Consumable Updates

Paul Kotturi

AGENDA

- Welcome
- Field Advisory Notifications and User Bulletins
- Consumable Updates
- Application Updates
- Overloading Pros and Cons

WELCOME TO THE SAMPLE PREP WORKSHOP

Reviewing all aspects of PacBio and our Partners consumable workflows

- DO ASK QUESTIONS
  - During a talk if there is a something that you don't know, it's likely that it's unclear to many others
  - Your questions will foster interactions with other users
    - Have you been able to get >30 kb size-selection protocol with reasonable yields?
    - I have a tricky sample to do DNA extractions. What do you recommend?
- DO NOT LEAVE WITHOUT YOUR QUESTIONS ANSWERED
  - **We're here for 1.5 days and if we don't know immediately, we'll find out**

Field Advisory Notices and User Bulletins

PACBIO IS ENSURING THE BEST PERFORMANCE THROUGH OPEN AND THOROUGH COMMUNICATION OF INFORMATION

Notifications in the form of Field Advisory Notices and User Bulletins

- Heat Seals
  - Do not use adhesive foil seals
- Axygen Tubes and Tips are not compatible with SMRT Sequencing
- Sequel Tips
  - Tape the tip rack to the box
  - New version coming
- Sequel Sequencing Kit tracking
  - Use barcode found in the box
- Tube Septa
  - Portion of the septa do not have complete cut slit

If you're not getting emails, please work with your Field Application Scientist to clarify your notifications.

Consumable Updates

NEW CONSUMABLES AND PROTOCOLS

- Sequel Pipet Tips v2
  - Same price, new secure tip rack
- Asymmetric Template Prep Kit
  - Auxiliary kit for amplicons
- Sequel Sequencing Kit 2.0 (4 rxn)
- SMRTbell Clean Up Columns v2
  - New version

- Streamlined workflow for Iso-Seq on Sequel System
- Diffusion loading of long insert libraries on Sequel System
- Microbial multiplexing on Sequel System
- New barcoding schema
- Target enrichment without amplification using CRISPR/cas9

NEW BARCODING SCHEMA FOR SEQUEL AND RS II SYSTEMS

- Have developed ~450 barcodes
- Prioritized top 384
- Original list:
  - lbc0001, …, lbc0384
- New list:
  - bc1001, …, bc1384

- Barcode lbc0001 is NOT bc1001

- Rearranged sequence positive predictive value
- New order is compatible for both PacBio RS II and Sequel Systems

MICROBIAL MULTIPLEXING

- Both laboratory protocol and analysis procedures
- Examples on Sequel System
  - 8-plex with E. coli (5 Mb)
  - 12-plex with B. sub (4 Mb)
  - 16-plex with H.pylori (1.6 Mb)
- 8-plex dataset is available

- Level of multiplex is dependent on:
  - Completeness
  - Genomic size
  - Genomic complexity

- Receiving first data from first customer experiences

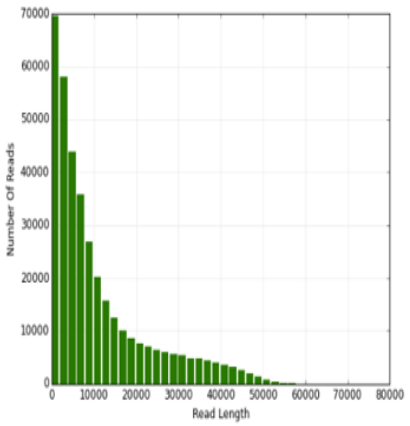
CUSTOMER EXAMPLE OF MICROBIAL MULTIPLEXING

- Goal:
  - Comparable performance to PacBio RS II
  - Low-plex (e.g., 4-plex)
  - Salmonella and Listeria
- gTube shear to 8-10 kb
- Determine the effect of pre-extension

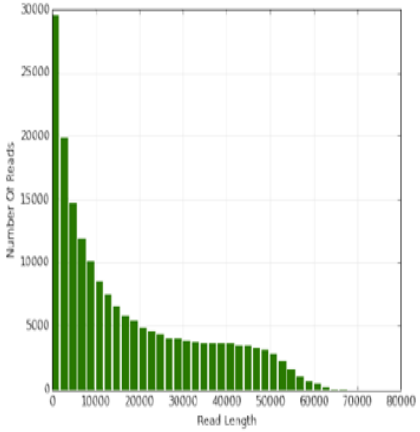
PRIMARY METRICS OF 4-PLEX

run nb	library		loading [pM]	Pre extension (mins)	Movie Length (mins)	Total Bases (GB)	Polymerase Reads (bp)	Longest subread (bp)	Productivity		
	size [bp]	Bp cut [Kb]							Empty (P0)	Productive (P1)	Other (P2)
Cell1	6200	4	10	0	600	4.15	11159	6104	30.20%	35.90%	33.90%
Cell2	6200	4	10	120	600	3.32	17712	6681	54%	18%	28%
Cell3	6200	4	15	120	600	4.89	16094	6545	39.01%	29.84%	32.88%
Cell4	6200	4	15	180	600	1.3	18296	6599	70.10%	6.90%	23.00%

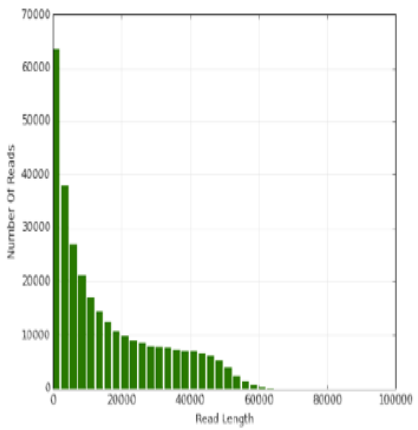
Cell1



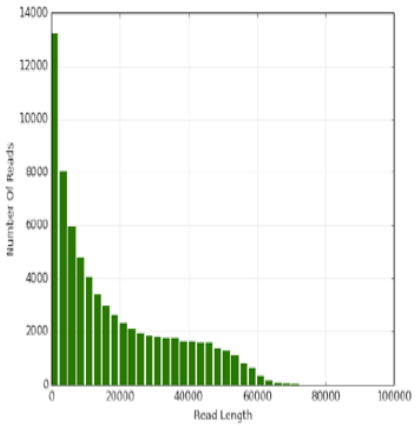
Cell2



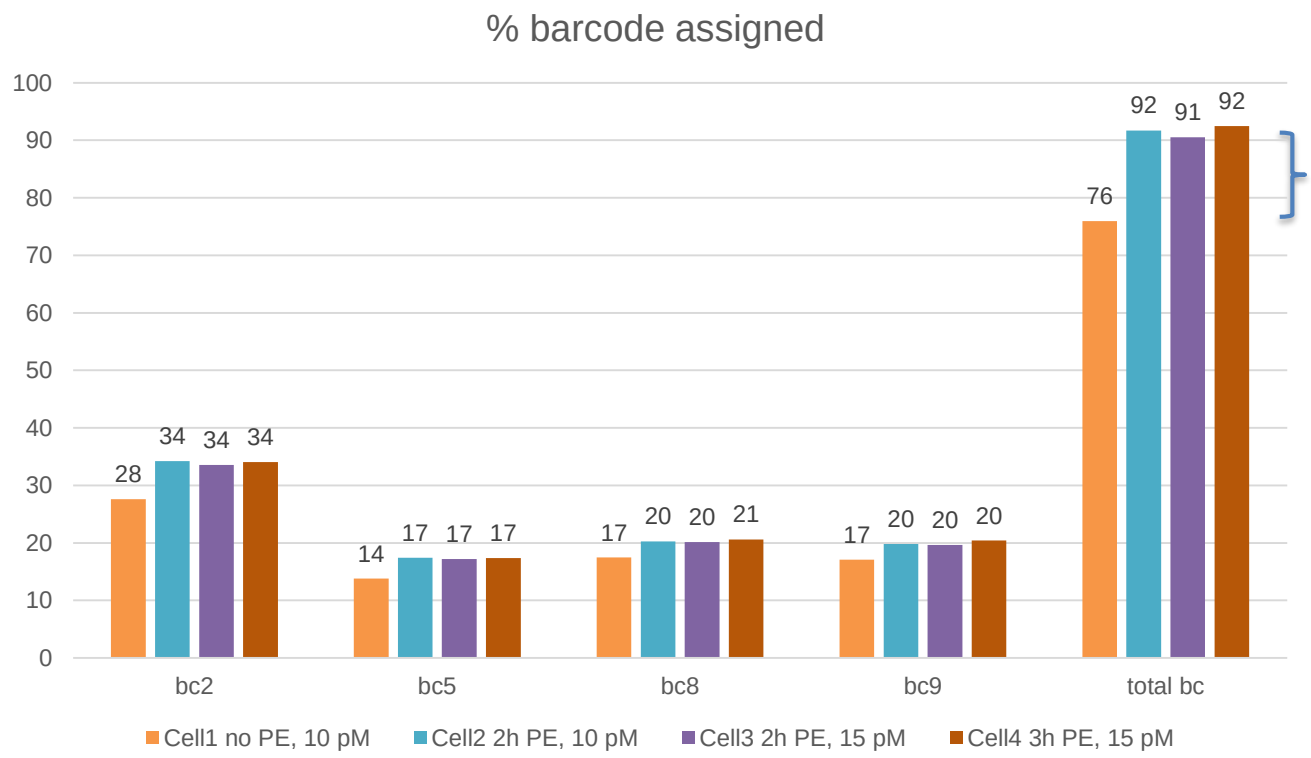
Cell3



Cell4



PRE-EXTENSION DRIVE BARCODE IDENTIFICATION > 90%

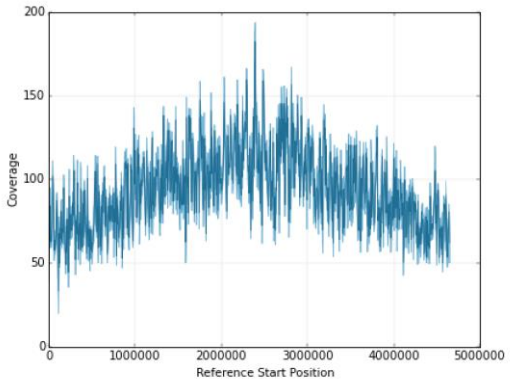


Increase of 15% of reads assigned with Pre Extension

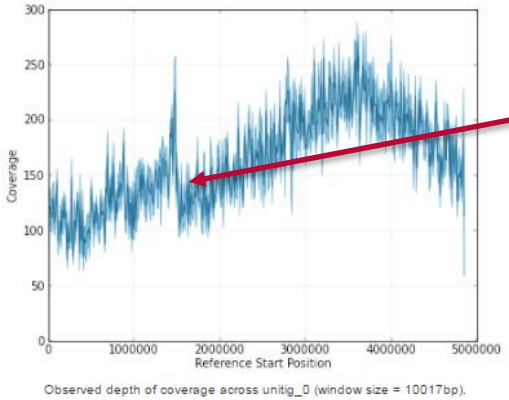
COMPARABLE EXPERIENCE WITH SEQUEL

1 Sequel SMRT Cell 1M compared to 4 PacBio RS II SMRT Cells

Strain	Barcode	Coverage - Sequel	Number of Contigs - Sequel	Coverage - RS II	Number of Contigs – RS II
Salmonella	2	193	4	164	2
Salmonella	5	95	10	192	4
Listeria	8	206	1	301	1



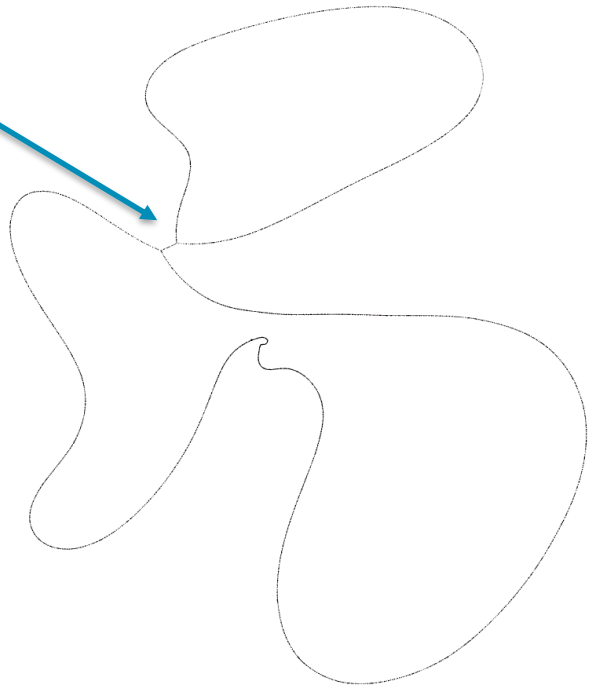
PacBio RS II



Discontinuity in the expected sigmoidal coverage

RS II - BASED ASSEMBLY OVERLAP GRAPH

Collapsed repeat in the graph, correct answer given the data is to generate three contigs



Plasmid

CUSTOMER'S FEEDBACK AND NEXT STEPS

- Sequel instrument performance meet expectations
- Salmonella has many phage insertion points
- HGAP4 produced more accurate assemblies compared to HGAP3

Further Optimization

- Increasing shear size to 8-10 kb inserts
- Assembly metrics

Effects of ‘Overloading’

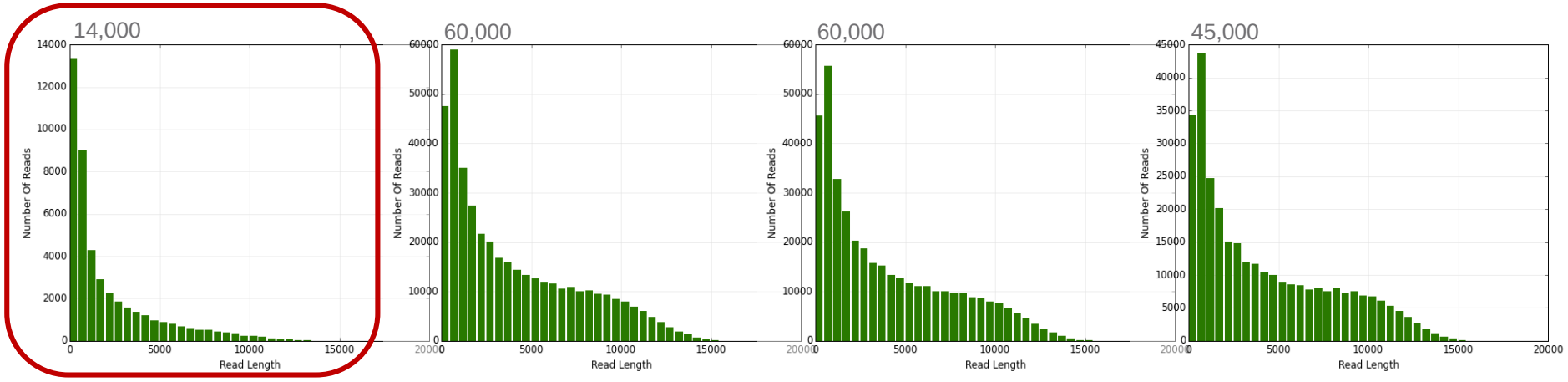
LOADING TITRATION, 1.6KB AMPLICON, V1.2.1 SEQUENCING KIT ON SEQUEL

Dashboard Results

Loading	Movie Length (mins)	Total Bases (GB)	Polymerase Reads		Reads Of Insert		Productivity		
			Length	Quality	Length	Quality	Empty (P0)	Productive (P1)	Other (P2)
350 pM	360	0.29	6504	100.00	1595	0.00	456846 (44%)	44968 (4%)	534941 (52%)
35 pM	360	5.04	12217	100.00	1715	0.00	586312 (57%)	412577 (40%)	37886 (4%)
25 pM	360	4.74	12174	100.00	1697	0.00	620121 (60%)	389036 (38%)	27630 (3%)
20 pM	360	3.96	12834	100.00	1704	0.00	719111 (69%)	308769 (30%)	8888 (1%)

* In SMRT Link Data Management

Polymerase Read Length



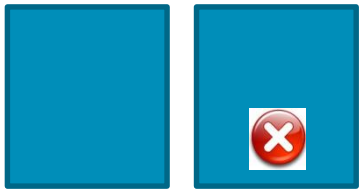
350 pM | 35 pM | 25 pM | 20 pM

LOADING REFERS TO # OF ACTIVE POLYMERASES IN A ZMW

Three loading classes: Empty, Single, Multi

Similar but not totally equivalent to P0, P1 and P2

“Empty”



Contains no active pols

May contain inactive pols

“Single load”

Contains one active pol

May contain inactive pols

“Multi load”

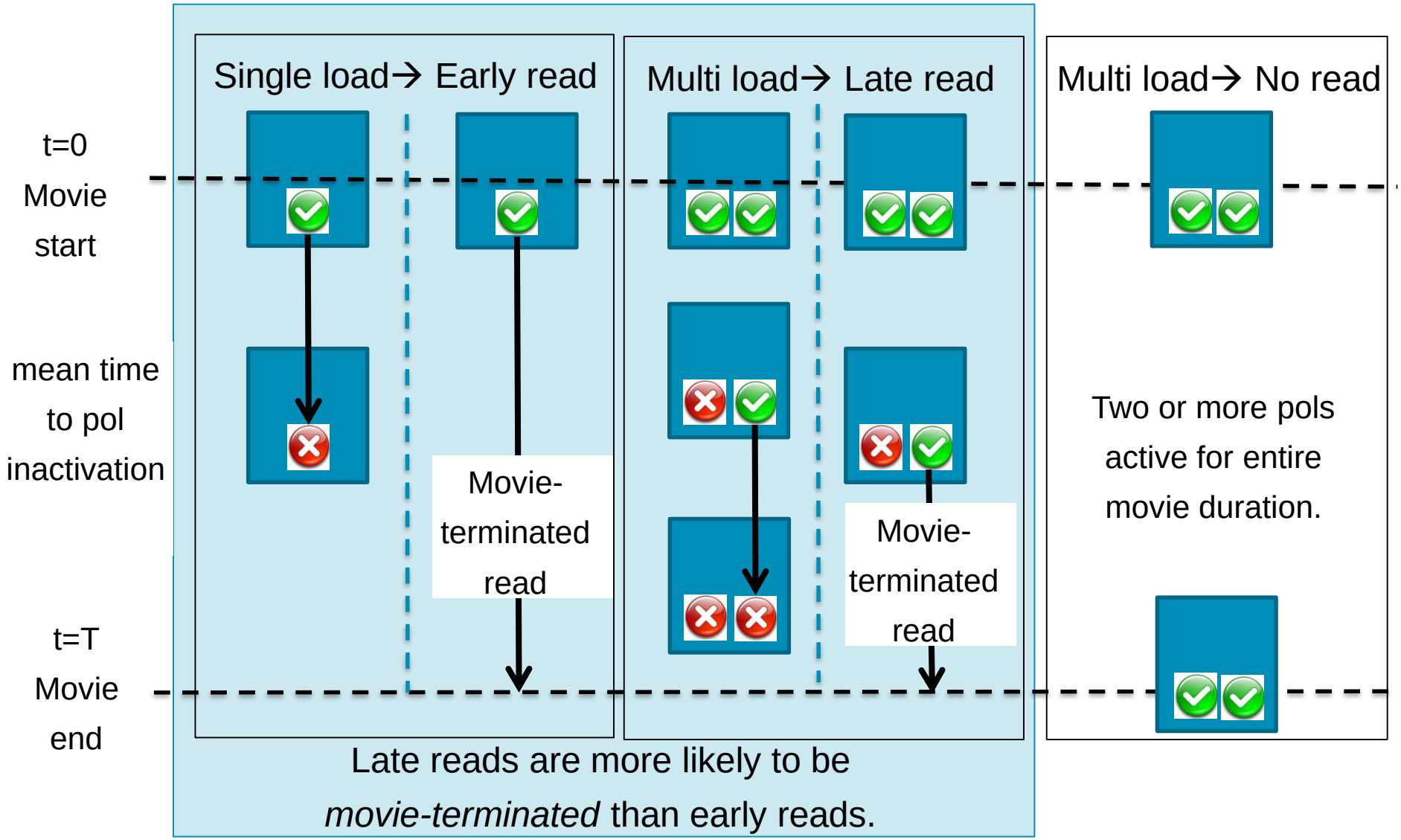
Contains >1 active pol

May contain inactive pols

- Active polymerase
- Inactive polymerase

Inactive pols have no effect on nReads, read length, but do reduce accuracy.

SINGLE LOADS → EARLY READS
MULTI LOADS → LATE READS OR NO READ



THERE ARE THREE COSTS TO OVERLOADING

Benefits of Overloading:

- Throughput increases
- Number of reads increases

Costs of Overloading:

- Mean (polymerase) read length decreases
- Mean (longest) subread length decreases
- Accuracy decreases

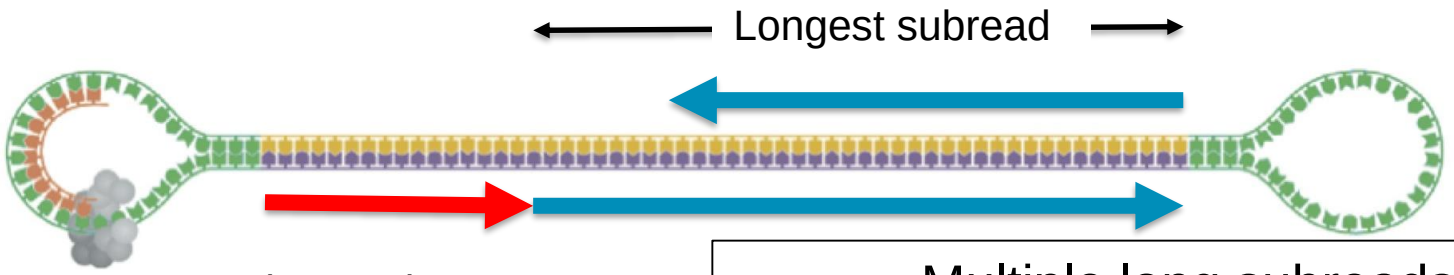
THERE ARE THREE COSTS TO OVERLOADING

Benefits of Overloading:

- Throughput increases
- Number of reads increases

Costs of Overloading:

- Mean (polymerase) read length decreases
- **Mean (longest) subread length decreases**
- Accuracy decreases



Lost bases due to
2 active polymerases

Multiple long subreads
spanning each long genomic repeat
→ contiguous assemblies

THERE ARE THREE COSTS TO OVERLOADING

Benefits of Overloading:

- Throughput increases
- Number of reads increases

Costs of Overloading:

- Mean (polymerase) read length decreases
- Mean (longest) subread length decreases
- **Accuracy decreases**

Finding: Accuracy drops by one percentage point (e.g. 87% → 86%) for each additional inactive polymerase present in a ZMW during sequencing.

PACBIO®

www.pacb.com