PACBIO®

Low Yield Troubleshooting Guide

Sequel® System:
The SMRT® Sequencer

For Research Use Only. Not for use in diagnostic procedures.

P/N 101-627-900-01

Pacific Biosciences of California, Inc.
1305 O'Brien Dr.
Menlo Park, CA 94025
www.pacificbiosciences.com

Introduction

This guide is a resource for troubleshooting sub-optimal yield of PacBio SMRT sequencing runs. The troubleshooting process has four stages: the evaluation of the Internal Control to address concerns with the instrument and the consumables, the evaluation of the SMRT Link Sample Setup or Binding Calculator used for the sequencing run to investigate generic sample issues, the evaluation of the SMRT Cell metrics for optimal loading, and application-specific troubleshooting such as evaluation of RNA quality for Iso-Seq or PCR troubleshooting for amplicons.

For thorough troubleshooting, all four stages should be performed. If the sample is suspect from the beginning, then troubleshooting can start at the relevant application-specific section. In addition, at the end of this guide, five example experiments are provided for reference. These are intended as practice for honing troubleshooting skills.

Step 1: Examine Internal Control

What is the Sequel DNA Internal Control Complex and why use it

The Sequel DNA Internal Control Complex, which comes in chemistry versions 2.0 and 2.1, is a fixed template of approximately 2 kilobases with symmetric SMRTbell adapters and a bound polymerase. It is for use as a spike-in sequencing control on the Sequel System.

This control is always loaded with Diffusion (even if the sample is loaded with Magbeads). It is spiked in at the end of Sample Setup and is intended to be an ideal sample for obtaining a good result on the Sequel System. Therefore, a poor result with this control can indicate potential issues with the system, reagents, or consumables.

The fixed template insert in the control does not have homology to any known organism and therefore can be easily separated from the sample data. Reimbursement for consumables on failed runs will be dependent on the use of the Sequel DNA Internal Control Complex.

How to use the Sequel DNA Internal Control Complex

The Sequel DNA Internal Control Complex comes in a single stock tube and requires a serial dilution prior to adding it to the sample. It is diluted in Complex Dilution Buffer. Figure 1 shows a schematic of this serial dilution. On average, this control represents approximately 1% or less of the total number of reads on a SMRT Cell. If the user sees the read count of this control exceeding 1%, then there may be a mistake in the serial dilution of the control or under-loading of the sample. It is critical to dilute the Sequel DNA Internal Control Complex appropriately. For each of the two serial dilutions the user should mix by gently vortexing or flicking the tube, and pulse spin to collect contents at the bottom of the tube prior to pipetting to ensure mixing. Avoid re-use of the diluted Sequel DNA Internal Control Complex from previous runs. The stability of the diluted Sequel DNA Internal Control Complex has not been extensively tested.

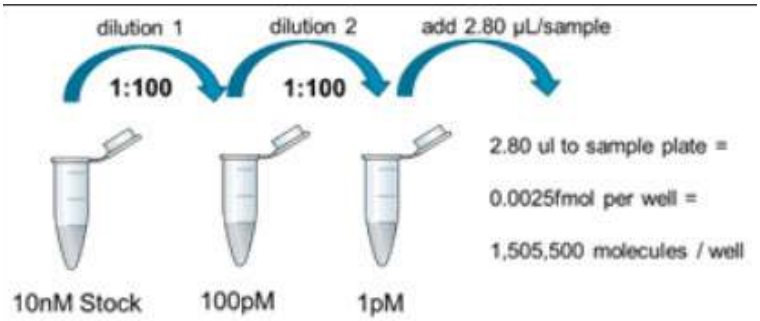
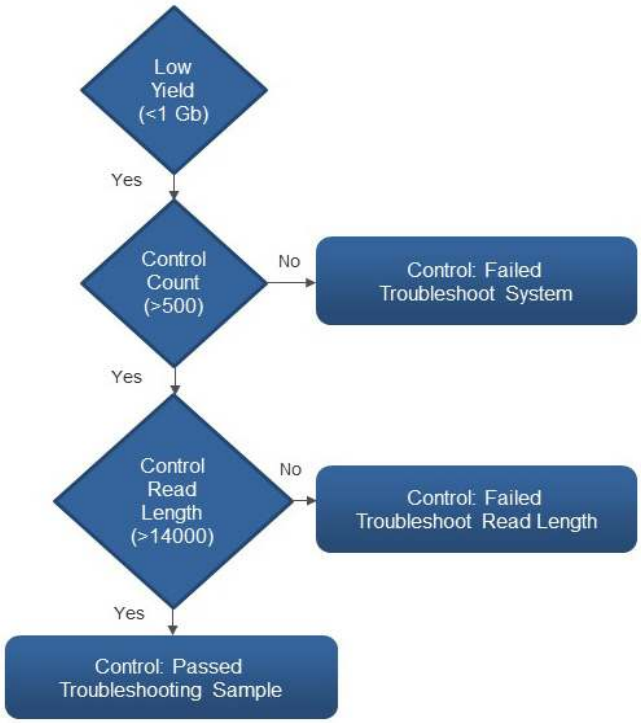


Figure 1. Serial dilution of the Sequel DNA Internal Control Complex starting with the 10nM stock (which can be used for multiple collections).

Using the Sequel DNA Internal Control Complex Metrics for Troubleshooting

The count and mean read length of the Sequel DNA Internal Control Complex are common metrics used to assess the performance of a sequencing run. While read length will vary as a function of movie time, sample handling and loading, and system variability, the control count is typically greater than 500 and the control read length is typically greater than 14,000 bases (when movie collection time is 10 hours or greater). If the control read count is less than 500, or read length is less than 14,000 bp, troubleshooting should be directed at the system and consumables. If not, then troubleshooting should be directed at the sample.

A case can be filed and additional sample information should be inspected according to the next step.



Troubleshooting Short Read Lengths

For issues involving short read lengths, file a case by emailing support@pacb.com. PacBio Technical Support and Field Application scientists will begin troubleshooting the sample when a case is received and assigned. Note that additional information is sometimes needed. Support personnel will inspect the run for conditions listed in the table below.

Cause	Description
Sample Overloading	Low P0, high P1 (>70%) values.
Sequencing Kit	Expired or improperly stored sequencing reagent plate. Opened sequencing reagent plates expire faster.
O_2 Exclusion	Low N_2 flow, empty or expired OS enzyme.
Damaged Complex	Exposure to heat, improper storage conditions, or incorrect buffers will damage the Sequel DNA Internal Control Complex.

Table 1. Troubleshooting Symptoms and Causes

Step 2 - Review Sample Setup

You can file a case by emailing support@pacb.com. PacBio Technical Support and Field Application scientists will begin troubleshooting the sample when a case is received and assigned. Note that additional information will always be needed. Before contacting Technical Support, be sure you do the following:

1. Use an up-to-date version of SMRT Link Sample Setup or Binding Calculator. Detailed information on how to check the version is found in the PacBio Operations Guide.
2. Pay attention to the warning messages in the Binding Calculator or SMRT Link Sample Setup. For example, if the Binding Calculator warns that the sample is too dilute, consider concentrating with AMPure PB beads or preparing fresh DNA template.
3. Double-check that the entry in Concentration matches the Qubit measurement. PacBio strongly recommends the Qubit for accurate quantification of double-stranded DNA.
4. Double-check that the entry in Insert Size matches the measurement of insert size. For libraries containing DNA fragments >20 kb, PacBio recommends measuring the insert size distribution with pulsed-field gel electrophoresis (PFGE) using the Bio-Rad® CHEF Mapper® XA Pulsed Field Electrophoresis system, the Sage Science's Pippin Pulse Electrophoresis Power Supply, or the FEMTO Pulse from Advanced Analytical Technologies, Inc. For users who do not have access to PFGE, PacBio Technical Support can help interpret the size distribution of the library.

   For Iso-Seq applications, PacBio recommends measuring the insert size distribution with the Bioanalyzer 12000 Kit by selecting the start and end point of the smear in the trace. The latest Iso-Seq procedure, found on the PacBio Documentation page, has further details.
   An accurate measurement of the insert size is important because inaccuracies will affect the estimation of molarity (as per the equation below) and can cause under-loading or over-loading of the sample.

In calculating SMRTbell molarity, we use the following formula:

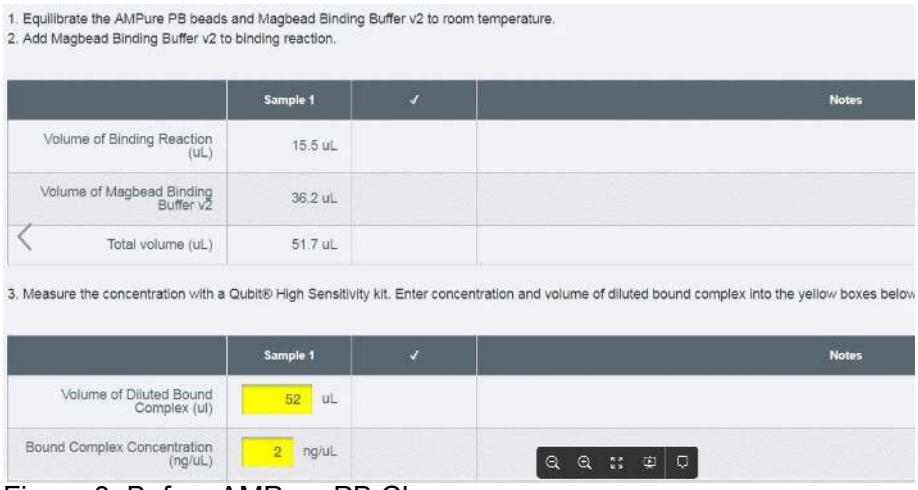
$$\frac{\text{Insert concentration (ng/µl)}}{\text{Mean insert size (bp)}} * \frac{1{,}000{,}000}{650} = \textbf{Insert concentration in (nM) (or fmol/µL)}$$

The detailed calculation is:

$$\frac{\text{Insert Concentration (ng/µL)}}{\text{Mean insert size (bp)}} * \frac{\text{1 mol basepairs}}{650\text{ g}} * \frac{\text{1 g}}{10^9\text{ ng}} * \frac{10^{15}\text{ fmol}}{\text{1 mol base pairs}} = \frac{\textbf{fmol}}{\textbf{µL}} = \textbf{nM}$$

Figure 2. SMRTbell Molarity Calculation

5. Double-check the yield after AMPure PB Cleanup, Column Cleanup, or Cleanup Beads. For example, below is a screen shot of Sample Setup Qubit measurement before AMPure PB Cleanup:



1. Equilibrate the AMPure PB beads and Magbead Binding Buffer v2 to room temperature.
2. Add Magbead Binding Buffer v2 to binding reaction.

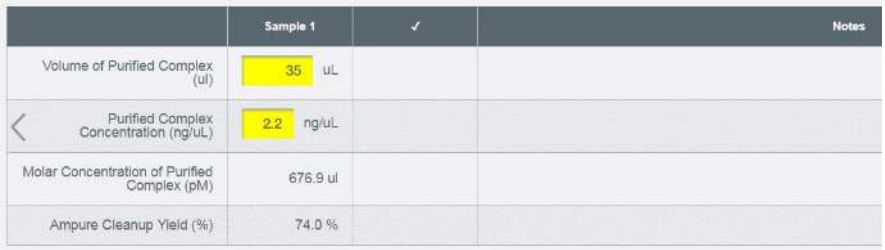
	Sample 1	✓	Notes
Volume of Binding Reaction (uL)	15.5 uL		
Volume of Magbead Binding Buffer v2	36.2 uL		
Total volume (uL)	51.7 uL		

3. Measure the concentration with a Qubit® High Sensitivity kit. Enter concentration and volume of diluted bound complex into the yellow boxes below

	Sample 1	✓	Notes
Volume of Diluted Bound Complex (ul)	52 uL		
Bound Complex Concentration (ng/uL)	2 ng/uL		

Figure 3. Before AMPure PB Cleanup

And below is the Qubit measurement after AMPure PB Cleanup with the automatic calculation of yield:



	Sample 1	✓	Notes
Volume of Purified Complex (ul)	35 uL		
Purified Complex Concentration (ng/uL)	2.2 ng/uL		
Molar Concentration of Purified Complex (pM)	676.9 ul		
Ampure Cleanup Yield (%)	74.0 %		

Figure 4. After AMPure PB Cleanup

If the recovery from AMPure PB bead purification is abnormally low, the sample may be degraded and it may be best to start with a fresh sample.

Step 3 - Check for Over-Loading or Under-Loading

You can file a case by emailing support@pacb.com. PacBio Technical Support and Field Application scientists will begin troubleshooting the system when a case is received and assigned. Note that in most cases, no additional information is needed.

Results of SMRT sequencing depend on having optimal loading of the SMRT Cell. You should avoid both under-loading and over-loading.

Under-loading of SMRT Cells

One of the most common and easiest-to-fix causes of low yield of Total Bases is under-loading of polymerase-bound complexes in the SMRT Cell. To diagnose whether this is the issue, check the P0 metric for the run in the Run QC module of SMRT Link. If P0 is substantially more than 20%, and the sample is high-quality, then the run is likely under-loaded. To fix this, re-run the sample at a substantially higher concentration.

Note that there may be some variability in loading from SMRT Cell to SMRT Cell. Also, increasing the sample concentration will only help if there is no DNA degradation, contamination, or other sample-related issues. If there is any doubt regarding sample quality, consider the other recommendations in this Troubleshooting Guide or contact PacBio Technical Support or your local Field Applications Scientist.

Over-loading of SMRT Cells

Over-loading of polymerase-bound complexes in the SMRT Cell may shorten the Polymerase Read Length of the sample and thus decrease the yield of Total Bases. There is a trade-off between the number of reads and the Polymerase Read Length. A larger number of short reads may result in less data than a smaller number of longer reads.

The best method for detecting over-loading is a titration of several values of sample concentration across SMRT Cells from the same lot. With increasing sample concentration, the P0 metric in the Run QC module of SMRT Link should decrease. Once P0 decreases to less than 10% and Polymerase Read Length starts to decrease, the sample concentration is enough to cause over-loading. Alternative methods of assessing over-loading are not as reliable. In particular, a high value of P2 for a particular SMRT Cell, taken by itself, should not be interpreted as necessarily being over-loaded.

If an individual SMRT Cell has shorter-than-expected Polymerase Read Length with concomitant lower-than-desired yield, and sample quality is not in doubt, then re-run the sample at a lower concentration.

Note that there may be some variability from SMRT Cell to SMRT Cell and that if there is any DNA damage, contamination, or other sample-related issues, then further troubleshooting may be necessary. For more information, refer to the other sections of this Troubleshooting Guide or contact PacBio Technical Support or your local Field Applications Scientist.

Step 4 - Troubleshoot the Specific Application

Iso-Seq

Sample issues specific to Iso-Seq typically occur before sample preparation begins. Therefore, this section describes how to check the quality of the RNA and cDNA sample.

How to Check the Quality of the RNA and cDNA Sample

1. Measure the RNA Integrity Number (RIN) of the RNA with the Bioanalyzer. We highly recommend RIN >8.0. Less than 8 may impact detection of full-length transcripts.
2. Check the Bioanalyzer trace or gel image for a smear of short RNA molecules. If it is present, the RNA is degraded, and it is best to start with a fresh aliquot.
3. Measure the size distribution of the RNA and/or cDNA with the Bioanalyzer or on a gel. If most of the RNA or cDNA molecules in the transcriptome are <2 kb, doing 0.40X AMPure PB bead purification will cause substantial sample loss. Contact PacBio Technical Support or the local Field Applications Scientist for suggestions on how to sequence short transcripts. Here is the transcriptome used to optimize the Iso-Seq procedure:

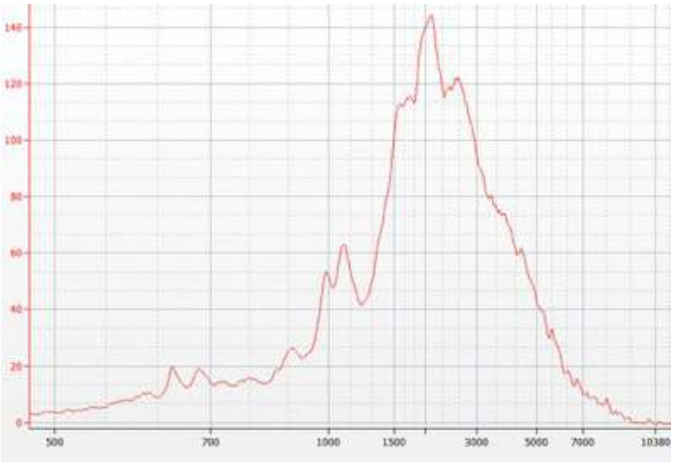


Figure 5. Optimal Transcriptome Size Distribution (Mouse Liver)

4. Measure the OD260/280 of the RNA. If the OD260/280 is below 2.0, the RNA is contaminated, typically with protein or organic solvents, and this contamination can inhibit sequencing. Troubleshoot the RNA extraction procedure to increase the RNA purity.
5. Make cDNA immediately after obtaining RNA. If the RNA is stored for longer than a few hours before reverse transcription, it may degrade. It is easier to store cDNA than RNA.
6. Check how RNA was stored. If RNA was stored at 4 degrees Celsius or -20 degrees Celsius for a prolonged period of time, it is more likely to degrade.

Long-Insert Libraries (> 10 kb)

This section describes common issues with samples which lead to low yield from long-insert libraries and how to troubleshoot each one. Issues presented are listed from more common to less common (as seen by the PacBio Technical Support team). Before contacting PacBio Technical Support or the local PacBio Field Applications Scientist, be sure you read the following:

The Long-Insert Library Contains RNA or Single-Stranded DNA

RNA and single-stranded DNA, often left-over from the DNA extraction process, contaminate a significant proportion of long-insert libraries prepared for PacBio SMRT sequencing. Low yield results because these contaminants inhibit the sequencing reaction. See below for tips on how to detect and fix the contamination:

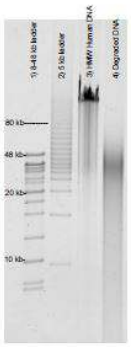
1. Measure the DNA concentration with both the Qubit and the Nano-Drop. The Qubit measurement and the NanoDrop measurement should be very similar. If the NanoDrop measurement is higher than the Qubit measurement, it is likely that the sample contains RNA and/or single-stranded DNA because the dye-based measurement method of the Qubit is more specific to double-stranded DNA than the absorbance-based measurement method of the NanoDrop.
2. To remove RNA, incubate the library with RNase A.
3. To remove single-stranded DNA, incubate it with ExoVII.
4. To remove these enzymes and the digested nucleic acids, purify the library using 0.45X AMPure PB.

The DNA Contains DNA Damage

Damage in the DNA molecule can prematurely stop the sequencing reaction, causing short polymerase read length and low number of total bases per SMRT Cell. Many forms of DNA damage manifest as a smear in a gel run with pulsed-field gel electrophoresis (PFGE). See Figure 7 for recommendations on how to use PFGE to detect DNA degradation.

A. BIO-RAD® CHEF Mapper® XA Pulsed Field Electrophoresis System

B. Sage Science™ Pippin Pulse Electrophoresis Power Supply System



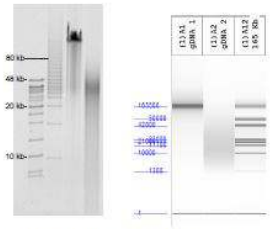
C. High Molecular Weight vs. Degraded DNA run on CHEF Mapper

A PFGE run provides information on sample quality and fragment size. PacBio recommends using either Bio-Rad's CHEF Mapper **(A)** or Sage Science Pippin Pulse **(B)** as PFGE instruments.

The gel image **(C)** shows high molecular weight DNA at ~150 kb which can be sheared to the desired size (>30 kb). Lane 4 shows a less than ideal gDNA with a smear up to 80 kb. Depending on the severity of degradation, the sample may be used directly for library construction. A size selection cutoff of 10 kb usually generates good subread lengths.

D. Advanced Analytical FEMTO Pulse™ Automated Pulsed-Field CE Instrument

While both CHEF Mapper and Pippin Pulse are reliable systems for characterizing genomic DNA, electrophoresis run times are intensive (>16 hrs) and require significant amount of DNA as input. Advanced Analytical's FEMTO Pulse instrument **(D)** is a fast high-resolution capillary based electrophoresis system able to resolve fragments up to 165 kb in one hour, ideal when constructing large-insert libraries. More importantly, the system requires picogram (pg) quantities of DNA.



E. CHEF Mapper **F.** FEMTO Pulse

Human genomic DNA was also loaded on the CHEF Mapper and FEMTO Pulse. Separation observed in CHEF Mapper (**E**) exhibits comparable performance as the Femto Pulse (**F**).

Figure 6. Recommendations for detecting degradation.

It is important to avoid introducing DNA damage. Below are some suggestions for handling the long-insert library to minimize DNA damage:

- Use a gentle extraction method. Avoid harsh chemicals, sonication, and bead-beating.
- When mixing, flick the tube, instead of pipetting or vortexing.
- For storage periods longer than a few hours, store the library frozen, but avoid more than three freeze-thaw cycles. Consider storing samples in small aliquots.

Contamination of the Long-Insert Library with Short DNA Fragments

Short DNA fragments will reduce the mean subread length, increase the P2 metric, and decrease the yield. If the Short Inserts and Adapter Dimers metrics in the Run QC module of SMRT Link are each significantly more than 0.01, the library may be contaminated with these short DNA fragments.

Consider re-purifying the library with 0.45X AMPure PB beads and re-doing the sequencing run or constructing a new library.

The DNA Contains Protein or Small-Molecule Contamination

Contamination of the library with protein and particularly with small molecules such as phenol or guanidine, often found in the reagents used for DNA extraction, can inhibit the sequencing reaction. To detect these types of contamination, measure the A260/280 and A260/230 of the library.

A260/280 should be between 1.8 and 2.0 for pure DNA. Note that this measurement is sensitive to pH, so ensure that the library is resuspended in PacBio Elution Buffer.

A260/230 should be between 2.0 and 2.2. Note that this measurement is sensitive to pH, so ensure that the library is resuspended in PacBio EB. Common causes of low A260/230 are polysaccharide carryover (often a problem with plants), residual phenol from the nucleic acid extraction, residual guanidine from a column-based kit, and residual glycogen from precipitation. Common causes of high A260/230 are making a measurement on a contaminated pedestal of a NanoDrop instrument, not having the measurement instrument well-calibrated, and using an erroneous solution as a blank measurement – the blank should have the same ionic strength and pH.

A few extra rounds of purification with 0.45X AMPure PB beads can often bring the absorbance ratios into the expected range. For samples that use the phenol-chloroform extraction, another round of this extraction can also serve as effective cleanup, provided that all traces of these chemicals are carefully removed.

The DNA Contains Polysaccharide Contamination

Polysaccharide contamination manifests as low yield and is difficult to detect. It is particularly common in plants but also possible in shrimp and bacteria. If the sample comes from one of these high-risk sources, please inquire with PacBio Technical Support or the local Field Applications Scientist about the high-salt phenol-chloroform cleanup as well as plant-specific and shrimp-specific DNA extraction procedures. DNA may need to be re-extracted.

Short-Insert Libraries (< 10 kb)

This section describes common issues with amplicons, or fragments, between 250 bp to 10,000 bp. Currently, inserts smaller than 250 bp are not supported for SMRT Sequencing because of the increased risk for creating chimeras.

Troubleshooting PCR

To generate optimal template for SMRT Sequencing it is highly recommended to start with high-quality double-stranded DNA (dsDNA), even for amplicon studies.

1. Measure the DNA concentration with both the Qubit and NanoDrop instruments.
2. Use the Qubit to estimate the concentration of dsDNA, and use the NanoDrop to assess the purity of the sample using the A280/260 and 260/230 ratios.
3. For pure DNA samples, the A280/260 should be between 1.8 and 2.0, and the A260/230 should be between 2.0 and 2.2. If the measurements fall outside these ranges, try to purify the sample with 0.45X AMPure PB beads.

The following are other steps to take to ensure successful amplicon template generation for SMRT Sequencing:

1. Dilute the DNA sample appropriately for PCR. Do not overload the PCR with too much template as this could inhibit the PCR, and increase the likelihood of non-specific inserts in the library.
2. Use a High-Fidelity Polymerase.
3. Optimize PCR conditions to minimize total time spent at 65 degrees Celsius or higher. That includes minimizing the extension and the number of cycles. Over-cycling can reduce DNA quality.
4. By carefully designing the PCR primers and quantifying the PCR product, ensure the PCR is specific with no off-target amplification, especially off-target products higher in molecular weight than the target.
5. If off-target amplification is observed and the PCR cannot be re-optimized, there are two clean-up options.
    - If the target is larger or equal to 1.5 kb and the non-specific amplicon is significantly smaller than 1.5 kb, use 0.45X AMPure PB beads to remove the non-specific amplicon.
    - If the non-specific amplicon is larger than 1.5 kb and separable from the target on a gel, it is possible to use gel extraction to purify the target. It is recommended to use an automated system, such as the BluePippin or the SageELF. It is also possible to use non-UV light and SYBR Safe for visualizing DNA on the gel. UV light and DNA-intercalating dyes, such as ethidium bromide, will cause damage to the DNA that will interfere with SMRT Sequencing. Refer to of the "Procedure & Checklist - Amplicon Template Preparation and Sequencing" for more detailed guidance on gel extraction.

6. Accurately size the amplicon on the Agilent 2100 Bioanalyzer or 2200 TapeStation, Advanced Analytical Technologies Fragment Analyzer, or agarose gel.
7. Accurately quantify the amplicons with the Qubit to ensure correct input for the adapter ligation step. Inaccurate quantification can lead to an incorrect adapter:template ratio and thus increase the risk of forming chimeras.
8. Nested, or iterative PCR is not recommended as it can increase the likelihood of off-target inserts.
9. Purify PCR template with the concentration of AMPure PB beads appropriate for the size of the target to avoid bringing primer dimers into the adapter ligation. See the table below.

Insert Size Range	Concentration of AMPure PB Beads
100 bp to 300 bp	1.8X
301 bp to 750 bp	1.0X
751 bp to 3 kb	0.6X
3 kb to 10 kb	0.45X

Table 2. Concentration of AMPure PB beads to use with various insert sizes.

Adapter Ligation for Amplicon Inserts (250 bp - 10kb)

To ensure a correct adapter:template ratio it is important to start with the appropriate amount of input DNA going into the ligation step. For insert size ranges from 250 to 750 bp use 250 ng, and for 750 bp to 10,000 bp use 500ng. See Table 3 below.

Insert Size Range	Input DNA Amount
250 bp to 750 bp	250 ng
750 bp to 10 kb	500 ng

Table 3: Amount of DNA to use in Adapter Ligation

The amount of template input to the ligation reaction is scalable, but the adapter should be present in excess relative to the molarity of the template. For inserts between 750 and 10,000 bp, the molar excess of adapter should be 30X to 50X. For inserts between 250 and 750 bp, it is necessary to increase the molar excess of adapter to 90X to 200X. For example, add 6.0 µL (3X) or 8.0 µL (4X) of the 20 µM adapter to the 40 µL ligation reaction for 250 ng of input template. When using more input DNA, scale the volume accordingly, so the molarity ratios remain constant.

AMPure PB Bead Purification of Short Insert Libraries

When purifying short insert SMRTbell Templates (250-750 bp) do three rounds of AMPure PB bead clean up. For inserts between 750 bp and 10,000 bp, it is sufficient to do two rounds of AMPure PB bead clean up. See the Table 2 above for the appropriate AMPure PB bead concentrations.

Pre-extension Time

For short-insert libraries and amplicons, a pre-extension is recommended to improve the number of ZMWs in CCS. The amount of pre-extension time can be calculated using the following equation: (insert size x 2)/ (1.5 bases/second x 60 minutes)

Use the longest inserts (if using a mixed-size pool) to determine pre-extension time.

Example Experiments

Good Yield but Short Read Length with a Long-Insert Library

Symptom/Problem

Table 4 below shows SMRT Link Run QC results with a 20 kb size-selected library loaded at 8 pM with Diffusion in two SMRT Cells where yield is more than 5GB per SMRT Cell. However, average polymerase read length is about 8674 bp for SMRT Cell A01 and 7558 bp for SMRT Cell B01. What can be done to increase the polymerase read length?

Well	Movie Time (min)	Loading	Total Bases (GB)	Read Length		Productivity			Control	
				Poly-merase AVG	Longest Subread AVG	P0	P1	P2	#	Poly RL
A01	600	Diffusion	6.9	8674	7859	7.2% (73,349)	78.0% (795,067)	14.8% (150,819)	3195	15939
B01	600	Diffusion	6.1	7060	7060	5.7% (57,680)	79.2% (807,364)	15.1% (154,191)	3458	13916

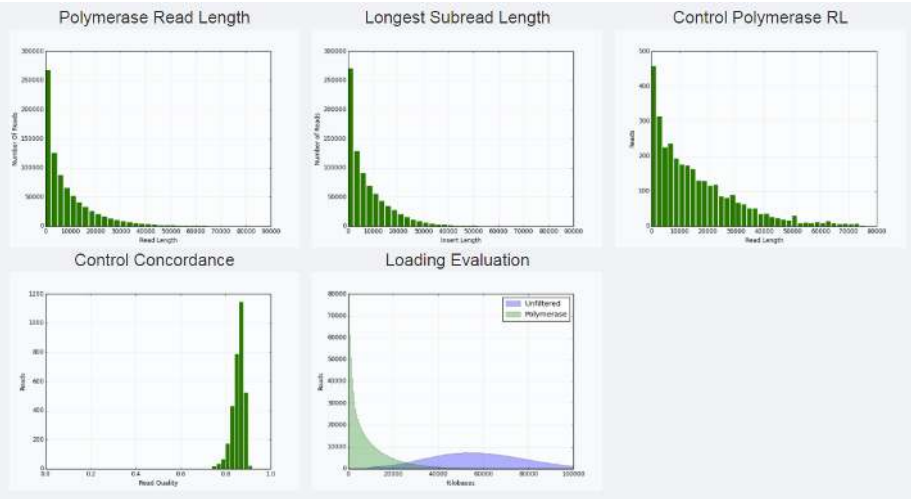


Figure 7. Screenshots of plots from Run QC for SMRT Cell A01

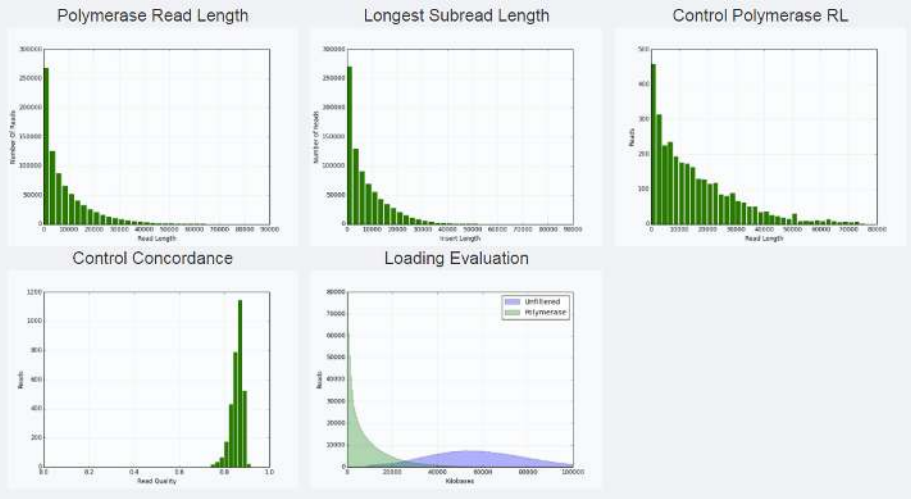


Figure 8. Screenshots of plots from Run QC for SMRT Cell B01

Diagnosis

1. Check the results in Table 4 for evidence of issues with the Sequel System, the SMRT Cells, or the reagents. In this case, there are no such issues because the Total Bases and the Poly RL of the Control are good. For up-to-date information on how to interpret the Total Bases and the Poly RL of the Control, please contact PacBio Technical Support.
2. Check the Productivity metrics in Table 4. Note that for every SMRT Cell P1 > P0 and also P2 > P0.
3. Check the loading concentration against the recommended range. In this instance, the loading concentration was at the upper bound of the recommended range.

Solution

The Diagnosis above indicates that the likely cause of short polymerase read length is overloading because P1 is much higher than P0, P2 is higher than P0, and the loading concentration is at the upper bound of the recommended starting range.

1. Verify that the measurements of DNA concentration and insert size of the library were accurate. Inaccurate measurements will result in inaccurate loading concentrations. In this case, this may not be necessary because the loading concentration was at the upper bound of the recommended starting range, and therefore overloading was likely.
2. Re-do the sequencing run with a reduced loading concentration. PacBio recommends performing a loading titration of a few loading concentrations to achieve optimal loading.

Variable Yield and Read Length from the Same Complex

Symptom/Problem

Table 5 below shows the SMRT Link Run QC results for three SMRT Cells of size-selected long-insert 15 kb library that went through primer annealing, polymerase binding and then Diffusion Loading. The loading concentration was below the lower bound of the recommended range. Internal Control was not used, therefore, there are no Control metrics.

Well	Movie Time (min)	Loading	Total Bases (GB)	Read Length		Productivity		
				Poly-merase AVG	Longest Subread AVG	P0	P1	P2
A01	360	Diffusion	1.2	5377	4909	16.7% (169,851)	21.9% (223,614)	61.4% (625,770)
C01	360	Diffusion	3.21	10330	8923	45.8% (466,342)	30.5% (310,780)	23.8% (242,113)
E01	360	Diffusion	4.47	13427	11152	59.5% (606,204)	32.6% (332,679)	7.9% (80,352)

Note that Total Bases vary from 4.47 GB to 1.2 GB and that average Polymerase Read Length varies from 13427 bp to 5377 bp.

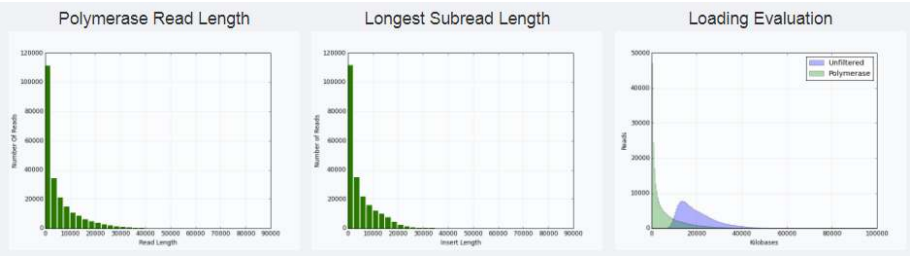


Figure 9. Screenshots of plots from Run QC for SMRT Cell A01

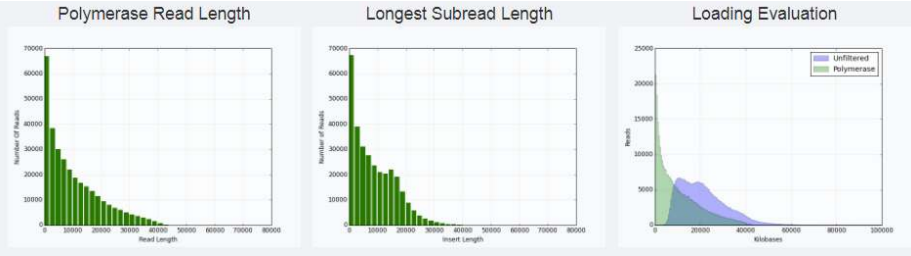


Figure 10. Screenshots of plots from Run QC for SMRT Cell C01

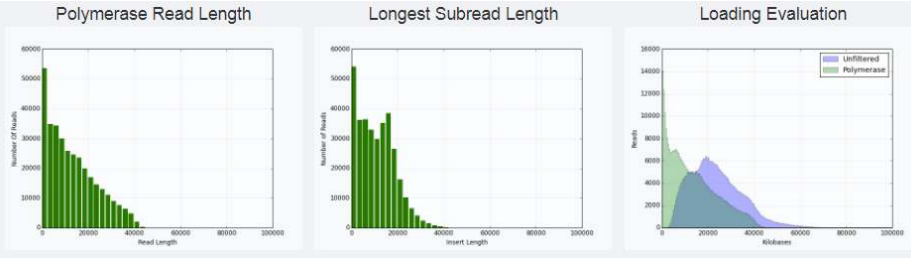


Figure 11. Screenshots of plots from Run QC for SMRT Cell E01

Diagnosis

This section requires a process of elimination, therefore, the steps can be done in any order.

1. Check the loading concentration against the recommended starting range of loading concentration to assess for overloading. Note that here, overloading is unlikely because the loading concentration was below the lower bound of the recommended range.
2. Check that the loading concentration was consistent. In this case all three SMRT Cells were loaded at 0.8 pM. Therefore, an error in loading concentration does not explain the variable yield and read length.
3. For SMRT Cells A01 and C01 where P2 is higher or comparable to P1, check that correct cleanup method of the polymerase-bound complex was done. In this case, SMRT Cell A01 was loaded without any cleanup, and this explains the very high P2. However, SMRT Cell C01 was loaded after the recommended Column Cleanup, and yet the P2 is still quite high.
4. Another common cause of high P2 is excess polymerase in the binding reaction. Check that the polymerase was diluted correctly and that the correct volume was added. In this case, all three SMRT Cells had 5 times more polymerase than recommended. Interestingly, SMRT Cell E01 was loaded after AMPure PB bead purification of the polymerase-bound complex. This cleanup is so effective that it removed most of the excess polymerase and produced a good P2 value, average polymerase read length, and yield of Total Bases.

Solution

Lower Total Bases and shorter polymerase read length correlated with higher P2. This combination of symptoms indicates contamination of the sample with free polymerase. In this case, it appears that too much polymerase was added into the binding reaction and, in SMRT Cell A01, there was no cleanup done. This sequencing run is also a side-by-side comparison of cleanup methods, and AMPure PB bead purification of the polymerase-bound complex is the most effective cleanup method for this long-insert library.

Low Yield and Short Read Length for Gel-Purified Sample

Symptom/Problem

Table 6 shows a gel-purified, 1 kb amplicon library that went through primer annealing, polymerase binding, and AMPure PB bead purification. The sample was loaded using Diffusion loading at 6 pM. Table 6 below shows the results of the Run QC. Internal Control was not used, therefore, Table 6 does not contain any Control metrics.

Well	Movie Time (min)	Loading	Total Bases (GB)	Read Length		Productivity		
				Poly-merase AVG	Longest Subread AVG	P0	P1	P2
A01	600	Diffusion	0.90	10817	1795	70% (727180)	8% (83531)	22% (226089)

Diagnosis

Note that the amplicon was gel-purified. Since gel-purification was used, start with investigating the gel-purification process.

1. Check the DNA quality guidelines below:
   - Verify that there was no exposure to intercalating agents, like ethidium bromide, or UV
   - OD 260/280 was between 1.8 and 2.0
   - OD 260/230 was between 2.0 and 2.2
   - Verify that the DNA concentration measurement by Qubit matches the NanoDrop measurement
   - Verify that no chelating agents (e.g., EDTA), divalent metal cations (e.g., Mg2+), denaturants (e.g., guanidinium salts, phenol), or detergents (e.g., SDS, Triton-X100, CTAB) were used
   - Verify that there was no exposure to pH extremes (<6 or >9)
   - Verify that there was no prolonged exposure to high temperatures (> 65ºC for 1 hour)
   - Verify that there was a minimum of freeze-thaw cycles, ideally fewer than 3
   - Verify that there was no carryover contamination from the starting organism
2. Note that the sample satisfies all the requirements, except that the sample was gel-purified from a gel containing ethidium bromide and visualized with UV.

Solution

Intercalating agents, such as ethidium bromide, and UV exposure both cause irreversible DNA damage that interferes with SMRT sequencing. The sample preparation has to be redone from the amplification step. If gel purification is needed, it should use a stain that does not intercalate into DNA, such as SYBR Safe, and that can be visualized with lower-energy light, such as blue light.

Variable Yield from the Same Library and Different Binding Reactions

Symptom/Problem

Table 7 shows a 20-kb library that went through primer annealing, polymerase binding, Column Cleanup, and loaded with Magbeads. In the first sequencing run (not shown), the yield of Total Bases was low. A fresh Binding Kit was then used, the polymerase binding, Column Cleanup, and Magbead loading was redone.

Table 7 below shows two SMRT Cells run side-by-side: SMRT Cell A01 used the fresh Binding Kit and SMRT Cell B01 used the older Binding Kit. Note the difference in Total Bases. Internal Control was not used, therefore, there are no Control metrics.

Well	Movie Time (min)	Loading	Total Bases (GB)	Read Length		Productivity		
				Poly-merase AVG	Longest Subread AVG	P0	P1	P2
A01	360	Diffusion	3.16	12869	11759	52% (537231)	24% (245200)	22% (226089)
C01	360	Diffusion	0.04	13480	11841	96% (996600)	0% (3292)	4% (36908)

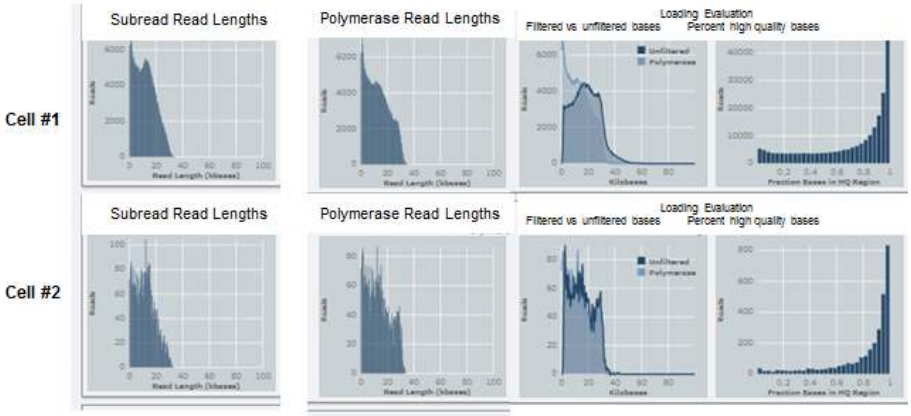


Figure 12. Plots of read length and loading evaluation for SMRT Cell A01 (Cell #1) and SMRT Cell C01 (Cell #2).

Diagnosis

1. Because the only difference between the two SMRT Cells is the Binding Kit, the polymerase should be the focus of troubleshooting.
2. Check the expiration date on the Binding Kits. In this case, neither the initial Binding Kit nor the fresh Binding Kit was expired.
3. Check the storage conditions. The Binding Kit should be stored at -20 degrees Celsius. The Binding Kits were stored frozen, but in a freezer that undergoes temperature cycles to automatically defrost.

Solution

The Binding Kit, as is any enzyme, is sensitive to temperature fluctuations. Sequencing reagents should be stored in a freezer that keeps stable temperature. Any remaining reagents in this Binding Kit that produced low yield should not be used.

Rescuing Read Length

Symptom/Problem

A 12-plex Microbial Multiplexing library was loaded at 4 pM with Diffusion and 120 minutes of pre-extension on both SMRT Cell A01 and SMRT Cell B01 in Table 8 below. Why is there longer polymerase read length and average longest subread in SMRT Cell B01?

Well	Movie Time (min)	Total Bases (GB)	Read Length		Productivity		
			Poly-merase AVG	Longest Subread AVG	P0	P1	P2
A01	600	7.38	12918	4983	16.3% (166322)	56.0% (571077)	27.7% (281836)
B01	600	7.33	13460	5981	29.8% (304101)	53.4% (544467)	16.7% (170667)

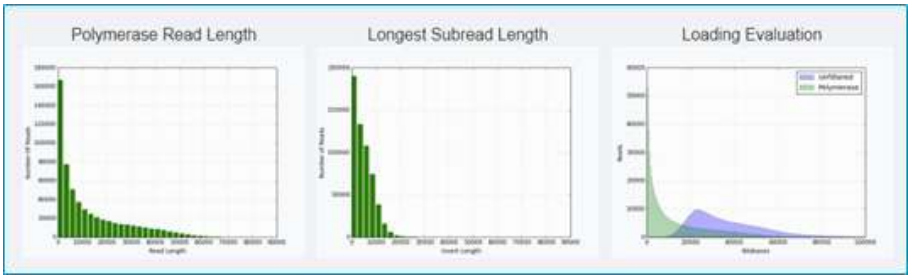


Figure 13. Screen shot of SMRT Link Run QC plots for SMRT Cell A01

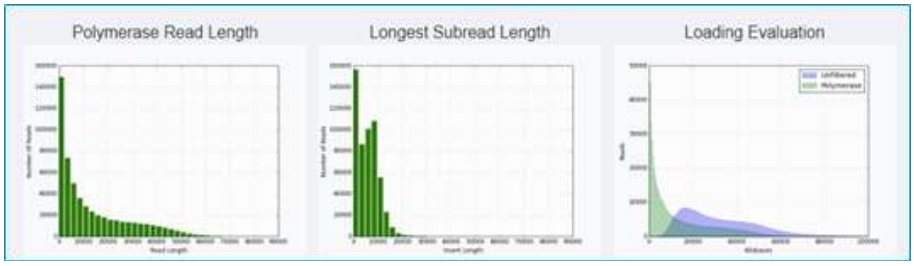


Figure 14. Screen shot of SMRT Link Run QC plots for SMRT Cell B01

Diagnosis

The diagnosis is through a process of elimination.

1. Pre-extension has an effect on polymerase read length, but both SMRT Cells had equal pre-extension duration.
2. SMRT Cell A01 does have higher % P1 and % P2, and overloading reduces the polymerase read length. However, SMRT Cell A01 also has a shorter Longest Subread, and overloading would not explain this difference. Also, the loading concentration is the same for both SMRT Cells.

Solution

The 12-plex Microbial Multiplexing library in SMRT Cell B01 underwent size-selection with a minimum cut-off of 6 kb. This led to longer polymerase read length and longest subread. The 12-plex Microbial Multiplexing library in SMRT Cell A01 did not undergo size-selection.

Revision History	Version	Date
Initial release	01	August 2018