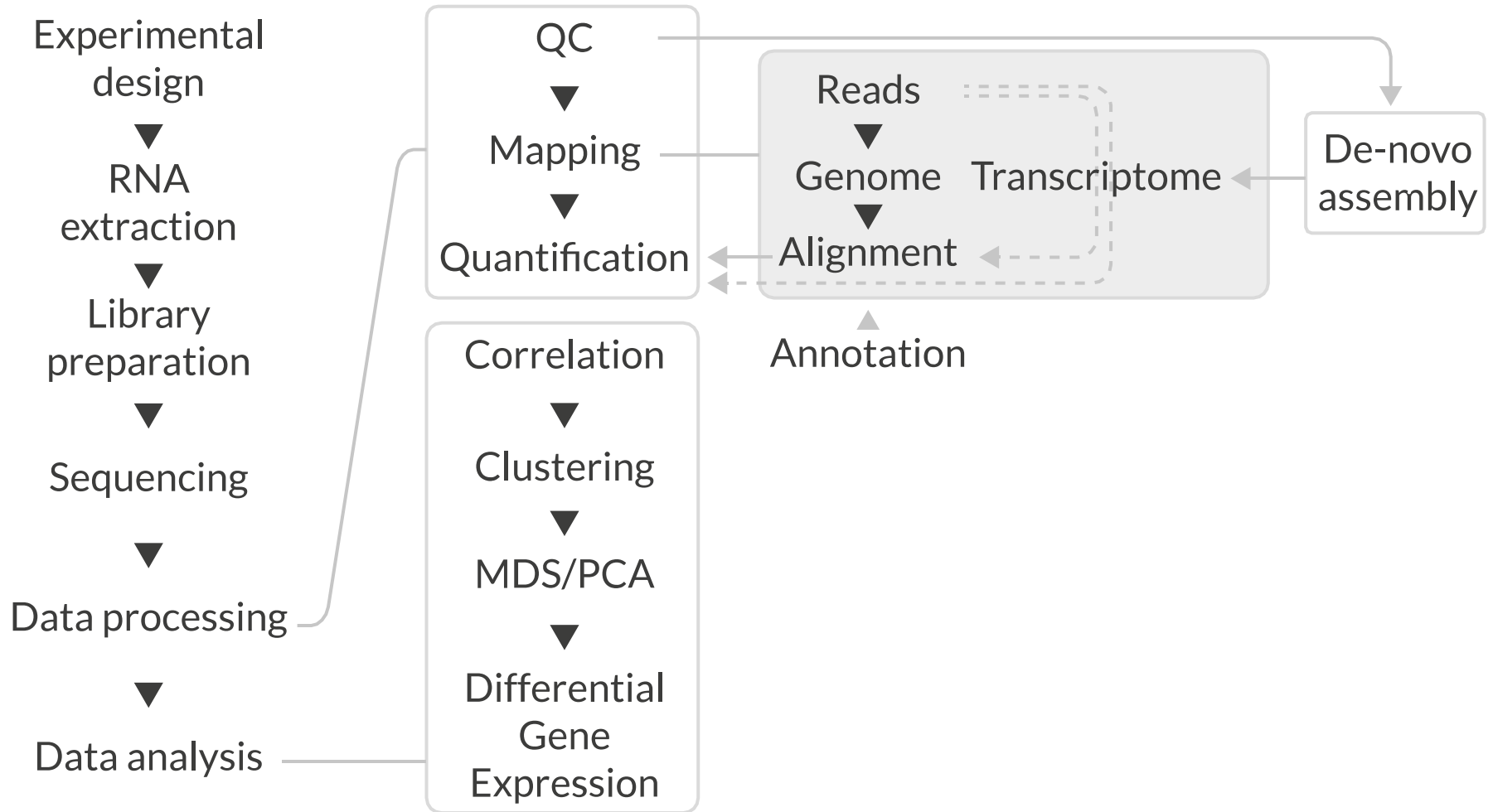# Summary

Workflow

Some considerations

- Sound experimental design to avoid confounding
- Plan carefully about lib prep, sequencing, etc based on experimental objective
- Biological replicates may be more important than paired-end reads or long reads
- Discard low quality bases, reads, genes, and samples
- Verify that tools and methods align with data assumptions
- Experiment with multiple pipelines and tools

Depending on the different steps you will get different results
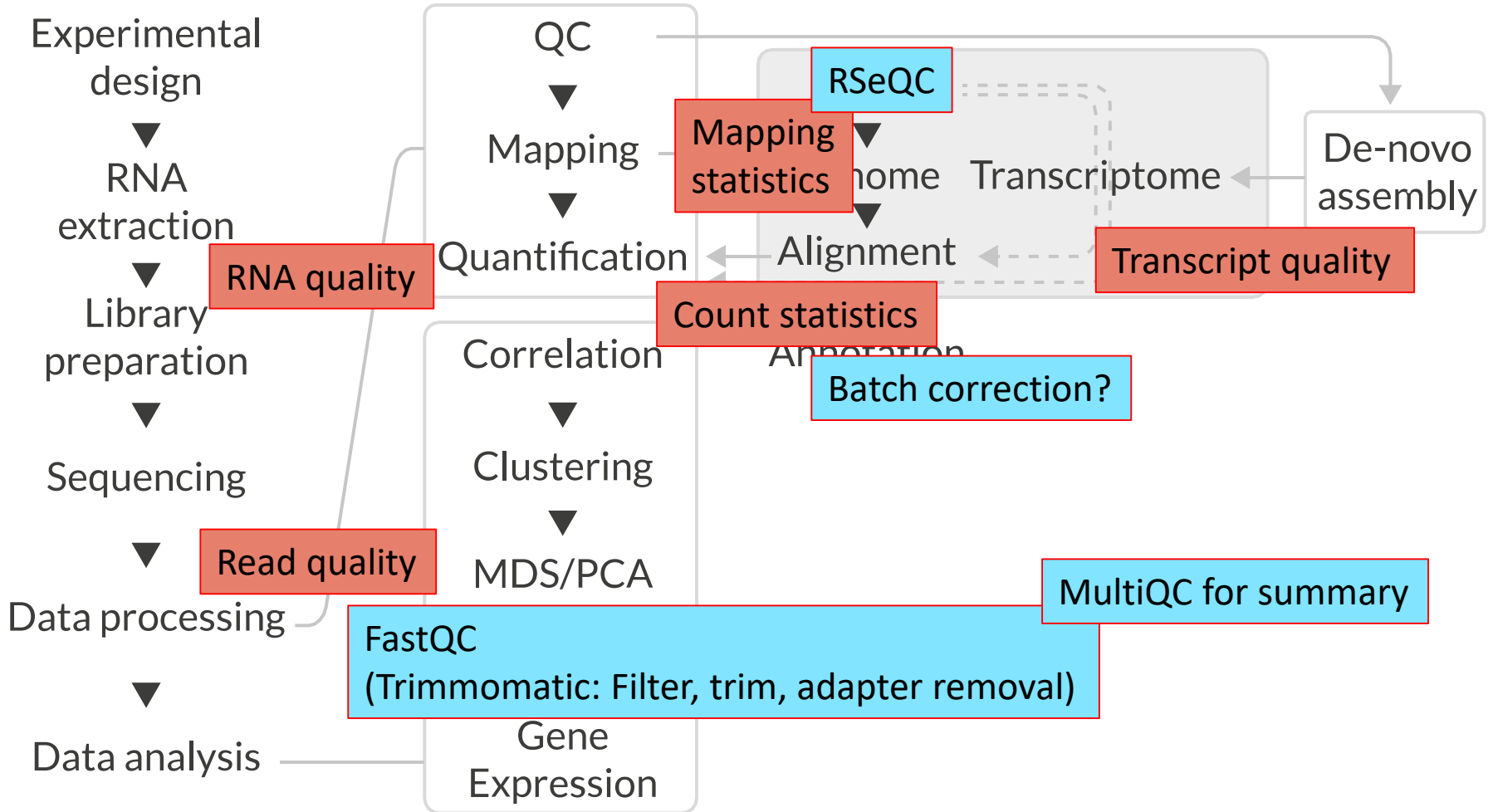
RNA->

enrichments ->

extraction of poly-A RNAs
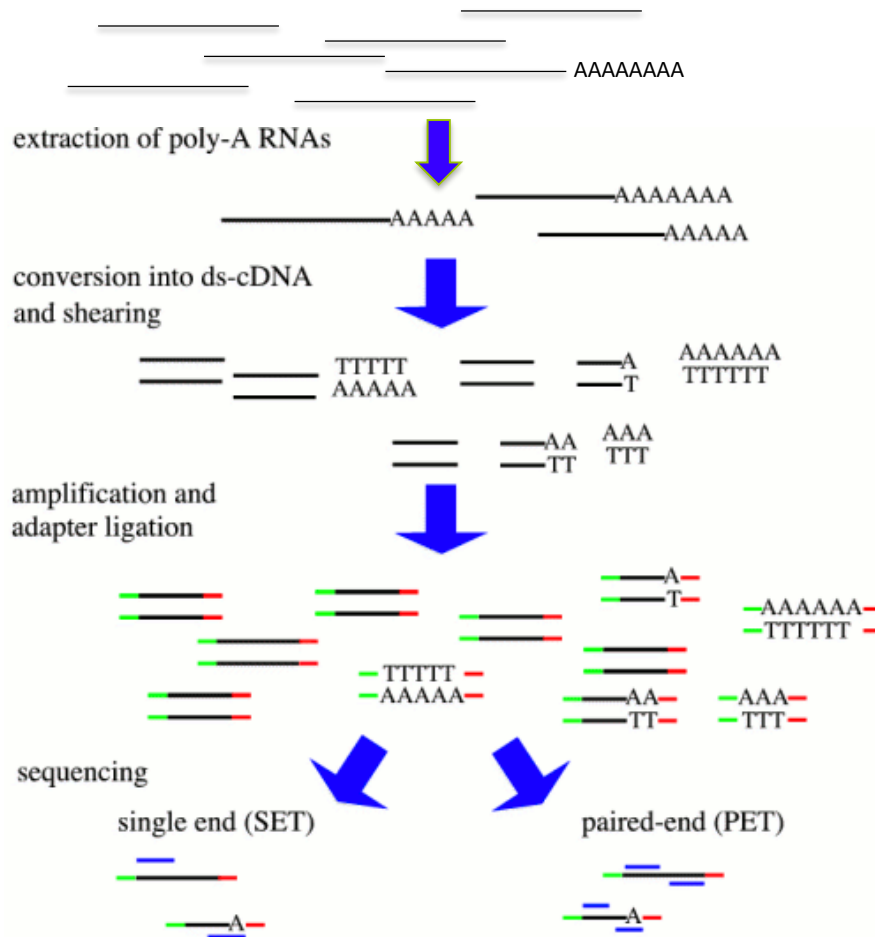
PolyA	(mRNA)
RiboMinus	(- rRNA)
Size <50 nt	(miRNA)
…..	

conversion into ds-cDNA and shearing

Size of fragment
Strand specific
5' end specific
3' end specific
…..

amplification and adapter ligation

library ->

sequencing

single end (SET)   paired-end (PET)

reads ->

Single end (1 read per fragment)
Paired end (2 reads per fragment)

Workflow

Experimental design
▼
RNA extraction
▼
Library preparation
▼
Sequencing
▼
Data processing
▼
Data analysis

QC
▼
Mapping
▼
Quantification

Correlation
▼
Clustering
▼
MDS/PCA

Gene Expression

Genome   Transcriptome

Alignment

Annotation

De-novo assembly

RSeQC

Mapping statistics

RNA quality

Transcript quality

Count statistics

Batch correction?

Read quality

MultiQC for summary

**FastQC
(Trimmomatic: Filter, trim, adapter removal)**

Workflow

Experimental design ▼ RNA extraction ▼ Library preparation ▼ Sequencing ▼ Data processing ▼ Data analysis

QC
▼
Mapping
▼
Quantification

Reads
▼
Genome   Transcriptome
▼
Alignment

De-novo assembly

Annotation

Correlation
▼
Clustering
▼
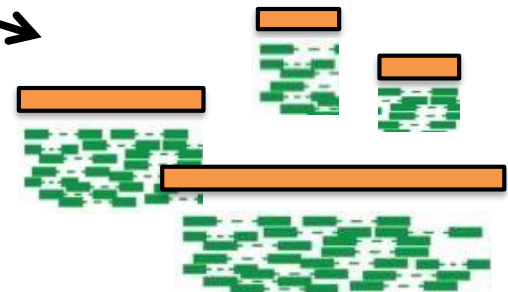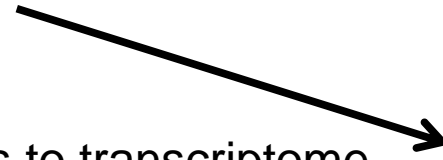MDS/PCA
▼
Differential Gene Expression
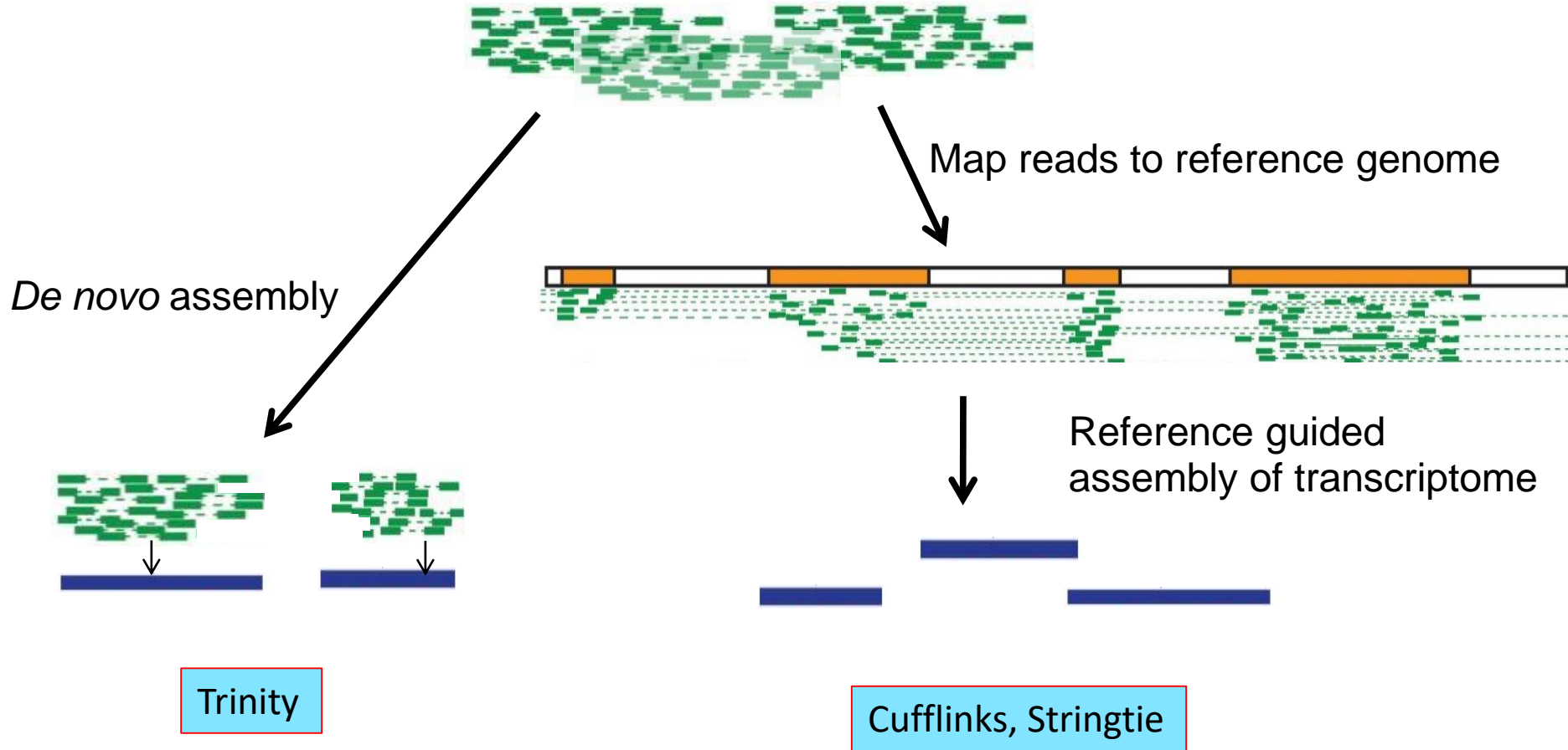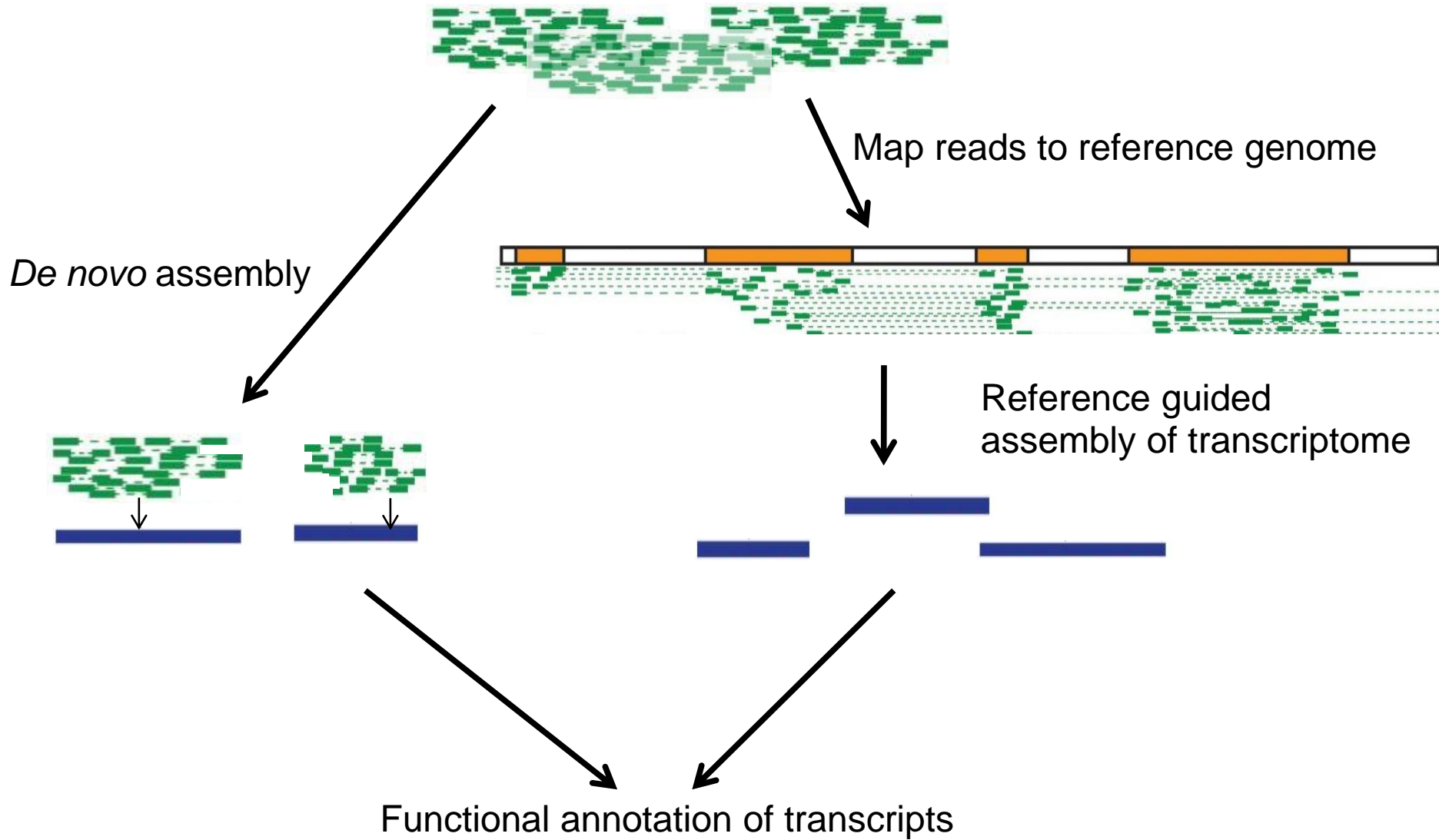
From reads to bamfile

Map reads to reference genome
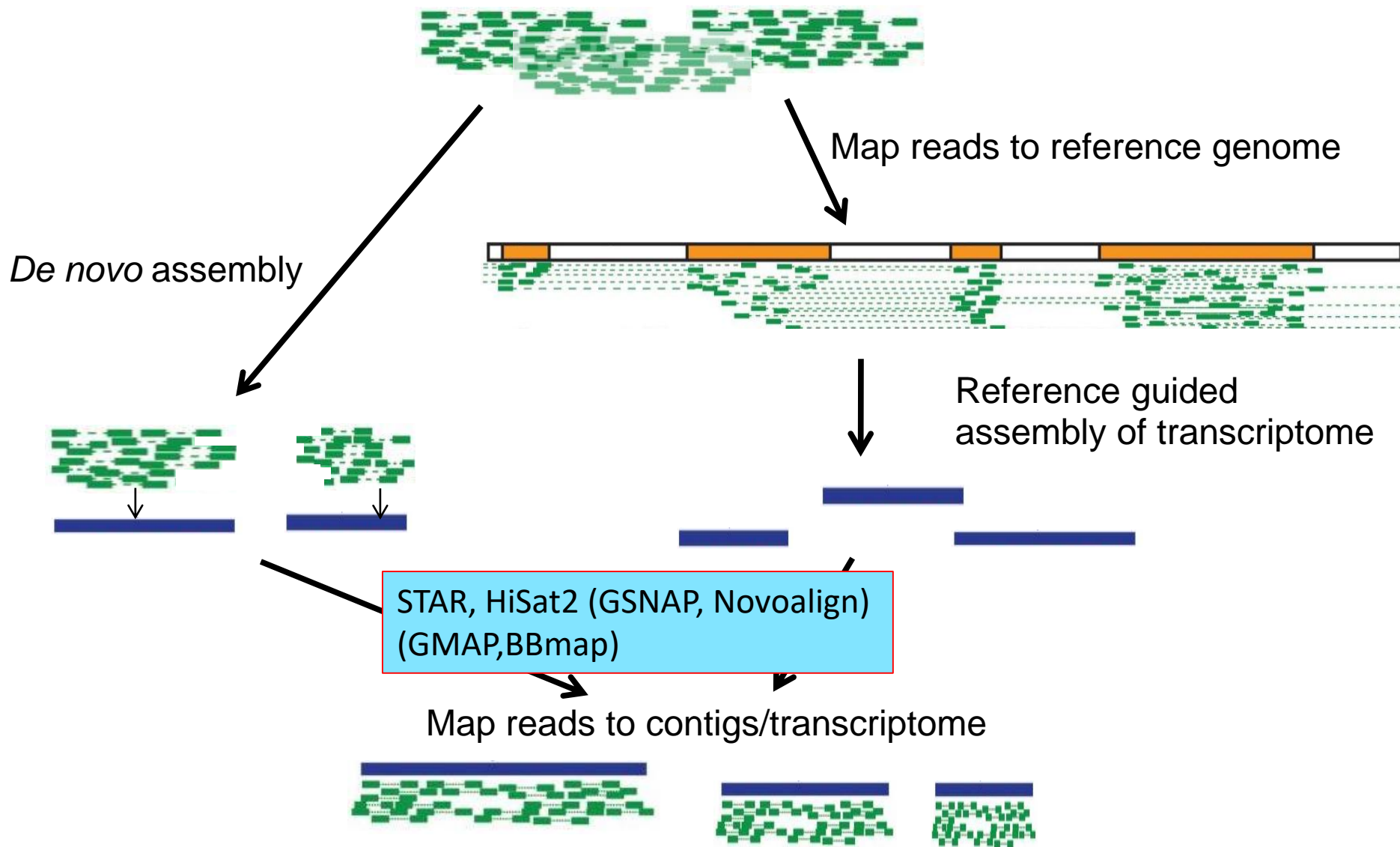
STAR, HiSat2 (GSNAP, Novoalign)
(GMAP,BBmap)

Map reads to transcriptome

From reads to bamfile

De novo assembly

Map reads to reference genome

Reference guided
assembly of transcriptome

Trinity

Cufflinks, Stringtie

De novo assembly

Map reads to reference genome

Reference guided assembly of transcriptome

Functional annotation of transcripts

From reads to bamfile

Map reads to reference genome

De novo assembly

Reference guided assembly of transcriptome

STAR, HiSat2 (GSNAP, Novoalign) (GMAP,BBmap)

Map reads to contigs/transcriptome

Bam to counts
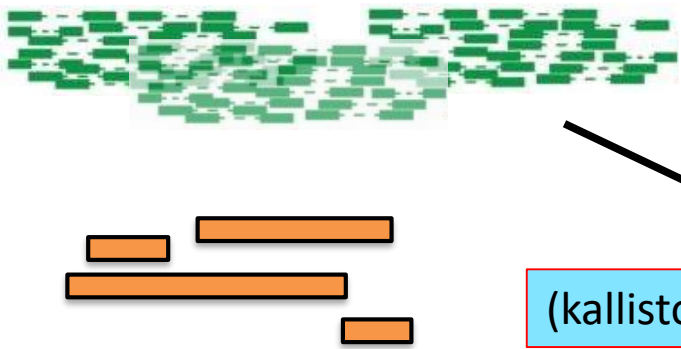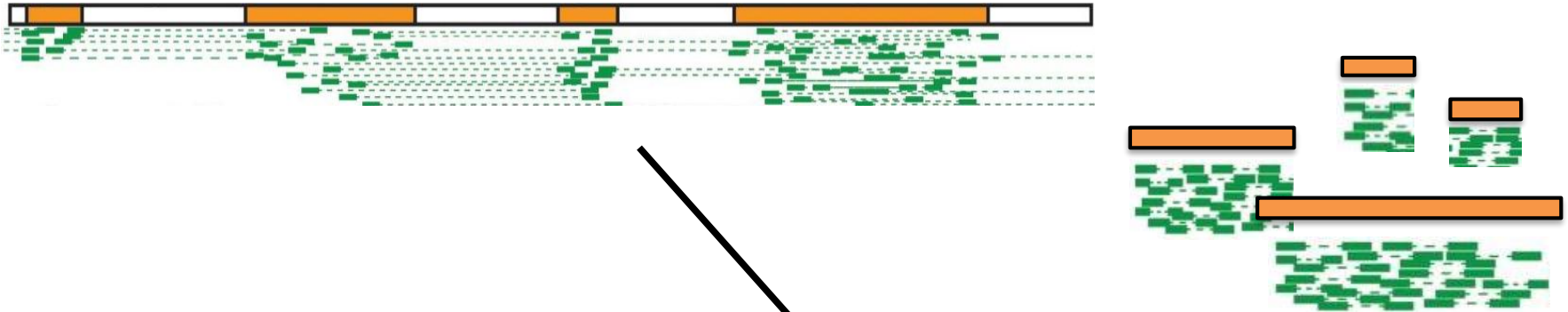
Differential expression analysis

```
              0h_1      0h_2      0h_3 ...
ENSG00000000003 321     303       204    ...
ENSG00000000005 0       0         0      ...
ENSG00000000419 696     660       472    ...
...
```
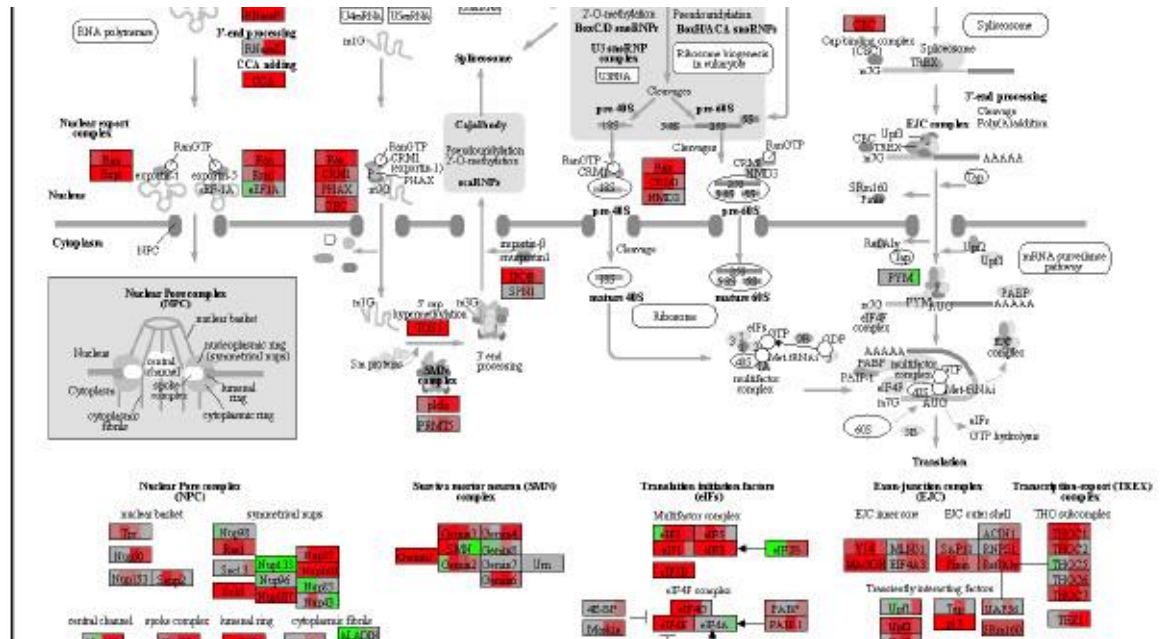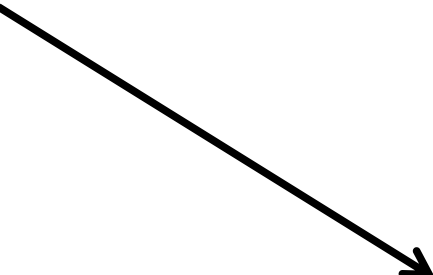
DESeq2 (EdgeR, sleuth, …)

samples

genes

Gene set analysis

samples

genes

All steps will affect the results
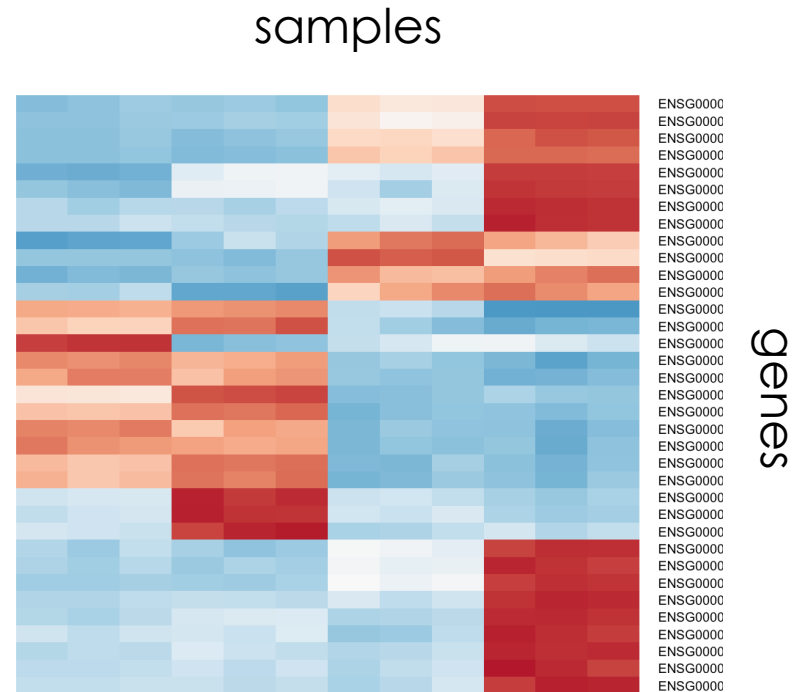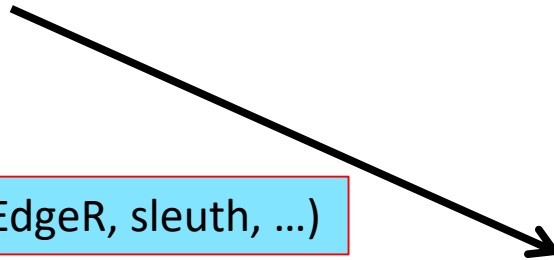
All RNA

All steps will affect the results
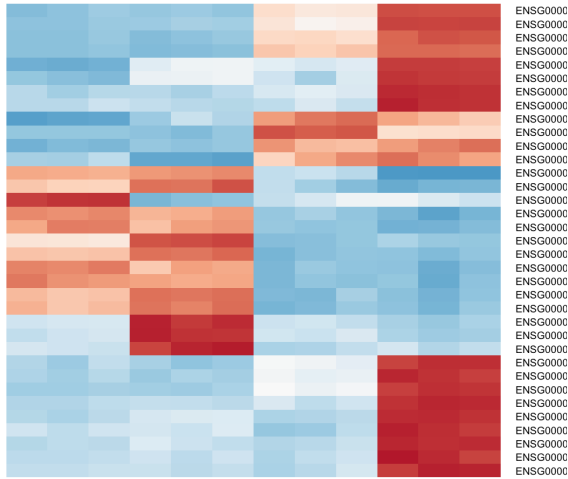
Experimental setup

All steps will affect the results

Lab work + RNA extraction

RNA enrichment protocoll
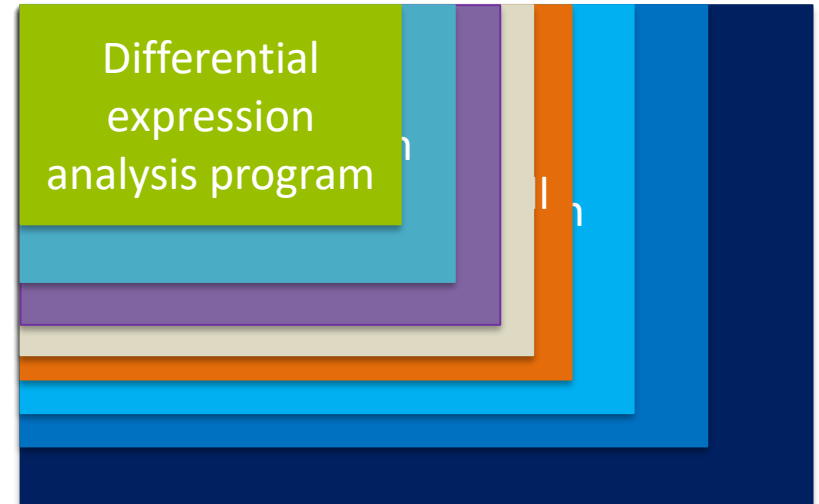
Sequencing machine

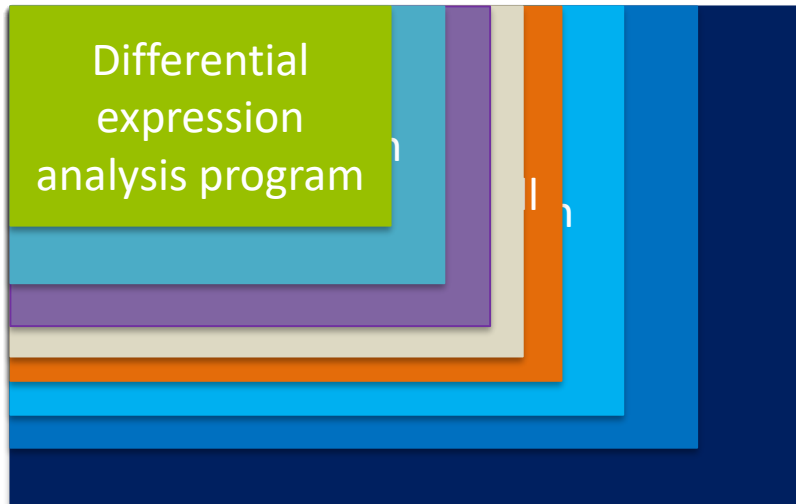Reference

Mapping program

All steps will affect the results
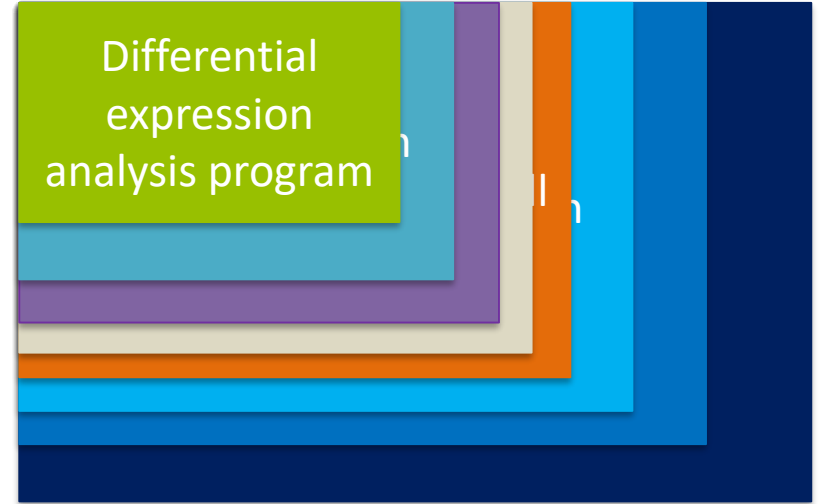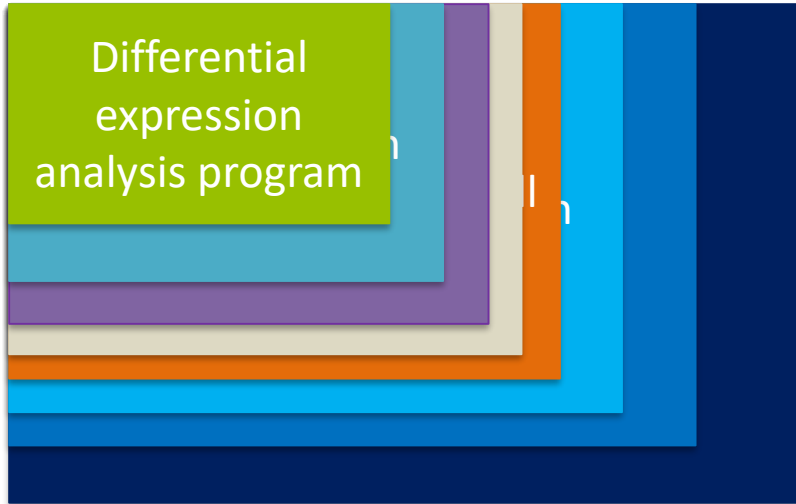
Differential expression analysis program

Try to be as consistent as possible

Finally…

- Thanks for joining the workshop!

- Please fill out the evaluation form
- Please leave name tags in the box – if you still have them…

- Good luck with your analysis! If you need help, we are here for you!

http://nbis.se/support