

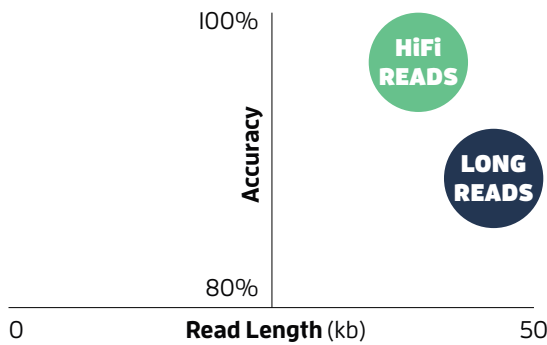
WHOLE GENOME SEQUENCING FOR *DE NOVO* ASSEMBLY

BEST PRACTICES



Single Molecule, Real-Time (SMRT®) Sequencing on the Sequel® II System enables easy and affordable generation of high-quality *de novo* assemblies. With megabase size contig N50s, accuracies >99.99%, and phased haplotypes, you can do more biology – capturing undetected SNVs, fully intact genes, and regulatory elements embedded in complex regions.

CHOOSE THE SEQUENCING MODE THAT'S RIGHT FOR YOUR PROJECT



PacBio is the only sequencing technology to offer highly accurate long reads (HiFi reads)!

HiFi Reads

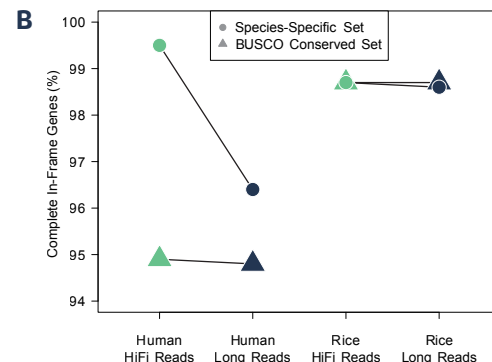
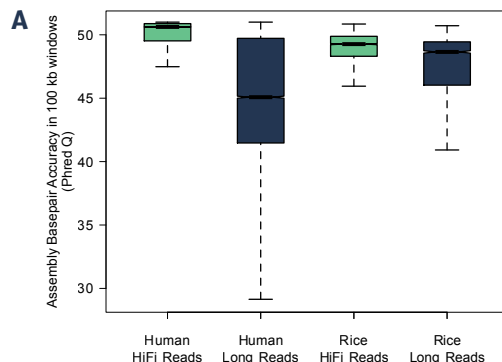
- Highly accurate long reads with minimum accuracy of Q20 (99%)
- Small file sizes and fast analysis time
- Assemble up to a 2 Gb genome in a single SMRT Cell 8M for ~\$1,300* or scale up for larger genomes
- Run up to 200 samples (2 Gb) per year, per system*

Long Reads

- Longest reads, with half of data >50 kb and maximum read lengths up to 175 kb
- Shorter sequencing run times
- Assemble up to a 3 Gb genome in a single SMRT Cell 8M for ~\$1,300* or scale up for larger genomes
- Run up to 400 samples (3 Gb) per year, per system*




*Read lengths, reads/data per SMRT Cell 8M, 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 System and does not include instrument amortization or other reagents.

GENERATE COMPLETE AND ACCURATE GENOME ASSEMBLIES



HiFi reads and long reads both provide assemblies with high base pair accuracy and gene completeness. (A) Base pair accuracy as measured against a reference is >Q40, even outside of known gene regions. (B) BUSCO and larger, species-specific gene sets are >94% complete in both human and rice genome assemblies.

WORKFLOW RECOMMENDATIONS – From DNA to Reference-quality Genome Assembly

HiFi Reads Workflow ²	Long Reads Workflow ²
 <p>LIBRARY PREP</p> <ul style="list-style-type: none"> - Start with >15 µg of unamplified genomic DNA - Prepare a library in ~5 hours with the SMRTbell® Express Template Prep Kit 2.0³ - Enrich for 15-20 kb inserts with size selection 	<ul style="list-style-type: none"> - Start with >5 µg of unamplified genomic DNA - Prepare a library in ~3 hours with the SMRTbell Express Template Prep Kit 2.0⁴ - Enrich for >30 kb inserts with size selection
 <p>SMRT SEQUENCING</p> <ul style="list-style-type: none"> - Sequence using the circular consensus sequencing (CCS) mode - Sequence up to a 2 Gb genome in a single SMRT Cell 8M* - Coverage recommendations: <ul style="list-style-type: none"> - 15-fold HiFi coverage for typical diploid genomes - >20-fold HiFi coverage for repeat-rich genomes 	<ul style="list-style-type: none"> - Sequence using the continuous long read (CLR) sequencing mode - Sequence up to a 3 Gb genome in a single SMRT Cell 8M* - Coverage recommendation: <ul style="list-style-type: none"> - >30-fold long-read coverage per haplotype
 <p>DATA ANALYSIS</p> <ul style="list-style-type: none"> - Generate HiFi reads with CCS analysis in SMRT Link or at the command line^{5,6,7} - Use FALCON and FALCON-Unzip3 to assemble and phase HiFi reads⁸ 	<ul style="list-style-type: none"> - Utilize the PacBio analytical portfolio, including community developed tools, to assemble and phase long reads⁹

HiFi assembly of a ~3 Gb genome in ~31 hours vs ~62 hours with long reads

Long read data of a ~3 Gb genome in ~15 hours vs ~60 hours with HiFi

Studying Small-bodied Organisms?

Check out the low DNA input workflow for high-quality genome assemblies from as little as 150 ng of gDNA¹⁰



KEY REFERENCES

1. SMRT Sequencing modes. PacBio Webpage.
2. Overview – Sequel Systems application options and sequencing recommendations. PacBio Documentation.
3. Procedure & Checklist – Preparing HiFi SMRTbell libraries using SMRTbell express template prep kit 2.0. PacBio Documentation.
4. Procedure & Checklist – Preparing gDNA libraries using the SMRTbell express template preparation kit 2.0. PacBio Documentation.
5. SMRT Link user guide (v8.0). PacBio Documentation.
6. SMRT Tools reference guide (v8.0). PacBio Documentation.
7. CCS: Generate highly accurate single-molecule consensus reads (HiFi reads). PacBio GitHub.
8. Assembling HiFi data: FALCON-Unzip3. PacBio GitHub.
9. Analytical software whole genome sequencing. PacBio Webpage.
10. Procedure & Checklist – Preparing SMRTbell libraries using express template prep kit 2.0 with low DNA input.