# MICROBIOME AND METAGENOME SEQUENCING WITH HIFI READS

Highly accurate long reads — HiFi reads — are the new benchmark for full-length 16S, metagenome profiling, and metagenome assembly

- Determine community composition at the species or strain level with competitively priced full-length 16S sequencing

- Profile metagenomes efficiently and cost-effectively with an average of eight full-length genes in every HiFi read

- Generate hundreds of high-quality (HQ) metagenome assembled genomes (MAGs), many of which are circular single-contigs

- Leverage epigenomic data to associate contigs and plasmids from closely related strains
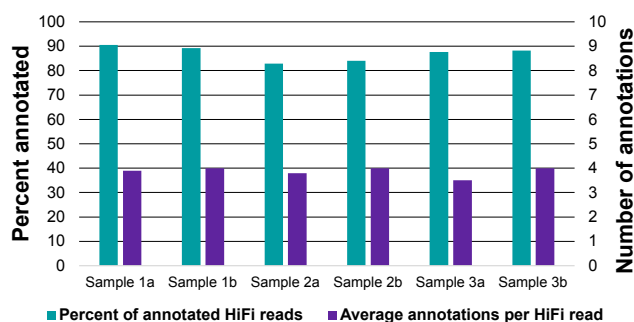
Full-length 16S rRNA sequencing

Species-level phylogenetic resolution

Percent unclassified

0   50   100

V1–V2   V1–V3   V1–V9

V3–V5   V4   V6–V9
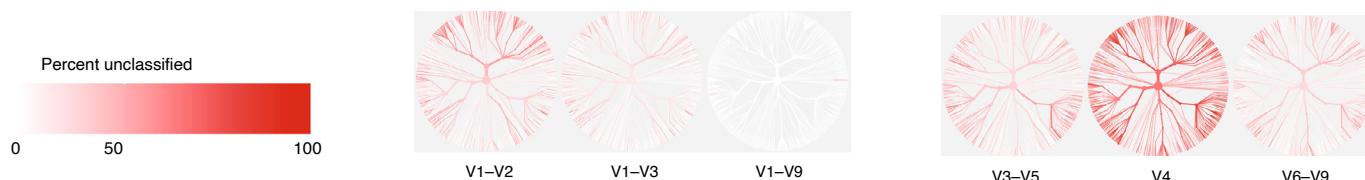
The proportion of 16S sequences from each bacterial genus that cannot be identified at the species level varies significantly depending on which variable region is used. Since the human gut can harbor a broad diversity of bacterial clades, only full-length sequences (V1–V9) can provide unbiased resolution of all the species that may be present.[1]

HiFi metagenomics

Precise species profiling, more functional annotations, more HQ MAGs, and more circular MAGs, even at lower coverage

Percent of annotated HiFi reads   Average annotations per HiFi read

High-quality MAGs   Single contig MAGs

With up to nine complete genes per HiFi read, PacBio® data provides rich functional information; nearly every read contributes to your understanding of the biological functions present in your microbial community.[2]
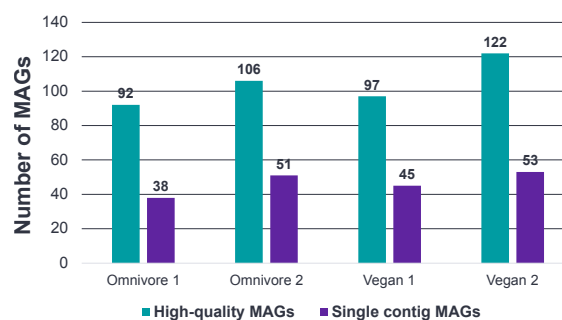
The unique combination of long read lengths and high accuracy overcomes many challenges involved with metagenome assembly such as distinguishing closely related strains in the same sample and yielding single-contig MAGs. Analysis of four human gut microbiome samples from *The BioCollective*.[3] High-quality MAGs ≥70% completeness, <10% contamination, <10 contigs.

Obtain more and richer metagenome functional information

- ~80–90% of HiFi reads are functionally annotatable

- Each HiFi read typically has an average of four functional annotations

Achieve standout metagenome assemblies

- ~90–125 HQ-MAGs per sample with ~17 Gb data; many are single-contig with ~50 per sample

- 417 HQ-MAGs in total across four samples

PacBiO

From DNA to resolved microbial communities

	Protocol/library prep[4]		Sequel II/IIe SMRT Cell 8M		Revio SMRT Cell 25M		Analysis tools
Full-length 16S rRNA High-resolution, cost-effective screening for microbial community studies	Amplification of bacterial full-length 16S gene with barcoded primers[5]	Preparing Kinnex™ libraries from 16S rRNA amplicons[6]	**Standard 16S** **192 samples** ~10k reads/ sample ~$9/sample	**Kinnex 16S** **768 samples** ~30k reads/ sample ~$5/sample	**Standard 16S** **384 samples** ~10k reads/ sample ~$5/sample	**Kinnex 16S** **1,536 samples** ~30k reads/ sample ~$4/sample	Analyze full-length 16S data with either *DADA2*,[8] *QIIME 2*,[9] *microbiomehelper*,[10] *OneCodex*,[11] *EZBiome*,[12] or with the PacBio GitHub pipeline[13]
Metagenome profiling Unbiased compositional and functional characterization of microbial communities	**Create ~10–15 kb libraries for HiFi metagenome sequencing** Preparing whole genome and metagenome libraries using SMRTbell® prep kit 3.0[7]		**48 samples** ~0.5 Gb/sample ~$109/sample		**96 samples** ~0.75 Gb/sample ~$90/sample		Study metagenome taxa and functions using *DIAMOND* and *MEGAN-LR*[14] or *Sourmash*[15] with the PacBio GitHub pipeline[16] or with *BugSeq*[17]
Metagenome assembly Generation of complete or near-complete MAGs from microbial populations			**4 samples** ~6 Gb/sample ~$400/sample		**12 samples** ~6 Gb/sample ~$160/sample		*De novo* assemble high-quality metagenomes with *hifiasm-meta*,[18] *metaFlye*,[19] or *metaMDBG*[20] and perform QC, binning, classification, and evaluation with the PacBio pipeline[16]

[*] Read lengths, reads/data, and number of samples per SMRT Cell and other sequencing performance results vary based on sample quality/type and insert size.

[†] Prices, listed in USD, are approximate and may vary by region. Pricing includes library and sequencing reagents run on a Sequel® II / IIe or Revio™ system and does not include instrument amortization or other reagents.

KEY REFERENCES

1. Johnson, J. S., et al. (2019) Evaluation of 16S rRNA gene sequencing for species and strain-level microbiome analysis. *Nature Communications.* 10(1), 5029.
2. Gehrig, J., et al. (2022) Finding the right fit: evaluation of short-read and long-read sequencing approaches to maximize the utility of clinical microbiome data. *Microbial Genomics*, 8(3), 10.1099/mgen.0.000794.
3. https://github.com/PacificBiosciences/pb-metagenomics-tools/blob/master/docs/PacBio-Data.md
4. Overview – Sequel systems application options and sequencing recommendations. PacBio documentation.
5. Procedure & checklist – Amplification of bacterial full-length 16S gene with barcoded primers. PacBio documentation
6. Procedure & checklist – Preparing Kinnex libraries from 16S rRNA amplicons. PacBio documentation.
7. Procedure & checklist – Preparing whole genome and metagenome libraries using SMRTbell prep kit 3.0. PacBio documentation.
8. Callahan, B. J., et al. (2019) High-throughput amplicon sequencing of the full-length 16S rRNA gene with single-nucleotide resolution. *Nucleic Acids Research*, 47(18), e103.
9. qiime2.org
10. https://github.com/LangilleLab/microbiome_helper/wiki/PacBio-CCS-Amplicon-SOP-v1-(qiime2)
11. onecodex.com
12. ezbiome.com
13. https://github.com/PacificBiosciences/pb-16S-nf
14. Huson, D. H., et al. (2018) MEGAN-LR: new algorithms allow accurate binning and easy interactive exploration of metagenomic long reads and contigs. *Biology direct*, 13(1), 6.
15. Pierce, N. T., et al. (2019) Large-scale sequence comparisons with sourmash. *F1000Research*, 8, 1006.
16. https://github.com/PacificBiosciences/pb-metagenomics-tools
17. bugseq.com and bugseq.com/pacbio
18. Feng, X., et al. (2022) Metagenome assembly of high-fidelity long reads with hifiasm-meta. *Nature methods,* 10.1038/s41592-022-01478-3
19. Kolmogorov, M., et al. (2020) metaFlye: scalable long-read metagenome assembly using repeat graphs. *Nature methods*, 17(11), 1103–1110.
20. https://github.com/GaetanBenoitDev/metaMDBG

Learn about metagenome sequencing:
pacb.com/microbial-pop

PacBio