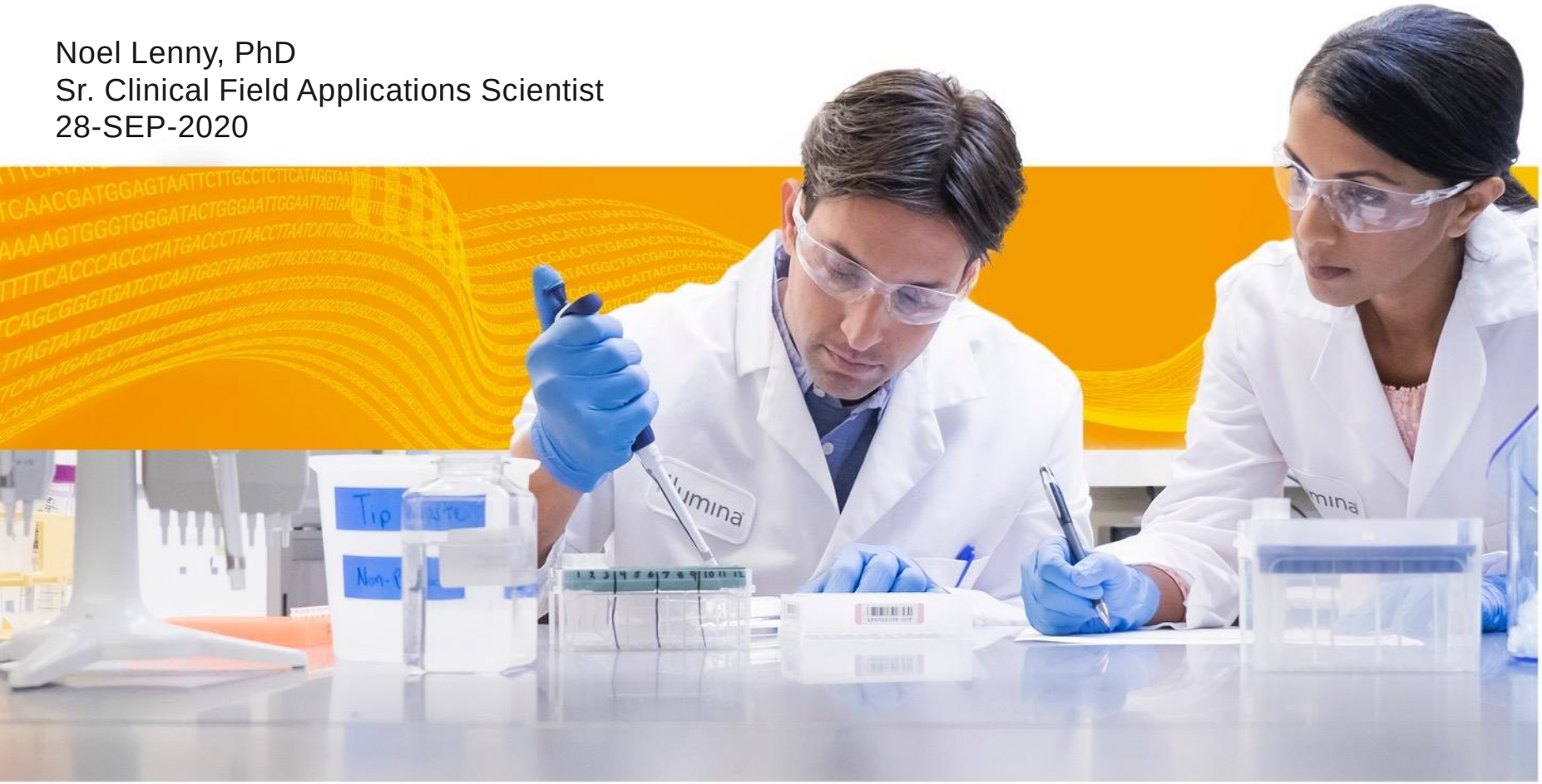
Best Practices to pass Library QC

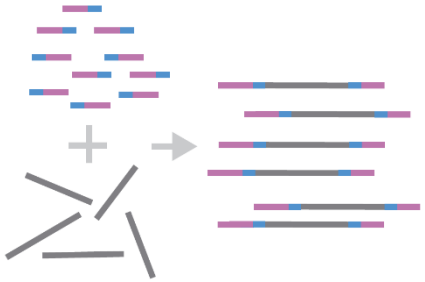
Noel Lenny, PhD
Sr. Clinical Field Applications Scientist
28-SEP-2020



illumina®

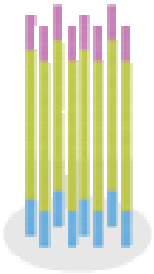
Library QC

Library Preparation



Modify nucleic acid to proper insert size & add adapters

Cluster Generation



Attach library to the flow cell & clone to increase intensity

Sequencing by Synthesis

Gather intensity information at each SBS cycle

Data Analysis

Convert intensity into base calls & Q-scores; write reads to FASTQ

illumina®

Importance of library QC

Garbage in

Garbage out

Bad Sample → Bad Library

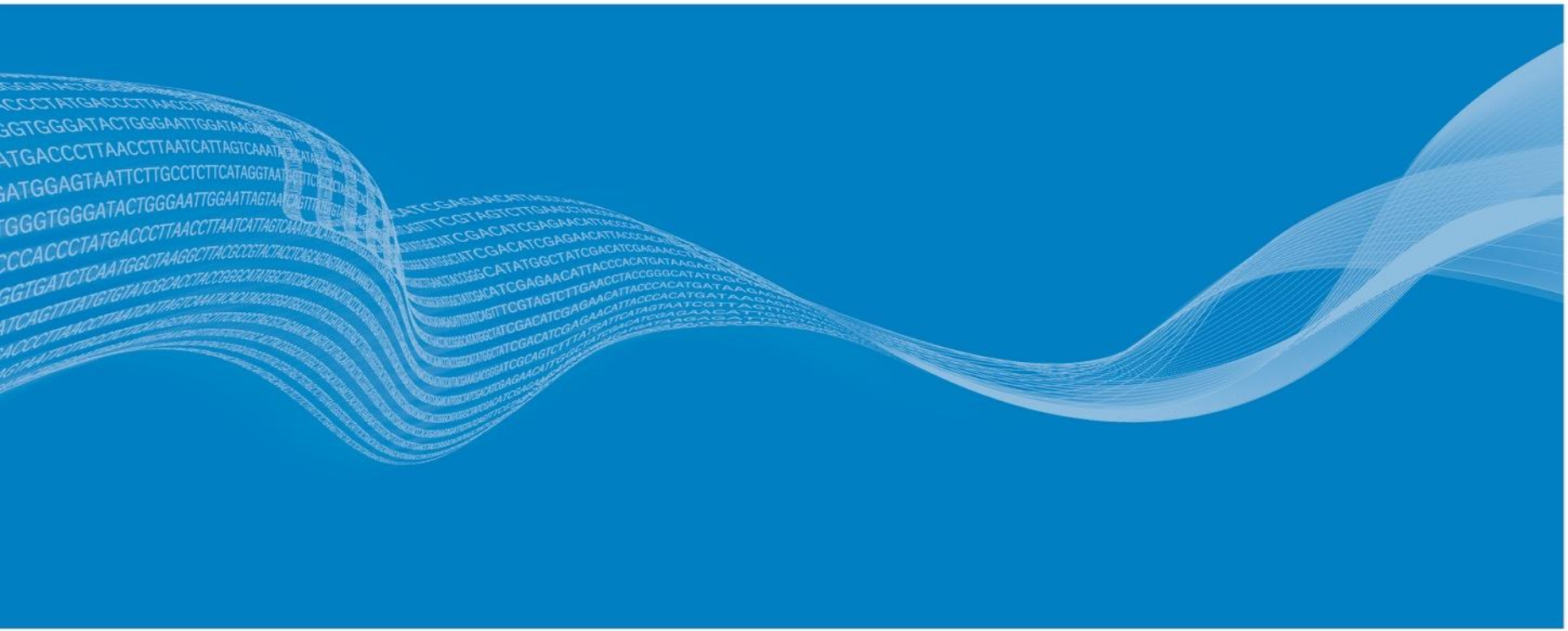
Bad Library → Bad Sequencing Data

illumina®

Importance of library QC

- **Library size**
  - Inserts too short or too long will lead to improper coverage
  - Insert size too short will lead to sequencing into the adapter
- **Contaminants**
  - Negatively impact data quantity (yield in Mb) and quality (Q-scores)
  - Can lead to index mis-assignment on some sequencing platforms
- **Library yield**
  - Sequencing data quality depends on the loading concentration
  - Pooling requires normalization to achieve equal amount of data/ sample
  - Too little or too much library: reflects input quality issue or library prep biases that also impact data coverage (random loss of input RNA)

Library quantification



illumina®

Quantification Methods of RNA-Seq Libraries

Fluorometric dsDNA assay

- Specifically detects double-stranded DNA
- Does not discriminate incomplete TruSeq libraries

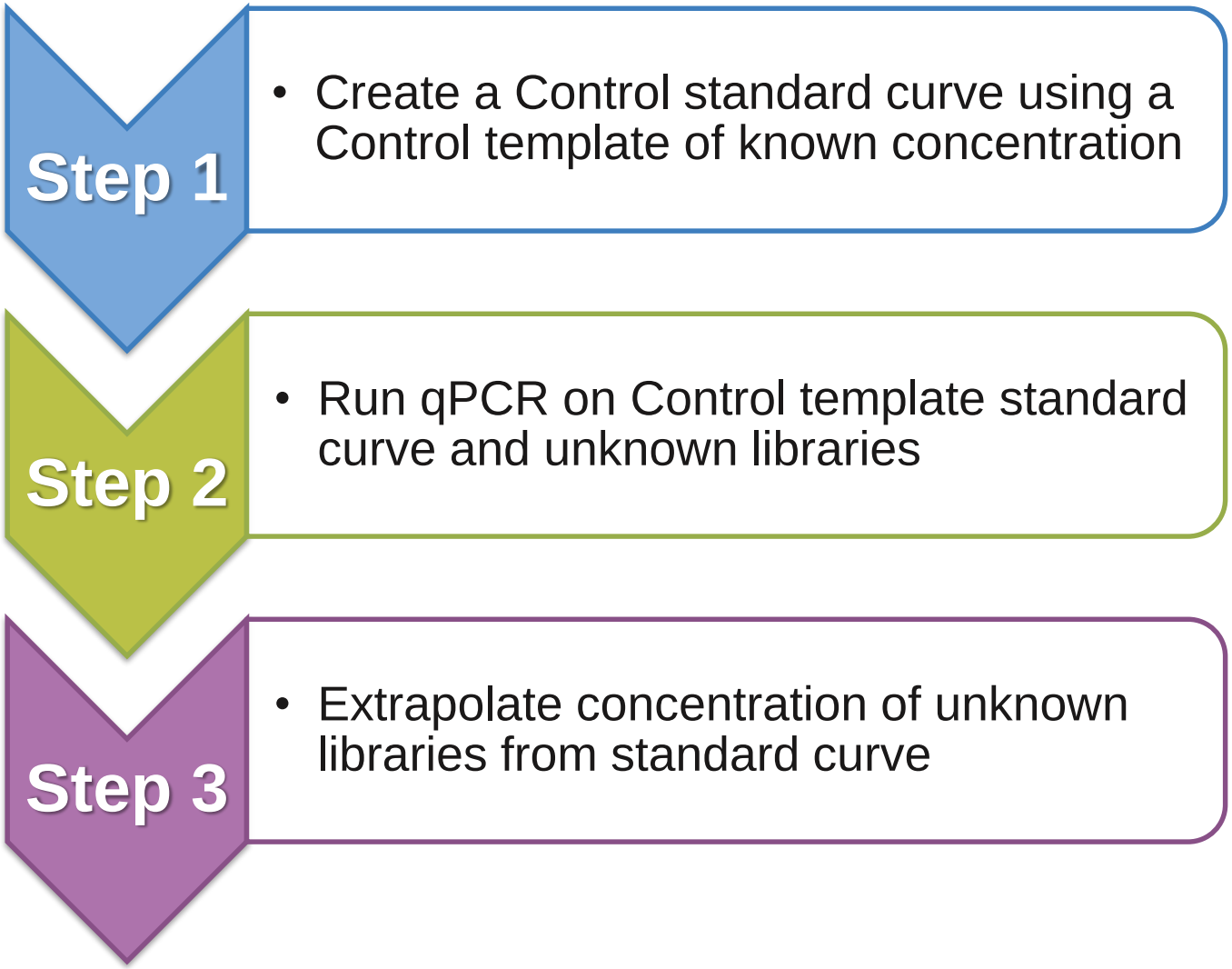
qPCR

- Specifically measures full-length libraries
- Detection very sensitive

UV spectrophotometer (Nanodrop)

- Non-specific nucleic acid detection
- Contaminants elevate values

Library qPCR workflow



illumina®

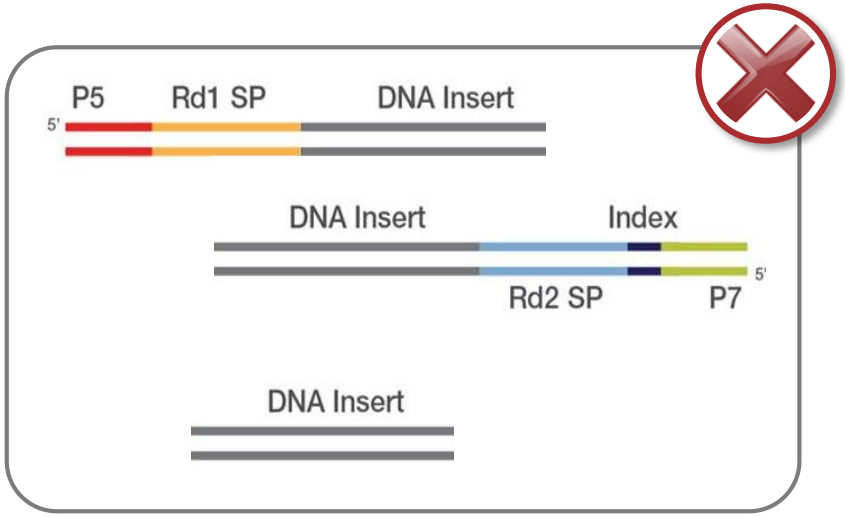
Why is qPCR the most accurate quantification?

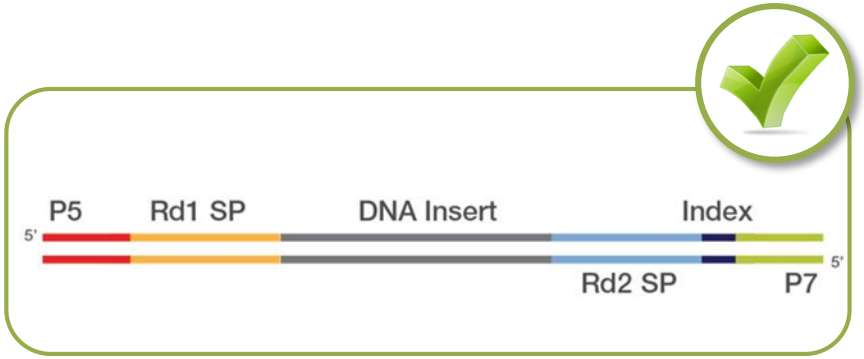
qPCR

Designed to quantify only cluster-forming fragments in the samples

Uses primers complementary to adapters to mimic amplification on the flow cell

Only amplifies and quantifies library fragments with proper adapters at both ends





illumina®

Library Quantification Resource

KAPA Biosystems Library Quantification Kit

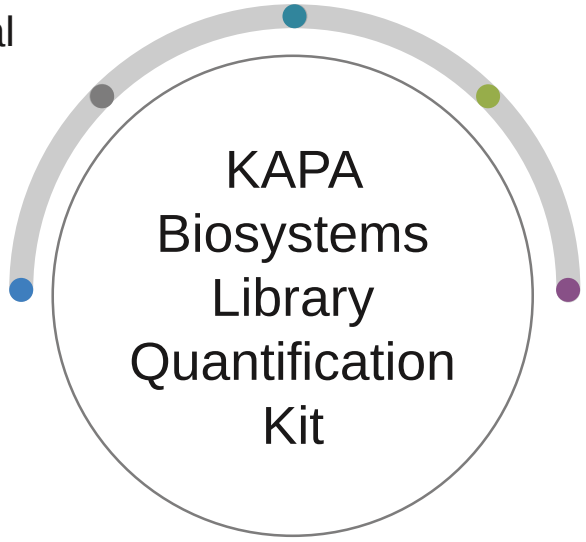
Best Practice

Use control library that has an insert size and GC content similar to your sample library

6 DNA Standards

Primers work with Illumina Universal Adapters

No dilutions needed



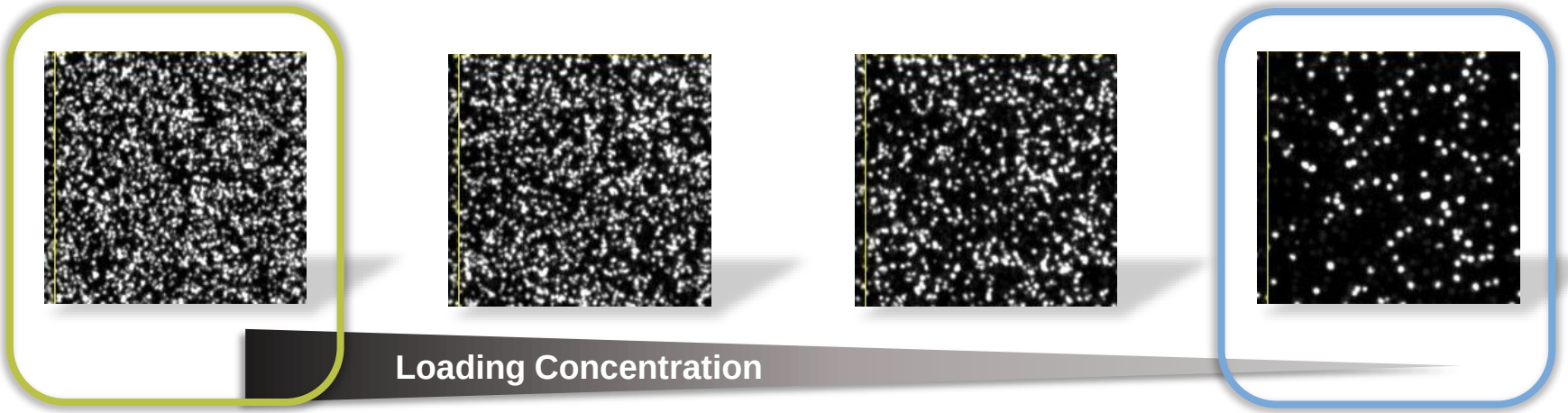
Designed specifically for libraries prepared with Illumina library prep assays

Reliable, accurate standard curve

illumina®

Maximize Data Quality and Quantity

Optimized flow cell clustering determines data quality and overall data yield



Overclustering can result in:

- Loss of data quality and data output
- Loss of focus
- Reduced base calls and Q30 scores
- Complete run failure

Underclusterring can result in:

- Loss of time and money
- Loss of focus
- Complete run failure

illumina®

Accurate Quantification Is Critical When Multiplexing

Calculated concentration is 10X higher for one library in pool

Sample	Expected Output	Actual Output
1	16%	66%
2	16%	6%
3	16%	6%
4	16%	6%
5	16%	6%
6	16%	6%

Sample	Expected Output	Actual Output
1	16%	20%
2	16%	20%
3	16%	20%
4	16%	20%
5	16%	20%
6	16%	2%

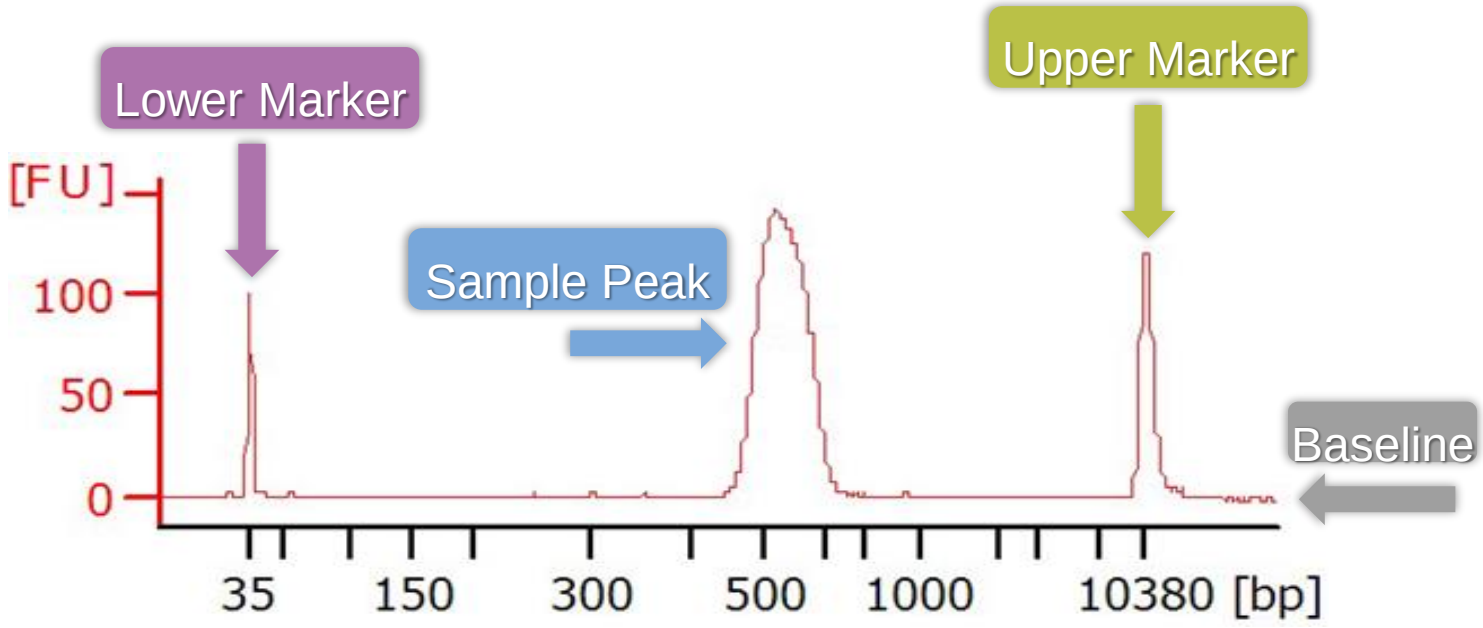
Calculated concentration is 10X lower for one library in pool

illumina®

Library Size

illumina®

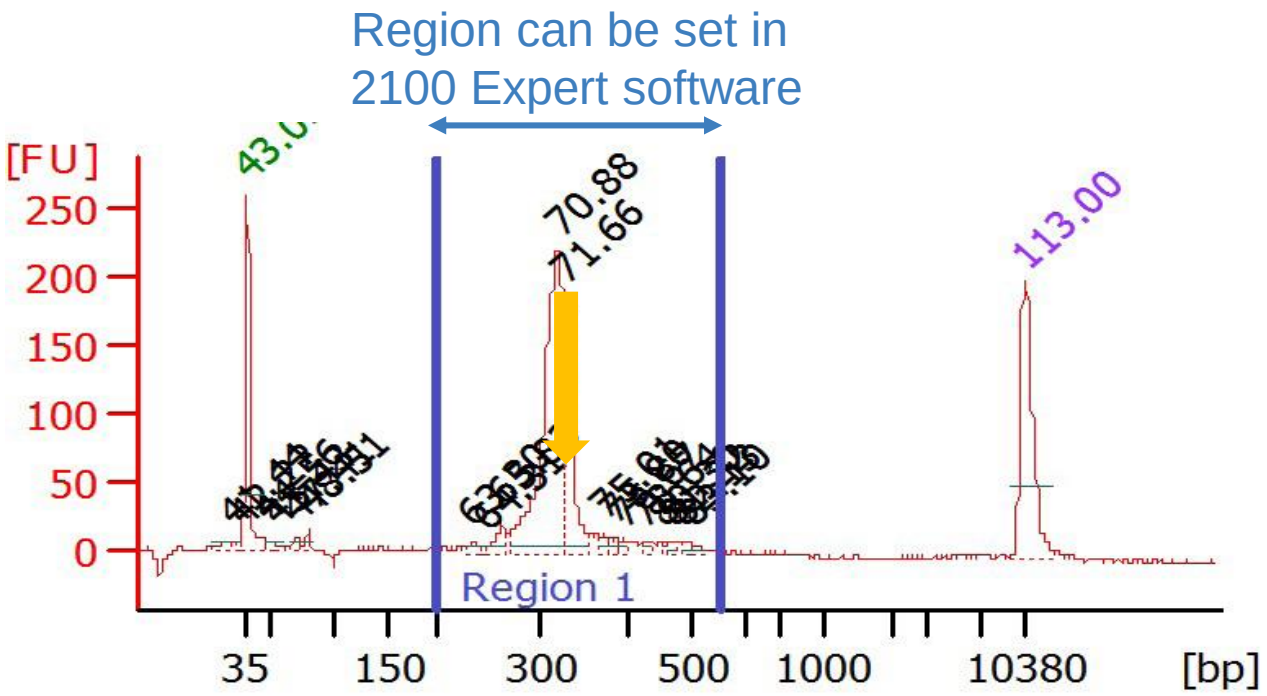
Understanding a Bioanalyzer Trace



	BioAnalyzer Assay Options	
	DNA 1000 Chip	High Sensitivity DNA Chip
Dynamic Range	25-1000 bp	50-7000 bp
Input required	0.5-50ng	5-500pg

illumina®

Bioanalyzer Details



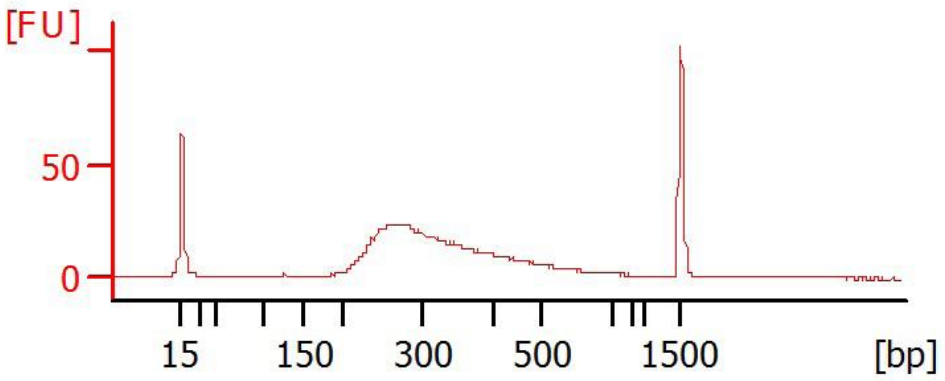
Average library size

illumina®

Expected Bioanalyzer Traces

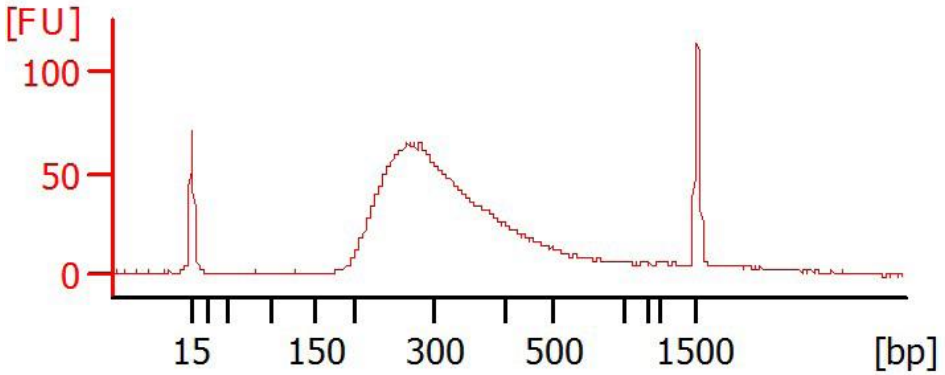
TruSeq Stranded mRNA:

100ng total RNA input



TruSeq Stranded Total RNA:

1µg total RNA input

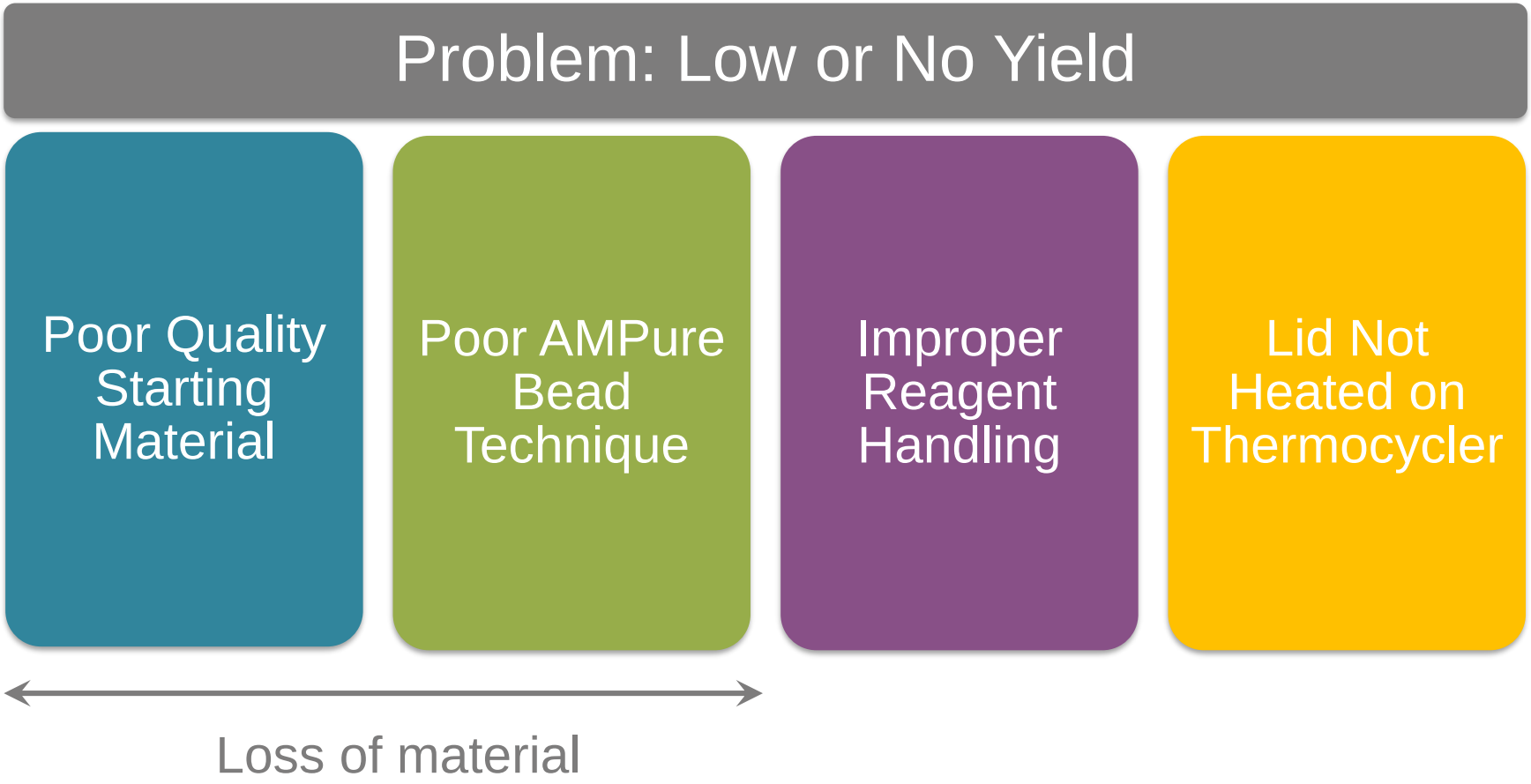


illumina®

Troubleshooting Library QC results

illumina®

Troubleshooting scenario 1



illumina®

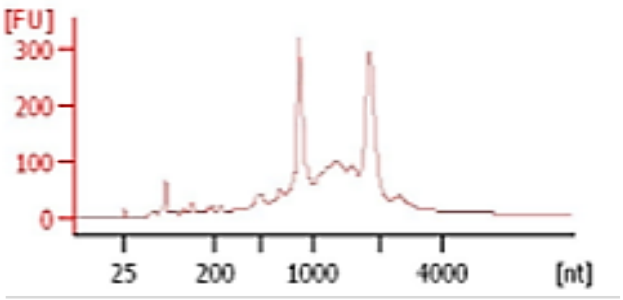
RNA input recommendations

Assess RNA quality (Bioanalyzer)

Quantify RNA (Qubit or PicoGreen)

Normalize total RNA samples

RIN value



TruSeq Stranded mRNA

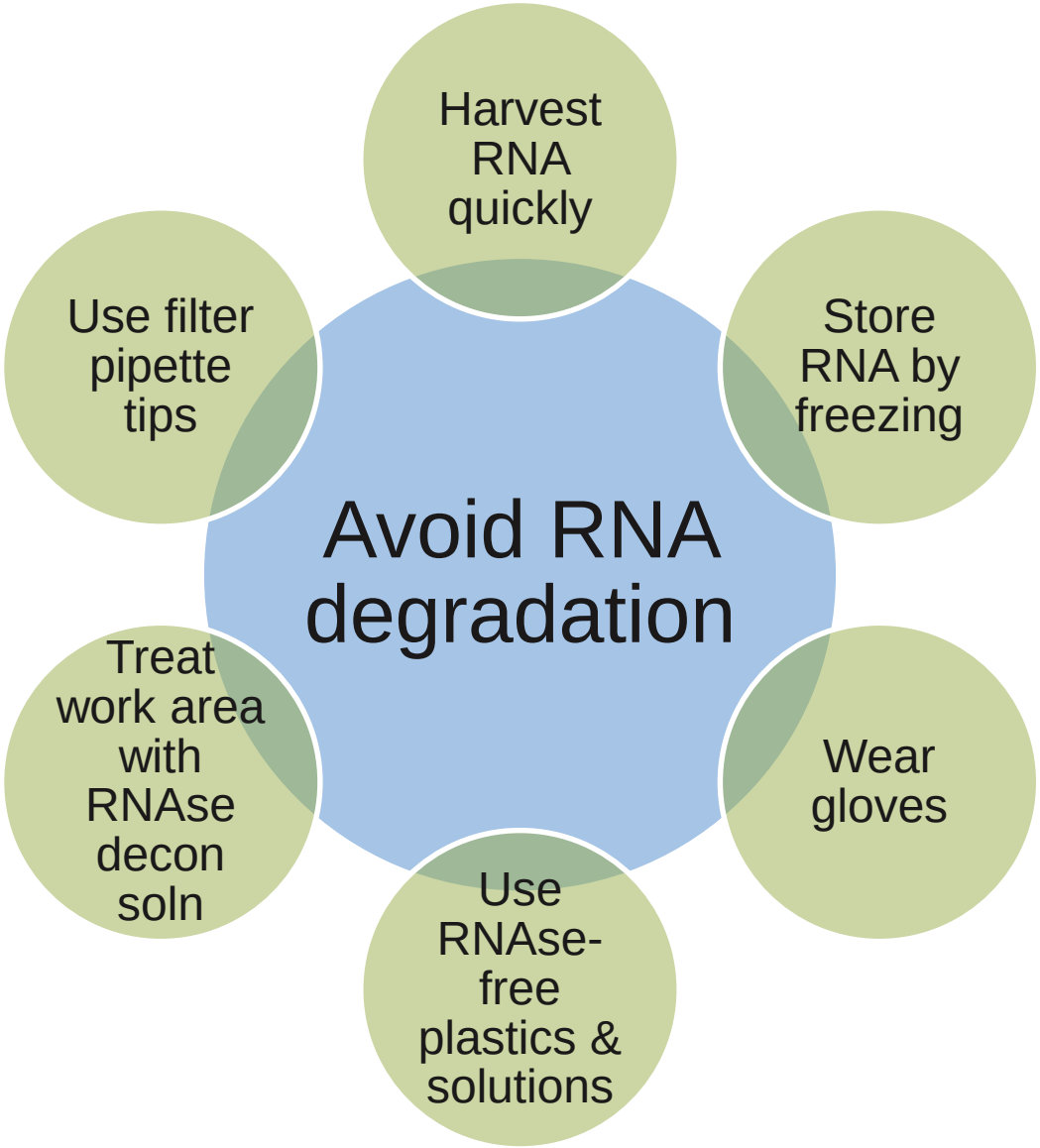
100 ng – 1 μg total RNA
High quality (RIN >7)

TruSeq Stranded Total RNA

100 ng – 1μg total RNA
Degraded RNA can be used

illumina

RNA Handling Best Practices



illumina®

Bead handling

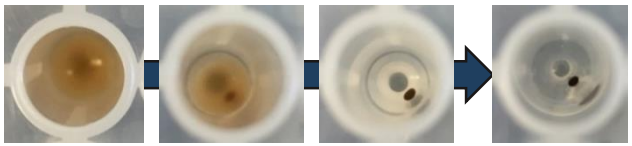
AMPure XP beads best practices

- Bring beads to room temp
- Resuspend beads thoroughly
- Mix well: use shaker and follow protocol
- Pipet carefully

  Minimize bead loss
- Use correct magnetic stand & follow protocol
- Dry pellet on plate

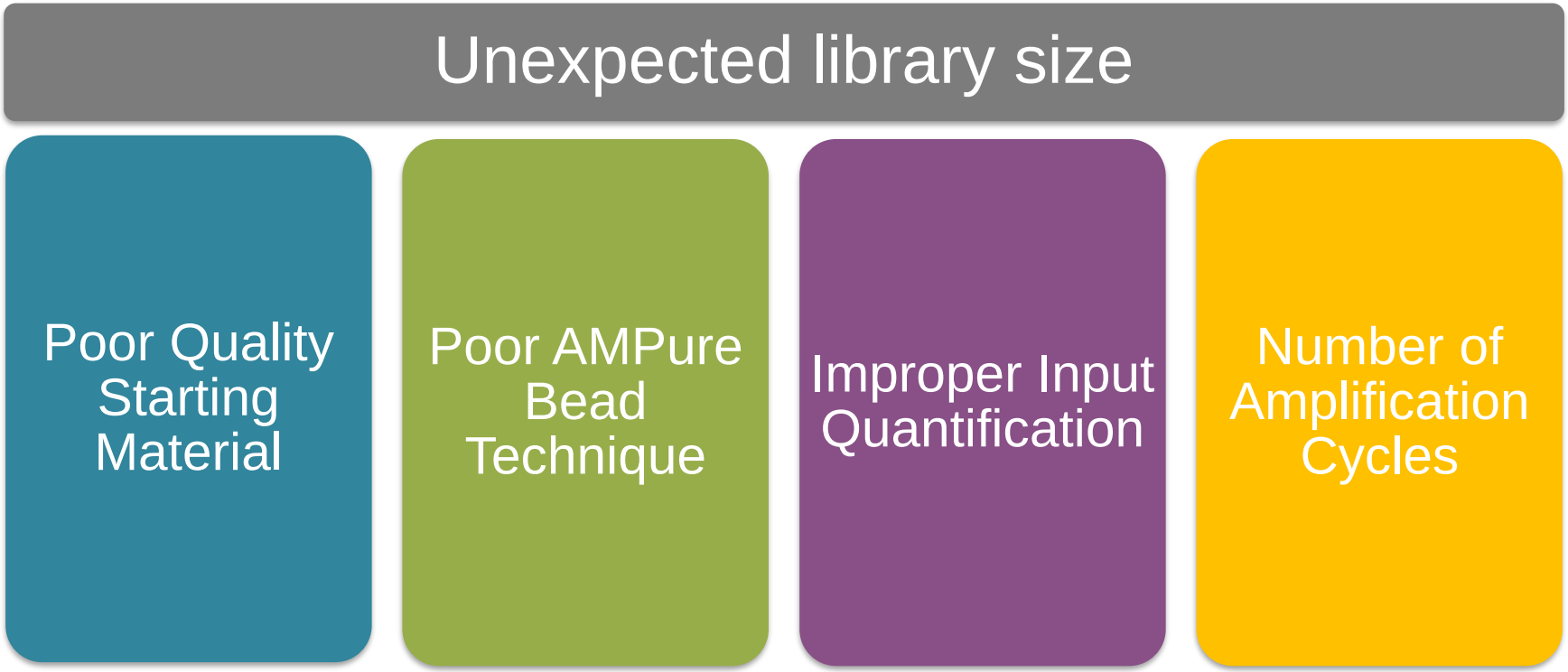
Vortex clumps

No Sunsets



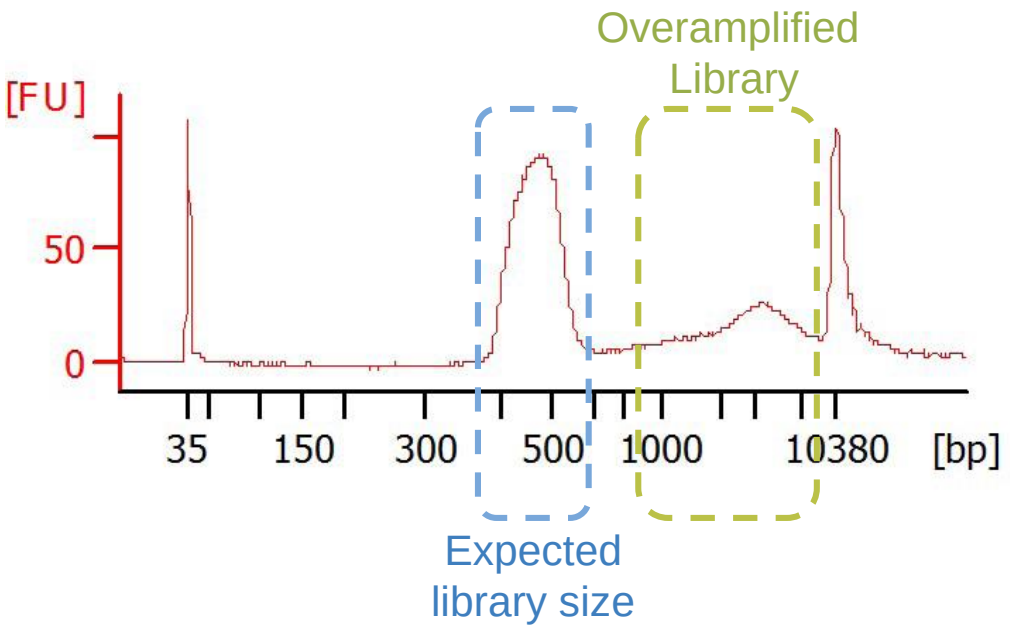
illumina®

Troubleshooting scenario 2



illumina®

Library over-amplification

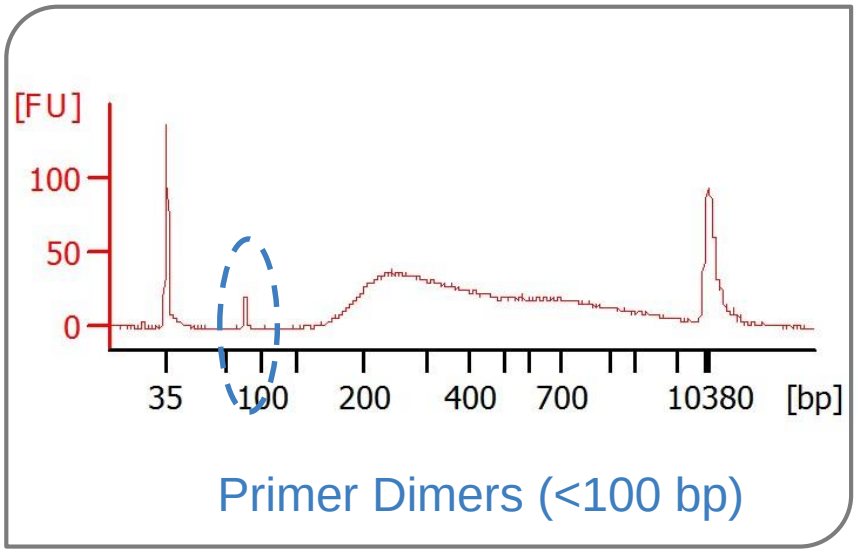


- **Occurs when PCR reagents (primers) are exhausted**
  - Do not increase PCR cycles
  - Quantify input RNA using fluorometric assay and do not exceed recommended input range (100 ng – 1 μg total RNA)
  - Will create duplicate sequencing reads and possible amplification bias

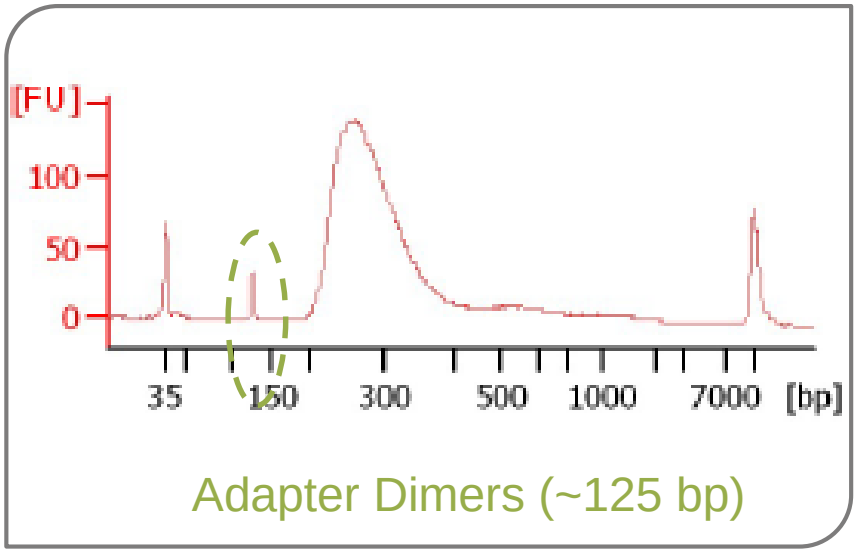
illumina®

Troubleshooting scenario 3

Presence of Library Contaminants



Primer Dimers (<100 bp)



Adapter Dimers (~125 bp)

illumina®

Why are dimers problematic?

- **Primer dimers cannot cluster, but they can bind to the flow cell surface**
  - Occupy real estate and reduce sequencing yield
- **Adapter dimers will form clusters and be sequenced**
  - Lower yield (some reads will be adapter sequence)
  - Lower passing filter (loss of quality)
  - On patterned flow cells (HiSeq 4000, NovaSeq), adapter dimers even at low levels cluster more efficiently than larger libraries
- **Inefficient removal of adapter dimers correlate with higher rates of index hopping on patterned flow cells**
  - Incorrect assignment of reads to their index

Adapter Dimer Sources

- **Adapter dimers are formed when the P5 and P7 adapters ligate to one another rather than to an inserted library**

Adapter Dimers Sources

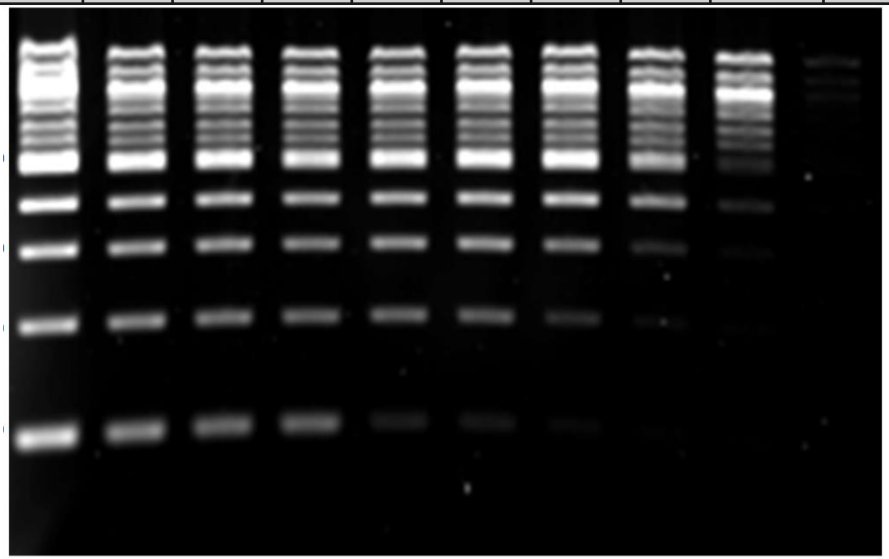
- Poor quality input
- Bead Handling Error
- Insufficient input amount
- Enzymatic failure

illumina®

Clean up steps

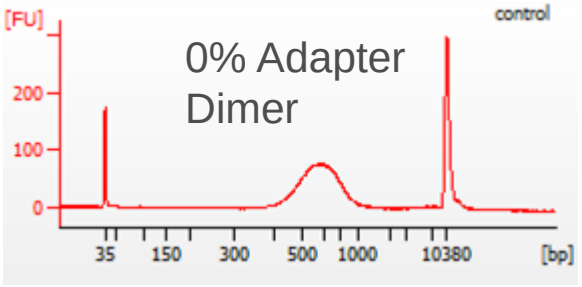
- **Beads are used to remove small contaminants**
- **Ratio of beads to DNA is critical**
  - 2 rounds of bead clean-up after ligation
  - Be mindful of extra beads on the outside of pipet tips

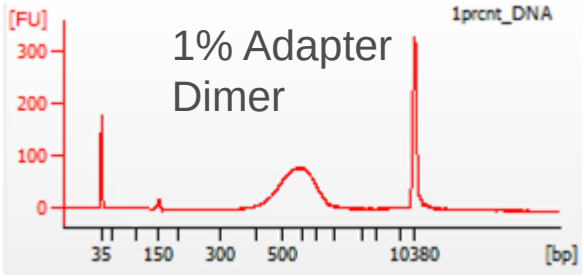
Input	10	13	17	25	31	36	42	50	45	µl DNA
	25	25	25	25	25	25	25	25	17.5	µl beads

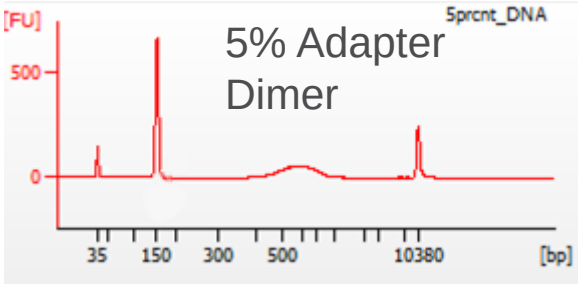


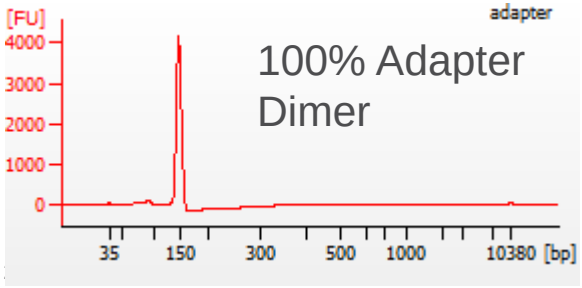
illumina®

Removal of contaminants





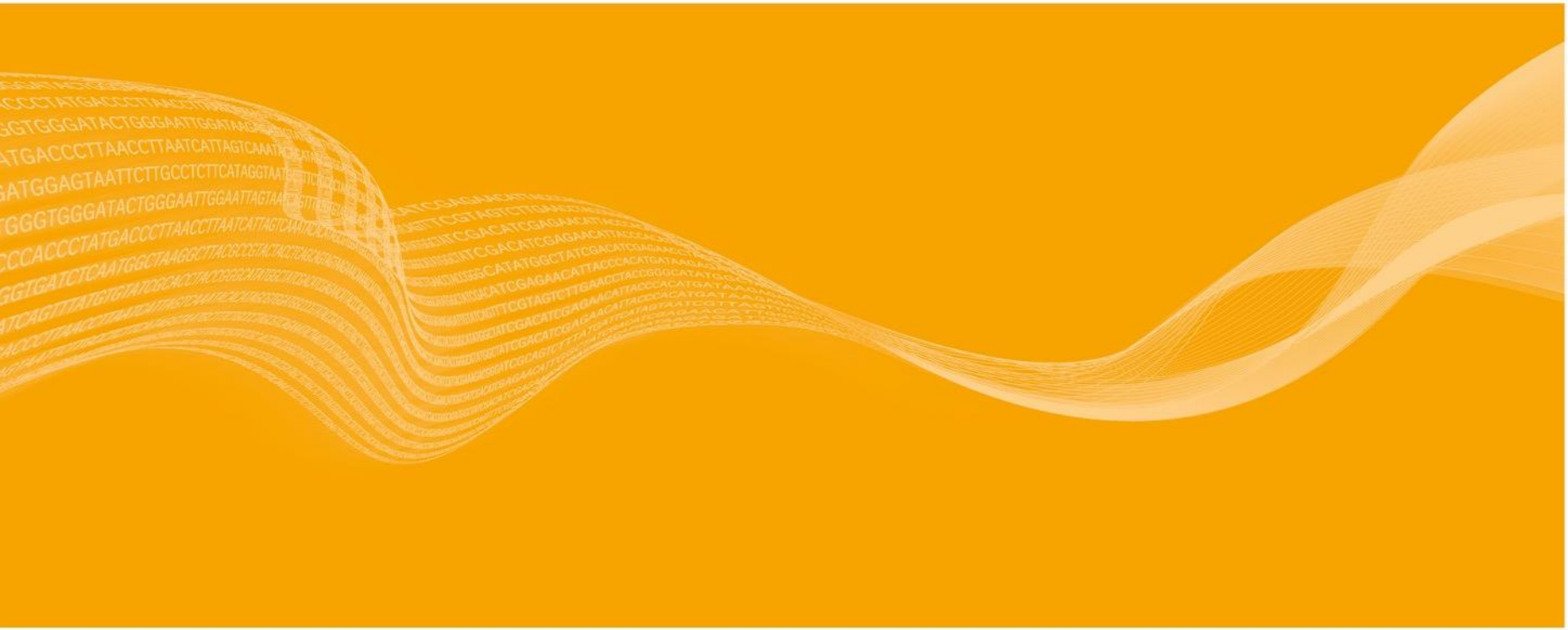




If you see adapter dimers, perform additional 1:1 bead clean-up step

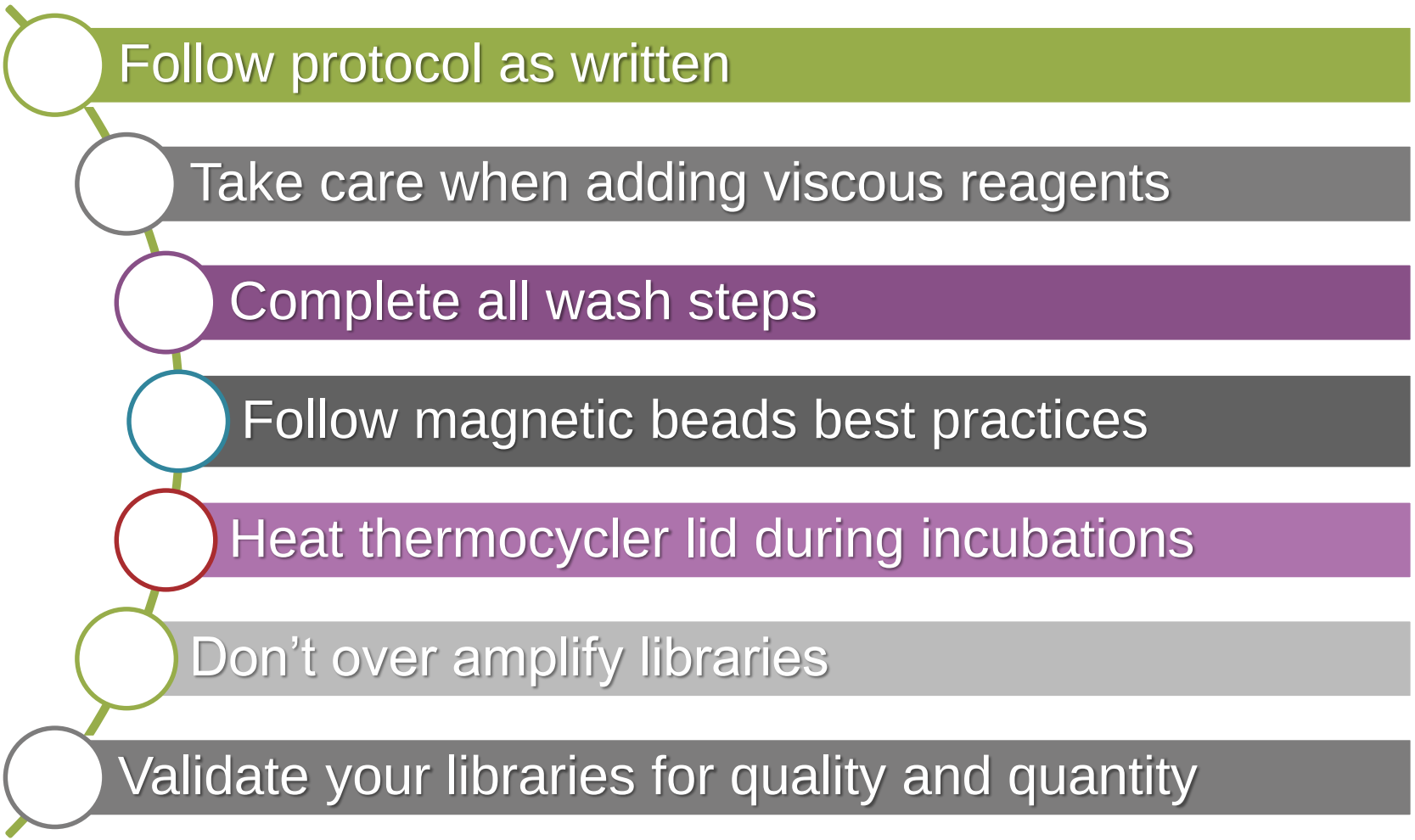
illumina®

Summary



illumina®

Best Practices Summary



illumina®

Before sequencing …

- **Make sure that**:
  - The inserts are of the correct size and have adapters on each end
  - The library concentration is appropriate
  - The libraries are clean and free of major contaminants

- **Furthermore, the sequencing facility needs to know**:
  - The sequence of the sequencing-priming site
  - The length and type of the read you desire
  - Whether there is a barcode or index sequence
  - If the inserts contain a region of low sequence complexity immediately after the sequencing-priming site
  - What buffer the libraries are resuspended in