WoPPER Tutorial 2

Salmonella enterica

(24/03/2017)

Introduction

This tutorial will show you the analysis of an RNA-seq experiment on *Salmonella enterica*, an organism with one chromosome using the NCBI annotation of the genome, without separating the two strands in the analysis.

For performing this tutorial, you will need to download the GED file for this specific experiment from the "**Tutorials**" section of WoPPER.

Tutorials

#	Descriptions	Organism	Type	# Chr	Annotation	Separated Strands	GED Files
1		*Acinetobacter baumannii*	RNA-Seq	1	NCBI	Yes	
2		*Salmonella enterica*	RNA-Seq	1	NCBI	No	
3		*Burkholderia thailandensis*	RNA-Seq	2	NCBI	Yes	
4		*Escherichia coli*	Microarray	1	NCBI	No	
5		*Helicobacter pylori*	RNA-Seq	1	Custom	Yes	

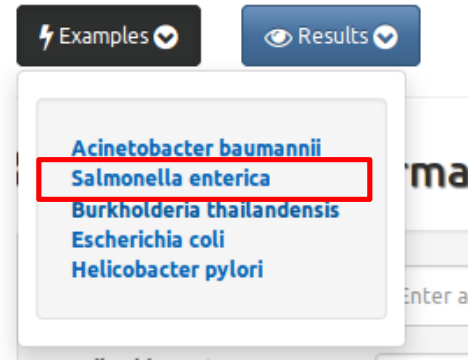
The file should be named: "**GEDfile_RNAseq-NOT-StrandSpec_Dataset_Salmonella-LT2.txt**".

An example on how the Input form should appear once completely filled in with the necessary files and parameters can be activated:

1. Clicking on the corresponding radio-button in "Examples" column of the "Tutorials" page

#	Descriptions	Organism	Type	# Chr	Annotation	Separated Strands	GED Files	Results	Examples
2		*Salmonella enterica*	RNA-Seq	1	NCBI	No			

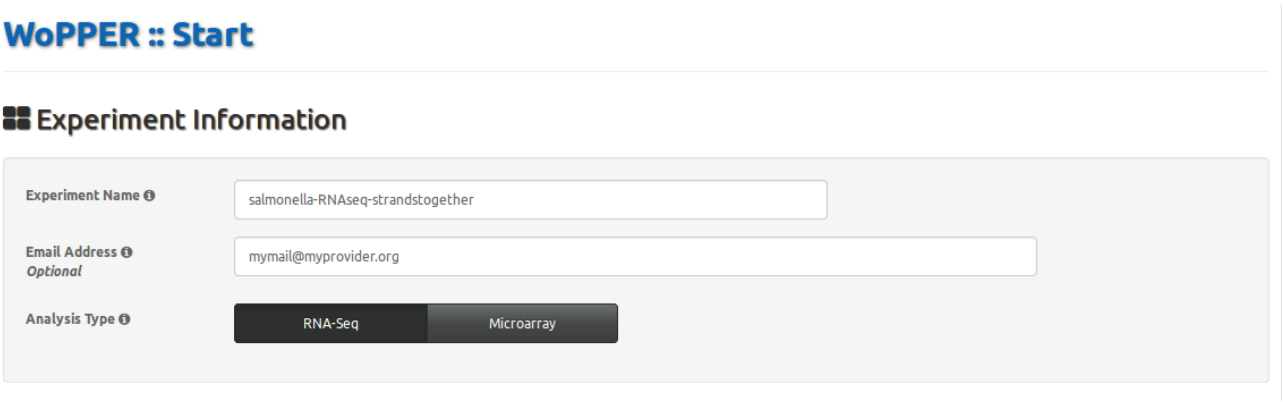
2. Selecting "Salmonella enterica" from the "Examples" drop-down menu available at the top of WoPPER "Start" page



Step-by-step procedure

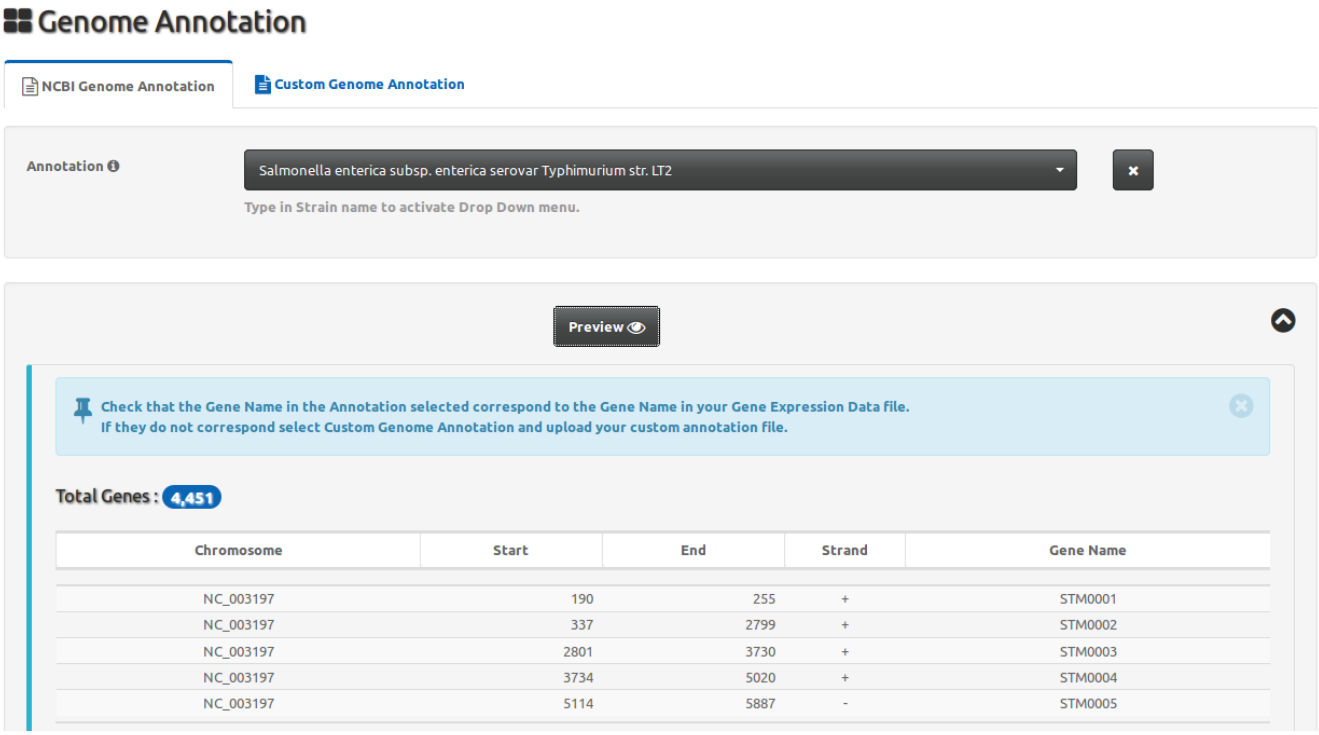
Step 1: Insert Experiment information

- Fill in the "Experiment Name" field with a suitable name for your experiment
- Fill in the "Email address" field with your preferred email address. An email pointing to the results page for your experiment will be sent to this address
- Select RNA-seq as the Analysis Type



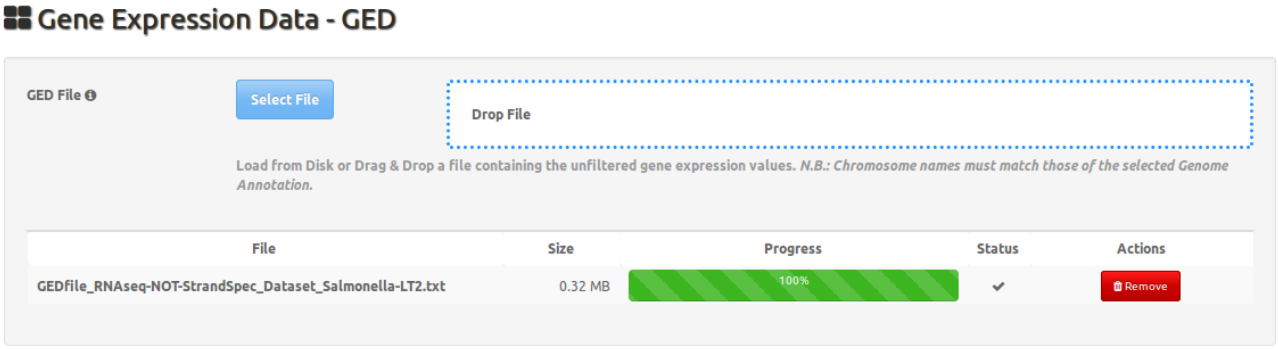
Step 2: Select Genome Annotation

- Select the "NCBI Genome Annotation" tab
- Type in "Salmo" in then web form to start the auto-fill based on a search in the internal database of bacterial genomes
- From the drop-down menu, select "Salmonella enterica subsp. Enterica serovar Typhimurium LT2" (please, make sure to select the main genome and not those annotated as "plasmid pSLT")
- Click on the"Preview" button to see the selected genome annotation

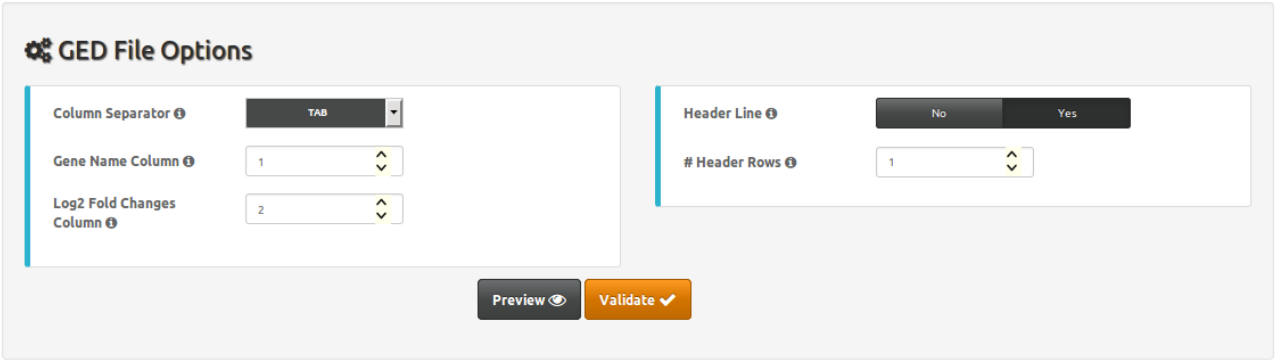


Step 3: Load Gene Expression Data

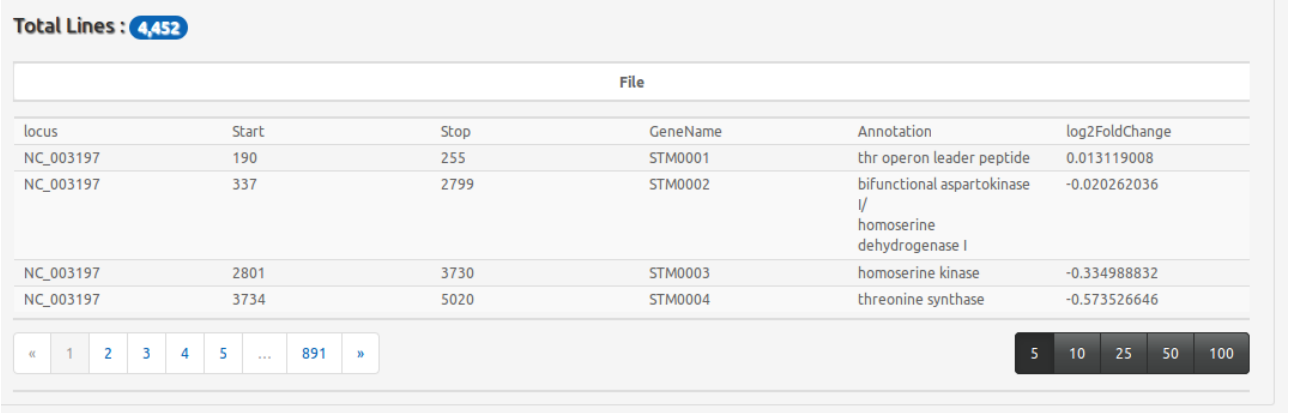
- Click on the "Select file" button and load the file named "GEDfile_RNAseq-NOT-StrandSpec_Dataset_Salmonella-LT2.txt" (alternatively, you can drag and drop the same file into the area named "Drop File")



- In the "GED File Options" section, select the values as in the reported screenshot



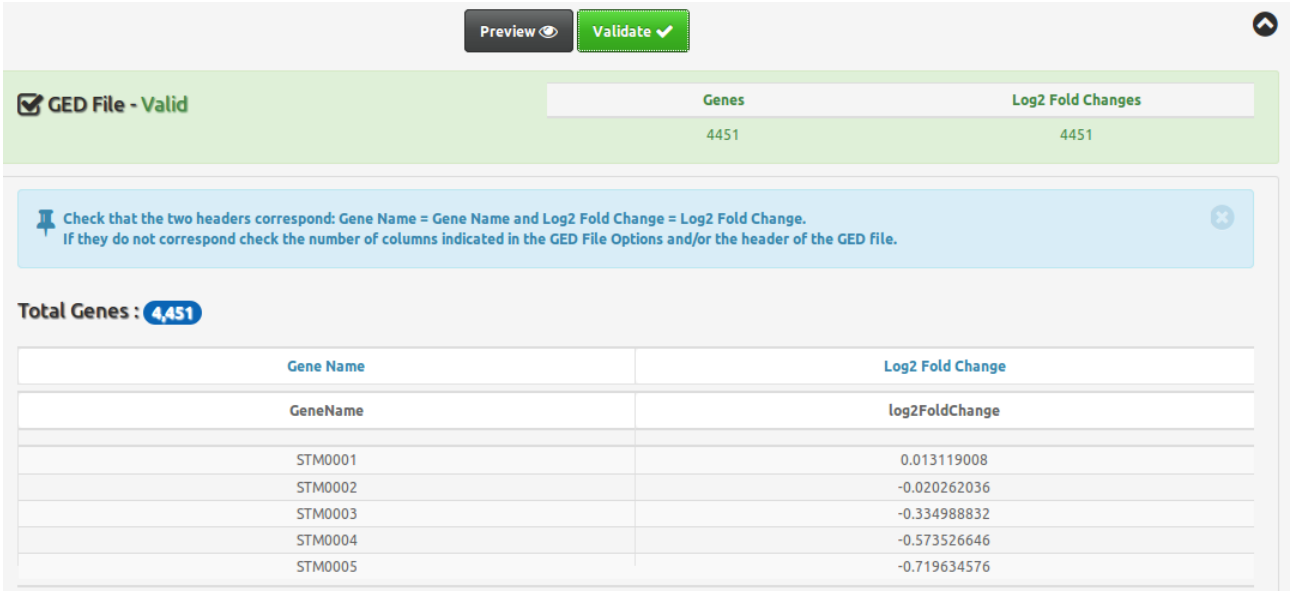
- Click on the "Preview" button to see the selected genome annotation



- Click on the "Validate" button

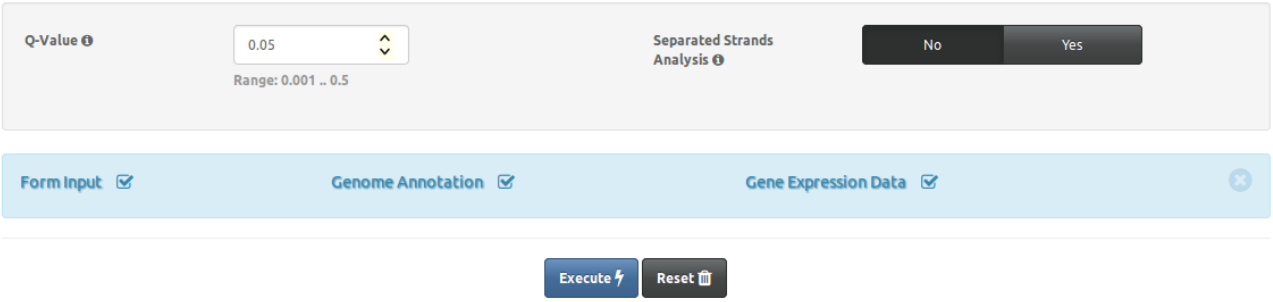
Step 4: Check Gene Expression Data

- Check that the box "GED File" is shaded in green and has a "Valid" flag. The numbers reported under "Genes" and "Log2 Fold Changes" must be the same
- Check that the two headers correspond (i.e.: the right columns were selected as those containing gene name and Log2 fold change information)
- Note that the "Gene Name" field contains the same values as the field of the same name in Genome Annotation section



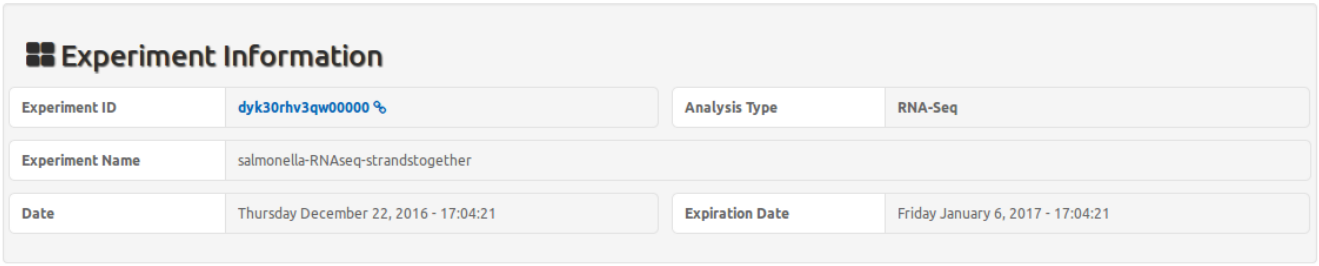
Step 5: Q-value and Separated Strands Analysis

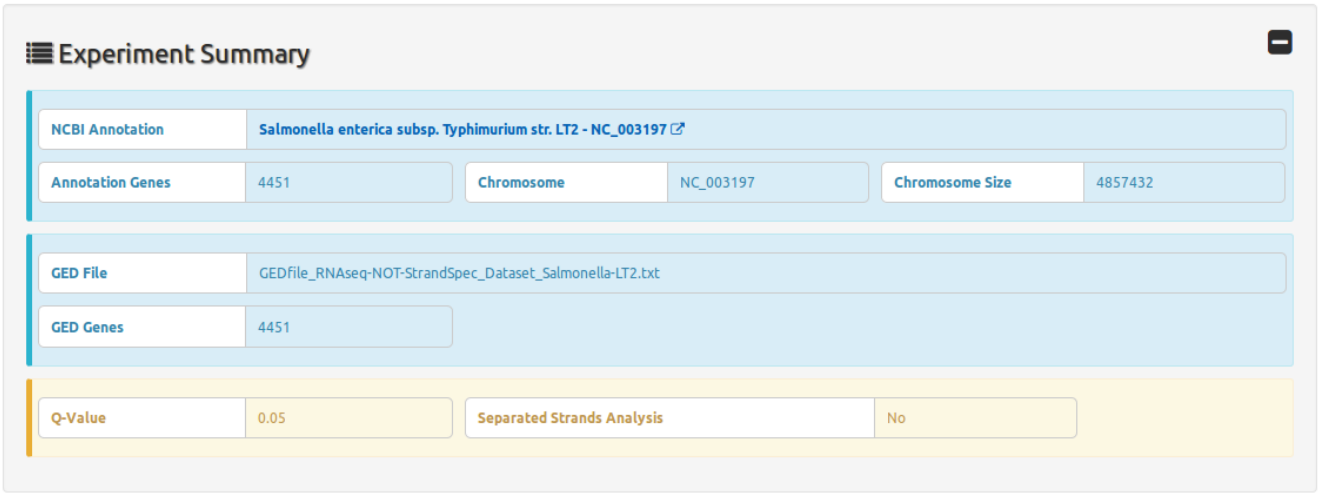
- Enter a q-value of "0.05" (default) or one in the range 0.001-0.5
- Select "No" button for "Separated strands analysis" selector
- Check that all check boxes in the light blue box are ☑
- Click the "Execute" button

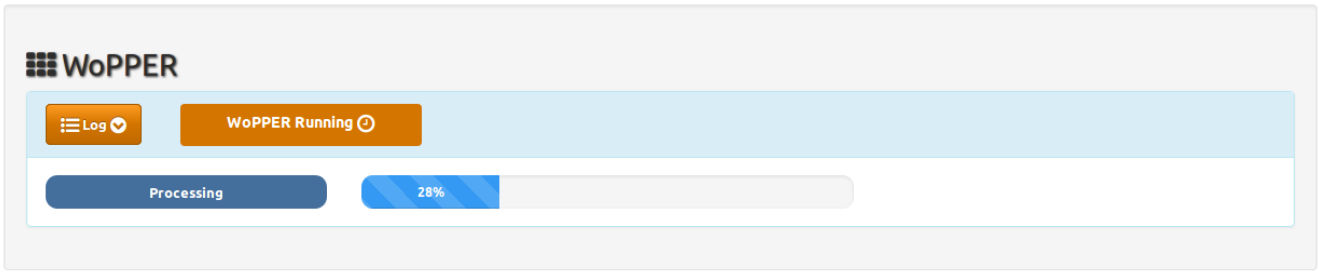


Step 6: Monitor the progress

- "Experiment Information" and "Experiment Summary" contain all the information about the current experiment, including the parameters and the overall number of genes to be tested
- "Experiment ID" contains the unique identifier of the WoPPER job, which can be bookmarked and used for redirection to the results page once WoPPER has finished analyzing the data.
- The progress bar updates regularly as the analysis proceeds


Experiment Information
Experiment ID
dyk30rhv3qw00000
Analysis Type
RNA-Seq
Experiment Name
salmonella-RNAseq-strandstogether
Date
Thursday December 22, 2016 - 17:04:21
Expiration Date
Friday January 6, 2017 - 17:04:21


Experiment Summary
NCBI Annotation
Salmonella enterica subsp. Typhimurium str. LT2 - NC_003197
Annotation Genes
4451
Chromosome
NC_003197
Chromosome Size
4857432
GED File
GEDfile_RNAseq-NOT-StrandSpec_Dataset_Salmonella-LT2.txt
GED Genes
4451
Q-Value
0.05
Separated Strands Analysis
No


WoPPER
Log
WoPPER Running
Processing
28%

Step 7: Check the output

Once WoPPER has completed its execution, you can have a look at the different outputs, which should appear as follows:

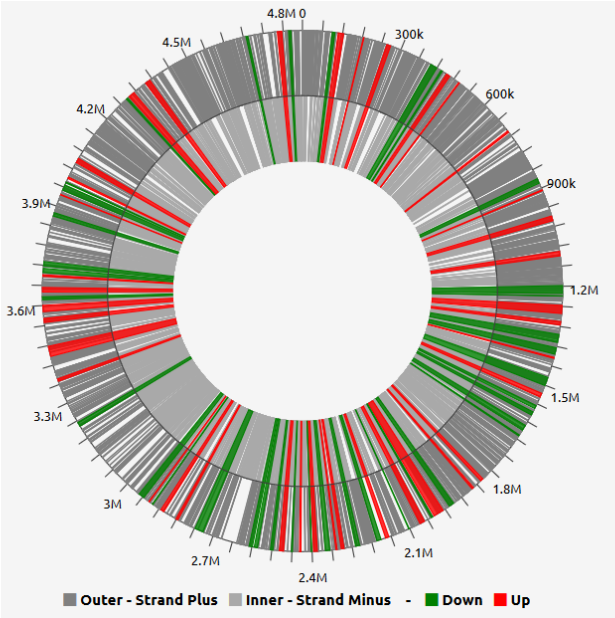
Tabular Output

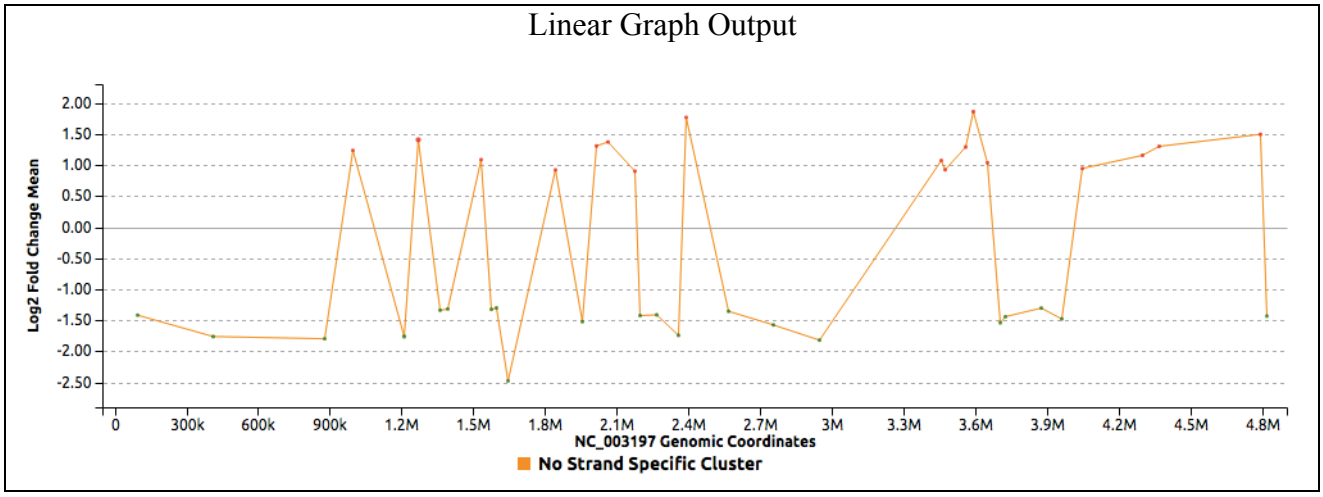
Total Clusters : 75

Clear filtering | Clear sorting | Columns

ID	Cluster Start	Cluster End	Cluster Width	# Genes	Genes in Cluster	Log2 FC Mean	Log2 FC SD	Expression Trend
1	89069	96903	7834	9	STM0075 STM0076 STM0077 STM0078 STM0079 STM0080 STM0081 STM0082 STM0084	-1.9585	1.8024	
2	399704	419571	19867	17	STM0353 STM0354 STM0355 STM0356 STM0357 STM0358.5 STM0359 STM0360 STM0361 STM0362 STM0363 STM0364 STM0365 STM0366 STM0367 STM0368 STM0369	-1.7338	2.1465	
3	434090	438685	4595	6	STM0380 STM0381 STM0382 STM0383 STM0384 STM0385	-2.1094	2.0189	
4	869787	884184	14397	18	STM0801 STM0802 STM0803 STM0804 STM0805 STM0806 STM0807 STM0808 STM0809 STM0810 STM0811 STM0812 STM0813 STM0814 STM0815 STM0816 STM0817 STM0818	-1.8080	1.7046	
5	1197859	1221012	23153	21	STM1110 STM1111 STM1112 STM1113 STM1114 STM1115 STM1116 STM1117 STM1118 STM1119 STM1121 STM1122 STM1123 STM1125 STM05010 STM1126 STM1127 STM1128 STM1129 STM1130 STM1131	-1.5855	2.2906	

Circular Plot Output





Congratulations! You have successfully completed WoPPER Tutorial 2

Conclusion

Once you have completed running WoPPER Tutorial 2, you can have a look also to the pre-computed results page. This page should contain exactly the same results and outputs you have just obtained.

Pre-computed results can be viewed clicking on the corresponding radio-button in “Results” column of the “Tutorials” page

#	Descriptions	Organism	Type	# Chr	Annotation	Separated Strands	GED Files	Results	Examples
2		*Salmonella enterica*	RNA-Seq	1	NCBI	No			