



ФЕДЕРАЛЬНАЯ СЛУЖБА
ПО ИНТЕЛЛЕКТУАЛЬНОЙ СОБСТВЕННОСТИ,
ПАТЕНТАМ И ТОВАРНЫМ ЗНАКАМ

C12Q 1/04 (2006.01)
C12Q 1/68 (2006.01)
C12N 1/10 (2006.01)
C12N 15/09 (2006.01)
C12N 1/21 (2006.01)
C12N 15/11 (2006.01)
A61K 39/00 (2006.01)
C07K 14/255 (2006.01)

(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ПАТЕНТУ

(21), (22) Заявка: 2003135645/13, 11.12.1995

(24) Дата начала отсчета срока действия патента:
11.12.1995

(30) Конвенционный приоритет:
11.12.1995 GB 95/02875

(43) Дата публикации заявки: 27.05.2005

(45) Опубликовано: 20.10.2009 Бюл. № 29

(56) Список документов, цитированных в отчете о
поиске: WO 9201056 A1, 23.01.1992. HENSEL M.
et al. Simultaneous identification of bacterial
virulence genes by negative selection. Science. -
1995, Jul 21; 269(5222): 400-3. SLAUCH J.M. et
al. In vivo expression technology for selection
of bacterial genes specifically induced in host
tissues. Methods Enzymol. - 1994; 235: 481-92.
PASCOPELLA L. et al. Use (см. прод.)

(62) Номер и дата подачи первоначальной заявки,
из которой данная заявка выделена: 97111781
11.12.1995

Адрес для переписки:
103735, Москва, ул. Ильинка, 5/2, ООО
"Союзпатент", пат.пов. А.П.Агуреву

(72) Автор(ы):

ХОЛДЕН Дэвид Уильям (GB),
ШИА Жаклин Элизабет (GB),
ХЕНЗЕЛЬ Михаэль (DE)

(73) Патентообладатель(и):

ИМЕРДЖЕНТ ПРОДАКТ
ДИВЕЛОПМЕНТ ЮК ЛИМИТЕД (GB)

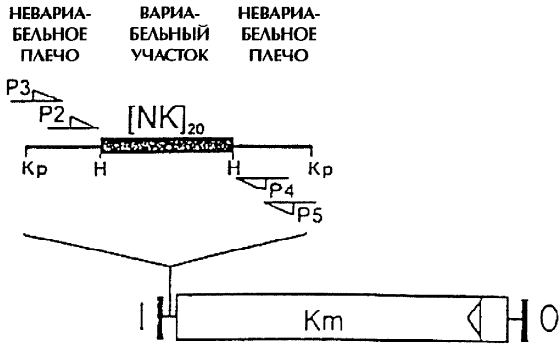
(54) VGC2 ДНК SALMONELLA TYPHIMURIUM, МУТАНТНАЯ БАКТЕРИЯ, ОБЛАДАЮЩАЯ
ПОНИЖЕННОЙ СПОСОБНОСТЬЮ К АДАПТАЦИИ К УСЛОВИЯМ ОКРУЖАЮЩЕЙ
СРЕДЫ, И СПОСОБ ЕЕ ПОЛУЧЕНИЯ

(57) Реферат:

Изобретение относится к генетической
инженерии, в частности к идентификации
генов, участвующих в адаптации организма в
среде его обитания, и может быть
использовано для создания рекомбинантной
бактерии с пониженной адаптацией к
конкретной окружающей среде. Инактивация
определенной последовательности VGC2
ДНК Salmonella typhimurium, которая
расположена между генами ydhE и rykF, или ее

части, состоящей, по меньшей мере, из 50
нуклеотидов, или варианта указанной ДНК с,
по крайней мере, 85% идентичностью,
представляющего собой VGC2 ДНК любого из
микроорганизмов Salmonella aberdeen,
Salmonella gallinarum, Salmonella cubana
и Salmonella typhi, позволяет получить
микроорганизм с пониженной способностью к
адаптации к конкретным условиям
окружающей среды. 3 н. и 3 з.п. ф-лы, 12 ил.

МЕТКА ПОСЛЕДОВАТЕЛЬНОСТЬ ДНК



МИНИ - Tn5 - ТРАНСПОЗОН
ФИГ. 1а

(56) (продолжение):
of in vivo complementation in Mycobacterium tuberculosis to identify a genomic fragment associated with virulence. Infect Immun. - 1994, Apr; 62(4): 1313-9. US 5397697, 14.03.1995. SU 1751208 A1, 30.07.1992.

RU 2 3 7 0 5 4 1 C 2

RU 2 3 7 0 5 4 1 C 2



FEDERAL SERVICE
FOR INTELLECTUAL PROPERTY,
PATENTS AND TRADEMARKS

(19) **RU** (11) **2 370 541** (13) **C2**

(51) Int. Cl.

C12Q 1/04 (2006.01)
C12Q 1/68 (2006.01)
C12N 1/10 (2006.01)
C12N 15/09 (2006.01)
C12N 1/21 (2006.01)
C12N 15/11 (2006.01)
A61K 39/00 (2006.01)
C07K 14/255 (2006.01)

(12) ABSTRACT OF INVENTION

(21), (22) Application: **2003135645/13**, **11.12.1995**

(24) Effective date for property rights:
11.12.1995

(30) Priority:
11.12.1995 GB 95/02875

(43) Application published: **27.05.2005**

(45) Date of publication: **20.10.2009 Bull. 29**

(62) Number and date of filing of the initial application, from which the given application is allocated: **97111781 11.12.1995**

Mail address:

**103735, Moskva, ul. Il'inka, 5/2, OOO
"Sojuzpatent", pat.pov. A.P.Agureevu**

(72) Inventor(s):

**KhOLDEN Dehvid Uil'jam (GB),
ShIA Zhaklin Ehlizabet (GB),
KhENZEL' Mikhaehl' (DE)**

(73) Proprietor(s):

**IMERDZhENT PRODAKT DIVELOPMENT JuK
LIMITED (GB)**

(54) **VGC2 DNA OF SALMONELLA TYPHIMURIUM, MUTANT BACTERIUM WITH REDUCED ADAPTABILITY TO ENVIRONMENT, AND METHOD OF ITS OBTAINMENT**

(57) Abstract:

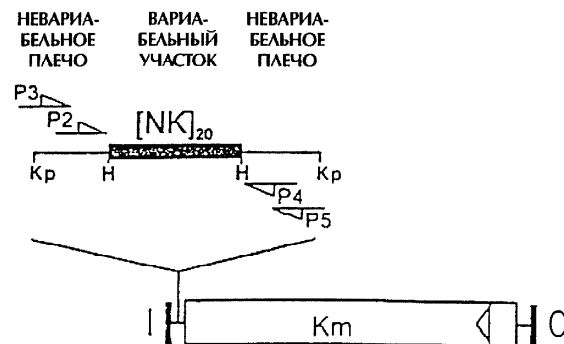
FIELD: medicine.

SUBSTANCE: method involves deactivation of definite VGC2 DNA sequence of Salmonella typhimurium positioned between ydhE and pykF genes or its part containing at least 50 nucleotides, or the DNA version of at least 85% identity, representing VGC2 DNA of any microbe out of Salmonella aberdeen, Salmonella gallinarum, Salmonella cubana and Salmonella typhi.

EFFECT: obtainment of microbe with reduced adaptability to specific environmental conditions.

6 cl, 12 dwg, 8 ex

МЕТКА ПОСЛЕДОВАТЕЛЬНОСТЬ ДНК



МИНИ - Тп5 - ТРАНСПОЗОН

ФИГ. 1a

Текст описания приведен в факсимильном виде.

Область техники, к которой относится изобретение

5 Настоящее изобретение относится к способам
идентификации генов, участвующих в адаптации микроорганизма в
среде его обитания, особенно к способам идентификации генов,
10 ответственных за вирулентность патогенного микроорганизма.

Уровень техники

15 Антибиотическая резистентность бактериальных и других
патогенов становится все более важной проблемой. В связи с
этим важны новые терапевтические подходы к борьбе с
20 патогенными микроорганизмами.

25 Патогенные микроорганизмы способны избегать
воздействия защитных механизмов хозяина и могут расти в
плохой питательной среде, вызывая заражение. Для этого
30 микроорганизм должен содержать ряд "вирулентных" генов.

Вирулентные гены могут быть определены с
использованием приемов классической генетики. С помощью
35 транспозонного мутагенеза осуществляют идентификацию
бактериальных вирулентных генов. Так, например, мутанты

подвергали скринингу на определенные физиологические
дефекты такие, как потеря железо-регулируемых белков
5 (Holland с сотр./ 1992) или использовали в анализах для
изучения пенетрации эпителиальных клеток (Finlay с сотр.,
1988) и выживания в макрофагах (Fields с сотр., 1989;
10 Miller с сотр., 1989a; Groisman с сотр., 1989).
Транспозонные мутанты также тестировали на измененную
15 вирулентность живых животных инфицированных моделей
(Miller с сотр., 1989b). Такой подход имеет преимущество в
том, что могут идентифицироваться гены, играющие важную
20 роль на различных стадиях заражения, но имеет ограничения,
связанные с необходимостью индивидуального тестирования
25 большого числа мутантов на изменения вирулентности.
Miller с сотр. (1989b) использовали группы из 8-10 мы-
шей и инфицировали орально 95 отдельных групп различными
30 мутантами, вследствие чего использовали 760-950
мышей. В связи с тем, что требуется чрезвычайно большое
35 количество животных,

полный скрининг бактериального генома на вирулентные гены не представляется возможным.

5 Недавно была описана генетическая система (технология экспрессии in vivo (IVET)), которая позволяет осуществлять пол-
жительную селекцию специфически инфицированных генов Salmonel-
10 la (Mahan с сотр., 1993). Такая техника позволяет идентифици-
ровать гены, которые экспрессируются на конкретной стадии про-
цесса инфицирования. Однако эта техника не позволяет идентифи-
15 цировать вирулентные гены, которые регулируются посттранскрип-
ционно и, что более важно, не обеспечивает информацией о том,
действительно ли идентифицированные гены требуются для процес-
20 са инфицирования или способствуют ему.

Lee и Falkow (1994) в Methods Enzymol. 236, 531-545 описы-
вают способ идентификации факторов, влияющих на инвазию Salmo-
25 nella в клетки млекопитающих in vitro, путем выделения гиперин-
вазивных мутантов.

Walsh и Серко (1992) Science 255, 434-440 описывают спо-
30 соб отслеживания пространственного расположения недефференци-
рованных клеток коры головного мозга в ходе развития коры го-
ловного мозга крыс. В методе Walsh и Серко используется метка,
35 содержащая уникальную последовательность нуклеиновой кислоты и
Iac Z ген, однако не имеется указаний о возможности детекции
этим методом ценных мутантов или генов.
40

В WO 94/26933 и в статье Smith с сотр., (1995) Proc.Natl.
Acad.Sci.USA 92, 6479-6483 описываются методы, нацеленные на
45 идентификацию функциональных участков известного гена или, по
крайней мере, молекулы ДНК, о последовательности которой име-
ется некоторая информация.

Groisman с сотр. (1993) в *Proc Natl. Acad. Sci*, США 90, 1033-1037 описывают молекулярный, функциональный и эволюционный анализ последовательностей специфичный на Salmonella.

Hensel с сотр. (1995) в *Science* 269, 400-403 описывают одновременную идентификацию бактериальных вирулентных генов путем негативного отбора. Эта статья была опубликована в июле 1995 г.

Slauch с сотр. (1994) в *Methods in Enzymol.* 235, 481-492 описывают технологию экспрессии in vivo для селекции бактериальных генов специфически индуцированных в тканях хозяина.

Pascopeella с сотр. (1994) в *Inf. Immun.* 62, 1313-1319 описывают использование in vivo комплементации в Mycobacterium tuberculosis для идентификации геномного фрагмента, связанного с вирулентностью.

В патенте США 5,397,697 описывается идентификация растительно-чувствительных генов бактерий.

Некоторые вирулентные гены уже известны для таких патогенных микроорганизмов, как *Escherichia coli*, *Salmonella typhimurium*, *Salmonella typhi*, *Vibrio cholerae*, *Clostridium botulinum*, *Yersinia pestis*, *Shigella flexneri* и *Listeria monocytogenes*, однако во всех случаях идентифицировано лишь небольшое их число.

Заболевание, которое *Salmonella typhimurium* вызывает у мышей, обеспечивает хорошую экспериментальную модель тифоидной лихорадки (Carter и Collins, 1974). К настоящему времени идентифицировано примерно сорок два гена, влияющих на вирулентность *Salmonella* (Groisman и Ochman, 1994). Это составляет примерно одну треть от общего числа предсказанных вирулентных

генов (Groisman и Saier, 1990).

Цель настоящего изобретения заключается в идентификации генов, принимающих участие в адаптации микроорганизма к среде его обитания, особенно в идентификации дополнительных вирулентных генов в патогенных микроорганизмах, с повышенной эффективностью. Другая цель изобретения состоит в уменьшении числа экспериментальных животных, используемых для идентификации генов. Другие цели настоящего изобретения состоят в разработке вакцин, и методов скрининга лекарственных средств, уменьшающих вирулентность.

Краткое изложение сущности изобретения

Первый аспект изобретения относится к разработке способа идентификации микроорганизма с пониженной адаптацией к конкретной среде, включающему стадии:

(1) обеспечения множества микроорганизмов, каждый из которых независимо мутирован инсерционной инактивацией гена с помощью нуклеиновой кислоты, включающей уникальную маркерную последовательность так, что каждый мутант содержит различную маркерную последовательность, или клонов, указанных микроорганизмов;

(2) индивидуального получения хранимого образца каждого мутанта, полученного на стадии (1) и обеспечения индивидуально накопленной нуклеиновой кислоты, включающей особую маркерную последовательность от каждого индивидуального мутанта;

(3) введения множества мутантов, полученных на стадии (1), в указанную конкретную среду и культивирования микроорганизмов в указанной среде;

(4) извлечение микроорганизмов из указанной среды или ее выбранной части и выделение нуклеиновой кислоты из извлеченных микроорганизмов;

(5) сравнения любой маркерной последовательности нуклеиновой кислоты, выделенной на стадии (4), с особой маркерной последовательностью каждого индивидуального мутанта, полученного на стадии (2) и

(6) отбора индивидуального мутанта, не содержащего какую-либо из маркерных последовательностей, выделенных на стадии (4).

Таким образом, в таком способе используется негативная селекция для идентификации микроорганизмов с пониженной способностью к пролиферации в окружающей среде.

Микроорганизм может жить в различных средах и известно, что конкретные гены и их продукты позволяют микроорганизму адаптироваться к окружающей среде обитания. Так, например, для выживания таких патогенных микроорганизмов, как патогенные бактерии или патогенные грибки, в их хозяине требуется продукт одного или более вирулентных генов. Таким образом, в предпочтительном техническом решении изобретения ген микроорганизма, позволяющий ему адаптироваться к конкретной окружающей среде, представляет собой вирулентный ген.

Предпочтительная конкретная окружающая среда представляет собой дифференцированный многоклеточный организм, такой, как растение или животное. Известно, что многие бактерии и грибки инфицируют растения, они способны выживать в нем и вызывать заболевание из-за наличия вирулентных генов или их экспрессии. Подходящие микроорганизмы, в том случае, когда конкретная ок-

ружающая среда представляет собой растение, включают бактерии Agrobacterium tumefaciens, которые образуют опухоли (галлы) особенно на винограде; Erwina amylovora; Pseudomonas solanacearum, которая вызывает увядание большого числа растений; Rhizobium leguminosarum, которая вызывает заболевание бобов; Xanthomonas campestris p.у. citri, которая вызывает язву цитрусовых фруктов; и включают грибок Magnaporthe grisea, который вызывает пирикулярриоз риса; Fusarium spp., который вызывает множество болезней растений; Erysiphe spp., Colletotrichum gloeosporioides; Gaeumannomyces graminis, который вызывает заболевания корневой системы и кроны хлебных злаков и трав; Glo-
mus spp., Laccaria spp.; Leptosphaeria maculans; Phoma trachei-
phila; Phytophthora spp., Pyrenophora teres; Verticillium al-
boatrum и V.dahliae; а также Mycosphaerella musicola и M.fi-
jiensis. Как будет подробно описано ниже, в том случае, когда микроорганизм представляет собой грибок, для его жизненного цикла требуется гаплоидная фаза.

Так известно, что многие микроорганизмы, включающие бактерии, грибы, простейшие животные и трипаносомы, инфицируют животных, особенно млекопитающих, включая людей. Выживание микроорганизмов в организме животного и его способность вызывать заболевания в основном связаны с наличием вирулентных генов и их экспрессии. Такие бактерии включают Bordetella spp., особенно B.pertussis, Campylobacter, особенно C.jejuni, Clostridium spp., особенно C.botulinum, Enterococcus spp., особенно E.faecalis, Escherichia spp., особенно E.coli, Haemophilus spp., особенно H.ducreyi и H.influenzae, Helicobacter spp., особенно H.pylori, Klebsiella spp., особенно K.pneumoniae,

Legionella spp., особенно, L.pneumophila, Listeria spp., осо-
 бенно, L.monocytogenes, Mycobacterium spp., особенно, M.smegma-
 5 tis и M.tuberculosis, Neisseria spp., особенно, N.gonorrhoeae
 и N.meningitidis, Pseudomonas spp., особенно, Ps.aeruginosa,
 10 Salmonella spp., Shigella spp., Staphylococcus spp., особенно, S.
aureus, Streptococcus spp., особенно, S.pyogenes и pneumoniae,
Vibrio spp., и Yersinia spp., особенно, Y.pestis. Все такие
 бактерии вызывают болезни людей и животных. В случае приме-
 15 ния бактерий в способе изобретения, конкретной окружающей сре-
 дой является животное, которое они способны инфицировать и у
 которого они вызывают болезнь. Так, например, в случае инфици-
 20 рования мышей Salmonella typhimurium у мышей развивается забо-
 левание, служащее моделью тифоидной лихорадки людей. Staphylo-
coccus aureus вызывает бактериемию и образование ренального
 25 абсцесса у мышей (Albus с сотр. (1991) Infect.Immun. 59, 1008-
 1014) и эндокардиты у кроликов (Perlmann и Freedman (1971)
Yale J.Biol.Med. 44, 206-213).

30 Необходимо, чтобы грибы или высший эукариотный паразит
 представляли собой гаплоиды в отдельные периоды его жизненного
 цикла (например, роста в окружающей среде). Предпочтительно,
 35 чтобы имелась ДНК-опосредованная интегрированная трансформаци-
 онная система, в том случае, когда микроорганизм представляет
 собой человеческий патоген, удобно, чтобы была доступна живот-
 40 ная модель заболевания человека. Грибы, патогенные в отношении
 людей, включают некоторые Aspergillus spp. (например, A.fumi-
gatus), Cryptococcus neoformans и Histoplasma capsulatum. Оче-
 45 видно, что указанные выше грибы имеют гаплоидную фазу, и для
 них имеется в распоряжении ДНК-опосредованная интегрированная

трансформационная система. Может также использоваться Тохо-
plasma, представляющая собой паразит с гаплоидной фазой в ходе
 5 инфицирования. Бактерии имеют гаплоидный геном.

Наиболее доступными животными моделями человеческих забо-
 леваний, является мышь, крыса, кролик, собака или обезьяна.
 10 Предпочтительно, когда такое животное представляет собой мышь.
 Вирулентные гены, детектируемые способом настоящего изобре-
 тения с использованием животной модели человеческого заболевания,
 15 являются генами, определяющими вирулентность микроорганизма в
 организме человека.

Особенно предпочтительными микроорганизмами для приме-
 20 ния в способах настоящего изобретения являются Salmonella ty-
phimurium, Staphylococcus aureus, Streptococcus pneumoniae, En-
terococcus faecalis, Pseudomonas aeruginosa и Aspergillus
 25 fumigatus.

Далее описывается предпочтительное техническое решение
 настоящего изобретения.

30 Нуклеиновую кислоту, содержащую уникальную маркерную пос-
 ледовательность, получают следующим образом. Комплексный пул
 последовательности двунитевой ДНК "меток" создают с использо-
 35 ванием олигонуклеотидного синтеза и полимеразной реакции син-
 теза цепи (PCR). Каждая ДНК "метки" имеет последовательность
 размером в 20-80 п.о., предпочтительно, 40 п.о., которая флан-
 40 кирована последовательностями, "плечами" размером примерно 15-
 30 п.о., предпочтительно около 20 п.о., которые являются об-
 45 щими для всех "меток". Число пар основания п.о. последовате-
 льности является достаточным для создания большого количества
 (например, $> 10^{10}$) уникальных последовательностей в ходе бес-

порядочного олигонуклеотидного синтеза, но не достаточным для образования вторичных структур, которые могут препятствовать
5 PCR. Аналогично, длина последовательности плечей должна быть достаточной для обеспечения эффективного примирования олигонуклеотидов в PCR.

10 Хорошо известно, что последовательность на 5' окончании олигонуклеотида не нуждается в совместимости с последовательностью-мишенью для амплифицирования.

15 Обычно PCR праймеры не содержат каких-либо комплементарных структур, каждая из которых была бы длиннее 2 пар оснований, особенно на их 3' концах, поскольку это может способствовать
20 образованию искусственных продуктов, называемых "праймерные димеры". В случае гибридизации 3' концов двух таких праймеров образуется "примированный матричный" комплекс, и праймерная
25 инсерция приводит в результате к образованию короткого дуплексного продукта, называемого "праймерным димером".

30 Следует избегать образования в праймерах внутренней вторичной структуры. Для симметричной PCR, рекомендуется для обоих праймеров содержание G+C-40-60% с короткой длиной любого
35 основания. Классические расчеты температуры плавления, используемые совместно с исследованиями гибридизации зонда ДНК, часто позволяют предсказать, что данный праймер должен ренатурироваться при конкретной температуре, или, что при повышении
40 температуры до 72°C будет происходить преждевременная диссоциация гибрида праймер/матрица. На практике такие гибриды оказываются более эффективными в PCR процессе, чем это обычно предсказывается простыми расчетами T_m.

45 Оптимальные температуры отжига могут быть определены эм-

пирически и могут превышать предположительные значения. При 37-55⁰С Taq ДНК полимераза является активной, при этом появля-
5 ется праймерная вставка в ходе стадии отжига и происходит ста-
билизация гибрида. Концентрации праймеров равны в типичной
(симметричной) PCR и, как правило, составляют 0,1-1 мкМ.

10 "Метки" лигируются в транспозон или транспозон-подобный
элемент с образованием нуклеиновой кислоты, содержащей особую
маркерную последовательность. В целях удобства транспозон пе-
15 реносится на вектор-"самоубийцу", который сохраняется в виде
плазмиды в организме "помощника", но теряется после переноса
на микроорганизм, используемый в способе настоящего изобре-
20 тения. Так, например, организм "помощник" может представлять со-
бой штамм Escherichia coli, микроорганизм способа может пред-
25 ставлять собой Salmonella, а передача происходит в ходе конъю-
гации. Хотя транспозон может теряться после переноса в опреде-
ленной части клеток он претерпевает транспозицию даже, если он
30 интегрируется рандомизированно, совместно с его уникальной
меткой в геном микроорганизма, используемого в способе изобре-
тения. Наиболее предпочтительно, чтобы транспозон или транспо-
35 зон подобный элемент, мог быть подвергнут селекции. Так, напри-
мер, в случае Salmonella, канамицин-устойчивый ген может при-
сутствовать в транспозоне, при этом эксконъюганты подвергают
40 отбору на среде, содержащей канамицин. Существует также воз-
можность комплементации ауксотрофного маркера в клетке-реципи-
енте функциональным геном в нуклеиновой кислоте, содержащей
45 уникальный маркер. Такой способ особенно удобен при использо-
вании грибов в способе изобретения.

Предпочтительно, чтобы комплементирующий функциональный

ген не был получен из тех же видов микроорганизмов, что и реципиентный микроорганизм иначе могут произойти неслучайные интеграционные события. Так же является предпочтительным перенос транспозона или транспозон-подобного элемента на вектор, который эписомально сохраняется (т.е. не является частью хромосомы) в микроорганизме, используемом в способе по первому аспекту изобретения в начальном заданном состоянии, принимая во внимание, что при его изменении, эписом не удерживается в состоянии, при котором возможно осуществлять селекцию клетки, в которой транспозон или транспозон-подобный элемент подвергаются транспозиции, даже если они интегрируются случайным образом, совместно с уникальной меткой, в геном микроорганизма используемого в способе изобретения. Такое техническое решение является выгодным, поскольку после получения микроорганизма, несущего эписомальный вектор, транспозиционное событие каждый раз отбирается или индуцируется изменением состояния микроорганизма (или его клона) от первого данного состояния во второе данное состояние, и транспозон может интегрироваться на различных сайтах генома микроорганизма. Так, после получения основной коллекции микроорганизмов, каждый член которой содержит особую меченную последовательность в транспозоне или транспозон-подобном элементе, переносимых на эписомальный вектор (в первом данном состоянии), такая коллекция может использоваться повторно для образования пулов рандомизированных инсерционных мутантов, каждый из которых содержит различные меченные последовательности (т.е. особые последовательности в рамках пула). Такое техническое решение особенно полезно поскольку (а) оно уменьшает число и сложность манипуляций, требуемых для создания множест-

ва ("пула") независимо мутированных микроорганизмов на стадии (1) настоящего способа; и (6) требуется лишь тоже число меток, что и число микроорганизмов во множестве микроорганизмов на стадии (1) способа изобретения. Пункт (а) упрощает такой способ при его использовании в организмах, в которых наиболее трудно осуществить транспозонный мутагенез (например, Staphylococcus aureus), а пункт (6) означает, что могут отбираться меченные последовательности с особенно хорошими гибридизационными характеристиками, в связи с чем облегчается контроль их качества. Как подробно описано ниже, размер "пула" обычно составляет 100 или 200 независимо мутированных микроорганизмов и, поэтому основную коллекцию микроорганизмов удобно хранить на одном или двух 96-луночных титрационных микропланшетах.

В соответствии с особенно предпочтительным техническим решением, первое заданное состояние представляет собой первую конкретную температуру или такой температурный интервал, как 25-32°C, наиболее предпочтительно, около 30°C, а второе заданное состояние представляет собой вторую конкретную температуру или такой температурный интервал, как 35-45°C, наиболее предпочтительно, 42°C. В дополнительных предпочтительных технических решениях первое заданное состояние представляет собой наличие такого антибиотика, как стрептомицин, а второе заданное состояние представляет собой отсутствие указанного антибиотика; либо первое заданное условие представляет собой отсутствие антибиотика, а второе заданное состояние - присутствие указанного антибиотика.

Транспозоны, подходящие для интеграции в геном Грам-отрицательной бактерии включают Tn5, Tn10 и их производные. Тран-

спозоны, подходящие для интеграции в геном Грамположительных бактерий включают Tn916, его производные или аналоги. Транспозоны, особенно подходящие для использования с Staphylococcus aureus, включают Tn917 (Cheung с сотр.) 1992, Proc. Natl. Acad. Sci. США, 89, 6462-6466 (и Tn918) Albus с сотр. (1991) Infect. Immun., 59, 1008-1014).

Особенно предпочтительно, если транспозоны обладают свойствами Tn917 производных, описанных Camilli с сотр. (1990) J. Bacteriol. 172, 3738-3744 и переносятся термо-чувствительными векторами такими, как pE194T (Villafane с сотр. (1978) J. Bacteriol. 169, 4822-4829).

Следует иметь в виду, что хотя транспозоны и удобны для инсерционной инактивации гена, но также возможно применение любых других способов. Другим удобным способом инсерционной инактивации гена, особенно в некоторых бактериях, таких как Streptococcus, является инсерционно-дупликационный мутагенез, подобный тому, что описан в статье Morrison с сотр. (1984) J. Bacteriol. 159, 870 в отношении S. pneumoniae. Общий метод может применяться на других микроорганизмах, особенно на бактериях.

В случае грибков инсерционные мутации создают путем трансформации с использованием фрагментов ДНК или плазмид, несущих "метки" и, предпочтительно, селективируемый маркер кодирующий, например, устойчивость к гигромицину В или флеомицину (см. Smith с сотр. (1994) Infect. Immunol. 62, 5247-5254). Неупорядоченная, единичная интеграция фрагментов ДНК, кодирующих устойчивость к гигромицину В, в геном гифомицеты с использованием интеграции под действием рестриктазы (REMI; Schiestl и

Petes (1991); Lu с сотр. (1994) Proc.Natl.Acad.Sci.США, 91, 12649-12653) является известным явлением.

5 Простая методика инсерционного мутагенеза гриба описана Schiestl и Petes (1994), на которых ссылаются в настоящем изобретении, и она включает, например, использование Ту элементов
10 и рибосомной ДНК в дрожжах.

Неупорядоченная интеграция транспозона или других последовательностей ДНК позволяет выделить множество независимо мутированных микроорганизмов, в которых различные гены инсерционно инактивированы в каждом мутанте, и каждый мутант содержит
15 маркерную последовательность.

20 Библиотека таких инсерционных мутантов ранжируется в микротитрационных планшетах так, что каждая лунка содержит различные мутантные микроорганизмы. Аккумулируют ДНК, содержащую
25 уникальную маркерную последовательность от каждого индивидуального мутантного микроорганизма (удобно использовать общую ДНК из клона). Такую операцию удобно проводить путем удаления
30 образца микроорганизма из микротитрационного планшета, нанесения его штрихами на мембрану для гибридизации нуклеиновой кислоты (например, на нитроцеллюлозную или найлоновую мембраны),
35 лизирования микроорганизма в щелочи и фиксации нуклеиновой кислоты на мембране. Таким способом получают копию содержимого лунок.

40 Из полученных на микротитрационных планшетах пулов микроорганизмов выделяют ДНК. Такую ДНК используют в качестве мишени для PCR с применением праймеров, которые осуществляют ренатурацию с общими участками, фланкирующими "метки", амплифицированную ДНК метят, например, с помощью ³²P. Продукт такой PCR
45

используют для зондирования ДНК, полученной из каждого индивидуального мутанта, с целью обеспечения эталонной гибридизационной карты для реплик микротитрационных планшетов. Такую операцию проводят с целью определения, действительно, каждый из индивидуальных микроорганизмов фактически содержит маркерную последовательность, и может ли маркерная последовательность быть эффективно амплифицирована и мечена.

Получают пулы транспозонных мутантов для введения в определенную среду. Удобно использовать 96-луночные планшеты для микротитрования, при этом пул содержит 96 транспозонных мутантов. Однако, нижний предел размера пула составляет два мутанта; не имеется теоретического верхнего предела размера пула, однако, как обсуждается ниже, верхний предел может быть определен с учетом окружающей среды, в которую введены мутанты.

При введении микроорганизмов в указанную конкретную окружающую среду микроорганизмам создают условия для их роста. Период нахождения микроорганизмов определяется его природой и средой его обитания. По истечении надлежащего времени пребывания микроорганизмы выводят из окружающей среды, выделяют ДНК и такую ДНК используют в качестве матрицы для РСР с применением праймеров, которые осуществляют ренатурацию с "плечами", фланкирующими "метки". Продукт РСР метят, например, с помощью ^{32}P , и используют для зондирования ДНК, запасенной от каждого индивидуального мутанта, реплицированного из микротитрационного планшета. Идентифицировали аккумулярованную ДНК, которая слабо гибридизируется или совсем не гибридизируется с зондом, полученным из ДНК, выделенной из микроорганизмов, извлеченных из окружающей среды. Негибридизирующиеся ДНК соответствуют

мутантам, адаптация которых в конкретной окружающей среде ослаблена инсерцией транспозона или другой последовательности ДНК.

Наиболее предпочтительно, когда "плечи" не содержат или содержат очень мало меченных атомов, по сравнению с "метками". Так, например, праймеры PCR строят таким образом, чтобы они не содержали или содержали только единственный G остаток, ^{32}P -меченный нуклеотид представляет собой dCTP и, в этом случае ни одного или один радиомеченный C-остаток вводится в каждое "плечо", а основное количество радиомеченных C-остатков вводится в "метку". Предпочтительно, чтобы "метка" содержала, по крайней мере, в десять раз большее количество меченных атомов, чем соответствующее количество, введенное в "плечи"; предпочтительно, такое количество должно быть больше в двадцать или более раз; более предпочтительно, в пятьдесят и более раз. Удобно, чтобы "плечи" могли удаляться из "метки" с использованием соответствующей рестриктазы, сайт для которого может вводиться в праймерную конструкцию.

Как отмечалось выше, наиболее предпочтительно, когда микроорганизм представляет собой патогенный микроорганизм, а конкретная окружающая среда представляет собой животное. В соответствии с таким техническим решением размер пула мутантов вводимого в животное определяется (а) числом клеток каждого мутанта, которые вероятно выживают в животном (предполагая, что вирулентный ген не инактивируется) и (б) общим количеством инокулята микроорганизма. Если число по пункту (а) слишком низкое, то могут быть получены ложные положительные результаты, а если число по пункту (б) слишком велико, то живот-

ное может погибнуть до того, как достаточное количество мутантов получит возможность для нормального роста. Число клеток по пункту (а) может быть определено для каждого используемого микроорганизма, но, предпочтительно, оно имеет значение более 50, более, предпочтительно, выше 100.

Число различных мутантов, которые могут вводиться в одно животное, предпочтительно составляет 50-500, наиболее предпочтительно, около 100. Предпочтительно, когда общее количество инокулята не превышает 10^6 клеток (и предпочтительно, составляет 10^5 клеток), хотя размер инокулята может быть выше или ниже указанного значения в зависимости от природы микроорганизма и животного.

В соответствии с наиболее удобным методом используют инокулят порядка 10^5 , содержащий 1000 клеток каждого из 100 различных мутантов, в расчете на одно животное. Следует отметить, что согласно такому способу может использоваться одно животное для скрининга 100 мутантов по сравнению с известными способами, в которых требуется, по крайней мере, 100 животных для отбора 100 мутантов.

Однако, удобнее всего инокулировать трех животных одним и тем же пулом мутантов так, чтобы, по крайней мере, на двух из них можно было проводить исследования (одно из животных в этом случае служит копией для проверки достоверности способа), в то время как третье животное служит объектом обратного контроля. Несмотря на это, такой способ все еще обеспечивает более, чем 30-кратную экономию числа используемых животных.

Промежуток времени между введением пула мутантов в животное и извлечением микроорганизмов может изменяться в за-

висимости от природы используемого микроорганизма и животного. Так, например, если животное представляет собой мышь, а микроорганизм представляет собой Salmonella typhimurium, время между инокуляцией и выделением составляет примерно три дня.

Согласно одному из воплощений настоящего изобретения, микроорганизмы извлекаются из среды обитания на стадии (5) на сайте, отдаленном от сайта введения на стадии (4) так, что подлежащие исследованию вирулентные гены включают те, что распространились между двумя такими сайтами.

Так, например, в случае растения, микроорганизм может вводиться в повреждение на стебле или на один участок листа, а микроорганизм извлекается из другого участка листа, где обнаружено болезненное состояние.

В случае животного, микроорганизм может вводиться орально, внутрибрюшинно, внутривенно или интраназально и с течением времени извлекаться из такого внутреннего органа, как селезенка. Может оказаться полезным сравнение вирулентных генов, идентифицированных при оральном применении и тех, что идентифицированы в результате внутрибрюшинного применения, поскольку некоторые гены вызывают заражение одним путем, но не другим. Предпочтительно, чтобы Salmonella вводилась внутрибрюшинно.

Другими предпочтительными окружающими средами, которые могут использоваться для идентификации вирулентных генов являются животные клетки в культуре (особенно макрофаги и эпителиальные клетки) и растительные клетки в культуре. Хотя использование клеток в культуре полезно само по себе, оно может также дополнять применение целого животного или растения, как это имеет место в определенных случаях, в качестве окружающей сре-

ды.

Также предпочтительно, когда окружающая среда представляет собой часть тела животного. В рамках данного взаимодействия хозяин-паразит возможен ряд различных сред обитания, включающих различные органы и ткани, а также их части, например, очаг *Peyer*.

Число индивидуальных микроорганизмов (т.е. клеток), выделенных из окружающей среды должно, по крайней мере, в два раза, предпочтительно, по крайней мере, в десять раз, более предпочтительно, в 100 раз превышать число различных мутантов, введенных в окружающую среду. Так, например, при инокулировании животного 100 различными мутантами примерно 10,000 индивидуальных микроорганизмов должно извлекаться и выделяться их маркерная ДНК.

Дополнительное предпочтительное техническое решение включает стадии:

(1A) удаления ауксотрофов из множества мутантов, продуцированных на стадии:

(1); или

(6A) определения ауксотрофной природы мутанта, отобранного на стадии (6); или

как (1A), так и (6A).

Необходимо различать ауксотрофы (представляющие собой мутантный микроорганизм, требующие наличия факторов роста, не являющихся необходимыми для дикого типа или для прототрофов) и мутантный микроорганизм, в котором ген, обеспечивающий адаптацию такого микроорганизма в конкретной окружающей среде, находится в инактивированном состоянии. Удобно осуществлять такую операцию между стадиями (1) и (2) или после стадии (6).

Предпочтительно, при идентификации вирулентных генов аук-

сотрофы не удалять.

В соответствии со вторым аспектом настоящего изобретения обеспечивается способ идентификации гена, позволяющего микроорганизму адаптироваться в конкретной окружающей среде, причем такой способ включает способ по первому аспекту изобретения, после которого проводят дополнительную стадию:

(7) выделения инсерционно-инактивированного гена или его части из индивидуального мутанта, отобранного на стадии (6).

Способы выделения гена, содержащего особый маркер хорошо известны в области молекулярной биологии.

Еще одно предпочтительное техническое решение включает следующую дополнительную стадию:

(8) выделения из микроорганизма дикого типа соответствующего гена дикого типа с использованием инсерционно-инактивированного гена, выделенного на стадии (7) или его части в качестве зонда. Способы генного зондирования хорошо известны в области молекулярной биологии.

Методы молекулярной биологии, подходящие для использования на практике настоящего изобретения, описаны в работе Sambrook с сотр. (1989), на которую ссылаются в тексте.

В том случае, когда микроорганизм представляет собой микроорганизм, патогенный в отношении животного, а ген представляет собой вирулентный ген, транспозон применяют для инсерционной инактивации гена. Удобно клонировать вирулентные гены перевариванием геномной ДНК из индивидуального мутанта, отобранного на стадии (6), с рестриктазой, осуществляющей разрез вне транспозона, лигированием размерно-фракционированной ДНК, содержащей транспозон, в плазмиду. Селекцию плазмидных реком-

бинантов осуществляют на основе антибиотической устойчивости, приобретенной от транспозона, а не от плазмиды. Геномную ДНК микроорганизма, примыкающую к транспозону, секвенируют с использованием двух праймеров, которые ренатурируются в терминальные области транспозона и двух праймеров, которые отжигаются вблизи полилинкерных последовательностей плазмиды. Такие последовательности могут быть подвергнуты исследованию с использованием базы данных ДНК с целью установления влияния транспозона на известный вирулентный ген. Так, например, последовательность, полученную таким способом, сравнивают с последовательностями, присутствующими в таких общедоступных базах данных, как EMBL и GenBank. Наконец, если оказывается, что прерванная последовательность присутствует в новом вирулентном гене, мутацию переносят на новый генетический фон (например, с помощью фаг Р22-управляемой трансдукции в случае Salmonella) и определяют LD_{50} мутантного штамма с целью подтверждения того, что неvirulentный фенотип обусловлен транспозиционным событием, а не вторичной мутацией.

Число индивидуальных мутантов, подвергнутых скринингу для детекции всех вирулентных генов в микроорганизме, зависит от числа генов в геноме микроорганизма. Так, например, вероятно, что 3000-5000 мутантов Salmonella typhimurium необходимо подвергнуть скринингу для детекции основного количества вирулентных генов, тогда как для Aspergillus spp., которая имеет более крупный геном, чем Salmonella, скринингу подвергают около 20 000 мутантов. Примерно 4% не являющихся незаменимыми генов S.typhimurium требуется для определения вирулентности (Grossman и Saier, 1990), в этом случае, геном S.typhimurium

содержит примерно 150 вирулентных генов. Однако, способы настоящего изобретения предусматривают более быстрый, более удобный и гораздо более практичный путь идентификации вирулентных генов.

Третий аспект настоящего изобретения обеспечивает микроорганизм, полученный с использованием способа по первому аспекту изобретения.

Значимость таких микроорганизмов заключается в том, что они не адаптируются к выживанию в конкретной окружающей среде.

В соответствии с предпочтительным техническим решением патогенный микроорганизм не адаптируется к выживанию в организме хозяина (среде обитания), и в случае микроорганизма, являющегося патогенным по отношению к животным, особенно млекопитающим, более конкретно, к людям, мутант полученный способом настоящего изобретения может использоваться в качестве вакцины. Такой мутант не является вирулентным, поэтому его можно использовать для введения пациенту, предполагается, что он является антигенным и вызывает защитный иммунный ответ.

В соответствии с еще одним предпочтительным техническим решением патогенный микроорганизм, не адаптированный на выживание в организме хозяина, полученный способами изобретения, модифицируют, предпочтительно, путем введения подходящей последовательности ДНК для экспрессии антигенного эпитопа от другого патогена. Такой модифицированный микроорганизм может выполнять функции вакцины для такого другого патогена.

В соответствии с четвертым аспектом изобретения предусматривается микроорганизм, содержащий мутацию в гене, идентифицированном с использованием способа по второму аспекту изоб-

ретения.

Таким образом, в микроорганизм по третьему аспекту изобретения специально вводят мутацию в идентифицированный ген. В соответствии с предпочтительным техническим решением, особенно в том случае, когда микроорганизм предназначен для использования в качестве вакцины, мутация в гене представляет собой делецию или сдвиг рамки, или какую-либо иную мутацию, которая в значительной степени неспособна к реверсии. Такие ген-специфичные мутации могут создаваться с использованием таких стандартных методов, как введение в микроорганизм копии мутантного гена на автономный репликон (такой, как плазида или вирусный геном) и использование гомологической рекомбинации для введения мутации в копию гена в геноме микроорганизма.

Пятый и шестой аспекты изобретения обеспечивают подходящий микроорганизм, предназначенный для использования в вакцине и вакцину, содержащую подходящий микроорганизм и фармацевтически приемлемый носитель.

Подходящий микроорганизм представляет собой упомянутый выше невирулентный мутант.

Предпочтительна активная иммунизация пациента. В соответствии с таким подходом получают один или более мутантных микроорганизмов в иммуногенной рецептуре, содержащей приемлемые адъюванты и носители, которые известными методами вводят пациенту. Такие адъюванты включают полный или неполный адъюванты Фрейнда, мурамоил дипептид, "Iscoms" из EP 109 942, EP 180 564 и EP 231 039, гидроксид алюминия, сапонин, ДЕАЕ-декстран, нейтральные масла (например, микглицол), растительные масла (например, арахисовое масло), липосомы, полиолы Pluronic или

адьювантную систему Ribi (см. например, GB-A-2 189 141). "Plu-
ronic" представляет собой зарегистрированную торговую марку.
5 Пациент, подвергаемый иммунизации, является пациентом, которо-
му необходима защита от заболевания, вызванного вирулентной
формой микроорганизма.

10 Упомянутые выше авирулентные микроорганизмы изобретения
или их рецептуры могут применяться любым традиционным спосо-
бом, включая оральное применение и парентеральную (например,
15 подкожную или внутримышечную) инъекцию. Лечение может состоять
из введения единичной дозы или многократных доз, применяемых в
течение определенного времени.

20 Хотя авирулентный микроорганизм изобретения может приме-
няться сам по себе, предпочитают, чтобы он присутствовал в фар-
мацевтической рецептуре совместно с одним или более приемлемы-
25 ми носителями. Носитель (ли) должен быть "приемлемым" в смысле
совместимости с авирулентным микроорганизмом изобретения без
пагубного воздействия на его реципиента. Обычно, в качестве но-
30 сителей используют воду или физиологический раствор, которые
должны быть стерильными и непирогенными.

35 Следует иметь в виду, что вакцина изобретения, в зависи-
мости от входящего в ее состав микроорганизма, может исполь-
зоваться в области медицины человека и ветеринарной медицины.

40 Болезни, вызываемые микроорганизмами, известны для многих
животных, например, для домашних животных. Вакцины изобре-
тения, содержащие соответствующий авирулентный микроорганизм,
особенно авирулентные бактерии, могут применяться на людях, но
45 также и, например, на коровах, овцах, свиньях, лошадях, соба-
ках и кошках, а также таких домашних птицах, как куры, индей-

50

ки, утки и гуси.

Седьмой и восьмой аспект изобретения обеспечивает ген, полученный по методу второго аспекта изобретения и закодированный полипептид. Под термином "ген" подразумеваются не только участки ДНК, кодирующие полипептид, но также регуляторные области ДНК, которые регулируют транскрипцию, трансляцию и, для некоторых микроорганизмов, сплайсинг РНК. Таким образом, такой ген охватывает промоторы, терминаторы, рибосомо-связывающие последовательности и, для некоторых организмов, интроны и сайты узнавания сплайсинга.

Обычно получают информацию о последовательности инактивированного гена, полученного на стадии 7. В целях удобства, последовательности, близкие к концам транспозона используются в качестве гибридизационного сайта секвенирующего праймера. Полученная последовательность или рестрикционный фрагмент ДНК, соседствующий с самим инактивированным геном, используются для получения гибридизационного зонда, с помощью которого идентифицируют и выделяют из организма дикого типа соответствующий ген дикого типа.

Предпочтительно осуществлять гибридизационное зондирование в строгих условиях с целью гарантии получения именно целевого гена. Под термином "строгие" подразумевается, что ген гибридизируется зондом при иммобилизации гена на мембране, а зонд (который в таком случае имеет длину > 200 нуклеотидов) находится в растворе, и систему иммобилизованный ген/гибридизированный зонд промывают в $0,1\times\text{SSC}$ при 65°C в течение 10 минут. SSC обозначает $0,15\text{M NaCl}/0,015\text{M}$ цитрата натрия.

Предпочтительные последовательности зонда для клонирования

ния вирулентных генов разновидности Salmonella показаны на фиг.5 и 6 и описаны в примере 2.

В особенно предпочтительном техническом решении вирулентные гены Salmonella содержат последовательность, показанную на фиг.5 и 6 и описанную в примере 2.

Еще более предпочтительные гены, которые содержатся, или, по крайней мере, часть которых содержится в последовательностях, показанных на фиг.11 и 12, и которые идентифицированы способом по второму аспекту изобретения. Последовательности, показанные на фиг.11 и 12, представляют собой кластер гена из Salmonella typhimurium, названный вирулентным генным кластером 2 (UGC2). Положение транспозонных инсерций указаны в последовательности и такие транспозонные инсерции инактивируют вирулентную детерминанту организма. Как более подробно обсуждается ниже и особенно в примере 4, в том случае, когда способ по второму аспекту изобретения используется для идентификации вирулентных генов в Salmonella typhimurium, большинство нуклеиновокислотных инсерций (и поэтому идентифицированных генов) классифицируются в относительно небольшой части генома. Такая область, UGC2, содержит другие вирулентные гены, которые, как описано ниже, составляют часть настоящего изобретения.

Ген, выделенный по способу настоящего изобретения, может экспрессироваться в подходящей клетке-хозяине. Так, например, ген (ДНК) в соответствии с известными методиками, может быть соответственно модифицирован в рамках указаний, содержащихся в описании, с целью конструирования вектора экспрессии, который далее используют для трансформации соответствующей клетки-хозяина для экспрессии и продуцирования полипептида настоящего

изобретения. Такие методы описаны в патентах США №№ 4,440,859, выданном 3 апреля 1984 г. на имя Rutter с сотр., 4,530,901 выданном 23 июля 1985 г. на имя Weissman, 4,582,800 выданном 15 апреля 1986 г. на имя Crowl, 4,677,063, выданном 30 июня 1987 на имя Mark с сотр., 4,678,751, выданном 7 июля 1987 г. на имя Goeddel, 4,704,362, выданном 3 ноября 1987 г. на имя Itakura с сотр., 4,710,463, выданном 1 декабря 1987 г. на имя Murray, 4,757,006, выданном 12 июля 1988 г. на имя Toole, Jr., с сотр., 4,766,075, выданном 23 августа 1988 г. на имя Goeddel с сотр., и 4,810,648, выданном 7 марта 1989 г. на имя Stalker, причем на все эти патенты ссылаются в настоящем описании.

ДНК, кодирующая полипептид настоящего изобретения, может быть соединена с большим числом других последовательностей ДНК для введения в соответствующего хозяина. ДНК-компаньон будет зависеть от природы хозяина, способа введения такой ДНК в организм хозяина, и от того, необходимо ли эписомальное содержание или интеграция.

Обычно ДНК встраивают в вектор экспрессии, например, в плазмиду, в надлежащей ориентации и правильной рамке считывания для экспрессии. Если в этом есть необходимость, то ДНК может быть связана с соответствующими транскрипционными и трансляционными регуляторными контрольными нуклеотидными последовательностями, распознаваемыми желаемым хозяином, хотя, как правило, такие контроли уже присутствуют в векторе экспрессии. Такой вектор далее вводят в организм хозяина стандартными методами. Как правило, не все хозяева будут трансформироваться таким вектором. Поэтому необходимо проводить селекцию трансформированных клеток-хозяев. Один из приемов такой

селекции включает введение в вектор экспрессии последовательности ДНК с какими-либо необходимыми контрольными элементами, которая кодирует селектируемый признак в трансформированной клетке, например, устойчивость к антибиотику. С другой стороны, ген для такого отбираемого признака может представлять собой другой вектор, который используется для сотрансформации желаемой клетки-хозяина.

Клетки-хозяева, трансформированные рекомбинантной ДНК настоящего изобретения, затем культивируют в течение достаточного времени и при соответствующих условиях, известных специалисту в данной области, на основании указаний, приведенных в настоящем описании, с целью осуществления экспрессии полипептида, который затем может быть выделен.

Известны многие экспрессионные системы, включающие бактерии (например, *E.coli* и *Bacillus Subtilis*), дрожжи (например, *Saccharomyces cerevisiae*), грибомицеты (например, *Aspergillus*), растительные клетки, животные клетки и клетки насекомых.

Векторы включают прокариотный репликон, например, *Coli* для экспрессии в прокариоте в других непкариотных, клеточных типах. Векторы могут также включать прокариотный промотор, способный направлять экспрессию (транскрипцию и трансляцию) генов в таких бактериальных клетках-хозяевах, как *E.coli*, трансформированных этим геном.

Промотор представляет собой контролирующей экспрессию элемент, образованный последовательностью ДНК, способствующей связыванию РНК-полимеразы и осуществлению транскрипции. Промоторные последовательности, совместимые с упомянутыми бактериальными хозяевами обычно предусматриваются в плазмидных век-

торах, содержащих удобные рестрикционные сайты, для инсерции сегмента ДНК настоящего изобретения.

5 Типичные прокариотические векторные плазмиды представляют собой pUC18, pUC19, pBR322 и pBR329, приобретенные у Биорад
Лабораториз (Ричмонд, КА, США), а также pTrc99A и pKK223-3,
10 приобретенные у Фармация, Пискатавэй, Нью-Джерси, США.

Типичная векторная плазида клетки млекопитающего пред-
ставляет собой pSVL, полученная от Фармация, Пискатавэй, Нью-
15 Джерси, США. В таком векторе используется SV 40 поздний промо-
тор для управления экспрессией клонированных генов, причем на-
ивысший уровень экспрессии обнаружен в таких Т-антиген - про-
20 дущирующих клетках, как COS-1 клетки.

Примером индуцибельного вектора экспрессии млекопитающего
может служить pMSG, также полученный от Фармации. В таком век-
25 торе используется глюкокортикоид-индуцибельный промотор длин-
ного конечного повтора вируса опухоли мышины грудной железы
для управления экспрессией клонированного гена.

Используемые дрожжевые плазмидные векторы представляют
собой pRS403-406 и pRS413-416, и их обычно получают от Страта-
35 ген Клонинг Системс, Ла Йолла, КА 92037, США. Плазмиды pRS 403,
pRS404, pRS405 и pRS406 представляют собой дрожжевые интегриру-
ющие плазмиды (Vlps) и включают дрожжевые селектируемые марке-
40 ры HIS3, TRP1, LEU2 и URA3. Плазмиды pRS413-416 представляют
собой дрожжевые центромерные плазмиды (VCps).

Известны различные способы для связывания ДНК с вектора-
45 ми через комплементарные липкие концы. Так, например, компле-
ментарные гомополимерные отрезки могут добавляться к сегменту
ДНК для ее инсерции в векторную ДНК. Затем такой вектор и сег-

мент ДНК соединяют через водородную связь между комплементарными гомополимерными хвостами с образованием молекул рекомбинантной ДНК.

Синтетические линкеры, содержащие один или более рестрикционных сайтов, обеспечивают альтернативный способ соединения сегмента ДНК с векторами. Сегмент ДНК, образовавшийся в результате переваривания с рестриктазой, как описано выше, обрабатывают ДНК полимеразой бактериофага Т4 или *E.coli* ДНК полимеразой 1, ферментами, которые удаляют выступающие 3'-однонуклеотидные окончания с их 3'-5'-экзонуклеолитическими активностями и заполняют рецессивные 3'-окончания с их полимеризующими активностями.

Поэтому комбинация таких активностей приводит к образованию сегментов ДНК с тупыми концами. Затем сегменты с тупыми концами инкубируют с большим молярным избытком линкерных молекул в присутствии фермента, способного катализировать лигирование молекул ДНК с тупыми концами, например, такого, как лигаза ДНК бактериофага Т4. Таким образом, продукты реакции представляют собой сегменты ДНК, несущие на своих концах полимерные линкерные последовательности. Затем такие сегменты ДНК расщепляют с помощью соответствующей рестриктазы и лигируют с вектором экспрессии, который был подвергнут ферментативному расщеплению в результате которого образовались концы, комплементарные концам сегмента ДНК.

Синтетические линкеры, содержащие множество рестрикционных эндонуклеазных сайтов, доступны коммерчески из ряда источников, включающих Интернейшонал Биотекнолоджиз Инк., Нью-Хайвен, ПН, США.

Предпочтительным способом модификации ДНК, кодирующей полипептид изобретения, является использование полимеразной цепной реакции в соответствии с описанным Saiki с сотр. (1988) Science 239, 487-491. В таком способе ДНК, подвергающаяся ферментативной амплификации, фланкируется двумя специфическими олигонуклеотидными праймерами, которые сами включаются в амплифицированную ДНК. Указанные специфические праймеры могут содержать сайты распознавания рестрикционной эндонуклеазы, которые могут использоваться для клонирования в векторы экспрессии с использованием методов известных в этой области.

Варианты генов также составляют часть изобретения. Предпочтительно, когда такие варианты имеют, по крайней мере, 70% гомологичность последовательности, более предпочтительно, по крайней мере, 85% наиболее предпочтительно, по крайней мере 95% гомологичность последовательности с генами, выделенными способом настоящего изобретения. Разумеется, что допустимы замены, делеции и инсерции. Степень гомологичности между той или иной последовательностями нуклеиновых кислот может быть определена с использованием программы GAP компьютерной группы Висконсинского Университета.

Аналогичным образом, сфера изобретения охватывает варианты белков, закодированные генами.

Под термином "варианты" подразумеваются инсерции, делеции и замены как консервативные, так и неконсервативные, в том случае, когда такие изменения не оказывают существенного влияния на нормальное функционирование белка.

Под термином "консервативные" замены подразумеваются такие комбинации, как Gly, Ala, Val, Ile, Leu, Asp, Glu, Asn,

Gln, Ser, Thr, Lys, Arg и Phe, Tyr.

Такие варианты могут быть получены с использованием хорошо известных способов белковой инженерии и сайт-направленного мутагенеза.

Девятый аспект изобретения предусматривает способ идентификации соединения, снижающего способность микроорганизма к адаптации в конкретной окружающей среде, включающий стадии селекции соединения, которое препятствует функционированию (1) гена, полученного способом по второму аспекту изобретения или (2) полипептида, закодированного таким геном.

Парные скрининги соединений, которые влияют на клетку дикого типа, но не клетку, сверхсинтезирующую ген, выделенный способами настоящего изобретения, составляют часть такого аспекта изобретения.

Так, например, в соответствии с одним из технических решений, одна клетка представляет собой клетку дикого типа, а вторая клетка представляет собой Salmonella, которую получают для сверхсинтеза гена, выделенного способом изобретения. Жизнеспособность и/или рост каждой клетки в конкретной окружающей среде определяется в присутствии соединения, подлежащего тестированию для идентификации соединения, понижающего жизнеспособность или рост клетки дикого типа, но не клетки, сверхсинтезирующей указанный ген.

Предпочтительно, чтобы такой ген представлял собой вирулентный ген.

Так, например, согласно одному из технических решений готовят микроорганизм (например, S.typhimurium) для сверхпродукции вирулентного гена, идентифицированного способом

по первому аспекту изобретения. Каждый из (а) "сверхпродуцирующего" микроорганизма и (б) эквивалентного микроорганизма (который не способен к сверхсинтезу вирулентного гена) используются для инфицирования клеток в культуре. Способность конкретного испытуемого соединения к селективному ингибированию функции вирулентного гена определяется оценкой количества испытуемого соединения, которое требуется для предотвращения инфицирования клеток-хозяев (а) сверхпродуцирующим микроорганизмом и (б) эквивалентным микроорганизмом (по крайней мере, для некоторых вирулентных генных продуктов предполагается, что испытуемое соединение способно инактивировать их и инактивироваться само по себе в результате связывания с продуктом вирулентного гена). Если значительное количество соединения требуется для предотвращения инфицирования микроорганизмами по (а) и (б), то предполагается, что такое соединение селективно. Предпочтительно, чтобы микроорганизмы (например, Salmonella) разрушались внеклеточно под действием таких мягких антибиотиков, как гентамицин (который не проникает в клетки-хозяев), и чтобы действие испытуемого соединения по предотвращению инфицирования клетки микроорганизмом заключалось в лизисе указанной клетки и определялось число присутствующих микроорганизмов (например, путем посева на агар).

Парные скрининги и другие скрининги соединений обобщены Kirsch и Di Domenico (1993) в книге "Обнаружение природных продуктов с терапевтическим потенциалом" (изд. У.Р. Gallo), глава 6, стр.177-221, Баттерворс, У.К. (причем на эту работу ссылаются в настоящем описании).

Парные скрининги осуществляют путем сравнения относитель-

ной чувствительности соединения, используя два генетически родствен-
ственных штамма. Если штаммы отличаются друг от друга в одном
5 локусе, то соединение, специфичное в отношении мишени, может
идентифицироваться сравнением чувствительности каждого из штам-
мов к ингибитору. Так, например, ингибиторы, специфичные в от-
10 ношении мишени, будут более активными в отношении сверхчувст-
вительного испытуемого штамма по сравнению с другим изогенным
родственным штаммом. При применении метода диффузии на агаре
это определяется измерением диаметра зоны подавления роста
15 тест-организма, окружающей диск или лунку с таким соединением.
В процессе диффузии создается непрерывный градиент concentra-
ции соединения, и чувствительность штамма к ингибиторам про-
20 порциональна расстоянию от диска или лунки до края зоны. Обыч-
ные антимикробные агенты или антимикробиальные агенты, с отлич-
ным способом функционирования, обычно являются агентами, обла-
25 дающими одинаковыми активностями по отношению к двум таким
штаммам.

30 Другой тип молекулярного генетического скрининга включает
пары штаммов, в которых клонированный генный продукт сверхпро-
дуцируется в одном штамме по сравнению с контрольным штаммом.
35 В таком анализе штамм, содержащий повышенное количество бел-
ка-мишени, должен быть более устойчивым к ингибиторам, специ-
фичным в отношении клонированного генного продукта, чем изо-
40 генный штамм, содержащий обычные количества белка-мишени. При
применении метода диффузии на агаре, размер зоны подавления
роста, окружающей специфическое соединение, будет меньшим у
45 штамма сверхпродуцирующего белок-мишень, чем у другого изо-
генного штамма.

Дополнительный или альтернативный отбор соединения проводят следующим образом:

5 1. В качестве контроля используют мутантный микроорганизм, полученный с использованием способа по первому аспекту изобретения. (Он имеет заданный фенотип, например, является авиарупентным) .
10

2. Испытуемое соединение смешивают с микроорганизмом дикого типа.

15 3. Микроорганизм дикого типа вводят в окружающую среду (в присутствии или отсутствии испытуемого соединения) .

4. Если микроорганизм дикого типа неспособен адаптироваться в окружающей среде (после обработки соединением или в его присутствии) , то такое соединение рассматривается, как понижающее способность микроорганизма к адаптации или выживанию в
20 окружающей среде.
25

Если окружающая среда представляет собой организм животного, а микроорганизм представляет собой патогенный микроорганизм, то соединения, идентифицированные указанным способом, могут использоваться в качестве лекарственного средства, предотвращающего или снижающего заражение микроорганизмом.
30

35 В связи с этим, десятый аспект изобретения предусматривает соединение, идентифицируемое с помощью способа по девятому аспекту.
40

Следует иметь в виду, что применения соединения по десятому аспекту относятся к способу его идентификации, в особенности в отношении хозяина патогенного микроорганизма. Так, например, если соединение идентифицируется с помощью способа, в
45

50

котором используется вирулентный ген или закодированный им полипептид из бактерии, инфицирующей млекопитающее, такое соединение может использоваться для его лечения.

Аналогичным образом, если соединение, идентифицируется способом, в котором используется вирулентный ген или закодированный им полипептид, из гриба, который заражает растение, такое соединение может применяться для его лечения.

Согласно одиннадцатому аспекту настоящего изобретения предусматривается молекула, которая селективно взаимодействует и существенно ингибирует функцию гена по седьмому аспекту изобретения или его нуклеиновокислотного продукта.

Под термином "его нуклеиновокислотный продукт" подразумевается любая РНК, особенно мРНК, транскрибированная из гена.

Предпочтительно, молекула, селективно взаимодействующая и существенно ингибирующая функцию указанного гена или указанного нуклеиновокислотного продукта, представляет собой антисмысловую нуклеиновую кислоту или производное нуклеиновой кислоты.

Более предпочтительно, указанная молекула представляет собой антисмысловой олигонуклеотид.

Антисмысловые олигонуклеотиды представляют собой однонитевую нуклеиновую кислоту, которая может специфично связываться с последовательностью комплементарной нуклеиновой кислоты. В результате связывания с последовательностью соответствующей мишени образуются дуплексы РНК-РНК, ДНК-ДНК или РНК-ДНК. Такие нуклеиновые кислоты часто называют "антисмысловыми", поскольку они комплементарны к смысловой или кодирующей нити гена. Недавно удалось получить тройную спираль, в которой олигонуклеотид связан с дуплексом ДНК. Было установлено, что олигонуклеотиды

способны распознавать последовательности в главном желобке двойной спирали ДНК. В результате образуется тройная спираль. Это подтверждает тот факт, что существует возможность синтеза молекул, специфичных к последовательности, которые специально связывают двужнестую ДНК посредством распознавания сайтов связывания водорода в главном желобке.

Очевидно, что последовательность антисмысловой нуклеиновой кислоты или олигонуклеотида может быть легко определена со ссылкой на нуклеотидную последовательность рассматриваемого гена. Так, например, могут быть сконструированы антисмысловая нуклеиновая кислота или олигонуклеотид, которые комплементарны к части последовательности, показанной на фиг.11 или 12, особенно к последовательностям, которые образуют часть вирулентного гена.

Олигонуклеотиды представляют собой объект, который деградируется или инактивируется клеточными эндогенными нуклеазами. Для решения такой проблемы можно использовать модифицированные олигонуклеотиды, например, содержащие измененные интернуклеотидные связи, в которых фосфодиэфирные связи природного происхождения заменены на другую связь. Так, например, Agrawal с сотр. (1988) Proc.Natl.Acad.Sci. США, 85, 7079-7083 продемонстрировали повышенное ингибирование в тканевой культуре H1Y-1 с использованием фосфорамидатов и фосфоротиоатов олигонуклеотида. Sarin с сотр. (1988) в Proc.Natl.Acad.Sci. США, 85, 7448-7451 продемонстрировали повышенное ингибирование H1Y-1 с использованием метилфосфонатов олигонуклеотида. Agrawal с сотр. (1989) в Proc.Natl.Acad.Sci. США, 86, 7790-7794 продемонстрировали ингибирование репликации H1Y-1 как в ранее инфициро-

ванных, так и в хронически инфицированных клеточных культурах с использованием фосфоротиоатов олигонуклеотида специфичного к последовательности нуклеотида. Leither с сотр. (1990) в Proc. Natl. Acad. Sci. США, 87, 3430-3434 сообщают об ингибировании в тканевой культуре репликации вируса гриппа с помощью фосфотиоатов олигонуклеотида. Было показано, что олигонуклеотиды, содержащие искусственные связи устойчивы к деградации in vivo. Так например, известно из Shaw с сотр. (1991) в Nucleic Acids Res. 19, 747-750, что немодифицированные олигонуклеотиды становятся более устойчивыми к действию нуклеаз in vivo в том случае, когда они блокируются на 3'-конце некоторыми кэпированными структурами и что некэпированные олигонуклеотидные фосфотиоаты не деградируются in vivo.

Подробное описание Н-фосфонатного подхода к синтезу олигонуклеозидных фосфотиоатов приведено Agrawal и Tanq (1990) в Tetrahedron Lett. 31, 7541-7544 на которое ссылаются в тексте настоящего описания. Синтезы олигонуклеозидных метилфосфонатов, фосфородитиоатов, фосфорамидатов, сложных фосфатных эфиров, мостиковых фосфорамидатов и мостиковых фосфотиоатов известны в данной области. См., например, статьи Agrawal и Goodchild (1987) Tetrahedron Lett. 28, 3539; Nielsen с сотр. (1988) в Tetrahedron Lett. 29, 2911; Jager с сотр. (1988) Biochemistry 27, 7237; Uznanski с сотр., (1987) в Tetrahedron Lett. 28, 3401; Bannwarth (1988) Helv. Chim. Acta 71, 1517; Crosstick и Vyle (1989) в Tetrahedron Lett. 30, 4693; Agrawal с сотр. (1990) в Proc. Natl. Acad. Sci. США 87, 1401-1405, на содержание которых ссылаются в тексте. Возможны и другие способы получения. В соответствии с предпочтительным техническим решением

олигонуклеотид представляет собой дезоксирибонуклеиновую кислоту (ДНК), хотя могут также синтезироваться и применяться по-
следовательности рибонуклеиновой кислоты (РНК).

Олигонуклеотиды, используемые в настоящем изобретении, предпочтительно предназначены для сопротивления деградации под действием эндогенных нуклеолитидных ферментов. Деградация олигонуклеотидов in vivo приводит к образованию продуктов распада олигонуклеотида меньшей длины. Такие продукты распада, по всей вероятности, принимают участие в неспецифической гибридизации, и маловероятно, что они являются эффективными в отношении их копий полной длины. Таким образом, желательно использовать олигонуклеотиды, обладающие устойчивостью к деградации в организме и способные достигать клеток-мишеней. Олигонуклеотидам настоящего изобретения может быть придана повышенная устойчивость к деградации in vivo путем замены нативных фосфодиэфирных связей на одну или более внутренних искусственных интернуклеотидных связей, например, путем замещения в связи фосфата серой. Примерами связей, которые могут использоваться в таких операциях могут служить фосфоротиааты, метилфосфонаты, сульфон, сульфат, кетил, фосфородитиоаты, различные фосфорамидаты, сложные фосфатные эфиры, мостиковые фосфоротиааты и мостиковые фосфорамидаты. Приведенные примеры носят иллюстрационный характер, не ограничивающие объем настоящего изобретения. См., например, работу Cohen (1990) в Trends in Biotechnol. Синтез олигонуклеотидов с одной или более таких связей, которыми заменены фосфодиэфирные интернуклеотидные связи, хорошо известен в данной области, и он включает синтетические маршруты получения олигонуклеотидов со смешанными интернуклеотидными связя-

ми.

Могут быть получены олигонуклеотиды, устойчивые к удлинению под действием эндогенных ферментов путем "кэппирования", или включения аналогичных групп на 5' или 3' терминальных нуклеотидах. Реагент для кэппирования коммерчески доступен в виде препарата Amino-Link 11TM от Эпплайд Биосистемс Инк., Фостер Сити, КА. Способы кэппирования описаны, например, Shaw с сотр. (1991) в Nucleic Acids Res. 19, 747-750 и Agrawal с сотр. (1991) в Proc. Natl. Acad. Sci США 88(17), 7595-7599, на содержание которых ссылаются в настоящем описании.

Еще один способ получения олигонуклеотидов, устойчивых к воздействию нуклеазы, представляет собой их "само-стабилизацию", как описано Tang с сотр. (1993) Nucl. Acids. Res. 21, 2729-2735, причем на содержание этой статьи ссылаются в настоящем описании. Самостабилизированные олигонуклеотиды имеют шпилькообразные петлевые структуры на их 3'-окончаниях и проявляют повышенную устойчивость к деградации под действием фосфодиэстеразы змеиного яда, ДНК полимеразы 1 и фетальной коровьей сыворотки. Самостабилизированная область олигонуклеотида не препятствует гибридизации с комплементарными нуклеиновыми кислотами, а исследования фармакокинетики и стабильности на мышах показало повышение in vivo устойчивости самостабилизированных олигонуклеотидов в отношении их линейных копий.

В соответствии с настоящим изобретением, свойственная антисмысловым олигонуклеотидам, связующая специфичность, характерная для объединения оснований, усиливается ограничением доступности антисмыслового соединения к его предполагаемому покусу in vivo, что позволяет использовать пониженные дози-

ровки и минимизировать системные эффекты. Так, например, олигонуклеотиды применяют локально для достижения желаемого эффекта. Концентрация олигонуклеотидов в желаемом локусе много
5 выше, чем при системном применении олигонуклеотидов, и терапевтический эффект может достигаться с использованием значительно меньшего общего количества. Высокая локальная концентрация олигонуклеотидов усиливает проникновение клеток-мишеней и эффективно блокирует трансляцию последовательностей нуклеино-
10 во-кислотных мишеней.

Олигонуклеотиды могут доставляться в локус с помощью любых средств, подходящих для локализованного применения лекарства. Так, например, раствор олигонуклеотидов может быть инъе-
20 цирован непосредственно в определенное место или может доставляться вливанием с использованием насоса для вливания. Олигонуклеотиды могут также вводиться в имплантируемое устройство, которое при помещении на желаемый участок обеспечивают выделе-
25 ние олигонуклеотидов в окружающий локус.

Наиболее предпочтительно, вводить такие олигонуклеотиды с помощью гидрогельного материала. Гидрогель является огнестой-
30 ким и биodeградируемым материалом. В настоящее время известно много таких материалов, включая те, что получают из природных и синтетических полимеров. В соответствии с предпочтительным техническим решением в способе используется гидрогель, представляющий собой жидкость при температуре ниже температуры те-
40 ла, но загущающийся с образованием сохраняющего форму полутвердого гидрогеля при температуре тела или ниже этой температуры. Предпочтительными гидрогелями являются полимеры из повторяющихся звеньев оксид этилена - оксид пропилена. Свой-
50

ства такого полимера зависят от его молекулярного веса и относительного содержания в нем полиэтиленоксида и полипропиленоксида. Предпочтительные гидрогели содержат от примерно 10 до, примерно, 80% вес.оксида этилена и от примерно 20 до примерно 90% вес.оксида пропилена. Особенно предпочтительный гидрогель содержит около 70% полиэтиленоксида и 30% полипропиленоксида. Используемые для этих целей гидрогели могут быть получены, например, от BASF Corp., Парсиппани, Нью-Джерси, под торговым наименованием Плюроник (Pluronic®) .

В соответствии с таким техническим решением гидрогель охлаждают до жидкого состояния и олигонуклеотиды примешивают к жидкости до концентрации около 1 мг олигонуклеотида на грамм гидрогеля. Полученную в результате смесь затем применяют на поверхности подлежащей обработке, например, путем распыления или с помощью мазков в ходе хирургической операции или с использованием катетера или методов эндоскопии. По мере нагревания полимера он загустевает с образованием геля, и олигонуклеотиды диффундируют из такого геля в окружающие клетки в течение периода времени, определяемого точным составом геля.

Олигонуклеотиды могут применяться с помощью других имплантантов, выпускаемых промышленностью, или описанных в научной литературе, включающих липосомы, микрокапсулы и имплантируемые устройства. Так, например, имплантанты из таких биodeградируемых материалов, как полиангидриды, полиортоэфиры, полимолочная кислота и полигликолевая кислота, а также их сополимеры, коллаген, белковые полимеры, или такие небiodeградируемые материалы, как этиленвинил ацетат (EVAc), поливинилацетат, этилен виниловый спирт и их производные, могут использо-

ваться для локальной доставки олигонуклеотидов. Олигонуклеотиды могут вводиться в материал в ходе его полимеризации или отвердевания с использованием методов испарения расплава или растворителя, либо механически смешиваться с материалом. Согласно одному из технических решений, олигонуклеотиды примешивают или применяют на покрытиях для имплантируемых средств таких, как шарики из оксида кремния, покрытые декстраном, стенты или катетеры.

Доза олигонуклеотидов зависит от их размера и цели их применения. Как правило, интервал дозировок рассчитывают на основе площади поверхности ткани, подлежащей обработке. Эффективная дозировка олигонуклеотида в некоторой степени зависит от длины и химического состава олигонуклеотида, но, как правило, она составляет 30–3000 мкг на квадратный сантиметр площади поверхности ткани.

Олигонуклеотиды могут применяться на пациенте системно, как в терапевтических, так и в профилактических целях. Олигонуклеотиды могут применяться любым эффективным методом, например, парентерально (например, внутривенно, подкожно, внутримышечно) или орально, назально, либо другими методами, которые обеспечивают доступ и циркуляцию олигонуклеотидов в кровеносной системе пациента. Олигонуклеотиды, применяемые системно, предпочтительно, дополняют местное применение олигонуклеотидов, однако их можно использовать и без местного применения. Дозировка в интервале примерно 0,1–10 г при применении на взрослом человеке обычно оказывается эффективной для такой цели.

Следует иметь в виду, что молекулы, согласно этому аспекту изобретения, используются для лечения или профилактики лю-

бого заражения, вызываемого микроорганизмом, из которого выделен указанный ген, или он находится в близком родстве с указанным микроорганизмом. Таким образом, указанная молекула представляет собой антибиотик.

Таким образом, двенадцатый аспект настоящего изобретения предусматривает молекулу одиннадцатого аспекта изобретения, предназначенную для применения в медицинских целях.

Тринадцатый аспект настоящего изобретения охватывает способ лечения хозяина, который инфицирован или подвержен инфицированию микроорганизмом, причем такой способ включает применение эффективного количества молекулы в соответствии с одиннадцатым аспектом изобретения, в которой указанный ген присутствует в указанном микроорганизме или является близко родственным указанному микроорганизму.

Под термином "эффективное количество" подразумевается количество, которое в значительной мере предотвращает или улучшает состояние при заражении. Под термином "хозяин" подразумевается любое животное или растение, которые могут инфицироваться микроорганизмом.

Следует принимать во внимание, что фармацевтические рецептуры молекулы по одиннадцатому аспекту изобретения составляют его часть. Такие фармацевтические рецептуры включают указанную молекулу в сочетании с одним или более приемлемых носителей. Носитель должен быть "приемлемым" в том смысле, что он совместим с указанной молекулой изобретения и не оказывает вредного влияния на его реципиента. Обычно в качестве носителей используют воду или физиологический раствор, которые должны быть стерильными и непирогенными.

Как отмечалось выше и более подробно описано ниже в примере 4, было обнаружено, что некоторые вирулентные гены характеризуются в Salmonella typhimurium в области хромосомы, названной VGC2. ДНК-ДНК гибридизационные эксперименты позволили установить, что последовательности гомологичные, по крайней мере, части VGC2 обнаруживаются во многих видах и штаммах Salmonella, но не присутствуют в E.coli и Shigella штаммах, подвергнутых тестированию (см. пример 4). Такие последовательности почти точно соответствуют консервативным генам, по крайней мере, в Salmonella, и лишь некоторые из них являются вирулентными генами. Предполагается, что эквивалентные гены в других разновидностях Salmonella и эквивалентные гены, в случае их наличия, в других кишечных, либо других бактериях также являются вирулентными генами.

Можно легко определить, является ли ген в области VGC2 вирулентным геном. Так, например, те гены в VGC2, которые были идентифицированы способом по второму аспекту изобретения (при использовании Salmonella typhimurium, и в котором окружающая среда представляла собой такое животное, как мышь) являются вирулентными генами. Вирулентные гены также идентифицируются осуществлением мутации в гене (предпочтительно неполярной мутации) и установлением факта авирулентности мутантного штамма. Способы осуществления мутаций в отобранном гене хорошо известны и описаны ниже.

Четырнадцатый аспект изобретения предусматривает VGC2 ДНК Salmonella typhimurium или ее часть, либо вариант указанной ДНК или вариант его части.

VGC2 ДНК Salmonella typhimurium изображена в виде диаг-

раммы на фиг.8, и может быть легко получена из Salmonella ty-
phimurium ATCC 14028 (полученного из коллекции американских ти-
повых культур, 12301 Парлаун Драйв, Роквил, Мэрилэнд 20852,
США; и также депонированного в NCTC лаборатория службы здраво-
охранения, Колиндэйл, Великобритания под каталожным номером
NCTC 12021) с использованием информации, предоставленной в при-
мере 4. Так, например, зонды, являющиеся производными последо-
вательностей, показанных на фиг.11 и 12, могут использоваться
для идентификации λ клонов из Salmonella typhimurium геномной
библиотеки. Стандартные способы скрининга генома могут исполь-
зоваться для получения всех VGC2 ДНК. Рестрикционная карта, по-
казанная на фиг.8, может использоваться для идентификации и
определения местоположения ДНК фрагментов из VGC2.

Под термином "часть VGC2 ДНК Salmonella typhimurium" под-
разумевается любая последовательность ДНК, содержащая, по край-
ней мере, 10 нуклеотидов, предпочтительно, по крайней мере, 20
нуклеотидов, более предпочтительно, по крайней мере, 50 нукле-
отидов, и еще более предпочтительно, по крайней мере, 100 нук-
леотидов, и наиболее предпочтительно, по крайней мере, 500 нук-
леотидов VGC2. Особенно предпочтительной частью VGC2 ДНК явля-
ется последовательность, показанная на фиг.11, или ее часть.
Другой, особенно предпочтительной частью VGC2 ДНК, является
последовательность, показанная на фиг.12, или ее часть.

Выгодно, когда часть VGC2 ДНК представляет собой ген или
его часть.

Гены могут быть идентифицированы в области VGC2 методами
статистического анализа открытых рамок считывания генетической
информации с использованием известных компьютерных программ.

Если открытая рамка считывания имеет размер более 100 кодонов, то, по всей вероятности, она является геном (хотя известны бо-
лее мелкие гены). Соответствие открытой рамки считывания поли-
пептид-кодирующей области гена может быть установлено экспери-
ментально. Так, например, часть ДНК, соответствующая открытой
рамке считывания, может использоваться в качестве зонда в Но-
зерн (РНК) блоттинге для установления факта экспрессирования
мРНК, которая гибридизируется в указанную ДНК; альтернативно
или дополнительно мутация может быть введена в открытую рамку
считывания, и может быть определен эффект мутации на фенотип
микроорганизма. Если фенотип изменяется, то открытая рамка
считывания соответствует гену. Способы идентификации генов в
последовательности ДНК известны в этой области.

Под термином "вариант указанной ДНК или вариант ее час-
ти" подразумевается любой вариант, подпадающий под термин
"вариант" по седьмому аспекту изобретения.

Таким образом, варианты VGC2 ДНК Salmonella typhimurium
включают эквивалентные гены или их части из других видов, та-
ких как Salmonella typhi и Salmonella enterica, а также экви-
валентные гены или их части из таких других бактерий, как ки-
шечные бактерии.

Под термином "эквивалентный ген" подразумеваются гены, ко-
торые функционально эквивалентны, а также те, в которых мута-
ция приводит к аналогичному фенотипу (например, авирулентнос-
ти). Следует принять во внимание, что до настоящего изобре-
тения VGC2 или, содержащиеся в нем гены не были идентифицирова-
ны и, разумеется, не охарактеризовывались с точки зрения ви-
рулентности.

Таким образом, дальнейшие аспекты настоящего изобретения предусматривают мутантную бактерию, в которой, если в обычном состоянии содержит ген, являющийся идентичным или эквивалентным гену в VGC2, то указанный ген мутируется или отсутствует в указанной мутантной бактерии; способы получения мутантной бактерии согласно которым, если бактерия в естественном состоянии содержит ген, идентичный или эквивалентный гену в VGC2, то указанный ген мутируется или отсутствует в указанной мутантной бактерии. Далее описывается предпочтительный способ инактивации VGC2 гена. Вначале проводят субклонирование гена на фрагменте ДНК из библиотеки ДНК Salmonella λ или другой библиотеки ДНК с использованием фрагмента VGC2 в качестве зонда в экспериментах по гибридизации, затем ген картируют относительно сайтов рестрикционного фермента и характеризуют ген секвенированием ДНК в Escherichia coli. С использованием рестрикционных ферментов, затем осуществляют введение в кодирующую область гена сегмента ДНК, кодирующего устойчивость к антибиотикам (например, к канамицину), возможно после делеции части, кодирующей области клонированного гена с помощью рестрикционных ферментов. Имеются в распоряжении способы и конструкции ДНК, содержащие маркер антибиотической устойчивости, гарантирующие тот факт, что инаktivация гена, предпочтительно, является неполярной, т.е. не оказывает влияния на экспрессию генов ниже по ходу транскрипции от интересующего гена. Затем мутантную версию гена переносят с E.coli на Salmonella typhimurium с использованием трансдукции фага P22 и трансдуктантов, проведенных Саузерн гибридизацией на гомологическую рекомбинацию мутантного гена в хромосому.

Такой подход обычно используется на Salmonella (и может применяться в S.typhi), и дополнительные подробности могут быть найдены во многих статьях, например, в работе Galan с сотр. (1992) 174, 4338-4339.

Еще одним аспектом изобретения является применение указанной мутантной бактерии в вакцине, фармацевтические композиции, содержащие указанную бактерию, и фармацевтически приемлемый носитель, полипептид, закодированный VGC2 ДНК Salmonella typhimurium или ее частью, или вариант его части, способ идентификации соединения, снижающего способность бактерии инфицировать или вызывать заболевание хозяина; соединение, идентифицируемое указанным способом, молекулу, которая селективно взаимодействует и существенно ингибирует функцию гена в VGC2 или его нуклеиновом продукте; а также применение в медицине фармацевтической композиции.

VGC2 содержит гены, которые идентифицированы способами по первому и второму аспектам изобретения, а также гены, которые были идентифицированы по их местоположению (хотя они могут быть идентифицированы способами по первому и второму аспектам изобретения). Эти дополнительные аспекты изобретения весьма близки к четвертому, пятому, шестому, седьмому, восьмому, девятому, десятому, одиннадцатому, двенадцатому и тринадцатому аспектам изобретения и, в соответствии с этим, информация, касающаяся этих аспектов и предпочтения, отмеченные в отношении этих аспектов изобретения, применимы к таким дополнительным аспектам.

Предпочтительно, если используемый ген является геном VGC2, или представляет собой эквивалентный ген из других видов Salmonella, например, S.typhi. Предпочтительно, когда мутант-

ная бактерия представляет собой мутант S.typhimurium или мутант других видов Salmonella, например, S.typhi.

Предполагается, что, по крайней мере, некоторые гены из VGC2 обеспечивают способность таких бактерий, как S.typhimurium к проникновению клеток.

Далее настоящее изобретение будет описано со ссылкой на следующие примеры и фигуры, где:

Фиг.1 иллюстрирует диаграмму одного из наиболее предпочтительных способов настоящего изобретения.

Фиг.2 демонстрирует Саузерн гибридизационный анализ ДНК из 12 S.typhimurium эксконъюгантов с последующим перевариванием с EcoRV. Фильтр зондировали с помощью канамицин-устойчивого гена мини-Tn5-транспозона.

На фиг.3 продемонстрирован блоттинг-гибридизационный анализ ДНК на колонии из 48 S.typhimurium эксконъюгантов с половины микротитрационной чаши (A1-H6). Фильтр гибридизировали с зондом, содержащим меченные амплифицированные метки ДНК, выделенной из пула первых 24 колоний (A1-D6).

На фиг.4 изображен блоттинг-гибридизационный анализ ДНК на колонии из 95 S.typhimurium эксконъюгантов с микротитрационной чаши (A1-H11), которые инъецировали в мышей. Репликатные фильтры гибридизировали с мечеными, амплифицированными метками из пула (инокулятная структура), или с мечеными амплифицированными метками из ДНК, изолированной из более, чем 10,000 объединенных колоний, которые выделяли из селезенки инфицированного животного (образец селезенки). Колонии B6, A11 и C8 давали сигналы слабой гибридизации на обеих сериях фильтров. Гибридизационные сигналы от колоний A3, C5, G3 (azOA)

и F10 присутствуют в инокулянтном образце, но отсутствуют в образце селезенки.

5 На фиг.5 изображена последовательность гена Salmonella, выделенного с использованием способа изобретения, и данные сравнения с Escherichia coli c1p протеазным геномом.

10 Фиг.6 изображает частичные последовательности дополнительного гена Salmonella, выделенного с использованием способа изобретения (SEQ ID N°N° 8-36).

15 На фиг.7 изображено картирование VGC2 на хромосоме S.typhimurium. (A) ДНК-зонды из трех областей VGC2 использовали в Саузерн гибридизационном анализе лизатов из серии штаммов S.typhimurium, закрепленных в Mud-P22 профагах. Показаны лизаты, которые гибридизировались с PstI фрагментом размером в 7,5 т.п.о. (зонд А на фиг.8). Два других используемых зонда гибридизировались с такими же лизатами. Точки инсерции и направления упаковки фага показаны вдоль положения карты в минутах (вариант УШ, ссылка 22 в примере 4). Обозначения фагов соответствуют следующим штаммам: 18Р, ТТ15242; 18Q, 15241; 19Р, ТТ15244; 19Q, ТТ15243; 20Р, ТТ15246 и 20Q, ТТ15245 (см.ссылки в примере 4). Местоположения картированных генов представлены горизонтальными полосами и указаны примерные локализации других генов.

40 На фиг.8 представлена физическая и генетическая карта VGC2. (A) Положения 16 транспозоновых вставок показаны верхней линией. Протяженность VGC2 указана толстой линией. Положение и направление транскрипции ORF, описанных в тексте примера 4, показаны стрелками, которые находятся ниже линии вместе с названиями аналогичных генов, за исключением ORF 12 и 13, продукты

которых аналогичны сенсорным и регуляторным компонентам соответственно, из множества двухкомпонентных регуляторных систем.

(В) местоположение перекрывающихся клонов и EcoRI/XbaI рестрикционного фрагмента из Mud-P22 профагового штамма TT15244 показаны жирными линиями. Показаны лишь части λ клонов, которые были картированы и такие клоны могут распространяться за рамки указанных пределов. (С) Указаны положения рестрикционных сайтов: B, BamHI; E, EcoRI; Y, EcoRY; H, Hind III; P, PstI и X, XbaI. Положения 7,5 т.п.о. PstI фрагмента (зонд А), используемого в качестве зонда на фиг.7 и 2,2 т.п.о PstI/Hind III фрагмента (зонд В), используемого в качестве зонда на фиг.10, указаны внизу рестрикционной карты. Положения последовательности 1 (описанной на фиг.11) и последовательности 2 (описанной на фиг.12) показаны тонкими стрелками (меченые последовательность 1 и последовательность 2).

На фиг.9 представлено картирование границ VGC2. (А) положения картированных генов на 37-38 минутах на E.coli K12 хромосоме приведены в соответствии с соответствующей областью S. typhimurium LT2 хромосомы (минуты 30-31). Расширенная карта VGC2 области показана в виде ДНК фрагментов 11 S.typhimurium/S.t., используемых в качестве зондов (толстые полосы) и рестрикционными сайтами, используемыми для их генерации: B, BamHI; C, ClaI; H, Hind II; K, KpnI; P, PstI; N, NsiI и S, SalI. Зонды, которые гибридизуются геномной ДНК E.coli K12 (E.c.) обозначены символом +; зонды, неспособные к гибридизации обозначены символом -.

Фиг.10 показывает, что VGC2 законсервирована и специфична к Salmonella. Геномная ДНК из Salmonella сероваров и дру-

гих патогенных бактерий подвергалась рестрикции с PstI (A),
Hind III или EcoRY (B) и подвергалась Саузерн гибридизационному
анализу с использованием 2,2 kb PstI/Hind III фрагмента из λ
клона 7 в качестве зонда (зонд B на фиг.2). Фильтры гибридизи-
ровали и промывали в строгих (A) и не строгих (B) условиях.

На фиг.11 показана последовательность ДНК "Последователь-
ности 1" VGC2 от центра до левого конца (см.отмеченную стрел-
кой последовательность 1 на фиг.2). ДНК транслировали во всех
шести открытых рамках считывания и указаны стартовые и оста-
навливающие положения предполагаемых генов, а также положения
инсерции транспозона для различных мутантов идентифицированных
с помощью STM (SEQ ID N° 37).

Как общепринято, символ^{*} обозначает стоп-кодон и где не-
обходимо, использованы коды нуклеотидной неопределенности.

На фиг.12 показана последовательность ДНК "последователь-
ности 2" VGC2 (кластер C) (см.отмеченную стрелкой последовате-
льность 2 на фиг.2). ДНК транслировали во всех шести рамках
считывания и отмечали положения старта и остановки предполага-
емых генов, а также положения инсерции транспозона для различ-
ных мутантов идентифицированных STM (SEQ ID N°38).

Как общепринято, индекс^{*} обозначает стоп-кодон, и где не-
обходимо, используются стандартные коды нуклеотидной неопреде-
ленности.

Фиг.7-12 большей частью относятся к примеру 4.

ПРИМЕР 1

Идентификация вирулентных генов в Salmonella typhimurium

Материалы и методы

Бактериальные штаммы и плазмиды

Salmonella typhimurium штамм 12023 (эквивалентный штамму 14028 из американской коллекции типовых культур (ATCC)) был получен от Национальной коллекции типовых культур (NCTC), Сервисной лаборатории здравоохранения, Колиндэйл, Лондон, Великобритания. Спонтанный мутант, устойчивый к налидиксовой кислоте указанного выше штамма (12023 Nal^r), был отобран в лаборатории авторов изобретения. Другое производное штамма 12023, CL 1509 (arOA:Tn10) было получено от Fred Heffron. Escherichia coli штаммы CC118 λ pir(Δ [ara -leu], araD, Δ lac X74, galE, galK, phoA20, thi-1, rpsE, rpoB, argE/Am/, recA1, λ pir фаговый лизоген/ и S17-1 λ pir /Tr^r, Sm^r, recA, thi, pro, hsdR⁻M⁺, RP4:2-Tc:Mu:Km Tn7, λ pir / были выделены фирмой Kenneth Timmis. E.coli DH5 $\mathcal{L}$ использовали для размножения pUC18 /Гибко-BRL/ и Bluescript(Стратаген) плазмид, содержащих S. typhimurium ДНК. Плазмида pUTмини-Tn5Km2(de Lorenzo с сотр., 1990) была получена от Kenneth Timmis.

Конструкция меток с полурандомизированнымипоследовательностями и лигирования

Олигонуклеотидный пул RT1 (5'-СТАГГТАССТАСААССТСААГСТТ-/NK/20-AAGСТТGGТТАГААТGGGTACCATG-3'//SEQ 1Д № 1) и праймеры P2 (5'-ТАССТАСААССТСААГСТ-3') (SEQ 1Д № 2), P3 (5'-CATGGТАССАТТСТААС-3') (SEQ 1Д № 3), P4 (5'-ТАССАТТСТААССААГС-3') (SEQ 1Д № 4) и P5 (5'-СТАГГТАССТАСААССТС-3') (SEQ 1Д № 5) синтезировали на олигонуклеотидном синтезаторе (Эпплайд Биосистемс, модель 380В). Метки двухнитевой ДНК готовили из RT1 в

100 мкл объеме PCR, содержащем 1,5 мМ $MgCl_2$, 50 мМ KCl и 10 мМ Трис-Cl (pH 8,0) в присутствии 200 пг RT1 в качестве мишени; по 250 мкМ каждого из dATP, dCTP, dGTP, dTTP; 100 пМ праймеров P3 и P5; и 2,5 ед. Ampli¹taq (Перкин-Элмер Петус). Условия циклической термообработки были следующими: 30 циклов при 95°C в течение 30 сек., 50°C в течение 45 сек. и 72°C в течение 10 сек. Продукт PCR очищали на геле (Sambrook с сотр. 1989), пропускали через колонку elutip D (полученную от Sch¹eicher и Schull) расщепляли с помощью KpnI перед лигированием в pUC18 или pUTмини-Tn5Km2. Для лигирований, плазмиды переваривали с KpnI и дефосфорилировали с телячьей кишечной щелочной фосфатазой (Гибко-BRL). Молекулы линеаризованной плазмиды очищали на геле (Sambrook с сотр., 1989) перед лигированием для удаления любой остаточной неразрезанной в процессе расщепления плазмидной ДНК. Реакционные среды для лигирования содержали примерно 50 нг каждой из плазмид и двухнитевую ДНК-метку в 25 мкл объеме в присутствии 1 единицы T4 ДНК лигазы (Гибко-BRL) в буфере, содержащем фермент. Лигирования проводили в течение 2 часов при 24°C. Для определения количества бактериальных колоний, возникающих в результате самолигирования плазмидной ДНК, или из неразрезанной плазмидной ДНК, проводили контрольную реакцию, в которой в реакционную среду для лигирования не вводили двухнитевую ДНК-метку. В результате этого не получали ампицилин-устойчивых бактериальных колоний после трансформации E.coli CC118 (Sambrook с сотр., 1989) по сравнению с образованием 185 колоний в среде для лигирования, содержащей двухнитевую ДНК-метку.

Бактериальная трансформация и спаривания

Продукты нескольких лигирований между рUT мини-Tn5Km2 и
5 двухнитевой ДНК-меткой использовали для трансформации E.coli
CC118 (Sambrook с сотр., 1989). Составляли пул из примерно
10,300 трансформантов и плазмидную ДНК, экстрагированную из
10 пула, использовали для трансформации E.coli S-17 λ pir (de
Lorenzo, Timmis, 1994). Для экспериментов по спариванию пул
15 из примерно 40,000 ампициллин устойчивых E.coli S-17 λ pir
трансформантов и S.typhimurium 12023 Nal^r индивидуально куль-
тивировали до оптической плотности $(\text{ОД})_{580}$ порядка 1,0. Алик-
20 воты каждой культуры (0,4 мл) смешивали в 5 мл 10 мМ MgSO_4 и
фильтровали через мембрану Миллипор (с диаметром 0,45 мкм).
Такие фильтры помещали на поверхность агара, содержащего М9
25 соли (de Lorenzo и Timmis, 1994) и инкубировали при 37°C в
течение 16 часов. Бактерии были выделены встряхиванием филь-
тров в жидкой LB среде в течение 40 минут при 37°C, и эксконъю-
30 ганты отобраны путем посева суспензии на LB среду, содержащую
100 мкг мл⁻¹ налидиксовой кислоты (для селекции от донорного
штамма) и 50 мкг мл⁻¹ канамицина (для селекции на реципиент-
35 ный штамм). Каждый эксконъюгант проверяли путем переноса на-
лидиксово-кислотно устойчивых (nal^r), канамицин устойчивых
(kan^r) колоний на МакКонки Лактозную индикаторную среду (для
40 различия между E.coli и S.typhimurium) и на LB среду, содер-
жащую ампициллин. Примерно 90% nal^r , kan^r колоний оказались
45 чувствительными к ампициллину, что является указанием на их
образование в результате аутентичных транспозиционных собы-
тий (de Lorenzo, Timmis, 1994). Индивидуальные ампициллин-

чувствительные эксконъюганты были помещены на 96-луночные микротитрационные планшеты, содержащие LB среду. При длительном хранении при -80°C в среду вводили 7% ДМСО или 15% глицерина.

Характеристика фенотипа мутантов

Мутанты пересевали при помощи штампа ζ микротитрационных планшетоу на твердую среду, содержащую M9 соли и 0,4% глюкозы (Sambrook с сотр., 1989) с целью идентификации ауксотрофов. Мутанты, колонии которых имеют ярко выраженную морфологию, определяют путем исследований колоний на агаровых чашках под микроскопом низкого разрешения.

Блоттинги колоний, экстракции ДНК, РС R, мечения

ДНК и гибридизации

Для блоттинг-гибридизаций колоний использовали 48-луночный металлический репликатор (Сигма) для переноса эксконъюгантов с микротитрационных планшетоу на нейлоновые фильтры Хибонд 11 (Амершам, Великобритания), которые помещали на поверхность LB агара, содержащего 50 мкг мл^{-1} канамицина. После инкубации в течение ночи при 37°C фильтры с колониями бактерий извлекали и сушили при комнатной температуре в течение 10 минут. Бактерии лизировали 0,4 н NaOH и фильтры промывали 0,5 н Трис-Cl pH 7,0 в соответствии с инструкциями производителей фильтров. ДНК бактерии фиксировали на фильтрах путем облучения в УФ-свете от Stratalinker (Стратаген). Гибридизации ^{32}P -меченными зондами осуществляли в строгих условиях, как было описано выше (Holden с сотр., 1989). Для выделения ДНК, *S. typhimurium* транспозоновые мутантные штаммы выращивали в жидкой

LB среде в микротитрационных планшетах или повторно суспендировали в LB среде с последующим выращиванием на твердой среде. Общую ДНК получали методом с использованием бромид гексадецилтриметиламмония (СТАВ), согласно методике Ausubel с сотр. (1987). Вкратце, клетки в объеме 150-1000 мкл осаждали центрифугированием и ресуспендировали в 576 мкл ТЕ. К полученной смеси добавляли 15 мкл 20% SDS (додecil сульфат Na) и 3 мкл 20 мг мл⁻¹ протеиназы К. После инкубирования при 37° в течение 1 часа добавляли 166 мкл 3М NaCl и смесь тщательно перемешивали, после чего добавляли 80 мкл 10% (вес./об.) СТАВ и 0,7М NaCl. После тщательного перемешивания раствор инкубировали при 65°С в течение 10 минут. После экстракции фенолом и смесью фенол-хлороформ ДНК осаждали добавлением изопропанола, промывали 70% этанолом и ресуспендировали в ТЕ при концентрации примерно 1 мкг мкл⁻¹.

Образцы ДНК подвергали двум циклам реакции PCR (полимеразно-цепьевая реакция) с целью получения меченных зондов. Первую PCR проводили в реакционной среде объемом 100 мкл, содержащей 20 мМ Трис-Cl pH 8,3; 50 мМ KCl; 2 мМ MgCl₂; 0,01% Твина 80; 200 мкМ каждого из dATP, dCTP, dGTP, dTTP; 2,5 единицы Amlitaq полимеразы (Перкин-Элмер Петус); по 770 нг каждого из праймеров Р2 и Р4; и 5 мкг ДНК-мишени. После начальной денатурации в течение 4 минут при 95°С проводили циклическую термообработку, включающую 20 циклов в течение 45 секунд при 50°С, 10 сек. при 72°С и 30 сек. при 95°С. Продукты PCR экстрагировали смесью хлороформ/изоамиловый спирт (24/1) и осаждали этанолом. ДНК ресуспендировали в 10 мкл ТЕ и продукты PCR очищали электрофорезом через 1,6% Seaplaque(FMC Био-

продуктс) гель в буфере ТАЕ. Тонкие срезы геля, содержащие фрагменты размером примерно 80 т.п.о. и использовали для второй PCR. Эту реакцию проводили в общем объеме 20 мкл, содержащем 20 мМ Трис-Cl pH 8,3; 50 мМ KCl; 2 мМ MgCl₂; 0,01% Твина 80; 50 мкМ каждого из dATP, dTTP, dGTP; 10 мкл ³²P-dCTP (3000 кюри/ммоль, Амершам); по 150 нг каждого из праймеров P2 и P4; примерно 10 нг ДНК-мишени (1-2 мкл 1,6% Seaplaque агарозы, содержащей продукт первого цикла PCR); 0,5 единиц Amp-litaq полимеразы. Реакционную смесь покрывали 20 мкл минерального масла и осуществляли циклическую термообработку, как описано выше, количество введенной радиоактивной метки определяли поглощением на Whatman DE81 бумагу (Sambrook с сотр. 1989).

Исследования заражения

Индивидуальные эксконъюганты Salmonella, содержащие меченные транспозоны, выращивали в 2% триптона, 1% дрожжевого экстракта, 0,92% об./об. глицерина, 0,5% Na₂PO₄, 1% KNO₃ (TVGPN среда) (Ausubel с сотр. 1987) в микротитрационных планшетах в течение ночи при 37°C.

Металлический репликатор использовали для переноса небольшого объема, выдержанных в течение ночи, культур на свежий микротитрационный планшет и культуры инкубировали при 37°C до значения OD₅₈₀ (оптическая плотность при длине волны 580 нм) (с использованием многоканального спектрофотометра для прочтения микротитрационного планшета Titertec), равного примерно 0,2 в каждой лунке. Затем культуры из отдельных лунок объединяли и с помощью спектрофотометра определяли значение OD₅₅₀. Культуру разбавляли стерильным физиологическим раствором до

примерно 5×10^5 КОЕ/мл. Дополнительные разбавления высевали на TYGPN, содержащую налидиксовую кислоту (100 мг мл^{-1}) и канамицин (50 мг мл^{-1}) для подтверждения значения КОЕ в инокуляте.

Группам из трех самок BAL B/c мышей (20-25 г) делали внутрибрюшинные инъекции 0,2 мл суспензии бактерий, содержащей примерно 1×10^5 КОЕ мл^{-1} . Мышей умертвляли через три дня после инокуляции и удаляли их селезенки для выделения бактерий. Половину каждой селезенки гомогенизировали в 1 мл стерильного физиологического раствора в пробирке микрофуги. Клеточному дебрису позволяли осесть и 1 мл физиологического раствора, содержащего клетки все еще в суспендированном состоянии, переносили в свежую пробирку и центрифугировали в течение двух минут на микрофуге. Супернатант отделяли, а осадок после центрифугирования ресуспендировали в 1 мл стерильной дистиллированной воды. Получали серии разбавлений в стерильной дистиллированной воде и 100 $\mu\text{л}$ каждого разбавления помещали на TYGPN агар, содержащий налидиксиновую кислоту ($100 \text{ }\mu\text{г мл}^{-1}$) и канамицин ($50 \text{ }\mu\text{г мл}^{-1}$). Бактерии выделяли с планшетов, содержащих 1000-4000 колоний, и создавали общий пул из примерно 10,000 колоний, выделенных с каждой селезенки и этот пул использовали для получения ДНК для PCR генерации зондов, которые применяли для скрининга пятен колоний.

Клонирование вирулентного гена и ДНК секвенирование

Общую ДНК выделяли из S.typhimurium эсконъюгантов и переваривали по отдельности с SstI, SalI, PstI и SphI. Перевары фракционировали через агароза гели, переносили на Ну-

bond N⁺ мембраны (Амершам) и подвергали саузерн-гибридизаци-
 онному анализу с использованием канамицин устойчивого гена из
 5 рU мини-Tn5Km2 в качестве зонда. Зонд метили дигоксигенином
 (Бюэрингер-Маннхейм) и осуществляли хемилюминесцентную детек-
 цию в соответствии с инструкциями производителя. Использовали
 10 описанные выше условия гибридизации и промывки. Рестрикцион-
 ные ферменты, обеспечивающие гибридизационные фрагменты раз-
 мером 3-5 т.п.о. использовали для переваривания ДНК на препа-
 15 ративном агарозовом геле и фрагменты ДНК, соответствующие раз-
 мерам сигналов гибридизации, вырезали, очищали и лигировали
 на рU C18. Реакции лигирования использовали для трансформации
 20 E.coli DH5a в канамициновую устойчивую форму. Плазмиды из ка-
 намицин-устойчивых трансформантов очищали пропусканием через
 25 колонку clutip Д и проверяли путем расщепления рестриктазами.
 Плазмидные вставки частично секвенировали ди-докси методом
 (Sanger с сотр., 1977) с использованием -40 праймера и обрат-
 30 ного секвенирующего праймера (Биохимическая корпорация США),
 а также праймеров Р6 (5'-CCTAGGCGGCCAGATCTCAT-3') (SEQ 1Д №6)
 и Р7 (5'-GCACTTGTGTGTATAAGAGTCAG-3') (SEQ 1Д № 7), которые от-
 35 жигались в 1 и 0 окончания Tn5, соответственно. Нуклеотидные
 последовательности и выведенные аминокислотные последователь-
 ности с использованием метода упаковки по программе Macvector
 40 3,5 на компьютере Макинтош SE/30. Последовательности сравни-
 вали с базами данных EMBL и Genbank ДНК с применением компью-
 терной системы UNIX/SUN в Исследовательском центре проекти-
 45 рования карты генома человека, Харроу, Великобритания.

РЕЗУЛЬТАТЫ

Строение метки

Структура ДНК-меток показана на фиг.1а. Каждая метка состоит из вариабельной центральной области, фланкированной "плечами" инвариантной последовательности. Последовательность центральной области ($(/NK/_{20})$) конструировалась таким образом, чтобы предотвратить возникновение сайтов для обычно применяемых п.о. распознающих рестрикционных ферментов, но она была достаточно вариабельной для статистического обеспечения того, что такая же последовательность может возникать лишь раз в 2×10^{11} молекул (ДНК секвенирование 12 случайно отобранных меток показало, что ни одна из них не имеет более, чем 50% идентичности в рамках вариабельного участка). (N обозначает любое основание (A, G, C или T), а K обозначает G или T). Плечи содержат KpnI сайты вблизи окончаний для облегчения стадии начального клонирования, а Hind III сайты, граничащие с вариабельным участком, используются для выделения радиомеченных вариабельных участков из плеч перед гибридизационным анализом. Кроме этого, плечи конструировали так, что каждый из праймеров P2 и P4 содержит лишь один гуаниновый остаток. Поэтому в ходе PCR с использованием таких праймеров лишь один цитозин будет вводиться в каждое вновь синтезированное плечо по сравнению с, в среднем, десятью такими остатками в уникальной последовательности. В том случае, когда радиомеченный dCTP включается в PCR, в среднем, в десять раз большее количество метки будет присутствовать в уникальной последовательности по сравнению с каждым плечом. Целью такой операции является минимизация фоновых гибридизационных сигналов от плеч после их выделения из уникальных последовательностей в результате перемаркирования с Hind III. Двухнуклеотидные метки лигировали в KpnI сайт

мини-Tn5 транспозона Km2 переносимого на плазмиду PUT (de Lorenzo и Timmis 1994). Репликация такой плазмиды зависит от R6K-специфического продукта *pir* гена. Он несет *oriT* последовательность плазмиды RP4, что обеспечивает перенос на множество бактериальных разновидностей (Miller и Mekalanos, 1988), а *tnp* ген необходим для транспозиции мини-Tn5-элемента. Меченные мини-Tn5 транспозоны переносили на S. typhimurium путем конъюгирования и 288 эксконъюгантов, образовавшихся в результате транспозиционных событий, хранили в лунках микротитрационных планшетов. Общую ДНК, выделенную из 12 таких конъюгантов переваривали с *EcoRV* и подвергали саузерн-гибридизационному анализу с использованием канамицин-устойчивого гена мини-Tn5 транспозона в качестве зонда. В каждом случае эксконъюгант возникал в результате единичной интеграции транспозона в различные сайты бактериального генома (фиг.2).

Исследования специфичности и чувствительности

Далее определялась эффективность и однородность амплификации ДНК-меток в реакциях PCR с использованием пулов эксконъюгантных ДНК в качестве мишеней реакций. С целью минимизации неравной амплификации меток в PCR, определялось максимальное количество ДНК-мишени, которое следует использовать в 100 мкл реакционной среды и минимальное число циклов PCR, в результате которых образуются продукты, которые могут визуализироваться путем окрашивания бромистым этидием на агарозовом геле (5 мкг ДНК и 20 циклов, соответственно).

S. typhimurium эксконъюганты, которые достигли фазы стационарного роста в микротитрационных планшетах объединяли и

использовали для экстракции ДНК. Ее подвергали PCR с использованием праймеров P2 и P4. PCR продукты размером 80 п.о. подвергали очистке на геле и использовали в качестве мишеней для второй PCR с применением тех же праймеров, но в присутствии ^{32}P -меченного СТР. В результате этого около 60% радиомеченого dСТР вводилось в PCR продукты. Радиомеченные продукты переваривали с HindIII и использовали в качестве зонда для колон-блоттированной ДНК из их соответствующих микротитрационных планшето́в. Из 1510 мутантов, тестированных этим методом, 358 оказались неспособными к проявлению отчетливого сигнала на автордиограмме после ночного облучения пятна колонии. Для такого явления имеется три потенциальных объяснения. Во-первых, возможно, что часть транспозонов не несет меток. Однако, при сравнении частот трансформации в результате реакций лигирования с участием транспозонов в присутствии или отсутствии меток, представляется маловероятным, что немеченные транспозоны могут составлять более примерно 0,5% от общего их числа (см. Материалы и методы). Более вероятные причины состоят в том, что переменная последовательность процессируется в некоторых метках и/или, что некоторые из последовательностей образуют вторичные структуры, прием обе таких причины должны предотвращать амплификацию. Мутанты, не дающие отчетливых сигналов, исключены из дальнейших исследований. Специфичность эффективно амплифицированных меток демонстрировали образованием зонда из 24 колоний микротитрационного планшета и его использованием для зондирования блота из 48 колоний, включающих 24 колонии, которые использовали для образования зонда. Отсутствие какого-либо гибридизационного сигнала от 24 колоний, не

использованных для образования зонда (фиг.3), показывает, что используемые условия гибридизации были достаточно строгими для предотвращения кросс-гибридизации среди радиомеченных меток и позволяет предположить, что каждый эксконъюгант не повторяется в микротитрационном планшете.

Имеются дополнительные соображения, касающиеся определения максимального размера пула, который может использоваться в качестве инокулянта в экспериментах на животных. Поскольку количество радиомеченных меток для каждого транспозона обратно пропорционально сложности меченного пула, имеется предел для размера пула, выше которого сигналы гибридизации становятся слишком слабыми для детекции после ночного облучения автордиограммы. Более важным обстоятельством является то, что поскольку вариабельность пула повышается, должна реализоваться вероятность невозможности присутствия вирулентной составляющей пула в достаточных количествах в селезенке инфицированного животного для того, чтобы получить в достаточной степени меченный зонд. Верхний предел размер пула для использованной модели мышинного сальмонеллоза не определялся, но он должен быть выше 96.

Тесты на вирулентность транспозоновых мутантов

Всего 1152 уникально меченных инсерционных мутантов (с двух микротитрационных планшетов) тестировали на вирулентность у BALB/с мышей в двенадцати пулах, каждый из которых представлял собой 96-луночный микротитрационный планшет. Животным делали внутрибрюшинную инъекцию примерно 10^3 клеток каждого из 96 транспозоновых мутантов с микротитрационного

планшета (всего 10^5 организмов). Через три дня после инъекции
 мышей умервляли и бактерии выделяли посевом гомогенатов селе-
 5 зенки на лабораторную среду. Получали пул из примерно 10,000
 колоний, выделенных от каждой мыши, и экстрагировали ДНК.
 Метки, присутствующие в таком образце ДНК, амплифицировали и
 10 метили с помощью PCR, а колониальные блоты зондировали и срав-
 нивали с гибридизационным образцом, полученным с использова-
 нием меток, амплифицированных из инокулянта (фиг.3). В качес-
 15 тве контроля ago A мутант S.typhimurium метили и использова-
 ли в качестве одного из 96 мутантов в инокулянте. Не следова-
 ло ожидать, что такой штамм будет регенерироваться в селезен-
 20 ке поскольку его вирулентность значительно ослаблена (Buch-
 meier с сотр., 1993). Идентифицировали 41 мутант, ДНК которых
 25 была гибридизирована с радиомечеными метками из инокулянта,
 но не с метками из бактерий, выделенных из селезенки. Экспе-
 римент повторяли и снова идентифицировали аналогичные 41 му-
 30 тант. Как и ожидалось два из них представляли собой agoA му-
 тант (один на пул). Другой представлял собой ауксотрофный му-
 тант (он не способен расти на минимальной среде). Все такие
 35 мутанты имели нормальную морфологию колоний.

ПРИМЕР 2

Клонирование и частичное охарактеризование последователь- ностей, фланкирующих транспозон

45 ДНК экстрагировали из одного из мутантов, описанных в
 Примере 1 (пул 1, FlO), переваривали с SstI и субклонировали
 на базе канамициновой устойчивости. Последовательность в 450
 50 п.о., фланкирующую один конец транспозона, определяли с испо-

льзованием праймера Р7. Такая последовательность демонстрировала 80% идентичность E.coli clp (lon) гену, который кодирует термо-регулируемую протеазу (фиг.5). Насколько известно, такой ген ранее не рассматривался, как вирулентная детерминанта. Неполные последовательности тринадцати Salmonella typhimurium вирулентных генов показаны на фиг.6 (последовательности А2-А9 и В1-В5). Установленные аминокислотные последовательности Р2Д6, S4C3, РЗР4, Р7 G2 и Р9В7 имели подобия семейству секреторно-ассоциированных белков, которые законсервированы бактериальными патогенами животных и растений, и которые известны в Salmonella, как семейство inv. В S. typhimurium inv гены требуются для бактериальной инвазии в кишечную ткань. Вирулентность inv мутантов уменьшается при их инокулировании оральным путем, но не при их внутрибрюшинном применении. Обнаружение inv-родственных генов, которые требуются для вирулентности после внутрибрюшинной инокуляции, предполагает новое секреторное устройство, которое может быть необходимым для инвазии нефагоцитных клеток селезенки и других органов. Продукты таких новых генов могут представлять собой лучшие лекарственные мишени, чем inv белки при лечении установленных инфекций.

Дополнительные характеристики генов идентифицированных в этом примере представлены в примере 4.

ПРИМЕР 3

Определения LD₅₀ и изучение вакцинации мышей

Мутации, идентифицированные способом изобретения, ослабляют вирулентность.

Пять из мутаций в генах, которые ранее не подразумевали вирулентность, переносили P22-регулируемой трансдукцией на родственный штамм Salmonella typhimurium 12028, чувствительный к нали-
диксиновой кислоте. Трансдуктанты проверяли рестрикционным картированием и затем внутрибрюшинно инъецировали в группы
BALB/c мышей для определения их 50% летальной дозы (LD_{50}). Значения LD_{50} для мутантов S4C3, P7G2, P3F4 и P9B7 были на несколько порядков выше соответствующих значений для штамма дикого типа. Различий в значениях LD_{50} не детектировалось для мутанта P1F10: однако имело место статистически значимое уменьшение количества P1F10 клеток, выделенных из селезенок мышей, получивших инъекции инокулянта, состоящего из равных количеств такого штамма и штамма дикого типа. Это означает, что мутация действительно, ослабляет вирулентность, однако степень такого ослабления не детектируется с помощью LD_{50} .

Мутанты P3F4 и P9B7 также применяли оральным путем при уровне инокулянта порядка 10^7 клеток/мышь. Ни одна из мышей не заболела, что указывает на то, что оральные LD_{50} таких мутантов, по крайней мере, на порядок выше соответствующих значений для штамма дикого типа.

При исследовании вакцинации мышей группы из пяти самок BALB/c мышей массой 20-25 г, вначале инокулировали орально (р.о.) или внутрибрюшинно (i.p.) серийными десятикратными разбавлениями Salmonella typhimurium мутантных штаммов P3F4 и P9B7. Затем, через четыре недели мышей инокулировали 500 коЕ родственного штамма дикого типа. Далее, в течение четырех недель регистрировали смертность.

Группу из двух мышей того же возраста и партии, что и

мыши инокулированные мутантными штаммами, также инокулирова-
ли i.p. 500 коЕ штамма дикого типа в качестве положительного
5 контроля. Как и ожидалось обе неиммунизированные мыши погибали
ли в течение четырех недель.

Полученные результаты приведены ниже в виде таблицы:

1) p.o. начальная инокуляция мутантным штаммом P3F4

начальная инокуляция, коЕ	число мышей, выживших	
	при первом заражении	живших при зара-
	:	жении диким ти-
	:	:пом
5×10^9	5	2 (40%)
5×10^8	5	2 (40%)
5×10^7	5	0 (0%)

2) i.p. начальная инокуляция мутантным штаммом P3F4

5×10^6	3	3 (100%)
5×10^5	5	4 (80%)
5×10^4	6	5 (83%)
5×10^3	5	4 (80%)

3) p.o. начальная инокуляция мутантным штаммом P9B7

5×10^9	5	0 (0%)
-----------------	---	--------

4) i.p. начальная инокуляция мутантом P9B7

5×10^6	4	2 (50%)
-----------------	---	---------

Из представленных экспериментальных данных можно сделать вывод, что мутант P3F4, по-видимому, обеспечивает некоторую защиту от последующего контрольного заражения дикого типа. Такая защита, по-видимому, выше для мышей, иммунизированных внутрибрюшинно (i.p.)

ПРИМЕР 4

Идентификация вирулентного локуса, кодирующего секреторную систему второго типа III в *Salmonella typhimurium*

В этом примере используются следующие сокращения: VGC1, вирулентный генный кластер 1; VGC2, вирулентный генный кластер 2.

Объяснение описанных экспериментов

Salmonella typhimurium является главным агентом гастроэнтерита людей и вызывает системные заболевания у мышей, которые служат моделью тифоидной лихорадки людей (1). После оральной инокуляции мышей *S. typhimurium* бактерии проходят из просвета тонкой кишки через слизистую кишечную оболочку, через энтероциты или М клетки фолликул бляшек Пейера (2). Далее бактерии поражают макрофаги и нейтрофилы, входят в ретикуло-эндотелиальную систему и распространяются по другим органам, включая селезенку и печень, дальнейшая репродукция в которых приводит к ярковыраженной и фатальной бактериемии (3). Для поражения клеток-хозяев, выживания и репликации в разнообразных физиологически стрессовых внутриклеточных и клеточных окружающих средах и для преодоления специфических антибактериальных активностей иммунной системы, *S. typhimurium* использует

сложный набор вирулентных факторов (4).

С целью более полного понимания механизмов вирулентного действия S.typhimurium и других патогенов, в примере 1 описана система мутагенеза транспозона, который для удобства назван "сигнатура-меченным мутагенезом" (STM), который объединяет возможности мутационного анализа со способностью одновременного прослеживания судьбы большого числа различных мутантов в одном животном (ссылка 5 и пример 1; ссылка 5 опубликована после даты приоритета изобретения). С использованием такого подхода авторы изобретения идентифицировали 43 мутанта с пониженной вирулентностью из общего количества в 1152 мутанта, которые были подвергнуты скринингу. Нуклеотидные последовательности ДНК, фланкирующие точки инсерции транспозонов в 5 из таких мутантов показали, что они относятся к генам, кодирующим секреторные системы типа III большого числа бактериальных патогенов (6, 7). Продукты *inv/spa* генного кластера S.typhimurium (8, 9) представляют собой белки, образующие секреторную систему типа III, требующуюся для сборки поверхностных придатков, обеспечивающей вхождение в эпителиальные клетки (10). Следовательно, вирулентность штаммов, несущих мутации в *inv/spa* кластере, ослабляется лишь в том случае, когда инокулят применяют орально, а не внутрибрюшинно (8). В отличие от этого 5 мутантов, идентифицированных методом STM, являются авирулентными после внутрибрюшинной инокуляции (5).

В настоящем примере показано, что сайты инсерции транспозона 5 таких мутантов и 11 дополнительных мутантов, идентифицированных с помощью STM, картируются в одну и ту же область S.typhimurium хромосомы. Дополнительный анализ такой

области позволил обнаружить дополнительные гены, установленные продукты которых, имеют последовательность, подобную другим компонентам секреторной системы типа Ш. Такая хромосомная область, на которую ссылаются, как на вирулентный генный кластер 2 (VG C2), отсутствует в ряде других кишечных бактерий и представляет собой важный локус для S.typhimurium вирулентности.

Материалы и методы

Бактериальные штаммы, трансдукция и растительная среда.

Salmonella enterica серотипы 5791 (aberdeen), 423180 (gallinarum), 7101 (cubana) и 12416 (typhimurium LT2) были получены из Национальной коллекции типовых культур, Служба лаборатории здравоохранения, Великобритания. Salmonella typhi BRD 123 геномная ДНК подарена G.Dougan энтерпатогенная Escherichia coli (EPEC), энтерогеморрагическая E.coli (EHEC), Vibrio cholera биотип El Tor, Shigella flexneri серотип 2 и Staphylococcus aureus представляли собой клинические изоляты, полученные из департамента инфекционных болезней и бактериологии, Королевская медицинская аспирантура, Великобритания. Геномная ДНК из Yersinia pestis получена J.Heesemann. Однако, геномная ДНК может быть выделена стандартными методами. Бактериальные штаммы и способы применяемые для создания сигнатур меченных мини-Tn5 транспозоновых мутантов S.typhimurium NCTC штамма 12023 были описаны ранее (5, 11). Общепринятое размножение плазмид осуществляли в E.coli DH5 α . Бактерии выращивали в LB бульоне (12), дополненном соответствующими антибиотиками. Перед оценкой уровней вирулентности индивидуальных мутан-

тных штаммов мутации вначале переносили фаг P22 - регулируемой трансдукцией (12) на чувствительный к налидиксовой кислоте родительский штамм S.typhimurium 12023. Трансдуктанты анализировали рестрикционным перевариванием и саузерн-гибридизацией перед использованием в качестве инокулянта.

Скрининг лямбда библиотеки. Лямбда (λ) клоны с перекрывающимися ДНК-вставками, покрывающими VGC2, получали стандартными методами (13) из λ 1059 библиотеки (14), содержащей вставки неполного Sau 3A перевара S.typhimurium LT2 геномной ДНК. Библиотеку получали через K.Sanderson из Центра по генетическому семейству Salmonella (SGSC), Калгари, Канада.

Mud-P22 лизогены. Радиомеченные ДНК-зонды гибридизировали Hybond N (Амершам) фильтрами, несущими ДНК, полученную из лизатов ряда S.typhimurium штаммов с якорно-закрепленными Mud-P22 профагами, в известных положениях в S.typhimurium генома. Получение митомицин-индуцированных Mud-P22 лизатов осуществляли, как описано в (12, 15). Набор Mud-P22 профагов был вначале собран Benson и Goldman (16) и был получен из SGSC.

Гельэлектрофорез и саузерн-гибридизация. Гельэлектрофорез проводили в 1% или 0,6% агарозных гелях в 0,5 x TBE. Гель-фракционированную ДНК переносили на Hybond N или N⁺ мембраны (Амершам) и гибридизацию в строгих условиях и процедуры промывки (позволяющими проводить гибридизацию между нуклеотидными последовательностями с 10% или менее неправильных спариваний) проводили как описано Holden с сотр. (17). При нестрогих условиях (обеспечивающих гибридизацию между последовательностями с 50% неправильных спариваний) фильтры гидризировали в течение ночи при 42°C в 10% формамид (0,25M

Na₂HPO₄/7% SDS и наиболее строгую стадию осуществляли с помощью 20 мМ Na₂HPO₄/1% SDS при 42⁰С. ДНК-фрагменты, используемые в качестве зондов метили с помощью (³²P)dCTP с использованием "Radprime" системы (Гибко-BRL) или с (дигоксигенин-11/dU TP и детектировали с применением Дигоксигениновой системы (Бюэрингер Маннхейм) в соответствии с инструкциями производителей за исключением того, что гибридизацию проводили в том же растворе, что использовался для радиоактивно-меченных зондов. Геномную ДНК получали для саузерн-гибридизации, как описано ранее (13).

Молекулярное клонирование и нуклеотидное секвенирование.

Рестрикционные эндонуклеазы и T4 ДНК-лигазу получали из Гибко-BRL. Использовали общие методы молекулярной биологии, описанные Sambrook (с сотр. (18). Секвенирование нуклеотидов осуществляли с помощью метода дидеокси цепной-терминации (19) с использованием T7 секвенаторного набора (Фармация). Последовательности собирали с помощью программного обеспечения Mac Vector 3,5 или AssemblyLIGN упаковок. Нуклеотидную и соответствующие аминокислотные последовательности сравнивали с образцами из баз данных Европейской лаборатории молекулярной биологии (EMBI) и Swiss Prot с применением программ BLAST и FASTA GCG упаковки из Висконсинского Университета (версия 8) (20) в сетевой службе при Проектно-исследовательском центре картирования человеческого генома, Хинкстон, Великобритания.

Тесты на вирулентность. Группы из пяти самок BAL B/c мышей (20-25 г) инокулировали орально (р.о.) или внутрибрюшинно (i.p.) 10-кратными разбавлениями бактерий, суспендированных в физиологическом растворе. Для получения инокулята бак-

терии выращивали в течение ночи при 37°C в LB бульоне при
встряхивании (50 об./мин) и затем использовали для инокуляции
свежей среды в течение различных промежутков времени до дос-
тижения оптической плотности (ОД) при 560 нм порядка 0,4-0,6.
Для плотностей клеток порядка 5×10^8 колониеобразующих еди-
ниц (коЕ) на мл и выше, культуры концентрировали центрифуги-
рованием и ресуспендировали в физиологическом растворе. Кон-
центрацию коЕ/мл проверяли посевом серий разбавлений инокуля-
та на LB агаровые чашки. Мышей инокулировали i.p. 0,2 мл объ-
емами и p.o. путем проглатывания того же объема инокулята.
Значения LD₅₀ рассчитывали через 28 дней по методу Reed и Me-
unch (21).

Результаты

Локализация транспозоновых инсерций. Образование банка
Salmonella typhimurium мини Tn5 транспозоновых мутантов и
скрининг, используемый для идентификации 43 мутантов с ослаб-
ленной вирулентностью были описаны ранее (5). Транспозоны и
фланкирующие ДНК области клонировали из эксконъюгантов путем
селекции на устойчивость к канамицину или обратной PCR. Полу-
чали нуклеотидные последовательности ДНК, имеющей протяжен-
ность 300-600 п.о., фланкирующей транспозоны для 33 мутантов.
Сравнение таких последовательностей с соответствующими после-
довательностями из баз данных по ДНК и белкам показало, что
14 мутантов образовалось в результате инсерций транспозона в
ранее известные вирулентные гены, 7 из мутантов образуются в
результате инсерций в новые гены, обладающие подобием извест-
ным генам энтеробактерий, и 12 образуются в результате инсер-

ций в последовательности, необладающие подобием элементам ДНК и белковой баз данных (ссылка 5, пример 1 и настоящий пример).

Три основания позволяют предположить, что 16 из 19 транспозоновых инсерций в новые последовательности кластеризуются в трех областях генома, первоначально обозначенных символами А, В и С. Во-первых, сравнение нуклеотидных последовательностей из участков, фланкирующих сайты инсерции транспозонов, друг с другом и с последовательностями из базы данных показывает, что некоторые последовательности перекрываются одна другой или обладают сильным сходством с различными областями того же гена. Во-вторых, саузерн-анализ геномной ДНК, переваренной с несколькими рестрикционными ферментами и зондированной рестрикционными фрагментами, фланкирующими сайты инсерции транспозона, показывает, что некоторые транспозоновые инсерции находятся на одних и тех же рестрикционных фрагментах. В-третьих, в том случае, когда зонды ДНК гибридизированы в бляшки из библиотеки λ ДНК *S.typhimurium*, зонды из мутантов, которые, как предполагается на двух предыдущих стадиях, способны связываться, на самом деле гибридизируются с теми же клонами λ ДНК. Таким образом, два мутанта (P9B7 и P12F5) приписаны кластеру А, пять мутантов (P2D6, P9B6, P11C3, P11D10 и P11H10) - кластеру В, а девять мутантов (P3F4, P4F8, P7A3, P7B8, P7G2, P8G12, P9G4, P10E11 и P11B9) приписаны кластеру С (фиг.8).

Гибридизация ДНК-зондов из трех этих кластеров с лизатами из ряда *S.typhimurium* штаммов с закрепленными замкнутыми Mud-P22 профагами (15,16) показала, что все три локуса находятся в области 30-31 минут (вариант УШ, ссылка 22) (фиг.7),

что является указанием на тот факт, что три этих локуса тесно связаны или составляют один крупный вирулентный локус. Для установления того, что какой-либо из λ клонов, покрывающих кластеры А, В и С, содержит перекрывающиеся ДНК-вставки, фрагменты ДНК из терминальных участков каждого клона использовали в качестве зондов в Саузерн-гибридизационном анализе других λ клонов. Гибридизация фрагментов ДНК показала, что несколько клонов перекрываются и, что кластеры А, В и С содержат один смежный участок (фиг.8). Затем ДНК фрагменты из концов такого участка использовали для зондирования λ библиотеки с целью идентификации дополнительных клонов, содержащих вставки, представляющих собой соседние участки. Не были идентифицированы λ клоны, которые покрывают крайнее правое окончание локуса, в связи с чем такую область получали клонирование 6,5 т.п.о. EcoRI/XbaI фрагмента лизата Mud-P22 профагового штамма TT15244 (16).

Рестрикционное картирование и Саузерн-гибридизационный анализ далее использовали для конструирования физической карты такого локуса (фиг.8). Для того, чтобы отличить такой локус от хорошо охарактеризованного inv/spa генного кластера на 63 минуте (вариант УШ, ссылка 22) (8, 9, 23, 24, 25, 26), на последний ссылаются, как на вирулентный генный кластер 1 (VGC1), новый вирулентный локус обозначали, как VGC2. На фиг.2 показано положение двух частей ДНК с установленной нуклеотидной последовательностью ("последовательность 1" и "последовательность 2"). Нуклеотидная последовательность представлена на фиг.11 и 12.

Картирование границ VGC2 на хромосоме S.typhimurium

Нуклеотидное секвенирование λ клона 7 с левой стороны VGC2 позволило обнаружить наличие открытой рамки считывания (ORF), установленная аминокислотная последовательность которой на примерно 90% идентична полученному продукту сегмента $ydhE^+$ гена E.coli, а секвенирование клонированного фрагмента E.coRI/XbaI из 6,5 т.п.о. с правой стороны VGC2 обнаружило присутствие ORF, что предсказанная аминокислотная последовательность примерно на 90% идентична пируваткиназе 1 E.coli, кодирующей $pykF$ геном (27). На E.coli хромосоме $ydhE$ и $pykF$ расположены вблизи друг друга в интервале 37-38 минут (28). Одиннадцать неперекрывающихся ДНК фрагментов, распределенных по длине VGC2, использовали в качестве зондов в нестрогом Саузерн-гибридизационном анализе E.coli и S.typhimurium геномной ДНК. Гибридизация фрагментов ДНК показала, что участок в примерно 40 т.п.о., содержащий VGC2, отсутствует в E.coli геноме и локализован на границах VGC2 в области 1 т.п.о. (фиг.9). Сравнение местоположения XbaI сайта вблизи с правым концом VGC2 (фиг.8) с картой известных XbaI сайтов (29) в области 30 минут хромосомы (22) позволяет определить позицию карты 30,7 минуты для VGC2.

Структура VGC2. Нуклеотидное секвенирование частей VGC2 обнаружило присутствие 19 ORF (фиг.8). G+C содержание примерно 26 т.п.о. нуклеотидной последовательности в VGC2 составляет 44,6%, по сравнению с 47% для VGC1 (9) и 51-53% установленному для всего Salmonella генома (30).

Полные, установленные аминокислотные последовательности

ORF 1-II аналогичны последовательностям белков секретионных систем типа III (6, 7), которые, как известно, необходимы для экспорта вирулентных детерминант в большое число бактериальных патогенов растений и животных (7). Предсказанные белки ORF I-8 (фиг.8) похожи по организации и последовательности продуктам *yscN-U* генов *Yersinia pseudotuberculosis* (31), *invC/spaS* кластера *inv/spa* в VGC1 *Salmonella typhimurium* (8, 9) и *spa 47/spa 40* кластера *spa/mxiA* *Shigella flexneri* (32, 33, 34, 35). Так, например, предсказанная аминокислотная последовательность ORF 3 (фиг.8) на 50% идентична *YscS* *Y.pseudotuberculosis* (31), на 34% идентична *spa9* из *S.flexneri* (35) и на 37% идентична *SpaQ* VGC1 *S.typhimurium* (9). Предсказанный белковый продукт ORF9 близкородственен *LcrD* семейству белков с 43% идентичностью к *LcrD* *Y.enterocolitica* (36), 39% идентичностью к *MxiA* *S.flexneri* (32) и 40% идентичностью к *InvA* VGC1 (23). Неполные нуклеотидные последовательности для оставшихся ORF, показанных на фиг.8 указывают на то, что предсказанный белок из ORF10 наиболее близок *Y.enterocolitica* *VscJ* (37) липопротеину, расположенному на бактериальной внешней мембране, с ORF11, аналогичной *S.typhimurium* *InvG*, члену *PulD* семейства транслоказ (38). ORF12 и ORF13 демонстрируют значительное подобие сенсорной и регуляторной субединицам, соответственно, из большого числа белков, содержащих двухкомпонентные регуляторные системы (39). Имеется достаточная кодирующая способность для дополнительных генов между ORF 9 и 10, ORF 10 и 11 и между ORF 19 и правым концом VGC2.

VGC2 законсервирован в и специфичен к *Salmonellae*

2,2 т.п.о. Pst 1/Hind III фрагмент, находящийся в центре VGC2 (зонд В, фиг.8), не обладающий подобием последовательности к элементам баз данных ДНК и белков, использовали в качестве зонда в Саузерн-гибридизационном анализе геномной ДНК из Salmonella сероваров и других патогенных бактерий (фиг.10А).

Фрагменты ДНК, гибридизирующиеся в нестрогих условиях, показывают, что VGC2 присутствует в S.aberdeen, S.gallinarum, S.cubana, S.typhi и отсутствует в ЕРЕС, ЕНЕС, V.pestis, S.flexneri, V.cholera и S.aureus. Таким образом, VGC2 законсервирован в и, по-видимому, специфичен к Salmonellae. Для установления факта консервации организации локуса среди испытанных Salmonella сероваров, проводили Саузерн-гибридизации в строгих условиях геномной ДНК, переваренной с двумя дополнительными рестрикционными ферментами. Гибридизация фрагментов ДНК показала, что имеется некоторая гетерогенность в расположении рестрикционных сайтов между S.typhimurium LT2 и S.gallinarum, S.cubana и S.typhi (фиг.10В). Более того, S.gallinarum и S.typhi содержат дополнительные гибридизирующиеся фрагменты по отношению к тем, что присутствуют в других изученных Salmonellae, что позволяет предположить дупликацию участков VGC2 в таких разновидностях.

VGC2 необходим для вирулентности мышей. Предварительные эксперименты показали, что значения LD_{50} для i.p. инокуляции транспозоновых мутантов P3F4, P7G2, P9B7 и P11C3, по крайней мере, в 100 раз выше, чем для штамма дикого типа (5). Для выяснения важности VGC2 в процессе инфицирования определяли значения LD_{50} для орального и внутрибрюшинного применений для мутантов P3F4 и P9B7. (Таблица 1). Оба мутанта демонстрирова-

ли понижение вирулентности, на, по крайней мере, пять поряд-
ков при любом пути инокуляции в сравнении с родительским штам-
мом. Сильное ослабление вирулентности в результате обоих пу-
тей инокуляции демонстрирует тот факт, что VGC2 необходим для
событий в инфекционном процессе после проникновения эпители-
альной клетки в организм BAL B/c мыши.

ТАБЛИЦА 1

Значения LD_{50} штаммов *S.typhimurium*

$LD_{50}/коЕ$

Штамм	i.p.	: p.o.
12023 дикого типа	4,2	$6,2 \times 10^4$
P3F4	$1,5 \times 10^6$	$>5 \times 10^9$
P9B7 - колониево- образующие единицы	$>1,5 \times 10^6$	$>5 \times 10^9$

Обсуждение

Был идентифицирован ранее неизвестный вирулентный локус
в *S.typhimurium* размером примерно 40 т.п.о., расположенный в
области 30,7 минуты на хромосоме путем картирования инсерци-
онных сайтов группы сигнатурно-меченных транспозоновых мутан-
тов с ослабленной вирулентностью. (5). Этот локус отнесен к
вирулентному генному кластеру 2 (VGC2) для того, чтобы отли-
чить его от inv/sra вирулентных генов на 63 минуте (вариант

УП, ссылка 22), который по предложению авторов был переименован в VGC1. Как VGC1, так и VGC2 кодируют компоненты секреторных систем типа III. Однако, такие секреторные системы отличаются друг от друга.

Из 19 мутантов, образующихся в результате инсерций в новые гены (ссылка 5 и настоящий пример), 16 картировались в одинаковый участок хромосомы. Возможно, что мини-Tn5 инсерция осуществляется преимущественно в VGC2. С другой стороны, поскольку негативный отбор, используемый для идентификации мутантов с ослабленной вирулентностью (5) был очень строгим (что отражается на высоких значениях LD_{50} для VGC2 мутантов), возможно, что среди ранее неизвестных генов лишь мутации в таких VGC2 обеспечивают степень ослабления достаточную для обнаружения при отборе. Неудача предшествующих поисков S.typhimurium вирулентных детерминант для идентификации VGC2 может быть связана со снижением достоверности анализов на клеточной культуре по сравнению с моделью инфицирования на живом животном. В предыдущем исследовании по идентификации участков S.typhimurium LT2 хромосомы уникальной в отношении Salmonellae (40) один из таких участков (RF333) был отнесен к области 30,5-32 минут. Поэтому, RF333 может соответствовать VGC2, хотя неизвестно, что RF333 принимает участие в определении вирулентности.

Сравнения с секреторными системами типа III, кодированными вирулентными плазмидами Yersinia и Shigella, а также с VGC1 Salmonella показывает, что VGC2 кодирует основные структурные компоненты секреторного устройства. Кроме этого, порядок ORF 1-8 в VGC2 одинаков с генными порядками в гомологиях

в Yersinia, Shigella и VGC1 S.typhimurium. Тот факт, что
организация и структура VGC2 секреторной системы не более
родственна с VGC1, чем с соответствующими генами Yersinia, в
совокупности с низким G+C содержанием VGC2 позволяет утверж-
дать, что VGC2 подобно VGC1 (40, 41, 42) независимо приобре-
тается S.typhimurium в результате горизонтальной передачи.
Белки, кодируемые ORF 12 и 13, обладают сильным сходством
с бактериальными двухкомпонентными регуляторами (39) и спо-
собны регулировать как ORF 1-II так и/или секретированные бел-
ки такой системы.

Было показано, что многие гены в VGC1 являются важными
для проникновения S.typhimurium в эпителиальные клетки. Такой
процесс требует бактериального контакта (2) и приводит в ре-
зультате к цитоскелетным перегруппировкам, приводящим к лока-
лизированным активным перемещениям в мембране клетки (43, 44).
Роль VGC1 и его рестрикции на этой стадии инфицирования отра-
жается примерно 50-кратным ослаблением вирулентности у BALB/c
мышей, инокулированных р.о. мутантами VGC1 и тем фактом, что
VGC1 мутанты не демонстрируют потери вирулентности при внут-
рибрюшинном применении (8). Второе наблюдение также объясняет,
почему VGC1 мутанты не получают в нашем скрининге (5). В от-
личие от этого, мутанты в VGC2 сильно ослабляются как после
р.о., так и i.p. инокуляции. Этот факт показывает, что, в от-
личие от VGC1, VGC2 необходим для вирулентности у мышей после
пенетрации эпителиальной клетки, однако такие факты не исклю-
чают роли VGC1 на такой ранней стадии инфицирования.

Таким образом, в резюмирующем картировании инсерционных
сайты 16 сигнатура-меченных транспозоновых мутантов на Salmo-

nella typhimurium хромосоме обеспечивают идентификацию *40 т.по.* вирулентного генного кластера на 30,7 минуте. Такой локус законсервирован среди всех других изученных видов Salmonella, но отсутствует в большом числе других патогенных бактерий или в Escherichia coli K12. Нуклеотидное секвенирование части такого локуса обнаружило 11 открытых рамок считывания, предсказанные белки которых кодируют компоненты секреторной системы типа III. Для проведения различия между ней и секреторной системой типа III, кодированной локусом инвазии *inv/spa*, на *inv/spa* локус ссылаются, как на вирулентный генный кластер 1 (VGC1), а на новый локус, как на VGC2. VGC2 имеет более низкое G+C содержание, чем у Salmonella генома и фланкирован генами, продукты которых имеют более, чем 90% идентичность с продуктами *E.coli* *udhE* и *rykF* генов. Таким образом, вероятно, что VGC2 приобретает горизонтально путем инсерции в участок, соответствующий области между *udhE* и *rykF* генами *E.coli*. Исследования вирулентности VGC2 мутантов показали, что она ослабляется по крайней мере, на пять порядков по сравнению со штаммом дикого типа после оральной или внутрибрюшинной инокуляции.

Литература к этому примеру

1. Carter, P.B. & Collins, F.M. (1974) *J. Exp. Med.* 139, 1189-1203.
2. Takeuchi, A. (1967) *Am. J. Pathol.* 50, 109-136.
3. Finlay, B.B. (1994) *Curr. Top. Microbiol. Immunol.* 192, 163-185.
4. Groisman, E.A. & Ochman, H. (1994) *Trends Microbiol.* 2, 289-293.

5. Hensel. M., Shea. J.E., Gleeson. C., Jones. M.D., Dalton. E. & Holden, D.W. (1995) *Science* **269**, 400-403.
6. Salmond, G.P.C. & Reeves, P.J. (1993) *Trends Biochem. Sci.* **18**, 7-12.
7. Van Gijsegem, F., Genin., S. & Boucher, C. (1993) *Trends Microbiol.* **1**, 175-180.
8. Galan. J.E. & Curtiss. R. (1989) *Proc. Natl. Acad. Sci. U.S.A.* **86**, 6383-6387.
9. Groisman, E.A. & Ochman. H. (1993) *EMBO J.* **12**, 3779-3787.
10. Ginocchio. C.C., Olmsted. S.B., Wells. C.L. & Galan. J.E. (1994) *Cell* **76**. 717-724.
11. de Lorenzo. V. & Timmis. K.N. (1994) *Methods Enzymol.* **264**. 386-405.
12. Davis, R.H., Botstein, D. & Roth, J.R. (1980) *Advanced Bacterial Genetics*, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
13. Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A. and Struhl, K. (1987) *Current Protocols in Molecular Biology* Vol 4 John Wiley and Sons, Inc, New York.
14. Maurer, R., Osmond, B.C., Shekhtman, E., Wong, A. & Botstein, D. (1984) *Genetics* **108**, 1-23.
15. Youderain, P., Sugiono, P., Brewer, K.L., Higgins, N.P. & Elliott, T. (1988) *Genetics* **118**, 581-592.
16. Benson, N.R. & Goldman, B.S. (1992) *J. Bacteriol.* **174**, 1673-1681.
17. Holden, D.W., Kronstad, J.W. & Leong, S. (1989) *EMBO J.* **8**, 1927-1934.
18. Sambrook, J., Fritsch, E.F. & Maniatis, T. (1989) *Molecular cloning: a laboratory manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, New York).

19. Sanger, F., Nicklen, S. & Coulson, A.R. (1977) *Proc. Natl. Acad. Sci. U.S.A.* **74**, 5463-5467.
- 5 20. Devereux, J., Hearberli, P. & Smithies, O. (1984) *Nucl. Acids Res.* **12**, 387-399.
21. Reed, L.J. & Muench, H. (1938) *Am. J. Hyg.* **27**, 493-497.
- 10 22. Sanderson, K.E., Hessel, A. & Rudd, K.E. (1995) *Microbiol. Rev.* **59**, 241-303.
23. Galan, J.E., Ginocchio, C. & Costeas, P. (1992) *J. Bacteriol.* **174**, 4338-4349.
- 15 24. Ginocchio, C., Pace, J. & Galan, J.E. (1992) *Proc. Natl. Acad. Sci. U.S.A.* **89**, 5976-5980.
- 20 25. Eichelberg, K., Ginocchio, C.C. & Galan, J.E. (1994) *J. Bacteriol.* **176**, 4501-4510.
26. Collazo, C.M., Zierler, M.K. & Galan, J.E. (1995) *Mol. Microbiol.* **15**, 25-38.
- 25 27. Ohara, O., Dorit, R.L. & Gilbert, W. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 6883-6887.
- 30 28. Bachman, B. (1990) *Micro. Rev.* **54**, 130-197.
29. Liu, S.L., Hessel, A. & Sanderson, K.E. (1993) *J. Bacteriol.* **175**, 4104-4120.
- 35 30. Fasman, G.D. (1976) *CRC Handbook of Biochemistry and Molecular Biology*, CRC Press, Cleveland.
31. Bergman, T., Erickson, K., Galyov, E., Persson, C. & Wolf Watz, H. (1994) *J. Bacteriol.* **176**, 2619-2626.
- 40 32. Andrews, G.P. & Maurelli, A.T. (1992) *Infect. Immun.* **60**, 3287-3295.
- 45 33. Allaoui, A., Sansonetti, P.J. & Parsot, C. (1993) *Mol. Microbiol.* **7**, 59-68.
- 50 34. Venkatesan, M., Buysse, J.M. & Oaks, E.V. (1992) *J. Bacteriol.* **174**, 1990-2001.

35. Sasakawa, C., Komatsu, K., Tobe, T., Suzuki, T. & Yoshikawa, M. (1993) *J. Bacteriol.* **175**, 2334-2346.
- 5 36. Plano, G.V., Barve, S.S. & Straley, S.C. (1991) *J. Bacteriol.* **173**, 7293-7303.
37. Michiels, T., Vanooteghem, J.C., Lambert de Rouvroit, C., China, B., Gustin, A., Boudry, P. & Cornelis, G.R. (1991) *J. Bacteriol.* **173**, 4994-5009.
- 10 38. Kaniga, K., Bossio, J.C. & Galan, J.E. (1994) *Mol. Microbiol.* **13**, 555-568.
- 15 39. Ronson, C.W., Nixon, B.T. & Ausubel, F.M. (1987) *Cell* **49**, 579-581.
- 20 40. Groisman, E.A., Sturmoski, M.A., Solomon, F.R., Lin, R. & Ochman, H. (1993) *Proc. Natl. Acad. Sci. U.S.A.* **90**, 1033-1037.
- 25 41. Altmeyer, R.M., McNern, J.K., Bossio, J.C., Rosenshine, I., Finlay, B.B. & Galan, J.E. (1993) *Mol. Microbiol.* **7**, 89-98.
42. Li, J., Ochman, H., Groisman, E.A., Boyd, E.F., Solomon, F., Nelson, K. & Selander, R.K. (1995) *Proc. Natl. Acad. Sci. USA* **92**, 7252-7256.
- 30 43. Finlay, B.B. & Rauschkowski, S. (1991) *J. Cell Sci.* **99**, 283-296.
44. Francis, C.L., Starnbach, M.N. & Falkow, S. (1992) *Mol. Microbiol.* **6**, 3077-3087.
- 35
- 40
- 45
- 50

ПРИМЕР 5Идентификация вирулентных генов в *Streptococcus pneumoniae*(а) Мутагенез

В отсутствии удобной транспозоновой системы наиболее эффективный путь создания меченных мутантов *Streptococcus pneumoniae* заключается в использовании инсерционно-дупликатного мутагенеза (Morrison с сотр. (1984) J.Bacteriol. 159, 870). Случайные *S.pneumoniae* ДНК фрагменты, размером в 200-400 п.о. могут быть получены перевариванием геномной ДНК рестрикционным ферментом или физическим распределением под действием ультразвука с последующим гель-фракционированием и репарацией конца ДНК с использованием Т4 ДНК полимеразы. Такие фрагменты лигируют в плазмиду рJDC9 (Pearce с сотр. (1993) Mol.Microbiol 9, 1037, которая несет erm ген для селекции эритромицина в *E.coli* и *S.pneumoniae*), предварительно модифицированную введением меток последовательности ДНК в один из полипинкерных клонирующих сайтов. Размер клонированной *S.pneumoniae* ДНК достаточен для обеспечения гомологической рекомбинации и уменьшает возможность возникновения нерепрезентативной библиотеки *E.coli* (экспрессия *S.pneumoniae* белков может быть токсичной для *E.coli*). Имеются в распоряжении альтернативные векторы, несущие селективируемые маркеры, и они могут использоваться вместо рJDC9. Меченные плазмиды, несущие ДНК фрагменты, вводили в соответствующий штамм *S.pneumoniae*, отобранный на основе серотипа и вирулентности в мышинной модели пневмококковой пневмонии. Регуляция компетентности генетической трансформации в *S.pneumoniae* управлялась фактором компетентности, пеп-

тидом из 17 аминокислот, который недавно был охарактеризован группой Don Morrison в Иллинойском Университете в Чикаго и описан Havarstein, Coomaraswamy и Morrison (1995) Proc.Natl. Acad.Sci. США 92, 11140-11144. Введение малых количеств такого пептида в трансформационные эксперименты приводило к проявлению очень эффективных трансформационных частот в некоторых инкапсулированных клинических изолятах S.pneumoniae. Это позволяет преодолеть главное препятствие в области пневмококковой молекулярной генетики и доступность пептида значительно облегчает конструирование S.pneumoniae мутантных банков и обеспечивает гибкость в выборе штаммов, подлежащих мутации. Определенное количество трансформантов анализировали с целью подтверждения гомологической интеграции плазмидных последовательностей и проверяли на стабильность. Очень низкий уровень реверсии, связанной с мутантами, полученными в результате инсерции-дупликации, минимизируется благодаря тому факту, что дублицированные участки являются короткими (200-400 п.о.); однако, если уровень реверсии неприемлемо высок, то осуществляют антибиотическую селекцию в ходе роста трансформантов в культуре и в ходе роста в животном.

(b) Животная модель

S.pneumoniae мутантный банк организовывали в пулы для инокуляции в Swiss и/или C57Bl/6 мышей. Проводили предварительные эксперименты для определения оптимальной вариабельности пулов и оптимального уровня инокуляции. В одной из моделей использовали инокулят из 10^5 коЕ, который орально доставляли в трахею (Veber с сотр. (1993) J.Antimicrobial Chemother. 32,

473). Мыши разновидности Swiss проявляли симптомы острой пневмонии за 3-4 дня, а мыши C57Bl/6 демонстрировали подострую пневмонию через 8-10 дней. Такие легочные модели инфицирования обеспечивали 10^8 cfu/легкое (Veber с сотр. (1993) J. Anti-microbial Chemotherapy 32, 473) ко времени смерти. Если требовалось, то мышей также инъецировали внутрибрюшинно для идентификации генов, требуемых для инфицирования кровеносной системы (Sullivan с сотр. (1993) Anti microbial Agents and Chemotherapy 37, 234).

(с) Идентификация вирулентного гена

После оптимизации параметров инфекционной модели мутантный банк, состоящий из нескольких тысяч штаммов, подвергали тестам на вирулентность. Мутанты с ослабленной вирулентностью идентифицировали гибридизационным анализом с использованием радиомеченных меток из "начального" и "регенерированного" пулов в качестве зондов. Если S. pneumoniae ДНК неспособна к легкому блоттингу колонии, то хромосомальную ДНК высвобождали химически или ферментативно в лунках титрационных микропланшетов перед переносом на нейлоновые мембраны с использованием устройства для дот-блоттинга. ДНК фланкирующую интегрированную плазмиду клонировали плазмидным "спасением" в E.coli (Morrison с сотр. (1984) J. Bacteriol. 159:870) и секвенировали. Библиотеки геномных ДНК конструировали в соответствующих векторах, находящихся либо в E.coli или в Грам-положительном штамме-хозяине и зондировали рестрикционными фрагментами, фланкирующими интегрированную плазмиду, для выделения клонированных вирулентных генов, которые затем подвергали полному секвенированию и

детальному функциональному анализу.

ПРИМЕР 6

Идентификация вирулентных генов в *Enterococcus faecalis*

(a) Мутагенез

Мутагенез *E. faecalis* осуществляли с использованием плазмиды рАТ112 или производного созданного для этой цели. рАТ112 несет гены для селекции как в Грам-отрицательных, так и в Грам-положительных бактериях и att сайт Tn1545. Поэтому, она требует присутствия в штамме-хозяине интегразы для транспозиции. Стабильные, однокопиевые инсерции получали в том случае, если хозяин не содержал гена эксцизионной гидролазы (Trieu-Cuot с сотр. (1991) Gene 106, 21). Выделение ДНК, фланкирующей интегрированную плазмиду, осуществляли рестрикционным перевариванием геномной ДНК, внутримолекулярным лигированием и трансформацией *E. coli*. Присутствие единичных сайтов для рестрикционных ферментов в рАТ112 и ее производных будет (Trieu-Cuot с сотр. (1991) Gene 106, 21) обеспечивать введение меток последовательности ДНК до переноса на вирулентный штамм *E. faecalis*, несущий плазмиду рАТ145 (для обеспечения интегразной функции) в результате конъюгации, электропорации или трансформации (Trieu-Cuot с сотр. (1991) Gene 106, 21; Wirth с сотр. (1986) J. Bacteriol. 165, 831).

(b) Животная модель

Большое число инсерционных мутантов анализировали на рандомизированную интеграцию плазмиды путем выделения ДНК из

транспициентов, рестрикционно-энзимного перевара и продуктов Саузерн-гибридизации. Индивидуальные мутанты хранили в лунках микротитрационных планшетов, вариабельность и размер инокулянтных пулов оптимизировали перед скринингом мутантного банка. Использовали две различные модели инфицирования, вызываемого E. faecalis. Первая из которых представляет собой хорошо установленную модель эндокардита на крысах, предусматривающую инъекцию до 10^8 коЕ E. faecalis в хвостовую вену животных, имеющих катетер, вставленный поперек клапана аорты (Whitman с сотр. (1993) Антимикробные агенты и химиотерапия 37, 1069). Животных умерщвляли через различные промежутки времени после инокуляции и бактериальные наросты на клапане аорты вырезали, гомогенизировали и высевали на культурную среду для выделения бактериальных колоний. Вирулентные бактерии также выделяли из крови через различные промежутки времени после инокуляции. Вторая модель представляла собой перитонит мышей, возникающий после внутрибрюшинной инъекции до 10^9 коЕ E. faecalis/Chenoweth с сотр. (1990) Антимикробные агенты и химиотерапия 34, 1800). Как и в случае модели S. pneumoniae проводили предварительные эксперименты для установления оптимальной вариабельности пулов и оптимального уровня инокуляции перед скринингом мутантного банка.

(с) Идентификация вирулентного гена

Выделение ДНК, фланкирующей сайт интеграции рAT112, используя ее E.coli начало репликации, упрощается в связи с отсутствием сайтов большинства рестриктаз, обычно используемых для распознавания 6 п.о. в векторе. В связи с этим, ДНК из

интересующих штаммов переваривали с одним из таких ферментов, самолигировали, трансформировали в *E.coli* и секвенировали с использованием праймеров, базирующихся на последовательностях соседних с att сайтами плазмиды. Геномную ДНК библиотеку *E. faecalis* зондировали интересующими последовательностями для идентификации интактных копий вирулентных генов, которые затем секвенировали.

ПРИМЕР 7

Идентификация вирулентных генов в *Pseudomonas aeruginosa*

(a) Мутагенез

Поскольку транспозон Tn5 использовался другими авторами для мутагенеза *Pseudomonas aeruginosa*, а мини-Tn5 производное, которое используется для идентификации *Salmonella typhimurium* вирулентных генов (пример 1), как сообщается, находит широкое применение среди Грам-отрицательных бактерий, включая некоторые псевдомонады (De Lorenzo и Timaris (1994) Methods Enzym, 264, 386), *P.aeruginosa* мутантный банк конструировали с использованием существующего пула сигнатура-меченных мини-Tn5 транспозонов путем конъюгального переноса вектора "самоубийцы" на один или более вирулентных (и возможно микобактериальных) реципиентных штаммов. Такой подход обеспечивает значительную экономию времени. Другие производные Tn5, сконструированные специально для *P.aeruginosa* мутагенеза (Rella с сотр. (1985) Gene 33, 293), могут альтернативно использоваться с мини Tn5 транспозонами.

(6) Животная модель и идентификация вирулентного гена

Банк P.aeruginosa инсерционных мутантов подвергали скринингу на ослабленную вирулентность в модели хронической легочной инфекции у крыс. Суспензии P.aeruginosa вводили в бронхи после трахеотомии и заболевание развивалось в течение 30 дневного периода (Woods с сотр. (1982) Infect.Immun. 36, 1223). Бактерии выделяли путем посева легочных гомогенатов на лабораторную среду и их меченные последовательности использовали для зондирования бактериальных пятен колоний ДНК, применяемых в качестве инокулянта. Можно также подвергать мутантный банк тестам на вирулентность в модели эндогенной бактериемии (Hirakata с сотр. (1992) Antimicrobial Agents and Chemother. 36, 1198), и циститного фиброза (Davidson с сотр. (1995) Nature Genetics 9, 351) у мышей. Клонирование и секвенирование ДНК, фланкирующую транспозоны, проводили, как описано в примере 1. Библиотеки геномных ДНК для выделения и секвенирования интактных копий генов конструировали в лаборатории стандартными методами.

ПРИМЕР 8Идентификация вирулентных генов в Aspergillus fumigatus(a) Мутагенез

Функциональный эквивалент транспозоновому мутагенезу в грибах представляет собой регулируемую рестриктазами интеграцию (REMI) трансформирующей ДНК (Schiestl и Petes (1991) Proc.Natl.Acad.Sci. 88, 7585). В таком процессе клетки грибов трансформировали фрагментами ДНК, несущими селектируемый мар-

кер, в присутствии рестрикционного фермента и однокопиевые интеграции осуществляли на различных геномных сайтах, ограни-
ченных последовательностью-мишенью рестрикционного фермента.
REMI уже с успехом использовали для выделения вирулентных ге-
нов Cochliobolus/Lu с сотр. (1994) Proc.Natl.Acad.Sci США 91,
12649) и Ustilago/Bolker с сотр. (1995) Mol.Gen.Genet. 248,
547) и было показано, что введение активного рестрикционного
фермента с плазмидой, кодирующей гигромициновую устойчивость,
приводит к единичной и, по-видимому, неупорядоченной интегра-
ции линейной плазмиды в A.fumigatus геном. Меченные последо-
вательности вводили в удобный сайт в одном или двух векто-
рах гигромициновой устойчивости и использовали для трансфор-
мации клинического изолята A. fumigatus.

(6) Животная модель и идентификация вирулентного гена

Низкодозовая модель аспергиллеза у нейрOPENиЕВЫХ мышей
особенно близко совместима с ходом легочного заболевания лю-
дей (Smith с сотр. (1994) Infect.Immun. 62, 5247). Мышей ино-
кулировали интраназально, введением до 1,000,000 конидиоспор/
мышь. Вирулентные мутанты грибов выделяли спустя 7-10 дней с
использованием легочных гомогенатов для инокуляции жидкой
среды. Через несколько часов собирали гифы, из которых экст-
рагировали ДНК для амплификации и мечения меток, для зондиро-
вания колониевых пятен ДНК из пула трансформантов, включающих
инокулят. ДНК из участков, фланкирующих точки REMI инсерции,
клонировали перевариванием трансформантной ДНК с рестрикцион-
ным ферментом, осуществляющим расщепление вектора вне REMI,
самолигированием и трансформацией E.coli. Праймеры на основе
известной последовательности плазмиды использовали для опре-

деления соседних A.fumigatus ДНК последовательностей. Для доказательства того, что инсерция вектора является причиной возникновения авирулентного фенотипа, выделенную плазмиду повторно расщепляли тем же рестрикционным ферментом, что использовался для клонирования, и подвергали обратной трансформации в родственный штамм A.fumigatus дикого типа. Затем трансформанты, возникшие в результате гомологической рекомбинации подвергали тестам на вирулентность.

Литература (помимо приведенной для примера 4)

Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A. and Struhl, K. (1987) *Current Protocols in Molecular Biology*, New York: John Wiley and Sons.

Buchmeier, N.A., Lipps, C.J., So, M.Y. and Heffron, F. (1993) *Recombination-deficient mutants of Salmonella typhimurium are avirulent and sensitive to the oxidative burst of macrophages. Mol. Microbiol.* 7, 933-936.

Carter, P.B. and Collins, F.M. (1974) The route of enteric infection in normal mice. *J. Exp. Med.* 139, 1189-1203.

de Lorenzo, V. and Timmis, K.N. (1994) Analysis and construction of stable phenotypes in Gram-negative bacteria with Tn5 and Tn10-derived minitransposons. *Methods Enzymol.* 264, 386-405.

de Lorenzo, V., Herrero, M., Jakubzik, U. and Timmis, K.N. (1990) Mini-Tn5 transposon derivatives for insertion mutagenesis, promoter probing, and chromosomal insertion of cloned DNA in gram-negative eubacteria. *J. Bacteriol.* 172, 6568-6572.

Fields, P.I., Groisman, E.A. and Heffron, F. (1989) A *Salmonella* locus locus

that controls resistance to microbicidal proteins from phagocytic cells.

контролирующий устойчивость к микробицидным белкам из фагоцитных клеток.
Science 243, 1059-1062.

5

Finlay, B.B., Starnbach, M.N., Francis, C.L., Stocker, B.A., Chatfield,

S., Dougan, G. and Falkow, S. (1988) Identification and characterization

Идентификация и характеристики мутантов неспособные
of *TnpH* mutants of *Salmonella* that are unable to pass through a

10

polarized MDCK epithelial cell monolayer. *Mol. Microbiol.* 2, 757-766.

проходить через поляризованный МДСК эпителиальный клеточный монослой

15

Groisman, E.A., Chiao, E., Lipps, C.J., Heffron, F. (1989) *Salmonella*

typhimurium *phoP* virulence gene is a transcriptional regulator. *Proc.*

Natl. Acad. Sci. USA. 86, 7077-7081.

20

Groisman, E.A. and Ochman, H. (1994) How to become a pathogen.

Trends Microbiol. 2, 289-293.

25

Groisman, E.A. and Saier, M.H., Jr. (1990) *Salmonella* virulence: new

clues to intramacrophage survival. *Trends Biochem. Sci.* 15, 30-33.

30

Herrero, M., de Lorenzo, V. and Timmis, K.N. (1990) Transposon Транспозоновые

vectors containing non-antibiotic resistance selection markers for cloning

векторы содержащие не-антибиотически устойчивые селекционные маркеры

and stable chromosomal insertion of foreign genes in Gram-negative
для клонирования и стабильная хромосомальная инсерция посторонних
bacteria. *J. Bacteriol.* 172, 6557-6567.
генов в Грам-отрицательные бактерии.

35

Holden D.W., Kronstad J.W., Leong S.A. (1989) Mutation in a heat- мутация в

regulated hsp70 gene of *Ustilago maydis*. *EMBO J.* 8, 1927-1934.

терморегулируемом 70 гене

40

Holland J., Towner K.J., Williams P. (1992) Tn916 insertion инсерционный

mutagenesis in *Escherichia coli* and *Haemophilus influenzae* type b

мутационез в и тип b
following conjugative transfer. *J. Gen. Microbiol.* 138, 509-515.

после конъюгативного переноса.

45

50

Mahan, M.J., Slauch, J.M., Mekalanos, J.J. (1993) Selection of bacterial Селекция

virulence genes that are specifically induced in host tissues. *Science* 259,
бактериальных вирулентных генов специфически индуцируемых в тканях
686-688.
хозяина.

Miller, S.I., Kukral, A.M. and Mekalanos, J.J. (1989a) A two-
Двухкомпонентные регуляторные системы/...../ для контроля
component regulatory system (*phoP phoQ*) controls *Salmonella*
вирулентности

typhimurium virulence. *Proc. Natl. Acad. Sci. USA.* 86, 5054-5058.

Miller, I., Maskell, D., Hormaeche, C., Johnson, K., Pickard, D. and
Выделение орально ослабленной
Dogan, G. (1989b) Isolation of orally attenuated *Salmonella*
после мутагенеза

typhimurium following *TnphoA* mutagenesis. *Infect. Immun.* 57, 2758-
2763.

Miller, V.L. and Mekalanos, J.J. (1988) A novel suicide vector and its Новый вектор
"самоубийца" и его использование в конструировании инверсионных мута-
use in construction of inversion mutations: osmoregulation of outer
ций: осморегуляция внешних мембранных белков и вирулентных детер-
membrane proteins and virulence determinants in *Vibrio cholerae*
инзент в требует
requires *toxR*. *J. Bacteriol.* 170, 2575-2583.

Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) *Molecular cloning:*
молекулярное клонирование:
лабораторное руководство
a laboratory manual, Cold Spring Harbor, New York: Cold Spring
Harbor Laboratory.

Sanger, F., Nicklen, S. and Coulson, A.R. (1977) DNA sequencing with
Секвенирование ДНК с помощью
ингибиторов обрыва цепей
chain terminating inhibitors. *Proc. Natl. Acad. Sci. USA.* 74, 5463-5467.

Schiestl R.H., and Petes T.D. (1991) Integration of DNA fragments by Интеграция
illegitimate recombination in *Saccharomyces cerevisiae*. *Proc. Natl. Acad.*
фрагментов ДНК логически неправильной рекомбинацией в
Sci USA. 88, 7585-7589.

Формула изобретения

1. VGC2 ДНК *Salmonella typhimurium*, где VGC2 ДНК определяется как
ДНК *Salmonella typhimurium*, расположенная между генами *udhE* и *rukF* и имеющая
последовательность, содержащуюся в SEQ ID NO: 37 и 38, или ее часть, состоящая, по
меньшей мере, из 50 нуклеотидов, или вариант указанной ДНК с, по крайней
мере, 85%-ной идентичностью, представляющий собой VGC2 ДНК любого из
микроорганизмов *Salmonella aberdeen*, *Salmonella gallinarum*, *Salmonella cubana*
и *Salmonella typhi*, причем инактивация указанных ДНК, ее части или варианта
сообщает микроорганизму пониженную способность к адаптации к конкретным
условиям окружающей среды.

2. ДНК по п.1, в которой указанная часть ее представляет собой промотор.

3. Мутантная бактерия, обладающая пониженной способностью к адаптации к конкретным условиям окружающей среды, полученная из бактерии, содержащей ген, выбранный из группы, включающей:

а) ген, характеризующийся последовательностью, которая, по крайней мере, на 85% идентична последовательности гена, содержащегося в последовательности, представленной в SEQ ID NO: 37 или 38;

б) ген в VGC2 ДНК *Salmonella typhimurium* или аналогичный ему ген, где VGC2 ДНК *Salmonella typhimurium* представляет собой ДНК *Salmonella typhimurium*, расположенную между генами *udhE* и *rykF*, путем удаления или инактивации указанного гена.

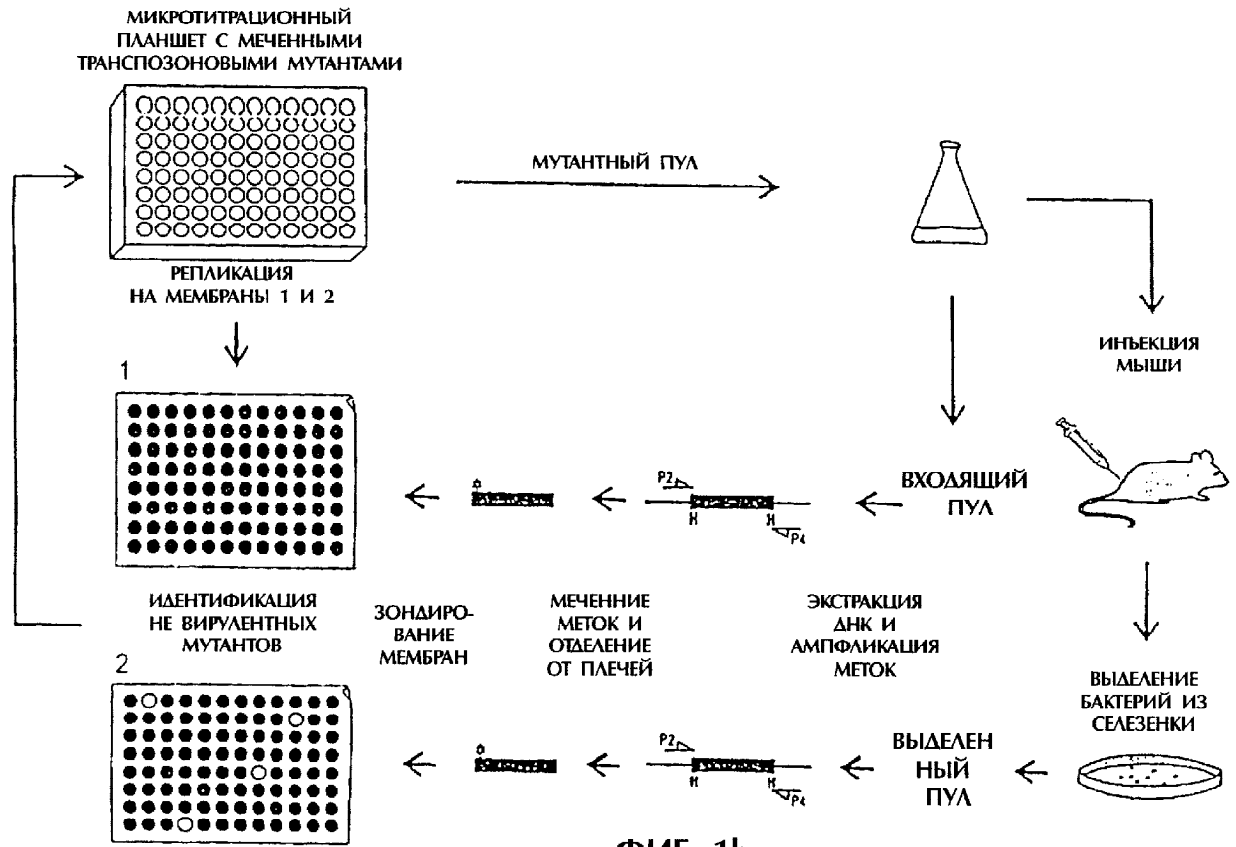
4. Мутантная бактерия по п.3, в которой ген, аналогичный или, по крайней мере, имеющий на 85% идентичную последовательность гену в VGC2, мутирован посредством делеции части кодирующей области.

5. Мутантная бактерия по п.3 или 4, которая представляет собой *Salmonella*, предпочтительно любую из *Salmonella typhimurium*, *Salmonella aberdeen*, *Salmonella gallinarum*, *Salmonella cubana* и *Salmonella typhi*.

6. Способ получения мутантной бактерии, охарактеризованной в п.3, включающий инактивацию гена в бактерии, выбранного из группы, включающей:

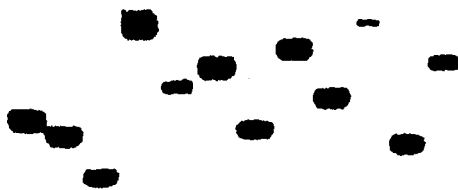
а) ген, характеризующийся последовательностью, которая, по крайней мере, на 85% идентична последовательности гена, представленной в SEQ ID NO: 37 или 38;

б) ген в VGC2 ДНК *Salmonella typhimurium* или аналогичный ему ген, где VGC2 ДНК *Salmonella typhimurium* представляет собой ДНК *Salmonella typhimurium*, расположенную между генами *udhE* и *rykF*, путем инсерции в кодирующую область гена сегмента ДНК или путем делеции всей или части кодирующей области гена.

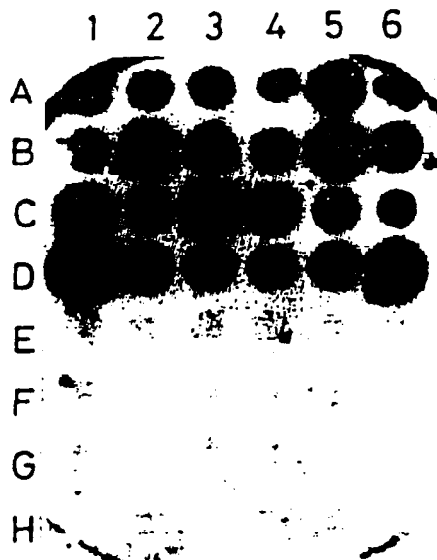


ФИГ. 1b

1 2 3 4 5 6 7 8 9 10 11 12

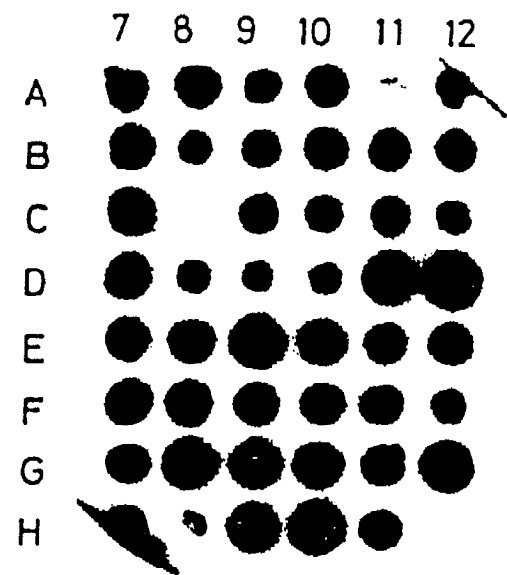
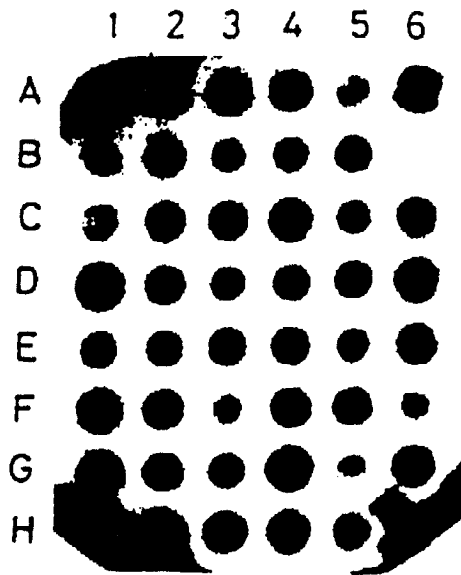


ФИГ. 2

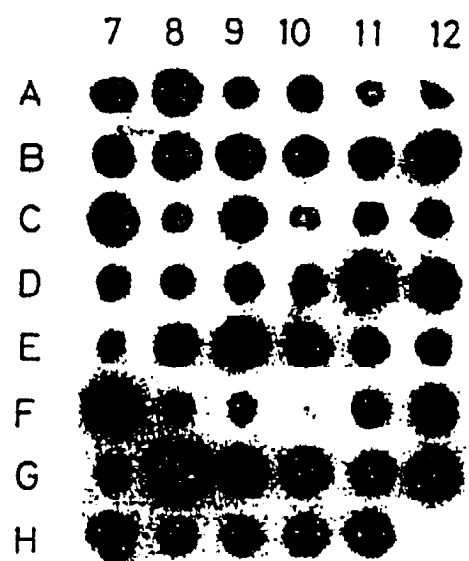
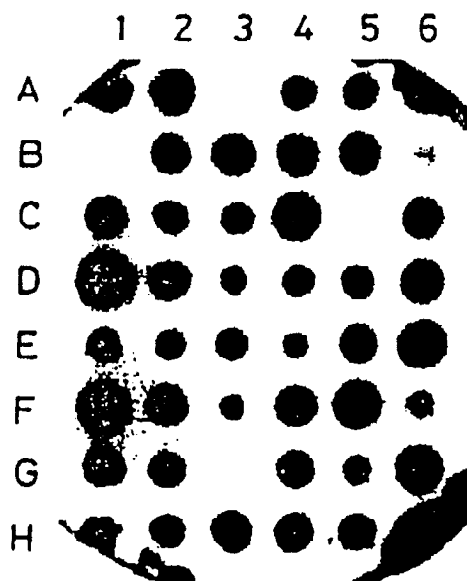


ФИГ. 3

ТИП ИНОКУЛЯНТА



ОБРАЗЕЦ СЕЛЕЗЕНКИ



ФИГ. 4

А) НОВЫЕ ВИРУЛЕНТНЫЕ ФАКТОРЫ С ПОДОБИЕМ К СЕКВЕНИРОВАННЫМ ГЕНАМ:

1. p1F10

ПОДОБИЕ *clpP* (*E.coli*)

(ФИГ. 5 ЗАЯВКИ)

2. p2D6

ПОДОБИЕ *lcrD* (*Yersinia spp.*)

ПОСЛЕДОВАТЕЛЬНОСТЬ p2D6_1_I

SEQ ID №8

GGTCTTAATGTACGGGCATGGTCTGCATCGATAACTCCGGCACGCAAATCGCCATCGATACTCATTTGT
TTGGCTGGCATCCCATCAAGCGAGAAACGTGCGCTAACTTCCGCCACCCCTCTCGATACCTTTTGTAATG
ACAAATAAATTGCACGATAGTAATGATGGTAAATACGACCAACCCAAACGGTGAGATTTCTCCTACGACA
AACTTACCGAAAGCATCCACAAATATTACCGGCATTATGTTGTAACAGTACCCAGCCGTGATGTGCTGA
TTGGGGAGTTAACAACCGATTTAT

3. s4C3

ВЕРОЯТНО ТОТ ЖЕ ГЕН, ЧТО p2D6, НО С ДРУГИМ УЧАСТКОМ

ПОДОБИЕ *S. typhimurium invA* and *Yersinia spp.lcrD*

ПОСЛЕДОВАТЕЛЬНОСТЬ s4C3_1_U

SEQ ID №9

GCGCGGACGCTAGTGTGGTGGGTGACAGCCAGACGTTACCGAACGGGATGGGGCAGATCTGTTGGCTTA
CAAAAGACATGGCCCATAGGGCGCAAGGTTTTGGGACTGGACGTTTTTCGCGGGCAGACAACGTATCTCT
GTCTTATTAATAATGTGTCTGCTTCGGCATATGTATCGAACCCCTCGGAGCAAAGTCGTTTGGGCGCAGA
ATTAGTACGTTTGGGTCGGTTCGCTGTATTTCCTTGGGCTCGGAAAAAGAGTGCCAGCGTGAAGGAGTGG
GATTTGGCAGACTGGCCGCCTAAT

ПОСЛЕДОВАТЕЛЬНОСТЬ s4C3_1_R

SEQ ID №10

CACTATAGGGAAAGCTTGCATGCCTGCAGGTCGACTCTAGAGGATCTACTAGTCATATGGATTGCACTT
GTGTATAAGAGTCAGGATTAGAGGACATGCGCCGGGAACCATACTATCTTTTTCCGGTGCTTCGACGCC
ATTTGCGGAAACCAAGACTTTTTGCGGCGAATGAGGATAATTGGCAATGCTAACCAACGCTGAAAAGAA
AGCGAGAGTGATAAAAAGGAAAGCCAGGAATTAAGCGAGGAGCATTAAACCAACAGCGGCTAATATGAG
CGACTGAGGTTGTCTGGCAATTTG

4. p3F4

ПОДОБИЕ *invG* (*S.typhimurium*)

ПОСЛЕДОВАТЕЛЬНОСТЬ p3F4_1_U

SEQ ID №11

TGCAGGCCGACTCTAGAGGATCCCCGGGTACCGGTAATTTCTTTAACCTCGCATCCCCGGTGGATGAAAG
GATATTCTGGCTGCGTAAGTAATGAATGAACCGCCAGTAGATAAAATATTGAAAGTGATAACCTGATG
TTTTAATAACGATGCAGGATATACATATAACATGCTGGCATCAAACCAGGTAAGCAAATCATATTGTGC
TGCCAGGTTATTCAAACATATCGACCGGTGGTCCAGGCGGGAATTTTCCACTAAATGTAGGTGGGATCA
ATGGGCTAATTGGTATAGGCGGAT

ФИГ. 6

5. p7G2

ПОДОБИЕ *yscC* (*Yersinia spp.*)

ПОСЛЕДОВАТЕЛЬНОСТЬ p7G2_1_U

SEQ ID №12

CCTGTGATTCCGGATGAAATAGCTTTTACGAAAGCTGTCAGACNTGCTGAAGAATACGCTGCAAATGGT
 AAGCTTGTAACCTTTTGGGTATTGTTCCAACGCATGCTGAAACGGGTATGGATATATTTCGTGCGGGTGA
 GTTGATAGGAAATGACGCTTATGCAGTGGCTGAATTTGTGGAGAAACCGGATATCGATACCGCCCGTGA
 CTAATTTCAAATCAGGGGAAATATTACTGGCCTAGCGGCGATGTTTTATTTCGCGCAAAGCCCTTATTT
 AAACGAATTAACGTATCTATCACCCCCAAATTCATACAGCTTGTGAA

ПОСЛЕДОВАТЕЛЬНОСТЬ p7G2_3_O

SEQ ID №13

TTACTAAACAGGGCCCCGGACCcTGTAACACCACGCTTGCCAAcACTAAAAACGATGCTtGCcGTAA
 AAAAAATTGAAcGTTATTTACTTAATAcGCCTATTTTATTTACATTATGCACGGACAGAGGGTGAGGATT
 AAATGGATAATATTGATAATAAGTATAcTCCACAGCTATGTAAAATTTTgGGGGcTATATCgGATtTGg
 TTGtTTtTAATTTAGCCtTATGGcTTtCACTAGGATGTGTCTATTTTTTTtGTGGtCAAGCACAGAGAT
 TTATTTCCCCaACCACC

ПОСЛЕДОВАТЕЛЬНОСТЬ p7G2_1_I

SEQ ID №14

TTTCCTTGCCGTGACAGTCCGGGATGCGAGGTAAACGAAATTACCGGCACCAAAGCTGTGGAGGTGAGC
 GGTGTCCCCAGCTGCCTGACTCGTATTAGTCAATTAGCTTCAGTGTCTGGATAATGCGTTAATCAAACGA
 AAAGACAGTGGCGTGAGTGTAAGTATATACACGCTTAAGTATGCCACTGCGATGGATACCCAGTACCAT
 TATCGCGATCACTCCGTGCTGGTTCCAGGGGTGCGCTAGTGTATTGCGTGAGATGAGTAACACCAGCGT
 CCCGACGTCAATCGACGAACAATGG

6. p9B7

ПОДОБИЕ *fliQ*, *invX* (*E.coli*)

ПОСЛЕДОВАТЕЛЬНОСТЬ p9B7_1_I

SEQ ID №15

CATGAGTAACCTACCCAACCTGTAATCTTTACCAATATGCATCATAATCTTCTGCTGGTAAATGATTGGT
 AATATCGGAAAGGTAAGTGACATAAGCACGCCATTACGTAAGTGCAGGCCCTAACTGCCACTTTTT
 AATAAGGGAAGTAATAAAGAAAGGCTCAATGGTGAATAAAGCCACAGCCAATGCAATAAGCCACTCA
 TTACCTGTTGTGCCATTCAACCATGCTCTCCAATTCGTAACATTATCTGCCGGGTATAATTCAACAGG
 ATACCGCTAAGCCATGGGTAG

ПОСЛЕДОВАТЕЛЬНОСТЬ p9B7_3_O

SEQ ID №16

ATTCCAGCCCCCGGGCCATCTAACCACTATGAACAATCATCTTCTGGGTGGACAATCATTGGTACCATC
 GGCCAGGCTTGTGCAATATGTATGTCATCACGTAAAAGCGCGGCCCTTAATCTCCCATCTCTCCTTA
 AGGGCAGTTATCACGGCTGGCTCAATGGCCGGCTTAACAGCCACAG

7. s6F5

ПОДОБИЕ *yscU* (*Y. enterocolitica*)

ПОСЛЕДОВАТЕЛЬНОСТЬ s6F5_1_O

SEQ ID №17

GAGGCGCGTCTTCGGTTGAGGGTCGCCCTCCAGATCTTTATGCTCCTGTTTTACGTCATCTTTACTCAT
 TTAAAGATCTTTTCTAATCTTATAATATTGAAAAGAATAGTCCAGTATGCCAACGACGAAATAAAGAAA
 CATCACCCCAACCCATAACCATTTTTTCAATGATGAAAGCAACAAGCACGCCACAGGCTAeACCACAGCC
 CGGAGGGGGCCGAAAGTGTGGGATCTTGATTAATGAAAAAGGCAAGGGAAGAGATAGGATGATGCA
 TGCTGGTTGGAGGCAGATTATTATCTTCCG

ФИГ. 6 (продолжение)

В) НОВЫЕ ПОСЛЕДОВАТЕЛЬНОСТИ БЕЗ ПОДОБИЯ ВХОДЯЩИМ В ДНК И ПРОТЕИНОВЫЕ БАЗЫ ДАННЫХ**1. s4D10**

ПОСЛЕДОВАТЕЛЬНОСТЬ s4D10_1_U

SEQ ID №18

AGTTGCCGTATTTATTAAATATTCACCTCAGGTCAATATGGAGGTCTTCCCGGCTAAAAATCATTGCTT
 TACTAGAGATATCACTCCCTGGGTTGCAATACAGTACGATTAGTTATCTTGATGCAGCCTGCTGATTTC
 AGAATGGCAGCTGACGTACCCGCGAGACAAACATTCCTGGATTATGGACGTTATCAACGCCAATATAGGG
 AAGGTGGTGAAGTGGTTGATGAAATACCCCTATCCCTTGATGTTATCGCTGACAGGACTGTTATCAGG
 AGCGGGCATCCTCGATCGGCT

ПОСЛЕДОВАТЕЛЬНОСТЬ s4D10_1_R

SEQ ID №19

CAAGAGACAGATCCAACTCGGGCCGATCGCCATAACGCCAGCAGTTTGAAAGATGAAAGCCCAGCTTAT
 CCAGCCATTCCGGTACAGCGTAACGAGCAGGTTGCCAGAAATAACGATAAAGTTGCAACACCTCGGGAT
 CAGGTCCGGCTCAAAAACGGGGTCTCAGGCAAAAATAGCCGATCAGGATGCCCACTCCTAATAACAGTCC
 TGTCAACGATAACATCAACGGATAAGGGTATTTTCATCAACCACTTCACCACCTTCCCTTTATTGGCGTT
 GGATAACGTCCATAATCCAGA

2. s4H10

ПОСЛЕДОВАТЕЛЬНОСТЬ s4H10_1_U

SEQ ID №20

AGGGCTTTATTGATTCCATTTTTTACTGATGAATGTTCCGTTGCGCTGCCCGGATTACAGCCGGATCC
 TCTAGAGTCGACCTGCAGAACCGAGCCAGGAGCAAATTAATTTTTTTGGGCAATTGCTGAAAGATGAAG
 CATCCACCAAGTAACGCCAGTGCTTTATTACCGCAGGTTATGTTGACCAGACAAATAGATTATATGCAGT
 TAACGGTAGGGCTCGATTATCTTGTTCAGAAATATCAGGCGCAGCATCGCAAGCGCTTAATAAGCTGGGTA
 ACATGGCATGAAGGGGCAACCC

ПОСЛЕДОВАТЕЛЬНОСТЬ s4H10_1_R

SEQ ID №21

CACTATAGGGAAAGCTTGATGCCTGCAGGTGCACTCTAGAGGATCTACTAGTCATATGGATTCTTAGG
 CGGCCAGATCTGATCAAGAGACAGATCCAACTCGGGCCGATCGCCATAACGCCAGCAGTTTGAAAGATG
 AAGCCCAGCTTATCCAGCCATTCCGGTACAGCGTAACGAGCAGGTTGCCAGAAATAACGATAAAGTTG
 CAACACCTCGGGATCAGGTCCGGCTCAAAAACGGGGTCTCAGGCAAAAATAGCCGATCAGGATGCCCACT
 CCTAATAACAGTCCTGTCAACG

3. p4G5

ПОСЛЕДОВАТЕЛЬНОСТЬ p4G5_1_O

SEQ ID №22

CCCCCCCCCTTCTCCTGGCTTACACAGCCCCAGACCGGCGCTGGAAAAGGCCATTCCCGCCATACAGGA
 GGCCAGCAACATATTTTCACGCGCCGCCAGATCGTGGCCGTAAACCCACGGCTTTCCGGCAGCGATTGCGC
 AATCATCGCTATCGCGCCAATCGCCAGGCTGTGCGTAACGGCGTGGCGTTGAGCGCGCTGTAGGCCTC
 AATCGCATGCGTCAACGCATCGATACCGGTCTCGCCGTACGTTTGCGGGAACGCCTTCGGTCACGGA
 AGCATCAAGAATCGCCACGTCCGCGC

ПОСЛЕДОВАТЕЛЬНОСТЬ p4G5_1_U

SEQ ID №23

CGCGAACGTGCGCCGCAACTGCTTGTGGACGGTGAATTGCAGTTTGACGCGGCTTTTCGTGCCGGAGGTC
 GCGCGCAAAAAGCGCCTGACAGCCCGCTGCAAGGCCGCGCCAACGTGATGATTTTCCCGTTCGCTGGAG
 GCGGGCAATATTGGCTACAAAATCACTCAGCGTCTGGGAGGCTATCGCGCTGTTGGGCGGCTAATTTCAG
 GGGCTTGGCGCGCGCTTCACGACCTCTCCCGAGGCTGTAGCGTGCAGGAAATTATCGAACTGCGGTTG
 GTGAGAAAAACCA

ФИГ. 6 (продолжение)

4. p7A3

ПОСЛЕДОВАТЕЛЬНОСТЬ p7A3_1_U

SEQ ID №24

CGCCCTAGCATGCCTGGCGTTGTCCGGTTATTGCTCGTCAAGCGAACAGATGCAAAAGGTGAGAGCGAC
TCTCGAATCATGGGGGTCATGTATCGGGATGGTGTAACTGTGATGACTTATTGGTACGAGAAGTGCA
GGATGTTTTGGATAAAAATGGGTTACCCGCATGCTGAAGTATCCAGCGAAGGGCCGGGGAGCGTGTTAA
TTCATGATGATATACAAATGGATCAGCAATGGCGCAAGGTTCAACCATTACTTGCAGATATTCCCGGGT
TATTGCACTGGCAGATTAGTCACTCTC

ПОСЛЕДОВАТЕЛЬНОСТЬ p7A3_1_I

SEQ ID №25

CCCTTCCCAGGCTCGACAGGTACACAGCCAGCCACTGGTGCAGGCAGTTACTTGCTTTCATCATGGGAA
GGAGCAATATCCTGATATATTAAAGAAAGAGCGGGATCCCCCTTTCTTTACTGCTGCTAACGTTTCTTGC
AAAATGCGTTGATGAGATTCATCCAGCACACCACTGATAACAAAAGAGCGCCGCATTGGCGTAACATTG
ACAAGCCCCACTAAACCGCTCTCTATTATCGCAGAAATAATATCATCCCCCTGAGACTGATGAGAGTGA
CTATTCTGCCAGCGCAAATAACCC

5. p10E11

ПОСЛЕДОВАТЕЛЬНОСТЬ p10E11_1

SEQ ID №26

ATACCGAGTATTAAGCGGCTGTGTAACATCGTCATCCAACAACATACGCAGCGAGCCGCCACGCCGGAA
AAACCGCATCGTGTATGTGCTGTTGTAGGGTCGGGTCTTTTTTTCATGAGTACGTTTCTCGCGCTATC
ATACTGGAAATTTCCCCCACTTACTGATAAGCCCTGTCAAGTTGGGTAAGGACAGAGTTAAGCTCCTGA
GACATTTTTTGGAAATGGTTATCTTTCCCGACTCATAAAATCGGTATTCCCGCTGGGGGCAATATCCAA
AGACGCTTTGGTCGCCCGTAGGGCACC

ПОСЛЕДОВАТЕЛЬНОСТЬ p10E11_U

SEQ ID №27

GCCGTATGCCTGCAGTTGCCCGGTTATTGCTCGTCAAGCGAACCGATGCCAAAGGTGAGAGCGACTCTC
GAATCATGGGGGTCATGTATCGGGATGGTGTAACTGTGATGACTTATTGGTACGAGAAGTGCAAGAT
GTTTTGGTAAAAATGGGTTACCCCATGCTGAAGTATCCAGCGAAGGGCGGGAGCGGTGTTAATTCAC
GATGATATTCAAATGGGTCAGCAATGGGCAAGGTTCAACCCCACTTGCAGATATTCCCCCCCCTATT
GGACTGGCAGATTAGTCACTCTCA

6. s4B9

ПОСЛЕДОВАТЕЛЬНОСТЬ s4B9_1_O

SEQ ID №28

GGGCGACCTGCCCGCGCGCAACTTTCCCGAAGCGTTTTCCATTTCTTGTCTTAAATGACCTGGAA
AGCTTACCTAAGCCTTGTCTTGCCTATGTGACAATACTGCTTGGAGAACACCCGGACGTCCATGATTAT
GCTATACAGATCACAGCGGATGGGGATGGTGAATCGGTTATTATACCACAAGTCGCAGCTCTGAGCTT
ATTGCTATTGAGATAGAAAAACACCCGCTTCAACTTGGATTTTGAATAATGTAATACGCAATCACCAT
ACACTATATTCCGGTGGCGTATAA

ПОСЛЕДОВАТЕЛЬНОСТЬ s4B9_1_R

SEQ ID №29

TTCCAGCTGGGGCACCGCTAATATCTTTAACCTCGCATCCCGGTGATGAAAGGATATTCTGGCTGCGTA
AGTAATGAATGAACCGCCAGCAGATAAAATATTGACAGTGATAACCCGATGTTTTTTTAAACGATGCAG
GCTATACATATAACATAGCTGGCCACCAACACAGCTGAAGTAAATCATATTGTTGCTGCCAGGCTACTT
CACACTATTGTCCGGCGGGCCAGCGGGGATTTTCCCCCTAAATCTCGCTGGTTCTCAA

7. p4F8

ПОСЛЕДОВАТЕЛЬНОСТЬ p4F8_1_I

SEQ ID №30

AGTCTACGATTTTCGCTATATCTTCTCTTAATCATGGCCGCCATTTGTGGATGCGATTTTAAATATCCG
GGCGATCTTTTCATTAAAAAATAAGATTCCCCATGACTTCACAGATAAAGGTATCGGTATTTTGAGTGA
TACGTAACAATTCGTTCTCTTCGTGTGGGTCCATGATGCGAAGAATAATGGTGGCATCATTTTCATGAG

ФИГ. 6 (продолжение)

GATTATGAACCCGAAATCTTTCTCTTTGCGATGCGCAGGCTAACTCTTTCAACTCAAAAAAATCTCTG
TAAGCCGCTCTCGTGTGGGGGCGC

8. p7B8

ПОСЛЕДОВАТЕЛЬНОСТЬ p7B8_1_0

SEQ ID №31

CGCCCCCTTAATTGGTTGAGGCGGCTGGTATTCTTGTAAGGGTAATACTAGCGAGACCCAGGTTCCAC
CCCCGGGACACTTTTGTAGTGTACGATTACCGCCCATCATTTTAGCCAGGCTTGACGCAATAGTCAGTC
CAATTCCTGTACCTTGCGAATTTGTGTCTGCTTGATAAAAAGCAGAAAAGATTTGAGACTGCTGCTGTT
TTTCAATCCCCCACCGCTATCGCTAACCAGAAATATTAATTGTTCTCACCAGATTGAGCGCCAGAC
GTATCCCTCCCCCTCGGGAAAT

9. p8G12

ПОСЛЕДОВАТЕЛЬНОСТЬ p8G12_1_I

SEQ ID №32

GGATAAGATCCCGGATAAGTATGTCAGGCTCGTATGCACAACAGGCATTATAAACCTCTAGACCATTTT
TAACATGCTCTACTATTTTAAATGAGGCCAGGGTAATAAGGCATTATAATGCCGTAAATGATGATTT
CATGATCGTCTACTAATAAGATCTTATATTCTTTTATTGCTGCCCTCGCGAAAATTAAGATAATATT
AAGTAATGGTGTAGGTTGTGGAGATCATACGTATTTTCTGGCGTAAGTCGGTTAGTTCTCAGCGCGA
TGATTTTCCCATTTTTTACGCGAT

10. p9G4

ПОСЛЕДОВАТЕЛЬНОСТЬ p9G4_1_0

SEQ ID №33

TTCCATATTGCTCGTCCGGGGAGCGTGTTAATTCTTGATGATATACCAATGGATCTGCAATGGCGCAAG
GTTCAACCATTACTTGGAGATATCCCGGGTTATTGTACTGGGAGATTAGTCACTCTCATCAGTCTCAG
GGGGGTGATGTTATTTCTGGGATAATAGAGCAACGGCGTTAGCAGGGGTGGGTGAGTACAGGCCAA
CTTCGGTGCACTTTTGGGTATCACTGGGGTATCATACTGAATCTCATCCCCCCTTTGGTAATCAC
AC

ПОСЛЕДОВАТЕЛЬНОСТЬ p9G4_1_U

SEQ ID №34

AATTCTTTTACCTCCATAAGCTGCGTGGCATAGCGATACAGAGTATTAAGCGGGTGTGTACATCGTCA
TCCAACAACATACGCAGCGAGCCGCCACGCCGAAAAACCGCATCGTGTATGTGCCTGTTGTAGGGTC
GGGTCTTTTTTTCATGAGTACGTGTTCTGCGCTATCATACTGGAAATTTCCCCCCTTACTGATAAGC
CCTGTAGTTGGGTAAGGACAGCGTTAAGCTCCTGAGACATTTTTTGGATTGTTATCTGCCCCCGACT
CATAAGATCGGGTATTCCGCGGTGG

11. p9B6

ПОСЛЕДОВАТЕЛЬНОСТЬ p9B6_1

SEQ ID №35

ATATCCCTAATGCTTTTCCCTTAAATAAATACCACGGAAGGATACTGGCCACCTAGCCAAATTTAGAAA
GCAATGAACATCCGGTTTATTCCTGAAAACGATTACTCCGGCGCACGTTGTTCTGGCGTTACCTGAGCC
AGCAAACGATATAATGGGGTGGTGACCCGCATACCGGTCATTGGCATCCCATCCACACCGGAGGGAGTA
AAACTCATTAGGCCATAGGTAATATCATTAAGACCGCTAATAAATGAGGGTGGGGGGCCAACTACC
ACTCCAGTATGTATTGAGTCA

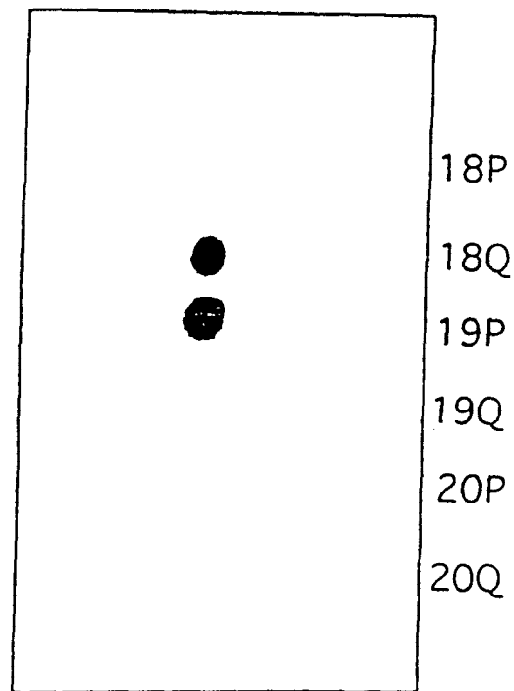
12. p6G5

ПОСЛЕДОВАТЕЛЬНОСТЬ p6G5_2_I

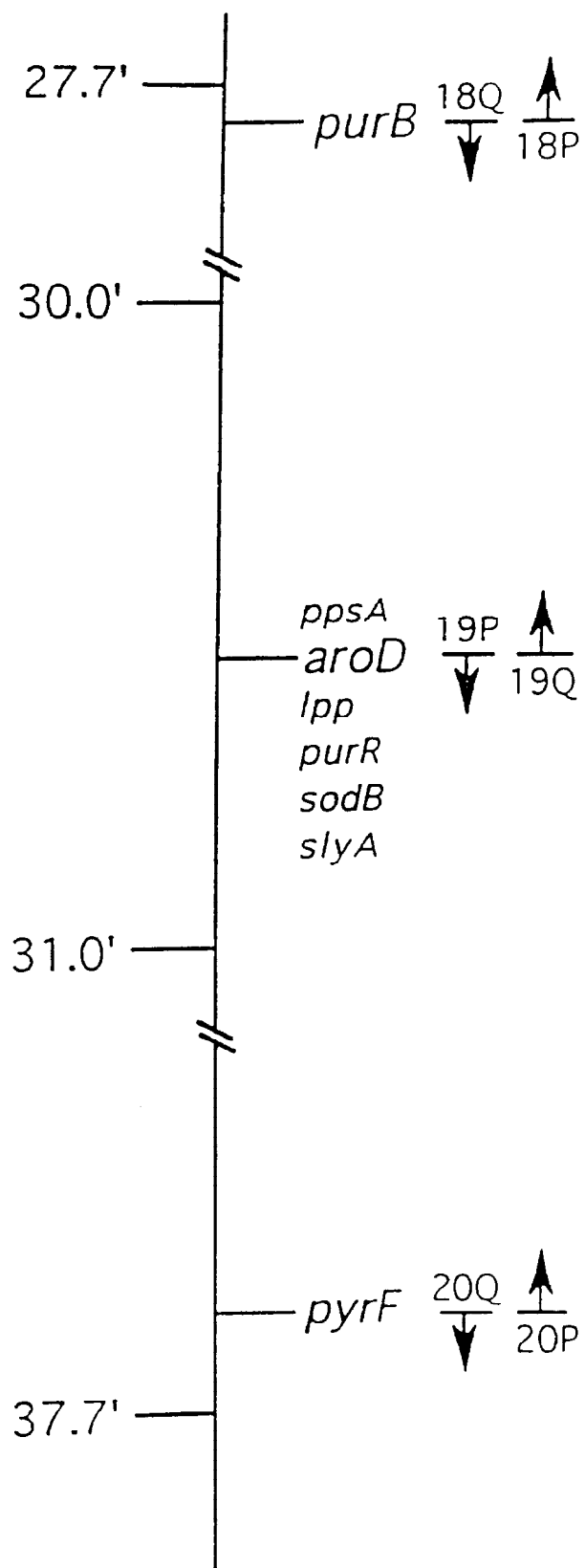
SEQ ID №36

CCCATGGGCGCAATTTGTTGCGCAGCGTTTACCCGACCATCGCGTTTATGAGCTGTAATTCATGGGGG
TAAAAACGGGCGTGACGACCCCAACGGAAGATAAGCCCGGGCTTAAACAGGAGATTATTGCTAATGCGC
AGCGCAAAGTGTGCTGGCGGACAGCAGTAAGTATGGCGCGCATTCGCTCTTTAATGTGGTGCCGCTTG
AGCGCTTTAATGACGTGATTACCGACGTCAATCTGCCGCGCTCAGCGCAGGTTGAACTGAAAGGGCGCG
CTTTTTGCGCTAACC

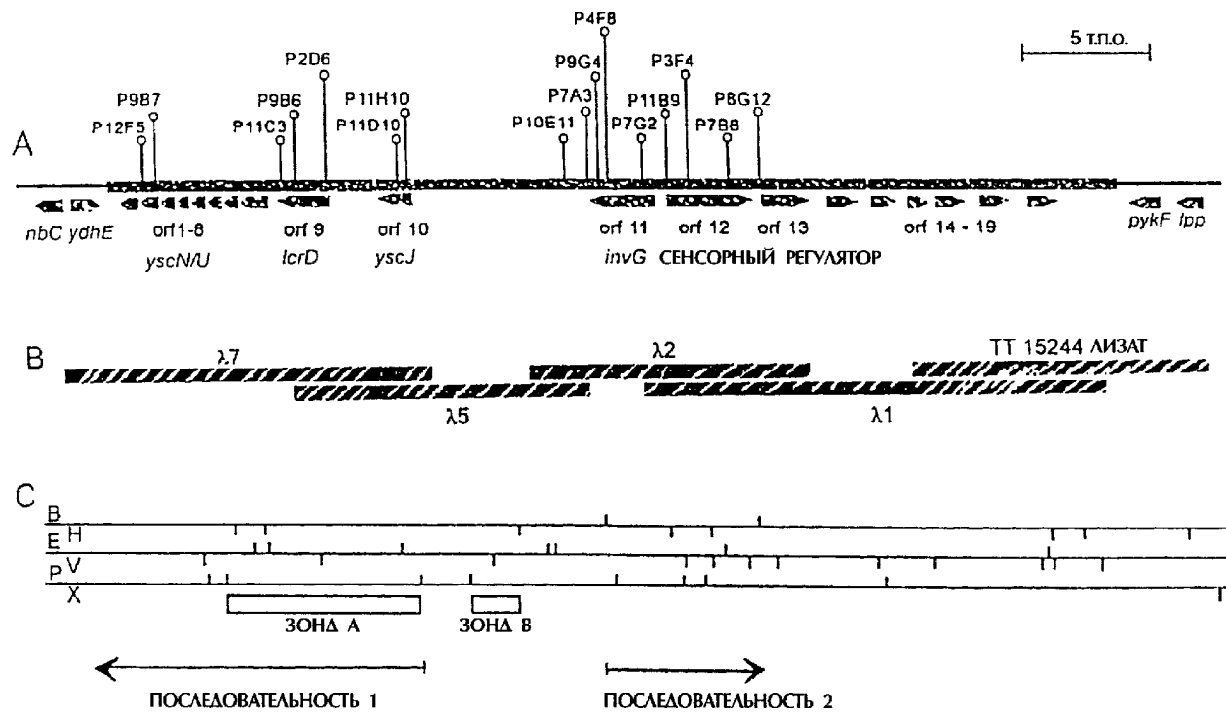
ФИГ. 6 (продолжение)



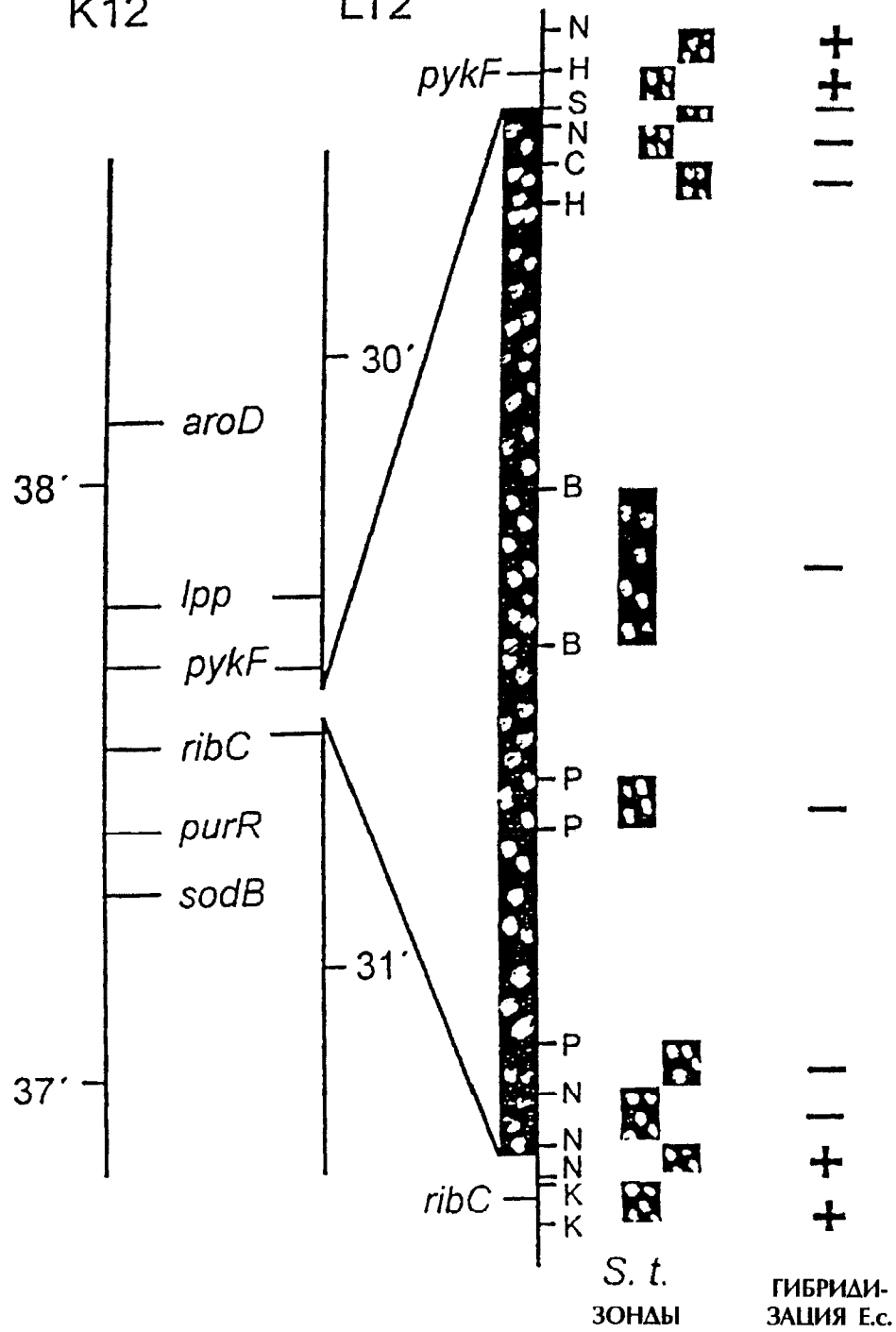
ФИГ. 7А



ФИГ. 7В



ФИГ. 8



ФИГ. 9

A

Pst I

S. typhimurium
S. aberdeen
S. gallinarum
S. cubana
S. typhi
 EPEC
 EHEC
Y. pestis
S. flexneri
V. cholera
S. aureus

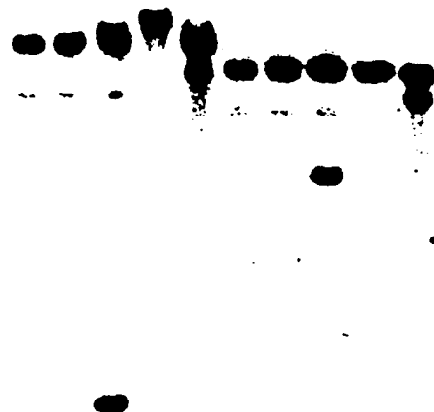


B

Hind III

Eco RV

S. typhimurium
S. aberdeen
S. gallinarum
S. cubana
S. typhi
S. typhimurium
S. aberdeen
S. gallinarum
S. cubana
S. typhi



ФИГ. 10

ДНК-ПОСЛЕДОВАТЕЛЬНОСТЬ ОТ ЦЕНТРА ДО ЛЕВОГО КОНЦА

CTGCAGAACCGAGCCAGGAGCAAAATTAATTTTTTTGAACAATTGCTGAAGATGAAGCATCCACCAGTAACGCCAGTGCT
 1 GACGCTCTGGCTCGGTCTCGTTTAATTAATAAACTTGTTAACGACTTCTACTTCGTAGGTGGTCATTGCGGTACCGA 80

L Q N R A R S K L I F L N N C * K H K H P P V T P V L -
 C R T E P G A N * F F * T I A E R * S I H Q * R Q C F -
 A E P S O E Q I N F F E Q L L K D E A S T S N A S A -

TTATTACCGCAGGTTATGTTGACCAGACAAATGGATTATATGCAGTTAAGGTTAGGCGTCGATTATCTTGCAGAAATATC
 81 AATAATGGCGTCCAATACAACCTGGTCTGTTTACCTAATAATACGTCAATTGCCATCCGCAGCTAATAGAACGGTCTTATAG 160

Y Y R R L C * P D K M I I C S * R * A S * I L P E Y H -
 I T A G Y V D Q T N G L Y A V N G R R R L S C Q N I -
 L L P Q V H L T R Q H D Y M Q L T V G V D Y L A R I S -

ACGGCGCAGCATGCCAAGCGCTTAATAAGCTGGATAACATGGCATGAAGGTTTCATCGTATAGTATTTCTTACTGTCTTA
 161 TCGCGCGTCGTACGGTTCGGCAATTATTGACCTATTGTACGTAAGTCCCAAGTAGCATATCATAAAGAATGACAGGAAT 240

G A A C O A L N K L D N H A * R F I V * Y F L L S L -
 T A O H A K R L I S W I T H H E G S S Y S I S Y C P Y -
 R R S H P S A * * A G * H G H K V N R I V F L T V L * -

CGTCTTTCTTACGGCATGTGATGTGGATCTTTATCGCTCATTCGCAGAGATGAAGCGAATCAATGCTGGCATTACTT
 241 GCAAGAAAGAATGCCGTACACTACACCTAGAAATAGCGAGTAACGGTCTTCTACTTCGCTTAGTTACGACCGTAATGAA 320

НАЧАЛО *yscJ*?
 R S F L R H V H W I F * A H C Q K M K R I K C W H Y L -
 V L S Y G H * C G S L S L I A R R * S E S N A G I T Y -
 F F L T A C D V D L Y R S L P E D E A N Q M L A L L -

ATGCAGCATCATATTGATGCGAAAAAACAGGAAGAGGATGGTGAACCTTACGTGTCGAGCAGTGGCGAGTTATTAA
 321 TACGTCGTAGTATAACTACGCTTTTTTTTGCTCTTCTCTACCACATTGGAATGCACAGCTCGTCAGCGCTCAAAATATT 400

НАЧАЛО *yscJ*?
 C S * I L H R K K T G R G M C M L T C R A V G S L L M -
 A A S Y * C E K K K Q E E D G V T L R V E Q S A V Y * -
 M Q H H I D A K K N R K R M V * P Y V S S S R Q F I N -

TGCGGTTGAGGCTACTTAGACTTAACGGTTATCCGCATAGGGCAGTTTACAACGGCGGATAGATGTTTCCGGCTAATCA
 401 ACGGCACTCCGATGAATCTGAATTGCCAATAGGCGATCCCGTCAATGTTGCCGCTATTCTACAAAGCCGATTAGT 480

R L R L L R L N G Y P H R A V Y M G G * D V S G * S -
 C G * G Y L D L Y V I R I G Q F T T A D K H F P A N Q -
 A V E A T * T * R L S A * G S L Q R R I R C F R L I S -

GTTAGTGGTATCACCCAGGAAGAAGAGGCAAGATTATTTTAAAGAACAAGAATTGAAGGAATGCTGAGTCAG
 481 CAATCACCATAGTGGGGTCTTCTTGTCCGTCTCTAATTAATAATTTCTTGTTCTTAACCTCCTTACGACTCAGTC 560

V S G I T P G R T G R R L I F * K N K E L K E C * V R -
 L V V S P Q E E Q A E D * F F K R T K N * R N A E S D -
 * H Y H P R K N R Q K I M F L K E Q R I E G H L S Q -

ATGGAGGGCGTGATTAATGGCAAAAGTGACCATTCGCGTACCGACTTATGATGAGGGAAGTAACGCTTCTCCGAGCTCA
 561 TACCTCCCGCACTAATTACCGTTTTCACTGGTAACGGGATGGCTGAATACTACTCCCTTCATTGCGAAGAGGCTCGAGT 640

H R G V I N G K S D H C A T D L * * G K * R F S E L S -
 C G A * L H A K V T I A L P T Y D E G S N A S P S S -
 H E G R D * M Q K * P L R Y R L M H R E V T L L R A Q -

GTTGCCGATTTATAAAATATTACCTCAGGTCAATATGGAGGCCCTTCCGGTAAAAATTAAAGATTAAATAGAGATGTC
 641 CAACGGCATAAATATTTTATAAGTGCAGTCCAGTTATAGCTCCGGAAGCCCATTTTAAATTTCTAAATTTATCTCTACAG 720

C R I Y K I F T S G Q Y G G L S G K N * R F N R D * V -
 V A V F I K Y S P Q V N H E A F R V K I K D L I E M S -
 L P Y L * N I H L R S I W R P F G * K L K I * * R C Q -

ПОСЛЕДОВАТЕЛЬНОСТЬ 1

SEQ ID №37

ФИГ. 11

SATCCCTGGGTGCAATACAGTAAGATTAGTATCTTGATGCAGCCTGCTGAATTCAGAAATGGTAGCTGACGTACCCGGG/ 800
 ITAGGGACCCCAACGTTATGTCATTCTAATCATAGAATACGTCGGGACGACTTAAGTCTTACCATCGACTGCATGGGCGCT
 N F W V A I Q D Y L D A A C I O N G S R T R E -
 I P G L Q Y S K I S I L M Q P A E F R M V A D V F A A -
 S L G C N T V R L V S C S L L N S E W L T Y P R -
 GACAACCATTTCTGGATTATGGACGTTATCAACGCCAATAAAGGGAAGGTGGTGAAGTGGTTGATGAAATACCCCTTATCCG 880
 CTGTTTGTAAAGACCTAATACCTGCAATAGTTGCGGTTATTTCCCTCCACCACTTCACCAACTACTTTATGGGAATAGCG
 T N I L D Y G R Y Q R Q R E G G E V V D E I P L S V -
 Q T F W I M D V I N A N K G K V V K W L M K Y P Y P -
 D K E S G L W T L S T F I K G R W S G N T L I R -
 ИНСОПЦИЯ P11H11
 TTGATGTTATCGTTGACAGGACTGTTATTAGGAGTGGGCATCCTGATCGGCTATTTTGCCTGAGACGCCGTTTTTGAGC 960
 AACTACATAGCAACTGTCCTGACAATAATCCTCACCCTAGGACTAGCCGATAAAACGGACTCTGCGGCAAAACACTCG
 D V E V D R T V I R S G H P D R L F L P E T P F L S -
 L M L S L T G L L G V G I L I G Y F C L R R P F A -
 C R Q D C Y E W A S S A I F A D A V F E F -
 CGACCTGATCCCGAGGTGTTGCAACTTTATCGTTATTTCTGGCAACCTGCTCGTTACGCTGTACCGGAATGGCTGGATA 1040
 GCTGGACTAGGGCTCCACAACGTTGAAATAGCAATAAAGACCGTTGGACGAGCAATGCGACATGGCCTTACCGACCTATT
 R P D P E V L Q L Y R Y F W O P A R Y A P E W L D K -
 D L I P R C C N F I V I S G N L L V T I R N G W I S -
 T S R G V A T L S L F L A T C S I K C T G M A G -
 ИНСОПЦИЯ P11D10
 GCTGGGCTTTTCATCTTCAAACCTGCTGGCGTTATGGCGATCGGCCCGAGTTGGATCGTCTTCTTGACAGAGCGTTAAATAG 1120
 CGACCCGAAAGTAGAAGTTTGACGACCCGAATACCGCTAGCCGGGCTCAACCTAGCAGAAGACTGTCTCGCAATTTATC
 L G F H L Q T A G V M A I G P S W I V F L T E R I D -
 W A F I F K L L A L W R S A R V G S S F Q S V K -
 A G L S S S N C W R Y G D R P E L D R L L D R A E N F -
 ACTAAGAGGAAGCTCTGTTATTCAGCGCTTTTAAATGACAGGCAAAACGGCAGGTTCTGCTTGCCCGCGGTATATCGG 1200
 TGATTCCTCTCGAGACAATAAGGTCGGACAAATTTACTGTCCGTTTTTGCCGTCCAAGCGAAGCGGCGCATATAGCC
 E E A L L F Q P V M T G K N G E F L K R V Y F -
 T K F F L C Y S S L F H Q A K T A G S E C A A Y I G -
 L R G S S V I P A C L N D R Q K R Q V F I P P I S A -
 CATTTGCCTTTTGGGCTGGGATTATTCAAACCTCAGGTGTAGTGACTATTTTATGCTACCAGAGTATCGGCAATTGCTTCTA 1280
 GTAAACGGAACCCGACCTAATAAGTTTGAGTCCACATCACTGATAAAATACGATGGTCTCATAGCCGTTAACGAAGAT
 НАЧАЛО lcrE?
 H L F L G W D Y S N S G V V T I L C Y Q F I G N C F Y -
 I C L W A G I I Q T O V L F Y A T P V S A I A S T -
 F A F G L G L F K L R C S D Y F H I P E S S O L L L -
 CAGTGGTTTAGCGAGGATGAGATCTGGCAGCTATATGGTTGGTTGGGGCAAGAGATGGCAATTACTTCTCCGCAAGT 1360
 GTCACCAATCGCTCCTACTCTAGACCGTCGATATACCAACCAACCCCGTTTCTCTACCGTTTAAATGAAGGAGGCGGTTCA
 S G L A R M R S G S Y M V G W G K E M A N Y F L R K -
 V V R C D L A A I W L V G A K R W Q I T S S A S -
 Q W F S E D E I W Q L Y G W L G Q R D G L L P P Q -
 GATGCAACAAACTGCATTGCAGATCGGTACCGCCATTCTTAATCGGGAAGCGCATGACGATCGGGGTTTTACATGCGCTA 1440
 CTACGTTGTTTGACGTAACGCTAGCCATGGCGGTAAAGATTAGCCCTTCGGGTACTGCTACGCCCAAAATGTACGGCAT
 C N E L H C E S P F E L I G K R M T N V L H A I -
 D A T N C I A D F Y R H S S G S A K C G E Y M R -
 H Q Q T A L Q I G T A I L N R E A P D D A G F T C A C -

ФИГ. 11 (продолжение)

```

TTAGTATTATTACCCCTCCGAGCGTATACTTTGGCCGAAGACTTCTCTTACCGAGATTATCTTCATGGAGCATTTGCT
1441 -----+----- 1520
AATCATAATAATGGGGGAGGCGTCGCATATGAAACCGGCTTCTGAAGAGAATGGCTCTAATAGAAGTACCTCGTAAACGA
L V L L P P F O R I L W P K T S L T E I I F M E H L L -
Y Y Y P L K S V Y F G R R L L L P A L S S W S I C Y -
S I I T F S A A Y T L A E D F S Y R D Y L H G A F A -
ATGAGTTTTACTTCACTTCTCTGACGGAATTAACCATAAGCTACCCGCTCGAAATATTATTGAGTCACAGTGGATAAC
1521 -----+----- 1600
TACTCAAAATGAAGTGAAGGAGACTGCCTTTAATTGGTATTGATGGGCGAGCTTTATAATAACTCAGTGTACCTATTG
V L L H F L R K L T I S Y P L E I L L S H S G H -
E F Y F T S S D G N P A T R S K Y Y V T V D N -
M S F T S L P L T E I N H K L F A R N I I E S Q W I T -
ATTACAATTAACCTTTATTTGCGCAAGAGCAACAAGCTAAGAGAGTTTACATGCTATTGTGAGCTCCGCTTACCGTAAGG
1601 -----+----- 1680
TAATGTTAATTGAAATAAACCGGTTCTCGTTGTTTCGATTCTCTCAAAGTGTACGATAACACTCGAGGCGAATGGCATTCC
Y N L Y L R K S N K L R E F H M L L A P L T V R -
T I N F I C A R A T S E S F T C Y C E L R L P G -
L Q L T L F A Q E Q Q A K R V S H A I V S S A Y R K A -
CTGAAATAATCATCCGAGACGCTATCGTTATCAGCGTGAACAGAAAGTTGAGCAGCAACAAGAACTAGCGTGCTTGGCT
1681 -----+----- 1760
GACTTTTTTAGTAGGCTCTGCGGATAGCAATAGTCGCACTTGTCTTCAACTCGTCGTTGTTCTTGATCGCAGCAACGCA
L K K S S E T P I V I S V N R K L S S H E N R A C V -
K N H P R R L S L S A T E S A A T R T S V L A -
E V I I R D A Y R Y Q R E Q K V E Q Q Q E L A C L R -
AAAAATACGCTGGAAAAATGGAAGTGGAAATGGCTGGAACAGCATGTAAACATTTACAAGACGATGAAATCAATTTCCG
1761 -----+----- 1840
TTTTTATGCGACCTTTTTTACCTTACCTTACCGACCTTGTGCTACATTTTGTAATGTTCTGCTACTTTTAGTTAAAGC
K I R W K K W K W N G W N S M N I Y K T M K I N F V -
K Y A G K N G S G M A G T A C K T F T R R K S I S -
K N T L E K M E V E W L E Q H V K H L Q D D E N Q F R -
TTCATTGGTCGATCAGCGAGCGCATCATATTAATAATAGTATAGAACAGGTTCTGTTGGCCTGGTTCGACCAACAGTCGG
1841 -----+----- 1920
AAGTAACCGAGCTAGTGGCTCGCGTAGTATAATTTTATCATATCTTGTCCAAGACACCCGACCAAGCTGTTGTACAGCC
H K S I T Q R I I L F I V N R F C W P G S T H S R -
F I G R S R S A S Y K Y R T G S V G L V R P T V G -
S L L D H A A H H I K N S I E Q V L L A W F D Q Q S V -
TAGACAGTGTATGTGCCATCGTCTGGCAGCCAGGCCACGGCTATGGCGGAAGAGGGAGCGCTTTATTTGCGTATTCAT
1921 -----+----- 2000
ATCTGTCAATAACCGGTAGCAGACCGTGGGTCGGTGCCGATACCGCCTTCTCCCTCGCGAATAAACGATAAGTA
T V L C A I V W H A R P R L W R K E E R F I C F I -
R Q C V P S S G T P G H G Y G G R G S L L F A S S -
D S V M C H R L A R Q A T A M A E E G A L I L A I H -
CCTGAATAAGAGGCGATTGATGCGAGAACTTTTGGCAAGCGGTTTACGTTGATTATCGAGCCTGGTTTCTCTCCGATCA
2001 -----+----- 2080
GGACTTTTTCTCCGTAACCTACGCTCTTTGAAAACCGTTGCGCAATGCAACTAATAGCTCGGACCAAGAGAGGGCTAGT
L K K R H C E K L L A S G L R L S S L V S L F I R -
K R G I D A R N F W Q A V Y V D Y R A W F L S R S -
P E K E A L H R E T F G K R F T L I I E P G F S P D Q -
GGCTGAACCTTTCTCAACACGATATGCCGTTGAATTTTCACTTTCTCGTCATTTCAACGCGTTACTGAAATGGTTACGTA
2081 -----+----- 2160
CCGACTTGAAAGGAGTTGTGCTATACGGCAACTTAAAGTGAAAGAGCAGTAAAGTTGCGCAATGACTTTACCAATGCAT
L N F P Q H D M P L N F H F L V I S T R Y N G Y V -
G T F L N T I C R I F T F S S F Q R V T E M V T -
A E L S T R Y A V E F S L S R H R N A L L K W L P N -
ATGGTGAAGATAAAGAGGTAGCGATGAATATTAATAATGAGATAAAATGACGCCCCCTACAGCATTTACCCCTGG
2161 -----+----- 2240
TACCACCTCTATTTTCTCATCGCTACTTATAATTTTAATTACTCTATTTTACTGCGGGGGATGTCGTAATGGGGACC
M V R C A E V A M N I E I N E I K M T P F T A I T P G -
W R K F R I L K L M K K R P L O H L P L A -
G E D R R G S D E Y N D K N D A P Y S I Y P W -

```

Фиг. 11 (продолжение)

CCAGGTTATAGAGGAACAAGAGGTTATTTTCGCCCTTCAATGTTAGCTCTCCAGGAGTTACAGGAAACGACGGGGCAGCGC
 2241 2320
 GGTCGATATCTCCTTGTCTCCAATAAAGCGGAAGTTACAATCGAGAGGTCCTCAATGTCTTTGCTGCCCCGTCGGC
 a O V I E E Q E V I S P S M L A L O E L O E T T G A A L -
 b R L R N K R L F R L Q C L S R S Y R K R R G Q R -
 c P G S G T R G Y F A F N V S S P G V T G N D G G S A -

TCTATGAGACGATGGAAGAAATAGGAATGGCGCTGAGTGGTAAACTGCGCGAAAATTATAAATTCAGTATGCTGAGAAA
 2321 2400
 AGATACTCTGCTACCTTCTTTATCCTTACCGCGACTCACCATTGACGCGCTTTTAATATTTAAGTGACTACGACTCTTT
 a Y E T H E E I G M A L S G K L R E N Y K F T D A E K -
 b S M R R W K K E W R V V N C A K I I N S L M L R N -
 c L D D G R N R N G A E W T A R K L I H C E T -

CTGGAGCGCAGACAGCAGGCTTTGCTGCGTTTGATAAAACAAATACAGGAGGATAATGGGGCAACGTTGCGTCCGCTTAC
 2401 2480
 GACCTCGCGCTGTGCTGCCGAAACGACGCAAACTATTTTGTATGTCTCTCTATTACCCCGTTGCAACGCAGCGCAATG
 a L E R R Q Q A L L R L I K Q I Q E D N G A T L R P L T -
 b W S L Q S R L C C V N K Y R R I M G Q R C V R L P -
 c G A Q T A G F A A F D K T N T G G W G H V A S A Y -

CGAAGGAGATAGTGATCCTGATTACAGAATGCGTATCAAATTATCGCTCTTGCAATGGCGCTTACTGCCGGCGGGTTGT
 2481 2560
 GCTTCTCTTATCACTAGGACTAAATGTCTTACGCATAGTTTAAATAGCGAGAAGCTTACCGCGAATGACGGCGCGCAACA
 a E E N S D F D L O N A Y Q I I A L A M A L T A G S L S -
 b K R V I L I Y R M R I K L S L L Q W R L L P A G C -
 c R E S F T E C V S N Y R S C N G A Y C K R V V -

CAAAAAGGAAAAACGCGATTTGCAATCGCAACTGGATACGTTACAGCGGAGGAGGATGGGAAGTTCGCGTTTGTAGTT
 2561 2640
 GTTTTTCTTTTTGCGCTAAACGTTAGCGTTGACCTATGCAATGTGCGCTCTCCCTACCCTTGAACGGCAAAATCAA
 a K K K K R D L Q S Q L D T L Q R R R D G H L P F L V -
 b Q K R F N A I C N R N W I R Y S G G G M G T C R F F -
 c K K E K T R F A I A T G Y V T A E E G W E L A V F S L -

TACTGGAAGTTGGCGAAGTGGATACCGTACGCTGTCTCTGTAAGCGTTTATGCAACAGGCGATAGACAACGATGAAA
 2641 2720
 ATGACCTTGAACCGCTTACCTATGGCATGCGACAGGAGAGACTTCGCAAAATACGTTGTCCGCTATCTGTGCTACTTT
 a Y W N E A K W I P Y A V L S E A F Y A T G D R Q R N -
 b T G T W F S G Y R T L S S L K R F M Q Q A I U N D E M -
 c L E L G E V D T V R C P L S V L C N R R T T M K -

TGCCCTTATCGCAGTGGTTCAGACGCGTGGCAGACTGGCCGGATCGCTGTGAACGGGTCCGTATTTGTGAAGAGCAGTA
 2721 2800
 ACGGGAATAGCGTCACCAAGTCTGCGCACCGTCTGACCGGCTAGCGACACTTGCCAGGCATAAACGATTCTCGTCAT
 a A L E A V V Q T R G R L A G S L T G P F A K S S S -
 b P L S Q W F R R V A D W P D R C E R V R L L R A V -
 c C P Y R S G S D A W Q T G R I A V N G S V F C E Q -

GCCTTTGAAGTTAGCATATGCATCGAACCCTCGGAGCAAAGTCGTTTGGCCGCGAGCATTAGTACGTTTGGCTCGTTTGTCT
 2801 2880
 CGGAAACTTGAATCGTATACGTAGCTTGGGAGCCTCGTTTCAGCAAACCGGCGTGTATCATGCAAACGCAGCAAACGA
 a L T H M H R T L G A K S F G R S I S T F A S F A -
 b A F E L S I C I E P S E Q S R L A A L V R L R L L -
 c P L N L A Y A S N P R S K V V W P Q H Y V C V Y C C -

GTTATTCCTTGGCCTTGAAAAAGAGTGCCAGCGTGAGGAGTGGATTTGCCAGTTGCCGCTAATACATTACTGCCGCTAC
 2881 2960
 CAATAAGGAACCGGAACTTTTCTCACGGTTCGCACTCCTCACCTAACGGTCAACGGCGGATTATGTAATGACGGCGATG
 a V I P W P K R V P A G V D L P V A A Y I T A A T -
 b L F L G L E K E C Q R E E W I C Q L P P N T L L P L L -
 c Y S L A L K K S A S V R S G F A S C K L I H Y C R Y -

TACTCGATATTTTGTGAGCGCTGGCTTTTTCAGTGATTGGTTGCTTGATAGACTTACCGCTATAGTTTCTTTCATCGAAG
 2961 3040
 ATGAGCTATAATAACACTCGCGACCGAAAAGTCACTAACCAACGAATATCTGAATGGCGATATCAAAGAAGTAGCTTC
 a T P Y Y I A L A F Q L V A T Y R Y S F F I E D -
 b L D I I C E R W L F S D W L L D R L T A I V S S S K -
 c Y S I L F V S A G F S V I G C L I D L P L F L H R R -

Фиг. 11 (продолжение)

ATGTTCAATCGGTTACTCCAACAACCTTGATGCGCAGTTTATGCTGATACCGGATAACTGTTTTAACGACGAAGATCAACG
 3041 ----- 3120
 TACAAGTTAGCCAATGAGGTTGTTGAACCTACGCGTCAAATACGACTATGGGCTATTGACAAAAATTGCTGCTTCTAGTTGC
 V Q S V T P T T * C A V Y A D T R * L E * R R R S T -
 M F N K L L Q Q L D A Q F M L I P D N C F N D E D Q R -
 C S I G Y S N N L M R S L C * Y P I T V L T T K I N V -
 TGAACAAATTCTCGAAACGCTTCGTGAAGTAAAGATAAATCAGGTTTTATTCTGATACCTGGCTTTCAATATTTAGGTAA
 3121 ----- 3200
 ACTTGTTTAAGAGCTTTGCGAAGCACTTCATTCTATTTAGTCCAAAATAAGACTATGGACCGAAAGTTATAAATCCATT
 * T N S R N A S * S K D K S G F I L I P G F Q Y L G K -
 E Q I L E T L R E V K I N Q V L F * Y L A F N I * V N -
 N K F S K R F V K * R * I R F Y S D T W L S I F R * -
 ATTGCGCTTTCTGGCTCATCATGAGGCGTCAGGATGGATTGGGATCTCATTACTGAACGTAATATTCAGCTTTTTATTCAA
 3201 ----- 3280
 TAACCGAAAGACCGAGTAGTACTCCGAGTCCTACCTAACCTAGAGTAATGACTTGCATTATAAGTCGAAAATAAGTT
 L A F W L I M R R Q D G L G S H Y * T * Y S A F Y S I -
 W L S G S S * G V R M D W D L I T E R N I Q L F I Q -
 I G F L A H H E A S G W I G I S L L N V I F S F L F N -
 TTAGCAGGATTAGCTGAACGGCCTTTAGCAACCAATATGTTCTGGCGGCAAGGACAATATGAACTATCATAACGGTCGT
 3281 ----- 3360
 AATCGTCTAATCGACTTGCCGGAATCGTTGGTTATACAAGACCGCCGTTTCTGTTATCTTTGATAGTATTGCCAGCA
 S R I S * T A F S N Q Y V L A A R T I * H Y H N G R -
 L A G L A E R P L A T N M F W R Q G Q Y E T I I T V V -
 * Q D * L N G L * Q P I C S G G K D N M K L S * R S Y -
 ATTCTCTTATGTCAGATACTCAAGCAAACCTTCTTAGACGAAGAACTGCTTTTTAAAGCGTTGGCTAACTGGAAACCCGC
 3361 ----- 3440
 TAAGAGAATACAGTCTATGAGTTCGTTTGGAGAATCTGCTTCTTGACGAAAAATTTCGCAACCGATTGACCTTTGGCGC
 I L L C Q I L K Q T F L D E E L L F K A L A N W K P A -
 F S Y V R Y S S K P S * T K N C F L K R W L T G N P Q -
 S L M S D T Q A N L L R R R T A F * S V G * L E T R -
 AGCGTTCAGGGTATTCTCAACGATTATTTTTGTTGCGCGATGGGCTTGCAATGAGTTGTTCTCCACCTCTTTCCAGCT
 3441 ----- 3520
 TCGCAAGGTCCCATAAGGAGTTGCTAATAAAAACAACGCGCTACCCGAACGTTACTCAACAAGAGGTGGAGAAAGGTGCA
 A F Q G I P Q R L F L L R D G L A M S C S P P L S S S -
 S R V F L N D Y F C C A M G L Q * V V L H L F P A -
 S V F G Y S S T I I F V A R W A C N E L F S T S F Q L -
 CCGCCGAGCTCTGGTTACGATTACATCATCGACAATAAAATTTCTGAGTTCGCAATGCGTTTATGTTAGTGAGGGA
 3521 ----- 3600
 GCGGGCTCGAGACCAATGCTAATGTAGTAGCTGTTATTTTAAAGXACCTCAGCGTTACGCAAGTACCAATCCACTCCCT
 A E L W L R L H H R Q I K F ? G V A M R S W L S E G -
 P P S S G Y D Y I I D K * N F ? E S Q C V H G * V R E -
 R R A L V T I T S S T N K I S W S R N A F M V R * G S -
 GTCAGGGCGCAACAGTGGCTCAGTGTATGCGGGGTGGCAGGATATGGTTCTGGCGACGGTGTATTAAATCGCTATTGT
 3601 ----- 3680
 CAGTCCCGCGTTGTACCGAGTCACATACGCGCCAGCCGTCCTATACCAAGACCGCTGCCACAATAATTAGCGATAACA
 начало *lcrD*
 V R A Q Q W L S V C A G R Q D M V L A T V L L I A I V -
 S G R N S G S V Y A R V G R I W F W R R C Y * S L L -
 Q G A T V A Q C M R G S A G Y G S G D G V I N S Y C -
 GATGATGCTGTTACCGTTGCCGACCTGGATGGTTGATATCCTGATTACTATCAACCTTATGTTTTAGTGATCCTGCTCT
 3681 ----- 3760
 CTACTACGACAATGGGAACGGCTGGACCTACCAACTATAGGACTAATGATAGTTGGAATACAAAAGTCACTAGGACGAGA
 M M L L P L P T W M V D I L I T I N L M F S V I L L L -
 * C C Y P C R P G W L I S * L L S T L C F Q * S C S -
 D D A V T L A D L D G * Y P O Y Y Q P Y V S D F A L -

Фиг. 11 (продолжение)

TAATIGCTATTTATCTTAGTGACCCTCTCGATTTATCGGTATTTCCGTCTTTATTACTTATTACTACATTATATCGTTTG
 3761 ----- 3841
 ATTAACGATAPATAGAATCACTGGGAGAGCTAATAGCCATAAAGGCAGAAATAATGAATAATGATGTAATATAGCAAC
 a I A I Y L S D P L D L S V F P S L L L I T T L Y P L -
 b L L F I L V T L S I Y R Y F R L Y Y L L H Y I V C -
 c N C Y L S * * P S R F I G I S V F I T Y Y Y I I S F A -
 TCACTCACAAATCAGCACATCAGCGGTGGTACTGTTACAACATAATGCCGGTAATATTGTGGATGCTTTCCGTAAGTTTGT
 3841 ----- 3920
 AGTGAGTGTAGTCGTGTAGTGCCGACCATGACAATGTTGTATTACGGCCATTATAACACCTACGAAAGCCATTCAAACA
 a S L T I S T S R L V L L Q H N A G N I V D A F G K F V -
 b H S Q S A H H G W Y C Y N I M P V I L W M L S V S L S -
 c T H N Q H I T A G T V T T * C R * Y C G C F R * V C -
 CGTAGGAGGAAATCTCACCGTTGGGTGGTGGTATTTACCATCATTACTATCGTGCAATTTATTGTCAATACAAAAGGTA
 3921 ----- 4000
 GCATCCTCCTTAGAGTGGCAACCCAACCAAGCATAAATGGTAGTAATGATAGCACGTTAAATAACAGTAATGTTTTCCAT
 a V G G N L T V G L V V F T I I T I V Q F I V I T K G -
 b * E E I S P L G W S Y L P S L L S C N L L S L Q K V -
 c R R R K S H R W V G K I Y H H Y Y R A I Y C H Y K R Y -
 TCGAGAGGGTGGCGAAGTTAGCGCACGTTTCTCGCTTGATGGGATGCCAGGCAAAACAATGAGTATCGATGGCGATTTC
 4001 ----- 4080
 AGCTCTCCCGCGCTTCAATCGCGTGCAAAGAGCGAACTACCTACGGTCCGTTTGTCTACTCATAGCTACCGCTAAAC
 a E R V A E V S A R F S L D G M P G K Q M S I D G D L -
 b S R G W R K L A H V S R L M G C Q A N K * V S M A I C -
 c R E G G G S * R T F L A * W D A R Q T N E Y R W R F A -
 инсерция P2D6
 CGTGCCGGAGTTATCGATGCAGACCATGCCCGTACATTAAGACAGCATGTCCAGCAGGAAAGCCGCTTTCTCGTGCGAT
 4081 ----- 4160
 GCACGGCCTCAATAGCTACGTCTGGTACGGGCATGTAATCTGTCTGACAGGTGTCCTTTCCGGGAAAGAGCCACGCTA
 a R A G V I D A D H A R T L R Q H V Q Q E S R F L G A M -
 b V P E L S M Q T M P V H * D S M S S R K A A F S V R W -
 c C R S * R C R P C P Y I K T A C P A G K P L S R C D -
 GGACGGTGGGATGAAATTTGTTAAAGGCGATACGATTGCCGGTATTATTGTTGTTCTGGTGAACATTATCGGCGGTATCA
 4161 ----- 4240
 CCTGCCACGCTACTTTAAACAATTTCCGCTATGCTAACGGCCATAATAACAACAAGACCACTTGTAAATAGCCGCATAGT
 a D G A M K F V K G D T I A G I I V V L V N I I G D I -
 b T V F * N L L K A I R L P V L L L F W * T L S A V S -
 c G R C C E I C * R P Y E C R Y Y C C S G E H Y R A Y H -
 TTATCGCTATCGTACAATATGATATGTCGATGAGTGAGGCTGTTACACCTTATAGCGTACTGTCAATCGGAGATGGTTTA
 4241 ----- 4320
 AATAGCGATAGCATGTTATACTATACAGCTACTCACTCCGACAAGTGTAATATCGCATGACAGTTAGCCTCTACCAAAAT
 a I A I * Q Y D M S M S E A V H T Y S V L S I G D G L -
 b L S L S * N M I C R * V R L F T L I A Y C Q S E M V Y -
 c Y R Y S T I * Y V D E * G C S H L * R T V N R R W F M -
 TGTGGGCAAATTCATCGCTGCTGATTTCCCTTAGCGCGGGAATTATTGTCACCCGTGTCCGGGTGAGAAACGCCAGAA
 4321 ----- 4400
 ACACCCGTTTAAAGGTAGCGACGACTAAAGGGAATCGCGCCCTTAATAACAGTGGGCACAGGGCCCACTCTTTGCGGTCTT
 a C G Q I F S L L I S L S A G I I V T R V P G E K R Q N -
 b V G K F H K C * F P L A R E L L S P V S R V R N A R T -
 c W A N S I A A D F P * R G N Y C H P C P G * E T P E -
 CCTGGCGACAGAGTTGAGTTCTCAAATTCAGACACAACCTCAGTCGCTCATATTAACCGCTGTGGTTTTAATGCTCCTCG
 4401 ----- 4480
 GGACCCGCTGTCTCAACTCAAGAGTTTAAACGGTCTGTTGGAGTCAGCGAGTATAATTGGCGACACCAAAATACGAGGAGC
 a L A T E L S S Q I A R Q P Q S L I L T A V V L M L L A -
 b W R Q S * V L K L P D N L S R S Y * P L W F * C S S -
 c P G D R V E F S H C Q T T S V A H I N R C G F H A P R -

Фиг. 11 (продолжение)

CTTTAATTCCTGGCTTTCTTTTATCACTCTCGCTTTCTTTTCAGCGTTGTTAGCATTGCCAATTATCCTCATTCCGCCG
 4481 ----- 4560
 GAAATTAAGGACCGAAAGGAAATAGTGAGAGCGAAAGAAAAGTCGCAACAATCGTAACGGTTAATAGGAGTAAGCGCG

a L I P G F P F I T L A F F S A L L A L P I I L I R R -
 b L * F L A F L L S L S L S F Q R C * H C Q L S S F A A -
 c F N S W L S F Y H S R F L F S V V S I A N Y P H S P Q -

инсерция P11C3
 u

AAAAAGTCTGTGGTTTCCGCAATGGCGTCGAAGCACCGGAAAAAGATAGTATGGTTCCCGGCGCATGCTCTAATCTT
 4561 ----- 4640
 TTTTTCAGACACCAAAGGCGTTTACCGCAGCTTCGTGGCCTTTTCTATCATACCAAGGGCGCGTACAGGAGATTAGAA

a K K S V V S A N G V E A P E K D S M V P G A C P L I L -
 b K S L W F P O M A S K H R K K I V W F P A H V L * S Y -
 c K V C G F R K W R R S T G K R * Y G S R R M S S N L -

ACGCTTAGCCCGACGTTACATTCTGCCGACCTGATTCTGTATATTGACGCCATGAGATGGTTTTATTTAGGATACCG
 4641 ----- 4720
 TGCAGAATCGGGCTGCAATGTAAGACGGCTGGACTAAGCACTATAACTGCGGTACTCTACCAAAAAATAACTCCTATGGC

a R L S P T L H S A D L I R D I D A M R W F L F E D T G -
 b V L A R R Y I L P T * F V I L T P * D G F I L R I P -
 c T S * P D V T F C R P D S * Y * R H E M V F I * G Y P -

CGCTCCCTCTCCCTGAGGTGAATATTGAGGTTTTGCCTGAACCCACCGAAAAATTGACGGTACTGCTATATCAGGAACCC
 4721 ----- 4800
 CGCAGGGAGAGGGACTCCACTTATAACTCCAAAAAGGACTTGGGTGGCTTTTTAACTGCCATGACGATATAGTCTTTGGG

a V P L P E V N I E V L P E P T E K L T V L L Y Q E P -
 b A S L S L R * I L R F C L N P P K N * R Y C Y I R N P -
 c R P S P * G E Y * G F A * T H R K I D G T A I S G T R -

GTATTTAGTTTATCTATTCCCGCTCAGGCGGATTATTTATTGATAGGCGCGGACGCTAGTGTGGTGGGTGACAGCCAGAC
 4801 ----- 4880
 CATAAATCAATAGATAAGGGCGAGTCCGCTAATAAATAACTATCCGCGCTGCGATCACACCACCCACTGTCGGTCTG

a V F S L S I P A Q A D Y L L I G A D A S V V G D S Q T -
 b Y L V Y L F P L R R I I Y * * A R T L V W W V T A R R -
 c I * F I Y S R S G G L F I D R R G R * C G G * Q P D -

GTTACCGAACGGGATGGGGCAGATCTGTGGCTTACAAAAGACATGGCCCATAGGGCGCAAGGTTTTGGACTGGACGTTT
 4881 ----- 4960
 CAATGGCTTGCCCTACCCGCTAGACAACCGAATGTTTTCTGTACCGGGTATTCCGCGTTCCAAAACCTGACCTGCAAA

a L P N G M G Q I C W L T K D M A H K A Q G F G L D V F -
 b Y R T G W G R S V G L Q K T W P I R R K V L D W T F -
 c V T E R D G A D L L A Y K R H G P * G A R F W T G R F -

TCGCGGGCAGCCAACGTATCTCTGCCTTATTAATATGTCTCTGCTTCGGCATATGGGAGAGTTTATTGGTGTTCAGGAA
 4961 ----- 5040
 AGCGCCCGTCGGTTGCATAGAGACGGAATAATTTTACACAGGACGAAGCGGTATACCTCTCAATAACCACAAGTCTTT

a A G S Q R I S A L L K C V L L R H M G E F I G V Q E -
 b S R A A N V S L P Y * N V S C F G I W E S L L V F R K -
 c R G Q P T Y L C L I K M C P A S A Y G R V Y W C S G N -

ACGCGTTATCTAATGAATGCGATGGAAAAAACTACTCTGAGCTGGTGAAGAGCTTCAGCGCCAGTTACCCATTAATAA
 5041 ----- 5120
 TGCGCAATAGATTACTTACGCTACCTTTTTTTGATGAGACTCGACCATTCTCGAAGTCGCGGTCAATGGGTAAATTATT

a T R Y L M N A M E K N Y S E L V K E L Q R Q L P I N F -
 b R V I * * M R W K K T T L S W * K S F S A S Y F L I F -
 c A L S N E C D G K K L L * A G E R A S A P V T H * * -

AATCGCTGAAACTTTGCAACGGCTTGATCAGAGCGGGTTTCTATTAGAGATTTACGCTTATTTTCGGCACCTTAATTG
 5121 ----- 5200
 TTAGCGACTTTGAAACGTTGCCGAACATAGTCTCGCCCAAAGATAATCTCTAAATGCAGAATAAAAGCCGTGGAATTAAC

a I A E T L Q R L V S E R V S I R D L R L I F G T L I D -
 b S L K L C N G L Y Q S G F L L E I Y V L F S A P * L -
 c N K * N F A T A C I R A G F Y * P F T S * F E H L N * -

Фиг. 11 (продолжение)

5201 ACTGGGCGCCACGTGAAAAAGATGTCCTGATGTTGACAGAATATGTCGGTATCGCGCTTCGTGCGTCATATTCTGCGTCGT
 5280 TGACCCGCGGTGCACCTTTTCTACAGGACTACAACTGCTTTATACAGGCATAGCGCGAAGCAGCAGTATAAGACCGCAGCA
 a W A P R E K D V L H I T E Y V R I A L R R H I L K R -
 b T G R H V K K M S * C * Q N M S V S R F V V I F C V V -
 c L G A T * K R C P D V D R I C P Y R A S S S Y S A S S -
 CTTAATCCGGAAGGAAAACCGCTGCCGATTTTGCGGATCGGCGAAGGTATTGAAAACCTCGTGGCTGAATCCATTGCGCA
 5281 GAATTAGGCCTTCCTTTTGGCGACGGCTAAACCGCTAGCCGCTTCCATAACTTTTGAGCAGCGCACTTAGGTAGCGGT
 5360
 a L N P E G K P L P I L R I G E G I E N L V R E S I R Q -
 b L I R K E N R C R F C G S A K V L K T S C V N P F A R -
 c * S G R K T A A D F A D R R H Y * K P R A * I H S P -
 GACGGCAATGGGACCTATACTGCGCTGTCGTCTCGTCATAAGACGCAGATCCTGCAACTTATCGAGCAGGCGCTGAAGC
 5361 CTGCGCTTACCCCTGGATATGACGCGACAGCAGAGCAGTATTCTGCGTCTAGGACGTTGAATAGCTCGTCCGCGACTTCG
 5440
 a T A H G T Y T A L S S R H K T Q I L Q L I E Q A L K Q -
 b R Q W G P I L R C R L V I P R R S C N L S S R R * S -
 c D G N G D L Y C A V V S S * D A D P A T Y R A G A E A -
 AGTCAGCCAAATTATTGCTGCTGCTGTCGACACCCGACGTTTCTTGGGAAAAATTACAGAAGCCACCTTGTTCGAC
 5441 TCAGTCGGTTTAATAAGTAACAGTGAAGACAGCTGTGGGCTGCAAGAACGCTTTTAAATGCTTCGGTGGAACAGCTG
 5520
 a S A H L F I V T S V D T R R F L R K I T E A T L F D -
 b S Q P H Y S L S L L S T P D V S C E K L Q K P P C S T -
 c V S Q I I H C H F C R H P T F L A K N Y R S H L V R R -
 GTACCGATTTTGTGTCGCGCAGGAATTAGGAGAGGAGAGCCTTATACAAGTGGTAGAAAGTATTGACCTTAGCGAAGAGGA
 5521 CATGGCTAAACAGTACCGCTCCTTAATCCTCTCCTCTCGGAATATGTTACCATCTTTCATAACTGGAATCGCTTCTCT
 5600
 a V P I L S W Q E L G E E S L I Q V V E S I D L S E E E -
 b Y R F C H G R N * E R R A L Y K W * K V L T L A K P S -
 c T D F V H A G I R R G E P Y T S G R K Y * P * R R G -
 GTTGGCGGACAATGAAGAATGAATTGATGCAACGTCTGAGGCTGAAATATCCGCCCCCGATGGTTATTGTGATGGGGG
 5601 CAACCGCCTGTTACTTCTTACTTAACTACGTTGCAGACTCCGACTTTATAGCGGGGGGCTACCAATAACAGCTACCCCG
 5680
 KOHEL *lcrD** HAHAO *yscN*?
 a L A D N F E * I D A T S E A E I S A P R W L L S M G P -
 b W R T H K N E L * Q R L R L K Y P P P D G Y C R W G -
 c V G G Q * R M N * C N V * G * N I R P P M V I V D G A -
 CGAATTCAGGATGTCAGCGCAACGTTGTTAAATGCGTGGTTGCCTGGGGTATTTATGGGCGAGTTGTGCTGTATAAGCT
 5681 GCTTAAGTCTACAGTCGCGTTGCAACAATTTACGCACCAACGGACCCCATAAATACCGCTCAACACGACATATTTCCG
 5760
 a N S G C Q R N V V K C V V A W G I Y G R V V L Y R A -
 b R I Q D V S A T L L N A W L P G V F M G E L C C I K F -
 c E F R M S A Q R C * M R G C L G Y L W A S C A V * S L -
 TGGAGAAGAACTTGCTGAAGTCGTGGGGATTAATGGCAGCAAAGCTTTGCTATCTCCTTTACGAGTACAATCGGGCTTC
 5761 ACCTCTTCTGAACGACTTCAGCACCCCTAATTACCGTCGTTTCGAAACGATAGAGGAAAATGCTCATGTTAGCCCGAAG
 5840
 yscN*
 a W R R T C * S R G D * W Q Q S F A I S F Y E Y N R A S -
 b G E E L A E V V G I N G S K A L I S P F T S T I G L H -
 c E K N L L K S W G L N A A K L C Y L L L R V Q S G F -
 ACTGCGGGCAGCAAGTGATGGCCTTAAGCGACGCCATCAGGTTCCCGTGGGCGAAGCGTTATTAGGGCGAGTTATTGATG
 5841 TGACGCCCGTCGTTCACTACCGGAATTCGCTGCGGTAGTCCAAGGGCACCGCTTCGCAATAATCCCGCTCAATAACTAC
 5920
 a L R A A S D G L K R R H Q V P V G E A L L G R V I D G -
 b C G Q Q V N A L S D A I R F P W A K R Y * G E L L M -
 c T A G S K * W P * A T T S G S R G R S V I K A S Y * W -

ФИГ. 11 (продолжение)

GCTTTGGTCGTCCTTGTATGGCCGCGAACTGCCCGACGCTCTGCTGGAAAGACTATGATGCAATGCCTCCTCCCGCAATG
 5921 ----- 6000
 CGAAACCAGCAGGGGAACTACCGGCGCTTGACGGGCTGCAGACGACCTTTCTGATACTACGTTACGGAGGAGGGCGTTAC

a F G R P L D G R E L P D V C W K D Y D A M P P P A M
 b A L V V P L M A A N C P T S A G K T M M Q C L L P Q W
 c L W S S P * W P R T A R R L L E R L * C N A S S R N G

GTTCGACAGCCTATCACTCAACCATTAAATGACGGGGATTGCGGCTATTGATAGCGTTGCGACCTGTGGCGAAGGGCAACG
 6001 ----- 6080
 CAAGCTGTCGGATAGTGAGTTGGTAATTACTGCCCTAAGCGCGATAACTATCGCAACGCTGGACACCGCTTCCCGTTG

a V R Q P I T Q P L M T G I R A I D S V A T C G E G Q R
 b F D S L S L N H * * R G F A L L I A L R P V A K G H E
 c S T A Y H S T I N D G D S R Y * * R C D L W R R A T

AGTGGGTATTTTTCTGCTCCTGGCGTGGGGAAAAAGCACGCTTCTGGCGATGCTGTGTAATGCCCGACGCGACAGCA
 6081 ----- 6160
 TCACCCATAAAAAAGACGAGGACCGCACCCCTTTTCGTGCGAAGACCGCTACGACACATTACGCGGTCTGCGTCTGCTG

a V G I F S A P G V G K S T L L A M L C N A P D A D S H
 b W V F F L L L A W G K A R F W R C C V M R Q T Q T A
 c S G Y F F C S W R G E K H A S G D A V * C A R R R R Q Q

ATGTTCTGGTGTAAATTGGTGAACGTGGACGAGAAGTCCGCGAATTCATCGATTTTACACTGTCTGAAGAGACCCGAAA
 6161 ----- 6240
 TACAAGACCAATTAACCACTTGACCTGCTCTTCAGGCGCTTAAGTAGCTAAAATGTGACAGACTTCTCTGGGCTTTT

a V L V L I G E R G R E V R E F I D F T L S E E T R
 b M F W C * L V N V D E K S A N S S I L H C L K R P E N
 c C S G V N W * T W T R S P R I H R F Y T V * R D P K T

CGTTGTGTGCTATTGTTGTGCGAACCTCTGACAGACCCGCTTAGAGCGCGTGAGGGCGCTGTTTGTGGCCACCACGATAGC
 6241 ----- 6320
 GCAACACAGTAACAACAGCGTTGGAGACTGTCTGGGCGGAATCTCGCGCACTCCCGGACAAACACCGGTGGTGTATCG

a R C V I V V A T S D R P A L E R V R A L F V A T T I A
 b V V S L L S Q P L T D P P * S A * G R C L W P P R * Q
 c L C H C C R N L * Q T R L R A R E G A V C G H H D S

AGAATTTTTGCGGATAATGGAAAGCGAGTCGTCTTGCTTGCCGACTCACTGACGCGTTATGCCAGGGCCGCACGGAAT
 6321 ----- 6400
 TCTTAAAAAGCGCTATTACCTTTGCTCAGCAGAACGAACGGCTGAGTGACTGCGCAATACGGTCCCGCGCTGCCTTTA

a E F F E D N G K R V V L L A D S L T R Y A R A A R K S
 b N F F L I M E S E S S C L P T H * R V M F G P H G N
 c R I F S S * W K A S R L A C R L T D A L C Q G R T E I

CGCTCTGGCGCGGAGAGACCGCGGTTTCTGGAGAATATCGCCAGGCGTATTAGTGCAATGCCACGACTTTTAGAAGCT
 6401 ----- 6480
 GCGAGACCGCGGCTCTCTGGGCCCAAAGACCTCTTATAGCGGTCCGCATAAATCACGTAACGGTGCTGAAAATCTTGCA

a L W R R R D R G F W R I S P G V F S A L P R L L E R
 b R S G A G E T A V S G E Y R Q A Y L V H C H D F * N V
 c A L A F E R P R F L E N I A R R I * C I A T T F R T Y

ACGGGAATGGGAGAAAAAGGCAGTATTACCGCATTTTATACGGTACTGGTGAAGGCGATGATGTAATGAAGCCGTTGG
 6481 ----- 6560
 TGCCCTTACCTCTTTTTCCGTCATAATGGCGTAAATATGCCATGACCACCTTCCGCTACTATACTTACTTCGGCAACC

yscN*

a T G M G E E G S I T A F Y T V L V E G D D M N E A V G
 b R E W E K K A V L P H F I R Y W W K A M I * M K P L A
 c G N G R K R Q Y Y R I L Y G T G G R R * Y E * S R W

CGGATGAAGTCCGTTCACTGCTTGATGGACATATTGTAATATCCCGACGGCTTGACAGAGAGGGGCATTATCCTGCCATT
 6561 ----- 6640
 GCCTACTTCAGGCAAGTGACGAACCTGTATAACATGATAGGGCTGCCGAACGCTCTCTCCCGTAATAGGACGGTAA

a G * S F F T A * W T Y C T I P T A C R E G A L S C H
 b D E V R S L L D G H I V L S R R L A E R G H Y P A
 c R M K S V H C L M D I L Y Y P D G L O R G G I I L P L

Фиг. 11 (продолжение)

```

6641  GACGTGTTGGCAACGCTCAGCCGCGTTTTTCCAGTCGTTACCAGCCATGAGCATCGTCAACTGGCGGCGATATTGGCAGG
-----
CTGCACAACCGTTGCGAGTCGGCGCAAAAAGGTCAGCAATGGTCGGTACTCGTAGCAGTTGACCGCCGCTATACGCTGC
6720
a   R V G N A Q P R F S S R Y Q P * A S S T G G D I A T -
b   D V L A T L S R V F P V V T S H E H R Q L A A I L R E -
c   T C W Q R S A A F F Q S L P A M S I V N W R R Y C D G -

6721  GTGCCTGGCGCTTTACCAGGAGGTTGAACTGTTAATACGCATTGGGGAATACCAGCGAGGAGTTGATACAGATACTGACA
-----
CACGGACCGCGAAATGGTCTCCAACCTTGACAATTATGCGTAACCCCTTATGGTCGCTCCTCAACTATGTCTATGACTGT
6800
a   V P G A L P G G * T V N T H W G I P A R S * Y R Y * Q -
b   C L A L Y Q E V E L L I R I G E Y Q R G V D T D T D K -
c   A W R F T R R L N C * Y A L G N T S E E L I Q ! L T -

6801  AAGCCATTGATACCTATCCGGATATTTGCACATTTTTGCGACAAAGTAAGGATGAAGTATGCGGACCGGAGCTACTTATA
-----
TTCCGTAACATATGGATAGGCCTATAAACGTGTAAAAACGCTGTTTCATTCTACTTCATACGCCCTGGGCTCGATGAATAT
6880
a   S H * Y L S G Y L H I F A T K * G * S M R T R A T Y R -
b   A I D T Y P D I C T F L R Q S K D E V C G P E L L I -
c   K P L I P I R I F A H F C D K V R M K Y A D P S Y L * -

6881  GAAAAATTACACCAATACTCACCGAGTGATCATGGAACCTTTGCTGGAGATAATCGCGCGGCTGAAAAGCAATTACGCG
-----
CTTTTTAATGTGGTTTATGAGTGGCTCACTAGTACCTTTGAAACGACCTCTATTAGCGCGCGGACTTTTCGTTAATGCGG
6960
a   K I T P N T H R V I M E T L L E I I A R L K S N Y A -
b   E K L H Q I I T E * S W K L C W R * S R G * K A I T R -
c   K N Y T K Y S P S D H G N E A G D N R A A E K Q L R G -

6961  GCAAGCTTACCGTACTTGATCAGCAGCAACAGGCGATTATTACGGAACAGCAGATTGCGCAGACGCGCGCTTTAGCAGTG
-----
CGTTCGAATGGCATGAACCTAGTCGTCGTTGTCGCTAATAATGCCTTGTCTGCTAAACGGTCTGCGCGCGAAATCGTCAC
7040
a   A S L P Y L I S S N R R L L R N S R F A R R A L * Q C -
b   Q A Y R T * S A A T G D Y Y G T A D L P D A R F S S V -
c   K L T V L D Q Q Q A Q I I T E Q Q I C Q T R A L A V -

7041  TCTACCAGACTGAAAGAATTAAATGGGCTGGCAAGGTACGTTATCTTGTCACTTTATTGTTGGATAAGAAACAACAATGGC
-----
AGATGGTCTGACTTTCTTAATTACCGACCGTTCCATGCAATAGAACAGTAATAACAACCTATTCTTTGTTGTTTACCG
7120
a   L P D * K N * W A G K V R Y L V I Y C W I R N N K W F -
b   Y Q T E R I N G L A R Y V I L S F I V G * E T T N G -
c   S T R L K E L M G W Q G T L S C H L L L D K K Q Q M A -

7121  CGGGTTATTCACTCAGGCGCAGAGCTTTTTGACGCAACGCGCAAGCAGTTAGAGAATCAGTATCAGCAGCTTGTCTCCCGG
-----
GCCCCAATAAGTGAGTCCGCGTCTCGAAAACTGCGTTGCGGTCGTCATCTCTTAGTCATAGTCGTCGAACAGAGGGGCG
7200
a   G Y S L R R R A F * R N G K Q L E N Q Y Q Q L V S R -
b   R V I H S G A E L F D A T A S S * R I S I S S L S P G -
c   G L F T Q A Q S F L T Q R Q A V R E S V S A A C L P -

7201  CGAAGCGAATTACAGAAGAATTTAATGCGCTTATGAAAAAGAAAGAAAAATTACTATGGTATTAAGCGATGCGTATTA
-----
GCTTCGCTTAATGTCTTCTTAAATACGCGAATACTTTTTCTTTCTTTTAAATGATACCATAATTCGCTACGCATAAT
7280
a   R S E L O K N F N A L M K K K E K I T M V L S D A Y Y -
b   E A N Y R R I L M R L * K R K K K L L W Y * A M P I T -
c   K R I T E E E * C A Y E K E R K N Y Y G I K R C V L -

7281  CCAAAGTTGAGGGAAGTCTTGGGTTGCCATGCCAGTCTTATCAGGATGATAACGAGGCGGAGGCGGAACGTATGGACTTT
-----
GGTTTCAACTCCCTTCAGAACCCAACGGTACGGTCAGAATAGTCCTACTATTGCTCCGCTCCGCTTGCATACCTGAAA
7360
a   Q S * G K S W V A M P V L S G * * R G G G G T Y G L * -
b   K V E G S L G L P C Q S Y Q D D N E A E A E R M D F -
c   P K L R E V L G C H A S L I R M I T R R K S N V W T L -

```

Фиг. 11 (продолжение)

GAACAACTCATGCACCAGGCATTACCCATTGGTGAGAATAATCCTCCTGCAGCATTGAATAAGAACGTGGTTCACGCA
 7361 ----- 7440
 CTTGTTGAGTACGTGGTCCGTAATGGGTAACCACTCTTATTAGGAGGACGTCGTAACCTATTCTTGCACCAAAAGTGCCT
 T T H A P G I T H W * E * S S C S I E * E R G F H A -
 E Q L M H Q A L P I G E N N P P A A L N K N V V F T Q -
 N N S C T R H Y P L V R I I L L Q H * I R T W F S R N -
 ACGTTATCGTGTAGTGGCGGTATCTTGACGGTGTAGAGTGTGAAGTATGTGAATCAGGGGGGCTAATCCAGTTAAGAA
 7441 ----- 7520
 TGCAATAGCACAATCACCGCCAATAGAAGTCCACATCTCACACTTCATACACTTAGTCCCCCGATTAGGTCAATTCTT
 T L S C * W R L S * R C R V * S H * I R G A N P V K H -
 R Y R V S G G Y L D G V E C E V C E S G G L I Q L R I -
 V I V L V A V I L T V * S V K Y V N Q G G * S S * E -
 TCAATGTCCTCATCATGAAATTTACCGTTCGATGAAAGCGCTAAAGCAGTGGCTGGAGTCTCAGTTGCTGCATATGGGG
 7521 ----- 7600
 AGTTACAGGGAGTAGTACTTTAAATGGCAAGCTACTTTCGCGATTTCGTACCGACCTCAGAGTCAACGACGTATACCCC
 Q C P S S * N L P F D E S A K A V A G V S V A A Y G V -
 N V P H H E I Y R S H K A L K Q W L E S Q L L H M G -
 S M S L I M K F T V R * K R * S S G W S L S C C I W G -
 TATATAATTTCCCTGGAGATATTCTATGTTAAGAATAGCGAATGAAGAGCGTCCGTGGGTGGAGATACTTCCAACGCAAG
 7601 ----- 7680
 ATATATTAAAGGGACCTCTATAAGATACAATTCTTATCGTTACTTCTCGCAGGCACCCACCTCTATGAAGGTTSCGTTCT
 конец *yscP** начало *yscQ**?
 Y N F F G D I L C * E * R M K S V R G W R Y F O R K -
 Y I I S L E I F Y V K N S E * R A S V G G D T S N A R -
 I * F P W R Y S M L R I A N E E R P W V E I L P T Q G -
 GCGCTACCATTTGGTGAGCTGACATTGAGTATGCAACAATATCCAGTACAGCAAGGGACATTATTTACCATAAATTATCAT
 7681 ----- 7760
 CGCGATGGTAACCACTCGACTGTAACCTCATACGTTGTTATAGGTCATGTCGTTCCCTGTAATAAATGGTATTTAATAGTA
 начало *yscQ**?
 A L F L V S * H * V C N N I Q Y S K G H Y L P * I I I -
 R Y H W * A D I E Y A T I S S T A R D I I Y H K L S * -
 A T I G E L T L S M Q Q Y P V Q Q G T L F T I N Y H -
 AATGAGCTGGGTAGGGTGTGGATTGCAGAACAAATGCTGGCAGCGCTGGTGTGAAGGGCTAATTGGCACCGCTAATCGATC
 7761 ----- 7840
 TTACTCGACCCATCCCACACCTAACGTCCTTGTACGACCGCTCGCGACCACACTTCCCGATTAAACCGTGGCGATTAGCTAG
 M S W V G C G L Q N N A G S A G V K G * L A P L I D R -
 * A G * G V D C R T M L A A L V * R A N W H R * S I -
 N E L G R V W I A E Q C W Q R W C E G L I G T A N E S -
 GGCTATCGATCCTGAATTGCTATATGGAATAGCTGAATGGGGGCTGGCGCGTTATTGCAAGCCAGTGATGCAACCCCTCT
 7841 ----- 7920
 CCGATAGCTAGGACTTAACGATATACCTTATCGACTTACCCCCGACCGCGCAATAACGTTGCGTCACTACGTTGGGAGAT
 L S I L N C Y H E * L N G G W R Y C K P V M Q P S -
 G Y R S * I A I W N S * M G A G A V I A S Q * C N P L -
 A I D P E L L Y G I A E W G L A F L L Q A S D A T L T -
 GTCAGAACGAGCCGCCAACATCCTGCAGTAATCTACCACATCAGCTAGCGTTGCATATTAATGGACAGTTGAAGAGCAT
 7921 ----- 8000
 CAGTCTTGCTCGGCGGTTGTAGGACGTCATTAGATGGTGTAGTCGATCGCAACGTATAATTTACCTGTCAACTTCTCGTA
 V R T S R Q H P A V I Y H I S * R C L L N G Q L K S M -
 S E R A A N I L Q * S T T S A S V A Y * M D S * R A -
 Q N E P P T S C S N L P H Q L A L H I K W T V E E H -
 GAGTTCCATAGCATTATTTTACATGGCCAACGGGTTTTTTCGCGCAATATAGTCGGAGAGCTTTCTGCTGAGCGACAACA
 8001 ----- 8080
 CTCAAGGTATCGTAATAAAATGTACCGGTTGCCCAAAAACCGGTTATATCAGCCTCTCGAAGACGACTCGCTGTTCT
 S S I A L F L H G Q P V F C A I * S E S F L L S D N R -
 V P * H Y F Y M A N G F F A Q Y S R R A F C * A T T -
 E F H S I I F T W P T G F L R N I V G E L S A E R Q Q -

Фиг. 11 (продолжение)

8081 GATTTATCCTGCCCTCCTGTGGTAGTCCCTGTATATTCAGGCTGGTGCCAGCTTACATTAATCGAACTTGAGTCTATCG
 CTAAATAGGACGGGGAGGACACCATCAGGGACATATAAGTCCGACCAGGTGCAATGTAATTAGCTTGAACCTCAGATAGC 8160

a F I L P L L W * E L Y I Q A G A S L H * S N L S I S -
 b D L S C P S C G S P C I F R L V P A Y I N R T * V Y R -
 c I Y P A P P V V V P V Y S G W C Q L T L I E L E S I E -

8161 AAATCGGCATGGGCGTTCCGATTCAATGCTTCGGCGACATCAGACTCGGTTTTTTTCTATTCAACTACCTGGGGGAATC
 TTTAGCCGTACCGCAAGCCTAAGTAACGAAGCCGCTGTAGTCTGAGCCAAAAAACGATAAGTTGATGGACCCCTTAG 8240

a K S A W A F G F I A S A T S D S V F L L F N Y L G E S -
 b N R H G R S D S L L R R H Q T R F F C Y S T T W G N L -
 c I G M G V R I H C F G D I R L G F F A I Q L P G G I -

8241 TACGCAAGGGTGTGTGCTGACAGAGGATAACACGATGAAATTTGACGAATTAGTCCAGGATATCGAAACGCTACTTGCCTC
 ATGCGTTCCCAACGACTGTCTCCTATTGTGCTACTTTAAACTGCTTAATCAGGTCTATAGCTTTGCGATGAACGCGAG 8320

a T Q G C C * Q R I T R * N L T N * S R I S K R Y L R Q -
 b R K G V A D R G * H D E I * R I S P G Y R N A T C V -
 c Y A R V L L T E D N T M K F D E L V Q D I E T L L A S -

8321 AGGGAGCCCAATGTCAAGAGTGACGGAACGCTCTTCAGTCGAACCTTGAGCAGATACCACAACAGGTGCTCTTTGAGGTCCG
 TCCCTCGGGTTACAGTTTCTCACTGCCTTGCAAGTCAGCTTGAACCTCGTCTATGGTGTGTCCACGAGAACTCCAGC 8400

a G A Q C Q R V T E R L Q S N L S R Y H N R C S L R S -
 b R E P N V K E * R N V F S R T * A D T T T G A L * G R -
 c G S P M S K S D G T S S V E L E Q I P Q Q V L F E V G -

8401 GACGTGCGAGTCTGGAATTTGGACAATTACGACAACTTAAACCGGGGACGTTTTGCCTGTAGGTGGATGTTTTGCGCCA
 CTGCACGCTCAGACCTTTAACCTGTTAATGCTGTTGAATTTTGCCCTGCAAAACGGACATCCACCTACAAAACGCGGT 8480

a D V R V W K L D N Y D N L K R G T F C L * V D V L R Q -
 b T C E S G N W T I T T T * N G G R F A C R W H F C A R -
 c R A S L E I G Q L R Q L K T G D V L P V G G C F A P -

8481 GAGGTGACGATAAGAGTAAATGACCGTATTATTGGGCAAGGTGAGTTGCTTGCCTGTGGCAATGAATTTATGGTGCCTAT
 CTCCACTGCTATTCTCATTTACTGGCATAATAACCCGTTCCACTCAACTAACGGACACCGTTACTTAAATACCAACGCATA 8560

a R * R * E * M T V L L G K V S * L P V A M N L W C V L -
 b G D D K S K * F Y Y W A R * V D C L W Q * I Y G A Y -
 c E V T I R V N D R I : G Q G E L I A C G N E F M V R I -

8561 TACACGTTGGTATCTTTGCAAAAATACAGCGTAACCTGATAAGAAAAATAATATGCGAACAATATAATAGCGTTCCAGG
 ATGTGCAACCATAGAAACGTTTTTATGTGCGATTTGGACTATTCTTTTATTATACGCTTGTTATATTATCGCAAGGTCC 8640

конец *uscQ**

a H V G I F A K I Q R K P D K K N N M R T I * * R S R -
 b Y T L V S L Q K Y S V N L I R K I I C E Q Y N S V P G -
 c T R W Y L C K N T A * T * * E K * Y A N N I I A F Q V -

8641 TCGTGTCAAGAGATACAGTATGTCTTTACCCGATTGCGCTTTGCAACTGATTGGTATATTGTTTCTGCTTCAACTACT
 AGCACAGTACTCTCTATGTCATACAGAAATGGGCTAAGCGGAAACGTTGACTAACCATATAACAAAGACGAAAGTTATGA 8720

начало *uscR*?

a S C H E R Y S M S L P D S F L Q L I G I L F L L S I L -
 b R V M R D T V C L Y P I R L C N * L V Y C F C F Q Y C -
 c V S * E I Q Y V F T R F A F A T D W Y I V S A F N T -

8721 GCCTCTCATTATCGTCATGGGAACCTCTTTCTTAAACTGGCGTGGTATTTTCGATTTACGAATGCTCTGGGTATTC
 CGGAGAGTAATAGCAGTACCCTTGAAGAAAGGAATTTGACCGCCACCATAAAAGCTAAAATGCTTTACGAGACCCATAAG 8800

a P L I : V M G T S F L F L A V V F S I I R N A L G I Q -
 b L S L S S W E L L S L N W R W Y F R F I E M L W V F -
 c A S H Y R H G N F F P * T G G G I F D F T E C S G Y S -

Фиг. 11 (продолжение)

```

AAGAAAGTCCCGCCAAATATCGCACTGTATGGCCTTGCCTTGTACTTTCCTTATTCATTATGGGGCCGACGCTATTAGCT
8801 -----+----- 8800
TTGTTTCAGGGGGGTTTATAGCGTGACATACCGGAACGCGAACATGAAAGGAATAAGTAATACCCCGCTGCGATAATCGA

a   Q V P P H I A L Y G L A L V L S L F I M G P T L L A -
b   N K S P Q I S H C M A L R L Y F P Y S L W G R R Y * L -
c   T S P P K Y R T V W P C A C T F L I H Y G A D A I S C -

GTAAAGAGCGCTGGCATCCGGTTCAGGTTCGCTGGCGCTCCTTCTGGACGTCTGAGTGGGACAGTAAAGCATTAGCGCC
8881 -----+----- 8960
CATTTCCTCGCGACCGTAGGCCAAGTCCAGCGACCGCGAGGAAAGACCTGCAGACTCACCTGTCAATTCGTAATCGCGG

a   V K E R W H P V Q V A G A P F W T S E W D S K A L A P -
b   * K S A G I R F R S L A L L S G R L S G T V K H * R L -
c   K R A L A S G S G R W R S F L D V * V G Q * S I S A -

TTATCGACAGTTTTTGCAAAAAACTCTGAAGAGAAGGAAGCCAATTATTTTCGGAATTTGATAAAACGAACCTGGCCTG
8961 -----+----- 9040
AATAGCTGTGCAAAACGTTTTTTTGAGACTTCTCTTCCTTCGGTTAATAAAAGCCTTAAACTATTTTGCTTGGACCGGAC

a   Y R Q F L Q K N S E E K E A N Y F R N L I K R T W E E -
b   I D S F C K K T L K R R K P I I F G I * * N E P G L -
c   L S T V F A K K L * R E G S Q L F S E F D K T N L A * -

AAGACATAAAAGAAAGATAAAACCTGATTCTTTGCTCATATTAATTCGGGCATTTACGGTGAGTCAGTTAACGCGAGGCA
9041 -----+----- 9120
TTCTGTATTTTTTTTTCTATTTTGGACTAAGAAACGAGTATAATTAAGGCCGTAAATGCCACTCAGTCAATTGCGTCCGT

a   D I K R K I K P D S L L I L I P A F T V S Q L T Q A -
b   K T * K E R * N L I L C S Y * F R H L R * V S * R R H -
c   R H K K K D K T * F F A H I N S G I Y G E S V N A G I -

TTTCGGATTGGATTACTTATTTATCTTCCCTTTCTGGCTATTGACCTGCTTATTTCAAATATACTGCTGGCTATGGGGAT
9121 -----+----- 9200
AAAGCCTAACCTAATGAATAAATAGAAGGGAAGACCGATAACTGGACGAATAAAGTTTATATGACGACCGATACCCCTA

a   F R I G L L I Y L P F L A I D L L I S N I L L A M G M -
b   F G L E Y L F I F P F H L L T C L F Q I Y C W L W G * -
c   S D W I T Y L S S L S G Y * P A Y F K Y T A G Y G D -

GATGATGGTGTGCGCGATGACCATTTTACCTACCGTTTAAGCTGCTAATATTTTACTGGCAGGCGGTTGGGATCTGACAC
9201 -----+----- 9280
CTACTACCACAGCGGCTACTGGTAAAGTAATGGCAATTTCGACGATTATAAAAATGACCGTCCGCCAACCTTAGACTGTG

a   M M V S P M T I S L P F K L L I F L L A G G W D L T L -
b   * W C R R * P F H Y R L S C * Y F Y W Q A V G I * H -
c   D D G V A D D H F I T V * A A N I F T G R R L G S C T -

TGGCGCAATTGGTACAGAGCTTTTCATGAATGATTCTGAATTGACGCAATTGTAAACGCACTTTTATGGATCGTCCTTT
9281 -----+----- 9360
ACCGCGTTAACCATGTCTCGAAAAGTACTTACTAAGACTTAACTGCGTTAAACATTGCGTTGAATAACCTAGCAGGAAA

      конец yscR*      начало yscS*
a   A Q L V Q S F S * M I L N * R N L * R N F Y G S S F -
b   W R N W Y R A F H E * F * I D A I C N A T F M D R F F -
c   G A I G T E L F M N D S E L T O F V T Q L L W I V L F -

TTACGCTATGCGGCTAGTGTGGTGGCATCGGTAGTTGGTGTATCGTAAGCCTTGTTTCAAGCCTTGACTCAAATACAG
9361 -----+----- 9440
AATGCAGATACGGCCATCACAACCCGCTAGCCATCAACCACAGTAGCATTGGAACAAGTCCGGAACCTGAGTTTATGTC

a   L R L C R * C W W H R * L V S S * A L F R P * L K Y R -
b   Y V Y A G S V G G I G S W C H R K P C S G L D S N T G -
c   T S M P V V L V A S V V G V I V S L V Q A L T Q I Q -

GACCAACGCTACAGTTCATGATTAAATTATTGGCAATTGCAATAACCTTAATGGTCAGCTACCCATGGCTTAGCGGTAT
9441 -----+----- 9520
CTGGTTTGGCATGTCAAGTACTAATTTAATAACCGTTAACGTTATTGGAATTACCAGTCGATGGGTACCGAATCGCCATA

a   T K R Y S S * L N Y W Q L Q * P * W S A T H G L A V S -
b   P N A T V H D * I I G N C N N L N G Q L P M A * R Y -
c   D Q T L Q F M I K L L A I A I T L M V S Y P W L S G I -

```

Фиг. 11 (продолжение)

CCTGTTGAATTATACCCGGCAGATAATGTTACGAATTGGAGAGCATGGTTGAATGGCACAAACAGGTAAATGAGTGGCTTA
 9521 ----- 9600
 GGACAACCTTAATATGGGCGTCTATTACAATGCTTAACCTCTCGTACCACTTACCGTGTGTCCATTACTCACCGAAT

КОНЕЦ *yscS** начало *yscT**
 C * I I P G R * C Y E L E S M V E W H N R * M S G L -
 a P V E L Y P A D N V T N W R A W L N G T T G K * V A Y -
 b L L N Y T R Q I M L R I G E H G * M A O O V N E W L I -
 c

TTGCATTGGCTGTGGCTTTTATTGACCATTTGAGCCTTTCTTTATTACTTCCCTTATTAAGTGGCAGTTTATGGGGCC
 9601 ----- 9680
 AACGTAACCGACACCGAAAATAAGCTGGTAACCTCGGAAAGAAATAATGAAGGGAATAATTTTACCCTGCAATCCCGG

L H W L W L L F D H * A F L Y Y F P Y * K V A V * G P -
 a C I G C G F Y S T I E P F F I T S L I K K W Q F R G R -
 b A L A V A F I R P L S L S L L L P L L K S G S L G A -
 c

GCACTTTTACGTAATGGCGTGCTTATGTCACTTACCTTTCCGATATTACCAATCATTTACCAGCAGAAGATTATGATGCA
 9681 ----- 9760
 CGTGAAATGCATTACCGCAGCAATACAGTGAATGGAAAGGCTATAATGGTTAGTAAATGGTCGTCTTCTAATACTACGT

H F Y V M A C L C H L P F R Y Y Q S E T S R R L * C I -
 a T F T * W R A Y V T Y L S D I T N H L P A E D Y D A -
 b A L L R N G V L M S L T F P I L P I I Y Q Q K I M M H -
 c

ИНСЕРЦИЯ P9B7
 U

TATTGGTAAAGATTACAGTTGGTTAGGGTTAGTCACTGGAGAGGTGATTATTGGTTTTTCAATTGGGTTTTGTGCGGCGG
 9761 ----- 9840
 ATAACCATTTCTAATGTCAACCAATCCCAATCAGTGACCTCTCCACTAATAACCAAAAGTTAACCCTAAACACGCCGCG

L V K I T V G * G * S L E R * L L V F Q L G F V R R -
 a Y W * R L Q L V R V S H W R G D Y W F F N W V L C G G -
 b I G K D Y S W L G L V T G E V I I G F S I G F C A A V -
 c

TTCCCTTTTGGGCGGTTGATATGGCGGGGTTCTGCTTGATACTTTACGTGGCGGACAAATGGGTACGATATTCAATTCT
 9841 ----- 9920
 AAGGGAAAACCGGCAACTATACCGCCCCAAAGACGAAGTATGAAATGCACCGCGCTGTACCCATGCTATAAGTTAAGA

F P F G P L I W R G F C L I L Y V A R Q W V R Y S I L -
 a S L L G R * Y G G V S A * Y F T W R D N G Y D I Q F Y -
 b P F W A V D M A G F L L D T L R G A T M G T I F N S -
 c

ACAATAGAAGCTGAAACCTCACTTTTGGCTTGCTTTTCAGCCAGTTCTTGTGTGTATTTCTTTATAAGCGGCGGCAT
 9921 ----- 10000
 TGTTATCTTCGACTTTGGAGTGAAAAACCGAAGCAAAAGTCGGTCAAGAACACACAATAAAAGAAATATTCGCCGCCGTA

Q * K L K P H F L A C F S A S S C V L F S L * A A A W -
 a N R S * N L T F W L A F Q P V L V C Y F L Y K R R H -
 b T I E A E T S L F G L L F S Q F L C V I F F I S G G M -
 c

GGAGTTTATATTAAACATTCTGTATGAGTCATATCAATATTTACCACCAGGGCGTACTTTATTATTGACCAGCAATTTT
 10001 ----- 10080
 CCTCAAATATAATTTGTAAGACATACTCAGTATAGTTATAAATGGTGGTCCCGCATGAAATATAAACTGGTCGTAAAA

S L Y * T F C M S H I N I Y H Q G V L Y Y L T S N F -
 a G V Y I K H S V * V I S I F T T R A Y F I I * P A I F -
 b E F I L N I L Y E S Y Q Y L P P G R T L L F D O O F L -
 c

TAAATATATCCAGGCAGAGTGGAGAACGCTTTATCAATTATGTATCAGCTTCTCTCTCTGCCATAATATGTATGGTA
 10081 ----- 10160
 ATTTTATATAGGTCGGTCTCACCTCTTGGGAAATAGTTAATACATAGTCGAAGAGAGAAGGACGGTATTATACATACCAT

N I S R Q S G E R F I N Y V S A S L F L P * Y V W Y -
 a K I Y P G R V E N A L S I M Y Q L L S S C H N M Y G I -
 b K Y I Q A E W R T L Y Q L C I S F S L P A I I C M V -
 c

TTAGCCGATCTGGCTTTAGGTCTTTTAAATCGGTGGGCACAACAATTGAATGTGTTTTTCTTCTCAATGCCGCTCAAAAG
 10161 ----- 10240
 AATCGGCTAGACCGAAATCCAGAAAATTTAGCCAGCCGTGTTGTTAACTTACACAAAAGAGAGTTACGGCGAGTTTTTC

F I W L * V F * I G R H N N * M C F S S Q C R S K V -
 a S R H G F R S F K S V G T T I E C V F L L N A A Q K -
 b L A D L A L G L L N R S A Q Q L N V F F F S M P L N S -
 c

Фиг. 11 (продолжение)

10241 TATATTGGTTCTACTGACGYCCTGATCTCATTCCCTTATGCTCTTCATCACTATTTGGTTGAAAGCGATAAATTTTATAT
 ATATAACCAAGATGACTGCRGGACTAGAGTAAGGGAATACGAGAAGTAGTGATAAACCAACTTTCGCTATTTAAATATA 10320
 a Y W F Y * R P D L I P L C S S S L F G * K R * I L Y -
 b Y I G S T D ? L I S F P Y A L H H Y L V E S D K F Y I -
 c I L V L L T ? * S H S L M L F I T I W L K A I N F I F -
 10321 TTATCTAAAAGACTGGTTTCCATCTGTATGAGCGAGAAAAACAGAACAGCCTACAGAAAAGAAATTACGTGATGGCCGTAA
 AATAGATTTTCTGACCAAAGGTAGACATACTCGCTCTTTTGTCTTGTGGATGTCTTTTCTTAAATGCACTACCGGCATT 10400
 конец *yscT** начало *yscU**
 a L S K R L V S I C M S E K T E Q P T E K K L R D G R K -
 b Y L K D W F P S V * A R K Q N S L Q K R N Y V M A V R -
 c I * K T G F H L Y E R E N R T A Y R K E I T * W P * -
 10401 GGAAGGGCAGGTTGTCAAAGTATTGAAATAACATCATTATTTTCAGCTGATTGCGCTTTATTTGTATTTTCATTTCTTTA
 CCTTCCCGTCCAACAGTTTTCATACTTTATTGTAGTAATAAAGTCGACTAACGCGAAATAAACATAAAAGTAAGAAAT 10480
 a E G Q V V K S I E I T S L F Q L I A L Y L Y F H F E T -
 b K G R L S K V L K * H H Y F S * L R F I C I F I S L -
 c G R A G C Q K Y * N N I I I S A D C A L F V F S F L Y -
 10481 CTGAAAAGATGATTTTGATACTGATTGAGTCAATACTTTTCACATTACAATTAGTAAATAAACCATTTTCTTATGCATTA
 GACTTTTCTACTAAAACATGACTAACTCAGTTATGAAAGTGTAATGTTAATCATTATTTTGGTAAAAGATACGTAAT 10560
 a E K M I L I L I E S I T F T L Q L V N K P F S Y A L -
 b L K R * F * Y * L S Q * L S H Y N * * I N H F L M H * -
 c * K D D F D T D * V N N F H I T I S K * T I F L C I N -
 10561 ACGCAATTGAGTCATGCTTTAATAGAGTCACTGACTTCTGCACTGCTGTTTCTGGGCGCTGGGGTAATAGTTGCTACTGT
 TGGCTTAACTCAGTACGAAATTATCTCAGTGACTGAAGACGTGACGACAAAGACCCGCGACCCCATATCAACGATGACA 10640
 a T Q L S H A L I E S L T S A L L F L G A G V I V A I V -
 b R N * V M L * * S H * L L H C C F W A L G * * L L L W -
 c A I E S C F N R V T D F C T A V S G R W G N S C Y C -
 10641 GGGTAGCGTGTTTCTTCAGGTGGGGTGTTTATTGCCAGCAAGGCCATTGGTTTTAAAAGCGAGCATATAAATCGGGTAA
 CCCATCGCACAAAGAAGTCCACCCCAACCAATAACGGTCGTTCCGGTAACCAAAATTTTCGCTCGTATATTAGGACATT 10720
 a G S V F L Q V G V V I A S K A I G F K S E H I N F S -
 b V A C F F R W G W L L P A R P L V L K A S I * I S * -
 c G * R V S S G G G G Y C Q Q G H W F * K R A Y K S S K -
 10721 GTAATTTAAGCAGATATTCTCTTTACATAGCGTAGTAGAATTATGTAATCCAGCCTAAAAGTTATCATGCTATCTCTT
 CATTAAATTCGTCTATAAGAGAAATGTATCGCATCATCTTAATACATTTAGGTGCGATTTTCAATAGTACGATAGAGAA 10800
 a N F K Q I F S L H S V V E L C K S S L K V I M L S L -
 b V I L S R Y S L Y I A * * N Y V N P A * K L S C S L L -
 c * F * A D I L F T * R S R I M * I Q P K S Y H A I S Y -
 10801 ATCTTTGCCTTTTCTTTTATTATTATGCCAGTACTTTTGGGCGCTACCGTACTGTGGGTTAGCCTGTGGCGTCTTGT
 TAGAAACGGAAGAAAGAAATATAATACGGTCATGAAAAGCCCGCATGGCATGACACCCAATCGGACACCGCACGAACA 10880
 a I F A F F F Y Y Y A S T F R A L P Y C G L A C G V L V -
 b S L P F S F I I M P V L F G R Y R T V G * P V A C L W -
 c L C L F L L L L C Q Y F S G A T V L W V S L W R A C -
 10881 GGTTTCTTCTTAATAAAATGGTTATGGGTAGGGTGATGGTTTTTATATCGTGGTGGCATACTGGACTATTCTTTTC
 CCAAAGAAGAAATATTTTACCAATACCCATCCCACTACCAAAAAATATAGCAGCAACCGTATGACCTGATAAGAAAAG 10960
 a V S S L I K W L W V G V M V F Y I V V G I L D Y S F Q -
 b F L L * * N G Y G * G * W F F I S S L A Y W T I L F -
 c G F F F H K M V M G R G D G F L Y R R W H T G L F F S -

Фиг. 11 (продолжение)

AATATTATAAGATTAGAAAAGCTATCTAAAAATGAGTAAAGATGACGTAAACAGGAGCATAAAGATCTGGAGGGCGACC
 10961 ----- 11040
 TTATAATATTCTAATCTTTTCGATAGATTTTACTCATTTCTACTGCATTTTGTCTCGTATTTCTAGACCTCCCGCTGG
 a Y Y K I R E A I * K * V K M T * N R S I K I W R A T -
 b N I I R L E F L S K N E * R * R K T G A * R S G G R P -
 c I L * D * K S Y L K M S K D D V K Q E H K D L E G D P -
 инсерция P12F5
 CTCAAATGAAGACGCGCGCTCGGAAATGCAGAGTGAAATACAAAGTGGGAGTTTAGCTCAATCTGTTAAACAATCTGTTG
 11041 ----- 11120
 GAGTTTACTTCTGCGCCGCGAGCCTTTACGTCTCACTTTATGTTTACCCTCAAATCGAGTTAGACAATTTGTTAGACAAC
 a L K * R R G V G N A E * N T K W E F S S I C * T I C C -
 b S N E D A A S E M Q S E I Q S G S L A Q S V K Q S V A -
 c Q M K T R R R K C R V K Y K V G V * L N L L N N L L -
 CGGTAGTGCGTAATCCAACGCATATTGCGGTTTGTCTTGGCTATCATCCCACCGATATGCCAATACCACGCGTCTGGAA
 11121 ----- 11200
 GCCATCAGCATTAGGTTGCGTATAACGCCAAACAGAACCGATAGTAGGGTGGCTATACGGTTATGGTGGCAGGACCTT
 a G S A * S N A Y C G L S W L S S H R Y A N T T R P G K -
 b V V R N P T H I A V C L G Y H P T D M P I P R V L E -
 c R * C V I Q R I L R F V L A I I P P I C Q Y H A S W K -
 AAAGGCGAGTGCTCAAGCTAACTATATTGTTAATCATCGCTGAACGCAACTGCATCCCCGTTGTTGAAATGTTGAGCT
 11201 ----- 11290
 TTTCGCTCACTACGAGTTGATTGATATAACAATTGTAGCGACTTGGCTTGACGTAGGGGCAACAACCTTTTACAACCTCGA
 a R Q * C S S * L Y C * H R * T Q L H P R C * K C * A -
 b K G S D A Q A N Y I V N I A E R N C I P V V E N V E L -
 c K A V M L K L T I L L T S L N A T A S P L L K M L S W -
 GGCCCGCTCATTATTTTTTGAAGTGGAAACGCGGAGATAAAATTCCTGAAACGTTATTTGAACCCGTTGCGAGCCTTGTAC
 11281 ----- 11360
 CCGGGCGAGTAATAAAAACTTCACCTTGCCTCTATTTTAAGGACTTTGCAATAAACTTGGGCAACGTCGGAACAATG
 a G P L I I F * S G T R R * N S * N V I * T R C S L V T -
 b A R S L F F E V E R G D K I P E T L F E P V A A L L R -
 c P A H Y F L K W N A E I K F L K R Y L N P L Q P C Y -
 GTATGGTGATGAAGATAGATTATGCGCATTCTACCGAAACACCATAAATGCTTTTGGTATGCTTCTCAGGCCACTGCGA
 11361 ----- 11440
 CATACCACTACTTCTATCTAATACGCGTAAGATGGCTTTGTGGTATTTACGAAAACCATACGAAGAAGTCGGTGACGCT
 КОНЕЦ *yscU**
 Y G D E D R L C A F Y R N T I N A F G M L L Q A T A K -
 a M V M K I D Y A H S T E T P * M L L V C F F R P L R -
 b V W * * R * I M R I L F K H H K C F W Y A S S G H C E -
 c V W * * R * I M R I L F K H H K C F W Y A S S G H C E -
 AGGTTAAGAGGGTAATAGCGTATAGAGCAGTGCTTGACGATAAAGGTGAGAGACTGAAAATAATCGCTTTTAGCCTGGCA
 11441 ----- 11520
 TCCAATTCTCCCATATCGCATATCTCGTCACGAAGTGTATTTCCACTCTCTGACTTTTATTAGCGAAAATCGGACCGT
 a V K R V I A Y R A V L D D K G E R L K I I A F S L A -
 b R L R G * * R I E Q C L T I K V R D * K * S L L A W H -
 c G * E G N S V * S S A * R * R * E T E N N R F * P G T -
 CAAGCACCAGATAGCGTATTATAAAATTAACAAGATAATGGATTGGTGGCTCTGAATGGACTCGAACCACTCGACCCCC
 11521 ----- 11600
 GTTCGTGGTCTATCGCATAATATTTAATTTGTTCTATTACCTAACCGCAGACTTACCTGAGCTTGGTgAGCTGGGGG
 a Q A P D S V L * N * T R * W I G A S E W T R T T R P P -
 b K H Q I A Y Y K I K Q D N G L V R L N G L E P L D P H -
 c S T R * R I I K L N K I M D W C V * M D S N H S T P -
 ACCATGTCAAGGTGGTGTCTAACCAACTGAGCTATGAACGGCAACGTTGTAGGTGACAACGGGGACGAATATTAGCGTC
 11601 ----- 11680
 TGGTACAGTTCCACCACGAGATTGGTTGACTCGATACTTGCCGTTGCAACATCCACTGTTGCCCTGCTTATAATCGCAG
 a P C Q G G A L T N * A M N G N V V G D N G D E Y * R H -
 b H V E V V L * P T E L * T A T L * V T T G T N I S V * -
 c T M S R W C S N Q L S Y E R Q R C F * Q R G R I L A S -

Фиг. 11 (продолжение)

```

ACAACCGCAATGAGGCAAGAGGGAAATCGCAATTTCTTCTGAAATCACCTGATTGCGGTGGAAATATGCAACATGTCC
11691 ----- 11760
TGTTGGCGTTACTCCGTTCTCCCTTTAGCGTTAAAGAAGGACTTTAGTGACTAACGCCACCTTTATACGTTGTACAGC
.
a N R N E A R G K S Q F S S * N H L I A V E I C N H S -
b T T A M R Q E G N R N F L P E I T * L R W K Y A T C R -
c Q P Q * G K R E I A I F F L K S P D C G G N M Q H V E -

AGAAAATAGCCGCCATGCGACGGCTATCGTCGTATTATCGGAGCGCGCTGCAAAATGATGGCGGACGGCTGACGTTGTAG
11761 ----- 11840
TCTTTTATCGGCGGTACGCTGCCGATAGCAGCATAATAGCCTCGCGGACGTTTACTACCGCCTGCCGACTGCAACATC

a R K * P P C D G Y R R I I G A R C K M M A D G * R C R -
b E N S R H A T A I V V L S E R A A K * W R T A D V V D -
c K I A A M R R L S S Y Y R S A L Q N D G G R L T L * -

ATAGCGCATCCGTAGCATCATTAAACACCGCGCGGAGGTGAGCGCGATGATGAACCCCATCCAGAAGCCTGCCGGTCCCA
11841 ----- 11920
TATCGCGTAGGCATCGTAGTAATTGTGGCGGCGGCTCCAGTCCGGCTACTACTTGGGGTAGGTCTTCGGACGGCCAGGGT

a * R I R S I I N T A A E V R P H M N P I Q K P A G P I -
b S A S * A S L T P P P R S G R * * T P S R S L P V P -
c I A H F * H H * H R R R G Q A D D E P H P E A C R S H -
11921 ----- 12000
TACGATCCACCCACAAATCCGTTAACGCCAGGATATAACCGCTGGGTAAACCTAACCCAGTAGGCGGTAAAGGTGATA
ATGCTAGGTGGTGGTTAGGCAATTGCGGTCCCTATATTGGCGACCCATTTGGATTGTGGGTATCCGCCATTTCCACTAT

a R S T T K S V N A R I * P L G K P N T Q * A V K V I -
b D P P P N P L T P G Y N R W V N L T P S R R * R * -
c T I H H Q I R * R Q D I T A G * T * H P V G G K G D K -

AAAAAGATGGAAACGCGTATCTTTTATAACCGCGCAGAATACCGCTGCCGATAACCTGTATAGAGTCGGAAATCTGGTAAAC
12001 ----- 12080
TTTTTCTACCTTGCGCATAGAAATATTGGCGCGTCTTATGGCGACGGCTATTGGACATATCTCAGCCTTTAGACCATTTC

a K K M E R V S L * P R R I P L P I T C I E S E I W * T -
b K R W N A Y L Y N R A E Y R C R * P V * S R K S G K P -
c K D G T R I F I T A Q N T A A D N L Y R V G N L V N -

CGCAGCGAGCAGCATTAAATTGCGGCAAGCGCCACGACCTCAGGGTTGTCTATTGTAGAGCAAAGCAATATGCTTACGCAGA
12081 ----- 12160
GCGTCGCTCGTCTAATTAAACGCGTTCGCGGTGCTGGAGTCCCAACAGTAACATCTCGTTTCTGTTATACGAATGCGTCT

a A A S S I N C G K R H D L R V V I V E Q S N M L T Q S -
b Q R A L I A A S A T T S G L S L * S K A I C L R -
c R S E Q H * L R Q A P R P Q G C H C R A K Q Y A E -
12161 ----- 12240
GTAACGGTAAATAGCGGTAAACACAGCCATACAAATGCCGACGCTAAACCGGTACGCGCTGCGTTTGGCGATCCAGC
CATTGCCATTTTATCGCCATTGGTGTGCGGTATGTTTACGGCTGCGGATTGGCCATGCGGACGCAACGCGTAGGTCC

a N G K N S G N H S H T N A D A * T G T R C V C A S S -
b V T V E I A V T T A I Q M P T P K P V R A A F A H F A -
c * R * K * R * P Q P Y K C R R L N R Y A L R L R I Q R -

GTTGAGCCCTGGCCGACCGGATAACCCACTCGAATCGTTACCGCGCGAGCCAGCGACATCGGCAGTACGAACATCAGCG
12241 ----- 12320
CAACTCGGGACCGGCTCTGGCTATTGGGTGAGCTTAGCAATGGCGGCGTCCGTCGCTGTAGCCGTCATGCTTGTAGTCGC

a V E P W P R P I T H S N R Y R R S Q R H R O Y E H O R -
b L S F G P D R * P T R I V T A A A S D I G S T N I S E -
c * A I A Q T D N P L E S L P P Q P A T S A V R T S A -
12321 ----- 12400
AGCTAAAGTTAAGCGCAATCTGATGACCGGCGACATCCACAATACCTAATGGCGAACCAGCAGCGCAACGACCGCAAT
TCGATTTCAATTCCGCTTAGACTACTGGCGCTGTAGGTGTTATGGATTACCGCTTTGGTCGTCGCTTGGTGGCGTTTA

a A K V K R N L M T G D I H N T * W R N Q Q R N D R K * -
b L K L S A I * * P A T S T I P N G E T S S A T T A N -
c S * S * Q S D D R R H P Q Y I M A K P A A Q R P Q I -
12401 ----- 12480
AACGTCACCTTCAAGAACAGCGCAATCGGCAACCCAGTTGAATCAGGCGCTTCATGACGACGCTATCGGGTTTGC
TTGCAGTGAAGTTCTTGTCGGTCCGCTTAGCCGTGGGGTCAACTTAGTCCGGAAGTACTGCTGCGATAGCCCAACG

a R H F K E Q P A Q S A T P V E S G A S * R R Y R V C -
b N V T S * N S Q R N R Q P Q L N Q A L H D D A I G F A -
c T S L Q R T A S A I G N P S * I R R F M T T L S G L P -

```

Фиг. 11 (продолжение)

12481 CAAAGCCTTTTTCATTACGAATATCAGCATTGAACGCGCGTGTAAATGTAAGAAAGCATGGCGATAACATCACCCAA
 ----- 12560
 GTTTCGGAAAAAGTAATGCTTATAGTGCCTAACTTGCAGCACAAATTACATTCTTTCGTACCGCTATTGTAGTGGGTT
 .
 a Q S L F H Y E Y H A L N A R V . C K K A W R . T S P N -
 b K A F I T N I T H . T R V F N V R K H G D K H H P I -
 c K P F S L R I S R I E R A C L M . E S M A I N I T Q -
 TAGACCGCGCAGTCGCAACGCGCGAGCGGATACCGCGAGTTCCGGCATACCAAAATGGCCATAGATAAAAAATAGTT
 ----- 12640
 12561 ATCTGGCGCGTCAGCGTTGCGGCGTCGGCTATGGCGGCTCAAGCGCGTATGGTTTACCGGTATCTATTTTATATCAA
 .
 a R P P Q S O R R S R Y R R V P A Y Q N G H R . K Y S S -
 b D R R S R N A A A D T A E F R H T K M A I D K N I V -
 c . T A A V A T P Q P I P P S S G I P K W P . I K I . F -
 CACCGGAATATTCACCAGCAGGCCCAAAATCCCATCACCATACCGGTTTGGTTTGGCCAGACCTTCGCACTGGTTTC
 ----- 12720
 12641 GTGGCCTTATAAGTGGTCGTCCGGGTTTTAGGGTAGTGGTATGGGCCAAACCAAAACCGGTCTGGAAGCGTGACCAAG
 .
 a P E Y S P A G P K I P S P Y P V W F W P D L R T G F . -
 b H R N I H Q Q A Q K S H H H T R F G F G O T F A L V S -
 c T G ! F T S R P K N P I T I P G L V L A R P S H W F R -
 GCGCTACCTGAAAGAAAAGGTATCCTGCGCCCCACAGCAGCGCGGAAGATAACCCACGGCTTTATCGGCCAGCGCCGA
 ----- 12800
 12721 CGCGATGGACTTTCTTTCCATAGGACGCGGGGTGTCGTGCGCGCTTCTATTGGGTGCCGAAATAGCCGGTCGCGGCCT
 .
 a A L P E R K G I L R P T A A R E D N P R L Y R P A P D -
 b R Y L K E K V S C A P Q Q R A K I T H G F I G Q R R I -
 c A T . K K R Y P A P H S S A R R . P T A L S A S A G -
 TCAATATTATGCATAGAGCGGATAATGTATCCGGCATTCCACAGGACGATCATCACCAGCAGGAGACAAAGCCCGCCAG
 ----- 12880
 12801 AGTTATAATACGTATCTCGCTATTACATAGGCCGTAAGGTGTCCTGCTAGTAGTGGTCGTGCCTCTGTTTCGGGCGGTC
 .
 a Q Y I A . S G . C I R H S T G R S S P A R R Q S P P A -
 b N I M H R A D N V S G I P Q D D H H Q H G D K A R Q -
 c S I L C I E R I M Y P A F H R T I I T S T E T K P A S -
 CCAGAACCTTGTGGAACCTGATGCGCGATACGCTCAGCAGCGCGGAGCCATTGAGTTGCGCAATCACAGGCGTCAAGG
 ----- 12960
 12881 GGTCTTGGGAACAGCTTGGACTACGCGCTATGCGAGTGCTGCCGGCCTCGGTAACCAACGCGTTAGTGTCCGAGTTCC
 .
 a R T L V E P D A R Y A H D G R S H . V A O S O A S . -
 b P E F L S N L M R D T L T T A G A I E L R N H R R Q G -
 c Q N F C R T . C A I R S R R R P E P L S C A I T G V K A -
 CCAGCAGTAAGCCGTGACCAAAACAAATGGCGGAAGCAGATAGAGGTGCCGATAGCGCAGGCAGCCATGTCCGTAGCGC
 ----- 13040
 12961 GGTGCTCATTCGGCACTGGTTTGTGTTTACCGCCCTTCGTCTATCTCCACGGCTATCGCTGCCGTCGGTACAGGCATCGCG
 .
 a P A V S R D Q T K W R E A D R G A D S D G S H V R S A -
 b Q Q . A V T K Q N G G K Q I E V P I A T A A M S V A L -
 c S S H P . P N K M A G S R . R C R . R R Q P C P . R -
 TATAGCTCCCGCATGACGGTATCGACGAATCCATTGCGGTCTATACCACTTGCAGCAAGGATACCGGTATCTGAACGC
 ----- 13120
 13041 ATATCGGAGGGCGGTACTGCCATAGCTGCTTAGGTAAGCCAGATATGGTGAACGCGTTCTAGTGGCCATAGACTTGCG
 .
 a I A S R H D G I D E S I A V Y T T C A R ! T G I . T L -
 b . P P A M T V S T N P L R S I P L A Q G S P V S E R -
 c Y S L P P . R Y R R I H C G L Y H L R K D H R Y L N A -
 TAATAACTGACGCGCTTCACTGGTATACTTCTGCACGTATTCACCTTTTATTTTGTGTTATATGAAAGACTAAAAAGCC
 ----- 13200
 13121 ATTATTGACTGCGCGAAGTGACCATATGAAGACGTGCATAAGTGAAAAATAAAACAACATATACTTTCTGATTTTTCGG
 .
 a I T D A L H W Y T S A R I H L L F C C Y M K D . K A -
 b . L T R F T G I L L H V F T F Y F V V I . K T K K P -
 c N N . R A S L V Y F C T Y S P F I L L L Y E R L K S R -
 GCCGAAGTGGCAGCCAAAGAAATAGCAGGGGAAATTTCACTCTATTGTAGCGGGGTATTACTATTTCTCCAGTGAAAAA
 ----- 13280
 13201 CGGCTTCACCGTCGGTTTTCTTTATCGTCCCTTTAAAGTCAGATAACATCGCCCCATAATGATAAAGAGGTCACCTTTT
 .
 a A E V A A A R N S H G N I S L L . R G I T I S P V K K -
 b P K W Q P K E I A G E I S V Y C S G V L L F L Q . K N -
 c R S G S Q K K . Q G K F Q S I V A G Y Y Y F S S F K -

Фиг. 11 (продолжение)

Страница: 136

ДНК-ПОСЛЕДОВАТЕЛЬНОСТЬ VGC II КЛАСТЕРА С

ИНСЕРЦИЯ P9B4

GGATCCTTTTCTTTAATGCTGCTAACGTTTCTTGCAAAATGCGTTGATGAGATTCACTCCAGTACACCACTGATAACAAA
 CCTAGGAAAAAGAAATTACGACGATTGCAAGAACGTTTACGCACTACTCTAAGTAGGTCATGTGGTGACTATTGTTT

ИНСЕРЦИЯ P7A3

AGAGCGNCGCATTGGCENNAHMTKNNHNNNSCNHNACTAAACCGTTCTCTATTATCGCAGAAATAATATCATCCCCCTG
 TCTCGNCGGTAACCGNNTKKHMYNNKYNHSGHNTGATTGGCAAGAGATAATAGCGTCTTTATTATAGTAGGGGAC
 AGACTGATGAGAGTGAATACTGCCAGTGCAATAACCCGGGAATATCTGCAAGTAATGGTTGAACCTTGCGCCATTGCT
 TCTGACTACTCTCACTGATTAGACGGTCACGTTATTGGGCCCTTATAGACGTTCACTACCAACTTGGAAACCGGTAACGA

ИНСЕРЦИЯ P964

GATCCATTTGTATATCATCATGAATTAACACGCTCCCGGCCCTTCGCTGGTACTTCAGCATNSGGTAACCCATTTT
 CTAGGTAAACATATAGTAGTACTTAATTGTGCGAGGGGCGGGAAGCGACCTATGAAGTCGTANSGCCATTGGGTAAAA
 ATCAAAACATCCTGCACTTCTCGTACCAATAAGTCATCACAGATTACACCATCCGATACATGACCCCCCATGATCGAG
 TAGTTTTGTAGGACGTGAAGAGCATGGTTATTCACTAGTGTCTAATGTGGTAGGGCTATGTAAGTGGGGGTACTAAGCTC
 AGTCGCTCTCACCTTTGCACTCTGTTGCTTGACGAGCAATAACCGGACAAGTCAGGCTGCCATCTCTTTCCATTGCG
 TCAGCGAGAGTGGAAAACGTAGACAAGCGAAGCTGCTGTTATTGGCTGTTGACGTCGACGGTAGAAGAAAGGTAAACGC
 CCGGCACATAATGAATATTGCTTTTGTCTAATAAAAACTTAACCCGCAAGGTAAGTCATTTACCGTTTCAGGCTGACCA
 GGGCGTGTATTACTTATAACGAAACAGATTATTTTGAATTGGCGTTCCATTCACTAAATGGCAAAGTCCGACTGGT
 CTAATACTTAACAGGACACCCATTCCACCGATGAATAACGAATACGCCAGCCAAACACCACTACCTGATCTGGAAC
 GATTATGAATGTCTGTGGGTAAAGGTGGCTACTTTTAGTTCTTATGCGGTGCGTTGGTGGTCACTGGGACTAGACCTTTG
 GGGTATTTGATAATCAGCAAGTTCACAATCCTGTTTACCAACCGCATASSCACTCCCGCAACCTGCAAAACCCCACTGG
 CCCATAAACTATTAGTCGTTCAAGTGTTAGGACAAATGGTTTGGCTATSSGTGAGGGCGTTGGACGTTTGGGGTGACC
 ATGGTAGCGGCTTATTGGATTAAATCTGCGGCCATTAACTCTAACTCTGGCTTTCCCGGCATCAACAAATAAATATCT
 TACCATCGCGAATAAACCTAATTTAGACGCCGTAATTGAGATTGAGACCGAAAGGGCGTACTGTTTATTGATAGA
 GCCTGTTCTCTCAGAATAATTTTTCATTATAGCCAGCGAATACAAATATCGCATCCCTCTCCCCAGTGACAGGTTA
 CGGACAAGAGACTCTTATTAATAAAGTAAATATCGGTGCGTTATGTTTATAGCGTAGGGAAGAGGGGTCACTGTCCAAT
 CCTTCATTGAGCCATCTTCCCGGCTTGTAAACGTTGACCTAAAAACGTTATTTCCAGGAAGCTCTTGGATTAAACAT
 GGAAGTAAGTCGGTATGAAGGGCCGGAACATTTTGCAGTGGATTTTTCATAAAAGGTCCTTGAGAAACCTAATTGGTA
 GAGATATGCCATTATTTACTACTGAGGCTTAAATCAAAAAAGCCTGATTACACTATGTACTTGAGTCGTATCATGGGA
 CTCATACCGTAATAAATGATGACTCCGAAATTAGTTTTTTCGGACTAATGTGATACATGAACCTCAGCATAGTAACGCT
 AACAAATGACCTACACAGGAATATCGCCCAATAAAGGGATTTGTTTTGCGAGTGGATTGTTTACCTTGTTTTAAACCC
 TTGTTTACTGGATGTTGTCTTATAGCGGTTATTTCCCTAAAAACAAACGCTCACCTAAACAAATGGAACAAATTTGGG
 TCCCAGCAATHAGACTTTGCCCGGCCAATAATGTGGCTTGGGAANCRAATTCAGAAATTTGCACTTCGGGCAGCGGGTCT
 AGGGTCGTTAATCTGAAACGGGCGGTTATTACACCGAACGCTTNGYTAAGTCTTAAACGTTGAAGCCCGTCGCCACGA
 GTNTYGCYTTKGNSTATCACTTTGTTGTCCATCCTGAANTATTAAATTAAGCATTATTTTTCGCTGCCATTGTCAATTT
 CANACGAAACNSTATAGTGAAACAAACAGGTAGGACTTAAATTTCTAATTCGTAAATAAAAAACGCACGGTAACAGTAAA
 AACAAAGCGAGGTGTAAACGCGWNAACAAAGAACCCGTAGTGATGGATTCAAGTTTACCCACTTTTTCTCCCTGCAGTTTGG
 TTGTTCCCTCCACATGCGCMTTGTCTTGGGCATCACTACCTAAGTTCAAAATCGGTGAAAAAGAGGGACGTCAAAAC

ПОСЛЕДОВАТЕЛЬНОСТЬ 2

SEQ ID №38

ФИГ. 12

[illegible]

Фиг. 12 (продолжение)

NYYKSSSCYSWKATMYYSWRWWTTAATGGAATGCCTTTTAAACTGCCAGCATGAATCCCTCCTCAGACATAAATGGGAG 2800
 2801 -----+-----+-----+-----+-----+-----+-----+-----+
 NRRMSSSGRSWMTAKRRSWYWWAATTACCTTACGGAATAATTTGACGGTCGTACTTAGGGAGGAGTCTGTATTTACCCCTC
 TTTCTATCAAATTCGCTCACAACCACATCCGTAAGAGCCTGATTCACATTATTTTCGACTATACTCTTCTGTACAATA 2960
 2881 -----+-----+-----+-----+-----+-----+-----+-----+
 AAAGATAGTTTAAGCGAGTGTGGTGTAGGCATTTTCGGACTAAGTGTAATAAAGCTGATATGAGAAGAACATGTTAT
 TCAGGATGCTGTCTACATATACCTTGTACAGGCGATTCTATCATTCGGATTTTCGGATAAATTNMMCAATTACATTTTC 3040
 2961 -----+-----+-----+-----+-----+-----+-----+-----+
 AGTCCTACGACAGATGTATATGGAACAGTGTCCGCTAAGATAGTAAGCCTAAAAGGCTATTTAANKKGTTAATGTAAAAG
 AGCATTGACATAAAAACTTACAATTTGNAAAAATTATTTATTAATAAACTGTTACGATGTTTTTACATCGCCATCTTATT 3120
 3041 -----+-----+-----+-----+-----+-----+-----+-----+
 TCGTAACTGTATTTTGAATGTTAAACNTTTTAATAAATAATTTATTTGACAATGCTACAAAAATGTAGCGGTAGAATAA
 AAAAAGTAATTGTAGTCATCGACTNGGTTATATATGAAGAAATTTATCTTCTAATGATAACACCATCGATTAAATCWCT 3200
 3121 -----+-----+-----+-----+-----+-----+-----+-----+
 TTTTTCATTAAACATCAGTAGCTGANCCAATATATACTTCTTTAAATAGAAGGATTACTATTGTGGTAGCTAATTAGWWGA
 GATGAACTATATGTACTGCGATAGTGATCAAGTGCCAAAGATTTTGAACAGGCAACTGGAGGGAAGCATTATGAATTT 3280
 3201 -----+-----+-----+-----+-----+-----+-----+-----+
 CTACTTTGATATACATGACGCTATCACTAGTTCACGGTTTCTAAAACGTTGTCCGTGACCTCCCTTCGTAATACTTAAA
 SSTCAATCTCAAGAATACSSYSYRNNNNNTCTTTAGTAATCAGGCTAACTTTTTTATTTTATTAACAACAATAATTWT 3360
 3281 -----+-----+-----+-----+-----+-----+-----+-----+
 SSAGTTAGAGTTCTTATGSSRSYRNNNNNAGAAATCATTAGTCCGATTGAAAAATAAAAAATAATTGTTGTTATTAAWA
 TTGGCTGCTATCTGTGCTTACCGCAGCTTATATATCAATGGTTTCGAAACGGCAGCATATAATAGAGGATTTATCCGTTT 3440
 3361 -----+-----+-----+-----+-----+-----+-----+-----+
 AACCACGATAGACACGAATGGCGTCGAATATATAGTTACCAAGYCTTTGCCGTGCTATATTATCTCTAAATAGGCAAG
 TATCCGAGATGAATATTGTACTAAGCAATCAACGGTTTGAAGAAGCTGAACGTGACGCTAAAAATTTAATGTATCAATGC 3520
 3441 -----+-----+-----+-----+-----+-----+-----+-----+
 ATAGGCTCTACTTATAACATGATTCGTTAGTTGCCAAACTTCTTCGACTTGCACTGCGATTTTAAATTACATAGTTACG
 TCATTAGCGACTGAGATTCATCATAACGATATTTTCCCTGAGGTGAGCCGGCATCTATCTGTGCGTCTTCAAATTCAC 3600
 3521 -----+-----+-----+-----+-----+-----+-----+-----+
 AGTAATCGCTGACTCTAAGTAGTATTGCTATAAAAGGGACTCCACTCGGCCGTAGATAGACAGCCAGGAAGTTTAAACGTG
 MGCCGACGCTNAACGGAGAGAAGCACCGTCTCTTCTGCAGTCCTCTGATATCGATGAAAAATAGCTTTCGTCGCGATAGT 3680
 3601 -----+-----+-----+-----+-----+-----+-----+-----+
 KCGGCTGCGANTTGCCTCTCTTCGTGGCAGAGAAAGACGTGAGGAGCTATAGCTACTTTTATCGAAAGCAGCGCTATCA
 инсерция P3F4
 ↓
 TTTATTCTTAATCATAAAAATGAGATTTTCGTTATTATCTACTGATAACCCTTCAGATTATTCAACTCTACAGCCTTTAAC 3760
 3681 -----+-----+-----+-----+-----+-----+-----+-----+
 AAATAAGAATTAGTATTTTACTCTAAAGCAATAATAGATGACTATTGGGAAGTCTAATAAGTTGAGATGTCGGAAATTG
 GCGAAAAAGCTTTCCTTTATACCCAACCCATGCCGGGTTTACTGGAGTGAACAGAAATACATAAACGGCAAAGGATGGC 3840
 3761 -----+-----+-----+-----+-----+-----+-----+-----+
 CGCTTTTTCGAAAGGAAATATGGGTTGGGTACGGCCCAAAATGACCTCACTTGGTCTTATGTATTGCGGTTTCCTACCG
 AACGCTTCCGTTGCGGTTGCCGATCAGGCAAGGCGTATTTTTCGAGTGACGGTTAACTTCCCGATCTCATTACTAAGA 3920
 3841 -----+-----+-----+-----+-----+-----+-----+-----+
 TTGCGAAGGCAACGCCAACGGCTAGTCCGTTCCGCATAAAAACTCCACTGCCAATTGGAAGGGCTAGAGTAATGATTCT
 GCCACCTGCCATTAGATGATAGTATTTCGAGTATGGCTGGATCAAAACAACCACTTATTGCCGTTTTCATACATCCCGGCA 4000
 3921 -----+-----+-----+-----+-----+-----+-----+-----+
 CGGTGGACGGTAATCTACTATCATAAGCTCATACCGACCTAGTTTTGTTGGTGAATAACGGCAAAAGTATGTAGGGCCGT
 AAAAATACGTACACAGTTAGAAAATGTAAACGCTGCATGATGGATGGCAGCAAAATCCCGGATTTCTGATATTACGCACAA 4080
 4001 -----+-----+-----+-----+-----+-----+-----+-----+
 TTTTATGCATGTGTCAATCTTTACATTGCGACGTACTACCTACCGTCGTTTAAAGGCTTAAAGACTATAATGCGTGTT
 CCTTGCATGGCCCCGGATGGAGTCTGGTTACGCTGTACCCATACGGTAATCTACATAATCGCATCTTAAAAATTATCCTT 4160
 4081 -----+-----+-----+-----+-----+-----+-----+-----+
 GGAACGTACCGGGGCCCTACCTCAGACCAATGCGACATGGGTATGCCATTAGATGTATTAGCGTAGAATTTTAAATAGGAA
 CAACAAATCCCTTTACATTAAACAGCATGGGTGTTGATGACGTGGCTTTTGTGGTTACTACATCGCTCACTGGCCAA 4240
 4161 -----+-----+-----+-----+-----+-----+-----+-----+
 GTTGTTTAGGGGAAATGTAATTGTCGTAAACCAACTACTGCAGCCGAAAAACGACCAATGATGTAGCGAGTGACCGGTT

Фиг. 12 (продолжение)

ACCGTTATGGCGTTTTGTCGATGTCATTAATAAAACCGCAACTGCACCGCTGAGCACACGTTTACCAGCACAACTGG
 4241 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 4320
 TGGCAATACCGCAAAACAGCTACAGTAATTATTTTGGCGTTGACGTGGCGACTCGTGTGCAATGGTCGTGTGCTGACC
 ATGAATTAGATAGTATTGCCGGTGCTTTTAACCAACTGCTTGATACTCTACAAGTCCAATACGACAATCTGGAAAAACAA
 4321 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 4400
 TACTTAATCTATCATAACGGCCACGAAAATTGGTTGACGAACTATGAGATGTTTCAAGTTATGCTGTTAGACCTTTGT
 GTCGACAGCGACCCAGGCGCTAAATGAAGCAAAAAACGCGCTGAGCNAGCTAACAAACGTAAAAGCATTCATCTTACG
 4401 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 4480
 CAGCGTCTGCGTGGGTCCGCGATTTACTTCGTTTTTTGCGCGACTCGNTCGATTGTTTGCATTTTCGTAAGTAGAATGC
 GTAATAAGTCATGAGTTACGTACTCCGATGAATGGCGTACTCGGTGCAATTGAATTATTACAAACACCCCTTTAAACAT
 4481 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 4560
 CATTATTCACTACTCAATGCATGAGGCTACTTACCGCATGAGCCACGTTAACTTAATAATGTTTGGTGGGGAATTTGTA
 AGAGCAACAAGGATTAGCTGATACCGCCAGAAATGTACACTGTCTTTGTTAGCTATTATTAATAATCTGCTGGATTTT
 4561 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 4640
 TCTCGTTGTTCCATAATCGACTATGGCGGTCTTTAACATGTGACAGAAACAATCGATAATAATTATTAGACGACCTAAAA
 CACGCATCGAGTCTGGTCATTTACATTACATATGGAAGAAACAGCGTTACTGCCGTTACTGGACAGGCAATGCAAAAC
 4641 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 4720
 GTGCGTAGCTCAGACCAGTAAAGTGTAATGTATACCTTCTTTGTCGCAATGACGGCAATGACCTGGTCCGTTACGTTTGG
 ATCCAGGGGCCAGCGCNAAGCAAAAACTGTCAATACGTACTTTTGTGCGGTCAACATGTCCCTCTCTATTTTCATACCG
 4721 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 4800
 TAGGTCCCCGGTCGCGNTTTCGTTTTTTGACAGTAATGCATGAAACAGCCAGTTGTACAGGGAGAGATAAAAGTATGGC
 ACAGTATCCGTTTACNNCAAATTTTGGTTAATTTACTCGGGAACGCGGTAAAATTTACCGAAACCGGAGGATACGTCTGA
 4801 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 4880
 TGT CATAGGCAATGNNGTTTAAAACCAATTAATGAGCCCTTGCGCCATTTTAAATGGCTTTGGCCTCCTATGCAGACT
 CGGTCAAGCGTCATGAGGAACAATTAATATTTCTGGTTAGCGATAGCGGTAAAGGGATTGAAATACAGCAGCAGTCTCAA
 4881 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 4960
 GCCAGTTCGCGACTCTCCTTGTAAATTATAAGACCAATCGCTATCGCCATTTCCCTAACTTTATGTCGTCGTCAGAGTT
 ATCTTTACTGCTTTTTATCAAGCAGACACAAATTCGCAAGGTACAGGAATTGGACTGACTATTGCGTCAAGCCTGGCTAA
 4961 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 5040
 TAGAAATGACGAAAAATAGTTCGTCTGTGTTTAAAGCGTCCATGTCCTTAACCTGACTGATAACGCAGTTCGGACCGATT
 AATGATGGGCGGTAATCTGACACTAAAAAGTGTCGCCGGGGTTGGAACCTGTGTCTCGCTAGTATTACCTTTACAAGAAT
 5041 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 5120
 TTAATACCCGCCATTAGACTGTGATTTTTCACAGGGGGCCCAACCTTGGACACAGAGCGATCATAATGGGAATGTTCTTA
 инсерция
 ↓
 ACCAGCCGCTCAACCAATTAAGGGACGCTGTGAGNNNCCGTTCTGCCTGCATCGGCAACTGGCTTGCTGGGGAATACG
 5121 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 5200
 TGGTCGGCGGAGTTGGTTAATTTCCCTGCGACAGTCNNNGGCAAGACGGACGTAGCCGTTGACCGAAGACCCCTTATGC
 CGGTGAACCACCCACAGCAAAATGCGCTTCTCAANNNAGAGCTTTTGTATTTCTCCGGAAAACTCTACGACCTGGCG
 5201 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 5280
 GCCACTTGGTGGGGTGGTCTTTTACGCGAAGAGTTNNNGTCTCGAAAACATAAGAGGCCCTTTGAGATGCTGGACCGC
 CAACAGTTAATATTGTGTACACCAATATGCCAGTAATAAATAATTTGTTACCACCTGGCAGTTGCAGATTCCTTTGGT
 5281 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 5360
 GTTGTCAATTATAACACATGTGGTTTATACGTCATTTATTTAATAAATGGTGGGACCGTCAACGTCTAAGAAAACCA
 TGATGATGCCGATATTAATCGGGATATCATCGGCAAAATGCTTGTGAGCCTGGGCCAACACGTCACTATTGCCGCCAGTA
 5361 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 5440
 ACTACTACGGCTATAATTAGCCCTATAGTAGCCGTTTTACGAACAGTCCGACCCGTTGTGAGTGATAACGCGGTCAT
 GTAACGAGGCTCTGACTTTATCACAACAGCAGCGATTTCGATTTAGTACTGATTGACATTAGAATGCCAGAAATAGATGGT
 5441 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 5520
 CATGCTCCGAGACTGAAATAGTGTGTCGTCGCTAAGCTAAATCATGACTAACTGTAATCTTACGGTCTTTATCTACCA
 ATTGAATGTGTACGATTATGGCATGATGAGCCGAATAATTTAGATCCTGACTGCATGTTTGTGGCACTATCCGCTAGCGT
 5521 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 5600
 TAACTTACACATGCTAATACCGTACTACTCGGCTTATTAATCTAGGACTGACGTACAAACACCGTGATAGGCGATCGCA
 ASCVNMAGAWRWMTWTCRTYGTDDAAAAAARWDRKDHWTCAHAYANNTTACAAAACAGTGACATTGGCTACCTTAGC
 5601 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 5680
 TSGBNKCTWYAKWAGYARCAHNTTTTWTWYCYMHDWAGTADTRTNNAATGTTTGGTCACTGTAACCGATGGAATCG

Фиг. 12 (продолжение)

5681 TCGCTACATCAGTATTGCCGCGAGAATACCAACTTTTACGAAATATAGAGCTACAGGAGCAGGATCC
 -----+-----+-----+-----+-----+-----+-----+-----+----- 5746
 AGCGATGTAGTCATAACGGCGTCTTATGGTTGAAAAATGCTTTATATCTCGATGTCCTCGTCCTAGG

Фиг. 12 (продолжение)