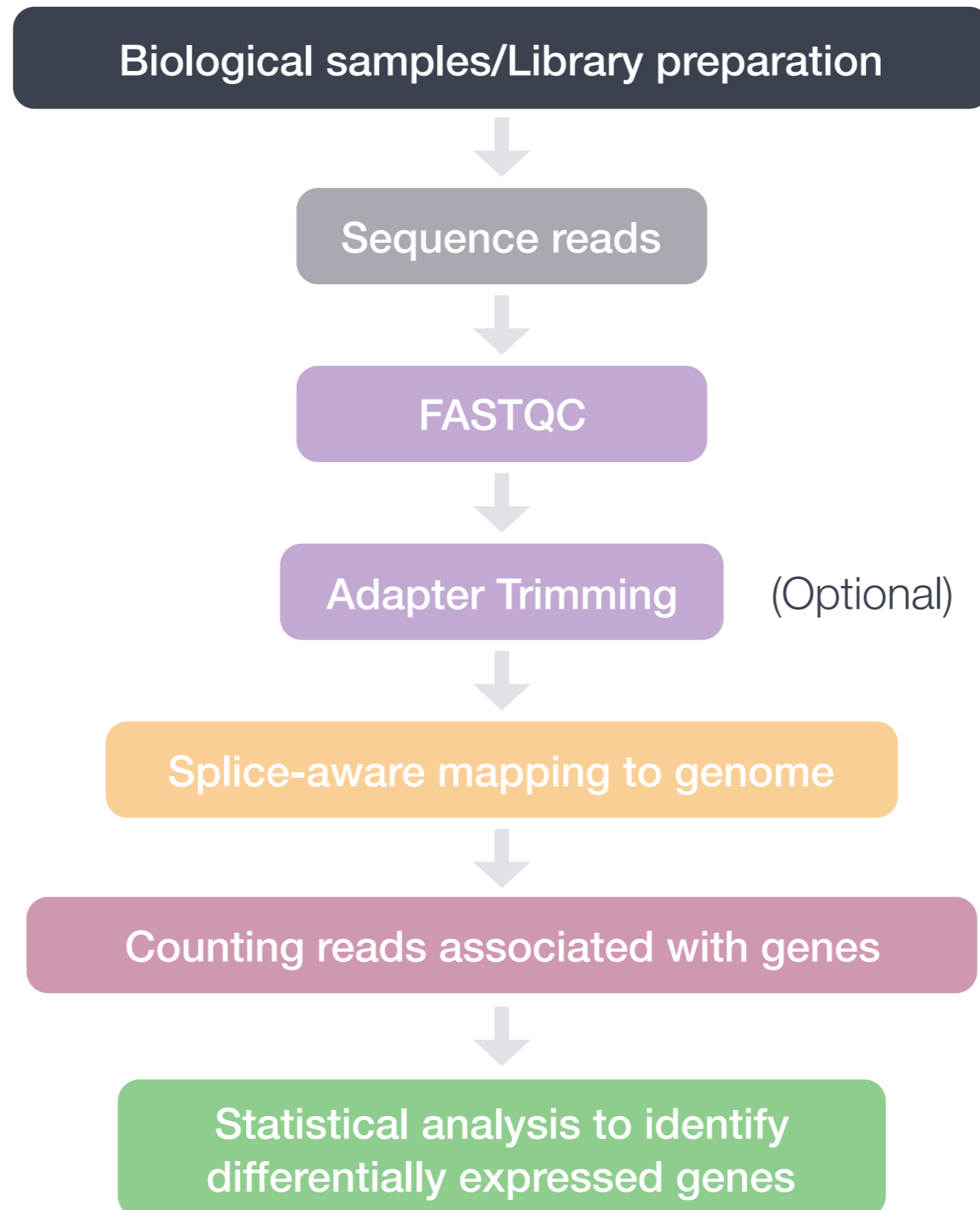


Sequencing Library Preparation

Slides courtesy of **Sarah Boswell** <http://scholar.harvard.edu/saboswell>



RNA-seq Workflow



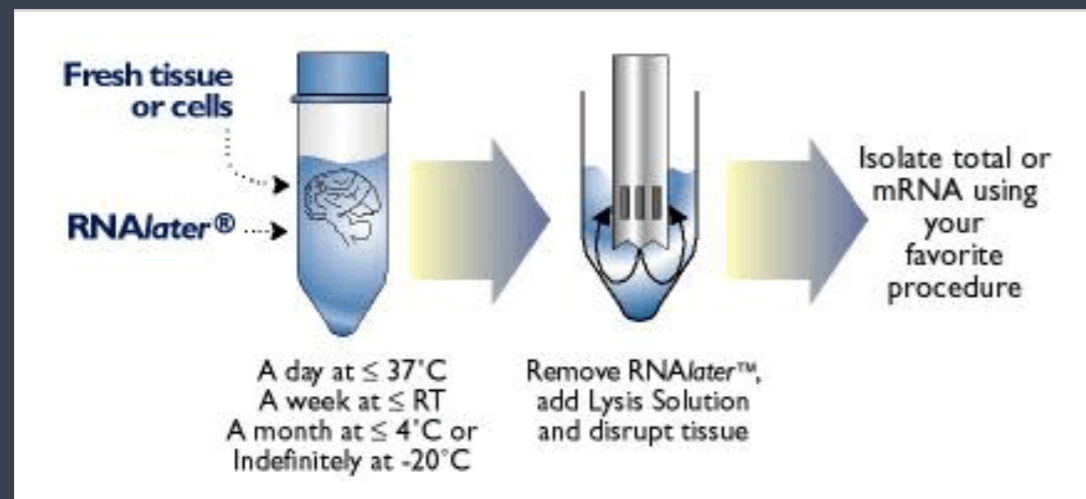
Key steps in library preparation

- ✓ Starting Material
- Library Preparation
- Library Amplification Bias
- Sequencing Terminology & Multiplexing

Purification and QC of RNA

- Start with highest quality RNA possible
- Perform DNase digestion
- Accurately quantify RNA
- Assess quality of RNA

RNA Extraction: Tissues / Trizol



RNAlater®-ICE

Thaw tissue not frozen in RNAlater® in this solution prior to RNA extraction.

- Keep tissues as cold as possible
- Work in cold room
- After homogenization suggest column based cleanup
 - Particularly important if used Trizol for lysis
 - DNase 10ug then column cleanup

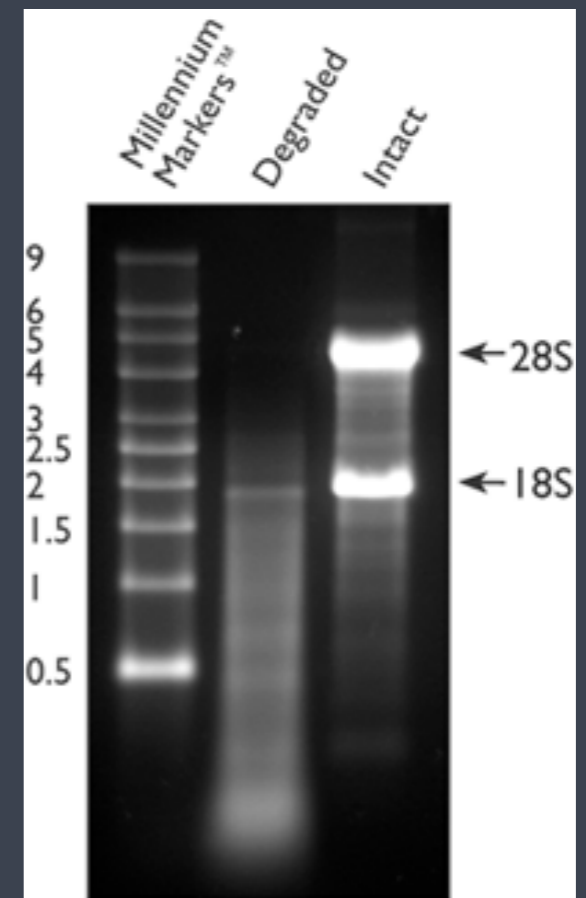
RNA Quantification & Quality

➤ Quantification

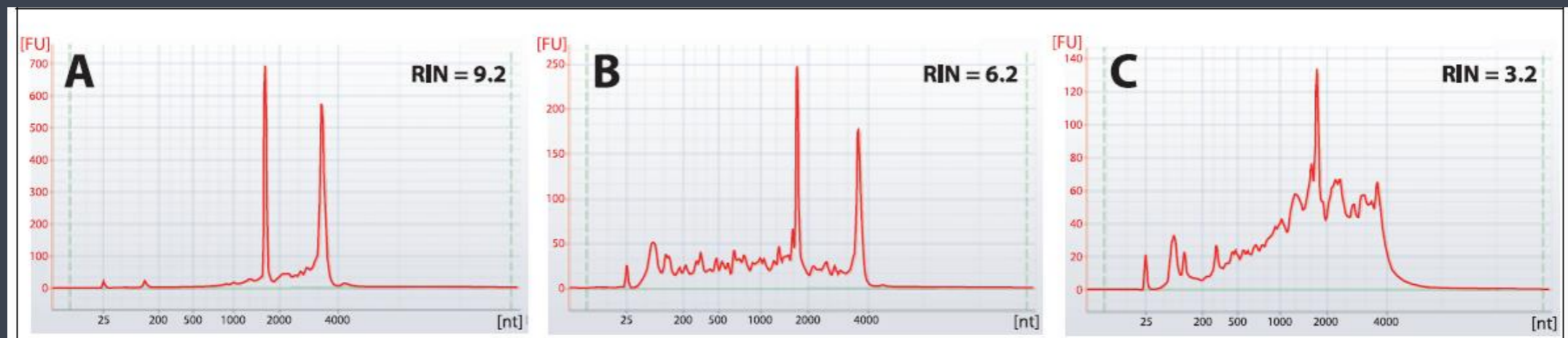
- Absorbance: Nano-drop (50-500 ng/ul)
 - Theoretically can read to 3000 ng/ul. Empirically find it is only accurate within range above.
- Dye based
 - RiboGreen
 - Qubit / Quant-IT

➤ Quality

- Visualize on gel
- Agilent Bioanalyzer (RIN)

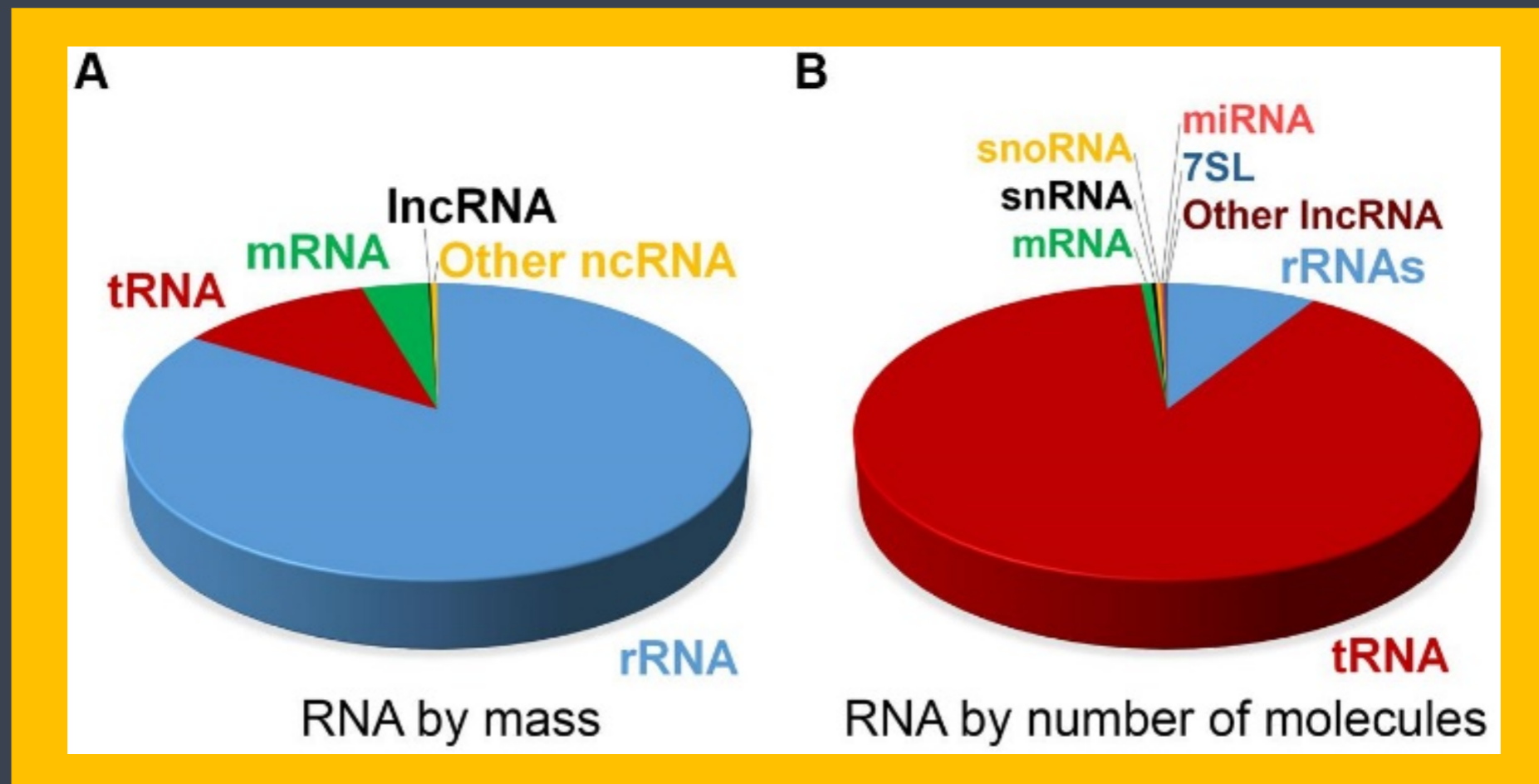


RNA quality



- High quality RNA needed for mRNA libraries
- Degraded samples should only be used to make a “total” RNA-seq library – rRNA removal
 - FFPE & Archival Samples

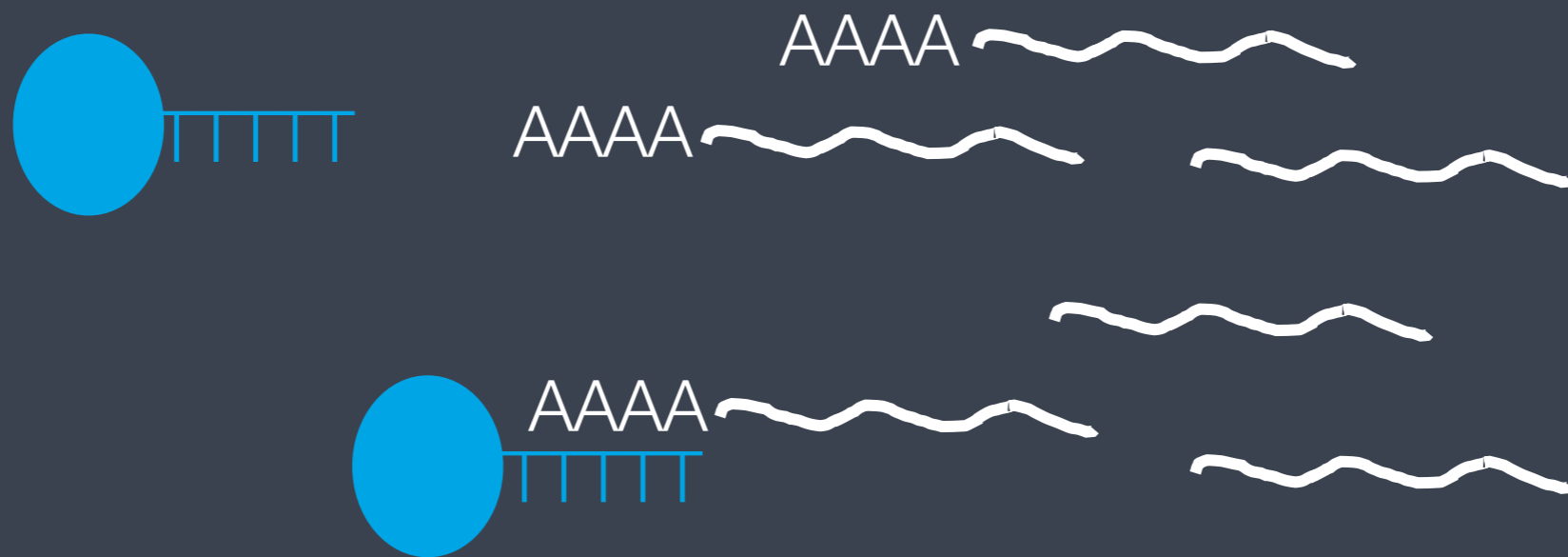
RNA enrichment



- PolyA tailed messenger RNA: mRNA-Seq
- Total RNA (rRNA removed): “total” RNA-Seq

mRNA (polyA) Purification

- mRNA enrichment
 - mRNA binds beads coated with oligo dT primer
 - Non-polyadenylated transcripts are washed away

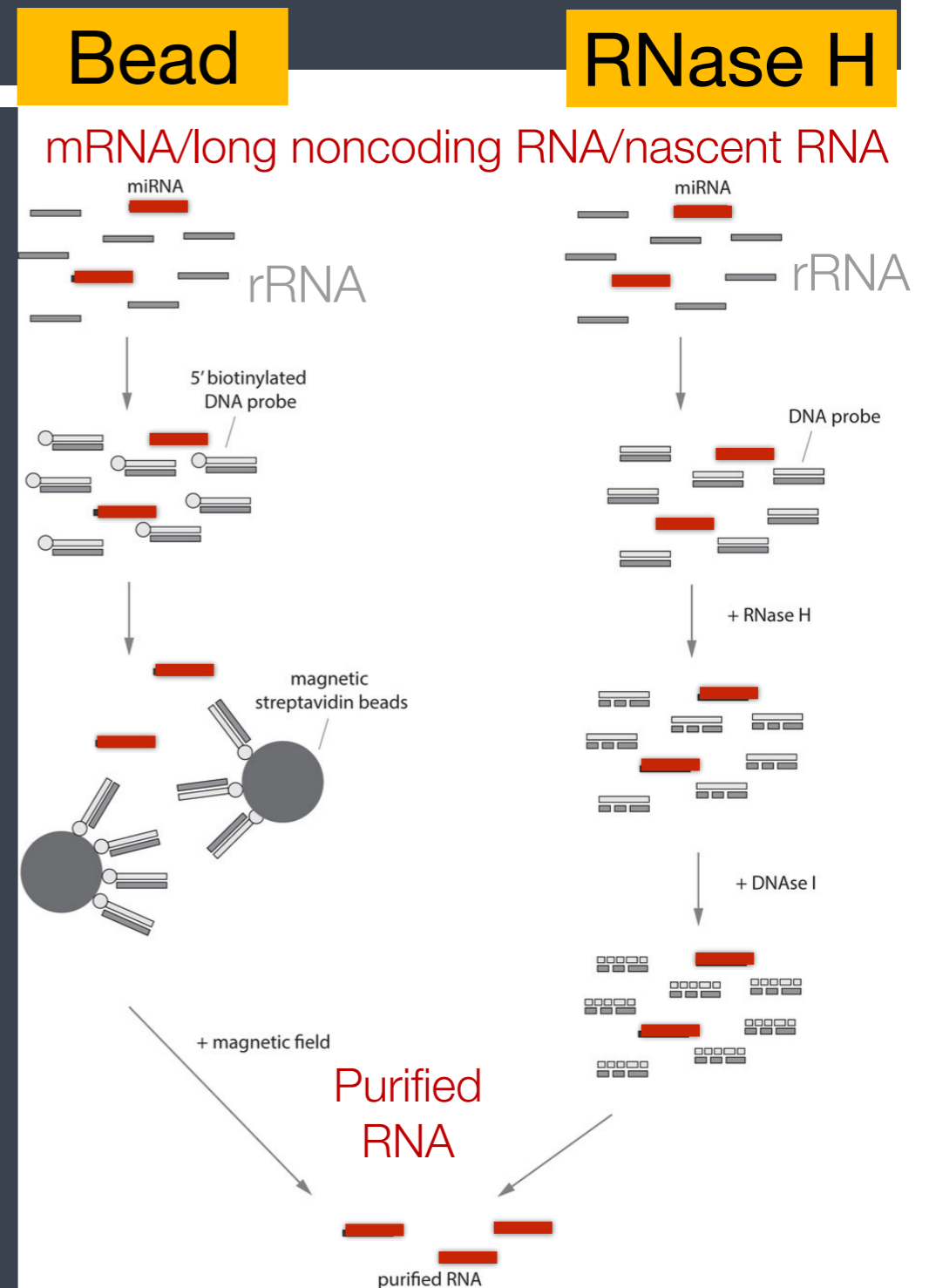


Transcripts Lost in polyA Purification

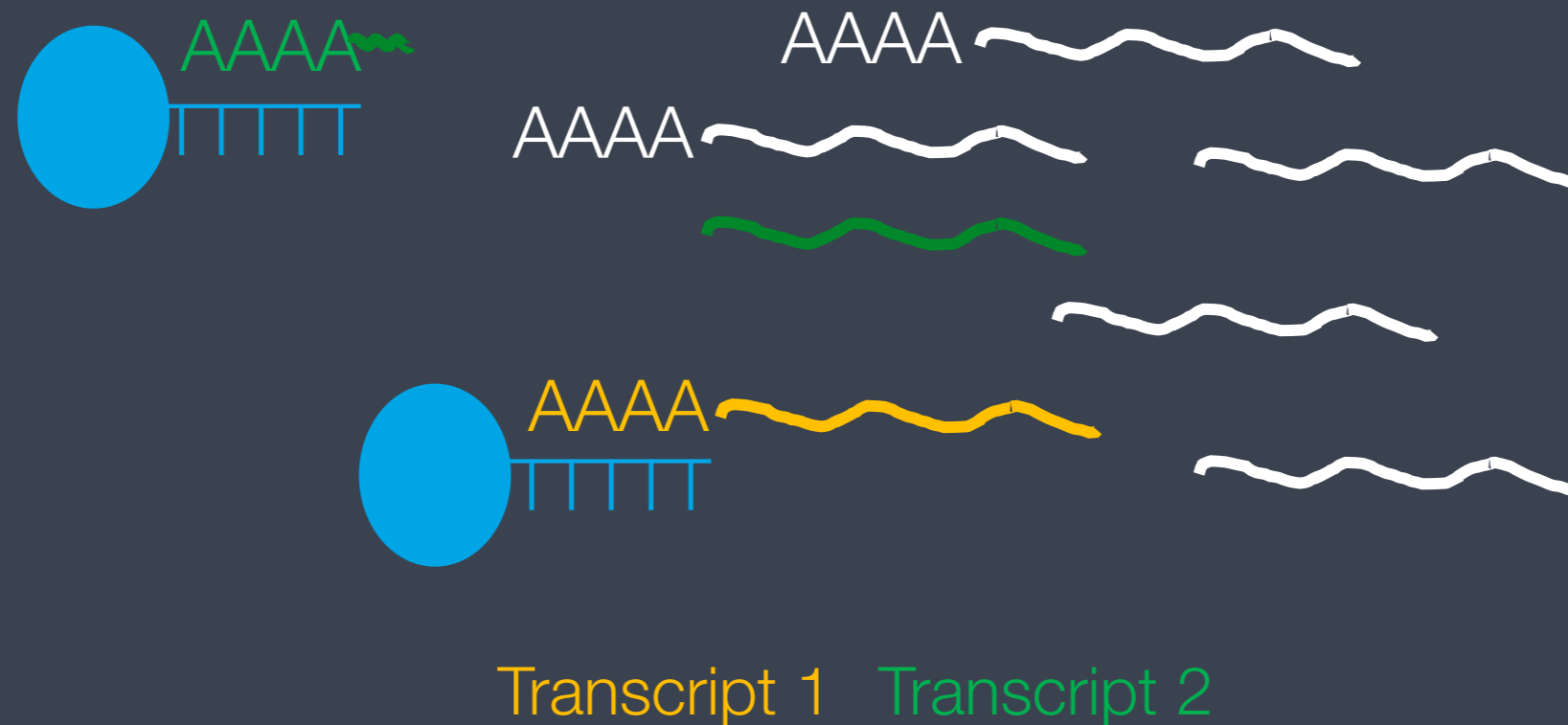
- Ribosomal/Transfer RNA
- Histone mRNA
- Long-noncoding RNA
- Nascent intron containing transcripts
- Micro RNA
- Degraded RNA
- Many viral transcripts
- Prokaryote/Bacterial transcripts
 - polyA is the degradation signal

rRNA Depletion

- Illumina: TruSeq
 - Probes hybridize rRNA on magnetic beads
 - RNA of interest remains in supernatant
- KAPA: RiboErase
 - Probes hybridize rRNA in solution
 - Hybrids are digested with RNase H
 - Probes digested with DNase I



mRNA Purification of Degraded Samples

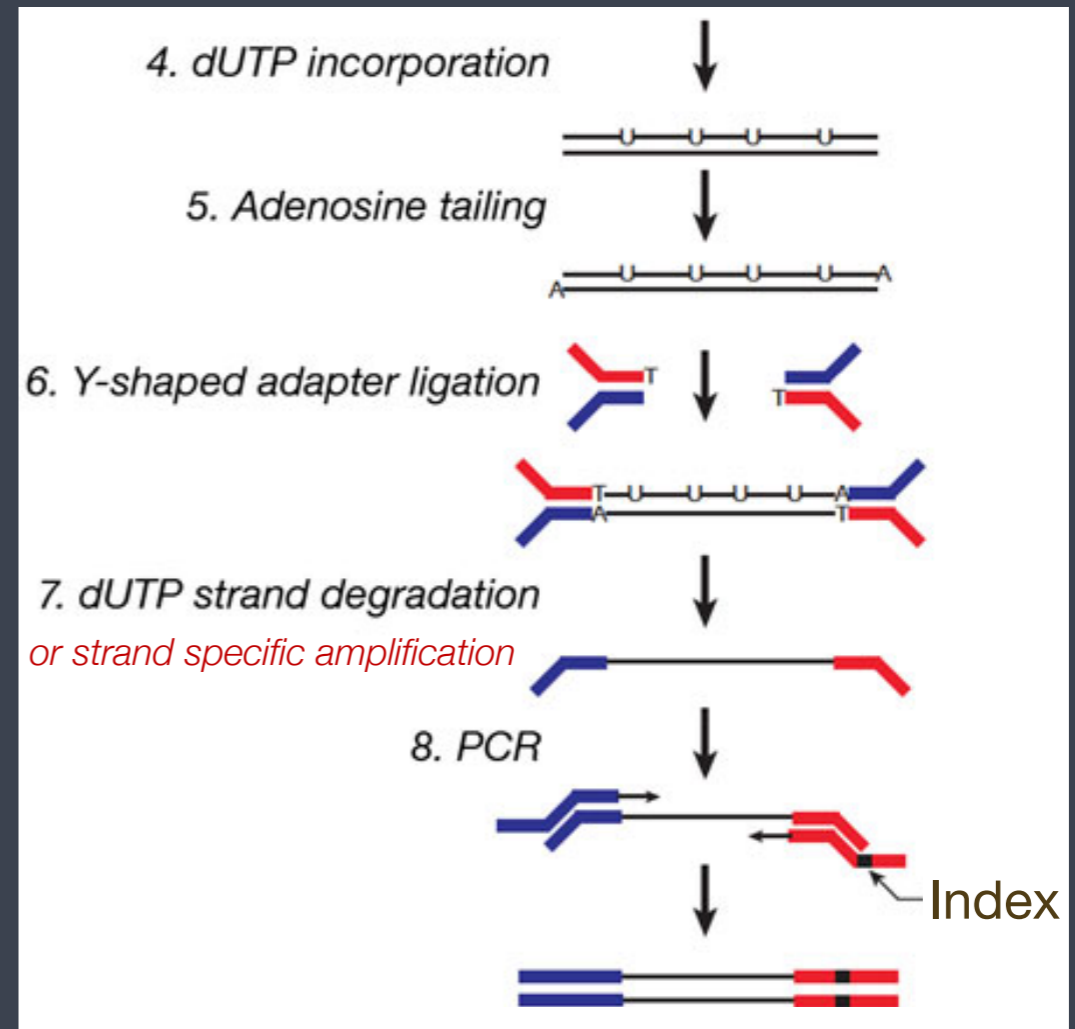
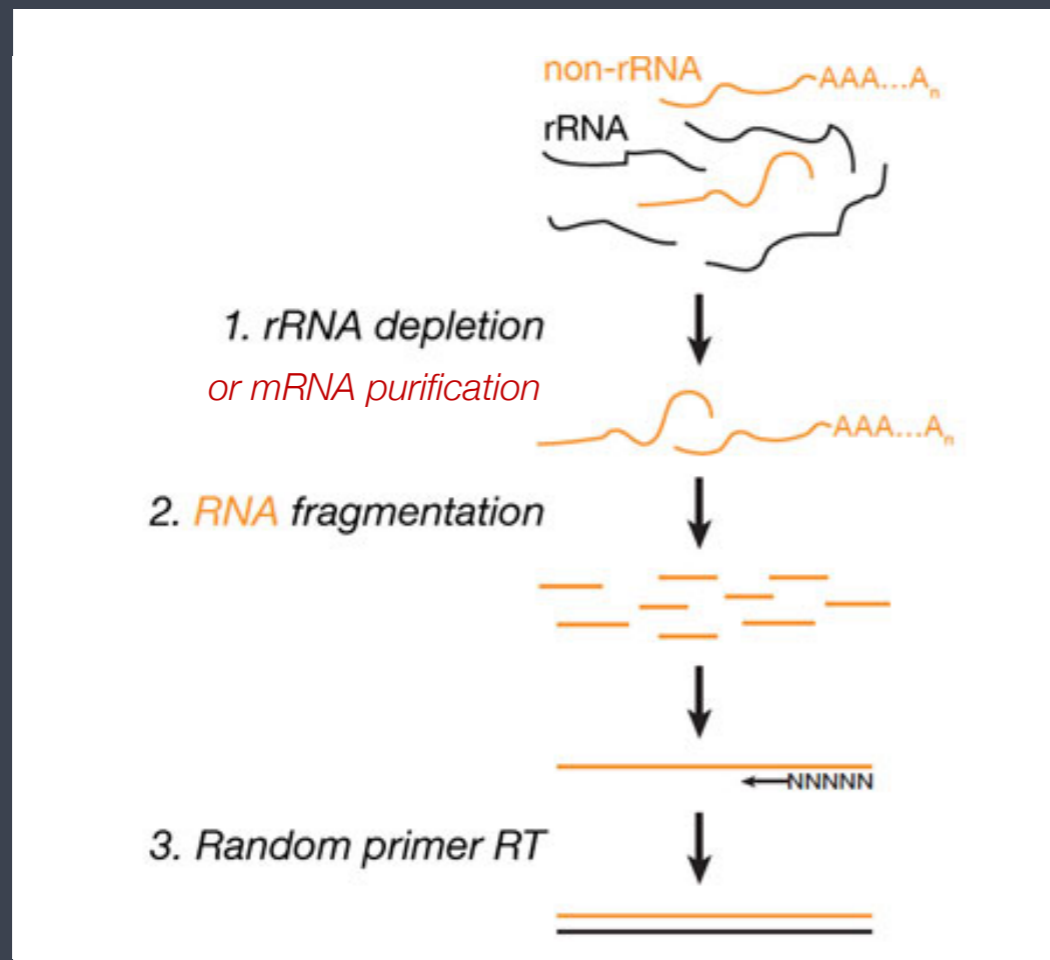


- PolyA tail no longer attached to transcript.
- Results in differential loss of transcripts between samples.

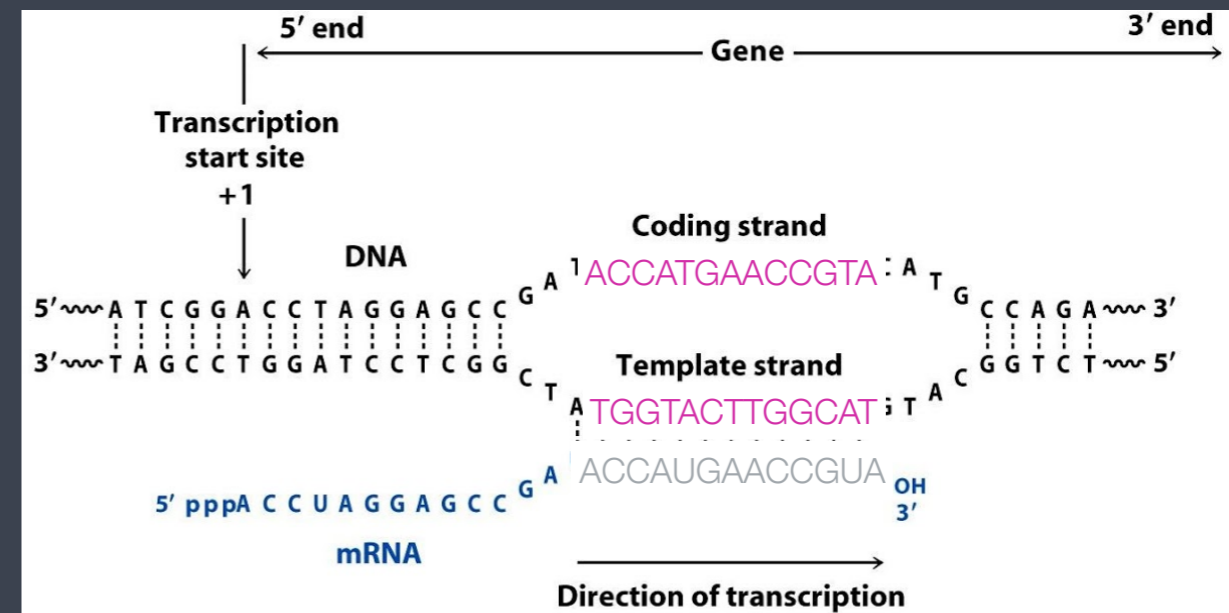
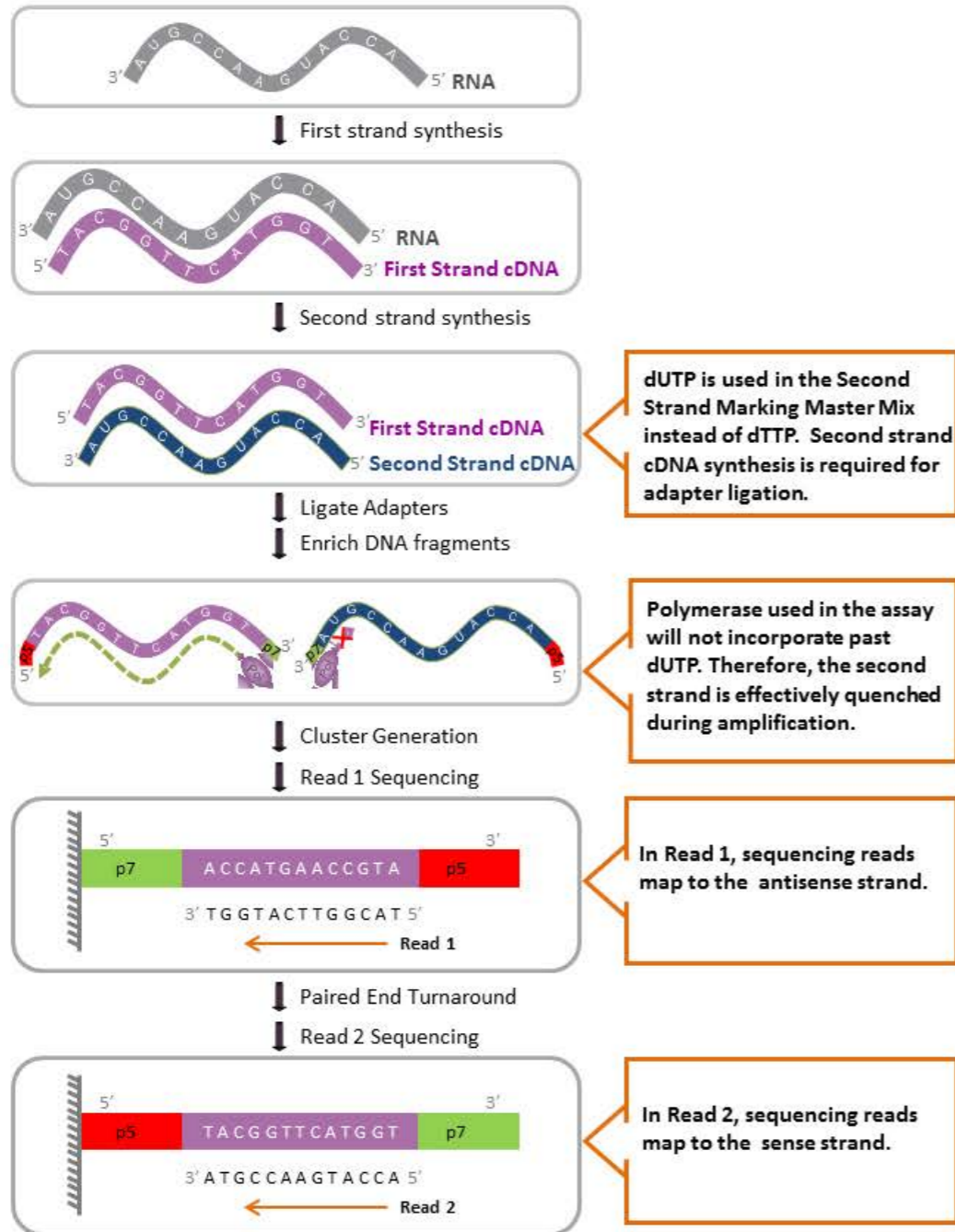
Key steps in library preparation

- ✓ Starting Material
- ✓ Library Preparation
 - Library Amplification Bias
 - Sequencing Terminology & Multiplexing

RNA-seq Stranded Library Prep (dUTP method)

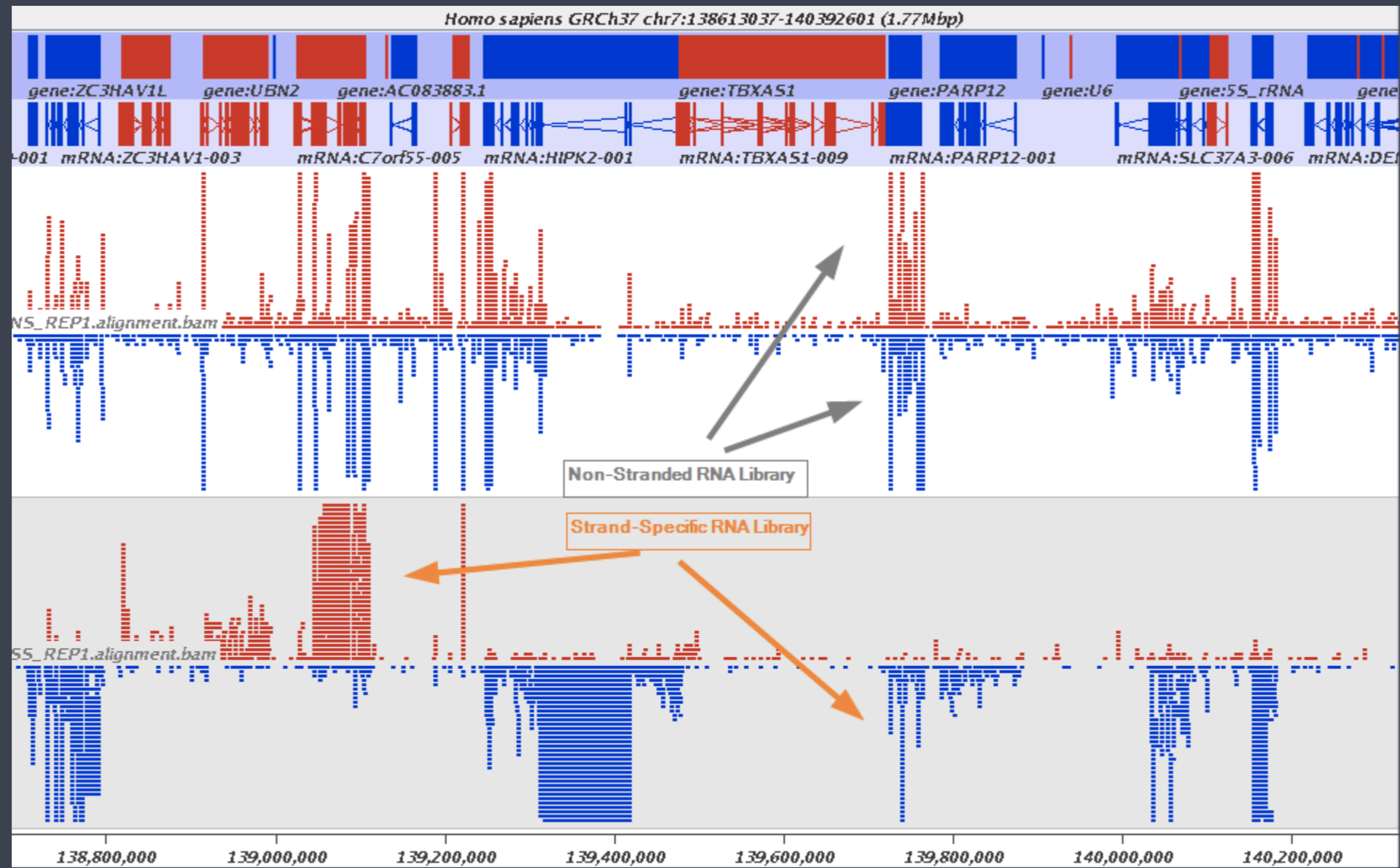


Library Strandedness



- Read alignment depends on direction of transcription
- “sense” strand of transcript can be on either the sense or antisense strand of the DNA

Library Strandedness



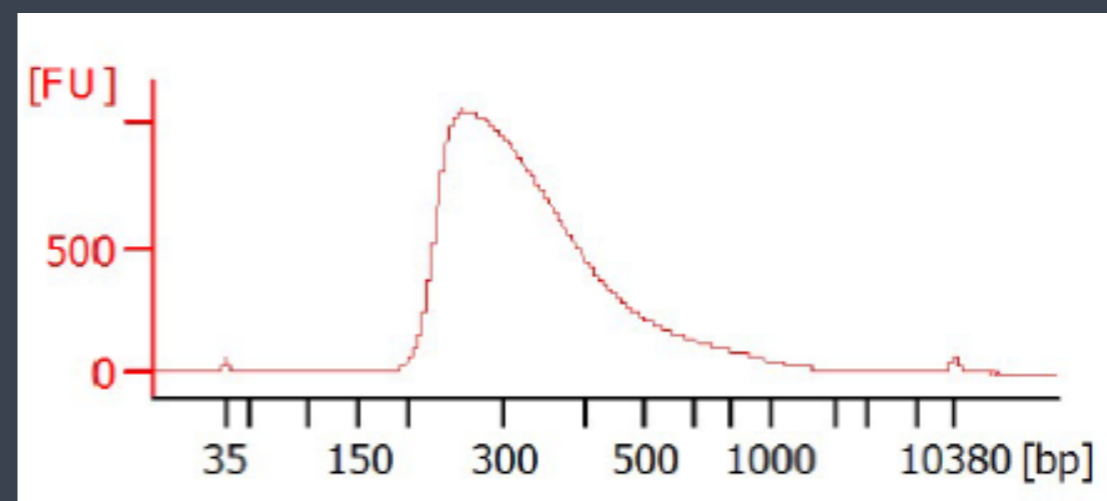
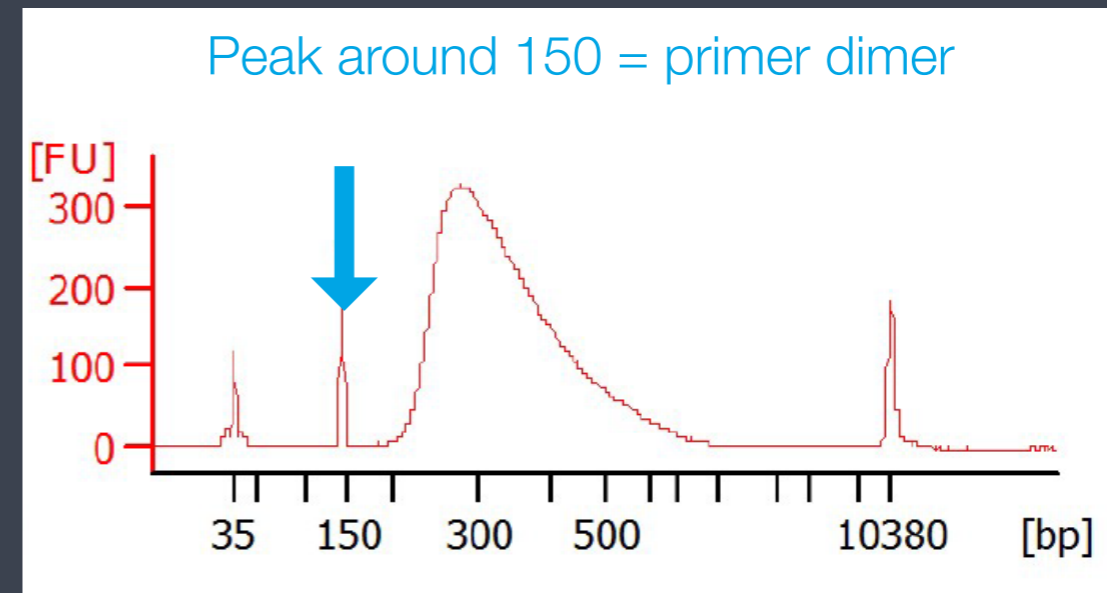
Library QC

➤ Quantification

- Dye based
 - SYBR Green
 - Qubit / Quant-IT

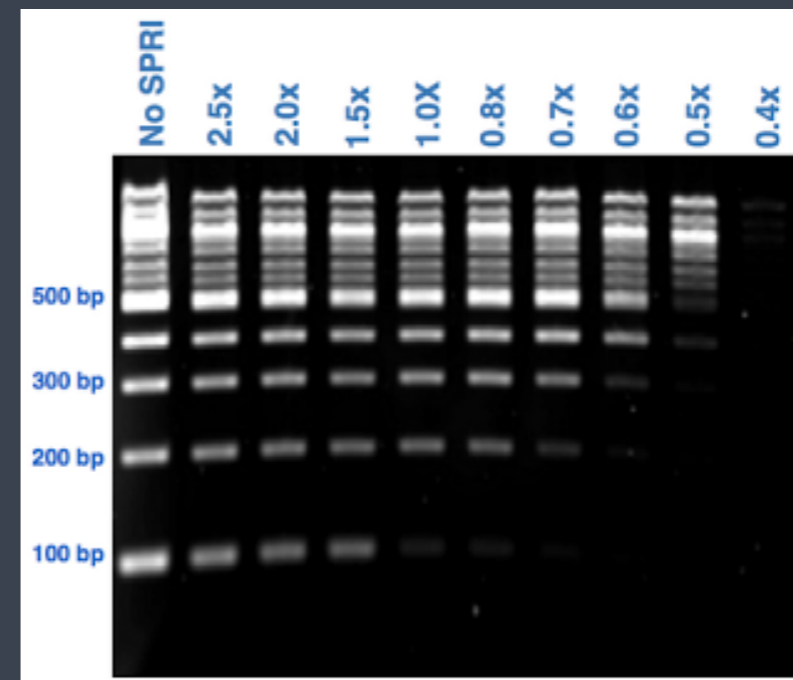
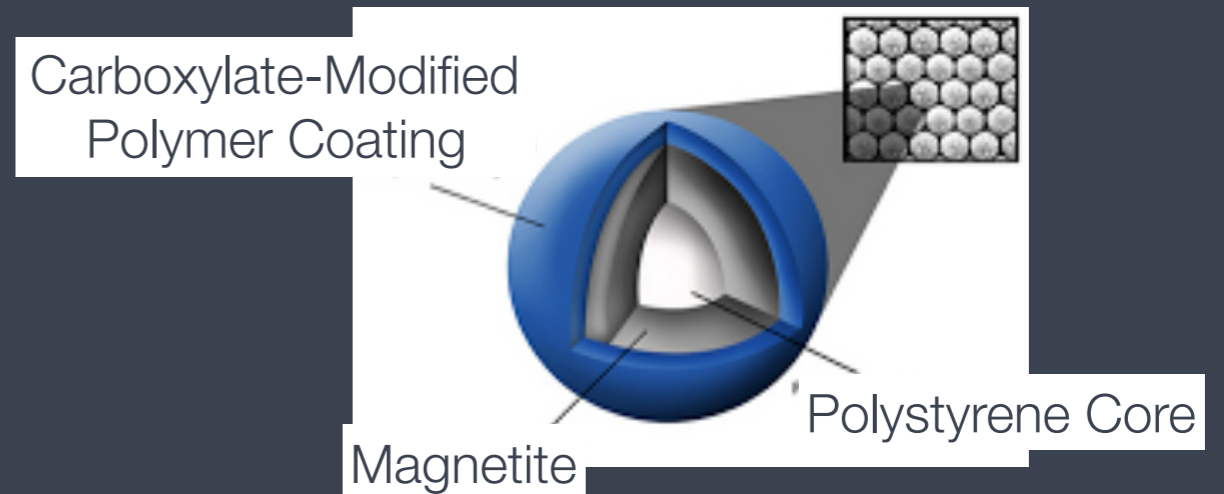
➤ Size & Quality

- Agilent Bioanalyzer
- Size determination
- Do not use for quantification



Size selection with SPRI beads

- Solid Phase Reverse Immobilization beads
- Carboxyl groups on surface bind DNA in the presence of crowding agents (PEG & NaCl)

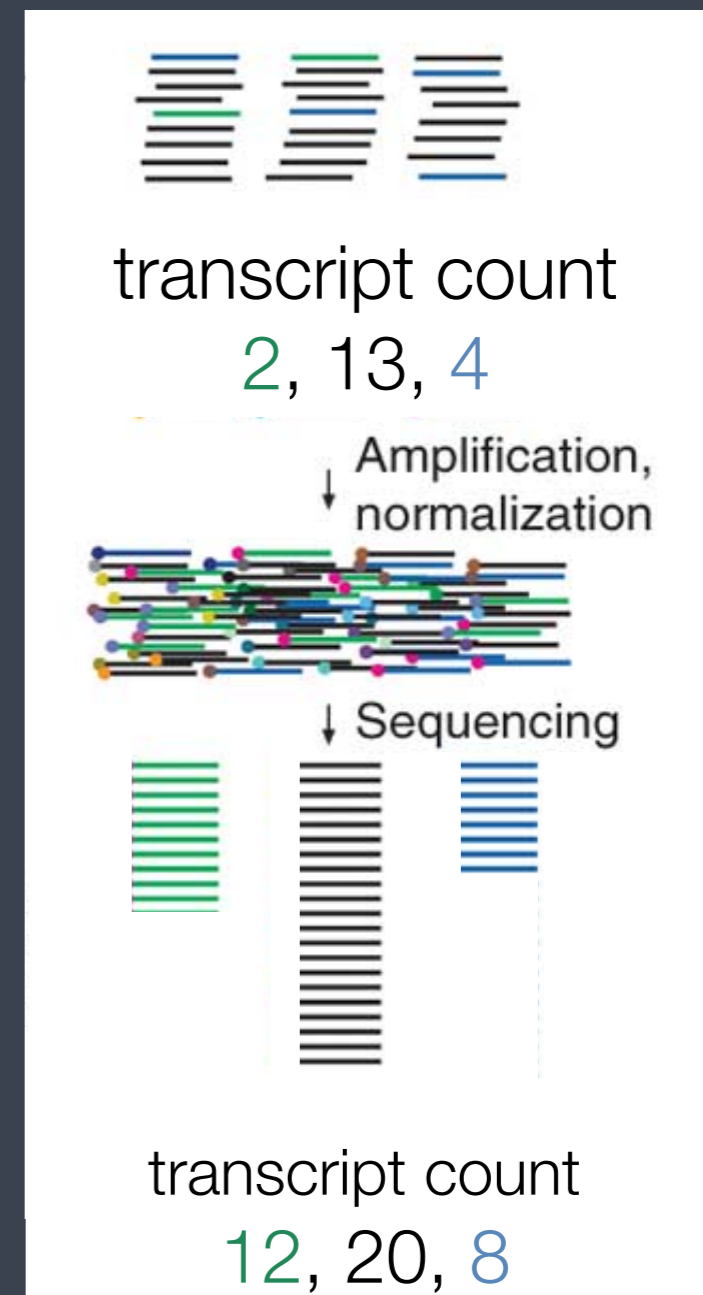


Key steps in library preparation

- ✓ Starting Material
- ✓ Library Preparation
- ✓ Library Amplification Bias
- Sequencing Terminology & Multiplexing

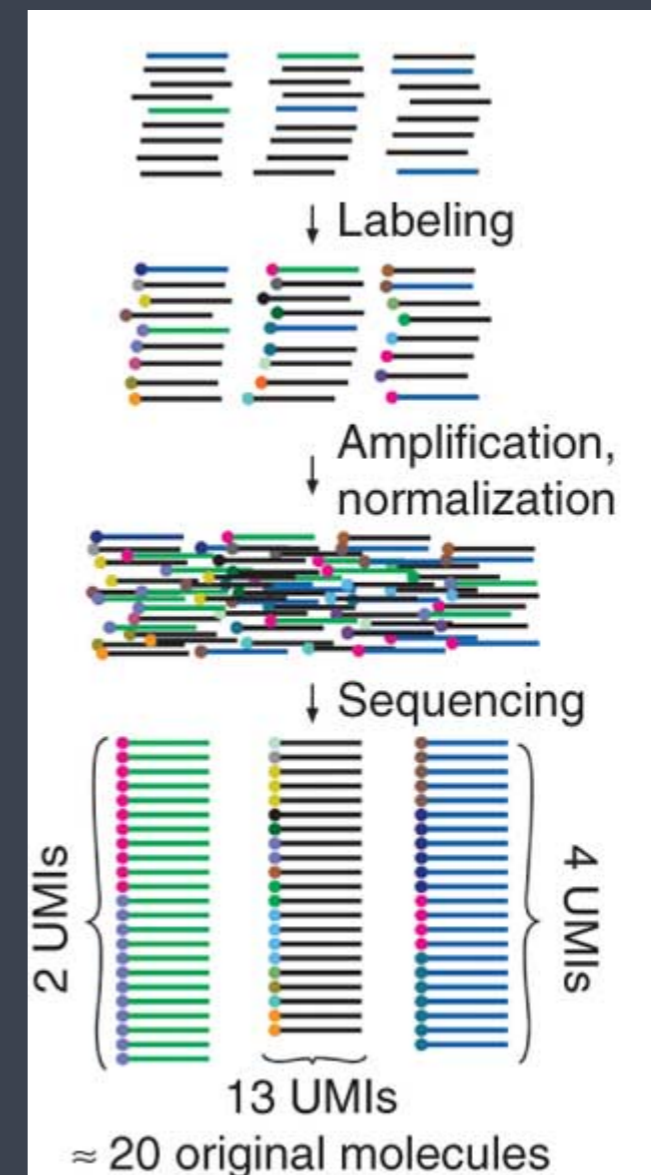
Library Amplification Bias

- Final step of library prep is amplification
- Introduces library bias
 - Some products preferentially amplified
- Fewer cycles = less bias



Library Amplification Bias

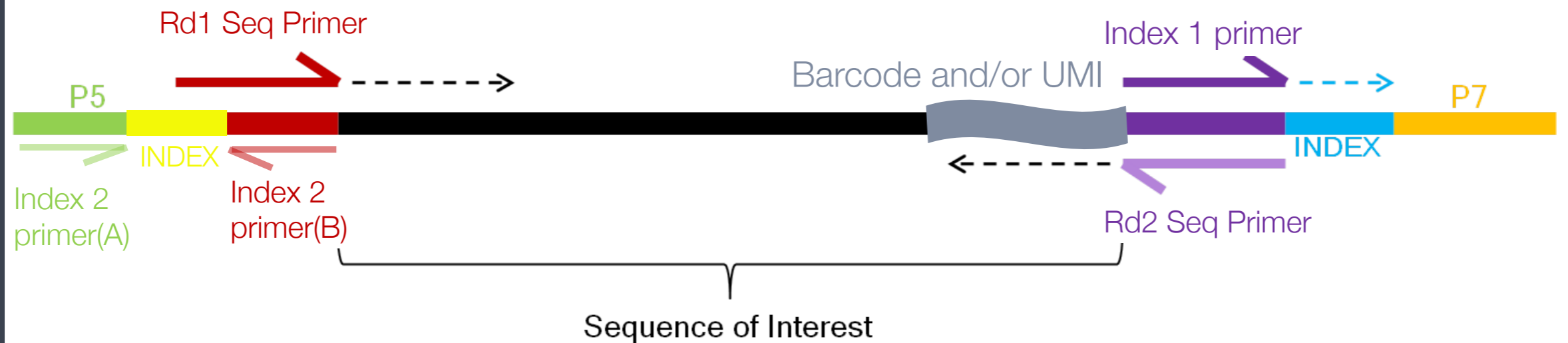
- Final step of library prep is amplification
- Introduces library bias
 - Some products preferentially amplified
- Fewer cycles = less bias
- **Unique Molecular Index (UMI)**
 - Accurately quantitates samples



Key steps in library preparation

- ✓ Starting Material
- ✓ Library Preparation
- ✓ Library Amplification Bias
- ✓ Sequencing Terminology & Multiplexing

Sequencing Read Order



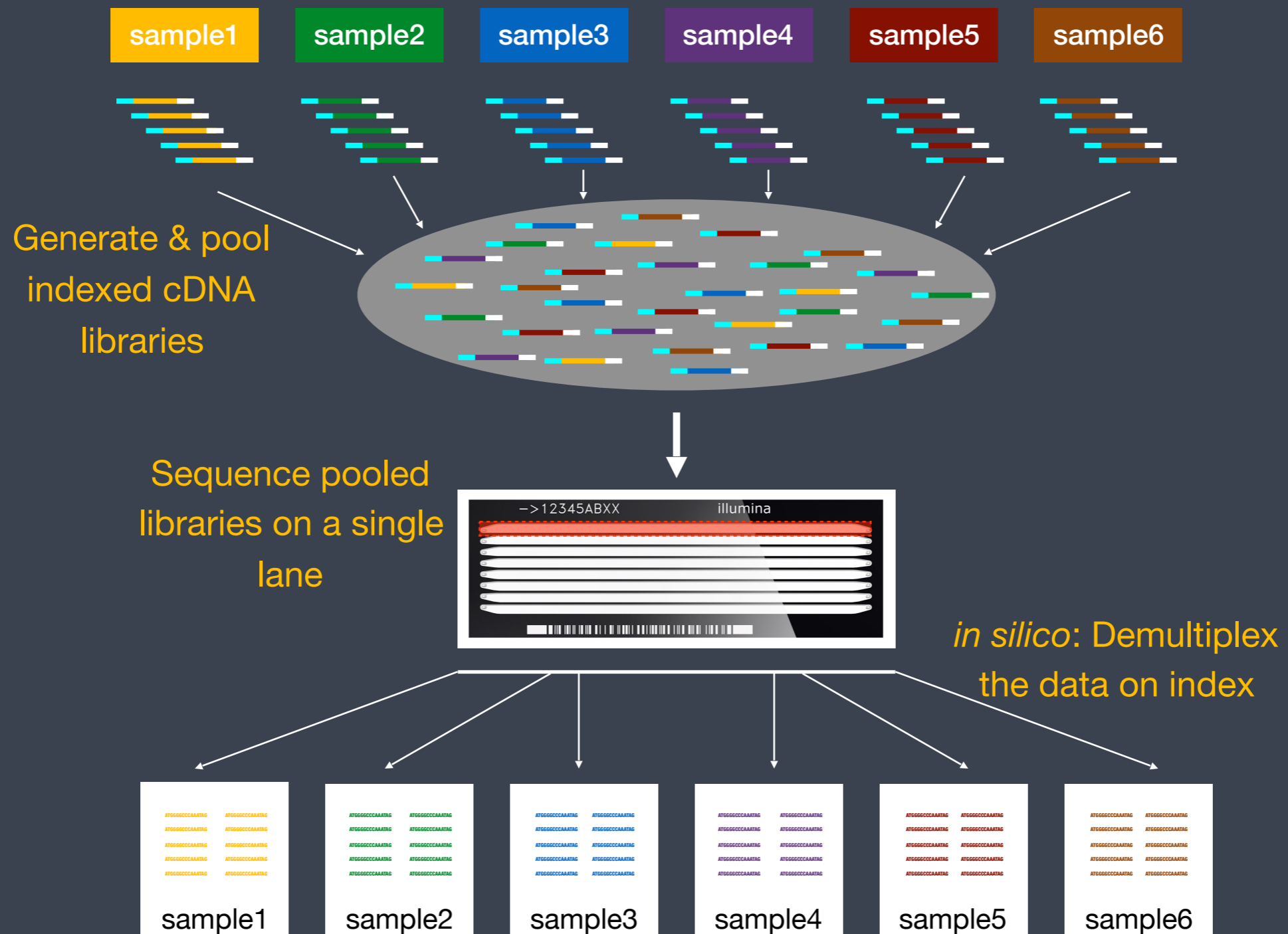
1. Read 1
2. Index Read 1 (i7)
3. Index Read 2 (i5)
4. Read 2

HiSeq/MiSeq (4 color)

- A&C read on one camera
- G&T read on other

NextSeq (2 color)

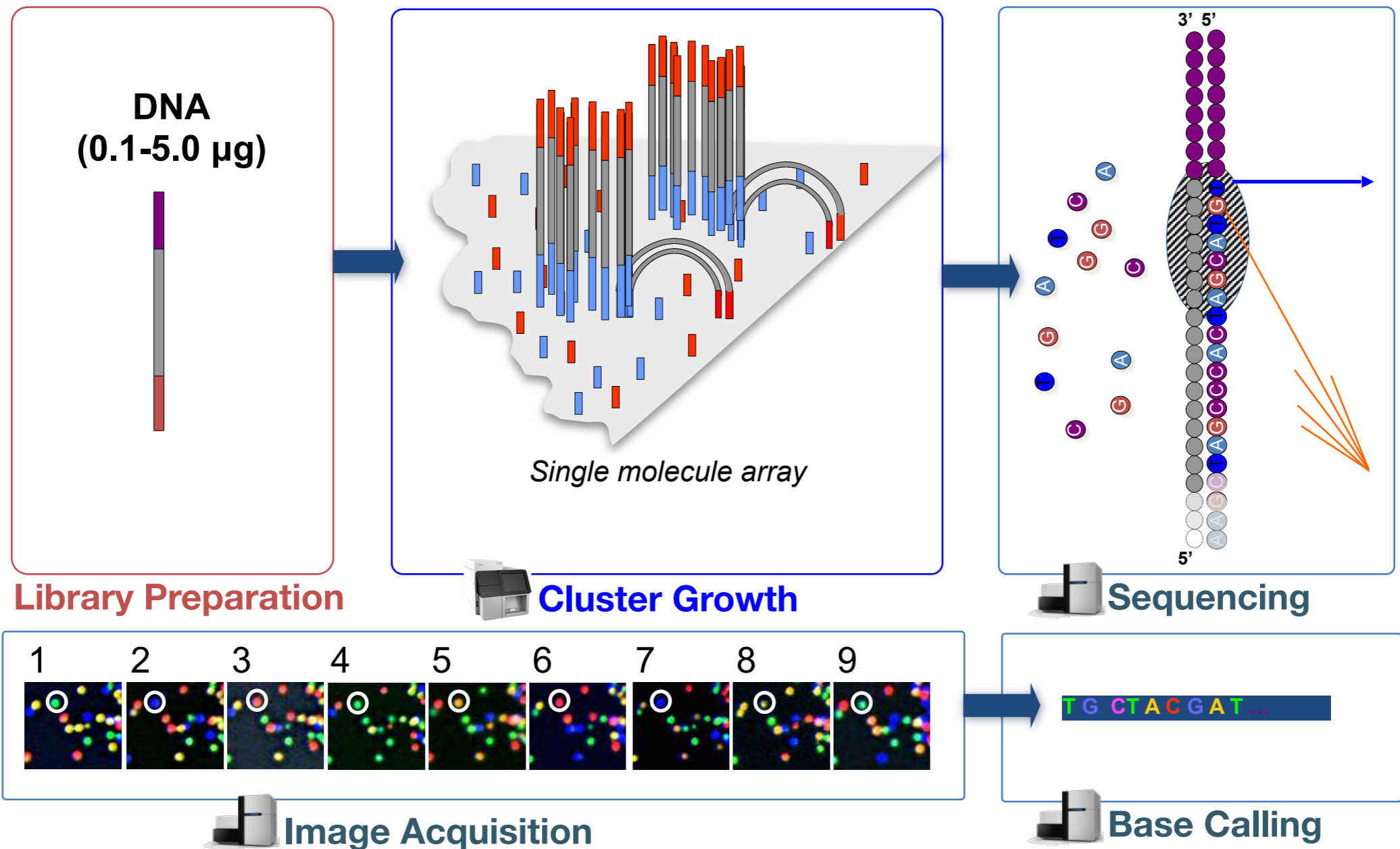
Multiplexing (barcodes and indices)



Multiplexing (barcodes and indices)

- Multiplexing allows optimal use of sequencing reads.
- Charges for sequencing are usually per lane of the flow cell
 - HiSeq generates ~150 million reads per lane
 - NextSeq generates ~ 450 million reads (one lane instrument)
- For RNA-seq number of reads you need will depend on your experiment
 - 15 million standard for transcriptome (polyA selected)
 - 20 million standard for total RNA (rRNA depleted)

Make sure multiplexing libraries of similar size



Consider Cluster Size in Multiplexing

Multiplexing

- Pool samples based on dye based quantification
- Submit pool to core facility for sequencing
- Make all sequencing libraries in one batch



qPCR quantification
before sequencing

Final Thoughts

- Practice your library prep on a control sample.
- Be sure you understand each step in library prep.
- Talk to someone who has done the protocol before starting.

qPCR

Precise quantitation is key to effective sequencing!

Useful Websites

- support.illumina.com/
- seqanswers.com/
- core-genomics.blogspot.com/2012/04/how-do-spri-beads-work.html