PACBIO® GUIDELINES FOR SUCCESSFUL SMRTbell™ LIBRARIES

I. Sample requirements for PacBio sequencing

The Pacific Biosciences® library preparation process does not utilize amplification techniques and resulting library molecules are directly used as templates for the sequencing process. As such, the quality of the DNA starting material will be directly reflected in the sequencing results. Any irreversible DNA damage present in the input material (e.g., interstrand crosslinks, etc.) will result in impaired performance in the system. High-quality, high-molecular-weight genomic DNA is imperative for obtaining long read lengths and optimal sequencing performance.

II. General guidelines for handling high-molecular-weight DNA

In general, the following precautions need to be taken when handling DNA:

- Avoid overdrying of genomic DNA. Allow the DNA to air dry. Do not heat when drying in a speed-vac.
- DNA should be eluted in neutral, buffered solution (e.g., 10 mM Tris Acetate or Tris-HCl, pH 8) and stored in TE (10 mM Tris, pH 8, 1mM EDTA)*. Avoid eluting in RNAse-free H_2O or unbuffered solutions.
- PCR products should be clean amplicons, without non-specific products or multiple bands.
- If gel purification is required, avoid using ethidium/UV based visualization methods. One alternative is to use SYBR® Safe (Invitrogen) and visualize with blue light.
- To help resuspend the DNA, carefully invert the tube several times after adding buffer and/or tap the tube gently.
- Alternatively, allow the DNA to stand in buffer overnight at 25°C to resuspend.
- Overheating can introduce DNA damage. Inactivate DNAase as recommended by the vendor kit. It is best to avoid heat inactivation when possible. An alternative is AMPure® purification.
- Avoid vortexing genomic DNA when possible as vortexing can cause shearing of the DNA.
- DNA storage conditions: 4°C (short-term); –20°C / –80°C (long-term).
- Repeated freezing and thawing of genomic DNA should be avoided as this will lead to DNA shearing.

*Note: EDTA must be removed prior to library preparation. This can be achieved during the initial AMPure purification.

III. Important measures impacting DNA quality

To maximize read length and quality, it is **essential** that the DNA sample:

- is double-stranded; single-stranded DNA cannot be used to generate the sequencing template.
- has not undergone multiple freeze-thaw cycles as they can lead to DNA damage.
- has not been exposed to high temperatures (e.g. > 65° C for 1 hour) or pH extremes (< 6 or > 9).
- has an OD_{260}/OD_{280} ratio of 1.8 to 2.0.
- has an OD_{260}/OD_{230} ratio of ~2.0.
- does not contain insoluble material.
- does not contain RNA contamination.
- has not been exposed to intercalating fluorescent dyes or ultraviolet radiation. SYBR dyes are not DNA damaging, but avoid ethidium bromide.
- does not contain denaturants (e.g., guanidinium salts or phenol) or detergents (e.g., SDS or Triton-X100).
- does not contain carryover contamination from the original organism/tissue (e.g., heme, humic acid, polyphenols, etc.)

IV. DNA sample quality assessment:

A thorough DNA quality check is required prior to submitting DNA for PacBio sequencing. The following recommendations to ascertain DNA integrity, purity, and concentration are recommended:

1) **Gel images of DNA sample:** Genomic DNA integrity can be assessed by agarose gel electrophoresis. For best results, DNA samples must show no signs of degradation, which is evidenced by smeared DNA bands. The presence of one predominant band showing high MW DNA with no degradation is optimal. A good practice is to indicate relevant marker sizes, and the amount of sample loaded in the agarose gel. For amplicon or cDNA samples, a Bioanalyzer® trace can be used as an alternative.

   If gel purification of your DNA sample is required, we recommend using SYBR Gold or SYBR Safe coupled with blue light for visualization. Do not use ethidium bromide and/or UV light since they can induce DNA damage. The SYBR stains can be easily removed from nucleic acids during the gel-extraction process, e.g., using Qiagen® gel-extraction kits.

2) **Purity of your DNA sample:** DNA purity can be determined by using the NanoDrop® instrument or other spectrophotometers. Readings of both A260:A280 and A260:A230 ratios need to be obtained:

   260/280:
   The ratio of absorbance at 260 nm and 280 nm is used to assess the purity of DNA. A ratio of ~1.8 is generally accepted as "pure" for DNA, but is dependent on the nucleotide composition of the submitted sample.

   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.

   Ensure DNA measurements are conducted in a buffered environment such as (TE or Tris HCl, pH8). Measurements are sensitive to small changes in the pH of the solution which will cause the 260/280 ratio to vary. Acidic solutions will skew the 260/280 ratio lower, while basic solutions will skew the ratio higher.

   260/230:
   The 260/230 ratios provides a secondary measurement of DNA purity to make inferences about the quality of sample extraction. Readings to determine purity are often higher than the respective 260/280 values. Expected 260/230 values are commonly in the range of 2.0-2.2. Abnormal 260/230 values may indicate a problem with the sample extraction procedure.

   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.

   The blank solution should be the same pH and of a similar ionic strength as the sample solution.

3) **Concentration of your DNA sample:** Accurate quantitation of DNA concentration is critical for the PacBio® template preparation procedures. Traditional spectrophotometric assays cannot determine DNA concentrations <5 ng/µl. More importantly, almost all spectrophotometric assays do not distinguish between different types of

nucleotides (e.g., double-stranded DNA, RNA, dNTPs, and single-stranded DNA). Therefore, while the presence of single-stranded DNA will not impair library preparation, this will result in inaccurate yield quantitation.

For PacBio library preparation, it is critical to determine the concentration of the double-stranded DNA, since only double-stranded DNA will be converted into sequencing templates. RNA, dNTPs, and single-stranded DNA included in the concentration measurement will skew the concentration reading. Therefore, it is highly recommended to use the PicoGreen® assay or a Qubit® fluorimeter for quantitation purposes.

Please closely follow the recommended guidelines provided by the respective vendors when carrying out quantitation assays. In particular, the following steps should be observed: 1) periodically have the instrument calibrated (preferably by the vendor); 2) conduct a standard curve alongside samples when using the Picogreen or Qubit assays; 3) perform replicate readings of concentration and use the average of replicates as the final concentration whenever possible. Accurate, consistent pipetting skill is needed to obtain reliable quantitation information.

V. Estimating Library yield on the PacBio® RS II (Table 1)

The PacBio Sample Calculator should be used to estimate yield for all samples. The table below provides estimates of expected yields for various DNA libraries. Please note the assumptions used to generate the table.

Library Insert Size*	Recommended Quantity for submission	Min Concentration Required (Post-Shearing)	Est. Total Yield (Range)	
			MIN	MAX
250 bp	600 ng	250 ng	60 GB	125 GB
500 bp	600 ng	250 ng	10 GB	20 GB
1 kb	1.2 µg	500 ng	90 GB	180 GB
2 kb	1.2 µg	500 ng	45 GB	90 GB
5 kb	2.4 µg	1 µg	45 GB	91 GB
10 kb	2.4 µg	1 µg	20 GB	45 GB
10 kb (AMPure kit)	10 µg	5 µg	90 GB	182 GB
20 kb (AMPure kit)	15 µg	5 ug	45 GB	91 GB
20 kb (BluePippin™ kit)	15 µg	5 µg	9 GB	18 GB

* Amounts recommended for submission represent quantities needed for one SMRTbell library prep and includes extra quantity needed for any additional QC and conservative excess. Reported library yield is based on an assumption of a DNA loading concentration of 50 ng/µl and throughput of 200 MB per SMRT Cell using P4-C2 chemistry. For insert sizes ≥1 KB, a magnetic bead loading protocol is used in the SMRT Cell calculation. Two size-selection protocols for large-insert libraries are available using either AMPure or BluePippin strategies. Actual results may vary.

VI. Accepted Volume and Buffers

- The maximum volume should not exceed 130 µl as this is the upper limit for some of the DNA shearing protocol. The upper limit for 20kb G-Tube® shearing is 200 µl.
- When resuspending DNA prior to generating SMRTbell libraries, such as during AMPure purification, avoid buffers containing EDTA to prevent enzymatic inhibition during downstream sample library preparation.
- Upon receipt of sample, performing AMPure purification to transfer DNA into a stable and appropriate buffer for downstream processing of PacBio SMRTbell templates is recommended. The additional benefit of this step is to further purify genomic DNA and remove carry-over contaminants.
- DNA can be dissolved in Tris buffer (e.g., 10 mM Tris, pH 7.0 – pH 8.0). Do not use nuclease-free water as this is insufficient for long-term DNA stabilization.

VII. Guidelines for SMRT® library preparation with high molecular weight and clean DNA

These are general recommendations to help obtain high molecular weight DNA.

Listed third party products are not officially endorsed by PacBio and are only provided as possible options.

1. Before DNA extraction:
    a. Avoid incubation in complex or rich media.
    b. Harvesting from several replicate cultures rather than a single, high-density culture during early- to mid-logarithmic growth phase is preferred.
    c. Extraction of small volumes is preferred over large volumes to avoid accumulating high concentrations of potentially inhibiting secondary components.
2. Options for DNA Extraction:
    a. Qiagen® MagAttract® HMW kit (100-200 kb) (Product Information)
    b. Qiagen Genomic-tip kit (50-100 kb) (Product Information)
    c. Qiagen Gentra® Puregene® kit (100-200 kb) (Product Information)
    d. Phenol-chloroform extraction (PacBio SampleNet Protocol)
        i. Ensure phenol is fresh and not oxidized; use within three months of opening of reagent bottle.
3. DNA cleanup before library prep
    a. Purification of DNA with AMPure kit (Default recommendation, for all users)
    b. Mo Bio® kit (MOBIO PowerClean® 50 prep kit, for highly contaminated samples)(Product Information)
        i. Protocol requires approximately 30-45 min. Multiple samples can be prepped simultaneously.
        ii. Some modifications to the protocol are suggested to maintain high-molecular-weight DNA and minimize damage:
            1. After adding Buffer 2, quickly vortex and add Buffer 3. Minimize time exposure of DNA in Buffer 2 to prevent damage.
            2. After adding Buffer 3 and vortexing, add 1 µl of glycogen and vortex before incubation on ice.
            3. During the final elution step, elute at 50 µl, spin for a second in the microcentrifuge, and incubate for 1 min at room temp. Proceed as stated in the protocol by adding 50 µl of elutant to the membrane for a minute incubation and final spin of 2 minutes.

    WARNING: DNA recovery is low following this procedure; use as necessary. A 10 µg genomic DNA sample per column resulted in 30-50% recovery. A lower sample input will result in higher recovery, whereas a high sample input results in lower recovery.

4. Shearing:
    a. G-Tube purification is recommended and is a preferred method due to the ease of use.
    b. Hydroshear is alternative which allows shearing up to 20 kb. Please note use of Hydroshear requires maintenance to prevent frequent clogging of samples.
    c. Covaris® E220 for fragments <5 kb.

I. Estimating no. of SMRT Cells for Iso-Seq Application (Table 1)

The table below provides general guidelines for the number of SMRT Cells needed for experimental design goals. Number of SMRT Cells for each project is defined per size-selected library and depends on the research goals and transcriptome complexity. For insert sizes >3 kb, the estimated number of SMRT Cells recommended is based on implementing the double BluePippin size-selection protocol, which improves sequencing performance for these longer transcript lengths. The number of SMRT Cells required generally increases for libraries that are not cleanly size-selected (due to the presence of shorter transcripts that were not removed from each size bin). Approximately 25,000 full-length transcripts can be sequenced with each SMRT Cell under optimized sample loading conditions.

Number of SMRT Cells (per sample)	Experimental Goals
1 SMRT Cell	Targeted, gene-specific isoform characterization. (e.g., using PCR products as input). Multiple genes may be analyzed on a single SMRT Cell by employing an appropriate PacBio Barcoding strategy.
1-4 SMRT Cells	General survey of full-length isoforms in a transcriptome using a no-size-selection protocol. Resulting transcript sizes may be skewed towards representation in favor of shorter isoforms (i.e., 1-3 kb).
4-8 SMRT Cells	A focused study on one or two size fractions of full-length transcripts with 4 SMRT Cells per fraction (e.g. 2-3 kb and 3-6 kb).
16 SMRT Cells	A more comprehensive survey of full-length isoforms in the whole transcriptome ranging from 1 kb to 10 kb. General guidance is to run 4 SMRT Cells per size-selected library: e.g., (a) 1-2 kb; (b) 2-3 kb; (c) 3-6 kb; and (d) 6-10 kb library size bins are recommended for enabling coverage of highly expressed isoforms.
>16 SMRT Cells	Add additional sequencing of specific size-selected libraries post-analysis to saturate isoform discovery, or increase the dynamic range for very low abundance transcript discovery and detection.

II. RNA sample requirements for PacBio sequencing

The Iso-Seq Application from Pacific Biosciences provides researchers the ability to sequence intact, full-length transcripts with the demonstrated detection of 5'/3' ends and a polyA+ tail. Project success is highly dependent on the quality of the starting RNA / cDNA material since any damages to transcripts will be directly reflected in the sequencing results. High-quality RNA extractions and clean size-selection of samples are imperative for obtaining long read lengths and optimal sequencing performance.

III. Important measures impacting RNA input sample quality

For optimal sequencing performance, it is **essential** that the RNA sample:

- Has not undergone multiple freeze-thaw cycles as they can lead to additional RNA damage.
- Has not been exposed to high temperatures (*e.g.*: > 65°C for 1 hour can cause a detectable decrease in sequence quality) or pH extremes (< 6 or > 9).
- Has an OD_{260}/OD_{280} ratio between 2.0 and 2.2.

- Has an OD_{260}/OD_{230} ratio between 1.8 and 2.1.
- Has a RIN number ≥ 9 (Recommended).
- Does not contain insoluble material.
- Does not contain DNA contamination.
- Has not been exposed to intercalating fluorescent dyes or ultraviolet radiation. SYBR dyes are not RNA damaging, but do avoid ethidium bromide.
- Does not contain denaturants (e.g., guanidinium salts or phenol) or detergents (e.g., SDS or Triton-X100).
- Does not contain carryover contamination from the original organism/tissue (e.g., heme, humic acid, polyphenols, etc.).
- Note: RNA samples should only be shipped with dry ice.

IV. Important measures impacting cDNA input sample quality

The most important factor impacting the quality of cDNA libraries is the quality of the starting RNA (see Section III above).

Please take all precautions against additional damages and ensure that the cDNA sample:

- Has not been exposed to the same types of damaging agents, conditions, and contaminants listed above for RNA input samples (see Section III above).
- Has an OD_{260}/OD_{280} ratio of 1.8 to 2.0.

V. Recommendations for RNA extraction/purification (not an official endorsement from PacBio)

The minimum starting amount is 2 ng of total RNA or 1 ng of polyA+ RNA. The following kits are recommended for polyA+ extraction/purification:

- Ambion® Poly(A) Purist™ MAG Kit (http://products.invitrogen.com/ivgn/product/AM1922)
- Qiagen® RNeasy Plus Kits (http://www.qiagen.com/qdm/rna/rneasy-plus-kits?cmpid=Qven10GARneasy)

VI. Recommendations for cDNA library construction

- Clontech® SMARTer® PCR cDNA Synthesis Kits (http://www.clontech.com/US/Products/cDNA_Synthesis_and_Library_Construction/cDNA_Synthesis_Kits/SMARTer_Kits)
- PacBio User Bulletin – Guidelines for Preparing cDNA Libraries for Isoform Sequencing (Iso-Seq™)

PACBIO® GUIDELINES FOR DNA SHIPPING AND HANDLING

Typically, freezing DNA samples is recommended. However, individual samples can be shipped at room temperature without compromising results as long as the samples are free of contaminants and the shipping temperature is controlled. If a customer prefers instead to ship samples as a dried pellet at ambient temperature, in solution at ambient temperature, or on dry ice, these are also acceptable. Please note that in many cases, the selection of shipping method depends on the choice of couriers and custom clearance procedures.

If isolating DNA via ethanol precipitation, exposure to extended heat should be minimized. Incubation at 65° C for 1 hr may incur DNA damage resulting in impaired sequencing performance. Air drying of pellets is preferred over heat drying.

Shipment with dry ice can sometimes be a challenge for long transports due to evaporation or shipping regulations. An alternative is to use cold packs (e.g., ice (-4° C) or gel (-20° C)) depending on the number of DNA samples and duration of transport. Samples will remain at a sufficiently low temperature until arrival when an ample quantity of cold packs has been used in a Styrofoam® box, assuming no delays in transport.

If DNA samples are shipped in tubes, these should be capped tightly to prevent accidental spillage or cross contamination. Sealing tops by wrapping Parafilm® wrap as an additional safeguard is recommended. Please ship DNA samples in secondary containment (e.g., a cardboard freezer box) with adequate padding.

Submission of samples in 96-well, fully skirted plates is also acceptable, provided that plates are appropriately sealed and placed in secondary containment (e.g., a cardboard freezer box). Customers should ensure each well is tightly sealed to withstand shipping conditions and avoid accidental spillage or cross contamination.

The following are available for sealing plates for shipping:

Listed third party products are not officially endorsed by PacBio and are only provided as possible options.

- MicroAmp® Clear Adhesive Films (Applied Biosystems, cat# 4306311)
- MicroSeal® 'F' Foil (Bio-Rad, cat# MSF-1001)
- Adhesive foil seal (Beckman Coulter, #BK538619)

Customers should ensure each individual well is sealed prior to shipping. Freeze the sealed plate in a secondary containment (e.g., a cardboard freezer box) at -20° C prior to shipment with dry ice. Do not ship plates without secondary containment as these may crack when placed directly on dry ice.

For shipment, ensure there is enough dry ice at the top, bottom, and sides of container with the plate (e.g., cardboard freezer box). If shipment contains multiple plates, please ensure there is padding between the plates to prevent puncture of seals. Pack plates snugly to minimize shifting or jostling during shipment.

Finally, when using courier companies, please ensure detailed and correct contact information for the receiving lab has been provided, including the phone number of the point of contact. Providing the tracking number and phone number of the courier company to the receiving party is also a good practice to track delays or mishandling of sample delivery.

Last updated: July 25, 2013