Setting Expectations and Sifting Through the Stinkers: A Core Perspective

Shana McDevitt
Managing Director, QB3 Genomics
Director, VJC Genomics Sequencing Lab
UC Berkeley

California Institute for Quantitative Biosciences

WAKE UP!!!!!

A Walk Down Memory Lane
Featuring: Candy and Genomes

The Penis Worm…

Setting Expectations: PacBio's First Pinterest Fail

Expectation

Reality

Setting Expectations:
Submission Requirements are not Suggestions

Three main components (in order of importance)

1. Integrity (size)
2. Purity
3. Total Yield

Set them up for success or don't take their money!!

Don’t Make Paul’s Family Sad!!!

We must be the ambassadors of the brand!

Setting Expectations: Sample Integrity

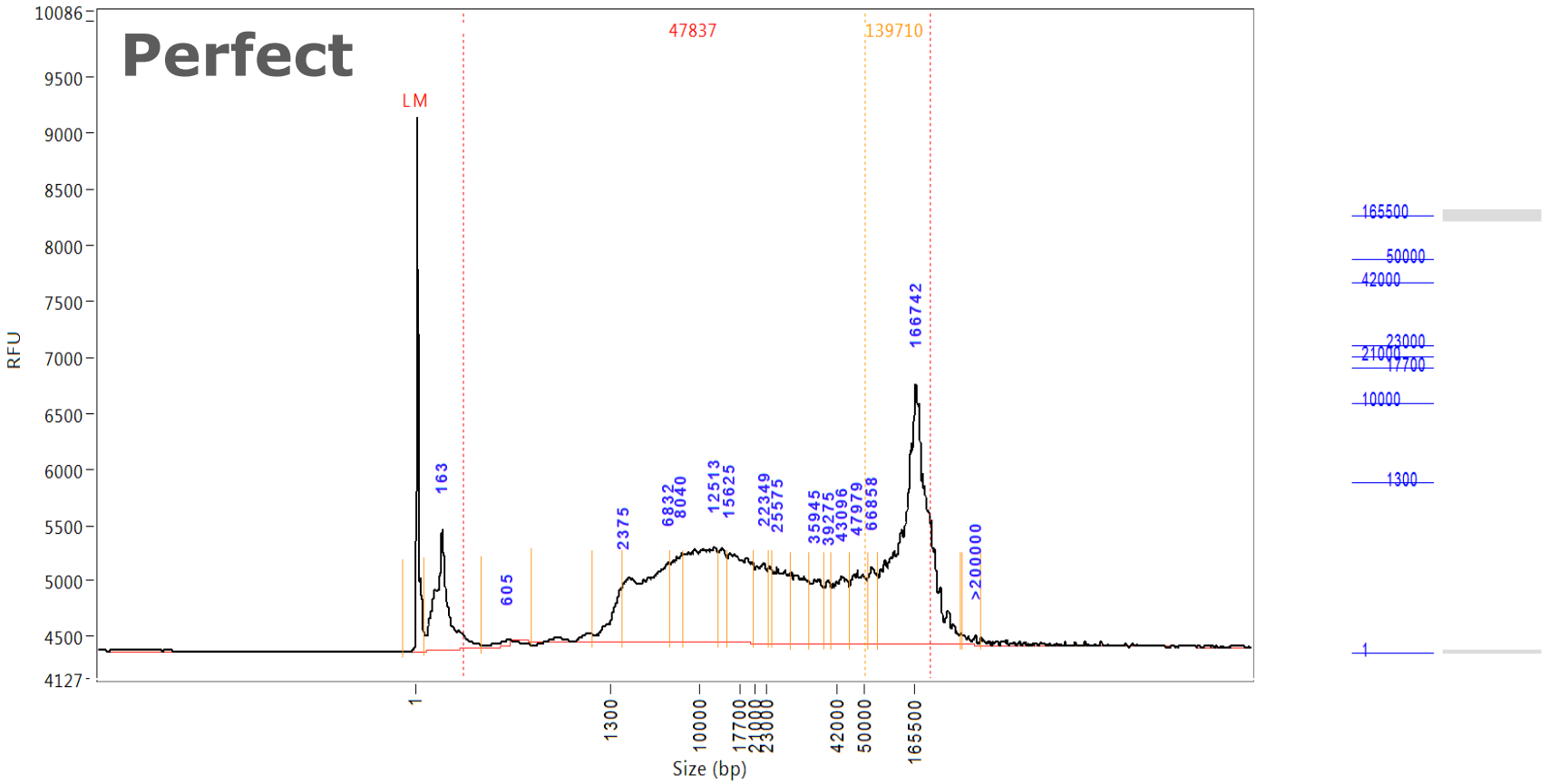
Why does it matter?

- If you don't have long DNA you can't do long read sequencing (yes even for HiFi)
- If you don't have nick free DNA you can't do single molecule sequencing

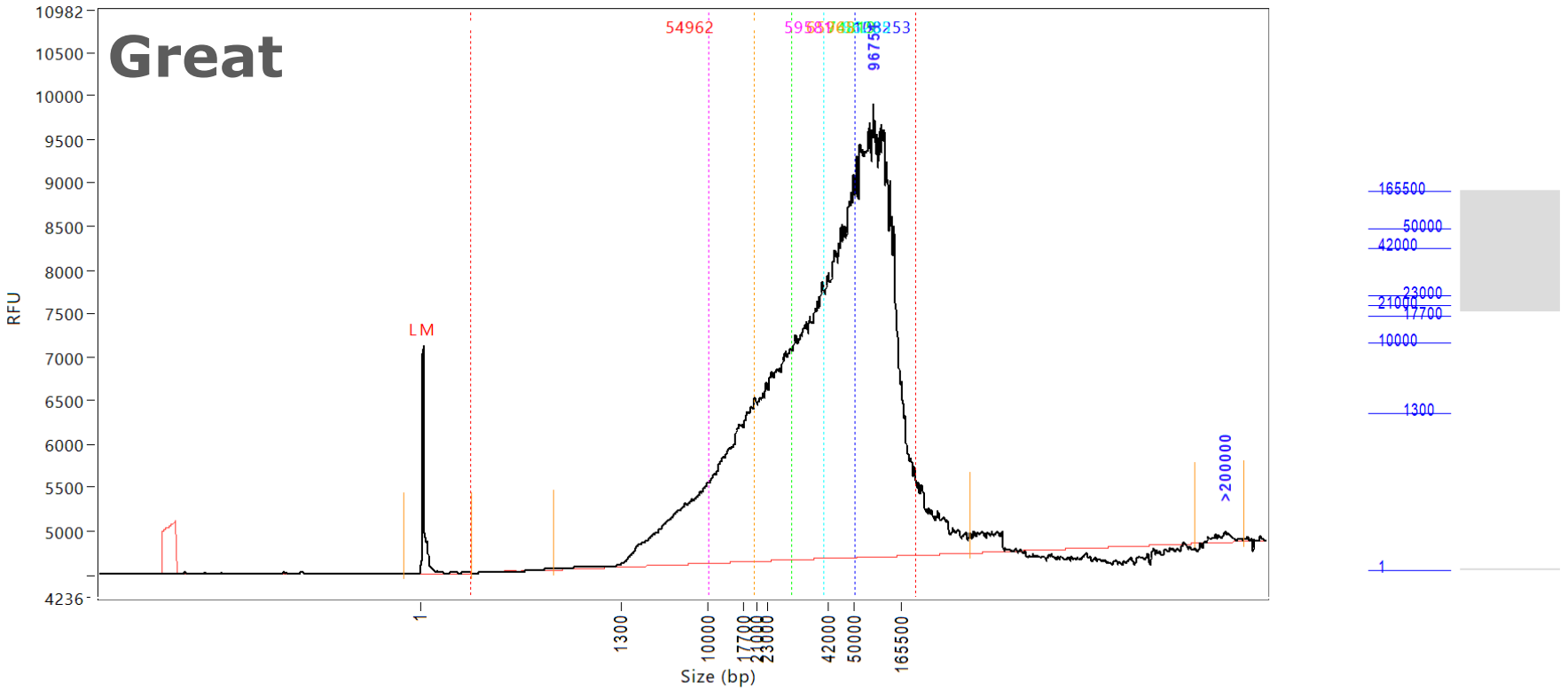
Requirements:

- Perfect HMW: 100kb+, sharp peak
- Required Minimum:50Kb+ average size without a trailing tail
  - Measured on PFGE or Femto Pulse
  - Regular gels need not apply

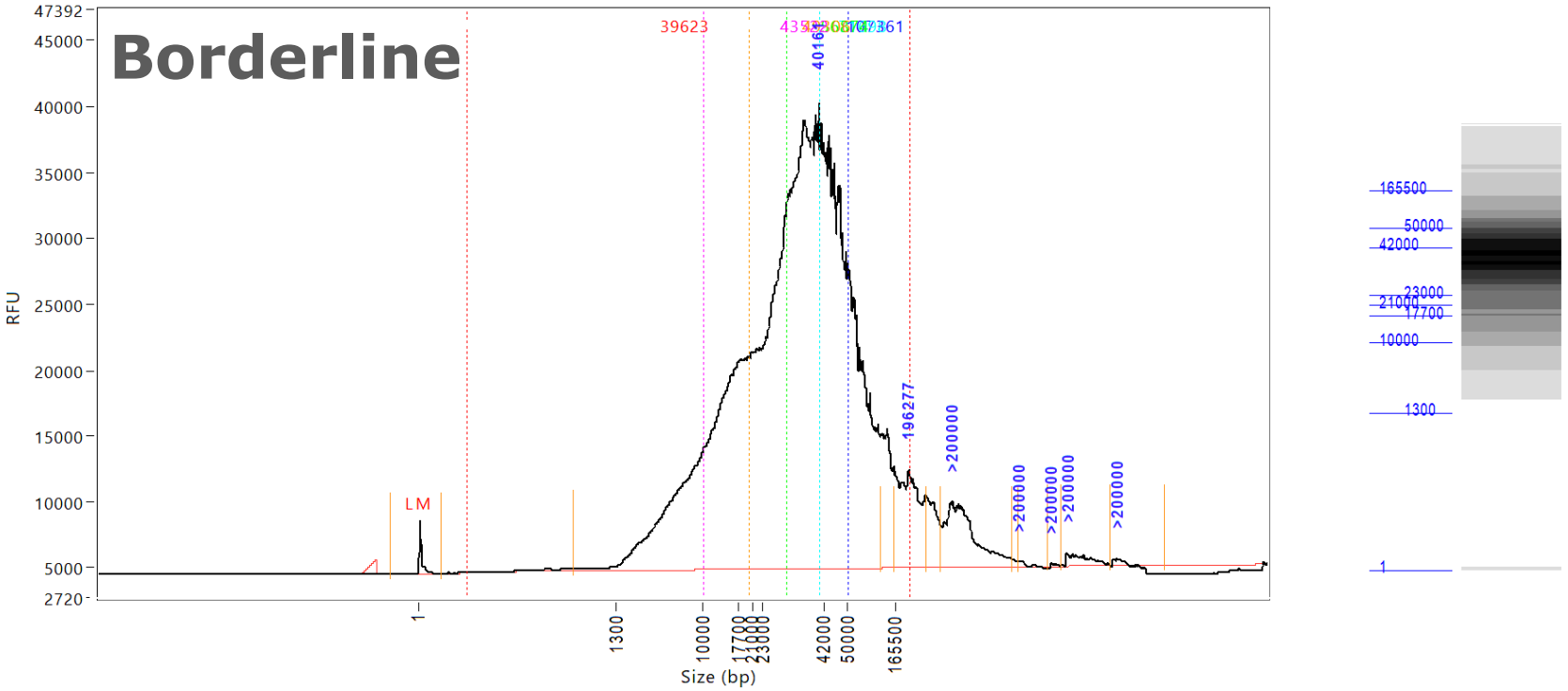
Setting Expectations: Sample Integrity



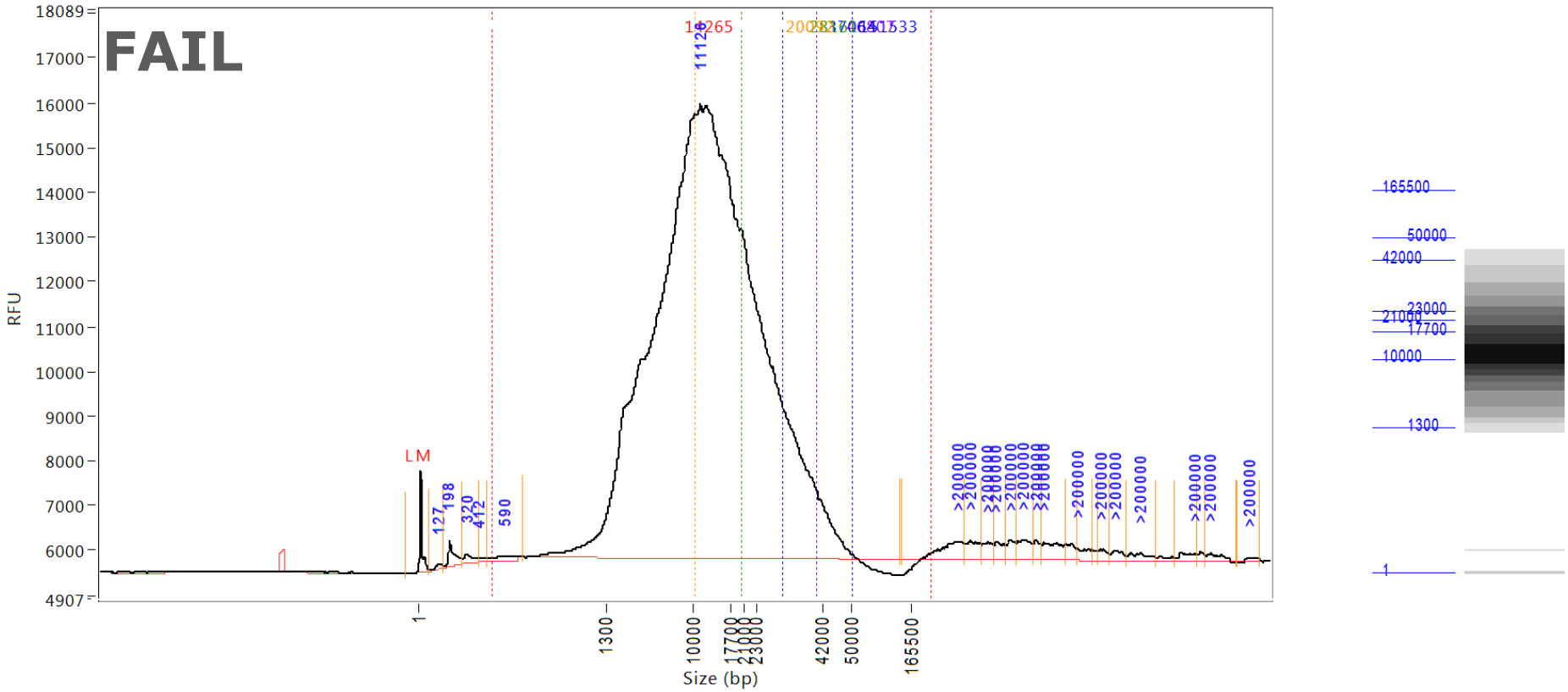
Setting Expectations: Sample Integrity



Setting Expectations: Sample Integrity



Setting Expectations: Sample Integrity



Smear Analysis	500 bp to 200000 bp	1.6768 ng/ul	95.2 %Total	0.1935 nmole/L	14265 Avg. Size (b.p.)	69.54 %CV
	10000 bp to 200000 bp	0.9656 ng/ul	54.8 %Total	0.0791 nmole/L	20092 Avg. Size (b.p.)	46.44 %CV
	20000 bp to 200000 bp	0.3954 ng/ul	22.5 %Total	0.0231 nmole/L	28161 Avg. Size (b.p.)	33.99 %CV
	30000 bp to 200000 bp	0.1288 ng/ul	7.3 %Total	0.0057 nmole/L	37064 Avg. Size (b.p.)	32.79 %CV
	40000 bp to 200000 bp	0.0285 ng/ul	1.6 %Total	0.0010 nmole/L	46807 Avg. Size (b.p.)	48.67 %CV
	50000 bp to 200000 bp	0.0009 ng/ul	0.1 %Total	0.0000 nmole/L	151533 Avg. Size (b.p.)	43.25 %CV

Setting Expectations: Purity

This is the biggest source of failure!!

- Contaminants can nick the gDNA and inhibit polymerase binding and processivity
- The longer they exist the more they degrade, get them out quick!

Requirements:

- 260/280 = 1.8-2 and **260/230 = 2 – 2.2**
  - No exceptions!
  - Never trust a users QC

Setting Expectations: Sample Yield

CLR

Need to start with 10ug of gDNA to fully load an 8M SMRTcell

- Assumes a clean, shear, and size selection may be needed
- Can go in with less but will not be able to do some or all techniques required in library prep

HiFI

For actual protocol, 15-20ug for 1 sample per SMRTcell, less if multiplexing but not by much.

- Very low final recoveries after exo steps even with HMW

For Express version (I cheat), 10ug (Our own protocol, not PB)

- Fewer laps around the SMRTbell…but not soon!

Order of Operations for Failing QC

QC Threshold 1: DNA is not HMW

Feedback

Hi Cheryl, Your PacBio submission shows evidence of degradation based on it's low size distribution, if I proceed with this sample you will most likely get short read lengths and a failure to generate ccs data. I would recommend you…

- Redo your extraction (review protocol with them)
- Review shipping and post extraction storage protocols

Extractions are Everything

- If you want to martyr yourself for science do extractions for users!
- Charge an arm and a leg! It's a time sink
- Get LOTS of tissue
  - Get tissue with an expectation of having easily accessible DNA/RNA content
  - Unknown Bacteria or complex metagenomics samples are hard

<u>Cheap:</u>

Our Protocol: Qiagen Buffers (with all of the ProtK) & PCI extraction

Qiagen Genomic Tips (~50kb max) & MagAttract (HMW but lower yield)

<u>Not as cheap but very good:</u>

Plants: Circulomix Big Plant Kit (requires nuclei isolation)

Bionano: High yield, very HMW kits

Tissue Storage and Shipping

- Flash freeze and ship on dry ice (must stay frozen)
  - Store in -80C at all other times up until extraction
- Ethanol is not ideal
  - I have literally never had a good extraction from an ethanol stored sample
- RNA Later is for RNA!
  - Use Zymo DNA/RNA Shield instead
- Can’t ship on dry ice?
  - Send gDNA in an agarose plug (Bionano protocol available)
  - gDNA Stabilization matrices do not work in my experience

Sample Shipping: Idealistic

- Avoid freeze thaw
- If their sample is at 4C ship at 4C overnight
  - QC aliquot immediately and store at -20 until ready to proceed if not immediately which is ideal
  - If it will be awhile ship on dry ice
    - Only thaw and proceed with QC when ready to do the full prep
- Recommend the user make a QC aliquot when they do their extractions. Then you only have to thaw the aliquot and can do QC immediately without jeopardizing the sample.

Sample Shipping: Reality

- gDNA will degrade at 4C if contaminants are present
- Always have the user immediately freeze their extraction after they do their internal QC (qubit and nanodrop)
- Contaminants are almost always present, we clean everything at least once
  - Accept a freeze thaw to potentially salvage a sample
- QC aliquot becomes very important

Order of Operations for Failing QC

QC Threshold 2: 260/230 is out of range

Feedback

Hi Cheryl, Your PacBio submission shows evidence of contaminants in your sample, left uncleaned your DNA will degrade and residual contaminants may inhibit sequencing. I would recommend you…

- Redo your extraction (review protocol with them)
- Charge them to clean the sample (High Salt PCI)
  - Set expectations, really poor sample may be too far gone
  - Some tissue types need 2x cleans
  - You may lose up to 50% yield each clean
- This is often an explanation for poor size integrity as well

Order of Operations for Failing QC

- QC Threshold 3: Yield

If a sample passed the other 2 thresholds we can generally make it work but there are some ifs

Scenario 1: very HMW, clean

- Loss to consider: sample will need to be sheared and size selected, loss will be significant

Scenario 2: 50kb average, clean with a degraded tail

- Loss to consider: sample will need to be size selected, loss will be less significant but still ~40% minimum (5ug ideal)

Scenario 3: 50kb average, clean, no tail

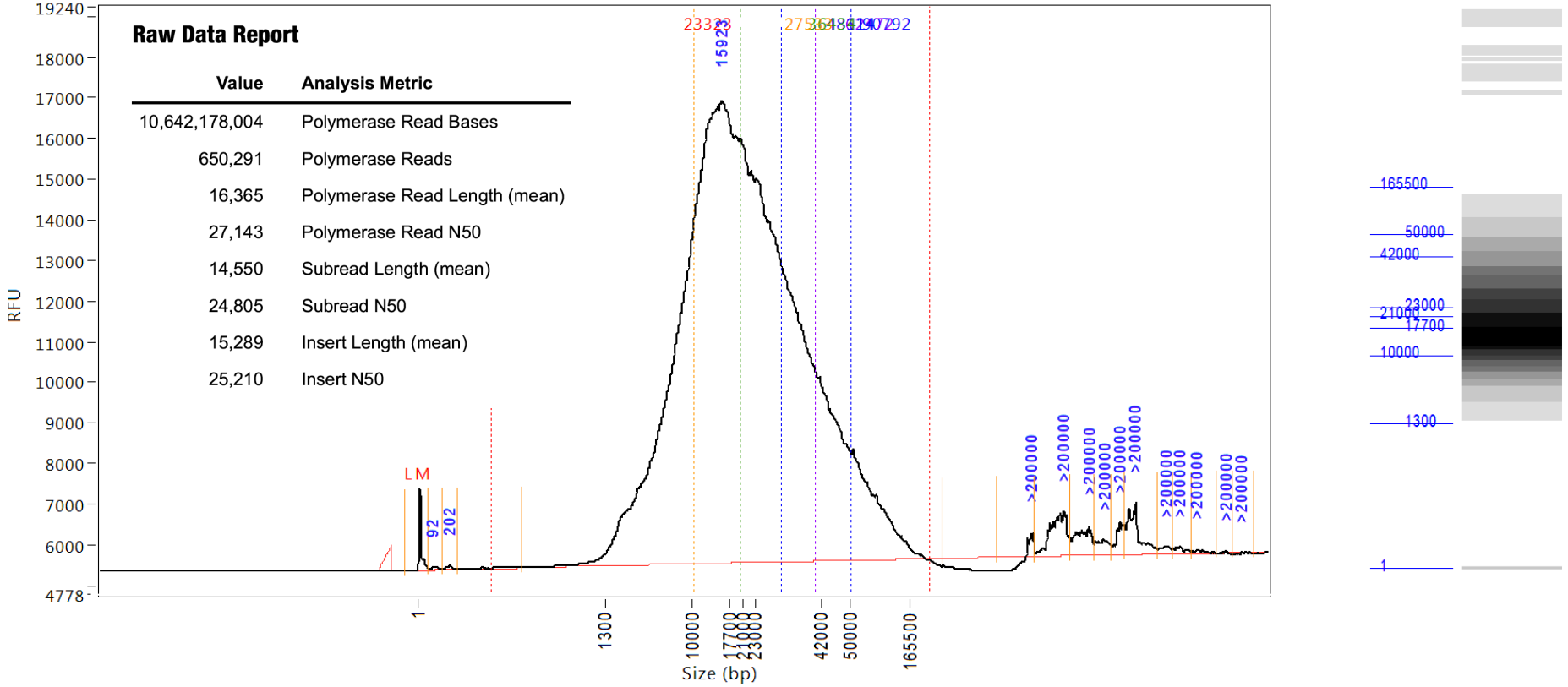
This is the only scenario where we can go low input! Minimal loss with express protocol, must be CLR

Some Tricks

- Have a lot of clean degraded sample with some HMW content?
  - Run the sample on the BluPippin to collect only 40-50kb+, pool the lane elutions, Express protocol, no size selection again.
    - Too expensive to do it in reverse, but if $$ doesn't matter that's a better plan.
    - Circulomix Short Read Eliminator beads could also be used in place of BluPippin

- Expectations:
  - You will see shorter read lengths as even the HMW fragments will have started to degrade
  - Only do if precious samples or don't expect improvements upon re-extraction

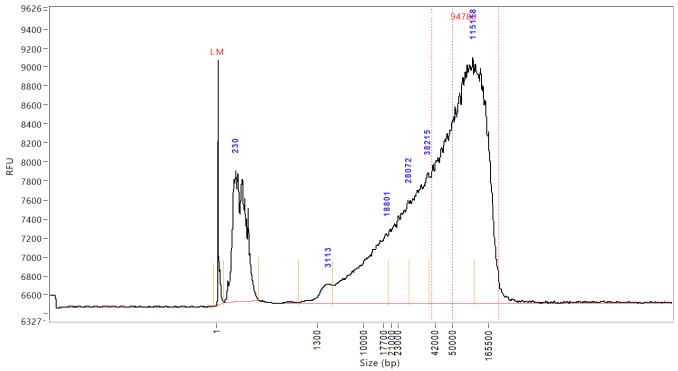
Some Tricks:Example



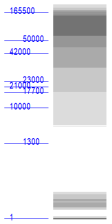
8 BluPippin lanes cut at 40kb into Express Prep
16 Sequel 1M cells

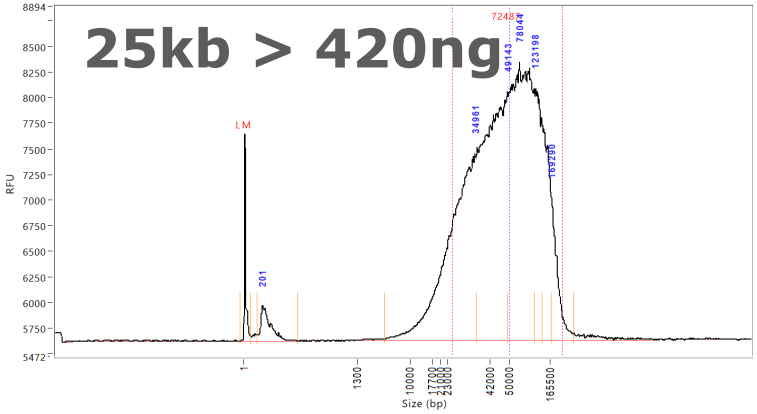
Short Read Eliminator in lieu of BluPippin

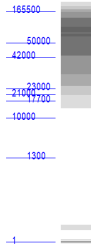
- Kits for 10kb, 25Kb, and 40Kb progressive cuts
- 3ug input

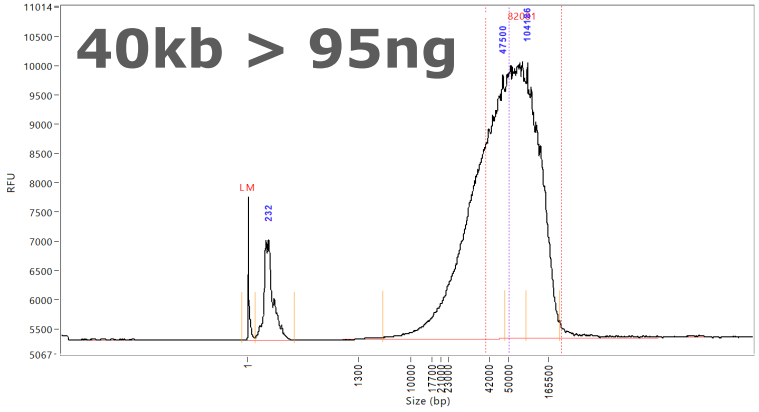


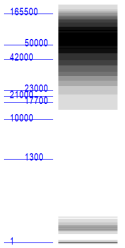
Original



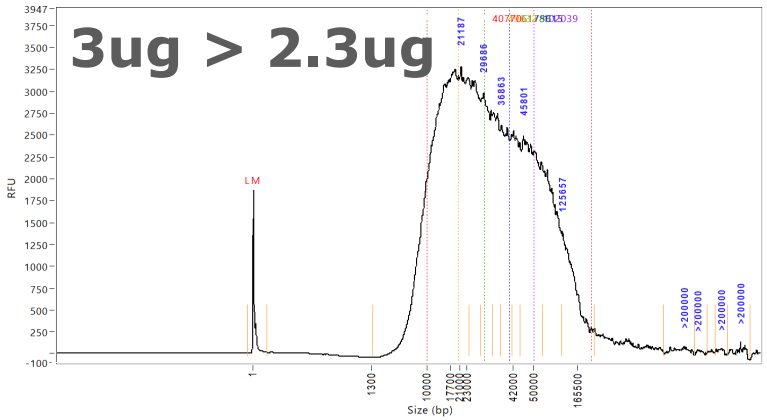


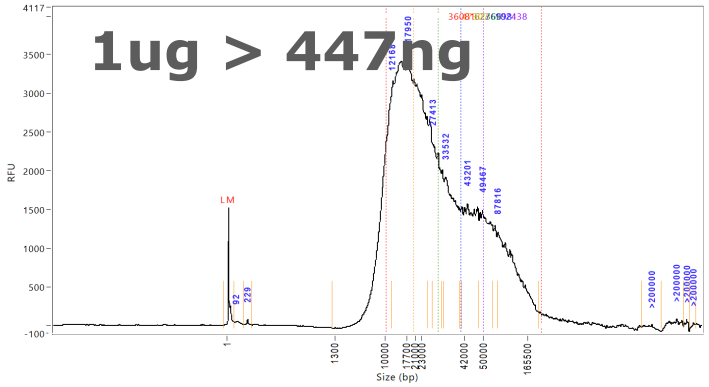


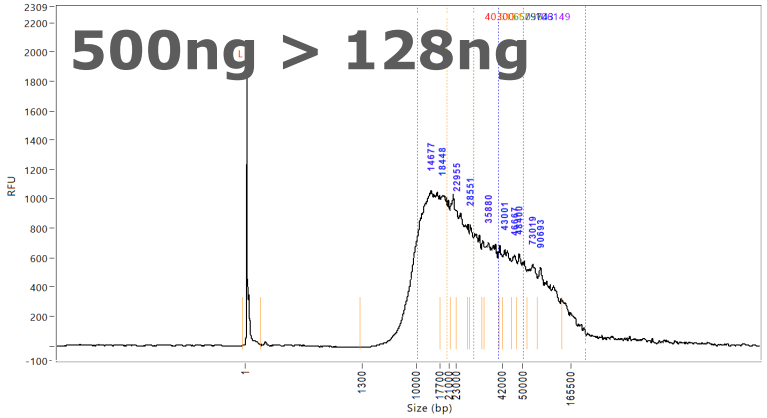


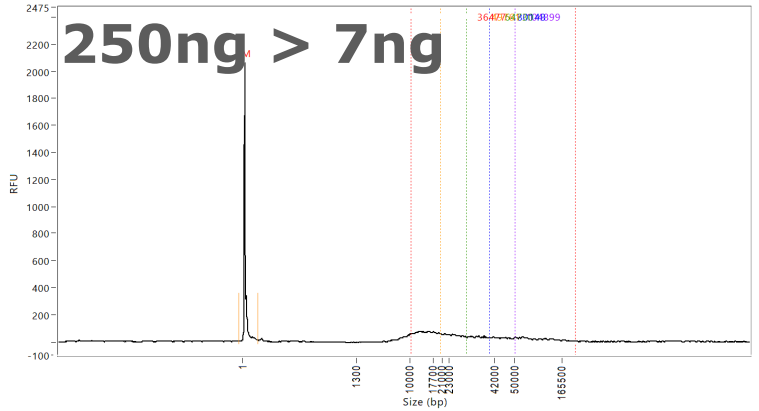


Short Read Eliminator Not for Low Input









In Summary: Set Proper Expectations

Everything is a maybe until you have in-house QC

Thanks to my Peeps

Rob Munch

←He does the actual work

Kristin Mars

Josiah Wilcots

Nick Sisneros

Paul Kotturi

qb3
ucb·ucsc·ucsf

THANK YOU!!!!

shana.mcdevitt@berkeley.edu