

omics.data.edu.au

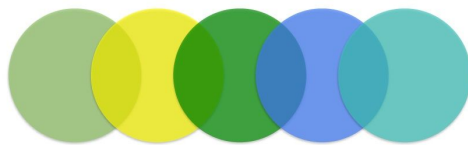
How to exploring differences in gene expression levels between two bacterial strains

STEP	Task				
1	Go to http://omics.data.edu.au/use/				
2	Launch the platform and log in				
3	<i>Add the "ABPRI-Data" and "Degust" apps to the toolbar.</i>				
5	Open the "ABPRI-Data" app.				
6	Log-into the "ABPRI-Data" app using the username and password given to you (you will need to enter omics as the domain).				
Steps 7-10 are to locate a file containing differential gene expression level data for two strains of <i>Klebsiella pneumoniae</i> (AJ218 and KPC2 (6 replicates of each)). This file has been generated by the ABRPI consortium.					
7	Search for "processed" "transcriptomic" data from " <i>Klebsiella pneumoniae</i> ".				
8	Send the resultant data file to the GenomeSpace part of the OMICs platform.				
9	Locate the data you just sent to GenomeSpace.				
10	Download the file onto your desktop.				
The next steps are to import the file just downloaded into Degust, a differential gene expression exploration, analysis and visualisation tool.					
11	Launch Degust by clicking on the Degust app you added to the GenomeSpace toolbar in step 4. Degust is a public server and log-in is not necessary.				
12	Upload the counts file to the Degust server that you have previously downloaded.				
The counts file has 19 columns. Bowtie2 and featureCounts were used respectively by the ABPRI bioinformaticians for mapping and generating the quantification information in this file.					
<table border="1"> <thead> <tr> <th>Column Header</th> <th>Description</th> </tr> </thead> <tbody> <tr> <td><i>GeneID</i></td> <td>A project-specific gene identifier used in the ABPRI project (and common across the genomics and transcriptomics analysis)</td> </tr> </tbody> </table>		Column Header	Description	<i>GeneID</i>	A project-specific gene identifier used in the ABPRI project (and common across the genomics and transcriptomics analysis)
Column Header	Description				
<i>GeneID</i>	A project-specific gene identifier used in the ABPRI project (and common across the genomics and transcriptomics analysis)				



omics.data.edu.au

<i>Chr</i>	chromosome/plasmid identifier in the species (<i>Klebsiella pneumoniae</i>)
<i>Start</i>	Gene start position on the chromosome sequence
<i>End</i>	Gene end position on the chromosome sequence
<i>Strand</i>	Strand (+ or -)
<i>Length</i>	Gene length
25861_PE_230bp_SEP_AGRF_CA3FUANXX_CTGAAGCT-CAGGACGT	Counts of gene expression in BPA sample 25861 (<i>K. pneumoniae</i> , strain AJ218)
g25862_PE_230bp_SEP_AGRF_CA3FUANXX_TAATGCGC-CAGGACGT	Counts of gene expression in BPA sample 25862 (<i>K. pneumoniae</i> , strain AJ218)
25863_PE_230bp_SEP_AGRF_CA3FUANXX_CGGCTATG-CAGGACGT	Counts of gene expression in BPA sample 25863 (<i>K. pneumoniae</i> , strain AJ218)
25864_PE_230bp_SEP_AGRF_CA3FUANXX_TCCGCGAA-CAGGACGT	Counts of gene expression in BPA sample 25864 (<i>K. pneumoniae</i> , strain AJ218)
25865_PE_230bp_SEP_AGRF_CA3FUANXX_TCTCGCGC-CAGGACGT	Counts of gene expression in BPA sample 25865 (<i>K. pneumoniae</i> , strain AJ218)
25866_PE_230bp_SEP_AGRF_CA3FUANXX_AGCGATAG-CAGGACGT	Counts of gene expression in BPA sample 25866 (<i>K. pneumoniae</i> , strain AJ218)
25867_PE_230bp_SEP_AGRF_CA3FUANXX_ATTACTCG-CAGGACGT	Counts of gene expression in BPA sample 25867 (<i>K. pneumoniae</i> , strain KPC2)
25868_PE_230bp_SEP_AGRF_CA3FUANXX_TCCGGAGA-CAGGACGT	Counts of gene expression in BPA sample 25868 (<i>K. pneumoniae</i> , strain KPC2)
25869_PE_230bp_SEP_AGRF_CA3FUANXX_CGCTCATT-CAGGACGT	Counts of gene expression in BPA sample 25869 (<i>K. pneumoniae</i> , strain KPC2)
25870_PE_230bp_SEP_AGRF_CA3FUANXX_GAGATTCC-CAGGACGT	Counts of gene expression in BPA sample 25870 (<i>K. pneumoniae</i> , strain KPC2)
25871_PE_230bp_SEP_AGRF_CA3FUANXX_ATTCAGAA-CAGGACGT	Counts of gene expression in BPA sample 25871 (<i>K. pneumoniae</i> , strain KPC2)
25872_PE_230bp_SEP_AGRF_CA3FUANXX_GAATTCGT-CAGGACGT	Counts of gene expression in BPA sample 25872 (<i>K. pneumoniae</i> , strain KPC2)
<i>Annotation</i>	Gene symbol, EC_number (when known)
13	<p>Configure Degust as follows:</p> <p>Name: Enter a name for your analysis</p> <p>Format Type: TSV</p>

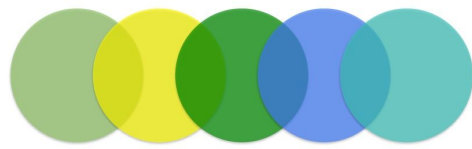


omics.data.edu.au

	<p>Info Columns: select GeneID and Annotation EC Number Column: Leave blank Gene Link Column: Leave blank Gene link URL: Leave blank Analyse server side: Check the box Min gene read count: 10 Min gene CPM: Leave Blank</p>
14	Define the two conditions (which are strains in this case) and replicates by clicking on 'Add condition'.
15	Name the first condition as AL218 , and select the columns starting with numbers 25861–25866 as the 6 replicates.
16	Name the second condition as KPC2 , and select the columns starting with numbers 25867–25872 as the 6 replicates .
17	Click on View. The most over/underexpressed genes in one strain relative to the other can be interactively explored using various tools (e.g. the volcano plot as shown below)

The screenshot shows the Degust web application interface. At the top, there are browser tabs and navigation links. The main content area is titled "Degust : test2" and includes a "Conditions" panel on the left with checkboxes for "AJ218" and "KPC2", and a "Method" dropdown set to "Voom/Limma". The central part of the interface is a "Volcano" plot showing $-\log_{10} \text{FDR}$ on the y-axis and $\log \text{FC}$ on the x-axis. A legend below the plot shows "AJ218_avg" in blue and "KPC2_avg" in red. To the right of the plot is an "Options" panel with settings for "FDR cut-off", "abs log FC", "FC relative to", "MA-Plot FC", and "Show Counts". Below the plot is a "Genes" section with a table showing the top 4 genes.

Geneid	Annotation	FDR	P value	AJ218	KPC2
KpAJ218_chr_03840		6.14e-12	1.22e-15	0.00	6.39
KpAJ218_chr_03839	gene_name=soxS_2;gene=soxS_2;	3.18e-11	1.26e-14	0.00	5.40
KpAJ218_chr_04520		1.35e-8	1.58e-10	0.00	4.97
KpAJ218_chr_01159	gene_name=rob_1;gene=rob_1;	1.82e-7	3.26e-9	0.00	6.46



omics.data.edu.au

If you need further assistance, please contact us at omicsdataservices@lists.unimelb.edu.au