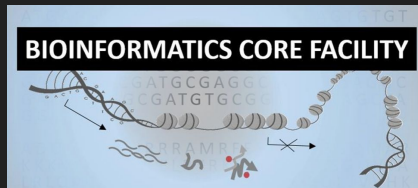


Welcome to ABC.10

12. December 2024

<https://abc.au.dk>

abcafe@au.dk



Health
Data Science
Sandbox

Agenda

- What's new
- Interesting things from you?
- Topic presentation
- Tutorial and/or open coding

What's new

Some updates

- We have a unified mail for the ABC:
abcafe@au.dk
- Use our newsletter and suggestion forms at the homepage
abc.au.dk

What's new

Some updates

- Next cafe dates

Jan 16

Jan 30

Feb 6

Feb 27

Mar 13

- You will always get an invitation if you are on the newsletter

Interesting things from you?

Topic presentation

Bulk RNA seq data alignment and
quality control (QC)

Tutorial topic

~~Bulk RNA seq data alignment and
quality control (QC)~~

Previous tutorials

RNA-seq

What is it?

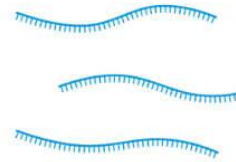
It's a technology used to study RNA molecules in a biological sample

The purpose of RNA-seq:

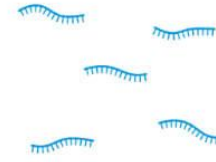
- to quantify gene expression (what genes are active or not)
- to compare conditions or changes over time (healthy vs diseased, treated/stimulated vs untreated)

RNA Sequencing

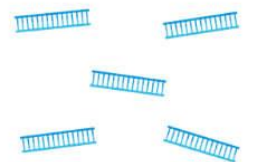
① Isolate RNA from samples



② Fragment RNA into short segments



③ Convert RNA fragments into cDNA



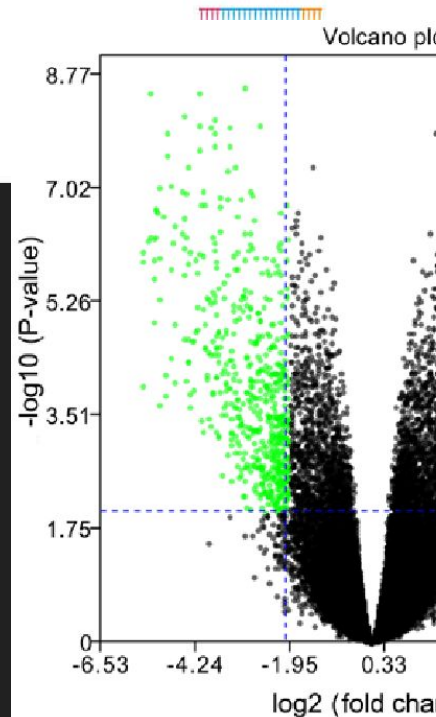
④ Ligate sequencing adapters and amplify



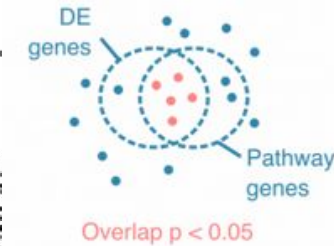
⑤ Perform NGS sequencing



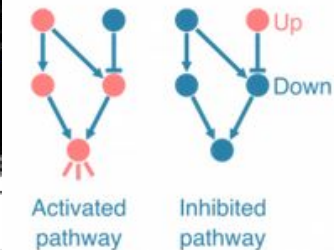
⑥ Map sequencing reads to the transcriptome/genome



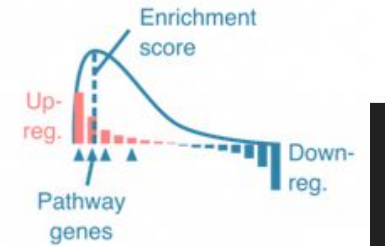
Over-representation analysis



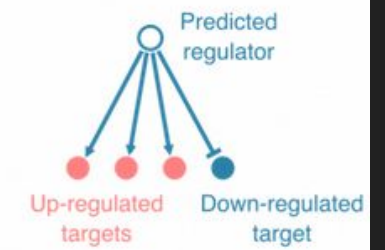
Regulatory pathway analysis



Gene set enrichment analysis



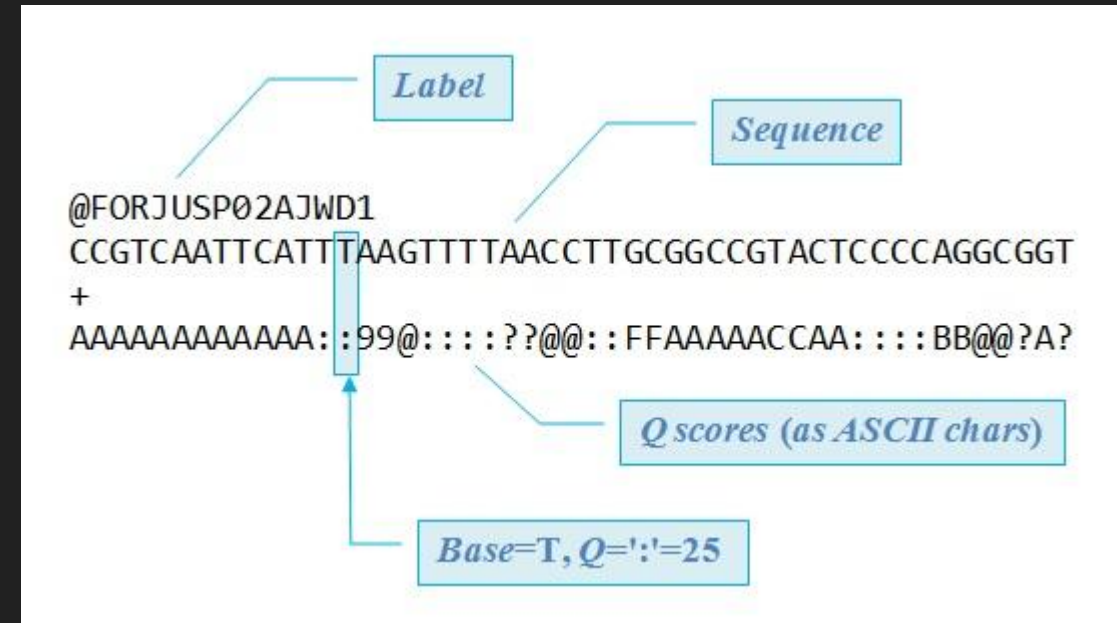
Upstream regulator analysis



Raw RNA seq data

What are they?

- unprocessed results from sequencing RNA molecules from a sample (sometimes with a focus on a specific type of RNA, like mRNA)
- they're stored in FASTQ files, which contain
 - line 1: the sequence identifier
 - line 2: the RNA sequence
 - line 3: a separator
 - line 4: the quality scores
 - extra: noise (adaptors, sequencing errors, or contamination)



How to process RNA seq data

Alignment:

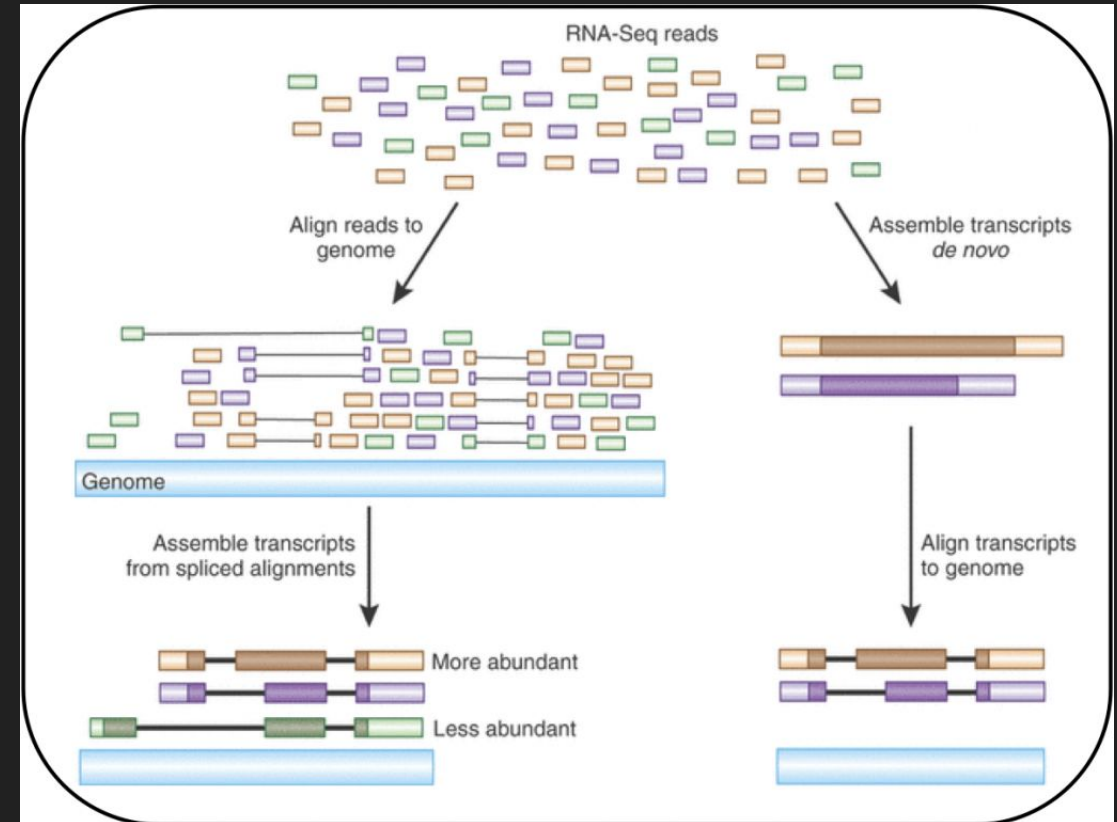
- preassembled reference genome
 - .fasta/.fa (reference genome) and .gft (annotation file)
- de novo assembly

Alignment tools:

- STAR, HISAT2, bowtie2

QC

- fastqc
- multiqc



Tutorial and open coding

At our home <https://abc.au.dk/Documentation> you can find

- Previous tutorials
- 1 conference workshop

Or you can code and ask for help or generic coding/bioinf/data science questions

