Balancing cost, DNA input, and performance in long-read sequencing workflows for microbial and metagenomic applications

Jeremy E. Wilkinson[1], Juniper A. Lake[1], Christine Lambert[1], Aaron Wenger[1], Ian McLaughlin[1], Daniel M. Portik[1], Jacob Brandenburg[1]

[1]PacBio, Menlo Park, CA, USA

Background

High-quality long-read sequencing has become essential for genome assembly, variant detection, and metagenomic profiling and assembly. While the PacBio HiFi workflow delivers industry-leading read accuracy and contiguity, researchers increasingly require solutions that reduce input requirements, simplify preparation, and lower per-sample cost. To address these needs, we evaluated three PacBio workflows: standard HiFi library preparation[1], the Ampli-Fi protocol[2], and the seqWell LongPlex multiplexing kit[3] method (Fig. 1) on genomic DNA from microbial isolates and metagenomic samples. Libraries were sequenced on the PacBio Vega system and analyzed appropriately (Fig. 2). We assessed key metrics including library simplicity, sequence data quality, HiFi read length, HiFi data yield, barcoded reads, duplicate reads, genome assembly metrics, per-sample reagent costs, and preparation time (Fig. 1,3,4; Tbl. 1,2). This comparison illustrates the flexibility of PacBio's HiFi sequencing ecosystem and quantifies trade-offs among quality results, input requirements, cost, and workflow simplicity. This study provides researchers with a practical framework for selecting the optimal PacBio workflow to suit their needs.

PacBio HiFi WGS library prep solutions

SMRTbell prep kit 3.0, HiFi prep kit 96, HiFi plex prep kit 96

- Benefits:
  - Maximum mean HiFi read lengths
  - Methylation information
  - Multiplexing up to 384
  - Prep time is ~3.5 – 6.5 hours*
- Limitations:
  - Higher library prep cost per sample (~$40+ depending on prep kit used**)
  - Requires 50 - 500 ng gDNA per sample depending on plexity

Ampli-Fi protocol

- Benefits:
  - HiFi sequencing with as little as 1 ng of gDNA
  - Hi-C on HiFi with CiFi application
  - Access to challenging sample types
  - Multiplexing up to 384 x 96
  - Prep time is ~7 hours*
- Limitations:
  - Loss of methylation information
  - Slightly shorter mean HiFi read lengths
  - Higher library prep cost per sample (~$46+ depending on SMRTbell prep kit used and plexity**)

seqWell LongPlex multiplexing kit

- Benefits:
  - Shearing and barcoding in one step
  - Easier workflow and faster TAT (prep time is ~5.5 hours for 96 samples*)
  - Lowest library prep cost per sample (~$23+ depending on plexity and SMRTbell prep kit used**)
  - Multiplexing up to 384 x 96
- Limitations:
  - If PCR+ approach, loss of methylation information
  - Shorter mean HiFi read lengths
  - For metagenomes/fragmented/lower MW DNA, PCR-free has low demux rate
  - Requires ≥150 ng gDNA per sample

Figure 1. HiFi WGS library prep solutions outlining benefits and limitations for each workflow. *Manual preparation time without using SRE. Time does not include optional QC checkpoints as that will vary by lab equipment used. **All prices are listed in USD and cost may vary by region. Pricing includes library reagents and does not include instrument amortization, other reagents, sequencing, or DNA extraction.

Experimental design & analysis workflow

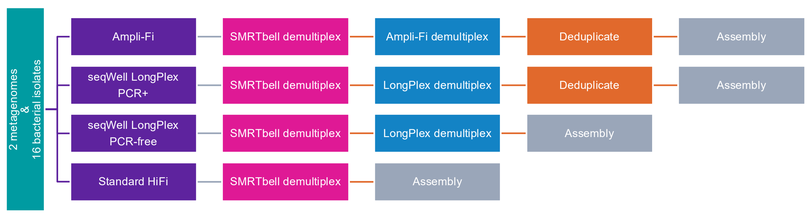

Figure 2. Study experimental design with sample type, library prep workflow, and analysis workflow. All Vega runs were 24 hours. HiFi data for both isolates and metagenomes were downsampled to approximately match samples with the lowest number of reads. The assembler used for bacterial isolates was hifiasm[4] and for metagenomes were hifiasm-meta[5] and myloasm[6]. Metagenome assembled contigs were then processed with the HiFi-MAG-Pipeline[7] for binning, QC, and metagenome-assembled genome (MAG) construction.

HiFi sequencing results and references

Sample	System	Read length (mean)	Read quality (median)	SMRTbell barcoded reads	% of movie reads	Ampli-Fi/ seqWell barcoded reads	% of Ampli-Fi/ seqWell demux reads	Unique reads	% duplicate reads	Final HiFi yield	% unique reads
Standard HiFi prep											
ZymoBIOMICS Gut Microbiome Standard (D6331)	Vega	6.32 kb	Q51	0.4 M	45.93%	–	–	–	–	2.52 Gb	–
ZymoBIOMICS Fecal Reference (D6323)	Vega	7.74 kb	Q47	4.18 M	4.38%	–	–	–	–	32.4 Gb	–
16 Bacterial isolates	Vega	15.29 kb	Q35	4.22 M	99.53%	–	–	–	–	64.6 Gb	–
Ampli-Fi											
ZymoBIOMICS Gut Microbiome Standard (D6331)	Vega	5.76 kb	Q50	4.53 M	49.69%	0.41 M	9.17%	0.41 M	0.02%	2.01 Gb	99.98%
ZymoBIOMICS Fecal Reference (D6323)	Vega	6.55 kb	Q48	4.53 M	49.69%	3.80 M	83.97%	3.79 M	0.16%	20.9 Gb	99.84%
16 Bacterial isolates	Vega	7.98 kb	Q42	8.13 M	100%	7.29 M	89.7%	7.29 M	0.007%	57.4 Gb	89.7%
seqWell LongPlex PCR+											
ZymoBIOMICS Gut Microbiome Standard (D6331)	Vega	4.84 kb	Q48	11.43 M	99.84%	2.03 M	17.77%	2.02 M	0.34%	9.80 Gb	99.66%
ZymoBIOMICS Fecal Reference (D6323)	Vega	5.38 kb	Q46	11.43 M	99.84%	9.29 M	81.30%	9.26 M	0.35%	49.8 Gb	99.65%
16 Bacterial isolates	Vega	5.93 kb	Q44	10.88 M	99.87%	10.82 M	99.5%	10.71 M	1.06%	62.3 Gb	98.4%
seqWell LongPlex PCR-free											
ZymoBIOMICS Gut Microbiome Standard (D6331)	Vega	5.15 kb	Q50	5.48 M	52.31%	0.68 M	12.48%	–	–	3.52 Gb	–
ZymoBIOMICS Fecal Reference (D6323)	Vega	5.63 kb	Q49	5.48 M	52.31%	2.45 M	44.66%	–	–	13.8 Gb	–
16 Bacterial isolates	Vega	7.84 kb	Q46	4.20 M	50.82%	2.97 M	70.6%	–	–	21.8 Gb	–

Table 1. General data metrics for the different library prep workflows on the samples in this study. NOTE: Demux rates don't accurately portray success due to differing pooling strategies across Vega runs.

Bacterial isolate	ATCC ID	Genome size (Expected)	Contigs (Expected)	GC%
Acinetobacter baumannii strain AYE	BAA-1710D-5	3.96 Mb	3.9M; 9.7k; 5.6k; 2.7k	39.4%
Campylobacter jejuni subsp. jejuni strain CIP 702	33560D-5	1.77 Mb	1.79M	30.2%
Clostridium perfringens	13124D-5	3.26 Mb	3.26M	28.4%
Cronobacter sakazakii strain 2001-10-01	BAA-894D-5	4.51 Mb	4.3M; 131.2k; 31.2k; 2.7k	56.7%
Enterobacter cloacae subsp. cloacae strain CDC 442-68	13047D-5	5.61 Mb	5.3M; 199.6k; 86.3k; 5.1k	54.6%
Enterococcus faecium strain MMC4	51559D-5	3.16 Mb	2.8M; 215.9k; 63.7k; 43.0k; 12.5k; 4.4k; 2.9k; 2.0k	37.7%
Escherichia coli strain MG1655	700926D-5	4.64 Mb	4.6M	50.8%
Escherichia coli strain H10407	35401D-5	5.39 Mb	5.2M; 94.8k; 67.9k; 66.6k; 5.8k; 5.5k	50.7%
Klebsiella pneumoniae strain 1000527, 7561	BAA-2146D-5	5.78 Mb	5.4M; 140.8k; 117.8k; 84.4k; 2.0k	57.0%
Listeria monocytogenes strain Li 2	19115D-5	2.95 Mb	3.0M	38.0%
Pseudomonas aeruginosa strain PAO1-LAC	47085D-5	6.30 Mb	6.3M; 12.3k; 11.7k	66.5%
Salmonella enterica subsp. Enterica serovar Typhimurium strain LT2	700720D-5	4.95 Mb	4.9M; 93.9k	52.2%
Shigella flexneri strain 2457T	700930D-5	4.82 Mb	4.6M; 165.7k; 4.1k; 3.2k	50.7%
Staphylococcus aureus subsp. aureus strain Seattle 1945	25923D-5	2.81 Mb	2.8M; 27.5k	32.9%
Vibrio harveyi strain BB120	BAA-1116D-5	6.07 Mb	3.8M; 2.2M; 89.0k	45.4%
Vibrio parahaemolyticus strain EB101	17802D-5	5.15 Mb	3.3M; 1.9M	45.3%

Table 2. Bacterial isolate references used in this study.

Isolate assembly results

Figure 3. Contig-level coverage of microbial isolates. After downsampling microbial isolates to 200 Mb*, HiFi reads were mapped to the respective reference genome to assess the mean coverage per contig (right) and the percent of each contig that achieved at least 10x coverage (left). The seqWell library preps (both PCR+ and PCR-free) had the fewest contig coverage dropouts with only *K. pneumoniae* and *E. cloacae* showing very low coverage on the smallest contig of each genome. However, we believe that *E. cloacae* is exhibiting reference divergence in contigs 2 and 4 since all library preps missed the same parts of the reference. *One sample, the standard library prep of *E. cloacae*, only sequenced to 170Mb and so wasn't downsampled.

Metagenome-assembled genome results

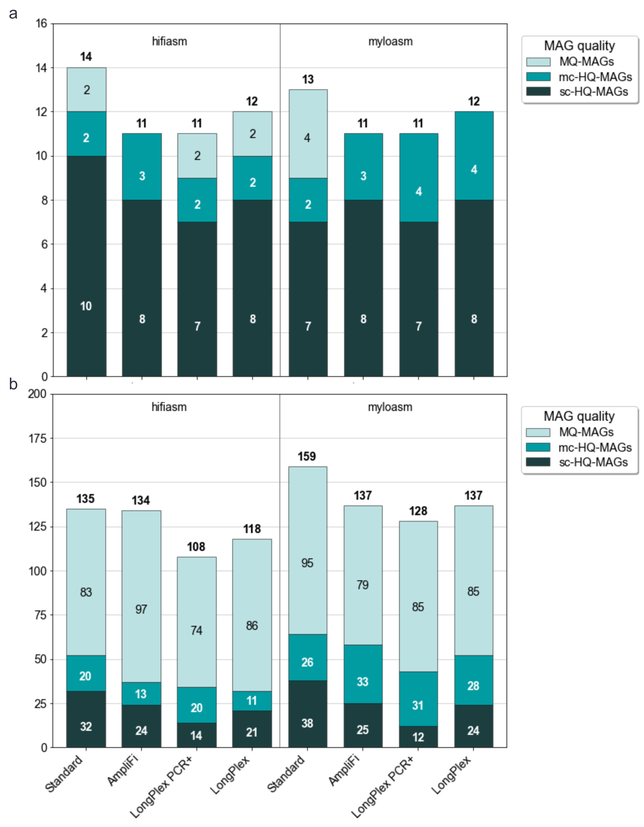

Figure 4. Metagenome-assembled genome (MAG) results from the Vega system for a) Zymo gut microbiome standard and b) Zymo pooled human fecal reference, each library prep workflow, and two assemblers used. Data for each sample was downsampled to approximately match the lowest number of reads.

sc-HQ-MAGs: single-contig high-quality MAGs (>90% completeness, <5% contamination); mc-HQ-MAGs: multiple-contig high-quality MAGs; MQ-MAGs: medium-quality MAGs (>50% completeness, <10% contamination).

Summary

- **All library prep methods** delivered similar assembly results for both isolate and metagenome whole genome sequencing.
- This result means that **a suitable choice of library prep method for HiFi sequencing can be made** given the specific use cases for the needs of each project or laboratory.
- **Standard HiFi prep** yielded the longest mean HiFi read lengths and maintains methylation information.
- The **Ampli-Fi protocol** supported the lowest DNA amounts, down to 1 ng, and enabled barcoding and pooling before SMRTbell library prep.
- The **seqWell LongPlex multiplexing kit** achieved the lowest per-sample cost depending on plexity, offered the simplest workflow via streamlined shearing and upfront barcoding before SMRTbell library prep, shears/recovers more plasmids, and maintains methylation information with the PCR-free workflow.
- **All workflows achieved per-sample costs below $50 USD**, depending on genome size and/or multiplexing, enabling high-throughput applications such as microbial isolate assembly, metagenome profiling, and metagenome assembly.

References

1. HiFi library preparation kits: https://www.pacb.com/products-and-services/consumables/library-prep-and-barcoding-kits/
2. Ampli-Fi protocol: https://www.pacb.com/wp-content/uploads/Procedure-checklist-Amplifying-genomic-DNA-for-SMRTbell-library-preparation-and-HiFi-sequencing.pdf
3. seqWell LongPlex multiplexing kit: https://seqwell.com/longplex-multiplexing/
4. Cheng, H., Concepcion, G. T., Feng, X., Zhang, H., & Li, H. (2021). Haplotype-resolved de novo assembly using phased assembly graphs with hifiasm. *Nature methods*, 18(2), 170–175. https://doi.org/10.1038/s41592-020-01056-5
5. Feng, X., Cheng, H., Portik, D., & Li, H. (2022). Metagenome assembly of high-fidelity long reads with hifiasm-meta. *Nature methods*, 19(6), 671–674. https://doi.org/10.1038/s41592-022-01478-3
6. Shaw, J., Marin, M. G., & Li, H. (2026). High-resolution metagenome assembly for modern long reads with myloasm. *Nature biotechnology*, https://doi.org/10.1038/s41587-026-03053-z
7. HiFi-MAG-Pipeline: https://github.com/PacificBiosciences/pb-metagenomics-tools/tree/master/HiFi-MAG-Pipeline