PacBio Americas User Group Meeting Sample Prep Workshop Breakout Session: *Service Provider Operations*

June.27.2017 / http://programs.pacificbiosciences.com/l/1652/2017-03-25/3sn5p2

AGENDA

Introduction

- PacBio Service Provider Operational Considerations
- Sample Processing and Sequencing Workflows
- Run QC
- Troubleshooting SMRTbell Library Sequencing Performance
- Technical resources for PacBio Service Providers
- PacBio Certified Service Provider Program

Service Provider Panel Discussion

- ***Erin Bernberg**, Associate Scientist, University of Delaware Sequencing and Genotyping Center*
- ***David Corney**, Ph.D., GENEWIZ*
- ***Melissa Laird Smith**, Ph.D., Assistant Professor, Dept. of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai and Assistant Director of Technology Development, Icahn Institute of Genomics and Multiscale Biology*
- ***Luke Tallon**, Institute for Genome Sciences at the University of Maryland School of Medicine*

Q&A and Open Discussion

PacBio Service Provider Operational Considerations

PACBIO SERVICE PROVIDER OPERATIONAL CONSIDERATIONS

What general 'best practices' can be adopted by PacBio service providers to be more effective partners to their customers?

What types of resources would you like PacBio to offer to enhance service provider deliverables?

What are your opinions on the current state of demand for SMRT Sequencing?

Which applications are you most likely to add to your list of services?

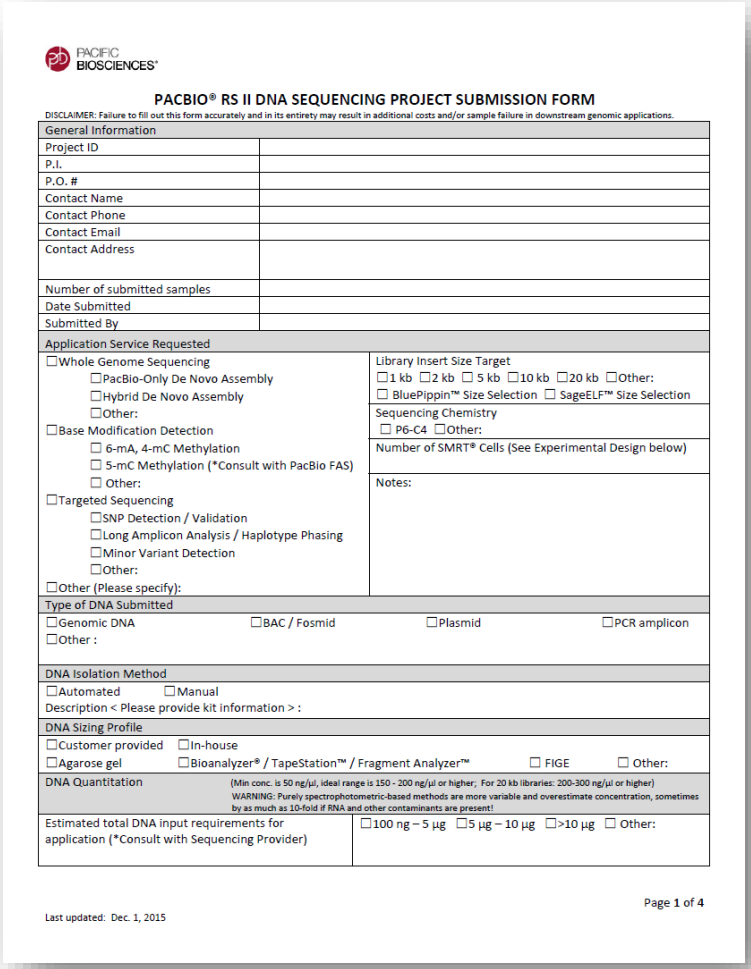
Which PacBio applications are most frequently employed?

- *Setting deliverables and expectations with clients*
- *Nucleic acid isolation*
- *Sample QC workflows*
- *Sample library processing workflows*
- *Run QC performance metrics*
- *Real time monitoring of run status*
- *Troubleshooting sequencing performance issues*
- *Turnaround time*
- *Instrument reliability and uptime*
- *Laboratory information management (LIMS)*
- *Application programming interfaces (APIs)*
- *Bioinformatics data analysis services*
- *IT support and data storage*
- *Report Outputs*
- *Automation*
- *Technical staff training (Library Sample Prep, Instrument Operators)*
- *Open Houses and Applications Seminars*
- *Technical Project Consultations*

Sample Processing and Sequencing Workflows

PacBio RS II Project Submission Form Templates

PacBio RS II DNA Sequencing Project Submission Form (December 2015)



PACIFIC BIOSCIENCES®

PACBIO® RS II DNA SEQUENCING PROJECT SUBMISSION FORM

DISCLAIMER: Failure to fill out this form accurately and in its entirety may result in additional costs and/or sample failure in downstream genomic applications.

General Information	
Project ID	
P.I.	
P.O. #	
Contact Name	
Contact Phone	
Contact Email	
Contact Address	
Number of submitted samples	
Date Submitted	
Submitted By	

Application Service Requested

☐Whole Genome Sequencing
- ☐PacBio-Only De Novo Assembly
- ☐Hybrid De Novo Assembly
- ☐Other:

☐Base Modification Detection
- ☐ 6-mA, 4-mC Methylation
- ☐ 5-mC Methylation (*Consult with PacBio FAS)
- ☐ Other:

☐Targeted Sequencing
- ☐SNP Detection / Validation
- ☐Long Amplicon Analysis / Haplotype Phasing
- ☐Minor Variant Detection
- ☐Other:

☐Other (Please specify):

Library Insert Size Target
☐1 kb ☐2 kb ☐ 5 kb ☐10 kb ☐20 kb ☐Other:
☐ BluePippin™ Size Selection ☐ SageELF™ Size Selection

Sequencing Chemistry
☐ P6-C4 ☐Other:

Number of SMRT® Cells (See Experimental Design below)

Notes:

Type of DNA Submitted

☐Genomic DNA ☐BAC / Fosmid ☐Plasmid ☐PCR amplicon
☐Other :

DNA Isolation Method

☐Automated ☐Manual
Description < Please provide kit information > :

DNA Sizing Profile

☐Customer provided ☐In-house
☐Agarose gel ☐Bioanalyzer® / TapeStation™ / Fragment Analyzer™ ☐ FIGE ☐ Other:

DNA Quantitation (Min conc. is 50 ng/µl, ideal range is 150 - 200 ng/µl or higher; For 20 kb libraries: 200-300 ng/µl or higher)
WARNING: Purely spectrophotometric-based methods are more variable and overestimate concentration, sometimes by as much as 10-fold if RNA and other contaminants are present!

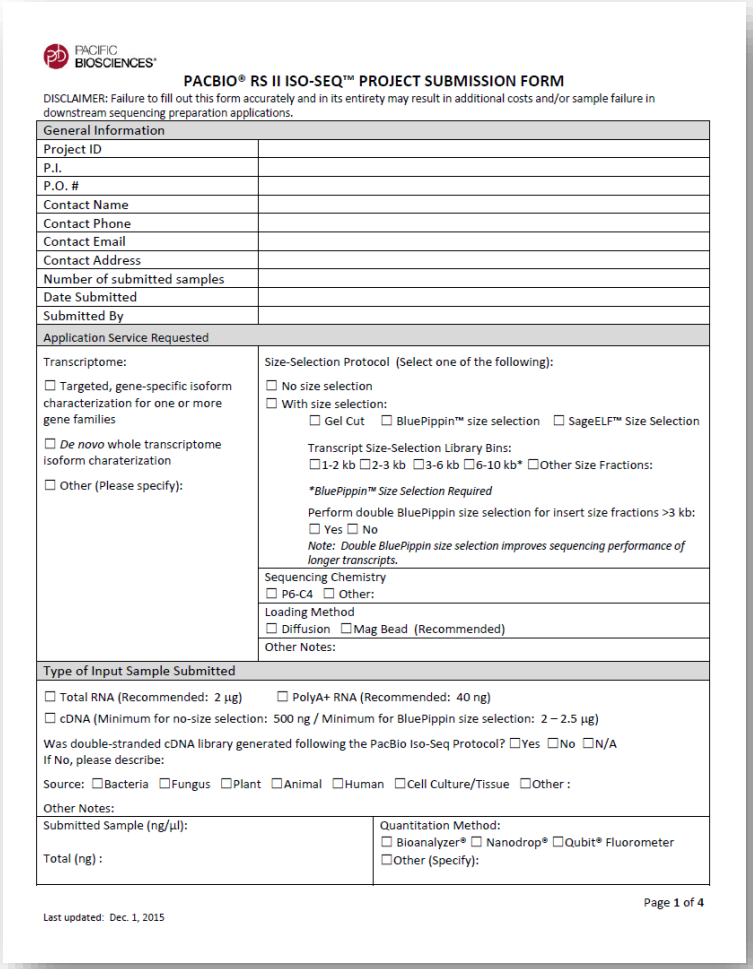
Estimated total DNA input requirements for application (*Consult with Sequencing Provider)	☐100 ng – 5 µg ☐5 µg – 10 µg ☐>10 µg ☐ Other:

Page 1 of 4

Last updated: Dec. 1, 2015

http://www.pacb.com/wp-content/uploads/Pacbio-Guidelines-Project-Submission-SMRT-Sequencing-DNA-v2.1.pdf

PacBio RS II Iso-Seq Project Submission Form (December 2015)



PACIFIC BIOSCIENCES®

PACBIO® RS II ISO-SEQ™ PROJECT SUBMISSION FORM

DISCLAIMER: Failure to fill out this form accurately and in its entirety may result in additional costs and/or sample failure in downstream sequencing preparation applications.

General Information	
Project ID	
P.I.	
P.O. #	
Contact Name	
Contact Phone	
Contact Email	
Contact Address	
Number of submitted samples	
Date Submitted	
Submitted By	

Application Service Requested

Transcriptome:

☐ Targeted, gene-specific isoform characterization for one or more gene families

☐ *De novo* whole transcriptome isoform charaterization

☐ Other (Please specify):

Size-Selection Protocol (Select one of the following):

☐ No size selection
☐ With size selection:
- ☐ Gel Cut ☐ BluePippin™ size selection ☐ SageELF™ Size Selection

Transcript Size-Selection Library Bins:
☐1-2 kb ☐2-3 kb ☐3-6 kb ☐6-10 kb* ☐Other Size Fractions:

**BluePippin™ Size Selection Required*

Perform double BluePippin size selection for insert size fractions >3 kb:
☐ Yes ☐ No
Note: Double BluePippin size selection improves sequencing performance of longer transcripts.

Sequencing Chemistry
☐ P6-C4 ☐ Other:

Loading Method
☐ Diffusion ☐Mag Bead (Recommended)

Other Notes:

Type of Input Sample Submitted

☐ Total RNA (Recommended: 2 µg) ☐ PolyA+ RNA (Recommended: 40 ng)
☐ cDNA (Minimum for no-size selection: 500 ng / Minimum for BluePippin size selection: 2 – 2.5 µg)

Was double-stranded cDNA library generated following the PacBio Iso-Seq Protocol? ☐Yes ☐No ☐N/A
If No, please describe:

Source: ☐Bacteria ☐Fungus ☐Plant ☐Animal ☐Human ☐Cell Culture/Tissue ☐Other :

Other Notes:

Submitted Sample (ng/µl): Total (ng) :	Quantitation Method: ☐ Bioanalyzer® ☐ Nanodrop® ☐Qubit® Fluorometer ☐Other (Specify):

Page 1 of 4

Last updated: Dec. 1, 2015

http://www.pacb.com/wp-content/uploads/Pacbio-Guidelines-Project-Submission-Iso-Seq-v2.1.pdf

EQUIPMENT AND MATERIALS NEEDED FOR SAMPLE LIBRARY PREPARATION AND SEQUENCING

PACBIO®

Equipment and Materials Needed for Sequencing
Quick Reference Card

This Quick Reference Card lists examples of equipment and other lab materials that may be needed for sequencing on PacBio® instruments. Note that although vendor part numbers are listed herein, they may not always be up to date. Verify and obtain the latest part numbers and relevant documentation from the vendor or manufacturer's website.

Quantification

Vendor	Item
Agilent Technologies (Agilent)	Bioanalyzer® 2100 platform 
Agilent	Consumable: DNA 7500 chip for Bioanalyzer 2100 (PN 5067-1506)
Agilent	Consumable: DNA 12000 chip for Bioanalyzer 2100 (PN 5067-1508) 
Invitrogen	Qubit® Consumable: Assay Tubes (PN Q32856)
Life Technologies	Qubit Quantitation Platform - Fluorometer 
Thermo-Scientific	Nanodrop® (2000, 2000c, 8000) spectrophotometer 

http://www.pacb.com/wp-content/uploads/qrc-equipment-and-materials-needed-for-sequencing.pdf

USER BULLETINS AND FIELD ADVISORIES TO ENSURE OPTIMAL PERFORMANCE OF YOUR SMRT SEQUENCING SYSTEMS

User Bulletins & Field Advisories (May 2017)

User Bulletin for PacBio RS II and Sequel Systems: Centrifuge Tube and Pipet Tip Recommendations ***(NEW!)*** *(May 2017)*

- PacBio advises against the use of Axygen MAXYMum Recovery™ tubes and pipet tips. Please discontinue use of these products immediately. PacBio recommends alternatives in the User Bulletin.
- http://www.pacb.com/wp-content/uploads/User-Bulletin-Centrifuge-Tube-and-Pipet-Tip-Recommendations.pdf

Field Advisory for Sequel System: Securing Sequel Pipet Tip Rack ***(NEW!)*** *(May 2017)*

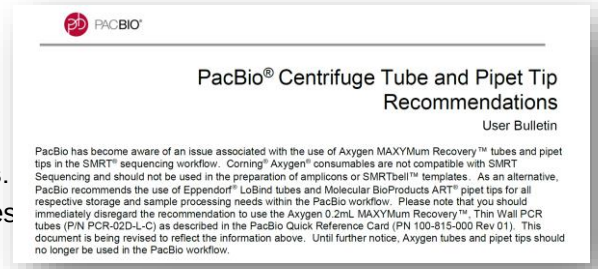
- PacBio recommends a simple procedure to ensure that the Sequel Pipet Tip rack is firmly affixed to the tip box.
- http://www.pacb.com/wp-content/uploads/Field-Advisory-Notice-Securing-Sequel-Pipet-Tip-Rack.pdf

User Bulletin for Sequel System: Heat Seal Advisory (Adhesive Seal Warning) ***(NEW!)*** *(May 2017)*

- PacBio advises against the use of adhesive foils and recommends the use of Sequel Sample Plate Foil.
- http://www.pacb.com/wp-content/uploads/User-Bulletin-Heat-Seal-Advisory-Adhesive-Seal-Warning.pdf

User Bulletin for Sequel System: Barcode Scanning of Sequel Sequencing Kit 2.0 ***(NEW!)*** *(May 2017)*

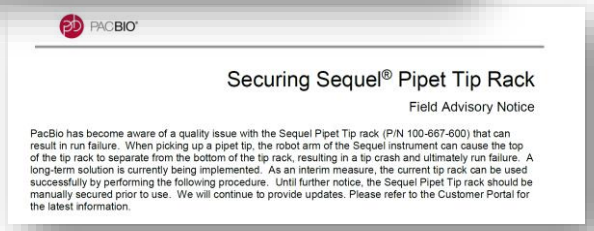
- PacBio is providing clarity on which barcode to scan to ensure the Sequel System has the correct information and that all the consumables are compatible.
- http://www.pacb.com/wp-content/uploads/User-Bulletin-Barcode-Scanning-of-Sequel-Sequencing-Kit-2.0.pdf


PACBIO

PacBio® Centrifuge Tube and Pipet Tip Recommendations

User Bulletin

PacBio has become aware of an issue associated with the use of Axygen MAXYMum Recovery™ tubes and pipet tips in the SMRT® sequencing workflow. Corning® Axygen® consumables are not compatible with SMRT Sequencing and should not be used in the preparation of amplicons or SMRTbell™ templates. As an alternative, PacBio recommends the use of Eppendorf® LoBind tubes and Molecular BioProducts ART® pipet tips for all respective storage and sample processing needs within the PacBio workflow. Please note that you should immediately disregard the recommendation to use the Axygen 0.2mL MAXYMum Recovery™, Thin Wall PCR tubes (P/N PCR-02D-L-C) as described in the PacBio Quick Reference Card (PN 100-815-000 Rev 01). This document is being revised to reflect the information above. Until further notice, Axygen tubes and pipet tips should no longer be used in the PacBio workflow.


PACBIO

Securing Sequel® Pipet Tip Rack

Field Advisory Notice

PacBio has become aware of a quality issue with the Sequel Pipet Tip rack (P/N 100-667-600) that can result in run failure. When picking up a pipet tip, the robot arm of the Sequel instrument can cause the top of the tip rack to separate from the bottom of the tip rack, resulting in a tip crash and ultimately run failure. A long-term solution is currently being implemented. As an interim measure, the current tip rack can be used successfully by performing the following procedure. Until further notice, the Sequel Pipet Tip rack should be manually secured prior to use. We will continue to provide updates. Please refer to the Customer Portal for the latest information.

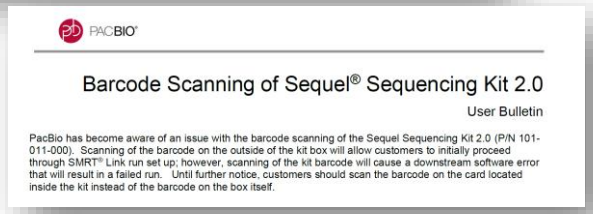

PACBIO

Heat Seal Advisory (Adhesive Seal Warning)

User Bulletin

PacBio has recently become aware of an issue associated with the use of adhesive foil seals. The use of adhesive foil seals can lead to instrument damage. Until further notice, PacBio advises against the use of adhesive foil seals and recommends the of use of the Sequel® Sample Plate Foil (P/N 100-667-400). The of use of the Sequel Sample Plate Foil requires the purchase of a heat sealer. PacBio recommends the Thermo Scientific™ ALPS™ 50 V Semi automated Microplate Heat Sealer (P/N AB-1443).

The use of adhesive foil seals will void existing instrument warranties. Customers will be responsible for any damage resulting from the use of products discordant with official PacBio recommendations.


PACBIO

Barcode Scanning of Sequel® Sequencing Kit 2.0

User Bulletin

PacBio has become aware of an issue with the barcode scanning of the Sequel Sequencing Kit 2.0 (P/N 101-011-000). Scanning of the barcode on the outside of the kit box will allow customers to initially proceed through SMRT® Link run set up; however, scanning of the kit barcode will cause a downstream software error that will result in a failed run. Until further notice, customers should scan the barcode on the card located inside the kit instead of the barcode on the box itself.

GUIDELINES FOR DNA SAMPLE SUBMISSIONS

High-quality, high-molecular-weight genomic DNA is imperative for obtaining long read lengths and optimal SMRT sequencing performance

Important General Considerations

- Submit samples of the highest possible quality and purity. DNA of high molecular weight (peak at ≥45 kb) is ideal for large insert libraries. If submitting PCR products: they should be clean amplicons (without non-specific products, carryover contaminants or multiple bands)
- 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.
- DNA should be eluted in neutral, buffered solution (e.g., 10 mM Tris, pH 8.0). Do not use nuclease-free water or unbuffered solutions as this is insufficient for long-term DNA stabilization. Avoid buffers containing EDTA to prevent enzymatic inhibition during downstream sample library preparation.
- Avoid resuspending samples in buffers containing detergents (e.g. SDS) or other additives that can inhibit enzymatic reactions during the library preparation and/or sequencing reaction.
- Quantify samples using the PicoGreen or Qubit method. Nanodrop or other purely spectrophotometric-based methods tend to overestimate sample concentration, resulting in an inadequate amount of starting material.

Human Biomedical Research

Plant and Animal Sciences

Microbiology and Infectious Disease

Desired Properties of High-Quality Starting Input DNA Samples

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

- Is double-stranded as single-stranded DNA is not compatible with the library preparation process.
- 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 can cause a detectable decrease in sequence quality), pH extremes (<6 or >9). Air drying of pellets is preferred over heat drying. Avoid overdrying of genomic DNA. Do not heat when drying in a speed-vac.
- Has an OD260/OD280 ratio of 1.8 to 2.0.
- Has an OD260/OD230 ratio of ~2.0 or higher
- Does not contain insoluble material.
- Does not contain RNA contamination.
- Has not been exposed to intercalating fluorescent dyes or ultraviolet radiation.
- Does not contain chelating agents (e.g., EDTA), divalent metal cations (e.g., Mg^{2+}), denaturants (e.g., guanidinium salts, phenol), or detergents (e.g., SDS, Triton-X100).
- Does not contain carryover contamination from the starting organism/tissue (heme, humic acid, polyphenols, polysaccharides, etc.).

Input DNA Sample Quality Assessment

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

DNA Sizing Characterization

- For >30 kb libraries, genomic DNA integrity should ideally be assessed by pulse-field gel electrophoresis (e.g., Bio-Rad CHEF Mapper, Sage Science Pippin Pulse); a Bioanalyzer / TapeStation analysis is insufficient
  - 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.
  - While both CHEF Mapper and Pippin Pulse are reliable systems for characterizing genomic DNA, electrophoresis run times are intensive (>16 hrs) and require a significant amount of DNA as input.
  - As an alternative to PFGE, Advanced Analytical's FEMTO Pulse™ is a fast high-resolution capillary based electrophoresis system able to resolve fragments up to 165 kb in 1 hour, ideal when constructing large insert libraries. More importantly, the system requires only picogram (pg) sample input quantities.
- For PCR amplicon products or cDNA samples, agarose gel electrophoresis or Bioanalyzer / TapeStation can be used to assess fragment size

http://www.sagescience.com/products/pippin-pulse/

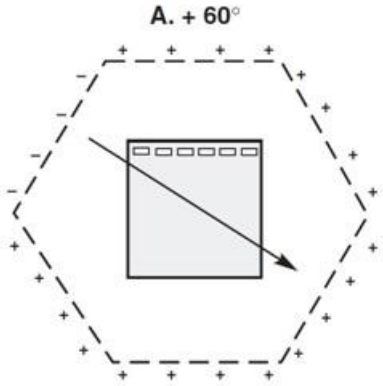
http://www.bio-rad.com/en-us/category/pulsed-field-gel-electrophoresis-systems

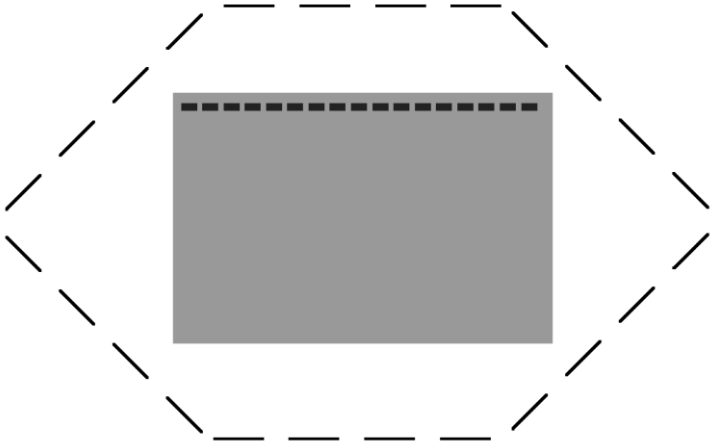
https://www.aati-us.com/instruments/femto-pulse/

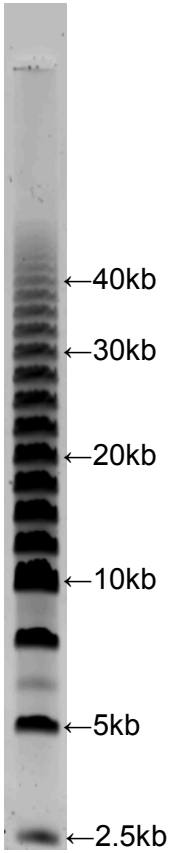
http://www.genomics.agilent.com/en/Bioanalyzer-System/2100-Bioanalyzer-Instruments/?cid=AG-PT-106

http://www.genomics.agilent.com/en/TapeStation-System/4200-TapeStation-Instrument/?cid=AG-PT-181&tabId=prod2420037

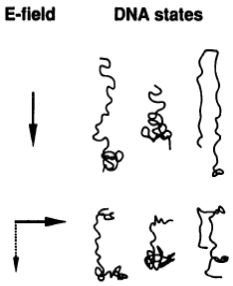
Bio-Rad CHEF Mapper XA PFGE System
(Contour-clamped Homogeneous Electric Field)







CHEF Run Parameters:
Low MW: 1 kb
High MW: 50 kb
Calibration Factor: 1.00
PFG Agarose: 1%
Buffer: 0.5X TBE
Fwd V Gradient: 9 V/cm
Rev V Gradient: 6 V/cm
Initial Switch Time: 0.05 sec
Final Switch Time: 0.92 sec
Fwd Ramp: Linear
Rev Ramp: Linear
Run Time: 13 hr, 47 min
Temperature: 14˚C



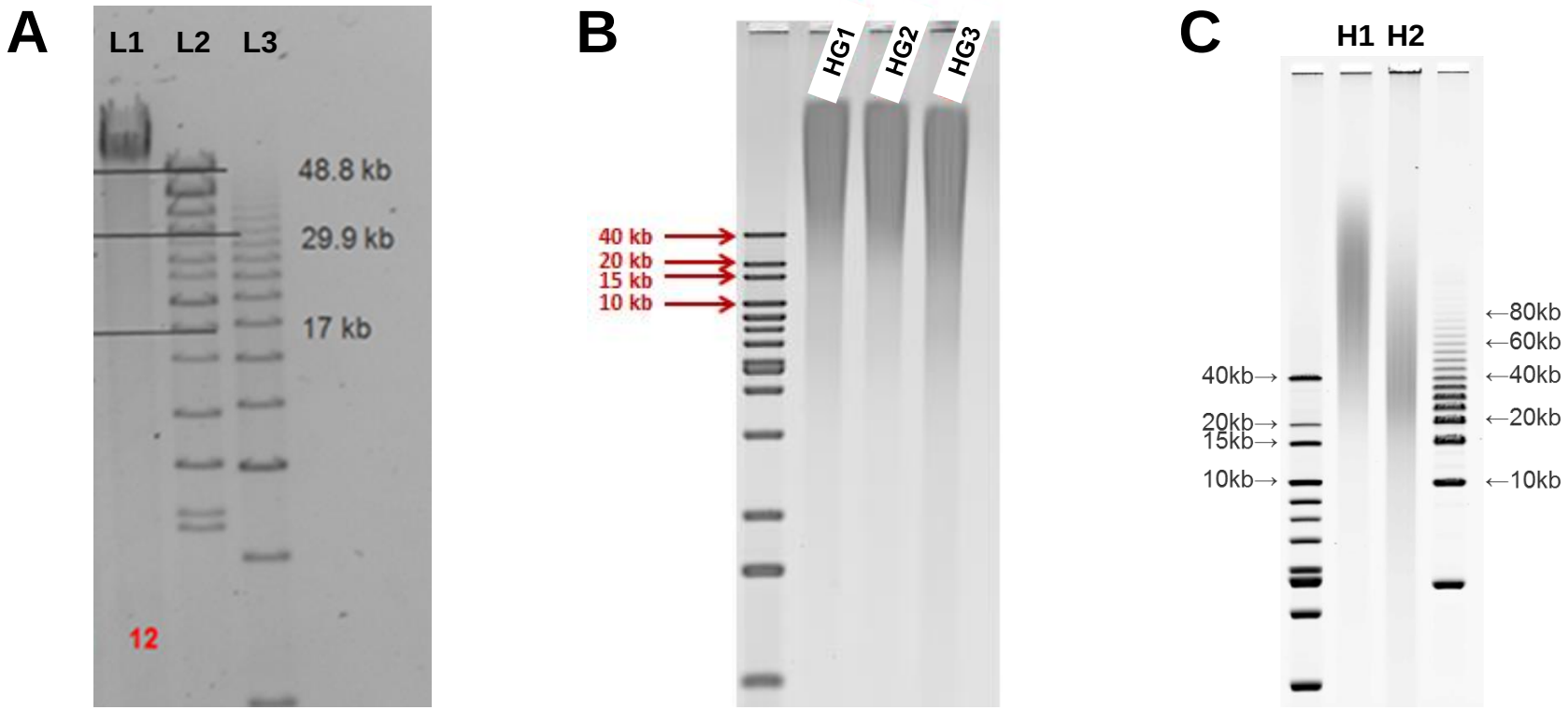
Chu, G. 1991, PNAS 88: 11071-5

Theory – In an alternating electric field, larger DNA fragments take longer to re-orient and migrate more slowly
Observation – Resolution of DNA fragments of up to 10 Mb
Limitation – PFGE requires ~100 ng DNA, takes 14-20 hours per run, and requires instrumentation hardware that can be relatively costly and not widely available

http://www.bio-rad.com/en-ca/product/chef-mapper-xa-system
http://www.bio.davidson.edu/genomics/method/pulse_field.html

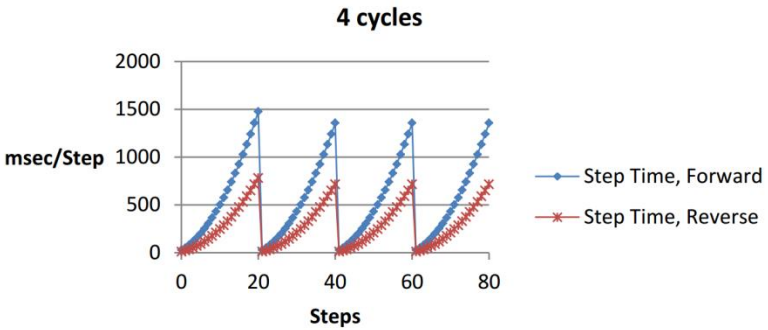
Example: QC of Human Genomic DNA Samples Using the CHEF Mapper® System

- Running pulsed-field gel electrophoresis is a common QC practice at Pacific Biosciences to verify the integrity of the starting gDNA material
- Very high-quality, high-molecular gDNA will tend to migrate as a pronounced band of approximately 50 kb or higher on a pulsed-field gel image



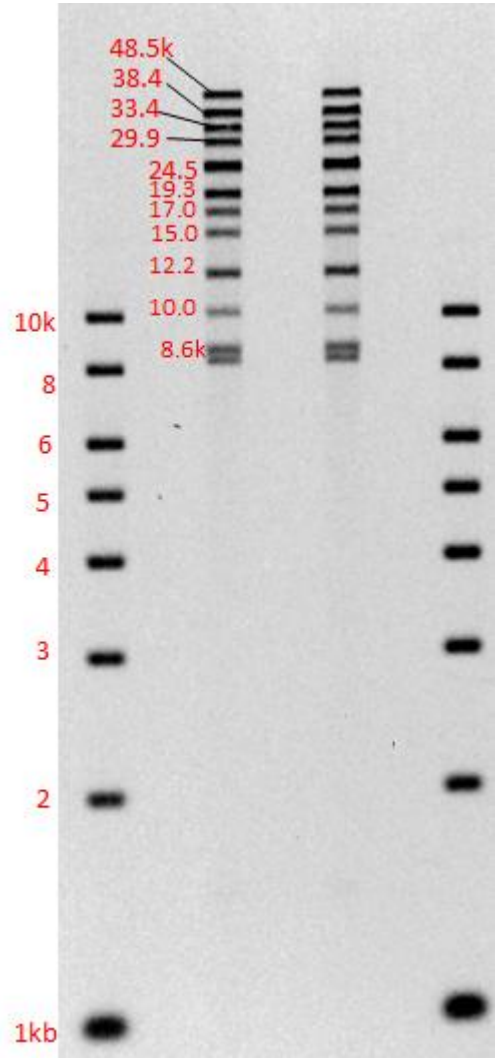
A. PFGE gel for a human gDNA sample shows that the sample (L1) is a high-quality, high-molecular gDNA sample
B. PFGE gel for three human gDNA samples shows that the bulk of the gDNA in each sample migrates above the 40-kb marker, suggesting that the input DNA is still suitable for proceeding with the shearing step.
C. PFGE gel for two human gDNA samples shows that Sample H2 is significantly degraded

Sage Science Pippin Pulse™ Field Inversion Gel Electrophoresis (FIGE) Power Supply System

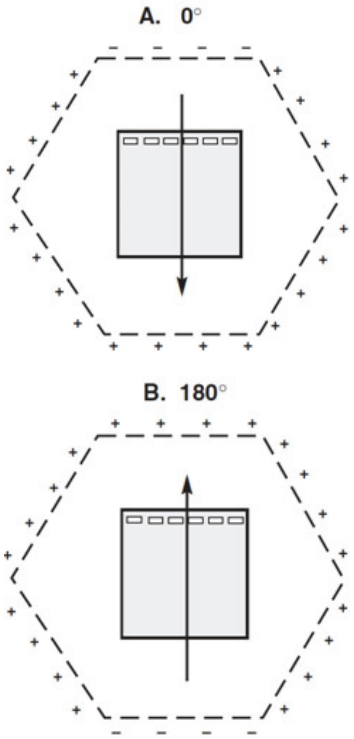


Example of running four cycles of a protocol employing ramped forward and reverse time steps. In this example, each cycle is about 6 seconds in length. Typical runs require 8-16 hours.

- In FIGE, large DNA molecules can be separated by periodically inverting a uniform electric field in one dimension (i.e., using a 180 degree re-orientation angle).
- To achieve a net forward migration, FIGE employs a difference in ether the duration or the voltage of the forward and reverse fields
- Pippin Pulse system resolves DNA fragments up to 100 kb and beyond
- Pippin Pulse system is less costly than other types of pulsed-field gel electrophoresis systems



FIGE Run Parameters: A gel image showing resolution of fragments from a gel run with Pippin Pulse using the pre-set 1-50 kb protocol. Values shown are kilobase pairs of DNA. The gel was run on a 12 X 14 cm gel for 9 hours cast with 0.75% Lonza SeaKem® GOLD agarose and 0.5X KBB buffer (Sage Science Part No. KBB1001).



http://www.sagescience.com/products/pippin-pulse/#tab-1391659714-1-79

FEMTO *Pulse*™ (*NEW!*)

Automated Pulsed-Field Capillary Electrophoresis Instrument Features and Benefits

Rapid Separation of High Molecular Weight DNA

- Separate large DNA fragments or smears through 200,000 bp in ~1 hour

Ultra-High Femtogram Level Sensitivity Detection

- 100 times more sensitivity for fragments (5 fg/uL)
- 10 times more sensitivity for smears (200 fg/uL)

Reduce Sample Input Quantity

- Use less sample for QC (20 uL min sample vol.)

Separate Samples in Parallel

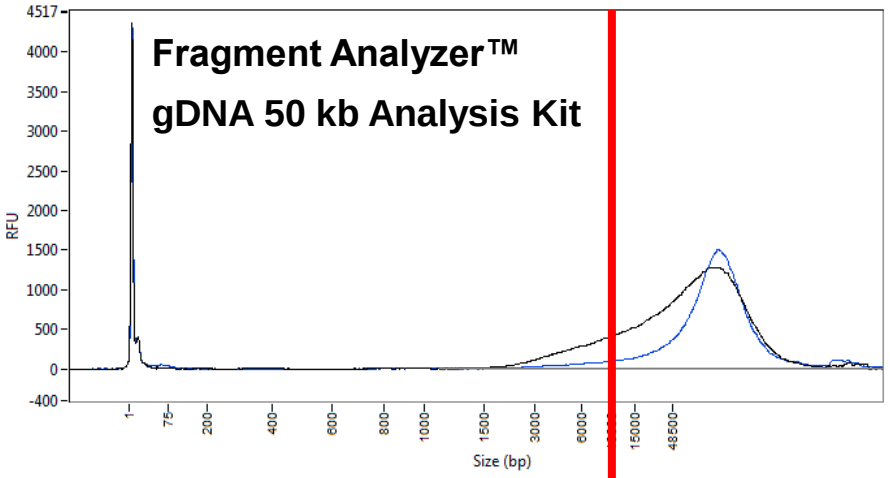
- Sample throughput capacity = 12 samples per run

Automated Operation

- Max unattended sample capacity = 288 samples

2016 R&D 100 WINNER

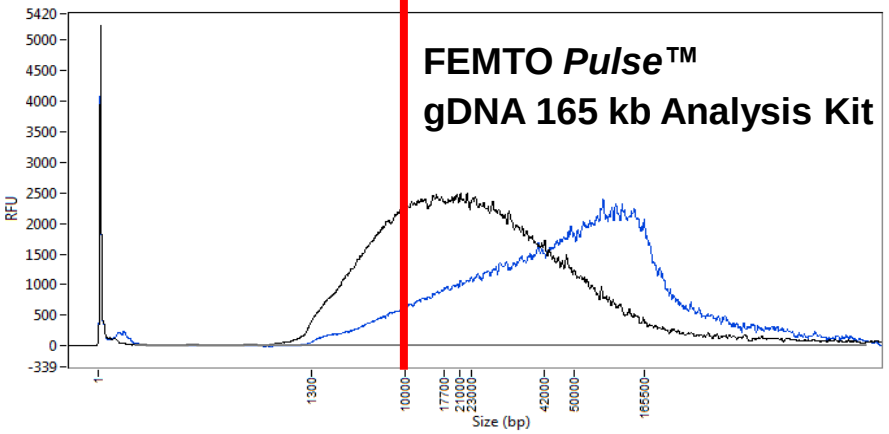
Comparison of HMW DNA Separation Using FA (Constant Field) *vs.* FEMTO *Pulse* (Pulsed Field)



Average Size

Sample 1 = 59,023bp

Sample 2 = 79,535bp



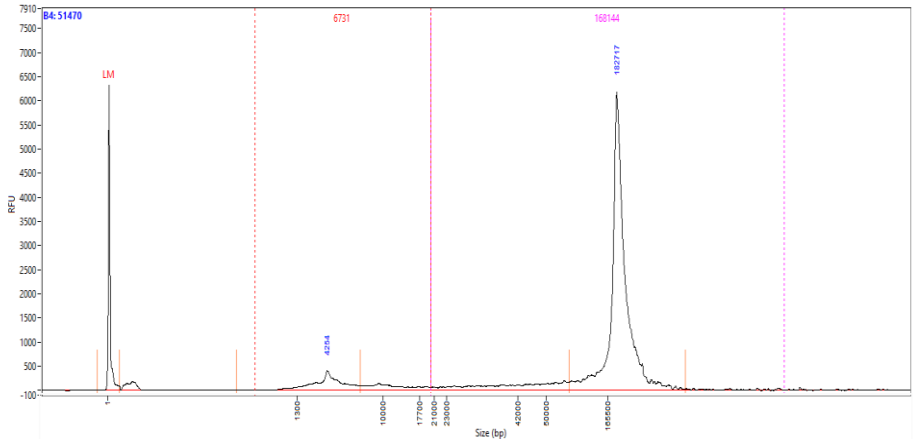
Average Size

Sample 1 = 37,396bp

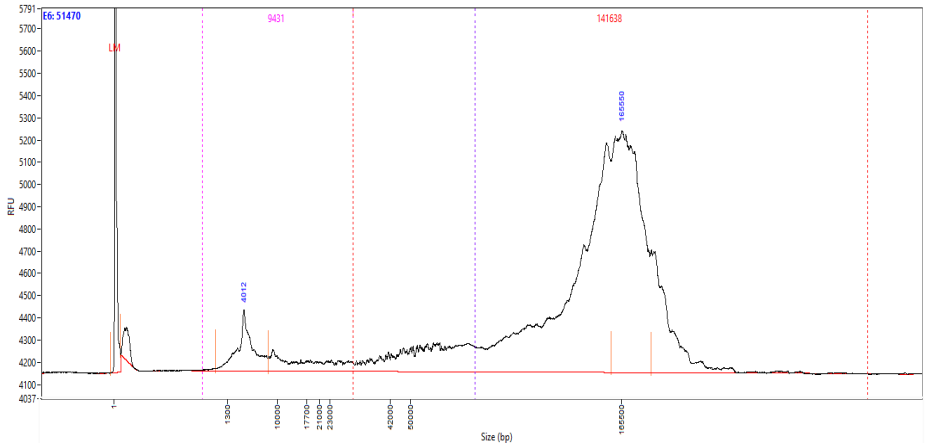
Sample 2= 69,257bp

Two Methods, One Kit to Size and Quantify Example: Genomic DNA QC

Standard Method
70 minutes



Extended Method
210 minutes



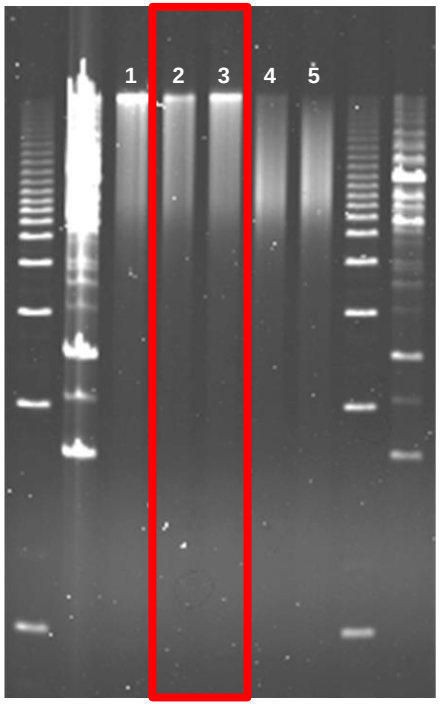
Method	Avrg Peak Size, bp, n=9	%CV	Smear Range Kb	Avrg Smear Size, bp, n=9	%CV	Qubit Concentration, pg/µL, n=2	FEMTO *Pulse* Concentration, pg/µL, n=9	%CV	% Difference From Qubit
SM	171,231	6.8	20 Kb - 500 Kb	160,341	7.2	63.6	67.2	5.3	5.6
EM	160,587	1.8	30 Kb - 300 Kb	137,200	2.1	63.6	69.2	4.9	8.9

gDNA (165 kb) Analysis Kit (P/N FP-1002-0275)

Example: Analysis of Large Fragment DNA Test Shears

Pulsed-Field Gel Electrophoresis (PFGE) Results

K12 ECOLI MEGARUPTOR TEST SHEARS

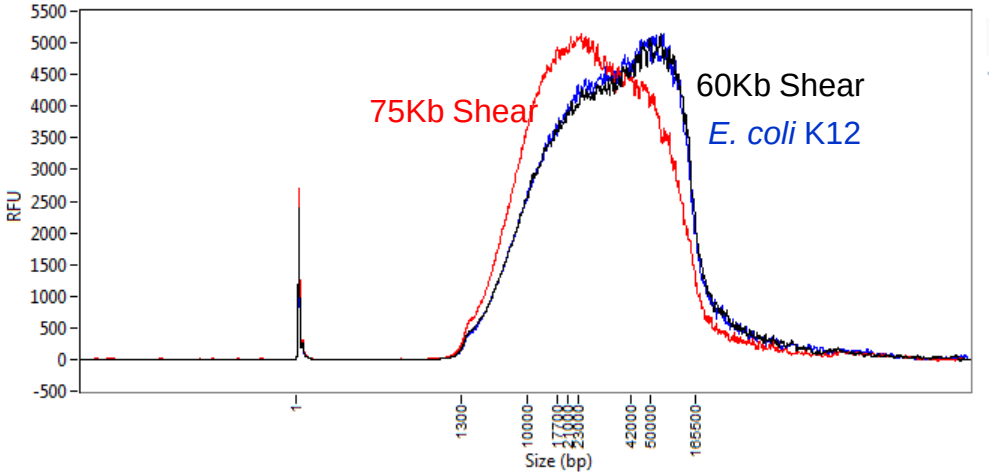


1. K12 Ecoli genomic DNA
2. 75 kb Megaruptor shear
3. 60 kb Megaruptor shear
4. 50 kb Megaruptor shear
5. 40 kb Megaruptor shear

FEMTO *Pulse* Results

gDNA (165 kb) Analysis Kit (P/N FP-1002-0275)

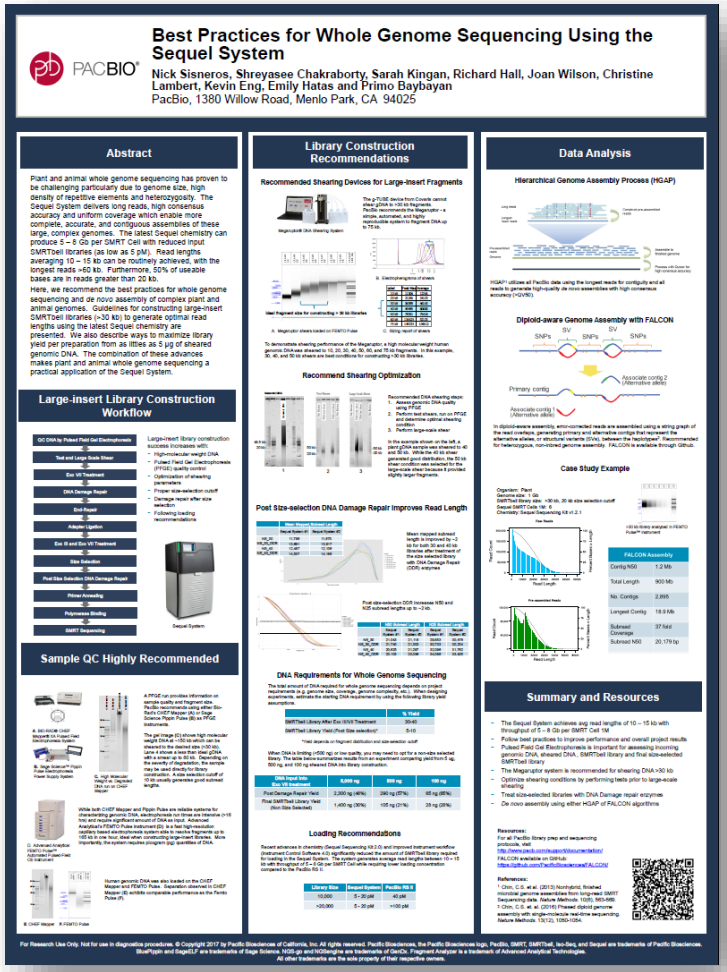
- https://www.aati-us.com/consumables/reagent-kits/femto-pulse/fp-1002-0275/
- Enables the separation of gDNA and DNA fragments through 165 kb at low concentrations.
- Enables replacement of PFGE with pulsed-field capillary electrophoresis at critical QC checkpoints to conserve sample and obtain results faster.



Sample ID	Avg Size, bp N=18	% Rel. Std.	DQN 40,000 N=18
EC K12	59,234	8.4	4.0
75Kb Shear	44,134	5.1	2.9
60Kb Shear	59,730	7.1	4.0
50Kb Shear	40,594	3.5	2.7
40Kb Shear	33,019	3.4	2.0

PacBio Poster: Best practices for Whole Genome Sequencing Using the Sequel System

- Presented at PAG 2017 in San Diego



Best Practices for Whole Genome Sequencing Using the Sequel System

Nick Sisneros, Shreyasee Chakraborty, Sarah Kingan, Richard Hall, Joan Wilson, Christine Lambert, Kevin Eng, Emily Hatas and Primo Baybayan
PacBio, 1380 Willow Road, Menlo Park, CA 94025

Abstract

Plant and animal whole genome sequencing has proven to be challenging particularly due to genome size, high density of repetitive elements and heterozygosity. The Sequel System delivers long reads, high consensus accuracy and uniform coverage which enable more complete, accurate, and contiguous assemblies of these large, complex genomes. The latest Sequel chemistry can produce 5 – 8 Gb per SMRT Cell with reduced input SMRTbell libraries (as low as 5 pM). Read lengths averaging 10 – 15 kb can be routinely achieved, with the longest reads >60 kb. Furthermore, 50% of useable bases are in reads greater than 20 kb.

Here, we recommend the best practices for whole genome sequencing and *de novo* assembly of complex plant and animal genomes. Guidelines for constructing large-insert SMRTbell libraries (>30 kb) to generate optimal read lengths using the latest Sequel chemistry are presented. We also describe ways to maximize library yield per preparation from as little as 5 µg of sheared genomic DNA. The combination of these advances makes plant and animal whole genome sequencing a practical application of the Sequel System.

Large-insert Library Construction Workflow

Sample QC Highly Recommended

Library Construction Recommendations

Recommended Shearing Devices for Large-insert Fragments

Recommend Shearing Optimization

Post Size-selection DNA Damage Repair Improves Read Length

DNA Requirements for Whole Genome Sequencing

Loading Recommendations

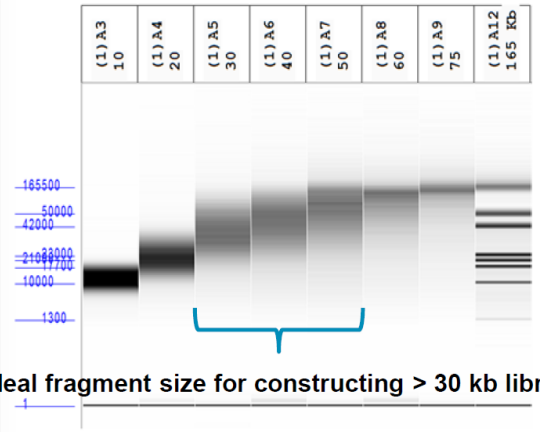
Data Analysis

Hierarchical Genome Assembly Process (HGAP)

Diploid-aware Genome Assembly with FALCON

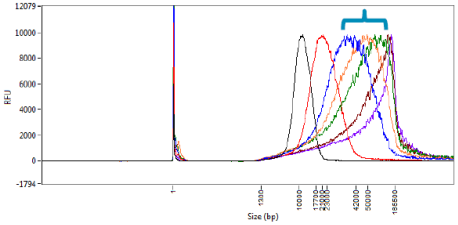
Case Study Example

Summary and Resources



Ideal fragment size for constructing > 30 kb libraries

A. Megaruptor shears loaded on FEMTO Pulse



B. Electropherograms of shears

Label	Peak Max	Average
10 kB	11304	12245
20 kB	21245	24123
30 kB	36399	46143
40 kB	49495	60600
50 kB	78351	75416
60 kB	134525	92133
75 kB	149225	104512

C. Sizing report of shears

To demonstrate shearing performance of the Megaruptor, a high molecular weight human genomic DNA was sheared to 10, 20, 30, 40, 50, 60, and 75 kb fragments. In this example, 30, 40, and 50 kb shears are best conditions for constructing >30 kb libraries.

http://www.pacb.com/wp-content/uploads/Baybayan-PAG-2017-Best-Practices-for-Whole-Genome-Sequencing-Using-the-Sequel-System.pdf

PACBIO®

Determination of DNA Purity

- 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:

<u>260/280 Ratio</u>

- 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, pH 8). 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 Ratio

- 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.

Determination of DNA Concentration

- 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.

DNA Sample Extraction

General recommendations to help obtain high molecular weight DNA:

Before DNA Extraction:

- Avoid culture incubation in complex or rich media.
- Harvesting from several replicate cultures rather than a single, high-density culture during early- to mid-logarithmic growth phase is preferred.
- Extraction of small culture volumes is preferred over large volumes to avoid accumulating high concentrations of potentially inhibiting secondary components

After DNA Extraction:

- 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. Avoid vortexing genomic DNA when possible as vortexing can cause shearing of the DNA. It is also recommended to use wide bore tips in sample handling.
- 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® beads purification.
- DNA storage conditions: 4 °C (short-term); –20 °C / –80 °C (long-term).

Kits and Methods for Isolating High Molecular Weight Genomic DNA:

Bacteria (Gram Negative and Gram Positive)

QIAGEN Genomic-tip 20/100/500/G Kit

- https://www.qiagen.com/ca/shop/sample-technologies/dna/dna-preparation/QIAGEN-Genomic-tip-500G#productdetails

QIAGEN Gentra Puregene Yeast/Bact. Kit

- https://www.qiagen.com/us/shop/sample-technologies/dna/dna-preparation/gentra-puregene-yeastbact-kit/#productdetails

Geneaid™ DNA Isolation (Bacteria) Kit

- http://www.geneaid.com/products/reagent-genomic-dna-extraction/genomic-dna-precipitation-extraction-kit-bacteria-gram-positive-gram-negative

Yeast and Fungi

QIAGEN Genomic-tip 20/100/500/G Kit
QIAGEN Gentra Puregene Yeast/Bact. Kit

GeneJET Plant Genomic DNA Purification Kit

- https://www.thermofisher.com/order/catalog/product/K0791

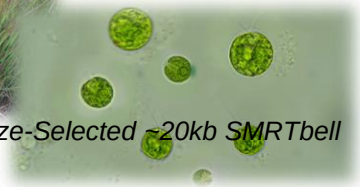
Zymo Research Fungal/Bacterial DNA MidiPrep™ Kit

- https://www.zymoresearch.com/dna/microbial-environmental-dna-isolation-1/bacterial-fungal-dna/zr-fungal-bacterial-dna-midiprep

Phenol Chloroform Extraction (Neurospora crassa)

- Reference: Long-read, whole-genome shotgun sequence data for five model organisms. *Scientific Data* **1**:140045 (2014) DOI: 10.1038/sdata.2014.45 http://www.nature.com/articles/sdata201445

Plant and Algae Tissue



Unsupported Protocol – Preparing Arabidopsis Genomic DNA for Size-Selected ~20kb SMRTbell Libraries (CTAB Treatment and Chloroform-Isoamyl Extraction)

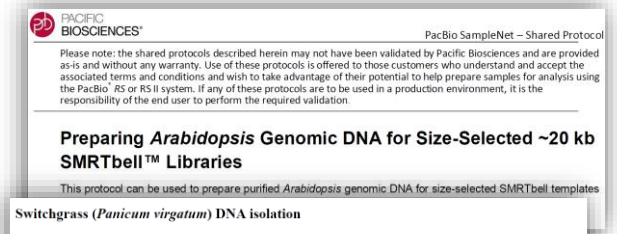
- http://www.pacb.com/wp-content/uploads/2015/09/Shared-Protocol-Preparing-Arabidopsis-DNA-for-20-kb-SMRTbell-Libraries.pdf

Unsupported Protocol – Switchgrass (Panicum virgatum) DNA isolation [USDA]

- http://www.pacb.com/wp-content/uploads/2015/09/Switchgrass-DNA-isolation.pdf

Unsupported Protocol – DNA extraction of Chlamydomonas using CTAB [JGI]

- http://www.pacb.com/wp-content/uploads/2015/09/DNA-extraction-chlamy-CTAB-JGI.pdf

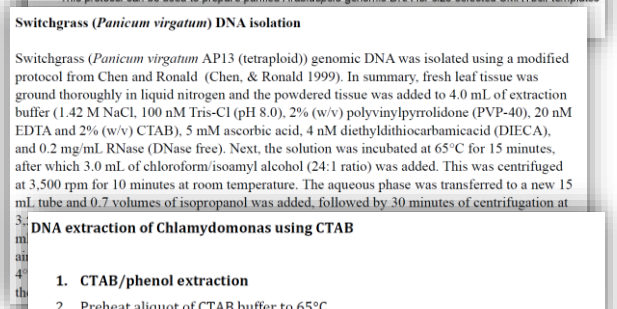

PACIFIC BIOSCIENCES

PacBio SampleNet – Shared Protocol

Please note: the shared protocols described herein may not have been validated by Pacific Biosciences and are provided as-is and without any warranty. Use of these protocols is offered to those customers who understand and accept the associated terms and conditions and wish to take advantage of their potential to help prepare samples for analysis using the PacBio® RS or RS II system. If any of these protocols are to be used in a production environment, it is the responsibility of the end user to perform the required validation.

Preparing *Arabidopsis* Genomic DNA for Size-Selected ~20 kb SMRTbell™ Libraries

This protocol can be used to prepare purified *Arabidopsis* genomic DNA for size-selected SMRTbell templates


Switchgrass (*Panicum virgatum*) DNA isolation

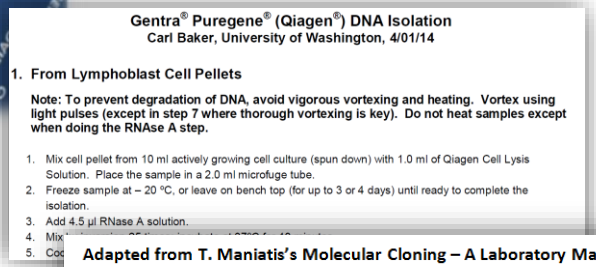
Switchgrass (*Panicum virgatum* AP13 (tetraploid)) genomic DNA was isolated using a modified protocol from Chen and Ronald (Chen, & Ronald 1999). In summary, fresh leaf tissue was ground thoroughly in liquid nitrogen and the powdered tissue was added to 4.0 mL of extraction buffer (1.42 M NaCl, 100 nM Tris-Cl (pH 8.0), 2% (w/v) polyvinylpyrrolidone (PVP-40), 20 nM EDTA and 2% (w/v) CTAB), 5 mM ascorbic acid, 4 nM diethyldithiocarbamicacid (DIECA), and 0.2 mg/mL RNase (DNase free). Next, the solution was incubated at 65°C for 15 minutes, after which 3.0 mL of chloroform/isoamyl alcohol (24:1 ratio) was added. This was centrifuged at 3,500 rpm for 10 minutes at room temperature. The aqueous phase was transferred to a new 15 mL tube and 0.7 volumes of isopropanol was added, followed by 30 minutes of centrifugation at

DNA extraction of Chlamydomonas using CTAB

1. **CTAB/phenol extraction**
2. Preheat aliquot of CTAB buffer to 65°C.

CTAB buffer	Stock	Vol
50 mM Tris-HCl (pH 8.0)	1M	10ml
20 mM EDTA	0.5M	4ml

Human Tissue


Gentra® Puregene® (Qiagen®) DNA Isolation
Carl Baker, University of Washington, 4/01/14

1. **From Lymphoblast Cell Pellets**

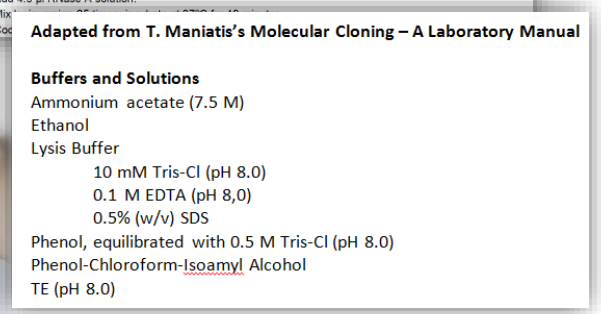
Note: To prevent degradation of DNA, avoid vigorous vortexing and heating. Vortex using light pulses (except in step 7 where thorough vortexing is key). Do not heat samples except when doing the RNAse A step.

1. Mix cell pellet from 10 ml actively growing cell culture (spun down) with 1.0 ml of Qiagen Cell Lysis Solution. Place the sample in a 2.0 ml microfuge tube.
2. Freeze sample at – 20 °C, or leave on bench top (for up to 3 or 4 days) until ready to complete the isolation.
3. Add 4.5 µl RNase A solution.

Unsupported Protocol – Gentra Puregene Cell Kit (Qiagen) DNA Isolation [Univ. Washington]

- http://www.pacb.com/wp-content/uploads/2015/09/Gentra-Puregene-Qiagen-DNA-Isolation.pdf

Reference: Resolving the complexity of the human genome using single-molecule sequencing. *Nature* **517**, 608-611 (29 January 2015) doi:10.1038/nature13907 http://www.nature.com/nature/journal/v517/n7536/full/nature13907.html

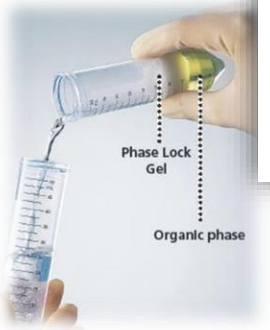

Adapted from T. Maniatis's Molecular Cloning – A Laboratory Manual

Buffers and Solutions

Ammonium acetate (7.5 M)
Ethanol
Lysis Buffer
10 mM Tris-Cl (pH 8.0)
0.1 M EDTA (pH 8,0)
0.5% (w/v) SDS
Phenol, equilibrated with 0.5 M Tris-Cl (pH 8.0)
Phenol-Chloroform-Isoamyl Alcohol
TE (pH 8.0)

Phenol Choloroform Extraction with Phase Lock Gel™ (PLG) [Sick Kids Hospital]

- Please see Sample Prep Workshop Customer Presentation *"Comparison of g-Tube and Needle Shearing"* by Karen Ng (OICR) or contact your local FAS



Insect Tissue

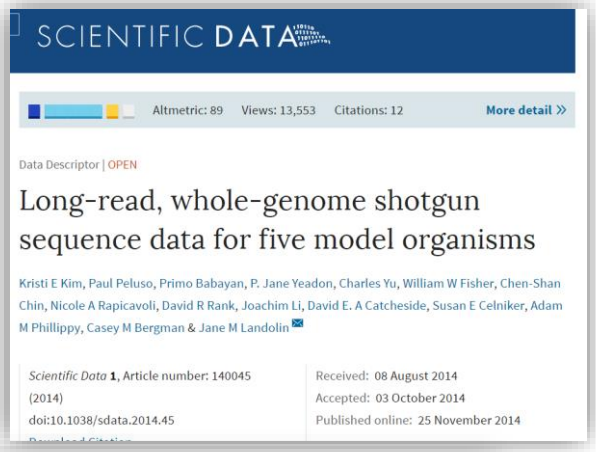
Phenol Chloroform Extraction (Drosophila melanogaster)

- Reference: Long-read, whole-genome shotgun sequence data for five model organisms. *Scientific Data* **1**:140045 (2014) DOI: 10.1038/sdata.2014.45 (Phenol Chloroform Method) http://www.nature.com/articles/sdata201445

Fish Tissue

QIAGEN Genomic-tip 20/100/500/G Kit

- https://www.qiagen.com/us/shop/sample-technologies/dna/dna-preparation/qiagen-genomic-tip-500g/#productdetails
- Reference: Chromosomal-Level Assembly of the Asian Seabass Genome Using Long Sequence Reads and Multi-layered Scaffolding. *PLOS Genetics* 12(4): e1005954. (2016) doi: 10.1371/journal.pgen.1005954 http://journals.plos.org/plosgenetics/article?id=10.1371%2Fjournal.pgen.1005954


SCIENTIFIC DATA

Altmetric: 89 Views: 13,553 Citations: 12 More detail »

Data Descriptor | OPEN

Long-read, whole-genome shotgun sequence data for five model organisms

Kristi E Kim, Paul Peluso, Primo Babayan, P. Jane Yeadon, Charles Yu, William W Fisher, Chen-Shan Chin, Nicole A Rapicavoli, David R Rank, Joachim Li, David E. A Catcheside, Susan E Celniker, Adam M Phillippy, Casey M Bergman & Jane M Landolin

Scientific Data **1**, Article number: 140045 (2014) doi:10.1038/sdata.2014.45

Received: 08 August 2014 Accepted: 03 October 2014 Published online: 25 November 2014

Other Application-Specific DNA Extraction Protocols

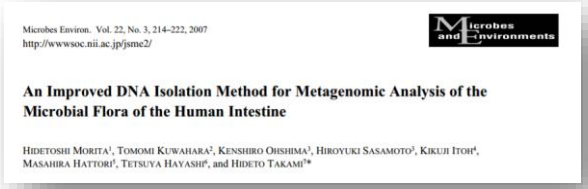
Metagenomic DNA Extraction Protocol (Tighe lab, University of Vermont)

Fecal Microbiome DNA Extraction Protocol (Microbes Environ. Vol. 22, No. 3, 214-222, 2007)

Stool and tissue DNA Extraction Protocol (Christian Hoffman, Instituto de Ciências Biológicas, Universidade Federal de Goiás)

Brain DNA extraction (Taylor Lab, Toronto Hospital for Sick Children)

Bacterial Genomic DNA Extraction (Ehrlich Lab, Allegheny Singer Research Institute)


Microbes Environ. Vol. 22, No. 3, 214–222, 2007
http://wwwsoc.nii.ac.jp/jsme2/

Microbes and Environments

An Improved DNA Isolation Method for Metagenomic Analysis of the Microbial Flora of the Human Intestine

HIDETOSHI MORITA[1], TOMOMI KUWAHARA[2], KENSHIRO OHSHIMA[3], HIROYUKI SASAMOTO[3], KIKUJI ITOH[4], MASAHIRA HATTORI[5], TETSUYA HAYASHI[6], and HIDETO TAKAMI[7]*

- Contact your Local Field Applications Scientist for details

DNA Sample Cleanup

General recommendations to help cleanup high molecular weight gDNA and SMRTbell Libraries:

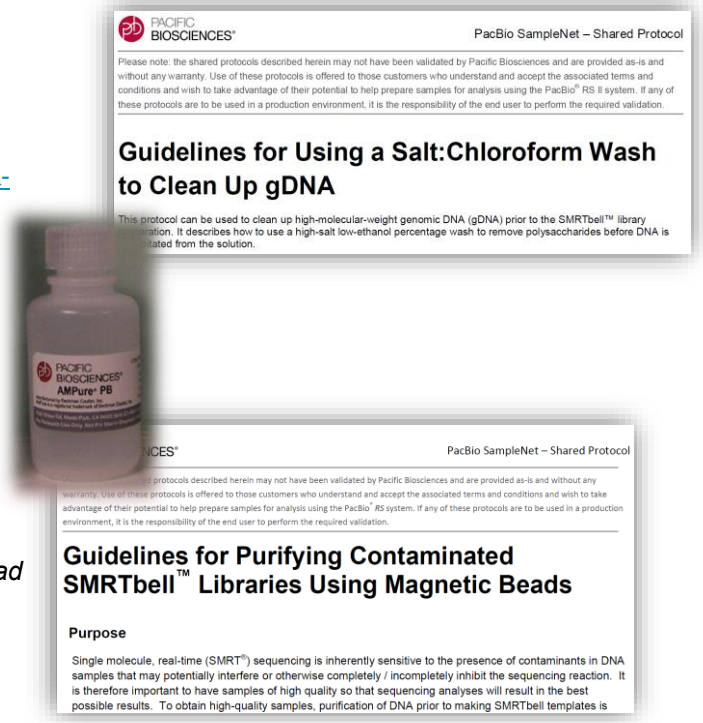
High Molecular Weight gDNA Cleanup

Unsupported Protocol – High Salt Phenol Chloroform Cleanup

- http://www.pacb.com/wp-content/uploads/2015/09/Shared-Protocol-Guidelines-for-Using-a-Salt-Chloroform-Wash-to-Clean-Up-gDNA.pdf

Ampure PB Bead Wash

- http://www.pacb.com/wp-content/uploads/2015/09/Guide-Pacific-Biosciences-Template-Preparation-and-Sequencing.pdf



SMRTbell Library Cleanup

Unsupported Protocol – Purification of Contaminated SMRTbell™ Library Using Magnetic Bead Capture

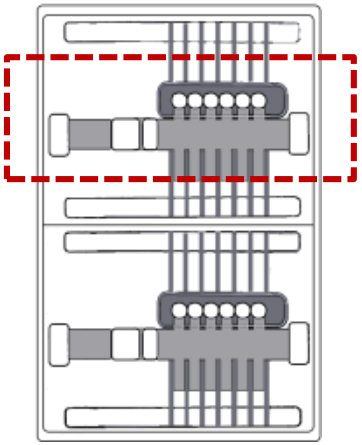
- http://www.pacb.com/wp-content/uploads/2015/09/Purifying-Contaminated-SMRTbell-Libraries-Using-MagBeads-052013.pdf

SageHLS: HMW Library System Overview

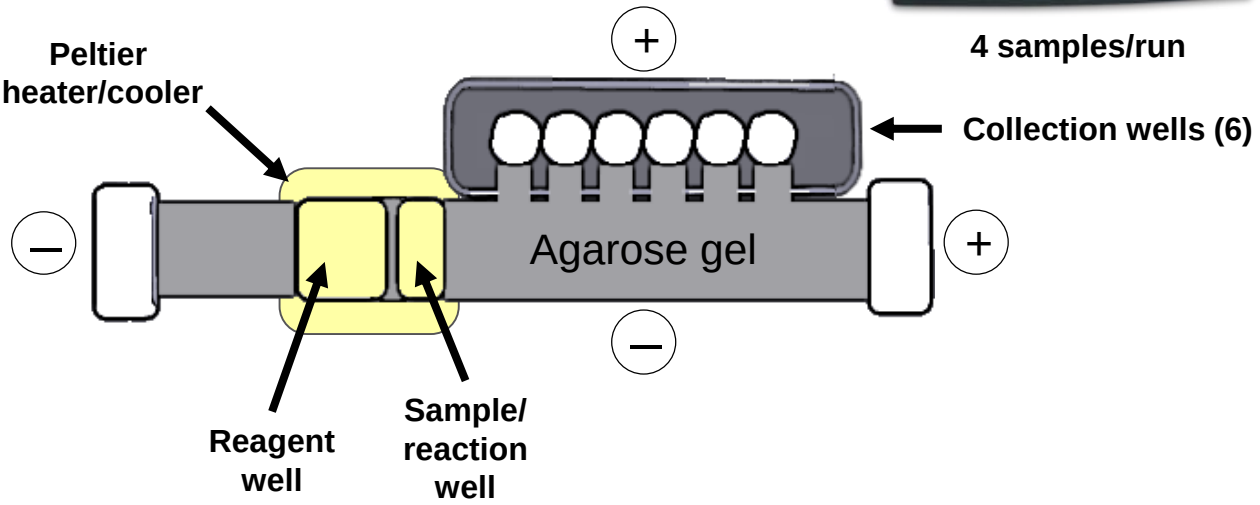
SageHLS (*NEW!*) provides a new approach to DNA isolation that enables extraction of DNA up to 2 Mbp directly from cells

- Users load intact cell or nuclei suspensions into agarose gel cassettes to perform lysis under electrophoretic conditions
- Degraded and solubilized proteins are removed but intact DNA is left immobilized in purified form within the cassette
- Processing enzymes are added to treat the DNA
- DNA products are size selected by electrophoresis, and recovered by electroelution into 6 size bins
- System capacity is 2 samples per cassette, 2 cassettes (4 samples) per run
- ~3 – 8 h run time (depending on the application and target DNA size)

4 samples/run



2 samples/cassette



Gel Cassettes enable sample processing *and* size-selection

sage science

Workflow: DNA Extraction and Enzymatic Reaction

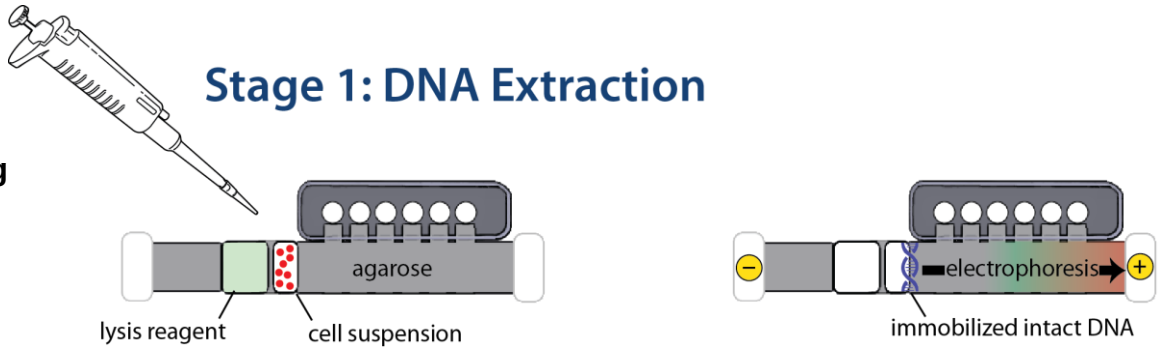
Current Applications:

- Ultra-High MW DNA extraction from cells
- ***C***as-9 ***A***ssisted ***T***argeting of ***CH***romosomal fragments (CATCH)

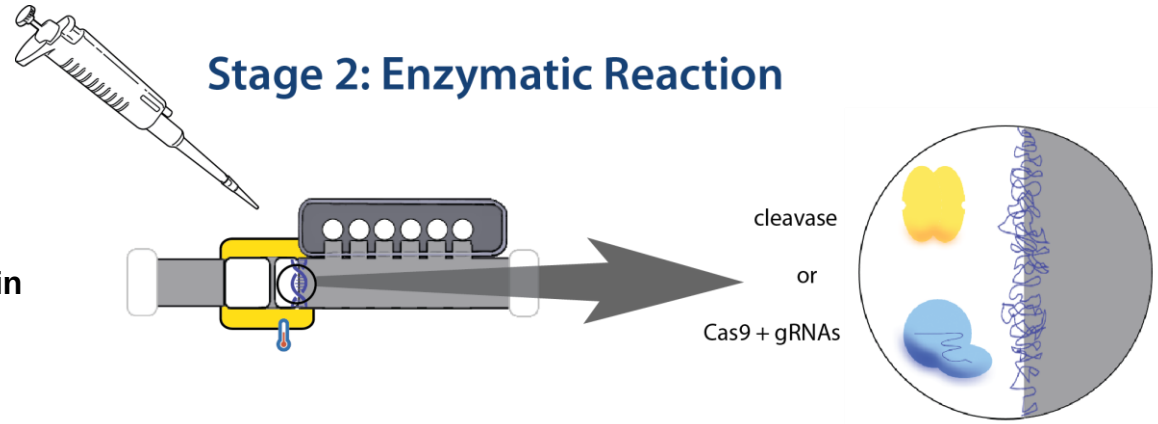
DNA Equiv. Load: 10 μg

Volume: 70 μl

Extraction Time: 1 h



Treatment Time: 30 min

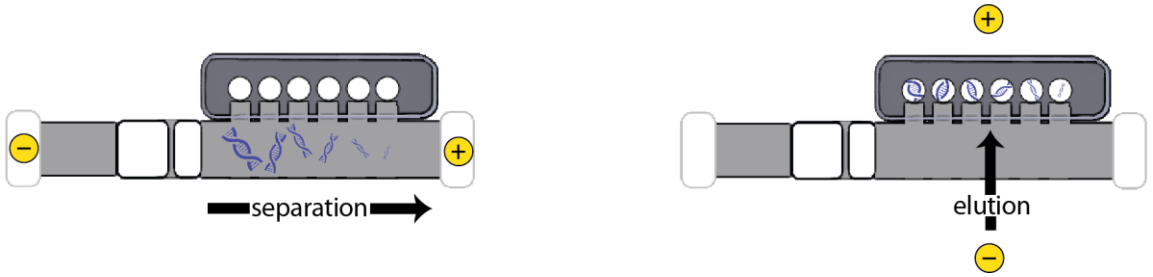


- **HMW DNA extraction** – DNA is purified and then digested lightly with NEB dsDNA Fragmentase to non-specifically fragment the DNA. HLS products can range from 10 kb to 2 Mb in size depending on digestion conditions and electrophoresis parameters.
- **CATCH** – DNA is purified and then digested with custom Cas9 cleavases designed to cut around genomic regions of interest. During size selection the targeted fragments are enriched since uncleaved genomic DNA remains entangled in the wall of the sample well during size selection and elution.

Workflow: DNA Size Selection

Stage 3: DNA Size Selection

Collection Time: 1 – 6 h



Example: Extraction of HMW DNA from White Blood Cells

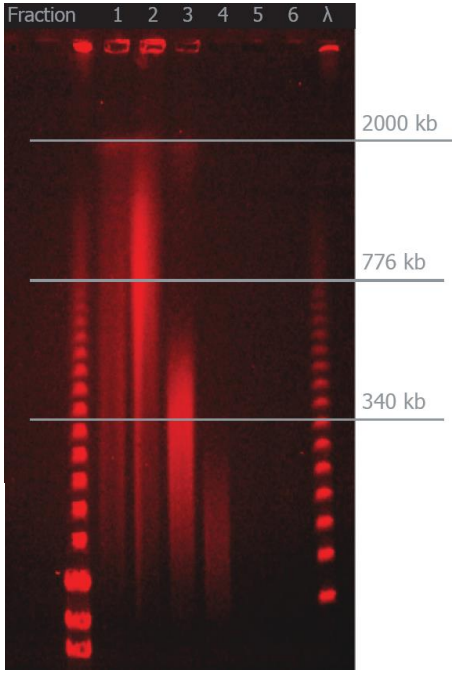
- White blood cells were prepared from whole blood using standard centrifugation techniques
- Cells (~6.6 x 10^4) were resuspended in 70 ml and loaded onto a SageHLS gel cassette
- Purification and cleaving processes (pipetting, electrophoresis, and incubation steps): ~90 min
- Size selection: ~3.5 h
- Total hands-on time: ~15 min

Input DNA: 8µg DNA Equivalents

Fraction	DNA (ng)	% yield
1	428	5.3%
2	1581	19.8%
3	1691	21.1%
4	314	3.9%
5	24	0.3%
6	26	0.3%
Total	**4064**	**51%**

DNA Recovery per Collection Well

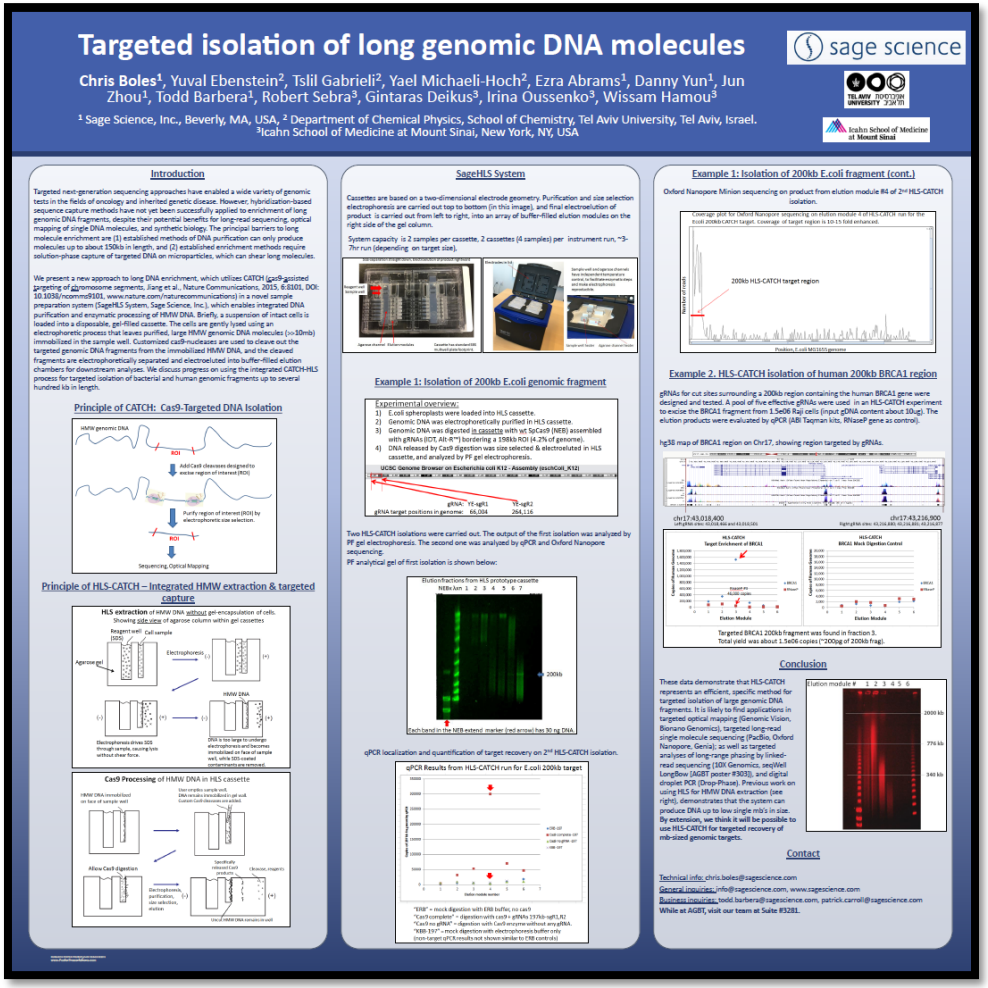
Pulsed-Field Gel Analysis



➢ Large size of the recovered DNA fragments potentially enables downstream analyses of genomic structure (e.g., structural rearrangements, copy number variation, and haplotype phasing)

SageHLS Product Information & Resources

http://www.sagescience.com/wp-content/uploads/2017/06/SageHLS-4page-brochure.pdf



http://www.sagescience.com/wp-content/uploads/2017/02/AGBT-Boles-2017.pdf

AGBT 2017

sage science

DNA SHEARING RECOMMENDATIONS

Recommendations for Shearing High-MW Genomic DNA for Preparing 20kb SMRTbell Libraries

PacBio's large insert size protocols have been validated using DNA fragmented with the Covaris® g-TUBE® device. With any system, there will be some variation in the distribution of the sheared fragments. In addition, some DNA will be lost during the shearing process itself.

- Depending on the quality of your starting material and the selected method of shearing, you may expect to lose 20% of the starting mass of your DNA sample.
- Shearing DNA Using a Covaris® g-TUBE® Shearing Device (>5 kb Insert Sizes):
  - The most up-to-date guidance on how to use the g-TUBE device, along with recommended centrifuges and centrifugation speeds, can be found in the g-TUBE device user manual available for download from the Covaris web site.
  - After the first centrifuge spin, check the upper chamber for residual liquid. Re-spin if necessary.
  - If there is still liquid in the chamber after 2 spins, use a 20 µL pipettor and pipette up and down several times. Then spin the tube down again.

http://covarisinc.com/products/g-tube/

Recommendations for Shearing High-MW Genomic DNA for Preparing >30 kb SMRTbell Libraries

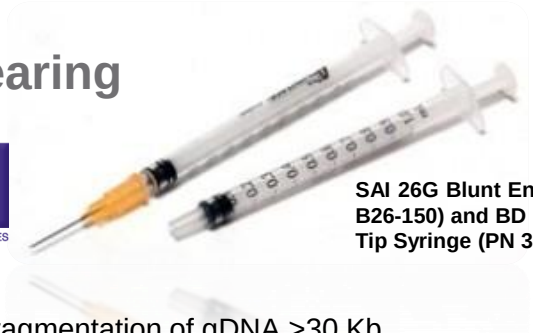
diagenode
Innovating Epigenetic Solutions

Megaruptor® Instrument

- Simple, automated, and reproducible device for the fragmentation of gDNA from 2 kb - 75 kb
- Uses mechanical shearing to fragment DNA by pumping a DNA solution through a disposable Hydropore shearing device containing a uniform array of pores
- Passage of the DNA molecules several times through the pores ensures that they will reach a minimum and uniform length as compared to a single pass through the Hydropore; thus a relatively tighter shearing size distribution profile is generated

Needle Shearing

BD SAI INFUSION TECHNOLOGIES



SAI 26G Blunt End Needles (PN B26-150) and BD 1-mL Luer-Lok Tip Syringe (PN 309628)

- Simple, manual shearing method using 26G needles for fragmentation of gDNA >30 Kb
- DNA solution is passed through the syringe needle multiple times (e.g., 5, 10, 20 times or more); the higher the number of passes, the smaller the resulting fragment size
- Relatively inexpensive shearing method – but the shearing size distribution profile is typically broader than that obtained with Megaruptor-treated samples

Performing Test Shears is Highly Recommended for Creating >30-kb SMRTbell Libraries

- Starting with high-quality, high-molecular-weight gDNA is required
- Perform Test Shears:
  - Perform test shears to target a fragment size distribution between ~40 kb – 60 kb
  - Shear by Megaruptor System or by Needle Shearing
- Verify fragment size distribution of test shears using pulsed-field gel electrophoresis prior to carrying out large-scale shears
- Perform large-scale shears using predetermined test shear conditions
  - Large-scale shears are carried out using the same concentration and conditions – but in a larger volume – to achieve a similar shearing distribution profile as the test shears
- Verify fragment size distribution of scale-up shears using pulsed-field gel electrophoresis prior to starting SMRTbell library construction

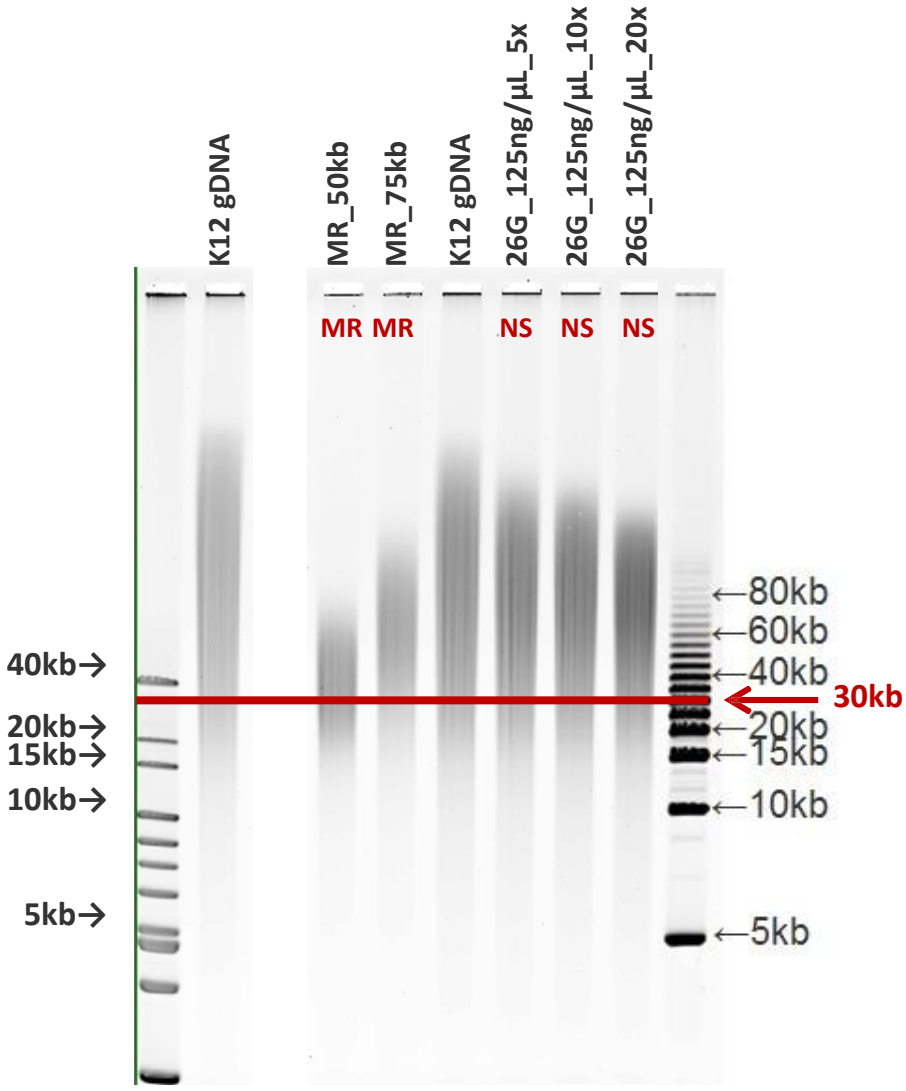
Comparison of Megaruptor vs. Needle Shearing Methods for Constructing >30 kb SMRTbell Libraries

Shearing Methods

- **MR** = Megaruptor System *(Setting 50, 75)*
- **NS** = Needle shearing *(26G needle, 125 ng/µL, 5-20 passes)*

Shearing Results

- Both MR and NS generated sheared gDNA samples with peak mode sizes >30 kb



NEEDLE SHEARING FOR >30 KB LIBRARIES

Supplies Recommended

- 26G Blunt End Needle (SAI Infusion B26-150)
- 1 mL Luer-Lok Tip Syringe (BD 309628)

Test shear first

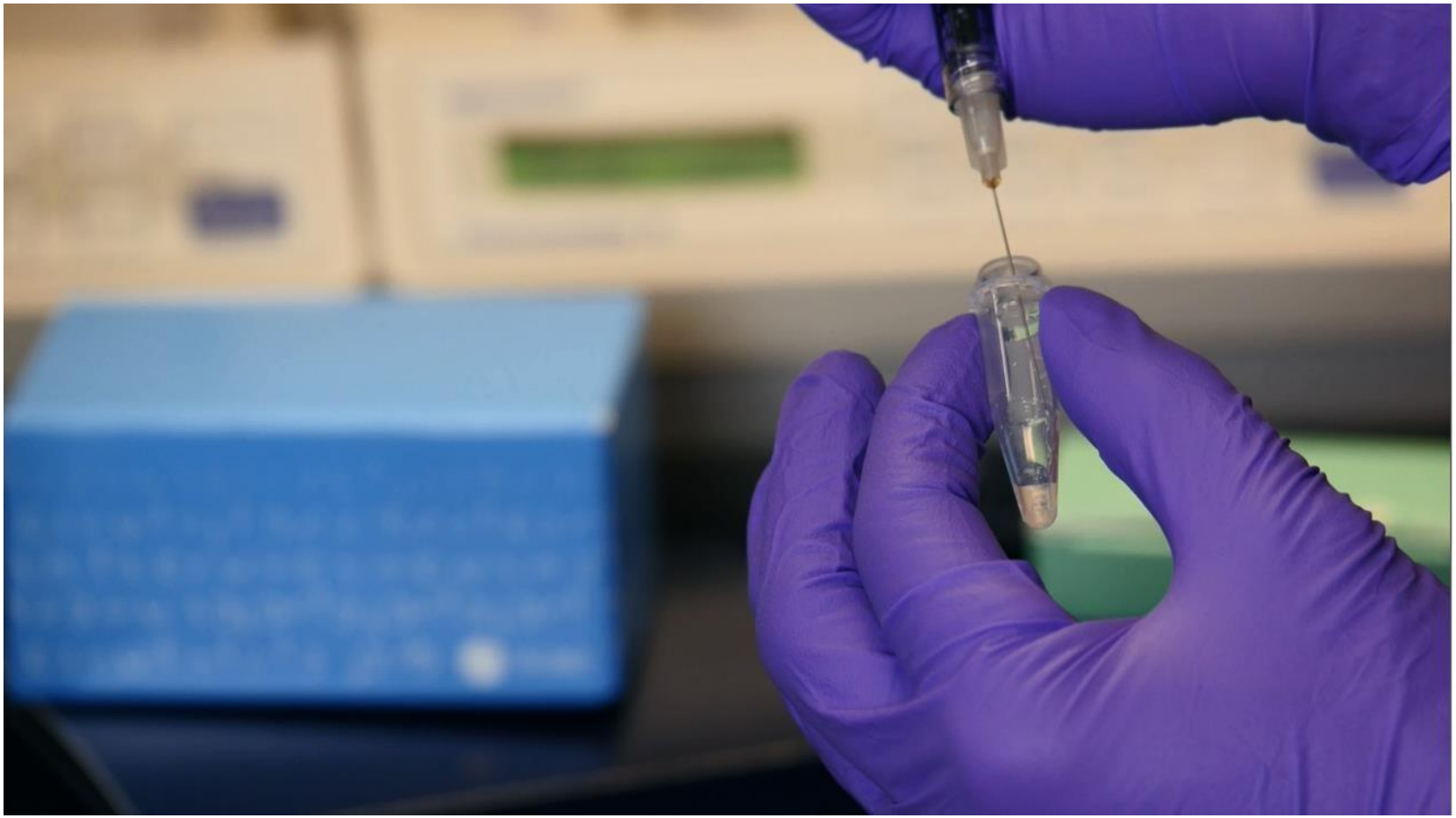
- Adjust sample conc. to 250 ng/µL
- Reserve 1 µL for QC at start of shearing
- Use 50µL sample in 1.5 mL Lo-Bind tube
- Aspirate entire sample volume through the needle 5 times, reserve 1 µL for QC
- Repeat aspiration/reservation for 10 to 20 times
- Run 1 µL reservations on PFGE to assess shearing

Scale up to mass shear remainder of gDNA sample

Video Demonstration of needle shearing available at: http://bit.ly/needle_shearing

VIDEO DEMONSTRATION OF NEEDLE SHEARING

See http://bit.ly/needle_shearing



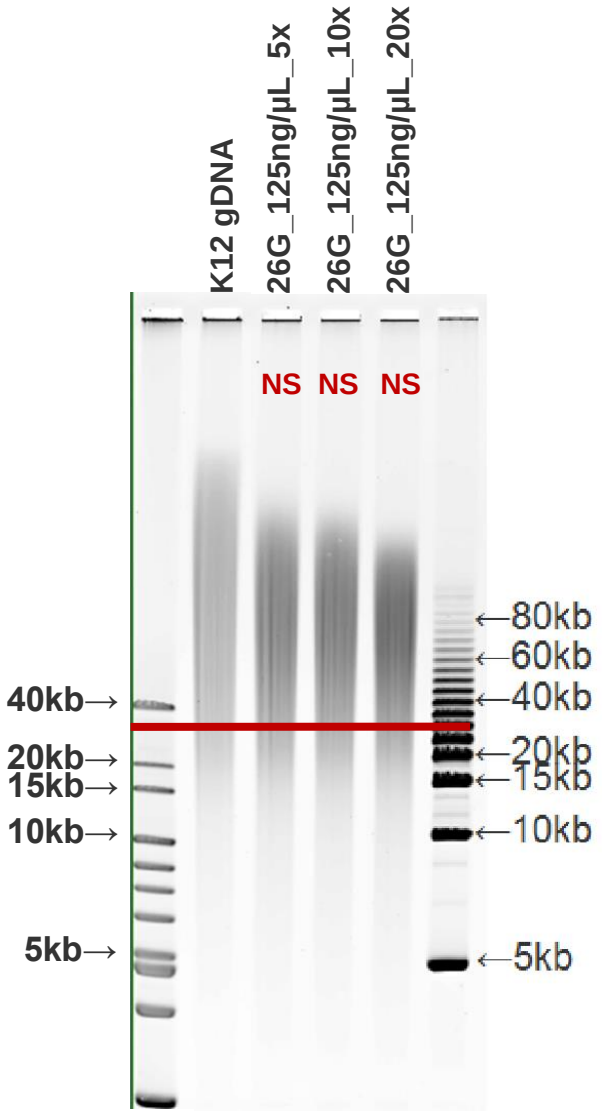
NEEDLE SHEARING FOR >30 KB LIBRARIES

Assess the Test Shear by PFGE

- If sample appears under-sheared:
  - decrease DNA concentration
    (for example, try 125 ng/µL)
  - and/or increase the number of passes until you achieve a similar distribution of fragmented gDNA.
- If sample appears over-sheared:
  - reduce the number of passes through the needle *(e.g., try 1 to 2 times)*

Scale up to mass shear remainder of gDNA sample

- Increase the volume while maintaining same concentration and ideal number of passes



Video Demonstration of needle shearing available at: http://bit.ly/needle_shearing

MEGARUPTOR DNA SHEARING SYSTEM OVERVIEW

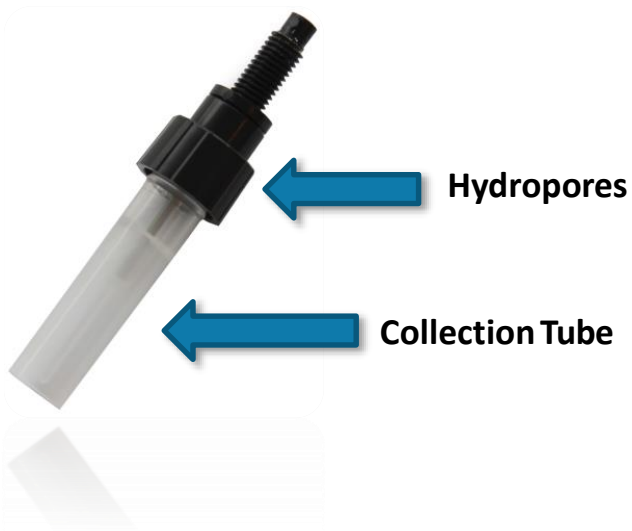
The Megaruptor is designed to provide researchers with a simple, automated, and reproducible device for the fragmentation of DNA from 2 kb - 75 kb.

- Uses mechanical shearing to fragment DNA
- As the DNA in solution is pumped through a Hydropore shearing device, it passes through an array of uniform pores.
- The resulting turbulent flow stretches and breaks the DNA strands.
- The length of the resulting fragments is dependent on the fluid flow rate and the size of the pores
- Passage of the DNA molecules several times through the pores ensures that they will reach a minimum and uniform length as compared to a single pass through the Hydropore device.

Diagenode, http://www.diagenode.com/
http://www.diagenode.com/en/catalog/shearing-technologies-54/megaruptor--101/

MEGARUPTOR SYSTEM TECHNICAL SPECIFICATIONS

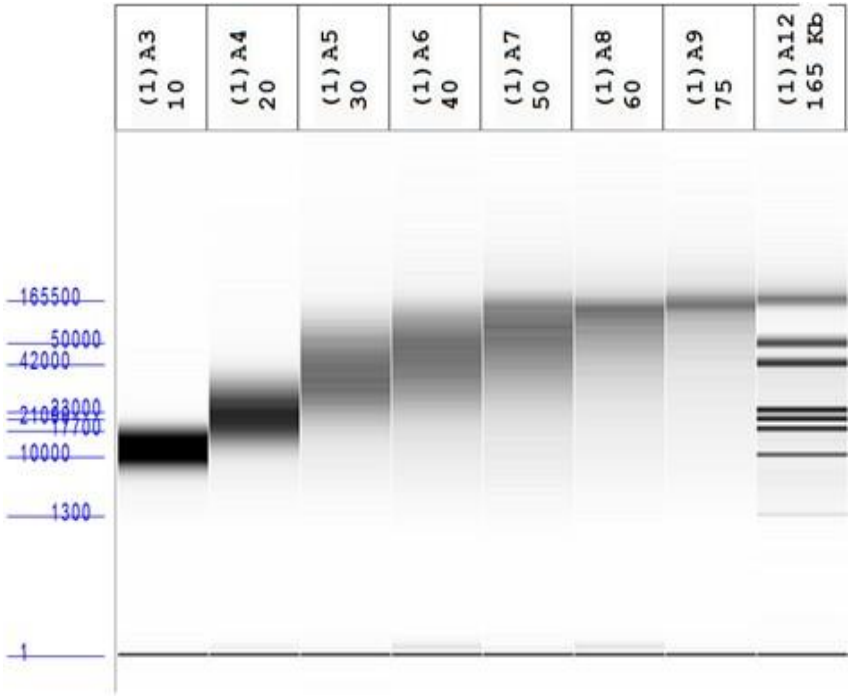
- Two Hydropore shearing devices available:
  - Short Fragment Hydropore: 2 kb – 9 kb
  - Large Fragment Hydropore: 10 kb – 75 kb
  - They each produce narrow distributions of fragments with the majority of molecules lying within a 2- to 3-fold size distribution.
- Hydropore devices are disposable to eliminate cross contamination
- Sample Volumes Accepted: 50 – 400 µl (Max.)
- Sample Concentration: Up to 50 ng/µl
- Time per sample: 10 - 20 min (depending on target size); includes sample processing and system washing



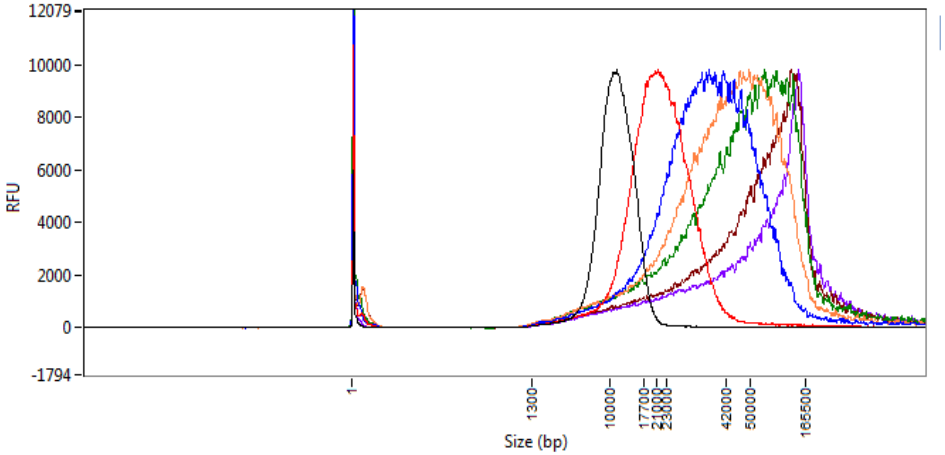
http://www.diagenode.com/en/catalog/shearing-technologies-54/megaruptor--101/
http://www.diagenode.com/media/catalog/file/Megaruptor_manual.pdf

MEGARUPTOR SHEARING EXAMPLE: HUMAN GDNA

To demonstrate shearing performance of the Megaruptor, a high molecular weight human genomic DNA was sheared to 10 kb, 20 kb, 30 kb, 40 kb, 50 kb, 60 kb, 75 kb. In this example, 30 kb, 40 kb and 50 kb shears are best conditions for constructing >30 kb libraries.

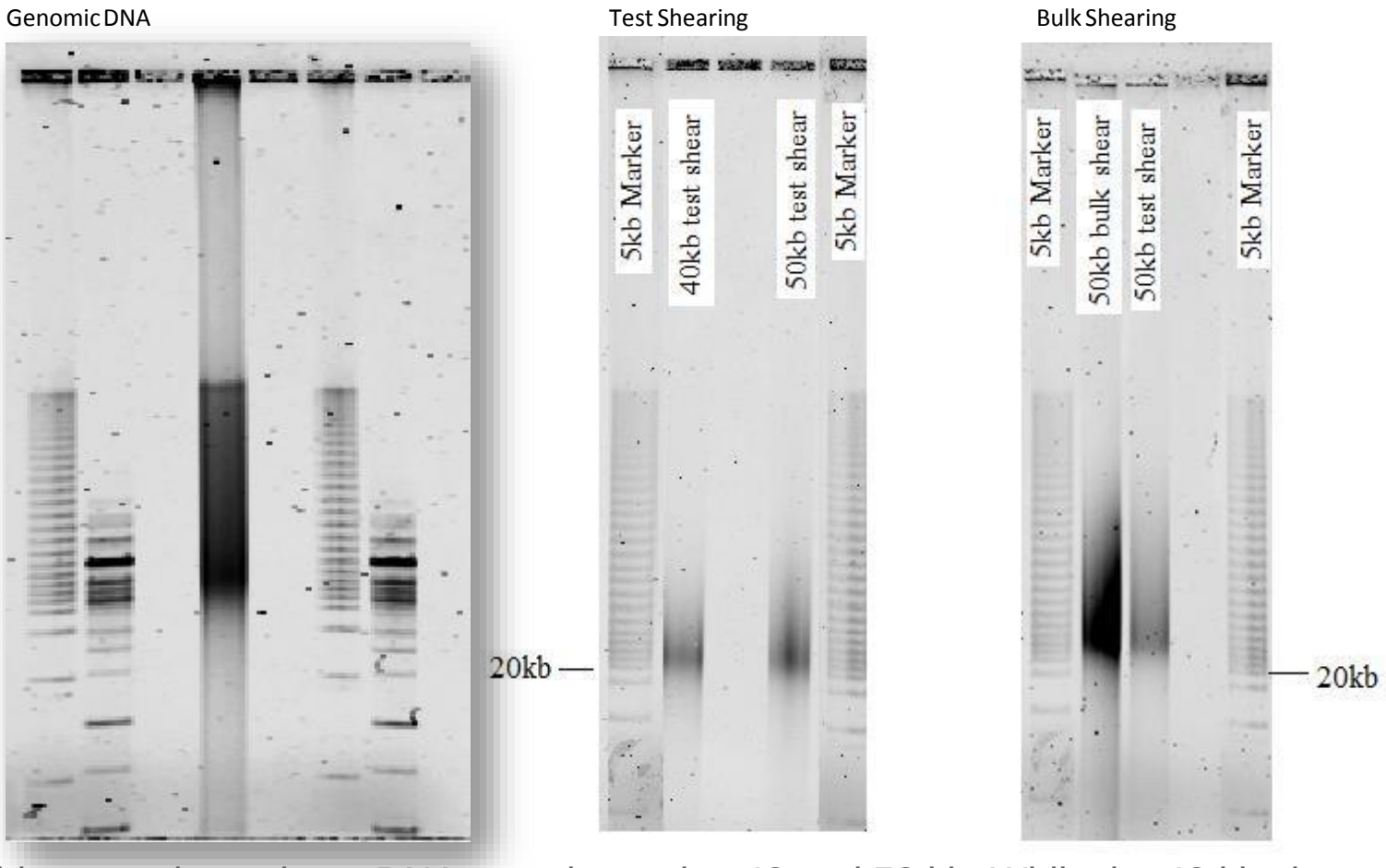


Megaruptor shears loaded on FEMTO Pulse



Label	Peak Max	Average
10 kB	11304	12245
20 kB	21245	24123
30 kB	36399	46143
40 kB	49495	60600
50 kB	78351	75416
60 kB	134525	92133
75 kB	149225	104512

MEGARUPTOR SHEARING EXAMPLE: PLANT GDNA



In this example, a plant gDNA was sheared to 40 and 50 kb. While the 40 kb shear generated good distribution, the 50 kb shear condition was selected for the large-scale shear because it provided slightly larger fragments

SMRTBELL LIBRARY CONSTRUCTION

Special Handling Recommendations for Preparing >30 SMRTbell Libraries

- Handle samples gently throughout extraction and library preparation
- Wide-bore pipette tips used throughout may be beneficial
- Finger-flick to mix is preferred over pipetting where possible
  - Tap the tube by hand until contents are mixed
  - Does not need to be overly rough, just make sure it's mixed
- End-over-end rotate samples to mix during AMPure PB steps
  - Rotate for 20 min (longer may be beneficial)
- Protocol directs ~2 min elution
  - Longer time elution time (20 min +) may increase recovery from AMPure PB beads
- Post-Size Selection DNA Damage Repair Step Helps Improve Read Length
  - Example: Mean mapped subread length is improved by ~ 2 kb for both 30 kb and 40 kb libraries after treatment of the size selected library with DNA Damage Repair enzymes

Mapped Subread Length	Inst ethan	Inst richard
NS_30	11796	11670
NS_30_DDR	13691	13917
NS_40	12487	12106
NS_40_DDR	14537	14186

DNA SIZE SELECTION FOR SMRTBELL LIBRARY CONSTRUCTION

sage science
BluePippin™

http://www.sagescience.com/product-support/bluepippin/
Cassette Definition: 0.*75% DF marker U1 high-pass 30-40kb v3*
Cassette Part PN: PAC30kb or BUF7510 (10-pk), BUF7503 (4-pk)

- Uses pulsed-field electrophoresis for resolving and collecting high molecular weight DNA samples (1 kb up to >50 kb)
- Up to 5 lanes (4 samples plus the External Marker lane) may be run on each gel cassette, with no cross contamination.
- Min. / Max. Load: 5 ng / 5 μg (sheared gDNA) per lane
- Typical run time for collection of >30 kb fragments is ~10 hours

sage science
PippinHT™

http://www.sagescience.com/product-support/pippinht-support/
Cassette Definition: *0.75% Agarose 30-40kb high-pass 75F*
Cassette Part PN: HPF7510 (10-pk), HPF7504 (4-pk)

- Flexible programming allows collection of multiple DNA size ranges across 2 gel cassettes using pulse-field electrophoresis
- Up to 24 lanes (22 samples plus 2 External Marker lanes) may be simultaneously run with 2 cassettes (12 lanes / cassette)
- Min. / Max. Load: Low ng / 1.5 μg (sheared gDNA) per lane
- Typical run time for collection of >30 kb fragments is ~7-9 hours

sage science
SageELF™

http://www.sagescience.com/product-support/pippinht-support/
Cassette Definition: *0.75% 10kb-40kb*
Cassette Part PN: ELD4010 (10-pk), ELD4004 (4-pk)

- Pulsed-field electrophoresis system separates large DNA samples by size, and then fractionates the whole sample into 12 fractions
- One sample is fractionated on a single precast agarose cassette, and 1 or 2 cassettes may be processed at one time.
- Min. / Max. Load: 100 ng / 5 μg (sheared gDNA) per lane
- Typical run time for collection of >30 kb fragments is ~9 hours

Size Selection Using Blue Pippin for >30 kb SMRTbell Library Construction

- Prior to size selection, perform qualitative analysis for size
- Blue Pippin: Prepare up to 4 - 5 µg SMRTbell templates in a final volume of 30 µL EB for each lane
- Follow Guidelines in the protocols to set up BluePippin System
- Must use BluePippin Software v6.20
- Must use “0.75%, DF Marker U1 high-pass 30 kb – 40 kb v3”
- A new size standard is required for this protocol: Use the U1 Marker
- Using the “Range” selection mode, enter “BPstart” value of 30000 or 40000 bp. A “BP End” value of 80000 bp should automatically appear
- Typical run times are ~10 hours
- To maximize recovery of eluted DNA, wait at least 45 minutes after the run terminates before removing the sample from the elution chamber
  - Collect the eluate into a 1.5 mL DNA LoBind tube
  - Wash elution well with 40 µL of Sage Science’s 0.1% Tween-20 Wash Solution, and add wash to eluted sample
  - Washing the elution well may increase yield 10-20%
- Determine average size of the library after BluePippin selection and DDR using PFGE or other sizing method

Cassette Definition: 0.*75% DF marker U1 high-pass 30-40kb v3*
Cassette Part PN: PAC30kb or BUF7510 (10-pk), BUF7503 (4-pk)

http://www.sagescience.com/product-support/bluepippin/

Size Selection Using PippinHT for >30 kb SMRTbell Library Construction

- Flexible programming allows collection of multiple DNA size ranges across 2 gel cassettes using pulse-field electrophoresis
- Up to 24 lanes (22 samples plus 2 External Marker lanes) may be simultaneously run with 2 cassettes (12 lanes / cassette)
- Min. / Max. Load: Low ng / 1.5 μg (sheared gDNA) per lane
- Typical run time for collection of >30 kb fragments is ~7-9 hours

Cassette Definition: *0.75% Agarose 30-40kb high-pass 75F*
Cassette Part PN: HPF7510 (10-pk), HPF7504 (4-pk)
http://www.sagescience.com/product-support/pippinht-support/

sage science
BluePippin™

SEQUENCING RECOMMENDATIONS

Annealing and Binding Recommendations

- In the Binding Calculator, use the average size determined after size selection and DDR treatment
- **PacBio RS II:**
  - Anneal the sequencing primer for 30 min at 20°C (thermocycler recommended)
  - Bind the polymerase for at least 30 min (up to 4 hours) at 30°C, then return to 4°C.
- **Sequel System:**
  - Anneal the sequencing primer for 60 min at 20°C (thermocycler recommended)
  - Bind the polymerase for 4 hours at 30°C, then return to 4°C

MagBead Binding Recommendations

- Always employ gentle treatment of MagBead-bound complexes to minimize any potential
- shearing damage
- For >30 kb SMRTbell libraries: During wash steps, do not pipet up and down to mix – instead, gently flow the MagBead Wash Buffer/MagBead Binding Buffer solution onto the bead pellet while the tube is still on the magnetic rack
- Process only one MagBead sample at a time, and as quickly as practicable, so that complexes never dry out during handling
- Keep MagBeads and MagBead-bound complexes on ice as much as possible
- **RSII:** 30 min MagBead binding at 4°C
- **Sequel:** 1 hour MagBead binding at 4°C

Loading Recommendations

- The table below compares loading requirements for large insert libraries (≥10 kb) for both PacBio RS II and the Sequel System, showing that the Sequel System requires lower on-plate loading concentrations than PacBio RS II.

Library Size	Sequel	RSII
10000	5 - 20 pM	40 pM
>20000	5 - 20 pM	>100 pM

Binding Coefficients and Loading Concentration Recommendations for PacBio RS II P6-C4 Chemistry

Diffusion vs. MagBead	Insert Size Range (bp)	Primer: Template Ratio		Polymerase: Template Ratio		P6v2 Loading Concentration (pM)	
		Small Scale	Large Scale	Small Scale	Large Scale	Small Scale - Size Selected	Small Scale - no Size Selection
Diffusion	101 bp - 300 bp	20	2	2:1	2:1	Not Tested	112.5
	301 bp - 750 bp	20	2	2:1	2:1	Not Tested	187.5
	751 bp - 1,500 bp	20	5	3:1	3:1	Not Tested	97.5
	1,501 bp - 3,000 bp	20	10	3:1	3:1	Not Tested	150
	3,001 bp - 7,500 bp	20	5	3:1	3:1	Not Tested	Not Tested
	7,501 bp - 50,000 bp	20	5	3:1	3:1	Not Tested	Not Tested
MagBead	751 bp - 1,500 bp	20	5	10:1	3:1	Not Tested	Not Tested*
	1,501 bp - 3,000 bp	20	10	10:1	3:1	Not Tested	10
	3,001 bp - 7,500 bp	20	5	10:1	3:1	Not Tested	10
	7,501 bp - 15,000 bp	20	5	10:1	3:1	40	10
	15,001 bp - 50,000 bp	20	5	10:1	3:1	100	Not Tested

For RSII, PacBio recommends using MagBead loading for insert sizes > 1 kb.

Binding Coefficients and Loading Concentration Recommendations for RSII Amplicon Samples 100 bp – 10 kb (P6-C4 Chemistry)

Insert Size Range	100 bp - 300 bp	301 bp - 999 bp	1 kb - 5 kb	5 kb - 10 kb
Run Protocol	Standard (Diffusion)	Standard (Diffusion)	MagBead OCPW or MagBead Standard	MagBead OCPW or MagBead Standard
Stage Start	No	No	1 kb - 3 kb (No) 3 kb - 5 kb (Yes)	Yes
On-Plate Loading Concentration (nM)	0.1 - 0.2 *(custom)*	0.2 - 0.45 *(custom)*	0.010 - 0.025 *(custom)*	0.025 - 0.040 *(custom)*
Primer:Template Ratio	5 *(custom)*	5 *(custom)*	20 (standard)	20 (standard)
Polymerase:Template Ratio	2 (standard) or 3 *(custom)*	2 (standard) or 3 *(custom)*	10 (standard)	10 (standard)

- The on-plate loading concentrations listed here are generally higher than the Binding Calculator recommendations, which were largely determined with sheared templates containing some smaller fragments.
- A range is provided since the optimal concentration for a given size range varies, depending on several factors, including application or project (target yield) and the relative sizes and abundances of contaminating amplicons.
- For projects with multiple SMRT Cells, a loading titration with a project-specific sample is recommended to identify the ideal on-plate loading concentration.
- As a general rule, 30-45% P1 is a good target, with P2 ideally <10%, maximally 20%, with P2<P0.
- For more information about using the Binding Calculator, see the Pacific Biosciences Template Preparation and Sequencing Guides and QRC - Annealing and Binding Recommendations.

Binding Coefficients and Loading Concentration Recommendations for Sequel System v2.0 Chemistry

Insert (bp)	Insert Range(bp)	Sample Type	Loading Method	Cleanup Method	Primer:Temp Ratio	Pol:Temp ratio	Pre-Extension*	Pre-Extension Time(min)	Immobilization Time	On-Plate Range (pM)
250	100-300	amplicon	Diffusion	n/a	10:1	3:1		n/a	120	0.66 – 2.6
500	301-750	amplicon	Diffusion	n/a	10:1	3:1		n/a	120	0.66 – 2.6
1000	751-1500	amplicon	Diffusion	Loading Clean up Beads	10:1	3:1	Recommended	TBD	120	1.3 - 2.6
2000	1501-3000	library	Diffusion	Loading Clean up Beads	10:1	3:1		n/a	120	3.3 - 5.2
2000	1501-3000	amplicon	Diffusion	Loading Clean up Beads	10:1	3:1	Recommended	TBD	120	2.6 - 5.2
5000	3001-7500	library	Diffusion		10:1	10:1		n/a	120	4-8
5000	3001-7500	amplicon	Diffusion	Loading Clean up Beads	10:1	10:1	Recommended	90**	120	4-8
10000	7501-15000	library	MagBead	Column	10:1	10:1		n/a	120	6.6 - 13.3
20000	15001-50000	library	MagBead	Column	10:1	10:1		n/a	120	6.6 - 25

*Tested on limited amplicons only

**Estimate pre-extension time by insert size (bp): (insert size x 2 / [(1.5 bases / s) x (60 s / min)]

Product Shelf Life and Storage Recommendations

- Pacific Biosciences' reagent kits are guaranteed a minimum shelf life of 30 days upon customer receipt
- For specific kit storage conditions, please see the product packaging and insert
- This table assists in determining shelf life and storage conditions of intermediate reagents and workflow products of Pacific Biosciences' Sample Preparation and Sequencing protocols

Workflow Category	Product	Shelf Life	Storage Conditions
Sample Preparation	AMPure® PB Beads	18 months	4ºC
	Sheared/Concentrated DNA	> 6 months	-20ºC
	70% Ethanol (For AMPure PB preparation)	< 2 weeks	RT
Template/ Complex Preparation	SMRTbell™ Template	> 6 months	-20ºC
	Diluted Sequencing Primer (< 1 µM)	Single Use	Ice
	Annealed SMRTbell Template	> 6 months	-20ºC
	Diluted Polymerase	Use Immediately After Diluting	Dispose Residuals
	Bound SMRTbell/Polymerase Complex:* – Standard Preparation – Concentrated Preparation in Storage Buffer (small scale and large scale)	 3 days > 3 months	 4ºC (never frozen) -20ºC
	MagBead Bound Template-Polymerase Complex	Keep On Ice - Use Immediately	Keep On Ice - Use Immediately
	Diluted On-Plate Complexes (< 3 nM)	Single Use	Ice
	Diluted Controls	Single Use	Ice

*For Sequel System, Bound SMRTbell/Sequel Polymerase v2 Complex has a shelf life of 7 days and should be stored at 4°C

http://www.pacb.com/wp-content/uploads/2015/09/QRC-Product-Shelf-Life-and-Storage-Recommendations.pdf

PACBIO RS II INSTRUMENT RUN TIMES AND SUGGESTED CELL MOVIE ACQUISITION TIMES FOR LARGE INSERT LIBRARIES (≥20 KB)

	8 Cells	16 Cells
4 Hour Movies	37 hours (1.5 days)	72 hours (3 days)
6 Hour Movies	53 hours (2.2 days)	104 hours (4.3 days)

Library Insert	Cut	Suggested Movie Time
20 kb	10 kb	4 hr
20 kb	>15 kb	4-5 hr
>30 kb	>15 kb	5-6 hr
>40 kb	>30 kb	6 hr

SEQUEL INSTRUMENT RUN TIMES: 6 HR *VS* 8 HR *VS* 10 HR

6hr MagBead		Process	Time to complete	Total Time (Min)	Total Time (Hr)	Days
	1	Up Front time (1st immobilization etc)	130	130		
	2	Chemistry	30	30		
	3	Movie time	360	2880		
	4	Post Primary	108	108		
		Number of SMRT Cells to run	8			
				3148	52	2.2
		MAX CELLS/Week	26			
		Througput per week (Gb)	**128**		*Est /Cell*	*5*

8 hr MagBead		Process	Time to complete	Total Time (Min)	Total Time (Hr)	Days
	1	Up Front time (1st immobilization etc)	130	130		
	2	Chemistry	30	30		
	3	Movie time	480	3840		
	4	Post Primary	144	144		
		Number of SMRT Cells to run	8			
				4144	69	2.9
		MAX CELLS/Week	19			
		Througput per week (Gb)	**97**		*Est /Cell*	*5*

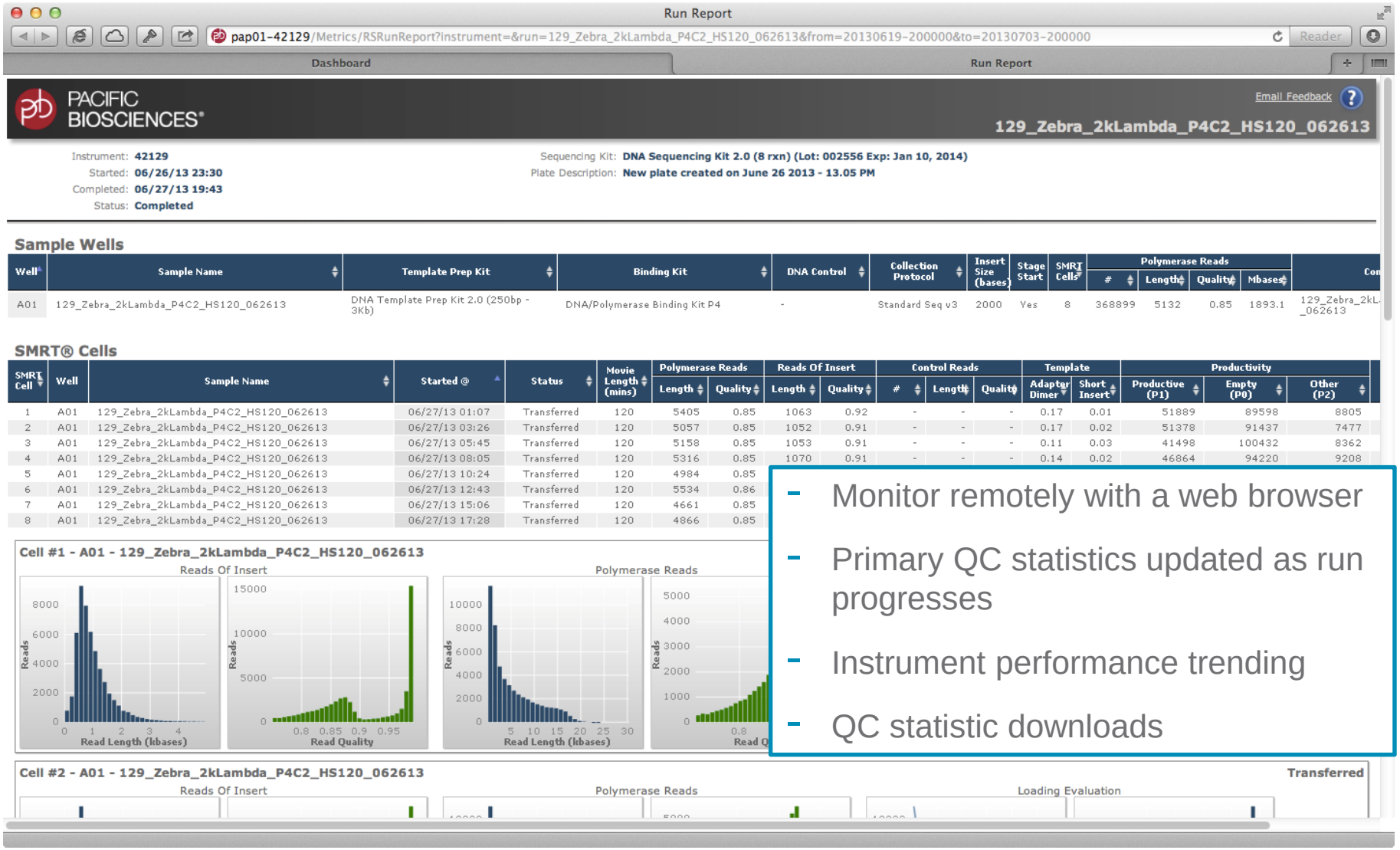
10 hr MagBead		Process	Time to complete	Total Time (Min)	Total Time (Hr)	Days
	1	Up Front time (1st immobilization etc)	130	130		
	2	Chemistry	30	30		
	3	Movie time	600	4800		
	4	Post Primary	180	180		
		Number of SMRT Cells to run	8			
				5140	86	3.6
		MAX CELLS/Week	16			
		Througput per week (Gb)	**78**		*Est /Cell*	*5*

- Sequel instrument can continuously run up 72 hours of movie collection times.
  - For example, the 72 hours would include running seven (7) SMRT Cells for 10 hours, 12 SMRT Cells for 6 hours, etc.
- You must verify that your network can accommodate the data transfer space for these runs.
- If all runs have successfully transferred off of the instrument, the instrument can store approximately 24 hours of run space.

Sequel Instrument Run Time Estimator Excel spreadsheet is available upon request; please contact your local Field Applications Scientist

Run QC

RS DASHBOARD: MONITOR AND QC PACBIO RS II RUNS IN A WEB BROWSER

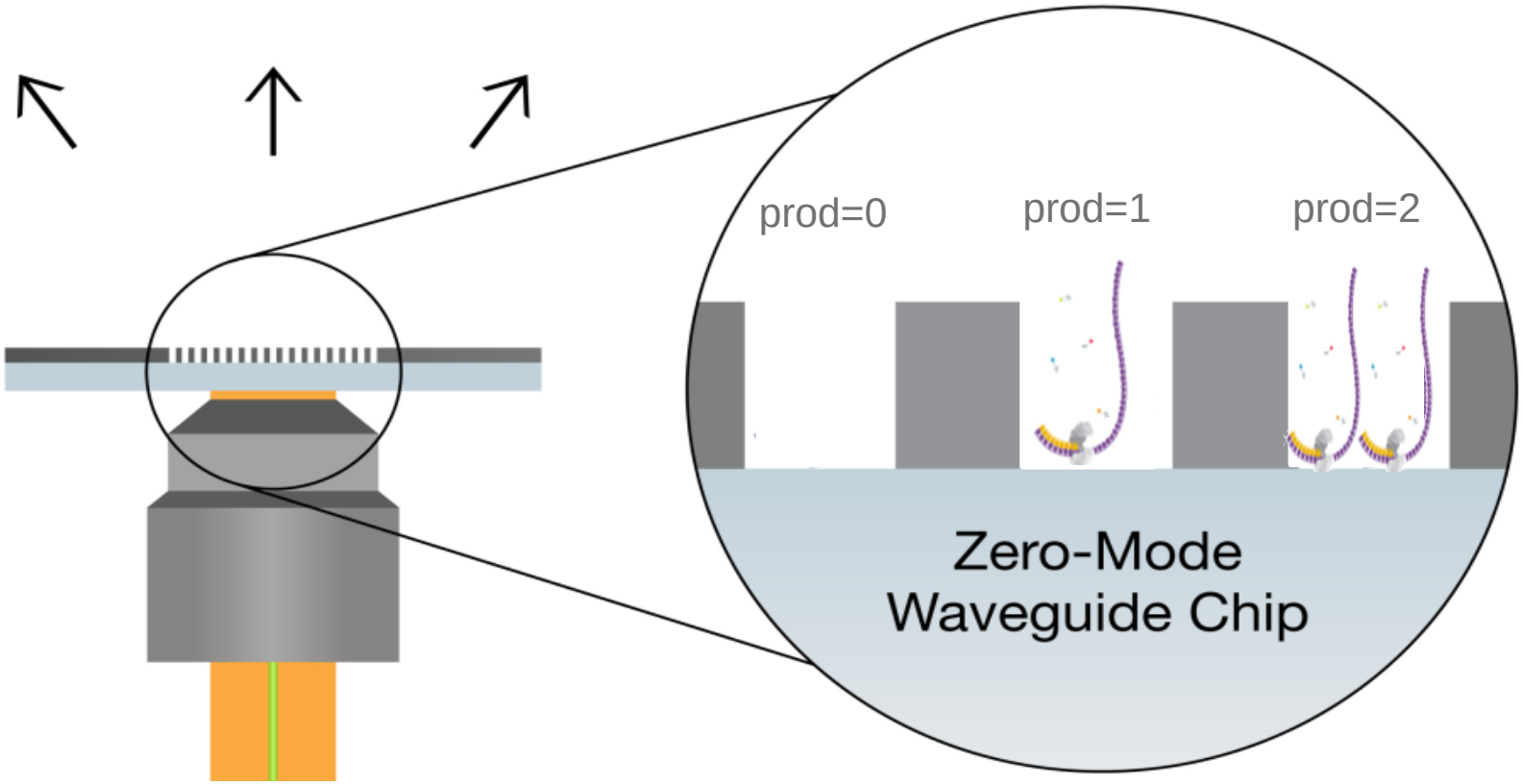


- Monitor remotely with a web browser
- Primary QC statistics updated as run progresses
- Instrument performance trending
- QC statistic downloads

RS Dashboard Overview: *http://www.pacb.com/training/RSDashboardOverview/story.html*

PRODUCTIVITY

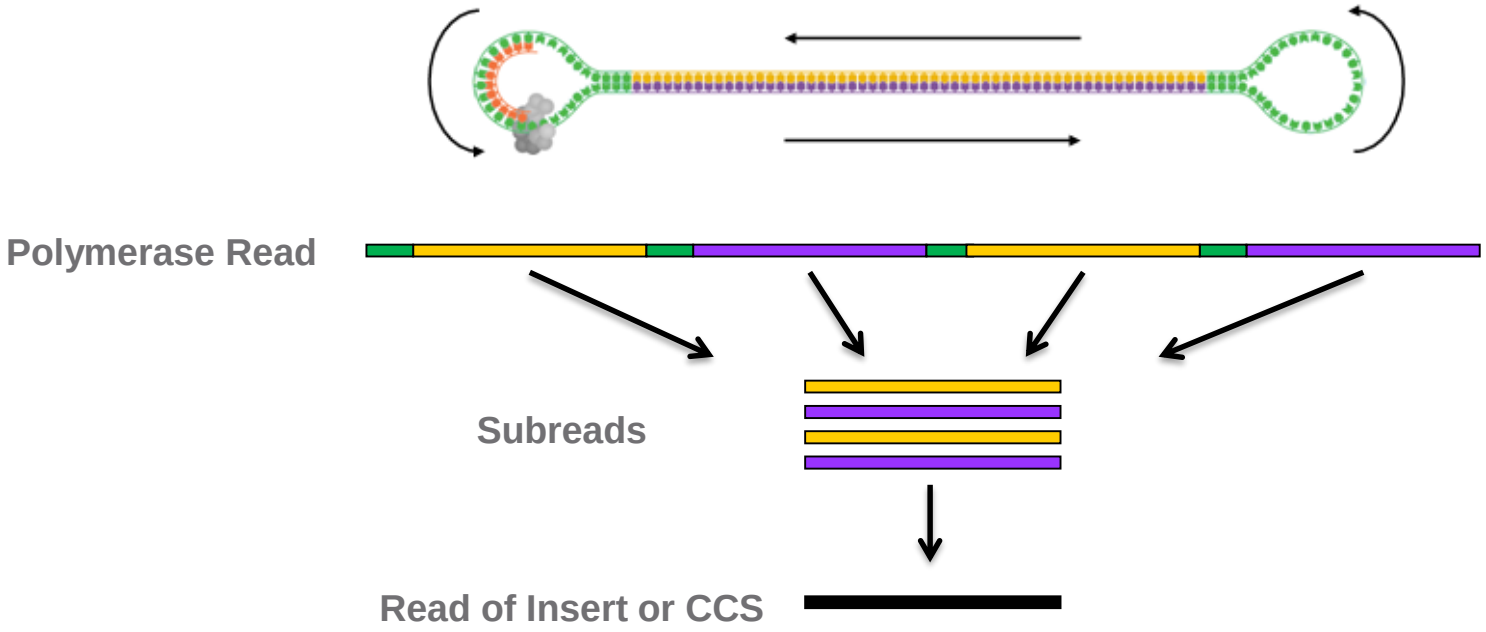
- Productivity: An estimate of the number of active polymerases in a ZMW
- P=0 means that a ZMW did not produce a read and is presumed to be lacking a polymerase.
- P=1 means that there is a polymerase read from that ZMW.
- P=2 means “other” and the signal collected from the ZMW was not conducive to efficient base calling, possibly due to multiple template-polymerase complexes bound in the ZMW, high background signal, etc.
- Raw reads generated by P=1 ZMWs are trimmed to include only the High Quality Region (HQR)



READ QUALITY (RQ) VALUE

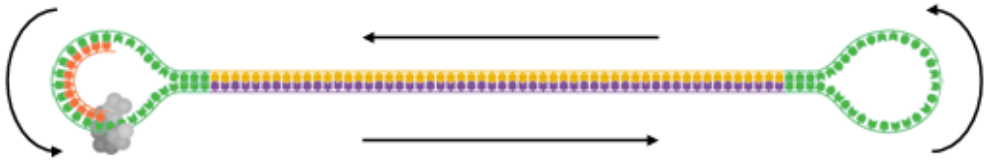
- A trained prediction (*de novo* estimate) of a read's mapped subread accuracy
- Based on its pulse and base-file characteristics, such as:
  - Peak signal-to-noise ratio
  - Average base QV
  - Inter-pulse duration
- Used during secondary analysis filtering

FROM POLYMERASE READS TO SUBREADS AND READ OF INSERT

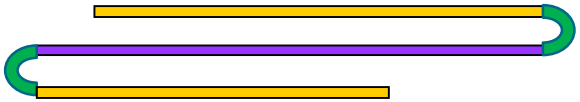


- **Polymerase Read**: A sequence of nucleotides incorporated by the DNA polymerase while reading a SMRTbell template. They can include sequences from adapters and from one or multiple passes around a circular template, which includes the insert of interest.
- **Subread** (purple and gold): Each polymerase read is partitioned to form one or more subreads, which contain sequence from a single pass of a polymerase on a single strand of an insert within a SMRTbell template and **no** adapter sequences (green)
- **Read of Insert:** Represents the highest quality single-sequence for an insert, regardless of number of passes
  - CCS read requires ≥2 full passes (Both adapters must be detected for a read to be identified as “full pass”)
- Either individual subread, read of insert or CCS can be used for subsequent secondary analysis depending on application needs

RS DASHBOARD READ METRICS DEFINITIONS AND THEIR UTILITY



SMRTbell Template



Polymerase Read

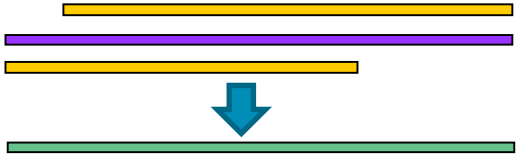
Definition:
- Sequence of nucleotides incorporated by polymerase while reading a template
- High Quality region only
- Includes adapters
- 1 molecule, 1 pol. Read

Purpose:
- QC of instrument run
- Benchmarking

Subreads

Definition:
- Adapters removed
- 1 molecule, 1 or more subreads

Purpose:
- Used for data analysis

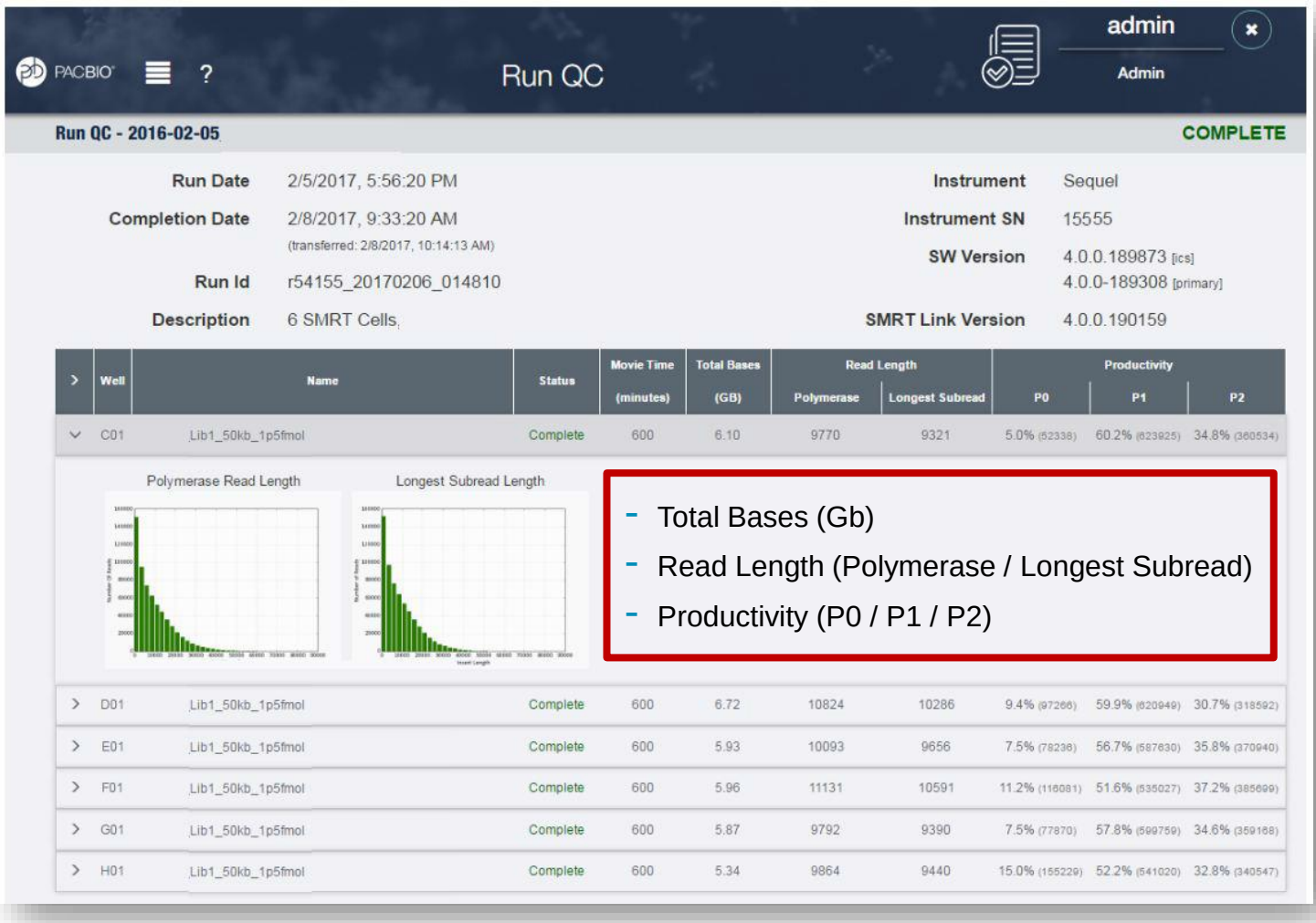


Read of Insert

Definition:
- Represents highest-quality single-sequence for an insert, regardless of number of passes
- 1 or more passes
- 1 molecule, 1 read of insert

Purpose:
- Library size QC
- Used for data analysis

SMRT LINK RUN QC MODULE: MONITOR AND QC SEQUEL SYSTEM RUNS IN A WEB BROWSER



Sequel System Operations Guide: *http://www.pacb.com/wp-content/uploads/Operations-Guide-Sequel-System.pdf*

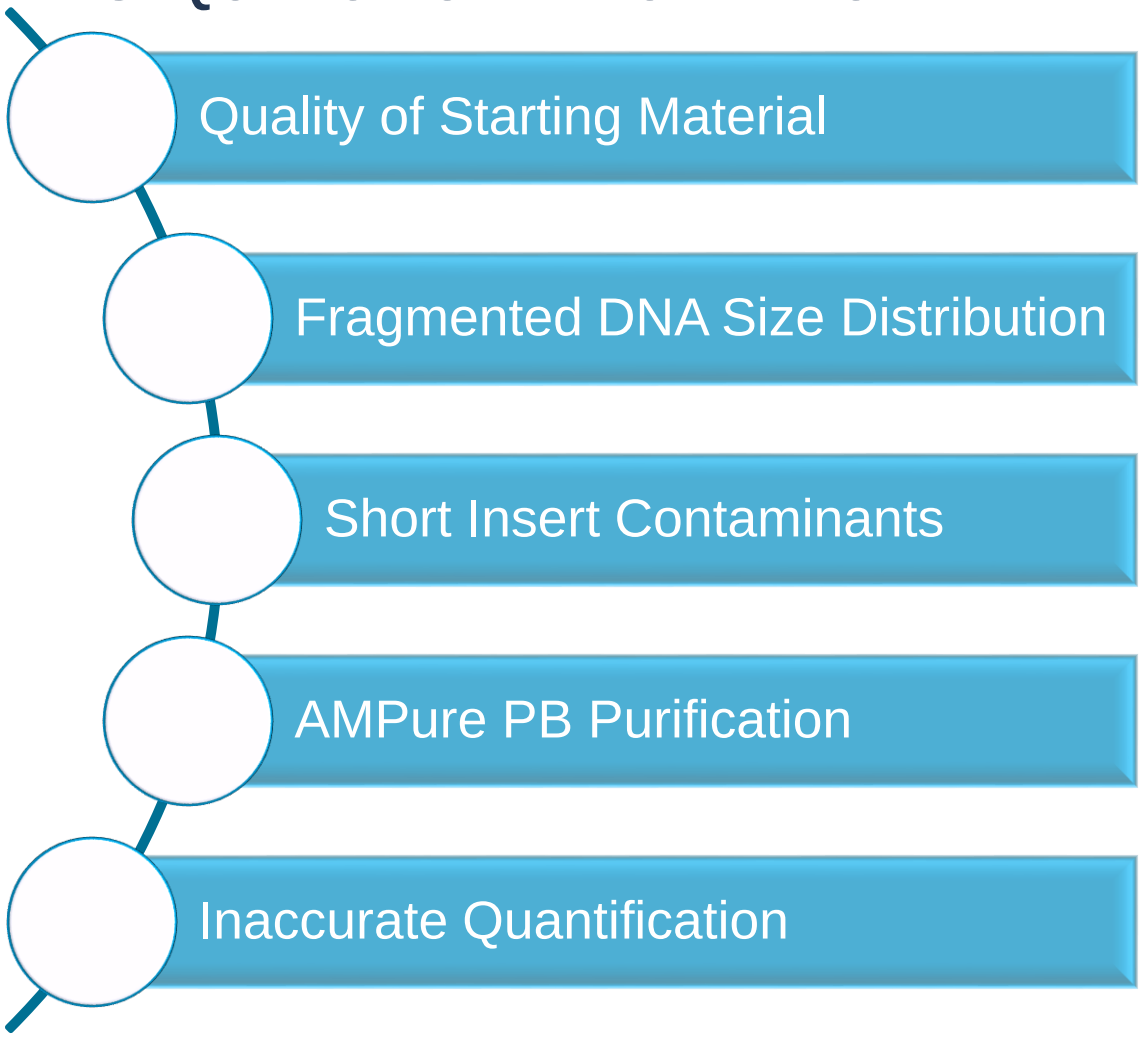
SEQUEL SYSTEM RUN QC METRICS TERMINOLOGY

- **Total Bases:** Sum all of bases from high quality reads.
- **Polymerase Read Length:** Mean polymerase read length (bp)
- **Longest Subread Length:** Mean of the maximum subread length (bp) per ZMW
- **Productivity:**
  - P0: Empty ZMW; no high quality read detected.
  - P1: ZMW with a high quality read detected.
  - P2: Other, signal detected but no high quality read.

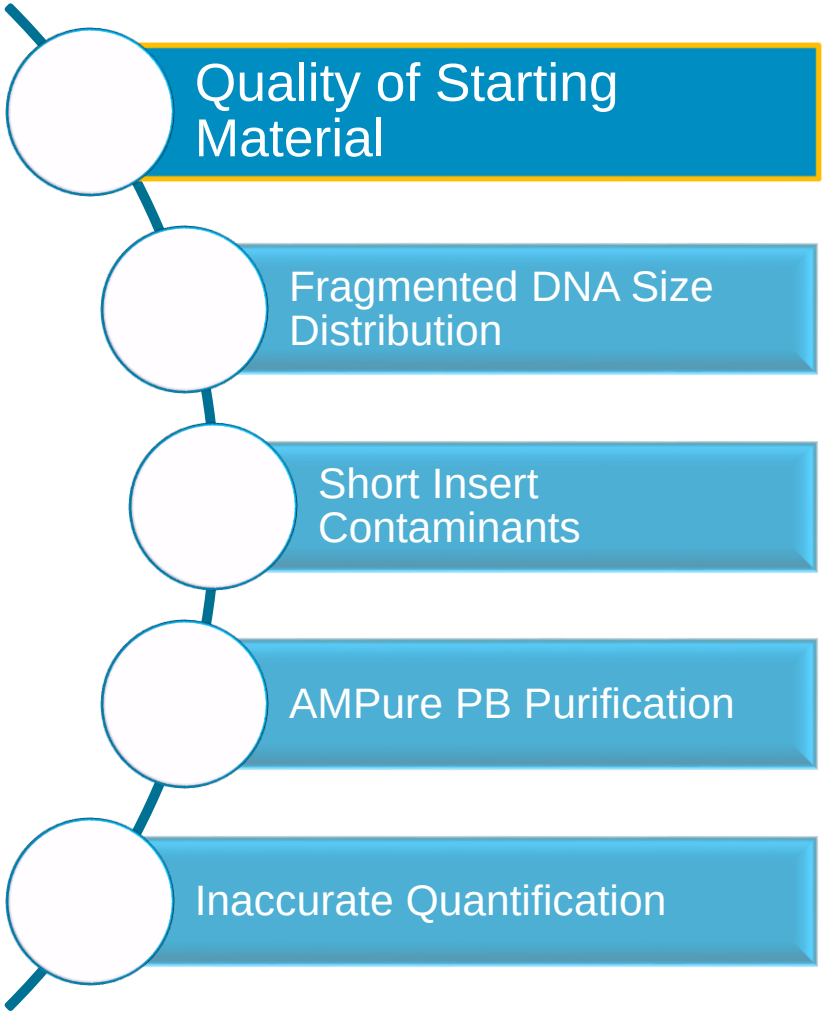
SMRT Link User Guide: *http://www.pacb.com/wp-content/uploads/SMRT-Link-User-Guide-v4.0.0.pdf*

Troubleshooting SMRTbell Library Sequencing Performance

FACTORS AFFECTING SMRTBELL LIBRARY PREPARATION AND SEQUENCING PERFORMANCE



FACTORS AFFECTING SMRTBELL LIBRARY PREPARATION AND SEQUENCING PERFORMANCE



Causes

- DNA damage (nicked DNA, abasic sites, etc.)
- Carry over contaminants

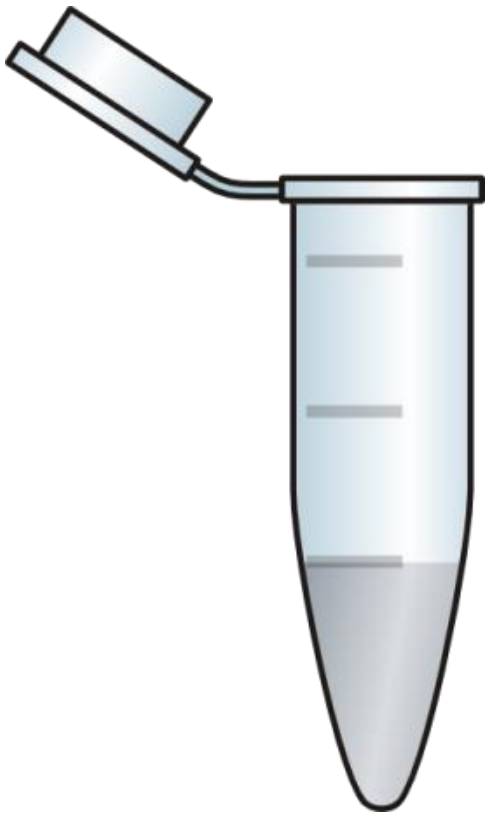
Impact on Performance

- Low SMRTbell template yields
- Lower sequencing yields
- Shorter read lengths

Recommendations

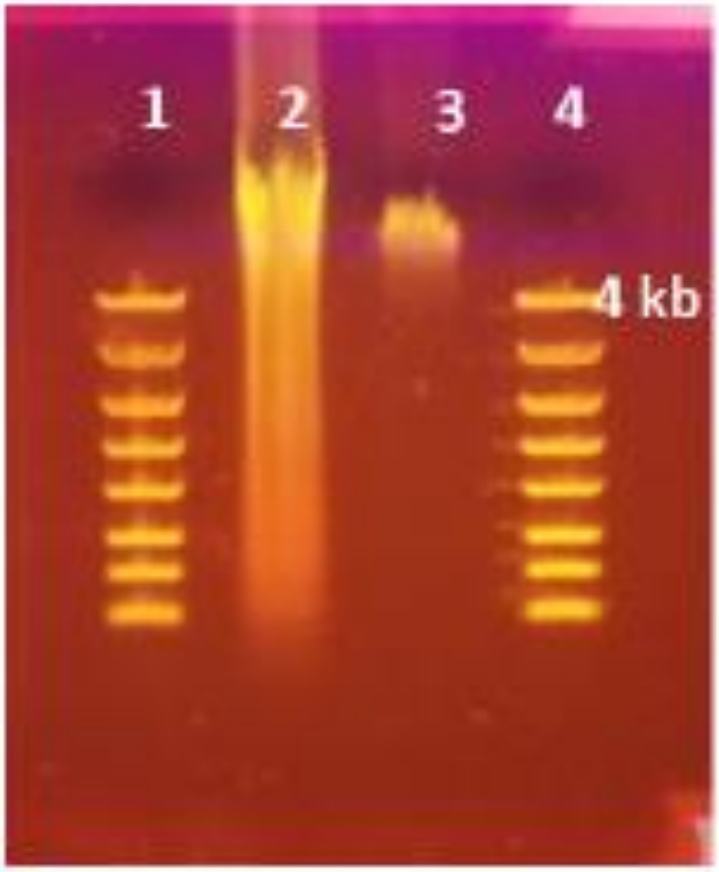
- Treat your gDNA gently during extraction
- Always perform DNA damage repair
- Purify starting DNA material to remove contaminants

SAMPLE CONDITIONS THAT LEAD TO HIGHER QUALITY LIBRARIES



- Double-stranded DNA Sample (dsDNA)
- Minimized freeze-thaw cycles
- No exposure to high temp (>65° C)
- No exposure to pH extremes (<6 or >9)
- OD260/280 between 1.8 and 2.0
- OD260/230 between 2.0 and 2.2
- No insoluble material
- No RNA contamination
- No exposure to UV or intercalating fluorescent dyes
- No chelating agents, divalent metal cations, denaturants, or detergents
- No carryover contamination (e.g. polysaccharides) from starting organism

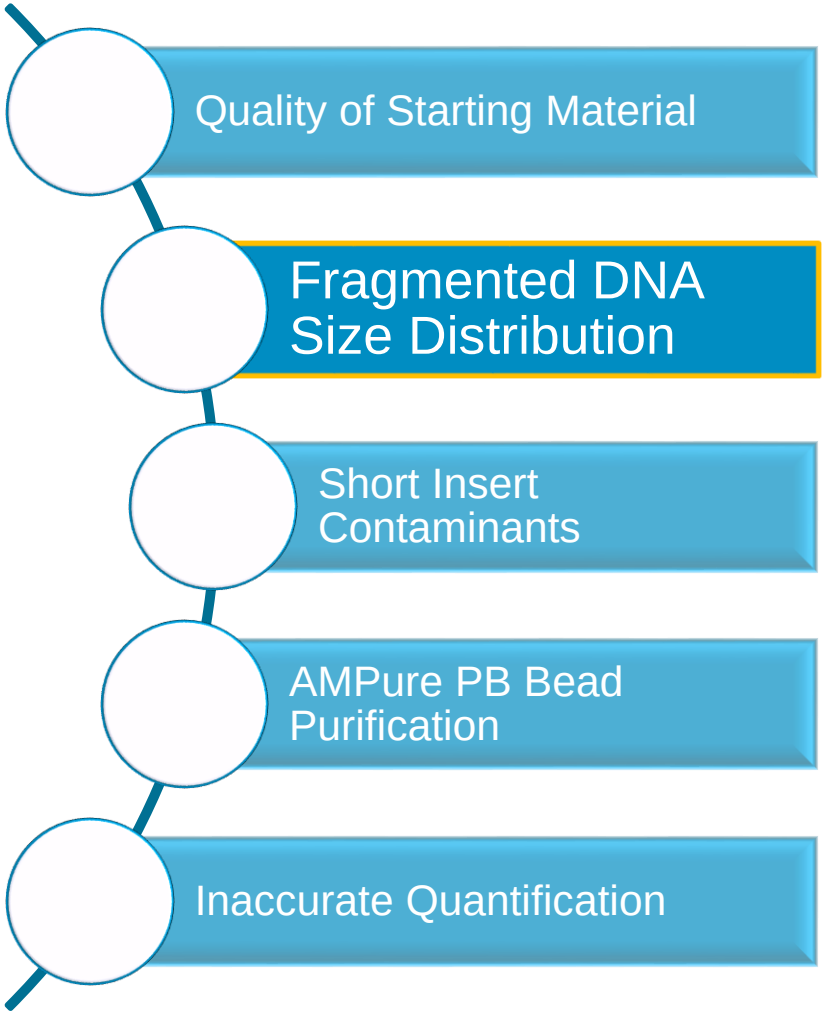
ALWAYS ASSESS THE QUALITY OF DNA



1.2% Lonza FlashGel System
Lanes 1 and 4: Lonza FlashGel Ladder (100 bp – 4 kb
Lane 2: degraded gDNA
Lane 3: intact DNA

- Always assess the quality of the gDNA prior to library construction
- Degraded DNA will adversely affect the quality and fragment distribution of the library
- For PCR, non-specific PCR products generate sequencing noise

FACTORS AFFECTING SMRTBELL LIBRARY PREPARATION AND SEQUENCING PERFORMANCE



Causes

- Insert size distribution
- Loading bias

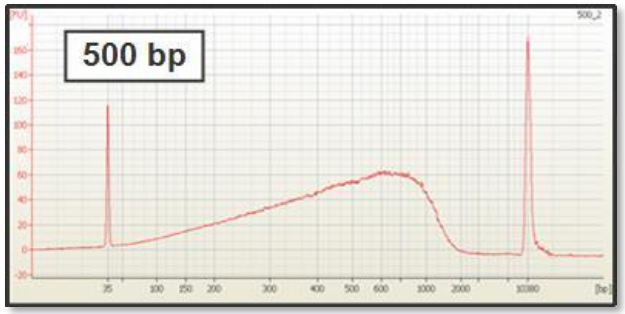
Impact on Performance

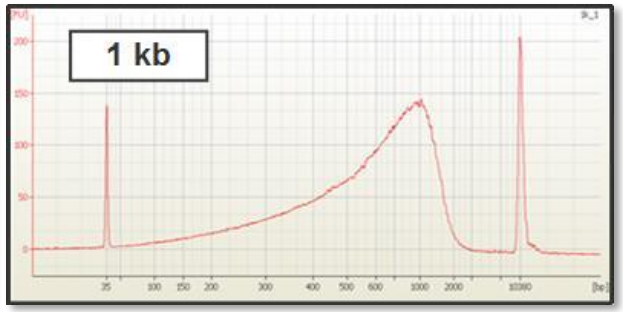
- Shorter subread lengths
- Preferential loading of shorter templates

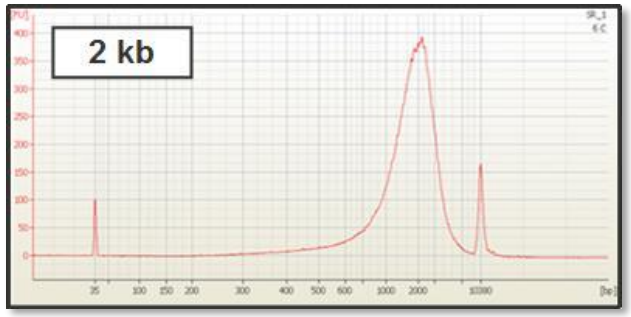
Recommendations

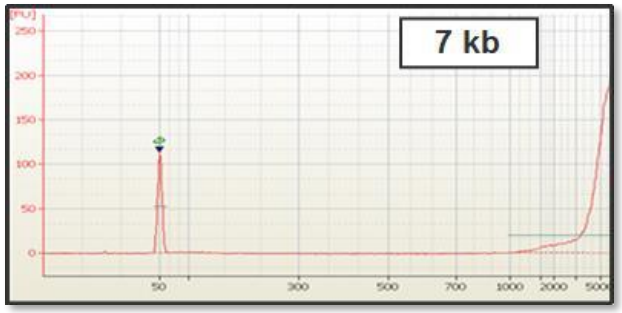
- Target larger insert shear sizes to maximize subread length
- Size selection
- Proper AMPure PB purification
- Removal of shorter fragments
- When pooling amplicons, aim for similar sizes

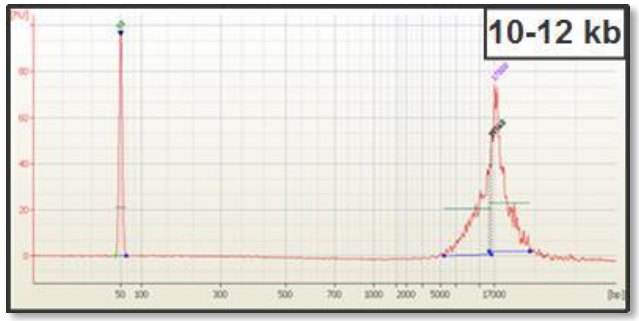
CHARACTERIZATION OF FRAGMENT SIZE IS IMPORTANT

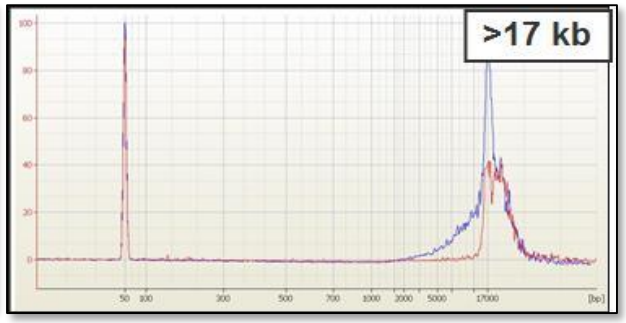






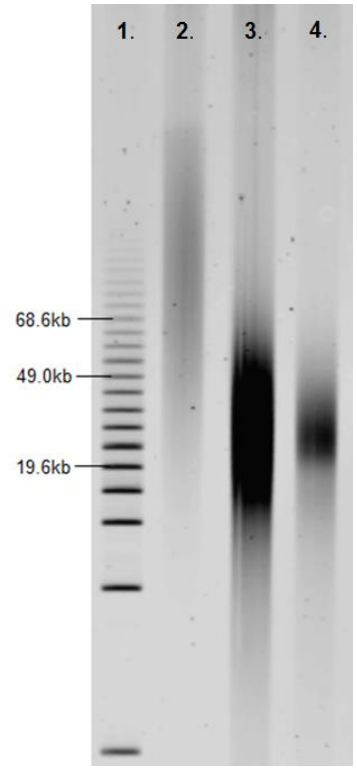






- Only ~30 to 40% of sheared DNA is in the desired size range
- Larger fragments do not have any advantage in loading, however, they affect library quantitation
- Small-size fraction within a shear has a higher loading advantage leading to reduced subread length
- Shorter insert sizes contain more individual molecules in a given quantity compared to larger inserts
- Accurate sizing of fragments larger than 17 kb is difficult to achieve on the BioAnalyzer instrument
- Recommend using pulse-field gel electrophoresis for more accurate sizing

FOR LARGE INSERT LIBRARIES, USE PFGE FOR ACCURATE SIZING OF HIGH MOLECULAR WEIGHT DNA



- Lane 1: Bio-Rad CHEF 5 kb ladder
- Lane 2: starting gDNA
- Lane 3: 20 kb sheared gDNA
- Lane 4: size-selected 20 kb SMRTbell template

- Pulse Field Gel Electrophoresis (CHEF Mapper or Bio-Rad system) or Pippin Pulse (Sage Science) are recommended for sizing >12 kb fragments

Pulsed-Field Gel Electrophoresis is the gold standard

BIO-RAD

CHEF Mapper

sage science

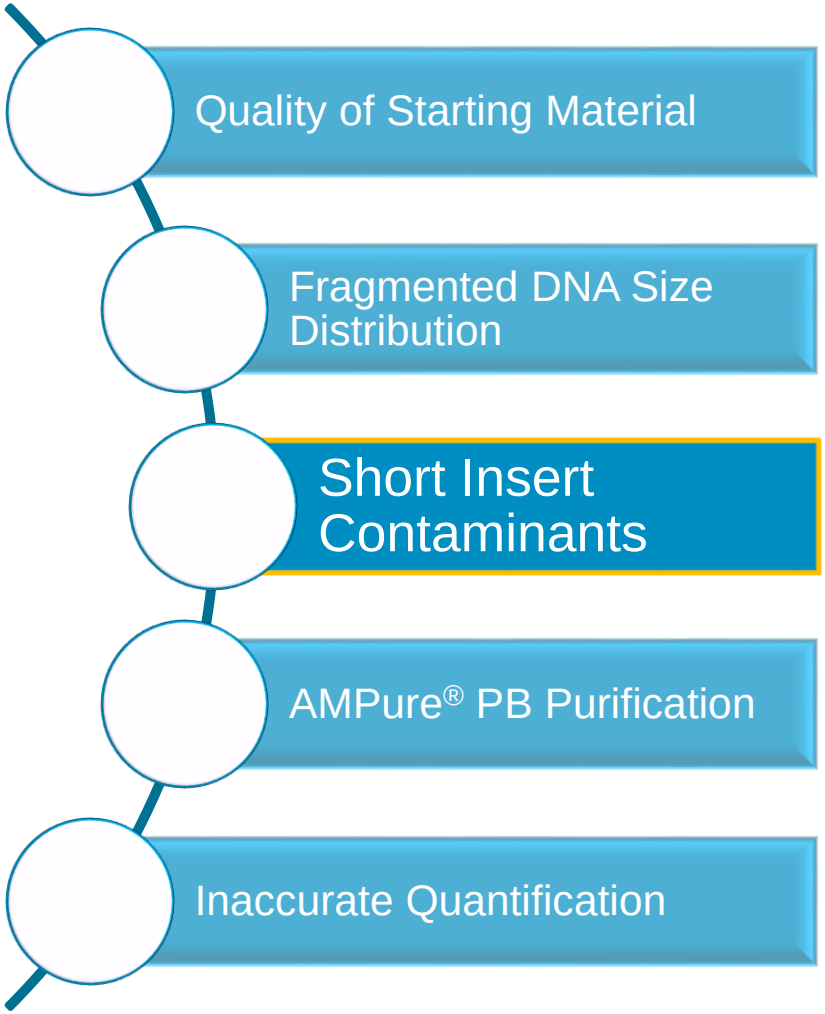
Pippin Pulse

Fragment Analyzer

ADVANCED ANALYTICAL

- Fragment Analyzer™ Automated CE System provides effective sizing of DNA fragments from 10 bp up to 50 kb

FACTORS AFFECTING SMRTBELL LIBRARY PREPARATION AND SEQUENCING PERFORMANCE



Causes

- Sub-optimal removal of
  - Adapter Dimers
  - Short Inserts

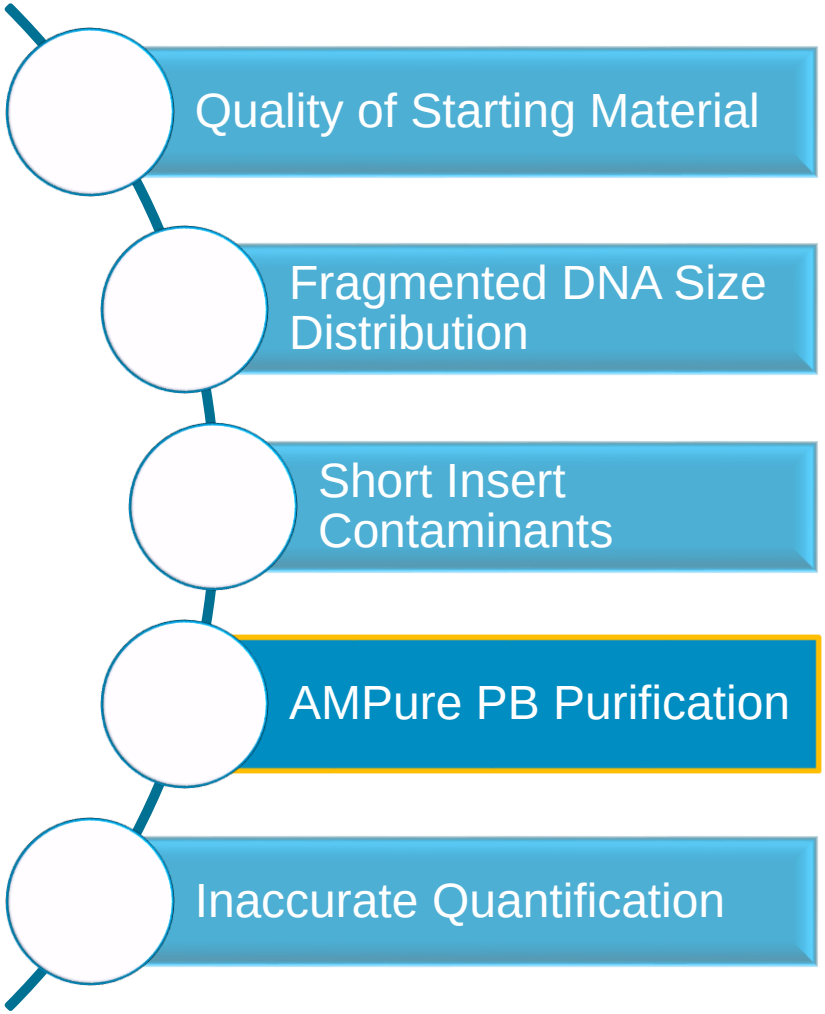
Impact on Performance

- Higher sequencing yields
- Higher polymerase read lengths
- Shorter subread lengths

Recommendations

- Use MagBeads for loading to remove short inserts and adapter dimers
- Thorough AMPure PB purification (3x ethanol washes)
- Ensure correct ratio of adapters to inserts
- If adapter dimer continues to be a problem, A/Tailing ligation maybe an option

FACTORS AFFECTING SMRTBELL LIBRARY PREPARATION AND SEQUENCING PERFORMANCE



Causes

- Carry-over contaminants from AMPure XP

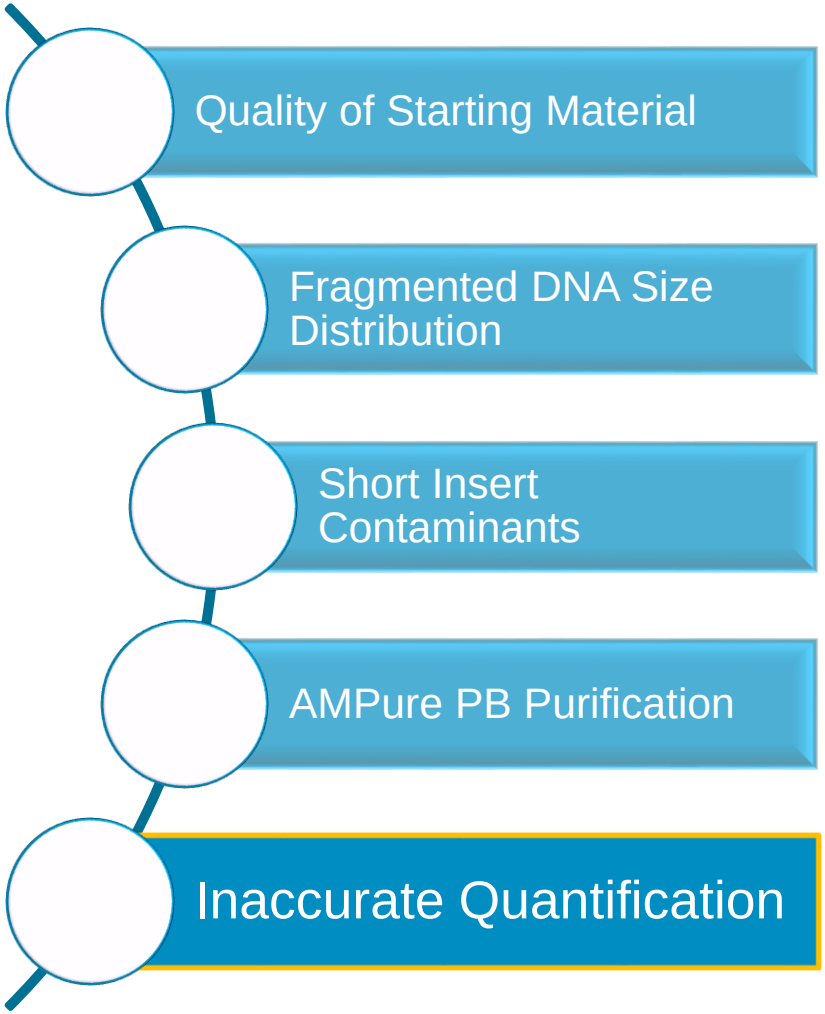
Impact on Performance

- Lower SMRTbell library yields
- Lower sequencing yields
- Shorter subread lengths
- Sample to sample variability

Recommendations

- Use AMPure PB
- Perform additional rounds of SMRTbell purification (2 – 3 times) using AMPure PB
- Use MagBeads for loading

FACTORS AFFECTING SMRTBELL LIBRARY PREPARATION AND SEQUENCING PERFORMANCE



Causes

- Variability in quantification tools
- Contaminants that affects readings

Impact on Performance

- Inaccurate binding reaction conditions resulting in over/under loading
- Impact on read length, accuracy and yield

Recommendations

- Highly recommend using Qubit systems
- Additional extraction methods to remove contaminants, RNA, short inserts, etc.

COMMON PITFALLS TO AVOID DURING AMPURE PB PURIFICATION

- Incorrect concentration of AMPure PB used in purification
  - Will retain undesired short inserts
  - Follow cut-off recommendations listed in the procedures
- Beads not thoroughly washed during purification
  - Will result in retention of short and adapter dimers
    - If short insert and adapter dimers are not washed thoroughly, they are preferentially loaded using diffusion loading
  - Wash beads thoroughly by adding 70% ethanol to the rim of the tube
  - Wash beads 3 times with 70% ethanol
- Over-drying of beads
  - Can result in low yield due to difficulties with bead suspension during sample elution
  - Do not let beads to dry more than 60 seconds (30-60 seconds recommended in the procedure)

ASSESSING SAMPLE QUALITY IS KEY TO SUCCESS

- What is the source of gDNA?
  - Understanding the source of gDNA is critical in upfront QC steps (plants, bacterial, tissues, blood, etc.)

- What methods were used in DNA isolation?
  - Carry-over contaminants can impact sequencing performance (CTAB, phenol/chloroform, others)

- What methods were used in DNA quantification?
  - Similar readings from Qubit and Nanodrop instruments provide higher confidence in the sample
  - Use of intercalating dyes is more accurate (Qubit instrument)

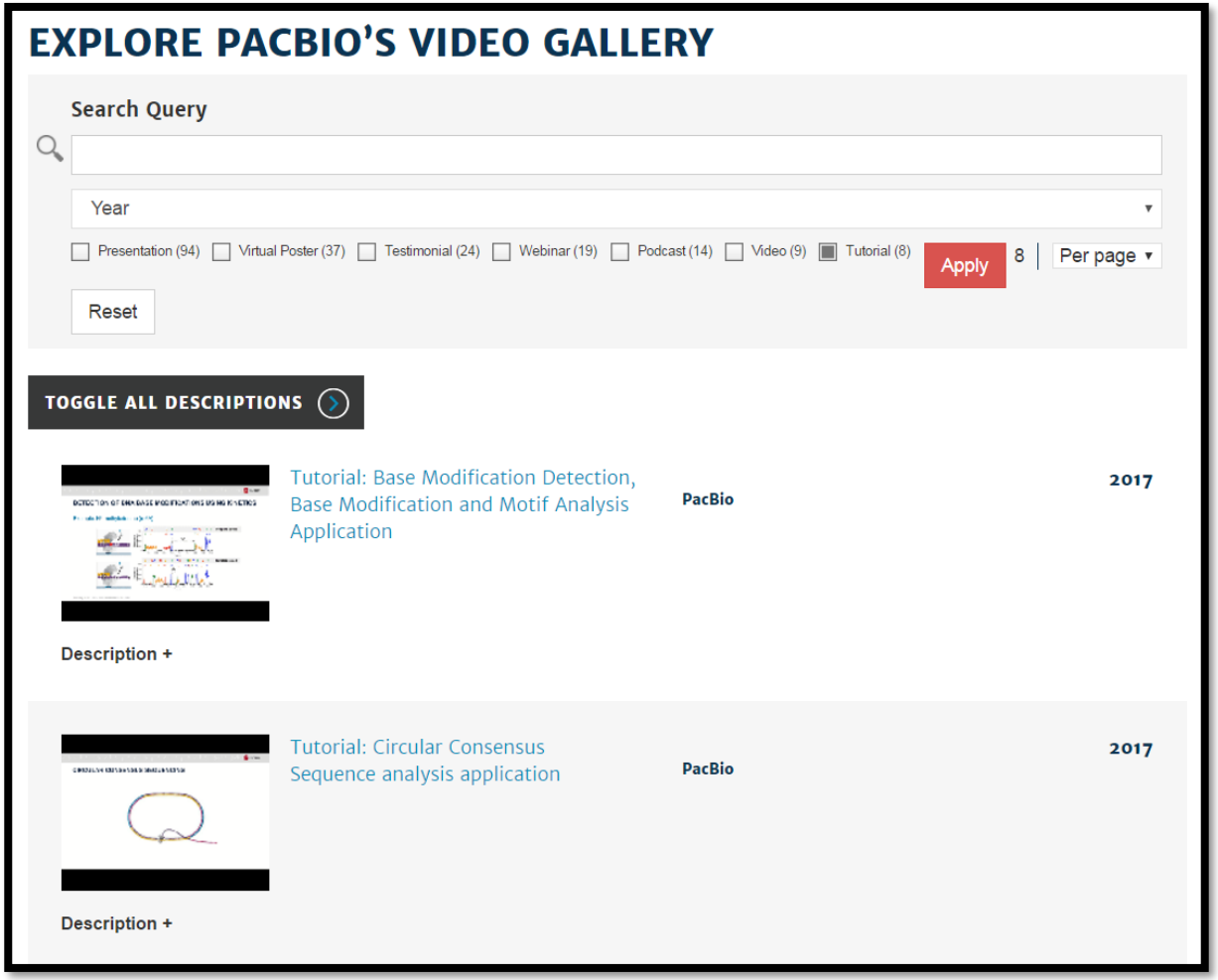
- Have you run gels to assess quality of the gDNA?
  - Running gels provides a clear picture of the integrity of the sample (degraded vs. RNA contamination)

EXAMPLES OF CASE STUDIES ON SAMPLE PREP AND SAMPLE QC

- Importance of Sample QC (PacBio Poster)
  - http://www.pacb.com/wp-content/uploads/2015/09/Importance-of-Sample-QC.pdf
- Sample Quality and Contamination (PacBio Poster)
  - http://www.pacb.com/wp-content/uploads/2015/09/Sample-Quality-and-Contamination.pdf

Technical Resources for PacBio Service Providers

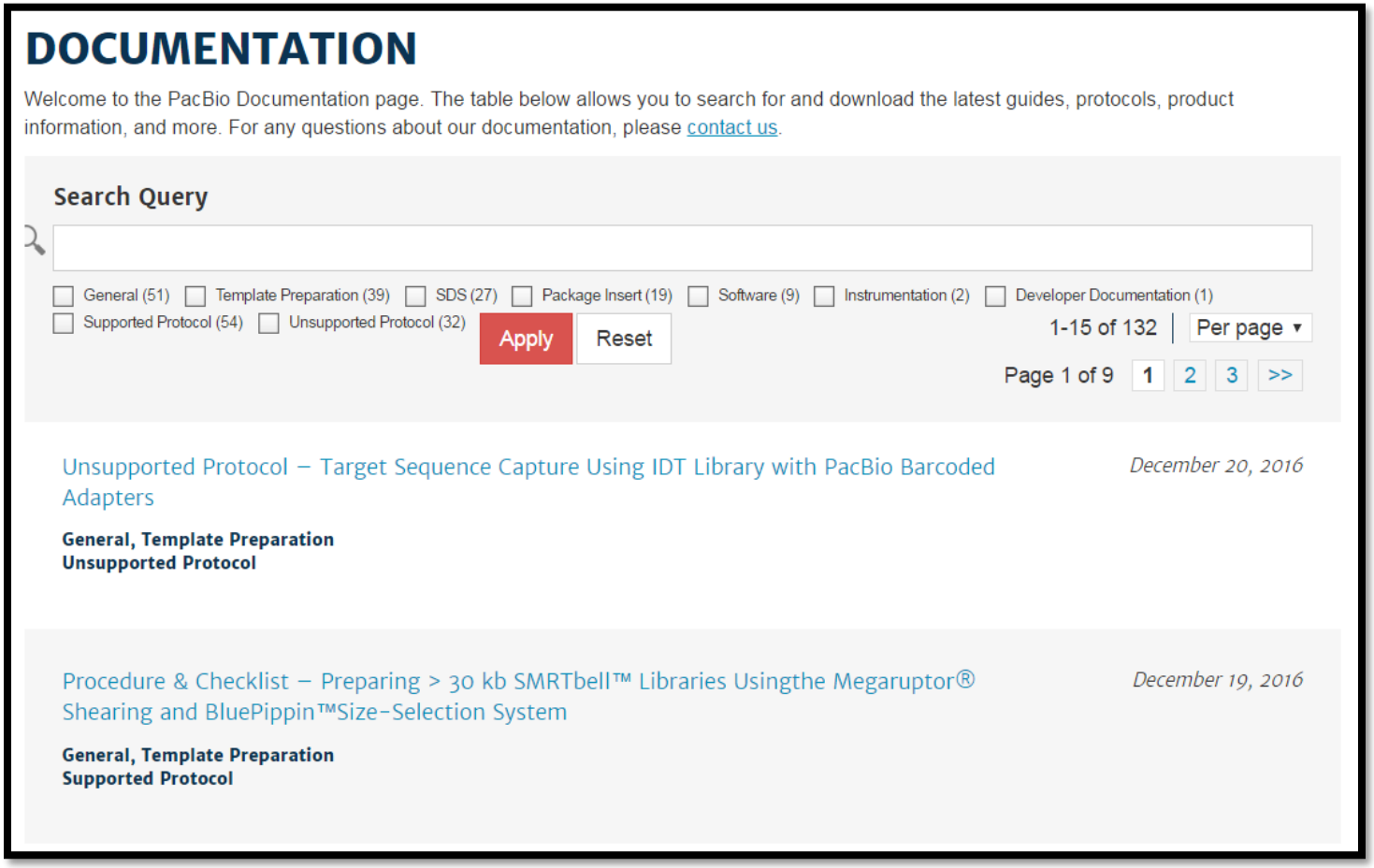
PACBIO VIDEO GALLERY



- PacBio Video Gallery page allows you to search for and download our collection of application-specific video tutorials, conference presentations, webinars and podcasts

http://www.pacb.com/smrt-science/smrt-resources/video-gallery/

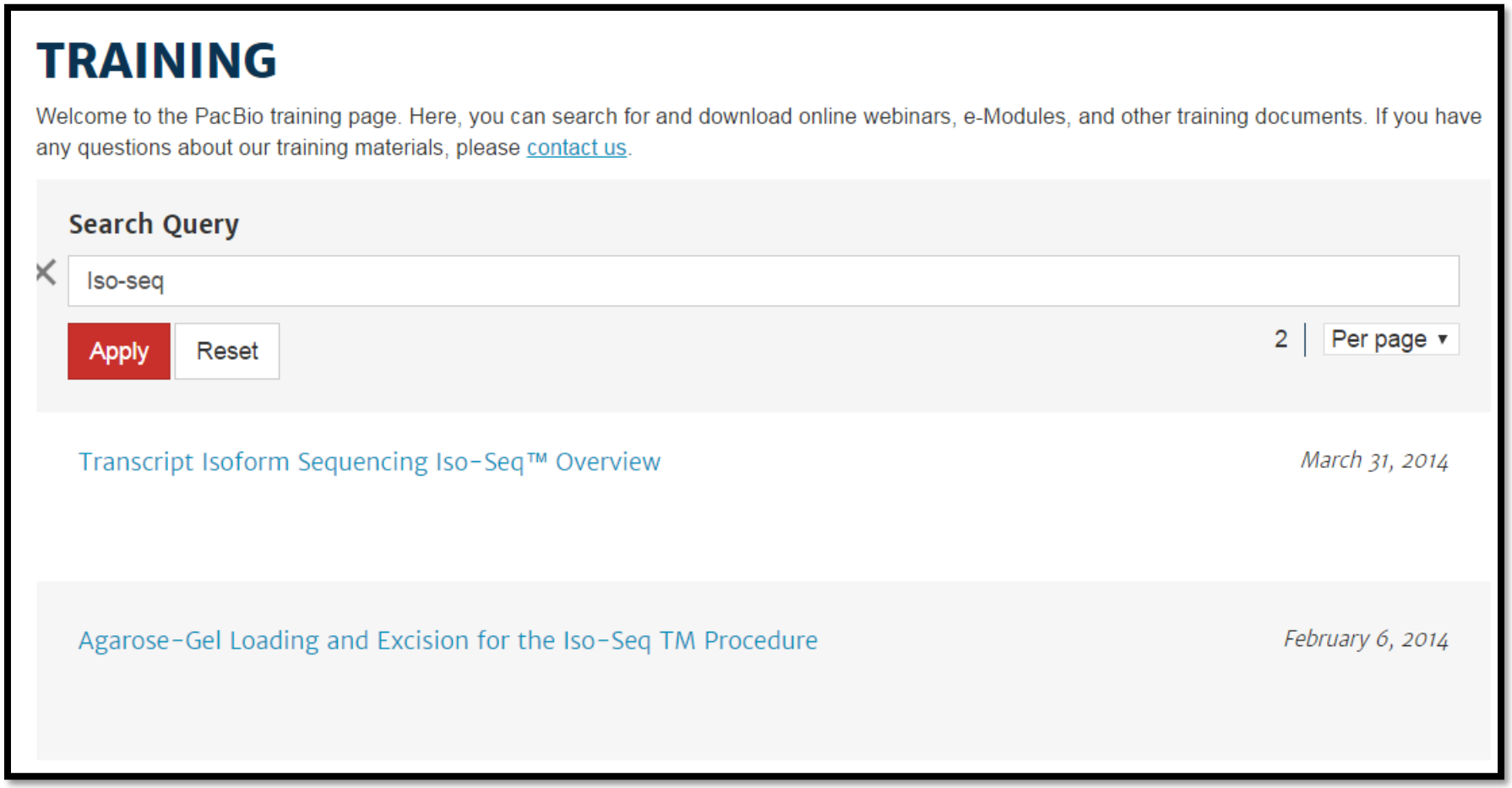
PACBIO DOCUMENTATION



- PacBio Documentation page allows you to search for and download the latest guides, protocols, product information, and more.

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

PACBIO TRAINING RESOURCES



- PacBio Training page allows you to search for and download online webinars, e-Modules, and other applications-specific training documents.

http://www.pacb.com/support/training/

PACBIO TRAINING RESOURCES

DNA/Polymerase Binding Kit P6 v2 (January 16, 2015)

http://www.pacb.com/wp-content/uploads/2015/01/2015.01_P6v2_Training_PolymeraseBindingKitv2Overview.pptx

Gain New Insights in Genome and Transcriptome Research with > 10,000 bp Reads (May 27, 2014)

http://aa314.gondor.co/webinar/gain-new-insights-in-genome-and-transcriptome-research-with-10000-bp-reads/

Post Run QC Analysis (May 9, 2014)

http://www.pacb.com/training/PostRunQCAnalysis/story.html

Template Preparation (April 11, 2014)

http://www.pacb.com/wp-content/uploads/2014/04/TemplatePreparation.pdf

Technology Overview (April 11, 2014)

http://www.pacb.com/training/TechnologyOverview.pdf

PacBio Workflow (April 11, 2014)

http://www.pacb.com/wp-content/uploads/2015/09/PacBioWorkflow.pdf

Transcript Isoform Sequencing Iso-Seq™ Overview (March 31, 2014)

http://www.pacb.com/training/IsoformSequencingIsoSeqOverview/story.html

Introduction to SMRTbell Template Preparation (March 3, 2014)

http://www.pacb.com/training/IntroductiontoSMRTbellTemplatePreparation/story.html

How to Load a Reference Sequence (March 3, 2014)

http://www.pacb.com/training/HowtoLoadaReferenceSequence/story.html

Agarose-Gel Loading and Excision for the Iso-Seq™ Procedure (February 6, 2014)

http://www.pacb.com/training/AgaroseGelLoadingExcisionforIsoSeqProcedure/story.html

How to Run the PacBio RS II Instrument (January 24, 2014)

http://www.pacb.com/training/HowtoRunPacBioInstrument/story.html

How to Run the PacBio RS II Instrument (January 24, 2014)

http://www.pacb.com/training/HowtoRunPacBioInstrument/story.html

Circular Contig Confirmation (January 13, 2014)

http://www.pacb.com/training/CircularContigConfirmationGepard/story.html

Introduction to SMRT Portal (September 27, 2013)

http://www.pacb.com/training/IntroductiontoSMRTPortal/story.html

RS Dashboard Overview (July 15, 2013)

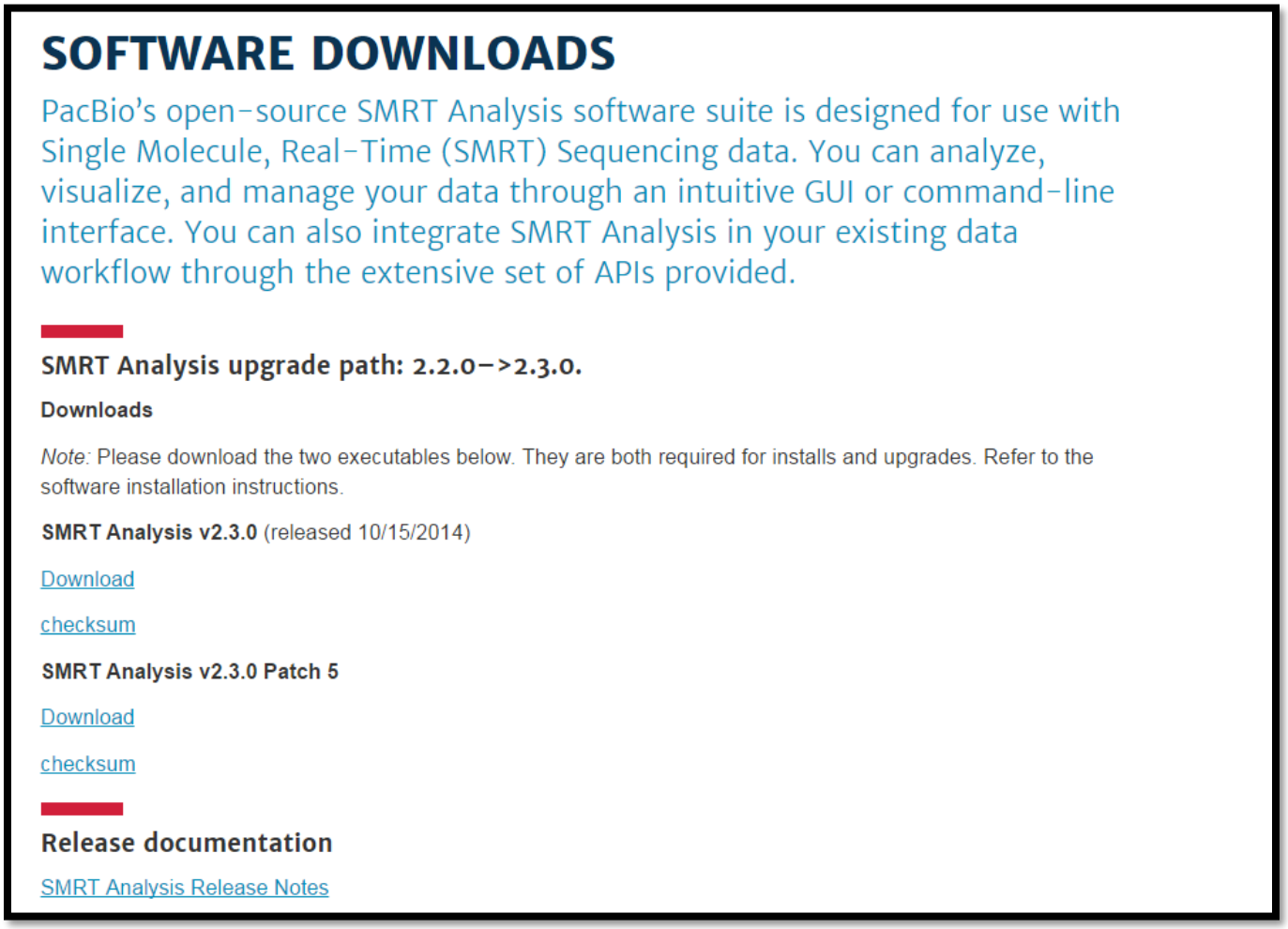
http://www.pacb.com/training/RSDashboardOverview/story.html

Bacterial Assembly and Epigenetic Analysis (July 12, 2013)

http://www.pacb.com/training/BacterialAssemblyandEpigeneticAnalysis/story.html

http://www.pacb.com/support/training/

PACBIO SOFTWARE DOWNLOADS



SOFTWARE DOWNLOADS

PacBio's open-source SMRT Analysis software suite is designed for use with Single Molecule, Real-Time (SMRT) Sequencing data. You can analyze, visualize, and manage your data through an intuitive GUI or command-line interface. You can also integrate SMRT Analysis in your existing data workflow through the extensive set of APIs provided.

SMRT Analysis upgrade path: 2.2.0->2.3.0.

Downloads

Note: Please download the two executables below. They are both required for installs and upgrades. Refer to the software installation instructions.

SMRT Analysis v2.3.0 (released 10/15/2014)

Download

checksum

SMRT Analysis v2.3.0 Patch 5

Download

checksum

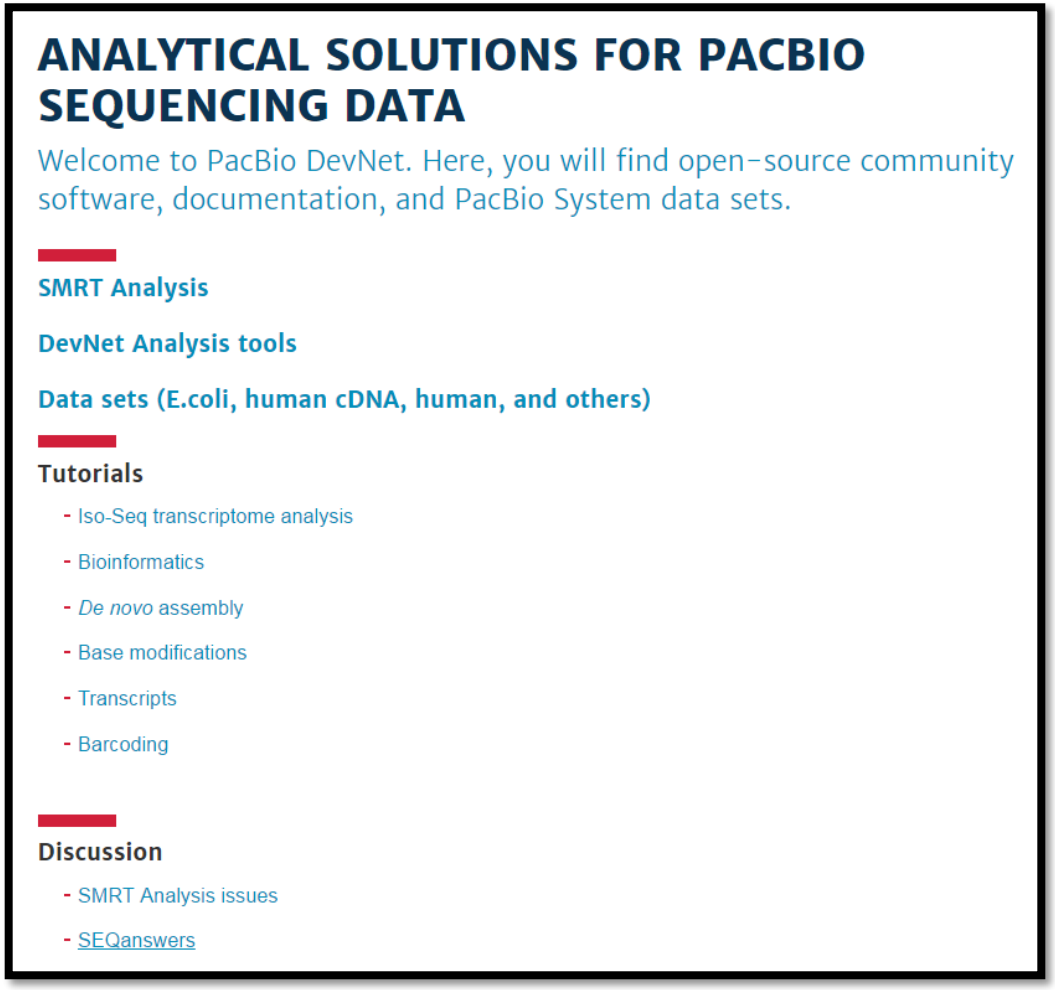
Release documentation

SMRT Analysis Release Notes

- Software Downloads page allows you to search for and download PacBio's open-source software tools and release documentation

http://www.pacb.com/support/software-downloads/

PACBIO DEVNET



ANALYTICAL SOLUTIONS FOR PACBIO SEQUENCING DATA

Welcome to PacBio DevNet. Here, you will find open-source community software, documentation, and PacBio System data sets.

SMRT Analysis

DevNet Analysis tools

Data sets (E.coli, human cDNA, human, and others)

Tutorials

- Iso-Seq transcriptome analysis
- Bioinformatics
- *De novo* assembly
- Base modifications
- Transcripts
- Barcoding

Discussion

- SMRT Analysis issues
- SEQanswers

- PacBio DevNet site contains open-source community software tools, software technical documentation, data analysis tutorials and PacBio System example data sets.

http://www.pacb.com/products-and-services/analytical-software/devnet/

SELF-HELP BIOINFORMATICS TRAINING

PacificBiosciences / Bioinformatics-Training

Watch 88 | Star 148 | Fork 60

Code | Issues 8 | Pull requests 0 | Projects 0 | Wiki | Pulse | Graphs

Home

Frank Ripp edited this page on Jul 17, 2014 · 47 revisions

Welcome to the PacBio Bioinformatics Training Wiki! (http://www.pacb.com/bfx-wiki)

This page includes resources for learning more about PacBio data and bioinformatics analysis, and includes content suitable for both beginners and experts. Below are links to training modules (webinars and PowerPoint presentations) to help you get started with your data processing, as well as information for specialized applications.

Training Resources:

- Bioinformatics Workshop (Webinars)
- Bioinformatics Training Slides

Specialized Applications:

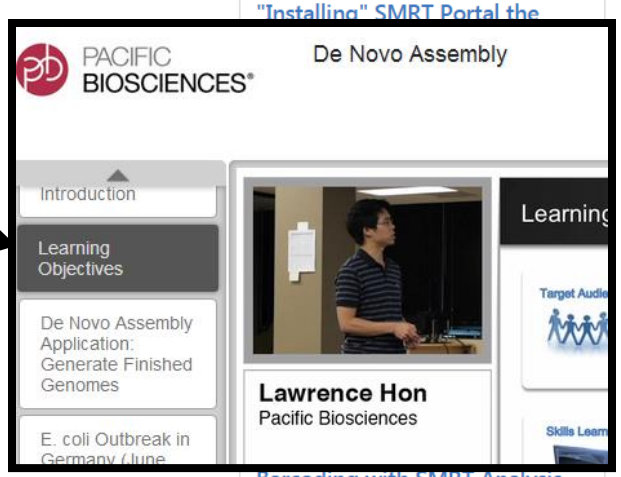
- De Novo Assembly
- Transcriptome analysis
- Base Modification Analysis
- Barcoding
- Data Analysis Tools
- Minor Variants and Phasing Analysis

Pages 52

Find a Page...

Home

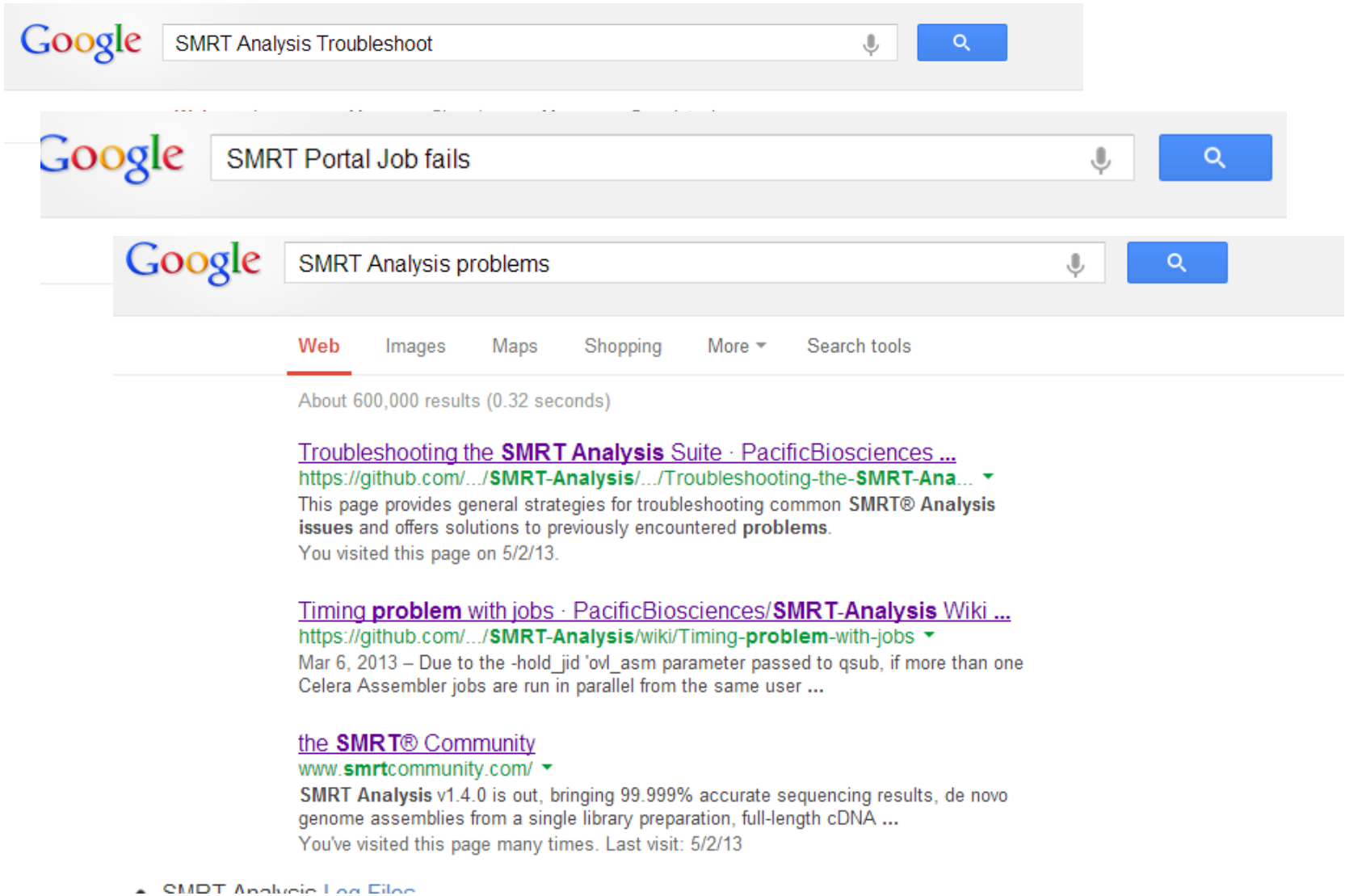
"Installing" SMRT Portal the



Barcoding with SMRT Analysis 2.0.1

https://github.com/PacificBiosciences/Bioinformatics-Training/wiki

TROUBLESHOOTING ARTICLES ARE SEARCHABLE ONLINE



HOW TO CONTACT PACBIO FOR TECHNICAL SUPPORT

- Tech Support Toll-free Telephone Number (Mon. – Fri., 9 a.m. to 5 p.m.)
  +1 (877) 920-PACB [+1 (877) 920-7222]

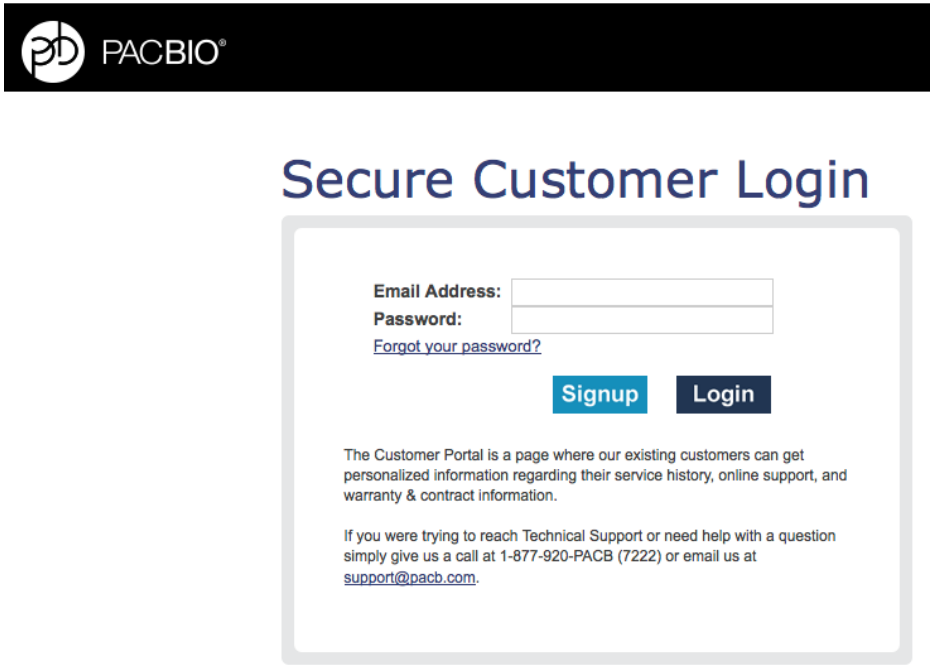
- Email: techsupport@pacificbiosciences.com

- Login to the PacBio® Customer Portal (http://www.pacbioportal.com/) and create a new case
  - Preferred route for most customers because your field team (FAS and FSE) is immediately notified via email of your issue
  - Facilitates the exchange of troubleshooting and diagnostic data and maintains a record of all communication between our support team and yourself

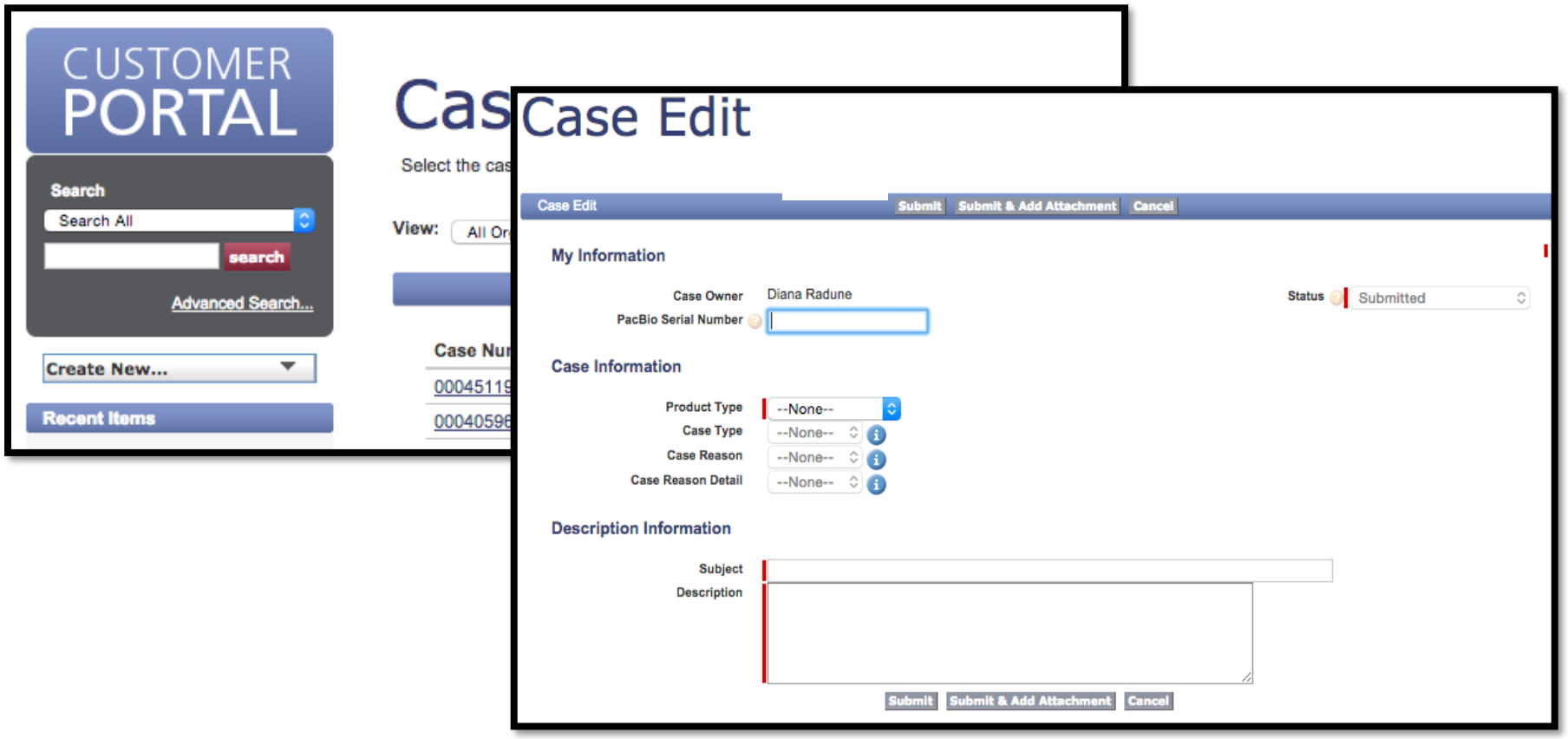
http://www.pacb.com/support/technical-support/

HOW TO CREATE A CASE IN THE CUSTOMER PORTAL

1. Go to www.pacbioportal.com (or click on the 'Customer Login' link on the top of any PacBio webpage)
2. Enter your Customer Portal login credentials.
   - For future Users needing accounts, simply click "**Signup**" and the request will be routed to your local PacBio Support Team for activation within 24 hours.
   - If you don't remember your password, click the "**Forgot your password**?" link.



CREATE AND SUBMIT A NEW CUSTOMER SUPPORT CASE



- From the **Cases** tab, you can view all the Cases submitted by your organization, or you can create a new Case.
- Upon submitting your new Case request, the PacBio Technical Support team and your local account team (FAS and FSE) are immediately notified of your issue via email.

CUSTOMER SUPPORT SUMMARY

- **Technical Support** (http://www.pacb.com/support/technical-support/)
  - Contact PacBio Technical Support for instrument and sequencing performance issues
  - Tech Support Toll-free Telephone Number: +1 (877) 920-PACB [+1 (877) 920-7222]
  - Tech Support Email: techsupport@pacificbiosciences.com
  - Login to the PacBio Customer Portal (http://www.pacbioportal.com/) and create a new Case

- **Customer Support Pages** (http://www.pacb.com/support/)
  - Documentation: http://www.pacb.com/support/documentation/
  - Training: http://www.pacb.com/support/training/
  - Software Downloads: http://www.pacb.com/support/software-downloads/

- **DevNet** (http://www.pacb.com/products-and-services/analytical-software/devnet/)
  - Latest PacBio data analysis software releases and third-party code
  - Bioinformatics Documentation
  - Bioinformatics Self-Help and Training Materials
  - Example PacBio Datasets

TECHNICAL RESOURCES FOR LARGE INSERT SMRTBELL LIBRARY SAMPLE PREPARATION AND SEQUENCING

User Bulletins and Field Advisories

User Bulletin for PacBio RS II and Sequel Systems: Centrifuge Tube and Pipet Tip Recommendations ***(NEW!)*** *(May 2017)*

- PacBio advises against the use of Axygen MAXYMum Recovery™ tubes and pipet tips. Please discontinue use of these products immediately. PacBio recommends alternatives in the User Bulletin.
- http://www.pacb.com/wp-content/uploads/User-Bulletin-Centrifuge-Tube-and-Pipet-Tip-Recommendations.pdf

Field Advisory for Sequel System: Securing Sequel Pipet Tip Rack ***(NEW!)*** *(May 2017)*

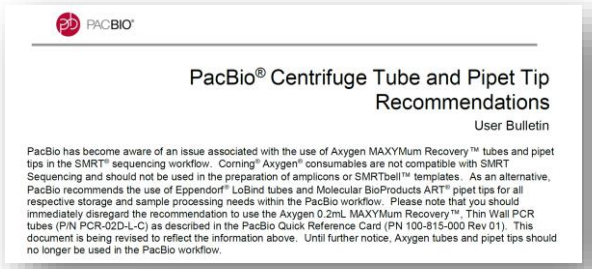
- PacBio recommends a simple procedure to ensure that the Sequel Pipet Tip rack is firmly affixed to the tip box.
- http://www.pacb.com/wp-content/uploads/Field-Advisory-Notice-Securing-Sequel-Pipet-Tip-Rack.pdf

User Bulletin for Sequel System: Heat Seal Advisory (Adhesive Seal Warning) ***(NEW!)*** *(May 2017)*

- PacBio advises against the use of adhesive foils and recommends the use of Sequel Sample Plate Foil.
- http://www.pacb.com/wp-content/uploads/User-Bulletin-Heat-Seal-Advisory-Adhesive-Seal-Warning.pdf

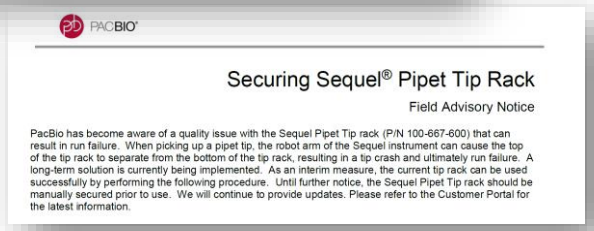
User Bulletin for Sequel System: Barcode Scanning of Sequel Sequencing Kit 2.0 ***(NEW!)*** *(May 2017)*

- PacBio is providing clarity on which barcode to scan to ensure the Sequel System has the correct information and that all the consumables are compatible.
- http://www.pacb.com/wp-content/uploads/User-Bulletin-Barcode-Scanning-of-Sequel-Sequencing-Kit-2.0.pdf


PacBio® Centrifuge Tube and Pipet Tip Recommendations

User Bulletin

PacBio has become aware of an issue associated with the use of Axygen MAXYMum Recovery™ tubes and pipet tips in the SMRT® sequencing workflow. Corning® Axygen® consumables are not compatible with SMRT Sequencing and should not be used in the preparation of amplicons or SMRTbell™ templates. As an alternative, PacBio recommends the use of Eppendorf® LoBind tubes and Molecular BioProducts ART® pipet tips for all respective storage and sample processing needs within the PacBio workflow. Please note that you should immediately disregard the recommendation to use the Axygen 0.2mL MAXYMum Recovery™, Thin Wall PCR tubes (P/N PCR-02D-L-C) as described in the PacBio Quick Reference Card (PN 100-815-000 Rev 01). This document is being revised to reflect the information above. Until further notice, Axygen tubes and pipet tips should no longer be used in the PacBio workflow.


Securing Sequel® Pipet Tip Rack

Field Advisory Notice

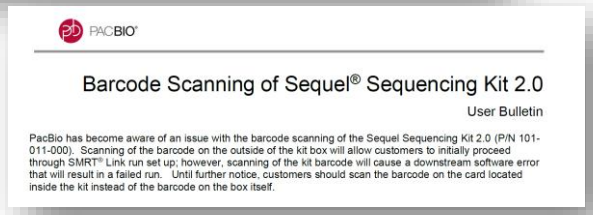
PacBio has become aware of a quality issue with the Sequel Pipet Tip rack (P/N 100-667-600) that can result in run failure. When picking up a pipet tip, the robot arm of the Sequel instrument can cause the top of the tip rack to separate from the bottom of the tip rack, resulting in a tip crash and ultimately run failure. A long-term solution is currently being implemented. As an interim measure, the current tip rack can be used successfully by performing the following procedure. Until further notice, the Sequel Pipet Tip rack should be manually secured prior to use. We will continue to provide updates. Please refer to the Customer Portal for the latest information.


Heat Seal Advisory (Adhesive Seal Warning)

User Bulletin

PacBio has recently become aware of an issue associated with the use of adhesive foil seals. The use of adhesive foil seals can lead to instrument damage. Until further notice, PacBio advises against the use of adhesive foil seals and recommends the of use of the Sequel® Sample Plate Foil (P/N 100-667-400). The of use of the Sequel Sample Plate Foil requires the purchase of a heat sealer. PacBio recommends the Thermo Scientific™ ALPS™ 50 V Semi automated Microplate Heat Sealer (P/N AB-1443).

The use of adhesive foil seals will void existing instrument warranties. Customers will be responsible for any damage resulting from the use of products discordant with official PacBio recommendations.


Barcode Scanning of Sequel® Sequencing Kit 2.0

User Bulletin

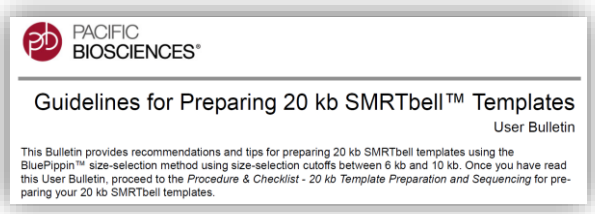
PacBio has become aware of an issue with the barcode scanning of the Sequel Sequencing Kit 2.0 (P/N 101-011-000). Scanning of the barcode on the outside of the kit box will allow customers to initially proceed through SMRT® Link run set up; however, scanning of the kit barcode will cause a downstream software error that will result in a failed run. Until further notice, customers should scan the barcode on the card located inside the kit instead of the barcode on the box itself.

Find all protocols at http://www.pacb.com/support/documentation/

User Bulletins (cont.)

User Bulletin – Guidelines for Preparing 20 kb SMRTbell™ Templates

- http://www.pacb.com/wp-content/uploads/2015/09/User-Bulletin-Guidelines-for-Preparing-20-kb-SMRTbell-Templates.pdf



Webinars

Webinar - DNA Quality Requirements for Single Molecule, Real-Time (SMRT®) Sequencing (September 30, 2015)

- Recorded Webinar: http://programs.pacificbiosciences.com/e/1652/Uppsala-Recording-Sept2015-mp4/3gr93b/490791521
- PowerPoint Presentation: http://programs.pacificbiosciences.com/e/1652/ala-Presentation-Sept2015-pptx/3gr93d/490791521
- Q&A: http://programs.pacificbiosciences.com/e/1652/PacBio-Uppsala-QA-Sept2015-pdf/3gr93g/490791521

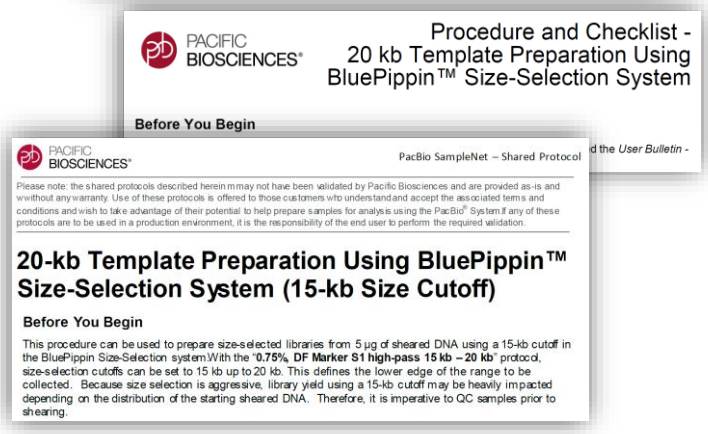
20 kb SMRTbell Library Preparation Protocols

Procedure & Checklist – 20 kb Template Preparation Using BluePippin Size-Selection

- http://www.pacb.com/wp-content/uploads/2015/09/Procedure-Checklist-20-kb-Template-Preparation-Using-BluePippin-Size-Selection.pdf

Unsupported Protocol – 20 kb Template Preparation with 15kb Cutoff Using BluePippin™ Size-Selection

- http://www.pacb.com/wp-content/uploads/2015/09/Shared-Protocol-20-kb-Template-Preparation-Using-BluePippin-Size-Selection-System-15-kb-Size-Cutoff.pdf



Find all protocols at http://www.pacb.com/support/documentation/

>20 kb SMRTbell Library Preparation and Column Purification Protocols (Sequel)

*Procedure & Checklist >20 kb Template Preparation Using BluePippin Size-Selection System (15 – 20 kb Cutoff) for Sequel Systems **(NEW!)** (Jan. 2017)*

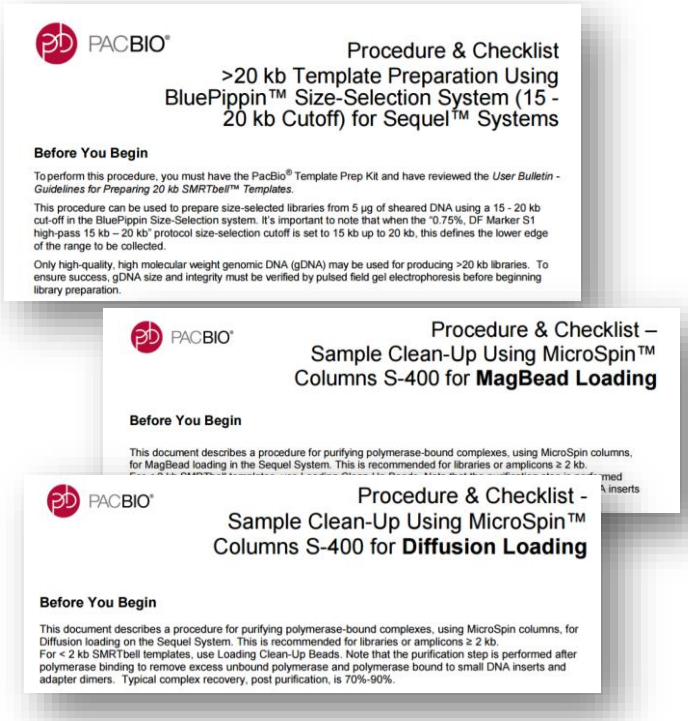
- http://www.pacb.com/wp-content/uploads/Procedure-Checklist-20-kb-Template-Preparation-Using-BluePippin-Size-Selection-System-Sequel-Systems.pdf

*Procedure & Checklist – Sample Clean-Up Using MicroSpin™ Columns S-400 for MagBead Loading **(NEW!)** (June 2017)*

- http://www.pacb.com/wp-content/uploads/Procedure-Checklist-Sample-Clean-Up-Using-MicroSpin%E2%84%A2-Columns-S-400-for-MagBead-Loading.pdf

*Procedure & Checklist – Sample Clean-Up Using MicroSpin™ Columns S-400 for Diffusion Loading **(NEW!)** (June 2017)*

- http://www.pacb.com/wp-content/uploads/Procedure-Checklist-Sample-Clean-Up-Using-MicroSpin%E2%84%A2-Columns-S-400-for-Diffusion-Loading-1.pdf



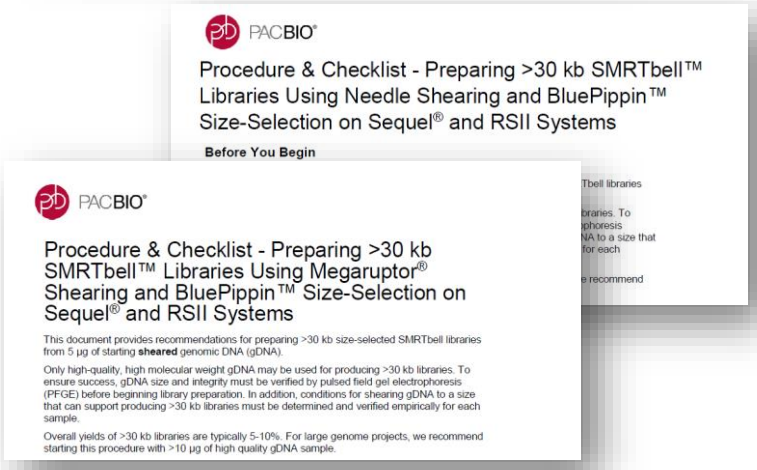
>30 kb SMRTbell Library Preparation Protocols

*Procedure & Checklist - Preparing >30 kb SMRTbell™ Libraries Using Needle Shearing and BluePippin™ Size-Selection on Sequel® and RSII Systems (**NEW!**) (June 2017)*

- TBD

*Procedure & Checklist - Preparing >30 kb SMRTbell™ Libraries Using Megaruptor® Shearing and BluePippin™ Size-Selection on Sequel® and RSII Systems (**NEW!**) (June 2017)*

- TBD



Find all protocols at http://www.pacb.com/support/documentation/

Multiplexed SMRTbell Library Preparation and Data Analysis Protocols for *De Novo* Assembly

Procedure & Checklist – Preparing SMRTbell Libraries using PacBio Barcoded Adapters for Multiplex SMRT Sequencing ***(NEW!)*** *(May 2017)*

- http://www.pacb.com/wp-content/uploads/Procedure-Checklist-Preparing-SMRTbell-Libraries-PacBio-Barcoded-Adapters-Multiplex-SMRT-Sequencing.pdf

Analysis Procedure – Multiplexed Microbial Assembly with SMRT Link ***(NEW!)*** *(Feb. 2017)*

- http://www.pacb.com/wp-content/uploads/Analysis-Procedure-Multiplexed-Microbial-Assembly-SMRT-Link.pdf

DNA Fragment Size Characterization Protocols and App Notes

Unsupported Protocol – Guidelines for Using the Sage Science™ Pippin Pulse

- http://www.pacb.com/wp-content/uploads/Unsupported-Guidelines-Using-Sage-Science-Pippin-Pulse.pdf

Unsupported Protocol – Guidelines for Using the BIO-RAD® CHEFMapper® XA Pulsed Field Electrophoresis

- http://www.pacb.com/wp-content/uploads/Unsupported-Guidelines-Using-BIO-RAD-CHEFMapper-XA-Pulsed-Field-Electrophoresis.pdf

Application Note – Fast High Resolution DNA Sizing with the Fragment Analyzer System

- http://www.pacb.com/wp-content/uploads/fast-high-resolution-dna-sizing-fragment-analyzer-system.pdf

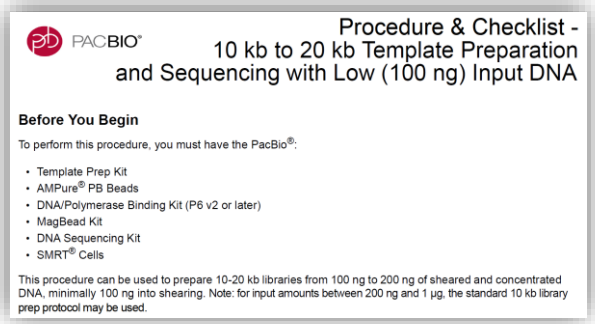
Low Input DNA Template Preparation Protocols

Procedure & Checklist – 10 kb to 20 kb Template Preparation and Sequencing with Low (100 ng) Input DNA

- http://www.pacb.com/wp-content/uploads/Procedure-Checklist-10-20kb-Template-Preparation-and-Sequencing-with-Low-Input-DNA.pdf

[For shorter insert sizes: *Procedure & Checklist – Very Low (10 ng) Input 2 kb Template Preparation and Sequencing with Carrier DNA*

- http://www.pacb.com/wp-content/uploads/Procedure-Checklist-Very-Low-ng-Input-2kb-Template-Preparation-and-Sequencing-with-Carrier-DNA.pdf]


PACBIO®

Procedure & Checklist - 10 kb to 20 kb Template Preparation and Sequencing with Low (100 ng) Input DNA

Before You Begin

To perform this procedure, you must have the PacBio®:

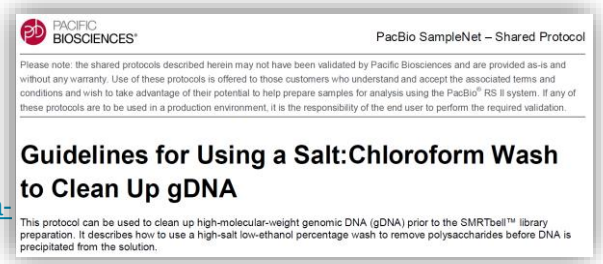
- Template Prep Kit
- AMPure® PB Beads
- DNA/Polymerase Binding Kit (P6 v2 or later)
- MagBead Kit
- DNA Sequencing Kit
- SMRT® Cells

This procedure can be used to prepare 10-20 kb libraries from 100 ng to 200 ng of sheared and concentrated DNA, minimally 100 ng into shearing. Note: for input amounts between 200 ng and 1 µg, the standard 10 kb library prep protocol may be used.

High Molecular Weight gDNA Cleanup

Unsupported Protocol – High Salt Phenol Chloroform Cleanup

- http://www.pacb.com/wp-content/uploads/2015/09/Shared-Protocol-Guidelines-for-Using-a-Salt-Chloroform-Wash-to-Clean-Up-gDNA.pdf


PACIFIC BIOSCIENCES®

PacBio SampleNet – Shared Protocol

Please note: the shared protocols described herein may not have been validated by Pacific Biosciences and are provided as-is and without any warranty. Use of these protocols is offered to those customers who understand and accept the associated terms and conditions and wish to take advantage of their potential to help prepare samples for analysis using the PacBio® RS II system. If any of these protocols are to be used in a production environment, it is the responsibility of the end user to perform the required validation.

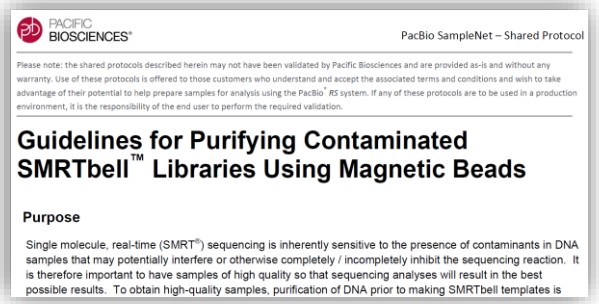
Guidelines for Using a Salt:Chloroform Wash to Clean Up gDNA

This protocol can be used to clean up high-molecular-weight genomic DNA (gDNA) prior to the SMRTbell™ library preparation. It describes how to use a high-salt low-ethanol percentage wash to remove polysaccharides before DNA is precipitated from the solution.

SMRTbell Library Cleanup

Unsupported Protocol – Purification of Contaminated SMRTbell™ Library Using Magnetic Bead Capture

- http://www.pacb.com/wp-content/uploads/2015/09/Purifying-Contaminated-SMRTbell-Libraries-Using-MagBeads-052013.pdf


PACIFIC BIOSCIENCES®

PacBio SampleNet – Shared Protocol

Please note: the shared protocols described herein may not have been validated by Pacific Biosciences and are provided as-is and without any warranty. Use of these protocols is offered to those customers who understand and accept the associated terms and conditions and wish to take advantage of their potential to help prepare samples for analysis using the PacBio® RS system. If any of these protocols are to be used in a production environment, it is the responsibility of the end user to perform the required validation.

Guidelines for Purifying Contaminated SMRTbell™ Libraries Using Magnetic Beads

Purpose

Single molecule, real-time (SMRT®) sequencing is inherently sensitive to the presence of contaminants in DNA samples that may potentially interfere or otherwise completely / incompletely inhibit the sequencing reaction. It is therefore important to have samples of high quality so that sequencing analyses will result in the best possible results. To obtain high-quality samples, purification of DNA prior to making SMRTbell templates is

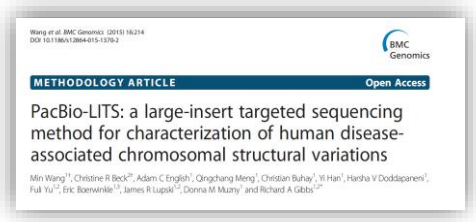
WHERE TO FIND OTHER SMRT RESOURCES

http://www.pacb.com/smrt-science/smrt-resources/

Explore our collection of resources and learn how scientists use SMRT Sequencing to advance their research.

Scientific publications

Explore our database of scientific publications featuring PacBio data.



http://www.biomedcentral.com/content/pdf/s12864-015-1370-2.pdf

Conference proceedings

Access conference posters and presentations our customers, collaborators, and internal scientists have presented at various scientific meetings.

PacBio literature

View case studies, brochures, application notes, and more.

Video gallery

Watch our collection of videos, webinars, customer testimonials, and more.

Blog

Read our blog featuring new research, publications, conference summaries, and SMRT Sequencing updates.

Product documentation and training

Visit user documentation for our entire documentation library and training for user training materials.

Find all protocols at http://www.pacb.com/support/documentation/

PacBio Certified Service Provider Program

PACBIO CERTIFIED SERVICE PROVIDER PROGRAM

The PacBio Certified Service Provider Program is a global network of validated service companies that provides access to Single Molecule, Real-Time (SMRT) Sequencing

- To achieve this certification, each provider must undergo a rigorous process that includes the most up-to-date training on PacBio sample processing and data analysis workflows.
- From these processes, our service providers are equipped to deliver the highest quality services for our advanced sequencing technology.
- Along with current training on all PacBio sequencing technologies, this certification includes:
  - Website presence
  - PacBio CSP designation on the PacBio Service Provider Directory page
  - Inclusion on a dedicated PacBio CSP-specific website that features core lab profiles, media coverage, and related PacBio publications
- Certified service providers are also included in co-marketing initiatives that may include:
  - Press releases
  - Custom flyers
  - Feature articles in PacBio's Core Lab Profile newsletter series
  - Co-promotional packages at trade shows
  - Conference co-sponsorships
  - Email blasts
  - E-newsletter presence

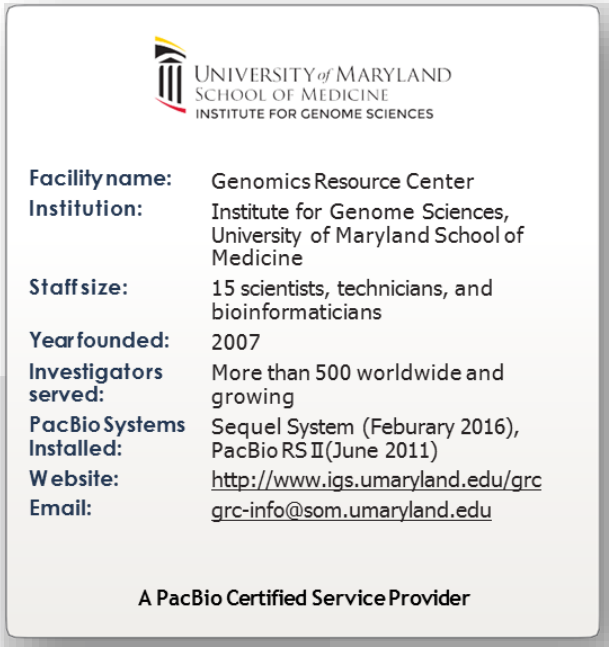
North American Certified Service Providers

PacBio RS II

+ Sequel™ System

Become a Certified PacBio Service Provider:

- If you are interested in becoming a partner with the industry leader in long-read sequencing, contact our PacBio Certified Service Provider Program at 1.650.521.8140 or portalsupport@pacb.com to learn more.



Arizona Genomics Institute
Tucson, AZ
http://www.genome.arizona.edu/
email: dkudrna@email.arizona.edu
ph: +1.520.626.9596

City of Hope
Duarte, CA
http://www.cityofhope.org/research/shared-resources/integrative-genomics-core
email: Jtermini@coh.org
ph: +1 626-301-8169

Icahn Institute for Genomics and Multiscale Biology at Mount Sinai
New York, NY
http://icahn.mssm.edu/genomics
email: robert.sebra@mssm.edu

McGill University and Génome Québec Innovation Centre
Montreal, Canada
http://gqinnovationcenter.com/index.aspx
email: infoservices@genomequebec.com
ph: +1.514.398.7211

Genomics Resource Center Institute for Genome Sciences, University of Maryland
Baltimore, MD
http://www.igs.umaryland.edu/resources/grc/index.php
email: grc-info@som.umaryland.edu
ph: +1.410.706.5668

Research and Testing Laboratory, LLC
Lubbock, TX
http://www.rtlgenomics.com
email: info@researchandtesting.com
ph: +1 806-771-1134

Washington State University (WSU) - Laboratory for Biotechnology and Bioanalysis - DNA Sequencing Core
Pullman, WA
http://lbb.wsu.edu/
email: dnaguy@mail.wsu.edu
ph: +1.509.335.1174

National Center for Genome Resources (NCGR) - Sequencing Center
Sante Fe, NM
http://www.ncgr.org
email: seq@ncgr.org
ph: +1.505.995.4449

UNIVERSITY OF MARYLAND CSP CORE LAB PROFILE (2016)

INSTITUTE FOR GENOME SCIENCES EXPANDS LONG-READ SEQUENCING SERVICES WITH NEW SEQUEL™ SYSTEM

At the University of Maryland's Genomics Resource Center, SMRT® Sequencing has become an integral tool for generating complete microbial genomes, improving plant and animal genome assemblies, and exploring human genome variation.

Strength of Experience

The Genomics Resource Center (GRC) at the Institute for Genome Sciences (IGS) has a scientific pedigree and a sample-to-interpretation service commitment that place it in a league of its own. The team operates under a simple mantra: 'If it can be sequenced, we can do it.'

Both the GRC and IGS were founded in 2007 when a high-powered team of investigators formerly at The Institute for Genomic Research (TIGR), led by Claire Fraser, joined the University of Maryland School of Medicine. "The team of faculty and staff that came here to start the institute was heavily focused on infectious disease research," says Luke Tallon, scientific director and founding leader of the GRC. "Our primary goal in joining a medical school was to extend our pathogen genomics expertise into host-pathogen studies and direct clinical genomics applications."

In addition to its infectious disease and genomics expertise, TIGR was also renowned for its bioinformatics talent – a trait that continues with the group at the GRC. Their team of 15 staff members is evenly split between wet lab and bioinformatics, and more than half of the institute's 100-plus employees are bioinformaticians.

"One of our strengths is that we go beyond generating efficient, high-quality sequence data. We have teams of bioinformatics analysts and software engineers who can assist investigators with downstream analysis and interpretation," Tallon says. Prior to project initiation, the GRC team consults with each investigator, recommending a custom solution for each scientist's particular goals and needs.

The GRC was formed both to serve the genomics institute and as a university core. "We serve investigators throughout the University of Maryland system as well as across the country and around the world," says Lisa Sadzewicz, administrative director of the facility. The GRC works with hundreds of investigators and has become more visible by presenting and exhibiting at conferences such as the annual meetings of the American Society for Microbiology and the American Society of Human Genetics, and engaging the community through social media, including the GRC blog.

"Our strength is not just our deep history and experience in sequencing and genomics, but our end-to-end service level from the initial project consultation through to publication, including all of the informatics," Sadzewicz says.

Over the past five years, the GRC has applied these strengths to the PacBio® platform. As early adopters of the technology, they have dedicated significant resources to the development of both laboratory and data analysis processes to leverage SMRT Sequencing. Since its adoption of the original PacBio *RS* in 2011, the GRC has steadily increased its utilization of the platform. In the past year alone, they have constructed more than 400 libraries and sequenced more than 1,200 SMRT Cells. These have spanned projects ranging from whole genome sequencing to metagenomics, Iso-Seq™ transcriptomes, and custom amplicons.

Microbial Genomics Expertise

Because of the institute's strong focus on infectious disease, the GRC conducts many sequencing projects for pathogen genomes, human microbiome samples, and other microbial genome applications. The PacBio platform has been a workhorse for generating finished or nearly complete genomes for the past five years. The inclusion of genome-wide methylation data provided in each run is an added bonus. "We're now routinely analyzing methylation patterns for our genomes," Tallon says.

Among the largest of the GRC's current microbial projects is the development of the FDA ARGOS database hosted at NCBI. Initiated in 2013, the database is supported by two contracts totaling over $4M to fund the sequencing, assembly, and analysis of more than 2,000 viral, bacterial, fungal, and parasite genomes. Designed to promote a less burdensome regulatory approach for devices that incorporate infectious disease NGS diagnostic technology, the ARGOS database supplies a set of validated regulatory-grade microbial genomic sequences and associated metadata.

"We partnered with the GRC because of their long-standing experience and excellent end-to-end service to generate high-quality microbial genomes," says Heike Sichtig, principal investigator at FDA. "The pipeline was customized to fit our needs to produce and make publically available FDA-vetted microbial regulatory-grade genomes, including raw reads, assemblies, annotations, and methylation data. The PacBio System has been extremely helpful in this project to generate these near-complete regulatory-grade microbial genomes."

Exploring Larger Genomes and Human Variation

As read lengths and throughput have steadily increased, the PacBio System has expanded its application base. The GRC team is forging ahead with larger genomes, transcriptomes, and targeted sequencing. "We've sequenced a number of agriculturally and medically relevant plant and animal genomes, and we are rapidly expanding our sequencing of fungal and parasite genomes," Sadzewicz comments.

> "We partnered with the GRC because of their long-standing experience and excellent end-to end service …" Heike Sichtig, principal investigator at FDA.

Among these are the GRC's contributions to the cichlid fish genome sequencing projects led by Tom Kocher at the University of Maryland. Following up on a short-read-based assembly of *Metriaclima zebra* generated as part of the Assemblathon2 contest, Tom and his team used 16-fold coverage of SMRT Sequencing reads to close 68% of the assembly gaps and add more than 90 Mbp of sequence to the genome. Additional analysis and *de novo* assembly of 50-fold coverage data sets from this and other cichlid genomes is underway. "The high-quality PacBio sequencing performed by the IGS core has been critical to our efforts to improve the reference sequences for cichlid fish, including tilapia and species from Lake Malawi, Africa," notes Dr. Kocher.

They are also using PacBio long reads for haplotype phasing, structural variation detection, and variant validations in human genomes. While sequencing large-scale whole human genomes using PacBio remains expensive in the near term, it is an ideal platform for these targeted applications. "We have developed custom amplicon approaches to detect indels and other structural variations," says Tallon. "We are also using a barcoded, pooled amplicon approach on the PacBio RS II to validate SNVs from our cancer exome projects."

Early Adoption of the Sequel™ System

The GRC's newest PacBio sequencer arrived in February 2016, making them one of the first PacBio Certified Service Providers to take delivery of a Sequel System. "Given our history of early adoption and success with the PacBio *RS*, and the promise of increased and scalable throughput, we were excited to be among the first centers to acquire a Sequel instrument," says Dr. Fraser. The Sequel System represents the newest generation of SMRT Sequencing, providing 5-10 Gb of throughput per SMRT Cell, more scalability and lower sequencing project costs compared to the PacBio RS II.

Development of processes and applications for the Sequel System is well underway at the GRC. The team plans to use the increased throughput to expand their services for Iso-Seq transcriptome sequencing and amplicon projects. In collaboration with Dr. Jacques Ravel, associate director of IGS, they are also developing a full-length 16S sequencing pipeline to complement and expand their human microbiome and metagenomic research portfolios. As read lengths and throughput on the Sequel instrument improve, Tallon's team will shift whole genome sequencing projects onto the platform.

For other teams considering whether SMRT Sequencing is the right choice for them, Tallon says: "If you value complete genome sequences, *de novo* transcript discovery, and are looking at epigenetics in addition to the genome sequence, there's no better technology out there."

Luke Tallon and Lisa Sadzewicz are expanding their Iso-Seq transcriptome, amplicon, and whole genome sequencing services with the new Sequel System.

University of Maryland School of Medicine
Institute for Genome Sciences

Facility name:	Genomics Resource Center
Institution:	Institute for Genome Sciences, University of Maryland School of Medicine
Staff size:	15 scientists, technicians, and bioinformaticians
Year founded:	2007
Investigators served:	More than 500 worldwide and growing
PacBio Systems Installed:	Sequel System (Feburary 2016), PacBio RS II (June 2011)
Website:	http://www.igs.umaryland.edu/grc
Email:	grc-info@som.umaryland.edu

A PacBio Certified Service Provider

www.pacb.com

PN: PF105-052716

http://www.pacb.com/wp-content/uploads/2015/09/CoreLabProfile-LongReads-Institute-for-Genome-Sciences.pdf

Q&A and Open Discussion

Q&A AND OPEN DISCUSSION

Frequently Asked Questions

How long can I store my polymerase-bound sample?

- ***PacBio RS II:***
  - PacBio recommends that polymerase-bound samples be stored at 4°C and used within 3 days.
- ***Sequel System:***
  - PacBio recommends that polymerase-bound samples be stored at 4°C and used within 7 days.

How do I dissociate my polymerase-bound sample from MagBeads?

- Dissociating polymerase-bound sample from MagBeads may damage the sample and is not recommended. PacBio recommends binding sample to MagBeads immediately before sequencing and proceeding with sequencing as soon as possible. If a delay between MagBead binding and sequencing is unavoidable, Customers can store the sample in the dark at 4°C, but delaying sequencing will be at the Customer's own risk. If a MagBead sample has already been aliquoted into a sample plate, the sample plate should be sealed upon storage at 4°C. For Sequel samples, the sample plate should be heat-sealed with the Sequel Sample Plate Foil (P/N 100-667-400). For PacBio RS II samples, the sample plate should be temporarily sealed with an adhesive microplate sealing film and then the sealing film should be replaced with the PacBio RS II Sample Plate Septum (P/N 000-882-901) before sequencing.

How long can I store my MagBead bound sample?

- PacBio recommends that MagBead samples be stored at 4°C in the dark and sequenced as soon as possible.

My MagBeads were accidentally left at room temperature for several hours. Can they still be used?

- In most cases, MagBeads should still be useable by first chilling them at 4°C before use.

My MagBeads / AMPure beads were accidentally stored at -20℃. Is it still okay to use the beads?

- PacBio does not recommend using AMPure PB beads or MagBeads that have been accidentally stored at -20°C because the beads may become damaged and may leach after being frozen. However, Customers *may* use them at their own risk after bringing the MagBeads to 4°C and AMPure PB beads to room temperature.

When preparing >30 kb SMRTbell libraries, can (AMPure-purified and concentrated) sheared gDNA be stored at 4℃ for longer than 24 hours?

- PacBio generally recommends that AMPure-purified and concentrated sheared gDNA be stored for up to 24 hours at 4°C or at -20°C for longer durations. However, if the gDNA is relatively pure (i.e., free of endonucleases), it should be acceptable to store the sheared gDNA sample for 2-3 days at 4°C.

Conditions for shearing gDNA to a size that can support producing ≥30 kb libraries must be determined and verified empirically for each sample. When preparing ≥30 kb SMRTbell libraries using Megaruptor, what is the recommended target shear size if the desired size selection lower cutoff is, for example, 15-20 kb, 30 kb, or 40 kb?

- When preparing ≥30 kb SMRTbell libraries using Megaruptor, the recommended target shear size depends on the size selection lower cutoff to be employed. The Table below may be considered a useful starting point; but empirical optimization and accurate size quantitation are essential:

Library Insert Size (kb)	Size Selection Lower Cut (kb)	Target gDNA Shear Size (kb)
30	15 - 20	30
30 - 40	15 - 20	50
40 - 50	30	60
50 - 60	40	75

Where can I find the Plate Map and sequences of all the primers in the Barcoded Universal F/R Primers Plate - 96 (P/N 100-466-100) product and Barcoded Adapter Plate - 96 (P/N 100-466-000) product?

- To obtain the sequences of the primers used in the Barcoded Universal F/R Primers Plate - 96 Kit, please contact your local Field Applications Scientist, or submit your inquiry through the PacBio Customer Portal (http://www.pacbioportal.com/) or email techsupport@pacificbiosciences.com.
- The Barcode Plate Map Diagram can be downloaded from PacBio's Documentation webpage (http://www.pacb.com/support/documentation/) here: http://www.pacb.com/wp-content/uploads/2015/09/User-Bulletin-Barcode-Plate-Mapping.pdf

There is a 'Barcoding - RSII and SMRT Analysis 2.3.0 or older' webpage on GitHub (https://github.com/PacificBiosciences/Bioinformatics-Training/wiki/Barcoding). Where can I find the latest guidance on PacBio Barcoding recommendations for multiplexed sample preparation for Sequel / SMRT Link v4.0 (or later)?

- The most up to date information on PacBio multiplexing applicable to SMRT Link v4.0 (or later) can be found here: https://github.com/PacificBiosciences/SMRT-Link/wiki/SMRT-Analysis-Barcoding-Primer

Can I use Illumina 8-bp barcode index sequences for preparing multiplexed samples for PacBio sequencing?

- No; PacBio does **not** recommend using Illumina 8-bp barcode index sequences for preparing multiplexed samples for PacBio SMRT sequencing applications.

How are the 16-bp PacBio barcodes incorporated into the SMRTbell DNA template?

- PacBio uses two approaches:
  - Adding a barcode to end of the standard SMRTbell adapter. The combined adapter is called a Barcoded Adapter.
  - Adding a barcode to the PCR amplicon. This approach involves a two-step PCR reaction workflow. The internal primers for the first PCR are augmented at the 5' end by universal sequences to the target-specific primers. The external primers contain the 16bp barcode at the 5' end and the universal sequences. This approach is called Barcoded Universal Primers.

What are the supported applications for using PacBio Barcoded Adapters and PacBio Barcoded Universal Primers with multiplexed samples? What are not supported applications?

- Supported applications are sequencing of **one species per sample or loci**. Examples of supported applications include: Confirmation of SNPs, resequencing, most Long Amplicon Analysis (LAA) applications, and Sanger sequencing replacement. An exception is HLA typing, which may have 2 species per loci. Multiplexing of HLA has also been demonstrated with the use of additional custom analyses (see PacBio's AGBT 2015 Poster: http://files.pacb.com/pdf/Poster_MultiplexingHumanHLAGenotyping_DNABarcodeAdapters_HighThroughputResearch.pdf)
- Note: The product specifications for the PacBio Barcoded Adapter Kit and PacBio Barcoded Universal Primer Kit are such that the level of barcode oligo contamination in the 96-plate wells should not exceed 5%. Therefore it is possible, though unlikely, to have 1 other contaminant barcode primer/adapter sequence present at levels up to 5%. PacBio does not recommend using the PacBio Barcoded Adapter Kit and PacBio Barcoded Universal Primer Kit for minor variant detection < 10%.

Does PacBio have any specific DNA polymerase enzyme or Kit recommendations for long-range PCR (LR PCR) for generating long DNA amplicon samples for sequencing?

- While PacBio does not recommend a specific enzyme, a high-fidelity enzyme is generally preferred. For example, PrimeStart GXL from Takara and ThermoFisher Phusion Hot Start II DNA Polymerase have given good results to our internal scientists.

Other Discussion Points

What general 'best practices' can be adopted by PacBio service providers to be more effective partners to their customers?

What types of resources would you like PacBio to offer to enhance service provider deliverables?

What are your opinions on the current state of demand for SMRT Sequencing?

Which applications are you most likely to add to your list of services?

Which PacBio applications are most frequently employed?

- *Setting deliverables and expectations with clients*
- *Nucleic acid isolation*
- *Sample QC workflows*
- *Sample library processing workflows*
- *Run QC performance metrics*
- *Real time monitoring of run status*
- *Troubleshooting sequencing performance issues*
- *Turnaround time*
- *Instrument reliability and uptime*
- *Laboratory information management (LIMS)*
- *Application programming interfaces (APIs)*
- *Bioinformatics data analysis services*
- *IT support and data storage*
- *Report Outputs*
- *Automation*
- *Technical staff training (Library Sample Prep, Instrument Operators)*
- *Open Houses and Applications Seminars*
- *Technical Project Consultations*

www.pacb.com