Contents

Accessibility

IRIS can be freely accessed directly through this link (**change link later**) or through R using the following commands (**not available yet**):

```
if (!require("shiny")) install.packages("shiny")
shiny::runGitHub("vidger-shiny", "btmonier")
```

Typically, the link will provide an easier route to using IRIS. In circumstances where internet connections will be limited (such as during travel), loading IRIS through R while internet is still available will allow users to utilize IRIS without an internet connection later on.

Input data

IRIS requires two pieces of information for analysis. The first is an expression estimation matrix, also referred to as a count matrix, displaying the gene expression estimates for each sample. The format requires a CSV file with the row names to list the gen IDs and column names to list the sample IDs. The second required input is a condition matrix, wherein the factor levels for each sample are provided.

This file requires a CSV format and row names to be the sample IDs matching the sample IDs from the expression estimation matrix and the column names to be the condition factors.

For the following tutorial, the data used for thhis tutorial are derived from 28 *Vitis vinifera* (i.e. grape) samples with three distinct factors (Rootstock, row, and block).

Expression matrices

Typically, an expression matrix, also known as count data or a count matrix referes to data where every i-th row and j-th column refer to how many reads are assigned to gene (ID) i in sample j. For example, if we have simplified count data for 4 samples and three genes, the R output will look something like this:

```
         sample1 sample2 sample3 sample4
gene001      23       3      45       2
gene002       6       7       7       8
gene003       0      34       3      42
```

NOTE: When loading count data into IRIS, make sure that the first column is your gene IDs and that sample names are short, concise, and avoid the use of mathematical operators (+, −, /, ⋆, ^, etc.) and spaces between words. If a space is necessary for legibility, please consider using an underscore (_)

Condition matrices

Condition matrices, also known as metadata, details the design of your experiment. In this type of matrix, every i-th row and j-th column refer to factor levels assigned to sample i and factor j. For example, if were to look at the samples given in the count data section, the metadata R output will look something like this:

```
         condition time
sample1   treated   0h
sample2 untreated   0h
sample3   treated  24h
sample4 untreated  24h
```

NOTE 1: When loading metadata into IRIS, make sure that the first column is your sample names and that column names and treatment levels are short, concise, and avoid the use of mathematical operators (+, −, /, ⋆, ^, etc.) and spaces between words. If a space is necessary for legibility, please consider using an underscore (_)

NOTE 2: Metadata can be expanded to fit the nature of your experiment (i.e. multiple factors can be added). The only thing that must remain consistent between these two matrices, is the sample information. Column names in count data **must** be the same as row names in the metadata.

Expedited analysis

Overview

Expedited analysis is for users who want a quick and efficient method of producing DGE results using the default parameters and tools in IRIS.

Submit and QC / DGE Analysis

1. Load user data by selecting “Load my own data”
    - User data requires one count matrix and one condition matrix
2. Click “Submit” to load the data.
    - **insert image**
3. After submitting the data, proceed to the “DGE Analysis” tab at the top.
    - If a specific experimental design other than the basic two-group comparison is required, select it accordingly.
4. Select the factor of interest and check the boxes for the comparisons of interest.

    NOTE: you must choose at least one comparison for “two-group comparisons” or “multiple factor comparisons” experimental design options or else an error will occur!
5. Submit the parameters to perform DGE analysis.
    - **insert image here**
6. Select the “Plots” subtab to view the DGE results table towards the bottom.

Data export

1. To export results data from the “DGE Analysis” tab, click on the “Download All Data” button at the bottom of the page to download the results file.
    - **inser image here**

In-depth Analysis

Overview

By choosing the in-depth analysis route, you will be given more information about your data including:

- Count data distribution;
- Total number of reads/sample;
- Sample correlation analysis;
- Biclustering;
- Principal components and multidimensional scaling;
- Identifiication of most variable IDs in data.

Submit and QC

Preliminary Analysis

Differential Expression Analysis