**INSTRUCTION FOR RUNNING THE CODE**

The code for Co-expression network has been written in RStudio, an IDE which is a powerful and productive user interface for R, the statistical programming language.

**R/RStudio Setup Guide**

**Prerequisites for RStudio** : Any version of R

This document details the step-by-step installation of R and RStudio (IDE), one of the free and open source integrated development environment for R. These software packages can be downloaded from [http://www.r-project.org/](http://www.r-project.org/%20) and [http://rstudio.org/](http://rstudio.org/%20) respectively and are available on the Windows, Linux and Mac OS X platforms. For the sake of simplicity, this guide shows the installation for the Windows platform.

**Installation of R on a Windows operating system:**

* Download the latest precompiled binary distributions from CRAN website [<http://www.rproject.org/>].
* Only the base package is required for this installation. (At the time of writing the latest version of R is R-3.2.2) .
* Follow the instructions on the website to complete the installation of R.

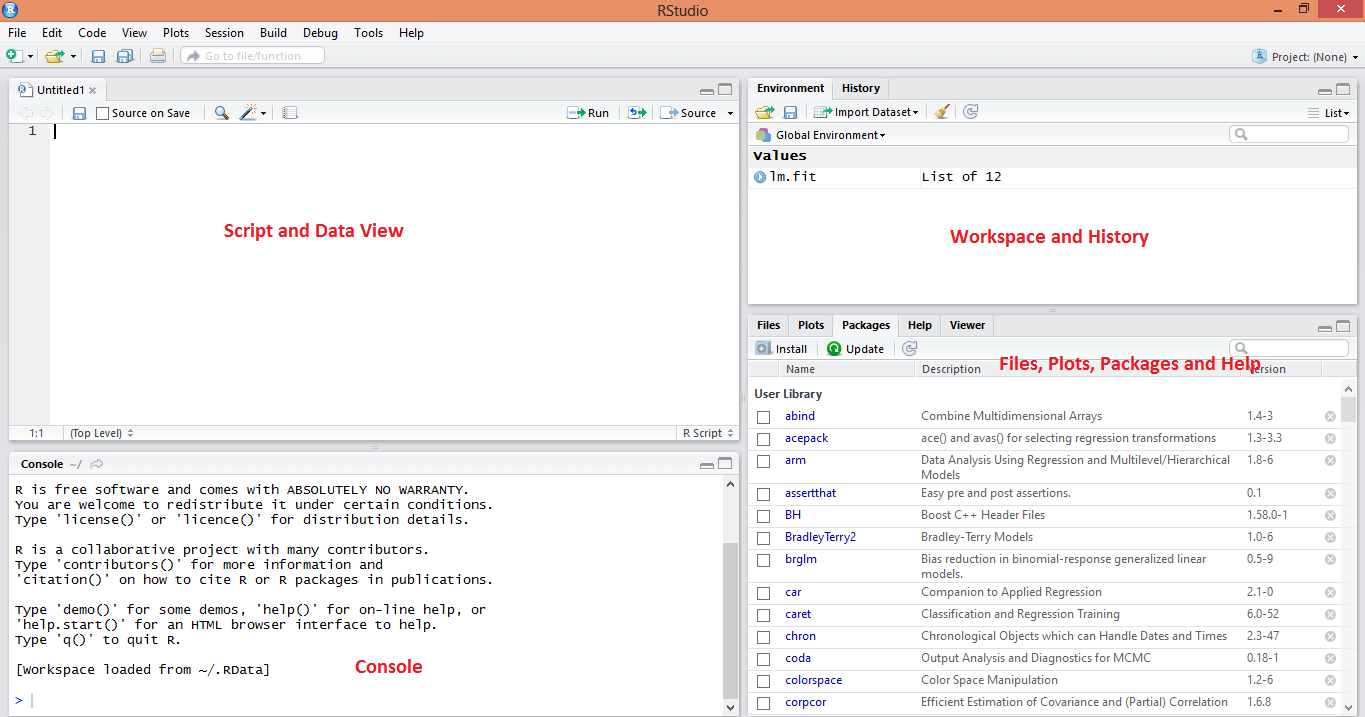
**Installation of RStudio IDE on Windows operating system :**

* Download the latest version of RStudio IDE for your Windows platform from [http://rstudio.org/download/desktop](http://rstudio.org/download/desktop%20) (At the time of writing the latest available version of RStudio is v0.99.489)
* Start the installation and follow the steps required by the Setup Wizard.
* Once completed, launch RStudio IDE from Start -> All Programs-> RStudio -> RStudio.exe or from your custom installation directory. The default installation directory for RStudio IDE is "C:\Program Files\RStudio\bin\rstudio.exe".

**RStudio screen:**

The usual Rstudio screen has four windows:

1. Console
2. Workspace and history
3. Files, plots, packages and help
4. The R script(s) and data view.



* Set your working directory to your R-scripts using the command, setwd(dir) For example, setwd("C:\EMTS\R").

Alternatively you can use RStudio's Tools Menu: Tools -> Set Working Directory -> Choose Directory...

**Installing required packages with RStudio IDE for Co-Expression Network**

Packages ggplot2, reshape2 and qgraph are required for drawing the Heatmap and constructing the co-expression Network.

* All required packages can be downloaded from the CRAN Repository or installed directly, along with any dependencies, by issuing the following commands on the R-console,

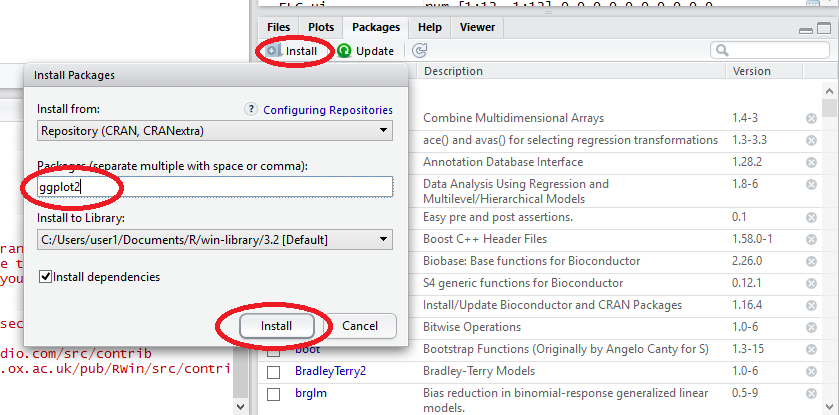
>install.packages("ggplot2")

>install.packages("reshape2")

>install.packages("qgraph")

Alternatively, you can easily install each of these packages using RStudio IDE by :

* Clicking the “Install Packages" button present at the bottom right window of the screen(Files, plots, packages and help window) .
* You will be asked to create a personal library directory.
* Click on Yes.



* However you need to be connected to the internet for installing the packages.

**Running the script**

* To run a script, open the script in RStudio's script editor and choose "Source" from the menu or type the following command at the Console prompt,

> source('C:/myproject/R/final\_code.R')

> co\_ex()

Alternatively you can run the script by :

* Copy the code and paste it in the script window.
* Click on the “Run” button at top of the script screen to compile it.
* Call the function by typing co\_ex() in the R-console.

**Demonstration using a Sample data set**

The file which is to be read into the function should be in tab delimited text format with variable names as the topmost row. There is a file “testdata.txt” in the same repository to give an idea how the input file should look like. This data set was created using random numbers generated from RStudio.

Here I will demonstrate the process of arriving at the required plot using the file “testdata.txt”. So let’s get started.

Type the following in the R-console and press Enter

>co\_ex()

After this it will ask you to select the input file. Choose it from where it is stored. After this the following lines will appear

Enter an integer

Pearson Correlation Matrix: Enter 1

Cross-Correlation Matrix : Enter 2

Choose the correlation matrix you want by typing the asked integer and press Enter. I will show it using the Pearson Correlation matrix. Cross-Correlation matrix works in the same way.

After this the following line will prompt up

Enter cut-off value(0-1):

This absolute value will prune the correlation matrix so that only significant correlation values remain. However this is for only drawing the network diagram as too many trivial lines will make the diagram difficult to interpret. The Heatmap will not be affected by it. I will continue by setting the cut-off value as 0.8

Then another prompter will ask you to enter an integer to choose between the type of diagram/ plot you want. I will continue with 3(Both type of plots).

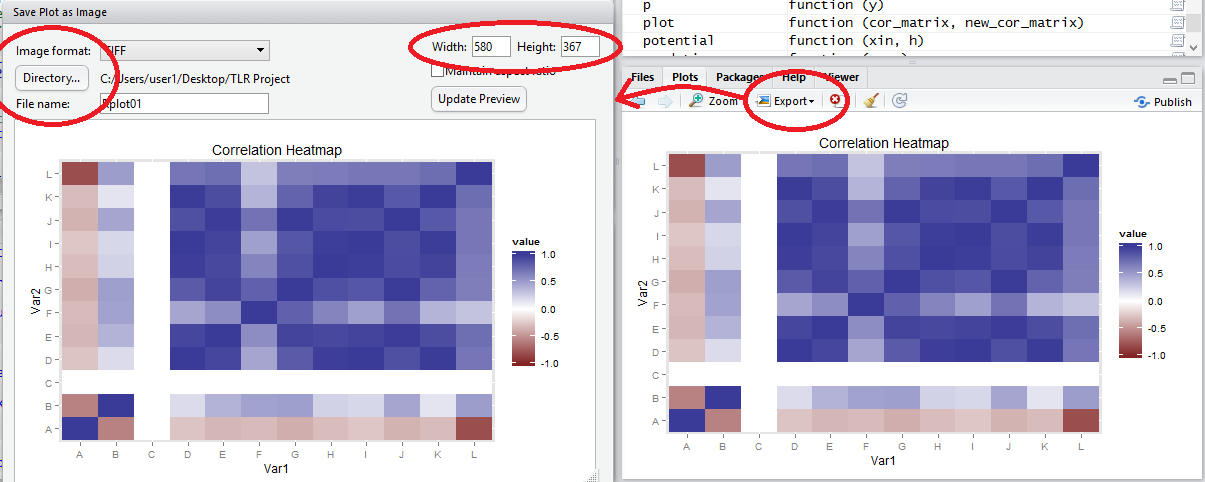
Enter an integer

For Correlation Heatmap : Enter 1

For Co-Expression Network : Enter 2

For Both : Enter 3

The Heatmap will appear at the Files, Plots, Packages and Help window. You can save it using the Export option in the same window. Another window will pop up. You can choose the format, file name and dimension of the image and the directory in which it will be saved.



Here the row and column corresponding to the C variable is empty because the fold change is between 0.5 and 2 for all the time points. Hence it is unresponsive to the experiment.

Meanwhile you can see the R-console is already asking you to Press Enter for displaying the Co-Expression Network.

Press Enter for Co-Expression Network

C:\Users\user1\Desktop\TLR Project\Rplot01.tiff

You can save the image using the same method as discussed in the above Heatmap example.

The color coding of borders of the variables is based on the following notation:

* Red: If at least one data point < 0.5 and no data point > 2
* Green: If at least one data point > 2 and no data point < 0.5
* Blue: If at least one data point > 2 and at least one data point < 0.5
* Black: If all data points < 2 and > 0.5

Color coding of the Network edges is based on the following notation:

* Green: Positive correlation value between the two variables
* Red: Negative correlation value between the two variables

The intensity of the color of the edges denotes the magnitude of the correlation value between the two variables.