PACIFIC BIOSCIENCES®

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

PN 101-998-000 Ver 2021-02-01-A (February 2021)

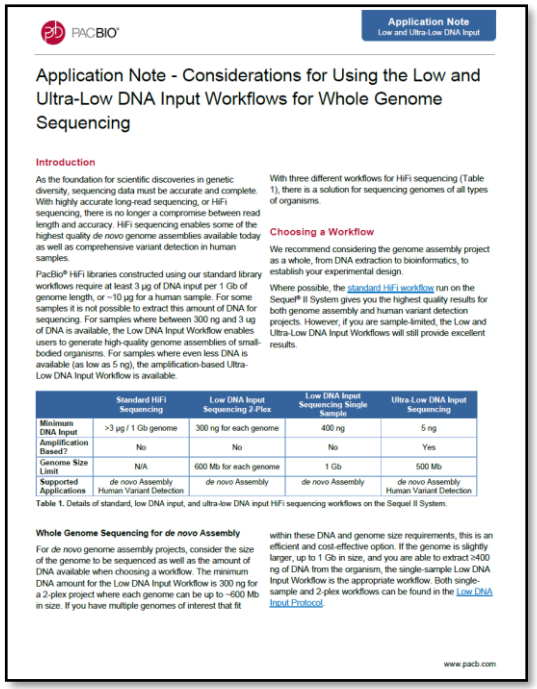
Ultra-Low DNA Input Library Preparation Using SMRTbell Express Template Prep Kit 2.0

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

Low- / Ultra-Low DNA Input Application Note → **Application-Specific Procedure & Checklist** → **Application Consumable Bundle Purchasing Guide** → **Library Construction, Sequencing & Analysis**



PACBIO

Application Note
Low and Ultra-Low DNA Input

Application Note - Considerations for Using the Low and Ultra-Low DNA Input Workflows for Whole Genome Sequencing

Introduction

As the foundation for scientific discoveries in genetic diversity, sequencing data must be accurate and complete. With highly accurate long-read sequencing, or HiFi sequencing, there is no longer a compromise between read length and accuracy. HiFi sequencing enables some of the highest quality *de novo* genome assemblies available today as well as comprehensive variant detection in human samples.

PacBio® HiFi libraries constructed using our standard library workflows require at least 3 µg of DNA input per 1 Gb of genome length, or ~10 µg for a human sample. For some samples it is not possible to extract this amount of DNA for sequencing. For samples where between 300 ng and 3 ug 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.

With three different workflows for HiFi sequencing (Table 1), there is a solution for sequencing genomes of all types of organisms.

Choosing a Workflow

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

Where possible, the standard HiFi workflow run on the Sequel® II System 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 will still provide excellent results.

	Standard HiFi Sequencing	Low DNA Input Sequencing 2-Plex	Low DNA Input Sequencing Single Sample	Ultra-Low DNA Input Sequencing
Minimum DNA Input Amplification Based?	>3 µg / 1 Gb genome	300 ng for each genome	400 ng	5 ng
Genome Size Limit	No	No	No	Yes
Supported Applications	N/A	600 Mb for each genome	1 Gb	500 Mb
	de novo Assembly Human Variant Detection	*de novo* Assembly	*de novo* Assembly	*de novo* Assembly Human Variant Detection

Table 1. Details of standard, low DNA input, and ultra-low DNA input HiFi sequencing workflows on the Sequel II System.

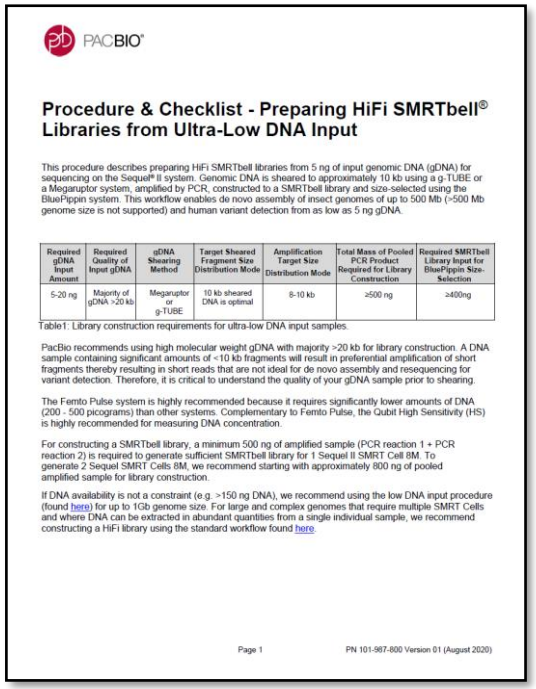
Whole Genome Sequencing for *de novo* Assembly

For *de novo* genome assembly projects, consider the size of the genome to be sequenced as well as the amount of DNA available when choosing a workflow. The minimum DNA amount for the Low DNA Input Workflow is 300 ng for a 2-plex project where each genome can be up to ~600 Mb in size. If you have multiple genomes of interest that fit within these DNA and genome size requirements, this is an efficient and cost-effective option. If the genome is slightly larger, up to 1 Gb in size, and you are able to extract ≥400 ng of DNA from the organism, the single-sample Low DNA Input Workflow is the appropriate workflow. Both single-sample and 2-plex workflows can be found in the Low DNA Input Protocol.

www.pacb.com

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



PACBIO

Procedure & Checklist - Preparing HiFi SMRTbell® Libraries from Ultra-Low DNA Input

This procedure describes preparing HiFi SMRTbell libraries from 5 ng of input genomic DNA (gDNA) for sequencing on the Sequel® II system. 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 gDNA.

Required gDNA Input Amount	Required Quality of Input gDNA	gDNA Shearing Method	Target Sheared Fragment Size Distribution Mode	Amplification Target Size Distribution Mode	Total Mass of Pooled PCR Product Required for Library Construction	Required SMRTbell Library Input for BluePippin Size-Selection
5-20 ng	Majority of gDNA >20 kb	Megaruptor or g-TUBE	10 kb sheared DNA is optimal	8-10 kb	≥500 ng	≥400ng

Table1: Library construction requirements for ultra-low DNA input samples.

PacBio recommends using high molecular weight gDNA with majority >20 kb for library construction. A DNA sample containing significant amounts of <10 kb fragments will result in preferential amplification of short fragments thereby resulting in short reads that are not ideal for de novo assembly and resequencing for variant detection. Therefore, it is critical to understand the quality of your gDNA sample prior to shearing.

The Femto Pulse system is highly recommended because it requires significantly lower amounts of DNA (200 - 500 picograms) than other systems. Complementary to Femto Pulse, the Qubit High Sensitivity (HS) is highly recommended for measuring DNA concentration.

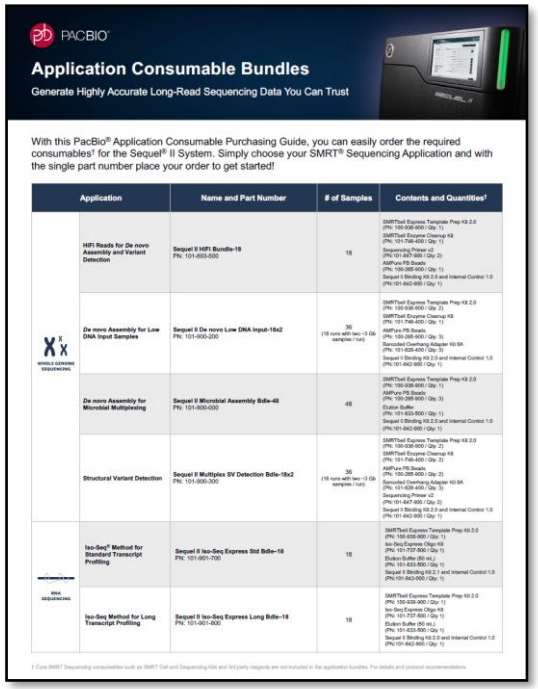
For constructing a SMRTbell library, a minimum 500 ng of amplified sample (PCR reaction 1 + PCR reaction 2) is required to generate sufficient SMRTbell library for 1 Sequel II SMRT Cell 8M. To generate 2 Sequel SMRT Cells 8M, we recommend starting with approximately 800 ng of pooled amplified sample for library construction.

If DNA availability is not a constraint (e.g. >150 ng DNA), we recommend using the low DNA input procedure (found here) for up to 1Gb genome size. For large and complex genomes that require multiple SMRT Cells and where DNA can be extracted in abundant quantities from a single individual sample, we recommend constructing a HiFi library using the standard workflow found here.

Page 1 PN 101-987-800 Version 01 (August 2020)

Procedure & Checklist – Preparing HiFi Libraries from Ultra-Low DNA Input (101-987-800)

Technical documentation containing sample library construction and sequencing preparation protocol details



PACBIO

Application Consumable Bundles

Generate Highly Accurate Long-Read Sequencing Data You Can Trust

With this PacBio® Application Consumable Purchasing Guide, you can easily order the required consumables† for the Sequel® II System. Simply choose your SMRT® Sequencing Application and with the single part number place your order to get started!

Application	Name and Part Number	# of Samples	Contents and Quantities‡
HiFi Reads for *De novo* Assembly and Variant Detection	Sequel II HiFi Bundle-1B PN: 101-893-500	1B	SMRTbell Express Template Prep Kit 2.0 (PN: 100-938-900 / Qty: 1) SMRTbell Enzyme Cleanup Kit (PN: 101-746-400 / Qty: 1) Sequencing Primer v2 (PN: 101-847-900 / Qty: 2) AMPure PB Beads (PN: 100-265-000 / Qty: 1) Sequel II Binding Kit 2.0 and Internal Control 1.0 (PN: 101-842-900 / Qty: 1)
De novo Assembly for Low DNA Input Samples	Sequel II De novo Low DNA Input-16x2 PN: 101-900-200	36 (16 runs with two ~3 Gb samples / run)	SMRTbell Express Template Prep Kit 2.0 (PN: 100-938-900 / Qty: 2) SMRTbell Enzyme Cleanup Kit (PN: 101-746-400 / Qty: 2) AMPure PB Beads (PN: 100-265-000 / Qty: 3) Barcoded Overhang Adapter Kit 8A (PN: 101-828-400 / Qty: 3) Sequel II Binding Kit 2.0 and Internal Control 1.0 (PN: 101-842-900 / Qty: 1)
De novo Assembly for Microbial Multiplexing	Sequel II Microbial Assembly 8dle-4B PN: 101-900-000	48	SMRTbell Express Template Prep Kit 2.0 (PN: 100-938-900 / Qty: 1) AMPure PB Beads (PN: 100-265-000 / Qty: 3) Elution Buffer (PN: 101-633-500 / Qty: 1) Sequel II Binding Kit 2.0 and Internal Control 1.0 (PN: 101-842-900 / Qty: 1)
Structural Variant Detection	Sequel II Multiplex SV Detection 8dle-16x2 PN: 101-900-300	36 (16 runs with two ~3 Gb samples / run)	SMRTbell Express Template Prep Kit 2.0 (PN: 100-938-900 / Qty: 2) SMRTbell Enzyme Cleanup Kit (PN: 101-746-400 / Qty: 2) AMPure PB Beads (PN: 100-265-000 / Qty: 2) Barcoded Overhang Adapter Kit 8A (PN: 101-828-400 / Qty: 3) Sequencing Primer v2 (PN: 101-847-900 / Qty: 2) Sequel II Binding Kit 2.0 and Internal Control 1.0 (PN: 101-842-900 / Qty: 1)
Iso-Seq® Method for Standard Transcript Profiling	Sequel II Iso-Seq Express Std 8dle-1B PN: 101-901-700	1B	SMRTbell Express Template Prep Kit 2.0 (PN: 100-938-900 / Qty: 1) Iso-Seq Express Oligo Kit (PN: 101-737-500 / Qty: 1) Elution Buffer (50 mL) (PN: 101-633-500 / Qty: 1) Sequel II Binding Kit 2.1 and Internal Control 1.0 (PN: 101-843-000 / Qty: 1)
Iso-Seq Method for Long Transcript Profiling	Sequel II Iso-Seq Express Long 8dle-1B PN: 101-901-800	1B	SMRTbell Express Template Prep Kit 2.0 (PN: 100-938-900 / Qty: 1) Iso-Seq Express Oligo Kit (PN: 101-737-500 / Qty: 1) Elution Buffer (50 mL) (PN: 101-633-500 / Qty: 1) Sequel II Binding Kit 2.0 and Internal Control 1.0 (PN: 101-842-900 / Qty: 1)

‡ Core SMRT Sequencing consumables such as SMRT Cell and Sequencing Kits and 3rd party reagents are not included in the application bundles. For details and protocol recommendations

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 Systema*

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.

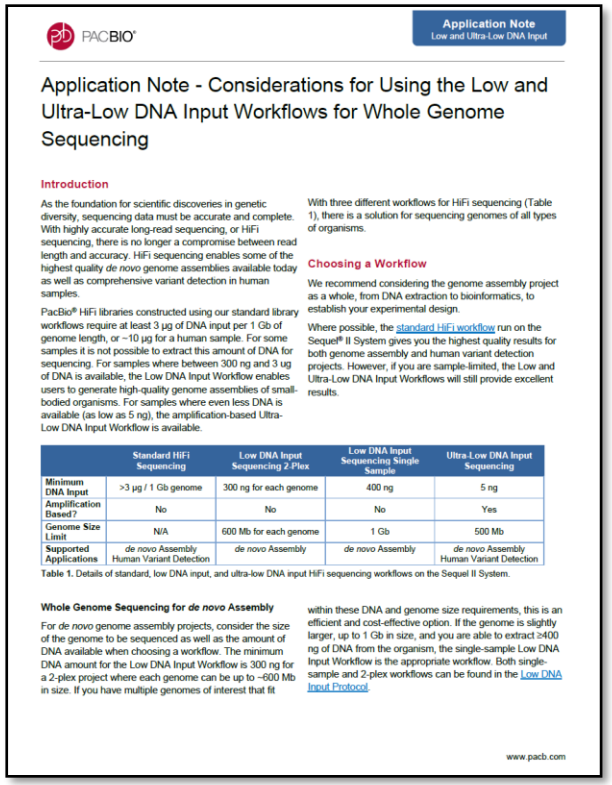
PacBio's standard HiFi library preparation workflow requires at least ~3 µg of DNA input per 1 Gb of genome length, or ~15 µg for a human sample – but for some samples, it is not possible to extract this amount of DNA for sequencing.

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 IIe 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



PACBIO

Application Note
Low and Ultra-Low DNA Input

Application Note - Considerations for Using the Low and Ultra-Low DNA Input Workflows for Whole Genome Sequencing

Introduction

As the foundation for scientific discoveries in genetic diversity, sequencing data must be accurate and complete. With highly accurate long-read sequencing, or HiFi sequencing, there is no longer a compromise between read length and accuracy. HiFi sequencing enables some of the highest quality *de novo* genome assemblies available today as well as comprehensive variant detection in human samples.

PacBio® HiFi libraries constructed using our standard library workflows require at least 3 µg of DNA input per 1 Gb of genome length, or ~10 µg for a human sample. For some samples it is not possible to extract this amount of DNA for sequencing. For samples where between 300 ng and 3 ug 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.

With three different workflows for HiFi sequencing (Table 1), there is a solution for sequencing genomes of all types of organisms.

Choosing a Workflow

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

Where possible, the standard HiFi workflow run on the Sequel® II System 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 will still provide excellent results.

	Standard HiFi Sequencing	Low DNA Input Sequencing 2-Plex	Low DNA Input Sequencing Single Sample	Ultra-Low DNA Input Sequencing
Minimum DNA Input	>3 µg / 1 Gb genome	300 ng for each genome	400 ng	5 ng
Amplification Based?	No	No	No	Yes
Genome Size Limit	N/A	600 Mb for each genome	1 Gb	500 Mb
Supported Applications	*de novo* Assembly Human Variant Detection	*de novo* Assembly	*de novo* Assembly	*de novo* Assembly Human Variant Detection

Table 1. Details of standard, low DNA input, and ultra-low DNA input HiFi sequencing workflows on the Sequel II System.

Whole Genome Sequencing for *de novo* Assembly

For *de novo* genome assembly projects, consider the size of the genome to be sequenced as well as the amount of DNA available when choosing a workflow. The minimum DNA amount for the Low DNA Input Workflow is 300 ng for a 2-plex project where each genome can be up to ~600 Mb in size. If you have multiple genomes of interest that fit within these DNA and genome size requirements, this is an efficient and cost-effective option. If the genome is slightly larger, up to 1 Gb in size, and you are able to extract ≥400 ng of DNA from the organism, the single-sample Low DNA Input Workflow is the appropriate workflow. Both single-sample and 2-plex workflows can be found in the Low DNA Input Protocol.

www.pacb.com

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.

SAMPLE & PROJECT CONSIDERATIONS	STANDARD HIFI SEQUENCING	LOW DNA INPUT SEQUENCING (2-PLEX)	LOW DNA INPUT SEQUENCING (SINGLE SAMPLE)	ULTRA-LOW DNA INPUT SEQUENCING
Minimum DNA Input	>3 µg / 1 Gb genome	300 ng for each genome	400 ng	5 ng
Amplification Based?	No	No	No	Yes
Genome Size Limit	N/A	600 Mb for each genome	1 Gb	500 Mb
Supported Applications	*De novo* Assembly Human Variant Detection	*De novo* Assembly	*De novo* Assembly	*De novo* Assembly Human Variant Detection

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

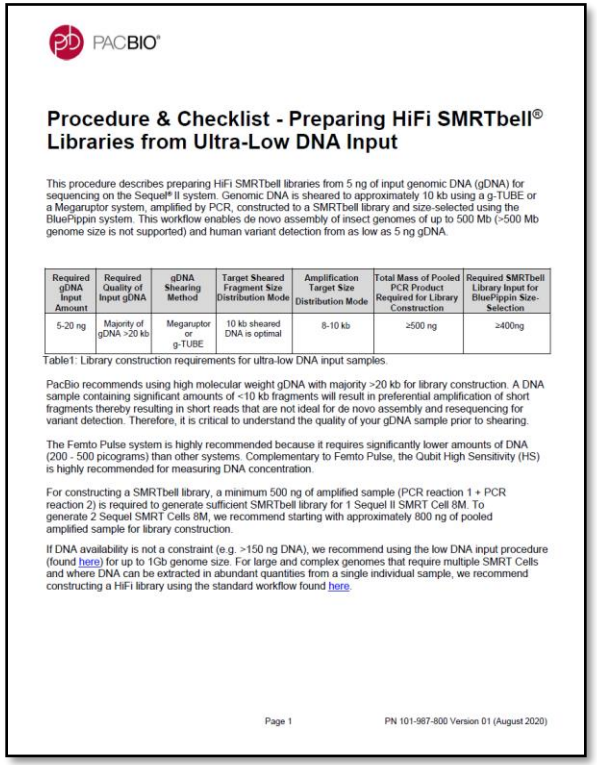
***De novo* assembly for microbes, plants, vertebrates, or other non-DNA limited sample types**

COMPLEX POPULATIONS

Metagenomics sequencing

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 IIe 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



PACBIO®

Procedure & Checklist - Preparing HiFi SMRTbell® Libraries from Ultra-Low DNA Input

This procedure describes preparing HiFi SMRTbell libraries from 5 ng of input genomic DNA (gDNA) for sequencing on the Sequel® II system. 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 gDNA.

Required gDNA Input Amount	Required Quality of Input gDNA	gDNA Shearing Method	Target Sheared Fragment Size Distribution Mode	Amplification Target Size Distribution Mode	Total Mass of Pooled PCR Product Required for Library Construction	Required SMRTbell Library Input for BluePippin Size-Selection
5-20 ng	Majority of gDNA >20 kb	Megaruptor or g-TUBE	10 kb sheared DNA is optimal	8-10 kb	≥500 ng	≥400ng

Table1: Library construction requirements for ultra-low DNA input samples.

PacBio recommends using high molecular weight gDNA with majority >20 kb for library construction. A DNA sample containing significant amounts of <10 kb fragments will result in preferential amplification of short fragments thereby resulting in short reads that are not ideal for de novo assembly and resequencing for variant detection. Therefore, it is critical to understand the quality of your gDNA sample prior to shearing.

The Femto Pulse system is highly recommended because it requires significantly lower amounts of DNA (200 - 500 picograms) than other systems. Complementary to Femto Pulse, the Qubit High Sensitivity (HS) is highly recommended for measuring DNA concentration.

For constructing a SMRTbell library, a minimum 500 ng of amplified sample (PCR reaction 1 + PCR reaction 2) is required to generate sufficient SMRTbell library for 1 Sequel II SMRT Cell 8M. To generate 2 Sequel SMRT Cells 8M, we recommend starting with approximately 800 ng of pooled amplified sample for library construction.

If DNA availability is not a constraint (e.g. >150 ng DNA), we recommend using the low DNA input procedure (found here) for up to 1Gb genome size. For large and complex genomes that require multiple SMRT Cells and where DNA can be extracted in abundant quantities from a single individual sample, we recommend constructing a HiFi library using the standard workflow found here.

Page 1 PN 101-987-800 Version 01 (August 2020)

https://www.pacb.com/support/documentation/

APPLICATIONS

WHOLE GENOME SEQUENCING

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

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

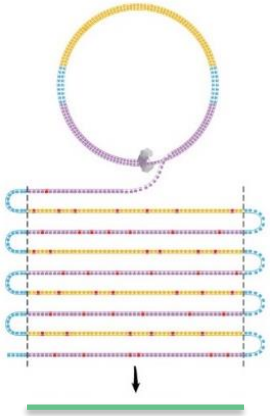
- BluePippin size selection (Collect 8 – 17 kb fraction)

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

DNA QUALITY AND QUANTITY REQUIREMENTS FOR ULTRA-LOW DNA INPUT SMRTBELL LIBRARY PREPARATION

- 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 IIe SYSTEMS.

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

* 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

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

Femto Pulse System

GENOMIC DNA SIZE EVALUATION METHODS AND PROCEDURES.

DNA SIZING METHOD	COMMENTS	PROCEDURE
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

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.

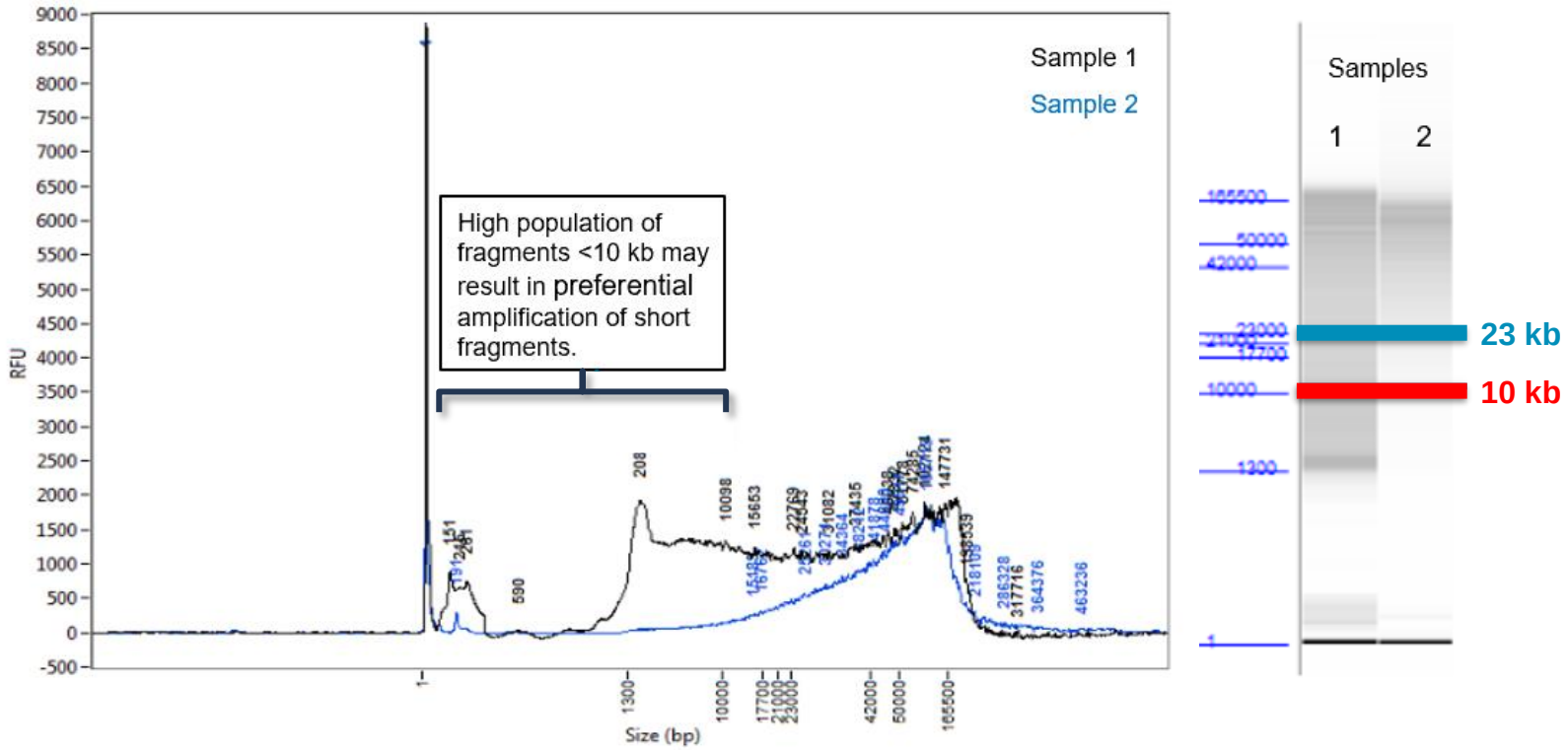
- PacBio recommends working with **high-quality samples** where the majority of the input gDNA is **>20 kb**
- Use of low-quality gDNA samples in the ultra-low DNA input procedure will negatively impact the PCR amplification step

  → 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.

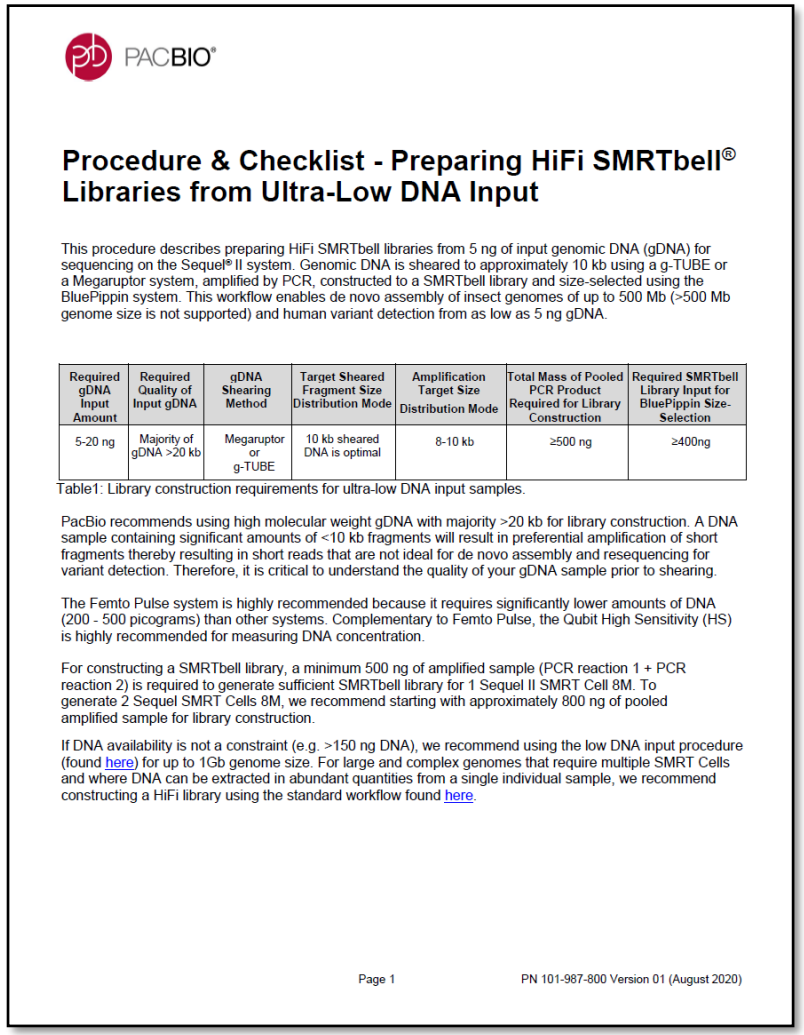


Example Femto Pulse sizing QC analysis of individual mosquito gDNA samples of varying quality. gDNA samples with a majority of fragments >30 kb are suitable for HiFi library construction using this procedure. Sample 1 is lower quality gDNA such that a significant proportion of the fragments are ≤10 and is not suitable constructing SMRTbell libraries using the ultra-low DNA input procedure. Sample 2 show 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 IIe 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 IIe Systems



PACBIO®

Procedure & Checklist - Preparing HiFi SMRTbell® Libraries from Ultra-Low DNA Input

This procedure describes preparing HiFi SMRTbell libraries from 5 ng of input genomic DNA (gDNA) for sequencing on the Sequel® II system. 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 gDNA.

Required gDNA Input Amount	Required Quality of Input gDNA	gDNA Shearing Method	Target Sheared Fragment Size Distribution Mode	Amplification Target Size Distribution Mode	Total Mass of Pooled PCR Product Required for Library Construction	Required SMRTbell Library Input for BluePippin Size-Selection
5-20 ng	Majority of gDNA >20 kb	Megaruptor or g-TUBE	10 kb sheared DNA is optimal	8-10 kb	≥500 ng	≥400ng

Table1: Library construction requirements for ultra-low DNA input samples.

PacBio recommends using high molecular weight gDNA with majority >20 kb for library construction. A DNA sample containing significant amounts of <10 kb fragments will result in preferential amplification of short fragments thereby resulting in short reads that are not ideal for de novo assembly and resequencing for variant detection. Therefore, it is critical to understand the quality of your gDNA sample prior to shearing.

The Femto Pulse system is highly recommended because it requires significantly lower amounts of DNA (200 - 500 picograms) than other systems. Complementary to Femto Pulse, the Qubit High Sensitivity (HS) is highly recommended for measuring DNA concentration.

For constructing a SMRTbell library, a minimum 500 ng of amplified sample (PCR reaction 1 + PCR reaction 2) is required to generate sufficient SMRTbell library for 1 Sequel II SMRT Cell 8M. To generate 2 Sequel SMRT Cells 8M, we recommend starting with approximately 800 ng of pooled amplified sample for library construction.

If DNA availability is not a constraint (e.g. >150 ng DNA), we recommend using the low DNA input procedure (found here) for up to 1Gb genome size. For large and complex genomes that require multiple SMRT Cells and where DNA can be extracted in abundant quantities from a single individual sample, we recommend constructing a HiFi library using the standard workflow found here.

Page 1 PN 101-987-800 Version 01 (August 2020)

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

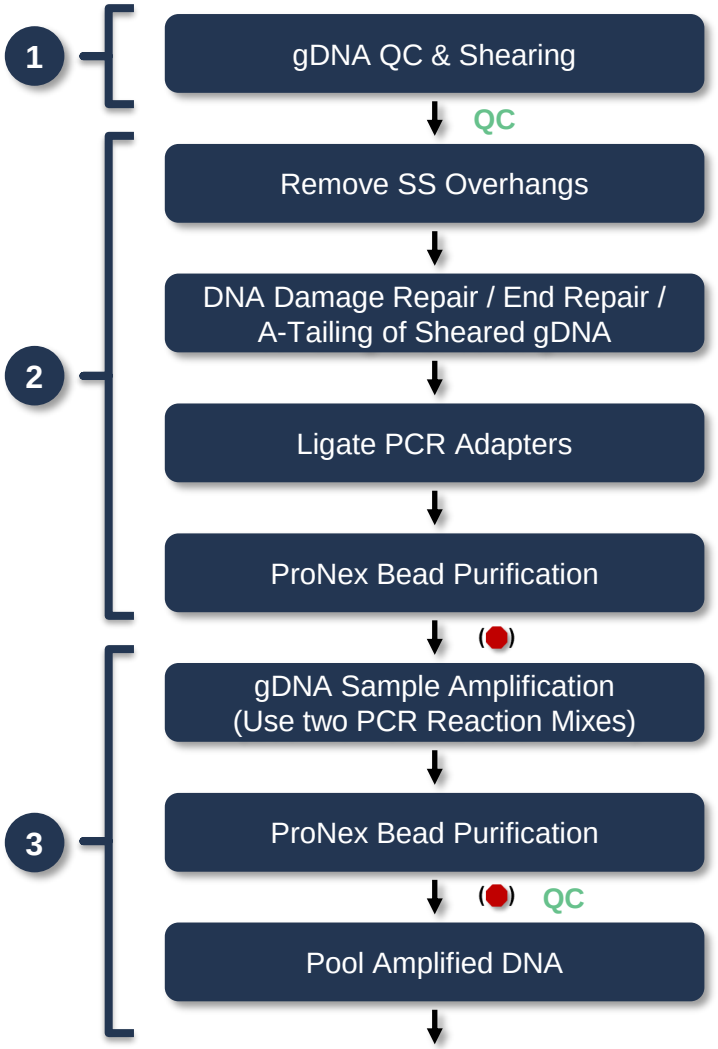
3. Sample Amplification & Pooling (2.5 h)

- 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**.

QC = Perform Femto Pulse DNA Sizing & Qubit Quantitation QC

(●) = Optional Stop Point



ULTRA-LOW DNA INPUT WORKFLOW DETAILED OVERVIEW (CONT.)

QC = Perform Femto Pulse DNA Sizing & Qubit Quantitation QC

(●) = Optional Stop Point

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)

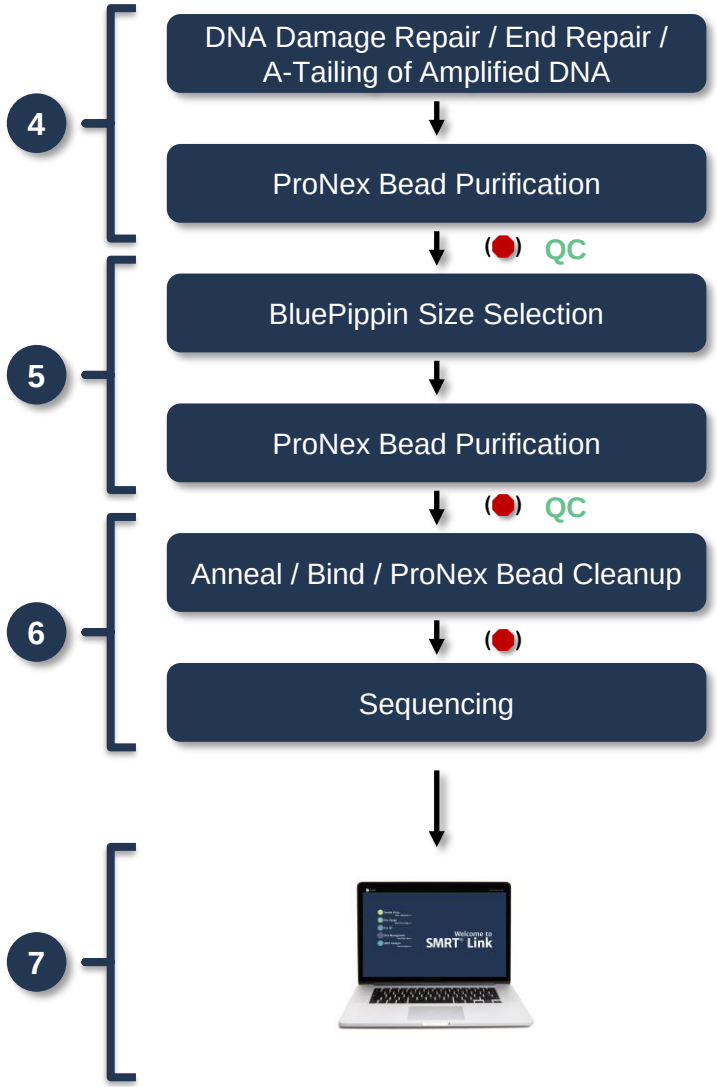
- Perform BluePippin size selection (8 – 17 kb fraction collection)
- Typical recovery yield of size-selected library: >20 – 30%
- The average size distribution of the final size-selected library is approximately **10 – 11 kb**.

6. Sequencing Preparation (Sequel II and IIe 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

ITEM	VENDOR	PART NUMBER
DNA Sizing QC (One of the following)		
Femto Pulse System (**Highly Recommended**)	Agilent	M5330AA
Fragment Analyzer	Agilent	M5310AA
DNA Quantitation		
Qubit Fluorometer	Thermo Fisher Scientific	Q33226
Qubit 1X dsDNA HS Assay Kit	Thermo Fisher Scientific	Q33230
DNA Shearing		
Megaruptor 1 or Megaruptor 2 System	Diagenode	B06010001 / B06010002
Long Hydropores	Diagenode	E07010002
Hydrotubes	Diagenode	C30010018
g-TUBE	Covaris	520104
Eppendorf MiniSpin Plus or other equivalent benchtop centrifuge model	Eppendorf	22620100
SMRTbell Library Preparation		
SMRTbell Express Template Prep Kit 2.0	PacBio	100-938-900
SMRTbell gDNA Sample Amplification Kit	PacBio	101-980-000
Elution Buffer	PacBio	101-633-500
HDPE 8 place Magnetic Separation Rack for 0.2 ml PCR Tubes	V&P Scientific Inc.	VP772F4-1
0.2 mL PCR 8-strip tubes	USA Scientific	1402-4708
Wide Orifice Tips (Tips LTS W-O 200UL Fltr RT-L200WFLR)	Rainin	17014294
ProNex Beads	Promega	NG2002 – 125 mL

LIST OF REQUIRED MATERIALS AND EQUIPMENT (CONT.)

ITEM	VENDOR	PART NUMBER
SMRTbell Library Preparation (Cont.)		
100% Ethanol, Molecular Biology Grade	Any Major Life Science Supply Vendor	Vendor-specific
Thermal Cycler that is 100 µL and 8-tube strip compatible	Any Major Life Science Supply Vendor	Vendor-specific
SMRTbell Library Size Selection		
BluePippin Size-Selection System	Sage Science	BLU0001
BluePippin with dye free, 0.75% Agarose Cassettes and S1 Marker	Sage Science	BLF7510

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

SMRTbell Express TPK 2.0

SMRTbell gDNA Sample Amplification Kit

LIST OF TEMPERATURE-SENSITIVE REAGENTS INCLUDED IN SMRTBELL gDNA SAMPLE AMPLIFICATION KIT AND SMRTBELL EXPRESS TPK 2.0

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

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.

KIT COMPONENT	PURPOSE
Duplex Buffer	Buffer for adapter ligation
Amplification Adapters	Adapters ligated to double-stranded fragments
Sample Amplification PCR Primer	Primer for amplification
PCR Master Mix 1	PCR Master mix 1
PCR Master Mix 2	PCR Master mix 2

SMRTbell gDNA Sample Amplification Kit

BEST PRACTICES RECOMMENDATIONS FOR PREPARING ULTRA-LOW DNA INPUT SMRTBELL LIBRARIES

Sample Processing Recommendations

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

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

Megaruptor System

g-TUBE

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

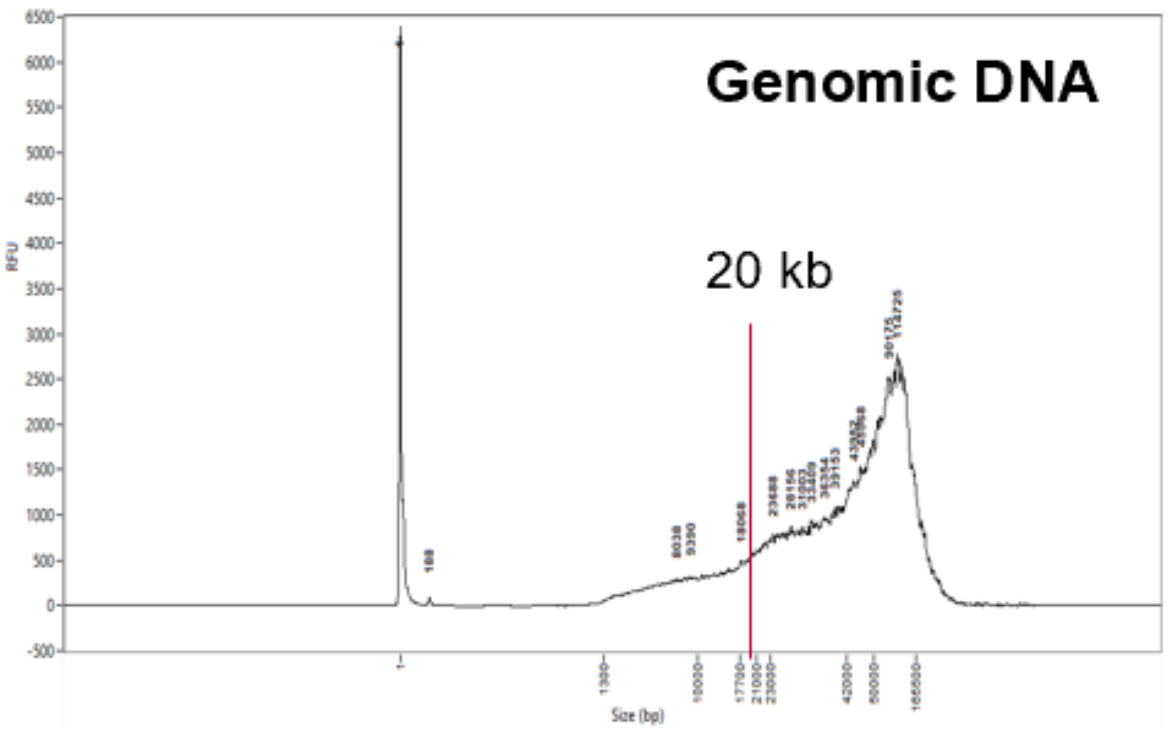
ULTRA-LOW DNA INPUT LIBRARY RECOMMENDATIONS	MEGARUPTOR SYSTEM OR g-TUBE DEVICE
Recommended Starting Input gDNA Amount for Shearing	**5 – 20 ng**
Target gDNA Shear Size	**10 kb**

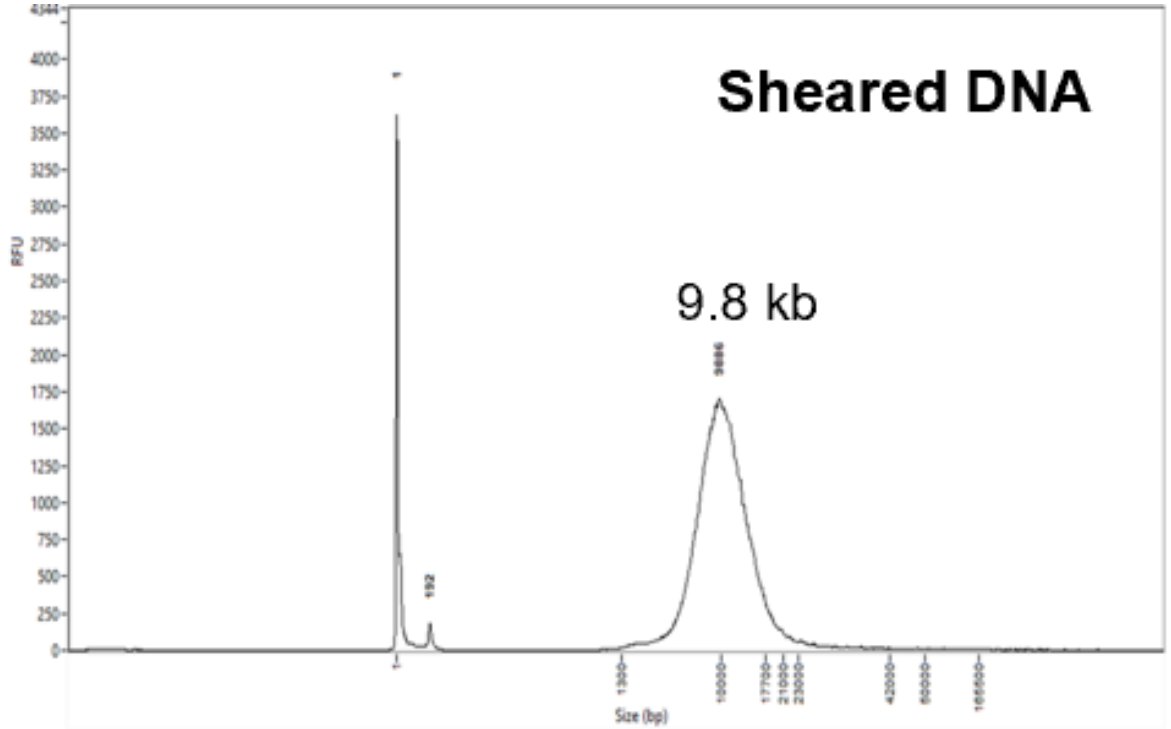
For shearing, it is important to note that **5 ng is the minimum starting DNA input**. If you have more than 5 ng DNA, we recommend shearing more than 5 ng of input DNA to ensure sufficient sample to perform accurate DNA sizing and quantification.

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.



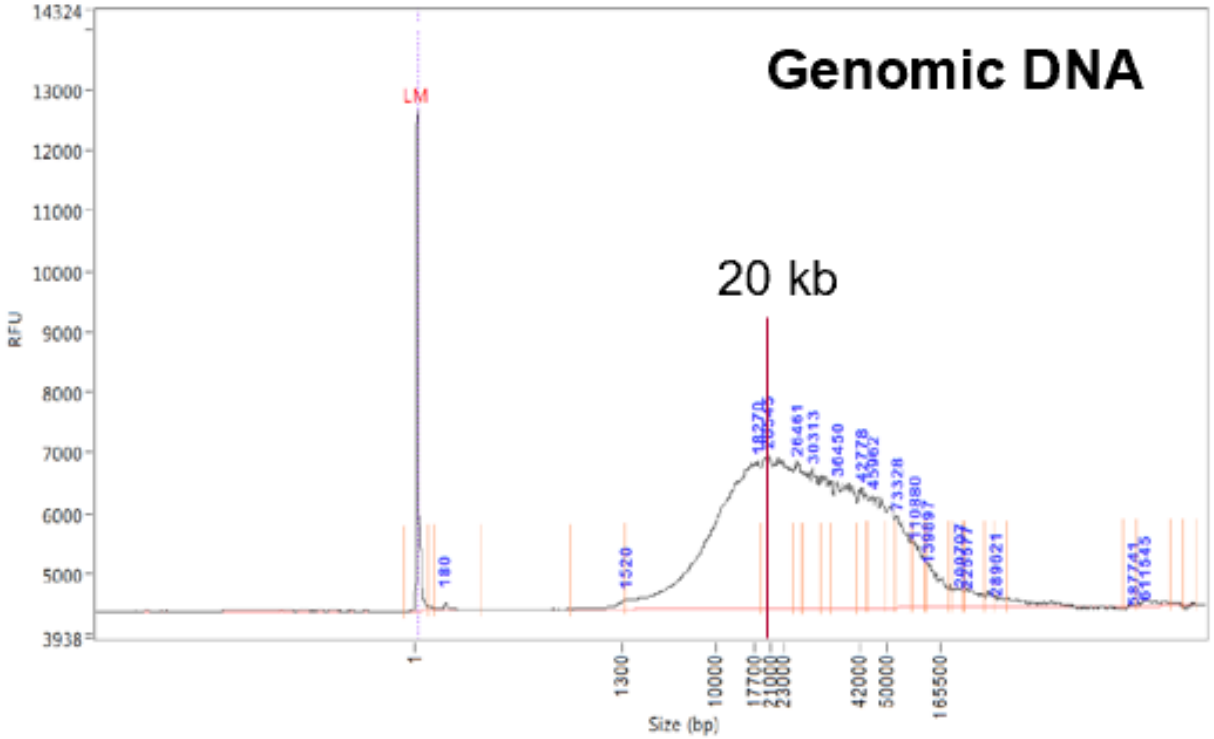


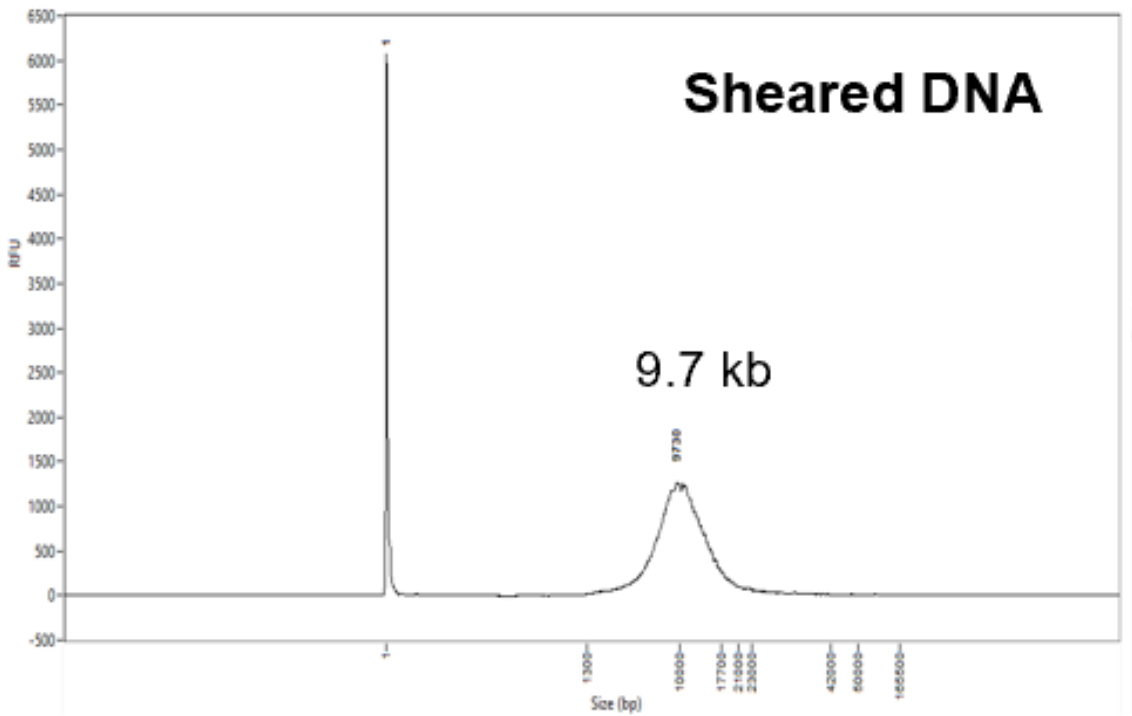
Example Femto Pulse sizing QC analysis of input gDNA and sheared gDNA samples. 5 ng of human genomic DNA was sheared to approximately 10 kb using a Covaris g-TUBE (1677 *x g* in a Eppendorf MIniSpin Plus centrifuge). The size distribution of the starting input genomic DNA is >20 kb.

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 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.





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	Tube Cap Color	Volume	✓
PCR Master Mix 1		50.0 µL	
Sample Amplification PCR Primer		2.0 µL	
Total Volume		52.0 µL	

Reaction Mix 5B	Tube Cap Color	Volume	✓
PCR Master Mix 2		50.0 µL	
Sample Amplification PCR Primer		2.0 µL	
Total Volume		52.0 µL	

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.

PCR Program for Reaction Mix 5A	
45 seconds at 98°C	**1 cycle**
10 seconds at 98°C	**13 cycles**
15 seconds at 62°C	
7 minutes at 72°C	
5 minutes at 72°C	**1 cycle**
Hold at 4°C	

PCR Program for Reaction Mix 5B	
30 seconds at 98°C	**1 cycle**
10 seconds at 98°C	**13 cycles**
15 seconds at 60°C	
10 minutes at 68°C	
5 minutes at 68°C	**1 cycle**
Hold at 4°C	

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.

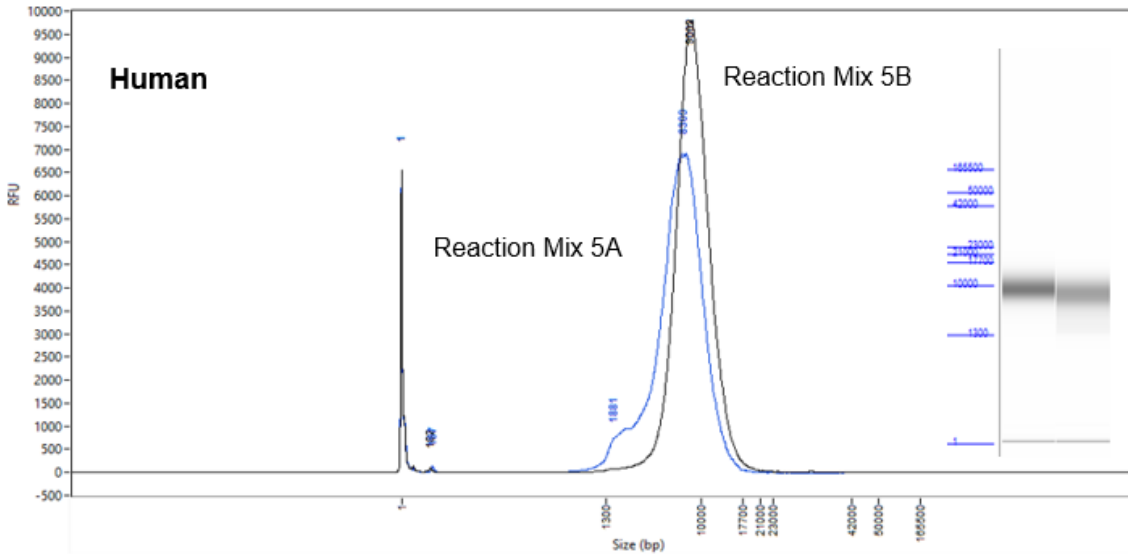
IMPORTANT: You must have the required yield (mass) of purified PCR-amplified DNA per reaction to proceed with "Pool Amplified DNA"; see guidelines in the table below.

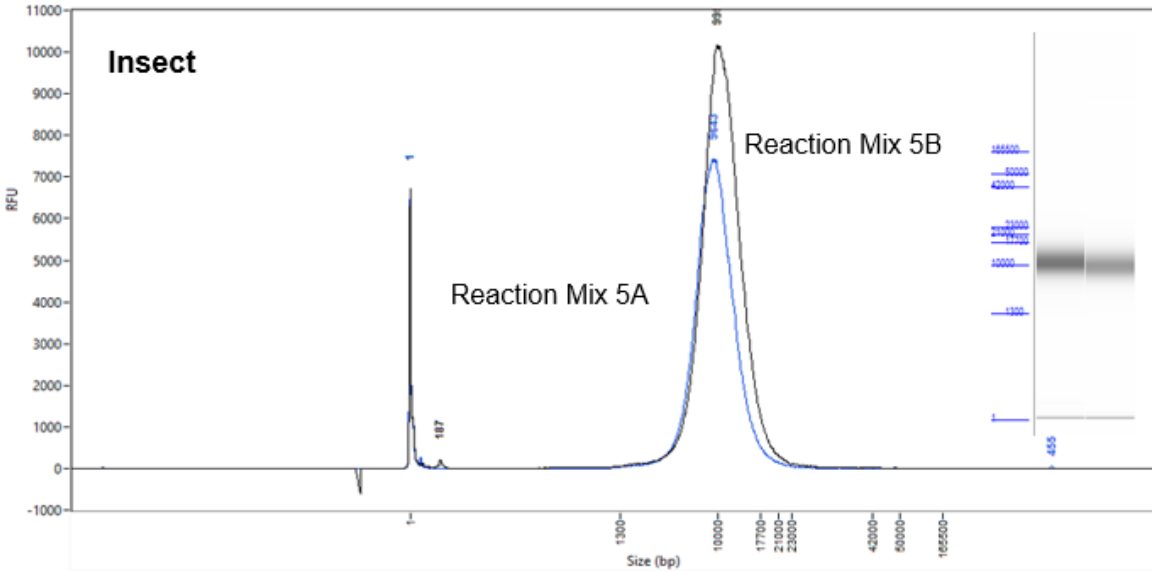
RECOMMENDATIONS FOR RE-AMPLIFYING SAMPLES WITH LOW PCR PRODUCT YIELDS.

DESIRED # OF SMRT CELLS TO RUN	TOTAL YIELD OF AMPLIFIED DNA REQUIRED FOR LIBRARY CONSTRUCTION	FOR SAMPLES WITH LOW PCR PRODUCT YIELD (PCR REACTION A OR B)	RECOMMENDED # OF ADDITIONAL PCR CYCLES (SEE APPENDIX 1)
1 SMRT Cell 8M	**500 ng (PCR Reaction A + B)**	**If the mass of amplified DNA for PCR Reaction A or B is <275 ng (<11 ng/µL)**	Perform **2** additional PCR cycles
		If the mass of amplified DNA for PCR Reaction A or B is <130 ng (<5 ng/µL)	Perform **3** additional PCR cycles
		If the mass of amplified DNA for PCR Reaction A or B is < 65 ng (<2.5 ng/µL)	Perform **5** additional PCR cycles
2 SMRT Cells 8M	**800 ng (PCR Reaction A + B)**	**If the mass of amplified DNA for PCR Reaction A or B is <400 ng (<15 ng/µL)**	Perform **4 to 5** additional PCR cycles

PURIFICATION OF PCR-AMPLIFIED DNA USING PRONEX BEADS (CONT.)

Example Femto Pulse sizing QC analysis of PCR-amplified human and insect DNA samples after purification with ProNex Beads





Example Femto Pulse sizing QC analysis of analysis of PCR-amplified 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.

STEP	✓	Pooling	Notes
1		Pool amplified DNA into a single PCR tube of an 8-tube strip.	
2		Mix and spin down the contents of the tube with a quick spin in a microfuge.	
3		Proceed to the "Repair DNA Damage" section below.	

SMRTBELL LIBRARY SIZE SELECTION USING THE BLUEPIPPIN SYSTEM

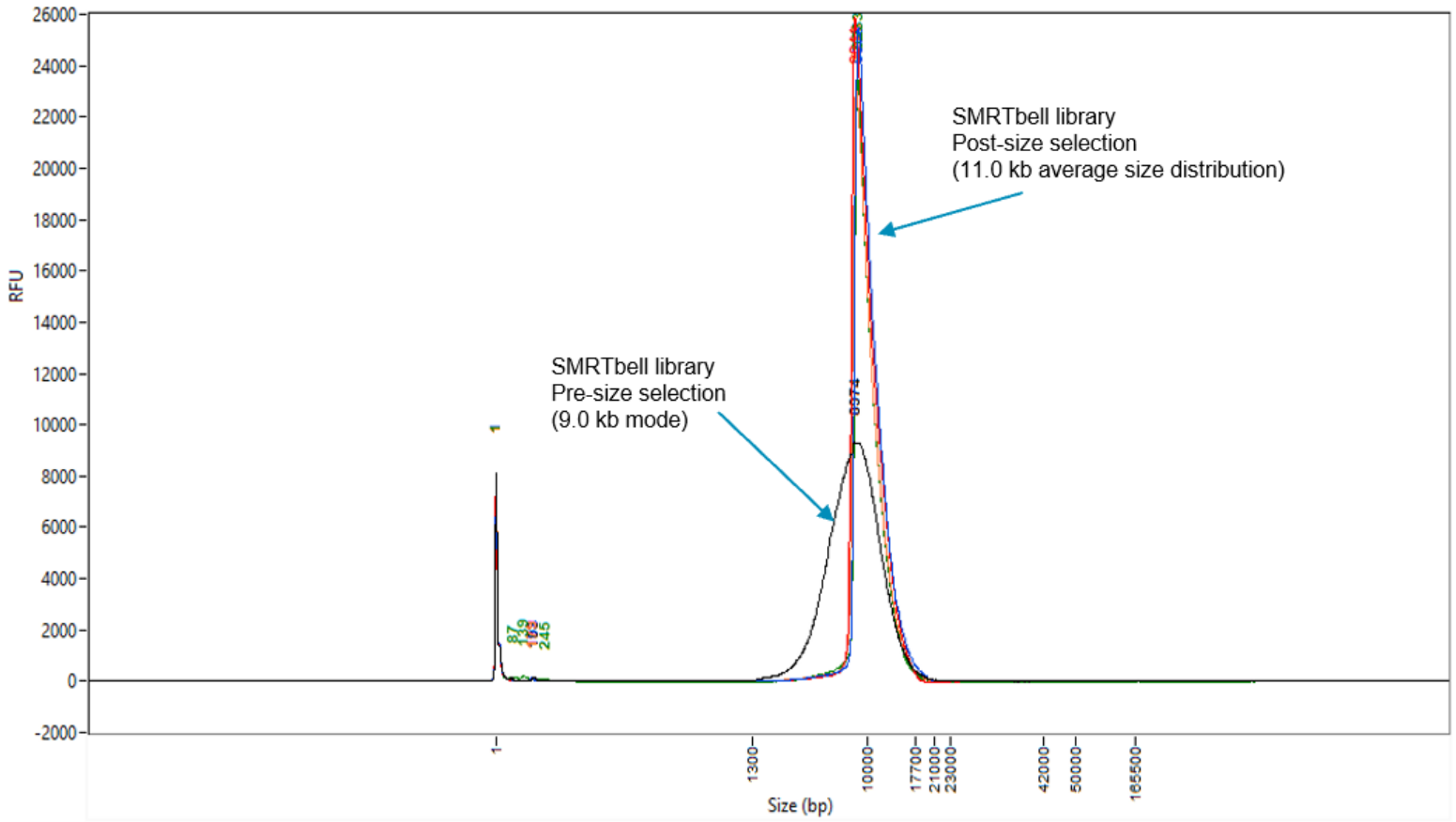
For constructing SMRTbell libraries from ultra-low DNA input, PacBio highly recommends size-selection using the BluePippin system.

- 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.

SMRTBELL LIBRARY SIZE SELECTION USING THE BLUEPIPPIN SYSTEM (CONT.)

Example Femto Pulse sizing QC analysis of final ultra-low DNA input SMRTbell library after size selection with a BluePippin System



Example Femto Pulse sizing QC analysis of an ultra-low DNA input SMRTbell library after size-selection using Sage Science's 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:

SMRT LINK SAMPLE SETUP PARAMETER	VALUE TO ENTER
Sequencing Primer	Sequencing Primer v4
Binding Kit	Sequel II Binding Kit 2.0
Sequencing Mode	CCS
Iso-Seq Experiment*	Yes
Iso-Seq Version*	Yes
Cleanup Anticipated Yield**	50%

* 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:

SMRT LINK RUN DESIGN PARAMETER	VALUE TO ENTER
Sequencing Mode	CCS
On-Plate Loading Concentration	50 – 60 pM
Pre-Extension Time	2 hours
Movie Time per SMRT Cell	30 hours
Generate CCS Data*	YES

* 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.

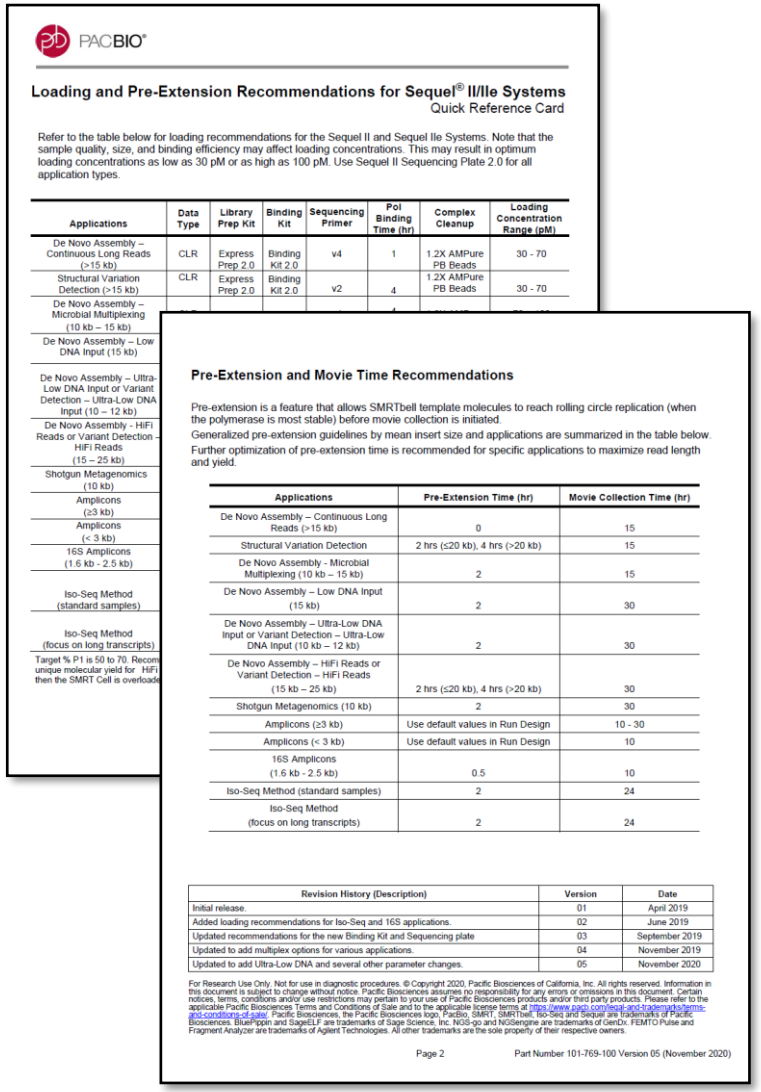
SAMPLE SETUP RECOMMENDATIONS FOR ULTRA-LOW DNA INPUT LIBRARIES – SEQUEL II AND IIe SYSTEMS (CHEMISTRY 2.0 / SMRT LINK V10.0)

- 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/IIe 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

Applications	Data Type	Library Prep Kit	Binding Kit	Sequencing Primer	Pol Binding Time (hr)	Complex Cleanup	Loading Concentration Range (pM)
De Novo Assembly – Ultra-Low DNA Input or Variant Detection – Ultra-Low DNA Input (10 – 12 kb)	CCS	Express Prep 2.0	Binding Kit 2.0	v4	1	1.2X ProNex Beads	50 - 70

Applications	Pre-Extension Time (hr)	Movie Collection Time (hr)
De Novo Assembly – Ultra-Low DNA Input or Variant Detection – Ultra-Low DNA Input (10 kb – 12 kb)	2	30


PACBIO®

Loading and Pre-Extension Recommendations for Sequel® II/IIe Systems
Quick Reference Card

Refer to the table below for loading recommendations for the Sequel II and Sequel IIe Systems. Note that the sample quality, size, and binding efficiency may affect loading concentrations. This may result in optimum loading concentrations as low as 30 pM or as high as 100 pM. Use Sequel II Sequencing Plate 2.0 for all application types.

Applications	Data Type	Library Prep Kit	Binding Kit	Sequencing Primer	Pol Binding Time (hr)	Complex Cleanup	Loading Concentration Range (pM)
De Novo Assembly – Continuous Long Reads (>15 kb)	CLR	Express Prep 2.0	Binding Kit 2.0	v4	1	1.2X AMPure PB Beads	30 - 70
Structural Variation Detection (>15 kb)	CLR	Express Prep 2.0	Binding Kit 2.0	v2	4	1.2X AMPure PB Beads	30 - 70
De Novo Assembly – Microbial Multiplexing (10 kb – 15 kb)							
De Novo Assembly – Low DNA Input (15 kb)							
De Novo Assembly – Ultra-Low DNA Input or Variant Detection – Ultra-Low DNA Input (10 – 12 kb)							
De Novo Assembly - HiFi Reads or Variant Detection – HiFi Reads (15 – 25 kb)							
Shotgun Metagenomics (10 kb)							
Amplicons (≥3 kb)							
Amplicons (< 3 kb)							
16S Amplicons (1.6 kb - 2.5 kb)							
Iso-Seq Method (standard samples)							
Iso-Seq Method (focus on long transcripts)							

Target % P1 is 50 to 70. Recom... unique molecular yield for HiFi ... then the SMRT Cell is overloade...

Pre-Extension and Movie Time Recommendations

Pre-extension is a feature that allows SMRTbell template molecules to reach rolling circle replication (when the polymerase is most stable) before movie collection is initiated.
Generalized pre-extension guidelines by mean insert size and applications are summarized in the table below.
Further optimization of pre-extension time is recommended for specific applications to maximize read length and yield.

Applications	Pre-Extension Time (hr)	Movie Collection Time (hr)
De Novo Assembly – Continuous Long Reads (>15 kb)	0	15
Structural Variation Detection	2 hrs (≤20 kb), 4 hrs (>20 kb)	15
De Novo Assembly - Microbial Multiplexing (10 kb – 15 kb)	2	15
De Novo Assembly – Low DNA Input (15 kb)	2	30
De Novo Assembly – Ultra-Low DNA Input or Variant Detection – Ultra-Low DNA Input (10 kb – 12 kb)	2	30
De Novo Assembly – HiFi Reads or Variant Detection – HiFi Reads (15 kb – 25 kb)	2 hrs (≤20 kb), 4 hrs (>20 kb)	30
Shotgun Metagenomics (10 kb)	2	30
Amplicons (≥3 kb)	Use default values in Run Design	10 - 30
Amplicons (< 3 kb)	Use default values in Run Design	10
16S Amplicons (1.6 kb - 2.5 kb)	0.5	10
Iso-Seq Method (standard samples)	2	24
Iso-Seq Method (focus on long transcripts)	2	24

Revision History (Description)	Version	Date
Initial release.	01	April 2019
Added loading recommendations for Iso-Seq and 16S applications.	02	June 2019
Updated recommendations for the new Binding Kit and Sequencing plate	03	September 2019
Updated to add multiplex options for various applications.	04	November 2019
Updated to add Ultra-Low DNA and several other parameter changes.	05	November 2020

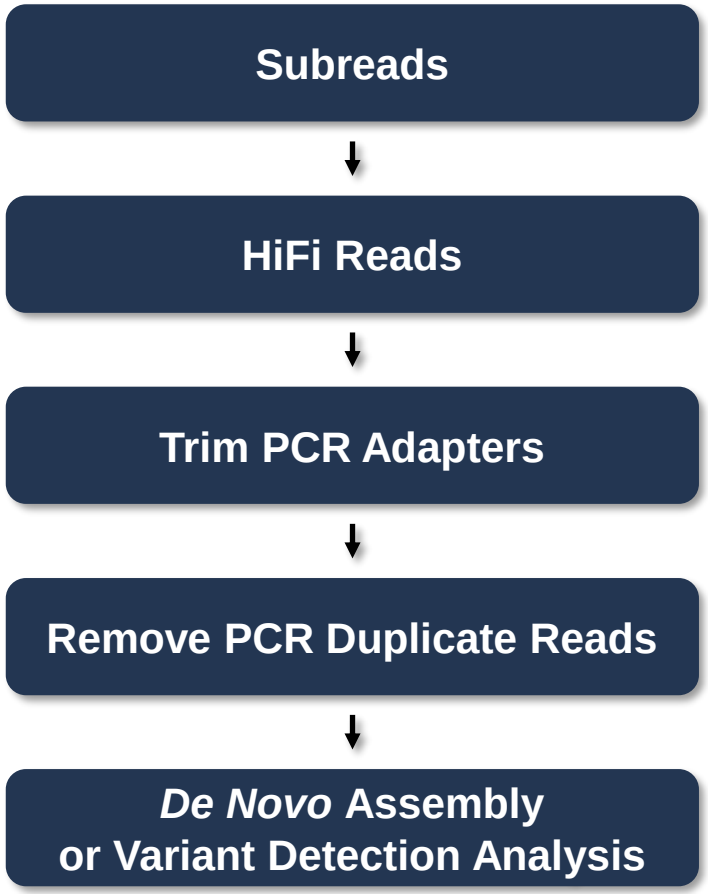
Part Number 101-769-100 Version 05 (November 2020)

https://www.pacb.com/support/documentation/

HiFi Sequencing Data Analysis Recommendations for Ultra-Low DNA Input Libraries

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***.



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 IIe 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

  → $Target\ HiFi\ Base\ Yield = [Sample\ Haploid\ Genome\ Size\ (Gb)]\ x\ [Target\ HiFi\ Coverage\ per\ Sample]$

  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

HIFI SEQUENCING DATA ANALYSIS RECOMMENDATIONS FOR ULTRA-LOW DNA INPUT LIBRARIES (CONT.)

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 IIe 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
  - → $Target\ HiFi\ Base\ Yield = [Sample\ Haploid\ Genome\ Size\ (Gb)]\ x\ [Target\ Coverage\ per\ Sample]$

    E.g., For variant detection analysis of a human genome (3 Gb):

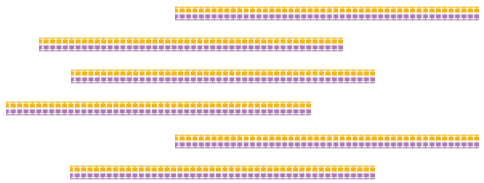
    Recommended Minimum Target HiFi Base Yield = 3 Gb x 15 = 45 Gb
- For detection of structural variants (>50 bp):
  - Start with the trimmed, deduplicated HiFi data set and run the **Structural Variant Calling** application in SMRT Analysis.
  - Under the advanced settings make the following parameter 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.

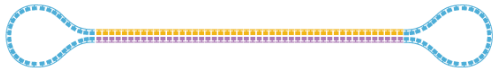
Ultra-Low DNA Input Library Example Performance Data

EXAMPLE 1: *DROSOPHILA* ULTRA-LOW DNA INPUT LIBRARY FOR *DE NOVO* GENOME ASSEMBLY

Sample Library Preparation and Sequencing Workflow

Image Credit: Shutterstock





Drosophila melanogaster insect sample collection

↓

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

↓

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)

↓

SMRTbell library construction with 500 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

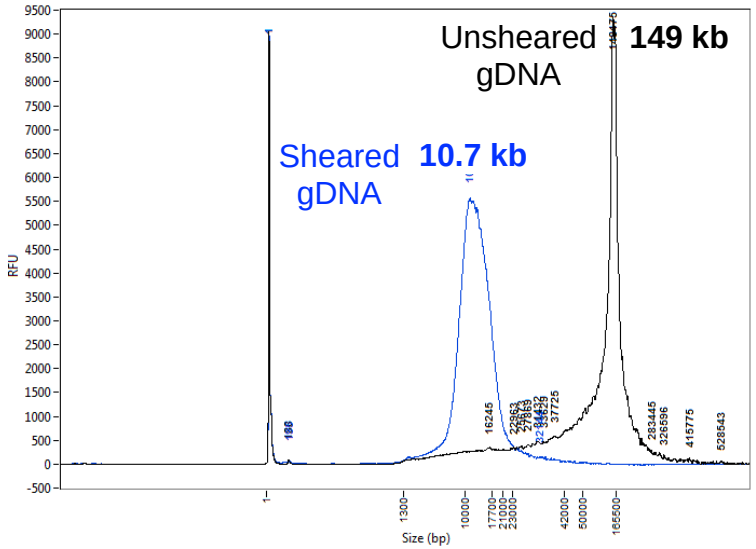
↓

A single SMRT Cell 8M produced **26.8 Gb of HiFi data**, or ~183-fold processed HiFi coverage of the *Drosophila* genome

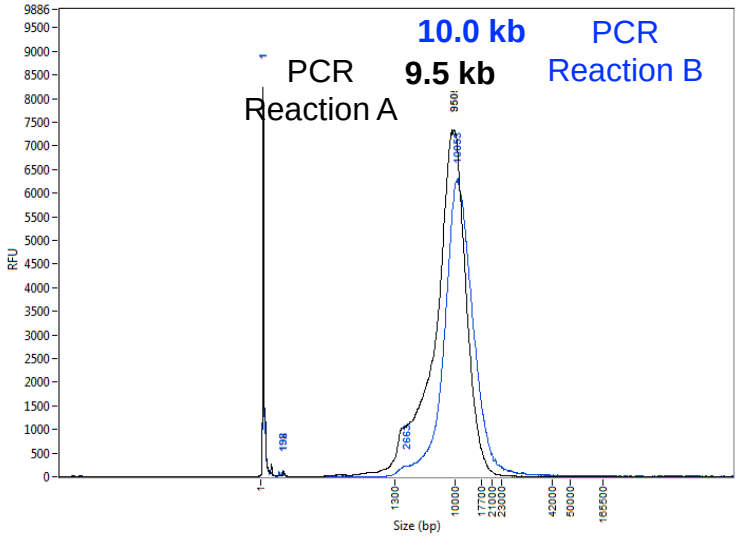
PacBio Application Note: Considerations for Using the Low and Ultra-Low DNA Input Workflows for Whole Genome Sequencing (PN 101-995-900)

EXAMPLE 1: *DROSOPHILA* ULTRA-LOW DNA INPUT LIBRARY FOR *DE NOVO* GENOME ASSEMBLY

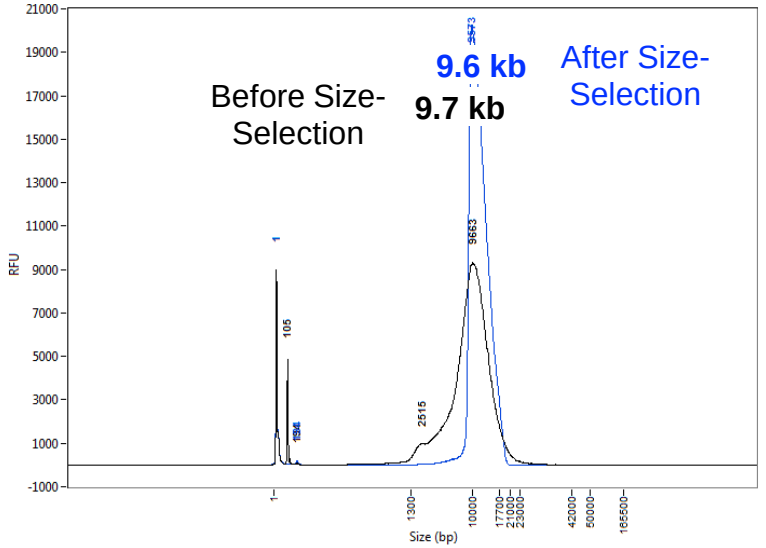
Sample Library Preparation DNA Quantitation & Sizing QC Results for *Drosophila* Ultra-Low DNA Input Sample



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.

Sample	PCR Reaction	PCR Yield	# of Cycles	Mass Used for Pooling	Total Pooled Mass	Library Yield (Pre-BP Size Selection)		Library Yield (Post-BP Size Selection)		Mean Library Insert Size
Drosophila melanogaster	A	343 ng	17	250 ng	500 ng	432 ng	86%	170 ng	40%	11.5 kb
	B	260 ng	13	250 ng						

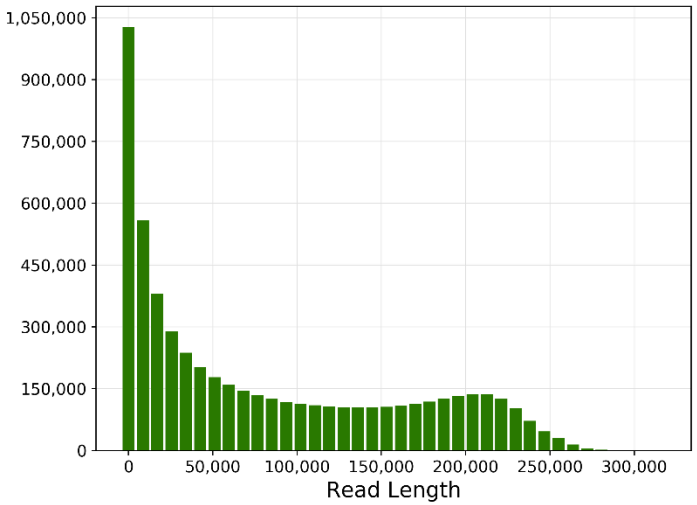
PacBio Application Note: Considerations for Using the Low and Ultra-Low DNA Input Workflows for Whole Genome Sequencing (PN 101-995-900)

EXAMPLE 1: *DROSOPHILA* ULTRA-LOW DNA INPUT LIBRARY FOR *DE NOVO* GENOME ASSEMBLY

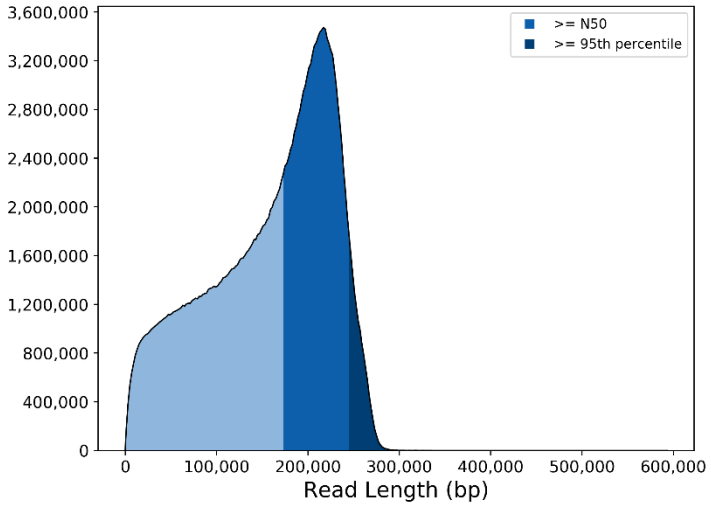
Primary Sequencing Metrics for *Drosophila* Ultra-Low DNA Input Library

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

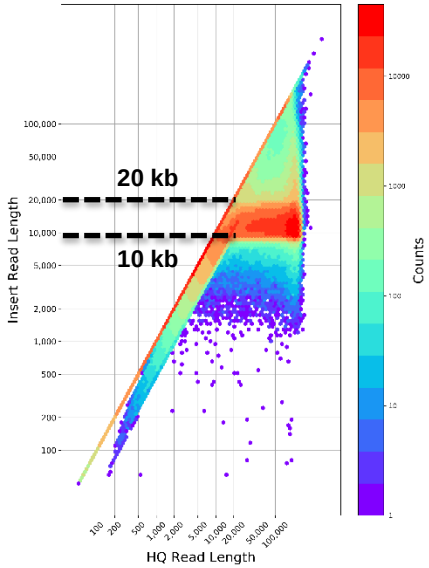
Polymerase Read Length Distribution



Base Yield Density



Insert Read Length Density



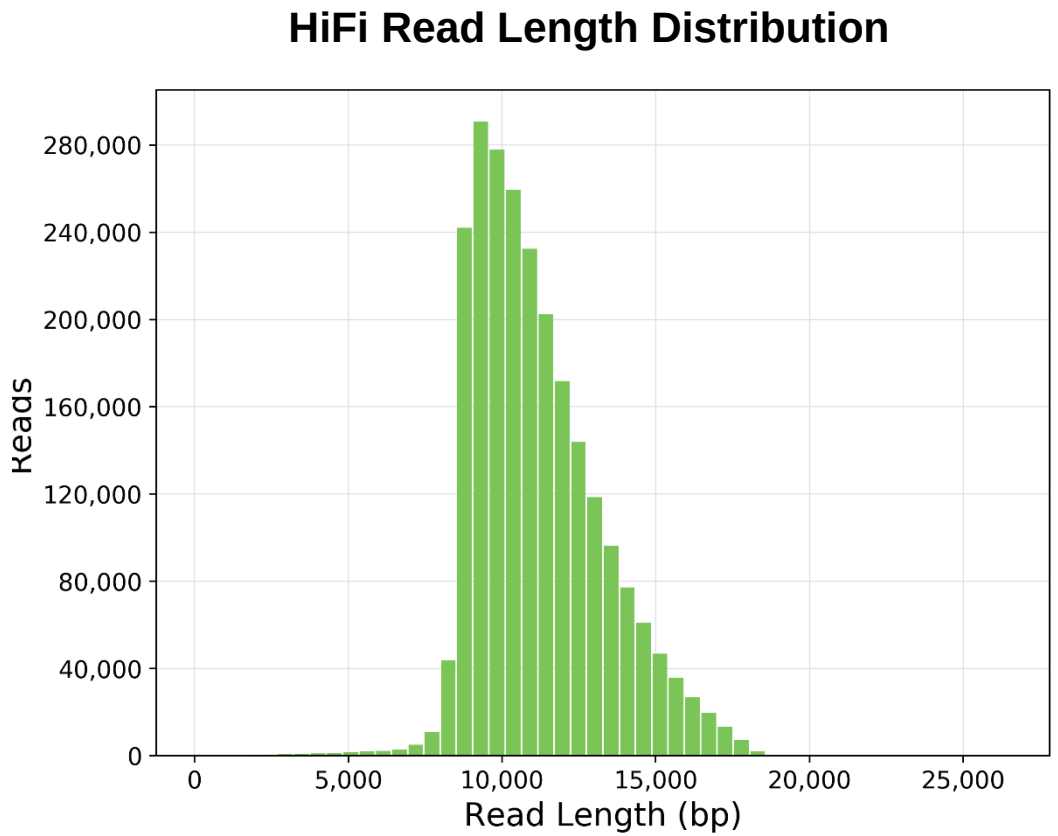
Example sequencing results shown were generated using Sequel II System Chemistry 2.0.

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)**

METRIC	VALUE
HiFi Data Yield	26.8 Gb
Mean HiFi Read Length	11,153 bp
Median HiFi Read Quality	Q34
PCR Adapter Percentage	99.85%
PCR Duplication Rate	3.78%
Processed HiFi Coverage for 140 Mb Genome	183-Fold



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

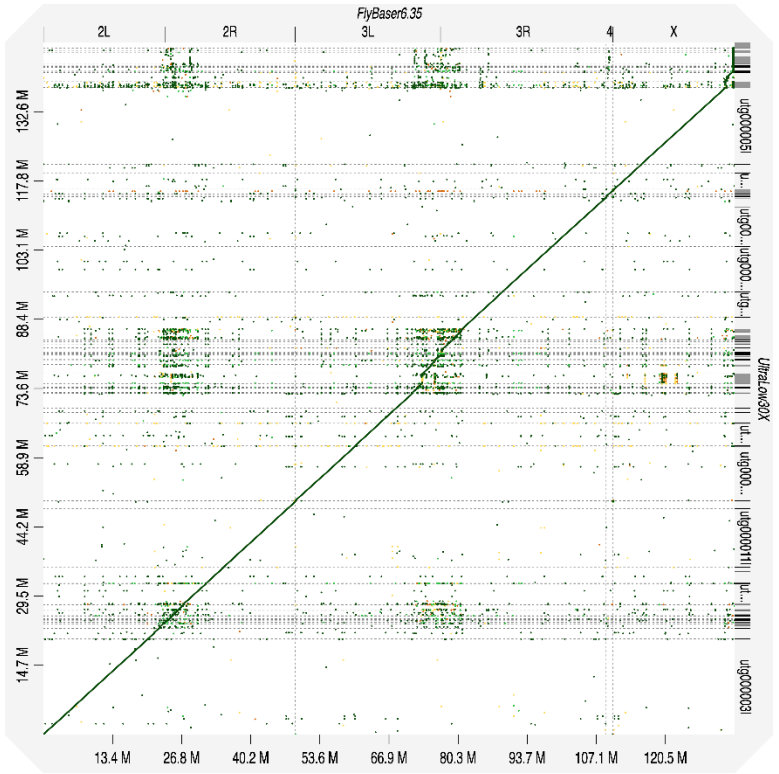
HiFi *De Novo* Assembly Performance Results for *Drosophila* Ultra-Low DNA Input Library

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

METRIC	VALUE
Down sampled HiFi Coverage	30-fold
Assembly Size	147 Mb
Contig N50	8.3 Mb
BUSCO Complete	98.6%

HiFi assembly using Hifiasm.

***Drosophila* ULI Assembly vs. FlyBaseRef 6.35**

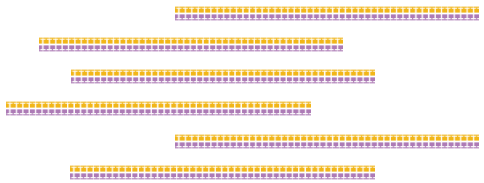


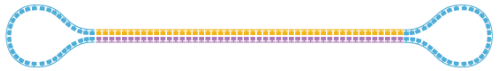
Dot plot analysis of *Drosophila* ultra-low DNA input genome assembly versus FlyBaseRef 6.35 reference sequence. The ULI assembly shows a high level of congruence with the FlyBase reference sequence.

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

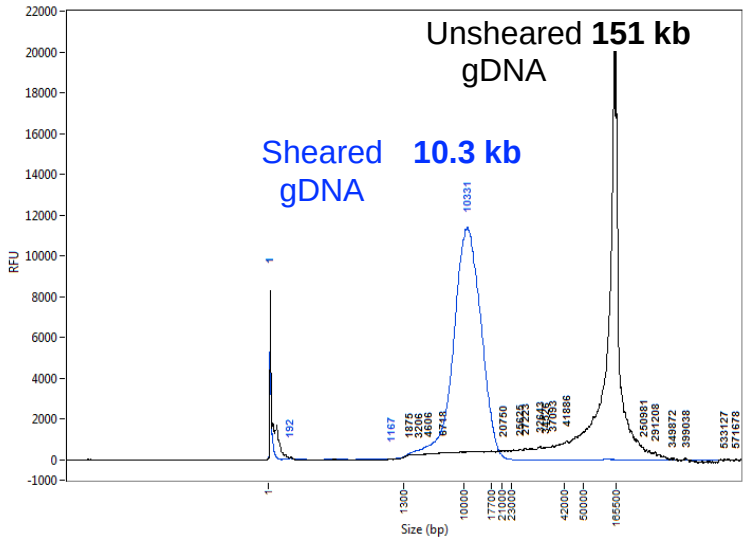
↓

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

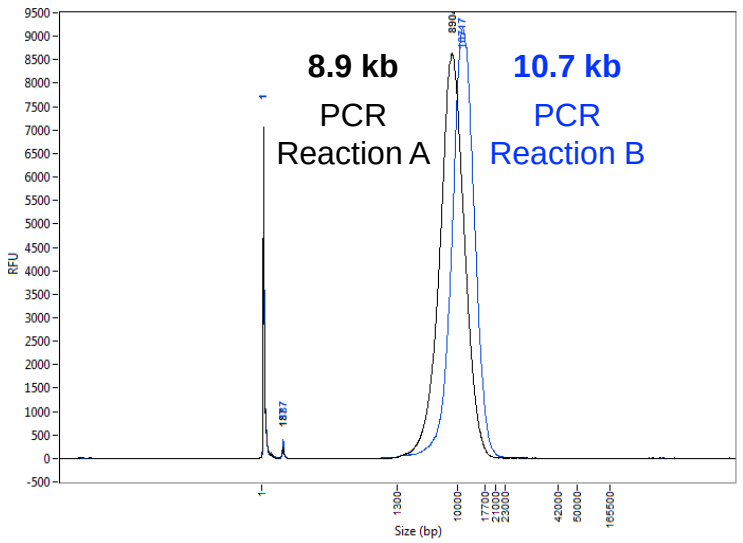
PacBio Application Note: Considerations for Using the Low and Ultra-Low DNA Input Workflows for Whole Genome Sequencing (PN 101-995-900)

EXAMPLE 2: CORIELL HG002 ULTRA-LOW DNA INPUT LIBRARY FOR HUMAN VARIANT DETECTION (CONT.)

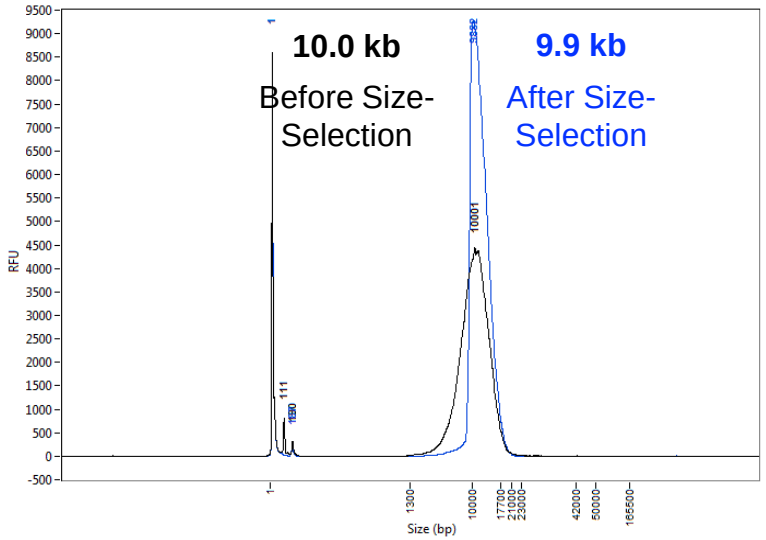
Sample Library Preparation DNA Quantitation & Sizing QC Results for HG002 Ultra-Low DNA Input Library



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.

Sample	PCR Reaction	PCR Yield	# of Cycles	Mass Used for Pooling	Total Pooled Mass	Library Yield (Pre-BP Size Selection)		Library Yield (Post-BP Size Selection)		Mean Library Insert Size	# of SMRT Cells 8M
HG002	A	811 ng	13	405 ng	808 ng	636	78%	236 ng	37%	11.0 kb	2
	B	806 ng	13	403 ng							

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

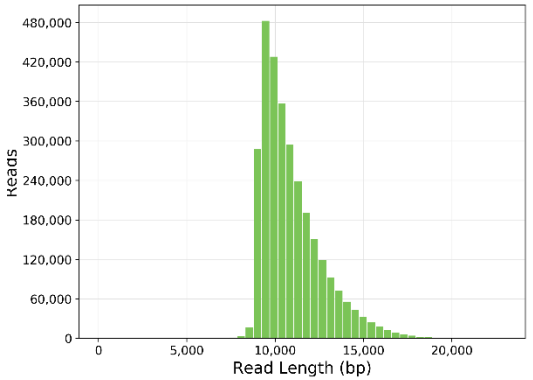
EXAMPLE 2: CORIELL HG002 ULTRA-LOW DNA INPUT LIBRARY FOR HUMAN VARIANT DETECTION (CONT.)

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

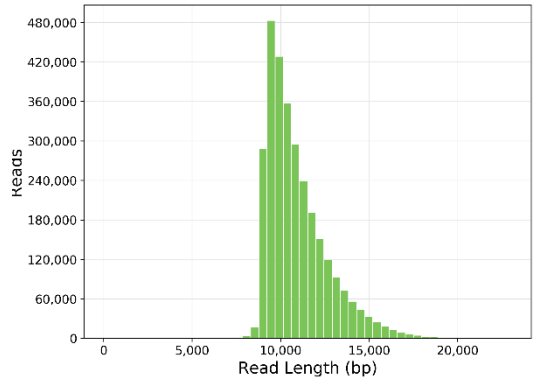
HG002 HIFI SEQUENCING YIELD METRICS (2 SMRT CELLS 8M)

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

HiFi Read Length Distribution



HiFi Read Length Distribution



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

HG002 VARIANT DETECTION BENCHMARKING RESULTS (2 SMRT CELLS 8M).

VARIANT TYPE	VARIANT CALLER	PRECISION	RECALL
Single Nucleotide Variants	DeepVariant	99.6 %	99.0%
Indels (<49 bp)	DeepVariant	84.7%	90.3%
Structural Variants (≥50 bp)	PBSV	95.4%	84.3%

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

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/IIe 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)

TECHNICAL DOCUMENTATION & APPLICATIONS SUPPORT RESOURCES FOR ULTRA-LOW DNA INPUT LIBRARY PREPARATION, SEQUENCING & DATA ANALYSIS (CONT.)

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

WHOLE GENOME SEQUENCING APPLICATION	DATASET	DATA TYPE	PACBIO SYSTEM
Assembly (Ultra-Low DNA Input)	*Phlebotomus papatasi* (Sand Fly) – Male [SRR12454518]	HiFi Reads	Sequel II System
	Drosophila melanogaster (Fruit Fly) – Female [SRR12473480]	HiFi Reads	Sequel II System
Variant Detection (Ultra-Low DNA Input)	*Homo sapiens* (HG002) – Male [SRR12454519]	HiFi Reads	Sequel II System
	Homo sapiens (HG002) – Male [SRR12454520]	HiFi Reads	Sequel II System

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

- Schneider, C. et al. (2020) Biodiversity genomics of small metazoa: high quality *de novo* genomes from single specimens of field-collected and ethanol-preserved springtails. BioRxiv Preprint. doi: https://doi.org/10.1101/2020.08.10.244541. Posted August.11.2020.

Webinars

- PacBio Virtual Global Summit Presentation (2020): Technical Tutorial – Choosing the best library prep for HiFi sequencing. [Webinar Recording]
- PacBio Webinar (2020): No Organism Too Small: Build High-Quality Genome Assemblies of Small Organisms with HiFi Sequencing [Webinar Recording]
- SMRT Leiden Presentation (2020): Revealing the Genome Biology of Earth's Smallest Animals. [Webinar Recording]

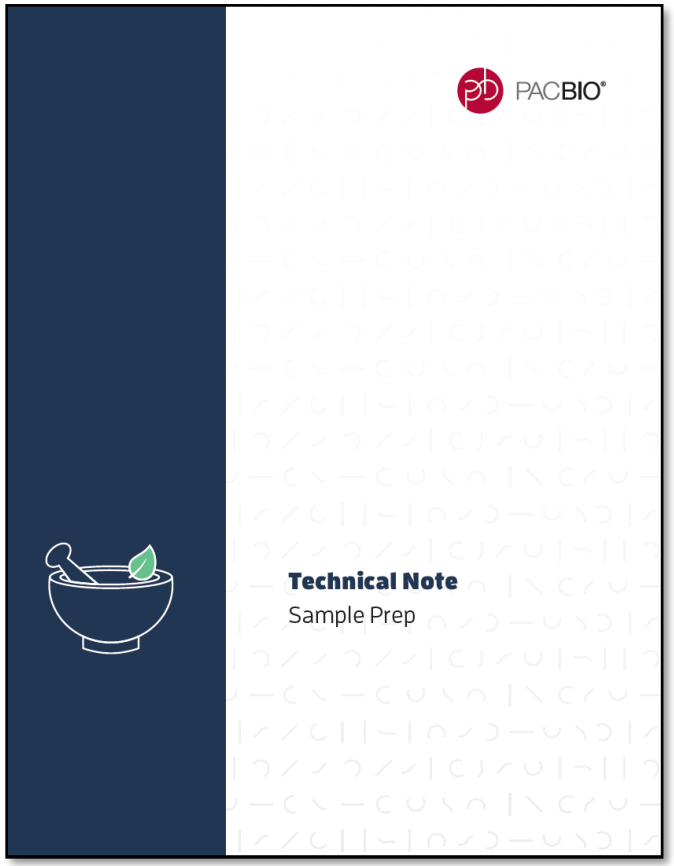
Appendix: General Recommendations for High-Molecular Weight gDNA QC and Handling for SMRTbell Library Construction

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



Technical Note (TN100-040518): Preparing samples for PacBio whole genome sequencing for *de novo* assembly – Collection and storage

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

Technical Note TN101-061920: Preparing DNA for PacBio HiFi sequencing – Extraction and quality control

DNA SAMPLE PREPARATION ONLINE RESOURCE

Literature resource for sample collection and DNA extraction protocol references

The listing below is a collection of publications by the scientific community describing extraction protocols for high-molecular weight DNA followed by PacBio sequencing. When possible, the links point directly to the methods section (or supplementary information).

Animals | Plants | Fungi | Protists

If you have protocols you would like to share, or have questions about DNA extraction for PacBio sequencing, contact ExtractDNA@pacb.com.

Animals

- Invertebrates
  - Panova2016 – DNA extraction protocols for whole-genome sequencing in marine organisms
  - *microinvertebrates*
    - Laumer2020 – protocols for diverse meiofauna species, including *C. elegans* & *Castrella truncata* (SMRT Leide… confer… presentation)
  - *arthropods*
    - arachnids
      - Guerrero2019 – The Pacific Biosciences de novo assembled genome dataset from a parthenogenetic New Zealand wild population of the longhorned tick, *Haemaphysalis longicornis* Neumann, 1901
      - Liu2019 – DNA Methylation Patterns in the Social Spider, *Stegodyphus dumicola*
      - Burgess2018 – Draft genome assembly of the sheep scab mite, *Psoroptes ovis*
      - Randall2018 – The draft genome assembly of *Dermatophagoides pteronyssinus* supports identification of novel allergen isoforms in *Dermatophagoides* species

…ous, Annotated Genome
…the Sex Chromosome
…phnia

Plants

…t genomic DNA for long-read sequencing of single molecules
…ce Genome With Single-Molecule Sequencing Uncovers a Recent Burst of …Genes
…ssembly reveals the structure of the *Arabidopsis thaliana* Nd-1 genome and its gene set
…sive, high-…
…consiste… …ng and se…
…ect run at PacBio (using the Circulomics Nanobind Plant Nuclei Big DNA Kit)
…olly pine mega-genome using long-read single-molecule sequencing
…ice for the Fast-Growing Microalga *Picochlorum celeri*
Undaria pinnatifida: Chromosome-Level Assembly Using PacBio and Hi-C
…e suggests convergent functions of homeobox genes in algae and land plants (algae)
- Panova2016 – DNA extraction protocols for whole-genome sequencing in marine organisms (algae)

Fungi

- Faure2019 – Long-Read Genome Sequence of the Sugar Beet Rhizosphere Mycoparasite *Pythium oligandrum*
- Nagappan2018 – Improved nucleic acid extraction protocols for *Ganoderma boninense*, *G. miniatocinctum* and *G. tornatu*
- Schwessinger2017 – Extraction of high molecular weight DNA from fungal rust spores for long read sequencing
- Solomon2016 – Robust and effective methodologies for cryopreservation and DNA extraction from anaerobic gut fungi
- Sonnenberg2016 – A detailed analysis of the recombination landscape of the button mushroom *Agaricus bisporus var. bisporus*

Methods

Sample collection

A female yellowbelly pufferfish (Fig. 2), reared in the fish breeding centre of Fujian Normal University in Fuzhou City of Fujian Province was used for genome sequencing and assembly. Fresh white muscle, eye, skin, gonad, gut, liver, kidney, blood, gall bladder and air bladder tissues were collected and quickly frozen in liquid nitrogen for one hour. White muscle tissues were used for DNA sequencing for genome assembly, while all tissues were used for transcriptome sequencing.

Fig. 2

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

DNA and RNA sequencing

Genomic DNA from white muscle tissue was extracted using the

www.ExtractDNAforPacBio.com

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!

METHODS FOR EVALUATION OF GENOMIC DNA QUALITY

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

A. Genomic DNA Sizing Characterization

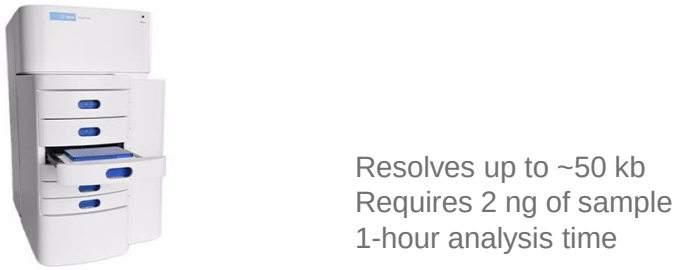
Femto Pulse System (Agilent Technologies)

Highly Recommended

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

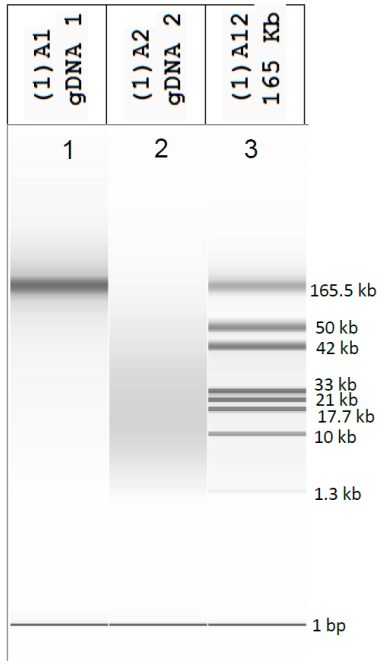
https://www.agilent.com/en/product/automated-electrophoresis/femto-pulse-systems

Fragment Analyzer System (Agilent Technologies)



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

https://www.agilent.com/en/product/automated-electrophoresis/fragment-analyzer-systems



Lane 1: High MW gDNA
Lane 2: Degraded gDNA
Lane 3: 165 kb Ladder

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..

www.pacb.com