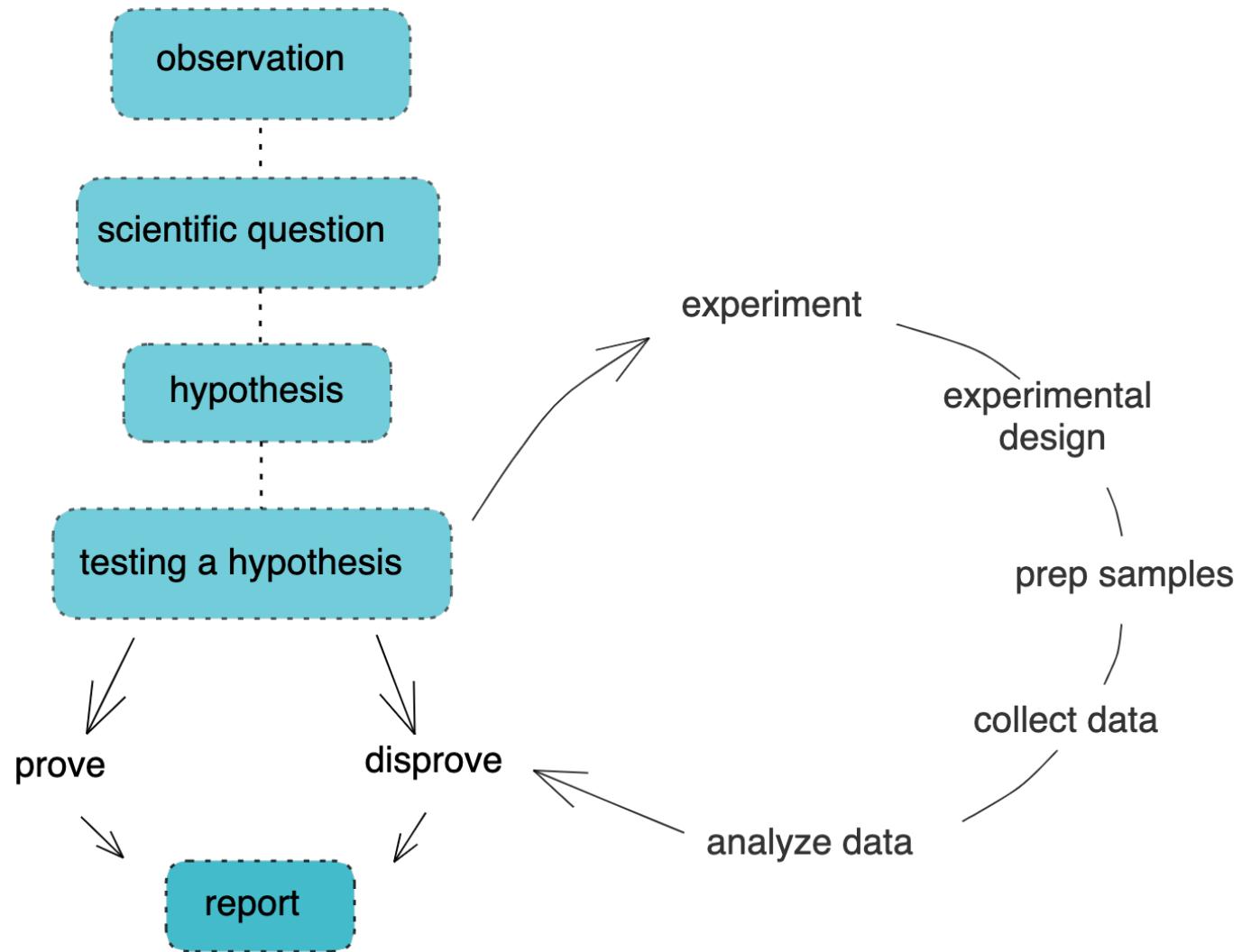


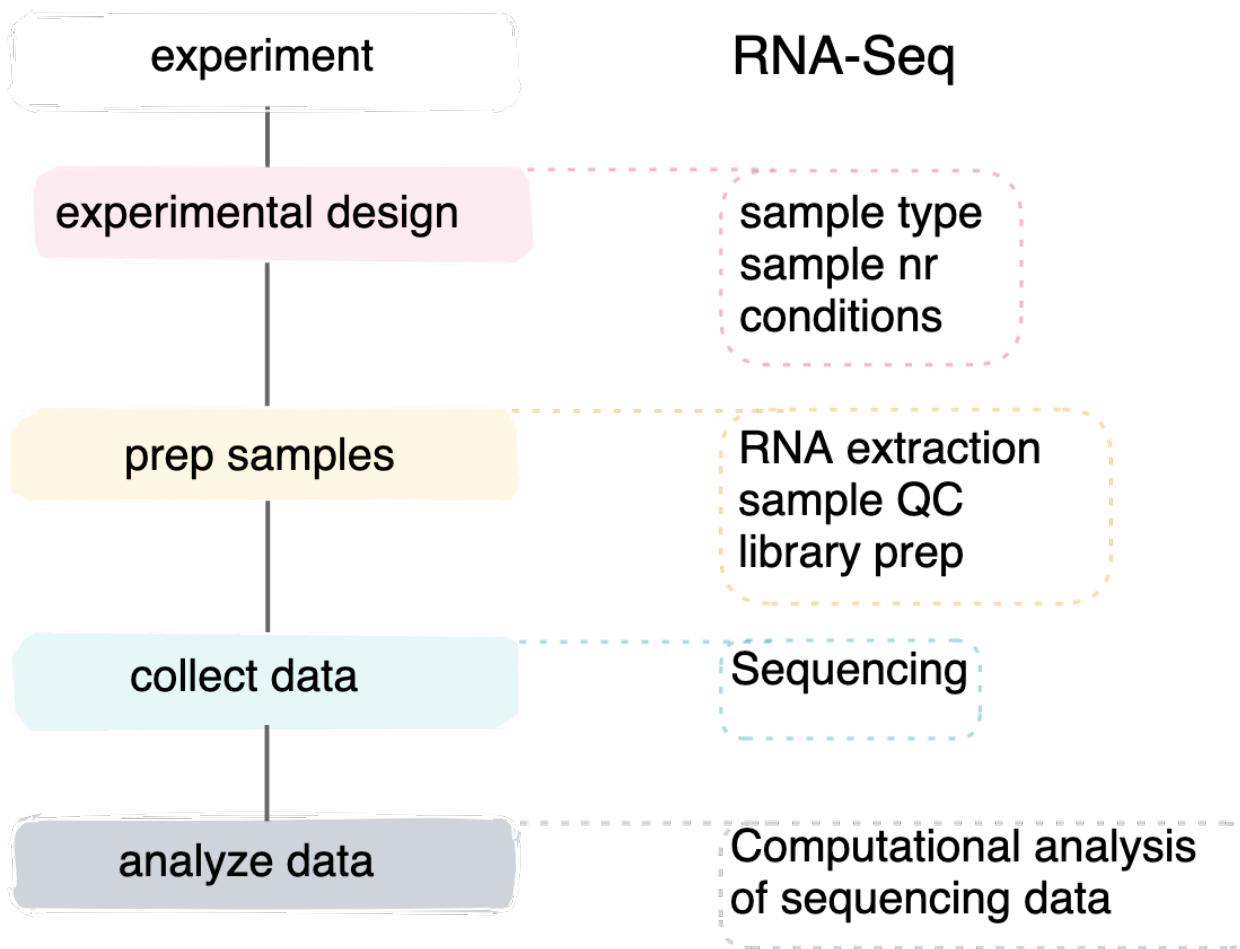
From Sample Preparation to Sequencing:

Experimental Design, Library Preparation and cDNA Sequencing

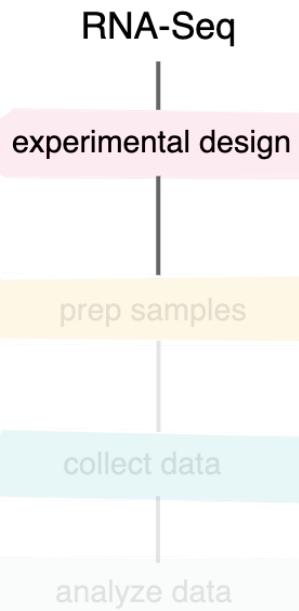
Annelien Verfaillie

29th sept 2020



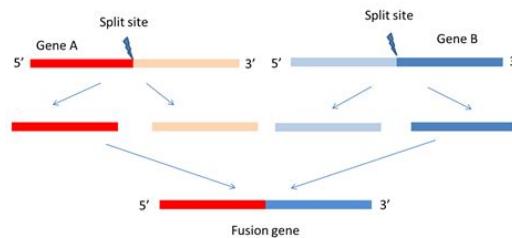


What is my scientific question?



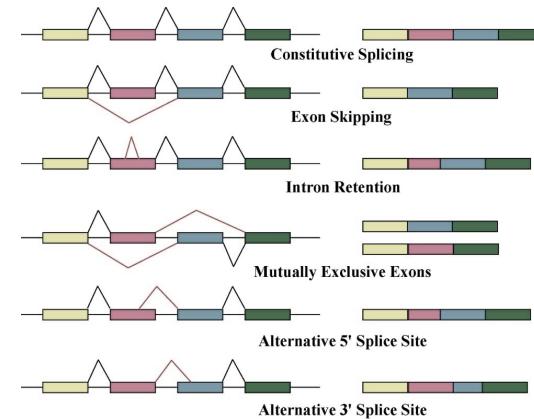
Quantify

- ✓ Difference between conditions
- ✓ Change over time

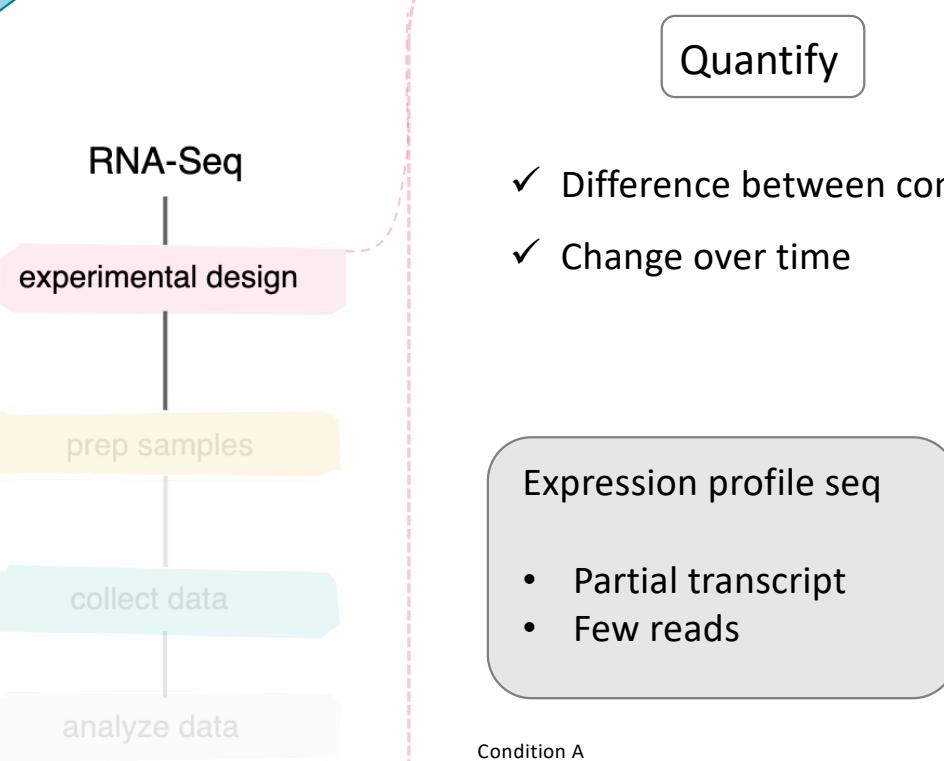


Identify

- ✓ De novo transcript assembly
- ✓ Fusion genes
- ✓ Isoform



What is my scientific question?

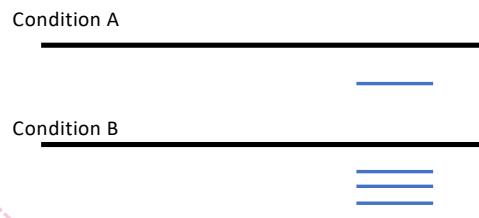


Quantify

- ✓ Difference between conditions
- ✓ Change over time

Expression profile seq

- Partial transcript
- Few reads



Identify

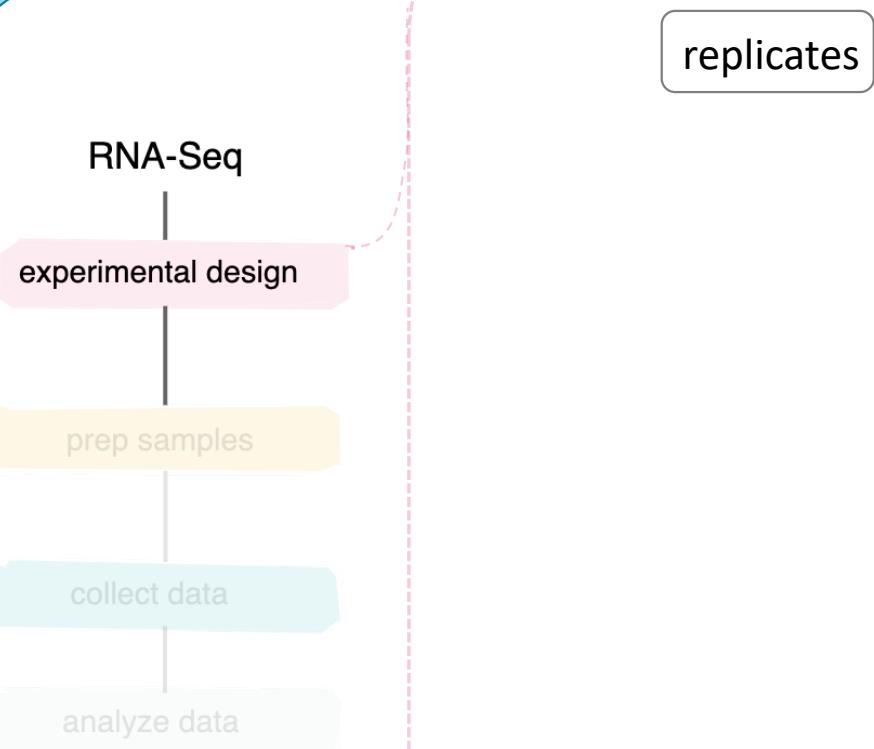
- ✓ De novo transcript assembly
- ✓ Fusion genes
- ✓ Isoform

Whole transcriptome seq

- Full transcript
- Many reads

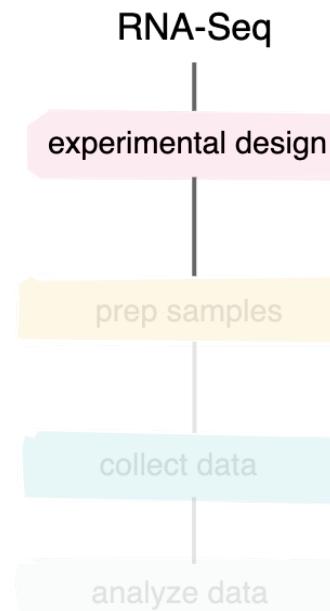


What do I need?



replicates

control



What do I need?

replicates

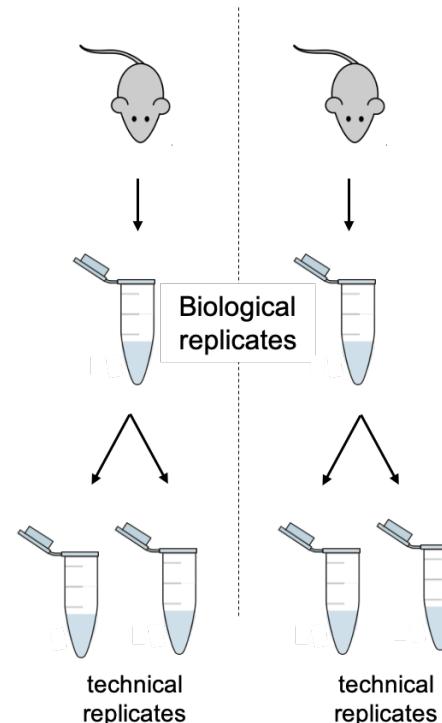
Technical

- Correct for variation due to sample handling
 - ✓ doing sample prep twice on same RNA
 - ✓ sequencing the same library twice

Biological

- Correct for variation in the population
 - ✓ Depends on sample origin

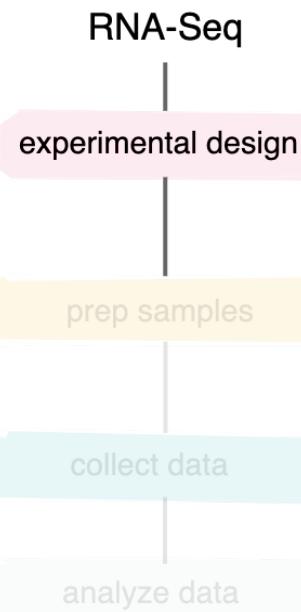
Cell lines:	3
Inbred mice:	4
Human:	6
Other species:	4-5



What do I need?

replicates

control



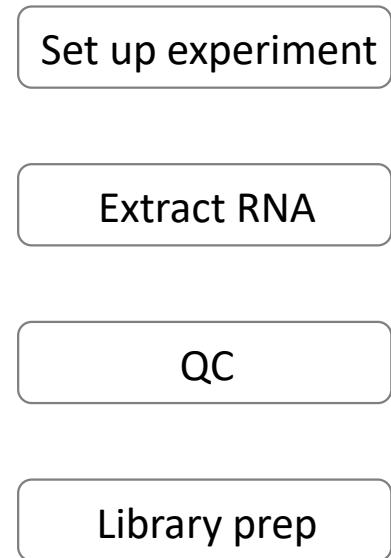
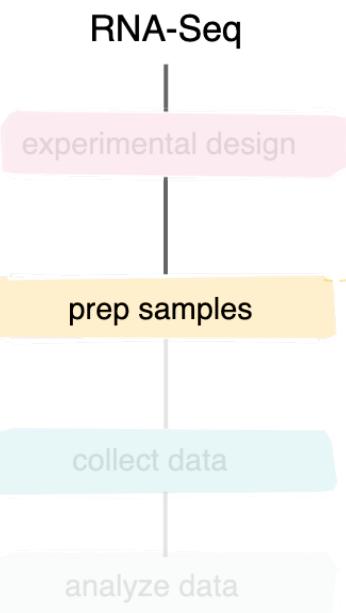
Technical

- Correct for variation due to sample handling
 - ✓ doing sample prep twice on same RNA
 - ✓ sequencing the same library twice

Biological

- Correct for variation in the population
 - ✓ Depends on sample origin

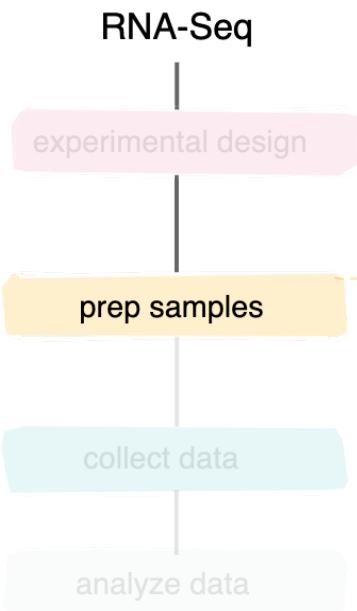
Cell lines:	3
Inbred mice:	4
Human:	6
Other species:	4-5



} You

You

You or the genomics core



RNA extraction

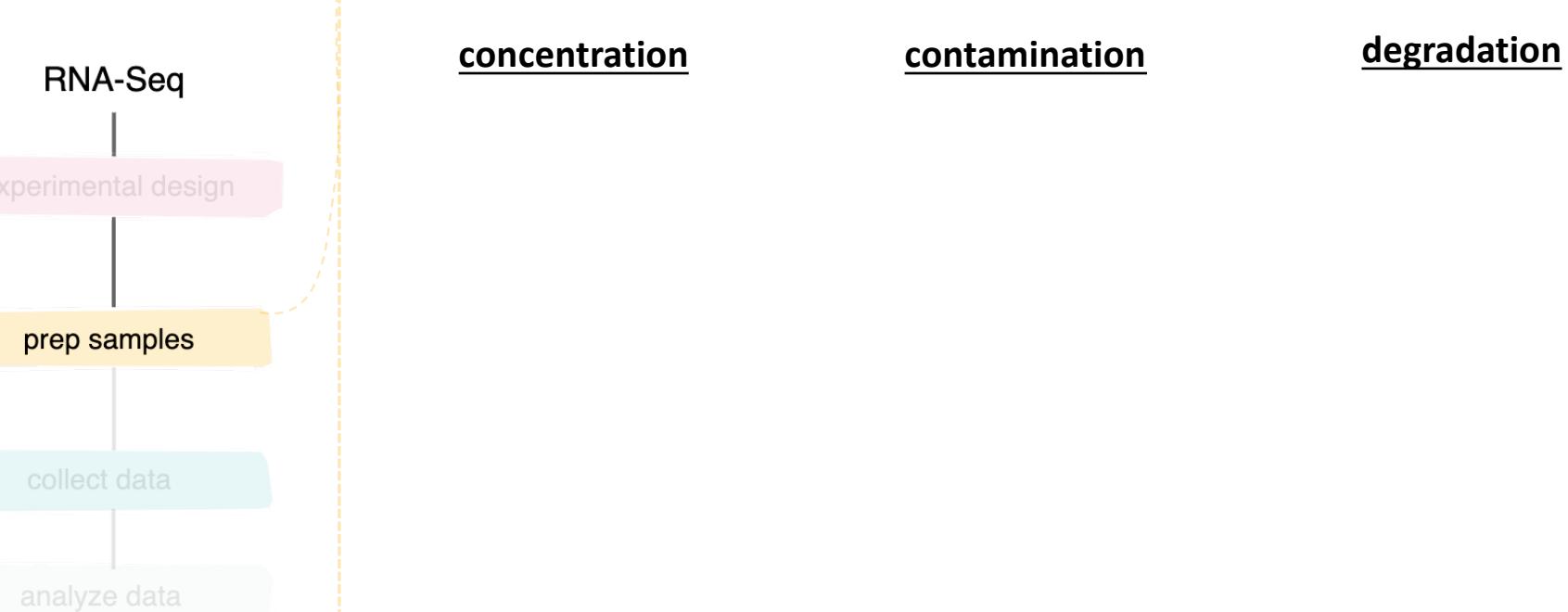
- **Total RNA:**

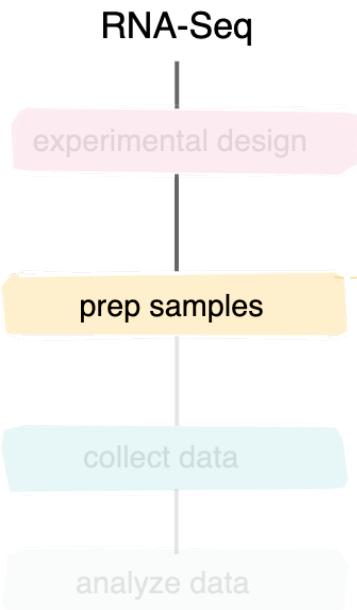
- ✓ mRNA
- ✓ noncoding RNA:
tRNA,
rRNA
small RNAs: microRNA, lincRNA, snoRNA, piRNA,....

- **mRNA:**

- ✓ precursor to a protein
- ✓ polyA tail
- ✓ represents the classical transcribed genes

→ Choice depends on experimental question and downstream library prep protocol





- Measure on a nanodrop
 - ✓ 260/280
 - ✓ 260/230
- Input required depends on downstream protocol
 - ✓ More is better
- Avoid gDNA contamination
- specific contaminations
 - ✓ globin mRNA from blood samples
 - ✓ rRNA for small RNA detection

RNA-Seq

experimental design

prep samples

collect data

analyze data

QC

concentration

contamination

degradation

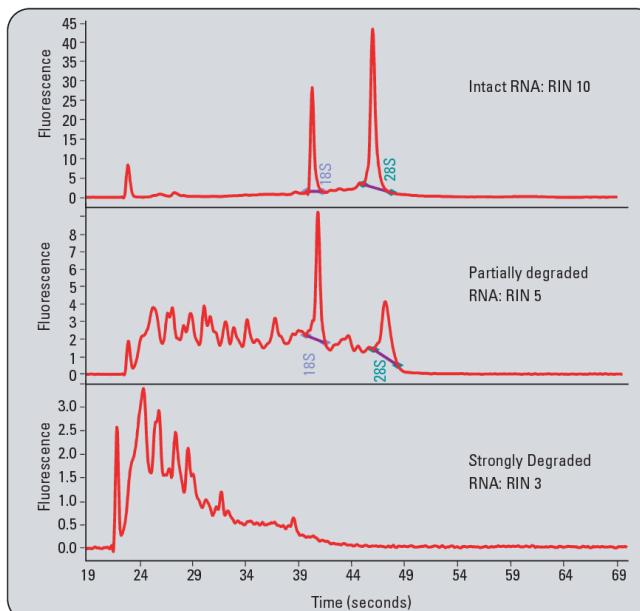
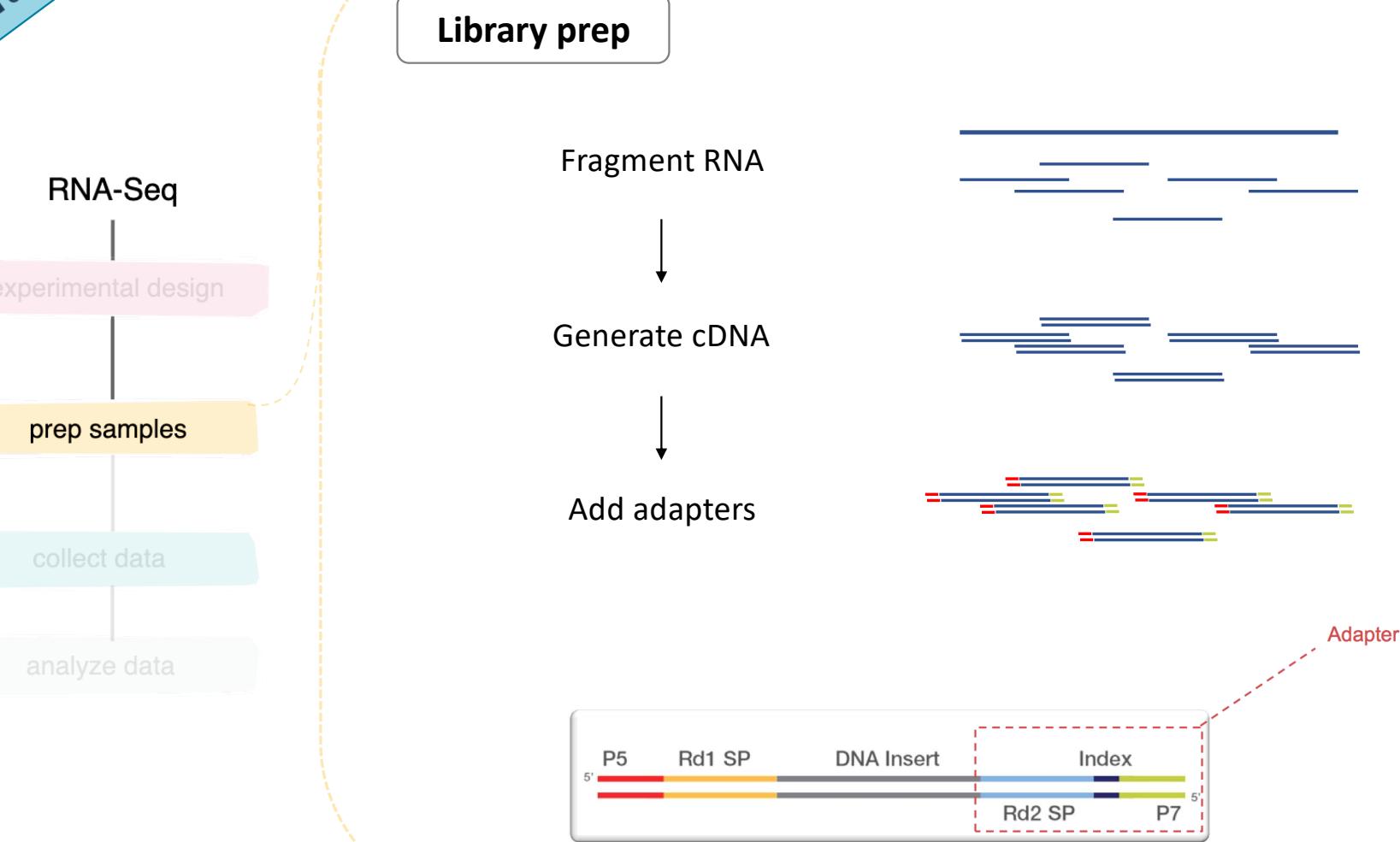


Figure 6

- RIN value
(RNA integrity number)
- Tolerance for low quality is protocol dependent.
- Full transcript methods require high quality RNA



RNA-Seq

experimental design

prep samples

collect data

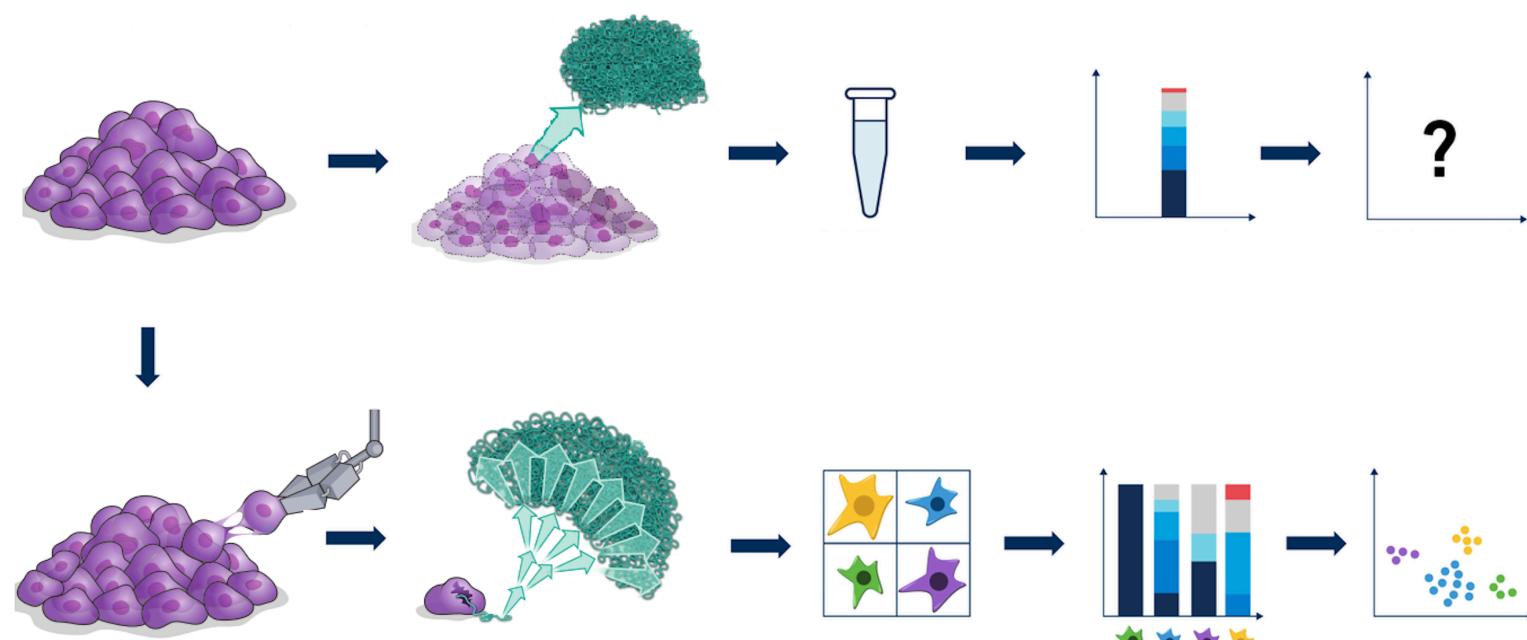
analyze data

Library prep

	Differential expression	Whole transcript, fusion, isoforms	Small RNA	Illumina compatible
Lexogen QuantSeq 3' mRNA	✓			✓
Lexogen Small RNA seq	✓		✓	✓
Illumina TruSeq stranded mRNA	✓	✓		✓
Illumina TruSeq stranded total RNA	✓	✓	✓	✓
IsoSeq	(✓)	✓		

Single cell RNA-seq as a complementary technique to bulk RNA-seq

Side note



Single cell RNA-seq as a complementary technique to bulk RNA-seq

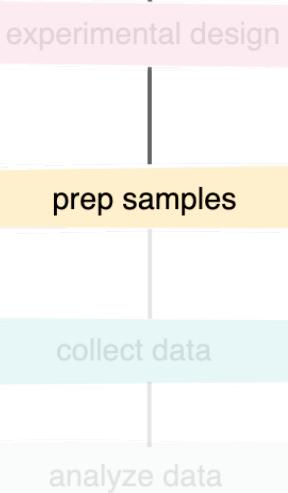
Side note

	RNA type	Transcript targeted	sensitivity	throughput	sequencing	Specific property
SMART SEQ2	mRNA	full transcript	sensitive	low	deep	FACS sorting specific populations
10x genomics 3' RNA seq	mRNA	3' end	medium	High (> 10000)	shallow	Can be combined with surface markers

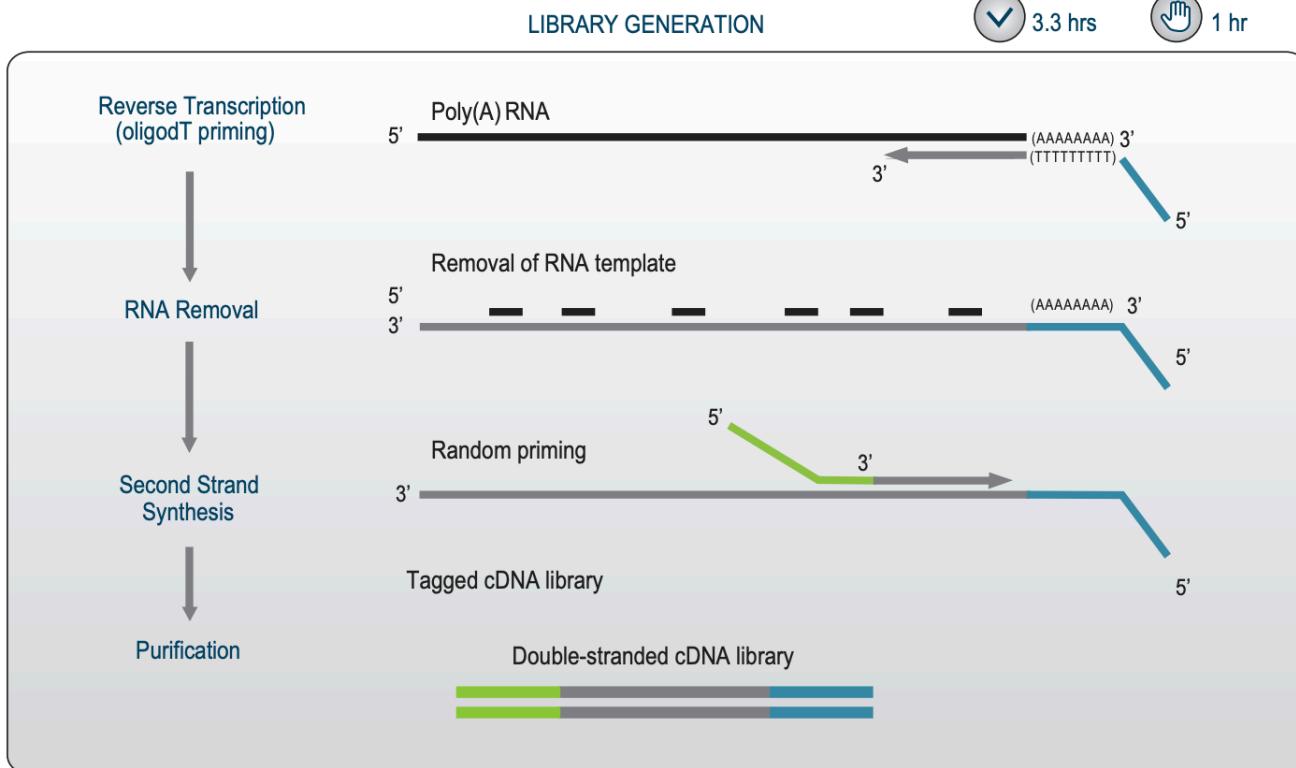
10x variations

- ATAC + RNA - seq
- Spatial transcriptomics

RNA-Seq



Lexogen QuantSeq 3' mRNA



RNA-Seq

experimental design

prep samples

collect data

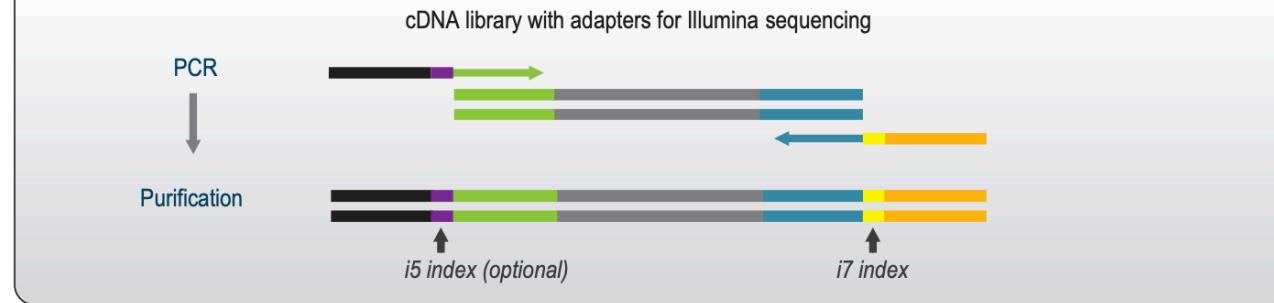
analyze data

Lexogen QuantSeq 3' mRNA

LIBRARY AMPLIFICATION

✓ 1.2 hrs

🕒 45 min



SEQUENCING - Read orientation for QuantSeq FWD



Lexogen QuantSeq 3' mRNA

RNA-Seq

experimental design

prep samples

collect data

analyze data

requirements

		Input	Sequencing	conditions	depth
	quantity			0.5 ng* – 1 µg	
	purity			Total RNA, (rRNA depletion not needed)	
	integrity			RIN: no minimum (tolerant for degradation (FFPE)) $260/230 > 0.8$ (indication for extra purification)	
				Single end, 50 bp	
				1-2 million per sample	

* Depending on contaminants

RNA-Seq

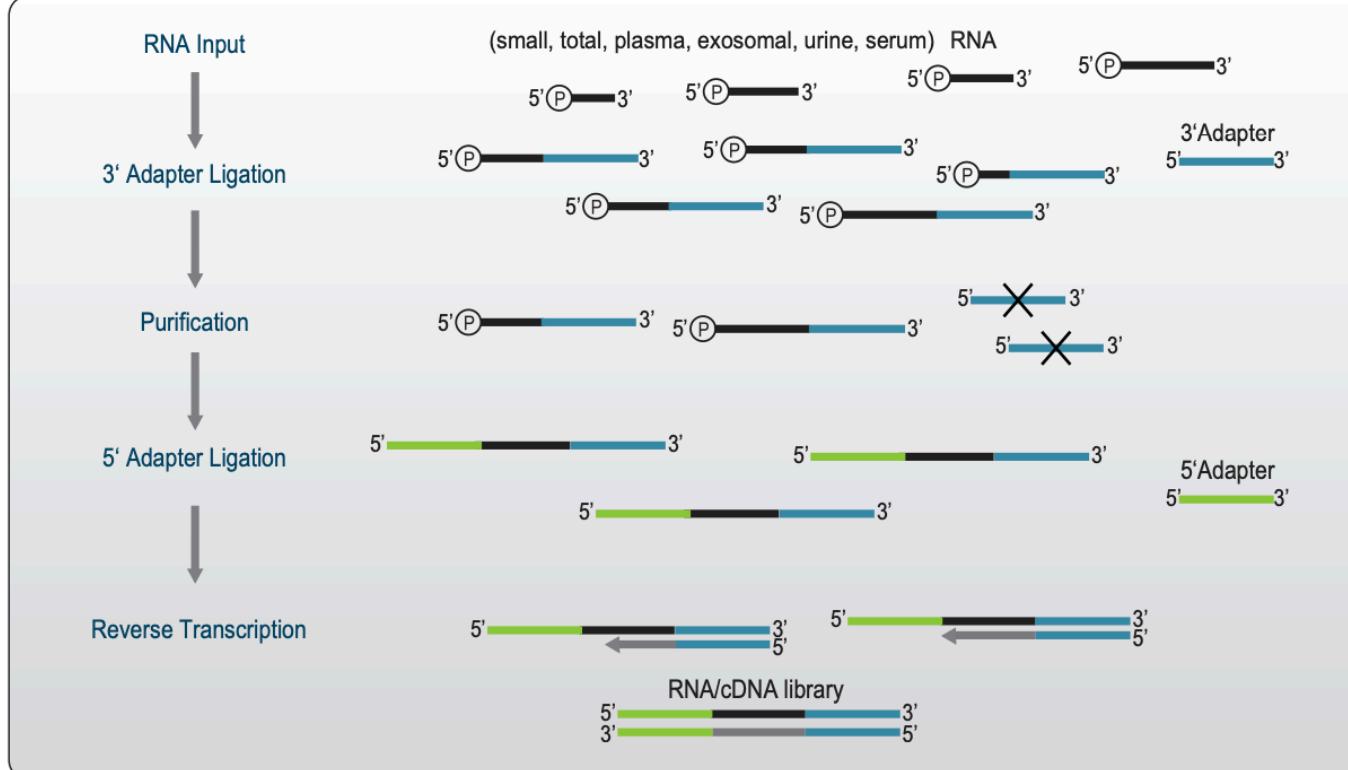
- experimental design
- prep samples
- collect data
- analyze data

Lexogen Small RNA seq

LIBRARY GENERATION

3.8 hrs

50 min



RNA-Seq

experimental design

prep samples

collect data

analyze data

Lexogen Small RNA seq

LIBRARY AMPLIFICATION

1 hr

25 min

cDNA Library with Adapters for Illumina Sequencing

PCR

Purification

i7 Index

SEQUENCING - Read orientation

Read 1

Sequencing

Primer

Read 1

Index 1 (*i7*)

Sequencing

Primer

i7

Read 2

Sequencing Primer

Lexogen Small RNA seq

RNA-Seq

experimental design

prep samples

collect data

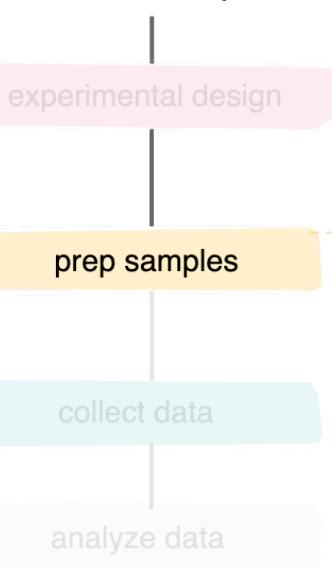
analyze data

		requirements
Input	quantity	0.1 – 1µg total 50pg- 1µg enriched small RNA
	purity	Total RNA* (rRNA depletion not needed)
	integrity	RIN: no minimum (tolerant for degradation (FFPE)) $260/230 > 0.8$ (indication for extra purification)
Sequencing	conditions	Single end, 50 bp
	depth	1-2 million per sample**

* extraction kits for total RNA should be able to retain small RNAs!

** additional size selection of library might be needed if focus on smallest RNAs

RNA-Seq

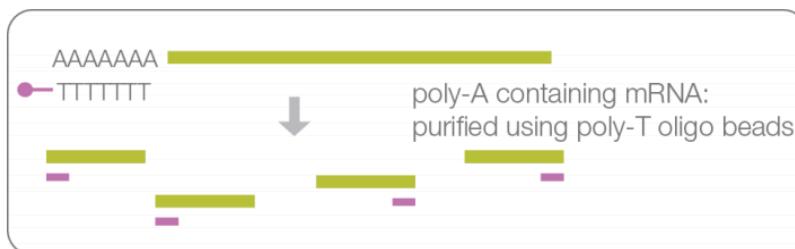


Illumina TruSeq stranded RNA (total or mRNA)

Figure 1 Ribo-Zero Depleting and Fragmenting RNA



Figure 1 Purifying and Fragmenting mRNA



RNA-Seq

experimental design

prep samples

collect data

analyze data

Illumina TruSeq stranded RNA (total or mRNA)



First strand cDNA

Random primers are used during first strand synthesis

Figure 3 Synthesizing Second Strand cDNA



to keep stranded info, dUTPs are incorporated during second strand synthesis

RNA-Seq

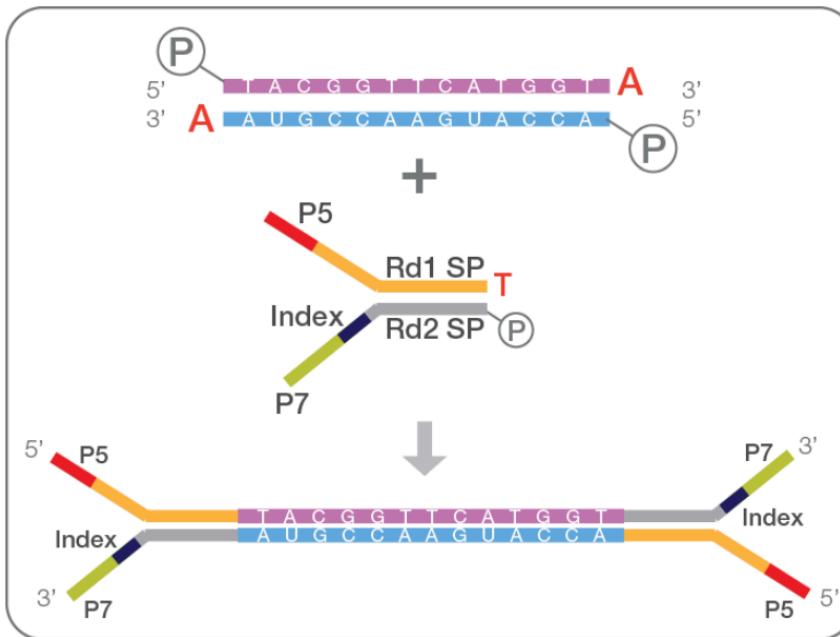
experimental design

prep samples

collect data

analyze data

Illumina TruSeq stranded RNA (total or mRNA)



Illumina works with Y adaptors to increase fraction of fragments with P5+P7

RNA-Seq

experimental design

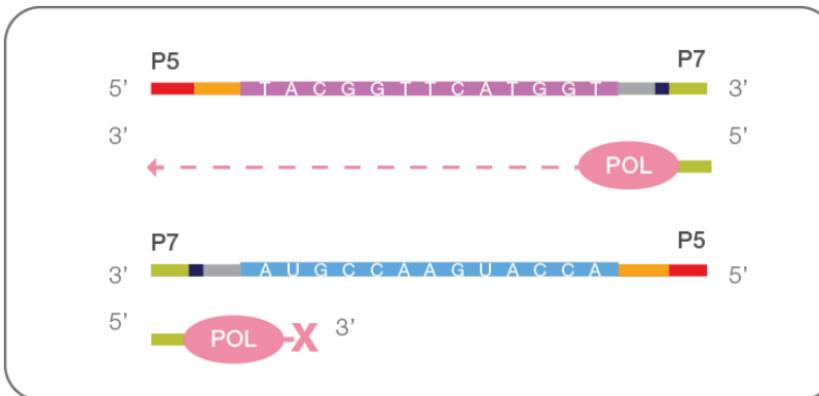
prep samples

collect data

analyze data

Illumina TruSeq stranded RNA (total or mRNA)

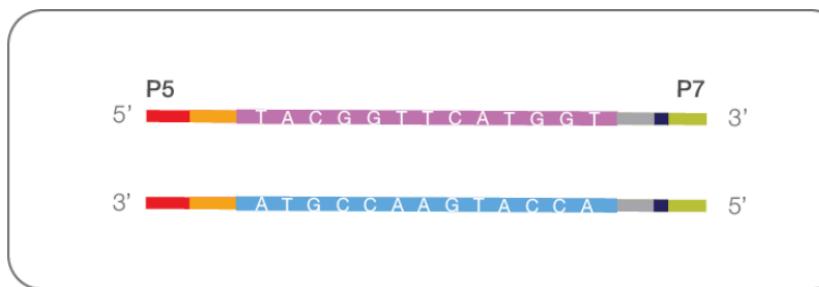
Figure 6 Enriching DNA Fragments



Polymerase cannot pass dUTPs

No amplification of the second strand

Figure 7 LS Final Library



In your final library the direction of the adapters ensures correct stranded information

RNA-Seq

experimental design

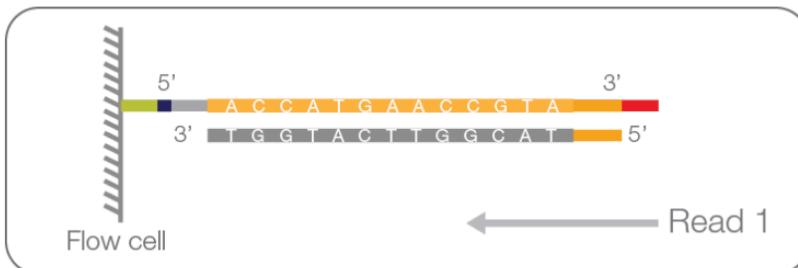
prep samples

collect data

analyze data

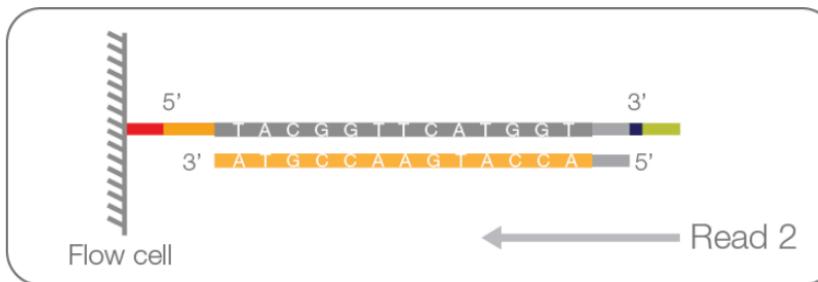
Illumina TruSeq stranded RNA (total or mRNA)

Figure 8 Cluster Generation and Read 1 Sequencing



Read 1 will map to the antisense strand

Figure 9 Paired-end Turnaround and Read 2 Sequencing



Read 2 will map to the sense strand

RNA-Seq

experimental design

prep samples

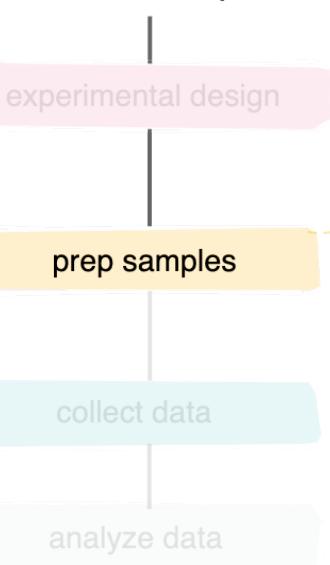
collect data

analyze data

Illumina TruSeq stranded RNA (total or mRNA)

requirements		
Input	quantity	0.2 – 2 µg total, 0.02-1 µg mRNA
	purity	Total RNA, (rRNA depletion optional)
	integrity	RIN > 8 (High integrity is crucial) 260/230 > 1.5 (indication for purification)
Sequencing	conditions	Single end, 50 bp (diff. exp) Paired end, 150 bp (other)
	depth	10-20 million (diff exp) 40 million (fusion etc...)

RNA-Seq

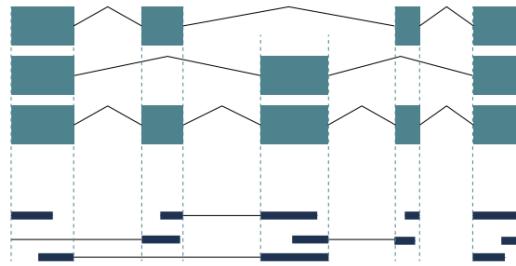


IsoSeq (PacBio)

- Long read sequencing
- Transcript is kept intact
- Easier to obtain fusion and splicing information

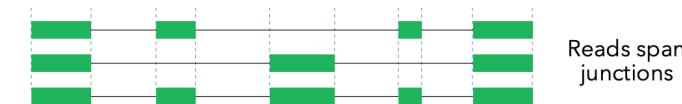
Gene with Several Transcript Isoforms

Few reads spanning junctions



RNA-seq (Short reads)

Additional inference required to parse isoforms



Full-length transcripts capture all splice junctions with no assembly required



RNA-Seq

experimental design

prep samples

collect data

analyze data

IsoSeq (PacBio)

Reverse Transcription



Template Switching

Template Switch
Oligo (TSO)



PCR Amplification

Amplified DNA

RNA-Seq

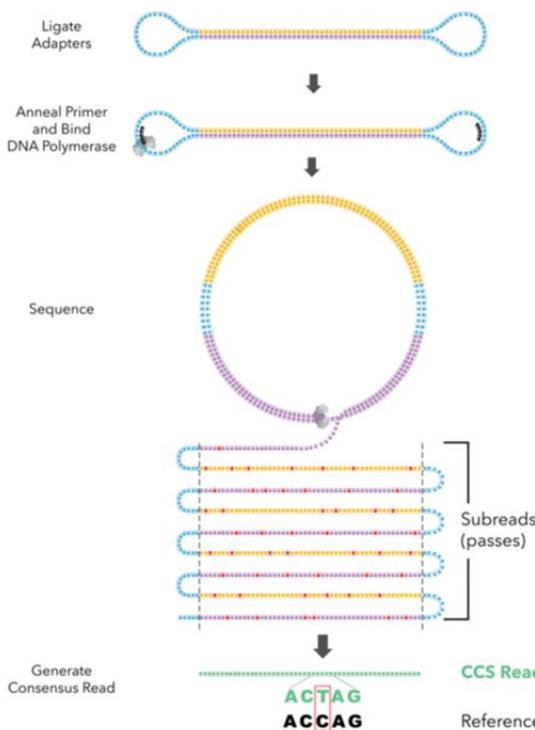
experimental design

prep samples

collect data

analyze data

IsoSeq (PacBio)



Final library is a circular molecule containing long stretches of cDNA

Repeated sequencing of each molecule allows for reliable analysis of intact RNA fragments

RNA-Seq

experimental design

prep samples

collect data

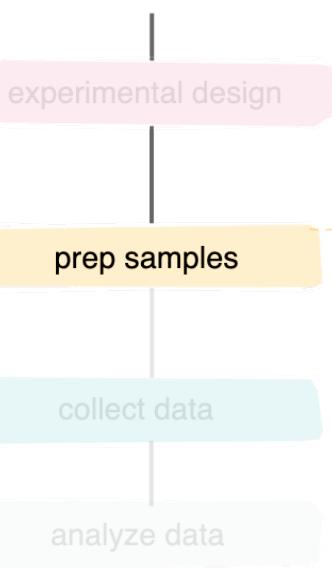
analyze data

IsoSeq (PacBio)

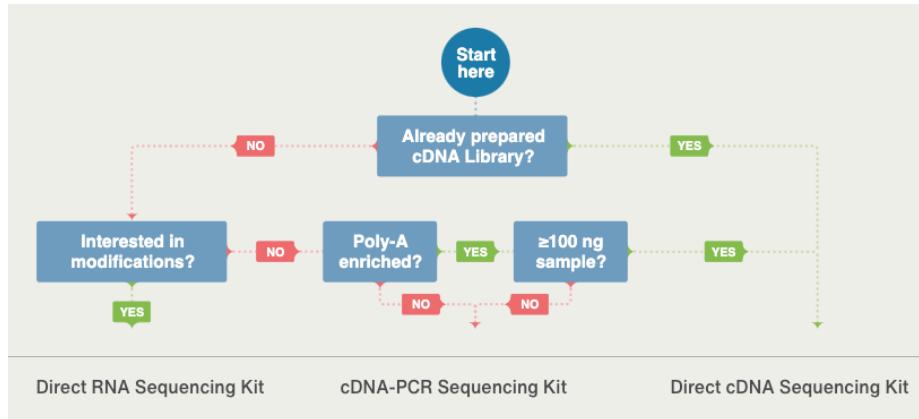
requirements

		Input	Sequencing	requirements
		quantity	depth	> 0.6 µg total
		purity		Total RNA
		integrity		RIN > 8 (High integrity is crucial) $260/230 > 1.8$
				1 human transcriptome = > 1 million full length non-concatemer reads

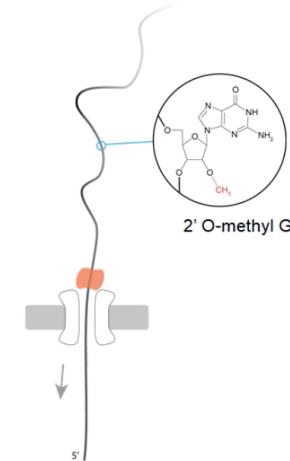
RNA-Seq



Oxford nanopore as an alternative long read RNA seq method



- Long read information
- detect modification by sequencing RNA directly!



Experimental design

- Quantify or identify:
differential expression vs full transcript
- Nr of samples:
replicates and controls

Library prep

- | | |
|---------|-----------------------------|
| Lexogen | ✓ diff exp |
| | ✓ mRNA or small RNA |
| Truseq: | ✓ diff exp or transcriptome |
| | ✓ full transcript |
| | ✓ mRNA or Total RNA |
| Isoseq: | ✓ long read information |

RNA extraction and QC

- Quantity: low or high yield?
- Quality: low RIN, contaminants?
- Purity: rRNA depletion, globin
mRNA, enrichment for small RNA?

Sequencing

- Lexogen: shallow and short reads
- Truseq: deep and medium length
- IsoSeq: deep and long reads

You are all set to start your first RNA-seq experiment!

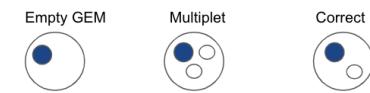
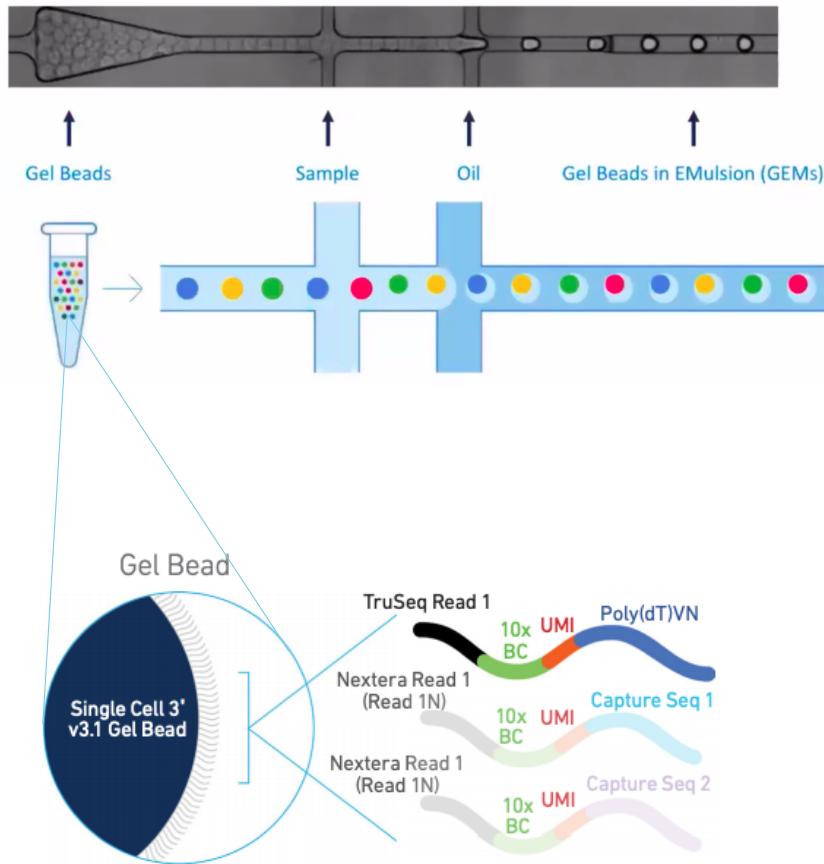
Still not sure on how to proceed?



Contact the Genomics Core for more info

info@genomicscore.be

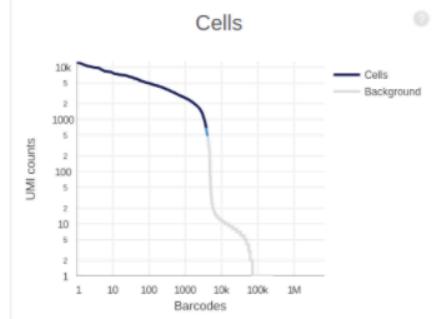
10x RNA-Seq



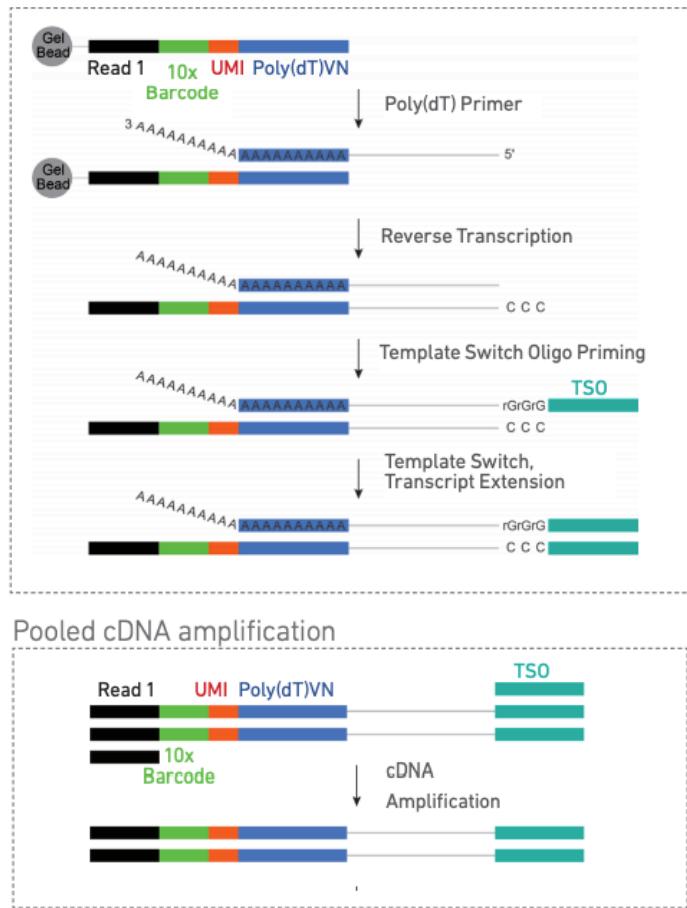
Estimated Number of Cells
3,868

Mean Reads per Cell Median Genes per Cell
4,939 **722**

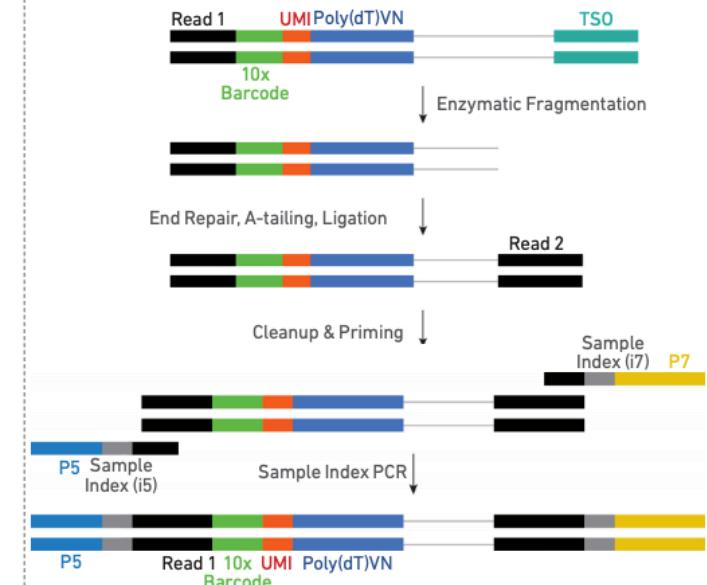
Sequencing
 Number of Reads: 19,107,635
 Valid Barcodes: 97.7%
 Sequencing Saturation: 4.7%



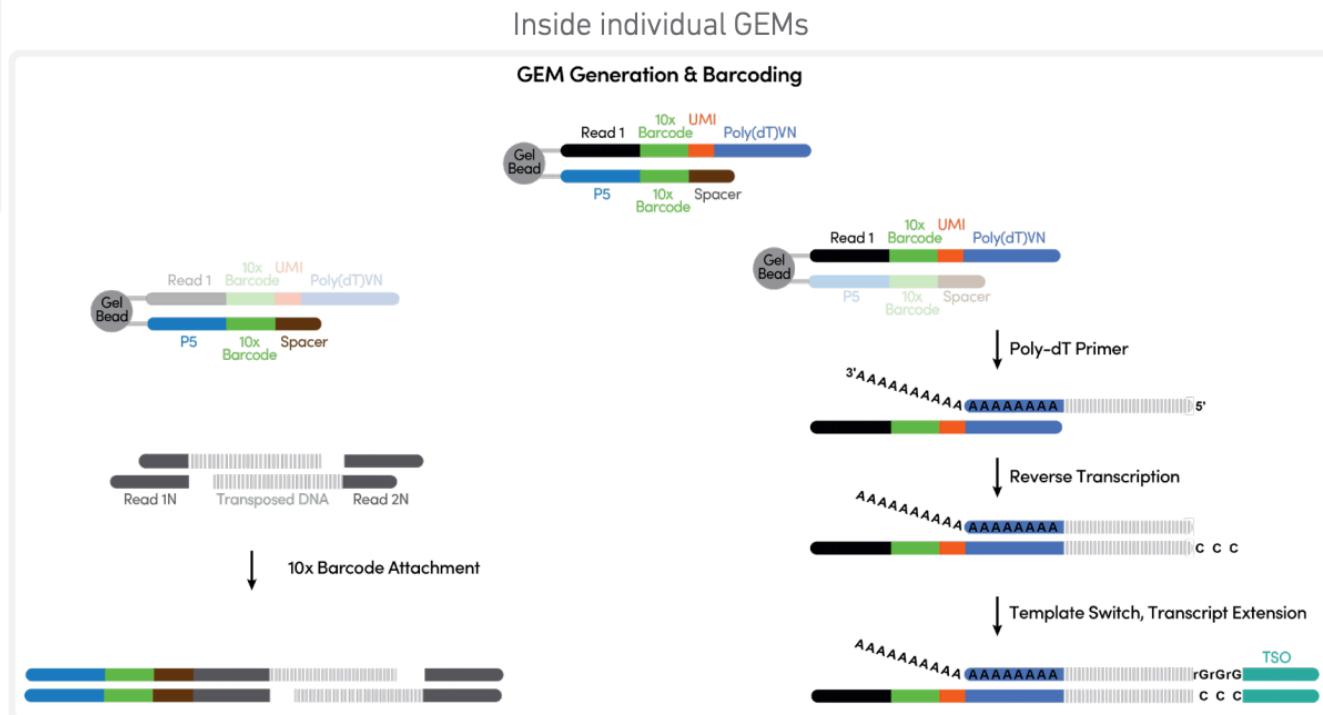
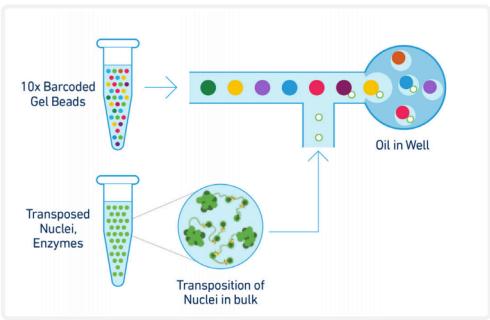
10x RNA-Seq



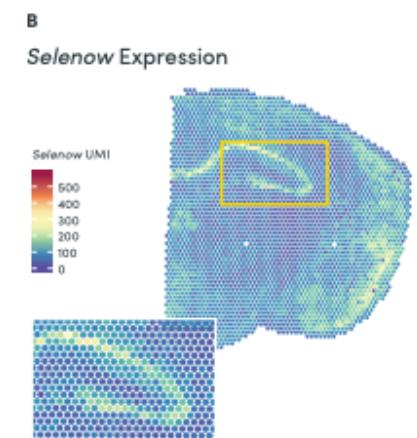
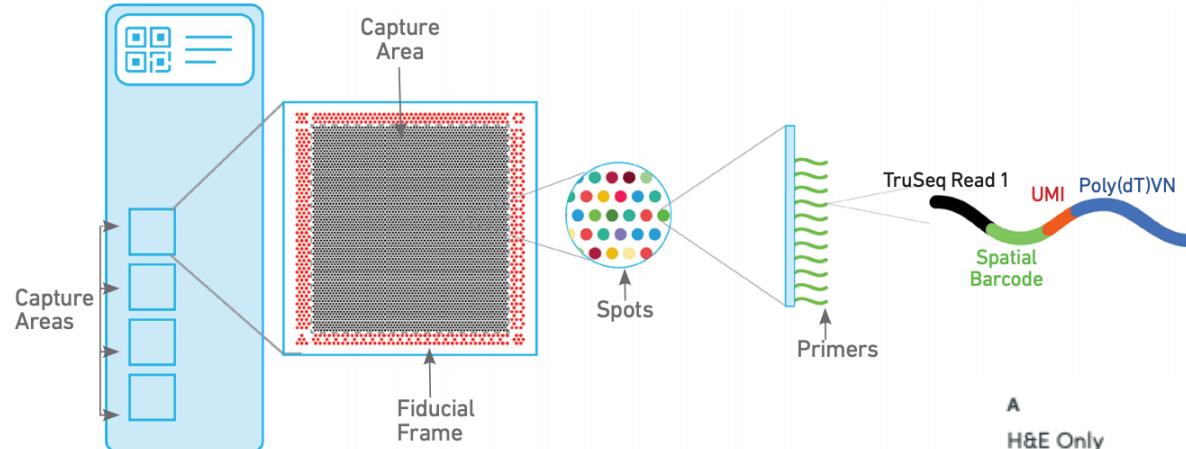
Amplified cDNA processing (dual index)



10x RNA-Seq variation: ATAC + RNA seq



10x RNA-Seq variation: Spatial transcriptomics



SMART-Seq2

