WoPPER Tutorial 3

Burkholderia thailandensis

(24/03/2017)

Introduction

This tutorial will show you the analysis of an RNA-seq experiment on *Burkholderia thailandensis*, an organism with two chromosomes using the NCBI annotation of the genome, using the separated strands analysis option.

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-StrandSpec_Dataset_Burkholderia-thailandensis-E264.txt**"

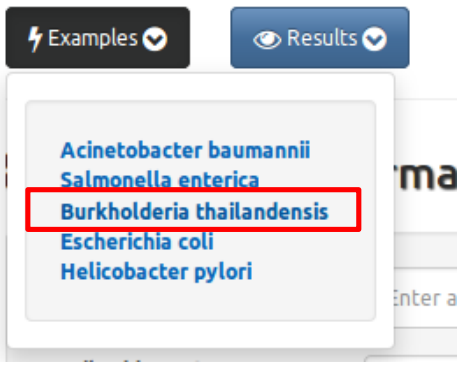
In order to process both chromosomes, you will have to proceed with separate analyses for each of the two chromosomes, choosing the proper reference from the drop down menu.

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
3		*Burkholderia thailandensis*	RNA-Seq	2	NCBI	Yes			

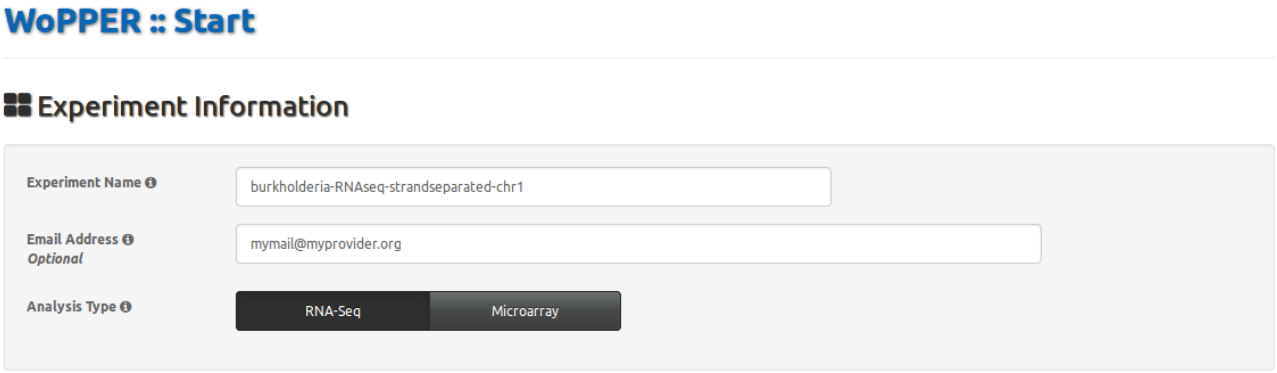
2. Selecting "Burkholderia thailandensis" from the "Examples" drop-down menu available at the top of WoPPER "Start" page



Step-by-step procedure

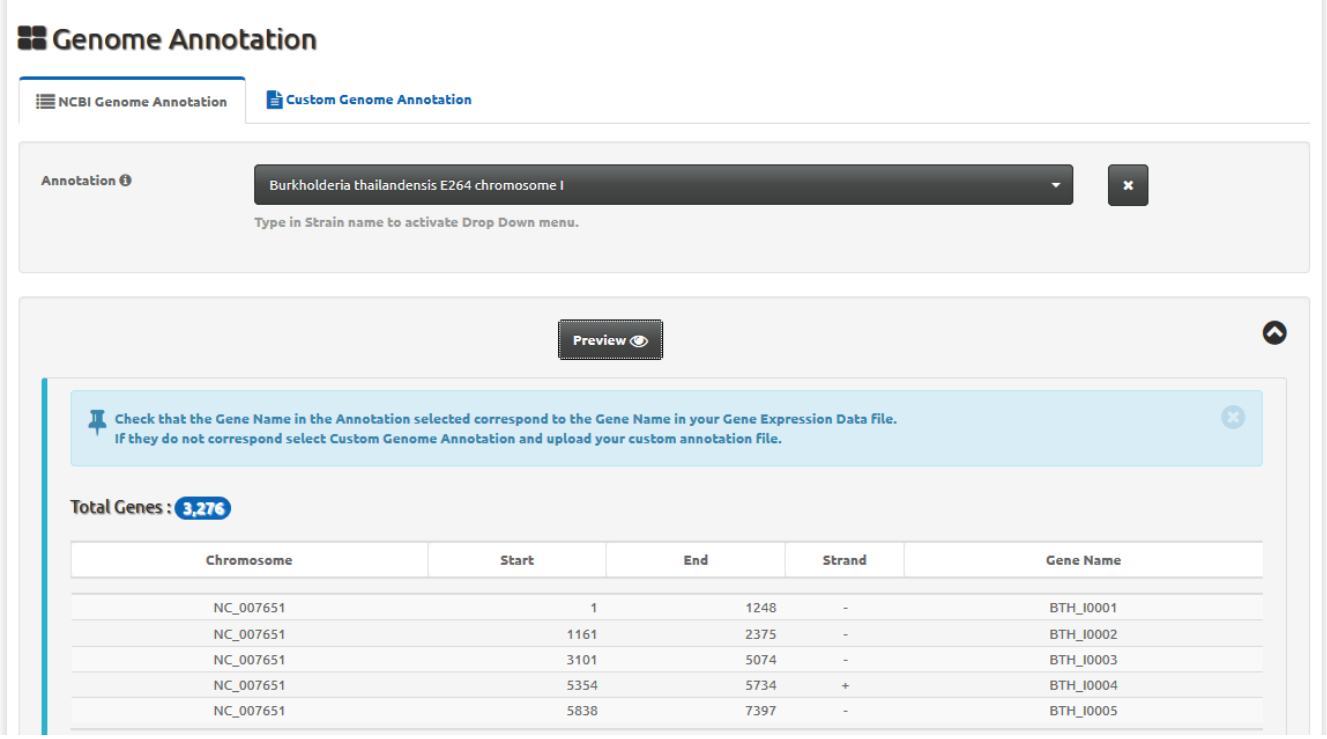
Step 1: Insert Experiment information chromosome I

- Fill in the "Experiment Name" field with a suitable name for your experiment. Since you will have to duplicate the analysis steps for each of the two chromosomes, you should name this experiment accordingly.
- 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



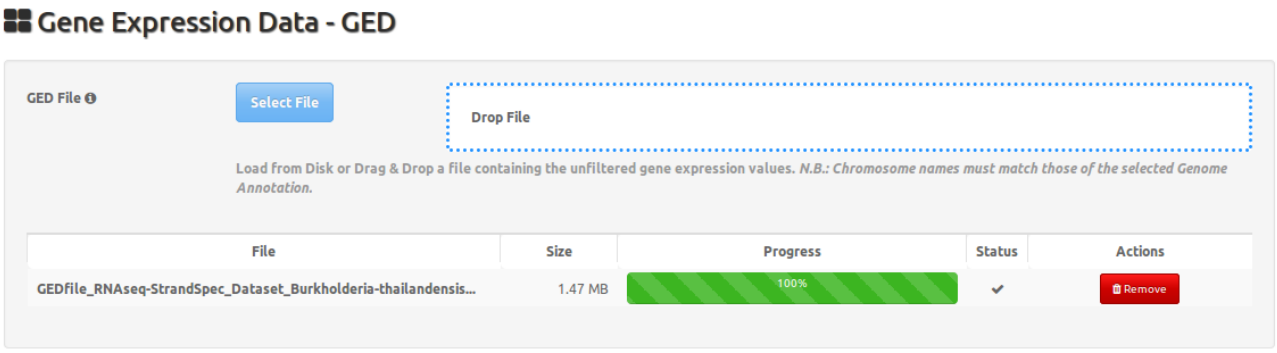
Step 2: Select Genome Annotation chromosome I

- Select the "NCBI Genome Annotation" tab
- Type in "Burkho" in then web form to start the auto-fill based on a search on the internal database of bacterial genomes
- From the drop down menu, select "Burkholderia thailandensis E264 chromosome I" (please note that there is a separate record for chromosome II, which will be employed later)
- Click on the "Preview" button to see the selected genome annotation

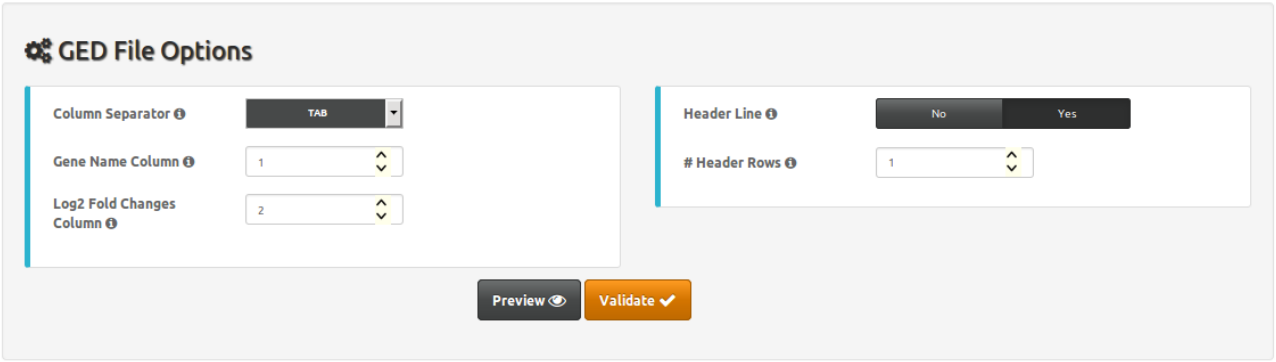


Step 3: Load Gene Expression Data

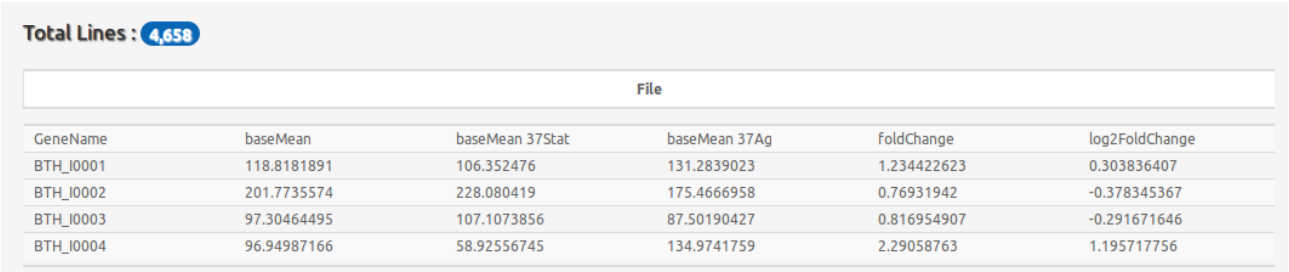
- Click on the "Select file" button and load the file named "GEDfile_RNAseq-StrandSpec_Dataset_Burkholderia-thailandensis-E264.txt" (alternatively, you can drag and drop the same file into the area named "Drop File")
  It is not necessary to have separated files for gene expression data per each chromosome, since WoPPER will correctly analyze only expression values for genes with names matching those reported in the chosen annotation.



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



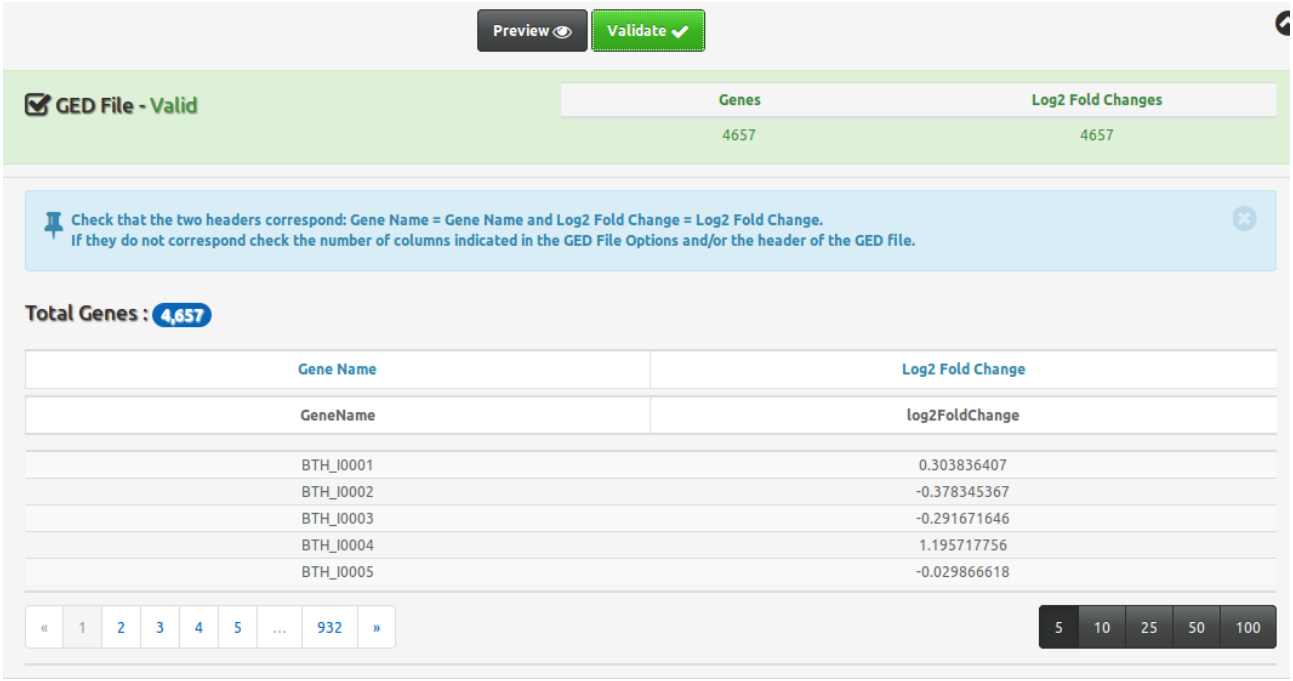
- Click on "Preview" button to check the GED file content and columns



- Click on the "Validate" button

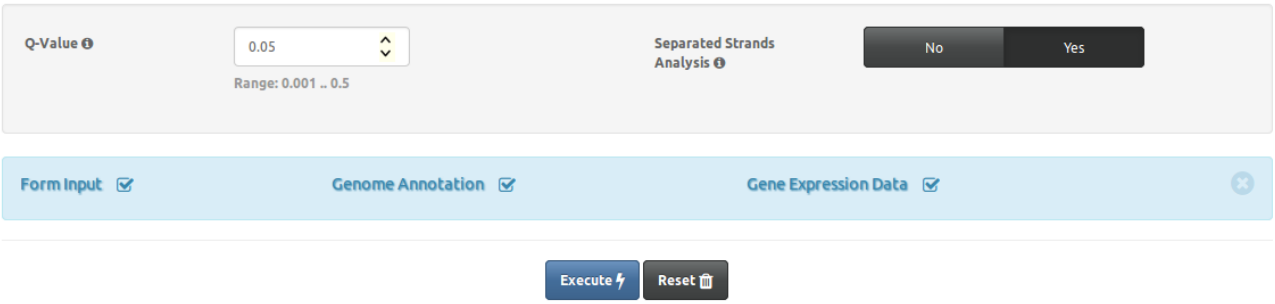
Step 4: Check Gene Expression Data

- 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



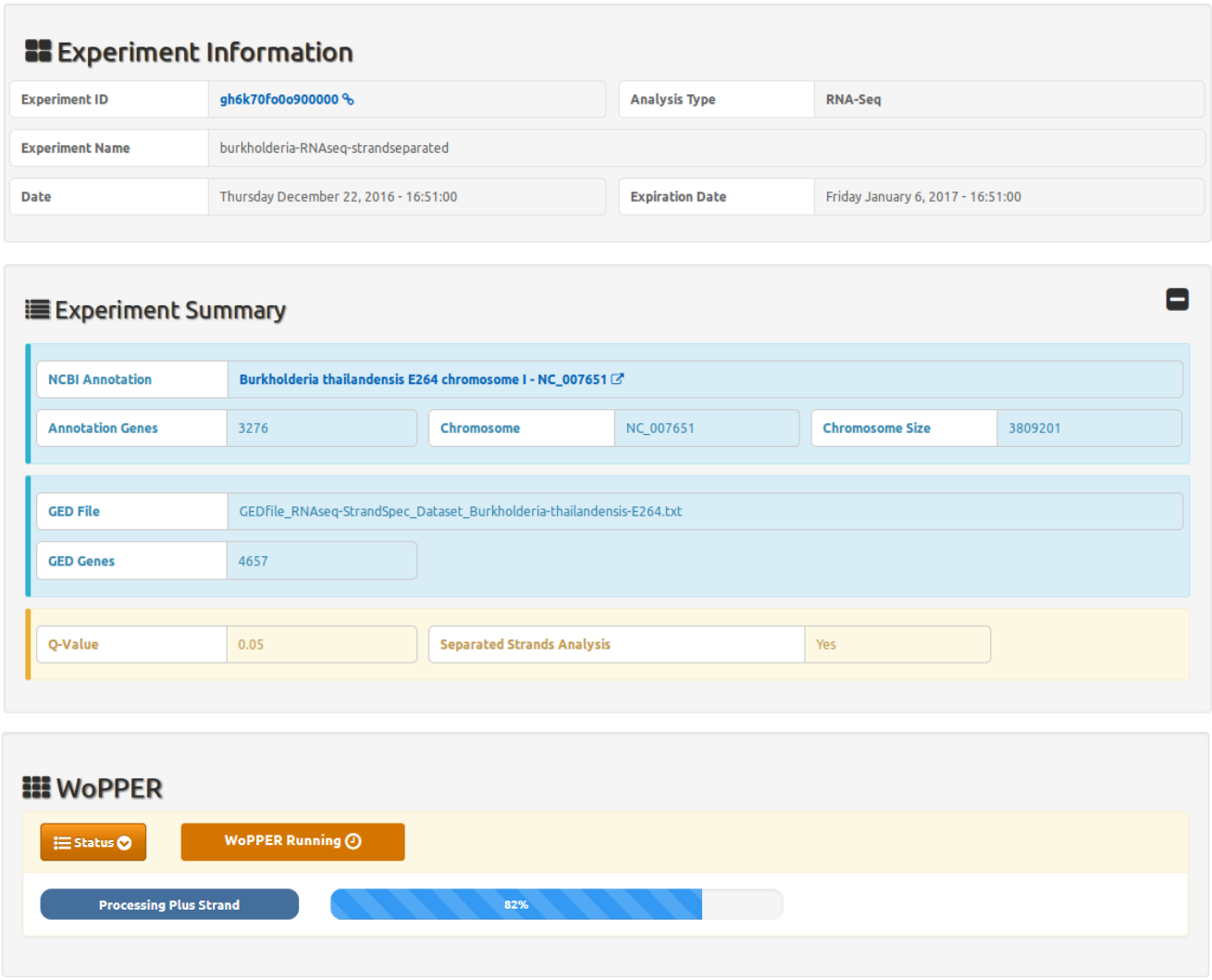
Step 5: Q-value and Separated Strands Analysis

- Enter a q-value of "0.05" (default) or one in the range 0.001-0.5
- Select "Yes" 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



Step 7: Check the output chromosome I

After the analysis has finished, you can download:

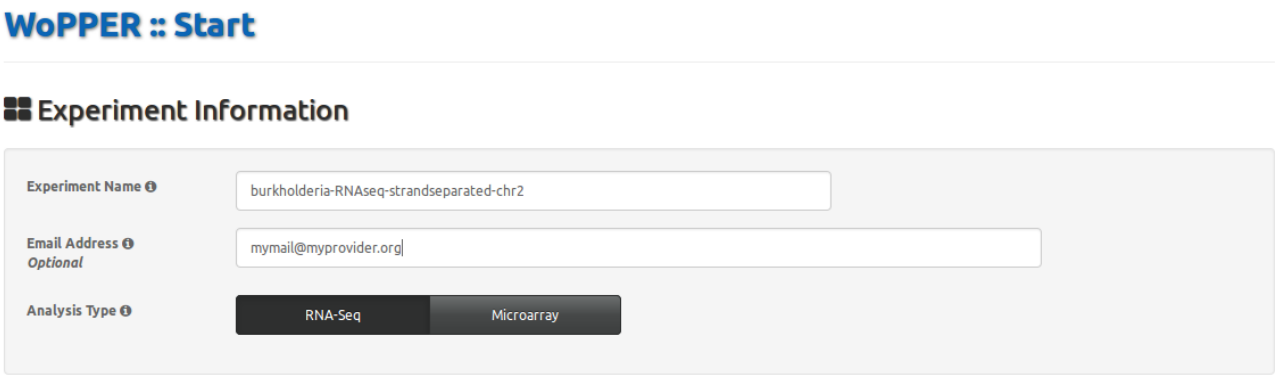
- Tabular output (in txt format)
- Circular Plot Output (in PNG or SVG format)

Then, you can proceed to the WoPPER analysis of chromosome II. Tabular and graphical outputs for both chromosomes will be presented together at the end of the second processing.

Step 8: Insert Experiment information for chromosome II

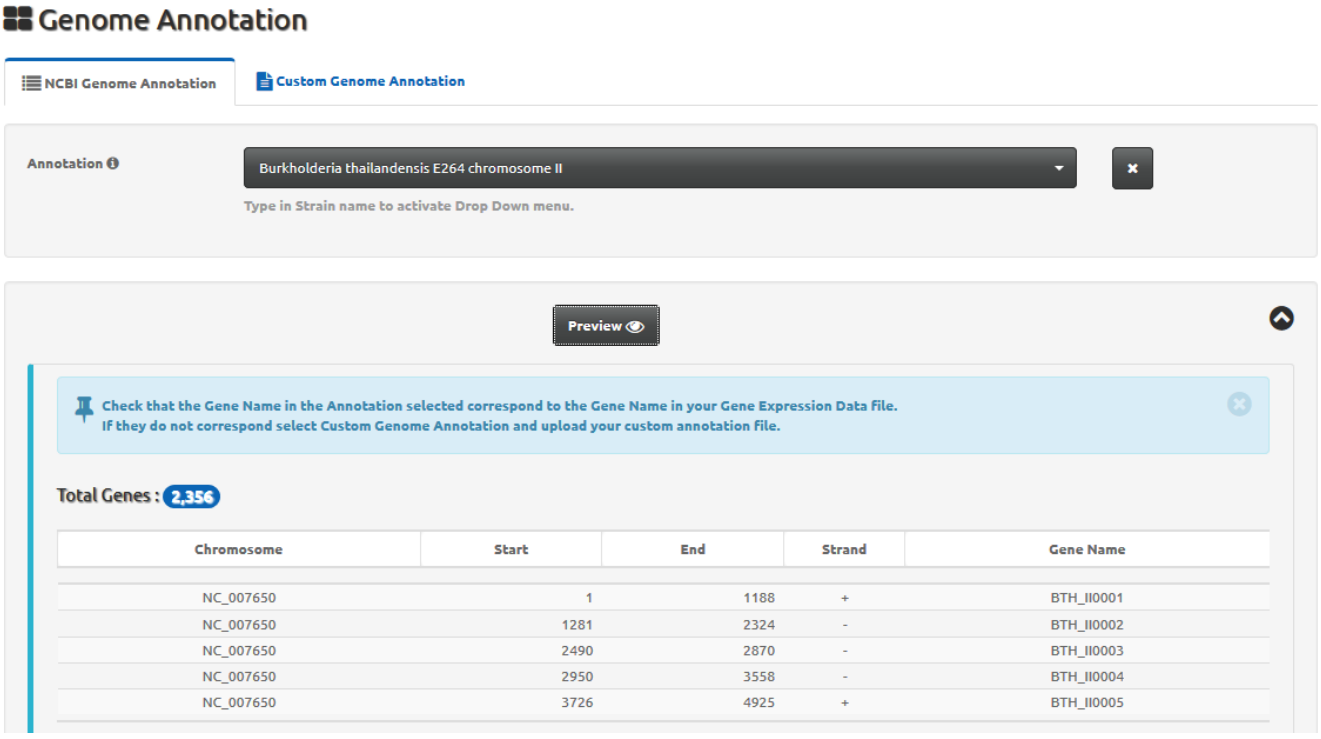
The very same procedure must be repeated for chromosome II, in order to have a complete view of the spatial clustering of up- and down-regulated gene clusters of the experiment

- Fill in “Experiment Name” field with a suitable name for your experiment. Since you will have to duplicate the analysis steps for each of the two chromosomes, you should name this experiment accordingly.
- Fill in “Email address” 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 9: Select Genome Annotation chromosome II

- Select the "NCBI Genome Annotation" tab
- Type in "Burkho" in then web form
- From the drop down menu, select "Burkholderia thailandensis E264 chromosome II" to activate the autofill based on an automatic search of the genomes available in the internal database
- Click on the "Preview" button to see the selected genome annotation



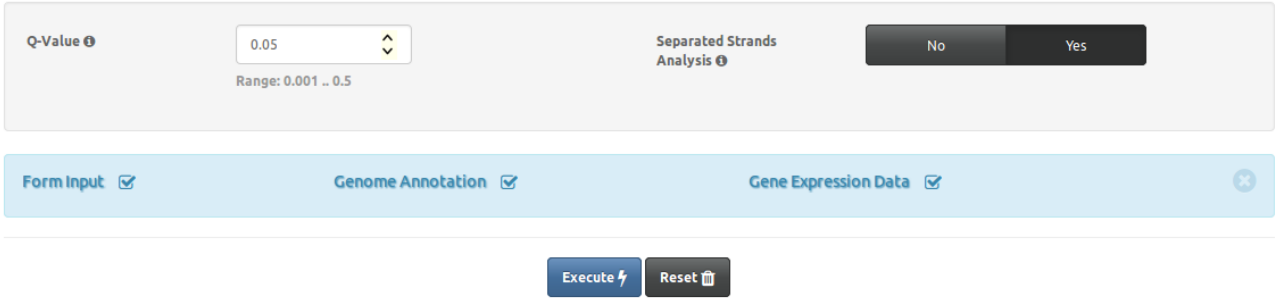
Step 10: Load Gene Expression Data and check Gene Expression Data

- Click on the "Select file" button and load again the file named "GEDfile_RNAseq-StrandSpec_Dataset_Burkholderia-thailandensis-E264.txt" (alternatively, you can drag and drop the same file into the area named "Drop File")

- In the "GED File Options" section, select the same values as done above for chromosome I (which is: TAB-separated file, select column 1 for Gene Name and column 6 for Log2 Fold Change values, row 1 contains the header)
- Click on the "Validate" button

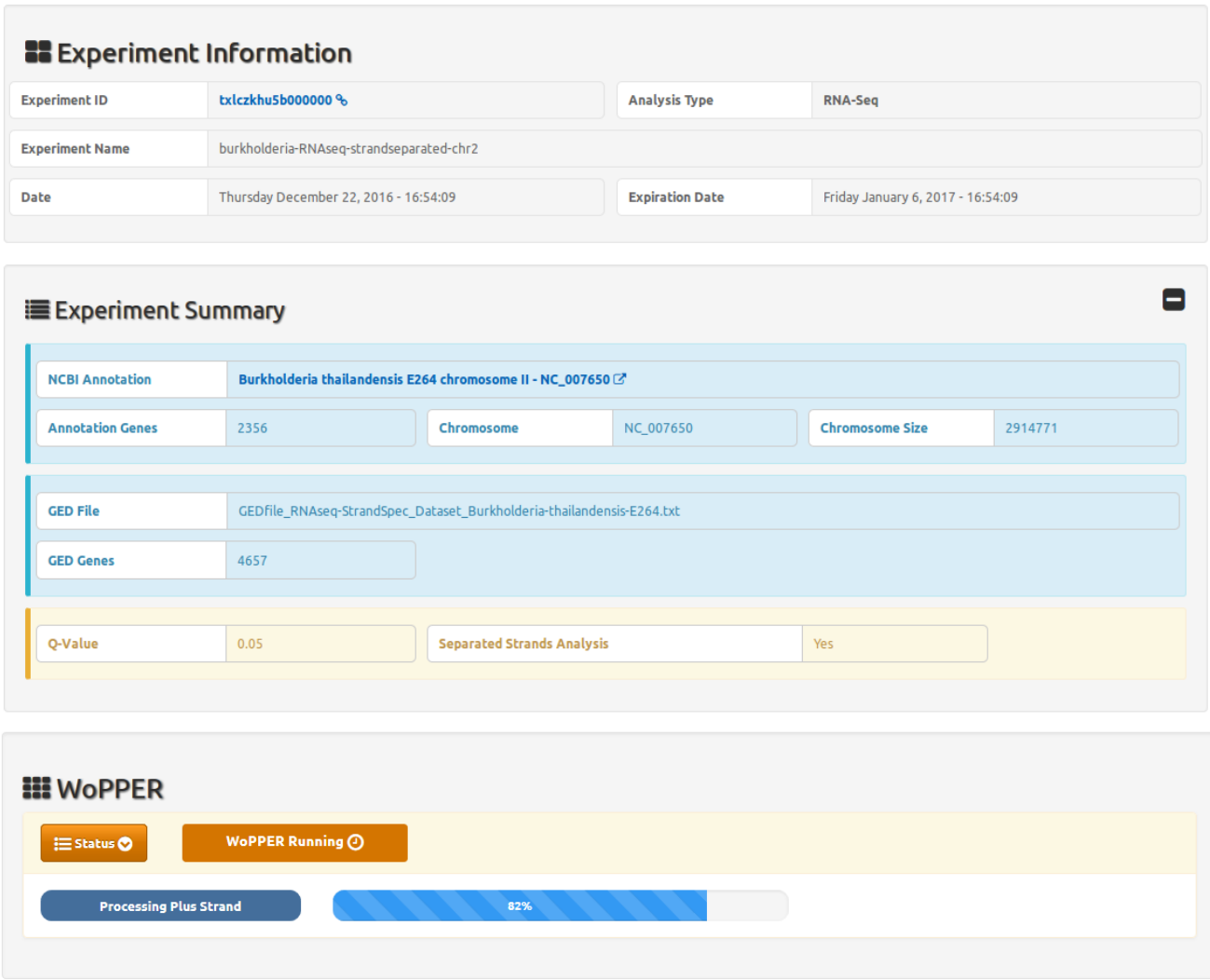
Step 11: Q-value and Separated Strands Analysis

- Enter a q-value of "0.05" (default) or one in the range 0.001-0.5
- Select "Yes" button for "Separated strands analysis" selector
- Check that all check boxes in the light blue box are ☑
- Click the "Execute" button



Step 12: 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



Step 13: Check the output chromosome II

After the analysis has finished, you can download:

- Tabular output (in txt format)
- Circular Plot Output (in PNG or SVG format)

Once you have completed the analysis on both chromosomes, you can see the outputs of the two WoPPER analyses, as follows

Tabular Output chr I

Total Clusters : 27

Clear filtering | Clear sorting | Columns

ID	Cluster Start	Cluster End	Cluster Width	# Genes	Genes in Cluster	Strand	Log2 FC Mean	Log2 FC SD	Expression Trend
1	2341214	2375523	34309	15	BTH_I2067 BTH_I2068 BTH_I2069 BTH_I2070 BTH_I2073 BTH_I2076 BTH_I2078 BTH_I2082 BTH_I2083 BTH_I2087 BTH_I2089 BTH_I2090 BTH_I2095 BTH_I2096 BTH_I2097	+	-0.5782	0.4504	↓
2	2592023	2733606	141583	54	BTH_I2299 BTH_I2300 BTH_I2301 BTH_I2303 BTH_I2304 BTH_I2305 BTH_I2306 BTH_I2308 BTH_I2311 BTH_I2313 BTH_I2326 BTH_I2327 BTH_I2328 BTH_I2329 BTH_I2335 BTH_I2338 BTH_I2339 BTH_I2340 BTH_I2341 BTH_I2342 BTH_I2343 BTH_I2344 BTH_I2345 BTH_I2346 BTH_I2348 BTH_I2351 BTH_I2352 BTH_I2353 BTH_I2354 BTH_I2355 BTH_I2356 BTH_I2369 BTH_I2371 BTH_I2372 BTH_I2376 BTH_I2378 BTH_I2379 BTH_I2380 BTH_I2381 BTH_I2382 BTH_I2383 BTH_I2384 BTH_I2385 BTH_I2386 BTH_I2389 BTH_I2390 BTH_I2392 BTH_I2393 BTH_I2398 BTH_I2399 BTH_I2400 BTH_I2401 BTH_I2402 BTH_I2403	+	-0.5749	0.6224	↓
3	5544	12688	7144	4	BTH_I0004 BTH_I0007 BTH_I0008 BTH_I0009	+	0.7405	0.5059	↑
4	1189948	1228555	38607	32	BTH_I1048 BTH_I1049 BTH_I1050 BTH_I1054 BTH_I1055 BTH_I1056 BTH_I1057 BTH_I1058 BTH_I1059 BTH_I1061 BTH_I1062 BTH_I1063 BTH_I1064 BTH_I1065 BTH_I1066 BTH_I1067 BTH_I1068 BTH_I1069 BTH_I1070 BTH_I1071 BTH_I1072 BTH_I1073 BTH_I1074 BTH_I1075 BTH_I1076 BTH_I1077 BTH_I1078 BTH_I1079 BTH_I1080 BTH_I1081 BTH_I1082 BTH_I1087	+	0.8581	0.7363	↑
5	1286801	1299873	13072	10	BTH_I1140 BTH_I1141 BTH_I1142 BTH_I1143 BTH_I1146 BTH_I1147 BTH_I1150 BTH_I1151 BTH_I1152 BTH_I1153	+	0.7451	1.0219	↑

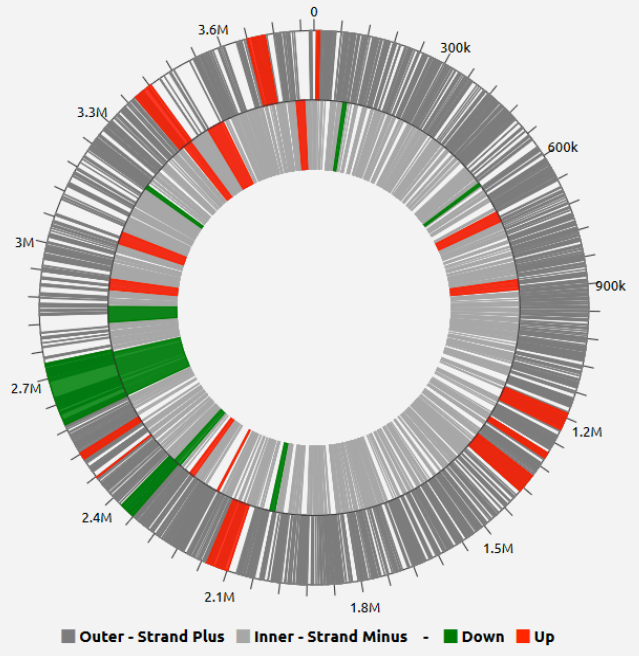
Tabular Output chr II

Total Clusters : 21

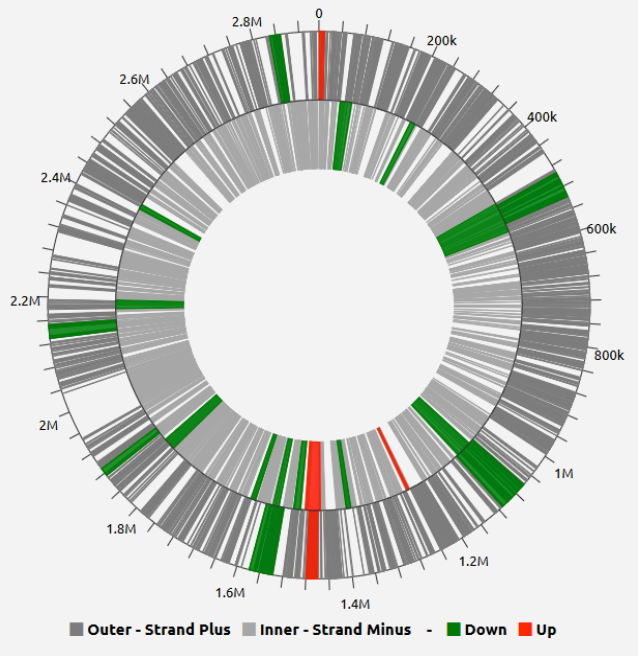
Clear filtering | Clear sorting | Columns

ID	Cluster Start	Cluster End	Cluster Width	# Genes	Genes in Cluster	Strand	Expression Trend	Log2 FC Mean	Log2 FC SD
1	493977	539245	45268	29	BTH_II0409 BTH_II0410 BTH_II0412 BTH_II0413 BTH_II0414 BTH_II0415 BTH_II0419 BTH_II0420 BTH_II0421 BTH_II0422 BTH_II0423 BTH_II0424 BTH_II0425 BTH_II0426 BTH_II0427 BTH_II0428 BTH_II0429 BTH_II0434 BTH_II0435 BTH_II0436 BTH_II0441 BTH_II0442 BTH_II0443 BTH_II0444 BTH_II0445 BTH_II0446 BTH_II0447 BTH_II0449 BTH_II0450	+	↓	-0.6114	0.4266
2	1059845	1112326	52481	25	BTH_II0904 BTH_II0906 BTH_II0907 BTH_II0908 BTH_II0909 BTH_II0910 BTH_II0914 BTH_II0915 BTH_II0917 BTH_II0920 BTH_II0921 BTH_II0922 BTH_II0923 BTH_II0927 BTH_II0928 BTH_II0929 BTH_II0930 BTH_II0931 BTH_II0933 BTH_II0934 BTH_II0935 BTH_II0936 BTH_II0937 BTH_II0940 BTH_II0941	+	↓	-0.7508	0.8353
3	1537215	1578262	41047	26	BTH_II1289 BTH_II1291 BTH_II1292 BTH_II1293 BTH_II1294 BTH_II1295 BTH_II1296 BTH_II1297 BTH_II1298 BTH_II1301 BTH_II1302 BTH_II1303 BTH_II1304 BTH_II1305 BTH_II1306 BTH_II1308 BTH_II1309 BTH_II1310 BTH_II1313 BTH_II1314 BTH_II1315 BTH_II1316 BTH_II1317 BTH_II1318 BTH_II1323 BTH_II1324	+	↓	-0.8565	0.9168
4	1874674	1889206	14532	4	BTH_II1595 BTH_II1596 BTH_II1600 BTH_II1611	+	↓	-0.9814	0.5202
5	2129487	2153045	23558	8	BTH_II1761 BTH_II1762 BTH_II1765 BTH_II1766 BTH_II1776 BTH_II1778 BTH_II1779 BTH_II1781	+	↓	-1.0590	0.8006

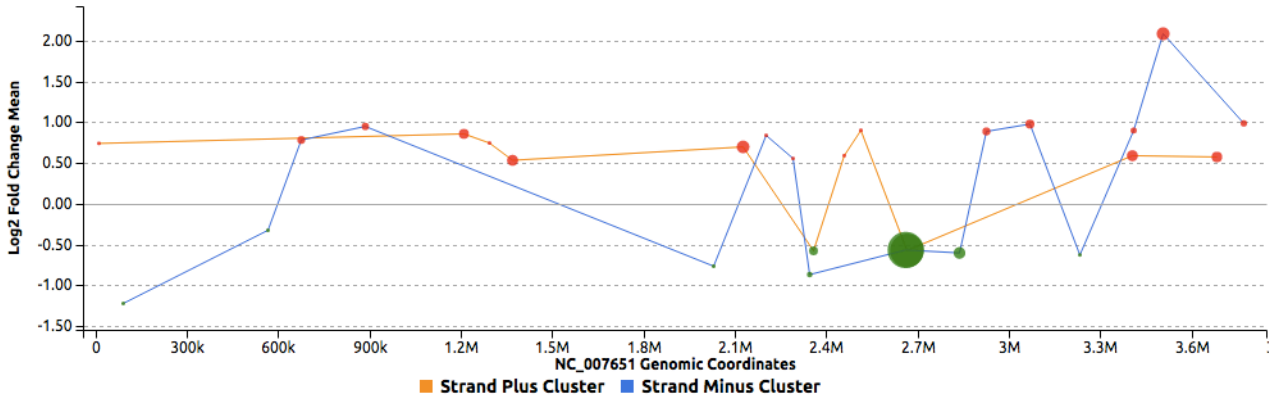
Circular Plot Output chr I



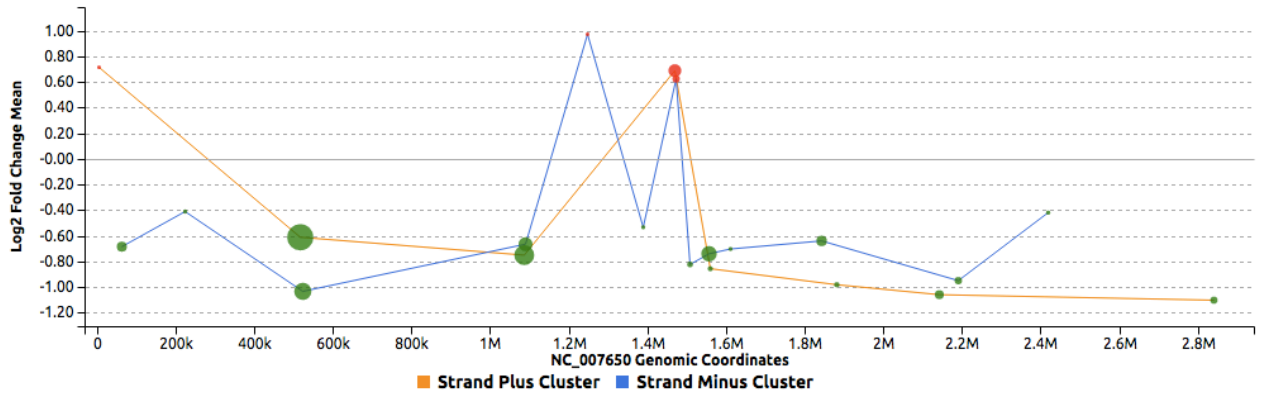
Circular Plot Output chr II



Linear Graph Output chr I



Linear Graph Output chr II



Congratulations! You have successfully completed WoPPER Tutorial 3

Conclusion

Once you have completed running WoPPER Tutorial 3, 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
3		*Burkholderia thailandensis*	RNA-Seq	2	NCBI	Yes			