WoPPER Tutorial 4

Escherichia coli

(24/03/2017)

Introduction

This tutorial will show you the analysis of a microarray experiment on *Escherichia coli*, 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.

The file should be named: "**GEDfile_Microarray_Dataset_E.coli-MG1655.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

2. Selecting "Escherichia coli" 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
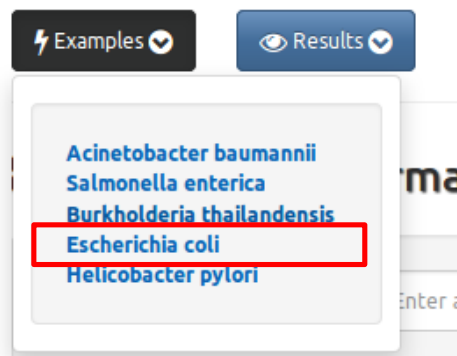- Select Microarray as the Analysis Type

WoPPER :: Input

Experiment Information

Experiment Name	escherichia-microarray
Email Address	mymail@myprovider.org
Analysis Type	RNA-Seq / Microarray

Step 2: Select Genome Annotation

- Select the "NCBI Genome Annotation" tab
- Type in "Escher" 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 "Escherichia coli str. K-12 substr. MG1655"
- Click on the "Preview" button to see the selected genome annotation

Genome Annotation

NCBI Genome Annotation | Custom Genome Annotation

Annotation	Escherichia coli str. K-12 substr. MG1655	✕

Type in Strain name to activate Drop Down menu.

Preview 👁

Check that the Gene Name in the Annotation selected correspond to the Gene Name in your Gene Expression Data File.
If they do not correspond select Custom Genome Annotation and upload your custom annotation file.

Total Genes : 4,140

Chromosome	Start	End	Strand	Gene Name
NC_000913	190	255	+	b0001
NC_000913	337	2799	+	b0002
NC_000913	2801	3733	+	b0003
NC_000913	3734	5020	+	b0004
NC_000913	5234	5530	+	b0005

Step 3: Load Gene Expression Data

- Click on the "Select file" button and load the file named "GEDfile_Microarray_Dataset_E.coli-MG1655.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 "Preview" button to check the GED file content and columns

- 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	fr9qscr90gp00000 ⚲	Analysis Type	Microarray
Experiment Name	escherichia-microarray		
Date	Thursday December 22, 2016 - 17:32:43	Expiration Date	Friday January 6, 2017 - 17:32:43

☰ Experiment Summary　　　　　　　　　　　　　　　　　　　　　　▬

NCBI Annotation	Escherichia coli str. K-12 substr. MG1655 - NC_000913 ↗				
Annotation Genes	4140	Chromosome	NC_000913	Chromosome Size	4641652

GED File	GEDfile_Microarray_Dataset_E.coli-MG1655.txt
GED Genes	4069

Q-Value	0.05	Separated Strands Analysis	No

▦ WoPPER

☰ Status ✅	WoPPER Running ⏱

Processing	14%

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

ID	Cluster Start	Cluster End	Cluster Width	# Genes	Genes in Cluster	Expression Trend	Log2 FC Mean	Log2 FC SD
1	120864	130333	9469	6	b0112 b0113 b0114 b0115 b0116 b0117	↓	-0.3195	0.4063
2	219240	223120	3880	6	b0195 b0196 b0197 b0198 b0199 b0200	↓	-0.3377	0.1723
3	376064	386590	10526	12	b0353 b0354 b0355 b0356 b0357 b0358 b0360 b0361 b0363 b0364 b0365 b0366	↓	-0.2515	0.2232
4	399699	411889	12190	15	b0380 b0381 b0382 b0383 b0384 b0385 b0386 b0387 b0388 b0389 b0390 b0391 b0393 b0394 b0396	↓	-0.3251	0.3787
5	433226	437648	4422	7	b0413 b0414 b0415 b0416 b0417 b0418 b0419	↓	-0.2784	0.4076

Total Clusters : 87

« 1 2 3 4 5 ... 18 »   5 10 25 50 100

Circular Plot Output

Linear Graph Output
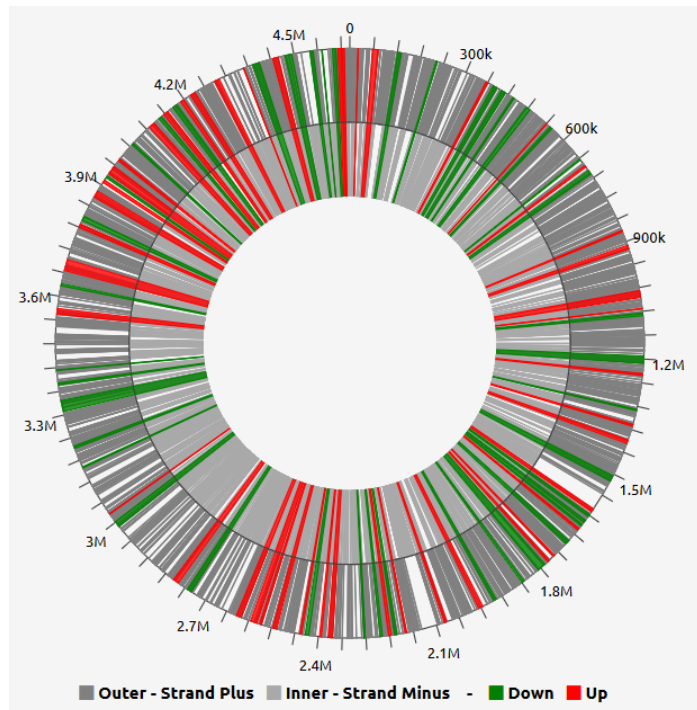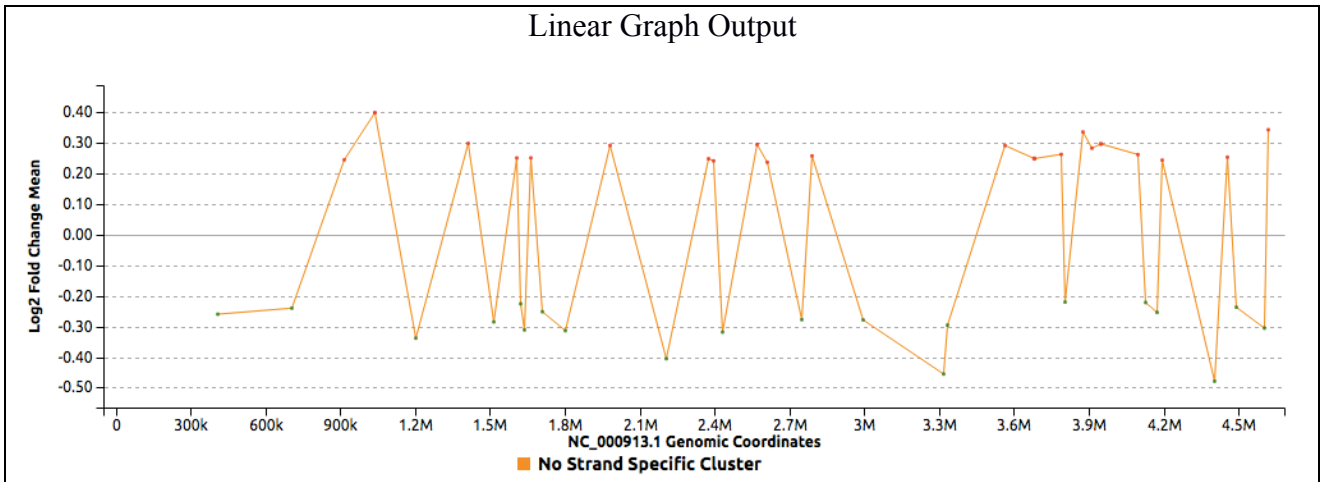
Congratulations! You have successfully completed WoPPER Tutorial 4

Conclusion

Once you have completed running WoPPER Tutorial 4, 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
4	📄	*Escherichia coli*	Microarray	1	NCBI	No	📄	👁	⚡