Biomedical Journal of Indonesia
Journal Homepage: https://bji-fk.ejournal.unsri.ac.id

ITS1 and ITS2 Region Use for Mycobiome Profiling In Deep Sequencing

Ridha Sukma Dewi[1], Conny Riana Tjampakasari[2*], Mardiastuti Wahid[2,3], Ibnu Agus Ariyanto[2], Yeva Rosana[2,3], Ika Ningsih[2]

[1]Master's of Biomedical Science Program, Faculty of Medicine, Universitas Indonesia, Jakarta, Indonesia
[2]Department of Microbiology, Faculty of Medicine, Universitas Indonesia, Jakarta, Indonesia
[3]Cipto Mangunkusumo General Hospital, Jakarta, Indonesia

ARTICLE INFO

Keywords:
Mycobiome
Fungi
ITS
Sequencing

Corresponding author:
Conny Riana Tjampakasari
E-mail address:
connyrianat@yahoo.com

All authors have reviewed and approved the final version of the manuscript.

https://doi.org/10.32539/BJI.v12i3.352

ABSTRACT

The human mycobiome is an important component of the microbiota and contributes to host health, metabolism, and immune responses. Accurate characterization of fungal communities requires appropriate molecular markers, with the internal transcribed spacer (ITS) region being the most widely used fungal barcode. However, the choice between ITS1 and ITS2 may influence mycobiome profiles due to differences in sequence variability, primer performance, and reference database coverage. This review aimed to critically compare ITS1 and ITS2 for human mycobiome profiling and evaluate their advantages, limitations, and methodological biases. Relevant literature was reviewed with emphasis on marker characteristics, primer selection, sequencing platforms, taxonomic resolution, reference databases, and bioinformatics approaches. Evidence indicates that ITS1 and ITS2 provide complementary information rather than one being universally superior. ITS1 can capture substantial fungal diversity but may be more affected by primer bias and sequence-length variation. ITS2 generally provides more consistent amplification and a broader reference database representation, although performance varies across fungal taxa, primer design, sequencing strategy, and analytical workflow. PCR amplification, sequencing platforms, databases, and bioinformatics pipelines can further influence observed profiles. Combining ITS1 and ITS2 may improve taxonomic coverage and reduce marker-specific detection bias. Overall, marker selection should be guided by study objectives, target fungal communities, primer characteristics, sequencing strategy, and bioinformatics workflow.

1. Introduction

The human mycobiome, comprising the fungal community residing in the human body, is an important component of the human microbiome and contributes to immune regulation, metabolism, inflammation, and host-microbial interactions. Fungi are present at various anatomical sites, including the oral cavity, gastrointestinal tract, respiratory tract, and skin. Although fungi constitute only a small proportion of the human microbiome, alterations in fungal community composition may disturb microbial homeostasis and contribute to disease processes. However, compared with the bacterial microbiome, the composition and functional role of the human mycobiome remain-less well understood.[1,2]

Characterizing the human mycobiome is challenging because many fungi cannot be cultivated using conventional culture-based methods. This limitation has led to the development of culture-independent molecular approaches, including sequencing-based methods, which provide broader characterization of fungal communities. Next-generation sequencing (NGS) has increased the feasibility and throughput of mycobiome profiling and has been applied using various sequencing platforms, including Illumina, Ion Torrent, Oxford Nanopore Technologies (ONT), and Pacific Biosciences (PacBio). However, the accuracy and comprehensiveness of fungal community characterization depend not only on the sequencing platform but also on the molecular target selected for fungal

identification.[3,4]

The Internal Transcribed Spacer (ITS) region of the fungal ribosomal RNA gene cluster is widely recognized as a primary DNA barcode for fungal identification. The ITS region consists of two variable subregions, ITS1 and ITS2, separated by the conserved 5.8S rRNA gene and flanked by the conserved 18S and 28S rRNA genes. The conserved regions facilitate primer design, whereas the variability of the ITS regions provides discriminatory power for fungal identification. Therefore, ITS1 and ITS2 have become commonly used molecular targets in NGS-based mycobiome studies. However, these two subregions differ in sequence variability, length characteristics, primer-binding regions, and taxonomic performance, which may influence the representation of fungal communities obtained from sequencing.[5–7]

ITS1 generally exhibits greater sequence variability than ITS2, which may provide high discriminatory power for differentiating fungal taxa. However, ITS1 is more susceptible to primer and amplification bias, sequence-length variation, GC-content variation, and intragenomic polymorphism. These characteristics may affect amplification efficiency and potentially lead to the underrepresentation or overestimation of specific fungal taxa. Conversely, ITS2 has been reported to provide more consistent amplification and broader support in reference database. Studies comparing the two regions have demonstrated differences in the number and types of fungal taxa detected; for example, one study reported of substantially more fungal genera detected using ITS2 than using ITS1. ITS2 has also been reported to improve the detection of certain fungal taxa, including *Malassezia* and *Yarrowia*, that may be missed with ITS1. [2,8–12]

These methodological differences are important because the choice of ITS subregion can influence estimates of fungal richness, diversity, and taxonomic composition. Both ITS1 and ITS2 are affected by primer bias and may differ in their ability to detect particular fungal groups. Consequently, studies using different ITS targets may produce different representations of the same fungal community, potentially affecting ecological interpretation and comparison between studies. Although ITS2 is frequently recommended for human mycobiome studies because of its consistency and taxonomic coverage, ITS1 may provide broader estimates of diversity and has also demonstrated advantages in certain applications. Thus, the relative performance of ITS1 and ITS2 cannot be considered in isolation from the fungal community, sample type, sequencing approach, and analytical objective.[2,5,12,13]

Despite the increasing use of ITS-based sequencing for human mycobiome profiling, differences between ITS1 and ITS2 remain an important methodological consideration. Existing studies have reported distinct advantages and limitations for each subregion, but these findings are not uniformly applicable across fungal communities and study settings. Therefore, a critical comparison is needed to clarify how the choice between ITS1 and ITS2 affect fungal detection, taxonomic resolution, diversity estimates, and the overall characterization of the human mycobiome.

This literature review aims to compare the use of the ITS1 and ITS2 regions for human mycobiome profiling using deep sequencing approaches. The review focuses on the characteristics, advantages, and limitations of each subregion, including taxonomic performance, primer and amplification bias, fungal diversity and richness estimation, detection of specific fungal taxa, and their application across different sequencing and bioinformatic approaches. By identifying the respective strengths and limitations of ITS1 and ITS2, this review aims to provide a clearer basis for selecting an appropriate ITS target based on the objectives and context of human mycobiome studies.

2. Mycobiome

The mycobiome is the collection of all fungal genomes in the human body, including unicellular yeast and filamentous fungi. Studies on human microbiomes have been increasing lately. Resident microbiomes play an important role in development of the immune system, metabolism, and inflammation, and are associated with various conditions, such as type 2 diabetes, autism spectrum disorder, depression, and inflammation of the intestine, heart, and respiration. However, most microbiome studies have focused on bacterial composition of the human microbiome, whereas the composition and functional role of fungi, viruses, and archaea remain poorly understood.[1,2]

Mycobiomes can be found in the healthy human body. As part of the human body's interconnected ecology, microorganisms establish both interdomain and intradomain interactions with other colonies. The biofilm that forms contain a bacterial and fungal cohabitatiation that is associated with many diseases.[1]

Fungi cause various human diseases, and, in an invasive form, can be associated with mortality, especially in high-risk groups such as immunocompromised individuals. The role of fungi in immunomodulation and their influence on inflammatory disease is of increasing interest nowadays. Similar to the study on the bacterial microbiome, the proportion of fungal dysbiosis in the microbiota can alter disease pathogenesis, with dynamic relationships among bacteria, viruses, and archaea.[2]

The mycobiome can be an important source of antigens to train the immune system and its response. Fungi activate fungus-specific pattern recognition receptors and other immune mechanisms that mediate defense against pathogens and tolerance to beneficial commensals. Compared with the bacteriome, interactions between the host and mycobiome remain poorly understood. Although fungi represent only a small proportion of the human microbiome, they play an important role in maintaining the microbial ecosystem. The interaction between fungi and other microorganisms may contribute to various biological processes. Therefore, understanding changes in the composition and its interaction with other microbes may provide important insights into the pathogenesis of disease at the body sites where these microorganisms colonize.[1]

3. ITS region as a molecular marker to identify the mycobiome

Ribosomal RNA (rRNA) is an important macromolecule present in living cells, including bacteria, archaea, and eukaryotes. Ribosomal RNA and proteins form ribosomes that function as mRNA translators into proteins. In 2012, the International Fungal Barcoding Consortium recommended using ITS from the ribosomal RNA gene cluster as a barcode to identify fungi. The total size of the fungal ITS region is approximately 600bp, consisting of ITS1 and ITS2 that divide the 5.8S rRNA. This ITS region lies between the 18s rRNA genome at the 5' end and the 28s rRNA gene at the 3' end, from the ITS2 spacer.[6,7]

The 18S, 5.8S, and 28S rRNA genes are highly conserved regions that enable the design of universal primers for amplification of ITS1, ITS2, or the entire ITS regions across a wide range of fungal taxa. Furthermore, ITS, as a DNA barcode, is a DNA region composed of multiple repeated copies of the ribosomal rRNA gene cluster, including the ITS region, within the haploid fungal genome. This high copy number facilitates PCR amplification from relatively small amounts of biological material.[7]

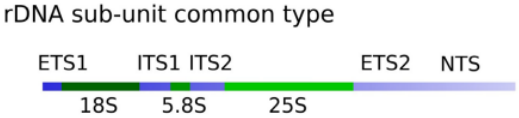


Figure 1. Genomic order of rDNA unit[6]

ITS1 characteristic

Internal transcribed spacer 1 (ITS1) has recently become an important molecular marker for fungal identification and diversity studies. ITS1 was part of the rDNA complex that became a vital component in the molecular identification of fungi. In general, ITS1 has more variability than ITS2 in various fungal taxa. This makes ITS1 a more powerful tool for differentiating species between complex communities.[8,9]

Appropriate primer selection based on the target taxa is essential for accurate metagenomic analysis. One of the most important considerations when using ITS1 as a marker is that amplification bias can alter the representation of fungal diversity. This bias could result in an unproportional representation fungal group in ecological analyses, which can alter the diversity analysis.[9,10]

The advantage of using ITS1 as a marker is its ability to be integrated into an automated sequencing workflow that could produce high-throughput fungal diversity. ITS1 has contributed to the development of computational pipelines for fungal community analysis, enhancing the accuracy of ecological analyses. The application of ITS1 as a marker has primarily focused on specific environments such as the gut microbiome. [10,14]

ITS1 also has several limitations, including variations in sequence length and GC content that may reduce PCR amplification and sequencing efficiency, as well as intragenomic polymorphism that can lead to an overestimation of fungal diversity. In some fungal groups, ITS1 alone may not provide sufficient taxonomic resolution. Therefore, combining it with other genetic markers or using the entire ITS region can improve identification accuracy.[5]

ITS2 characteristics

The use of the ITS2 region in analyzing gut mycobiome profiles shows that it can identify more genera than ITS1. The use of ITS2 allows the detection of up to 479 genera, compared to ITS1, which only detects 59 genera. These findings indicate a significant increase in the resolution of fungal community identification at the genus level when using ITS2. This finding suggests role for ITS2 in the study of fungal dysbiosis and its potential relationship to human disease.[11]

Using ITS2 allows the detection of more fungal taxa in human mycobiome samples, which can decrease the probability of missing important species. A universal ITS2 primer has been validated and supported by a broad reference database, facilitating automatic identification and comparison across studies. ITS2 exhibits minimal cross-amplification with human or bacterial DNA, thereby improving the specificity of fungal profiling. Studies of various human specimens, including fecal, oral, and respiratory samples, demonstrated that ITS2 enables highly sensitive and specific characterization of mycobiome diversity and composition and facilitates the detection of fungal species that are not identifiable using conventional culture-based methods or alternative molecular markers.[2]

Comprehensive comparison of ITS1 and ITS2

Comparison of ITS1 and ITS2 as molecular markers to identifying the fungal mycobiome highlights their advantages and disadvantages, with performance depending on the purpose and sample type. ITS2 has demonstrated advantages over ITS1 in several human mycobiome studies, particularly for detecting taxa such as Malassezia and Yarrowia. However, marker performance is context-dependent and influenced by primer design, fungal composition, specimen type, sequencing platform, and reference database. Therefore, ITS2 should not be considered universally superior to ITS1. ITS2 provides more accurate and consistent estimates of fungal richness and taxonomic composition across different environments. The advantages of ITS1 are higher estimated diversity and wide taxonomic coverage, with a taxonomic profile similar to the metagenomic shotgun

result, making it suitable for surveillance or when the source is limited.[2,12]

Table 1. Comparison of ITS1 and ITS2 mycobiome identification

Aspect	ITS1	ITS2	Source
Taxonomic accuracy	Higher, the bias on certain taxa	Higher, more consistent	2,5,15
Primer or amplification bias	Higher	Lower	5,15
Reference database	Enough, less than ITS2	More and wide	15
Richness estimation	Higher, overestimation of potency	More conservative, more specific	5
Detecting specific taxa	Less in *Malassezia and Yarrowia*	Better in a complex community	2,15
Recommendation	Good for surveillance	Recommended for the human mycobiome study	2,5,15

Both ITS1 and ITS2 can yield different results in detected taxa; both markers have primer bias and limitation in detecting certain fungal groups. Therefore, several studies recommend using both markers together to obtain more comprehensive results and reduce identification bias.[5,12] Compared with ITS1, ITS2 demonstrates lower length variability and more conserved primer-binding sites, thereby minimizing primer and amplification biases and improving fungal detection and taxonomic coverage. ITS1 is more susceptible to these biases, which may lead to the omission or underrepresentation of certain fungal taxa.[5,15] The comparison of ITS1 and ITS2 from various aspects can be seen in Table 1.

4. Deep sequencing for mycobiome profiling on the Illumina platform for the ITS amplicon

The Illumina sequencing method is a sequencing technique based on sequencing by synthesis using reversible dye terminators. The fluorescent dye can be released after the base-reading process is completed. Each nucleotide base (A, T, G, C) is dyed with a unique fluorescent dye so that it can be differentiated optically during the detection process. This technology enables the high-accuracy parallel sequencing of thousands of DNA fragments. The process includes DNA fragmentation, library preparation, library hybridization to the flow cell, and bridge amplification until a DNA cluster forms and is sequenced. Illumina sequencing also supports paired-end sequencing, which increases the data accuracy and reduces analysis cost.[15]

The Illumina platform for fungal ITS sequencing has gained considerable attention, particularly for mycobiome profiling. The Illumina MiSeq platform is widely used for ITS amplicon sequencing because of its high-throughput capacity and cost-effectiveness. Primer selection plays a critical role in achieving accurate genus-level identification and comprehensive characterization of fungal diversity. Some studies have shown that specific ITS primers outperform others in recovering diverse fungal taxa across multiple datasets. This finding highlights the importance of careful selection of methods to produce good data and comparable results consistently.[16]

Ion Torrent for mycobiome study

The Ion Torrent method is a sequencing technique that works by gradually adding nucleotide bases, so that the base sequence is complete as the step is added. After five ligations cycles, the information gathered to arrange the complete base sequence was sufficient. These methods can be applied to several analyses, such as small RNA sequencing, gene expression analysis, whole-genome sequencing, and region-targeted sequencing. Ion-Torrent utilizes a combination of metal conductivity and software based on a complementary metal-oxide-semiconductor (CMOS). When a nucleotide base is added, the sensitive layer of the metal oxide experiences a change in the surface potential, which is then converted into an electrical signal.[15]

DNA fragments are connected to

microbeads using special adaptor sequences, and then the PCR emulsion is performed, so that each bead contains one DNA fragment. DNA emulsion in beads is placed in a chip with millions of microwells, one bead in each well. This sequencing reaction occurs in a flow cell that is equipped with electrical sensors. When a nucleotide was added and bound to the DNA strand, a hydrogen ion (proton) was released, causing a pH change. Sensors sensitive to ions will detect pH changes and convert them into a voltage signal proportional to the number of nucleotides added. Every small change in electrical potential can represent the addition of a new base in the DNA synthesis process.[15]

The combination of Ion Torrent sequencing and optimized ITS-specific primers improves fungal community analysis resolution by enhancing fungal DNA amplification and enabling high-throughput sequencing for mycobiome profiling. Primer combinations can be used to simultaneously amplify bacterial and fungal genes in metagenomic studies. These methodological improvements demonstrate the value of Ion Torrent sequencing as a practical tool for microbial diagnostics, particularly in complex ecosystems where mycobiome interactions are important. [17,18]

Long-read sequencing applications

Long-read sequencing applications in human mycobiome profiling have shown significant potential in revealing the complexity of the microbiome that was previously difficult to access. This technology could enhance the accuracy of microbial amplicon sequencing using synthetic-based methods to obtain more accurate long-read data than common short-read sequencing approaches. For example, PacBio and Oxford Nanopore methods extend read lengths and enhance accuracy, thereby supporting applications in microbial taxon classification.[19]

Long-read sequencing overcomes the limitations of short-read sequencing by resolving repetitive and highly conserved genomic regions, thereby improving metagenomic assembly and microbial community characterization. Further exploration of long-read sequencing shows that this technology can be applied to microbiome-based diseases with higher species resolution, which is important for diagnosis and predicting therapy response.[20,21]

The effectiveness of long-read sequencing in clinical applications from combining approaches, such as shotgun metagenomic and transcriptomic approaches, can reveal the microbiome structure and function. The application of this technology to microbiome profiling demonstrates that an integrated metagenomic approach can provide deeper insights into microbiome dysbiosis associated with specific disease states.[22]

Optimal Platform Selection for Mycobiome Profiling

The selection of an optimal platform for mycobiome profiling depends on the study objectives, required taxonomic resolution, and available budget and resources. Long-read sequencing platforms, such as Pacific Biosciences (PacBio) and Oxford Nanopore Technologies (ONT), generally provide higher species-level resolutions than short-read platforms, such as Illumina. Comparative studies have shown that PacBio accurately reconstructs complex fungal communities, particularly for ITS regions with variable lengths, whereas the latest ONT platforms achieve >99% accuracy and produce comparable community profiling results. Despite these advantages, long-read platforms are generally more expensive and require more advanced bioinformatic analysis.[18,23,24]

Meanwhile, short-read platforms such as Illumina remain the preferred choice for large-scale studies since their low per-sample cost and well-established bioinformatics ecosystem, despite their limited species-level resolution and bias toward shorter ITS fragments. Ion Torrents are also suitable for mycobiome profiling but may exhibit platform-specific biases

depending on the fungal genome composition. In addition to platform selection, the choice of amplicon target (e.g., ITS1, ITS2, 18S, and 28S) and primers is critical, as variation in ITS length and copy number can introduce bias in species detection and abundance estimates. The use of curated fungal reference databases, such as UNITE and ntF-ITS1, together with a well-documented bioinformatics pipeline, is recommended to improve taxonomic identification accuracy.[13,18,23,24]

5. Mycobiome Profiling

Sample collection and DNA extraction are crucial steps in mycobiome profiling because they can alter the identification results and fungal community composition. Several studies show that DNA extraction methods, such as QIAmp, PowerSoil, DNeasy and manual protocols like the IHMS protocols Q, yield different efficiencies, quality, and mycobiome representations from feces, saliva, urine, and tissue samples. The addition of a bead-beating step or mechanical lysis could increase fungal DNA yield, especially for thick-walled fungi, and enrich the detection of filamentous species. For long-read sequencing, such as ONT, bead-beating methods can reduce DNA length and quality; therefore, chemical-based protocols like CTAB/SDS are more commonly recommended. Contamination from the reagent or kit can occur, especially in low-biomass samples, which is why a negative control is needed to differentiate the target DNA from the contaminant. [25,26,27,28]

PCR amplification is a critical step in mycobiome profiling because it can influence the representation of the microbial communit. DNA extraction methods, primer design, and PCR protocol can alter the taxonomic diversity and bias potency. The selection of the right primer is crucial because primer mismatch can prevent some taxa from being amplified. Some strategies that can be used to decrease PCR bias are PCR cycle optimization, statistical model for bias correction, and application of a two-step protocol with double barcoding to enhance accuracy and decrease wrong sequences.[29,30,31]

Library preparation can alter taxonomy and functional characterization in silico-constructed mocked community. However, the effects of library preparation are modest compared with the variability associated with DNA extraction methods, so the inter- and intra-sample variation. Furthermore, the selection of a sequencing platform has a substantial impact on the overall microbiome characterization. Both library preparation and sequencing platform selection introduce systematic biases in microbiome profiling. Variability attributable to differences in sequencing platforms is typically greater than that associated with library preparation. Biases related to GC content, amplification efficiency, and sequencing coverage can affect the detection and relative representation of specific fungal and microbial taxa.[33,33]

6. Bioinformatics Pipeline

The widespread application of metagenomic approaches remains limited by the substantial sequencing depth required to detect relatively rare fungal taxa. The specialized bioinformatics workflow and the limited development of fungal reference genome databases make metagenomic analysis relatively complex and resource-intensive. Early culture-independent studies of mycobiota targeted the eucaryotic 18S ribosomal RNA (rRNA) gene, analogous to the use of the procaryotic 16S rRNA gene, to enable PCR amplification of fungal DNA followed by taxonomic characterization through sequence analysis. However, the 18S rRNA gene provides relatively low taxonomic resolution for fungi and is often used to reliably distinguish taxa at lower taxonomic levels, particularly at the genus and species levels. The application of next-generation sequencing is still too costly due to the sequencing effort to capture the relatively rare fungal biosphere, and the special needs for bioinformatics analysis,

paired with underdeveloped fungal reference databases, making metagenomic approaches still challenging.[34]

Raw data processing is an essential first step in mycobiome analysis to ensure high-quality sequencing data. This process typically includes quality control (QC) to remove low-quality reads, adapter contamination, and poor-quality sequence ends, followed by demultiplexing and when applicable, paired-end read merging. Tools such as fastp and SOAPnuke efficiently perform QC, trimming, and deduplication, while integrated pipelines, including Natrix, PipeCraft, EasyAmplicon, and Dix-seq, automate downstream processing, including quality filtering, chimera removal, and OTU or ASV tables generation for subsequent analyses.[35,36]

ITS amplicon sequencing

The most commonly used high-throughput sequencing (HTS) methods for microbiome analysis are amplicon and metagenomic sequencing. Amplicon sequencing is the most widely used approach because it can be applied to nearly all sample types and employs marker genes, including 16S rRNA for prokaryotes and 18S rRNA or the ITS region for eukaryotes. Library preparation is typically performed using a two-step polymerase chain reaction PCR protocol, followed by sequencing on the Illumina PE250 platform, which generally generates 50,000–100,000 reads per sample. Amplicon sequencing is particularly effective for low-biomass samples and specimens contaminated with host DNA. However, its taxonomic resolution is generally limited to the genus level and is influenced by primer selection and the number of PCR amplification cycles.[37]

Amplicon analyses are commonly performed using USEARCH or QIIME, in which representative sequences are generated either by clustering reads into operational taxonomic units (OTUs) or by denoising them into amplicon sequence variants (ASVs) with algorithms such as DADA2, Deblur, or UNOISE3. ASVs provide higher taxonomic resolution than OTUs by distinguishing closely related species or strains. The resulting OTUs or ASVs are then quantified to generate feature tables and taxonomic annotations. Shotgun metagenomic analysis provides direct functional information and higher taxonomic resolution, often to the species or strain level, but requires substantially larger datasets, Linux-based software, and greater computational resources. Package managers, such as BioConda, and parallel computing tools like GNU Parallel, are commonly used to streamline software installation and improve analytical efficiency.[37]

Shotgun metagenomic sequencing

Taxonomic and mycobiome profiling depend on the availability of appropriate software and comprehensive reference databases. Shotgun metagenomic analysis remains challenging because fungal-specific tools and databases are still limited. The commonly used tools, such as Kraken2, MetaPhlAn4, EukDetect, FunOMIC, MiCoP, and HumanMycobiomeScan, have their own performance characteristics. Evaluations using mock communities have shown that no single tool consistently identifies all fungal species, although MetaPhlAn4 performs well at the genus level, FunOMIC identifies most species, and EukDetect, MiCoP, and FunOMIC generally provide the highest identification accuracy and relative abundance estimates. Furthermore, using reference databases restricted to fungi may increase the number of false-positive identifications because of conserved sequences and contamination. Therefore, reference databases covering all life domains are recommended for shotgun metagenomic profiling, particularly when analyzing microbial eukaryotes.[36]

Shotgun metagenomic analysis provides direct functional information and higher taxonomic resolution, often to the species or strain level, but requires substantially larger datasets, Linux-based software, and greater computational resources. Package

managers, such as BioConda, and parallel computing tools like GNU Parallel, are commonly used to streamline software installation and improve analytical efficiency.[37]

Next, clean reads are converted into taxonomic and functional profiles using either read- or assembly-based approaches. Read-based methods annotate individual reads by mapping them to reference databases, with MetaPhlAn2 using marker genes for taxonomic profiling, Kraken2 applying *k*-mer matching and LCA algorithm for classification, HUMAnN2 performing functional profiling and pathway analysis, and MEGAN providing graphical visualization of taxonomic and functional data. .[37]

Statistical and visualization analyses are then used to identify species, genes, and pathways that distinguish study groups and to evaluate microbial diversity and community composition. Alpha diversity, which measures species richness and evenness within samples, is commonly calculated using QIIME, vegan, or USEARCH and compared between groups using statistical tests such as analysis of variance, the Mann–Whitney U test, or the Kruskal–Wallis test.[37]

Diversity measures used in microbiota analysis consist of two community structures: richness and evenness. Richness represents the number of distinct taxa detected within a community, regardless of their relative frequencies. Evenness describes the equitability of the taxa frequencies in a community. Three commonly used alpha diversity measures include species observed, Shannon index and Simpson index.[38]

The Chao index is used to estimate the number of OTUs within a sample. It is widely applied in ecological studies to estimate species richness by accounting for rare taxa. A higher Chao index indicates a greater number of OTUs, reflecting higher species richness and greater microbial diversity within the sample. The Shannon index is used to measure microbial species diversity within a sample. The value is altered by the number and abundance of species found. A higher Shannon index indicates greater community diversity. The Simpson index is a measure of microbial diversity based on the degree of species dominance within a sample. It represents the probability that two randomly selected individuals belong to the same species, thereby reflecting both the richness and evenness of the microbial community.[39]

Diversity can be partitioned not only into richness and evenness but also across group samples. Alpha diversity represents the diversity within an individual sample, whereas gamma diversity describes the overall diversity across a collection of samples within a region or community. The relationship between alpha and gamma diversity is represented by beta diversity, which is determined by comparing pairs of samples and expresses the degree of similarity or dissimilarity between them.[38]

Beta diversity analyzes differences between samples and visualizes them using dimension-reduction methods such as PCoA or NMDS, with PERMANOVA as the statistical analyzer. The taxonomy composition is presented in a stacked bar plot at the genus or filum level to simplify interpretation. Comparative analyses between groups are performed to identify significantly different features (e.g., species, genes, or strains) using statistical tests or tools such as ALDEx2, edgeR, STAMP, or LEfSe, with results commonly visualized as volcano plots or Manhattan plots. This abundance-based approach is prone to false-positive findings because an increase in one feature can proportionally decrease another.[37]

7. Conclusion

The internal transcribed spacer (ITS) region, particularly ITS1 and ITS2, is the primary molecular marker for high-resolution characterization of the human mycobiome because of its high genetic variability and ability to distinguish fungi at the species level. Although ITS1 provides

broad taxonomic coverage and higher diversity estimates, it is more susceptible to primer bias and sequence-length variation, whereas ITS2 offers more consistent amplification, improved taxonomic resolution, and broader reference-database support, making it the preferred marker for many human mycobiome studies. Advances in next-generation sequencing (NGS) platforms, including Illumina, Ion Torrent, Pacific Biosciences (PacBio), and Oxford Nanopore Technologies (ONT), together with bioinformatics pipelines, such as QIIME2 and DADA2, and curated databases such as UNITE, have greatly improved the accuracy and comprehensiveness of fungal community profiling. While the combined use of ITS1 and ITS2 is often recommended to maximize taxonomic coverage and minimize identification bias, the study objectives, sample type, and target fungal taxa should ultimately guide the choice of marker.

8. Author Contribution

R.S.D. conceptualized the idea and concept, collected data, and wrote the manuscript. C.R.T., M.W., I.A.A., Y.R., and I.N. provided supervision and guidance on idea development, revised the manuscript, and funded it.

9. Acknowledgments

We express our gratitude to the Microbiology Department as part of the Biomedical Science Master's Program at the Faculty of Medicine, Universitas Indonesia, for facilitating this manuscript writing.

10. References

1. Belvoncikova P, Splichalova P, Videnska P, Gardlik R. *The Human Mycobiome: Colonization, Composition and the Role in Health and Disease*. *J Fungi*. 2022;8(10):1046.
2. Hoggard M, Vesty A, Wong G, Montgomery JM, Fourie C, Douglas RG, *et al*. *Characterizing the human mycobiota: A comparison of small subunit rRNA, ITS1, ITS2, and large subunit rRNA genomic targets*. *Front Microbiol*. 2018;9:2208.
3. Cui L, Morris A, Ghedin E. *The human mycobiome in health and disease*. *Genome Med*. 2013;5(7):63.
4. Satam H, Joshi K, Mangrolia U, Waghoo S, Zaidi G, Rawool S. *et al*. *Next-Generation Sequencing Technology: Current Trends and Advancements*. *Biology*. 2023;12(7):997.
5. Yang RH, Su JH, Shang JJ, Wu YY, Li Y, Bao DP. *et al*. *Evaluation of the ribosomal DNA internal transcribed spacer (ITS), specifically ITS1 and ITS2, for the analysis of fungal diversity by deep sequencing*. *PLoS One*. 2018;13(10):e0206428.
6. Blechert O, Zhan P. *Fungal rDNA Analysis in the NGS Era: Progress, Insights and Challenges*. *Fungal Genom Biol*. 2022;12(6):1-13.
7. Xu J. *Fungal DNA barcoding*. *Genome*. 2016;59(11):913-932.
8. Pieczul K, Świerczyńska I, Wójtowicz A. *Advanced rDNA-Based Detection of Wheat Pathogens in Grain Samples Using Next-Generation Sequencing (NGS)*. *Pathogens*. 2025;14(2):164.
9. Abdillah A, Kodio A, Ranque S. *Malian Children's Core Gut Mycobiome*. *Microorganisms*. 2024;12(5):926.
10. Yang P, Xu R, Chen F, Chen S, Khan A, Li L. *et al*. *Fungal gut microbiota dysbiosis in systemic lupus erythematosus*. *Front Microbiol*. 2023;14:1149311.
11. Yadav M, Ali S, Shrode RL, et al. *Multiple sclerosis patients have an altered gut mycobiome and increased fungal to bacterial richness*. *PLoS One*. 2022;17(4): :e0264556.
12. Mbareche H, Veillette M, Bilodeau G, Duchaine C. *Comparison of the performance of ITS1 and ITS2 as barcodes in amplicon-based sequencing of bioaerosols*. *PeerJ*. 2020;8:e8523.
13. Ali NABM, Mac Aogáin M, Morales RF, Tiew PY, Chotirmall SH. *Optimisation and Benchmarking of Targeted Amplicon Sequencing for Mycobiome Analysis of Respiratory Specimens*. *Int J Mol Sci*. 2019;20(20):4991.
14. Rolling T, Zhai B, Frame J, Hohl TM, Taur

Y. *Customization of a DADA2-based pipeline for fungal internal transcribed spacer 1 (ITS1) amplicon data sets*. *JCI Insight*. 2022;7(1):e151663.
15. Eren K, Taktakoğlu N, Pirim I. *DNA Sequencing Methods: From Past to Present*. *Eurasian J Med*. 2023;54(1):S47-S56.
16. Yu Y, Yang Q, Petropoulos E, Zhu T. *ITS3/ITS4 outperforms other ITS region and 18S rRNA gene primer sets for amplicon sequencing of soil fungi*. *Eur J Soil Sci*. 2022;73(6):e13329.
17. Motooka D, Fujimoto K, Tanaka R, Yaguchi T, Gotoh K, Maeda Y. *et al*. *Fungal ITS1 Deep-Sequencing Strategies to Reconstruct the Composition of a 26-Species Community and Evaluation of the Gut Mycobiota of Healthy Japanese Individual*. *Front Microbiol*. 2017;8;238.
18. Chen MF, Chang CH, Chiang-Ni C, Hsieh PH, Shih HN, Ueng SWN, *et al*. *Rapid analysis of bacterial composition in prosthetic joint infection by 16S rRNA metagenomic sequencing*. *Bone Joint Res*. 2019;8(8):367-377.
19. Callahan BJ, Grinevich D, Thakur S, Balamotis MA, Yehezkel TB. *Ultra-accurate microbial amplicon sequencing with synthetic long reads*. *Microbiome*. 2021;9(1):130.
20. Gehrig JL, Portik DM, Driscoll MD, Jacson E, Chakraborty S, Gratalo, *et al*. *Finding the right fit: evaluation of short-read and long-read sequencing approaches to maximize the utility of clinical microbiome data*. *Microb Genom*. 2022;8(3):000794.
21. Kim HS, Kim BH, Nam B, Oh SJ, Park SK. *et al*. *Oral–gut microbiome axis in a Korean cohort with inflammatory bowel disease and ankylosing spondylitis (INTEGRATE): a prospective and observational study protocol*. *BMJ Open*. 2025;15(8):e092075.
22. Radaschin DS, Iancu AV, Ionescu AM, Gurau G, Niculet E Bujoreanu FC, *et al*. *Comparative Analysis of the Cutaneous Microbiome in Psoriasis Patients and Healthy Individuals—Insights into Microbial Dysbiosis: Final Results*. *Int J Mol Sci*. 2024;25(19):10583.
23. Castaño C, Berlin A, Durling MB, Ihrmark K, Lindahl BD, Stenlid J. *et al*. *Optimized metabarcoding with Pacific biosciences enables semi-quantitative analysis of fungal communities*. *New Phytol*. 2020;228(3):1149-1158.
24. Zhang T, Li H, Ma S, Cao J, Liao H, Huang Q, *et al*. The newest Oxford Nanopore R10.4.1 full-length 16S rRNA sequencing enables the accurate resolution of species-level microbial community profiling. *Appl Environ Microbiol*. 2023;89(10):e0060523.
25. Portela ND, Mena C, Martín MG, Burstein VL, Chiapello LS, Pesoa SA. *Effect of DNA extraction method in gut fungal community assessment*. *Rev Argent Microbiol*. 2025;57(4):309-317.
26. Rintarhat P, Cho YJ, Koh H, Park S, Lee EJ, Lim H, *et al*. *Assessment of DNA extraction methods for human gut mycobiome analysis*. *R Soc Open Sci*. 2024;11(1):231129.
27. Thu MS, Sawaswong V, Chanchaem P, Klomkliew P, Campbell BJ, Hirankarn N, *et al*. *Optimization of a DNA extraction protocol for improving bacterial and fungal classification based on Nanopore sequencing*. *Access Microbiol*. 2024;6(10):000754.
28. Fiedorová K, Radvanský M, Němcová E, Grombiříková H, Bosák J, Černochová M, *et al*. *The Impact of DNA Extraction Methods on Stool Bacterial and Fungal Microbiota Community Recovery*. *Front Microbiol*. 2019;10:821.
29. Silverman JD, Bloom RJ, Jiang S, Durand, HK, Dallow E, Mukherjee S, *et al*. *Measuring and mitigating PCR bias in microbiota datasets*. *PLoS Comput Biol*. 2021;17(7):e1009113.
30. McNichol J, Berube PM, Biller SJ, Fuhrman JA. *Evaluating and Improving Small Subunit rRNA PCR Primer Coverage for Bacteria, Archaea, and Eukaryotes Using Metagenomes from Global Ocean Surveys*. *mSystems*. 2021;6(3):e0056521.

31. Hong BY, Driscoll M, Gratalo D, Jarvie T, Weinstock GM. *Improved DNA Extraction and Amplification Strategy for 16S rRNA Gene Amplicon-Based Microbiome Studies*. *Int J Mol Sci*. 2024;25(5):2966.
32. Poulsen CS, Ekstrøm CT, Aarestrup FM, Pamp SJ. *Library Preparation and Sequencing Platform Introduce Bias in Metagenomic-Based Characterizations of Microbiomes*. *Microbiol Spectr*. 2022;10(2):e0009022.
33. Chen Z, Ong CT, Nguyen LT, Lamb HJ, González-Recio O, Gutiérrez-Rivas M, *et al*. *Biases from Oxford Nanopore library preparation kits and their effects on microbiome and genome analysis*. *BMC Genomics*. 2025;26(1):504.
34. Halwachs B, Madhusudhan N, Krause R, Nilsson RH, Moissl-Eichinger C, Hogenauer C, *et al*. *Critical Issues in Mycobiota Analysis*. *Front Microbiol*. 2017;8:180.
35. Chen S. *Ultrafast one-pass FASTQ data preprocessing, quality control, and deduplication using fastp*. *iMeta*. 2023;2(2):e107.
36. Welzel M, Lange A, Heider D, Schwarz M, Freisleben B, Jensen M, *et al*. *Natrix: a Snakemake-based workflow for processing, clustering, and taxonomically assigning amplicon sequencing reads*. *BMC Bioinformatics*. 2020;21(1):526.
37. Liu YX, Qin Y, Chen T, Lu M, Qian X, Guo X. *et al*. *A practical guide to amplicon and metagenomic analysis of microbiome data. Protein Cell*. *Higher Education Press Limited Company*. 2021;12(5):315-330.
38. Wagner BD, Grunwald GK, Zerbe GO, Mikulich-Gilbertson SKM, Robertson CE, Zemanick ET, *et al*. *On the Use of Diversity Measures in Longitudinal Sequencing Studies of Microbial Communities*. *Front Microbiol*. 2018;9:1037.
39. CD Genomic. *The Use and Types of Alpha-Diversity Metrics in Microbial NGS*. 2026.