

микрофлоры верхних дыхательных путей поросят-отъемышей и подсвинков при неспецифической бронхопневмонии / А. В. Савинков, В. В. Ермаков, А. В. Лямин, Д. Д. Исмагуллин, А. В. Жёстков, К. М. Садов // Известия Оренбургского государственного аграрного университета. – 2020. – № 4 (84). С. 217–221.

DOI: 10.48612/sbornik-2023-1-46

УДК 574.24 579.262:578.4:636.5

### **ИССЛЕДОВАНИЕ РАСПРОСТРАНЕНИЯ ГЕНА ЭНДОЛИЗИНА БАКТЕРИОФАГА RB43 В ПРИРОДЕ С ПОМОЩЬЮ АНАЛИЗА БАЗ ДАННЫХ ГЕНЕТИЧЕСКИХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ**

**Зимин Андрей Антонович**<sup>1</sup>, канд. биол. наук

**Никулина Александра Николаевна**<sup>1</sup>, аспирант

**Никулин Никита Алексеевич**<sup>1</sup>, аспирант

**Кощаев Андрей Георгиевич**<sup>2</sup>, д-р. биол. наук, профессор

**Осепчук Денис Васильевич**<sup>2,3</sup>, д-р. с.-х. наук

<sup>1</sup>Институт биохимии и физиологии микроорганизмов им. Г. К. Скрыбина РАН – обособленное подразделение ФИЦ «Пушчинский научный центр биологических исследований РАН»,

г. Пушкино, Российская Федерация

<sup>2</sup>ФГБНУ «Краснодарский научный центр по зоотехнии и ветеринарии»,

г. Краснодар Российская Федерация

<sup>3</sup>ФГБОУ ВО «Кубанский государственный аграрный университет имени И. Т. Трубилина»,

г. Краснодар, Российская Федерация

Было найдено 75 последовательностей гомологов гена эндолизина бактериофага RB43 при сравнении с базой данных nr GenBank, семейства Straboviridae (taxid 2946170). С использованием алгоритма UPGMA было показано, что гомологи данного гена встречаются редко среди T4-фагов кишечной палочки, но имеются у ряда других энтеробактерий. С помощью BLASTn, настроенной на поиск слабых гомологий, в базе данных всех вирусных последовательностей за исключением семейства Straboviridae было найдено 22 последовательности. Это были последовательности из бактериофагов Salmonella схожие с P22. Картирование расположения этого сходства показало, что 3'- и 5'- концевые последовательности исследуемого гена не имеют гомологии с ДНК фагов Salmonella и могут быть использованы для выбора праймеров для идентификационного ПЦР.

**Ключевые слова:** ген эндолизина бактериофага RB43; семейство вирусов Straboviridae

### **STUDY OF THE DISTRIBUTION OF THE BACTERIOPHAGE RB43 ENDOLYSIN GENE IN NATURE USING THE ANALYSIS OF GENETIC SEQUENCE DATABASES**

**Zimin Andrei Antonovich**<sup>1</sup>, PhD Biol. Sci.

**Nikulina Aleksandra Nikolaevna**<sup>1</sup>, PhD student

**Nikulin Nikita Alekseevich**<sup>1</sup>, PhD student

**Koshchaev Andrei Georgievich**<sup>2</sup>, Dr. Biol. Sci., professor

**Osepchuk Denis Vasilyevich**<sup>2,3</sup>, Dr. Agr. Sci.

<sup>1</sup>Institute of Biochemistry and Physiology of Microorganisms named after G. K. Scriabin RAS – a separate subdivision of the Federal Research Center "Pushchino Scientific Center for Biological Research of the Russian Academy of Sciences", Pushchino, Russian Federatuin

<sup>2</sup>Kuban State Agrarian University named after I. T. Trubilin, Krasnodar, Russian Federation

<sup>3</sup>Krasnodar Research Centre for Animal Husbandry and Veterinary Medicine, Krasnodar, Russian Federation

75 sequences of homologues of the bacteriophage RB43 endolysin gene were found when compared with the nr GenBank database, Straboviridae family (taxid 2946170). Using the UPGMA algorithm, it was shown that homologues of this gene are rare among T4 phages of *Escherichia coli*, but are present in a number of other enterobacteria. Using BLASTn, configured to search for weak homologies, 22 sequences were found in the database of all viral sequences except for the Straboviridae family. These were sequences from *Salmonella* bacteriophages similar to P22. Mapping the location of this similarity showed that the 3'- and 5'-terminal sequences of the studied gene have no homology with *Salmonella* phage DNA and can be used to select oligonucleotides for identification PCR.

**Key words:** bacteriophage RB43 endolysin gene, Straboviridae family of viruses

*Escherichia coli* – широко распространенная в окружающей среде грамотрицательная неподвижная бактерия. Микроорганизмы этого вида традиционно считают комменсалами, их можно обнаружить на коже человека, в желудочно-кишечном тракте. Это существенный по распространенности сельскохозяйственный патоген в мире, способный вызывать широкий спектр инфекций значимых для животноводства [1]. Широчайшее и достаточно часто нерациональное использование антибиотиков во второй половине 20 века привело к распространению бактерий с генами устойчивости. Вследствие распространения устойчивых штаммов, большую актуальность приобретает разработка новых методов антибактериальной терапии сельскохозяйственных животных и птицы. Одной из важнейших перспектив является контроль патогенных бактерий вирулентными и не трансдуцирующими бактериофагами. Фаги обладают способностью быстро разрушать бактерии за счет фаголизиса. Им безразлично обладают эти бактерии устойчивостью к антибиотикам или не обладают. Фаговую терапию применяют в клинической практике с начала 20 века. Существенным ограничением этого подхода к антибактериальной терапии является способность ряда бактериофагов осуществлять горизонтальный перенос генов бактерий за счет фаговой трансдукции [5].

Геном T4-фагов содержит разные функциональные участки, обеспечивающие его репликацию, транскрипцию, трансляцию и сборку вириона. Изучения бактериофагов является существенным этапом для получения биопрепаратов на их основе и их применения в ветеринарии и зоотехнии. В настоящее время в отношении кишечной палочки, специфичные бактериофаги, в основном, были выделены из сточных вод, а также из речных, озёрных вод и почвы. Первичные фаговые изоляты обычно изучают на круг специфич-

ности их штаммов-хозяев, по морфологии негативных колоний - бляшек, а также морфологии вириона с помощью трансмиссионной электронной микроскопии (ТЭМ). Следующим рутинным этапом является обычно ПЦР с праймерами к консервативным генам для той или иной группы бактериофагов. Для вирулентных T4-родственных бактериофагов это обычно ген 23, кодирующий основной белок капсида бактериофагов этой группы [5].

Бактериофаг T4 не способен к осуществлению трансдукции. Эта его особенность связана с отсутствием цитозина в его ДНК, где последний заменен на 5'- гидроксиметилцитозин. В конце 20, начале 21 веков Тяняшиным и соавторами было показано, что сходные с T-четными псевдо – T – четные, бактериофаги RB43, RB49, RB42, положительные по ПЦР к гену 23, способны осуществлять трансдукцию плазмид [9, 10]. В процессе работы по выделению фагов, для формирования лечебных препаратов, важно отделить бактериофаги не способные к общей трансдукции и тем самым безопасные для применения в практике фаговой терапии. Подобный опыт с помощью прямых экспериментов по трансдукции был проведен нами для отбора нетрансдуцирующих бактериофагов против колибактериозов порослят в постотъемном периоде их развития [6]. В этой работе мы рассматриваем альтернативный подход. С помощью изучения геномов T4-бактериофагов выявлена значительная гетерогенность геномов и существенное отличие T-четных бактериофагов и псевдо- T-четных. Это позволяет произвести достаточно широкий выбор гена для формирования идентификационного ПЦР, с помощью которого можно отделить трансдуцирующие бактериофаги от нетрансдуцирующих T4-фагов, оптимальных для включения в препараты для контроля патогенных штаммов кишечной палочки в животноводстве и птицеводстве.

В основном предполагается использовать наличие/отсутствие того или иного генетического маркера или очень малое сходство последовательностей ДНК для данной селекции псевдо – Т – четных бактериофагов. Результаты анализа распространения этих фагов и их аллелей в природных популяциях бактериовирусов, отраженные в базах данных генетических последовательностей, указывают на сравнительную редкость этих фагов, но данная ситуация тоже требует наличия методов быстрой идентификации этих трансдуцирующих фагов после изоляции их из природных и искусственных водоемов.

**Методика исследований.** С помощью методов биоинформатики предполагалось исследовать базы данных генетических последовательностей бактериофагов, имеющих отростки семейства *Straboviridae* с целью получения представлений о распространении в природе гена эндолизина бактериофага RB43. Бактериофаги этого семейства могут выделяться вместе с T4 – родственными бактериофагами и сравнение последовательности этого гена позволит оценить специфичность обнаружения псевдо – Т – четного бактериофага RB43 среди изолированных бактериофагов с помощью ПЦР. Этот подход можно применить для отсека их из препаратов для бактериофаговой терапии. Для исследования баз данных генетических последовательностей использовались подходы, выработанные нами ранее [2, 3]. Для поиска сходств генетических последовательностей использовался алгоритм BLASTn [8], настроенный на разную степень сходства.

**Результаты исследований.** Было проведено сравнение нуклеотидной последовательности гена эндолизина бактериофага RB43 с базой данных nr GenBank, семейства *Straboviridae* (taxid 2946170). Всего было найдено 75 нуклеотидных последовательностей. Все последовательности были из геномов фагов, родственных T4, но заражающих различные семейства бактерий. Был сформирован файл нуклеотидных последовательностей в формате Fasta. Множественное наложение этих последовательностей было проведено с помощью алгоритма MUSCL в пакете программ MegaX [8]. Для выявления наиболее близких последовательностей и исследования возможности ложноположительного ПЦР при исполь-

зовании проб из водоемов и других проб из природы или фекалий сельскохозяйственных животных мы провели филогенетический анализ найденных последовательностей с использованием алгоритма UPGMA также в пакете программ MegaX (рис. 1) [8].

Консенсусное дерево начальной загрузки, полученное из 1000 повторов, использовано для представления эволюционной истории анализируемых таксонов. Ветви, соответствующие разделам, воспроизведенным менее чем в 50 % репликах начальной загрузки, свернуты. Рядом с ветвями показан процент повторяющихся ветвей, в которых связанные таксоны сгруппированы вместе в тесте начальной загрузки (1000 повторов). Эволюционные расстояния были рассчитаны с использованием метода максимального составного правдоподобия и выражены в единицах количества замен оснований на сайт. В этом анализе использовались 75 нуклеотидных последовательностей. Все неоднозначные позиции были удалены для каждой пары последовательностей (опция попарного удаления). Всего в окончательном наборе данных было 438 позиций. Эволюционный анализ проводился в MEGA X [4]

Этот алгоритм основывается в первую очередь на простом сравнении идентичности нуклеотидов в ДНК. Этот метод реализует и эволюционную близость, но за счет учета в первую очередь поиска идентичности среди последовательностей. Было показано, что гомологи данного гена встречаются весьма редко среди T4-бактериофагов кишечной палочки, но имеются у ряда других энтеробактерий. Эти фаги, вероятно, не смогут заразить используемые нами штаммы за счет фаговой специфичности и будут отделены на стадии выделения изолятов из природы на двухслойном агаре. Для того чтобы исключить наличие этой нуклеотидной последовательности у других вирусов был произведен поиск сходных последовательностей среди всех вирусов, за исключением бактериофагов семейства *Straboviridae* (*viruses* (taxid 10239) and exclude *Straboviridae* (taxid 2946170)). Результаты этого исследования в виде графического изображения множественно наложения найденных нуклеотидных последовательностей представлены на рисунках 2.

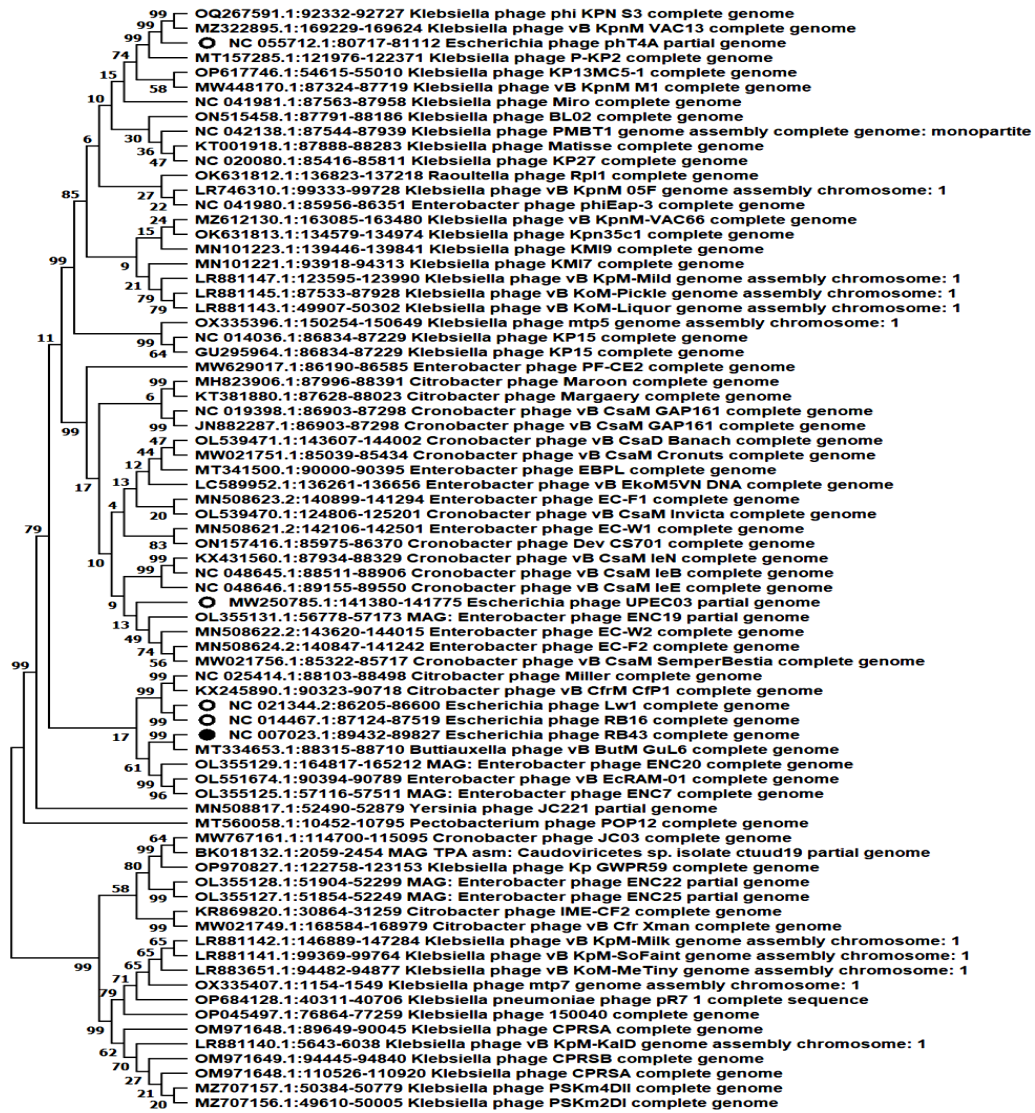


Рисунок 1 – Филогенетическое дерево генов гомологов гена *lys* RB43, полученное с использованием метода UPGMA

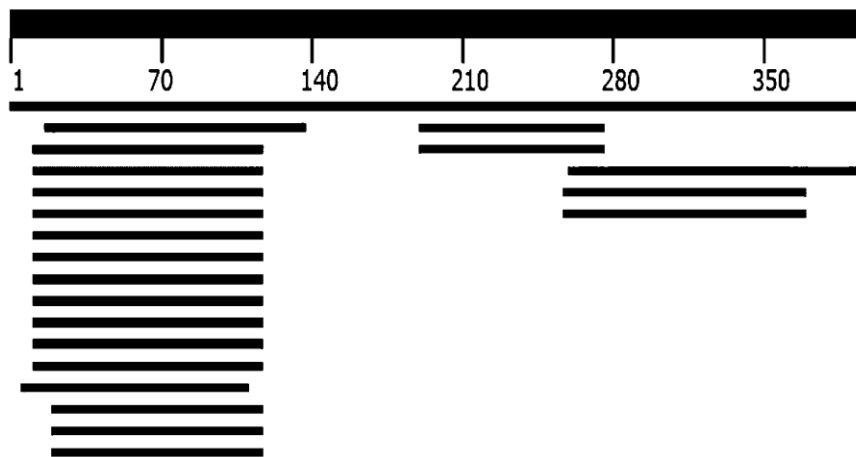


Рисунок 2 – Графическое представление распределения участков средней гомологии среди находок нуклеотидных последовательностей сходных с геном эндолизина бактериофага RB43 в базе данных всех вирусных последовательностей за исключением семейства *Straboviridae* (*viruses* (taxid 10239) and exclude *Straboviridae* (taxid 2946170))

С помощью программы BLASTn, настроенной на поиск слабых гомологий, мы искали последовательности сходные с геном эндолизина бактериофага RB43 в базе данных всех вирусных последовательностей за исключением семейства *Straboviridae* (*viruses (taxid 10239)* and *exclude Straboviridae (taxid 2946170)*). Всё-таки было найдено 22 последовательности. Одна из них представляла геном бактериофага семейства *Straboviridae* пока не отнесенный к этому семейству в базе данных. Большинство остальных располагались в положении 20 – 110 по последовательности гена эндолизина псевдо – Т – четного фага. Это были последовательности из бактериофагов *Salmonella phage ST-32*, *Salmonella phage ST160*, *Enterobacteria phage ST64T*, *Bacteriophage PS3 lysis genes 13, 19, 15, and packaging gene 3*, *Proteus phage vB\_PmiM\_ZX7*, *Salmonella phage FSL SP-062*, *Salmonella phage ST-87*, *Salmonella phage ST-29* и ряд других фагов сальмонелл. Большинство этих фагов ранее были охарактеризованы как схожие с P22 [10]. Фаги сальмонелл могут встречаться в тех же биотопах, что и T4-фаги кишечной палочки и стоит обратить внимание, что их гены

лизиса имеют небольшое сходство с геном эндолизина бактериофага RB43. В этом случае нельзя исключить ложноположительную ПЦР при низких температурах отжига в реакции. Картирование расположения этого сходства показало, что 3'- концевая и 5' - концевая последовательности исследуемого гена не имеют гомологии с ДНК этих фагов сальмонелл и могут быть использованы для выбора в праймеров для идентификационного ПЦР.

Поиск и выделение бактериофагов перспективно производить не только из пресноводных водоемов, но из морских также. С целью исследования распространенности этой нуклеотидной последовательности в море мы провели сравнение с помощью алгоритма BLASTn среди первичных ридов, полученных профессором К. Вентером и его коллегами в ходе экспедиции на яхте Сорсерер в тропических экваториальных водах Атлантического и Тихого океанов по программе Global Ocean Sampling (GOS). Результаты этого исследования в виде графического изображения множественно наложения найденных нуклеотидных последовательностей представлены на рисунке 3.

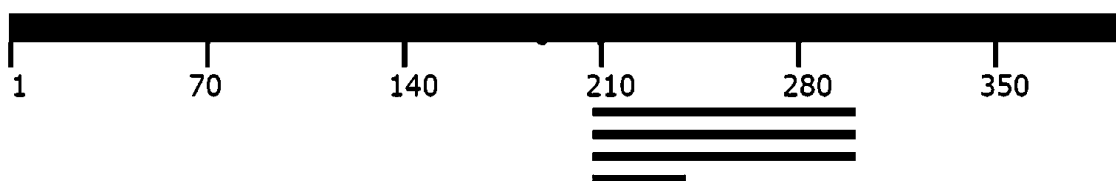


Рисунок 3 – Графическое представление распределения участков средней гомологии среди находок нуклеотидных последовательностей сходных с геном эндолизина бактериофага RB43 в базе данных метагенома океана, сделанного по программе Global Ocean Sanpling.

Использовался аналогичный предыдущему подход и программное средство. Было найдено лишь четыре последовательности из ридов Global Ocean Sampling. Распределение этих ридов почти совпадало с положением двух ридов вирусного происхождения. Эта находка требует дополнительного изучения с позиций фундаментальной науки, но вряд ли может внести коррективы при анализе с помощью ПЦР содержания псевдо – Т – четных бактериофагов в пробах из океана. Эти находки были весьма неожиданными, но имели слабую гомологию ( $E > 10^{-12}$ ) и вряд ли они проявятся в ПЦР даже при анализе вирионов. Хотя стоит обратить внимание на распределение этих находок по длине исследуемой последовательности. Большинство находок рас-

положено в районе 210 – 300 нуклеотидов. При конструировании праймеров следует провести проверку конкретной выбранной для этой цели короткой последовательности путем её сравнения с базами данных вирусных последовательностей программой BLASTn, настроенной на поиск слабых гомологий. Скорее всего, найденные на этом этапе работы океанические последовательности, для которых процент идентичности около 40% и менее, не могут повлиять на специфичность узнавания праймеров.

#### **Выводы.**

1. Было найдено 75 нуклеотидных последовательностей гомологов гена эндолизина бактериофага RB43 при сравнении с базой данных nr GenBank, семейства *Straboviridae*. С

использованием алгоритма UPGMA было показано, что гомологи данного гена встречаются весьма редко среди T4-бактериофагов кишечной палочки, но имеются у ряда других энтеробактерий. Эти фаги, вероятно, не смогут заразить используемые нами штаммы за счет фаговой специфичности, и будут отделены на стадии выделения изолятов фагов из природы на двухслойном агаре.

2. С помощью BLASTn, настроенной на поиск слабых гомологий, в базе данных всех вирусных последовательностей за исключением семейства *Straboviridae* (*viruses (taxid 10239)* and *exclude Straboviridae (taxid 2946170)*) было найдено 22 последовательности. Это были последовательности из бактериофагов *Salmonella* схожие с P22. Картирование расположения этого сходства показало, что 3'- и 5'- концевые последовательности исследуемого гена не имеют гомологии с ДНК сальмонелльных фагов и могут быть использованы для выбора олигонуклеотидов для идентификационного ПЦР.

**Благодарности.** Исследование выполнено за счет гранта Российского научного фонда № 22-25-00669, <https://rscf.ru/project/22-25-00669/>

### Список литературы

1. Зимин А. А. Использование бактериофагов для борьбы с колибактериозом и кампилобактериозом в птицеводстве / А. А. Зимин, Ф. В. Кочетков, С. И. Кононенко, Д. В. Осепчук, Н. Э. Скобликов // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ). – Краснодар: КубГАУ. – 2016. – № 09(123). – С. 421–432.
2. Зимин А. А. Коронавирусы и животноводство / А. А. Зимин, Д. В. Осепчук // Сборник научных трудов Краснодарского научного центра по зоотехнии и ветеринарии. – 2020. – Том – 9. – № 1. – С. 8–14.
3. Зимин А. А. Сравнение структурного белка денсовируса BmDENV-1 тутового шелкопряда с белками вирусов бактерий и архей для изучения возможности ложноположительных ответов при ИФА-тестировании гусениц / А. А. Зимин, Н. Э. Скобликов, Н. Н. Назипова, Д. В. Осепчук, А. Г. Кощаев // Политематический сетевой электронный научный журнал КубГАУ. – Краснодар: КубГАУ. – 2020. – № 161. – С. 150-160
4. Зимин А. А. Изучение вариабельности хвостовых шипиков бактериофагов энтеробактерий P22-типа методами сравнительного биоинформационного анализа молекул белков для разработки препаратов для лечения молодняка сельскохозяйственной птицы / А. А. Зимин, С. И. Кононенко, Н. Э. Скобликов, Н. Н. Назипова // Известия Горского государственного аграрного университета. – 2017. – Том – 54. № 4. – С. 160–165.
5. Никулин Н. А. Конструирование терапевтических фаговых коктейлей на основе бактериофагов: преимущества и недостатки / Н. А. Никулин, С. И. Кононенко, А. Г. Кощаев, А. А. Зимин // Политематический сетевой электронный научный журнал Кубанского государственного аграрного университета (Научный журнал КубГАУ). – Краснодар: КубГАУ. – 2017. – №09(133).
6. Скобликов Н. Э. Выделение и отбор нетрансдуцирующих бактериофагов *E. coli* для противоклибактериозных препаратов / Н. Э. Скобликов, С. И. Кононенко, Д. В. Осепчук, Е. А. Москаленко, В. В. Авдиенко, А. А. Зимин // Политематический сетевой электронный научный журнал КубГАУ. – 2016. – №08(122). – С. 554–566.
7. Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W. & Lipman, D.J. (1997) "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs." *Nucleic Acids Res.* 25:3389–3402.
8. Kumar S., Stecher G., Li M., Knyaz C., Tamura K. (2018). MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms. *Molecular Biology and Evolution* 35:1547–1549.
9. Tanyashin V. I., Zimin A. A., Shlyapnikov M. G., Boronin A. M. Transduction of Plasmid Antibiotic Resistance Determinants with Pseudo-T-Even Bacteriophages. *Russian Journal of Genetics*, Vol. 39. – No. 7. – 2003, pp. 761–772. DOI: 10.1023/A:1024748903232.
10. Tanyashin V. I., Zimin A. A., Boronin A. M. The Cotransduction of pET System Plasmids by Mutants of T4 and RB43 Bacteriophages. *Microbiology*. – Vol. 72. – No. 6. – 2003. – pp. 694–700. DOI:10.1023/B:MICI.0000008372.06477.43.