WHOLE GENOME SEQUENCING FOR *DE NOVO* ASSEMBLY – BEST PRACTICES

PacBio® HiFi reads provide long read lengths (up to 25 kb), high accuracy (>99.9%), and detection of 5-methylcytosine (5mC) methylation to quickly and affordably generate contiguous, complete, and correct *de novo* genome assemblies of even the most complex genome.

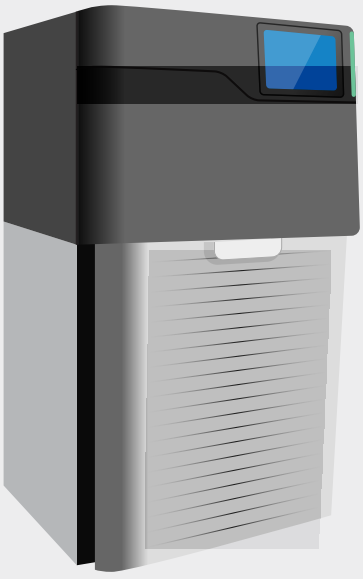
Contiguity
High contig N50

Completeness
No missing bases or fragmented genes

Correctness
High base accuracy and phased alleles

Compute
Small file sizes and fast analysis time

CpG methylation
Accurate epigenome from standard sequencing runs



The Sequel® IIe system provides cost-effective and scalable HiFi sequencing of any genome

Large or complex genomes

Fast and efficient assembly of even the largest genomes with haplotype resolution of complex polyploids

The 27 Gb hexaploid genome of the redwood tree was sequenced and assembled in under two weeks

Human genomes at scale

Flexible and scalable workflows for sequencing 100s–1000s of human genomes per year from a variety of sample types

Assemble a human genome in one day

PacBio

HIFI SEQUENCING WORKFLOW RECOMMENDATIONS

from DNA to reference-quality genome assembly

Sample + library prep

- Extract HMW DNA with kits available for different sample types
- Prepare a library with SMRTbell® prep kit 3.0[1]

SMRT® sequencing

- Use the Sequel II or IIe system and SMRT® Cell 8M to sequence to desired coverage depth for complexity of genome
- 10- to 15-fold coverage per haplotype recommended

Data analysis

- Use SMRT® Link genome assembly[4], or open-source tools[5] including IPA, HiCanu, or hifiasm to assemble and phase the genome and epigenome
- Obtain base-level methylation profiles with 5mC detection
- Example datasets available at **pacb.com/dataset**

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 Sequel II or IIe 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 or IIe system and does not include instrument amortization or other reagents.

Flexible options for DNA input quantities

	DNA input	Genome size limit
Standard HiFi sequencing[1]	1 µg/Gb	None
Ultra-low DNA input sequencing[6] (amplification-based)	5 ng	500 Mb

Learn about whole genome sequencing for *de novo* assembly: **pacb.com/wgs**

Connect with PacBio for more info:
North America: **nasales@pacb.com**
South America: **sasales@pacb.com**
EMEA: **emea@pacb.com**
Asia Pacific: **apsales@pacb.com**

KEY REFERENCES

1. **Procedure & checklist — Preparing whole genome and metagenome libraries using SMRTbell prep kit 3.0** PacBio documentation.
2. **Overview – Sequel systems application options and sequencing recommendations.** PacBio documentation.
3. **Technical note: Preparing DNA for PacBio HiFi sequencing – Extraction and quality control.** PacBio literature.
4. **Product brochure: SMRT Link – Explore and analyze your data with confidence.** PacBio literature.
5. Recommended Open-source Genome Assembly Tools: **IPA, HiCanu, hifiasm.**
6. **Application note: Considerations for using the low and ultra-low DNA input workflows for whole genome sequencing.** PacBio literature.

102-193-627 REV02 15APR2022