

ПОИСК И АНАЛИЗ CRISPR-CAS СИСТЕМЫ В ШТАММЕ *ESCHERICHIA COLI* HS И ДЕТЕКТИРУЕМЫХ СПЕЙСЕРАМИ ЕГО CRISPR-КАССЕТЫ ФАГОВЫХ РАС МЕТОДАМИ БИОИНФОРМАТИКИ

Е. И. Иванова¹✉, Ю. П. Джиоев², А. Ю. Борисенко², Н. П. Перетолчина², Л. А. Степаненко², А. И. Парамонов¹, Е. В. Григорова¹, У. М. Немченко¹, Т. В. Туник¹, Е. А. Кунгурцева¹

¹ Научный центр проблем здоровья семьи и репродукции человека, Иркутск

² Институт биомедицинских технологий, Иркутский государственный медицинский университет, Иркутск

CRISPR-Cas система — это иммунная система прокариот, обеспечивающая защиту от чужеродных репликонов, в первую очередь вирусов и плазмид. Устойчивость к повторным инфекциям приобретает в результате включения в состав CRISPR-кассет коротких последовательностей, или спейсеров, комплементарных участкам соответствующих вирусных или плазмидных геномов. В настоящее время эшерихиозные инфекции остаются серьезной проблемой практической медицины. Вследствие их крайней устойчивости к терапии с использованием антибиотиков необходима разработка новых подходов лечения. Целью исследования был поиск структур CRISPR-Cas систем в геномной последовательности штамма *Escherichia coli* HS. Использовали методы программного моделирования MacSyFinder (Macromolecular System Finder, ver. 1.0.2.). Поиск точной гомологии последовательностей осуществляли посредством установленных вспомогательных пакетов makeblastdb (ver. 2.2.28), HMMER (ver. 3.0). В результате методами биоинформатики была выявлена одна CRISPR-Cas система и расшифрованы спейсерные последовательности CRISPR-кассеты у штамма *Escherichia coli* HS. С помощью последовательностей спейсеров CRISPR-кассеты были определены комплементарные им протоспейсерные участки нескольких типов фагов, что позволяет оценить степень их устойчивости к этим чужеродным генетическим элементам.

Ключевые слова: биоинформатика, CRISPR-Cas система, *Escherichia coli* HS, бактериофаги

✉ Для корреспонденции: Елена Иннокентьевна Иванова
ул. Тимирязева, д. 16, г. Иркутск, 664003; ivanova.iem@gmail.com

Статья получена: 15.03.18 Статья принята к печати: 24.03.18

DOI: 10.24075/vrgmu.2018.019

THE SEARCH AND ANALYSIS OF A CRISPR-CAS SYSTEM IN *ESCHERICHIA COLI* HS WITH SUBSEQUENT SCANNING FOR THE CORRESPONDING PHAGE RACES BASED ON THE SPACERS OF THE DETECTED CRISPR ARRAY USING BIOINFORMATIC METHODS

Ivanova EI¹✉, Dzhioev YuP², Borisenko AY², Peretolchina NP², Stepanenko LA², Paramonov AI¹, Grigorova EV¹, Nemchenko UM¹, Tunik TV¹, Kungurtseva EA¹

¹ Scientific Center for Family Health and Human Reproduction Problems, Irkutsk

² Research Institute for Biomedical Technologies of Irkutsk State Medical University, Irkutsk

CRISPR-Cas is an immune system of prokaryotes that protects them against alien replicons, mainly viruses and plasmids. Short sequences (spacers) complementary to the regions of a viral or plasmid genome are inserted into a CRISPR array conferring resistance to reinfection. Infections caused by *Escherichia coli* still present a serious challenge for clinical medicine. The aim of this study was to scan the genome of *Escherichia coli* HS for CRISPR-Cas components. The search was conducted using MacSyFinder (Macromolecular System Finder, ver. 1.0.2.), a program for bioinformatic modelling. Sequence homology searches were done using makeblastdb (ver. 2.2.28) and HMMER (ver. 3.0) tools. Bioinformatics-based methods allowed us to detect one CRISPR-Cas system in the studied genome of *Escherichia coli* HS and read the spacer sequences of its CRISPR array. The protospacer regions complementary to the spacer sequences of the detected CRISPR array are typical for a few types of phages. Based on these findings, one can assess the degree of bacterial resistance to alien genetic elements.

Keywords: bioinformatics, CRISPR-Cas system, *Escherichia coli* HS, bacteriophage

✉ Correspondence should be addressed: Elena Ivanova
Timiryazeva 16, Irkutsk, 664003; ivanova.iem@gmail.com

Received: 15.03.18 Accepted: 24.03.18

DOI: 10.24075/brsmu.2018.019

Escherichia coli — это вид, состоящий из многих биотипов. Представители некоторых из них являются комменсальными колонизаторами кишечного биотопа млекопитающих (в том числе человека), другие вызывают различные типы заболеваний. Среди патогенных возбудителей кишечных инфекций, важную роль играет энтерогеморрагическая

кишечная палочка *Escherichia coli* O157:H7, а среди комменсальных выделяются представители биотипа *E. coli* HS. Вызванная штаммами *E. coli* O157:H7 инфекция может привести к развитию гемолитико-уремического синдрома (ГУС) с формированием почечной недостаточности. *E. coli* O157:H7 — это один из серотипов, способных

продуцировать шига-токсины [1–3]. Лечения с доказанной эффективностью не существует, и во время острой фазы заболевания рекомендуется исключительно поддерживающая терапия. Полного согласия по вопросу о необходимости применения антибиотиков для лечения инфекции, вызванной шига-токсин-продуцирующими *E. coli* (Stx-*E. coli*), до сих пор не достигнуто [4, 5]. Показано, что антибиотикотерапия на стадии гастроинтестинальной инфекции Stx-*E. coli* повышает (приблизительно в 17 раз) риск развития развернутой картины ГУС [6]. Был сделан вывод о том, что повреждение мембраны бактерий, индуцируемое антибиотиками, может способствовать развитию острой стадии заболевания из-за выделения токсины в больших количествах [7].

В связи с этим необходима разработка альтернативных антибиотикам подходов борьбы с бактериальными патогенами, среди которых вновь становится актуальной фаговая терапия [8–10]. Для развития этого направления необходимы фундаментальные знания генетической природы взаимоотношений между бактериями и фагами, что требует наличия у исследователя их геномных структур и новых методов их анализа. На сегодняшний день уже применяются компьютерные программные методы биоинформатики, позволяющие оперировать большими массивами геномных данных и получать новую информацию о генетической природе бактерий [11].

Кроме развития методов биоинформатики значительным событием за последние несколько лет стало открытие специфической адаптивной защитной системы у прокариот. Длительное время считалось, что бактерии беззащитны в отношении бактериофагов, но еще в 1987 г. в геноме *E. coli* был обнаружен странный участок, состоящий из многочисленных повторов [12]. Лишь в 2005 г. стало известно, что разделяющие эти повторы промежуточные последовательности зачастую бывают идентичными последовательностям, найденным в геномах бактериофагов и плазмид [13, 14]. Обнаруженная структура, называемая CRISPR-Cas (Clustered Regularly Interspaced Short Palindromic Repeats — CRISPR-associated proteins, или короткие палиндромные повторы, регулярно расположенные группами с CRISPR-ассоциированными белками), является специфической адаптивной защитной системой бактерий и архей от чужеродного генетического материала (в большинстве случаев от фагов и плазмид) [15–18]. CRISPR-касеты представляют собой набор палиндромных повторов длиной 21–47 пар нуклеотидов (п.н.), разделенных уникальными спейсерными сайтами. Спейсеры комплементарно соответствуют участкам генов фагов и плазмид, к которым бактерия демонстрирует устойчивость [13]. В непосредственной близости от CRISPR-локуса находятся *cas*-гены, продукты которых обеспечивают функционирование CRISPR-локусов. На сегодняшний день выделяют 3 типа CRISPR-Cas систем, различающихся по механизму действия и *cas*-генам, представленным в геноме [19].

Использование методов биоинформатики позволяет обнаруживать и определять CRISPR-Cas системы в геномах бактерий [20, 21]. Так, тестирование фаговых рас через структуры спейсеров с помощью биоинформационных программных методов помогает в короткие сроки определить степень устойчивости бактерий к специфичным фагам и плазмидам [22–24]. Исследования в этом направлении крайне актуальны как для решения практических задач терапии инфекционных заболеваний, так и для изучения внутривидовых и межвидовых

эволюционных процессов у бактерий [17, 22]. Однако для многих видов бактерий механизм взаимоотношений между бактериями и фагами посредством их CRISPR-Cas систем и анти-CRISPR-Cas систем соответственно остается совершенно неизученным. Поэтому на первом этапе исследований необходима разработка эффективного алгоритма биоинформационных методов поиска и анализа локусов и структур CRISPR-Cas систем в геномах бактерий и последующего скрининга фаговых рас, тестируемых через их CRISPR-касеты. Исходя из этого, целью данной работы был поиск и исследование сайтов CRISPR-Cas систем в геноме *Escherichia coli* HS и последующий скрининг фаговых рас через CRISPR-касеты бактерий посредством разработанного алгоритма поисковых методов биоинформатики.

МАТЕРИАЛЫ И МЕТОДЫ

Объектом исследования послужил штамм *Escherichia coli* HS, два генома которого представлены в базе данных GenBank (№ NC_009800.1, секвенирован в 2017 г. и № CP000802, секвенирован в 2014 г.). Штамм *E. coli* HS, представленный геномом в базе данных GenBank под № NC_009800.1, был культивирован из эталонного штамма коллекции Центра разработки вакцин (США) [25]. Для поиска CRISPR-Cas систем нами была использована последовательность генома эталонного штамма под № CP000802 [26]. Данный штамм был выделен от здорового человека, желудочно-кишечный тракт которого был колонизирован этим штаммом без каких-либо видимых клинических симптомов [25].

Для поиска сайтов CRISPR-Cas систем использовали методы программного моделирования MacSyFinder (Macromolecular System Finder, ver. 1.0.2.) [27]. Программа основана на поиске через белковый профиль геномных последовательностей, закодированных как скрытые марковские модели (СММ), которые доступны в базах данных: PFAM, TIGRFAM или PRODOM. Поиск точной гомологии последовательностей осуществляли при помощи установленных вспомогательных пакетов makeblastdb (ver. 2.2.28) и HMMER (ver. 3.0), а также выявлении структурных и функциональных характеристик обнаруженных *cas*-генов каждого анализируемого генома [28]. Визуализацию полученных результатов в MacSyFinder осуществляли через веб-интерфейс MacSyView. В качестве языка программирования использовали Python (ver. 2.7) [29]. Поиск расшифровки CRISPR-касет производили при помощи онлайн-приложения «CRISPR: a CRISPR RInteractive database» на Gen Ouest Bioinformatics Platform (<http://genouest.org/>). Геномы бактерий и архей в приложении загружены из NCBI FTP Server и обработаны в программах C и Java (ver. 1.5.0.12.) [30]. Метод обнаружения данного приложения был основан на ограничении числа максимальных совпадений. Во избежание детекции ошибочных структур фактический минимальный процент идентичности, необходимый для обнаружения, был фиксирован и составлял 60%. Работу веб-страницы осуществляли с помощью PHP (ver. 4.3.9), Java (ver. 1.5.0.12). Идентифицировали фаги по расшифрованным спейсерным последовательностям с помощью алгоритма поиска BLASTn по базе данных GenBank-Phage [31]. Были использованы онлайн-приложения: CRISPRTarget (http://bioanalysis.otago.ac.nz/CRISPRTarget/crispr_analysis.html), Mycobacteriophage Database (<http://phagesdb.org/blast/>).

РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЯ

В результате поиска в геноме штамма *E. coli* HS (№ CP000802) был выявлен локус одной CRISPR-Cas системы в позиции 2920652-2921839 нуклеотидных оснований (н.о.), т. е. его размер составлял 1187 н.о. По структуре эта CRISPR-Cas система была отнесена к типу CAS-Type-IE.

При помощи программы MacSyFinder были обнаружены и визуализированы участки генома *E. coli* HS — гены, кодирующие Cas-белки:

- обязательные (mandatory), присутствие которых в геноме указывает на наличие CRISPR-Cas системы (рис. 1);
- дополнительные (accessory), соответствующие генам, которые могут быть найдены в нескольких системах и которые трудно идентифицировать с помощью одного профиля белка, но также указывающие на наличие CRISPR-Cas системы у бактерии.

При помощи MacSyFinder удалось не только обнаружить cas-гены в CRISPR-Cas системе анализируемого генома *E. coli* HS, но и визуализировать полученную информацию XML через MacSyView. Пример обнаруженных cas-генов и их расположение в геноме данного штамма представлен на рис. 1.

При помощи программных пакетов HMMER (ver. 3.0) и makeblastdb (ver. 2.2.28) у обнаруженных cas-генов в каждом

анализируемом геноме, были получены структурные и функциональные характеристики: gene (профиль), system (система, к которой принадлежит ген), hitid (идентификатор), hit seq length (длина последовательности совпадения), replicon name (имя репликона), position hit (позиция во входном файле), i-eval ("независимый Evaluate"), score (оценка hit), profile coverage (процент профиля, который соответствует последовательности hit), sequence coverage (процент от последовательности hit, который соответствует профилю), begin match (начало совпадения с профилем в последовательности), end match (конец совпадения с профилем в последовательности) (рис. 2).

Расшифровку структур CRISPR-кассет осуществляли в режиме реального времени при помощи онлайн-приложения «CRISPI: a CRISPR Interactive database», основанного главным образом на гомологии повторяющихся участков. В результате в CRISPR-кассете исследуемого штамма было идентифицировано 11 повторяющихся (repeat) последовательностей, консенсусная схема которых представлена на рис. 3. После обнаружения повторов в приложении «CRISPI: a CRISPR Interactive database» удалось расшифровать в CRISPR-кассете 10 «спейсерных» (spacer) структур (табл. 1). При помощи Java были визуализированы CRISPR-кассета и cas-гены в геноме бактерии (рис. 4).

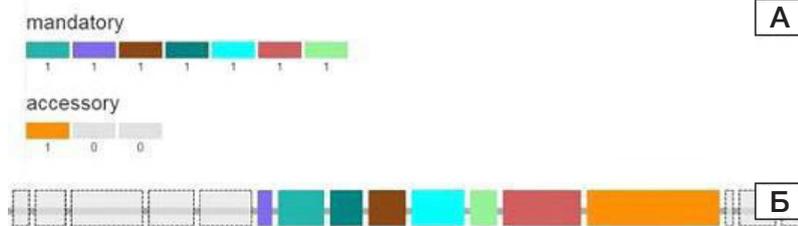


Рис. 1. Cas-гены (А) и их расположение в геноме (Б) *E. coli* HS (№ CP000802), обнаруженные при помощи MacSyFinder и визуализированные через MacSyView

| Color | Sequence Id | Position | Profile Match | Function | Gene status | System | Protein length (aa) | Score | i-value | Profile coverage | Sequence coverage | Begin match | End match |
|-------------|--------------------------------------|----------|---------------|-------------|-------------|------------|---------------------|-------|----------|------------------|-------------------|-------------|-----------|
| Blue | kl NC_002695.1_prot_NP_311635.1_3467 | 3467 | cas2_TypeIE | cas2_TypeIE | mandatory | CAS-TypeIE | 97 | 133.6 | 5.5e-40 | 1.00 | 0.89 | 3 | 88 |
| Green | kl NC_002695.1_prot_NP_311636.1_3468 | 3468 | cas1_TypeIE | cas1_TypeIE | mandatory | CAS-TypeIE | 307 | 380.1 | 1.4e-114 | 0.99 | 0.86 | 8 | 272 |
| Red | kl NC_002695.1_prot_NP_311637.1_3469 | 3469 | cas6_TypeIE | cas6_TypeIE | mandatory | CAS-TypeIE | 216 | 292.6 | 7e-88 | 1.00 | 0.98 | 1 | 212 |
| Orange | kl NC_002695.1_prot_NP_311638.1_3470 | 3470 | cas5_TypeIE | cas5_TypeIE | mandatory | CAS-TypeIE | 248 | 159.4 | 3e-47 | 0.99 | 0.93 | 3 | 233 |
| Yellow | kl NC_002695.1_prot_NP_311639.1_3471 | 3471 | cas7_TypeIE | cas7_TypeIE | mandatory | CAS-TypeIE | 351 | 447.9 | 9.3e-135 | 1.00 | 0.92 | 3 | 324 |
| Purple | kl NC_002695.1_prot_NP_311640.1_3472 | 3472 | cse2_TypeIE | cse2_TypeIE | mandatory | CAS-TypeIE | 178 | 127.7 | 1.3e-37 | 1.00 | 0.89 | 12 | 169 |
| Light Blue | kl NC_002695.1_prot_NP_311641.1_3473 | 3473 | cse1_TypeIE | cse1_TypeIE | mandatory | CAS-TypeIE | 520 | 620.3 | 7.7e-187 | 1.00 | 0.97 | 5 | 509 |
| Light Green | kl NC_002695.1_prot_NP_311642.1_3474 | 3474 | cas3_TypeI | cas3_TypeI | accessory | CAS | 885 | 216.2 | 1.9e-64 | 0.90 | 0.42 | 292 | 662 |

Рис. 2. Структурные и функциональные характеристики белков Cas *E. coli* HS (№ CP000802), обнаруженные при помощи MacSyFinder

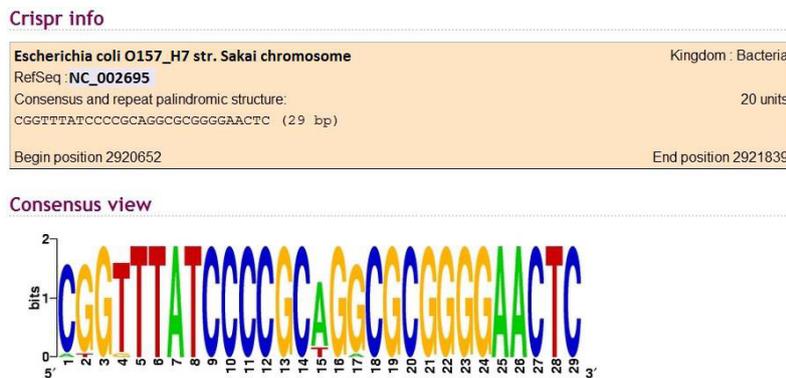


Рис. 3. Консенсусная схема чередующихся повторов в геноме штамма *E. coli* HS (№ CP000802), полученных при помощи приложения «CRISPI: a CRISPR Interactive database». Размер нуклеотидных символов свидетельствует о величине варибельности нуклеотида в последовательности повтора — чем меньше его высота, тем он более изменчив

ОБСУЖДЕНИЕ РЕЗУЛЬТАТОВ

В прошлом году в базе данных GenBank был задокументирован геном штамма *Escherichia coli* HS (№ NC_009800.1) и появилась информация о наличии в его геноме трех локусов CRISPR-Cas систем. В базе данных CRISPR-Cas систем (<http://crispr.i2bc.paris-saclay.fr/crispr/>) эти локусы представлены в нескольких вариантах. Структуры CRISPR-кассеты, полученные

в нашем исследовании генома штамма *E. coli* HS (№ CP000802, секвенирован в 2014 г.) в основном совпадали со структурами штамма *E. coli* HS (вариант NC_009800_6, секвенирован в 2017 г.).

В ходе работы была проведена идентификация фагов по расшифрованным спейсерным последовательностям, выявленным в CRISPR-кассете этого штамма (табл. 2). Из 10 спейсерных последовательностей только 4 спейсера (1, 5, 7, 10) комплементарно связались с протоспейсерными

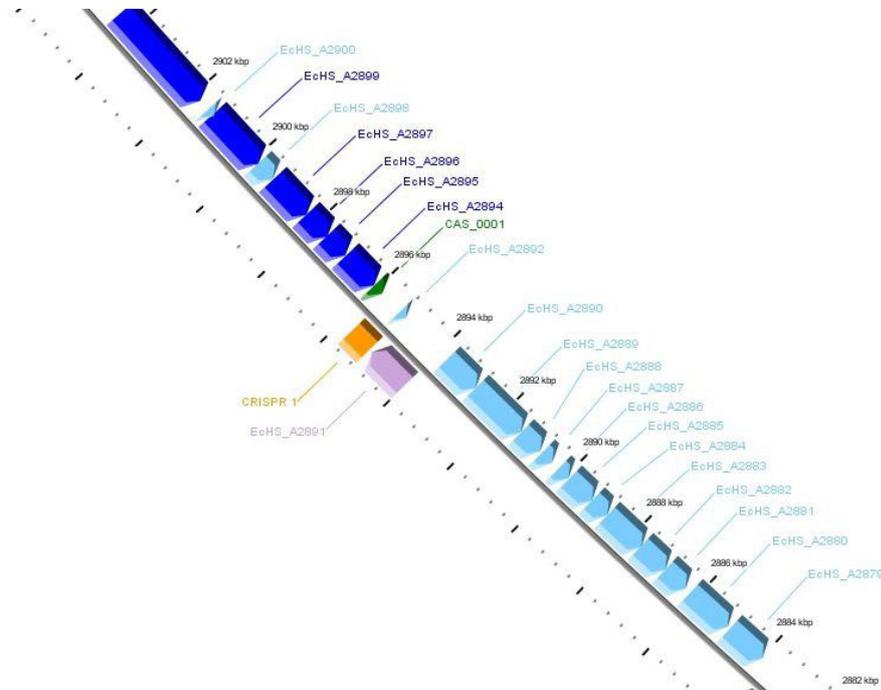


Рис. 4. Расположение cas-генов и CRISPR-кассеты в геноме штамма *E. coli* HS (№ CP000802)

Таблица 1. Список последовательностей в CRISPR-кассете: спейсерные участки (spacer) разделенные повторами (repeat-unit), обнаруженные при помощи «CRISPI: a CRISPR Interactive database» в геноме штамма *E. coli* HS (№ CP000802)

| Спейсеры/повторы | Начало | Конец | Нуклеотидные последовательности | Размер |
|------------------|---------|---------|----------------------------------|--------|
| unit 1 | 2920652 | 2920680 | ATGGTTATCCCCGCTGACGCGGGGAATC | 29 |
| spacer 1 | 2920681 | 2920712 | TCGTCCAGACTGAATACGTTGCCCCAAAATCT | 31 |
| unit 2 | 2920713 | 2920741 | CGGTTTATCCCCGCTGGCGCGGGGAATC | 29 |
| spacer 2 | 2920742 | 2920773 | CTATTGATGAGGTGCACCATCAGAAGCGAGAT | 31 |
| unit 3 | 2920774 | 2920802 | CGGTTTATCCCCGCTGGCGCGGGGAATC | 29 |
| spacer 3 | 2920803 | 2920834 | GACGTACAGATTGGCTGCGGCACCTCAAACAC | 31 |
| unit 4 | 2920835 | 2920863 | CGGTTTATCCCCGAGCGCGGGGAATC | 29 |
| spacer 4 | 2920864 | 2920895 | TTAATTCGCGTACCTGCGCATCCATTGCCGCG | 31 |
| unit 5 | 2920896 | 2920924 | CGGTTTATCCCCGAGCGCGGGGAATC | 28 |
| spacer 5 | 2920925 | 2920956 | CGCAATCATGTTTTTCATTGGGTTTACGTCTT | 31 |
| unit 6 | 2920957 | 2920985 | CGGTTTATCCCCGAGCGCGGGGAATC | 28 |
| spacer 6 | 2920986 | 2921017 | TTTTTATGACTGAATCCACTACGCCTTCATAG | 31 |
| unit 7 | 2921018 | 2921046 | CGGTTTATCCCCGAGCGCGGGGAATC | 28 |
| spacer 7 | 2921047 | 2921078 | TTTACGTCGTTGATGACATCGTTCAGGTGTTT | 31 |
| unit 8 | 2921079 | 2921107 | CGGTTTATCCCCGAGCGCGGGGAATC | 28 |
| spacer 8 | 2921108 | 2921139 | GTGATTTTCGTACCCGGCGGATCGCGATATG | 31 |
| unit 9 | 2921140 | 2921168 | CGGTTTATCCCCGAGCGCGGGGAATC | 28 |
| spacer 9 | 2921169 | 2921200 | GATAACCGCTTCGCGGTCAATATCTGCCGCAC | 31 |
| unit 10 | 2921201 | 2921229 | CGGTTTATCCCCGAGCGCGGGGAATC | 28 |
| spacer 10 | 2921230 | 2921261 | GCCCATCGCTGCGCCACTGTTAAAAAGTT | 31 |
| unit 11 | 2921262 | 2921290 | CGGTTTATCCCCGAGCGCGGGGAATC | 28 |
| spacer 11 | 2921291 | 2921322 | TCATTGCAATCATCCACTGACTCAGGGGCTG | 31 |

Таблица 2. Спект фаговых рас, выявленных по комплементарным структурам спейсерных последовательностей CRISPR-кассеты штамма *E. coli* HS (№ CP000802)

| № | Спейсер | Бактериофаги | Количество замен |
|---|------------------------------|--|------------------|
| 1 | спейсер 1 (2920681-2920712) | <i>Aeromonas phage phiAS4</i> , (HM452125) положение: 100313-100337, | 8 |
| | | <i>Cronobacter phage vB_CsaP_Ss1</i> , (KM058087) положение: 19880-19863 | 10 |
| 2 | спейсер 5 (2920925-2920956) | <i>Salmonella phage PVP-SE1</i> , (GU070616) положение: 124932-124959 | 7 |
| | | <i>Salmonella phage SSE-121</i> , (JX181824) положение: 87806-87779 | 7 |
| | | <i>Bacillus phage Bp8p-T</i> , (KJ010548) положение: 144792-144820 | 8 |
| | | <i>Bacillus phage Bp8p-C</i> , (KJ010547) положение: 144790-144818 | 8 |
| 3 | спейсер 7 (2921047-2921078) | <i>Rhizobium phage vB_RleM_P10VF</i> , (KM199770) положение: 93101-93076 | 8 |
| | | <i>Burkholderia phage phiE255</i> , (CP000622) положение: 17180-17211 | 7 |
| | | <i>Burkholderia cenocepacia phage BcepMu</i> , (AY539836) положение: 30887-30918 | 7 |
| | | <i>Gordonia phage GTE5</i> , (JF923796) положение: 49708-49734 | 8 |
| | | <i>Dickeya phage vB_DsoM_LIMEstone1</i> (HE600015) положение: 52018-52038 | 8 |
| | | <i>Dickeya phage RC-2014</i> , (KJ716335) положение: 27496-27516 | 8 |
| | | <i>Synechococcus phage S-CAM1</i> (HQ634177) положение: 189041-189018 | 9 |
| | | <i>Cyanophage S-SSM6b</i> (HQ316603) положение: 161353-161374 | 10 |
| | | <i>Cyanophage S-SSM4</i> (HQ316583) положение: 103276-103255 | 10 |
| 4 | спейсер 10 (2921230-2921261) | <i>Bacteriophage RTP</i> , (AM156909) положение: 34535-34554 | 10 |

структурами фаговых рас, представленными в данной таблице. Выявленные фаговые расы характерны для широкого круга разнообразных бактериальных хозяев. Возможно, это следствие горизонтального переноса CRISPR-Cas систем между типами бактерий за длительный период формирования и развития у них «адаптивной иммунной системы». Необходимы дальнейшие исследования, в ходе которых будут получены новые фундаментальные знания о природе антагонистических взаимоотношений между бактериями и фагами. По детектируемым фаговым расам можно судить о степени защищенности и живучести бактерий в период эволюционного развития.

ВЫВОДЫ

Проведенные поиск и анализ расшифрованной структуры CRISPR-кассеты в геноме штамма *E. coli* HS (№ CP000802, секвенирован в 2014 г.) позволяют считать используемые биоинформационные программные методы поиска структур

CRISPR-Cas систем в расшифрованных геномах бактерий весьма эффективными, а результаты высоко информативными. Наличие Cas-белков, относящихся к категории mandatory, свидетельствует о высокой антифаговой активности CRISPR-Cas системы изученного штамма. Количество выявленных спейсеров свидетельствует о длительности его эволюционной истории. Сравнительный анализ структур спейсеров в двух CRISPR-кассетах, выявленных в геноме штамма *E. coli* HS (№ CP000802, секвенированного в 2014 г. и № NC_009800.1, секвенированного в 2017 г.), показывает, что количество спейсеров в CRISPR-кассете генома штамма № NC_009800.1 увеличилось до 19. При этом количество спейсеров в геноме штамма № CP000802 равно всего 10. Можно предположить, что такое увеличение количества спейсеров стало возможным за счет их накопления в процессах частого пассирования этого штамма либо вследствие большого числа случаев его контаминации фагами. В любом случае оно может свидетельствовать о высокой активности CRISPR-Cas системы штамма *E. coli* HS.

Литература

- Trachtman H, Austin C, Lewinski M, Stahl R. A. Renal and neurological involvement in typical Shiga toxin-associated HUS. *Nature Reviews Nephrology*. 2012; (8): 658–69.
- Croxen MA, Law RJ, Scholz R, Keeney KM, Wlodarska M, Finlay BB. Recent advances in understanding enteric pathogenic *Escherichia coli*. *Clinical Microbiology Reviews*. 2013; (26): 822–80.
- Tarr G, Shringi S, Phipps AI, Besser TE, Mayer J, Oltean HN, et al. Geogenomic Segregation and Temporal Trends of Human Pathogenic *Escherichia coli* O157:H7, Washington, USA, 2005–2014. *Emerg Infect Dis*. 2018; 24 (1): 32–9.
- Ivanova EI, Rychkova LV, Nemchenko UM, Bukharova EV, Savelkaeva MV, Dzhioeva YuP. The Structure of the Intestinal Microbiota of the Intestine and the Frequency of Detection of Pathogenicity Genes (*stx1*, *stx2*, *bfp*) in *Escherichia coli* with Normal Enzymatic Activity Isolated from Children during the First Year of Life. *Molecular Genetics, Microbiology and Virology*. 2017; 32 (1): 42–8.
- Holmes A, Dallman TJ, Shabaan S, Hanson M, Allison L. Validation of Whole-Genome Sequencing for Identification and Characterization of Shiga Toxin-Producing *Escherichia coli* To Produce Standardized Data To Enable Data Sharing. *J Clin Microbiol*. 2018; 56 (3): e01388-17.
- Wong CS, Mooney JC, Brandt JR, Staples AO, Jelacic S, Boster DR, et al. Risk factors for the hemolytic uremic syndrome in children infected with *Escherichia coli* O157:H7: a multivariable analysis. *Clin Infect Dis*. 2012; 55 (1): 33–41.
- Freedman SB, Xie J, Neufeld MS, Hamilton WL, Hartling L, Tarr PI. Shiga toxin-producing *Escherichia coli* infection, antibiotics, and risk of developing hemolytic uremic syndrome: A meta-analysis. *Clinical Infectious Diseases*. 2016; (62): 1251–58.
- Pirnay JP, Blasdel BG, Bretaudeau L, Buckling A, Chanishvili N, Clark JR, et al. Quality and safety requirements for sustainable phage therapy products. *Pharm Res*. 2015; 32 (7): 2173–79.
- Qadir MI. Review: phage therapy: a modern tool to control bacterial infections. *Pak J Pharm Sci*. 2015; 28 (1): 265–70.
- Назаров П. А. Альтернативы антибиотикам: литические ферменты бактериофагов и фаговая терапия. *Вестник РГМУ*. 2018; (1).
- Rusconi B, Sanjar F, Koenig SS, Mammel MK, Tarr PI, Eppinger M. Whole Genome Sequencing for Genomics-Guided Investigations of *Escherichia coli* O157:H7 Outbreaks. *Front Microbiol*. 2016; 7: 985.
- Nakata A, Amemura M, Makino KJ. Unusual nucleotide arrangement with repeated sequences in the *Escherichia coli* K-12 chromosome. *Bacteriol*. 1989; 171 (6): 3553–56.
- Bolotin A, Quinquis B, Sorokin A, Ehrlich SD. Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of

- extrachromosomal origin. *Microbiology* 2005; 151: 2551–61.
14. Brouns SJ, Jore MM, Lundgren M, Westra ER, Slijkhuys RJ, Snijders AP, et al. CRISPR RNAs guide antiviral defense in prokaryotes. *Science*. 2008; (321): 960–4.
 15. Джагаров Д. Э. Умные ножницы для ДНК. *Химия и жизнь*. 2014; (7): 6–9.
 16. Westra ER, Buckling A, Fineran PC. CRISPR-Cas systems: beyond adaptive immunity. *Nat Rev Microbiol*. 2014; 12 (5): 317–26.
 17. Makarova KS, Koonin EV. Annotation and Classification of CRISPR-Cas Systems. *Methods Mol Biol*. 2015; (1311): 47–75.
 18. Modell JW, Jiang W, Marraffini LA. CRISPR-Cas systems exploit viral DNA injection to establish and maintain adaptive immunity. *Nature*. 2017; 544 (7648): 101–4.
 19. Gasiunas G, Sinkunas T, Siksnys V. Molecular mechanisms of CRISPR-mediated microbial immunity. *Cellular and Molecular Life Sciences*. 2014; 71 (3): 449–65.
 20. Peretolchina NP, Dzhioev YP, Borisenko AY, Paramonov AI, Voskresenskaya EA, Stepanenko LA, et al. Bioinformatic search and screening of phages and plasmids via spacer sites of *Yersinia pseudotuberculosis* YPIII CRISPR/Cas System. *Materials 15th Medical Biodefence Conference*, 26–29 April 2016; Minich: 43–44.
 21. Zlobin VI, Dzhioev YuP, Peretolchina NP, Borisenko AY, Stepanenko LA, Yingchen Wang, et al. Prospects to Enhance Phage Therapy by Looking At CRISP Fingerprints in Bacterial Populations. *Current Trends in Biomedical Engineering & Biosciences*. 2018; 10 (5): 1–3.
 22. Abedon S, Kuhl S, Blasdel B. Phage treatment of human infections. *Bacteriophage*. 2011; (1): 66–85.
 23. Степаненко Л. А., Джиоев Ю. П., Борисенко А. Ю., Колбасеева О. В., Злобин В. И., Малов И. В. Поиск фагов и плазмид через спейсерные сайты CRISPR/CAS-системы *Neisseria meningitidis* fdaargos_214. *Журнал инфектологии*. 2018; 10 (S1): 30–31.
 24. Борисенко А. Ю., Джиоев Ю. П., Перетолчина Н. П., Степаненко Л. А., Кузьминова В. М., Колбасеева О. В. и др. Биоинформационный поиск и скрининг бактериофагов через спейсеры CRISPR/CAS-системы штамма *Staphylococcus aureus* Mu3. *Актуальные проблемы науки прибайкалья*. 2017; (2): 45–9.
 25. Levine M. M., Rennels M. B. *E. coli* colonization factor antigen in diarrhoea. *Lancet*. 1978; 2 (8088): 534.
 26. Rasko DA, Rosovitz MJ, Myers GSA, Mongodin EF, Fricke WF, Gajer P, et al. The Pangenome Structure of *Escherichia coli*: Comparative Genomic Analysis of *E. coli* Commensal and Pathogenic Isolates. *J Bacteriol*. 2008; 190 (20): 6881–93.
 27. Abby SS, Néron B, Ménager H, Touchon M, Rocha Eduardo PC. MacSyFinder: A Program to Mine Genomes for Molecular Systems with an Application to CRISPR-Cas Systems. *PLoS One*. 2014; 9 (10): e110726.
 28. Gaj T, Gersbach CA, Barbas CF. ZFN, TALEN and CRISPR/Cas-based methods for genome engineering. *Trends Biotechnol*. 2013; 31: 397–405.
 29. Qi LS, Larson MH, Gilbert LA, Doudna JA, Weissman JS, Arkin AP, et al. Repurposing CRISPR as an RNA guided platform for sequence-specific control of gene expression. *Cell*. 2013; 152 (5): 1173–83.
 30. Grissa I, Vergnaud G, Poursel C. CRISPRFinder: a web tool to identify clustered regularly interspaced short palindromic repeats. *Nucleic Acids Res*. 2007; (35) (Web Server issue): W52–7.
 31. Johnson M, Zaretskaya I, Raytselis Y, Merezhuk Y, McGinnis S, Madden TL. NCBI BLAST: a better web interface. *NCBI BLAST: a better web interface*. *Nucleic Acids Res*. 2008; (1); 36 (Web Server issue): W5–9.

References

1. Trachtman H, Austin C, Lewinski M, Stahl R. A. Renal and neurological involvement in typical Shiga toxin-associated HUS. *Nature Reviews Nephrology*. 2012; (8): 658–69.
2. Croxen MA, Law RJ, Scholz R, Keeney KM, Wlodarska M, Finlay BB. Recent advances in understanding enteric pathogenic *Escherichia coli*. *Clinical Microbiology Reviews*. 2013; (26): 822–80.
3. Tarr G, Shringi S, Phipps AI, Besser TE, Mayer J, Oltean HN, et al. Geogenomic Segregation and Temporal Trends of Human Pathogenic *Escherichia coli* O157:H7, Washington, USA, 2005–2014. *Emerg Infect Dis*. 2018; 24 (1): 32–9.
4. Ivanova EI, Rychkova LV, Nemchenko UM, Bukharova EV, Savelkaeva MV, Dzhioeva YuP. The Structure of the Intestinal Microbiota of the Intestine and the Frequency of Detection of Pathogenicity Genes (*stx1*, *stx2*, *bfp*) in *Escherichia coli* with Normal Enzymatic Activity Isolated from Children during the First Year of Life. *Molecular Genetics, Microbiology and Virology*. 2017; 32 (1): 42–8.
5. Holmes A, Dallman TJ, Shabaan S, Hanson M, Allison L. Validation of Whole-Genome Sequencing for Identification and Characterization of Shiga Toxin-Producing *Escherichia coli* To Produce Standardized Data To Enable Data Sharing. *J Clin Microbiol*. 2018; 56 (3): e01388–17.
6. Wong CS, Mooney JC, Brandt JR, Staples AO, Jelacic S, Boster DR, et al. Risk factors for the hemolytic uremic syndrome in children infected with *Escherichia coli* O157:H7: a multivariable analysis. *Clin Infect Dis*. 2012; 55 (1): 33–41.
7. Freedman SB, Xie J, Neufeld MS, Hamilton WL, Hartling L, Tarr PI. Shiga toxin-producing *Escherichia coli* infection, antibiotics, and risk of developing hemolytic uremic syndrome: A meta-analysis. *Clinical Infectious Diseases*. 2016; (62): 1251–58.
8. Pirnay JP, Blasdel BG, Bretaudeau L, Buckling A, Chanishvili N, Clark JR, et al. Quality and safety requirements for sustainable phage therapy products. *Pharm Res*. 2015; 32 (7): 2173–79.
9. Qadir MI. Review: phage therapy: a modern tool to control bacterial infections. *Pak J Pharm Sci*. 2015; 28 (1): 265–70.
10. Nazarov PA. Аlтернативу антибиотикам: литические ферменты бактериофагов и фаговая терапия. *Vestnik RGMU*. 2018; (1).
11. Rusconi B, Sanjar F, Koenig SS, Mammel MK, Tarr PI, Eppinger M. Whole Genome Sequencing for Genomics-Guided Investigations of *Escherichia coli* O157:H7 Outbreaks. *Front Microbiol*. 2016; 7: 985.
12. Nakata A, Amemura M, Makino KJ. Unusual nucleotide arrangement with repeated sequences in the *Escherichia coli* K-12 chromosome. *Bacteriol*. 1989; 171 (6): 3553–56.
13. Bolotin A, Quinquis B, Sorokin A, Ehrlich SD. Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiology* 2005; 151: 2551–61.
14. Brouns SJ, Jore MM, Lundgren M, Westra ER, Slijkhuys RJ, Snijders AP, et al. CRISPR RNAs guide antiviral defense in prokaryotes. *Science*. 2008; (321): 960–4.
15. Dzhagarov DJe. Умные ножницы для ДНК. *Химия и жизнь*. 2014; 7: 6–9.
16. Westra ER, Buckling A, Fineran PC. CRISPR-Cas systems: beyond adaptive immunity. *Nat Rev Microbiol*. 2014; 12 (5): 317–26.
17. Makarova KS, Koonin EV. Annotation and Classification of CRISPR-Cas Systems. *Methods Mol Biol*. 2015; (1311): 47–75.
18. Modell JW, Jiang W, Marraffini LA. CRISPR-Cas systems exploit viral DNA injection to establish and maintain adaptive immunity. *Nature*. 2017; 544 (7648): 101–4.
19. Gasiunas G, Sinkunas T, Siksnys V. Molecular mechanisms of CRISPR-mediated microbial immunity. *Cellular and Molecular Life Sciences*. 2014; 71 (3): 449–65.
20. Peretolchina NP, Dzhioev YP, Borisenko AY, Paramonov AI, Voskresenskaya EA, Stepanenko LA, et al. Bioinformatic search and screening of phages and plasmids via spacer sites of *Yersinia pseudotuberculosis* YPIII CRISPR/Cas System. *Materials 15th Medical Biodefence Conference*, 26–29 April 2016; Minich: 43–44.
21. Zlobin VI, Dzhioev YuP, Peretolchina NP, Borisenko AY, Stepanenko LA, Yingchen Wang, et al. Prospects to Enhance

- Phage Therapy by Looking At CRISP Fingerprints in Bacterial Populations. *Current Trends in Biomedical Engineering & Biosciences*. 2018; 10 (5): 1–3.
22. Abedon S, Kuhl S, Blasdel B. Phage treatment of human infections. *Bacteriophage*. 2011; (1): 66–85.
 23. Stepanenko LA, Dzhioev JuP, Borisenko AJu, Kolbaseeva OV, Zlobin VI, Malov IV. Poisk fagov i plazmid cherez spejsernye sajty CRISPR/CAS-sistemy *Neisseria meningitidis* fdaargos_214. *Zhurnal infektologii*. 2018; 10 (S1): 30–31.
 24. Borisenko AJu, Dzhioev JuP, Peretolchina NP, Stepanenko LA, Kuz'minova VM, Kolbaseeva OV i dr. Bioinformacionnyj poisk i skringing bakteriofagov cherez spejsery CRISPR/CAS-sistemy shtamma *Staphylococcus aureus* Mu3. *Aktual'nye problemy nauki pribajkal'ja*. 2017; 2: 45–9.
 25. Levine M. M., Rennels M. B. *E. coli* colonization factor antigen in diarrhoea. *Lancet*. 1978; 2 (8088): 534.
 26. Rasko DA, Rosovitz MJ, Myers GSA, Mongodin EF, Fricke WF, Gajer P, et al. The Pangenome Structure of *Escherichia coli* Pathogenic Isolates. *J Bacteriol*. 2008; 190 (20): 6881–93.
 27. Abby SS, Néron B, Ménager H, Touchon M, Rocha Eduardo PC. MacSyFinder: A Program to Mine Genomes for Molecular Systems with an Application to CRISPR-Cas Systems. *PLoS One*. 2014; 9 (10): e110726.
 28. Gaj T, Gersbach CA, Barbas CF. ZFN, TALEN and CRISPR/Cas-based methods for genome engineering. *Trends Biotechnol*. 2013; 31: 397–405.
 29. Qi LS, Larson MH, Gilbert LA, Doudna JA, Weissman JS, Arkin AP, et al. Repurposing CRISPR as an RNA guided platform for sequence-specific control of gene expression. *Cell*. 2013; 152 (5): 1173–83.
 30. Grissa I, Vergnaud G, Pourcel C. CRISPRFinder: a web tool to identify clustered regularly interspaced short palindromic repeats. *Nucleic Acids Res*. 2007; (35) (Web Server issue): W52-7.
 31. Johnson M, Zaretskaya I, Raytselis Y, Merezhuk Y, McGinnis S, Madden TL. NCBI BLAST: a better web interface. *NCBI BLAST: a better web Interface*. *Nucleic Acids Res*. 2008; (1); 36 (Web Server issue): W5-9.