PacBio

Application note

Full-length HiFi sequencing of eukaryotic ITS and rRNA amplicons

Introduction

Taxonomic precision is critical for community ecology studies on microbial eukaryotes. For taxa such as fungi that are difficult to culture, the taxonomic precision of culture-free approaches affords greater understanding of their role in the mycobiome. Molecular-based approaches for classification include sequencing target regions chosen based on their rates of evolution and ease of amplification and detection. The regions most commonly used as barcode sequences for eukaryotes are the internal transcribed spacer region (ITS1 and ITS2) and the 18S and 28S rRNA genes (Figure 1).

PacBio® HiFi long-read amplicon sequencing has demonstrated advantages for high-throughput, high-resolution eukaryotic barcode sequencing.

In contrast to short reads, in which coverage is limited to the ITS1 or ITS2 regions only, HiFi reads span the full length of fungal ITS regions. This encompasses 800 bp, with the opportunity for amplicons up to ~5 kb if 18S and/or 28S genes are included to span the operon. Along with more conserved rRNA genes, these regions allow for reduced sequence similarity and higher-resolution taxonomic information at the species level (Tedersoo et al., 2018, Tedersoo et al., 2021, Figure 2). In addition to yielding greater taxonomic value, HiFi full-length amplicon sequencing is cost comparable to short-read, partial amplicon sequencing. In this Application note, we present general guidance for the amplification of full-length eukaryotic ITS and rRNA regions from complex community DNA samples.
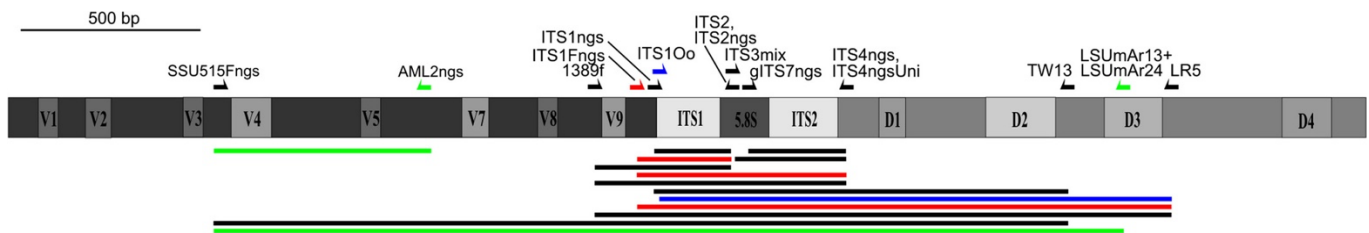

Figure 1. (from Tedersoo et al., 2018) Map of primers used to amplify eukaryotic rRNA genes and the internal transcribed spacer (ITS) therein. Lines below the map indicate amplicons. Colors of primers and lines indicate specificity: black, universal; red, Fungi; green, Glomeromycota; blue, Oomycetes.

PacBio

Figure 2. (from Tedersoo et al., 2018.) Relative operational taxonomic unit (OTU) richness of different primer pairs of (a) full internal transcribed spacer (ITS), (b) ITS1 and (c) ITS2 data subsets. White bars, PacBio MagBead loading on RSII instrument; light grey bars, PacBio Diffusion loading on RSII instrument; orange bars, Diffusion loading on PacBio Sequel® instrument; dark grey bars, Illumina MiSeq. This primer design can be easily adapted for the Sequel IIe instrument using the same steps as outlined in the 16S amplificiation protocol.

Workflow overview

Following amplification of the ITS and/or rRNA primers, amplification workflows (Figure 3) proceed according to the PacBio 16S bacterial amplification protocol.

Note: The target region primer sequences need to be replaced in this protocol to suit that of the project's needs and some PCR conditions may need to be adjusted.

The amplification workflow should begin with extracted, normalized DNA. PCR amplification of the desired target regions should be performed with barcoded primers for multiplexing, and should include extraction, negative, and positive controls. Amplification results should be confirmed with agarose gel or another approach. Next, barcoded amplicons are pooled together, and multiplexed amplicon libraries should be prepared for HiFi sequencing according to the protocol for the SMRTbell® prep kit 3.0. Finally, HiFi microbiome amplicon data can be analyzed with several bioinformatics tools (see Data analysis section).

Multiplexing approach

This assay uses a one-step PCR workflow with barcodes added to the target region primer sequences.

This protocol currently supports up to 192 unique sample barcodes with dual indexing. Amplicons are pooled prior to SMRTbell library construction following the PacBio procedure and the library is run on a SMRT® Cell.

Figure 3. HiFi sequencing ITS and rRNA amplicon workflow.

Amplicon selection and PCR conditions

There are many primer options available for eukaryotic rRNA and ITS amplicons. To date, many studies have compared the ability of multiple primer sets to best determine eukaryotic diversity with PacBio long-read sequencing. Here we give example primer sets with PCR conditions available in the references (Table 1).

Note: These primer sequences and PCR conditions have neither been verified nor are supported by PacBio.

It is recommended that you do further review of the literature provided here (see References) to determine which primers will best suit your needs. For example, primer performance can vary depending on amplicon size, sample environment, and ability to amplify certain taxa.

SMRTbell library construction and sequencing

It is recommended to QC the PCR products for pooling prior to SMRTbell library construction. The guidance for constructing multiplexed amplicon libraries can be found here.

The resulting library can be prepared for sequencing on PacBio long-read systems following instructions in the Sample Setup module of SMRT® Link.

Data analysis tools

HiFi microbiome amplicon data can be analyzed with either DADA2, QIIME2, Microbiome Helper, NextITS, One Codex, EzBiome, or the PacBio pipeline available on GitHub.

Conclusion

Studying fungi and other eukaryotes in complex communities is challenging because they can be difficult to identify and are therefore underrepresented in environmental microbiome studies. Full-length sequencing of ITS and/or rRNA amplicons provide a convenient approach to barcoding eukaryotes and yields greater taxonomic resolution compared to short reads. Here we demonstrate a modified HiFi 16S sequencing protocol for use with ITS and rRNA amplicon regions. This adapted workflow allows for the taxonomic precision needed to improve our understanding of microbial community ecology.

Target region and approximate size	Primer	Sequence (5'–3')	Primary target	Reference
ITS* ~800 bp	F: ITS1catta‖ R: ITS4ngsUni	5'-ACCWGCGGARGGATCATTA-3' 5'-CCTSCSCTTANTDATATGC-3'	Fungi, oomycetes, and stramenophiles	Tedersoo and Anslan (2019) Tedersoo and Lindahl (2016)
18S-ITS† ~1 kb	F: ITS9MUNngs R: ITS4ngsUni	5'-TACACACCGCCCGTCG-3' 5'-CCTSCSCTTANTDATATGC-3'	Most eukaryotes	Tedersoo and Lindahl (2016) Tedersoo and Lindahl (2016)
ITS-28S‡ ~1.8 kb	F: ITS1catta‖ R: LR5_TW14ngs	5'-ACCWGCGGARGGATCATTA-3' 5'-TCCTGAGGGAAACTTCG-3'	Most eukaryotes	Tedersoo and Anslan (2019) Tedersoo and Lindahl (2016)
18S-ITS-28S§ ~5 kb	F: V4F R: D11_3143R	5'-CCAGCASCYGCGGTAATTCC-3' 5'- RCCACAAGCYARTTATCC-3'	Most eukaryotes	Stoeck et al. (2010) Latz et al. (2022)

Table 1. Example primer sets by target region and taxa. *Flanks at the extreme 3' terminus of 18S at position 1790, spans the entire ITS regions 1 and 2, and flanks at the extreme 5' terminus of 28S at position 40. †Flanks in 18S at position 1635 before V9 (last variable region), spans the entire ITS regions 1 and 2, and flanks at the extreme 5' terminus of 28S at position 40. ‡Flanks at the extreme 3' terminus of 18S at position 1790, spans the entire ITS regions 1 and 2, and flanks in 28S at position 968 after D4. §Flanks in 18S before V4 at position 565, spans the entire ITS regions 1 and 2, and flanks in 28S at position 3143 after D11. ‖ITS1catta primer selects against amplifying plant DNA.

PacBio

Resources

[Procedure & checklist – Amplification of bacterial full-length 16S gene with barcoded primers](#)

[Procedure & checklist – Preparing multiplexed amplicon libraries using SMRTbell® prep kit 3.0](#)

References

These references represent a subset of publications that use **PacBio full-length amplicon sequencing**, as categorized by gene/region and taxa.

16S and fungal ITS

Wang, Q., Li, J., Yang, J., Zou, Y., & Zhao, X. Q. (2022). Diversity of endophytic bacterial and fungal microbiota associated with the medicinal lichen *Usnea longissima* at high altitudes. *Frontiers in Microbiology*, *13*, 958917. https://doi.org/10.3389/fmicb.2022.958917

Zhou, J. C., Wang, Y. F., Zhu, D., & Zhu, Y. G. (2023). Deciphering the distribution of microbial communities and potential pathogens in the household dust. *Science of The Total Environment*, *872*, 162250. https://doi.org/10.1016/j.scitotenv.2023.162250

16S and fungal ITS-28S

Li, J., He, Z., Yan, L., He, Y., & Yang, J. (2023). Analysis of the microbial community structure and flavor components succession during salt-reducing pickling process of *zhacai* (preserved mustard tuber). *Food Science & Nutrition*. https://doi.org/10.1002/fsn3.3297

18S and fungal ITS

Latz, M. A., et al. (2022). Short-and long-read metabarcoding of the eukaryotic rRNA operon: evaluation of primers and comparison to shotgun metagenomics sequencing. *Molecular Ecology Resources*, *22*(6), 2304-2318. https://doi.org/10.1111/1755-0998.13623

Runnel, K., et al. (2022). DNA barcoding of fungal specimens using PacBio long-read high-throughput sequencing. *Molecular Ecology Resources*, *22*(8), 2871-2879. https://doi.org/10.1111/1755-0998.13663

16S and 18S

Zhang, X., Bi, L., Gentekaki, E., Zhao, J., Shen, P., & Zhang, Q. (2023). Culture-Independent Single-Cell PacBio Sequencing Reveals Epibiotic *Variovorax* and Nucleus Associated *Mycoplasma* in the Microbiome of the Marine Benthic Protist *Geleia* sp. YT (Ciliophora, Karyorelictea). *Microorganisms*, *11*(6), 1500. https://doi.org/10.3390/microorganisms11061500

18S

Zhu, H., Li, S., Hu, Z., & Liu, G. (2018). Molecular characterization of eukaryotic algal communities in the tropical phyllosphere based on real-time sequencing of the 18S rDNA gene. *BMC Plant Biology*, *18*(1), 1-14. https://doi.org/10.1186/s12870-018-1588-7

16S, 18S, and ITS

Tedersoo, L., Albertsen, M., Anslan, S., & Callahan, B. (2021). Perspectives and benefits of high-throughput long-read sequencing in microbial ecology. *Applied and environmental microbiology*, *87*(17), e00626-21. https://doi.org/10.1128/AEM.00626-21

18S, 28S, and ITS

Tedersoo, L., Tooming-Klunderud, A., & Anslan, S. (2018). PacBio metabarcoding of Fungi and other eukaryotes: errors, biases and perspectives. *New Phytologist*, *217*(3), 1370-1385. https://doi.org/10.1111/nph.14776

PacBio