

WOPPER

Tutorial 5

Helicobacter pylori












(24/03/2017)

Introduction

This tutorial will show you the analysis of a strand-specific RNA-seq experiment on *Helicobacter pylori*, an organism with one chromosome, using a custom annotation file, using the separated strand analysis option.

For performing this tutorial, you will need to download the GED and the custom annotation file for this specific experiment from the “**Tutorials**” section of WoPPER.

Tutorials






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

The two files should be named:

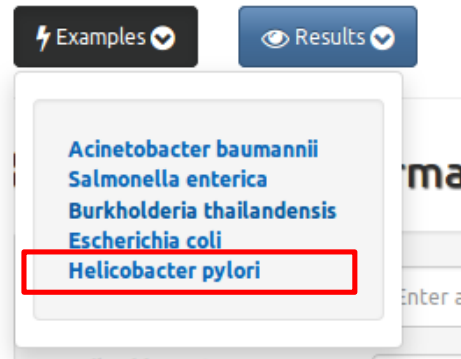
- “GEDfile_RNAseq_StrandSpec_Dataset_Acinetobacter-baumannii-ATCC17978.txt”
- “HP_G27_GCF_000021165.1_ASM2116v1_genomic.gff”

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
5		<i>Helicobacter pylori</i>	RNA-Seq	1	Custom 	Yes			

2. Selecting “*Helicobacter pylori*” 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

WoPPER :: Start

Experiment Information

Experiment Name ⓘ	<input type="text" value="helicobacter-RNAseq-strandseparated"/>
Email Address ⓘ <i>Optional</i>	<input type="text" value="mymail@myprovider.org"/>
Analysis Type ⓘ	<input checked="" type="radio"/> RNA-Seq <input type="radio"/> Microarray

Step 2: Select Genome Annotation

- Select the “Custom Genome Annotation” tab
- Click on “Select file” button and load file named “HP_G27_GCF_000021165.1_ASM2116v1_genomic.gff” (alternatively, you can drag’n’drop the same file into the area named “Drop File”)
- Set the “Annotation Name” and “Chromosome size” as indicated in screenshot below
- In “Annotation File Options”, set:
 - “File format” to GFF
 - “Column Separator” to “TAB”
 - “Header Line” to “Yes”
 - “#Header Rows” to 7
- Click on “Validate” button to see the selected genome annotation
- In the green box, select Chromosome NC_011333.1, corresponding to the main genome (the other “chromosome” is a plasmid)

NCBI Genome Annotation | Custom Genome Annotation

Annotation File Drop File

Load from Disk or Drag & Drop a custom Annotation File. *N.B.: Chromosome names must match those of Gene Expression Data File.*

File	Size	Progress	Status	Actions
HP_G27_GCF_000021165.1_ASM2116v1_genomic.gff	0.60 MB	<div style="width: 100%; background-color: green;">100%</div>	✓	<input type="button" value="Cancel"/> <input type="button" value="Remove"/>

Annotation Name Chromosome Size

Annotation File Options

File Format

Column Separator

Header Line

Header Rows

✓

Annotation File - Valid

Active	Chromosome	Genes
<input checked="" type="radio"/>	NC_011333.1	1544
<input type="radio"/>	NC_011334.1	10

✕


Total Genes : 1,554

Chromosome	Start	End	Strand	Gene Name
NC_011333.1	1	417	-	PG27_RS00005
NC_011333.1	419	889	-	PG27_RS00010
NC_011333.1	899	1729	-	PG27_RS00015
NC_011333.1	1716	2381	-	PG27_RS00020
NC_011333.1	2503	3186	+	PG27_RS00025

Step 3: Load Gene Expression Data

- Click on the “Select file” button and load the file named “GEDfile_RNAseq-StrandSpec_Dataset_Helicobacter_Pylori_G27.txt” (alternatively, you can drag and drop the same file into the area named “Drop File”)

Gene Expression Data - GED


GED File  Select File Drop File


Load from Disk or Drag & Drop a file containing the unfiltered gene expression values. *N.B.: Chromosome names must match those of the selected Genome Annotation.*


File	Size	Progress	Status	Actions
GEDfile_RNAseq-StrandSpec_Dataset_Helicobacter_Pylori_G27.tx...	0.39 MB	<div style="width: 100%; background-color: green;">100%</div>	✓	Remove


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


GED File Options

Column Separator  TAB

Gene Name Column  1

Log2 Fold Changes Column  6

Header Line  No Yes

Header Rows  1

Preview Validate

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

Total Lines : **1,557**

File					
GeneName	baseMean	WT_dipy_mean	WT_fe_mean	foldChange	log2FoldChange
HPG27_RS00005	371.918054286308	395.615682898184	348.220425674432	0.880198740159779	-0.190023320069296
HPG27_RS00010	1202.39368816412	1287.10620427781	1117.68117205043	0.868367480737581	-0.1981593966117
HPG27_RS00015	2582.97565284308	2886.39384649423	2279.55745919192	0.789759672596531	-0.336921884623864
HPG27_RS00020	686.774128877469	728.342995914613	645.205261840326	0.885853595708864	-0.176733138440953

« 1 2 3 4 5 ... 312 » 5 10 25 50 100

- 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

Preview Validate

GED File - Valid

Genes	Log2 Fold Changes
1556	1556

Check that the two headers correspond: Gene Name = Gene Name and Log2 Fold Change = Log2 Fold Change.
If they do not correspond check the number of columns indicated in the GED File Options and/or the header of the GED file.

Total Genes : **1,556**

Gene Name	Log2 Fold Change
GeneName	log2FoldChange
HPG27_RS00005	-0.190023320069296
HPG27_RS00010	-0.1981593966117
HPG27_RS00015	-0.336921884623864
HPG27_RS00020	-0.176733138440953
HPG27_RS00025	0.0374892394951457

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

Q-Value
Range: 0.001 .. 0.5

Separated Strands Analysis

Form Input Genome Annotation Gene Expression Data

"/> "/>

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	rhfnngtye000000	Analysis Type	RNA-Seq
Experiment Name	helicobacter-RNAseq-strandseparated		
Date	Thursday December 22, 2016 - 18:07:25	Expiration Date	Friday January 6, 2017 - 18:07:25

Experiment Summary

Annotation File	HP_G27_GCF_000021165.1_ASM2116v1_genomic.gff		
Annotation Genes	1544	Chromosome	NC_011333.1
Chromosome Size	1652858		
GED File	GEDfile_RNAseq-StrandSpec_Dataset_Helicobacter_Pylori_G27.txt		
GED Genes	1556		
Q-Value	0.05	Separated Strands Analysis	Yes

WoPPER

Status	WoPPER Running
Processing	14%

Step 7: Check the output

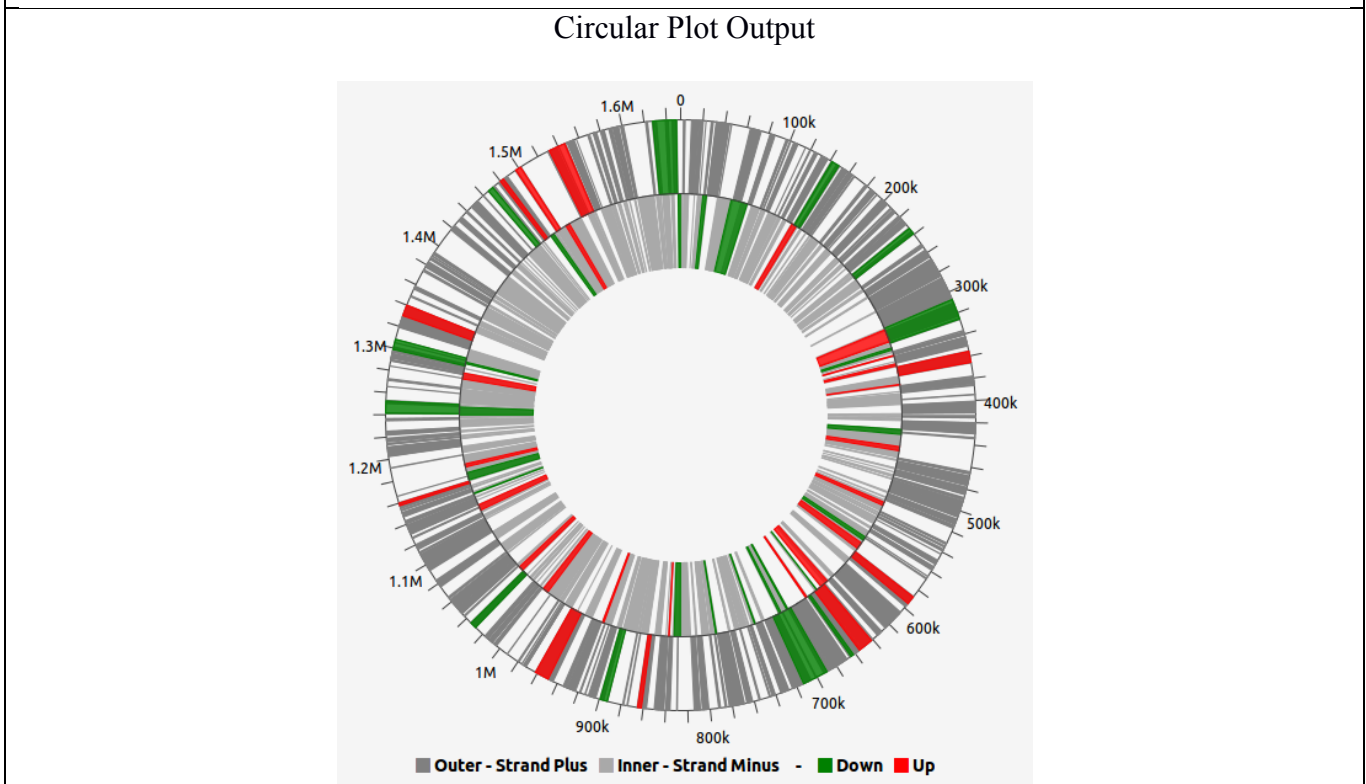
Once WoPPER has completed its execution, you can have a look at the different outputs, which should appear as follows:

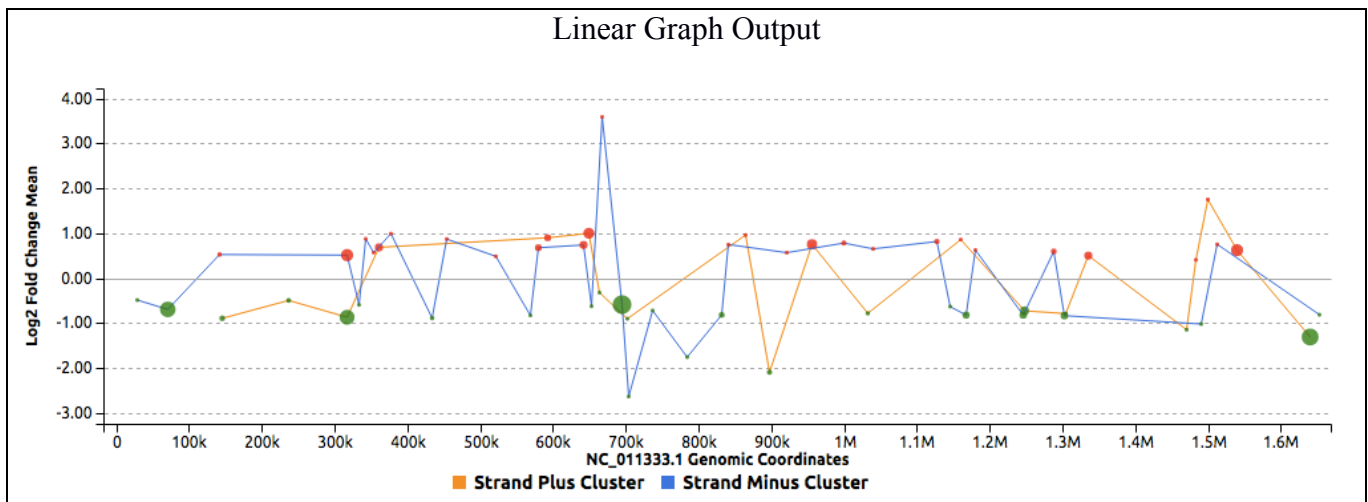
Tabular Output

Total Clusters : 56 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	26584	30779	4195	3	HPG27_RS00145 HPG27_RS00160 HPG27_RS00165	-	-0.4911	0.6523	↓
2	60714	80384	19670	11	HPG27_RS00300 HPG27_RS00355 HPG27_RS00360 HPG27_RS00365 HPG27_RS00370 HPG27_RS00375 HPG27_RS00380 HPG27_RS00385 HPG27_RS00390 HPG27_RS00395 HPG27_RS00400	-	-0.6993	0.6079	↓
3	138789	144713	5924	5	HPG27_RS00670 HPG27_RS00675 HPG27_RS00690 HPG27_RS00695 HPG27_RS00700	-	0.5241	0.2803	↑
4	141569	148950	7381	4	HPG27_RS00680 HPG27_RS00685 HPG27_RS00710 HPG27_RS00715	+	-0.8986	0.7947	↓
5	233009	239233	6224	5	HPG27_RS01140 HPG27_RS01160 HPG27_RS01165 HPG27_RS01170 HPG27_RS01175	+	-0.4996	0.1997	↓

« 1 2 3 4 5 ... 12 » 5 10 25 50 100





Congratulations! You have successfully completed WoPPER Tutorial 5

Conclusion

Once you have completed running WoPPER Tutorial 5, 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
5		<i>Helicobacter pylori</i>	RNA-Seq	1	Custom	Yes		<input checked="" type="radio"/>	