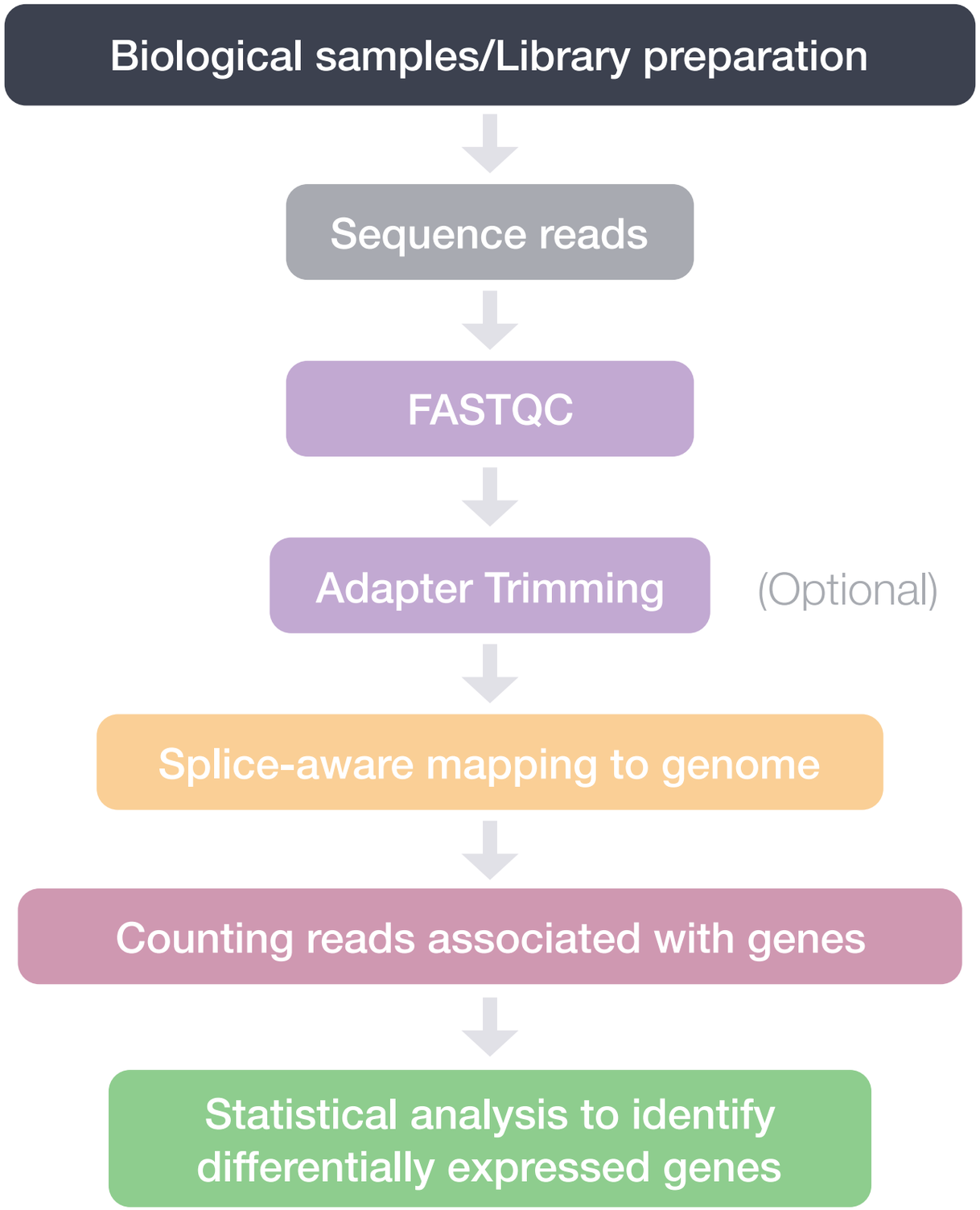
Sequencing Library Preparation

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

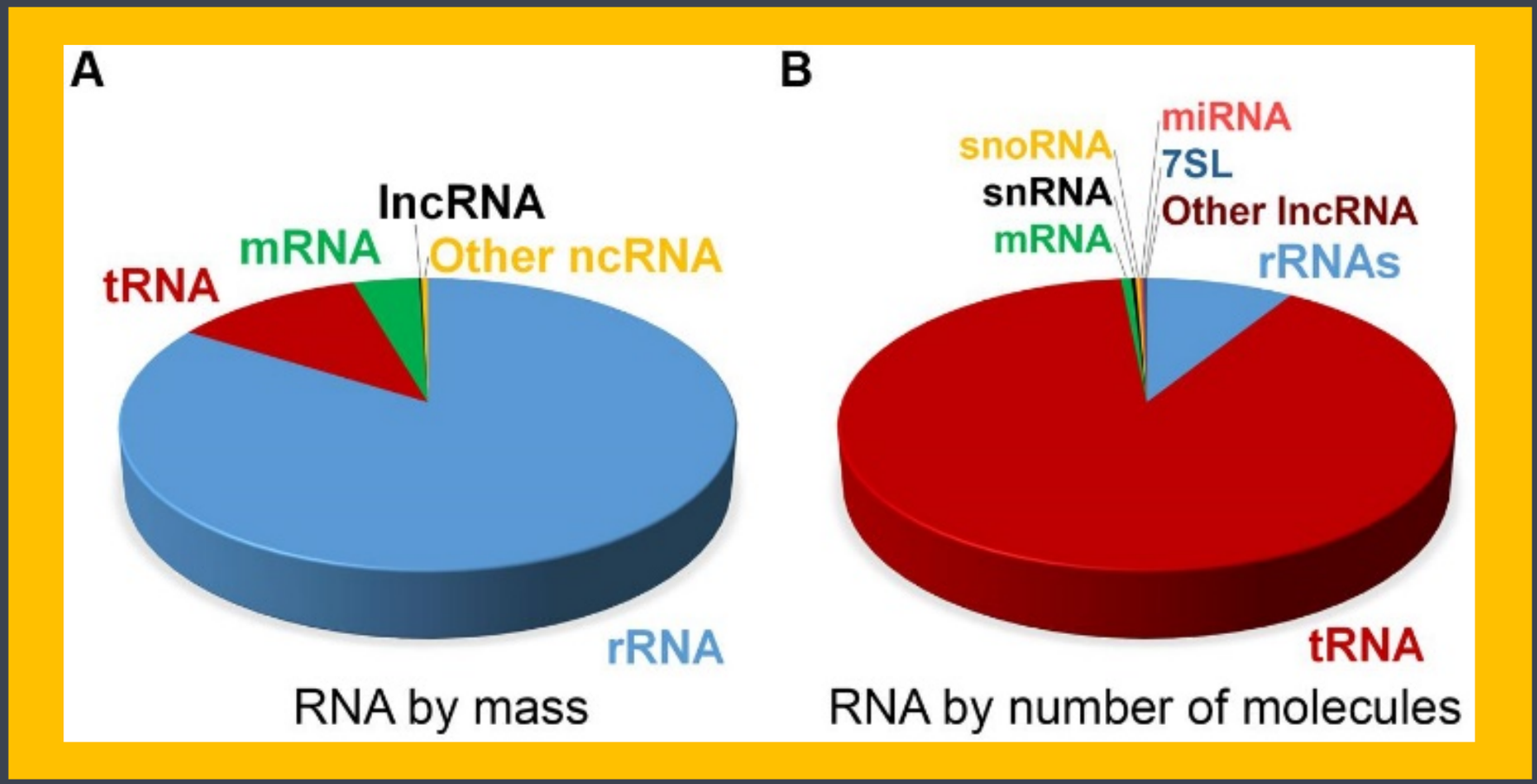
RNA-seq Workflow



Key steps in library preparation

- ✓ Starting Material
- ➢ Library amplification bias
- ➢ Multiplexing
- ➢ Sequencing read order & terminology

RNA enrichment



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

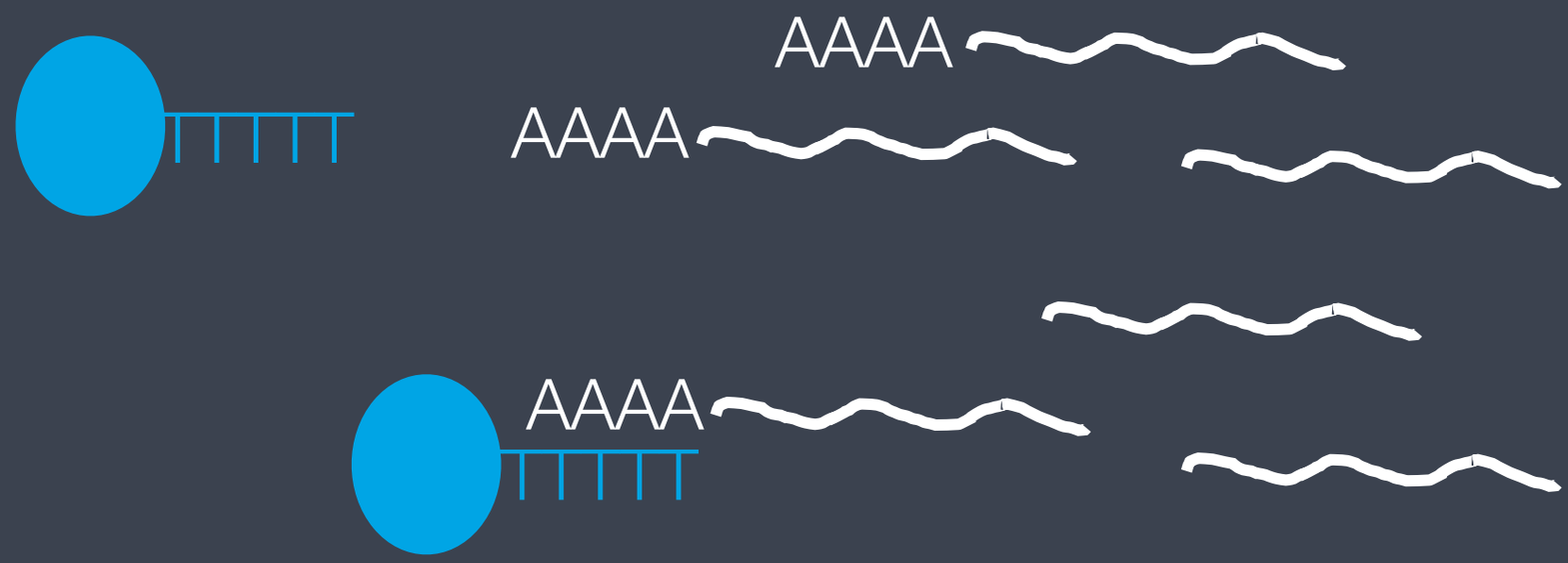
Front Genet. 2015 Jan 26;6:2

Purification and QC of RNA

- Start with highest quality RNA possible
- Accurately quantify RNA
- Assess quality of RNA

mRNA (polyA) Purification

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

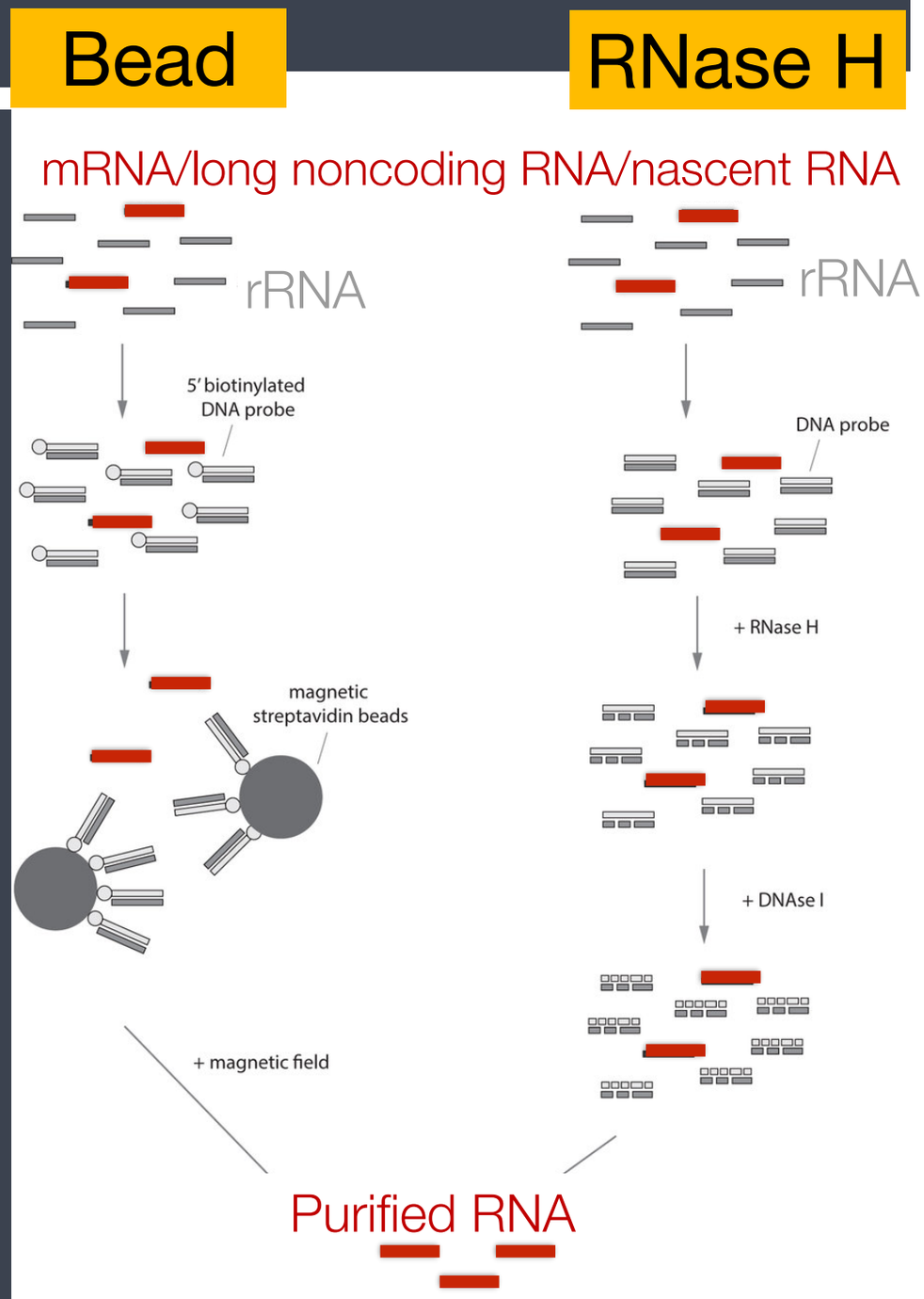


Transcripts Lost in polyA Purification

- 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

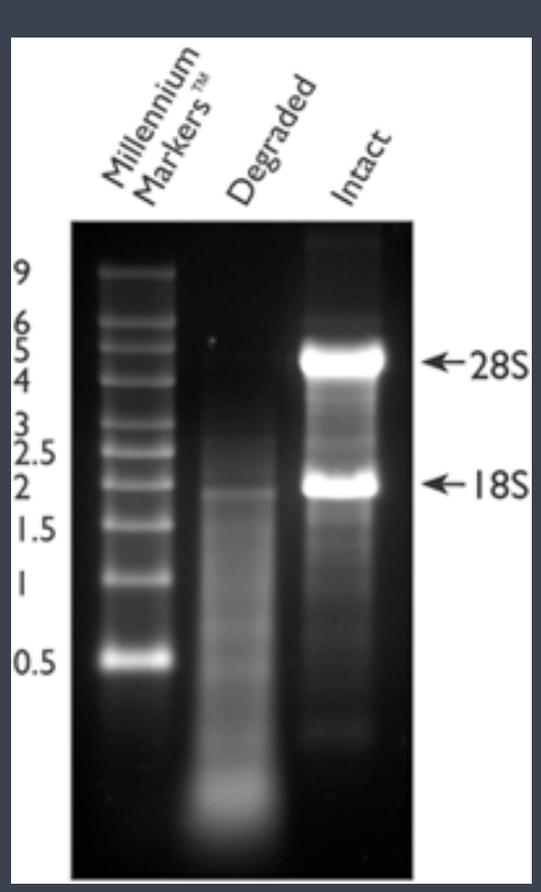
- Illumina: TruSeq
  - Probes hybridize rRNA on magnetic beads
  - RNA of interest remains in supernatant
- KAPA: RiboErase
  - Probes hybridize rRNA in solution
  - Hybrids are digested with RNase H
  - Probes digested with DNAse I



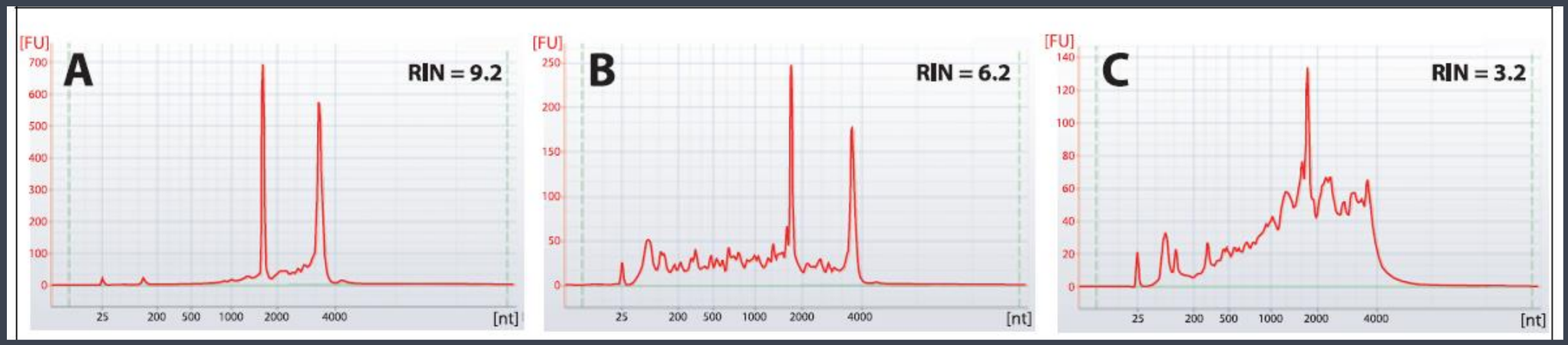
Modified from: Scientific Reports 6, article 37876 (2016)

RNA Quantitation & Quality

- Quantitation
  - Absorbance: Nano-drop (50-500 ng/ul)
    - Theoretically should 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)

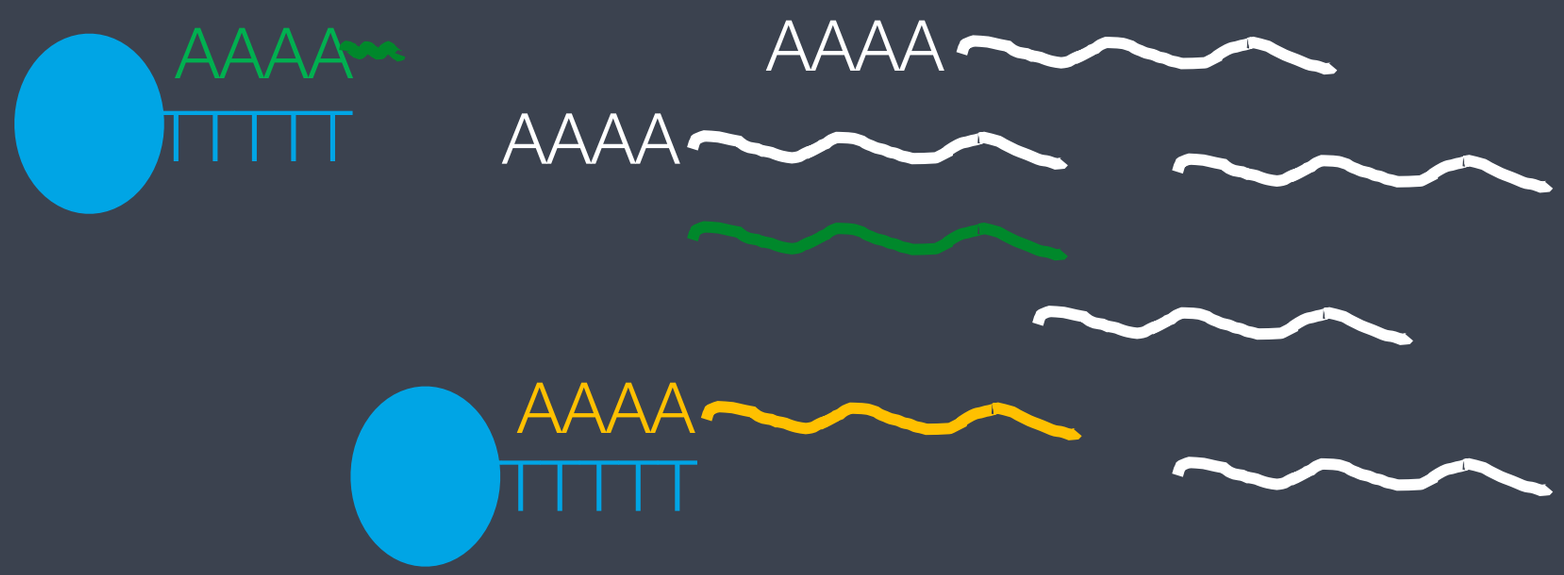


RNA quality



- 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

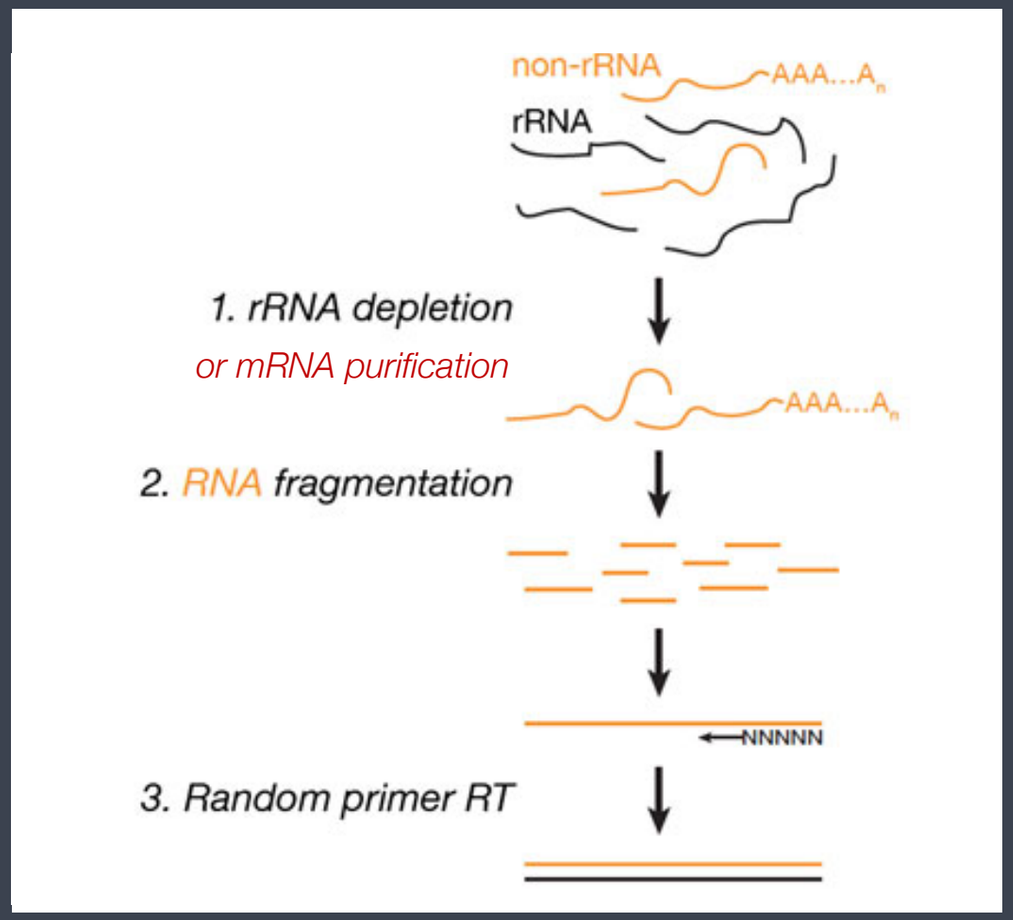
mRNA Purification of Degraded Samples

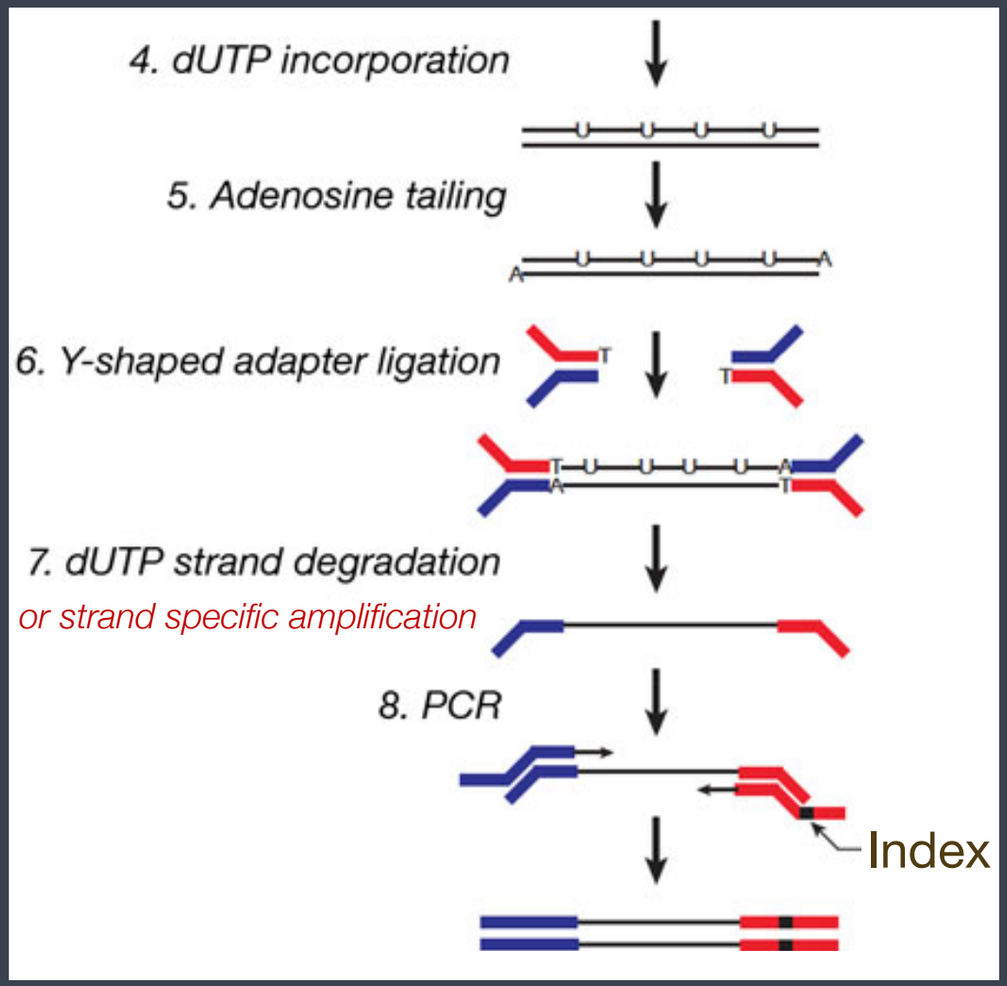


Transcript 1 Transcript 2

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

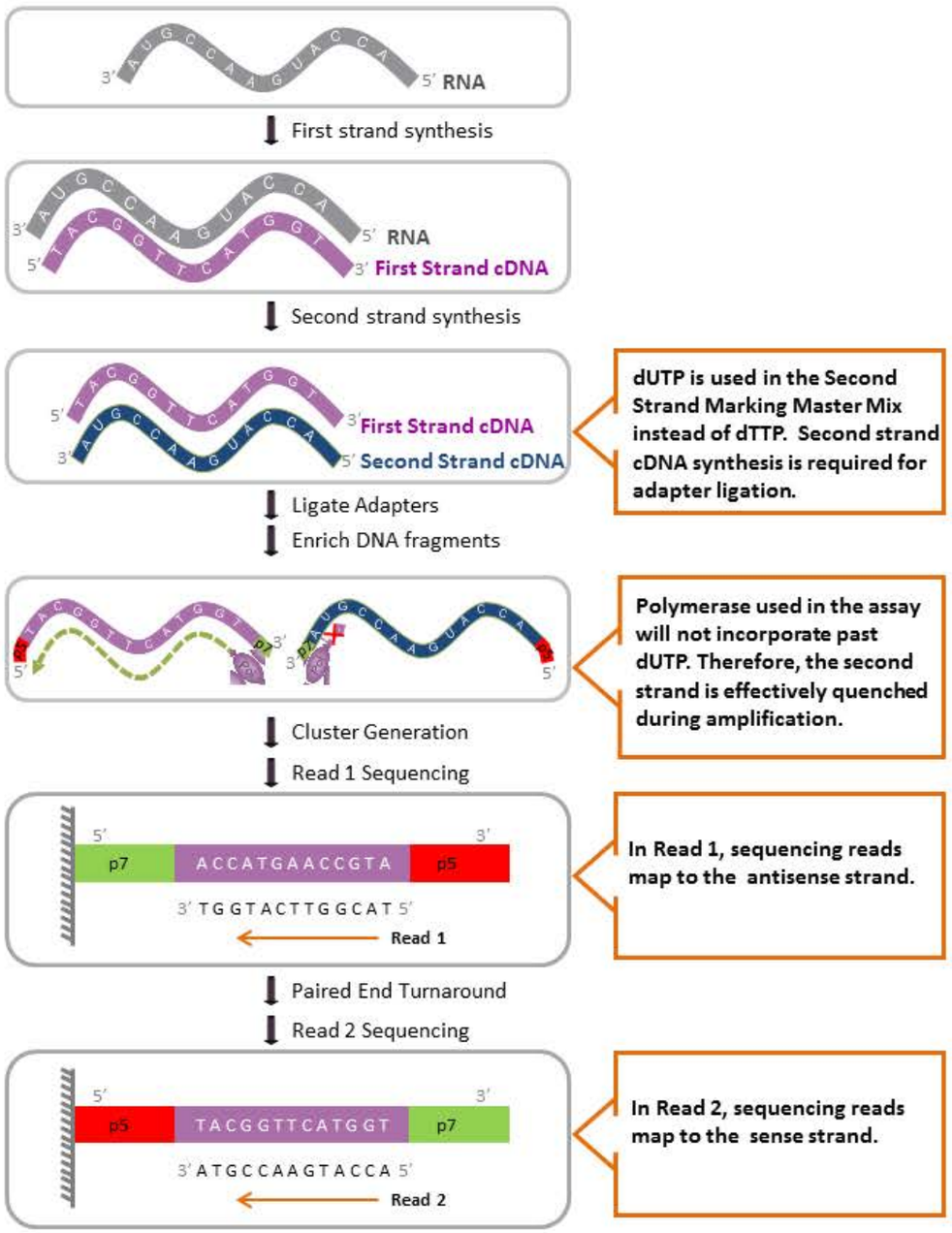
RNA-seq Stranded Library Prep (dUTP method)

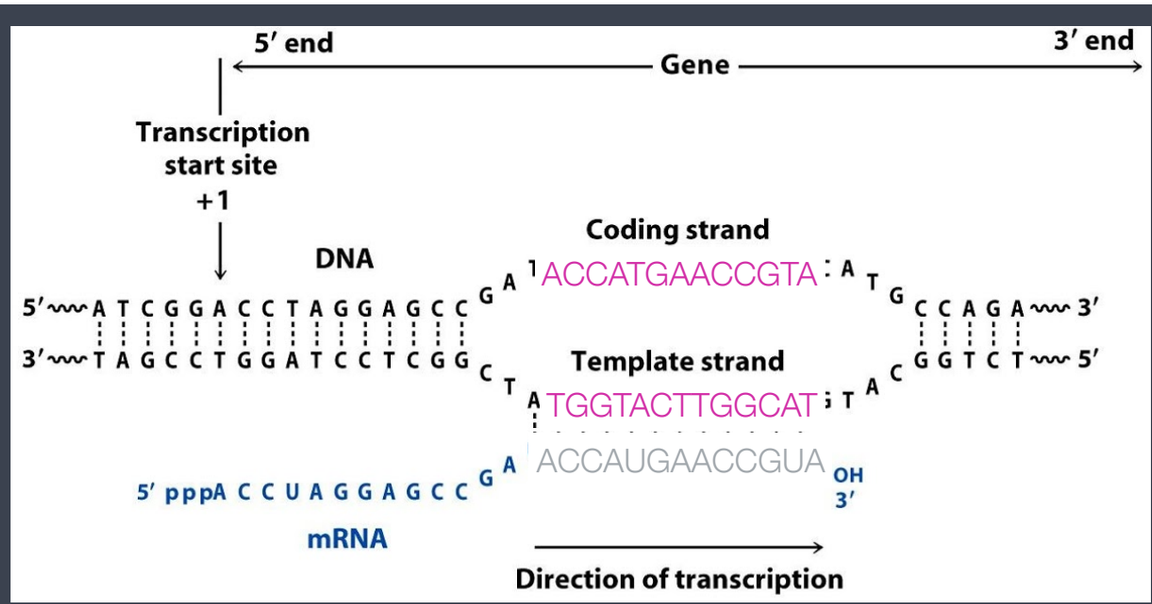




http://www.rna-seqblog.com/wp-content/uploads/2012/12/library-preparation.jpg

Library Strandedness

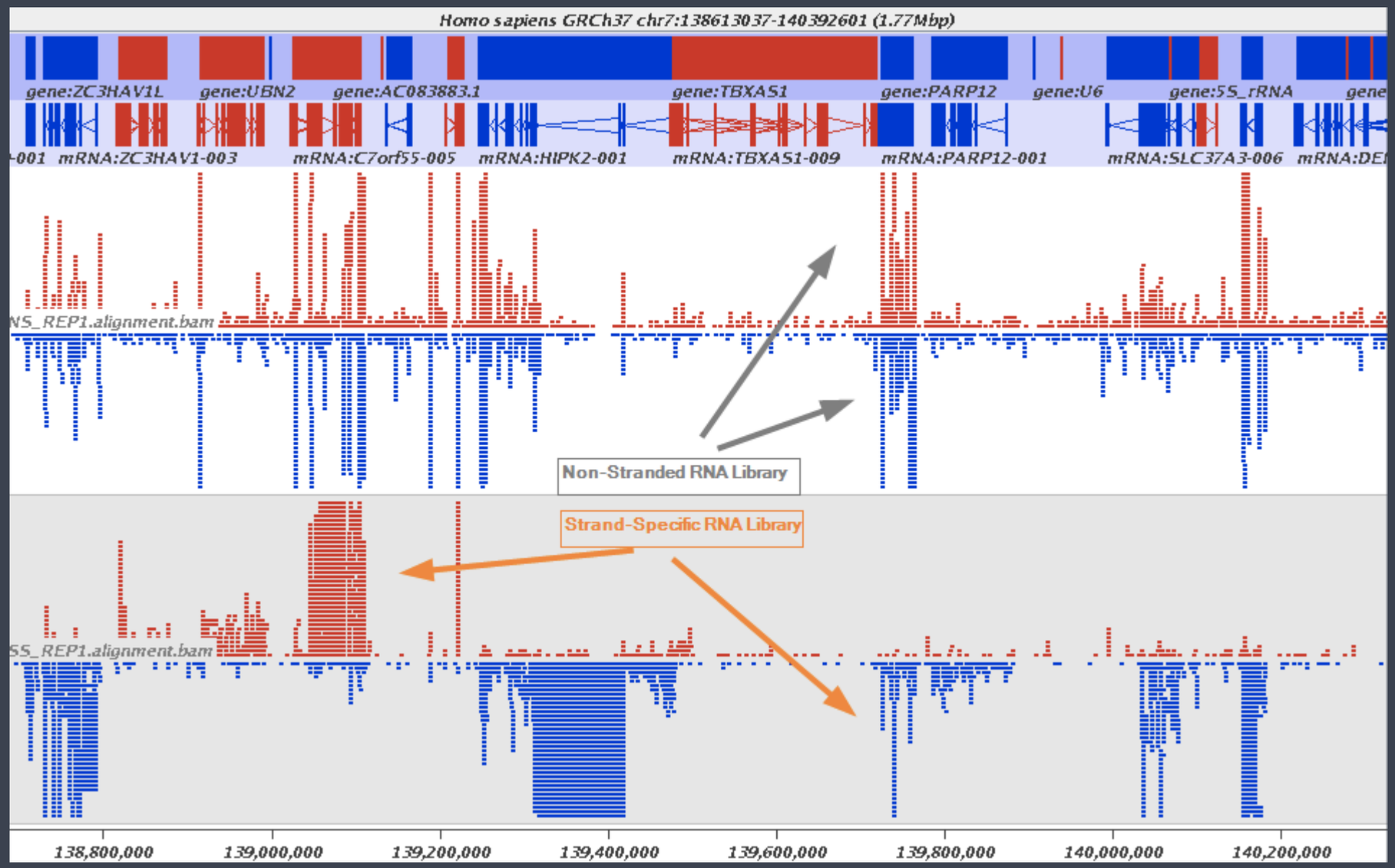




- Read alignment depends on direction of transcription
- “sense” strand of transcript can be on either the sense or antisense strand of the DNA

http://seqanswers.com/forums/showthread.php?t=44220

Library Strandedness



https://galaxyproject.org/tutorials/rb_rnaseq/

Key steps in library preparation

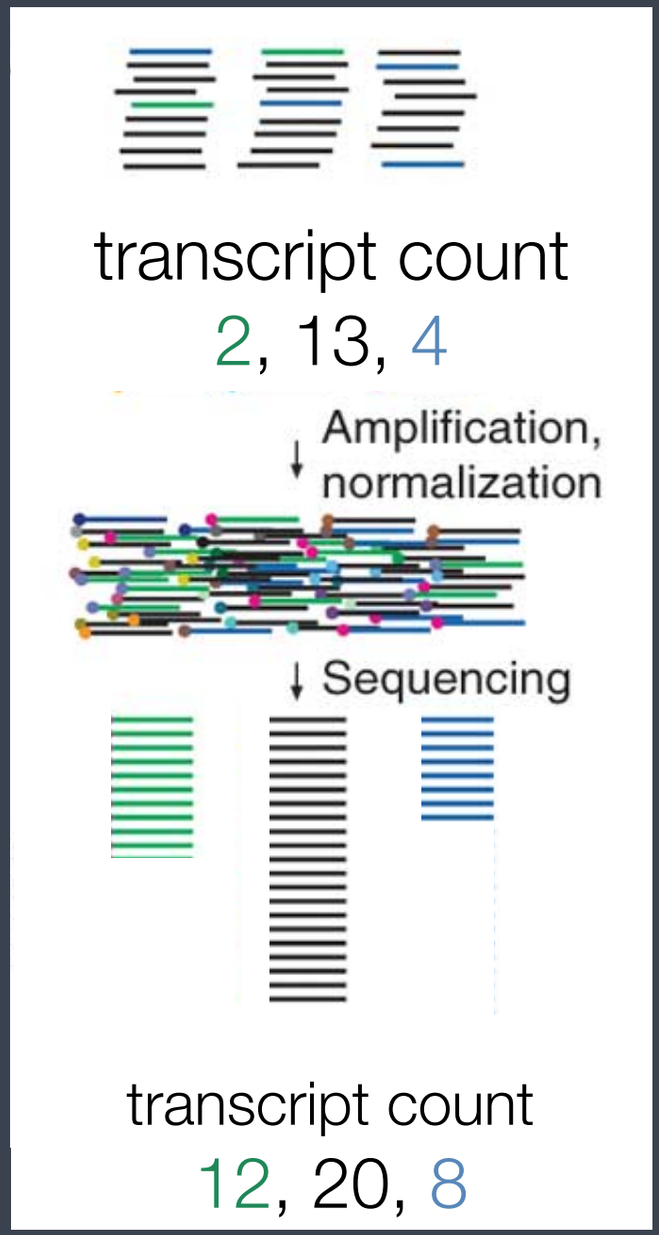
- ✓ Starting Material
- ✓ Library amplification bias
- ➢ Multiplexing
- ➢ Sequencing read order & terminology

Library Amplification Bias

- Final step of library prep is amplification

- Introduces library bias
  - Some products preferentially amplified

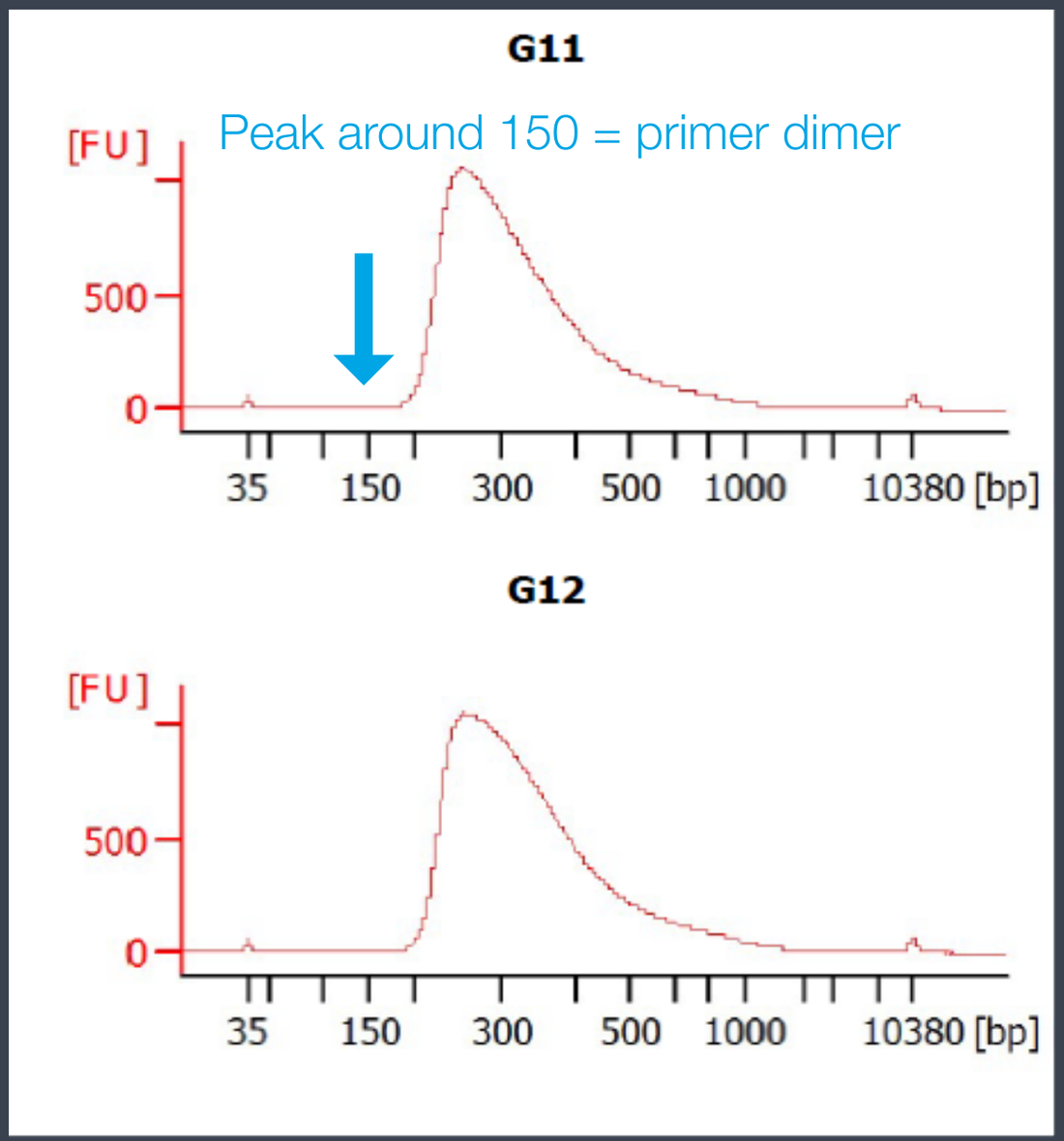
- Fewer cycles = less bias



Modified from: Nature Methods 9, 72-74 (2012)

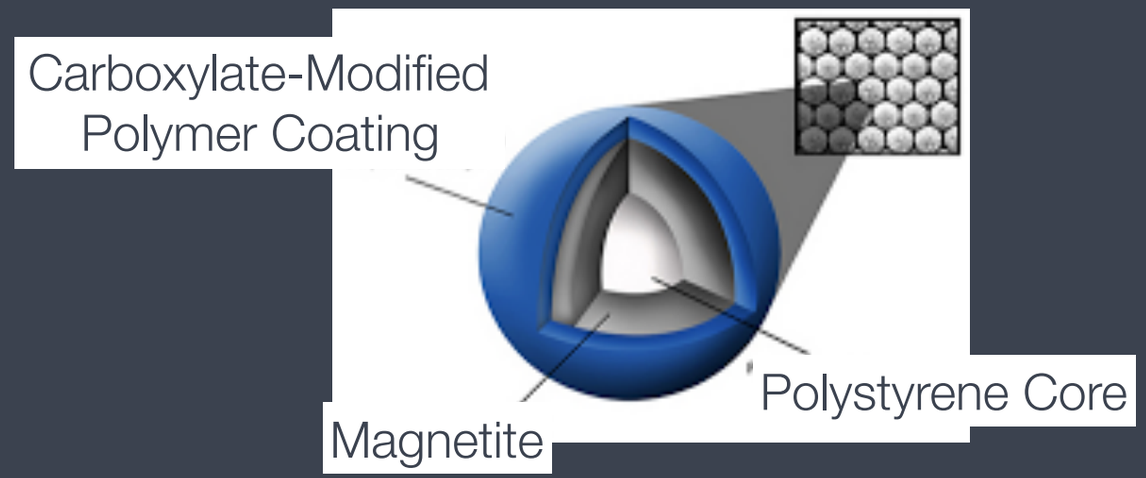
Library QC

- Quantification
  - Dye based
    - SYBR Green
    - Qubit / Quant-IT
- Size & Quality
  - Agilent Bioanalyzer
  - Size determination
  - Do not use for quantification

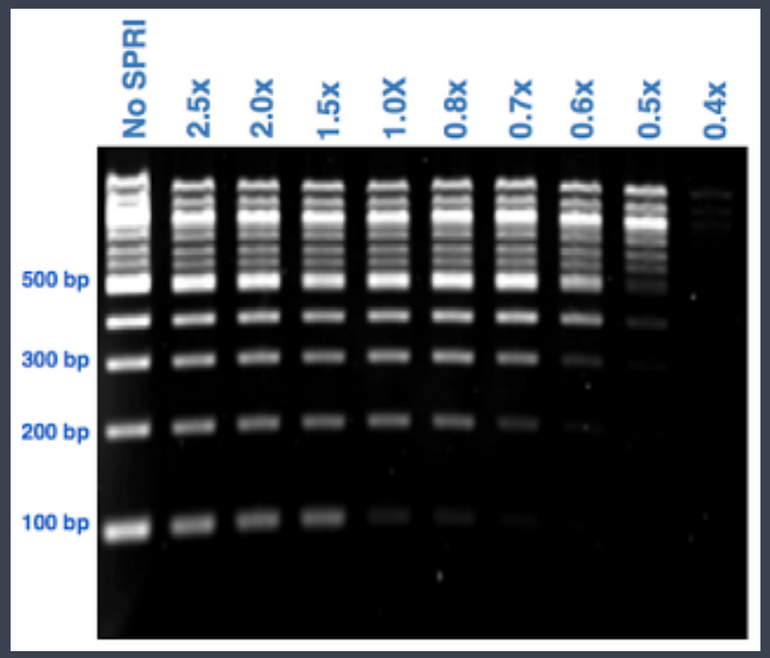


Size selection with SPRI beads

- Solid Phase Reverse Immobilization beads



- Carboxyl groups on surface bind DNA in the presence of crowding agents (PEG & NaCl)



http://core-genomics.blogspot.com/2012/04/how-do-spri-beads-work.html

Key steps in library preparation

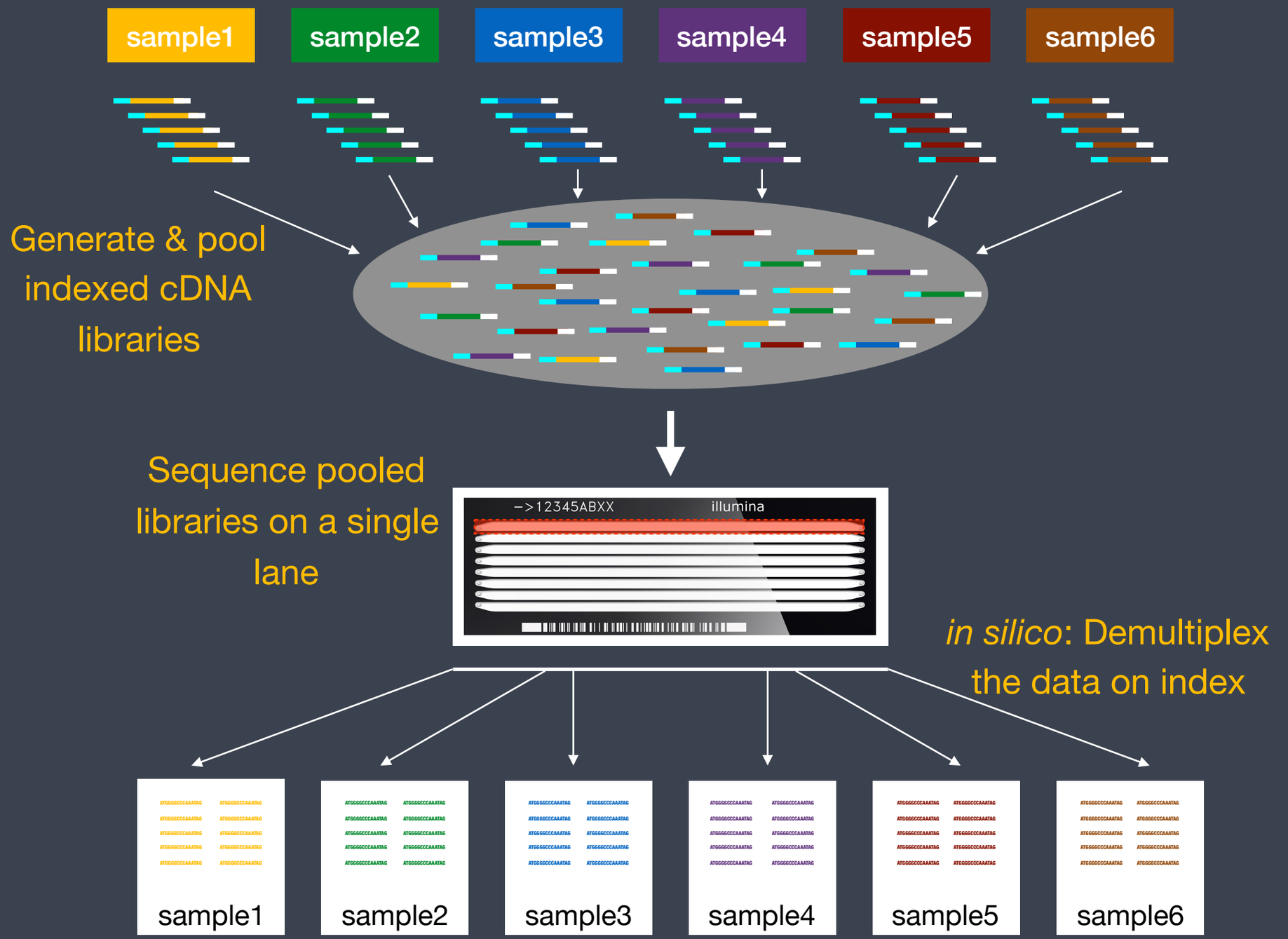
- ✓ Starting Material
- ✓ Library amplification bias
- ✓ Multiplexing
- ➢ Sequencing read order & terminology

Multiplexing (barcodes and indices)

- Multiplexing allows optimal use of reads you will get
- 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

Multiplexing (barcodes and indices)



Multiplexing

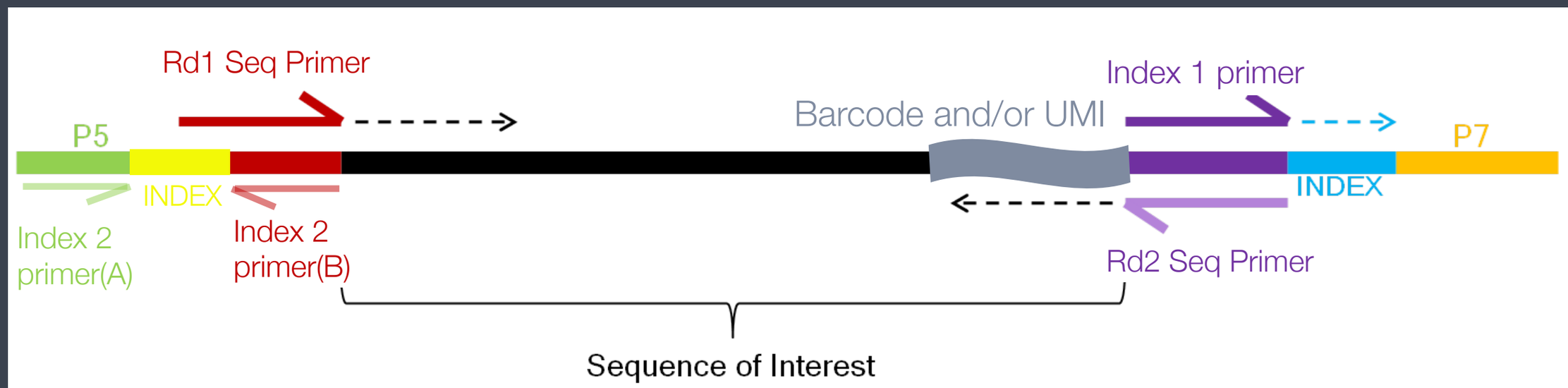
- Pool samples based on dye based quantification
- Submit pool to core facility for sequencing
- Make all sequencing libraries in one batch

qPCR quantification
before sequencing

Key steps in library preparation

- ✓ Starting Material
- ✓ Library amplification bias
- ✓ Multiplexing
- ✓ Sequencing read order & terminology

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)

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