BIBLIOGRAPHIC INFORMATION SYSTEM

Journal Full Title: Journal of Biomedical Research & Environmental Sciences
Journal NLM Abbreviation: J Biomed Res Environ Sci
Journal Website Link: https://www.jelsciences.com
Journal ISSN: 2766-2276
Category: Multidisciplinary
Subject Areas: Medicine Group, Biology Group, General, Environmental Sciences
Topics Summation: 133
Issue Regularity: Monthly
Review Process: Double Blind
Time to Publication: 21 Days
Indexing catalog: IndexCopernicus ICV 2020: 53.77 | GoogleScholar | View more
Publication fee catalog: Visit here

DOI: 10.37871 (CrossRef)
Plagiarism detection software: iThenticate
Managing entity: USA
Language: English
Research work collecting capability: Worldwide
Organized by: SciRes Literature LLC

IndexCopernicus ICV 2020: 53.77

Manuscript should be submitted in Word Document (.doc or .docx) through

Online Submission

form or can be mailed to support@jelsciences.com

Vision: Journal of Biomedical Research & Environmental Sciences main aim is to enhance the importance of science and technology to the scientific community and also to provide an equal opportunity to seek and share ideas to all our researchers and scientists without any barriers to develop their career and helping in their development of discovering the world.

JOURNAL OF BIOMEDICAL RESEARCH & ENVIRONMENTAL SCIENCES
issn: 2766-2276
SUBJECT AREA(S): GENOMICS | GENETICS | MICROBIOLOGY
DOI: 10.37871

MINI REVIEW

Decoding Microbial Networks: An Insight into 16S rRNA and Whole Genome Sequencing Approaches in Metagenomic Studies

Cinzia Alfonsi, Fabio Di Pietro, Filomena Tiziana Papa and Federico Gabrielli*

Biolab srl, Molecular Genetics and Genomics Laboratory, Ascoli Piceno, Italy

Abstract

The exploration of microbial communities is pivotal in understanding environmental, human, and animal ecosystem dynamics. Advances in high-throughput sequencing technologies have significantly enriched our insights into microbiomes, with 16S rRNA and Whole Genome Sequencing (WGS) being central methodologies. This review delineates a comparative analysis of these sequencing techniques, particularly focusing on different hypervariable regions of 16S rRNA (V3-V4 and V2-3-4-6-7-8-9) and WGS. The 16S rRNA sequencing, despite being cost-effective and less computationally demanding, often limits identification to the genus level. In contrast, WGS, while being resource-intensive, provides a broader spectrum of microbial identification including bacteria, viruses, fungi, and parasites, alongside a deeper insight into microbial functional attributes. The challenges associated with sequencing depth in WGS are discussed, alongside emerging mitigating strategies like host DNA depletion. The choice between these methodologies hinges on the project objectives and available resources. This comparative assessment aims to guide researchers in selecting the apt sequencing approach for their metagenomic studies, thereby facilitating a more extensive understanding of microbial ecosystems.

Introduction

High-throughput sequencing technologies have dramatically enhanced our comprehension of microbial ecosystems. Central to these advancements are methodologies like 16S rRNA sequencing and Whole Genome Sequencing (WGS), which are employed to explore microbial diversity and metagenomic functionality. The 16S ribosomal RNA (rRNA) gene is a fundamental component of the bacterial and archaeal ribosome, crucial for protein synthesis. It contains nine hypervariable regions (V1-V9) interspersed with conserved sequences. The hypervariable regions exhibit substantial sequence diversity across different microbial species, making them ideal targets for identifying and differentiating microbial taxa. This review aims to compare the 16S rRNA sequencing techniques, focusing on the V3-V4 and V2-3-4-6-7-8-9 regions, and WGS, elucidating their advantages and limitations in metagenomic analyses [1].

16S rRNA sequencing

16S rRNA sequencing capitalizes on the sequence variability in the

***Corresponding author(s)**

Federico Gabrielli, Biolab srl, Molecular Genetics and Genomics Laboratory, Ascoli Piceno, Italy

Email: federico.gabrielli@laboratoriobiolab.it

DOI: 10.37871/jbres1816

Submitted: 09 October 2023

Accepted: 20 October 2023

Published: 21 October 2023

Keywords

- WGS
- NGS
- Microbiome
- 16S
- rRNA

BIOLOGY GROUP

GENOMICS | GENETICS | MICROBIOLOGY

VOLUME: 4 ISSUE: 10 - OCTOBER, 2023

How to cite this article: Alfonsi C, Di Pietro F, Papa FT, Gabrielli F. Decoding Microbial Networks: An Insight into 16S rRNA and Whole Genome Sequencing Approaches in Metagenomic Studies. J Biomed Res Environ Sci. 2023 Oct 21; 4(10): 1443-1446. doi: 10.37871/jbres1816, Article ID: JBRES1816, Available at: https://www.jelsciences.com/articles/jbres1816.pdf

hypervariable regions of the 16S rRNA gene to characterize microbial communities. The V3-V4 region is widely utilized for phylogenetic analyses and determination of microbial composition at the genus level. One of the primary advantages of targeting the V3-V4 region is the simplicity in bioinformatic analysis. The lesser degree of sequence complexity compared to whole genome data often results in lower computational power requirements, making it a more accessible choice for researchers with limited computational resources. Additionally, the availability of well-established protocols and extensive reference databases further facilitates the analysis [2].

However, a significant drawback of the V3-V4 region sequencing is its limitation in achieving species-level resolution. Often, the nucleotide variability within this region is insufficient to distinguish between closely related species, thereby providing identification mostly at the genus level. This limitation may hinder the comprehensive understanding of microbial community structures and functions, particularly in ecosystems where species-level identification is critical for interpreting ecological or clinical significance. Moreover, the primer design for this region may exhibit biases towards certain microbial groups, potentially skewing community composition insights [3]. The choice of primers for 16S rRNA sequencing is crucial as it significantly impacts the representation of microbial diversity. Different primers may exhibit varying affinities towards certain microbial taxa, thereby potentially introducing biases in community profiling [4-6]. Additionally, the PCR amplification step inherent in 16S rRNA sequencing may introduce biases such as preferential amplification of certain sequences and the generation of chimeric sequences, which can further skew the representation of microbial diversity [5]. These biases underline the importance of careful primer design and validation to ensure accurate microbial community profiling. The V2-3-4-6-7-8-9 regions provide a broader insight into microbial diversity by encompassing multiple hypervariable regions. A significant advantage of targeting these regions is the potential to achieve species-level identification, enhancing the resolution of microbial community profiling. While primer design can be complex due to the necessity of targeting multiple regions, commercial kits are available with pre-designed primers that ease this task, making the approach more accessible to researchers [7].

Furthermore, despite the increased complexity in covering multiple regions, the bioinformatic analysis remains relatively straightforward, retaining one of the advantages of 16S rRNA sequencing in requiring less computational power compared to WGS. This approach makes the V2-3-4-6-7-8-9 regions an attractive choice for researchers seeking a balance between resolution and computational demand [7].

Other advantages include potentially higher phylogenetic resolution and a broader representation of microbial diversity, as previously mentioned. However, the main disadvantages include increased complexity in primer design (mitigated by commercial kits), amplification, and data analysis. Moreover, the increased sequencing depth required to cover multiple regions may lead to higher costs. Additionally, the lack of well-established protocols and comprehensive reference databases compared to the V3-V4 region might pose challenges in data interpretation and validation [7,8].

Whole Genome Sequencing (WGS)

WGS offers a panoramic view of the microbial genome, facilitating not only species-specific identification but also functional analysis.

WGS offers much higher resolution compared to 16S rRNA sequencing, enabling species and even strain-level identification enabling a more accurate depiction of the microbial community structure. A distinct advantage of WGS is its ability to identify and characterize a wide range of microorganisms including bacteria, viruses, fungi, and parasites, making it a highly versatile tool for comprehensive microbial community analysis. It also empowers the detection of genetic elements like plasmids, phages, and antibiotic resistance genes, as well as virulence factors, providing critical insights into microbial pathogenicity and interactions within the community. However, disadvantages include the considerably higher cost, computational requirements, and data complexity. The requirement for high-quality DNA and the challenges in assembling and annotating microbial genomes, particularly in the presence of host DNA or in highly complex microbial communities, can also be limiting factors [9-11]. The assembly process in WGS can introduce various errors and biases. Misassemblies can occur due to the presence of repetitive sequences, leading to incorrect genome reconstructions. Coverage biases, often stemming from the DNA extraction or library

preparation stages, can also affect the accuracy of the assembly. These challenges necessitate the use of robust assembly algorithms and the availability of high-quality reference genomes for accurate genome reconstruction and annotation [12-14].

Unlike 16S rRNA sequencing, WGS allows for the analysis of metabolic pathways and prediction of microbial functionality. The advantages include the ability to infer microbial metabolic capabilities, interactions, and the potential to discover novel biochemical pathways. Moreover, WGS can provide insights into the functional potential and metabolic interactions within a microbial community. On the downside, the functional prediction is often dependent on the quality of available genomic databases and annotations. Also, the high cost, extensive computational resources, and expertise required for data analysis can be significant hurdles. A notable challenge associated with WGS is the requirement for a significantly higher sequencing depth compared to 16S rRNA sequencing to achieve adequate resolution. For instance, while 100k reads might be sufficient in 16S rRNA sequencing, WGS often needs a much deeper sequencing depth, ranging from 4 to 10 million reads, to achieve a comparable level of resolution. This demand for higher sequencing depth translates to increased costs and computational requirements, making WGS a more resource-intensive approach. However, advancements in sample preparation methods, such as host DNA depletion techniques, are beginning to mitigate this challenge. By effectively reducing the host DNA content in the samples, WGS techniques allow a more focused analysis of microbial DNA. This, in turn, allows for a reduction in the required sequencing depth, potentially lowering the costs and computational demands associated with WGS, and making it a more feasible choice for a broader range of research projects [15-17].

Conclusion

Both 16S rRNA sequencing and WGS are instrumental in exploring microbial ecosystems. While 16S rRNA sequencing is more cost-effective and simpler, WGS offers exceptional resolution and rich functional information, extending beyond bacterial communities to encompass a wide spectrum of microorganisms including viruses, fungi, and parasites. The choice between these approaches will depend on the specific objectives of the project and the resources available. The emerging methods to deplete host resent a promising avenue to overcome some of the challenges associated with WGS, particularly the requirement for higher sequencing depth. These advancements are likely to bridge the gap between these methodologies, enabling a more nuanced and comprehensive analysis of microbial communities across diverse ecosystems. Through continued innovation in sequencing technologies and bioinformatic tools, researchers are poised to delve deeper into the microbial world, unlocking novel insights that could significantly impact environmental, medical, and industrial fields.

References

1. Sanschagrin S, Yergeau E. Next-generation sequencing of 16S ribosomal RNA gene amplicons. J Vis Exp. 2014 Aug 29;(90):51709. doi: 10.3791/51709. PMID: 25226019; PMCID: PMC4828026.

2. Fadeev E, Cardozo-Mino MG, Rapp JZ, Bienhold C, Salter I, Salman-Carvalho V, Molari M, Tegetmeyer HE, Buttigieg PL, Boetius A. Comparison of Two 16S rRNA Primers (V3-V4 and V4-V5) for Studies of Arctic Microbial Communities. Front Microbiol. 2021 Feb 16;12:637526. doi: 10.3389/fmicb.2021.637526. PMID: 33664723; PMCID: PMC7920977.

3. Kameoka S, Motooka D, Watanabe S, Kubo R, Jung N, Midorikawa Y, Shinozaki NO, Sawai Y, Takeda AK, Nakamura S. Benchmark of 16S rRNA gene amplicon sequencing using Japanese gut microbiome data from the V1-V2 and V3-V4 primer sets. BMC Genomics. 2021 Jul 10;22(1):527. doi: 10.1186/s12864-021-07746-4. PMID: 34246242; PMCID: PMC8272389.

4. Tremblay J, Singh K, Fern A, Kirton ES, He S, Woyke T, Lee J, Chen F, Dangl JL, Tringe SG. Primer and platform effects on 16S rRNA tag sequencing. Front Microbiol. 2015 Aug 4;6:771. doi: 10.3389/fmicb.2015.00771. PMID: 26300854; PMCID: PMC4523815.

5. Nelson MC, Morrison HG, Benjamino J, Grim SL, Graf J. Analysis, optimization and verification of Illumina-generated 16S rRNA gene amplicon surveys. PLoS One. 2014 Apr 10;9(4):e94249. doi: 10.1371/journal.pone.0094249. PMID: 24722003; PMCID: PMC3983156.

6. Thijs S, Op De Beeck M, Beckers B, Truyens S, Stevens V, Van Hamme JD, Weyens N, Vangronsveld J. Comparative Evaluation of Four Bacteria-Specific Primer Pairs for 16S rRNA Gene Surveys. Front Microbiol. 2017 Mar 28;8:494. doi: 10.3389/fmicb.2017.00494. PMID: 28400755; PMCID: PMC5368227.

7. Treangen TJ, Salzberg SL. Repetitive DNA and next-generation sequencing: computational challenges and solutions. Nat Rev Genet. 2011 Nov 29;13(1):36-46. doi: 10.1038/nrg3117. Erratum in: Nat Rev Genet. 2012 Feb;13(2):146. PMID: 22124482; PMCID: PMC3324860.

8. Barb JJ, Oler AJ, Kim HS, Chalmers N, Wallen GR, Cashion A, Munson PJ, Ames NJ. Development of an Analysis Pipeline Characterizing Multiple Hypervariable Regions of 16S rRNA

Using Mock Samples. PLoS One. 2016 Feb 1;11(2):e0148047. doi: 10.1371/journal.pone.0148047. PMID: 26829716; PMCID: PMC4734828.

9. Malczynski M, Zhu A, Zembower T, Qi C. Diagnostic performance of Ion 16S metagenomics kit and Ion reporter metagenomics workflow for bacterial pathogen detection in culture-negative clinical specimens from sterile sources. Diagn Microbiol Infect Dis. 2021 Oct;101(2):115451. doi: 10.1016/j.diagmicrobio.2021.115451. Epub 2021 Jun 11. PMID: 34237647.

10. Wood DE, Lu J, Langmead B. Improved metagenomic analysis with Kraken 2. Genome Biol. 2019 Nov 28;20(1):257. doi: 10.1186/s13059-019-1891-0. PMID: 31779668; PMCID: PMC6883579.

11. Wensel CR, Pluznick JL, Salzberg SL, Sears CL. Next-generation sequencing: insights to advance clinical investigations of the microbiome. J Clin Invest. 2022 Apr 1;132(7):e154944. doi: 10.1172/JCI154944. PMID: 35362479; PMCID: PMC8970668.

12. Shi Y, Wang G, Lau HC, Yu J. Metagenomic Sequencing for Microbial DNA in Human Samples: Emerging Technological Advances. Int J Mol Sci. 2022 Feb 16;23(4):2181. doi: 10.3390/ijms23042181. PMID: 35216302; PMCID: PMC8877284.

13. Ross MG, Russ C, Costello M, Hollinger A, Lennon NJ, Hegarty R, Nusbaum C, Jaffe DB. Characterizing and measuring bias in sequence data. Genome Biol. 2013 May 29;14(5):R51. doi: 10.1186/gb-2013-14-5-r51. PMID: 23718773; PMCID: PMC4053816.

14. Robins HS, Campregher PV, Srivastava SK, Wacher A, Turtle CJ, Kahsai O, Riddell SR, Warren EH, Carlson CS. Comprehensive assessment of T-cell receptor beta-chain diversity in alphabeta T cells. Blood. 2009 Nov 5;114(19):4099-107. doi: 10.1182/blood-2009-04-217604. Epub 2009 Aug 25. PMID: 19706884; PMCID: PMC2774550.

15. Feldgarden M, Brover V, Haft DH, Prasad AB, Slotta DJ, Tolstoy I, Tyson GH, Zhao S, Hsu CH, McDermott PF, Tadesse DA, Morales C, Simmons M, Tillman G, Wasilenko J, Folster JP, Klimke W. Validating the AMRFinder Tool and Resistance Gene Database by Using Antimicrobial Resistance Genotype-Phenotype Correlations in a Collection of Isolates. Antimicrob Agents Chemother. 2019 Oct 22;63(11):e00483-19. doi: 10.1128/AAC.00483-19. Erratum in: Antimicrob Agents Chemother. 2020 Mar 24;64(4): PMID: 31427293; PMCID: PMC6811410.

16. Blin K, Shaw S, Augustijn HE, Reitz ZL, Biermann F, Alanjary M, Fetter A, Terlouw BR, Metcalf WW, Helfrich EJN, van Wezel GP, Medema MH, Weber T. antiSMASH 7.0: new and improved predictions for detection, regulation, chemical structures and visualisation. Nucleic Acids Res. 2023 Jul 5;51(W1):W46-W50. doi: 10.1093/nar/gkad344. PMID: 37140036; PMCID: PMC10320115.

17. Oyola SO, Gu Y, Manske M, Otto TD, O'Brien J, Alcock D, Macinnis B, Berriman M, Newbold CI, Kwiatkowski DP, Swerdlow HP, Quail MA. Efficient depletion of host DNA contamination in malaria clinical sequencing. J Clin Microbiol. 2013 Mar;51(3):745-51. doi: 10.1128/JCM.02507-12. Epub 2012 Dec 5. PMID: 23224084; PMCID: PMC3592063.

How to cite this article: Alfonsi C, Di Pietro F, Papa FT, Gabrielli F. Decoding Microbial Networks: An Insight into 16S rRNA and Whole Genome Sequencing Approaches in Metagenomic Studies. J Biomed Res Environ Sci. 2023 Oct 21; 4(10): 1443-1446. doi: 10.37871/jbres1816, Article ID: JBRES1816, Available at: https://www.jelsciences.com/articles/jbres1816.pdf