FIGS USER'S MANUAL

Atif Khan PhD. and Juilee Thakar Ph.D.

Department of Microbiology and Immunology
Department of Biostatistics and Computational Biology
University of Rochester, NY USA

October, 2016

Revision Sheet

Release No.	Date	Revision Description
Version 1.0	10/31/16	Published FIGS User's Manual

FIGS USER'S MANUAL

TABLE OF CONTENTS

1.0   GENERAL INFORMATION

1.1   System Overview

Fuzzy inference of gene sets (FIGS) is a Matlab-based standalone executable package that uses dendritic cell and A549 epithelial cell influenza response to represent data driven fuzzy gene-sets. This interactive graphic user interface (GUI) package comes with several control parameters to probe fuzzy gene-sets where you can also perform enrichment analysis by providing your own gene set. The results of overlapping gene-sets are presented as circular graph that can be inspected interactively.

1.2   Authorized Use Permission

FIGS is developed solely for the research purpose to benefit immunologist and broader bioinformatics community. FIGS is freely available for the personal use. However, unauthorized or commercial redistribution of FIGS is highly prohibited.

The use of FIGS package is subject to the following conditions:

(1) The package and portions thereof may not be sold. Users may modify the software for their own use, but may not redistribute any version of the package other than the original version without the express permission of the authors.
(2) This package is provided on an "as is" basis. The authors in no way warrant either this package or results it may produce.
(3) The authors are under no obligation to provide any services by way of maintenance, updates or corrections for this package.
(4) Reports or publications resulting from use of this package must contain an acknowledgment in the form commonly used in academic research.

By downloading this package, you agree to the above conditions.

1.3   Points of Contact

For further information regarding development of FIGS, assistance in using FIGS or potential collaborative opportunities please contact:

Dr. Juilee Thakar
601 Elmwood Avenue,
Rochester, NY 14642.
Phone: 5852766925
juilee_thakar@urmc.rochester.edu

1.4   Citing FIGS

Thank you for using FIGS. Please cite the FIGS package as:

- Atif Khan, Dejan Katanic and Juilee Thakar, "FIGS: Fuzzy Inference of Gene-Sets: A data driven clustering approach", BMC Bioinformatics, **(Under Review, to be updated later).**

  FIGS is an advanced tool developed subsequently after PathCellNet that can be accessed at http://www.bio-networks.com/. Please cite PathCellNet as:

- Dejan Katanic, Atif Khan and Juilee Thakar. "PathCellNet: Cell-type specific pathogen-response network explorer." Journal of Immunological Methods (2016).

1.5   Acronyms and Abbreviations

FIGS: Fuzzy inference of gene-sets

PathCellNet: Pathogen specific cell type network

DC: Dendritic cells

EC: Epithelial cells

FCM: Fuzzy c-means clustering

ISGs: Interferon stimulated genes

KEGG: Kyoto encyclopedia of genes and genomes

U: Unified distance matrix

MV: fuzzy membership value

GUI: Graphic user interface

2.0   ISNTALLATION AND SYSTEM REQUIRMENTS

FIGS is a computational package that allows user to identify groups of gene signatures in human monocyte derived dendritic cell (DC) and A549 epithelial cells (EC) responsive to influenza infection.

2.1   System Requirements

FIGS is developed in Matlab R2016a environment and deployed as installable package. FIGS package does not explicitly require Matlab but it requires 'Matlab Runtime' during the installation process (in case you haven't installed it previously).

The 'Matlab Runtime' is a freely available standalone set of shared libraries that enables the execution of compiled Matlab applications or components on computers that do not have Matlab installed.

2.2   Installation Procedure

FIGS can be downloaded from https://github.com/Thakar-Lab/FIGS

Please make sure you downloaded FIGS.exe file (~20 Mb) from GitHub. Double click FIGS.exe to install FIGS package on the desired directory.
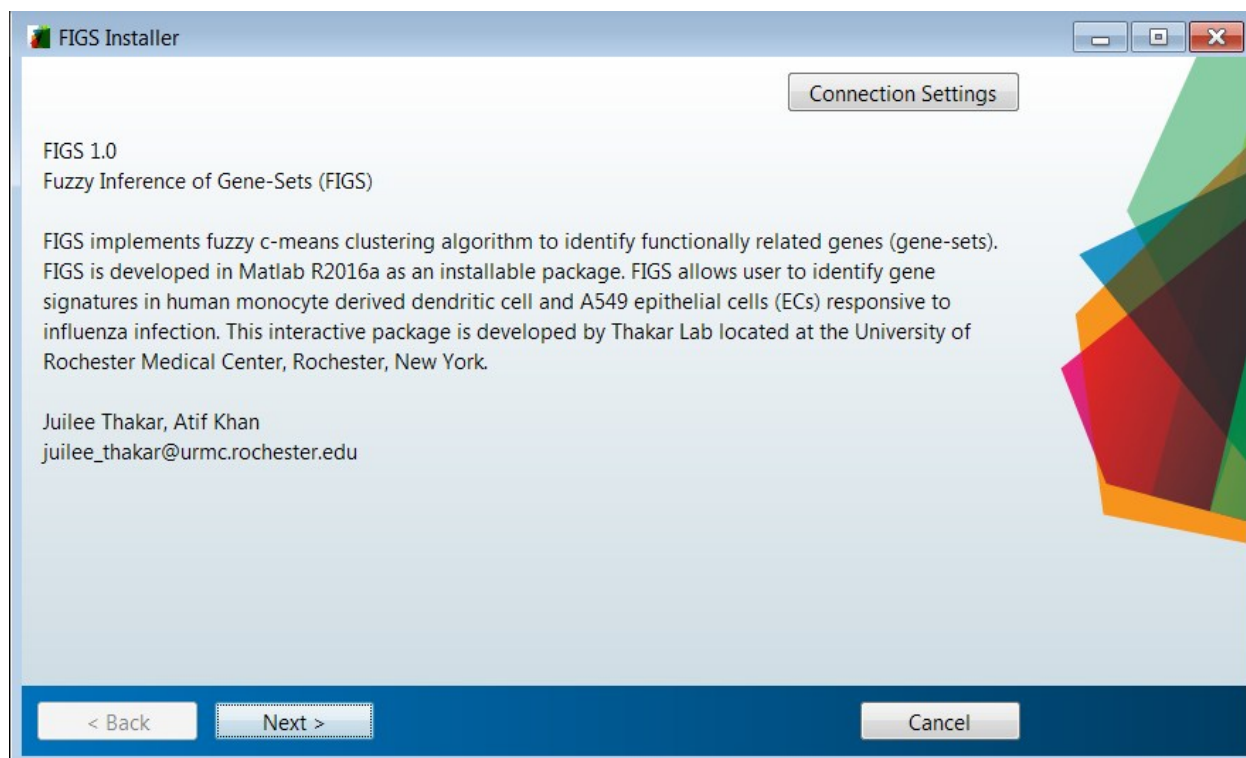
Figure 1: Installation of FIGS Package

If the FIGS installer finds the appropriate version of 'Matlab Runtime' already installed on your PC, the following information will be displayed (Figure 2) and you will no longer be required to install it again.

Figure 2: Detection of Previously Installed Matlab Runtime

In case the 'Matlab Runtime' is not found on your PC, the following information will be displayed (Figure 3). For the ease of use, 'Matlab Runtime' compiler installation is packaged with FIGS.exe. You will be automatically redirected to the Matlab website for downloading and installation of Matlab compiler (Figure 4).
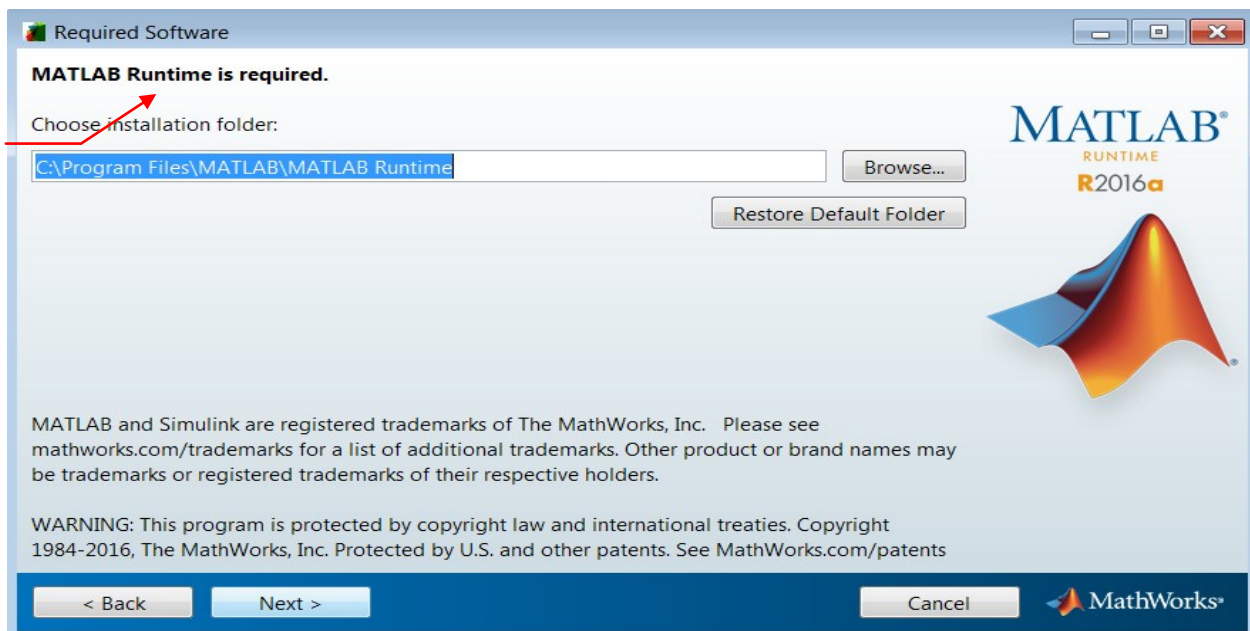
Figure 3: Matlab Runtime Installation Requirement
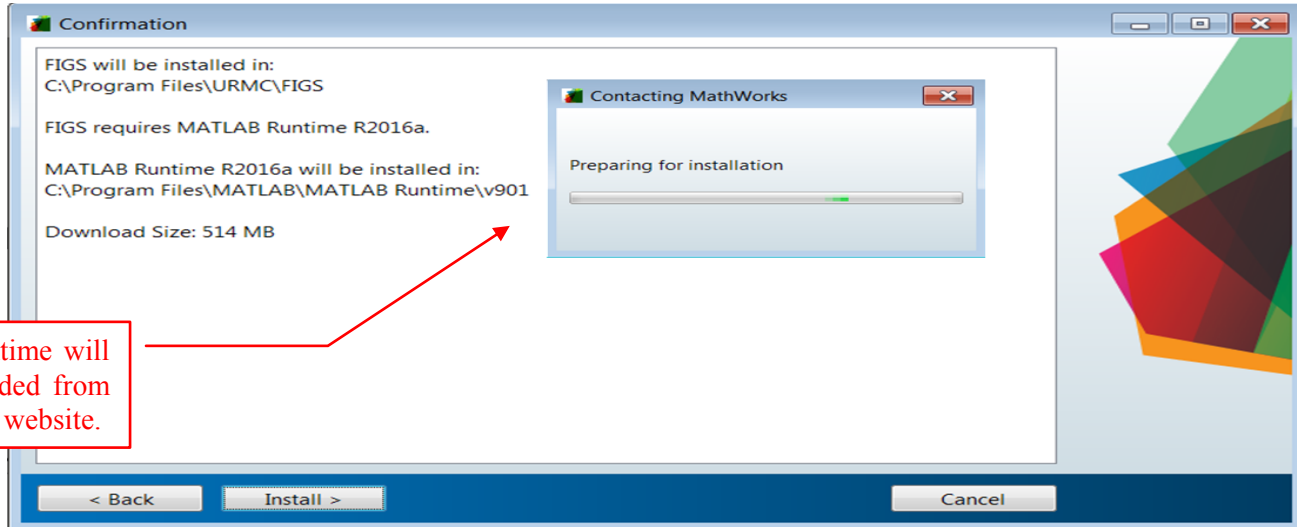
Figure 4: Automatic Downloading of Matlab Runtime

Please beware that Matlab Runtime (size ~500 Mb) will be downloaded from Matlab website (www.MathWorks.com). Downloading and installation (Figure 5) might take some time depending on your connection and internet speed.
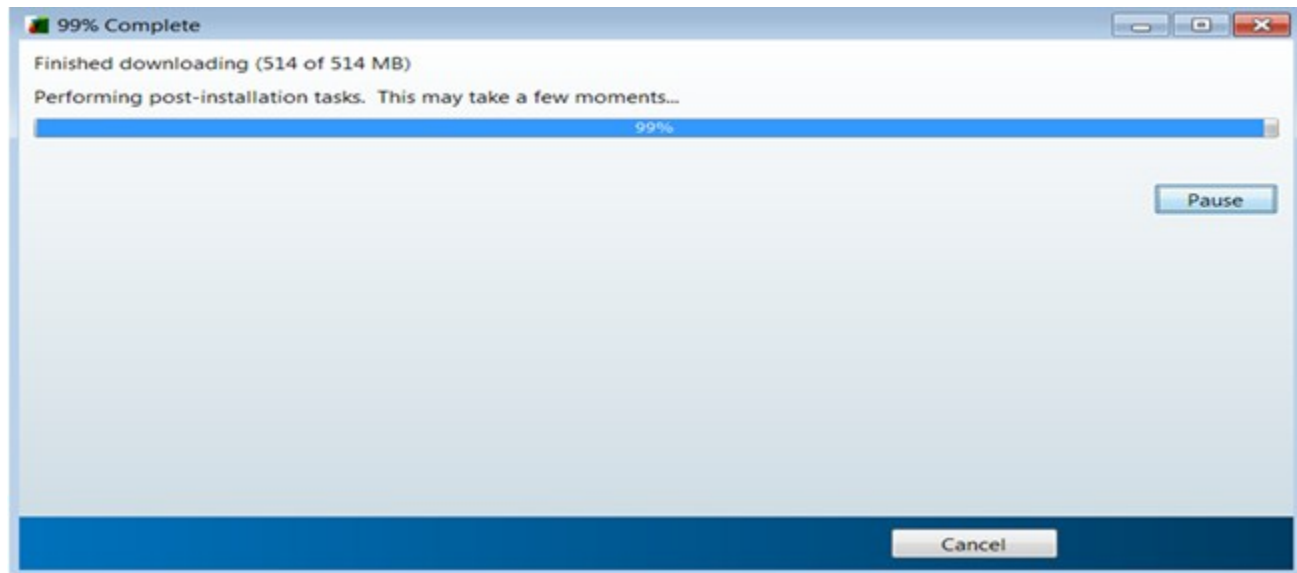
Figure 5: Automatic Installation of Matlab Runtime

2.3   Getting Started

Once the installation is finished, FIGS package can be accessed by running FIGS application from the program list (in case you have not created a desktop shortcut).

3.0 MAIN SYSTEM MENU

This section provides a general overview of the FIGS package with the help of screen shots.

3.1 Graphic User Interface

Run FIGS from the program list (in case you have not created a desktop shortcut). The main menu will appear with several control parameters as shown and described in figure 6.
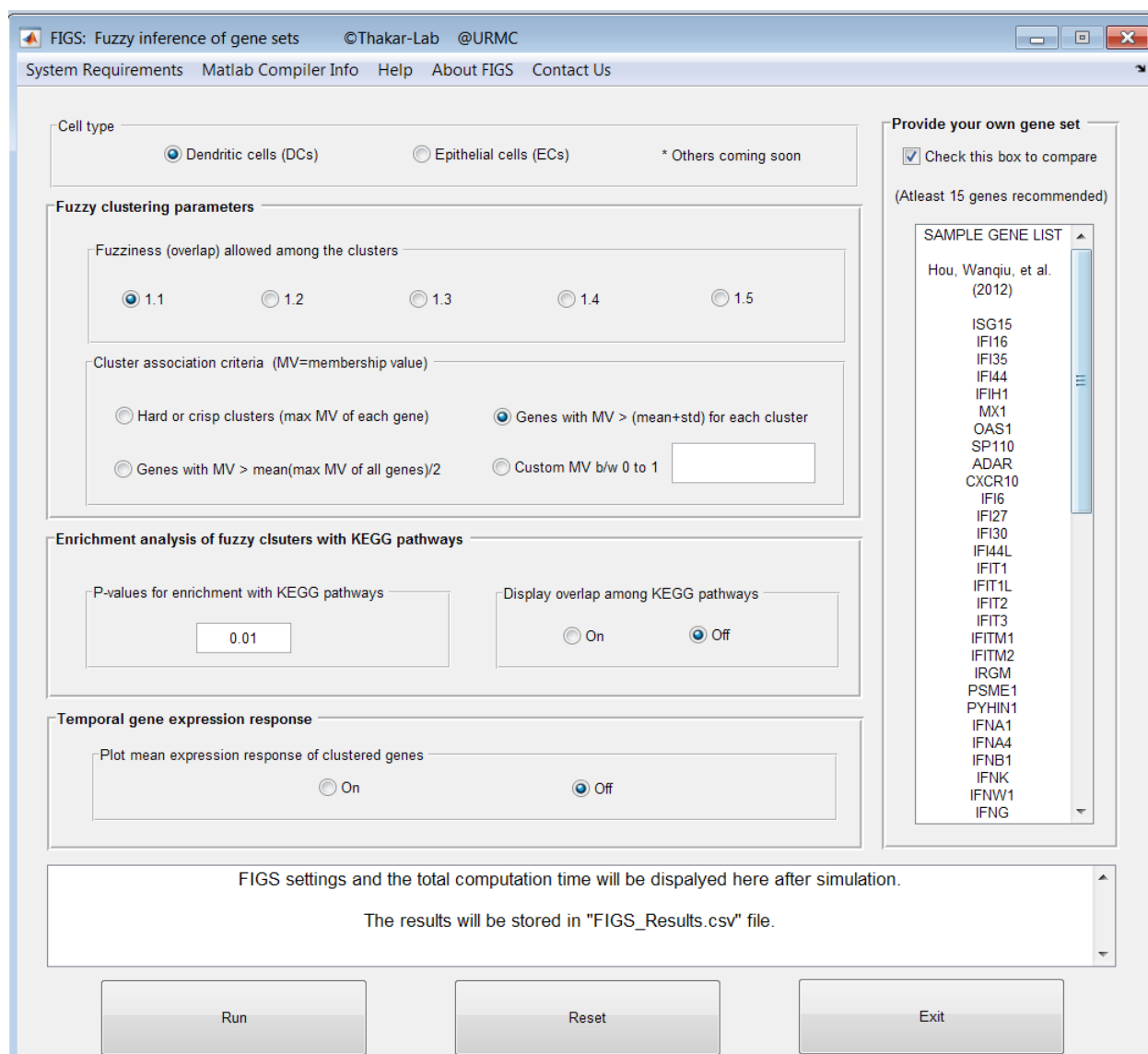
Figure 6: Main GUI of the FIGS Package

3.2   Menu Tabs

The main menu of FIGS that has several navigation tabs and menu bars. The description of tabs and the menu bars (figure 7) are self-explanatory.
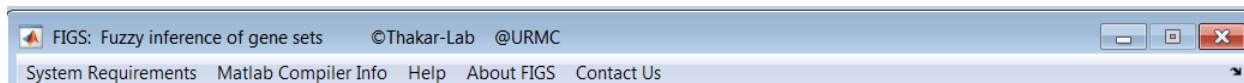
Figure 7: Main Menu Tabs

3.3   Cell Type Selection

Currently, FIGS package database includes Dendritic Cells (DCs) and Epithelial Cells (ECs). The results of FCM clustering for these datasets are stored in database to save computational time and facilitate quick retrieval. In future, more cell types will be added to the database and FIGS package will be updated accordingly.
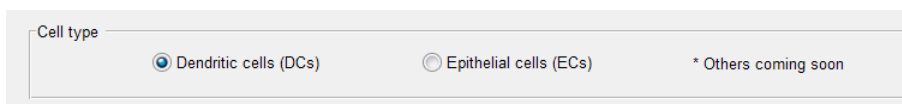
Figure 8: Cell Type Selection Tabs

3.4   Panel of Controllable Parameters

FIGS package comes with several user defined control parameters. We recommend using the default parameters but for the exploratory purposes, parameters can be modified depending on your requirement.
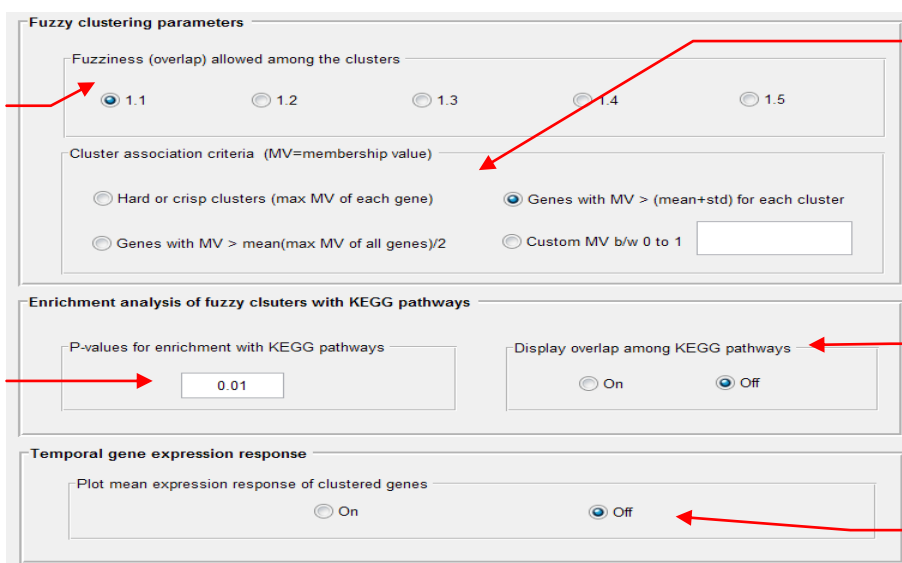
The extent of overlap allowed between the clusters. (A high value means more overlap).

Four different criteria for associating genes to the clusters, 1: Hard clusters 2: The distribution based assignment. 3: Mean-Max $(\sum_{i=1}^{n} \max(U_i)/2n)$ and 4. User defined threshold.

Provides an additional output to represent overlap among ISGs and KEGG pathways.

Provides an additional output plot to represent temporal expression of the clustered genes.

Figure 9: Controllable Parameters of FIGS Package

3.5   Upload User Gene-Sets

There is an option to provide your own gene-sets for enrichment with FCM produced gene-sets. Gene names can be manually entered or can be copy-pasted in the panel shown below (Figure 10). To do the enrichment of your own gene-sets, please check the box "Check this box to compare".
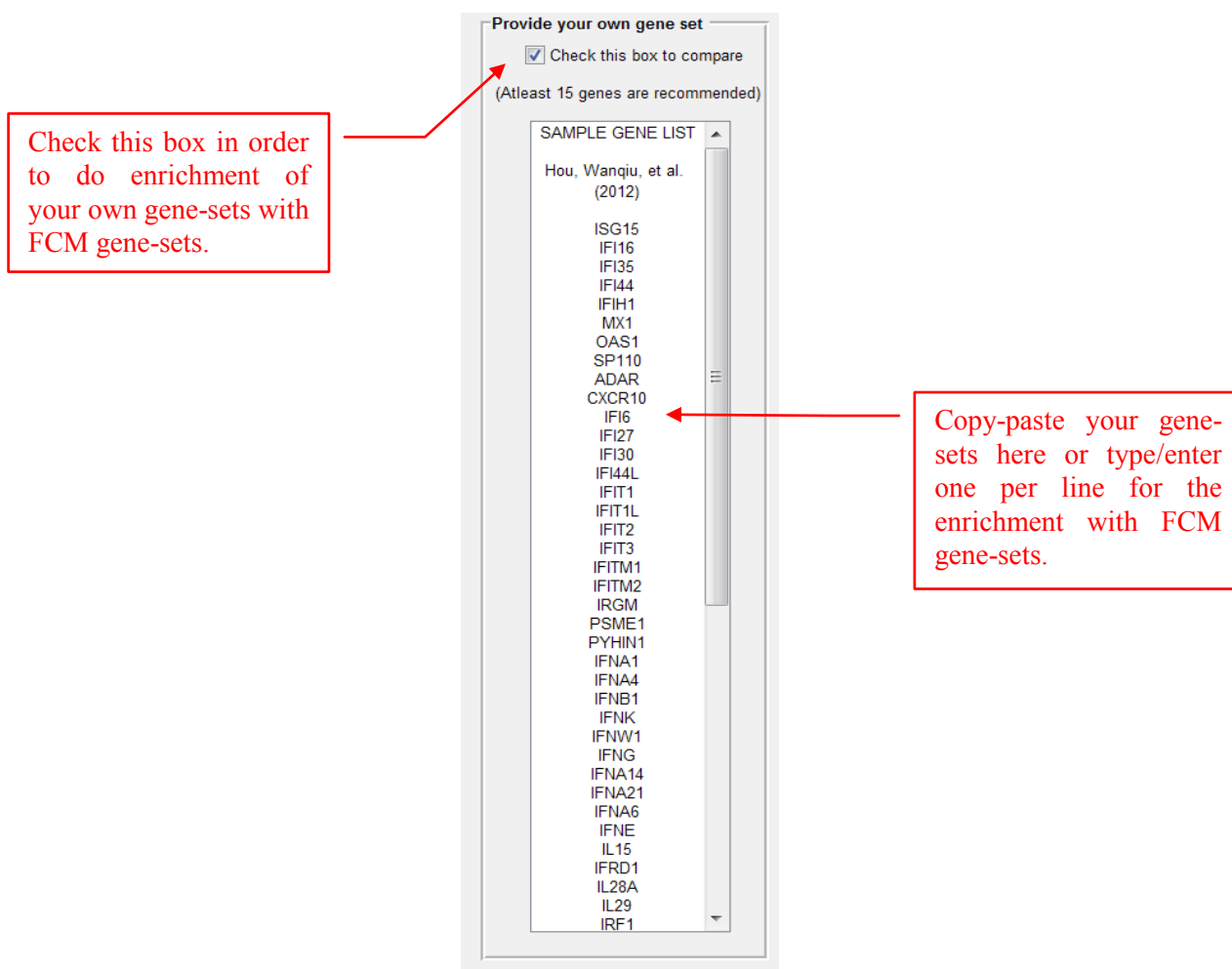
Check this box in order to do enrichment of your own gene-sets with FCM gene-sets.

Copy-paste your gene-sets here or type/enter one per line for the enrichment with FCM gene-sets.

Figure 10: User Provided Gene-Sets for Enrichment with FCM Gene-Sets

3.6   Run, Reset and Exit

Pressing the run tab will start computation and exit tab will quit the package. Pressing reset tab will clear the setting of all the parameters to the default state.
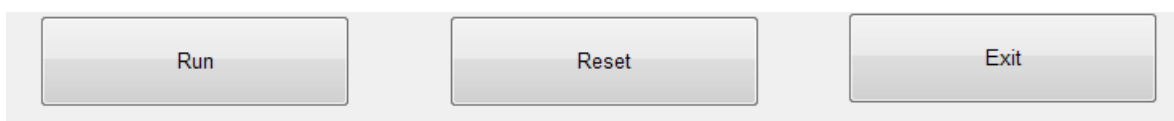
Figure 11: Control Tabs for Computation

3.7   Error Dialogs

The package will show an error dialog if the cell type, cluster association criteria and/or fuzziness values are not selected. Additionally, there will be an error message if the value of the user provided membership value threshold is not between 0 and 1.
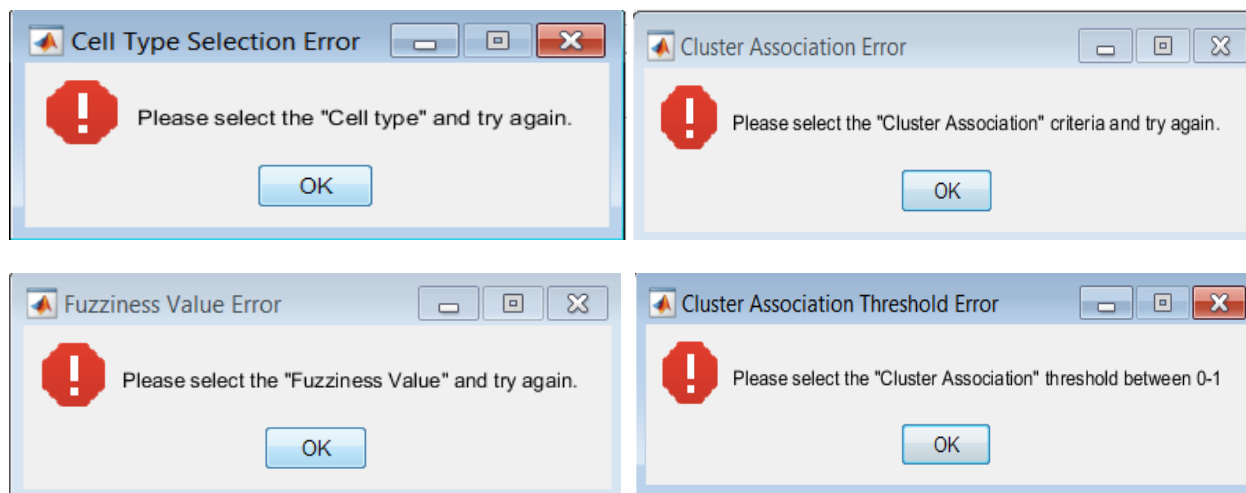
Figure 12: Error Dialogs for FIGS Parameter Settings

3.8   Progress Bars

The package computes output in three steps as shown in Figure 13. The progress at each step will be shown by the green color progress bar.
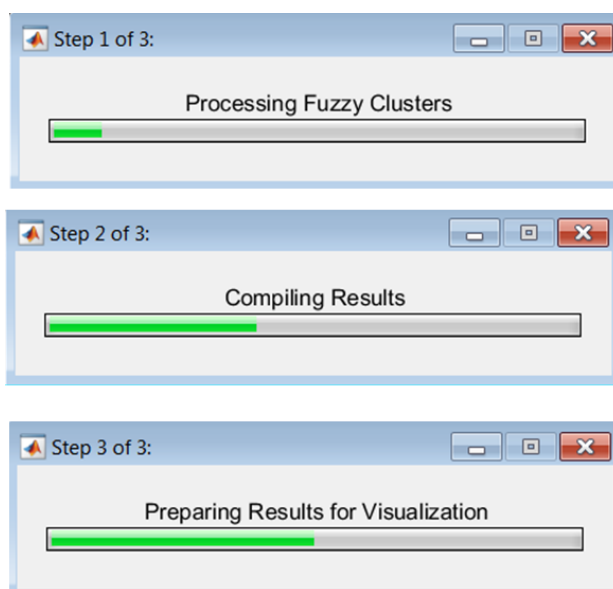
Figure 13: Automatic Installation of Matlab Runtime

4.0 USING THE SYSTEM

This section provides a descriptive usage of the FIGS package.

4.1 Fuzzy Gene-Sets

The results of FCM gene-sets are displayed in tabular as well as interactive circular graphs. Figure 14 shows the composition of FCM gene-sets and their enrichment with KEGG pathways (where red color shows significant overlap based on user selected p-value).

Figure 14: Tabular Representation of FCM Gene-Sets and Their Enrichment with KEGG Pathways

4.2   Enrichment of User Provided Genes

If there is a user provided gene set, the results for the enrichment of user genes with FCM gene-sets will be displayed as shown in the Figure 15 below.

P-values from hypergeometric test to represent significant overlap between FCM gene-sets and the user provided gene-sets. The p-values that are less than the user defined p–value threshold are highlighted with red color.

Figure 15: Enrichment of User Genes with FCM Gene-Sets

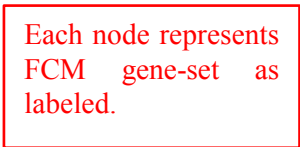## 4.3   Overlap among KEGG Pathways

There overlap found among the KEGG pathways are displayed as interactive circular graph (Figure 16). Each node of the circular graph represents KEGG pathway as labelled where a line or an arc between the two nodes shows the existence of common genes between the two pathways. The width of the line/arc corresponds to the number of common genes between the two pathways.

Figure 16: Overlap among KEGG Pathways

4.4 Overlap between Fuzzy Gene-Sets

The overlap among FCM gene-sets are displayed a separate interactive circular graph (Figure 17). In addition to the representation of an overlap, each cluster is labelled with the KEGG that was most significantly enriched by that cluster.

Figure 17: Overlap among FCM Gene-Sets

4.5 Interactive Circular Graphs

The overlap among KEGG pathways and among FCM gene-sets are represented by interactive circular graphs (Figure 18). User can select/hide the nodes by mouse clicks. A single mouse click on a node will change it state from 'show connection' to 'hide connections' and vice versa.

The interaction of a node with others can be inspected by single clicking that node.

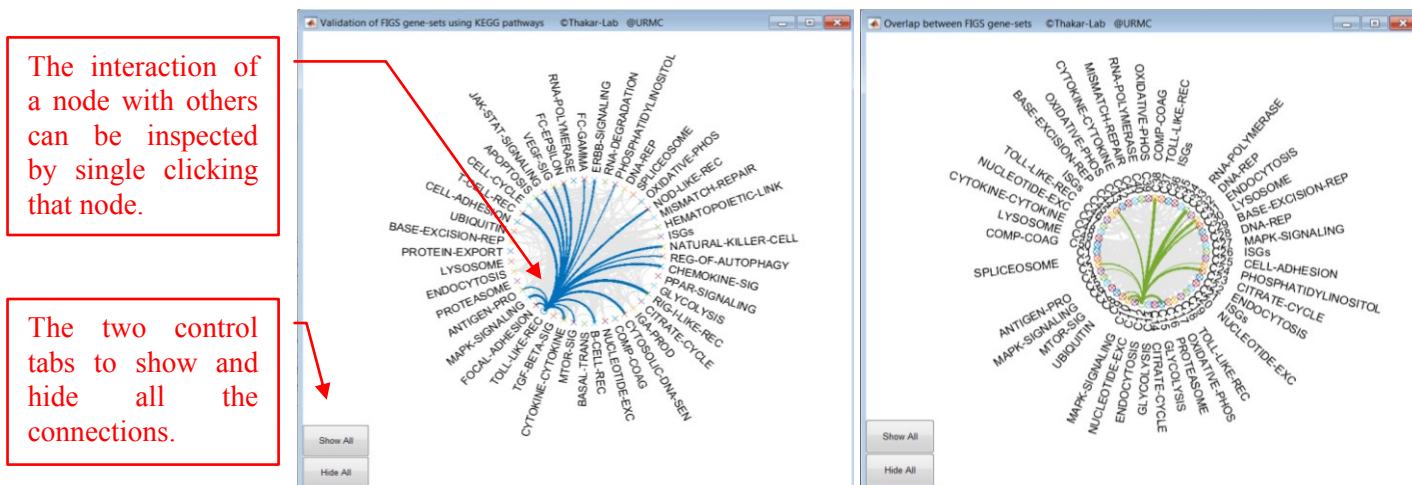The two control tabs to show and hide all the connections.

Figure 18: Overlap among FCM Gene-Sets

4.6 Temporal Expression Profile of Gene-Sets

From the drop down menu, the temporal expression of any two clusters can be compared at a time.

Drop-down menu for selecting any two clusters for analyzing temporal expression patterns.

The mean temporal expression of the clustered gene-sets.

Box plot repressing distribution of genes within the cluster.
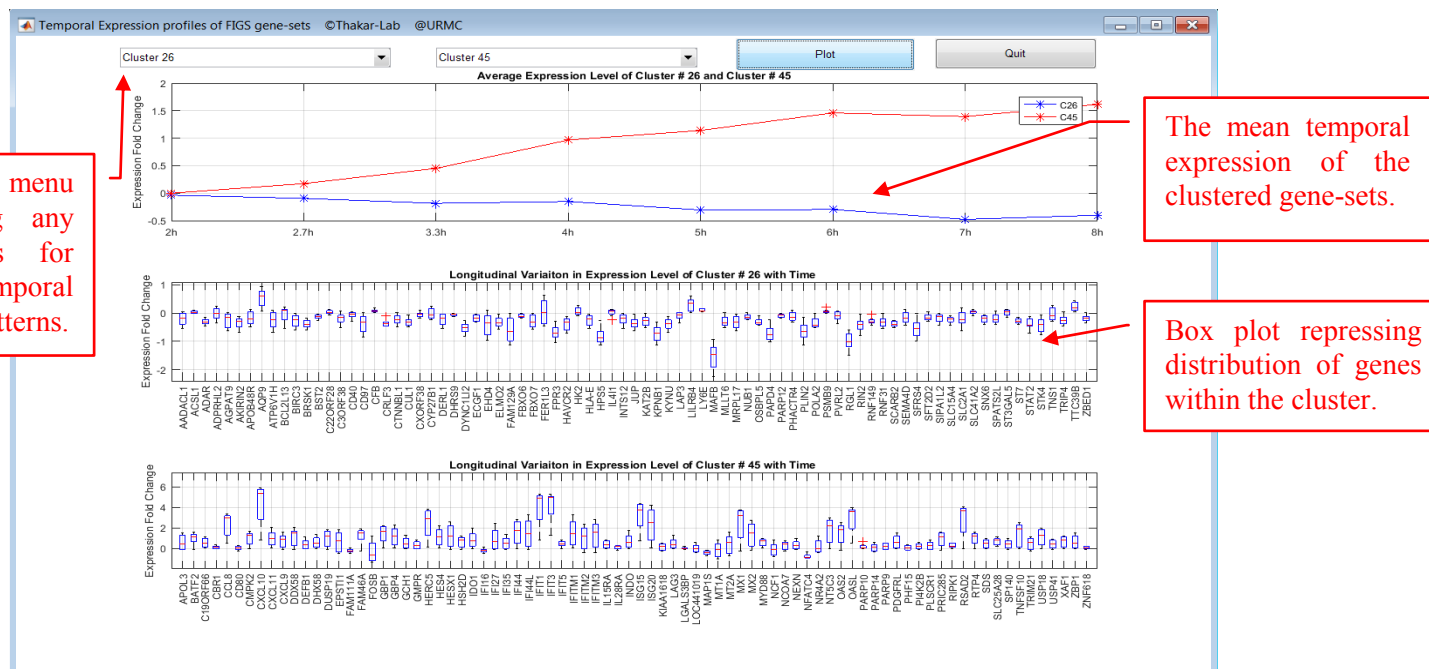

Figure 19: Temporal Expression Pattern of FCM Gene-Sets

4.7   Saving the Results

The detailed results are displayed in the tabular form (see section 4.1). A copy of the data-driven FCM gene-sets will be generated and saved as "FIGS_Results.csv". This results file can be found in the installation folder of FIGS.
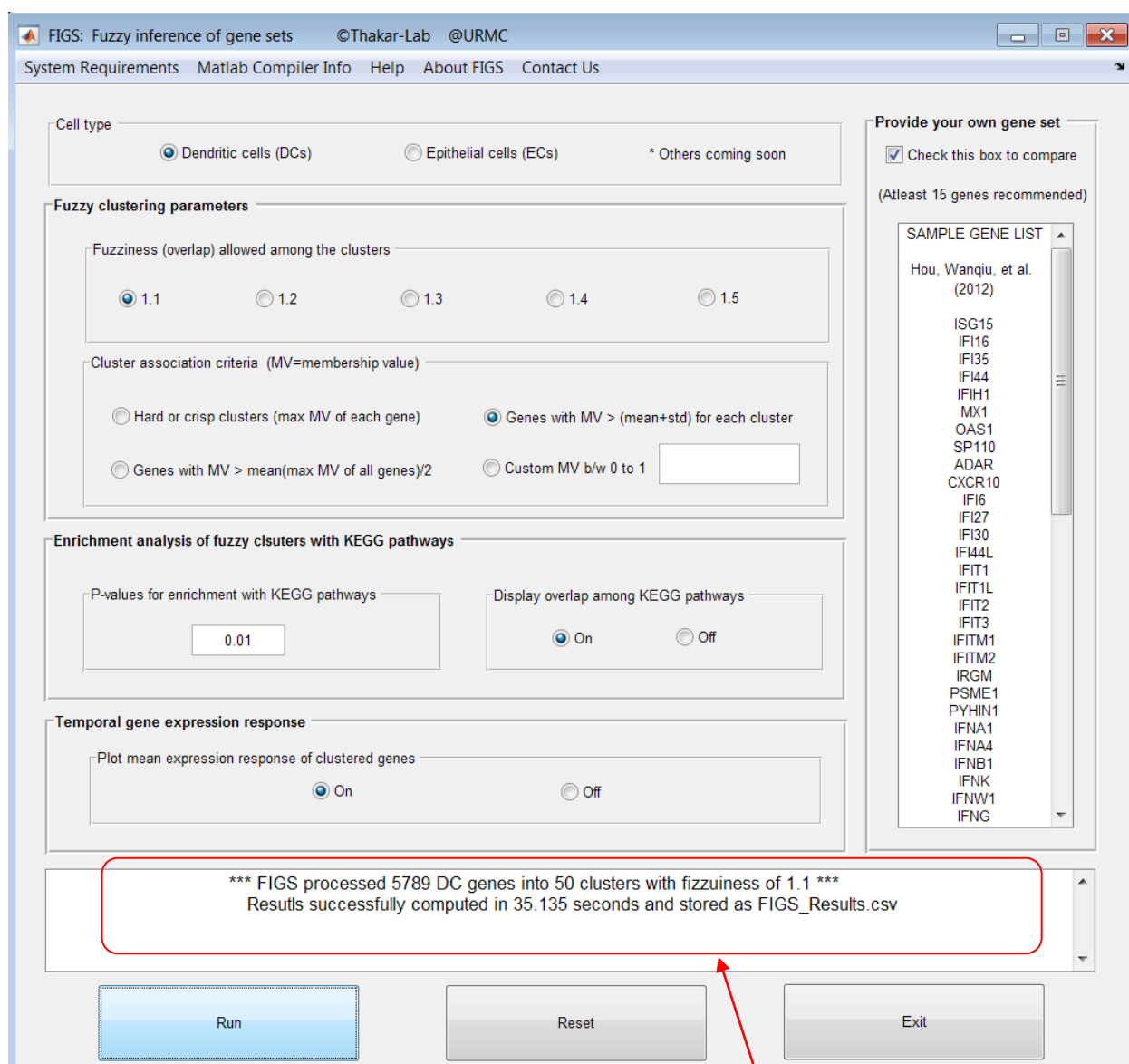
Figure 20: FIGS Results

The gene-sets produced by FIGS will be stored as FIGS_Resutls.csv file.