

УДК 579.842.14:577.21:616.074/077

DOI:10.14427/jipai.2025.4.67

Молекулярно-генетическое типирование штаммов *Salmonella enterica* Enteritidis на основе маркерных SNP

А.М. Горох, А.С. Водопьянов, А.А. Герасименко, Р.В. Писанов

ФКУЗ Ростовский-на-Дону противочумный институт Роспотребнадзора, Ростов-на-Дону

Molecular genetic typing of *Salmonella enterica* Enteritidis strains by marker SNP

A.M. Gorokh, A.S. Vodopyanov, A.A. Gerasimenko, R.V. Pisanov

«Rostov-on-Don Antiplague Scientific Research Institute» of Rospotrebnadzor, Rostov-on-Don, Russia

Аннотация

В работе были исследованы 2868 геномов штаммов *Salmonella enterica* Enteritidis, выделенных в 2007-2024 гг. из клинического материала, продуктов питания и окружающей среды. Проведён SNP-анализ с построением филогенетического дерева исследуемых штаммов и определены маркерные мутации для каждой отдельной генетической линии.

Ключевые слова

Salmonella, кластер, SNP-анализ, генотип, классификация.

Summary

The study examined 2868 genomes of *Salmonella enterica* Enteritidis strains isolated in 2007-2024 from clinical material, food and the environment. A SNP analysis was performed with the construction of a phylogenetic tree of the studied strains, identifying marker mutations for each individual genetic line.

Keywords

Salmonella, cluster, SNP-analysis, genotype, classification.

Широкое практическое применение для субвидовой характеристики микроорганизмов получили методы молекулярно-генетического анализа, обладающие большей дифференцирующей способностью по сравнению с фенотипическими методами исследования.

Salmonella enterica серотипа Enteritidis – один из наиболее распространённых возбудителей пищевых инфекций во всём мире. Это обусловлено способностью данного серотипа контаминировать продукты питания (особенно куриные яйца и мясо птицы), что ведёт к частым вспышкам пищевых отравлений [1]. Мониторинг и оперативный анализ вспышек, вызванных штаммами *S. Enteritidis*, являются важной проблемой общественного здравоохранения. Для достоверной дифференциации штаммов важной проблемой остаётся генетическая однородность данного патогена.

Применение SNP-типирования на практике уже позволило благополучно использовать его при расследовании ряда сложных вспышек,

вызванных *S. Enteritidis*. Jiang и соавт. (2020) описали вспышку из 10 случаев *S. Enteritidis* в городе Шэньчжэнь, где все заболевшие заказывали блюда через популярное приложение доставки еды. Предполагаемым источником стала курица, из которой были приготовлены блюда, из одного ресторана. SNP-анализ выявил, что клинические штаммы образуют единый кластер, близкий к изолятам, полученным с кухонного инвентаря ресторана. Это помогло достоверно связать вспышку с деятельностью конкретного заведения общественного питания [1]. В крупной вспышке, произошедшей летом 2014 года в Великобритании и ряде стран ЕС (Германия, Австрия, Франция, Люксембург), было зарегистрировано более 350 случаев гастроэнтерита, вызванного *S. Enteritidis* фаготипа 14b. Традиционные методы типирования указывали на общий источник инфицирования – заражённую партию яиц, однако установить конкретного поставщика было непросто, поскольку яйца распространялись через международную торговую сеть. В ходе

дальнейшего расследования было собрано и секвенировано свыше 400 изолятов *S. Enteritidis* из клинического материала, а также продуктов питания – образцов яиц с птицеводческих предприятий. Проведённый по SNP-данным филогенетический анализ выявил чёткую корреляцию между формированием генетических кластеров штаммов с распределением поставок яиц. Было показано, что вспышка состояла из нескольких близкородственных линий *S. Enteritidis*, каждая из которых соответствовала поставкам партии яиц, полученных с отдельной фермы крупного производителя, находящейся в Баварии. После этого были приняты необходимые меры для предотвращения дальнейшего распространения инфекции [2]. Этот пример демонстрирует, как SNP-типирование позволило связать случаи из разных стран и обеспечить своевременное выявление заражённого продукта.

Спорадические случаи сальмонеллёза, вызванные *S. Enteritidis* ST11, поначалу без выявления очевидной связи между собой, регистрировались также в Великобритании в период 2012-2015 гг. Анализ с использованием полногеномного секвенирования показал, что многие из этих штаммов генетически практически идентичны и формируют единую вспышку [3]. В результате детального расследования стало известно, что все пациенты были владельцами рептилий и использовали замороженных мышей в качестве корма для своих питомцев. Источником инфекции оказались импортированные замороженные кормовые мыши, заражённые *S. Enteritidis*. Клональный характер инфекции *S. Enteritidis* у 147 заболевших в течение нескольких лет был связан не с пищевым продуктом для людей, а с кормом для животных. Такой путь передачи без методов молекулярного типирования выявить было бы крайне трудно. Таким образом, SNP-типирование является мощным инструментом анализа штаммов *S. enterica* в эпидемиологических расследованиях.

Цель исследования заключалась в изучении молекулярно-генетического разнообразия с определением маркерных SNP и выявлении филогенетических связей штаммов *S. Enteritidis*, выделенных в Российской Федерации и за рубежом в 2007-2024 гг.

Материалы и методы

В работе были использованы 2868 геномов штаммов *S. Enteritidis*, изолированные в 2007-2024 гг., из которых 62 были секвенированы во ФКУЗ Ростовский-на-Дону противочумный институт Роспотребнадзора, 1933 было получено из

базы данных NCBI и 873 из базы данных VGARus. Геномную ДНК выделяли набором Рибо-Преп (Амплисенс, Россия). Библиотеку фрагментов ДНК готовили с использованием набора реагентов Illumina Nextera Flex (Illumina, США) согласно инструкции производителя. Ампликоны метили с использованием Nextera CD Index Kit (Illumina, США). Секвенирование проводили на секвенаторе MiSeq (Illumina, США) с набором реагентов MiSeq Reagent Kit. Сборку геномов осуществляли с помощью пакета SPAdes v.3.15.4 [4] с параметрами по умолчанию. SNP-анализ был проведён с помощью пакета snippy [5]. Филогенетическое дерево построено в программе MEGA v.11.0.11 [6] по алгоритму максимального правдоподобия (bootstrap=1000). Для визуализации дендрограммы использовали пакет ete3 [7]. Маркерные мутации были определены с помощью авторского скрипта на языке Python.

Результаты

Первым этапом работы стало формирование репрезентативной коллекции геномов *S. Enteritidis*. В эту выборку были включены три основные категории данных: все геномы, секвенирование которых и сборка выполнены в рамках настоящего исследования; все доступные на момент анализа геномы из отечественной базы данных VGARus; часть геномов из международной базы данных NCBI. При отборе из базы данных NCBI особое внимание уделялось геномам, опубликованным за последние десять лет, что позволило сосредоточиться на наиболее актуальных изолятах и одновременно обеспечить широкий географический охват. Таким образом, коллекция включала геномы, полученные из различных стран, что делает её подходящей для дальнейших сравнительных и эпидемиологических анализов.

В результате исследования отобраны полные геномы штаммов, сборка которых соответствовала критериям оценки качества. Для этой цели был использован показатель N50 – в финальную выборку включались только те геномы, у которых это значение превышало 25 000 п.н. Выбор данного порогового значения согласуется с подходами, использованными в других геномных исследованиях штаммов *Salmonella*, где также применялся аналогичный критерий для отбора качественных сборок [8]. Такой подход позволил исключить из рассмотрения геномы низкого качества или неполные сборки, что является критически важным для анализа. В результате была сформирована итоговая коллекция, состоящая из 2868 геномов *S. Enteritidis*.

На основе анализа SNP было проведено изучение филогенетических взаимоотношений между исследуемыми геномами. В результате кластерного анализа было идентифицировано 528 иерархически вложенных кластеров. Из них 21 кластер был определён как основной, включающий в себя остальные более мелкие субкластеры. В качестве ключевого критерия для определения статистической значимости ветвления на дендрограмме использовалась бутстреп-поддержка (bootstrap support). Кластером считалась только та группа геномов, узел которой на дендрограмме имел значение бутстреп 90% и выше, что указывает на высокую статистическую поддержку конкретной ветви филогенетического дерева. Так, к примеру, если группа штаммов из одной вспышки образует один кластер с бутстрепом более 90% – это надёжно доказывает их общее происхождение и, наоборот, в тех случаях, когда показатель будет низким – это может указывать на ошибки выравнивания геномов или ложное родство штаммов [9]. Такой высокий порог обеспечивает большую надёжность и достоверность выделяемых филогенетических групп. На основе вышеописанного SNP анализа штаммов построена дендрограмма, отражающая филогенетические взаимоотношения между исследуемыми геномами (рис. 1).

Для каждого из выделенных кластеров были определены маркерные мутации (характерные SNP), позволяющие однозначно идентифици-

ровать принадлежность генома к той или иной группе. Система наименований была построена по географическому принципу. Каждый кластер получал условное обозначение, основанное на месте выделения первого идентифицированного в нём штамма или штамма, являющегося мажорным (преобладающим) представителем данной филогенетической группы. В результате дальнейших исследований молекулярно-генетического разнообразия микроорганизмов *S. Enteritidis* классификация может быть дополнена в связи с появлением новых сублиний. В случае если конечная генетическая линия не определена, вывод о генетическом родстве штаммов делается по ближайшему родительскому кластеру.

Полученные данные по филогенетическому анализу были включены в авторскую программу *Salmonella Analyzer 2* (рис. 2).

Программа позволяет определять принадлежность штаммов *S. Enteritidis* к определённой генетической линии без построения дендрограммы и представлять результаты в виде таблицы (табл. 1).

Проведённый филогенетический анализ позволяет оценить частоту и плотность распределения штаммов, так, на примере кластера Rostov-2, где подавляющее количество изолятов распределены на Юге Российской Федерации и ветви EastRus-19, характерную для штаммов, выделенных в Восточной Сибири (рис. 3а,б).

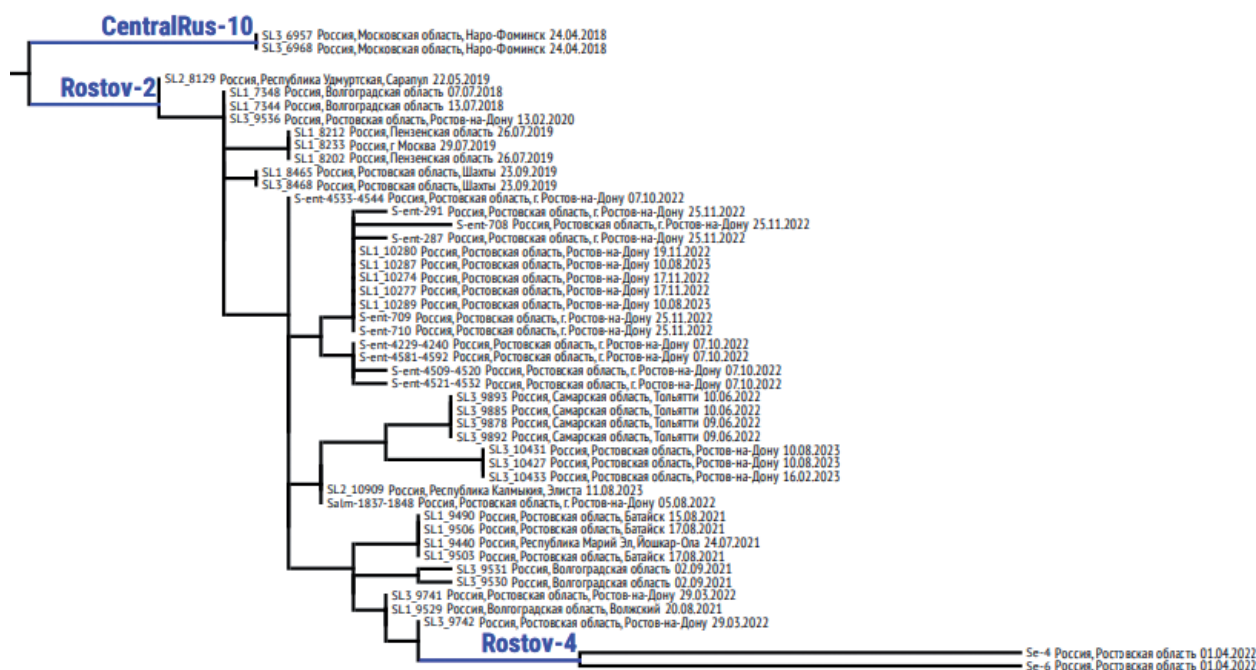


Рис. 1. Фрагмент дендрограммы, построенной с помощью SNP-анализа российских штаммов

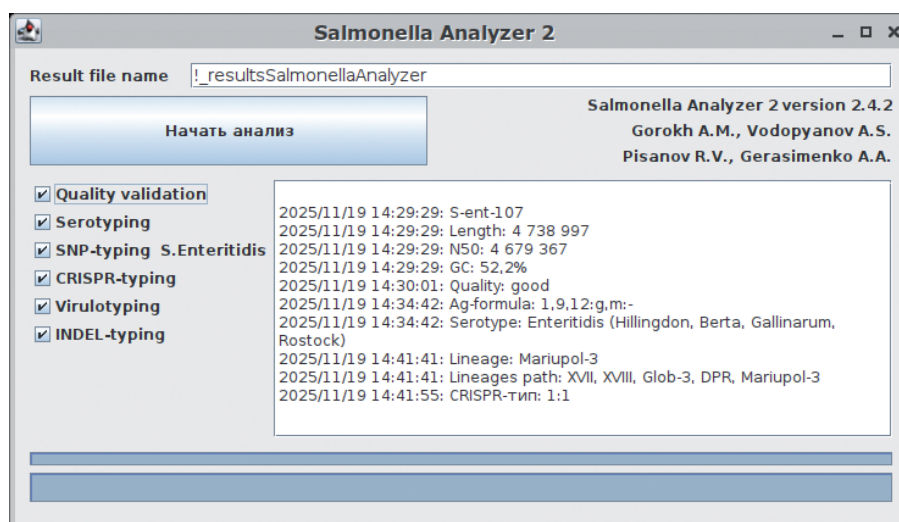


Рис. 2. Интерфейс работы программы

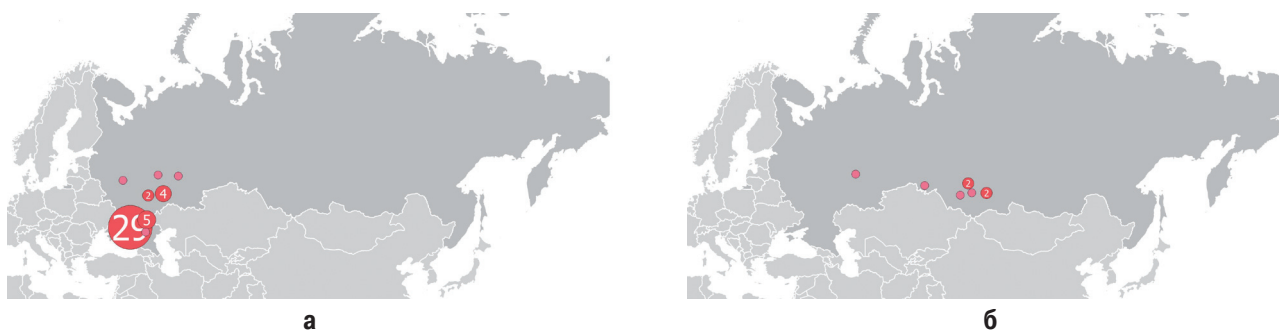


Рис. 3. Распределение штаммов, относящихся к ветви Rostov-2 и EastRus-19 (размер маркера соответствует количеству штаммов)

SNP-типирование микроорганизмов рода *Salmonella* проводится для изучения генетического разнообразия и филогенетического анализа вновь выделяемых штаммов с помощью различных программ, специализированных платформ и баз данных, которые позволяют анализировать геномы и определять SNP-маркеры, необходимые для эпидемиологических расследований. Рассмотрим EnteroBase – онлайн-платформу для анализа геномов микроорганизмов рода *Salmonella* и других энтеробактерий, содержащую инструменты для SNP-типирования, MLST и серотипирования [10]. Это интегрированная веб-платформа, которая комплексно объединяет SNP-анализ с другими методами, что является одновременно и преимуществом, и недостатком. Из положительных моментов – бесплатна, автоматизирована, не требует установки локального программного обеспечения; из отрицательных – зарубежный сервис, который в какое-то время может стать недоступным; анализ в онлайн-ре-

жиме; ограничение на количество запросов и анализов. В работе Katz et al. 2017 года, где были использованы методы SNP-анализа, включая подходы, используемые в EnteroBase, приведены основные проблемы и недостатки. Среди них трудности с установлением истинного филогенетического родства, а также проблемы с наличием ложнородственных групп. Также полученные результаты не всегда коррелируют с эпидемиологическими данными [11].

Обсуждение

Стоит также привести в пример PubMLST – открытую базу данных, которая специализируется на методах мультилокусного типирования, но также интегрирует другие подходы, включая SNP для анализа бактериальных штаммов, в том числе *S. Enteritidis* [12]. Однако в этой базе данных SNP-анализ не оптимизирован и требует дополнительной обработки данных через инструменты EnteroBase.

Таблица 1. Определение генетических линий отдельных штаммов с помощью программы Salmonella Analyzer 2

Штамм	Серотип	SNP-тип	Филогенетическая линия
S-ent-2363	Enteritidis (Hillingdon, Berta, Gallinarum, Rostock)	Mariupol-1	I, II, X, CentralRus-11, EvroRus-2, SouthRus-8, Mariupol-1
S-ent-2719	Enteritidis (Hillingdon, Berta, Gallinarum, Rostock)	Mariupol-1	I, II, X, CentralRus-11, EvroRus-2, SouthRus-8, Mariupol-1
S-ent-2965-2976	Enteritidis (Hillingdon, Berta, Gallinarum, Rostock)	Norilsk-1	I, II, X, CentralRus-11, Norilsk-1
S-ent-213	Enteritidis (Hillingdon, Berta, Gallinarum, Rostock)	Rostov-2	I, II, X, GlobRus-2, SouthRus-7, Rostov-2
S-ent-237	Enteritidis (Hillingdon, Berta, Gallinarum, Rostock)	Rostov-2	I, II, X, GlobRus-2, SouthRus-7, Rostov-2
S-ent-1serv	Enteritidis (Hillingdon, Berta, Gallinarum, Rostock)	DPR-2	XVII, XVIII, Glob-3, DPR-2
S-ent-107	Enteritidis (Hillingdon, Berta, Gallinarum, Rostock)	Mariupol-3	XVII, XVIII, Glob-3, DPR, Mariupol-3
S-ent-2362	Enteritidis (Hillingdon, Berta, Gallinarum, Rostock)	Mariupol-3	XVII, XVIII, Glob-3, DPR, Mariupol-3
S-ent-108	Enteritidis (Hillingdon, Berta, Gallinarum, Rostock)	Mariupol-4	XVII, XVIII, Glob-3, DPR, Mariupol-3, Mariupol-4
S-ent-2849	Enteritidis (Hillingdon, Berta, Gallinarum, Rostock)	Mariupol-4	XVII, XVIII, Glob-3, DPR, Mariupol-3, Mariupol-4
S-ent-3125-3136	Enteritidis (Hillingdon, Berta, Gallinarum, Rostock)	USA-2	XVII, XVIII, Glob-3, USA-2

Отличие разработанной нами программы от используемых в приведённых выше ресурсах в том, что при анализе геномов в ней есть система иерархичности в классификации по данным SNP-анализа, что делает итоговый результат удобным для определения филогенетических связей между штаммами. Программа легко устанавливается, работает офлайн как в консольном, так и в графическом режиме. Количество выполняемых анализов ограничено только временем и мощностью компьютера исследователей.

Определённые нами уникальные комбинации маркерных мутаций позволяют достоверно классифицировать штаммы *S. Enteritidis* (в отличие от других методик типирования, где производится выборка гетерогенных наборов SNP у каждой группы исследователей), что делает результат практически невозпроизводимым другими авторами. Также в других методах отсутствует иерархичность классификации генетических линий, что не отражает эволюционный процесс и происхождение штаммов по типу родитель-потомок [13].

Заключение

В настоящем исследовании в результате SNP-анализа 2868 геномов штаммов *S. Enteritidis*, построена дендрограмма, выявляющая филогенетические связи исследованных штаммов. Определены маркерные мутации, характерные для каждой отдельной генетической линии дендрограммы. Полученные данные заложены в авторскую программу Salmonella Analyzer 2, которая обеспечивает классификацию любой выборки штаммов *S. Enteritidis*. Определение маркерных мутаций анализ и иерархичность используемой в программе классификации минимизирует получение ошибочных либо сомнительных результатов. При помощи классификации генетических линий по маркерным мутациям решается проблема гомоплазии в случаях, когда одинаковые SNP могут возникать независимо в неродственных линиях, создавая ложные кластеры. Разработанная программа использует эффективную систему идентификации и типирования штаммов *S. Enteritidis*, что позволяет её использовать не только при мониторинговых исследованиях, но и для прогнозирования развития эпидемической ситуации.

Литература

1. Jiang M, Zhu F, Yang C, et al. Whole-genome analysis of *Salmonella enterica* Serovar Enteritidis isolates in outbreak linked to online food delivery, Shenzhen, China, 2018. *Emerg Infect Dis.* 2020;26(4):789. doi:10.3201/eid2604.191446.
2. Dallman T, Inns T, Jombart T, et al. Phylogenetic structure of European *Salmonella* Enteritidis outbreak correlates with national and international egg distribution network. *Microb Genom.* 2016;2(8):e000070. doi:10.1099/mgen.0.000070.
3. Chattaway MA, Painset A, Godbole G, et al. Evaluation of genomic typing methods in the *Salmonella* reference laboratory in Public Health, England, 2012–2020. *Pathogens.* 2023;12(2):223. doi:10.3390/pathogens12020223.
4. Prijbelski A, Antipov D, Meleshko D, et al. Using SPAdes De Novo Assembler. *Curr Protoc Bioinformatics.* 2020;70(1):e102. doi:10.1002/cpbi.102.
5. GitHub – tseemann/snippy: Rapid haploid variant calling and core genome alignment [Электронный ресурс]. Режим доступа: <https://github.com/tseemann/snippy>. Дата обращения: 27.11.2023.
6. Tamura K, Stecher G, Kumar S. MEGA11: molecular evolutionary genetics analysis version 11. *Mol Biol Evol.* 2021;38(7):3022–3027. doi:10.1093/molbev/msab120.
7. Huerta-Cepas J, Serra F, Bork P. ETE 3: reconstruction, analysis, and visualization of phylogenomic data. *Mol Biol Evol.* 2016;33(6):1635–1638. doi:10.1093/molbev/msw046.
8. Fenske GJ, Thachil A, McDonough PL, et al. Geography shapes the population genomics of *Salmonella enterica* Dublin. *Genome Biol Evol.* 2019;11(8):2220–2231. doi:10.1093/gbe/evz158.
9. Hillis DM, Bull JJ. An Empirical Test of Bootstrapping as a Method for Assessing Confidence in Phylogenetic Analysis. *Syst Biol.* 1993;42(2):182–192. doi:10.1093/sysbio/42.2.182.
10. Zhou Z, Alikhan NF, Mohamed K, et al. The Enterobase user's guide, with case studies on *Salmonella* transmissions, *Yersinia pestis* phylogeny, and *Escherichia* core genomic diversity. *Genome Res.* 2020;30(1):138–152. doi:10.1101/gr.251678.119.
11. Katz LS, Griswold T, Williams-Newkirk AJ, et al. A comparative analysis of the Lyve-SET phylogenomics pipeline for genomic epidemiology of foodborne pathogens. *Front Microbiol.* 2017;8:375. doi:10.3389/fmicb.2017.00375.
12. Егорова СА, Кулешов КВ, Кафтырева ЛА. Современные методы субтипирования сальмонелл при расследовании вспышек сальмонеллеза. *Имунопатология, аллергология, инфектология.* 2019;3:36–42. doi:10.14427/jipai.2019.3.33.
13. Adhikari Y, Bailey MA, Kitchens S, et al. Whole-genome sequencing and phylogenetic analysis of *Salmonella* isolated from pullets through final raw product in the processing plant of a conventional broiler complex: a longitudinal study. *Microbiol Spectr.* 2025;13(2):e0209024. doi:10.1128/spectrum.02090-24.

Сведения об авторах

Горох Алефтина Михайловна – младший научный сотрудник ФКУЗ Ростовский-на-Дону противочумный институт Роспотребнадзора. E-mail: gorokh_am@antiplague.ru.
Водопьянов Алексей Сергеевич – кандидат медицинских наук, ведущий научный сотрудник ФКУЗ Ростовский-на-Дону противочумный институт Роспотребнадзора. E-mail: vodopyanov_as@antiplague.ru.
Герасименко Артем Александрович – младший научный сотрудник ФКУЗ Ростовский-на-Дону противочумный институт Роспотребнадзора. E-mail: gerasimenko_aa@antiplague.ru.
Писанов Руслан Вячеславович – кандидат биологических наук, ведущий научный сотрудник ФКУЗ Ростовский-на-Дону противочумный институт Роспотребнадзора. E-mail: pisanov_rv@antiplague.ru.

Поступила 26.12.2025.