User manual demo version of Transcriptome-wide outlier detection and filtering application

Important information before start analysis:

This is a demo version of the web application, in which only part of the data is loaded to ensure stable and smooth operation via MyBinder. Therefore, results presented in this version of the web browser application may slightly deviate from those reported in the Dekker & Schot et al. manuscript. Complete exonic and intronic Z-score data are made available only for genes discussed in the accompanying manuscript and a set of randomly chosen genes (~1.5% of all) . Genes available for analysis (n=295) are listed in Table S6.

User Manual

1. Starting up web browser application (Chrome or Firefox) and sample selection

  1.1 Go to: https://mybinder.org/v2/gh/KlinGenErasmusMC/rnaseq-clingendiag/HEAD?urlpath=voila%2Frender%2Frnaseq-filtering-app.ipynb
(Generally takes a few minutes to load)

  1.2 Select experiments and species

meta dekker_et_al_rnaseq_metadata ⌄	➜ Leave as default
species fib_untreated ⌄	➜ Select sample type (e.g. "*fib_untreated*")
fragment genes ⌄	➜ Z-score type (e.g. "*exons*")
Run Interact	

Select *"Run Interact"* to start the analysis.

  1.3 Select sample

Run Interact

ID refresh ⌄	➜ Select sample (e.g."*Dekker_et_al_individual4_untreated*")

Select *"Run Interact"*.

1.4 **Show PCA plot**

Run Interact

Select *"Run Interact"* to generate PCA plot. Sample selected for analysis is indicated in red, the other 8 samples are indicated in blue. Cursor hovering provides details on sample name.

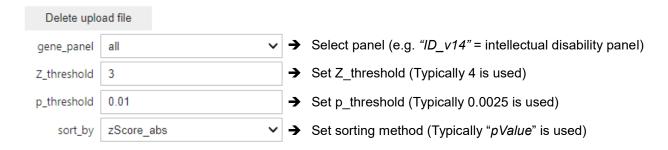**1.5** **Select results-file**

Loading the data might take a while (even up to 5 minutes).

Run Interact

`<function __main__.path_results()>`

Select *"Run Interact"*. Z-scores will now be loaded into the application.

2. RNA-seq analysis: filtering on gene panel, Z-score and p-value

2.1 **Filter on Z-scores and p-value, HPO terms, ROH, genomic position**

Delete upload file		
gene_panel	all ⌄	➔ Select panel (e.g. *"ID_v14"* = intellectual disability panel)
Z_threshold	3	➔ Set Z_threshold (Typically 4 is used)
p_threshold	0.01	➔ Set p_threshold (Typically 0.0025 is used)
sort_by	zScore_abs ⌄	➔ Set sorting method (Typically *"pValue"* is used)

The app offers various other filtering options in this section, for example filtering on specific gene lists, HPO terms or ROH regions.

Select *"Run Interact"* at the bottom of this section. A Z-score ranking plot and a list of genes, exons or introns that meet the filtering settings will be generated.
This table can be downloaded by selecting *"Save Table"*

2.2

Volcano plot

Z score threshold:	5	→ Set Z-threshold (e.g. 4)
p-value threshold:	0.01	→ Set p-threshold (e.g. 0.0025)
max number of annotated genes:	100	→ Total number of significant genes/exon/introns shown

show only panelgenes? ● yes ○ no

→ By selecting "*yes*" only genes, exons or introns in selected gene panel are highlighted in red
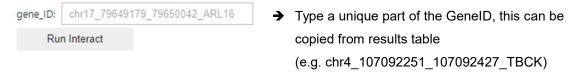
Run Interact

Select *"Run Interact"* to generate the Volcano plot. Volcano plots highlight all genic, exonic or intronic Z-score outliers present in the selected gene panel in red. Genes, exons or introns, which are not an outlier or not present in the selected gene panel are indicated in blue. Cursor hovering provides Z-score and p-value details.

2.3

All fragments in Gene of Interest

search gene: RALGAPA1

→ Type gene name (e.g. TBCK)

NB: This field is case sensitive

Run Interact

Select *"Run Interact"* to generate a fragment plot of one gene. Cursor hovering provides chromosome coordinate and Z-score details when analyzing at exon-level.

2.4

Exon ranking plot

gene_ID: chr17_79649179_79650042_ARL16

→ Type a unique part of the GeneID, this can be copied from results table
(e.g. chr4_107092251_107092427_TBCK)

Run Interact

Select *"Run Interact"* to generate an exon ranking plot (may take a few minutes).

3. Additional features

3.1

Z-score on chromosome

N_Z_scores 100000

Run Interact

This will generate plots showing genic, exonic or intronic Z-scores plotted across all chromosomes.

3.2 Panels of filtered genes

```
    Run Interact
<function __main__.panel_filtered_genes()>

  Save table panels
```

Select *"Run interact"*. This option shows in which gene panel(s) the filtered genes from the results table are present.

This table can be downloaded by selecting *"Save table panels"*