

METHOD FOR QUANTITATIVE ASSESMENT OF GUT MICROBIOTA: A COMPARATIVE ANALYSIS OF 16S NGS AND qPCR

Zlobovskaya OA¹ ✉, Kurnosov AS¹, Sheptulina AF², Glazunova EV¹

¹ Centre for Strategic Planning and Management of Biomedical Health Risks of the Federal Medical Biological Agency, Moscow, Russia

² National Medical Research Center for Therapy and Preventive Medicine of the Ministry of Healthcare of the Russian Federation, Moscow, Russia

Recently, considerable commercialization of services for quantification of gut microbiota aimed to diagnose dysbiosis, the microbial imbalance, is observed. In the context of growing interest to the personalized approaches in medicine and preventive therapy, the diagnosis of dysbiosis is becoming increasingly important. The results of such screening are used to adjust guidelines on correction of the diet, lifestyle modification, or, where necessary, drug therapy prescription. Such assessment requires a reliable and accurate method for evaluation of microbiota, since validity of further recommendations and therapeutic interventions depends on the quality of the data obtained. The paper reports the main aspects of the two approaches used for microbiota quantification: 16S rRNA next-generation sequencing (16S NGS) and real-time PCR (qPCR). The strengths (from our perspective) and weaknesses of the approaches are also provided.

Keywords: 16S NGS, real-time PCR, metagenomic analysis, quantitative analysis, microbiome, personalized medicine

Author contribution: Zlobovskaya OA — concept, literature review, manuscript writing; Kurnosov AS, Sheptulina AF, Glazunova EV — manuscript editing.

✉ **Correspondence should be addressed:** Olga A. Zlobovskaya
Pogodinskaya, 10, str. 1, Moscow, 119121, Russia; OZlobovskaya@cspsfmba.ru

Received: 15.10.2024 **Accepted:** 25.10.2024 **Published online:** 30.10.2024

DOI: 10.24075/brsmu.2024.047

ВЫБОР МЕТОДА КОЛИЧЕСТВЕННОЙ ОЦЕНКИ МИКРОБИОТЫ КИШЕЧНИКА: СРАВНИТЕЛЬНЫЙ АНАЛИЗ 16S NGS И ПЦР-РВ

О. А. Злобовская¹ ✉, А. С. Курносов¹, А. Ф. Шептулина², Е. В. Глазунова¹

¹ Центр стратегического планирования и управления медико-биологическими рисками здоровью Федерального медико-биологического агентства, Москва, Россия

² Национальный медицинский исследовательский центр терапии и профилактической медицины Министерства здравоохранения Российской Федерации, Москва, Россия

В последнее время наблюдается значительная коммерциализация услуг по количественной оценке микробиоты кишечника с целью диагностики дисбиоза — нарушения микробного баланса. В условиях растущего интереса к персонализированным подходам в медицине и профилактической терапии диагностика дисбиоза приобретает все большее значение. Результаты подобного скрининга используют для рекомендаций по коррективке питания, изменению образа жизни или, при необходимости, назначения медикаментозного лечения. Для подобной оценки необходим надежный и точный метод оценки микробиоты, поскольку от качества полученных данных зависит корректность последующих рекомендаций и терапевтических вмешательств. В статье рассмотрены основные аспекты двух подходов, применяемых для количественной оценки микробиоты, — высокопроизводительного секвенирования гена 16S рНК (16S NGS) и ПЦР в реальном времени (ПЦР-РВ), а также представлены их сильные, на наш взгляд, и слабые стороны.

Ключевые слова: 16S NGS, ПЦР-РВ, метагеномный анализ, количественный анализ, микробиом, персонализированная медицина

Вклад авторов: О. А. Злобовская — идея, анализ литературы, написание рукописи; А. С. Курносов, А. Ф. Шептулина, Е. В. Глазунова — редактирование рукописи.

✉ **Для корреспонденции:** Ольга Анатольевна Злобовская
ул. Погодинская, д. 10, с. 1, г. Москва, 119121, Россия; OZlobovskaya@cspsfmba.ru

Статья получена: 15.10.2024 **Статья принята к печати:** 25.10.2024 **Опубликована онлайн:** 30.10.2024

DOI: 10.24075/vrgmu.2024.047

16S NGS: Broad Capabilities and Significant Limitations

The 16S NGS method has become an essential tool for studying the microbiota. Its main advantage is the ability to simultaneously sequence multiple samples and detect a wide range of microorganisms. Comprehensive assessment of the taxonomic composition of microbial communities and their diversity makes 16S NGS indispensable for fundamental research. However, despite its benefits, this method has several significant limitations that can lead to distorted quantitative results.

Uneven amplification (dependence on primers)

Universal primers are used to amplify the variable regions of the 16S rRNA gene. They exhibit different affinities for the DNA of various taxa, resulting in unequal amplification efficiency during library preparation [1]. As a result, the microbiota structure data can be skewed, with some taxa being overestimated, while others are underestimated or entirely missed.

Uneven amplification (dependence on taxonomic composition)

The most abundant taxa gain a significant advantage during the early stages of amplification [2], thus reducing the likelihood of accurately detecting rare taxa (up to 10% of the total community). Since each sample has a unique microbiota composition, it is impossible to apply a systematic correction for all samples, even when using the same protocols [3, 4].

Low sensitivity

On average, between 5,000 and 50,000 reads are obtained per sample when using the 16S NGS method. However, according to Poisson distribution, quantitative assessment of a taxon can only be considered statistically reliable when there are at least 100 reads for that taxon in the sample [5]. This limits the ability to reliably quantify taxa that make up less than 0.2–2% of the total reads (depending on the total number of reads). Increasing the number of reads per sample is not

always effective, as the amplification of dominant taxa occurs in the early stages, leading to a significant underrepresentation or even loss of minor taxa. Consequently, the taxonomic diversity saturation curve reaches a plateau at 20,000–50,000 reads, meaning that further increasing the number of reads will not improve data representativeness. This is especially important for minor opportunistic microorganisms that may have clinical significance at low concentrations but are often either undetected or inaccurately quantified. Additionally, there is no consensus among researchers on whether it is more accurate to compare samples with different numbers of reads or to introduce bias by unifying the number of reads [6, 7].

Reduced specificity

When analyzing short regions (V1-V3, V3-V4, V6, etc.), the high degree of conservation in the 16S region often prevents taxonomic resolution at the species level, and sometimes even at the genus level [1, 8]. Using the full-length 16S gene increases the resolution of sequencing but is only available on such platforms as ONT, PacBio, and LoopSeq. A significant drawback of these platforms is their higher error rate compared to short-read platforms like Illumina.

Limitations of relative quantification of taxa

The 16S NGS method evaluates only the relative abundance of taxa, not their absolute quantity. This means that an increase in the relative abundance of one taxon, for example, due to dietary changes, will automatically reduce the proportion of other taxa in the analysis. Simultaneous changes in multiple taxa in either direction makes the reconstruction of the true dynamics of the community impossible [4–6].

Impact of 16S rRNA gene copy number

Each microbial species has a unique number of 16S rRNA gene copies, which is rarely considered during analysis, particularly when identifying sequences to the genus or family level. Even when using specialized plugins for QIIME 2, biases usually persist. One reason is that in cases where the copy number data for a specific taxonomic group is absent from the rrnDB database, the algorithm automatically assigns a copy number of one.

Uneven phylogenetic resolution

Different regions of the 16S rRNA gene have varying levels of phylogenetic resolution [1, 8–10]. This leads to inconsistent classification accuracy, complicating the comparison of data across different studies.

Differences in sequencing platforms and data processing methods

The choice of sequencing platforms and library preparation methods can lead to significant variations in results [1, 11, 12]. As mentioned above, this makes it more challenging to compare data across various studies.

Dependence on databases

Different databases (RDP, SILVA, Greengenes, etc.) can yield different quantitative assessments for the same sample [1, 13]. Additionally, databases are updated every few years, which means that newly introduced taxa may be missing.

qPCR: Specialized Tasks, High Accuracy

Unlike NGS, specific DNA fragments are amplified in real-time PCR (qPCR). This results in several advantages.

High sensitivity and a broad quantitative range

qPCR enables the detection and quantification of even a few target copies in a reaction with high precision. This is especially important when studying rare clinically significant taxa, which may be missed by 16S NGS. Additionally, qPCR can reliably quantify up to 10^7 – 10^8 target copies in a reaction.

High specificity

Oligonucleotides are designed to distinguish even closely related microorganisms with high accuracy.

Improved Precision

Unlike 16S NGS, the absence of simultaneous amplification of hundreds of different targets leads to a more reliable individual assessment of a specific taxon abundance.

Fast and simple interpretation

Unlike 16S NGS, qPCR does not require complex bioinformatics methods for data interpretation. This makes it more accessible and convenient for clinical research and diagnostics, where speed and accuracy are critical.

High reproducibility

qPCR provides higher reproducibility compared to 16S NGS due to the simplicity of the method and data analysis. This is particularly important for clinical diagnostics and long-term studies, and also facilitates data comparison between different studies and laboratories.

Absolute quantification

qPCR allows for both relative and absolute quantification of taxa. Thus qPCR enables analysis of microbiota dynamics under different conditions, unlike the relative approach of NGS.

Reduced dependency on sample quality

qPCR analysis is less dependent on the initial quality of the sample (e.g., quantity, presence of PCR inhibitors) compared to the 16S NGS method, where these factors significantly impact the library preparation stage.

Nevertheless, the qPCR method also has certain limitations. However, unlike NGS, many of potential issues can be minimized if addressed properly.

Selection of target microorganisms

Preselected genetic targets are amplified in qPCR, which requires prior knowledge of the microbiota key representatives in the given study.

Target region selection

The most commonly studied region for the majority of bacteria is the 16S rRNA gene, making it the typical target for qPCR

assay development. However, this is a highly conserved genomic region, so for some taxonomic units at the species level (and occasionally at the genus level, e.g., *Oscillibacter/Dysosmobacter*), it may not be possible to develop specific systems that amplify 16S region. For some microorganisms, whole-genome data are available, allowing the selection of another region for detection. However, these organisms are in the minority, so the chosen target may be nonspecific, or the system may fail to amplify all members of the given taxonomic group.

Limitation on the number of taxa

High qPCR specificity limits the number of taxa that can be analyzed simultaneously. For accurate quantitative assessment, it is recommended to combine no more than two targets (if they exhibit a broad range and are consistently present in most samples) or three targets (for rare taxa) in a single tube. Moreover, due to the limited number of taxa analyzed in this method, qPCR does not provide information on the structure of the entire microbial community or its diversity, which may also hold clinical significance.

Biases related to gene copy number

This issue can arise if the system is designed to detect a taxonomic group at a higher level (e.g., family), where different

genera/species within the group possess significantly varying numbers of 16S rRNA gene copies.

Need for data standardization

Converting the data obtained through qPCR into absolute values requires the use of calibration standards. For maximum accuracy, it is essential to pre-assess the standards using droplet digital PCR. In addition, the sensitivity and linear range of oligonucleotide systems should preferably be tested not on model samples (e.g., plasmid or amplicon titration) but on the genomic DNA of the corresponding taxon, ideally against a background of fecal DNA in clinically relevant quantities.

CONCLUSION

A comparison of the 16S NGS and qPCR methods shows that NGS is better suited for studying the overall composition and diversity of the microbiota. However, its use for quantitative assessment is limited by several factors that currently lack practical solutions. Meanwhile, qPCR offers more accurate and reliable quantitative assessment, making it the preferred method for studies where high precision is required, and the target markers are well-defined.

References

- Abellan-Schneyder I, Matchado MS, Reitmeier S, Sommer A, Sewald Z, Baumbach J, et al. Primer, Pipelines, Parameters: Issues in 16S rRNA Gene Sequencing. Tringe SG, editor. mSphere. 2021; 6 (1): e01202–20.
- Gonzalez JM, Portillo MC, Belda-Ferre P, Mira A. Amplification by PCR Artificially Reduces the Proportion of the Rare Biosphere in Microbial Communities. Gilbert JA, editor. PLoS ONE. 2012; 7 (1): e29973.
- Boers SA, Jansen R, Hays JP. Understanding and overcoming the pitfalls and biases of next-generation sequencing (NGS) methods for use in the routine clinical microbiological diagnostic laboratory. Eur J Clin Microbiol Infect Dis. 2019; 38 (6): 1059–70.
- Vaginal Microbiome Consortium (additional members), Brooks JP, Edwards DJ, Harwich MD, Rivera MC, Fettweis JM, et al. The truth about metagenomics: quantifying and counteracting bias in 16S rRNA studies. BMC Microbiol. 2015; 15 (1): 66.
- Barlow JT, Bogatyrev SR, Ismagilov RF. A quantitative sequencing framework for absolute abundance measurements of mucosal and luminal microbial communities. Nat Commun. 2020; 11 (1): 2590.
- Vandeputte D, Kathagen G, D'hoë K, Vieira-Silva S, Valles-Colomer M, Sabino J, et al. Quantitative microbiome profiling links gut community variation to microbial load. Nature. 2017; 551 (7681): 507–11.
- Nearing JT, Douglas GM, Hayes MG, MacDonald J, Desai DK, Allward N, et al. Microbiome differential abundance methods produce different results across 38 datasets. Nat Commun. 2022; 13 (1): 342.
- Johnson JS, Spakowicz DJ, Hong BY, Petersen LM, Demkowicz P, Chen L, et al. Evaluation of 16S rRNA gene sequencing for species and strain-level microbiome analysis. Nat Commun. 2019; 10 (1): 5029.
- Rintala A, Pietilä S, Munukka E, Eerola E, Pursiheimo JP, Laiho A, et al. Gut Microbiota Analysis Results Are Highly Dependent on the 16S rRNA Gene Target Region, Whereas the Impact of DNA Extraction Is Minor. J Biomol Tech JBT. 2017; 28 (1): 19–30.
- Kameoka S, Motooka D, Watanabe S, Kubo R, Jung N, Midorikawa Y, et al. Benchmark of 16S rRNA gene amplicon sequencing using Japanese gut microbiome data from the V1–V2 and V3–V4 primer sets. BMC Genomics. 2021; 22 (1): 527.
- Tremblay J, Singh K, Fern A, Kirton ES, He S, Woyke T, et al. Primer and platform effects on 16S rRNA tag sequencing. Front Microbiol. 2015; 6.
- Chuang HH, Huang CG, Chou SH, Li HY, Lee CC, Lee LA. Comparative analysis of gut microbiota in children with obstructive sleep apnea: assessing the efficacy of 16S rRNA gene sequencing in metabolic function prediction based on weight status. Front Endocrinol. 2024; 15: 1344152.
- Ceccarani C, Severgnini M. A comparison between Greengenes, SILVA, RDP, and NCBI reference databases in four published microbiota datasets. 2023 [cited 2024 Oct 4]. Available from: <http://biorxiv.org/lookup/doi/10.1101/2023.04.12.535864>.

Литература

- Abellan-Schneyder I, Matchado MS, Reitmeier S, Sommer A, Sewald Z, Baumbach J, et al. Primer, Pipelines, Parameters: Issues in 16S rRNA Gene Sequencing. Tringe SG, editor. mSphere. 2021; 6 (1): e01202–20.
- Gonzalez JM, Portillo MC, Belda-Ferre P, Mira A. Amplification by PCR Artificially Reduces the Proportion of the Rare Biosphere in Microbial Communities. Gilbert JA, editor. PLoS ONE. 2012; 7 (1): e29973.
- Boers SA, Jansen R, Hays JP. Understanding and overcoming the pitfalls and biases of next-generation sequencing (NGS) methods for use in the routine clinical microbiological diagnostic laboratory. Eur J Clin Microbiol Infect Dis. 2019; 38 (6): 1059–70.
- Vaginal Microbiome Consortium (additional members), Brooks JP, Edwards DJ, Harwich MD, Rivera MC, Fettweis JM, et al. The truth about metagenomics: quantifying and counteracting bias in 16S rRNA studies. BMC Microbiol. 2015; 15 (1): 66.

5. Barlow JT, Bogatyrev SR, Ismagilov RF. A quantitative sequencing framework for absolute abundance measurements of mucosal and lumenal microbial communities. *Nat Commun.* 2020; 11 (1): 2590.
6. Vandeputte D, Kathagen G, D'hoë K, Vieira-Silva S, Valles-Colomer M, Sabino J, et al. Quantitative microbiome profiling links gut community variation to microbial load. *Nature.* 2017; 551 (7681): 507–11.
7. Nearing JT, Douglas GM, Hayes MG, MacDonald J, Desai DK, Allward N, et al. Microbiome differential abundance methods produce different results across 38 datasets. *Nat Commun.* 2022; 13 (1): 342.
8. Johnson JS, Spakowicz DJ, Hong BY, Petersen LM, Demkowicz P, Chen L, et al. Evaluation of 16S rRNA gene sequencing for species and strain-level microbiome analysis. *Nat Commun.* 2019; 10 (1): 5029.
9. Rintala A, Pietilä S, Munukka E, Eerola E, Pursiheimo JP, Laiho A, et al. Gut Microbiota Analysis Results Are Highly Dependent on the 16S rRNA Gene Target Region, Whereas the Impact of DNA Extraction Is Minor. *J Biomol Tech JBT.* 2017; 28 (1): 19–30.
10. Kameoka S, Motooka D, Watanabe S, Kubo R, Jung N, Midorikawa Y, et al. Benchmark of 16S rRNA gene amplicon sequencing using Japanese gut microbiome data from the V1–V2 and V3–V4 primer sets. *BMC Genomics.* 2021; 22 (1): 527.
11. Tremblay J, Singh K, Fern A, Kirton ES, He S, Woyke T, et al. Primer and platform effects on 16S rRNA tag sequencing. *Front Microbiol.* 2015; 6.
12. Chuang HH, Huang CG, Chou SH, Li HY, Lee CC, Lee LA. Comparative analysis of gut microbiota in children with obstructive sleep apnea: assessing the efficacy of 16S rRNA gene sequencing in metabolic function prediction based on weight status. *Front Endocrinol.* 2024; 15: 1344152.
13. Ceccarani C, Severgnini M. A comparison between Greengenes, SILVA, RDP, and NCBI reference databases in four published microbiota datasets. 2023 [cited 2024 Oct 4]. Available from: <http://biorxiv.org/lookup/doi/10.1101/2023.04.12.535864>.