

2013•2014
FACULTY OF SCIENCES
Master of Statistics

Master's thesis
Software development for biclustering methods: The biclust GUIR package

Promotor :
Prof. dr. Ziv SHKEDY
De heer Martin OTAVA

Ewoud De Troyer
Thesis presented in fulfillment of the requirements for the degree of Master of Statistics

Transnational University Limburg is a unique collaboration of two universities in two countries:
the University of Hasselt and Maastricht University.



Universiteit Hasselt | Campus Hasselt | Martelarenlaan 42 | BE-3500 Hasselt
Universiteit Hasselt | Campus Diepenbeek | Agoralaan Gebouw D | BE-3590 Diepenbeek



2013•2014
FACULTY OF SCIENCES
Master of Statistics

Master's thesis

Software development for biclustering methods: The
biclust GUIR package

Promotor :
Prof. dr. Ziv SHKEDY
De heer Martin OTAVA

Ewoud De Troyer

*Thesis presented in fulfillment of the requirements for the degree of Master of
Statistics*

Acknowledgments

First of all I would like to thank my parents for always being there for me and providing me with everything I needed to be able to follow my dreams. Without your support and love and I would never have been able to find the energy time after time, in good and bad moments, to chase my ambitions and always give it the best I have. Thank you mom for all the moments we shared. Thank you dad for being there for me every step of the way.

Next I would also like to extend a lot of gratitude to professor Ziv Shkedy. It was a pleasure having you as my promotor for both my bachelor and master thesis. Your teachings, kindness and advice have each time inspired me to go the extra mile! You were an incredible mentor these past 2 years and thanks to your unrelenting support I will be able to learn even more from you the coming years during my PhD.

Further I would also like thank Martin Otava for his great guidance during the stages of R software development. Your crash courses were greatly appreciated and of immense help to me! More thanks go out to Tatsiana Khamiakova whose insight in superbiclust was of great aid to me in improving my understanding of the package as well as its implementation in the GUI.

Next I am also very grateful for the quick response and help of Mengsteab Aregay in getting the bcdiag package back online on CRAN!

I would also like to greatly thank my landlady and landlord, Maria and Kamiel, who were always concerned about my well-being. They were always there to lend a helping hand if necessary. Your Sinterklaas days will not soon be forgotten!

Last, but certainly not least, I have to thank 2 more persons who were of great importance to me during these past years at U Hasselt. Marijke and Kevin, without your support, laughs, tears and discussions, I would have never made it as far as I have. You were always there for me, in prosperity and adversity. You kept me on my feet the entire way through. You cannot fathom how happy I am I can share 4 more close years with the both of you, not only as fantastic friends, but also as colleagues.

Abstract

This thesis is centered around the subject of *biclustering* which is the simultaneous clustering of genes and conditions in a gene expression matrix. The focus will not be on the actual methods themselves, but on the software development of their implementation in a user-friendly *Graphical User Interface*.

This GUI is packaged as `RcmdrPlugin.BiclustGUI` which is a plug-in for the already existing `R-commander`. The report will handle the implementation of several biclustering methods as well as ways to extract diagnostics and plots from the results, all by using the GUI.

The second part of the report will handle the implementation of future new biclustering methods or diagnostic tools. Because the framework of `RcmdrPlugin.BiclustGUI` was created specifically for biclustering, this is can be achieved very easily. Thanks to easy-to-use scripts, developers can freely design their own GUI windows for their methods which can then later be added to `RcmdrPlugin.BiclustGUI` very quickly by the maintainer. The strength of these scripts lies in the fact that they do not rely on any knowledge of `tcltk` or `Rcmdr` which is normally required for creating such windows. This enables `RcmdrPlugin.BiclustGUI` to quickly gain a vast collection of biclustering methods which will then be accessible by non-R users through a simple point-and-click system. Further, it is also an attractive approach to easily connect biclustering methods with diagnostic and plotting packages to further investigate and visualize bicluster results.

Contents

1	Introduction	1
1.1	R Commander	1
1.2	Biclustering	2
2	A New Concept in Software Development using R	5
3	The BiclustGUI R Package	7
3.1	Installing and Loading	7
3.1.1	R Commander	7
3.1.2	RcmdrPlugin.BiclustGUI	7
3.1.3	Data Input	7
3.2	GUI Structure	8
3.3	biclust-package	9
3.3.1	Plaid Biclustering	9
3.3.2	CC(δ) Biclustering	10
3.3.3	XMotifs Biclustering	11
3.3.4	Spectral Biclustering	12
3.3.5	QuestMotif Biclustering	13
3.3.6	Bimax Biclustering	13
3.3.7	Biclust: Plots & Diagnostics	14
3.4	fabia-package	16
3.4.1	Laplace Prior	17
3.4.2	Post-Projection	18
3.4.3	Sparseness Projection	18
3.4.4	SPARSE	19
3.4.5	Fabia: Plots & Diagnostics	20
3.5	isa2-package	22
3.6	iBBiG-package	24
3.7	Diagnostic Packages	25
3.7.1	BcDiag-package	25
3.7.2	superbiclust-package	27
4	A Guideline to New Implementations	29
4.1	Implementing a New Method	29
4.1.1	New Method Script - ClusterTab & PlotDiagTab	29
4.1.2	The Frame Scripts	34
4.1.3	A quick example - Plaid	40
4.2	Implementing a New Tool	41
4.2.1	New Tool Script	41
4.2.2	A quick example - BcDiag	42
4.3	Testing your Script	44
5	References	45
6	Appendix	47
6.1	Introduction	47
6.2	The BiclustGUI R Package	48
6.2.1	biclust-package	48
6.2.2	fabia-package	51
6.2.3	isa2-package	55
6.2.4	iBBiG-package	55
6.2.5	BcDiag-package	56
6.2.6	superbiclust-package	57
6.3	Guideline - Template Scripts	57
6.3.1	newmethod.script	57
6.3.2	frames.script	59
6.3.3	Quick Example - Plaid	61
6.3.4	newtool.script	71
6.3.5	Quick Example - BcDiag	72
6.3.6	Extra Example - Superbiclust	78

1 Introduction

Data obtained from microarrays has been subject to a variety of studies. The art of identifying differentially expressed genes, finding high predictive genes or groups of genes with respect to a response or conditions has been of major interest in order to explore the structure of these *gene expression datasets*.

Another way to investigate this kind of data is through the help of *clustering*. Through clustering gene expression matrices can be analyzed in both dimensions, the gene and condition dimension. This translates into the grouping of genes according to their expression under multiple conditions and the grouping of conditions based on the expression of a number of genes. These results can then for example be utilized for classification afterwards.

This procedure can be extended to clustering on the genes and conditions simultaneously which is called *biclustering*.

There exist a great deal of methods to do biclustering for which many R packages have been developed. For example `biclust` (Kaiser and Leisch, 2008), `fabia` (Hochreiter *et al.*, 2010), `isa2` (Csardi *et al.*, 2010) and `iBBiG` (Gusenleitner *et al.*, 2011) are some of these.

But while there are a lot of R packages available, not a lot of user-friendly graphical user interfaces (GUI) exist to execute these biclustering methods. This is particularly helpful for scientists with limited knowledge of R as they can apply the multitude of methods through simple point-and-click dialogs. Further, since the GUI is a plugin of `Rcmdr`, the user will also be exposed to the actual R commands of the implemented packages, making it a learning experience on the use of R.

The `RcmdrPlugin.BiclustGUI` is a continuation of the same-named package available on R-Forge, made by Setia Pramana. The already implemented biclustering algorithms were completely redone, adding extra parameter options in the process. New methods have been added as well as more options to graphically present the results of all methods. Further all biclustering procedures have been implemented in the GUI in a very specific framework which makes adding new packages in the future a quick and easy task with minimal interference of the maintainer. In short, the method/package developer will be able to create his own dialogs/windows for his procedure *without* having to rely on any knowledge of the `Rcmdr` or `tcltk` syntax which is normally necessary to create these.

Thanks to this, the package has the potential to become a GUI from which a vast collection of biclustering methods can be accessed in the future.

Following on the introduction, there will be two major sections about the `RcmdrPlugin.BiclustGUI` package. In the first, the already implemented biclustering methods will be briefly explained followed by showcasing the functionality of the GUI. This will contain the executing of the procedure itself as well as the visualisation of the results through the appropriate plots.

In the second section, an extensive guideline will be presented with explanatory examples on how to create the appropriate scripts to add new biclustering procedures. The main idea is that these scripts can then be send forward to the maintainer of `RcmdrPlugin.BiclustGUI` who can then easily add these new dialogs to the already implemented procedures.

It should be stressed that over the entirety of the report the focus will be on the software development of the GUI. While a multitude of biclustering methods will be visited, the goal is not to investigate into detail how they work, but how they are implemented in the GUI. What will be of interest though is the structure of the GUI, namely how the several methods are automatically linked to the plots and diagnostics without unnecessary interference of the user. Therefore the description of the implemented methods will be very brief and basic.

1.1 R Commander

R Commander, `Rcmdr`(2005), is a GUI developed by John Fox from McMaster University, Canada. Originally it was conceived as a basic-statistics graphical user interface for R, but its capabilities have been extended substantially since. The `Rcmdr` package is based on the `tcltk` package (Dalgaard, 2001) which provides an R interface to the *Tcl/Tk* GUI builder. Since `tcltk` is available on all the operating systems on which R is commonly run, the R Commander GUI will also run on all of these platforms.

The GUI is also very easy to start to use for beginners who do not have any or little experience with R. It will protect beginners from errors as the dialog boxes only have limited options related to the current context which minimizes the errors made by users. Further, since the users are exposed to the actual R commands through a script and output window, besides analyzing and managing the data in R easily, they can also learn how to do it in R without a GUI. Another advantage is that the script will be generated on the fly as the user applies the desired statistics through the point-and-click GUI. This means it can be easily saved at the end of a

session which enables the user afterwards to recreate the results by running the R script without going through all the dialogs again. Advanced users can even adapt the created script to do some more detailed analysis. These are the main advantages Rcmdr has over other available RGUI packages.

Starting with version 1.3-0, Rcmdr also provides the possibility of *plug-in* packages which can augment the R Commander menus. These packages are developed, maintained, distributed and installed indepently of Rcmdr and can provide a wide variety of new dialog boxes and statistical functionality. More information on developing such a plug-in can be found in Fox (2007).

It is through this functionality that the RcmdrPlugin.BiclustGUI has been created which brings a new menu to the R Commander GUI containing a collection of biclustering procedures.

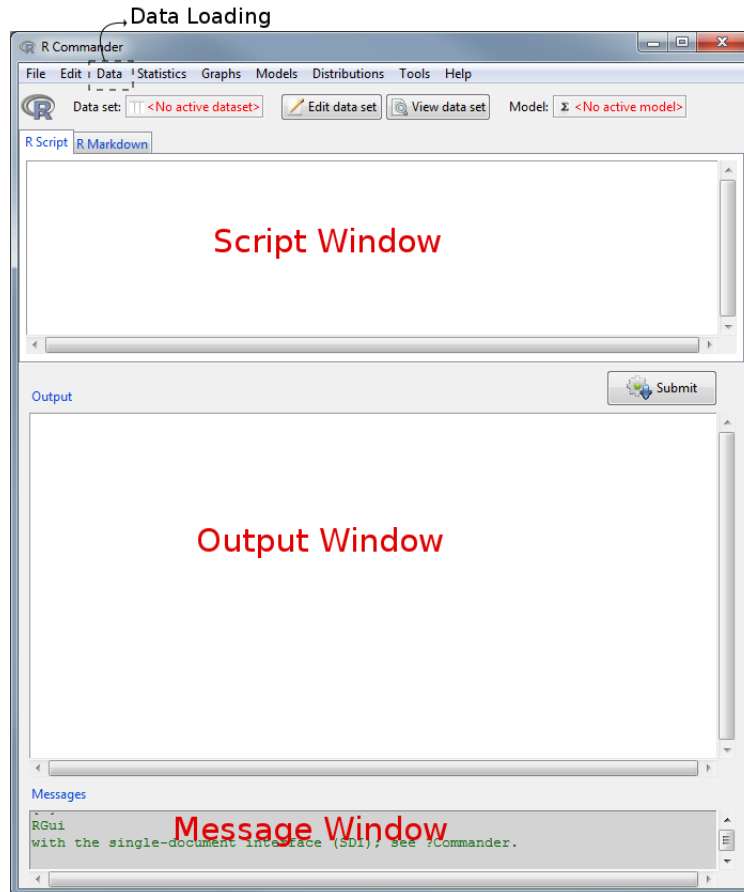


Figure 1: Default R Commander

1.2 Biclustering

Let \mathbf{Y} be a $m \times n$ matrix. The goal of *biclustering* now is to find subgroups of rows and columns which are as similar as possible to each other and as different as possible to the rest (Kaiser and Leisch, 2008). This basically comes down to clustering on both the row and column dimension simultaneously and while clustering methods on 1 dimension derive a *global model*, biclustering algorithms will produce a *local model*. For example in clustering algorithms each row in a rowcluster is defined over all the columns, however a row in a bicluster is selected using only a subset of columns.

Going back to the matrix \mathbf{Y} , this corresponds to looking for submatrices with a high similarity of elements. This submatrix is what is called a *bicluster*.

$$Y = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

While biclustering has its applications in many areas such as marketing and behavior science, a popular use of it is for the analysis of genetic data. The \mathbf{Y} matrix would then be a gene expression matrix in which the rows correspond with the genes and the columns with the conditions/samples. In this setting a_{mn} is the expression

level of the m th gene under the n th condition and biclustering algorithms will identify groups of genes that show similar activity patterns under a specific subset of the conditions. Therefore, biclustering is the key technique in situations where a cellular process, in which only a small set of the genes participate, is of interest or in the situation that an interesting cellular process is active only in a subset of the conditions.

There exist a great deal of different biclustering algorithms today and several authors have provided extensive reviews, discussion and comparisons of these (Maderia and Oliveira, 2004, Tanay *et al.*, 2004, Prelić *et al.*, 2006).

All these different algorithms are able to identify several types of biclusters which can be identified into four major classes:

1. Biclusters with constant values
2. Biclusters with constant values on rows or columns
3. Biclusters with coherent values
4. Biclusters with coherent evolutions

To briefly elaborate on this, coherent values would mean that each row and column can be obtained by adding a constant to each of the others or by multiplying each of the others by a constant value. Further in the evolution approaches, the elements of the matrix are viewed as symbolic values and biclusters with coherent behaviour will be discovered regardless of the exact numeric values in the matrix. These types of biclusters can be found in Figure 37 in the Appendix.

It is also interesting to know that while many biclustering algorithms perform simultaneous clustering on both dimensions of the data matrix, there also exist two-way clustering approaches. These use one-way clustering to produce clusters on both dimensions of the data matrix separately.

Finally, if a biclustering algorithm assumes the existence of several biclusters, several underlying structures of the data matrix can be considered. Madeira and Oliveira (2004) listed the following eight major types of underlying bicluster structures:

1. Exclusive row and column biclusters (rectangular diagonal blocks after row and column reorder).
2. Non-Overlapping biclusters with checkerboard structure.
3. Exclusive-rows biclusters.
4. Exclusive-columns biclusters.
5. Non-Overlapping biclusters with tree structure.
6. Non-Overlapping non-exclusive biclusters.
7. Overlapping biclusters with hierarchical structure.
8. Arbitrarily positioned overlapping biclusters.

These different structures are visualized in figure 37 in the Appendix.

2 A New Concept in Software Development using R

There exist a great number of packages for R. On CRAN for example there are almost 6000 packages available which handle many topics. Due to the fact that for each topic or area many packages can be installed, it can sometimes be challenging for new users to quickly execute a certain analysis of interest (e.g. dose-response modelling, biclustering methods, etc.).

The creation of ‘*envelope packages*’, related to a common data analysis problem, can be a solution to this often-occurring problem. An attractive way to implement this, is with the help of a Graphical User Interface (= GUI).

The idea is to design a joint development programming environment in which methodological developers can include their packages in this ‘envelope package’ without too much programming effort.

Not only can multiple people contribute to this ‘envelope package’, the users will be able to make use of all of these packages in one clear environment. To clarify, all the implemented packages in the ‘*envelope package*’ are still independent entities on CRAN or Bioconductor. This means the ‘*envelope*’ package is simply depending on these other packages, not incorporating the actual code behind them. As a result, any updates by the authors on an outsourced package will also automatically update the ‘*envelope package*’. The main goal of the ‘*envelope package*’ is to provide a solid and intuitive ‘module-like’ structure in which the external packages can fit in, while also providing the possibility of interconnecting these packages.

So while the packages are still independently developed, the end user will only need this ‘envelope package’ to carry out a specific analysis for which multiple methods exist. The ‘envelope package’ can make sure the output of implemented packages is processed in a similar way, providing a user-friendly environment to execute each step of an analysis. An example of this could be that after applying a certain analysis method, the results can immediately be visualized through an appropriate button in the GUI.

Further, such an ‘envelope package’ also introduces a lot of flexibility in the way several packages can be connected, providing an easier workflow. For example a package focused at the visualisation of results could easily be connected with other packages in a GUI, alleviating the work of the user in figuring the right data input for each function.

Also, by creating the GUI as a plug-in for R-Commander, it will not deny the user access to the actual R code which is used for applying a certain analysis. The R-Commander window will generate the actual R code on the fly which can be saved, re-executed and even edited. Before, in other GUI systems for R, this was not yet possible. They did not provide any insight for the user about the R code behind the functionalities.

It is this concept which was the basis for the development of the `RcmdrPlugin.BiclustGUI` package. The goal of this package is to serve as a platform from which several biclustering algorithms as well as the appropriate diagnostics and graphs can be accessed.

3 The BiclustGUI R Package

3.1 Installing and Loading

3.1.1 R Commander

Like a lot of R packages, R Commander can simply be downloaded from CRAN at <http://cran.r-project.org/web/packages/Rcmdr/index.html> or simply by using `install.packages("Rcmdr")` in the R console. If you use this last command, R will automatically give you the option to install the packages required for R commander to work correctly.

After the installation, the R commander main window can simply be opened by using `library(Rcmdr)`. It could very well be that at first launch the GUI will still prompt to install some suggested packages. Note that if you close down the main window it can be reopened by using the `Commander()` command.

As can be seen in Figure 1, the window is separated in three parts: a *script window*, an *output window* and a *messages window*. In the first the generated R commands from R Commander (and plug-ins) will appear. Users are also able to edit, enter and re-execute commands from this window. The second window is simply what would normally appear in the R-console window and the third displays error messages and warnings.

More information on the use of R commander can be found in ‘*Getting Started With the R Commander*’ by J. Fox (2007).

3.1.2 RcmdrPlugin.BiclustGUI

The latest version of the package can be obtained by sending a mail to ewoud.detroyer@student.uhasselt.be. You will get a zip-file which contents will have to be placed inside the library folder of R.

Also note that RcmdrPlugin.BiclustGUI depends on several other biclustering packages which should also be installed. Some can be found on **CRAN**, others on **Bioconductor**.

- **CRAN:** biclust, BcDiag, superbiclust, isa2
- **Bioconductor:** iBBiG, fabia

Note that installing from Bioconductor is slightly different. First you need to use the `source("http://bioconductor.org/biocLite.R")`, followed by a `biocLite()` command. (e.g. `biocLite("fabia")`)

After succesfully installing the plug-in, there are two ways to load it: directly or through R Commander.

- Simply use `library(RcmdrPlugin.BiclustGUI)` in the R console.
- Launch R Commander first, then go to *Tools, Load Rcmdr plug-in(s)...* and simply select RcmdrPlugin.BiclustGUI in the newly opened window.

3.1.3 Data Input

There are a couple of ways on how to load data into the Rcmdr package:

- Enter *new* data directly with: *Data –> New data set ...*
- Load an existing data set from the *R workspace*: *Data –> Load data set ...*
- Import existing data from a plain-text file, other statistical packages (SPSS, SAS, Minitab, STATA) or even an Excel file: *Data –> Import data –> (choose required option)*
- Use a data set which is included in an R package: *Data –> Data in packages –> Read data set from an attached package ...*

Note that in most cases the biclustering procedures require a data matrix with the rows as genes and the columns as conditions/samples.

3.2 GUI Structure

The **general structure** of a biclustering window is always the same for each method. This is shown in Figure 2. The method dialog consists out of two tabs: the *biclustering* tab and the *plots & diagnostics* tab.

In the first tab, the user will be able choose the appropriate parameters for a biclustering algorithm and then click the *Show Results* button to execute it. If relevant, there will also be a small seed box present to set your seed for this run in order to obtain reproducible results (this will appear as `set.seed(...)` in the script). In the second tab, one can find the *specific* plots and diagnostic tools for the chosen method or package. However, from this second tab you are also able to access more general diagnostics and plots obtained from an external package. Two which are already implemented are *BcDiag* and *superbiclust*. They can be easily accessed through a button, provided they are compatible with the chosen method. Clicking one of these buttons will open up a new window entirely with additional plotting and diagnostic tools.

When these diagnostics are opened up from a specific biclustering window, the plots and diagnostics will automatically choose the right object to apply their functions to. This way no extra interference of the user is necessary when applying their functionality and the user will not have to worry about the communication between these windows.

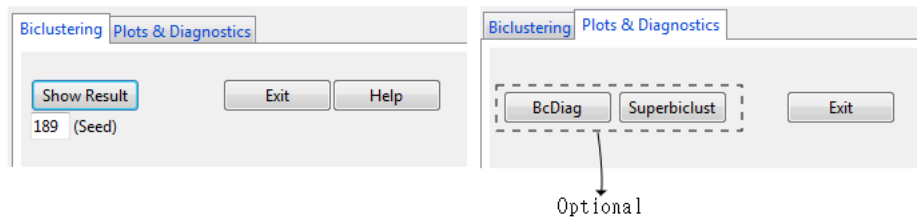


Figure 2: *Default Window Structure*

Note that when the graphs are produced they will appear in a separate graphics device window. Due to the variety of graphical parameters some diagnostic plots use, it might sometimes be necessary to close this graphics device down before utilizing a certain graph.(e.g. when you observe the size of a plot is considerably smaller and multiple are appearing in the same device). Further, if you would like to save a graph appearing in the device, simply select the device window and go to the working bar of R itself (not R Commander). Here, select *File, Save As* and then choose the desired extension (png, pdf,...).

Finally, each biclustering method can be accessed from the *Biclustering* menu (see Figure 3) in R commander which will appear after loading the plug-in.

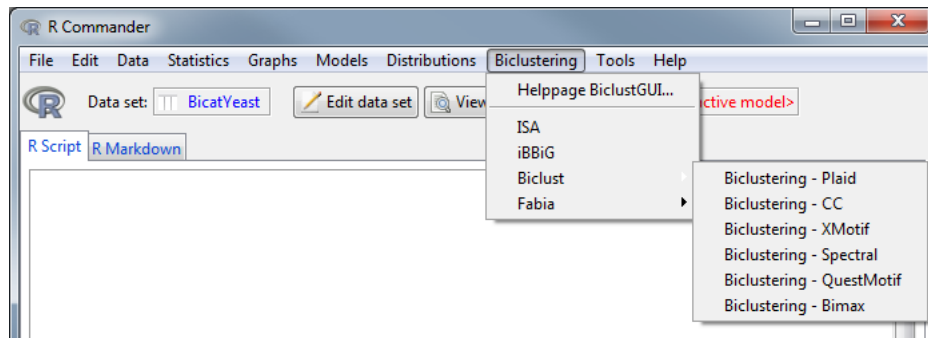


Figure 3: *Biclustering Menu (Biclust submenu)*

The implemented algorithms and diagnostic tools can be found in Table 1.

Type	Package	Method/Description
<i>Biclustering Algorithms</i>	biclust	Plaid
		CC
		XMotif
		Spectral
		QuestMotif
		Bimax
	fabia	Laplace Prior
		Post-Projection
		Sparseness Projection
		SPARSE
	isa2	The Iterative Signature Algorithm
	iBBiG	Iterative Binary Biclustering of Genesets
<i>General Plots/Diagnostics</i>	BcDiag	Bicluster Diagnostics Plots
	superbiclust	Generating Robust Biclusters from a Bicluster Set

Table 1: *Table of implemented packages*

3.3 biclust-package

3.3.1 Plaid Biclustering

The plaid biclustering in the GUI implements the plaid algorithm by Turner *et al.* (2005) which was proposed as an improvement of the plaid model discussed by Lazzeroni and Owen (2000). The plaid model is a biclustering method which takes the interactions between biclusters into consideration by defining the data structure (e.g. expression level) as a sum of layers. This model includes a background layer to capture the global effects and afterwards the method will construct a series of layers that represent the biclusters.

Plaid Model:

$$Y_{mn} = \theta_{mn0} + \sum_{p=1}^P \theta_{mnp} \gamma_{mp} \eta_{np} + \epsilon_{mn} \quad (1)$$

$$\gamma_{mp} = \begin{cases} 1 & m \in p \\ 0 & \text{otherwise} \end{cases}, \text{ and } \eta_{np} = \begin{cases} 1 & n \in p \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

$$\theta_{mnp} = \begin{cases} \mu_p & (3.1) \\ \mu_p + \alpha_{mp} & (3.2) \\ \mu_p + \beta_{np} & (3.3) \\ \mu_p + \alpha_{mp} + \beta_{np} & (3.4) \end{cases} \quad (3)$$

In (1), Y_{mn} is the expression level of gene m in condition n with $m = 1, \dots, M$ and $n = 1, \dots, N$. Further, p is the layer index, P is the number of biclusters, θ_{mn0} is a sum of overall mean and ϵ_{mn} is a random error with mean zero. The model also contains two indicator variables, γ_{mp} and η_{np} which represent the membership of the gene/condition in a bicluster p as formulated in (2). Finally θ_{mnp} is the mean gene expression which can take four possible forms as shown in (3). In this formula, (3.1) implies a constant bicluster whereas (3.2) and (3.3) respectively imply biclusters with constant rows or columns. The last one, (3.4) implies a bicluster with coherent values across the genes and conditions in a bicluster.

The estimation of the plaid method is done with an iterative algorithm. First the background layer is fitted, then the bicluster-specific layers are added one at a time. In each iteration the algorithm will estimate the parameters with binary least squares after which a permutation test is performed which is a built-in protection against the discovery of random biclusters. This procedure is repeated until no layer is found anymore or until the maximum amount of layers as been reached. More detailed information about these steps can be found in Lazzeroni and Owen (2000) and Turner *et al.* (2005).

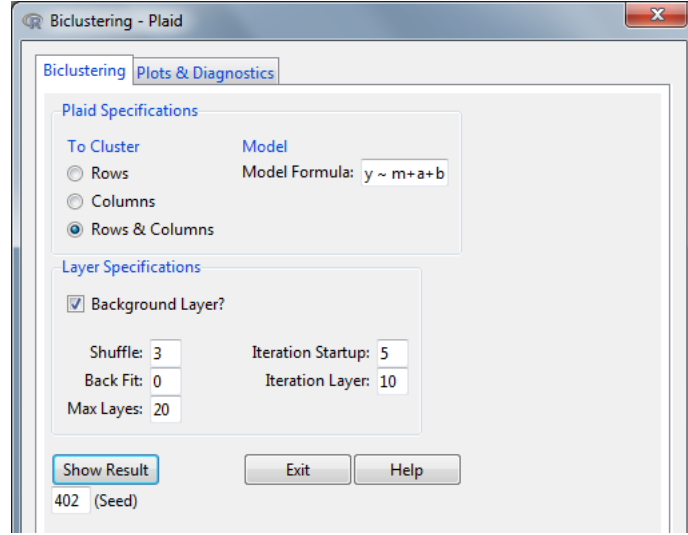


Figure 4: *Plaid Window*

In Figure 4, the standard *Plaid window* can be found which contains all the necessary parameters to apply plaid biclustering. First the user is able to decide they only want to cluster the rows and columns, or on both (recommended). Next the model which is fitted to each layer can be specified in the *Model Box*. This coincides with formula (3) and is defaulted to $y \sim m + a + b$ (m =constant for all elements in bicluster; a =constant for all rows in bicluster; b =constant for all columns in bicluster).

The remaining parameters are there to further specify the layer options in the algorithm. The *background* check specifies if there is a background layer present in the data. *Iteration startup* and *layer* define the number of iterations to find respectively the starting values and each layer. And as already explained earlier, *max layers* will determine the maximum number of layers to include in the model and therefore the maximum number of biclusters.

Finally, *Back Fit* specifies the additional iterations to refine the fitting of the layer and *Shuffle* is a parameter connected with the permutation test. Before a layer is added, its statistical significance is compared against a number of random obtained layers, defined by this parameter.

3.3.2 CC(δ) Biclustering

The δ biclustering, also known as CC algorithm is based on the framework by Cheng and Church (2000). The algorithm discovers biclusters one at time and considers a bicluster as a subset of rows and columns that show coherent values. The method is a combination of data analysis based on an *ANOVA model* and a *node deletion algorithm*.

Let \mathbf{A}_{IJ} be a submatrix, i.e. a δ -bicluster, in the data matrix \mathbf{A} ($I = (i_1, \dots, i_k); J = (j_1, \dots, j_k)$). Note that a_{ij} is the expression level of gene i in condition j .

Now Cheng and Church defined a *mean residual score* (MSR) as follows

$$H_{IJ} = \frac{1}{|I||J|} \sum_{i \in I, j \in J} r_{ij}^2$$

where $r_{ij} = a_{ij} - a_{iJ} - a_{IJ} + a_{IJ}$, $i \in I, j \in J$.

A submatrix is now called a bicluster if the MSR is less than a pre-defined threshold δ .

In order to find these δ -biclusters, the algorithm will start with the full matrix and calculates the MSR. Now the MSR will be minimized by deleting/including rows and columns in the matrix. Since the *brute-force approach* is computationally not time-efficient, *node deletion* algorithms were developed (single & multiple) which is a *greedy* algorithm. These algorithms iterate the process of choosing a row and column with the largest MSR and removing them from the data matrix until the desired submatrix is found. Note that after the node deletion, the bicluster may not be maximal (some rows/columns may be added without increasing the MSR) so therefore node addition is performed by again adding rows and columns one by one (if it does not increase the MSR).

More about both algorithms and the CC biclustering can be found in Cheng and Church (2000).

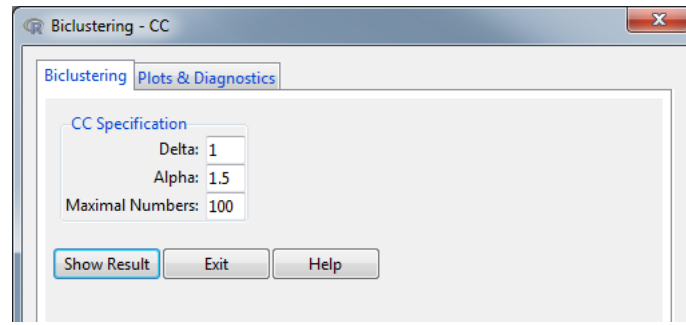


Figure 5: *CC Window*

As can be seen in Figure 5, there are not a lot of input parameters in the *CC window* for the δ -biclustering. The most important parameter is of course *Delta*, the maximal accepted score which will be compared with the MSR. The choice of this variable should depend on the total variability of the data, taking into account both the assumed variability of the noise and bicluster values. Next, *Alpha* is a scaling factor. It is a parameter for the multiple node deletion and takes part in the three major steps of the algorithm:

1. Deleting rows and columns with a score larger than *Alpha* times the matrix score.
2. Deleting the rows and columns with the largest scores.
3. Adding rows or columns until *Alpha* level is reached.

Finally, the user is also able to set a maximum number of clusters to be found with the *Maximal Numbers* input parameter. The algorithm will stop until no bicluster is found or until this threshold is reached.

3.3.3 XMotifs Biclustering

The XMotifs biclustering algorithm was proposed by Murali and Kasif (2003) and looks for conserved gene expression motifs in a discretized version of the data matrix. This is achieved by searching for rows with constant values over a set of columns. The authors assume in their model that the gene can be expressed in a finite number of states (e.g. 2 states: up- and downregulated). The states of the gene expression matrix can also be defined by a fold change, represented by quantile discretization of the original matrix with log-transformed values. A conserved gene expression motif is now defined as a submatrix of *maximum size* for which the values within each row are equal to the same level.

Following the discretization, the biclusters are discovered with an iterative procedure which will be briefly touched upon down below while explaining the input parameters of the window. More detailed information can be found in Murali and Kasif (2003).

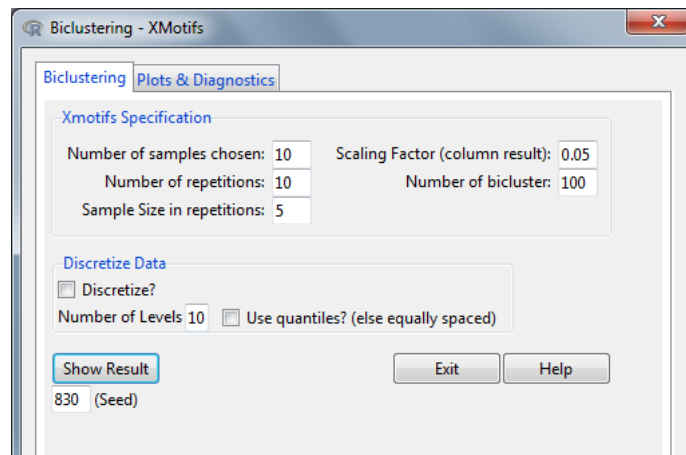


Figure 6: *XMotifs Window*

In Figure 6, the *Xmotifs window* is displayed which contains all the necessary input parameters for the algorithm. Note that in the bottom box, the user is able to discretize the active data set in the R commander session if necessary. This is done with the `discretize` function in the `biclust` package. For plotting however, the original matrix will be used.

After the discretization, the algorithm first needs a number of samples/columns (= ‘*Number of samples chosen*’) to be randomly selected as a seed. Then a number of sets of samples (= ‘*Number of repetitions*’) of a defined size (= ‘*Sample size in repetitions*’) needs to be inputted which will be randomly selected from the samples that were not in the *number of samples chosen*.

After this is done, the following steps are performed as described in Kaiser and Leisch (2008):

1. Choose a subset from these columns and collect all rows with equal state in this subset.
2. Collect all columns where these rows have the same state.
3. Return the bicluster if it has the most rows from all found and is also larger than a α (= ‘*Scaling factor*’) fraction of the data.

To collect more than one bicluster, the calculation is rerun without the rows and columns already found (or return the smaller combinations found). This is done until the maximum number of biclusters (= ‘*Number of biclusters*’) is achieved or until no clusters can be found anymore.

3.3.4 Spectral Biclustering

Spectral biclustering is a method developed by Kluger *et al.* (2003), used for discovering multiplicative biclusters of coherent values. First of all it assumes that the gene expression matrix assumes a checkerboard structure after normalization therefore resulting in non-overlapping biclusters. These biclusters are called multiplicative because each bicluster element (a_{ij} , expression level gene i , condition j) can be defined as a product of three terms: overall mean (μ), row-specific (α_i) and column-specific (β_j) means ($\alpha_{ij} = \mu \times \alpha_i \times \beta_j$).

The spectral biclustering is mostly based on a singular value decomposition (SVD) of the normalized data matrix and consists out of the following steps (Kaiser and Leisch, 2008):

1. Re-order the data matrix and apply one of the three normalization methods (independent rescaling, bistochasticization or log-interactions).
2. Obtaining eigenvalues and eigenvectors using SVD.
3. Depending on the normalization method, the biclusters are constructed beginning from the largest or second largest eigenvalue. The eigenvectors (left & right) corresponding to the largest eigenvalues are expected to provide optimal clustering of rows and columns.
4. The data is projected on the the best two or three eigenvectors and k -means clustering is run to get the grouping.

More detailed descriptions of the above steps can be read in Kluger *et al.* (2003).

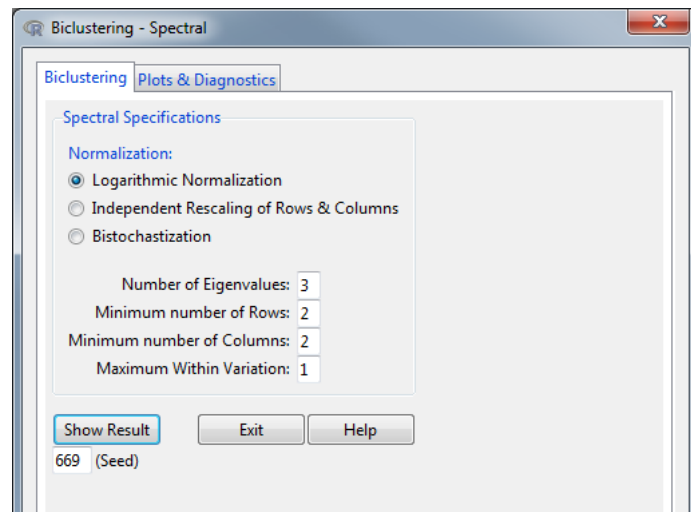


Figure 7: *Spectral Window*

In Figure 7 above, the *Spectral window* is shown which contains the major steps of the spectral biclustering, namely the normalization and the SVD.

As already explained earlier, the data matrix need to be normalized first as this is necessary for a checkboard structure to be discovered by the use of Singular Value Decomposition. The user is able to choose out of three

different options for this step. The *Independent Rescaling of Rows & Columns* assumes the non-normalized matrix is obtained by multiplying each row and column with a scalar. *Bistochastization* works by repeating the independent scaling of rows and columns until stability is reached. For this normalization, the final matrix has all rows sum to a constant and all columns sum to a different constant. The final method, *Logarithmic Normalization* (= log-interactions) assumes that after taking the logarithm, the original rows/columns differ by additive constants. Further each row and column is expected to have mean zero which is achieved with a transformation.

Finally, the user will have to specify the input parameters connected with the SVD. These include: *Number of Eigenvalues*, *Minimum number of Rows*, *Minimum number of columns*, *Maximum Within Variation*. Note the number of eigenvalues coincides with the number of biclusters that should be discovered.

3.3.5 QuestMotif Biclustering

The Questmotif Biclustering is based on the framework by Murali and Kasif (2003) and developed by Kaiser (2012). The algorithm will search for biclusters of questioners which have similar answer to the questions.

Following now is a short description of the algorithm with respect to the *Questmotif window* in Figure 8.

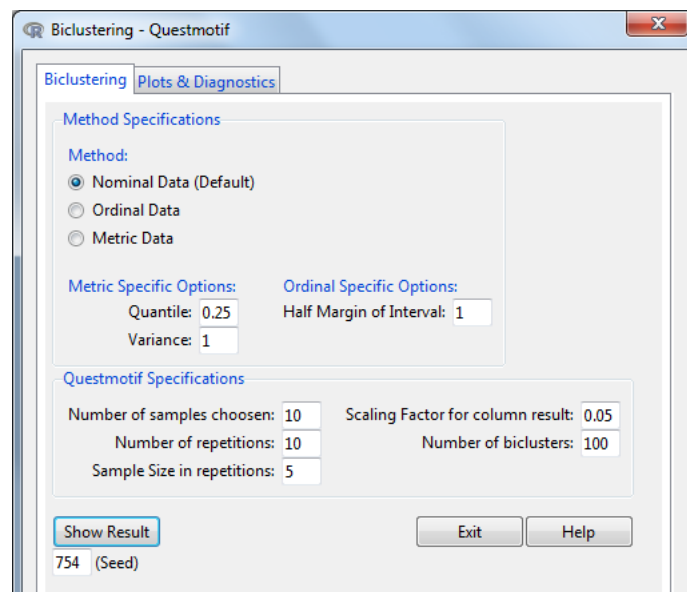


Figure 8: *QuestMotif Window*

The Quest algorithm contains three methods to deal with different scale levels data, especially for biclustering questionnaire data. All of the three methods can be selected by the user and depending on the choice, some additional parameters might have to be defined.

If the answers are given on a *nominal scale*, the algorithm simply works like the Xmotifs algorithm. For the *ordinal scale*, the algorithm will search for similar answers in an interval of a size set by the parameter '*Half Margin of Interval*' (= d). This implies that the interval will be of the form $[mean - d, mean + d]$. In the continuous case, namely *metric data*, this interval is set by the quantile of a chosen distribution. It uses all previously found values to calculate the mean value and uses a given parameter for the variance of this distribution. Both the *Quantile* and *Variance* will have to be provided by the user in this case. (Since the normal scores are used in such data, the normal distribution is commonly used in this case.)

Finally, a couple of general input parameters, used by the algorithm, will have to be set by user in the *Questmotif Specifications* box. These include: *number of samples choosen*, *number of repetitions*, *sample size in repetitions*, *scaling factor for column result* and *number of biclusters*. Note that these are the same input parameters used by the Xmotif algorithm.

More insight and details about the algorithm itself can be found in Kaiser, S. (2012).

3.3.6 Bimax Biclustering

The last implemented biclustering method from the `biclust` package is the *Bimax* (= binary inclusion-maximal) biclustering algorithm which was developed by Prelić *et al.* (2006). They advocated its use as a preprocessing

step to identify potentially relevant biclusters that can be used as input for other methods. According to the authors, the main benefit of the method is the relatively small computation time while still providing relevant biclusters using a simple data model.

The Bimax algorithm works on a binarized data matrix in which the expression value is set to 1 if there is a change with respect to the control setting and to 0 otherwise. If a control setting is unavailable, one can simply take a threshold based on the distribution of the data values. The goal of the Bimax method is to find maximal inclusion biclusters. This means a Bimax bicluster spans a submatrix of 1's which cannot be part of a larger submatrix of 1's. The algorithm achieves this by applying a divide-and-conquer strategy in which the rows and columns are rearranged (to concentrate ones in the upper right corner of the matrix) before dividing the matrix into two submatrices.

More detailed information about the algorithm can be found in Prelić *et al.* (2006).

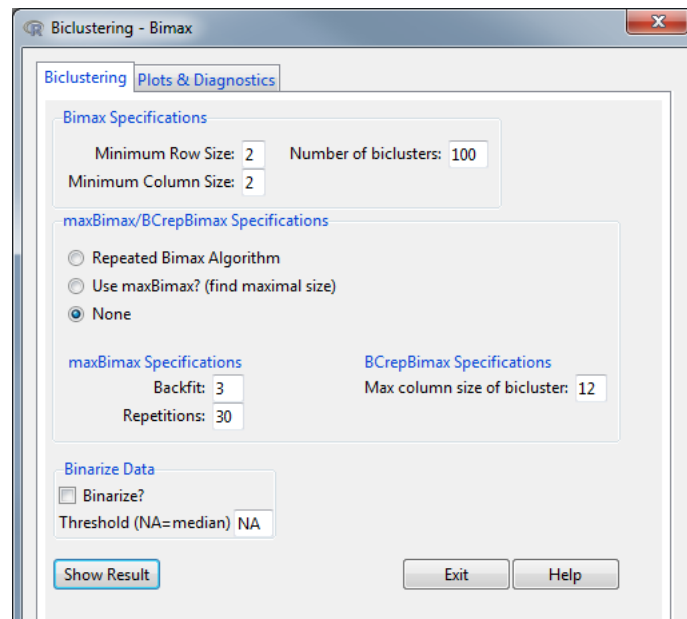


Figure 9: *Bimax Window*

The first box in the *Bimax window*, shown in figure 9, are general input parameters required by the algorithm. These include the *minimum row & column size* of a bicluster (for it to be included) and the *number of biclusters*. The algorithm will terminate when this boundary level of number of biclusters is achieved or when no more maximal inclusion biclusters can be found.

Further, Kaiser and Leisch (2008) suggested that in order to get satisfying results, the method should be iterated several times with different starting points. This can be accomplished with the ‘*Use maxBimax*’ option in the second box for which the *number of repetitions* and *backfit* should be defined. With this option biclusters of maximal size will be discovered.

Finally, also the *Repeated Bimax Algorithm* is implemented in this window for which the *max column size of biclusters* should be inputted. More information about this last variation of Bimax can be found in Dolnicar *et al.* (2012).

Finally in the last box, the active data set in your R Commander session can be binarized if this is not a binary matrix by default. This is done through a available function `binarize` in the `biclust` package. A threshold can be set for this transformation or it can be left on the default option, namely the median. Do note that when plotting graphs (e.g. the biclust plots), the original data matrix will be used which is the active data set in your session.

3.3.7 Biclust: Plots & Diagnostics

After executing any of the biclustering methods available in the `biclust` package, the user is now able to proceed to the second tab ‘*Plots & Diagnostics*’ which is identical for all of these methods. No additional work is needed by the user for the plotting and diagnostic functions to use the correct results from the method which was applied in the first tab. An example (for the Plaid Biclustering) of the *Plots & Diagnostics window* is given in Figure 10. Note that the second window in this figure is accessible through the *Extra Biclust Plots* button.

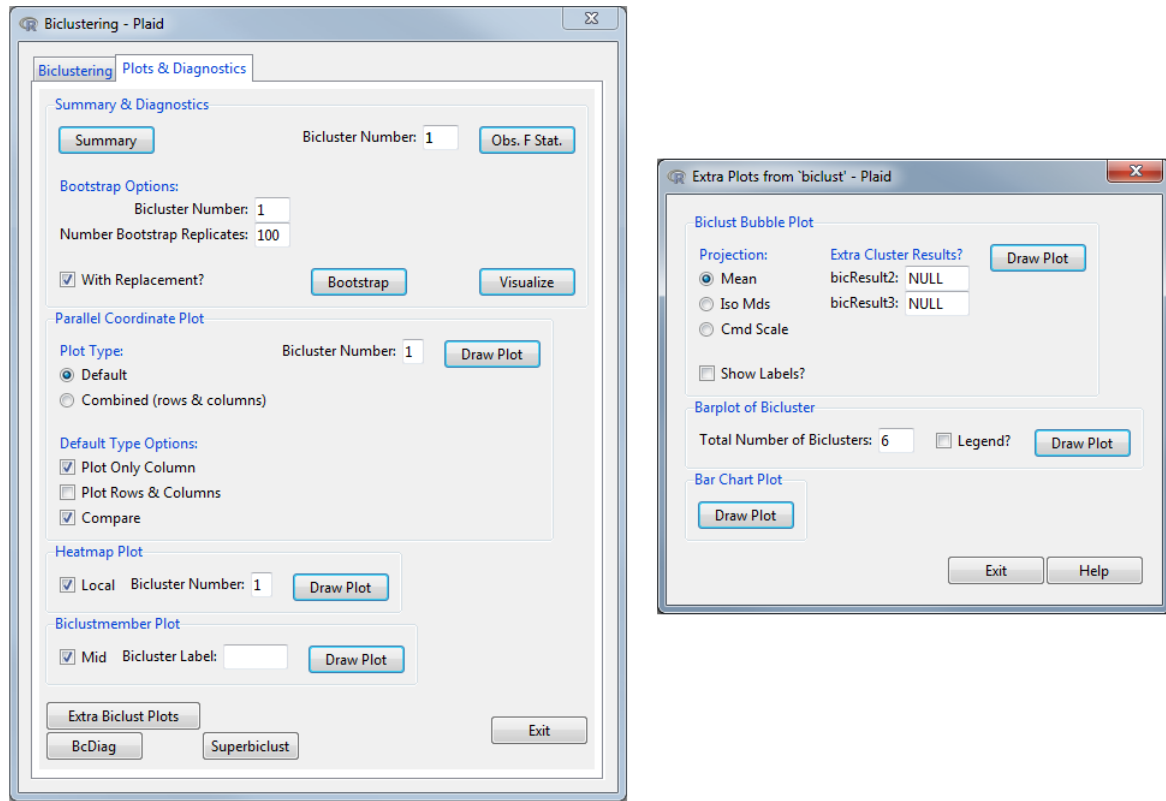


Figure 10: *Biclust - Plots & Diagnostics Tab*

Summary & Diagnostics

In the first box, the summary of the biclustering result as well as some basic diagnostics can be called by the user. The *Summary* button will provide the user with the number of clusters found as well with the number of rows and columns each bicluster contains. Next, the *Obs. F Stat.* button will enable the user to compute some F-statistics about a specific bicluster, defined by the number in the *Cluster Number* entry box. These F-statistics include the main effects, namely the row and column, and also the interaction effect. The first two F-statistics (main effects) are calculated from a two-way ANOVA with such row and column effect. Because the full model with interaction is unidentifiable, Tukey's test for non-additivity is used to detect an interaction within a bicluster. Lastly, the p-values are obtained from asymptotic F distributions.

The last option in this first box, is the ability to calculate p-values of earlier F-statistics but now with the help of bootstrapping. This means the p-values are now calculated by taking the number of bootstrap replicates which are larger than the observed F-statistics and divide it by the total number of replicates plus one. This number can be set by the *Number Bootstrap Replicates* entry box and the user can also decide to bootstrap with or without replacement with the corresponding checkbox. The bootstrapping on the defined *Cluster Number* is executed through the *Bootstrap* button after which the user can use the *Visualize* button to obtain histograms of the bootstrap results of the row and column tests. An example of this graph (and bootstrapping output) can be found in the Appendix in figure 39. Note that the vertical green line represents the observed F-statistic.

Parallel Coordinate Plot

In the next box the user is able to produce a variety of parallel coordinate plot. The graphs make use of the expression levels in the original data matrix (unaffected by discretization or binarization). The plot represents these levels through gene and/or condition profiles in a bicluster as lines. The bicluster which will be drawn is once again defined by a *Cluster Number* entry box.

The first *Plot Type* is the default one for which extra options are available in the form of checkboxes. By checking only the *Plot Only Column*, the expression levels for the columns in the selected bicluster will be drawn. This means each line is a column profile with the genes on the x-axis. By checking *Plot Rows & Columns* a second plot will be added to the graphics device, but now with the row/gene profiles with the columns/conditions on the x-axis. The last checkbox, *Compare*, will make sure the other profiles, not in the chosen bicluster, are also plotted in a light-gray colour.

Finally the second type, *Combined (rows & columns)*, in a way combines the information about the bicluster grouping of rows and columns. Again, each line is a gene/row profile, but the columns/conditions on the x-axis are reordered in such a way that the columns belonging to the bicluster come first. This is visualized by showing a red line when belonging to the selected bicluster and black one when not.

An example of these parallel coordinate plots (or profile plots), can be found in the Appendix in Figure 40.

Heatmap Plot

In this box, the user can visualize the gene expression data matrix as a heatmap. The rows and columns will be reordered so that the inputted bicluster in the *Cluster Number* entry box will appear in the top-left of the matrix. However, by checking the *Local* option, only the heatmap of the bicluster will be drawn, omitting the rest of the matrix.

An example is given in the Appendix in Figure 41.

Biclustmember Plot

The last box on this window contains the options to draw a Biclustmember Plot. This plot can primarily be used to compare the discovered biclusters against each other. On this graph, as given in the Appendix in figure 42, one can find multiple columns of stacked rectangles. Each such column is a representation of a bicluster and each rectangle inside represents a column/condition/sample of the data matrix. Basically if in a column of these stacked rectangles, a rectangle is coloured, it means that this condition is part of that particular bicluster. Now, if the *Mid* box is not checked, a coloured rectangle consists out of two parts, left and right. The left colour represents the mean of this condition for all the genes within the biclusters. However, the right colour contains the global mean value for this condition. If the *Mid* option is checked though, the rectangle exist out of three colours with the global mean in the middle and the bicluster mean on the left and right.

Finally, the user is also able to set a label which will come in front of the cluster number with the *Cluster Label* entry box.

Biclust Bubble Plot

The first plot the user can create from the extra window, is the Biclust Bubble Plot. The bubbleplot is a 2D-projection of the biclusters, done through multidimensional scaling based on the gene and condition profiles. It is used as a tool to help understand the overall behaviour of biclustering methods, detect trends, outliers, etc. Each bicluster is represented as a circle of which the brightness represents the homogeneity (darker, less homogeneous). The size on the other hand represents the size of the biclusters, as rows \times columns. The user is able to add up to three bicluster results in the *Extra Cluster Results* entry boxes, obtained from earlier runs of methods from **biclust**. Note that each bicluster set will get a different colour in the plot.

Further, the user is also able to choose between three different kind of projections, namely *mean*, *Iso Mds* and *CmdScale* of which more information can be found in the help files (note that this *Help* button is linked to the bubbleplot help page).

Lastly the *Show Labels* checkbox will give each bicluster the corresponding bicluster number if checked. An example of this type of graph can be found in the Appendix in Figure 43.

Barplot of Bicluster

The graph available in this box is a barplot of biclusters which is used to compare the values inside a bicluster with the values outside of the bicluster. For each bicluster, three bars are drawn per column/condition part of the bicluster. The darkest represents the values inside of the bicluster and the other two the mean and median of the values outside of the cluster. The user is able to draw a legend yes or no with the *Legend* checkbox and can also determine the number of biclusters which should be drawn with the *Total Number of Biclusters* entry box. However this works in a slightly different way, namely on each graphics output device, only 6 bicluster barplots can fit. Further the spot such barplot will get on this 'grid' will always be the same. For example, if you would take as input the number 14, only bicluster 13 and 14 would appear on the device. If you would put in the number 12, bicluster 7 to 12 would appear.

An example is given in the Appendix in Figure 44.

Bar Chart Plot

The final graph in this window will create a barchart for all the biclusters, representing the columns. Each block represents one bicluster and the bars inside of it represent the means of bicluster values for the corresponding column.

An example of this is also given in the Appendix in figure 45.

3.4 fabia-package

The biclustering algorithm FABIA or Factor Analysis for Bicluster Acquisition was proposed by Hochreiter *et al.* (2010). A couple of variations are available in the **fabia** packages of which the *Laplace Prior*, *Post-Projection*, *Sparseness Projection* and *SPARSE* are implemented in the GUI.

The main description of fabia will be given for the Laplace prior implementation. For the others, the differences between the windows will be briefly touched upon.

3.4.1 Laplace Prior

Factor Analysis for Bicluster Acquisition with Laplace Prior is an algorithm based on factor analysis where the homogeneity is based on the latent relationship between the variables in the data. The method will discover multiplicative biclusters which are found by sparse factor analysis where both the factors and loadings are sparse. This assumption of sparseness comes from the gene expression data, where normally only a small fraction of the genes is active under a small subset of conditions (Khamiakova, 2009). Further, the model assumes non-Gaussian signal distributions with heavy tails.

Now, a factor model for data matrix Y with P factors can be described as follows

$$Y = \sum_{p=1}^P \lambda_p Z_p + \epsilon$$

where Z_p is the p th factor, λ_p is the vector of factor loadings for Z_p and where additive random noise is assumed to be normally distributed, $\epsilon \sim N(0, \Psi)$. Furthermore, the model assumes that Ψ is a diagonal matrix, i.e. the error terms ϵ are independently and normally distributed given the p factors in the model. Another assumption is that Z and Ψ are independent which implies that the noise is independent of the signal strength. Lastly, as said before, the factor model assumes sparseness of factors and their loadings and this is reflected by the choice of the corresponding prior on loadings and factors (i.e. a Laplace Distribution).

Using this factor model, biclusters can be obtained as following. On a side note, it is important to mention that the method works on normalized data. Now, the algorithm will estimate the parameters and factors through maximizing the posterior of λ , Ψ and Z . Using the estimates for Λ and Z , the denoised data is obtained and the biclusters are derived from $\lambda_p Z_p$. This means that this component can be seen as a bicluster of which the non-zero genes and samples are members of the bicluster.

More intricate details and information about the method and its variations, can be found in Hochreiter *et al.* (2010 and 2014).

The *Fabia with Laplace Prior window* is shown in figure 11 in which the fabia specifications as well as some data manipulation can be decided upon.

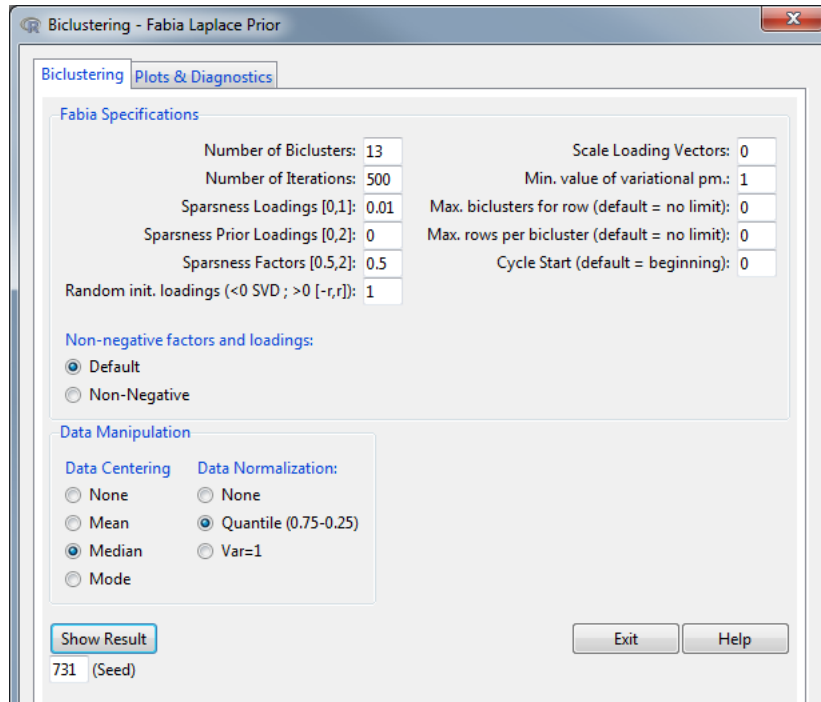


Figure 11: *FABIA (Laplace Prior) Window - Clustering Tab*

As one can see there are a lot of input parameters which can be specified. The most important ones are the *Number of Biclusters* which is equal to the number of factors and can be set to the upper boundary. The other parameters of importance are responsible for the specification of the sparseness such as the *Loadings*, *Prior Loadings*, *Factors*, etc. These depend on the noise level in the data and also the size of the data set.

The user can also select if the factors and loadings are non-negative or not.

Finally, the user can also apply some data manipulation before the method is executed, namely *Centering* and *Normalizing* the data matrix.

3.4.2 Post-Projection

For *Fabia Post-Projection*, some post-processing is present. Namely the final results of the loadings and the factors are projected to a sparse vector according to Hoyer (2004). This means: given an l_1 -norm and an l_2 -norm, minimize the Euclidean distance to the original vector (currently the l_2 -norm is fixed to 1). The *Fabia Post-Projection window* is given in figure 12.

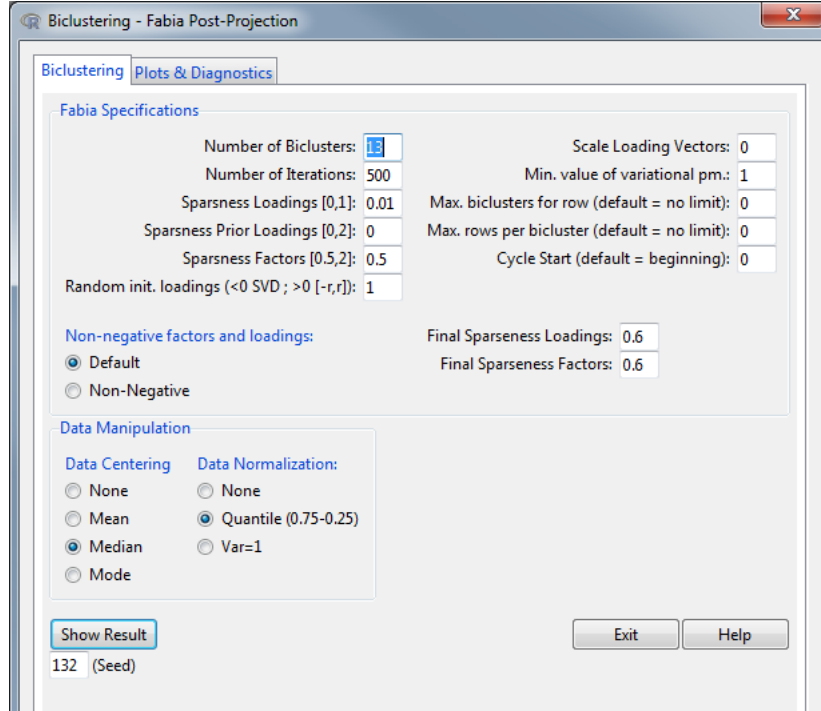


Figure 12: *FABIA (Post-Projection) Window - Clustering Tab*

As one can see, the window is primarily the same as the *Laplace Prior* one, but the user can now also define the *Final Sparseness Loadings* and *Final Sparseness Factors*.

3.4.3 Sparseness Projection

The next implemented fabia algorithm is fabia with *Sparseness Projection*. In this version, the prior has finite support, therefore after each update of the loadings they are projected to this finite support. This projection is again done according to Hoyer (2004) (See *Post-Projection*).

Figure 13 shows the *Fabia Sparseness Projection window*.

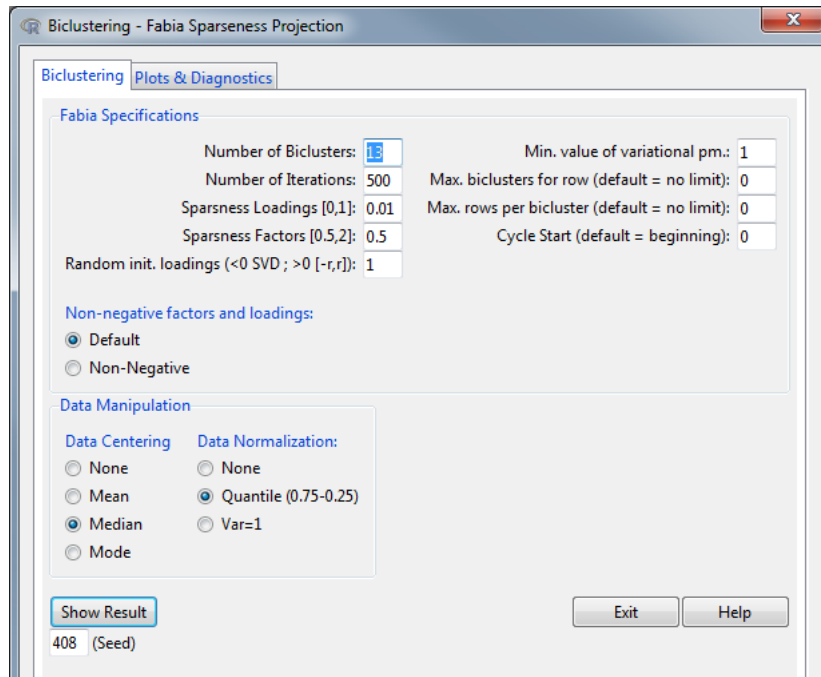


Figure 13: *FABIA (Sparseness Projection) Window - Clustering Tab*

The only difference with *Laplace Prior* is that some options in the specifications have disappeared. These include the *Sparseness Prior Loadings* and the *Scale Loading Vectors*.

3.4.4 SPARSE

This is a version of *fabia* for a sparse data matrix. The matrix is directly scanned by C-code and therefore must be in sparse matrix format as described in Hochreiter *et al.* (2010) and Hochreiter (2014).

Again biclusters are discovered through sparse factor analysis and the model selection is performed by a variational approach according to Girolami (2001) and Palmer *et al.* (2006). Further a prior on the parameters is included and a lower bound on the posterior of the parameters is minimized, given the data. The update of the loadings includes an additive term which pushes the loadings towards zero (Gaussian prior leads to a multiplicate factor).

More detailed information about this algorithm and its methodology can be found in Hochreiter *et al.* (2010) and Hochreiter (2014). The *Fabia SPARSE window* is shown in Figure 14.

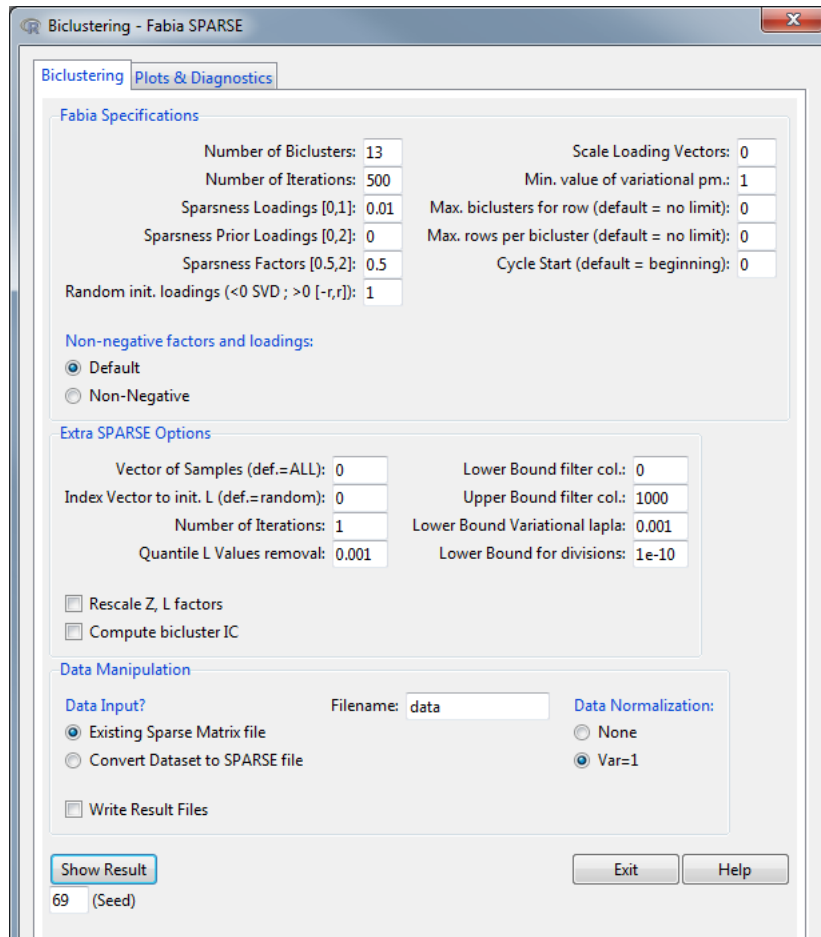


Figure 14: *FABIA (SPARSE) Window - Clustering Tab*

Once again, the same specifications as for the *Laplace Prior* are available in this window. However the options for *Data Centering* and *Quantile Normalization* are not available anymore.

In the second box, *Extra SPARSE Options*, more specifications are given, specific for the SPARSE algorithm. More information about these can be found in the reference manual and vignette of **fabia** (Hochreiter, 2014).

As explained earlier, the algorithm requires the data matrix in a special sparse format so therefore, the data input for this method works slightly different. Namely the function requires the data to be located in a plain text file in said format. There are two options of inputting the data into this algorithm which will now be explained.

The first option assumes that the user already has this text file which contains the data matrix in sparse format. In this case the *Existing Sparse Matrix file* option should be used and the *Filename* entry box should contain the name of this file (without the extension). Upon pressing the *Show Result* button, the GUI will prompt the user with a directory window in which they have to select the folder where this file is located.

The second option, *Convert Dataset to SPARSE file*, will -as the option states- transform the active dataset in the R commander session to a data matrix in sparse format. With this option a plain text file will be generated, containing the data matrix in sparse format with the name inputted in the *Filename* entry box. Now upon pressing the *Show Result* button, the user chooses the folder where this file will be saved. Note that the code behind this transformation was based on example R-code, available in the **fabia** vignette (Hochreiter 2014).

Finally, the user can also check the *Write Result Files* which will enable the results being saved in the chosen folder location in the form of plain text files.

3.4.5 Fabia: Plots & Diagnostics

Similar to the '*Plots & Diagnostics*' of **biclust**, also for **fabia** the second tab is the same for all of the methods implemented from this package. The idea is to proceed to this tab after executing the biclustering method in the first one so that the biclustering result object is available for the plotting functions. The structure of this *Plots & Diagnostics window* is given in Figure 15 which, in this example, is part of the Laplace Prior dialog.



Figure 15: *Fabia Window - Plots & Diagnostics*

Summary & Summary Plots

By clicking the *Summary* button, the user will receive some general information about the *fabia* result. Information such as the number of rows and columns in the data matrix, the number of clusters, the information content of the biclusters & samples and finally some summary statistics of the column clusters/factors and row clusters/loadings, will be printed in the output window.

Further in this first box, the user can also plot several graphs containing the information which was outputted by the *Summary* button. The first two are histograms containing the *Information Content of Biclusters* and *Information Content of Samples*. The other two options are boxplots of the *Loadings of the Biclusters* and the *Factors of the Biclusters*. Examples of these four types of graphs can be found in the Appendix in figures 46 and 47.

More information about the calculation of the information content can be found in the *fabia* vignette (Hochreiter, 2014). Basically it is a way to rank the discovered biclusters analogously to principal components which are ranked according to the data variance they explain. In this case the biclusters are ranked according to the information they contain about the data.

Extract, Extract Plot & Bicluster Plot

First of all, in the top of this box, the user is able to extract the bicluster results from the *fabia* results object so it can be used for other applications. This is achieved by using the *Extract* button. Before doing this, 2 thresholds should be set, namely the '*Threshold bicluster sample*' which is a threshold determining when a sample belongs to a bicluster and the '*Threshold bicluster loading*' which is, unsurprisingly, the threshold determining when a loading (=row) belongs to a bicluster. Note that by default this last one is estimated. The result of this extraction is a list object which contains the following items: the extracted biclusters and their indices, the extracted opposite biclusters and their indices, the scaled and centered data matrix and lastly the number of biclusters. (Opposite means that the negative pattern is present.)

The next option in this box is the ability to extract more useful plots from the *fabia* result. Note that again the earlier mentioned threshold for samples and loadings apply to these graphs as well. The user is able to pick out of a multiple of plots, namely the *Data*, *Reconstructed Data*, *Error*, *Absolute Factors*, *Absolute Factors* and (if available) the *Noise Free data*. For the latter the object name of the noise free matrix should be entered in the *Noise Free Data Matrix* entry box. Interesting to know is that to achieve sorting, k-means is performed so that the vectors belonging to the same cluster can be put together. However, in general this sorting is not able to visualize all biclusters as blocks, namely if they overlap. Several examples of these plots are given in Figure 48 in the Appendix.

The final plot in this particular box is the *Bicluster Plot* which is basically the same as the heatmap available in *biclust*. The user is able to choose a specific bicluster in the *Cluster Number* entry box and then draw only the bicluster (*Only Bicluster*) or the entire data matrix with the bicluster in the top-left (*Data Matrix (with bicluster)*). Further the user can also decide to plot the opposite bicluster with the *Opposite Bicluster* checkbox. Note that the bicluster plot will *only* work if the extraction in the first part of this box is executed as it requires this output to draw the biclusters. An example of both graphs can be found in figure 49 in the Appendix.

BiPlot

The final implemented plot in the last box is the *BiPlot* for the matrix factorization result. The user will have to specify the *Principal Factors* that are plotted along the horizontal (*First*) and vertical (*Second*) axis. On this biplot, the row-items/genes are represented as circles with their areas proportional to the row weighting. The most informative genes are those that are the most distant from the center of the plot. Note that the column-items are represented by squares. Also for this last plot type, an example is available in the Appendix in Figure 50.

3.5 isa2-package

The ISA biclustering or *Iterative Signature Algorithm* (Bergmann *et al.*, 2003) is a semi-supervised method, designed to decompose a large set of data into *modules*. These modules or biclusters consist of subsets of genes (rows) that exhibit a coherent expression profile only over a subset of samples/conditions/experiments (columns). ISA does allow for overlapping modules (rows and columns belonging to multiple modules) and it is developed to find biclusters that have correlated rows and columns which are identified through an iterative procedure. A standard ISA procedure starts with normalizing the data first and then generates random input seeds which correspond to some set of genes or samples. This is refined at each iteration by adding and/or removing genes and/or samples until the process converges to a stable set, the *transcription module*. Following now will be the explanation of the *ISA window* in Figure 16 with some more basic information about the important parameters of the algorithm. More detailed information can however be found in Bergmann *et al.* (2003) and the *isa2* documentation.

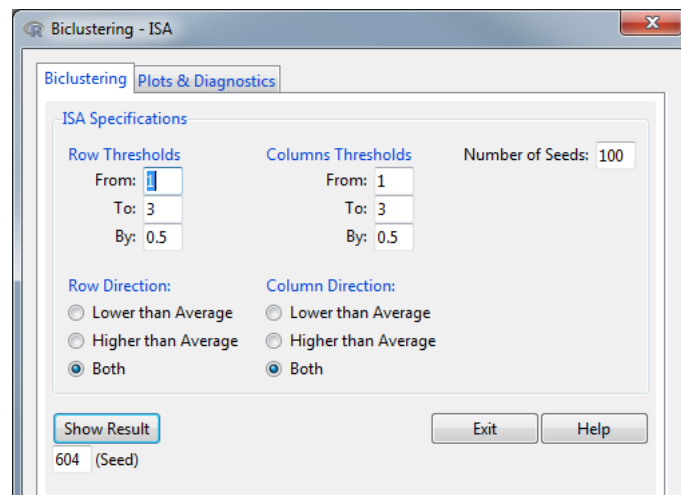


Figure 16: *ISA Window - Clustering Tab*

The two main parameters of ISA are the two thresholds (for rows & columns). If the row threshold is high, then the modules will have very similar rows, if the threshold is mild, the modules will be bigger but with less similar rows (analogous for columns). The user is able to set a sequence of thresholds for both the rows and the columns (default: $c(1, 1.5, 2, 2.5, 3)$) and ISA will run over all combinations of these sequences. For each threshold combination the similar modules will be merged and as a last step again similar modules will be

merged but now across all threshold combinations. In this last step, if two modules are similar, then the larger one (with milder thresholds) is kept.

Another interesting parameter the user can set is the direction of rows and columns. This will determine if you are interested in rows/columns that are higher or lower than the average or even both.

The final input value is the number of seeds which are generated to start the ISA algorithm with. For now the GUI only allows for random seeding, but advanced user can use the `isa2` package to set non-random seeds which are based on knowledge of the data (e.g. gene sets).

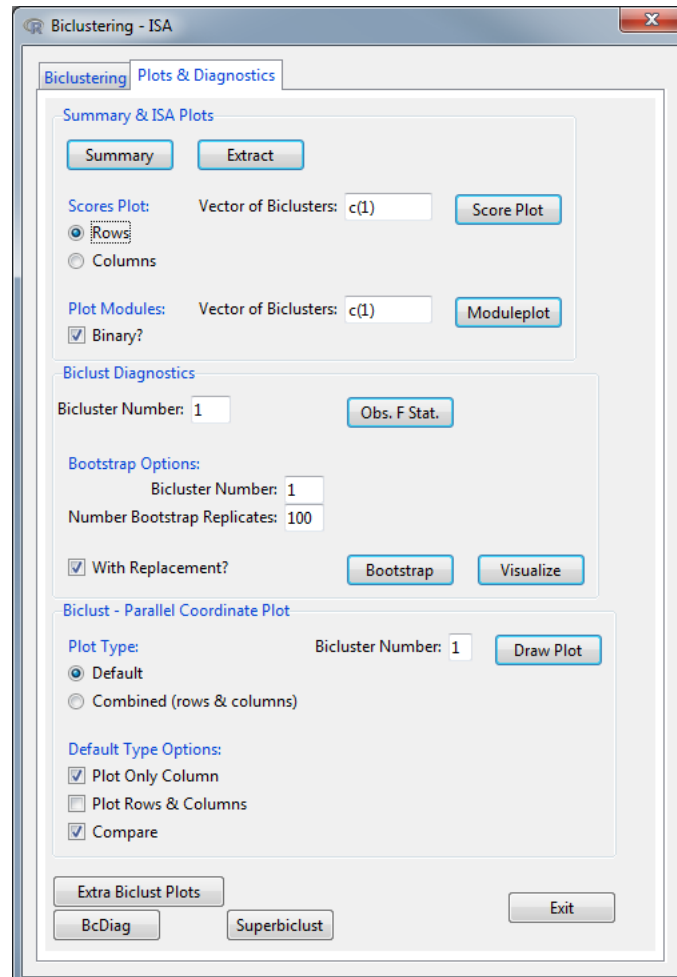


Figure 17: ISA Window - Plots & Diagnostics Tab

In Figure 17 above, the the second tab for the ISA method can be found. Apart from the first box ‘Summary & ISA Plots’ the plotting and diagnostic options are the same as for the `biclust` methods therefore they will not be elaborated on.

The *Summary* button will give some information about the bicluster result of `isa2`. It will give the number of found clusters and report the top 10 clusters based on the robustness score which is a measure of how well the rows and columns are correlated. The next button *Extract* will transform the result of `isa` to a list of biclusters and save it in an list-object called **Extract**. Each entry of this list will have two sublists, **rows** and **columns** which contain the indices of the subset of rows and columns.

The last two buttons in this box will create graphs based on a user-defined vector of biclusters (e.g. `c(1,2,3)` = bicluster 1, 2 and 3). The first, *Scores Plot*, will plot either the row or columns scores (depending on the radio button selection) which is a number between -1 and 1. The further this number is from zero, the stronger the association of the given *row* or *column* to the bicluster. The last, *Moduleplot*, will create image plots for the chosen set of modules as well as the matrix of the original data. The binary check box will determine whether to binarize the biclusters before plotting or use the actual ISA scores. An example of these last two graphs can be find in Figure 51 in the Appendix.

On a final note it should be mentioned that while the *extract* and *score plot* functions were not available in the `isa2` package, they were based on the author’s code in the Csárdi, G. (2013).

3.6 iBBiG-package

iBBiG or *Iterative Binary Bi-clustering of Gene Sets* is a biclustering algorithm proposed by Gusenleitner *et al.* (2012) which, similar to *Bimax*, will look for submatrices (= modules) of 1's in a sparse binary matrix. But because it works under the assumption of noisy data (non-perfect binarization), a number of 0's are tolerated within a bicluster. Further, it also allows for discovery of overlapping biclusters and the method is optimized for finding modules in matrices of discretized p-values from gene set enrichment analysis (GSA). However this does not prevent one to apply it to any kind of binary matrix. Another attractive feature of iBBiG is that it does not require prior knowledge or limit the number or size of clusters.

In short, the iBBiG algorithm consists out of three main components: 1) a module *fitness score* 2) a heuristic search algorithm to identify and grow modules in a high dimensional search space (*Genetic Algorithm*) and 3) an iterative extraction method to mask the signal of modules that have already been discovered (*Iterative Module Extraction*). Further information can be found in Gusenleitner *et al.* (2012).

The *iBBiG window* is shown in Figure 18 which contains the parameters which iBBiG needs to execute its algorithm.

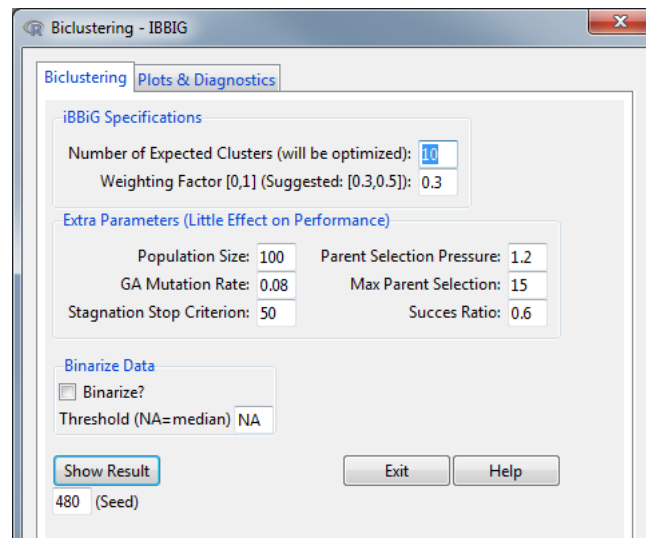


Figure 18: *iBBiG - Clustering Tab*

The most important parameter is α , the weighting factor, which is responsible for the balance between module *homogeneity* and module *size* when computing the fitness score (number of phenotypes versus number of gene-sets, consequently tradeoff between specificity and sensitivity). A small α will give more weight to the module size while a large one gives more weight to the homogeneity. However the authors showed in simulated studies that a range of 0.3-0.5 is appropriate with an optimal default value of 0.3.

The other important parameter is the number of expected values, but because the algorithm is optimized to find a minimal number, this parameter can be larger than the expected value. This means it is recommended to choose an upper boundary for this parameter.

All the other parameters are linked to the genetic algorithm which is a class of heuristic search algorithms based on evolutionary concepts. It is however recommended to keep these on default as these input parameters were reported to have little effect on the results.

Finally, the active data set in your R Commander session can be binarized if this is not a binary matrix by default. (This uses the same function available in `biclust`.) Do note that when plotting graphs (e.g. the `biclust` plots), the original data matrix will be used which is the active data set in your session.

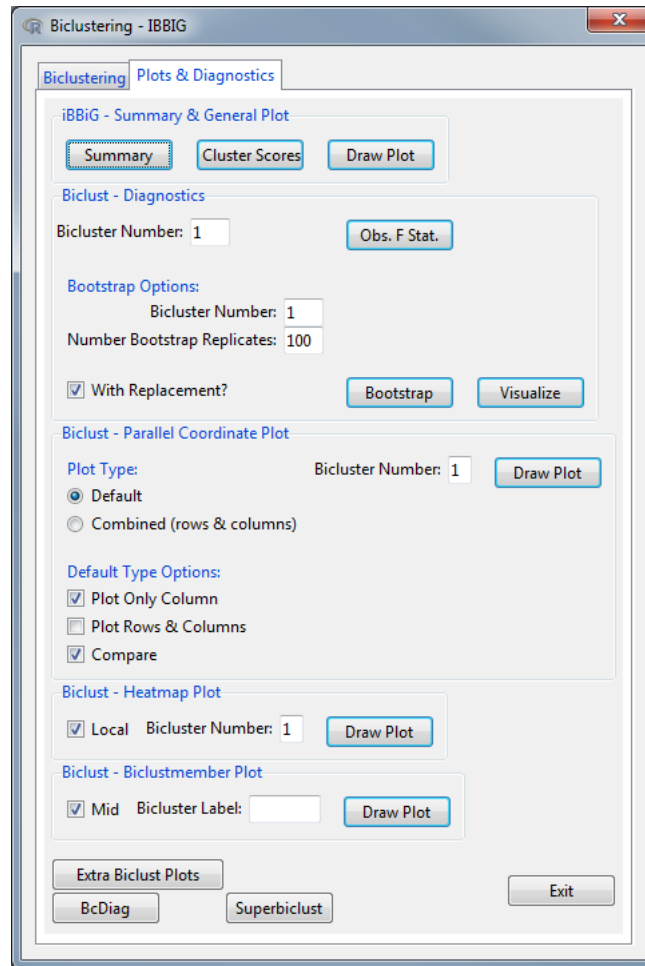


Figure 19: *iBBiG - Plots & Diagnostics Tab*

As can be seen in Figure 19, a lot of the functionality is the same as for the `biclust` algorithms so it therefore will not be explained again here.

On the top, three `iBBiG` specific buttons can be found. The *summary* button will give some basic information about the number of clusters found and first five biclusters of which the number of rows and columns will be reported. The next button, *Cluster Scores*, will give the module score of the clusters and the third one, *Draw Plot*, will create a general graph of the biclustering result. This graph contains a plot of all the modules colour-labeled in the matrix, aswell as some histograms with the module sizes, module scores and weighted scores. An example of this can be found in the Appendix in Figure 52.

3.7 Diagnostic Packages

3.7.1 BcDiag-package

The first implemented diagnostics package is `BcDiag` which is compatible with `biclust`, `fabia` and `isa2`. Since the output of `iBBiG` is an extension of `biclust`, the diagnostics will also work on this package. `BcDiag`'s task is the visualisation of bicluster data which can be categorized in three sections:

1. Profiling and Summarising the biclustered versus the clustered data simultaneously.
2. Profiling and Summarising the biclustered data only.
3. Exploring the biclustered data using *anova* and *median polish* techniques.

A general overview will follow, but more detailed descriptions can be found in Aregay et al. (2014).

In Figure 20, the *BcDiag Window* with the accompanying *output summary* window are shown. Note that in this example the diagnostics window was called from a *Plaid Window* as can be read in the window title.

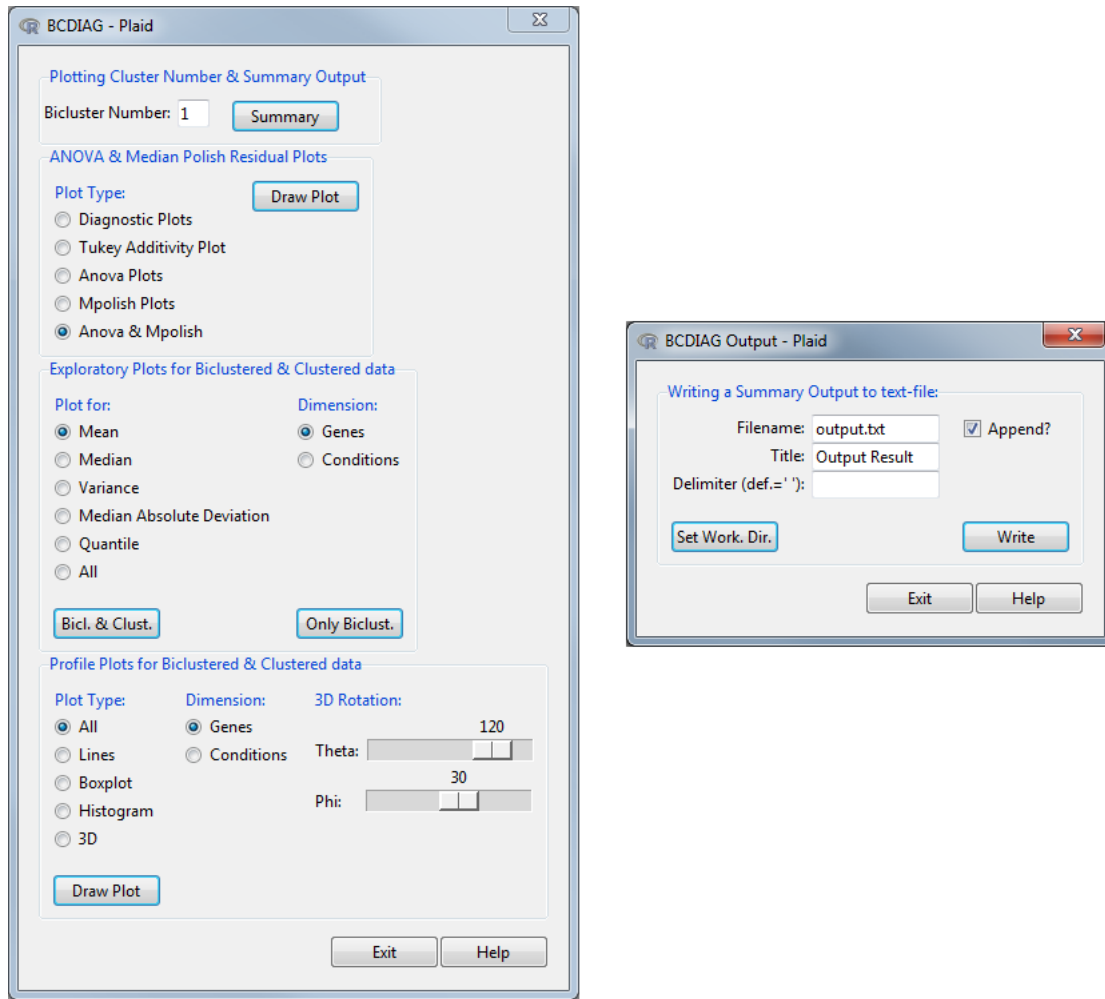


Figure 20: *BcDiag Window with Summary Output Window*

Summary Output

Using the *Summary* button will open up a small new window which can be seen at the right in Figure 20. With the use of this window a summary output in text format can be created. The user needs to provide *filename* and *title* and can also decide on which *delimiter* to use (default=space). Further the user is also able to set the working directory in which the file will be saved before using the *write* button which will create the text file. The original function was developed in the *biclust* package and will create a text file with the total number of biclustered, the dimension and the name of biclustered rows and columns.

Cluster Number

At the top of the *BcDiag Window* the user can set the number of the bicluster for which the graphs should be created. This value is used for all the plots down in the window apart from of course the summary button which extracts all biclusters.

ANOVA & Median Polish Residual Plots

Here residual plots or residual box plots for a bicluster can be created. These can be diagnostic plots (fitted vs residual, QQ,...), *Tukey additivity plots*, *anova* and *mpolish* plots. An example of some of these graphs can be found in the Appendix in Figure 53.

Profile Plots

With this graph, profile plots for biclustered data can be created. These profile plots can either be in *lines*, *boxplot*, *histogram* or even *3D* format. For this last type the rotation can be controlled by the *theta* and *phi* parameter, respectively the *z* and *y* axis. Further, the user can decide to either to make these profile plots on the *genes* (=rows) or *conditions* (=columns) dimension. If for example you would choose the conditions dimension and lines as type, you would get the same result as the *parallel coordinates plot* (combined type) from the *biclust* plots. In this scenario each line is a gene in the bicluster and it gets a red colour labeling if the column also is a part of the bicluster. The other way around, namely choosing the genes dimension, would mean that each line is one condition which is red when the gene on the x-axis is part of the bicluster. Some

examples of these are shown in Figure 54 in the Appendix.

Exploratory Plots

These plots are summary plots for the *mean*, *median*, *variance*, *MAD* and *quantiles* of the data. Just like for the profile plots, the dimension can be chosen here as well. Lastly the user can decide to show both the in-bicluster and out-bicluster information with the left button *Bicl. & Clust.* (as is always the case for the profile plots) or only the in-bicluster information with the *Only Biclust.* button. This is basically just the first part of the graph with both the in- and out-bicluster information as can be seen in figure 55 in the Appendix.

3.7.2 superbiclust-package

A central issue of biclustering is its stability which can highly influence the ability to interpret the results of a biclustering analysis. The stability can be affected by initialisation, parameter settings and perturbations such as random noise. Shi *et al.* (2010) introduced a novel procedure for obtaining robust biclusters from a set of initial values which is exactly what **superbiclust** will accomplish. It will construct a hierarchy of biclusters based on for example the Jaccard index from which the robust biclusters based on this ‘superbiclust’ concept can be extracted.

Starting now, the functionality of the *superbiclust window* will be explained, following the default structure of a superbiclust analysis. However, for more information about the architecture of **superbiclust**, see Khamiakova, T. (2013, 2014) and see Shi *et al.* (2010) for more details about the ‘superbiclust’ concept itself.

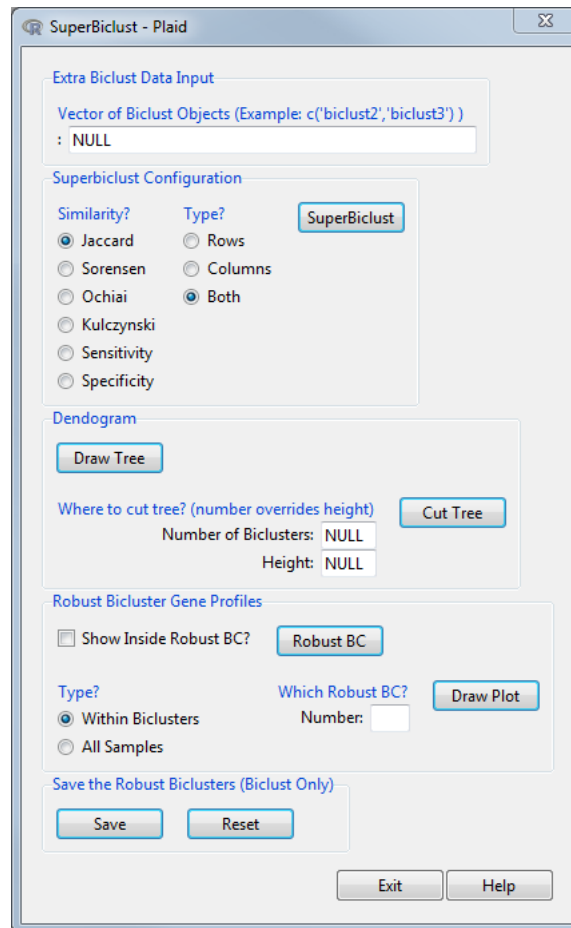


Figure 21: *Superbiclust Window*

Extra Biclust Data Input

First of all, at the top of the window in Figure 21, there is room to add other biclust objects which are loaded in the R workspace. The format should be as in the given example: `c('biclust2','biclust3')`. This box will only be available for **biclust** methods and **iBBiG**.

Superbiclust Configuration

In this box, the first part of a superbiclust analysis is executed, namely the computation of the similarity matrix of the biclusters. In this figure, we are working with the results of the plaid method as can be seen in the title

window. This matrix can be based on several types of similarity and the user can also decide to base it on the rows, columns or on both (as is suggested for bicluster results).

Dendrogram

In the next step, the similarity matrix will be used to construct a hierarchical tree. This is visualised in a dendrogram which can be drawn with the *Draw Tree* button of which an example is given in Figure 56 (Appendix). By cutting this tree now, the *robust* or *super biclusters* can be obtained. This user can do this either by setting a *number* of desired *clusters* or by choosing an appropriate *height* to cut the tree at before using the *Cut Tree* button.

Robust Bicluster Gene Profiles

After cutting the tree, some information can be called with the *Robust BC* button with or without the *Show Inside* checkbox. The result of this can be found in Figure 57 in the Appendix. In first part a matrix is presented with in the first row the index of the robust bicluster and in the second row the number of original biclusters it contains. Note that only those robust biclusters created from more than one bicluster will be displayed. The second part, which will appear if you check the *Show Inside* option, will tell you the indices of the original biclusters in a robust bicluster.

In the last part of this box, the user is able to quickly draw a profile plot of a robust bicluster of choice. Just like in the *Bcdiag Window*, either all the data (in-bicluster + out-bicluster) or only the in-bicluster data can be drawn.

Save the Robust Biclusters

The last box gives the user the option to *save* the robust biclusters. By doing this, the user can go back to the previous window (plots & diagnostics tab) or even to the *bcdiag* window. If the user creates the graphs now, they will be based on the robust biclusters instead of the original ones. However, take note that this only applies to the *biclust*-plots/diagnostics and the plots created by *bcdiag*. For example in the case of *isa2*, the user will not be able to create the correct scores or modules plots as this information is lost after combining biclusters into a robust one. Another example would be the buttons in the '*summary & general plot*' box of *iBBiG*. If you would like to go back to the original bicluster result, simply click the *Reset Button*. Also note that currently this save-option is not available for the *fabia*-methods.

4 A Guideline to New Implementations

Thanks to the easy plug-in extensibility of R Commander, adding new dialogs is very simple. One only needs to make the desired window function and then add this to the `menus.txt` file situated in the plugin package folder (`package -> inst -> etc`) as explained in Fox (2007).

If you are familiar with the `Rcmdr` and `tcltk` syntax, you can make your own customized dialog functions and then send these to the maintainer of `RcmdrPlugin.BiclustGUI`.

However, if you are new to these two packages or simply want an easier and quicker way to construct your biclustering dialogs, *template functions* have been created to serve this purpose. In the following sections, an explanation will be given on two kinds of scripts in which these functions are called to create your desired window. One script will address the creating of a dialog to execute biclustering methods, the other one addresses the creation of an extra tool window. This last one can serve as either an extension of the first or can be a general diagnostic window compatible with multiple methods. This will be further discussed in a later section. It should also be noted that these templates are specifically tailored to biclustering methods and that all of the already implemented methods are done through these scripts.

In the end, what should be given to the maintainer to implement a new method are the following items:

- The necessary window scripts (manually or through the templates)
- A description of how the menu should look like in R-Commander (e.g. Biclust menu opens up into several subitems like Plaid, Xmotif, ...)

4.1 Implementing a New Method

For starters, each biclustering window is made out of two tabs. One to execute the method and one to show diagnostics and plots. Both tabs have some standard buttons which will appear at the bottom as well as some optional ones as depicted in Figure 22.

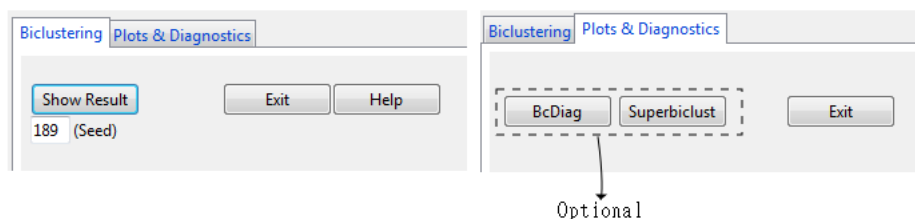


Figure 22: *Standard Tab Buttons*

As will become clear later, the clustering tab will contain the parameters of the biclustering method while the second tab will contain diagnostics and plots *specific* to this method. The concept is that the optional buttons will lead to new windows of general diagnostic packages, compatible with multiple biclustering methods.

Secondly, it is also important to realise that the idea is to have one biclustering function/method for each biclustering dialog and tailor it towards this. Why this is important will become clear once we go over the details of the template script.

4.1.1 New Method Script - ClusterTab & PlotDiagTab

We will now start going through `newmethod_script.R` which can be found in the Appendix.

General Method Information

The script starts by making a function which will be called on to create a window. First thing you should do of course is to rename this function to your own liking. Next, some objects are initialized that will be used to store information in about your window you are about to create.

```
newmethod_WINDOW <- function(){  
  
  new.frames <- .initialize.new.frames()  
  grid.config <- .initialize.grid.config()  
  grid.rows <- .initialize.grid.rows()  
}
```

```
#####
## GENERAL INFORMATION ABOUT THE NEW METHOD/WINDOW ##
#####

methodname <- "A new method"

methodfunction <- "methodfunction"
data.arg <- "d"
data.matrix <- TRUE
methodshow <- TRUE
other.arg <- ""
methodhelp <- ""

# Extra Data Conversion Boxes
data.discr <- FALSE
data.bin <- FALSE

# Possibility to give a seed ?
methodseed <- TRUE

## COMPATIBILITY? ##

# BcDiag
bcdiag.comp <- FALSE

# SuperBiclust
superbiclust.comp <- FALSE

} # ignore
```

The scripts starts be filling in some information about your biclustering method. A clarifying example follows later, filling in this information for the Plaid method.

- **methodname:** The title of the window which will be shown on top. It may contain no special characters besides '-'. It is important to know that the result of your biclustering function will be saved in a global variable, named `methodname` without the spaces and '-' symbols. Therefore, do not make this string too elaborate.
- **methodfunction:** A string which contains the name of your biclustering method's function. By this the actual R function, corresponding with the method, is meant. Without going into too much detail, it might be interesting to know that the R command to execute the biclustering actually starts with this small string. New arguments are simply appended to this string until you have the full command you need.
- **data.arg:** The name of the argument which needs the data.
- **data.matrix:** Logical value which will determine if the data needs to be transformed to a matrix with `as.matrix()`.
- **method.show:** Logical value which decides if the the object in which the clustering result is saved, should be printed. See figure 23.
- **other.arg:** A string containing extra arguments you do *not* want the user to change. For example in the Plaid window, it is used as following: since the `biclust` package uses only one function to execute all of the different methods, for the Plaid method the string here was `",method='BCPlaid'"`. Also note the use of the comma in the beginning! This is because this string is not restricted to only one argument, it could contain several of them. They simply need to be added in this string as they would be added inside the function itself.
- **methodhelp:** The name of the helppage the help button should be directed to. (`help(methodhelp)`)
- **data.discr:** Logical value determining if a frame should be added above the *Show Results* button with the option to discretize the data. The `discretize` function from the `biclust` package is used to achieve this. See Figure 24.

- **data.bin:** Logical value determining if a frame should be added above the *Show Results* button with the option to binarize the data. The **binarize** function from the **biclust** package is used to achieve this. See Figure 24.
- **methodseed:** Logical value determining if there should be a seed box below the *Show Results* button. This will make sure a **set.seed()** is executed before the biclustering function.
- **bcdiag.comp & superbiclust.comp :** Logical value determining compatibility with the **BcDiag** and **superbiclust** package respectively. Please note that this only enables the appearance of the buttons in the second tab. For a button to function properly, some minor coding needs to be carried out by the maintainer.

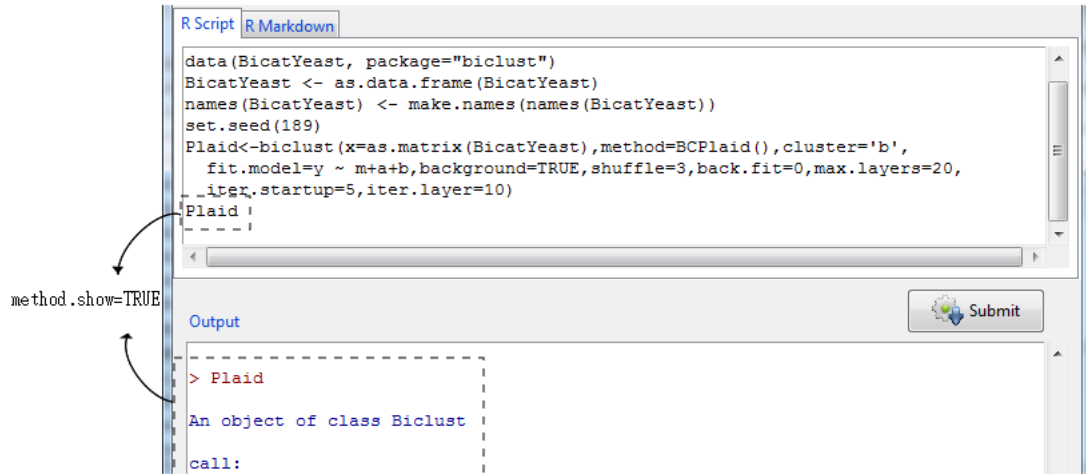


Figure 23: *Example of Plaid Biclustering - With method.show=TRUE*

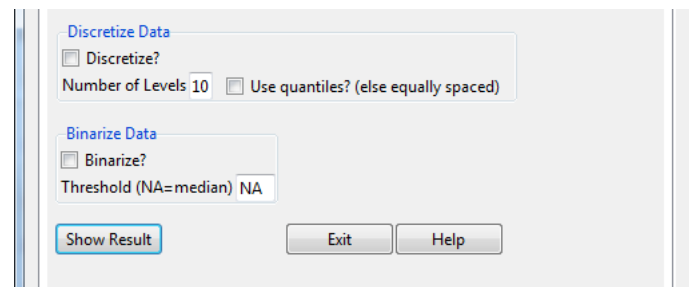


Figure 24: *Example of discretize and binarize frames.*

Cluster Tab

After providing the information about the method we can finally start making windows! Both the clustering tab and the plotting & diagnostics tab are created in three easy steps as shown in Figure 25:

1. Making the frames
2. Configuring the frames into a grid
3. Combining rows into a box

1. Making the Frames

2. Configuring the Grid

3. Combining Rows

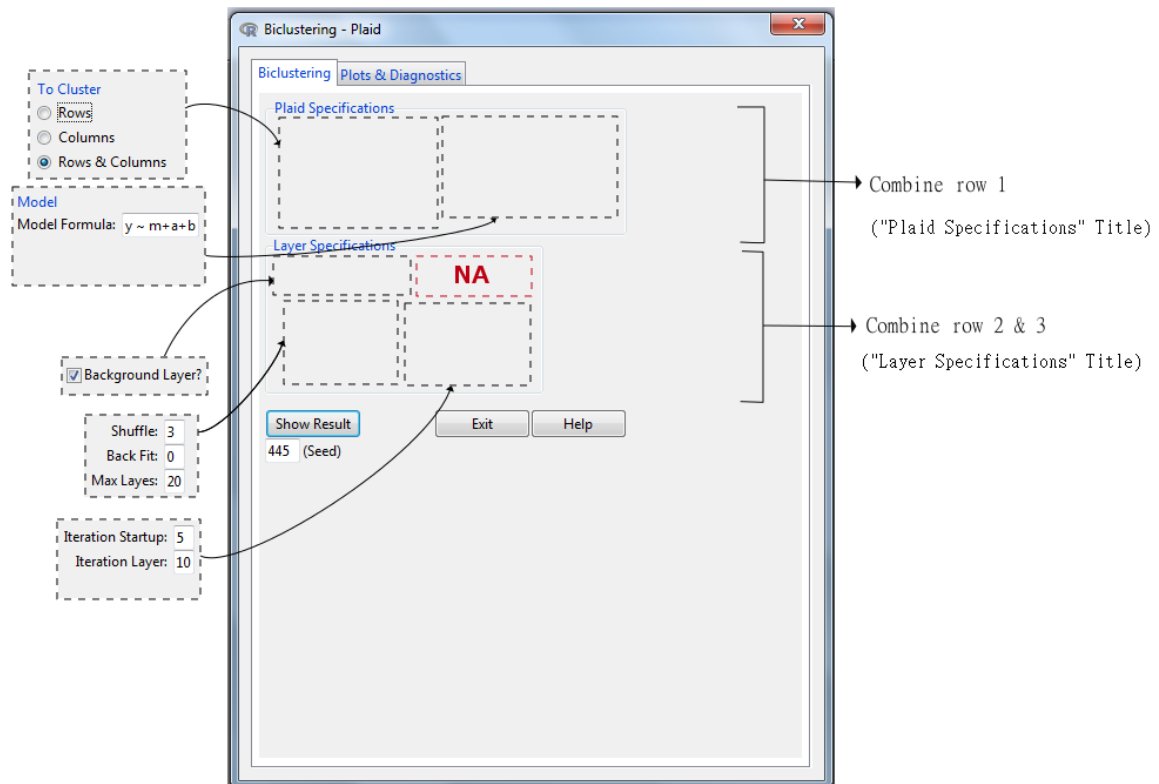


Figure 25: Making windows in 3 steps

```
#####
#### CLUSTERING TAB ####
#####

input <- "clusterTab"

### 1. ADDING THE FRAMES ###

# Add frames here

### 2. CONFIGURING THE GRID ###

grid.config <- .grid.matrix(input=input,c("frame1","frame2","frame3",NA,"frame4",NA),
                             nrow=3,ncol=2,byrow=TRUE,grid.config=grid.config)

### 3. COMBING THE ROWS ###

grid.rows <- .combine.rows(input=input,rows=c(1),title="A nice box: ",
                             border=TRUE,grid.rows=grid.rows,grid.config=grid.config)
```

```
grid.rows <- .combine.rows(input=input,rows=c(2,3),title="A nice box: ",
                           border=TRUE,grid.rows=grid.rows,grid.config=grid.config)
```

Looking at the script, you can see it starts with putting the `input` to *clusterTab*. This will make sure everything you are creating and saving now will be done for the first tab.

Step 1:

As already explained earlier, the first step will be to create the frames in which you want to put your function arguments. A variety of frames can be created, but these will be explained in more detail in the following section. To give a quick summary, here is the list of the types of frames which can be generated in the clustering tab.

- Check Boxes
- Radio Buttons
- Entry Fields
- Sliders
- Spinboxes

In future updates, there is still the possibility to add even more types if required.

Step 2:

During the creation of the frames in the previous step, you will have given each of them a unique name. Using these framenames, the next step will be to simply order them into a matrix grid, filling in the empty spots with NA's. This is achieved with the `.grid.matrix` function. The function accepts the exact same arguments as the `matrix` function apart from two new ones, namely `input` and `grid.config`. The first is to make sure the template function knows we are adding frames in the first tab, while second one is there to ensure that the new information is added to the old `grid.config` object and that old information is not lost.

Further, it is important to know that the inserted frames will *always* be pulled towards the north-west as much as possible. Therefore in a 1-row matrix, something like `c(NA,"frame1")` or `c("frame1",NA)` would give exactly the same result.

Step 3:

The final step will enable you to put one or multiple rows in a separate box which can serve two different purposes. The first, being the most obvious one, is simply to add some visual distinction between rows with the help of a title with or without a border around the row(s).

The second purpose is connected to the way frames are added in this *grid*. Sometimes if frames have a large difference in size, other frames might seem to be jumping to the right, trying to fit in one general grid. In general if you see this happening, putting this row(s) in a box will solve this problem and the frames will again be pulled towards the left.

Creating these boxes by combining rows is again very easy, one simply needs to use the `.combine.rows` function which will save the necessary information in the `grid.rows` object. The function only has three arguments you should change: `rows` which is a vector containing the rows you wish to combine, `title` to give the box a title ("" means no title) and `border` to decide if there should be a border.

Note that in contrast to the grid configuration, you can call this function multiple times until the desired result is obtained.

PlotDiag Tab

The clustering tab configuration is immediately followed by making the second tab.

```
#####
#### PLOTTING & DIAGNOSTICS TAB ####
#####

input <- "plotdiagTab"

### 1. ADDING THE FRAMES ###

# Add frames here

### 2. CONFIGURING THE GRID ###

grid.config <- .grid.matrix(input=input,c("frame5","frame6"),nrow=1,ncol=2,
  byrow=TRUE,grid.config=grid.config)

### 3. COMBINING THE ROWS ###

grid.rows <- .combine.rows(input=input,rows=c(1),title="Plot 1",border=TRUE,
  grid.rows=grid.rows,grid.config=grid.config)

#####
## USE ALL THE ARGUMENTS ABOUT IN THE GENERAL CLUSTERTEMPLATE FUNCTION ##
#####

cluster_template(methodname=methodname,methodfunction=methodfunction,
  methodhelp=methodhelp,data.arg=data.arg,other.arg=other.arg,
  methodseed=methodseed,grid.config=grid.config,grid.rows=grid.rows,
  new.frames=new.frames,superbiclust.comp=superbiclust.comp,
  bcdiag.comp=bcdiag.comp,data.matrix=data.matrix,
  data.discr=data.discr,data.bin=data.bin,methodshow=methodshow)
```

Analogous to the first tab, this part of the script starts by putting the `input` variable to `plotdiagTab`. Next the same three steps are repeated as for the clustering tab. There is only one difference and that is the addition of one new type of frame which can be created in the plots & diagnostics tab. This is the *manual button frame* which will also be explained in more detail down below. Basically a manual button can serve two purposes, it can be tied to a plot or diagnostics function which uses some arguments or it can lead to a new window entirely which can be created with the *new tool script*.

Finally at the end of the script, all the variables and created objects are used in the `cluster_template` function which will create the actual window.

4.1.2 The Frame Scripts

In this section, the several types of frames which can be used in the `newmethod_script` will be showcased. The idea is that these parts of the R-code (which are also in the Appendix) are copy-pasted into the `newmethod_script` and are adjusted as deemed necessary.

All the frame types have the `title` and `border` option in common. The results of these options can be seen in figure 26. Also note that for each frametype the information is saved in one object, namely `new.frames`. Just as the grid and row configuration earlier, new information will keep on getting added to this object, now with the help of the `.add.frame` function. Lastly, at the start of each frame script, a `type` variable will be set to determine the type of frame for this previous mentioned `.add.frame` function.

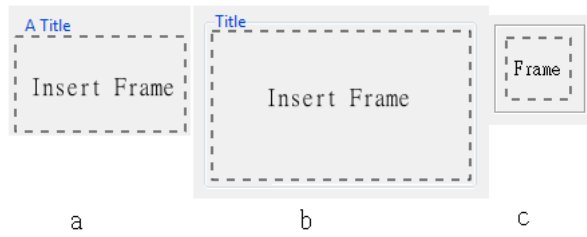


Figure 26: **a.** Title & No Border **b.** Title & Border **c.** No Title & No Border

Entry Fields

The first type of frame is the entry fields frame. It can be used for both numerical arguments and character arguments of your biclustering function. Multiple entries can be added in one frame which will be placed below each other.

```
#### ENTRY FIELDS FRAME ####

type <- "entryfields"

# Change variables accordingly:
frame.name <- "entryframe1"
argument.names <- c("Argument 1","Argument 2","Argument 3")
argument.types <- c("num","num","char")
arguments <- c("arg1","arg2","arg3")
initial.values <- c(1,2,"a")
title <- "A Title"
border <- FALSE
entry.width <- c("2","2","6")

# Do not change this line:
new.frames <- .add.frame(input=input,type=type
,frame.name=frame.name,argument.names=argument.names
,arguments=arguments,initial.values=initial.values
,title=title,border=border,entry.width=entry.width
,argument.types=argument.types ,new.frames=new.frames)
```

Figure 27: Entry Fields: Code + Example

Entry Fields Variables:

- **frame.name:** The unique name of this frame. (Which is used in the grid matrix)
- **argument.names:** The argument names how they will appear in the window.
- **argument.types:** A vector defining if the argument is "num" or "char". This basically just means if there should be a ' ' around the value when filling it in in the biclustering function. (e.g. In Figure 27 the arguments would be filled in as ,arg1=1,arg2=2,arg3='a')
- **arguments:** The actual argument names, used for the biclustering function.
- **initial.values:** A vector containing the initial values in the entry fields.
- **title:** Optional title for the frame (" " means no title)
- **border:** Logical value determining the presence.
- **entry.width:** A vector containing the width of the entry fields (1 width = 1 number/character).

Check Boxes

The second type of frame is the check boxes frame which is used for TRUE/FALSE arguments. Just like for entry fields, multiple check boxes can be added below each other.



Figure 28: Check Boxes: Code + Example

Check Boxes Variables:

- **frame.name:** The unique name of this frame. (Which is used in the grid matrix)
- **argument.names:** The argument names how they will appear in the window.
- **arguments:** The actual argument names, used for the biclustering function.
- **initial.values:** A vector containing the initial values in the entry fields. (0 for FALSE, 1 for TRUE)
- **title:** Optional title for the frame (" " means no title)
- **border:** Logical value determining the presence.

Radio Buttons

The next type is radio buttons, which is used for only one argument with a finite number of values.



Figure 29: Radio Buttons: Code + Example

Radio Buttons Variables:

- **frame.name:** The unique name of this frame. (Which is used in the grid matrix)
- **argument.names:** The names of the buttons how they will appear in the window.
- **arguments:** The actual argument name, used for the biclustering function.
- **argument.types:** Just as for the entry fields, this will determine if the values are filled in with or without ". The two options are again "num" and "char", but in contrast with the entry fields it is now only one value and not a vector..

- **argument.values:** The actual values of the radio buttons that correspond to the values passed to biclustering function.
- **initial.values:** The initial value of the radio buttons. It will determine which button is selected on opening the window.
- **title:** Optional title for the frame ("" means no title)
- **border:** Logical value determining the presence.

Value Sliders

The following type will create value sliders which can only be used for numerical values. Again multiple sliders can be placed under each other. The current value of the slider will always appear on top of it.

```
#### VALUE SLIDER FRAME ####

type <- "valuesliders"

# Change variables accordingly:
frame.name <- "sliderframe1"
argument.names <- c("Slider 1 ", "Slider 2 ", "Slider 3 ")
arguments <- c("sliderarg1", "sliderarg2", "sliderarg3")
initial.values <- c(1,5,10)
from <- c(1,1,1)
to <- c(5,50,500)
by <- c(1,10,50)
length <- c(50,100,150)
title <- "Title"
border <- TRUE

# DO NOT CHANGE THIS LINE:
new.frames <- .add.frame(input=input,type=type,
  title=title,border=border,frame.name=frame.name,
  argument.names=argument.names,arguments=arguments,
  initial.values=initial.values,from=from,to=to,by=by,
  length=length,new.frames=new.frames)
```

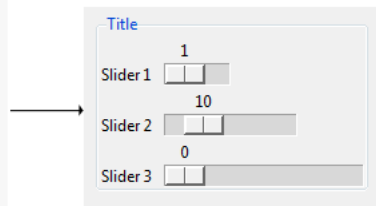


Figure 30: Value Slider: Code + Example

Value Sliders Variables:

- **frame.name:** The unique name of this frame. (Which is used in the grid matrix)
- **argument.names:** The argument names how they will appear in the window.
- **arguments:** The actual argument names, used for the biclustering function.
- **initial.values:** Vector of initial values of the sliders.
- **from:** Vector of starting points of the sliders.
- **to:** Vector of ending points of the sliders.
- **by:** Vector with the values determining how one movement of the sliders will change the current value.
- **length:** Vector containing the lengths of the sliders.
- **title:** Optional title for the frame ("" means no title)
- **border:** Logical value determining the presence.

Spin Boxes

This type will create spin boxes which are again solely used for numerical values. Just as for sliders, multiple spin boxes can be placed below each other.

```
#### SPIN BOX FRAME ####

type <- "spinboxes"

# Change variables accordingly:
frame.name <- "spinboxframe1"
argument.names <- c("Spin Box 1: ", "Spin Box 2: ", "Spin Box 3: ")
arguments <- c("spinarg1", "spinarg2", "spinarg3")
initial.values <- c(5, 10, 20)
from <- c(1, 5, 10)
to <- c(10, 20, 30)
by <- c(1, 1, 1)
entry.width <- "2"
title <- "Spin Box !"
border <- TRUE

# DO NOT CHANGE THIS LINE:
new.frames <- .add.frame(input=input, type=type,
  frame.name=frame.name, argument.names=argument.names,
  arguments=arguments, initial.values=initial.values,
  from=from, to=to, by=by, entry.width=entry.width,
  title=title, border=border, new.frames=new.frames)
```

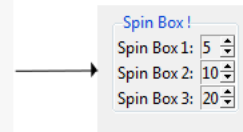


Figure 31: *Spin Boxes: Code + Example*

Spin Boxes Variables:

- **frame.name:** The unique name of this frame. (Which is used in the grid matrix)
- **argument.names:** The argument names how they will appear in the window.
- **arguments:** The actual argument names, used for the biclustering function.
- **initial.values:** Vector of initial values of the spin boxes.
- **from:** Vector of starting points of the spin boxes.
- **to:** Vector of ending points of the spin boxes.
- **by:** Vector with the values determining how much one click will change the current value.
- **entry.width:** Width of all the spinboxes (one value which applies to all of them)
- **title:** Optional title for the frame (" " means no title)
- **border:** Logical value determining the presence.

Manual Buttons

The last type of frame which can be utilized, is making manual buttons which will *only* work in the plot & diagnostics tab. As already explained earlier, there are two primary uses for these buttons. The first use is to simple execute a plot or diagnostic function, based on the arguments of other frames in the window. The second application is to tie the button to another window function (created with `newtool_script`) to open up more options.



Figure 32: *Manual Button: Code + Example*

Manual Button Variables:

- **frame.name:** The unique name of this frame. (Which is used in the grid matrix)
- **button.name:** The text which will appear on the button. No special characters are allowed and if the result of the button is saved, it will be in an object with this name without the spaces.
- **button.function:** A string of the function which should be tied to this button. This could either be a plot, diagnostic or simple summary. Another useful practice is to actually make an entire new function for this manual button. This new function could then for example contain a series of diagnostic functions which would then be carried out all at the same time when clicking on this button.
- **button.data:** The name of the data argument the button function. The data which is loaded in R Commander will then be pasted after this argument. (Simply put "" when this is not necessary)
- **button.biclust:** The name of the biclustering result argument in the button function. Here the saved result after using the *Show Results* button will be pasted.
- **button.otherarg:** Just like for the biclustering function (**other.arg**), this is the variable in which one can add arguments to the function which should not be controlled by the user (e.g. ",type='biclust']"). Again, do not forget to add a comma in the beginning. The only time when the comma should be excluded is when both **button.data** and **button.biclust** are put to "".
- **arg.frames:** A vector containing the names of those frames from which this button function should pull its arguments.
- **save:** Logical value determining if the result of the button function should be saved. For example for a plotting function this is mostly likely not necessary, however for a diagnostic result it is. The difference between a TRUE and FALSE option is shown in figure 33.
- **show:** Logical value determining if the button function should be shown in R Commander. It is good practice to do this for the plotting and diagnostics functions however if is a function to create a new window, it is probably not necessary to show it.

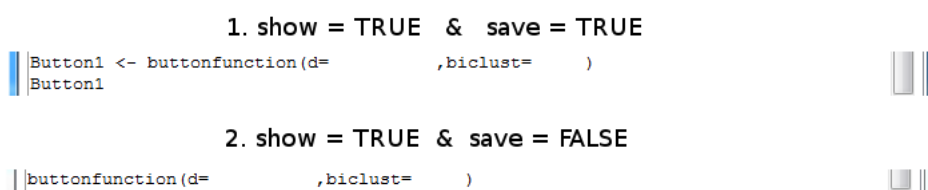


Figure 33: *Manual Buttons - save option*

4.1.3 A quick example - Plaid

The entire script to construct the *Plaid Window* can be found in the Appendix. In this small section, some parts will be highlighted and explained in their context.

```
#####
## GENERAL INFORMATION ABOUT THE NEW METHOD/WINDOW ##
#####

methodname <- "Plaid"
methodfunction <- "biclust"
data.arg <- "x"
data.matrix <- TRUE
other.arg <- ",method=BCPlaid()"
methodhelp <- "BCPlaid"
methodseed <- TRUE
data.discr <- FALSE
data.bin <- FALSE
bcdiag.comp <- TRUE
superbiclust.comp <- TRUE

# Biclust only (Not for public use)
extrabiclustplot <- TRUE
```

First of all, the general information is filled in for the plaid method as shown in the script above. Please note that the `extrabiclustplot` is a `biclust`-only variable and will not be elaborated on.

Next, in Figure 34 we can see the code which makes the two frames in the top row.

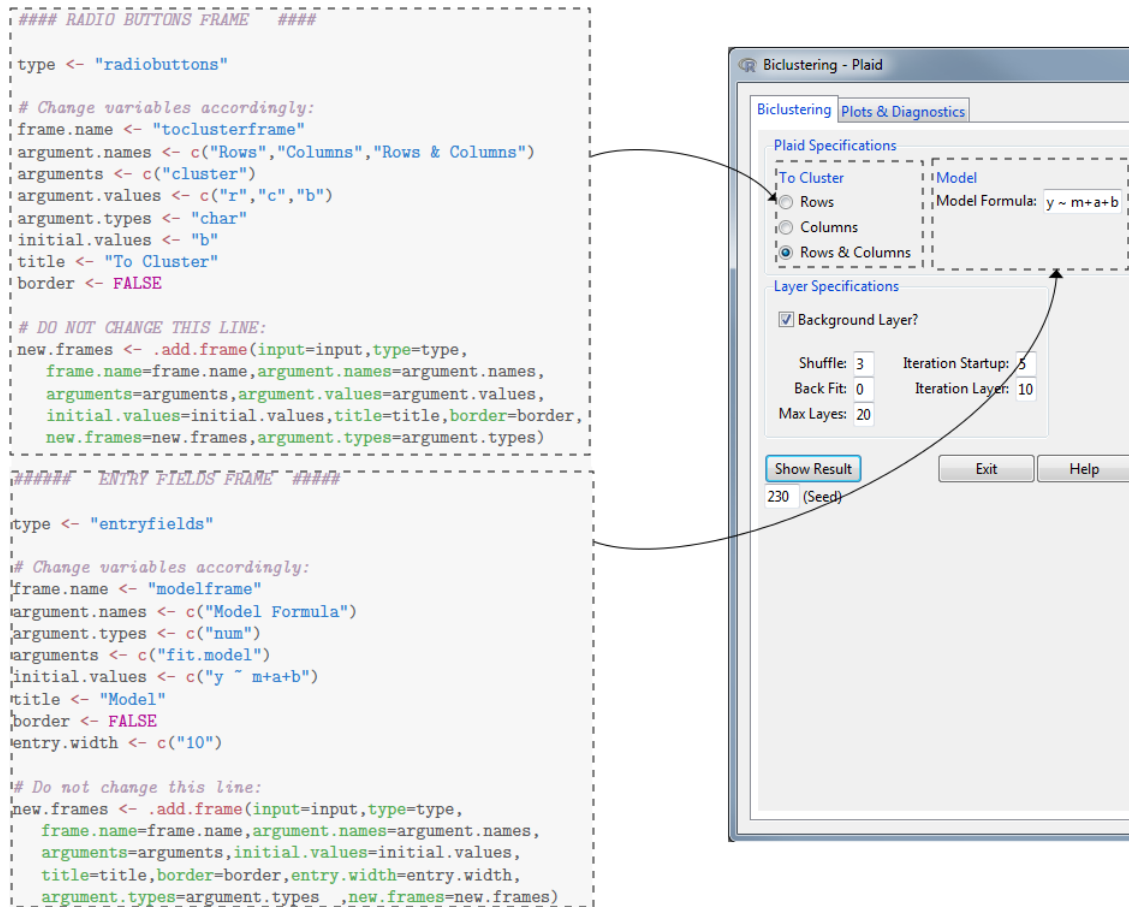


Figure 34: Building the Plaid Window - ClusterTab

Following the rest of the frame creations is of course the grid configuring and the row combining of the cluster

tab. In this extract of the script, one can see the two frames from Figure 34 being placed in the top row of the matrix after which the first row is made into a box with border and *Plaid Specifications* title.

```
### 2. CONFIGURING THE GRID ###
grid.config <- .grid.matrix(input=input,c("toclusterframe","modelframe",
  "backgroundcheckframe",NA,"backgroundentryframe1","backgroundentryframe2"),
  byrow=TRUE,nrow=3,ncol=2,grid.config=grid.config)

### 3. COMBINING THE ROWS ###
grid.rows <- .combine.rows(input=input,rows=c(1),title="Plaid Specifications",
  border=TRUE,grid.rows=grid.rows,grid.config=grid.config)
grid.rows <- .combine.rows(input=input,rows=c(2,3),title="Layer Specifications",
  border=TRUE,grid.rows=grid.rows,grid.config=grid.config)
```

Finally, an example of the creation of manual buttons is shown in Figure 35. Note that the line of code containing the `.add.frame` has been omitted here for better clarity of code. Naturally the creations of the other frames, together with the configuring of the grid and rowcombining follows after.

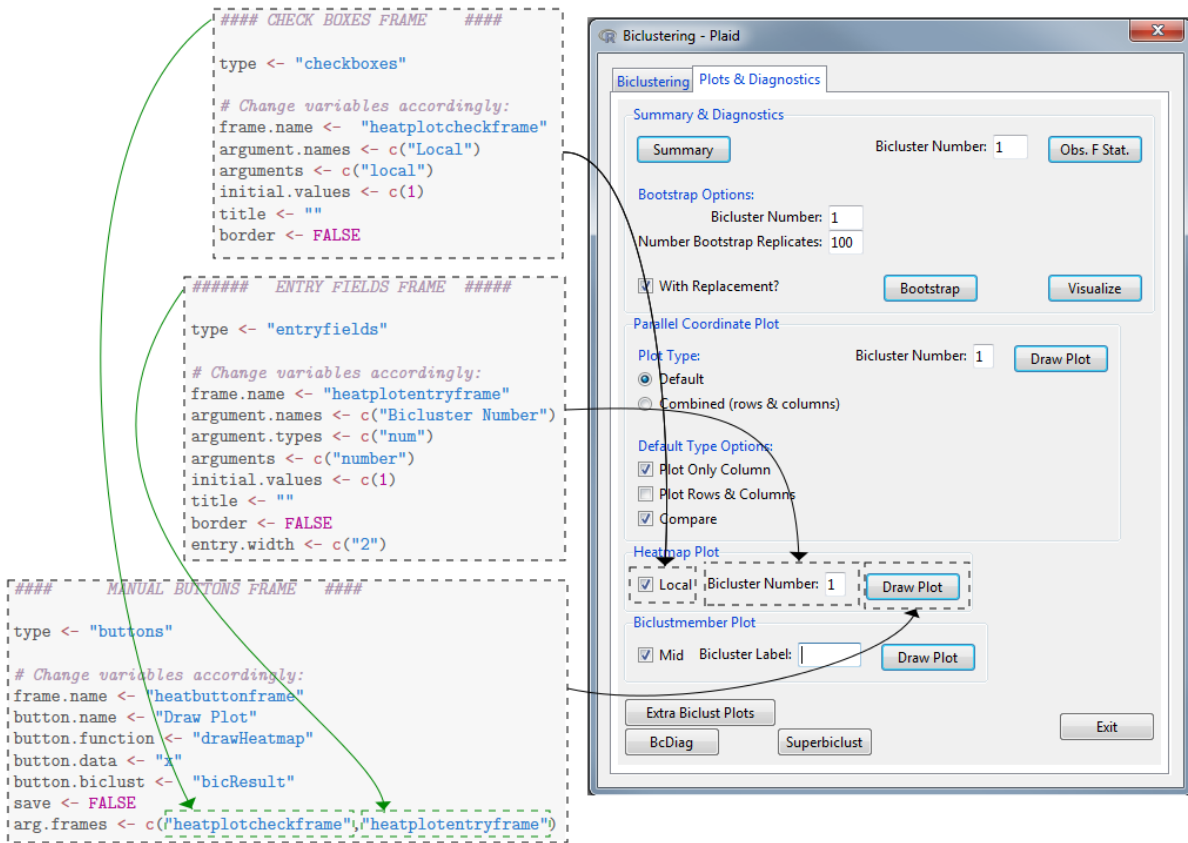


Figure 35: Building the Plaid Window - PlotDiagTab

4.2 Implementing a New Tool

The second template script (which is also available in the Appendix) is `newtool_script.R`. With the help of this script one can make a new window which can for example be tied to a manual button in the second tab of a clustering window. Another possibility is to make another general diagnostics window compatible with several other biclustering packages. This is for example already done with the `BcDiag` and `superbiclust` package for which the button appears in the second tab of the compatible methods. Also just as for these two packages, a compatibility variable could then be added in the begin of the `newmethod_script` to decide if the button for your new general diagnostics package should appear for a specific method.

4.2.1 New Tool Script

```

newtool_WINDOW <- function(methodname){

  new.frames <- .initialize.new.frames()
  grid.config <- .initialize.grid.config()
  grid.rows <- .initialize.grid.rows()

  #####
  ## GENERAL INFORMATION ABOUT THE NEW TOOL ##
  #####

  toolname <- "A new tool"
  toolhelp <- "helppage"
  data.matrix <- TRUE

  # Do not change this line:
  input <- "plotdiagTab"

  ### ADDING FRAMES ###

  # Analogous to plotdiag tab.

  ### CONFIGURING GRID ###
  grid.config <- .grid.matrix(input=input,c(),nrow=1,ncol=2,
    byrow=TRUE,grid.config=grid.config)

  ### COMBINING ROWS ###
  grid.rows <- .combine.rows(input=input,rows=c(),title="Plot 1",
    border=TRUE,grid.rows=grid.rows,grid.config=grid.config)

  #####
  ## USE ALL THE ARGUMENTS ABOUT IN THE GENERAL NEW TOOL FUNCTION ##
  #####

  newtool_template(toolname=toolname,methodname=methodname,
    toolhelp=toolhelp,data.matrix=data.matrix,grid.config=grid.config,
    grid.rows=grid.rows,new.frames=new.frames)
}

```

Just as for `newmethod_script` the adjusting of the template starts with changing `newtool_WINDOW` to your own desired windowfunction name. Next at the start of the script again the necessary objects are initialized followed by some variables containing general information about the tool:

- **toolname:** The name of the tool appearing at the top of the window.
- **toolhelp:** The helppage the help button should be linked with. (`help(toolhelp)`)
- **data.matrix:** Logical value determining if the active data set should be transformed to a matrix when using any of the functionality of the tool window.

After filling in these couple of variables, making this window is *exactly* the same as creating the second tab in `newmethod_script`. This means it is also possible to create manual buttons in this window. Note that even the `input` variable is again set to `plotdiagTab`. The only difference can be found in the last line where now the template function has changed to `newtool_template`.

Finally the attention should also be pointed to the fact that by default the `newtool_WINDOW` function has `methodname` as an argument. This means if this function is used inside a `newmethod_script`, the new tool window can behave differently depending on the method it is being called by. This will become clear in the example shortly.

4.2.2 A quick example - BcDiag

We will now go through a couple of parts of the script responsible for the BcDiag window. The full script can once again be found in the Appendix.

```

bcdiag_WINDOW <- function(methodname){

  new.frames <- .initialize.new.frames()
  grid.config <- .initialize.grid.config()
  grid.rows <- .initialize.grid.rows()

  # Some extra code to determine the input type: "biclust", "fabia", "isa2"
  biclust.names <- c("Bimax","CC","Plaid","Questmotif","Spectral",
    "XMotifs","IBBIG")
  fabia.names <- c("Fabia Laplace Prior","Fabia Post-Projection",
    "Fabia Sparseness Projection","Fabia SPARSE")
  isa.names <- c("ISA")

  if(methodname %in% biclust.names){
    extra.arg <- ",mname='biclust'"
  }
  if(methodname %in% fabia.names){
    extra.arg <- ",mname='fabia'"
  }
  if(methodname %in% isa.names){
    extra.arg <- ",mname='isa2'"
  }

  #####
  ## GENERAL INFORMATION ABOUT THE NEW TOOL ##
  #####

  toolname <- "BCDIAG"
  toolhelp <- "BcDiag-package"
  data.matrix <- TRUE
}#ignore

```

In the code above the variables containing the general information are filled in as expected, however just above some extra code has been added. These extra lines of code are simply determining what string should be saved in `extra.arg` based on the `methodname` given to `bcdiag_WINDOW`. This `extra.arg` variable will then be used later in the creation of a manual button as we will see now in Figure 36. Note that just as for the plaid example the `.add.frame` line has been omitted to maintain clarity.

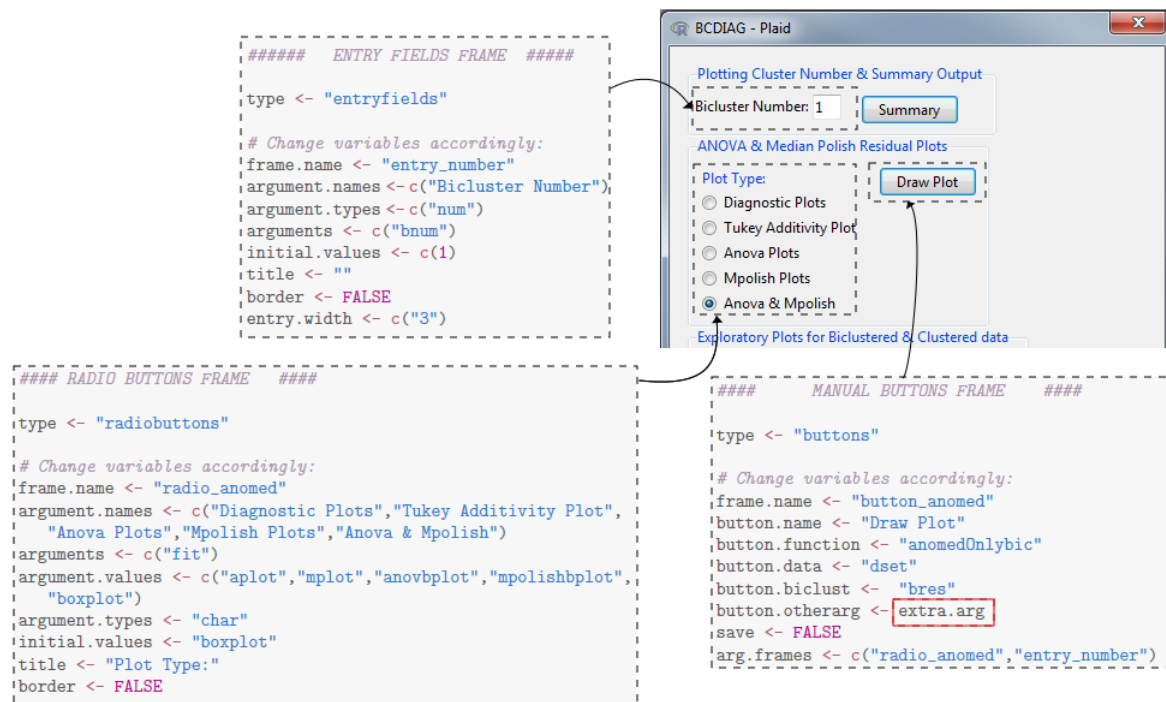


Figure 36: Building the BcDiag Window

Apart from the usual framebuilding, one can see that this earlier mentioned `extra.arg` is now used for the `button.otherarg` argument in the manual button setup.

This is only one of many ways in which you can tailor the tool window to a specific biclustering method. A more complicated example would be the implementation of the `superbiclust` window in which `methodname` is also transformed to `method_result` (`methodname` without spaces and `'-'`s) in order to make some distinctions. Further in this script a different `grid.config` and `grid.rows` will be created depending on `methodname`. Since apart from these additions the script is still very much the same as the others, it will not be elaborated on. It can however still be found in the Appendix.

4.3 Testing your Script

Once you have completed your script for a certain *windowfunction*, it is very simple to test it to check if it is to your liking or if any errors were made. Simply load in the `RcmdrPlugin.BiclustGUI` package and wait for the R Commander window to open up. Then run the entire script and execute the created *windowfunction*. A new window should appear on top of the R Commander one, containing whatever you specified in the script.

5 References

- Aregay, M., Otava, M., Khamiakova, T. and De Troyer, E. (2014). Package ‘BcDiag’. *Diagnostics plots for Bicluster Data*. Available: <http://cran.stat.ucla.edu/web/packages/BcDiag/BcDiag.pdf>. Last accessed 21th August 2014.
- Bergman, S., Ihmels, J. and Barkai, N. (2003). Iterative signature algorithm for the analysis of large-scale gene expression data. *Physical Review*, **67**.
- Cheng, Y. and Church, G.M. (2000). Biclustering of expression data. *Proc International Conference on Intelligent Systems for Molecular Biology*, **8**, 93-103.
- Csárdi, G. (2013). The Iterative Signature Algorithm - Tutorial. Available: http://raptor1.bizlab.mtsu.edu/s-drive/TEFF/Rlib/library/isa2/doc/ISA_tutorial.pdf. Last accessed 3rd September 2014.
- Csárdi, G. (2014). Package ‘isa2’. *The Iterative Signature Algorithm*. Available: <http://cran.r-project.org/web/packages/isa2/isa2.pdf>. Last accessed 21th August 2014.
- Culhane, A. and Gusenleitner, D. (2014). Introduction to iBBiG. Available: <http://www.bioconductor.org/packages/release/bioc/vignettes/iBBiG/inst/doc/tutorial.pdf>. Last accessed 21th August 2014.
- Dalgaard, P. (2001). The R-Tcl/Tk interface. *DSC 2001 Proceedings of the 2nd International Workshop on Distributed Statistical Computing, Vienna, Austria*.
- Dalgaard, P. (2001). A Primer on the R-Tcl/Tk Package. *R-News*. **1/3**.
- Dalgaard, P. (2002). Changes to the R-Tcl/Tk package. *R-News*. **2/3**.
- Dolnicar, S., Kaiser, S., Lazarevski, K. and Leisch, F. (2011). Biclustering: overcoming data dimensionality problems in market segmentation. *Journal of Travel Research*
- Fox, J. (2005). The R Commander: A basic-statistics graphical user interface to R. *Journal of Statistical Software*. **14(9)**.
- Fox, J. (2007). Extending the R Commander by “Plug-In” Packages. *R News*. **7/3**.
- Fox, J. and Bouchet-Valat, M. (2013). Getting Started With the R Commander. Available: <http://socserv.mcmaster.ca/jfox/Misc/Rcmdr/Getting-Started-with-the-Rcmdr.pdf> Last accessed 3rd September 2014.
- Girolami, M. (2001). A Variational Method for Learning Sparse and Overcomplete Representations. *Neural Computation*, **13(11)**, 2517-2532.
- Gusenleitner, D., Howe, E.A., Betink, S., Quackenbush, J. and Culhane, A.C. (2012). iBBiG: Iterative Binary Bi-clustering of Gene Sets. *Bioinformatics*, **28**, 2484-2492.
- Gusenleitner, D. and Culhane, A. (2014). Package ‘iBBiG’. *Iterative Binary Biclustering of Genesets*. Available: <http://www.bioconductor.org/packages/release/bioc/manuals/iBBiG/man/iBBiG.pdf>. Last accessed 21th August 2014.
- Hochreiter, S., Bodenhofer, U., Heusel, M., Mayr, A., Mitterecker, A., Kasim, A., Khamiakova, T., Sanden, S.V., Lin, D., Talloen, W., Bijmans, L., Göhlmann, H.W.H., Shkedy, Z. and Clevert, D.-A. (2010). FABIA: Factor analysis for bicluster acquisition. *Bioinformatics*, **26(12)**, 1520-1527.
- Hochreiter, S. (2014). Package ‘fabia’. *FABIA: Factor Analysis for Bicluster Acquisition*. Available: <http://www.bioconductor.org/packages/release/bioc/manuals/fabia/man/fabia.pdf>. Last accessed 21th August 2014.
- Hochreiter, S. (2014). FABIA: Factor Analysis for Bicluster Acquisition -Manual for the R package-. Available: <http://www.bioconductor.org/packages/release/bioc/vignettes/fabia/inst/doc/fabia.pdf>. Last accessed 21th August 2014.

- Hoyer, P.O. (2004). Non-negative Matrix Factorization with Sparseness Constraints. *Journal of Machine Learning Research*, **5**, 1457-1469.
- Kaiser, S. and Leisch, F. (2008). A toolbox for bicluster analysis in R. *Tech. rep.*. Department of Statistics University of Munich.
- Kaiser, S. (2011). *Biclustering: Methods, Software and Application*. Ph.D. thesis, Ludwig-Maximilians-Universität München.
- Kaiser, S., Santamaria, R., Khamiakova, T., Sill, M., Theron, R., Quintales, L. and Leisch, F. (2014). Package ‘biclust’. *BiCluster Algorithms*. Available: <http://cran.r-project.org/web/packages/biclust/biclust.pdf>. Last accessed 21th August 2014.
- Khamiakova, T. (2013). Statistical Methods for Analysis of High Throughput Experiments in Early Drug Development. *The R Package superbiclust*. PhD Thesis. Hasselt University
- Khamiakova, T. (2014). Package ‘superbiclust’. *Generating Robust Biclusters from a Bicluster Set (Ensemble Biclustering)*. Available: <http://cran.r-project.org/web/packages/superbiclust/superbiclust.pdf>. Last accessed 21th August 2014.
- Kluger, Y., Basri, R., Chang, J.T. and Gerstein, M. (2003). Spectral biclustering of microarray data: Co-clustering genes and conditions. *Genome Research*, **13**, 703-716.
- Lazzeroni, L. and Owen, A. (2000). Plaid models for gene expression data. *Tech. rep.*, Stanford University.
- Madeira, S.C. and Oliveira, A.L. (2004). Biclustering Algorithm for Biological Data Analysis: A Survey. *IEEE/ACM Transactions, Computational Biology and Bioinformatics*. **1(1)**, 24-45
- Murali, T. and Kasif, S. (2003). Extracting conserved gene expression motifs from gene expression data. In: *Pacific Symposium on Biocomputing* **8**, 77-88.
- Palmer, J., Wipf, D., Kreutz-Delgado, K. and Rao, B. (2006). Variational EM algorithms for non-Gaussian latent variable models. *Advances in Neural Information Processing Systems*, **18**, 1059-1066.
- Pramana, S. (2011). Statistical Methods for Microarray Experiments: Analysis Dose-response studies and Software Development in R. PhD Thesis. UHasselt University.
- Prelić, A., Bleuler, S., Zimmermann, P., Wille, A., Bühlmann, P., Gruissem, W., Henning, L., Thiele, L. and Zitzler, E. (2006). A systematic comparison and evaluation of biclustering methods for gene expression data. *Bioinformatics*, **22**, 1122-1129. Available: <http://view.ncbi.nlm.nih.gov/pubmed/16500941>. Last accessed 21th August 2014.
- R Core Team. (2014). R: A language and environment for statistical computing. *R Foundation for Statistical Computing, Vienna, Austria*. Available: <http://www.R-project.org/>. Last accessed 2nd September 2014.
- Shi, F., Leckie, C., MacIntyre, G., Haviv, I., Boussioutas, A. and Kowalczyk, A. (2010). A bi-ordering approach to linking gene expression with clinical annotations in gastric cancer. *BMC Informations*, **11**, 477.
- Tanay, A., Sharan, R. and Shamir, R. (2004). Biclustering algorithms: A Survey. Available: www.cs.tau.ac.il/~roded/bicrev.ps. Last accessed 21th August 2014.
- Turner, H., Bailey, T. and Krzanowski, W. (2005). Improved biclustering of microarray data demonstrated through systematic performance tests. *Computational Statistics and Data Analysis*, **48**, 235-254.

6 Appendix

6.1 Introduction

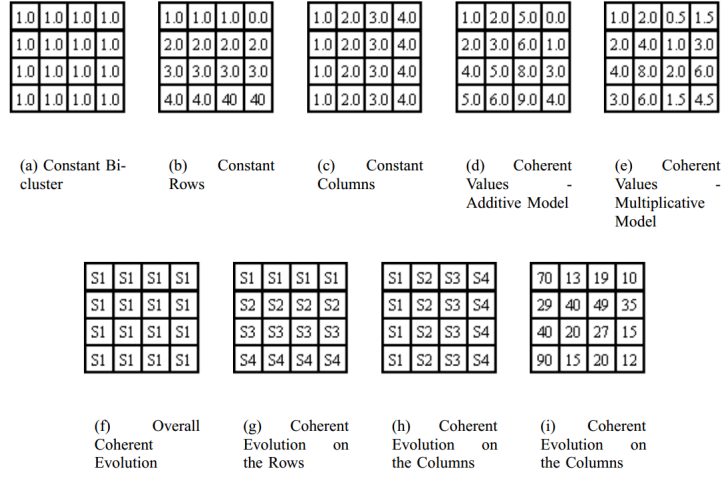


Figure 37: *Several Types of Biclusters*

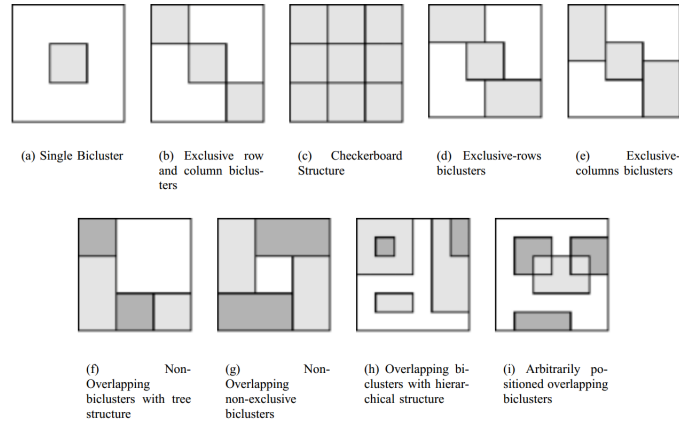


Figure 38: *Underlying Bicluster Structures*

6.2 The BiclustGUI R Package

6.2.1 biclust-package

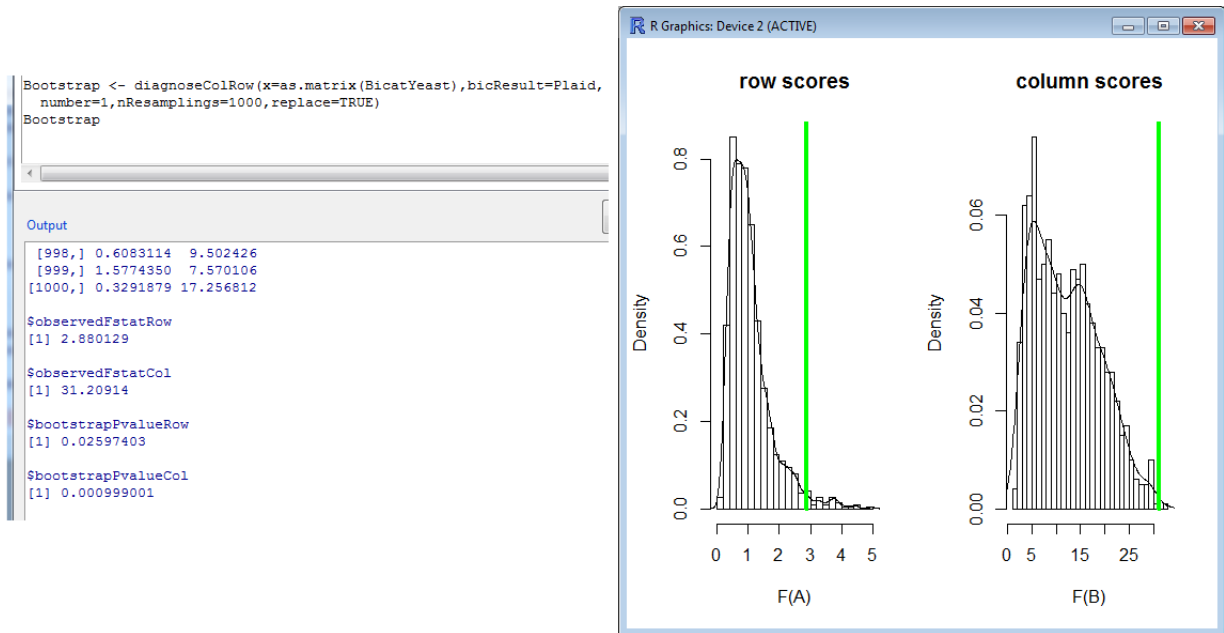


Figure 39: An example of bootstrap output & Histograms

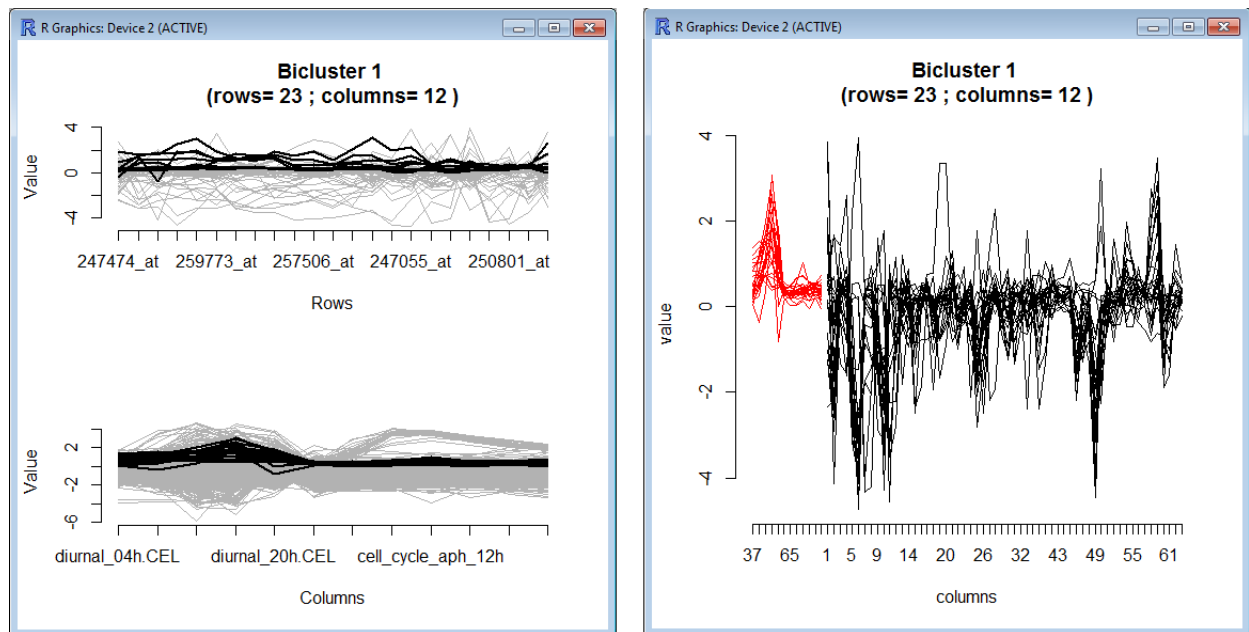


Figure 40: An example of Parallel Coordinates Plot - Default Type (row & column + Compare) & Combined Type

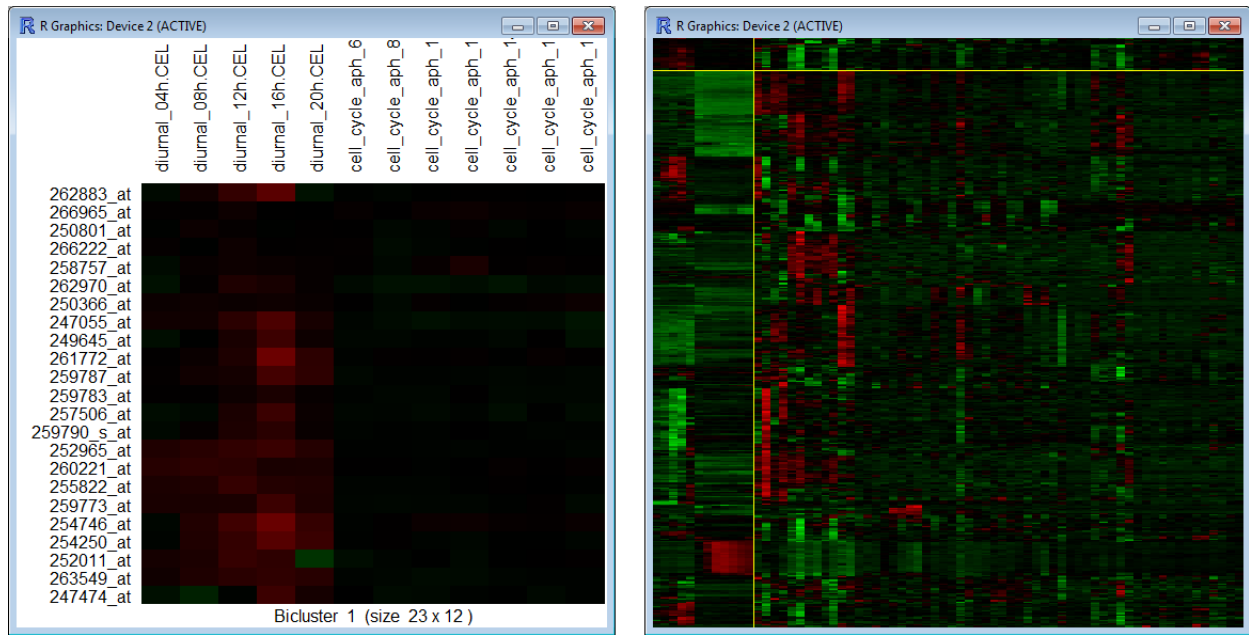


Figure 41: An example of the Heatmap - Local & Full Matrix

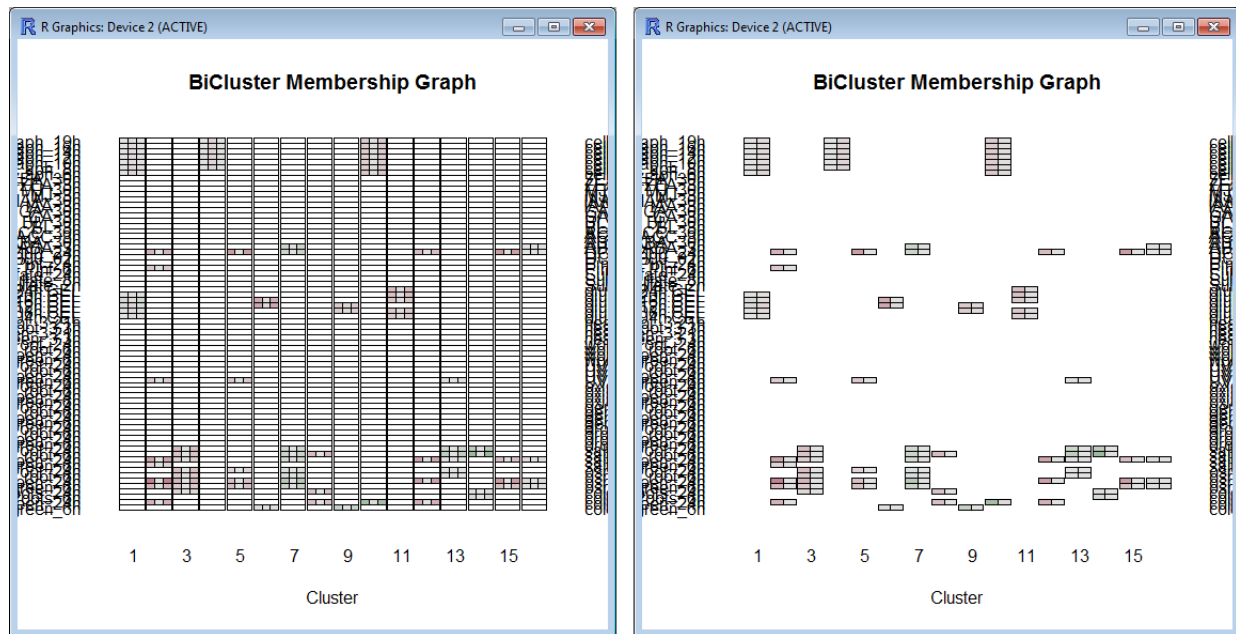


Figure 42: An example of BiclustMember Plot - Mid Checked & Mid Unchecked

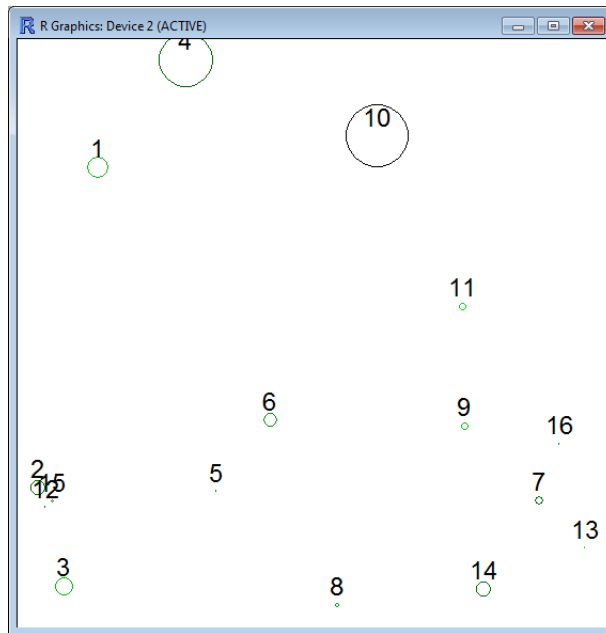


Figure 43: An example of Biclust Bubble Plot with Mean Projection

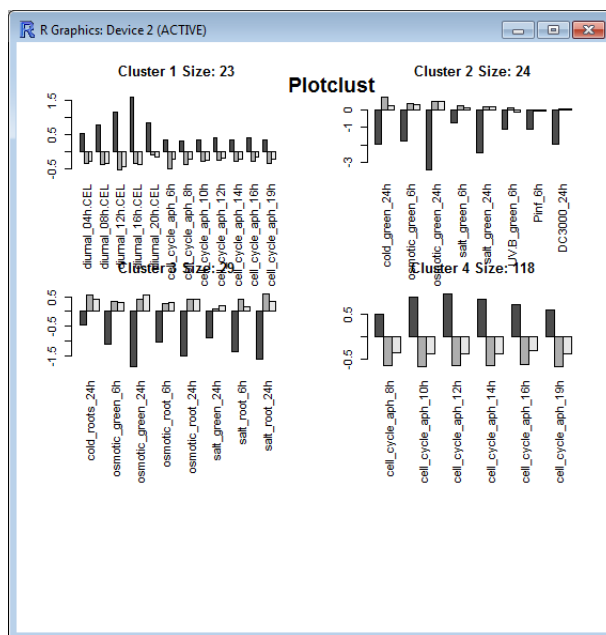


Figure 44: An example of the Barplots

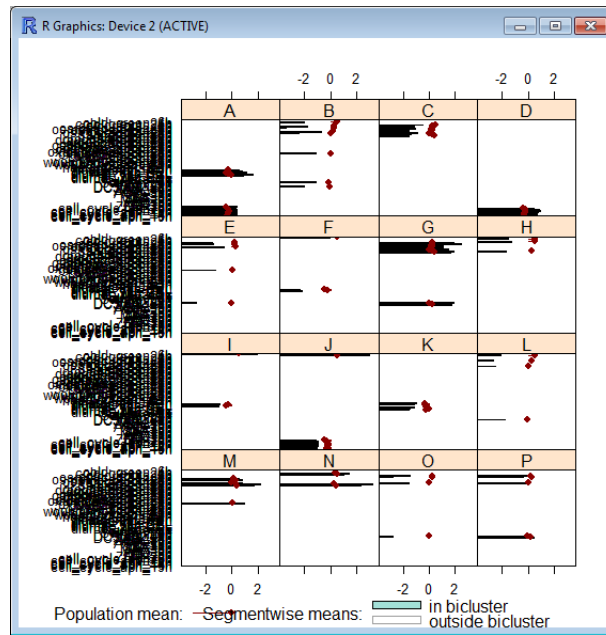


Figure 45: *An example of Bicluster Barchart*

6.2.2 fabia-package

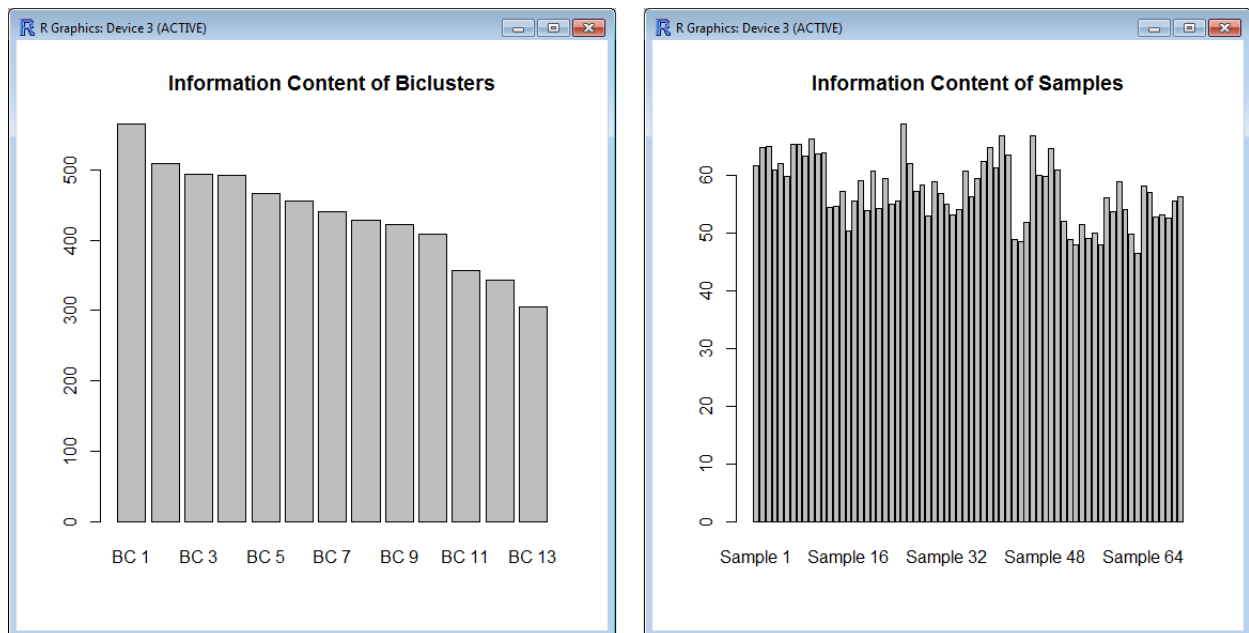


Figure 46: *Examples of Information Content Histogram of biclusters and samples*

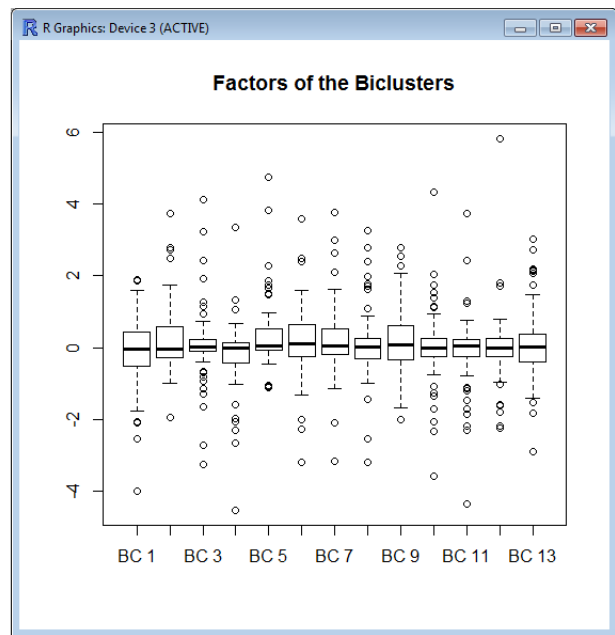
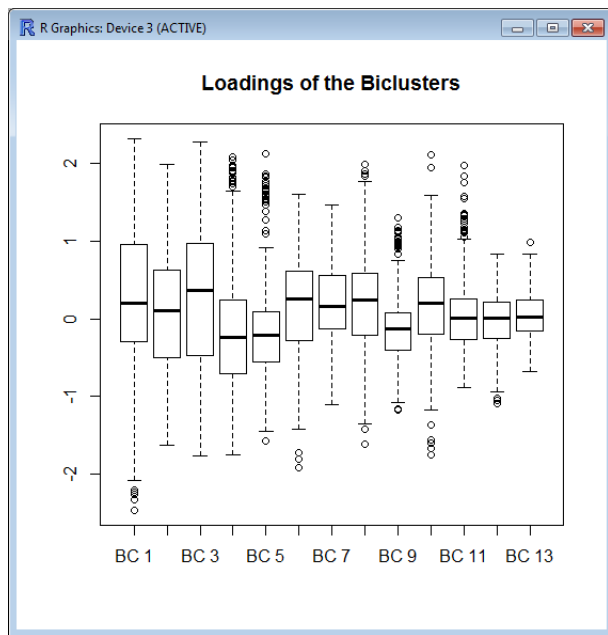


Figure 47: *Examples of boxplots of Loadings & Factors*

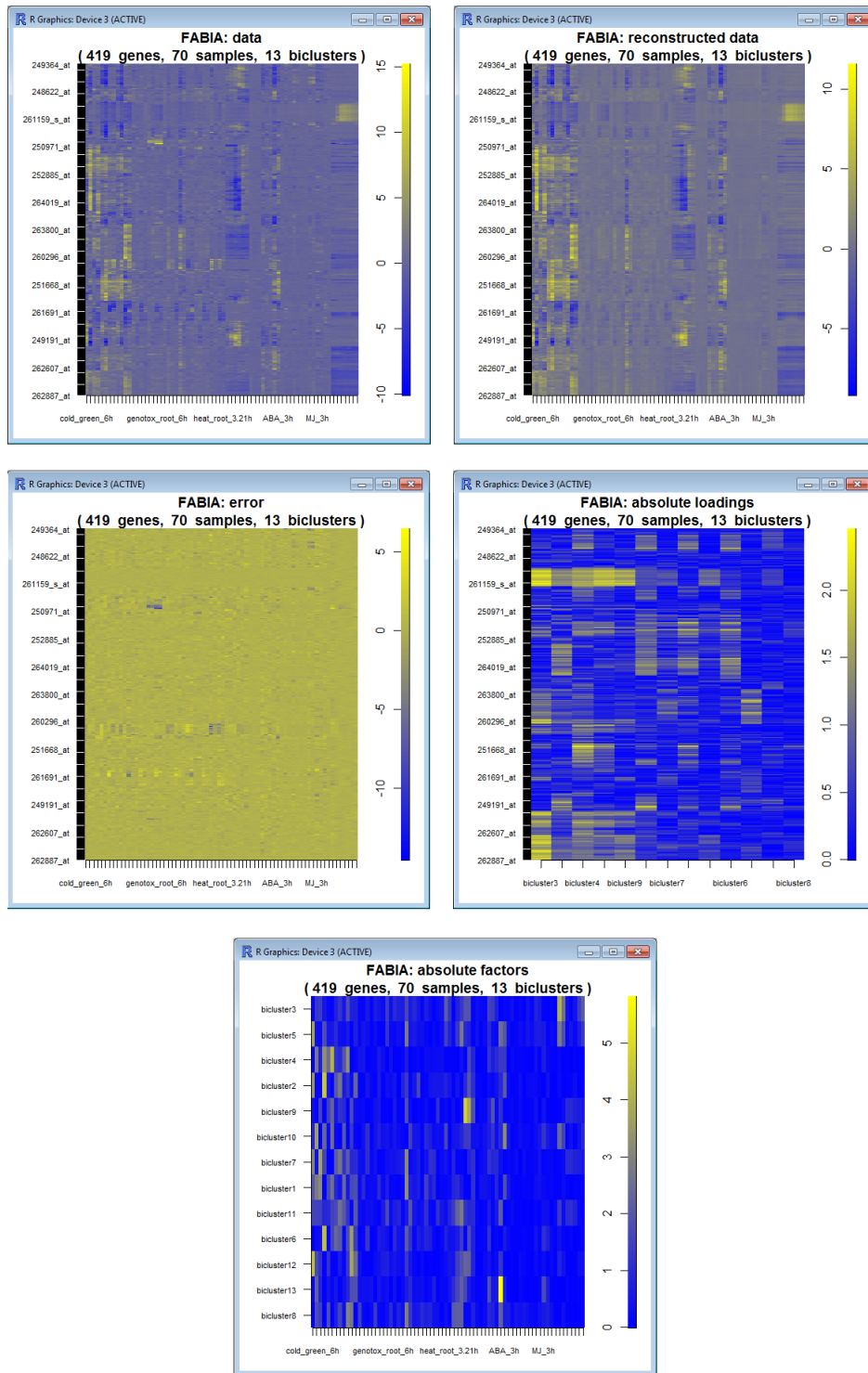


Figure 48: *Examples Extract Plots*

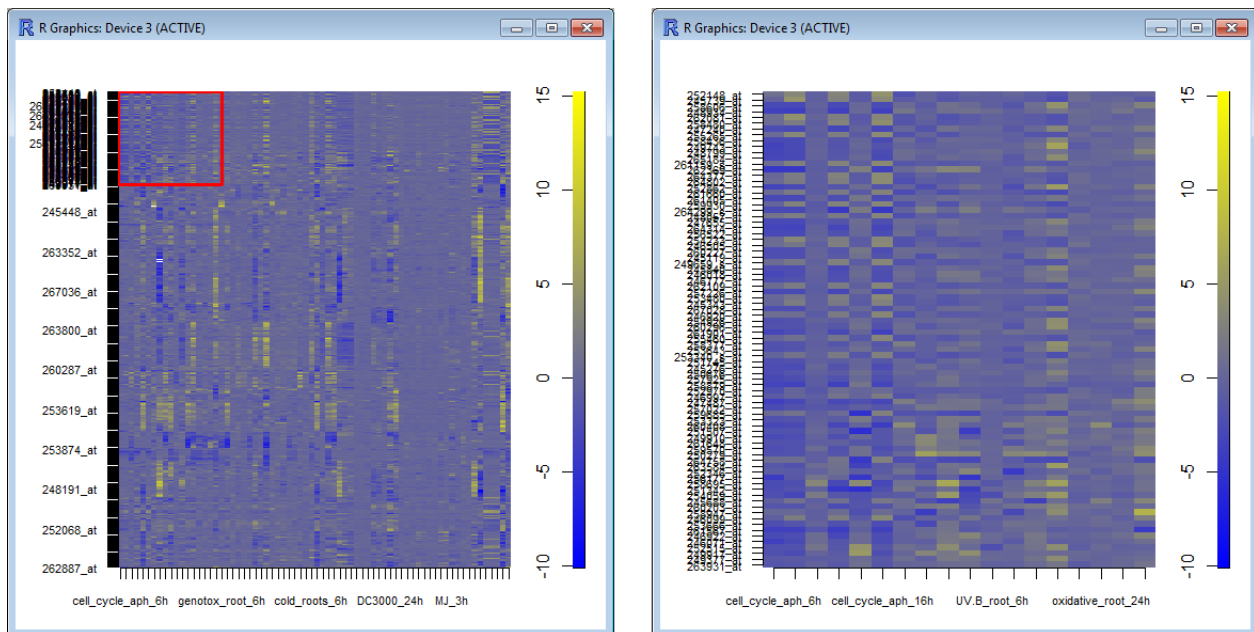


Figure 49: *Example of Bicluster Plot - Full & Only Bicluster*

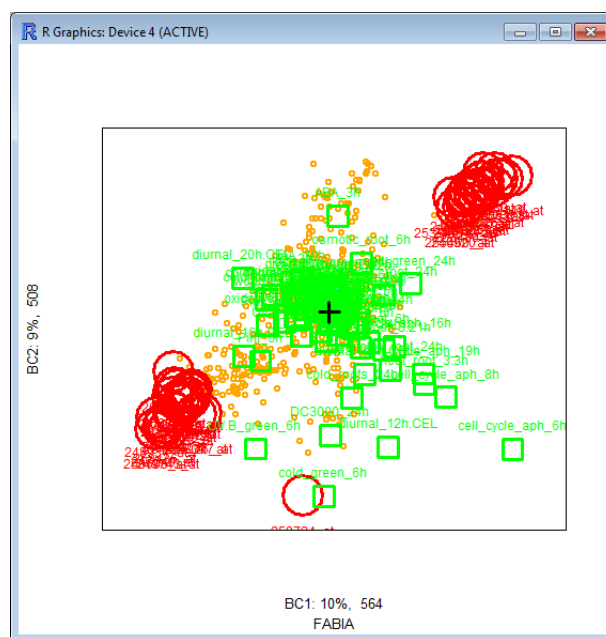


Figure 50: *Example of BiPlot*

6.2.3 isa2-package

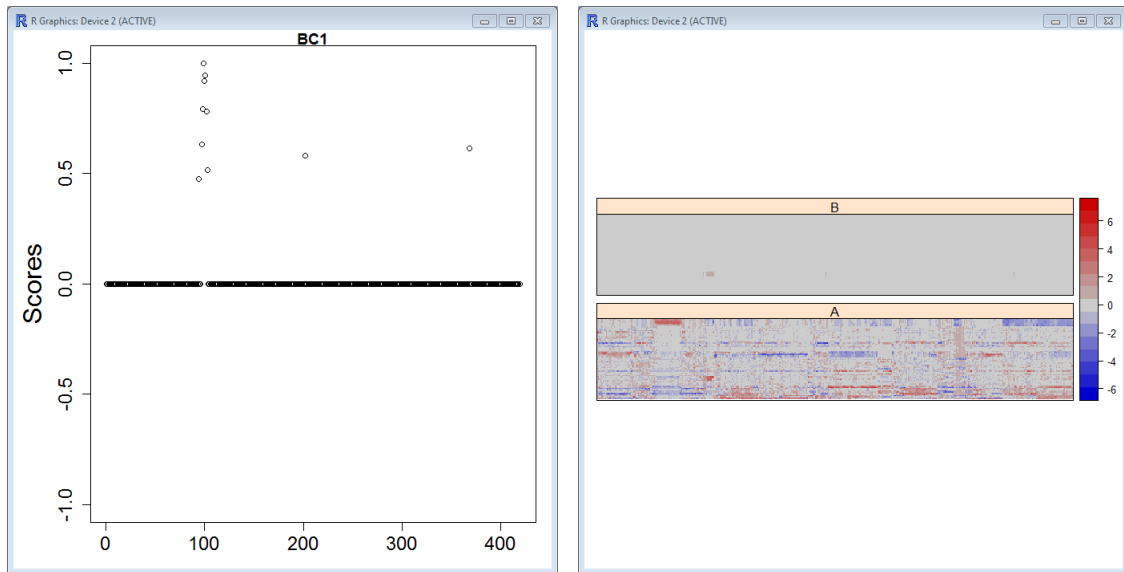


Figure 51: An example of score \mathcal{E} module plot

6.2.4 iBBiG-package

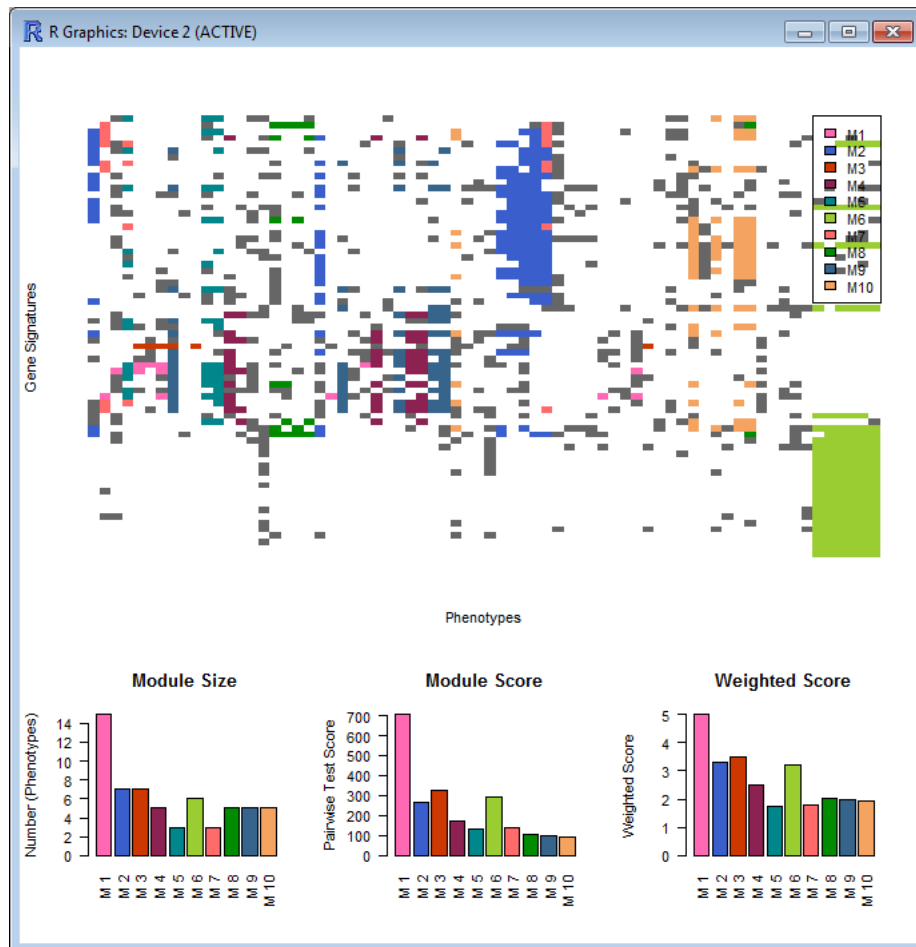


Figure 52: An example of the general iBBiG plot

6.2.5 BcDiag-package

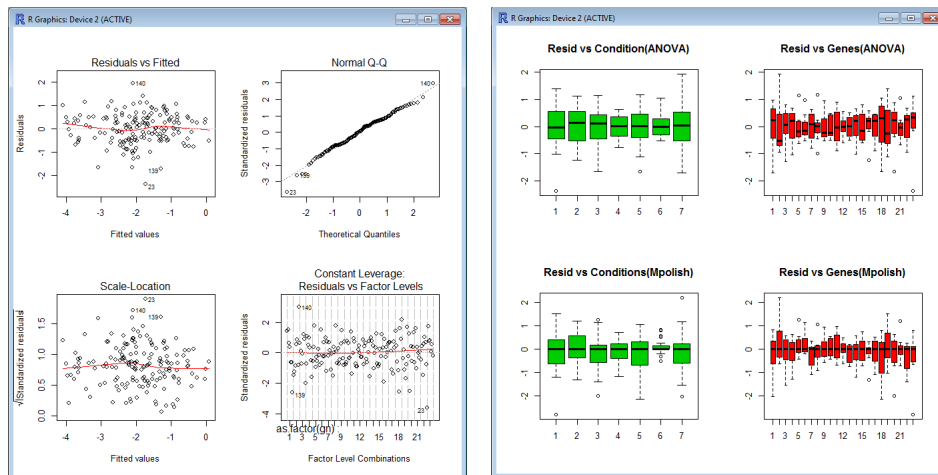


Figure 53: *ANOVA & Residual Plots: Diagnostic Plots and ANOVA & Mpolish*

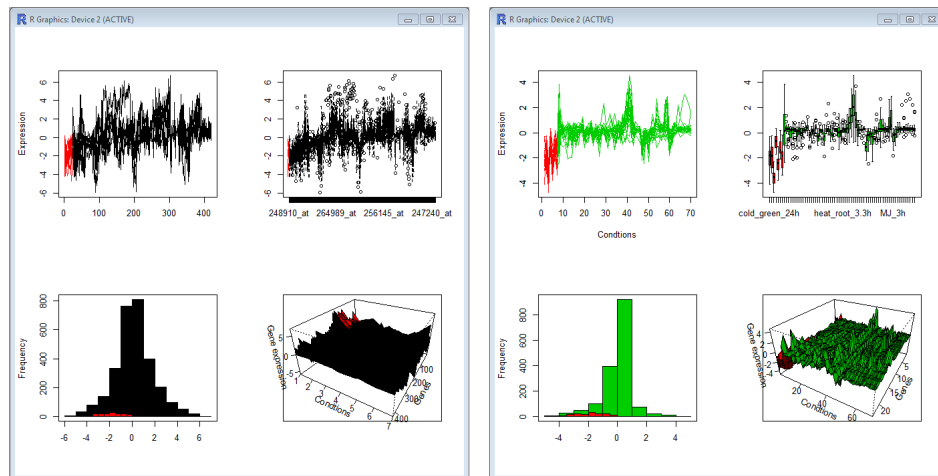


Figure 54: *Profile Plots: Example of all types for gene & condition dimension*

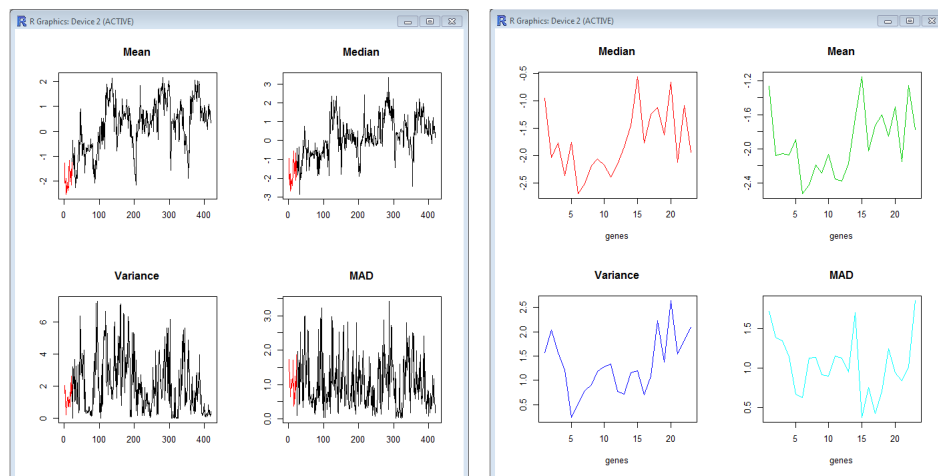


Figure 55: *Exploratory Plots: Left and right button for gene dimension*

6.2.6 superbiclust-package

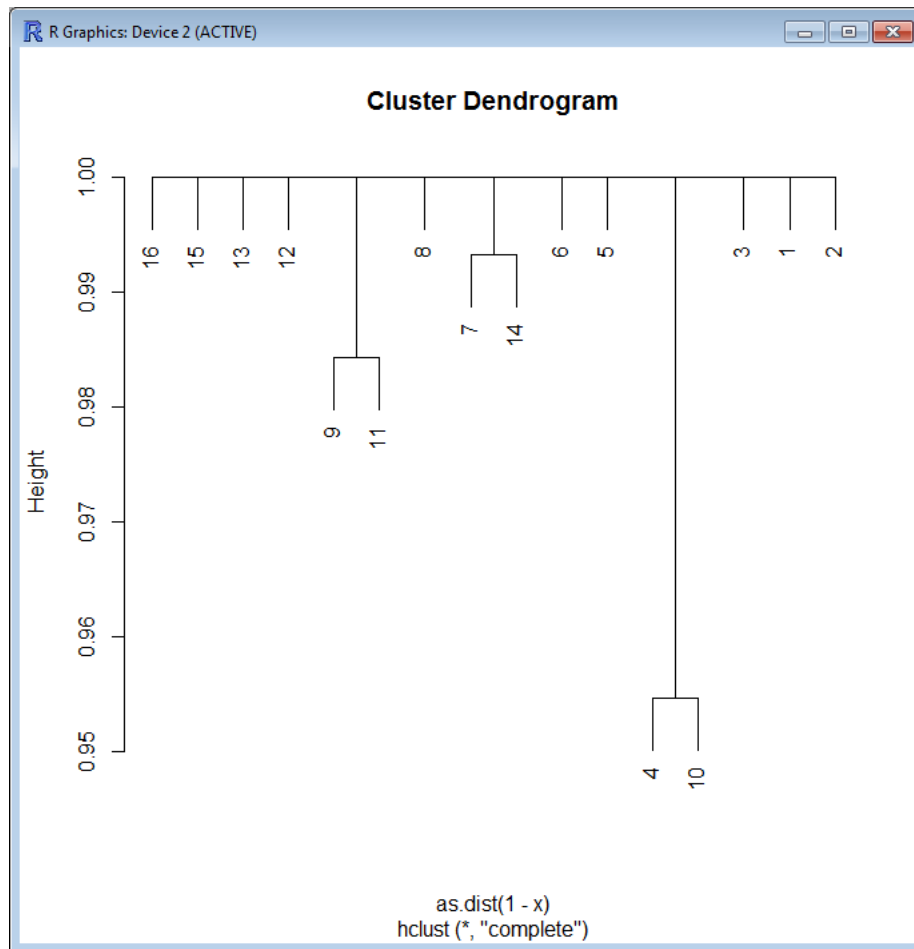


Figure 56: Example of a superbiclust tree (Note that this tree is only here for example purposes. All the biclusters here seem to be unique.)

```
> superbiclust.robust.GUI(CutTree=CutTree, show.which=TRUE)

Robust Biclusters created from more than 1 bicluster:
      <NA> <NA>
Robust Bicluster:   4   9
#BC's inside:      2   2

Which Biclusters Inside?
Robust BC 4 :   4 10
Robust BC 9 :   9 11
```

Figure 57: Example of Robust BC button with Show Inside checked

6.3 Guideline - Template Scripts

6.3.1 newmethod_script

This script can also be found in the doc subdirectory of the package.

```
newmethod_WINDOW <- function(){      # Change newmethod to your own method name

new.frames <- .initialize.new.frames()
grid.config <- .initialize.grid.config()
grid.rows <- .initialize.grid.rows()
```

```
#####
## GENERAL INFORMATION ABOUT THE NEW METHOD/WINDOW ##
#####

methodname <- "A new method"

methodfunction <- "methodfunction"
data.arg <- "d"
data.matrix <- FALSE
methodshow <- TRUE
other.arg <- ""
methodhelp <- ""

# Extra Data Conversion Boxes
data.discr <- FALSE
data.bin <- FALSE

# Possibility to give a seed ?
methodseed <- TRUE

## COMPATIBILITY? ##

# BcDiag
bcdiag.comp <- FALSE

# SuperBiclust
superbiclust.comp <- FALSE

#####
#### CLUSTERING TAB ####
#####

input <- "clusterTab"

### 1. ADDING THE FRAMES ###

# Add frames here

### 2. CONFIGURING THE GRID ###

grid.config <- .grid.matrix(input=input,c("frame1","frame2","frame3",NA
    ,"frame4",NA),nrow=3,ncol=2,byrow=TRUE,grid.config=grid.config)

### 3. COMBING THE ROWS ###

grid.rows <- .combine.rows(input=input,rows=c(1),title="A nice box: "
    ,border=TRUE,grid.rows=grid.rows,grid.config=grid.config)
grid.rows <- .combine.rows(input=input,rows=c(2,3),title="A nice box: "
    ,border=TRUE,grid.rows=grid.rows,grid.config=grid.config)

#####
#### PLOTTING & DIAGNOSTICS TAB ####
#####

input <- "plotdiagTab"

### 1. ADDING THE FRAMES ###
```

```

# Add frames here

### 2. CONFIGURING THE GRID ###

grid.config <- .grid.matrix(input=input,c("frame5","frame6"),nrow=1,
                             ncol=2,byrow=TRUE,grid.config=grid.config)

### 3. COMBINING THE ROWS ###

grid.rows <- .combine.rows(input=input,rows=c(1),title="Plot 1",
                             border=TRUE,grid.rows=grid.rows,grid.config=grid.config)

#####
## USE ALL THE ARGUMENTS ABOUT IN THE GENERAL CLUSTERTEMPLATE FUNCTION ##
#####

cluster_template(methodname=methodname,methodfunction=methodfunction,
                  methodhelp=methodhelp,data.arg=data.arg,other.arg=other.arg
                  ,methodseed=methodseed,grid.config=grid.config,grid.rows=grid.rows,
                  new.frames=new.frames,superbiclust.comp=superbiclust.comp,
                  bcdiag.comp=bcdiag.comp,data.matrix=data.matrix,data.discr=data.discr,
                  data.bin=data.bin,methodshow=methodshow)

}

```

6.3.2 frames_script

```

#### ENTRY FIELDS FRAME ####

type <- "entryfields"

# Change variables accordingly:
frame.name <- "entryframe1"
argument.names <- c("Argument 1","Argument 2","Argument 3")
argument.types <- c("num","num","char")
arguments <- c("arg1","arg2","arg3")
initial.values <- c(1,2,"a")
title <- "A Title"
border <- FALSE
entry.width <- c("2","2","6")

# Do not change this line:
new.frames <- .add.frame(input=input,type=type
                        ,frame.name=frame.name,argument.names=argument.names
                        ,arguments=arguments,initial.values=initial.values
                        ,title=title,border=border,entry.width=entry.width
                        ,argument.types=argument.types ,new.frames=new.frames)

#### RADIO BUTTONS FRAME ####

type <- "radiobuttons"

# Change variables accordingly:
frame.name <- "radioframe1"
argument.names <- c("Button 1","Button 2","Button 3")
arguments <- c("buttonarg")
argument.types <- "char"

```

```

argument.values <- c("b1","b2","b3")
initial.values <- "b3"
title <- "Button Options"
border <- TRUE

# DO NOT CHANGE THIS LINE:
new.frames <- .add.frame(input=input,type=type
  ,frame.name=frame.name,argument.names=argument.names
  ,arguments=arguments,argument.values=argument.values
  ,initial.values=initial.values,title=title,border=border
  ,new.frames=new.frames,argument.types=argument.types)

#### CHECK BOXES FRAME ####

type <- "checkboxes"

# Change variables accordingly:
frame.name <- "checkboxframe1"
argument.names <- c("Check 1","Check 2","Check 3")
arguments <- c("checkarg1","checkarg2","checkarg3")
initial.values <- c(0,1,1)
title <- "title"
border <- FALSE

# DO NOT CHANGE THIS LINE:
new.frames <- .add.frame(input=input,type=type
  ,frame.name=frame.name,argument.names=argument.names
  ,arguments=arguments,initial.values=initial.values
  ,title=title,border=border,new.frames=new.frames)

#### VALUE SLIDER FRAME ####

type <- "valuesliders"

# Change variables accordingly:
frame.name <- "sliderframe1"
argument.names <- c("Slider 1 ","Slider 2 ","Slider 3 ")
arguments <- c("sliderarg1","sliderarg2","sliderarg3")
initial.values <- c(1,5,10)
from <- c(1,1,1)
to <- c(5,50,500)
by <- c(1,10,50)
length <- c(50,100,150)
title <- "Title"
border <- TRUE

# DO NOT CHANGE THIS LINE:
new.frames <- .add.frame(input=input,type=type,
  title=title,border=border,frame.name=frame.name,
  argument.names=argument.names,arguments=arguments,
  initial.values=initial.values,from=from,to=to,by=by,
  length=length,new.frames=new.frames)

#### SPIN BOX FRAME ####

type <- "spinboxes"

# Change variables accordingly:

```

```

frame.name <- "spinboxframe1"
argument.names <- c("Spin Box 1: ", "Spin Box 2: ", "Spin Box 3: ")
arguments <- c("spinarg1", "spingarg2", "spingarg3")
initial.values <- c(5, 10, 20)
from <- c(1, 5, 10)
to <- c(10, 20, 30)
by <- c(1, 1, 1)
entry.width <- "2"
title <- "Spin Box !"
border <- TRUE

# DO NOT CHANGE THIS LINE:
new.frames <- .add.frame(input=input, type=type,
  frame.name=frame.name, argument.names=argument.names,
  arguments=arguments, initial.values=initial.values,
  from=from, to=to, by=by, entry.width=entry.width,
  title=title, border=border, new.frames=new.frames)

#### MANUAL BUTTONS FRAME ####

type <- "buttons"

# Change variables accordingly:
frame.name <- "buttonframe1"
button.name <- "Button 1"
button.function <- "buttonfunction"
button.data <- "d"
button.biclust <- "biclust"

arg.frames <- c("frame1", "frame2")

save <- TRUE
show <- TRUE
button.otherarg <- ""

# Do not change this line:
new.frames <- .add.frame(input=input, frame.name=frame.name,
  type=type, button.name=button.name,
  button.function=button.function, button.data=button.data,
  button.biclust=button.biclust, button.otherarg=button.otherarg,
  arg.frames=arg.frames, save=save, show=show, new.frames=new.frames)

```

6.3.3 Quick Example - Plaid

```

biclustplaid_WIN <- function(){

new.frames <- .initialize.new.frames()
grid.config <- .initialize.grid.config()
grid.rows <- .initialize.grid.rows()

#####
## GENERAL INFORMATION ABOUT THE NEW METHOD/WINDOW ##
#####

methodname <- "Plaid"
methodfunction <- "biclust"
data.arg <- "x"
data.matrix <- TRUE

```

```

other.arg <- ",method=BCPlaid()"
methodhelp <- "BCPlaid"

# Possibility to give a seed ?
methodseed <- TRUE
# Add a discretize box?
data.discr <- FALSE
# Add a binarize box?
data.bin <- FALSE

## COMPATIBILITY? ##

# BcDiag
bcdiag.comp <- TRUE

# SuperBiclust
superbiclust.comp <- TRUE

# Biclust only (Not for public use)
extrabiclustplot <- TRUE

#####
#### CLUSTERING TAB ####
#####

### 1. ADDING THE FRAMES ###

input <- "clusterTab"

#### RADIO BUTTONS FRAME ####

type <- "radiobuttons"

# Change variables accordingly:
frame.name <- "toclusterframe"
argument.names <- c("Rows","Columns","Rows & Columns")
arguments <- c("cluster")
argument.values <- c("r","c","b")
argument.types <- "char"
initial.values <- "b"
title <- "To Cluster"
border <- FALSE

# DO NOT CHANGE THIS LINE:
new.frames <- .add.frame(input=input,type=type,
  frame.name=frame.name,argument.names=argument.names,
  arguments=arguments,argument.values=argument.values,
  initial.values=initial.values,title=title,border=border,
  new.frames=new.frames,argument.types=argument.types)

##### ENTRY FIELDS FRAME #####

type <- "entryfields"

# Change variables accordingly:
frame.name <- "modelframe"
argument.names <- c("Model Formula")
argument.types <- c("num")
arguments <- c("fit.model")

```

```

initial.values <- c("y ~ m+a+b")
title <- "Model"
border <- FALSE
entry.width <- c("10")

# Do not change this line:
new.frames <- .add.frame(input=input,type=type,frame.name=frame.name,
  argument.names=argument.names,arguments=arguments,initial.values=
  initial.values,title=title,border=border,entry.width=entry.width,
  argument.types=argument.types ,new.frames=new.frames)

#### CHECK BOXES FRAME      ####

type <- "checkboxes"

# Change variables accordingly:
frame.name <- "backgroundcheckframe"
argument.names <- c("Background Layer?")
arguments <- c("background")
initial.values <- c(1)
title <- ""
border <- FALSE

# DO NOT CHANGE THIS LINE:
new.frames <- .add.frame(input=input,type=type,frame.name=
  frame.name,argument.names=argument.names,arguments=arguments,
  initial.values=initial.values,title=title,border=border,new.frames=
  new.frames)

##### ENTRY FIELDS FRAME #####

type <- "entryfields"

# Change variables accordingly:
frame.name <- "backgroundentryframe1"
argument.names <- c("Shuffle","Back Fit","Max Layes")
argument.types <- c("num","num","num")
arguments <- c("shuffle","back.fit","max.layers")
initial.values <- c(3,0,20)
title <- ""
border <- FALSE
entry.width <- c("2","2","2")

# Do not change this line:
new.frames <- .add.frame(input=input,type=type,frame.name=frame.name,
  argument.names=argument.names,arguments=arguments,initial.values=
  initial.values,title=title,border=border,entry.width=entry.width,
  argument.types=argument.types ,new.frames=new.frames)

##### ENTRY FIELDS FRAME #####

type <- "entryfields"

# Change variables accordingly:
frame.name <- "backgroundentryframe2"
argument.names <- c("Iteration Startup","Iteration Layer")
argument.types <- c("num","num")
arguments <- c("iter.startup","iter.layer")
initial.values <- c(5,10)

```



```

title <- ""
border <- FALSE
entry.width <- c("2","2")

# Do not change this line:
new.frames <- .add.frame(input=input,type=type,frame.name=frame.name,
  argument.names=argument.names,arguments=arguments,initial.values=
  initial.values,title=title,border=border,entry.width=entry.width,
  argument.types=argument.types ,new.frames=new.frames)

### 2. CONFIGURING THE GRID ###

grid.config <- .grid.matrix(input=input,c("toclusterframe","modelframe",
  "backgroundcheckframe",NA,"backgroundentryframe1","backgroundentryframe2"),
  byrow=TRUE,nrow=3,ncol=2,grid.config=grid.config)

### 3. COMBING THE ROWS ###
grid.rows <- .combine.rows(input=input,rows=c(1),title="Plaid
  Specifications",border=TRUE,grid.rows=grid.rows,grid.config=grid.config)
grid.rows <- .combine.rows(input=input,rows=c(2,3),title=
  "Layer Specifications",border=TRUE,grid.rows=grid.rows,
  grid.config=grid.config)

#####
#### PLOTTING & DIAGNOSTICS TAB ####
#####

input <- "plotdiagTab"

### 1. ADDING THE FRAMES ###

###                                     ###
## Parallel Coordinates Plot ##
###                                     ###

#### RADIO BUTTONS FRAME   ####

type <- "radiobuttons"

# Change variables accordingly:
frame.name <- "pplottypeframe"
argument.names <- c("Default","Combined (rows & columns)")
arguments <- c("type2")
argument.values <- c("default","combined")
argument.types <- "char"
initial.values <- "default"
title <- "Plot Type:"
border <- FALSE

# DO NOT CHANGE THIS LINE:
new.frames <- .add.frame(input=input,type=type,frame.name=frame.name,
  argument.names=argument.names,arguments=arguments,argument.values=

```

```

argument.values,initial.values=initial.values,title=title,border=border,
new.frames=new.frames,argument.types=argument.types)

```

```

#### CHECK BOXES FRAME      ####

```

```

type <- "checkboxes"

```

```

# Change variables accordingly:
frame.name <- "pplotcheckframe"
argument.names <- c("Plot Only Column","Plot Rows & Columns","Compare")
arguments <- c("plotcol","plotBoth","compare")
initial.values <- c(1,0,1)
title <- "Default Type Options:"
border <- FALSE

```

```

# DO NOT CHANGE THIS LINE:

```

```

new.frames <- .add.frame(input=input,type=type,frame.name=frame.name,
    argument.names=argument.names,arguments=arguments,initial.values=
    initial.values,title=title,border=border,new.frames=new.frames)

```

```

##### ENTRY FIELDS FRAME #####

```

```

type <- "entryfields"

```

```

# Change variables accordingly:
frame.name <- "pplotentryframe"
argument.names <- c("Biccluster Number")
argument.types <- c("num")
arguments <- c("number")
initial.values <- c(1)
title <- ""
border <- FALSE
entry.width <- c("2")

```

```

# Do not change this line:

```

```

new.frames <- .add.frame(input=input,type=type,frame.name=frame.name,
    argument.names=argument.names,arguments=arguments,initial.values=
    initial.values,title=title,border=border,entry.width=entry.width,
    argument.types=argument.types ,new.frames=new.frames)

```

```

####      MANUAL BUTTONS FRAME      ####

```

```

type <- "buttons"

```

```

# Change variables accordingly:
frame.name <- "parallelbuttonframe"
button.name <- "Draw Plot"
button.function <- "parallelCoordinates3"
button.data <- "x"
button.biclust <- "bicResult"
button.otherarg <- ""
save <- FALSE
arg.frames <- c("pplotcheckframe","pplotentryframe","pplottypeframe")

```

```

# Do not change this line:

```

```

new.frames <- .add.frame(input=input,frame.name=frame.name,save=save,
  type=type,button.name=button.name,button.otherarg=button.otherarg,
  button.function=button.function,button.data=button.data,button.biclust=
  button.biclust,arg.frames=arg.frames,new.frames=new.frames)

#####

###          ###
##      Heatmap Plot      ##
###          ###

#### CHECK BOXES FRAME      ####

type <- "checkboxes"

# Change variables accordingly:
frame.name <- "heatplotcheckframe"
argument.names <- c("Local")
arguments <- c("local")
initial.values <- c(1)
title <- ""
border <- FALSE

# DO NOT CHANGE THIS LINE:
new.frames <- .add.frame(input=input,type=type,frame.name=frame.name,
  argument.names=argument.names,arguments=arguments,initial.values=
  initial.values,title=title,border=border,new.frames=new.frames)

#####      ENTRY FIELDS FRAME      #####

type <- "entryfields"

# Change variables accordingly:
frame.name <- "heatplotentryframe"
argument.names <- c("Bicluster Number")
argument.types <- c("num")
arguments <- c("number")
initial.values <- c(1)
title <- ""
border <- FALSE
entry.width <- c("2")

# Do not change this line:
new.frames <- .add.frame(input=input,type=type,frame.name=frame.name,
  argument.names=argument.names,arguments=arguments,initial.values=
  initial.values,title=title,border=border,entry.width=entry.width,
  argument.types=argument.types ,new.frames=new.frames)

#####      MANUAL BUTTONS FRAME      #####

type <- "buttons"

# Change variables accordingly:
frame.name <- "heatbuttonframe"
button.name <- "Draw Plot"
button.function <- "drawHeatmap"
button.data <- "x"
button.biclust <- "bicResult"

```

```

save <- FALSE
arg.frames <- c("heatplotcheckframe","heatplotentryframe")

# Do not change this line: (without button.otherarg)
new.frames <- .add.frame(input=input,frame.name=frame.name,type=
  type,save=save,button.name=button.name,button.function=button.function,
  button.data=button.data,button.biclust=button.biclust,arg.frames=
  arg.frames,new.frames=new.frames)

#####

###          ###
##      Biclust Member Plot      ##
###          ###

#### CHECK BOXES FRAME      ####

type <- "checkboxes"

# Change variables accordingly:
frame.name <- "mplotcheckframe"
argument.names <- c("Mid")
arguments <- c("mid")
initial.values <- c(1)
title <- ""
border <- FALSE

# DO NOT CHANGE THIS LINE:
new.frames <- .add.frame(input=input,type=type,frame.name=frame.name,
  argument.names=argument.names,arguments=arguments,initial.values=
  initial.values,title=title,border=border,new.frames=new.frames)

#####      ENTRY FIELDS FRAME      #####

type <- "entryfields"

# Change variables accordingly:
frame.name <- "mplotentryframe"
argument.names <- c("Bicluster Label")
argument.types <- c("char")
arguments <- c("cl_label")
initial.values <- c("")
title <- ""
border <- FALSE
entry.width <- c("8")

# Do not change this line:
new.frames <- .add.frame(input=input,type=type,frame.name=frame.name,
  argument.names=argument.names,arguments=arguments,initial.values=
  initial.values,title=title,border=border,entry.width=entry.width,
  argument.types=argument.types ,new.frames=new.frames)

####      MANUAL BUTTONS FRAME - EXAMPLE      ####

type <- "buttons"

# Change variables accordingly:
frame.name <- "memberbuttonframe"
button.name <- "Draw Plot"
button.function <- "biclustmember"

```

```

button.data <- "x"
button.biclust <- "bicResult"
save <- FALSE
arg.frames <- c("mplotcheckframe","mplotentryframe")

# Do not change this line: (without button.otherarg)
new.frames <- .add.frame(input=input,frame.name=frame.name,save=save,
  type=type,button.name=button.name,button.function=button.function,
  button.data=button.data,button.biclust=button.biclust,arg.frames=
  arg.frames,new.frames=new.frames)

#####

###                ###
## Summary & Diagnostics Box ##
###                ###

####      MANUAL BUTTONS FRAME      ####

type <- "buttons"

# Change variables accordingly:
frame.name <- "summarybuttonframe"
button.name <- "Summary"
button.function <- "summary"
button.data <- ""
button.biclust <- "object"
arg.frames <- c()
save <- FALSE

# Do not change this line: (without button.otherarg)
new.frames <- .add.frame(input=input,frame.name=frame.name,save=save,
  type=type,button.name=button.name,button.function=button.function,
  button.data=button.data,button.biclust=button.biclust,arg.frames=
  arg.frames,new.frames=new.frames)

#####

####      MANUAL BUTTONS FRAME      ####

type <- "buttons"

# Change variables accordingly:
frame.name <- "fstatbuttonframe"
button.name <- "Obs. F Stat."
button.function <- "computeObservedFstat"
button.data <- "x"
button.biclust <- "bicResult"
arg.frames <- c("fstatentryframe")

# Do not change this line: (without button.otherarg)
new.frames <- .add.frame(input=input,frame.name=frame.name,type=type,
  button.name=button.name,button.function=button.function,button.data=
  button.data,button.biclust=button.biclust,arg.frames=arg.frames,
  new.frames=new.frames)

#####

```

```

##### ENTRY FIELDS FRAME #####

type <- "entryfields"

# Change variables accordingly:
frame.name <- "fstatentryframe"
argument.names <- c("Biccluster Number")
argument.types <- c("num")
arguments <- c("number")
initial.values <- c("1")
title <- ""
border <- FALSE
entry.width <- c("4")

# Do not change this line:
new.frames <- .add.frame(input=input,type=type,frame.name=frame.name,
  argument.names=argument.names,arguments=arguments,initial.values=
  initial.values,title=title,border=border,entry.width=entry.width,
  argument.types=argument.types ,new.frames=new.frames)

##### ENTRY FIELDS FRAME #####
#      #

type <- "entryfields"

# Change variables accordingly:
frame.name <- "bootstrapentryframe"
argument.names <- c("Biccluster Number","Number Bootstrap Replicates")
argument.types <- c("num","num")
arguments <- c("number","nResamplings")
initial.values <- c(1,100)
title <- "Bootstrap Options:"
border <- FALSE
entry.width <- c("4","4")

# Do not change this line:
new.frames <- .add.frame(input=input,type=type,frame.name=frame.name,
  argument.names=argument.names,arguments=arguments,initial.values=
  initial.values,title=title,border=border,entry.width=entry.width
  ,argument.types=argument.types ,new.frames=new.frames)

#### CHECK BOXES FRAME #####

type <- "checkboxes"

# Change variables accordingly:
frame.name <- "bootstrapreplacementframe"
argument.names <- c("With Replacement?")
arguments <- c("replace")
initial.values <- c(1)
title <- ""
border <- FALSE

# DO NOT CHANGE THIS LINE:
new.frames <- .add.frame(input=input,type=type,frame.name=frame.name,
  argument.names=argument.names,arguments=arguments,initial.values=
  initial.values,title=title,border=border,new.frames=new.frames)

```

```

####      MANUAL BUTTONS FRAME      ####

type <- "buttons"

# Change variables accordingly:
frame.name <- "bootstrapbuttonframe"
button.name <- "Bootstrap"
button.function <- "diagnoseColRow"
button.data <- "x"
button.biclust <- "bicResult"
arg.frames <- c("bootstrapentryframe","bootstrapreplacementframe")

# Do not change this line: (without button.otherarg)
new.frames <- .add.frame(input=input,frame.name=frame.name,type=type,
  button.name=button.name,button.function=button.function,button.data=
  button.data,button.biclust=button.biclust,arg.frames=arg.frames,
  new.frames=new.frames)

####      MANUAL BUTTONS FRAME      ####

type <- "buttons"

# Change variables accordingly:
frame.name <- "bootstrapvisualbuttonframe"
button.name <- "Visualize"
button.function <- "diagnosticPlot"
button.data <- ""
button.biclust <- ""
button.otherarg <- "bootstrapOutput=Bootstrap"
save <- FALSE
arg.frames <- c()

# Do not change this line: (without button.otherarg)
new.frames <- .add.frame(input=input,frame.name=frame.name,save=save,
  type=type,button.name=button.name,button.function=button.function,
  button.data=button.data,button.biclust=button.biclust,arg.frames=
  arg.frames,new.frames=new.frames,button.otherarg=button.otherarg)

### 2. CONFIGURING THE GRID ###

grid.config <- .grid.matrix(input=input,c("summarybuttonframe",
  "fstatentryframe","fstatbuttonframe","bootstrapentryframe",NA,NA,
  "bootstrapreplacementframe","bootstrapbuttonframe",
  "bootstrapvisualbuttonframe","pplottypeframe","pplotentryframe",
  "parallelbuttonframe","pplotcheckframe",NA,NA,"heatplotcheckframe",
  "heatplotentryframe","heatbuttonframe","mplotcheckframe",
  "mplotentryframe","memberbuttonframe"),byrow=TRUE,nrow=7,ncol=3,
  grid.config=grid.config)

### 3. COMBING THE ROWS ###

grid.rows <- .combine.rows(input=input,rows=c(1,2,3),title=
  "Summary & Diagnostics",border=TRUE,grid.rows=grid.rows,grid.config=
  grid.config)
grid.rows <- .combine.rows(input=input,rows=c(4,5),title=
  "Parallel Coordinate Plot",border=TRUE,grid.rows=grid.rows,grid.config=

```

```

    grid.config)
grid.rows <- .combine.rows(input=input,rows=c(6),title="Heatmap Plot"
    ,border=TRUE,grid.rows=grid.rows,grid.config=grid.config)
grid.rows <- .combine.rows(input=input,rows=c(7),title="Biclustmember
    Plot",border=TRUE,grid.rows=grid.rows,grid.config=grid.config)

#####
## USE ALL THE ARGUMENTS ABOUT IN THE GENERAL CLUSTERTEMPLATE FUNCTION ##
#####

cluster_template(methodname=methodname,methodfunction=methodfunction,
    methodhelp=methodhelp,data.arg=data.arg,other.arg=other.arg,
    methodseed=methodseed,grid.config=grid.config,grid.rows=grid.rows,
    new.frames=new.frames,superbiclust.comp=superbiclust.comp,bcdiag.comp=
    bcdiag.comp,data.matrix=data.matrix,data.discr=data.discr,data.bin=
    data.bin,extrabiclustplot=extrabiclustplot)

}

```

6.3.4 newtool_script

This script can also be found in the doc subdirectory of the package.

```

newtool_WINDOW <- function(methodname){

new.frames <- .initialize.new.frames()
grid.config <- .initialize.grid.config()
grid.rows <- .initialize.grid.rows()

#####
## GENERAL INFORMATION ABOUT THE NEW TOOL    ##
#####

toolname <- "A new tool"

toolhelp <- "helppage"

data.matrix <- TRUE

# Do not change this line:
input <- "plotdiagTab"

#####
## MAKING THE WINDOW ##
#####

### ADDING FRAMES ###

# Analogous to plotdiag tab.

### CONFIGURING GRID ###
grid.config <- .grid.matrix(input=input,c(),nrow=1,ncol=2,byrow=TRUE
    ,grid.config=grid.config)

### COMBINING ROWS ###
grid.rows <- .combine.rows(input=input,rows=c(),title="Plot 1",

```



```
border=TRUE,grid.rows=grid.rows,grid.config=grid.config)
```

```
#####
## USE ALL THE ARGUMENTS ABOUT IN THE GENERAL NEW TOOL FUNCTION ##
#####
```

```
newtool_template(toolname=toolname,methodname=methodname,toolhelp=toolhelp,
  data.matrix=data.matrix,grid.config=grid.config,grid.rows=grid.rows,
  new.frames=new.frames)
```

```
}
```

6.3.5 Quick Example - BcDiag

```
bcdiag_WINDOW <- function(methodname){
```

```
  new.frames <- .initialize.new.frames()
  grid.config <- .initialize.grid.config()
  grid.rows <- .initialize.grid.rows()
```

```
  # Some extra code to determine the input type: "biclust", "fabia", "isa2"
  biclust.names <- c("Bimax","CC","Plaid","Questmotif","Spectral",
    "XMotifs","IBBIG")
  fabia.names <- c("Fabia Laplace Prior","Fabia Post-Projection",
    "Fabia Sparseness Projection","Fabia SPARSE")
  isa.names <- c("ISA")
```

```
  if(methodname %in% biclust.names){
    extra.arg <- ",mname='biclust'"
  }
  if(methodname %in% fabia.names){
    extra.arg <- ",mname='fabia'"
  }
  if(methodname %in% isa.names){
    extra.arg <- ",mname='isa2'"
  }
```

```
#####
## GENERAL INFORMATION ABOUT THE NEW TOOL      ##
#####
```

```
toolname <- "BCDIAG"
toolhelp <- "BcDiag-package"
data.matrix <- TRUE
```

```
# Do not change this line:
input <- "plotdiagTab"
```

```
### ADDING FRAMES ###
```

```
####                                ####
## The anomedOnlybic Function ##
####                                ####
```

```
#### RADIO BUTTONS FRAME ####
```

```

type <- "radiobuttons"

# Change variables accordingly:
frame.name <- "radio_anomed"
argument.names <- c("Diagnostic Plots","Tukey Additivity Plot",
  "Anova Plots","Mpolish Plots","Anova & Mpolish")
arguments <- c("fit")
argument.values <- c("aplot","mplot","anovbplot","mpolishbplot",
  "boxplot")
argument.types <- "char"
initial.values <- "boxplot"
title <- "Plot Type:"
border <- FALSE

# DO NOT CHANGE THIS LINE:
new.frames <- .add.frame(input=input,type=type,frame.name=frame.name,
  argument.names=argument.names,arguments=arguments,argument.values=
  argument.values,initial.values=initial.values,title=title,border=border
  ,new.frames=new.frames,argument.types=argument.types)

##### ENTRY FIELDS FRAME #####

type <- "entryfields"

# Change variables accordingly:
frame.name <- "entry_number"
argument.names <- c("Biccluster Number")
argument.types <- c("num")
arguments <- c("bnum")
initial.values <- c(1)
title <- ""
border <- FALSE
entry.width <- c("3")

# Do not change this line:
new.frames <- .add.frame(input=input,type=type,frame.name=
  frame.name,argument.names=argument.names,arguments=arguments,
  initial.values=initial.values,title=title,border=border,
  entry.width=entry.width,argument.types=argument.types ,
  new.frames=new.frames)

#### MANUAL BUTTONS FRAME ####

type <- "buttons"

# Change variables accordingly:
frame.name <- "button_anomed"
button.name <- "Draw Plot"
button.function <- "anomedOnlybic"
button.data <- "dset"
button.biclust <- "bres"
button.otherarg <- extra.arg
save <- FALSE
arg.frames <- c("radio_anomed","entry_number")

# Do not change this line:
new.frames <- .add.frame(input=input,frame.name=frame.name,
  save=save,type=type,button.name=button.name,button.otherarg=
  button.otherarg,button.function=button.function,button.data=

```

```

button.data,button.biclust=button.biclust,arg.frames=arg.frames,
new.frames=new.frames)

```

```

####                                     ####
##      EXPLORATORY PLOTS (biclustered & Clustered Data)      ##
####                                     ####

#### RADIO BUTTONS FRAME      ####

type <- "radiobuttons"

# Change variables accordingly:
frame.name <- "radio1_explore"
argument.names <- c("Mean","Median","Variance","Median Absolute
  Deviation","Quantile","All")
arguments <- c("pfor")
argument.values <- c("mean","median","variance","mad","quant","all")
argument.types <- "char"
initial.values <- "mean"
title <- "Plot for:"
border <- FALSE

# DO NOT CHANGE THIS LINE:
new.frames <- .add.frame(input=input,type=type,frame.name=frame.name,
  argument.names=argument.names,arguments=arguments,argument.values=
  argument.values,initial.values=initial.values,title=title,border=border
  ,new.frames=new.frames,argument.types=argument.types)

#### RADIO BUTTONS FRAME      ####

type <- "radiobuttons"

# Change variables accordingly:
frame.name <- "radio2_explore"
argument.names <- c("Genes","Conditions")
arguments <- c("gby")
argument.values <- c("genes","conditions")
argument.types <- "char"
initial.values <- "genes"
title <- "Dimension:"
border <- FALSE

# DO NOT CHANGE THIS LINE:
new.frames <- .add.frame(input=input,type=type,frame.name=frame.name,
  argument.names=argument.names,arguments=arguments,argument.values=
  argument.values,initial.values=initial.values,title=title,border=
  border,new.frames=new.frames,argument.types=argument.types)

####      MANUAL BUTTONS FRAME      ####

type <- "buttons"

# Change variables accordingly:
frame.name <- "button_explore"
button.name <- "Bicl. & Clust."
button.function <- "exploreBic"
button.data <- "dset"
button.biclust <- "bres"

```

```

button.otherarg <- extra.arg
save <- FALSE
arg.frames <- c("radio1_explore","radio2_explore","entry_number")

# Do not change this line:
new.frames <- .add.frame(input=input,frame.name=frame.name,save=
  save,type=type,button.name=button.name,button.otherarg=button.otherarg,
  button.function=button.function,button.data=button.data,button.biclust=
  button.biclust,arg.frames=arg.frames,new.frames=new.frames)

####      MANUAL BUTTONS FRAME      ####

type <- "buttons"

# Change variables accordingly:
frame.name <- "button2_explore"
button.name <- "Only Biclust."
button.function <- "exploreOnlybic"
button.data <- "dset"
button.biclust <- "bres"
button.otherarg <- extra.arg
save <- FALSE
arg.frames <- c("radio1_explore","radio2_explore","entry_number")

# Do not change this line:
new.frames <- .add.frame(input=input,frame.name=frame.name,
  save=save,type=type,button.name=button.name,button.otherarg=
  button.otherarg,button.function=button.function,button.data=
  button.data,button.biclust=button.biclust,arg.frames=arg.frames,
  new.frames=new.frames)

####      PROFILE PLOTS      ####
##      PROFILE PLOTS      ##
####      ####

#### RADIO BUTTONS FRAME      ####

type <- "radiobuttons"

# Change variables accordingly:
frame.name <- "radio1_profile"
argument.names <- c("All","Lines","Boxplot","Histogram","3D")
arguments <- c("bplot")
argument.values <- c("all","lines","boxplot","histogram","threeD")
argument.types <- "char"
initial.values <- "all"
title <- "Plot Type:"
border <- FALSE

# DO NOT CHANGE THIS LINE:
new.frames <- .add.frame(input=input,type=type,frame.name=frame.name,
  argument.names=argument.names,arguments=arguments,argument.values=
  argument.values,initial.values=initial.values,title=title,border=border
  ,new.frames=new.frames,argument.types=argument.types)

#### RADIO BUTTONS FRAME      ####

type <- "radiobuttons"

# Change variables accordingly:

```

```

frame.name <- "radio2_profile"
argument.names <- c("Genes","Conditions")
arguments <- c("gby")
argument.values <- c("genes","conditions")
argument.types <- "char"
initial.values <- "genes"
title <- "Dimension:"
border <- FALSE

# DO NOT CHANGE THIS LINE:
new.frames <- .add.frame(input=input,type=type,frame.name=frame.name
,argument.names=argument.names,arguments=arguments,argument.values=
argument.values,initial.values=initial.values,title=title,border=
border,new.frames=new.frames,argument.types=argument.types)

#### VALUE SLIDER FRAME - EXAMPLE ####
# Only for numerical values
type <- "valuesliders"

# Change variables accordingly:
frame.name <- "slider_profile"
argument.names <- c("Theta: ","Phi: ")
arguments <- c("teta","ph")
initial.values <- c(120,30)
from <- c(-180,-180)
to <- c(180,180)
by <- c(10,10)
length <- c(125,125)
title <- "3D Rotation:"
border <- FALSE

# DO NOT CHANGE THIS LINE:
new.frames <- .add.frame(input=input,type=type,title=title,
border=border,frame.name=frame.name,argument.names=argument.names,
arguments=arguments,initial.values=initial.values,from=from,to=to,
by=by,length=length,new.frames=new.frames)

#### MANUAL BUTTONS FRAME ####

type <- "buttons"

# Change variables accordingly:
frame.name <- "button_profile"
button.name <- "Draw Plot"
button.function <- "profileBic"
button.data <- "dset"
button.biclust <- "bres"
button.otherarg <- extra.arg
save <- FALSE
arg.frames <- c("radio1_profile","radio2_profile","entry_number",
"slider_profile")

# Do not change this line:
new.frames <- .add.frame(input=input,frame.name=frame.name,
save=save,type=type,button.name=button.name,button.otherarg=
button.otherarg,button.function=button.function,button.data=

```

```

        button.data,button.biclust=button.biclust,arg.frames=arg.frames,
        new.frames=new.frames)

####      MANUAL BUTTONS FRAME      ####

type <- "buttons"

# Change variables accordingly:
frame.name <- "button_summaryoutput"
button.name <- "Summary"
button.function <- "bcdiagwrite_WINDOW"
button.data <- ""
button.biclust <- ""

temp <- paste("methodname=",methodname,"",sep="")
button.otherarg <- temp
save <- FALSE
show <- FALSE
arg.frames <- c()

# Do not change this line:
new.frames <- .add.frame(input=input,frame.name=frame.name,
        show=show,save=save,type=type,button.name=button.name,
        button.otherarg=button.otherarg,button.function=button.function,
        button.data=button.data,button.biclust=button.biclust,arg.frames=
        arg.frames,new.frames=new.frames)

### CONFIGURING GRID ###

grid.config <- .grid.matrix(input=input,c("entry_number",
        "button_summaryoutput",NA,"radio_anomed","button_anomed",NA,
        "radio1_explore","radio2_explore",NA,"button_explore","button2_explore"
        ,NA,"radio1_profile","radio2_profile","slider_profile","button_profile"
        ,NA,NA),byrow=TRUE,nrow=6,ncol=3,grid.config=grid.config)

### COMBINING ROWS ###
grid.rows <- .combine.rows(input=input,rows=c(1),title=
        "Plotting Cluster Number & Summary Output",border=TRUE,grid.rows=
        grid.rows,grid.config=grid.config)
grid.rows <- .combine.rows(input=input,rows=c(2),title=
        "ANOVA & Median Polish Residual Plots",border=TRUE,grid.rows=grid.rows,
        grid.config=grid.config)
grid.rows <- .combine.rows(input=input,rows=c(3,4),title=
        "Exploratory Plots for Biclustered & Clustered data",border=TRUE,
        grid.rows=grid.rows,grid.config=grid.config)
grid.rows <- .combine.rows(input=input,rows=c(5,6),title=
        "Profile Plots for Biclustered & Clustered data",border=TRUE,
        grid.rows=grid.rows,grid.config=grid.config)

#####
## USE ALL THE ARGUMENTS ABOUT IN THE GENERAL NEW TOOL FUNCTION ##
#####

newtool_template(toolname=toolname,methodname=methodname,toolhelp=
        toolhelp,data.matrix=data.matrix,grid.config=grid.config,grid.rows=
        grid.rows,new.frames=new.frames)
}

```

6.3.6 Extra Example - Superbiclust

```
superbiclust_WINDOW <- function(methodname){

new.frames <- .initialize.new.frames()
grid.config <- .initialize.grid.config()
grid.rows <- .initialize.grid.rows()


method_result <- gsub(" ", "", methodname, fixed=TRUE)
method_result <- gsub("-", "", method_result, fixed=TRUE)


biclust.names <- c("Bimax", "CC", "Plaid", "Questmotif", "Spectral",
  "XMotifs", "IBBIG")
fabia.names <- c("Fabia Laplace Prior", "Fabia Post-Projection",
  "Fabia Sparseness Projection", "Fabia SPARSE")
isa.names <- c("ISA")


if(methodname %in% biclust.names){
#extra.arg <- paste("x=", method_result, sep="")
extra.arg <- paste("method_result=", method_result, "", sep="")
#isa2 <- FALSE
biclust.combine <- TRUE
make.save.button <- TRUE
save.type <- "biclust"
}
if(methodname %in% fabia.names){
#extra.arg <- paste("x=", method_result, sep="")
extra.arg <- paste("method_result=", method_result, "", sep="")
#isa2 <- FALSE
biclust.combine <- FALSE
make.save.button <- FALSE
save.type <- "fabia"
}
if(methodname %in% isa.names){
#extra.arg <- paste("x=isa.biclust", method_result, "", sep="")
extra.arg <- paste("method_result=", method_result, "", isa2=TRUE",
  , sep="")
#isa2 <- TRUE
biclust.combine <- FALSE
make.save.button <- TRUE
save.type <- "isa"
}


#####
## GENERAL INFORMATION ABOUT THE NEW TOOL ##
#####


toolname <- "SuperBiclust"

toolhelp <- "superbiclust"

data.matrix <- TRUE

# Do not change this line:
input <- "plotdiagTab"

#####
## MAKING THE WINDOW ##
```

```
#####

### ADDING FRAMES ###

####          ####
## EXTRA DATA INPUT ##
#####          #####

##### ENTRY FIELDS FRAME #####

type <- "entryfields"

# Change variables accordingly:
frame.name <- "biclustcombine"
argument.names <- c("")
argument.types <- c("num")
arguments <- c("extra.biclust")
initial.values <- c("NULL")
title <- "Vector of Biclust Objects (Example: c('biclust2','biclust3'))"
border <- FALSE
entry.width <- c("50")

# Do not change this line:
new.frames <- .add.frame(input=input,type=type,frame.name=
  frame.name,argument.names=argument.names,arguments=arguments,
  initial.values=initial.values,title=title,border=border,entry.width=
  entry.width,argument.types=argument.types ,new.frames=new.frames)

####          ####
## SUPERBICLUST CONFIGURATION ##
####          #####

#### RADIO BUTTONS FRAME #####

type <- "radiobuttons"

# Change variables accordingly:
frame.name <- "indexradio"
argument.names <- c("Jaccard","Sorensen","Ochiai","Kulczynski",
  "Sensitivity","Specificity")
arguments <- c("index")
argument.values <- c("jaccard","sorensen","ochiai","kulczynski",
  "sensitivity","specificity")
argument.types <- "char"
initial.values <- "jaccard"
title <- "Similarity?"
border <- FALSE

# DO NOT CHANGE THIS LINE:
new.frames <- .add.frame(input=input,type=type,frame.name=frame.name,
  argument.names=argument.names,arguments=arguments,argument.values=
  argument.values,initial.values=initial.values,title=title,border=border
  ,new.frames=new.frames,argument.types=argument.types)

#### RADIO BUTTONS FRAME #####

type <- "radiobuttons"
```



```

# Change variables accordingly:
frame.name <- "typeradio"
argument.names <- c("Rows","Columns","Both")
arguments <- c("type")
argument.values <- c("rows","cols","both")
argument.types <- "char"
initial.values <- "both"
title <- "Type?"
border <- FALSE

# DO NOT CHANGE THIS LINE:
new.frames <- .add.frame(input=input,type=type,frame.name=frame.name,
  argument.names=argument.names,arguments=arguments,argument.values=
  argument.values,initial.values=initial.values,title=title,border=
  border,new.frames=new.frames,argument.types=argument.types)

####      MANUAL BUTTONS FRAME      ####

type <- "buttons"

# Change variables accordingly:
frame.name <- "superbiclust"
button.name <- "SuperBiclust"
button.function <- "superbiclust.GUI"
button.data <- ""
button.biclust <- "x"
button.otherarg <- extra.arg
save <- FALSE
show <- FALSE

if(biclust.combine==TRUE){
arg.frames <- c("indexradio","typeradio","biclustcombine")
}
else{
arg.frames <- c("indexradio","typeradio")
}

# Do not change this line:
new.frames <- .add.frame(input=input,frame.name=frame.name,show=
  show,save=save,type=type,button.name=button.name,button.otherarg=
  button.otherarg,button.function=button.function,button.data=
  button.data,button.biclust=button.biclust,arg.frames=arg.frames,
  new.frames=new.frames)

####      ####
## DENDOGRAM ##
####      ####

####      MANUAL BUTTONS FRAME      ####

type <- "buttons"

# Change variables accordingly:
frame.name <- "drawtree"
button.name <- "Draw Tree"
button.function <- "plot"
button.data <- ""
button.biclust <- ""
button.otherarg <- "x=superbiclust.tree"

```

```

save <- FALSE
show <- TRUE
arg.frames <- c()

# Do not change this line:
new.frames <- .add.frame(input=input,frame.name=frame.name,show=show,
  save=save,type=type,button.name=button.name,button.otherarg=
  button.otherarg,button.function=button.function,button.data=button.data,
  button.biclust=button.biclust,arg.frames=arg.frames,new.frames=new.frames)

##### ENTRY FIELDS FRAME #####

type <- "entryfields"

# Change variables accordingly:
frame.name <- "cutentry"
argument.names <- c("Number of Biclusters","Height")
argument.types <- c("num","num")
arguments <- c("k","h")
initial.values <- c("NULL","NULL")
title <- "Where to cut tree? (number overrides height)"
border <- FALSE
entry.width <- c("6","6")

# Do not change this line:
new.frames <- .add.frame(input=input,type=type,frame.name=frame.name,
  argument.names=argument.names,arguments=arguments,initial.values=
  initial.values,title=title,border=border,entry.width=entry.width,
  argument.types=argument.types ,new.frames=new.frames)

#### MANUAL BUTTONS FRAME ####

type <- "buttons"

# Change variables accordingly:
frame.name <- "cuttree"
button.name <- "Cut Tree"
button.function <- "cuttree"
button.data <- ""
button.biclust <- ""
button.otherarg <- "tree=superbiclust.tree"
save <- TRUE
show <- TRUE
arg.frames <- c("cutentry")

# Do not change this line:
new.frames <- .add.frame(input=input,frame.name=frame.name,show=show,
  save=save,type=type,button.name=button.name,button.otherarg=
  button.otherarg,button.function=button.function,button.data=button.data,
  button.biclust=button.biclust,arg.frames=arg.frames,new.frames=new.frames)

#####

## Plotting Robust Biclusters ##
#####

#### CHECK BOXES FRAME ####

```

```

type <- "checkboxes"

# Change variables accordingly:
frame.name <- "showrobust"
argument.names <- c("Show Inside Robust BC?")
arguments <- c("show.which")
initial.values <- c(0)
title <- ""
border <- FALSE

# DO NOT CHANGE THIS LINE:
new.frames <- .add.frame(input=input,type=type,frame.name=frame.name,
  argument.names=argument.names,arguments=arguments,initial.values=
  initial.values,title=title,border=border,new.frames=new.frames)

####      MANUAL BUTTONS FRAME      ####

type <- "buttons"

# Change variables accordingly:
frame.name <- "robustbutton"
button.name <- "Robust BC"
button.function <- "superbiclust.robust.GUI"
button.data <- ""
button.biclust <- ""
button.otherarg <- "CutTree=CutTree"
save <- FALSE
show <- TRUE
arg.frames <- c("showrobust")

# Do not change this line:
new.frames <- .add.frame(input=input,frame.name=frame.name,
  show=show,save=save,type=type,button.name=button.name,button.otherarg
  =button.otherarg,button.function=button.function,button.data=button.data,
  button.biclust=button.biclust,arg.frames=arg.frames,new.frames=new.frames)

#### RADIO BUTTONS FRAME      ####

type <- "radiobuttons"

# Change variables accordingly:
frame.name <- "plottyperadio"
argument.names <- c("Within Biclusters","All Samples")
arguments <- c("type")
argument.values <- c("within","all")
argument.types <- "char"
initial.values <- "within"
title <- "Type?"
border <- FALSE

# DO NOT CHANGE THIS LINE:
new.frames <- .add.frame(input=input,type=type,frame.name=frame.name,
  argument.names=argument.names,arguments=arguments,argument.values=
  argument.values,initial.values=initial.values,title=title,border=border,
  new.frames=new.frames,argument.types=argument.types)

#####      ENTRY FIELDS FRAME      #####

```

```

type <- "entryfields"

# Change variables accordingly:
frame.name <- "whichrobust"
argument.names <- c("Number")
argument.types <- c("num")
arguments <- c("which.robust")
initial.values <- c("")
title <- "Which Robust BC?"
border <- FALSE
entry.width <- c("4")

# Do not change this line:
new.frames <- .add.frame(input=input,type=type,frame.name=frame.name,
  argument.names=argument.names,arguments=arguments,initial.values=
  initial.values,title=title,border=border,entry.width=entry.width,
  argument.types=argument.types ,new.frames=new.frames)

####      MANUAL BUTTONS FRAME      ####

type <- "buttons"

# Change variables accordingly:
frame.name <- "superplot"
button.name <- "Draw Plot"
button.function <- "plotSuper.GUI"
button.data <- ""
button.biclust <- ""
button.otherarg <- "CutTree=CutTree"
save <- FALSE
show <- FALSE
arg.frames <- c("plottyperadio","whichrobust")

# Do not change this line:
new.frames <- .add.frame(input=input,frame.name=frame.name,show=show,
  save=save,type=type,button.name=button.name,button.otherarg=
  button.otherarg,button.function=button.function,button.data=button.data,
  button.biclust=button.biclust,arg.frames=arg.frames,new.frames=new.frames)

###      ####
## SAVE BUTTON ##
####      ####

# ONLY FOR BICLUST

####      MANUAL BUTTONS FRAME      ####
#                                     #

type <- "buttons"

# Change variables accordingly:
frame.name <- "savebutton"
button.name <- "Save"
button.function <- "biclust.robust.fuse"
button.data <- ""
button.biclust <- ""
button.otherarg <- paste("CutTree=CutTree","method_result=",
  method_result,"",",type=",save.type,"",sep="")
save <- FALSE

```

```

show <- TRUE
arg.frames <- c()

# Do not change this line:
new.frames <- .add.frame(input=input,frame.name=frame.name,show=show,
  save=save,type=type,button.name=button.name,button.otherarg=
  button.otherarg,button.function=button.function,button.data=
  button.data,button.biclust=button.biclust,arg.frames=arg.frames,
  new.frames=new.frames)

####      MANUAL BUTTONS FRAME      ####

type <- "buttons"

# Change variables accordingly:
frame.name <- "resetbutton"
button.name <- "Reset"
button.function <- "robust.reset"
button.data <- ""
button.biclust <- ""
button.otherarg <- paste("method_result=",method_result,"",sep="")
save <- FALSE
show <- FALSE
arg.frames <- c()

# Do not change this line:
new.frames <- .add.frame(input=input,frame.name=frame.name,show=show,
  save=save,type=type,button.name=button.name,button.otherarg=
  button.otherarg,button.function=button.function,button.data=
  button.data,button.biclust=button.biclust,arg.frames=arg.frames,
  new.frames=new.frames)

### CONFIGURING GRID ###

if(biclust.combine==TRUE){
  if(make.save.button==TRUE){
    grid.config <- .grid.matrix(input=input,c("biclustcombine",
      NA,NA,"indexradio","typeradio","superbiclust","drawtree",NA,NA,
      "cutentry","cutree",NA,"showrobust","robustbutton",NA,"plottyperadio",
      "whichrobust","superplot","savebutton","resetbutton",NA),byrow=TRUE
      ,nrow=7,ncol=3,grid.config=grid.config)
  }
  else{
    grid.config <- .grid.matrix(input=input,c("biclustcombine",NA
      ,NA,"indexradio","typeradio","superbiclust","drawtree",NA,NA,"cutentry"
      ,"cutree",NA,"showrobust","robustbutton",NA,"plottyperadio",
      "whichrobust","superplot"),byrow=TRUE,nrow=6,ncol=3,grid.config=grid.config)
  }
}
else{
  if(make.save.button==TRUE){
    grid.config <- .grid.matrix(input=input,c("indexradio",
      "typeradio","superbiclust","drawtree",NA,NA,"cutentry","cutree",NA,
      "showrobust","robustbutton",NA,"plottyperadio","whichrobust","superplot",
      "savebutton","resetbutton",NA),byrow=TRUE,nrow=6,ncol=3,

```

```

    grid.config=grid.config)
}
else{
grid.config <- .grid.matrix(input=input,c("indexradio",
    "typeradio","superbiclust","drawtree",NA,NA,"cutentry","cuttree",NA,
    "showrobust","robustbutton",NA,"plottyperadio","whichrobust","superplot")
    ,byrow=TRUE,nrow=5,ncol=3,grid.config=grid.config)
}
}

### COMBINING ROWS ###

if(biclust.combine==TRUE){

grid.rows <- .combine.rows(input=input,rows=c(1),title=
    "Extra Biclust Data Input",border=TRUE,grid.rows=grid.rows,
    grid.config=grid.config)
grid.rows <- .combine.rows(input=input,rows=c(2),
    title="Superbiclust Configuration",border=TRUE,grid.rows=grid.rows,
    grid.config=grid.config)
grid.rows <- .combine.rows(input=input,rows=c(3,4),
    title="Dendogram",border=TRUE,grid.rows=grid.rows,grid.config=grid.config)
grid.rows <- .combine.rows(input=input,rows=c(5,6),title=
    "Robust Bicluster Gene Profiles",border=TRUE,grid.rows=grid.rows,
    grid.config=grid.config)
if(make.save.button==TRUE){
grid.rows <- .combine.rows(input=input,rows=c(7),
    title="Save the Robust Biclusters (Biclust Only)",border=TRUE,
    grid.rows=grid.rows,grid.config=grid.config)
}

}
else{
grid.rows <- .combine.rows(input=input,rows=c(1),
    title="Superbiclust Configuration",border=TRUE,grid.rows=grid.rows,
    grid.config=grid.config)
grid.rows <- .combine.rows(input=input,rows=c(2,3),
    title="Superbiclust Configuration",border=TRUE,grid.rows=grid.rows,
    grid.config=grid.config)
grid.rows <- .combine.rows(input=input,rows=c(4,5),
    title="Robust Bicluster Gene Profiles",border=TRUE,grid.rows=grid.rows,
    grid.config=grid.config)
if(make.save.button==TRUE){
grid.rows <- .combine.rows(input=input,rows=c(6),
    title="Save the Robust Biclusters (Biclust Only)",border=TRUE,
    grid.rows=grid.rows,grid.config=grid.config)
}

}

#####
## USE ALL THE ARGUMENTS ABOUT IN THE GENERAL NEW TOOL FUNCTION ##
#####

newtool_template(toolname=toolname,methodname=methodname,
    toolhelp=toolhelp,data.matrix=data.matrix,grid.config=grid.config,

```

```
grid.rows=grid.rows,new.frames=new.frames)
```

```
}
```

Auteursrechtelijke overeenkomst

Ik/wij verlenen het wereldwijde auteursrecht voor de ingediende eindverhandeling:

Software development for biclustering methods: The biclust GUI R package

Richting: **Master of Statistics-Bioinformatics**

Jaar: **2014**

in alle mogelijke mediaformaten, - bestaande en in de toekomst te ontwikkelen - , aan de Universiteit Hasselt.

Niet tegenstaand deze toekenning van het auteursrecht aan de Universiteit Hasselt behoud ik als auteur het recht om de eindverhandeling, - in zijn geheel of gedeeltelijk -, vrij te reproduceren, (her)publiceren of distribueren zonder de toelating te moeten verkrijgen van de Universiteit Hasselt.

Ik bevestig dat de eindverhandeling mijn origineel werk is, en dat ik het recht heb om de rechten te verlenen die in deze overeenkomst worden beschreven. Ik verklaar tevens dat de eindverhandeling, naar mijn weten, het auteursrecht van anderen niet overtreedt.

Ik verklaar tevens dat ik voor het materiaal in de eindverhandeling dat beschermd wordt door het auteursrecht, de nodige toelatingen heb verkregen zodat ik deze ook aan de Universiteit Hasselt kan overdragen en dat dit duidelijk in de tekst en inhoud van de eindverhandeling werd genotificeerd.

Universiteit Hasselt zal mij als auteur(s) van de eindverhandeling identificeren en zal geen wijzigingen aanbrengen aan de eindverhandeling, uitgezonderd deze toegelaten door deze overeenkomst.

Voor akkoord,

De Troyer, Ewoud

Datum: **10/09/2014**