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

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

1. Ultra-Low DNA Input Workflow Overview
2. Ultra-Low DNA Input Sample QC Requirements
3. Ultra-Low DNA Input Library Sample Preparation Workflow Details
4. Ultra-Low DNA Input Library Sequencing Workflow Details
5. HiFi Sequencing Data Analysis Recommendations for Ultra-Low DNA Input Libraries
6. Ultra-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
Ultra-Low DNA Input Workflow Overview
ULTRA-LOW DNA INPUT SEQUENCING: HOW TO GET STARTED

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 SMRTbell Libraries from Ultra-Low DNA Input (101-987-800)

Technical documentation containing sample library construction and sequencing preparation protocol details

PacBio Application Consumable Bundle Purchasing Guide (PN PG100-051320)

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

Library Construction, Sequencing & Analysis

gDNA QC & Shearing

10 kb Target DNA Shear Size

Sample Amplification & Library Construction

SMRTbell gDNA Amplification Kit

SMRTbell Express TPK 2.0

HiFi Sequencing (Sequel II and IIe Systems)

Aim for >30-fold HiFi Read Coverage per Diploid Sample for Assembly

Aim for ≥15-fold HiFi Read Coverage for Human Variant Detection

Data Analysis

Genome Assembly

Variant Detection

* 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 ULTRA-LOW DNA INPUT WORKFLOW?

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.

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.

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.
Ultra-Low DNA Input: SUPPORTED APPLICATIONS

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

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

Ultra-Low DNA Input: UNSUPPORTED APPLICATIONS

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

- **COMPLEX POPULATIONS**
  - Metagenomics sequencing

---

**SAMPLE & PROJECT CONSIDERATIONS**

<table>
<thead>
<tr>
<th></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 Human Variant Detection</td>
</tr>
</tbody>
</table>

**WHEN IS IT APPROPRIATE TO CONSIDER THE ULTRA-LOW DNA INPUT WORKFLOW? (CONT.)**

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

- **Procedure & Checklist - Preparing HiFi Libraries from Ultra-Low DNA Input** (PN 101-995-900) protocol document describes how to prepare SMRTbell libraries with ultra-low DNA input amounts for sequencing on the **Sequel II and Ile Systems** for WGS *de novo* assembly and variant detection applications using HiFi reads.

- With this procedure, HiFi libraries can be constructed from as little as 5 ng of input genomic DNA (gDNA) using SMRTbell Express Template Prep Kit 2.0.

- Genomic DNA is sheared to approximately 10 kb using a g-TUBE device or a Megaruptor system, amplified by PCR, constructed to a SMRTbell library and size-selected using the BluePippin system.
  - For *de novo* assembly of small insect/arthropod genomes, a minimum of 500 ng of amplified gDNA is required to generate sufficient SMRTbell library to run 1 Sequel II SMRT Cell 8M.
  - For variant detection of human genomes, a minimum of 800 ng of amplified gDNA is required to generate sufficient SMRTbell library to run 2 Sequel II SMRT Cells 8M.

**APPLICATIONS**

**WHOLE GENOME SEQUENCING**

Ultra-Low DNA Input for *De Novo* Assembly and Variant Detection Applications

https://www.pacb.com/support/documentation/
ULTRA-LOW DNA INPUT LIBRARY SAMPLE PREPARATION & SEQUENCING WORKFLOW

Workflow summary for constructing ultra-low DNA input SMRTbell libraries suitable for HiFi sequencing on the Sequel II and IIe Systems for small insect *de novo* assembly and human variant detection applications

**Genomic DNA QC & Shearing**
- Shear ≥5 ng of gDNA to 10 kb target size

**Sample Amplification (5.5 hrs)**
- Ligate PCR adapters and perform two complementary PCR reactions to amplify sheared gDNA

**SMRTbell Library Construction (4.5 hrs)**
- Procedure & Checklist – Preparing HiFi Libraries from Ultra-Low DNA Input (PN 101-987-800)

**SMRTbell Library Size Selection**
- BluePippin size selection (Collect 8 – 17 kb fraction)

**HiFi Sequencing (Sequel II and IIe Systems)**
- Aim for >30-fold HiFi read coverage per diploid sample for *de novo* assembly
- Aim for ≥15-fold HiFi read coverage for human variant detection

**HiFi Data Analysis**
- For variant detection, can use DeepVariant for small variants <50 bp and SMRT Link PBSV for larger variants >50 bp
- For *de novo* assembly, can use SMRT Link Genome Assembly or other third-party software

PacBio HiFi reads achieve >99.9% accuracy
Ultra-Low DNA Input Library Sample QC Requirements
Table below summarizes the required DNA quality and quantity for processing samples using the ultra-low DNA input workflow.

### DNA QUALITY AND QUANTITY REQUIREMENTS FOR ULTRA-LOW DNA INPUT SAMPLES RUN ON THE SEQUEL II AND Ile SYSTEMS.

<table>
<thead>
<tr>
<th>ULTRA-LOW DNA INPUT LIBRARY WORKFLOW (SEQUEL II AND Ile SYSTEMS)</th>
<th>RECOMMENDED INPUT gDNA AMOUNT</th>
<th>REQUIRED INPUT gDNA QUALITY</th>
<th>gDNA SHEARING METHOD</th>
<th>TARGET SHEARED gDNA SIZE DISTRIBUTION MODE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Supports up to a 500 Mb genome size for <em>de novo</em> assembly or up to a 3 Gb genome size for variant detection applications*</td>
<td>5 – 20 ng</td>
<td>Majority of gDNA &gt;20 kb</td>
<td>Megaruptor System or g-TUBE</td>
<td>10 kb sheared gDNA fragment size is optimal</td>
</tr>
</tbody>
</table>

* Multiplexed sample preparation is currently **unsupported** for the ultra-low DNA input procedure.
RECOMMENDED TOOLS FOR GENOMIC DNA QUANTIFICATION AND QUALIFICATION

When working with small amounts of input DNA, accurate sizing and quantification is critical to enable generation of sufficient coverage of long reads to produce a high-quality genome assembly.

DNA Quantification

- For quantification of gDNA to be used with the ultra-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**
- Commercially available systems that may be used to evaluate gDNA size distribution are listed below with links to the corresponding third-party vendor website

<table>
<thead>
<tr>
<th>GENOMIC DNA SIZE EVALUATION METHODS AND PROCEDURES.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>DNA SIZING METHOD</strong></td>
</tr>
</tbody>
</table>
| Femto Pulse System (Agilent) | ▪ **Highly recommended**  
▪ Requires 200 – 500 pg | Agilent Femto Pulse Website |
| Fragment Analyzer System (Agilent) | ▪ Requires 2 ng | Agilent Fragment Analyzer Website |

We highly recommend the use of the **Femto Pulse System (Agilent)** for ultra-low DNA input applications because of its ability to evaluate size distributions using only ~200 – 500 picograms of DNA.
EVALUATION OF GENOMIC DNA FOR ULTRA-LOW DNA INPUT LIBRARY CONSTRUCTION

The size distribution of the starting input gDNA sample is critical to successful HiFi SMRTbell library construction and sequencing on the Sequel II and IIe Systems. Always evaluate the quality of the gDNA samples before proceeding with library construction.

- PacBio recommends working with high-quality samples where the majority of the input gDNA is >20 kb
- Use of low-quality gDNA samples in the ultra-low DNA input procedure will negatively impact the PCR amplification step
  → If the starting gDNA sample contains a high amount of short fragments (<10 kb), then these will become preferentially amplified during PCR – and thus lead to mostly short-insert read length sequencing data
EVALUATION OF GENOMIC DNA FOR ULTRA-LOW DNA INPUT LIBRARY CONSTRUCTION (CONT.)

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

- **Sample 1** is not suitable for using the ultra-low DNA input procedure.
  - The gDNA is severely fragmented such that a significant proportion of the fragments are ≤10 kb resulting in preferential amplification of short fragments. We recommend re-extraction of the gDNA to obtain a higher-quality sample for shearing and SMRTbell library construction.

- **Samples 2** shows a size distribution with the majority of the fragments >20 kb.
  - This sample is appropriate for shearing and constructing a HiFi SMRTbell library for *de novo* assembly.

Example Femto Pulse sizing QC analysis of individual mosquito gDNA samples of varying quality. gDNA samples with a majority of fragments >30 kb are suitable for HiFi library construction using this procedure. Sample 1 is lower quality gDNA such that a significant proportion of the fragments are ≤10 and is not suitable constructing SMRTbell libraries using the ultra-low DNA input procedure. Sample 2 shows size distribution with the majority of the fragments >20 kb and are appropriate for shearing and constructing SMRTbell libraries.
Ultra-Low DNA Input Library Sample Preparation Workflow Details
PROCEDURE & CHECKLIST – PREPARING HIFI LIBRARIES FROM ULTRA-LOW DNA INPUT USING SMRTBELL EXPRESS TEMPLATE PREP KIT 2.0

- This document (PN 101-987-800) protocol document describes how to prepare SMRTbell libraries: HiFi SMRTbell libraries from ≥5 ng of input genomic DNA (gDNA) for sequencing on the Sequel II and Ile Systems.

- Genomic DNA is sheared to approximately 10 kb using a g-TUBE or a Megaruptor system, amplified by PCR, constructed to a SMRTbell library and size-selected using the BluePippin system.

- This workflow enables *de novo* assembly of insect genomes of up to 500 Mb (>500 Mb genome size is not supported) and human variant detection from as low as 5 ng of gDNA.

- Protocol document contains:
  1. General laboratory best practices recommendations
  2. Guidelines for QC evaluation of gDNA samples for ultra-low DNA Input library construction
  3. Enzymatic steps for preparation ultra-low DNA input SMRTbell libraries
  4. Instructions for size-selection of ultra-low DNA input libraries using the BluePippin System
  5. Sample setup guidance for preparing ultra-low DNA input libraries for sequencing on the Sequel II and Ile Systems

---

https://www.pacb.com/support/documentation
ULTRA-LOW DNA INPUT WORKFLOW DETAILED OVERVIEW

1. Genomic DNA QC and Shearing
   - Recommended starting input gDNA quality: >20 kb
   - Shear gDNA using a Megaruptor System or g-TUBE device
     ▪ Minimum required input gDNA mass for shearing: 5 ng
     ▪ Target DNA shear size: 10 kb

2. DNA Repair & PCR Adapter Ligation (3 h)
   - Perform DNA Damage Repair / End Repair / A-Tailing of sheared gDNA
   - Ligate linear PCR adapters to repaired gDNA

3. Sample Amplification & Pooling (2.5 h)
   - Perform two complementary PCR reactions (PCR Reaction Mix 1 and PCR Reaction Mix 2) in parallel to amplify the gDNA sample using SMRTbell gDNA Sample Amplification Kit (PN 101-980-000)

   ▪ For de novo assembly of small insect/arthropod genomes, a minimum of 500 ng of pooled amplified gDNA is required to generate sufficient SMRTbell library to run 1 Sequel II SMRT Cell 8M.
   ▪ For variant detection of human genomes, a minimum of 800 ng of pooled amplified gDNA is required to generate sufficient SMRTbell library to run 2 Sequel II SMRT Cells 8M.
4. SMRTbell Express TPK 2.0 Library Construction (3 h)
   - Perform DNA Damage Repair / End Repair / A-Tailing of amplified gDNA
   - Typical library construction yield (pre-size selection): >70%

5. SMRTbell Library Size Selection (4 h)
   - Perform BluePippin size selection (8 – 17 kb fraction collection)
   - Typical recovery yield of size-selected library: >20 – 30%
   - The average size distribution of the final size-selected library is approximately 10 – 11 kb.

6. Sequencing Preparation (Sequel II and Ile Systems) (2.5 h)
   - Perform sequencing primer annealing, polymerase binding and complex cleanup
   - Movie collection time: 30 hours

7. Data Analysis
   - Perform CCS analysis, trim PCR adapter sequences and remove duplicate PCR reads using SMRT Link or the command line
   - 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)
   - For HiFi variant detection, can use SMRT Link Structural Variant Calling analysis application (powered by PBSV) to call structural variants and Google DeepVariant to call small variants

* Read lengths, reads/data per SMRT Cell 8M and other sequencing performance results vary based on sample quality/type and insert size.
# 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 (Highly Recommended)</td>
<td>Agilent</td>
<td>M5330AA</td>
</tr>
<tr>
<td>Fragment Analyzer</td>
<td>Agilent</td>
<td>M5310AA</td>
</tr>
<tr>
<td>DNA Quantitation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Qubit Fluorometer</td>
<td>Thermo Fisher Scientific</td>
<td>Q33226</td>
</tr>
<tr>
<td>Qubit 1X dsDNA HS Assay Kit</td>
<td>Thermo Fisher Scientific</td>
<td>Q33230</td>
</tr>
<tr>
<td>DNA Shearing</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Megaruptor 1 or Megaruptor 2 System</td>
<td>Diagenode</td>
<td>B06010001 / B06010002</td>
</tr>
<tr>
<td>Long Hydropores</td>
<td>Diagenode</td>
<td>E07010002</td>
</tr>
<tr>
<td>Hydrotubes</td>
<td>Diagenode</td>
<td>C30010018</td>
</tr>
<tr>
<td>g-TUBE</td>
<td>Covaris</td>
<td>520104</td>
</tr>
<tr>
<td>Eppendorf MiniSpin Plus or other equivalent benchtop centrifuge model</td>
<td>Eppendorf</td>
<td>22620100</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>SMRTbell gDNA Sample Amplification Kit</td>
<td>PacBio</td>
<td>101-980-000</td>
</tr>
<tr>
<td>Elution Buffer</td>
<td>PacBio</td>
<td>101-633-500</td>
</tr>
<tr>
<td>HDPE 8 place Magnetic Separation Rack for 0.2 ml PCR Tubes</td>
<td>V&amp;P Scientific Inc.</td>
<td>VP772F4-1</td>
</tr>
<tr>
<td>0.2 mL PCR 8-strip tubes</td>
<td>USA Scientific</td>
<td>1402-4708</td>
</tr>
<tr>
<td>Wide Orifice Tips (Tips LTS W-O 200UL Fltr RT-L200WFLR)</td>
<td>Rainin</td>
<td>17014294</td>
</tr>
<tr>
<td>ProNex Beads</td>
<td>Promega</td>
<td>NG2002 – 125 mL</td>
</tr>
<tr>
<td>ITEM</td>
<td>VENDOR</td>
<td>PART NUMBER</td>
</tr>
<tr>
<td>-----------------------------------------------------------</td>
<td>---------------------------------------------</td>
<td>-------------</td>
</tr>
<tr>
<td><strong>SMRTbell Library Preparation (Cont.)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100% Ethanol, Molecular Biology Grade</td>
<td>Any Major Life Science Supply Vendor</td>
<td>Vendor-specific</td>
</tr>
<tr>
<td>Thermal Cycler that is 100 μL and 8-tube strip compatible</td>
<td>Any Major Life Science Supply Vendor</td>
<td>Vendor-specific</td>
</tr>
<tr>
<td><strong>SMRTbell Library Size Selection</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BluePippin Size-Selection System</td>
<td>Sage Science</td>
<td>BLU0001</td>
</tr>
<tr>
<td>BluePippin with dye free, 0.75% Agarose Cassettes and S1 Marker</td>
<td>Sage Science</td>
<td>BLF7510</td>
</tr>
</tbody>
</table>
SMRTBELL gDNA SAMPLE AMPLIFICATION KIT AND SMRTBELL EXPRESS TEMPLATE PREP KIT 2.0 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

<table>
<thead>
<tr>
<th>PACBIO KIT</th>
<th>REAGENT</th>
<th>WHERE USED</th>
</tr>
</thead>
<tbody>
<tr>
<td>SMRTbell gDNA Sample Amplification Kit (PN 101-980-000)</td>
<td>Amplification Adapters</td>
<td>Ligation</td>
</tr>
<tr>
<td></td>
<td>PCR Master Mix 1</td>
<td>Amplification</td>
</tr>
<tr>
<td></td>
<td>PCR Master Mix 2</td>
<td>Amplification</td>
</tr>
<tr>
<td></td>
<td>Sample Amplification PCR Primer</td>
<td>Amplification</td>
</tr>
<tr>
<td>SMRTbell Express Template Prep kit 2.0 (PN 100-938-900)</td>
<td>DNA Prep Additive</td>
<td>Remove Single-Strand Overhangs</td>
</tr>
<tr>
<td></td>
<td>DNA Prep Enzyme</td>
<td>Remove Single-Strand Overhangs</td>
</tr>
<tr>
<td></td>
<td>DNA Damage Repair Mix v2</td>
<td>DNA Damage Repair</td>
</tr>
<tr>
<td></td>
<td>End Prep Mix</td>
<td>End-Repair/A-tailing</td>
</tr>
<tr>
<td></td>
<td>Overhang Adapter v3</td>
<td>Ligation</td>
</tr>
<tr>
<td></td>
<td>Barcoded Overhang Adapters</td>
<td>Ligation</td>
</tr>
<tr>
<td></td>
<td>Ligation Mix</td>
<td>Ligation</td>
</tr>
<tr>
<td></td>
<td>Ligation Additive</td>
<td>Ligation</td>
</tr>
<tr>
<td></td>
<td>Ligation Enhancer</td>
<td>Ligation</td>
</tr>
</tbody>
</table>
SMRTBELL gDNA SAMPLE AMPLIFICATION KIT

DNA sample amplification step in the Ultra-Low DNA Input procedure is performed using the SMRTbell gDNA Sample Amplification Kit

- PacBio SMRTbell gDNA Sample Amplification Kit (PN 101-980-000)

- Kit contains enough reagents to support 9 – 18 ultra-low DNA input sample preparation reactions
  - Actual number of samples supported depends on whether any additional PCR re-amplification is needed for SMRTbell library construction

PACBIO SMRTBELL gDNA SAMPLE AMPLIFICATION KIT COMPONENTS.

<table>
<thead>
<tr>
<th>KIT COMPONENT</th>
<th>PURPOSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Duplex Buffer</td>
<td>Buffer for adapter ligation</td>
</tr>
<tr>
<td>Amplification Adapters</td>
<td>Adapters ligated to double-stranded fragments</td>
</tr>
<tr>
<td>Sample Amplification PCR Primer</td>
<td>Primer for amplification</td>
</tr>
<tr>
<td>PCR Master Mix 1</td>
<td>PCR Master mix 1</td>
</tr>
<tr>
<td>PCR Master Mix 2</td>
<td>PCR Master mix 2</td>
</tr>
</tbody>
</table>
BEST PRACTICES RECOMMENDATIONS FOR PREPARING ULTRA-LOW DNA INPUT SMRTBELL LIBRARIES

Sample Processing Recommendations

<table>
<thead>
<tr>
<th>PROTOCOL STEP</th>
<th>RECOMMENDATIONS</th>
<th>TARGET DNA SIZE OR AMOUNT</th>
</tr>
</thead>
<tbody>
<tr>
<td>DNA Input Quality</td>
<td>Freshly extracted high-quality genomic DNA</td>
<td>Majority of starting input gDNA &gt;20 kb</td>
</tr>
<tr>
<td>DNA Sizing QC</td>
<td>Femto Pulse System</td>
<td>Majority of starting input gDNA &gt;20 kb</td>
</tr>
<tr>
<td>DNA Shearing</td>
<td>g-TUBE (5 – 20 ng in 50 µl); or</td>
<td>10 kb sheared DNA size distribution</td>
</tr>
<tr>
<td></td>
<td>Megaruptor System (5 – 20 ng in 50 µl)</td>
<td></td>
</tr>
<tr>
<td>DNA Amplification</td>
<td>PCR Master Mix 1 (from SMRTbell gDNA Sample Amplification Kit); and</td>
<td>8 – 10 kb amplified DNA size distribution</td>
</tr>
<tr>
<td></td>
<td>PCR Master Mix 2 (from SMRTbell gDNA Sample Amplification Kit)</td>
<td>Amount of pooled amplified DNA required to proceed with SMRTbell library construction:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>≥ 500 ng for 1 SMRT Cell 8M</td>
</tr>
<tr>
<td></td>
<td></td>
<td>≥ 800 ng for 2 SMRT Cells 8M</td>
</tr>
<tr>
<td>SMRTbell Library Construction</td>
<td>SMRTbell Express Template Prep Kit 2.0</td>
<td>8 – 10 kb SMRTbell library size distribution</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Amount of SMRTbell library required to proceed with BluePippin size selection: ≥400 ng</td>
</tr>
<tr>
<td>Library Size Selection</td>
<td>BluePippin System (0.75% DF 3-10 kb Marker S1 Improved Recovery cassette definition; 8 – 17 kb fraction collection)</td>
<td>Final size-selected SMRTbell library size distribution: 11 kb – 12 kb</td>
</tr>
</tbody>
</table>
BEST PRACTICES RECOMMENDATIONS FOR PREPARING ULTRA-LOW DNA INPUT SMRTBELL LIBRARIES (CONT.)

General Equipment Setup & Reagent Handling Recommendations

1. Use wide-bore tips for all mixing steps.
2. Never vortex tubes containing high-molecular weight genomic DNA samples.
3. Minimize the number of freeze/thaw cycles the gDNA undergoes to reduce DNA damage.
4. Allowing sufficient time for thawing aliquots of DNA, as partially frozen DNA is prone to shearing.
5. Always set your heat blocks or thermocyclers to the appropriate temperature for incubations before proceeding with the procedure.
6. Ensure that ProNex Beads are at room temperature prior to performing the purification steps.
7. When performing ProNex 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.
8. 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 ULTRA-LOW DNA INPUT LIBRARY CONSTRUCTION

The success of this procedure is highly dependent on the size distribution of the sheared DNA

- The Megaruptor System (Diagenode) or g-TUBE device (Covaris) is recommended for shearing gDNA to a **target size distribution mode ~10 kb** for ultra-low DNA input applications.

- Over-shearing gDNA may impact amplification and the yield of the final size-selected SMRTbell library.
  - Therefore, it is critical to work with samples where the majority of the starting input genomic DNA is >20 kb and optimize the shearing conditions to obtain the target DNA fragment size.

- After shearing the gDNA samples, evaluate the size distribution using a Femto Pulse or Fragment Analyzer system.

### RECOMMENDED INPUT GENOMIC DNA AMOUNTS AND TARGET DNA SHEAR SIZE FOR ULTRA-LOW DNA INPUT SAMPLES.

<table>
<thead>
<tr>
<th>ULTRA-LOW DNA INPUT LIBRARY RECOMMENDATIONS</th>
<th>MEGARUPTOR SYSTEM OR g-TUBE DEVICE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recommended Starting Input gDNA Amount for Shearing</td>
<td>5 – 20 ng</td>
</tr>
<tr>
<td>Target gDNA Shear Size</td>
<td>10 kb</td>
</tr>
</tbody>
</table>

For shearing, it is important to note that **5 ng is the minimum starting DNA input**. If you have more than 5 ng DNA, we recommend shearing more than 5 ng of input DNA to ensure sufficient sample to perform accurate DNA sizing and quantification.
DNA SHEARING RECOMMENDATIONS FOR ULTRA-LOW DNA INPUT LIBRARY CONSTRUCTION (CONT.)

Example Femto Pulse sizing QC analysis of gDNA samples sheared to 10 kb using a Covaris g-TUBE device.

- If necessary, optimize the shearing conditions to obtain the optimal DNA size.

Example Femto Pulse sizing QC analysis of input gDNA and sheared gDNA samples. 5 ng of human genomic DNA was sheared to approximately 10 kb using a Covaris g-TUBE (1677 x g in a Eppendorf MiniSpin Plus centrifuge). The size distribution of the starting input genomic DNA is >20 kb.
Example Femto Pulse sizing QC analysis of gDNA samples sheared to 10 kb using a Megaruptor 1 System.

- **Important**: If using the Megaruptor 1 or Megaruptor 2 System, ensure that the system is thoroughly washed before shearing. Since this workflow requires amplification, contamination introduced during shearing will be amplified and may complicate downstream analysis.

---

**Genomic DNA**

20 kb

**Sheared DNA**

9.7 kb

---

Example Femto Pulse sizing QC analysis of input gDNA and sheared gDNA samples. 5 ng of human genomic DNA was sheared to approximately 10 kb using the Megaruptor 1 System (Long Hydropores with the Pre-Load Hydropores setting turned OFF). The size distribution of the starting input genomic DNA is >20 kb.
gDNA SAMPLE AMPLIFICATION BY PCR

The Ultra-Low DNA Input procedure utilizes PCR amplification to increase the amount of DNA material available for SMRTbell library construction

PCR Reaction Mix Preparation

- After the PCR adapter ligation step, the purified eluted sample (97 µL) is divided and used in two complementary amplification reactions (Reaction Mix 5A and Reaction Mix 5B) run in parallel
  - Sample Amplification PCR Primers, PCR Master Mix 1 and PCR Master Mix 2 are found in the SMRTbell gDNA Sample Amplification Kit (PN 101-980-000)
  - PCR Master Mix 1 and PCR Master Mix 2 have been optimized to enable amplification across a broad range of genomic regions with different %GC composition
- When working with multiple samples, prepare enough Reaction Mix for all reactions, plus 10% of the total reaction mix volume
- Pipette mix the prepared Reaction Mixes 10 times with wide-bore pipette tips and then perform a quick spin to collect all liquid from the sides of the tube

### Reaction Mix 5A

<table>
<thead>
<tr>
<th>Component</th>
<th>Tube Cap Color</th>
<th>Volume</th>
</tr>
</thead>
<tbody>
<tr>
<td>PCR Master Mix 1</td>
<td></td>
<td>50.0 µL</td>
</tr>
<tr>
<td>Sample Amplification PCR Primer</td>
<td></td>
<td>2.0 µL</td>
</tr>
<tr>
<td>Total Volume</td>
<td></td>
<td>52.0 µL</td>
</tr>
</tbody>
</table>

### Reaction Mix 5B

<table>
<thead>
<tr>
<th>Component</th>
<th>Tube Cap Color</th>
<th>Volume</th>
</tr>
</thead>
<tbody>
<tr>
<td>PCR Master Mix 2</td>
<td></td>
<td>50.0 µL</td>
</tr>
<tr>
<td>Sample Amplification PCR Primer</td>
<td></td>
<td>2.0 µL</td>
</tr>
<tr>
<td>Total Volume</td>
<td></td>
<td>52.0 µL</td>
</tr>
</tbody>
</table>
gDNA SAMPLE AMPLIFICATION BY PCR (CONT.)

PCR Thermal Cycling Conditions

- On ice:
  - Add 52 µl of Reaction Mix 5A + 48 µL of purified sample = 100 µL
  - Add 52 µl of Reaction Mix 5B + 48 µL of purified sample = 100 µL
- Pipette mix prepared PCR reactions 10 times with wide-bore pipette tips and then perform a quick spin to collect all liquid from the sides of the tube.
- Place the reaction tubes in a thermal cycler and run the PCR programs as shown on Page 13 of the procedure (set heated lid temp. = 105°C).
- The PCR reactions may be held at 4°C overnight.

<table>
<thead>
<tr>
<th>PCR Program for Reaction Mix 5A</th>
</tr>
</thead>
<tbody>
<tr>
<td>45 seconds at 98°C</td>
</tr>
<tr>
<td>10 seconds at 98°C</td>
</tr>
<tr>
<td>15 seconds at 62°C</td>
</tr>
<tr>
<td>7 minutes at 72°C</td>
</tr>
<tr>
<td>5 minutes at 72°C</td>
</tr>
<tr>
<td>Hold at 4°C</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>PCR Program for Reaction Mix 5B</th>
</tr>
</thead>
<tbody>
<tr>
<td>30 seconds at 98°C</td>
</tr>
<tr>
<td>10 seconds at 98°C</td>
</tr>
<tr>
<td>15 seconds at 60°C</td>
</tr>
<tr>
<td>10 minutes at 68°C</td>
</tr>
<tr>
<td>5 minutes at 68°C</td>
</tr>
<tr>
<td>Hold at 4°C</td>
</tr>
</tbody>
</table>
PURIFICATION OF PCR-AMPLIFIED DNA USING PRONEX BEADS

- **ProNex Beads** must be brought to *room temperature* for 30 to 60 minutes prior to use.
- After performing ProNex Bead purification, perform DNA quantitation with the Qubit dsDNA HS assay kit and DNA sizing with the Femto Pulse System.

**IMPORTANT:** You must have the required yield (mass) of purified PCR-amplified DNA per reaction to proceed with “Pool Amplified DNA”; see guidelines in the table below.

### RECOMMENDATIONS FOR RE-AMPLIFYING SAMPLES WITH LOW PCR PRODUCT YIELDS.

<table>
<thead>
<tr>
<th>DESIRED # OF SMRT CELLS TO RUN</th>
<th>TOTAL YIELD OF AMPLIFIED DNA REQUIRED FOR LIBRARY CONSTRUCTION</th>
<th>FOR SAMPLES WITH LOW PCR PRODUCT YIELD (PCR REACTION A OR B)</th>
<th>RECOMMENDED # OF ADDITIONAL PCR CYCLES (SEE APPENDIX 1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 SMRT Cell 8M</td>
<td>500 ng (PCR Reaction A + B)</td>
<td>If the mass of amplified DNA for PCR Reaction A or B is &lt;275 ng (&lt;11 ng/µL)</td>
<td>Perform 2 additional PCR cycles</td>
</tr>
<tr>
<td></td>
<td></td>
<td>If the mass of amplified DNA for PCR Reaction A or B is &lt;130 ng (&lt;5 ng/µL)</td>
<td>Perform 3 additional PCR cycles</td>
</tr>
<tr>
<td></td>
<td></td>
<td>If the mass of amplified DNA for PCR Reaction A or B is &lt; 65 ng (&lt;2.5 ng/µL)</td>
<td>Perform 5 additional PCR cycles</td>
</tr>
<tr>
<td>2 SMRT Cells 8M</td>
<td>800 ng (PCR Reaction A + B)</td>
<td>If the mass of amplified DNA for PCR Reaction A or B is &lt;400 ng (&lt;15 ng/µL)</td>
<td>Perform 4 to 5 additional PCR cycles</td>
</tr>
</tbody>
</table>
Example Femto Pulse sizing QC analysis of PCR-amplified human and insect DNA samples after purification with ProNex Beads.

Human and insect samples were PCR-amplified with Reaction Mix 5A and 5B. The size distribution of amplified products are approximately 10 kb and the samples are appropriate for proceeding with SMRTbell library construction.
BEST PRACTICES FOR POOLING PCR-AMPLIFIED DNA FOR SMRTBELL LIBRARY CONSTRUCTION

The amplified DNA generated from PCR Reaction A and B are pooled together in equal mass quantities. The pooled DNA can then be constructed into a SMRTbell library as a single sample.

- Always quantify samples before pooling. Since DNA amounts may be limited at this step, PacBio recommends using the Qubit dsDNA High Sensitivity Assay Kit for concentration measurements.

- The total mass of the pooled amplified DNA must meet the minimum mass requirements below to proceed with SMRTbell library construction.
  - For 1 SMRT Cell 8M: ≥500 ng of pooled amplified DNA in 47.4 µL (PCR Reaction A + PCR Reaction B)
  - For 2 SMRT Cells 8M: ≥800 ng of pooled amplified DNA in 47.4 µL (PCR Reaction A + PCR Reaction B)

- If the total mass (resulting from equal-mass pooling) of the two PCR reactions is <500 ng, use any remaining sample to adjust the total combined mass to ≥500 ng. This is to ensure that there is sufficient material for library construction and size selection.

<table>
<thead>
<tr>
<th>STEP</th>
<th>Pooling</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Pool amplified DNA into a single PCR tube of an 8-tube strip.</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Mix and spin down the contents of the tube with a quick spin in a microfuge.</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Proceed to the “Repair DNA Damage” section below.</td>
<td></td>
</tr>
</tbody>
</table>
SMRTBELL LIBRARY SIZE SELECTION USING THE BLUEPIPPIN SYSTEM

For constructing SMRTbell libraries from ultra-low DNA input, PacBio highly recommends size-selection using the BluePippin system.

- BluePippin size-selection requires starting with a SMRTbell library size mode of approximately 8 kb – 10 kb.

- SMRTbell templates <8 kb are efficiently removed using BluePippin size selection
  ▪ Note: A high presence of DNA fragments <8 kb may result in high number of short reads which may impact de novo assembly quality.

- The average size distribution of the final size-selected library is approximately 10 kb – 11 kb.

- Perform size-selection using Sage Sciences’ BluePippin system according to the manufacturer’s recommendations.
  ▪ Select the 0.75% DF 3-10 kb Marker S1 - Improved Recovery cassette definition file
  ▪ Using “Range” selection mode / BP Start = 8000 bp / BP End = 17000 bp
  ▪ Important: Adjust the BluePippin size selection parameters if needed based on the actual size distribution of your SMRTbell library sample as determined by Femto Pulse DNA sizing QC analysis
  ▪ Be sure to assign a marker lane

- Typical recovery yields after size-selection are 20 – 30% (from input of purified SMRTbell library) and are highly dependent on the size distribution of the starting SMRTbell library

- For the latest BluePippin User Manual and guidance on the size-selection protocol, contact Sage Science (www.sagescience.com)

Note: We do not recommend running lanes with <400 ng of SMRTbell library material.
Example Femto Pulse sizing QC analysis of final ultra-low DNA input SMRTbell library after size selection with a BluePippin System.

After size-selection, the mean insert size of the library is 11.0 kb.
ULTRA-LOW DNA INPUT SMRTBELL LIBRARY CONSTRUCTION YIELDS

- Starting with PCR-amplified DNA, SMRTbell library construction yields in this ultra-low DNA input workflow are typically >70% (before size-selection)
- Typical recovery yields after size-selection are ~20 – 30% (from input of purified SMRTbell library) and are highly dependent on the size distribution of the starting SMRTbell library
- For library construction, a minimum 500 ng of amplified gDNA (PCR reaction 1 + PCR reaction 2) is required to generate sufficient SMRTbell library for 1 Sequel II SMRT Cell 8M.
- To generate 2 Sequel II SMRT Cells 8M, we recommend starting with approximately 800 ng of amplified gDNA for library construction.
Ultra-Low DNA Input Library Sequencing Workflow Details
SAMPLE SETUP AND RUN DESIGN RECOMMENDATIONS FOR ULTRA-LOW DNA INPUT LIBRARIES – SEQUEL II SYSTEM (CHEMISTRY 2.0 / SMRT LINK V9.0)

A. Use SMRT Link Sample Setup to prepare ultra-low DNA input libraries for sequencing on the Sequel II System

- In SMRT Link v9.0 Sample Setup, select the following:

<table>
<thead>
<tr>
<th>SMRT LINK SAMPLE SETUP PARAMETER</th>
<th>VALUE TO ENTER</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sequencing Primer</td>
<td>Sequencing Primer v4</td>
</tr>
<tr>
<td>Binding Kit</td>
<td>Sequel II Binding Kit 2.0</td>
</tr>
<tr>
<td>Sequencing Mode</td>
<td>CCS</td>
</tr>
<tr>
<td>Iso-Seq Experiment*</td>
<td>Yes</td>
</tr>
<tr>
<td>Iso-Seq Version*</td>
<td>Yes</td>
</tr>
<tr>
<td>Cleanup Anticipated Yield**</td>
<td>50%</td>
</tr>
</tbody>
</table>

* Select ‘Yes’ for Iso-Seq Experiment to enable 20:1 primer-to-template ratio15:1 polymerase-to-template ratio.

** Cleanup Anticipated Yield for ultra-low DNA input libraries is 50% when using ProNex Bead for the complex cleanup step.
SAMPLE SETUP AND RUN DESIGN RECOMMENDATIONS FOR ULTRA-LOW DNA INPUT LIBRARIES – SEQUEL II SYSTEM (CHEMISTRY 2.0 / SMRT LINK V9.0) (CONT.)

B. Use SMRT Link Run Design to specify instrument run parameters for sequencing ultra-low DNA input libraries on the Sequel II System

- In SMRT Link v9.0 Run Design, select the following:

<table>
<thead>
<tr>
<th>SMRT LINK RUN DESIGN PARAMETER</th>
<th>VALUE TO ENTER</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sequencing Mode</td>
<td>CCS</td>
</tr>
<tr>
<td>On-Plate Loading Concentration</td>
<td>50 – 60 pM</td>
</tr>
<tr>
<td>Pre-Extension Time</td>
<td>2 hours</td>
</tr>
<tr>
<td>Movie Time per SMRT Cell</td>
<td>30 hours</td>
</tr>
<tr>
<td>Generate CCS Data*</td>
<td>YES</td>
</tr>
</tbody>
</table>

* Note: If the SMRT Link instance connected to your sequencing instrument does not communicate directly with your compute cluster, specify “Generate CCS Data” = ‘NO’. In this case, manually start a CCS analysis job in SMRT Link to generate HiFi reads after the sequencing data are transferred to your local storage server.
SAMPLE SETUP RECOMMENDATIONS FOR ULTRA-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 Ultra-Low DNA Input library samples for sequencing

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

### Table 1: Loading and Pre-Extension Time Recommendations

<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 – Ultra-Low DNA Input or Variant Detection – Ultra-Low DNA Input (10 kb – 12 kb)</td>
<td>2</td>
<td>30</td>
</tr>
</tbody>
</table>

### Table 2: De Novo Assembly – Ultra-Low DNA Input or Variant Detection – Ultra-Low DNA Input (19 – 12 kb)

<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 – Ultra-Low DNA Input or Variant Detection – Ultra-Low DNA Input (19 – 12 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 ProNex Beads</td>
<td>50 - 70</td>
</tr>
</tbody>
</table>

https://www.pacb.com/support/documentation/
HiFi Sequencing Data Analysis Recommendations for Ultra-Low DNA Input Libraries
HiFi reads generated from Ultra-Low DNA Input libraries must go through two preliminary processing steps before they can be used for de novo assembly or variant calling: **Trimming of PCR adapter sequences** and **removal of PCR-duplicate reads**.

**HIFI SEQUENCING DATA ANALYSIS RECOMMENDATIONS FOR ULTRA-LOW DNA INPUT LIBRARIES**

Utilize SMRT Link (or the command line via pbbioconda) 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.

Run ‘Trim gDNA Amplification Adapters’ analysis application in SMRT Link (v9.0+) or use lima command line tool (pbbioconda / SMRT Tools).

Run ‘Mark PCR Duplicates’ analysis application in SMRT Link (v9.0+) or use pbmarkdups command line tool (pbbioconda).

Successful datasets typically have >98% of HiFi reads with the PCR adapter sequence present and less than 10% PCR duplicate rates. Samples with higher coverage may have higher observed PCR duplication rates and will consequently have a higher percentage of data removed during processing.
HIFI SEQUENCING DATA ANALYSIS RECOMMENDATIONS FOR ULTRA-LOW DNA INPUT LIBRARIES (CONT.)

A. Using Ultra-Low DNA Input HiFi Reads for De Novo Assembly Applications

- 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

- >30-fold (trimmed, de-duplicated) HiFi read coverage per diploid sample is recommended for most de novo assembly projects using ultra-low DNA input samples

  \[ \text{Target HiFi Base Yield} = \text{[Sample Haploid Genome Size (Gb)]} \times \text{[Target HiFi Coverage per Sample]} \]

  E.g., For de novo assembly analysis of a 500 Mb diploid genome:

  Recommended Minimum Target HiFi Base Yield = 0.5 Gb x 30 = 15 Gb

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

- Can use 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
B. Using Ultra-Low DNA Input HiFi Reads for Variant Detection Applications

- 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

- ≥15 – 20-fold (trimmed, de-duplicated) HiFi read coverage per sample is sufficient for most human variant detection projects using ultra-low DNA input samples

  \[ \text{Target HiFi Base Yield} = \left[ \frac{\text{Sample Haploid Genome Size (Gb)}}{\text{Target Coverage per Sample}} \right] \]

  E.g., For variant detection analysis of a human genome (3 Gb):
  
  Recommended Minimum Target HiFi Base Yield = 3 Gb x 15 = 45 Gb

- For detection of structural variants (>50 bp):
  - Start with the trimmed, deduplicated HiFi data set and run the Structural Variant Calling application in SMRT Analysis.
  - Under the advanced settings make the following parameter changes to optimize performance from this amplified data type:
    - Minimum % of Reads that Support Variant (any one sample): 30
    - Minimum Reads that Support Variant (any one sample): 2
    - Minimum Reads that Support Variant (total over all samples): 2

- For detection of small variants (<50 bp):
  - Start with the trimmed, deduplicated HiFi dataset and run the “Mapping” application in SMRT Analysis
  - The resulting aligned BAM is compatible with the Google DeepVariant PacBio HiFi model for germline small variant calling.
Ultra-Low DNA Input Library Example Performance Data
EXAMPLE 1: DROSOPHILA ULTRA-LOW DNA INPUT LIBRARY FOR DE NOVO GENOME ASSEMBLY

Sample Library Preparation and Sequencing Workflow

1. Drosophila melanogaster insect sample collection

2. Genomic DNA (>20 kb) extraction using QIAGEN MagAttract HMW DNA Kit

3. 10 ng of gDNA sheared to 10 kb and amplified using SMRTbell gDNA Sample Amplification Kit (17 cycles PCR Rx A; 13 cycles PCR Rx B)

4. SMRTbell library construction with 500 ng of amplified gDNA using SMRTbell Express TPK 2.0 and size-selected (11 kb mean insert size)

5. Sequencing on the Sequel II System using a 30-hour collection time

A single SMRT Cell 8M produced 26.8 Gb of HiFi data, or ~183-fold processed HiFi coverage of the Drosophila genome
EXAMPLE 1: DROSOPHILA ULTRA-LOW DNA INPUT LIBRARY FOR DE NOVO GENOME ASSEMBLY

Sample Library Preparation DNA Quantitation & Sizing QC Results for Drosophila Ultra-Low DNA Input Sample

Femto Pulse trace of Drosophila melanogaster gDNA sample before and after shearing with a g-TUBE device to a target fragment size of 10 kb.

Femto Pulse trace of Drosophila melanogaster sample after amplification of the sheared gDNA using PCR Reaction A and PCR Reaction B.

Femto Pulse trace of Drosophila melanogaster SMRTbell library before and after BluePippin size selection.

<table>
<thead>
<tr>
<th>Sample</th>
<th>PCR Reaction</th>
<th>PCR Yield</th>
<th># of Cycles</th>
<th>Mass Used for Pooling</th>
<th>Total Pooled Mass</th>
<th>Library Yield (Pre-BP Size Selection)</th>
<th>Library Yield (Post-BP Size Selection)</th>
<th>Mean Library Insert Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Drosophila melanogaster</td>
<td>A</td>
<td>343 ng</td>
<td>17</td>
<td>250 ng</td>
<td>500 ng</td>
<td>432 ng</td>
<td>170 ng</td>
<td>11.5 kb</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>260 ng</td>
<td>13</td>
<td>250 ng</td>
<td></td>
<td>86%</td>
<td>40%</td>
<td></td>
</tr>
</tbody>
</table>
EXAMPLE 1: DROSOPHILA ULTRA-LOW DNA INPUT LIBRARY FOR DE NOVO GENOME ASSEMBLY

Primary Sequencing Metrics for Drosophila Ultra-Low DNA Input Library

<table>
<thead>
<tr>
<th>Sample</th>
<th>OPC</th>
<th>Movie Time</th>
<th>Pre-Extension Time</th>
<th>Total Bases</th>
<th>Unique Molecular Yield</th>
<th>P0 %</th>
<th>P1 %</th>
<th>P2 %</th>
<th>Pol RL Mean (bp)</th>
<th>Pol RL N50 (bp)</th>
<th>Longest Subread (bp)</th>
<th>Longest Subread N50 (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Drosophila melanogaster</td>
<td>50 pM</td>
<td>30 h</td>
<td>2 h</td>
<td>465.1 Gb</td>
<td>69.8 Gb</td>
<td>29.3</td>
<td>69.1</td>
<td>1.6</td>
<td>84,036</td>
<td>173,698</td>
<td>13,732</td>
<td>13,946</td>
</tr>
</tbody>
</table>

Example sequencing results shown were generated using Sequel II System Chemistry 2.0.
EXAMPLE 1: DROSOPHILA ULTRA-LOW DNA INPUT LIBRARY FOR DE NOVO GENOME ASSEMBLY

HiFi Sequencing Yield Metrics for Drosophila Ultra-Low DNA Input Library

**DROSOPHILA HIFI SEQUENCING YIELD METRICS (1 SMRT CELL 8M)**

<table>
<thead>
<tr>
<th>METRIC</th>
<th>VALUE</th>
</tr>
</thead>
<tbody>
<tr>
<td>HiFi Data Yield</td>
<td>26.8 Gb</td>
</tr>
<tr>
<td>Mean HiFi Read Length</td>
<td>11,153 bp</td>
</tr>
<tr>
<td>Median HiFi Read Quality</td>
<td>Q34</td>
</tr>
<tr>
<td>PCR Adapter Percentage</td>
<td>99.85%</td>
</tr>
<tr>
<td>PCR Duplication Rate</td>
<td>3.78%</td>
</tr>
<tr>
<td>Processed HiFi Coverage for 140 Mb Genome</td>
<td>183-Fold</td>
</tr>
</tbody>
</table>

PacBio Application Note: Considerations for Using the Low and Ultra-Low DNA Input Workflows for Whole Genome Sequencing (PN 101-995-900)
EXAMPLE 1: **DROSOPHILA ULTRA-LOW DNA INPUT LIBRARY FOR DE NOVO GENOME ASSEMBLY**

HiFi *De Novo* Assembly Performance Results for *Drosophila* Ultra-Low DNA Input Library

### **DROSOPHILA GENOME ASSEMBLY METRICS (1 SMRT CELL 8M)**

<table>
<thead>
<tr>
<th>METRIC</th>
<th>VALUE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Down sampled HiFi Coverage</td>
<td>30-fold</td>
</tr>
<tr>
<td>Assembly Size</td>
<td>147 Mb</td>
</tr>
<tr>
<td>Contig N50</td>
<td>8.3 Mb</td>
</tr>
<tr>
<td>BUSCO Complete</td>
<td>98.6%</td>
</tr>
</tbody>
</table>

HiFi assembly using Hifiasm.

**Drosophila ULI Assembly vs. FlyBaseRef 6.35**

Dot plot analysis of *Drosophila* ultra-low DNA input genome assembly versus FlyBaseRef 6.35 reference sequence. The ULI assembly shows a high level of congruence with the FlyBase reference sequence.

*PacBio Application Note*: Considerations for Using the Low and Ultra-Low DNA Input Workflows for Whole Genome Sequencing (PN 101-995-900)
EXAMPLE 2: CORIELL HG002 ULTRA-LOW DNA INPUT LIBRARY FOR HUMAN VARIANT DETECTION

Sample Library Preparation and Sequencing Workflow

Coriell HG002 cell line pellets were collected

Genomic DNA (>20 kb) extraction using Lucigen MasterPure DNA Purification Kit

10 ng of gDNA sheared to 10 kb and amplified using SMRTbell gDNA Sample Amplification Kit (18 cycles PCR Rx A; 17 cycles PCR Rx B)

SMRTbell library construction with 808 ng of amplified gDNA using SMRTbell Express TPK 2.0 and size-selected (11 kb mean insert size)

Sequencing on the Sequel II System using a 30-hour collection time

Sequencing on 2 SMRT Cells 8M produced >60 Gb of HiFi data, or ~20-fold processed HiFi coverage of the HG002 genome

PacBio Application Note: Considerations for Using the Low and Ultra-Low DNA Input Workflows for Whole Genome Sequencing (PN 101-995-900)
EXAMPLE 2: CORIELL HG002 ULTRA-LOW DNA INPUT LIBRARY FOR HUMAN VARIANT DETECTION (CONT.)

Sample Library Preparation DNA Quantitation & Sizing QC Results for HG002 Ultra-Low DNA Input Library

HG002 gDNA sample before and after shearing with a g-TUBE device to a target fragment size of 10 kb. HG002 sample after amplification of the sheared gDNA using PCR Reaction A and PCR Reaction B. HG002 SMRTbell library before and after BluePippin size selection.

<table>
<thead>
<tr>
<th>Sample</th>
<th>PCR Reaction</th>
<th>PCR Yield</th>
<th># of Cycles</th>
<th>Mass Used for Pooling</th>
<th>Total Pooled Mass</th>
<th>Library Yield (Pre-BP Size Selection)</th>
<th>Library Yield (Post-BP Size Selection)</th>
<th>Mean Library Insert Size</th>
<th># of SMRT Cells 8M</th>
</tr>
</thead>
<tbody>
<tr>
<td>HG002</td>
<td>A</td>
<td>811 ng</td>
<td>13</td>
<td>405 ng</td>
<td>808 ng</td>
<td>636</td>
<td>236 ng</td>
<td>11.0 kb</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>806 ng</td>
<td>13</td>
<td>403 ng</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
EXAMPLE 2: CORIELL HG002 ULTRA-LOW DNA INPUT LIBRARY FOR HUMAN VARIANT DETECTION (CONT.)

HiFi Sequencing Yield Metrics for HG002 Ultra-Low DNA Input Library

HG002 HIFI SEQUENCING YIELD METRICS (2 SMRT CELLS 8M)

<table>
<thead>
<tr>
<th>METRIC</th>
<th>SMRT CELL 1</th>
<th>SMRT CELL 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>HiFi Data Yield</td>
<td>32.22 Gb</td>
<td>30.77 Gb</td>
</tr>
<tr>
<td>Mean HiFi Read Length</td>
<td>10,909 bp</td>
<td>10,999 bp</td>
</tr>
<tr>
<td>Median HiFi Read Quality</td>
<td>Q33</td>
<td>Q37</td>
</tr>
<tr>
<td>PCR Adapter Percentage</td>
<td>99.90%</td>
<td>99.89%</td>
</tr>
<tr>
<td>PCR Duplication Rate</td>
<td>6.89%</td>
<td>4.32%</td>
</tr>
<tr>
<td>Processed HiFi Coverage for 3 Gb Genome</td>
<td>10-Fold</td>
<td>10-Fold</td>
</tr>
</tbody>
</table>

HiFi Sequencing Yield Metrics for HG002 Ultra-Low DNA Input Library

PacBio Application Note: Considerations for Using the Low and Ultra-Low DNA Input Workflows for Whole Genome Sequencing (PN 101-995-900)
## EXAMPLE 2: CORIELL HG002 ULTRA-LOW DNA INPUT LIBRARY FOR HUMAN VARIANT DETECTION (CONT.)

### Variant Calling Performance Results for HG002 Ultra-Low DNA Input Library

<table>
<thead>
<tr>
<th>VARIANT TYPE</th>
<th>VARIANT CALLER</th>
<th>PRECISION</th>
<th>RECALL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single Nucleotide Variants</td>
<td>DeepVariant</td>
<td>99.6 %</td>
<td>99.0%</td>
</tr>
<tr>
<td>Indels (&lt;49 bp)</td>
<td>DeepVariant</td>
<td>84.7%</td>
<td>90.3%</td>
</tr>
<tr>
<td>Structural Variants (≥50 bp)</td>
<td>PBSV</td>
<td>95.4%</td>
<td>84.3%</td>
</tr>
</tbody>
</table>
Technical Documentation & Applications Support Resources
TECHNICAL DOCUMENTATION & APPLICATIONS SUPPORT RESOURCES FOR ULTRA-LOW DNA INPUT LIBRARY PREPARATION, SEQUENCING & DATA ANALYSIS

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 – Preparing HiFi Libraries from Ultra-Low DNA Input (PN 101-987-800)
- 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: Ultra-Low DNA Input Library Preparation Using SMRTbell Express Template Prep Kit 2.0 (PN 101-998-000)
Data Analysis Resources

- SMRT Link User Guide (v9.0: PN 101-908-200)
  - Describes how to run CCS, Trim gDNA Amplification Adapters and Mark PCR Duplicates analysis applications in SMRT Link

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 (Ultra-Low DNA Input)</td>
<td><em>Phlebotomus papatas</em> (Sand Fly) – Male [SRR12454518]</td>
<td>HiFi Reads</td>
<td>Sequel II System</td>
</tr>
<tr>
<td></td>
<td><em>Drosophila melanogaster</em> (Fruit Fly) – Female [SRR12473480]</td>
<td>HiFi Reads</td>
<td>Sequel II System</td>
</tr>
<tr>
<td>Variant Detection (Ultra-Low DNA Input)</td>
<td><em>Homo sapiens</em> (HG002) – Male [SRR12454519]</td>
<td>HiFi Reads</td>
<td>Sequel II System</td>
</tr>
<tr>
<td></td>
<td><em>Homo sapiens</em> (HG002) – Male [SRR12454520]</td>
<td>HiFi Reads</td>
<td>Sequel II System</td>
</tr>
</tbody>
</table>
TECHNICAL DOCUMENTATION & APPLICATIONS SUPPORT RESOURCES FOR ULTRA-LOW DNA INPUT LIBRARY PREPARATION, SEQUENCING & DATA ANALYSIS (CONT.)

Posters
- PacBio PAG 2020 Poster: A High-Quality PacBio Insect Genome from 5 ng of Input DNA.

Publications

Webinars
- PacBio Virtual Global Summit Presentation (2020): Technical Tutorial – Choosing the best library prep for HiFi sequencing. [Webinar Recording]
- SMRT Leiden Presentation (2020): Revealing the Genome Biology of Earth’s Smallest Animals. [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](#)) 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

Methods

Sample collection

A picture of the yellow belly pufferfish used in the genome sequencing and assembly.

DNA and RNA sequencing

Tissue DNA from white muscle tissue was extracted using the...

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

### A. Genomic DNA Sizing Characterization

**Femto Pulse System (Agilent Technologies)**

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


**Fragment Analyzer System (Agilent Technologies)**

- Resolves up to ~50 kb
- Requires 2 ng of sample
- 1-hour analysis time


**Evaluation of gDNA quality using a Femto Pulse System.** Lane 1B shows an example of a high quality, high-molecular weight genomic DNA sample. Lane 2B shows an example of a degraded gDNA sample.
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.