



SMRTbell Library Preparation & SMRT Sequencing Workflow Updates

PacBio Americas User Group Meeting Sample Prep Workshop

June.27.2017

AGENDA

- Asymmetric SMRTbell Templates for Amplicon Sequencing
 - Asymmetric SMRTbell Template Overview
 - Asymmetric Library Template Prep Workflow
 - Example Data
 - Conclusions
- Sequel System Pre-Extension Run Design Feature
- Diffusion Loading for Sequencing Large Insert Libraries on Sequel System



Asymmetric SMRTbell Templates for Amplicon Sequencing

OPTIONS FOR INCREASING READ LENGTH

- Longer movie collection times
 - Can specify movie lengths up to 10 hours (Sequel System)
- Pre-Extension Time Feature (Sequel System) [**NEW!**]
 - Pre-extension feature enables the start of movie acquisition to be delayed until the polymerase enters rolling circle replication and is in its most stable (processive) phase of sequencing
 - Pre-extension effectively helps increase the mean Polymerase Read Length metric by reducing the number of early-termination reads collected during primary analysis
- SMRTbell template format: Symmetric vs Asymmetric [**NEW!**]
 - With Asymmetric SMRTbell templates, only **one** polymerase is active during sequencing



- Asymmetric Template Prep workflow has only been validated for the Sequel instrument (Update to instrument control software is not required)

OVERVIEW OF ASYMMETRIC SMRTBELL TEMPLATE USE CASES & EXPECTED SEQUENCING PERFORMANCE BENEFITS

Appropriate Sample Types

- For PCR products in the ~2 – 10 kb size range
- Not applicable to sheared gDNA samples

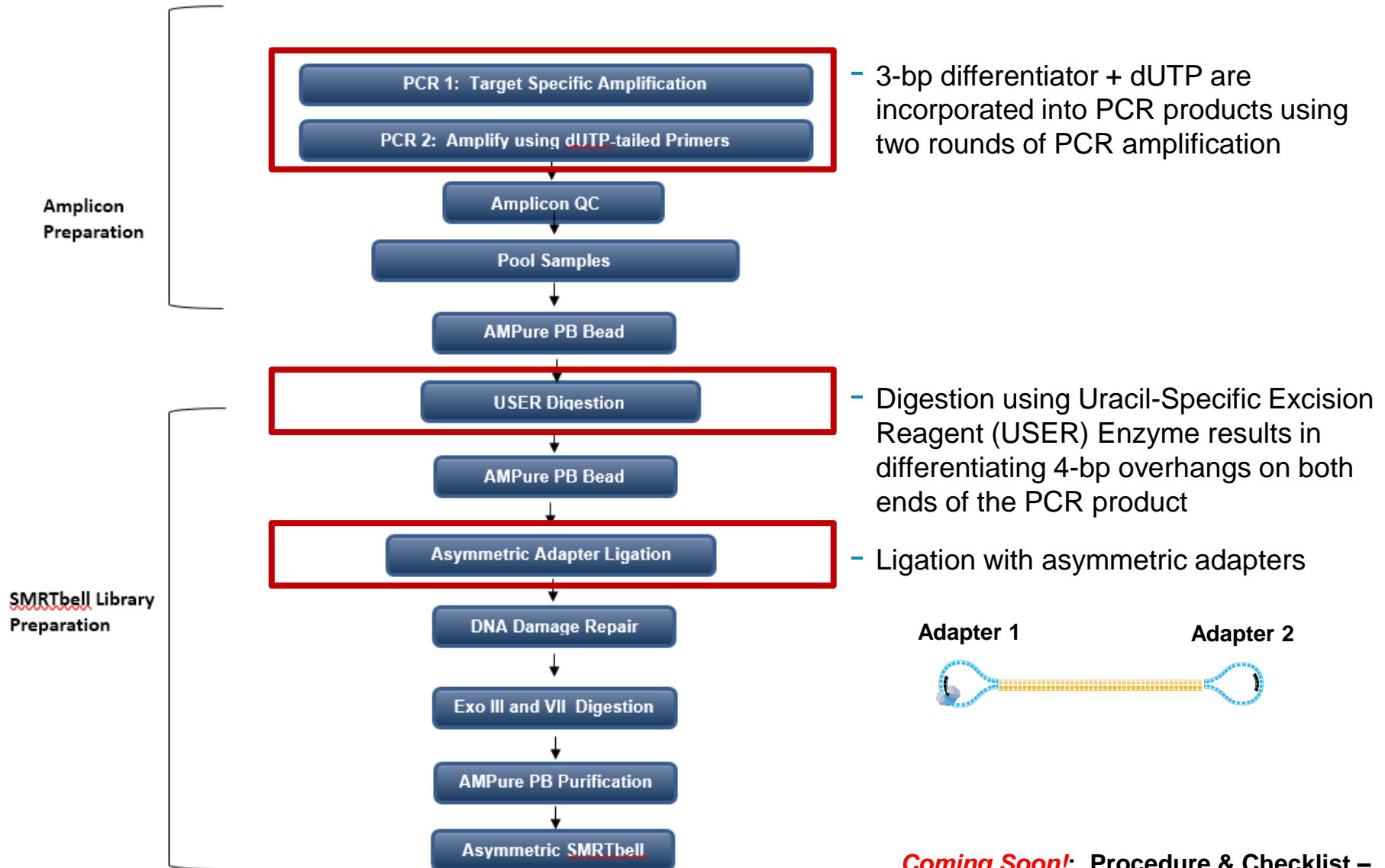
Expected Sequencing Performance Improvements

- More robust towards overloading
 - At higher loading concentrations, impact on polymerase read length is less with Asymmetric templates relative to Symmetric templates
- Can enable achievement of longer mean polymerase read lengths
 - Increased number of subreads
 - Higher yield of circular consensus sequencing (CCS) reads

Example Use Case

- High-throughput sequencing of PCR products where achieving equimolar pooling is a challenge (e.g. HLA amplicons 3 - 5 kb)

ASYMMETRIC SMRTBELL TEMPLATE PREP WORKFLOW OVERVIEW



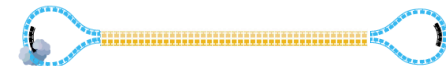
- 3-bp differentiator + dUTP are incorporated into PCR products using two rounds of PCR amplification

- Digestion using Uracil-Specific Excision Reagent (USER) Enzyme results in differentiating 4-bp overhangs on both ends of the PCR product

- Ligation with asymmetric adapters


Adapter 1

Adapter 2



Coming Soon! Procedure & Checklist – Preparing Asymmetric SMRTbell Templates

PROCEDURE & CHECKLIST – PREPARING ASYMMETRIC SMRTBELL TEMPLATES



Procedure & Checklist - Preparing Asymmetric SMRTbell™ Templates


Before You Begin

In this procedure, PCR products are generated using two rounds of amplification. The first round uses target specific primers tailed with a universal sequence and the second uses barcoded universal primers tailed with a 4 bp sequence containing dUTP. The PCR product is treated with USER (Uracil-Specific Excision Reagent) Enzyme resulting in 4 bp overhangs differentiated by a 3 bp sequence, followed by ligation with Asymmetric SMRTbell adapters. The resulting SMRTbell templates are annealed to two different primers: the first primer has a 3' poly A tail for binding oligo dT magnetic beads, and the second primer is available for polymerase binding.

General Workflow for Asymmetric Template Preparation

Amplicon Preparation

SMRTbell Library Preparation



```

graph TD
    A[PCR 1: Target Specific Amplification] --> B[PCR 2: Amplify using dUTP-tailed Universal Primers]
    B --> C[Amplicon QC]
    C --> D[Pool Samples]
    D --> E[AMPure PB Bead]
    E --> F[USER Digestion]
    F --> G[AMPure PB Bead]
    G --> H[Asymmetric Adapter Ligation]
    H --> I[DNA Damage Repair]
    I --> J[Exo III and VII Digestion]
    J --> K[AMPure PB Purification]
    K --> L[Asymmetric SMRTbell Template]
        
```

Page 1

Document Contents

- Guidance for designing primers
- PCR reaction cycling parameter recommendations
- Asymmetric Adapter ligation conditions
- Primer annealing and binding recommendations
- Sequencing recommendations

Required Materials

Amplification:

- KAPA HiFi HotStart Uracil+ PCR Kit from Kapa Biosystems (Kit Codes KK2801 or KK2802)
- Primers for 1st and 2nd PCR:
 1. Target Specific Primers tailed with Universal sequence
 2. Forward and Reverse Barcoded Universal Primers tailed with 4 bp containing dUTP

Library Construction:

- PacBio® SMRTbell™ Template Prep Kit (100-259-100)
- NEB USER™ Enzyme (Catalog #M5505)
- PacBio SMRTbell Asymmetric Auxiliary TPK (101-080-400)
- PacBio AMPure® PB Beads (100-265-900)

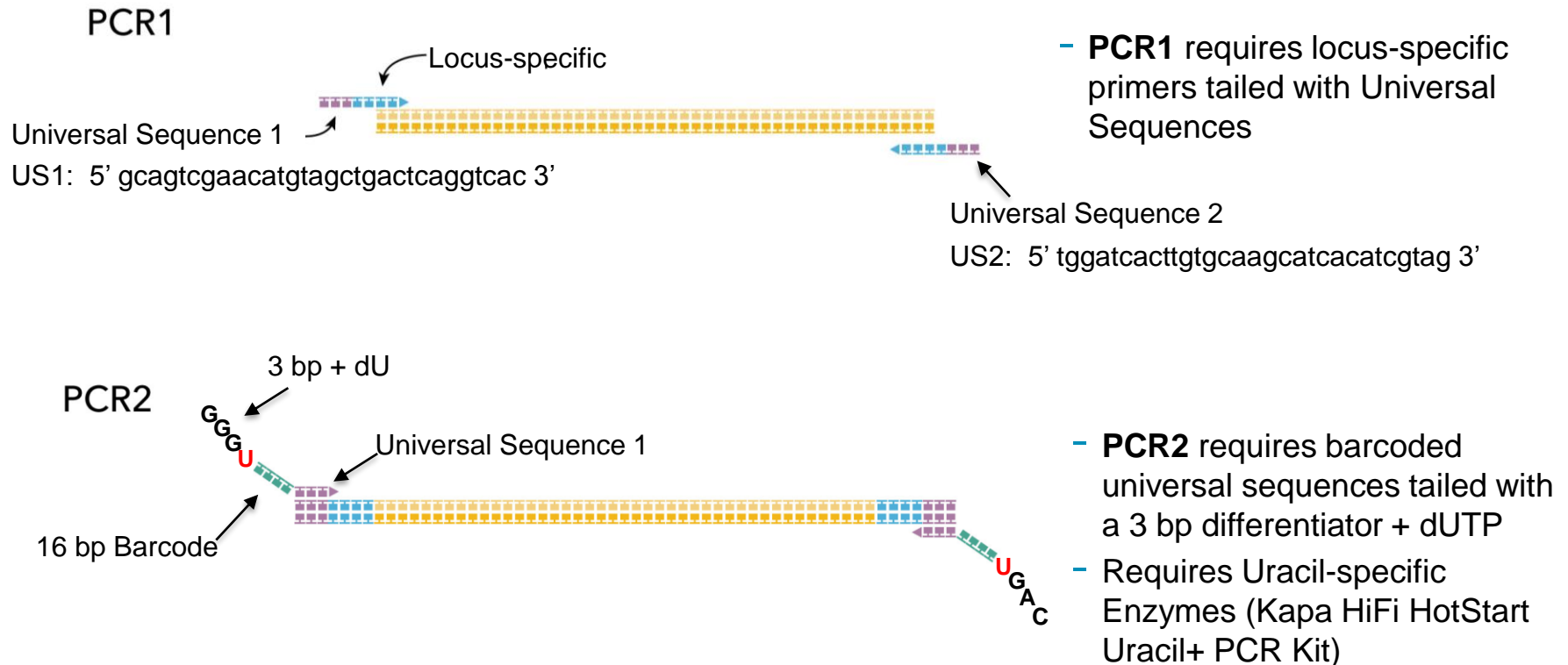
<http://www.pacb.com/support/documentation/>

(Coming Soon!)

ASYMMETRIC TEMPLATE PREP KIT DETAILS

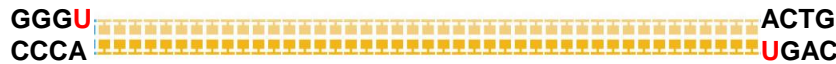
Part Number	Description
101-113-000	SMRTbell™ Asymmetric TPK 1.0
100-259-100	➤ SMRTbell™ Template Prep Kit 1.0 (Standard template prep kit)
101-080-400	➤ SMRTbell™ Asymmetric Auxiliary TPK (Asymmetric adapters PLUS sequencing & capture primers)
101-113-100	SMRTbell™ Bulk Asym TPK 1.0 100 rxn
100-946-300	➤ SMRTbell™ Bulk Temp Prep Kit 1.0 100 rxn (Bulk version of the standard template prep kit)
101-080-800	➤ SMRTbell™ Bulk Asym Auxiliary TPK (Bulk version of the asymmetric auxiliary template prep kit)

INCORPORATION OF dUTP VIA TWO-STEP PCR AMPLIFICATION



- Primers can be ordered from any oligo synthesis provider and should be desalted

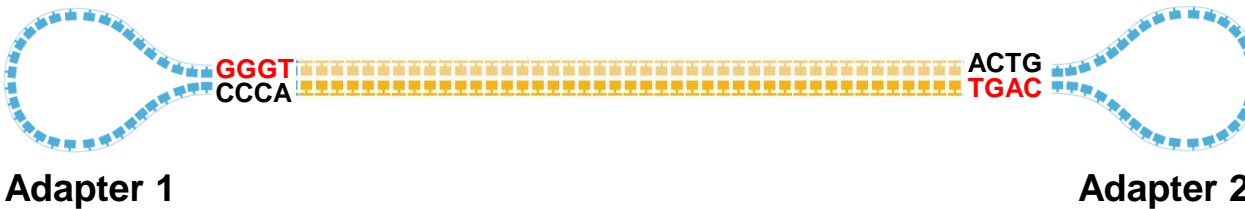
USER (URACIL-SPECIFIC EXCISION REAGENT) ENZYME DIGESTION GENERATES DIFFERENT 4-BASE PAIR OVERHANGS AT EACH END



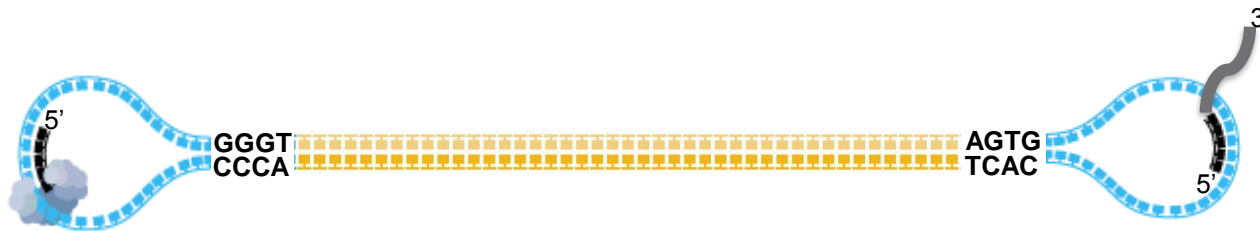
- USER Enzyme is a mixture of Uracil DNA glycosylase (UDG) and Endonuclease VIII (a DNA glycosylase-lyase)
- Digestion with USER Enzyme generates a different 4-bp overhang on both ends of PCR product
- 37°C for 15 minutes, 25°C for 15 minutes



- Ligate to two different adapters

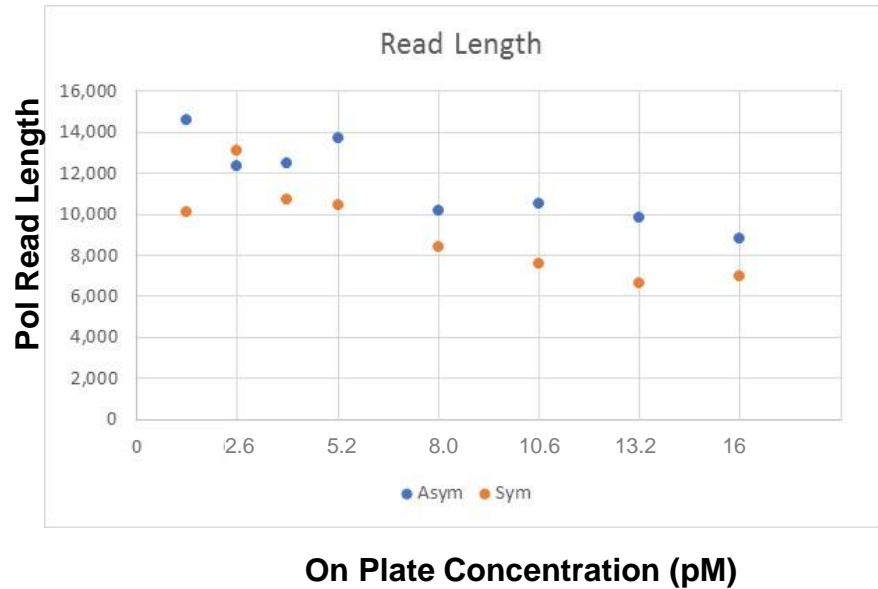
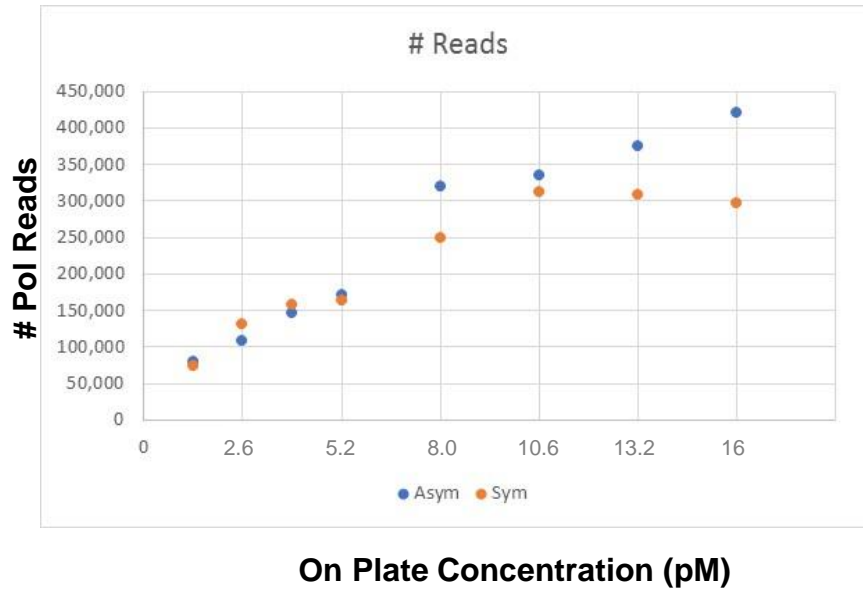


ASYMMETRIC TEMPLATE PRIMER ANNEALING REACTION



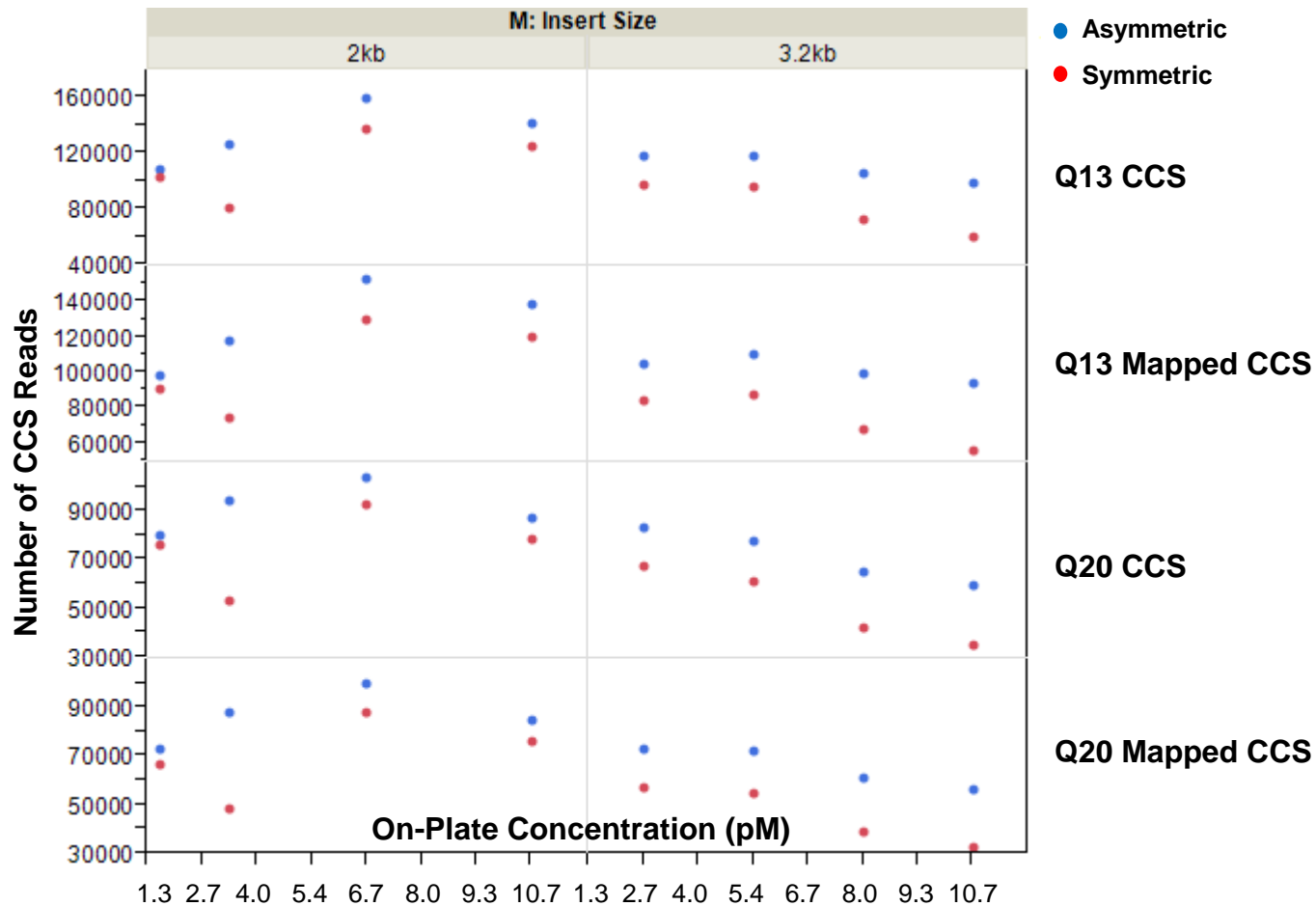
- The two different hairpin adapters are primed with two different oligos
- **Asymmetric Seq Primer A**
 - No poly-A tail
 - Facilitates polymerase binding to 3' end of primer
- **Asymmetric Capture Primer B**
 - Poly-A tail at 3' end of primer
 - Does not facilitate polymerase binding
- As a result, only **one end** of the SMRTbell template is immobilized to the bottom of the zero mode waveguide (ZMW) and only **one end** initiates a sequencing reaction

LONGER POLYMERASE READ LENGTHS ARE ACHIEVABLE WITH ASYMMETRIC TEMPLATES COMPARED TO SYMMETRIC TEMPLATES (5 KB AMPLICON)



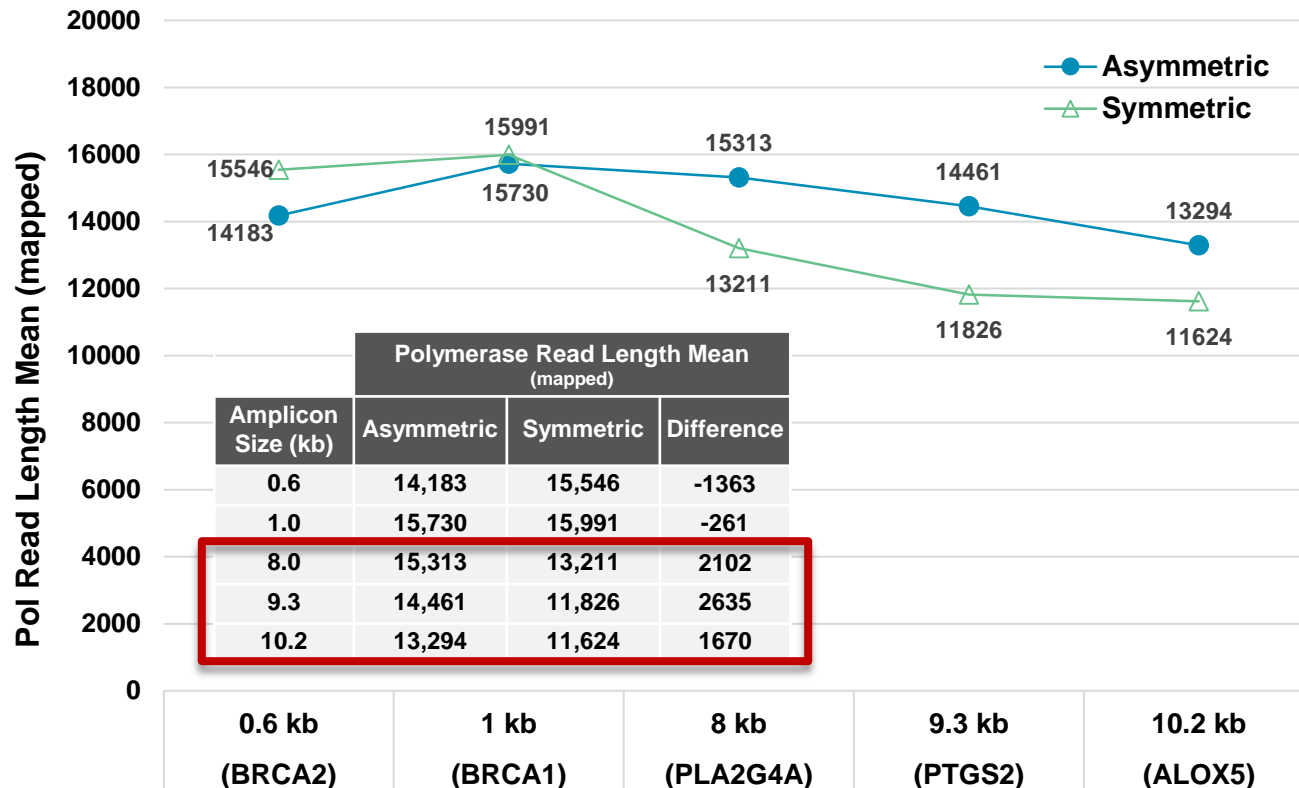
- Overloading negatively impacts Polymerase Read Lengths for both Asymmetric and Symmetric templates
- At higher loading concentrations, longer Polymerase Read Lengths are achievable with Asymmetric templates compared to Symmetric templates

HIGHER YIELD OF CCS READS CAN BE ACHIEVED WITH ASYMMETRIC TEMPLATES VS. SYMMETRIC TEMPLATES (2 KB AND 3.2 KB AMPLICONS)



- Because of the longer Polymerase read lengths, more circular consensus sequencing (CCS) reads are achievable with Asymmetric templates at higher loading concentrations
- Asymmetric template is more robust to overloading compared to Symmetric templates

EFFECT OF ASYMMETRIC VS SYMMETRIC SMRTBELL TEMPLATE FORMAT ON MAPPED READ LENGTH PERFORMANCE FOR AMPLICONS OF VARIED SIZE (0.6 KB – 10.2 KB)



- Resequencing analysis performed with SMRT Analysis v4.0
- In this example, longer mean mapped polymerase read lengths were achieved with Asymmetric templates compared to Symmetric Templates for larger 8.0 kb, 9.3 kb and 10.2 kb amplicons

DNA INPUT REQUIREMENTS ASYMMETRIC SMRTBELL LIBRARY CONSTRUCTION FOR DIFFERENT AMPLICON SIZE BINS

- The amount of amplified DNA required for Asymmetric library construction depends on the size of the PCR product

Insert size range	2 - 3 kb	3 - 4 kb	4 - 5 kb	>5 kb
Input DNA amount	2-3 μ g	3-4 μ g	4-5 μ g	5 μ g

- If you are constructing a single PCR product, prepare enough PCR products to meet the above input DNA requirements.
- For multiplexing, prepare enough PCR products such that when pooled, the total amount of amplified product meets the above input DNA requirements.

CONCLUSIONS

Asymmetric SMRTbell Library Template Features

- For use with PCR amplicon products only
- More robust to overloading
 - Impact on read length is less, relative to symmetric templates
- Can enable achievement of longer polymerase reads
 - More subreads
 - More CCS reads
- May be suited for high-throughput sequencing of PCR products where achieving equimolar pooling is a challenge (e.g. HLA amplicons 3-5 kb)
- Update to instrument control software is not required




Sequel System Pre-Extension Run Design Feature

OVERVIEW OF PRE-EXTENSION CONCEPT FOR SEQUEL RUN DESIGNS

Sequencing initiation:



Sequencing through the second adapter and back onto the initial strand:

- Polymerase activity is most stable during rolling circle replication mode
 - Start data collection **here** 
- Pre-extension time value depends on the insert size



- *Pre-extension (PE) enables the start of movie acquisition to be **delayed** until the polymerase enters rolling circle replication and is in its **most stable** (processive) phase of sequencing*

- PE effectively helps increase the mean Polymerase Read Length metric by reducing the number of early-termination reads collected during primary analysis
- Because early terminations that occur during PE are not collected, a reduction in %P=1 reads may be observed; however, this can be compensated by using a higher sample on-plate loading concentration (e.g., >2X over default recommendations for non-PE run designs)

RECOMMENDED PRE-EXTENSION TIME SETTINGS

Mean Insert Size	Distance (bases)	Time (minutes)
1 kb	2000	≥30
2 kb	4000	45
4 kb	8000	90
6 kb	12000	135

- Pre-extension times are estimated based on a base incorporation rate of 1.5 bases sec⁻¹ during polymerase extension:

$$\text{Pre-Extension Time (min)} = [\text{Insert Size (bp)} \times 2] / 1.5 \text{ bases sec}^{-1} / 60 \text{ sec min}^{-1}$$

RECOMMENDED USE CASES FOR PRE-EXTENSION FEATURE AND EXPECTED SEQUENCING PERFORMANCE BENEFITS

Sequel System Amplicon sequencing

- PCR products ~1 - 5 kb insert size range
- Pre-extension improves mean Polymerase Read Length of amplicon samples
- Diffusion loading

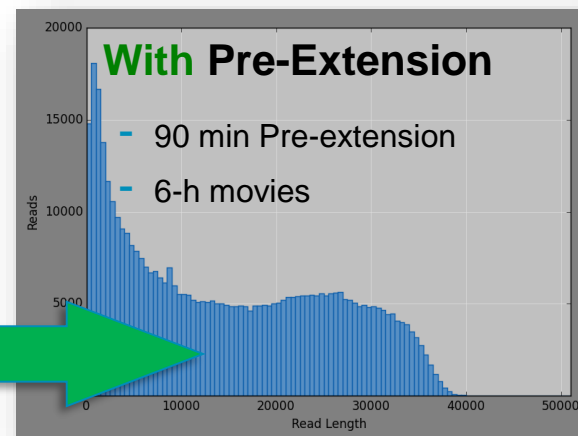
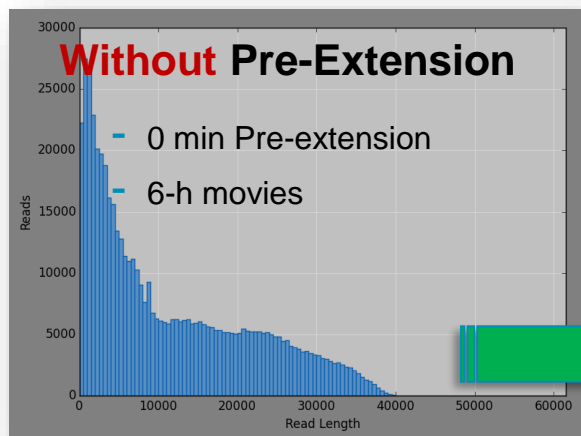
Sequel System Iso-Seq analysis

- PCR-amplified non-size selected cDNA library
- Default recommendation is to use Pre-extension feature for Iso-Seq analysis
- MagBead loading

Sequel Microbial Multiplexing

- Non-size selected sheared gDNA library (~10 kb insert size range)
- Pre-extension improves yield of barcoded reads for multiplexed samples
- MagBead loading

EXAMPLE: 3 KB AMPLICON DATA SET



Mapping Report

Value	Analysis Metric
86.33%	Mean Concordance (mapped)
1,848,101	Number of Subreads (mapped)
6,100,507,913	Number of Subread Bases (mapped)
3,301	Subread Length Mean (mapped)
4,307	Subread Length N50 (mapped)
4,560	Subread Length 95% (mapped)
6,025	Subread Length Max (mapped)
539,719	Number of Polymerase Reads (mapped)
11,578	Polymerase Read Length Mean (mapped)
21,297	Polymerase Read N50 (mapped)
31,280	Polymerase Read Length 95% (mapped)
61,119	Polymerase Read Length Max (mapped)

+3.2 kb

+3.1 kb

Mapping Report

Value	Analysis Metric
86.69%	Mean Concordance (mapped)
1,844,327	Number of Subreads (mapped)
6,431,249,275	Number of Subread Bases (mapped)
3,487	Subread Length Mean (mapped)
4,379	Subread Length N50 (mapped)
4,730	Subread Length 95% (mapped)
5,980	Subread Length Max (mapped)
445,874	Number of Polymerase Reads (mapped)
14,765	Polymerase Read Length Mean (mapped)
24,393	Polymerase Read N50 (mapped)
33,200	Polymerase Read Length 95% (mapped)
50,946	Polymerase Read Length Max (mapped)

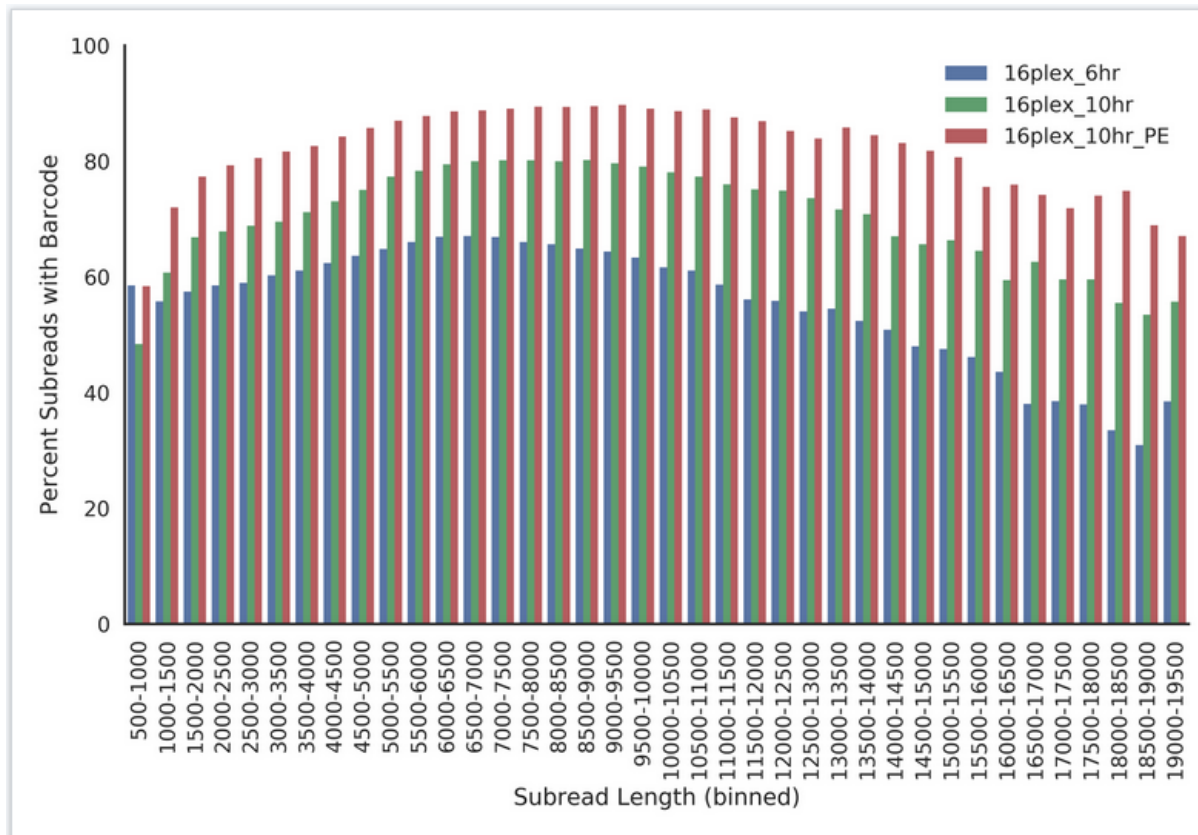
EXAMPLE SEQUEL SYSTEM ISO-SEQ ANALYSIS DATASET: SECONDARY ANALYSIS METRICS (ISO-SEQ PROTOCOL)

Sample	[On-Plate] (pM)	Movie Time (min)	#CCS	CCS RL (bp)	#FLNC (%)	FLNC RL (bp)	# Polished HQ Isoforms	# Polished LQ Isoforms
Cancer Cell Line	50 pM	360	415,539	2602	202,328 (48.7%)	2892	14,722	96,755
Cancer Cell Line	40 pM	600	545,724	2535	264,779 (48.5%)	2867	20,386	132,119
Cancer Cell Line	50 pM	600	595,533	2597	244,521 (41.1%)	3019	17,765	125,467

Pre-extension = 120 min, MagBead loading

- PacBio's current default recommendation is to run Sequel System Iso-Seq analysis samples with MagBead Loading and Pre-extension enabled
- 1 Sequel SMRT Cell typically yields $\geq 200,000$ Full-length non-chimeric (FLNC) reads

EXAMPLE: MICROBIAL MULTIPLEXING DATA SET (10 KB, 16-PLEX, 1.6 MB GENOMES)



Pre-extension = 120 min, MagBead loading

- Using longer movies + Pre-extension increased the percentage yield of barcoded reads for this multiplexed sheared gDNA library sample (~10 kb insert size)

CONCLUSIONS

Pre-Extension Can Improve Sequencing Performance of Amplicons and Other Types of Small Insert Libraries

- Sequel System Amplicon sequencing (~1 – 5 kb or larger)
 - Increases mean Polymerase Read Length
- Sequel System Iso-Seq analysis (PCR-amplified non-size selected cDNA library)
 - Default recommendation is to run samples for Iso-Seq analysis with Pre-extension enabled
- Sequel System Microbial Multiplexing (Non-size selected sheared gDNA library ~10 kb)
 - Pre-extension improves yield of barcoded reads



Diffusion Loading for Sequencing Large Insert Libraries on Sequel System

DIFFUSION LOADING FOR LARGE INSERT LIBRARIES (SEQUEL)

With Sequel Sequencing Kit v2.0, large insert libraries can be loaded efficiently using Diffusion loading

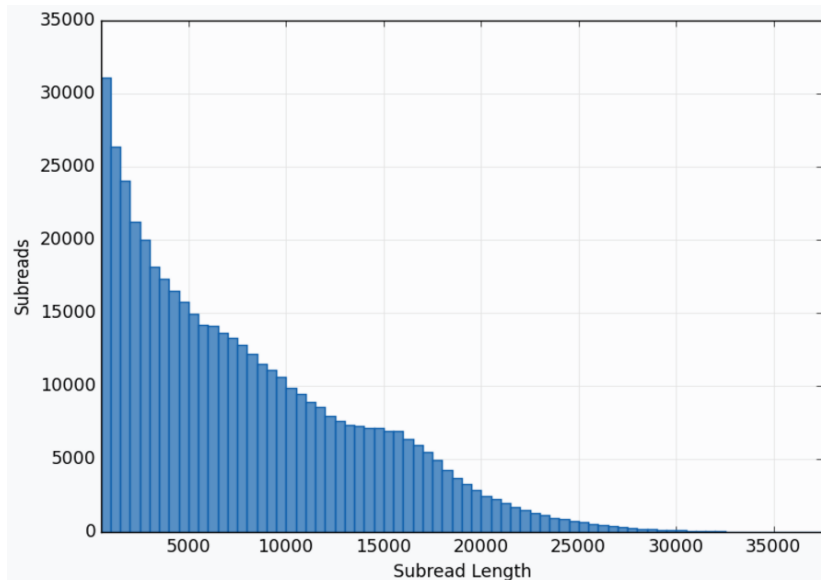
Advantages

- Diffusion loading requires less input SMRTbell library material per SMRT Cell 1M compared to MagBead loading
- Using Diffusion loading reduces overall cost by eliminating MagBeads
- Total time needed to prepare samples for loading onto instrument is reduced with Diffusion loading
 - Eliminates time required for MagBead binding incubations
- A more simplified workflow for preparing samples for Diffusion loading
 - New cleanup protocol uses MicroSpin Columns instead of Loading Cleanup Beads to remove excess polymerase and sequencing primer

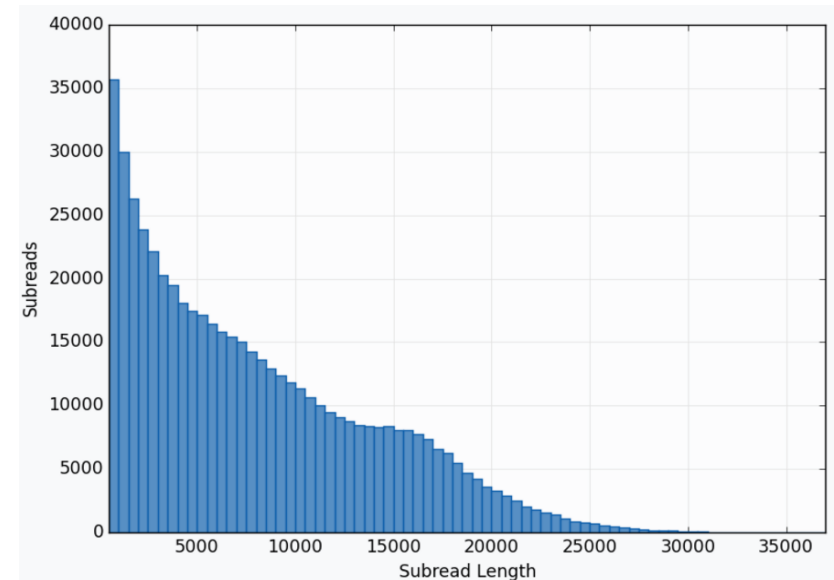


DIFFUSION VS. MAGBEAD LOADING EXAMPLE: SIZE-SELECTED 20 KB *E. COLI* LIBRARY (15 KB BLUE PIPPIN LOWER CUT)

Diffusion Loading, MicroSpin Column Cleanup



MagBead Loading, MicroSpin Column Cleanup



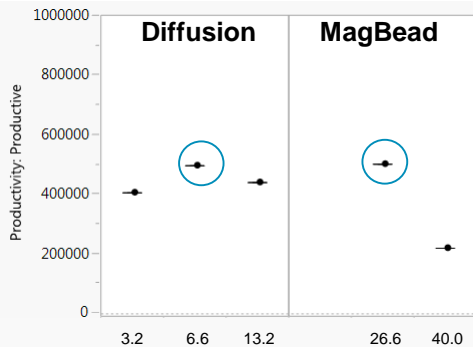
- Comparable Subread Read Length performance was achieved for long insert libraries loaded by Diffusion or MagBeads

COMPARISON OF (SMRT ANALYSIS) MAPPING STATISTICS FOR DIFFUSION VS. MAGBEAD LOADED 20-KB *E. COLI* LIBRARY

Mapping Report	MagBead	Diffusion
Mean Concordance (mapped)	86.02%	86.15%
Number of Subreads (mapped)	535,145	504,233
Number of Subread Bases (mapped)	4,028,736,571	3,574,948,545
Subread Length Mean (mapped)	7,528	7,090
Subread Length N50 (mapped)	12,315	11,626
Subread Length 95% (mapped)	18,850	18,170
Subread Length Max (mapped)	36,180	35,766
Number of Polymerase Reads (mapped)	476,102	433,766
Polymerase Read Length Mean (mapped)	8,476	8,259
Polymerase Read N50 (mapped)	13,893	13,698
Polymerase Read Length 95% (mapped)	22,600	22,960
Polymerase Read Length Max (mapped)	44,240	52,444

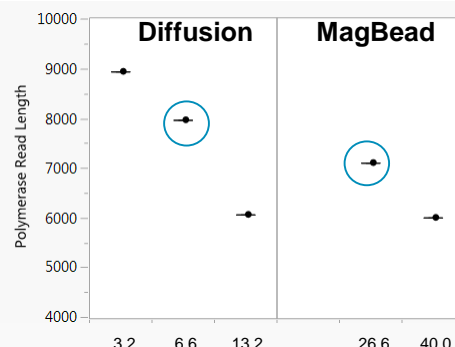
DIFFUSION LOADING REQUIRES LESS SMRTBELL LIBRARY SAMPLE THAN MAGBEAD LOADING: 20-KB *E. COLI*/LIBRARY EXAMPLE

P1 Loading



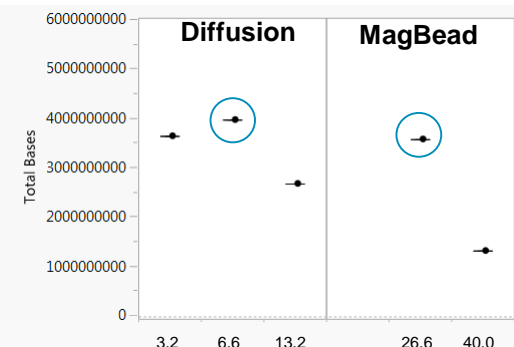
[On-Plate]

Pol Read Length



[On-Plate]

Yield/SMRT Cell

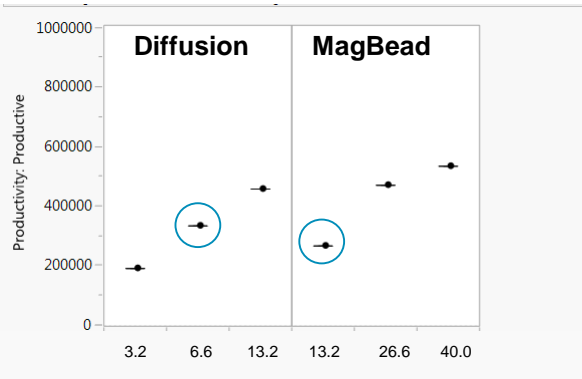


[On-Plate]

- Diffusion loading required 3- to 4-fold less DNA library amount compared with MagBead loading
 - Blue circles indicate 6.6 pM [on-plate] for Diffusion loading and 26.6 pM [on-plate] for MagBead loading
- Mean Polymerase Read Length is negatively impacted at higher on-plate concentrations for both loading conditions
- ~4 Gb/SMRT Cell is achievable by using Diffusion loading with 6-hour movies

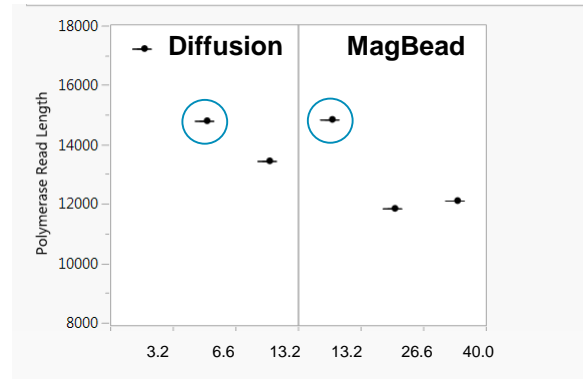
DIFFUSION LOADING REQUIRES LESS SMRTBELL LIBRARY SAMPLE THAN MAGBEAD LOADING: 48-KB LAMBDA BELL EXAMPLE

Loading



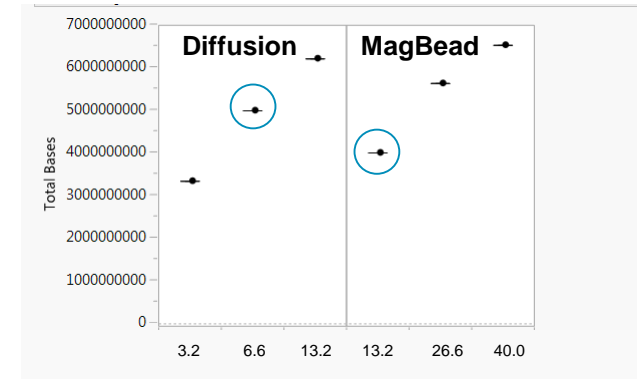
[On-Plate]

Pol Read Length



[On-Plate]

Yield/SMRT Cell



[On-Plate]

- Diffusion loading required a lesser amount of DNA library compared with MagBead loading
 - Blue circles indicate 6.6 pM [on-plate] for Diffusion loading and 13.2 pM [on-plate] for MagBead loading
- Mean Polymerase Read Length is negatively impacted at higher on-plate concentrations for both loading conditions
- >5 Gb/SMRT Cell is achievable by using either Diffusion loading or MagBead loading with 10-hour movies

SAMPLE PURIFICATION USING MICROSPIN COLUMNS IS NECESSARY FOR DIFFUSION LOADING WITH SEQUEL SYSTEM



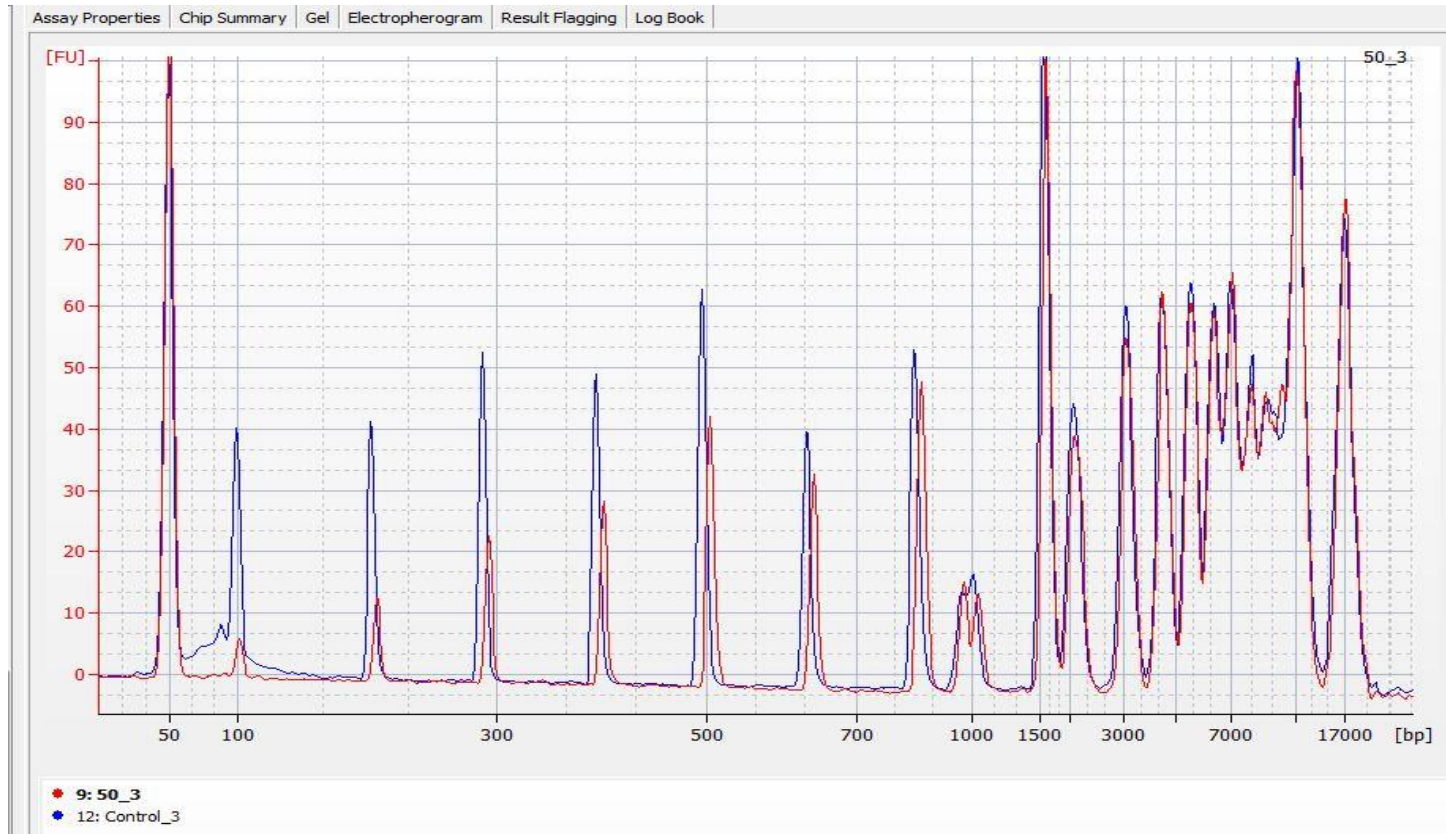
- MicroSpin columns are easier to use than the existing Loading Cleanup Beads or SMRTbell™ Clean Up Columns (v1)
 - Swinging bucket centrifuge equipment is not required
- Recommended spin column type: **MicroSpin S-400 HR**
 - GE HealthCare Life Science (P/N 27-5140-01)

MICROSPIN S-400 HR LIBRARY RECOVERY YIELD PERFORMANCE

Control vs 50 μ l
(100 ng/ μ l Input)


Size [bp]	% Recovery
100	*
200	*
300	37.4%
400	47.0%
500	57.4%
650	66.8%
850	74.8%
1650	82.9%
2000	82.1%
5000	83.8%
12,000	86.2%

* Not called a peak by BioAnalyzer



- MicroSpin S-400 HR columns are recommended for use with inserts **≥ 2.0 kb**
- Removes free polymerase, sequencing primer and adapter dimers

MICROSPIN COLUMNS SHOW SIMILAR RECOVERY YIELDS AT VARIED DNA INPUT LEVELS



ul Filtered	% Recovery	
	10ng/ul	100ng/ul
50	74.9%	88.7%
40	75.2%	79.6%
30	65.9%	73.9%
20	66.8%	55.6%

- 1kb Plus ladder was diluted to 10 ng/ μ l or 100 ng/ μ l
- Various volumes of diluted ladder were filtered through MicroSpin columns
- Optimal recovery was obtained from using **50 μ l** of sample as input
 - Recommend that samples be diluted to at least 50 μ l total volume for use with the MicroSpin purification procedure

PROCEDURE & CHECKLIST DOCUMENTS FOR SAMPLE LIBRARY CLEAN-UP USING MICROSPIN S-400 COLUMNS

Procedure & Checklist – Sample Clean-Up Using MicroSpin™ Columns S-400 for Diffusion Loading

<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>

Procedure & Checklist – Sample Clean-Up Using MicroSpin™ Columns S-400 for MagBead Loading

<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**

Before You Begin

This document describes a procedure for purifying polymerase-bound complexes, using MicroSpin columns, for Diffusion loading on the Sequel System. This is recommended for libraries or amplicons ≥ 2 kb. For < 2 kb SMRTbell templates, use Loading Clean-Up Beads. Note that the purification step is performed after polymerase binding to remove excess unbound polymerase and polymerase bound to small DNA inserts and adapter dimers. Typical complex recovery, post purification, is 70%-90%.



Procedure & Checklist - Sample Clean-Up Using MicroSpin™ Columns S-400 for **MagBead Loading**

Before You Begin

This document describes a procedure for purifying polymerase-bound complexes, using MicroSpin columns, for MagBead loading in the Sequel System. This is recommended for libraries or amplicons ≥ 2 kb. For < 2 kb SMRTbell templates, use Loading Clean-Up Beads. Note that the purification step is performed after polymerase binding to remove excess unbound polymerase and polymerase bound to small DNA inserts and adapter dimers. Typical complex recovery, post purification, is 70%-90%.

CONCLUSIONS

- With Sequel Sequencing Kit v2.0, large insert libraries can be loaded efficiently using Diffusion loading
- Sequencing performance metrics for Diffusion-loaded samples are comparable to MagBead-loaded samples
- Advantages
 - Diffusion loading requires less input SMRTbell library material
 - Reduces overall cost by eliminating MagBeads
 - Sample preparation time before sample is loaded onto Sequel instrument is reduced
 - Eliminates time for MagBead Binding



New 3rd Party Technology Offerings

NEW Product Offering from AATI: FEMTO Pulse (NEW!)

FEMTO *Pulse* is the only pulsed field capillary electrophoresis instrument available that can quantify, qualify, and size low-concentration and / or large-size nucleic acid samples with high accuracy and precision.

FEMTO *Pulse* easily analyzes diverse sample types including: cfDNA, total RNA, genomic DNA, large fragment DNA, messenger RNA, and more.

- Using a pulsed-field power supply, the FEMTO *Pulse* is the first parallel capillary electrophoresis instrument able to resolve DNA smears and DNA fragments through 200,000 bp
- Optical detection platform enables unprecedented sensitivity, detecting nucleic acids into the lower femtogram range
- DNA and RNA samples can be separated on the same capillary array. Furthermore, two different gel matrices can be loaded, enabling the unattended, sequential separation of RNA and DNA samples (e.g. mRNA followed by genomic DNA), without cumbersome array swaps or cleaning



NEW Product Offering from Sage Science: SageHLS (NEW!)

The SageHLS (HMW Library System) provides a new approach to DNA extraction that enables extraction of DNA up to 2 MB in size directly from cells

- Users load cell suspensions on Sage Science gel cassettes to perform lysis under electrophoretic conditions
- Degraded and solubilized proteins are removed but intact DNA is left behind, bound in agarose
- A cleavase is then used to carefully fragment the DNA into electrophoretically-mobile sizes
- Automated DNA size selection collects the DNA in 6 size bins



 sage science



www.pacb.com

For Research Use Only. Not for use in diagnostics procedures. © Copyright 2017 by Pacific Biosciences of California, Inc. All rights reserved. Pacific Biosciences, the Pacific Biosciences logo, PacBio, SMRT, SMRTbell, Iso-Seq, and Sequel are trademarks of Pacific Biosciences. BluePippin and SageELF are trademarks of Sage Science. NGS-go and NGSengine are trademarks of GenDx. FEMTO Pulse and Fragment Analyzer are trademarks of Advanced Analytical Technologies. All other trademarks are the sole property of their respective owners.