**Bioinformatics Analysis Request Template**

**revised on 2025-02-21**

The sole purpose of this document is to align the wet-lab researchers and the bioinformaticians in terms of the analysis requests and preferred output.

**Project name**: A not-so-long name/abbreviation to refer to the project. This will be the “nickname” of the project and will be used as name of the Teams channel and the git repo that will be used throughout the project.

Example: *Roel’s project on fetal radiation exposure is called “FRE”*

**WBS code:**

**Stakeholders**: Researchers other than the scientist that make the analysis request so they can be included in the communications.

**Project description**: One or two sentences explaining the project. The point is not to increase bureaucracy but to make sure the analyst comprehends the general aim of the project.

Example: *Understanding whether different cell lines respond differently to irradiation.*

**Raw data path**: The location of the data if they are already on the server/Store9 or the link from the sequencing facility/company.

**Sample meta data**: Information about all the samples in a table where rows are samples and columns are attributes. This information will be provided via filling in the “samplesheet.csv”. Using MS Excel to fill in this file is OK as long as one does not use hidden/merged cells, formatting or whitespaces or special characters a column/feature names. A minimal example is shown below:

*sample\_name,treatment\_type,treatment\_dose,batch,age\_weeks,sex*

*sample1,control,0,batch1,8,M*

*sample2,drug1,10,batch1,7,F*

*sample3,drug1,20,batch1,8,F*

**Hypotheses to be tested**: Unlike the “project description” part, this is crucial for the analysis as these will be converted to “contrasts” (case vs control while controlling for batch) in the case of differential expression. This information will be provided via the “contrasts.csv” or as bullet points. Basically the researchers need to specify what are to be “contrasted” and what variables need to be “accounted for”.

Bullet point example:

* *10 Gy vs 0 Gy while controlling for the cell line*
* *10 Gy vs 0 Gy specifically for cell line LBT005*

contrastst.csv example:

*id,variable,reference,target,blocking*

*condition\_control\_treated,condition,control,treated,*

*condition\_control\_treated\_blockrep,condition,control,treated,replicate;batch*

“id” is an arbitrary name for the contrast, “variable” is the name of the column within the samplesheet.csv that contain levels, “reference” and “target”. “blocking” is/are the column(s) used to account for variables.

**Organism**: Human, mouse, …

**MSigDB Collections** (In case of DEA from RNA-seq data): Gene signatures of interest to perform GSEA:

* Human: <https://www.gsea-msigdb.org/gsea/msigdb/human/collections.jsp>
* Mouse : <https://www.gsea-msigdb.org/gsea/msigdb/mouse/collections.jsp>

**Result Output:** Expected/desired output

Example: *For a typical differential expression analysis that would be a table of differential expression results, a heatmap, a volcano plot, and for an alternative splicing that would be a Sashimi plot*