Baskin Engineering
UC SANTA CRUZ

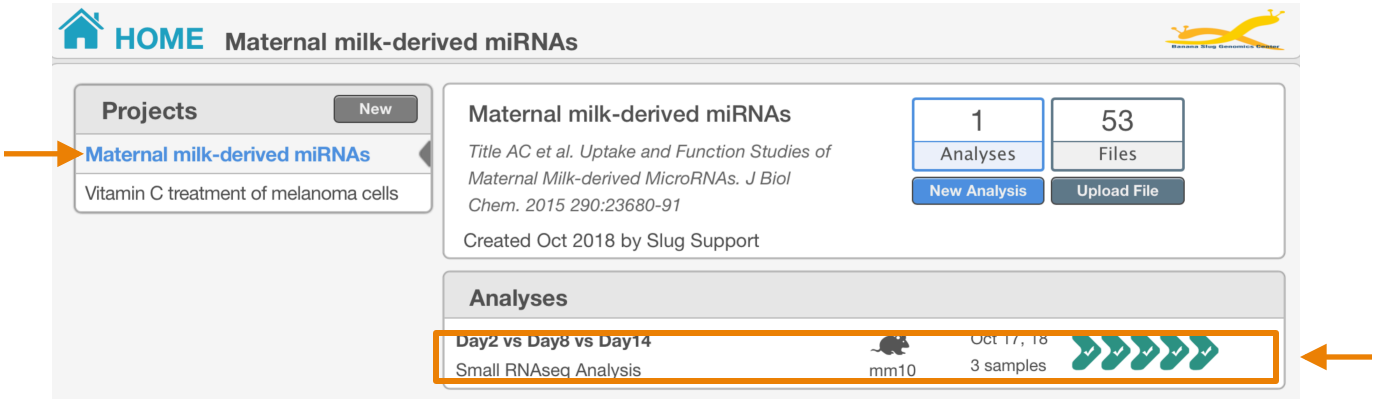
Banana Slug Analytics Platform

Small RNA-Seq Analysis Results

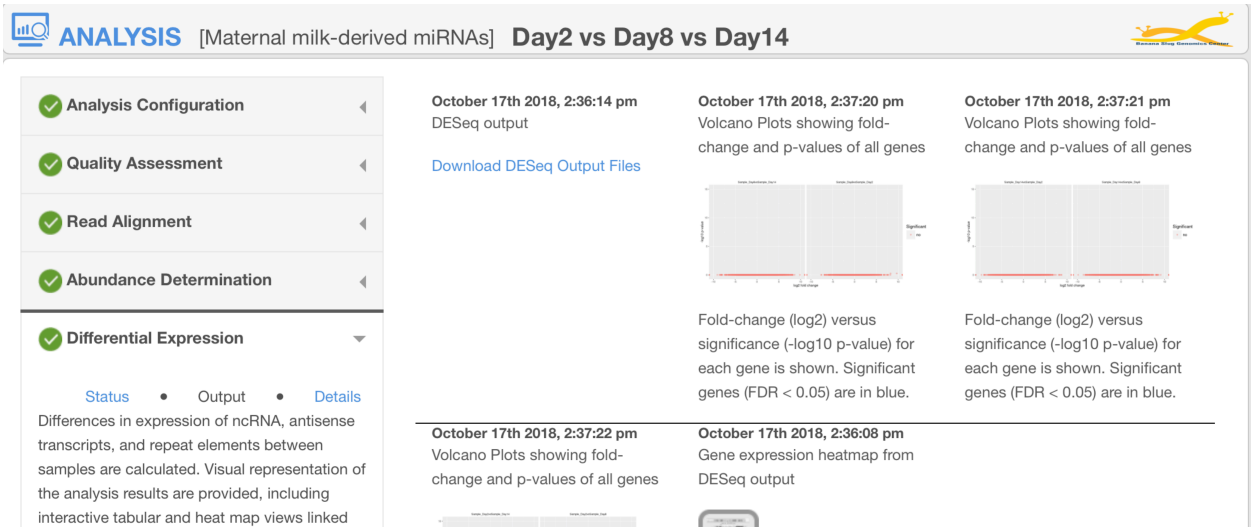
Rev. 01/2018

This document provides a guide to explore the Small RNA-Seq analysis results on the Banana Slug Analytics Platform. This should be expected to be used after the analysis has been performed.

At the HOME page of your account, select the project in the Projects panel on the left that has the analysis performed. Then click on the analysis listed under the Analyses panel on the right.



The ANALYSIS page will be opened with all the results displaced at each corresponding stage of the analysis workflow: Analysis Configuration, Quality Assessment, Read Alignment, Abundance Determination, Differential Expression. You can expand the desired stage on the left to explore the data.

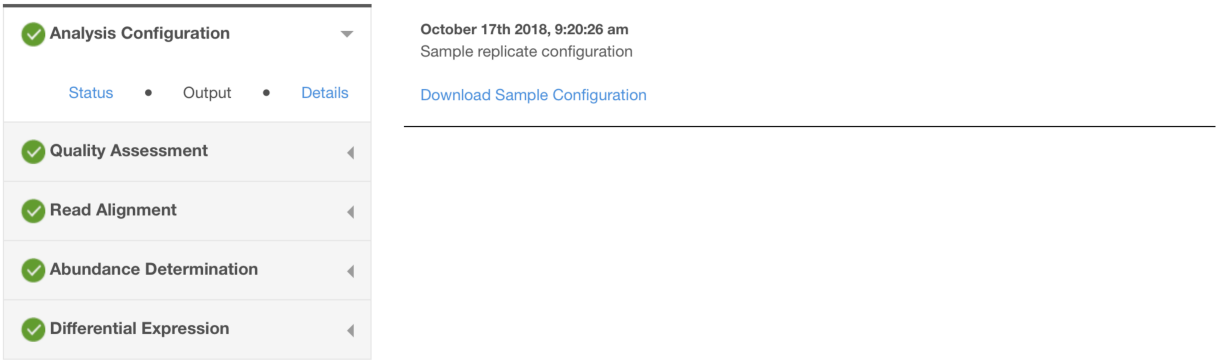


Each stage has three types of presented information:

- Status: Current status of the analysis stage with start and completion timestamps
- Output: Results produced at the corresponding analysis stage
- Details: Detailed information of analysis stage, for example, major software command and options used during the analysis

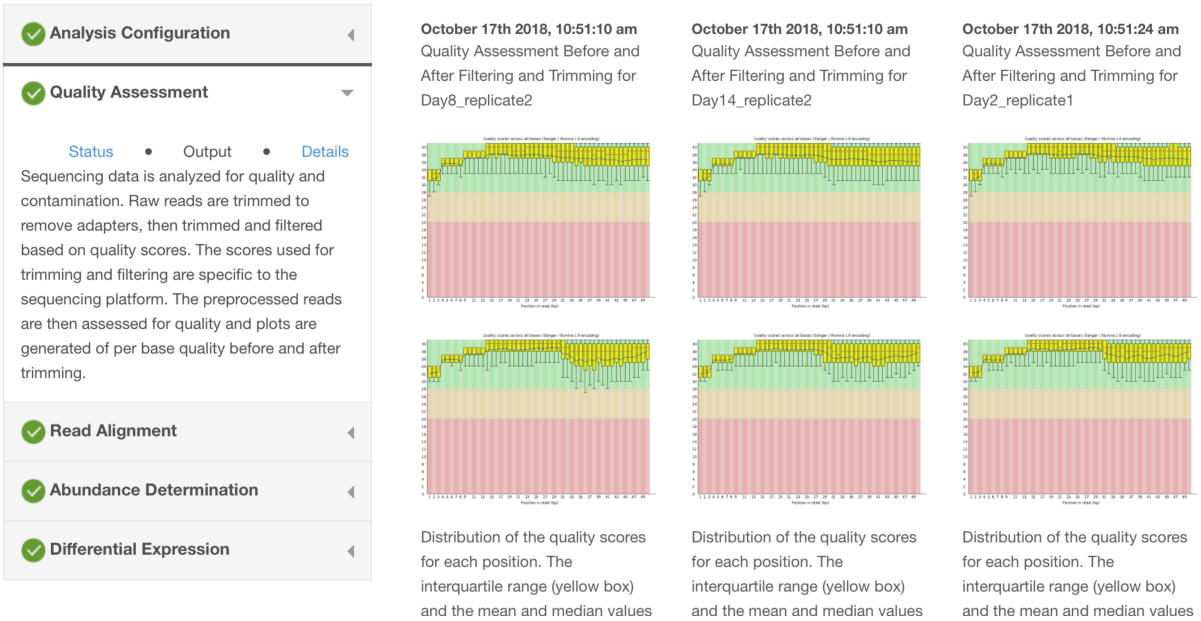
Analysis Configuration

Users can download the sample configuration file that includes the sample/replicate names and their corresponding FASTQ files used in the analysis. This is especially useful for analysis that includes multiple replicates for each treatment or tissue sample.



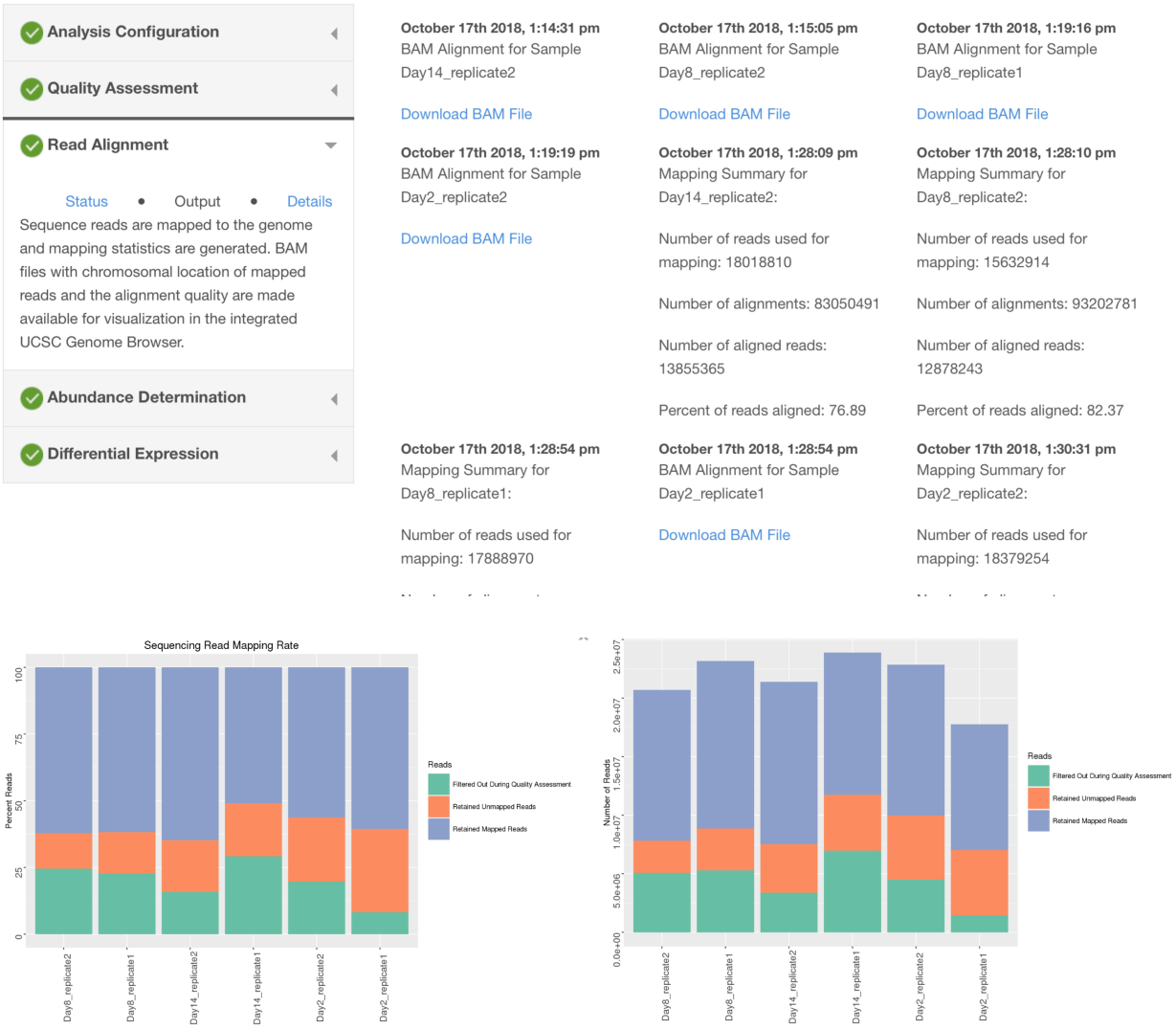
Quality Assessment

This stage includes checking the quality of the raw sequencing data, trimming low quality reads, and computing statistics of read trimming results. Users can enlarge plots for visualization by clicking on the plots or designated links or download plots on the FILES page of the platform.



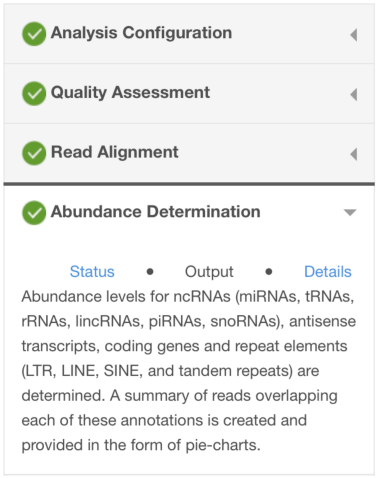
Read Alignment

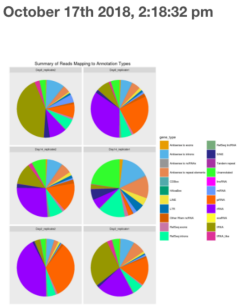
This stage includes aligning trimmed reads from the quality assessment stage to the reference genome and computing read mapping statistics. The outputs provide the number of reads used for mapping, number of alignments, number of aligned reads, and percent of aligned reads in each sample/replicate. Summary bar charts are included to show the mapping rates of all the samples in the analysis.



Abundance Determination

This stage includes estimating the abundance of each annotated gene feature in the reference genome for each sample/replicate. Users can obtain the summary information through the pie charts or download the raw read counts.





October 17th 2018, 2:18:32 pm

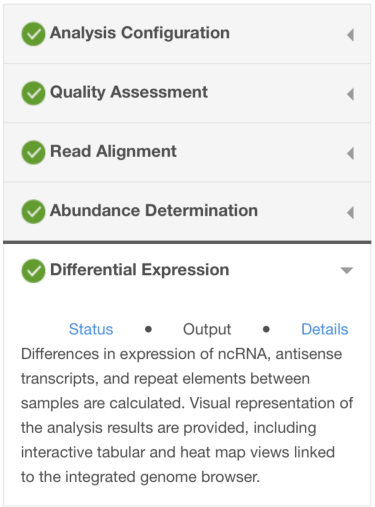
Pie-chart summarizing the number of reads mapping to each annotation type in each sample

October 17th 2018, 2:18:45 pm
Raw Read Counts

Download Raw Read Counts File

Differential Expression

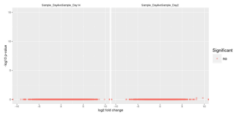
This stage includes the differential expression analysis results using DESeq. Users can download the DESeq output files, visualize the pairwise analysis volcano plots of fold-changes and adjusted P-values of all the genes, and launch the interactive DESeq heatmap with the UCSC Genome Browser.



October 17th 2018, 2:36:14 pm
DESeq output

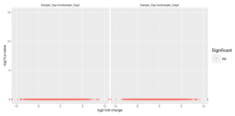
Download DESeq Output Files

October 17th 2018, 2:37:20 pm
Volcano Plots showing fold-change and p-values of all genes



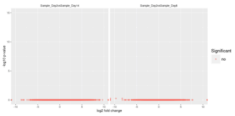
Fold-change (log2) versus significance (-log10 p-value) for each gene is shown. Significant genes (FDR < 0.05) are in blue.

October 17th 2018, 2:37:21 pm
Volcano Plots showing fold-change and p-values of all genes



Fold-change (log2) versus significance (-log10 p-value) for each gene is shown. Significant genes (FDR < 0.05) are in blue.

October 17th 2018, 2:37:22 pm
Volcano Plots showing fold-change and p-values of all genes

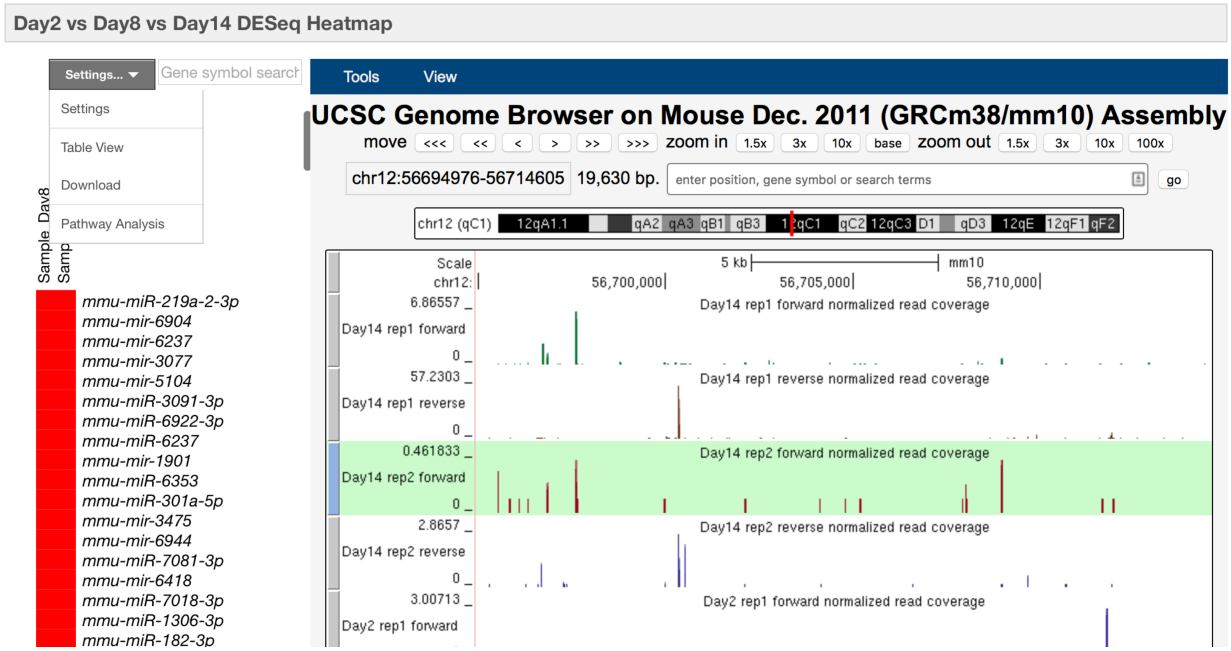


Fold-change (log2) versus significance (-log10 p-value) for each gene is shown. Significant genes (FDR < 0.05) are in blue.

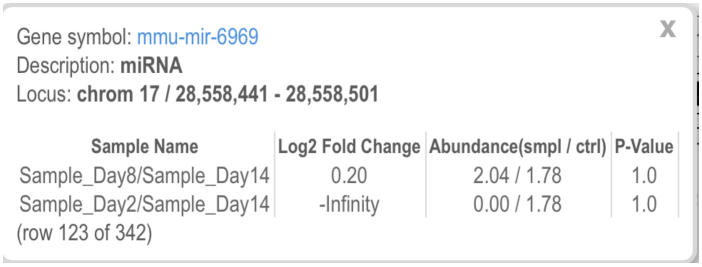
October 17th 2018, 2:36:08 pm
Gene expression heatmap from DESeq output

View Day2 vs Day8 vs Day14 DESeq Heatmap

Interactive DESeq Heatmap

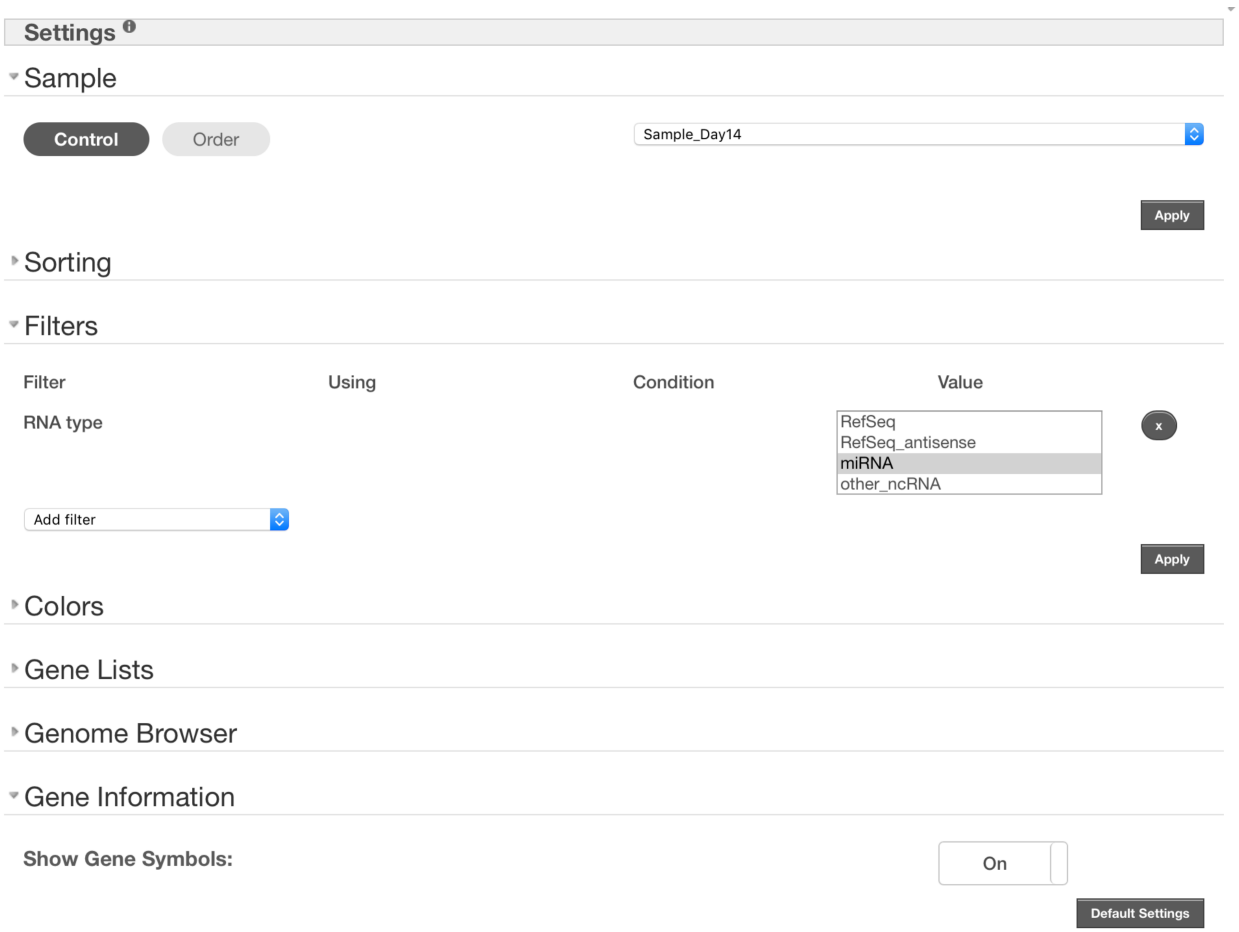


Users can select the desired gene feature on the interactive heatmap to explore its expression profile and read coverage on the UCSC Genome Browser. By hovering a gene feature on the heatmap, the expression data including log2 fold change, abundance, and adjusted P-value will be displayed on a popup window.



To view the read coverage of a desired gene feature on the genome browser, click the gene on the heatmap and the corresponding genomic region will be shown.

The heatmap data can be displayed in a table format by selecting Table View under the Settings menu. User can also download the expression data in display by selecting the Download option. To change the display and filtering options of the heatmap, choose the Settings menu item.



By default, P-value < 0.05 is set as a filter to display data on the heatmap. If too few genes are shown, users can select to relax or remove the filter. Pairwise differential expression analysis is included. Users can select the appropriate sample as the control for comparison on the heatmap. Users can also upload a pre-determined gene list to be displayed on the heatmap.

For details on how to use the UCSC Genome Browser, please go to https://genome.ucsc.edu/training/index.html.