgg	ttatcaatca	4860
caagaaacgt	aaaaatagcc	cgcgaatcgt	ccagtcaaac	gacctcactg	aggcggcata	4920
tagtctctcc	cgggatcaaa	aacgtatgct	gtatctgttc	gttgaccaga	tcagaaaatc	4980
tgatggcacc	ctacaggaac	atgacgggat	ctgcgagatc	catgttgcta	aatatgctga	5040
aatattcgga	ttgacctctg	cgggaagccag	taaggatata	cggcaggcat	tgaagagttt	5100
cgcggggaag	gaagtgggtt	tttatcgccc	tgaagaggat	gccggcgatg	aaaaaggcta	5160
tgaatctttt	ccttggttta	tcaaacgtgc	gcacagtcca	tccagagggc	tttacagtgt	5220
acatatcaac	ccatatctca	ttcccttctt	tatcgggtta	cagaaccggg	ttacgcagtt	5280
tcggcttagt	gaaacaaaag	aaatcaccaa	tccgtatgcc	atgcgtttat	acgaatccct	5340
gtgtcagtat	cgtaagccgg	atggctcagg	catcgtctct	ctgaaaatcg	actggatcat	5400
agagcgttac	cagctgcctc	aaagttacca	gcgtatgcct	gacttccgcc	gccgcttcc	5460
gcaggtctgt	gttaatgaga	tcaacagcag	aactccaatg	cgcctctcat	acattgagaa	5520
aaagaaaggc	cgccagacga	ctcatatcgt	atcttccttc	cgcgatatca	cttccatgac	5580
gacaggatag	tctgagggtt	atctgtcaca	gatttgaggg	tggttcgtca	catttgttct	5640
gacctactga	gggtaatttg	tcacagtttt	gctgtttcct	tcagcctgca	tggattttct	5700
catacttttt	gaactgtaat	ttttaaggaa	gccaaatttg	agggcagttt	gtcacagtgt	5760
atcttcctct	ctttcccttc	gtcatgtgac	ctgatatcgg	gggttagttc	gtcatcattg	5820
atgagggttg	attatcacag	tttattactc	tgaattggct	atccgcgtgt	gtacctctac	5880
ctggagtttt	tcccacggtg	gatatttctt	cttgcgctga	gcgtaagagc	tatctgacag	5940
aacagtctct	ctttgcttcc	tcgccagttc	gctcgtctatg	ctcggttaca	cggtgcgggc	6000
gagcgctagt	gataataagt	gactgaggta	tgtgctcttc	ttatctcctt	ttgtagtgtt	6060
gctcttattt	taaacaactt	tgcgggtttt	tgatgacttt	gcgattttgt	tgttgctttg	6120
cagtaaattg	caagatttaa	taaaaaaacg	caaagcaatg	attaaaggat	gttcagaatg	6180
aaactcatgg	aaacacttaa	ccagtgcata	aacgctggtc	atgaaatgac	gaaggctatc	6240
gccattgcac	agtttaatga	tgacagcccg	gaagcgagga	aaataacccg	gcgctggaga	6300
ataggtgaag	cagcggattt	agttgggggt	tcttctcagg	ctatcagaga	tgccgagaaa	6360
gcagggcgac	taccgcaccc	ggatatggaa	attcgaggac	gggttgagca	acgtgttggt	6420
tatacaattg	aacaaattaa	tcatatgcgt	gatgtgtttg	gtacgcgatt	gcgacgtgct	6480
gaagacgtat	ttccaccggg	gatcgggggt	gctgcccata	aaggtggcgt	ttacaaaacc	6540

tcagtttctg	ttcatcttgc	tcaggatctg	gctctgaagg	ggctacgtgt	tttgctcgtg	6600
gaaggtaacg	acccccaggg	aacagcctca	atgtatcacg	gatgggtacc	agatcttcat	6660
attcatgcag	aagacactct	cctgcctttc	tatcttgggg	aaaaggacga	tgtcacttat	6720
gcaataaagc	ccacttgctg	gccggggcct	gacattattc	cttcctgtct	ggctctgcac	6780
cgtattgaaa	ctgagttaat	gggcaaattt	gatgaaggta	aactgcccac	cgatccacac	6840
ctgatgctcc	gactggccat	tgaaactgtt	gctcatgact	atgatgtcat	agttattgac	6900
agcgcgccta	acctgggtat	cggcacgatt	aatgtcgtat	gtgctgctga	tgtgctgatt	6960
gttcccacgc	ctgctgagtt	gtttgactac	acctccgcac	tgcagttttt	cgatatgctt	7020
cgtgatctgc	tcaagaacgt	tgatcttaaa	gggttcgagc	ctgatgtacg	tattttgctt	7080
accaaataca	gcaatagtaa	tggctctcag	tccccgtgga	tggaggagca	aattcgggat	7140
gcctggggaa	gcatggttct	aaaaaatgtt	gtacgtgaaa	cggatgaagt	tggtaaaggt	7200
cagatccgga	tgagaactgt	ttttgaacag	gccattgatc	aacgctcttc	aactggtgcc	7260
tggagaaatg	ctctttctat	ttgggaacct	gtctgcaatg	aaattttcga	tcgtctgatt	7320
aaaccacgct	gggagattag	ataatgaagc	gtgcgcctgt	tattccaaaa	catacgctca	7380
atactcaacc	ggttgaagat	acttcgttat	cgacaccagc	tgccccgatg	gtggattcgt	7440
taattgcgcg	cgtaggagta	atggctcgcg	gtaatgcat	tactttgcct	gtatgtggtc	7500
gggatgtgaa	gtttactctt	gaagtgtctc	ggggtgatag	tgttgagaag	acctctcggg	7560
tatggtcagg	taatgaacgt	gaccaggagc	tgcttactga	ggacgcactg	gatgatctca	7620
tcccttcttt	tctactgact	ggtcaacaga	caccggcggt	cggtcgaaga	gtatctgggtg	7680
tcatagaaat	tgccgatggg	agtcgccgtc	gtaaagctgc	tgcacttacc	gaaagtgatt	7740
atcgtgttct	ggttggcgag	ctggatgatg	agcagatggc	tgcattatcc	agattgggta	7800
acgattatcg	ccaacaagt	gcttatgaac	gtggtcagcg	ttatgcaagc	cgattgcaga	7860
atgaatttgc	tggaaatatt	tctgcgctgg	ctgatgcgga	aaatatttca	cgtaagatta	7920
ttacccgctg	tatcaacacc	gccaaattgc	ctaaatcagt	tgttgctctt	ttttctcacc	7980
ccggtgaact	atctgcccgg	tcaggatgat	cacttcaaaa	agcctttaca	gataaagagg	8040
aattacttaa	gcagcaggca	tctaaccctc	atgagcagaa	aaaagctggg	gtgatatttg	8100
aagctgaaga	agttatcact	cttttaactt	ctgtgcttaa	aacgtcatct	gcatcaagaa	8160
ctagtttaag	ctcacgacat	cagtttgctc	ctggagcgac	agtattgtat	aagggcgata	8220
aaatggtgct	taacctggac	aggtctcgtg	ttccaactga	gtgtatagag	aaaattgagg	8280
ccattcttaa	ggaacttgaa	aagccagcac	cctgatgcga	ccacgtttta	gtctacgttt	8340
atctgtcttt	acttaatgtc	ctttgttaca	ggccagaaag	cataactggc	ctgaatatct	8400
tctctggggc	cactgttcca	cttgtatcgt	cggctctgata	atcagactgg	gaccacgggtc	8460

ccactcgtat	cgtcggtctg	attattagtc	tgggaccacg	gtcccactcg	tatcgtcggt	8520
ctgattatta	gtctgggacc	acggtcccac	tcgtatcgtc	ggtctgataa	tcagactggg	8580
accacggtcc	cactcgtatc	gtcgggtctga	ttattagtc	gggaccatgg	tcccactcgt	8640
atcgtcggtc	tgattattag	tctgggacca	cgggtcccact	cgtatcgtcg	gtctgattat	8700
tagtctggaa	ccacgggtccc	actcgtatcg	tcgggtctgat	tattagtc	ggaccacggt	8760
cccactcgta	tcgtcgggtct	gattattagt	ctgggaccac	gateccactc	gtgttgctgg	8820
tctgattatc	ggtctgggac	cacgggtccca	cttgtattgt	cgatcagact	atcagcgtga	8880
gactacgatt	ccatcaatgc	ctgtcaaggg	caagtattga	catgtcgtcg	taacctgtag	8940
aacggagtaa	cctcgggtgtg	cggttgtatg	cctgctgtgg	attgctgctg	tgtcctgctt	9000
atccacaaca	ttttgcgcac	ggttatgtgg	acaaaatacc	tggttacca	ggcgtgccg	9060
gcacgttaac	cggtctgcat	ccgatgcaag	tgtgtcgtcg	tcgacgagct	cgcgagctcg	9120
gacatgaggt	tgccccgtat	tcagtgtcgc	tgatttgtat	tgtctgaagt	tgtttttacg	9180
ttaagttgat	gcagatcaat	taatacgata	cctgcgtcat	aattgattat	ttgacgtggt	9240
ttgatggcct	ccacgcacgt	tgtgatatgt	agatgataat	cattatcact	ttacgggtcc	9300
tttcgggtga	tccgacaggt	tacggggcgg	cgcctcgcg	ggttttcgct	atztatgaaa	9360
attttcgggt	ttaaggcgtt	tccgttcttc	ttcgtcataa	cttaatgttt	ttatttaaaa	9420
taccctctga	aaagaaagga	aacgacaggt	gctgaaagcg	agcttttttg	cctctgtcgt	9480
ttcctttctc	tgtttttgtc	cgtggaatga	acaatggaag	tccgagctca	tcgctaataa	9540
cttcgtatag	catacattat	acgaagtatt	attcgatgcg	gccgcaaggg	gttcgcgtca	9600
gcgggtgttg	gcgggtgtcg	gggctggctt	aactatgcgg	catcagagca	gattgtactg	9660
agagtgcacc	atatgcggtg	tgaaataccg	cacagatgcg	taaggagaaa	ataccgcac	9720
aggcgccatt	cgccattcag	gctgcgcaac	tggtgggaag	ggcgatcggg	gcgggcctct	9780
tcgctattac	gccagctggc	gaaaggggga	tgtgctgcaa	ggcgattaag	ttgggtaacg	9840
ccagggtttt	cccagtcacg	acgttgtaaa	acgacggcca	gtgaattgta	atacgactca	9900
ctatagggcg	aattcgagct	cggatcccgg	ggatcctcgt	ttaaaccatc	atcaataata	9960
taccttattt	tggtattgaag	ccaatatgat	aatgagatgg	gcggcgcggg	gcggggcgcg	10020
gggcgggagg	cgggttttgg	ggcgggcccg	cgggcggggc	ggtgtggcgg	aagtggactt	10080
tgtaagtgtg	gcggatgtga	cttgctagt	cgggcgcggg	taaaagtga	gttttcctgt	10140
cgcgacaacg	ccccgggaa	gtgacatttt	tcccgcggtt	tttaccggat	gttgtagtga	10200
atltgggcgt	aaccaagtaa	gatttgcca	ttttcgcggg	aaaactgaaa	cggggaagt	10260
aaatctgatt	aattttgcgt	tagtcatacc	gcgtaatat	tgtctagggc	cagaggactt	10320
tggccgatta	cgtggaggac	tcgcccaggt	gttttttgag	gtgaatttcc	gcgttcgggg	10380

tcaaagtctg	cgttttatta	ttataggata	tcccattgca	tacgttgtat	ccatatcata	10440
atatgtacat	ttatattggc	tcatgtccaa	cattaccgcc	atgttgacat	tgattattga	10500
ctagttatta	atagtaatca	attacggggt	cattagttca	tagcccatat	atggagttcc	10560
gcgttacata	acttacggta	aatggccccg	ctggctgacc	gcccacacgac	ccccgcccac	10620
tgacgtcaat	aatgacgtat	gttcccatag	taacgccaat	agggactttc	cattgacgtc	10680
aatgggtgga	gtattttacgg	taaactgccc	acttggcagt	acatcaagtg	tatcatatgc	10740
caagtacgcc	ccctattgac	gtcaatgacg	gtaaatggcc	cgcttgccat	tatgcccagt	10800
acatgacctt	atgggacttt	cctacttggc	agtacatcta	cgtattagtc	atcgctatta	10860
ccatggtgat	gcggttttgg	cagtacatca	atgggcgtgg	atagcggttt	gactcacggg	10920
gatttccaag	tctccacccc	attgacgtca	atgggagttt	gttttggcac	caaaatcaac	10980
gggactttcc	aaaatgtcgt	aacaactccg	ccccattgac	gcaaatgggc	ggtaggcgtg	11040
tacggtggga	ggtctatata	agcagagctc	tccctatcag	tgatagagat	ctccctatca	11100
gtgatagaga	tcgtcgacga	gctcgtttag	tgaaccgtca	gatcgccctg	agacgccatc	11160
cacgctgttt	tgacctccat	agaagacacc	gggaccgatc	cagcctccgc	ggccgggaac	11220
ggtgcattgg	aacgcggatt	ccccgtgcca	agagttagat	cttcgcgttt	tctaggtacc	11280
gggccccccc	tcgaggtcga	cggtatcgat	aagcttcacg	ctgccgcaag	cactcagggc	11340
gcaagggctg	ctaaaggaag	cggaacacgt	agaaagccag	tccgcagaaa	cggtgctgac	11400
cccggatgaa	tgtcagctac	tgggctatct	ggacaaggga	aaacgcaagc	gcaaagagaa	11460
agcaggtagc	ttgcagtggg	cttacatggc	gatagctaga	ctgggcgggt	ttatggacag	11520
caagcgaacc	ggaattgcca	gctggggcgc	cctctggtaa	ggttgggaag	ccctgcaaag	11580
taaactggat	ggctttcttg	ccgccaagga	tctgatggcg	caggggatca	agatctaacc	11640
aggagctatt	taatggcaac	agttaaccag	ctggtacgca	aaccacgtgc	tcgcaaagtt	11700
gcgaaaagca	acgtgcctgc	gctggaagca	tgcccgcata	aacgtggcgt	atgtactcgt	11760
gtatatacta	ccactcctaa	aaaaccgaac	tccgcgctgc	gtaaagtatg	ccgtgttcgt	11820
ctgactaacg	gtttcgaagt	gacttcctac	atcggtggtg	aaggtcacaa	cctgcaggag	11880
cactccgtga	tcctgatccg	tggcggtcgt	gttaaagacc	tcccgggtgt	tcgttaccac	11940
accgtacgtg	gtgcgcttga	ctgctccggc	gttaaagacc	gtaagcaggc	tcgttccaag	12000
tatggcgtga	agcgtcctaa	ggcttaatgg	tagatctgat	caagagacag	gatgacggtc	12060
gtttcgcatg	cttgaacaag	atggattgca	cgcaggttct	ccggccgctt	gggtggagag	12120
gctattcggc	tatgactggg	cacaacagac	aatcggtcgc	tctgatgccg	ccgtgttcgg	12180
gctgtcagcg	caggggcgcc	cggttctttt	tgtcaagacc	gacctgtccg	gtgccctgaa	12240
tgaactgcag	gacgaggcag	cgcggtatct	gtggctggcc	acgacggggc	ttccttgcgc	12300

agctgtgctc	gacgttgtca	ctgaagcggg	aagggactgg	ctgctattgg	gcgaagtgcc	12360
ggggcaggat	ctcctgtcat	ctcaccttgc	tcttgccgag	aaagtatcca	tcatggctga	12420
tgcaatgcgg	cggctgcata	cgcttgatcc	ggctacctgc	ccattcgacc	accaagcgaa	12480
acatcgcatc	gagcgagcac	gtactcggat	ggaagccggt	cttgtcgata	aggatgatct	12540
ggacgaagag	catcaggggc	tcgcgccagc	cgaactgttc	gccaggctca	aggcgcgcat	12600
gcccgcaggc	gaggatctcg	tcgtgaccca	tggcgatgcc	tgcttgccga	atatcatggg	12660
ggaaaatggc	cgcttttctg	gattcatcga	ctgtggccgg	ctgggtgtgg	cggaccgcta	12720
tcaggacata	gcgttggcta	cccgtgatat	tgctgaagag	cttggcggcg	aatgggctga	12780
ccgcttcctc	gtgctttacg	gtatcgccgc	tcccgattcg	cagcgcatcg	ccttctatcg	12840
ccttcttgac	gagttcttct	gagcgggact	ctggggttcg	aaatgaccga	ccaagcgacg	12900
cccaacctgc	catcacgaga	tttcgattcc	accgccgcct	tctatgaaag	gttgggcttc	12960
ggaatcgttt	tccgggacgc	cggctggatg	atcctccagc	gcggggatct	catgctggag	13020
ttcttcgccc	accccgggct	cgatccctc	ggggggaatc	agaattcagt	cgacagcggc	13080
cgcgatctgc	tgtgccttct	agttgccagc	catctgttgt	ttgcccctcc	cccgtgcctt	13140
ccttgaccct	ggaaggtgcc	actcccactg	tcttttccta	ataaaatgag	gaaattgcat	13200
cgcatgtctc	gagtaggtgt	cattctattc	tgggggggtg	ggtggggcag	gacagcaagg	13260
gggaggattg	ggaagacaat	agcaggcatg	ctggggatgc	ggtgggctct	atggccgata	13320
agcgatcgct	gaggtgggtg	agtgggcgtg	gcctgggggtg	gtcatgaaaa	tatataagtt	13380
gggggtctta	gggtctcttt	atltgtgttg	cagagaccgc	cggagccatg	agcgggagca	13440
gcagcagcag	cagtagcagc	agcgccttgg	atggcagcat	cgtgagccct	tatttgacga	13500
cgcggatgcc	ccactgggcc	ggggtgcgtc	agaatgtgat	gggctccagc	atcgacggcc	13560
gaccgctcct	gcccgcaaat	tccgccacgc	tgacctatgc	gaccgtcgcg	gggacgccgt	13620
tggacgccac	cgccgccgcc	gccgccaccg	cagccgcctc	ggcgtgcgc	agcctggcca	13680
cggactttgc	attcctggga	ccactggcga	caggggctac	ttctcggggc	gctgctgccg	13740
ccgttcgcga	tgacaagctg	accgccctgc	tggcgcagtt	ggatgcgctt	actcgggaac	13800
tgggtgacct	ttctcagcag	gtcatggccc	tcgccagca	ggtctcctcc	ctgcaagctg	13860
gcgggaatgc	ttctcccaca	aatgccgttt	aagggcgcgc	ctagggataa	cagggtaata	13920
cccctatttg	tttatttttc	taaatacatt	caaatatgta	tccgctcatg	agacaataac	13980
cctgataaat	gcttcaataa	tattgaaaaa	ggaagagtat	gagtattcaa	catttccgtg	14040
tcgcccttat	tccctttttt	gcggcatttt	gccttcctgt	ttttgctcac	ccagaaacgc	14100
tggtgaaagt	aaaagatgct	gaagatcagt	tgggtgcacg	agtgggttac	atcgaactgg	14160
atctcaacag	cggtaagatc	cttgagagtt	ttcgccccga	agaacgtttt	ccaatgatga	14220

gcacttttta	agttctgcta	tgtggcgcg	tattatcccg	tattgacgcc	gggcaagagc	14280
aactcgggtc	ccgcatacac	tattctcaga	atgacttgg	tgagtactca	ccagtcacag	14340
aaaagcatct	tacggatggc	atgacagtaa	gagaattatg	cagtgtctgc	ataaccatga	14400
gtgataacac	tgcggccaac	ttactttctg	caacgatcgg	aggaccgaag	gagctaaccg	14460
cttttttgca	caacatgggg	gatcatgtaa	ctcgccttga	tcgttgggaa	ccggagctga	14520
atgaagccat	accaaagcac	gagcgtgaca	ccacgatgcc	tgtagcaatg	gcaacaacgt	14580
tgcgcaaa	attaactggc	gaactactta	ctctagcttc	ccggcaacaa	ttaatagact	14640
ggatggaggc	ggataaagtt	gcaggaccac	ttctgcgctc	ggcccttc	gctggctggt	14700
ttattgctga	taaatctgga	gccggtgagc	gtgggtctcg	cggtatcatt	gcagcactgg	14760
ggccagatgg	taagccctcc	cgtatcgtag	ttatctacac	gacggggagt	caggcaacta	14820
tggatgaacg	aaatagacag	atcgctgaga	taggtgcctc	actgattaag	cattggtaac	14880
tgtcagacca	agtttactca	tatatacttt	agattgattt	aaaatacgta	tatatgtatt	14940
agtcatcgct	attaccatgg	ttaatgcgcc	gctacagggc	gcgtccattc	gccattcagg	15000
ctgcgcaact	gttgggaagg	gcgatcggtg	cgggcctctt	cgctattacg	ccagctggcg	15060
aaaggggat	gtgctgcaag	gcgattaagt	tgggtaacgc	cagggttttc	ccagtcacga	15120
cgttgtaaaa	cgacggccag	tgaattgtaa	tacgactcac	tatagggcga	attgggccct	15180
ctagatgcat	gctcgagcgg	ccgccagtgt	gatggatata	tgcagaattc	cagcacactg	15240
gcggccgtta	ctagtggatc	cgagctcggt	accaagcttg	gcgtaatcat	ggtcatagct	15300
gtttcctgtg	tgaaattgtt	atccgctcac	aattccacac	aacatacgag	ccggaagcat	15360
aaagtgtaaa	gcctgggggtg	cctaattgagt	gagctaactc	acattaattg	cgttgcgctc	15420
ac						15422

<210> 39
 <211> 225
 <212> DNA
 <213> Bos taurus

<400> 39	
ctgtgccttc	tagttgccag ccatctgttg tttgccctc ccccgtcct tccttgacct 60
tggaaggtgc	cactcccact gtcctttcct aataaaatga ggaaattgca tcgcattgtc 120
tgagtaggtg	tcattctatt ctgggggggtg ggggtggggca ggacagcaag ggggaggatt 180
gggaagacaa	tagcaggcat gctgggggatg cgggtgggctc tatgg 225

<210> 40
 <211> 3983
 <212> DNA
 <213> Pan troglodytes

<400> 40

catcatcaat aatatacctt attttggatt gaagccaata tgataatgag atgggcggcg	60
cggggcgggg cgcggggcgg gaggcgggtt tgggggcggg cgggcgggcg gggcggtgtg	120
gcggaagtgg actttgtaag tgtggcggat gtgacttgct agtgccgggc gcggtaaaag	180
tgacgttttc cgtgcgcgac aacgcccccg ggaagtgaca tttttcccgc ggtttttacc	240
ggatgttgta gtgaatttgg gcgtaaccaa gtaagatttg gccattttcg cgggaaaact	300
gaaacgggga agtgaaatct gattaatttt gcgttagtca taccgcgtaa tatttgtcta	360
gggccgaggg actttggccg attacgtgga ggactcgccc aggtgttttt tgaggtgaat	420
ttccgcgttc cgggtcaaag tctccgtttt attattatag gatatcccat tgcatacgtt	480
gtatccatat cataatatgt acaggcgcgc caaagcatga cactgatgtt catttctgat	540
tcttatttta ttattttcaa acacaacaaa atcattcaag tcattcttcc atcttagctt	600
aatagacaca gtagcttaat agaccagta gtgcaaagcc ccattctagc ttataactag	660
tggagaagta ctgcctaca tgggggtaga gtcataatcg tgcatacagga tagggcggtg	720
gtgctgcagc agcgcgcgaa taaactgctg ccgccgcgc tccgtcctgc aggaatacaa	780
catggcagtg gtctcctcag cgatgattcg caccgcccgc agcataaggc gccttgtcct	840
ccgggcacag cagcgcaccc tgatctcact taaatcagca cagtaactgc agcacagcac	900
cacaatattg ttcaaaatcc cacagtgcaa ggcgtgtat ccaaagctca tggcggggac	960
cacagaacct acgtggccat cataccacaa gcgcaggtag attaagtggc gacccctcat	1020
aaacacgctg gacataaaca ttacctctt tggcatgttg taattcacca cctcccggta	1080
ccatataaac ctctgattaa acatggcgcc atccaccacc atcctaaacc agctggccaa	1140
aacctgcccg ccggctatac actgcagggg accgggactg gaacaatgac agtggagagc	1200
ccaggactcg taacctgga tcatcatgct cgtcatgata tcaatgttgg cacaacacag	1260
gcacacgtgc atacacttcc tcaggattac aagctcctcc cgcgttagaa ccatatccca	1320
gggaacaacc cattcctgaa tcagcgtaaa tcccacactg caggggaagac ctgcgcagta	1380
actcacgttg tgcattgtca aagtgttaca ttcgggcagc agcggtgat cctccagtat	1440
ggtagcgcgg gtttctgtct caaaaggagg tagacgatcc ctactgtacg gagtgcgccg	1500
agacaaccga gatcgtgttg gtcgtagtgt catgccaaat ggaacgccgg acgtagtcac	1560
atttctgaa gtcttagatc tctcaacgca gcaccagcac caacacttcg cagtgtaaaa	1620
ggccaagtgc cgagagagta tatataggaa taaaaagtga cgtaaacggg caaagtccaa	1680
aaaacgcca gaaaaaccgc acgcgaacct acgccccgaa acgaaagcca aaaaacacta	1740
gacactccct tccggcgtca acttccgctt tcccacgcta cgtcacttgc cccagtcaaa	1800
caaactacat atcccgaact tccaagtcgc cagcccaaaa acaccgccta cacctccccg	1860
cccgccggcc cgccccaaa cccgcctccc gccccgcgc ccgccccg cgccccatct	1920

cattatcata	ttggcttcaa	tccaaaataa	ggtatattat	tgatgatggt	ttaaacggat	1980
ccaattcttg	aagacgaaag	ggcctcgtga	tacgcctatt	tttatagggt	aatgtcatga	2040
taataatggt	ttcttagacg	tcaggtggca	cttttcgggg	aaatgtgcgc	ggaacccta	2100
tttgtttatt	tttctaaata	cattcaaata	tgtatccgct	catgagacaa	taaccctgat	2160
aaatgcttca	ataatattga	aaaaggaaga	gtatgagtat	tcaacatttc	cgtgtcgccc	2220
ttattccctt	ttttgcggca	ttttgccttc	ctgtttttgc	tcaccagaa	acgctgggtga	2280
aagtaaaaga	tgctgaagat	cagttgggtg	cacgagtggg	ttacatcgaa	ctggatctca	2340
acagcggtaa	gatccttgag	agttttcgcc	ccgaagaacg	ttttccaatg	atgagcactt	2400
ttaaagttct	gctatgtggc	gcggtattat	cccgtgttga	cgccgggcaa	gagcaactcg	2460
gtcgccgcat	acactattct	cagaatgact	tggttgagta	ctcaccagtc	acagaaaagc	2520
atcttacgga	tggtcatgaca	gtaagagaat	tatgcagtgc	tgccataacc	atgagtgata	2580
acactgcggc	caacttactt	ctgacaacga	tcggaggacc	gaaggagcta	accgcttttt	2640
tgcaacaat	gggggatcat	gtaactcgcc	ttgatcgttg	ggaaccggag	ctgaatgaag	2700
ccataccaaa	cgacgagcgt	gacaccacga	tgctgtagc	aatggcaaca	acgttgcgca	2760
aactattaac	tggtgaacta	cttactctag	cttcccgga	acaattaata	gactggatgg	2820
aggcggataa	agttgcagga	ccacttctgc	gctcgccct	tccggctggc	tggtttattg	2880
ctgataaatc	tggtgcccgt	gagcgtgggt	ctcgcggtat	cattgcagca	ctggggccag	2940
atggtaagcc	ctcccgtatc	gtagttatct	acacgacggg	gagtcaggca	actatggatg	3000
aacgaaatag	acagatcgct	gagatagggt	cctcactgat	taagcattgg	taactgtcag	3060
accaagttta	ctcatatata	cttttagattg	atttaaaagg	atctagggtga	agatcctttt	3120
tgataatctc	atgacaaaaa	tcccttaacg	tgagttttcg	ttccactgag	cgtcagaccc	3180
cgtagaaaag	atcaaaggat	cttcttgaga	tccttttttt	ctgcgcgtaa	tctgctgctt	3240
gcaaacaaaa	aaaccaccgc	taccagcggg	ggtttgtttg	cgggatcaag	agctaccaac	3300
tctttttccg	aaggtaactg	gcttcagcag	agcgcagata	ccaaatactg	tccttctagt	3360
gtagccgtag	ttaggccacc	acttcaagaa	ctctgtagca	ccgcctacat	acctcgctct	3420
gctaatacctg	ttaccagtgg	ctgctgccag	tggtcgataag	tcgtgtctta	ccgggttgga	3480
ctcaagacga	tagttaccgg	ataaggcgca	gcggtcgggc	tgaacggggg	gttcgtgcac	3540
acagcccagc	ttggagcgaa	cgacctacac	cgaactgaga	tacctacagc	gtgagctatg	3600
agaaagcgcc	acgcttcccg	aaggagaaaa	ggcggacagg	tatccggtaa	gcggcagggg	3660
cggaacagga	gagcgcacga	gggagcttcc	agggggaaac	gcctgggtatc	tttatagtcc	3720
tgtcgggttt	cgccacctct	gacttgagcg	tcgatttttg	tgatgctcgt	cagggggggcg	3780
gagcctatgg	aaaaacgcc	gcaacgcggc	ctttttacgg	ttcctggcct	tttgcgtggcc	3840

ttgaagctgt ccctgatggg cgtcatctac ctgcctggac agcatggcct gcaacgcggg	3900
catcccgatg ccgccggaag cgagaagaat cataatgggg aaggccatcc agcctcgcgt	3960
cgcagatccg aattcgttta aac	3983

<210> 41
 <211> 25
 <212> DNA
 <213> Human cytomegalovirus

<400> 41	
acatcaatgg gcgtggatag cggtt	25

<210> 42
 <211> 1187
 <212> DNA
 <213> Human cytomegalovirus

<400> 42	
ccattgcata cgttgtatcc atatcataat atgtacattt atattggctc atgtccaaca	60
ttaccgccat gttgacattg attattgact agttattaat agtaatcaat tacgggggtca	120
ttagttcata gcccatatat ggagttccgc gttacataac ttacggtaaa tggcccgct	180
ggctgaccgc ccaacgaccc ccgcccattg acgtcaataa tgacgtatgt tcccatagta	240
acgccaatag ggactttcca ttgacgtcaa tgggtggagt atttacggta aactgcccac	300
ttggcagtac atcaagtgt tcatatgcca agtacgcccc ctattgacgt caatgacggg	360
aaatggcccc cctggcatta tgcccagtac atgaccttat gggactttcc tacttggcag	420
tacatctacg tattagtcac cgctattacc atggtgatgc ggttttggca gtacatcaat	480
gggcgtggat agcggtttga ctcacgggga tttccaagtc tccaccccat tgacgtcaat	540
gggagtttgt tttggcacca aaatcaacgg gactttccaa aatgtcgtaa caactccgcc	600
ccattgacgc aaatgggcgg taggcgtgta cgggtgggagg tctatataag gcgaagcgct	660
ccctatcagt gatagagatc tccctatcag tgatagagat cgtcgacgag ctcgcggcgg	720
gcgggagtcg ctgcgcgctg ccttcgcccc gtgccccgct ccgccgccgc ctcgcgccgc	780
ccgccccggc tctgactgac cgcgttacta aaacaggtaa gtccggcctc cgcgcgggt	840
tttggcgcct cccgcgggcg ccccccctct cacggcgagc gctgccacgt cagacgaagg	900
gcgcagcgag cgtcctgac cttccgcccc gacgctcagg acagcgcccc gctgctcata	960
agactcggcc ttagaacccc agtatcagca gaaggacatt ttaggacggg acttgggtga	1020
ctctagggca ctggttttct ttccagagag cggaacaggc gaggaaaagt agtcccttct	1080
cggcgattct gcggagggat ctccgtgggg cgggtgaacgc cgatgatgcc tctactaacc	1140
atgttcatgt tttctttttt tttctacagg tcctgggtga cgaacag	1187

ФОРМУЛА ИЗОБРЕТЕНИЯ

1. Выделенный полинуклеотид, который кодирует полипептид, выбранный из группы, состоящей из:

- (а) полипептида, имеющего аминокислотную последовательность SEQ ID NO: 1, и
- (б) функционального производного полипептида, имеющего аминокислотную последовательность SEQ ID NO: 1, где функциональное производное имеет аминокислотную последовательность, которая по меньшей мере на 99,8% идентична на всем своем протяжении аминокислотной последовательности SEQ ID NO: 1.

2. Рекombинантный полинуклеотид, содержащий полинуклеотид, выбранный из группы, состоящей из:

- (а) полинуклеотида, кодирующего полипептид, имеющий аминокислотную последовательность SEQ ID NO: 1, и
- (б) полинуклеотида, кодирующего функциональное производное полипептида, имеющего аминокислотную последовательность SEQ ID NO: 1, где функциональное производное имеет аминокислотную последовательность, которая по меньшей мере на 99,8% идентична на всем своем протяжении аминокислотной последовательности SEQ ID NO: 1.

3. Рекombинантный вектор, содержащий полинуклеотид, выбранный из группы, состоящей из:

- (а) полинуклеотида, кодирующего полипептид, имеющий аминокислотную последовательность SEQ ID NO: 1, и
- (б) полинуклеотида, кодирующего функциональное производное полипептида, имеющего аминокислотную последовательность SEQ ID NO: 1, где функциональное производное имеет аминокислотную последовательность, которая по меньшей мере на 99,8% идентична на всем своем протяжении аминокислотной последовательности SEQ ID NO: 1.

4. Рекombинантный аденовирус, содержащий по меньшей мере один полинуклеотид или полипептид, выбранный из группы, состоящей из:

- (а) полинуклеотида, кодирующего полипептид, имеющий аминокислотную последовательность SEQ ID NO: 1,
- (б) полинуклеотида, кодирующего функциональное производное полипептида, имеющего аминокислотную последовательность SEQ ID NO: 1, где функциональное производное имеет аминокислотную последовательность, которая по меньшей мере на 99,8% идентична на всем своем протяжении аминокислотной последовательности SEQ ID NO: 1,
- (в) полипептида, имеющего аминокислотную последовательность SEQ ID NO: 1, и
- (г) функционального производного полипептида, имеющего аминокислотную последовательность SEQ ID NO: 1, где функциональное производное имеет аминокислотную последовательность, которая по меньшей мере на 99,8% идентична на всем своем протяжении аминокислотной последовательности SEQ ID NO: 1.

5. Композиция, содержащая по меньшей мере одно из следующего:

- (а) полинуклеотид, кодирующий полипептид, имеющий аминокислотную последовательность SEQ ID NO: 1,
- (б) полинуклеотид, кодирующий функциональное производное полипептида, имеющего аминокислотную последовательность SEQ ID NO: 1, где функциональное производное имеет

аминокислотную последовательность, которая по меньшей мере на 99,8% идентична на всем своем протяжении аминокислотной последовательности SEQ ID NO: 1,

(в) полипептид, имеющий аминокислотную последовательность SEQ ID NO: 1,

(г) функциональное производное полипептида, имеющего аминокислотную последовательность SEQ ID NO: 1, где функциональное производное имеет аминокислотную последовательность, которая по меньшей мере на 99,8% идентична на всем своем протяжении аминокислотной последовательности SEQ ID NO: 1,

(д) вектор по п. 3 и

(е) рекомбинантный аденовирус по п. 4

и фармацевтически приемлемый эксципиент.

6. Клетка, содержащая по меньшей мере одно из следующего:

(а) полинуклеотид, кодирующий полипептид, имеющий аминокислотную последовательность SEQ ID NO: 1,

(б) полинуклеотид, кодирующий функциональное производное полипептида, имеющего аминокислотную последовательность SEQ ID NO: 1, где функциональное производное имеет аминокислотную последовательность, которая по меньшей мере на 99,8% идентична на всем своем протяжении аминокислотной последовательности SEQ ID NO: 1,

(в) полипептид, имеющий аминокислотную последовательность SEQ ID NO: 1,

(г) функциональное производное полипептида, имеющего аминокислотную последовательность SEQ ID NO: 1, где функциональное производное имеет аминокислотную последовательность, которая по меньшей мере на 99,8% идентична на всем своем протяжении аминокислотной последовательности SEQ ID NO: 1,

(д) вектор по п. 3 и

(е) рекомбинантный аденовирус по п. 4.

7. Выделенный аденовирусный полипептид, выбранный из группы, состоящей из:

(а) полипептида, имеющего аминокислотную последовательность SEQ ID NO: 1, и

(б) функционального производного полипептида, имеющего аминокислотную последовательность SEQ ID NO: 1, где функциональное производное имеет аминокислотную последовательность, которая по меньшей мере на 99,8% идентична на всем своем протяжении аминокислотной последовательности SEQ ID NO: 1.

8. Полинуклеотид, вектор, аденовирус, композиция или клетка по любому из пп. 1-6, где полинуклеотид кодирует функциональное производное полипептида, имеющего аминокислотную последовательность SEQ ID NO: 1, где функциональное производное имеет аминокислотную последовательность, которая по меньшей мере на 99,8% идентична на всем своем протяжении аминокислотной последовательности SEQ ID NO: 1.

9. Полинуклеотид, вектор, аденовирус, композиция или клетка по любому из пп. 1-6, где полинуклеотид кодирует полипептид, имеющий аминокислотную последовательность SEQ ID NO: 1.

10. Полинуклеотид, вектор, аденовирус, композиция или клетка по п. 9, где полинуклеотид имеет последовательность SEQ ID NO: 2.

11. Полинуклеотид, вектор, аденовирус, композиция или клетка по любому из пп. 1-10, дополнительно содержащие полинуклеотид, кодирующий:

(a) полипептид, имеющий аминокислотную последовательность SEQ ID NO: 3; или

(б) функциональное производное полипептида, имеющего аминокислотную последовательность SEQ ID NO: 3, где функциональное производное имеет аминокислотную последовательность, которая по меньшей мере на 60% идентична на всем своем протяжении аминокислотной последовательности SEQ ID NO: 3.

ИЛИ

(a) полипептид, имеющий аминокислотную последовательность SEQ ID NO: 5; или

(б) функциональное производное полипептида, имеющего аминокислотную последовательность SEQ ID NO: 5, где функциональное производное имеет аминокислотную последовательность, которая по меньшей мере на 60% идентична на всем своем протяжении аминокислотной последовательности SEQ ID NO: 5.

12. Полинуклеотид, вектор, аденовирус, композиция или клетка по п. 11, дополнительно содержащие полинуклеотид, кодирующий:

(a) полипептид, имеющий аминокислотную последовательность SEQ ID NO: 3; или

(б) функциональное производное полипептида, имеющего аминокислотную последовательность SEQ ID NO: 3, где функциональное производное имеет аминокислотную последовательность, которая по меньшей мере на 98% идентична на всем своем протяжении аминокислотной последовательности SEQ ID NO: 3.

ИЛИ

(a) полипептид, имеющий аминокислотную последовательность SEQ ID NO: 5; или

(б) функциональное производное полипептида, имеющего аминокислотную последовательность SEQ ID NO: 5, где функциональное производное имеет аминокислотную последовательность, которая по меньшей мере на 98% идентична на всем своем протяжении аминокислотной последовательности SEQ ID NO: 5.

13. Полинуклеотид, вектор, аденовирус, композиция или клетка по любому из пп. 1-10, дополнительно содержащие полинуклеотид, кодирующий:

(a) полипептид, имеющий аминокислотную последовательность SEQ ID NO: 3; или

(б) функциональное производное полипептида, имеющего аминокислотную последовательность SEQ ID NO: 3, где функциональное производное имеет аминокислотную последовательность, которая по меньшей мере на 98% идентична на всем своем протяжении аминокислотной последовательности SEQ ID NO: 3.

И

(a) полипептид, имеющий аминокислотную последовательность SEQ ID NO: 5; или

(б) функциональное производное полипептида, имеющего аминокислотную последовательность SEQ ID NO: 5, где функциональное производное имеет аминокислотную последовательность, которая по меньшей мере на 98% идентична на всем своем протяжении аминокислотной последовательности SEQ ID NO: 5.

14. Полинуклеотид, вектор, аденовирус, композиция или клетка по п. 13, дополнительно содержащие полинуклеотид, кодирующий:

(a) полипептид, имеющий аминокислотную последовательность SEQ ID NO: 3; или

(б) функциональное производное полипептида, имеющего аминокислотную последовательность

SEQ ID NO: 3, где функциональное производное имеет аминокислотную последовательность, которая по меньшей мере на 98% идентична на всем своем протяжении аминокислотной последовательности SEQ ID NO: 3,

и

(а) полипептид, имеющий аминокислотную последовательность SEQ ID NO: 5; или

(б) функциональное производное полипептида, имеющего аминокислотную последовательность SEQ ID NO: 5, где функциональное производное имеет аминокислотную последовательность, которая по меньшей мере на 98% идентична на всем своем протяжении аминокислотной последовательности SEQ ID NO: 5.

15. Полинуклеотид, вектор, аденовирус, композиция или клетка по любому из пп. 11-14, где полинуклеотид содержит последовательность SEQ ID NO: 4.

16. Полинуклеотид, вектор, аденовирус, композиция или клетка по любому из пп. 11-15, где полинуклеотид содержит последовательность SEQ ID NO: 6.

17. Полинуклеотид, вектор, аденовирус, композиция или клетка по любому из пп. 1-16, где полинуклеотид содержит по меньшей мере одно из следующего:

(а) аденовирусный 5'-конец, предпочтительно аденовирусный 5'-инвертированный концевой повтор;

(б) аденовирусную область E1A или ее фрагмент, выбранный из областей E1A_280R и E1A_243R;

(в) аденовирусную область E1B или IX или ее фрагмент, выбранный из группы, состоящей из областей E1B_19K, E1B_55K или IX;

(г) аденовирусную область E2B или ее фрагмент, выбранный из группы, состоящей из областей E2B_pTP, E2B_полимераза и E2B_IVa2;

(д) аденовирусную область L1 или ее фрагмент, где указанный фрагмент кодирует аденовирусный белок, выбранный из группы, состоящей из белков L1_13.6k, L1_52k и L1_IIIa;

(е) аденовирусную область L2 или ее фрагмент, где указанный фрагмент кодирует аденовирусный белок, выбранный из группы, состоящей из белка L2_пентон по п. 3, L2_pVII, L2_V и белка L2_pX;

(ж) аденовирусную область L3 или ее фрагмент, где указанный фрагмент кодирует аденовирусный белок, выбранный из группы, состоящей из белка L3_pVI, белка L3_гексон по п. 2 и L3_протеазы;

(з) аденовирусную область E2A;

(и) аденовирусную область L4 или ее фрагмент, где указанный фрагмент кодирует аденовирусный белок, выбранный из группы, состоящей из белка L4_100k, белка L4_33k и белка L4_VIII;

(к) аденовирусную область E3 или ее фрагмент, выбранный из группы, состоящей из E3 ORF1, E3 ORF2, E3 ORF3, E3 ORF4, E3 ORF5, E3 ORF6, E3 ORF7, E3 ORF8 и E3 ORF9;

(л) аденовирусную область L5 или ее фрагмент, где указанный фрагмент кодирует белок фибриллы L5_фибрилла по п. 1;

(м) аденовирусную область E4 или ее фрагмент, выбранный из группы, состоящей из E4 ORF7, E4 ORF6, E4 ORF4, E4 ORF3, E4 ORF2 и E4 ORF1;

(н) аденовирусный 3'-конец, предпочтительно аденовирусный 3'-инвертированный концевой повтор; и/или

(о) аденовирусную область VAI или VAII PHK, предпочтительно аденовирусную область VAI или VAII PHK из аденовируса, отличного от ChAd157, более предпочтительно из Ad5.

18. Полинуклеотид, вектор, аденовирус, композиция или клетка по п. 17, где полинуклеотид

содержит по меньшей мере одно из следующего:

- (а) аденовирусный 5'-конец, предпочтительно аденовирусный 5'-инвертированный концевой повтор;
- (д) аденовирусную область L1 или ее фрагмент, где указанный фрагмент кодирует аденовирусный белок, выбранный из группы, состоящей из белков L1_13.6k, L1_52k и L1_IIIa;
- (е) аденовирусную область L2 или ее фрагмент, где указанный фрагмент кодирует аденовирусный белок, выбранный из группы, состоящей из белка L2_пентон по п. 3, L2_pVII, L2_V и белка L2_pX;
- (ж) аденовирусную область L3 или ее фрагмент, где указанный фрагмент кодирует аденовирусный белок, выбранный из группы, состоящей из белка L3_pVI, белка L3_гексон по п. 2 и L3_протеазы;
- (и) аденовирусную область L4 или ее фрагмент, где указанный фрагмент кодирует аденовирусный белок, выбранный из группы, состоящей из белка L4_100k, белка L4_33k и белка L4_VIII;
- (л) аденовирусную область L5 или ее фрагмент, где указанный фрагмент кодирует белок фибриллы L5_фибрилла по п. 1;
- (н) аденовирусный 3'-конец, предпочтительно аденовирусный 3'-инвертированный концевой повтор.

19. Полинуклеотид, вектор, аденовирус, композиция или клетка по любому из пп. 1-16, где полинуклеотид содержит аденовирусную область VAI или VAII РНК.

20. Полинуклеотид, вектор, аденовирус, композиция или клетка по п. 19, где область VAI или VAII РНК имеет происхождение из аденовируса, отличного от ChAd157.

21. Полинуклеотид, вектор, аденовирус, композиция или клетка по п. 20, где область VAI или VAII РНК имеет происхождение из Ad5.

22. Полинуклеотид, вектор, аденовирус, композиция или клетка по любому из пп. 1-16, где полинуклеотид содержит или состоит из полинуклеотида, который идентичен по меньшей мере на 95% на всем своем протяжении эталонной последовательности, по существу состоящей из SEQ ID NO: 15 или 22.

23. Полинуклеотид, вектор, аденовирус, композиция или клетка по п. 22, где полинуклеотид содержит или состоит из полинуклеотида, который идентичен по меньшей мере на 99% на всем своем протяжении эталонной последовательности.

24. Полинуклеотид, вектор, аденовирус, композиция или клетка по п. 23, где полинуклеотид содержит или состоит из полинуклеотида, который идентичен по меньшей мере на 99,5% на всем своем протяжении эталонной последовательности.

25. Полинуклеотид, вектор, аденовирус, композиция или клетка по любому из пп. 22-24, где полинуклеотид содержит или состоит из полинуклеотида, который идентичен на всем своем протяжении эталонной последовательности.

26. Полинуклеотид, вектор, аденовирус, композиция или клетка по п. 25, где эталонная последовательность представляет собой SEQ ID NO: 15.

27. Полинуклеотид, вектор, аденовирус, композиция или клетка по п. 25, где эталонная последовательность представляет собой SEQ ID NO: 22.

28. Полинуклеотид, вектор, аденовирус, композиция или клетка по любому из пп. 1-16, где полинуклеотид содержит мутацию или делецию, которая делает нефункциональным по меньшей мере один ген в области генома, выбранной из группы, состоящей из E1A, E1B, E2A, E2B, E3 и E4.

29. Полинуклеотид, вектор, аденовирус, композиция или клетка по п. 28, где у полинуклеотида отсутствует по меньшей мере один ген в области генома, выбранной из группы, состоящей из E1A, E1B, E2A, E2B, E3 и/или E4.

30. Полинуклеотид, вектор, аденовирус, композиция или клетка по пп. 28 или 29, где области генома представляют собой E1A и/или E1B.

31. Полинуклеотид, вектор, аденовирус, композиция или клетка по любому из пп. 1-16, где полинуклеотид содержит делецию области генома E1.

32. Аденовирус по любому из пп. 4 и 8-31, где рекомбинантный аденовирус является репликативно-компетентным.

33. Аденовирус по любому из пп. 4 и 8-32, где рекомбинантный аденовирус является репликативно-некомпетентным.

34. Аденовирус по любому из пп. 4 и 8-33, где рекомбинантный аденовирус содержит последовательность нуклеиновой кислоты, кодирующую белок, где последовательность нуклеиновой кислоты функционально связана с одной или более чем одной последовательностью, направляющей экспрессию указанного белка в клетке-хозяине.

35. Аденовирус по п. 34, где белок представляет собой антигенный белок или его фрагмент.

36. Аденовирус по п. 35, где белок представляет собой гетерологичный белок или его фрагмент.

37. Аденовирус по п. 36, где белок имеет происхождение из вируса.

38. Аденовирус по любому из пп. 34-37, где одна или более чем одна последовательность, направляющая экспрессию указанного продукта в клетке-хозяине, включает последовательность, выбранную из одной или более последовательностей из группы, состоящей из: последовательности инициации транскрипции, последовательности терминации транскрипции, последовательности промотора и последовательности энхансера.

39. Аденовирус по п. 38, где одна или более чем одна последовательность, направляющая экспрессию указанного продукта в клетке-хозяине, включает последовательность промотора.

40. Аденовирус по п. 39, где последовательность промотора выбрана из группы, состоящей из внутреннего промотора, нативного промотора, промотора LTR RSV (длинные концевые повторы респираторного синцитиального вируса), промотора CMV (цитомегаловируса), промотора SV40, промотора дигидрофолатредуктазы, промотора β -актина, промотора PGK (фосфоглицеролкиназы), промотора EF1a (фактора элонгации 1a) и промотора CASI.

41. Аденовирус по п. 39, где последовательность промотора представляет собой усиленный промотор hCMV, такой как представленный в SEQ ID NO: 42.

42. Аденовирус по любому из пп. 4 и 8-41, где аденовирус имеет доминирование серотипа менее 10% у субъектов-людей, и предпочтительно доминирование серотипа у субъектов-людей отсутствует.

43. Аденовирус по любому из пп. 4 и 8-42, где аденовирус способен инфицировать клетку млекопитающего.

44. Композиция по любому из пп. 5 и 8-43, содержащая адъювант, выбранный из перечня, состоящего из: неорганических адъювантов (например, неорганических солей металлов, таких как фосфат алюминия или гидрат окиси алюминия), органических адъювантов (например, сапонинов, таких как QS21, или сквалена), адъювантов на основе масел (например, полного адъюванта Фрейнда и неполного адъюванта Фрейнда), цитокинов (например, IL-1 β (интерлейкин 1 β), IL-2, IL-7, IL-12, IL-18, GM-CSF и INF- γ (интерферон γ)), адъювантов в виде частиц (например, иммуностимулирующих комплексов (ISCOMS), липосом или биodeградируемых микросфер), виросом, бактериальных адъювантов (например, монофосфориллипида A, такого как 3-де-О-ацетилированный монофосфориллипид A (3D-MPL), или

мурамилпептидов), синтетических адъювантов (например, неионных блок-сополимеров, аналогов мурамилпептидов или синтетического липида А), синтетических полинуклеотидных адъювантов (например, полиаргинина или полилизина) и иммуностимулирующих олигонуклеотидов, содержащих неметилированные динуклеотиды CpG («CpG»).

45. Композиция по п. 44, где адъювант представляет собой 3D-MPL и/или QS21.

46. Клетка по любому из пп. 6 и 8-31, где клетка представляет собой клетку-хозяина, которая экспрессирует по меньшей мере один аденовирусный ген, выбранный из группы, состоящей из E1A, E1B, E2A, E2B, E3, E4, L1, L2, L3, L4 и L5.

47. Клетка по п. 45, где клетка-хозяин выращена в суспензии.

48. Полинуклеотид, полипептид, вектор, аденовирус, композиция или клетка по любому из пп. 1-47 для применения в качестве лекарственного средства.

49. Полинуклеотид, полипептид, вектор, аденовирус, композиция или клетка по п. 48 для применения в качестве вакцины.

50. Применение полинуклеотида, полипептида, вектора, аденовируса, композиции или клетки по любому из пп. 1-47 для терапии или профилактики заболевания.

51. Способ индуцирования иммунного ответа у субъекта, включающий введение субъекту полинуклеотида, полипептида, вектора, аденовируса, композиции или клетки по любому из пп. 1-47.

52. Выделенный полинуклеотид, содержащий или состоящий из последовательности SEQ ID NO: 2.

53. Рекombинантный аденовирусный вектор по любому из пп. 4 и 8-43, кодирующий трансген, для введения субъекту, который ранее находился в контакте с рекombинантным аденовирусным вектором, не содержащим фибриллу ChAd157 или ее функциональное производное, как описано в данном документе.

54. Рекombинантный аденовирусный вектор по п. 53, где субъект ранее находился в контакте с рекombинантным аденовирусным вектором, содержащим фибриллу, гексон и/или пентон ChAd155.

55. Рекombинантный аденовирусный вектор по п. 54, где субъект ранее находился в контакте с рекombинантным аденовирусным вектором, содержащим фибриллу, гексон и пентон ChAd155.

56. Рекombинантный аденовирусный вектор по п. 54 или 55, где рекombинантный аденовирусный вектор, содержащий фибриллу, гексон и/или пентон ChAd155, кодирует трансген, направленный на иное медицинское показание или показания, нежели рекombинантный аденовирусный вектор по любому из пп. 4 и 8-43, кодирующий трансген.

57. Рекombинантный аденовирусный вектор по п. 54 или 55, где рекombинантный аденовирусный вектор, содержащий фибриллу, гексон и/или пентон ChAd155, кодирует трансген, направленный на то же медицинское показание или показания, что и рекombинантный аденовирусный вектор по любому из пп. 4 и 8-43, кодирующий трансген.

58. Рекombинантный аденовирусный вектор по п. 54 или 55, где рекombинантный аденовирусный вектор, содержащий фибриллу, гексон и/или пентон ChAd155, кодирует такой же трансген, как и рекombинантный аденовирусный вектор по любому из пп. 4 и 8-43, кодирующий трансген.

59. Рекombинантный аденовирусный вектор по любому из пп. 53-58, где трансгены кодируют иммуноген, который полезен для иммунизации человека или животного, не являющегося человеком, против патогена, такого как бактерии, грибы, паразитические микроорганизмы или многоклеточные паразиты, инфицирующие человека и позвоночных, не являющихся человеком, или против раковой клетки или опухолевой клетки.

60. Рекombинантный аденовирусный вектор по любому из пп. 4 и 8-43, кодирующий трансген, для введения субъекту, который может впоследствии оказаться в контакте с рекombинантным аденовирусным вектором, не содержащим фибриллу ChAd157 или ее функциональное производное, описанные в данном документе.

61. Рекombинантный аденовирусный вектор по п. 60, где субъект может впоследствии оказаться в контакте с рекombинантным аденовирусным вектором, содержащим фибриллу, гексон и/или пентон ChAd155.

62. Рекombинантный аденовирусный вектор по п. 61, где субъект может впоследствии оказаться в контакте с рекombинантным аденовирусным вектором, содержащим фибриллу, гексон и пентон ChAd155.

63. Рекombинантный аденовирусный вектор по пп. 61 или 62, где рекombинантный аденовирусный вектор, содержащий фибриллу, гексон и/или пентон ChAd155, кодирует трансген, направленный на иное медицинское показание или показания, нежели рекombинантный аденовирусный вектор по любому из пп. 4 и 8-43, кодирующий трансген.

64. Рекombинантный аденовирусный вектор по пп. 61 или 62, где рекombинантный аденовирусный вектор, содержащий фибриллу, гексон и/или пентон ChAd155, кодирует трансген, направленный на то же медицинское показание или показания, что и рекombинантный аденовирусный вектор по любому из пп. 4 и 8-43, кодирующий трансген.

65. Рекombинантный аденовирусный вектор по п. 61 или 62, где рекombинантный аденовирусный вектор, содержащий фибриллу, гексон и/или пентон ChAd155, кодирует такой же трансген, как и рекombинантный аденовирусный вектор по любому из пп. 4 и 8-43, кодирующий трансген.

66. Рекombинантный аденовирусный вектор по любому из пп. 60-65, где трансгены кодируют иммуноген, который полезен для иммунизации человека или животного, не являющегося человеком, против патогена, такого как бактерии, грибы, паразитические микроорганизмы или многоклеточные паразиты, инфицирующие человека и позвоночных, не являющихся человеком, или против раковой клетки или опухолевой клетки.

67. Способ индуцирования иммунного ответа у субъекта, включающий:

(а) введение субъекту рекombинантного аденовирусного вектора по любому из пп. 4 и 8-43, кодирующего первый трансген, и

(б) введение субъекту рекombинантного аденовирусного вектора, который не содержит фибриллу ChAd157, где вектор кодирует второй трансген;

где стадии (а) и (б) можно осуществлять в любом порядке и первый и второй трансгены могут быть одинаковыми или различными.

68. Способ профилактики или лечения субъекта, включающий:

(а) введение субъекту рекombинантного аденовирусного вектора по любому из пп. 4 и 8-43, кодирующего первый трансген, кодирующий иммуноген, который полезен для иммунизации человека или животного, не являющегося человеком, против патогена, такого как бактерии, грибы, паразитические микроорганизмы или многоклеточные паразиты, инфицирующие человека и позвоночных, не являющихся человеком, или против раковой клетки или опухолевой клетки; и

(б) введение субъекту рекombинантного аденовирусного вектора, не содержащего фибриллу ChAd157 или ее функциональное производное, как описано в данном документе, где вектор кодирует второй трансген, кодирующий иммуноген, который полезен для иммунизации человека или животного, не

являющегося человеком, против иного патогена, такого как бактерии, грибы, паразитические микроорганизмы или многоклеточные паразиты, инфицирующие человека и позвоночных, не являющихся человеком, или против раковой клетки или опухолевой клетки;

где стадии (а) и (б) можно осуществлять в любом порядке.

69. Рекombинантный аденовирусный вектор или способ по любому из пп. 53-68, где рекombинантный аденовирусный вектор, не содержащий фибриллу ChAd157 или его функциональное производное, как описано в данном документе, имеет низкую перекрестную реактивность с рекombинантным аденовирусным вектором по любому из пп. 4 или 8-43.

70. Рекombинантный аденовирусный вектор или способ по п. 69, где иммунизация первым вектором вызывает нейтрализующий титр, который в среднем составляет менее 50% от уровня, развивающегося после иммунизации вторым вектором.

71. Рекombинантный аденовирусный вектор или способ по любому из пп. 53-70, где рекombинантный аденовирусный вектор, не содержащий фибриллу ChAd157, не содержит фибриллу ChAd157, гексон ChAd157 или фибриллу ChAd157 или ее функциональные производные, имеющие по меньшей мере 98% идентичности с ней.

Фиг. 1А

	10	20	30	40	50	60	70	80	
	-----+	-----+	-----+	-----+	-----+	-----+	-----+	-----+	
12_ChAd157	MKRTKTSDESNPVV	PYDTEGGPSVP	FLTPPFVSP	DGFGQESPPG	VLSLNLAEP	LVTS	SHGMLALKMG	SGLSLDDAGN	LTS80
01_ChAd3	MKRTKTSDESNPVV	PYDTEGGPSVP	FLTPPFVSP	DGFGQESPPG	VLSLNLAEP	LVTS	SHGMLALKMG	SGLSLDDAGN	LTS80
02_PanAd3	MKRKTSDESNPVV	PYDTEGGPSVP	FLTPPFVSP	DGFGQESPPG	VLSLNLAEP	LVTS	SHGMLALKMG	SGLSLDDAGN	LTS80
03_ChAd17	MKRTKTSDESNPVV	PYDTEGGPSVP	FLTPPFVSP	DGFGQESPPG	VLSLNLAEP	LVTS	SHGMLALKMG	SGLSLDDAGN	LTS80
04_ChAd19	MKRTKTSDESNPVV	PYDTEGGPSVP	FLTPPFVSP	DGFGQESPPG	VLSLNLAEP	LVTS	SHGMLALKMG	SGLSLDDAGN	LTS80
05_ChAd24	MKRTKTSDESNPVV	PYDTEGGPSVP	FLTPPFVSP	DGFGQESPPG	VLSLNLAEP	LVTS	SHGMLALKMG	SGLSLDDAGN	LTS80
06_ChAd155	MKRTKTSDESNPVV	PYDTEGGPSVP	FLTPPFVSP	DGFGQESPPG	VLSLNLAEP	LVTS	SHGMLALKMG	SGLSLDDAGN	LTS80
07_ChAd11	MKRTKTSDESNPVV	PYDTEGGPSVP	FLTPPFVSP	DGFGQESPPG	VLSLNLAEP	LVTS	SHGMLALKMG	SGLSLDDAGN	LTS80
08_ChAd20	MKRTKTSDESNPVV	PYDTEGGPSVP	FLTPPFVSP	DGFGQESPPG	VLSLNLAEP	LVTS	SHGMLALKMG	SGLSLDDAGN	LTS80
09_ChAd31	MKRTKTSDESNPVV	PYDTEGGPSVP	FLTPPFVSP	DGFGQESPPG	VLSLNLAEP	LVTS	SHGMLALKMG	SGLSLDDAGN	LTS80
10_PanAd1	MKRKTSDESNPVV	PYDTEGGPSVP	FLTPPFVSP	DGFGQESPPG	VLSLNLAEP	LVTS	SHGMLALKMG	SGLSLDDAGN	LTS80
11_PanAd2	MKRKTSDESNPVV	PYDTEGGPSVP	FLTPPFVSP	DGFGQESPPG	VLSLNLAEP	LVTS	SHGMLALKMG	SGLSLDDAGN	LTS80
	90	100	110	120	130	140	150	160	
	-----+	-----+	-----+	-----+	-----+	-----+	-----+	-----+	
12_ChAd157	QDSTTAPPLKKT	KTNLSLETSP	PLTVSTSGALT	VAAAAPLAVAGT	SLTMQSEAP	PLTVQDAKL	TLATKG	PLTVSE	GGKLAL160
01_ChAd3	QDSTTAPPLKKT	KTNLSLETSP	PLTVSTSGALT	VAAAAPLAVAGT	SLTMQSEAP	PLTVQDAKL	TLATKG	PLTVSE	GGKLAL160
02_PanAd3	QDSTTAPPLKKT	KTNLSLETSP	PLTVSTSGALT	VAAAAPLAVAGT	SLTMQSEAP	PLTVQDAKL	TLATKG	PLTVSE	GGKLAL159
03_ChAd17	QDSTTAPPLKKT	KTNLSLETSP	PLTVSTSGALT	VAAAAPLAVAGT	SLTMQSEAP	PLTVQDAKL	TLATKG	PLTVSE	GGKLAL160
04_ChAd19	QDSTTAPPLKKT	KTNLSLETSP	PLTVSTSGALT	VAAAAPLAVAGT	SLTMQSEAP	PLTVQDAKL	TLATKG	PLTVSE	GGKLAL160
05_ChAd24	QDSTTAPPLKKT	KTNLSLETSP	PLTVSTSGALT	VAAAAPLAVAGT	SLTMQSEAP	PLTVQDAKL	TLATKG	PLTVSE	GGKLAL160
06_ChAd155	QDSTTAPPLKKT	KTNLSLETSP	PLTVSTSGALT	VAAAAPLAVAGT	SLTMQSEAP	PLTVQDAKL	TLATKG	PLTVSE	GGKLAL160
07_ChAd11	QDSTTAPPLKKT	KTNLSLETSP	PLTVSTSGALT	VAAAAPLAVAGT	SLTMQSEAP	PLTVQDAKL	TLATKG	PLTVSE	GGKLAL160
08_ChAd20	QDSTTAPPLKKT	KTNLSLETSP	PLTVSTSGALT	VAAAAPLAVAGT	SLTMQSEAP	PLTVQDAKL	TLATKG	PLTVSE	GGKLAL160
09_ChAd31	QDSTTAPPLKKT	KTNLSLETSP	PLTVSTSGALT	VAAAAPLAVAGT	SLTMQSEAP	PLTVQDAKL	TLATKG	PLTVSE	GGKLAL160
10_PanAd1	QDSTTAPPLKKT	KTNLSLETSP	PLTVSTSGALT	VAAAAPLAVAGT	SLTMQSEAP	PLTVQDAKL	TLATKG	PLTVSE	GGKLAL159
11_PanAd2	QDSTTAPPLKKT	KTNLSLETSP	PLTVSTSGALT	VAAAAPLAVAGT	SLTMQSEAP	PLTVQDAKL	TLATKG	PLTVSE	GGKLAL159

Полнуклеотиды и полнентитиды аденовируса

Фиг. 1Б

	170	180	190	200	210	220	230	240	
12_ChAd157	QTSAPLTAADSSTLTVSATPP	ES	SGSLG	DMEDP	MYTNGKLG	INIGGPLV	VVDSLH	LTWVTG	GLTVDNALQTRV 240
01_ChAd3	QTSAPLTAADSSTLTVSATPP	ES	SGSLG	DMEDP	MYTNGKLG	INIGGPLV	VVDSLH	LTWVTG	GLTVDNALQTRV 240
02_PanAd3	QTSAPLTAADSSTLTVSATPP	ES	SGSLG	DMEDP	MYTNGKLG	INIGGPLV	VVDSLH	LTWVTG	GLTVDNALQTRV 239
03_ChAd17	QTSAPLTAADSSTLTVSATPP	ES	SGSLG	DMEDP	MYTNGKLG	INIGGPLV	VVDSLH	LTWVTG	GLTVDNALQTRV 240
04_ChAd19	QTSAPLTAADSSTLTVSATPP	ES	SGSLG	DMEDP	MYTNGKLG	INIGGPLV	VVDSLH	LTWVTG	GLTVDNALQTRV 240
05_ChAd24	QTSAPLTAADSSTLTVSATPP	ES	SGSLG	DMEDP	MYTNGKLG	INIGGPLV	VVDSLH	LTWVTG	GLTVDNALQTRV 240
06_ChAd155	QTSAPLTAADSSTLTVSATPP	ES	SGSLG	DMEDP	MYTNGKLG	INIGGPLV	VVDSLH	LTWVTG	GLTVDNALQTRV 240
07_ChAd11	QTSAPLTAADSSTLTVSATPP	ES	SGSLG	DMEDP	MYTNGKLG	INIGGPLV	VVDSLH	LTWVTG	GLTVDNALQTRV 240
08_ChAd20	QTSAPLTAADSSTLTVSATPP	ES	SGSLG	DMEDP	MYTNGKLG	INIGGPLV	VVDSLH	LTWVTG	GLTVDNALQTRV 240
09_ChAd31	QTSAPLTAADSSTLTVSATPP	ES	SGSLG	DMEDP	MYTNGKLG	INIGGPLV	VVDSLH	LTWVTG	GLTVDNALQTRV 240
10_PanAd1	QTSAPLTAADSSTLTVSATPP	ES	SGSLG	DMEDP	MYTNGKLG	INIGGPLV	VVDSLH	LTWVTG	GLTVDNALQTRV 239
11_PanAd2	QTSAPLTAADSSTLTVSATPP	ES	SGSLG	DMEDP	MYTNGKLG	INIGGPLV	VVDSLH	LTWVTG	GLTVDNALQTRV 239
	250	260	270	280	290	300	310	320	
12_ChAd157	TGALGYDTS	GNLQLRAAGGM	RIDANGQLIL	VAYPFDAQ	NNLSLRLGQG	PLYV	IAHNLDLNCNRGL	ETTS	GNTKKLE 320
01_ChAd3	TGALGYDTS	GNLQLRAAGGM	RIDANGQLIL	VAYPFDAQ	NNLSLRLGQG	PLYV	IAHNLDLNCNRGL	ETTS	GNTKKLE 320
02_PanAd3	TGALGYDTS	GNLQLRAAGGM	RIDANGQLIL	VAYPFDAQ	NNLSLRLGQG	PLYV	IAHNLDLNCNRGL	ETTS	GNTKKLE 319
03_ChAd17	TGALGYDTS	GNLQLRAAGGM	RIDANGQLIL	VAYPFDAQ	NNLSLRLGQG	PLYV	IAHNLDLNCNRGL	ETTS	GNTKKLE 320
04_ChAd19	TGALGYDTS	GNLQLRAAGGM	RIDANGQLIL	VAYPFDAQ	NNLSLRLGQG	PLYV	IAHNLDLNCNRGL	ETTS	GNTKKLE 320
05_ChAd24	TGALGYDTS	GNLQLRAAGGM	RIDANGQLIL	VAYPFDAQ	NNLSLRLGQG	PLYV	IAHNLDLNCNRGL	ETTS	GNTKKLE 320
06_ChAd155	TGALGYDTS	GNLQLRAAGGM	RIDANGQLIL	VAYPFDAQ	NNLSLRLGQG	PLYV	IAHNLDLNCNRGL	ETTS	GNTKKLE 320
07_ChAd11	TGALGYDTS	GNLQLRAAGGM	RIDANGQLIL	VAYPFDAQ	NNLSLRLGQG	PLYV	IAHNLDLNCNRGL	ETTS	GNTKKLE 320
08_ChAd20	TGALGYDTS	GNLQLRAAGGM	RIDANGQLIL	VAYPFDAQ	NNLSLRLGQG	PLYV	IAHNLDLNCNRGL	ETTS	GNTKKLE 320
09_ChAd31	TGALGYDTS	GNLQLRAAGGM	RIDANGQLIL	VAYPFDAQ	NNLSLRLGQG	PLYV	IAHNLDLNCNRGL	ETTS	GNTKKLE 320
10_PanAd1	TGALGYDTS	GNLQLRAAGGM	RIDANGQLIL	VAYPFDAQ	NNLSLRLGQG	PLYV	IAHNLDLNCNRGL	ETTS	GNTKKLE 319
11_PanAd2	TGALGYDTS	GNLQLRAAGGM	RIDANGQLIL	VAYPFDAQ	NNLSLRLGQG	PLYV	IAHNLDLNCNRGL	ETTS	GNTKKLE 319

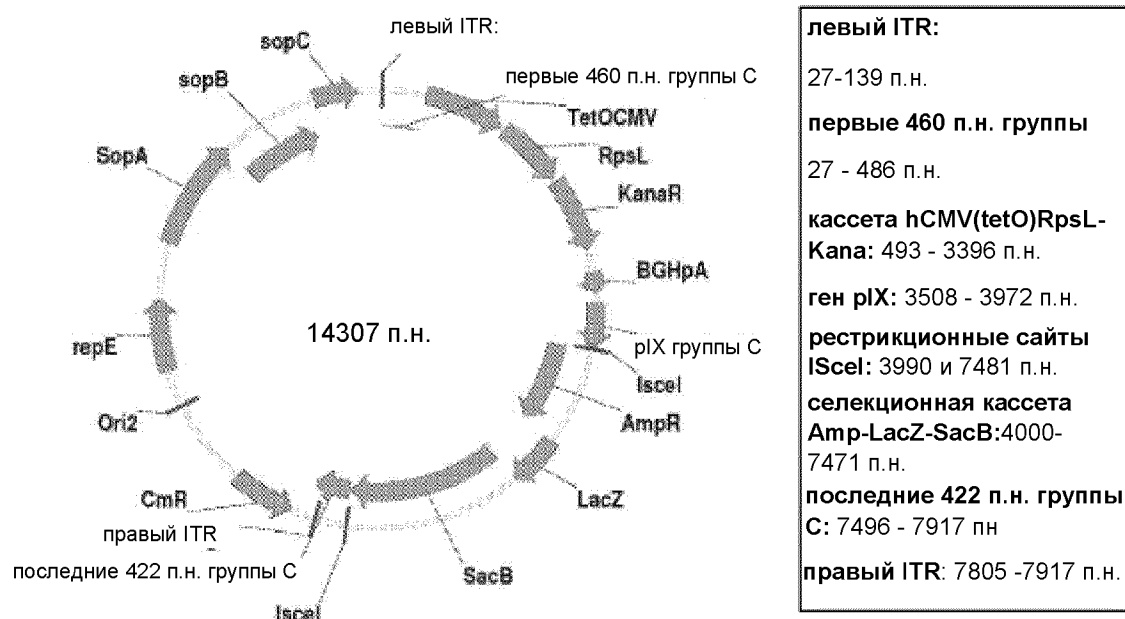
	330	340	350	360	370	380	390	400
12_ChAd157	KTSGLDYDTNGAIIKLGTGLSFNDTGATVGNTEDDKLTLLWT							
01_ChAd3	KTSGLDYDTNGAIIKLGTGLSFNDTGATVGNTEDDKLTLLWT							
02_PanAd3	KIDSGLDYDTNGAIIKLGTGLSFNDTGATVGNTEDDKLTLLWT							
03_ChAd17	KTSGLDYDTNGAIIKLGTGLSFNDTGATVGNTEDDKLTLLWT							
04_ChAd19	KGGGLDYDTNGAIIKLGTGLSFNDTGATVGNTEDDKLTLLWT							
05_ChAd24	KGGGLDYDTNGAIIKLGTGLSFNDTGATVGNTEDDKLTLLWT							
06_ChAd155	NIKTAGGLIYDDTAIAINAGDGLQFSGSDDTNFLKTKGGGLDYDSSRAIIKLGTGLSFNDTGATVGNTEDDKLTLLWT							
07_ChAd11	NIKTAGGLIYDDTAIAINAGDGLQFSGSDDTNFLKTKGGGLDYDSSRAIIKLGTGLSFNDTGATVGNTEDDKLTLLWT							
08_ChAd20	NIKTAGGLIYDDTAIAINAGDGLQFSGSDDTNFLKTKGGGLDYDSSRAIIKLGTGLSFNDTGATVGNTEDDKLTLLWT							
09_ChAd31	NIKTAGGLIYDDTAIAINAGDGLQFSGSDDTNFLKTKGGGLDYDSSRAIIKLGTGLSFNDTGATVGNTEDDKLTLLWT							
10_PanAd1	NIKTAGGLFYDGTATAINAGDGLQFSGSDDTNFLKTKGGGLDYDSSRAIIKLGTGLSFNDTGATVGNTEDDKLTLLWT							
11_PanAd2	NIKTAGGLFYDGTATAINAGDGLQFSGSDDTNFLKTKGGGLDYDSSRAIIKLGTGLSFNDTGATVGNTEDDKLTLLWT							

	410	420	430	440	450	460	470	480
12_ChAd157	TPDPSPNCRIHSKDKKFTLVLTCKGSQLASVAILVVGSLAIIIGTVTSAGIIFLRFDNGVLLNSSLDPQYWNERRG							
01_ChAd3	TPDPSPNCRIHSKDKKFTLVLTCKGSQLASVAILVVGSLAIIIGTVTSAGIIFLRFDNGVLLNSSLDPQYWNERRG							
02_PanAd3	TPDPSPNCRIHSKDKKFTLVLTCKGSQLASVAILVVGSLAIIIGTVTSAGIIFLRFDNGVLLNSSLDPQYWNERRG							
03_ChAd17	TPDPSPNCRIHSKDKKFTLVLTCKGSQLASVAILVVGSLAIIIGTVTSAGIIFLRFDNGVLLNSSLDPQYWNERRG							
04_ChAd19	TPDPSPNCRIHSKDKKFTLVLTCKGSQLASVAILVVGSLAIIIGTVTSAGIIFLRFDNGVLLNSSLDPQYWNERRG							
05_ChAd24	TPDPSPNCRIHSKDKKFTLVLTCKGSQLASVAILVVGSLAIIIGTVTSAGIIFLRFDNGVLLNSSLDPQYWNERRG							
06_ChAd155	TPDPSPNCRIHSEKDAKFTLVLTCKGSQLASVAILVVGSLAIIIGTVTSAGIIFLRFDNGVLLNSSLDPQYWNERRG							
07_ChAd11	TPDPSPNCRIHSEKDAKFTLVLTCKGSQLASVAILVVGSLAIIIGTVTSAGIIFLRFDNGVLLNSSLDPQYWNERRG							
08_ChAd20	TPDPSPNCRIHSEKDAKFTLVLTCKGSQLASVAILVVGSLAIIIGTVTSAGIIFLRFDNGVLLNSSLDPQYWNERRG							
09_ChAd31	TPDPSPNCRIHSEKDAKFTLVLTCKGSQLASVAILVVGSLAIIIGTVTSAGIIFLRFDNGVLLNSSLDPQYWNERRG							
10_PanAd1	TPDPSPNCRIHSEKDAKFTLVLTCKGSQLASVAILVVGSLAIIIGTVTSAGIIFLRFDNGVLLNSSLDPQYWNERRG							
11_PanAd2	TPDPSPNCRIHSEKDAKFTLVLTCKGSQLASVAILVVGSLAIIIGTVTSAGIIFLRFDNGVLLNSSLDPQYWNERRG							

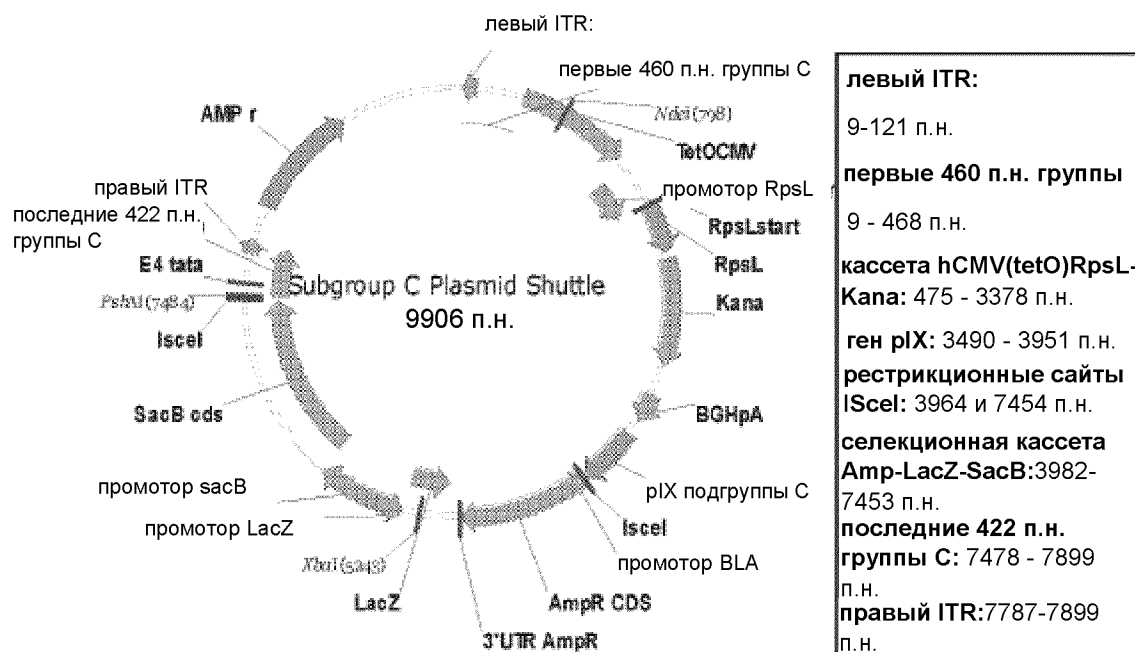
Фиг. 1Г

	490	500	510	520	530	540	550	560	
12_ChAd157	STNAAAYTNAVGFMPNLAAYPKTQSQTA	KNIVSQVYLNGDKSKPMTLTITLNGTNE	SETSQVSYYSMSFQWAWESGG						524
01_ChAd3	STNAAAYTNAVGFMPNLAAYPKTQSQTA	KNIVSQVYLNGDKSKPMTLTITLNGTNE	SETSQVSYYSMSFQWAWESGG						524
02_PanAd3	STNAAAYTNAVGFMPNLAAYPKTQSQTA	KNIVSQVYLNGDKSKPMTLTITLNGTNE	SETSQVSYYSMSFQWAWESGG						522
03_ChAd17	STNAAAYTNAVGFMPNLAAYPKTQSQTA	KNIVSQVYLNGDKSKPMTLTITLNGTNE	SETSQVSYYSMSFQWAWESGG						524
04_ChAd19	STNAAAYTNAVGFMPNLAAYPKTQSQTA	KNIVSQVYLNGDKSKPMTLTITLNGTNE	SETSQVSYYSMSFQWAWESGG						524
05_ChAd24	STNAAAYTNAVGFMPNLAAYPKTQSQTA	KNIVSQVYLNGDKSKPMTLTITLNGTNE	SETSQVSYYSMSFQWAWESGG						524
06_ChAd155	STETGTAYTNAVGFMPNLTAYPKTQSQTA	KNIVSQVYLNGDKSKPMTLTITLNGTNE	GG-DATVSYYSMSFQWNNNGGN						559
07_ChAd11	STETGTAYTNAVGFMPNLTAYPKTQSQTA	KNIVSQVYLNGDKSKPMTLTITLNGTNE	GG-DATVSYYSMSFQWNNNGGN						559
08_ChAd20	STETGTAYTNAVGFMPNLTAYPKTQSQTA	KNIVSQVYLNGDKSKPMTLTITLNGTNE	GG-DATVSYYSMSFQWNNNGGN						559
09_ChAd31	STETGTAYTNAVGFMPNLTAYPKTQSQTA	KNIVSQVYLNGDKSKPMTLTITLNGTNE	GG-DATVSYYSMSFQWNNNGGN						559
10_PanAd1	STETGTAYTNAVGFMPNLTAYPKTQSQTA	KNIVSQVYLNGDKSKPMTLTITLNGTNE	GG-DATVSYYSMSFQWNNNGGN						558
11_PanAd2	STETGTAYTNAVGFMPNLTAYPKTQSQTA	KNIVSQVYLNGDKSKPMTLTITLNGTNE	GG-DATVSYYSMSFQWNNNGGN						558
	570								
12_ChAd157	YATETFTNSFTFSYIAAG								543
01_ChAd3	YATETFTNSFTFSYIAAG								543
02_PanAd3	YATETFTNSFTFSYIAAG								541
03_ChAd17	YATETFTNSFTFSYIAAG								543
04_ChAd19	YATETFTNSFTFSYIAAG								543
05_ChAd24	YATETFTNSFTFSYIAAG								543
06_ChAd155	YATETFTNSFTFSYIAAG								578
07_ChAd11	YATETFTNSFTFSYIAAG								578
08_ChAd20	YATETFTNSFTFSYIAAG								578
09_ChAd31	YATETFTNSFTFSYIAAG								578
10_PanAd1	YATETFTNSFTFSYIAAG								577
11_PanAd2	YATETFTNSFTFSYIAAG								577

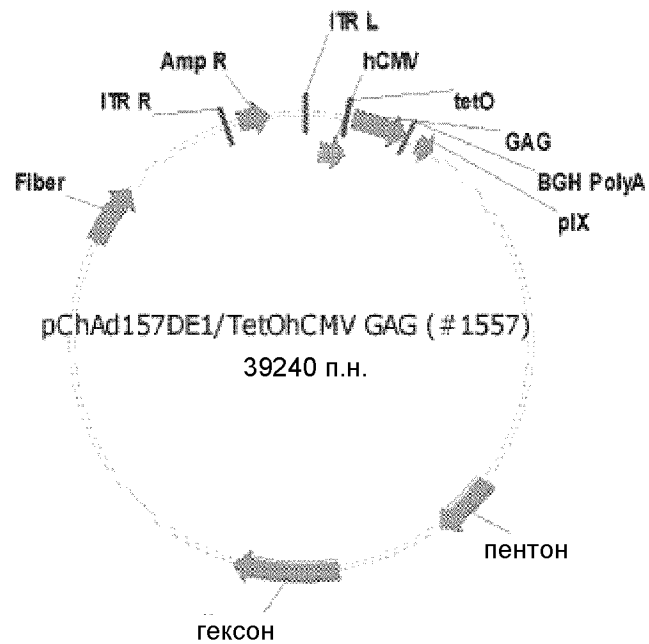
Фиг. 2



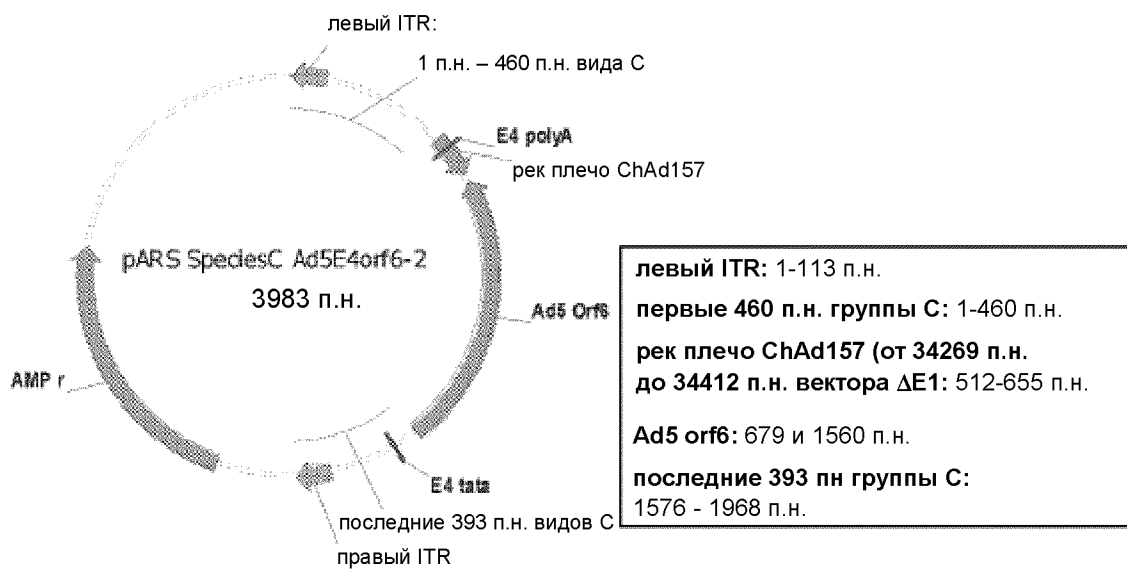
Фиг. 3



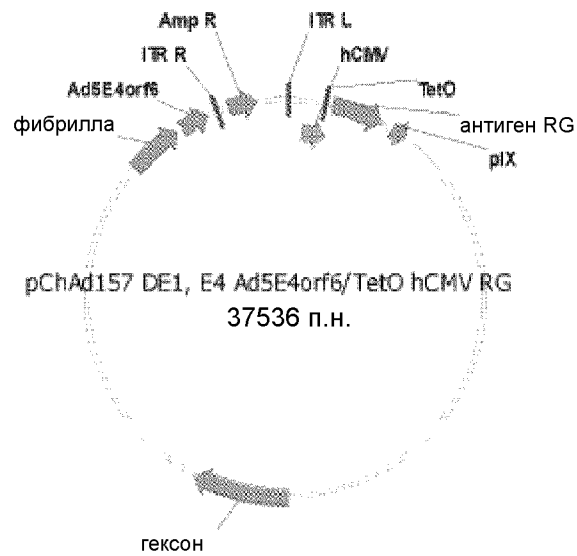
Фиг. 4



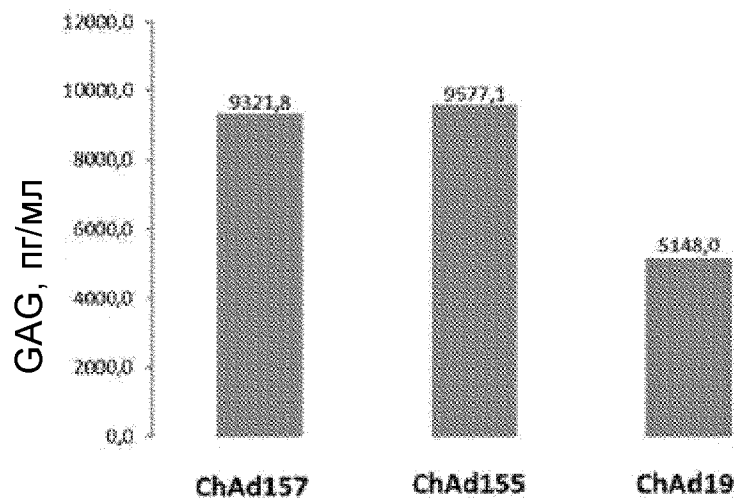
Фиг. 5



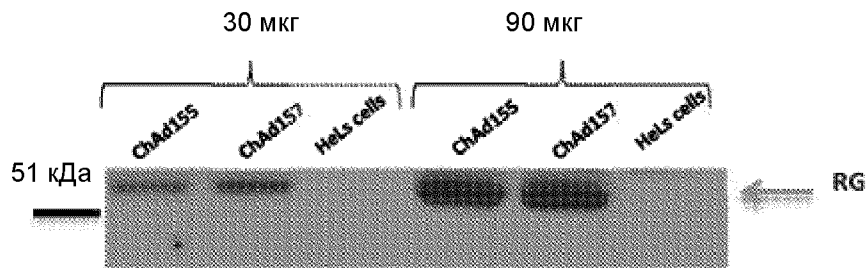
Фиг. 6



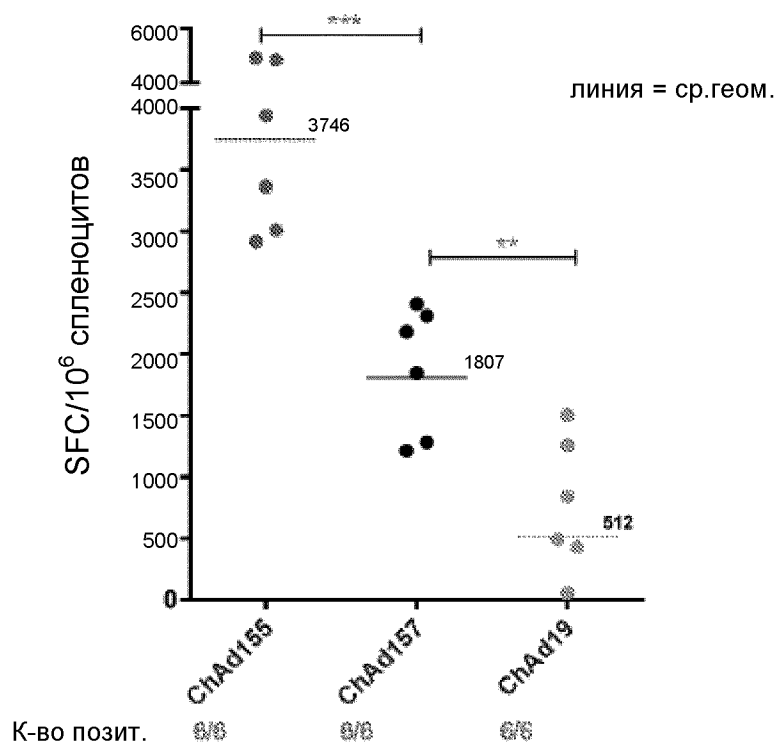
Фиг. 7



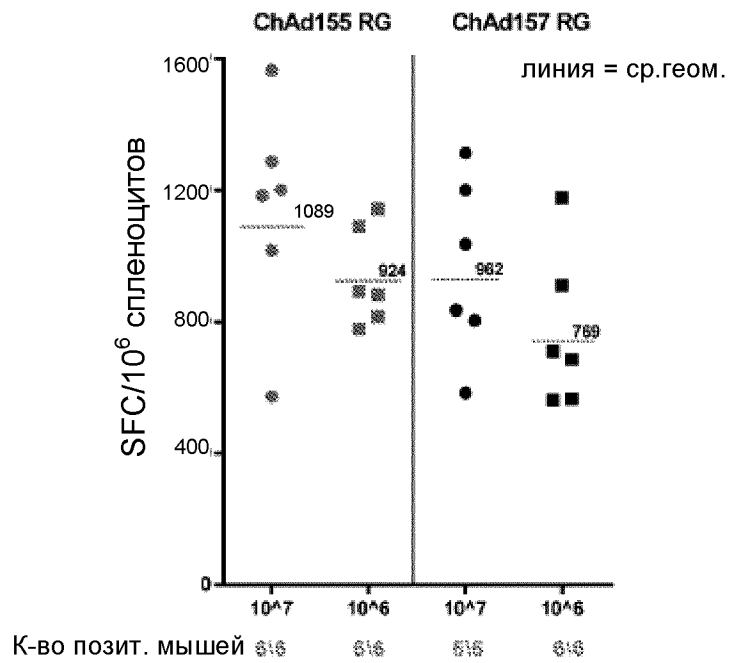
Фиг. 8



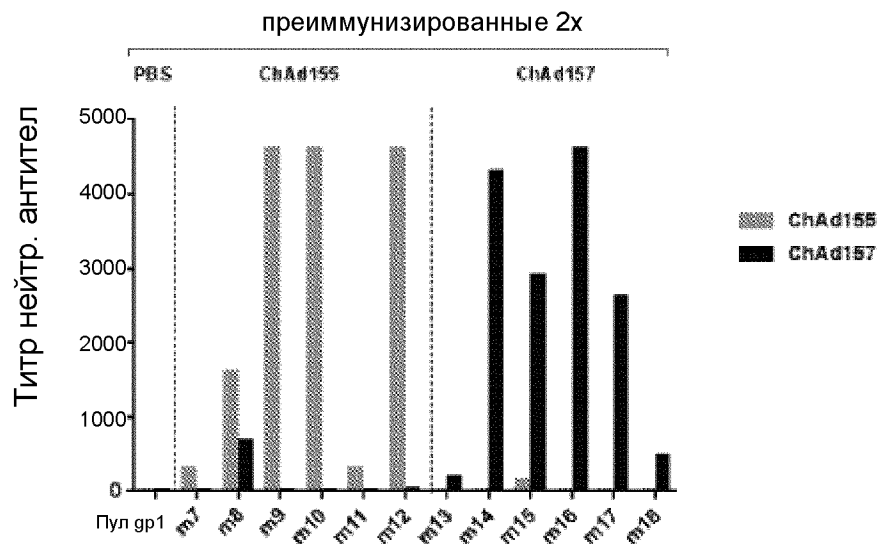
Фиг. 9



Фиг. 10



Фиг. 11



Фиг. 12

