Differential gene expression (DGE) analysis from bulk RNA-seq data

Consultation guidelines

The following guidelines outline a few points you should consider before discussing your project with the Next Generation Sequencing Platform (NGSP) and/or the Interfaculty Bioinformatics Unit (IBU). This will ensure that all important points are discussed and we can provide you with the best possible advice on your project. **Please note that if you require bioinformatics support, it is best to include NGSP and IBU in the same meeting.**

Bulk RNA-seq experiments are typically aimed at identifying differentially expressed (DE) genes between two or more experimental groups.

1. *General experimental design*

   a) What is the biological question you want to address?
   b) What is your study species?
   c) What experimental design do you have in mind? This includes, for example, the number of experimental groups, potential confounding variables (e.g. sex, age, BMI etc), number of time points etc. **It is generally very helpful to have a schematic representation of the design for discussion.**
   d) Can you estimate how large the difference between your biological replicates (i.e. individuals) is likely to be? For example, we expect inbred mice in a lab setting to be more similar to each other than different human patients which differ in genetic background, age, diet etc.
   e) Is gene-level information sufficient to address your research question or do you need isoform-level resolution?

2. *Library preparation and sequencing*

   a) Which tissue(s) will be used and how will the tissue(s) be stored before RNA extraction?
   b) What is the expected quality and quantity of RNA that will be available? This information is critical to choose the optimal library preparation protocol for your study and should, ideally, be tested in a small pilot.
   c) Will you be able to collect sufficient material to have a backup aliquot for all samples?
   d) Are you specifically interested in transcripts that are not polyadenylated?
   e) Are you interested in miRNAs?

3. *Bioinformatic processing*

   Who will be doing the bioinformatics processing of your data? The following models are possible:

   i. No bioinformatics. In this case, you will receive fastq files for your samples.

ii. Standard DGE analysis as outlined here:
https://www.bioinformatics.unibe.ch/services/service_mode/bulk_rna_seq_differential_gene_expression/index_eng.html
Here, you will receive a basic analysis, and additional downstream analysis will be run by your group. This will require some basic experience with software for plotting and statistical analysis such as R, but does not usually need a high performance computing infrastructure.

iii. Standard DGE analysis plus further customized downstream analyses. Here, please prepare a description of additional analyses/visualisations you would need. For example, you could refer us to selected publications that include the type of analysis you have in mind.