WHOLE GENOME SEQUENCING FOR DE NOVO ASSEMBLY — BEST PRACTICES

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

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

The Sequel® Ile 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

“If your genome isn’t HiFi, it’s no longer reference grade.”
Kevin McKernan, Medicinal Genomics
HIFI SEQUENCING WORKFLOW RECOMMENDATIONS
from DNA to reference-quality genome assembly

Library prep
- Prepare a library in ~5 hours with the SMRTbell® express template prep kit 2.0 from ≥5 μg of unamplified genomic DNA.
- Enrich for longer inserts with size selection

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, or open-source tools including IPA, HiCanu, or hifiasm to assemble and phase the genome
- 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

Flexible options for DNA input quantities

<table>
<thead>
<tr>
<th>DNA input</th>
<th>Genome size limit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard HiFi sequencing¹</td>
<td>≥5 μg</td>
</tr>
<tr>
<td>Low DNA input sequencing⁶</td>
<td>300 ng</td>
</tr>
<tr>
<td>Ultra-low DNA input sequencing⁶ (amplification-based)</td>
<td>5 ng</td>
</tr>
</tbody>
</table>

*Scales with DNA input amount

KEY REFERENCES
2. Overview — Sequel systems application options and sequencing recommendations. PacBio documentation.
5. Recommended Open-source Genome Assembly Tools: IPA, HiCanu, hifiasm.

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

Information in this document is subject to change without notice. PacBio assumes no responsibility for any errors or omissions in this document. Certain notices, terms, conditions, and/or use restrictions may pertain to your use of PacBio products and/or third party products. Refer to the applicable PacBio terms and conditions of sale and to the applicable license terms at http://www.pacb.com/legal-and-trademarks/terms-and-conditions-of-sale/. PacBio, the PacBio logo, SMRT, SMRTbell, Iso-Seq, and Sequel are trademarks of PacBio. All other trademarks are the sole property of their respective owners.