



**(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ЕВРАЗИЙСКОМУ ПАТЕНТУ**

**(45)** Дата публикации и выдачи патента  
**2021.09.21**

**(21)** Номер заявки  
**201791343**

**(22)** Дата подачи заявки  
**2015.12.15**

**(51)** Int. Cl. **C12N 15/113** (2010.01)  
**C12N 9/22** (2006.01)  
**C12N 7/01** (2006.01)

**(54) КОМПОЗИЦИИ И СПОСОБЫ ДЛЯ КОНСТРУИРОВАНИЯ ВИРУСНОГО ГЕНОМА IN VITRO**

**(31)** **62/092,707; 62/102,362; 62/242,811**

**(32)** **2014.12.16; 2015.01.12; 2015.10.16**

**(33)** **US**

**(43)** **2018.03.30**

**(86)** **PCT/US2015/065891**

**(87)** **WO 2016/100389 2016.06.23**

**(71)(73)** Заявитель и патентовладелец:  
**СиЗДжей ТЕРАПЬЮТИКС, ИНК.**  
**(US)**

**(72)** Изобретатель:  
**Кэди Кайл К., Барбу Э. Магда,**  
**Дипетрилло Кристен Г. (US)**

**(74)** Представитель:  
**Медведев В.Н. (RU)**

**(56)** LIU B.L. et al. ICP34.5 Deleted Herpes Simplex Virus with Enhanced Oncolytic, Immune Stimulating, and Anti-tumour Properties. *Gene Therapy*. 2003, Vol. 10, pages 292-303; abstract; page 293, second column, second paragraph; 298, first column, second paragraph; page 301, first column, third paragraph - second column, first paragraph; Figures 1, 2, 4, DOI: 10.1038/sj.gt.3301885

MARTEL B. et al. CRISPR-Cas: An Efficient Tool for Genome Engineering of Virulent Bacteriophages. *Nucleic Acids Research*. 24 July 2014, Vol. 14, No. 42, pages 1-10; abstract. DOI: 10.1093/nar/gku628

WO-A1-1999023107

US-A1-20140079671

CEYSSSENS P.J. et al. Genomic Analysis of *Pseudomonas aeruginosa* Phages LKD16 and LKA1: Establishment of the KMV Subgroup within the T7 Supergroup. *Journal of Bacteriology*. October 2006, Vol. 188, No. 19, pages 6924-6931; DOI: 10.1128/JB.00831-06. NCBI, Supplement, page 1

WO-A1-1999023107

LAMMENS E. et al. Representational Difference Analysis (RDA) of Bacteriophage Genomes. *Journal of Microbiological Methods*. 20 February 2009, Vol. 77, pages 207-213; DOI:10.1016/j.mimet.2009.02.006. NCBI Supplement

CEYSSSENS P.J. Isolation and Characterization of Lytic Bacteriophage Infecting *Pseudomonas aeruginosa* [online]. University in Leuven. December 2009 [retrieved on 2016-04-14]. Retrieved from the Internet; <URL: <http://lirias.kuleuven.be/bitstream/123456789/250877/1/phdAERUGINOSA>>; pages 1-166; page VI, page 64, third paragraph; Tables 4.2, 4.4

US-A1-20130225451

SUENAGA T. et al. Engineering Large Viral DNA Genomes Using the CRISPR-Cas9 System. *Microbiology and Immunology*. September 2014, Vol. 50, pages 513-522; abstract; page 514, second column, second paragraph; page 515, second column, first paragraph. DOI: 10.1111/1348-0421.12180.

ENGLER C. et al. A One Pot, One Step, Precision Cloning Method with High Throughput Capability. *PLoS ONE*. 05 November 2008, Vol. 3, No. 11, pages 1-7; page 4, column 2, first paragraph. DOI: 10.1371/journal.pone.0003647.

PINGOUD A. Spermidine Increases the Accuracy of Type II Restriction Endonucleases. Suppression of Cleavage at Degenerate, Non-Symmetrical Sites. *European Journal of Biochemistry*, 1985, Vol. 147, pages 105-109; abstract; page 106, second column, first-second paragraphs

WO-A1-1999041397

KIM S. et al. Highly Efficient RNA-Guided Genome Editing in Human Cells via Delivery of Purified Cas9 Ribonucleoproteins. *Genome Research*. June 2014, Vol. 24, No. 6, pages 1012-1019; abstract. DOI: 10.1101/gr.171322.113

GLONTI T. et al. Tail Tubular Protein B [*Pseudomonas* phage PT2]. National Center for Biotechnology Information. Genbank Entry. 19 October 2007 [retrieved on 14 April 2016]; Retrieved from the Internet <URL: <http://www.ncbi.nlm.nih.gov/protein/195546741>>; pages 1-2.

GLONTI T. et al. Hypothetical Protein PT2\_gp16 [*Pseudomonas* phage PT2]. National Center for Biotechnology Information. Genbank Entry. 19 October 2007 [retrieved on 14 April 2016]; Retrieved from the Internet <URL: <http://www.ncbi.nlm.nih.gov/protein/195546717>>; page 1.

US-A1-20110104119

BI Y. et al. High-Efficiency Targeted Editing of Large Viral Genomes by RNA-Guided Nucleases. *PLoS Pathogens*. 1 May 2014, Vol. 10, No. 5; e1004090; DOI: 10.1371/journal.ppat.1004090

**(57)** Изобретение относится к способу конструирования нуклеиновых кислот в условиях *in vitro*. Изобретение дополнительно относится к *in vitro* конструированию вирусных геномов и к улучшению вирусных свойств с помощью геномной инженерии *in vitro* вирусных геномов.

В частности, изобретение относится к расщеплению вирусного генома в условиях *in vitro* с использованием РНК-направляемой Cas9, сборки рекомбинантного генома путем вставки фрагмента ДНК или РНК в расщепленный вирусный геном и трансформации клетки-хозяина рекомбинантным геномом. Данный способ также относится к конструированию *in vitro* для исправления ошибок в нуклеиновых кислотах.

038595 B1

038595 B1

---

### **Перекрестная ссылка на родственные заявки**

Настоящая заявка испрашивает приоритет в соответствии с USC 119 (e) по предварительной заявке на патент США № 62/092707, поданной 16 декабря 2014 г., предварительной заявке на патент США № 62/102362, поданной 12 января 2015 г., и предварительной патентной заявке США № 62/242811, поданной 16 октября 2015 г., полное содержание каждой из которых включено в настоящий документ посредством ссылки.

### **Перечень последовательностей**

Настоящая заявка содержит ссылки на последовательности нуклеиновой кислоты, которые одновременно представлены в виде текстового файла с перечнем последовательностей "SGI1840\_3WO\_Sequence\_Listing\_ST25.txt", с размером файла (139 кБ) в килобайтах, созданного 15 декабря 2015 г., который включен в качестве ссылки в полном объеме в соответствии со Сводом федеральных нормативных актов США, раздела 37, пункта 1.52(e) (iii) (5).

### **Область техники**

Изобретение в основном относится к быстрому конструированию геномов и, более конкретно, к конструированию вирусных геномов *in vitro*.

### **Уровень техники**

Вирусы используют во многих научных целях, особенно при разработке профилактических, терапевтических и диагностических средств. Для этих целей вирусы часто подвергают генной инженерии. Инженерия *in vitro* требует легко поддающегося обработке организма-хозяина и часто может занять от нескольких недель до нескольких месяцев для создания модифицированных вирусов и вирусных векторов (Levin and Bull, Nat. Rev. Microbiol., 2004 Feb; 2(2): 166-73, включенный в настоящий документ посредством ссылки). Кроме того, существуют проблемы токсичности, связанные с манипулированием многими вирусными геномами в клетках. Усилия по разработке методов генной инженерии *in vitro* крупных вирусных геномов до сих пор были ограничены наличием уникальных последовательностей целевых рестрикционных ферментов и низкой эффективностью, полученной для расщепления генома и последующей рекомбинантной сборки. Кроме того, многие усилия в области генной инженерии пресекаются неправильно предсказанными вирусными геномными точками окончания репликации. Например, общедоступные PB1-подобные вирусные геномы неправильно размещают концевые последовательности в середине генома, что является часто встречающейся ошибкой при использовании современных методов секвенирования и сборки генома *in silico* (Ceysens et al., Environ Microbiol. 2009 Nov; 11(11): 2874-83).

Остается потребность в быстрой генной инженерии вирусных геномов, особенно для вирусов, инфицирующих генетически неприемлемых хозяев. Изобретение использует Cas9-опосредованное расщепление *in vitro* и сборку сайтспецифических модифицированных целых вирусных геномов. Настоящий способ значительно увеличивает точность, простоту и скорость, с которой вирусные геномы могут быть генетически модифицированы. Кроме того, настоящий способ преодолевает общепризнанную сложность манипулирования часто токсичными вирулентными вирусными геномами внутри нативных и гетерологичных клеток-хозяев. Использование раскрытого способа конструирования *in vitro* также позволяет идентифицировать подходящие вирусные геномные концы, что облегчает последующее конструирование при помощи настоящего раскрытия.

Коррекция ошибок *in vitro* представляет собой бесценный метод получения желаемых последовательностей после клонирования или методов сборки. Стандартные методы коррекции ошибок, основанные на ПЦР, которая имеет две неотъемлемые проблемы: 1) ПЦР может вводить дополнительные нежелательные мутации в нуклеиновую кислоту; и 2) ПЦР в этом контексте имеет ограничение по размеру приблизительно 5 т.п.н. (тысяча пар нуклеотидов), прежде чем она становится все более подверженной ошибкам (руководство по набору Quick Change site-directed mutagenesis kit, New England Biolabs, США). Поэтому стандартные методы исправления ошибок на основе ПЦР не могут быть надежно выполнены на плазидах размером более 5 т.п.н., либо из-за дополнительных создаваемых ПЦР мутаций, либо из-за невозможности амплификации всей матрицы.

### **Сущность изобретения**

Среди различных аспектов настоящего раскрытия представлены композиции и способы для конструирования последовательностей нуклеиновых кислот *in vitro* с использованием РНК-направляемой нуклеазы. В одном аспекте раскрытие относится к улучшению специфических вирусных свойств при помощи генетической инженерии *in vitro* последовательностей вирусных нуклеиновых кислот и к улучшенным вирусным композициям или частицам. В другом аспекте раскрытие относится к расщеплению *in vitro* последовательностей вирусных нуклеиновых кислот с использованием эндонуклеазы, направляемой РНК, например Cas9, с последующей сборкой последовательности рекомбинантной нуклеиновой кислоты путем вставки фрагмента(ов) ДНК или РНК в расщепленную вирусную нуклеиновую кислоту.

В некоторых вариантах реализации настоящее изобретение относится к сконструированному вирусу, содержащему сконструированную вирусную нуклеиновую кислоту, способную при введении в клетку-хозяина продуцировать не встречающиеся в природе вирусные частицы с двумя или более улучшенными вирусными свойствами по сравнению с вирусными частицами, полученными путем введения сконструированной вирусной нуклеиновой кислоты в клетку-хозяина.

В некоторых аспектах полученные вирусные частицы имеют по меньшей мере три улучшенных вирусных свойства.

В некоторых аспектах каждое улучшенное вирусное свойство выбрано из группы, состоящей из круга хозяев, вирусного литического цикла, адсорбции, прикрепления, инъекции, репликации и сборки, лизиса, величины выхода фага, ускользания от иммунологического надзора, иммунной стимуляции, иммунной дезактивации, дисперсии биопленки, устойчивости бактерий к фагам, антибиотической сенсбилизации бактерий, модуляции вирулентности и направленного расщепления или редактирования генома хозяина.

В некоторых аспектах сконструированная вирусная нуклеиновая кислота представляет собой сконструированный вирусный геном.

В некоторых аспектах сконструированный вирусный геном представляет собой сконструированный геном бактериофага. В некоторых аспектах по меньшей мере одно из улучшенных вирусных свойств представляет собой специфичность к хозяевам.

В некоторых аспектах каждое улучшенное вирусное свойство представляет собой результат по меньшей мере одной модификации в сконструированной вирусной нуклеиновой кислоте.

В некоторых аспектах по меньшей мере одно улучшенное вирусное свойство представляет собой результат по меньшей мере двух модификаций в сконструированной вирусной нуклеиновой кислоте.

В некоторых аспектах по меньшей мере одна модификация в сконструированной вирусной нуклеиновой кислоте представляет собой результат одного этапа конструирования.

В некоторых аспектах по меньшей мере одна модификация в сконструированной вирусной нуклеиновой кислоте представляет собой результат неоднократно повторяемых этапов конструирования.

В некоторых аспектах по меньшей мере одна из модификаций находится в последовательности нуклеиновой кислоты, имеющей по меньшей мере 85% идентичности с SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 50 или SEQ ID NO: 25.

В некоторых аспектах по меньшей мере одна из модификаций находится в последовательности нуклеиновой кислоты, кодирующей аминокислотную последовательность, имеющую по меньшей мере 85% идентичности с SEQ ID NO: 34, SEQ ID NO: 35, SEQ ID NO: 36, SEQ ID NO: 5, SEQ ID NO: 48 или SEQ ID NO: 49.

В некоторых аспектах сконструированный вирусный геном содержит весь или часть вирусного генома, имеющего по меньшей мере 85% идентичности с геномом LUZ19. В некоторых аспектах сконструированный вирусный геном дополнительно содержит весь или часть гетерологичного гена gp18. В некоторых аспектах гетерологичный ген gp18 имеет по меньшей мере 85% идентичности с SEQ ID NO: 26. В некоторых аспектах гетерологичный ген gp18 кодирует аминокислотную последовательность с по меньшей мере 85% идентичностью с SEQ ID NO: 38.

В некоторых аспектах сконструированный вирусный геном содержит весь или часть вирусного генома, имеющего по меньшей мере 85% идентичности с геномом LUZ19. В некоторых аспектах сконструированный вирусный геном дополнительно содержит весь или часть сконструированного гена gp34. В некоторых аспектах сконструированный ген gp34 кодирует аминокислотную последовательность, содержащую мутацию в положении, соответствующем положению 55 аминокислоты в SEQ ID NO: 5.

В некоторых аспектах сконструированный вирусный геном содержит весь или часть вирусного генома, имеющего по меньшей мере 85% идентичности с геномом LUZ19. В некоторых аспектах сконструированный вирусный геном дополнительно содержит модификацию в одной или более последовательностях, имеющих по меньшей мере 85% идентичности с последовательностью, выбранной из группы, состоящей из SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3 и SEQ ID NO: 50. В некоторых аспектах сконструированный вирусный геном дополнительно содержит модификацию в каждой последовательности, имеющей по меньшей мере 85% идентичности SEQ ID NO: 1, последовательности, имеющей по меньшей мере 85% идентичности SEQ ID NO: 2, последовательности, имеющей по меньшей мере 85% идентичности с SEQ ID NO: 3 и последовательности, имеющей по меньшей мере 85% идентичности с SEQ ID NO: 50. В некоторых аспектах модификации включают замену G на A в положении, соответствующем положению 50 нуклеиновой кислоты SEQ ID NO: 1, замену G на T в положении, соответствующем положению 160 нуклеиновой кислоты в SEQ ID NO: 50, замену A на G в положении, соответствующем положению 245 нуклеиновой кислоты в SEQ ID NO: 2, замену AT на TC в положениях, соответствующих положениям 247-248 нуклеиновой кислоты в SEQ ID NO: 2 и замену A на G в положении, соответствующем положению 757 нуклеиновой кислоты в SEQ ID NO: 3.

В некоторых аспектах сконструированный вирусный геном содержит весь или часть вирусного генома, имеющего по меньшей мере 85% идентичности с геномом LUZ19. В некоторых аспектах сконструированный вирусный геном дополнительно включает модификацию в одной или более последовательностях нуклеиновой кислоты, кодирующих аминокислотную последовательность, имеющую по меньшей мере 85% идентичности, с последовательностью, выбранной из группы, состоящей из SEQ ID NO: 34, SEQ ID NO: 35, SEQ ID NO: 36 и SEQ ID NO: 48. В некоторых аспектах сконструированный вирусный геном включает модификацию в последовательности нуклеиновой кислоты, кодирующей каждую из аминокислотной последовательности, имеющей по меньшей мере 85% идентичности с SEQ ID NO: 34,

аминокислотной последовательности, имеющей по меньшей мере 85% идентичности с SEQ ID NO: 35, аминокислотной последовательности, имеющей по меньшей мере 85% идентичности с SEQ ID NO: 36, и аминокислотной последовательности, имеющей по меньшей мере 85% идентичности с SEQ ID NO: 48. В некоторых аспектах модификации включают замену С на У в положении, соответствующем аминокислотному положению 17 в SEQ ID NO: 34, замену D на У в положении, соответствующем аминокислотному положению 36 в SEQ ID NO: 48, замену D на G в положении, соответствующем аминокислотному положению 82 в SEQ ID NO: 35, замену I на S в положении, соответствующем аминокислотному положению 83 в SEQ ID NO: 35 и замену N на D в положении, соответствующем аминокислотному положению 253 в SEQ ID NO: 36.

В некоторых аспектах сконструированный вирусный геном содержит весь или часть вирусного генома, имеющего по меньшей мере 85% идентичности с геномом LUZ19. В некоторых аспектах сконструированный вирусный геном дополнительно содержит модификацию в пределах последовательности, имеющей по меньшей мере 85% идентичности с SEQ ID NO: 25. В некоторых аспектах модификация представляет собой введение гетерологичной молекулы нуклеиновой кислоты в последовательность, имеющую по меньшей мере 85% идентичности с SEQ ID NO: 25, или замену последовательности, содержащейся в последовательности, имеющей по меньшей мере 85% идентичности с SEQ ID NO: 25 с гетерологичной молекулой нуклеиновой кислоты. В некоторых аспектах гетерологичная молекула нуклеиновой кислоты содержит гетерологичную последовательность нуклеиновой кислоты, имеющую по меньшей мере 85% идентичности с последовательностью, выбранной из группы, состоящей из SEQ ID NO: 6, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19 и SEQ ID NO: 20.

В некоторых аспектах сконструированный вирусный геном содержит весь или часть вирусного генома, имеющего по меньшей мере 85% идентичности с геномом LUZ19. В некоторых аспектах сконструированный вирусный геном дополнительно содержит модификацию в пределах последовательности нуклеиновой кислоты, кодирующей аминокислотную последовательность, имеющую по меньшей мере 85% идентичности с SEQ ID NO: 49. В некоторых аспектах модификация представляет собой введение гетерологичной молекулы нуклеиновой кислоты в последовательность нуклеиновой кислоты, кодирующей аминокислотную последовательность, имеющую по меньшей мере 85% идентичности с SEQ ID NO: 49, или замену последовательности нуклеиновой кислоты, содержащейся в последовательности нуклеиновой кислоты, кодирующей аминокислотную последовательность, имеющую по меньшей мере 85% идентичности с SEQ ID NO: 49 с гетерологичной молекулой нуклеиновой кислоты. В некоторых аспектах гетерологичная молекула нуклеиновой кислоты содержит гетерологичную последовательность нуклеиновой кислоты, кодирующую аминокислотную последовательность, имеющую по меньшей мере 85% идентичности с последовательностью, выбранной из группы, состоящей из SEQ ID NO: 37, SEQ ID NO: 39, SEQ ID NO: 40, SEQ ID NO: 41, SEQ ID NO: 43, SEQ ID NO: 44, SEQ ID NO: 45, SEQ ID NO: 46 и SEQ ID NO: 47.

В некоторых аспектах сконструированная вирусная нуклеиновая кислота содержит гетерологичную последовательность нуклеиновой кислоты, функционально связанную с промотором, содержащим последовательность нуклеиновой кислоты, содержащуюся в SEQ ID NO: 21, или ее часть.

В некоторых аспектах сконструированная вирусная нуклеиновая кислота содержит гетерологичную последовательность нуклеиновой кислоты, функционально связанную с терминатором, содержащим последовательность нуклеиновой кислоты SEQ ID NO: 22, или ее часть.

В некоторых вариантах реализации настоящее раскрытие представляет способ создания сконструированного вируса, представляющего интерес, обладающего двумя или более необходимыми вирусными свойствами, включающими: (а) предоставление первого вирусного генома; и (б) создание сконструированного вирусного генома путем объединения по меньшей мере одного фрагмента первого вирусного генома по меньшей мере с одной молекулой восстанавливающей нуклеиновой кислоты для создания второго вирусного генома, содержащего по меньшей мере одну модификацию по сравнению с первым вирусным геномом; при этом второй вирусный геном, будучи введенным в клетку-хозяина, способен продуцировать вирусные частицы с двумя или более улучшенными вирусными свойствами.

В некоторых аспектах способ дополнительно содержит (с) повторение этапов (а)-(б) в одном или более повторениях.

В некоторых аспектах каждое улучшенное вирусное свойство выбрано из группы, состоящей из круга хозяев, вирусного литического цикла, адсорбции, прикрепления, инъекции, репликации и сборки, лизиса, величины выхода фага, ускользания от иммунологического надзора, иммунной стимуляции, иммунной дезактивации, дисперсии биопленки, устойчивости бактерий к фагам, антибиотической сенситивности бактерий, модуляции вирулентности и направленного расщепления или редактирования генома хозяина.

В некоторых аспектах улучшенное свойство или улучшенные свойства и улучшенное вирусное свойство или улучшенные вирусные свойства используются взаимозаменяемо.

В некоторых аспектах создание сконструированного вирусного генома на стадии (б) включает: (1) расщепление *in vitro* области первого вирусного генома с использованием эндонуклеазы; и (2) сборка по

меньшей мере одного фрагмента расщепленного первого вирусного генома с по меньшей мере одной молекулой восстанавливающей нуклеиновой кислоты.

В некоторых аспектах первый вирусный геном выделяют из вирусных частиц.

В некоторых аспектах первый вирусный геном или по меньшей мере одна восстанавливающая молекула нуклеиновой кислоты синтезируется в условиях *de novo*.

В некоторых аспектах синтез *de novo* включает объединение химически синтезированных молекул нуклеиновой кислоты, ПЦР-амплифицированных последовательностей нуклеиновой кислоты, расщепленных фрагментов выделенных молекул нуклеиновой кислоты или любых их комбинаций.

В некоторых аспектах первый вирусный геном или по меньшей мере одну восстанавливающую молекулу нуклеиновой кислоты амплифицируют перед расщеплением *in vitro*.

В некоторых аспектах первый вирусный геном по меньшей мере 3 т.п.н., по меньшей мере 10 т.п.н., по меньшей мере 18 т.п.н., по меньшей мере 25 т.п.н. или по меньшей мере 30 т.п.н.

В некоторых аспектах сборку выполняют *in vitro* или *in vivo*.

В некоторых аспектах сборку выполняют *in vitro* со смесью, содержащей: (a) выделенную 5'-3'-экзонуклеазу, которая не обладает 3'-экзонуклеазной активностью; (b) выделенную невязкую замещающую ДНК-полимеразу с 3'-экзонуклеазной активностью или смесь указанной ДНК-полимеразы со второй ДНК-полимеразой, которая не обладает 3'-экзонуклеазной активностью; (c) выделенную лигазу; и (d) смесь дНТФ в условиях, эффективных для вставки фрагмента в расщепленную вирусную нуклеиновую кислоту с образованием рекомбинантной нуклеиновой кислоты, содержащей сконструированный вирусный геном.

В некоторых аспектах эндонуклеаза представляет собой РНК-направляемую нуклеазу.

В некоторых аспектах способ дополнительно включает по меньшей мере одну гидовую РНК.

В некоторых аспектах нуклеаза, управляемая РНК, представляет собой Cas9 или фермент, полученный из Cas9, и в котором по меньшей мере одна гидовая РНК включает 1) химерную гРНК или 2) crРНК и tracrРНК.

В некоторых аспектах эндонуклеазу инактивируют нагреванием или удаляют перед сборкой.

В некоторых аспектах система расщепления *in vitro* дополнительно содержит спермидин.

В некоторых аспектах способ дополнительно включает трансформирование сконструированного вирусного генома в клетку-хозяина.

В некоторых аспектах способ дополнительно включает использование упаковочного набора *in vitro* для упаковки сконструированного вирусного генома в вирусные частицы.

В некоторых вариантах реализации настоящее раскрытие представляет собой сконструированный вирус, полученный любым из способов, описанных в настоящем документе. В некоторых аспектах сконструированный вирус представляет собой любой из сконструированных вирусов, описанных в настоящем документе.

В некоторых вариантах реализации настоящее раскрытие представляет набор для конструирования молекул вирусной нуклеиновой кислоты, содержащий: (a) очищенную рекомбинантную РНК-направляемую нуклеазу; (b) выделенную 5'-3'-экзонуклеазу, которая не обладает 3'-экзонуклеазной активностью; (c) выделенную невытесняющую ДНК-полимеразу с 3'-экзонуклеазной активностью или смесь указанной ДНК-полимеразы со второй ДНК-полимеразой, которая не обладает 3'-экзонуклеазной активностью; и (d) выделенную термостабильную лигазу.

В некоторых аспектах набор дополнительно содержит один или более из: (1) краудинг-агента; (2) смеси дНТФ; и (3) подходящего буфера.

В некоторых аспектах набор дополнительно содержит специально разработанные гидовые РНК.

В некоторых аспектах набор дополнительно содержит специально разработанные синтезированные молекулы нуклеиновой кислоты, предназначенные для использования в качестве вставленного фрагмента ДНК в реакции сборки.

В некоторых аспектах набор дополнительно содержит компетентные клетки-хозяева для трансформации.

В некоторых аспектах набор дополнительно содержит выделенные вирусные геномные нуклеиновые кислоты.

В некоторых вариантах реализации изобретение представляет собой систему сконструированной вирусной нуклеиновой кислоты *in vitro*, содержащую выделенную вирусную нуклеиновую кислоту, рекомбинантную РНК-направляемую нуклеазу, по меньшей мере одну гидовую РНК и фрагмент нуклеиновой кислоты, который должен быть вставлен в сайт расщепления выделенной нуклеиновой кислоты.

В некоторых аспектах система такова, что рекомбинантная РНК-направляемая нуклеаза и по меньшей мере одна гидовая РНК образуют комплекс, способный расщеплять выделенную вирусную нуклеиновую кислоту.

В некоторых аспектах система дополнительно содержит спермидин.

В некоторых аспектах система дополнительно содержит выделенную 5'-3'-экзонуклеазу, которая не обладает 3'-экзонуклеазной активностью; выделенную невытесняющую ДНК-полимеразу с 3'-экзонуклеазной активностью или смесь указанной ДНК-полимеразы со второй ДНК-полимеразой, которая не

обладает 3'-экзонуклеазной активностью; выделенную лигазу; и смесь дНТФ, причем система находится в условиях, которые являются эффективными для вставки фрагмента нуклеиновой кислоты в выделенную вирусную нуклеиновую кислоту в месте расщепления нуклеазой, управляемой РНК, с образованием рекомбинантной вирусной нуклеиновой кислоты.

В некоторых аспектах система, описанная в настоящем документе, такова, что рекомбинантная вирусная нуклеиновая кислота способна продуцировать вирусные частицы, не встречающиеся в природе, по меньшей мере с двумя улучшенными вирусными свойствами по сравнению с вирусными частицами, полученными из несконструированной вирусной нуклеиновой кислоты. В некоторых примерах улучшенное вирусное свойство или свойства выбраны из группы, состоящей из круга хозяев, вирусного литического цикла, адсорбции, прикрепления, инъекции, репликации и сборки, лизиса, величины выхода фага, ускользания от иммунологического надзора, иммунной стимуляции, иммунной дезактивации, дисперсии биопленки, устойчивости бактерий к фагам, антибиотической сенсбилизации бактерий, модуляции вирулентности и направленного расщепления или редактирования генома хозяина.

В некоторых аспектах в системе, описанной в настоящем документе, нуклеаза, управляемая РНК, представляет собой Cas9 или фермент, полученный из Cas9. В некоторых аспектах РНК-направляемую нуклеазу инактивируют или удаляют перед сборкой.

В некоторых вариантах реализации настоящее раскрытие представляет способ конструирования последовательности нуклеиновой кислоты, включающий: (a) обеспечение нуклеиновой кислоты; (b) расщепление *in vitro* области нуклеиновой кислоты с использованием нуклеазы, управляемой РНК; и (c) сборку рекомбинантной нуклеиновой кислоты путем вставки фрагмента ДНК в расщепленную нуклеиновую кислоту, причем сборку проводят *in vitro* в одном сосуде со смесью компонентов, включающих: (i) выделенную 5'-3'-экзонуклеазу, которая не обладает 3'-экзонуклеазной активностью; (ii) выделенную невымещающую ДНК-полимеразу с 3'-экзонуклеазной активностью или смесь указанной ДНК-полимеразы со второй ДНК-полимеразой, которая не обладает 3'-экзонуклеазной активностью; (iii) выделенную лигазу; и (iv) смесь дНТФ в условиях, эффективных для вставки фрагмента в расщепленную нуклеиновую кислоту с образованием рекомбинантной нуклеиновой кислоты.

В некоторых аспектах нуклеаза, управляемая РНК, представляет собой Cas9 или фермент, полученный из Cas9. В некоторых примерах нуклеазу, управляемую РНК, инактивируют воздействием тепла или удаляют до сборки.

В некоторых аспектах способ дополнительно включает: (d) трансформацию рекомбинантной нуклеиновой кислоты в клетку-хозяина.

В некоторых аспектах настоящее раскрытие относится к способу получения нуклеиновой кислоты, причем нуклеиновая кислота представляет собой плазмиду, выделенную из клетки-хозяина. В некоторых аспектах плазида составляет по меньшей мере 5 т.п.н. В некоторых аспектах плазида составляет по меньшей мере 6 т.п.н. В некоторых аспектах плазида составляет по меньшей мере 10 т.п.н. В некоторых аспектах плазида составляет по меньшей мере 15 т.п.н. В некоторых аспектах плазида составляет по меньшей мере 20 т.п.н.

#### **Краткое описание графических материалов**

На фиг. 1A-1E показана схема процесса *in vitro* для прямого конструирования вирусных геномов. А) Геномы выделяют из очищенных вирусных частиц с использованием способов, известных специалистам в данной области техники. Серыми линиями показан пример дцДНК вирусного генома. Светло-серыми линиями на концах генома обозначены прямые концевые повторы, обычно встречающиеся во многих вирусных геномах. Б) Вирусные геномы затем расщепляют на одном или нескольких участках местоположения, в частности, используя нуклеазу, управляемую РНК, такую как Cas9, в сочетании с очищенными нацеливающими РНК, такими как химерные гРНК, сгРНК и трагРНК или только сгРНК. На иллюстрации изображена РНК-направляемая нуклеаза, нацеленная на определенные вирусные геномные локализации, в соответствии с указанной РНК. В) РНК-направляемую нуклеазу инактивируют с использованием способов, известных в данной области техники, включая, но без ограничений, воздействие тепла и/или удаление с использованием классической экстракции фенол-хлороформом. Г) Вставку ДНК или РНК выполняют с использованием способов, известных в данной области техники, включая, но без ограничений, синтез *in vitro*, амплификацию (ПЦР) или фермент-опосредуемое освобождение от плазмид, вирусов или бактериальной геномной ДНК (гДНК). На диаграмме изображены вновь созданные вставки (темно-серые линии) с областями гомологии, соответствующими вирусным последовательностям, фланкирующим сайт(ы) расщепления РНК-направляемой нуклеазой (серые концевые области). Д) Расщепленные вирусные геномы и очищенную вставку собирают *in vitro* с использованием методов, известных в данной области техники, включая, но без ограничений, метод сборки Gibson Assembly, SLIC (метод безлигазного клонирования, в присутствии белка RecA) и/или сборку Golden Gate. На иллюстрации показан собранный рекомбинантный геном, который теперь содержит новую последовательность вставки (темно-серые линии) в нужной локализации. Е) Рекомбинантные вирусные геномы трансформируют непосредственно в клетки-хозяева с использованием методов, известных в данной области техники, включая, но без ограничений, электропорацию или химическую трансформацию. На фигуре показано восстановление функциональных вирусных частиц после трансформации инфекционного вирусного генома в восприим-

чивые клетки-хозяева;

на фиг. 2А-2Е - конструирование *in vitro* вирусного генома. А) Очистка ~ 43 т.п.н. дцДНК вирусного генома LUZ19 непосредственно от вирусных частиц. Б) Сайтспецифическое расщепление очищенного вирусного генома в двух независимых локализациях для удаления фрагмента гена *gp7* с использованием РНК-зависимой нуклеазы Cas9 и *in vitro* транскрибируемых гРНК. В) ПЦР использовали для амплификации гена *gp7* из вируса ФКФ77. Г) Метод сборки *in vitro* Gibson Assembly использовали для последовательной специфической интеграции ПЦР-амплифицированного фрагмента гена *gp7* ФКФ77 в расщепленный геном LUZ19. Д) Инфекционные геномы, собранные *in vitro*, трансформировали непосредственно в клетки-хозяева для восстановления функциональных вирусных частиц, что подтверждается образованием бактерий. Е) Внутренние и внешние праймеры использовали в ПЦР для проверки того, что вирусы содержат новый фрагмент ДНК в правильном геномном сайте. Все тестируемые вирусные клоны были ПЦР-положительными для новой вставки фрагмента *gp7* ФКФ77 (правые 7 полос);

на фиг. 3А-3Б - создание вируса с улучшенными вирусными свойствами после *in vitro* вирусного геномного конструирования. А) На диаграмме показаны геномы естественного вируса LUZ19 и сконструированного производного, несущего ген *gp18* вируса LKD16 вместо естественной последовательности *gp18* LUZ19. Черными стрелками обозначены открытые рамки считывания LUZ19, а серой стрелкой указан недавно встроенный ген *gp18* LKD16. Б) Слева на диаграмме Вена показаны общие и отдельные бактерии-хозяева, инфицированные вирусами LUZ19 и LKD16. Были протестированы разнообразные коллекции из 282 клинических изолятов *P.aeruginosa*. Справа на диаграмме Вена показано, что сконструированный вирус LUZ19, содержащий ген *gp18* LKD16, имеет расширенный круг хозяев, включая 3 из 6 штаммов, ранее зараженных только LKD16;

на фиг. 4А-4В - схемы, демонстрирующие процесс, используемый для идентификации и выбора генетических элементов и точечных мутаций, необходимых для расширения круга хозяев и разработки способности вирусов с широким кругом хозяев заражать полный круг хозяев рода вирусов А) Схематическое представление процесса, используемого для идентификации мутаций, ответственных за специфичность к кругу хозяев. Б) На схематическом представлении изображены модификации генома, необходимые для создания вируса LUZ19 с широким кругом хозяев (ШКХ-LUZ19); звездочки (\*) определяют локализацию каждой точечной мутации, связанной с кругом хозяев. Метки *gp13* C17Y, *gp18* D36Y, *gp38* D82G и 183S, и *gp40* N253D описывают генные продукты и аминокислотную точечную мутацию, связанную с расширением круга хозяев LUZ19. PA7245, PA7255, PA7410, PA7427, PA7503 и PA7686 представляют собой клинические изоляты *P.aeruginosa*, восприимчивые только к LKD16 и ШКХ-LUZ19; PA7649 представляют собой клинические изоляты *P.aeruginosa*, чувствительные только к ФКМВ и ШКХ-LUZ19. Клинические изоляты, инфицированные после добавления данной мутации, изображены выше данной мутации. В) Слева на диаграмме Вена показаны общие и отдельные бактерии-хозяева, инфицированные вирусами LUZ19, LKD16 и ФКМВ. Справа на диаграмме Вена показано, что сконструированный вирус ШКХ-LUZ19, содержащий описанные выше точечные мутации, способен заразить все 67 штаммов, восприимчивых к роду вирусов ФКМВ;

на фиг. 5А-5Д показано, что мутация белка *Gp34* LUZ19 улучшает литическую активность. А) Белок *Gp34* LUZ19 является членом комплекса трубок вируса (см. вставленное изображение). Б) Анализ бляшкообразования в мягком агаре для двух связанных фагов, экспрессирующих или *Gp34* LUZ19 дикого типа или *Gp34* с разницей в лейцине 55 (L55Δ) (фаг \*). Снимки были сделаны в течение двух суток и было показано, что фаг, экспрессирующий мутацию L55 *Gp34*, имеет зоны повышенного лизиса. В) Анализ биопленки с окрашиванием кристаллическим фиолетовым, экстраполирующий биомассу биопленки в виде измерения включения кристаллического фиолетового. Фаг LUZ19\*, экспрессирующий *Gp34* L55Δ, был способен лучше разрушать биопленку *P.aeruginosa*, предварительно сформированную в течение 8 ч по сравнению с LUZ19 дикого типа. Гентамицин в десятикратном размере минимальной ингибирующей концентрации (МИК) использовали для полного удаления биопленки. Г) На иллюстрации показано расположение мутации *gp34* по сравнению с геномом LUZ19 дикого типа. Д) На таблице продемонстрирована разница в величине поглощения и величине выхода фага между LUZ19 и LUZ19, экспрессирующим *Gp34* L55Δ;

на фиг. 6А-6Е представлены схемы, демонстрирующие итеративное конструирование вируса с улучшением до двух независимых свойств. А) Схематическое представление LUZ19<sub>LKD16gp18</sub> вирусной гДНК, в которой ген *gp18* LUZ19 дикого типа заменен гомологом LKD16. Черным цветом обозначена последовательность генома LUZ19 дикого типа; серым цветом - *gp18* от LKD16. Б) Восприимчивость лабораторных и МЛР клинических штаммов к очищенному исходному (LKD16 и LUZ19) и сконструированному вирусу LUZ19\*<sub>LKD16gp18</sub>, демонстрирующему объединение круга хозяев. В) Схематическое представление вирусной гДНК LUZ19\*<sub>LKD16gp18</sub>, в которой и лейцин, закодированный в позиции 55 *Gp34*, был удален, а *gp18* LUZ19 заменен на *gp18* вируса LKD16. Черным цветом обозначена геномная последовательность LUZ19 дикого типа; серым цветом - *gp18* от LKD16; серая звезда обозначает *gp34*Δ<sub>leu55</sub>. Г) Восприимчивость лабораторных и МЛР клинических штаммов к очищенному исходному (LKD16, LUZ19 и LUZ19<sub>LKD16gp18</sub>, содержащим *gp18* из вируса LKD16) и сконструированному вирусу (LUZ19,



несущему делецию лейцина, закодированную в позиции 55 GP34 и LUZ19\*<sub>LKD16gp18</sub>), демонстрирующему объединение круга хозяев у вирусов LUZ19<sub>LKD16gp18</sub> и LUZ19\*<sub>LKD16gp18</sub>. Д) Оценка литической активности фага дикого типа и сконструированного фага против бактерий, прикрепленных к монослоям кератиноцитов. Количество бактерий PA01K и PA7245, прикрепленных к клеткам, отмечали как процент от общего количества бактерий, инкубированных с монослоями кератиноцитов. Данные демонстрируют улучшенную литическую активность вирусов LUZ19\* и LUZ19\*<sub>LKD16gp18</sub>. Е) Улучшенное раннее разрушение 8-часовой биопленки PA01K и PA7245 при помощи сконструированного фага по сравнению с исходными вирусами. Гентамицин использовали в десятикратном размере минимальной ингибирующей концентрации для полного удаления биопленки. Приведенные данные представляют собой три независимых эксперимента, выполненных в трех повторениях. Столбцы представляют собой среднее  $\pm$  СОС; \* P < 0,01; \*\* P < 0,001; \*\*\* p < 0,0001;

на фиг. 7А-7Е показаны схемы, демонстрирующие второй пример итеративного конструирования вируса с улучшением до двух независимых свойств. А) Схематическое представление LUZ19, сконструированного для экспрессии различных генетических закодированных полезных нагрузок из улучшенного локуса gp49. Ген gp49 заменяли касетой, содержащей ген интереса (ГИ), фланкированный основным промотором и терминатором капсида (gp32) (P<sub>gp32</sub> и T<sub>gp32</sub>). Использовали биопленку, диспергирующую ГИ: ЭПС-деполимеразы (Pp15gp44 - хвостовой шип gp44 из cp15 *Pseudomonas putida*; NTUgp34 - хвостовой шип gp34 из фага *Klebsiella pneumoniae* NTUH-K2044-K1-1 (NTU); LKA1gp49 - хвостовой шип gp49 из фага *P.aeruginosa* LKA1), поверхностно-активные фенол-растворимые модулины из *Staphylococcus epidermidis* (PSMa) и *Staphylococcus aureus* (PSMa3 и PSMb2) и DspB-сурфактин (дисперсин В-сурфактин) из *Aggregatibacter actinomycetemcomitans*. Б) Анализ дисперсии биопленки, демонстрирующий активность сконструированного фага LUZ19 против 24 ч биопленки PA01K *P.aeruginosa*, обработанной 100 фагами в течение 3 ч. Гентамицин использовали в десятикратном размере минимальной ингибирующей концентрации (МИК). В) Схематическое представление ранее сконструированного фага ШКХ-LUZ19, дополнительно сконструированного для экспрессии ГИ из сконструированного локуса gp49. Г) Анализ дисперсии биопленки, демонстрирующий сконструированный ШКХ-LUZ19, дополнительно модифицировали для экспрессии ферментов и сурфактинов с активностью против 24 ч биопленки *P.aeruginosa* PA01K, обработанной 100 фагами в течение 3 ч. Сконструированная полезная нагрузка: ЭПС-деполимераза Pp15gp44 и SePSMa. Гентамицин использовали в десятикратном размере МИК. Д) Восприимчивость лабораторных и клинических штаммов к очищенным исходным (LKD16 и LUZ19) и производным LUZ19, демонстрирующим объединение и сохранение круга хозяев при дальнейшей разработке для экспрессии фрагментов, диспергирующих биопленку. Е) На диаграмме Венна показано сохранение круга хозяев ШКХ-LUZ19 после добавления биопленки, диспергирующей полезную нагрузку Pp5gp44 и SePSMa;

на фиг. 8А-8В представлены схемы, демонстрирующие создание вируса, способного помешать клеткам-хозяевам получить резистентность к вирусу в сочетании с субингибирующей концентрацией антибиотика. А) Схематическое изображение LUZ19 дикого типа, сконструированного для экспрессии лизинов из фага или MS2 или PRR1. Б) и В) Анализ зависимости "время - летальное действие", демонстрирующей сенсбилизацию *P.aeruginosa* PA01K к субингибирующим концентрациям карбенициллина (Cb - 1/5  $\times$  МИК) при помощи LUZ19, экспрессирующих лизины из оцПНК бактериофага. Эти данные демонстрируют, что сконструированный бактериофаг, экспрессирующий ненативные лизины в сочетании с субингибирующими концентрациями антибиотиков, может препятствовать быстрому приобретению резистентности бактерий к единственному вирусу;

на фиг. 9А-9Г - схемы, демонстрирующие создание 2-го вируса, способного помешать клеткам-хозяевам получить вирусную резистентность. А) Схематическое изображение LUZ19 дикого типа, сконструированного для экспрессии белка бактериоцина P<sub>yoS5</sub> из сконструированного локуса gp49. Б) Анализ зависимости "время - летальное действие", демонстрирующие, что рост штамма PA7416 XRSPA был первоначально ингибирован диким типом LUZ19, однако бактерии быстро избегали вируса, что приводило к повторному росту бактерий. Приблизительно  $1 \times 10^7$  КОЕ добавляли на лунку. Высокую МИ (множественность инфицирования)=10 БОЕ/КОЕ и низкую МИ=0,01 БОЕ/КОЕ указанного вируса или носителя добавляли на 0 ч. В) Тесты на поражение, демонстрирующие, что LUZ19, кодирующий P<sub>yoS5</sub>, способен ингибировать рост и повторный рост штамма PA7416 XDRPA по сравнению с вирусом дикого типа. Приблизительно  $1 \times 10^7$  КОЕ добавляли на лунку. В течение 0 ч добавляли высокую МИ=10 БОЕ/КОЕ и низкую МИ=0,01 БОЕ/КОЕ указанного вируса или носителя. Г) Сравнение роста PA7416 через 24 ч в присутствии или LUZ19 дикого типа или LUZ19+P<sub>yoS5</sub>. На графике отображены данные для эксперимента с низким МИ;

на фиг. 10 - схематический чертеж системы, включающей редактирование целевого вирусного генома с фенотипическим скринингом бактериофагов для создания генетически сконструированного бактериофага с улучшением двух или более свойств. Система опирается на неоднократное повторение этапов скрининга и секвенирования мутантных или естественных вирусов с желаемыми фенотипическими признаками и интегрированием данных признаков в одно или более вирусных "шасси" на одном или бо-

лее этапах конструирования. Данный процесс обеспечивает прямой и рациональный способ быстрой идентификации генетических элементов, лежащих в основе специфического фенотипического признака бактериофага, интегрирования множественных независимых мутаций или аллелей в один геном бактериофага и создания искусственного вируса, сочетающего два или более улучшенных признака;

на фиг. 11А-11Ж показано конструирование *in vitro* генома фага *E.coli* M13. А) Схематическое изображение умеренного фага M13mp18 и M13<sub>паприка</sub> *E. coli*. Б) Электрофорез в геле расщепленной кольцевой геномной ДНК M13 *in vitro* с использованием гРНК 1 и 2 в независимых реакциях с РНК-направляемой эндонуклеазой Cas9. На диаграмме ниже на геле показана кольцевая и линейная природа нерасщепленных и дважды расщепленных геномов M13 соответственно. Эти данные показывают, что обе гРНК точно и полностью расщепляют дцДНК M13 в заданных локализациях. В) Гель-электрофорез расщепленной кольцевой геномной ДНК M13 *in vitro* с использованием как гРНК, так и РНК-направляемой эндонуклеазы Cas9, в той же реакции (двойное расщепление). На диаграмме ниже на геле показана кольцевая и линейная природа нерасщепленных и дважды расщепленных геномов M13 соответственно. Г) Гель-электрофорез, визуализирующий ПЦР полученной вставки, содержащей флуоресцентный репортер паприка. Д) Трансформация расщепленной и собранной *in vitro* сконструированной гДНК M13<sub>паприка</sub> в клетки *E.coli* для восстановления функциональных вирусных частиц. Вирусные бляшки неотчетливы и скрыты, потому что M13 является умеренным бактериофагом, который не лизирует клетки-хозяева, что приводит к плохому образованию бляшек. В качестве положительного контроля использовали нерасщепленную гДНК M13. ГДНК M13, расщепленная Cas9, но собранная в отсутствие вставки (без вставки) демонстрирует как полноту расщепления, так и низкий уровень фона. Е) Проверка при помощи ПЦР бляшек сконструированного M13<sub>паприка</sub>. Прямые и обратные праймеры были разработаны вне областей гомологии вставки. Несконструированная гДНК M13 продуцировала 0,9 т.п.н. продукт и использовалась в качестве отрицательного контроля для реакций ПЦР. Ж) Флуоресцентные (нижние) и яркие (верхние) изображения исходной и сконструированной M13<sub>паприка</sub> во время образования бляшек;

на фиг. 12А-12Д - конструирование *in vitro* второго генома фага *E.coli*. А) Схематическое представление фага *E.coli* X АсII. Линейный фаговый геном имеет размер 48,5 т.п.н. Б) Гель-электрофорез расщепленной геномной ДНК  $\lambda$  *in vitro* с использованием гРНК 1 и 2 в независимых реакциях с РНК-направляемой эндонуклеазой. На диаграмме ниже на геле показаны линейные нерасщепленные и ожидаемые продукты расщепления. Эти данные показывают, что обе гРНК точно и полностью расщепляют дцДНК  $\lambda$  в заданных локализациях. В) Гель-электрофорез дважды расщепленной геномной ДНК  $\lambda$  *in vitro* с использованием как гРНК, так и РНК-направляемой эндонуклеазы в той же реакции. На диаграмме ниже на геле показаны линейные нерасщепленные и ожидаемые продукты двойного расщепления. Г) На схеме показано использование фагового упаковочного буфера фага  $\lambda$  для упаковки генома дикого типа и рекомбинантного фага *in vitro*. Cas9 с дважды расщепленным и собранным геномом фага были упакованы *in vitro* в соответствии с протоколом изготовителя и покрыты *E.coli* для восстановления недавно сконструированного фага  $\lambda$   $\Delta$ СII. Д) ПЦР-проверка гена  $\Delta$ СII  $\lambda$ . Прямой праймер был расположен снаружи от области конструирования. Делеции положительных клонов имеют ожидаемый размер 300 п.о;

на фиг. 13А-13Г - конструирование *in vitro* последовательностей из цитомегаловируса человека (ЦМВ). А) Схематическое изображение вирусного генома ЦМВ длиной 235 т.п.н. Сверху геном в форме сигары представляет собой полноразмерный геном, а черный раздел обозначает область манипуляции. Небольшая белая секция означает 235 п.о. вставку, добавляемую с использованием метода инженерии *in vitro*, описанного в настоящем документе. Б) Гель-электрофорез с дважды расщепленной плазмидой *in vitro*, содержащей 17,8 т.п.н. области генома ЦМВ, с использованием двух гРНК и РНК-направляемой эндонуклеазы Cas9. На диаграмме ниже на геле показаны кольцевые нерасщепленные и линейные дважды расщепленные продукты. Эти данные показывают, что обе гРНК точно и полностью расщепляют последовательность дцДНК ЦМВ в заданной локализации. В) Гель-электрофорез, визуализирующий ПЦР созданной вставки, содержащей новую последовательность вставки RL13. Г) ПЦР-проверка сконструированной последовательности ЦМВ. Прямой праймер был расположен снаружи от области конструирования. Вставки положительных клонов имеют ожидаемый размер 500 п.о;

на фиг. 14А-14Е - быстрая идентификация концов фага. А) Выделение геномной ДНК из очищенных вирусных частиц. Б) Секвенирование нового поколения гДНК (MiSeq или PacBio) и автоматическое слияние высококачественных считываний ДНК в более длинные сборки для реконструирования исходной последовательности. В светло-сером цвете - ПТП - прямые терминальные повторы. Программное обеспечение автоматической сборки неправильно размещает ПТП терминально повторяющихся геномов во внутренней области вирусной последовательности. Геномные физические концы подтверждены целевым расщеплением Cas9 предсказанной последовательности. В) Прогнозирование *in silico* физических концов генома на основе идентификации областей секвенирования двойного охвата и поиска BLAST, который соответствует близкородственному терминально повторяющемуся геному. Физические концы подтверждаются при помощи расщепления эндонуклеазой Cas9 предсказанных физических концов. Г) После инактивации Cas9 фрагменты ДНК, соответствующие геномным физическим концам, очищают и секвенируют. Д) Точная сборка генома на основе последовательности физических концов. Е)

Пример картирования геномных физических концов фагов LBL3 и 14-1 (терминально повторяющихся геномов) с использованием Cas9-направленного расщепления в определенном положении, предсказанного при помощи реанжировки генома *in silico*. Светлосерыми стрелками указано на очищенные и секвенированные фрагменты ДНК;

на фиг. 15А-15В - схематическое изображение конструирования и стратегия синтеза химерной оgРНК. А) На иллюстрации показано расположение мотивов NGG РАМ (темно-серые подчеркнутые последовательности) и целевых сайтов оgРНК (светлосерые последовательности), фланкирующих представляющий интерес ген (ГИ). Черные последовательности обозначают оставшиеся вирусные геномные последовательности. Б) Конструирование олигонуклеотидов, используемых в качестве матриц для транскрипции оgРНК *in vitro*. Последовательности, составляющие промотор Т7, последовательность направляющей оgРНК и консервативную область химерной оgРНК, обозначены соответственно подчеркнутым темно-серым, светло-серым и черным текстом. В) Схема *in vitro* транскрибируемой химерной оgРНК. Светло-серыми и черными последовательностями указаны нацеливающие и консервативные химерные области, составляющие каждую функциональную оgРНК соответственно. Все Ns обозначают вариабельные последовательности, используемые для изменения целевой специфичности каждой оgРНК.

#### **Подробное описание иллюстративных вариантов реализации изобретения**

Изобретение обеспечивает композиции и способы для конструирования *in vitro* и дополнительно относится к улучшению вирусных свойств. Изобретение дополнительно обеспечивает способ конструирования нуклеиновых кислот *in vitro*.

До того как будут описаны настоящие композиции и способы, следует понимать, что это раскрытие не ограничивается конкретными композициями, способами и описанными экспериментальными условиями, поскольку такие композиции, способы и условия могут варьироваться. Следует также понимать, что используемая в настоящем документе терминология предназначена только для описания конкретных вариантов реализации изобретения и не предназначена для ограничения, поскольку объем настоящего раскрытия будет ограничен только в прилагаемой формуле изобретения.

Если не указано иное, все технические и научные термины, используемые в настоящем документе, имеют то же значение, которое обычно понимается специалистом в данной области техники, к которому относится настоящее раскрытие. Хотя любые способы и материалы, аналогичные или эквивалентные описанным в настоящем документе, могут быть использованы в практике или тестировании раскрытия, ниже описаны предпочтительные способы и материалы. Определения, изложенные ниже, предназначены для понимания раскрытия, но ни в коем случае не должны считаться заменяющими понимание терминов, которыми обладают специалисты в данной области техники.

Как используется в данном описании и приложенной формуле изобретения, формы единственного числа включают ссылки на множественное число, если из контекста явно не следует иное. Так, например, ссылки на "способ" включают один или более способов и/или этапы, описанные в настоящем документе, которые станут очевидными для специалистов в данной области техники после прочтения настоящего раскрытия и т.д.

Используемые в настоящем документе термины "около" или "приблизительно" при обращении к любому численному значению обозначают значение плюс или минус 10% от указанного значения. Например, "около 50°C" (или "приблизительно 50°C") охватывает диапазон температур от 45 до 55°C включительно. Аналогично, "около 100 ммоль/л" (или "приблизительно 100 ммоль/л") охватывает диапазон концентраций от 90 до 110 ммоль/л включительно. В качестве альтернативы "около" или "приблизительно" может означать в пределах 5% от заявленного значения, или в некоторых случаях в пределах 2,5% от заявленного значения, или "около" может означать округление до ближайшей значащей цифры. Все диапазоны, предусмотренные в заявке, включают значения верхнего и нижнего концов диапазона.

Термины "клетки", "культуры клеток", "клеточная линия", "клетки рекомбинантного хозяина", "клетки-реципиенты" и "клетки-хозяева", используемые в настоящем документе, включают первичные клетки-субъекты и любое их потомство независимо от количества пассажей. Следует понимать, что не все потомства точно идентичны исходной клетке (из-за преднамеренных или непреднамеренных мутаций или различий в окружающей среде); однако такое измененное потомство включено в эти термины, при условии, что потомство сохраняет те же функциональные свойства, что и у изначально трансформированной клетки.

Термин "сборка" или "сборка", используемый в настоящем документе, относится к соединению молекул ДНК или РНК.

Термин "восстанавливающая молекула нуклеиновой кислоты", используемый в настоящем документе, относится к молекуле нуклеиновой кислоты, которая может быть собрана с одним или более фрагментами ДНК или расщеплена, или отщеплена ДНК-плазмидой или молекулой нуклеиновой кислоты ДНК, с тем чтобы создать молекулу непрерывной последовательности нуклеиновой кислоты или закрытой плазмидной ДНК.

Термины "синтез *de novo*", "сборка *de novo*", "химический синтез" и "синтез ДНК" относятся к способам создания последовательностей нуклеиновых кислот без необходимости использования ранее существовавшей матрицы предшественника.

В тех способах изобретения, которые выполняют "in vitro", все белковые компоненты выделяют и/или практически очищают. Реакции сборки in vitro не проводят в живой клетке или с необработанным клеточным экстрактом; реакции проводят в бесклеточной среде.

"Функциональная молекула РНК" представляет собой молекулу РНК, которая может взаимодействовать с одним или более белками, или молекулами нуклеиновой кислоты для осуществления или участия в структурной, каталитической или регуляторной функции, которая влияет на экспрессию или активность гена, или продукта гена, отличного от гена, который продуцирует функциональную РНК. Функциональной РНК может быть, например, транспортной РНК (тРНК), рибосомальной РНК (рРНК), антисмысловой РНК (асРНК), микроРНК (микро РНК), короткой шпилечной РНК (кшРНК), малой интерферирующей РНК (миРНК), гидовой РНК (гРНК), crisp РНК (сгРНК), или трансактивирующей РНК (tracrРНК) системы CRISPR, малые ядрышковые РНК (мякРНК), РНК, взаимодействующие с PIWI, (пиРНК) или рибозимом.

Термин "ген" широко используется для обозначения любого сегмента молекулы нуклеиновой кислоты (обычно ДНК, но необязательно РНК), кодирующего полипептид или экспрессируемую РНК. Таким образом, гены содержат последовательности, кодирующие экспрессируемую РНК (которая может содержать полипептидные кодирующие последовательности или, например, функциональные РНК, такие как рибосомные РНК, тРНК, антисмысловые РНК, микроРНК, короткие шпилечные РНК, гРНК, сгРНК, tracrРНК, рибозимы и т.д.). Гены могут дополнительно содержать регуляторные последовательности, необходимые для их экспрессии или влияющие на их экспрессию, а также последовательности, связанные с белковой или РНК-кодирующей последовательностью в ее естественном состоянии, таком как, например, интронные последовательности, 5'- или 3'-нетранслируемые последовательности и т.д. В некоторых примерах ген может относиться только к кодирующей белок части молекулы ДНК или РНК, которая может содержать или не содержать интроны. Ген предпочтительно составляет более 50 нуклеотидов в длину, более предпочтительно более 100 нуклеотидов и может быть, например, от 50 нуклеотидов до 500000 нуклеотидов в длину, например от 100 нуклеотидов до 100000 нуклеотидов в длину или от около 200 нуклеотидов до около 50 000 нуклеотидов в длину, или от около 200 нуклеотидов до около 20 000 нуклеотидов в длину. Гены могут быть получены из различных источников, включая клонирование из источника, представляющего интерес, или путем синтеза из известной или предсказанной информации о последовательности.

Термин "нуклеиновая кислота" или "молекула нуклеиновой кислоты" относится к сегменту ДНК или РНК (например, мРНК), а также включает нуклеиновые кислоты, имеющие сконструированные остовы (например, пептидные нуклеиновые кислоты, закрытые нуклеиновые кислоты) или сконструированные или неестественные нуклеотидные основания. Молекулы нуклеиновой кислоты могут быть двухцепочечными или одноцепочечными; одноцепочечная нуклеиновая кислота, которая содержит ген или его часть, может быть кодирующей (смысловой) цепью или некодирующей (антисмысловой) цепью.

Термины "кодирующая последовательность" или "кодирующая область", используемые в настоящем документе, относятся к областям последовательности нуклеиновой кислоты, которые могут быть транскрибированы для получения функциональной РНК или транскрипта РНК, которые могут быть переведены в полипептид, когда они находятся под контролем соответствующих последовательностей контроля экспрессии и в присутствии соответствующих клеточных механизмов или ферментов. Термин "некодирующая последовательность" или "некодирующая область" относится к областям последовательности нуклеиновой кислоты, которые не транскрибируются и не транслируются в аминокислоты (например, интроны, нетранслируемые области и т.д.) или не транскрибируются или не образуют по меньшей мере часть зрелой функциональной последовательности РНК.

Используемый в настоящем документе термин "белок" или "полипептид" предназначен для охвата отдельного "полипептида", а также множественных "полипептидов" и относится к молекуле, состоящей из мономеров (аминокислот), линейно связанных амидными связями (также известные как пептидные связи). Термин "полипептид" относится к любой цепи или цепям из двух или более аминокислот и не относится к определенной длине продукта. Таким образом, пептиды, дипептиды, трипептиды, олигопептиды, "белок", "аминокислотная цепь" или любой другой термин, используемый для обозначения цепи или цепей из двух или более аминокислот, включены в определение "полипептид", и термин "полипептид" может использоваться вместо или взаимозаменяемо с любым из этих терминов.

Молекула нуклеиновой кислоты может быть "получена из" указанного источника, что включает выделение (полностью или частично) сегмента нуклеиновой кислоты из указанного источника. Молекула нуклеиновой кислоты также может быть получена из указанного источника путем, например, прямого клонирования, ПЦР-амплификации или искусственного синтеза из указанного полинуклеотидного источника или на основе последовательности, связанной с указанным полинуклеотидным источником. Гены или молекулы нуклеиновой кислоты, полученные из конкретного источника или вида, также включают гены или молекулы нуклеиновой кислоты, имеющие модификации последовательности относительно исходных молекул нуклеиновой кислоты. Например, ген или молекула нуклеиновой кислоты, полученные из источника (например, конкретный упоминаемый ген), могут включать одну или более мутаций относительно исходного гена или молекулы нуклеиновой кислоты, которые непреднамеренно

или намеренно введены, и если одна или более мутаций, включая замены, делеции или вставки, преднамеренно введены, изменения последовательности могут быть введены случайной или целевой мутацией клеток или нуклеиновых кислот, путем амплификации или другими методами молекулярной биологии или путем химического синтеза или любой их комбинации. Ген или молекула нуклеиновой кислоты, которая получена из упомянутого гена или молекулы нуклеиновой кислоты, которая кодирует функциональную РНК или полипептид, может кодировать функциональную РНК или полипептид, имеющий по меньшей мере 75%, по меньшей мере 80%, по меньшей мере 85%, по меньшей мере 90% или по меньшей мере 95% идентичности последовательности с эталонной или исходной функциональной РНК или полипептидом или с его функциональным фрагментом. Например, ген или молекула нуклеиновой кислоты, которая получена из упомянутого гена или молекулы нуклеиновой кислоты, которая кодирует функциональную РНК или полипептид, может кодировать функциональную РНК или полипептид, имеющий по меньшей мере 85%, по меньшей мере 90%, по меньшей мере 95%, по меньшей мере 96%, по меньшей мере 97%, по меньшей мере 98% или по меньшей мере 99% идентичности последовательности с эталонной или исходной функциональной РНК или полипептидом или с его функциональным фрагментом.

Используемую в настоящем документе "выделенную" нуклеиновую кислоту или белок удаляют из его естественной среды или окружения, в котором нуклеиновая кислота или белок существуют в природе. Например, выделенный белок или молекулу нуклеиновой кислоты удаляют из клетки или организма, с которым они связаны в своей естественной или природной среде. Выделенная нуклеиновая кислота или белок может быть в некоторых случаях частично или практически очищена, но для выделения не требуется определенный уровень очистки. Так, например, выделенная молекула нуклеиновой кислоты может быть последовательностью нуклеиновой кислоты, которая была вырезана из хромосомы, генома или эписомы, которая была встроена в природе.

"Очищенная" молекула нуклеиновой кислоты или нуклеотидная последовательность, или белковая или полипептидная последовательность, по существу, свободна от клеточного материала и клеточных компонентов. Например, очищенная молекула нуклеиновой кислоты или белка может быть свободна от химических веществ, если не считать буфер или растворитель. "Существенно свободный" не означает, что другие компоненты за пределами новых молекул нуклеиновой кислоты не поддаются определению.

Термины "естественного происхождения" и "дикого типа" относятся к форме, встречаемой в природе. Например, молекула нуклеиновой кислоты естественного происхождения или дикого типа, нуклеотидная последовательность или белок могут присутствовать и могут быть выделены из природного источника и ненамеренно модифицироваться при помощи человеческих манипуляций.

Термин "экспрессия", используемый в настоящем документе, включает экспрессию гена, по меньшей мере, на уровне продуцирования РНК, а "продукт экспрессии" включает полученный продукт, например полипептид или функциональную РНК (например, рибосомальную РНК, тРНК, антисмысловую РНК, микроРНК, кшРНК, рибозим и т.д.) экспрессируемого гена. Термин "повышенная экспрессия" включает изменение в экспрессии генов посредством повышения продукции мРНК и/или увеличения экспрессии полипептида. "Повышенное продуцирование", когда речь идет о количестве белка или количестве активного белка в результате экспрессии генов, скорости текучести белка, состояниях активации белка и т.п., включает увеличение количества полипептидной экспрессии на уровне ферментативной активности полипептида или их комбинации по сравнению с естественным продуцированием или ферментативной активностью полипептида.

"Молекула экзогенной нуклеиновой кислоты" или "экзогенный ген" относится к молекуле нуклеиновой кислоты или гену, который был введен ("трансформирован") в клетку или вирус. Трансформированным организмом можно назвать рекомбинантную клетку или вирус, в который можно ввести дополнительный экзогенный ген(ы). Потомство клетки или вируса, трансформированного молекулой нуклеиновой кислоты, также упоминается как "трансформированное" или "рекомбинантное", если оно унаследовало молекулу экзогенной нуклеиновой кислоты. Экзогенный ген может быть от другого вида (и, следовательно, "гетерологичного") или от одного и того же вида (и, следовательно, "гомологичного") по отношению к трансформируемому организму. "Эндогенная" молекула нуклеиновой кислоты, ген или белок представляет собой нативную молекулу нуклеиновой кислоты, ген или белок, когда она встречается в организме или естественным образом продуцируется организмом.

Кроме того, термин "экзогенный", используемый в настоящем документе, в контексте гена или белка, относится к генам или белкам, которые не получают из видов организма-хозяина.

Термин "трансген", используемый в настоящем документе, относится к экзогенному гену, т.е. к гену, введенному в микроорганизм или предшественник путем вмешательства человека.

Используемый в настоящем документе термин "ортолог" гена или белка относится к его функциональному эквиваленту у другого вида.

Номера доступа генов и белков, обычно представленные в настоящем документе в скобках после имени гена или вида, являются уникальными идентификаторами для записи последовательности, общедоступной на веб-сайте Национального центра биотехнологической информации (NCBI) ([ncbi.nlm.nih.gov](http://ncbi.nlm.nih.gov)), поддерживаемого Национальным институтом здравоохранения Соединенных Штатов. Идентификационный номер последовательности "GenInfo Identifier" (GI) специфичен для нуклеотидной

или аминокислотной последовательности. Если последовательность изменяется каким-либо образом, назначается новый номер GI. Инструмент "История изменений последовательности" доступен для отслеживания различных номеров GI, номеров версий и дат обновления для последовательностей, которые отображаются в определенной записи GenBank. Поиск и получение последовательностей нуклеиновых кислот или генов, или последовательностей белка на основе номеров доступа и номеров GI хорошо известны в данной области техники, например, в клеточной биологии, биохимии, молекулярной биологии и молекулярной генетике.

Используемые в настоящем документе термины "процент идентичности" или "гомология" в отношении нуклеиновых кислот или полипептидных последовательностей определяются как процент нуклеотидных или аминокислотных остатков в последовательности-кандидате, которые идентичны с известными полипептидами после выравнивания последовательностей для максимального процента идентичности и введения гэпов, если необходимо, для достижения максимального процента гомологии. N-концевая или C-концевая вставка, или делеция не должны истолковываться как влияющие на гомологию, а внутренняя делеция и/или вставки в последовательность полипептида менее чем около 30, менее чем около 20 или менее чем около 10 аминокислотных остатков не должны быть истолкованы как влияющие на гомологию. Гомологию или идентичность на уровне нуклеотидов или аминокислотных последовательностей можно определить с помощью анализа BLAST (средство поиска основного локального выравнивания) с использованием алгоритма, используемого программами blastp, blastn, blastx, tblastn и tblastx (Altschul (1997), *Nucleic Acids Res.* 25, 3389-3402, и Karlin (1990), *Proc. Natl. Acad. Sci. USA* 87, 2264-2268), которые предназначены для поиска сходства последовательностей. Подход, используемый программой BLAST, заключается в том, чтобы сначала рассмотреть аналогичные сегменты с гэпами и без них между запрашиваемой последовательностью и последовательностью базы данных, затем оценить статистическую значимость всех идентифицируемых совпадений и, наконец, суммировать только те соответствия, которые удовлетворяют предварительно выбранному порогу значимости. Для обсуждения основных проблем поиска сходства баз данных последовательностей см. Altschul (1994), *Nature Genetics* 6, 119-129. Параметры поиска для гистограмм, описаний, выравниваний, ожиданий (т.е. порог статистической значимости для сопоставления отчетов по последовательностям базы данных), обрезание, матрица и фильтр (низкая сложность) могут быть в настройках по умолчанию. Матрица выравнивания весов по умолчанию, используемая blastp, blastx, tblastn и tblastx, представляет собой матрицу BLOSUM62 (Henikoff (1992), *Proc. Natl. Acad. Sci. USA* 89, 10915-10919), рекомендованную для запрашиваемых последовательностей длиной более 85 (нуклеотидных оснований или аминокислот).

Для blastn, предназначенного для сравнения нуклеотидных последовательностей, матрица подсчета определяется отношениями M (т.е. оценкой вознаграждения для пары сопоставимых остатков) к N (т.е. оценкой штрафа за несоответствующие остатки), где значения по умолчанию для M и N могут быть +5 и -4 соответственно. Четыре параметра blastn можно настроить следующим образом: Q=10 (штраф за создание гэпа); R=10 (штраф за продолжение гэпа); wink=1 (генерирует попадания слов в каждом просматриваемом в данный момент положении вдоль запроса); и gapw=16 (задает ширину окна, в которой создаются выравнивания гэпов). Эквивалентные параметры параметра Blastp для сравнения аминокислотных последовательностей могут быть: Q=9; R=2; wink=1; и gapw=32. Сравнение Bestfit между последовательностями, доступными в пакете GCG версии 10.0, может использовать параметры ДНК GAP=50 (штраф за создание гэпа) и LEN=3 (штраф за расширение гэпа), а эквивалентные настройки в сравнении с белками могут быть GAP=8 и LEN=2.

Таким образом, когда речь идет о последовательностях полипептида или нуклеиновой кислоты согласно настоящему изобретению включают идентичности последовательностей по меньшей мере 40%, по меньшей мере 45%, по меньшей мере 50%, по меньшей мере 55%, по меньшей мере 70%, по меньшей мере 65%, по меньшей мере 70%, по меньшей мере 75%, по меньшей мере 80% или по меньшей мере 85%), например, по меньшей мере 86%, по меньшей мере 87%, по меньшей мере 88%, по меньшей мере 89%, по меньшей мере 90%, по меньшей мере 91%, по меньшей мере 92%, по меньшей мере 93%, по меньшей мере 94%, по меньшей мере 95%, по меньшей мере 96%, по меньшей мере 97%, по меньшей мере 98%, по меньшей мере 99% или около 100% идентичности последовательности с полноразмерным полипептидом или последовательностью нуклеиновой кислоты или их фрагментами, содержащими непрерывную последовательность по меньшей мере в 100, по меньшей мере в 125, по меньшей мере в 150 или более аминокислотных остатков всего белка; варианты таких последовательностей, например, в которых по меньшей мере один аминокислотный остаток был вставлен в N-и/или C-конец и/или в пределах раскрытой последовательности(ей), которые содержат(ит) вставку и замену. Рассмотренные варианты могут дополнительно или поочередно включать те, которые содержат predetermined мутации, например, путем гомологичной рекомбинации или сайт-направленного или ПЦР-мутагенеза и соответствующих полипептидов или нуклеиновых кислот других видов, включая, но без ограничений, описанные в настоящем документе аллели или другие природные варианты семейства полипептидов или нуклеиновых кислот, которые содержат вставку и замену; и/или производные, в которых полипептид ковалентно модифицирован путем замены, химическим, ферментативным или другим подходящим способом с остатком, отличным от природной аминокислоты, которая содержит вставку и замену (например, детекти-

руемый фрагмент, такой как фермент).

Термин "нативный", используемый в настоящем документе, для обозначения последовательностей нуклеиновых кислот или аминокислотных последовательностей, поскольку они естественно происходят в организме хозяина, в организме или вирусе. Термин "ненативный", используемый в настоящем документе, для обозначения последовательностей нуклеиновых кислот или аминокислотных последовательностей, которые не встречаются естественным образом у хозяина, в организме или вирусе. Последовательность нуклеиновой кислоты или аминокислотная последовательность, которая была удалена из клетки или вируса, подвергнута лабораторной манипуляции и введена или повторно введена в клетку-хозяин или вирус, считается "ненативной". Синтетические или частично синтетические гены, вводимые в клетку-хозяин или вирус, являются "ненативными". Ненативные гены дополнительно включают гены, эндогенные к вирусу, функционально связанному с одной или более гетерологичными регуляторными последовательностями, которые были рекомбинированы в геном хозяина.

"Рекомбинантная" или "сконструированная" молекула нуклеиновой кислоты представляет собой молекулу нуклеиновой кислоты, которая была изменена путем манипулирования, осуществленным человеком. В качестве неограничивающих примеров рекомбинантная молекула нуклеиновой кислоты включает любую молекулу нуклеиновой кислоты, которая: 1) была частично или полностью синтезирована или модифицирована *in vitro*, например с использованием химических или ферментативных методик (например, с использованием синтеза химической нуклеиновой кислоты или с использованием ферментов для репликации, полимеризации, расщепления (экзонуклеолитического или эндонуклеолитического) лигирования, обратной транскрипции, транскрипции, модификации оснований (включая, например, метилирование), интеграции или рекомбинации (включая гомологичную и сайт-специфическую рекомбинацию) молекул нуклеиновых кислот); 2) включает соединенные нуклеотидные последовательности, которые не соединены в природе; 3) была сконструирована с использованием методик молекулярного клонирования, так что она не имеет одного или более нуклеотидов по отношению к естественной последовательности молекулы нуклеиновой кислоты; и/или 4) была подвергнута манипулированию с использованием методик молекулярного клонирования, так что она имеет одно или более изменений последовательности или перегруппировок по отношению к естественной последовательности нуклеиновой кислоты. В качестве неограничивающих примеров кДНК представляет собой рекомбинантную молекулу ДНК, как и любая молекула нуклеиновой кислоты, которая была получена полимеразной реакцией(иями) *in vitro* полимеразы или к которой присоединены линкеры или которые были интегрированы в вектор, такой как вектор клонирования или вектор экспрессии.

Используемый в настоящем документе термин "рекомбинантный белок" относится к белку, полученному при помощи генной инженерии.

При применении к организмам или вирусам термин рекомбинантный, сконструированный или генетически сконструированный относится к организмам или вирусам, на которых воздействовали путем вставки гетерологичной или экзогенной (например, ненативной) рекомбинантной последовательности нуклеиновой кислоты в организм или вирус и включает, без ограничений, нокауты генов, целевые мутации и замену гена, замену, делецию или вставку промотора или перенос молекулы нуклеиновой кислоты, например трансгена, синтетического гена, промотора или другой последовательности в организм или вирус. Рекомбинантные или генетически сконструированные организмы или вирусы также могут быть организмами или вирусами, в которые были введены конструкции для "нокдауна" гена. Такие конструкции включают, но без ограничений, одну или более гидовых РНК, РНКи, микроРНК, кшРНК, миРНК, антисмысловых и рибозимных конструкций. Также включены организмы или вирусы, чьи геномы были изменены действием нуклеаз Cas, мегануклеаз или цинковопальцевых нуклеаз. Молекула экзогенной или рекомбинантной нуклеиновой кислоты может быть интегрирована в рекомбинантный/генетически сконструированный вирус или геном организма, или в других случаях не интегрирована в рекомбинантный/генетически сконструированный вирус или геном организма. Используемый в настоящем документе термин "рекомбинантный вирус" или "рекомбинантная клетка-хозяин" включает потомство или производные рекомбинантного вируса согласно настоящему раскрытию. Поскольку определенные изменения могут произойти в последующих поколениях из-за мутации или воздействия, или влияния окружающей среды, такое потомство или производные могут фактически не совпадать с родительской клеткой, но все еще включены в объем термина, как он используется в настоящем документе.

Используемый в настоящем документе термин "этап конструирования" относится к выполнению любого способа инженерии, раскрытого в настоящем документе или известного в данной области техники. Например, и "этап конструирования" может представлять собой один раунд интересующего инженерного способа, такого как, например, один раунд описанного в настоящем изобретении способа конструирования *in vitro*, один ПЦР-опосредованный мутагенез или одна реакция лигирования, объединяющая две части ДНК вместе. Аналогичным образом, "неоднократное повторение этапов конструирования" означает выполнение способа конструирования два или более последовательных раз.

Термин "гетерологичный" при использовании в отношении полинуклеотида, гена, нуклеиновой кислоты, полипептида или фермента относится к полинуклеотиду, гену, нуклеиновой кислоте, полипептиду или ферменту, который не является производным от вида хозяина. Например, "гетерологичный ген" или

"гетерологичная последовательность нуклеиновой кислоты", как используется в настоящем документе, относится к последовательности гена или нуклеиновой кислоты от разных видов, чем к видам организма-хозяина или вируса, в который он введен. Когда речь идет о регуляторной последовательности гена или к вспомогательной последовательности нуклеиновой кислоты, используемой для регуляции экспрессии генной последовательности (например, 5'-нетранслируемая область, 3'-нетранслируемая область, последовательность добавления поли-А, последовательность интрона, сайт сплайсинга, сайт связывания рибосом, внутренняя входная последовательность рибосом, область гомологии генома, сайт рекомбинации и т.д.) или последовательности нуклеиновой кислоты, кодирующей домен белка или последовательность локализации белка или белка, "гетерологичная" означает, что регуляторная или вспомогательная последовательность или последовательность, кодирующая белковый домен или последовательность локализации, из другого источника, чем ген, с которым регуляторная или вспомогательная последовательность нуклеиновой кислоты или последовательность нуклеиновой кислоты, кодирующая белковый домен или последовательность локализации, сопоставляются в геноме, хромосоме или эписоме. Таким образом, промотор, функционально связанный с геном, к которому он не функционально связан в своем естественном состоянии (например, в геноме генетически несконструированного организма или вируса), в настоящем документе упоминается как "гетерологичный промотор", даже хотя промотор может быть получен из одного и того же вида (или, в некоторых случаях, одного и того же организма или вируса) в качестве гена, с которым он связан. Аналогично, когда речь идет о последовательности локализации белка или белкового домена сконструированного белка, "гетерологичный" означает, что последовательность локализации или белковый домен получают из белка, отличного от белка, в который он включен при помощи методов генной инженерии.

"Регуляторная последовательность", "регуляторный элемент" или "последовательность регуляторных элементов" относится к нуклеотидной последовательности, расположенной выше (5'), в пределах или ниже (3') кодирующей последовательности. Транскрипция кодирующей последовательности и/или трансляция молекулы РНК в результате транскрипции кодирующей последовательности обычно зависит от наличия или отсутствия регуляторной последовательности. Данные последовательности регуляторных элементов могут содержать промоторы, цис-элементы, энхансеры, терминаторы или интроны. Регуляторные элементы могут быть выделены или идентифицированы из нетранслируемых областей (НТО) из конкретной полинуклеотидной последовательности. Любой из описанных в настоящем документе регуляторных элементов может присутствовать в химерном или гибридном регуляторном экспрессирующем элементе. Любой из описанных в настоящем документе регуляторных элементов может присутствовать в рекомбинантной конструкции согласно настоящему изобретению.

Термины "промотор", "область промотора" или "последовательность промотора" относятся к последовательности нуклеиновой кислоты, способной связывать РНК-полимеразу с инициацией транскрипции гена в направлении от 5' к 3' ("в прямом направлении"). Ген находится "под контролем" или "регулируется" промотором, в том случае, когда связывание РНК-полимеразы с промотором является непосредственной причиной транскрипции указанного гена. Промотор или промоторная область обычно обеспечивает сайт распознавания для РНК-полимеразы и другие факторы, необходимые для правильной инициации транскрипции. Промотор может быть выделен из 5'-нетранслируемой области (5' НТО) геномной копии гена. Альтернативно, промотор может быть синтетически получен или сконструирован путем изменения известных элементов ДНК. Также рассмотрены химерные промоторы, которые объединяют последовательности одного промотора с последовательностями другого промотора. Промоторы могут быть определены по их профилю экспрессии гена, основанной, например, на метаболических, экологических или связанных с развитием условиях. Промотор может быть использован как регуляторный элемент для модуляции экспрессии функционально связанной транскрибируемой полинуклеотидной молекулы, например, кодирующей последовательности. Промоторы могут содержать, помимо последовательностей, распознаваемых РНК-полимеразой и, предпочтительно, другие факторы транскрипции, элементы регуляторной последовательности, такие как цис-элементы или энхансерные домены, которые влияют на транскрипцию функционально связанных генов. "Вирусный промотор" является нативным или ненативным промотором, который инициирует транскрипцию одного или более генов, расположенных в вирусном геноме.

Используемый в настоящем документе термин "конститутивный" промотор относится к промотору, который активен в большинстве условий окружающей среды и развития. Конститутивный промотор активен независимо от окружающей среды, такой как освещение и композиция культуральной питательной среды. В некоторых примерах конститутивный промотор активен в присутствии и в отсутствие питательного вещества. Например, конститутивный промотор может быть промотором, который является активным (опосредует транскрипцию гена, с которым он функционально связан), в условиях истощения азота, а также в условиях, когда азот не ограничен (условия насыщения азотом). Напротив, "индуцируемый" промотор является промотором, который активен в ответ на конкретные условия окружающей среды, такие как наличие или отсутствие питательного вещества или регулятора, наличие света и т.д.

Используемый в настоящем документе термин "функционально связанный" обозначает конфигурацию, в которой контрольная последовательность помещается в соответствующее положение относительно



но кодирующей последовательности полинуклеотидной последовательности, так что контрольная последовательность направляет или регулирует экспрессию кодирующей последовательности полипептида и/или функциональной РНК). Таким образом, промотор находится в функциональной связи с последовательностью нуклеиновой кислоты, если он может опосредовать транскрипцию последовательности нуклеиновой кислоты. При введении в клетку-хозяина каскада экспрессии может приводить к транскрипции и/или трансляции закодированной РНК или полипептида в соответствующих условиях. Антисмысловые или смысловые конструкторы, которые не транслируются или не могут быть транслированы, не исключены данным определением. В случае как экспрессии трансгенов, так и подавления эндогенных генов (например, антисмысловыми или РНКи) любой специалист определит, что вставленная полинуклеотидная последовательность не обязательно должна быть идентичной, но может быть только по существу идентичной последовательности гена, из которого она была получена. Как поясняется в настоящем документе, эти по существу идентичные варианты конкретно покрыты ссылкой на конкретную последовательность нуклеиновой кислоты.

Термин "селективный маркер" или "селективный маркерный ген", используемый в настоящем документе, включает любой ген, который придает фенотип клетке, в которой он экспрессируется, для облегчения отбора клеток, которые трансфецированы или трансформированы конструкцией нуклеиновой кислоты согласно изобретению. Данный термин также может быть использован для обозначения генных продуктов, которые осуществляют указанные фенотипы. Неограничивающие примеры выбираемых маркеров включают: 1) гены, придающие устойчивость к антибиотикам, таким как амикацин (*aphA6*), ампициллин (*ampR*), бластицидин (*bis*, *bsr*, *bsd*), блеомицин или флеомицин (*ZEOCIN™*) (*ble*), хлорамфеникол (*cat*), эметин (*RBS14r* или *cry 1-1*), эритромицин (*ermE*), G418 (*GENETICIN™*) (*neo*), гентамицин (*aac3* или *aacC4*), гигромицин В (*aphIV*, *hph*, *hpt*), канамицин (*ncplI*), метотрексат (*DHFR mtxR*), пенициллин и другие β-лактамы (β-лактамазы), стрептомицин или спектиномицин (*aadA*, *spec/strep*) и тетрациклин (*tetA*, *tetM*, *tetQ*); 2) гены, придающие толерантность к гербицидам, таким как аминотриазол, амитрол, анримид, арилоксифеноксипропионаты, атразины, бипиридины, бромоксинил, циклогександиеноксима далапон, дикамба, дикфоп, дихлорфенилдиметилмочевина (ДКМ), дифунон, дикетонитрилы, диурон, флуридон, глюфозинат, глифосат, галогенированные гидробензонитрилы, галоксифоп, 4-гидрокси-пиридины, имидазолиноны, изоксасфлутол, изоксазолы, изоксазолидиноны, мирамид В, п-нитродифениловые эфиры, норфлуразон, оксадиазолы, м-феноксипропионамиды, N-фениламины, пиноксадин, ингибиторы протопорфириногенаксидазы, пиридазины, пиразолины, сульфонилмочевины, 1,2,4-триазолпиридин, трикетоны или мочевины; ацетил-КоА-карбоксилаза (АККаза); синтаза ацетогидроксикислоты (*ahas*); ацетолактатсинтаза (*a/s*, *csrl-1*, *csrl-2*, *imr1*, *imr2*), аминогликозид-фосфотрансфераза (*apt*), антрацилатсинтаза, бромоксинилнитрилаза (*bxn*), цитохром P450-НАДФ-цитохром P450-оксидоредуктаза, далапон-дегалогеназа (*dehal*), дигидроптероат-синтаза (*sul*), 5-енолпируватшикимат-3-фосфатсинтаза класса I (EPSPS), класс II EPSPS (*agoA*), не I/II классов EPSPS, глутатионредуктаза, глифосат-ацетилтрансфераза (*gat*), глифосат-оксидоредуктаза (*gox*), гидроксифенилпируватдегидрогеназа, гидроксифенилпируватдиоксигеназа (*hprp*), изопренилпирофосфат-изомераза, ликопенициклаза, фосфинотрицин-ацетилтрансфераза (*pat*, *bar*), фитоэндезатураза (*crtI*), пренилтрансфераза, протопорфириноксидаза, полипептид *psbA* фотосистемы II (*psbA*) и эстераза SMM (*SulE*) супероксиддисмутаза (*sod*); 3) гены, которые могут быть использованы в ауксотрофных штаммах или для придания других метаболических эффектов, таких как *arg7*, *his3*, *hisD*, *hisG*, *lysA*, *manA*, *metE*, *nit1*, *trpB*, *ura3*, *xyIA*, ген дигидрофолатредуктазы, манноза-6-фосфат-изомеразы, ген нитратредуктазы или ген орнитиндекарбоксилазы; фактор негативного отбора, такой как тимидинкиназа; или факторы устойчивости к токсину, такие как ген устойчивости к дезоксилюкозе.

"Репортерный ген" представляет собой ген, кодирующий белок, который обнаруживается или обладает активностью, которая продуцирует детектируемый продукт. Репортерный ген может кодировать визуальный маркер или фермент, который продуцирует детектируемый сигнал, такой как *cat*, *lacZ*, *uidA*, *xyIE*, ген щелочной фосфатазы, ген α-амилазы, ген α-галактозидазы, ген β-глюкуронидазы, ген β-лактамазы, гена пероксидазы хрена, ген люциферина/люциферазы, ген R-локуса, гена тирозиназы или ген, кодирующий флуоресцентный белок, включая, но без ограничений, синий, голубой, зеленый, красный, паприка или желтый флуоресцентный белок, фотоотверждаемый, фотографируемый или оптический флуоресцентный белок с подсветкой или любым из его вариантов, включая, без ограничений, оптимизированный по кодону, с быстрым сворачиванием, мономерный, с повышенной стабильностью и варианты с усиленной флуоресценцией.

Термин "РНК-направляемая нуклеаза" или "РНК-направляемая эндонуклеаза", используемый в настоящем документе, относится к ферменту, расщепляющему нуклеиновую кислоту, который нацелен на целевой сайт расщепления при помощи одной или более гидовых РНК. Неограничивающие примеры РНК-направляемых нуклеаз включают *Cas9*, *Cpf1*, *C2c1*, *C2c2* и *C2c3*.

Термин "терминатор" или "терминаторная последовательность" или "терминатор транскрипции", используемый в настоящем документе, относится к регуляторной части генетической последовательности, которая вызывает прекращение транскрипции РНК-полимеразой.

Термины "введение в клетку-хозяин" и "трансформация", используемые в настоящем документе, относятся к введению одной или более экзогенных последовательностей нуклеиновых кислот или "поли-нуклеотидов" в клетку-хозяин или организм с использованием одного или более физических, химических или биологических методов. Физические и химические методы трансформации (т.е. "трансфекция") включают в качестве неограничивающего примера электропорацию, бомбардировку частицами, химически-индуцированную компетентность и доставку липосом. Биологические методы трансформации (т.е. "трансдукция") включают передачу ДНК с использованием вирусов или микроорганизмов (например, *Agrobacterium*).

Используемый в настоящем документе термин "конструирование" генома относится к определению желаемой последовательности нуклеиновой кислоты конечного генома, представляющего интерес. Конструирование может быть основано на базовых знаниях, литературных источниках, экспериментальных данных или любой их комбинации.

Используемый в настоящем документе термин "рекомбинантный" или "сконструированный", когда речь идет о молекуле нуклеиновой кислоты, белке, вирусной частице или их комбинации, означает не имеющую природное происхождение молекулу нуклеиновой кислоты, белок, вирусную частицу или их комбинацию, полученную посредством манипуляции, осуществляемой человеком. В качестве неограничивающих примеров рекомбинантная или сконструированная молекула нуклеиновой кислоты включает любую молекулу нуклеиновой кислоты, которая: 1) была частично или полностью синтезирована или модифицирована *in vitro*, например, с использованием химических или ферментативных методик (например, с использованием синтеза химической нуклеиновой кислоты или с использованием ферментов для репликации, полимеризации, расщепления (экзонуклеолитического или эндонуклеолитического) лигирования, обратной транскрипции, транскрипции, модификации оснований (включая, например, метилирование), интеграции или рекомбинации (включая гомологичную и сайт-специфическую рекомбинацию) молекул нуклеиновых кислот); 2) включает соединенные нуклеотидные последовательности, которые не соединены в природе, 3) была сконструирована с использованием методик молекулярного клонирования, так что она не имеет одного или более нуклеотидов по отношению к естественной последовательности молекулы нуклеиновой кислоты; и/или 4) была подвергнута воздействию с использованием методик молекулярного клонирования, так что она имеет одно или более изменений последовательности или перегруппировок по отношению к естественной последовательности нуклеиновой кислоты. В качестве неограничивающих примеров кДНК представляет собой рекомбинантную молекулу ДНК, как и любая молекула нуклеиновой кислоты, которая была получена при помощи полимеразной реакции(ий) *in vitro* или к которой присоединены линкеры или которые были интегрированы в вектор, такой как вектор клонирования или вектор экспрессии. Рекомбинантная или сконструированная РНК или белок представляет собой рекомбинантную или сконструированную РНК или белок, которая транскрибируется или транслируется соответственно из рекомбинантной или сконструированной молекулы нуклеиновой кислоты. Рекомбинантная или сконструированная вирусная частица или вирус получают из сконструированной вирусной последовательности или вирусного генома.

Термин "вирусный геном" относится к полному генетическому комплексу, содержащемуся в одной или более молекулах ДНК или РНК в вирусной частице, включая гены и некодирующие последовательности. Термин "сконструированный вирусный геном" относится к не встречающемуся в природе вирусному геному, который является результатом манипуляции, осуществляемой человеком, и способен продуцировать вирусные частицы, не встречающиеся в природе, после введения в совместимую клетку-хозяина.

Термин "вирусная нуклеиновая кислота" относится к нуклеиновой кислоте, содержащей последовательность, полученную из вирусного генома. "Вирусная нуклеиновая кислота" может содержать целый вирусный геном или часть вирусного генома. Вирусные нуклеиновые кислоты могут кодировать аминокислотные последовательности, содержащие вирусные белки. В некоторых случаях нативные, зрелые белки или полипептидные последовательности, кодируемые данной открытой рамкой считывания, не могут быть определены или охарактеризованы. Аминокислотные последовательности, представленные в настоящем документе, которые кодируются последовательностями вирусной нуклеиновой кислоты, которые могут содержать сайт, подходящий для мутации (такой как изменение, делеция или замена) или вставку гетерологичных последовательностей, могут быть раскрыты в настоящем документе как кодирующие аминокислотные последовательности, которые могут содержать весь или часть вирусного полипептида или белка.

Термины "вирусная частица" и "вирион" относятся к независимой форме, в которой существует вирус, когда он не находится внутри инфицированной клетки или в процессе заражения клетки. Данные вирусные частицы (вирионы) состоят или из генома ДНК, или из РНК, окруженного белковой оболочкой, называемой капсидом. Некоторые вирионы также имеют дополнительную липидную оболочку, как внутри, так и снаружи оболочки капсидного белка. Термины "вирусная частица", "вирион" и "вирус" могут использоваться взаимозаменяемо.

Используемый в настоящем документе термин "вирусное свойство" относится к любому аспекту репликации вируса или жизненного цикла или к аспекту, который возникает в результате репликации

вируса или жизненного цикла. Используемый в настоящем документе термин "вирусное свойство" часто относится к свойствам, которые могут быть изменены или сконструированы посредством вмешательства человека для достижения желаемого результата. Неограничивающие примеры вирусных свойств включают специфичность к кругу хозяев, вирусный литический цикл, адсорбцию, прикрепление, инъекцию, репликацию и сборку, лизис, величину выхода фага, ускользание от иммунологического надзора, иммунную стимуляцию, иммунную дезактивацию, дисперсию биопленки, устойчивость к бактериальному фагу, антибиотической сенсбилизации бактерий, модуляции факторов вирулентности и направленного расщепления или редактирования генома хозяина. В некоторых аспектах улучшенное свойство или улучшенные свойства и улучшенное вирусное свойство или улучшенные вирусные свойства используются взаимозаменяемо.

Термины "бактериофаг" и "фаг" могут использоваться взаимозаменяемо и относятся к вирусу, который заражает бактерии.

Системы CRISPR.

CRISPR (короткие палиндромные повторы, регулярно расположенные группами) представляют собой локусы ДНК, содержащие короткие повторы последовательностей оснований. За каждым повтором следуют короткие сегменты "спейсерной ДНК" от предыдущих воздействий на мобильные генетические элементы. CRISPR обнаруживают в около 40% секвенированных геномов бактерий и 90% секвенированных архей. CRISPR часто ассоциируются с CRISPR-ассоциированными (cas) генами, которые кодируют белки, связанные с функцией CRISPR. Система CRISPR-Cas представляет собой прокариотическую иммунную систему, которая обеспечивает устойчивость к чужеродным генетическим элементам, таким как плазмиды и фаги, и обеспечивает форму приобретенного иммунитета. CRISPR-спейсеры кодируют малые стРНК последовательность которых специфически ориентирует эндонуклеазы Cas на целевые последовательности и разрезает эти экзогенные генетические элементы способом, аналогичным РНКи в эукариотических организмах.

Системы типа II CRISPR-Cas использовали для редактирования генов и регуляции генов у многих видов. Эти системы особенно полезны, потому что они требуют только одну эндонуклеазу Cas (Cas9) и нацеливающую стРНК. В природных системах эндонуклеаза Cas9 требует две независимо транскрибируемые РНК для активности, однако эти две РНК также могут быть ковалентно связаны с образованием одной химерной гРНК. Обеспечивают белком Cas9 и соответствующими гРНК клетку, геном организма можно разрезать в любом необходимом месте. Системы CRISPR-Cas представляют собой РНК-направляемую систему защиты, которая защищает от вирусов, плазмид и других мобильных генетических элементов. Данный защитный путь имеет три этапа. Во-первых, копия встраиваемой нуклеиновой кислоты интегрируется в массив CRISPR. Затем массив CRISPR транскрибируется в большой транскрипт CRISPR и затем процессируется в зрелые стРНК. Затем стРНК вводят в эффекторные комплексы, причем стРНК направляет комплекс к встраиваемой нуклеиновой кислоте, а белки Cas деградируют данную нуклеиновую кислоту. Как указано выше, для нативной системы CRISPR-Cas II типа требуется как транс-активирующая стРНК (tracrРНК), так и пре-стРНК для активации Cas9. TracrРНК является комплементарной к и комплементарной парами оснований с пре-стРНК, образующей дуплекс РНК. Он расщепляется РНКазой III, РНК-специфической рибонуклеазой, с образованием гибрида стРНК/tracrРНК. Этот гибрид выступает в качестве направляющего для эндонуклеазы Cas9, которая расщепляет встраиваемую нуклеиновую кислоту, образующую двухцепочечный разрыв в встраиваемой ДНК для защиты клетки-хозяина. Cas9-опосредованное расщепление строго зависит от наличия мотива, смежного с протоспейсером, (PAM) в целевой нуклеиновой кислоте. Возможность программировать Cas9 для расщепления определенных сайтов, определяемых гидовыми РНК, привела к принятию его в качестве универсальной платформы для конструирования генома и геномной регуляции. Данный способ конструирования генома описан в публикации заявки на патент США № 2014/0068797, опубликованной 6 марта 2014 года, 2014/0170753, опубликованной 19 июня 2014 года и 2014/0273037 и 2014/0273226, обе из которых опубликованы 18 сентября 2014 года, все из которых включены в качестве ссылки.

Были описаны другие программируемые системы CRISPR-Cas, которые могут быть использованы для геномной инженерии, включая системы Cpf1, C2c1, C2c2 и C2c3. Система Cpf1 представляет собой систему CRISPR V типа и опосредует расщепление ДНК с липкими концами с помощью одной РНК, ориентированной на нацеливание (Zetsche et al., Cell (2015) 163, 1-13) (включен посредством ссылки). C2c1 и C2c3 являются одновременно системами CRISPR V типа, тогда как C2c2 предлагается быть системой CRISPR VI типа (Shmakov et al., Molecular Cell (2015) 60, 1-13) (включен в качестве ссылки).

Сборка ДНК.

Существуют различные методы, известные в данной области техники для сборки ДНК во время геномной инженерии. Двухстадийный метод на основе термоциклов использовали для сборки частей гена M.genitalium, как описано в Gibson D.G. et al., "Complete chemical synthesis, assembly, and cloning of a Mycoplasma genitalium genome". Science (2008) 319: 1215-1220 (включена в качестве ссылки) и публикации PCT WO 2009/103027 (включены в качестве посредством ссылки). Другой подход описан Li M.Z. et al., Nature Meth. (2007) 4: 251-256 (включен посредством ссылки). Одностадийный способ сборки с исполь-

зованием 5' экзонуклеазы T7 и одноцепочечного ДНК-связывающего белка раскрыт в публикации РСТ WO 2006/021944 (включена посредством ссылки). В настоящее время хорошо известны комбинаторные методики сборки химических соединений для использования при высокопроизводительном скрининге. Кроме того, в течение ряда лет практикуют методики перегасовки генов, в которых кодирующие последовательности случайным образом фрагментируют и повторно отжигают. Например, протоколы для создания библиотек фрагментов химерных генов описаны в Meyer M. и др. "Combinatorial Recombination of Gene Fragments to Construct a Library of Chimeras" Current Protocols in Protein Science (2006) 26.2.1-26.2.17; McKee A.E. et al., JBEI abstract. Методики сборки различных компонентов в полные или минимальные геномы были установлены. Например, публикация патента США 2000/0264688 (включена посредством ссылки), опубликованная 15 ноября 2007 г., описывает способы конструирования синтетического генома путем создания и сборки кассет, содержащих части генома. Поэтапный иерархический способ сборки нуклеиновых кислот описан в публикации патента США № 2007/004041 (включен посредством ссылки), опубликованной 4 января 2007 года.

Кроме того, метод одного сосуда для сборки ДНК описан в публикации заявки на патент США № 2010/0035768 и 2012/0053087, опубликованы 11 февраля 2010 года и 1 марта 2012 года соответственно, обе из которых включены посредством ссылки. Данный метод был назван методом сборки Gibson Assembly и позволяет успешно собрать несколько фрагментов ДНК, независимо от длины фрагмента или конечной совместимости. Реакция сборки Gibson Assembly проводится в одном сосуде в изотермических условиях с использованием трех ферментативных активностей: 5'-экзонуклеаза образует длинные липкие концы, полимеразы заполняют промежутки отожденных одноцепочечных областей, а ДНК-лигаза лигирует разрывы отжигом и заполнением промежутков. Данный метод был широко принят и является основной рабочей лошадкой проектов в области синтетической биологии во всем мире. Применяя эту методологию, митохондриальный геном мыши размером 16,3 т.п.н. был собран из 600 перекрывающихся 60 мономеров. В комбинации со сборкой *in vivo* у дрожжей, метод сборки Gibson Assembly использовали для синтеза 1,1 млн пар оснований генома *Mycoplasma mycoides*. Синтезированный геном трансплантировали в клетку-реципиент *M. carficolum*, создавая новые самовоспроизводящиеся клетки *M. mycoides*. 5'-экзонуклеазная активность возвращает обратно 5'-концевые последовательности и раскрывает комплементарную последовательность для отжига. Затем полимерная активность заполняет промежутки в отоженных областях. ДНК-лигаза затем заполняет одноцепочечные разрывы и ковалентно связывает фрагменты ДНК вместе. Перекрывающаяся последовательность смежных фрагментов намного длиннее, чем те, которые используются в сборке Golden Gate, и, следовательно, приводит к увеличению процента правильных сборок.

#### Вирусы.

Вирус представляет собой ультрамикроскопический и метаболически инертный инфекционный агент, который реплицируется только внутри клеток живых хозяев. Вирусы могут заражать все типы жизненных форм, включая животных, растения, грибы, водоросли, бактерии и археи. Хотя они не находятся внутри инфицированной клетки или в процессе заражения клетки, вирусы существуют в виде независимых частиц. Данные вирусные частицы (вирионы) состоят или из генома ДНК, или из РНК, окруженного белковой оболочкой, называемой капсидом. Некоторые вирионы также имеют дополнительную липидную оболочку, как внутри, так и снаружи оболочки капсидного белка.

Существует два цикла вирусной репликации, однако терминология варьирует между прокариотическими и эукариотическими вирусными полями. Скрытые или лизогенные вирусы объединяют вирусный генетический материал в геном клетки-хозяина или образуют эписомальный репликон. Когда клетка-хозяин реплицируется, вирусный генетический материал также копируется и продолжает разделяться с геномом хозяина до начала вирусного продуцирования. Начало вирусного продуцирования и гибель клеток являются маркерами литического или вирулентного цикла. Во время литического цикла вирусный геном реплицируется отдельно от гена хозяина и захватывает механизмы репликации и трансляции клеток, чтобы создавать больше вирусов. После накопления достаточного количества вирусов специализированные вирусные белки растворяют клеточную стенку и/или мембрану хозяина. Клетка-хозяин лопается из-за высокого внутреннего осмотического давления, этот процесс называется лизисом. Это высвобождает потомство вирусов в среду, где они могут заражать другие клетки и повторять процесс. Вирулентные вирусы - это те, которые не входят в латентное или лизогенное состояние, но вместо этого реплицируются только путем захвата механизма клеток-хозяев (в отличие от умеренных вирусов, которые входят в латентное состояние).

#### Исследования вирусной мутации.

Вирусные мутационные исследования, используемые в настоящем документе, относятся к исследованиям быстрой эволюции, адаптации и/или случайному или направленному мутагенезу, и данные термини могут использоваться взаимозаменяемо.

Исследования в области эволюции и/или адаптации включают выбор вирусов по определенным признакам или при определенных условиях. Данные методы особенно полезны для вирусов из-за естественной высокой скорости мутаций, присущей вирусной репликации, что приводит к большому количеству вирусного разнообразия. Например, штаммы можно выращивать в условиях высокой температуры,

чтобы наблюдать молекулярные изменения, которые способствуют выживанию и размножению в этих условиях. В качестве неограничивающих примеров можно выделить эволюцию или адаптацию эксперимента вируса или бактериофага для выбора вариантов с изменением круга хозяев, вирусного литического цикла, адсорбции, прикрепления, инъекции, репликации и сборки, лизиса, величины выхода фага, ускользания от иммунологического надзора, иммунной стимуляции, иммунной дезактивации, дисперсии биопленки, устойчивости бактерий к фагу, антибиотической сенсibilизации бактерий, модуляцию факторов вирулентности, или направленного расщепления или редактирования гена хозяина. Неограничивающие примеры экспериментов по эволюции вируса или адаптации включают эксперименты по коинфекции, коэволюции или контрансформации. Коинфекция относится более чем к одному вирусу, заражающему один и тот же хозяин одновременно, что часто приводит к обмену генами между двумя или более вирусами. Коэволюция относится к исследованию, в котором рекомбинация между двумя или более вирусами или бактериофагами происходит в перmissive или неpermissive хозяине, что приводит к сборке нового вируса или бактериофага с различными вирусными свойствами, такими как, например, более широким кругом хозяев. Котрансформация относится к тому, когда два "оголенных" генома трансформируются вместе в перmissive или неpermissive штамме. Любое из этих исследований эволюции или адаптации может быть выполнено в перmissive (восприимчивом) или неpermissive (резистентном) хозяине. Данные типы экспериментов часто включают передачу вируса несколько раз выбранному хозяину в отсутствие или присутствии одного, или более других выбранных вирусов. Вирусы будут получать мутации, которые приводят к множественным вариантам. На протяжении всего пассирования некоторые варианты будут обогащаться в зависимости от условий пассирования и отбора.

Мутагенез может быть выполнен любым способом, например инсерционным мутагенезом, химическим мутагенезом,  $\gamma$ -облучением или ультрафиолетовым облучением, или ПЦР-опосредуемым мутагенезом. Способы получения мутантов или варианты геномных последовательностей хорошо известны. Например,  $\gamma$ -облучение, УФ-облучение и обработка любым из большого числа возможных химических мутагенов (например, 5-бромдезоксипуридин, этилметансульфонат (EMS), метилметансульфонат (MMS), диэтилсульфат (DES), нитрозогуанидин (NTG), соединения ICR и т.д.) или лечение соединениями, такими как эндеиновые антибиотики, которые вызывают разломы хромосом (например, блеомицин, адриамицин, неокарзиностин) представляют собой методы, которые применяют для мутагенеза водорослей, грибов и хитридиомицетов (см., например, патент США 8232090, заявка на патент США 20120088831, заявка на патент США 20100285557, заявка на патент США 20120258498). В данной области техники известно большое количество химических мутагенов, включая, но без ограничений, интеркалирующие агенты, алкилирующие агенты, дезаминирующие агенты, аналоги оснований. Интеркалирующие агенты включают в качестве неограничивающих примеров производные акридина или производные фенантридина, такие как бромид этидия (также известный как 2,7-диамино-10-этил-6-фенилфенантридбромид или 3,8-диамино-5-этил-6-фенилфенантридинбромид). Неограничивающие примеры алкилирующих агентов включают производные нитрозогуанидина (например, N-метил-N'-нитронитрозогуанидин), этилметансульфонат (EMS), этилэтансульфонат, диэтилсульфат (DES), метилметансульфонат (MMS), азотистую кислоту или  $\text{HNO}_2$ , и азотистый иприт или соединения ICR. Неограничивающие примеры аналогов оснований, которые могут быть использованы в качестве мутагенов, включают соединения 5-бромурацил (также известный как дезоксинуклеозид 5-бромдезоксипуридин), 5-бромдезоксипуридин и 2-аминопурин. Методы мутагенеза на основе ПЦР хорошо известны в данной области техники и часто включают условия реакции и/или ДНК-полимеразу, которая увеличивает частоту ошибок в течение ПЦР-амплификации.

Мутагенез может дополнительно или альтернативно включать введение экзогенных молекул нуклеиновой кислоты непосредственно в вирусный геном или в клетку-хозяина для последующей рекомбинации в представляющий интерес вирусный геном. Например, экзогенная молекула нуклеиновой кислоты, введенная в клетку-хозяина, может интегрироваться в вирусный генетический локус путем случайной или целевой интеграции, влияя на экспрессию генов, в которые вставляются чужеродные ДНК-вставки или гены, которые являются проксимальными по отношению к чужеродной ДНК, вставленной в геном (например, патент США 7019122; патент США 8216844). Как правило, введенная молекула нуклеиновой кислоты содержит селективируемый маркерный ген для отбора трансформантов, которые интегрировали конструктор молекулы экзогенной нуклеиновой кислоты. Молекула экзогенной нуклеиновой кислоты в некоторых вариантах реализации изобретения может включать транспонируемый элемент или его компонент, такой как, например, инвертированные повторы, которые могут быть распознаны транспозазой и/или геном, кодирующим транспозазу, или молекула экзогенной нуклеиновой кислоты может быть основана, по меньшей мере, частично на вирусе, таком как интегрирующий вирус.

Для случайного инсерционного мутагенеза конструкция предпочтительно содержит селективируемый маркер, который может быть использован для выбора трансформантов, имеющих интегрированный конструктор, и необязательно также может служить маркером сегрегации и молекулярной меткой для выделения и идентификации гена, прерванного интегрированным селективируемым маркерным геном. Селективные маркеры не ограничиваются генами устойчивости к антибиотикам, но также включают любой ген,

который может обеспечить преимущество роста вирусу (оба гена установленной и гипотетической функции). Альтернативно, специфический генетический локус может быть нацелен. Конструкция для разрушения гена может включать, например, селектируемый маркерный ген, фланкированный последовательностями из интересующего генетического локуса, например по меньшей мере часть гена, который кодирует регулятор, и, необязательно, дополнительные геномные последовательности, окружающие ген. Такие фланкирующие последовательности могут содержать, например, по меньшей мере 50 нуклеотидов, по меньшей мере 100 нуклеотидов, по меньшей мере 500 нуклеотидов или по меньшей мере 1 тысяча пар нуклеотидов геномной последовательности.

Сбор вирусных вариантов может быть создан любым из вышеперечисленных способов, другими способами, хорошо известными в данной области техники или любой их комбинацией. Затем набор вариантов можно скринировать для обеспечения необходимого фенотипа. Выделенные вирусы с желаемым фенотипом(ами) могут быть подвергнуты дополнительным этапам исследований мутаций. Выделенные вирусы, демонстрирующие желаемые свойства или фенотипы, могут быть дополнительно или альтернативно секвенированы для идентификации генетической мутации, ответственной за желаемое свойство или фенотип. Эти идентифицированные генетические поражения могут быть подтверждены повторением мутации в чистом эталонном фоне и тестированием желаемого свойства или фенотипа.

Вирусные полезные нагрузки.

Литические ферменты.

"Литический фермент" включает любой литический фермент бактериальной клеточной стенки, который убивает одну или более бактерий в подходящих условиях и в течение соответствующего периода времени. Примеры литических ферментов включают, без ограничений, различные амидазы клеточной стенки. Литическим ферментом может быть литический фермент бактериофага, который относится к литическому ферменту, экстрагированному или выделенному из бактериофага или синтезированного литического фермента с аналогичной структурой белка, которая поддерживает литическую функциональность фермента.

Литический фермент способен специфически расщеплять связи, которые присутствуют в пептидогликане бактериальных клеток, чтобы разрушить клеточную стенку бактерий. В настоящее время также постулируется, что пептидогликан бактериальной клеточной стенки очень консервативен среди большинства бактерий, и расщепление только нескольких связей может нарушить клеточную стенку бактерий. Примерами литических ферментов, которые расщепляют эти связи, представляют собой мурамидазы, глюкозаминидазы, эндопептидазы или N-ацетил-мурамоил-L-аланиновые амидазы. Fischetti et al. (1974) сообщили, что фермент лизин C1 стрептококкового бактериофага представляет собой амидазу. Garcia et al. (1987, 1990) сообщили, что лизин Cpl из *S. pneumoniae* из фага Cp-1 представляет собой лизоцим. Caldentey and Bamford (1992) сообщили, что литический фермент из фага ф6 *Pseudomonas* представляет собой эндопептидазу, разделяющей пептидный мост, образованный меламинампилиновой кислотой и D-аланином. T1 и T6 литические ферменты фага *E. coli* представляют собой амидазы, как и литический фермент из фага *Listeria* (ply) (Loessner et al., 1996). Существуют также другие литические ферменты, известные в данной области техники, которые способны расщеплять клеточную стенку бактерий.

Литический фермент, генетически кодируемый бактериофагом, содержит полипептид, способный убивать бактерию-хозяина, например, путем, по меньшей мере, разрушения клеточной стенки или ингибирования активности клеточной стенки против бактерий-хозяев. Полипептид может иметь последовательность, которая охватывает нативные литические ферменты и их варианты. Полипептид может быть выделен из множества источников, например из бактериофага ("фага"), или получен путем рекомбинантных или синтетических методов. Полипептид может, например, содержать холинсвязывающую часть на карбоксильной концевой стороне и может быть охарактеризован ферментативной активностью, способной расщеплять пептидогликан клеточной стенки (такой как активность амидазы действовать на амидные связи в пептидогликане) на аминоконцевой стороне. Были описаны литические ферменты, которые включают множественные активности ферментов, например два ферментативных домена, такие как лизин PlyGBS. Кроме того, описаны другие литические ферменты, содержащие только каталитический домен и домен, не связывающий клеточную стенку.

Полипептиды для подавления кворума.

Автоиндукторы представляют собой небольшие химические сигнальные молекулы, продуцируемые и используемые бактериями, участвующими в определении кворума. Чувство кворума позволяет бактериям ощущать друг друга посредством наличия аутоиндукторов и регулировать самые разнообразные поведения на уровне группы. Такое поведение включает симбиоз, вирулентность, подвижность, продуцирование антибиотиков и образование биопленки. Автоиндукторы входят в различные химические формы в зависимости от вида, но во многих случаях эффект, который у них есть, во многом аналогичен, что позволяет генетически сконструированным бактериофагам воздействовать на самые разнообразные бактерии, использующие подобные автоиндукторы. В целом, грамотрицательные бактерии используют AHL в качестве аутоиндукторов, а грамположительные бактерии используют процессированные олигопептиды для коммуникации, в то время как автоиндуктор 2 (AI-2) является универсальным для грамотрицательных и грамположительных бактерий.

АНЛ, продуцируемые различными видами граммотрицательных бактерий, различаются по длине и составу ацильной боковой цепи, которая часто содержит от 4 до 20 атомов углерода. АНЛ способны диффундировать в и из клеток как с помощью пассивного транспорта, так и с помощью активных механизмов транспорта. Рецепторы для определения АНЛ включают ряд регуляторов транскрипции, таких как LuxR, которые функционируют как ДНК-связывающие факторы транскрипции, которые могут активировать разнообразную экспрессию генов, регулиующую поведение популяции бактерий.

Автоиндукторы могут быть ингибированы полипептидами, подавляющими кворум. Полипептиды для подавления кворума могут модифицировать или деградировать автоиндукторы, чтобы сделать их менее активными или неактивными. Определенные полипептиды для подавления кворума представляют собой ферменты, которые инактивируют автоиндуктор (например, путем модификации или деградации), такой как описанный в настоящем документе белок лактоназы AiiA, который расщепляет лактонные кольца из ацильных фрагментов АНЛ с широкой субстратной специфичностью для инактивации АНЛ из различных бактерий (Wang et al. (2004) *J. Biol. Chem.* 279 (14): 136.45-51).

Предлагаемый в настоящем документе способ конструирования *in vitro* может быть использован для получения синтетического бактериофага, сконструированного для кодирования, например, полипептида, подавляющего кворум, полученного из *Pseudomonas aeruginosa*. Полипептиды, подавляющие кворум, могут быть экспрессированы в виде свободных белков, которые высвобождаются в область, окружающую фаг и/или бактерии, например при фаговой инфекции и лизиса бактерий-хозяев. Не исключено также, что полипептиды, подавляющие кворум, также могут быть экспрессированы и активно секретированы из клетки бактериального хозяина с использованием способов, известных в данной области техники. Аналогичным образом, полипептиды, подавляющие кворум, могут быть трансляционно слиты с белком бактериофага, например капсидным, хвостовым или шейным белком.

Хвостовые волокна.

Раскрытие изобретения предусматривает в некоторых вариантах реализации изменение круга хозяев бактериофага при помощи сконструированного рекомбинантного бактериофага. В некоторых вариантах реализации изобретения изменение круга хозяев вируса включает конструирование вируса с гетерологичными, нативными, ненативными хвостовыми волокнами и любой их комбинацией. Специфичность клеток-хозяев бактериофага может зависеть от хвостовых волокон(а) вирусной частицы. Путем изменения (например, замены и/или мутации) хвостовых волокон или частей хвостовых волокон бактериофага-хозяина круг хозяев может быть изменен (например, расширен).

Хвостовые волоконные белки обычно содержат детерминанты антигенности и детерминанты круга хозяев. Гетерологичное хвостовое волокно может быть закодировано набором геномных фрагментов, выделенных или синтезированных на основе генома одного типа бактериофага. Набор фрагментов гена хвостового волокна может содержать подмножества геномных фрагментов, выделенных или полученных на основе геномов нескольких бактериофагов. Например, консервативные области хвостового волокна могут быть закодированы геномными фрагментами, выделенными из генома шасси бактериофага, в то время как области детерминант круга хозяев могут быть закодированы геномными фрагментами, выделенными из генома другого типа бактериофага.

Антимикробные пептиды.

Раскрытие изобретения рассматривает в качестве неограничивающего примера бактериофаг, сконструированный для экспрессии антимикробного пептида, который необязательно секретируется клеткой-хозяином. Например, сконструированные бактериофаги могут экспрессировать антимикробный агент, такой как антимикробный пептид (АМР) или антимикробный полипептид, включая, но без ограничений, естественные пептиды для предотвращения развития и/или распространения устойчивости бактерий-хозяев к бактериофагу и для обеспечения более быстрого и эффективного уничтожения бактерий при бактериальных инфекциях, таких как бактериальные инфекции, включающие более одного вида бактерий.

Бактериофаги обеспечивают аттрактивный противомикробный агент для устранения бактериальных инфекций из-за их усиления и механизма хозяина-хищника, например, путем распространения в бактериях-хозяевах, а затем уничтожения бактерий в форме лизиса, что приводит к высвобождению размноженных бактериофагов, которые впоследствии заражают и убивают окружающие бактерии тем же механизмом. Практическое использование бактериофага в устранении бактериальных инфекций обусловлено значительными ограничениями, такими как (i) очень узкий круг бактерий-хозяев как внутри, так и между видами, и (ii) очень быстрое развитие устойчивости к бактериофагу популяцией бактерии-хозяина. Таким образом, как это часто бывает во многих областях науки, теоретический результат трудно достичь в реальных жизненных ситуациях. Поэтому, в то время как бактериофаги кажутся полезными в качестве антимикробных агентов в теории, на практике они сдерживают антимикробные свойства, и их использование для устранения бактериальных инфекций очень трудно достичь из-за быстрого развития устойчивости хозяина к бактериофагу. Следовательно, бактериофаги были неэффективны при длительной элиминации бактерий-хозяев.

Соответственно, настоящее изобретение относится к сконструированному бактериофагу, кодирующему антимикробный агент, причем бактериофаг модифицирован или сконструирован для экспрессии антимикробного пептида (АМР), который необязательно секретируется клеткой-хозяином, по меньшей

мере одну или любую комбинацию различных антимикробных агентов сконструированного бактериофага могут использовать отдельно или в любой комбинации для устранения или уничтожения бактериальной инфекции. В некоторых вариантах реализации изобретения сконструированный бактериофаг, кодирующий антимикробный агент, может быть использован с дополнительными агентами, такими как другой антимикробный агент сконструированного бактериофага, очищенный антимикробный пептид(ы) или низкомолекулярный антибиотик. Сконструированные бактериофаги, кодирующие антимикробные пептиды (или сконструированные бактериофаги, кодирующие AMP), могут кодировать любой антимикробный агент, известный специалисту в данной области техники.

В некоторых вариантах реализации аспектов изобретения сконструированный бактериофаг, кодирующий антимикробный агент, может экспрессировать и секретировать антимикробный агент, который представляет собой нуклеиновую кислоту, например антимикробный агент, который действует посредством "сайлесинга генов" общеизвестных бактериальных генов, известных обычным специалистам в данной области техники. Антимикробный агент на основе нуклеиновой кислоты включает, например, но без ограничений, молекулы, индуцирующие РНК-интерференцию (РНКи), например, но без ограничений, миРНК, дцРНК, мвРНК, кшРНК, микроРНК, и их сконструированные версии, где ген молекулы РНК-интерференции подавляет экспрессию гена, экспрессированного и важного для жизнеспособности (т.е. выживания) бактерий. Антимикробный агент на основе нуклеиновой кислоты может представлять собой антисмысловую олигонуклеиновую кислоту или аналог нуклеиновой кислоты, например, но без ограничений, ДНК, РНК, пептид-нуклеиновую кислоту (ПНК), псевдокомплементарную ПНК (пк-ПНК), или закрытую нуклеиновую кислоту (ЗНК) и тому подобное. Альтернативно, антимикробным агентом на основе нуклеиновой кислоты может быть ДНК или РНК и аналоги нуклеиновых кислот, например ПНК, пкПНК и ЗНК. Нуклеиновая кислота может быть одно- или двухцепочечной и может быть выбрана из группы, включающей нуклеиновую кислоту, кодирующую представляющий интерес белок, олигонуклеотиды, ПНК и т.д. Такие ингибиторы нуклеиновой кислоты включают, например, последовательность нуклеиновой кислоты, кодирующую белок, который является транскрипционным репрессором, или антисмысловую молекулу, или рибозим, или небольшую ингибирующую последовательность нуклеиновой кислоты, такую как РНКи, кшРНКи, миРНК, микроРНКи (микроРНК), антисмысловый олигонуклеотид и т.д.

Антимикробные пептиды могут дополнительно или альтернативно быть антибактериальными ферментами. Иллюстративные антибактериальные активности могут включать, но без ограничений, литический фермент, ацилазу, аминопептидазу, амилазу, карбогидразу, карбоксипептидазу, каталазу, целлюлазу, хитиназу, кутиназу, циклодекстрингликозилтрансферазу, дезоксирибонуклеазу, эстеразу, альфа-галактозидазу,  $\beta$ -галактозидазу, глюкоамилазу,  $\alpha$ -глюкозидазу,  $\beta$ -глюкозидазу, галопероксидазу, инвертазу, лакказу, липазу, маннозидазу, оксидазу, пектинолитический фермент, пептидоглутаминазу, пероксидазу, фитазу, полифенолоксидазу, протеолитический фермент, рибонуклеазу, транслутаминазу, ксиланазу, РНКазу, ДНКазу, лизостафин или порообразующие пептиды.

Антимикробные пептиды или антимикробные полипептиды могут непосредственно разрушать бактериальную мембрану путем связывания с отрицательно заряженной микробной мембраной и разрушения мембраны путем образования водных каналов, заставляя липидный бислой сворачиваться назад на себя или покрывая мембрану с образованием мицелл. В дополнение к их прямым бактерицидным эффектам любые антимикробные пептиды и полипептиды могут также активировать TLR-сигналинг и дополнительные иммунные ответы, служить в качестве хемоаттрактантов лейкоцитов, увеличивать бактерицидную опсонизацию путем вторжения фагоцитов, очищать жизненно важные питательные вещества, необходимые бактериям для роста и ингибировать бактериальные протеазы, или любую комбинацию его.

Биосурфактанты.

Образование бактериальной биопленки может приводить к локальным инфекциям, а также к тяжелому лечению, а иногда и к смертельным, системным инфекциям, таким как бактериемия (наличие бактерий в крови) и бактериальный сепсис (множественная органная недостаточность, вызванная распространением бактерий или их продуктов через кровотоки). Внеклеточные вещества, которые составляют матрицу биопленки, могут действовать как барьер, который защищает и изолирует бактерии, находящиеся в биопленке от нормальных иммунологических защитных механизмов, таких как антитела и фагоциты, а также от антимикробных агентов, включая антибактериальные ферменты и антибиотики. Биопленка также способствует росту и распространению бактерий, обитающих в биопленке.

Изобретение относится к способам получения и композициям сконструированных вирусов, экспрессирующих дополнительный агент, используемый для облегчения удаления или ослабления биопленки, нанесенной на поверхность. Например, композиции могут включать биосурфактант. Типичные биосурфактанты включают, но без ограничений, гликолипиды, липопептиды, депсипептиды, фосфолипиды, замещенные жирные кислоты, липополисахариды, сурлактин, сурфактин, висконсин и рамнолипиды.

Вирусная инженерия.

Методы генетического конструирования вирусных частиц являются трудоемкими и длительными из-за отсутствия широко применимых и целевых методов конструирования *in vitro*. Современные методы *in vivo* могут занять недели или месяцы для создания модифицированных вирусов и вирусных векторов (Levin and Bull, Nat. Rev. Microbiol., 2004, февраль, 2 (2): 166-73, включены в настоящий документ по-



средством ссылки). Кроме того, существует токсичность, связанная с манипуляцией вирусными геномами в клетках. До этого раскрытия усилия по разработке широко применимых методов для точной генной инженерии вирусов *in vitro* были в значительной степени безуспешными. В настоящем документе описан широко применимый процесс для быстрого создания вирусных геномов полностью *in vitro*.

Представленные в настоящем документе системы и методы генетического конструирования *in vitro* имеют несколько преимуществ по сравнению с существующими методами генной инженерии вирусов: 1) позволяют легко манипулировать токсичными генами/продуктами полностью *in vitro*; 2) быстрые, т.е. могут быть выполнены в течение дня по сравнению с неделями или месяцами для методов *in vivo*; 3) позволяют удерживать геномную модификацию большей части вирусного генома; 4) не требует рекомбинации сигнальных путей хозяина; 5) более прямой и менее подвержен ошибкам, чем предыдущие методы; и 6) применим к нескольким вирусам без изменений в протоколе.

Изобретение представляет собой способы расщепления, опосредуемой РНК-направляемой нуклеазой, и сборку *in vitro* сайт-специфических сконструированных целых геномов. Изобретение значительно увеличивает точность, простоту и скорость, с которой вирусные геномы могут быть генетически модифицированы. Кроме того, данный способ преодолевает сложившуюся сложность манипулирования часто ядовитыми вирусными геномами вируса внутри клеток-хозяев. Также данный подход, выполняемый полностью в условиях *in vitro*, устраняет потребность в генетически приемлемом штамме хозяина для конструирования, что препятствует манипулированию многими важными и интересными вирусами архей, прокариот и эукариот. Данный подход не усиливает манипуляции вирусными геномами и поэтому позволяет сохранять большинство модификаций вирусного генома, таких как метилирование. Хорошо известно, что модификации генома могут иметь глубокое влияние на пригодность вирусов, и поэтому сохранение данных модификаций генома дает четкое преимущество перед другими инженерными методами. Кроме того, этот метод отличается от других методов, относящихся к конструированию генома РНК-направляемой нуклеазой *in vivo*, поскольку он не сосредоточен на использовании РНК-направляемой нуклеазы, такой как Cas9 и гРНК для редактирования эукариотического генома, но вместо этого относится к преодолению известных проблем вирусной инженерии целиком в условиях *in vitro*.

В некоторых аспектах, описанные в настоящем документе новые способы могут включать модификацию вирусной нуклеиновой кислоты или вирусного генома, например, используя РНК-направляемую нуклеазу и сборку, описанную в настоящем документе, и введение сконструированной вирусной нуклеиновой кислоты или сконструированного вирусного генома непосредственно в хозяин, который будет продуцировать сконструированные вирусные частицы или сконструированные вирусы, которые содержат сконструированную вирусную нуклеиновую кислоту или сконструированный вирусный геном. Например, в некоторых аспектах способы включают конструирование вирусной нуклеиновой кислоты или вирусного генома без введения сконструированной вирусной нуклеиновой кислоты или сконструированного вирусного генома в клонирующем хозяине для амплификации сконструированной вирусной нуклеиновой кислоты или сконструированного вирусного генома, например, посредством репликации в векторе. Например, в некоторых способах сконструированная вирусная нуклеиновая кислота или сконструированный вирусный геном не вводятся в дрожжи, *E.coli* или другие известные клонирующие хозяева, такие как, но без ограничений, *Bacillus* или *Vibrio*, до введения сконструированной вирусной нуклеиновой кислоты или сконструированного вирусного генома в клетку-хозяина, которая будет продуцировать сконструированные вирусные частицы или сконструированные вирусы.

Новые способы, предложенные в настоящем документе, позволяют проводить целенаправленную модификацию двух, трех, четырех, пяти или более сайтов в вирусном геноме. Способы могут быть выполнены целиком в условиях *in vitro*, что позволяет производить вирусные геномы, измененные в нескольких сайтах, что не достигается с использованием обычных методов конструирования.

Приведенные в настоящем документе сконструированные вирусы, содержащие сконструированную вирусную нуклеиновую кислоту и/или сконструированные вирусные геномы, которые имеют две, три, четыре, пять или более модификаций в отношении несконструированной вирусной нуклеиновой кислоты или несконструированного вирусного генома. Две или более модификаций могут быть вставкой, делецией, заменой или любой их комбинацией. Две или более модификаций могут приводить к одному, двум или более улучшенным вирусным свойствам, таким как любое раскрытое в настоящем документе. Сконструированные вирусы можно получать полностью посредством методов конструирования в условиях *in vitro*, раскрытых в настоящем документе. Способы конструирования *in vitro*, описанные в настоящем документе, приводят к целенаправленным модификациям, в отличие от классического или случайного мутагенеза. В отличие от модификаций, получаемых классическим или случайным мутагенезом, целевые модификации могут быть удобно скринированы для использования в стандартных молекулярно-генетических лабораторных методах, таких как ПЦР и/или секвенирование до любых анализов фенотипа.

Также в настоящем документе раскрыта система получения синтетических вирусов с улучшенными вирусными свойствами (например, см., фиг. 10). Система включает идентификацию последовательностей нуклеиновых кислот, ответственных за обеспечение требуемого свойства, и затем включение этих изменений последовательности в выбранный вирусный геном для получения вирусных частиц с улучшенными вирусными свойствами. Последовательности нуклеиновой кислоты, способные придавать желаемое

вирусное свойство, могут быть идентифицированы с помощью фундаментальных научных знаний, поиска литературы, эмпирического тестирования, исследований мутаций или любой их комбинации. Исследования мутаций могут включать эволюционные исследования, исследования адаптации, исследования мутагенеза и/или другие экспериментальные подходы, хорошо известные в данной области техники. Исследования мутагенеза могут включать ультрафиолетовый (УФ), химический и/или инсерционный мутагенез. Инсерционный мутагенез может включать транспозон и/или селективный маркер инсерционного мутагенеза. Эксперименты по мутации, используемые для идентификации интересующих последовательностей нуклеиновых кислот, могут быть выполнены с использованием вируса или вирусного генома вируса, который станет отправной точкой для конструирования *in vitro*. Дополнительно или альтернативно вместо выбранного вируса или вирусного генома, связанный или гетерологичный вирус, или вирусный геном может быть использован в исследовании мутаций с целью идентификации последовательностей рекомбинантных нуклеиновых кислот для включения в первоначально выбранный вирус или вирусный геном для придания дополнительных свойств выбранного вируса.

Желаемые свойства могут включать одно или более из круга хозяев, вирусного литического цикла, адсорбции, прикрепления, инъекции, репликации и сборки, лизиса, величины выхода фага, ускользания от иммунологического надзора, иммунной стимуляции, иммунной дезактивации, дисперсии биопленки, устойчивости бактерий к фагу, антибиотической сенсбилизации бактерий, модуляции факторов вирулентности и направленного расщепления или редактирования генома хозяина, других необходимых свойств, которые могут быть хорошо известны специалисту в данной области техники или в любой их комбинации. Идентифицированные последовательности нуклеиновых кислот, обладающие необходимым свойством, могут быть включены в выбранный вирусный геном, используя описанный в настоящем документе метод конструирования *in vitro* для включения одного или более изменений в один вирусный геном в течение одного или более циклов повторного конструирования и тестирования до тех пор, пока не будет подтвержден желаемый набор одного или более улучшенных вирусных свойств. Конечный вирусный геном, представляющий интерес, может представлять собой комбинацию естественных и синтезированных молекул нуклеиновых кислот или может быть полностью синтезирован *de novo* с использованием способов, описанных в настоящем документе, и/или известных в данной области техники. Получение вирусов или вирусных частиц с улучшенными вирусными свойствами может включать введение сконструированного вирусного генома, представляющего интерес, в совместимую клетку, причем геном активируется, тем самым создавая вирусные частицы или вирусы. Для получения молекулы нуклеиновой кислоты, идентифицированной для придания желаемого свойства для включения в выбранный вирусный геном, последовательность, представляющая интерес, может быть выделена или амплифицирована из вирусного генома, из которого она была идентифицирована путем расщепления, ПЦР-амплификации, синтеза, других методов известных в данной области техники, или в любой их комбинации. Синтезированную последовательность нуклеиновой кислоты можно химически синтезировать или собрать из химически синтезированных перекрывающихся олигонуклеотидов. Дополнительно или альтернативно, молекула нуклеиновой кислоты, которая должна быть включена в выбранный вирусный геном для того, чтобы придать желаемый фенотип, может быть комбинацией последовательностей естественных и синтезированных нуклеиновых кислот. В зависимости от конструкции молекулы нуклеиновой кислоты, которая должна быть включена в выбранный вирусный геном, полученный сконструированный вирусный геном может содержать последовательности нуклеиновых кислот, замененные альтернативными последовательностями или любой их комбинацией для придания необходимого вирусного свойства. Способы конструирования молекул нуклеиновой кислоты для того, чтобы изменить последовательность таким образом, что последовательности выделяют, удаляют, заменяют или любую их комбинацию хорошо известны специалистам в данной области техники. Сконструированные вирусные геномы, получаемые при помощи системы и методами, описанными в настоящем документе, могут быть использованы для получения вирусов или вирусных частиц с улучшенными вирусными свойствами. Получение вирусов или вирусных частиц с улучшенными вирусными свойствами может включать введение сконструированного вирусного генома в совместимую клетку, причем геном активируется, тем самым создавая вирусные частицы или вирусы. Введение сконструированного генома в клетку может быть осуществлено путем электропорации, трансформации, конъюгации, контакта клетки с предварительно упакованными вирусными геномами и т.д. или другими способами, хорошо известными в данной области техники.

Изобретение дополнительно относится к открытию способа для конструирования нуклеиновой кислоты *in vitro* с использованием РНК-направляемой эндонуклеазы. Настоящее раскрытие дополнительно относится к улучшению вирусных свойств при помощи генной инженерии *in vitro* вирусных нуклеиновых кислот. В частности, раскрытие относится к расщеплению *in vitro* последовательностей вирусов с использованием эндонуклеаз, таких как РНК-направляемая эндонуклеаза, например, Cas9, с последующей сборкой рекомбинантной нуклеиновой кислоты путем вставки фрагмента(ов) ДНК или РНК в расщепленный вирусный геном.

В некоторых аспектах настоящее раскрытие обеспечивает способ конструирования *in vitro* вирусной нуклеиновой кислоты, включающий выделение вирусной нуклеиновой кислоты; расщепление *in vitro* области вирусной нуклеиновой кислоты с использованием РНК-направляемой нуклеазы; и сборку

рекомбинантной нуклеиновой кислоты путем вставки фрагмента ДНК или РНК в расщепленную вирусную нуклеиновую кислоту. В некоторых примерах расщепление *in vitro* представляет собой ферментативное расщепление, основанное на РНК. В некоторых примерах ферментативное расщепление выполняется при помощи РНК-направляемой нуклеазы. В некоторых примерах РНК-направляемая нуклеаза представляет собой Cas9, фермент, полученный из Cas9, фермент, связанный с Cas9, или любую очищенную программируемую РНК-направляемую нуклеазу. В некоторых примерах расщепление дополнительно включает нацеливающие РНК. В некоторых примерах система расщепления дополнительно содержит спермидин. В некоторых примерах нацеливающие РНК представляют собой гРНК, сгРНК и/или tracrРНК. В некоторых примерах после расщепления РНК-направляемую нуклеазу инактивируют стандартными методами, такими как воздействие тепла и/или удалением стандартными методами, такими как, например, фенольно-хлороформной экстракцией. В некоторых примерах нагревание в процессе активации достигается путем воздействия на раствор, содержащий белок, нагревая, например, по меньшей мере на 80°C.

Любая программируемая РНК-направляемая нуклеаза может быть использована в способах и композициях настоящего изобретения, например Cas1, Cas1B, Cas2, Cas3, Cas4, Cas5, Cas6, Cas7, Cas8, Cas9 (также известна как Csn1 и Csx12), Cas10, Csy1, Csy2, Csy3, Cse1, Cse2, Cse1, Csc2, Csa5, Csn2, Csm2, Csm3, Csm4, Csm5, Csm6, Cmr1, Cmr3, Cmr4, Cmr5, Cmr6, Csb1, Csb2, Csb3, Csx17, Csx14, Csx10, Csx16, CsaX, Csx3, Csx1, Csx15, Csf1, Csf2, Csf3, Csf4, Cpf1, C2c1, C2c2, C2c3 или их гомологи, или их сконструированные варианты. Любую программируемую систему CRISPR можно использовать в способах и композициях в настоящем документе, включая тип I, тип II, тип III, тип IV, тип V, тип VI или любую их комбинацию. РНК-направляемая нуклеаза может быть белком Cas9, таким как белок Cas9 из *Staphylococcus pyogenes*, *S.thermophilus*, *S.pneumonia*, *S.aureus* или *Neisseria meningitidis*, в виде неограничивающих примеров. Также рассматриваются белки Cas9, представленные в виде SEQ ID NO: 1-256 и 795-1346 в публикации заявки на патент США № US 2014/0068797, включенной в настоящий документ посредством ссылки, и химерных белков Cas9, которые могут объединять домены из более чем одного белка Cas9, а также варианты и мутанты идентифицированных белков Cas9. В дополнение к Cas9, специалист в данной области техники легко узнает, что любой известный функциональный эквивалент будет достаточным альтернативным примером.

Вирусные частицы могут быть архей-, прокариот- или эукариот- специфическими вирусами. Например, вирус может быть таким, который может заразить *Pseudomonas aeruginosa*, *E.coli* или *Homo sapiens*. В некоторых примерах вирус может быть таким, который заражает виды возбудителей, такие как те, что относятся к роду *Acinetobacter*, *Clostridium*, *Enterobacter*, *Enterococcus*, *Escherichia*, *Klebsiella*, *Mycobacterium*, *Neisseria*, *Pseudomonas*, *Salmonella*, *Staphylococcus* или *Streptococcus*. В некоторых примерах вирус может инфицировать виды архей, такие как те, что относятся к роду *Acidianus*, *Aeropyrum*, *Haloarcula*, *Haloferax*, *Halogulbum*, *Methanobacterium*, *Pyrobaculum*, *Pyrococcus*, *Stygiolobus*, *Sulfolobus* или *Thermoproteus*. В некоторых примерах вирус может инфицировать эукариотические хозяева, такие как люди, млекопитающие, животные, растения, водоросли или грибы. Вирусная нуклеиновая кислота может представлять собой ДНК или РНК. В некоторых примерах вирусная нуклеиновая кислота состоит из цельного вирусного генома, части вирусного генома, или одного или более вирусных генов. В некоторых примерах часть вирусного генома субклонирована в плазмиду до конструирования.

Вирусная нуклеиновая кислота может быть одиночной или двойной (или более) расщепленной РНК-направляемой нуклеазой, такой как Cas9, в сочетании с нацеливающей(ими) РНК *in vitro* для удаления одного или более нуклеотидов, одного гена, множества генов или любого размера геномной области или открытой ДНК для вставки новой последовательности. В дополнение к Cas9 специалисту в данной области техники понятно, что любая программируемая РНК-направляемая нуклеаза или другой механизм расщепления ДНК, пригодный для ДНК, будет достаточным и будет функционально эквивалентным. Множественные расщепления могут быть выполнены одновременно; однако было обнаружено, что последовательное расщепление РНК-направляемой Cas9 может повысить эффективность. Кроме того, спермидин может быть добавлен в реакционную смесь для увеличения диссоциации Cas9 от ДНК, что позволяет увеличить доступность Cas9 для ферментативной активности. Вирусная последовательность, удаленная расщеплением Cas9, не рекомбинируется обратно в геном, потому что Cas9 представляет собой тупой режущий фермент и фрагменты не содержат гомологии с сайтом встраивания. Кроме того, дезактивация теплом Cas9 обеспечивает непосредственный переход от расщепления к реакциям сборки, упрощая протокол.

Используемый в настоящем документе термин "нацеливающие РНК" или "направляющие РНК" относится к CRISPR РНК (сгРНК), транс-активирующим сгРНК (сгРНК), модифицированным химерным гидовым РНК (гРНК), включающим как сгРНК, так и tracrРНК, или одиночные гРНК, совместимые с выбранной системой CRISPR. CRISPR РНК (сгРНК) транскрибируют из локуса CRISPR, включают в эффективные комплексы и направляют комплекс к последовательностям встраивающейся нуклеиновой кислоты, что приводит к расщеплению нуклеиновой кислоты, опосредованному РНК-направляемой нуклеазой. TracrРНК являются комплементарными с и комплементарной парами оснований с pre-сгРНК, образующими дуплекс РНК, необходимый для Cas9-опосредованного расщепления. Гибридные гРНК

представляют собой химерные РНК, которые связывают нацеливающую sgРНК с tracrРНК, что позволяет использовать одну РНК для Cas9-опосредованного расщепления. Cas9-опосредованное расщепление могут проводить как с транскрибированной *in vitro* смесью sgРНК-tracrРНК или с химерной gРНК.

Вставка ДНК или РНК может быть получена любыми способами, известными в данной области техники, а именно путем синтеза *in vitro*, химического синтеза, синтеза *de novo*, сборки *de novo*, амплификации (ПЦР), фермент-опосредуемого освобождения от плазмид, вирусов или бактерий, или любой их комбинации. В одном аспекте вставка ДНК или РНК создается сборкой олигонуклеотидов или ПЦР с праймерами, содержащими перекрывающиеся последовательности с сайтом встраивания. Вставка ДНК или РНК может быть комбинацией естественных и синтезированных нуклеиновых кислот, или полученных полностью естественным или синтетическим путем.

Сборка вставки ДНК или РНК и расщепленной вирусной нуклеиновой кислоты может быть проведена любым известным в данной области техники способом, таким как реакции клонирования *in vitro* или любым из ранее обсуждавшихся способов. В одном аспекте сборка вставки ДНК или РНК в расщепленный вирусный геном выполняют с использованием метода сборки Gibson Assembly. В одном аспекте сборку вставки ДНК или РНК в расщепленный вирусный геном выполняют *in vivo* с использованием механизмов рекомбинации клеток-хозяев. Сборка вставки ДНК или РНК может приводить к вставке, делеции, замене или любой их комбинации последовательности нуклеиновой кислоты. Процесс конструирования последовательности ДНК или РНК такой, что сборка в расщепленную вирусную нуклеиновую кислоту приводит к вставке, делеции, замене или любой их комбинации нуклеиновых кислот, представляющих интерес, хорошо известны в данной области техники.

В некоторых аспектах настоящее раскрытие обеспечивает способ конструирования *in vitro* вирусной последовательности, включающий выделение вирусной нуклеиновой кислоты; расщепления *in vitro* области вирусной нуклеиновой кислоты с использованием РНК-направляемой нуклеазы; и сборку рекомбинантной нуклеиновой кислоты путем вставки фрагмента ДНК или РНК в расщепленную вирусную нуклеиновую кислоту. В некоторых примерах сборку проводят *in vitro* в одном сосуде со смесью компонентов, включающих: (a) выделенную не термостабильную 5'-3'-экзонуклеазу, которая не обладает 3'-экзонуклеазной активностью; (b) краудинг-агент; (c) выделенную термостабильную невытесняющую ДНК-полимеразу с 3'-экзонуклеазной активностью или смесью указанной ДНК-полимеразы со второй ДНК-полимеразой, которая не обладает 3'-экзонуклеазной активностью; (d) выделенную термостабильную лигазу; (e) смесь ДНТФ; и (f) подходящий буфер в условиях, эффективных для вставки фрагмента в расщепленную вирусную нуклеиновую кислоту с образованием рекомбинантной нуклеиновой кислоты. В некоторых аспектах экзонуклеаза представляет собой экзонуклеазу T5, и приведение в контакт происходит в изотермических условиях, и/или краудинг-агент представляет собой ПЭГ, и/или невытесняющая ДНК-полимераза представляет собой ДНК-полимеразу Phusion™ или ДНК-полимеразу VENT® и/или лигазу представляет собой Taq-лигазу. В некоторых примерах сборку *in vitro* выполняют при помощи одноступенчатой или изотермической сборки методом Gibson Assembly. В некоторых примерах сборку *in vitro* выполняют при помощи двухступенчатой сборки методом Gibson Assembly. В некоторых примерах расщепленная нуклеиновая кислота и фрагмент ДНК или РНК могут быть собраны *in vitro* лигированием тупых концов с использованием фермента лигазы.

В некоторых аспектах настоящее раскрытие обеспечивает метод *in vitro* для конструирования вирусной последовательности, включающий этап сборки. В некоторых примерах сборку выполняют *in vivo* в совместимой клетке-хозяине с использованием механизма рекомбинации клеток-хозяев. Хотя рекомбинантная нуклеиновая кислота может быть собрана целиком в условиях *in vitro* с использованием очищенных ферментов, как описано в настоящем документе, этот процесс также может быть осуществлен с использованием естественных или сконструированных рекомбинантных сигнальных путей в чувствительном штамме хозяина. В некоторых случаях совместимыми клетками-хозяевами могут быть *S.cerevisiae*, *E.coli*, *P.aeruginosa*, *B.subtilis*, *V.natrigens* или другие организмы, доступные в данной области техники. Трансформация очищенных и обработанных *in vitro* вирусных геномов вместе с фрагментом вставки восстановления, содержащим концевые области гомологии, достаточна для того, чтобы некоторые клетки-хозяева собирали рекомбинантный вирусный геном *in vivo*. Фрагменты вставки восстановления могут быть синтезированы или амплифицированы стандартными методами, известными в данной области техники, или могут находиться внутри плазмид, стабильно реплицирующихся в выбранной клетке-хозяине. Данный способ, вероятно, будет иметь более низкую эффективность по сравнению со сборкой *in vitro* из-за клеток-хозяев, имеющих как гомологичные, так и не гомологичные пути восстановления ДНК, задача совместной доставки достаточного количества вставки и расщепленного генома в клетку-хозяина и более низкой эффективности большинства гомологичных сигнальных путей рекомбинации хозяина. Поскольку только расщепленные геномы не образуют функциональных вирусных частиц и в последующем бляшек без опосредованной хозяином рекомбинации, бляшки, полученные после трансформации и выращивания, могут быть скринированы с помощью ПЦР для данной вставки, чтобы подтвердить корректную сборку желаемой сконструированной вирусной нуклеиновой кислоты.

В некоторых аспектах настоящее раскрытие обеспечивает метод *in vitro* для конструирования вирусной последовательности, содержащей РНК-направляемую нуклеазу. В некоторых примерах РНК-

направляемая нуклеаза представляет собой Cas9 II типа. В некоторых примерах РНК-направляемая нуклеаза представляет собой Cas9 или фермент, полученный из Cas9. В некоторых примерах РНК-направляемая нуклеаза представляет собой выделенный рекомбинантный фермент Cas9 или Cas9. В некоторых примерах существует по меньшей мере одна нацеливающая РНК. В некоторых примерах существует две нацеливающие РНК. В некоторых примерах нацеливающая РНК представляет собой химерную гидовую РНК (гРНК) или набор crРНК и tracrРНК. В некоторых примерах в реакции расщепления *in vitro* использует две гРНК. В некоторых примерах в реакции расщепления *in vitro* использует два набора crРНК и tracrРНК, чтобы, например, одновременно нацеливать на две последовательности.

В некоторых аспектах настоящее раскрытие обеспечивает способ конструирования *in vitro* вирусной последовательности, включающий стадию расщепления *in vitro*. В некоторых примерах после расщепления РНК-направляемую нуклеазу инактивируют при помощи стандартных методов, таких как воздействие теплом, например, по меньшей мере на 80°. В некоторых примерах после расщепления РНК-направляемую нуклеазу удаляют фенольно-хлороформной экстракцией. В некоторых примерах, после расщепления, РНК-направляемую нуклеазу удаляют другими способами экстракции, хорошо известными в данной области техники.

В некоторых аспектах настоящее раскрытие обеспечивает способ *in vitro* для конструирования вирусной последовательности, которая приводит к созданию сконструированной вирусной нуклеиновой кислоты. В некоторых примерах сконструированную вирусную нуклеиновую кислоту затем трансформируют в клетку-хозяин. В некоторых примерах клетка-хозяин представляет собой *E.coli*, *P.aeruginosa*, *S.cerevisiae*, *V.natriegens*, *B.subtilis* или другой организм, хорошо известный в данной области техники. В некоторых примерах трансформацию выполняют путем теплового шока, электропорации, биолистики, бомбардировки частиц, конъюгации, трансдукции, липофекции или другого известного метода, хорошо известного в данной области техники. В некоторых примерах сконструированную вирусную нуклеиновую кислоту трансформируют в клетку-хозяина и затем снова выделяют после репликации. В некоторых примерах выделенную сконструированную вирусную нуклеиновую кислоту используют в качестве исходной вирусной нуклеиновой кислоты для другого цикла конструирования *in vitro*, который в настоящем документе называется итеративное конструирование *in vitro*. В некоторых примерах представлен один цикл итеративного конструирования *in vitro*. В других примерах представлен по меньшей мере один цикл итеративного конструирования *in vitro*. В других примерах существует два или более циклов итеративного инженерии *in vitro*.

В некоторых аспектах настоящее раскрытие обеспечивает способ *in vitro* конструирования вирусной последовательности, который приводит к созданию сконструированной вирусной нуклеиновой кислоты. В некоторых примерах сконструированную вирусную нуклеиновую кислоту упаковывают в вирусные частицы с использованием набора для упаковки *in vitro*, который может быть коммерчески доступен. В некоторых примерах комплект для упаковки *in vitro* представляет собой набор Maxplax lambda packaging extract.

В некоторых аспектах настоящее раскрытие обеспечивает способ *in vitro* для конструирования вирусной последовательности, который приводит к образованию рекомбинантной сконструированной вирусной нуклеиновой кислоты. В некоторых примерах сконструированная вирусная нуклеиновая кислота улучшает или изменяет свойство вируса по сравнению с контрольным и/или несконструированным вирусом. В некоторых примерах улучшенное или измененное вирусное свойство представляет собой свойство, такое как круг хозяев, вирусный литический цикл, адсорбция, прикрепление, инъекция, репликация и сборка, лизис, величина выхода фага, ускользание от иммунологического надзора, иммунная стимуляция, иммунная дезактивация, дисперсия биопленки, устойчивость к бактериальному фагу, антибиотическая сенсibilизация бактерий, модуляция факторов вирулентности, направленное расщепление или редактирование геном хозяина или любая их комбинация. В некоторых примерах улучшение свойства может быть увеличением, уменьшением или изменением свойства. Например, улучшенное вирусное свойство может быть расширенным или уменьшенным кругом хозяев, измененным вирусным литическим циклом, увеличенной или уменьшенной адсорбцией в клетке-хозяине, увеличенным или уменьшенным прикреплением к клетке-хозяина, увеличенным или уменьшенным инъекцией, увеличенной или уменьшенной или измененной репликацией и сборкой, повышенным или сниженным лизисом, увеличенной или уменьшенной величиной выхода фага, увеличенным или уменьшенным или измененным ускользанием от иммунного надзора, увеличенной или уменьшенной или измененной иммунной стимуляцией, увеличенной или уменьшенной или измененной иммунной дезактивацией, увеличенной или уменьшенной или измененной дисперсией биопленки, увеличенной или уменьшенной или измененной устойчивости бактерий к фагу, повышенной или уменьшенной или измененной антибиотической сенсibilизацией бактерий, увеличенной или уменьшенной или измененной модуляцией факторов вирулентности, увеличением или уменьшением или изменением направленного расщепления или редактирования генома или любой их комбинации.

В некоторых аспектах настоящее раскрытие обеспечивает способ конструирования вирусной нуклеиновой кислоты, который приводит к улучшению вирусного свойства, такого как, например, увеличенный круг хозяев. Круг хозяев представляет собой количество типов клеток, штаммов или видов хозя-

ев, которые вирус может заразить. Увеличение круга хозяев представляет собой расширение абсолютно-го числа различных типов клеток, штаммов или видов, которые вирус способен инфицировать по сравнению с контрольным и/или несконструированным вирусом. В некоторых примерах увеличенный круг хозяев представляет собой увеличение количества бактериальных штаммов или вариантов внутри бактериальных видов, которые вирус способен инфицировать. Увеличение в круге хозяев может быть увеличением по меньшей мере одного или более штаммов, типов клеток или видов. Круг хозяев могут анализировать, например, стандартным анализом бляшкообразования, который хорошо известен в данной области техники.

В некоторых аспектах настоящее раскрытие обеспечивает способ для конструирования вирусной нуклеиновой кислоты, который приводит к улучшению вирусного свойства, такого как, например, вирусный литический цикл. Вирусный литический цикл представляет собой один из двух циклов вирусной репликации, другой является лизогенным циклом. Литический цикл приводит к разрушению инфицированной клетки и инфицированной клеточной мембраны. Литический цикл состоит из шести этапов, каждый из которых может быть индивидуально сконструирован. Шесть этапов вирусного литического цикла представляют собой адсорбцию, прикрепление, инъекцию, репликацию и сборку, лизис и величину выхода фага.

В некоторых аспектах настоящее раскрытие обеспечивает способ конструирования вирусной нуклеиновой кислоты, который приводит к улучшению вирусного свойства, такого как, например, адсорбция. Адсорбция представляет собой действие вируса, контактирующего с клеткой-хозяином. Вирусная адсорбция характеризуется как сродство вируса к данной клетке-хозяину и может быть проанализировано с помощью стандартных анализов адсорбции, например, описанных Хайманом и Абедоном (*Methods in Molecular Biology*, 2009). Кроме того, или, альтернативно, вирусная адсорбция может быть определена другими стандартными анализами аффинности, широко используемыми в биохимии для анализа взаимодействий рецептор-лиганд.

В некоторых аспектах настоящее раскрытие обеспечивает способ конструирования вирусной нуклеиновой кислоты, который приводит к улучшению вирусного свойства, такого как, например, прикрепление. Вирусное прикрепление - это когда вирус сильно прикрепляется к клетке-хозяина. Вирусное прикрепление является необратимым взаимодействием между вирусом и рецептором клетки-хозяина.

В некоторых аспектах настоящее раскрытие обеспечивает способ конструирования вирусной нуклеиновой кислоты, который приводит к улучшению вирусного свойства, такого как, например, инъекция. Введение относится к инъекции вирусного генома, и когда вирус вставляет свой генетический материал в клетку-хозяина. В качестве примера можно привести инъекцию вирусного генома путем измерения оттока ионов калия (Cady et al., *J. Bacteriol.*, 2012, 194 (21): 5728-38; Leavitt et al., *PLoS ONE*, 2013 8 (8): E70936, которые включены в настоящее описание посредством ссылки во всей их полноте).

В некоторых аспектах настоящее раскрытие обеспечивает способ конструирования вирусной нуклеиновой кислоты, который приводит к улучшению вирусного свойства, такого как, например, репликация и сборка. Вирусная репликация и сборка относятся к созданию новых вирусов в клетке-хозяине. После инъекции вирусного генома механизм клеток-хозяев захватывается и вирусные гены транскрибируются, вирусные белки транслируются, а вирусные частицы представляют собой сборку, содержащую реплицированные вирусные геномы. Вирусная репликация и сборка в конечном итоге приводят к лизису клеток-хозяев, поэтому репликацию и сборку можно анализировать, контролируя скорость роста вируса стандартным анализом бляшкообразования или анализом бляшкообразования в двойном агаровом слое. Скорости вирусной репликации могут дополнительно или альтернативно определяться путем измерения величины выхода фага в стандартном анализе бляшкообразования, кривой одиночного цикла размножения или других стандартных анализах на репликативную способность вируса, которые хорошо известны в данной области техники.

В некоторых аспектах настоящее раскрытие обеспечивает способ конструирования вирусной нуклеиновой кислоты, который приводит к улучшению вирусного свойства, такого как, например, лизис. Лизис относится к лизису клеток-хозяев. После репликации и сборки новых вирусных частиц образуется фермент, который разрушает стенку клетки-хозяина и/или клеточную мембрану изнутри и позволяет вводить жидкость, что в конечном итоге приводит к лизису клеток-хозяев. Способность увеличивать или ингибировать вирулентную репликацию вируса может увеличивать или уменьшать время, необходимое для того, чтобы данный вирус убивал клетку-хозяина путем лизиса. Вирусную вирулентность можно исследовать, анализируя время между заражением и лизисами клеток-хозяев, контролируя скорость роста вируса стандартным анализом бляшкообразования или анализом бляшкообразования в двойном агаровом слое. Кроме того, или альтернативно, усиленный бактериальный лизис сконструированного вируса по сравнению с контрольным и/или несконструированным вирусом может быть определен при помощи колониеобразующих единиц (КОЕ) после анализа, количества или диаметра бляшкообразующих единиц (БОЕ) или после анализа бляшкообразования, из анализа биопленки или других стандартных анализов, которые хорошо известны в данной области техники.

В некоторых аспектах настоящее раскрытие обеспечивает способ модификации вирусной нуклеиновой кислоты, который приводит к улучшению вирусного свойства, такого как, например, величина

выхода фага. Величина выхода фага относится к числу вирусов, продуцируемых зараженной клеткой. Величина выхода фага может быть проанализирована стандартными анализами величины выхода фага, такими как те, которые описаны Эллисом и Делбрейком (J. Gen. Physiol. 1939 Jan 20; 22(3): 365-384, включенный в настоящий документ посредством ссылки) и Delbriick (Delbriick J. Gen. Physiol, 1940, 23; 643, включенный в настоящий документ посредством ссылки)

В некоторых аспектах настоящее раскрытие обеспечивает способ модификации вирусной нуклеиновой кислоты, который приводит к улучшению вирусного свойства, такого как, например, ускользание от иммунного надзора. Ускользание от иммунного надзора представляет собой способность вируса избегать клиренса врожденной или адаптивной иммунной системы. Ускользание от иммунного надзора может быть проанализировано путем изучения уровня или скорости нейтрализации продуцирования антител. Кроме того, или, альтернативно, ускользание от иммунного надзора можно измерить, анализируя период полужизни или времени пребывания данного вируса внутри животного.

В некоторых аспектах настоящее раскрытие обеспечивает собой способ модификации вирусной нуклеиновой кислоты, который приводит к улучшению вирусного свойства, такого как, например, иммунная стимуляция. Иммунная стимуляция представляет собой способность вируса индуцировать иммунный ответ, который обычно не ассоциируется с вирусом дикого типа или несконструированным вирусом. Это может быть проанализировано путем анализа иммунных факторов, продуцируемых в присутствии вируса, с использованием стандартных наборов ИФА, проточной цитометрии, гистологии или других стандартных иммунологических анализов, известных специалистам в данной области техники.

В некоторых аспектах настоящее раскрытие обеспечивает способ модификации вирусной нуклеиновой кислоты, который приводит к улучшению вирусного свойства, такого как, например, иммунная дезактивация. Иммунная дезактивация представляет собой способность вируса уменьшать иммунный ответ, обычно связанный с вирусом дикого типа или несконструированным вирусом. Это может быть проанализировано путем анализа иммунных факторов, продуцируемых в присутствии вируса, с использованием стандартных наборов ИФА, проточной цитометрии, гистологии или других стандартных иммунологических анализов, известных специалистам в данной области техники.

В некоторых аспектах настоящее раскрытие обеспечивает способ модификации вирусной нуклеиновой кислоты, который приводит к улучшению вирусного свойства, такого как, например, дисперсия биопленки. Дисперсия биопленки представляет собой способность деградировать, ослаблять или увеличивать проницаемость биопленки. Мероприятия, которые могут приводить к дисперсии биопленки, включают, но без ограничений, деградацию экзополисахарида (ЭПС), модуляцию молекул бактериального чувства кворума, и деградацию внеклеточной ДНК или РНК в биопленке или месте бактериальной инфекции. "Деградация экзополисахарида" представляет собой способность вируса продуцировать белок или фермент, способный разрушать или диссоциировать высокомолекулярные соединения, секретируемые микроорганизмами в их среду для формирования структурной целостности биопленок. ЭПС-деградирующая активность может включать, но без ограничений, поверхностно-активные вещества, гликозидазы и протеазы. Их активность может быть измерена с использованием стандартных биохимических анализов, известных специалистам в данной области техники. Модуляция молекул бактериального чувства кворума также может приводить к дисперсии биопленки. Известно, что молекулы бактериального чувства кворума являются высококонсервативными регуляторами вирулентности у ряда патогенных бактерий человека. Были идентифицированы белки с ферментативной активностью, способные разрушать молекулы бактериального чувства кворума, и их активность измеряли при помощи различных анализов микробных репортеров, биохимических репортерных анализов или путем анализа продуктов расщепления с использованием TLC (Rajesh and Rai, Microbiological Research, July-August 2014, Volume 169, Issues 7-8, Pages 561-569, включенные в настоящий документ посредством ссылки). Деградация внеклеточной ДНК или РНК в биопленке или месте бактериальной инфекции также может привести к дисперсии биопленки. Активность вирусной закодированной ДНКазы или РНКазы может быть измерена с помощью имеющихся в продаже наборов, известных специалистам в данной области техники, таких как те, что доступны от Jena Bioscience или ThermoFisher в качестве неограничивающих примеров. Предотвращение, проникновение, разрушение или дисперсия биопленки также может быть оценена путем количественной оценки биопленки, присутствующей после обработки, и сравнения ее с контрольным условием. Измерения биопленки хорошо известны в данной области техники и включают в качестве неограничивающего примера окрашивание биопленки красителем, таким как кристаллический фиолетовый, и количественное определение поглощения на спектрофотометре.

В некоторых примерах настоящее раскрытие обеспечивает способ модификации вирусной нуклеиновой кислоты, который приводит к улучшенному вирусному свойству, такому как, например, устойчивость бактериального фага. Фаг или бактериофаг представляют собой термины, которые можно использовать взаимозаменяемо и которые относятся к вирусам, которые заражают бактерии. Бактериальная устойчивость к фагу относится к появлению бактериофаг-устойчивых бактерий из популяции, обработанной или подвергнутой воздействию конкретного вируса. Это происходит либо посредством случайных мутаций внутри бактерий, либо из-за того, что некоторые бактерии внутри популяции не способны заразиться вирусом. Когда данные резистентные бактерии распространяются, новая популяция устойчива к

вирусу или бактериофагу, изначально подвергаемому воздействию. Неограничивающим примером оценки устойчивости бактерий является отслеживание скорости роста бактерий после вирусной обработки, поскольку количество резистентных бактерий непосредственно влияет на скорость повторного роста популяции. Бактериофаг может быть модифицирован таким образом, чтобы бактерии не могли получить вирусную резистентность по меньшей мере тремя способами, включая 1) ингибирование известных систем устойчивости к вирусам, 2) кодирование вторичного токсина и/или 3) увеличение вирулентности за счет увеличения литической способности. Бактериофаг может избегать или ингибировать известные системы устойчивости к вирусам посредством экспрессии известных или синтетических ингибирующих белков, как один пример. Активность данных ингибирующих белков можно контролировать с помощью классического метода титрования на двухслойной бляшке и/или анализа эффективности культивирования. Системы устойчивости к вирусам могут включать, но без ограничений, системы CRISPR-Cas и рестрикции-модификации. Предотвращение устойчивости к вирусам также может быть достигнуто путем экспрессии вторичных токсинов, таких как бактерицидные полезные нагрузки. Активность данных вторичных токсинов не зависит от естественной литической активности данного вируса и может быть измерена анализом кривой роста/гибели. Кроме того, или альтернативно, генетически закодированный токсический белок может быть очищен и охарактеризован с использованием установленных биохимических и/или фенотипических анализов, обычно используемых для характеристики белковых токсинов и которые хорошо известны специалисту в данной области техники.

В некоторых примерах настоящее раскрытие обеспечивает способ модификации вирусной нуклеиновой кислоты, который приводит к улучшению вирусного свойства, такого как, например, антибиотическая сенсibilизация бактерий. "Антибиотическая сенсibilизация бактерий" относится к способности вируса экспрессировать генетически закодированную полезную нагрузку, чтобы сделать инфицированные или соседние клетки более чувствительными к антимикробному агенту. Полезная нагрузка может быть генетически закодирована на вирусе или бактериофаге, а затем экспрессирована внутри клетки-хозяина. Экспрессированная полезная нагрузка может быть необязательно секретирована клеткой-хозяином или высвобождена при лизисе клеток-хозяев. Активность антибиотической сенсibilизации бактерий можно наблюдать при помощи тестирования синергии, используя, например, хорошо известный метод микроразведений с использованием серийных разведений.

В некоторых примерах настоящее раскрытие обеспечивает способ конструирования вирусной нуклеиновой кислоты, который приводит к улучшению вирусного свойства, такого как, например, модуляция факторов вирулентности. "Модуляция факторов вирулентности" относится к вирусным генетически закодированным белкам или соединениям, способным модулировать экспрессию или активность известных факторов вирулентности. Неограничивающими примерами модуляторов факторов вирулентности являются факторы транскрипции, антитела и белки иммунитета. Экспрессию или активность факторов вирулентности и модуляторов фактора вирулентности можно наблюдать, например, на животных моделях, биохимических тестах или репортерных анализах.

В некоторых примерах настоящее раскрытие обеспечивает способ конструирования вирусной нуклеиновой кислоты, который приводит к улучшению вирусного свойства, такого как, например, направленное расщепление или редактирование генома хозяина. "Расщепление или редактирование целевого генома хозяина" относится к способности вируса генетически кодировать специфичную для последовательности нуклеазу, способную целенаправленно расщеплять геном в данном генетическом локусе, и, необязательно, редактировать, например, путем вставки восстанавливающей молекулы ДНК. Активность целевого расщепления можно наблюдать путем секвенирования, определения количества жизнеспособных микроорганизмов, подтверждения интеграции новой последовательности и/или других стандартных методов, известных специалистам в данной области техники.

В некоторых аспектах настоящее раскрытие обеспечивает способ конструирования *in vitro* вирусной последовательности, включающий этап расщепления *in vitro*. В некоторых примерах расщепленную вирусную нуклеиновую кислоту выделяют и секвенируют вместо использования в реакции сборки *in vitro* или *in vivo*. В некоторых примерах результаты секвенирования из фрагмента вирусной нуклеиновой кислоты используют для определения концов вирусных геномов. В некоторых примерах скорректированные последовательности вирусного генома используют для планирования и разработки дальнейших подходов и этапов конструирования *in vitro*.

В некоторых аспектах настоящее изобретение относится к способу конструирования *in vitro* вирусной последовательности, включающей выделение вирусной нуклеиновой кислоты. В некоторых примерах вирусная нуклеиновая кислота представляет собой полный вирусный геном. В некоторых примерах полный вирусный геном выделяют из вирусной частицы. В некоторых примерах вирусная нуклеиновая кислота представляет собой часть вирусного генома. В некоторых примерах вирусная нуклеиновая кислота представляет собой часть вирусного генома, содержащегося в плазмиде. В некоторых примерах плазмиду, содержащую часть вирусного генома, выделяют из клетки-хозяина. В некоторых примерах часть вирусного генома клонировали в плазмиду, трансформировали в клетку-хозяина и изолировали перед конструированием *in vitro*. В некоторых примерах вирусную нуклеиновую кислоту синтезируют *de novo*. Синтез *de novo* может включать синтез олигонуклеотидов и сбор их *in vitro* или *in vivo* с использо-



ванием стандартных методов, известных в данной области техники. В некоторых примерах вирусную нуклеиновую кислоту амплифицируют перед расщеплением, таким как, например, с ПЦР-амплификацией.

В некоторых аспектах настоящее раскрытие обеспечивает набор для конструирования вирусной последовательности, включающий: (a) выделенную нетермостабильную 5'-3'-экзонуклеазу, которая не обладает 3'-экзонуклеазной активностью; (b) краудинг-агент; (c) выделенную термостабильную непересекающаяся ДНК-полимеразу с 3'-экзонуклеазной активностью или смесь указанной ДНК-полимеразы со второй ДНК-полимеразой, которая не обладает 3'-экзонуклеазной активностью; (d) выделенную термостабильную лигазу; (e) смесь дНТФ; (f) подходящий буфер; и (g) очищенную рекомбинантную РНК-направляемую нуклеазу. В некоторых примерах РНК-направляемая нуклеаза представляет собой Cas9 или фермент, полученный из Cas9. В некоторых примерах набор дополнительно содержит специально разработанные нацеливающие РНК. В некоторых примерах нацеливающие РНК представляют собой химерные гРНК или crРНК и tracrРНК. В некоторых примерах набор дополнительно содержит специально разработанные синтезированные молекулы нуклеиновой кислоты, предназначенные для использования в качестве вставленного фрагмента ДНК в реакции сборки. В некоторых примерах набор дополнительно содержит компетентные клетки-хозяева. В некоторых примерах набор дополнительно содержит выделенные вирусные нуклеиновые кислоты.

В некоторых аспектах настоящее раскрытие обеспечивает систему для конструирования *in vitro* вирусной нуклеиновой кислоты, содержащей выделенную вирусную нуклеиновую кислоту, рекомбинантную РНК-направляемую нуклеазу, по меньшей мере одну нацеливающую РНК и фрагмент ДНК или РНК, которые будут собраны в выделенную вирусную нуклеиновую кислоту в сайте расщепления. В некоторых примерах выделенная вирусная нуклеиновая кислота представляет собой полный геном, выделенный из вирусных частиц. В некоторых примерах выделенная вирусная нуклеиновая кислота представляет собой часть вирусного генома, который субклонирован в плазмиду и выделяется из клетки-хозяина. В некоторых примерах РНК-направляемая нуклеаза представляет собой Cas9 или фермент, полученный из Cas9. В некоторых примерах нацеливающая РНК представляет собой crРНК и tracrРНК. В некоторых примерах нацеливающая РНК представляет собой химерную гидовую РНК (гРНК). В некоторых примерах существует две нацеливающие РНК или гРНК. В некоторых примерах имеется два набора crРНК и tracrРНК.

В некоторых аспектах изобретение обеспечивает систему сконструированной вирусной нуклеиновой кислоты *in vitro*, содержащую выделенную вирусную нуклеиновую кислоту, рекомбинантную РНК-направляемую нуклеазу, по меньшей мере одну нацеливающую РНК и фрагмент нуклеиновой кислоты, который должен быть вставлен в сайт расщепления выделенной нуклеиновой кислоты. В некоторых примерах система такова, что рекомбинантная РНК-направляемая нуклеаза и по меньшей мере одна нацеливающая РНК образуют комплекс, способный расщеплять выделенную вирусную нуклеиновую кислоту. В некоторых примерах система дополнительно содержит спермидин. В некоторых примерах система дополнительно содержит выделенную нетермостабильную 5'-3'-экзонуклеазу, которая не обладает 3'-экзонуклеазной активностью; краудинг-агент; выделенную термостабильную невтыскающую ДНК-полимеразу с 3'-экзонуклеазной активностью или смесь указанной ДНК-полимеразы со второй ДНК-полимеразой, которая не обладает 3'-экзонуклеазной активностью; выделенную термостабильную лигазу; смесь дНТФ; и подходящий буфер, в котором система находится в условиях, эффективных для вставки фрагмента нуклеиновой кислоты в выделенную вирусную нуклеиновую кислоту в месте расщепления РНК-направляемой нуклеазы с образованием рекомбинантной вирусной нуклеиновой кислоты.

В некоторых аспектах система, описанная в настоящем документе, такова, что рекомбинантная вирусная нуклеиновая кислота способна продуцировать не встречающиеся в природе вирусные частицы по меньшей мере с одним улучшенным вирусным свойством по сравнению с эталонной и/или несконструированной вирусной нуклеиновой кислотой. В некоторых примерах улучшенное вирусное свойство выбирают из группы, состоящей из круга хозяев, вирусного литического цикла, адсорбции, прикрепления, инъекции, репликации и сборки, лизиса, величины выхода фага, ускользания от иммунного надзора, иммунной стимуляции, иммунной дезактивации, дисперсии био пленки, устойчивости бактерий к фагам, антибиотикочувствительности бактерий, модуляции факторов вирулентности и направленного расщепления или редактирования генома хозяина.

В некоторых аспектах в системе, описанной в настоящем документе, РНК-направляемая нуклеаза представляет собой Cas9 или фермент, полученный из Cas9. В некоторых примерах РНК-направляемую нуклеазу инактивируют или удаляют после расщепления.

Способ, описанный в настоящем документе, может быть использован в нескольких других вирусных геномах и вирусных векторных конструкциях, используемых для модификации геномов РНК путем прямого редактирования генома РНК или ДНК-матрицы, который затем будет транскрибироваться *in vitro* в вирусную РНК, используемую для конструирования и непосредственного изменения вирусов как прокариотов, так и эукариотов, и используется для непосредственной модификации вирусных геномов, используемых для отображения фагов, фаговой терапии, вирусной диагностики или разработки/производства вакцин.

В некоторых аспектах настоящее раскрытие обеспечивает рекомбинантную вирусную нуклеиновую кислоту, вырабатываемую любым из способов, описанных в настоящем документе. В некоторых примерах рекомбинантная вирусная нуклеиновая кислота способна продуцировать не встречающиеся в природе вирусные частицы по меньшей мере с одним улучшенным вирусным свойством по сравнению с несконструированной вирусной нуклеиновой кислотой. В некоторых примерах улучшенное вирусное свойство выбрано из группы, состоящей из круга хозяев, вирусного литического цикла, адсорбции, прикрепления, инъекции, репликации и сборки, лизиса, величины выхода фага, ускользания от иммунного надзора, иммунной стимуляции, иммунной дезактивации, дисперсии биопленки, устойчивости бактерий к фагам, антибиотикочувствительности бактерий, модуляции факторов вирулентности и направленного расщепления или редактирования генома хозяина.

В некоторых аспектах настоящее раскрытие обеспечивает сконструированную вирусную композицию, содержащую рекомбинантную нуклеиновую кислоту, способную продуцировать не встречающиеся в природе вирусные частицы по меньшей мере с одним улучшенным вирусным свойством по сравнению с несконструированной вирусной нуклеиновой кислотой. В некоторых примерах улучшенное вирусное свойство выбирают из группы, состоящей из круга хозяев, вирусного литического цикла, адсорбции, прикрепления, инъекции, репликации и сборки, лизиса, величины выхода фага, ускользания от иммунного надзора, иммунной стимуляции, иммунной дезактивации, дисперсии биопленки, устойчивости бактерий к фагам, антибиотикочувствительности бактерий, модуляции факторов вирулентности и направленного расщепления или редактирования генома хозяина. В некоторых примерах сконструированную вирусную нуклеиновую кислоту в соответствии с настоящим изобретением получают любым из этапов, описанных в настоящем документе, способов.

Данный способ может быть использован для изменения нуклеотида, гена или цельной геномной области. Например, как описано в приведенных ниже примерах, данный способ, как было показано, заменяет ген *gp18 LKD16* на *LUZ19*, что приводит к улучшению чувствительности к кругу вирусных хозяев. Кроме того, данный способ может быть использован для вставки единственной мутации в комплекс трубок вируса для улучшения репликации вируса. Данный способ также может быть использован для конструирования антимикробных пептидов; пиоцинов; ЭПС-деполимераз; белков, ингибирующих CRISPR/Cas; хвостовых волокон из бактериофага; репортерных генов (т.е. *Lux*, *GFP*); генов для подавления кворума; нуклеаз; нуклеаз TALEN; белков системы CRISPR типа I, типа II, типа III, типа IV, типа V и типа VI (т.е. *Cas9*); CRISPR РНК, факторов транскрипции и иммуномодулирующих факторов человека в бактериофаге для улучшения активности бактериофага в бактериофаговой терапии или связанных с ним применениях. Данные элементы могут быть функционально связаны с нативными или гетерологичными регуляторными элементами, такими как нативный промотор, гетерологичный промотор, индуцируемый промотор или любой их комбинацией.

В некоторых вариантах реализации настоящее изобретение относится к сконструированному вирусу, содержащему сконструированную вирусную нуклеиновую кислоту, способную при введении в клетку-хозяина продуцировать вирусные частицы, не встречающиеся в природе, с двумя или более улучшенными вирусными свойствами по сравнению с несконструированной вирусной нуклеиновой кислотой. В некоторых аспектах полученные вирусные частицы имеют по меньшей мере три улучшенных вирусных свойства. В некоторых аспектах каждое улучшенное вирусное свойство выбрано из группы, состоящей из круга хозяев, вирусного литического цикла, адсорбции, прикрепления, инъекции, репликации и сборки, лизиса, величины выхода фага, ускользания от иммунного надзора, иммунной стимуляции, иммунной дезактивации, дисперсии биопленки, устойчивости бактерий к фагам, антибиотикочувствительности бактерий, модуляции факторов вирулентности и направленного расщепления или редактирования генома хозяина.

В некоторых вариантах реализации настоящее изобретение относится к сконструированному вирусу, содержащему сконструированную вирусную нуклеиновую кислоту. В некоторых аспектах сконструированная вирусная нуклеиновая кислота представляет собой сконструированный вирусный геном. В некоторых аспектах сконструированный вирусный геном представляет собой сконструированный геном бактериофага. В некоторых аспектах у сконструированного бактериофага по меньшей мере одним из улучшенных вирусных свойств является круг хозяев.

В некоторых вариантах реализации настоящее раскрытие обеспечивает сконструированный вирус с двумя или более улучшенными вирусными свойствами, который содержит сконструированную вирусную нуклеиновую кислоту. В некоторых аспектах каждое улучшенное вирусное свойство является результатом по меньшей мере одной модификации в сконструированной вирусной нуклеиновой кислоте. В некоторых аспектах по меньшей мере одно улучшенное вирусное свойство является результатом по меньшей мере двух модификаций в сконструированной вирусной нуклеиновой кислоте. В некоторых аспектах модификации, содержащиеся в сконструированной вирусной нуклеиновой кислоте, являются результатом одного этапа конструирования. В некоторых аспектах модификации, содержащиеся в сконструированной вирусной нуклеиновой кислоте, являются результатом неоднократно повторяемых этапов конструирования.

В некоторых вариантах реализации настоящее раскрытие обеспечивает сконструированный вирус с





меньшей мере 70%, по меньшей мере 75%, по меньшей мере 80%, по меньшей мере 85%, по меньшей мере 90%, по меньшей мере 95, 100% или полную идентичность с SEQ ID NO: 49 или замену последовательности нуклеиновой кислоты, содержащейся в последовательности нуклеиновой кислоты, кодирующей аминокислотную последовательность, имеющую по меньшей мере 50%, по меньшей мере 60%, по меньшей мере 65%, по меньшей мере 70%, по меньшей мере 75%, по меньшей мере 80%, по меньшей мере 85%, по меньшей мере 90%, по меньшей мере 95, 100% или полную идентичность с SEQ ID NO: 49 с гетерологичной молекулой нуклеиновой кислоты. В некоторых аспектах гетерологичная молекула нуклеиновой кислоты содержит гетерологичную последовательность нуклеиновой кислоты, кодирующую аминокислотную последовательность, имеющую по меньшей мере 50%, по меньшей мере 60%, по меньшей мере 65%, по меньшей мере 70%, по меньшей мере 75%, по меньшей мере 80%, по меньшей мере 85%, по меньшей мере 90%, по меньшей мере 95, 100% или полную идентичность с последовательностью, выбранной из группы, состоящей из SEQ ID NO: 37, SEQ ID NO: 39, SEQ ID NO: 40, SEQ ID NO: 41, SEQ ID NO: 43, SEQ ID NO: 44, SEQ ID NO: 45, SEQ ID NO: 46 и SEQ ID NO: 47.

В некоторых аспектах сконструированная вирусная нуклеиновая кислота содержит гетерологичную последовательность нуклеиновой кислоты, функционально связанную с промотором, содержащим последовательность нуклеиновой кислоты, содержащуюся в SEQ ID NO: 21 или ее часть.

В некоторых аспектах сконструированная вирусная нуклеиновая кислота содержит гетерологичную последовательность нуклеиновой кислоты, функционально связанную с терминатором, содержащим последовательность нуклеиновой кислоты, содержащуюся в SEQ ID NO: 22, или ее часть.

В некоторых вариантах реализации настоящее раскрытие обеспечивает способ получения сконструированного вируса, представляющего интерес, обладающего двумя или более необходимыми вирусными свойствами, включающий: (а) обеспечение первого вирусного генома; и (b) конструирование второго вирусного генома путем объединения по меньшей мере одного фрагмента первого вирусного генома с по меньшей мере одной молекулой восстанавливающей нуклеиновой кислоты, так что полученный второй вирусный геном содержит по меньшей мере одну модификацию по сравнению с первым вирусным геномом и при этом при введении в клетку-хозяина второй вирусный геном способен продуцировать вирусные частицы с двумя или более улучшенными вирусными свойствами. В некоторых аспектах способ, описанный в настоящем документе, дополнительно включает (с) повторение этапов (а)-(b) в одном или более циклах. В некоторых аспектах каждое улучшенное вирусное свойство выбрано из группы, состоящей из круга хозяев, вирусного литического цикла, адсорбции, прикрепления, инъекции, репликации и сборки, лизиса, величины выхода фага, ускользания от иммунного надзора, иммунной стимуляции, иммунной дезактивации, дисперсии биопленки, устойчивости бактерий к фагам, антибиотикочувствительности бактерий, модуляции факторов вирулентности и направленного расщепления или редактирования генома хозяина.

В некоторых вариантах реализации настоящее раскрытие обеспечивает способ получения сконструированного вируса, представляющего интерес, обладающего двумя или более необходимыми вирусными свойствами, описанными в настоящем документе. В некоторых аспектах конструирование второго вирусного генома на стадии (b) дополнительно включает: (1) расщепление *in vitro* области первого вирусного генома с использованием эндонуклеазы; и (2) сборку по меньшей мере одного фрагмента расщепленного первого вирусного генома по меньшей мере с одной молекулой восстанавливающей нуклеиновой кислоты. В некоторых аспектах первый вирусный геном выделяется из вирусных частиц. В некоторых аспектах первый вирусный геном или по меньшей мере одну восстанавливающую молекулу нуклеиновой кислоты синтезируют *de novo*. В некоторых аспектах синтез *de novo* включает объединение химически синтезированных молекул нуклеиновой кислоты, последовательностей нуклеиновой кислоты с амплифицированной ПЦР, расщепленных фрагментов выделенных молекул нуклеиновой кислоты или любой их комбинации. В некоторых аспектах первый вирусный геном или по меньшей мере одну восстанавливающую молекулу нуклеиновой кислоты амплифицируют перед расщеплением *in vitro*.

В некоторых вариантах реализации настоящее раскрытие обеспечивает способ получения сконструированного вируса, представляющего интерес, обладающего двумя или более желательными вирусными свойствами, описанными в настоящем документе. В некоторых аспектах первый вирусный геном составляет по меньшей мере 18 т.п.н. В некоторых аспектах первый вирусный геном составляет от по меньшей мере 2 т.п.н. до по меньшей мере 4 млн п.н. В некоторых аспектах первый вирусный геном составляет от по меньшей мере 18 т.п.н. до по меньшей мере 4 млн п.н. В некоторых аспектах первый вирусный геном составляет по меньшей мере 5 т.п.н., по меньшей мере 10 т.п.н., по меньшей мере 15 т.п.н., по меньшей мере 18 т.п.н., по меньшей мере 20 т.п.н., по меньшей мере 25 т.п.н., по меньшей мере 30 т.п.н., по меньшей мере 35 т.п.н., по меньшей мере 40 т.п.н., по меньшей мере 45 т.п.н., по меньшей мере 50 т.п.н., по меньшей мере 55 т.п.н., по меньшей мере 60 т.п.н., по меньшей мере 65 т.п.н., по меньшей мере 70 т.п.н., по меньшей мере 75 т.п.н., по меньшей мере 80 т.п.н., по меньшей мере 85 т.п.н., по меньшей мере 90 т.п.н., по меньшей мере 100 т.п.н., по по меньшей мере 125 т.п.н., по меньшей мере 150 т.п.н., по меньшей мере 175 т.п.н., по меньшей мере 200 т.п.н., по меньшей мере 250 т.п.н., по меньшей мере 300 т.п.н., по меньшей мере 400 т.п.н., по меньшей мере 500 т.п.н., по меньшей мере 600 т.п.н., по меньшей мере 700 т.п.н., по меньшей мере 800 т.п.н., по меньшей мере 900 т.п.н., по меньшей мере 1 млн

п.н., по меньшей мере 1,5 млн п.н., по меньшей мере 2 млн п.н., по меньшей мере 2,5 млн п.н., по меньшей мере 3 млн п.н. или по меньшей мере 3,5 млн п.н.

В некоторых вариантах реализации настоящее раскрытие обеспечивает способ получения сконструированного вируса, представляющего интерес, обладающего двумя или более желательными вирусными свойствами, описанными в настоящем документе. В некоторых аспектах сборку выполняют *in vitro* или *in vivo*. В некоторых аспектах сборку выполняют *in vitro* при помощи смеси, содержащей: (a) выделенную нетермостабильную 5'-3'-экзонуклеазу, которая не обладает 3'-экзонуклеазной активностью; (b) краудинг-агент; (c) выделенную термостабильную невымесняющую ДНК-полимеразу с 3'-экзонуклеазной активностью или смесь указанной ДНК-полимеразы со второй ДНК-полимеразой, которая не обладает 3'-экзонуклеазной активностью; (d) выделенную термостабильную лигазу; (e) смесь дНТФ; и (f) подходящий буфер в условиях, эффективных для вставки фрагмента в расщепленную вирусную нуклеиновую кислоту с образованием рекомбинантной нуклеиновой кислоты, содержащей сконструированный вирусный геном.

В некоторых вариантах реализации настоящее раскрытие обеспечивает способ получения сконструированного вируса, представляющего интерес, обладающего двумя или более желательными вирусными свойствами, описанными в настоящем документе. В некоторых аспектах сборку выполняют *in vitro* или *in vivo*. В некоторых аспектах сборку выполняют *in vivo* в клетке-хозяина.

В некоторых вариантах реализации настоящее раскрытие обеспечивает способ получения сконструированного вируса, представляющего интерес, обладающего двумя или более желательными вирусными свойствами, описанными в настоящем документе. В некоторых аспектах эндонуклеаза представляет собой РНК-направляемую нуклеазу. В некоторых аспектах способ дополнительно включает одну или две направляющие РНК. В некоторых аспектах РНК-направляемая нуклеаза представляет собой Cas9 или фермент, полученный из Cas9. В некоторых аспектах направляющие РНК включают: 1) химерную гРНК или; 2) sgРНК и tracrРНК.

В некоторых вариантах реализации настоящее раскрытие обеспечивает способ получения сконструированного вируса, представляющего интерес, обладающего двумя или более желательными вирусными свойствами, описанными в настоящем документе. В некоторых аспектах эндонуклеазу инактивируют нагреванием или удаляют. В некоторых аспектах система расщепления *in vitro* дополнительно содержит спермидин.

В некоторых вариантах реализации настоящее раскрытие обеспечивает способ получения сконструированного вируса, представляющего интерес, обладающего двумя или более желательными вирусными свойствами, описанными в настоящем документе. В некоторых аспектах способ дополнительно включает трансформацию сконструированного вирусного генома в клетку-хозяина. В некоторых аспектах способ дополнительно включает использование упаковочного набора *in vitro* для упаковки сконструированного вирусного генома в вирусные частицы.

В некоторых вариантах реализации настоящее раскрытие обеспечивает сконструированный вирус, полученный любым из способов, описанных в настоящем документе.

В некоторых вариантах реализации настоящее раскрытие обеспечивает композиции любого из сконструированных вирусов, описанных в настоящем документе, получаемых любым из способов конструирования, описанных в настоящем документе.

В некоторых вариантах реализации настоящее раскрытие обеспечивает набор для конструирования молекул вирусной нуклеиновой кислоты, включающий очищенную рекомбинантную РНК-направляемую нуклеазу; выделенную нетермостабильную 5'-3'-экзонуклеазу, которая не обладает 3'-экзонуклеазной активностью; краудинг-агент; выделенную термостабильную невымесняющую ДНК-полимеразу с 3'-экзонуклеазной активностью или смесь указанной ДНК-полимеразы со второй ДНК-полимеразой, которая не обладает 3'-экзонуклеазной активностью; выделенную термостабильную лигазу; смесь дНТФ; и подходящий буфер. В некоторых аспектах набор дополнительно содержит специально разработанные гидовые РНК. В некоторых аспектах набор дополнительно содержит специально разработанные синтезированные молекулы нуклеиновой кислоты, предназначенные для использования в качестве вставленного фрагмента ДНК в реакции сборки. В некоторых аспектах набор дополнительно содержит компетентные клетки-хозяева для трансформации. В некоторых аспектах набор дополнительно содержит выделенные вирусные геномные нуклеиновые кислоты.

В некоторых аспектах настоящее раскрытие обеспечивает систему сконструированной вирусной нуклеиновой кислоты *in vitro*, содержащую выделенную вирусную нуклеиновую кислоту, нуклеазу, рекомбинантную РНК-направляемую нуклеазу, по меньшей мере одну нацеливающую РНК и фрагмент нуклеиновой кислоты, который должен быть вставлен в сайт расщепления выделенной нуклеиновой кислоты. В некоторых примерах система такова, что рекомбинантная РНК-направляемая нуклеаза и по меньшей мере одна нацеливающая РНК образуют комплекс, способный расщеплять выделенную вирусную нуклеиновую кислоту. В некоторых примерах система дополнительно содержит спермидин. В некоторых примерах система дополнительно содержит выделенную нетермостабильную 5'-3'-экзонуклеазу, которая не обладает 3'-экзонуклеазной активностью; краудинг-агент; выделенную термостабильную невымесняющую ДНК-полимеразу с 3'-экзонуклеазной активностью или смесь указанной ДНК-полиме-

разы со второй ДНК-полимеразой, которая не обладает 3'-экзонуклеазной активностью; выделенную термостабильную лигазу; смесь дНТФ; и подходящий буфер, в котором система находится в условиях, эффективных для вставки фрагмента нуклеиновой кислоты в выделенную вирусную нуклеиновую кислоту в месте расщепления РНК-направляемой нуклеазы с образованием рекомбинантной вирусной нуклеиновой кислоты.

В некоторых аспектах система, описанная в настоящем документе, такова, что рекомбинантная вирусная нуклеиновая кислота способна продуцировать вирусные частицы, не встречающиеся в природе, по меньшей мере с одним улучшенным вирусным свойством по сравнению с несконструированной вирусной нуклеиновой кислотой. В некоторых примерах улучшенное вирусное свойство выбирают из группы, состоящей из круга хозяев, вирусного литического цикла, адсорбции, прикрепления, инъекции, репликации и сборки, лизиса, величины выхода фага, ускользания от иммунного надзора, иммунной стимуляции, иммунной дезактивации, дисперсии биопленки, устойчивости бактерий к фагам, антибиотикочувствительности бактерий, модуляции факторов вирулентности и направленного расщепления или редактирования генома хозяина.

В некоторых аспектах в системе, описанной в данном документе, РНК-направляемая нуклеаза представляет собой Cas9 или фермент, полученный из Cas9. В некоторых примерах РНК-направляемую нуклеазу инактивируют или удаляют после расщепления.

В некоторых аспектах способ, описанный в настоящем документе, используют в качестве способа коррекции ошибок для коррекции последовательностей в выделенных нуклеиновых кислотах. Стандартные методы коррекции ошибок основаны на ПЦР, имеют две неотъемлемые проблемы: 1) ПЦР может вводить дополнительные нежелательные мутации в нуклеиновую кислоту; и 2) ПЦР в этом контексте имеет ограничение по размеру приблизительно 5 т.п.н. Поэтому стандартные методы коррекции ошибок на основе ПЦР не могут быть надежно выполнены на плаزمидях размером более 5 т.п.н., либо в результате мутаций, вызванных ПЦР, либо в случае невозможности усиления. Описанный в настоящем документе способ конструирования *in vitro* последовательности нуклеиновой кислоты обходит необходимость ПЦР-амплификации, которая устраняет ограничение по размеру и исключает возможность возникновения мутаций, вызванных ПЦР.

В некоторых аспектах настоящее раскрытие обеспечивает способ конструирования *in vitro* последовательности нуклеиновой кислоты, включающей выделение нуклеиновой кислоты; расщепление *in vitro* области нуклеиновой кислоты с использованием РНК-направляемой нуклеазы; и сборку рекомбинантной нуклеиновой кислоты путем вставки фрагмента ДНК или РНК в расщепленную нуклеиновую кислоту. В одном аспекте расщепление *in vitro* представляет собой ферментативное расщепление, основанное на РНК. В другом аспекте ферментативное расщепление проводят с использованием Cas9 или фермента, полученного из Cas9. В дополнительном аспекте расщепление дополнительно включает нацеливающие РНК. В другом аспекте система расщепления дополнительно содержит спермидин. В конкретном аспекте нацеливающие РНК представляют собой гРНК, сгРНК и/или tracrРНК. В дополнительном аспекте после расщепления РНК-направляемую нуклеазу инактивируют стандартными методами, таким как воздействием тепла, например, такими как по меньшей мере 80°C. Дополнительно или альтернативно, РНК-направляемую нуклеазу удаляют стандартными способами, такими как, например, фенольно-хлороформной экстракцией.

В некоторых аспектах настоящее раскрытие обеспечивает способ конструирования *in vitro* последовательности нуклеиновой кислоты, включающий выделение нуклеиновой кислоты; расщепления *in vitro* области нуклеиновой кислоты с использованием РНК-направляемой нуклеазы; и сборку рекомбинантной нуклеиновой кислоты путем вставки фрагмента ДНК или РНК в расщепленную нуклеиновую кислоту. В некоторых примерах сборку проводят *in vitro* в одном сосуде со смесью компонентов, включающих: (a) выделенную нетермостабильную 5'-3'-экзонуклеазу, которая не обладает 3'-экзонуклеазной активностью; (b) краудинг-агент; (c) выделенную термостабильную невытесняющую ДНК-полимеразу с 3'-экзонуклеазной активностью или смесью указанной ДНК-полимеразы со второй ДНК-полимеразой, которая не обладает 3'-экзонуклеазной активностью; (d) выделенную термостабильную лигазу; (e) смесь дНТФ; и (f) подходящий буфер в условиях, эффективных для вставки фрагмента в расщепленную вирусную нуклеиновую кислоту с образованием рекомбинантной нуклеиновой кислоты. В некоторых аспектах экзонуклеаза представляет собой экзонуклеазу T5, и приведение в контакт происходит в изотермических условиях, и/или краудинг-агент представляет собой ПЭГ, и/или невытесняющая ДНК-полимераза представляет собой ДНК-полимеразу Phusion™ или ДНК-полимеразу VENT® и/или лигаза представляет собой Taq-лигазу. В некоторых примерах сборку *in vitro* выполняют при помощи одноступенчатой или изотермической сборки методом Gibson Assembly. В некоторых примерах сборку *in vitro* выполняют при помощи двухступенчатой сборки методом Gibson Assembly.

В некоторых аспектах настоящее раскрытие обеспечивает способ конструирования *in vitro* последовательности нуклеиновой кислоты, содержащей РНК-направляемую нуклеазу. В некоторых примерах РНК-направляемая нуклеаза представляет собой Cas9 II типа. В некоторых примерах РНК-направляемая нуклеаза представляет собой Cas9 или фермент, полученный из Cas9. В некоторых примерах РНК-направляемая нуклеаза представляет собой выделенный рекомбинантный фермент Cas9 или Cas9. В не-

которых примерах существует по меньшей мере одна нацеливающая РНК. В некоторых примерах существуют две нацеливающие РНК. В некоторых примерах нацеливающая РНК представляет собой химерную гидовую РНК (гРНК) или набор сгРНК и трасРНК. В некоторых примерах в реакции расщепления *in vitro* используют две гРНК. В некоторых примерах в реакции расщепления *in vitro* используют два набора сгРНК и трасРНК.

В некоторых аспектах настоящее раскрытие обеспечивает способ конструирования *in vitro* последовательности нуклеиновой кислоты, включающий этап расщепления *in vitro*. В некоторых примерах после расщепления РНК-направляемую нуклеазу инактивируют стандартными методами, такими как воздействие теплом, таким как, например, по меньшей мере 80°. В некоторых примерах после расщепления РНК-направляемую нуклеазу удаляют фенольно-хлороформной экстракцией. В некоторых примерах после расщепления РНК-направляемую нуклеазу удаляют другими способами экстракции, хорошо известными в данной области техники.

В некоторых аспектах настоящее раскрытие обеспечивает способ конструирования *in vitro* последовательности нуклеиновой кислоты, приводящей к образованию сконструированной нуклеиновой кислоты. В некоторых примерах сконструированную нуклеиновую кислоту затем трансформируют в клетку-хозяин. В некоторых примерах клетка-хозяин представляет собой *E.coli*, *P.aeruginosa*, *S.cerevisiae*, *V.natriegens*, *B.subtilis* или другой микроорганизм, хорошо известный в данной области техники. В некоторых примерах трансформацию выполняют путем теплового шока, электропорации, биолистики, бомбардировки частиц, конъюгации, трансдукции, липофекции или другого известного метода, хорошо известного в данной области техники.

В некоторых аспектах настоящее раскрытие обеспечивает способ конструирования *in vitro* последовательности нуклеиновой кислоты, содержащей выделенную нуклеиновую кислоту. В некоторых примерах нуклеиновая кислота представляет собой полный геном, выделенный из клетки-хозяина. В некоторых примерах клетка-хозяин представляет собой *E.coli*, *S.cerevisiae*, *B.subtilis*, *V.natriegens*, *P.aeruginosa* или другой известный микроорганизм. В некоторых примерах нуклеиновая кислота представляет собой плазмиду. В некоторых примерах плазмиду выделяют из клетки-хозяина. В некоторых примерах нуклеиновую кислоту, представляющую интерес, клонировали в плазмиду, трансформировали в клетку-хозяин и выделяли до конструирования по способу, описанному в настоящем документе.

В некоторых аспектах настоящее раскрытие обеспечивает способ конструирования *in vitro* последовательности нуклеиновой кислоты, включающей выделение нуклеиновой кислоты. В некоторых примерах выделенная нуклеиновая кислота представляет собой геном или плазмиду. В некоторых примерах выделенный геном или плазмиду составляет по меньшей мере 6 т.п.н., по меньшей мере 7 т.п.н., по меньшей мере 8 т.п.н., по меньшей мере 9 т.п.н., по меньшей мере 10 т.п.н., по меньшей мере 12 т.п.н., по меньшей мере 15 т.п.н., по меньшей мере 20 т.п.н., по меньшей мере 25 т.п.н. или по меньшей мере 28 т.п.н. В некоторых примерах выделенный геном или плазмиду составляет от 6 т.п.н. до 1 млн п.н. В некоторых примерах выделенный геном или плазмиду составляет: от 6 до 10 т.п.н., от 8 до 15 т.п.н., от 12 до 20 т.п.н., от 15 до 22 т.п.н., от 20 до 25 т.п.н., от 22 до 28 т.п.н., от 25 до 30 т.п.н., от 25 до 50 т.п.н. или от 40 до 100 т.п.н.

Дополнительно или, альтернативно, к любому из вышеописанных вариантов реализации раскрытие включает следующие варианты реализации изобретения.

Вариант реализации изобретения 1 представляет собой сконструированный вирус, содержащий сконструированную вирусную нуклеиновую кислоту, способную при введении в клетку-хозяина продуцировать не встречающиеся в природе вирусные частицы с двумя или более, или, возможно, тремя или более улучшенными вирусными свойствами по сравнению с вирусными частицами, продуцируемыми при введении несконструированной вирусной нуклеиновой кислоты в клетку-хозяина.

Вариант реализации изобретения 2 представляет собой сконструированный вирус согласно варианту реализации изобретения 1, при этом каждое улучшенное вирусное свойство выбрано из группы, состоящей из круга хозяев, вирусного литического цикла, адсорбции, прикрепления, инъекции, репликации и сборки, лизиса, величины выхода фага, ускользания от иммунного надзора, иммунной стимуляции, иммунной дезактивации, дисперсии биопленки, устойчивости бактерий к фагу, антибиотической сенсibilизации бактерий, модуляции факторов вирулентности и направленного расщепления и редактирования генома хозяина.

Вариант реализации изобретения 3 представляет собой сконструированный вирус согласно варианту реализации изобретения 1 или 2, причем вирусная нуклеиновая кислота представляет собой одну или более из следующих вирусных нуклеиновых кислот: вирусный геном, фрагмент вирусного генома, генома бактериофага, фрагмент генома бактериофага, геном литического бактериофага, фрагмент генома литического бактериофага или любую их комбинацию.

Вариант реализации изобретения 4 представляет собой сконструированный вирус любого из вариантов реализации изобретения 1-3, причем сконструированная вирусная нуклеиновая кислота представляет собой геном бактериофага и необязательно при этом по меньшей мере одно из улучшенных вирусных свойств представляет собой круг хозяев.

Вариант реализации изобретения 5 представляет собой сконструированный вирус любого из вари-



антов реализации изобретения 1-4, причем выполняется по меньшей мере одно из следующего: 1) каждое улучшенное вирусное свойство является результатом по меньшей мере одной модификации в сконструированной вирусной нуклеиновой кислоте; 2) по меньшей мере одно улучшенное вирусное свойство является результатом по меньшей мере двух модификаций в сконструированной вирусной нуклеиновой кислоте; 3) модификации, содержащиеся в вирусной нуклеиновой кислоте, являются результатом одного этапа конструирования, 4) модификации, содержащиеся в сконструированной вирусной нуклеиновой кислоте, являются результатом итерации этапов конструирования или 5) любой их комбинации.

Вариант реализации изобретения 6 представляет собой сконструированный вирус любого из вариантов реализации изобретения 1-5, в котором по меньшей мере одна из модификаций находится внутри:

1) последовательности нуклеиновой кислоты, имеющей по меньшей мере 50%, по меньшей мере 60%, по меньшей мере 65%, по меньшей мере 70%, по меньшей мере 75%, по меньшей мере 80%, по меньшей мере 85%, по меньшей мере 90%, по меньшей мере 95, 100% или полную идентичность с последовательностью, содержащейся в SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 50 или SEQ ID NO: 25, или

2) последовательности нуклеиновой кислоты, кодирующей аминокислотную последовательность, имеющую по меньшей мере 50%, по меньшей мере 60%, по меньшей мере 65%, по меньшей мере 70%, по меньшей мере 75%, по меньшей мере 80%, по меньшей мере 85%, по меньшей мере 90%, по меньшей мере 95, 100% или полную идентичность с SEQ ID NO: 34, SEQ ID NO: 35, SEQ ID NO: 36, SEQ ID NO: 5, SEQ ID NO: 48 или SEQ ID NO: 49, или

3) любой их комбинации.

Вариант реализации изобретения 7 представляет собой сконструированный вирус согласно любому из вариантов реализации изобретения 1-6, причем сконструированная вирусная нуклеиновая кислота содержит сконструированный вирусный геном, содержащий весь или часть вирусного генома, имеющего по меньшей мере 50%, по меньшей мере 60%, по меньшей мере 65%, по меньшей мере 70%, по меньшей мере 75%, по меньшей мере 80%, по меньшей мере 85%, по меньшей мере 90%, по меньшей мере 95, 100% или полную идентичность с геномом LUZ19.

Вариант реализации изобретения 8 представляет собой сконструированный вирус любого из вариантов реализации изобретения 1-7, причем сконструированный вирусный геном дополнительно содержит по меньшей мере одно из следующего:

1) весь или часть гетерологичного гена gp18 и, необязательно, в котором гетерологичный ген gp18 имеет по меньшей мере 50%, по меньшей мере 60%, по меньшей мере 65%, по меньшей мере 70%, по меньшей мере 75%, по меньшей мере 80%, по меньшей мере 85%, по меньшей мере 90%, по меньшей мере 95, 100% или полную идентичность с SEQ ID NO: 26;

2) весь или часть гетерологичного гена gp18 и, необязательно, в котором гетерологичный ген gp18 кодирует аминокислотную последовательность с по меньшей мере 50%, по меньшей мере 60%, по меньшей мере 65%, по меньшей мере 70%, по меньшей мере 75%, по меньшей мере 80%, по меньшей мере 85%, по меньшей мере 90%, по меньшей мере 95, 100% или полной идентичностью с SEQ ID NO: 38;

3) весь или часть сконструированного гена gp34 и, необязательно, в котором гетерологичный ген gp34 кодирует аминокислотную последовательность, содержащую мутацию в положении, соответствующем положению аминокислоты SEQ ID NO: 5, или необязательно, в котором гетерологичный ген gp34 имеет по меньшей мере 50%, по меньшей мере 60%, по меньшей мере 65%, по меньшей мере 70%, по меньшей мере 75%, по меньшей мере 80%, по меньшей мере 85%, по меньшей мере 90%, по меньшей мере 95, 100% или полную идентичность с SEQ ID NO: 4;

4) модификацию в одной или более последовательностях, имеющих по меньшей мере 50%, по меньшей мере 60%, по меньшей мере 65%, по меньшей мере 70%, по меньшей мере 75%, по меньшей мере 80%, по меньшей мере 85%, по меньшей мере 90%, по меньшей мере 95, 100% или полную идентичность с последовательностью, выбранной из группы, состоящей из SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3 и SEQ ID NO: 50,

и, необязательно, модификацию в каждой последовательности, имеющей по меньшей мере 50%, по меньшей мере 60%, по меньшей мере 65%, по меньшей мере 70%, по меньшей мере 75%, по меньшей мере 80%, по меньшей мере 85%, по меньшей мере 90%, по меньшей мере 95, 100% или полную идентичность с SEQ ID NO: 1, последовательность, имеющую по меньшей мере 50%, по меньшей мере 60%, по меньшей мере 65%, по меньшей мере 70%, по меньшей мере 75%, по меньшей мере 80%, по меньшей мере 85%, по меньшей мере 90%, по меньшей мере 95, 100% или полную идентичность SEQ ID NO: 2, последовательность, имеющую по меньшей мере 50%, по меньшей мере 60%, по меньшей мере 65%, по меньшей мере 70%, по меньшей мере 75%, по меньшей мере 80%, по меньшей мере 85%, по меньшей мере 90%, по меньшей мере 95, 100% или полную идентичность с SEQ ID NO: 3, и последовательность, имеющую по меньшей мере 50%, по меньшей мере 60%, по меньшей мере 65%, по меньшей мере 70%, по меньшей мере 75%, по меньшей мере 80%, по меньшей мере 85%, по меньшей мере 90%, по меньшей мере 95, 100% или полную идентичность с SEQ ID NO: 50,

и, необязательно, в которой модификации включают замену G на A в положении, соответствующем положению 50 нуклеиновой кислоты SEQ ID NO: 1, замену G на T в положении, соответствующем по-



по меньшей мере 75%, по меньшей мере 80%, по меньшей мере 85%, по меньшей мере 90%, по меньшей мере 95, 100% или полную идентичность с последовательностью, выбранной из группы, состоящей из SEQ ID NO: 37, SEQ ID NO: 39, SEQ ID NO: 40, SEQ ID NO: 41, SEQ ID NO: 43, SEQ ID NO: 44, SEQ ID NO: 45, SEQ ID NO: 46 и SEQ ID NO: 47,

8) любую их комбинацию.

Вариант реализации изобретения 9 представляет собой сконструированный вирус по любому из вариантов реализации изобретения 1-8, причем сконструированная вирусная нуклеиновая кислота содержит гетерологичную последовательность нуклеиновой кислоты, функционально связанную с 1) промотором, содержащим последовательность нуклеиновой кислоты, содержащуюся в SEQ ID NO: 21, или ее часть, 2) терминатором, содержащим последовательность нуклеиновой кислоты, содержащуюся в SEQ ID NO: 22, или ее часть, или 3) любую их комбинацию.

Вариант реализации изобретения 10 представляет собой способ получения сконструированного вируса, имеющего два или более желательных вирусных свойства, включающий: (a) обеспечение первого вирусного генома; и (b) создание сконструированного вирусного генома путем объединения по меньшей мере одного фрагмента первого вирусного генома с по меньшей мере одной молекулой восстанавливающей нуклеиновой кислоты для получения второго вирусного генома, содержащего по меньшей мере одну модификацию по сравнению с первым вирусным геномом; при этом второй вирусный геном, будучи введенным в клетку-хозяин, способен продуцировать вирусные частицы с двумя или более улучшенными вирусными свойствами и необязательно (c) повторять этапы (a)-(b) в одной или более итерациях.

Вариант реализации изобретения 11 представляет собой вариант реализации изобретения 10, в котором каждое улучшенное вирусное свойство выбрано из группы, состоящей из круга хозяев, вирусного литического цикла, адсорбции, прикрепления, инъекции, репликации и сборки, лизиса, величины выхода фага, ускользания от иммунного надзора, иммунной стимуляции, иммунной дезактивации, дисперсии биопленки, устойчивости бактерий к фагу, антибиотической сенсбилизации бактерий, модуляции факторов вирулентности и направленного расщепления или редактирования генома хозяина.

Вариант реализации изобретения 12 представляет собой способ согласно варианту реализации изобретения 10 или 11, в котором создание сконструированного вирусного генома на стадии (b) включает: (1) расщепление *in vitro* области первого вирусного генома с использованием эндонуклеазы; и (2) сборку по меньшей мере одного фрагмента расщепленного первого вирусного генома по меньшей мере с одной молекулой восстанавливающей нуклеиновой кислоты.

Вариант реализации изобретения 13 представляет собой способ согласно любому из вариантов реализации изобретения 10-12, в котором выполняется по меньшей мере один из следующих элементов: 1) первый вирусный геном выделяют из вирусных частиц, 2) первую вирусную и/или по меньшей мере одну восстанавливающую молекулу нуклеиновой кислоты синтезируют *de novo* и, необязательно, в котором синтез *de novo* включает объединение химически синтезированных молекул нуклеиновой кислоты, ПЦР-амплифицированных последовательностей нуклеиновых кислот, расщепленных фрагментов изолированных молекул нуклеиновой кислоты или любой их комбинации, 3) первый вирусный геном и/или по меньшей мере одну восстанавливающую молекулу нуклеиновой кислоты амплифицируют перед расщеплением *in vitro* или 4) любую их комбинацию.

Вариант реализации изобретения 14 представляет собой способ согласно любому из вариантов реализации изобретения 10-13, причем первый вирусный геном составляет по меньшей мере одно из следующего:

1) по меньшей мере 3 т.п.н., по меньшей мере 10 т.п.н., по меньшей мере 18 т.п.н., по меньшей мере 25 т.п.н. или по меньшей мере 30 т.п.н.;

2) по меньшей мере 18 т.п.н.;

3) от по меньшей мере 2 т.п.н. до по меньшей мере 4 млн п.н.;

4) от по меньшей мере 18 т.п.н. до по меньшей мере 4 млн п.н.; или

5) по меньшей мере 5 т.п.н., по меньшей мере 10 т.п.н., по меньшей мере 15 т.п.н., по меньшей мере 18 т.п.н., по меньшей мере 20 т.п.н., по меньшей мере 25 т.п.н., по меньшей мере 30 т.п.н., по меньшей мере 35 т.п.н., по меньшей мере 40 т.п.н., по меньшей мере 45 Kb, по меньшей мере 50 т.п.н., по меньшей мере 55 т.п.н., по меньшей мере 60 т.п.н., по меньшей мере 65 т.п.н., по меньшей мере 70 т.п.н., по меньшей мере 75 т.п.н., по меньшей мере 80 т.п.н., по меньшей мере 85 т.п.н., по меньшей мере 90 т.п.н., по меньшей мере 100 т.п.н., по меньшей мере 125 т.п.н., по меньшей мере 150 т.п.н., по меньшей мере 175 т.п.н., по меньшей мере 200 т.п.н., по меньшей мере 250 т.п.н., по меньшей мере 300 т.п.н., по меньшей мере 400 т.п.н., по меньшей мере 500 т.п.н., по меньшей мере 600 т.п.н., по меньшей мере 700 т.п.н., по меньшей мере 800 т.п.н., по меньшей мере 900 т.п.н., по меньшей мере 1 млн п.н., по меньшей мере 1,5 млн п.н., по меньшей мере 2 млн п.н., по меньшей мере 2,5 млн п.н., по меньшей мере 3 млн п.н. или по меньшей мере 3,5 млн п.н.

Вариант реализации изобретения 15 представляет собой способ согласно любому из вариантов реализации изобретения 10-14, в котором сборку выполняют *in vitro* и, необязательно, в котором сборку выполняют *in vitro* с помощью смеси, содержащей: (a) выделенную 5'-3'-экзонуклеазу, которая не обладает 3'-экзонуклеазной активностью, которая необязательно не термостабильна; (b) необязательно краудинг-

агент; (с) выделенную невытесняющую ДНК-полимеразу с 3'-экзонуклеазной активностью, которая, обязательно, является термостабильной, или смесь указанной ДНК-полимеразы со второй ДНК-полимеразой, которая не обладает 3'-экзонуклеазной активностью; (d) выделенную лигазу, которая, обязательно, является термостабильной; (е) смесь дНТФ; и (f) необязательно подходящий буфер в условиях, эффективных для вставки фрагмента в расщепленную вирусную нуклеиновую кислоту с образованием рекомбинантной нуклеиновой кислоты, содержащей сконструированный вирусный геном.

Вариант реализации изобретения 16 представляет собой способ согласно любому из вариантов реализации изобретения 10-14, в котором сборку выполняют *in vivo* и, необязательно, в которой сборку *in vivo* выполняют в клетке-хозяине.

Вариант реализации изобретения 17 представляет собой способ согласно любому из вариантов реализации изобретения 10-16, в котором выполняется по меньшей мере один из следующих элементов: 1) эндонуклеаза представляет собой РНК-направляемую нуклеазу; 2) способ дополнительно включает по меньшей мере одну направляющую РНК; 3) РНК-направляемая нуклеаза представляет собой Cas9 или фермент, полученный из Cas9, и при этом по меньшей мере одна направляющая РНК содержит: (а) химерную гРНК; или (б) crРНК и tracrРНК; 4) эндонуклеазу подвергают инактивации теплом или удалению перед сборкой; 5) система расщепления *in vitro* дополнительно содержит спермидин; 6) способ дополнительно включает трансформацию сконструированного вирусного генома в клетку-хозяина; 7) способ дополнительно включает использование упаковочного набора *in vitro* для упаковки сконструированного вирусного генома в вирусные частицы; или 8) любую их комбинацию.

Вариант реализации изобретения 18 представляет собой модифицированный вирус, полученный способом согласно любому из вариантов реализации изобретения 10-17, и необязательно, в котором сконструированный вирус представляет собой сконструированные вирусы согласно любому из вариантов реализации изобретения 1-9.

Вариант реализации изобретения 19 представляет собой набор для сконструированных молекул нуклеиновых кислот, которые являются необязательно молекулами вирусной нуклеиновой кислоты, которые включают: (а) очищенную рекомбинантную РНК-направляемую нуклеазу; (b) выделенную 5'-3'-экзонуклеазу, которая не обладает 3'-экзонуклеазной активностью, которая необязательно нетермостабильна; (с) выделенную невытесняющую ДНК-полимеразу с 3'-экзонуклеазной активностью, которая необязательно является термостабильной, или смесь указанной ДНК-полимеразы со второй ДНК-полимеразой, которая не обладает 3'-экзонуклеазной активностью; (d) выделенную лигазу, которая обязательно является термостабильной; и необязательно дополнительно включает любое из следующего: 1) краудинг-агент, 2) смесь дНТФ, 3) подходящий буфер, 4) специально разработанные направляющие РНК, 5) специально разработанные синтезированные молекулы нуклеиновой кислоты, предназначенные для использования в качестве вставленного фрагмента ДНК в реакции сборки, 6) компетентные клетки-хозяева для трансформации, 7) выделенную вирусную геномную нуклеиновую кислоту или 8) любую их комбинацию.

Вариант реализации изобретения 20 представляет собой способ модификации последовательности нуклеиновой кислоты, включающий: (а) предоставление нуклеиновой кислоты; (b) расщепление *in vitro* области нуклеиновой кислоты с использованием РНК-направляемой нуклеазы; и (с) сборку рекомбинантной нуклеиновой кислоты путем вставки фрагмента ДНК в расщепленную нуклеиновую кислоту, причем сборку проводят *in vitro* в одном сосуде со смесью компонентов, включающих: (i) выделенную 5'-3'-экзонуклеазу, которая не обладает 3'-экзонуклеазной активностью, которая необязательно не термостабильна; (ii) выделенную невытесняющую ДНК-полимеразу с 3'-экзонуклеазной активностью, которая необязательно является термостабильной, или смесь указанной ДНК-полимеразы со второй ДНК-полимеразой, которая не обладает 3'-экзонуклеазной активностью; (iii) изолированную лигазу, которая необязательно является термостабильной; (iv) смесь дНТФ в условиях, которые являются эффективными для вставки фрагмента в расщепленную нуклеиновую кислоту с образованием рекомбинантной нуклеиновой кислоты и, необязательно, в которой сборка *in vitro* дополнительно содержит (v) краудинг-агент или (vi) подходящий буфер.

Вариант реализации изобретения 21 представляет собой способ согласно варианту реализации изобретения 20, в котором, по меньшей мере, выполняется один из следующих элементов: 1) РНК-направляемая нуклеаза представляет собой Cas9 или фермент, полученный из Cas9; 2) РНК-направляемая нуклеаза подвергается инактивации теплом или удалению перед сборкой; 3) способ дополнительно включает трансформацию рекомбинантной нуклеиновой кислоты в клетку-хозяина; 4) нуклеиновая кислота представляет собой плазмиду, выделенную из клетки-хозяина, и необязательно, при этом плазида составляет по меньшей мере 6 т.п.н., по меньшей мере 10 т.п.н., по меньшей мере 15 т.п.н. или по меньшей мере 20 т.п.н. или 5) любую их комбинацию.

Раскрытие во всех его аспектах проиллюстрировано дополнительно в следующих примерах. Однако примеры не ограничивают объем настоящего раскрытия, который определен прилагаемой формулой изобретения. Обсуждение общих способов, приведенных в настоящем документе, предназначено только для иллюстративных целей. Другие альтернативные способы и варианты реализации настоящего изобретения будут очевидны специалистам в данной области техники после рассмотрения настоящего описания

и должны быть включены в сущность и сферу действия настоящей заявки.

### Примеры

#### Пример I.

Модификация вирусного генома *in vitro*.

43 т.п.н. дцДНК вирусного генома (номер доступа NC\_010326.1) было выделено из вирусных частиц, например, с использованием набора для выделения ДНК фагов Norgen Biotek или любых других способов, известных специалистам в данной области техники (фиг. 2А). Сайт-специфическое расщепление выполняли с использованием РНК-направляемой нуклеазы Cas9 и транскрибируемых *in vitro* гРНК в двух независимых местах. Нерасщепленная геномная ДНК размером 43 т.п.н. мигрирует значительно меньше, чем самая большая маркерная полоса ДНК (10 т.п.н.). Расщепление линейного генома дает фрагменты трех размеров: ~ 39 т.п.н., ~ 4,3 т.п.н. и ~ 200 п.о. Нацеливающие гРНК использовали в избытке и блокировали фрагмент в 200 п.о. (фиг. 2Б). Фрагмент gp7 из ФКФ77 был ПЦР-амплифицирован (фиг. 2В) с использованием праймеров, содержащих 5'-хвосты с гомологией 100 п.о. к областям, непосредственно расположенным в прямом и обратном направлении от сайтов расщепления LUZ19. Метод сборки Gibson Assembly был использован для бесшовной вставки ПЦР-амплифицированного фрагмента gp7 ФКФ77 (SEQ ID NO: 8) в расщепленный геном LUZ19 для замены нативной области gp7 (SEQ ID NO: 23) (фиг. 2Г). Наблюдали небольшой фон, так как Cas9-расщепление приводит к двухцепочечному разрыву ДНК с тупыми концами, который не имеет гомологии, необходимой для сборки методом Gibson Assembly *in vitro*. Отредактированные *in vitro* геномы трансформировали непосредственно в клетки-хозяева, чтобы получить функциональные вирусные частицы (фиг. 2Д). Вставку фрагмента гена ФКФ77 в восстановленные вирусы проверяли с использованием ПЦР с праймерами, внутренними и внешними по отношению к области модификации. Неотредактированную геномную ДНК LUZ19 использовали как отрицательный контроль, тогда как все экспериментальные вирусы содержали новый фрагмент гена ФКФ77 (последние 7 полос).

Эти данные представляют собой пример применения *in vitro* вирусной инженерии для редактирования генома литического фага *P.aeruginosa*. Сконструированный фаг, такой как LUZ19, не может быть получен стандартными методами из-за эффектов токсичности в гетерологичных бактериальных хозяевах, таких как *E.coli*, отсутствия селективных маркеров, подходящих для вирулентных вирусов, и отсутствия уникальных сайтов для стандартных ферментов рестрикции в геноме LUZ19. Поэтому эти данные демонстрируют, как описанный в настоящем документе, способ модификации *in vitro* позволяет осуществлять прямое и быстрое конструирование других генетически несовместимых вирусных геномов.

Для трансформаций в *P.aeruginosa* были получены химически компетентные клетки *P.aeruginosa*, как описано в Irani and Rowe (Irani V.R. & Rowe J.J. *BioTechniques* 1997, 22, 54-56). По существу, 3 мл стартовой культуры клеток *P.aeruginosa* разбавляли в 400 мл свежей LB (среда Лурия-Бертани). Культуру выращивали при 37°C при встряхивании (220 об/мин) до  $OD_{600}=0,6$ , если не указано иное. Клетки охлаждали в течение 10 мин на льду, переносили в 500 мл колбу для центрифугирования и осаждали в центрифуге с охлаждением (4°C) при 5000 g в течение 20 мин. Бактериальный осадок промывали 100 мл ледяного 150 ммоль/л  $MgCl_2$  перед разделением на две 50 мл конические пробирки и осаждали при 5000 g в центрифуге с охлаждением (4°C). Клетки промывали еще один раз с использованием 30 мл 150 ммоль/л  $MgCl_2$  перед центрифугированием и ресуспендировали в 15 мл холодного 150 ммоль/л  $MgCl_2$ . Клеточную суспензию инкубировали на льду в течение 1 ч перед центрифугированием при 4°C и ресуспендировали в 4 мл охлажденного 150 ммоль/л  $MgCl_2$ . Аликвоты в 200 мкл помещали в отдельные 1,5 мл микроцентрифужные пробирки и выдерживали на льду в течение 2 сут. Очищенную ДНК добавляли к каждой аликвоте клеток, кратковременно встряхивали и инкубировали на льду еще 1 ч. Клетки подвергали тепловому шоку при 50°C в течение 3 мин и помещали непосредственно на лед в течение 5 мин до культивирования. Каждую трансформацию добавляли к 4 мл 50°C агара для верхнего слоя LB и помещали на предварительно нагретую чашку Петри с LB. Чашки Петри переворачивали и инкубировали при 37°C, чтобы обеспечить образование бактерий.

#### Пример II.

Сконструированный вирус с расширенным кругом хозяев.

Большая клиническая библиотека (282 штамма *P.aeruginosa*) была подвергнута скринингу на предмет чувствительности к фагам LUZ19 и LKD16 с использованием анализа бляшкообразования в двойном агаровом слое. Шестьдесят шесть штаммов были инфицированы хотя бы одним из двух вирусов, причем 18 и 6 штаммов были инфицированы исключительно LUZ19 и LKD16 соответственно. Таким образом, LUZ19 был выбран в качестве шасси для тестирования генетических элементов LKD16, ответственных за расширение круга хозяев. Сравнительная геномика между двумя вирусами показала, что продукт гена 18 LKD16 (gp18) имел отличную последовательность от гомолога gp18 LUZ19, указывая на то, что он может отвечать за определение круга хозяев. Вирусный геном был выделен из вирусных частиц LUZ19, как описано выше. Сайтспецифическое расщепление проводили с использованием РНК-зависимой нуклеазы и транскрибируемых *in vitro* гРНК для экспрессии гена gp18 LUG19. Gp18 из LKD16 был амплифицирован ПЦР с гомологичными концами LUZ19 для интеграции. Метод сборки Gibson Assembly использова-

ли для бесшовной вставки ПЦР-амплифицированного gp18 LKD16 (SEQ ID NO: 7) в расщепленный геном LUZ19, чтобы заменить нативную последовательность gp18 (SEQ ID NO: 50). Сконструированные *in vitro* геномы трансформировали непосредственно в клетки-хозяева, чтобы получить функциональные вирусные частицы. Сконструированный вирус LUZ19, содержащий ген gp18 LKD16, смог заразить все штаммы, обычно инфицированные фагом LUZ19, а также 3 штамма, ранее инфицированные только LKD16, демонстрируя расширение круга хозяев (фиг. 3Б и 6Б). Это демонстрирует, что gp18 является генетическим элементом, ответственным за различающийся круг хозяев LKD16, и что сконструированный вирус LUZ19, несущий данный ген, лучше способен реплицировать в большем количестве штаммов хозяина.

Эти данные демонстрируют реализацию способа конструирования *in vitro*, описанного в настоящем документе, в отношении генетически неприемлемого вирусного генома, который приводит к улучшению вирусного свойства расширенного круга хозяев. Способность рационально модифицировать бактериофаг с расширенным кругом хозяев является большой ценностью при разработке вирусов для уничтожения бактерий.

#### Пример III.

Сконструированный вирус с кругом хозяев рода вирусов LUZ19 и/или производное LUZ19 использовали в качестве исходного материала для экспериментов по эволюции или коинфекции для определения целей для объединения круга хозяев рода вирусов ФКМВ в один репрезентативный вирус. Эксперименты по котрансформации или коинфекции выполняли либо в перmissive (РАОІК), либо в неpermissive (устойчивом) хозяине (РА7410 или РА7649) (фиг. 4А). Как коинфекцию, так и котрансформацию выполняли в присутствии LKD16 или ФКМВ соответственно. Круг хозяев тестировали с применением анализа бляшкообразования в двойном агаровом слое на указанных бактериальных штаммах. После скрининга на расширенный круг хозяев в интересующем штамме усовершенствованный фаг пасировали 3-5 раз альтернативно через permissive и селективные штаммы (штамм, который инфицирован только LUZ19-РА7632). Усовершенствованный фаг амплифицировали в РАОІК, гДНК очищали и секвенировали. Сравнительную геномику между LUZ19 и усовершенствованным LUZ19, способным заражать штаммы, ранее чувствительные только к LKD16 или ФКМВ, использовали для определения точечной мутации, ответственной за расширение круга хозяев (фиг. 4Б).

Большую клиническую библиотеку (282 штамма *P.aeruginosa*) подвергли скринингу на предмет восприимчивости к роду вирусов ФКМВ с использованием анализа бляшкообразования в двойном агаровом слое. Три фага (LUZ19, LKD16 и ФКМВ) отображали различающийся круг хозяев и были способны заразить 67 штаммов, при этом LUZ19 заразил большинство клинических штаммов (фиг. 4В). Шесть клинических штаммов (РА7245, РА7255, РА7427, РА7503, РА7686 и РА7410) восприимчивы только к LKD16, и один клинический штамм восприимчив только к ФКМВ (РА7649). Таким образом, LUZ19 был выбран в качестве шасси для экспериментов по эволюции/коинфекции/котрансформации для получения варианта, который способен инфицировать все клинические штаммы, чувствительные к роду ФКМВ. Сравнительной геномикой показано, что для того, чтобы LUZ19 заразил штаммы, восприимчивые только к LKD16 или ФКМВ, необходимо было указать несколько точечных мутаций: (i) gp13 C17Y (положение 50 SEQ ID NO: 1) необходимо для инфицирования РА7427; (ii) gp18 D36Y (положение 106 SEQ ID NO: 50) необходимо для инфицирования РА7245, РА7503 и РА7686; gp38 D82G и 183S (положение 245 и 247-248 SEQ ID NO: 2 соответственно) позволяет инфицировать РА7410 и РА7649; (iv) gp40 N253D (положение 757 SEQ ID NO: 3) позволяет инфицировать РА7255 (фиг. 4Б). Итеративное конструирование вышеупомянутых мутаций в шасси LUZ19 с использованием способа конструирования *in vitro*, описанного в настоящем документе, дало LUZ19 с широким кругом хозяев (ШКХ-LUZ19), способному заражать все клинические штаммы, восприимчивые к фактору рода ФКМВ (фиг. 4В).

Эти данные представляют собой пример использования способа конструирования *in vitro*, описанного в настоящем документе, для объединения круга хозяев рода вируса в один вирусный геном путем первой идентификации генетических мутаций, ответственных за различия в круге хозяев после эволюционных экспериментов, скрининга, секвенирования, сравнительной геномики, и любой их комбинации.

#### Пример IV.

Улучшенная вирусная репликация улучшает раннее разрушение биопленки.

В другом примере вирусная эволюция и сравнительная геномика показали, что усовершенствованный фаг LUZ19 с мутацией L55 в пределах хвостового трубчатого белка В (Gp34) реплицируется с большей скоростью из-за увеличенного выхода фага (фиг. 5Б). Чтобы подтвердить, что мутация Gp34 L55Δ улучшила вирусные свойства, вирусный геном LUZ19 выделяли из вирусных частиц. Сайт-специфическое расщепление выполняли с использованием РНК-зависимой нуклеазы и транскрибируемых *in vitro* гРНК для удаления гена gp34 (SEQ ID NO: 4). Ген gp34 L55Δ, который несет делецию кодона лейцина в аминокислотном положении 55 (Gp34 L55Δ, положение 163-165 SEQ ID NO:4), ПЦР-амплифицировали из усовершенствованного фага LUZ19. Метод сборки Gibson Assembly использовали для вставки ПЦР-амплифицированного гена gp34 L55Δ в расщепленный геном LUZ19. Трансформированные *in vitro* геномы трансформировали непосредственно в клетки-хозяева, чтобы получить функцио-

нальные вирусные частицы.

Сконструированный вирус LUZ19, содержащий Gr34 L55Δ, способен рассеять и лизировать бактерии. Анализ бляшкообразования в двойном агаровом слое использовали для того, чтобы показать, что у фага LUZ19, несущего мутацию gr34 L55Δ5 (фаг\*), была большая область очистки (прозрачное пятно), чем LUZ19 дикого типа. Были сделаны снимки и области очистки измеряли в течение двух суток (фиг. 5Б). Расширяющийся диаметр лизиса указывает на то, что вирусы, несущие мутации Gr34 L55Δ, лучше способны рассеивать и лизировать бактерии. Анализом биопленки с окрашиванием кристаллическим фиолетовым измеряют накопление биопленки в виде измерения включения кристаллического фиолетового (фиг. 5В). Образцы, обработанные вирусами, содержащими мутации gr34 L55Δ, имели значительное уменьшение биопленки по сравнению с вирусами с геном gr34 дикого типа. На иллюстрации показано расположение мутации gr34 (звездочка) по сравнению с геномом LUZ19 дикого типа (фиг. 5Г). Стандартные анализы, известные в данной области техники, должны были измерять адсорбцию вируса, латентный период и величину выхода фага как у дикого типа, так и мутантов gr34 L55Δ. Эти данные указывают на то, что вирусы, содержащие мутацию gr34 L55Δ, имели значительно увеличенный выход фага (фиг. 5Д).

Эти данные обеспечивают пример использования способа конструирования *in vitro*, описанного в настоящем документе, для создания вируса с улучшенными вирусными свойствами повышенного бактериального лизиса, выхода фага, репликации и раннего разрушения биопленки.

Пример V.

Вирус, полученный путем итеративного конструирования, с ранним разрушением биопленки и расширенным кругом хозяев.

Расширенный круг хозяев LUZ19<sub>LKD16gp18</sub> рекомбинантного вирусного генома, созданный в примере II, был выделен из вирусных частиц. Для удаления gr34 (SEQ ID NO: 4) с использованием РНК-зависимой нуклеазы и транскрибируемых *in vitro* гРНК выполняли сайтспецифическое расщепление. Мутацию gr34 ΔLeu55, повышающую литическую активность, (положение 163-165 SEQ ID NO: 4), описанную в примере IV, затем ПЦР-амплифицировали и собирали в расщепленный вирусный геном LUZ19<sub>LKD16gp18</sub> с применением метода сборки Gibson Assembly. Геномы, полученные путем итеративного конструирования *in vitro*, трансформировали непосредственно в клетки-хозяева для получения функциональных вирусных частиц, т.е. сконструированный вирус LUZ19, содержащий как ген gp18 LKD16, так и мутацию gr34 ΔLeu55 (LUZ19\*<sub>LKD16gp18</sub>).

Вирус LUZ19\*<sub>LKD16gp18</sub> анализировали на улучшенные вирусные свойства, используя анализы бляшкообразования в двойном агаровом слое, биопленки и анализы на прикрепление кератиноцитов человека *in vitro*. На фиг. 6Г показано, что у LUZ19\*<sub>LKD16gp18</sub> увеличен круг хозяев. LUZ19\*<sub>LKD16gp18</sub> по сравнению с нативным LUZ19 для способности разрушать предварительно сформированные МЛР биопленки *P.aeruginosa*. В частности, LUZ19\*<sub>LKD16gp18</sub> и LUZ19 дикого типа инкубировали с биопленкой *P.aeruginosa*, и разрушение измеряли с использованием кристаллического фиолетового. На фиг. 6Д показано, что LUZ19\*<sub>LKD16gp18</sub> обладает повышенной способностью разрушать предварительно сформированные МЛР биопленки *P.aeruginosa* по сравнению с LUZ19 дикого типа. Вирус LUZ19\*<sub>LKD16gp18</sub> анализировали на эффективность фаговой терапии против бактерий, прикрепленных к кератиноцитам человека. В частности, *P.aeruginosa* прикрепляли к монослою клеток HaCaT. Затем клетки инкубировали с LUZ19\*<sub>LKD16gp18</sub> или LUZ19 дикого типа. Результаты показали, что фаг LUZ19\*<sub>LKD16gp18</sub> лучше способен устранять множественную лекарственную устойчивость (IVIDR) Клетки *P.aeruginosa* прикрепляли к кератиноцитам человека (см. фиг. 6Е).

Эти данные приводят пример того, как способ конструирования *in vitro*, описанный в настоящем документе, использовали в системе для итеративного конструирования бактериофага с несколькими независимыми улучшенными вирусными свойствами, такими как расширенный круг хозяев и увеличенный выход фага. Важно отметить, что данные этапы конструирования не могут быть выполнены как напрямую, так и с использованием стандартных методов. Кроме того, эти данные демонстрируют, что способ конструирования *in vitro*, описанный в настоящем документе, использовали последовательно в итерационных циклах конструирования, что является важным свойством для применений в синтетической биологии.

Пример VI.

Вирусы, полученные путем итеративного конструирования, с биофильм-диспергирующей полезной нагрузкой, и расширенный круг хозяев, охватывающий полный вирусный род.

Любые экзополисахариды (ЭПС) деполимеразы или фенол-растворимые модулины (PSM) клонировали в LUZ19 путем замены gr49 (SEQ ID NO: 25) с использованием способа конструирования *in vitro*, описанного в настоящем документе, для определения их способности диспергировать зрелую биопленку (фиг. 7). Чтобы сконструировать LUZ19 и WHR LUZ19 для экспрессии внеклеточной деполимеразы или поверхностно-активных полипептидов, gr49 (SEQ ID NO: 25) LUZ19 или WHR LUZ19 удаляли путем расщепления с использованием РНК-зависимой нуклеазы, в этом случае Cas9 и *in vitro* транскрибируемые гРНК и затем заменяли на ген интереса (ГИ), фланкированный основным промотором капсида P<sub>gp32</sub>

(SEQ ID NO: 21) и терминатором T<sub>gp32</sub> (SEQ ID NO: 22) с использованием метода сборки Gibson Assembly (фиг. 7А и 7В). В случае LUZ19 дикого типа, ГИ представляют собой ЭПС-деполимеразы (Pp15gp44 - хвостовой шип gp44 из Ф15 *Pseudomonas pudita* (SEQ ID NO: 14); NTUgp34 - хвостовой шип gp34 из фага NTU *Klebsiella pneumoniae* (SEQ ID NO: 13); LKA1gp49 - хвостовой шип gp49 фага из LKA1 *P.aeruginosa* (SEQ ID NO: 12)), поверхностно-активные фенол-растворимые модулины из *Staphylococcus epidermidis* (PSMa, SEQ ID NO: 18) и *Staphylococcus aureus* (PSMa3 (SEQ ID NO: 16) и PSMb2 (SEQ ID NO: 17)) и DspB-сурфактин из *Aggregatibacter actinomycetemcomitans* (SEQ ID NO: 15) (фиг. 7Б). В случае ШКХ-LUZ19, ГИ представлял собой ЭПС-деполимеразу Pp15gp44 (SEQ ID NO: 14) и сурфактин SePSMa (SEQ ID NO: 18) (фиг. 7Г). Сконструированный фаг амплифицировали в соответствующей клетке-хозяина, выделяли и проверяли путем секвенирования.

Способность сконструированного фага к диспергированию зрелой биопленки была протестирована на 24-часовой биопленке, выращенной в устройстве MBEC, с использованием 100 фагов на лунку в течение 3 ч. Вкратце, ночные культуры *P.aeruginosa* разбавляли (1:100) в минимальной среде M63, дополненной сульфатом магния (1 мМ), глюкозой (0,2%) и казиминокислотами (0,5%), а затем добавляли в стерильные микротитрационные планшеты (150 мкл на лунку). Крышку с заглушками вставляли в планшет для микротитрования. После 24 ч инкубации при 37°C крышку с заглушками перемещали на микротитрационные планшеты (150 мкл на лунку). Крышку с заглушками вставляли в планшет для микротитрования. После 24 ч инкубации при 37°C крышку с заглушками перемещали на микротитрационный планшет, содержащий 160 мкл полной MG63, содержащей 100 фагов на лунку. После 3 ч инкубации при 37°C крышку с заглушками промывали 3 раза в воде, сушили и окрашивали 200 мкл 0,5% кристаллического фиолетового. Затем планшеты промывали водой для удаления несвязанного кристаллического фиолетового и сушили. Краситель растворяли в 200 мкл 30% уксусной кислоты и измеряли оптическую плотность при ОП=550 нм.

DspB, который является поверхностно активным против биопленок *E.coli*, служил в качестве отрицательного контроля, поскольку он не имеет активности против *P.aeruginosa*. Две полезные нагрузки (Pp15gp44 и SePSMa) показали заметную активность против биопленки (фиг. 7Б). Примечательно, что PSM, которые являются сурфактинами с известной антибиопленочной активностью у грамположительных бактерий, никогда ранее не демонстрировали, что они диспергируют биопленку *P.aeruginosa*. Данные полезные нагрузки были сконструированы в ШКХ-LUZ19, чтобы определить, может ли фаг с широким кругом хозяев быть дополнительно модифицирован с целью придания ему способности диспергировать биопленку. Результаты показывают, что ШКХ-LUZ19, кодирующие Pp15gp44 или SePSMa, поддерживают их биофильм-диспергирующую активность (фиг. 7Г) и способность инфицировать все клинические штаммы, восприимчивые к роду вирусов O-KMV (фиг. 7Д, 7Е).

Эти данные представляют собой пример того, как способ конструирования *in vitro*, описанный в настоящем документе, может использоваться в системе для итеративного конструирования бактериофага с несколькими независимыми улучшенными вирусными свойствами, такими как неограничивающие свойства дисперсии биопленки и круг хозяев.

#### Пример VII.

Сконструированные вирусы, экспрессирующие антибиотическую сенсibiliзирующую полезную нагрузку.

С использованием способа конструирования *in vitro*, описанного в настоящем документе, LUZ19 был сконструирован для экспрессии лизинов из оцПНК-вирусов PRR1 и MS2. Лизины как из PRR1 (SEQ ID NO: 20), так и из MS2 (SEQ ID NO: 19) оцПНК-фага были встроены в локус gp49 LUZ19 (SEQ ID NO: 25), фланкированный основным капсидным промотором Pgp32 (SEQ ID NO: 21) и терминатором Tgp32 (SEQ ID NO: 22), чтобы определить их способность ингибировать появление бактерий, устойчивых к фагу (фиг. 8А). Эти лизины ингибируют образование новых клеточных стенок путем связывания и инактивации ферментов, важных для синтеза клеточной стенки, и предположительно сенсibiliзируют бактерии к другим антимикробным препаратам, особенно к нацеленным на клеточные стенки антибиотикам, таким как карбенициллин.

Конструкт получали, как описано выше, с использованием способа конструирования *in vitro*, описанного в настоящем документе. Сконструированный фаг амплифицировали в соответствующей клетке-хозяина, выделяли и проверяли путем секвенирования. Способность сконструированного фага ингибировать появление бактерий, устойчивых к фаговой терапии, в присутствии карбенициллина при 1/5 × МИК тестировали в стандартном анализе зависимости "время - летальное действие" (фиг. 8Б, 8В). Результаты показывают, что сконструированные LUZ19, экспрессирующие лизины из оцПНК фага в сочетании с карбенициллином в субингибирующих концентрациях (1/5 × МИК) предотвращают повторный рост бактерий после фаговой терапии.

Эти данные представляют собой пример использования способа конструирования *in vitro*, описанного в настоящем документе, для создания вируса с улучшенными вирусными свойствами, в частности в данном случае, предотвращения развития устойчивости к фагу у бактерий.



## Пример VIII.

Сконструированный вирус, экспрессирующий видоспецифическую противомикробную белковую нагрузку.

Используя способ конструирования *in vitro*, описанный в настоящем документе, LUZ19 был сконструирован для экспрессии антимикробного белка P<sub>yoS5</sub>, продуцируемого *P.aeruginosa*. Бактериоцин P<sub>yoS5</sub> является видоспецифическим антимикробным белком, продуцируемым одним штаммом *P.aeruginosa*, чтобы препятствовать росту конкурирующих штаммов *P.aeruginosa*. ГДНК штамма PA01 *P.aeruginosa* использовали в качестве матрицы для ПЦР-амплификации p<sub>yoS5</sub> (SEQ ID NO: 6) до клонирования в локус gp49 LUZ19 (SEQ ID NO: 25), фланкированный основным промотором капсида P<sub>gp32</sub> (SEQ ID NO: 21) и терминатором T<sub>gp32</sub> (SEQ ID NO: 22) (фиг. 9А). P<sub>yoS5</sub> связывается с широко диспергированным пиохелиновым рецептором FptA перед конформационными изменениями для создания пор в мембране *P.aeruginosa*.

LUZ19+p<sub>yoS5</sub> был создан, как описано выше, с использованием способа конструирования *in vitro*, описанный в настоящем документе. Сконструированный фаг амплифицировали внутри восприимчивого хозяина PA01, выделяли и проверяли путем секвенирования. Бактериальный штамм PA7416 выбирали для анализа, потому что лабораторный штамм PA01, как известно, устойчив к P<sub>yoS5</sub>, однако в анализе *in silico* показано, что штамм PA7416 МЛР *P.aeruginosa* был восприимчивым к фагу LUZ19 и кодировал P<sub>yoS5</sub>-рецептор FptA.

Способность фага ингибировать появление бактерий PA7416, устойчивых к фаговой терапии, тестировали в стандартном анализе зависимости "время - летальное действие". Результаты показывают, что в то время как LUZ19 дикого типа первоначально ингибирует рост PA7416, бактерии быстро становятся устойчивыми, а повторный рост происходит через 8-12 часов (фиг. 9Б). Однако сконструированный UJZ/9+p<sub>yoS5</sub> ингибирует повторный рост бактерий PA7416 на более чем 24 ч после обработки фагом (фиг. 9В, 9Г).

Эти данные представляют собой пример использования способа конструирования *in vitro*, описанного в настоящем документе, для получения вируса с улучшенными вирусными свойствами, в частности в данном случае, предотвращения развития устойчивости к фагу у бактерий.

## Пример IX.

Система итеративного конструирования бактериофага для создания антимикробного продукта.

Используя описанный в настоящем документе способ конструирования *in vitro*, геномы бактериофагов могут быть быстро сконструированы без обширной генетической манипуляции штаммом хозяина. Объединение вирусных мутационных исследований и методов селекции, хорошо известных специалистам в данной области техники, с полным секвенированием генома, сравнительной геномикой и раскрытием в настоящем изобретении способом конструирования *in vitro* создает новую и улучшенную систему для разработки новых и улучшенных антимикробных препаратов. Система основана на итеративном улучшении 1, 2 или более 2 различных свойств в одном вирусном шасси для создания антимикробного препарата на основе вирусов. Последовательная очистка и редактирование генома LUZ19 для улучшения отдельных вирусных свойств были раскрыты (фиг. 6, 7 и 10), однако данный способ можно распространить на множество других бактериофагов *P.aeruginosa* или другого бактериофага, заражающего любые другие штаммы или виды бактерии. Кроме того, данный способ можно использовать для улучшения свойств нескольких отдельных бактериофагов, заражающих одни и те же виды бактерий, для создания превосходного коктейля бактериофагов, предотвращающего или лечащего бактериальные инфекции, заражения или изменения микробиомы.

Эти данные демонстрируют, как конструирование *in vitro* в сочетании с секвенированием генома, сравнительной геномикой и вирусными мутационными/селекционными исследованиями могут выполняться последовательно для достижения поэтапных улучшений или инженерных изменений для включения улучшенных вирусных свойств, представляющих интерес (фиг. 10).

## Пример X.

## Способы.

Гидовые РНК (гРНК) синтезировали и очищали с использованием коммерчески доступного набора для транскрипции *in vitro*, такого как набор MEGAscript T7 kit (Thermo Fisher). Гидовые РНК были разработаны с использованием методов, хорошо известных в данной области техники (фиг. 15).

Разбавьте *in vitro* транскрибируемые гРНК на рабочий материал 500 нг/мкл.

Реакции сборки без очищенной РНК-направляемой нуклеазы, такой как Cas9. Очищенный Cas9 (SEQ ID NO: 31) был получен из экспрессирующей плазмиды, содержащей последовательность генов, кодирующую His-меченый Cas9 (SEQ ID NO: 27) и очищающей ее с помощью известных способов никель-аффинной очистки. Необязательно используйте гРНК, которая сначала разрезает самую внутреннюю часть генома для итеративных расщеплений.

Полная реакционная смесь:

мкл:

10X Cas9 буфер\* \* 4;

50 ммоль/л MgCl<sub>2</sub> 8;

100 ммоль/л спермидин 4;  
 гРНК1 2;  
 гРНК2 2;  
 Фермент Cas9 (0,45 мг/мл) 8 (общее количество);  
 гДНК X (общая масса 2 мкг);  
 дистиллированная H<sub>2</sub>O X (до общего объема 40 мкл).

\* Полная реакционная смесь может быть использована на одном этапе для одновременного вырезания нескольких сайтов (совместное расщепление), однако это может привести к низкой эффективности вырезания вирусной гДНК. Реакции совместного расщепления выполняют на льду перед добавлением Cas9 и инкубацией при 37°C в течение 30 мин. Сконструированная 2-стадийная (или более) реакция также может быть выполнена, что обеспечивает более полное расщепление (описано ниже).

\*\*10x Cas9 буфер содержит 200 ммоль/л HEPES pH 7.4, 1.5M KCl, 5 ммоль/л DTT и 1 ммоль/л EDTA pH8.

1 стадия реакции сборки и инкубация при комнатной температуре в течение 5 мин.

Реакционная смесь 1 стадии:

мкл:  
 10X Cas9 буфер 4;  
 50 ммоль/л MgCl<sub>2</sub> 8;  
 100 ммоль/л спермидина 4;  
 гРНК1 2;  
 гДНК (общее количество 2 мкг);  
 дистиллированная H<sub>2</sub>O X (до общего объема 36 мкл).

Инкубируйте на льду в течение 10 мин.

Инкубируйте при 37°C в течение 2 мин.

Добавьте 4 мкл фермента Cas9 (0,45 мг/мл). Инкубируйте при 37°C в течение 30 мин.

2 стадия реакции, добавление второй гРНК и дополнительного фермента Cas9.

Реакционная смесь 2 стадии:

мкл:  
 Реакционная стадия 1 стадии 40;  
 гРНК2 2;  
 Фермент Cas9 (0,45 мг/мл) 4;  
 10X Cas9 буфер 1;  
 дистиллированная H<sub>2</sub>O 3 (общий объем 50 мкл).

Инкубируйте реакцию 2 стадии при 37°C в течение 30 мин. Дополнительные шаги могут быть добавлены для расщепления генома в более чем 2 местах.

Инактивируйте фермент Cas9 путем инкубации при 80°C в течение 10 мин. Необязательно проведите очистку с использованием фенольно-хлороформной экстракции (повышает эффективность сборки фрагментов в сборке методом Gibson Assembly) или других способов инактивации, дезактивации или очистки, хорошо известных в данной области техники.

Исследуйте 5 мкл образца на агарозном геле для проверки правильности разрезания.

Для сборки *in vitro* с использованием метода сборки Gibson Assembly соответствующую концентрацию продуктов расщепления и *in vitro* сконструированной вставки использовали в соответствии с протоколом метода сборки Gibson Assembly NEB.

После сборки *in vitro*, возможно, трансформируют в клетки-хозяева для амплификации сконструированного генома, части генома или восстановления сконструированного вируса.

Пример XI.

Конструирование фага M13 E.coli.

Используя описанный в настоящем документе способ конструирования *in vitro*, был сконструирован вирус, заражающий *Escherichia coli*, для экспрессии флуоресцентного репортера паприка (SEQ ID NO: 5). На фиг. 11А показана схема подхода конструирования *in vitro* для включения гена флуоресцентного белка паприка в геном фага M13 E.coli. Этот процесс конструирования был разработан для создания лизогенного фага, экспрессирующего флуоресцентный репортер, который будет представлять собой улучшенное вирусное свойство, поскольку аналогичные вирусы использовали в качестве диагностики. Вирусный геном M13 (номер доступа X02513) был выделен из вирусных частиц. Поскольку экспериментальная конструкция включает использование двух гРНК, функциональность каждой отдельной гРНК была впервые подтверждена в отдельных реакциях Cas9-расщепления *in vitro* (фиг. 11Б). Зная, что каждая гРНК была функциональной, сайт-специфическое расщепление выполняли с использованием РНК-зависимой нуклеазы и обеих *in vitro* транскрибируемых гРНК (фиг. 11В). Флуоресцентный репортерный ген паприка (SEQ ID NO: 29) ПЦР-амплифицировали (фиг. 11Г) с использованием праймеров, которые добавляли 5'- и 3'-последовательностям, гомологичным последовательностям, фланкирующим ген *LacZa*, который был выделен из генома M13 с использованием расщепления РНК-зависимой нуклеазой, например Cas9. Метод сборки Gibson Assembly использовали для бесшовной вставки ПЦР-амплифицирован-

ного гена паприка в расщепленный геном M13, заменяя ген LacZa (SEQ ID NO: 28). Сконструированные геномы трансформировали непосредственно в хозяина *E.coli* для получения функциональных вирусных частиц, кодирующих ген паприка. Сконструированный фаг оценивали по его способности образовывать бляшки у *E.coli* (фиг. 11Д). Вирусную ДНК выделяли из бляшек и ПЦР-амплифицировали для подтверждения присутствия вставленного гена паприка (фиг. 11Е). Присутствие и функции рекомбинантного белка паприка подтвердили при помощи флуоресцентной визуализации (фиг. 11Ж).

Эти данные демонстрируют успешное использование описанного в настоящем документе способа конструирования *in vitro* для создания репортерного гена в геноме фага *E.coli*. Демонстрируя, что раскрываемый способ можно расширить на другой род вирусов, включая тот, который заражает другой род бактерий.

#### Пример XII.

##### Конструирование фага X *E.coli*.

Используя описанный в настоящем документе способ конструирования *in vitro*, был отредактирован второй вирус, заражающий *Escherichia coli*. На фиг. 12А показана схема подхода конструирования *in vitro* для удаления гена *ell* (SEQ ID NO: 30) из выделенного генома фага X (доступ NC\_001416.1). Данный технический процесс был разработан для создания конститутивно литического вируса, который будет представлять собой улучшенное вирусное свойство. Вирусный геном  $\lambda$  был выделен из вирусных частиц. Поскольку экспериментальная конструкция предполагает использование двух гРНК, функциональность каждой отдельной гРНК была впервые подтверждена в отдельных реакциях Cas9-расщепления *in vitro* (фиг. 12Б). Зная, что каждая гРНК была функциональной, сайт-специфическое расщепление проводили с использованием РНК-зависимой нуклеазы и обеих *in vitro* транскрибируемых гРНК (фиг. 12В). Две синтезированные одноцепочечные молекулы ДНК были отождествлены *in vitro* для получения шаблона для восстановления двухцепочечной ДНК (SEQ ID NO: 9), содержащего 5'- и 3'-последовательности, гомологичные последовательностям, фланкирующим Cas9-направленные вырезанные сайты в выделенный вирусный геном  $\lambda$ . Метод сборки Gibson Assembly использовали для бесшовной вставки ПЦР-амплифицированной восстанавливающей матрицы в расщепленный геном  $\lambda$ . Затем сконструированные геномы были упакованы *in vitro* с использованием набора Maxplax lambda packaging extraction kit из Epicentre в соответствии со способом изготовителя (фиг. 12Г). После *in vitro* упаковки сконструированные геномы  $\lambda$  были извлечены из анализа бляшкообразования с двойным агаровым слоем с использованием рекомендованных изготовителем клеток-хозяев *E.coli*. Было определено, что сконструированный фаг функционирует на основе его способности образовывать бляшки у *E.coli*. Вирусную ДНК выделяли из образовавшихся бляшек и ПЦР-амплифицировали, чтобы подтвердить отсутствие гена *ell* (фиг. 12Д).

Эти данные демонстрируют успешное использование описанного в настоящем документе способа конструирования *in vitro* для удаления нежелательного гена из генома фага *E.coli*. Эти данные также обеспечивают пример упаковки сконструированных вирусных геномов *in vitro*, что повышает эффективность восстановления вируса и обеспечивает альтернативу прямой трансформации в клетку-хозяина. Кроме того, эти данные обеспечивают пример использования отождествленных *in vitro* синтезированных олигонуклеотидов в качестве вставки для конструирования. Кроме того, эти данные являются еще одним примером использования этого подхода для конструирования генома фага, чтобы получить улучшенное вирусное свойство, а именно конститутивно литический фенотип. Наконец, эти данные показывают, что второй род вируса, заражающий *E.coli*, может быть сконструирован с использованием описанного в настоящем документе способа конструирования *in vitro*.

#### Пример XIII.

##### Коррекция ошибок ЦМВ человека.

Используя раскрытый в настоящем документе способ конструирования *in vitro*, была отредактирована часть человеческого вируса. На фиг. 13А показана схема подхода конструирования *in vitro*, используемого для исправления ошибок. 18 т.п.н. часть ~230 т.п.н. вирусного генома ЦМВ человека содержалась в пределах плазмиды, реплицирующейся в *E.coli*. Настоящая часть генома ЦМВ человека (SEQ ID NO: 10) содержала начало вирусного генома и несла мутантный аллель RL13 (SEQ ID NO: 33). Вместе фрагмент ЦМВ человека и плаزمиды *E.coli* имели размер примерно 28 т.п.н., что превышает спецификации большинства современных методов коррекции ошибок. Для исправления ошибок плазмиду 28 т.п.н. выделяли из *E.coli* и проводили сайт-специфическое расщепление с использованием РНК-зависимой нуклеазы и двух *in vitro* транскрибируемых гРНК (фиг. 13Б). Cas9-опосредованное расщеплением вырезали область гена RL13 непосредственно в прямом и обратном направлении от сайта мутации. Исправленная область гена RL13 (SEQ ID NO: 32) была синтезирована и ПЦР-амплифицирована с дополнительными 5'- и 3'-фланкирующими последовательностями, гомологичными областям, граничащим с каждым сайтом расщепления, специфичным для РНК (фиг. 13В). Метод сборки Gibson Assembly использовали для бесшовной вставки синтезированной восстанавливающей матрицы в расщепленную плазмиду. Исправленный RL13-содержащий фрагмент ЦМВ человека (SEQ ID NO: 11), содержащийся внутри плазмиды, затем трансформировали в *E.coli* и восстанавливали на антибиотикосодержащих питательных средах. Колонии *E.coli* подвергали скринингу с помощью ПЦР для подтверждения наличия исправленного

гена RL13, который содержал дополнительную последовательность по сравнению с геном RL13, содержащим ошибки, что позволяет отличать его от гена RL13, содержащего ошибки (фиг. 13Г). Геномный фрагмент с исправленной ошибкой затем амплифицировали в *E.coli* с использованием стандартных методов для последующего использования в обратных приложениях.

Эти данные демонстрируют успешное использование описанного в настоящем документе способа конструирования *in vitro* для конструирования генов из генома вируса человека и дополнительно обеспечивают способ использования синтезированной ДНК в качестве восстанавливающей матрицы в реакции сборки *in vitro*. Эти данные также демонстрируют использование этого способа конструирования *in vitro* для коррекции ошибок ДНК или плазмид, которые слишком велики для стандартных методов коррекции ошибок. Стандартная методика коррекции ошибок имеет ограничение по размеру около 5 т.п.н. и основана на ПЦР, которое по своей природе может создавать более нежелательные ошибки. Представленный в настоящем документе способ конструирования *in vitro* не зависит от ПЦР-амплификации всей или даже большей части плазмиды или вирусного генома и поэтому поддается применению при исправлении ошибок последовательностей, превышающих 5 т.п.н.

Пример XIV.

Быстрая идентификация терминально избыточных вирусных концов.

Описанный в настоящем документе способ расщепления *in vitro* также может быть адаптирован для идентификации точных концов терминально избыточных вирусных геномов. На фиг. 14 показана схема подхода расщепления *in vitro*, который использовался для определения концов фаговых геномов LBL3 и 14-1. LBL3 и 14-1 (номер доступа NC\_011703.1) фаговую геномную ДНК очищали от вирусных частиц (фиг. 14А). Секвенирование нового поколения выполняли с использованием платформы MiSeq или PacBio, а затем при помощи автоматического слияния высококачественных считываний ДНК в более длинные сборки для восстановления исходной последовательности (фиг. 14Б). Как правило, программное обеспечение автоматической сборки неправильно собирает вирусные или бактериофаговые геномы в круговые контиги и размещает ПТП терминально повторяющихся геномов во внутреннюю область вирусной последовательности. Прогнозирование *in silico* физических концов генома выполняют на основе идентификации областей секвенирования двойного охвата и результатов поиска BLAST, которые соответствуют близкородственному терминально повторяющемуся геному (фиг. 14В). Данные прогнозируемые концы были подтверждены расщеплением эндонуклеазой Cas9. После инактивации Cas9 фрагменты ДНК, соответствующие геномным физическим концам, очищали и секвенировали (фиг. 14Г). Эти результаты секвенирования привели к точной сборке генома на основе истинных физических концевых последовательностей (фиг. 14Д).

Одной из самых больших технических проблем, связанных с секвенированием генома фага, является точное отображение геномных физических концов из-за их повторяющегося характера. Эти сегменты могут простираются от 4-14 п.о. в циклически перегруппированных геномах (например, большинство фагов *Mycobacterium* и *Propionibacterium acnes*) до нескольких сотен пар оснований в терминально повторяющихся геномах (например, ФКМV-подобные, РВ1-подобные и N4-подобные фаговые роды *P.aeruginosa*) и даже до нескольких тысяч пар оснований (например, T5 и DTR *E.coli*). Сопоставление повторяющихся концов (или ПТП - прямых терминальных повторов) в настоящее время выполняется комбинацией глубокого анализа последовательностей (для идентификации фрагментов ДНК с двойным покрытием), праймер-опосредованной прогулкой (секвенирование по Сэнгеру), идентификацией основных разрывов ДНК и анализом рестрикционных эндонуклеаз. Однако каждый из этих подходов часто ограничен в использовании или безрезультативный: I) плохо определенные двойные секвенсы границы охвата двойного упорядочения в данных NGS (секвенирования следующего поколения); II) считывание праймер-опосредованной прогулки с прохождением через конкатэмеры ПТП, дающая неубедительные результаты; III) низкая частота мест рестрикции вблизи фаговых концов или обструкция рестрикционных участков вследствие модификаций ДНК, таких как метилирование. Использование направленного Cas9-расщепления ДНК фага в определенных положениях устраняет необходимость в ненадежных или громоздких анализах или процедурах и значительно упрощает идентификацию фаговых геномных физических концов. Такой подход имеет потенциал для точного отображения концов уже секвенированных фаговых геномов (как показано на примере картирования ПТП LBL3 и 14-1), а также для быстрой идентификации ПТП новых идентифицированных вирусов.

Использование направленного Cas9-расщепления в пределах описанного в настоящем документе способа конструирования *in vitro* для сопоставления физических концов терминально повторяющихся фаговых геномов представляет собой явное преимущество по сравнению с существующими подходами, поскольку оно не зависит от тонких изменений в охвате секвенирования и может выполняться независимо от образования конкатэмера. Кроме того, активность Cas9 менее чувствительна к модификациям ДНК, чем многие рестрикционные ферменты.

Эти данные показывают успешное применение РНК-направляемого Cas9-расщепления *in vitro*, что дает возможность идентифицировать правильное положение последовательности фагового генома. Эта информация затем может быть использована для разработки нижерасположенных подходов модификации *in vitro* для конструирования этих фагов, что было ранее невозможным из-за отсутствия истинных

границ генома.

Пример XV.

Способ конструирования со сборкой *in vivo*.

Настоящее раскрытие обеспечивает способ *in vitro* сайт-специфического расщепления очищенной вирусной нуклеиновой кислоты с использованием РНК-направляемой нуклеазы; и сборку сконструированной нуклеиновой кислоты путем вставки фрагмента ДНК или РНК в расщепленную вирусную нуклеиновую кислоту. Хотя рекомбинантная нуклеиновая кислота может быть собрана целиком в условиях *in vitro* с использованием очищенных ферментов, как описано в настоящем документе, этот процесс также может быть осуществлен с использованием естественных или сконструированных рекомбинантных сигнальных путей в чувствительном штамме хозяина. Трансформация очищенных и обработанных *in vitro* вирусных геномов вместе с фрагментом вставки восстановления, содержащим концевые области гомологии, достаточна для того, чтобы некоторые клетки-хозяева собирали рекомбинантный вирусный геном *in vivo*. Фрагменты вставки восстановления могут быть синтезированы или амплифицированы стандартными методами, известными в данной области техники, или могут находиться внутри плазмид, стабильно реплицирующихся в выбранной клетке-хозяине. Данный способ, вероятно, будет иметь более низкую эффективность по сравнению со сборкой *in vitro* из-за клеток-хозяев, имеющих как гомологичные, так и не гомологичные пути восстановления ДНК, задача совместной доставки достаточного количества вставки и расщепленного генома в клетку-хозяина и более низкой эффективности большинства гомологичных сигнальных путей рекомбинации хозяина. Поскольку только расщепленные геномы не образуют функциональных вирусных частиц и в последующем бляшек без опосредованной хозяином рекомбинации, бляшки, полученные после трансформации и выращивания, могут быть скринированы с помощью ПЦР для данной вставки, чтобы подтвердить корректную сборку желаемой сконструированной вирусной нуклеиновой кислоты.

Пример XVI.

Сконструированные вирусы, описанные в настоящем документе.

В табл. 1 представлены сконструированные вирусы, полученные в соответствии с настоящим изобретением в способе модификации *in vitro*. В табл. 2 представлены сконструированные вирусы, описанные в настоящем документе, вместе с соответствующим примером и фигурой. В табл. 3 перечислены вирусы дикого типа, описанные в настоящем документе, и номера доступа для их полной геномной последовательности. В табл. 4 перечислены некоторые из последовательностей нуклеиновых кислот дикого типа, описанных в настоящем документе, и соответствующие аминокислотные последовательности.

Таблица 1. Сконструированные вирусы, описанные в изобретении

Сконструированный вирус	Исходный Вирус	Тип мутации	Целевая область	Номер доступа, положение нуклеотида (SEQ ID NO.)	Изменение нуклеотида и аминокислоты или добавление последовательности
LUZ19+ фрагмент <i>gp7</i> ФКФ77	LUZ19	Замена	<i>gp7</i> (SEQ ID NO:23)	Доступ NC_010326.1, 4288- 4491 (SEQ ID NO:23)	Фрагмент <i>gp7</i> ФКФ77 (SEQ ID NO:8)
LUZ19+LKD16 <i>gp18</i>	LUZ19	Замена	<i>gp18</i> (SEQ ID NO:50)	Доступ NC_010326.1, 11368-11688 (SEQ ID NO:24)	<i>gp18</i> LKD16 (SEQ ID NO:7)
ШКХ-LUZ19	LUZ19	TM	<i>gp13</i> (SEQ ID NO:1)	Доступ NC_010326.1, 7325 (положение 50 SEQ ID NO:1)	G на A, C17Y
		TM	<i>gp18</i> (SEQ ID NO:50)	Доступ NC_010326.1, 11469 (положение 106 SEQ ID NO:50)	G на T, D36Y
		TM	<i>gp38</i>	Доступ	A на G, D82G

			(SEQ ID NO:2)	NC_010326.1, 36462 (положение 245 SEQ ID NO:2)	
		TM	gp38 (SEQ ID NO:2)	Доступ NC_010326.1, 36464; 36465 (положение 247, 248 SEQ ID NO:2)	AT на TC, I83S
		TM	gp40 (SEQ ID NO:3)	Доступ NC_010326.1, 38180 (положение 757 of SEQ ID NO:3)	A на G, N253D
LUZ19+gp34 L55Δ (LUZ19*)	LUZ19	Делеция	gp34 (SEQ ID NO:4)	Доступ NC_010326.1, 26664-26666 (положение 163- 165 SEQ ID NO:4)	CTG на -, L55
LUZ19+LKD16 gp18+gp34 L55A	LUZ19+ LK D16 gp18	Делеция	gp34 (SEQ ID NO:4)	Доступ NC_010326.1, 26664-26666 (положение 163- 165 SEQ ID NO:4)	CTG на -, L55
LUZ19+ gp49 LKA1	LUZ19	Вставка	gp49 (SEQ ID NO:51)	Доступ NC_010326.1, 42719-42943 (SEQ ID NO:25)	P32 (SEQ ID NO:21) gp49 LKA1 (SEQ ID NO:12) T32 (SEQ ID NO:22)
LUZ19+ gp34 NTU	LUZ19	Вставка	gp49 (SEQ ID NO:51)	Доступ NC_010326.1, 42719-42943 (SEQ ID NO:25)	P32 (SEQ ID NO:21) NJ\Jgp34 (SEQ ID NO:13) T32 (SEQ ID

					NO:22)
LUZ19+Pp15gp44	LUZ19	Вставка	gp49 (SEQ ID NO:51)	Доступ NC_010326.1, 42719-42943 (SEQ ID NO:25)	P32 (SEQ ID NO:21) Pp15gp44 (SEQ ID NO:14) T32 (SEQ ID NO:22)
LUZ19+SaPSMa3	LUZ19	Вставка	gp49 (SEQ ID NO:51)	Доступ NC_010326.1, 42719-42943 (SEQ ID NO:25)	P32 (SEQ ID NO:21) SaPSMa3 (SEQ ID NO:16) T32 (SEQ ID NO:22)
LUZ19+SaPSMb2	LUZ19	Вставка	gp49 (SEQ ID NO:51)	Доступ NC_010326.1, 42719-42943 (SEQ ID NO:25)	P32 (SEQ ID NO:21) SaPSMb2 (SEQ ID NO:17) T32 (SEQ ID NO:22)
LUZ19+SePSMa	LUZ19	Вставка	gp49 (SEQ ID NO:51)	Доступ NC_010326.1, 42719-42943 (SEQ ID NO:25)	P32 (SEQ ID NO:21) SePSMa (SEQ ID NO:18) T32 (SEQ ID NO:22)
LUZ19+dspS	LUZ19	Вставка	gp49 (SEQ ID NO:51)	Доступ NC_010326.1, 42719-42943 (SEQ ID NO:25)	P32 (SEQ ID NO:21) dspB (SEQ ID NO:15) T32 (SEQ ID NO:22)
LUZ19+Pp15gp44	ШКХ LUZ19	Вставка	gp49 (SEQ ID NO:51)	Доступ NC_010326.1, 42719-42943 (SEQ ID NO:25)	Pp15gp44 (SEQ ID NO:14)
LUZ19+SePSMa	ШКХ LUZ19	Вставка	gp49 (SEQ ID NO:51)	Доступ NC_010326.1, 42719-42943	SePSMa (SEQ ID NO:18)



				(SEQ ID NO:25)	
LUZ19+PPR1L	LUZ19	Вставка	gp49 (SEQ ID NO:51)	Доступ NC_010326.1, 42719-42943 (SEQ ID NO:25)	P32 (SEQ ID NO:21) PRR1 L (SEQ ID NO:20) T32 (SEQ ID NO:22)
LUZ19+MSR L	LUZ19	Вставка	gp49 (SEQID NO:51)	Доступ NC_010326.1, 42719-42943 (SEQ ID NO:25)	P32 (SEQ ID NO:21) MS2 L (SEQ ID NO:19) T32 (SEQ ID NO:22)
LUZ19+pyoS5	LUZ19	Вставка	gp49 (SEQID NO:51)	Доступ NC_010326.1, 42719-42943 (SEQ ID NO:25)	P32 (SEQ ID NO:21) pyoS5 (SEQ ID NO:6) T32 (SEQ ID NO:22)
M13MP18+пап рика	M13MP1 8	Замена	lacZ (SEQID NO:28)	Доступ X02513, 6216-6722 (SEQID NO:28)	паприка (SEQID NO:29)
Лямбда c11 удаление	лямбда	Делеция	e11 (SEQID NO:30)	Доступ NC_001416.1, 38390-28623 (SEQ ID NO:30)	e11 удаленная (SEQID NO:9)
ЦМВ человека+от редактирова нный RL13	фрагме нт ЦМВ Челове ка	Замена	RL13 (SEQID NO:33)	Неотредактирова нный полноразмерный фрагмент (SEQ ID NO:10) и неотредактирова нный RL13 (SEQ ID NO:33)	Отредактированн ый RL13 (SEQID NO:32) и отредактированн ый полноразмерный фрагмент (SEQ ID NO:11)

TM - точечная мутация, замена - замена, делеция - удаление, вставка - инсерция.

Таблица 2. Сконструированные вирусы, описанные в изобретении

Сконструированный фaг	Пример	Фигура	Свойство
LUZ19+ фрагмент gp7 ФКФ77	1	2	Сконструированный ОП
LUZ19+LKD16 gp18	II	3	Расширение круга хозяев
ШКХ-LUZ19	III	4	Расширение круга хозяев
LUZ19+gp34 L55Δ (LUZ19*)	IV	5	Улучшенная литическая активность
LUZ19+LKD16 gp18+gp34 L55Δ	V	6	Итеративное конструирование
LUZ19+LKA1gp49	VI	7	Дисперсия биопленки
LUZ19+ gp34 NTU	VI	7	Дисперсия биопленки
LUZ19+Pp15gp44	VI	7	Дисперсия биопленки
LUZ19+SaPSMa3	VI	7	Дисперсия биопленки
LUZ19+SaPSMb2	VI	7	Дисперсия биопленки
LUZ19+SePSMa	VI	7	Дисперсия биопленки
LUZ19+dsps	VI	7	Дисперсия биопленки
ШКХ-LUZ19+Pp15gp44	VI	7	Итеративное конструирование
ШКХ-LUZ19+SePSMa	VI	7	Итеративное конструирование
LUZ19+PPR1L	VII	8	Антибиотическая сенсibilизация /предовращение устойчивости к фагу
LUZ19+MS2 L	VII	8	Антибиотическая сенсibilизация
			/предовращение устойчивости к фагу
LUZ19+pyoS5	VIII	9	Предовращение устойчивости к фагу
M13MP18+паприка	XI	11	Сконструированный ОП
делеция c11 λ	XII	12	Сконструированный ОП
ЦМВ человека+отредактиро ванный RL13	XIII	13	Коррекция ошибок/сконструированный ОП

ОП - опытный образец.

Таблица 3. Вирусы дикого типа, описанные в изобретении

Название вируса дикого типа	Геномная последовательность
Фага LUZ19 <i>P. aeruginosa</i>	Доступ NC 010326.1
Фаг <i>E. coli</i> \ dl 857 SAM7	Доступ NC 001416.1
Фаг M13 <i>E. coli</i>	Доступ X02513
Фаг 14-1 <i>P. aeruginosa</i>	Доступ NC_011703.1

Таблица 4. Последовательности дикого типа, описанные в изобретении

Название	Последовательность нуклеиновой кислоты	Аминокислотная последовательность
gp13 LUZ19	SEQ ID NO: 1	SEQ ID NO:34
gp38 LUZ19	SEQ ID NO:2	SEQ ID NO:35
gp40 LUZ19	SEQ ID NO:3	SEQ ID NO:36
gp34 LUZ19	SEQ ID NO:4	SEQ ID NO:5
gp49 LUZ19	SEQ ID NO:51	SEQ ID NO:49
gp18 LUZ19	SEQ ID NO:50	SEQ ID NO:48
gp18 LKD16	SEQ ID NO:26	SEQ ID NO:38
gp49 LKA1	SEQ ID NO: 12	SEQ ID NO:39
PyoS5	SEQ ID NO:6	SEQ ID NO:37
gp34 NTU	SEQ ID NO: 13	SEQ ID NO:40
gp44 Pp15	SEQ ID NO: 14	SEQ ID NO:41
DspB	SEQ ID NO: 15	SEQ ID NO:42
SaPSMa3	SEQ ID NO: 16	SEQ ID NO:43
SaPAMb2	SEQ ID NO: 17	SEQ ID NO:44
SePSMa	SEQ ID NO: 18	SEQ ID NO:45
MS2L	SEQ ID NO: 19	SEQ ID NO:46
PRR1 L	SEQ ID NO:20	SEQ ID NO:47

Из вышеприведенного описания специалист в данной области техники может легко определить основные характеристики изобретения и не отходя от его сущности и объема может вносить изменения и модификации в изобретение, чтобы адаптировать его к различным условиям применения и использования, а также реализовывать изобретение в полном объеме. Предшествующие конкретные варианты реализации изобретения должны толковаться как просто иллюстративные, а не ограничивающие объем изобретения каким-либо образом. Полное раскрытие всех заявок, патентов и публикаций (включая справочные руководства), упомянутых выше и на фигурах, полностью включено в настоящее описание посредством ссылки.

**Перечень последовательностей**

SEQ ID NO: 1

ДНК

Род/вид- Phikmvlikevirus LUZ19

Описательное название- *gpl3* LUZ19 дикого типа

GTGCTGGCCCTCGGTGCCTTCGACCTGTCCGGCCTGATGGTAGGTTTCCTGCCTCGTAGTAGGTG  
GTGAGCTG  
AAGGCCCTGTGCGTTGATGACCGGCACAGCAGGCAGGGTATCGGCGCTGAGCTGGTACGGGCCG  
CTGAGCT  
GGCTGGTGCCGAGTATCTGACCTGCTTCGAGTTCCTGGAGCCGTTCTACGCCGACTTGGGCTGG  
AGCACCAC  
CCACCGGAGGCCGAACCTGGACAGCAGGAGAGCCGGACGTGCTGCACATGAGGGCACCCGGTCAT  
GACGTAT

GA

SEQ ID NO:2

ДНК

Род/вид- Phikmvlikevirus LUZ19

Описательное название- *gp38* LUZ19 дикого типа

GTGGCTCGGTTCAAGAATCCCGAGACCATCCACGTTGCAGATGGGGTCGAGGCTGTCTTCAGTC  
 TCGACTTC  
 CCGTTCCTGCGGCGTGAGGACGTATTCGTCCAGGTCGATAAGATACTCGTCACCGACTATACGT  
 GGGTAGAC  
 GACACCAACATCCAATTGGCCGTGGTGCCGAAGAAGGACCAAGAGGTCCGCATCTTCCGCGACA  
 CGCCCGC  
 CCAGGTCCCGGACACACAGTTCAGCCAGGACATCCCGTTCCTGCCTCGATAACATCGACGCGAAC  
 AACAAAGC  
 AGCTCCTGTACGCTGTGCAGGAAGGCATCAACACCGCGAACCTCGCTCTCGATGGCGTACTCGA  
 CGCGATCC  
 GTATCGCCGAGGAGGCTCGTCGCCTGGCGCAGGAAGCACTCGACGCCGCCAATGAGGCGCTTCG  
 CCGTGCC  
 CTGGGCTTCGCTGAGATTCGCACCGTGACCGAGGACTCGGACATCGATCCGAGCTGGCGCGGTT  
 ACTGGAAC  
 CGTTGCATCACCGCCGATAAACCTCTGACCCTGACCATGCAGATGGAAGACCCGGATGCACCGT  
 GGGTCGA  
 GTTCAGCGAGGTTCACTTCGAGCAGGCCGGTGTGCGTGACCTAAACATCGTAGCCGGTCTGGC  
 GTTACCAT  
 CAACCGTTTGCAGAACACCACCATGCAGCTCTACGGCGAGAATGGCGTGTGTACTCTCAAGCGG  
 CTGGGCGC  
 TAACCACTGGATCGTGTTCCGGGCCATGGAGGACGAATAA

SEQ ID NO:3

ДНК

Род/вид- Phikmvlikevirus LUZ19

Описательное название- *gp40* LUZ19 дикого типа

ATGTTTAAGACCGAAGTAAAGGGACGTTACACCCTGATTCGCCGCAAGGCGGACGGCACTCCGG  
 TGGAGAC  
 TCTGGAGTTCGACAACATCATTACGAATGCGGGCCTGGATTGGATCGCCGCTATGGATACCGAC  
 CTCATGGG  
 CGAACCCGTAGCGGTCAGCACTTCTACAGCCGATCCCAACCCGAGCGCACCCGCCATCCCGGAG  
 GTTGTGCA

ACGCACGTCCGCATCTGCCCTGGTGGAGGTACTACGTCGGGCCTGGATGGCGAGTGGCTGTTC  
 TGGCGGAG  
 GCGTTGGAGATTCCCGCAGGGCACCTAGCTGGTCAAGTCTGGCCACCGTGGGCCTCATCTGC  
 AACTCGGA  
 TCGTCGCTTCGAGAGTAACACGGGTGAGCTGATCCCGAAGGATACCCCGCTGTCGTACACTCGC  
 ATCAAGGA  
 CGCCGCCGGGCAGCCTACTACTCTGGTGGTGGCCGCTGACGAGATTCTGGATGTCCAGTACGAG  
 TTCCGCAG  
 CCGGCCGTAGGAACGGCTGAGGCCAAGTTCGTGATCTCCGGCGTGGAACGCACCTTCCGGCTG  
 ATCCCAA  
 GCCTTTTGCGAACCGTGCTAATCTCTCCGGGAACGCTACATCTTCTACAACACCAACCCCTAC  
 ATCAACGG  
 CAAGGACGCCCTCCGGCGGCAATGTCCGAGACGGTCAGTGGCAGAAGAAATATCCCAAGTACGTG  
 CGCGGCT  
 CCTACAAGGCGCAGATCACGCTGCTGGCCCAGGTCCAGAACGGCAATATGGCTGGCGGCATCAC  
 CGGCACC  
 GAGGAACTCCAGATTTACAATGGACGTAACATATGTGCTCGATATCAACCCGCCTGTTGTGAAGA  
 ACAATACC  
 CAGGAGTTCACCGTGACCCTGGAGTTTACGGTGGCGAGGGCATAA

SEQ ID NO:4

ДНК

Род/вид- Phikmvlikevirus LUZ19

Описательное название- gp34 LUZ19 дикого типа

ATGAGСТАСААGСААТССGCGTATCCCAATCTGCTGATGGGTGTGAGCCAGCAGGTGCCCTTCG  
 AGCGCCTG  
 CCGGGCCAGCTCAGCGAGCAGATCAACATGGTATCCGATCCCGTGTСAGGACTTCGGCGGGCGCA  
 GCGGTAT  
 CGAGCTGATGGCCCACCTGCTGCATACCGACCAGCCCTGGCCGAGGCCGTTCTCTACCACAG  
 AACCTCGG  
 TGGCCGCAGCATTGCGATGCTGGTGGCGCAGCACCGTGGCGAGCTGTACCTGTTСGACGAGCGG  
 GACGGTC  
 GCCTGCTGATGGGTСAGCCCCTGGTGCATGACTACCTCAAGGCCAACGATTACAGGCAGCTACG  
 GGCCGCCA  
 CGGTGGCCGATGACCTGTTСATCGCCAACCTGAGTGТAAAGCCCGAGGCCGACCGCACCACAT  
 CAAGGGC  
 GTAGACCCCAACAAGGCCGGCTGGCTGTACATCAAGGCAGGCCAGTATTCGAAGGCATTCTCCA

TGACCATC  
AAGGTCAAGGACAACGCCACCGGCACCACCTACAGCCACACGGCCACCTACGTGACGCCGGACA  
ACGCCAG  
CACGAACCCCAACCTCGCTGAGGCGCCATTCCAAACGAGCGTAGGCTACATCGCGTGGCAGCTC  
TACGGCA  
AGTTCTTCGGTGCGCCGGAGTACACTCTGCCAACTCGACGAAGAAGTACCCGAAGGTAGACCC  
GGACGCC  
AACGCGGCAACCATAGCCGGTTACCTCAACCAACGGGGCGTGCAGGACGGGTACATCGCGTTCC  
GTGGCGA  
CGCCGATATCCACGTTGAAGTGTCCACGGACATGGGCAACAACCTACGGCATAGCCTCCGGCGGT  
ATGAGCCT  
CAACGCCACGGCAGACCTGCCGGCCTTACTGCCGGGCGCGGGTGCTCCTGGCGTGGGTGTGCAG  
TTCATGGA  
CGGCGCTGTCATGGCCACCGGCTCCACCAAGGCCCGGTATACTTCGAGTGGGATTCCGCTAAC  
CGCCGCTG  
GGCAGAGCGGGCCCTACGGCACCGATTGGGTCTGAAGAAGATGCCACTGGCCCTGCGCTGG  
GATGAGG  
CTACCGACACCTACAGCTTGAACGAGCTGGAGTATGATCGACGTGGCTCCGGCGACGAGGATAC  
GAACCCC  
ACGTTCAACTTCGTACCCGAGGCATCACCGGCATGACGACCTTCCAGGGTCGCTCGTCTCC  
TGTCGCAG  
GAGTACGTCTGCATGTGCGCCAGTAACAATCCACACCGCTGGTTCAAGAAGTCGGCAGCCGCGC  
TGAACGA  
CGATGATCCTATCGAGATCGCAGCCCAGGGGAGCCTGACTGAACCGTACGAGCACGCGGTACCC  
TTCAACA  
AGGACTTGATCGTCTTCGCCAAGAAGTATCAGGCCGTGGTCCCCGGTGGCGGCATTGTAACCTCC  
CCGGACGG  
CGGTTATCAGCATCACACGCAGTACGACCTCGATACCAGGGCGGCACCTGCCGTGACTGGCCG  
CAGTGTGT  
ACTTCGCTGCGGAGCGTGCCCTGGGTTTTTCATGGGCTGCATGAGATGGCCCCGTCTCCGTCCAC  
GGACAGCC  
ACTACGTGCGCGAAGACGTTACCAGCCACATCCCAGCTACATGCCGGGGCCTGCTGAGTACAT  
CCAGGCG  
GCGGCCCTCCAGCGGCTACCTGGTGTTCGGCACCAGCACGGCGGACGAGATGATCTGCCACCAGT  
ACCTCTGG  
CAGGGCAACGAGAAAGTGCAGAACGCGTTTTTCATCGCTGGACGTTGCGGCATCAGATCATCGGCG

CCTACTTC  
 ACTGGTGACAACCTGATGGTTCTGATTGAGAGGGCCAGGAGATCGCCCTGGGACGGATGCACC  
 TGAACAG  
 CCTGCCAGCCCGTGAGGGTCTGCAATACCCTAAATACGACTACTGGCGGCGTATCGAGGCGACC  
 GTCGATGG  
 TGAGCTGGAAGTGAACCAAGCAGCATTGGGACCTGATCAAGGATGCCTCTGCCGTGTACCAGCTA  
 CAGCCTGT  
 GGCCGGCGCCTACATGGAGCGTACCCATCTCGGCGTGAAGCGCGAGACGAATACGAAGGTGTTC  
 CTCGACG  
 TGCCCGAGGCCGTGGTTCGGGGCGGTGTATGTGGTTCGGCTGCGAGTTCTGGTTCGAAGGTGGAGTT  
 CACTCCGC  
 CGGTTCTCCGGGACCAATGGCCTGCCCATGACCTCGACCCGTGCAGTGCTTCATCGGTACAA  
 CGTAAACT  
 TCGGCTGGACCGGCGAGTTCCTGTGGCGCATCAGCGACACGGCTCGACCCAACCAGCCGTGGTA  
 CGACACG  
 ACGCCCTCCGGTTGTTTCAGCCGGCAACTCAATGCCGGGAGCCTCTGGTGGATAGCGCTGTGG  
 TGCCCGCTG  
 CCGGCACGGGTTCGATATGGCCACGTCCAAGTTCGAGCTGAGCTGTCACAGTCCGTACGACATGA  
 ACGTTCGG  
 GCTGTTCGAGTACAACCTCAAGTCCAACCAACCTACAGGAGGGTGTGA

SEQ ID NO:5

Белок

Род/вид- Phikmvlikevirus LUZ19

Описательное название- белок Gp34 LUZ19 дикого типа

MSYKQSAYPNLLMGVSQQVPPERLPGQLSEQINMVSDPVSGLRRRSGIELMAHLLHTDQPWPRP  
 FLYHTNLGGR  
 SIAMLVAQHRGELYLFDERDGRLLMGQPLVHDYLNKANDYRQLRAATVADDLFIANLSVKPEADR  
 TDIKGVDPN  
 KAGWLYKAGQYSKAFSMTIKVKDNATGTTYSHATATYVTPDNASTNPNLAEAPFQTSVGYIAWQL  
 YGKFFGAPE  
 YTLPNSTKKYPKVDPDANAATIAGYLNQRGVQDGYIAFRGDADIHVEVSTDMGNNYGIASGMS  
 LNATADLPA  
 LLPGAGAPGVGVQFMDGAVMATGSTKAPVYFEWDSANRRWAERAAYGTDWVLKKMPLALRWDEA  
 TDTYSL  
 NELEYDRRGSGDEDNPTFNFVTRGITGMTTFQGRVLVLLSQEYVCMSASNNPHRWFKKSAAALN  
 DDDPIEIAAQ



GSLTEPYEHAVTFNKDLIVFAKKYQAVVPGGGIVTPRTAVISITTYDLDTRAAPAVTGRSVYF  
 AAERALGFMGL  
 HEMAPSPSTD SHYVAEDVTSHIPSYMPGPAEYIQAAASSGYLWGTSTADEMICHQYLWQGNEKV  
 QNAFHRWTL  
 RHQIIIGAYFTGDNLMVLIQKGQEIALGRMHLNSLPAREGLQYPKYDYWRRIEATVDGELELTKQ  
 HWDLKDASA  
 VYQLQPVAGAYMERTHLGVKRETNTKVFLDVPEAVVGAVYVVGCEFWKVEFTPPVLRDHNGLP  
 MTSTRAVL  
 HRYNVNFGWTGEFLWRISDTARPNQPYDTPPLRLFSTRQLNAGEPLVDSAVVPLPARVDMATSK  
 FELSCHSPYD  
 MNVRAVEYNFKSNQTYRRV

SEQ ID NO:6

ДНК

Род/вид- *Pseudomonas aeruginosa*

Описательное название- последовательность PyoS5

ATGTCCAATGACAACGAAGTACCTGGTTCCATGGTTATTGTCGCACAAGGTCCAGACGATCAAT  
 ACGCATACT  
 GAGGTTCCCCCTATCGATAGCGCGCCGTTGCCGGAATATGTTTGGCGACTTAATTCAAAGAG  
 AAATATAT  
 CTACAGAAAAACATTTATTATCCAGTCCGATCTATTTTTGAACAAGGAACAAAAGAAAAGAAGG  
 AGATCAA  
 CAAGAAAGTATCTGATCAAGTCGATGGCTTGCTAAAGCAGATCACTCAAGGAAAAAGGGAGGCC  
 ACAAGGC  
 AAGAGCGAGTCGATGTCATGTTCGGCAGTCCTGCACAAGATGGAATCTGATCTTGAAGGATACAA  
 AAAGACC  
 TTTACCAAAGGCCATTCATTGACTACGAAAAGCAGTCAAGCCTCTCCATCTATGAGGCCTGGG  
 TCAAGATC  
 TGGGAGAAGAACTCTTGGGAAGAAAGAAAGAAGTACCCTTTTCAGCAGCTTGTTAGAGATGAAC  
 TGGAGCG  
 GCGGTTGCCTACTACAAACAAGATTCCTCTCTGAAGCGGTAAAAGTGCTAAGACAGGAGCTC  
 AACAAGC  
 AAAAAGCGCTAAAAGGAAAAAGAGGACCTCTCTCAACTGGAGCGGGACTACAGAACCCGAAAGGC  
 GAATCT  
 CGAGATGAAAGTACAATCCGAGCTTGATCAAGCGGGAAGTGCTTTGCCTCCATTGGTCAGTCCA  
 ACGCCAGA  
 GCAATGGCTTGAACGTGCCACAAGACTGGTTACGCAAGCAATTGCTGATAAAAAAGCAGCTGCAG

ACCACAA  
 ACAATACTCTTATCAAGAATTCCTCAACCCCTCTAGAAAAGCAGAAAGCCATCTACAATGGTGA  
 GCTACTTG  
 TGGATGAGATAGCCAGTCTACAGGCCCGCTTAGTTAAGCTGAACGCCGAAACGACACGACGCAG  
 GACAGAA  
 GCAGAACGCAAGGGCGCCGAGGAACAAGCGTTGCAAGATGCTATTAAATTTACTGCCGACTTTT  
 ATAAGGA  
 AGTAACTGAGAAATTTGGCGCACGAACATCGGAGATGGCGGCCAACTGGCCGAAGGCGCCAGG  
 GGGAAA  
 AATATCAGGAGTTCGGCGGAAGCAATCAAGTCGTTTAAAAGCACAAGGATGCGTTAAATAAAA  
 AACTTAG  
 CCTTAAAGATAGGCAAGCCATTGCCAAAGCCTTTGATTCTCTAGACAAGCAGATGATGGCGAAG  
 AGCCTTGA  
 GAAATTTAGCAAAGGCTTTGGAGTTGTAGGCAAAGCTATTGACGCCGCCAGCCTGTACCAAGAG  
 TTCAGAT  
 ATCTACGGAAACCGGGGACTGGAAACCATTCTTTGTAATAATTGAAACACTAGCTGCTGGTGCG  
 GCCGCCA  
 GTTGGCTTGTGGGTATTGCATTTGCCACGGCAACAGCCACTCTATAGGCATTCTGGGGTTCGC  
 ACTGGTAAT  
 GGCAGTTACCGGGGCGATGATTGACGAAGACCTTCTAGAAAAGCAAACAATCTTGTAATATCC  
 ATTTAA

SEQ ID NO:7

ДНК

Род/вид- Phikmvlikevirus LKD16

Описательное название - добавлена последовательность gp18  
LKD16

GAGTACCAACTGAACACGAGCGCACCCCTGCGCTGCCTGCTCCAAGACATCCACGGGCCGCTGAA  
 TCTGCTGT  
 TCCCAGGTATCCGGGTGAAGGTGGAGGAGGCGTGCCTCGGATACTTGGGCTACAGGGAGCGGGG  
 CTATTGG  
 GAGCTGCGCCTCCAGGTGGACTACGACCACCCGAAGCTTGGGCACCTCCGCTACAGTCAGGCCG  
 TGCCGGA  
 GTACGTGCTGATCAACGACCGCGACAGCATCATCAAGTACCTGATGGAAGCAGTCCCTCGGCAG  
 GТАCTAG  
 AGGGCATGCTCAATAAGGCCAGGAATTCGTAACCAAGAACTGGTATTCCCTATGACGAC

SEQ ID NO:8

ДНК

Род/вид- Phikmvlikevirus phi-KF77

Описательное название - добавлена последовательность gp7

ФКФ77

TACAAGGTGGTGACGCCTAGCTCGGCAGAGGGCGCCGTTGTGCTGGCGACCAAGCAGACGCCTG  
 CCCTCGC  
 TCAGGCAGTCAATCGTACTGCACAGCATGAACCCCGCGCAGTACGCGGTGGGCACGGCCATACTA  
 AACACAG  
 ACTGGCGGTGCCGCCGCTGGGTGCCGGCGAGTACATCAAGCTCGTTCAAGGGGAGGCCGAC

SEQ ID NO:9

ДНК

Род/вид- *Lambdavirus* лямбда

Описательное название- ell фага  $\lambda$  *E. coli*

ATGGTTCGTGCAAACAACGCAACGAGGCTCGTTCGAAACAAATCCAGATGGAGTTCTGA

SEQ ID NO: 10

ДНК

Род/вид - цитомегаловирус ЦМВЧ

Описательное название - предварительное редактирование  
 фрагмента ЦМВЧ

ACGACGGCCAGTGAATTGTAATACGACTCACTATAGGGCGAATTCGAGCTCGGTACCCGATTAC  
 CCTGTTAT  
 CCCTACCATTCCGGGCCGTGTGCTGGGTCCCCGAGGGCGGGGGGTGTTTTTAGCGGGGGGT  
 GAAATTTG  
 GAGTCTTGAGACCGCGTGTGCTGTGGAGGACGGTGACGGTGGTAAGAGTGTGCTGCGGTGCGGT  
 TGGGACG  
 GCGGCGCGAATAAAAGCGGCGTGCGGCGCGCACGGCGAAAAGCAGACGCGCGTCTGTGTTGTG  
 TGTCTTT  
 GACCGCGGCGGAACACACGCGGAAAAGCGAGTCCCAGGGGACACACGACGAGCGAGTCCCAGGG  
 GGGGAC  
 GACGACGGCCAGGGACGCGGAAAACGACGCGGAAAAGAGGAAGTCCCAGGGGGACGGGCGGAAA  
 AGAGG  
 AAGCGCCTAGGGGACCGCGGGGGCAGGAACAGACGAAGTACGCCGCAACCCGCGTTCGAGGACAC  
 ACGCAG  
 AAGCGCCGCCAGGGGAGGGGGGGGGGGGACTCGCGGGCCCCGGGGCACACTTGTGTTCC  
 TCCGGCC

GCCGACACGCACCCCGAAGCCGCGCACACCCGCCGACACACCCCTGACACACCCGCGACACACCC  
GCCACAC  
GCCCCGACACACGCCCCGCGACACACCCGACCGACACACCCCTGACACACCCCGCCAACACACCCAG  
CCGCACC  
CGCCCCGCCAACACACCCCCGACACACCCGACACACGCCCGCGACACACCCGGCACACACCCAC  
CCACCCA  
GCCGCGCCCCGACACACCCCGAACGGCGCCGGTGCGGGACAGGGCTCACGGAGGTTTGCGGGC  
CGTGAGC  
ACGCCTCCCTTTGTACACACTACCGGTGCGTGCGGTCCCACGCTATTTGTTGCGGAGACCGGAC  
TAAGGGAG  
GTTTGCGGTGCGTCAGCGCGGGGCGGGCTTTGCGGGCGTGTTCGACCAGCGCTTTGTGCGCGCT  
GCCTGTGC  
GTGTCGTCCCATGGTCTTTGTGTCAGCGGCACGGCGCTGGGGACGGGGTTTCACCGCGCTGAGGGA  
TCTTTCTG  
CGGGTGTGAGGGACGGAGCTTTTTTTCGCACGCTGGGCACCGGGCTGGGGGACGGGGGGTGTGCG  
GGACGGC  
GGTGGGGCCGGGGCGTTGCGGGTACGGGGATTACGCTGGGAACGGGGACTCGCGGACCCGGGCT  
GAGGGA  
CGGGGGTGGCGGGGGTGTTCGCGGCGAGGACGGGGGCTTTTTCGCGGGGGACGGGGACTCACC  
CTCGCT  
ATTTAACCTCCACCCACTTCAACACACACATGCCGCACAATCATGCCAGCCACAGACACAAACA  
GCACCCAC  
ACCACGCCGCTTACCCAGAGTACCAACACACGTTACCCTTACACCACAGCAACACACAACCCG  
CTATCAA  
ACCTCGGACAAACACGCCAACGAAGAACCACGCAGATGGAGCTCGACGCCGCGGATTACG  
CTGCTTG  
CGCGCAGGCCCCGCCAACACCTCTACGCTCAAACACAACCCCAACTACACGCATACCCCAACGCC  
AACCCCTCA  
GGAAAGCGCTCATTTTTCCACAGAAAATCAACATCAACTCACGCATCTACTTCAACAATTGGC  
GAAGGCGC  
AGCGCTCGGCTACCCCGTCCCCCGCGCGAAATCCGCCGCGGGTGGCGACTGGGCGGACAGC  
GCGAGCG  
ACTTCGACGCCGACTGCTGGTGCATGTGGGGACGCTTCGGAACCATGGGCGCCAACCTATCGT  
GACCTTAC  
TGTTGGCGCGCCAACGCGACGGCCTCGCTGACTGGAACGTGCTACGCTGCCGCGGCACAGGCTT  
TCGCGCAC

ACGATTCCGAGGACGGCGTCTCTGTCTGGCGTCAGCACTTGGTTTTTTTTACTCGGAGGCCACGG  
CCGCCGTGT  
ACAGTTAGAACGTCCATCCGCGGGAGAAGCCCAAGCTCGAGGCCTATTGCCACGCATCCGGATC  
ACCCCCAT  
CTCCACATCTCCACGCCCAAACCACCCAGCCACCATATCCACCGCATCGCACCCACATGCT  
ACGACTCG  
CCCACATCACACGCTCTTTCCTATCCCTTCTACACCCTCAGCCACGGTTCACAATCCCCGAAAC  
TACGCCGTC  
CAACTTCACGCCGAAACGACCCGCACATGGCGCTGGGCACGACGCGGTGAACGTGGCGCGTGGA  
TGCCGGC  
CGAGACATTTACATGTCCAAGGATAAACGTCCCTGGTAGACGGGGTAGGGGGATCTACCAGCC  
CAGGGAT  
CGCGTATTTGCGCGCCACGCTGCTTACCAGATATCCAATAAAACCCATCCCCTCGCCACGACGTC  
TCCGCGTAT  
CTTTGTAGCCTCAGGAATCCGTCCCCACGTCCATCCATCCCGAGCACTCCACACGCTATAACAG  
ACCACGGA  
CACGGCAAATGCATGCAAACCTTCTCATTTATTGTGTCTACTACTCTGTGTTGCTACAGGGAGTG  
AAGGGGGT  
GAAGGCAAAGAAAAAAAAAAGGAACAAAATAATAGATTAGCAGAAGGAATAATCCGTGCGACCG  
AGCTTG  
TGCTTCTTTTCTTATAAGGAGGCAAATATACTAGGGAAAACCTAAGAATAGGAAGAAACCGAGG  
TTTGGA  
GAAAAGCTGAGATAAAATAGCGCATTTTCCATACAGAGGTTGTTGTTTTTGTGGATCCTAAGAG  
GTTTCAAG  
TGCGAATCTCAAAGTTCTCACGAGAATATTGTCTTCAAGAATCGACAACCTGTGGTCCAAGATTT  
TTTTTTGGT  
CTTTTTAGGTTCTGCGAGGGACATCACGATGGATCGTTGCGATGAAGTCACGCGTACGCCTCTG  
GTGTGGCG  
CGGTGTCGTGACAGGAGAGTGTGTTTTCAGTGCAGAGCTGTCTTGATTCCCTATATCCGAGTATC  
TGTTTTCTC  
GTAAGGACGGTAATCTTCTTTGGTGAAGTACATCTAAAAGCTGCAAACCTATATTTTAAGGGCT  
GTCTCTAG  
GTGTACTTTGATGCTGGAGTTTTTCGCTGTGTTGATGTGAATAAATCTACTACTACTATTATAT  
GCAGAAAGA  
GTGATTATGCCGAGACAAGATTGCATTGGCTGAACTGTTTCAAAAACGCCTACACTCTACTTAT  
CCGTAAAC

CTAAGGTAATACTATGTGTAAGTTGTTTTTTTTCTTTTTGTAGTAAAATGGTGATACGTGCAA  
TTAAAAC TG  
TATTCCATGTTTCCATCCTTTCATTTCAACTTTAAAGGCGGCTTTGAGAGCGAAGAAGTGCGAG  
GATAAAAA  
TGGATGACTCCTTCGTGTCCAGGGAGTCGACTACTGCAACGCTGATTGATTAAAAGATGGTCTC  
CGATGATG  
ATGTTGTTATTGATCGAATCATGGTGCAGAACGGCGACGGAGAGGAGCGTGTCCGCCCGCGGA  
AGGTGGT  
CTCTTCTCTTTTTCTTTTTCAAGAAATCTTCCATGTGTTTATCGTAGTGATCGAAATCGACTG  
ATCTCGGGTT  
CTTTTTGTTGGTTTCTTTTCGGTTAATCATGTATTGTTTTCTTTTTTACAGAAAGATACTTTT  
TTCATGAGCAA  
TTCCTCGCCCCGGCGCCGGCATGCCGAGGTGGGGCCACTGCGATCAGCGGCATGCCGACGCCGAC  
CCGGGGA  
TCTTGATTACCGTTTTCTCTCTCTCTCTACATACAGACCGGGTGGCAGGAGCGGTAAGG  
AATCATCGT  
CGTCTTTCATTCTTCGATGATTATGGTAATACTAAATCTTATCTAGGAGCATATACATCTAAGA  
TTGGAGTAC  
TAGTAGTCGTTTGTGGTTTCTATTTTTTTTTATATTTATCTATGACAGTTTTTCTGTTTTTCGTT  
TTGATAATAAT  
ATAATAAAAACTCATGGACGTGAAATCTGGCTTGGTTGTGGTGATTTCAATCTCATTATTGTTG  
TTTTCTTTCC  
GTCTTGCGGATGAAGATGTTGCGATGCGGTTGTTGTTGGTGTGCTATACACCGAGAGAGATGA  
TCTTTTTGT  
TCTTCTGGTTCAATTCCTATGATTGTTTGGCTGCTGACCGACGCGTCAGGATGTGCAGGGCATG  
CGGGGAATC  
AGGACCGGACACGGGATAATTTTCTCTACCTATACGGAGATCGCGGTCTCGCCATGAGGATCG  
CGACAGG  
CGCGTCGAGGGGGCAGGAACACCCTTGCAGATTGACATTTCTGGTGGTGTTCGTTGTTGTCGG  
TAGTTGTTG  
TTGACGATGAGGATAAATAAAAATGACCTTGTTTTTGTTCTGTTTTCTCTTGTGGGAATCGTC  
GACTTTGAA  
TTCTTCGAGTTATCGGAAAGCTGAGGTACCCAAATGTCTGTAGCTTTTTTCTTTTTACCCTCTT  
GTTTATCATC  
TGCGATTTCGTGGTAGGTAGGAGAGGAAATGATAATCCGAGATTAAGGAAAGGAGAAGATAAAA  
AATAAA

AAAAAAATAATAAAACAGAAGCCGACCGGCCGCCGACCCGTTCCCCAGGACCAGCCTACGAGGA  
ATGGATA  
ACGCGGTGGCGACGGCAGCGGTGGTGGCGCTGGGGGTGGCGGCAGTGGTACTGCTGATGGTAGT  
CGGGACG  
GAGGAGAGGCGATGCATACATACACGCGTGCATGCTGCATGGGTGGATGGTACGGCCGGGAGAC  
GCGGAA  
GAGAACTCACATAAAAAGGTGACAAAAAGAGCGGTTGAAAAAGAAAACGAGATTCGACCAGA  
CAGAAG  
AGAAGGACCGGGGCTTGGCGACCCCTCCACGACTGCTGTTGTCATCTCGGCTCCCCCGTCTTCT  
CCCGGCCAC  
GGGCGGCTAAGTCACCGCGTCTCCCCATCCGTCCGAGCGCCGACCGACCAGCCGGCCGATTC  
GCCCCCG  
GGGCTTCTGGAGAACGCCGGGGCAGCAGCGATCTGGGAAGCCGCTAAACCCCTGCGTTTTTAT  
ATGGTAGC  
TCTGCCGAGCGGGCTGACGCGTTGAGTAAGCGAAAGACGTGTGTGACGAAAAGGGGTCCCA  
TGGTATT  
TCACGTGACGATGAGGAGATGCGGTTTTGGAGCACATACGTTTTAGAAAAAGGGAGTTGTCTGTA  
CAAGGGC  
TGAGGGACCTCTGTCTCCATGTGTGTATAAAAAGCAAGGCACGTTCATAATGTAAAAAGAACA  
CGTTGTAA  
ACAAGCTATTGCTGTATCATTCCGGCTGACTATGCTTCATTCGGACTGATTTTCTTTTCTAACG  
GCGTAACTT  
AAAGTGATTAACGTATGATATTTGTTCCCCAGAGTTATACTATAGTCATCATCCTAAAATTCAG  
ATATAAATG  
AACACATGTCGTATGGGATTATTAAGAAACCGAAACTCTCCACAGTTCACCATCTTCTTCGTCA  
TTCAACCG  
ATGACCCACTCCGTACAACGAATCAGTCTGCTGTGTACACTGCAAACCTACTAGCGACGTATGC  
AAACAACCT  
TGAAACACGGGCTGTTGTATTGACGACCGTTGTACCATTACTAGTCACATTGCATAGAGACCAT  
CCACCGTC  
ATCCCATCTTTCCACCCGATGGAAAACCGTCTTCTATCATCAACTATGGTAAGATTTTCGACCC  
TGCGAGGTA  
TTCAGTTTCCCACATATCCATAACCTGGATTTTATCATTAACCCCAATATTAAACACTTTTTTA  
GTACCCCCC  
ACCCACCAAAAAATGTGACTGGACCGGTTCTTAGCAGCTCTGGGAGCCATGTTTCAGGTTGAACC  
ACAGCTAC

AGCGAAACCGAGTCCAGTGACCGGTAACCACGTCCAGCCCCCTGCGTATGTACCAGTCCAAGCAC  
GTCCGGTC  
ATTGTTCTACACAGGAAATCTAACTAGGTCAACGCAATTTTATTCACCGTTACGCAGAATACT  
AACAAACA  
AACACACAAATTTAACGAATTACACGTAGTTTATTACATGAAAACGTGAAGAACACCAATTCAC  
TAAGCGAT  
ACAACATTTAGCTGACTTCCAAGTGCCACACATCACCCTGTATTCATCCATGTTTTCCACCGAA  
CCAACGAG  
ACAGATCGAAGAAGCCAGAATCTCCCGACTTTAAATTACATAAAATCCAACGTATTATGACCACA  
GCTCGACA  
CACAAATAGTTGCGTTACTATTACAGTAGCATTACCTATACCCGTAACGTTGCACAACCACTG  
ATCACCATT  
GTTACCAAAAACGGTTTTCCACTTAGTTGTCAACGGATCTTTCCCATGCGTAATGGTCAAATTA  
CTACCAGTC  
GTCGCTTTTAGCTCATTACGAGTATTATCCGCATCCACATATATCAACGTCATAGCTAGGCACG  
CTATAAGTA  
CCCCCCCCCACAATGGAATGTTGCCAAACCGGTTCTTTCCCGTTATAGCCATAGCGTTCCAG  
GCAAAAGC  
AAACGCCAAACCTAATGCAGTGAAAAGCGCTTGCAGCCAGAACCAGCTTATGTACCAGCCACAA  
TCACATC  
CGGTTATTGTTTTCCACAGGAAATCCTACCAGGCAAAGCCCCGCTTGTTTTGTTCTGACCATCT  
TGTTTAGCA  
ATTCGTAAACTGTCAGCCTAGCGACGTCCGTTTAGATCAAAGTCACGTATATAGCGACGCTGT  
TTCCACCC  
GTTTCCCGTCCCGCCGTTTCCGAACAACCCACCCGGGTTTCAGACAACCGACCACCAACAGAAA  
TATACACA  
CAGACCACCGGAGTTCAGTTAAAGATTTTCATCAGGTTTATTTTGGCTGCTGCTAGTCTTTTGC  
TTCTTAGAA  
AAAAAATACCCATATAGAGAAATAATGATAGTTTGACAACACATATGGCAGGGATTTCTTCTTC  
ATCAATAA  
GATATGCAATTTCCCCAGGGAGAGACTTTCAACAATTGAATTTACAAAAACAAAATTACATCAG  
GAGAAAG  
AGAGGATACATTAATAAATATATTATATCTGGTGTATATACTGAATGCTGCTGGTTCATAAGGT  
AACGATGC  
TACTTTTTTTAATTCCAAGATGGTTTTTCTTTGTTAGTCTTTTGTGACTTGCTGGTTCCTAAA  
AGTTCGCAAA



AACGATTGTGTGAAGATTATGACGTTGGTTGACTAGTTCATGAGATTCTGCTGTACGTGTGATG  
GTTATTCGC  
TGGTTCGTTCTAAGATGAGTATCGTACTGTGTCTGCGATGGTCGTCTCTTACTGGCATTCTCTC  
GGCTGCCTCT  
TGTTTTCATGATTGAAAAGGAAAAAGGACTCCGAGGGCGCGGTCATCTTTACTTTTCGGTTT  
TCTCGTTGG  
CGGGTCAGAGGTAGTCAGATCATGAGACTGTCGTGGTCGATGAAACTGTGTCTGCTCAAGTGAC  
GTCCATTT  
CTGTACGGAGAAAAAAGTCATCGGGATAAATAAGGCTATACAAGGCGTTGTCAAGCGTGCGGC  
TCTAAAC  
AAATTAAGCGATACAAAATTACAGTGATACGAATAATAAATTACCCCTCCCCCTGTGGTCCCC  
CCGAGGCG  
AGAGCCACCCATCGTGTACTCTCGCACCACCACGACCACAGGGGGAGACGGGACGAAGAGACG  
ACGCAG  
AGGCCATCTCCTCCTGGAGGCCGGCGGCTTAAGTGTACAGCTGCGGCGGCGACGACAGCTG  
CGATTTGT  
CGGCCGACATGCCGATGGTATGGGCGGCGGCGGCGGTGGCCGCGGCGAGCGGGGAGGAGAGAGA  
GAGAAG  
AGGAGCGGGGCGTCCGAAGGCGAGGATGGCATGGTCTCGCCGAGCGCCCGCTTTTATGGAAC  
ACTCGCG  
TCCGGTTGGGTATCACCCACAGGAAGATGAATCACAACCTCCAAACCATCTTGAGACCCGAGTA  
ACGGTTTA  
CAGGTCGACGCCAGTCTCAGCTAAAAACAGCGGACAGTCCCACGCTGTTTCTGTTGTGGCTCT  
CTCCAGTTT  
CCTCATCGCCGCTCTGGTCTCCGTATCATCGGAAGAATACCACCCGCTCTCATGCGGCAGTCG  
ATCAGCCTC  
GATGAACGAGACCGGGCGACGCCTTTCTACGGCCGACTGGTTGTGGTGGTGAAAGAAGAGCACC  
AGCAATC  
CCAGGAGGAGCAACAAGCCCTCACATGTCCAGGAGGTCGGGGAGAGGGCCTGTCGGAGATGACC  
GTGAGG  
CATCACGTACGGCAGCTGAGGAGAAACGGAGAAGAAAGGAAAATTACCGTCAGGGGCCGGGTT  
CTTATTA  
GAGAAACAGCACGTAGGTCAGGATCCAGATGCTAATGGCAATCATGATGACGATGATCATGCAG  
GCCAAGA  
CGCGGCGACCAATGCAGAATCCAATAGCCGCCGTGCCTCCGTTGGTGGCCGGCGGCATCTAG  
AGACATG

ATTTGGGGGGGGACCGGCGGCGCAAAAAGACAGGGAGATGGACAGTGCCACGGTGT TTTGTTA  
TGATTAG  
GACATGGGGACCGGAAGCCGAGACAGAGTACTACAGGGTGTGAAGGGTAACGTGAGGGAGATC  
ATGTCAT  
GGCGGGCTGAAGACCGTGCGGGGAGGATCGACGTGTGCGGTGCTTGTGGAACACGGTGT TTTA  
ATATGTA  
TCCGCGTGAATGCACGCGGTGTGCTTTTTAGCACTCGGCTTGATAAGCTACGTGACCGTCTGC  
GCTGAAAC  
CATGGTCGCCACCAACTGTCTCGTGAAAACAGAAAATACCCACCTAGCATGTAAGTGCAATCCG  
AATAGTAC  
ATCTACCAATGGCAGCAAGTGCCACGCGATGTGCAAATGCCGGGTCACAGAACCCATTACCATG  
CTAGGCG  
CATACTCGGCCTGGGGCGGGGCTCGTTCGTGGCCACGCTGATAGTCCGTGGTGGTCTTCTT  
CGTAATTTA  
CGCGCGGAGGAGAGAAAAACAACACGGGCACCGAGGTAGATCAATGTCTGGCCTATCGGAGC  
CTGACAC  
GCAAAAAGCTGGAACAACACGCGGCTAAAAAGCAGAACATCTACGAACGGATTCCATACCGACC  
CTCCAGA  
CAGAAAATAACTCCCCGTTGATCGAACCGACGGGCACAGACGACGAAGAGGACGAGGACGACG  
ACGTTT  
AACGAGGAAGACGAGAACGTGT TTTGCACCATGCAGACCTACAGCAACTCCCTCACGCTTGTC  
TAGTCACG  
TCGCTGT TTTTATTACAGCTCAGGGAAGTTTATCGAATGCCGTGCAACCAATCAAAAAACCC  
TAAAGCTC  
GCCAACTACCGGCCACTTGCGAAAACCGTACACGCACGCTGGTTACCAGGCTTAACACTAGCC  
ATCACAGC  
GTAGTCTGGCAACGTTATGATATCTACAGCAGATACATGCGTCGTATGCCGCCACTTTGCATCA  
TTACAGAC  
GCCTATAAAGAAACCACGCGTCAGGGTGGCGCAACTTTCACGTGCACGCGCCAAAATCTCACGC  
TGTACAAT  
CTTACGGTTAAAGATACGGGAGTCTACCTTCTACAGGATCAGTATACCGGCGATGTCGAAGCTT  
TCTACCTC  
ATCATCCACCACGAGCTTCTGCCGAGCCTTGGAACGCGTCGATGCTTTTATCCGGGACCAG  
GCAGAGTC  
GGTGTGGTCACGGATTCCCAAGAGGCAGACCGAGCAATTATCTCGGATTTAAAACGCCAGTGGT  
CCGGCCTC

TCACTCCATTGCGCCTGGGTTTTCGGGACTGATGATCTTTGTTGGCGCACTGGTCATCTGCTTTC  
TGCGATCGC  
AACGAATCGGAGAACAGGACGTTGAACATCTGCGGACGGACCTGGATACGGAACCTTTGTTGTT  
GACGGTG  
GACGGGAATTTGGAATAAAAGATGCGTAACACCTGTGGAAGATGCGATAACTTTACATACAGGC  
AAACAGT  
GTATACAATTATAGTATTTTGTATGTTGCATAAAGTTACATGCAACAGTACTGCTAACAGTACT  
GCATCCATT  
ACGCTATCCAACACTGCCTCTACCACTTTTGTAAACCAACATATATTCAACTCCGAATAACAACA  
CATCAACG  
ACGCCACACACATCTGTCACCTCACAAGCGTCAACCATTGGCAACATCACCAACGTTACCTCCG  
ACTTGAGT  
ACTTTCACAACCGTATATTCTACATTCAATACATCATTGCCAATATATCTAATACGGCTGTCA  
CTACAGAAT  
TGATTTCAACAAATACCAACACTATCTCATCTTTTACCAACGTAACAGCAAACGCTACATCATC  
TTATAACAC  
AACAATCACCGTAACTGTCACGTCAGATGAACTTCGCACAACGTATCCACTAATAATGCACTT  
ATAAGCAC  
ACCATGGCCTACAAATTGCAGCGCCACAACATACACCACGTACAACCTTACTAACTCTTCCAAC  
GCTTGTC  
CACAGAGACAACAATCATACTTTCAAGGAAACCAATACAACAGGAATAGAAGGGAGTAATGTC  
ACCATAA  
AGGGTAATTCTACGTGGGACTGTCTTTCAGTCGCCTGGATACGACATTACAATAGATCCACACA  
CGGACATC  
ATCTAGGTTATCGTAAGAACGCACATACCCAATCTTGGTATTGGCTACGCATCCTTACCTCTCA  
CACTGTATG  
TCATTCTCAACATGAAAGACCTTCACTGTACCATGACTTATGTGTTTCGTGCAACAACACAGAA  
TTACATCTG  
TACGATCTAAATATCACCAATTCCGGCAGGTACAGCAGACGTTGTTTTAAAGAAAATTACTTCA  
CAGGACAT  
CACGAAGATGAAAATTTCTACCTATTAGTAACACCAAAAAATCATACTGAAGCTATTAATGCTA  
CTTTCGTT  
TGCCCTAGATACAACACCGATATCGAAAATGAAGATAGAGAGAAAGGAAGTCAACATACTAACA  
ATACACA  
TCACCACAAACGTAATCTCTATCATAGCTCGCAAAGAAGCCGCACCGTATGGACCATCGTGTTG  
GTTTGTAT

GGCCTGCATAGTTCTGTTTTTTGCACGACGAGCCTTTAACAAAAAGTATCATATGTTACAAGAC  
ACCGTCAG  
TGAATCAGAATTCATTGTTTCGATATCACCCAGAACATGAAGATTGAGCTACGTTTCCGGGCAGA  
CATCTTAT  
GAAGCTGAACAATAAACTAAAACATTCTGTAAGACTCAGCGTTCAAAGGAATATTAATGCCCAT  
TGAGCGA  
AAACTAATATTGCAATGGACTGGCGATTTACGGTTACGTGGACCGTTACTTGTGATGGTTTCAA  
TTATACAGT  
CCATAAAAGATGCGATCGCAGTTACGAGGTAATCAACGTAACAGGATACGTTGGTAGCAACATA  
ACTCTAA  
AAAAATGCAATCAGACTGAGAAATGGCACAATGTAGACTGGATTCAATTATGAGTACCCACGCA  
TAAAATG  
TGCGAATTAGGCAACTATCACCAAACCACACCACGGCACGACATATGTTTTGACTGCAACGACA  
CCTCCCTA  
ACTATCTACAACCTAACCCAAAAAACGCTGGAAAATATACCAGGCGTCACCGTGATAACGGTC  
AAGAAGA  
AAATTACTACGTAACGGTGTTAATTGGAGACACAACGTTATTCACCTTGGCACATGCCCTGTA  
AGATATAA  
AGAATCTACGAACTGAAAACACCATTGGAAGTAGCATCATAGAAACCATTGAGAAAGCTAAC  
ATTCCCC  
TGGGAATTCATGCTGTATGGGCAGGCGTAGTGGTATCAGTGGCGCTTATAGCGTTGTACATGGG  
TAGCCATC  
GCATTCCCAAAAAGCCGCATTACACCAAACCTCCCAAATATGATCCAGATGAATTTTGGACTAA  
GGCTTAAC  
ATGCTGATCAATAAACTTTTTTTAACCAATAACATGTCTCCGTTTTTTTTTGTTAACAACCTAT  
GATATAAAG  
CGTTATATTCAGTCGTTACTAAACAAAAAACATGGGCATGCAATGCAACACTAAATTGTTATT  
GCCAGTCG  
CACTAATACCGGTTGCAATCATCCTAATTGGTACTCTAGTGCCGATACTTTTACATGAACAAAA  
AAAGGCGT  
TTTACTGGCGACTTTTTCTGCAAAGTCAACATGTAGAAGCACCCATTACAGTAACGCAGGGAGA  
CACAGTCT  
ACCTAGACGCTAGCAATAATCCCTGTAATTATTCCAGCTTTTGGTACCACGGTAATTGCGAACT  
TTGTGGATG  
GAACGGATATCTACGCAATGTTACACATTACTACACAAACACATCGTGTTCCCGCAATTCATC  
TGCATAAA

CGAAACTAAAGGTCTGCAGTTATATAATGTAACATTAACGATTCAGGCGCTTATACTGAACAC  
GTTTACGA  
ATGTGACCTTTCGTGTAACATTACTACTAATAACGAATATGAAATACTCAATTATTTTGATAAC  
TGTAECTAC  
ACCATAAATAGCACCAAGCATATTATCACCGTGGTGTCTTCACGTCATTCTAAACAAACAAATT  
CCCACGTA  
TCCACTCACGCTGGTTGGGCAGTCGCCGTGGTGACGGTAATTATGATCTACGTTCTGATCCACT  
TTAACGTCC  
CGGCAACTCTGAGACACAAACTACGAACTAGAAACAACGTAAATCGCATAGCGTGATTATAAAG  
TATCGAC  
GCTAATTTCTCCAAGATAAAAATTTGATTACTCCGTGCAGTTCTCAAAAAGTAAAGGCCCGCT  
TTTCCACTC  
CGTCATGAAGGATCGCAATAGAATACTGCTATGTATCATCTTTATTTGCATTATGTGCCTCATT  
TGTATTTAC  
TTTAAACGTCGTTGTGTTTTTACTCCGTCTCCAGACAAAGCAGATCTGCGAGTGGAAATTTCCCT  
CGTTACCCC  
CGTGTATTGGCATAACAGTGCCTGCATGAGAACACGCGTGACACATAGCGTACCCCTGGACGGT  
ACAGTTTA  
TGATAACGTAATTCAGGGAAAGTATACATTCATACCAACATGTTATCACATAACACACAGATTT  
TCTGCGTG  
TTTTATAAAAGAGCGTCTCGAAGCAGCTTGAGCCACACTACGGTCCAGATGACGAGCGTAATTA  
AAAATATG  
CCGCGCAGTATTCGAAAGCCGTACTGAGCGTGCGAGGCGGGTAGGGTGCCGAACGACGGATATG  
CGTCGTT  
GTCATCTTCGACTATAAGGATCGCGACCGAGTCTTCGGCCATGGTAAACGTCACCCTGTGTGGC  
TGGTATGT  
AGCGTATCCGGTTTGGAAATGTTCTGCTCCAGCTCGGGGATAGTGAGGAATTCTCAAGGGATA  
CGGGACCC  
AATGACTGGATAAGAGAAGGGTTTTTCCCCGTAAGATGATCCTCGTATCACATGAGGTCTGGAT  
ATGTATAA  
ATGAAGAGTGAAATAGGCACAGGAATCAGATGCCAGCCTCGTGATGCAGCCGCTGGTTCTCTC  
GGCGAAG  
AAACTGTCGCTTTGCTGACTTGCAAATACATCCCGCCTTAAGCGATGAGTCTATAAAGCACCG  
TTGCCCGA  
GTACGGTAAAAGTGACCCGATTGTAGAACGTCCTTTTTTTTTGTTTTTGCATCGTTTATCGTC  
ACTACTAGT

GCAATATTTTGGATTGTAAGGCTGAAAGAGTATCGTTATGATGCTTAGAACGTGGAGATTATTAC  
AGATGGTA  
CTGCTTGCCGCGTACTGTTATTATGTTTTTGGCAGTTGTTCAATCAGCACGACGACTGCTCCTG  
TGAATGGA  
AGTCTCCCGACCGTCAGATTCCCAAGAATATTACCTGCGCTAATTACTCAGGGACCGTCAACGG  
CAACGTTA  
GTTGTTCTTCCCGAAACTCCAGGGTAACTATGACGAACAACATTACAGATATGAAGTAGCGAA  
CCTGACGT  
ATAACTGCACCTATAACCGCTTGACGTTGCTGAATCTGACGACGGAAAACAGCGGAAAGTACTA  
TTTCAAAA  
GGGAAGATGCGAATTTACCTTCTATTACTCTTGTACAACTTGACCGTGTCTAAAGATCGCA  
CGTGAAGTT  
TCACAGAGCCGCGTGGCTGTAGCTATTGTGTTTACGTTGCTTTTGAAATGTTAAGCGTCCCTAC  
GGCGCTAAC  
ATGTTTCTAGGCTACTCTGACTGTGTAGATCCCGCCTTGCTGTGTATCGTGTATCTAGATCAC  
GCTTAAAGC  
TCATGTTGTCTTTTGTGTGGTTGGTCGGTTTTCGCTTTCTATGATTGTGCCGCGTTTCGAGTCCTG  
CTGTTACGAC  
ATCACCGAGGCGGAGAGTAACAAGGCTATATCAAGGGACGAAGCAGCATTACCTCCAGCGTGA  
GCACCCG  
TACACCGTCCCTGGCGATCGCGCCTCCTCCTGACCGATCGATGCTGTTGTCGCGAGAGGAAGAA  
CTCGTTCC  
GTGGAGTCGTCTCATCATCACTAAGCAGTTCTACGGAGGCTGATTTTCCACACCACCTGGGTC  
ACCGGCTTC  
GTCTGCTAGGACTCTTGACGCTTTTCGCCAGCCTGTTTCGCGTACCGCAATCCATCTGTGCTT  
TCTGCATAG  
ACCGTCTCCGGGACATCGCCCGTCTCTGAAATACCGCTATCAACGTCTTGTGCTACCGTGTA  
GCTAGTTAG  
CCAGCTGTGTAGTGTGTTTTGCTTTTGCATATTTGTTTTTCAGTCAGAGAGTCTGAAACGGGGTG  
GGAGGGACT  
TTTGCGGGTAGTGCATGCTAAGATGAACGGGTGGGCTGGGGTGTGCTTGATAACTCACTGTTTG  
AATACGCG  
CTCACGCACATATGTAGCACTCAACATGTTAGCTTTTGGCCGCACGCCCCGGGGCGTGCCGAGC  
TGCCTTTTT  
AATAAAGTCTGGGTTTTCCAGATACGCGCTGGTTCTGATTTTGATGGTTTGTGCCTCTGAAAGCT  
CTACGAGCT

GGGCCGTGACATCCAATGGACTGCCTAACTGTAGCACGGTAACTAGAACAGCGGGTCAAGACGC  
TGAATTG  
CACGGTCCGGCACCCTAAGCTGTAATGTGACCCAGTGGGGACGTTACGAGAATGGAAGCACAC  
CCGTGTT  
ATGGTGCACCTTACGGGGATCAAGCATGCGAGTCTCATTAGGACACCGTGTAGCGTTTGGCTGT  
TCTTGGA  
AACATTTTTTATTTATAACGTTTCTGAAAGTAGCGGTGGCACTTACTATCAAAAAGGTTACAAC  
TGCACCGA  
CAAACATATAACACTATCTTGTTCACCTTAACGGTGGTTCCTCGAGCGGTTCAAAGCACAACC  
ACCGTAAT  
GACACCCACGCTGGTTACAACTCCACATTCAGTGTGTCACCTGTTCCGTTGAGACTGACGACA  
AATTCCAG  
CGCGTTTGGACACGCTATTTATCAACGACAACAGCGTGTGAAAACGGGACGTTATCCAAGAAC  
ATAACTAA  
CTTGGCATTACCTATGGCAGCTGGGGCGTTGCGATGCTGCTGTTTGCCGCCGTGATGGTGCTC  
GTTGATTTG  
GGTTTGCCTCAATCGGCTTGGCGACGCTGGCGAAGCCACGTGGACGATGAAGAACGTGGTTTGT  
TAATGTAG  
GAAATAAAAGGCAGTTTGGAGCATGACTGTTTCCAAACCGTAACGTGGTAAATAAATCATGGCTT  
CCGACGTG  
GGTTCTCATCCTCTGACGGTTACACGATTCGCTGCAGAGTGCATTATGTGTACAATAAACTGT  
TGATTTTAA  
CTTTGTTTGCCCCCGTGATTCTGGAATCCGTCATCTACGTGTCCGGGCCACAGGGAGGGAACGT  
TACCCTGGT  
ATCCAACCTCACTTCAAACATCAGCGCACGGTGGTTCGCTGGGACGGCAACGATAGCCATCTC  
ATTTGCTT  
TTACAAACGTGGAGAGGGTCTTTCTACGCCCTATGTGGGTTTAAGCCTAAGTTGTGCGGCTAAC  
CAAATCAC  
CATCTTCAACCTCACGTTGAACGACTCCGGTCGTTACGGAGCAGAAGGTTTTACGAGAAGCGGC  
GAAAATG  
AAACGTTCCCTGTGGTATAATTTGACCGTGAAACCCAAACCTTTGGAACTACTCCAGCTAGTAA  
CGTAACAA  
CCATCGTCACGACGACATCGACGATGATCGACGCGAAAAGTAACGTTACAGGGAACGCCAGTTT  
AGCACCA  
CAATTACGTGCCGTCGCTGGATTCTCCAATCAGACGCCTTTGGAAAACAACACGCACCTGGCCT  
TGGTAGGT

GTTGTTGTGTTTTTAGTTCTGATAGTTGTTTGCATTATGGGGTGGTGGAAATTGTTGTGTGGTA  
AACCAGAGT  
TATAGTAATGTGCTTTTTATCAGGGAGAAGTTTTGTGCCAACCAATGACTAGCCCGGGACTATC  
TGCGTCAG  
AAAATTATGACGGAAATTATGAATTCACGGAAACCGCCAATACAACCGGTACAAATAGAAGTGA  
CTGGACA  
ACGTTAGAAAACAGTGCATTGCTATTGAAAAACACGGAGACTGCAGTGAACCTCAGCAACGCGA  
CTACGGT  
CATCCACAACCTGTAGAATACCCGGCTGGGGAAGTACAATATCAAAGAACGGCAACGCATTAT  
TCTTGGAT  
GCTAATCATTGTCAATCATTCTCATCATTTTTATTATCATCTGTCTACGAGCACCTCGAAAAATC  
TACCATCACT  
GGAAAGACAGTAAACAGTACGGACAAGTGTATGACAGACACGGAACTGTGACAGTGATGTCT  
AAGCGTT  
TGCAGGTATTTCCATGGATAACAATTTTTATTTTACACATCAAATCCCAGTATTGGAACATAT  
GGCAATACC  
ATGTACCCCTACAGTTGGATACGGCAGTCATAATATTAGCTTGCATCCGCTTAATAACTCATTA  
TTTCAAGAC  
GATGTTTTTGAATGGTACATAGACAAACCAATGGTTACAAGTTATGTCTTTATCAAAGTAATGA  
ACGCACAA  
AATCCAATCTAGACTCTCCAAATATTGTGTGGCAATGCACAGATAATCGTACACTAATTCTCAT  
GAACTTAA  
CCACAACATACAGTAGAACTATTATTTTCAATCCTTTAAATATCTCGGACGAGGAGTACAAA  
ACCGAATA  
ACTTGTGTTATAACGTTAGTGTACACTTTACCCACCAAACACATTGCCATACAACACTACATCAT  
CCTGTATCC  
ACCTACATCTGTACACGATTCATTAGAAATATCACAGTCATTCACCTCAACCAACTTCACACAT  
ACCGCGGT  
CCACTACGCCACCGGTAACGTTGAAGCACAAACACGACACTACCACTCCACATACAATGTGGATC  
ATACCCCT  
AGTTATCGTTATAACAATCATCGTTTTAACTTGTTTTCAAATCCCCCAGAAAGCTTGAATAAA  
TTCACACAA  
TACAGATACAGCGGTATGCTCGCCGCGCTTAAAGAATCAACGCCAAGGAAACCAAACGTAA  
AAGAATA  
GATATGTACGTTTTATTTTTCAGCTCACTGTTTGAATACCGTAAACATAATGACGTACATATACG  
TGGTTATAC



AACAGGTGTTTGTGTTATGCGGCGACTGATTAACCATATCGTGAACCATGATCTTTTCCGATGG  
TCCGTCTGTG  
ACCGCAATGATATTTTACAGATATTTCCGAAACCTGTATGGAGGTCACCTGTGAGAGTAGGTGATC  
CAGTTACC  
CTCGGTAGTGGACATGGTTATCATCCAGGTAGGGATAACAGGGTAATGATCCTCTAGAGTCGAC  
CTGCAGGC  
ATGCAAGCTTGAGTATTCTATAGTCTCACCTAAATAGCTTGG

SEQ ID NO: 11

ДНК

Род/вид - цитомегаловирус ЦМВЧ

Описательное название- пост-редактирование фрагмента ЦМВЧ

ACGACGGCCAGTGAATTGTAATACGACTCACTATAGGGCGAATTCGAGCTCGGTACCCGATTAC  
CCTGTTAT  
CCCTACCATTCCGGGCCGTGTGCTGGGTCCCCGAGGGGCGGGGGGTGTTTTTAGCGGGGGGT  
GAAATTTG  
GAGTCTTGGAGCCCGTGTGCTGTGGAGGACGGTGACGGTGGTAAGAGTGTGCTGCGGTGCGGT  
TGGGACG  
GCGGCGCGAATAAAAAGCGGCGTGCGGCGCGCACGGCGAAAAGCAGACGCGCTCTGTGTTGTG  
TGTCTTT  
GACCGGCGGAACACACGCGGAAAAGCGAGTCCCAGGGGACACACGACGAGCGAGTCCCAGGG  
GGGGAC  
GACGACGGCCAGGGACGCGGAAACGACGCGGAAAAGAGGAAGTCCCCAGGGGGACGGGCGGAAA  
AGAGG  
AAGCGCCTAGGGGACCGCGGGGGCAGGAACAGACGAAGTACGCCGCAACCCGCGTCGAGGACAC  
ACGCAG  
AAGCGGCCCGCCAGGGGAGGGGGGGGGGGGACTCGCGGGCCCCGGGGCACACTTGTGTTCCC  
TCCGGCC  
GCCGACACGCACCCCGAAGCCGCGCACACCGCCGACACACCCCTGACACACCCGCGACACACCC  
GCCACAC  
GCCCACACACGCCCCGCGACACACCCGACCGACACACCCCTGACACACCCCGCCAACACACCCAG  
CCGCACC  
CGCCCCGCAACACACCCCCGACACACCCGACACACGCCCGCGACACACCCGGCACACACCCAC  
CCACCCA  
GCCGCGCCCCGACACACCCCGAACGGCGCCGGTGCGGGACAGGGCTCACGGAGGTTTGCGGGC  
CGTGAGC  
ACGCCTCCCTTTGTACACACTACCGGTGCGTGGCGTCCCACGCTATTTGTTTCGCGAGACCGGAC

TAAGGGAG  
GTTTGCGGTGCCTCAGCGCGGGGCGGCGTTTTCGGCGTGTTCGACCAGCGCTTTGTGCGCGCT  
GCCTGTGC  
GTGTCGTCCCATGGTCTTTGTTCAGCGGCACGGCGCTGGGGACGGGGTTTCACCGCGCTGAGGGA  
TCTTTCTG  
CGGGTGTGAGGGACGGAGCTTTTTTCGCACGCTGGGCACCGGGCTGGGGACGGGGGTGTGCG  
GGACGGC  
GGTGGGGCCGGGGCGTTGCGGGTACGGGGATTACGCTGGGAACGGGGACTCGCGGACCCGGGCT  
GAGGGA  
CGGGGTGGCGGGGGTGTTCGCGCGAGGACGGGGGCCTTTTGCGGCGGGGACGGGGACTCACC  
CTCGCCT  
ATTTAACCTCCACCCACTTCAACACACACATGCCGCACAATCATGCCAGCCACAGACACAAACA  
GCACCCAC  
ACCACGCCGCTTCACCCAGAGTACCAACACACGTTACCCTTACACCACAGCAACACACAACCGC  
CTATCCAA  
ACCTCGGACAAACACGCCAACGAAGAACCCGCACGCAGATGGAGCTCGACGCCGGGATTACG  
CTGCTTG  
CGCGCAGGCCCGCCAACACCTCTACGCTCAAACACAACCCCAACTACACGCATACCCCAACGCC  
AACCCCTCA  
GGAAAGCGCTCATTTTTCCACAGAAAATCAACATCAACTCACGCATCTACTTCACAACATTGGC  
GAAGGCGC  
AGCGCTCGGCTACCCCGTCCCCCGCGCGAAATCCGCCGCGGCGGTGGCGACTGGGCCGACAGC  
GCGAGCG  
ACTTCGACGCCGACTGCTGGTGCATGTGGGGACGCTTCGGAACCATGGGCCGCCAACCTATCGT  
GACCTTAC  
TGTTGGCGCGCCAACGCGACGGCCTCGCTGACTGGAACGTCGTACGCTGCCGCGGCACAGGCTT  
TCGCGCAC  
ACGATTCGAGGACGGCGTCTCTGTCTGGCGTCAGCACTTGGTTTTTTTTACTCGGAGGCCACGG  
CCGCCGTGT  
ACAGTTAGAACGTCCATCCGCGGGAGAAGCCCAAGCTCGAGGCCTATTGCCACGCATCCGGATC  
ACCCCAT  
CTCCACATCTCCACGCCAAAACCACCCAGCCACCATATCCACCGCATCGCACCCACATGCT  
ACGACTCG  
CCCACATCACACGCTCTTTCCTATCCCTTCTACACCCCTCAGCCACGGTTCACAATCCCCGAAAC  
TACGCCGTC  
CAACTTCACGCCGAAACGACCCGCACATGGCGCTGGGCACGACGCGGTGAACGTGGCGCGTGGA

TGCCGGC  
 CGAGACATTTACATGTCCCAAGGATAAACGTCCCTGGTAGACGGGGTAGGGGATCTACCAGCC  
 CAGGGAT  
 CGCGTATTTGCGCGCCACGCTGCTTACCGATATCCAATAAACCCATCCCCTCGCCACGACGTC  
 TCCGCGTAT  
 CTTTGTAGCCTCAGGAATCCGTCCCCACGTCCATCCATCCCGAGCACTCCACACGCTATAACAG  
 ACCACGGA  
 CACGGCAAATGCATGCAAACCTTCTCATTTATTGTGTCTACTACTCTGTGTTGCTACAGGGAGTG  
 AAGGGGGT  
 GAAGGCAAAGAAAAAAAAAAGGAACAAAATAATAGATTAGCAGAAGGAATAATCCGTGCGACCG  
 AGCTTG  
 TGCTTCTTTTCTTATAAGGAGGCAAATATACTAGGAAAACCTAAGAATAGGAAGAAACCGAGG  
 TTTGGGA  
 GAAAAGCTGAGATAAAATAGCGCATTTTCCATACAGAGGTTGTTGTTTTTGTGGATCCTAAGAG  
 GTTTCAAG  
 TGCGAATCTCAAAGTTCTCACGAGAATATTGTCTTCAAGAATCGACAACCTGTGGTCCAAGATTT  
 TTTTTTGGT  
 CTTTTTAGGTTCTGCGAGGGACATCACGATGGATCGTTGCGATGAAGTCACGCGTACGCCTCTG  
 GTGTGGCG  
 CGGTGTCGTGACAGGAGAGTGTGTTTTTTCAGTGCAGAGCTGTCTTGATTCCCTATATCCGAGTATC  
 TGTTTTCTC  
 GTAAGGACGGTAATCTTCTTTGGTGTAAGTACATCTAAAAGCTGCAAACCTATATTTTAAGGGCT  
 GTCTCTAG  
 GTGTACTTTGATGCTGGAGTTTTTCCrCTGTGTTGATGTGAATAAATCTACTACTACTATTATA  
 TGCAGAAAGA  
 GTGATTATGCCGAGACAAGATTGCATTGGCTGAACTGTTTCAAAAACGCCTACACTCTACTTAT  
 CCGTAAAC  
 CTAAGGTAATACTATGTGTAAGTTGTTTTTTTTTCTTTTTGTAGTAAAATGGTGATACGTGCAA  
 TTAAAACCTG  
 TATTCCATGTTTCCATCCTTTTCATTTCAACTTTAAAGGCGGCTTTGAGAGCGAAGAAGTGCGAG  
 GATAAAAA  
 TGGATGACTCCTTCGTGTCCAGGGAGTCGACTACTGCAACGCTGATTGATTAAAAGATGGTCTC  
 CGATGATG  
 ATGTTGTTATTGATCGAATCATGGTGCAGAACGGCGACGGAGAGGAGCGTGTCCGCCCGGGGA  
 AGGTGGT  
 CTCTTTCTCTTTTCTTTTTTCAAGAAATCTTCCATGTGTTTATCGTAGTGATCGAAATCGACTG

ATCTCGGGTT  
 CTTTTTGGTTGGTTTCTTTTCGGTTAATCATGTATTGTTTTCTTTTTTTACAGAAAGATACTTTT  
 TTCATGAGCAA  
 TTCTCGCCCCGGCGCCGGCATGCCGAGGTGGGGCCACTGCGATCAGCGGCATGCCGACGCCGAC  
 CCGGGGA  
 TCTTGGATTACCGTTTTCTCTCTCTCTCTACATACAGACCGGGTGGCAGGAGCGGTAAGG  
 AATCATCGT  
 CGTCTTTCATTCTTCGATGATTATGGTAATACTAAATCTTATCTAGGAGCATATACATCTAAGA  
 TTGGAGTAC  
 TAGTAGTCGTTTGTGGTTTCTATTTTTTTTTATATTTATCTATGACAGTTTTTCTGTTTTTCGTT  
 TTGATAATAAT  
 ATAATAAAACTCATGGACCrTGAAATCTCrCrCTTGGTGTGGTGATTCATTCTCATTATTG  
 TTCrTTTTCTTTCC  
 GTCTTGC GGATGAAGATGTTGCGATGCGGTTGTTGTTGGTGTGCTATACACCGAGAGAGATGA  
 TCTTTTTGT  
 TCTTCTGGTTCATTTCTATGATTGTTTGGCTGCTGACCGACGCGTCAGGATGTGCAGGGCATG  
 CGGGGAATC  
 AGGACCGGACACGGGATAATTTCTACCTATACGGAGATCGCGGTCTCGCCATGAGGATCG  
 CGACAGG  
 CGCGTCGAGGGGGCAGGAACACCCTTGCGGATTGACATTCTTGGTGGTGTTCGTTGTTGTCGG  
 TAGTTGTTG  
 TTGACGATGAGGATAAATAAAAATGACCTTGTTTTTGTCTGTTTTCTCTTGTGGGAATCGTC  
 GACTTTGAA  
 TTCTTCGAGTTATCGGAAAGCTGAGGTACCCAAATGTCTGTAGCTTTTTTCTTTTTACCCTCTT  
 GTTTATCATC  
 TGCGATTTCGTGGTAGGTAGGAGAGGGAAATGATAATCCGAGATTAAGGAAAGGAGAAGATAAAA  
 AATAAA  
 AAAAAAATAAATAAACAGAAGCCGACCGCCGCCGACCCGTTCCCCAGGACCAGCCTACGAGGA  
 ATGGATA  
 ACGCGGTGGCGACGGCAGCGGTGGTGGCGCTGGGGTGGCGGCAGTGGTACTGCTGATGGTAGT  
 CGGGACG  
 GAGGAGAGGGCATGCATACATACACGCGTGCATGCTGCATGGGTGGATGGTACGGCCGGGAGAC  
 GCGGAA  
 GAGAACTCACATAAAAAGGTGACAAAAGAGCGTTGAAAAAGAAAACGAGATTCGACCAGA  
 CAGAAG  
 AGAAGGACCGGGGCTTGGCGACCCTTCCACGACTGCTGTTGTCATCTCGGCTCCCCCGTCTTCT

CCCGGCCAC  
GGGCGGCTAAGTCACCGCCGTTCTCCCCATCCGTCCGAGCGCCGACCGACCAGCCGGCCGATTCC  
GCCCCGCCG  
GGGCTTCTGGAGAACGCCGGGGCAGCAGCGATCTGGGGAAGCCGCTAAACCCCTGCGTTTTTTAT  
ATGGTAGC  
TCTGCCGAGCGCGGGCTGACGCGTTGAGTAAGCGGAAAGACGTGTGTGACGAAAAGGGGTCCCA  
TGGTATT  
TCACGTGACGATGAGGAGATGCGGTTTTGGAGCACATACGGTTTTAGAAAAAGGGAGTTGTCGTGA  
CAAGGGC  
TGAGGGACCTCTGTCTCCATGTGTGTATAAAAAGCAAGGCACGTTTCATAATGTAAAAAGAACA  
CGTTGTAA  
ACAAGCTATTGCTGTATCATTCCGGCTGACTATGCTTCATTCCGACTGATTTTTCTTTTCCTAACG  
GCGTAACTT  
AAAGTGATTAACGTATGATATTTGTTCCCCAGAGTTATACTATAGTCATCATCTAAAATTCAG  
ATATAAATG  
AACACATGTCGTATGGGATTATTAAGAAACCGAAACTCTCCACAGTTCACCATCTTCTTCGTCA  
TTCAACCG  
ATGACCCACTCCGTACAACGAATCAGTCTGCTGTGTCCACTGCAAACACTAGCGACGTATGC  
AAACAAC  
TGAAACACGGGCTGTTGTATTGACGACCGTTGTACCATTACTAGTCACATTGCATAGAGACCAT  
CCACCGTC  
ATCCCATCTTTCCCACCCGATGGAAAACCGTCTTCTATCATCAACTATGGTAAGATTTTCGACCC  
TGCGAGGTA  
TTCAGTTTCCCCATATCCATAACCTGGATTTTATCATTAAACCCCAATATTAACACTTTTTTTA  
GTACCCCCC  
ACCCACCAAAAAATGTGACTGGACCGGTTCCCTAGCAGCTCTGGGAGCCATGTTTCAGGTTGAACC  
ACAGCTAC  
AGCGAAACCGAGTCCAGTGACCGGTAACCACGTCCAGCCCCTGCGTATGTACCAGTCCAAGCAC  
GTCCGGTC  
ATTGTTCTACACAGGAAATCTAACTAGGTCAACGCAATTTTATTCCACCGTTACGCAGAATACT  
AACAAACA  
AACACACAAATTTAACGAATTACACGTAGTTTATTACATGAAAACCTGTAAGAACACCAATTCAC  
TAAGCGAT  
ACAACATTTAGCTGACTTCCAAGTGCCACACATCACCAGTGTATTTCATCCATGTTTTACCGAA  
CCAACGAG  
ACAGATCGAAGAAGCCAGAATCTCCCGACTTTAAATTACATAAATCCAACGTATTATGACCACA

GCTCGACA  
CACAAATAGTTGCGTTACTATTACAGTAGCATTACCTATACCCGTAACGTTGCACAACCACTG  
ATCACCATT  
GTTACCAAAAACGGTTTTTCCACTTAGTTGTCAACGGATCTTTCCCATGCGTAATGGTCAAATTA  
CTACCAGTC  
GTCGCTTTTAGCTCATTACGAGTATTATCCGCATCCACATATATCAACGTCATAGCTAGGCACG  
CTATAAGTA  
CCCCCCCCCACAATGGAATGTTGCCAAACCGTTCTTTCCCGTTATAGCCATAGCGTTCCCAG  
GCAAAAAGC  
AAACGCCAAACCTAATGCAGTGAAAAGCGCTTGCAGCCAGAACCAGCTTATGTACCAGCCACAA  
TCACATC  
CGGTTATTGTTTTCCACAGGAAATCCTACCAGGCAAAGCCCCGCTTGTTTTGTTCTGACCATCT  
TGTTTAGCA  
ATTCGTAAACTGTCAGCCTAGCGACGTCCGTTTAGATCAAAAGTCACGTATATAGCGACGCTGT  
TTCCACCC  
GTTTCCCCGTCCCGCCGTTTTCCGAACAACCCACCCGGGTTGAGACAACCGACCACCAACAGAAA  
TATACACA  
CAGACCACCGGGAGTTCAGTTAAAGATTTTCATCAGGTTTATTTTTGGCTGCTGCTAGTCTTTTGC  
TTCTTAGAA  
AAAAAATACCCATATAGAGAAATAATGATAGTTTGACAACACATATGGCAGGGATTTCTTCTTC  
ATCAATAA  
GATATGCAATTTCCCCAGGGAGAGACTTTCAACAATTGAATTTACAAAAACAAAATTACATCAG  
GAGAAAAG  
AGAGGATACATTAATAAATATATTATATCTGGTGTATATACTGAATGCTGCTGGTTCATAAGGT  
AACGATGC  
TACTTTTTTTAATTCCAAGATGGTTTTTCTTTGTTAGTCTTTTGTGACTTGCTGGTTCCTAAA  
AGTTCGCAAA  
AACGATTGTGTGAAGATTATGACGTTGGTTGACTAGTTCATGAGATTCTGCTGTACGTGTGATG  
GTTATTTCGC  
TGGTTCGTTCTAAGATGAGTATCGTACTGTGTCTGCGATGGTCGTCTCTTACTGGCATTCTCTC  
GGCTGCCTCT  
TGTTTTTCATGATTGAAAAGGAAAAAGGACTCCGAGGGCGCGTCATCTTTTACTTTTCGGTTT  
TCTCGTTGG  
CGGGTCAGAGGTAGTCAGATCATGAGACTGTGCTGGTCGATGAAACTGTGTCTGCTCAAGTGAC  
GTCCATTT  
CTTGACGGAGAAAAAAGTCATCGGGATAAATAAGGCTATACAAGGCGTTGTCAAGCGTGCGGC

TCTAAAC  
AAATTAAGCGATACAAAATTACAGTGATACGAATAATAAATTACCCCCTCCCCCTGTGGTCCCC  
CCGAGGCG  
AGAGCCACCCATCGTGTACTCTCGCACCACCCACGACCACAGGGGGAGACGGGACGAAGAGACG  
ACGCAG  
AGCGCCATCTCCTCCTGGAGGCCGGCGGGCTTAACTGCTACAGCTGCGGCGGGCAGCAGAGCTG  
CGATTTGT  
CGGCCGACATGCCGATGGTATGGGCGGCGGGCGGGTGGCCGCGGCAGCGGGGAGGAGAGAGA  
GAGAAG  
AGGAGCGGGGCGTCCGAAGGCGAGGATGGCATGGTCTCGCCGGAGCGCCCGGCTTTTATGGAAC  
ACTCGCG  
TCCGGTTGGGTATCACCCACAGGAAGATGAATCACAACCTCCAAACCATCTTGAGACCCGAGTA  
ACGGTTTA  
CAGGTGCGACGCCAGTCTCAGCTAAAAACAGCGGACAGTCCCACGCTGTTTCTGTTGTGGCTCT  
CTCCAGTTT  
CCTCATCGCCGTCTTGGTCTCCGTCTCATCGGAAGAATACCACCCGCTCTCATGCGGCAGTCG  
ATCAGCCTC  
GATGAACGAGACGCGGCGACGCCTTTCTACGGCCGACTGGTTGTGGTGGTGAAGAAGAGCACC  
AGCAATC  
CCAGGAGGAGCAACAAGCCCTCACATGTCCAGGAGGTGCGGGAGAGGGCCTGTCGGAGATGACC  
GTGAGG  
CATCACGTACGGCAGCTGAGGAGAAACGGAGAAGAAAGAAAATTACCGTCAGGGGCCGGGGTT  
CTTATTA  
GAGAAACAGCACGTAGGTCAGGATCCAGATGCTAATGGCAATCATGATGACGATGATCATGCAG  
GCCAAGA  
CGCGGCGCACCAATGCAGAATCCAATAGCCGCCGTGCCTCCGGTTGGTGGCCGGCGGCATCTAG  
AGACATG  
ATTTGGGGGGGGACCGGCGGCGCAAAAAGACAGGGAGATGGACAGTGCCACGGTGTTTTGTTA  
TGATTAG  
GACATGGGGACCGGAAGCCGAGACAGAGTACTACAGGGTGTGAAGGGTAACGTGAGGGAGATC  
ATGTCAT  
GGGCGGGCTGAAGACCGTGCAGGGAGGATCGACGTGTGCGGTGCTTGTGGAACACGGTGTTTTA  
ATATGTA  
TCCGCGTGAATGCACGCGGTGTGCTTTTTAGCACTCGGCTTGATAAGCTACGTGACCGTCTGC  
GCTGAAAC  
CATGGTCGCCACCAACTGTCTCGTGAAAACAGAAAATACCCACCTAGCATGTAAGTGCAATCCG

AATAGTAC  
ATCTACCAATGGCAGCAAGTGCCACGCGATGTGCAAATGCCGGGTACAGAACCCATTACCATG  
CTAGGCG  
CATACTCGGCCTGGGGCGGGGCTCGTTCGTGGCCACGCTGATAGTCCTGCTGGTGGTCTTCTT  
CGTAATTTA  
CGCGCGGAGGAGAGAAAAACAACACGGGCACCGAGGTAGATCAATGTCTGGCCTATCGGAGC  
CTGACAC  
GCAAAAAGCTGGAACAACACGCGGCTAAAAAGCAGAACATCTACGAACGGATTCCATACCGACC  
CTCCAGA  
CAGAAAGATAACTCCCCGTTGATCGAACCGACGGGCACAGACGACGAAGAGGACGAGGACGACG  
ACGTTT  
AACGAGGAAGACGAGAACGTGTTTTGCACCATGCAGACCTACAGCAACTCCCTCACGCTTGTC  
TAGTCACG  
TCGCTGTTTTTATTACAGCTCAGGGAAGTTTATCGAATGCCGTGCAACCAATCAAAAACCCC  
TAAAGCTC  
GCCAACTACCGCGCCACTTGCGAAAACCGTACACGCACGCTGGTTACCAGGCTTAACACTAGCC  
ATCACAGC  
GTAGTCTGGCAACGTTATGATATCTACAGCAGATACATGCGTCGTATGCCGCCACTTTGCATCA  
TTACAGAC  
GCCTATAAAGAAACCACGCGTCAGGGTGGCGCAACTTTCACGTGCACGCGCCAAAATCTCACGC  
TGTACAAT  
CTTACGGTTAAAGATACGGGAGTCTACCTTCTACAGGATCAGTATACCGGCGATGTCGAAGCTT  
TCTACCTC  
ATCATCCACCCACGCAGCTTCTGCCGAGCCTTGAAACGCGTCGATGCTTTTATCCGGGACCAG  
GCAGAGTC  
GGTGTGGTCACGGATTCCCAAGAGGCAGACCGAGCAATTATCTCGGATTTAAAACGCCAGTGGT  
CCGGCCTC  
TCACTCCATTGCGCCTGGGTTTTCGGGACTGATGATCTTTGTTGGCGCACTGGTCATCTGCTTTC  
TGCGATCGC  
AACGAATCGGAGAACAGGACGTTGAACATCTGCGGACGGACCTGGATACGGAACCTTTGTTGTT  
GACGGTG  
GACGGGAATTTGGAATAAAAGATGCGTAACACCTGTGCAAGATGCGATAACTTTACATACAGGC  
AAACAGT  
GTATACAATTATAGTATTTTGTATGTTGCATAAAGTTACATGCAACAGTACTGCTAACAGTACT  
GCATCCATT  
ACGCTATCCAACACTGCCTCTACCACTTTTGTAAACCATATATTCAACTCCGAATAACAACA



CATCAACG  
ACGCCACACACATCTGTACCTCACAAGCGTCAACCATTGGCAACATCACCAACGTTACCTCCG  
ACTTGAGT  
ACTTTCACAACCGTATATTCTACATTCAATACATCATTTGCCAATATATCTAATACGGCTGTCA  
CTACAGAAT  
TGATTTCAACAAATACCAACACTATCTCATCTTTTACCAACGTAACAGCAAACGCTACATCATC  
TTATAACAC  
AACAAATCACCGTAACTGTCACGTCAGATGAAACTTCGCACAACGTATCCACTAATAATGCACTT  
ATAAGCAC  
ACCATGGCCTACAAATTGCAGCGCCACAACATAACCCACGTACAACCTTACTAACTCTTCCAAC  
GCTTGTC  
CACAGAGACAACAATCATAACGTTTCAAGGAAACCAATACAACAGGAATAGAAGGGAGTAATGTC  
ACCATAA  
AGGGTAATTCTACGTGGGACTGTCTTTCAGTCGCCTGGATACGACATTACAATAGATCCACACA  
CGGACATC  
ATCTAGGTTATCGTAAGAACGCACATAACCAATCTTGGTATTGGCTACGCATCCTTACCTCTCA  
CACTGTATG  
TCATTCTCAACATGAAAGACCTTCACTGTACCATGACTTATGTCGTTTCGTGCAACAACACAGAA  
TTACATCTG  
TACGATCTAAATATCACCAATTCCGGCAGGTACAGCAGACGTTGTTTTAAAGAAAATTACTTCA  
CAGGACAT  
CACGAAGATGAAAATTTCTACCTATTAGTAACACCAAAAAATCATACTGAAGCTATTAATGCTA  
CTTTCGTT  
TGCCCTAGATACAACACCGATATCGAAAATGAAGATAGAGAGAAAGGAAGTCAACATACTAACA  
ATACACA  
TCACCACAAACGTAATCTCTATCATAGCTCGCAAAGAAGCCGCACCGTATGGACCATCGTGTTG  
GTTTGTAT  
GGCCTGCATAGTTCTGTTTTTTGCACGACGAGCCTTTAACAAAAAGTATCATATGTTACAAGAC  
ACCGTCAG  
TGAATCAGAATTCATTGTTTCGATATCACCCAGAACATGAAGATTGAGCTACGTTTCCGGGCAGA  
CATCTTAT  
GAAGCTGAACAATAAACTAAAACATTCTGTAAGACTCAGCGTTCAAAGGAATATTAATGCCCAT  
TGAGCGA  
AAACTAATATTGCAATGGACTGGCGATTTACGGTTACGTGGACGATACTAATGTCCGCGTTGTC  
AGAAAGCT  
GCAATCAAACCTGTTCTTGTCAATGTCCCTGTAGTACTACCGTTAACTATTCAACTAGTACTGA

GACAGCCAC  
ATCAACATACAGTACAACAGTTATCAGCAATAAAAGCACTTCAGAATCTATAAATTGCTCTACT  
GCAACTAC  
ACCAGCAAACACCGTTTCTACAAAACCGTCGGAAACAACCACACAGATATCCACAACGACGAAC  
ACAAACG  
TTGAGACTACCACATGTACCAACACCACCACGACCGTTACTTGTGATGGTTTCAATTATACAGT  
CCATAAAA  
GATGCGATCGCAGTTACGAGGTAATCAACGTAACAGGATACGTTGGTAGCAACATAACTCTAAA  
AAAATGC  
AATCAGACTGAGAAATGGCACAATGTAGACTGGATTCAATTATGAGTACCCACGCATAAAATGT  
GCGAATT  
AGGCAACTATCACCAAACCACACCACGGCAGCATATGTTTTGACTGCAACGACACCTCCCTA  
ACTATCTA  
CAACTTAACCACAAAAACGCTGGAAAATATACCAGGCGTCACCGTGATAACGGTCAAGAAGAA  
AATTACT  
ACGTAACGGTGTTAATTGGAGACACAACGTTATTCCTCTTGGCACATGCCCTGTAAGATATAA  
AGAATCTA  
CGAACACTGAAAACACCATTGGAAGTAGCATCATAGAAACCATTGAGAAAGCTAACATTCCCCT  
GGGAATT  
CATGCTGTATGGGCAGGCGTAGTGGTATCAGTGGCGCTTATAGCGTTGTACATGGGTAGCCATC  
GCATTCCC  
AAAAAGCCGCATTACACCAAACCTCCCAAATATGATCCAGATGAATTTTGGACTAAGGCTTAAC  
ATGCTGAT  
CAATAAACTTTTTTTAACCAATAACATGTCTCCGTTTTTTTTTTGTTAACAACCTATGATATAAA  
GCGTTATATT  
CAGTCGTTACTAAACAAAAAACATGGGCATGCAATGCAACACTAAATTGTTATTGCCAGTCGC  
ACTAATAC  
CGGTTGCAATCATCCTAATTGGTACTCTAGTGCCGATACTTTTACATGAACAAAAAAGGCGTT  
TTACTGGC  
GACTTTTTCTGCAAAGTCAACATGTAGAAGCACCCATTACAGTAACGCAGGGAGACACAGTCTA  
CCTAGACG  
CTAGCAATAATCCCTGTAATTATTCCAGCTTTTGGTACCACGGTAATTGCGAACTTTGTGGATG  
GAACGGAT  
ATCTACGCAATGTTACACATTACTACACAAACACATCGTGTCCCCGCAATTCATCTGCATAAA  
CGAAACTA  
AAGGTCTGCAGTTATATAATGTAACATTAACGATTTCAGGCGCTTATACTGAACACGTTTACGA

ATGTGACC  
TTTCGTGTAACATTACTACTAATAACGAATATGAAATACTCAATTATTTTGATAACTGTAAC  
TAACTA  
CACCATAAA  
TAGCACCAAGCATATTATCACCGTGGTGTCTTCACGTCATTCTAAACAAACAAATTTCCACGTA  
TCCACTCAC  
GCTGGTTGGGCAGTCGCCGTGGTGACGGTAATTATGATCTACGTTCTGATCCACTTTAACGTCC  
CGGCAACTC  
TGAGACACAAACTACGAACTAGAAACAACGTAAATCGCATAGCGTGATTATAAAGTATCGACGC  
TAATTTCT  
CCAAGATAAAAATTTGATTACTCCGTGCAGTTCTCAAAAACGTAAAGGCCCGCTTTTCCACTCC  
GTCATGAA  
GGATCGCAATAGAAATACTGCTATGTATCATCTTTATTTGCATTATGTGCCTCATTGTATTTAC  
TTTAAACGTC  
GTTGTGTTTTTACTCCGTCTCCAGACAAAGCAGATCTGCGAGTGGAAATTTCCCTCGTTACCCCC  
GTGTATTGG  
CATACAGTGCGCTGCATGAGAACACGCGTGACACATAGCGTACCCCTGGACGGTACAGTTTATG  
ATAACGTA  
ATTCAGGGAAAGTATACATTCATACCAACATGTTATCACATAACACACAGATTTTCTGCGTGTT  
TTATAAAA  
GAGCGTCTCGAAGCAGCTTGAGCCACACTACGGTCCAGATGACGAGCGTAATTA AAAATATGCC  
GCGCAGT  
ATTCGAAAAGCCGTACTGAGCGTGCGAGGCGGGTAGGGTGCCGAACGACGGATATGCGTCGTTGT  
CATCTTCG  
ACTATAAGGATCGCGACCGAGTCTTCGGCCATGGTAAACGTCACCCTGTGTGGCTGGTATGTAG  
CGTATCCG  
GTTTGGAAATGTTCTGCTCCAGCTCGGGGATAGTGAGGAATTC CAAGGGATACGGGACCCAA  
TGACTGGA  
TAAGAGAAGGGTTTTTCCCGTAAGATGATCCTCGTATCACATGAGGTCTGGATATGTATAAAT  
GAAGAGTG  
AAATAGGCACAGGGAATCAGATGCCAGCCTCGTGATGCAGCCGCTGGTTCTCTCGGCGAAGAAA  
CTGTCGT  
CTTTGCTGACTTGCAAATACATCCCGCCTTAAGCGATGAGTCTATAAAGCACCGTTGCCCGAGT  
ACGGTAAA  
AGTGACCCGGATTGTAGAACGTCCTTTTTTTTTTTGTTTTTGCATCGTTTATCGTCACTACTAGTG  
CAATATTTTG  
ATTGTAAGGCTGAAAGAGTATCGTTATGATGCTTAGAACGTGGAGATTATTACAGATGGTACTG

CTTGCCGC  
GTACTGTTATTATGTTTTTGC GACTTGTTCAATCAGCACGACGACTGCTCCTGTGGAATGGAAG  
TCTCCCGAC  
CGTCAGATTCCCAAGAATATTACCTGCGCTAATTACTCAGGGACCGTCAACGGCAACGTTACAT  
TTCGAGGT  
CTTCAGAACAAAACGGAAGACTTTTTGTACTGGTTGTTAGGATGGGGTCATAAGTCCATTTGTT  
CGTTCTTCC  
CGAAACTCCAGGGTAACTATGACGAACAACATTACAGATATGAAGTAGCGAACCTGACGTATAA  
CTGCACC  
TATAACCGCTTGACGTTGCTGAATCTGACGACGAAAACAGCGGAAAGTACTATTTCAAAGGG  
AAGATGC  
GAATTTACCTTCTATTACTCTTGTTACAACCTGACCGTGTCTAAAGATCGCACGTGAAGTTT  
CACAGAGCC  
GCGTGGCTGTAGCTATTGTGTTTACGTTGCTTTTTGAAATGTTAAGCGTCCCTACGGCGCTAACA  
TGTTTCTAG  
GCTACTCTGACTGTGTAGATCCCGGCCTTGCTGTGTATCGTGTATCTAGATCACGCTTAAAGCT  
CATGTTGTC  
TTTTGTGTGGTTGGTCGGTTTTGCGTTTTCTATGATTGTGCCGCTTCGAGTCCCTGCTGTTACGAC  
ATCACCGAG  
GCGGAGAGTAACAAGGCTATATCAAGGGACGAAGCAGCATTACCTCCAGCGTGAGCACCCGTA  
CACCGTC  
CCTGGCGATCGCGCCTCCTCCTGACCGATCGATGCTGTTGTGCGGAGAGGAAGAACTCGTTCCG  
TGGAGTCG  
TCTCATCATCACTAAGCAGTTCTACGGAGGCCTGATTTTCCACACCACCTGGGTCACCGGCTTC  
GTCTGCTA  
GGACTCTTGACGCTTTTTCGCCAGCCTGTTTCGCGTACCGCAATCCATCTGTGCTTTCTGCATAG  
ACCGTCTCC  
GGGACATCGCCCGTCTCTGAAATACCGCTATCAACGTCTTGTGCTACCGTGTAGCTAGTTAG  
CCAGCTGT  
GTGTAGTGTTTTGCTTTTGCATATTTGTTTTAGTCAGAGAGTCTGAAACGGGGTGGGAGGGAC  
TTTTGCGGG  
TAGTGCATGCTAAGATGAACGGGTGGGCTGGGGTGTGCTTGATAACTCACTGTTTGAATACGCG  
CTCACGCA  
CATATGTAGCACTCAACATGTTAGCTTTTTGCCCGCACGCCCCGGGGCGTGCCGAGCTGCCTTTT  
TAATAAAGT  
CTGGGTTTCCAGATACGCGCTGGTTCTGATTTTGATGGTTTGTGCCTCTGAAAGCTCTACGAGC

TGGGCCGTG  
ACATCCAATGGACTGCCTAACTGTAGCACGGTAACTAGAACAGCGGGTCAAGACGCTGAATTGC  
ACGGTCC  
GGCACCGTTAAGCTGTAATGTGACCCAGTGGGGACGTTACGAGAATGGAAGCACACCCGTGTTA  
TGGTGCA  
CTTTACGGGGATCAAGCATGCGAGTCTCATTAGGACACCGTGTAGCGTTTGGCTGTTCTTGAA  
AACATTTTT  
TATTTATAACGTTTCTGAAAGTAGCGGTGGCACTTACTATCAAAAAGGTTACAACGCACCGAC  
AACATAT  
AACACTATCTTGTTCACCTAACGGTGGTTCCTCGAGCGGTTCAAAGCACACCACCGTAATG  
ACACCCAC  
GCTGGTTACAACTCCACATTCAGTGTGTCACTTGTTCGGTTGAGACTGACGACAAATTCCAGC  
GCGTTTGG  
ACACGCTATTTATCAACGACAACAGCGTGTGAAAACGGGACGTTATCCAAGAACATAACTAAC  
TTGGCATT  
CACCTATGGCAGCTGGGGCGTTGCGATGCTGCTGTTGCCGCCGTGATGGTGCTCGTTGATTTG  
GGTTTGCCT  
CAATCGGCTTGGCGACGCTGGCGAAGCCACGTGGACGATGAAGAACGTGGTTTGTAAATGTAGG  
AAATAAA  
AGGCAGTTTGGAGCATGACTGTTTCCAAACCGTAACGTGGTAAATAAATCATGGCTTCCGACGTG  
GGTTCTCA  
TCCTCTGACGGTTACACGATTTGCTGCAGAGTGCATTATGTGTACAATAAACTGTTGATTTTA  
ACTTTGTTT  
GCCCCGTGATTCTGGAATCCGTCATCTACGTGTCCGGGCCACAGGGAGGGAACGTTACCCTGG  
TATCCAAC  
TTCACTTCAAACATCAGCGCACGGTGGTTCGCTGGGACGGCAACGATAGCCATCTCATTGCT  
TTTACAAA  
CGTGGAGAGGGTCTTTCTACGCCCTATGTGGGTTAAGCCTAAGTTGTGCGGCTAACCAAATCA  
CCATCTTC  
AACCTCACGTTGAACGACTCCGGTCGTTACGGAGCAGAAGGTTTTACGAGAAGCGGGCGAAAATG  
AAACGTT  
CCTGTGGTATAATTTGACCGTGAAACCCAAACCTTTGGAACTACTCCAGCTAGTAACGTAACA  
ACCATCGT  
CACGACGACATCGACGATGATCGACGCGAAAAGTAACGTTACAGGGAACGCCAGTTTAGCACCA  
CAATTAC  
GTGCCGTGCTGGATTCTCCAATCAGACGCCTTTGGAAAACAACACGCACCTGGCCTTGGTAGG

TGTTGTTG  
TGTTTTTAGTTCTGATAGTTGTTTGCATTATGGGGTGGTGGAAATTGTTGTGTGGTAAACCAGA  
GTTATAGTA  
ATGTGCTTTTTTATCAGGGAGAAGGTTTTGTGCCAACAATGACTAGCCCGGGACTATCTGCGTCA  
GAAAATTA  
TGACGGAAATTATGAATTCACGGAAACCGCCAATACAACGCGTACAAATAGAAGTGACTGGACA  
ACGTTAG  
AAACCAGTGCATTGCTATTGAAAAACACGGAGACTGCAGTGAACCTCAGCAACGCGACTACGGT  
CATCCCA  
CAACCTGTAGAATACCCGGCTGGGGAAGTACAATATCAAAGAACGGCAACGCATTATTCTTGGA  
TGCTAATC  
ATTGTCATCATTCTCATCATTTTTTATTATCATCTGTCTACGAGCACCTCGAAAAATCTACCATC  
ACTGGAAAG  
ACAGTAAACAGTACGGACAAGTGTATGACAGACACGGAACGTGACAGTGTCTAAGCGT  
TTGCAGG  
TATTTCCATGGATAACAATTTTTATTTTACACATCAAATCCCAGTATTGGAACTATATGGCAAT  
ACCATGTAC  
CCCTACAGTTGGATACGGCAGTCATAATATTAGCTTGCATCCGCTTAATAACTCATTATTTCAA  
GACGATGTT  
TTTGAATGGTACATAGACAAACCAATGGTTACAAGTTATGTCTTTATCAAAGTAATGAACGCAC  
AAAATCCA  
ATCTAGACTCTCCAAATATTGTGTGGCAATGCACAGATAATCGTACACTAATTCTCATGAACTT  
AACCACAA  
CATACAGTAGAACTATTATTTTCAATCCTTTAAATATCTCGGACGAGGAGTACCAAACCGAA  
TAACTTGT  
GTTATAACGTTAGTGTACACTTTACCCACCAAACACATTGCCATACAACCTACATCATCCCTGTA  
TCCACCTAC  
ATCTGTACACGATTCATTAGAAATATCACAGTCATTCACCTCAACCAACTTCACACATACCGG  
GTCCACTA  
CGCCACCGGTAACGTTGAAGCACAACACGACACTACCACTCCACATACAATGTGGATCATACCC  
CTAGTTAT  
CGTTATAACAATCATCGTTTTAACTTGTTTTCAAATTCCCCCAGAAAGCTTGAATAAATTCACA  
CAATACAG  
ATACAGCGGTATGCTCGCCGCCGCTTAAAGAATCAACGCCAAGGAAACCAAACGTAAAAAGAA  
TAGATAT  
GTACGTTTATTTTTTTCAGCTCACTGTTTGAATACCGTAAACATAATGACGTACATATACGTGGTT

ATACAACAG  
GTGTTTGTGTTATGCGGCGACTGATTAACCATATCGTGAACCATGATCTTTTCCGATGGTCCGT  
CGTGACCGC  
AATGATATTTTACAGATATTCGGAAACCTGTATGGAGGTCACGTGCAGAGTAGGTGATCCAGTT  
ACCCTCGG  
TAGTGGACATGGTTATCATCCAGGTAGGGATAACAGGGTAATGATCCTCTAGAGTCGACCTGCA  
GGCATGCA  
AGCTTGAGTATTCTATAGTCTCACCTAAATAGCTTGG

SEQ ID NO: 12

ДНК

Род/вид- Phikmvlikevirus LKA1

Описательное название - последовательность gp49 LKA1

ATGGCGCAAACACCCAGTACATGGGCGGACTACGTAGGCGACGGCGTAGAGGATACGTTCCAAG  
TCACAT  
TCCCGTACCAGAAGCAGCAAGAGGTGTTTGTGACTGTGGGCGGCGATCCGGCAGCTTTCACATT  
CATCTC  
GGCAGGTTGGATTCAACTGGCAGCGGTCCCGGTAATGGGGCCGCAATCCGTGTACGGCGCAGC  
ACTGAG  
GCATTCGAGCCTCGGCACGAGTTCCGCAACGGCGTGCCATTACTGCCGCGATTTCATAGACGAGA  
ATAATA  
CCCAGTTCTTGTACACTGTACAAGAGGCAGTGAATGAGACACATGGCATTGCTTCCGAAGCGCT  
GAGTGT  
CGCAGAGGAGGCCAGAGGCATTGCGCAGGCGGCATCGGATAAAGTGGATGCTGCCACCATTGAC  
TCCGCA  
CACCAGTTGCGTCTAGACCTCGCCGACCCGGCGAAGGGCCTGGGCTGCTAGGCTACGACCGAG  
ACGTAA  
GTTATCCGGTCCGGTCCGGTCAAAGCCTACAGTTTCTGGAAATGGGTCCGGTCCACACCAGC  
GCAATT  
TGGCGCCGTTGGTGATGGCGCCAGCCACCCCTCTCTGAGCGATACGCAACTCTAGCGGAAGCT  
CAGACT  
GTCTATCCGCATGCAGTCGCACTCTCCGACGAAATAGACTGGGCCGCATTGCAAGCTGCCGTGG  
ATTCAG  
GGGCACCTGTACACATACCGTCTGGGGACTATCAGATAAATAGGGGGATTAGCAGTACGGGCTC  
TCTACA  
GATTGCGGGTGATGGCGCTACATCTATTATACGCCCCGACTGCTGCGTTCCTGGTACATCGGTC  
CTCAGT

TGTGTGGGGAGCTTAGTTGCCTTGCCGAATATATCCTCCGTGTCGGCTGGGTCCCTAACCATG  
ACTTTG  
CCAGCACCCCTAATCTTGTAGCGGGGATGTATTCATCATCTACAACCCGACTGATAGCAGCTT  
CTCGGG  
ATTTTCGGACGAGCTATCGCGCAGGAGAGTTCTGTGAGGTCAGGGCGGTTTCTGGGAACACCGTG  
ACAATC  
CGTTCCGCACTCTATGCCGCATACGACGGGGCTACTGTTGCTATTTACAAAGTAGTCTCTGGTG  
TAGTTG  
ATATAGCTAGCATCCAAATCGTTGGCGGGACAGTCCCAATGAATGGACTGTTAGTGGAGGCTGT  
CGTTTC  
ACCGCGCGTFCGATGACGTGACGGTCACCCTTGCAAACAACGCCGGTGTGTATTTTGCCCGCTGC  
TATGAC  
GCTAAGATCACAAACAGTAATATATCGAACATCGGCGACGGTGGCGATGACTATGGAATCATCT  
TTGGGA  
ACTGTCACGACGGTGGGGCAGACAACGTAAAGTCTACGCTAGGCGACATGCCATCGCCACGGG  
CGGCGA  
TGCAGAAGTAGGCTGCGTTCCGGTCCGTAATGTGCGTATGCGTAACTGCACACTTAGGAATGAT  
ATTACC  
TCTGGTACACACTGCGCAGACTTCCACGGTAACGCCGAGGATTGCAGCTACGAAAACCTGCACAA  
TCTACG  
GTGGTGCAACTTGGCAGGGGAAGGATATCAGCTACAGACACTGTACAATCACTAACGCGTCGGG  
TGGTTG  
GATTGTTATATCCGCTGAGATTCTTGGTGGTACATTCTTCTCGACCAATGCACATTGTACACA  
ACCGGC  
GATCCGACGCTGGTAACCGTGGGGTTATAGATGTAGGTGGGAACTCCGCAGTCTCTACTACAA  
ATACAA  
CGCAACCCGTGAACTTCTTATAACAAGGCGGCAGTCTGCGAGCGCCAGCTTAAGTACGTCTAG  
TTACCT  
ACTGCGCGCACGTCTTGAGGGTAGTACAGTTCCAGTAAACATACAGTACAGCGGACAGGCTATT  
GATGTA  
GGCTCTCTGGGCAAGGTACTACAACCTCGATATTACCTCGGGCAGTACCTCTCTGAGTATTTGA  
TCGTGG  
AGAATTTAGCGGGGTTGCCATCTGGCATCACGCTGGCGTCTGCTGCTGGTGGTTTCGCAAGTGC  
CCCGAT  
GCGTATGCCTGTGCTGGGTGGTAGGGTTCAAGTAACTACGGCAACCAACGCGAGTAGCGTTACT  
GCTCCA



038595

GTAACGTT CAGGTACATTTATCCTAAGGCCCAACCGTCCAGGTCACAAAGACGGACAGGAGCT  
ACGCCG

GTAACAGGGT CGGCGTTGCTATCGCCAATCCGACCTCTGCGTCTGGGGCGACGTTGGGTCTGTT  
CACGGA

CGACGGGACAAACTTTAGCTCAGCCGTTACTAACCAGTTGAACTGGCAGGCAGGTATTTATGAG  
GTGTAA

SEQ ID NO: 13

ДНК

Род/вид- Phikmvlikevirus NTUH-K2044-K1-1

Описательное название- gp34 NTUH-K2044-K1-1

ATGGCCSTGATCCGGCTCGTGGCGCCCGAGCGCGTTCAGCGACCTGGCCAGCATGGTCGCCT  
ATCCGAAC

TTCCAGGTGCAGGACAAGATCACCCCTGCTGGGCTCGGCCGGCGGCGACTTCACCTTCACCACCA  
CCGCGTCG

GTGGTGGACAACGGCACCGTGTTCGCCGTGCCCGGGCTATCTCCTGCGGAAGTTCGTCCGCC  
CGGCGTAT

AGCTCGTGGTTCAGCAACTGGACCCGGGATCGTACGTTTCATGAGCGCGCCGAACCGGCACCTGG  
TGGTGGA

CACCGTGTGCAGGCCACGAGCGTGTGAACATCAAGAGCAACAGCACGCTGGAATTCACGGAC  
ACGGGCC

GCATCCTGCCCGACGCCGCCGTGGCCCGCCAGGTGCTGAACATCACCGGCTCCGCGCCCTCGGT  
GTTTCGTGC

CCCTCGCCGCCGACGCCGCCGGGGTGAAGGTGATCACCGTGGCCGCCGGCGCGCTGTCCGC  
GGTGAAA

GGCACCTACCTCTATCTGCGCTCCAACAAGCTGTGCGACGGCGGGCCGAACACCTATGGCGTCA  
AGATCAGC

CAAATCCGTAAGGTGGTTCGGCGTGAGCACCAGCGGGGGCGTGACGTCCATCCGCCTCGACAAAAG  
CCCTGCA

CTATAACTACTACCTCTCGGATGCCGCCGAAGTGGGCATCCCACCATGGTGGAGAACGTCACC  
CTGGTGAG

CCCGTACATCAACGAGTTCGGCTACGACGACCTGAACCGCTTCTTCACCAGCGGCATCTCCGCG  
AACTTCGC

GGCCGACCTGCACATCCAGGACGGCGTCATCATCGGCAACAAGCGTCCGGGGCGCCTCCGACATC  
GAGGGCC

GCAGCGCCATCAAGTTCAACAACCTGCGTGGATAGCACCGTGAAGGGCACCTGCTTCATAATAT  
CGGCTGGT

ACGGCGTGGAGGTCCTCGGCTGCTCGGAGGACACCGAGGTGCACGACATCCACGCCATGGACGT  
GCGCCAT  
GCCATCTCCCTGAACTGGCAAAGCACCGCCGACGGCGATAAGTGGGGCGAACCGATCGAGTTCC  
TGGGCGT  
GAACTGTGAGGCGTACAGCACCCAGGCCGGCTTCGACACCCACGACATCGGGAAGCGTGT  
AAATTCG  
TCCGCTGCGTGTCTACGACAGCGCGGATGACGGCTTCCAGGCCCGCACCAACGGCGTGGAGTA  
CCTCAACT  
GCCGCGCTACCGCGCCGCCATGGACGGCTTCGCCTCGAACACGGGCGTGCCTTCCCGATCTA  
CCGCGAAT  
GCCTGGCTACGACAACGTGCGCAGCGGGTTCAACTGCAGCTACGGCGGGGATGTGTACGA  
CTGCGAG  
GCGCACGGCAGCCAGAACGGCGTCCGCATCAACGGCGGCCGGGTCAAAGCGGGCGCTACACCC  
GCAACTC  
GTCGAGCCACATCTTCGTGACGAAAGATGTGGCGAAACCGCCAAACCAGCTCGAGATCGAC  
GGCGTCT  
CCATGCGGTACGACGGCACCGGCCGCGCCGTGTACTTCCACGGCACCGTGGGCATCGATCCGAC  
GCTCGTGA  
GCATGTCCAACAACGACATGACCGGCCACGGCCTGTTCTGGGCCCTGCTGTCCGGCTATAACCGT  
GCAGCCGA  
CCCCGCCGCGCATGTGCGGCAACCTGCTCGACGATACCGGCATCCGCGGGCGTCCGACCCCTGGT  
CGCGGGCG  
AAGCGACCGTCAATGCCCCGCTCCGCGGGAAC TTCGGCAGCGTGGCCAACAGCTTCAAGTGGGT  
GTCGGAG  
GTGAAGCTGACGCGCCTCACGTTCCCGTCGTCGGCCGGCGCCCTCACGGTCACCAGCGTCGCCC  
AAAACCAG  
GACGTGCCGACCCCCAACCCGGACCTGAACAGCTTCGTCATCCGCAGCAGCAACGCCCGCAGC  
TGTCCCA

AGTC GC CT GGG AGGT CT ACCTGT G A

SEQ ID NO: 14

ДНК

Род/вид - T7-подобный Pp15

Описательное название - добавлена последовательность gp44  
Pp15

ATGGCACGAACTATCGTCCAGAACGCCCTAACAGGGGACAAACAGGACTTCGAGGTACSTTTTCG  
ACTACATC

TTGCAGCGCTTCGTTAAGCTTACCCTGATCGGTGACGGTAACCGACAAGAGCTGGTCCTCGGTA  
CCGACTTC  
CGGTTTCATCGGTCTTCGCACCGTTCGCACTAACGTCTTCTGGGGACCAGCGCAGGGGTATACCT  
CCATCGAG  
ATCCGACGAGTTACCAGCGCTTCTGATCGTCGCGTAGAGTTCCTCGGACGGTCCATCCTGACCG  
CAGGTGAT  
CTGAACATCGCCAGCTTCAGGCCATCCACATTGCCGAAGAAGCGCGAGACTCTGCCACTGAGA  
ACCTGAG  
CCCAGATGCTGATGGCAACTACGATGCACGTGGTGC CGCATT TACAACCTCGGTGACGCTGTT  
CAGCCGAA  
GGATGCGGTCAACCGGTACACTCTTGACCTCGCTATCGCAGCCGCTCTGGCCATGAATACCGGC  
AACCCGAA  
CAACGCCCAGAACATCTCGTACACCCCTAACGGGCCTGGTCAGTCGATCCGAAGTGTGAAGGC  
CGTCTGCG  
GGATGCTGTGTTCTCGTCTCGGACTACATGACCACTCCACGTGATGGAGTTACCAGTAACCAGCAG  
GACCTCGA  
AAAGGCACTCGCTGCGGCGAACGCTAAAGGTGCCGACCTATTCTGGCCTGACGACATCCCGTTC  
TTCTCCAC  
GTCCCCGCTGGCACTGATCCACGCGGTCTACCATGTTGGACGTGGTGTCAACCGGAACGGT  
ACGCTGTT  
CTACGTGAACCCGAAGAACGGCCAACACAACAGGCTACACGTGTCTCCCGGGGGCACC GGGGAT  
GGTCTGG  
CAGCTGGCCGCCACTGGGGACCATCTGGAGTGCACCTCGCGGCCCTTAACATGCGAGCCCCACT  
GACCACGC  
GCTGGTCCCTGGAGATGACCGCTGGCGCCTATAATGAAGCCGTTACACTTCCGAACTACCTGAC  
CAGCTGTA  
ACGACTACTTGGCGTTTAACTGGCCGAACACCGGTGAGAACGTATGGAGCCCACTGCGTACCC  
ATCAGCTC  
TCGACGGCACAGGCCAGACCGGCCTCACAGGTTTCCACACTGGCATCGGCAACCGCATTACCAT  
CAACAAC  
GTGTGCATGTCCAACCTGGTACGACACTGCGCTGACTCCTACCCAACAGGTGCGAAGAGCGTTTCG  
TTGTAGGT  
GCGTATTGACTGCCTACGTGGTCAACTGCGCGTTCATTTACAACGGCATCGCGAGCGTGTCTG  
TGCTGCCC  
GGTGGCACTGCTATCGTAACCGGTGGCATCGTCGATGGTGGGCGGTTTCGGCCTCGACAACACTG  
GCGGTCGC

CTGTCCCTGACGGCAACCAAGAGCAATTATACGCAGGTCCGGAAGTGCCTCGAATATGGACTGT  
ACTCGAAG  
CATGACGCATCGACCGTAATGGACAACACCGAGTTCCGCAACTGCGGTAATCACCCCTGCGGCTG  
TTGCGTAT  
GGTGCTGCAATCTTCGCGTACAAGTTCAACTGTTCTGTTGACACTCGTGGGGTCAAGTTCTACG  
GCAACAAC  
ATCGCCAGCACTGCCGTGGCGGTATCACCTCGGACAATCCGGGCGATCCGGACATCTACGGTA  
CCGGCGCA  
GATGCTAATAAGCGTCTATTCTGTGCACCGGTGGTGGCTCTGACGACATCCAGTTCTACGAAG  
CTCGGCGC  
GTCATGGACATCACGAAGCGCACTGGTGGCGGCTCAACTACTGCCAGCGTATCGTCGCTGCTAC  
TGGCTGCC  
GTTGCGTCTGTCCGTAAGGGCTACTTTGCGCACAACGATCAGGTGATCCGGATGACCCTGATGT  
TCCGCGCT  
ACAGGCTCGGCTGGCATCTTCACGCCGACCTTGCGCACACCTCTGGGGACTATCCCTCTGGGTA  
GCTTCAGG  
GTCGCATCGGGACAGTACGGCGAGATCAAGTTGACCATTCGACCTACTCTGACATCTGATGGTC  
TCATAGTC  
GGGTTCTCCTGCATCAACGCCGTGCAGAATCTTGGGTCTCTGTTGGTCAAATCATCGTCAGCG  
GCACCGTA  
GACCTCCGCACCGTCGACCAGCTGGTTCGAGATGTGGGGCTATTTCGGAAGCTGGTGGCACCCTT  
CGTACATT  
CAAGGCCT GATCGAGCT GGTCGGGT GA

SEQ ID NO: 15

ДНК

Род/вид - *Aggregatibacter actinomycetemcomitans*

Описательное название - добавлена последовательность dspB

ATGAACTGTTGCGTCAAGGGCAATTCCATCTACCCCCAGAAGACCTCCACCAAGCAGACCGGCC  
TGATGCTC  
GATATCGCCCGGCATTTCTACAGCCCCGAGGTGATCAAGAGCTTCATCGATACGATCAGCCTGA  
GCGGCGGC  
AACTTCCTCCACCTGCACTTCTCGACCATGAAAACCTATGCCATCGAGTCGCACCTGCTCAACC  
AGCGGGCG  
GAGAACGCCGTCCAGGGGAAGGATGGCATCTACATCAATCCGTACACCGGAAACCGTTCTCTGA  
GCTACCG  
CCAGCTGGACGACATCAAGGCCTACGCCAAGGCCAAGGGCATCGAACTGATCCCGGAGCTGGAC

AGCCCGA  
 ACCATATGACGGCCATCTTCAAACCTGGTCCAGAAGGACCGCGCGTCAAGTACCTGCAGGGGCT  
 GAAATCC  
 CGCCAGGTGGACGACGAGATCGACATCACCAACGCCGATAGCATCACCTTCATGCAGAGCCTGA  
 TGAGCGA  
 GGTCATCGATATCTTCGGCGACACGAGCCAGCACTTCCACATCGGCGGCGACGAATTTCGGCTAC  
 TCCGTCGA  
 GAGCAACCACGAGTTCATCACCTACGCCAACAAAGCTGTCTACTTCTGGAGAAGAAGGGGCTC  
 AAGACCC  
 GCATGTGGAACGACGGCCTCATCAAGAACACCTTCGAGCAGATCAATCCCAACATCGAAATCAC  
 GTACTGG  
 TCGTACGACGGCGACACCCAGGATAAGAACGAAGCGGCCGAGCGCCGCGACATGCGCGTGAGCC  
 TGCCGGA  
 GCTGCTGGCGAAGGGCTTCACCGTGTGAACTACAACAGCTACTACCTCTACATCGTGCCGAAG  
 GCGAGCCC  
 GACGTTCTCGCAGGACGCCGCTTCGCCGCCAAAGACGTGATCAAGAAGTGGGATCTGGGCGTC  
 TGGGATG  
 GCCGGAACACCAAGAACCGCGTGCAGAACACCCATGAGATCGCCGGGGCGGCGTGTCTGATCTG  
 GGGCGAG  
 GATGCGAAGGGCTCAAGGACGAGACGATCCAGAAGAACACCAAAAGCCTGCTCGAGGCCGTCA  
 TCCACAA  
 GACCAACGGCGACGAGTGA

SEQ ID NO: 16

ДНК

Род/вид- *Staphylococcus aureus*

Описательное название- добавлена последовательность SaPSMa3  
 ATGGAGTTCGTGGCGAAGCTCTTCAAGTCTTCAAGGACCTGCTCGGGAAGTTCCTGGGGAATA  
 ACTGA

SEQ ID NO: 17

ДНК

Род/вид- *Staphylococcus aureus*

Описательное название- добавлена последовательность SaPAMb2  
 ATGACCGCCTGGCCGAGGCGATCGCGAATACCGTCCAGGCGGCCAGCAGCACGACAGCGTCA  
 AGCTGGG  
 CACCTCGATCGTGGACATCGTTCGCCAACGGCGTGGGCCTGCTGGGCAAACCTCTTCGGCTTCTGA

SEQ ID NO: 18

ДНК

Род/вид- Staphylococcus epidermidis

Описательное название - добавлена последовательность SePSMa  
 ATGGCGGACGTCATCGCCAAGATCGTCGAGATCGTGAAGGGCCTGATCGACCAGTTCACCCAGA  
 AGTGA

SEQ ID NO: 19

ДНК

Род/вид- Levivirus MS2

Описательное название- добавлена последовательность L MS2  
 ATGGAGACCCGGTTCCCGCAGCAGTCCCAGCAAACCCCGGCCAGCACCAACCGCCGCCGCCCT  
 TCAAGCA  
 CGAGGACTACCCGTGCCGCCGGCAGCAGCGCAGCTCCACCCTGTACGTGCTGATCTTCTGGCG  
 ATCTTCT  
 GAGCAAGTTCACCAACCAGCTGCTGCTGTCCCTGCTGGAGGCGGTCATCCGGACCGTCAACCACC  
 CTGCAGCA  
 GCTGCTGACCTGA

SEQ ID NO:20

ДНК

Род/вид- Levivirus PRR1

Описательное название- добавлена последовательность L PRR1  
 ATGTGCAAGGTGTСТАСТААGGTAGACTСТАААСТGACTGAGTCAGTTGGACAАСТCACCАТАА  
 GGAGCTAC  
 STATGGCTACGGAATATCСТАGCAТTAGCAGGACTTCTTTTCGTAATCCTTCTTGGACCAATC  
 АТТТАТССА  
 TCGСТАТСТАCAGTCCGTAA

SEQ ID NO:21

ДНК

Род/вид- Phikmvlikevirus LUZ19

Описательное название - промотор gp32 (P<sub>32</sub>) LGZ19  
 CGACCSTGCCСТАТCCGGCCTTAAACCCACATCCAAAAGAGAGAGAATCGC

SEQ ID NO:22

ДНК

Род/вид- Phikmvlikevirus LUZ19

Описательное название - терминатор gp32 (T<sub>32</sub>) LGZ19  
 TGCCACGAAACCCCGCACTTCGGTGTGGGGTTTCTTCAAAGCCTAACGACCCGCGCAGATTCCC  
 TCGGTGGG

TTTTTGCCTTTAGGAGAAACCT

SEQ ID NO:23

ДНК

Род/вид- Phikmvlikevirus LUZ19

Описательное название - область gp7 LUZ19 дикого типа

TACAAGGTGGTGGCACCCAGCTCGGCGGAAGGTATCATTTGTGCTGGCGACCAAGCAGACGCCGG  
CGCTAGC

CCAAGCAGCCGTCGTA CTGCACAGCATGAACCTGCGCAGTATCCCGCAGGTTCCGGCTATCCTC  
AACACGGC

CTGGAAGTGCCGCCGCCTGGGAGTGGGCGAGTACGTCAAGCTCGTCCAAGGGGAGGAGGAC

SEQ ID NO:24

ДНК

Род/вид- Phikmvlikevirus LUZ19

Описательное название- область gp18 LUZ19 дикого типа

GAATGCCAACC GAAGAAGACGCATGATCCGCTGTTTACTGGCGGATATCCACGAGCCACTGGA  
CCTGCTGT

TCCCCGGCCTCCGTACCAAGGCCATATGGACCCGCAAGCAGAGGAACTGTGATTCGAATTGA  
CTACGACC

ATGCGAAGCTGGGCCGTATGGGATTCTGCCACGCGGTATCCCTATATCAACTGTCCATATATGG  
CCGCGAGG

GGATGGTCCGCTACCTGATGCAGGAGATCCCCCGCGTGCTGGAAGGTCTGCTGGTCAAGGC  
GCAGCAGT

ACAGCCAAAGCAACTGGTACAGCAAATGACGAC

SEQ ID NO:25

ДНК

Род/вид- Phikmvlikevirus LUZ19

Описательное название - gp49 и межгенная область gp48-gp49  
LUZ19 дикого типа

GGGGACACCATGAGCAAAGCCAAACTACGAGTCATCGCCGACACCCCGGAGCTGGAGTCAGTGC  
TAAAAGC

ATTGCTGACCGCCACCTACGCTATCGAGGACCTGCTCAACGAGGCCGTGGCTAGCAAGGTGCTA  
AACTCCCG

CCTGGGCTGGTCCGCAGTCGGCGAGTATGTCGAACTGTTCAACCGCACGCAATCCCGCGTGGCC  
GGTTGAT

TCCCGAGTAG

SEQ ID NO:26

038595

ДНК

Род/вид- Phikmvlikevirus LKD16

Описательное название - ген gp18 LKD16 дикого типа

GTGCGAGTACCCAACGTAACACGAGCGCACCCCTGCGCTGCCTGCTCCAAGACATCCACGGGCCGC  
TGAATCTG  
CTGTTCCCAGGTATCCGGGTGAAGGTGGAGGAGGCGTGCCTCGGATACTTGGGCTACAGGGAGC  
GGGGCTA  
TTGGGAGCTGCGCCTCCAGGTGGACTACGACCACCCGAAGCTTGGGCACCTCCGCTACAGTCAG  
GCCGTGCC  
GGAGTACGTGCTGATCAACGACCGCGACAGCATCATCAAGTACCTGATGGAAGCAGTCCCTCGG  
CAGGTAC  
TAGAGGGCATGCTCAATAAGGCCCAGGAATTCGTAACCAAGAACTGGTATTCCSTATGA

SEQ ID NO:27

ДНК

Синтетический (искусственный/неизвестный)

Описательное название - ген, кодирующий NLS-FLAG-CAS9-His

ATGCCCAAGAAAAAGCGGAAGGTCGGCGACTACAAGGATGACGATGACAAGTTGGAGCCTGGAG  
AGAAGC  
CCTACAAATGCCCTGAGTGCGGAAAGAGCTTCAGCCAATCTGGAGCCTTGACCCGGCATCAACG  
AACGCAT  
ACACGAGACAAGAAGTACTCCATCGGGCTGGACATCGGGACGAACTCCGTGGGATGGGCCGTGA  
TCACAGA  
CGAATACAAGGTGCCTTCCAAGAAGTTCAAGGTGCTGGGGAACACGGACAGACACTCCATCAAG  
AGAACC  
TCATCGGGGCCTTGCTCTTCGACTCCGGAGAAACCGCCGAAGCAACGCGATTGAAAAGAACCGC  
CAGAAGA  
CGATACACACGACGGAAGAACCGCATCTGCTACCTCCAGGAGATCTTCAGCAACGAGATGGCCA  
AGGTGGA  
CGACTCGTCTTTTCATCGCCTGGAGGAGAGCTTCCTGGTGGAGGAAGACAAGAAACATGAGCGC  
CACCCGA  
TCTTCGGGAACATCGTGGACGAAGTGGCCTACCACGAGAAATACCCACGATCTACCACTTGGC  
CAAGAAA  
CTCGTGGACTCCACGGACAAAGCGGACTTGC GGTTGATCTACTTGGCCTTGGCCCACATGATCA  
AATTTCCG  
GGCCACTTCTGATCGAGGGCGACTTGAATCCCGACAATTCGACGTGGACAAGCTCTTCATCC  
AGCTGGTG



CAGACCTACAACCAGCTCTTCGAGGAGAACCCCATCAATGCCTCCGGAGTGGACGCCAAAGCCA  
TCTTGTCCT  
GCCCCGATTGTCCAAATCCAGACGCTTGGAGAACTTGATCGCACAACTTCCTGGCGAGAAGAAGA  
ACGGCCT  
CTTCGGCAACTTGATCGCGCTGTCGCTGGGATTGACGCCTAACTTCAAGTCCAACCTTCGACTTG  
GCCGAGGA  
CGCCAAGTTGCAACTGTCCAAGGACACCTACGACGACGACCTCGACAACCTGCTGGCCCAAATT  
GGCGACC  
AATACGCGGACTTGTTTTTGGCGGCCAAGAAGTTGAGCGACGCCATCTTGTTGAGCGACATCTT  
GCGCGTGA  
ATACGGAGATCACCAAAGCCCCTTTGTCCGCTCTATGATCAAGCGGTACGACGAGCACCACCA  
AGACTTGA  
CCCTGTTGAAAGCCCTCGTGCGGCAACAATTGCCCGAGAAGTACAAGGAGATCTTCTTCGACCA  
GTCCAAGA  
ACGGGTACGCCGGCTACATCGACGGAGGAGCCTCCCAAGAAGAGTTCTACAAGTTCATCAAGCC  
CATCCTG  
GAGAAGATGGACGGCACCGAGGAGTTGCTCGTGAAGCTGAACCGCGAAGACTTGTTGCGAAAAC  
AGCGGAC  
GTTTCGACAATGGCAGCATCCCCACCAAATCCATTTGGGAGAGTTGCACGCCATCTTGCGACGG  
CAAGAGG  
ACTTCTACCCGTTCTGAAGGACAACCGCGAGAAAATCGAGAAGATCCTGACGTTTCAGAATCCC  
CTACTACG  
TGGGACCCCTTGGCCCGAGGCAATTCCCGTTTTGCATGGATGACGCGCAAAGCGAAGAGACGAT  
CACCCCC  
TGGAACCTCGAAGAAGTGGTCGACAAAGGAGCATCCGCACAGAGCTTCATCGAGCGAATGACGA  
ACTTCGA  
CAAGAACCTGCCCAACGAGAAGGTGTTGCCAAGCATTGCTGCTGTACGAGTACTTCACGGTG  
TACAACGA  
GCTGACCAAGGTGAAGTACGTGACCGAGGGCATGCGCAAACCCGCGTTCCTGTGGGAGAGCAA  
AAGAAGG  
CCATTGTGGACCTGCTGTTCAAGACCAACCGGAAGGTGACCGTCAAACAGCTGAAAGAGGACTA  
CTTCAAG  
AAGATCGAGTGCTTCGACTCCGTGGAGATCTCCGGCGTGGAGGACCGATTCAATGCCTCCTTGG  
GAACCTAC  
CATGACCTCCTGAAGATCATCAAGGACAAGGACTTCTTGACAACGAGGAGAACGAGGACATCC  
TGGAGGA

CATCGTGCTGACCCTGACCCTGTTTCGAGGACCGAGAGATGATCGAGGAACGGTTGAAAACGTAC  
GCCCACTT  
GTTTCGACGACAAGGTGATGAAGCAGCTGAAACGCCGCCCTACACCGGATGGGGACGATTGAGC  
CGCAAAC  
TGATTAATGGAATTCGCGACAAGCAATCCGGAAAGACCATCCTGGACTTCTGAAGTCCGACGG  
GTTTCGCCA  
ACCGCAACTTCATGCAGCTCATCCACGACGACTCCTTGACCTTCAAGGAGGACATCCAGAAGGC  
CCAAGTGT  
CCGGACAAGGAGACTCCTTGACGAGCACATCGCCAATTTGGCCGGATCCCCGCAATCAAAA  
AGGCATC  
TTGCAAACCGTGAAAGTGGTTCGACGAACTGGTGAAGGTGATGGGACGGCACAAGCCCGAGAACA  
TCGTGAT  
CGAAATGGCCCGGAGAACCAAACCACCCAAAAAGGACAGAAGAACTCCCGAGAGCGCATGAAG  
CGGATC  
GAAGAGGGCATCAAGGAGTTGGGCTCCAGATCCTGAAGGAGCATCCCGTGGAGAATACCCAAT  
TGCAAAA  
CGAGAAGCTCTACCTCTACTACCTCCAGAACGGGCGGGACATGTACGTGACCAAGAGCTGGAC  
ATCAACC  
GCCTCTCCGACTACGATGTGGATCATATTGTGCCCCAGAGCTTCTCAAGGACGACAGCATCGA  
CAACAAGG  
TCCTGACGCGCAGCGACAAGAACCAGGGCAAGTCTGACAATGTGCCTTCCGAAGAAGTCGTGAA  
GAAGATG  
AAGAACTACTGGCGGCAGCTGCTCAACGCCAAGCTCATCACCCAACGGAAGTTCGACAACCTGA  
CCAAGGC  
CGAGAGAGGAGGATTGTCCGAGTTGGACAAAGCCGGCTTCATTAAACGCCAACTCGTGGAGACC  
CGCCAGA  
TCACGAAGCACGTGGCCCAAATCTTGGACTCCCGGATGAACACGAAATACGACGAGAATGACAA  
GCTGATC  
CGCGAGGTGAAGGTGATCACGCTGAAGTCCAAGCTGGTGAAGGACTTCCGGAAGGACTTCCAGT  
TCTACAA  
GGTGCAGGAGATCAACAACCTACCATCACGCCCATGACGCCTACCTGAACGCCGTGGTCCGGAACC  
GCCCTGA  
TCAAGAAATACCCCAAGCTGGAGTCCGAATTCGTGTACGGAGATTACAAGGTCTACGACGTGCG  
GAAGATG  
ATCGCGAAGTCCGAGCAGGAGATCGGCAAAGCCACCGCCAAGTACTTCTTTTACTCCAACATCA  
TGAACTTC

038595

TTCAAGACCGAGATCACGCTCGCCAACGGCGAGATCCGCAAGCGCCCCCTGATCGAGACCAACG  
GCGAGAC  
GGGAGAGATTGTGTGGGACAAAGGAAGAGATTTTGCACAGTGCGCAAGGTGCTGTCCATGCCT  
CAGGTGA  
ACATCGTGAAGAAGACCGAGGTGCAAACAGGAGGGTTTTCCAAAGAGTCCATTTTGCCTAAGAG  
GAATTCC  
GACAAGCTCATCGCCCGCAAGAAGGACTGGGACCCCAAGAAGTACGGGGGCTTCGACTCCCCCA  
CGGTGGC  
CTACTCCGTGTTGGTGGTGGCCAAAGTGGAGAAAGGGAAGAGCAAGAAGCTGAAATCCGTGAAG  
GAGTTGC  
TCGGAATCACGATCATGGAACGATCGTCGTTTCGAGAAAAACCCCATCGACTTCCTCGAAGCCAA  
AGGGTAC  
AAAGAGGTGAAGAAGGACCTGATCATCAAGCTGCCCAAGTACTCCCTGTTTCGAGCTGGAGAACG  
GCCGCAA  
GCGGATGCTGGCCTCCGCCGGGGAAGTGCAGAAAGGGAACGAATTGGCCTTGCCCTCCAAATAC  
GTGAACT  
TCCTCTACTTGGCCTCCCATTACGAAAAGCTCAAAGGATCCCTGAGGACAATGAGCAGAAGCA  
ACTCTTCG  
TGGAACAACACAAGCACTACCTGGACGAGATCATCGAGCAGATCAGCGAGTTCTCCAAGCGCGT  
GATCCTC  
GCCGACGCCAACCTGGACAAGGTGCTCTCCGCCTACAACAAGCACCCGCGACAAGCCTATCCGCG  
AGCAAGC  
CGAGAATATCATTCACCTGTTTACCCTGACGAATTTGGGAGCCCCTGCCGCCTTTAAATACTTT  
GACACCACC  
ATCGACCGCAAAAGATACACCTCCACCAAGGAAGTCTTGGACGCCACCCCTCATCCACCAGTCCA  
TCACGGGC  
CTCTACGAGACGCGCATCGACCTCTCCCAATTGGGCGGCGACCATCATCACCACCACCACTAA

SEQ ID NO:28

ДНК

Род/вид - Inovirus M13MP18

Описательное название - заменена область M13MP18 дикого  
типа

ATGACCATGATTACGAATTCGAGCTCGGTACCCGGGGATCCTCTAGAGTCGACCTGCAGGCATG  
CAAGCTTG  
GCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCCTGGCGTTACCCAACTTAATCGCC  
TTGCAGCA

038595

CATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGT  
TGCGCAGC  
CTGAATGGCGAATGGCGCTTTGCCTGGTTTTCCGGCACCAGAAGCGGTGCCGAAAGCTGGCTGG  
AGTGCGA  
TCTTCCTGAGGCCGATACGGTCGTCTGCCCTCAAACCTGGCAGATGCACGGTTACGATGCGCCC  
ATCTACAC  
CAACGTAACCTATCCCATTACGGTCAATCCGCCGTTTGTTCACGGAGAATCCGACGGGTGTG  
TACTCGCTC  
ACATTTAATGTTGATGAAAGCTGGCTACAGGAAGGCCAGACGCGAATTATTTTTGATGGCGTTC  
CTATTGGT  
TAA

SEQ ID NO:29

ДНК

Неизвестный/искусственный - коммерчески доступный из DNA2.0

Описательное название - последовательность паприка

ATGGTGTCAAAGGGAGAAGAACTGATCAAAGAGAATATGAGGATGAACTCTACATGGAAGGAA  
CTGTGA  
ACAACCACCATTTCAAGTGCACGAGCGAGGGTGAAGGGAAACCTTACGAAGGTACCCAGACCAT  
GCGGATT  
AAGGTCGTGCAAGGAGGACCACTCCCCTTCGCATTTCGACATCCTGGCCACTTCCTTCATGTACG  
GGTCGCGC  
ACTTTCATCAAGTACCCAAAAGGGATCCCCGACTTCTTCAAGCAGTCCTTCCGGAGGGATTCA  
CTTGGGAA  
CGCGTCACTAGATACGAGGATGGCGGAGTGGTCACCGTGATGCAAGACACCTCTTTGGAAGATG  
GATGCCT  
GGTGTAACACGTGCAAGTCAGAGGAGTGAACCTTCCGAGCAATGGGCCGGTGATGCAGAAGAAA  
ACCAAGG  
GCTGGGAACCGAACACCGAAATGCTGTATCCAGCAGACGGAGGCTTGGAGGGCCGGTCCGACAT  
GGCTCTG  
AAGCTTGTGGAGGAGGACATCTGTCCTGCTCGTTCGTGACGACCTACCGGAGCAAGAAGCCGG  
CGAAAAA  
CCTTAAGATGCCGGGGATCCACGCGGTGGATCATCGCCTGGAAAGGCTCGAGGAGTCAGACAAC  
GAGATGT  
TTGTCGTGCAACCGGAGCACGCCGTGGCCCCTACTGTGATCTCCCTTCAAAGCTGGGCCACAA  
GCTGAATT  
CCGGCCTCCGGTCGAGAGCCCAGGCTTCGAATTCAGCCGTGGACGGAACCTGCGGGCCCTGGTTC

GACCGGA

AGCCGATGA

SEQ ID NO:30

ДНК

Род/вид- *Lambdavirus* лямбда

Описательное название - последовательность *ell* фага X  
дикого типа *E. coli*

ATGGTTCGTGCAAACAACGCAACGAGGCTCTACGAATCGAGAGTGCCTTGCTTAACAAAATCG  
CAATGCTT

GGAACAGAGAAGACAGCGGAAGCTGTGGGCGTTGATAAGTCGCAGATCAGCAGGTGGAAGAGGG  
ACTGGA

TTCCAAAGTTCTCAATGCTGCTTGCTGTTCTTGAATGGGGGTCGTTGACGACGACATGGCTCG  
ATTGGCGC

GACAAGTTGCTGCGATTCTACCAATAAAAAACGCCCGGCGGCAACCGAGCGTTCTGAACAAAT  
CCAGATG

GAGTTCTGA

SEQ ID NO:31

Белок

Синтетический (искусственный/неизвестный)

Описательное название - белок NLS-FLAG-CAS9-His,  
транслируемый из SEQ ID NO: 27

MPKKRRKVG DYKDDDDKLEPGEKPYKCECGKSFQSGALTRHQ RTHTRDKKYSIGLDIGTNSV  
GWAVITDEYK

VPSKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEM  
AKVDDSFHRLE

ESFLVEEDKKHERHP IFGNIVDEVAYHEKYPTIYHLRKKLVDS TDKADLR LIYLALAHMIKFRG  
HFLIEGDLNPDN

SDVDKLF IQLVQTYNQLFEENPINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLI  
ALSLGLTPNFKSN

FDLAEDAKLQLSKD TYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSA  
SMIKRYDEHHQ

DLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKNL  
REDLLRKQRTFD

NGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPIYVGPLARGNSRFAMTRKSE  
ETITPWNFEVVD

KGASAQSFIERMTNFDKNLPNEKVLPKHSLLEYEFTVYNELTKVKYVTEGMRKPAFLSGEQKKA

IVDLLFKTNR  
 KVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKI IKDKDFLDNEENEDILEDIVL  
 TLTLEFEDREMIEE  
 RLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTI LDFLKSDFANRNFMLIH  
 DDSLTFKEDIQ  
 KAQVSGQGDSLHEHIANLAGSPAIKKILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKG  
 QKNSRERMKR  
 IEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFL  
 KDDSIDNKVLTR  
 SDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGELSELDKAGFIKRQLVE  
 TRQITKHVA  
 QILDSRMNTKYDENDKLIREVKVI TLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTA  
 LIKKYPKLESE  
 FVYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGE  
 IVWDKGRDFATV  
 RKVLSMPQVNI VKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVA  
 KVEKGKSKKL  
 KSVKELLGITIMERSSEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQ  
 KGNELALPSKYV  
 NFLYLASHYEKLGKSPEDNEQKQLFVEQHKHYLDEIEQISEFSKRVI LADANLDKVL SAYNKH  
 RDKPIREQAENII  
 HLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVL DATLIHQSI TGLYETRIDLSQLGGDHHHHHH

SEQ ID NO: 32

ДНК

Род/вид - цитомегаловирус ЦМВЧ

Описательное название - фрагмент RL13 ЦМВЧ после редактирования

ATGGACTGGCGATTTACGGTTACGTGGACGATACTAATGTCCGCGTTGTCAGAAAAGCTGCAATC  
 AAACCTGT  
 TCTTGTC AATGTCCCTGTAGTACTACCGTTAACTATTC AACTAGTACTGAGACAGCCACATCAA  
 CATACAGTA  
 CAACAGTTATCAGCAATAAAAAGCACTTCAGAATCTATAAATTGCTCTACTGCAACTACACCAGC  
 AAACACCG  
 TTTCTACAAAACCGTCGGAAACAACCACACAGATATCCACAACGACGAACACAAAACGTTGAGAC  
 TACCACA  
 TGTACCAACACCACCACGACCGTTACTTGTGATGGTTTCAATTATACAGTCCATAAAAAGATGCG

ATCGCAGT  
TACGAGGTAATCAACGTAACAGGATACGTTGGTAGCAACATAACTCTAAAAAATGCAATCAGA  
CTGAGAA  
ATGGCACAATGTAGACTGGATTCAATTATGAGTACCCACGCATAAAATGTGCGAATTAGGCAAC  
TATCACCA  
AACCACACCACGGCAGACATATGTTTTGACTGCAACGACACCTCCCTAACTATCTACAACCTTA  
ACCACAAA  
AAACGCTGAAAAATATACCAGGCGTCACCGTGATAACGGTCAAGAAGAAAAATTA CTACGTAACG  
GTGTAA  
TTGGAGACACAACGTTATTCACTCTTGGCACATGCCCTGTAAGATATAAAGAATCTACGAACAC  
TGAAAAACA  
CCATTGGAAGTAGCATCATAGAAACCATTGAGAAAGCTAACATTCCCCTGGGAATTCATGCTGT  
ATGGGCAG  
GCGTAGTGGTATCAGTGGCGCTTATAGCGTTGTACATGGGTAGCCATCGCATTCCCAAAAAGCC  
GCATTACA  
CCAAACTCCCAAATATGATCCAGATGAATTTTGGACTAAGGCTTAA

SEQ ID NO:33

ДНК

Род/вид - цитомегаловирус ЦМВЧ

Описательное название - предварительное редактирование  
фрагмента RL13 ЦМВЧ

ATGGACTGGCGATTTACGGTTACGTGGACCGTТАCTTGTGATGGTTTCAATTATACAGTCCATA  
AAAGATGC  
GATCGCAGTTACGAGGTAATCAACGTAACAGGATACGTTGGTAGCAACATAACTCTAAAAAAT  
GCAATCA  
GACTGAGAAAATGGCACAATGTAGACTGGATTCAATTATGAGTACCCACGCATAAAATGTGCGAA  
TTAGGCA  
ACTATCACCAAACCACACCACGGCAGACATATGTTTTGACTGCAACGACACCTCCCTAACTAT  
CTACAACCT  
TAACCACAAAAAACGCTGAAAAATATACCAGGCGTCACCGTGATAACGGTCAAGAAGAAAAATTA  
CTACGTA  
ACGGTGTТАATTGGAGACACAACGTTATTCACTCTTGGCACATGCCCTGTAAGATATAAAGAAT  
CTACGAAC  
ACTGAAAACACCATTGGAAGTAGCATCATAGAAACCATTGAGAAAGCTAACATTCCCCTGGGAA  
TTCATGCT  
GTATGGGCAGGCGTAGTGGTATCAGTGGCGCTTATAGCGTTGTACATGGGTAGCCATCGCATTС

ССААААААГ

ССGCATTACACCAAААСТТСССAAATATGATCCAGATGAATTTTGGACTAAGGCTTAA

SEQ ID NO:34

Белок

Род/вид- Phikmvlikevirus LUZ19

Описательное название - последовательность белка Gp13 LUZ19  
дикого типа

MLALGAFDLSGLMVGSCLVVGGELKALCVDDRHSRQIGIAELVRAAELAGAEYLTCFEFLEPFY  
ADLGWSTTH

REANWTAGEPDVLMRAPGHDV

SEQ ID NO:35

Белок

Род/вид- PMkmvlikevirus LUZ19

Описательное название - последовательность белка Gp38 LUZ19  
дикого типа

MARFKNPEТИHVADGVEAVFSLDFPFLRREDV FVQVDKIILVTDYTWVDDTNIQLAVVPKKDQEV  
RIFRDTPAQVP

DTQFSQDIPFLPRYIDANNKQLLYAVQEGINTANLALDGVLDAIRIAEEARRLAQEALDAANEA  
LRRALGFAEIRT

VTEDSDIDPSWRGYWNRCITADKPLTLTQMEDPDAPWVEFSEVHFEQAGVRDLNIVAGPGVTI  
NRLQNTTMQL

YGENGVCTLKRLGANHVIWGAMEDE

SEQ ID NO:36

Белок

Род/вид- PMkmvlikevirus LUZ19

Описательное название - последовательность белка Gp40 LUZ19  
дикого типа

MFKTEVKGRYTLIRRKADGTPVETLEFDNIIITNAGLDWIAAMDТDLMGEPVAVSTSTADPNPSA  
PAIPEVVQRTS

ASAPGGGTTSGLDGEWLFWRRRRFRFPQGTLAGQVLATVGLICNSDRRFESNTGELIPKDTPLSY  
TRIKDAAGQPT

TLVVAADEILDVQYEFRRSRPVGTAЕAKFVISGVERTFRLIPKPFANRANLSGERYIFYNTNPYI  
NGKDASGGNVRD

GQWQKKYPKYVRGSYKAQITLLAQVQNGNMAGGITGTEELQIYNGRNYVLDINPPVVKNNТQEF  
TVTLEFTVAR

A



038595

SEQ ID NO:37

Белок

Род/вид- *Pseudomonas aeruginosa*

Описательное название - последовательность белка PycS5

MSNDNEVPGSMVIVAQGPDDQYAYEVPPIDSAAVAGNMFGDLIQREIYLQKNIYYPVRSIFEQG  
TKEKKEINKKV  
SDQVDGLLKQITQGKREATRQERVDVMSAVLHKMESDLEGYKKTFTKGPFFIDYEKQSSLSIYEA  
WVKIWEKNSW  
EERKKYPFQQLVRDELERAVAYYKQDSLSEAVKVLRLQELNKQKALKEKEDLSQLERDYRTRKAN  
LEMKVQSEL  
DQAGSALPPLVSPPTPEQWLERATRLVTQAIADKKQLQTTNNTLIKNSPTPLEKQKAIYNGELLV  
DEIASLQARLVK  
LNAETTRRRTEAERKAAEEQALQDAIKFTADFYKEVTEKFGARTSEMARQLAEGARGKNIRSSA  
EAIKSFEEKHD  
ALNKKLSLKDQAIKAFDSLDKQMMAKSLEKFSKGFVVGKAI DAASLYQEFKISTETGDWKP  
FFVKIETLAA  
GAAASWLVGIAFATATATPIGILGFALVMAVTGAMIDEDLLEKANNLVISI

SEQ ID NO:38

Белок

Род/вид- PMkmvlikevirus LKD16

Описательное название - последовательность белка Gp18 LKD16

MRVPTHEHERTLRCLLQDIHGPLNLLFPGIRVKVEEACLGYLGYRERGYWELRLQVDYDHPKLGH  
LRYSQAVPEY  
VLINDRDSIIKYLMEAVPRQVLEGM LNKAQEFVTKNWYSL

SEQ ID NO:39

Белок

Род/вид- PMkmvlikevirus LKA1

Описательное название - последовательность белка Gp49 LKA1

MAQTPSTWADYVGDGVEDTFQVTFPPYQKQEVFVTVGGDPAAFTFISAGWIQLAAVPVNGAAIR  
VRRSTEAFEP  
RHEFANGVPLLPRFIDENNTQFLYTVQEAVNETHGIASEALSVAEEARGIAQAASDKVDAATID  
SAHQRLRLDLAD  
PAKGPGLLGYDRDVSYPVGSVQSLQFLEMGRVTPAQFGAVGDGASHPLSERYATLAEAQTVYP  
HAVALSDEID  
WAALQAAVDSGAPVHIPSGDYQINRGISSTGSLQIAGDGATSIIRPTAAFTGTSLSCVGSLSVA  
LPNISSVSAGSLTI

038595

DFASTPNLVAGDVFI IYNPTDSSFSGFRTSYRAGEFCEVRAVSGNTVTIRSALYAAAYDGATVAI  
YKVVSGVVDIASI  
QIVGGTVPMNGLLVEAVVSPRVDDVTVTLANNAGVYFARCYDAKITNSNISNIGDGGDDYGIIF  
GNCHDGGADN  
CKVYARRHAIATGGDAEVGCVVVRNVRMRNCTLRNDITSGTHCADFHGNAEDCSYENCTIYGGA  
TWQKDISY  
RHCTITNASGGWIVISAEILGGTFLLDQCTLYTTGDPQPGNRGVIDVGGNSAVLTTNTTQPCNF  
LIQGGSLRAPSL  
TSSYLLRARLEGSTVPVNIQYSGQAIDVGS LGKVLQLDITSGSTSPEYLIVENLAGLPSGITLA  
SAAGGFASAPMRM  
PVLGGRVQVTTATNASSVTAPVTFRYIYPKAPTQVTKTDRSYAGNRVGVAIANPTSASGATLG  
LFTDDGTNFSS  
AVTNQLNWQAGIYEV

SEQ ID NO:40

Белок

Род/вид- PMkmvlikevirus NTUH-K2044-K1-1

Описательное название - последовательность белка Gp34 NTUH-  
K2044-K1-1

MALIRLVAPERVFSDLASMVAYPNFQVQDKITLLGSAGGDFTFTTTASVVDNGTVFAVPGGYLL  
RKFBVGPAYSS  
WFSNWTGIVTFMSAPNRHLVVDTVLQATSVLNIKSNSTLEFTDTGRILPDAAVARQVLNITGSA  
PSWVPLAADA  
AAGSKVITVAAGALSAVKGTLYLRSNKLCDGGPNTYGVKISQIRKVVGVSTSGGVTSIRLDKA  
LHYNYYLSDA  
AEVGIPTMVENVTLVSPYINEFGYDDLNRFFTSGISANFAADLHIQDGVIIIGNKRPASDIEGR  
SAIKFNNCVDSTV  
KGTCFYNIGWYGVEVLGCSEDETVHDIHAMDVRHAISLNWQSTADGDKWGEPIEFLGVNCEAYS  
TTQAGFDTH  
DIGKRVKRVRCVSYDSADDGFQARTNGVEYLNCRAYRAAMDGFASNTGVAFPIYRECLAYDNVR  
SGFNCSYGG  
GYVYDCEAHGSQNGVRINGGRVKGGRYTRNSSSHIFVTKDVAETAQTSLEIDGVSMRYDGTGRA  
VYFHGTVGID  
PTLVSMNNDMTGHGLFWALLSGYTVQPTPPRMSRNLDDTGIRGVATLVAGEATVNARVRGNF  
GSVANSFKW  
VSEVKLTRLTFPSSAGALTVTSVAQNQDVPTPNPDLNSFVIRSSNAADVSQVAWEVYL

SEQ ID NO:41

038595

Белок

Род/вид - T7-подобный Pp15

Описательное название - последовательность белка Gp44 Pp15

MARTIVQNALTTGGQQDFEVPFDYILQRFVKLTLIGDGNRQELVLGTDFRFIGPRTVTRTNVFWGP  
AQQYTSIEIRRV  
TSASDRRVEFSDGSILTAGDLNIAQLQAIHIAEEARDSATENLSPDADGNYDARGARIYNLGD  
VQPKDAVNRYT  
LDLAIATAALAMNTGNPNNAQNI SYTPNGPGQS IRSVEGRLRDAV FVSDYMTTPRDGVT SNQQDL  
EKALAAANAK  
GADLFWPDDIPFFSTSPALIHAVYHVGRGVINANGTLFYVNPKNQHNRLHVSPGGTGDGLAA  
GRPLGTIWSAL  
AALNMRAPLTTTRWSLEMTAGAYNEAVTLPNYLTSNDYLAFNWPNTGQERMEPTAYPSALDGTG  
QTGLTGFHT  
GIGNRITINNVCM SNWYDTALTPTQQVRRAFVVGAYSTAYVNC AFIYNGIASVSVLPGGTAIV  
TGGIVDGGFRFG  
LDNTGGRLSLTATKSNYTQVRNCLEYG LYSKHDASTVMDNTEFRNCGNH PAAVAYGAAI FAYKF  
NCSVDTRGV  
KFGYGNIAQHCRGGITSDNPGDPDIYGTGADANKRLF LCTGGGSDDIQFYEARVMDITKRTGG  
GSTTASVSSLL  
LAAVASVRKGYFAHNDQVIRMTLMFRATGSAGIF TPTLRTPLGTIPLGSFRVASGQYGEIKLTI  
RPTLTSDGLIVGF  
SCINAVQNLGSSVGOIIVSGTVDLRTVDQLVEMWGYSEAGGTASYIQGLIELVG

SEQ ID NO:42

Белок

Род/вид - *Aggregatibacter actinomycetemcomitans*

Описательное название - последовательность белка DspB

MNCCVKGNSTYPQKTSTKQTGI.MI,DTARHFYSPF.VTKSFTDTTSLSGGNFI.HI,HFSDHF  
.NYATF.SHI,1,NQRAF.NAV  
QKGDGIYINPYTGKPF LSYRQLDDIKAYAKAGIELIPELDSPNHMTAIFKLVQKDRGVKYLQ  
LKSQRQVDDEIDI  
TNADSITFMQSLMSEVIDIFGDTSQHFHIGGDEF GYSVESNHEFIT YANKLSYFLEKKGLKTRM  
WNDGLIKNTFEQI  
NPNI EITYWSYDGDTQDKNEAAERRDMRVSLPELLAKGFTVLNYSYLYIVPKASPTFSQDAA  
FAAKDVIKNW  
DLGVWDGRNTKNRVQNTHEIAGAALS IWGEDAKALKDETIQKNTKSLLEAVIHKTNNGDE

SEQ ID NO:43

Белок

Род/вид- Staphylococcus aureus

Описательное название - последовательность белка SaPSMa3

MEFVAKLFFKFFKDLLGKFLGNN

SEQ ID NO:44

Белок

Род/вид- Staphylococcus aureus

Описательное название - последовательность белка SaPAMb2

MTGLAEAIANTVQAAQQHDSVKLGTSIVDIVANGVLLGKLFGE

SEQ ID NO:45

Белок

Род/вид- Staphylococcus epidermidis

Описательное название - последовательность белка SEPSMa

MADVIAKIVEIVKGLIDQFTQK

SEQ ID NO:46

Белок

Род/вид- Levivirus MS2

Описательное название - последовательность белка L MS2

METRFPPQSSQOTPASTNRRRPFKHEDYPCRRQQRSTLYVLI FLAIFLSKFTNQ LLLSLLEAVI  
RTVTTLQQLLT

SEQ ID NO:47

Белок

Род/вид- Levivirus PRR1

Описательное название - последовательность белка L PRR1

MCKVSTKVDSKLTESVGQLTIRSYLWLRN LALAGLLFVILLATNHL SIAIYSP

SEQ ID NO:48

Белок

Род/вид- Phikmvlikevirus LUZ19

Описательное название - последовательность белка Gp18 LUZ19

MRMPTEEERMIRCLLADIHEPLDLLFPGLRTKAHMDPQAEELSIRIDYDHAKLGRMGFCHAVSL  
YQLSIYGREGM

VRYLMQEIPRRVLEGLLVKAQQYSQSNWYSK

SEQ ID NO:49

Белок

Род/вид- Phikmvlikevirus LUZ19

Описательное название - последовательность белка Gp49 LUZ19

MSKAKLRVIADTPELESVLKALLTATYAIEDLLNEAVASKVLNSRLGWSAVGEYVELFNRTQSR  
VAGLIPE

SEQ ID NO:50

ДНК

Род/вид- Phikmvlikevirus LUZ19

Описательное название - последовательность гена gp18 LUZ19

ATGAGAATGCCAACCGAAGAAGAACGCATGATCCGCTGTTTACTGGCGGATATCCACGAGCCAC  
TGGACCT  
GCTGTTCCCCGGCCTCCGTACCAAGGCCATATGGACCCGCAAGCAGAGGAACTGTCGATTCGA  
ATTGACTA  
CGACCATGCGAAGCTGGGCCGTATGGGATTCTGCCACGCGGTATCCCTATATCAACTGTCCATA  
TATGGCCG  
CGAGGGGATGGTCCGCTACCTGATGCAGGAGATCCCCGCCGCTGCTGGAAGGTCTGCTGGTC  
AAGGCGC  
AGCAGTACAGCCAAAGCAACTGGTACAGCAAATGA

SEQ ID NO:51

ДНК

Род/вид- Phikmvlikevirus LUZ19

Описательное название - последовательность белка Gp49 LUZ19

ATGAGCAAAGCCAAACTACGAGTCATCGCCGACACCCCGGAGCTGGAGTCAGTGCTAAAAGCAT  
TGCTGAC  
CGCCACCTACGCTATCGAGGACCTGCTCAACGAGGCCGTGGCTAGCAAGGTGCTAAACTCCCGC  
CTGGGCTG  
GTCCGCAGTCGGCGAGTATGTCGAACTGTTCAACCGCACGCAATCCCGCGTGGCCGGGTTGATT  
CCCGAGTA

G

## Перечень последовательностей

- <110> SYNTHETIC GENOMICS, INC.  
DIPETRILLO, Christen  
CADY, Kyle C.  
BARBU, Elana M.
- <120> КОМПОЗИЦИИ И СПОСОБЫ ДЛЯ КОНСТРУИРОВАНИЯ ВИРУСНОГО ГЕНОМА IN VITRO
- <130> SGI1840-3W0
- <150> US 62/092,707  
<151> 2014-12-16
- <150> US 62/102,362  
<151> 2015-01-15
- <150> US 62/242,811  
<151> 2015-10-16
- <160> 51
- <170> PatentIn версии 3.5
- <210> 1  
<211> 288  
<212> ДНК  
<213> Phikmvlikevirus LUZ19
- <220>  
<221> misc\_feature  
<223> gp13 LUZ19 дикого типа
- <400> 1  
gtgctggccc tcggtgcctt cgacctgtcc ggcctgatgg taggttcctg cctcgtagta 60  
ggtggtgagc tgaaggccct gtgcgtgat gaccggcaca gcaggcaggg tatcggcgct 120  
gagctggtac gggccgctga gctggctggt gccgagtatc tgacctgctt cgagttcctg 180  
gagccgttct acgccgactt gggctggagc accaccacc gcgaggcгаа ctggacagca 240  
ggagagccgg acgtgctgca catgagggca cccggatcatg acgtatga 288
- <210> 2  
<211> 756  
<212> ДНК  
<213> Phikmvlikevirus LUZ19
- <220>

## 038595

<221> misc\_feature  
 <223> gp38 LUZ19 дикого типа  
  
 <400> 2

gtggctcggg tcaagaatcc cgagaccatc cacgttcgag atggggtcga ggctgtcttc	60
agtctcgact tcccgttcct gcggcgtgag gacgtattcg tccaggtcga taagatactc	120
gtcaccgact atacgtgggt agacgacacc aacatccaat tggccgtggt gccgaagaag	180
gaccaagagg tccgcatctt ccgcgacacg cccgcccagg tcccggacac acagttcagc	240
caggacatcc cgttcctgcc tcgatacatc gacgcgaaca acaagcagct cctgtacgct	300
gtgcaggaag gcatcaacac cgcgaaacct gctctcgatg gcgtactcga cgcgatccgt	360
atgccgagg aggctcgtcg cctggcgcag gaagcactcg acgccccaa tgaaggcctt	420
cgccgtgccc tgggcttcgc tgagattcgc accgtgaccg aggactcgga catcgatccg	480
agctggcgcg gttactggaa ccgttgcatc accgccgata aacctctgac cctgaccatg	540
cataggaag accccgatgc accgtgggtc gagttcagcg aggttcactt cgagcaggcc	600
ggtgtgcgtg acctaaacat cgtagccggt cctggcgtta ccatcaaccg tttgcagaac	660
accaccatgc agctctacgg cgagaatggc gtgtgtactc tcaagcggct gggcgctaac	720
cactggatcg tgttcggggc catggaggac gaataa	756

<210> 3  
 <211> 906  
 <212> ДНК  
 <213> Phikmvlikevirus LUZ19

<220>  
 <221> misc\_feature  
 <223> gp40 LUZ19 дикого типа  
  
 <400> 3

atgtttaaga ccgaagtaaa gggacgttac accctgattc gccgcaaggc ggacggcact	60
ccggtggaga ctctggagtt cgacaacatc attacgaatg cgggcctgga ttggatcgcc	120
gctatggata ccgacctcat gggcgaaccg gtagcggatc gcacttctac agccgatccc	180
aaccgagcg caccgccat cccggagggt gtgcaacgca cgtccgcatc tgcccctggt	240
ggaggtaacta cgtcgggcct ggatggcgag tggctgttct ggccggaggcg ttggagattc	300

## 038595

ccgcagggca ccctagctgg tcaagtcctg gccaccgtgg gcctcatctg caactcggat 360  
 cgtcgccttcg agagtaacac gggtagctg atcccgagg ataccgct gtcgtacct 420  
 cgcatcaagg acgccgccgg gcagcctact actctggagg tggccgctga cgagattctg 480  
 gatgtccagt acgagttccg cagccggccc gtaggaacgg ctgaggccaa gttcgtgatc 540  
 tccggcgtgg aacgcacctt ccggctgatc ccaaagcctt ttgcaaccg tgctaattctc 600  
 tccggggaac gctacatctt ctacaacacc aaccctaca tcaacggcaa ggacgcctcc 660  
 ggcggaatg tccgagacgg tcagtggcag aagaaatatt ccaagtacgt gcgcggctcc 720  
 tacaaggcgc agatcacgct gctggcccag gtccagaacg gcaatatggc tggcggcctc 780  
 accggcaccg aggaactcca gattacaat ggacgtaact atgtgctcga tatcaaccg 840  
 cctgtgtga agaacaatac ccaggagttc accgtgacc tggagtttac ggtggcgagg 900  
 gcataa 906

<210> 4  
 <211> 2481  
 <212> ДНК  
 <213> Phikmvlikevirus LUZ19

<220>  
 <221> misc\_feature  
 <223> gp34 LUZ19 дикого типа

<400> 4  
 atgagctaca agcaatccgc gtatcccaat ctgctgatgg gtgtgagcca gcagggtccc 60  
 ttcgagcgcc tgccgggcca gctcagcgag cagatcaaca tggatccga tcccgtgtca 120  
 ggacttcggc ggcgagcgg tatcagctg atggcccacc tgctgcatac cgaccagccc 180  
 tggccgaggc cgttcctcta ccacacgaac ctcggggcc gcagcattgc gatgctggtg 240  
 gcgcagcacc gtggcgagct gtacctgtt gacgagcggg acggtcgcct gctgatgggt 300  
 cagcccctgg tgcatgacta cctcaaggcc aacgattaca ggagctacg ggccgccacg 360  
 gtggccgatg acctgttcat cgccaacctg agtgtaaagc ccgaggccga ccgcaccgac 420  
 atcaagggcg tagaccccaa caagccggc tggctgtaca tcaaggcagg ccagtattcg 480  
 aaggcattct ccatgacat caaggtcaag gacaacgcca ccggcaccac ctacagccac 540



acggccacct acgtgacgcc ggacaacgcc agcacgaacc ccaacctcgc tgaggcgcca 600  
 ttccaaacga gcgtaggcta catcgcgtgg cagctctacg gcaagttctt cgggtgcgcc 660  
 gagtacactc tgcccaactc gacgaagaag taccgaagg tagaccgga cgccaacgcg 720  
 gcaaccatag ccggttacct caaccaacgg ggcgtgcagg acgggtacat cgcgttcctg 780  
 ggcgacgccg atatccacgt tgaagtgtcc acggacatgg gcaacaacta cggcatagcc 840  
 tccggcggta tgagcctcaa cgccacggca gacctgccgg cttactgcc gggcgcgggt 900  
 gtccttggcg tgggtgtgca gttcatggac ggcgctgtca tggccaccgg ctccaccaag 960  
 gccccggtat acttcgagtg ggattccgct aaccgccgct gggcagagcg ggccgcctac 1020  
 ggcaccgatt gggctctgaa gaagatgcca ctggccctgc gctgggatga ggctaccgac 1080  
 acctacagct tgaacgagct ggagtatgat cgacgtggct ccggcgacga ggatacgaac 1140  
 cccacgttca acttcgtcac ccgaggcatc accggcatga cgaccttcca gggtcgcctc 1200  
 gtcctcctgt cgcaggagta cgtctgcatg tcggccagta acaatccaca ccgctggttc 1260  
 aagaagtcgg cagccgcgct gaacgacgat gatcctatcg agatcgcagc ccaggggagc 1320  
 ctgactgaac cgtacgagca cgcggtcacc ttcaacaagg acttgatcgt cttcgccaag 1380  
 aagtatcagg ccgtgggtccc cgggtggcggc attgtaactc cccggacggc ggttatcagc 1440  
 atcaccacgc agtacgacct cgataccagg gcggcacctg ccgtgactgg ccgcagtgtg 1500  
 tacttcgctg cggagcgtgc cctgggtttc atgggcctgc atgagatggc cccgtctccg 1560  
 tccacggaca gccactacgt cgccgaagac gttaccagcc acatcccag ctacatgccg 1620  
 gggcctgctg agtacatcca ggcggcggcc tccagcggct acctggtggt cggcaccagc 1680  
 acggcggacg agatgatctg ccaccagtac ctctggcagg gcaacgagaa agtgcagaac 1740  
 gcgtttcatc gctggacggt gcgcatcag atcatcggcg cctacttcac tggtgacaac 1800  
 ctgatggttc tgattcagaa gggccaggag atcgccctgg gacggatgca cctgaacagc 1860  
 ctgccagccc gtgaggtct gcaataacct aaatacact actggcggcg tatcgaggcg 1920  
 accgtcgatg gtgagctgga actgaccaag cagcattggg acctgatcaa ggatgcctct 1980  
 gccgtgtacc agctacagcc tgtggcggc gcctacatgg agcgtacca tctcggcgtg 2040  
 aagcgcgaga cgaatacga ggtgttcctc gacgtgcccg aggccgtggt cggggcggtg 2100

038595

tatgtggtcg gctgcgagtt ctggtcgaag gtggagtca ctccgccggt tctccgggac 2160  
 cacaatggcc tgcccatgac ctcgaccctg gcagtgcttc atcgggtacaa cgtaaacttc 2220  
 ggctggaccg gcgagttcct gtggcgcatc agcgacacgg ctcgacccaa ccagccgtgg 2280  
 tacgacacga cggccctccg gttgttcagc cggcaactca atgccgggga gcctctggtg 2340  
 gatagcgtg tgggtccgct gccggcacgg gtcgatatgg ccacgtcaa gttcgagctg 2400  
 agctgtcaca gtccgtacga catgaacgtt cgggctgtcg agtacaactt caagtccaac 2460  
 caaacctaca ggagggtgtg a 2481

<210> 5  
 <211> 826  
 <212> белок  
 <213> Phikmvlikevirus LUZ19

<220>  
 <221> misc\_feature  
 <223> белок Gp34 LUZ19 дикого типа

<400> 5

Met Ser Tyr Lys Gln Ser Ala Tyr Pro Asn Leu Leu Met Gly Val Ser  
 1 5 10 15

Gln Gln Val Pro Phe Glu Arg Leu Pro Gly Gln Leu Ser Glu Gln Ile  
 20 25 30

Asn Met Val Ser Asp Pro Val Ser Gly Leu Arg Arg Arg Ser Gly Ile  
 35 40 45

Glu Leu Met Ala His Leu Leu His Thr Asp Gln Pro Trp Pro Arg Pro  
 50 55 60

Phe Leu Tyr His Thr Asn Leu Gly Gly Arg Ser Ile Ala Met Leu Val  
 65 70 75 80

Ala Gln His Arg Gly Glu Leu Tyr Leu Phe Asp Glu Arg Asp Gly Arg  
 85 90 95

Leu Leu Met Gly Gln Pro Leu Val His Asp Tyr Leu Lys Ala Asn Asp

## 038595

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

## 038595

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

038595

515 520 525

Glu Asp Val Thr Ser His Ile Pro Ser Tyr Met Pro Gly Pro Ala Glu  
 530 535 540

Tyr Ile Gln Ala Ala Ala Ser Ser Gly Tyr Leu Val Phe Gly Thr Ser  
 545 550 555 560

Thr Ala Asp Glu Met Ile Cys His Gln Tyr Leu Trp Gln Gly Asn Glu  
 565 570 575

Lys Val Gln Asn Ala Phe His Arg Trp Thr Leu Arg His Gln Ile Ile  
 580 585 590

Gly Ala Tyr Phe Thr Gly Asp Asn Leu Met Val Leu Ile Gln Lys Gly  
 595 600 605

Gln Glu Ile Ala Leu Gly Arg Met His Leu Asn Ser Leu Pro Ala Arg  
 610 615 620

Glu Gly Leu Gln Tyr Pro Lys Tyr Asp Tyr Trp Arg Arg Ile Glu Ala  
 625 630 635 640

Thr Val Asp Gly Glu Leu Glu Leu Thr Lys Gln His Trp Asp Leu Ile  
 645 650 655

Lys Asp Ala Ser Ala Val Tyr Gln Leu Gln Pro Val Ala Gly Ala Tyr  
 660 665 670

Met Glu Arg Thr His Leu Gly Val Lys Arg Glu Thr Asn Thr Lys Val  
 675 680 685

Phe Leu Asp Val Pro Glu Ala Val Val Gly Ala Val Tyr Val Val Gly  
 690 695 700

Cys Glu Phe Trp Ser Lys Val Glu Phe Thr Pro Pro Val Leu Arg Asp  
 705 710 715 720

His Asn Gly Leu Pro Met Thr Ser Thr Arg Ala Val Leu His Arg Tyr

038595

	725		730		735
Asn Val Asn Phe Gly Trp Thr Gly Glu Phe Leu Trp Arg Ile Ser Asp	740		745		750
Thr Ala Arg Pro Asn Gln Pro Trp Tyr Asp Thr Thr Pro Leu Arg Leu	755		760		765
Phe Ser Arg Gln Leu Asn Ala Gly Glu Pro Leu Val Asp Ser Ala Val	770		775		780
Val Pro Leu Pro Ala Arg Val Asp Met Ala Thr Ser Lys Phe Glu Leu	785		790		795
Ser Cys His Ser Pro Tyr Asp Met Asn Val Arg Ala Val Glu Tyr Asn	805		810		815
Phe Lys Ser Asn Gln Thr Tyr Arg Arg Val	820		825		

<210> 6

<211> 1497

<212> ДНК

<213> Pseudomonas aeruginosa

<400> 6

atgtccaatg acaacgaagt acctggttcc atggttattg tcgcacaagg tccagacgat	60
caatacgcac acgaggttcc ccctatcgat agcgcggccg ttgccgggaa tatgtttggc	120
gacttaattc aaagagaaat atatctacag aaaaacattt attatccagt ccgatctatt	180
tttgaacaag gaacaaaaga aaagaaggag atcaacaaga aagtatctga tcaagtcgat	240
ggcttgctaa agcagatcac tcaaggaaaa agggaggcca caaggcaaga gcgagtcgat	300
gtcatgtcgg cagtcctgca caagatggaa tctgatcttg aaggatacaa aaagaccttt	360
accaaaggcc cattcattga ctacgaaaag cagtcaagcc tctccatcta tgaggcctgg	420
gtcaagatct gggagaagaa ctcttgggaa gaaagaaaga agtacccttt tcagcagctt	480
gttagagatg aactggagcg ggcggttgcc tactacaac aagattcact ctctgaagcg	540
gtaaaagtgc taagacagga gctcaacaag caaaaagcgc taaaggaaaa agaggacctc	600

## 038595

tctcaactgg agcgggacta cagaaccgga aaggcgaatc tcgagatgaa agtacaatcc 660  
 gagcttgatc aagcgggaag tgctttgcct ccattgggtca gtccaacgcc agagcaatgg 720  
 cttgaactgt ccacaagact ggttacgcaa gcaattgctg ataaaaagca gctgcagacc 780  
 acaacaata ctcttatcaa gaattcccca acccctctag aaaagcagaa agccatctac 840  
 aatggtgagc tacttgtgga tgagatagcc agtctacagg cccgcttagt taagctgaac 900  
 gccgaaacga cacgacgcag gacagaagca gaacgcaagg cggccgagga acaagcgttg 960  
 caagatgcta ttaaatttac tgccgacttt tataaggaag taactgagaa atttggcgca 1020  
 cgaacatcgg agatggcgcg ccaactggcc gaaggcgcca gggggaaaaa tatcaggagt 1080  
 tcggcggaag caatcaagtc gtttgaaaag cacaaggatg cgtaaataa aaaacttagc 1140  
 cttaaagata ggcaagccat tgccaaagcc tttgattctc tagacaagca gatgatggcg 1200  
 aagagccttg agaaatttag caaaggcttt ggagttgtag gcaaagctat tgacgccgca 1260  
 agcctgtacc aagagttcaa gatatctacg gaaaccgggg actggaacc attctttgta 1320  
 aaaattgaaa cactagctgc tgggtcggcc gccagttggc ttgtgggtat tgcatattgcc 1380  
 acggcaacag ccactcctat aggcattctg gggttcgcac tggtaatggc agttaccggg 1440  
 gcgatgattg acgaagacct tctagaaaaa gcaaacaatc ttgtaatatc catttaa 1497

<210> 7  
 <211> 345  
 <212> ДНК  
 <213> Phikmvlikevirus LKD16

<400> 7  
 gagtaccaac tgaacacgag cgcaccctgc gctgcctgct ccaagacatc cacgggccgc 60  
 tgaatctgct gttcccaggt atccgggtga aggtggagga ggcgtgcctc ggatacttgg 120  
 gctacagga gcggggctat tgggagctgc gcctccaggt ggactacgac caccgaagc 180  
 ttgggcacct ccgctacagt caggccgtgc cggagtacgt gctgatcaac gaccgcgaca 240  
 gcatcatcaa gtacctgatg gaagcagtcc ctcggcaggt actagagggc atgctcaata 300  
 aggccagga attcgaacc aagaactggt attccctatg acgac 345

<210> 8

## 038595

<211> 204  
 <212> ДНК  
 <213> Phikmvlikevirus phi-KF77

<220>  
 <221> misc\_feature  
 <223> добавлена последовательность gp7 Phi KF77  
 <400> 8  
 tacaagggtgg tgacgcctag ctccggcagag ggccgcttg tgctggcgcac caagcagacg 60  
 cctgccctcg ctccaggcagt catcgtactg cacagcatga accccgcgca gtacgcggtg 120  
 ggcacggcca tactaaacac agactggcgg tgccgcccgc tgggtgccgg cgagtacatc 180  
 aagctcgttc aaggggaggc cgac 204

<210> 9  
 <211> 60  
 <212> ДНК  
 <213> Lambdavirus лямбда

<220>  
 <221> misc\_feature  
 <223> фар E. coli  
 <400> 9  
 atggttcgtg caaacaacg caacaggcct cgttctgaac aaatccagat ggagtctga 60

<210> 10  
 <211> 17757  
 <212> ДНК  
 <213> Цитомегаловирус ЦМВ человека

<220>  
 <221> misc\_feature  
 <223> предварительное редактирование фрагмента ЦМВ человека  
 <400> 10  
 acgacggcca gtgaattgta atacgactca ctataggcgc aattcgagct cggtagccga 60  
 ttaccctggt atccctacca ttccgggccc tgtgctgggt ccccgagggg cggggggggtg 120  
 tttttagcgg gggggtgaaa tttggagtct tggagccgcg tgtgctgtgg aggacggtga 180  
 cggtaggtaag agtgtgctgc ggtgcggttg ggacggcggc ggcaataaa agcggcgtgc 240  
 ggccgcgacg gcgaaaagca gacgcgctc tgtgttgtgt gtctttgacc gcggcggaac 300



acacgcggaa aagcgagtcc caggggacac acgacgagcg agtcccaggg ggggacgacg 360  
 acggccaggg acgcggaaac gacgcggaaa agaggaagtc cccaggggga cgggcggaaa 420  
 agaggaagcg cctaggggac cgcgggggca ggaacagacg aagtacgccg caaccgcgt 480  
 cgaggacaca cgcagaagcg gcccccagg ggaggggggg ggggggactc gcgggccccg 540  
 gggcacactt gttgttcctt ccggccgccg acacgcaccc cgaagcccg cacaccgccg 600  
 acacaccctt gacacaccg cgacacacc gccacacgcc cgacacacgc ccgcgacaca 660  
 cccgaccgac acaccctgac acacccgcc aacacaccca gccgcacccg ccccgccaac 720  
 acacccccga cacaccgac acacgccgc gacacaccg gcacacacc acccaccag 780  
 ccgcgcccc gacacacccc gaacggcgcc ggtgaggac agggctcacg gaggtttgcg 840  
 ggccgtgagc acgcctcctt ttgtacacac taccggtgcg tggcgtcca cgctatttgt 900  
 tcgcgagacc ggactaaggg aggtttgagg tgctcagcg cggggcggcg tttgcggcgt 960  
 gtttcgacca gcgctttgtg cgcgctgcct gtgctgtgct tccatggct tttgtcagcg 1020  
 gcacggcgct ggggacgggg tttcaccgc ctgagggatc tttctgagg tgtgaggac 1080  
 ggagcttttt tcgcacgctg ggcaccgggc tgggggacgg ggggtgtgct ggacggcgg 1140  
 ggggccgggg cgttgagggt acggggatta cgctgggaac ggggactcgc ggaccgggg 1200  
 tgagggacgg ggggtggcgg ggtgtttgct gcgaggacgg gggcctttt cggcggggac 1260  
 ggggactcac cctcgcctat ttaacctca cccacttcaa cacacacatg ccgcacaatc 1320  
 atgccagcca cagacacaaa cagcaccac accacgccg ttcaccaga gtaccaaac 1380  
 acgttacctt tacaccacag caacacacaa ccgctatcc aaacctcga caaacacgc 1440  
 aacgaagaac accgcacgca gatggagctc gacgccggg attacgctgc ttgcgcgag 1500  
 gcccgccaac acctctacg tcaaacacaa cccaactac acgcataccc caacgccaac 1560  
 cctcaggaaa gcgctcattt ttccacagaa aatcaacatc aactcacgca tctacttcac 1620  
 aacattggcg aaggcgcagc gctcggctac cccgtcccc gcgcggaaat ccgcccggc 1680  
 ggtggcgact gggccgacag cgcgagcgac ttcgacgcc actgctggtg catgtgggga 1740  
 cgcttcgaaa ccatgggccg ccaacctatc gtgaccttac tgttggcgcg ccaacgcgac 1800  
 ggcctcgtg actggaacgt cgtacgctgc cgcggcacag gctttcgcgc acacgattcc 1860

gaggacggcg tctctgtctg gcgtcagcac ttggtttttt tactcggagg ccacggccgc 1920  
cgtgtacagt tagaacgtcc atccgcggga gaagcccaag ctcgaggcct attgccacgc 1980  
atccggatca cccccatctc cacatctcca cgcccaaac caccagcc caccatctc 2040  
accgcatcgc acccacatgc tacgactcgc ccacatcaca cgctctttcc tatcccttct 2100  
acaccctcag ccacggttca caatccccga aactacgccg tccaacttca cgccgaaacg 2160  
acccgcacat ggcgctgggc acgacgcggt gaacgtggcg cgtggatgcc ggccgagaca 2220  
tttacctgctc ccaaggataa acgtccctgg tagacggggt aggggggatct accagcccag 2280  
ggatcgcgta tttcgcgcc acgctgcttc accgatatcc aataaaccca tcccctcgc 2340  
acgacgtctc cgctatctt tgtagcctca ggaatccgtc cccacgtcca tccatcccga 2400  
gcactccaca cgctataaca gaccacggac acggcaaatg catgcaaact tctcatttat 2460  
tgtgtctact actctgtgtt gctacagggg gtgaaggggg tgaaggcaaa gaaaaaaaa 2520  
aggaacaaaa taatagatta gcagaaggaa taatccgtgc gaccgagctt gtgcttcttt 2580  
tcttataagg aggcaaatat actagggaaa acttaagaat aggaagaaac cgaggtttgg 2640  
gagaaaagct gagataaaat agcgcatttt ccatacagag gttgttgttt ttgtggatcc 2700  
taagaggttt caagtgcgaa tctcaaagtt ctcacgagaa tattgtcttc aagaatcgac 2760  
aactgtggtc caagatTTTT ttttggcttt tttaggttct gcgagggaca tcacgatgga 2820  
tcgttcgat gaagtcacgc gtacgcctct ggtgtggcg ggtgtcgtga caggagagtg 2880  
tgTTTTcagt gcagagctgt cttgattcct atatccgagt atctgttttc tcgtaaggac 2940  
ggtaatcttc tttgggtgaa gtacatctaa aagctgcaaa ctatatttta agggctgtct 3000  
ctaggtgtac tttgatgctg gagtttttcg ctgtgttgat gtgaataaat ctactactac 3060  
tattatatgc agaaagagtg attatgccga gacaagattg cattggctga actgtttcaa 3120  
aaacgcctac actctactta tccgtaaacc taaggtaata ctatgtgtaa gttgtttttt 3180  
tttctttttg tagtaaaatg gtgatactg caattaaac tgtattccat gtttccatcc 3240  
tttcttttca actttaaagg cggctttgag agcgaagaag tgcgaggata aaaatggatg 3300  
actccttcgt gtccagggag tcgactactg caacgctgat tgattaaaag atggtctccg 3360  
atgatgatgt tgttattgat cgaatcatgg tgcagaacgg cgacggagag gagcgtgtcc 3420

gcccgggga aggtggtctc tttctctttt cttttttcaa gaaatcttcc atgtgtttat 3480  
cgtagtgatc gaaatcgact gatctcgggt tctttttggt ggtttctttt cggttaatca 3540  
tgtattgttt tcttttttta cagaaagata cttttttcat gagcaattcc tcgcccggcg 3600  
ccggcatgcc gaggtggggc cactgcatc agcggcatgc cgacgccgac ccggggatct 3660  
tggattcacc gttttctctc ttctctctct acatacagac cgggtggcag gagcggtaag 3720  
gaatcatcgt cgtctttcat tcttcgatga ttatggtaat actaaatctt atctaggagc 3780  
atatacatct aagattggag tactagtagt cgtttggtt ttctattttt ttatatatta 3840  
tctatgacag tttttctggt tttcgttttg ataataatat aataaaaact catggacgtg 3900  
aaatctggct tggttggtt gatttcattc tcattattgt tgttttcttt ccgtcttgcg 3960  
gatgaagatg ttgcatgacg gttggtgtt gtgttgctat acaccgagag agatgatctt 4020  
ttgtttcttc tggttcattt cctatgattg tttggctgct gaccgacgcg tcaggatgtg 4080  
cagggcatgc ggggaatcag gaccggacac gggataattt catctaccta tacggagatc 4140  
gcggtcctcg ccatgaggat cgcgacaggc gcgtcgaggg ggcaggaaca cccttgcgga 4200  
ttgacattct tgggtggtgtt tcgttgtgtt cggtagttgt tgttgacgat gaggataaat 4260  
aaaaatgacc ttgtttttgt tctgttttct ctgtttggga atcgtcgact ttgaattctt 4320  
cgagttatcg gaaagctgag gtaccctaat gtctgtagct tttttctttt taccctcttg 4380  
tttatcatct gcgattcgtg gtaggtagga gagggaaatg ataatccgag attaaggaaa 4440  
ggagaagata aaaaataaaa aaaaaataat aaaacagaag ccgaccggcc gccgaccctg 4500  
tccccaggac cagcctacga ggaatggata acgcggtggc gacggcagcg gtggtggcgc 4560  
tgggggtggc ggcagtggta ctgctgatgg tagtcgggac ggaggagagg cgatgcatac 4620  
atacacgcgt gcatgctgca tgggtggatg gtacggccgg gagacgcgga agagaaactc 4680  
acataaaaag gtgacaaaaa gagcggttga aaaaagaaaa cgagattcga ccagacagaa 4740  
gagaaggacc ggggcttggc gacccttcca cgactgctgt tgtcatctcg gctccccgt 4800  
cttctcccgg ccacgggagg ctaagtcacc gccgttctcc ccatccgtcc gagcggcagc 4860  
cgaccagccg gccgattcgc ccgccggggc ttctggagaa cgccggggca gcagcgatct 4920  
ggggaagccg ctaaaccctt gcgtttttat atggtagctc tgccgagcgc gggctgacgc 4980

gttgagtaag cggaaagacg tgtgtgacga aaaggggtcc catggtattht cacgtgacga 5040  
 tgaggagatg cggtttgag cacatacgggt ttagaaaaag ggagttgtcg tgacaagggc 5100  
 tgaggacact ctgtctccat gtgtgtataa aaagcaaggc acgttcataa tgtaaaaaag 5160  
 aacacgttgt aaacaagcta ttgtgtatc attcggctga ctatgcttca ttcggactga 5220  
 ttttcttttc ctaacggcgt aacttaaagt gattaacgta tgatatttgt tccccagagt 5280  
 tatactatag tcatcatcct aaaattcaga tataaatgaa cacatgtcgt atgggattat 5340  
 taagaaaccg aaactctcca cagttcacca tcttcttcgt cattcaaccg atgaccact 5400  
 ccgtacaacg aatcagtctg ctgtgtcaca ctgcaacta ctacgacgt atgcaaaaa 5460  
 cttgaaacac gggctgttgt attgacgacc gttgtaccat tactagtcac attgcataga 5520  
 gaccatccac cgtcatcca tctttccac ccgatggaaa accgtcttct atcatcaact 5580  
 atggtaaagat ttcgaccctg cgaggattc agtttccca tatccataac ctggatttta 5640  
 tcattaaacc ccaatattaa acactttttt agtaccctcc caccaccaa aaaatgtgac 5700  
 tggaccggtt cctagcagct ctgggagcca tgttcaggtt gaaccacagc tacagcgaaa 5760  
 ccgagtccag tgaccggtta ccacgtccag cccctgcgta tgtaccagtc caagcacgtc 5820  
 cggtcattgt tctacacagg aaatctaact aggtcaacgc aattttattc caccgttacg 5880  
 cagaatacta acaaaaaaac acacaaattt aacgaattac acgtagttha ttacatgaaa 5940  
 actgtaagaa caccaattca ctaagcgata caacatttag ctgacttcca agtgccacac 6000  
 atcaccactg tattcatcca tgttttcacc gaaccaacga gacagatcga agaagccaga 6060  
 atctcccacg tttaaattac ataaatcaa cgtattatga ccacagctcg acacacaaat 6120  
 agttgcgtha ctattcacag tagcattacc tataaccgta acgttgcaaa accactgatc 6180  
 accattgtha ccaaaaacgg ttttccactt agttgtcaac ggatctttcc catgcgtha 6240  
 ggtcaaattha ctaccagtcg tcgcttttag ctattacga gtattatccg catccacata 6300  
 tatcaacgtc atagctaggc acgctataag tcccccccc ccacaatgga atggtgcaaa 6360  
 accggttctt tcccgttata gccatagcgt tcccaggcaa aagcaaacgc caaacctaata 6420  
 gcagtgaaaa gcgcttgag ccagaaccag cttatgtacc agccacaatc acatccggtt 6480  
 attgtttcca caggaaatcc taccaggcaa agccccgctt gttttgttcc tgaccatctt 6540

gtttagcaat tcgtaaactg tcagcctagc gacgtccggt tagatcaaaa gtcacgtata 6600  
 tagcgacgct gtttccaccc gtttccccgt cccgccgttt ccgaacaacc cacccggggt 6660  
 cagacaaccg accaccaaca gaaatataca cacagaccac cgggagttca gttaaagatt 6720  
 tcatacagggt tattttggct gctgctagtc ttttgcttct tagaaaaaaa atacccatat 6780  
 agagaaataa tgatagtttg acaacacata tggcagggat ttcttcttca tcaataagat 6840  
 atgcaattcc cccagggaga gactttcaac aattgaattt acaaaaacaa aattacatca 6900  
 ggagaaagag aggatacatt aataaatata ttatatctgg tgtatatact gaatgctgct 6960  
 ggttcataag gtaacgatgc tacttttttt aattccaaga tggtttttct ttgttagtct 7020  
 tttgttgact tgctggttcc taaaagtctg caaaaacgat tgtgtgaaga ttatgacgtt 7080  
 ggttgactag ttcatagatg tctgctgtac gtgtgatggt tattcgctgg ttcgttctaa 7140  
 gatgagtatc gtactgtgtc tgcgatggtc gtctcttact ggcatctctt cggctgcctc 7200  
 ttgttttcat gattgaaaag gaaaaaagga ctccgagggc gcggtcatct tttacttttc 7260  
 ggttttctcg ttggcgggct agaggtagtc agatcatgag actgtcgtgg tcgatgaaac 7320  
 tgtgtctgct caagtgacgt ccatttcttg tacggagaaa aaagtcacg ggataaataa 7380  
 ggctatacaa ggcgttgca agcgtgcggc tctaaacaaa ttaagcgata caaaattaca 7440  
 gtgatacгаа taataaatta cccctcccc ctgtggtccc cccgaggcga gagccacca 7500  
 tcgtgtactc tcgaccacc cacgaccaca gggggagacg ggacgaagag acgacgcaga 7560  
 gcgccatctc ctctggagg ccggcggcgt taactgctac agctgcggcg gcgacgacag 7620  
 ctgcgatttg tcggccgaca tgccgatggt atgggcggcg gcggcgggtg ccgcccagc 7680  
 ggggaggaga ggagagagaa gaggagcggg gcgtccgaag gcgaggatgg catggtctcg 7740  
 ccggagcggc cggcttttat ggaacactcg cgtccggtg ggtatcacc acaggaagat 7800  
 gaatcacaac ttcaaacca tcttgagacc cgagtaacgg ttacaggtc gcacgccagt 7860  
 ctacagctaaa aacagcggac agtcccacgc tgtttctggt tggctctct ccagtttct 7920  
 catcgccgtc ttggtctccg tcatcatcgg aagaatacca cccgctctca tgcggcagtc 7980  
 gatcagcctc gatgaacgag acgcggcgac gcctttctac ggccgactgg ttgtggtggt 8040  
 gaaagaagag caccagcaat cccaggagga gcaacaagcc ctacatgctc caggaggtcg 8100

gggagagggc ctgtcggaga tgaccgtgag gcatcacgta cggcagctga ggagaaacgg 8160  
 agaagaaagg aaaattaccg tcaggggccg gggttcttat tagagaaaca gcacgtaggt 8220  
 caggatccag atgctaattg caatcatgat gacgatgatc atgcaggcca agacgcggcg 8280  
 caccaatgca gaatccaata gccgccgtgc ctccggttgg tggccggcgg catctagaga 8340  
 catgatttgg gggggggacc ggcggcgcaa aaagacaggg agatggacag tgccacggtg 8400  
 ttttgttatg attaggacat ggggaccgga agccgagaca gagtactaca ggggtgtgaa 8460  
 gggtaacgtg agggagatca tgtcatgggc gggctgaaga ccgtgcgggg aggatcgacg 8520  
 tgtgcgggtc ttgtggaaca cgggtgttta atatgtatcc gcgtgtaatg cacgcggtgt 8580  
 gctttttagc actcggcttg ataagctacg tgaccgtctg cgctgaaacc atggtcgcca 8640  
 ccaactgtct cgtgaaaaca gaaaataccc acctagcatg taagtgcaat ccgaatagta 8700  
 catctaccaa tggcagcaag tgccacgcga tgtgcaaatg ccgggtcaca gaaccatta 8760  
 ccatgctagg cgcatactcg gcctggggcg cgggctcgtt cgtggccacg ctgatagtcc 8820  
 tgctgggtgt cttcttcgta atttacgcgc gcgaggagga gaaaaacaac acgggcaccg 8880  
 aggtagatca atgtctggcc tatcggagcc tgacacgcaa aaagctggaa caacacgcgg 8940  
 ctaaaaagca gaacatctac gaacggattc cataccgacc ctccagacag aaagataact 9000  
 ccccgttgat cgaaccgacg ggcacagacg acgaagagga cgaggacgac gacgtttaac 9060  
 gaggaagacg agaacgtgtt ttgcaccatg cagacctaca gcaactccct cacgcttgtc 9120  
 atagtcacgt cgctgttttt attcacagct cagggaagtt tatcgaatgc cgtcgaacca 9180  
 atcaaaaaac ccctaaagct gcccaactac cgcgccactt gcgaaaaccg tacacgcacg 9240  
 ctggttacca ggcttaacac tagccatcac agcgtagtct ggcaacttta tgatatctac 9300  
 agcagataca tgcgtcgtat gccgccactt tgcatcatta cagacgccta taaagaaacc 9360  
 acgcgtcagg gtggcgcaac tttcacgtgc acgcgcaaaa atctcacgct gtacaatctt 9420  
 acggttaaag atacgggagt ctaccttcta caggatcagt ataccggcga tgtcgaagct 9480  
 ttctacctca tcatccacc acgcagcttc tgccgagcct tggaaacgcg tcgatgcttt 9540  
 tatccgggac caggcagagt cgggtgtggtc acggattccc aagaggcaga ccgagcaatt 9600  
 atctcggatt taaaacgcca gtggtccggc ctctcactcc attgcgcctg ggtttcggga 9660

## 038595

ctgatgatct ttgttggcgc actggtcatc tgctttctgc gatcgcaacg aatcggagaa	9720
caggacgttg aacatctgcg gacggacctg gatacggaac ctttgttgtt gacgggtggac	9780
gggaatttgg aataaaagat gcgtaacacc tgtcgaagat gcgataactt tacatacagg	9840
caaacagtgt atacaattat agtattttgt atgttgcata aagttacatg caacagtact	9900
gctaacagta ctgcatccat tacgctatcc aacctgcct ctaccacttt tgtaaccaac	9960
atatattcaa ctccgaataa caacacatca acgacgccac acacatctgt cacctcacia	10020
gcgtcaacca ttggcaacat caccaacgtt acctccgact tgagtacttt cacaaccgta	10080
tattctacat tcaatacatc atttgccaat atatctaata cggctgtcac tacagaattg	10140
atltcaacia ataccaacac tatctcatct tttaccaacg taacagcaaa cgctacatca	10200
tcttataaca caacaatcac cgtaactgtc acgtcagatg aaacttcgca caacgtatcc	10260
actaataatg cacttataag cacaccatgg cctacaaatt gcagcgccac aacatacacc	10320
acgtacaacc ttactaactc ttccaacgct tgtcacacag agacaacaat catacgtttc	10380
aaggaaacca atacaacagg aatagaaggg agtaatgtca ccataaaggg taattctacg	10440
tgggactgtc tttcagtcgc ctggatacga cattacaata gatccacaca cggacatcat	10500
ctaggttatc gtaagaacgc acatacccaa tcttggattt ggctacgcat cttacctct	10560
cacactgtat gtcatttctca acatgaaaga ctttcaactgt accatgactt atgtcgttcg	10620
tgcaacaaca cagaattaca tctgtacgat ctaaataatca ccaattccgg caggtagcgc	10680
agacgttgtt ttaaagaaaa ttacttcaca ggacatcacg aagatgaaaa tttctaccta	10740
ttagtaacac caaaaaatca tactgaagct attaattgcta ctttcgtttg ccctagatac	10800
aacaccgata tcgaaaatga agatagagag aaaggaagtc aacatactaa caatacacat	10860
caccacaaac gtaatctcta tcatagctcg caaagaagcc gcaccgtatg gaccatcgtg	10920
ttggtttgta tggcctgcat agttctgttt tttgcacgac gagcctttaa caaaaagtat	10980
catatgttac aagacaccgt cagtgaatca gaattcattg ttcgatatca cccagaacat	11040
gaagattgag ctacgtttcc gggcagacat cttatgaagc tgaacaataa actaaaacat	11100
tctgtaagac tcagcgttca aaggaatatt aatgccatt gagcgaaaac taatattgca	11160
atggactggc gatttacggt tacgtggacc gttacttgtg atggtttcaa ttatacagtc	11220

## 038595

cataaaagat gcgatcgag ttacgaggta atcaacgtaa caggatacgt tggtagcaac 11280  
ataactctaa aaaaatgcaa tcagactgag aatggcaca atgtagactg gattcattat 11340  
gagtacccca cgcataaaat gtgcgaatta ggcaactatc accaaaccac accacggcac 11400  
gacatatgtt ttgactgcaa cgacacctcc ctaactatct acaacttaac cacaaaaaac 11460  
gctggaaaat ataccaggcg tcaccgtgat aacggccaag aagaaaatta ctacgtaacg 11520  
gtgttaattg gagacacaac gttattcact cttggcacat gccctgtaag atataaagaa 11580  
tctacgaaca ctgaaaacac cattggaagt agcatcatag aaaccattga gaaagctaac 11640  
attcccctgg gaattcatgc tgtatgggca ggcgtagtgg tatcagtggc gcttatagcg 11700  
ttgtacatgg gtagccatcg cattcccaa aagccgatt acaccaaact tcccaaatat 11760  
gatccagatg aatthggac taaggcttaa catgctgac aataaacttt ttttaaccaa 11820  
taacatgtct ccgtttttt ttgttaacaa cctatgatat aaagcgttat attcagtcgt 11880  
tactaaacaa aaaaacatgg gcatgcaatg caacactaaa ttgttattgc cagtcgcact 11940  
aataccggtt gcaatcatcc taattggtac tctagtgccg atacttttac atgaacaaaa 12000  
aaaggcgttt tactggcgac tttttctgca aagtcaacat gtagaagcac ccattacagt 12060  
aacgcaggga gacacagtct acctagacgc tagcaataat ccctgtaatt attccagctt 12120  
ttggtaccac ggtaattgcg aactttgtgg atggaacgga tatctacgca atgttacaca 12180  
ttactacaca aacacatcgt gttccccgca attcatctgc ataaacgaaa ctaaaggtct 12240  
gcagttatat aatgtaacat taaacgattc aggcgcttat actgaacacg tttacgaatg 12300  
tgacctttcg tgtaacatta ctactaataa cgaatatgaa atactcaatt attttgataa 12360  
ctgtaactac accataaata gcaccaagca tattatcacc gtggtgtctt cacgtcattc 12420  
taaacaacaa aattcccacg tatccactca cgctggttgg gcagtcgccg tggtagcgg 12480  
aattatgatc tacgttctga tccactttaa cgccccgca actctgagac acaaactacg 12540  
aactagaaac aacgtaaadc gcatagcgtg attataaagt atcgacgcta atttctcaa 12600  
gataaaatth gattactccg tgcagttctc aaaaactgta aggccccgct tttccactcc 12660  
gtcatgaagg atcgcaatag aatactgcta tgtatcatct ttatttgcat tatgtgcctc 12720  
atthgtatth actthaaacg tcgttggtt tttactccgt ctccagacaa agcagatctg 12780



cgagtggaat ttccctcgtt acccccgtgt attggcatac agtgcgctgc atgagaacac 12840  
 gcgtgacaca tagcgtacct ctggacggta cagtttatga taacgtaatt cagggaaagt 12900  
 atacattcat accaacatgt tatcacataa cacacagatt ttctgctgtt ttataaaaag 12960  
 agcgtctcga agcagcttga gccacactac ggtccagatg acgagcgtaa ttaaaaatat 13020  
 gccgcgcagt attcгаааgс сgtactgagc gtgcgaggcg ggtaggggtgc cgaacgacgg 13080  
 atatgсgtсg ttgtcatcct cgactataag gatcсgсgacc gagtcttcgg ccatggtaaa 13140  
 cgtcaccctg tgtggctggg atgtagcgta tccggtttgg aattgttctg ctccagctcg 13200  
 ggggatagtg aggaattctc aagggatagc ggaccaatg actggataag agaagggttt 13260  
 ttccccgtaa gatgatcctc gtatcacatg aggtctggat atgtataaat gaagagtгаа 13320  
 ataggcacag ggaatcagat gccagcctcg tgatgcagcc gctggttctc tcggcгааа 13380  
 aactgtcgtc ttgtctgact tgcaaataca tcccgcctta agcgatgagt ctataaaгсa 13440  
 ccgttgcccg agtacggtaa aagtгacccg gattgtagaa cgtccttttt tttgttttt 13500  
 gcacgttta tcgtcactac tagtgcaata ttttgattgt aaggctгааа gagtatcgtt 13560  
 atgatgctta gaacgtggag attattacag atggtagctc ttgcccgta ctgttattat 13620  
 gtttttgсgа ctgttcaat cagcacgacg actgctcctg tggaatggaa gtctcccгac 13680  
 cgtcagattc ccaagaatat tacctgcgct aattactcag ggaccgtcaa cggcaacgtt 13740  
 acatttcgag gtcttcagaa caaacggaa gactttttgt actggttggtt aggatggggт 13800  
 cataagtcca ttgttcgtt cttcccгааа ctccagggtа actatgacгa аcaacattac 13860  
 agatatgаag tagcгaacct gacgtataac tgcacctata accgcttgac gttgctgаat 13920  
 ctgacgacgg aaaacagcgg aaagtactat ttcaaaaggг aagatgсгaa tttcaccttc 13980  
 tattactctt gttacaactt gaccgtgtcc taaagatcгc acgtgаagtt tcacagagcc 14040  
 gcgtggctgt agctattgtg tttagcttgc ttttgaaatg ttaagcgtcc ctacggcгct 14100  
 aacatgtttc taggctactc tgactgtgta gatcccggcc ttgctgtgta tcgtgtatct 14160  
 agatcacgct taaagctcat gttgtctttt gtgtggttgg tcggtttgсg tttctatgat 14220  
 tgtgccgсgt tcgagtcctg ctgttacgac atcaccgagg cggagagtaa caaggctata 14280  
 tcaaggгacg aagcagcatt cacctccagc gtgagcaccс gtacaccgtc cctggcгatc 14340

gcgcctctc ctgaccgatc gatgctgttg tcgcgagagg aagaactcgt tccgtggagt 14400  
 cgtctcatca tactaagca gttctacgga ggcttgattt tccacaccac ctgggtcacc 14460  
 ggcttcgtcc tgctaggact cttgacgctt ttcgccagcc tgtttcgcgt accgcaatcc 14520  
 atctgtcgtt tctgcataga ccgtctccgg gacatcgccc gtcctctgaa ataccgctat 14580  
 caacgtcttg tcgctaccgt gtagctagtt agccagctgt gtgtagtggt ttgcttttgc 14640  
 atatttgttt tcagtcagag agtctgaaac ggggtgggag ggacttttgc gggtagtgca 14700  
 tgctaagatg aacgggtggg ctgggggtgtg cttgataact cactgtttga atacgcgctc 14760  
 acgcacatat gtagcactca acatgttagc ttttgcccgc acgccccggg gcgtgccgag 14820  
 ctgccttttt aataaagtct gggtttccag atacgcgctg gttctgattt tgatggtttg 14880  
 tgctctgaa agctctacga gctgggccgt gacatccaat ggactgccta actgtagcac 14940  
 ggtaactaga acagcgggtc aagacgctga attgcacggt ccggcaccgt taagctgtaa 15000  
 tgtgaccagc tggggacgtt acgagaatgg aagcacaccg gtgttatggt gcactttacg 15060  
 gggatcaagc atgcgagtct cattaggaca ccgtgtagcg tttggctggt cttggaaaac 15120  
 attttttatt tataacgttt ctgaaagtag cgggtggcact tactatcaaa aaggttacia 15180  
 ctgcaccgac aaacatataa cactatcttg tttcaactta acggtggttc ctcgagcgtt 15240  
 tcaaagcaca accaccgtaa tgacaccac gctggttaca aactccacat tcagtgtgtc 15300  
 acttgttccg ttgagactga cgacaaattc cagcgcgttt ggacacgcta tttatcaacg 15360  
 acaacagcgt gttgaaaacg ggacgttatc caagaacata actaacttgg cattcaccta 15420  
 tggcagctgg ggcgttgca tgctgctggt tgccgccgtg atggtgctcg ttgatttggg 15480  
 tttgcctcaa tcggcttggc gacgctggcg aagccacgtg gacgatgaag aacgtggttt 15540  
 gttaatgtag gaaataaaag gcagtttgag catgactggt tccaaaccgt aacgtggtaa 15600  
 ataaatcatg gcttccgacg tgggttctca tcctctgacg gttacacgat ttcgctgcag 15660  
 agtgcattat gtgtacaata aactgttgat ttttaactttg tttgccccg tgattctgga 15720  
 atccgtcatc tacgtgtccg ggccacaggg agggaaactt accctggtat ccaactcac 15780  
 ttcaaacatc agcgcacggt ggttccgctg ggacggcaac gatagccatc tcatttgctt 15840  
 ttacaaacgt ggagagggtc tttctacgcc ctatgtgggt ttaagcctaa gttgtgcggc 15900

taaccaa	atc	accatcttca	acctcacg	tt	gaacgactcc	ggtcgttacg	gagcagaagg	15960				
ttttacgaga	agcggcgaaa	atgaaacg	tt	cctgtgg	tat	aatttgaccg	tgaaacccaa	16020				
acc	tttgaa	actactccag	ctagtaac	gt	aacaaccatc	gtcacgacga	catcgacgat	16080				
gatcgacg	cg	aaaagtaacg	ttacaggg	aa	cgccagttta	gcaccacaat	tacgtgccgt	16140				
cgctggattc	tccaatcaga	cg	cctttgga	aa	acaacacg	cacctggcct	tggtaggtgt	16200				
tgttg	tgttt	ttagttctga	tagttg	ttg	cattatgggg	tggtggaaat	tgttg	16260				
taaaccagag	ttatagtaat	gtg	cctttta	tcagggagaa	ggtttt	gtgc	caacaatgac	16320				
tagcccggga	ctatctg	cg	taata	tgacgg	aaat	tatgaattca	cggaaaccgc	16380				
caatacaacg	cgtaca	aaata	gaagtg	actg	gacaac	gta	gaaaccagtg	cattgctatt	16440			
gaaaaacacg	gagactg	cag	tgaac	ctcag	caacgcg	act	acgg	tcaccc	cacaacctgt	16500		
agaatacccg	gctggg	gaag	taca	aatca	aagaac	ggca	acgcattatt	cttggatgct	16560			
aatcattgtc	atcattctca	tc	at	ttttat	tatcat	ctgt	ctacgagcac	ctcgaaaaat	16620			
ctaccatcac	tggaaagaca	gtaa	acagta	cggaca	agtg	tttatgacag	acacggaact	16680				
gtgacagtga	tgtctaagcg	tttgcag	gta	ttccat	gga	taacaat	ttt	at	tttacaca	16740		
tcaaaatccc	agtattggaa	ctat	atggca	ataccat	gta	cccctacag	tggatacggc	16800				
agtcataata	ttagcttga	tccg	ctta	aat	cattat	ttcaag	acga	tgtttttgaa	16860			
tggtacatag	acaaaccaat	ggttaca	agt	tatgt	cctta	tcaaag	ta	gaacgc	acaa	16920		
aatccaatct	agactctcca	aatatt	gtgt	ggcaat	gcac	agataat	cg	ac	actaattc	16980		
tc	atgaactt	aaccaca	aca	tacag	tagaa	actatt	ttt	tcaat	ccttt	aaat	atctcg	17040
gacgaggagt	acaaaaaccg	aataact	gt	gtata	acgt	tagt	gtacac	ttt	accacc	17100		
aaacacattg	ccatacaact	acat	catccc	tgtat	ccacc	tacat	ctgta	cac	gattcat	17160		
tagaaat	atc	acag	tcattc	ac	ctcaacca	act	tcacaca	tacc	gcggtc	cact	acgcca	17220
ccg	gtaacgt	tgaag	cacaa	cacgac	acta	ccact	ccaca	taca	atgtgg	atcata	cccc	17280
tagttat	cg	tata	caatc	atcg	ttttaa	ctg	tttcaa	att	ccccag	aaag	cttgga	17340
ataaattcac	acaatacaga	tacag	cggt	tgct	cgccgc	cg	cttaaaga	atca	acgcca	17400		
aggaaaccaa	aacgtaaaaa	gaatag	at	gtac	gtttat	tttt	cagctc	act	gtttgaa	17460		

038595

taccgtaaac ataatgacgt acatatacgt ggttatacaa cagggtgttg tgttatgcgg 17520  
 cgactgatta accatatacgt gaaccatgat cttttccgat ggtccgtcgt gaccgcaatg 17580  
 atattttaca gatattccga aacctgtatg gaggtcactg tcagagtagg tgatccagtt 17640  
 accctcggta gtggacatgg ttatcatcca ggtagggata acagggtaat gatcctctag 17700  
 agtcgacctg caggcatgca agcttgagta ttctatagtc tcacctaaat agcttgg 17757

<210> 11  
 <211> 18036  
 <212> ДНК  
 <213> Цитомегаловирус ЦМВ человека

<220>  
 <221> misc\_feature  
 <223> предварительное редактирование фрагмента ЦМВ человека

<400> 11  
 acgacggcca gtgaattgta atacgactca ctatagggcg aattcgagct cgttaccgga 60  
 ttacctgtt atccctacca ttccgggccg tgtgctgggt ccccgagggg cgggggggtg 120  
 tttttagcgg gggggtgaaa tttggagtct tggagcccg tggtctgtgg aggacgggtga 180  
 cggtggtgta agtggtctgc ggtgctggtg ggacggcggc ggcaataaa agcggcgtgc 240  
 ggccgcgacg gcgaaaaagca gacgcgcgtc tgtgtttgtg gtctttgacc gcggcgggaa 300  
 acacgcggaa aagcaggtcc caggggacac acgacgagcg agtcccaggg ggggacgacg 360  
 acggccaggg acgcggaaaac gacgcggaaa agaggaagtc cccaggggga cgggcggaaa 420  
 agaggaagcg cctaggggac cgcgggggca ggaacagacg aagtacgccg caaccgcgt 480  
 cgaggacaca cgcagaagcg gccgcccagg ggaggggggg ggggggactc gcgggccccg 540  
 gggcacactt gttgttccct ccggcccgcc acacgcacc cgaagcccg caccgcccg 600  
 acacaccct gacacaccg cgacacacc gccacacgcc cgacacacgc ccgcgacaca 660  
 cccgaccgac acaccctgac acaccgcc aacacacca gccgcaccg ccccgccaac 720  
 acacccccga cacaccgac acacgccgc gacacaccg gcacacacc acccaccag 780  
 ccgcgcccc gacacacccc gaacggcgcc ggtgcgggac agggctcacg gagtttgcg 840  
 ggccgtgagc acgcctccct ttgtacacac taccggtgcg tggcgtcca cgctatttgt 900

tcgcgagacc ggactaaggg aggtttgcgg tgcgtcagcg cggggcggcg tttgcggcgt 960  
 gtttcgacca gcgctttgtg cgcgctgcct gtgcgtgtcg tcccatggtc tttgtcagcg 1020  
 gcacggcgct ggggacgggg tttcaccgcg ctgagggatc tttctgcggg tgtgagggac 1080  
 ggagcttttt tcgcacgctg ggcaccgggc tgggggacgg ggggtgtgcg ggacggcggg 1140  
 ggggccgggg cgttgcgggt acggggatta cgctgggaac ggggactcgc ggaccgggc 1200  
 tgagggacgg ggggtggcggg ggtgtttgcg gcgaggacgg gggccttttg cggcggggac 1260  
 ggggactcac cctcgcctat ttaacctca cccacttcaa cacacacatg ccgcacaatc 1320  
 atgccagcca cagacacaaa cagcaccac accacgccg ttcaccaga gtaccaaac 1380  
 acgttacct tacaccacag caacacaaa ccgcctatcc aaacctgga caaacacgcc 1440  
 aacgaagaac accgcacgca gatggagctc gacgccgcg attacgtgc ttgcgcgag 1500  
 gcccgcaac acctctacg tcaaacaaa ccccaactac acgcataccc caacgcaac 1560  
 cctcaggaaa gcgctcattt ttccacagaa aatcaacatc aactcacgca tctacttcac 1620  
 aacattggcg aaggcgcagc gctcggctac cccgtcccc gcgcggaaat ccgccgccc 1680  
 ggtggcgact gggccgacag cgcgagcgc ttcgacgcc actgctggtg catgtgggga 1740  
 cgcttcggaa ccatgggccc ccaacctatc gtgacctac tgttggcgcg ccaacgcgac 1800  
 ggcctcgtg actggaacgt cgtacgtgc cgcggcacag gctttcgcgc acacgattcc 1860  
 gaggacggcg tctctgtctg gcgtcagcac ttggttttt tactcggagg ccacggccc 1920  
 cgtgtacagt tagaacgtcc atccgaggga gaagccaag ctcgaggcct attgccacgc 1980  
 atccgatca ccccatctc cacatctca cgcacaaaac caccagcc caccatctc 2040  
 accgcatcgc acccatatgc tacgactcgc ccacatcaca cgctctttcc tatcccttct 2100  
 acaccctcag ccacggttca caatccccga aactagccc tccaacttca cgccgaaacg 2160  
 accgcacat ggcgctgggc acgacgcggg gaacgtggcg cgtggatgcc ggccgagaca 2220  
 tttacatgtc ccaaggataa acgtccctgg tagacggggt aggggatct accagcccag 2280  
 ggatcgcgta tttcggccc acgctgcttc accgatatcc aataaaccca tcccctcgc 2340  
 acgacgtctc cgcgtatctt tgtagcctca ggaatccgtc cccacgtcca tccatcccga 2400  
 gcactccaca cgctataaca gaccacggac acggcaaatg catgcaaact tctcatttat 2460

tgtgtctact actctgtgtt gctacagggg gtgaaggggg tgaaggcaaa gaaaaaaaaa 2520  
 aggaacaaaa taatagatta gcagaaggaa taatccgtgc gaccgagctt gtgcttcttt 2580  
 tcttataagg aggcaaatat actagggaaa acttaagaat aggaagaaac cgaggtttgg 2640  
 gagaaaagct gagataaaat agcgcatttt ccatacagag gttgttgttt ttgtggatcc 2700  
 taagaggttt caagtgcgaa tctcaaagtt ctcacagaaa tattgtcttc aagaatcgac 2760  
 aactgtggtc caagatTTTT ttttggctct ttttaggttct gcgagggaca tcacgatgga 2820  
 tcgttgcat gaagtacgc gtacgcctct ggtgtggcgc ggtgtcgtga caggagagtg 2880  
 tgttttcagt gcagagctgt cttgattcct atatccgagt atctgttttc tcgtaaggac 2940  
 ggtaatcttc tttgggtgaa gtacatctaa aagctgcaaa ctatatttta agggctgtct 3000  
 ctaggtgtac tttgatgctg gagtttttcg ctgtgttgat gtgaataaat ctactactac 3060  
 tattatatgc agaaagagtg attatgccga gacaagattg cattggctga actgtttcaa 3120  
 aaacgcctac actctactta tccgtaaacc taaggtaata ctatgtgtaa gttgtttttt 3180  
 tttctttttg tagtaaaatg gtgatacgtg caattaaac tgtattccat gtttccatcc 3240  
 tttcatttca actttaaagg cggctttgag agcgaagaag tgcgaggata aaaatggatg 3300  
 actccttctg gtccagggag tcgactactg caacgctgat tgattaaaag atggtctccg 3360  
 atgatgatgt tgttattgat cgaatcatgg tgcagaacgg cgacggagag gagcgtgtcc 3420  
 gccccggga aggtggctct tttctctttt cttttttcaa gaaatcttcc atgtgtttat 3480  
 cgtagtgatc gaaatcgact gatctcgggt tctttttgtt ggtttctttt cggttaatca 3540  
 tgtattgttt tcttttttta cagaaagata cttttttcat gagcaattcc tcgcccggcg 3600  
 ccggcatgcc gaggtggggc cactgcatc agcggcatgc cgacgccgac ccggggatct 3660  
 tggattcacc gttttctctc ttctctctct acatacagac cgggtggcag gagcggtaag 3720  
 gaatcatcgt cgtctttcat tcttcgatga ttatggtaat actaaatctt atctaggagc 3780  
 atatacatct aagattggag tactagtagt cgtttgggt ttctattttt tttatattta 3840  
 tctatgacag tttttctgtt tttcgttttg ataataatat aataaaaact catggacgtg 3900  
 aaatctggct tggttgtggt gatttcattc tcattattgt tgttttcttt ccgtcttgcg 3960  
 gatgaagatg ttgcgatgc gttgttgttg gtgttgctat acaccgagag agatgatctt 4020

tttgttcttc tggttcattt cctatgattg tttggctgct gaccgacgcg tcaggatgtg 4080  
 cagggcatgc ggggaatcag gaccggacac gggataatct catctaccta tacggagatc 4140  
 gcggtcctcg ccatgaggat cgcgacaggc gcgtcgaggg ggcaggaaca cccttgcgga 4200  
 ttgacattct tgggtggtgt tcggttgtgt cggtagttgt tgttgacgat gaggataaat 4260  
 aaaaatgacc ttgtttttgt tctgttttct cttgttggga atcgtcgact ttgaattctt 4320  
 cgagttatcg gaaagctgag gtaccctaat gtctgtagct tttttctttt taccctcttg 4380  
 tttatcatct gcgattcgtg gtaggttagga gagggaaatg ataatccgag attaaggaaa 4440  
 ggagaagata aaaaataaaa aaaaaataat aaaacagaag ccgaccggcc gccgaccgt 4500  
 tccccaggac cagcctacga ggaatggata acgcggtggc gacggcagcg gtggtggcgc 4560  
 tgggggtggc ggcatggta ctgctgatgg tagtcgggac ggaggagagg cgatgcatac 4620  
 atacacgcgt gcatgctgca tgggtggatg gtacggccgg gagacgcgga agagaaactc 4680  
 acataaaaag gtgacaaaaa gagcgggtga aaaaagaaaa cgagattcga ccagacagaa 4740  
 gagaaggacc ggggcttggc gacccttcca cgactgctgt tgtcatctcg gctccccgt 4800  
 cttctcccgg ccacgggagg ctaagtcacc gccgttctcc ccatccgtcc gagcgccgac 4860  
 cgaccagccg gccgattcgc ccgccggggc ttctggagaa cgccggggca gcagcgatct 4920  
 ggggaagccg ctaaaccctt gcgtttttat atggtagctc tgccgagcgc gggctgacgc 4980  
 gttgagtaag cggaaagacg tgtgtgacga aaaggggtcc catggtatct cacgtgacga 5040  
 tgaggagatg cggtttggag cacatacggg ttagaaaaag ggagttgtcg tgacaagggc 5100  
 tgagggacct ctgtctccat gtgtgtataa aaagcaaggc acgttcataa tgtaaaaaag 5160  
 aacacgttgt aaacaagcta ttgctgtatc attcggctga ctatgcttca ttcggactga 5220  
 ttttcttttc ctaacggcgt aacttaaagt gattaacgta tgatatttgt tccccagagt 5280  
 tatactatag tcatcatcct aaaattcaga tataaatgaa cacatgtcgt atgggattat 5340  
 taagaaaccg aaactctcca cagttcacca tcttctctgt cattcaaccg atgaccact 5400  
 ccgtacaacg aatcagtctg ctgtgtcaca ctgcaaaacta ctagcgacgt atgcaaaaca 5460  
 cttgaaacac gggctgttgt attgacgacc gttgtacat tactagtcac attgcataga 5520  
 gaccatccac cgtcatccca tctttccac ccgatggaaa accgtcttct atcatcaact 5580

atggtaagat ttcgaccctg cgaggatattc agtttcccca tatccataac ctggatttta 5640  
tcattaaacc ccaatattaa acactttttt agtaccctcc caccaccaa aaaatgtgac 5700  
tggaccggtt cctagcagct ctgggagcca tgttcagggtt gaaccacagc tacagcgaaa 5760  
ccgagtccag tgaccggtaa ccacgtccag cccctgcgta tgtaccagtc caagcacgtc 5820  
cggtcattgt tctacacagg aaatctaact aggtcaacgc aattttattc caccgttacg 5880  
cagaatacta acaaacaac acacaaattt aacgaattac acgtagtta ttacatgaaa 5940  
actgtaagaa caccaattca ctaagcgata caacatttag ctgacttcca agtgccacac 6000  
atcaccactg tattcatcca tgttttcacc gaaccaacga gacagatcga agaagccaga 6060  
atctcccgac tttaaattac ataaatccaa cgtattatga ccacagctcg acacacaaat 6120  
agttgcgta ctattcacag tagcattacc tatacccgta acgttgaca accactgatc 6180  
accattgtta ccaaaaacgg ttttcactt agttgtcaac ggatctttcc catgctaat 6240  
ggtaaaata ctaccagtc tcgcttttag ctattacga gtattatccg catccacata 6300  
tatcaacgtc atagctaggc acgctataag taccctccc ccacaatgga atgttgccaa 6360  
accggttctt tcccgttata gccatagcgt tcccaggcaa aagcaaacgc caaacctaat 6420  
gcagtgaaaa gcgcttcag ccagaaccag cttatgtacc agccacaatc acatccggtt 6480  
attgtttcca caggaaatcc taccaggcaa agccccgctt gttttgttcc tgaccatctt 6540  
gttttagcaat tcgtaaaactg tcagcctagc gacgtccggt tagatcaaaa gtcacgtata 6600  
tagcgacgct gtttccacc gtttcccctg cccgccgtt ccgaacaacc caccgggtt 6660  
cagacaaccg accaccaaca gaaatataca cacagaccac cgggagtca gttaaagatt 6720  
tcatcagggt tattttggct gctgctagtc ttttgcttct tagaaaaaa ataccatata 6780  
agagaaataa tgatagtttg acaacacata tggcagggat ttcttcttca tcaataagat 6840  
atgcaattcc cccagggaga gactttcaac aattgaattt acaaaaacaa aattacatca 6900  
ggagaaagag aggatacatt aataaatata ttatatctgg tgtatatact gaatgctgct 6960  
ggttcataag gtaacgatgc tactttttt aattccaaga tggtttttct ttgttagtct 7020  
ttgttgact tgctggttcc taaaagttcg caaaaacgat tgtgtgaaga ttatgacgtt 7080  
ggttgactag ttcatgagat tctgctgtac gtgtgatggt tattcgctgg ttcgttctaa 7140



gatgagtatc gtactgtgtc tgcgatggtc gtctcttact ggcatttctt cggctgcctc 7200  
ttgttttcat gattgaaaag gaaaaaagga ctccgagggc gcggtcatct tttacttttc 7260  
ggttttctcg ttggcgggtc agaggtagtc agatcatgag actgtcgtgg tcgatgaaac 7320  
tgtgtctgct caagtgcagt ccatcttctg tacggagaaa aaagtcatcg ggataaataa 7380  
ggctatacaa ggcgttgtca agcgtgcggc tctaaacaaa ttaagcgata caaaattaca 7440  
gtgatacгаа taataaatta cccctcccc ctgtgggtccc cccgaggcga gagccacca 7500  
tcgtgtactc tcgcaccacc cacgaccaca gggggagacg ggacgaagag acgacgcaga 7560  
gcgccatctc ctctggagg ccggcggcgt taactgctac agctgcggcg gcgacgacag 7620  
ctgcgatttg tcggccgaca tgccgatggt atgggcggcg gcggcgggtg ccgcggcagc 7680  
ggggaggaga ggagagagaa gaggagcggg gcgtccgaag gcgaggatgg catggtctcg 7740  
ccggagcgcc cggcttttat ggaacactcg cgtccggttg ggtatcacc acaggaagat 7800  
gaatcacaac ttccaacca tcttgagacc cgagtaacgg tttacaggtc gcacgccagt 7860  
ctcagctaaa aacagcggac agtcccacgc tgtttctggt ttggctctct ccagtttctt 7920  
catcgcctc ttggtctccg tcatcatcgg aagaatacca cccgctctca tgcggcagtc 7980  
gatcagcctc gatgaacgag acgcggcgac gcctttctac ggccgactgg ttgtggtggt 8040  
gaaagaagag caccagcaat cccaggagga gcaacaagc ctcacatgtc caggaggtcg 8100  
gggagagggc ctgtcggaga tgaccgtgag gcatcacgta cggcagctga ggagaaacgg 8160  
agaagaaagg aaaattaccg tcaggggccc gggttcttat tagagaaaca gcacgtaggt 8220  
caggatccag atgctaattg caatcatgat gacgatgatc atgcaggcca agacgcggcg 8280  
caccaatgca gaatccaata gccgccgtgc ctccggttgg tggccggcgg catctagaga 8340  
catgatttgg gggggggacc ggcggcгаа aaagacagg agatggacag tgccacggtg 8400  
ttttgttatg attaggacat ggggaccgga agccgagaca gagtactaca ggggtttgaa 8460  
gggtaacgtg agggagatca tgtcatgggc gggctgaaga ccgtgcgggg aggatcgacg 8520  
tgtgcggtgc ttgtggaaca cgggttttta atatgtatcc gcgtgtaatg cacgcggtgt 8580  
gctttttagc actcggcttg ataagctacg tgaccgtctg cgtgaaacc atggtcgcca 8640  
ccaactgtct cgtgaaaaca gaaaataccc acctagcatg taagtgcaat ccgaatagta 8700

catctacca tggcagcaag tgccacgcga tgtgcaaatg ccgggtcaca gaaccatta 8760  
ccatgctagg cgcatactcg gcctggggcg cgggctcggt cgtggccacg ctgatagtcc 8820  
tgctggtggt cttcttcgta atttacgcgc gcgaggagga gaaaaacaac acgggcaccg 8880  
aggtagatca atgtctggcc tatcggagcc tgacacgcaa aaagctggaa caacacgcgg 8940  
ctaaaaagca gaacatctac gaacggattc cataccgacc ctccagacag aaagataact 9000  
ccccgttgat cgaaccgacg ggcacagacg acgaagagga cgaggacgac gacgtttaac 9060  
gaggaagacg agaacgtggt ttgcacatg cagacctaca gcaactccct cacgcttgtc 9120  
atagtacagt cgctgttttt attcacagct cagggaagtt tatcgaatgc cgtcgaacca 9180  
atcaaaaaac ccctaaagct cgccaactac cgcgccactt gcgaaaaccg tacacgcacg 9240  
ctggttacca ggcttaacac tagccatcac agcgtagtct ggcaacgtta tgatatctac 9300  
agcagatata tgcgtcgtat gccgccactt tgcatcatta cagacgccta taaagaaacc 9360  
acgcgtcagg gtggcgcaac ttccacgtgc acgcgcaaaa atctcacgct gtacaatctt 9420  
acggttaaag atacgggagt ctaccttcta caggatcagt ataccggcga tgtcgaagct 9480  
ttctacctca tcatccacc acgcagcttc tgccgagcct tggaaacgcg tcgatgcttt 9540  
tatccgggac caggcagagt cgggtgtggtc acggattccc aagaggcaga ccgagcaatt 9600  
atctcggatt taaaacgcca gtgggtccggc ctctcactcc attgcgcctg ggtttcggga 9660  
ctgatgatct ttgttggcgc actggtcatc tgctttctgc gatcgcaacg aatcggagaa 9720  
caggacgttg aacatctgcg gacggacctg gatacggaac ctttgttgtt gacggtggac 9780  
gggaatttgg aataaaagat gcgtaacacc tgtcgaagat gcgataactt tacatacagg 9840  
caaacagtgt atacaattat agtattttgt atgttgcata aagttacatg caacagtact 9900  
gctaacagta ctgcatccat tacgctatcc aacactgcct ctaccacttt tgtaaccaac 9960  
atatattcaa ctccgaataa caacacatca acgacgccac acacatctgt cacctcacia 10020  
gcgtcaacca ttggcaacat caccaacgtt acctccgact tgagtacttt cacaaccgta 10080  
tattctacat tcaatacatc atttgccaat atatctaata cggctgtcac tacagaattg 10140  
atltcaaca ataccaacac tatctcatct ttaccaacg taacagcaaa cgctacatca 10200  
tcttataaca caacaatcac cgtaactgtc acgtcagatg aaacttcgca caacgtatcc 10260

## 038595

actaataatg cacttataag cacaccatgg cctacaaatt gcagcgccac aacatacacc 10320  
 acgtacaacc ttactaactc ttccaacgct tgtcacacag agacaacaat catacgtttc 10380  
 aaggaaacca atacaacagg aatagaaggg agtaatgtca ccataaaggg taattctacg 10440  
 tgggactgtc tttcagtcgc ctggatacga cattacaata gatccacaca cggacatcat 10500  
 ctaggttatc gtaagaacgc acatacccaa tcttggattt ggctacgcat ccttacctct 10560  
 cacactgtat gtcattctca acatgaaaga ccttcactgt accatgactt atgtcgttcg 10620  
 tgcaacaaca cagaattaca tctgtacgat ctaaataatca ccaattccgg caggtacagc 10680  
 agacgttgtt ttaaagaaaa ttacttcaca ggacatcacg aagatgaaaa tttctaccta 10740  
 ttagtaacac caaaaaatca tactgaagct attaatgcta ctttcgtttg ccctagatac 10800  
 aacaccgata tcgaaaatga agatagagag aaaggaagtc aacatactaa caatacacat 10860  
 caccacaac gtaatctcta tcatagctcg caaagaagcc gcaccgtatg gaccatcgtg 10920  
 ttggtttgta tggcctgcat agttctgttt tttgcacgac gagcctttaa caaaaagtat 10980  
 catatgttac aagacaccgt cagtgaatca gaattcattg ttcgatatca cccagaacat 11040  
 gaagattgag ctacgtttcc gggcagacat cttatgaagc tgaacaataa actaaaacat 11100  
 tctgtaagac tcagcgttca aaggaatatt aatgccatt gagcgaaac taatattgca 11160  
 atggactggc gatctacggg tacgtggacg atactaatgt ccgctgtgtc agaaagctgc 11220  
 aatcaaacct gttcttgtca atgtccctgt agtactaccg ttaactattc aactagtact 11280  
 gagacagcca catcaacata cagtacaaca gttatcagca ataaaagcac ttcagaatct 11340  
 ataaattgct ctactgcaac tacaccagca aacaccgttt ctacaaaacc gtcggaaca 11400  
 accacacaga tatccacaac gacgaacaca aacgttgaga ctaccacatg taccaacacc 11460  
 accacgaccg ttacttgtga tggtttcaat tatacagtcc ataaaagatg cgatcgcagt 11520  
 tacgaggtaa tcaacgtaac aggatacgtt ggtagcaaca taactctaaa aaaatgcaat 11580  
 cagactgaga aatggcacia tgtagactgg attcattatg agtaccaccac gcataaaatg 11640  
 tgcaatttag gcaactatca ccaaacaca ccacggcacg acatatgttt tgactgcaac 11700  
 gacacctccc taactatcta caacttaacc acaaaaaacg ctggaaaata taccaggcgt 11760  
 caccgtgata acggtcaaga agaaaattac tacgtaacgg tgtaattgg agacacaacg 11820

## 038595

ttattcactc ttggcacatg ccctgtaaga tataaagaat ctacgaacac tgaaaacacc	11880
attggaagta gcatcataga aaccattgag aaagctaaca ttcccctggg aattcatgct	11940
gtatgggcag gcgtagtggt atcagtggcg cttatagcgt tgtacatggg tagccatcgc	12000
attcccaaaa agccgcatta caccaaacct cccaaatatg atccagatga attttgact	12060
aaggcttaac atgctgatca ataaactttt ttaaccaat aacatgtctc cgTTTTTTTT	12120
tgtaaacaac ctatgatata aagcgttata ttcagtcggt actaaacaaa aaaacatggg	12180
catgcaatgc aacactaaat tgttattgcc agtcgcacta ataccggttg caatcatcct	12240
aattggtact ctagtgccga tacttttaca tgaacaaaaa aaggcgtttt actggcgact	12300
ttttctgcaa agtcaacatg tagaagcacc cattacagta acgcagggag acacagtcta	12360
cctagacgct agcaataatc cctgtaatta ttccagcttt tggtagcacg gtaattgcga	12420
actttgtgga tggaacggat atctacgcaa tgttacacat tactacacaa acacatcgtg	12480
ttccccgcaa ttcactctga taaacgaac taaaggtctg cagttatata atgtaacatt	12540
aaacgattca ggcgcttata ctgaacacgt ttacgaatgt gaccttctgt gtaacattac	12600
tactaataac gaatatgaaa tactcaatta ttttgataac tgtaactaca ccataaatag	12660
caccaagcat attatcaccg tgggtctctc acgtcattct aaacaaacaa attcccacgt	12720
atccactcac gctggttggg cagtcgccgt ggtgacggta attatgatct acgttctgat	12780
ccactttaac gtccccgcaa ctctgagaca caaactacga actagaaaca acgtaaatcg	12840
catagcgtga ttataaagta tcgacgctaa tttctccaag ataaaatttg attactccgt	12900
gcagttctca aaaactgtaa ggccccgctt ttccactccg tcatgaagga tcgcaataga	12960
atactgctat gtatcatctt tatttgatt atgtgcctca tttgtattta ctttaaactg	13020
cgttgtgttt ttactccgtc tccagacaaa gcagatctgc gagggaatt tccctcgta	13080
cccccgta ttggcatata gtgcgctgca tgagaacacg cgtgacacat agcgtacccc	13140
tggacggtac agtttatgat aacgtaattc agggaaagta tacattcata ccaacatggt	13200
atcacataac acacagattt tctgcgtggt ttataaaaga gcgtctcgaa gcagcttgag	13260
ccacactacg gtccagatga cgagcgtaat taaaaatatg ccgcgagta ttcgaaagcc	13320
gtactgagcg tgcgaggcgg gtaggggtgcc gaacgacgga tatgctcgtg tgcattctt	13380

gactataagg atcgcgaccg agtcttcggc catggtaaac gtcaccctgt gtggctggta 13440  
tgtagcgtat ccggtttgga attgttctgc tccagctcgg gggatagtga ggaattctca 13500  
agggatacgg gacccaatga ctggataaga gaagggtttt tccccgtaag atgatcctcg 13560  
tatcacatga ggtctggata tgtataaatg aagagtgaaa taggcacagg gaatcagatg 13620  
ccagcctcgt gatgcagccg ctggttctct cggcgaagaa actgtctgtct ttgctgactt 13680  
gcaaatacat cccgccttaa gcgatgagtc tataaagcac cgttgcccga gtacggtaaa 13740  
agtgaccggg attgtagaac gtcctttttt tttgtttttg catcgtttat cgtcactact 13800  
agtgcaatat tttgattgta aggctgaaag agtatcgta tgatgcttag aacgtggaga 13860  
ttattacaga tggactgct tgccgcgtac tgttattatg tttttgcgac ttgttcaatc 13920  
agcacgacga ctgctcctgt ggaatggaag tctcccgacc gtcagattcc caagaatatt 13980  
acctgcgcta attactcagg gaccgtcaac ggcaacgta catttcgagg tcttcagaac 14040  
aaaacggaag actttttgta ctggttgta ggatggggtc ataagtccat ttgttcgttc 14100  
ttcccgaac tccagggtaa ctatgacgaa caacattaca gatatgaagt agcgaacctg 14160  
acgtataact gcacctataa ccgcttgacg ttgctgaatc tgacgacgga aacagcggga 14220  
aagtactatt tcaaaagga agatgcgaat ttcaccttct attactcttg ttacaacttg 14280  
accgtgtcct aaagatcgca cgtgaagttt cacagagccg cgtggctgta gctattgtgt 14340  
ttacgttgct tttgaaatgt taagcgtccc tacggcgcta acatgtttct aggctactct 14400  
gactgtgtag atccccgct tgctgtgat cgtgtatcta gatcacgctt aaagctcatg 14460  
ttgtcttttg tgtggttggg cggtttgcgt ttctatgatt gtgcccggtt cgagtcctgc 14520  
tgttacgaca tcaccgaggc ggagagtaac aaggctatat caaggacga agcagcattc 14580  
acctccagcg tgagcaccg tacaccgtcc ctggcgatcg cgcctcctcc tgaccgatcg 14640  
atgctgttgt cgcgagagga agaactcgtt ccgtggagtc gtctcatcat cactaagcag 14700  
ttctacggag gcctgatttt ccacaccacc tgggtcaccg gcttcgtcct gctaggactc 14760  
ttgacgcttt tcgccagcct gtttcgcgta ccgcaatcca tctgtcgttt ctgcatagac 14820  
cgtctccggg acatcggccg tcctctgaaa taccgctatc aacgtcttgt cgctaccgtg 14880  
tagctagtta gccagctgtg tgtagtgttt tgcttttgca tatttgtttt cagtcagaga 14940

gtctgaaacg ggggtgggagg gacttttgcg ggtagtgcac gctaagatga acgggtgggc 15000  
tgggggtgtgc ttgataactc actgtttgaa tacgcgctca cgcacatatg tagcactcaa 15060  
catgttagct tttgcccgca cgccccgggg cgtgccgagc tgccctttta ataaagtctg 15120  
ggtttccaga tacgcgctgg ttctgatttt gatggtttgt gcctctgaaa gctctacgag 15180  
ctgggccgtg acatccaatg gactgcctaa ctgtagcacg gtaactagaa cagcgggtca 15240  
agacgctgaa ttgcacggtc cggcaccggt aagctgtaat gtgaccaggt ggggacgtta 15300  
cgagaatgga agcacacccg tgttatgggt cactttacgg ggatcaagca tgcgagtctc 15360  
attaggacac cgtgtagcgt ttggctgttc ttggaaaaca ttttttattt ataacgtttc 15420  
tgaaagtagc ggtggcactt actatcaaaa aggttacaac tgcaccgaca aacatataac 15480  
actatcttgt ttcaacttaa cgggtggttc tcgagcgggt caaagcaca ccaccgtaat 15540  
gacaccacg ctggttacia actccacatt cagtgtgtca cttgttccgt tgagactgac 15600  
gacaaattcc agcgcgtttg gacacgctat ttatcaacga caacagcgtg ttgaaaacgg 15660  
gacgttatcc aagaacataa ctaacttggc attcacctat ggacagctggg gcgttgcgat 15720  
gctgctgttt gccgccgtga tgggtcctgt tgatttgggt ttgcctcaat cggcttggcg 15780  
acgctggcga agccacgtgg acgatgaaga acgtggttt ttaatgtagg aaataaaagg 15840  
cagtttgagc atgactgttt ccaaaccgta acgtggtaaa taaatcatgg cttccgacgt 15900  
gggttctcat cctctgacgg ttacacgatt tcgctgcaga gtgcattatg tgtacaataa 15960  
actgttgatt ttaactttgt ttgccccgt gattctggaa tccgtcatct acgtgtccgg 16020  
gccacaggga gggaacgtta ccctggtatc caacttact tcaaacatca gcgcacggtg 16080  
gttccgctgg gacggcaacg atagccatct catttgcttt taaaaacgtg gagagggtct 16140  
ttctacgcc tatgtgggtt taagcctaag ttgtgcggct aaccaaata ccatcttcaa 16200  
cctcacgttg aacgactccg gtcgttacgg agcagaagg tttacgagaa gcggcgaaaa 16260  
tgaaacgttc ctgtggtata atttgaccgt gaaacccaaa ctttggaac ctactccagc 16320  
tagtaacgta acaaccatcg tcacgacgac atcgacgatg atcgacgcga aaagtaacgt 16380  
tacaggaac gccagtttag caccacaatt acgtgccgtc gctggattct ccaatcagac 16440  
gcctttggaa aacaacacgc acctggcctt ggtagggtgt gttgtgttt tagttctgat 16500

## 038595

agttgtttgc attatgggggt ggtggaaatt gttgtgtggt aaaccagagt tatagtaatg 16560  
 tgctttttat cagggagaag gttttgtgcc aacaatgact agcccgggac tatctgcgtc 16620  
 agaaaattat gacggaaatt atgaattcac ggaaaccgcc aatacaacgc gtacaaatag 16680  
 aagtgactgg acaacgttag aaaccagtgc attgctattg aaaaacacgg agactgcagt 16740  
 gaacctcagc aacgcgacta cggatcatccc acaacctgta gaatacccgg ctggggaagt 16800  
 acaatatcaa agaacggcaa cgcattattc ttggatgcta atcattgtca tcatttctcat 16860  
 cttttttatt atcatctgtc tacgagcacc tcgaaaaatc taccatcact ggaaagacag 16920  
 taaacagtac ggacaagtgt ttatgacaga cacggaactg tgacagtgat gtctaagcgt 16980  
 ttgcaggat ttccatggat aacaatttta ttttacacat caaaatccca gtattggaac 17040  
 tatatggcaa taccatgtac ccctacagtt ggatacggca gtcataatat tagcttgcat 17100  
 ccgcttaata actcattatt tcaagacgat gtttttgaat ggtacataga caaaccaatg 17160  
 gttacaagtt atgtctttat caaagtaatg aacgcacaaa atccaatcta gactctcaa 17220  
 atattgtgtg gcaatgcaca gataatcgta cactaattct catgaactta accacaacat 17280  
 acagtagaaa ctattatfff caatccttta aatatctcgg acgaggagta ccaaaaccga 17340  
 ataacttgtg ttataacggt agtgtacact ttaccacca aacacattgc catacaacta 17400  
 catcatccct gtatccacct acatctgtac acgattcatt agaaatatca cagtcatcca 17460  
 cctcaaccaa cttcacacat accgcgggcc actacgccac cggtaacggt gaagcacaac 17520  
 acgacactac cactccacat acaatgtgga tcataccctt agttatcgtt ataacaatca 17580  
 tcgttttaac ttgtttcaaa ttccccaga aagcttggaa taaattcaca caatacagat 17640  
 acagcgggat gctcggccc gcttaaagaa tcaacgcaa ggaaaccaa acgtaaaaag 17700  
 aatagatatg tacgtttatt tttcagctca ctgtttgaat accgtaaaca taatgacgta 17760  
 catatacgtg gttatacaac aggtgtttgt gttatgcggc gactgattaa ccatatcgtg 17820  
 aacatgatc ttttccgatg gtccgtcgtg accgcaatga tttttacag atattccgaa 17880  
 acctgatgg aggtcactgt cagagtaggt gatccagtta ccctcggtag tggacatggt 17940  
 tatcatccag gtagggataa cagggtaatg atcctctaga gtcgacctgc aggcacgcaa 18000  
 gcttgagtat tctatagtct cacctaaata gcttgg 18036

## 038595

<210> 12  
 <211> 2310  
 <212> ДНК  
 <213> последовательность gp49 LKA1

<400> 12  
 atggcgcaaa caccagtagc atgggcccgc tacgtaggagc acggcgtaga ggatacgttc 60  
 caagtcacat tcccgtacca gaagcagcaa gaggtgtttg tgactgtggg cggcgatccg 120  
 gcagctttca cattcatctc ggcaggttgg attcaactgg cagcgggtccc ggtaaatggg 180  
 gccgcaatcc gtgtacggcg cagcactgag gcattcgagc ctcggcacga gttcgccaac 240  
 ggctgtccat tactgccgagc attcatagac gagaataata cccagttctt gtacactgta 300  
 caagaggcag tgaatgagac acatggcatt gcttccgaag cgctgagtgt cgcagaggag 360  
 gccagaggca ttgcgcaggc ggcatcggat aaagtggatg ctgccacat tgactccgca 420  
 caccagttgc gtctagacct cggcgaccgc gcgaaggggc ctgggctgct aggctacgac 480  
 cgagacgtaa gttatccggt cgggtcggtc ggtcaaagcc tacagtttct ggaaatgggt 540  
 cgggtcacac cagcgcaatt tggcgccggt ggtgatggcg ccagccacc cctctctgag 600  
 cgatacgcaa ctctagcgga agctcagact gtctatccgc atgcagtcgc actctccgac 660  
 gaaatagact gggccgcatt gcaagctgcc gtggattcag gggcacctgt acacataccg 720  
 tctggggact atcagataaa tagggggatt agcagtagcg gctctctaca gattgcgggt 780  
 gatggcgcta catctattat acgcccgact gctgcgttca ctggtacatc ggtcctcagt 840  
 tgtgtgggga gcttagttgc cttgccgaat atatcctccg tgtcggctgg gtcccctaacc 900  
 attgactttg ccagcaccac taatcttgta gcgggggatg tattcatcat ctacaaccgc 960  
 actgatagca gcttctcggg atttcggagc agctatcgcg caggagagtt ctgtgaggtc 1020  
 agggcgggtt ctgggaacac cgtgacaatc cgttccgcac tctatgccgc atacgacggg 1080  
 gctactgttg ctatttaca agtagtctct ggtgtagttg atatagctag catccaaatc 1140  
 gttggcggga cagtcccaat gaatggactg ttagtggagg ctgtcgtttc accgcgcgtc 1200  
 gatgacgtga cggtcaccct tgcaacaac gccggtgtgt attttgcggc ctgctatgac 1260  
 gctaagatca caaacgtaa tatatcgaac atcggcgacg gtggcgatga ctatggaatc 1320  
 atctttggga actgtcacga cgggtgggca gacaactgta aagtctacgc taggcgacat 1380



## 038595

gccatcgcca cgggcggcga tgcagaagta ggctcgttc cggtcgtaa tgtgcgatg 1440  
cgtaactgca cacttaggaa tgatattacc tctggtacac actgcgcaga cttccacggt 1500  
aacgccgagg attgcagcta cgaaaactgc acaatctacg gtggtgcaac ttggcagggg 1560  
aaggatatca gctacagaca ctgtacaatc actaacgctg cgggtggttg gattgttata 1620  
tccgctgaga ttcttggtgg tacattcctt ctcgaccaat gcacattgta cacaaccggc 1680  
gatccgcagc ctggtaacgg tggggttata gatgtaggtg ggaactccgc agtcctcact 1740  
acaaatacaa cgcaaccctg taacttcctt atacaaggcg gcagtctgcg agcgcccagc 1800  
ttaagtacgt ctagttacct actgcgcgca cgtcttgagg gtagtacagt tccagtaaac 1860  
atacagtaca gcggacaggc tattgatgta ggctctctgg gcaaggctact acaactcgat 1920  
attacctcgg gcagtacctc tcctgagtat ttgatcgtgg agaatttagc ggggttgcca 1980  
tctggcatca cgctggcgtc tgctgctggt ggtttcgcaa gtgccccgat gcgtatgcct 2040  
gtgctgggtg gtagggttca agtaactacg gcaaccaacg cgagtagcgt tactgctcca 2100  
gtaacgttca ggtacattta tcctaaggcc ccaaccgtcc aggtcacaaa gacggacagg 2160  
agctacgccg gtaacagggt cggcgttgct atcgccaatc cgacctctgc gtctggggcg 2220  
acgttgggtc tgttcacgga cgacgggaca aactttagct cagccgttac taaccagttg 2280  
aactggcagg caggtattta tgaggtgtaa 2310

<210> 13  
<211> 1956  
<212> ДНК  
<213> Phikmvlikevirus NTUH-K2044-K1-1

<220>  
<221> misc\_feature  
<223> gp34 NTUH-K2044-K1-1

<400> 13  
atggccctga tccggctcgt ggcgcccag cgcggttca gcgacctggc cagcatggtc 60  
gcctatccga acttccaggt gcaggacaag atcaccctgc tgggctcggc cggcggcgac 120  
ttcaccttca ccaccaccgc gtcggtggtg gacaacggca ccgtgttcgc cgtgcccggc 180  
ggctatctcc tgcggaagtt cgtcggcccg gcgtatagct cgtggttcag caactggacc 240

gggatcgtca cgttcatgag cgcgccgaac cggcacctgg tggtaggacac cgtgctgcag 300  
 gccacgagcg tgctgaacat caagagcaac agcacgctgg aattcacgga cacgggcccgc 360  
 atcctgcccg acgccccgt ggccccccag gtgctgaaca tcaccggctc cgcgccctcg 420  
 gtgttcgtgc ccctcggccg cgacgccgcc gcggggctga aggtgatcac cgtggcccgc 480  
 ggcgctgt cgcggtgaa aggcacctac ctctatctgc gctccaacia gctgtgcgac 540  
 ggccggccga acacctatgg cgtcaagatc agccaaatcc gtaaggtagt cggcgtgagc 600  
 accagcgggg gcgtgacgtc catccgcctc gacaaagccc tgcaactataa ctactacctc 660  
 tcggatgccg ccgaagtggg catcccgacc atggtggaga acgtcacctt ggtgagcccg 720  
 tacatcaacg agttcggcta cgacgacctg aaccgcttct tcaccagcgg catctccgcg 780  
 aacttcgagg ccgacctgca catccaggac ggctcatca tcggcaacia gcgtccgggc 840  
 gcctccgaca tcgagggccc cagcgccatc aagttcaaca actgcgtgga tagcaccgtg 900  
 aagggcacct gcttctataa tatcggctgg tacggcgtgg aggtcctcgg ctgctcggag 960  
 gacaccgagg tgcacgacat ccacgccatg gacgtgcgcc atgccatctc cctgaactgg 1020  
 caaagcaccg ccgacggcga taagtggggc gaaccgatcg agttcctggg cgtgaactgt 1080  
 gaggcgtaca gcaccacca ggccggcttc gacaccacg acatcgggaa gcgtgtcaaa 1140  
 ttctcggct gcgtgtccta cgacagcgcg gatgacggct tccaggcccc caccaacggc 1200  
 gtggagtacc tcaactgccg cgcctaccgc gccgccatgg acggcttcgc ctcgaacacg 1260  
 ggctgcct tcccgatcta ccgcgaatgc ctggcctacg acaacgtgcg cagcgggttc 1320  
 aactgcagct acggcggcgg gtatgtgtac gactgcgagg cgcacggcag ccagaacggc 1380  
 gtccgcatca acggcggccc ggtcaaaggc gggcgctaca cccgcaactc gtcgagccac 1440  
 atcttcgtga cgaaagatgt ggcggaaacc gcccaaacca gcctcgagat cgacggcgtc 1500  
 tccatcgggt acgacggcac cggccgcgcc gtgtacttcc acggcaccgt gggcatcgat 1560  
 ccgacgctcg tgagcatgtc caacaacgac atgaccggcc acggcctggt ctgggccctg 1620  
 ctgtccggct ataccgtgca gccgaccccg ccgcgcatgt cgcgcaacct gctcgacgat 1680  
 accggcatcc gcggcgtcgc gaccctggtc gcgggcgaag cgaccgtcaa tgcccgcgtc 1740  
 cgcgggaact tcggcagcgt ggccaacagc ttcaagtggg tgtcggaggt gaagctgacg 1800

## 038595

cgctcacgt tcccgtcgtc ggccggcgcc ctcacggta cagcgtcgc ccaaaaccag 1860  
gacgtgccga ccccaacc ggacctgaac agcttcgtca tccgcagcag caacgccgcc 1920  
gacgtgtccc aagtcgctg ggaggtctac ctgtga 1956

<210> 14  
<211> 2184  
<212> ДНК  
<213> Т7-подобный Рр15

<220>  
<221> misc\_feature  
<223> добавлена последовательность gp44 Рр15

<400> 14  
atggcacgaa ctatcgtcca gaacgcccta acaggcggac aacaggactt cgaggtaacct 60  
ttcactaca tcttgacgcg cttcgttaag cttaccctga tcggtgacgg taaccgacaa 120  
gagctgggtcc tcggtaccga cttccggttc atcggctctc gcaccgttcg cactaacgtc 180  
ttctggggac cagcgcaggg gtatacctcc atcgagatcc gacgagttac cagcgcttct 240  
gatcgtcgcg tagagtcttc ggacgggtcc atcctgaccg caggtgatct gaacatcgcc 300  
cagcttcagg ccatccacat tgccgaagaa gcgcgagact ctgccactga gaacctgagc 360  
ccagatgctg atggcaacta cgatgcacgt ggtgcgcgca tttacaacct cggtgacgct 420  
gttcagccga aggatgcggt caaccggtag actcttgacc tcgctatcgc agccgctctg 480  
gccatgaata ccggcaacc gaacaacgcc cagaacatct cgtacacccc taacgggcct 540  
ggtcagtcga tccgaagtgt tgaaggccgt ctgcgggatg ctgtgttcgt ctcggactac 600  
atgaccactc cacgtgatgg agttaccagt aaccagcagg acctcgaaaa ggcaactcgt 660  
gcggcgaacg ctaaagggtg cgacctattc tggcctgacg acatcccgtt cttctccacg 720  
tccccgctgg cactgatcca cgcggcttac catgttggac gtgggtgtcat caacgcgaac 780  
ggtacgctgt tctacgtgaa cccgaagaac ggccaacaca acaggctaca cgtgtctccc 840  
gggggcaccg gggatggctt ggcagctggc cgcccactgg ggaccatctg gactgcactc 900  
gcggccctta acatgcgagc cccactgacc acgcgctggc ctttgagat gaccgctggc 960  
gcctataatg aagccgttac acttccgaac tacctgacca gctgtaacga ctacttggcg 1020

## 038595

tttaactggc cgaacaccgg tcaggaacgt atggagccca ctgctaccc atcagctctc	1080
gacggcacag gccagaccgg cctcacaggt ttccacactg gcatcggcaa ccgcattacc	1140
atcaacaacg tgtgcatgtc caactggtac gacactgcgc tgactcctac ccaacaggtg	1200
cgaagagcgt tcggttagg tgcgtattcg actgcctacg tggccaactg cgcgttcatt	1260
tacaacggca tcgcgagcgt gtctgtgctg cccggtggca ctgctatcgt aaccggtggc	1320
atcgtcgatg gtggcggtt cggcctcgac aacctggcg gtcgcctgtc cctgacggca	1380
accaagagca attatacgca ggtccggaac tgcctcgaat atggactgta ctcgaagcat	1440
gacgcatcga ccgtaatgga caacaccgag ttccgcaact gcgtaatca ccctgcggct	1500
gttgctatg gtgctgcaat cttcgcgtac aagttcaact gttctgttga cactcgtggg	1560
gtcaagttct acggcaaca catcggccag cactgccgtg gcggtatcac ctcggacaat	1620
ccggcgatc cggacatcta cggtagcggc gcagatgcta ataagcgtct attcctgtgc	1680
accggtggtg gctctgacga catccagttc tacgaagctc ggcgctcat ggacatcacg	1740
aagcgcactg gtggcggtc aactactgcc agcgtatcgt cgctgctact ggctgccgtt	1800
gcgtctgtcc gtaagggcta ctttgcgcac aacgatcagg tgatccggat gaccctgatg	1860
ttccgcgcta caggctcggc tggcatcttc acgccacct tgcgcacacc tctggggact	1920
atccctctgg gtagcttcag ggtcgcacg ggacagtacg gcgagatcaa gttgaccatt	1980
cgacctactc tgacatctga tggctcata gtcgggttct cctgcatcaa cgccgtgcag	2040
aatcttgggt cctctgttgg tcaaatcatc gtcagcggca ccgtagacct ccgcaccgtc	2100
gaccagctgg tcgagatgtg gggctattcg gaagctggtg gcaccgcttc gtacattcaa	2160
ggcctgatcg agctggtcgg gtga	2184

<210> 15  
 <211> 1089  
 <212> ДНК  
 <213> *Aggregatibacter actinomycetemcomitans*

<220>  
 <221> misc\_feature  
 <223> добавлена последовательность dspB

## 038595

<400> 15  
 atgaactgtt gcgtcaaggg caattccatc tacccccaga agacctccac caagcagacc 60  
 ggctgatgc tcgatatcgc ccggcatttc tacagccccg aggtgatcaa gagcttcatc 120  
 gatac gatca gcctgagcgg cggcaacttc ctccacctgc acttctcgga ccatgaaaac 180  
 tatgccatcg agtcgcacct gctcaaccag cgggcggaga acgccgtcca ggggaaggat 240  
 ggcatctaca tcaatccgta caccgggaaa ccgttctga gctaccgcca gctggacgac 300  
 atcaaggcct acgccaaggc caagggcacg gaactgatcc cggagctgga cagcccgaac 360  
 catatgacgg ccatcttcaa actgggtccag aaggaccgcg gcgtcaagta cctgcagggg 420  
 ctgaaatccc gccaggtgga cgacgagatc gacatcacca acgccgatag catcaccttc 480  
 atgcagagcc tgatgagcga ggtcatcgat atcttcggcg acacgagcca gcacttccac 540  
 atcggcggcg acgaattcgg ctactccgtc gagagcaacc acgagttcat cacctacgcc 600  
 aacaagctgt cgtacttctt ggagaagaag gggctcaaga cccgcatgtg gaacgacggc 660  
 ctcatcaaga acaccttca gcagatcaat cccaacatcg aatcacgta ctggtcgtac 720  
 gacggcgaca cccaggataa gaacgaagcg gccgagcgcc gcgacatgcg cgtgagcctg 780  
 ccggagctgc tggcgaaggg cttcaccgtg ctgaaactaca acagctacta cctctacatc 840  
 gtgccgaagg cgagcccgac gttctcgag gacgccgct tcgccgcaa agacgtgatc 900  
 aagaactggg atctgggcgt ctgggatggc cggaacacca agaaccgct gcagaacacc 960  
 catgagatcg ccggggcggc gctgtcgatc tggggcgagg atgcgaaggc gctcaaggac 1020  
 gagacgatcc agaagaacac caaaagcctg ctcgaggccg tcatccacaa gaccaacggc 1080  
 gacgagtga 1089

<210> 16  
 <211> 69  
 <212> ДНК  
 <213> Staphylococcus aureus

<220>  
 <221> misc\_feature  
 <223> добавлена последовательность SaPSMa3

<400> 16  
 atggagtctg tggcgaagct cttcaagttc ttcaaggacc tgctcgggaa gttcctgggg 60

aataactga	69
<210> 17	
<211> 135	
<212> ДНК	
<213> Staphylococcus aureus	
<220>	
<221> misc_feature	
<223> добавлена последовательность SaPAMb2	
<400> 17	
atgaccggcc tggccgaggc gatcgcgaat accgtccagg cggcccagca gcacgacagc	60
gtcaagctgg gcacctcgat cgtggacatc gtcgccaacg gcgtgggcct gctgggcaaa	120
ctcttcggct tctga	135
<210> 18	
<211> 69	
<212> ДНК	
<213> Staphylococcus epidermidis	
<220>	
<221> misc_feature	
<223> добавлена последовательность SePSMa	
<400> 18	
atggcggacg tcatcgccaa gatcgtcgag atcgtgaagg gcctgatcga ccagttcacc	60
cagaagtga	69
<210> 19	
<211> 228	
<212> ДНК	
<213> Levivirus MS2	
<220>	
<221> misc_feature	
<223> добавлена последовательность L MS2	
<400> 19	
atggagaccs ggttcccgsa gcagtcccag saaaccccgg ccagcaccaa ccgccgccgc	60
cccttcaagc acgaggacta cccgtgccgc cggcagcagc gcagctccac cctgtacgtg	120

## 038595

ctgatcttcc tggcgatcct cctgagcaag ttcaccaacc agctgctgct gtcctgctg 180  
gaggcgggtca tccggaccgt caccaccctg cagcagctgc tgacctga 228

<210> 20  
<211> 165  
<212> ДНК  
<213> Levivirus PRR1

<220>  
<221> misc\_feature  
<223> добавлена последовательность L PRR1

<400> 20  
atgtgcaagg tgtctactaa ggtagactct aaactgactg agtcagttgg acaactcacc 60  
ataaggagct acctatggct acggaatata ctagcattag caggacttct tttcgtaatc 120  
cttcttgсga cсаатсатт atccatcgct atctacagtc cgtaa 165

<210> 21  
<211> 52  
<212> ДНК  
<213> Phikmvlikevirus LUZ19

<220>  
<221> misc\_feature  
<223> промотор gp32 (P32) LUZ19

<400> 21  
cgaccctgcc ctactccggc cttaaaccса catccaaaag agagagaatc gc 52

<210> 22  
<211> 96  
<212> ДНК  
<213> Phikmvlikevirus LUZ19

<220>  
<221> misc\_feature  
<223> терминатор gp32 (T32) LUZ19

<400> 22  
tgccacgaaa ccccgactt cgggtgtgggg tttcttcaaa gcctaacgac ccgсgсagat 60  
tcctgcgtg ggtttttgсg ctttaggaga aaccct 96

## 038595

<210> 23  
 <211> 204  
 <212> ДНК  
 <213> Phikmvlikevirus LUZ19

<220>  
 <221> misc\_feature  
 <223> область gp7 LUZ19 дикого типа

<400> 23  
 tacaagggtgg tggcacccag ctcggcgga ggtatcattg tgctggcgac caagcagacg 60  
 ccggcgctag cccaagcagc cgctgtactg cacagcatga accctgcgca gtatcccgca 120  
 ggttcggcta tcctcaacac ggcctggaag tgcccggcc tgggagtggg cgagtacgtc 180  
 aagctcgtcc aaggggagga ggac 204

<210> 24  
 <211> 321  
 <212> ДНК  
 <213> Phikmvlikevirus LUZ19

<220>  
 <221> misc\_feature  
 <223> область gp18 LUZ19 дикого типа

<400> 24  
 gaatgссаас сгаагаагаа сgcatgatcc gctgtttact ggcggatata cagagccac 60  
 tggacctgct gttccccggc ctccgtacca aggcccatat ggacccgcaa gcagaggaac 120  
 tgtcgattcg aattgactac gaccatgcga agctgggccg tatgggattc tgccacgcgg 180  
 tatccctata tcaactgtcc atatatggcc gcgaggggat ggtccgctac ctgatgcagg 240  
 agattccccg ccgctgtctg gaaggctctg tgggtcaaggc gcagcagtac agccaaagca 300  
 actggtacag caaatgacga c 321

<210> 25  
 <211> 225  
 <212> ДНК  
 <213> Phikmvlikevirus LKD16



## 038595

<220>  
 <221> misc\_feature  
 <223> gp49 и межгенная область gp48-gp49 LUZ19 дикого типа  
 <400> 25  
 ggggacacca tgagcaaaagc caaactacga gtcacgcgac acaccccgga gctggagtca 60  
 gtgctaanaag cattgctgac cgcacactac gctatcgagg acctgctcaa cgaggccgtg 120  
 gctagcaagg tgctaaactc ccgctgggac tggctccgag tcggcgagta tgtcgaactg 180  
 ttcaaccgca cgcaatcccg cgtggccggg ttgattcccg agtag 225

<210> 26  
 <211> 345  
 <212> ДНК  
 <213> Phikmvlikevirus LKD16

<220>  
 <221> misc\_feature  
 <223> ген gp18 LKD16 дикого типа  
 <400> 26  
 gtgacgagta caactgaaca cgagcgacc ctgacgctgc tgctccaaga catccacggg 60  
 ccgctgaatc tgctgttccc aggtatccgg gtgaagggtg aggaggcgtg cctcggatac 120  
 ttgggctaca gggagcgggg ctattgggag ctgacgctcc aggtggacta cgaccacccg 180  
 aagcttgggc acctccgcta cagtcaggcc gtgcccggagt acgtgctgat caacgaccgc 240  
 gacagcatca tcaagtacct gatggaagca gtcacctcggc aggtactaga gggcatgctc 300  
 aataaggccc aggaattcgt аассаагаас tggatttccc tatga 345

<210> 27  
 <211> 4269  
 <212> ДНК  
 <213> Искусственная последовательность

<220>  
 <223> ген, кодирующий NLS-FLAG-CAS9-His  
 <400> 27  
 atgcccсаага аааагсggаа ggtcggcgac tacaaggatg acgatgacaа gttggagcct 60  
 ggagagaagc cctacaаatg ccctgagtgс ggaaagagct tcagccaatc tggagccttg 120  
 acccgcatc аасгаасгса tacacgagac аагаагтас ссатсgggct ggacatcggg 180

acgaactccg tgggatgggc cgtgatcaca gacgaataca aggtgccttc caagaagttc	240
aaggtgctgg ggaacacgga cagacactcc atcaagaaga acctcatcgg ggccttgctc	300
ttcgactccg gagaaaccgc cgaagcaacg cgattgaaaa gaaccgccag aagacgatac	360
acacgacgga agaaccgcat ctgctacctc caggagatct tcagcaacga gatggccaag	420
gtggacgact cgttcittca tgcctggag gagagcttcc tggtaggagga agacaagaaa	480
catgagcgcc acccgatctt cgggaacatc gtggacgaag tggcctacca cgagaaatac	540
cccacgatct accacttgcg caagaaactc gtggactcca cggacaaaagc ggacttgagg	600
ttgatctact tggccttggc ccacatgatc aaatttcggg gccacttctt gatcgagggc	660
gacttgaatc ccgacaatc cgacgtggac aagctcttca tccagctggt gcagacctac	720
aaccagctct tcgaggagaa ccccatcaat gcctccggag tggacgcaa agccatcttg	780
tcccccgat tgtccaaatc cagacgcttg gagaactga tcgcacaact tcctggcgag	840
aagaagaacg gcctcttcgg caacttgatc gcgctgtcgc tgggattgac gcctaacttc	900
aagtccaact tcgacttggc cgaggacgcc aagttgcaac tgtccaagga cacctacgac	960
gacgacctcg acaacctgct ggcccaaatt ggcgaccaat acgaggactt gtttttggcg	1020
gccaaagaact tgagcgacgc catcttgttg agcgacatct tgcgcgtgaa tacggagatc	1080
accaaagccc ctttgtccgc ctctatgatc aagcggtagc acgagcacca ccaagacttg	1140
accctgttga aagccctcgt gcggcaacaa ttgcccgaga agtacaagga gatcttcttc	1200
gaccagtcca agaacgggta cgccggctac atcgacggag gagcctcca agaagagttc	1260
tacaagttca tcaagcccat cctggagaag atggacggca ccgaggagt gctcgtgaag	1320
ctgaaccgcg aagacttggt gcgaaaacag cggacgttcg acaatggcag catccccac	1380
caaatccatt tgggagagtt gcacgccatc ttgcgacggc aagaggactt ctaccgttc	1440
ctgaaggaca accgcgagaa aatcgagaag atcctgacgt tcagaatccc ctactacgtg	1500
ggacccttgg cccgaggcaa ttcccgttt gcatggatga cgcgcaaaag cgaagagacg	1560
atcaccctcct ggaacttcga agaagtggtc gacaaaggag catccgcaca gagcttcatc	1620
gagcgaatga cgaacttcga caagaacctg cccaacgaga aggtgttgcc caagcattcg	1680
ctgctgtacg agtacttcac ggtgtacaac gagctgacca aggtgaagta cgtgaccgag	1740

ggcatgcgca aaccgcggtt cctgtcggga gagcaaaaga aggccattgt ggacctgctg	1800
ttcaagacca accggaaggt gaccgtgaaa cagctgaaag aggactactt caagaagatc	1860
gagtgccttcg actccgtgga gatctccggc gtggaggacc gattcaatgc ctcttggga	1920
acctaccatg acctcctgaa gatcatcaag gacaaggact tcctggacaa cgaggagaac	1980
gaggacatcc tggaggacat cgtgctgacc ctgacctgt tcgaggaccg agagatgac	2040
gaggaacggt tgaaaacgta cggccacttg ttcgacgaca aggtgatgaa gcagctgaaa	2100
cgccgccgct acaccgatg gggacgattg agccgcaaac tgattaatgg aattcgcgac	2160
aagcaatccg gaaagacat cctggacttc ctgaagtccg acgggttcgc caaccgcaac	2220
ttcatgcagc tcatccacga cgactccttg accttcaagg aggacatcca gaaggcccaa	2280
gtgtccggac aaggagactc cttgcacgag cacatcgcca atttggccgg atccccgca	2340
atcaaaaaag gcatcttgca aaccgtgaaa gtggtcgacg aactggtgaa ggtgatggga	2400
cggcacaagc ccgagaacat cgtgatcgaa atggcccgcg agaaccaaac cacccaaaaa	2460
ggacagaaga actcccgaga gcgcatgaag cggatcgaag agggcatcaa ggagttggc	2520
tcccagatcc tgaaggagca tcccgaggag aataccaat tgcaaacga gaagctctac	2580
ctctactacc tccagaacgg gcgggacatg tacgtcgacc aagagctgga catcaaccgc	2640
ctctccgact acgatgtgga tcatattgtg cccagagct tcctcaagga cgacagcatc	2700
gacaacaagg tcctgacgcg cagcgacaag aaccggggca agtctgaca tgtgccttcc	2760
gaagaagtcg tgaagaagat gaagaactac tggcggcagc tgctcaacgc caagctcatc	2820
acccaacgga agttcgacaa cctgaccaag gccgagagag gaggattgtc cgagttggac	2880
aaagccggtc tcattaaacg ccaactcgtg gagaccgcc agatcacgaa gcacgtggcc	2940
caaatcttgg actccggat gaacacgaaa tacgacgaga atgacaagct gatccgag	3000
gtgaaggatgac tcacgctgaa gtccaagctg gtgagcgact tccggaagga cttccagttc	3060
tacaaggtgc gggagatcaa caactacat cacgccatg acgcctacct gaacgccgtg	3120
gtcggaaaccg ccctgatcaa gaaatacccc aagctggagt ccgaattcgt gtacggagat	3180
tacaaggtct acgacgtgcg gaagatgatc gcgaagtccg agcaggagat cggcaaagcc	3240
accgccaagt acttctttta ctccaacatc atgaacttct tcaagaccga gatcacgctc	3300

## 038595

gccaacggcg agatccgcaa gcgccccctg atcgagacca acggcgagac gggagagatt 3360  
gtgtgggaca aaggaagaga ttttgccaca gtgcgcaagg tgctgtccat gcctcaggtg 3420  
aacatcgtga agaagaccga ggtgcaaaca ggagggtttt ccaaagagtc cattttgcct 3480  
aagaggaatt ccgacaagct catcgcccgc aagaaggact gggaccccaa gaagtacggg 3540  
ggcttcgact cccccacggt ggcctactcc gtgttggtgg tggccaaagt ggagaaaggg 3600  
aagagcaaga agctgaaatc cgtgaaggag ttgctcggaa tcacgatcat ggaacgatcg 3660  
tcgttcgaga aaaaccccat cgacttcctc gaagccaaag ggtacaaaga ggtgaagaag 3720  
gacctgatca tcaagctgcc caagtactcc ctgttcgagc tggagaacgg ccgcaagcgg 3780  
atgctggcct ccgccgggga actgcagaaa gggaacgaat tggccttgcc ctccaaatac 3840  
gtgaacttcc tctacttggc ctcccattac gaaaagctca aaggatcccc tgaggacaat 3900  
gagcagaagc aactcttctg ggaacaacac aagcactacc tggacgagat catcgagcag 3960  
atcagcgagt tctccaagcg cgtgatcctc gccgacgcca acctggacaa ggtgctctcc 4020  
gcctacaaca agcaccgga caagcctatc cgcgagcaag ccgagaatat cattcacctg 4080  
tttaccctga cgaatttggg agccccctgc gcctttaaat actttgacac caccatcgac 4140  
cgcaaaagat acacctccac caaggaagtc ttggacgcca ccctcatcca ccagtccatc 4200  
acgggcctct acgagacgcg catcgacctc tcccaattgg gcggcgacca tcatcaccac 4260  
caccactaa 4269

<210> 28  
<211> 507  
<212> ДНК  
<213> Inovirus M13MP18

<220>  
<221> misc\_feature  
<223> заменена область M13MP18 дикого типа

<400> 28  
atgacatga ttacgaattc gagctcggta cccggggatc ctctagagtc gacctgcagg 60  
catgcaagct tggcactggc cgtcgtttta caacgtcgtg actgggaaaa ccctggcggt 120  
accactta atcgccttgc agcacatccc cttttcgcca gctggcgtaa tagcgaagag 180

## 038595

gccccgaccg atcgcccttc ccaacagttg cgcagcctga atggcgaatg gcgctttgcc 240  
 tggtttccgg caccagaagc ggtgccggaag agctggctgg agtgcgatct tcctgaggcc 300  
 gatacggctg tcgtcccctc aaactggcag atgcacggtt acgatgcgcc catctacacc 360  
 aacgtaacct atcccattac ggtcaatccg ccgtttgttc ccacggagaa tccgacgggt 420  
 tgttactcgc tcacatttaa tgttgatgaa agctggctac aggaaggcca gacgcaatt 480  
 atttttgatg gcgttcctat tggtaa 507

<210> 29  
 <211> 792  
 <212> ДНК  
 <213> Искусственная последовательность

<220>  
 <223> Неизвестный

<220>  
 <221> misc\_feature  
 <223> последовательность Paprika

<220>  
 <221> misc\_feature  
 <223> коммерчески доступный из ДНК2.0

<400> 29  
 atggtgtcaa agggagaaga actgatcaaa gagaatatga ggatgaaact ctacatggaa 60  
 ggaactgtga acaaccacca tttcaagtgc acgagcagg gtgaaggga accttacgaa 120  
 ggtaccaga ccatcggat taaggctgc gaaggaggac cactcccctt cgcattcgac 180  
 atcctggcca cttccttcat gtacgggtcg cgcactttca tcaagtacc aaaagggatc 240  
 cccgacttct tcaagcagtc ctttccggag ggattcactt gggaacgcgt cactagatac 300  
 gaggatggcg gagtggcac cgtgatgcaa gacacctctt tggaagatgg atgcctggtg 360  
 taccacgtgc aagtcagagg agtgaacttt ccgagcaatg ggccggtgat gcagaagaaa 420  
 accaagggtc gggaaccgaa caccgaaatg ctgtatccag cagacggagg cttggagggc 480  
 cggctcgaca tggctctgaa gcttgttgga ggaggacatc tgtcctgctc gttcgtgacg 540  
 acctaccgga gcaagaagcc ggcgaaaaac cttaatgac cgggatcca cgcggtgat 600  
 catcgcctgg aaaggctcga ggagtcagac aacgagatgt ttgtcgtgca acgagcac 660

038595

gccgtggccc gctactgtga tctcccttca aagctgggcc acaagctgaa ttccggcctc 720  
 cggtcgagag cccaggcttc gaattcagcc gtggacggaa ctgcggggccc tggttcgacc 780  
 ggaagccgat ga 792

<210> 30  
 <211> 294  
 <212> ДНК  
 <213> Lambdavirus лямбда

<220>  
 <221> misc\_feature  
 <223> фаг дикого типа E. coli

<400> 30  
 atggttcgtg сааасааасг саасгaggct ctacgaatcg agagtgcggt gcttaacaaa 60  
 atcgcaatgc ttggaactga gaagacagcg gaagctgtgg gcgttgataa gtcgcagatc 120  
 agcaggtgga agagggactg gattccaaag ttctcaatgc tgcttgctgt tcttgaatgg 180  
 ggggtcgttg acgacgacat ggctcgattg gcgcgacaag ttgctgcgat tctaccaat 240  
 аааааасгсс сggcggсаас сgagcgttct гаасааатсс агатggagtt ctga 294

<210> 31  
 <211> 1422  
 <212> белок  
 <213> Искусственная последовательность

<220>  
 <223> Неизвестный

<220>  
 <221> misc\_feature  
 <223> NLS-МЕЧЕННЫЙ-CAS9-His белок, транслируемый с SEQ ID NO:27

<400> 31

Met Pro Lys Lys Lys Arg Lys Val Gly Asp Tyr Lys Asp Asp Asp Asp  
 1 5 10 15

Lys Leu Glu Pro Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys  
 20 25 30

## 038595

Ser Phe Ser Gln Ser Gly Ala Leu Thr Arg His Gln Arg Thr His Thr  
 35 40 45

Arg Asp Lys Lys Tyr Ser Ile Gly Leu Asp Ile Gly Thr Asn Ser Val  
 50 55 60

Gly Trp Ala Val Ile Thr Asp Glu Tyr Lys Val Pro Ser Lys Lys Phe  
 65 70 75 80

Lys Val Leu Gly Asn Thr Asp Arg His Ser Ile Lys Lys Asn Leu Ile  
 85 90 95

Gly Ala Leu Leu Phe Asp Ser Gly Glu Thr Ala Glu Ala Thr Arg Leu  
 100 105 110

Lys Arg Thr Ala Arg Arg Arg Tyr Thr Arg Arg Lys Asn Arg Ile Cys  
 115 120 125

Tyr Leu Gln Glu Ile Phe Ser Asn Glu Met Ala Lys Val Asp Asp Ser  
 130 135 140

Phe Phe His Arg Leu Glu Glu Ser Phe Leu Val Glu Glu Asp Lys Lys  
 145 150 155 160 165

His Glu Arg His Pro Ile Phe Gly Asn Ile Val Asp Glu Val Ala Tyr  
 165 170 175

His Glu Lys Tyr Pro Thr Ile Tyr His Leu Arg Lys Lys Leu Val Asp  
 180 185 190

Ser Thr Asp Lys Ala Asp Leu Arg Leu Ile Tyr Leu Ala Leu Ala His  
 195 200 205

Met Ile Lys Phe Arg Gly His Phe Leu Ile Glu Gly Asp Leu Asn Pro  
 210 215 220

Asp Asn Ser Asp Val Asp Lys Leu Phe Ile Gln Leu Val Gln Thr Tyr  
 225 230 235 240

## 038595

Asn Gln Leu Phe Glu Glu Asn Pro Ile Asn Ala Ser Gly Val Asp Ala  
 245 250 255

Lys Ala Ile Leu Ser Ala Arg Leu Ser Lys Ser Arg Arg Leu Glu Asn  
 260 265 270

Leu Ile Ala Gln Leu Pro Gly Glu Lys Lys Asn Gly Leu Phe Gly Asn  
 275 280 285

Leu Ile Ala Leu Ser Leu Gly Leu Thr Pro Asn Phe Lys Ser Asn Phe  
 290 295 300

Asp Leu Ala Glu Asp Ala Lys Leu Gln Leu Ser Lys Asp Thr Tyr Asp  
 305 310 315 320

Asp Asp Leu Asp Asn Leu Leu Ala Gln Ile Gly Asp Gln Tyr Ala Asp  
 325 330 335

Leu Phe Leu Ala Ala Lys Asn Leu Ser Asp Ala Ile Leu Leu Ser Asp  
 340 345 350

Ile Leu Arg Val Asn Thr Glu Ile Thr Lys Ala Pro Leu Ser Ala Ser  
 355 360 365

Met Ile Lys Arg Tyr Asp Glu His His Gln Asp Leu Thr Leu Leu Lys  
 370 375 380

Ala Leu Val Arg Gln Gln Leu Pro Glu Lys Tyr Lys Glu Ile Phe Phe  
 385 390 395 400

Asp Gln Ser Lys Asn Gly Tyr Ala Gly Tyr Ile Asp Gly Gly Ala Ser  
 405 410 415

Gln Glu Glu Phe Tyr Lys Phe Ile Lys Pro Ile Leu Glu Lys Met Asp  
 420 425 430

Gly Thr Glu Glu Leu Leu Val Lys Leu Asn Arg Glu Asp Leu Leu Arg  
 435 440 445



## 038595

Lys Gln Arg Thr Phe Asp Asn Gly Ser Ile Pro His Gln Ile His Leu  
 450 455 460

Gly Glu Leu His Ala Ile Leu Arg Arg Gln Glu Asp Phe Tyr Pro Phe  
 465 470 475 480

Leu Lys Asp Asn Arg Glu Lys Ile Glu Lys Ile Leu Thr Phe Arg Ile  
 485 490 495

Pro Tyr Tyr Val Gly Pro Leu Ala Arg Gly Asn Ser Arg Phe Ala Trp  
 500 505 510

Met Thr Arg Lys Ser Glu Glu Thr Ile Thr Pro Trp Asn Phe Glu Glu  
 515 520 525

Val Val Asp Lys Gly Ala Ser Ala Gln Ser Phe Ile Glu Arg Met Thr  
 530 535 540

Asn Phe Asp Lys Asn Leu Pro Asn Glu Lys Val Leu Pro Lys His Ser  
 545 550 555 560

Leu Leu Tyr Glu Tyr Phe Thr Val Tyr Asn Glu Leu Thr Lys Val Lys  
 565 570 575

Tyr Val Thr Glu Gly Met Arg Lys Pro Ala Phe Leu Ser Gly Glu Gln  
 580 585 590

Lys Lys Ala Ile Val Asp Leu Leu Phe Lys Thr Asn Arg Lys Val Thr  
 595 600 605

Val Lys Gln Leu Lys Glu Asp Tyr Phe Lys Lys Ile Glu Cys Phe Asp  
 610 615 620

Ser Val Glu Ile Ser Gly Val Glu Asp Arg Phe Asn Ala Ser Leu Gly  
 625 630 635 640

Thr Tyr His Asp Leu Leu Lys Ile Ile Lys Asp Lys Asp Phe Leu Asp  
 645 650 655

## 038595

Asn Glu Glu Asn Glu Asp Ile Leu Glu Asp Ile Val Leu Thr Leu Thr  
 660 665 670

Leu Phe Glu Asp Arg Glu Met Ile Glu Glu Arg Leu Lys Thr Tyr Ala  
 675 680 685

His Leu Phe Asp Asp Lys Val Met Lys Gln Leu Lys Arg Arg Arg Tyr  
 690 695 700

Thr Gly Trp Gly Arg Leu Ser Arg Lys Leu Ile Asn Gly Ile Arg Asp  
 705 710 715 720

Lys Gln Ser Gly Lys Thr Ile Leu Asp Phe Leu Lys Ser Asp Gly Phe  
 725 730 735

Ala Asn Arg Asn Phe Met Gln Leu Ile His Asp Asp Ser Leu Thr Phe  
 740 745 750

Lys Glu Asp Ile Gln Lys Ala Gln Val Ser Gly Gln Gly Asp Ser Leu  
 755 760 765

His Glu His Ile Ala Asn Leu Ala Gly Ser Pro Ala Ile Lys Lys Gly  
 770 775 780

Ile Leu Gln Thr Val Lys Val Val Asp Glu Leu Val Lys Val Met Gly  
 785 790 795 800

Arg His Lys Pro Glu Asn Ile Val Ile Glu Met Ala Arg Glu Asn Gln  
 805 810 815

Thr Thr Gln Lys Gly Gln Lys Asn Ser Arg Glu Arg Met Lys Arg Ile  
 820 825 830

Glu Glu Gly Ile Lys Glu Leu Gly Ser Gln Ile Leu Lys Glu His Pro  
 835 840 845

Val Glu Asn Thr Gln Leu Gln Asn Glu Lys Leu Tyr Leu Tyr Tyr Leu  
 850 855 860

038595

Gln Asn Gly Arg Asp Met Tyr Val Asp Gln Glu Leu Asp Ile Asn Arg  
865 870 875 880

Leu Ser Asp Tyr Asp Val Asp His Ile Val Pro Gln Ser Phe Leu Lys  
885 890 895

Asp Asp Ser Ile Asp Asn Lys Val Leu Thr Arg Ser Asp Lys Asn Arg  
900 905 910

Gly Lys Ser Asp Asn Val Pro Ser Glu Glu Val Val Lys Lys Met Lys  
915 920 925

Asn Tyr Trp Arg Gln Leu Leu Asn Ala Lys Leu Ile Thr Gln Arg Lys  
930 935 940

Phe Asp Asn Leu Thr Lys Ala Glu Arg Gly Gly Leu Ser Glu Leu Asp  
945 950 955 960

Lys Ala Gly Phe Ile Lys Arg Gln Leu Val Glu Thr Arg Gln Ile Thr  
965 970 975

Lys His Val Ala Gln Ile Leu Asp Ser Arg Met Asn Thr Lys Tyr Asp  
980 985 990

Glu Asn Asp Lys Leu Ile Arg Glu Val Lys Val Ile Thr Leu Lys Ser  
995 1000 1005

Lys Leu Val Ser Asp Phe Arg Lys Asp Phe Gln Phe Tyr Lys Val  
1010 1015 1020

Arg Glu Ile Asn Asn Tyr His His Ala His Asp Ala Tyr Leu Asn  
1025 1030 1035

Ala Val Val Gly Thr Ala Leu Ile Lys Lys Tyr Pro Lys Leu Glu  
1040 1045 1050

Ser Glu Phe Val Tyr Gly Asp Tyr Lys Val Tyr Asp Val Arg Lys  
1055 1060 1065

## 038595

Met Ile Ala Lys Ser Glu Gln Glu Ile Gly Lys Ala Thr Ala Lys  
1070 1075 1080

Tyr Phe Phe Tyr Ser Asn Ile Met Asn Phe Phe Lys Thr Glu Ile  
1085 1090 1095

Thr Leu Ala Asn Gly Glu Ile Arg Lys Arg Pro Leu Ile Glu Thr  
1100 1105 1110

Asn Gly Glu Thr Gly Glu Ile Val Trp Asp Lys Gly Arg Asp Phe  
1115 1120 1125

Ala Thr Val Arg Lys Val Leu Ser Met Pro Gln Val Asn Ile Val  
1130 1135 1140

Lys Lys Thr Glu Val Gln Thr Gly Gly Phe Ser Lys Glu Ser Ile  
1145 1150 1155

Leu Pro Lys Arg Asn Ser Asp Lys Leu Ile Ala Arg Lys Lys Asp  
1160 1165 1170

Trp Asp Pro Lys Lys Tyr Gly Gly Phe Asp Ser Pro Thr Val Ala  
1175 1180 1185

Tyr Ser Val Leu Val Val Ala Lys Val Glu Lys Gly Lys Ser Lys  
1190 1195 1200

Lys Leu Lys Ser Val Lys Glu Leu Leu Gly Ile Thr Ile Met Glu  
1205 1210 1215

Arg Ser Ser Phe Glu Lys Asn Pro Ile Asp Phe Leu Glu Ala Lys  
1220 1225 1230

Gly Tyr Lys Glu Val Lys Lys Asp Leu Ile Ile Lys Leu Pro Lys  
1235 1240 1245

Tyr Ser Leu Phe Glu Leu Glu Asn Gly Arg Lys Arg Met Leu Ala  
1250 1255 1260

## 038595

Ser Ala Gly Glu Leu Gln Lys Gly Asn Glu Leu Ala Leu Pro Ser  
 1265 1270 1275

Lys Tyr Val Asn Phe Leu Tyr Leu Ala Ser His Tyr Glu Lys Leu  
 1280 1285 1290

Lys Gly Ser Pro Glu Asp Asn Glu Gln Lys Gln Leu Phe Val Glu  
 1295 1300 1305

Gln His Lys His Tyr Leu Asp Glu Ile Ile Glu Gln Ile Ser Glu  
 1310 1315 1320

Phe Ser Lys Arg Val Ile Leu Ala Asp Ala Asn Leu Asp Lys Val  
 1325 1330 1335

Leu Ser Ala Tyr Asn Lys His Arg Asp Lys Pro Ile Arg Glu Gln  
 1340 1345 1350

Ala Glu Asn Ile Ile His Leu Phe Thr Leu Thr Asn Leu Gly Ala  
 1355 1360 1365

Pro Ala Ala Phe Lys Tyr Phe Asp Thr Thr Ile Asp Arg Lys Arg  
 1370 1375 1380

Tyr Thr Ser Thr Lys Glu Val Leu Asp Ala Thr Leu Ile His Gln  
 1385 1390 1395

Ser Ile Thr Gly Leu Tyr Glu Thr Arg Ile Asp Leu Ser Gln Leu  
 1400 1405 1410

Gly Gly Asp His His His His His His  
 1415 1420

<210> 32  
 <211> 909  
 <212> ДНК  
 <213> Цитомегаловирус ЦМВЧ

<220>  
 <221> misc\_feature

## 038595

<223> пост-редактирование фрагмента RL14 ЦМВЧ

<400> 32

```

atggactggc gatttacggt tacgtggacg atactaatgt ccgcttgtc agaaagctgc      60
aatcaaacct gttcttgtca atgtccctgt agtactaccg ttaactattc aactagtact      120
gagacagcca catcaacata cagtacaaca gttatcagca ataaaagcac ttcagaatct      180
ataaattgct ctactgcaac tacaccagca aacaccgttt ctacaaaacc gtcggaaaca      240
accacacaga tatccacaac gacgaacaca aacgttgaga ctaccacatg taccaacacc      300
accacgaccg ttacttgtga tggtttcaat tatacagtcc ataaaagatg c gatcgcagt      360
tacgaggtaa tcaacgtaac aggatacgtt ggtagcaaca taactctaaa aaaatgcaat      420
cagactgaga aatggcaca tgtagactgg attcattatg agtaccacc gcataaaatg      480
tgcgaattag gcaactatca ccaaacaca ccacggcacg acatatgttt tgactgcaac      540
gacacctccc taactatcta caacttaacc acaaaaaacg ctggaaaata taccaggcgt      600
cacctgata acggtcaaga agaaaattac tacgtaacgg tgtaattgg agacacaacg      660
ttattcactc ttggcacatg ccctgtaaga tataaagaat ctacgaacac tgaaaacacc      720
attggaagta gcatcataga aaccattgag aaagctaaca ttcccctggg aattcatgct      780
gtatgggcag gcgtagtggt atcagtggcg cttatagcgt tgtacatggg tagccatcgc      840
attcccaaaa agccgatta caccaactt cccaatatg atccagatga attttgact      900
aaggcttaa                                     909

```

<210> 33

<211> 630

<212> ДНК

<213> Цитомегаловирус ЦМВЧ

<220>

<221> misc\_feature

<223> пост-редактирование RL13 ЦМВЧ

<400> 33

```

atggactggc gatttacggt tacgtggacc gttacttgtg atggttcaa ttatacagtc      60
cataaaagat gcgatcgcag ttacgaggta atcaacgtaa caggatacgt tggtagcaac      120
ataactctaa aaaaatgcaa tcagactgag aaatggcaca atgtagactg gattcattat      180

```

## 038595

gagtacccca cgcataaaat gtgcgaatta ggcaactatc accaaaccac accacggcac 240  
gacatatggt ttgactgcaa cgacacctcc ctaactatct acaacttaac cacaaaaaac 300  
gctggaaaat ataccaggcg tcaccgtgat aacggtaag aagaaaatta ctacgtaacg 360  
gtgttaattg gagacacaac gttattcact cttggcacat gccctgtaag atataaagaa 420  
tctacgaaca ctgaaaacac cattggaagt agcatcatag aaaccattga gaaagctaac 480  
attcccctgg gaattcatgc tgtatgggca ggcgtagtgg tatcagtggc gcttatagcg 540  
ttgtacatgg gtagccatcg cattcccaaa aagccgcatt acaccaaact tcccaaatat 600  
gatccagatg aattttggac taaggcttaa 630

<210> 34  
<211> 95  
<212> белок  
<213> Phikmvlikevirus LUZ19

<220>  
<221> misc\_feature  
<223> последовательность белка Gr13 LUZ19 дикого типа

<400> 34

Met Leu Ala Leu Gly Ala Phe Asp Leu Ser Gly Leu Met Val Gly Ser  
1 5 10 15

Cys Leu Val Val Gly Gly Glu Leu Lys Ala Leu Cys Val Asp Asp Arg  
20 25 30

His Ser Arg Gln Gly Ile Gly Ala Glu Leu Val Arg Ala Ala Glu Leu  
35 40 45

Ala Gly Ala Glu Tyr Leu Thr Cys Phe Glu Phe Leu Glu Pro Phe Tyr  
50 55 60

Ala Asp Leu Gly Trp Ser Thr Thr His Arg Glu Ala Asn Trp Thr Ala  
65 70 75 80

Gly Glu Pro Asp Val Leu His Met Arg Ala Pro Gly His Asp Val  
85 90 95

## 038595

<210> 35  
 <211> 251  
 <212> белок  
 <213> Phikmvlikevirus LUZ19

<220>  
 <221> misc\_feature  
 <223> последовательность белка Gp38 LUZ19 дикого типа

<400> 35

Met Ala Arg Phe Lys Asn Pro Glu Thr Ile His Val Ala Asp Gly Val  
 1 5 10 15

Glu Ala Val Phe Ser Leu Asp Phe Pro Phe Leu Arg Arg Glu Asp Val  
 20 25 30

Phe Val Gln Val Asp Lys Ile Leu Val Thr Asp Tyr Thr Trp Val Asp  
 35 40 45

Asp Thr Asn Ile Gln Leu Ala Val Val Pro Lys Lys Asp Gln Glu Val  
 50 55 60

Arg Ile Phe Arg Asp Thr Pro Ala Gln Val Pro Asp Thr Gln Phe Ser  
 65 70 75 80

Gln Asp Ile Pro Phe Leu Pro Arg Tyr Ile Asp Ala Asn Asn Lys Gln  
 85 90 95

Leu Leu Tyr Ala Val Gln Glu Gly Ile Asn Thr Ala Asn Leu Ala Leu  
 100 105 110

Asp Gly Val Leu Asp Ala Ile Arg Ile Ala Glu Glu Ala Arg Arg Leu  
 115 120 125

Ala Gln Glu Ala Leu Asp Ala Ala Asn Glu Ala Leu Arg Arg Ala Leu  
 130 135 140

Gly Phe Ala Glu Ile Arg Thr Val Thr Glu Asp Ser Asp Ile Asp Pro  
 145 150 155 160



## 038595

Ser Trp Arg Gly Tyr Trp Asn Arg Cys Ile Thr Ala Asp Lys Pro Leu  
 165 170 175

Thr Leu Thr Met Gln Met Glu Asp Pro Asp Ala Pro Trp Val Glu Phe  
 180 185 190

Ser Glu Val His Phe Glu Gln Ala Gly Val Arg Asp Leu Asn Ile Val  
 195 200 205

Ala Gly Pro Gly Val Thr Ile Asn Arg Leu Gln Asn Thr Thr Met Gln  
 210 215 220

Leu Tyr Gly Glu Asn Gly Val Cys Thr Leu Lys Arg Leu Gly Ala Asn  
 225 230 235 240

His Trp Ile Val Phe Gly Ala Met Glu Asp Glu  
 245 250

<210> 36  
 <211> 301  
 <212> белок  
 <213> Phikmvlikevirus LUZ19

<220>  
 <221> misc\_feature  
 <223> послеовательность белка Gr40 LUZ19 дикого типа

<400> 36

Met Phe Lys Thr Glu Val Lys Gly Arg Tyr Thr Leu Ile Arg Arg Lys  
 1 5 10 15

Ala Asp Gly Thr Pro Val Glu Thr Leu Glu Phe Asp Asn Ile Ile Thr  
 20 25 30

Asn Ala Gly Leu Asp Trp Ile Ala Ala Met Asp Thr Asp Leu Met Gly  
 35 40 45

Glu Pro Val Ala Val Ser Thr Ser Thr Ala Asp Pro Asn Pro Ser Ala  
 50 55 60

## 038595

Pro Ala Ile Pro Glu Val Val Gln Arg Thr Ser Ala Ser Ala Pro Gly  
65 70 75 80

Gly Gly Thr Thr Ser Gly Leu Asp Gly Glu Trp Leu Phe Trp Arg Arg  
85 90 95

Arg Trp Arg Phe Pro Gln Gly Thr Leu Ala Gly Gln Val Leu Ala Thr  
100 105 110

Val Gly Leu Ile Cys Asn Ser Asp Arg Arg Phe Glu Ser Asn Thr Gly  
115 120 125

Glu Leu Ile Pro Lys Asp Thr Pro Leu Ser Tyr Thr Arg Ile Lys Asp  
130 135 140

Ala Ala Gly Gln Pro Thr Thr Leu Val Val Ala Ala Asp Glu Ile Leu  
145 150 155 160

Asp Val Gln Tyr Glu Phe Arg Ser Arg Pro Val Gly Thr Ala Glu Ala  
165 170 175

Lys Phe Val Ile Ser Gly Val Glu Arg Thr Phe Arg Leu Ile Pro Lys  
180 185 190

Pro Phe Ala Asn Arg Ala Asn Leu Ser Gly Glu Arg Tyr Ile Phe Tyr  
195 200 205

Asn Thr Asn Pro Tyr Ile Asn Gly Lys Asp Ala Ser Gly Gly Asn Val  
210 215 220

Arg Asp Gly Gln Trp Gln Lys Lys Tyr Pro Lys Tyr Val Arg Gly Ser  
225 230 235 240

Tyr Lys Ala Gln Ile Thr Leu Leu Ala Gln Val Gln Asn Gly Asn Met  
245 250 255

Ala Gly Gly Ile Thr Gly Thr Glu Glu Leu Gln Ile Tyr Asn Gly Arg  
260 265 270

## 038595

Asn Tyr Val Leu Asp Ile Asn Pro Pro Val Val Lys Asn Asn Thr Gln  
 275 280 285

Glu Phe Thr Val Thr Leu Glu Phe Thr Val Ala Arg Ala  
 290 295 300

<210> 37  
 <211> 498  
 <212> белок  
 <213> Pseudomonas aeruginosa

<220>  
 <221> misc\_feature  
 <223> последовательность белка PyoS5

<400> 37

Met Ser Asn Asp Asn Glu Val Pro Gly Ser Met Val Ile Val Ala Gln  
 1 5 10 15

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

Ala Val Ala Gly Asn Met Phe Gly Asp Leu Ile Gln Arg Glu Ile Tyr  
 35 40 45

Leu Gln Lys Asn Ile Tyr Tyr Pro Val Arg Ser Ile Phe Glu Gln Gly  
 50 55 60

Thr Lys Glu Lys Lys Glu Ile Asn Lys Lys Val Ser Asp Gln Val Asp  
 65 70 75 80

Gly Leu Leu Lys Gln Ile Thr Gln Gly Lys Arg Glu Ala Thr Arg Gln  
 85 90 95

Glu Arg Val Asp Val Met Ser Ala Val Leu His Lys Met Glu Ser Asp  
 100 105 110

Leu Glu Gly Tyr Lys Lys Thr Phe Thr Lys Gly Pro Phe Ile Asp Tyr  
 115 120 125

## 038595

Glu Lys Gln Ser Ser Leu Ser Ile Tyr Glu Ala Trp Val Lys Ile Trp  
 130 135 140

Glu Lys Asn Ser Trp Glu Glu Arg Lys Lys Tyr Pro Phe Gln Gln Leu  
 145 150 155 160

Val Arg Asp Glu Leu Glu Arg Ala Val Ala Tyr Tyr Lys Gln Asp Ser  
 165 170 175

Leu Ser Glu Ala Val Lys Val Leu Arg Gln Glu Leu Asn Lys Gln Lys  
 180 185 190

Ala Leu Lys Glu Lys Glu Asp Leu Ser Gln Leu Glu Arg Asp Tyr Arg  
 195 200 205

Thr Arg Lys Ala Asn Leu Glu Met Lys Val Gln Ser Glu Leu Asp Gln  
 210 215 220

Ala Gly Ser Ala Leu Pro Pro Leu Val Ser Pro Thr Pro Glu Gln Trp  
 225 230 235 240

Leu Glu Arg Ala Thr Arg Leu Val Thr Gln Ala Ile Ala Asp Lys Lys  
 245 250 255

Gln Leu Gln Thr Thr Asn Asn Thr Leu Ile Lys Asn Ser Pro Thr Pro  
 260 265 270

Leu Glu Lys Gln Lys Ala Ile Tyr Asn Gly Glu Leu Leu Val Asp Glu  
 275 280 285

Ile Ala Ser Leu Gln Ala Arg Leu Val Lys Leu Asn Ala Glu Thr Thr  
 290 295 300

Arg Arg Arg Thr Glu Ala Glu Arg Lys Ala Ala Glu Glu Gln Ala Leu  
 305 310 315 320

Gln Asp Ala Ile Lys Phe Thr Ala Asp Phe Tyr Lys Glu Val Thr Glu  
 325 330 335

## 038595

Lys Phe Gly Ala Arg Thr Ser Glu Met Ala Arg Gln Leu Ala Glu Gly  
 340 345 350

Ala Arg Gly Lys Asn Ile Arg Ser Ser Ala Glu Ala Ile Lys Ser Phe  
 355 360 365

Glu Lys His Lys Asp Ala Leu Asn Lys Lys Leu Ser Leu Lys Asp Arg  
 370 375 380

Gln Ala Ile Ala Lys Ala Phe Asp Ser Leu Asp Lys Gln Met Met Ala  
 385 390 395 400

Lys Ser Leu Glu Lys Phe Ser Lys Gly Phe Gly Val Val Gly Lys Ala  
 405 410 415

Ile Asp Ala Ala Ser Leu Tyr Gln Glu Phe Lys Ile Ser Thr Glu Thr  
 420 425 430

Gly Asp Trp Lys Pro Phe Phe Val Lys Ile Glu Thr Leu Ala Ala Gly  
 435 440 445

Ala Ala Ala Ser Trp Leu Val Gly Ile Ala Phe Ala Thr Ala Thr Ala  
 450 455 460

Thr Pro Ile Gly Ile Leu Gly Phe Ala Leu Val Met Ala Val Thr Gly  
 465 470 475 480

Ala Met Ile Asp Glu Asp Leu Leu Glu Lys Ala Asn Asn Leu Val Ile  
 485 490 495

Ser Ile

<210> 38  
 <211> 114  
 <212> белок  
 <213> Phikmvlikevirus LKD16

## 038595

<220>  
 <221> misc\_feature  
 <223> последовательность белка Gr18 LKD16

<400> 38

Met Arg Val Pro Thr Glu His Glu Arg Thr Leu Arg Cys Leu Leu Gln  
 1 5 10 15

Asp Ile His Gly Pro Leu Asn Leu Leu Phe Pro Gly Ile Arg Val Lys  
 20 25 30

Val Glu Glu Ala Cys Leu Gly Tyr Leu Gly Tyr Arg Glu Arg Gly Tyr  
 35 40 45

Trp Glu Leu Arg Leu Gln Val Asp Tyr Asp His Pro Lys Leu Gly His  
 50 55 60

Leu Arg Tyr Ser Gln Ala Val Pro Glu Tyr Val Leu Ile Asn Asp Arg  
 65 70 75 80

Asp Ser Ile Ile Lys Tyr Leu Met Glu Ala Val Pro Arg Gln Val Leu  
 85 90 95

Glu Gly Met Leu Asn Lys Ala Gln Glu Phe Val Thr Lys Asn Trp Tyr  
 100 105 110

Ser Leu

<210> 39  
 <211> 769  
 <212> белок  
 <213> Phikmvlikevirus LKA1

<220>  
 <221> misc\_feature  
 <223> последовательность белка Gr49 LKA1

<400> 39

Met Ala Gln Thr Pro Ser Thr Trp Ala Asp Tyr Val Gly Asp Gly Val  
 1 5 10 15

038595

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

038595

Ala Ala Leu Gln Ala Ala Val Asp Ser Gly Ala Pro Val His Ile Pro  
 225 230 235 240

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

Gln Ile Ala Gly Asp Gly Ala Thr Ser Ile Ile Arg Pro Thr Ala Ala  
 260 265 270

Phe Thr Gly Thr Ser Val Leu Ser Cys Val Gly Ser Leu Val Ala Leu  
 275 280 285

Pro Asn Ile Ser Ser Val Ser Ala Gly Ser Leu Thr Ile Asp Phe Ala  
 290 295 300

Ser Thr Pro Asn Leu Val Ala Gly Asp Val Phe Ile Ile Tyr Asn Pro  
 305 310 315 320

Thr Asp Ser Ser Phe Ser Gly Phe Arg Thr Ser Tyr Arg Ala Gly Glu  
 325 330 335

Phe Cys Glu Val Arg Ala Val Ser Gly Asn Thr Val Thr Ile Arg Ser  
 340 345 350

Ala Leu Tyr Ala Ala Tyr Asp Gly Ala Thr Val Ala Ile Tyr Lys Val  
 355 360 365

Val Ser Gly Val Val Asp Ile Ala Ser Ile Gln Ile Val Gly Gly Thr  
 370 375 380

Val Pro Met Asn Gly Leu Leu Val Glu Ala Val Val Ser Pro Arg Val  
 385 390 395 400

Asp Asp Val Thr Val Thr Leu Ala Asn Asn Ala Gly Val Tyr Phe Ala  
 405 410 415

Arg Cys Tyr Asp Ala Lys Ile Thr Asn Ser Asn Ile Ser Asn Ile Gly  
 420 425 430



038595

Asp Gly Gly Asp Asp Tyr Gly Ile Ile Phe Gly Asn Cys His Asp Gly  
 435 440 445

Gly Ala Asp Asn Cys Lys Val Tyr Ala Arg Arg His Ala Ile Ala Thr  
 450 455 460

Gly Gly Asp Ala Glu Val Gly Cys Val Pro Val Arg Asn Val Arg Met  
 465 470 475 480

Arg Asn Cys Thr Leu Arg Asn Asp Ile Thr Ser Gly Thr His Cys Ala  
 485 490 495

Asp Phe His Gly Asn Ala Glu Asp Cys Ser Tyr Glu Asn Cys Thr Ile  
 500 505 510

Tyr Gly Gly Ala Thr Trp Gln Gly Lys Asp Ile Ser Tyr Arg His Cys  
 515 520 525

Thr Ile Thr Asn Ala Ser Gly Gly Trp Ile Val Ile Ser Ala Glu Ile  
 530 535 540

Leu Gly Gly Thr Phe Leu Leu Asp Gln Cys Thr Leu Tyr Thr Thr Gly  
 545 550 555 560

Asp Pro Gln Pro Gly Asn Arg Gly Val Ile Asp Val Gly Gly Asn Ser  
 565 570 575

Ala Val Leu Thr Thr Asn Thr Thr Gln Pro Cys Asn Phe Leu Ile Gln  
 580 585 590

Gly Gly Ser Leu Arg Ala Pro Ser Leu Ser Thr Ser Ser Tyr Leu Leu  
 595 600 605

Arg Ala Arg Leu Glu Gly Ser Thr Val Pro Val Asn Ile Gln Tyr Ser  
 610 615 620

Gly Gln Ala Ile Asp Val Gly Ser Leu Gly Lys Val Leu Gln Leu Asp  
 625 630 635 640

038595

Ile Thr Ser Gly Ser Thr Ser Pro Glu Tyr Leu Ile Val Glu Asn Leu  
 645 650 655

Ala Gly Leu Pro Ser Gly Ile Thr Leu Ala Ser Ala Ala Gly Gly Phe  
 660 665 670

Ala Ser Ala Pro Met Arg Met Pro Val Leu Gly Gly Arg Val Gln Val  
 675 680 685

Thr Thr Ala Thr Asn Ala Ser Ser Val Thr Ala Pro Val Thr Phe Arg  
 690 695 700

Tyr Ile Tyr Pro Lys Ala Pro Thr Val Gln Val Thr Lys Thr Asp Arg  
 705 710 715 720

Ser Tyr Ala Gly Asn Arg Val Gly Val Ala Ile Ala Asn Pro Thr Ser  
 725 730 735

Ala Ser Gly Ala Thr Leu Gly Leu Phe Thr Asp Asp Gly Thr Asn Phe  
 740 745 750

Ser Ser Ala Val Thr Asn Gln Leu Asn Trp Gln Ala Gly Ile Tyr Glu  
 755 760 765

Val

- <210> 40
- <211> 651
- <212> белок
- <213> Phikmvlikevirus NTUH-K2044-K1-1

- <220>
- <221> misc\_feature
- <223> последовательность белка Gp34 NTUH-K2044-K1-1

<400> 40

Met Ala Leu Ile Arg Leu Val Ala Pro Glu Arg Val Phe Ser Asp Leu  
 1 5 10 15

## 038595

Ala Ser Met Val Ala Tyr Pro Asn Phe Gln Val Gln Asp Lys Ile Thr  
20 25 30

Leu Leu Gly Ser Ala Gly Gly Asp Phe Thr Phe Thr Thr Thr Ala Ser  
35 40 45

Val Val Asp Asn Gly Thr Val Phe Ala Val Pro Gly Gly Tyr Leu Leu  
50 55 60

Arg Lys Phe Val Gly Pro Ala Tyr Ser Ser Trp Phe Ser Asn Trp Thr  
65 70 75 80

Gly Ile Val Thr Phe Met Ser Ala Pro Asn Arg His Leu Val Val Asp  
85 90 95

Thr Val Leu Gln Ala Thr Ser Val Leu Asn Ile Lys Ser Asn Ser Thr  
100 105 110

Leu Glu Phe Thr Asp Thr Gly Arg Ile Leu Pro Asp Ala Ala Val Ala  
115 120 125

Arg Gln Val Leu Asn Ile Thr Gly Ser Ala Pro Ser Val Phe Val Pro  
130 135 140

Leu Ala Ala Asp Ala Ala Ala Gly Ser Lys Val Ile Thr Val Ala Ala  
145 150 155 160

Gly Ala Leu Ser Ala Val Lys Gly Thr Tyr Leu Tyr Leu Arg Ser Asn  
165 170 175

Lys Leu Cys Asp Gly Gly Pro Asn Thr Tyr Gly Val Lys Ile Ser Gln  
180 185 190

Ile Arg Lys Val Val Gly Val Ser Thr Ser Gly Gly Val Thr Ser Ile  
195 200 205

Arg Leu Asp Lys Ala Leu His Tyr Asn Tyr Tyr Leu Ser Asp Ala Ala  
210 215 220

## 038595

Glu Val Gly Ile Pro Thr Met Val Glu Asn Val Thr Leu Val Ser Pro  
 225 230 235 240

Tyr Ile Asn Glu Phe Gly Tyr Asp Asp Leu Asn Arg Phe Phe Thr Ser  
 245 250 255

Gly Ile Ser Ala Asn Phe Ala Ala Asp Leu His Ile Gln Asp Gly Val  
 260 265 270

Ile Ile Gly Asn Lys Arg Pro Gly Ala Ser Asp Ile Glu Gly Arg Ser  
 275 280 285

Ala Ile Lys Phe Asn Asn Cys Val Asp Ser Thr Val Lys Gly Thr Cys  
 290 295 300

Phe Tyr Asn Ile Gly Trp Tyr Gly Val Glu Val Leu Gly Cys Ser Glu  
 305 310 315 320

Asp Thr Glu Val His Asp Ile His Ala Met Asp Val Arg His Ala Ile  
 325 330 335

Ser Leu Asn Trp Gln Ser Thr Ala Asp Gly Asp Lys Trp Gly Glu Pro  
 340 345 350

Ile Glu Phe Leu Gly Val Asn Cys Glu Ala Tyr Ser Thr Thr Gln Ala  
 355 360 365

Gly Phe Asp Thr His Asp Ile Gly Lys Arg Val Lys Phe Val Arg Cys  
 370 375 380

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

Val Glu Tyr Leu Asn Cys Arg Ala Tyr Arg Ala Ala Met Asp Gly Phe  
 405 410 415

Ala Ser Asn Thr Gly Val Ala Phe Pro Ile Tyr Arg Glu Cys Leu Ala  
 420 425 430

## 038595

Tyr Asp Asn Val Arg Ser Gly Phe Asn Cys Ser Tyr Gly Gly Gly Tyr  
 435 440 445

Val Tyr Asp Cys Glu Ala His Gly Ser Gln Asn Gly Val Arg Ile Asn  
 450 455 460

Gly Gly Arg Val Lys Gly Gly Arg Tyr Thr Arg Asn Ser Ser Ser His  
 465 470 475 480

Ile Phe Val Thr Lys Asp Val Ala Glu Thr Ala Gln Thr Ser Leu Glu  
 485 490 495

Ile Asp Gly Val Ser Met Arg Tyr Asp Gly Thr Gly Arg Ala Val Tyr  
 500 505 510

Phe His Gly Thr Val Gly Ile Asp Pro Thr Leu Val Ser Met Ser Asn  
 515 520 525

Asn Asp Met Thr Gly His Gly Leu Phe Trp Ala Leu Leu Ser Gly Tyr  
 530 535 540

Thr Val Gln Pro Thr Pro Pro Arg Met Ser Arg Asn Leu Leu Asp Asp  
 545 550 555 560

Thr Gly Ile Arg Gly Val Ala Thr Leu Val Ala Gly Glu Ala Thr Val  
 565 570 575

Asn Ala Arg Val Arg Gly Asn Phe Gly Ser Val Ala Asn Ser Phe Lys  
 580 585 590

Trp Val Ser Glu Val Lys Leu Thr Arg Leu Thr Phe Pro Ser Ser Ala  
 595 600 605

Gly Ala Leu Thr Val Thr Ser Val Ala Gln Asn Gln Asp Val Pro Thr  
 610 615 620

Pro Asn Pro Asp Leu Asn Ser Phe Val Ile Arg Ser Ser Asn Ala Ala  
 625 630 635 640

## 038595

Asp Val Ser Gln Val Ala Trp Glu Val Tyr Leu  
645 650

<210> 41  
<211> 727  
<212> белок  
<213> T7-подобный Pp15

<220>  
<221> misc\_feature  
<223> последовательность белка Gr44 Pp15

<400> 41

Met Ala Arg Thr Ile Val Gln Asn Ala Leu Thr Gly Gly Gln Gln Asp  
1 5 10 15

Phe Glu Val Pro Phe Asp Tyr Ile Leu Gln Arg Phe Val Lys Leu Thr  
20 25 30

Leu Ile Gly Asp Gly Asn Arg Gln Glu Leu Val Leu Gly Thr Asp Phe  
35 40 45

Arg Phe Ile Gly Pro Arg Thr Val Arg Thr Asn Val Phe Trp Gly Pro  
50 55 60

Ala Gln Gly Tyr Thr Ser Ile Glu Ile Arg Arg Val Thr Ser Ala Ser  
65 70 75 80

Asp Arg Arg Val Glu Phe Ser Asp Gly Ser Ile Leu Thr Ala Gly Asp  
85 90 95

Leu Asn Ile Ala Gln Leu Gln Ala Ile His Ile Ala Glu Glu Ala Arg  
100 105 110

Asp Ser Ala Thr Glu Asn Leu Ser Pro Asp Ala Asp Gly Asn Tyr Asp  
115 120 125

Ala Arg Gly Ala Arg Ile Tyr Asn Leu Gly Asp Ala Val Gln Pro Lys  
130 135 140

## 038595

Asp Ala Val Asn Arg Tyr Thr Leu Asp Leu Ala Ile Ala Ala Ala Leu  
 145 150 155 160

Ala Met Asn Thr Gly Asn Pro Asn Asn Ala Gln Asn Ile Ser Tyr Thr  
 165 170 175

Pro Asn Gly Pro Gly Gln Ser Ile Arg Ser Val Glu Gly Arg Leu Arg  
 180 185 190

Asp Ala Val Phe Val Ser Asp Tyr Met Thr Thr Pro Arg Asp Gly Val  
 195 200 205

Thr Ser Asn Gln Gln Asp Leu Glu Lys Ala Leu Ala Ala Ala Asn Ala  
 210 215 220

Lys Gly Ala Asp Leu Phe Trp Pro Asp Asp Ile Pro Phe Phe Ser Thr  
 225 230 235 240

Ser Pro Leu Ala Leu Ile His Ala Val Tyr His Val Gly Arg Gly Val  
 245 250 255

Ile Asn Ala Asn Gly Thr Leu Phe Tyr Val Asn Pro Lys Asn Gly Gln  
 260 265 270

His Asn Arg Leu His Val Ser Pro Gly Gly Thr Gly Asp Gly Leu Ala  
 275 280 285

Ala Gly Arg Pro Leu Gly Thr Ile Trp Ser Ala Leu Ala Ala Leu Asn  
 290 295 300

Met Arg Ala Pro Leu Thr Thr Arg Trp Ser Leu Glu Met Thr Ala Gly  
 305 310 315 320

Ala Tyr Asn Glu Ala Val Thr Leu Pro Asn Tyr Leu Thr Ser Cys Asn  
 325 330 335

Asp Tyr Leu Ala Phe Asn Trp Pro Asn Thr Gly Gln Glu Arg Met Glu  
 340 345 350

038595

Pro Thr Ala Tyr Pro Ser Ala Leu Asp Gly Thr Gly Gln Thr Gly Leu  
 355 360 365

Thr Gly Phe His Thr Gly Ile Gly Asn Arg Ile Thr Ile Asn Asn Val  
 370 375 380

Cys Met Ser Asn Trp Tyr Asp Thr Ala Leu Thr Pro Thr Gln Gln Val  
 385 390 395 400

Arg Arg Ala Phe Val Val Gly Ala Tyr Ser Thr Ala Tyr Val Val Asn  
 405 410 415

Cys Ala Phe Ile Tyr Asn Gly Ile Ala Ser Val Ser Val Leu Pro Gly  
 420 425 430

Gly Thr Ala Ile Val Thr Gly Gly Ile Val Asp Gly Gly Arg Phe Gly  
 435 440 445

Leu Asp Asn Thr Gly Gly Arg Leu Ser Leu Thr Ala Thr Lys Ser Asn  
 450 455 460

Tyr Thr Gln Val Arg Asn Cys Leu Glu Tyr Gly Leu Tyr Ser Lys His  
 465 470 475 480

Asp Ala Ser Thr Val Met Asp Asn Thr Glu Phe Arg Asn Cys Gly Asn  
 485 490 495

His Pro Ala Ala Val Ala Tyr Gly Ala Ala Ile Phe Ala Tyr Lys Phe  
 500 505 510

Asn Cys Ser Val Asp Thr Arg Gly Val Lys Phe Tyr Gly Asn Asn Ile  
 515 520 525

Ala Gln His Cys Arg Gly Gly Ile Thr Ser Asp Asn Pro Gly Asp Pro  
 530 535 540

Asp Ile Tyr Gly Thr Gly Ala Asp Ala Asn Lys Arg Leu Phe Leu Cys  
 545 550 555 560



## 038595

Thr Gly Gly Gly Ser Asp Asp Ile Gln Phe Tyr Glu Ala Arg Arg Val  
565 570 575

Met Asp Ile Thr Lys Arg Thr Gly Gly Gly Ser Thr Thr Ala Ser Val  
580 585 590

Ser Ser Leu Leu Leu Ala Ala Val Ala Ser Val Arg Lys Gly Tyr Phe  
595 600 605

Ala His Asn Asp Gln Val Ile Arg Met Thr Leu Met Phe Arg Ala Thr  
610 615 620

Gly Ser Ala Gly Ile Phe Thr Pro Thr Leu Arg Thr Pro Leu Gly Thr  
625 630 635 640

Ile Pro Leu Gly Ser Phe Arg Val Ala Ser Gly Gln Tyr Gly Glu Ile  
645 650 655

Lys Leu Thr Ile Arg Pro Thr Leu Thr Ser Asp Gly Leu Ile Val Gly  
660 665 670

Phe Ser Cys Ile Asn Ala Val Gln Asn Leu Gly Ser Ser Val Gly Gln  
675 680 685

Ile Ile Val Ser Gly Thr Val Asp Leu Arg Thr Val Asp Gln Leu Val  
690 695 700

Glu Met Trp Gly Tyr Ser Glu Ala Gly Gly Thr Ala Ser Tyr Ile Gln  
705 710 715 720

Gly Leu Ile Glu Leu Val Gly  
725

<210> 42

<211> 362

<212> белок

<213> *Aggregatibacter actinomycetemcomitans*

## 038595

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;223&gt; последовательность белка DspB

&lt;400&gt; 42

Met Asn Cys Cys Val Lys Gly Asn Ser Ile Tyr Pro Gln Lys Thr Ser  
 1 5 10 15

Thr Lys Gln Thr Gly Leu Met Leu Asp Ile Ala Arg His Phe Tyr Ser  
 20 25 30

Pro Glu Val Ile Lys Ser Phe Ile Asp Thr Ile Ser Leu Ser Gly Gly  
 35 40 45

Asn Phe Leu His Leu His Phe Ser Asp His Glu Asn Tyr Ala Ile Glu  
 50 55 60

Ser His Leu Leu Asn Gln Arg Ala Glu Asn Ala Val Gln Gly Lys Asp  
 65 70 75 80

Gly Ile Tyr Ile Asn Pro Tyr Thr Gly Lys Pro Phe Leu Ser Tyr Arg  
 85 90 95

Gln Leu Asp Asp Ile Lys Ala Tyr Ala Lys Ala Lys Gly Ile Glu Leu  
 100 105 110

Ile Pro Glu Leu Asp Ser Pro Asn His Met Thr Ala Ile Phe Lys Leu  
 115 120 125

Val Gln Lys Asp Arg Gly Val Lys Tyr Leu Gln Gly Leu Lys Ser Arg  
 130 135 140

Gln Val Asp Asp Glu Ile Asp Ile Thr Asn Ala Asp Ser Ile Thr Phe  
 145 150 155 160

Met Gln Ser Leu Met Ser Glu Val Ile Asp Ile Phe Gly Asp Thr Ser  
 165 170 175

Gln His Phe His Ile Gly Gly Asp Glu Phe Gly Tyr Ser Val Glu Ser  
 180 185 190

## 038595

Asn His Glu Phe Ile Thr Tyr Ala Asn Lys Leu Ser Tyr Phe Leu Glu  
 195 200 205

Lys Lys Gly Leu Lys Thr Arg Met Trp Asn Asp Gly Leu Ile Lys Asn  
 210 215 220

Thr Phe Glu Gln Ile Asn Pro Asn Ile Glu Ile Thr Tyr Trp Ser Tyr  
 225 230 235 240

Asp Gly Asp Thr Gln Asp Lys Asn Glu Ala Ala Glu Arg Arg Asp Met  
 245 250 255

Arg Val Ser Leu Pro Glu Leu Leu Ala Lys Gly Phe Thr Val Leu Asn  
 260 265 270

Tyr Asn Ser Tyr Tyr Leu Tyr Ile Val Pro Lys Ala Ser Pro Thr Phe  
 275 280 285

Ser Gln Asp Ala Ala Phe Ala Ala Lys Asp Val Ile Lys Asn Trp Asp  
 290 295 300

Leu Gly Val Trp Asp Gly Arg Asn Thr Lys Asn Arg Val Gln Asn Thr  
 305 310 315 320

His Glu Ile Ala Gly Ala Ala Leu Ser Ile Trp Gly Glu Asp Ala Lys  
 325 330 335

Ala Leu Lys Asp Glu Thr Ile Gln Lys Asn Thr Lys Ser Leu Leu Glu  
 340 345 350

Ala Val Ile His Lys Thr Asn Gly Asp Glu  
 355 360

<210> 43

<211> 22

<212> белок

<213> Staphylococcus aureus

## 038595

<220>  
 <221> misc\_feature  
 <223> последовательность белка SaPSMa3

<400> 43

Met Glu Phe Val Ala Lys Leu Phe Lys Phe Phe Lys Asp Leu Leu Gly  
 1 5 10 15

Lys Phe Leu Gly Asn Asn  
 20

<210> 44  
 <211> 44  
 <212> белок  
 <213> Staphylococcus aureus

<220>  
 <221> misc\_feature  
 <223> последовательность белка SaPAMb2

<400> 44

Met Thr Gly Leu Ala Glu Ala Ile Ala Asn Thr Val Gln Ala Ala Gln  
 1 5 10 15

Gln His Asp Ser Val Lys Leu Gly Thr Ser Ile Val Asp Ile Val Ala  
 20 25 30

Asn Gly Val Gly Leu Leu Gly Lys Leu Phe Gly Phe  
 35 40

<210> 45  
 <211> 22  
 <212> белок  
 <213> Staphylococcus epidermidis

<220>  
 <221> misc\_feature  
 <223> последовательность белка SePSMa

<400> 45

Met Ala Asp Val Ile Ala Lys Ile Val Glu Ile Val Lys Gly Leu Ile  
 1 5 10 15

## 038595

Asp Gln Phe Thr Gln Lys  
20

<210> 46  
<211> 75  
<212> белок  
<213> Levivirus MS2

<220>  
<221> misc\_feature  
<223> последовательность белка L MS2

<400> 46

Met Glu Thr Arg Phe Pro Gln Gln Ser Gln Gln Thr Pro Ala Ser Thr  
1 5 10 15

Asn Arg Arg Arg Pro Phe Lys His Glu Asp Tyr Pro Cys Arg Arg Gln  
20 25 30

Gln Arg Ser Ser Thr Leu Tyr Val Leu Ile Phe Leu Ala Ile Phe Leu  
35 40 45

Ser Lys Phe Thr Asn Gln Leu Leu Leu Ser Leu Leu Glu Ala Val Ile  
50 55 60

Arg Thr Val Thr Thr Leu Gln Gln Leu Leu Thr  
65 70 75

<210> 47  
<211> 54  
<212> белок  
<213> Levivirus PRR1

<220>  
<221> misc\_feature  
<223> последовательность белка L PRR1

<400> 47

Met Cys Lys Val Ser Thr Lys Val Asp Ser Lys Leu Thr Glu Ser Val  
1 5 10 15

## 038595

Gly Gln Leu Thr Ile Arg Ser Tyr Leu Trp Leu Arg Asn Ile Leu Ala  
 20 25 30

Leu Ala Gly Leu Leu Phe Val Ile Leu Leu Ala Thr Asn His Leu Ser  
 35 40 45

Ile Ala Ile Tyr Ser Pro  
 50

<210> 48  
 <211> 106  
 <212> белок  
 <213> Phikmvlikevirus LUZ19

<220>  
 <221> misc\_feature  
 <223> последовательность белка Gp18 LUZ19

<400> 48

Met Arg Met Pro Thr Glu Glu Glu Arg Met Ile Arg Cys Leu Leu Ala  
 1 5 10 15

Asp Ile His Glu Pro Leu Asp Leu Leu Phe Pro Gly Leu Arg Thr Lys  
 20 25 30

Ala His Met Asp Pro Gln Ala Glu Glu Leu Ser Ile Arg Ile Asp Tyr  
 35 40 45

Asp His Ala Lys Leu Gly Arg Met Gly Phe Cys His Ala Val Ser Leu  
 50 55 60

Tyr Gln Leu Ser Ile Tyr Gly Arg Glu Gly Met Val Arg Tyr Leu Met  
 65 70 75 80

Gln Glu Ile Pro Arg Arg Val Leu Glu Gly Leu Leu Val Lys Ala Gln  
 85 90 95

Gln Tyr Ser Gln Ser Asn Trp Tyr Ser Lys  
 100 105

<210> 49  
 <211> 71  
 <212> белок  
 <213> Phikmvlikevirus LUZ19

<220>  
 <221> misc\_feature  
 <223> последовательность белка Gr49 LUZ19

<400> 49

Met Ser Lys Ala Lys Leu Arg Val Ile Ala Asp Thr Pro Glu Leu Glu  
 1 5 10 15

Ser Val Leu Lys Ala Leu Leu Thr Ala Thr Tyr Ala Ile Glu Asp Leu  
 20 25 30

Leu Asn Glu Ala Val Ala Ser Lys Val Leu Asn Ser Arg Leu Gly Trp  
 35 40 45

Ser Ala Val Gly Glu Tyr Val Glu Leu Phe Asn Arg Thr Gln Ser Arg  
 50 55 60

Val Ala Gly Leu Ile Pro Glu  
 65 70

<210> 50  
 <211> 321  
 <212> ДНК  
 <213> Phikmvlikevirus LUZ19

<220>  
 <221> misc\_feature  
 <223> последовательность гена gp18 LUZ19

<400> 50  
 atgagaatgc caaccgaaga agaacgcatg atccgctggt tactggcgga tatccacgag 60  
 ccactggacc tgctgttccc cggcctccgt accaaggccc atatggaccc gcaagcagag 120  
 gaactgtcga ttcgaattga ctacgacat gcgaagctgg gccgtatggg attctgccac 180  
 gcggtatccc tatatcaact gtccatatat ggccgagagg ggatgggtccg ctacctgatg 240

```

caggagattc cccgccgct gctggaaggt ctgctgtca aggcgcagca gtacagccaa      300
agcaactggg acagcaaatg a                                             321

<210> 51
<211> 216
<212> ДНК
<213> Phikmvlikevirus LUZ19

<220>
<221> misc_feature
<223> последовательность гена Gp49 LUZ19

<400> 51
atgagcaaag ccaaactacg agtcatcgcc gacaccccg agctggagtc agtgctaaaa      60
gcattgctga ccgccaccta cgctatcgag gacctgctca acgaggccgt ggctagcaag     120
gtgctaaact cccgcctggg ctggtccgca gtcggcgagt atgtcgaact gttcaaccgc     180
acgcaatccc gcgtggccgg gttgattccc gagtag                               216

```

#### ФОРМУЛА ИЗОБРЕТЕНИЯ

1. Сконструированный вирус, способный при введении в клетку-хозяина продуцировать не встречающиеся в природе вирусные частицы с одним или более улучшенными вирусными свойствами, содержащий:

(i) весь или часть сконструированного гена gp34, где сконструированный ген gp34 кодирует аминокислотную последовательность, содержащую делецию в положении, соответствующем аминокислотному положению 55 в SEQ ID NO: 5 (белок gp34); и/или

(ii) модификацию одной или более последовательностей нуклеиновых кислот, кодирующих аминокислотную последовательность, имеющую по меньшей мере 85% идентичность с последовательностью, выбранной из группы, состоящей из SEQ ID NO: 34 (белок gp13), SEQ ID NO: 35 (белок gp38), SEQ ID NO: 36 (белок gp40) и SEQ ID NO: 48 (белок gp18), где модификации включают замену С на У в положении, соответствующем аминокислотному положению 17 в SEQ ID NO: 34 (белок gp13), замену D на У в положении, соответствующем аминокислотному положению 36 в SEQ ID NO: 48 (белок gp18), замену D на G в положении, соответствующем аминокислотному положению 82 в SEQ ID NO: 35 (белок gp38), замену I на S в положении, соответствующем аминокислотному положению 83 в SEQ ID NO: 35 (белок gp38), и замену N на D в положении, соответствующем аминокислотному положению 253 в SEQ ID NO: 36 (белок gp40); и/или

(iii) вставку гетерологичной молекулы нуклеиновой кислоты в последовательность нуклеиновой кислоты, кодирующую аминокислотную последовательность, имеющую по меньшей мере 85% идентичность с SEQ ID NO: 49 (белок gp49), или замену последовательности нуклеиновой кислоты, содержащейся в последовательности нуклеиновой кислоты, кодирующей аминокислотную последовательность, имеющую по меньшей мере 85% идентичность с SEQ ID NO: 49 (gp49), молекулой гетерологичной нуклеиновой кислоты; и/или

(iv) весь или часть гетерологичного гена gp18, где гетерологичный ген gp18 кодирует аминокислотную последовательность по меньшей мере с 85% идентичностью с SEQ ID NO: 38 (последовательность белка gp18).

2. Сконструированный вирус по п.1, причем указанный вирус содержит сконструированный ген gp34, где сконструированный ген gp34 кодирует аминокислотную последовательность, содержащую делецию в положении, соответствующем аминокислотному положению 55 в SEQ ID NO: 5 (белок gp34), причем указанный фаг проявляет повышенную литическую активность по сравнению с LUZ19 дикого типа.

3. Сконструированный вирус по любому из пп.1 и 2, причем указанный вирус содержит модификацию одной или более последовательностей нуклеиновых кислот, кодирующих аминокислотную последовательность, имеющую по меньшей мере 85% идентичность с последовательностью, выбранной из группы, состоящей из SEQ ID NO: 34 (белок gp13), SEQ ID NO: 35 (белок gp38), SEQ ID NO: 36 (белок gp40) и SEQ ID NO: 48 (белок gp18), где модификации включают замену С на У в положении, соответствующем аминокислотному положению 17 в SEQ ID NO: 34 (белок gp13), замену D на У в положении, соответствующем аминокислотному положению 36 в SEQ ID NO: 48 (белок gp18), замену D на G в положе-



нии, соответствующем аминокислотному положению 82 в SEQ ID NO: 35 (белок gp38), замену I на S в положении, соответствующем аминокислотному положению 83 в SEQ ID NO: 35 (белок gp38), и замену N на D на положении, соответствующем аминокислотному положению 253 в SEQ ID NO: 36 (белок gp40), в результате чего указанный вирус демонстрирует улучшенный диапазон хозяев по сравнению с LUZ19 дикого типа.

4. Сконструированный вирус по любому из пп.1-3, причем указанный вирус содержит замену последовательности нуклеиновой кислоты, содержащей последовательность нуклеиновой кислоты, кодирующую аминокислотную последовательность, имеющую по меньшей мере 85% идентичность с SEQ ID NO: 49 (gp49), гетерологичной нуклеиновой кислотой.

5. Сконструированный вирус по п.4, в котором указанная гетерологичная нуклеиновая кислота содержит ген, который кодирует агент, диспергирующий биопленку.

6. Сконструированный вирус по п.5, в котором указанная гетерологичная нуклеиновая кислота кодирует экзополисахаридную (EPS) деполимеразу или феноластворимый модулин (PSM).

7. Сконструированный вирус по п.5, в котором указанная гетерологичная нуклеиновая кислота содержит ген, выбранный из группы, состоящей из деполимеразы EPS (Pp15gp44 хвостовой шип gp44 из *Pseudomonas pudita* Ф15 (SEQ ID NO: 14)), NTn34 хвостовой шип gp34 из фара *Klebsiella pneumoniae* NTU (SEQ ID NO: 13), LKA1gp49 хвостовой шип gp49 из фара *P.aeruginosa* (SEQ ID NO: 12), поверхностно-активные фенол-растворимые модулины из *Staphylococcus epidermidis* (PSMa, SEQ ID NO: 18), *Staphylococcus aureus* (PSMa3 (SEQ ID NO: 16) и/или PSMb2 (SEQ ID NO: 17) и сурфактин DspB из *Aggregatibacter actinomycetemcomitans* (SEQ ID NO: 15).

8. Сконструированный вирус по любому из пп.5-7, где гетерологичная последовательность нуклеиновой кислоты функционально связана с основным капсидным промотором P<sub>gp32</sub> (SEQ ID NO: 21).

9. Сконструированный вирус по любому из пп.5-8, в котором гетерологичная последовательность нуклеиновой кислоты функционально связана с терминатором T<sub>gp32</sub> (SEQ ID NO: 22).

10. Сконструированный вирус по любому из пп.1-9, причем сконструированный вирус представляет собой бактериофаг, сконструированный для экспрессии противомикробного пептида.

11. Сконструированный вирус по п.10, в котором противомикробный пептид выбран из группы, состоящей из литического фермента, ацилазы, аминопептидазы, амилазы, карбогидразы, карбоксипептидазы, каталазы, целлюлазы, хитиназы, кутиназы, циклодекстрин гликозилтрансферазы, дезоксирибонуклеазы, эстеразы, α-галактозидазы, β-галактозидазы, глюкоамилазы, α-глюкозидазы, β-глюкозидазы, галлопероксидазы, инвертазы, лакказы, липазы, маннозидазы, оксидазы, пектинолитического фермента, пептидоглутаминазы, пероксидазы, фитазы, полифенолоксидазы, протеолитического фермента, рибонуклеазы, транслугтаминазы, ксиланазы, РНКазы, ДНКазы, лизостафина или порообразующего пептида.

12. Сконструированный вирус по п.10, где противомикробный пептид представляет собой ДНКазу.

13. Сконструированный вирус по любому из пп.1-12, в котором указанный модифицированный геном представляет собой модифицированный phiKMV-подобный геном.

14. Сконструированный вирус по любому из пп.1-13, в котором указанный модифицированный геном представляет собой модифицированный геном LUZ19.

15. Сконструированный вирус по п.14, в котором gp18 LUZ19 дикого типа заменен на gp18 вируса LKD16 (SEQ ID NO: 7).

А) Очистка вирусного генома непосредственно из вирусных частиц



Б) Сайт-специфическое расщепление очищенного вирусного генома (ДНК или РНК) с очищенной РНК-управляемой нуклеазой и нацеливающей РНК(ами)



В) Деактивация РНК-управляемой нуклеазы и необязательная очистка расщепленных продуктов с использованием фенольно-хлороформной экстракции



Г) Создание вставленного фрагмента посредством прямого синтеза ДНК/РНК, сборки множества олигонуклеотидов, выделения рекомбинантной ДНК (такой как плазмиды), или ПЦР-амплификации из матрицы



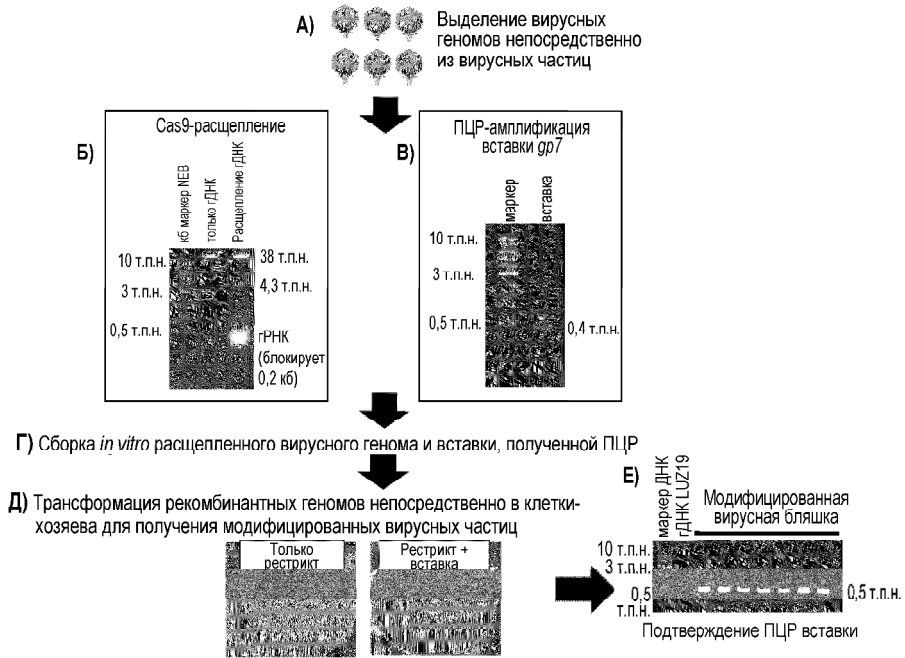
Д) Сборка *in vitro* геномных фрагментов с использованием методик *in vitro*, таких как, но без ограничений, сборка Гибсона, SLIC, лигирование и др.



Е) Трансформация рекомбинантного генома непосредственно в клетки хозяина для образования функциональных вирусных частиц

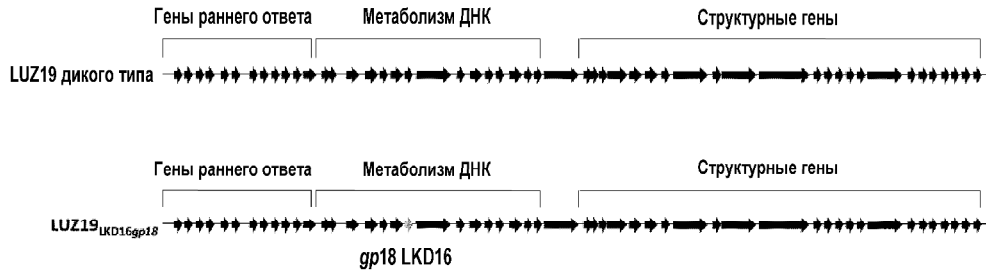


Фиг. 1

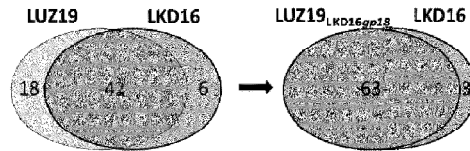


Фиг. 2

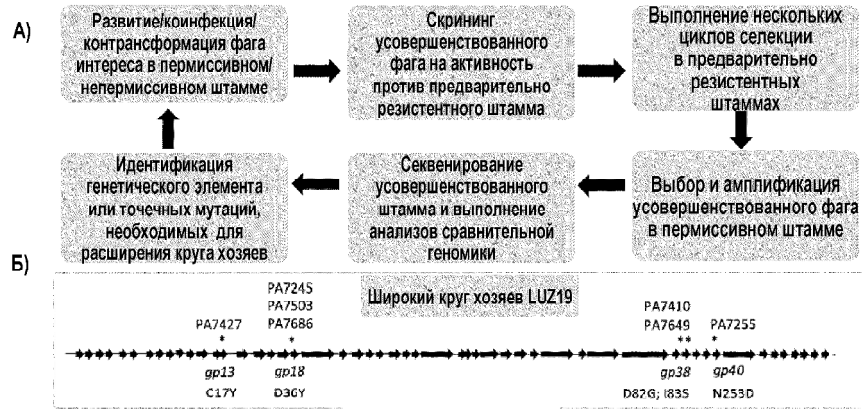
**А)** LUZ19 дикого типа и модифицированный *in vitro* геном LUZ19, содержащий ген *gp18* из вируса LKD16



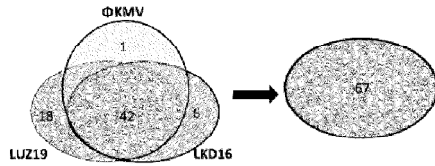
**Б)** Круг хозяев фагов LUZ19, LKD16 и LUZ19<sub>LKD16gp18</sub>



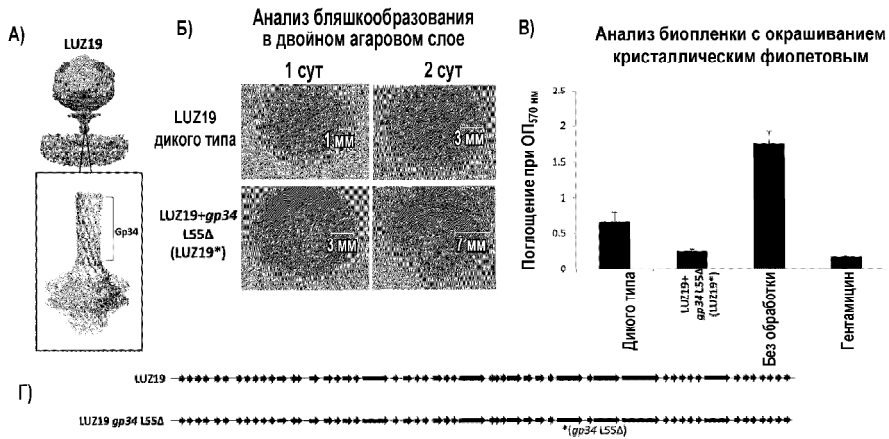
Фиг. 3



**В)** Круг хозяев рода фагов Phi-KMV среди штаммов *P. aeruginosa* клинической библиотеки



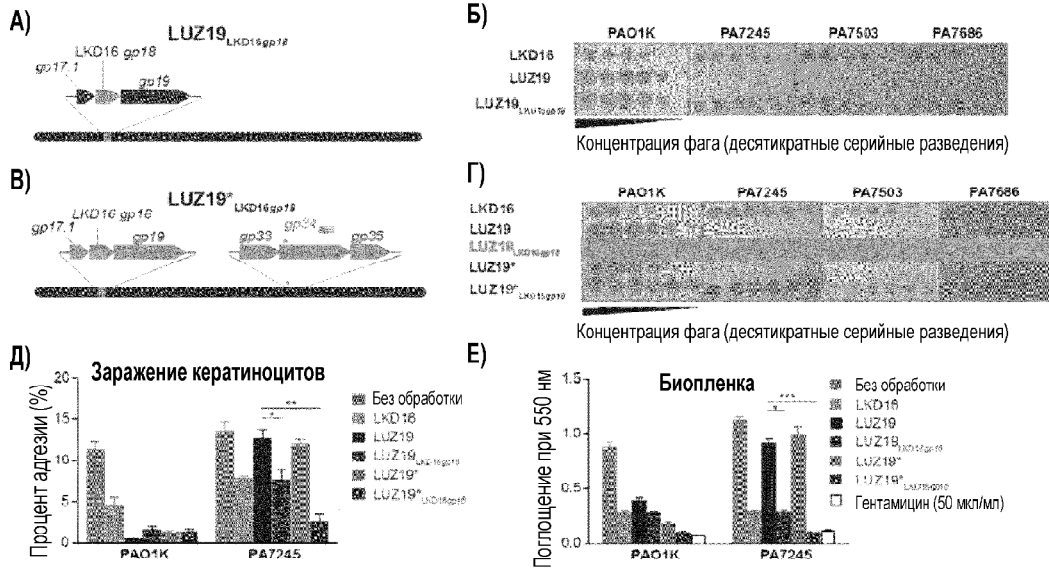
Фиг. 4



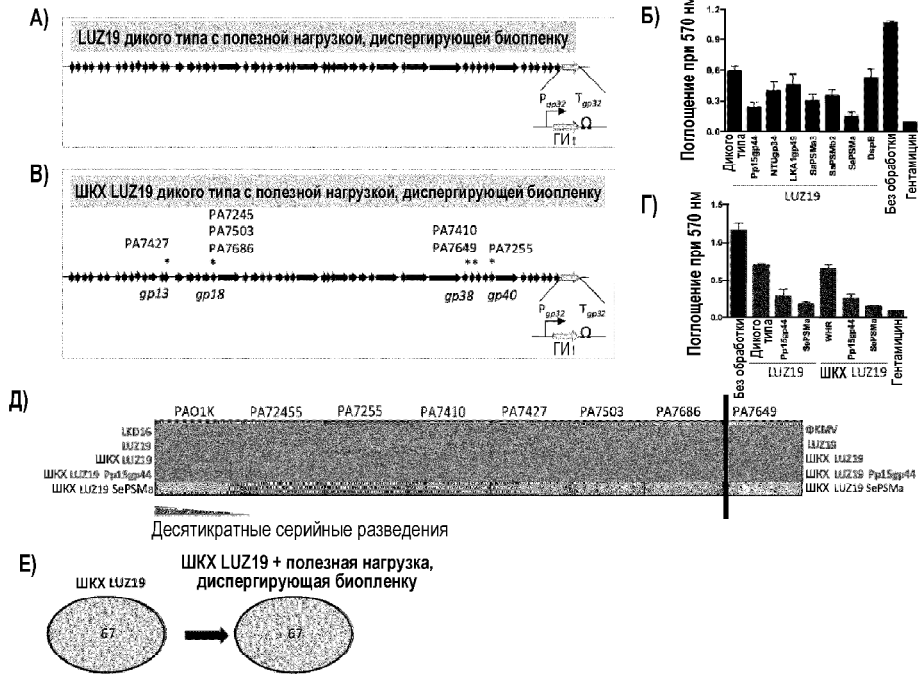
**Д)**

Фаг	Адсорбция	Латентный период	Выход фага
LUZ19	~17	30 мин	70
LUZ19+gp34 L55Δ (LUZ19*)	~12	30 мин	122

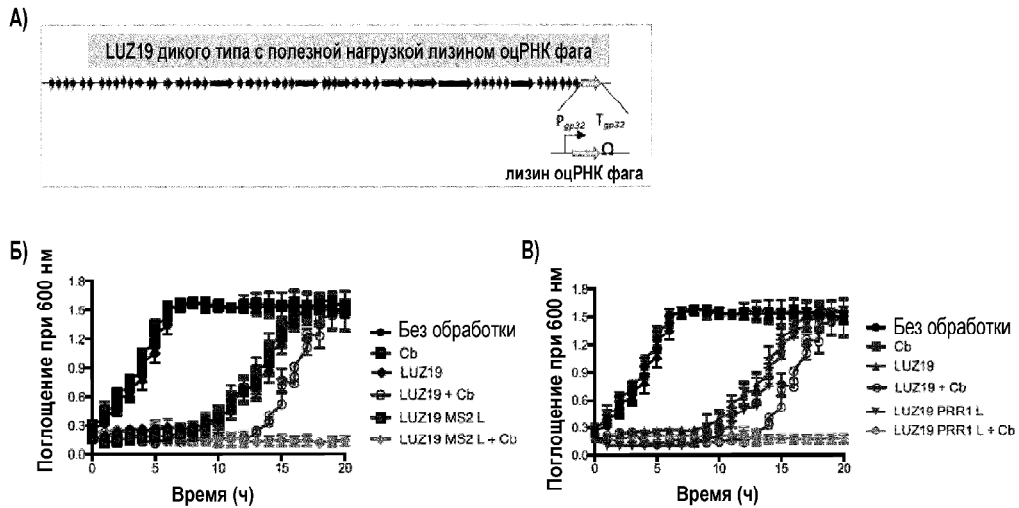
Фиг. 5



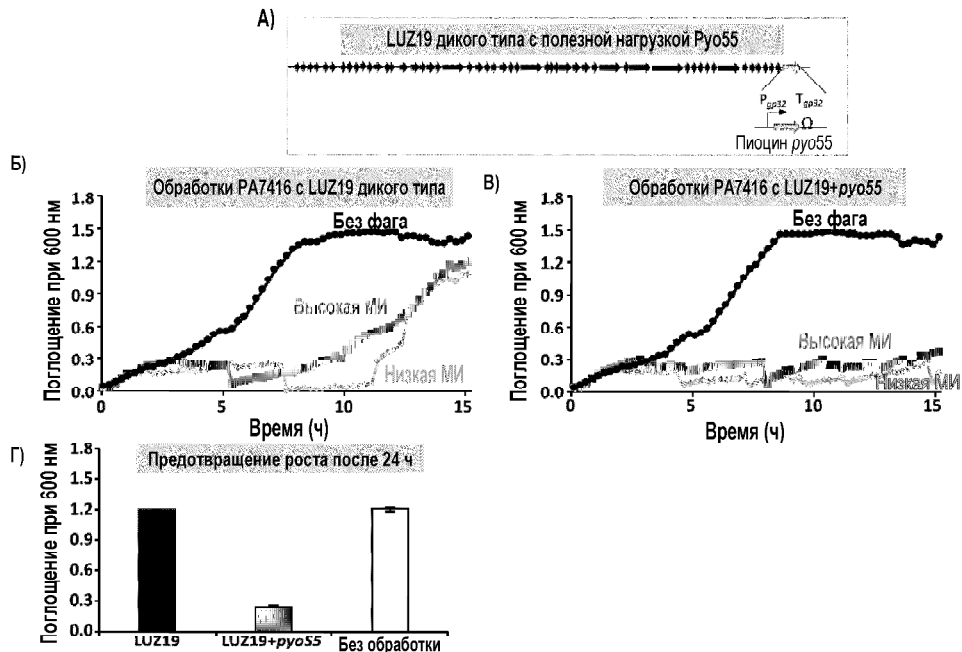
Фиг. 6



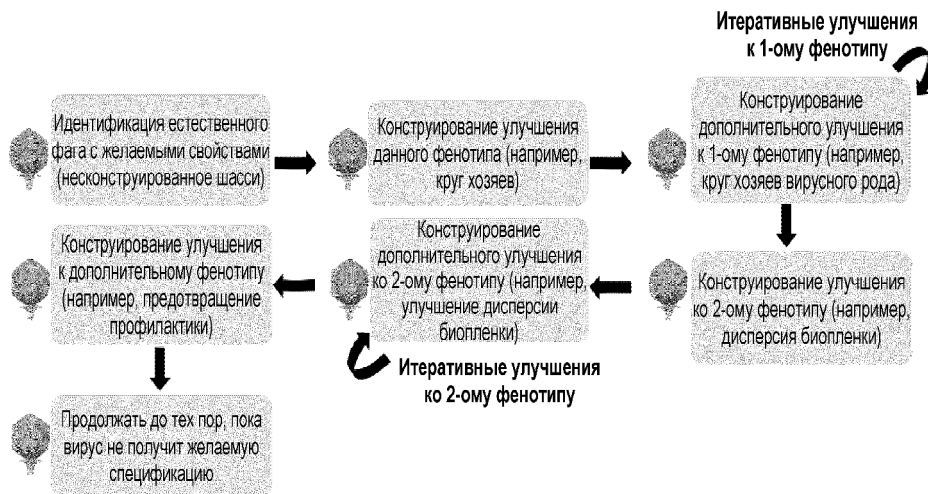
Фиг. 7



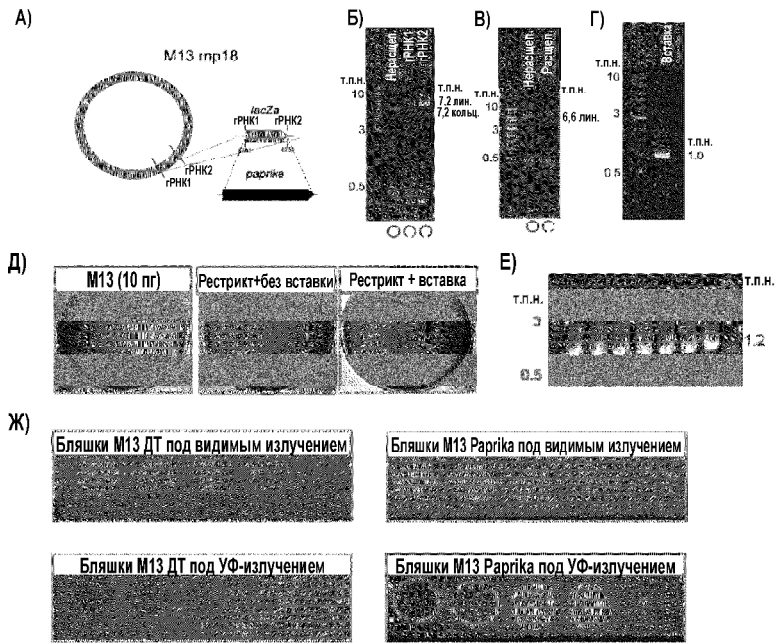
Фиг. 8



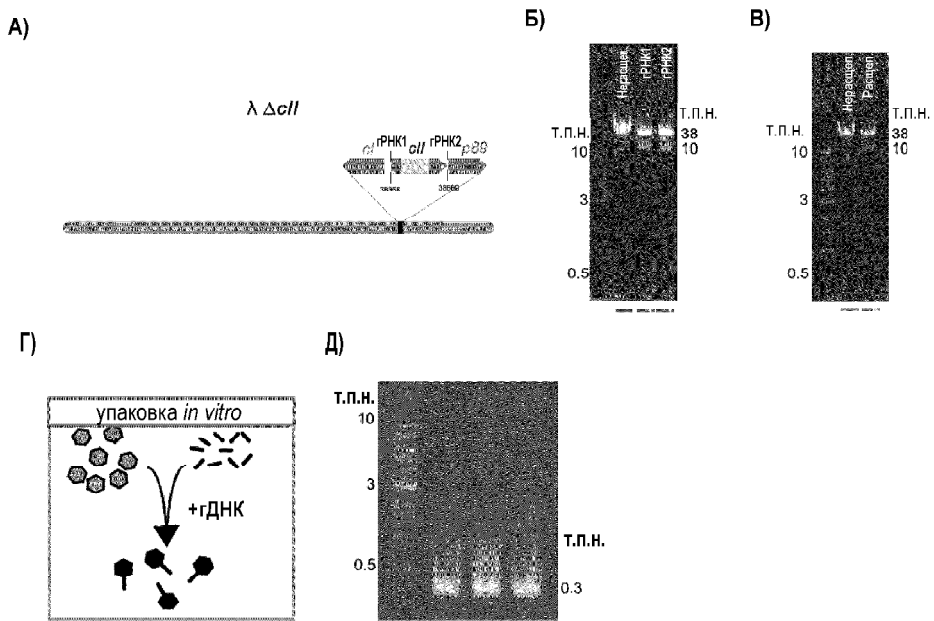
Фиг. 9



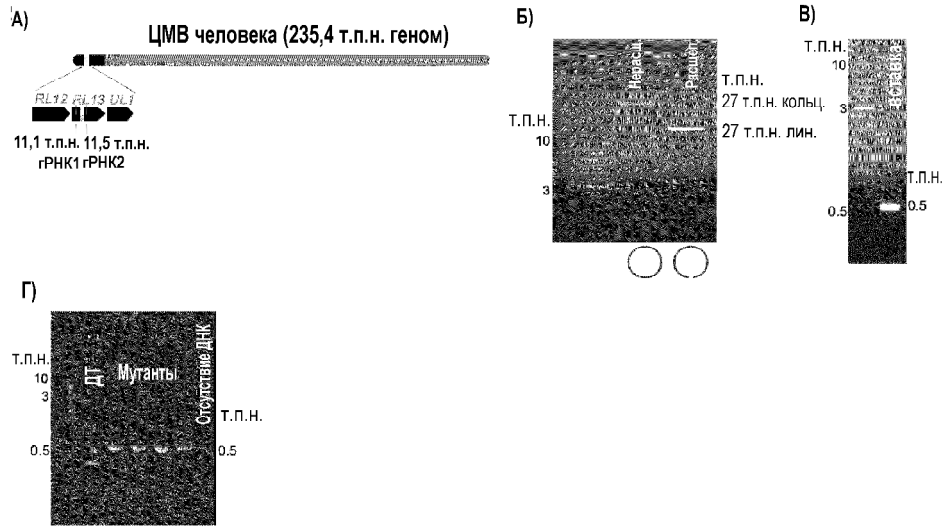
Фиг. 10



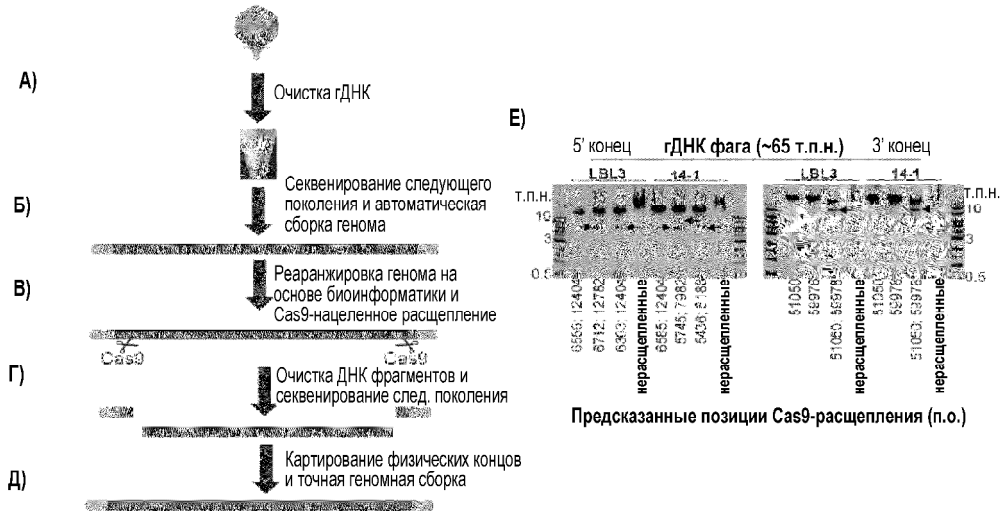
Фиг. 11



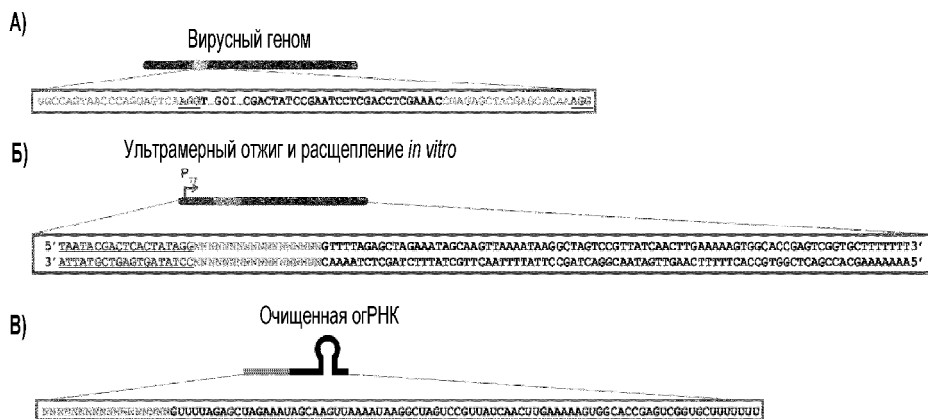
Фиг. 12



Фиг. 13



Фиг. 14



Фиг. 15