

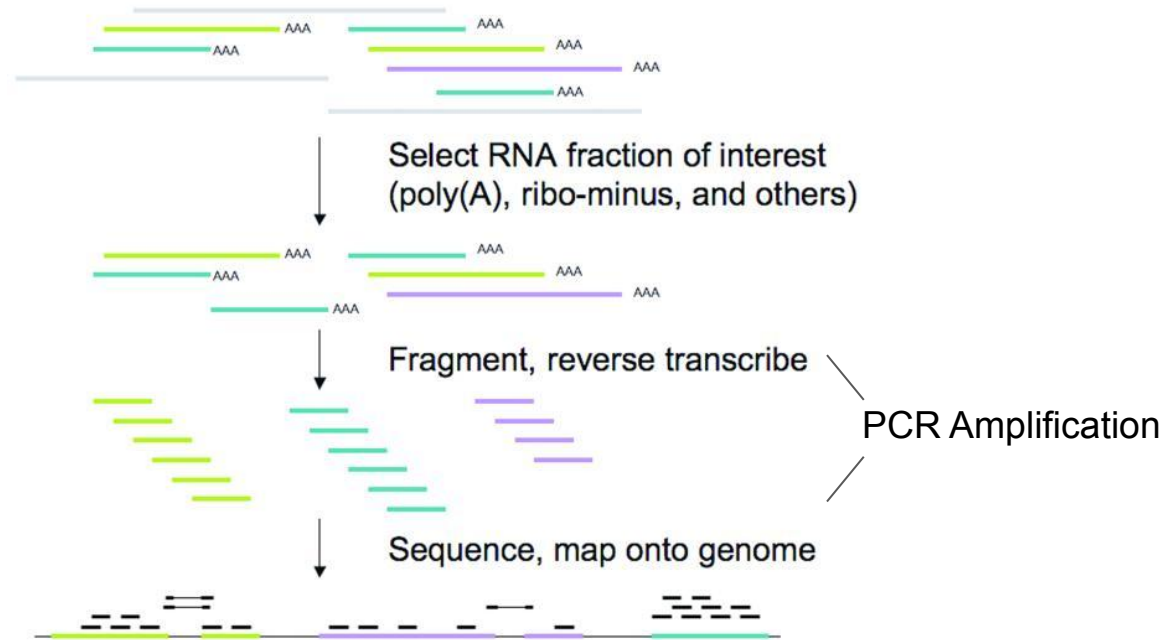


Introduction to RNA-seq

The CCDL

There is no optimal pipeline for the variety of different applications and analysis scenarios in which RNA-seq can be used. Scientists plan experiments and adopt different analysis strategies depending on the organism being studied and their research goals.

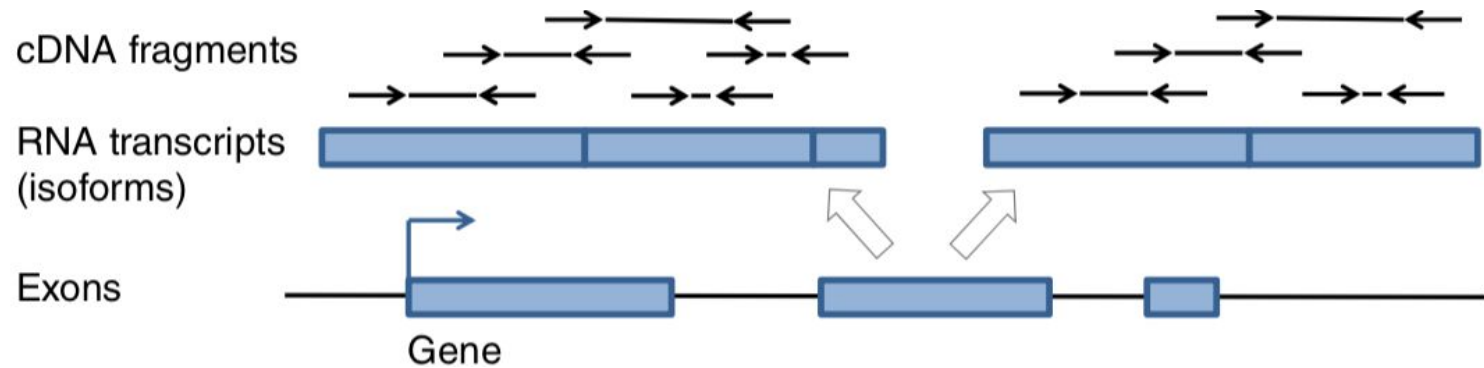
- [Conesa et al. 2016](#)



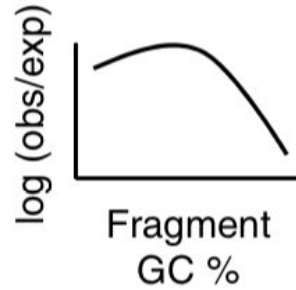
Biases to be aware of

- Library size or sequencing depth - the total number of reads is not always equivalent between samples
- Gene length - longer genes are more likely to be observed

Abundance measures like TPM (Transcripts Per Million)
take this into account



Fragment
sequence bias
(PCR amplification)

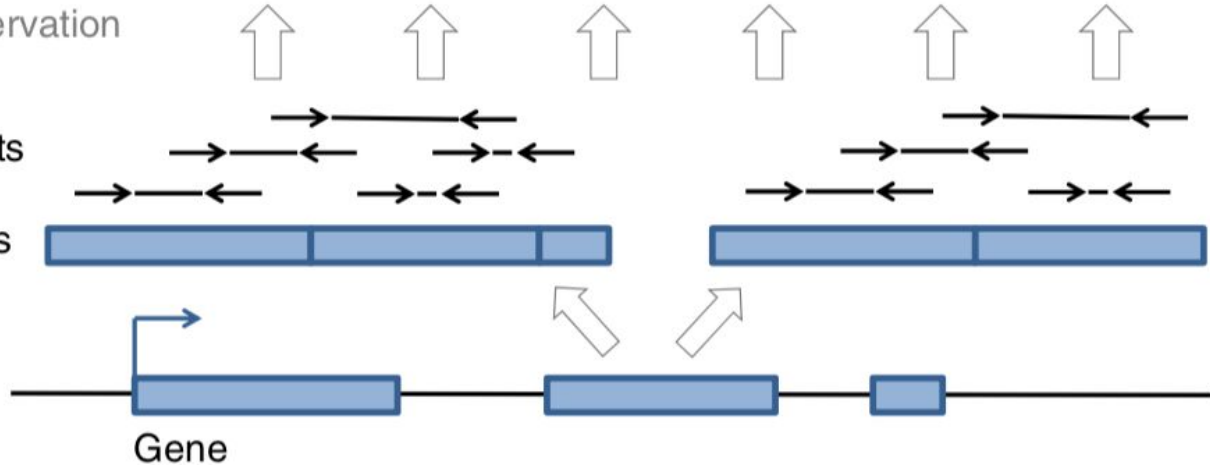


Biases on observation
of fragments

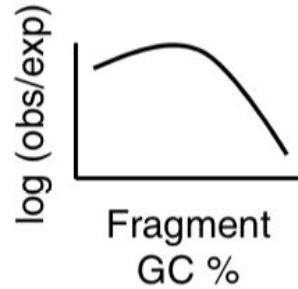
cDNA fragments

RNA transcripts
(isoforms)

Exons



Fragment
sequence bias
(PCR amplification)



Read start bias
(random hexamer priming)

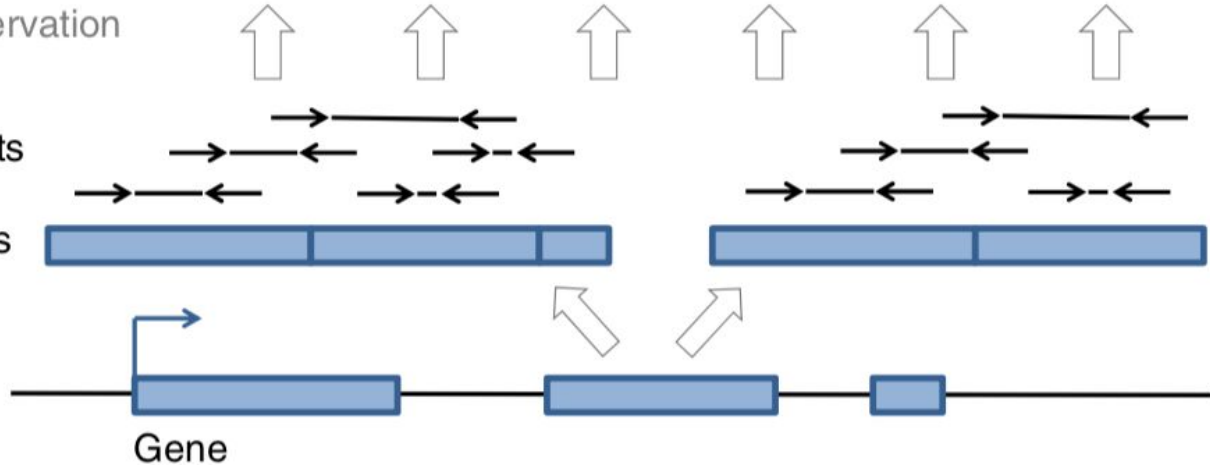


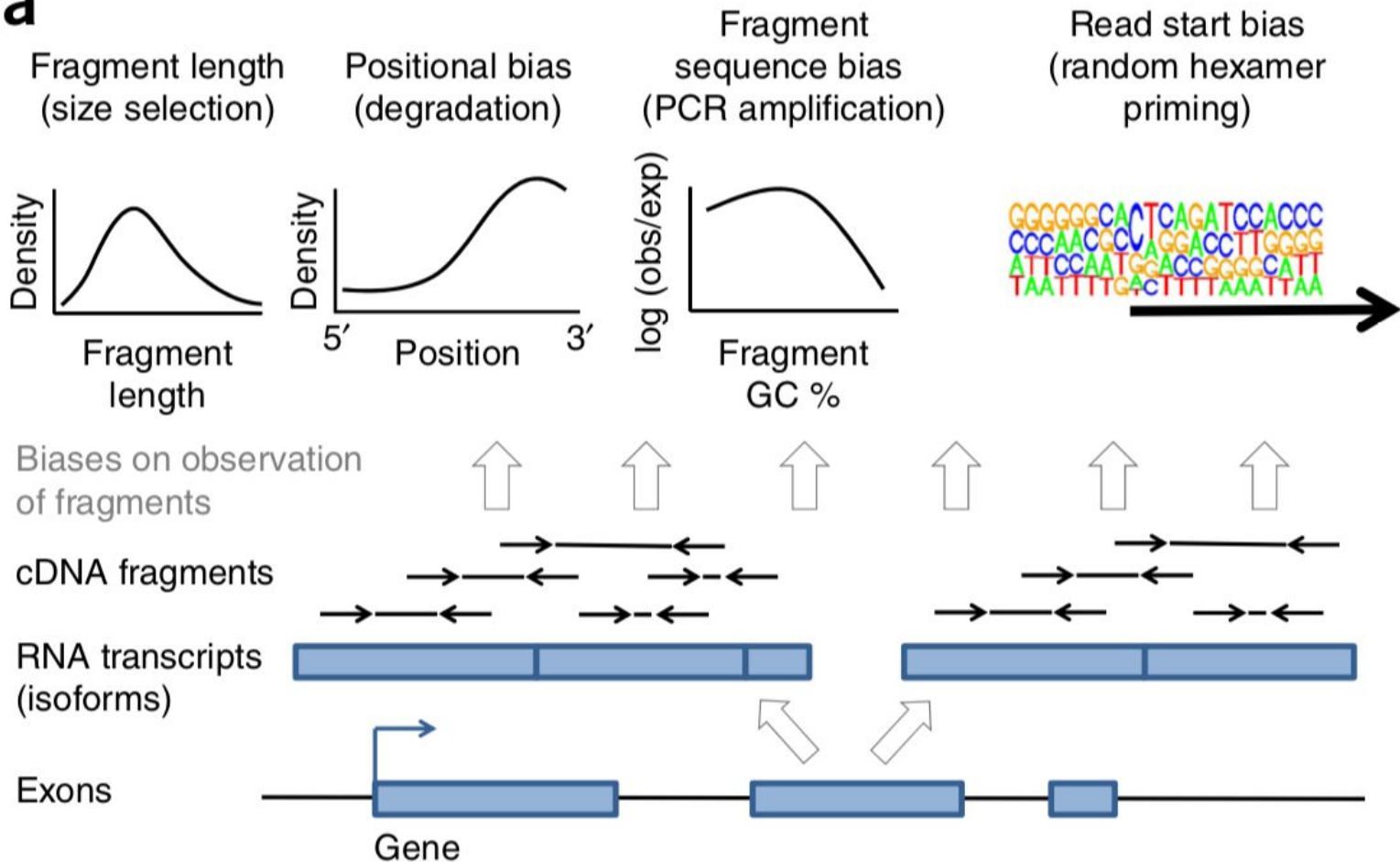
Biases on observation
of fragments

cDNA fragments

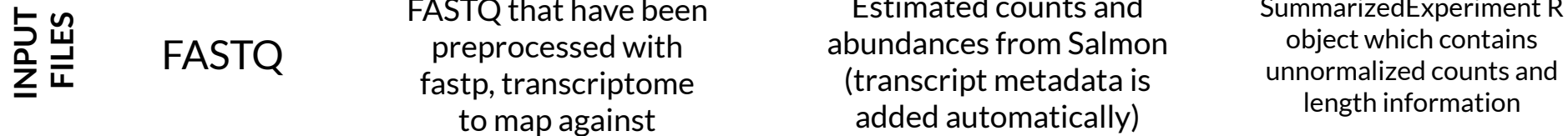
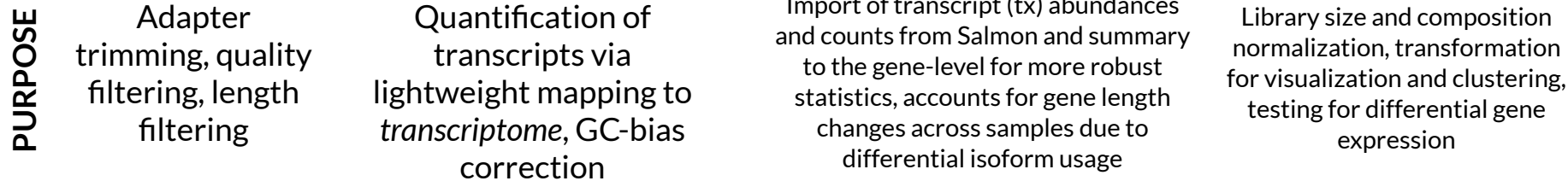
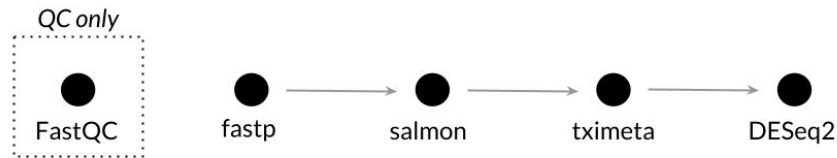
RNA transcripts
(isoforms)

Exons




a

Overview of pipeline



What you'll learn to do in this module

- Perform quality control checks with FastQC ([Andrews](#))
 - Perform FASTQ preprocessing with fastp ([Chen et al. 2018](#))
 - Quantify transcripts with Salmon ([Patro et al. 2017](#))
 - Import quantification estimates with tximeta and summarize to the gene level ([Love et al. 2020](#); [Soneson et al. 2015](#))
 - Perform exploratory data analysis with DESeq2 ([Love et al. 2014](#))
 - Perform differential expression analysis with DESeq2
 - Make fancy volcano plots and fancy heatmaps ([Blighe et al.](#); [Gu 2016](#))
- 

Tool-specific tutorials

[Getting Started with Salmon](#)

[Tximeta: transcript quantification import with automatic metadata](#)

Note: if you are not using Salmon, you can't use tximeta (for now) so you will want to look at tximport: [Importing transcript abundance datasets with tximport](#)

[Analyzing RNA-seq data with DESeq2](#)



Links to follow-up information

[StatQuest Video: A Gentle Guide to RNA-seq](#)

[StatQuest Video: RPKM, FPKM, and TPM](#)

[StatQuest Video: DESeq2, part 1, Library Normalization](#)

[Hansen et al. Biases in Illumina transcriptome sequencing caused by random hexamer priming. *Nucleic Acid Research*. 2010.](#)

[Michigan State University Research Technology Support Facility “FastQC Tutorial & FAQ”](#)