

Министерство науки и высшего образования РФ
Федеральное государственное автономное
образовательное учреждение высшего образования
«СИБИРСКИЙ ФЕДЕРАЛЬНЫЙ УНИВЕРСИТЕТ»

Институт фундаментальной биологии и биотехнологии
Кафедра геномики и биоинформатики

УТВЕРЖДАЮ:

Заведующая кафедрой

_____ И. Е. Ямских

«__» _____ 2024 г.

МАГИСТЕРСКАЯ ДИССЕРТАЦИЯ

«Сравнительное исследование ансамбля генов транспортных РНК бактерий
кишечной микрофлоры человека по их триплетному составу»

06.04.01 Биология

06.04.01.06 Геномика и биоинформатика

Руководитель

подпись, дата

проф., д.ф.-м.н.

М.Г. Садовский

Выпускник

подпись, дата

М.А. Мостафа

Рецензент

подпись, дата

доцент, к.ф.-м.н.

О. А. Золотов

Красноярск 2024

РЕФЕРАТ

Магистерская диссертация по теме «Сравнительное исследование ансамбля генов транспортных РНК бактерий кишечной микрофлоры человека по их триплетному составу» содержит 29 страниц текстового документа, 17 иллюстраций, 3 таблицы и 19 использованных источников.

Ключевые слова: ТРНК, БАКТЕРИИ, ТРИПЛЕТНЫЙ СОСТАВ, МИКРОФЛОРА ЧЕЛОВЕКА.

Объект исследования — взаимосвязь между триплетным составом тРНК и степенью патогенности бактерий.

Предмет исследования — статистические свойства тРНК бактерий (особенности триплетного состава).

Цель работы — выявление связи степени патогенности бактерий с частотным составом триплетов генов тРНК.

В ходе проделанной работы было проанализировано 2828 генов тРНК для 26 видов бактерий различной патогенности. В результате сравнительного анализа генов тРНК бактерий было обнаружено различие между патогенными и непатогенными бактериями на основании триплетного состава гена тРНК, которых участвует в транспортировке лейцина (tRNA-Leu-TAG). Наибольший вклад в различимость бактерий по патогенности вносят частоты следующих триплетов: TTT, TGT, GTT, GTG, GCA. Филогенетический анализ также подтвердил распределение на группы. Также с помощью метода упругих карт была обнаружена кластеризация, которая согласуется с распределением генов по типу аминокислот, их строению и свойствам.

Работа может быть полезна для определения патогенности/непатогенности неизвестных тРНК бактерий человека.

СОДЕРЖАНИЕ

ВВЕДЕНИЕ	4
1 Обзор литературы	5
1.1 Транспортная РНК	5
1.1.1 Важность тРНК в живых организмах	5
1.1.2 Структура тРНК	6
1.2 Особенности бактериальной микрофлоры человека.....	8
1.2.1 Непатогенные бактерии.....	8
1.2.2 Условно-патогенные бактерии	9
1.2.3 Патогенные бактерии.....	9
1.2.4 Проблема устойчивости к антибиотикам	10
2 Материалы и методы	13
3 Результаты и их обсуждение.....	15
3.1 Результаты анализа данных методом упругих карт	Ошибка! Закладка не определена.
3.1.1 Состав кластеров по степени патогенности бактерий	Ошибка! Закладка не определена.
3.1.2 Состав кластеров по видам аминокислот	Ошибка! Закладка не определена.
3.1.3 Состав кластеров по строению и свойствам аминокислот	Ошибка! Закладка не определена.
3.2 Результаты анализа данных методом главных компонент (РСА) ..	Ошибка! Закладка не определена.
3.2.1 Анализ частот триплетов генов тРНК	Ошибка! Закладка не определена.

3.2.2 Филогенетический анализ тРНК .. **Ошибка! Закладка не определена.**

ЗАКЛЮЧЕНИЕ	16
СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ	17

ВВЕДЕНИЕ

Бактерии имеют большое значение в жизни человека. Они могут как полезны, например участвуют в процессе пищеварения, синтезируют некоторые витамины, препятствуют деятельности болезнетворных бактерий. Также могут быть и вредны – патогенные и условно-патогенные бактерии являются причиной различных заболеваний, тем самым повышая воспалительные процессы в организме человека.

Для сравнительного анализа выбраны гены тРНК, так как тРНК играют ключевую роль в синтезе белков – жизненно необходимых молекул, которые участвуют в защитной, строительной, регулярной и других функциях в организме. Также известно, что вариации в тРНК могут оказывать влияние на фенотип. Например, у патогенного гриба рода кандиды один вид тРНК может транспортировать две различные аминокислоты. Это приводит к образованию разных вариантов белков, что проявляется в разнообразии строения гриба. А также благодаря этому повышается его выживаемость в кишечной микрофлоре человека [1].

Сравнительный анализ позволит ответить нам: «Можно ли с помощью профилирования бактериальной тРНК определить, принадлежит ли новая последовательность тРНК к патогенным или непатогенным бактериям?»

Целью данной работы является выявление связи степени патогенности бактерий с частотным составом триплетов генов тРНК.

Для достижения данной цели были поставлены следующие **задачи**:

- 1) Получение данных тРНК бактерий различной патогенности;
- 2) Подсчёт частот триплетов для последовательностей тРНК бактерий;
- 3) Кластеризация данных методом PCA и визуализация результатов в VidaExpert.

Объектом работы является взаимосвязь между триплетным составом тРНК и патогенностью бактерий. Предмет работы – статистические свойства тРНК бактерий (особенности триплетного состава).

1 Обзор литературы

1.1 Транспортная РНК

1.1.1 Важность тРНК в живых организмах

тРНК играет важную роль в синтезе белка, выполняя функцию адаптера, который переносит информацию от нуклеиновых кислот к белкам. Она помогает нам идентифицировать болезни и может быть использована для определения способов их предотвращения [2].

Природные белки состоят из 20 аминокислот, которые являются их основными компонентами. Белки имеют различный размер, последовательность и трёхмерную структуру, и многие мутации могут влиять на их функции. Например, одиночная мутация в белковой последовательности способна вызвать остановку его работы или наоборот – усилить его активность, либо может быть нейтральной, т.е. не оказывать влияния на функции. Недавние исследования позволили создать генетически модифицированные аминокислоты. Например, показано, что теперь возможно эффективно переназначить два стоп-кодона у *E. coli* для производства белков с двумя неестественными аминокислотами [3].

Авторы подчеркивают способность включать пару флуорофоров, обеспечивающих передачу энергии Фёрстера, в один белок. Эта технология позволит проводить точные измерения динамики белков в живых клетках [4].

Первое доказательство расширения генетического кода за пределы канонических 20 аминокислот было найдено в открытии 21-й генетически кодируемой аминокислоты — селеноцистеина [5].

тРНК участвует во многих клеточных процессах помимо трансляции, например, в аминокислотации липидов и конъюгации бактерий. Также бактерии разработали способы сопротивляться антибиотикам [6].

Изучение тРНК позволяет нам глубже понять её важную роль в клеточных процессах и её потенциал в биомедицинских приложениях, таких как генная терапия и идентификация бактерий [7].

1.1.2 Структура тРНК

Общая длина тРНК составляет от 76 до 90 оснований, что может создать впечатление, что это маленькие и лёгкие молекулы, но на самом деле их средняя масса превышает 26,000 г/моль.

Рассмотрим более подробно двумерное представление тРНК в форме клеверного листа (рис.1, А). От 5' до 3' области тРНК включают: акцепторный стебель (зелёный), D-руку (розовый), антикодонную руку (синяя) с антикодоном (тёмно-синий), переменную руку (оранжевая) и Т-руку (фиолетовая). Дискриминаторная база (база 73) — это непарная база на 3' конце (жёлтая) перед конечными остатками ССА. Аминоацил-тРНК синтетазы присоединяют аминокислоту к конечному 3' аденозину. Пронумерованные позиции — это постоянные основания [8].

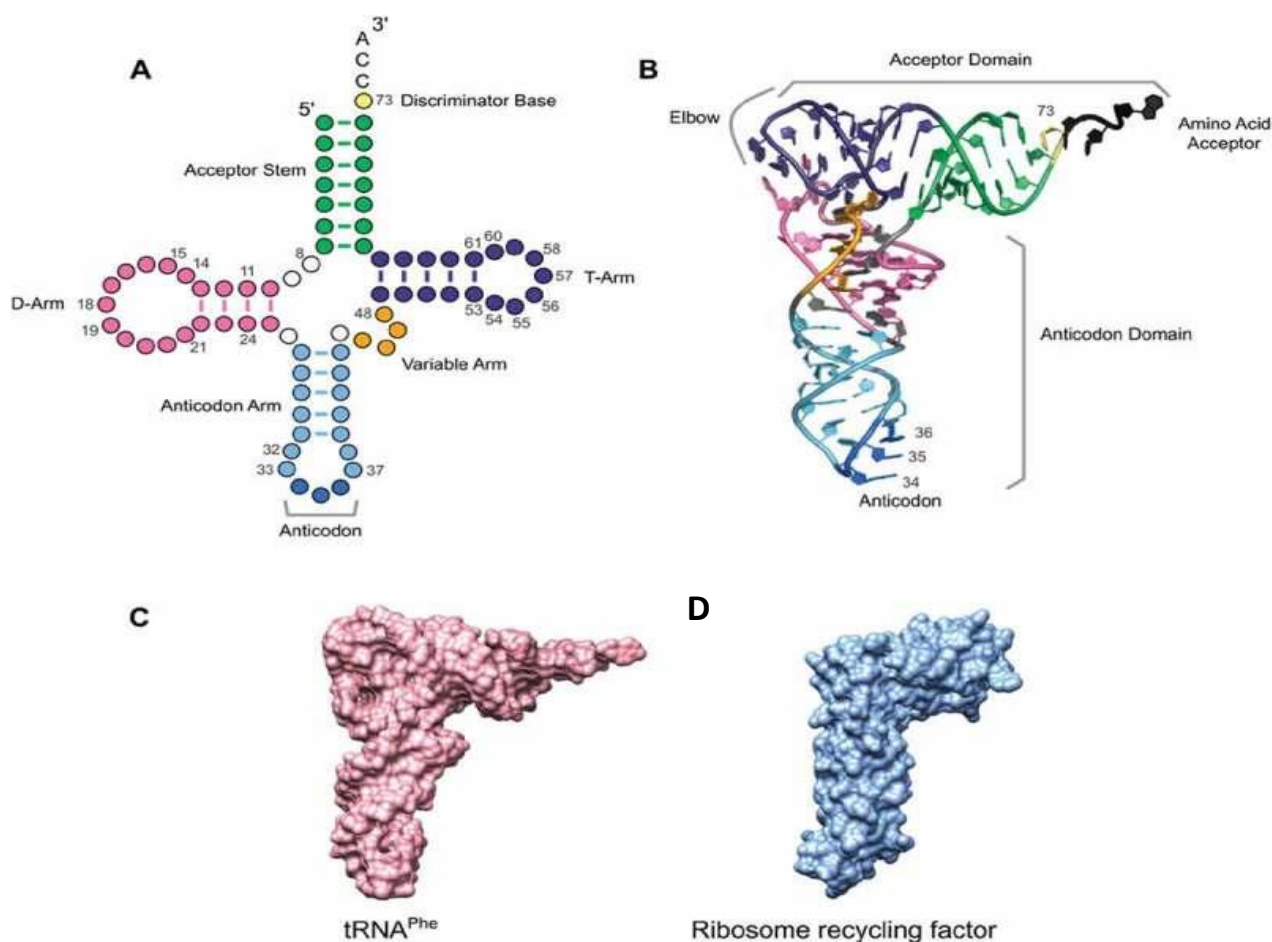


Рисунок 1 – Структура тРНК [8]

Д-петля названа в честь модифицированной базы дигидроуридина, присутствующей в петле, и отвечает за стабилизацию третичной структуры тРНК. Т-рука, также известная как ТΨС-рука, названа так из-за присутствия универсально сохранённых модифицированных оснований тимидина, псевдоуридина и цитидина [6].

Как следует из названия, длина переменной руки варьируется. У эукариот она наиболее выражена у тРНК серина (tRNA^{Ser}), селеноцистеина (tRNA^{Sec}) и лейцина (tRNA^{Leu}). Более длинные переменные руки также встречаются у эубактериальных и органеллярных тРНК тирозина (tRNA^{Tyr}) и в систематическом поиске других тРНК с переменными руками [6].

Рассмотрим трёхмерную структура тРНК на примере тРНК-Phe дрожжей (рис. 1, B). Цвет двухмерных элементов такой же, как и у двумерной структуры (рис. 1, A). Внутримолекулярные взаимодействия между Т- и Д-руками

(фиолетовая и розовая соответственно) в области локтя способствуют сворачиванию тРНК [8].

Поверхностное представление тРНК-Phe и фактора рециркуляции рибосомы (рис. 1, D) демонстрирует, как молекулы, взаимодействующие с трансляционным аппаратом, часто принимают структуры, похожие на тРНК [8].

тРНК обладают способностью переносить одни и те же аминокислоты и называются изоакцепторами. Они различаются по антикодонам, что справедливо для всех аминокислот, кроме триптофана и метионина. Изоакцепторы схожи, но не идентичны. Существует 21 тип изоакцепторов: по одному для каждой из 20 стандартных аминокислот и один для селеноцистеина. Три стоп-кодона представлены белками факторов освобождения, которые завершают трансляцию [9].

1.2 Особенности бактериальной микрофлоры человека

1.2.1 Непатогенные бактерии

Полезные бактерии, также известные как пробиотики, обитают в различных частях человеческого тела, особенно в желудочно-кишечном тракте, где они способствуют пищеварению, иммунной функции и общему здоровью. Эти бактерии помогают расщеплять сложные углеводы, синтезировать витамины и вытеснять патогены, способствуя тем самым здоровью кишечника. Помимо кишечника, полезные бактерии населяют такие области, как кожа и ротовая полость, где они оказывают защитное действие против инфекций и поддерживают здоровье кожи и зубов.

В последние годы исследования показали связь между человеческим микробиомом, состоящим из разнообразных бактериальных сообществ, и

различными аспектами здоровья, включая метаболизм, иммунитет и даже психическое благополучие. Кроме того, пробиотические бактерии всё чаще признаются за их потенциальное терапевтическое применение в лечении таких состояний, как синдром раздражённого кишечника, воспалительные заболевания кишечника и даже аллергии. Они также исследуются на предмет их роли в поддержке общей иммунной функции и предотвращении некоторых инфекций.

Понимание важности полезных бактерий в организме человека подчёркивает значимость поддержания сбалансированного микробиома для оптимального здоровья. Использование потенциала этих бактерий обещает значительные достижения в области персонализированной медицины и стратегий профилактического здравоохранения [10].

1.2.2 Условно-патогенные бактерии

К условно-патогенным микроорганизмам относятся представители нормальной микрофлоры человека и животных (*Escherichia coli*, *Enterococcus faecalis*, *Streptococcus epidermidis*, *Proteus vulgaris* и др.), обитающие на коже, слизистых оболочках органов и систем, сообщающихся с внешней средой. В здоровом организме нормальная микрофлора создает конкурентные условия для патогенных микробов, подавляя их рост и размножение, и оказывает стимулирующее влияние на развитие и функционирование иммунной системы. Присущие им потенциально-патогенные свойства условно-патогенные микробы проявляют при ослаблении защитных сил организма, когда они могут вызывать различные заболевания, главным образом болезни, поражающие дыхательные и пищеварительные тракты [11].

1.2.3 Патогенные бактерии

Патогенные бактерии представляют значительную угрозу здоровью человека, вызывая широкий спектр инфекционных заболеваний. Эти бактерии эволюционировали, приобретая различные механизмы для вторжения, колонизации и нанесения вреда человеческому организму, что приводит к заболеваниям, варьирующимся от лёгких инфекций до опасных для жизни состояний. В прошлом такие случаи были более распространены из-за отсутствия способов предотвращения бактериальных инфекций, таких как антибиотики.

Патогенные бактерии могут проникать в организм различными путями, включая приём пищи, вдыхание или прямой контакт с заражёнными поверхностями или людьми. Попад в организм, они могут нацеливаться на определённые ткани или органы, вызывая локализованные инфекции, такие как инфекции мочевыводящих путей, кожные инфекции или инфекции дыхательных путей. В тяжёлых случаях они могут распространяться по всему телу.

Распространённые примеры патогенных бактерий включают виды, такие как *Clostridium botulium* (вызывает ботулизм) *Mycobacterium bovis* (вызывает туберкулёз), *Pasteurella multocida* (вызывает пневмонию, менингит) и другие. Эти бактерии производят токсины, ферменты или другие факторы вирулентности, которые позволяют им избегать защитных механизмов хозяина и наносить ущерб тканям хозяина.

1.2.4 Проблема устойчивости к антибиотикам

Устойчивость к антибиотикам представляет значительную проблему в лечении бактериальных инфекций, так как некоторые патогенные бактерии развили устойчивость к нескольким антибиотикам. Эта устойчивость может осложнить лечение и увеличить риск неудачи лечения, что приводит к затяжным заболеваниям и повышению уровня смертности. Некоторые бактерии

развили эту устойчивость из-за неправильного использования антибиотиков людьми.

Понимание механизмов, с помощью которых патогенные бактерии вызывают заболевания, и факторов, способствующих их устойчивости к антибиотикам, имеет решающее значение для разработки новых стратегий борьбы с бактериальными инфекциями. Это включает как профилактические меры, такие как улучшение контроля за использованием антибиотиков, так и разработку новых терапевтических подходов [12].

Недавний отчёт Всемирной организации здравоохранения (ВОЗ) ясно указывает на то, что это явление происходит не только в бедных или развивающихся странах; проблема устойчивости к антибиотикам (AMR) сейчас встречается повсеместно по всему миру [13].

В развивающихся странах выявлены ключевые факторы, включая:

- 1) отсутствие мониторинга развития устойчивости;
- 2) низкое качество контроля за имеющимися антибиотиками;
- 3) клиническое неправильное использование;
- 4) лёгкость доступа.

В отличие от этого, в развитых странах факторы, способствующие устойчивости к антибиотикам, включают:

- 1) недостаточное регулирование на уровне больниц;
- 2) чрезмерное использование антибиотиков в продуктивном животноводстве.

Наконец, имеется значительный недостаток исследований новых антибиотиков, что затрагивает как развивающиеся, так и развитые страны, как показано на рисунке 2 [14].



Рисунок 2 – Основные социальноэкономические и политические факторы, способствующие устойчивости к антибиотикам в развивающихся (зелёный), развитых (красный) странах, а также в обеих развивающихся и развитых странах (серый)¹⁴

Профилактические меры, такие как вакцинация, соблюдение хорошей гигиены, правильная обработка пищи и санитария, играют ключевую роль в снижении передачи вредных бактерий и предотвращении инфекций. Понимание характеристик и поведения патогенных бактерий является важным для разработки эффективных стратегий борьбы с инфекционными заболеваниями и защиты общественного здоровья.

2 Материалы и методы

Данные последовательностей тРНК скачаны из базы данных GtRNAdb (<http://gtgnadb.ucsc.edu>). Систематическое положение исследуемых бактерий представлено в таблице 1.

Таблица 1 – Систематика исследуемых бактерий

Тип (phy)	Класс	Порядок	Семья	Род (Gen)
Actinobacteria	Actinobacteria	Bifidobacteriales	Bifidobacteriaceae	<i>Bifidobacterium</i>
Firmicutes	Bacilli	Lactobacillales	Lactobacillaceae	<i>Lactobacillus</i>
		Bacillales	Bacillaceae	<i>Bacillus</i>
			Staphylococcaceae	<i>Staphylococcus</i>
			Listeriaceae	<i>Listeria</i>
	Clostridia	Clostridiales	Clostridiaceae	<i>Clostridium</i>
Proteobacteria	Gammaproteobacteria	Enterobacteriales	Enterobacteraceae	<i>Escherichia</i>
				<i>Cronobacter</i>
				<i>Salmonella</i>
				<i>Yersinia</i>
		Vibrionales	Vibrionaceae	<i>Vibrio</i>
	Betaproteobacteria	Burkholderiales	Alcaligenaceae	<i>Bordetella</i>
	Epsilonproteobacteria	Campylobacteriales	Campylobacteraceae	<i>Campylobacter</i>
Spirochaetota	Spirochaetia	Spirochaetales	Treponemataceae	<i>Treponema</i>

Данные анализировались двумя методами. Первый – метод упругих карт. Для этого для каждой последовательности тРНК бактерий осуществлялся подсчёт частот триплетов с помощью программы python, затем производилась их кластеризация методом упругих карт в программе *VidaExpert* [15]. Алгоритм подсчёта частоты триплетов состоит в том, что для каждой последовательности тРНК считается количество вхождений триплетов в последовательность, затем значение нормализуется путём деления на сумму всех частот триплетов.

Второй – метод главных компонент. Сначала были выбраны такие типы тРНК, которые присутствуют во всех исследуемых организмах. Далее для каждого гена производился подсчёт частоты триплетов с помощью программы *ad hoc* на python. Затем для каждого гена тРНК было визуализировано распределение данных методом главных компонент с помощью программы Past3 [16]. Для тех тРНК, в которых группы патогенных и непатогенных бактерий не пересекаются, были построены филогенетические деревья с помощью программы MEGA-X [17] методом ближайших соседей (Neighbor-Joining). Для генов тРНК, для которых разделение на группы патогенности подтвердилось филогенетически, были определены ключевые триплеты, которые вносят наибольший вклад в распределение частот триплетов в пространстве главных компонент.

3 Результаты и их обсуждение

В связи с авторским правом изъято 13 страниц

ЗАКЛЮЧЕНИЕ

В ходе сравнительного анализа ансамбля генов транспортных РНК бактерий кишечной микрофлоры человека по их триплетному составу было обнаружено различие между патогенными и непатогенными бактериями на основании триплетного состава гена тРНК, которых участвует в транспортировке лейцина (tRNA-Leu-TAG). Наибольший вклад в различимость бактерий по патогенности вносят частоты следующих триплетов: TTT, TGT, GTT, GTG, GCA. Филогенетический анализ также подтвердил распределение на группы.

Таким образом, поставленная **цель достигнута** – выявлена связь степени патогенности бактерий с частотным составом триплетов генов тРНК.

Также с помощью метода упругих карт была обнаружена кластеризация, которая согласуется с распределением генов по типу аминокислот, их строению и свойствам.

Работа может быть полезна для определения патогенности/непатогенности неизвестных тРНК бактерий человека.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Ling, J., O'Donoghue, P., Söll, D. Genetic code flexibility in microorganisms: novel mechanisms and impact on physiology // *Nature reviews microbiology*. – 2015. – Vol. 13. – P. 707–721.
2. O'Donoghue, P., Ling, J., Söll, D. Transfer RNA function and evolution // *RNA Biol.* – 2018. – Vol. 15, Iss. 4–5. – P. 423–426.
3. Tharp, J. M., Ehnbohm, A., Liu, W.R. tRNAPyl: Structure, function, and applications // *RNA Biol.* – 2018. – Vol. 15, Iss. 4-5. – P. 441–452.
4. Wright, D. E., Altaany, Z., Bi, Y., Alperstein, Z., O'Donoghue, P. Acetylation Regulates Thioredoxin Reductase Oligomerization and Activity // *Antioxid Redox Signal.* – 2018. – Vol. 29, Iss. 4. – P. 377–388.
5. Zinoni, F., Birkmann, A., Leinfelder, W., Böck, A. Cotranslational insertion of selenocysteine into formate dehydrogenase from *Escherichia coli* directed by a UGA codon // *Proc Natl Acad Sci U S A.* – 1987. – Vol. 84, Iss. 10. P. 3156–3160.
6. Holley, R. W., Apgar, J., Everett, G. A., Madison, J. T., Marquisee, M., Merrill, S. H., Penswick, J. R., Zamir, A. Structure of a ribonucleic acid // *Science.* – 1965. – Vol. 147, Iss. 3664 – P. 1462–1465.
7. Hamashima, K., Tomita, M., Kanai, A. Expansion of Noncanonical V-Arm-Containing tRNAs in Eukaryotes // *Mol Biol Evol.* – 2016. – Vol. 33, Iss. 2. – P. 530–540.
8. Berg, M. D., Brandl, C. J. Transfer RNAs: diversity in form and function // *RNA Biol.* – 2021. – Vol. 18, Iss. 3. – P. 316–339.
9. Kisselev, L. Polypeptide release factors in prokaryotes and eukaryotes: same function, different structure // *Structure.* – 2002. – Vol. 10, Iss. 1. – P. 8–9.
10. Davis, C. P. Normal Flora. In: Baron S, editor. *Medical Microbiology* / C. P. Davis. – 4th edition. Galveston (TX): University of Texas Medical Branch at Galveston, 1996. – Chapter 6. – Available from: <https://www.ncbi.nlm.nih.gov/books/ NBK7617/>

11. Роль условно-патогенной микрофлоры в инфекционном процессе : Учебно-методическое пособие для самостоятельной работы студентов / А. К. Галиуллин, Ф. М. Нургалиев, Ф. М. Софронов, Р. А. Волков, А. Ю. Шаева. – Казань: КГАВМ, 2020. – с. 11.
12. Llor, C., Bjerrum L. Antimicrobial resistance: risk associated with antibiotic overuse and initiatives to reduce the problem // *Ther Adv Drug Saf.* – 2014. – Vol. 5, Iss. 6. – P. 229–241.
13. World Health Organization. WHO'S first global report on antibiotic resistance reveals serious, worldwide threat to public health. Antimicrobial resistance – global surveillance report. Virtual Press Conference. – 2014. – URL: <http://www.who.int/mediacentre/multimedia/amr-transcript.pdf?ua=1>.
14. Chokshi, A., Sifri, Z., Cennimo, D., Horng, H. Global Contributors to Antibiotic Resistance / *J Glob Infect Dis.* – 2019. – Vol. 11, Iss. 1. – P. 36–42.
15. Gorban, A. Zinovyev, A., Pitenko, A. ViDaExpert: user-friendly tool for nonlinear visualization and analysis of multidimensional vectorial data // *Proceedings of IEEE International Conference on Data Science and Advanced Analytics (DSAA).* – 2014.
16. Hammer, Ø., Harper, D. A.T., Paul, D. R. Past: Paleontological Statistics Software Package for Education and Data Analysis // *Palaeontologia Electronica.* – 2001. – Vol. 4, Iss. 1. – P. 1–9.
17. Kumar, S., Stecher, G., Li, M., Knyaz, C., Tamura, K. MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms // *Molecular Biology and Evolution.* – 2018. – Vol. 35. – P. 1547–1549.
18. СанПиН 3.3686-21 «Санитарно-эпидемиологические требования по профилактике инфекционных болезней». Приложение 1. ПБА по группам патогенности // КонсультантПлюс : справочная правовая система. – URL: <http://www.consultant.ru>. (дата обращения: 10.06.2024).
19. Березов Т. Т., Коровкин Б. Ф. Классификация аминокислот // *Биологическая химия.* – 3-е изд., перераб. и доп.. – М.: Медицина, 1998. – 704 с. – ISBN 5-225-02709-1.

Министерство науки и высшего образования РФ
Федеральное государственное автономное
образовательное учреждение высшего образования
«СИБИРСКИЙ ФЕДЕРАЛЬНЫЙ УНИВЕРСИТЕТ»

Институт фундаментальной биологии и биотехнологии
Кафедра геномики и биоинформатики

УТВЕРЖДАЮ:

Заведующая кафедрой

 И. Е. Ямских

«27» июня 2024 г.

МАГИСТЕРСКАЯ ДИССЕРТАЦИЯ

«Сравнительное исследование ансамбля генов транспортных РНК бактерий
кишечной микрофлоры человека по их триплетному составу»

06.04.01 Биология

06.04.01.06 Геномика и биоинформатика

Руководитель	 20.06.2024 подпись, дата	проф., д.ф.-м.н.	М.Г. Садовский
Выпускник	Мостафа Мехлуа подпись, дата		М.А. Мостафа
Рецензент	 подпись, дата	доцент, к.ф.-м.н.	О. А. Золотов

Красноярск 2024