

УДК:577.2:006
DOI: 10.52419/issn2072-2419.2025.3.563

ВЛИЯНИЕ МЕТОДОЛОГИЧЕСКИХ РЕШЕНИЙ НА РЕЗУЛЬТАТЫ МЕТАГЕНОМНОГО СЕКВЕНИРОВАНИЯ БАКТЕРИАЛЬНЫХ СООБЩЕСТВ

Прасолова О.В. – канд. ветеринар. наук, вед. науч. сотр. (ORCID 0000-0001-8924-2273)

ФГБУ «Всероссийский государственный центр качества
и стандартизации лекарственных средств для животных и кормов»

*o.prasolova@vgnki.ru

Ключевые слова: микробиом, стандартизация, валидация, секвенирование, генетика

Keywords: microbiome, standardization, validation, sequencing, genetics

Поступила: 20.08.2025

Принята к публикации: 26.08.2025

Опубликована онлайн: 15.09.2025

РЕФЕРАТ



Термин микробиом охватывает совокупность всех генов, принадлежащих микроорганизмам (бактериям, вирусам, грибам), которые населяют определённую среду. Секвенирование участка гена 16S рРНК позволяет изучить таксономический состав микробиоты, выявляя бактерии и археи, в том числе сложно поддающиеся культивированию. Однако данный метод до сих пор не имеет стандартизированной процедуры. Множество доступных методологических вариантов усложняют воспроизведение результатов и, как правило, ограничивают сопоставимость результатов между независимыми исследованиями, использующими разные методики, измерительные и биоинформатические конвейеры. Интерпретация, обсуждение и визуализация результатов исследования микробиома животных представляют собой сложную задачу из-за отсутствия единых стандартных параметров и справочных данных для сбора и сравнения. Стандартизированные методологии при исследовании сообществ микроорганизмов необходимы для упрощения и сравнения результатов различных исследований. В данном анализе представлены основные причины расхождений в получаемых данных на всех этапах работ. К ним относятся отбор проб (место отбора проб, метод, транспортировка и хранение образца), извлечение материала нуклеиновой кислоты, выбор праймера 16S рРНК, амплификация, подготовка библиотеки, секвенирование и конвейер биоинформатического анализа. Проблемы, связанные с таксономическим и функциональным профилированием на основе метагеномных данных, включают в себя вопросы точности классификации в условиях систематических ошибок в базах данных, сложности прогнозирования функций и метаболических взаимодействий в различных микробных сообществах, а также отсутствие единого методологического подхода к проведению исследования. Преодоление этих препятствий потребует междисциплинарного сотрудничества, развития вычислительных инструментов и совершенствования аналитических методов.

ВВЕДЕНИЕ / INTRODUCTION

Сложность культивирования большинства микроорганизмов, присутствующих в естественных или управляемых системах, вынуждает микробиологов использовать молекулярно-генетический подход, а именно ген 16S рРНК в качестве филогенетического маркера при изучении таксономического состава микрорекосистем [1]. Ген 16S рРНК, кодирует 30S субъединицу прокариотической рибосомы, имеет длину приблизительно 1600 нуклеотидов. Он высококонсервативен у всех видов бактерий и содержит девять гипервариабельных участков разной степени консервативности (V1–9). Последовательности генов 16S рРНК, полученные из микробиомов, обычно кластеризуются в таксономические единицы (OTU, operational taxonomic unit) на нескольких уровнях для определения видового богатства, разнообразия, состава и структуры сообщества. Более консервативные участки полезны для определения таксонов более высокого ранга, тогда как более быстро эволюционирующие участки помогают при идентификации рода или вида. Определение таксономии на основе ограниченного сегмента гена 16S рРНК, например, только области V4, может быть проблематичным. Во многих случаях классифицируемая последовательность практически идентична нескольким другим бактериальным последовательностям в референтной базе данных. Аналогично, при метагеномном анализе, когда извлекаются только части бактериального генома, классификация на таксономическом уровне будет зависеть от степени консервативности доступных последовательностей. Таким образом, определение таксономии видов, содержащих очень похожие последовательности, будет сложнее, и анализ будет накапливать большую долю прочтений на более высоких уровнях таксономического дерева. Для ~ 42% бактериальных родов будут пары последовательностей внутри рода, которые не могут быть достоверно разделены, потому что их последовательности генов 16S рРНК схожи более чем на 97% [2]. Использо-

вание для анализа, областей V3 и V4 совместно, является золотым стандартом для идентификации бактерий на уровне рода [1].

Интерпретация, обсуждение и визуализация результатов исследования микробиома животных представляют собой сложную задачу из-за отсутствия единых стандартных параметров и справочных данных для сбора и сравнения. Стандартизированные методологии при исследовании сообществ микроорганизмов необходимы для упрощения и сравнения результатов различных исследований [3].

Важной проблемой метагеномного исследования является валидация анализа с учетом используемой матрицы, который нельзя считать валидированным, пока он не будет практически использован большинством исследователей, что сделает возможным осуществление межлабораторной воспроизводимости результатов [4].

МАТЕРИАЛЫ И МЕТОДЫ / MATERIALS AND METHODS

Нами проведен анализ методологического ландшафта для исследования бактериальных сообществ и влияния методологических решений на результаты метагеномного секвенирования в рамках опубликованных исследований из различных баз данных Elibrary, PubMed, NCBI, Elsevier, Web of Science.

РЕЗУЛЬТАТЫ / RESULTS

При попытках стандартизации исследований микробиомных сообществ было предпринято множество усилий, направленных на получение сопоставимых данных аналитических характеристик измерений с точки зрения чувствительности, специфичности, точности, воспроизводимости и т. д.

Однако, существует необходимость исследований для определения приоритетов и изучения источников изменчивости и смещения в экспериментальном рабочем процессе, а также потребность в стандартизированных материалах и методах для улучшения сопоставимости и охвата результатов измерений.

Одной из возможных причин расхож-

дений в получаемых данных является введение системных смещений из-за различий в дизайне исследования, включая методологии и аналитические рабочие процессы в разных исследованиях. Эти смещения начинаются с момента сбора образцов и продолжают вноситься на протяжении всего эксперимента, что приводит к наблюдаемому сообществу, которое значительно отличается от истинного основного микробного состава. Выявление того, как факторы в подходе к работе с образцами по-разному способствуют этим микробным смещениям между исследованиями, имеет важное значение и требует использования высокочувствительных методов метагеномного секвенирования в большой когорте при различных условиях. Тщательно разработанный дизайн эксперимента увеличивает потенциал получения высококачественных, воспроизводимых данных и гарантирует научную обоснованность полученных выводов. Для обеспечения статистической достоверности необходимо учитывать уровень выборки в эксперименте. [5]. Дизайн исследования должен быть согласован с конкретными исследовательскими целями, с учетом исследуемого материала (от человека, животных или окружающей среды), поскольку он напрямую влияет на валидность и интерпретируемость результатов. [6]. В исследованиях микробиома животных использование материала, с помощью неинвазивных методов (образцы фекалий, ректальные мазки), значительно упрощает процесс, однако, например, содержимое слепой кишки предоставляет возможность получения полного профиля микробного сообщества кишечника, но требует эвтаназии животного. Поскольку микробиомные сообщества представляют собой живые и метаболически активные микроорганизмы, образцы следует замораживать или инактивировать сразу после сбора, а затем хранить в стабильных условиях до выделения ДНК. Следует избегать колебаний температуры, воздействия кислорода и многократных циклов замораживания и размораживания поскольку они

могут нарушить целостность ДНК и РНК, влияя на метагеномные профили микробных сообществ [7].

Золотым стандартом метода обработки образцов после стерильного сбора является быстрое замораживание в жидком азоте с последующим хранением при температуре -80°C . Однако это часто невозможно в полевых условиях, и предложены альтернативные варианты стабилизации и хранения образцов: консервирующие буферы, стабилизирующие агенты на основе этанола или гуанидинизотиоцианата, которые доступны в коммерческих наборах [8]. С целью обеспечения воспроизводимости результатов, рекомендуем указывать в метаданных исследования сведения об обработке и хранении образцов.

Из всех этапов подготовки образцов к исследованию способ выделения нуклеиновой кислоты может оказывать наибольшее влияние на вариабельность исследования из-за уникальных смещений в различных подходах к микробному лизису. Тип используемого метода извлечения ДНК зависит от типа платформы секвенирования, которая будет использоваться (короткое считывание или длинное считывание), а также от вида материала, однако проведенные сравнительные исследования многих ученых не смогли создать стандартизированную процедуру. Включение отрицательных контролей имеет решающее значение для выявления контаминации на этапах отбора проб и выделения ДНК, поскольку в противном случае это может привести к ошибочному фоновому выявлению микроорганизмов. Кроме того, следует включать положительные контроли (например, имитирующее сообщество бактерий, обработанное как отдельный образец) или внутренние стандарты (целые клетки или экзогенная ДНК, добавленные в матрицу образца) для выявления потенциальных смещений в концентрации образца, выделении ДНК и биоинформатическом анализе [6]. Имитационные референсные материалы сообщества представляют собой приготовленные в лаборатории смеси определенных

компонентов (обычно ДНК из индивидуально культивируемых бактерий; иногда смеси целых клеток) в указанных количествах. Таким образом, эти материалы полезны в качестве «истинных» для рабочих процессов анализа, позволяя количественно оценивать аналитические характеристики (например, точность, смещение, воспроизводимость и т. д.) [9]. Данные референсные материалы должны быть подробно изучены и задепонированы в государственные репозитории, с целью сохранения их свойств и стандартизации процедур с учетом межлабораторной воспроизводимости [10]. Получение материалов с целью дальнейшего использования в качестве референсных в микробиологических и молекулярно-генетических исследованиях осуществляется в рамках научно-исследовательских работ и как правило является результатом интеллектуальной собственности, оформленным в виде патента на изобретение.

После выделения ДНК процесс подготовки библиотеки преобразует выделенные образцы ДНК в маркированные библиотеки, готовые к секвенированию на выбранной платформе. Создание библиотеки коротких прочтений обычно включает четыре основных этапа: фрагментацию ДНК на более мелкие, случайные, перекрывающиеся фрагменты с использованием ультразвуковых или ферментативных методов; репарацию концов фрагментов; лигирование парных или одинарных адаптеров; и индексацию, которая позволяет объединять несколько образцов в одном цикле секвенирования, снижая материальные затраты на образец и увеличивая производительность за цикл. Индексированные библиотеки нормализуются и объединяются в эквимоллярные концентрации перед секвенированием. Смещение стандартной процедуры на данном этапе зависит использования/неиспользования ПЦР [11].

С ростом разнообразия биоинформатических инструментов и параметров, определяющих использование каждого из них, может быть сложно определить, какой инструмент даст наиболее точное

представление о составе образца. В условиях этой растущей проблемы использование контрольных материалов помогает исследователям в валидации и выборе процесса разработки. Эти материалы также могут быть использованы для оценки текущих и будущих версий программного обеспечения. Однако выбор наиболее подходящего контрольного материала сам по себе является на данный момент нерешенной задачей [12].

ВЫВОДЫ / CONCLUSION

Существует множество технических этапов, необходимых для проведения экспериментов по секвенированию ампликона гена 16S рРНК, которые могут повлиять на результаты изучения бактериальных сообществ. К ним относятся отбор проб (место отбора проб, метод, транспортировка и хранение образца), извлечение материала нуклеиновой кислоты, выбор праймера 16S рРНК, амплификация, подготовка библиотеки, секвенирование и конвейер биоинформатического анализа. Проблемы, связанные с таксономическим и функциональным профилированием на основе метагеномных данных, включают в себя вопросы точности классификации в условиях систематических ошибок в базах данных, сложности прогнозирования функций и метаболических взаимодействий в различных микробных сообществах, а также отсутствие единого методологического подхода к проведению исследования. Преодоление этих препятствий потребует междисциплинарного сотрудничества, развития вычислительных инструментов и совершенствования аналитических методов.

INFLUENCE OF METHODOLOGICAL DECISIONS ON THE RESULTS OF METAGENOMIC SEQUENCING OF BACTERIAL COMMUNITIES

Prasolova O.V. – Candidate of Sciences (Veterinary medicine), Leading researcher of the Molecular Biology Department (orcid.org/0000-0001-8924-2273)

The Russian State Center for Animal Feed and Drug Standardization and Quality

*o.prasolova@vgnki.ru

ABSTRACT

The term microbiome encompasses the totality of genes belonging to microorganisms (bacteria, viruses, fungi) that inhabit a particular environment. Sequencing a region of the 16S rRNA gene allows us to study the taxonomic composition of the microbiota, identifying bacteria and archaea, including those that are difficult to cultivate. However, this method still does not have a standardized procedure. The many available methodological options complicate the reproducibility of results and, as a rule, limit the comparability of results between independent studies using different methodologies, measurement and bioinformatics pipelines. Interpretation, discussion and visualization of animal microbiome research results are a complex task due to the lack of uniform standard parameters and reference data for collection and comparison. Standardized methodologies in the study of microorganism communities are needed to simplify and compare the results of different studies. This analysis presents the main reasons for discrepancies in the data obtained at all stages of the work. These include sampling (sampling location, method, sample transport and storage), nucleic acid extraction, 16S rRNA primer selection, amplification, library preparation, sequencing and bioinformatics pipeline. Challenges associated with metagenomic data-based taxonomic and functional profiling include issues of classification accuracy in the face of systematic database errors, difficulties in predicting functions and metabolic interactions in different microbial communities, and the lack of a unified methodological approach to conducting the study. Overcoming these obstacles will require interdisciplinary collaboration, development of computational tools and improvement of analytical methods.

СПИСОК ИСТОЧНИКОВ

1. The effect of 16S rRNA region choice on bacterial community metabarcoding results. / Y. Bukin, Y. Galachyants, I. Morozov [et al.] // Sci

Data 6, 190007 (2019). <https://doi.org/10.1038/sdata.2019.7>.
2. Characterization of the gut microbiome using 16S or shotgun metagenomics / J. Jovel, J. Patterson, W. Wang // *Frontiers in microbiology*. – 2016. – №7 – C.459.
3. Towards standards for human fecal sample processing in metagenomic studies / P.I. Costea, G. Zeller, S. Sunagawa [et al.] // *Nature biotechnology*. – 2017. – T. 35. – №. 11. – C. 1069-1076.
4. Variation in the metagenomic analysis of fecal microbiome composition calls for a standardized operating approach / Z. Xu, Y.K. Yeoh, H.M. Tun [et al.] // *Microbiology Spectrum*. – 2024. – T. 12. – №. 12. – C. e01516-24.
5. Determination of effect sizes for power analysis for microbiome studies using large microbiome databases / G. Rahman, D. McDonald, A. Gonzalez [et al.] // *Genes*. – 2023. – T. 14. – №. 6. – C. 1239.
6. Analysis of metagenomic data / S. Liu, J.S. Rodriguez, V. Munteanu [et al.] // *Nature Reviews Methods Primers*. – 2025. – T. 5. – №. 1. – C. 5. <https://doi.org/10.1038/s43586-024-00376-6>
7. Comparison of methods to collect fecal samples for microbiome studies using whole-genome shotgun metagenomic sequencing / D. A. Byrd, R. Sinha, K. L. Hoffman [et al.] // *Mosphere*. – 2020. – T. 5. – №. 1. – C. 10.1128/msphere. 00827-19, <https://doi.org/10.1128/msphere.00827-19>.
8. Reproducibility, stability, and accuracy of microbial profiles by fecal sample collection method in three distinct populations / D. A. Byrd, J. Chen, E. Vogtmann [et al.] // *PLoS One*. – 2019. – T. 14. – №. 11. – C. e0224757.
9. Variability and bias in microbiome metagenomic sequencing: an interlaboratory study comparing experimental protocols / S. P. Forry, S. L. Servetas, J. G. Kralj [et al.] // *Scientific reports*. – 2024. – T. 14. – №. 1. – C. 9785. <https://doi.org/10.1038/s41598-024-57981-4>.
10. Прасолова, О. В. Депонирование патогенных штаммов микроорганизмов как основа биобезопасности / О.В. Прасолова // *Нормативно-правовое регулирование*

в ветеринарии. — 2024. — № 3. — С. 39–43. — DOI 10.52419/issn2782-6252.2024.3.39.

11. Library construction for next-generation sequencing: overviews and challenges / S. R. Head, H. K Komori, S. A. LaMere [et al.] // *Biotechniques*. — 2014. — Т. 56. — №. 2. — С. 61-77.

12. Benchmarking short-read metagenomics tools for removing host contamination / Y. Gao, H. Luo, H. Lyu [et al.] // *GigaScience*. — 2025. — Т. 14. — С. giaf004.

REFERENCES

1. The effect of 16S rRNA region choice on bacterial community metabarcoding results. / Y. Bukin,

Y. Galachyants, I. Morozov [et al.] // *Sci Data* 6, 190007 (2019). <https://doi.org/10.1038/sdata.2019.7>.

2. Characterization of the gut microbiome using 16S or shotgun metagenomics / J. Jovel, J. Patterson, W. Wang // *Frontiers in microbiology*. — 2016. — №7— С.459.

3. Towards standards for human fecal sample processing in metagenomic studies / P.I. Costea, G. Zeller, S. Sunagawa [et al.] // *Nature biotechnology*. — 2017. — Т. 35. — №. 11. — С. 1069-1076.

4. Variation in the metagenomic analysis of fecal microbiome composition calls for a standardized operating approach / Z. Xu, Y.K. Yeoh, H.M. Tun [et al.] // *Microbiology Spectrum*. — 2024. — Т. 12. — №. 12. — С. e01516-24.

5. Determination of effect sizes for power analysis for microbiome studies using large microbiome databases / G. Rahman, D. McDonald, A. Gonzalez [et al.] // *Genes*. — 2023. — Т. 14. — №. 6. — С. 1239.

6. Analysis of metagenomic data / S. Liu, J.S. Rodriguez, V. Munteanu [et al.] // *Nature Reviews Methods Primers*. — 2025. — Т. 5. — №. 1. — С. 5. <https://doi.org/10.1038/s43586-024-00376-6>

7. Comparison of methods to collect fecal samples for microbiome studies using whole-genome shotgun metagenomic sequencing / D. A. Byrd, R. Sinha, K. L. Hoffman [et al.] // *Msphere*. — 2020. — Т. 5. — №. 1. — С. 10.1128/msphere.00827-19, <https://doi.org/10.1128/msphere.00827-19>.

8. Reproducibility, stability, and accuracy of microbial profiles by fecal sample collection method in three distinct populations / D. A. Byrd, J. Chen, E. Vogtmann [et al.] // *PLoS One*. — 2019. — Т. 14. — №. 11. — С. e0224757.

9. Variability and bias in microbiome metagenomic sequencing: an interlaboratory study comparing experimental protocols / S. P. Forry, S. L. Servetas, J. G. Kralj [et al.] // *Scientific reports*. — 2024. — Т. 14. — №. 1. — С. 9785. <https://doi.org/10.1038/s41598-024-57981-4>.

10. Prasolova, O. V., Deposit of pathogenic microbial strains as a basis for biosafety / O. V. Prasolova // *Legal regulation in veterinary medicine*. — 2024. — № 3. — P. 39–43. — DOI 10.52419/issn2782-6252.2024.3.39.

11. Library construction for next-generation sequencing: overviews and challenges / S. R. Head, H. K Komori, S. A. LaMere [et al.] // *Biotechniques*. — 2014. — Т. 56. — №. 2. — С. 61-77.

12. Benchmarking short-read metagenomics tools for removing host contamination / Y. Gao, H. Luo, H. Lyu [et al.] // *GigaScience*. — 2025. — Т. 14. — С. giaf004.