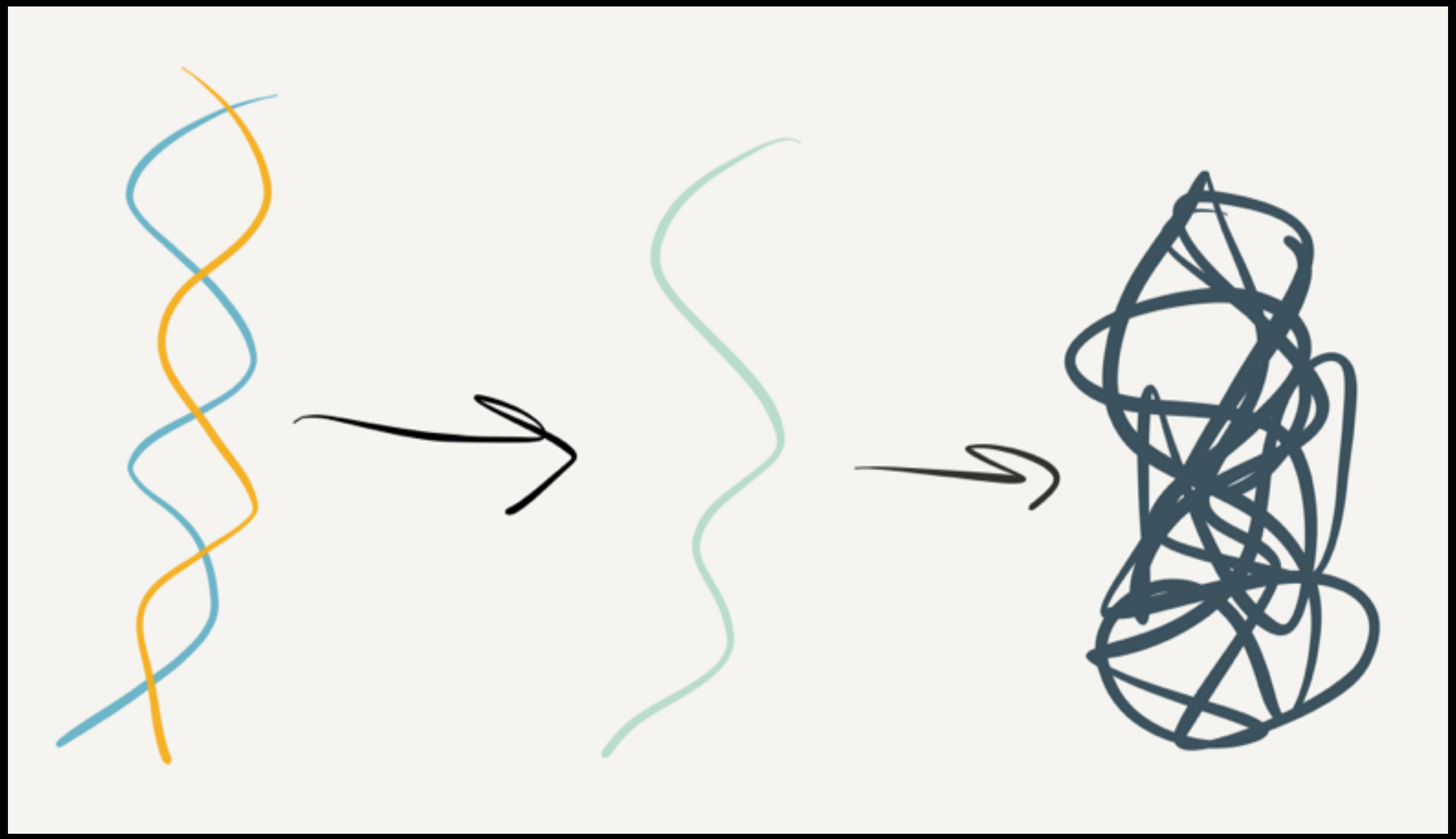
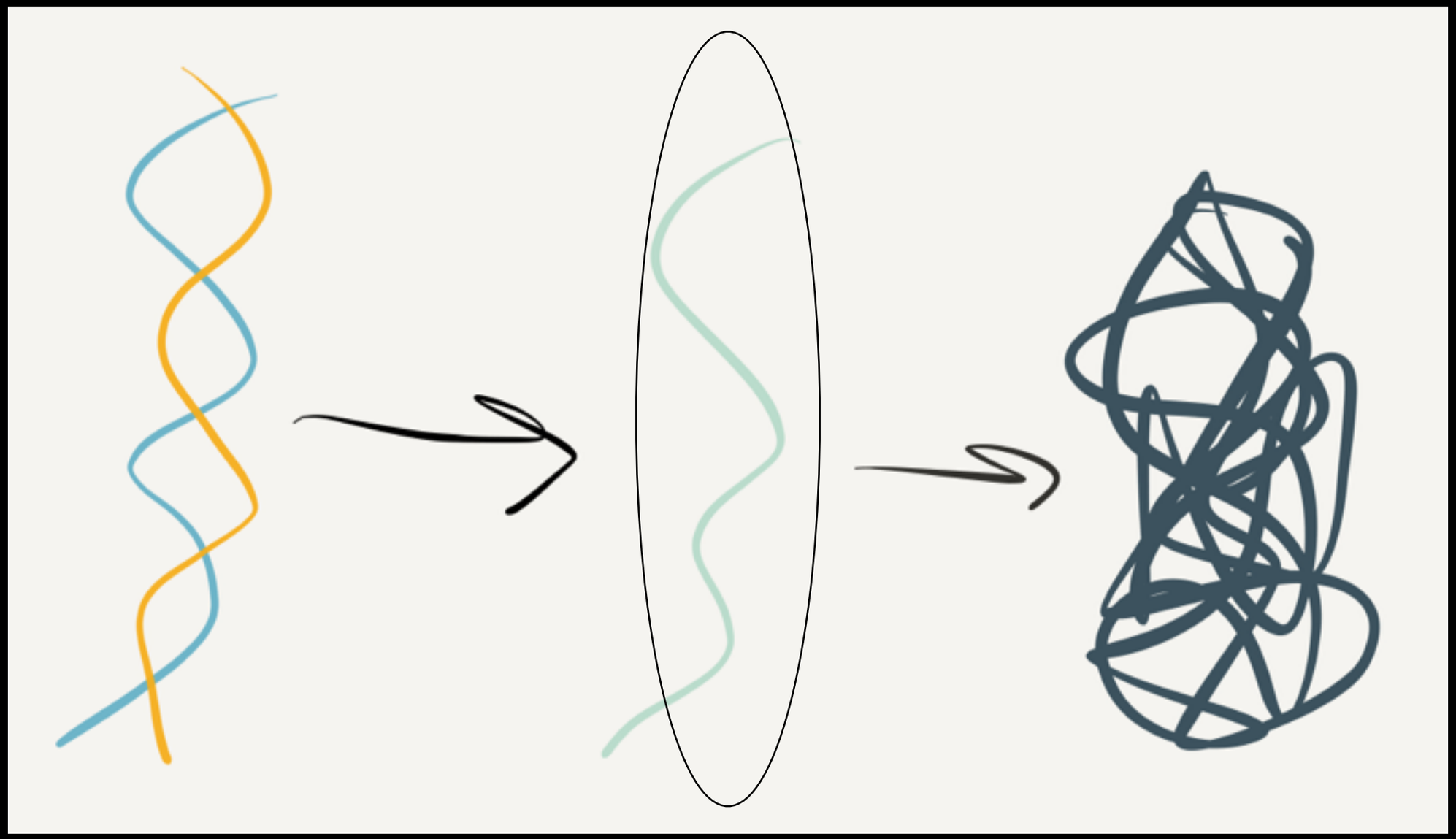
INTRO TO RNASEQ AND QUALITY CONTROL/TRIMMING

RNASEQ

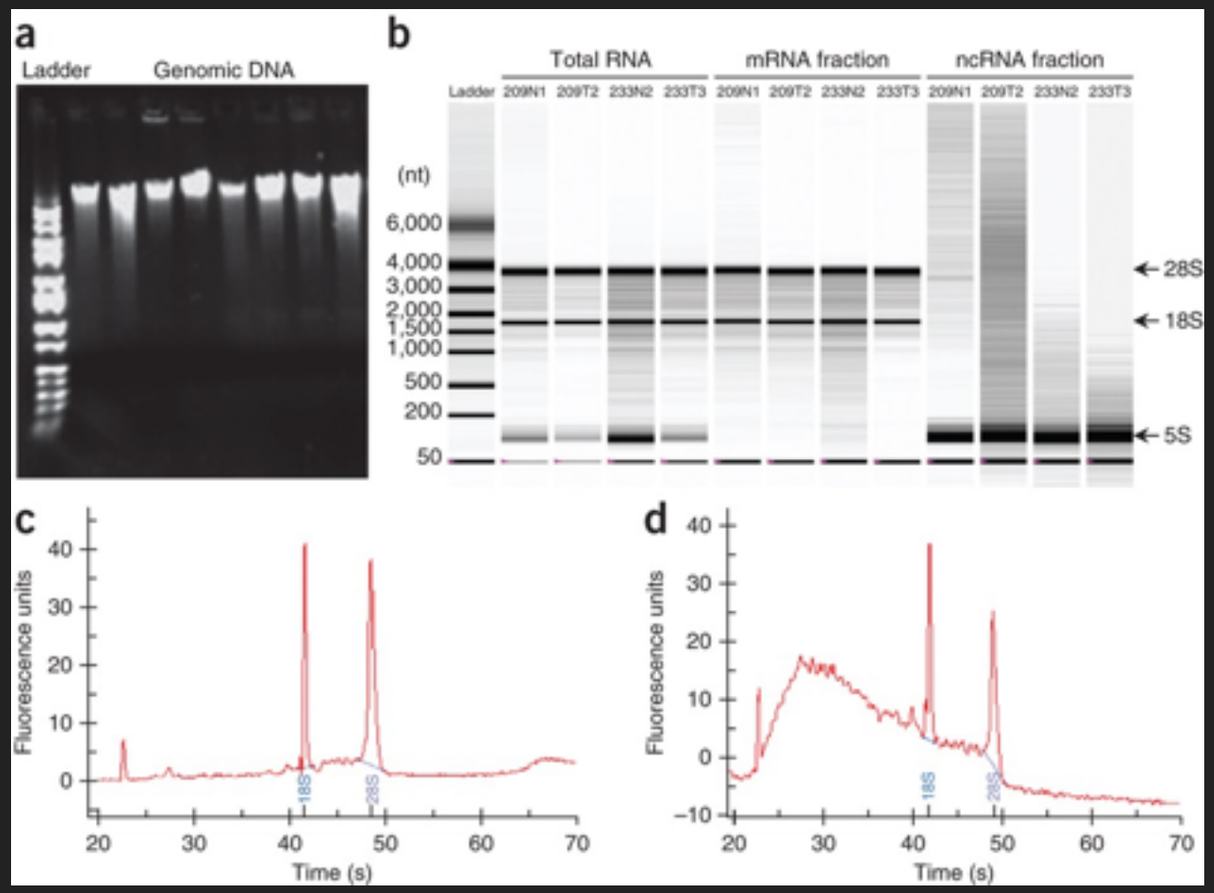


RNASEQ



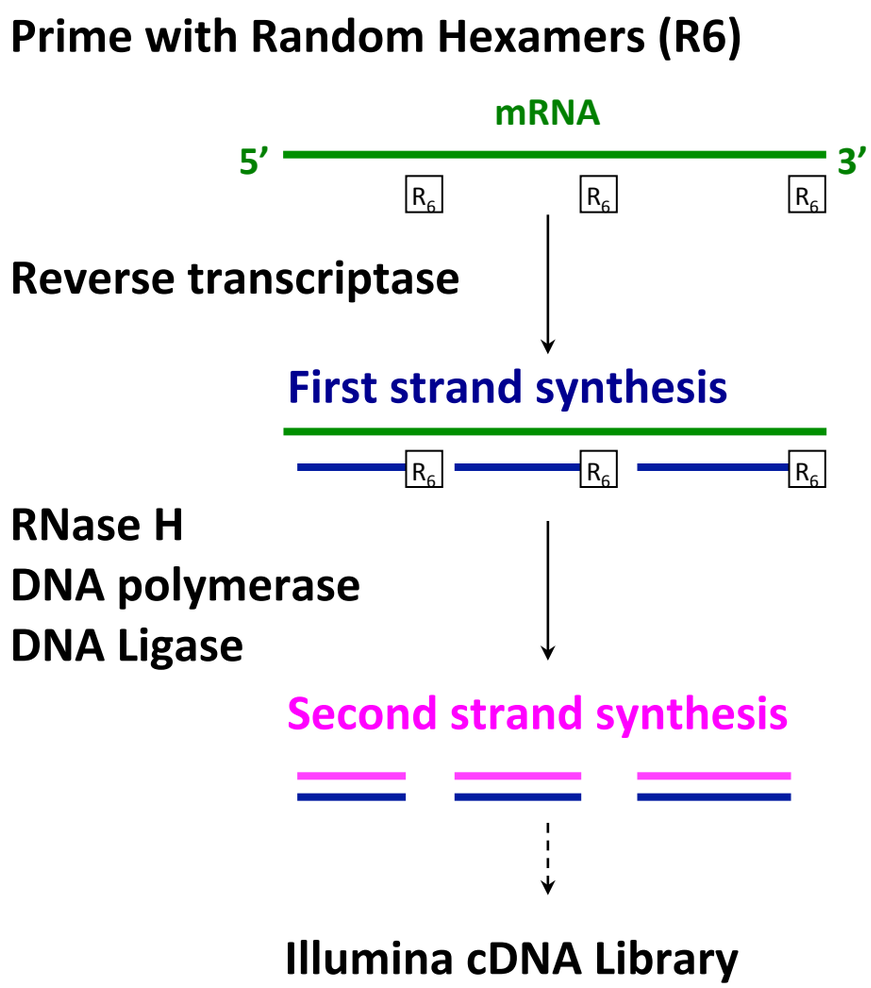
RNA EXTRACTION METHODS:

- Dependent on starting material: Tissue type, Amount
- End product: mRNA or total RNA?
  - Total RNA: Trizol vs. Columns (Promega, RNeasy, ect.)
  - mRNA: direct extraction or isolation from total RNA
- RNA from Bird Blood?
  - Contact: Loren Cassin Sackett <loren.sackett@gmail.com>



http://www.nature.com/nprot/journal/v8/n11/full/nprot.2013.141.html

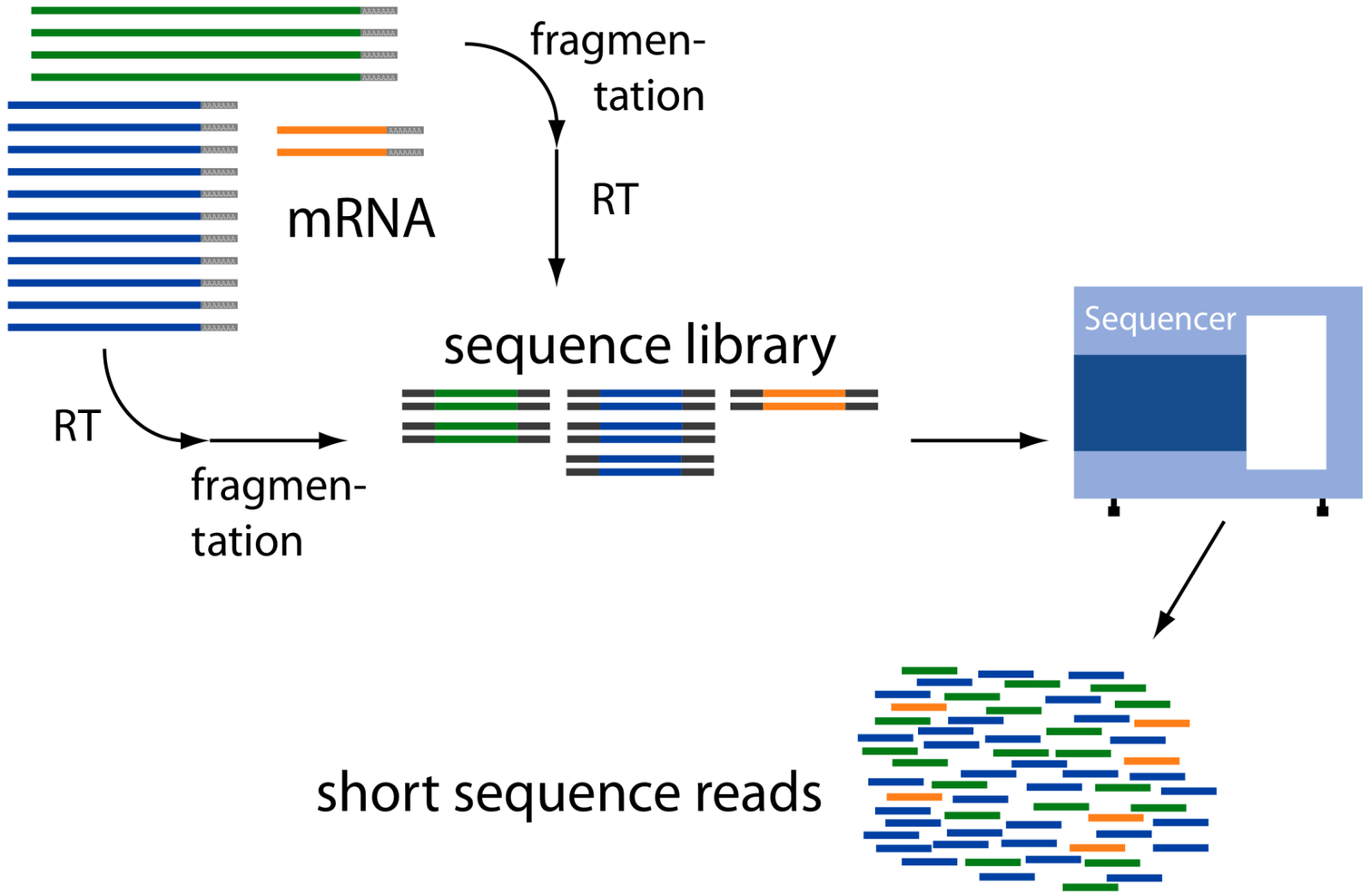
RNA-Seq: How do we make cDNA?



Slide courtesy of Joshua Levin, Broad Institute.

Slide from: https://github.com/trinityrnaseq/RNASeq_Trinity_Tuxedo_Workshop/blob/master/docs/rnaseq_workshop_slides.pdf

Overview of RNA-Seq



From: http://www2.fml.tuebingen.mpg.de/raetsch/members/research/transcriptomics.html

Slide from: https://github.com/trinityrnaseq/RNASeq_Trinity_Tuxedo_Workshop/blob/master/docs/rnaseq_workshop_slides.pdf

Common Data Formats for RNA-Seq

FASTA format:

```
>61DFRAAXX100204:1:100:10494:3070/1
AAACAACAGGGCACATTGTCACTCTTGTATTTGAAAAACACTTTCCGGCCAT
```

FASTQ format:

```
@61DFRAAXX100204:1:100:10494:3070/1
AAACAACAGGGCACATTGTCACTCTTGTATTTGAAAAACACTTTCCGGCCAT
+
ACCCCCCCCCCCCCCCCCCCCCCCCCCCCCCBC?CCCCCCCCC@@CACCCCCA
```

Read

Quality values

AsciiEncodedQual(x) = -10 * log10(Pwrong(x)) + 33

AsciiEncodedQual (‘C’) = **64**

So, Pwrong(‘C’) = 10^((**64**-33/ (-10)) = 10^-3.4 = **0.0004**

Slide from: https://github.com/trinityrnaseq/RNASeq_Trinity_Tuxedo_Workshop/blob/master/docs/rnaseq_workshop_slides.pdf

Paired-end Sequences

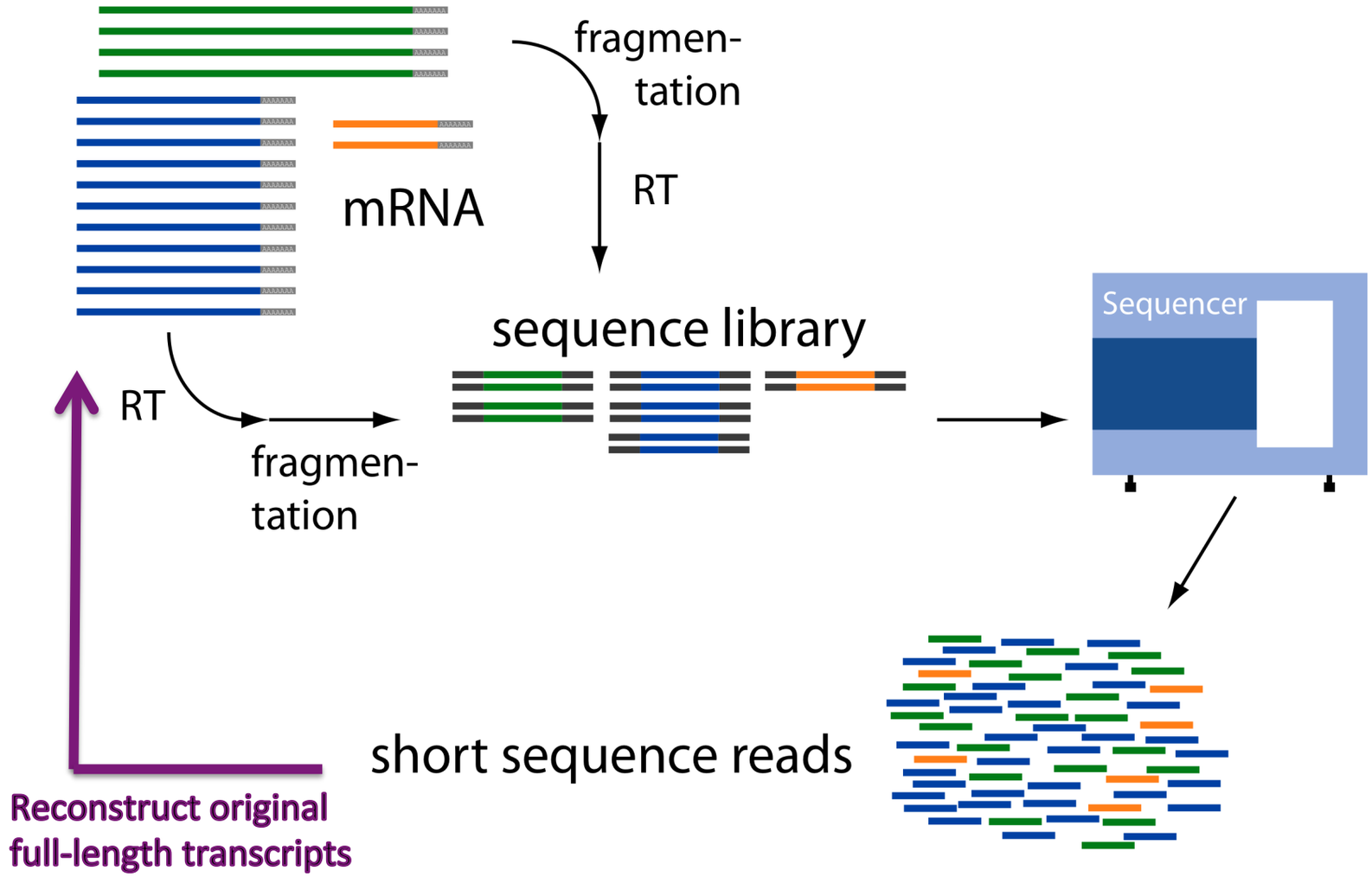
Two FastQ files, read name indicates left (/1) or right (/2) read of paired-end

```
@61DFRAAXX100204:1:100:10494:3070/1
AAACAACAGGGCACATTGTCACTCTTGTATTTGAAAAACACTTTCCGGCCAT
+
ACCCCCCCCCCCCCCCCCCCCCCCCCCCCCCBC?CCCCCCCCC@@CACCCCCA
```

```
@61DFRAAXX100204:1:100:10494:3070/2
CTCAAATGGTTAATTCTCAGGCTGCAAATATTCGTTCAGGATGGAAGAACA
+
C<CCCCCCCACCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCBCCCC
```

Slide from: https://github.com/trinityrnaseq/RNASeq_Trinity_Tuxedo_Workshop/blob/master/docs/rnaseq_workshop_slides.pdf

Overview of RNA-Seq

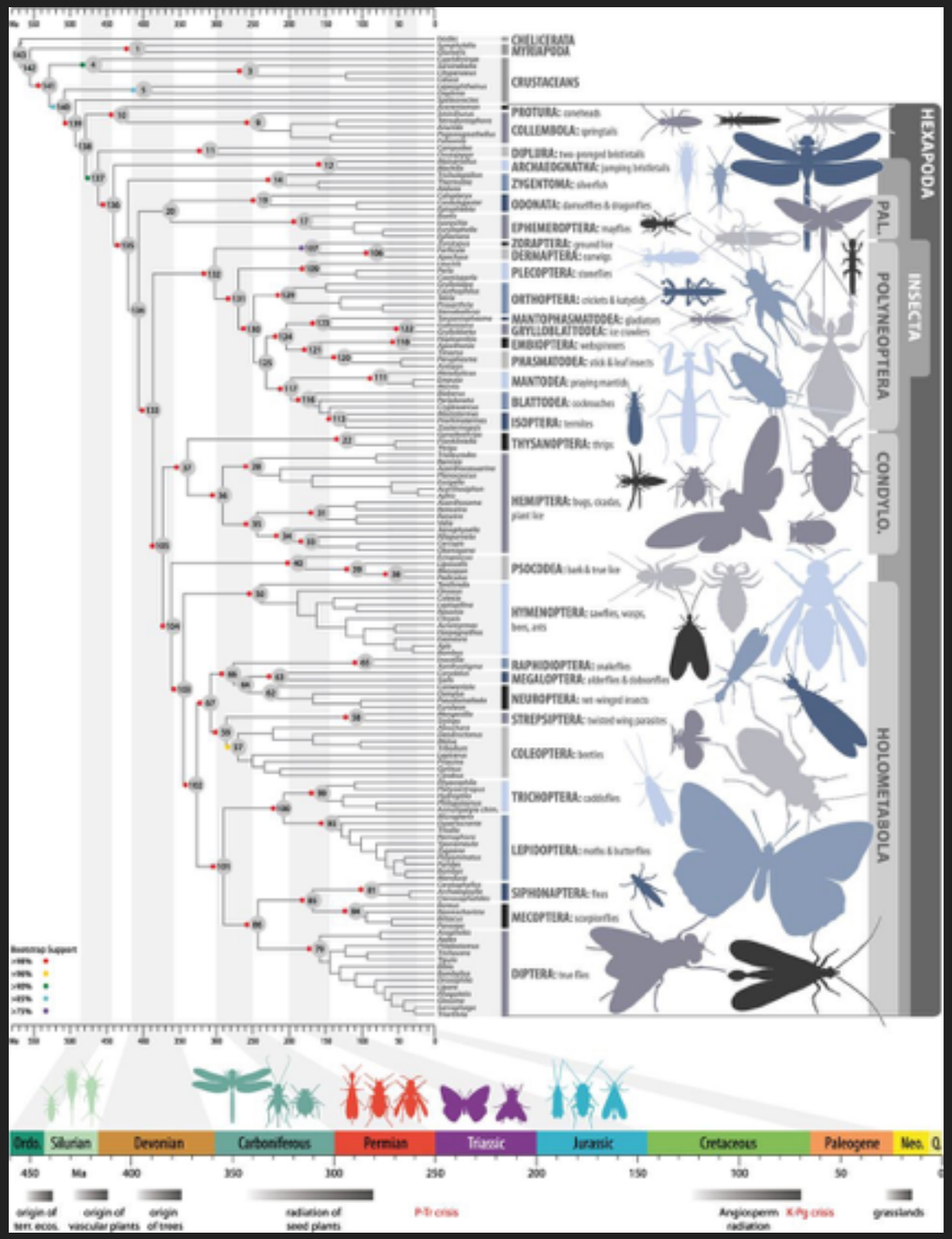


From: http://www2.fml.tuebingen.mpg.de/raetsch/members/research/transcriptomics.html

Slide from: https://github.com/trinityrnaseq/RNASeq_Trinity_Tuxedo_Workshop/blob/master/docs/rnaseq_workshop_slides.pdf

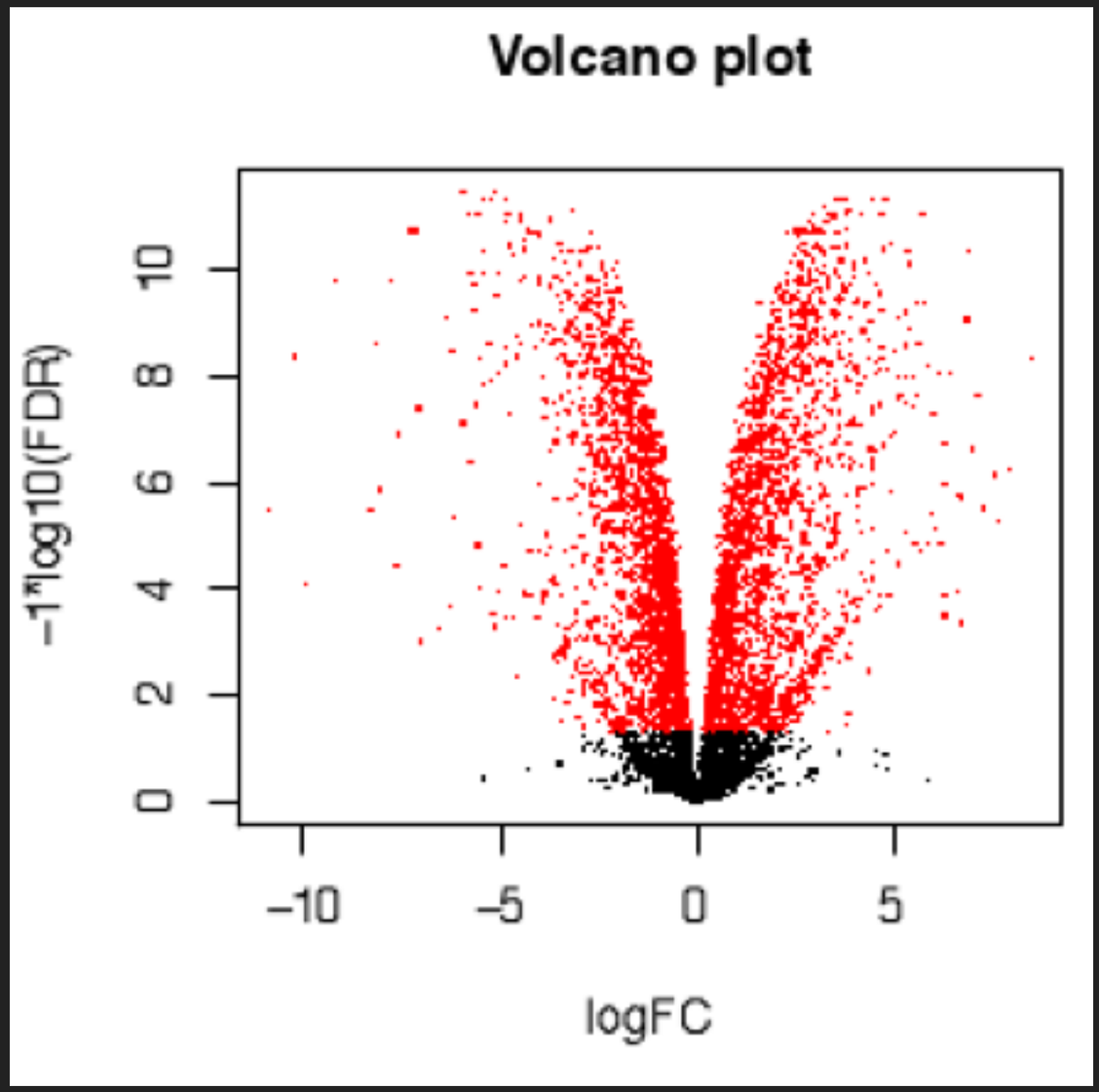
COMMON RNASEQ APPLICATIONS

- Raw genetic data
  - Population genomics
  - Phylogenomics
- Experimental biology
  - Differential expression analysis



COMMON RNASEQ APPLICATIONS

- Raw genetic data
  - Population genomics
  - Phylogenomics
- Experimental biology
  - Differential expression analysis



QUALITY CHECK SHORT READ SEQUENCE

- FASTQC "A quality control tool for high throughput sequence data"
  - http://www.bioinformatics.babraham.ac.uk/projects/fastqc/

TRIMMING

- Many different trimmers available
- Trim adapters/low quality reads
- Trim Galore! is one that works well on Hydra

HANDS-ON

- Go to https://github.com/SmithsonianWorkshops/RNA-seq
- Click on “2a_Raw Read QA-QC.md”