

УДК 577.152.314

# Клонирование и характеристика новой сайт-специфической метилзависимой эндонуклеазы ElmI, узнающей и расщепляющей C5-метилованную последовательность ДНК 5'-G(5mC)^NG(5mC)-3'

В. А. Чернухин\*, Д. А. Гончар, М. А. Абдурашитов, О. А. Беличенко, В. С. Дедков, Н. А. Михненкова, Е. Н. Ломаковская, С. Г. Удальева, С. Х. Дегтярёв

ООО «СибЭнзайм», 630117, Новосибирск, ул. Тимакова, 2/12

\*E-mail: valera@sibenzyme.ru

Поступила в редакцию 13.08.2015

Принята к печати 23.11.2015

**РЕФЕРАТ** В результате поиска аминокислотных последовательностей, гомологичных последовательности MD-эндонуклеазы BisI, в геномах энтеробактерий выявлены предполагаемые открытые рамки считывания MD-эндонуклеаз. Высокая консервативность нуклеотидных последовательностей данных рамок у энтеробактерий разных родов (*Escherichia*, *Klebsiella* и *Cronobacter*) позволила создать праймеры для проведения ПЦР-скрининга на ДНК энтеробактерий, собранных из природных источников. С использованием геномной ДНК *E. coli* LM N17 в качестве матрицы амплифицирован и встроен в вектор pMTL22 фрагмент ДНК длиной около 440 п.н. В штамме *E. coli* ER2267, трансформированном полученной конструкцией, выявлена эндонуклеазная активность. Из биомассы данного штамма хроматографическими методами получен препарат нового фермента ElmI. Оказалось, что ElmI подобно BisI специфично расщепляет метилированную последовательность 5'-GCNGC-3' перед центральным нуклеотидом «N» при условии, что она содержит два остатка 5-метилцитозина. Однако, в отличие от BisI, ElmI с большей эффективностью расщепляет эту последовательность при наличии в ней более двух метилированных остатков цитозина.

**КЛЮЧЕВЫЕ СЛОВА** анализ метилома, клонирование генов ДНК-эндонуклеаз, метилзависимые ДНК-эндонуклеазы, MD-эндонуклеазы, эпигенетика.

**СПИСОК СОКРАЩЕНИЙ** MD-эндонуклеаза – сайт-специфическая метилзависимая ДНК-эндонуклеаза.

## ВВЕДЕНИЕ

Сайт-специфические метилзависимые эндонуклеазы (MD-эндонуклеазы) узнают и расщепляют ДНК по определенным метилированным последовательностям и не расщепляют неметилированную ДНК. За последние 9 лет описано более 10 прототипов этих ферментов с разными сайтами узнавания, которые расщепляются лишь при условии C5-метилирования определенных остатков цитозина в них.

MD-эндонуклеазы, в отличие от эндонуклеаз рестрикции, узнают не только определенную нуклеотидную последовательность и положение гидролизуемой связи относительно этой последовательности, но и строго определенный узор метилирования. Поэтому разные MD-эндонуклеазы даже со сходным сайтом узнавания могут по-разному расщеплять

ДНК в зависимости от характера ее метилирования. Это хорошо видно на примере ферментов, узнающих метилированную последовательность 5'-GCNGC-3'. Так, среди ферментов с таким сайтом узнавания и с гидролизуемой позицией после центрального нуклеотида «N» BslI [1] расщепляет эту последовательность при наличии в ней, как минимум, двух, а PkrI [2] – трех остатков 5-метилцитозина. Среди ферментов, гидролизующих связь перед центральным нуклеотидом «N», MD-эндонуклеаза BisI [3] расщепляет 5'-GCNGC-3' при наличии двух 5-метилцитозинов (и намного менее эффективно в случае одного), тогда как для GluI [4] нужны четыре остатка 5-метилцитозина.

Нами описан новый представитель данной группы ферментов – MD-эндонуклеаза ElmI, узнающая

метилованную последовательность 5'-GCNGC-3' и расщепляющая ее перед центральным нуклеотидом «N» (с образованием 5'-выступающих однонуклеотидных концов), если содержит не менее двух остатков 5-метилцитозина, причем активность фермента увеличивается на порядок, если в сайте узнавания метилированы три или четыре остатка цитозина.

## ЭКСПЕРИМЕНТАЛЬНАЯ ЧАСТЬ

### Проведение ПЦР-скрининга на ДНК энтеробактерий из природных образцов и получение плазмиды pElmI, несущей ген новой MD-эндонуклеазы

Колиформные бактерии выделяли из природных образцов (сточных вод) на селективной среде Эндо согласно [5]. Из посевов каждого образца отбирали по 8–20 штаммов (серии LM, LT, LP и LV), отличающихся по морфологическим характеристикам.

Из полученных штаммов выделяли хромосомную ДНК, на которой проводили ПЦР-скрининг с использованием праймеров, приведенных ниже. Фрагмент, амплифицированный с ДНК одного из штаммов дикого типа, встраивали в плазмиду pMTL22 [6] по сайтам эндонуклеаз рестрикции FauNDI и BglII.

После трансформации клеток *E. coli* ER2267 клоны, несущие целевую плазмиду, названную pElmI, высевали на чашки Петри с агаризованной средой LB с добавлением ампициллина (50 мкг/мл). Клоны подращивали в течение ночи при 37°C, после чего пересевали на отдельные чашки с ампициллином (100 мкг/мл) и вновь выдерживали в течение ночи с целью дальнейшего анализа.

### Наработка биомассы рекомбинантного штамма-продуцента ElmI и тестирование целевой активности

Рекомбинантный клон штамма *E. coli* ER2267, трансформированного плазмидой pElmI, переносили с чашки Петри микробиологической петлей во флакон с 200 мл бульона LB, содержащего 100 мкг/мл ампициллина. Инокулят подращивали в течение ночи на термостатированной качалке при 37°C и 120 об/мин.

Далее по 5 мл инокулята засевали в 20 флаконов с 200 мл бульона LB, содержащего 100 мкг/мл ампициллина и 0.5 мМ изопрропил-β-D-тиогалактопиранозид (IPTG).

Культуру подращивали в термостатированной качалке при 120 об/мин в течение 10 ч, после чего для анализа активности фермента отбирали аликвоту выросшей культуры (1 мл) и переносили ее в пробирку Eppendorf на 1.5 мл. Клетки осаждали

в настольной центрифуге 5416 Eppendorf (Eppendorf GmbH, Германия) при 12000 об/мин в течение 2 мин. Супернатант убирали, а осадок суспендировали в 0.2 мл лизирующего буфера (10 мМ трис-HCl, pH 8.5, 0.1 мг/мл лизоцима, 0.5 М NaCl, 1 мМ EDTA, 0.1% Тритон X-100).

Активность фермента в лизате определяли в 20 мкл реакционной смеси, содержащей в качестве субстрата ДНК плазмиды pFsp4HI3 [4], предварительно линеаризованной рестриктазой DriI в SE-буфере «W» (10 мМ трис-HCl (pH 8.5 при 25°C), 10 мМ MgCl<sub>2</sub>, 100 мМ NaCl, 1 мМ дитиотреитол) в течение 2 ч при 37°C. За 1 ед. акт. ElmI принимали количество фермента, достаточное для исчерпывающего гидролиза 1 мкг ДНК pFsp4HI3 в SE-буфере «W» за 2 ч при 37°C. Наличие или отсутствие гидролиза субстратной ДНК определяли методом электрофореза в 1% агарозном геле.

Клетки всей наработанной биомассы осаждали на центрифуге J2-21 (Beckman, США) в течение 30 мин (ротатор JA-10, 8000 об/мин) и замораживали.

### Получение препарата фермента ElmI

Все процедуры выделения и очистки препарата фермента проводили при 4°C с использованием следующих растворов:

- буфер А – 10 мМ трис-HCl, pH 7.5; 0.1 мМ EDTA; 7 мМ 2-меркаптоэтанол;
- буфер В – 10 мМ К-фосфатный буфер, pH 7.4; 0.1 мМ EDTA; 7 мМ 2-меркаптоэтанол.

Ферментативную активность определяли с использованием в качестве субстрата ДНК pFsp4HI3 [4] в SE-буфере «W», которую гидролизовали в течение 15 мин в 20 мкл реакционной смеси путем добавления аликвот (1 мкл) из фракций хроматографического профиля.

**Получение грубого экстракта.** Биомассу (8 г) суспендировали в 30 мл буфера А, содержащего 0.2 М NaCl, 1 мМ фенолметилсульфонилфторид (PMSF), 0.1% Тритон X-100; 0.1 мг/мл лизоцима. Клетки разрушали ультразвуком на дезинтеграторе Soniprep 150 (MSE, Великобритания), снабженного адаптером диаметром 2 см, при амплитуде 20 мкм, 4 раза по 1 мин с интервалом в 1 мин для охлаждения суспензии в водяной бане.

Экстракт осветляли центрифугированием при 15000 об/мин в течение 30 мин (ротаторе JA-20, центрифуга J2-21, Beckman, США).

**Хроматография на фосфоцеллюлозе P-11.** Грубый экстракт разбавляли в 2 раза буфером А и наносили на колонку с фосфоцеллюлозой P-11 объемом 30 мл, уравновешенную буфером А, содержащим

0.1 М NaCl, промывали 160 мл этого же буфера. Адсорбированный материал вымывали линейным градиентом NaCl от 0.1 до 1 М объемом 800 мл в буфере А. Собирали 30 фракций, из которых объединяли фракции с 16 по 22 (от 0.35 до 0.47 М NaCl), содержащие пик целевой активности.

*Хроматография на гидроксилпатите.* Объединенные фракции наносили на колонку с гидроксилпатитом объемом 2 мл, уравновешенную буфером В, содержащим 0.05 М NaCl, промывали 10 мл этого же буфера. Адсорбированный материал вымывали линейным градиентом К-фосфатного буфера, рН 7.4, от 0.01 до 0.1 М, содержащим 0.05 М NaCl, объемом 30 мл. Собирали градиент из 20 фракций, из которых объединяли фракции с 8 по 12 (от 0.044 до 0.056 М калий-фосфата), содержащие пик активности ElmI. Полученные фракции диализовали в течение 1 ч в 300 мл буфера А.

*Хроматография на гепарин-агарозе.* Диализованные фракции наносили на колонку с гепарин-агарозой объемом 2 мл, уравновешенную в буфере В, содержащим 0.05 М NaCl, и промывали 4 мл этого же буфера. Адсорбированный материал вымывали линейным градиентом NaCl (0.05 до 1 М) в буфере В объемом 30 мл. Собирали 20 фракций, из которых объединяли фракции 12 и 13 (с градиентом от 0.62 до 0.67 М NaCl), содержащие целевую активность.

*Концентрирование, проверка активности и хранение препарата.* Полученные фракции диализовали в течение 20 ч против 15-кратного объема буфера В с 50% глицерином, 0.2 М NaCl и хранили при -20°C.

Активность проверяли добавлением восьми последовательных двукратных разведений препарата фермента (2, 1, 1/2 мкл и т.д.) к 20 мкл реакционной смеси, содержащей 0.5 мкг ДНК рFsp4HI3/DriI в SE-буфере «W» с последующей инкубацией в течение 2 ч при 37°C. Для получения разбавленных препаратов ферментов использовали SE-буфер для разбавления ферментов «В100» (10 мМ трис-НСl (рН 7.6 при 25°C), 50 мМ КСl, 0.1 мМ EDTA, 200 мкг/мл БСА, 1 мМ дитиотреитол, 50% глицерин). Реакцию останавливали, добавляя в каждую реакционную смесь 1 мкл стоп-буфера следующего состава: 50% глицерин, 10 мМ EDTA, 0.2% бромфеноловый синий.

**Нуклеотидную последовательность ДНК** определяли по методу Сэнгера на автоматическом секвенаторе ABI 3130xI Genetic Analyzer (Applied Biosystems, США) в соответствии с рекомендациями производителя.

**Препараты ферментов, ДНК, дезокси-нуклеозид-трифосфатов и синтетических олигонуклеотидов,**

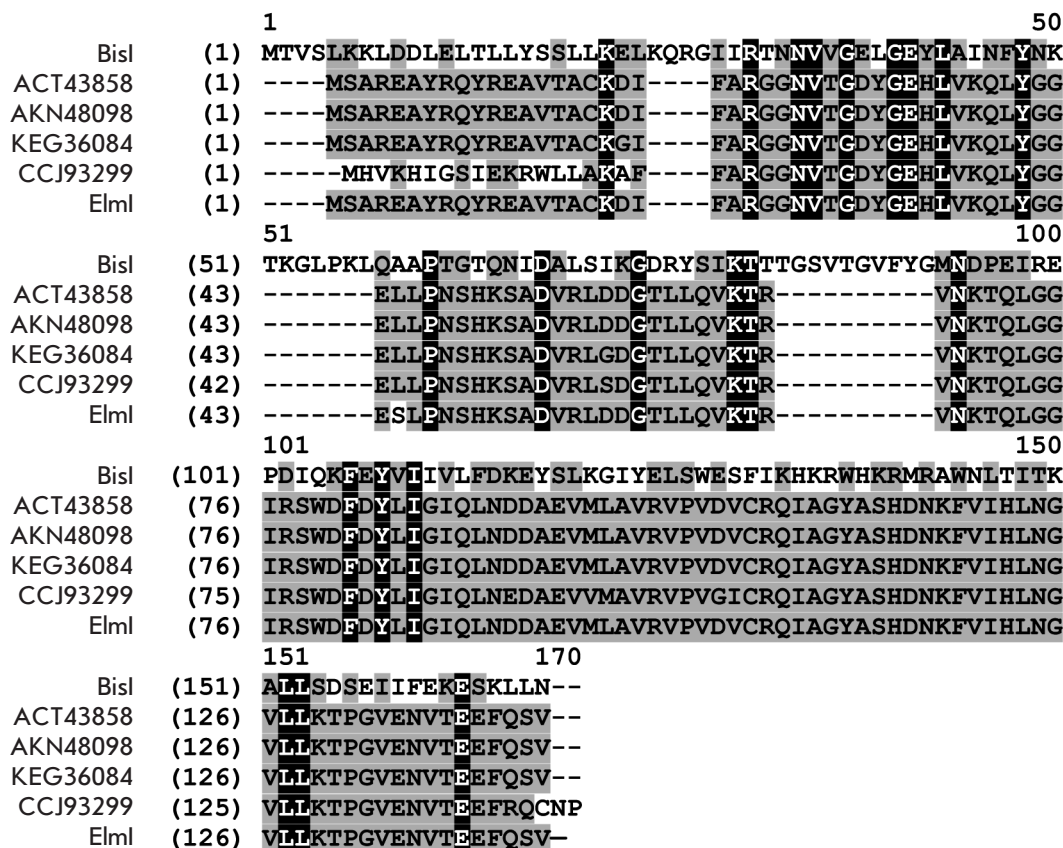
а также маркеры молекулярных масс (1 kb Ladder и Lambda/StyI), использованные в работе, произведены в ООО «СибЭнзайм» (Россия).

## РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

### **Клонирование гена новой MD-эндонуклеазы ElmI и сравнительный анализ нуклеотидных и аминокислотных последовательностей**

Описанные нами ранее MD-эндонуклеазы обнаружены в бактериях различных таксономических групп, но в основном среди представителей семейств Microbacteriaceae и Bacillaceae. Скрининг активности в лизатах клеток, проведенный ранее, не позволил нам выявить подобные сайт-специфические ферменты в штаммах энтеробактерий, что может свидетельствовать либо об отсутствии, либо о крайне низкой активности MD-эндонуклеаз у бактерий этой группы. Поэтому для решения этого вопроса была предпринята попытка найти гомологичные белки энтеробактерий не биохимическим, а биоинформатическим способом.

В базе данных аминокислотных последовательностей энтеробактерий с помощью программы PSI-BLAST (<https://blast.ncbi.nlm.nih.gov>) проведен поиск последовательностей, гомологичных ранее определенной нами MD-эндонуклеазе BisI (GenBank AJW87312) [7]. После выполнения двух итераций поиска у энтеробактерий обнаружили около 50 аминокислотных последовательностей, обладающих гомологией (32–48% сходства и 17–30% идентичности) с последовательностью BisI, причем функция ни одного из выявленных белков-гомологов еще не установлена. Нуклеотидные последовательности соответствующих генов извлечены из базы данных GenBank и сравнены между собой. Показано, что в выборке присутствуют две группы генов, нуклеотидные последовательности которых высокогомологичны между собой. Первая группа включает гены четырех гипотетических белков длиной 143–144 аминокислотных остатка бактерий родов *Escherichia* (ACT43858 и AKN48098), *Cronobacter* (CCJ93299) и *Klebsiella* (KEG36084), в скобках указаны идентификационные номера в базе данных GenBank. Сравнение этих генов между собой показало, что степень их идентичности составляет 93–99%. Вторая группа представлена генами энтеробактерий, кодирующими белки из 290 аминокислотных остатков, N-концевая часть которых гомологична белку BisI (номера белков в GenBank: KFC97828, WP\_000794335, WP\_000794336, WP\_000794337, WP\_001655794, WP\_004952390, WP\_008806407, WP\_021557167, WP\_025912430, WP\_032653240, WP\_032671961 и WP\_033070923). Степень идентич-



**Рис. 1.** Выравнивание аминокислотных последовательностей BisI и Elml и наиболее гомологичных белков энтеробактерий. Аминокислотные последовательности обозначены номерами в GenBank как описано в тексте. Белым цветом на черном фоне показаны аминокислотные остатки, идентичные во всех представленных аминокислотных последовательностях. Черным цветом на сером фоне выделены аминокислоты, сходные по физико-химическим свойствам

ности нуклеотидных последовательностей в данной группе составляет 83–99%.

На рис. 1 приведено множественное выравнивание аминокислотных последовательностей четырех высокогомологичных белков энтеробактерий из первой группы с последовательностью BisI. Также приведена последовательность обнаруженной путем ПЦР-скрининга эндонуклеазы ElmI (см. описание ниже).

Множественное выравнивание соответствующих генов энтеробактерий первой группы, приведенное на рис. 2, показало высокую степень идентичности и на уровне нуклеотидного состава, несмотря на принадлежность хозяев к разным родам. В это выравнивание добавлена также последовательность гена *elmI*, полученная в ходе дальнейшей работы.

Высокий уровень идентичности последовательностей генов из двух вышеупомянутых групп, который распространяется и на их концевые участки, позволил подобрать праймеры для ПЦР-скрининга

аналогичных генов в штаммах дикого типа. Для поиска генов, кодирующих белки, родственные белкам представителей первой группы, использовали праймеры, содержащие сайты узнавания эндонуклеаз рестрикции FauNDI и VamHI, для встраивания ПЦР-фрагментов в плазмидный вектор:

Esp-1 5'-CCCCATATGAGTGCACGTGAAGCAT-ATC-3'

Esp-2 5'-CGCGGATCCTTAGGGATTACACTGAC-TGAAACTCTTC-3'.

Для ПЦР-скрининга генов, аналогичных генам второй группы, были синтезированы праймеры:

Esp-3 – 5'-TTGAAAAATAATCATTTAACATCATATG-3'

Esp-4 – 5'-TCACTCCAGAACGCTGATAAGTTT-3'.

Из 64 штаммов колиформных бактерий, обнаруженных в сточных водах, была выделена геномная ДНК, которая служила матрицей для проведения ПЦР. Стоит отметить, что амплификация с исполь-

зованием праймеров Esp-3 и Esp-4 (вторая группа генов) не привела к появлению фрагмента ожидаемой длины (~870 п.н.) ни с одной из матричных ДНК. При использовании пары праймеров Esp-1 и Esp-2 (первая группа генов) положительный результат получен для одного из образцов ДНК – обнаружен ПЦР-фрагмент ожидаемой длины ~430 п.н.

Этот амплифицированный фрагмент был обработан рестриктазами FauNDI и BamHI и лигирован в вектор pMTL22, предварительно расщепленный FauNDI и BglII. Полученной плазмидой, названной pElmI, трансформировали клетки *E. coli* ER2267.

Определение таксономической специфичности исходного штамма, с геномной ДНК которого получен амплифицированный фрагмент, проведено с использованием традиционных биохимических и морфологических критериев [8], а также с помощью анализа структуры фрагмента 16S рРНК по программе BLAST [9]. Исходный природный штамм-продуцент обозначен как *E. coli* LM N17. Продуцируемая этим штаммом сайт-специфическая ДНК-эндонуклеаза названа ElmI.

Проведено секвенирование встроенного в плазмиду pElmI ПЦР-фрагмента. Нуклеотидная последовательность этого фрагмента длиной 432 п.н. депонирована в базу данных GenBank под номером LN869919. Последовательность начинается со стартового кодона ATG и заканчивается стоп-кодоном TAA (причем в гипотетической рамке трансляции отсутствуют другие стоп-кодона), поэтому может рассматриваться как гипотетическая рамка трансляции ДНК-эндонуклеазы ElmI, а ген, кодирующий этот белок, – *elmI*.

Согласно результатам сравнительного анализа секвенированного фрагмента ген *elmI* имеет практически такую же последовательность, как и гены, кодирующие полипептиды из ближайших гомологов – штаммов *E. coli* BL21(DE3) (ACT43858) и C41(DE3) (AKN48098). Единственная найденная замена – цитозин в положении 131 в гене *elmI* вместо тимина в гомологичных генах (рис. 2).

Выведенная аминокислотная последовательность эндонуклеазы ElmI соответственно отличается от ближайших гомологов из штаммов *E. coli* BL21(DE3) и C41(DE3) одним аминокислотным остатком: в позиции 44 у ElmI находится серин, тогда как у ближайших гомологов (*E. coli* BL21(DE3) и C41(DE3)) в аналогичной позиции стоит лейцин (рис. 1). В то же время сходство аминокислотной последовательности ElmI и BisI составляет около 50% при том, что количество идентичных аминокислот всего 112%. Таким образом, клонированный фрагмент ДНК, выявленный в результате ПЦР-скрининга и представленный предполагаемым геном метила-

зисимой ДНК-эндонуклеазы ElmI, оказывается высокоомологичным участкам геномных ДНК из хорошо известных штаммов *E. coli*.

### Определение специфичности новой MD-эндонуклеазы ElmI

В лизате клонов *E. coli* ER2267, несущих плазмиду pElmI, в отличие от исходного штамма, обнаружена эндонуклеазная активность, и один из клонов *E. coli* pElmI был выбран для наработки биомассы и выделения фермента.

По схеме, описанной в разделе «Экспериментальная часть», наработано 8 г биомассы *E. coli* pElmI. В результате хроматографической очистки из данной биомассы получено 3 мл препарата фермента ElmI с концентрацией 4 ед. акт./мкл.

С целью определения сайт-специфичности ElmI различные субстратные ДНК расщепили в предварительно установленных оптимальных условиях: 37°C, реакционный SE-буфер «W», 20 мкл реакционной смеси, содержащей 0.5 мкг субстратной ДНК, в течение 2 ч. ДНК расщепляли контрольным ферментом BisI в тех же условиях, но с использованием реакционного SE-буфера «Y».

В качестве субстрата для определения специфичности фермента ElmI использовали ДНК плазмид, содержащих гены различных ДНК-метилтрансфераз. Благодаря активности этих генов в штаммах *E. coli*, из которых выделены плазмиды, ДНК-субстраты оказывались модифицированными соответствующими ДНК-метилтрансферазами и несли определенный узор метилирования.

В качестве метилированных ДНК-субстратов использовали следующие плазмиды:

1) рМНраII, содержащая ген, кодирующий ДНК-метилтрансферазу НраII, вследствие чего в данной плазмиде метилирован первый цитозин во всех последовательностях 5'-CCGG-3' на обеих цепях ДНК [10];

2) рМНаеIII, содержащая ген, кодирующий ДНК-метилтрансферазу НаеIII, вследствие чего в данной плазмиде метилирован первый цитозин во всех последовательностях 5'-GGCC-3' на обеих цепях ДНК [11];

3) рНспAI2, содержащая ген, кодирующий ДНК-метилтрансферазу НспAI, вследствие чего в плазмиде метилирован первый цитозин во всех последовательностях 5'-GCGC-3' на обеих цепях ДНК [12]. Кроме того, данная плаزمида содержит гиперметилированный участок:

5'-G(5mC)G(5mC)G(5mC)GC-3'  
3'-CG(5mC)G(5mC)G(5mC)G-5';

4) рНспAI4, также содержащая ген, кодирующий ДНК-метилтрансферазу НспAI, вследствие чего в плазмиде метилирован первый цитозин во всех по-

ЭКСПЕРИМЕНТАЛЬНЫЕ СТАТЬИ

<i>elml</i>	(1)	ATGAGTGCACGTGAAGCATATC	GGCAGTATCGAGAAGCGGTGACTGCTTGCAAAG
ACT43858	(1)	ATGAGTGCACGTGAAGCATATC	GGCAGTATCGAGAAGCGGTGACTGCTTGCAAAG
KEG36084	(1)	ATGAGTGCACGTGAAGCATATC	GGCAGTATCGAGAAGCGGTGACTGCTTGCAAAG
AKN48098	(1)	ATGAGTGCACGTGAAGCATATC	GGCAGTATCGAGAAGCGGTGACTGCTTGCAAAG
CCJ93299	(1)	----GTGCACGTGAAGCATATC	GGCAGTATCGAGAAGCGGTGCTGCTTGCAAAG
<i>elml</i>	(56)	ACATTTTT-GCCCCGAGGTGGTAATGTC	ACTGGTGATTA_CGGTGAACATCTCGTAA
ACT43858	(56)	ACATTTTT-GCCCCGAGGTGGTAATGTC	ACTGGTGATTA_CGGTGAACATCTCGTAA
KEG36084	(56)	GCATTTTT-GCCCCGAGGTGGTAATGTC	ACTGGTGATTA_TGGTGAACATCTCGTAA
AKN48098	(56)	GCATTTTT-GCCCCGAGGTGGTAATGTC	ACTGGTGATTA_TGGTGAACATCTCGTAA
CCJ93299	(52)	GCATTTTTTGGCCCCGAGGTGGTAATGTC	ACTGGTGATTA_TGGTGAACATCTCGTAA
<i>elml</i>	(110)	AGCAACTTTTATGG_TGGTGAAT_C	ATTGCCAAACTCCCATAAAAAGCGCCGATGT_CAG
ACT43858	(110)	AGCAACTTTTATGG_TGGTGAAT_T	ATTGCCAAACTCCCATAAAAAGCGCCGATGT_CAG
KEG36084	(110)	AGCAACTTTTATGG_CGGTGAAT_T	ATTGCCAAACTCCCATAAAAAGCGCCGATGTAAAG
AKN48098	(110)	AGCAACTTTTATGG_CGGTGAAT_T	ATTGCCAAACTCCCATAAAAAGCGCCGATGTAAAG
CCJ93299	(107)	AGCAACTTTTATGG_GGGTGAAT_T	ATTGCCAAACTCCCATAAAAAGCGCCGATGTAAAG
<i>elml</i>	(165)	ATTAG_ACGATGG_CACGCT_GTTG	CAGGTCAAACCAGAGTCAATAAAAACCCAGTTA
ACT43858	(165)	ATTAG_ACGATGG_CACGCT_GTTG	CAGGTCAAACCAGAGTCAATAAAAACCCAGTTA
KEG36084	(165)	ATTAG_GCGATGG_CACGCT_ATTG	CAGGTCAAACCAGAGTCAATAAAAACCCAGTTA
AKN48098	(165)	ATTAG_GCGATGG_CACGCT_ATTG	CAGGTCAAACCAGAGTCAATAAAAACCCAGTTG
CCJ93299	(162)	ATTA_ACGATGGGACGCT_GTTG	CAGGTCAAACCAGAGTCAATAAAAACCCAGTTA
<i>elml</i>	(220)	GGCGGAATTCGCTCCTGGGATTTTCGATTACCTGATCGGT	TATTCAGCTTAATGACG
ACT43858	(220)	GGCGGAATTCGCTCCTGGGATTTTCGATTACCTGATCGGT	TATTCAGCTTAATGACG
KEG36084	(220)	GGCGGAATTCGCTCCTGGGATTTTCGATTACCTGATCGGT	TATTCAGCTTAATGACG
AKN48098	(220)	GGCGGAATTCGCTCCTGGGATTTTCGATTACCTGATCGGT	TATTCAGCTTAATGACG
CCJ93299	(217)	GGCGGAATTCGCTCCTGGGATTTTCGATTACCTGATCGG	CATTTCAGCTTAATGAAG
<i>elml</i>	(275)	ATGCCGAAGT_CATGTTGGCAGTTCGTGTCCCTGTAGATG	TTTGTGCGCCAAATTGC
ACT43858	(275)	ATGCCGAAGT_CATGTTGGCAGTTCGTGTCCCTGTAGATG	TTTGTGCGCCAAATTGC
KEG36084	(275)	ATGCCGAAGT_CATGTTGGCAGTTCGTGTCCCTGTAGATG	TTTGTGCGCCAAATTGC
AKN48098	(275)	ATGCCGAAGT_CATGTTGGCAGTTCGTGTCCCTGTAGATG	TTTGTGCGCCAAATTGC
CCJ93299	(272)	ATGCCGAAGT_AGTGATGGCAGTTCGTGTCCCTGTGGGT	ATTTGTGCGCCAAATTGC
<i>elml</i>	(330)	CGGATATGCCAGCCATGATAACAAATTTGTGATCCATCTTAAT	GGCGTGCTTCTT
ACT43858	(330)	CGGATATGCCAGCCATGATAACAAATTTGTGATCCATCTTAAT	GGCGTGCTTCTT
KEG36084	(330)	CGGATATGCCAGCCATGATAACAAATTTGTGATCCATCTTAAT	GGCGTGCTTCTT
AKN48098	(330)	CGGATATGCCAGCCATGATAACAAATTTGTGATCCATCTTAAT	AGCGTGCTTCTT
CCJ93299	(327)	CGGATATGCCAGCCA_CGATAACAAATTTGTGAT_TCATCTCAAT	GGCGTGCTTCTT
<i>elml</i>	(385)	AAAACGCCTGGCGTAGAAAACGTCAC	TGAAGAGTTTCAGTCAGTGTAATCCCTAA
ACT43858	(385)	AAAACGCCTGGCGTAGAAAACGTCAC	TGAAGAGTTTCAGTCAGTGTAATCCCTAA
KEG36084	(385)	AAAACGCCTGGCGTAGAAAACGTCAC	TGAAGAGTTTCAGTCAGTGTAAT-----
AKN48098	(385)	AAAACGCCTGGCGTAGAAAACGTCAC	TGAAGAGTTTCAGTCAGTGTAAT-----
CCJ93299	(382)	AAAACGCCTGGCGTAGAAAACGTCAC	TGAAGAGTTTC-GTCAATGTAATCCTTAA

Рис. 2. Выравнивание нуклеотидной последовательности предполагаемого гена, кодирующего *Elml* (ген *elml*), с ближайшими наиболее гомологичными последовательностями. Нуклеотидные последовательности обозначены по номерам кодируемых ими белков в GenBank. Серым фоном выделены нуклеотиды, идентичные у большинства, но не у всех из пяти проанализированных последовательностей. Нуклеотиды, не встречающиеся в большинстве последовательностей, выделены курсивом и подчеркнуты. Рамкой в последовательности гена *elml* выделен единственный нуклеотид, по которому ген отличается от ближайших гомологов из штаммов BL21(DE3) и C41(DE3) *E. coli*. Рамками с 5'- и 3'-концов выделены последовательности, соответствующие праймерам, с помощью которых проведен ПЦР-скрининг на геномных ДНК колиформных бактерий, выделенных из природных источников

следовательностях 5'-GCGC-3' на обеих цепях ДНК [12]. Кроме того, данная плазмида содержит гиперметилированный участок:

5'-G(5mC)G C AG(5mC)G C-3'  
3'-C G(5mC)GTC G(5mC)G-5';

5) pFsp4HI3, содержащая ген, кодирующий ДНК-метилтрансферазу Fsp4HI, в результате чего в плазмиде во всех последовательностях 5'-GCNGC-3' метилирован первый цитозин [4]. Данная плазмида содержит гиперметилированный участок:

5'-G(5mC)C G(5mC)G G(5mC)A G C-3'  
3'-C G G(5mC) G C(5mC)G T(5mC)G-5'.

Все плазмиды были предварительно переведены из кольцевой формы в линейную путем гидролиза эндонуклеазой рестрикции DriI по уникальному (для каждой плазмиды) сайту 5'-GACNNNNGTC-3'.

Результаты анализа сайт-специфичности на данных ДНК-субстратах приведены на рис. 3.

Как видно из рис. 3, ElmI не расщепляет последовательности, метилированные ДНК-метилтрансферазами HpaII (дорожка 2), HspAI2 (дорожка 4) и HaeIII (дорожка 6), т.е. последовательности 5'-C(5mC)GG-3'/3'-GG(5mC)C-5', 5'-G(5mC)-GC-3'/3'-CG(5mC)G-5' и 5'-GG(5mC)C-3'/3'-C(5mC)GG-5' соответственно не являются сайтами узнавания ElmI.

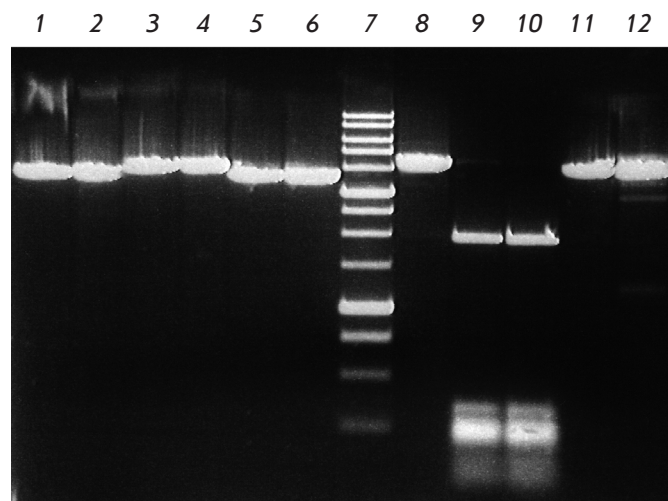


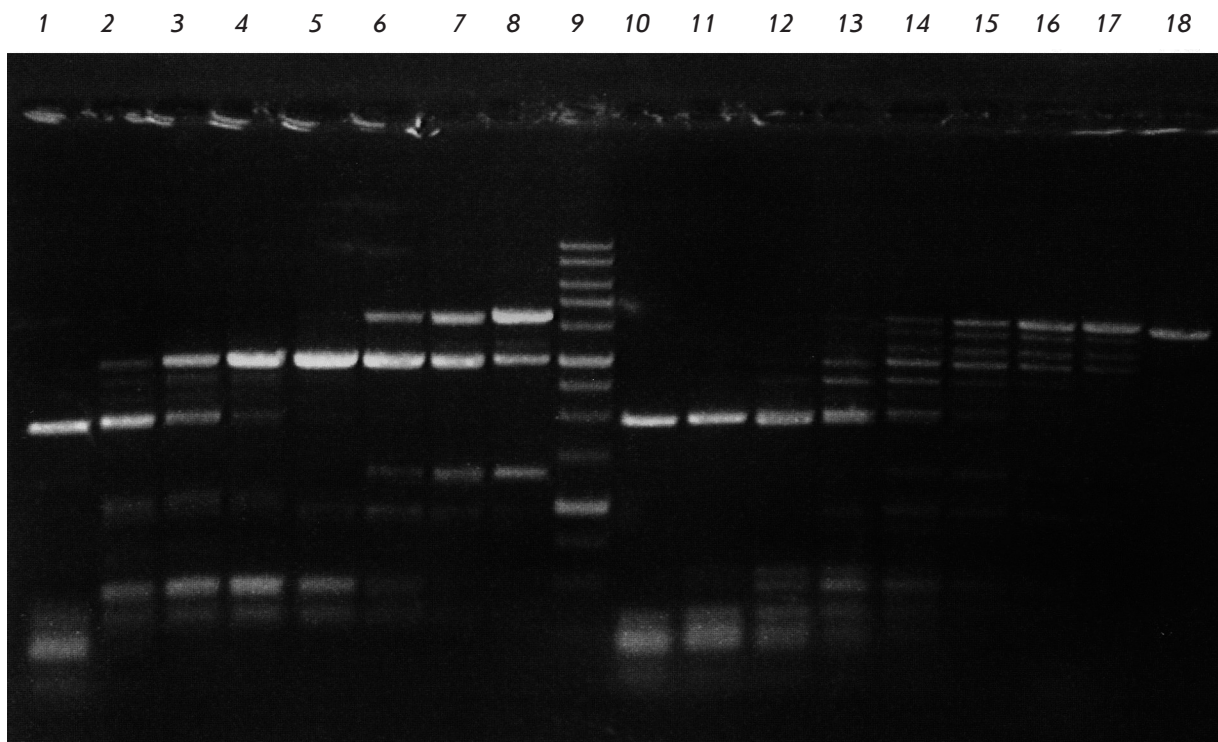
Рис. 3. Анализ сайт-специфичности ElmI на различных C5-метилированных ДНК-субстратах. Электрофорез в 1% агарозном геле. Дорожки: 1 – pMHpaII/DriI; 2 – pMHpaII/DriI + ElmI; 3 – pHspAI2/DriI; 4 – HspAI2/DriI + ElmI; 5 – pMHaelll/DriI; 6 – pMHaelll/DriI + ElmI; 7 – маркер молекулярной массы ДНК 1 kb; 8 – pFsp4HI3/DriI; 9 – pFsp4HI3/DriI + ElmI; 10 – pFsp4HI3/DriI + BisI; 11 – pHspAI4; 12 – pHspAI4 + ElmI

Из рис. 3 видно, что ElmI расщепляет ДНК pFsp4HI3/DriI (дорожка 9), образуя фрагменты ДНК такой же длины, как и после обработки этого же субстрата BisI (дорожка 10). Поскольку в этой плазмиде C5-метилированы все первые цитозины в последовательностях 5'-GCNGC-3' [4], и все они расщепляются MD-эндонуклеазой BisI, то данный факт свидетельствует о том, что ElmI узнает и расщепляет эти же метилированные сайты. Важно также, что ElmI расщепляет плазмиду pHspAI4 со значительно меньшей эффективностью, образуя два слабых фрагмента (дорожка 12). Это является следствием того, что указанный выше гиперметилированный участок плазмиды pHspAI4 включает уникальную последовательность 5'-GCNGC-3', у которой метилированы уже не внутренние (первые) цитозины на обеих цепях, как в плазмиде pFsp4HI3, а внешние (вторые).

Полученные данные означают, что ElmI узнает и расщепляет метилированную последовательность 5'-GCNGC-3', содержащую два остатка 5-метилцитозина, причем на порядок более эффективно, если метилированы два внутренних остатка цитозина на обеих цепях, чем при метилировании только двух внешних цитозинов.

Для того чтобы узнать, насколько эффективно ElmI расщепляет последовательности 5'-GCNGC-3' с большим количеством остатков 5-метилцитозина, pFsp4HI3/DriI обрабатывали ElmI в количестве, недостаточном для полного гидролиза, и сравнивали с картиной расщепления препаратом эндонуклеазы BisI (рис. 4).

Плазмида pFsp4HI3 содержит две последовательности 5'-GCNGCNGC-3' и одну 5'-GCNGCNGCNGC-3', в которые входят два и три соответственно пересекающихся сайта узнавания метилазы Fsp4HI. В результате метилирования этих последовательностей метилазой Fsp4HI образуются последовательности 5'-GCNGC-3' с тремя 5-метилцитозинами, в случае 5'-GCNGCNGCNGC-3' центральный сайт 5'-GCNGC-3' будет содержать уже четыре таких остатка. Анализ нуклеотидной последовательности плазмиды pFsp4HI3, проведенный с помощью программы Vector NTI Suite 7, показывает, что при расщеплении pFsp4HI3/DriI только по этим последовательностям, а также по сайту узнавания эндонуклеазы DriI, используемой для лиnearизации плазмиды, должны образоваться фрагменты длиной ~3000, ~490 (двойной фрагмент) и ~340 п.н. Как видно из рис. 4, эти фрагменты хорошо визуализируются при использовании 2-6 разведений ElmI (дорожки 2-5). Даже в последнем (дорожка 8) разведении фермента хорошо визуализируется фрагмент длиной ~1300 п.н., который должен образоваться при расщеплении уникальной гиперметилированной



**Рис. 4.** Неполное расщепление pFsp4HI3/Dril, показывающее более высокую активность ElmI на последовательностях 5'-GCNGC-3' с тремя или четырьмя 5-метилцитозинами по сравнению с 5'-GCNGC-3', содержащими лишь два 5-метилцитозина, которые предпочтительно расщепляются BisI. Электрофорез в 1% агарозном геле. Дорожки: 1–8 – pFsp4HI3/Dril, обработанная последовательными разведениями ElmI (исходное количество, добавленное в реакционную смесь, – 1 ед. акт.) в 2 раза; 9 – маркер молекулярной массы (ДНК 1 kb); 10–17 – pFsp4HI3/Dril, обработанная последовательными разведениями BisI (исходное количество, добавленное в реакционную смесь, – 1 ед. акт.) в 2 раза; 18 – pFsp4HI3/Dril

последовательности 5'-GCNGCNGCNGC-3', содержащей в том числе сайт 5'-GCNGC-3' с четырьмя остатками 5-метилцитозина.

В случае же BisI эти фрагменты визуализируются намного хуже. Несмотря на то что во 2–4 разведениях гидролиз плазмиды pFsp4HI3/Dril намного глубже по сравнению с ElmI, в последнем (восьмом) разведении BisI мы уже практически не видим фрагмент длиной ~1300 п.н. (дорожка 17).

Полученные данные говорят о том, что, в отличие от BisI, новая MD-эндонуклеаза ElmI расщепляет варианты последовательности 5'-GCNGC-3' с тремя или четырьмя 5-метилцитозинами на порядок лучше, чем в случае, когда таких метилированных остатков только два. Поэтому даже при использовании 1/16 ед. акт. препарата ElmI (дорожка 5) исходная ДНК полностью отсутствует благодаря более эффективному расщеплению 5'-GCNGC-3' с тремя или четырьмя остатками 5-метилцитозина. Напротив, BisI расщепляет такие гиперметилированные варианты менее эффективно, поэтому при использовании 1/16

ед. акт. препарата исходный фрагмент ДНК уже визуализируется (дорожка 14).

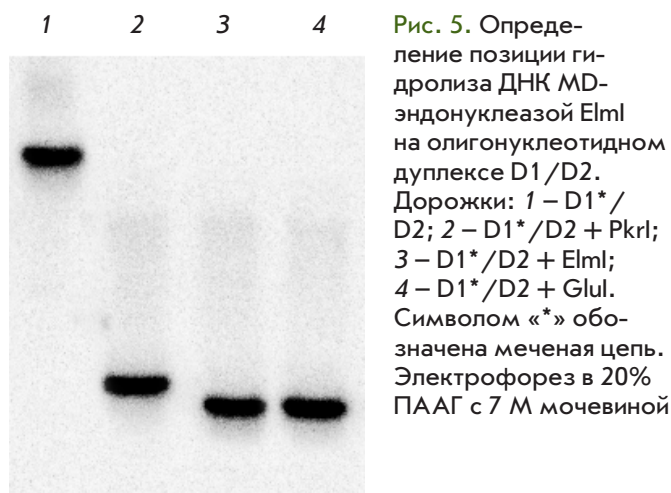
#### Определение позиции гидролиза ДНК в сайте узнавания ElmI

Места гидролиза ДНК ElmI определяли, сравнивая длины фрагментов, образуемых при расщеплении MD-эндонуклеазами ElmI, PkrI и GluI (последняя также узнает метилированную последовательность 5'-GCNGC-3' [4] и расщепляет ее подобно BisI перед центральным нуклеотидом) олигонуклеотидного дуплекса D1/D2, образованного из олигонуклеотидов D1 и D2 (предполагаемая узнаваемая ElmI последовательность подчеркнута):

D1: 5'-GAGTTTAG(5mC)GG(m5C)TATCGATCC-3'  
 D2: 5'-GGATCGATAG(5mC)CG(m5)-CTAAACTC-3'.

На рис. 5 приведен радиоавтограф электрофореграммы продуктов расщепления радиоактивно меченного дуплекса D1\*/D2 в 20% полиакриламидном геле с 7 М мочевиной.





**Рис. 5.** Определение позиции гидролиза ДНК MD-эндонуклеазой *ElmI* на олигонуклеотидном дуплексе D1/D2. Дорожки: 1 – D1\*/D2; 2 – D1\*/D2 + *PkrI*; 3 – D1\*/D2 + *ElmI*; 4 – D1\*/D2 + *GluI*. Символом «\*» обозначена меченая цепь. Электрофорез в 20% ПААГ с 7 М мочевиной

Как видно из рис. 5, фрагменты, образованные в результате гидролиза ферментами *PkrI* и *ElmI* (дорожки 2 и 3 соответственно) дуплекса D1\*/D2, имеют разную электрофоретическую подвижность, что свидетельствует о том, что эти ферменты имеют разные позиции гидролиза относительно сайта узнавания. В то же время электрофоретическая подвижность фрагментов ДНК, образованных при гидролизе *ElmI* и *GluI*, одинакова (дорожки 3 и 4 соответственно). Таким образом, *ElmI* и *GluI* имеют одинаковую гидролизуемую позицию относительно узнаваемой последовательности 5'-GCNGC-3'. Так как *GluI* расщепляет последовательность 5'-GC<sup>^</sup>NGC-3' перед центральным нуклеотидом «N» [4], то и *ElmI* также расщепляет ее перед центральным нуклеотидом.

### ЗАКЛЮЧЕНИЕ

Таким образом, впервые обнаруженная в энтеробактерии рекомбинантная MD-эндонуклеаза *ElmI* узнает нуклеотидную последовательность 5'-GC<sup>^</sup>NGC-3' и расщепляет обе цепи ДНК перед центральным нуклеотидом «N» с образованием 5'-выступающих однонуклеотидных концов.

Полученные результаты свидетельствуют о том, что геномы энтеробактерий содержат гены MD-

эндонуклеаз, аминокислотные последовательности которых имеют умеренную гомологию с *BisI*, причем лишь примерно половина аминокислотных остатков может рассматриваться как близких по физико-химическим свойствам. Тем не менее, несмотря на умеренную гомологию первичной структуры, сайты узнавания и позиции гидролиза ферментами *BisI* и *ElmI* в значительной мере схожи.

Стоит отметить, что с использованием пары праймеров *Esp-1* и *Esp-2* можно амплифицировать высокоомологичный гену *elmI* фрагмент ДНК ~430 п.н. из лабораторного штамма *E. coli* BL21(DE3) (ACT43858). Согласно базе данных GenBank этот фрагмент представляет собой рамку трансляции, которая кодирует полипептид с неизвестной функцией: Enterobact1 – WP\_001276099.1 hypothetical protein [*Escherichia coli*] >ref|YP\_003035796.1| hypothetical protein ECBD\_1551 [*Escherichia coli* 'BL21]. Однако, по нашим данным, эта рамка считывания представляет собой ген, кодирующий метилзависимую ДНК-эндонуклеазу. Мы обозначили ген, соответствующий этой рамке, как *ecoBLI*, а кодируемый им белок – *EcoBLI*. Его свойствам будет посвящена отдельная публикация.

Сайт-специфическая эндонуклеаза *ElmI* может быть использована в эпигенетических исследованиях, молекулярной биологии и генной инженерии для сайт-специфического расщепления метилированной ДНК, в том числе и для анализа статуса метилирования геномной ДНК растений [13], CNG-метилирование у которых считается эпигенетически важным. ●

*Работа проведена при финансовой поддержке Министерства образования и науки Российской Федерации по соглашению № 14.576.21.0075 от 06.11.2014 г. (уникальный идентификатор RFMEFI57614X0075), заключенному в рамках федеральной целевой программы «Исследования и разработки по приоритетным направлениям развития научно-технологического комплекса России на 2014–2020 годы».*

### СПИСОК ЛИТЕРАТУРЫ

1. Чернухин В.А., Томилова Ю.Э., Чмуж Е.В., Соколова О.О., Дедков В.С., Дегтярев С.Х. // Вестник биотехнологии и физико-химической биологии имени Ю.А. Овчинникова. 2007. Т. 3. № 1. С. 28–33.
2. Чернухин В.А., Наякшина Т.Н., Гончар Д.А., Томилова Ю.Э., Тарасова М.В., Дедков В.С., Михненкова Н.А., Дегтярев С.Х. // Вестник биотехнологии и физико-химической биологии имени Ю.А. Овчинникова. 2011. Т. 7. № 3. С. 35–42.
3. Чмуж Е.В., Каширина Ю.Г., Томилова Ю.Э., Мезенцева Н.В., Дедков В.С., Гончар Д.А., Абдурашитов М.А., Дегтярев С.Х. // Биотехнология. 2005. № 3. С. 22–26.
4. Чернухин В.А., Чмуж Е.В., Томилова Ю.Э., Наякшина Т.Н., Гончар Д.А., Дедков В.С., Дегтярев С.Х. // Вестник биотехнологии и физико-химической биологии имени Ю.А. Овчинникова. 2007. Т. 3. № 2. С. 13–17.
5. Manual of methods for general bacteriology / Edited by Gerhard P. Washington, D.C.: Am. Soc. Microbiol., 1981. 524 p.

6. Chambers S.P., Prior S.E., Barstow D.A., Minton N.P. // *Gene*. 1988. V. 68. P. 139–149.
7. Xu S.-Y., Boitano M., Clark T.A., Vincze T., Fomenkov A., Kumar S., Too P.H.-M., Gonchar D., Degtyarev S.Kh., Roberts R.J. // *Genome A*. 2015. V. 3. Issue 3. e00395–159.
8. *Bergey's manual of determinative bacteriology* / Ed. Holt J.G. et al. 9th ed. Baltimore: Williams and Williams, 1993. 787 p.
9. Madden T.L., Tatusov R.L., Zhang J. // *Meth. Enzymol*. 1996. V. 266. P. 131–141.
10. Чернухин В.А., Килева Е.В., Томилова Ю.Э., Болтенгаген А.А., Дедков В.С., Михненкова Н.А., Гончар Д.А., Голикова Л.Н., Дегтярев С.Х. // *Вестник биотехнологии и физико-химической биологии имени Ю.А. Овчинникова*. 2011. Т. 7. № 1. С. 14–20.
11. Чернухин В.А., Беличенко О.А., Тарасова Г.В., Гончар Д.А., Акишев А.Г., Дедков В.С., Михненкова Н.А., Дегтярев С.Х. Патент RU 2399663, Россия, C12N1/21, C12R1/06, 2009.
12. Чернухин В.А., Гончар Д.А., Килева Е.В., Соколова В.А., Голикова Л.Н., Дедков В.С., Михненкова Н.А., Дегтярев С.Х. // *Вестник биотехнологии и физико-химической биологии имени Ю.А. Овчинникова*. 2012. Т. 8. № 1. С. 16–26.
13. Vanyushin V.F. // *Curr. Top. Microbiol. Immunol*. 2006. V. 301. P. 67–122.