Technical Overview: Low DNA Input Library Preparation Using SMRTbell Express Template Prep Kit 2.0

Sequel System ICS v8.0 / Sequel Chemistry 3.0 / SMRT Link v9.0
Sequel II System ICS v9.0 / Sequel II Chemistry 2.0 / SMRT Link v9.0
Sequel Ile System ICS v10.0 / Sequel II Chemistry 2.0 / SMRT Link v10.0
Low DNA Input Library Preparation Using SMRTbell Express Template Prep Kit 2.0

1. Low DNA Input Workflow Overview
2. Low DNA Input Sample QC Requirements
3. Low DNA Input Library Sample Preparation Workflow Details
4. Low DNA Input Library Sequencing Workflow Details
5. HiFi Sequencing Data Analysis Recommendations for De Novo Assembly Applications
6. Low DNA Input Library Example Performance Data
7. Technical Documentation & Applications Support Resources
8. Appendix: General Recommendations for High-Molecular Weight gDNA QC and Handling for SMRTbell Library Construction
LOW DNA INPUT SEQUENCING: HOW TO GET STARTED

Low- / Ultra-Low DNA Input Application Note

Application-Specific Procedure & Checklist

Application Consumable Bundle Purchasing Guide

Library Construction, Sequencing & Analysis

Application Note: Considerations for Using the Low and Ultra-Low DNA Input Workflows for Whole Genome Sequencing (101-995-900)

Summary overview of experimental design considerations, sample preparation workflow and data analysis recommendations

Procedure & Checklist – Preparing HiFi Libraries from Low DNA Input Using SMRTbell Express Template Prep Kit 2.0 (101-730-400)

Technical documentation containing sample library construction and sequencing preparation protocol details

PacBio Application Consumable Bundle Purchasing Guide (PG100-051320)

Purchasing Guide enables users to easily order required consumables needed to prepare an SMRTbell library to run a specific type of application on the Sequel II and IIe Systems*

Application Consumable Bundles include reagents for library construction, primer annealing and polymerase binding. Core PacBio-branded SMRT Sequencing consumables (SMRT Cells, Sequencing Kits & SMRT Oil), plastics and other 3rd-party reagents are not included in the application bundles

Library Construction (SMRTbell Express TPK 2.0)

Multiplex Up To 2 low DNA input samples with the Sequel II and IIe Systems using Barcoded Overhang Adapters (BOA)

HiFi Sequencing

Aim for 10- to 15-fold HiFi Read Coverage per Haplotype for Assembly

Data Analysis (SMRT Link)

Demultiplexing Analysis

Genome Assembly Analysis

gDNA QC & Shearing

12 kb – 20 kb Target DNA Shear Size

* Application Consumable Bundles include reagents for library construction, primer annealing and polymerase binding. Core PacBio-branded SMRT Sequencing consumables (SMRT Cells, Sequencing Kits & SMRT Oil), plastics and other 3rd-party reagents are not included in the application bundles
WHEN IS IT APPROPRIATE TO CONSIDER THE LOW DNA OR ULTRA-LOW DNA INPUT WORKFLOWS?

We recommend considering the genome assembly project as a whole, from DNA extraction to bioinformatics, to establish your experimental design.

PacBio’s standard HiFi library preparation workflow requires at least ~3 μg of DNA input per 1 Gb of genome length, or ~15 μg for a human sample – but for some samples, it is not possible to extract this amount of DNA for sequencing.

For samples where between ~300 ng and ~3 μg of DNA is available, the Low DNA Input Workflow enables users to generate high-quality genome assemblies of small-bodied organisms.

For samples where even less DNA is available (as low as 5 ng), the amplification-based Ultra-Low DNA Input Workflow is available.

- Where possible, the standard HiFi workflow run on the Sequel II and Ile Systems gives you the highest quality results for both genome assembly and human variant detection projects.
- However, if you are sample-limited, the Low and Ultra-Low DNA Input Workflows can still provide excellent results.

See Application Note: Considerations for Using the Low and Ultra-Low DNA Input Workflows for Whole Genome Sequencing (PN 101-995-900) for further guidance on choosing a specific workflow for sequencing your particular sample type.
WHEN IS IT APPROPRIATE TO CONSIDER THE LOW DNA OR ULTRA-LOW DNA INPUT WORKFLOWS?

DETAILS OF STANDARD, LOW DNA INPUT, AND ULTRA-LOW DNA INPUT HIFI SEQUENCING WORKFLOWS ON THE SEQUEL II SYSTEM.

<table>
<thead>
<tr>
<th>SAMPLE &amp; PROJECT CONSIDERATIONS</th>
<th>STANDARD HIFI SEQUENCING</th>
<th>LOW DNA INPUT SEQUENCING (2-PLEX)</th>
<th>LOW DNA INPUT SEQUENCING (SINGLE SAMPLE)</th>
<th>ULTRA-LOW DNA INPUT SEQUENCING</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimum DNA Input</td>
<td>&gt;3 μg / 1 Gb genome</td>
<td>300 ng for each genome</td>
<td>400 ng</td>
<td>5 ng</td>
</tr>
<tr>
<td>Amplification Based?</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Genome Size Limit</td>
<td>N/A</td>
<td>600 Mb for each genome</td>
<td>1 Gb</td>
<td>500 Mb</td>
</tr>
<tr>
<td>Supported Applications</td>
<td>De novo Assembly</td>
<td>De novo Assembly</td>
<td>De novo Assembly</td>
<td>De novo Assembly</td>
</tr>
</tbody>
</table>

- **De novo Assembly**
  - Human Variant Detection

- **ASSEMBLY**
  - De novo assembly of insect/arthropod genomes (Up to 500 Mb)

- **VARIANT DETECTION**
  - Variant detection (SNPs, Indels, SVs) in human genomes (3 Gb)

- **Unsupported Applications**
  - Metagenomics sequencing

**Ultra-Low DNA Input:**

- **SUPPORTED APPLICATIONS**
  - De novo assembly for microbes, plants, vertebrates, or other non-DNA limited sample types

- **UNSUPPORTED APPLICATIONS**
  - De novo assembly for microbes, plants, vertebrates, or other non-DNA limited sample types
Low DNA Input Workflow Overview
LOW DNA INPUT SAMPLE PREPARATION PROCEDURE DESCRIPTION

- Procedure & Checklist - Preparing HiFi Libraries from Low DNA Input Using SMRTbell Express Template Prep Kit 2.0 (PN 101-730-400) protocol document describes how to prepare SMRTbell libraries with low DNA input amounts for sequencing on the Sequel, Sequel II and Sequel Ile Systems for WGS de novo assembly applications using HiFi reads

- The procedure describes preparing HiFi libraries from >150 ng of input genomic DNA (gDNA) for the Sequel System and from >400 ng of input gDNA for the Sequel II and Ile Systems using SMRTbell Express TPK 2.0

  - Using this non-multiplexed library preparation procedure, genome assemblies of up to 1 Gb can be achieved with the Sequel II and Ile Systems

- This procedure also provides recommendations for multiplying a maximum of 2 small genomes on the Sequel II or Ile System starting with >300 ng of gDNA per genome

  - Using this multiplexed library preparation procedure, genome assemblies of up to 600 Mb/genome can be achieved

### APPLICATIONS

WHOLE GENOME SEQUENCING

- Noctua fimbriata
  - Moth
- Anopheles coluzzii
  - Mosquito
- Drosophila erecta
  - West African Fruit Fly
- Oryza sativa subsp. Indica
  - Rice Plant
- Schistosoma mansoni
  - Parasitic Worm
- Vanessa atalanta
  - Red Admiral Butterfly
**WORKFLOW SUMMARY OVERVIEW:** SINGLE-SAMPLE (NON-MULTIPLEXED) LOW DNA INPUT LIBRARY PREPARATION (SEQUEL, SEQUEL II AND SEQUEL Ile SYSTEMS)

1. **Genomic DNA QC and Shearing**
   - Recommended starting input gDNA quality: >30 kb
   - Minimum required input DNA mass per sample for singleplex library construction
     - Sequel System: ≥150 ng
     - Sequel II and Ile Systems: ≥400 ng
   - Target DNA shear size: 12 – 20 kb

2. **SMRTbell Express TPK 2.0 Library Construction (~4 hours)**
   - Follow instructions in *Procedure & Checklist – Preparing HiFi Libraries from Low DNA Input Using SMRTbell Express Template Prep Kit 2.0 (PN 101-730-400)* for preparation of non-multiplexed samples

3. **Sequencing Preparation**
   - Generate ≥10 Gb of HiFi base data per Sequel II System SMRT Cell 8M* depending on low DNA input sample DNA quality and library insert size

4. **Data Analysis**
   - Utilize SMRT Link to generate highly accurate and long single-molecule reads (HiFi reads) using the Circular Consensus Sequencing (CCS) analysis application or perform CCS analysis on-instrument using the Sequel Ile System
   - For HiFi de novo assembly, can use SMRT Link Genome Assembly analysis application (powered by IPA) or other third-party software (e.g., HiCanu, hifiasm)

* Read lengths, number of reads, data per SMRT Cell, and other sequencing performance results vary based on sample quality/type and insert size, among other factors.
WORKFLOW SUMMARY OVERVIEW: MULTIPLEXED LOW DNA INPUT LIBRARY PREPARATION (SEQUEL II AND IIe SYSTEMS)

1. Genomic DNA QC and Shearing
   - Recommended starting input gDNA quality: >30 kb
   - Minimum required input DNA mass per sample for multiplexed library construction
     ▪ Sequel II and IIe Systems: ≥300 ng per sample
   - Target DNA shear size: 12 – 20 kb

2. SMRTbell Express TPK 2.0 Library Construction (~4 hours)
   - Follow instructions in Procedure & Checklist – Preparing HiFi Libraries from Low DNA Input Using SMRTbell Express Template Prep Kit 2.0 (PN 101-730-400) for preparation of multiplexed samples
     ▪ Use PacBio Barcoded Overhang Adapter Kit 8A/8B (PN 101-628-400/500) to multiplex up to 2 samples per SMRT Cell 8M

3. Sequencing Preparation
   - Generate ≥10 Gb of HiFi base data per Sequel II System SMRT Cell 8M* depending on low DNA input sample DNA quality and library insert size

4. Data Analysis
   - Utilize SMRT Link to generate highly accurate and long single-molecule reads (HiFi reads) using the Circular Consensus Sequencing (CCS) analysis application or perform CCS analysis on-instrument using the Sequel IIe System
     - For HiFi de novo assembly, can use SMRT Link Genome Assembly analysis application (powered by IPA) or other third-party software

* Read lengths, number of reads, data per SMRT Cell, and other sequencing performance results vary based on sample quality/type and insert size, among other factors.
Low DNA Input Library Sample QC Requirements
Table below summarizes the supported low DNA input workflows and the required DNA quality and quantity for each.

<table>
<thead>
<tr>
<th>LOW DNA INPUT LIBRARY TYPE</th>
<th>RECOMMENDED INPUT gDNA AMOUNT</th>
<th>REQUIRED INPUT gDNA QUALITY</th>
<th>gDNA SHEARING METHOD</th>
<th>REQUIRED SHEARED gDNA SIZE DISTRIBUTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-multiplexed low DNA input for the Sequel System 1 sample per SMRT Cell 1M Up to 300 Mb per genome</td>
<td>&gt;150 ng</td>
<td>Majority of gDNA &gt;30 kb</td>
<td>Megaruptor System</td>
<td>12-20 kb sheared gDNA mean fragment size is optimal</td>
</tr>
<tr>
<td>Non-multiplexed low DNA input for the Sequel II or IIe Systems 1 sample per SMRT Cell 8M Up to 1 Gb per genome</td>
<td>&gt;400 ng</td>
<td>Majority of gDNA &gt;30 kb</td>
<td>Megaruptor System</td>
<td>12-20 kb sheared gDNA mean fragment size is optimal</td>
</tr>
<tr>
<td>Multiplexed low DNA input for the Sequel II or IIe Systems 2 samples per SMRT Cell 8M Up to 600 Mb per genome</td>
<td>&gt;300 ng per sample</td>
<td>Majority of gDNA &gt;30 kb</td>
<td>Megaruptor System</td>
<td>12-20 kb sheared gDNA mean fragment size is optimal</td>
</tr>
</tbody>
</table>
RECOMMENDED TOOLS FOR GENOMIC DNA QUANTIFICATION AND QUALIFICATION

When working with small amounts of input DNA, accurate sizing and quantification are critical for enabling the generation of sufficient coverage of reads to produce a high-quality genome assembly.

DNA Quantification

- For quantification of gDNA to be used with the low DNA input library preparation workflow, we recommend using the Qubit fluorometer and Qubit High Sensitivity (HS) DNA assay reagents (Thermo Fisher Scientific)
  - Qubit dsDNA HS assay quantitation range: 0.2 – 100 ng
  - Measure the gDNA sample concentration as recommended by the manufacturer.

Qubit™ dsDNA HS Assay Kit

Qubit 4 Fluorometer

https://www.thermofisher.com/order/catalog/product/Q33230#Q33230
RECOMMENDED TOOLS FOR GENOMIC DNA QUANTIFICATION AND QUALIFICATION (CONT.)

DNA Sizing

- Three commercially available systems that may be used to evaluate gDNA size distribution are listed below with links to recommended procedures.
  - **Note:** The CHEF Mapper and Pippin Pulse systems may not be appropriate options in some cases since both tools require at least 50 ng of DNA sample for analysis.

<table>
<thead>
<tr>
<th>DNA SIZING METHOD</th>
<th>COMMENTS</th>
<th>PROCEDURE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Femto Pulse System (Agilent)</td>
<td>• <em>Highly recommended</em></td>
<td>Agilent Femto Pulse Website</td>
</tr>
<tr>
<td></td>
<td>• Requires 200 – 500 pg</td>
<td></td>
</tr>
<tr>
<td>CHEF Mapper XA PFGE System (Bio-Rad)</td>
<td>• Requires &gt;50 ng</td>
<td>Procedure &amp; Checklist - Using the BIO-RAD® CHEF Mapper XA Pulsed Field Electrophoresis System</td>
</tr>
<tr>
<td>Pippin Pulse System (Sage Science)</td>
<td>• Requires &gt;50 ng</td>
<td>Procedure &amp; Checklist - Using the Sage Science Pippin Pulse Electrophoresis Power Supply System</td>
</tr>
</tbody>
</table>

We highly recommend the use of the **Femto Pulse System (Agilent)** for low DNA input applications because of its ability to evaluate size distributions using only ~200 – 500 picograms of DNA.
The size distribution of the starting input gDNA sample is critical to successful HiFi SMRTbell library construction and sequencing on the Sequel, Sequel II and Sequel Ile Systems. Always evaluate the quality of the gDNA samples before proceeding with library construction.

- PacBio recommends working with samples where the majority of the input gDNA is >30 kb
- SMRTbell libraries constructed from low-quality, fragmented gDNA samples tend to generate shorter read lengths compared to libraries constructed from high-quality, high-molecular weight gDNA samples
  - This may be caused by sequencing termination events occurring at damaged sites introduced during DNA extraction that were not repairable by DNA Damage repair enzymes
  - Additionally, fragmented DNA usually contains excess levels of short fragments (<10 kb) that cannot be removed even with the most aggressive AMPure PB bead size-selection procedure
  - Such types of low-quality DNA samples typically lead to the generation of short subread lengths and – consequently – a more fragmented de novo genome assembly
EVALUATION OF GENOMIC DNA FOR LOW DNA INPUT LIBRARY CONSTRUCTION (CONT.)

Example Femto Pulse sizing QC analysis of individual mosquito gDNA samples of varying quality.

- **Samples 1** and **2** show size distributions with the majority of the fragments >30 kb.
  - Both samples are appropriate for shearing and constructing HiFi SMRTbell libraries for *de novo* assembly.

- If the gDNA is severely fragmented such that a significant proportion of the fragments are ≤10 kb (e.g., **Samples 3** and **4**), we recommend re-extraction of the gDNA to obtain a higher-quality sample for HiFi SMRTbell library construction.

- If the sample is moderately fragmented with a size distribution of ~10 – 20 kb with relatively few <10 kb fragments, a HiFi SMRTbell library may optionally be constructed directly without shearing (with the caveats noted on the previous slide).
Low DNA Input Library Sample Preparation Workflow Details
PROCEDURE & CHECKLIST - PREPARING HIFI LIBRARIES FROM LOW DNA INPUT USING SMRTBELL EXPRESS TEMPLATE PREP KIT 2.0

- This document (PN 101-730-400) describes preparing HiFi libraries from >150 ng of input genomic DNA (gDNA) for the Sequel System and from >400 ng of input gDNA for the Sequel II and IIe Systems using SMRTbell Express TPK 2.0.
  
  ▪ Using this procedure, genome assemblies of up to 1 Gb can be achieved with the Sequel II and IIe Systems. (Assemblies of up to 300 Mb can be achieved with the Sequel System.)

- This procedure also provides recommendations for multiplexing a maximum of 2 small genomes (up to 600 Mb/genome) on the Sequel II and IIe Systems, from >300 ng of gDNA per genome.

- Protocol document contains:
  1. General laboratory best practices recommendations
  2. Guidelines for QC evaluation of gDNA samples for low-Input library construction
  3. Enzymatic steps for preparation of non-multiplexed and multiplexed low DNA input SMRTbell libraries
  4. Instructions for size-selection of low DNA input libraries using the AMPure PB bead size selection method
  5. Sample setup guidance for preparing low DNA input libraries for sequencing on the Sequel, Sequel II and Sequel IIe Systems
## LIST OF REQUIRED MATERIALS AND EQUIPMENT

<table>
<thead>
<tr>
<th>ITEM</th>
<th>VENDOR</th>
<th>PART NUMBER</th>
</tr>
</thead>
<tbody>
<tr>
<td>DNA Sizing QC (One of the following)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Femto Pulse System <em>(Highly Recommended)</em></td>
<td>Agilent</td>
<td>M5330AA</td>
</tr>
<tr>
<td>Pippin Pulse Electrophoresis Power Supply</td>
<td>Sage Science</td>
<td>PP10200</td>
</tr>
<tr>
<td>Pulsed Field Gel Electrophoresis System: CHEF Mapper XA</td>
<td>Bio-Rad</td>
<td>170-3670</td>
</tr>
<tr>
<td>DNA Quantitation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Qubit Fluorometer</td>
<td>ThermoFisher Scientific</td>
<td>Q33226</td>
</tr>
<tr>
<td>Qubit 1X dsDNA HS Assay Kit</td>
<td>ThermoFisher Scientific</td>
<td>Q33230</td>
</tr>
<tr>
<td>DNA Shearing</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Megaruptor 3 System</td>
<td>Diagenode</td>
<td>10145</td>
</tr>
<tr>
<td>Megaruptor 3 Shearing Kit</td>
<td>Diagenode</td>
<td></td>
</tr>
<tr>
<td>SMRTbell Library Preparation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SMRTbell Express Template Prep Kit 2.0</td>
<td>PacBio</td>
<td>100-938-900</td>
</tr>
<tr>
<td>Barcoded Overhang Adapter Kit 8A (8 adapters) or Barcoded Overhang Adapter Kit 8B (8 adapters)</td>
<td>PacBio</td>
<td>101-628-400</td>
</tr>
<tr>
<td>AMPure PB Beads</td>
<td>PacBio</td>
<td>100-265-900</td>
</tr>
<tr>
<td>SMRTbell Enzyme Cleanup Kit</td>
<td>PacBio</td>
<td>101-746-400</td>
</tr>
<tr>
<td>Elution Buffer</td>
<td>PacBio</td>
<td>101-633-500</td>
</tr>
<tr>
<td>DNA Lo Bind microfuge tubes</td>
<td>PacBio</td>
<td>022431021</td>
</tr>
<tr>
<td>Wide Orifice Tips <em>(Tips LTS W-O 200UL Ftr RT-L200WFLR)</em></td>
<td>Rainin</td>
<td>17014294</td>
</tr>
<tr>
<td>Tube Rotator</td>
<td>VWR</td>
<td>10136-084</td>
</tr>
</tbody>
</table>
SMRTBELL EXPRESS TEMPLATE PREP KIT 2.0 AND SMRTBELL ENZYME CLEANUP KIT REAGENT HANDLING RECOMMENDATIONS

- Several reagents in the kit are sensitive to temperature and vortexing
- PacBio highly recommends:
  - Never leaving reagents at room temperature
  - Working on ice at all times when preparing master mixes
  - Finger tapping followed by a quick-spin prior to use

### LIST OF TEMPERATURE-SENSITIVE REAGENTS INCLUDED IN SMRTBELL EXPRESS TPK 2.0 AND SMRTBELL ENZYME CLEANUP KIT.

<table>
<thead>
<tr>
<th>PACBIO KIT</th>
<th>REAGENT</th>
<th>WHERE USED</th>
</tr>
</thead>
<tbody>
<tr>
<td>DNA Prep Additive</td>
<td>Remove Single-Strand Overhangs</td>
<td></td>
</tr>
<tr>
<td>DNA Prep Enzyme</td>
<td>Remove Single-Strand Overhangs</td>
<td></td>
</tr>
<tr>
<td>DNA Damage Repair Mix v2</td>
<td>DNA Damage Repair</td>
<td></td>
</tr>
<tr>
<td>End Prep Mix</td>
<td>End-Repair/A-tailing</td>
<td></td>
</tr>
<tr>
<td>Overhang Adapter v3*</td>
<td>Ligation</td>
<td></td>
</tr>
<tr>
<td>Ligation Mix</td>
<td>Ligation</td>
<td></td>
</tr>
<tr>
<td>Ligation Additive</td>
<td>Ligation</td>
<td></td>
</tr>
<tr>
<td>Ligation Enhancer</td>
<td>Ligation</td>
<td></td>
</tr>
<tr>
<td>SMRTbell Express Template Prep kit 2.0 (PN 100-938-900)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Enzyme A</td>
<td>Nuclease Treatment</td>
<td></td>
</tr>
<tr>
<td>Enzyme B</td>
<td>Nuclease Treatment</td>
<td></td>
</tr>
<tr>
<td>Enzyme C</td>
<td>Nuclease Treatment</td>
<td></td>
</tr>
<tr>
<td>Enzyme D</td>
<td>Nuclease Treatment</td>
<td></td>
</tr>
<tr>
<td>SMRTbell Enzyme Cleanup Kit</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Barcoded Overhang Adapters (not included with SMRTbell Express TPK 2.0) are also temperature-sensitive reagents.
## SMRTBELL EXPRESS TPK 2.0 LIBRARY CONSTRUCTION WORKFLOW COMPARISON FOR SUPPORTED LOW DNA INPUT LIBRARY TYPES

<table>
<thead>
<tr>
<th>LIBRARY CONSTRUCTION STEP</th>
<th>LOW DNA INPUT SINGLEPLEX SAMPLE (SEQUEL SYSTEM)</th>
<th>LOW DNA INPUT SINGLEPLEX SAMPLE (SEQUEL II AND Ile SYSTEM)</th>
<th>LOW DNA INPUT MULTIPLEXED SAMPLE (SEQUEL II AND Ile SYSTEM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recommended Starting Input gDNA Amount for Shearing</td>
<td>&gt;150 ng (Single Sample)</td>
<td>&gt;400 ng (Single Sample)</td>
<td>&gt;300 ng per sample (Up to 2 Samples)</td>
</tr>
<tr>
<td>gDNA Quality</td>
<td>&gt;30 kb</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DNA Shearing</td>
<td></td>
<td>Megaruptor 3 System</td>
<td></td>
</tr>
<tr>
<td>Minimum DNA Input for Remove Single-Stranded DNA Overhangs</td>
<td>&gt;150 ng (Sheared gDNA)</td>
<td>&gt;300 ng (Sheared gDNA)</td>
<td>&gt;200 ng per sample (Sheared gDNA)</td>
</tr>
<tr>
<td>Remove Single-Stranded DNA Overhangs</td>
<td></td>
<td>37°C / 15 min</td>
<td></td>
</tr>
<tr>
<td>DNA Damage Repair</td>
<td></td>
<td>37°C / 30 min</td>
<td></td>
</tr>
<tr>
<td>End-Repair/A-tailing</td>
<td></td>
<td>20°C / 10 min; 65°C / 30 min</td>
<td></td>
</tr>
<tr>
<td>Adapter</td>
<td>Overhang Adapter v3</td>
<td>Overhang Adapter v3</td>
<td>Barcoded Overhang Adapter</td>
</tr>
<tr>
<td>Adapter Ligation</td>
<td></td>
<td>20°C / 60 min (Overnight ligation optional)</td>
<td></td>
</tr>
<tr>
<td>Nuclease Treatment</td>
<td>Not required</td>
<td>Not required</td>
<td>37°C / 60 min</td>
</tr>
<tr>
<td>1st AMPure PB Bead Purification</td>
<td>0.45X AMPure PB Beads</td>
<td>0.45X AMPure PB Beads</td>
<td>0.80X AMPure PB Beads</td>
</tr>
<tr>
<td>2nd AMPure PB Bead Purification (Removal of &lt;3 kb SMRTbell Templates)</td>
<td></td>
<td>2.2X AMPure PB Beads (40% v/v)</td>
<td></td>
</tr>
<tr>
<td>Library Construction Yield</td>
<td>50%</td>
<td>50%</td>
<td>30%</td>
</tr>
</tbody>
</table>
BEST PRACTICES RECOMMENDATIONS FOR PREPARING LOW DNA INPUT SMRTBELL LIBRARIES

1. Use wide-bore tips for all mixing steps.

2. Throughout the procedure, **do not flick the tubes to mix**. Flicking induces damage to DNA. Instead, use gentle pipetting with wide-bore tips to mix reagents. Note that template preparation reagents should first be dispensed with a standard pipette tip (i.e., P10 or P20) and then a wide-bore tip should be used for pipette mixing.

3. Never vortex tubes containing high-molecular weight genomic DNA samples.

4. Minimize the number of freeze/thaw cycles the gDNA undergoes to reduce DNA damage.

5. Allowing sufficient time for thawing aliquots of DNA, as partially frozen DNA is prone to shearing.

6. Always set your heat blocks or thermocyclers to the appropriate temperature for incubations before proceeding with the procedure.

7. Ensure that the AMPure PB beads are at **room temperature** prior to performing the purification steps.

8. When performing AMPure PB bead purification steps, note that 80% ethanol is hygroscopic and should be prepared FRESH to achieve optimal results. Also, 80% ethanol should be stored in a tightly capped polypropylene tube for no more than 3 days.

9. Always follow best practices for DNA quantitation using a Qubit fluorometer system. Use the Qubit dsDNA High Sensitivity (HS) reagent kit.
DNA SHEARING RECOMMENDATIONS FOR LOW DNA INPUT LIBRARY CONSTRUCTION

- PacBio recommends the **Megaruptor System (Diagenode)** for shearing genomic DNA because of its ability to generate tight shear distributions and good post-shear recovery efficiencies.

- The Megaruptor 3 System is recommended for shearing gDNA to a **target size distribution mode ~12 – 20 kb** for low DNA input applications.

- To shear gDNA using the Megaruptor 3 system, generally follow the manufacturer’s recommendations.

- After shearing the gDNA samples, evaluate the size distribution using a Femto Pulse system and then proceed with the Concentrate DNA Using AMPure PB Beads step.

### RECOMMENDED STARTING INPUT GENOMIC DNA AMOUNTS FOR SHEARING.

<table>
<thead>
<tr>
<th>LOW DNA INPUT LIBRARY RECOMMENDATIONS</th>
<th>LOW DNA INPUT SINGLE-SAMPLE (SEQUEL SYSTEM)</th>
<th>LOW DNA INPUT SINGLE-SAMPLE (SEQUEL II AND IIe SYSTEM)</th>
<th>LOW DNA INPUT MULTIPLEXED SAMPLE (SEQUEL II AND IIe SYSTEM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recommended Starting Input gDNA Amount for Shearing</td>
<td>&gt;150 ng (Single Sample)</td>
<td>&gt;400 ng (Single Sample)</td>
<td>&gt;300 ng per sample (Up to 2 Samples)</td>
</tr>
</tbody>
</table>
DNA SHEARING RECOMMENDATIONS FOR LOW DNA INPUT LIBRARY CONSTRUCTION (CONT.)

Sheared gDNA QC considerations when multiplexing low DNA input samples on the Sequel II and IIe Systems

For constructing multiplexed low DNA input HiFi SMRTbell libraries for the Sequel II and IIe Systems, we recommend the following:

- It is necessary to shear the gDNA so that the fragment size distribution is 12 kb – 20 kb.
  - This insert size range maximizes the detection of barcodes during the demultiplexing analysis step.

- Samples for a multiplex experiment must have the same average size and distribution to avoid biased read representation.
  - It is highly recommended to work with samples that contain mostly high-molecular weight DNA so that they can be sheared to achieve a similar target size and distribution for HiFi SMRTbell library construction.

Example Femto Pulse sizing QC analysis of input gDNA and sheared gDNA samples. Samples 1 and 2 are examples of gDNA suitable for library construction using the low DNA input procedure for multiplexing. Both samples were sheared using a Megaruptor 3 system with speed setting 33. Sample 1 sheared DNA mode size is ~16 kb and Sample 2 is ~14 kb.
PACBIO BARCODED OVERHANG ADAPTERS FOR MULTIPLEXED LOW DNA INPUT LIBRARY CONSTRUCTION

- PacBio Barcoded Overhang Adapter Kit 8A (PN 101-628-400) or 8B (PN 101-628-500) are available for multiplexing low DNA input samples.
  - 8 barcodes in each kit; 6 reactions per BOA tube
  - Compatible with SMRTbell Express TPK 2.0
  - Includes Sequencing Primer v4

- To download the barcode FASTA sequences for BOA Kit 8A/8B, visit PacBio’s Multiplexing Resources webpage
  - FASTA filename: Sequel_16_Barcodes_v3.zip (Link)
SAMPLE POOLING BEST PRACTICES RECOMMENDATIONS FOR MULTIPLEXED LOW DNA INPUT LIBRARY CONSTRUCTION

- Always quantify libraries before pooling.
  - Since DNA is in limited quantity, PacBio recommends using the Qubit dsDNA High Sensitivity (HS) Assay Kit for concentration measurements.

- Equal mass pooling is recommended for samples that have a similar distribution (+/- 10%)
  - The total mass of the pooled library must be ≥150 ng (e.g. 75 sample 1 + 75 ng sample 2)
  - If the total mass of the two samples is <150 ng, use the remaining sample to adjust the total mass to 150 ng.
    - This is to ensure that there is enough library material for loading at least one SMRT Cell 8M
    - In such cases, it is possible that the resulting number of reads per sample may be slightly imbalanced

- After pooling, perform AMPure PB bead size selection to remove fragments <3 kb

<table>
<thead>
<tr>
<th>STEP</th>
<th>Pooling</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Pool barcoded libraries into a single 1.5 mL DNA Lo-Bind tube.</td>
</tr>
<tr>
<td>2</td>
<td>Mix and spin down the contents of the tube with a quick spin in a microfuge.</td>
</tr>
<tr>
<td>3</td>
<td>Bring the total volume of the pooled sample to 100 μL with Elution Buffer.</td>
</tr>
<tr>
<td>4</td>
<td>Proceed to the “Size-Selection with AMPure PB Beads to remove &lt;3 kb SMRTbell Templates” section below.</td>
</tr>
</tbody>
</table>
AMPURE PB BEAD SIZE SELECTION FOR LOW DNA INPUT LIBRARY CONSTRUCTION

- This final AMPure PB bead purification step removes SMRTbell templates <3 kb
- AMPure PB bead stock solution is first diluted to 40% (vol./vol.) with Elution Buffer and subsequently used for purification
- The final AMPure PB bead concentration is critical to the success of this procedure.
  - Therefore, accurate pipetting is of utmost importance to achieve a final 40% (v/v) AMPure PB bead working solution in EB

<table>
<thead>
<tr>
<th>Reagent</th>
<th>Volume</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elution Buffer</td>
<td>3.0 mL</td>
<td>✓</td>
</tr>
<tr>
<td>AMPure PB Bead (stock reagent, thoroughly mixed)</td>
<td>2.0 mL</td>
<td></td>
</tr>
<tr>
<td>Total Volume</td>
<td>5.0 mL</td>
<td></td>
</tr>
</tbody>
</table>

1. Bring the AMPure PB bead stock solution to room temperature.
2. Vortex the stock solution for 30 seconds to mix well.
3. Using a P1000 pipette, transfer 3.0 mL of Elution Buffer into a 15 mL conical tube.
4. Add 2.0 mL of the AMPure PB bead stock solution to the 3.0 mL of Elution Buffer. When pipetting the viscous AMPure PB bead stock solution, pipette slowly to ensure that the volume aspirated is as precise as possible. Large residual AMPure solution adhering to the tip should be removed prior to addition to the 3.0 mL of Elution Buffer.
5. Vortex the diluted AMPure PB bead solution for 30 seconds to mix well before use. This solution may be stored at 4°C for 2 weeks for future use.
LOW DNA INPUT SMRTBELL LIBRARY CONSTRUCTION YIELDS

- Overall, SMRTbell library yields in this low-input workflow are typically 50% (starting from sheared DNA input into the first enzymatic reaction) for the single-sample workflow and 30% for the multiplexing workflow.

- Depending on the final size of the library, sufficient amounts of SMRTbell template material to run approximately 3 or more SMRT Cells 1M can be generated for the Sequel System.

- The Sequel II and Ile Systems require higher on-plate loading concentrations and, as a result, the amount of SMRTbell library material generated in this procedure is enough for only one SMRT Cell 8M.
Low DNA Input Library Sequencing Workflow Details
SAMPLE SETUP RECOMMENDATIONS FOR LOW DNA INPUT LIBRARIES – SEQUEL SYSTEM (CHEMISTRY 3.0 / SMRT LINK V9.0)

- Follow SMRT Link v9.0 Sample Setup instructions using the recommendations provided in the Quick Reference Card – Loading and Pre-Extension Time Recommendations for the Sequel System for low DNA input samples with the additional guidance noted below.

### Note 1:
- Select “CLR” sequencing mode in SMRT Link v9.0 Sample Setup to enable a 20:1 Primer:Template ratio and 30:1 Polymerase:Template ratio.
- Select “CCS” sequencing mode in SMRT Link v9.0 Run Design to enable CCS auto-analysis.

### Note 2:
- For the Sequel System, we recommend using a 20-hour movie collection time for HiFi sequencing with low DNA input samples.

---

<table>
<thead>
<tr>
<th>Applications</th>
<th>Sequencing Mode</th>
<th>Library Prep Kit</th>
<th>Binding Kit</th>
<th>Sequencing Primer</th>
<th>Pol Binding Time (hr)</th>
<th>Complex Cleanup</th>
<th>Loading Concentration Range (pM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Large insert (&gt;15 kb size-selection cutoff)</td>
<td>CLR</td>
<td>Express Prep 2.0</td>
<td>Binding Kit 3.0</td>
<td>v4</td>
<td>1</td>
<td>1.2X AMPure PB Beads</td>
<td>2 - 8</td>
</tr>
<tr>
<td>Microbial Multiplex (10 kb)</td>
<td>CLR</td>
<td>Express Prep 2.0</td>
<td>Binding Kit 3.0</td>
<td>v4</td>
<td>1</td>
<td>1.2X AMPure PB Beads</td>
<td>4 - 8</td>
</tr>
<tr>
<td>Low DNA Input (&gt;10 kb, AMPure PB Bead size-selection)</td>
<td>CLR</td>
<td>Express Prep 2.0</td>
<td>Binding Kit 3.0</td>
<td>v4</td>
<td>1</td>
<td>1.2X AMPure PB Beads</td>
<td>4 - 8</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Applications</th>
<th>Pre-Extension Time (hr)</th>
<th>Movie Collection Time (hr)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Large insert (&gt;15 kb size-selection cutoff)</td>
<td>0</td>
<td>10</td>
</tr>
<tr>
<td>Microbial Multiplex (10 kb)</td>
<td>2</td>
<td>10</td>
</tr>
<tr>
<td>Low DNA Input (&gt;10 kb, AMPure PB Bead size-selection)</td>
<td>2</td>
<td>10</td>
</tr>
</tbody>
</table>

---

https://www.pacb.com/support/documentation/
SAMPLE SETUP RECOMMENDATIONS FOR LOW DNA INPUT LIBRARIES – SEQUEL II SYSTEM (CHEMISTRY 2.0 / SMRT LINK V9.0)

- Follow SMRT Link v9.0 Sample Setup instructions using the recommendations provided in the Quick Reference Card – Loading and Pre-Extension Time Recommendations for the Sequel II System for low DNA input samples with the additional guidance noted below.

**Note:**
- Select "CLR" sequencing mode in SMRT Link v9.0 Sample Setup to enable a 20:1 Primer:Template ratio and 30:1 Polymerase:Template ratio.
- Select “CCS” sequencing mode in SMRT Link v9.0 Run Design to enable CCS auto-analysis.

---

### Loading and Pre-Extension Time Recommendations

<table>
<thead>
<tr>
<th>Applications</th>
<th>Data Type</th>
<th>Library Prep Kit</th>
<th>Binding Kit</th>
<th>Sequencing Primer</th>
<th>Pot Binding Time (hr)</th>
<th>Complex Cleanup</th>
<th>Loading Concentration Range (pM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>De Novo Assembly – Continuous Long Reads (&gt;15 kb)</td>
<td>CLR</td>
<td>Express Prep 2.0</td>
<td>Binding Kit 2.0</td>
<td>v4</td>
<td>1</td>
<td>1.2X AMPure PB Beads</td>
<td>30 - 70</td>
</tr>
<tr>
<td>Structural Variation Detection (&gt;15 kb)</td>
<td>CLR</td>
<td>Express Prep 2.0</td>
<td>Binding Kit 2.0</td>
<td>v2</td>
<td>4</td>
<td>1.2X AMPure PB Beads</td>
<td>70 - 100</td>
</tr>
<tr>
<td>De Novo Assembly – Microbial Multiplexing (10 kb – 15 kb)</td>
<td>CLR</td>
<td>Express Prep 2.0</td>
<td>Binding Kit 2.0</td>
<td>v4</td>
<td>1</td>
<td>1.2X AMPure PB Beads</td>
<td>70 - 100</td>
</tr>
<tr>
<td>De Novo Assembly – Low DNA Input (15 kb)</td>
<td>CCS</td>
<td>Express Prep 2.0</td>
<td>Binding Kit 2.0</td>
<td>v4</td>
<td>1</td>
<td>1.2X AMPure PB Beads</td>
<td>30 - 70</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Applications</th>
<th>Pre-Extension Time (hr)</th>
<th>Movie Collection Time (hr)</th>
</tr>
</thead>
<tbody>
<tr>
<td>De Novo Assembly – Continuous Long Reads (&gt;15 kb)</td>
<td>0</td>
<td>15</td>
</tr>
<tr>
<td>Structural Variation Detection (&gt;20 kb)</td>
<td>2 hrs (±20 kb), 4 hrs (&gt;20 kb)</td>
<td>15</td>
</tr>
<tr>
<td>De Novo Assembly - Microbial Multiplexing (10 kb – 15 kb)</td>
<td>2</td>
<td>15</td>
</tr>
<tr>
<td>De Novo Assembly – Low DNA Input (15 kb)</td>
<td>2</td>
<td>30</td>
</tr>
</tbody>
</table>

[https://www.pacb.com/support/documentation/](https://www.pacb.com/support/documentation/)
SAMPLE SETUP RECOMMENDATIONS FOR LOW DNA INPUT LIBRARIES – SEQUEL II AND IIe SYSTEMS (CHEMISTRY 2.0 / SMRT LINK V10.0)

- Follow SMRT Link v10.0 Sample Setup instructions using the recommendations provided in the Quick Reference Card – Loading and Pre-Extension Time Recommendations for the Sequel II/Ile Systems for preparing Low DNA Input library samples for sequencing

→ For SMRT Link v10.0 (or higher): Select ‘Low DNA Input’ from the Application field drop-down menu in the SMRT Link Sample Setup and SMRT Link Run Design user interface

---

<table>
<thead>
<tr>
<th>Applications</th>
<th>Data Type</th>
<th>Library Prep Kit</th>
<th>Binding Kit</th>
<th>Sequencing Primer</th>
<th>Pot Binding Time (hr)</th>
<th>Complex Cleanup</th>
<th>Loading Concentration Range (pM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>De Novo Assembly – Continuous Long Reads (&gt;15 kb)</td>
<td>CLR</td>
<td>Express Prep 2.0</td>
<td>Binding Kit 2.0</td>
<td>v4</td>
<td>1</td>
<td>1.2x AMPure PB Beads</td>
<td>30 - 70</td>
</tr>
<tr>
<td>Structural Variation Detection (&gt;15 kb)</td>
<td>CLR</td>
<td>Express Prep 2.0</td>
<td>Binding Kit 2.0</td>
<td>v2</td>
<td>4</td>
<td>1.2x AMPure PB Beads</td>
<td>30 - 70</td>
</tr>
<tr>
<td>De Novo Assembly – Microbial Multiplexing (10 kb – 15 kb)</td>
<td>CLR</td>
<td>Express Prep 2.0</td>
<td>Binding Kit 2.0</td>
<td>v4</td>
<td>4</td>
<td>1.2x AMPure PB Beads</td>
<td>70 - 100</td>
</tr>
<tr>
<td>De Novo Assembly – Low DNA Input (15 kb)</td>
<td>CCS</td>
<td>Express Prep 2.0</td>
<td>Binding Kit 2.0</td>
<td>v4</td>
<td>1</td>
<td>1.2x AMPure PB Beads</td>
<td>30 - 70</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Applications</th>
<th>Pre-Extension Time (hr)</th>
<th>Movie Collection Time (hr)</th>
</tr>
</thead>
<tbody>
<tr>
<td>De Novo Assembly – Continuous Long Reads (&gt;15 kb)</td>
<td>0</td>
<td>15</td>
</tr>
<tr>
<td>Structural Variation Detection</td>
<td>2 hrs (≤20 kb), 4 hrs (&gt;20 kb)</td>
<td>15</td>
</tr>
<tr>
<td>De Novo Assembly - Microbial Multiplexing (10 kb – 15 kb)</td>
<td>2</td>
<td>15</td>
</tr>
<tr>
<td>De Novo Assembly – Low DNA Input (15 kb)</td>
<td>2</td>
<td>30</td>
</tr>
</tbody>
</table>

https://www.pacb.com/support/documentation/
IMPORTING THE BARCODE FASTA FILE INTO SMRT LINK FOR AUTOMATED DEMULTIPLEXING OF POOLED LOW DNA INPUT LIBRARY SAMPLES

- **Note:** SMRT Link v9.0 (and higher) software installations by default come pre-bundled with a FASTA file containing a list of PacBio barcodes recommended for use with multiplexed SMRT sequencing applications.

- If your SMRT Link installation does not already include an appropriate barcode FASTA file, the following steps describe how to import such a file for use in automated demultiplexing (refer to “Importing Data” section in the SMRT Link User Guide):

  1. Download the FASTA file containing the relevant barcode sequences from PacBio’s Multiplexing website, for example:
     - Sequel_16_Barcodes_v3.zip (contains a list of 16 PacBio barcodes for use with Barcoded Overhang Adapters)
2. Import the desired FASTA file into SMRT Link.
   i. On the SMRT Link Home Page, select **Data Management**.
   ii. Click **Import Data** and follow the steps below:
   
   A. Specify whether to import data from the **SMRT Link Server**, or from a **Local File System**. *(Note: Only references and barcodes are available if you select Local File System.)*
   
   B. Select the data type to import: **Barcodes** – FASTA (.fa or .fasta), XML (.barcodeset.xml), or ZIP files containing barcodes.
   
   C. Navigate to the appropriate file and click **Import**. The selected barcode file is imported and becomes available for viewing in the SMRT Link Data Management module home screen.
SMRT LINK RUN DESIGN SETUP PROCEDURE FOR AUTOMATED DEMULTIPLEXING OF POLLED LOW DNA INPUT LIBRARY SAMPLES

- Open the Run Design module in SMRT Link and click New Run Design.
- Fill in the Sample Information section, then click the small arrow to open Barcoded Sample Options.
- Specify the following options:

1. Sample is Barcoded: Yes
2. Barcode Set: <Select Barcode File Name> (e.g., Sequel_16_barcodes_v3)
3. Same Barcodes on Both Ends of Sequence: Yes
4. Assign a Biological Sample Name to each barcoded sample using one of two ways: or From a CSV File or Interactively (SMRT Link v10.0 or higher only)
Barcode Selection and Bio Sample Name Specification Using a CSV File:

1. Click the **From a File** button, then click **Download File**.
2. Edit the file and enter the biological sample names associated with the barcodes in the second column, then save the file.
   - Delete entire rows of barcodes not used
   - Allowed characters: Alphanumeric; space; dot; underscore; hyphen. Other characters will be automatically removed.
3. Browse for the Barcoded Sample File you just edited and click on Open.
4. You see **Upload Was Successful** appear on the line below, assuming the file is formatted correctly.

- Refer to “Run Design” section in the **SMRT Link User Guide** for further details

<table>
<thead>
<tr>
<th>Barcode</th>
<th>Bio Sample Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>bc1001_BAK8A_OA–bc1001_BAK8A_OA</td>
<td>Low DNA Input Sample 1</td>
</tr>
<tr>
<td>bc1002_BAK8A_OA–bc1002_BAK8A_OA</td>
<td>Low DNA Input Sample 2</td>
</tr>
</tbody>
</table>
Interactive Method for Barcode Selection and Bio Sample Name Specification (SMRT Link v10.0 Only):

1. Click the **Interactively** button, then drag barcodes from the Available Barcodes column to the Included Barcodes column.
2. (Optional) Click a Bio Sample field to edit the Bio Sample Name associated with a barcode.
3. (Optional) Click Download as a file for later use.
4. Click Save to save the edited barcodes/bio sample names. You see Success on the line below, assuming the file is formatted correctly.
HiFi Sequencing Data Analysis Recommendations for *De Novo* Assembly Applications
HiFi sequencing data analysis recommendations for low DNA input de novo assembly applications

HiFi reads are compatible with third-party de novo assembly tools.

- Utilize SMRT Link to generate highly accurate and long single-molecule reads (HiFi reads) using the Circular Consensus Sequencing (CCS) analysis application or perform CCS analysis on-instrument using the Sequel Ile System.

- Output data in standard file formats, (BAM and FASTA/Q) for seamless integration with downstream analysis tools.

- 10- to 15-fold HiFi read coverage per haplotype is recommended for Low DNA Input de novo assembly projects.

- Can use PacBio’s SMRT Link Genome Assembly analysis application (powered by IPA) or other third-party software for de novo assembly analysis using HiFi reads:
  - Hifiasm
  - HiCanu

- Contact PacBio Technical Support (support@pacb.com) or your local Bioinformatics Field Applications Scientist for additional information about data analysis recommendations.
Low DNA Input Library Example Performance Data
EXAMPLE 1: NON-MULTIPLEXED BUTTERFLY LOW DNA INPUT LIBRARY (SINGLE-SAMPLE, SEQUEL II SYSTEM)

Sample Preparation and Sequencing Workflow

- **Single** female Red Admiral butterfly (*Vanessa atalanta*) collected
- HMW DNA (>40 kb) extraction using Qiagen MagAttract HMW Kit with modifications as described in [Kingan et al. (2019)](https://doi.org/10.1093/bioinformatics/btz431)
- Low DNA input SMRTbell library construction with 400 ng of input gDNA
- Singleplex sequencing on the **Sequel II System** using a 30-hour collection time

A single SMRT Cell 8M produced **13 Gb of HiFi data**, or ~35-fold HiFi coverage of the ~350 Mb genome.
EXAMPLE 1: NON-MULTIPLEXED BUTTERFLY LOW DNA INPUT LIBRARY (SINGLE-SAMPLE, SEQUEL II SYSTEM) (CONT.)

De Novo Assembly Performance Results

<table>
<thead>
<tr>
<th>METRIC</th>
<th>VALUE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assembly Length</td>
<td>368 Mb</td>
</tr>
<tr>
<td>N Contigs</td>
<td>80</td>
</tr>
<tr>
<td>Contig N50</td>
<td>10.13 Mb</td>
</tr>
<tr>
<td>BUSCO Complete</td>
<td>99.1%</td>
</tr>
<tr>
<td>BUSCO Duplicate</td>
<td>0.3%</td>
</tr>
</tbody>
</table>

HiFi assembly using FALCON; polishing with Racon; haplotype deduplication with Purge Dups

Single-contig chromosomes with resolved telomere sequences revealed by alignment to *H. melpomene*

EXAMPLE 2: MULTIPLEXED MOSQUITO LOW DNA INPUT LIBRARY (2-PLEX SAMPLE, SEQUEL II SYSTEM)

Sample Preparation and Sequencing Workflow

Two individual female mosquitos (Anopheles coluzzii) collected

HMW DNA (>40 kb) extraction using Qiagen MagAttract HMW Kit with modifications as described in Kingan et al. (2019)

Low DNA input SMRTbell library construction with 230 ng of input gDNA per mosquito sample

Multiplex 2 samples / SMRT Cell 8M for sequencing on the Sequel II System using a 30-hour collection time

A single SMRT Cell 8M produced >6 Gb of HiFi data per sample, or ~25-fold HiFi coverage of the ~250 Mb genome

---

Kigan, S.B. et al. (2020) Every species can be a model: Reference-quality PacBio genomes from single insects. PAG 2020 Poster.
EXAMPLE 2: MULTIPLEXED MOSQUITO LOW DNA INPUT LIBRARY (2-PLEX SAMPLE, SEQUEL II SYSTEM) (CONT.)

De Novo Assembly Performance Results

<table>
<thead>
<tr>
<th>METRIC</th>
<th>MOSQUITO 1</th>
<th>MOSQUITO 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assembly Length</td>
<td>262 Mb</td>
<td>259 Mb</td>
</tr>
<tr>
<td>N Contigs</td>
<td>465</td>
<td>358</td>
</tr>
<tr>
<td>Contig N50</td>
<td>5.3 Mb</td>
<td>2.9 Mb</td>
</tr>
<tr>
<td>BUSCO Complete</td>
<td>98.8%</td>
<td>98.8%</td>
</tr>
<tr>
<td>BUSCO Duplicate</td>
<td>0.1%</td>
<td>0.3%</td>
</tr>
</tbody>
</table>

HiFi assembly using FALCON; polishing with Racon; haplotype deduplication with Purge Dups

MOSQUITO SAMPLE 1 SUBREADS ASSEMBLED AND COMPARED TO HI-FI ASSEMBLY.

<table>
<thead>
<tr>
<th>METRIC</th>
<th>HI-FI READ ASSEMBLY</th>
<th>LONG READ (CLR) ASSEMBLY</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coverage</td>
<td>25-Fold</td>
<td>40-Fold</td>
</tr>
<tr>
<td>N50 Read Length (N5)</td>
<td>11 kb (19 kb)</td>
<td>12 kb (22 kb)</td>
</tr>
<tr>
<td>Primary Assembly Length</td>
<td>262 Mb</td>
<td>243 Mb</td>
</tr>
<tr>
<td>Primary Contig N50</td>
<td>5.28 Mb</td>
<td>3.86 Mb</td>
</tr>
<tr>
<td>Primary Contigs</td>
<td>465</td>
<td>212</td>
</tr>
<tr>
<td>BUSCO</td>
<td>C:98.7%, D:0.1%, F:0.6%, M:0.7%</td>
<td>C:98.7, D:0.2%, F:0.6%, M:0.7%</td>
</tr>
<tr>
<td>CPU Hours (Consensus + Accuracy)</td>
<td>1604</td>
<td>1947</td>
</tr>
</tbody>
</table>

MOSQUITO GENOME ASSEMBLY RESULTS.

- MOSQUITO 1
  - Assembly Length: 262 Mb
  - N Contigs: 465
  - Contig N50: 5.3 Mb
  - BUSCO Complete: 98.8%
  - BUSCO Duplicate: 0.1%

- MOSQUITO 2
  - Assembly Length: 259 Mb
  - N Contigs: 358
  - Contig N50: 2.9 Mb
  - BUSCO Complete: 98.8%
  - BUSCO Duplicate: 0.3%

HiFi Assemblies Capture Satellites and Other Repeats

- 9 Mb of HiFi Read assembly does not map to Long Read assembly
- Primarily map to “UNKN” (96%) or sex chromosomes (3% Y, 1% X)
- A known satellite repeat (AgX367, L = 367 bp) maps across contig

FALCON + Purge-Dups

- Coverage: 25-Fold
- N50 Read Length (N5): 11 kb (19 kb)
- Primary Assembly Length: 262 Mb
- Primary Contig N50: 5.28 Mb
- Primary Contigs: 465
- BUSCO: C:98.7%, D:0.1%, F:0.6%, M:0.7%
- CPU Hours (Consensus + Accuracy): 1604

Assembly is Fast with HiFi Reads

HiFi assemblies using FALCON; polishing with Racon; haplotype deduplication with Purge Dups

Kingan, S.B. et al. (2020) Every species can be a model: Reference-quality PacBio genomes from single insects. PAG 2020 Poster.
EXAMPLE 3: MULTIPLEXED MOTH LOW DNA INPUT LIBRARIES (2-PLEX SAMPLE, SEQUEL II SYSTEM)

Sample Preparation and Sequencing Workflow

- Ten individual moth samples collected
- HMW DNA (>40 kb) extraction using Qiagen MagAttract HMW Kit with modifications as described in Kingan et al. (2019)
- Low DNA input SMRTbell library construction with 400 ng of input gDNA per moth sample
- Multiplex 2 samples / SMRT Cell 8M for sequencing on the Sequel II System using a 30-hour collection time

<table>
<thead>
<tr>
<th>MOTH SPECIES</th>
<th>COMMON NAME</th>
</tr>
</thead>
<tbody>
<tr>
<td>Autographa gamma</td>
<td>Silver Y</td>
</tr>
<tr>
<td>Cosmia trapezina</td>
<td>Dun-bar</td>
</tr>
<tr>
<td>Craniophora ligustri</td>
<td>Coronet</td>
</tr>
<tr>
<td>Endotricha flammealis</td>
<td>Rose-flounced tabby</td>
</tr>
<tr>
<td>Lymantria monacha</td>
<td>Black-arched tussock</td>
</tr>
<tr>
<td>Noctua fimbriata</td>
<td>Broad-bordered yellow underwing</td>
</tr>
<tr>
<td>Notocelia uddmanniana</td>
<td>Bramble shoot</td>
</tr>
<tr>
<td>Parapoynx stratiotata</td>
<td>Ringed china-mark</td>
</tr>
<tr>
<td>Recurvaria leucatella</td>
<td>White-barred groundling</td>
</tr>
<tr>
<td>Thyatira batis</td>
<td>Peach blossom</td>
</tr>
</tbody>
</table>
**EXAMPLE 3: MULTIPLEXED MOTH LOW DNA INPUT LIBRARIES**
(2-PLEX SAMPLE, SEQUEL II SYSTEM) (CONT.)

*De Novo Assembly Performance Results*

**MOTH GENOME ASSEMBLY RESULTS.**

<table>
<thead>
<tr>
<th>SAMPLE</th>
<th>ASSEMBLY SIZE (Mb)</th>
<th># CONTIGS</th>
<th>CONTIG N50 (Mb)</th>
<th>BUSCO COMPLETE (ARTHROPODA)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Autographa gamma</td>
<td>368</td>
<td>79</td>
<td>12.01</td>
<td>99.2%</td>
</tr>
<tr>
<td>Cosmia trapezina</td>
<td>862</td>
<td>1924</td>
<td>0.88</td>
<td>97.5%</td>
</tr>
<tr>
<td>Craniophora ligustri</td>
<td>436</td>
<td>267</td>
<td>3.50</td>
<td>98.9%</td>
</tr>
<tr>
<td>Endotricha flammealis</td>
<td>489</td>
<td>489</td>
<td>1.92</td>
<td>99.2%</td>
</tr>
<tr>
<td>Lymantria monacha</td>
<td>912</td>
<td>301</td>
<td>5.77</td>
<td>99.2%</td>
</tr>
<tr>
<td>Noctua fimbriata</td>
<td>572</td>
<td>307</td>
<td>3.75</td>
<td>98.9%</td>
</tr>
<tr>
<td>Notocelia uddmanniana</td>
<td>814</td>
<td>809</td>
<td>2.07</td>
<td>98.3%</td>
</tr>
<tr>
<td>Parapoynx stratiotata</td>
<td>481</td>
<td>129</td>
<td>6.92</td>
<td>99.4%</td>
</tr>
<tr>
<td>Recurvaria leucatella</td>
<td>746</td>
<td>1247</td>
<td>0.98</td>
<td>98.2%</td>
</tr>
<tr>
<td>Thyatira batis</td>
<td>316</td>
<td>88</td>
<td>7.15</td>
<td>98.9%</td>
</tr>
</tbody>
</table>

HiFi assembly using **Improved Phased Assembler (IPA)**; haplotype deduplication with Purge Dups.

Assembly results using **Improved Phased Assembler (IPA)** show contig N50 ≥1 Mb for most samples analyzed.
Technical Documentation & Applications Support Resources
Sample Preparation Literature

- Application Note: Considerations for Using the Low and Ultra-low DNA Input Workflows for Whole Genome Sequencing (PN 101-995-900)
- Procedure & Checklist - Procedure & Checklist - Preparing HiFi Libraries from Low DNA Input Using SMRTbell Express Template Prep Kit 2.0 (PN 101-730-400)
- Quick Reference Card – Loading and Pre-extension Recommendations for the Sequel System (PN 101-461-600)
- Quick Reference Card – Loading and Pre-extension Recommendations for the Sequel II/Ile Systems (PN 101-769-100)
- Overview – Sequel Systems Application Options and Sequencing Recommendations (PN 101-851-300)
- Application Consumable Bundles Purchasing Guide (PN PG100-051320)
- Technical Note: Preparing samples for PacBio whole genome sequencing for de novo assembly – Collection and storage (PN TN100-040518)
- Technical Note: Preparing DNA for PacBio HiFi sequencing – Extraction and quality control (PN TN101-061920)
- Technical Overview: Low DNA Input Library Preparation Using SMRTbell Express Template Prep Kit 2.0 (PN 101-781-000)
Data Analysis Resources

- **SMRT Analysis Barcoding Overview (v9.0)** (PN 101-923-200)
  - Contains detailed information on barcoding experimental design options and describes QC metrics for evaluation of barcoding performance using SMRT Link

- PacBio [Multiplexing Resources](https://www.pacb.com/products-and-services/analytical-software/multiplexing/)
  - Barcoding Overview documents for different SMRT Link software versions
  - PacBio barcode sequence files (compressed FASTA) for use with Sequel, Sequel II and Sequel IIe Systems
  - Barcoded oligo ordering sheets

Example PacBio Data Sets

<table>
<thead>
<tr>
<th>WHOLE GENOME SEQUENCING APPLICATION</th>
<th>DATASET</th>
<th>DATA TYPE</th>
<th>PACBIO SYSTEM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assembly (Low DNA Input)</td>
<td><em>Anopheles gambiae</em> (African malaria mosquito) – 2plex</td>
<td>HiFi Reads</td>
<td>Sequel II System</td>
</tr>
</tbody>
</table>
Posters
- PacBio PAG 2020 Poster: Every species can be a model: Reference-quality PacBio genomes from single insects
- PacBio ABRF 2020 Poster: A low DNA input protocol for high-quality PacBio de novo genome assemblies
- PacBio AGBT 2020 Poster: New advances in SMRT Sequencing facilitate multiplexing for de novo and structural variant studies

Publications

Webinars
- PacBio Virtual Global Summit Presentation (2020): Anopheles and the journey from a reference genome to a thousand genomes to spatiotemporal genomic observatories. [Webinar Recording]
Appendix: General Recommendations for High-Molecular Weight gDNA QC and Handling for SMRTbell Library Construction
SAMPLE COLLECTION, PREPARATION, AND STORAGE FOR SMRT SEQUENCING WHOLE GENOME DE NOVO ASSEMBLY PROJECTS

To obtain the highest quality genomic DNA, it is important to start with sample types compatible with high molecular weight (HMW) DNA extraction methods

- **Technical Note: Sample Prep** ([TN100-040518](#)) provides general guidance on biological sample collection, preparation, and storage across a range of commonly encountered sample types used for SMRT Sequencing whole genome projects

  - Includes sample and storage recommendations for:
    - Vertebrates - mammals, birds, fish, amphibians, reptiles
    - Invertebrates - marine, terrestrial
    - Arthropods - insects, crustaceans
    - Fungi - microorganisms, mushrooms, algae*
    - Plants - broad leaf plants, grasses

  *Algae is included with fungi due to similar growth and storage conditions

- This technical note also includes additional considerations for planning HMW DNA isolation
DNA EXTRACTION AND QC FOR SMRT SEQUENCING WHOLE GENOME DE NOVO ASSEMBLY PROJECTS

Starting with high-quality, high molecular weight (HMW) genomic DNA (gDNA) will result in longer libraries and better performance during sequencing.

- **Technical Note:** DNA Prep ([TN101-061920](https://example.com)) is intended to give recommendations, tips and tricks for the extraction of genomic DNA, as well as assessing and preserving the quality and size of your DNA sample to be used for PacBio HiFi sequencing for *de novo* assembly.

- Topics covered include:
  - DNA Extraction
    - Commercially available kits across a wide variety of input sample types
    - Resource for alternative DNA extraction methods
  - DNA Quality Control (QC)
    - DNA quantification, purity, size, and damage
    - Use of nucleic acid stabilizers
    - DNA storage and shipping
  - Best Practices for DNA Extraction for PacBio Sequencing

- This technical note also includes an example dataset for a California Redwood tree DNA sample that was isolated using a commercial DNA extraction kit.
DNA SAMPLE PREPARATION ONLINE RESOURCE

Literature resource for sample collection and DNA extraction protocol references

PacBio does not assume responsibilities/guarantees for these external publications/protocols, but we are happy to help as best as we can to guide / connect. Please contact ExtractDNA@pacb.com for more discussions around your particular species & sequencing project!
RECOMMENDED EXTRACTION KIT FOR ISOLATION OF HMW GENOMIC DNA FOR LOW DNA INPUT LIBRARY PREPARATION

- QIAGEN MagAttract HMW DNA Kit
  - Cat. No/ID: 67563

- MagAttract Magnetic Rack
  - Cat. No/ID: 19606

- MagAttract Protocol Versions
  - “Standard” protocol (as per QIAGEN instructions)
  - 10X Genomics-modified protocol (pp. 6-8 in “Chromium™ Genome Reagent Kits v2 User Guide”)
EXAMPLE RESULTS USING “STANDARD” QIAGEN MAG ATTRACT EXTRACTION PROTOCOL FOR MOSQUITO GENOMIC DNA EXTRACTION

- Figure shows typical size distribution profiles obtained for gDNA samples extracted from fresh or freshly killed and stored at -80°C mosquitos
- Recommend preparing multiple DNA extractions for evaluation of prior to library construction (Note: Majority of gDNA must be >30 kb and <100 kb)
- Typically obtain ~200 ng of HMW DNA per mosquito sample

METHODS FOR EVALUATION OF GENOMIC DNA QUALITY

Starting with high-quality genomic DNA will result in longer libraries and better *de novo* assembly performance.

- Input genomic DNA must be carefully QC’d to assess integrity.
- PFGE/FIGE or Femto Pulse sizing tool is highly recommended.
- High-quality, high-molecular weight DNA $\rightarrow$ longer read lengths.
- Low-quality, degraded/damaged DNA $\rightarrow$ shorter read lengths, lower library synthesis yields (dependent on size selection method and parameter settings employed).
- DNA purity can be determined by using a NanoDrop instrument or other spectrophotometer device.
- PacBio highly recommends using the Qubit High Sensitivity fluorometric assay for accurate dsDNA quantitation.
A. Genomic DNA Sizing Characterization

Recommended methods for determining gDNA size distribution:

PippinPulse System (Sage Science)

- Resolves up to ~80 kb
- Requires ≥50 ng of sample
- ~16-hour analysis time

Femto Pulse System (Agilent Technologies)

- Highly Recommended
- Resolves up to ~165 kb
- Requires <1 ng of sample
- <1.5-hour analysis time

CHEF Mapper XA System (Bio-Rad)

- Resolves up to ~10 Mb
- Requires ≥100 ng of sample
- ~16-h analysis time

Evaluation of gDNA quality using A) Bio-Rad CHEF Mapper System and B) Femto Pulse System. Lanes 3A and 1B are examples of high quality, high-molecular weight genomic DNA. Lanes 4A and 2B are examples of degraded gDNA.
B. DNA Purity Determination

- DNA purity can be determined by using a NanoDrop instrument or other spectrophotometers
- For ultrapure gDNA, A260/280 ratio is typically between ~1.8 - 2.0 and A260/230 ratio is ≥2.0
- If A260/280 and A260/230 readings are out of the range specified above, PacBio recommends performing an AMPure PB bead purification step followed by re-assessment of quantity and purity of the gDNA sample

260/280 Ratio

- A low A260/A280 ratio may indicate the presence of protein, phenol, or other contaminants that absorb strongly at or near 280 nm. Sometimes it may be caused by a very low concentration of nucleic acid.
- High 260/280 ratios are not indicative of an issue

260/230 Ratio

- A low A260/A230 ratio may be the result of:
  - Carbohydrate carryover (often a problem with plants)
  - Residual phenol from nucleic acid extraction
  - Residual guanidine (often used in column-based kits)
  - Glycogen used for precipitation
- A high A260/A230 ratio may be the result of:
  - Making a blank measurement on a dirty pedestal of a Nanodrop instrument
  - Using an inappropriate solution for the blank measurement
C. DNA Quantification

- Accurate quantitation of DNA concentration is critical for PacBio template preparation procedures.
  - Specifically, it is critical to determine the concentration of the double-stranded DNA, since only double-stranded DNA will be converted into sequencing templates.

- PacBio highly recommends using a Qubit fluorometer tool and Qubit dsDNA High Sensitivity (HS) Assay Kit (Thermo Fisher Scientific) for routine DNA quantitation during SMRTbell library construction.

- When assessing gDNA QC, PacBio recommends using both fluorometric and spectrophotometric methods – for example, using both the Qubit and NanoDrop instruments
  - If the sample is pure gDNA, free of any RNA contaminants and other small molecules, the two methods should converge to similar DNA concentration measurement values.

- If the measured NanoDrop concentration is significantly different (>50%) from the Qubit measurement, PacBio recommends doing an AMPure PB bead purification step (as specified by your chosen library preparation protocol), followed by a re-measurement with both methods. Typically, a single AMPure PB bead purification step resolves the discrepancy.
  - If the agreement does not improve after three rounds of AMPure PB bead purification, try using either a commercial kit, isopropanol precipitation, or a new DNA extraction method to obtain a cleaner DNA sample.