

(19)



**Евразийское
патентное
ведомство**

(11) **040402**(13) **B1**

(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ЕВРАЗИЙСКОМУ ПАТЕНТУ

(45) Дата публикации и выдачи патента
2022.05.27

(21) Номер заявки
201790680

(22) Дата подачи заявки
2015.10.05

(51) Int. Cl. **A61K 39/21** (2006.01)
C12N 15/869 (2006.01)
C07K 14/16 (2006.01)
A61P 31/18 (2006.01)

(54) ВАКЦИНЫ ПРОТИВ ВИРУСА ИММУНОДЕФИЦИТА ЧЕЛОВЕКА (ВИЧ), СОДЕРЖАЩИЕ ОДИН ИЛИ БОЛЕЕ ЭПИСЕНСУСНЫХ (EPISENSUS) АНТИГЕНОВ ДЛЯ ПОПУЛЯЦИИ

(31) **62/059,497; 62/059,506**

(32) **2014.10.03**

(33) **US**

(43) **2017.11.30**

(86) **PCT/US2015/054067**

(87) **WO 2016/054654 2016.04.07**

(71)(73) Заявитель и патентовладелец:
**ТРИАД НЭШНЛ СЕКЬЮРИТИ,
ЛЛС; ОРЕГОН ХЕЛС ЭНД
САЙЕНС ЮНИВЕРСИТИ; ВИР
БИОТЕХНОЛОДЖИ, ИНК. (US)**

(72) Изобретатель:
**Брюнинг Эрик, Фрюэ Клаус, Пикер
Луис, Корбер Бетт Т.М., Тейлер
Джеймс, Маршалл Эмили (US)**

(74) Представитель:
Нилова М.И. (RU)

(56) US-A1-20110123485
KORBER, B et al. LANL/New Mexico
Consortium HIV Vaccine Design and Analysis. Los
Alamos National Laboratory. Presentation. 7 July 2014
[Retrieved on 25 February 2016] Retrieved from the
Internet: <URL: <http://permalink.lanl.gov/object/tr?what=info:lanl-repo/lareport/LA-UR-14-25023>>

US-A1-20140073525

WO-A1-2014138209

US-A1-20100183651

US-A1-20140186384

KIROVSKI, D et al. Combinatorics of
The Vaccine Design Problem: Definition and An
Algorithm. Microsoft Research. Technical Report.
November 2007 [retrieved on 9 December 2015].
Retrieved from the Internet: <URL: <http://research.microsoft.com/pubs/70508/tr-2007-148.pdf>>

(57) В данном изобретении предложены вакцины против ВИЧ-1, содержащие носитель и эписенсусный антиген для популяции, определенный с применением подхода EpiGraph. Также предложены вакцины против ВИЧ-1, содержащие носитель, эписенсусный антиген для популяции и индивидуализированный антиген. Также предложены способы разработки и производства вакцины против ВИЧ-1 для субъекта, включающие разработку антигенов для вакцины для оптимального покрытия разнообразия в пределах географической области с использованием последовательности аминокислот антигена, полученной с применением подхода EpiGraph, и получения указанного разработанного антигена для вакцины. Также предложены способы стимулирования ответа эффекторных Т-клеток памяти, включающие разработку одной или более последовательностей аминокислот EpiGraph, получение вакцины, содержащей одну или более последовательностей аминокислот EpiGraph и вектор, и введение указанной вакцины субъекту. Дополнительно предложены способы лечения ВИЧ-1 у субъекта, включающие введение эффективного количества описанных вакцин против ВИЧ-1 нуждающемуся в этом субъекту.

B1**040402****040402****B1**

Ссылка на родственные заявки

Настоящая заявка испрашивает приоритет согласно предварительной заявке на патент США № 62/059497, поданной 3 октября 2014 г., и предварительной заявке на патент США № 62/059506, поданной 3 октября 2014 г., описания каждой из которых настоящим полностью включены в данную заявку посредством ссылки.

Ссылка на перечень последовательностей, поданный в электронном виде

Содержимое поданного в электронном виде перечня последовательностей в текстовом файле ASCII (название: 3525.010PC02 Seq listing_ST25.txt; размер 3634253 байт и дата создания 5 октября 2015 г.), поданного вместе с настоящей заявкой, полностью включено в настоящее описание посредством ссылки.

Область техники

Объект настоящего изобретения относится в целом к ВИЧ и в частности к вакцинам против ВИЧ.

Уровень техники

В 2013 г. было зарегистрировано приблизительно 2,3 млн вновь приобретенных инфекций вирусом иммунодефицита человека (ВИЧ), более 35 млн людей, живущих с ВИЧ, и 1,6 млн людей, погибших от синдрома иммунодефицита (СПИД). Несмотря на то, что был достигнут большой прогресс в лечении ВИЧ/СПИД, всех индивидов, живущих с ВИЧ, необходимо будет лечить с помощью антиретровирусной терапии (АРТ) до конца жизни, поскольку лекарственная терапия не способна устранить резервуары латентного вируса, которые находятся в покоящихся CD4⁺ Т-клетках с частотой, приблизительно составляющей 1/10⁶ клеток. См. Eriksson, S. 2013. PLoS Pathog 9:e1003174.

Главные стратегии по очистке латентных резервуаров ВИЧ, как правило, направлены на реактивацию латентного вируса с применением ингибиторов деацетилазы гистонов (HDAC), тем не менее, клинические исследования с ингибиторами HDAC не приводили к устойчивому уменьшению латентных резервуаров. Одна вероятная причина этого заключается в неспособности специфических к ВИЧ CD8⁺ Т-клеток уничтожать покоящиеся CD4⁺Т-клетки.

Было задокументировано лишь несколько случаев, когда ВИЧ-1 был устранен из организма индивида с существующей инфекцией. До тех пор, пока не будет разработана эффективная терапия, приблизительно 35 млн индивидов, живущих с ВИЧ-1, продолжат принимать противовирусные лекарственные средства для подавления вирусного резервуара, который устоял против всех попыток уничтожения.

Излечение от СПИДа труднодостижимо, но недавно проведенная работа позволяет предположить, что строгий иммунологический контроль может с течением времени устранить ВИЧ. В частности, обнаружили, что макаки-резус (RM), вакцинированные векторами на основе цитомегаловируса (CMV), экспрессирующими антигены вируса иммунодефицита обезьян (ВИО), изначально были инфицированы, но ВИО невозможно было обнаружить по нескольким строгим критериям через один - два года после инфекции. Данный результат еще более поразителен в свете того факта, что высоко вирулентный штамм SIVmac239, использованный в данных исследованиях, свел на нет все предыдущие усилия по созданию вакцины. Данные результаты расширили современную парадигму от сфокусированной на превентивной вакцине против ВИЧ до таковой, в которой иммунотерапия ВИЧ/СПИД в конечном счете станет незаменимой частью борьбы против этой пандемии. Таким образом, дополнительно к превентивной вакцине, остается потребность в эффективной терапии для лечения индивидов, живущих с ВИЧ-1. В частности, остается потребность в разработке, производстве и тестировании профилактических и терапевтических вакцин против ВИЧ в композиции для клинического испытания.

Краткое описание изобретения

В данной заявке предложены полипептиды ВИЧ/ВИО, содержащие одну или более последовательностей антигенов EpiGraph, содержащих последовательности аминокислот, соответствующие Gag, Nef, Pol, Env из ВИЧ/ВИО, включая полноразмерные последовательности, их части или любые комбинации. Также в данной заявке предложены полипептиды ВИЧ/ВИО, содержащие одну или более последовательностей эписенсусных антигенов для популяции, содержащих последовательности аминокислот, соответствующие Gag, Nef, Pol, Env из ВИЧ/ВИО, или их комбинацию. Также в данной заявке предложены один или более носителей, содержащих полипептиды ВИЧ/ВИО, содержащие одну или более последовательностей эписенсусных антигенов для популяции. Дополнительно в данной заявке предложены полипептиды ВИЧ/ВИО, содержащие одну или более последовательностей индивидуализированных антигенов, содержащих последовательности аминокислот, соответствующие Gag, Nef, Pol, Env из ВИЧ/ВИО, или их комбинацию. Указанные полипептиды ВИЧ/ВИО согласно настоящему описанию могут содержать один или более индивидуализированных антигенов ВИЧ-1, при этом указанный полипептид содержит последовательность аминокислот, выбранную из группы, состоящей из последовательностей SEQ ID NO: 692-696 и последовательностей SEQ ID NO: 754-789. Также в данной заявке предложены один или более носителей, содержащих полипептиды ВИЧ/ВИО, содержащие одну или более последовательностей индивидуализированных антигенов. В данной заявке предложены последовательности антигенов EpiGraph, представленные в SEQ ID NO: 691-789.

В данной заявке предложены вакцины против ВИЧ-1, содержащие один или более носителей и один или более эписенсусных антигенов для популяции. Также в данной заявке предложены вакцины против ВИЧ-1, содержащие вектор, способный экспрессировать эписенсусный антиген для популяции и

один или более индивидуализированных антигенов.

В некоторых вариантах реализации указанный эписенсусный антиген для популяции ВИЧ-1 содержит эпитопы из Gag, Pol, Nef, Env, Tat, Rev, Vpr, Vif или Vpu. В некоторых вариантах реализации указанный эписенсусный антиген для популяции ВИЧ-1 представляет собой слитый антиген, содержащий два или более эписенсусных антигенов для популяции ВИЧ-1. В некоторых вариантах реализации указанный эписенсусный антиген для популяции ВИЧ-1 является основным для эпидемического ВИЧ-1 клады В в Соединенных Штатах Америки. В некоторых вариантах реализации указанный эписенсусный антиген для популяции ВИЧ-1 является основным для эпидемического ВИЧ-1 клады С в Южной Африке. В некоторых вариантах реализации указанный эписенсусный антиген для популяции ВИЧ-1 является основным для регионально эпидемического ВИЧ-1 клады 2 в Таиланде. В некоторых вариантах реализации указанный эписенсусный антиген для популяции ВИЧ-1 является основным для ВИЧ-1 группы М в мировой популяции.

В некоторых вариантах реализации указанный эписенсусный антиген для популяции ВИЧ-1 содержит эпитопы из Gag. В некоторых вариантах реализации указанный эписенсусный антиген для популяции ВИЧ-1 содержит эпитопы из консервативной области ВИЧ-1. В некоторых вариантах реализации указанный эписенсусный антиген для популяции ВИЧ-1 содержит эпитопы из консервативной области Gag, Pol или Nef. В некоторых вариантах реализации указанные эпитопы из консервативной области Gag представляют собой эпитопы из капсидного белка р24 Gag.

В некоторых вариантах реализации указанный индивидуализированный антиген ВИЧ-1 содержит эпитопы из Gag, Pol, Nef, Env, Tat, Rev, Vpr, Vif или Vpu. В некоторых вариантах реализации индивидуализированный антиген ВИЧ-1 является основным для ВИЧ-1 группы М в мировой популяции. В некоторых вариантах реализации индивидуализированный антиген ВИЧ-1 является основным для эпидемического ВИЧ-1 клады С в Южной Африке. В некоторых вариантах реализации индивидуализированный антиген ВИЧ-1 является основным для эпидемического ВИЧ-1 клады В в Соединенных Штатах Америки.

Дополнительно предложены способы предупреждения или лечения инфекции ВИЧ-1 у субъекта, включающие введение эффективного количества описанных вакцин против ВИЧ-1 нуждающемуся в этом субъекту. Дополнительно предложены способы разработки и производства вакцины против ВИЧ-1 для субъекта, включающие секвенирование вирусов ВИЧ-1 у индивида, выбор антигенов для вакцины, разработанных таким образом, чтобы они оптимально покрывали разнообразие в пределах географической области, и вставку указанных антигенов для вакцины в вектор. Также в данной заявке предложены способы лечения инфекции ВИЧ-1 у субъекта, включающие введение эффективного количества описанных вакцин нуждающемуся в этом субъекту.

Также в данной заявке предложены вакцины против ВИЧ-1, содержащие один или более носителей и эписенсусный антиген для популяции, определенный с применением подхода EpiGraph. Дополнительно в данной заявке предложены способы разработки антигенов для вакцины для оптимального покрытия разнообразия в пределах географической области с применением последовательности аминокислот антигена для вакцины, полученной с помощью способа EpiGraph выбора последовательности аминокислот антигена, и производства указанного разработанного антигена для вакцины. Дополнительно в данной заявке предложены способы разработки и производства вакцины против ВИЧ-1 для субъекта, включающие определение последовательности аминокислот вирусов ВИЧ-1 у индивида путем секвенирования, выбор антигенов для вакцины, разработанных таким образом, чтобы они оптимально покрывали разнообразие в пределах географической области, с применением последовательности аминокислот антигена для вакцины, полученной с помощью способа EpiGraph выбора последовательности аминокислот антигена, и вставку указанных антигенов для вакцины в вектор.

Также в данной заявке предложены способы стимулирования ответа эффекторных Т-клеток памяти, включающие определение одной или более последовательностей аминокислот EpiGraph, получение вакцины, содержащей указанную одну или более последовательностей аминокислот антигенов EpiGraph и один или более носителей, и введение указанной вакцины нуждающемуся в этом субъекту. Дополнительно предложены способы лечения ВИЧ-1 у субъекта, включающие введение эффективного количества описанных вакцин против ВИЧ-1 нуждающемуся в этом субъекту. Также в данной заявке предложены способы защиты от инфекции ВИЧ-1 у субъекта, включающие введение эффективного количества описанных вакцин против ВИЧ-1 нуждающемуся в этом субъекту.

Краткое описание фигур

На фиг. 1А показана полная диаграмма для клады CRF01-AE белка Nef.

Прямоугольником выделен фрагмент, который показан на фигуре 1В.

На фиг. 1В представлен фрагмент из фиг. 1А. Узлы представлены серыми точками и представляют собой каждый k-мерный вариант, где k=9. Ребра представляют собой горизонтальные линии, которые соединяют эпитопы, последовательности которых перекрываются по k-1 аминокислоте, как показано для первых двух эпитопов ($e_a=VTSSNMNNA$ [SEQ ID NO: 790], $e^*=TSSNMNNAD$ [SEQ ID NO: 791]) в верхнем левом углу. Хотя топологические свойства диаграммы не зависят от положений узлов, в данном графике используют вертикальную ось, чтобы указать частоту эпитопа в наборе целевых последовательностей $y=f(e)$ для каждого узла.

Горизонтальное положение

$$x(e) = 1 + \max_{e' \in P(e)} x(e')$$

где $P(e)$ представляет собой набор предшествующих e , наделяет данный график таким свойством, что все ориентированные ребра соединяют слева направо. Идеальный путь через данную диаграмму проходит через как можно более высокие значения y . На данной фигуре показано два пути через узлы. Сплошная черная линия представляет собой оптимальный путь и соответствует последовательности

TSSNMNNADSVWLRAQEEE (SEQ ID NO: 792),

тогда как пунктирная линия соответствует

TSSNMNNADCVWLRAQEEE (SEQ ID NO: 793).

Пунктирная линия достигает более высоких значений $f(e)$ в 4 узлах, но сплошная линия достигает более высоких $f(e)$ в 5 узлах и $\sum f(e)$ больше. Обратите внимание, что нет пути, который содержит во всех значениях x узлы с наибольшими значениями.

На фиг. 2 показаны узлы диаграммы эпитопов, показывающие последовательность эпитопа из k символов и частоту (f), с которой данный эпитоп находится в данном положении в выровненном наборе последовательностей, и совокупный оптимальный балл S соответствующих путей, которые заканчиваются в данном узле. Узлы расположены в колонках, где каждая колонка соответствует положению t , связанному с данным эпитопом. Линии, соединяющие данные узлы, соответствуют соседним эпитопам, которые совместимы. Цель состоит в том, чтобы найти соответствующий путь через данную диаграмму, который позволяет максимизировать сумму значений частоты в каждом узле. Жирные линии показывают один путь, который позволяет получить оптимальный суммарный балл.

На фиг. 3 показано, что у узлов из первоначального решения (представленного $q1=ARCHSLM$ (SEQ ID NO: 794), показанного на фиг. 2) значения частоты установили равными нулю, и совокупный балл $S(t,e)$ перерасчитали на основе данных узлов с новыми частотами. Это привело к оптимальному решению данной задачи комплементарного покрытия: $q2=DEFGKLM$ (SEQ ID NO: 795). Жирные линии показывают один путь, который позволяет получить оптимальный суммарный балл.

На фиг. 4А, В показана диаграмма эпитопов для набора из 690 последовательностей белков Gag клады В из США (последовательности SEQ ID NO: 1-690), каждая из которых выровнена до положения 556. Горизонтальная ось представляет собой положение t , и колонки узлов указывают на различные эпитопы в каждом положении. Узлы расположены таким образом, что наиболее частые в каждом положении находятся внизу каждой колонки. Полная диаграмма показана на фиг. 4А; увеличенная часть той же диаграммы показана на фиг. 4В.

На фиг. 5А-С показано, что исключение редких вариантов уменьшает покрытие, но лишь немного для Gag (фиг. 5А), Nef (фиг. 5В) и Pol (фиг. 5С). Покрытие поливалентных ($m=2$) решений показано как функция минимального значения f_0 (т.е. популяционной частоты самого редкого эпитопа в вакцине). Пунктирные линии соответствуют покрытию, полученному с помощью прямого последовательного алгоритма; черные сплошные линии построены на основе наилучших решений после 100 произвольных перезапусков. Вертикальная ось на всех трех графиках ограничена диапазоном 0,015.

На фиг. 6А, В показано покрытие для двухантигенной вакцины.

Сравнения иллюстрируют среднее покрытие эпитопов на последовательность для современных последовательностей клады В, выделенных в Соединенных Штатах Америки, которые считают гипотетической целевой популяцией для терапевтической вакцины. Для того, чтобы проиллюстрировать покрытие потенциальных Т-клеточных эпитопов (РТЕ) с применением пары природных последовательностей внутри клады В в качестве антигенов для вакцины, 5000 произвольно выбранных пар природных последовательностей клады В (серый) оценивали как потенциальные вакцины, и распределение среднего покрытия 189 последовательностей клады В из Соединенных Штатов Америки показано на серой гистограмме. Его сравнивают со средним покрытием, которое обеспечивает набор 2 антигенов из группы М EpiGraph (база данных М), набор двух антигенов из мировой клады В EpiGraph (база данных В) и индивидуализированная вакцина клады В из США, для производства которой были выбраны 2 наилучших совпадения из набора из 6 типичных EpiGraph в качестве "индивидуализированного" совпадения для каждой из 189 природных последовательностей клады В из США. На фиг. 6А показано сравнение для полноразмерного белка Gag. На фиг. 6В показаны сравнения только для консервативной области р24.

На фиг. 7 показано картирование покрытия потенциальных эпитопов по всему протеому ВИЧ.

На фиг. 8 показано среднее покрытие потенциальных эпитопов Gag последовательностей ВИЧ-1 клады В из США различными антигенами для вакцины. На фиг. 8А показано среднее покрытие эпитопов Gag одиночными антигенами для вакцины, а на фиг. 8В показано среднее покрытие эпитопов Gag парами антигенов для вакцины.

На фиг. 9А-С показано, что разработанные с помощью EpiGraph синтетические антигены ВИЧ экспрессируются в виде полноразмерных белков. Клетки HeLa трансфицировали экспрессионными плазмидами, кодирующими слитые белки gag и nef для ВИЧ или ВНО (EpiGraph 1 группы М ВИЧ (SEQ ID NO: 705) и EpiGraph 2 группы М ВИЧ (SEQ ID NO: 707), полноразмерные белки SIVmac239, гибридные белки вариантов из ВНО (SEQ ID NO: 713) и консервативные части gag и nef ВНО (SEQ ID NO: 714)), как

представлено на фиг. 9А; белки полимеразы из ВИЧ или ВИО (полноразмерные (FL) из SIVmac239, гибридные белки вариантов из ВИО (SEQ ID NO: 715), (EpiGraph 1 группы М ВИЧ (SEQ ID NO: 709) и EpiGraph 2 группы М ВИЧ (SEQ ID NO: 711), и консервативные части pol для ВИО (SEQ ID NO: 716)), как представлено на фиг. 9В; и слитые белки gag и nef для ВИЧ (gag/nef EpiGraph 1 кланды В (SEQ ID NO: 701)) и gag/nef EpiGraph 2 кланды В (SEQ ID NO: 702)), и белки полимеразы для ВИЧ (pol EpiGraph 1 кланды М (SEQ ID NO: 709)); pol EpiGraph 1 кланды В (SEQ ID NO: 703)); pol EpiGraph 2 кланды В (SEQ ID NO: 704)), как представлено на фиг. 9С. Все конструкции gag/nef содержат карбоксиконцевую метку V5 (фиг. 9А), и все конструкции pol содержат карбоксиконцевую метку HA (фиг. 9В и 9С). Для фиг. 9В 1 и 2 указывают на несколько клонов каждой конструкции. Лизаты собирали через 48 ч после трансфекции и подвергали иммуноблоттингу, применяя антитела к V5 (фиг. 9А) или HA (фиг. 9В и 9С). Наблюдаемая молекулярная масса белков соответствовала предсказанной молекулярной массе для каждой из конструкций.

На фиг. 10А, В показано, что разработанные с помощью EpiGraph синтетические антигены экспрессировались с векторов CMV. На фиг. 10А показано, что штамм 68-1 RhCMV, экспрессирующий полноразмерную полимеразу (FL) SIVmac239, гибридные полимеразы ВИО из различных вариантов ВИО (SEQ ID NO: 715), синтетический ген полимеразы на основании EpiGraph 1 мировой группы М ВИЧ (SEQ ID NO: 709) или EpiGraph 2 мировой группы М ВИЧ (SEQ ID NO: 711) конструировали с помощью мутагенеза ВАС и воссоздавали в теломеризованных фибробластах макаки-резус. На фиг. 10В показано, что штамм 68-1 RhCMV, экспрессирующий консервативные области из конструкций полимеразы SIVmac239 (SEQ ID NO: 716), также конструировали путем мутагенеза ВАС и воссоздавали в теломеризованных фибробластах макаки-резус. Во всех векторах экспрессия антигена запускалась с эндогенного вирусного промотора Rh107. Осадки клеток собирали при полном цитопатическом действии (ЦПД) и подвергали иммуноблоттингу для обнаружения метки HA, экспрессированной на карбоксильном конце каждого белка. Для консервативных областей pol два независимых клон (2.1 и 2.2) показаны на фиг. 10В.

На фиг. 11А-С показаны консервативные области внутри: Nef (фиг. 11А), Gag (фиг. 11В) и Pol (фиг. 11С), определенные на основании возможности покрытия потенциальных Т-клеточных эпитопов (PTE) бивалентной (т.е. содержащей 2 антигена) вакциной.

На фиг. 12 проиллюстрировано среднее покрытие EpiGraph каждой из консервативных областей различными антигенами для вакцины, по сравнению с более вариабельными участками белков.

Подробное описание типичных вариантов реализации

Различные термины, относящиеся к аспектам настоящего описания, используются повсюду в настоящем описании и формуле изобретения. Такие термины имеют обычные значения, понимаемые в данной области, если не указано иначе. Другие термины с конкретными определениями должны истолковываться таким образом, чтобы они согласовались с определениями, приведенными в данном описании.

В настоящем описании и в прилагаемой формуле изобретения форма единственного числа включает ссылку на множественное число, если только в содержании явно не указано иное. Таким образом, например, ссылка на "клетку" включает комбинацию двух или более клеток и тому подобное.

Подразумевают, что в данном описании в объем термина "приблизительно" по отношению к измеримому значению, такому как количество, временная продолжительность и тому подобное, входят вариации до $\pm 20\%$ от указанного значения, так как такие вариации подходят для осуществления описанных способов. Если не указано иначе, то все числа, выражающие количества ингредиентов, свойства, такие как молекулярная масса, условия реакций и так далее, используемые в настоящем описании и формуле изобретения, во всех случаях следует понимать как дополненные термином "приблизительно". Соответственно, если не указано противоположное, численные параметры, изложенные в следующем описании и прилагаемой формуле изобретения, являются приблизительными и могут изменяться в зависимости от желательных свойств, которых стремиться добиться с помощью настоящего изобретения. По меньшей мере и без попытки ограничить применение доктрины эквивалентов к объему формулы изобретения, каждый численный параметр должен по меньшей мере истолковываться с учетом количества указанных значащих разрядов и с применением обычных способов округления.

Несмотря на то, что числовые диапазоны и параметры, устанавливающие общий объем настоящего изобретения, являются приблизительными, численные значения, изложенные в конкретных примерах, приведены настолько точно, насколько это возможно. Любое численное значение, тем не менее, по своему существу включает некоторые ошибки, которые обязательно возникают в результате стандартного отклонения в измерениях при соответствующем тестировании.

Выше и повсюду в настоящем описании термин "эффективное количество" относится к количеству, которое при необходимых дозировках и в течение необходимых периодов времени приводит к достижению желательного результата в отношении лечения соответствующего расстройства, состояния или побочного эффекта. Должно быть очевидно, что эффективное количество компонентов согласно настоящему изобретению будет изменяться от пациента к пациенту не только в зависимости от конкретной выбранной вакцины, компонента или композиции, пути введения и способности указанных компонентов приводить к желательному результату у индивида, но также и от таких факторов как конкретное болезненное состояние или тяжесть состояния, которое необходимо облегчить, уровни гормонов, возраст, пол, масса тела индивида, состояние пациента и тяжесть патологического состояния, от которого лечат, со-

путствующая лекарственная терапия или особая диета, которую соблюдает конкретный пациент, и другие факторы, которые очевидны для специалистов в данной области, при этом подходящую дозировку определяет лечащий врач. Режим дозирования можно изменять, чтобы добиться лучшего терапевтического ответа. Эффективное количество также представляет собой такое количество, при котором какие-либо токсичные или вредные действия указанных компонентов перевешиваются терапевтически полезными действиями.

Термин "введение" означает либо введение непосредственно соединения или композиции согласно настоящему изобретению, либо введение пролекарства, производного или аналога, из которого будет образовываться эквивалентное количество активного соединения или вещества внутри организма.

Термины "субъект", "индивид" и "пациент" используются взаимозаменяемо в данном описании и относятся к животному, например, к человеку, которого лечат, включая профилактическое лечение, фармацевтической композицией согласно настоящему изобретению. Термин "субъект" в данном описании относится к человеку и к не относящимся к человеку животным. Термины "не относящиеся к человеку животные" и "не относящиеся к человеку млекопитающие" используют взаимозаменяемо в данном описании, и они включают всех позвоночных, например млекопитающих, таких как не относящиеся к человеку приматы (особенно высшие приматы), овца, собака, грызун (например, мышь или крыса), морская свинка, коза, свинья, кошка, кролики, коровы, лошади, и не млекопитающих, таких как рептилии, амфибии, куры и индейки.

Термин "коктейль" относится к набору антигенов, которые предназначены для введения пациенту в комбинации.

Термин "эписенсусный" относится к консенсусной последовательности на основе эпитопов. Это последовательность, эпитопы в которой совпадают, насколько это возможно, с эпитопами в эталонном наборе природных последовательностей. Термины "эпитоп" и "потенциальный эпитоп" относятся к последовательности из k символов (обычно k находится в диапазоне от 8 до 12), часто в контексте подпоследовательности из k символов в гораздо более длинной природной последовательности или последовательности антигена для вакцины. Т-клетки могут распознавать такие пептиды при иммунном ответе.

Термин "EpiGraph" относится к новой вычислительной стратегии, разработанной для создания последовательностей, в которых предусмотрена оптимальная эписенсусная последовательность или набор последовательностей, которые в комбинации обеспечивают оптимальное покрытие популяции разнообразных вирусных последовательностей.

Термины "последовательность EpiGraph" или "эписенсусный антиген" относятся к компонентам вакцины, разработанным на основе алгоритма EpiGraph.

Термин "эписенсусный антиген для популяции" относится к последовательности, полученной с помощью алгоритма EpiGraph, которая является "центральной" для популяции последовательностей ВИЧ. Указанная популяция может представлять собой конкретную кладу ВИЧ, кластер последовательностей, полученный с применением разработанного нами алгоритма кластеризации на основе индивидуализированных эпитопов, или мировую эпидемическую популяцию. Термин "центральная" обозначает наличие общих потенциальных эпитопов, таким образом указанная последовательность представляет собой рассчитанную с помощью компьютера последовательность, которая обеспечивает максимальное среднее покрытие эпитопов в данной популяции.

Последовательности EpiGraph можно применять в качестве варианта профилактической, предупредительной вакцины или можно внедрить как часть более комплексных стратегий по разработке терапевтических вакцин, которые будут индивидуализированы, чтобы они подходили для лечения инфекций индивида. Термин "набор индивидуализированных вакцин" относится к набору последовательностей антигенов для вакцины, разработанному для производства, из которого можно выбрать подгруппу антигенов для наилучшего совпадения с вирусом пациента для доставки в качестве терапевтической вакцины.

Термин "индивидуализированный антиген" или "индивидуализированный эписенсусный антиген" относится к одной или более последовательностям аминокислот из "набора индивидуализированных вакцин", которые будут специально выбраны на основании наилучшего совпадения с инфицировавшим пациента штаммом ВИЧ-1 для доставки данному пациенту в качестве терапевтической вакцины.

В данном описании термины "лечение" или "терапия" (а также различные формы данных терминов, включая "лечебный" или "паллиативный") относятся к лечению инфицированного субъекта. В данном описании термин "лечение" включает облегчение или ослабление по меньшей мере одного неблагоприятного или отрицательного эффекта или симптома состояния, заболевания или расстройства. Указанное состояние, заболевание или расстройство может представлять собой ВИЧ-инфекцию.

В данном описании термины "предупреждение" или "профилактика" относятся к предупреждению инфицирования субъекта, или снижению риска инфицирования субъекта, или остановке передачи или снижению риска передачи, например, ВИЧ, ВЮО или родственного вируса.

"Фармацевтически приемлемый" относится к таким соединениям, материалам, композициям и/или лекарственным формам, которые, в рамках медицинской точки зрения, подходят для контактирования с тканями людей и животных, не вызывая чрезмерной токсичности, раздражения, аллергического ответа или других проблемных осложнений, в соответствии с приемлемым отношением польза/риск.

Векторы, которые можно применять, включают, но не ограничены перечисленными: плазмиды, бактериальные векторы и вирусные векторы. Вирусные векторы включают цитомегаловирусные (CMV) векторы. Преимущество данных векторов CMV для применения для доставки терапевтической вакцины состоит в том, что они создают новую модель CD8⁺ Т-клеточного эпитопа и вызывают более эффективные и продолжительные ответы. В моделях на животных было показано, что вакцины на основе данных вирусных векторов могут устранять вирусные инфекции (Hansen, S. G. 2013. Science 340:1237874), и следовательно, данные подходы перспективны для создания терапевтической вакцины, в схеме лечения, в которой могут быть полезны индивидуализированные вакцины.

Другие вирусные векторы могут включать вирус оспы (вирус осповакцины), включая вирус осповакцины Анкара и вирус оспы канареек; аденовирусы, включая аденовирус типа 5 (Ad5); вирус краснухи; вирус Сендай; рабдовирус; альфавирусы и аденоассоциированные вирусы. В качестве альтернативы антигены EpiGraph для вакцины можно доставлять в виде ДНК, РНК или белковых компонентов вакцины. Так как это стратегия разработки антигена, разработанные с помощью EpiGraph антигены для вакцины будут совместимы по существу с любым способом доставки антигенов для вакцины.

В некоторых вариантах реализации вакцины, разработанные с применением последовательностей аминокислот EpiGraph, содержат один антиген. В некоторых вариантах реализации вакцины, разработанные с применением последовательностей аминокислот EpiGraph, содержат наборы антигенов.

В некоторых вариантах реализации указанные последовательности антигенов EpiGraph можно применять для получения профилактических вакцин.

В некоторых вариантах реализации стратегию EpiGraph можно применять для получения последовательностей, в которых полностью оптимизировано покрытие эпитопов для профилактической вакцины. Это может быть одна последовательность или несколько последовательностей, разработанных таким образом, чтобы они дополняли друг друга в виде превентивной вакцины. Вакцины EpiGraph можно использовать с любым вектором, включая, но не ограничиваясь перечисленными: плазмидные, бактериальные и вирусные векторы.

В алгоритме EpiGraph природные последовательности описаны в виде большой диаграммы узлов, каждый узел соответствует эпитопу, который встречается в природных последовательностях. Ориентированные ребра соединяют два узла, если соответствующие две цепочки эпитопа "совместимы", то есть последние k-1 символов в первой цепочке соответствуют первым k-1 символам во второй цепочке (например, "ЕРТАРРАЕРТАР" (SEQ ID NO: 796) и "РТАРРАЕРТАРР" (SEQ ID NO: 797) являются совместимыми k=12-мерами). Если две цепочки совместимы, то цепочка длины k+1 (например, "ЕРТАРРАЕРТАРР") (SEQ ID NO: 798) содержит оба эпитопа. В более общем смысле, путь через данную диаграмму узлов и ребер соответствует одной цепочке, которая содержит k-мерные подцепочки, соответствующие каждому из узлов на диаграмме. Каждый узел взвешен в соответствии с количеством последовательностей в эталонном наборе, содержащих подцепочку, соответствующую данному узлу. В алгоритме EpiGraph используют схему динамического программирования для обнаружения пути через данную полную диаграмму, для которого сумма данных весов максимальна, и, следовательно, обеспечивают наилучшее покрытие.

В некоторых вариантах реализации предложена вакцина против ВИЧ-1, содержащая вектор и эписенсусный антиген для популяции или комбинацию оптимизированных антигенов EpiGraph, разработанную таким образом, чтобы использовать их в виде набора. Существует множество различных векторов, которые можно использовать для получения вакцины. Например, один тип вектора, который можно применять, представляет собой вирусный вектор. Данные вирусные векторы могут включать цитомегаловирус человека (HCMV), вирус оспы, аденовирусы, вирус краснухи, вирус Сендай, рабдовирус, альфавирус или аденоассоциированный вирус. Указанную вакцину также можно доставлять в виде гена, кодирующего белок EpiGraph, применяя либо ДНК, либо РНК, или в виде экспрессированного белка или части белка.

В некоторых вариантах реализации антиген EpiGraph из описанных вакцин может быть получен из белка Gag ВИЧ-1. В другом варианте реализации белок Gag ВИЧ-1 инактивировали путем удаления последовательности остатков миристиновой кислоты на N-конце белка Gag ВИЧ-1. В другом варианте реализации антиген EpiGraph может быть получен из белка Pol или Nef ВИЧ-1 или, в действительности, из любого другого белка ВИЧ, включая Env, Tat, Rev, Vif или Vpr.

В некоторых вариантах реализации эписенсусный антиген для популяции можно определить, применяя подход EpiGraph. Данный эписенсусный антиген для популяции затем можно применять для получения вакцины. В качестве альтернативы антигены EpiGraph можно разработать в виде комбинации последовательностей, разработанных таким образом, чтобы использовать их в виде набора. ВИЧ-1 можно разбить на клады на основании географического положения. В одном варианте реализации в описанных вакцинах эписенсусный антиген для популяции является основным для эпидемического ВИЧ-1 клады В в Соединенных Штатах Америки. В другом варианте реализации в описанных вакцинах эписенсусный антиген для популяции является основным для эпидемического ВИЧ-1 клады С в Южной Африке. В другом варианте реализации в описанных вакцинах эписенсусный антиген для популяции является основным для регионально эпидемического ВИЧ-1 клады 2 в Таиланде. В другом варианте реализации в

ВИЧ-1 клады С, содержащий эпитопы из Nef. В некоторых вариантах реализации эписенсусный антиген для популяции ВИЧ-1 клады С представляет собой слитый антиген, содержащий (1) эписенсусный антиген для популяции ВИЧ-1 клады С, содержащий эпитопы из капсидного белка р24 Gag, и (2) эписенсусный антиген для популяции ВИЧ-1 клады С, содержащий эпитопы из консервативной области Nef.

В некоторых вариантах реализации указанный эписенсусный антиген для популяции ВИЧ-1 является основным для регионально эпидемического ВИЧ-1 клады 2 в Таиланде. В некоторых вариантах реализации эписенсусный антиген для популяции ВИЧ-1 клады 2 содержит эпитопы из Gag. В некоторых вариантах реализации эписенсусный антиген для популяции ВИЧ-1 клады 2 содержит эпитопы из капсидного белка р24 Gag. В некоторых вариантах реализации эписенсусный антиген для популяции ВИЧ-1 клады 2 содержит эпитопы из Nef. В некоторых вариантах реализации эписенсусный антиген для популяции ВИЧ-1 клады 2 содержит эпитопы из консервативной области Nef. В некоторых вариантах реализации эписенсусный антиген для популяции ВИЧ-1 клады 2 содержит эпитопы из Pol. В некоторых вариантах реализации эписенсусный антиген для популяции ВИЧ-1 клады 2 представляет собой слитый антиген, содержащий (1) эписенсусный антиген для популяции ВИЧ-1 клады 2, содержащий эпитопы из Gag, и (2) эписенсусный антиген для популяции ВИЧ-1 клады 2, содержащий эпитопы из Nef. В некоторых вариантах реализации эписенсусный антиген для популяции ВИЧ-1 клады 2 представляет собой слитый антиген, содержащий (1) эписенсусный антиген для популяции ВИЧ-1 клады 2, содержащий эпитопы из капсидного белка р24 Gag, и (2) эписенсусный антиген для популяции ВИЧ-1 клады 2, содержащий эпитопы из консервативной области Nef.

В некоторых вариантах реализации указанный эписенсусный антиген для популяции ВИЧ-1 является основным для ВИЧ-1 группы М в мировой популяции. В некоторых вариантах реализации эписенсусный антиген для популяции ВИЧ-1 группы М содержит эпитопы из Gag и содержит последовательность аминокислот, представленную в SEQ ID NO: 718, SEQ ID NO: 720 или SEQ ID NO: 754. В некоторых вариантах реализации эписенсусный антиген для популяции ВИЧ-1 группы М содержит эпитопы из капсидного белка р24 Gag и содержит последовательность аминокислот, представленную в SEQ ID NO: 719, SEQ ID NO: 721 или SEQ ID NO: 755. В некоторых вариантах реализации эписенсусный антиген для популяции ВИЧ-1 группы М содержит эпитопы из Nef и содержит последовательность аминокислот, представленную в SEQ ID NO: 722 или SEQ ID NO: 724. В некоторых вариантах реализации эписенсусный антиген для популяции ВИЧ-1 группы М содержит эпитопы из консервативной области Nef и содержит последовательность аминокислот, представленную в SEQ ID NO: 723 или SEQ ID NO: 725. В некоторых вариантах реализации эписенсусный антиген для популяции ВИЧ-1 группы М представляет собой слитый антиген, содержащий (1) эписенсусный антиген для популяции ВИЧ-1 группы М, содержащий эпитопы из Gag, и (2) эписенсусный антиген для популяции ВИЧ-1 группы М, содержащий эпитопы из Nef, при этом указанный слитый антиген содержит последовательность аминокислот, представленную в SEQ ID NO: 705 или SEQ ID NO: 707. В некоторых вариантах реализации эписенсусный антиген для популяции ВИЧ-1 группы М представляет собой слитый антиген, содержащий (1) эписенсусный антиген для популяции ВИЧ-1 группы М, содержащий эпитопы из капсидного белка р24 Gag, и (2) эписенсусный антиген для популяции ВИЧ-1 группы М, содержащий эпитопы из консервативной области Nef, при этом указанный слитый антиген содержит последовательность аминокислот, представленную в SEQ ID NO: 706 или SEQ ID NO: 708. В некоторых вариантах реализации эписенсусный антиген для популяции ВИЧ-1 группы М содержит эпитопы из Pol и содержит последовательность аминокислот, представленную в SEQ ID NO: 709, SEQ ID NO: 711, SEQ ID NO: 726 или SEQ ID NO: 728. В некоторых вариантах реализации эписенсусный антиген для популяции ВИЧ-1 группы М содержит эпитопы из консервативной области Pol и содержит последовательность аминокислот, представленную в SEQ ID NO: 710, SEQ ID NO: 712, SEQ ID NO: 727 или SEQ ID NO: 729.

В некоторых вариантах реализации описанные вакцины могут дополнительно содержать вектор HCMV, содержащий остов HCMV и индивидуализированный антиген, выбранный как наилучшим образом совпадающий с таковым для природного штамма ВИЧ-1. В другом варианте реализации описанные вакцины могут дополнительно содержать вектор HCMV, содержащий остов HCMV и индивидуализированный антиген, выбранный как наилучшим образом совпадающий с таковым для отличного штамма ВИЧ-1.

В некоторых вариантах реализации индивидуализированный антиген ВИЧ-1 содержит эпитопы из Gag, Pol, Nef, Env, Tat, Rev, Vpr, Vif или Vpu. В некоторых вариантах реализации индивидуализированный антиген содержит последовательность, выбранную из: SEQ ID NO: 692-696; SEQ ID NO: 756-765; SEQ ID NO: 769-777 или SEQ ID NO: 780-789.

В некоторых вариантах реализации индивидуализированный антиген ВИЧ-1 содержит эпитопы из Gag и содержит последовательность аминокислот, представленную в SEQ ID NO: 692, SEQ ID NO: 693, SEQ ID NO: 694, SEQ ID NO: 695 или SEQ ID NO: 696.

В некоторых вариантах реализации индивидуализированный антиген ВИЧ-1 является основным для ВИЧ-1 группы М в мировой популяции. В некоторых вариантах реализации индивидуализированный антиген ВИЧ-1 группы М содержит эпитопы из Gag и содержит последовательность аминокислот, вы-

бранную из группы, состоящей из последовательностей SEQ ID NO: 756-765. В некоторых вариантах реализации индивидуализированный антиген ВИЧ-1 группы М содержит эпитопы из капсидного белка р24 Gag и содержит последовательность аминокислот, представленную в SEQ ID NO: 757, SEQ ID NO: 759, SEQ ID NO: 761, SEQ ID NO: 763 или SEQ ID NO: 765.

В некоторых вариантах реализации индивидуализированный антиген ВИЧ-1 является основным для эпидемического ВИЧ-1 клады С в Южной Африке. В некоторых вариантах реализации индивидуализированный антиген ВИЧ-1 клады С содержит эпитопы из Gag и содержит последовательность аминокислот, выбранную из группы, состоящей из последовательностей SEQ ID NO: 768-777. В некоторых вариантах реализации индивидуализированный антиген ВИЧ-1 клады С содержит эпитопы из капсидного белка р24 Gag и содержит последовательность аминокислот, представленную в SEQ ID NO: 769, SEQ ID NO: 771, SEQ ID NO: 773, SEQ ID NO: 775 или SEQ ID NO: 777.

В некоторых вариантах реализации индивидуализированный антиген ВИЧ-1 является основным для эпидемического ВИЧ-1 клады В в Соединенных Штатах Америки. В некоторых вариантах реализации индивидуализированный антиген ВИЧ-1 клады В содержит эпитопы из Gag и содержит последовательность аминокислот, выбранную из группы, состоящей из последовательностей SEQ ID NO: 780-789. В некоторых вариантах реализации индивидуализированный антиген ВИЧ-1 клады В содержит эпитопы из капсидного белка р24 Gag и содержит последовательность аминокислот, представленную в SEQ ID NO: 781, SEQ ID NO: 783, SEQ ID NO: 785, SEQ ID NO: 787 или SEQ ID NO: 789.

В некоторых вариантах реализации предложены способы лечения ВИЧ-1 у субъекта, включающие введение эффективного количества одной из описанных вакцин нуждающемуся в этом субъекту. В другом варианте реализации при выборе антигенов для вакцины для оптимального покрытия разнообразия в некоторой географической области используют последовательность антигена для вакцины, полученную с применением способа EpiGraph выбора последовательности антигена.

В вариантах реализации настоящего изобретения предложены способы лечения инфекции ВИЧ-1 у субъекта, включающие введение эффективного количества любой из данных описанных вакцин нуждающемуся в этом субъекту.

В вариантах реализации настоящего изобретения также предложены способы стимулирования ответа эффекторных Т-клеток памяти, включающие определение одной или более последовательностей EpiGraph, получение вакцины, содержащей одну или более последовательностей аминокислот EpiGraph и вектор; и введение указанной вакцины нуждающемуся в этом субъекту. В другом варианте реализации предложены способы индуцирования ответа эффекторных Т-клеток памяти, в которых одна или более последовательностей аминокислот EpiGraph включают SEQ ID NO: 697, SEQ ID NO: 698, SEQ ID NO: 699 или SEQ ID NO: 700. Данная вакцина может представлять собой профилактическую или терапевтическую вакцину.

Недавние открытия в области разработки вакцин против ВИЧ включают идею вакцины, индуцирующей эффекторные Т-клетки памяти (ТЕМ), чтобы предотвратить инфицирование ВИЧ. В отличие от центральных Т-клеток памяти (ТСМ), которые индуцируют с помощью обычных вакцин, ТЕМ постоянно находятся в лимфоидных тканях и экстралимфоидных эффекторных сайтах и готовы сразу опосредовать противовирусную эффекторную функцию, таким образом создавая постоянный иммунный щит на воротах проникновения вируса и сайтах реактивации вируса. Наиболее подходящую векторную систему для индуцирования и неограниченного поддержания ТЕМ получили из CMV. Предположительно благодаря непрерывному низкому уровню реактивации и/или экспрессии генов в постоянно инфицированных клетках CMV поддерживает как раз подходящий постоянный низкий уровень иммунной стимуляции, необходимый для поддержания ТЕМ без запуска истощения Т-клеток.

В некоторых вариантах реализации коктейль индивидуализированных антигенов можно применять в контексте терапевтических вакцин. Например, для создания индивидуализированной вакцины можно применить стратегию кластерного анализа методом k-средних к определенному набору из 6-10 последовательностей, которые обеспечивают очень хорошее покрытие эпитопов в популяции людей, инфицированных высоковариабельным патогеном, таким как ВИЧ. Вирус, инфицирующий субъекта, можно секвенировать, а затем доставить 2 или 3 индивидуализированных вакцины, которые наилучшим образом покрывают инфицирующий вирус. Покрытие эпитопов будет оптимизировано, тогда как несовпадения эпитопов между вакциной и инфицирующим штаммом будут минимизированы.

В некоторых вариантах реализации предложена вакцина против ВИЧ-1, содержащая остов вектора HCMV, в котором отсутствуют некоторые участки генов CMV, и эписенсусный антиген для популяции. В некоторых вариантах реализации в остоге HCMV может отсутствовать участок гена UL131A-128. В некоторые варианты реализации также может быть включена делеция белка тегумента pp71 (гена UL82) (публикации заявок на патент США № 2014-0141038; 2008-0199493; 2013-0142823; и публикация заявки на международный патент WO 2014/138209).

В некоторых вариантах реализации в настоящем описании предложены вакцины, которые могут содержать последовательность второго индивидуализированного антигена. В некоторых вариантах реализации указанные вакцины могут содержать последовательности второго, третьего или более индивидуализированных антигенов. Эписенсусный антиген может включать последовательность аминокислот,

представленную в SEQ ID NO: 691. Индивидуализированный антиген может включать последовательность аминокислот, представленную в SEQ ID NO: 692, SEQ ID NO: 693, SEQ ID NO: 694, SEQ ID NO: 695 или SEQ ID NO: 696. В одном варианте реализации последовательность первого индивидуализированного антигена и последовательность второго индивидуализированного антигена включают две из следующих последовательностей аминокислот SEQ ID NO: 692, SEQ ID NO: 693, SEQ ID NO: 694, SEQ ID NO: 695 и SEQ ID NO: 696. В другом варианте реализации последовательность первого индивидуализированного антигена, последовательность второго индивидуализированного антигена и последовательность третьего индивидуализированного антигена включают три из следующих последовательностей аминокислот: SEQ ID NO: 692, SEQ ID NO: 693, SEQ ID NO: 694, SEQ ID NO: 695 и SEQ ID NO: 696.

Также предложены последовательности индивидуализированных антигенов, которые могут включать SEQ ID NO: 692, SEQ ID NO: 693, SEQ ID NO: 694, SEQ ID NO: 695 или SEQ ID NO: 696.

В некоторых вариантах реализации описанные вакцины могут дополнительно содержать вектор HCMV, содержащий остов HCMV и индивидуализированный антиген, выбранный как наилучшим образом совпадающий с таковым для природного штамма ВИЧ-1. В другом варианте реализации описанные вакцины могут дополнительно содержать вектор HCMV, содержащий остов HCMV и индивидуализированный антиген, выбранный как наилучшим образом совпадающий с таковым для отличного (и далекого) штамма ВИЧ-1.

В некоторых вариантах реализации эписенсусный антиген и индивидуализированный антиген из описанных вакцин может быть получен из белка Gag ВИЧ-1. В другом варианте реализации белок Gag ВИЧ-1 инактивировали путем удаления последовательности остатков миристиновой кислоты на N-конце белка Gag ВИЧ-1. В другом варианте реализации эписенсусный антиген и индивидуализированный антиген может быть получен из белка Pol или Nef ВИЧ-1.

В некоторых вариантах реализации предложены способы лечения ВИЧ-1 у субъекта, включающие введение эффективного количества одной из описанных вакцин нуждающемуся в этом субъекту. В другом варианте реализации предложены способы разработки и производства вакцины против ВИЧ-1 для субъекта, включающие секвенирование вирусов ВИЧ-1 у индивида, выбор антигенов для вакцины, разработанных таким образом, чтобы они оптимально покрывали разнообразие в пределах географической области, и вставку указанных антигенов для вакцины в вектор.

В некоторых вариантах реализации предложены способы лечения инфекции ВИЧ-1 у субъекта, включающие введение эффективного количества любой из данных описанных вакцин нуждающемуся в этом субъекту.

Следующие примеры предложены для более подробного описания вариантов реализации, описанных в данном описании. Они предназначены для иллюстрирования, но не ограничения, вариантов реализации.

Все документы, патенты и заявки на патент, цитированные в данном описании, настоящим включены посредством ссылки и могут использоваться для осуществления настоящего изобретения.

Примеры

Пример 1. Графическая модель для оптимального покрытия эпитопов выровненных последовательностей.

Предполагается существование набора $S = \{s_1, s_2, \dots, s_N\}$ из N выровненных последовательностей. Каждая последовательность представляет собой цепочку из буквенных символов (соответствующих двадцати аминокислотам, возможно дополнительно к нескольким специальным символам, соответствующим гэпам, неизвестным и тому подобному). Поскольку указанные последовательности выровнены, все имеют одинаковую длину T , и положения $t = 1 \dots T$ четко определяются от последовательности к последовательности; $s_n[t]$ будет записан как t -й символ в n -й последовательности. Полезно ввести условное обозначение $s[t:u]$ для подпоследовательности s , которая начинается в положении t и заканчивается в положении u .

Потенциальный эпипотоп определили как короткую последовательность из k символов, обычно от 8 до 12. Интересующие эпипотопы представляют собой подпоследовательности последовательностей из S . В действительности, последовательность s можно представить как перечень эпипотопов: $e_1, \dots, e_{T-k+1}$, при $e_t = s[t:t+k-1]$. Тем не менее, стоит отметить, что для того, чтобы перечень эпипотопов был связан с последовательностью, необходимо, чтобы эпипотопы были совместимыми. В частности, для любой соседней пары e_t и e_{t+1} необходимо, чтобы последние $k-1$ символов из e_t совпадали с первыми $k-1$ символами из e_{t+1} ; то есть: $e_t[2:k] = e_{t+1}[1:k-1]$.

С каждым эпипотопом e может быть ассоциирована частота $f(e)$, которая учитывает фракцию последовательностей из S , в которых встречается указанный эпипотоп.

Для решения проблемы с основным коктейлем ищут набор искусственных последовательностей $Q = \{q_1, \dots, q_M\}$, эпипотопы в которых обладают полезными свойствами. Пусть E_Q представляет собой набор из всех эпипотопов, которые встречаются в Q ; то есть:

$$E_Q = \{q|q = q_m[t:t+k-1] \text{ для некоторых выбранных } m, t\}.$$

В частности, такие эпипотопы были разработаны таким образом, чтобы они покрывали насколько возможно больше последовательностей из набора S . Цель состояла в максимизировании балла покрытия

$\sum f(e)$, где указанная сумма по всем $e \in E_Q$

При $M < N$, и обычно $M \ll N$, все эпитопы обычно невозможно покрыть.

Для решения проблемы с выровненными последовательностями можно сделать некоторые упрощения. В этом случае каждое положение в данном выравнивании трактовали как отличную (но не независимую) задачу. Функция частоты, теперь зависящая от положения: $f(t, e)$, представляет собой количество последовательностей в S , для которых e представляет собой k -мерную подпоследовательность, которая начинается в положении t ; таким образом,

$$f(t, e) = \sum I(e = s_n[t : t + k - 1])$$

где I представляет собой индикаторную функцию: она равна единице, если ее аргумент - истина, и нулю в противном случае. Набор E_Q также можно разбить в соответствии с положением в последовательностях q . Таким образом: $E_Q[t] = \{e | e = q_m[t:t+k-1] \text{ для некоторых } m\}$.

Это позволяет получить произведение меры покрытия для каждого положения $c[t] = \sum_{e \in E_Q[t]} f(t, e)$, по которому можно определить общее покрытие как $c = \sum c[t]$.

Для обычного частного случая $k=1$ указанные эпитопы представляли собой лишь символы аминокислот. Но даже несмотря на то, что данную задачу было легко решить, решение, тем не менее, было полезным. Для $M=1$ наилучшим решением является консенсусная последовательность, где в качестве $q_1[t]$ выбирают любую аминокислоту, которая наиболее часто встречается в положении t . Для $M=2$ можно оптимизировать покрытие, взяв в качестве $q_2[t]$ вторую по встречаемости аминокислоту в положении t . И так далее для больших значений M .

Для $k>1$ такая задача становится нетривиальной, потому что эпитопы перекрываются, и, следовательно, каждый $c[t]$ более нельзя оптимизировать независимо. Случай $M=1$ назвали "эписенсусной" задачей, потому что она похожа на консенсус, за исключением того, что представляет собой искомым консенсус эпитопов.

В одном примере консенсус и $k=3$ эписенсус не совпадают. В консенсусе использовали наиболее часто встречающуюся букву в каждом положении. В эписенсусе использовали наиболее часто встречающийся "эпитоп", где мы называем потенциальным эпитопом цепочку из 3 символов. В табл. 1 подробно описан данный пример, чтобы проиллюстрировать перекрывающиеся цепочки эпитопов.

Таблица 1. Показано шесть последовательностей и соответствующие им $k=3$ эпитопы. В нижней строке показаны консенсусные последовательности (образованные из наиболее часто встречающегося символа в каждом положении) и наиболее часто встречающиеся эпитопы в каждом положении. Но данные эпитопы (в частности, EFG и CHS) не совместимы друг с другом, то есть их нельзя объединить в одно эписенсусное решение. В данном случае наилучший эписенсусный балл (хотя и неоднозначно) получили для последовательности ARCHSLM (SEQ ID NO: 794), которая покрывает $1+1+2+2+3=9$ из 30 возможных эпитопов в данных последовательностях. Консенсусная последовательность ARCGSLM (SEQ ID NO: 799) покрывает $1+1+1+1+3=7$. Примечательно, что верхнюю границу для данного балла можно получить из частоты наиболее часто встречающихся в каждом положении эпитопов: в данном случае она равна $2+2+2+2+3=11$.

	Последовательности			Маленькие эпитопы		
[SEQ ID NO: 800]	ARCGSPM	ARC	RCG	CGS	GSP	SPM
[SEQ ID NO: 801]	ARYGSNM	ARY	RYG	YGS	GSN	SNM
[SEQ ID NO: 802]	AYCHSLM	AYC	YCH	CHS	HSL	SLM
[SEQ ID NO: 803]	YRCHSLM	YRC	RCH	CHS	HSL	SLM
[SEQ ID NO: 804]	DEFGSLM	DEF	EFG	FGS	GSL	SLM
[SEQ ID NO: 805]	DEFGKLM	DEF	EFG	FGK	TKL	KLM

Решение эписенсусной задачи.

Сначала рассматривали случай $M=1$, в котором искали одну последовательность q , эпитопы которой оптимально покрывали эпитопы в перечне невыровненных последовательностей S . Алгоритм Epi-Graph при подходящих допущениях позволяет добиться оптимального решения. В данных сравнениях также рассматривали консенсусный алгоритм (очень простой и быстрый) и оптимизацию с помощью генетического алгоритма (очень медленный), описанного в (Fisher, Nat Med. 2007 Jan;13(1):100-6, включен в данную заявку посредством ссылки). Алгоритм EpiGraph показан на фиг. 1A и 1B.

Позже рассматривали задачу более общего коктейля вакцин с $M>1$ и показали, как эписенсусный алгоритм был модифицирован для данной более общей задачи.

На фиг. 2 показаны узлы диаграммы эпитопов, на которой каждый узел содержит последовательность эпитопа из k -символов, и частота (f), с которой указанный эпитоп наблюдают в данном положении в выровненном наборе последовательностей, и совокупный оптимальный балл S соответствующих путей, которые заканчиваются в данном узле. Узлы расположили в колонках, при этом каждая колонка соответствовала положению t , связанному с эпитопом. Линии, соединяющие данные узлы, соответствовала

соседним эпитомам, которые совместимы. Цель состояла в том, чтобы провести соответствующий путь через данную диаграмму, который дает максимальную сумму значений частоты в каждом узле.

Жирные линии на фиг. 2 показывают один путь, который позволяет получить оптимальный суммарный балл. Всегда будет по меньшей мере один такой путь, но он может быть не уникальным.

Совокупный балл $S(t, e)$ определяли как наиболее высокий балл, достижимый, если начать с $t=1$ и закончить на эпитопе e в положении t . Наблюдали, что $S(t=1, e)=f(e)$ и что значения для $t>1$ можно вычислить рекурсивно:

$$S(t, e) = f(e) + \max_{e'} S(t-1, e') \text{ при } e' \in E(t-1, e)$$

где $E(t-1, e)$ представляет собой набор эпитопов в положении $t-1$, которые совместимы с e , который находится в положении t . После вычисления данного совокупного балла для всех эпитопов, нашли суммарный балл для наилучшего пути как максимальный балл в последней колонке: $S_{\max} = \max_e S(T-k+1, e)$. Более того, можно пойти в обратном направлении от данного максимума, чтобы найти оптимальный путь:

$$e_{T-k+1}^* = \operatorname{argmax}_e S(T-k+1, e)$$

$$e_{t-1}^* = \operatorname{argmax}_{e'} S(t-1, e') \text{ при } e' \in E(t-1, e_{t-1}^*)$$

Последовательность

$$e_1^*, e_2^*, \dots, e_{T-k}^*$$

описывала результирующую последовательность эпитопов с наивысшим баллом. Эписенсусную цепочку q получали, взяв первый символ из каждого эпитопа $q[t]=e_t^*[1]$ и закончив на последнем эпитопе $q[T-k:T]=e_{T-k}^*[1-k]$.

Оператор argmax возможно не имеет уникального значения; и если не имеет, то будет несколько решений эписенсусной задачи, все из которых оптимальны в смысле покрытия.

Гэпы.

Для того, чтобы выравнивать последовательности, необходимо учитывать вставки и делеции, и для этого в выровненные последовательности введены гэпы. Например, последовательности ARCCDEGH (SEQ ID NO: 806) и ARCDEFGH (SEQ ID NO: 807) лучше выравнивались в виде ARCCDE-GH (SEQ ID NO: 808) и ARC-DEFGH (SEQ ID NO: 809). Эпитопы с заполнителем были разработаны, чтобы учитывать такие гэпы в алгоритме EpiGraph.

Эпитопы с заполнителем.

Эпитопы с $k=3$ представляли собой ARC, RCC, CCD, CDE, DEG, EGH и ABC, BCD, CDE, DEF, EFG, FGH соответственно; но когда эпитопы были выровнены по первой колонке, необходимо было ввести гэпы в данный перечень: ABC, BCC, CCD, CDE, DEG, EGH, -GH и ABC, BCD, CDE, -DE, DEF, EFG, FGH. Цепочки -GH и -DE представляли собой эпитопы с заполнителем. Заполнители не учитываются в функции покрытия; то есть: $f(t; -XY)=0$. Но они все же были полезны, так как их можно применять для определения совместимости соседних эпитопов.

Для последовательностей без гэпов два соседних эпитопа считают совместимыми, если последние $k-1$ символов первого эпитопа совпадают с первыми $k-1$ символами второго эпитопа. Таким образом, ABC и BCD являются совместимыми, а BCC и CDE - нет. Если вводят гэпы, то пару считают совместимой, если второй эпитоп начинается с символа гэпа, а остальные $k-1$ символов совпадают с последними $k-1$ символами последнего эпитопа. Таким образом, CDE и -DE являются совместимыми.

Для алгоритма заместительного размещения ("drop-in-place") использовали последовательность "субстрата", которая, как правило, представляла собой консенсусную последовательность. Затем все эпитопы во всех положениях брали и сортировали в соответствии с их частотой встречаемости. Начиная с наименее часто встречающегося эпитопа, каждый эпитоп размещается на субстрате путем замены символов в субстрате в положениях $[t:t+k-1]$ на символы в указанном эпитопе.

Когда наиболее часто встречающийся эпитоп размещается на субстрате, цепочку используют в качестве эписенсусного решения q . Поскольку наиболее часто встречающиеся эпитопы "перезаписывали" более редкие эпитопы, добивались более высокого покрытия эпитопов. И поскольку единая последовательность q постоянно обновляется, то конечное решение будет состоять из совместимых эпитопов.

Данный алгоритм может не быть полностью детерминированным, потому что у некоторых эпитопов могут быть идентичные частоты, и если их положения перекрываются, то конечный результат может зависеть от того, в каком порядке их размещали. При осуществлении берется порядок, который дает алгоритм сортировки, но есть возможность рандомизировать такой порядок и осуществить несколько запусков данного алгоритма, при этом в некоторых запусках можно получить более высокие баллы.

Эвристический алгоритм "заместительного размещения". В данном алгоритме в качестве последовательности "субстрата" брали консенсусную последовательность, но выбор субстрата редко имеет значение. Все эпитопы во всех положениях брали и сортировали в соответствии с их частотой встречаемости. Начиная с наименее часто встречающегося эпитопа.

Каждый эпитоп "размещали" на субстрате путем замены символов в субстрате в положениях $[t:t+k-1]$

на символы в указанном эпитопе. Когда наиболее часто встречающийся эпитоп, наконец, размещался на субстрате, получали цепочку, которую использовали в качестве эписенсусного решения q . Поскольку наиболее часто встречающиеся эпитопы перезаписывали более редкие эпитопы, добивались высокого покрытия эпитопов. Поскольку единая последовательность q постоянно обновлялась, то конечное решение состояло из совместимых эпитопов.

Данный алгоритм может не быть полностью детерминированным, потому что у некоторых эпитопов могут быть идентичные частоты, и если их положения перекрываются, то конечный результат может зависеть от того, в каком порядке их размещали. При осуществлении берется порядок, который дает алгоритм сортировки, но есть возможность рандомизировать такой порядок и осуществить несколько запусков данного алгоритма, при этом в некоторых запусках можно получить более высокие баллы. Польза данного рандомизированного подхода с несколькими запусками не изучалась.

Решение зависит только от наиболее часто встречающихся эпитопов в каждом положении, так для примера в табл. 1 оно будет представлять собой совместимую комбинацию DEF, EFG, CHJ, HJL, JLM. Если они размещены в данном порядке (частота первых четырех составляет 2, частота последнего составляет 3), то получают последовательность DECHJLM (SEQ ID NO: 810), для которой балл $0+0+2+2+3=7$ выше, чем у консенсусного, но меньше, чем оптимальный. Если они размещены в порядке HJL, CHJ, EFG, DEF, JLM, то получают последовательность DEFGJLM (SEQ ID NO: 811) с баллом $2+2+1+1+3=9$, который представляет собой оптимальный балл для данного примера. Данные два решения показаны на фиг. 3.

Задача для выровненного коктейля вакцин ($M>1$). В исходном мозаичном решении, в котором используют оптимизацию генетического алгоритма, все M из мозаичных последовательностей оптимизируют одновременно.

Последовательные решения.

Один способ расширения эписенсусных решений $M=1$ на задачи $M>1$ состоит в модификации алгоритмов для оптимизации общего покрытия, чтобы оптимизировать комплементарное покрытие. А именно: дано эписенсусное решение q_1 , нужно найти q_2 , которое покрывает как можно больше остальных эпитопов, не покрытых q_1 .

Итерационное уточнение последовательных решений.

Даны первичные решения $q_1, q_2, \dots, q_M$, можно заново рассчитать новую оценку q_1 . Это осуществляют, начиная с исходных значений частоты для каждого из эпитопов, но устанавливают на ноль значения таких эпитопов, которые покрываются $q_2, \dots, q_M$. Оптимизация данной задачи комплементарного покрытия позволяет получить новый q_1 . Можно заново проходить через все начальные решения данным способом, каждый раз оптимизируя комплементарное покрытие.

Подсчет баллов с занижением на единицу.

При уже показанном анализе доверяют покрытию, только если эпитоп в последовательности s точно совпадает с эпитопом в последовательности q . Но, особенно для более длинных эпитопов, например, при $k=12$, эпитоп в последовательности все еще может быть эффективным, если это приблизительное совпадение. Например, совпадение в 11 из 12 символов может представлять собой удовлетворительное покрытие.

Результаты.

Данные алгоритмы сравнивали, применяя набор данных из 690 последовательностей.

На фиг. 4 проиллюстрировано, как выглядит диаграмма для такого большого набора данных. На фиг. 4 показана диаграмма для набора данных из 690 последовательностей белков Gag клады В из США (последовательности SEQ ID NO: 1-690), каждая из которых выровнена до положения 556. Результаты показаны в табл. 2. Горизонтальная ось представляет собой положение t , и колонки узлов указывают на различные эпитопы в каждом положении. Узлы расположены таким образом, что наиболее частые находятся внизу. Полная диаграмма показана в (а); увеличенная часть той же диаграммы показана в (b).

В табл. 2 показано сравнение баллов покрытия; это фракция эпитопов ($k=12$) в последовательностях S , которые покрыты (при точном совпадении) эпитопами в последовательностях $Q=\{q_1, \dots, q_M\}$ вакцины.

Таблица 2

Алгоритм	Эписенсус ($M = 1$)	$M = 2$	$M = 3$	$M = 6$
Консенсус	0,5070	-	-	-
Генетический алгоритм	0,5072	0,6098	0,6643	0,7308

Для мировых решений EpiGraph определили отдельные наилучшие последовательности EpiGraph на основе последовательностей из базы данных для группы М, клады В и клады С (включая примерно до 2015 г.), а также комплементарную 2^{10} последовательность EpiGraph для бивалентной вакцины для Gag, Pol и Nef. Новое преимущество кода EpiGraph над мозаичной разработкой состоит в том, что он обеспечивает возможность взвешенного исключения редких эпитопов, это свойство включено в разработку новых последовательностей, см. на фигуре ниже пример вклада исключения редких вариантов из группы М. На фиг. 5А-С проиллюстрированы конечные значения для EpiGraph и индивидуализированных вакцин, ко-

торые были использованы.

Пример 2. Графическая модель для оптимального покрытия эпитопов невыровненных последовательностей.

Набор $S = \{s_1, s_2, \dots, s_N\}$ из N невыровненных белковых последовательностей брали, чтобы он символизировал вариабельность вируса по всей целевой популяции (например, филогенетической кладе, стране или миру). Потенциальный эпитоп представляет собой подпоследовательность из k аминокислот (обычно $k=9$). Каждому потенциальному эпитопу e присваивается целочисленная частота $f(e)$, соответствующая количеству последовательностей в S , в которых встречается e . Моновалентная задача состоит в разработке единой искусственной последовательности q , которая похожа на природный белок, но оптимально покрывает указанные потенциальные эпитопы в популяции S . Если записать в виде $E(q)$ набор эпитопов, которые встречаются в q , нашей мерой покрытия является:

$$\text{Покрывание}(q) = \sum_{e \in E(q)} f(e) / \sum_{e \in E(s)} f(e)$$

Числитель представляет собой сумму частот эпитопов, которые встречаются в q , и знаменатель нормирует с помощью суммы по всем эпитопам, которые встречаются в любой из последовательностей в S . Данную формулу можно представить в виде ориентированной диаграммы. Каждый узел в данной диаграмме соответствует отличному эпитопу e , и направленное ребро соединяет два эпитопа длины k ($e_a; e_b$), если эти эпитопы перекрываются на $k-1$ символ. Путь через диаграмму представляет собой последовательность узлов $e_1, e_2, \dots, e_L$, с ребром от e_i до e_{i+1} при $i=1, \dots, L-1$. Такой путь соответствует последовательности из $L+k-1$ символов, которая представляет собой искусственный антиген q .

Если данная ориентированная диаграмма не содержит циклы, то EpiGraph находит путь через диаграмму, который строго максимизирует покрытие, позволяя получить оптимальное решение. Более того, данную оптимизацию осуществляют с вычислительными затратами, которые масштабируются лишь линейно при изменении размера (который измеряют в узлах и ребрах) диаграммы. На практике ориентированная диаграмма, созданная по S , может не быть ациклической, хотя часто она очень близка к этому, особенно для больших значений k . Для данного случая диаграмму "дециклизируют" путем итерационной идентификации циклов, а затем удаления ребер с низкими значениями до тех пор, пока не останется циклов.

Поливалентный "коктейль" из $m > 1$ антигенов можно создать путем последовательного запуска EpiGraph и оптимизации комплементарного покрытия эпитопов. Этого добиваются путем рассмотрения эпитопов e , которые были включены в первые антигены, как если бы их частоты были равны нулю, а затем запуска EpiGraph с данными модифицированными частотами. Если любой из указанных эпитопов среди первых антигенов необходим для завершения пути (и создания полноразмерного белка), то они будут все еще доступны, но они будут невыгодны, поскольку они более не будут вносить вклад в балл покрытия. Данное последовательное решение можно улучшить, применяя итерационное уточнение.

Пример 3. Индивидуализированные терапевтические вакцины.

Тогда как невозможно создать индивидуально разработанную вакцину для каждого субъекта, возможно секвенировать вирус из данного субъекта, чтобы попытаться получить хорошее совпадение внутри малого эталонного набора вариантов вакцин. Первой из рассмотренных была популяция из клинического испытания на основе клады В из США, сфокусированного на белке Gag. Были разработаны набор эталонных вакцин из Южной Африки и набор мировых вакцин, а также усовершенствованная конструкция на основе клады В из США. p24 представляет собой наиболее высоко консервативный внутренний белок в Gag, и его можно вырезать из большего белка Gag, чтобы получить консервативную область длиной ~230 аминокислот. Подход с консервативной областью также рассматривали в качестве альтернативы Gag, возможно сосредоточиваясь на областях в Gag и Pol, возможно содержащих консервативный участок Nef, а также любых других интересующих белков.

Это совсем другая задача оптимизации, отличающаяся от попытки разработать набор, который обеспечивает оптимальное покрытие популяции для профилактической вакцины. В случае профилактического применения неизвестно, с какими вирусами может встретиться вакцинируемый. В случае терапевтического применения можно получить последовательность инфицирующего вируса и выровнять с ней.

Оптимизацию осуществляли, учитывая два обстоятельства: во-первых, необходимо максимизировать совпадения из инфицирующих субъекта вирусов и, во-вторых, необходимо минимизировать несовпадения, чтобы ответ на соответствующие эпитопы вакцины был настолько направленным, насколько это возможно.

Филогенез основных клад ВИЧ демонстрирует тенденцию к менее отчетливой структуре. Скорее он "звездообразный" с очень длинными внешними ветвями и очень короткими, неявно выраженными внутренними ветвями около основания. Часть данной структуры, вероятно, возникла вследствие рекомбинации внутри подтипа. Хотя это сложно подсчитать, рекомбинация безусловно встречается относительно часто, и по аналогии с тем, что наблюдают в рамках рекомбинации внутри подтипа, она, вероятно, обширна. Учитывая структуру дерева, простое применение кластеризации на филогенетическом дереве для определения эталонного набора возможных вакцин не будет настолько эффективным, так как ассоциации внутри клады имеют ограниченное значение с "перспективы эпитопа". Напротив, необходимо учитывать взаимосвязи последовательностей с помощью подходящей меры, и эталонный набор необходимо

выбирать на основании сходства потенциальных эпитопов между природными штаммами и предполагаемыми конструкциями вакцин.

12-меры оптимизировали с учетом эпитопов класса II, но в программном коде можно использовать k-меры любой длины в качестве эталонной точки, где k представляет собой длину предполагаемого потенциального эпитопа. Также использовали 9-меры. В прошлой работе с мозаичными эпитопами оптимальное решение для 9-меров было очень близким к оптимальному для других близких длин (8, 10, 11, 12), и ожидается, что это справедливо для новых алгоритмов, описанных в данном описании. 9-меры используются для некоторых наборов индивидуализированных вакцин, описанных в данном описании.

Оптимальность определили с точки зрения k-мерного покрытия. Ее определяют путем замены каждой последовательности на "комплект эпитопов", т.е. неупорядоченный перечень из всех k-меров, которые встречаются в данной последовательности. Для набора последовательностей можно определить комплект эпитопов путем составления перечня из всех k-меров, которые встречаются в любой из последовательностей в данном наборе. Покрытие данной последовательности S набором последовательностей $Q = \{q_1, q_2, \dots, q_N\}$ осуществляется фракцией эпитопов в комплекте S, которые также находятся в совокупном комплекте последовательностей Q, где Q может представлять собой поливалентную комбинацию в коктейле вакцин. Это покрытие, которое было оптимизировано. Другое интересное количество представляло собой фракцию эпитопов в комплекте Q, которые не находятся в комплекте S. Хотя (на сегодняшний день) это напрямую не используют при оптимизации, меньшая фракция данных посторонних эпитопов предпочтительна, и данные количества рассчитывают для сравнений и дизайна эксперимента.

Оценивали шесть (6) путей нахождения "центральных последовательностей" при проведении кластеризации для индивидуализированной вакцины, включенной в код анализа индивидуализированных вакцин (см. ниже). Решение EpiGraph сочли наилучшим и использовали для конечного кода. Также проверили несколько стратегий кластеризации.

Ниже представлены исследованные стратегии определения центральных последовательностей кластеров на основе аминокислот.

1) Консенсусное решение. Консенсус был общим стандартом, получаемым путем последовательного соединения наиболее часто встречающихся в выравнивании аминокислот. На основе потенциального эпитопа (k-мера).

2) Эписенсусное решение. Данный подход позволяет найти решение в виде центральной последовательности внутри популяции или внутри кластера. Исследовали два алгоритма для нахождения эписенсуса. Первый представлял собой алгоритм заместительного размещения. Он начинается с консенсуса в качестве "субстрата". Исключение очень редких k-меров (таких, которые встречаются в популяции только 1 раз) начинается с нахождения редко встречающихся вариантов и замены консенсусных k-меров на данный редкий вариант, затем продолжают заменять на варианты, продвигаясь вверх по перечню из всех k-меров, составленному на основании их частоты встречаемости, заменяя и переписывая на более и более часто встречающиеся варианты до тех пор, пока не останутся наиболее частые варианты из всех k-меров; перекрывающиеся k-меры с большей частотой встречаемости будут по своей природе стремиться заместить пептиды, которые перекрываются с меньшими частотами встречаемости. Второй алгоритм представлял собой EpiGraph, который описан в примере 1. Последовательности EpiGraph более быстро рассчитываются, чем мозаичные из генетического алгоритма (GA), и поэтому они легко были включены в алгоритм кластеризации, необходимый для разработки индивидуализированной вакцины. В эксперименте с набором из 690 выровненных последовательностей из 556 аминокислот, которые включают белок Gag (последовательности SEQ ID NO: 1-690), решение EpiGraph было очень близким к консенсусному, а также было очень близким к отдельному наилучшему мозаичному решению, и содержало отличие лишь по одной аминокислоте от других двух центроидов.

3) Последовательное решение. Данный подход позволяет найти решение для набора из N вариантов последовательностей вакцины. В данном подходе сначала определяли для данной популяции эписенсус, и он представлял собой q_1 . Потом все k-меры, которые уже покрывал q_1 , исключали, и подход искал решение второй последовательности в данной серии с помощью того же процесса заместительного размещения с получением q_2 . затем все k-меры, которые уже покрывали q_1 и q_2 , исключали, и подход искал решение q_3 и так далее до тех пор, пока не был найден набор из N последовательностей, каждая из которых содержала все более редкие варианты потенциальных эпитопов.

4) Итерационное решение. Это итерационный вариант последовательного алгоритма. Он также позволяет получить набор из N вариантов последовательностей вакцины, обычно с лучшим покрытием, чем при последовательном подходе, но без предпочтительного упорядочивания, которое получают в результате последовательного алгоритма. Начиная с последовательного решения $q_1, q_2, \dots, q_N$ идентифицируют все эпитопы, которые находятся в наборе данных последовательностей, а затем указанные эпитопы, которые покрываются $q_2, \dots, q_N$, исключают. В поиске решения для эписенсуса с использованием остальных (неисключенных) эпитопов, получают новое значение для q_1 . На следующем этапе то же самое продвигают с q_2 , исключая эпитопы в $q_1, q_3, \dots, q_N$. И так далее с q_3 до q_N .

5) Мозаичное (GA-мозаичное) решение. Определяют, применяя исходный генетический алгоритм (GA). Мозаичный относится к генетическому алгоритму, названному мозаичным, и/или к последователь-

ностям антигена, полученным с помощью GA. Здесь получали решение для набора сразу из всех мозаиков размера N , или можно получить решение для наилучшего отдельного мозаика, зафиксировать его и получить решение для дополнительного набора из 5 последовательностей, чтобы получить всего 6 вариантов последовательностей вакцин. Данную стратегию использовали, чтобы позволить сравнение напрямую со стратегиями кластеризации, описанными ниже.

б) Наилучший природный штамм. В данном подходе находят в качестве решения набор из N последовательностей вариантов вакцин. Находят природный штамм в наборе, который содержит наиболее "центричные эпитопы", т.е. он содержит наиболее часто встречающиеся k -меры. Для того, чтобы его найти, для каждого k -мера в этом природном штамме данной конкретной форме можно присвоить частоту на основании ее частоты в популяции, а затем суммировать полученные частоты, получив меру того, насколько хорошо указанный природный штамм покрывает данную популяцию. Природный штамм с наиболее высоким баллом представляет собой наилучший отдельный штамм Nat.1. Затем k -меры, которые покрываются Nat.1, можно убрать из схемы оценки, и можно выбрать наилучшее дополнение к Nat.2 путем нахождения среди природных штаммов наилучшего покрытия эпитопов, исключая такие эпитопы, которые уже покрываются Nat.1. Это осуществляют итерационно, то есть выбирают k природных штаммов, где k представляет собой количество вариантов вакцин, которые хотят рассмотреть, и их упорядочивают таким образом, что Nat.1 представляет собой наилучший отдельный штамм, Nat.2 представляет собой наилучшее дополнение к Nat.1, Nat.3 представляет собой наилучшее дополнение к (Nat.1+Nat.2).

Исследовали варианты вакцин, которые были разработаны в соответствии со следующими основными стратегиями, сравнивая новые идеи, чтобы конкретно рассмотреть последовательности популяции терапевтических вакцин, которые были разработаны в прошлом, для оптимизации покрытия популяции.

Если все субъекты должны были получить одну и ту же вакцину (это может привести к лучшему результату только с консервативными областями), то секвенирование и индивидуализацию не проводили, и это были универсальные разработки, не оптимизированные для каждого индивида.

Консенсус. Находят одну универсальную последовательность, которая наилучшим образом покрывает популяцию. Потенциально подходящая одна универсальная последовательность включает: консенсусную последовательность для популяции, наилучшее отдельное мозаичное решение GA, эпиконсенсус и природный штамм с наиболее "центричными эпитопами".

EpiGraph. Находят либо 2, либо 3 последовательности для данной популяции и дают их всем в этой популяции. Такую стратегию с учетом популяции сравнивали с применением мозаичного решения GA, наилучших природных штаммов и последовательного и итерационного мозаичного решения. Поскольку теперь доступен EpiGraph и улучшение мозаичных решений, так как мы можем также исключить редкие эпитопы, мы применяем их.

Напротив, для индивидуализированной вакцины каждый индивид получит только последовательности вакцины, которые наилучшим образом совпадают с его или ее инфекцией. Производят, например, 6 последовательностей антигенов для вакцины и выбирают наилучшую одну или наилучшую комбинацию 2 или 3 из этих 6 для доставки пациенту; т.е., выбирают такие последовательности, которые дают наилучшее покрытие инфицировавшего пациента штамма, с как можно меньшим количеством несовпадающих эпитопов. Для данного плана исследовали несколько стратегий.

а) Кластерные последовательности со стратегией, подобной алгоритму k -средних, для создания 6 кластеров, осуществляли 1500 итераций (каждая из данных итераций представляла собой этап разделения-слияния данного испытания, а затем осуществляли несколько обычных итераций) для получения конечных наборов, определяли центроидные последовательности из данных 6 кластеров для наборов вакцин. Расстояние между двумя последовательностями определяли как один минус покрытие эпитопов одной последовательности другой последовательностью. Изначально для рассредоточения кластеров использовали 6 произвольно выбранных природных штаммов. Это позволило получить набор из 6 очень высоко сходных центроидных последовательностей. Определили, что если было представлено большее природное разнообразие, то это позволяло получить лучший набор реагентов для создания индивидуализированных вакцин. Затем указанные 6 кластеров рассредотачивали по дополнительным 6 природным (Nat6) штаммам, так как они очень отличаются, а затем заново присваивали центр на основе данных кластеров как эпиконсенсуса, и итерационно заново проводили кластеризацию и заново центрировали. Данную стратегию сравнивали со стратегией, в которой начинали с 6 наилучших природных штаммов и использовали консенсус в качестве центроида кластера вместо эпиконсенсуса, и эпиконсенсусная стратегия позволила получить несколько лучшее покрытие. Принудительное назначение минимального размера кластера также позволило получить несколько лучшее покрытие (проверили минимальные размеры кластера 1, 5 и 20), поэтому минимальный размер кластера включили в качестве ограничения. Минимальный размер, равный 20, был лучше, чем 1 или 5. Для того, чтобы внедрить минимальный размер, если в данном цикле количество последовательностей в размере кластера оказывается ниже минимального размера, то каждый представитель данного "слишком маленького кластера" перераспределяется в наилучший центроид из других пяти кластеров. Для получения нового кластера, чтобы заменить один утраченный, кластер, в котором самые большие средние расстояния до центроида, разбивают, взяв две произвольные природные последовательности из данного кластера в качестве центроидов и перестроив два кластера вокруг

них, заново вычисляя центроиды и переходя к следующему этапу с данными новыми шестью кластерами. Указанные центроиды из данных кластеров были очень близки к центру дерева. По данным последовательностям определили, что причиной этого была повторная репликация консенсуса внутри кластеров, что приводило к доминированию сигнала.

б) Эписенсус + (5 центров кластера). В данном подходе сначала определили центральную последовательность для целой популяции, применяя алгоритм EpiGraph (эписенсус для популяции), фиксированную для включения в качестве реагента вакцины. Любые эпипептиды, которые совпадали с эписенсусом для популяции, исключали из рассмотрения для кластеризации. Последовательности кластеризовали, как и прежде, с минимальным размером, но на этот раз получали 5 кластеров на основе всех потенциальных эпипептидов, за исключением эпипептидов, которые находились в эписенсусе, таким образом, целевыми были такие кластеры, которые дополняли эписенсус для популяции. Определили, что это было наилучшее решение для Gag. При включении эписенсуса для популяции в индивидуализированную вакцину для каждого субъекта, даже если данный наиболее часто встречающийся k-мер не виден в полученных из них образцах последовательностей ВИЧ, он может скрываться или может представлять собой распространенную форму для реверсии, учитывая переключение частых аминокислот ВИЧ между распространенными формами. Вторая дополнительная последовательность из одного из 5 кластеров затем приумножит способность указанного варианта вызывать перекрестные реакции между вакциной и инфицирующим штаммом.

с) Фиксировали наилучший EpiGraph, добавляли 5 мозаичных дополнительных последовательностей с получением набора из шести последовательностей (5 добавленных не упорядочены), или фиксировали наилучшую природную последовательность, последовательно добавляли 5 природных последовательностей.

После обширных сравнений и уточнений мы отдали предпочтение применению алгоритма EpiGraph для определения центра кластеров для индивидуализированных вакцин. Предложен набор из 6 антигенов белка Gag для производства вакцин в рамках индивидуализированной разработки, нацеленной либо на вакцину мировой группы M, либо на вакцину современной клады B, либо на вакцину современной клады C для Южной Африки. Сравнение покрытия природных последовательностей современной клады B с использованием 2 природных последовательностей клады B, вариантов EpiGraph для группы M, вариантов EpiGraph для клады B или индивидуализированных последовательностей для клады B показаны на фиг. 6A - B; индивидуально разработанные последовательности из клады B позволяли получить наилучшее покрытие эпипептидов.

Можно резюмировать различные варианты с некоторыми числами, что видно в табл. 3, и они обозначены M, C, T, A и P. Это описано ниже с точки зрения количества "пилюль" (т.е. антигенов для вакцины) для каждой категории.

M = Произведенные - это суммарное количество произведенных пилюль, из которых для каждого индивида выбирают некоторое подмножество (как правило, мы предполагаем, что M = 6 или меньше).

C = Общие - это такие пилюли, которые получают все, возможно дополнительно к некоторым индивидуализированным пилюлям.

T/A = Индивидуализированные/альтернативные варианты - T представляет собой количество индивидуализированных пилюль (из A альтернативных вариантов), которые дают каждому индивиду, возможно, в комбинации с некоторыми обычными пилюлями.

P = Всего на субъекта (T+C) - количество пилюль, которые дают каждому индивиду.

Таблица 3

<u>Центроидный</u>	<u>P</u>	<u>C</u>	<u>T/A</u>	<u>M</u>	<u>Покрытие</u>	<u>Дополнит.</u>	<u>Дельта</u>
					<u>>хорошо</u>	<u>>плохо</u>	<u>> хорошо</u>

Консенсусн.	1	1	0/0	1	0,5070	0,4913	0,0157
1-GA мозаичн.	1	1	0/0	1	0,5072	0,4911	0,0161
Эписенсусн.	1	1	0/0	1	0,5065	0,4918	0,0147
Природн.	1	1	0/0	1	0,4793	0,5191	-0,0398
2-GA мозаичн.	2	2	0/0	2	0,5960	0,6692	-0,0732
2-итер. мозаичн.	2	2	0/0	2	0,5965	0,6346	-0,0381
2-природн.	2	2	0/0	2	0,5473	0,6584	-0,1110
3-GA мозаичн.	3	3	0/0	3	<u>0,6346</u>	<u>0,7424</u>	<u>-0,1078</u>
3-итер. мозаичн.	3	3	0/0	3	<u>0,6534</u>	<u>0,7190</u>	<u>-0,0656</u>
3-посл. мозаичн.	3	3	0/0	3	<u>0,6429</u>	<u>0,7119</u>	-0,0690
3-природн.	3	3	0/0	3	0,5941	<u>0,7308</u>	-0,1367
6-итер. мозаичн.	6	6	0/0	6	<u>0,7205</u>	<u>0,8324</u>	<u>-0,1119</u>
1+5-GA мозаичн.	6	6	0/0	6	<u>0,7048</u>	<u>0,8241</u>	<u>-0,1193</u>
6-природн.	6	6	0/0	6	<u>0,6628</u>	<u>0,8104</u>	<u>-0,1476</u>

Общая: все получают одинаковую вакцину, разработанную с учетом популяции

В табл. 3 однократно подчеркнутые числа укзывают на то, что в среднем >60% указанных 12-меров в природных последовательностях покрываются вакциной (Хорошо), тогда как дважды подчеркнутые числа укзывают на то, что в среднем более 70% указанных 12-меров не присутствует в природных штаммах.

Индивидуализированный - выбирали наилучший 1 из 6 для каждой из 690 популяционных последовательностей. Не сильно лучше (0,54 по сравнению с 0,51), чем просто получить один для всей популяции, но С6-эпи была наилучшей в классе, если собирались делать 6 вакцин и давать одну из данных шести субъекту на основании его последовательности.

С6-эпи	1	0	1/6	6	<u>0,5366</u>	<u>0,4616</u>	0,0750
1+5-эпи-С5	1	0	1/6	6	0,5079	0,4904	0,0175
1+5-GA мозаичн.	1	0	1/6	6	0,5083	0,4900	0,0183
1+5-итер. мозаичн.	1	0	1/6	6	0,5074	0,4909	0,0165
6-итер. мозаичн.	1	0	1/6	6	0,3979	0,6008	-0,2029
6-природн.	1	0	1/6	6	0,4863	0,5122	-0,0259
С6-природнЗатрав	1	0	1/6	6	0,5291	0,4692	0,0599

Индивидуализированная наилучшая пара в группе из шести

С6-эпи	2	0	2/6	6	0,5785	0,5313	0,0472
1+5-эпи-С5	2	0	2/6	6	<u>0,6255</u>	<u>0,6310</u>	-0,0055 **
6-итер. мозаичн.	2	0	2/6	6	0,5552	0,6859	-0,1307
6-посл. мозаичн.	2	0	2/6	6	<u>0,6075</u>	0,6331	-0,0256
1+5-GA мозаичн.	2	0	2/6	6	0,5932	0,6401	-0,0469
6-природн.	2	0	2/6	6	0,5759	0,6431	-0,0672

Гибридная индивидуализированная пара. Фиксировали центр популяции, добавляли 1 из 5 центроидов, чтобы наилучшим образом дополнить фиксированную последовательность, чтобы покрыть каждую из тестируемых последовательностей.

1+5-эпи-С5	2	1	1/5	6	<u>0,6255</u>	<u>0,6310</u>	-0,0055 **
1+5-GA мозаичн.	2	1	1/5	6	0,5922	0,6394	-0,0472
1+5-природн.	2	1	1/5	6	0,5724	0,6395	-0,0671
6-посл. мозаичн.	2	1	1/5	6	0,6071	0,6330	-0,0259
6-итер. мозаичн.	2	1	1/5	6	0,5280	<u>0,7039</u>	-0,1759

Гибридные наилучшие три. Фиксировали центр популяции и добавляли 2 из 5 центроидов, чтобы наилучшим образом дополнить фиксированную последовательность, чтобы покрыть тестируемую последовательность.

1+5-эпи-С5	3	1	2/5	6	<u>0,6590</u>	<u>0,6868</u>	-0,0278
НВР-6-итер. мозаичн.	3	1	2/5	6	0,6586	<u>0,7079</u>	-0,0493
НВР-1+5-GA мозаичн	3	1	2/5	6	0,6447	<u>0,7155</u>	-0,0708

Общая: все получают одну и ту же вакцину, разработанную с учетом популяции, доставляют либо 1, либо 2 антигена для вакцины.

Общая	Природн.	1	1	0,4793	0,5191
Общая	Мозаичн.	1	1	0,5072	0,4911
Общая	2-природн.	2	2	0,5473	0,6584
Общая	2-GA мозаичн.	2	2	0,5960	0,6692
Общая	2-итер. мозаичн.	2	2	0,5965	0,6346
Индивидуализир	С6-эпи	2	6	0,5785	0,5313
Индивидуализир	1+5ЭпиС	2	6	0,6341	0,6361
Индивидуализир	1+5ЭпиС5	3	6	0,6590	0,6868
Общая	3-GA мозаичн.	3	3	0,6346	0,7424

Точное совпадение 1+5-эпи-С5.

	Оптимизировали для 9		Оптимизировали для 12	
	ПОКРЫТИЕ ДОПОЛНИТ.		ПОКРЫТИЕ ДОПОЛНИТ.	
Оценивали для 9	0,7098	0,5744	0,7066	0,5535
Оценивали для 12	0,6301	0,6525	0,6341	0,6361

И занижение на единицу считали совпадением 1+5-эпи-С5.

	Оптимизировали для 9		Оптимизировали для 12	
	ПОКРЫТИЕ ДОПОЛНИТ.		ПОКРЫТИЕ ДОПОЛНИТ.	
Оценивали для 9	0,9207	0,1783	0,9183	0,1723
Оценивали для 12	0,8722	0,2590	0,8708	0,2474

** Это, вероятно, наилучшее решение. Получали эписенсус для популяции, 12-меры, обнаруженные в эписенсусе, из кластеризации были исключены. Когда эписенсус для популяции фиксировали, и наилучший эписенсус из других 5 дополнительных кластеров выбирали для получения пары с ним, то получали точно такой же ответ, как и при выборе наилучшей пары из этих 6 вариантов. Это означает, что эписенсус для популяции всегда составлял одну из наилучших пар.

Занижение на единицу. Для приведенных выше оценок рассматривали только идеальные совпадения, для того, чтобы эпитоп совпадал между коктейлем вакцин и природным штаммом, требовалось идеальное совпадение 12/12. Несовпадения часто вполне допустимы, особенно для эпитопов II класса, если для совпадения требуется совпадение 11/12, а для несовпадения требуется 10/12 или менее, то все выглядит более оптимистично. Истина, вероятно, где-то посередине, и 10/12 также может быть приемлемо в некоторых случаях. Здесь вероятно наилучший вариант сравнивают со сопоставимым наилучшим вариантом природного штамма с идеальными совпадениями: Идеальное совпадение (взяли из таблицы выше):

1+5-эпи-С5	2	1	1/5	6	0,6255	0,6310	-0,0055 **
1+5-природн.	2	1	1/5	6	0,5724	0,6395	-0,0671

И занижение на единицу считали совпадением:

1+5-эпи-C5	2	1	1/5	6	0,8700	0,2524	0,6176
1+5-природн.	2	1	1/5	6	0,8412	0,3112	0,5300

Три пиллюли: одна общая (одинаковая для всех) и две индивидуализированные пиллюли - наилучшая пара из остальных пяти.

Пример 4. Разработка и оптимизация "индивидуализированного" коктейля антигенов клады В ВИЧ-1.

Разнообразие ВИЧ на уровне популяции начинается с быстрой эволюции внутри каждого инфицированного хозяина. Большая часть разнообразия внутри хозяина является прямым следствием ускользания от иммунного ответа (Bar, K. J. 2012. PLoS Pathog 8:e1002721., Liu, M. K. 2013.7. Clin. Invest. 123:380-93), и варианты ускользания возникают вскоре после инфицирования (Fischer, W. 2010. PLoS One 5:e12303). Целью профилактических вакцин является быстрое уничтожение инфицирующего штамма или предупреждение изначального инфицирования им. Поскольку неизвестно, какой именно штамм будет передан, профилактические вакцины должны вызывать иммунные ответы, которые будут активны против любого варианта, который может встретиться. Напротив, иммунотерапия может быть направлена против известной вирусной популяции, присутствующей внутри данного хозяина. Профилактические и терапевтические вакцины против ВИЧ, следовательно, должны быть разработаны с учетом данных различных требований. На фиг. 7 показано, что мозаичные решения достигают максимально возможного покрытия эпитопов для данной популяции ВИЧ. На данной фигуре также описаны различия в консервативности белков ВИЧ: большие фрагменты Gag, Nef и Pol высоко консервативны, тогда как большинство Env, а также большинство малых вспомогательных белков (результаты не представлены) высоковариабельны. Тогда как мозаичные решения предлагают почти оптимальное решение для покрытия эпитопов в популяции для профилактической вакцины, в них не используется дополнительная информация об известной последовательности инфицирующего штамма, на которую нацелена терапевтическая вакцина.

Для реализации стратегии разработки индивидуализированных вакцин требуется производство удобного количества реагентов вакцины, которые максимально охватывают разнообразие эпитопов в целевой популяции, секвенирование ВИЧ из инфицированного реципиента вакцины, а затем выбор из пула реагентов вакцины подгруппы вставленных антигенов, которые наилучшим образом совпадают с антигенами у субъекта. Стратегия индивидуализированных вакцин, следовательно, концептуально отличается от стратегии разработки мозаичной вакцины.

98% из приблизительно 150000 последовательностей ВИЧ, которые были взяты в США, относятся к кладе В. Использовали взятые в США последовательности Gag клады В, которые ныне доступны в базе данных Los Alamos HIV. В данном наборе было 690 интактных последовательностей Gag (последовательности SEQ ID NO: 1-690), каждая была получена из отдельного инфицированного субъекта, представляющих профиль разнообразия штаммов ВИЧ в США. Для того, чтобы максимизировать покрытие Т-клеточных эпитопов в гипотетической вакцине, в то же время фиксируя количество различных последовательностей Gag, которые необходимо включить в коктейль вакцин, изобрели новую вычислительную стратегию, в которой совокупно применяют алгоритм консенсусной последовательности на основе эпитопов (или "эписенсусной") с кластеризацией подобно k-средним. В отличие от простой консенсусной последовательности, которая в каждом положении выравнивания содержит наиболее часто встречающуюся аминокислоту, в качестве эписенсуса ищут наиболее часто встречающийся эпитоп (например, 12-мер), начинающийся в каждом положении выравнивания. Тем не менее, нельзя просто взять наиболее часто встречающийся эпитоп в каждом положении, так как эпитопы перекрываются, и это может приводить к противоречиям в соседних аминокислотах. Цель состоит в том, чтобы урегулировать данные противоречия таким способом, который приводит к максимальному покрытию вакциной эпитопов в популяции.

Для того, чтобы определить малый набор антигенов для производства, который представляет разнообразие эпитопов в популяции, использовали подход, подобный k-средним, чтобы разбить указанные 690 последовательностей Gag на отдельные кластеры на основании подобия эпитопов, и получали отдельные эписенсусные последовательности для каждого кластера, чтобы они служили центральными последовательностями кластеров. Покрытие индивидуальных вариантов приближается к пределу после 5-6 кластеров, и производство 6 векторов вакцины с 6 вариантами антигенов ВИЧ вполне осуществимо, поэтому изначально планировали получить набор из 6 антигенов. Алгоритм начинается с 6 произвольно выбранных природных штаммов в качестве исходных центроидных последовательностей, и присваивает каждую из остальных последовательностей к наиболее близкой из данных 6 центроидных последовательностей. Внутри каждого из 6 кластеров вычисляют новую центроидную последовательность на основании эписенсуса для последовательностей в данном кластере. Затем все последовательности перераспределяют в соответствии с тем, какой из новых центроидов наиболее близок. Данный процесс сходится к подобной k-средним оптимизации для разработки 6 антигенов для вакцины, которые берут в качестве эписенсусных центроидов отдельных кластеров, и данные антигены можно применять в качестве промышленного набора, из которого выбирают для данной вакцины антигены, обеспечивающие наилучшее совпадение.

Обнаружили, что гибридный подход существенно улучшает покрытие эпитопов, при этом миними-

зируя усиление потенциальных специфических ответов на вакцину. В стратегии гибридизации центральную последовательность вычисляли для всей популяции (эписенсус для популяции), и последующая кластеризация была основана исключительно на эпитопах, которые не были найдены в центральной последовательности. Таким образом, эписенсусные последовательности кластеров дополняют эписенсус для популяции. Следовательно, набор объединенных реагентов из 6 последовательностей состоит из одного эписенсуса для популяции и пяти дополнительных центроидных последовательностей кластеров, и субъектам будут давать двухантигенную индивидуализированную вакцину: эписенсус для популяции в паре с наилучшей дополнительной последовательностью. Данная стратегия позволяла включить более разнообразные последовательности в исходный набор и улучшала потенциальное покрытие природных последовательностей Gag в исследуемой популяции, что видно из табл. 4.

Таблица 4. Среднее покрытие для каждого штамма потенциальных эпитопов Gag из клады В из США различными вакцинами

Вакцина	Доставленная	Произведенная	% совпаде ния с природны м Gag	% несовпаде ния с природным Gag	Фактор улучшен ия
1 природная	1	1	47,9	51,9	
1 мозаичная	1	1	50,7	49,1	1,06
1 эписенсусная	1	1	50,7	49,1	1,06
2 природных	2	2	54,7	65,8	1,14
2 мозаичных	2	2	59,6	66,9	1,24
2 индивидуализированных	2	6	63,4	63,1	1,32
3 природных	3	3	59,7	73,1	1,35
3 мозаичных	3	3	63,4	74,3	1,32
3 индивидуализированных	3	6	65,9	68,7	1,37
6 природных	6	6	66,3	81,0	1,38
6 мозаичных	6	6	72,1	83,2	1,51

Для каждого из 690 вариантов получали индивидуализированную вакцину с 2 или 3 антигенами, выбранными из эписенсуса для популяции и полученных из 5 кластеров дополнительных последовательностей. Это сравнивали с покрытием потенциальных эпитопов при стратегиях разработки вакцины с учетом популяции, при которых каждому вакцинируемому будут давать одну и ту же вакцину, либо мозаичную, либо наилучшие природные комбинации. "Доставленная" указывает на количество антигенов, которое включено в коктейль вакцин. "Произведенная" указывает на размер исходного объединенного набора антигенов для вакцины, который необходимо будет синтезировать, чтобы выбрать из него антигены для применения в индивидуализированном подходе; "% совпадения" указывает на среднее количество потенциальных эпитопов, идеально совпадающих между каждым из вариантов природного Gag и вакциной; чем выше данное значение, тем больше вероятность того, что ответы на вакцину будут приводить к перекрестным реакциям с природными вариантами; "% несовпадения" представляет собой фракцию потенциальных эпитопов в вакцине, которые не были обнаружены в данной природной последовательности Gag, рассчитанных для каждой из указанных 690 последовательностей отдельно, а затем усредненных. Чем выше данное значение, тем больше способность вакцины вызывать специфические ответы на вакцину, которые могут отвлекать от перекрестных ответов. Фактор улучшения указывает на увеличение покрытия при применении предложенных новых вариантов разработки вакцин по сравнению с применением одного наилучшего природного штамма.

Решения для индивидуализированных вакцин, оптимизированные для 12-меров, приблизительно оптимальны для 9-меров и наоборот, следовательно, кластеры на основе 12-меров должны хорошо работать для представления эпитопов как класса I, так и класса II, что видно из таблицы 5.

Таблица 5. Допущение одного несовпадения аминокислоты при оценке 9-мера или 12-мера, которое нужно рассматривать как положительное совпадение при вычислении среднего покрытия при индивидуализированной разработке вакцины из 2 белков

	Оптимизированный для 9-мера		Оптимизированный для 12-мера	
	% совпадения	% несовпадения	% совпадения	% несовпадения
Оценивали для 9-мера	92,07	17,83	91,83	17,23
Оценивали для 12-мера	87,22	25,90	87,08	24,74

Наконец, подход индивидуализированных вакцин теоретически очень хорошо работает, если в потенциальных эпитопах допускается 1 несовпадение из 12 вместо требования идентичности (табл. 5); поскольку это более мягкий и, возможно, биологически более реалистичный критерий, то приблизительно 90% ответов на индивидуализированные вакцины могут быть перекрестными с эпитопами в соответствующих природных белках Gag. Подход индивидуализированных вакцин оказался лучше, чем подход с мозаичными для популяции последовательностями, подход с консенсусными последовательностями и подход с наилучшими природными штаммами, с точки зрения как максимизирования покрытия эпитопов последовательностей Gag, так и минимизирования потенциально вредных специфических для вакцины эпитопов (Фиг. 8). Данный программный код также можно применять для индивидуализации вакцин с использованием множества последовательностей вместо использования по одной типичной из каждого инфицированного индивида, и применения стратегии разработки индивидуализированных вакцин к различным популяциям (кладе С, эпидемической в Южной Африке, кладе 2, регионально эпидемической в Таиланде, и группе М в мировой популяции).

Индивидуализированные антигены для вакцины могут предоставить лучшее покрытие (по сравнению с антигенами, разработанными на основе популяции) природных последовательностей, когда известен инфицирующий штамм.

Пример 5. Двойные векторы экспрессии.

Используя эписенсусные антигены для популяции и/или индивидуализированные антигены, описанные в данном описании, получают двойные векторы экспрессии, каждый из которых экспрессирует полный антиген Gag и второй антиген ВИЧ. Вторым антигеном ВИЧ может представлять собой, например, слитый белок обратной транскриптазы (RT) и центральной части Nef. Интегразу не включают, так как она довольно слабо стимулирует ответы Т-клеток. Применяя векторы CMV (например, векторы RhCMV или HCMV) в виде двойных векторов экспрессии, можно одновременно индуцировать Т-клетки против двух различных антигенов ВНО или ВИЧ.

Например, получали панели из шести или менее векторов HCMV, содержащих индивидуализированные последовательности Gag, разработанные на основе алгоритма EpiGraph (один вектор, экспрессирующий эписенсусный для популяции антиген Gag, и пять векторов, каждый из которых экспрессирует дополнительный разработанный на основе кластеров антиген Gag), а также разработали панели антигенов, покрывающих RT и центральный участок Nef. Предложены один вектор, экспрессирующий эписенсусный для популяции антиген Gag, плюс два вектора, каждый из которых экспрессирует отличный антиген Gag, выбранный из указанных пяти дополнительных разработанных на основе кластеров антигенов. Данные векторы также могут содержать одну из двух дополнительных последовательностей ВИЧ-1 EpiGraph RT/nef. Если прогнозируют, что индивидуализация не улучшит покрытие вследствие высокой консервативности данных последовательностей, то оставляют 2-мозаичное решение и используют его для получения вектора. Например, один мозаичный RT/nef включают в эписенсусный вектор для популяции, а другой используют в индивидуализированных векторах. Панель векторов вакцин на основе HCMV, которые можно включить в производство вакцин, получают путем секвенирования полученных векторов и определения экспрессии с них антигена и роста *in vitro*.

Получали синтетические кодон-оптимизированные вставки ДНК, соответствующие мозаичным Gag, RT и Nef и индивидуализированным антигенам, разработанным, как описано в примере 3.

Пример 6. Временная экспрессия вирусных антигенов, разработанных с применением подхода EpiGraph.

Антигены, разработанные с применением алгоритма EpiGraph таким образом, чтобы максимизировать частоту эпитопов, похожи на природные последовательности, но более не кодируют нативные белки. Хотя при конструировании данных последовательностей придерживались теоретических рекомендаций для экспрессии данных искусственных последовательностей, белки, кодируемые ими, могут проявлять непредвиденные профили экспрессии или не могут экспрессировать стабильный полноразмерный белок.

Для того чтобы оценить профиль экспрессии данных последовательностей в контексте клеток мле-

копитающих, синтезировали и клонировали последовательности EpiGraph для временной трансфекции. ДНК, кодирующую данные конструкции (Genscript, Пискатауэй, Нью-Джерси) синтезировали так, чтобы она содержала совместимые сайты клонирования для плазмидных векторов (pcDNA3.1 и pOri). Все вставки оптимизировали по кодонам для экспрессии в соответствующем хозяине (макак-резус для ВАО или человек для ВИЧ). Каждую конструкцию также модифицировали, чтобы устранить остаточную активность фермента нативной последовательности, что описано у Kulkarni и др. Vaccine (2011). Удаленные положения указаны на основе последовательности аминокислот, относящейся к EpiGraph-1 клады В. Удаленные аминокислоты включали: "DTG", связанные с активностью протеазы (положения 81-83), "YMDD", связанные с активностью обратной транскриптазы (положения 338-341), "E", связанная с активностью РНКазы H (положение 633), "D", связанная с активностью интегразы (положение 779), "D", связанная с активностью интегразы (положение 831), и "E", связанная с активностью интегразы (положение 867). Синтетическую ДНК разводили в воде и расщепляли эндонуклеазами рестрикции (5'-NheI, 3'-BamHI), а затем инактивировали нагреванием. Плазмидный вектор линейаризовали с помощью подходящих эндонуклеаз и обрабатывали фосфатазой кишечника телят, чтобы предотвратить рециркуляризацию пустого вектора. Фрагменты вектора и вставки разделяли с помощью электрофореза в агарозном геле, чтобы подтвердить размеры расщепленных фрагментов, и очищали для лигирования с помощью набора для очистки продуктов ПЦР (Thermo Scientific). Вставки лигировали с линейаризованным вектором при соотношении вставки к вектору, приблизительно равном 3:1, в течение 15 мин при комнатной температуре, применяя набор для быстрого лигирования (Roche, Индианаполис, Индиана), трансформировали полученными векторами химически компетентные *E. coli* (DH5-альфа), и высевали на чашки для селекции антибиотиком. ДНК из полученных колоний подвергали скринингу на наличие вставок путем расщепления рестриктазами.

Клоны, содержащие каждую из правильных вставок в подходящей ориентации по отношению к промотору и поли(А)-последовательности вектора, растили в суспензионной культуре для очистки плазмидной ДНК. Активно растущие субконфлюэнтные клетки HeLa в 12-луночных культуральных планшетах получали 500 мкл свежей среды (DMEM с 10% FBS) в то время, пока готовили липосомы. Для того, чтобы получить липосомы, содержащие плазмидную ДНК, 250 мкл бессывороточной среды смешивали с 500 нг плазмидной ДНК и 250 мкл бессывороточной среды смешивали с 2 мкл липида (Lipofectamine 2000, Invitrogen). После 5 мин инкубации при комнатной температуре эти растворы объединяли, перемешивали и инкубировали в течение 20 мин. Содержащие ДНК липосомы (500 мкл), сформировавшиеся в данном процессе, добавляли по каплям в культуру и оставляли инкубироваться в течение 12-16 ч, по прошествии данного промежутка времени трансфицирующую смесь заменяли на свежую среду. После дополнительного дня инкубации культуры собирали путем соскабливания и центрифугирования. Супернатанты удаляли путем аспирации и осадки клеток лизировали путем ресуспендирования в 100 мкл загрузочного буфера, содержащего краситель, 5% ДСН и 10% 2-меркаптоэтанол, и центрифугирования на колонке QiaShred (Qiagen, Валенсия, Калифорния).

Экспрессию белков EpiGraph продемонстрировали с помощью электрофореза в полиакриламидном геле с добавлением ДСН (ПААГ/ДСН) и вестерн-блоттинга с антителами к эпителиальной метке V5 или гемоглобулина (HA), вставленной в каждую конструкцию. Вкратце, получали 10% полиакриламидные гели, загружали на них по 10 мкл (10%) каждого образца и подвергали электрофорезу при 110-120 Вольт в течение 90 минут. Разделенные белки переносили на поливинилиденфторидные (PVDF) мембраны путем полусухого блоттинга при 20 В в течение 45-50 мин. Неспецифическое связывание блокировали с помощью раствора 10% обезжиренного сухого молока в фосфатно-солевом буфере с добавлением 0,1% tween-20 (ФБР-Т) в течение 60 мин. Антитела к HA (Sigma) или V5 (Santa Cruz) разбавляли в 5% растворе молока и инкубировали с мембранами в течение 1 ч, а затем трижды промывали ФБР-Т перед добавлением разведенного 1:2000 конъюгированного с пероксидазой хрена вторичного антитела козы против антител мыши (Santa Cruz) в течение 1 ч. Затем блоты промывали три раза в ФБР-Т, проявляли, используя набор с субстратом для усиленной хемилюминесценции (ECL) (Thermo-Pierce), и визуализировали с помощью рентгеновской пленки. Для всех исследованных конструкций продемонстрировали надежную временную экспрессию белков с предсказанной молекулярной массой и подтвердили их пригодность для тестирования в остоке вектора CMV (например, см. фиг. 9 А-С).

Пример 7. Встраивание разработанных с помощью EpiGraph антигенов в ВАС-конструкции вектора CMV и экспрессия с воспроизведенного вируса.

Антигены EpiGraph разработали таким образом, чтобы максимизировать покрытие Т-клеточных эпитопов, типичных для спектра вирусных последовательностей и клады ВИЧ, из которых они были получены. Для наиболее эффективного использования данных антигенов, их встраивали в векторы CMV, для которых продемонстрировали в три раза больший спектр CD8⁺ Т-клеток по сравнению с конкурирующими платформами. Широкое представление антигена и профили пожизненной экспрессии векторов CMV продемонстрировали способность защищать и лечить обезьян макак-резус, инфицированных ВАО. Алгоритм разработки антигенов EpiGraph в комбинации с векторами CMV может обеспечить даже большее покрытие ВИЧ внутри клады и по всем кладам, когда его применяют по отношению к профилактическим вакцинам широкого спектра действия или индивидуализированным направленным вакцинам.

Последовательности EpiGraph, экспрессию которых продемонстрировали в системах временной трансфекции, субклонировали в плазмиду для рекомбинации (pOri) и переносили в остовы CMV, применяя рекомбинационную инженерию ВАС (Messerle и др. *Proc Natl Acad Sci USA*. 1997 Dec 23;94(26):14759-63; и Borst и др. *J Virol*. 1999 Oct;73(10):8320-9).

Рекомбинационная инженерия ВАС упрощает манипулирование большими последовательностями ДНК, используя регулируемые температурой и метаболитами ферменты рекомбинации в контексте штамма EL250 *E. coli*, содержащего исходный ВАС. Рекомбинация представляет собой последовательный двухэтапный процесс, состоящий из вставки последовательности антигена с геном устойчивости к антибиотику (канамицину) в целевой участок, а затем удаления канамициновой кассеты. Фрагменты вставок амплифицируют с помощью ПЦР с матричной ДНК, содержащей интересующий антиген плюс канамицин, применяя праймеры с длинными (50+ п.о.) гомологичными плечами.

Для подготовки бактериальных клеток к этапу вставки, культуры объемом пять миллилитров растили в течение ночи при 30°C в бульоне Лурия (LB) с добавлением хлорамфеникола и разбавляли до 50 мл на следующее утро. Бактерии растили в течение приблизительно 3-4 дополнительных часов при 30°C (до оптической плотности OD=0,6), а затем подвергали тепловому шоку путем встряхивания при 42°C в течение 15 мин, чтобы индуцировать ферменты рекомбинации. После такой индукции бактерии осаждали (3000 об/мин, 10 мин, 4°C), а затем промывали три раза в охлажденной до температуры льда воде. Клетки *E. coli* делали электрокомпетентными, чтобы внедрить в них продукт ПЦР, и компетентными для рекомбинации для вставки последовательности в целевой участок ВАС. Очищенную вставку (500 нг) объединяли со 100 мкл компетентных *E. coli* на льду, переносили в 0,2 см кювету (Fisher) и подвергали электропорации, применяя устройство Bio-rad MicroPulser. После электропорации бактерии разбавляли путем добавления 900 мкл культуральной среды LB и позволяли им восстановиться при 30°C в течение 2 ч перед посевом на чашки с добавлением хлорамфеникола/канамицина. Чашки инкубировали при 30°C в течение двух дней и колонии подвергали скринингу путем расщепления рестриктазами и ПЦР для выявления событий рекомбинации.

ВАС-конструкции, для которых подтвердили рекомбинацию, подвергали второму этапу, на котором канамициновую кассету вырезали путем индукции арабинозой рекомбиназы Flip, опосредованной фланкирующими сайтами FRT. Культуры объемом 5 мл растили в течение ночи в LB + хлорамфеникол и разбавляли 1:10 на следующее утро. После 3 ч роста бактерии обрабатывали L-арабинозой (Arcos, конечная концентрация 0,1%) и индуцировали в течение 1,5 ч при 30°C. После индукции бактерии наносили штрихом на чашки с хлорамфениколом и инкубировали в течение двух дней при 30°C. Колонии затем высевали методом отпечатков на чашки с хлорамфениколом/канамицином и хлорамфениколом, чтобы выявить клоны, которые утратили устойчивость к канамицину. Данные клоны подвергали дополнительному скринингу путем расщепления рестриктазами и ПЦР, чтобы подтвердить наличие конструкции.

Воспроизведение вируса.

Для того чтобы регенерировать вирус, ДНК ВАС переносили в клетки-хозяева млекопитающих, допускающие размножение вирусов. ДНК ВАС, очищенную из 10 мл ночной культуры, внедряли посредством электропорации в теломеризованные фибробласты, находящиеся при приблизительно 1/5 от конфлюэнтности во флаконе (~200000 клеток). Вкратце, клетки осаждали (1500 об/мин, 5 мин) и ресуспендировали в 700 мкл Opti-Mem. Данную смесь клеток затем добавляли в 50 мкл ДНК ВАС и аккуратно перемешивали перед переносом в 4 мм кювету. Электропорацию осуществляли, применяя GenePulser II от Bio-rad при 0,25 кВ и 0,95 мкФ. После электропорации клетки высевали на 100 мм чашки, содержащие DMEM+10% FBS, и среды заменяли на следующий день, чтобы удалить обломки клеток. За клетками наблюдали ежедневно для выявления образования бляшек и собирали при полном ЦПД. Остальные присоединенные клетки собирали с помощью скребка для клеток и осаждали путем центрифугирования (1500 об/мин, 5 мин) и супернатант, содержащий воспроизведенный вирусный вектор, сохраняли для пассажа рекомбинантного вируса. Осадки клеток лизировали путем ресуспендирования в 100 мкл загрузочного буфера, содержащего краситель, 5% ДСН и 10% 2-меркаптоэтанол, и центрифугировали на колонке QiaShred (Qiagen, Валенсия, Калифорния).

Экспрессия вирусных белков EpiGraph.

Экспрессию белков EpiGraph исследовали с помощью электрофореза в полиакриламидном геле с добавлением ДСН (ПААГ/ДСН) и вестерн-блоттинга с антителами к эпителиальной метке V5 или гемагглютинина, вставленной в каждую конструкцию. Вкратце, получали 10% полиакриламидные гели, загружали на них по 10 мкл (10%) каждого образца и подвергали электрофорезу при 110-120 В в течение 90 мин. Разделенные белки переносили на мембраны PVDF путем полусухого блоттинга при 20 В в течение 45-50 мин. Неспецифическое связывание блокировали с помощью раствора 10% обезжиренного сухого молока в фосфатно-солевом буфере с добавлением 0,1% tween-20 (ФБР-Т) в течение 60 мин. Антитела к НА (Sigma) или V5 (Santa Cruz) разбавляли в 5% растворе молока и инкубировали с мембранами в течение 1 ч, а затем трижды промывали ФБР-Т перед добавлением разведенного 1:2000 конъюгированного с пероксидазой хрена вторичного антитела козы против антител мыши (Santa Cruz) в течение 1 ч. Затем блоты промывали три раза в ФБР-Т, проявляли, используя набор с субстратом для усиленной хемилюминес-

ценции (ECL) (Thermo-Pierce), и визуализировали с помощью рентгеновской пленки. Для всех исследованных конструкций продемонстрировали надежную стабильную экспрессию белков с предсказанной молекулярной массой, таким образом, подтвердив их пригодность для тестирования иммуногенности в модели вакцины CMV для макака-резус (например, см. фиг. 10A-B).

Пример 8. Вакцины EpiGraph для популяции.

Алгоритм EpiGraph использовали для создания набора антигенов для вакцины, применяя изначально векторы CMV, тем не менее, можно применять другие системы доставки вакцины.

Рассматривали группу М (мировую), а также клады В и С (применение географически ограничено регионами, в которых эндемичны данные конкретные клады). Получали антигены EpiGraph для вакцины Gag, Pol и Nef. Экспрессировали группы В и М.

Основные параметры разработки EpiGraph.

В EpiGraph применяется теория графов/подход динамического программирования для разработки антигенов, которые максимизируют покрытие потенциальных Т-клеточных эпитопов (РТЕ). При некоторых условиях они математически оптимальны, и они очень эффективны в вычислительном отношении. У вариантов EpiGraph есть дополнительное заметное преимущество по сравнению с мозаичными антигенами, состоящее в том, что пользу от исключения очень редких эпитопов из конструкций можно уравновесить допущением минимальной зоны покрытия РТЕ. Данные варианты EpiGraph были разработаны с учетом этого, допуская небольшую зону покрытия (0,005), чтобы удостовериться в том, что даже самые редкие эпитопы, представленные в антигенах EpiGraph, наблюдаются во многих из популяционных последовательностей (точное количество зависит от входного набора данных).

Входные наборы данных для этих вариантов EpiGraph получали из набора выровненных последовательностей из базы данных последовательностей ВИЧ для каждого из указанных белков: Gag, Pol и Nef, - включая одну последовательность на человека, приблизительно сентябрь 2014 г. Неполноразмерные последовательности исключали. После этого оставались следующие количества последовательностей для каждого набора белков, Nполс представляет собой количество последовательностей в исходном выравнивании.

Нполс	Клада	Белок
1729	В	Gag
1780	В	Nef
1072	В	Pol
940	С	Gag
749	С	Nef
414	С	Pol
4596	М	Gag
4040	М	Nef
2780	М	Pol

Затем находили последовательное решение для наборов парных антигенов EpiGraph для бивалентной вакцины с применением алгоритма EpiGraph для невыровненных последовательностей. Использовали последовательное решение, которое позволяло применение первого EpiGraph в качестве моновалентной вакцины в изоляте. Это означает, что их разрабатывали таким образом, что сначала находили лучший отдельный антиген EpiGraph ("эписенсусный"), чтобы обеспечить оптимальное покрытие РТЕ в популяции, а затем его фиксировали для включения в разработку бивалентной вакцины. Затем находили дополнительный антиген с получением наилучшего покрытия РТЕ в популяции с помощью бивалентной пары антигенов, которая содержала первый антиген EpiGraph - эписенсус. Зоны покрытия были минимальны по сравнению с одновременным нахождением антигенов.

Затем проводили анализ, чтобы определить зону покрытия при исключении редких вариантов. Результаты кратко представлены в следующей таблице. f_0 представляет собой порог редкого эпитопа. Последовательности получали после отбрасывания всех РТЕ, которые встречаются в f_0 или в более редких последовательностях. Другими словами, каждый РТЕ, который находится в вакцине, встретился в более чем f_0 последовательностях. Данные значения f_0 сделали как можно большими при достижении покрытия, которое находилось в пределах 0,005 от максимального покрытия, достигаемого, когда $f_0=0$. Nполс представляет собой количество последовательностей в исходном выравнивании.

Белок	Клада	Нполс	f_0
Gag	В	1729	41
Gag	С	940	21

Gag	M	4596	146
Nef	B	1780	50
Nef	C	749	14
Nef	M	4040	100
Pol	B	1072	34
Pol	C	414	11
Pol	M	2780	67

Основные антигены EpiGraph были разработаны как полноразмерные белки, и их экспрессировали в векторах CMV и исследовали либо в виде слитых белков Gag/Nef или Pol с делециями для безопасности, либо в виде слитых наиболее консервативных областей Gag и Nef или слитых наиболее консервативных областей Pol. Перечень каждой из указанных полноразмерных форм белков для экспрессии в векторе CMV включен ниже в примеры для группы M и клады B.

Консервативные области для получения антигенов для вакцины вырезали из полноразмерных белков EpiGraph.

Консервативные области в Gag, Pol и Nef определяли на основании возможности покрытия РТЕ бивалентной (т.е. содержащей 2 антигена) вакциной. То есть, их определяли на основании возможности обеспечения двумя оптимизированными антигенами покрытия РТЕ клады B. Последовательности консервативных областей показаны в перечне ниже. Границы выбирали, чтобы захватить наиболее консервативную половину каждого из трех указанных белков (Gag, Pol и Nef), в насколько это возможно наиболее длинным непрерывном фрагменте, таким образом, между консервативными областями разбросаны более вариабельные участки. В Gag такие границы просто представляют собой границы капсидного белка p24.

Пример 9. Индивидуализированные терапевтические антигены.

В качестве индивидуализированных антигенов использовали либо Gag, либо наиболее консервативную часть Gag - p24. В одной разработке для полной группы M использовали 2015 выравниваний 4596 последовательностей, чтобы получить решение для мировой популяции. Выравнивание 189 последовательностей современной клады B, выделенных в США, или выравнивание 199 последовательностей современной клады C из Южной Африки. "Современный" относится к вирусам, которые были выделены после 2004 г.; все последовательности, образцы которых были собраны за последнее десятилетие, использовали для получения достаточной выборки последовательностей из каждого региона. Последовательности, которые не были полностью отделены или были неполноразмерными, исключали. Современные вирусы конкретного регионального подтипа выбирали, потому что они представляют собой хорошие совпадения для популяций, которые будут подходящими для первоначальных исследований обоснованности концепции для оценки индивидуализированных вакцин; используя региональные наборы одного подтипа мы смогли ограничить разнообразие популяции, чтобы увеличить вероятность успеха терапевтической вакцины против ВИЧ.

Такие антигены для терапевтических вакцин разработаны для применения в планах лечения, в которых они могли совпадать, или быть специально "индивидуализированными", с секвенированным вирусом, инфицировавшим индивида, который будет получать данную вакцину. Будут производить 6 антигенов для вакцины; 2 или 3 из данного набора, наилучшим образом совпадающие с инфицировавшим индивида вирусом, будут давать индивиду, чтобы максимизировать совпадения между вакциной и инфицировавшим его штаммом.

Таким образом, ниже представлен полноразмерный Gag для каждого EpiGraph для производства 6 разработанных индивидуализированных антигенов, где p24 выделен жирным шрифтом. В индивидуализированной вакцине можно применять либо Gag, либо более консервативный p24.

Совпадающие баллы РТЕ и несовпадающие баллы (количество дополнительных) для доставки n наилучшим образом совпадающих антигенов из пула из m, для белков Gag или внутреннего участка p24, где n выбраны так, чтобы наилучшим образом покрывать каждую отдельную последовательность в целевой популяции. Целевые популяции представляют собой популяцию, инфицированную современной кладой C, образцы которой брали в Южной Африке, или популяцию, инфицированную современной кладой B, образцы которой брали в США. В наилучших решениях используется пул из m=6 антигенов, и они выделены жирным шрифтом; они основаны на шести последовательностях, показанных выше.

Покрывание эпитопов Gag клады С из Южной Африки

n m	Покрывание	Дополнительные	Вакцина, которую оценили для набора данных
1 1	0,39886	318437	М Epigraph
2 2	0,57907	681980	М Epigraph
1 1	0,59591	200809	С Epigraph
2 2	0,70758	550618	С Epigraph
2 3	0,71875	559206	С индивидуализированная
3 3	0,73852	778603	С индивидуализированная
2 4	0,72609	550397	С индивидуализированная
3 4	0,75294	805658	С индивидуализированная
2 5	0,73124	549452	С индивидуализированная
3 5	0,75975	800337	С индивидуализированная
2 6	0,73440	527623	С индивидуализированная
3 6	0,76513	767759	С индивидуализированная

Покрывание эпитопов p24 клады С из Южной Африки

n m	Покрывание	Дополнительные	Вакцина, которую оценили для набора данных
1 1	0,48548	114739	М Epigraph
2 2	0,76864	230593	М Epigraph
1 1	0,73525	59040	С Epigraph
2 2	0,84494	204578	С Epigraph
2 2	0,84494	204578	С индивидуализированная
2 3	0,86153	205387	С индивидуализированная
3 3	0,88154	338417	С индивидуализированная
2 4	0,87000	188307	С индивидуализированная
3 4	0,89175	290854	С индивидуализированная
2 5	0,87699	187849	С индивидуализированная
3 5	0,89945	295930	С индивидуализированная
2 6	0,87996	191538	С индивидуализированная
3 6	0,90373	290985	С индивидуализированная

Покрывание эпитопов Gag клады В из США

n m	Покрывание	Дополнительные	Вакцина, которую оценили для набора данных
1 1	0,55488	237815	М Epigraph
2 2	0,67323	629333	М Epigraph
1 1	0,61174	195720	В Epigraph
2 2	0,72471	553899	В Epigraph
2 3	0,73615	557254	В индивидуализированная
3 3	0,75734	786772	В индивидуализированная
2 4	0,74354	547725	В индивидуализированная
3 4	0,76973	762524	В индивидуализированная
2 5	0,74786	546111	В индивидуализированная
3 5	0,77777	779487	В индивидуализированная
2 6	0,75226	536619	В индивидуализированная
3 6	0,78669	773000	В индивидуализированная

Покрывание эпитопов p24 клады В из США

n	m	Покрывание	Дополнительные	Вакцина, которую оценили для набора данных
1	1	0,73131	59979	М EpiGraph
2	2	0,84204	214296	М EpiGraph
1	1	0,76421	52646	В EpiGraph
2	2	0,87373	201233	В EpiGraph
2	3	0,88878	194434	В индивидуализированная
3	3	0,90810	297571	В индивидуализированная
2	4	0,89779	190635	В индивидуализированная
3	4	0,91754	282386	В индивидуализированная
2	5	0,90461	191794	В индивидуализированная
3	5	0,92704	285005	В индивидуализированная
2	6	0,90859	187889	В индивидуализированная
3	6	0,93556	280704	В индивидуализированная

Пример 10. Тестирование вакцины.

Группы для первоначального тестирования вакцины с CMV включают: 1) один эписенсусный для популяции антиген Gag, центральный для клады В, эпидемической в США; 2) эписенсус для популяции плюс индивидуализированный белок Gag, выбранный как наилучшим образом совпадающий с таковым для природного штамма ВИЧ-1; 3) эписенсус для популяции плюс индивидуализированный белок Gag, выбранный как наилучшим образом совпадающий с отличным (и отдаленным) природным штаммом ВИЧ-1; 4) эписенсус для популяции плюс оба белка Gag из групп 2 и 3. Вызванные иммунные ответы анализировали, используя перекрывающиеся 15-мерные пептиды (перекрывание по 4 аминокислотам), соответствующие антигену для вакцины (чтобы определить совокупные вызванные вакциной Gag-специфические ответы), а затем анализировали иммунные ответы как на "целевой" штамм ВИЧ-1, так и на выбранные нецелевые штаммы ВИЧ (см. ниже) для измерения специфических ответов на штамм и уровень совпадения эпитопов (сравнение целевой и нецелевых последовательностей Gag ВИЧ). Определяли, действительно ли разработанные с помощью компьютера комбинации вставок и вектора вызывают более сильные и широкие ответы Т-клеток на целевой штамм, при этом минимизируя сопоставимые ответы на нецелевой штамм. Результаты данного анализа позволяют экспериментально проверить спрогнозированные совпадения эпитопов, полученные в примере 1.

Четыре группы по 5 макак-резус (RM) инокулировали 10^6 бляшкообразующих единиц (PFU) векторов HCMV, как описано далее: группа 1 получила один вектор, содержащий эписенсусную последовательность клады В, группа 2 получила эписенсусный вектор плюс один вектор индивидуализированной вакцины, группа 3 получила эписенсусный вектор плюс отличный вектор индивидуализированной вакцины и группа 4 получила эписенсусный вектор плюс оба вектора индивидуализированной вакцины. Данные вакцины "индивидуализированы" по отношению к 2 примерам передаваемых штаммов/штаммов-основателей ВИЧ, выбранных из небольшого набора из 9 ВИЧ-инфекций клады В США, которые были обширно секвенированы в долгосрочном исследовании; данная панель представляет собой спектр природных инфекций, сравнимый с таковым, который может встретиться при клиническом испытании у человека. По существу, данные 2 дивергентных природных штамма использовались в качестве прототипа пациентов. Пара вставок для эписенсусной/индивидуализированной вакцины, которая оптимизирована для одного штамма, будет субоптимальной для другого, и наоборот, что позволяет, путем определения перекрестного узнавания наборов пептидов, отражающих специфические последовательности каждого штамма, провести взаимный анализ того, есть ли польза от нашей стратегии совпадения с индивидуализированной последовательностью. Дополнительно к данному взаимному анализу последовательность из других 7 инфекций будет изменяться по относительному "совпадению" с выбранной вставкой(-ами) в вектор вакцины, и путем анализа ответов на наборы пептидов, отражающих последовательности данных других 7 штаммов ВИЧ, будет подробно проанализирована взаимосвязь между последовательностью вставки в вектор и комбинацией векторов и влияние на силу и широту совпадающих ответов по сравнению с несовпадающими ответами. Часть обоснования сосредоточена на данном набо-

ре из 9 штаммов состоит в том, что доступны тщательно секвенированные полноразмерные геномы из образцов из долгосрочных исследований. Изоляты также доступны в виде инфекционных молекулярных клонов, которые могут быть полезны для оценки ответов в последующих испытаниях на человеке. Таким образом, при проведении основополагающей работы на RM с применением данного набора, предусматривается особенно полезный совпадающий набор белков для непосредственного сравнения ответов макаки и человека, когда данные вакцины распространят на испытания на человеке.

Макак-резус (RM) инокулировали подкожно в день 0 и в неделю 12 и прослеживали в долгосрочном исследовании в течение одного года. Поскольку на вакцинацию векторами HCMV не влияет уже существующий иммунитет к RhCMV, то для данных экспериментов использовали животных, которые были естественным образом инфицированы RhCMV. Анализ внутриклеточных цитокинов методом проточной цитометрии (ICS) использовали для определения ответа CD4⁺ и CD8⁺ Т-клеток на отдельные последовательные 15-мерные пептиды, содержащие последовательности вакцины внутри вставок вакцины, которые вводили каждому животному (который будет представлять собой совокупный вызванный вакциной ответ). Затем определили, распознают ли данные специфические к эпитопу Т-клетки варианты эпитопа как в целевом штамме, так и в других 8 нецелевых штаммах. Для пептидов, для которых выявили ответы на специфические для штамма эпитопы, сравнивали величину, функциональную avidность и функциональные особенности (продукцию IFN- γ , TNF- α , IL-2 и MIP-1 β и экстернализацию CD107) данных ответов на "исходные" (последовательности вставки вакцины) варианты пептида, чтобы определить степень функциональной способности вступать в перекрестные реакции. В выбранных случаях использовали анализ укорачивания, чтобы определить основной эпитоп для сходного сравнительного анализа. Для того, чтобы определить процент присутствующих ограниченных MHC-II CD8⁺Т-клеток, "блокирующие" МАТ, специфические к MHC-I и MHC-II, и полученный из инвариантной цепи специфический к MHC-II связывающий пептид CLIP использовали для ингибирования специфических для гриппа ответов CD8⁺ Т-клеток в РВМС, что ранее осуществляли для ответов на ВНО.

Независимо от результатов примирения Т-клеток, связанных со специфическими для вектора генами, иммунологический анализ позволяет проверить гипотезу, состоящую в том, что Т-клетки, индуцированные индивидуализированными вакцинами, лучше в отношении покрытия эпитопов данного природного резервуара ВИЧ по сравнению с таковыми, индуцированными неиндивидуализированными вакцинами. Индивидуализированный коктейль из 2 антигенов мог, вероятно, вызывать ответы Т-клеток с по меньшей мере на 25% более высокой частотой перекрестных ответов на подходящий природный штамм по сравнению с эписенсусом отдельно (табл. 5). Также провели исследование, чтобы определить, будет ли включение в коктейль 3 дополнительных индивидуализированных антигенов вызывать большую способность вступать в перекрестные реакции, или будет ли разведение антигена или присутствие больших количеств специфических для вакцины эпитопов в результате добавления 3 вместо 2 антигенов фактически уменьшать величину или широту перекрестного ответа на природный белок. Ожидают, что ответы Т-клеток на CMV будут гораздо обширнее и, следовательно, будут покрывать гораздо больший процент последовательностей, чем описано ранее для других векторов. Таким образом, даже при относительно небольшом количестве животных, должны наблюдаться достаточные ответы на эпитоп для того, чтобы оценить влияние изменения последовательности на возможность вступать в перекрестные реакции в рамках данных ответов. Количество и величину всех ответов на вакцины определяли, применяя совпадающие с вакциной наборы пептидов. После того, как определили целевые пептиды, применяя только такие пептиды, которые положительны у каждого животного, определяли влияние природной изменчивости на каждый вызывающий ответ на вакцину пептид. Природные варианты, которые исследовали, были выбраны на основании вариантов, обнаруженных в исходной панели, включая как индивидуализированные, так и плохо совпадающие Gag. Непараметрические и вычислительные статистические способы с повторной выборкой использовали в качестве первичного способа оценки влияния изменчивости эпитопа на уменьшение величины или прекращение узнавания. Данные анализы, тем не менее, дополняются применением обобщенных линейных моделей, при необходимости, чтобы исследовать влияние более сложных взаимодействий на возможность перекрестного ответа Т-клеток.

SEQ ID NO: 691

Эписенсусная последовательность индивидуализированного антигена для вакцины.

MGARASVLSGGELDRWEKIRLRPGGKKYKLVHIVWASRELERFAVNPGLLETSEGCRQILGQLQPSLQTGS
 EELKSLYNTVATLYCVHQKIEVKDTKEALDKIEEEQNKS KKKAAQ-----AAADTGNSS-----
 QVSNQYPIVQNLQGGQMVHQAI SPRTLNAWVKVVEEKAFSPVIMFSALEGATPQDLNTMLNTVGGHQAAMQMLKET
 INEEAAEWDRLHPVHAGPIAPGQMREPRGSDIAGTTSTLQEQIGWMTNPNPIPVGEIYKRWIILGLNKIVRMYSPSTSI
 LDIKQGPKEPFRDYVDRFYKTLRAEQASQEVKNWMTETLLVQNSNPDCCKTILKALGPAATLEEMMTACQGVGGP SHKA
 RVLAEAMSQVTN--SATIMMQRGNFRNQRKTVKCFNCGKEGHIARNCRAPRKKGCWKCGKEGHQMKDC--
 TERQANFLGKIWPSH-KGRPGNFLQ-----SRPEPT-----APPEESFRFGEETTTPS-----
 QKQEPIDKE-----LYP-LASLRSFLGNDPSSQ

SEQ ID NO: 692

Последовательность индивидуализированного антигена для вакцины 1.

MGARASVLSGGELDKWEKIRLRPGGKKYRLKHIVWASRELERFAVNPGLLETSGGCRQILEQLQPSLQTGS
 EELRSLYNTVATLYCVHQKIDVKDTKEALEKIEEEQNKS KKKAAQ-----AAAAADTGNNS-----
 QVSNQYPIVQNMQGGQMVHQALS PRTLNAWVKVIEEKAFSPVIMFTALSEGATPQDLNTMLNSVGGHQAAMQMLKET
 INEEAAEWDRLHPVHAGPVAPGQMREPRGSDIAGTTSTLQEQIGWMTSNPPIPVGEIYKRWIIMGLNKIVRMYSPSTSI
 LDIKQGPKEPFRDYVDRFYKTLRAEQATQEVKNWMTETLLVQNSNPDCCKTILKALGPGATLEEMMTACQGVGGP SHKA
 RVLAEAMSQATN--SATIMMQKGNFRNQRKIVKCFNCGKEGHI AKNCRAPRKKGCWKCGREGHQMKDC--
 TERQVNFGLGIWPSH-KGRPGNFLQ-----NRPEPT-----APPAESFRFGEETTTPP-----
 QKQEPIDKE-----LYP-LASLKSFLGNDPSSQ

SEQ ID NO: 693

Последовательность индивидуализированного антигена для вакцины 2.

MGARASVLSGGKLDKWEKIRLRPGGKKYKLVHIVWASRELERFAVNPGLLETAEGCRQILGQLQPALQTGS
 EELKSLFNTVATLYCVHQRIDVKDTKEALEKIEEEQNKS KKKAAQ-----AAADTGSSS-----
 QVSNQYPIVQNMQGGQMVHQAI SPRTLNAWVKVIEEKAFSPVIMFAALSEGATPQDLNMLNTVGGHQAAMQMLKET
 INEEAAEWDRLHPVHAGPVAPGQMREPRGSDIAGTTSTLQEQIGWMTNPNPIPVGEIYKKWIIMGLNKIVRMYSPSTSI
 LDIKQGPKEPFRDYVDRFYKTLRAEQATQEVKNWMTETLLVQNSNPDCCKTILKALGPGATLEEMMTACQGVGGP SHKA
 RVLAEAMSQATN--SATIMMQRGNFRNQRKTIKCFNCGKEGHI AKNCRAPRKKGCWKCGREGHQMKDC--
 TERQANFLGKIWPSH-KGRPGNFLQ-----NRPEPT-----APPAESFRFGEETTTPP-----
 QKQEPIDKE-----LYP-LASLKSFLGNDPSSQ

SEQ ID NO: 694

Последовательность индивидуализированного антигена для вакцины 3.

MGARASVLSGGELDKWEKIRLRPGGKKYKLVHIVWASRELERFAVNPGLLETAEGCRQILEQLQPSLQTGS
 EELRSLYNTVAVLYCVHQRIDVKDTKEALDKIEEEQNKS KKKAAQ-----AAAAADTGNNS-----
 QVSNQYPIVQNLQGGQMVHQPI SPRTLNAWVKVIEEKAFSPVIMFTALSEGATPQDLNTMLNTVGGHQAAMQMLKDT
 INEEAAEWDRLHPVQAGPIAPGQIREPRGSDIAGTTSTLQEQIAWMTNPNPIPVGDIYKRWIIMGLNKIVRMYSPSTSI
 LDIKQGPKEPFRDYVDRFYKTLRAEQASQDVKNWMTETLLVQNSNPDCCKTILKALGPGATLEEMMTACQGVGGP SHKA
 RVLAEAMSQATN--SATIMMQKGNFRNQRKIVKCFNCGKEGHI AKNCRAPRKKGCWKCGREGHQMKDC--
 TERQANFLGKIWPSH-KGRPGNFLQ-----SRPEPT-----APPAESFRFGEETTTPP-----
 QKQEPIDKE-----LYP-LASLKSFLGNDPSSQ

SEQ ID NO: 695

Последовательность индивидуализированного антигена для вакцины 4.

MGARASVLSGGKLDKWEKIRLRPGGKKYKLVHIVWASRELERFAVNPGLLETAEGCRQILGQLQPALQTGS
 EELRSLYNTVATLYCVHQRIEVKDTKEALEKIEEEQNKS KKKVQ-----AAAAADTGNNS-----
 QVSNQYPIVQNIQGGQMVHQPL SPRTLNAWVKVIEEKAFSPVIMFTALSEGATPQDLNTMLNTIGGHQAAMQMLKET
 INEEAADWDLRLHPVHAGPVAPGQMREPRGSDIAGTTSNLQEQIGWMTSNPPIPVGEIYKRWIIMGLNKIVRMYSPSTSI
 LDIKQGPKEPFRDYVDRFYKTLRAEQASQDVKNWMTETLLVQNSNPDCCKTILKALGPGATLEEMMTACQGVGGP SHKA
 RVLAEAMSQVTN--SATIMMQKGNFRNQRKIVKCFNCGKEGHI AKNCRAPRKKGCWKCGKEGHQMKEC--
 TERQANFLGKIWPSY-KGRPGNFLQ-----SRPEPS-----APPEESFRFGEETATPS-----
 QKQEPIDKE-----LYP-LASLKSFLGNDPSSQ

SEQ ID NO: 696

Последовательность индивидуализированного антигена для вакцины 5.

MGARASVLSGGKLDKWEKIRLRPGGKKQYKLKHLVWASRELERFAVNPGLLETAEGCRQILGQLQPALQTGS
EELKSLFNTVATLYCVHQRIDVVDKDTKEALEKIEEEQNKSKKKAQQ-----AAADTGNNS-----
QVSQNYPIVQNIQGQMVHQALSPTLNAWVKVIEEKAFSPPEVIMFSALEGATPQDLNMLNTVGGHQAMQILKET
INEEAAEWDRHLHPVQAGPVAPGQMRPRGSDIAGTTSNLQEQIAWMTNHPPIPVGEIYKRWIIMGLNKIVRMYSPTSI
LDIRQGPKEPFRDYVDRFYKTLRAEQASQDVKNWMTETLLVQNSNPCKTILKALGPATLEEMMTACQGVGGPSHKA
RVLAEAMSQATN--PATIMMQRGNFKNQKIVKCFNCGKEGHIKNCRAPRKKGCWKCCKEGHQMKEC--
TERQANFLGKIWPSY-KGRPGNFLQ-----SRPEPS-----APPEESFRFGEETTTP-----
QKQEPIDKE-----LYP-LTSLRSLFGNDPSSQ

SEQ ID NO: 697

Последовательность антигена 1 EpiGraph.

MGARASVLSGGELDRWEKIRLRPGGKKKYKLKHIVWASRELERFAVNPGLLETSEGCRQILGQLQPSLQTGS
EELKSLYNTVATLYCVHQRIDVVDKDTKEALDKIEEEQNKSKKKAQQ-----AAADTGNSS-----
QVSQNYPIVQNLQGQMVHQAI SPRTLNAWVKVVEEKAFSPPEVIMFSALSEGATPQDLNMLNTVGGHQAMQMLKET
INEEAAEWDRHLHPVHAGPIAPGQMRPRGSDIAGTTSNLQEQI GWMTNHPPIPVGEIYKRWIILGLNKIVRMYSPTSI
LDIRQGPKEPFRDYVDRFYKTLRAEQASQEVKNWMTETLLVQANPDCKTILKALGPAATLEEMMTACQGVGGPGHKA
RVLAEAMSQVTN--SATIMMQRGNFRNQRKTVKCFNCGKEGHIARNCRAPRKKGCWKCCKEGHQMKEC--
TERQANFLGKIWPSH-KGRPGNFLQ-----SRPEPT-----APPEESFRFGEETTTPS-----
QKQEPIDKE-----LYP-LASLRSFLFGNDPSSQ

SEQ ID NO: 698

Антиген 1 коктейля EpiGraph.

MGARASVLSGGELDRWEKIRLRPGGKKKYKLKHIVWASRELERFAVNPGLLETSEGCRQILGQLQPSLQTGS
EELKSLYNTVATLYCVHQRIDVVDKDTKEALDKIEEEQNKSKKKAQQ-----AAADTGNSS-----
QVSQNYPIVQNLQGQMVHQAI SPRTLNAWVKVVEEKAFSPPEVIMFSALSEGATPQDLNMLNTVGGHQAMQMLKET
INEEAAEWDRHLHPVHAGPIAPGQMRPRGSDIAGTTSNLQEQI GWMTNHPPIPVGEIYKRWIILGLNKIVRMYSPTSI
LDIRQGPKEPFRDYVDRFYKTLRAEQASQEVKNWMTETLLVQANPDCKTILKALGPAATLEEMMTACQGVGGPGHKA
RVLAEAMSQVTN--SATIMMQRGNFRNQRKTVKCFNCGKEGHIARNCRAPRKKGCWKCCKEGHQMKEC--
TERQANFLGKIWPSH-KGRPGNFLQ-----SRPEPT-----APPEESFRFGEETTTPS-----
QKQEPIDKE-----LYP-LASLRSFLFGNDPSSQ

SEQ ID NO: 699

Антиген 2 коктейля EpiGraph.

MGARASVLSGGKLDKWEKIRLRPGGKKKYRLKHIVWASRELERFAINPGLLETAEGCRQILGQLQPALQTGS
EELKSLYNTVATLYCVHQRIDVVDKDTKEALEKIEEEQNKSKKKAQQ-----AAAAADTGNSS-----
QVSQNYPIVQNMQGQMVHQALSPTLNAWVKVIEEKAFSPPEVIMFTALSEGATPQDLNMLNTVGGHQAMQILKET
INEEAAEWDRHLHPVHAGPVAPGQMRPRGSDIAGTTSNLQEQI GWMTSNHPPIPVGEIYKRWIIMGLNKIVRMYSPTSI
LDIKQGPKEPFRDYVDRFYKTLRAEQASQDVKNWMTETLLVQNSNPCKTILKALGPATLEEMMTACQGVGGPSHKA
RVLAEAMSQATN--SATIMMQRGNFRNQRKIVKCFNCGKEGHIKNCRAPRKRGCWKCCKEGHQMKEC--
TERQANFLGKIWPSY-KGRPGNFLQ-----NRPEPT-----APPAESFRFGEETTTP-----
QKQEPIDKE-----LYP-LASLKSFLFGNDPSSQ

SEQ ID NO: 700

Антиген 3 коктейля EpiGraph.

ARASVLSGGELDKWEKIRLRPGGKKQYKLKHLVWASRELERFAINPGLLETSGGCRQILEQLQPSLQTGSEE
LRSLYNTVAVLYCVHQRIEVKDTKEALEKVVEEQNKSKKKVQQ-----AAADTGNNS-----
QVSQNYPIVQNIQGQMVHQFISPTLNAWVKVIEDKAFSPPEVIMFAALSEGATPQDLNMLNTIGGHQAMQMLKDT
INEEAAEWDRHLHPVQAGPVAPGQMRDPRGSDIAGTTSNLQEQIAWMTNHPPIPVGDIYKRWIILGLNKIVRMYSPTSI
LDIKQGPKEPFRDYVDRFYKTLRAEQATQEVKNWMTETLLVQANPDCTILKALGPAATLEEMMSACQGVGGPSHKA
RILAEAMSQVTN--
STAIMMQRGNFKNQKTVKCFNCGREGHIAKNCRAPRKKGCWKCCKEGHQMKEC--
KGRPGNFLQ-----SRPEPS-----APPEESFRFGEETATPS-----QKQEPIDKE-----
LYP-LTSLRSLFGNDPSSQ

SEQ ID NO: 701

EpiGraph 1 слитого gag/nef ВИЧ В

MGARASVLSGGELDRWEKIRLRPGGKKKYRLKHIVWASRELERFAVNPGLLETSEGCRQILGQLQPSLQTGS
EELKSLYNTVATLYCVHQRIEVKDTKEALDKIEEEQNKSKKKAQQAADTGNSSQVSQNYPIVQNLQGQMVHQAI SP
RTLNAWVKVVEEKAFSPPEVIMFSALSEGATPQDLNMLNTVGGHQAMQMLKETINEEAAEWDRHLHPVHAGPIAPGQ

REPRGSDIAGTTSTLQEIQGWMTNPPPIPVGEIYKRWIILGLNKIVRMYSPTSILDIRQGPKEPFRDYVDRFYKTLRA
EQASQEVKNWMTETLLVQANPDCKTILKALGPAATLEEMMTACQGVGGPGHKARVLAEAMSQVNTSATIMMQRGNFR
NQRKTVKCFNCGKEGHIACNCRAPRKKGCWCKGKEGHQMKDCTERQANFLGKIWPSHKGRPGNLFQSRPEPTAPPEES
FRFGEETTPSQKQEPIDKELYPLASLKS LFGNDPSSQGGKWSKSSIVGWPAVREMRRAEPAAEGVGAVSRDLEKHG
AITSSNTAATNADCAWLEAQEEEEVGFPVRPQVPLRPMTYKGAALDLSHFLKEKGGLEGLIYSQKRQDILDLWVYHTQG
YFPDWQNYTPGPGIRYPLTFGWCFKLVPVEPEKVEEANEENNSLLHPSMSQHGMDDPEKEVLWKFDSRLAFHHMARE
LHPEYYKDC

SEQ ID NO: 702

EpiGraph 2 слитого gag/nef ВМЧ В

MGARASVLSGGKLDKWEKIRLRPGGKKKYKLKHIVASRELERFALNPGLLETSEGCKQILGQLQPALQTGS
EELRSLYNTVAVLYCVHQRIDVKDTKEALEKIEEEQNKCKKKAQAAAAADTGNNSQVVSQNYPIVQNMQGMVHQALS
PRTLNAWVKVIEEKAFSPEVIMFTALSEGATPQDLNLTMLNTIGGHQAAMQMLKDTINEEAAEWDVHPVHAGPVPAG
QMREPRGSDIAGTTSNLQEIQAWMTNPPPIPVGEIYKRWIIMGLNKIVRMYSVPSILDIRQGPKEPFRDYVDRFYKTL
RAEQASQDVKNWMTETLLVQNSNDCKTILKALGPATLEEMMTACQGVGGPSSHARVLAEAMSQVNTSATIMMQRGNFR
FRNQRKIVKCFNCGKEGHIARNCRAPRKKGCWCKGKEGHQMKECTERQANFLGKIWPSYKGRPGNLFQNRPEPTAPPA
ESFRFGEETATPPQKQEPIDKEMYPLASLRS LFGNDPSSQGGKWSKRSVPGWNTIRERMRTTEPAEAGVGAASRDLEH
GAITSSNTAANNAACAWLEAQEDEEVGFPVKPQVPLRPMTYKAAVDLSHFLKEKGGLEGLIHSQKRQDILDLWVYNTQ
GYFPDWHNYTPGPGTRFPLTFGWCFKLVPVDPEQVEKANEGENNCLLHPSMSLHGMDDPEREVLWKFDSRLAFHHVAR
EKHPEYYKDC

SEQ ID NO: 703

EpiGraph 1 pol ВМЧ В

MFFRENLAFFQKGAREFSSEQTRANSPTRRELQVWGRDNNLSLSEAGADRQGTVSFSFPQITLWQRPPLVTIKI
GGQLKEALLADDTVLEEMNLPGRWKPKMIGGIGGFIVKVRQYDQIPIEICGHKAIGTVLVGPTPVNIIGRNLTLQIGCT
LNFPISPIETVPVKLPKGMGDPKVKQWPLTEEEKIKALVEICTEMEKEGKISKIGPENPYNTPVFAIKKKDSTKWRKLV
DFRELNRKTQDFWEVQLGIPHPAGLKKKKSVTVLVDGDAYFSVPLDKDFRKYTAFTIPINNTEPGIRYQYNVLPQGW
KGSPAIFQSSMTKILEPFRKQNPDIIVYQLYVGSLEIGQHRTKIEELRQHLLRWGFTTPDKKHQKEPPFLWMGYELH
PDKWTVQPIVLPEKDSWTVNDIQKLVGKLNWASQIYAGIKVKQLCKLLRGTKALTEVPLTEEALELAENREILKEP
VHGVYYPDSKDLIAEIQKQGGQWYQIYQEPFKNLKTGKYARMGAHTNDVQLTEAVQKIATESIWIWGKTPKFKL
PIQKETWEAWWTEYWQATWIPEWEFVNTPLVKLWYQLEKEPIVGAETFYVDGAANRETCLGKAGYVTDGRQKVVSL
TDTTNQKTQAIHLALQDSGLEVNIVTDSQYALGIIQAQPDKSESELVSIIEQLIKKEKVYLAWVPAHKIGGNEQVD
KLVSAGIRKVLFLDGDIDKAQEEHEKYHSNWRAMASDFNLPPVVAKEIVASCDKQCKLGEAMHGQVDCSPGIWQLCTHL
EGKIIIVAVHVASGYIEAEVI PAETGQETAYFLKLKLAGRWPKTIHTNGSNFTSTTVKAACWWAGIKQEFGIPYNPQS
QGVVSMNKELKKIIGQVRDQAEHLKTAVQMAVFIHNFKRKGGIGGYSAGERIVDIIATDIQTKELQKQITKIQNFRVY
YRDSRDLWKGPAKLLWKGEAVVIQDNSDIKVVPRRKAKIIRDYGKQMGAGDDCVASRQDED

SEQ ID NO: 704

EpiGraph 2 pol ВМЧ В

MFFREDLAFFQGEAREFPSEQTRANSPTRRELQVWGGDNNSPSEAGADRQGTVLSFPQITLWQRPPLVTVKI
GGQLKEALLADDTVLEEMSLPGWKPKMIGGIGGFIVKVRQYDQVPIEICGHKTIGTVLIGPTPVNIIGRNLTLQIGCT
LNFPISPIETVPVKLPKGMGDPVKQWPLTEEEKIKALIEICTEMEKEGKISRIGPENPYNTPIFAIKKKDSTKWRKLV
DFRELNRKTQDFWEVQLGIPHPGSLKKKKSVTVLVDGDAYFSVPLDKDFRKYTAFTIPSTNNTEPGIRYQYNVLPQGW
KGSPAIFQCSMTKILEPFRKQNPDIIVYQLYVGSLEIGQHRAKIEELRQHLLRWGFTTPDKKHQKEPPFLWMGYELH
PDKWTVQPIELPEKDSWTVNDIQKLVGKLNWASQIYPIKVRQLCKLLRGAKALTEVIPLTKEAELELAENREILREP
VHGVYYPDTKDLIAEIQKQGLGQWYQIYQEPFKNLKTGKYARTGAHTNDVRQLTEAVQKITTESIWIWGKTPKFRL
PIQKETWETWTEYWQATWIPEWEFVNTPLVKLWYQLEKEPIIIGAETFYVDGASNRETCLGKAGYVVTNRGRQKVISL
TDTTNQKTQAIYALQDSGSEVNIVTDSQYALGIIQAQPDQSESELVNQIIIEQLINKEKVYLAWVPAHKIGGNEQV
DKLVSTGIRKVLFLDGDIDRAQEEHEKYHNNWRAMASDFNLPPIVAKEIVASCDKQCKLGEAIIHGQVDCSPGIWQLCTH
LEGKVIIVAVHVASGYIEAEVI PAETGQETAYFILKLAGRWPKTVHTNGSNFTSATVKAACWWAGVKQEFGIPYNPQ
SQGVVSMNNELKKIIGQIRDQAEHLKTAVQMAVFIHNFKRKGGIGEYSAGERIIDIIATDIQTRRELQKQITKIQNFRVY
YYRDNDRDLWKGPAKLLWKGEAVVIQDENSEIKVVPRRKVKIIRDYGKQMGAGDDCVAGRQDED

SEQ ID NO: 705

EpiGraph 1 слитого gag/nef ВМЧ М

MGARASVLSGGKLDWEKIRLRPGGKKKYRLKHLVWASRELERFALNPGLLETSEGCRQILGQLQPSLQTGS
EELKSLYNTVATLYCVHQRIEVKDTKEALDKIEEEQNKSKKKAQAAAAADTGNSSQVVSQNYPIVQNLQGMVHQAI
SPRTLNAWVKVIEEKAFSPEVIMFSALESEGATPQDLNLTMLNTVGGHQAAMQMLKETINEEAAEWDRLHPVHAGPIAPGQM
REPRGSDIAGTTSTLQEIQGWMTNPPPIPVGEIYKRWIILGLNKIVRMYSVPSILDIRQGPKEPFRDYVDRFYKTLRA

EQATQEVKNWMTETLLVQNANPDCKTILKALGPAATLEEMMTACQGVGGPGHKARVLAEAMSQVNTSATIMMQRGNFK
GQKRIKCFNCGKEGHIARNCRAPRKKGCGKCGKEGHQMKDCTERQANFLGKIWPSHKGRPGNFLQSRPEPTAPPEESF
RFGEETTPSQKQEPIDKELYPLASLKSFLGNDPLSQGGKWSKSSIVGWPVAVRERMRAEPAEAGVGAVSRDLEKHGA
ITSSNTAATNADCAWLEAQEEEEVGFPVRPQVPLRPMTYKGAALDLSHFLKEKGGLLEGLIYSKKRQEILDLWVYHTQGY
FPDWQNYTPGPGIRYPLTFGWCFKLVPVDPREVEEANEENNCLLHPMSQHGMDDPEKEVLMWKFDSRLAFHHMAREL
HPEYYKDC

SEQ ID NO: 706

EpiGraph 1 консервативного gag/nef ВИЧ М

MPIVQNLQGGQMVHQAI SPRTLNAWVKVVEEKAFAFSEVI PMFSALSEGATPQDLNTMLNTVGGHQAAMQMLKE
TINEEAAEWDRHLPVHAGPIAPGQMREPRGSDIAGTTSTLQEQIGWMTNNPPIPVGEIYKRWIILGLNKIVRMYSPTS
ILDIRQGPKEPFRDYVDRFYKTLRAEQATQEVKNWMTETLLVQNANPDCKTILKALGPAATLEEMMTACQGVGGPGHK
ARVLVGGFVVRPQVPLRPMTYKGAALDLSHFLKEKGGLLEGLIYSKKRQEILDLWVYHTQGYFPDWQNYTPGPGIRYPLTF
GWCFKLVP

SEQ ID NO: 707

EpiGraph 2 слитого gag/nef ВИЧ М

MGARASVLSGGELDRWEKIRLRPGGKKYKLKHIVWASRELERFAVNPGLLETAEGCQIIEQLQSTLKTGS
EELKSLFNTVAVLYCVHQRIDVKDTKEALEKIEEEQNKSQQTQQAAGTGSSSKVSQNYPIVQNAQGMVHQPLSPR
TLNAWVKVIEEKAFAFSEVI PMFTALSEGATPQDLNMLNIVGGHQAAMQMLKDTINEEAAEWDRVHPVHAGPIPPGQM
REPRGSDIAGTTSNLQEQIGWMTSNPPIPVGDIYKRWIILGLNKIVRMYSPTSILDIKQGPKEPFRDYVDRFFKTLRA
EQASQEVKNWMTDTLLVQNANPDCKTILRALGPGATLEEMMTACQGVGGPSHKARVLAEAMSQVNTSATIMMQRGNFR
NQRKTVKCFNCGKEGHLARNCRAPRKKGCGKCGREGHQMKDCNERQANFLGKIWPSNKGPRGNFPQSREPTAPPAES
FRFEETTPAPKQEPKDRPLTSLKSLFGSDPLSQGSKWSKSSIVGWPVAVRERMRTPEAAEAGVGAASRDLEHGAITS
NEPTAANNADCAWLEAQEEEEVGFPVKPQVPLRPMTYKAAFDLSFFLKEKGGLDGLIYSQKRQDILDLWVYNTQGFPPD
WQNYTPGPGVRYPLTFGWCFKLVPVEPEKVEEANEENNSLLHPMSLHGMDDPEREVLWKFDSRLARRHMARELHPE
FYKDC

SEQ ID NO: 708

EpiGraph 2 консервативного gag/nef ВИЧ М

MPIVQNAQGGQMVHQPLSPRTLNAWVKVIEEKAFAFSEVI PMFTALSEGATPQDLNMLNIVGGHQAAMQMLKD
TINEEAAEWDRVHPVHAGPIPPGQMREPRGSDIAGTTSNLQEQIGWMTSNPPIPVGDIYKRWIILGLNKIVRMYSPTS
ILDIKQGPKEPFRDYVDRFFKTLRAEQASQEVKNWMTDTLLVQNANPDCKTILRALGPGATLEEMMTACQGVGGPSHK
ARVLVGGFVVRPQVPLRPMTYKAAFDLSFFLKEKGGLDGLIYSQKRQDILDLWVYNTQGFPPDWQNYTPGPGVRYPLTF
GWCFKLVP

SEQ ID NO: 709

EpiGraph 1 pol ВИЧ М

MFFRENLAFFPQGEAREFSSEQTRANSPTRRELQVWGRDNNSLSEAGADRQGTVSFSFPQITLWQRPLVTIKI
GGQLKEALLADDTVLEDINLPGKWPKMIGGIGGFIKVRQYDQILIEICGKKAIGTVLVGPTPVNIIIGRNLITQIGCT
LNFPISPIETVPVKLPGMDGPKVKQWPLTEEKIKALVEICTEMEKEGKISKIGPENPYNTPVFAIKKKDSTKWRKLV
DFRELNKRTQDFWEVQLGIPHPAGLKKKSVTVLDVGDAYFSVPLDKDFRKYTAFTIP SINNETPGIRYQYNVLPQGW
KGSPAIFQSSMTKILEPFRKQNPDIVIYQLYVGSDEIGQHRTKIEELRQHLLKWGFTTPDKKHQKEPFLWMGYELH
PDKWTVQPIVLPKDSWTVNDIQKLVGKLNWASQIYPGIVKQKLCCLLRGAKALTDIVPLTEEALELEAENREILKEP
VHGYYDPSKDLIAEIQKQGDQWYQIYQEPFKNLKTGKYARMGAHTNDVKQLTEAVQKIATESIWIWKT PKFRL
PIQKETWETWWTWYQATWIPWEFEVNTPLVKLWYQLEKEPIVGAETFYVDGAANRETKLGKAGYVTDGRGRQKVVS
TETTNQKTQLQAIHLALQDSGSEVNIVTDSQYALGIIQAQPDKSESELVNQIIEQLIKKEKVYLSWVPAHKGIGGNEQV
DKLVSSGIRKVLFLDGIDKAQEEHEKYHSNWRAMASDENLPPIVAKEIVASCDKCQLKGEAMHGQVDCSPGIWQLCTH
LEGKVLVAVHVASGYIEAEVIPAETGQETAYFLLKLGRWPVKVIHTNGSNFTSAAVKAACWWAGIKQEFPIPNPQ
SQGVVSMNKKELKKIIGQVRDQAEHLKTAVQMAVFIHNFKRKGIGGYSAGERIIDIIATDIQTKELQKITKIQNFV
YYRDSRDPWKGPAKLLWKGEAVVIQDNNDIKVVPRRKAKIIRDYKQMGAGDDCVAGRQDEDQ

SEQ ID NO: 710

EpiGraph 1 консервативного pol ВИЧ М

MPQITLWQRPLVTIKIGGQLKEALLADDTVLEDINLPGKWPKMIGGIGGFIKVRQYDQILIEICGKKAIGT
VLVGPTPVNIIIGRNLITQIGCTLNFPISPIETVPVKLPGMDGPKVKQWPLTEEKIKALVEICTEMEKEGKISKIGPE
NPYNTPVFAIKKKDSTKWRKLVDFRELNKRTQDFWEVQLGIPHPAGLKKKSVTVLDVGDAYFSVPLDKDFRKYTAFT
IP SINNETPGIRYQYNVLPQGWKSPAIFQSSMTKILEPFRKQNPDIVIYQLYVGSDEIGQHRTKIEELRQHLLKWG
FTTPDKKHQKEPFLWMGYELHLPDKWTVQPIVLPKDSWTVNDIQKLVGKLNWASQIYHGQVDCSPGIWQLCTHLEGK

VILVAVHVASGYIEAEVIPAETGQETAYFLLKLAGRWPVKVIHTNGSNFTSAAVKAACWWAGIKQEFGIPYNPQSQGV
VSMNKLKKIIGQVRDQAEHLKTAVQMAVFIHNFRRGGIGGYSAGERIIDIIATDIQTKLQKQITKIQNFRVYYRDS
RDPWKGPAKLLWKGEAVVIQDNSEIKVVP RRKAKIIRDYGKQ MAGDDCVAGRQDEDEQ

SEQ ID NO: 711

EpiGraph 2 pol ВЧ М

MPFFREDLAFPGKAREFPSEQTRANSPTRGELQVWGGDNNSPSEAGADRQGTVSFSFPQITLWQRPLVSIKV
GGQIKEALLADDTVLEEMNLPGRWKPKMIGGIGGFIKVRQYDQIPIEICGHKAIGTVLIGTPVNIIGRNMLTQLGCT
LNFPISPIDTVPVTLKPGMDGPRVKQWPLTEEEKIKALTEICKEMEKEGKITKIGPENPYNTPIFAIKKDDSTKWRKLV
DFRELNNKTQDFWEVQLGIPHPAGLKKKKSVTVLVDVGDAYFSVPLDESFRKYTAFTIPSTNNETPGIRYQYNVLPQGW
KGS PAIFQCSMTKILEPFRIKNPEIVYQLYVGS DLEIGQHRAKIEELREHLLRWGFTTPDKKHQKEPPFLWMGYELH
PDRWTVQPIELPEKDSWTVNDIQKLVGKLNWASQIYAGIKVRQLCKLLRGTKALTEVVPLTEEALELAENREILKTP
VHGVIYDPSKDLVAEIQKQGGQWYQIYQEPYKNLKTGKYARKRSAHTNDVRQLTEVVQKIATESIWIWGKTPKFKL
PIQKETWEAWWTDYQATWIPDWEFVNTPLVLKWLQLEKDFIVGAETFYVDGAASRETKLGKAGYVTRNGRQKVVSL
TDTTNQKTLHAIHLALQDSGLEVNIVTDSQYALGIIQAQPPDRSESEVVNQIIEELIKKEKVYLAWVPAHKGIGGNEQV
DKLVSAGIRKVLFLDGDIDKAEHEHRYHSNWRTMASDFNLPPVVAKEIVANDCKCQLKGEAIIHGQVDCSPGMWQLCTH
LEGKIIILVAVHVASGYMEAEVIPAETGQETAYFLLKLAGRWPVKTIHTNGSNFTSTTVKAACWWAGIQQEFGIPYNPQ
SQGVVSMNKLKKIIGQVREQAEHLKTAVQMAVFIHNFRRGGIGGYSAGERIVDIIATDIQTRQLQKQIIKIQNFRV
YYRDSRDPWKGPAKLLWKGEAVVIQDNSEIKVVP RRKVKI IKDYGKQ MAGDDCVASRQDEDE

SEQ ID NO: 712

EpiGraph 2 консервативного pol ВЧ М

MPQITLWQRPLVSIKVGQIKEALLADDTVLEEMNLPGRWKPKMIGGIGGFIKVRQYDQIPIEICGHKAIGT
VLIGTPVNIIGRNMLTQLGCTLNFPISPIDTVPVTLKPGMDGPRVKQWPLTEEEKIKALTEICKEMEKEGKITKIGPE
NPYNTPIFAIKKDDSTKWRKLVDFRELNNKTQDFWEVQLGIPHPAGLKKKKSVTVLVDVGDAYFSVPLDESFRKYTAFT
IPSTNNETPGIRYQYNVLPQGWKGS PAIFQCSMTKILEPFRIKNPEIVYQLYVGS DLEIGQHRAKIEELREHLLRWG
FTTPDKKHQKEPPFLWMGYELHPDRWTVQPIELPEKDSWTVNDIQKLVGKLNWASQIYHGQVDCSPGMWQLCTHLEGK
IILVAVHVASGYMEAEVIPAETGQETAYFLLKLAGRWPVKTIHTNGSNFTSTTVKAACWWAGIQQEFGIPYNPQSQGV
VSMNKLKKIIGQVREQAEHLKTAVQMAVFIHNFRRGGIGGYSAGERIVDIIATDIQTRQLQKQIIKIQNFRVYYRDS
RDPWKGPAKLLWKGEAVVIQDNSEIKVVP RRKVKI IKDYGKQ MAGDDCVASRQDEDE

SEQ ID NO: 713

Гибрид gag/nef ВНО

MGARGSVLSGKKTDELEKVRRLRPGGRKKYMLKHIVWAARELDRFGSAESLLESKEGCQRILAVLAPLMPGTGS
EDKLSLFSTVCVWCLHAEMKVVDTEEAKKTVQSHLVVESGTAETMPAQSRPTAPPSGRGGNYPVQQIGGNYVHPLPS
PRTLNAWVKLIEKKFGEAVVPGFQALSEGCTPYDINQMLNCVGDHQAAMQIIRDIINEEAADWDLQHPQPAPQGGQL
REPSGSDIAGTTSSVDEQIQWMYRQONPIPVGNIYRRWIQLGLQKVRMYNPTNILDVKQGPKEFPQSYVDRFYKSLR
AEQTDAAVKNWMTQTLLIQANPDCKLVLKGLGVNPTLEEMLTACQGVGGPGQKARLMAEALKDALTGPPIFFAAVQ
RGQRKIKCWNCGKTHSARQCKAPRRKGCKGKAGHVMAKCPERQAGFLGFGPWGKKPHNFMAQMPQGLTPTAPP
ADPAVMDLKNYMKMKRQREKQRENREPYKEVSEDLHLSSLFGEQDQPGGATSKRRSKPSGDLRQKLLRARGENYGR
LWGELEDGSSQSLGGLGKGLSSRSCEGQKYSQGGFMNTPWKNPAEEKEKLPYRKQNI DDVDEEDNDLVGVSVRPKVPL
RTMSYKLAI DMSHFIEKGGLEGIYYSARRHRI LDILEKEEGIIPDWQDYTS GPGIRYPKTFGWLWKLVPVDMNSNEA
QEDDTHYLVHPAQTHQWSDPWGEVLVWKFDP LLAHTYEA FVRHPEEFGWKSGLPKEEVERRLAARGLLKMDKKETR

SEQ ID NO: 714

Консервативный gag/nef ВНО

MPVQQIGGNYVHPLSPRTLNAWVKLIEKKFGEAVVPGFQALSEGCTPYDINQMLNCVGDHQAAMQIIRDI
INEEAADWDLQHPQPAPQGGQLREPSGSDIAGTTSSVDEQIQWMYRQONPIPVGNIYRRWIQLGLQKVRMYNPTNILD
VKQGPKEFPQSYVDRFYKSLRAEQTDAAVKNWMTQTLLIQANPDCKLVLKGLGVNPTLEEMLTACQGVGGPGQKAR
LMVGVSVRPKVPLRTMSYKLAI DMSHFIEKGGLEGIYYSARRHRI LDILEKEEGIIPDWQDYTS GPGIRYPKTFGW
LWKLVP

SEQ ID NO: 715

Гибрид pol ВНО

MPFFRAWPMGKEASQFPHGPDASGADTNCS PRGSSCGSTEELHEVGQKAERKAEGEQRETLOGGNGGFAAPQF
SLWRRPVVTAHIEGQPEVLLADDSIVTGIELGPHYTPKIVGGIGGINTKEYKNVEIEVLGKRIKGTIMTGDTPINI
FGRNLLTALGMSLNFPIAKVEPVKVALKPGKDGPKLQWPLSKEKIVALREICEKMEKDQGLEEAPPTNPYNTPTFAI
KKKDNKWRMLIDFRELNRVTQDFTEVQLGIPHPAGLAKRKRTVLDIGDAYFSIPLDEEFRQYATFTLPSVNNAEFG
KRYIYKVLPPQGWKGS PAIFQYTMRHVLEPFRKANPDVTLVQILIASDRDLEHDRVVLQSKELLNSIGFSTPEEKFQK

DPPFQWMGYELWPTKWKLOKIELPQRETWTVNDIQKLVGLNWAAQIYPGKTKNLCKMIRGKMALTEGVQWTELAEEA
ELEENRIILNQEQEGRYYREDKPLEATVLKNQDNQWYTIHQGDRIKLVGKYAKVKNTHNGIRLLANVVQKIGKESI
VIWQQTPTFHLPVREVDQWWTDYQATWIPDWDFVSTPLIRLVFNLVKEPIEKEEVYIDGSCNRNSKEGKAGYV
TDRGKEKVLVLEQATNQALQAFLLALKDSGPKANIVTDSQYVLGIITGQPTESDSRIVAQIEQMIKKSEVYIGWVP
AHKGLGGNQEVDRIVSQEIRQVLFLESI EPAQEDHDKYHSNIKELAFKFGLPRLVAKQIVDTCKNCQKQGEATHGQAN
SDLGTMQCTHLEGGKIIIVAVHVASGFIEAEVIPAETGRQTALFLLKLAGRWPITHLHTNGANFASQEVKMAVWAGI
EHTFGVPYNPQSQGVVAMNHLKNQIDRIREQANSVETIVLMAVHGMNFKRRGGIGDMTPAERLINMITTEQEIQFQQ
SKNSKFKNFRVYREGDQLWKGPGEELLWKGEAVILKVGTDIKVPRRKAKIIKDYGGGKEVDSSSHMEDTGEAREV
A

SEQ ID NO: 716

Консервативный pol BИO

MPQFSLWRRPVVTAHIEGQPVLEVLLADDSIVTGIELGPHYTPKIVGGIGGFINTKEYKNVEIEVLGKRKIGT
IMTGDTPINIFGRNLLTALGMSLNFPIAKVEPVKVALKPGKDGPKLKQWPLSKEKIVALREICEKMEKDGQLEEAPPT
NPYNTPTFAIKKKDKNKWRMLIDFRELNRVTQDFTEVQLGIPHPAGLAKRRITVLDIGDAYFSIPLDEEFQYTAFT
LPSVNNAEPGKRYIYKVLPPQGWKGSFAIFQYTMHVLEPFRKANPDVTLVQILIASDRDLEHDRVVLQSKELLNSIG
FSTPEEKFKDPPFQWMGYELWPTKWKLOKIELPQRETWTVNDIQKLVGLNWAAQIYHGQANSDLGTWQMCTHLEGG
KIIIVAVHVASGFIEAEVIPAETGRQTALFLLKLAGRWPITHLHTNGANFASQEVKMAVWAGI EHTFGVPYNPQSQGV
VAMNHLKNQIDRIREQANSVETIVLMAVHGMNFKRRGGIGDMTPAERLINMITTEQEIQFQQSKNSKFKNFRVYRE
GRDQLWKGPGEELLWKGEAVILKVGTDIKVPRRKAKIIKDYGGGKEVDSSSHMEDTGEAREVA

SEQ ID NO: 717

EpiGraph 1 pol BИЧ B

FFRENLAFFQGGKAREFSSEQTRANSPTRRELQVWGRDNNSLSEAGADRQGTVSFSFPQITLWQRPLVTIKIG
GQLKEALLDTGADDTVLEEMNLPGRWKPKMIGGIGGFIVRQYDQIPIEICGHKAIGTVLVGTPVNIIGRNLLTQIG
CTLNFPISPIETVPVKLPGMDGPKVKQWPLTEEKIKALVEICTEMEKEGKISKIGPENPYNTPVFAIKKKDSTKWRK
LVDFRELNRKTQDFWEVQLGIPHPAGLKKKSVTVLDVGDAYFSVPLDKDFRKYTAFTIPSINNETPGIRYQYNVLPQ
GWKGSFAIFQSSMTKILEPFRKQNPDIYIYQYMDLTVGSDLEIGQHRTKIEELRQHLLRWGFTTPDKKHQKEPPFLW
MGYELHDPKWTVPQIVLPEKDSWTVNDIQKLVGKLNWASQIYAGIKVKQLCKLLRGTKALTEVVPLTEEALELAENR
EILKEPVHGVYYPQSKDLIAEIQKQGQGWYQIYQEPFKNLKTGKYARMGAHTNDVKQLTEAVQKIATESIYIWK
TPKFKLPQKETWEAWWTEYQATWIPWEFVNTPLVLKWLWYQLEKEPIVGAETFYVDGAANRETKLGKAGYVTDGR
QKVVSLDTNTQKTELQAIHLALQDSGLEVNIVTDSQYALGIIQAQPDKSESELVSQIIIEQLIKKEKVYLAWVPAHKG
IGNEQVDKLVSAGIRKVLFLDGDIDKAQEEHEKYSNWRMASDFNLPPVVAKEIVASCDKQQLKGEAMHGQVDCSPG
IWQLDCTHLEGGKIIIVAVHVASGYIEAEVIPAETGQETAYFLLKLAGRWPVKTITHDNGSNFTSTTVKAACWWAGIKQ
EFGIPYNPQSQGVVSMNKELKKIIGQVRDQAEHLKTAVQMAVFIHNFKRKGGIGGYSAGERIVDIIATDQIKELQK
QITKIQNFRVYRDSRDLWKGPAPKLLWKGEAVIQDNSDIKVVPRRKAKIIRDYGKQMGDDCVASRQDED

SEQ ID NO: 718

EpiGraph 1 gag BИЧ M

MGARASVLSGGKLDAAWEKIRLRPGGKKYRLKHLVWASRELERFALNPGLLETSEGCRQILGQLQPSLQTS
EELKSLYNTVATLYCVHQRIEVKDTKEALDKIEEQNKSKKKAQQAADTGNSSQVSQNYPIVQNLQGMVHQAI
SPRTLNAWVKVVEKAFAFSEVPMFSALESEGATPDNLNTMLNTVGGHQAAMQMLKETINEEAAEWDRLHPVHAGPI
APGQMRPRGSDIAGTTSTLQEQIGWMTNPNPIPVGEIYKRWIILGLNKIVRMYSVPVSIILDIRQGPKEPFRDYVDRFYKTLRA
EQATQEVKNWMTETLLVQANPDCKTILKALGPAATLEEMMTACQGVGGPGHKARVLAEMSQVTNSATIMMQRGNFK
GQKRIKCFNCGKEGHIARNCRAPRKKGCWKCQKEGHQMKDCTERQANFLGKIWPSSHKGRPGNLFQSRPEPTAPPEESF
RFGEEETTPSQKQEPIDKELYPLASLSLFGNDPLSQ

SEQ ID NO: 719

EpiGraph 1 консервативного gag BИЧ M

PIVQNLQGMVHQAI SPRTLNAWVKVVEKAFAFSEVPMFSALESEGATPDNLNTMLNTVGGHQAAMQMLKET
INEEAAEWDRLHPVHAGPIAPGQMRPRGSDIAGTTSTLQEQIGWMTNPNPIPVGEIYKRWIILGLNKIVRMYSVPVSI
LDIRQGPKEPFRDYVDRFYKTLRAEQATQEVKNWMTETLLVQANPDCKTILKALGPAATLEEMMTACQGVGGPGHKA
RVL

SEQ ID NO: 720

EpiGraph 2 gag BИЧ M

MGARASVLSGGELDRWEKIRLRPGGKKYKLVHVASRELERFAVNPGLLETAEGCQIIIEQLQSTLKTGS
EELKS LNTVAVLYCVHQRIDVKDTKEALEKIEEQNKSQKQQAAGTGSSSKVSQNYPIVQNAQGMVHQPLSPR
TLNAWVKVIEKAFAFSEVPMFTALSEGATPDNLNMLNIVGGHQAAMQMLKDTINEEAAEWDRVHPVHAGPIPPGQM

REPRGSDIAGTTSNLQEQIGWMTSNPPPIPVGDIYKRWIILGLNKIVRMYSPTSILDIKQGPKEPFRDYVDRFFKTLRA
EQASQEVKNWMTDTLLVQNANPDCKTILRALGPGATLEEMMTACQGVGGPSHKARVLAEAMSQVTNSATIMMQRGNFR
NQRKTVKFCNCGKEGHLARNCRAPRKKGWCWKGREGHQMKDCNERQANFLGKIWPSNKGPRGNFPQSRPEPTAPPAES
FRFEETTPAPKQEPKDRPLTSLKSLFGSDPLSQ

SEQ ID NO: 721

EpiGraph 2 консервативного gag ВЧ М

PIVQNAQQQMVFHQLSPRTLNAWVKVIEEKAFSPVIMFTALSEGATPQDLNMLNIVGGHQAAMQMLKDT
INEEAAEWDRVHPVHAGPIPPGQMREPRGSDIAGTTSNLQEQIGWMTSNPPPIPVGDIYKRWIILGLNKIVRMYSPTS
ILDIKQGPKEPFRDYVDRFFKTLRAEQASQEVKNWMTDTLLVQNANPDCKTILRALGPGATLEEMMTACQGVGGPSHKA
RVL

SEQ ID NO: 722

EpiGraph 1 nef ВЧ М

MGGKWSKSSIVGWFAVRERMRRAEPAAEGVGAVERDLEKHGAITSSNTAATNADCAWLEAQEEEEVGFPVRP
QVPLRPMTYKGAALDLSHFLKEKGGLEGLIYSKKRQEIIDLWVYHTQGYFPDWQNYTPGPGIRYPLTFGWCFKLVPVDP
REVVEANEGENNNCLLHPMSQHGMDPEKEVLMWKFDSRLAFHHMARELHPEYKDC

SEQ ID NO: 723

EpiGraph 1 консервативного nef ВЧ М

VGFPVRPQVPLRPMTYKGAALDLSHFLKEKGGLEGLIYSKKRQEIIDLWVYHTQGYFPDWQNYTPGPGIRYPL
TFGWCFKLVP

SEQ ID NO: 724

EpiGraph 2 nef ВЧ М

MGSKWSKSSIVGWFAIRERMRTEPAAEGVGAASRDLEKHGAITSSNTAANNADCAWLEAQEDEEVGFVPK
QVPLRPMTYKAAFDLSFFLKEKGGLDGLIYSQKRQDIIDLWVYHTQGYFPDWQNYTPGPGVRYPLTFGWCFKLVPVDP
EKVEEANEGENNNLLHPMSLHGMDPEKEVLMWKFDSRLARRHMAELHPEYKDC

SEQ ID NO: 725

EpiGraph 2 консервативного nef ВЧ М

VGFPVKPQVPLRPMTYKAAFDLSFFLKEKGGLDGLIYSQKRQDIIDLWVYHTQGYFPDWQNYTPGPGVRYPL
TFGWCFKLVP

SEQ ID NO: 726

EpiGraph 1 pol ВЧ М

FFRENLAFFPQGEAREFSSEQTRANSPTRRELQVWGRDNNLSLSEAGADRQGTVSFSFPQITLWQRPLVTIKIG
GQLKEALLDTGADDTVLEDINLPGKWKPKMIGGIGGFIKVRQYDQILIEICGKKAIGTVLVGPTPVNIIIGRNLLTQIG
CTLNFPISPIETVPVKLPKPGMDGPKVKQWPLTEEKIKALVEICTEMEKEGKISKIGPENPNYTPVFAIKKKDSTKWRK
LVDFRELNRKTQDFWEVQLGIPHPAGLKKKKSVTVLDVGDAYFSVPLDKDFRKYTAFTIPSINNETPGIRYQYNVLPQ
GWKGSPAIFQSSMTKILEPFRKQNPDIVIYQYMDLYVGSDEIGQHRTKIEELRQHLLKWGFTTPDKKHQKEPPFLW
MGYELHPDKWTVQPIVLPEKDSWTVNDIQKLVGKLNWASQIYPGKVKQLCKLLRGAKALTDIVPLTEEEAELEAENR
EILKEPVHGVYDPSKDLIAEIQKQGQDQWYQIYQEPFKNLKTGKYARMGAHTNDVKQLTEAVQKIATESIWIWVK
TPKFRPLIQKETWETWTEYQWATWIPWEFEVNTPLVLKWLWYQLEKEPIVGAETFFYVDGAANRETCLGKAGYVTDGR
QKVSLSLTETTNQKTELQAIHLALQDSGSEVNIIVTDSQYALGIIQAQPDKSESELVNQIIEQLIKKEKVYLSWVPAHKG
IGNEQVDKLVSSGIRKVLFLDGDKAQEEHEKYHSNWRAMASDFNLPPIVAKEIVASCDKQCLKGEAMHGQVDCSPG
IWQLDCTHLEGKILVAVHVASGYIEAEVIPAETGQETAYFLLKLGRWPVKVIHTDNGSNFTSAAVKAACWWAGIKQ
EFGIPYNPQSQGVVESMNKELKKIIGQVRDQAEHLKTAQVQMAVFIHNFKRKGGIGGYSAGERIIDIIATDIQTKELOK
QITKIQNFVRYRDSRDPWKGPAKLLWKGEAVVQDNSDIKVVPRRKAIIIRDYKQMGAGDDCVAGRQDEDQ

SEQ ID NO: 727

EpiGraph 1 консервативного pol ВЧ М

PQITLWQRPLVTIKIGGQLKEALLDTGADDTVLEDINLPGKWKPKMIGGIGGFIKVRQYDQILIEICGKKAIG
TVLVGPTPVNIIIGRNLLTQIGCTLNFPISPIETVPVKLPKPGMDGPKVKQWPLTEEKIKALVEICTEMEKEGKISKIG
PENPNYTPVFAIKKKDSTKWRKLVDFRELNRKTQDFWEVQLGIPHPAGLKKKKSVTVLDVGDAYFSVPLDKDFRKYTA
FTIPSINNETPGIRYQYNVLPQGWKGSPAIFQSSMTKILEPFRKQNPDIVIYQYMDLYVGSDEIGQHRTKIEELRQ
HLLKWGFTTPDKKHQKEPPFLWMGYELHPDKWTVQPIVLPEKDSWTVNDIQKLVGKLNWASQIYHGQVDCSPGWQLD
CTHLEGKILVAVHVASGYIEAEVIPAETGQETAYFLLKLGRWPVKVIHTDNGSNFTSAAVKAACWWAGIKQEFGIP

YNPQSQGVVESMNNELKKIIGQVRDQAEHLKTAVQMAVFIHNFKRGGIGGYSAGERIIDIIATDIQTRELQKQITKI
QNFVRVYRDSRDLWKGPAKLLWKGEAVVIQDNSDIKVVPRKAKIIRDYGKQMGDDCVAGRQDEDED

SEQ ID NO: 728

EpiGraph 2 pol ВМЧ М

FFREDLAFPPQKGAREFFSEQTRANSPTRGELQVWGGDNNSPSEAGADRQGTVSFSFPQITLWQRPLVSIKVG
GQIKEALLDTGADDTVLEEMNLPGRWKPKMIGGIGGFIKVRQYDQIPIEICGHKAIGTVLIGPTPVNIIIGRNMLTQLG
CTLNFPISPIDTVPVTLKPGMDGPRVKQWPLTEEKIKALTEICKEMEKEGKITKIGPENPYNTPIFAIKKKDSTKWRK
LVDFRELNNKKTQDFWEVQLGIPHPAGLKKKKSVTVLDVGDAYFSVPLDESFRKYTAFTIPSTNNETPGIRYQYNVLPQ
GWKGS PAIFQCSMTKILEPFRIKNPEI VIYQYMDDLVYGS DLEIGQHRAKIEELREHLLRWGFTTPDKKHQKEPPFLW
MGYELHPDRWTVQPIELPEKDSWTVNDIQKLVGKLNWASQIYAGIKVRQLCKLLRGTKALTEVVPLTEEALELAENR
EILKTPVHGYYDPSKDLVAEIQKQGQGWTYQIYQEPYKNLKTGKYARKRSAHTNDVRQLTEVVQKIATESI VIW GK
TPKFKLPIQKETWEAWWTDYQATWIPDWEFVNTPLVLKWLQLEKDPVGAETFYVDGAASRETKLGKAGYVTRNGR
QKVVSLTDTTQKTELHAIHLALQDSGLEVNIVTDSQYALGIIQAQPDSESEVVNQIIEELIKKEKVYLAWVPAHKG
IGGNEQVDKLVASGIRKVLFLDGDIDKAQEEHERYHSNWRMTASDFNLPVVAKEIVANC DKCQLKGEAIHQVDCSPG
MWQLDCTHLEGKIIILVAVHVASGYMEA EVI PAETGQETAYFIIKLKAGRWPVKTIHTDNGSNFTSTTVKACWWAGIQQ
EFGIPYNPQSQGVVESMNNELKKIIGQVREQAEHLKTAVQMAVFIHNFKRGGIGGYSAGERIVDIIATDIQTRELQKQ
I IKIQNFVRVYRDSRDLWKGPAKLLWKGEAVVIQDNSEIKVVPRKVKI IKDYGKQMGDDCVASRQDEDED

SEQ ID NO: 729

EpiGraph 2 консервативного pol ВМЧ М

PQITLWQRPLVSIKVGQIKEALLDTGADDTVLEEMNLPGRWKPKMIGGIGGFIKVRQYDQIPIEICGHKAI
GTVLIGPTPVNIIIGRNMLTQLGCTLNFPISPIDTVPVTLKPGMDGPRVKQWPLTEEKIKALTEICKEMEKEGKITKIG
PENPYNTPIFAIKKKDSTKWRKLVDFRELNNKKTQDFWEVQLGIPHPAGLKKKKSVTVLDVGDAYFSVPLDESFRKYTA
FTIPSTNNETPGIRYQYNVLPQGWKGS PAIFQCSMTKILEPFRIKNPEI VIYQYMDDLVYGS DLEIGQHRAKIEELRE
HLLRWGFTTPDKKHQKEPPFLWMGYELHPDRWTVQPIELPEKDSWTVNDIQKLVGKLNWASQIYHGQVDCSPGMWQLD
CTHLEGKIIILVAVHVASGYMEA EVI PAETGQETAYFIIKLKAGRWPVKTIHTDNGSNFTSTTVKACWWAGIQQEF
GIPYNPQSQGVVESMNNELKKIIGQVREQAEHLKTAVQMAVFIHNFKRGGIGGYSAGERIVDIIATDIQTRELQKQI
IKIQNFVRVYRDSRDLWKGPAKLLWKGEAVVIQDNSEIKVVPRKVKI IKDYGKQMGDDCVASRQDEDED

SEQ ID NO: 730

EpiGraph 1 gag ВМЧ B

MGARASVLSGGELDRWEKIRLRPGGKKKYRLKHIVWASRELERFAVNPGLETSEGCRQILGQLQPSLQTS
EELKSLYNTVATLYCVHQRIEVKDTKEALDKIEEEQNKS KKKQAQAAADTGNSSQVSQNYPIVQN LQGQMVHQAI
SPRTLNAWVKVVEEKA FSP E V I P M F S A L S E G A T P Q D L N T M L N T V G G H Q A A M Q M L K E T I N E E A A E W D R L H P V H A G P I A P G Q M
REPRGSDIAGTTSTLQEQIGWMTNNPPIPVGEIYKRWIIILGLNKIVRMYSPTSILDIRQGPKEPFRDYVDRFYKTLRA
EQASQEVKNWMTETLLVQNaNPDCKTILKALGPAATLEEMMTACQGVGGPGHKARVLAEAMSQVTNSATIMMQRGNFR
NQRKTVKCFNCGKEGHI AKNCRAPRKKGCWCKGKEGHQMKDC TERQANFLGKIWPSHKGRPGNFLQSRPEPTAPPEES
FRFGEETTTTPSQKQEPIDKELYPLASLSLFGNDPSSQ

SEQ ID NO: 731

EpiGraph 1 консервативного gag ВМЧ B

PIVQN LQGQMVHQAI SPRTLNAWVKVVEEKA FSP E V I P M F S A L S E G A T P Q D L N T M L N T V G G H Q A A M Q M L K E T I N E E A A E W D R L H P V H A G P I A P G Q M
REPRGSDIAGTTSTLQEQIGWMTNNPPIPVGEIYKRWIIILGLNKIVRMYSPTSILDIRQGPKEPFRDYVDRFYKTLRA
EQASQEVKNWMTETLLVQNaNPDCKTILKALGPAATLEEMMTACQGVGGPGHKARVLAEAMSQVTNSATIMMQRGNFR
NQRKTVKCFNCGKEGHI ARNCRAPRKKGCWCKGREGHQMKDCNERQANFLGKIWPSYKGRPGNFLQNRPEPTAPPAES
FRFGEETTTTPSQKQEPIDKELYPLASLSLFGNDPSSQ

SEQ ID NO: 732

EpiGraph 2 gag ВМЧ B

MGARASVLSGGKLDKWEKIRLRPGGKKKYKLKHIVWASRELERFALNPGLETSEGCKQILGQLQPALQTS
EELRSYNTVAVLYCVHQRIDVKDTKEALEKIEEEQNKS KKKQAQAAADTGNSSQVSQNYPIVQNMQGQMVHQPI
SPRTLNAWVKVIEEKA FSP E V I P M F T A L S E G A T P Q D L N T M L N T V G G H Q A A M Q I L K E T I N E E A A D W D R L H P V H A G P V A P G Q M
REPRGSDIAGTTSNLQEQIGWMTSNPPIPVGEIYKRWIIIMGLNKIVRMYSVPSILDIKQGPKEPFRDYVDRFYKVLRA
EQASQDVKNWMTETLLVQNSNPDCCKTILKALGPGATLEEMMTACQGVGGP SHKARILAEAMSQVTNSATIMMQKGNFR
NQRKIVKCFNCGKEGHI ARNCRAPRKKGCWCKGREGHQMKDCNERQANFLGKIWPSYKGRPGNFLQNRPEPTAPPAES
FRFGEETTTTPSQKQEPIDKELYPLASLSLFGNDPSSQ

SEQ ID NO: 733

EpiGraph 2 консервативного gag БИЧ В

PIVQNMQGMVHQPI SPRTLNAWVKVIEEKAFSPEVIMFTALSEGATPQDLNTMLNTVGGHQAAMQILKET
INEEAADWDRHLHPVHAGPVAPGQMREPRGSDIAGTTSNLQEQIGWMTSNPPIPVGEIYKRWIIMGLNKIVRMYSVPSI
LDIKQGPKEFFRDYVDRFYKVLRAEQASQDVKNWMTETLLVQNSNPDKTILKALGPGATLEEMMTACQGVGGPSHKA
RIL

SEQ ID NO: 734

EpiGraph 1 nef БИЧ В

MGGKWSKSSIVGWPAVRERMRRAEPAAEGVAVSRDLEKHGAITSSNTAATNADCAWLEAQEEEEVGFPVVRP
QVPLRPMTYKGALDLSHFLKEKGGLEGLIYSQKRQDILDWVYHTQGYFPDWQNYTPGPGIRYPLTFGWCFKLVPVEP
EKVEEANEENNSLLHPMSQHGMDPEKEVLMWKFD SRLAFHHMARELHP EYYKDC

SEQ ID NO: 735

EpiGraph 1 консервативного nef БИЧ В

GFPVRPQVPLRPMTYKGALDLSHFLKEKGGLEGLIYSQKRQDILDWVYHTQGYFPDWQNYTPGPGIRYPLT
FGWCFKLVP

SEQ ID NO: 736

EpiGraph 2 nef БИЧ В

MGGKWSKSSVVGWPAIRERMRRAEPADGVGAASRDLEKHGAITSSNTAANNAACAWLEAQEEDVEVGFPVKP
QVPLRPMTYKAAVDLSHFLKEKGGLEGLIHSQKRQEILDWVYHTQGFFPDWQNYTPGPGTRFPLTFGWCFKLVPVDP
DKVEEANEENNSLLHPMSLHGMDPEREVLVWKFD SRLAFHHVARELHP EYYKNC

SEQ ID NO: 737

EpiGraph 2 консервативного nef БИЧ В

GFPVKPQVPLRPMTYKAAVDLSHFLKEKGGLEGLIHSQKRQEILDWVYHTQGFFPDWQNYTPGPGTRFPLTFGWCFK
LVP

SEQ ID NO: 738

EpiGraph 1 pol БИЧ В

FFRENLAFFQKGAREFSSEQTRANSPTRRELQVWGRDNNSLSEAGADRQGTVSFSFPQITLWQRPLVTIKIG
GQLKEALLDTGADDTVLEEMNLPGRWKPKMIGGIGGFIKVRQYDQIPIEICGHKAI GTVLVGP TPVNI IGRNLLTQIG
CTLNFPISPIETVPVKLPGMDGPKVKQWPLTEEKIKALVEICTEMEKEGKISKIGPENPYNTPVFAIKKDKSTKWRK
LVDFRELNKRQTQDFWEVQLGIPHPAGLKKKKSVTVLDVGDAYFSVPLDKDFRKYTAFTIP SINNETPGIRYQYNVLPQ
GWKGSPAIFQSSMTKILEPFRKQNPDI VIYQYMDDLVVGSDLEIGQHRTKIEELRQHLLRWGFTTPDKKHQKEPPFLW
MGYELHPDKWTVQPIVLP EKDSWTVNDIQKLVGKLNWASQIYAGIKVKQLCKLLRGTKALTEVVPLTEEALELAENR
EILKEPVHGVYDPSKDLIAEIQKQGGQWQTYQIYQEPFKNLKTGKYARMGAHTNDVKQLTEAVQKIATESIWIWGK
TPKFKLPIQKETWEAWWTEYWQATWIP EWEFVNT PPLVKLWYQLEKEPIVGAETFYVDGAANRET KL GKAGYVTDGR
QKVSLTDTNTQKTQLAIHLALQDSGLEVINVTDSQYALGIIQAQPDKSESELVSQIIIEQLIKKEKVYLAWVPAHKG
IGGNEQVDKLVSAIRKVLFLDGIDKAQEEHEKYHSNWRAMASDFNLPPVVAKEIVASCDKCQLKGEAMHGQVDCSPG
IWQLDCTHLEGKIIILVAVHVASGYIEAEVIPAETGQETAYFLLKLAGRWPVKTIHTDNGSNFTSTTVKAACWWAGIKQ
EFGIPYNPQSQGVVESMNKELKKIIGQVRDQAHLKTAVQMAVFIHNFKRKGGIGGYSAGERIVDIATDIQTKELQK
QITKI QNFRVYRDSRDLWKGP AKLLWKGE GAVVIQD NSDIKVVPRR KAKIIRDYGKQ MAGDDC VASRQDED

SEQ ID NO: 739

EpiGraph 1 консервативного pol БИЧ В

PQITLWQRPLVTIKIGGQLKEALLDTGADDTVLEEMNLPGRWKPKMIGGIGGFIKVRQYDQIPIEICGHKAI
GTVLVGP TPVNI IGRNLLTQIGCTLNFPISPIETVPVKLPGMDGPKVKQWPLTEEKIKALVEICTEMEKEGKISKIG
PENPYNTPVFAIKKDKSTKWRKLVDFRELNKRQTQDFWEVQLGIPHPAGLKKKKSVTVLDVGDAYFSVPLDKDFRKYTA
FTIP SINNETPGIRYQYNVLPQGWKGSPAIFQSSMTKILEPFRKQNPDI VIYQYMDDLVVGSDLEIGQHRTKIEELRQ
HLLRWGFTTPDKKHQKEPPFLWMGYELHPDKWTVQPIVLP EKDSWTVNDIQKLVGKLNWASQIYHGQVDCSPGIWQLD
CTHLEGKIIILVAVHVASGYIEAEVIPAETGQETAYFLLKLAGRWPVKTIHTDNGSNFTSTTVKAACWWAGIKQEF
GIPYNPQSQGVVESMNKELKKIIGQVRDQAHLKTAVQMAVFIHNFKRKGGIGGYSAGERIVDIATDIQTKELQKQITKI
QNFRVYRDSRDLWKGP AKLLWKGE GAVVIQD NSDIKVVPRR KAKIIRDYGKQ MAGDDC VASRQDED

SEQ ID NO: 740

EpiGraph 2 pol БИЧ В

FFREDLAFFQGEAREFPSEQTRANSPTRGELQVWGGDNNSPSEAGADRQGTVSLSPQITLWQRPLVTIKVG
GQLKEALLDTGADDTVLEDMNLPWKWKPKMIGGIGGFIKVRQYDQIPIEICGHKAI GTVLVGP TPVNI IGRNLLTQIG

CTLNFPISPIDTVFVKLPKPGMDGPRVKQWPLTEEEKIKALIEICTEMEKEGKISRIGPENPYNTPIFAIKKKDSTKWRK
 LVDFRELNKKTQDFWEVQLGIPHPISGLKKKSVTVLDVGDAYFSVPLDKEFRKYTAFTIPSTNNETPGIRYQYNVLPQ
 GWKGSIPAIFQCSMTKILEPFRKQNPFIYQYMDLTVGSDLEIEQHRTKIEELRQHLLKWGFTTPDKKHQKEPPFLW
 MGYELHPDKWTVQPIMLPEKDSWTVNDIQKLVGKLNWASQIYPGIKVRQLCKLLRGAKALTEVIPLTEEALELAENR
 EILREPVHGVYDPTKDLIAEIQKQGLGQWYQIYQEPYKLNKTGKYARTGAHTNDVRQLTEAVQKITTESIVIWGK
 TPKFRLPIQKETWETWWTEYQATWIPWEFEVNTPLVLKLWYQLEKEPIIGAETFYVDGAANRDTKLKGAGYVTNKGK
 QKVVTLLDTTNQKTELQAIYALQDSGSEVNIVTDSQYALGIIQAQPDKSESELVNQIIIEQLIKKEKIYLAWVPAHKG
 IGGNEQVDKLVSSGIRKVLFLDGDIDKAQEDHEKYHSNWKAMASDFNLPPIVAKEIVASCDKQKLGAEIHGQVDCSPG
 IWQLDCTHLEGGVILVAVHVASGYIEAEVIPAETGQETAYFILKLAGRWPVKTVHTDNGSNFTSATVKAACWWAGVKQ
 EFGIPYNPQSQGVVESMNNELKKIIGQIRDQAEHLKTAVQMAVFIHNFKRKGGIGGYSAGERIIDIIATDIQTRELQK
 QITKIQNFRVYYRDNRPDLWKGPALKLLWKGEAVVIQDNSEIKVVPFRKVKIIRDYGKQMGADDCVAGRQDED

SEQ ID NO: 741

EpiGraph 2 консервативного pol ВУС В

PQITLWQRPLVTIKVGGQLKEALLDGTADDTVLEDMLPGKWKPKMIGGIGGFIKVRQYDQILIEICGHKAI
 GTVLIGPTFVNIIGRNLTLQGLCTLNFPISPIDTVFVKLPKPGMDGPRVKQWPLTEEEKIKALIEICTEMEKEGKISRIG
 PENPYNTPIFAIKKKDSTKWRKLVDFRELNKKTQDFWEVQLGIPHPISGLKKKSVTVLDVGDAYFSVPLDKEFRKYTA
 FTIPSTNNETPGIRYQYNVLPQGWKGSIPAIFQCSMTKILEPFRKQNPFIYQYMDLTVGSDLEIEQHRTKIEELRQ
 HLLKWGFTTPDKKHQKEPPFLWMGYELHPDKWTVQPIMLPEKDSWTVNDIQKLVGKLNWASQIYHGQVDCSPGIWQLD
 CTHLEGGVILVAVHVASGYIEAEVIPAETGQETAYFILKLAGRWPVKTVHTDNGSNFTSATVKAACWWAGVKQEF
 YNPQSQGVVESMNNELKKIIGQIRDQAEHLKTAVQMAVFIHNFKRKGGIGGYSAGERIIDIIATDIQTRELQKQITKI
 QNFRVYYRDNRPDLWKGPALKLLWKGEAVVIQDNSEIKVVPFRKVKIIRDYGKQMGADDCVAGRQDED

SEQ ID NO: 742

EpiGraph 1 gag ВУС С

MGARASILRGGKLDKWEKIRLRPGGKKHYMLKHLVWASRELERFALNPGLLETSEGCKQIMKQLPALQTGT
 EELRSYNTVATLYCVHEKIEVRDTKEALDKIEEEQNKSQKQKQAKAADGKVSQNYPIVQNLQGQMVHQAI
 SPRTLNAWVKVIEEKAFSPEVIMFTALSEGATPQDLNLTMLNTVGGHQAAMQMLKDTINEEAAEWDRHPVHAGPIAPGQ
 MREPRGSDIAGTTSTLQEQIAWMTSNPPIPVGDIYKRWIIILGLNKIVRMYSVPSILDIKQGPKEPFRDYVDRFFKTLRAEQ
 TQDVKNWMTDTLLVQANANPDCKTILRALGPGATLEEMMTACQGVGGPGHKARVLAEAMSQANSNIMMQRSNFKGPKRI
 VKCFNCGKEGHIARNCRAPRKKGCWKCQEGHQMCDCTERQANFLGKIWPSHKGRPGNFLQSRPEPTAPPAEPTAPPA
 ESFRFEETTPAPKQEPKDPREPLTSLKSLFGSDPLSQ

SEQ ID NO: 743

EpiGraph 1 консервативного gag ВУС С

PIVQNLQGQMVHQAI
 SPRTLNAWVKVIEEKAFSPEVIMFTALSEGATPQDLNLTMLNTVGGHQAAMQMLKDT
 INEEAAEWDRHPVHAGPIAPGQ
 MREPRGSDIAGTTSTLQEQIAWMTSNPPIPVGDIYKRWIIILGLNKIVRMYSVPSILDIKQGPKEPFRDYVDRFFKTLRAEQ
 TQDVKNWMTDTLLVQANANPDCKTILRALGPGATLEEMMTACQGVGGPGHKARVLAEAMSQANSNIMMQRSNFKGPKRI
 VKCFNCGKEGHIARNCRAPRKKGCWKCQEGHQMCDCTERQANFLGKIWPSHKGRPGNFLQSRPEPTAPPAEPTAPPA
 ESFRFEETTPAPKQEPKDPREPLTSLKSLFGSDPLSQ

SEQ ID NO: 744

EpiGraph 2 gag ВУС С

MGARASVLRGEKLDKWERIRLRPGGKKRYMLKHIVWASRELEKFAALNPGLLETAEGCKQIIKQLHPALQTGT
 EELKSLFNTVATLYCVHKKIDVDRDTKEALDKIEEEQNKCQKQKQAEAAADGKVSQNYPIVQNLQGQMVHQALS
 PRTLNAWVKVIEEKAFSPEIIPMFTALSEGATPTDLNLTMLNTVGGHQAAMQMLKDTINEEAAEWDRVHPVHAGPVAPGQ
 MREPRGSDIAGTTSTLQEQIAWMTNNPPIPVGEIYKRWIIILGLNKIVRMYSVPSILDIRQGPKEPFRDYVDRFFKTLRAEQ
 ATQEVKNWMTETLLVQANANPDCKTILRALGPGASLEEMMTACQGVGGPSHKARVLAEAMSQANNANIMMQRSNFKGSK
 RIVKCFNCGKEGHIAKNCRAPRKKGCWKCQEGHQMCDCTERQANFLGKIWPSNKGPRPGNFLQNRPEPTAPPAEPTAP
 PAESFKFEETTPAPKQESKDPREPLISLKSFLGNDPLSQ

SEQ ID NO: 745

EpiGraph 2 консервативного gag ВУС С

PIVQNLQGQMVHQALS
 PRTLNAWVKVIEEKAFSPEIIPMFTALSEGATPTDLNLTMLNTVGGHQAAMQMLKDT
 INEEAAEWDRVHPVHAGPVAPGQ
 MREPRGSDIAGTTSTLQEQIAWMTNNPPIPVGEIYKRWIIILGLNKIVRMYSVPSILDIRQGPKEPFRDYVDRFFKTLRAEQ
 ATQEVKNWMTETLLVQANANPDCKTILRALGPGASLEEMMTACQGVGGPSHKA
 RVL

SEQ ID NO: 746

EpiGraph 1 nef ВИЧ C

MGGKWSKSSIVGWPVAVRERIRRTEPAAEGVGAASQDLDDKYGALTSSNTAHNNADCAWLQAQEEEEEVGFVVR
PQVPLRPMYKAAFDLSFFLKEKGGLEGLIYSKKRQEILDLWVYHTQGFFPDWQNYTPGPGVRYPLTFGWCFKLVFVD
PREVEEANGENNCLLHPSQHGMEDEDREVLKWQFDSLSLARRHMARELHPEYKDC

SEQ ID NO: 747

EpiGraph 1 консервативного nef ВИЧ C

GFPVVRPQVPLRPMYKAAFDLSFFLKEKGGLEGLIYSKKRQEILDLWVYHTQGFFPDWQNYTPGPGVRYPLT
FGWCFKLV

SEQ ID NO: 748

EpiGraph 2 nef ВИЧ C

MGSKWSKSSIVGWPVAVRERMRAEPAAEGVGAASRDLDKHGALTSSNTPANNADCAWLEAQEEEEGEVGFVVK
PQVPLRPMYKGAFDLGGFLKEKGGDLGLIYSKKRQDILDLWVYNTQGYFPDWQNYTPGPGIRYPLTFGWICYKLVFVD
PSEVEEANGENNCLLHPSLHGMEDEHREVLKWKFDSSLSARRHLAREKHPEFYKDC

SEQ ID NO: 749

EpiGraph 2 консервативного nef ВИЧ C

GFPVKPQVPLRPMYKGAFDLGGFLKEKGGDLGLIYSKKRQDILDLWVYNTQGYFPDWQNYTPGPGIRYPLT
FGWICYKLV

SEQ ID NO: 750

EpiGraph 1 pol ВИЧ C

FFRENLAFFPQGEAREFPSEQTRANSPTSRSNSPTSRELQVRGDNPRSEAGAERQGTINFPQITLWQRPLVSI
KVGGOIKEALLDTGADDTVLEEINLPGKWKPKMIGGIGGFIKVRQYDQILIEICGKKAIGTVLVGPTPVNIIGRNMLT
QLGCTLNFPISPIETVPVKLKPMDGPKVKQWPLTEEEKIKALTAICEEMEKEGKITKIGPENPYNTPVFAIKKKDSTK
WRKLVDFRELNRKTQDFWEVQLGIPHPAGLKKKKSVTVLVDGDAYFSVPLDEGFRKYTAFTIPINNTPGIRYQYNV
LPQGWKGSFAIFQSSMTKILEPFRAQNPEIIVYQYMDDLVYVGSLEIGQHRAKIEELREHLLKWGFTTPDKKHQKEPP
FLWMGYELHPDKWTVPQIQLPEKDSWTVNDIQKLVGKLNWASQIYPGIVKVRQLCKLLRGAKALTDIVPLTEEALELA
ENREILKEPVHGVYDPSKDLIAEIQKQGHQWYTYQIYQEPFKNLKTGKYAKMRTAHTNDVKQLTEAVQKIAMESIVI
WGKTPKFRLPQKETWETWTDYQATWIPWEFEVNTPLVLKWLQYQLEKEPIAGAETFYVDGAANRETKIGKAGYVTD
RGRQKIVSLTETTNQKTELQAIQLALQDSGSEVNIVTDSQYALGIIQAQPDKSESELVNQIIIEQLIKKERVYLSWVPA
HKGIGGNEQVDKLVSSGIRKVLFLDGDIDKAQEEHEKYHSNWRAMASEFNLPPIVAKEIVASCDKQQLKGEAIGHQVDC
SPGIWQLDCTHLEGKIIILVAVHVASGYIEAEVIPAETGQETAYYILKLAGRWPVKVIHTDNGSNFTSAAVKAACWWAG
IQQEFGIPYNPQSQGVVESMNKELKKIIGQVRDQAEHLKTAVQMAVFIHNFKRKGGIGGYSAGERIIDIIATDIQTKE
LQKQIIKIQNFRVYRDSRDPWKGPAKLLWKGEAVVIQDNSDIKVVPRRKVKIKDYGKQMGADCVAGRQDEBQ

SEQ ID NO: 751

EpiGraph 1 консервативного pol ВИЧ C

PQITLWQRPLVSIKVGGOIKEALLDTGADDTVLEEINLPGKWKPKMIGGIGGFIKVRQYDQILIEICGKKAIGTVLVGPTPVNIIGRNMLTQLGCTLNFPISPIETVPVKLKPMDGPKVKQWPLTEEEKIKALTAICEEMEKEGKITKIGPENPYNTPVFAIKKKDSTKWRKLVDFRELNRKTQDFWEVQLGIPHPAGLKKKKSVTVLVDGDAYFSVPLDEGFRKYTAFTIPINNTPGIRYQYNVLPQGWKGSFAIFQSSMTKILEPFRAQNPEIIVYQYMDDLVYVGSLEIGQHRAKIEELREHLLKWGFTTPDKKHQKEPPFLWMGYELHPDKWTVPQIQLPEKDSWTVNDIQKLVGKLNWASQIYHGQVDCSPGIWQLDCTHLEGKIIILVAVHVASGYIEAEVIPAETGQETAYYILKLAGRWPVKVIHTDNGSNFTSAAVKAACWWAGIQQEFGIPYNPQSQGVVESMNKELKKIIGQVRDQAEHLKTAVQMAVFIHNFKRKGGIGGYSAGERIIDIIATDIQTKELOKQIIKIQNFRVYRDSRDPWKGPAKLLWKGEAVVIQDNSDIKVVPRRKVKIKDYGKQMGADCVAGRQDEBQ

SEQ ID NO: 752

EpiGraph 2 pol ВИЧ C

FFRENLAFFPQGEAREFPSEQARANSPTSRSNSPTSRELQVRGDNPCSEAGAERQGTINFPQITLWQRPLVTI
KVGGOIKEALLDTGADDTVLEDINLPGKWKPRMIGGIGGFIKVRQYDQIPIEICGKKAIGSVLVGPTPVNIIGRNLLT
QLGCTLNFPISPIETIPVKLKPMDGPRVKQWPLTEEKIKALTEICEEMEKEGKISKIGPENPYNTPIFAIKKKDSTK
WRKLVDFRELNRKTQDFWEVQLGIPHPAGLKKKKSVTVLVDGDAYFSVPLDESFRKYTAFTIPSTNNTPGIRYQYNV
LPQGWKGSFAIFQSSMTRILEPFRAKNPEIIVYQYMDDLVYVGSLEIEQHRAKIEELREHLLRWGFTTPDKKHQKEPP
FLWMGYELHPDKWTVPQIQLPEKESWTVNDIQKLVGKLNWASQIYPGIVKVLCKLLRGTKALTDIVPLTEEALELA
ENREILREPVGYYDPSKELIAEIQKQGDQWYTYQIYQEPFKNLKTGKYAKRRTAHTNDVRQLTEAVQKIALESIVI
WGKIPKFRLPQKETWEIWWTDYQATWIPDWEFVNTPLVLKWLQYQLEKEPIAGVETFYVDGAANRETKLGKAGYVTD
KGRQKIVTLTETTNQKAELOAIQLALQDSGPEVNIVTDSQYALGIIQAQPDKSESEIVNQIIIEQLINKERIYLSWVPA

HKGIGGNEQVDKLVSNIRKVLFLDGDIDKAQEEHEKYHNNWRAMASDFNLPPVVAKEIVASCDQCQLKGEAMHGQVDC
SPGIWQLDCTHLEGKVLVAVHVASGYMEAEVIPAETGQETAYFILKLAGRWPVKI IHTDNGSNFTSTAVKAACWWAG
IKQEFGIPYNPQSQGVVESMNKELKKIIGQVREQAEHLKTAVQMAVFIHNFKRRGGIGGYSAGERIIDIIASDIQTKE
LQKQITKIQNFVYRDSRDIWKGPAKLLWKGEAVVLQDNSDIKVVPRRKAKIIRDYGKQMAGDDCVAGRQDED

SEQ ID NO: 753

EpiGraph 2 консервативного pol ВИЧ C

PQITLWQRPLVTIKVGGQIKEALLDTGADDTVLEDINLPGKWKPRMIGGIGGFIKVRQYDQPIEICGKKAI
GSVLVGPTPVNIIGRNLLTQLGCTLNFPISPIETIPVKLPGMDGPRVKQWPLTEEKIKALTEICEEMEKEGKISKIG
PENPYNTPIFAIKKKDSTKWRKLVDFRELNKRTQDFWEVQLGIPHPAGLKKKKSVTVLVDVGDYFVSPLDESFRKYTA
FTIPSTNNETPGIRYQYNVLPQGWKGSPIFQSSMTRILEPFRKNPEIIVYQYMDLLYVGSLEIEQHRAKIEELRE
HLLRWGFTTPDKKHQKEPPFLWMGYELHPDKWTVQPIQLPEKESWTVNDIQKLVGKLNWASQIYHGQVDCSPGIWQLD
CTHLEGKVLVAVHVASGYMEAEVIPAETGQETAYFILKLAGRWPVKI IHTDNGSNFTSTAVKAACWWAGIKQEFGIP
YNPQSQGVVESMNKELKKIIGQVREQAEHLKTAVQMAVFIHNFKRRGGIGGYSAGERIIDIIASDIQTKE
LQKQITKIQNFVYRDSRDIWKGPAKLLWKGEAVVLQDNSDIKVVPRRKAKIIRDYGKQMAGDDCVAGRQDE

SEQ ID NO: 754

Эписенсус EG-0 Gag ВИЧ M, индивидуализированный

MGARASVLSGGKLDWEKIRLRPGGKKKYRLKHLVWASRELERFALNPGLLETSEGCRQILGQLQPSLQTGS
EELKSLYNTVATLYCVHQRIEVDKTEALDKIEEQNKSKKKAQQAADTGNSSQVSNYPIVQNLQGMVHQAI SPR
TLNAWVKVVEEKAFSPEVI PMFSALSEGATPQDLNMTMLNTVGGHQAAMQMLKETINEEAAEWDRLHPVHAGPIAPGQM
REPRGSDIAGTTSTLQEQIGWMTNPNPIPVGEIYKRWIILGLNKIVRMYSFVSILDIRQGPKEPFRDYVDRFYKTLRA
EQATQEVKNWMTETLLVQANPDCCKTILKALGPAATLEEMMTACQGVGGPGHKARVLAEAMSQVTNSATIMMQRGNFK
GQKRIKCFNCGKEGHIARNCRAPRKKGCWKCQKEGHQMKDCNERQANFLGKIWPSHKGRPGNFLQSRPEPTAQSRPEP
TAPPAESFRPQPTAPPEESFRFGEETTPSQKQEPIDKELYPLASLSLFGNDPLSQY

SEQ ID NO: 755

Эписенсус EG-0 Gag ВИЧ M, консервативный, индивидуализированный

PIVQNLQGMVHQAI SPRTLNAWVKVVEEKAFSPEVI PMFSALSEGATPQDLNMTMLNTVGGHQAAMQMLKET
INEEAAEWDRLHPVHAGPIAPGQMREPRGSDIAGTTSTLQEQIGWMTNPNPIPVGEIYKRWIILGLNKIVRMYSFVSI
LDIRQGPKEPFRDYVDRFYKTLRAEQATQEVKNWMTETLLVQANPDCCKTILKALGPAATLEEMMTACQGVGGPGHKA
RVL

SEQ ID NO: 756

ЦЕН-1 gag ВИЧ M, индивидуализированный

MGARASVLTGGKLDAWERIRLRPGGKKKYRMKHLVWASRELERFAINPGLLETAEGCQOIIEQLQSTLKTGS
EELKSLFNTVATLYCVHQRIEIKDTKEALDKLEEVQNKSSQKTQQAAGTGSSSKVSNYPIVQNAQGMVHQPLSPR
TLNAWVKVVEEKGFNPEVI PMFSALSDGATPQDLNMTMLNIVGGHQAAMQMLKDTINEEAAEWDRLHPVHAGPIPPGQM
REPRESDIAGTTSTLQEQIGWMTSNPPIPVGDIYKRWIILGLHKIVRMYSFVGILDIRQGPKEPFRDYVDRFFKTLRA
EQASQEVKNWMTETLLIQANPDCCKSILKALGTGATLEEMMTACQGVGGPSHKARVLAEAMSAQHANIMMQRGNFKG
QRKIKCFNCGKEGHIARNCRAPRKRGCWKCQEGHQMKNERQANFLGKIWPSNKGPRGNFPQSRPEPTAPRTEPTA
PPARPEPTAPPLQSRLEPTAPPAEPTAPPAENWGMGEIITSLLKQEQKDKEHPPPLVSLKSLFGNDPLSQ

SEQ ID NO: 757

ЦЕН-1 gag ВИЧ M, консервативный, индивидуализированный

PIVQNAQGMVHQPLSPRTLNAWVKVVEEKGFNPEVI PMFSALSDGATPQDLNMTMLNIVGGHQAAMQMLKDT
INEEAAEWDRLHPVHAGPIPPGQMREPRESDIAGTTSTLQEQIGWMTSNPPIPVGDIYKRWIILGLHKIVRMYSFVGIL
DIRQGPKEPFRDYVDRFFKTLRAEQASQEVKNWMTETLLIQANPDCCKSILKALGTGATLEEMMTACQGVGGPSHKA
RVL

SEQ ID NO: 758

ЦЕН-2 gag ВИЧ M, индивидуализированный

MGARASILRGGKLDKWEKIRLRPGGKKHYMLKHIWASRELEKFALNPDLLETSEGCKQIIKQLQPALQTGT
EELRSLFNTVATLYCVHEKIEVRDTKEALDKVEEQNKSSQKTQQAADGKVSQNYPIVQNAQGMVHQALSPTLN
AWVKVIEEKAFSPEVI PMFTALSEGATPSDLNMTMLNTIGGHQAAMQMLKDTINEEAAEWDRLHPVHAGPVAPGQMREP
RGSDIAGTSTLQEQIAWMTSNPPIPVGDIYKRWIILGLNKIVKMYSPVSI LDIRQGPKEPFRDYVDRFFKTLRAEQ
TQDVKNWMTDTLLVQANPDCCKTILRALGPGATLEEMMTACQGVGGPSHKARVLAEAMSAQSNIMMQRNSNFKGPKRI
VKCFNCGKEGHIARNCRAPRKKGCWKCQREGHQMKNERQANFLGKIWPSNKGPRGNFLQNRPEPTAPPLQSRLEPT
APLEPTAPPEPTAPPAVPTAPPVEPTAPPAEPTAPPAESFRFEETTPAPKQEPKDRPLTSLKSLFGSDPLSQ

SEQ ID NO: 759

ЦЕН-2 gag ВИЧ М, консервативный, индивидуализированный

PIVQNAQGQMVHQALSPRTLNAWVKVIEEKAFSPEVIMFTALSEGATPSDLNTMLNTIGGHQAAMQMLKDT
INEEAAEWDVRVHPVHAGPVAPGQMRPRGSDIAGTSTLQEQIAWMTSNPPIPVGDIYKRWII LGLNKIVKMYSPVSI
LDIKQGPKEPFRDYVDRFFKTLRAEQATQDVKNWMTDTLLVQANANPDCKTILRALGPGATLEEMMTACQGVGGPSHKA
RVL

SEQ ID NO: 760

ЦЕН-3 gag ВИЧ М, индивидуализированный

MGSRASVLSGGELDRWEKIRLRPGGKKKYKLKHIVWASRELERFAVNPGLLETSEGCKQILGQLQFALQTGS
EELRSLYNTVAVLYCVHQRIDVKDTKEALEKIEEEQNCKCKKAQAAAAADTGNNNSQVSQNYPIVQNIQGQMVHQALS
PRTLNAWVKVIEEKAFSPEVIMFTALSEGATPSDLNTMLNTIGGHQAAMQMLKDTINEEAAADWDLRHPVQAGPVAPG
QMRDPRGSDIAGTTSNLQEQIGWMTSNPPIPVGDIYKRWIIMGLNKIVRMYSPTSILDIKQGPKEPFRDYVDRFFKTL
RAEQASQEVKNWMTETLLVQNSNPDCCKTILKALGPGATLEEMMTACQGVGGPSHKA RVLAEAMSQATNSAAIMMQRGN
FRNQRKIVKCFNCGKEGHI AKNCRAPRKRCWCWKGREGHQMKDCNERQANFLGRWPSNKGPRGNFLQNRPEPTAPNF
LQSRPEPSAPPEPTAPPEESFRFGEETATPSQKQEPDKELYPLASLRSLFGNDPSSQ

SEQ ID NO: 761

ЦЕН-3 gag ВИЧ М, консервативный, индивидуализированный

PIVQNIQGQMVHQALSPRTLNAWVKVIEEKAFSPEVIMFTALSEGATPSDLNTMLNTIGGHQAAMQMLKDT
INEEAAADWDLRHPVQAGPVAPGQMRDPRGSDIAGTTSNLQEQIGWMTSNPPIPVGDIYKRWIIMGLNKIVRMYSPTS
LDIKQGPKEPFRDYVDRFFKTLRAEQASQEVKNWMTETLLVQNSNPDCCKTILKALGPGATLEEMMTACQGVGGPSHKA
RVL

SEQ ID NO: 762

ЦЕН-4 gag ВИЧ М, индивидуализированный

MGTRASVLSGGELDRWEKIRLRPGGKKKYKLKHIVWASRELERFAVNPGLLETAEGCRQILEQLQFALQTGS
EELRSLYNTVAVLYCVHQRIDVKDTKEALEKIEEEQNCKCKKAQQTAAADTGNNNSQVSQNYPIVQNMQGQMVHQPI SPR
TLNAWVKVIEEKAFSPEVIMFTALSEGATPHDLNTMLNTIGGHQAAMQMLKDTINEEAAEWDVRVHPVHAGPVAPGQ
MRPKGSDIAGTTSNLQEQIGWMTSNPPIPVGDIYKRWIIMGLNKIVRMYSPTSILDIKQGPKEPFRDYVDRFFKTLRA
EQASQDVKNWMTETLLVQNSNPDCCKTILKALGPGATLEEMMSACQGVGGPSHKA RILAEAMSQATNSANIMMQRGNFR
NQKRTVKCFNCGKEGHI AKNCRAPRKRCWCWKGREGHQMKDCNERQANFLGKIWPSSYKGRPGNFLQNRPEPTAPPEPT
APPEESFGFGEETTTTPQKQEPIDKDLYPLASLRSLFGNDPSSQ

SEQ ID NO: 763

ЦЕН-4 gag ВИЧ М, консервативный, индивидуализированный

PIVQNMQGQMVHQPI SPRTLNAWVKVIEEKAFSPEVIMFTALSEGATPHDLNTMLNTIGGHQAAMQMLKDT
INEEAAEWDVRVHPVHAGPVAPGQMRPKGSDIAGTTSNLQEQIGWMTSNPPIPVGDIYKRWIIMGLNKIVRMYSPTS
LDIKQGPKEPFRDYVDRFFKTLRAEQASQDVKNWMTETLLVQNSNPDCCKTILKALGPGATLEEMMSACQGVGGPSHKA
RIL

SEQ ID NO: 764

ЦЕН-5 gag ВИЧ М, индивидуализированный

MGARASILSGGKLDAWERIRLRPGGKKKYRMKHLVWASRELDRLFALNPSLLETAEGCQQIMEQLQFALKTGT
EELRSLFNTVATLYCVHQRIDVKDTKEALDKIEEI QNKSKQKTQQAADTGNNSSKVSQNYPIVQNAQGQMIHQSLSPR
TLNAWVKVIEEKAFSPEVIMFTALSEGATPQDLNMLNIVGGHQAAMQMLKDTINEEAAEWDVRVHPVHAGPIPPGQ
MRPRGSDIAGTTSNPQEQIGWMTSNPPIPVGDIYKRWII LGLHKLVRMYSVPVSI LDIKQGPKEPFRDYVDRFFKTLRA
EQATQEVKGWMTETLLIQNANPDCKSILRALGPGATLEEMMTACQGVGGPSHKA RVLAEAMSQVQHTNIMMQRGNFRG
QKRIKCFNCGKEGHLARNCRAPRKRCWCWKGREGHQMKDCNERQANFLGKIWPSSYKGRPGNFPQSRPEPTAPQNRLEP
TAPPAEPTAPPAEI FGMGEEITSPPKQEQKDREQAPPLVSLKSLFGNDLLSQ

SEQ ID NO: 765

ЦЕН-5 gag ВИЧ М, консервативный, индивидуализированный

PIVQNAQGQMIHQSLSPRTLNAWVKVIEEKAFSPEVIMFTALSEGATPQDLNMLNIVGGHQAAMQMLKDT
INEEAAEWDVRVHPVHAGPIPPGQMRPRGSDIAGTTSNPQEQIGWMTSNPPIPVGDIYKRWII LGLHKLVRMYSVPVSI
LDIKQGPKEPFRDYVDRFFKTLRAEQATQEVKGWMTETLLIQNANPDCKSILRALGPGATLEEMMTACQGVGGPSHKA
RVL

SEQ ID NO: 766

Эписенсус EG-0 gag ВИЧ С, индивидуализированный

MGARASILRGGKLDKWEKIRLRPGGKKHYMLKHLVWASRELERFALNPGLLTSEGCKQIMKQLQPALQTGT
EELRSLYNTVATLYCVHEKIEVRDTKEALDKIEEEQNKSQQTQQAADGKVSQNYPIVQNLQGQMVHQALSPRTL
AWVKVIEEKAFSPEVI PMFTALSEGATPQDLNTMLNTVGGHQAAMQMLKDTINEEAAEWDRLHPVHAGPVAPGQMRP
RGSDIAGTSTLQEQIAWMTSNPPIPVGDIYKRWIILGLNKIVRMYSVPSILDIKQGPKEFFRDYVDRFFKTLRAEQ
TQDVKNWMTDTLLVQANPDCKTILRALGPGATLEEMMTACQGVGGPGHKARVLAEAMSQANSNIMMQRSNFKGPKRI
VKCFNCGKEGHIARNCRAPRKKGCWKCGKEGHQMKDCSTERQANFLGKIWPSHKGRPGNFLQSRPEPTAPPAEPTAPPA
ESFRFEETTPAPKQEPKDRPLTSLKSLFGSDPLSQ

SEQ ID NO: 767

Эписенсус EG-0 gag ВИЧ С, консервативный, индивидуализированный

PIVQNLQGQMVHQALSPRTLNAWVKVIEEKAFSPEVI PMFTALSEGATPQDLNTMLNTVGGHQAAMQMLKDT
INEEAAEWDRLHPVHAGPVAPGQMRPGRSDIAGTSTLQEQIAWMTSNPPIPVGDIYKRWIILGLNKIVRMYSVPSI
LDIKQGPKEFFRDYVDRFFKTLRAEQATQDVKNWMTDTLLVQANPDCKTILRALGPGATLEEMMTACQGVGGPGHKA
RVL

SEQ ID NO: 768

ЦЕН-1 gag ВИЧ С, индивидуализированный

MGARASILRGEKLDKWEKIKLRPGGKKRYMLKHLI WASRELERFALNPSLLETSEGCKQIIKQLQPALKTGT
EELRSLFNTVATLYCVHAGIEVRDTKEALDRIEEEQNKCQQTQQAADGKVSQNYPIVQNAQGQMVHQALSPRTL
NAWVKVVEEKAFSPEIIPMFTALSEGATPSDLNSMLNTVGGHQAAMQMLKDTINEEAAEWDRVHPVHAGPIAPGQMR
PRGSDIAGTSTNLQEQIAWMTNNPPIPVGEIYKRWIILGLNKIVRMYSVPSILDIRQGPKEFFRDYVDRFFKCLRAEQ
ATQEVKNWMTDTLLIQANPDCKTILKALGPGASLEEMMTACQGVGGPSHKARVLAEAMSVNNANIMMQRGNFKGPK
RIIKCFNCGKEGHLARNCRAPRKKGCWKCGQEGHQMKDCSNERQANFLGKLWPSHKGGRPGNFLQNRPEPTAPPVEPT
APPAEPTAPPAESFKFEETTPVPKQELKDREPLISLKSFLGNDPLSQ

SEQ ID NO: 769

ЦЕН-1 gag ВИЧ С, консервативный, индивидуализированный

IVQNAQGQMVHQALSPRTLNAWVKVVEEKAFSPEIIPMFTALSEGATPSDLNSMLNTVGGHQAAMQMLKDTI
NDEAAEWDRVHPVHAGPIAPGQMRPGRSDIAGTSTNLQEQIAWMTNNPPIPVGEIYKRWIILGLNKIVRMYSVPSI
LDIRQGPKEFFRDYVDRFFKCLRAEQATQEVKNWMTDTLLIQANPDCKTILKALGPGASLEEMMTACQGVGGPSHKAR
VL

SEQ ID NO: 770

ЦЕН-2 gag ВИЧ С, индивидуализированный

MGARASILRGEKLDKWERIRLRPGGKKHYMIKHLVWASRELEKFAFNPGLLTADGCKQIIKQLHPALQTGT
EELKSLYNTVATLYCVHERIEVRDTKEALDRIEEEQNKCQQTQQAADGKVSQNYPIVQNAQGQMVHQALSPRTL
NAWVKVVEEKAFSPEIIPMFTALSEGATPTDLNTMLNTIGGHQAAMQILKDTINEEAEWDRLHPVQAGPVAPGQIRE
PRGSDIAGTSTNLQEQIAWMTGNPPVPVGDYKRWIIMGLNKIVRMYSVPSILDIRQGPKEFFRDYVDRFFKVLRAEQ
ATQEVKNWMTETLLVQANPDCKIILKGLGPAATLEEMMTACQGVGGPSHKARVLAEAMSVNNNTNIMMQKSNFKGPK
RTVKCFNCGKEGHIAKNCRAPRKKGCWKCGREGHQMKDCNERQANFLGKIWPSQKGRPGNFLQNRPEPTAPRPEPTAP
PRLEPTAPPAEPSAPPAESFRFEGTTPAPKQESKDREPLISLKSFLGNDPLSQ

SEQ ID NO: 771

ЦЕН-2 gag ВИЧ С, консервативный, индивидуализированный

PIVQNAQGQMVHQALSPRTLNAWVKVVEEKAFSPEIIPMFTALSEGATPTDLNTMLNTIGGHQAAMQILKDT
INEEAEWDRLHPVQAGPVAPGQIREPRGSDIAGTSTNLQEQIAWMTGNPPVPVGDYKRWIIMGLNKIVRMYSVPSI
LDIRQGPKEFFRDYVDRFFKVLRAEQATQEVKNWMTETLLVQANPDCKIILKGLGPAATLEEMMTACQGVGGPSHKA
RVL

SEQ ID NO: 772

ЦЕН-3 gag ВИЧ С, индивидуализированный

MGARASILRGEKLDKWERIRLRPGGKKCYMLKHIVWASRELERFSLNPGLLTAEAGCKQIIKQLHPALQTGT
EELKSIFYNTVATLYCVHKKIDVRDTKEALDKVEEQNKQQTQQAADGKVSQNYPIVQNIQGQMVHQALSPRTL
AWVKVVEEKAFSPEIIPMFTALSEEATPQDLNTMLNAVGGHQAAMQMLKETINEEAEWDRVHPVHAGPIAPGQMRP
RGSDIAGTSTNLQEQIAWMTNNPPIPVGEIYKRWIILGLNKIVRMYSVPSILDIRQGPKEFFRDYVDRFFKVLRAEQ
TQEVKNWMTDTLLIQANPDCKTILKALGPGASLEEMMTACQGVGGPSHKARVLAEAMSVANNINIMMQRGNFKGPKR

TVKCFNCGKEGHIKNCRAPRKRCGCKGEGHQMDCNERQANFLGRIWPSHKGRPGNLFQNRPEPTAPSAESFRQN
RTEPTAPPARLEPTAPPAEPSAPPVESFRFEETTPALKQESKDREPLTSLRSLFGSDPLFQ

SEQ ID NO: 773

ЦЕН-3 gag ВИЧ С, консервативный, индивидуализированный

PIVQNIQGQMVHQALSPTLNAAWVKVVEEKAFSP E I I PMFTALSEEATPQDLNTMLNAVGGHQAAMQMLKET
INEEAAEWDRVHPVHAGPIAPGQMREPRGSDIAGTTSNLQEQIAWMTNNPPIPVGEIYKRWII LGLNKIVRMYSFVSI
LDIRQGPKEPFRDYVDRFVRVLAERQATQEVKNWMTDTLLIQNANPDCKTILKALGPASLEEMMTACQGVGGPSHKA
RVL

SEQ ID NO: 774

ЦЕН-4 gag ВИЧ С, индивидуализированный

MGARASVLRGEKLDKWERIRLRPGGKKRYMLKHIVWASRELEKFALNPGLLETAEGCKQIIKQLHPALQGTG
EELRSLFNTVATLYCVHGKIDVDRDTEALDKVEEQNKCCQKQQAADKKVSQNYPIVQNIQGQMVHQPLSPRTLNA
WVKVVEEKAFSP E I I PMFSALSEGATPGDLNTMLNTIGGHQAAMQMLKDTINDEAAEWDRVHPVHAGPIAPGQMREPR
GSDIAGTTSNLQEQIAWMTNNPPIPVGEIYKRWIVLGLNKIVRMYSFVSI LDIRQGPKEPFRDYVDRFFRTLRAEQAT
QEVKNWMTETLLVQANANPDCKNILRALGPASLEEMMTACQGVGGPSHKA RVLAEAMSQANNTNIMMQKSNFKGPRRI
VKCFNCGKEGHIKNCRAPRKRCGCKGEGHQMDCNERQANFLGKIWP SNKGRPGNLFQNRPEPTAPQSRPEPTAP
LEPTAPPAEPTAPPAESFKFEETTPAPKQEQKDREPLISLKS LFGNDPLSQ

SEQ ID NO: 775

ЦЕН-4 gag ВИЧ С, консервативный, индивидуализированный

PIVQNIQGQMVHQPLSPRTLNAAWVKVVEEKAFSP E I I PMFSALSEGATPGDLNTMLNTIGGHQAAMQMLKDT
INDEAAEWDRVHPVHAGPIAPGQMREPRGSDIAGTTSNLQEQIAWMTNNPPIPVGEIYKRWIVLGLNKIVRMYSFVSI
LDIRQGPKEPFRDYVDRFFRTLRAEQATQEVKNWMTETLLVQANANPDCKNILRALGPASLEEMMTACQGVGGPSHKA
RVL

SEQ ID NO: 776

ЦЕН-5 gag ВИЧ С, индивидуализированный

MGARASVLRGEKLDKWERIRLRPGGKKRYMLKHIVWASRELEKFALNPGLLETAEGCKQIIKQLQPALQGTG
EELKSLFNTVATLYCVHEKIDVDRDTEALDRIEEQNKCCQKQQAADKKVSQNYPIVQNAQGQMVHQALSPTLN
AAWVKVIEEKGFNP E I I PMFTALSDGATPQDLNSMLNTVGGHQAAMQILKDTINDEAAEWDRVHPVHAGPIAPGQMREP
RGSDIAGTTSNLQEQIAWMTGNPPIPVGEIYKRWII LGLNKIVRMYSFVSI LDIRQGPKEPFRDYVDRFFKALRAEQAT
QEVKNWMTETLLVQANANPDCKNILRALGPASLEEMMTACQGVGGPSHKA RVLAEAMSQANNTNIMQNNFKGPKR
I IKCFNCGKEGHIKNCRAPRKRCGCKGEGHQMDCNERQANFLGRIWPSHKGRPGNLFQNRPEPTAPPVEPTAP
PAEPTAPPAESFKFEETTPTPKQEQKDREPLISLKS LFGNDPLSQ

SEQ ID NO: 777

ЦЕН-5 gag ВИЧ С, консервативный, индивидуализированный

PIVQNAQGQMVHQALSPTLNAAWVKVIEEKGFNP E I I PMFTALSDGATPQDLNSMLNTVGGHQAAMQILKDT
INEEAAEWDRVHPVHAGPIAPGQMREPRGSDIAGTTSNLQEQIAWMTGNPPIPVGEIYKRWII LGLNKIVRMYSFVSI
LDIRQGPKEPFRDYVDRFFKALRAEQATQEVKNWMTETLLVQANANPDCKNILRALGPASLEEMMTACQGVGGPSHKA
RVL

SEQ ID NO: 778

Эписенсус EG-0 gag ВИЧ В, индивидуализированный

MGARASVLSGGEIDRWEKIRLRPGGKKKYKLKHIVWASRELERFAVNPGLLETSEGCRQIILGQLQPSLQGTGS
EELKSLYNTVATLYCVHQKIDVVDTEALDKIEEQNKSKKKAQQAADTGNSSQVSQNYPIVQNLQGQMVHQAI SPR
TLNAAWVKVVEEKAFSP E I I PMFSALSEGATPQDLNTMLNTVGGHQAAMQMLKET INEEAAEWDRVHPVHAGPIAPGQM
REPRGSDIAGTTSNLQEQIGWMTNNPPIPVGEIYKRWII LGLNKIVRMYSPTS I LDIRQGPKEPFRDYVDRFYKTLRA
EQASQEVKNWMTETLLVQANANPDCKTILKALGPAATLEEMMTACQGVGGPGHKARVLAEAMSQVNTSATIMMQGNFR
NQRKTVKCFNCGKEGHIARNCRAPRKRCGCKGEGHQMDCNERQANFLGKIWP SHKGRPGNLFQSRPEPTAPPEPT
APPEESFRFGEETTPSQKQEPIDKELYPLASLKS LFGNDPSSQ

SEQ ID NO: 779

Эписенсус EG-0 gag ВИЧ В, консервативный, индивидуализированный

PIVQNLQGQMVHQAI SPRTLNAAWVKVVEEKAFSP E I I PMFSALSEGATPQDLNTMLNTVGGHQAAMQMLKET
INEEAAEWDRVHPVHAGPIAPGQMREPRGSDIAGTTSNLQEQIGWMTNNPPIPVGEIYKRWII LGLNKIVRMYSPTS I

LDIRQGPKEPFRDYVDRFYKTLRAEQASQEVKNWMTETLLVQANPDCKTILKALGPAATLEEMMTACQGVGGPGHKA
RVL

SEQ ID NO: 780

ЦЕН-1 gag ВИЧ В, индивидуализированный

MGSRASVLSGGKLDQWEKIRLRPGGKKRYRLKHLVWASRELERFAVNPSLLETSEGCKQIILGQLPALQTGS
EELRSLYNTIAVLYCVHQRIEVKDTKEALEKIEEEQNKCKKKAQAAAAAADTGNNSQVSQNYPIVQNMQGMVHQAL
SPRTLNAWVKVIEEKAFSPEVIMFTALSEGATPADLNTMLNTIGGHQAAMQILKETINEEAAEWDVRVHPVHAGPVAP
GQMREPRGSDIAGSTSTLQEQIAWMTSNPPIPVGDIYKRWIIMGLNKIVRMYSPTSILDIKQGPKEPFRDYVDRFYKT
LRAEQATQEVKNWMTETLLVQNSNPDCCKTILKALGPGATLEEMMTACQGVGGPSHKARVLAEAMSQVTTTPAIMMQRG
NFKNQKRVKCFNCGKEGHLARNCRAPRKRGCWKCGREGHQMKDCSERQANFLGKIWPSYKGRPGNFLQNRPEPTAPP
AEPTAPPAESFRFGEETATPPQKQEPIDKEMYPLTSLRSLFGNDPSQ

SEQ ID NO: 781

ЦЕН-1 gag ВИЧ В, консервативный, индивидуализированный

PIVQNMQGMVHQALSPTLNWVKVIEEKAFSPEVIMFTALSEGATPADLNTMLNTIGGHQAAMQILKET
INEEAAEWDVRVHPVHAGPVAPGQMREPRGSDIAGSTSTLQEQIAWMTSNPPIPVGDIYKRWIIMGLNKIVRMYSPTSIL
LDIKQGPKEPFRDYVDRFYKTLRAEQATQEVKNWMTETLLVQNSNPDCCKTILKALGPGATLEEMMTACQGVGGPSHKA
RVL

SEQ ID NO: 782

ЦЕН-2 gag ВИЧ В, индивидуализированный

MGSRASVLSGGKLDKWEKIRLRPGGKKRYRLKHLVWASRELERYALNPGILLETAEGCRQIILGQLPALQTGS
EELKSLFNTVAVLYCVHQRIDVKTKEALEKIEEEQNKSKKKTQAAAAADTGNNSQVKVSQNYPIVQNIQGMVHQALS
PRTLNAWVKVIEEKAFSPEVIMFTALSEGATPHDLNTMLNTIGGHQAAMQILKETINDEAAEWDRTHPVHAGPVAPG
QMRDPRGSDIAGTTSNLQEQIGWMTHNPPIPVGDIYKRWIIMGLNKIVRMYSPTSILDIKQGPKEPFRDYVDRFYKIL
RAEQASQDVKNWMTETLLVQNSNPDCCKTILKALGPGATLEEMMTACQGVGGPSHKARVLAEAMSQMTNSATIMMQGN
FRNQKRTIKCFNCGKEGHLARNCRAPRKRGCWKCGQEGHQMKDCNERQANFLGKIWPSYKGRPGNFLQSRPESRPEPT
APPAEPTAPPAESFRFGEETATPPQKQEPIDKEMYPLASLRSFLGNDPSSK

SEQ ID NO: 783

ЦЕН-2 gag ВИЧ В, консервативный, индивидуализированный

PIVQNIQGMVHQALSPTLNWVKVIEEKAFSPEVIMFTALSEGATPHDLNTMLNTIGGHQAAMQILKET
INDEAAEWDRTHPVHAGPVAPGQMRDPRGSDIAGTTSNLQEQIGWMTHNPPIPVGDIYKRWIIMGLNKIVRMYSPTSIL
LDIKQGPKEPFRDYVDRFYKILRAEQASQDVKNWMTETLLVQNSNPDCCKTILKALGPGATLEEMMTACQGVGGPSHKA
RVL

SEQ ID NO: 784

ЦЕН-3 gag ВИЧ В, индивидуализированный

MGARASILSGGELDKWEKIRLRPGGKKQYRLKHIVWASRELERFALNPGILLETSGGCRQIILEQLPALQTGS
EELRSLYNTVAVLYCVHQRIEVKDTKEALEKIEEEQNKCKKKAQAAAAAADTGNNSQVSQNYPIVQNIQGMVHQAL
SPRTLNAWVKVIEEKAFSPEVIMFTALSEGATPGDLNLMLNAVGGHQAAMQMLKDTINEEAADWDRHLHPVQAGPVAP
GQLREPRGSDIAGTTSNLQEQIAWMTSNPPIPVGEIYKRWIIMGLNKIVRMYSPTSILDIKQGPKEPFRDYVDRFYRT
LRAEQASQDVKNWMTETLLIQANPDCCRTILKALGPGATLEEMMTACQGVGGPSHKARVLAEAMSQATSSATIMMQG
NFRNQKRVKCFNCGKEGHIKNCRAPRKRGCWKCGREGHQMKDCSERQANFLGKIWPSYKGRPGNFLQNRPEPTAPP
AEPTAPPAESFRFGEETATPPQKQEPIDKELYPLASLRSFLGNDPLSQ

SEQ ID NO: 785

ЦЕН-3 gag ВИЧ В, консервативный, индивидуализированный

PIVQNIQGMVHQALSPTLNWVKVIEEKAFSPEVIMFTALSEGATPGDLNLMLNAVGGHQAAMQMLKDT
INEEAADWDRHLHPVQAGPVAPGQLREPRGSDIAGTTSNLQEQIAWMTSNPPIPVGEIYKRWIIMGLNKIVRMYSPTSIL
LDIKQGPKEPFRDYVDRFYRTLRAEQASQDVKNWMTETLLIQANPDCCRTILKALGPGATLEEMMTACQGVGGPSHKA
RVL

SEQ ID NO: 786

ЦЕН-4 gag ВИЧ В, индивидуализированный

MGARASILSGGELDKWEKIRLRPGGKKRYRLKHIVWASNELEFALNPGILLETS DGCRQIILGQLHPSLQTGS
EELRSLYNTVAVLYCVHQRIEIKDTKEALEKIEEEQNKCKKKAQAAAAQAAAGTGNNSQVSQNYPIVQNMQGMVHQAL
SPRTLNAWVKVIEEKAFSPEVIMFTALSEGATPHDLNTMLNTIGGHQAAMQILKETINEEAAEWDVRVHPVHAGP

IAPGQIREPRGSDIAGTTSNLQEQIGWMTHNPPIPVGEIYKKWIIMGLNKIVRMYSVPSILDIKQGPKEPFRDYVDRF
YKTLRAEQATQEVKNWMTETLLVQNSNPDCCKTILKALGPGATLEEMMTACQGVGGPAHKARVLAEAMSQATNSAAIMM
QKGNFRNQRRTVKCFNCGKEGHIKNCRAPRKKGCWKCQGEHQMKDCNERQANFLGRSWPSLKGPRGNFLQNRPEPS
APPEESFKFGEETTPPQKQEPIDKDLPLASLRSFLGNDPSST

SEQ ID NO: 787

ЦЕН-4 gag ВИЧ В, консервативный, индивидуализированный

PIVQNMQGGQMVHQALSPRTLNAWVKVIEEKAFSPEVIMFTALSEGATPHDLNLTMLNTIGGHQAAMQILKET
INEEAAEWDRVHPVHAGPIAPGQIREPRGSDIAGTTSNLQEQIGWMTHNPPIPVGEIYKKWIIMGLNKIVRMYSVPSI
LDIKQGPKEPFRDYVDRFYKTLRAEQATQEVKNWMTETLLVQNSNPDCCKTILKALGPGATLEEMMTACQGVGGPAHKA
RVL

SEQ ID NO: 788

ЦЕН-5 gag ВИЧ В, индивидуализированный

MGSRASVLSGGKLDKWEKIRLRPGGKKKYQLKHIVWASRELEERYALNPGLLETAEGCRQILEQLQPALQTGS
EELRSLYNTVAVLYCVHQKIEVKDTKEALEKVEEQNKSKKRIQQAQQAADTGNSSKVSQNYPIVRNLQGQMVHQP
ISPRTLNAWVKVIEEKAFSPEVIMFSALEAGATPQDLNMLNAVGGHQAAMQMLKDTINEEAAEWDRMHPVHAGPVA
PGQMREPRGSDIAGTTSNLQEQIGWMTSNPPIPVGEIYKRWIIMGLNKIVRMYSVPSILDIKQGPKEPFRDYVDRFYR
TLRAEQASQDVKNWMTETLLVQNSNPDCCKTILKALGPGATLEEMMTACQGVGGPSHKARVLAEAMSQVTNSTAIMQR
GNFKNQRRKIVKCFNCGKEGHIKNCRAPRKRGCWKCGRGEGHQMECTERQVNFGLGIWPSYKGRPGNFLQNRPEPTAP
PAPPEESFRFGEETTPSQKQGTIDKELYPLTSLRSFLGNDPS

SEQ ID NO: 789

ЦЕН-5 gag ВИЧ В, консервативный, индивидуализированный

PIVRNLQGQMVHQPISPRTLNAWVKVIEEKAFSPEVIMFSALEAGATPQDLNMLNAVGGHQAAMQMLKDT
INEEAAEWDRMHPVHAGPVAPGQMREPRGSDIAGTTSNLQEQIGWMTSNPPIPVGEIYKRWIIMGLNKIVRMYSVPSI
LDIKQGPKEPFRDYVDRFYRTLRAEQASQDVKNWMTETLLVQNSNPDCCKTILKALGPGATLEEMMTACQGVGGPSHKA
RVL

SEQ ID NO: 790

VTSSNMNNA

SEQ ID NO: 791

TSSNMNNAD

SEQ ID NO: 792

TSSNMNNADSVWLRAQEEE

SEQ ID NO: 793

TSSNMNNADCVWLRAQEEE

SEQ ID NO: 794

ARCHSLM

SEQ ID NO: 795

DEFGKLM

SEQ ID NO: 796

EPTAPPAEPTAP

SEQ ID NO: 797

PTAPPAEPTAPP

SEQ ID NO: 798

EPTAPPAEPTAPP

SEQ ID NO: 799

ARCGSLM

SEQ ID NO: 800
ARCGSPM

SEQ ID NO: 801
ARYGSLM

SEQ ID NO: 802
AYCHSLM

SEQ ID NO: 803
YRCHSLM

SEQ ID NO: 804
DEFGSLM

SEQ ID NO: 805
DEFGKLM

SEQ ID NO: 806
ARCCDEGH

SEQ ID NO: 807
ARCDEFGH

SEQ ID NO: 808
ARCCDE-GH

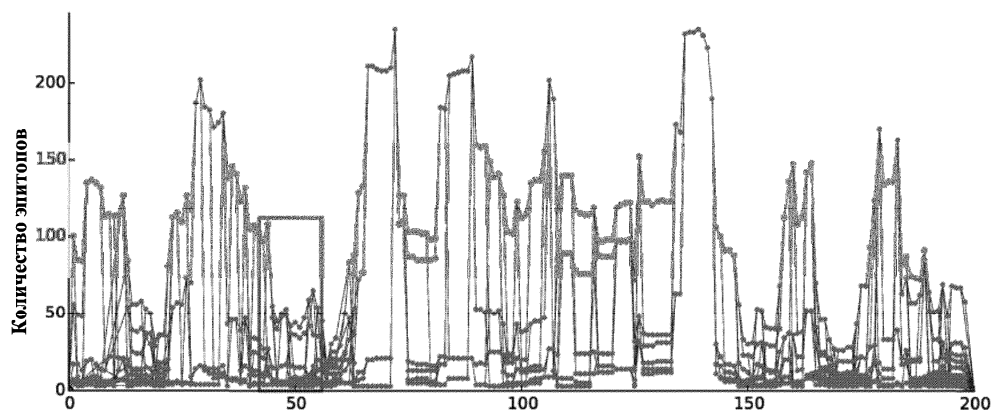
SEQ ID NO: 809
ARC-DEFGH

SEQ ID NO: 810
DECHJLM

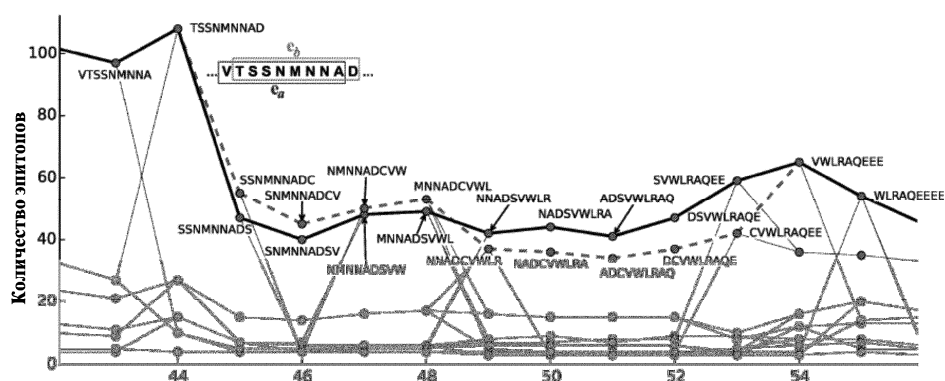
SEQ ID NO: 811
DEFGJLM

ФОРМУЛА ИЗОБРЕТЕНИЯ

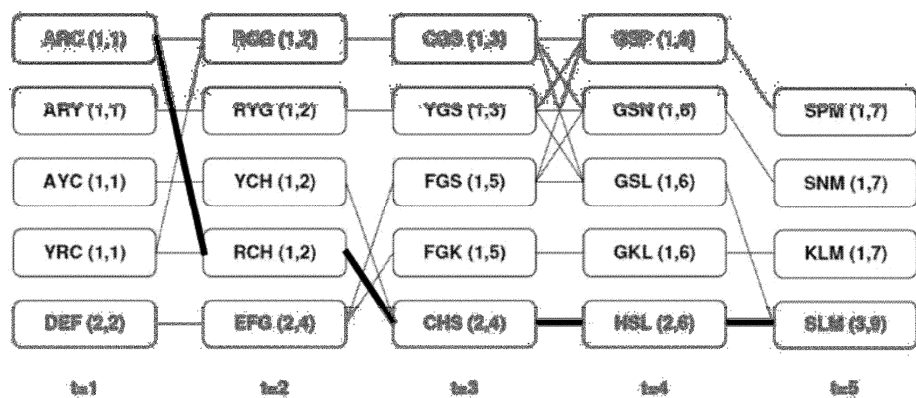
1. Полипептид, содержащий аминокислотную последовательность SEQ ID NO: 706.
2. Полипептид по п.1, где указанный полипептид содержит один или более эписенсусных антигенов для популяции ВИЧ-1, определенных с применением подхода EpiGraph.
3. Фармацевтическая композиция для индукции иммунного ответа против ВИЧ-1, содержащая полипептид по любому из пп.1, 2 и носитель.
4. Вакцина против ВИЧ-1, содержащая полипептид по любому из пп.1, 2 и носитель.
5. Молекула нуклеиновой кислоты, кодирующая полипептид по любому из пп.1, 2.
6. Вектор, содержащий остов цитомегаловируса человека (HCMV) или остов цитомегаловируса макаки-резус (RhCMV), причем указанный остов HCMV или RhCMV содержит последовательность нуклеиновой кислоты, кодирующую антиген, причем указанный антиген содержит полипептид по любому из пп.1, 2.
7. Вектор по п.6, отличающийся тем, что в осто́ве HCMV или RhCMV отсутствует участок гена UL130-128.
8. Вектор по п.6 или 7, отличающийся тем, что в осто́ве HCMV или RhCMV отсутствует ген UL82, кодирующий белок тегумента pp71.
9. Вектор по п.6, отличающийся тем, что в осто́ве HCMV или RhCMV отсутствует участок гена UL130-128 и ген UL82, кодирующий белок тегумента pp71.
10. Способ лечения ВИЧ-1 у субъекта или защиты субъекта от инфицирования ВИЧ-1, включающий введение эффективного количества полипептида, нуклеиновой кислоты, фармацевтической композиции, вакцины или вектора по любому из пп.1-9 нуждающемуся в этом субъекту.
11. Способ индукции иммунного ответа против ВИЧ-1 у субъекта, включающий введение эффективного количества полипептида, нуклеиновой кислоты, фармацевтической композиции, вакцины или вектора по любому из пп.1-9 указанному субъекту.
12. Способ индукции ответа эффекторных Т-клеток памяти против ВИЧ-1 у субъекта, включающий введение эффективного количества полипептида, нуклеиновой кислоты, фармацевтической композиции, вакцины или вектора по любому из пп.1-9 указанному субъекту.



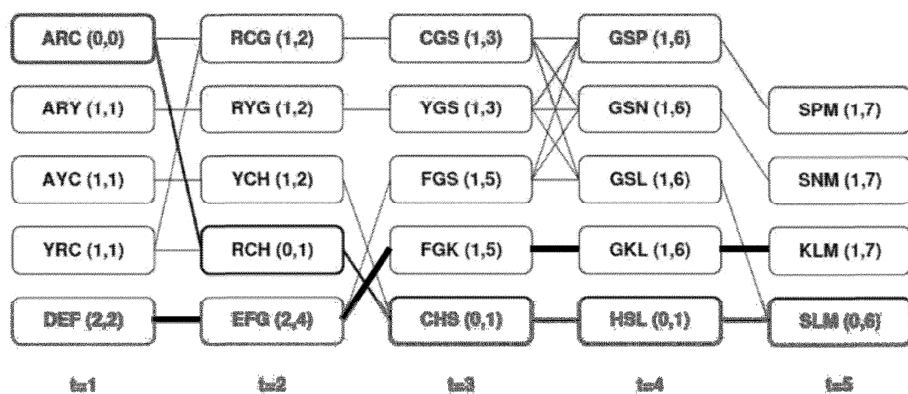
Фиг. 1А



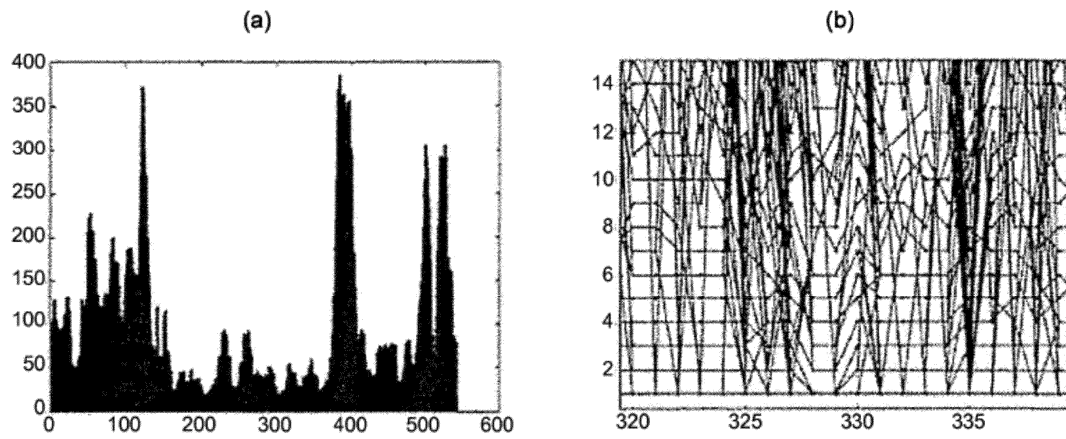
Фиг. 1В



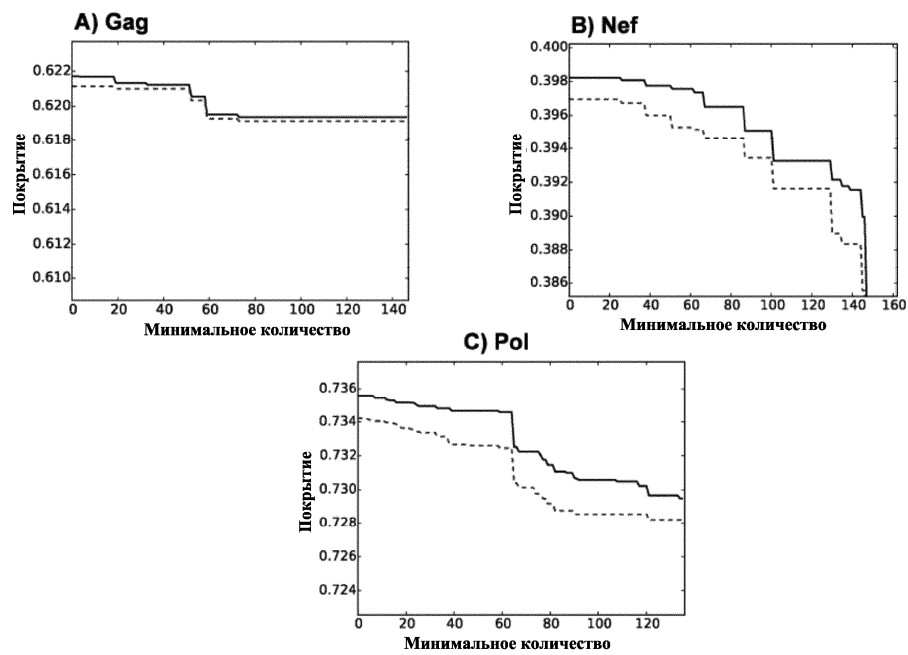
Фиг. 2



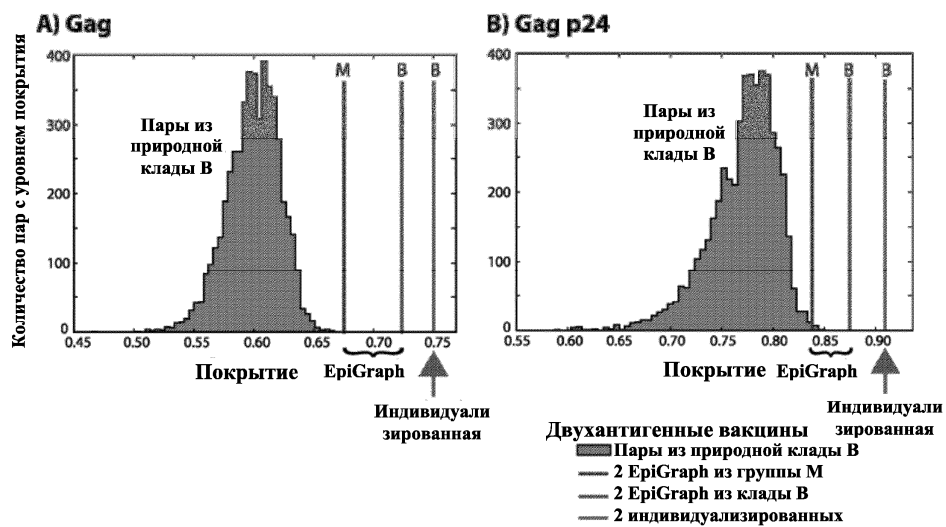
Фиг. 3



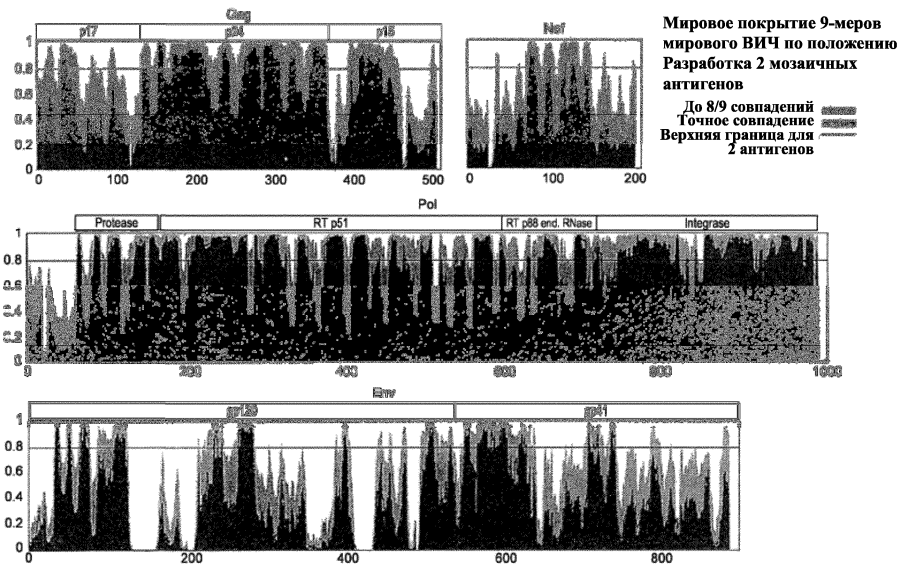
Фиг. 4А, В



Фиг. 5А-С

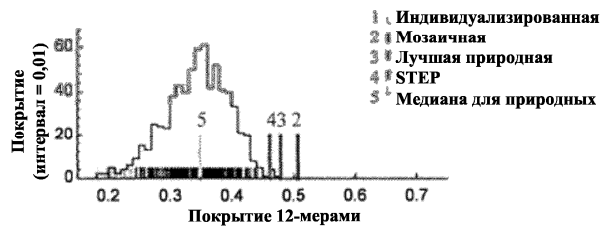


Фиг. 6

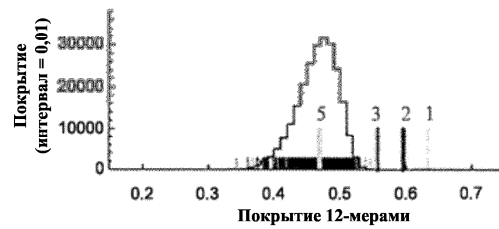


Фиг. 7

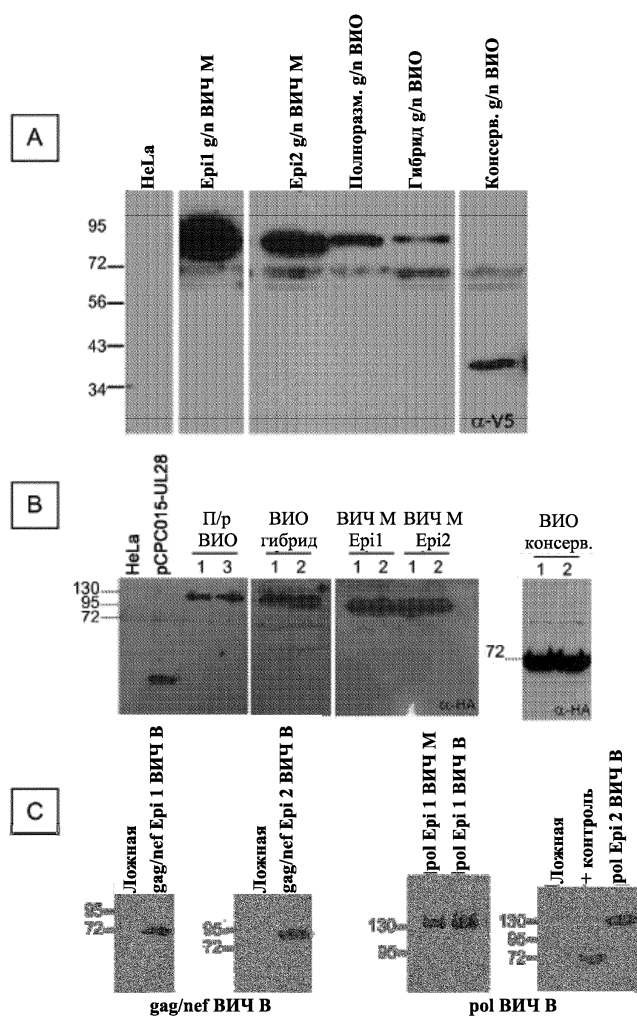
А) Среднее покрытие эпитопов Gag одиночными антигенами для вакцины



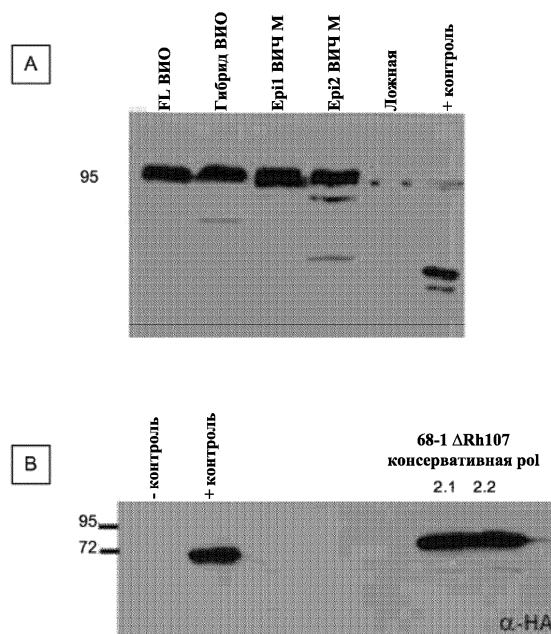
В) Среднее покрытие эпитопов Gag парами антигенов для вакцины



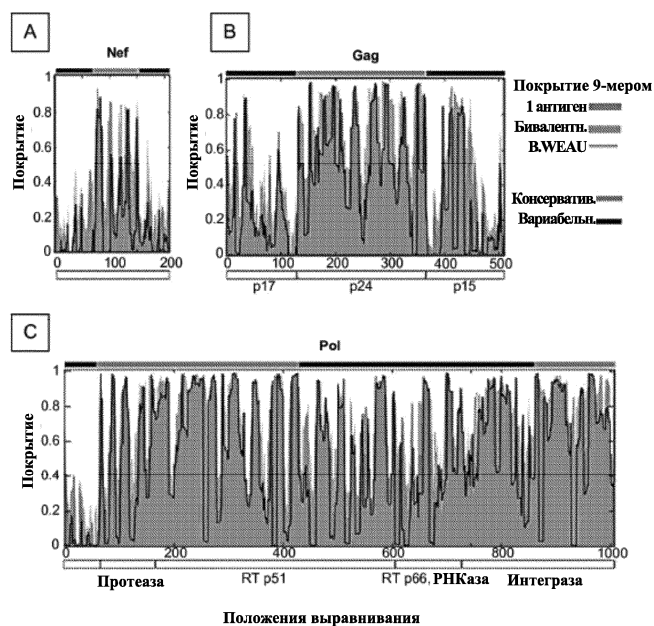
Фиг. 8А, В



Фиг. 9А-С



Фиг. 10А, В



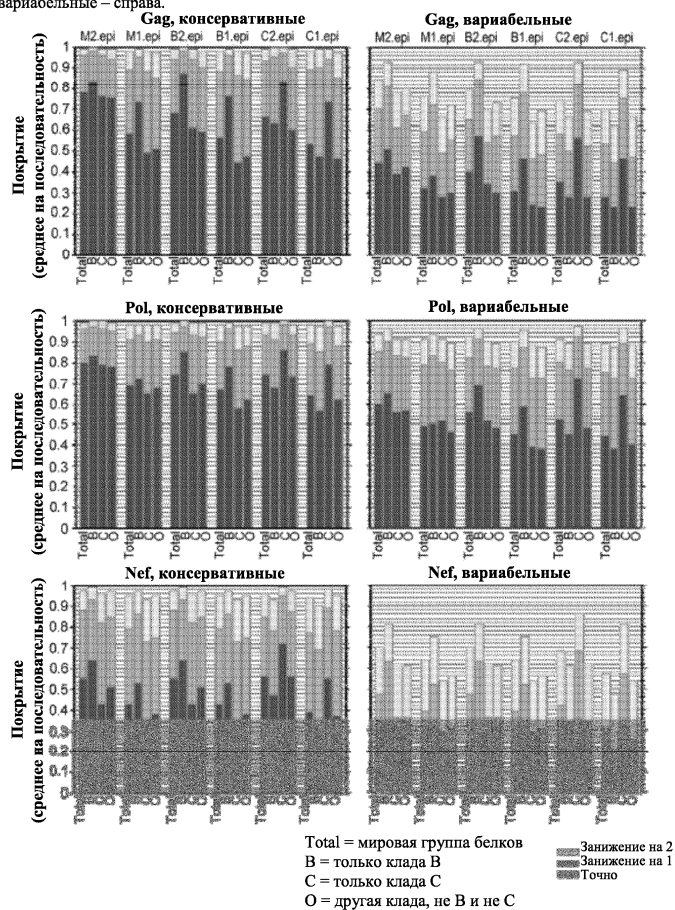
Фиг. 11

Потенциальный Т-клеточный эпитоп (9-мер)

Покрывание консервативных и переменных участков Gag, Pol и Nef:

M = группа M, мировая; B = клада B; C = клада C

M2.epi означает 2-валентное решение EpiGraph, консервативные части белков слева, переменные – справа.



Фиг. 12



Евразийская патентная организация, ЕАПВ

Россия, 109012, Москва, Малый Черкасский пер., 2