



RaNA-Seq

USER MANUAL

Version 1.0



CONTENT

ABOUT THIS MANUAL	3
WHY RANA-SEQ?	3
WHAT IS RANA-SEQ?	3
MAIN FEATURES OF RANA-SEQ	4
ANALYSIS PIPELINE OF RANA-SEQ	4
GRAPHICAL USER INTERFACE	6
LOGIN	6
NEW ANALYSIS	6
<i>Input Files</i>	6
<i>Upload New Files</i>	7
<i>Get ENA samples</i>	9
RESULTS	10
<i>Quantification</i>	11
<i>Quality Control</i>	12
<i>Differential Expression Results</i>	13
<i>Functional Enrichment</i>	15
<i>Gene Set Enrichment Analysis (GSEA)</i>	17
LOG	19
<i>Files</i>	19
<i>Tasks</i>	19

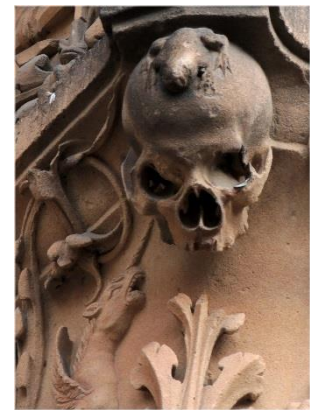


About This Manual

This is the user guide for the web application RaNA-Seq. It introduces to the graphical user interface for accessing RaNA-Seq functionality and describes the analysis protocols. If you are looking for a quick tutorial, please see our video tutorial on the Web.

Why RaNA-Seq?

Our University facade is a plateresque design (a type of intricate stonework seen on many of the famous buildings of Salamanca) containing many highly decorative stone carvings. There is a small frog sat atop a skull carved into the facade which is difficult to find. The story says that the students studying at university had to try and find the tiny frog amongst the carvings on the facade, and if they managed this without help they would pass their exams. Rana means frog in Spanish, and this is the explanation of the name. Once you have found the RaNA, we hope you will pass all your analyses.



What is RaNA-Seq?

RaNA-Seq is an open bioinformatics tool for the quick analysis of RNA-Seq data. It performs a full analysis in minutes quantifying FASTQ files, calculating quality control metrics, running differential expression analyses and enabling the interpretation of results with functional analyses. Our analysis pipeline integrates cutting edge bioinformatics tools and simplify its application with a friendly Web interface designed for non-experienced users in these analyses. Each analysis can be customized setting up input parameters and applies generally accepted and reproducible protocols. Analysis results are presented as interactive graphics and reports, ready for their interpretation and publication.



Main features of RaNA-Seq

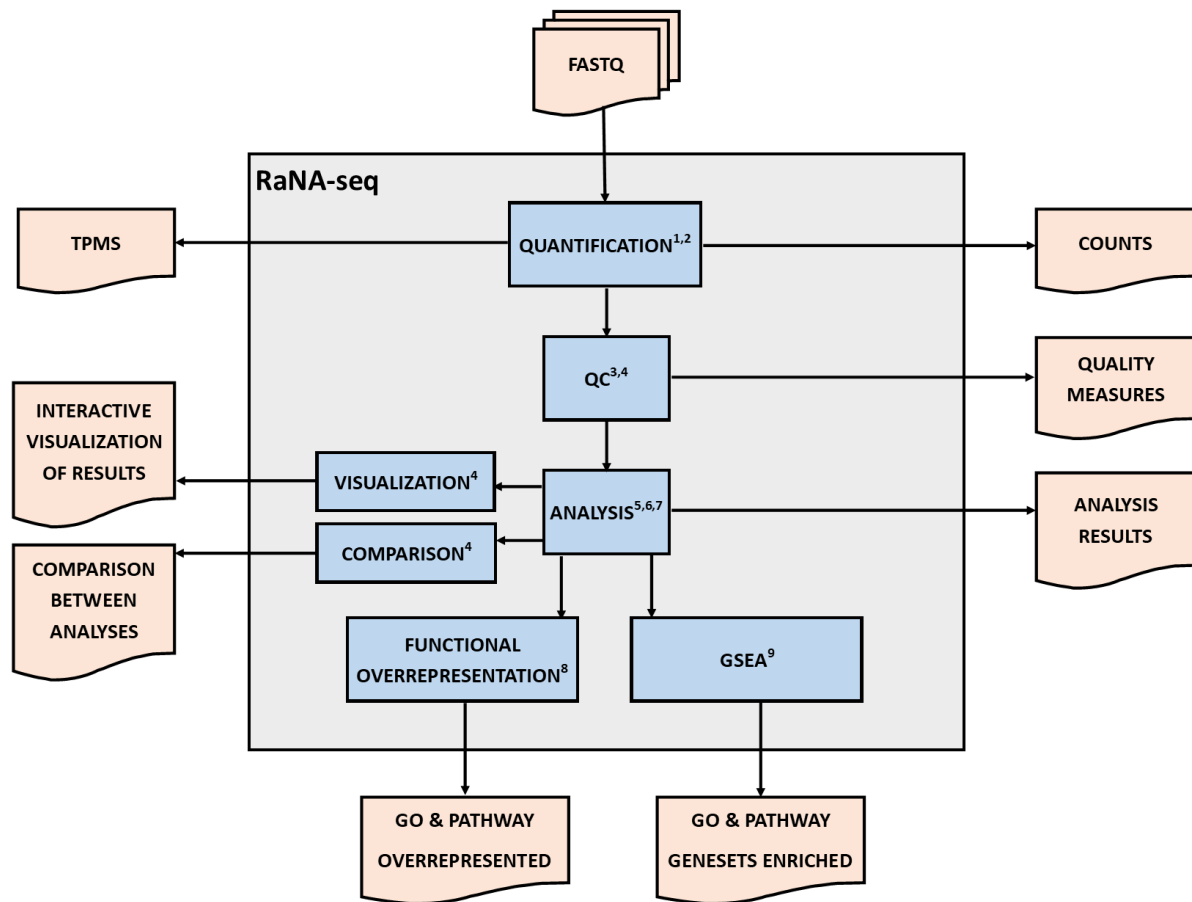
- Full RNA-Seq analysis from the FASTQ file to the functional analysis of results
- Stand alone website without any software installation required
- Connected with the ENA repository
- Auto-configuration of input parameters oriented to non-experienced users
- Quality control of input samples with several graphs and metrics
- Customization of analyses with input parameters and several differential expression methods
- Functional over-representation analysis of results
- Gene Set Enrichment Analysis of results (GSEA)
- Generation of reports with full information about the analysis, ready for its inclusion in publications
- Presentation of results with interactive graphics
- Generally accepted and reproducible protocols

Analysis Pipeline of RaNA-Seq

Figure 1 shows a flowchart which describe the analysis pipeline. It was designed to deal with raw sequencing data files in FASTQ format. Once the files are uploaded, reads are filtered based on their quality and are processed with Salmon software to obtain the quantification of reads per gene. These reads are also normalized as TPM (Transcripts Per Million, it should be read as "for every million of RNA molecules in the sample, x came from this gene/transcript"). Next steps use counts and TPM matrix for the analysis and FASTQ files are deleted from RaNA-Seq. The next step of the analysis is the differential expression analysis between conditions. It can be performed with DESeq2, EdgeR and limma analysis tools. We recommend repeating the differential expression analyses modifying the analysis tools. Several reviews have concluded that it is necessary to adapt the analysis to each study design. RaNA-Seq allows the quick execution of different analysis methods and the comparison or results with Venn diagrams and heatmaps. The final step is the functional analysis of results. RaNA-Seq performs a functional overrepresentation analysis and a Gene Set Expression Analysis (GSEA) of results.



Figure 1 Flow chart of RaNA-Seq analysis protocol.



1. Chen,S. et al. (2018) Fastp: An ultra-fast all-in-one FASTQ preprocessor. In, Bioinformatics., pp. i884–i890.
2. Patro,R. et al. (2017) Salmon provides fast and bias-aware quantification of transcript expression. Nat. Methods, 14, 417–419.
3. Schulze,S.K. et al. (2012) SERE: Single-parameter quality control and sample comparison for RNA-Seq. BMC Genomics, 13.
4. Barrios,D. and Prieto,C. (2018) RJSplot: Interactive Graphs with R. Mol. Inform., 37.
5. Love,M.I. et al. (2014) Differential analysis of count data - the DESeq2 package. Genome Biol., 15, 550.
6. Robinson,M.D. et al. (2010) edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics, 26, 139–140.
7. Law,C.W. et al. (2014) Voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. Genome Biol., 15.
8. Young,M.D. et al. (2010) Gene ontology analysis for RNA-seq: accounting for selection bias. Genome Biol., 11.
9. Sergushichev,A. (2016) An algorithm for fast preranked gene set enrichment analysis using cumulative statistic calculation. bioRxiv, 60012.



Graphical User Interface

Login

RaNA-Seq allows the registration of user accounts. Registered users will be notified when a dataset is uploaded or an analysis is finished. Moreover, uploaded files, results and logs will be stored in the user data and will be available for a further exploration. Registration process is simple, the user must fill all the form fields and accept the privacy policy.

Figure 2: Login interface

Login

Email

Password

Login

[Lost password](#)
[Anonymous access](#)

Register a new account

Full Name

Position

Research group

Department

Institution

Country

Email

Password

I have read and accept the [privacy policy](#)

Register

New Analysis

Input Files

Figure 3 shows the web interface of “input files”. The table shows information of the files which has been uploaded to RaNA-Seq:

- Filename: name of each uploaded file.
- File name 2: When the sequencing is pair-end, it contains the file name of the second file on the paired sequencing.
- Sample name: name of each sequenced sample.
- Project name: sequencing project name.
- : This button opens a window with quantification statistics.
- : Delete the sample.



The user can order the table clicking on the arrows, show more samples selecting the number on the Show entries menu, select all samples clicking on the “Select all” samples, deselect samples clicking on the “Select none” samples, delete selected samples clicking on the “Delete” button and change to a Projects view which group samples into projects clicking on “File/Projects view”. Moreover, test samples can be loaded with the button “Load Test Samples”.

Figure 3: Web interface of user uploaded files

Input Files

Select **input files** for the analysis or [upload new files](#)

Select all
Select none
Delete
File/Projects View

Load Test Samples

Show 10 entries
Search:

	Filename	Filename 2	Sample name	Project name	
<input type="checkbox"/>	SRR5615379_1 ✓	SRR5615379_2 ✓	Nrf2_KO_2 ↗	PRJNA388152 ↗	🔍 ✕
<input type="checkbox"/>	SRR5615378_1 ✓	SRR5615378_2 ✓	Nrf2_KO_1 ↗	PRJNA388152 ↗	🔍 ✕
<input type="checkbox"/>	SRR5615377_1 ✓	SRR5615377_2 ✓	WT_2 ↗	PRJNA388152 ↗	🔍 ✕
<input type="checkbox"/>	SRR5615376_1 ✓	SRR5615376_2 ✓	WT_1 ↗	PRJNA388152 ↗	🔍 ✕
<input type="checkbox"/>	SRR4016943_1.fastq.gz ✓	SRR4016943_2.fastq.gz ✓	SRR4016943 ↗	PRJNA338329 ↗	🔍 ✕
<input type="checkbox"/>	SRR4016944_1.fastq.gz ✓	SRR4016944_2.fastq.gz ✓	SRR4016944 ↗	PRJNA338329 ↗	🔍 ✕
<input type="checkbox"/>	SRR4016945_1.fastq.gz ✓	SRR4016945_2.fastq.gz ✓	SRR4016945 ↗	PRJNA338329 ↗	🔍 ✕
<input type="checkbox"/>	SRR4016946_1.fastq.gz ✓	SRR4016946_2.fastq.gz ✓	SRR4016946 ↗	PRJNA338329 ↗	🔍 ✕
<input type="checkbox"/>	SRR4016947_1.fastq.gz ✓	SRR4016947_2.fastq.gz ✓	SRR4016947 ↗	PRJNA338329 ↗	🔍 ✕
<input type="checkbox"/>	SRR4016948_1.fastq.gz ✓	SRR4016948_2.fastq.gz ✓	SRR4016948 ↗	PRJNA338329 ↗	🔍 ✕

Previous
1
2
3
4
5
Next

Quality Control

Analysis >

Upload New Files

The user can upload new files in the “New Analysis” web page. It is necessary to set the reference genome assembly against the sample will be analyzed and select the sequencing type (single-end or pair-end). Once the reference genome and the sequencing type are set, the user can select or drag and drop the FASTQ files which will be uploaded. The window must be opened until the upload is finished.



Figure 4: Upload files interface

Upload New Files

1) Sample type

<p>Reference genome assembly ⓘ</p> <input style="width: 90%; height: 20px;" type="text" value="Write a species"/>	<p>Sequencing type ⓘ</p> <p><input type="radio"/> Single-end</p> <p><input type="radio"/> Paired-end</p>	<p>Project name ⓘ</p> <input style="width: 90%; height: 20px;" type="text" value="20190726"/>
--	---	--

Advanced ▾

2) Select your FASTQ Files ⓘ

Select or drag and drop one or more files:

No files selected.

If the number of matched read-pairs is low (under 50%), we recommend the manual setting of quantification options or the application of reads filtering before quantification (“Apply a quality filter before quantification”, see below). These are the advanced options which could be set up:

- Library Type: The pseudoaligner (Salmon) can determine the library type automatically. Uncheck the check box if you want to want to specify the library type.
- Relative Orientation: The user can specify the relative orientation of the pairs in the library. This is only possible if the library is paired-end.
- Strandness: The user can specify whether the protocol is stranded or unstranded.
- Strandness Direction: The user can specify the strand from which the read originates in a strand-specific protocol. The possible values are forward or reverse.
- Correct sequence-specific biases: Passing the 'seqBias' flag to Salmon will enable it to learn and correct for sequence-specific biases in the input data. Specifically, this model will attempt to correct for random hexamer priming bias, which results in the preferential sequencing of fragments starting with certain nucleotide motifs. By default, Salmon learns the sequence-specific bias parameters using 1,000,000 reads from the beginning of the input.
- Correct fragment-level GC biases: Passing the 'gcBias' flag to Salmon will enable it to learn and correct for fragment-level GC biases in the input data. Specifically, this model will attempt to correct for biases in how likely a sequence is to be observed based on its internal GC content.
- Fragment length distribution mean: This option is only important when running Salmon with single-end reads. Since the empirical fragment length distribution cannot be estimated from the mappings of single-end reads, the 'fldMean' allows the user to set the expected mean fragment length of the sequencing library. This value will affect the effective length correction, and hence the estimated effective lengths of the transcripts and the TPMs. The value passed to 'fldMean'



will be used as the mean of the assumed fragment length distribution (which is modeled as a truncated Gaussian with a standard deviation given by 'fldSD').

- **Fragment length distribution SD:** This option is only important when running Salmon with single-end reads. Since the empirical fragment length distribution cannot be estimated from the mappings of single-end reads, the 'fldSD' allows the user to set the expected standard deviation of the fragment length distribution of the sequencing library. This value will affect the effective length correction, and hence the estimated effective lengths of the transcripts and the TPMs. The value passed to 'fldSD' will be used as the standard deviation of the assumed fragment length distribution (which is modeled as a truncated Gaussian with a mean given by 'fldMean').
- **Quality filter before quantification:** This option applies a quality filter of reads in the FASTQ files before quantification. It is performed with the Fastp software. It filters low quality and low complexity reads and performs an automatic detection and filtering of adapters (<https://github.com/OpenGene/fastp>). For a personalized trimming of adapters and read quality filtering, please follow the [Fastp guide](#) which explains how to execute it in your computer.

Figure 5: Advanced options of quantification



Get ENA samples

RaNA-Seq allows the direct upload of SRA/ENA projects. The user has to write the project identifier of SRA/ENA and RaNA-Seq will download and quantify the input samples. In this case, the user can close the web page and our server will continue with the download. An e-mail will be sent when the download is finished.

Figure 6: Download data from ENA



Results

Results menu shows a table which summarize finished, running and failed analysis processes. When a process is finished (“done”), the user can access to static  and dynamic  results clicking on the icons. Analysis name can be renamed by clicking on the pen. Analyses can be also deleted by selecting them and clicking on the Delete button. Moreover, resulting differential expressed genes can be compared between analysis, selecting desired analyses and clicking on the compare button.

The table contains the following information:





























- Date: This column shows the date when the analysis was performed.
- Analysis Name: It shows the name of each analysis.
- Quantification: It shows if the quantification has finished. Click on the icon  to show quantification statistics, gene counts and TPMs (Transcript Per Millions).
- QC: It shows the quality control graphs of the quantified samples.
- DE Results: It shows if the differential expression analysis (DE) has finished. Click on the icon  to show a dynamic report of results. Click on the icon  to show a pdf report and a differential expression table with results.
- Func. Enrichment: This column shows if the functional enrichment analysis has finished. Click on the icon  to show a dynamic report of results. Click on the icon  to show a pdf report and Excel tables with results.
- GSEA: This column shows if the Gene Set Enrichment Analysis (GSEA) has finished. Click on the icon  to show a dynamic report of results. Click on the icon  to show a PDF report and Excel tables with results.

Figure 7: Results table

Select all Select none Compare			Search: <input type="text"/>				
	Date	Analysis Name	Quantification	QC	DE Results	Func. Enrichment	GSEA
<input type="checkbox"/>	02/07/2019	preview_1562054003 	done 	done 	-	-	-
<input type="checkbox"/>	20/12/2018	Analysis_limma 	done 	done 	done 	done 	done 
<input type="checkbox"/>	20/12/2018	Analysis_edgeR 	done 	done 	done 	done 	done 
<input type="checkbox"/>	20/12/2018	Analysis_DESeq2 	done 	done 	done 	done 	done 

Previous | 1 | Next



Quantification

Quantification section allows to download the number of counts per gene in each sample and the Transcripts Per Millions. Moreover, it shows a table with quantification statistics:

- Total reads/read pairs: Total reads/read pairs found in FASTQ files
- Matched reads/read pairs: This is the total number of fragments that could be assigned to at least one transcript (including orphan mappings).
- Properly matched reads/read pairs: This is the number of actual fragments compatible with the library type.

Figure 8: Quantification summary

Download files

- Counts [tsv](#) [xlsx](#)
- TPMs [tsv](#) [xlsx](#)
- Quantification stats [tsv](#) [xlsx](#)

Quantification statistics

Samples	Total fragments ⓘ	Matched fragments ⓘ	Properly matched fragments ⓘ
SRR5615377_1	29010267	25438035 (87.7%)	24928944 (98%)
SRR5615376_1	28916063	25438996 (88%)	25040777 (98.4%)
SRR5615379_1	23558778	20860668 (88.5%)	20561278 (98.6%)
SRR5615378_1	24523124	21562710 (87.9%)	21187518 (98.3%)

Alignment code [\[sh\]](#)

Quality control reports show five graphs which could detect errors on the input samples:

- **Box-and-whisker plots (boxplots):** it shows the distribution of expression values (normalized as TPMs) for each sample. This representation is useful to identify samples which have a different expression distribution which could be caused by a sequencing or sample collection problem.
- **Barplot:** It shows the estimated expressed genes per sample. A gene is considered as expressed when it has at least one transcript sequenced per million of sequenced transcripts. Big differences on the number of expressed genes would be caused by a library construction or a sample collection problem. Low numbers of expressed genes could be caused by a low sequencing deep. It is also recommendable to check the number of matched fragments to discard quantification problems.
- **Cluster heatmap:** Heatmap of expression similarity between samples. It shows the distance between samples calculating the SERE coefficient between each pair of samples. Similar samples have lower values in the graph. Samples with problems may have a high distance against the others. Paired samples and samples with the same input condition should tend to be clustered together.
- **PCA plot of samples:** It shows the two first PCs (principal components) values of each sample. Samples with errors may be separated from the others and paired samples and samples with the same input condition should tend to be near in the graph.

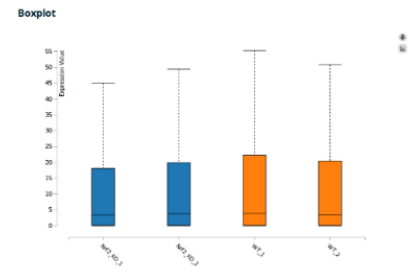


Figure 1 Box and-whisker plots (boxplots) of the expression values (normalized as TPMs) for each sample.

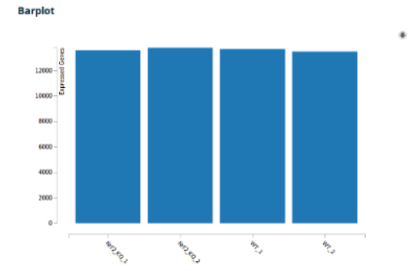


Figure 2 Barplot with the estimated expressed genes per sample.

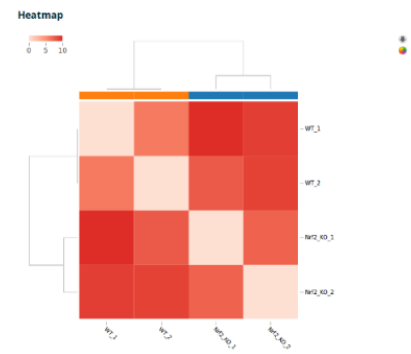


Figure 3 Heatmap of expression similarity between samples. It shows the distance between samples calculating the SERE coefficient between each pair of samples. Similar samples have lower values in the graph.

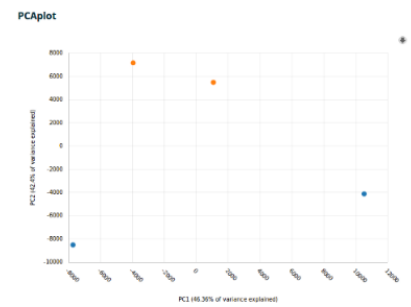


Figure 4 PCA plot of samples. It shows the two first PCs (principal components) values of each sample.

Quality control code [9]

Figure 9: Quality Control Report



Differential Expression Results

Differential expression results are showed as an interactive web page and with static reports.

Results table contains the following columns in the Excel file and the interactive web:

- ID: This column shows the Ensembl Gene Identifier of each gene.
- Symbol: This column shows the official Gene Symbol of each gene.
- ExpMean: This column shows the expression mean of each gene in the input samples. This mean was calculated with the TPMs values.
- Exp Samples A: This column shows the expression mean of each gene in the first group samples. This mean was calculated with the TPMs values.
- Exp Samples B: This column shows the expression mean of each gene in the second group samples. This mean was calculated with the TPMs values.
- log2FC: This column shows the log₂ scaled fold change of each gene. This fold change was calculated with the TPMs values.
- lfcSE: This column gives the standard error of the log₂ scaled fold change.
- Stat: This column shows the statistic calculated by the selected analysis method for each gene.
- p-value: This column shows the unadjusted p value calculated by the selected analysis method for each gene.
- padj: This column shows the adjusted p value calculated by the selected analysis method for each gene.

Results reports contains the following figures which could help to explore and interpret results:

- Volcano plot: It shows the log₂ scaled fold change (x axis) and the minus log₁₀ p-value (y axis) of each gene in the differential expression analysis. Genes with a significant expression change are highlighted as red dots.
- MA plot: This plot visualizes the differences between measurements taken in two group of samples, by transforming the data onto M (log ratio) and A (mean average) scales. It shows the log₂ scaled mean of expression (x axis) and the log₂ scaled fold change (y axis) of each gene in the differential expression analysis. Genes with a significative expression change are highlighted as red dots.
- Line chart: It shows the expression value (normalized as TPMs) of selected genes for each sample.
- Boxplot: It shows the expression value distribution (normalized as TPMs) of selected genes in each group of samples.
- Heatmap: It shows the gene expression values (normalized as TPMs) of selected genes in each sample.



Figure 10: Interactive report of differential expression results



Figure 1 (top-left) Volcano plot of differential expression results. It shows the log₂ scaled fold change (x axis) and the minus log₁₀ p-value (y axis) of each gene in the differential expression analysis. Genes with a significant expression change are highlighted as red dots. Figure 2 (top-right) MA of differential expression results. It shows the log₂ scaled mean of expression (x axis) and the log₂ scaled fold change (y axis) of each gene in the differential expression analysis. Genes with a significant expression change are highlighted as red dots. Table 1 Summary of differential expression analysis.

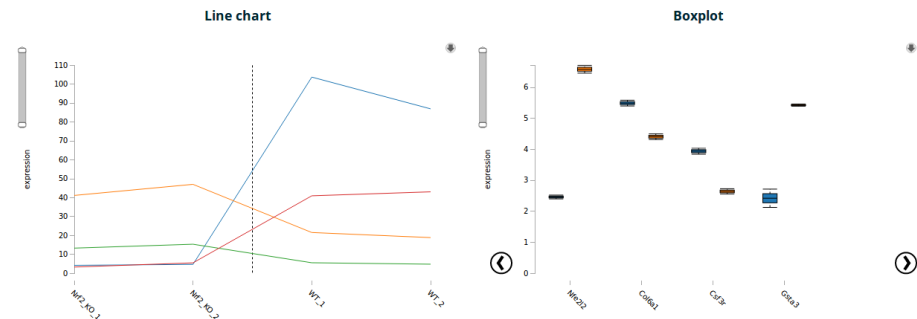


Figure 3 Line chart which shows the expression value (normalized as TPMs) of selected genes for Figure 4 Boxplot which shows the expression value distribution (normalized as TPMs) of selected genes in each group of samples defined.

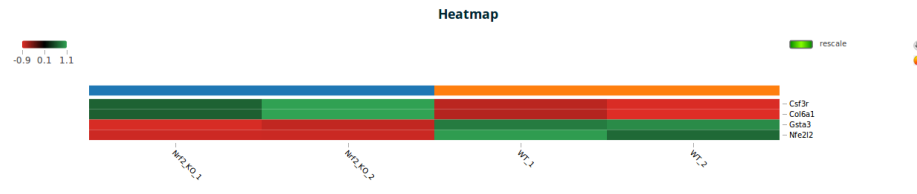


Figure 5 Heatmap of expression values (normalized as TPMs) of selected genes in each sample.



Functional Enrichment

Functional enrichment results are showed as an interactive web page and with static reports. Results table contains the following columns in the Excel file and the interactive web:

- Pathway_ID: This column shows the Pathway identifier
- Pathway: This column shows the Pathway name
- Database: This column shows the Pathway source database
- GO_ID: This column shows the Gene Ontology identifier
- GO_TERM: This column shows the Gene Ontology term. It describes the functional category in the analysis
- Category: This column describes the type of term in the Gene Ontology. Three values are possible: Functional (if it belongs to the Molecular Function category), Process (if it belongs to the Biological Process category), Location (if it belongs to the cellular location category)
- NumDEInCat: This column shows the number of differential expressed genes annotated to each term
- NumDEAnot: This column shows the number of differential expressed genes annotated to Gene Ontology
- DEPercent: This column shows the percentage of differential expressed genes annotated to each GO term
- NumInCat: This column shows the number of genes annotated to each GO term
- NumGenesAnot: This column shows the number of genes annotated to Gene Ontology
- AnotPercent: This column shows the percentage of genes annotated to each term
- p-value: p-value which measure the statistical significance of a possible functional enrichment for each term. It was calculated with GoSEQ.

Results reports contains the following figures which could help to explore and interpret results:

- Barplot: It shows the percentage of significant differential expressed genes (in blue) annotated to each functional category against whole genes (in orange).
- Bubble plot: It shows the percentage of significant differential expressed genes (in blue) annotated to each functional category against whole genes (in orange).
- Network: Each node represents a functional category. Its color represents the percentage of significant DE genes annotated this category. Edges show the functional categories with genes in common. Edge line width shows the number of genes in common. Edges can be filtered based on the number of genes in common with the slider in the bottom part of the graph.
- Symmetric heatmap: It shows the functional categories with genes in common. Blue intensity shows the number of genes in common.



GO ID	GO TERM	Category	NumOfGenes	NumOfGenes	DEPercent	NumOfCet	NumOfGenes	Assessment	p-value
GO:0006955	immune response	Process	4	76	22.222	113	10145	1.225	1.7e-2
GO:0002177	immunoglobulin production	Process	5	36	13.889	58	10145	0.51	5.8e-7
GO:0002115	defense response to fungus	Process	2	36	5.556	3	10145	0.017	1.2e-5
GO:0006032	dear leukocyte process	Process	2	36	5.556	19	10145	0.030	7.5e-4
GO:0006004	inflammatory response	Process	5	76	14.893	70	10145	1.504	2.9e-4
GO:0002083	negative regulation of immunoglobulin production	Process	2	36	5.556	13	10145	0.072	3.2e-4
GO:0047474	defense response to bacterium	Process	4	36	11.111	181	10145	0.036	3.5e-4
GO:0002753	positive regulation of immunoglobulin production	Process	2	36	5.556	19	10145	0.185	6.3e-4
GO:0006970	negative regulation of immunoglobulin synthesis process	Process	2	76	5.766	19	10145	0.145	0.002
GO:0006945	defense response	Process	5	36	5.556	49	10145	0.075	0.005

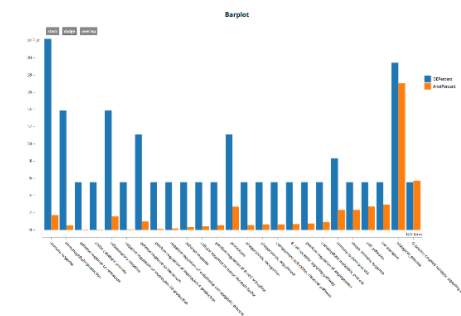


Figure 1 Barplot of percentages of significant genes in each functional category against whole genes in orange.

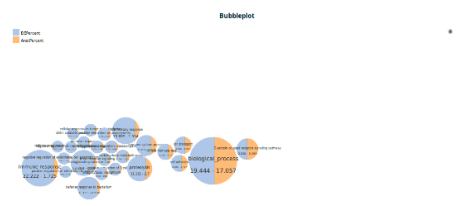


Figure 2 Bubble plot of percentages of significant DE genes in each functional category against whole genes in orange.

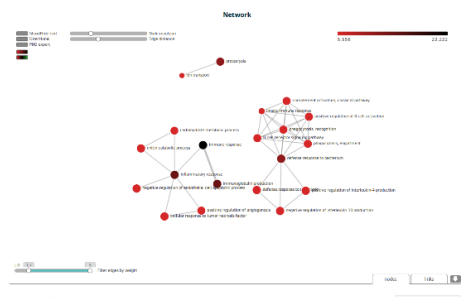


Figure 3 Network of functional analysis results. Each node represent a functional category. Its color represents the percentage of significant DE genes annotated this category. Edges show the between category number of genes in common. Edge line width shows the number of genes in common.

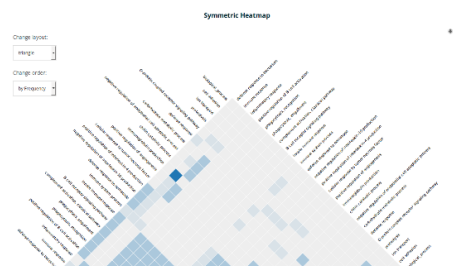


Figure 4 Symmetric heatmap which shows the functional categories with genes in common. Its color intensity shows the number of genes in common.

Figure 11: Functional Enrichment Results



Gene Set Enrichment Analysis (GSEA)

Functional enrichment results are showed as an interactive web page and with static reports. Results table contains the following columns in the Excel file and the interactive web:

- Pathway_ID: This column shows the Pathway identifier
- Pathway: This column shows the Pathway name
- Database: This column shows the Pathway source database
- GO_ID: This column shows the Gene Ontology identifier
- GO_TERM: This column shows the Gene Ontology term. It describes the functional category in the analysis
- Category: This column describes the type of term in the Gene Ontology. Three values are possible: Functional (if it belongs to the Molecular Function category), Process (if it belongs to the Biological Process category), Location (if it belongs to the cellular location category)
- numGenes: This column shows the leading-edge genes that drive the enrichment of each category
- ES: This column shows the Enrichment Score (ES), which reflects the degree to which a gene set is overrepresented at the top or bottom of a ranked list of genes.
- NES: This column shows the Normalized Enrichment Score (NES). By normalizing the enrichment score, GSEA accounts for differences in gene set size and in correlations between gene sets and the expression dataset; therefore, the normalized enrichment scores (NES) can be used to compare analysis results across gene sets.
- nMoreExtreme: This column shows the number of times that a random gene set had a more extreme enrichment score
- size: This column shows the size of the pathway after removing missing genes in input gene names
- p-value: This column shows an enrichment p-value of each category
- padj: This column shows a BH-adjusted p-value of each category

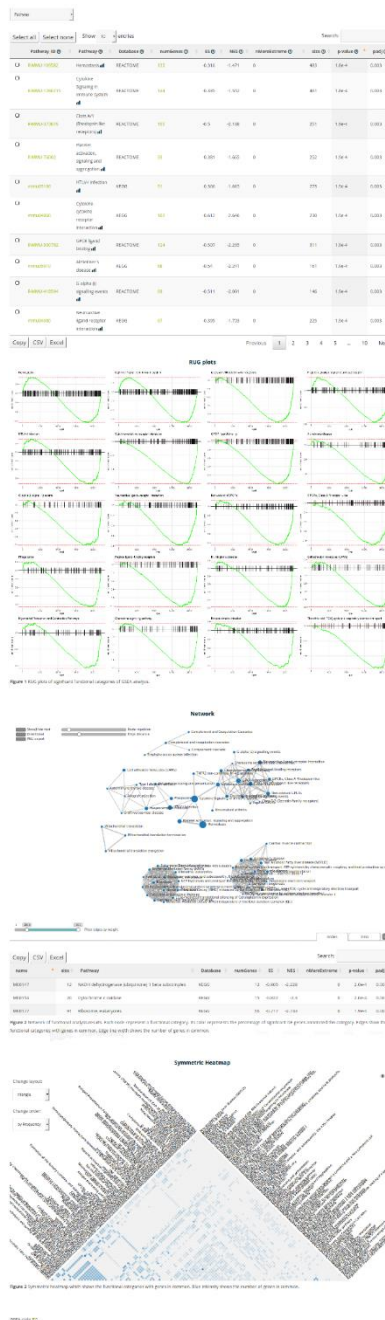
Results reports contains the following figures which could help to explore and interpret results:

- RUG plots: It shows the enrichment score against the gene rank. Genes annotated to each category are displayed as marks along the X axis.
- Barplot: It shows the percentage of significant differential expressed genes (in blue) annotated to each functional category against whole genes (in orange).
- Bubble plot: It shows the percentage of significant differential expressed genes (in blue) annotated to each functional category against whole genes (in orange).



- Network: Each node represents a functional category. Its color represents the percentage of significant DE genes annotated this category. Edges show the functional categories with genes in common. Edge line width shows the number of genes in common. Edges can be filtered based on the number of genes in common with the slider in the bottom part of the graph.
- Symmetric heatmap: It shows the functional categories with genes in common. Blue intensity shows the number of genes in common.

Figure 12: GSEA results





LOG

Files

The files log shows the information of each file that the user has uploaded to RaNA-seq. It shows if the sample is paired, the library type, the assembly and the current status. This log also allows the deletion of samples.

<input type="button" value="Select all"/> <input type="button" value="Select none"/> <input type="button" value="Delete"/>			Search: <input type="text"/>					
Date	Name	Samplename	Projectname	Paired	Libtype	Assembly	Status	
<input type="checkbox"/>	29/07/2019	SRR5615376_1	WT_1	PRJNA388152	true	A	GRCm38	quantified
<input type="checkbox"/>	29/07/2019	SRR5615376_2	WT_1	PRJNA388152	true	A	GRCm38	quantified
<input type="checkbox"/>	29/07/2019	SRR5615377_1	WT_2	PRJNA388152	true	A	GRCm38	quantified
<input type="checkbox"/>	29/07/2019	SRR5615377_2	WT_2	PRJNA388152	true	A	GRCm38	quantified
<input type="checkbox"/>	29/07/2019	SRR5615378_1	Nrf2_KO_1	PRJNA388152	true	A	GRCm38	quantified
<input type="checkbox"/>	29/07/2019	SRR5615378_2	Nrf2_KO_1	PRJNA388152	true	A	GRCm38	quantified
<input type="checkbox"/>	29/07/2019	SRR5615379_1	Nrf2_KO_2	PRJNA388152	true	A	GRCm38	quantified
<input type="checkbox"/>	29/07/2019	SRR5615379_2	Nrf2_KO_2	PRJNA388152	true	A	GRCm38	quantified

Tasks

The tasks log shows the parameters and the status of the analysis processes that the user has launched. It has links to the quantification, quality control and differential expression results. It also allows to reanalyse the samples with the link “Reanalyse”. It loads all the input parameters on the analysis web page and allow their modifications and the reanalysis.

<input type="button" value="Select all"/> <input type="button" value="Select none"/> <input type="button" value="Delete"/>			Search: <input type="text"/>									
Date	Name	Groupa	Groupb	Samples	Groups	Method	Test	Paired	Cutoff	Assembly	actions	
<input type="checkbox"/>	29/07/2019	Analysis_1564418013	Samples A	Samples B	Nrf2_KO_1, Nrf2_KO_2, WT_1, WT_2	0, 0, 1, 1	DESeq2	Wald	false	0.05	GRCm38	View counts View QCResults View DResults Reanalyse