HOUSTON Methodist LEADING MEDICINE

CCCExplorer

Cell-Cell Communication Explorer Software version 1.0 User Guide

Revision History

Version	Date	Description	Created By	Reviewed By
1.0	9/6/2017	User Guide	Xiaohui Yu/ Jianteng Sheng	Stephen Wong

HOUSTON Methodist LEADING MEDICINE

Contents

Introduction

CCCExplorer (Cell-Cell Communication Explorer) is a Java-based software for modeling, predicting, and visualizing crosstalk signaling networks among different types of cells within a microenvironment, e.g., stroma-tumor or tumor-immune cell crosstalk identification in the tumor microenvironment. CCCExplorer incorporates computational models and matching algorithms for uncovering cell-cell communication as a directed and connected network from ligand-receptor interactions to transcription factors (TFs) and their target genes. The next version of CCCExplorer will include exosomal communications among different types of cells.

If you have any questions, please contact Jianting Sheng (jsheng@houstonmethodist.org) or Stephen Wong (STWong@houstonmethodist.org)

System requirements

The minimum of system requirements for CCCExplorer

Processor 1GHz
Memory 512MB
Graphics Card On board Video
Monitor XGA (1024X768)

Getting Started

Installing CCCExplorer on Windows OS

- **Install Java JRE**
  CCCExplorer is a Java-based application. If Java is not installed on your computer, please download and install Java SE 7 or higher. The JRE package is available from:

  http://www.oracle.com/technetwork/java/javase/downloads/jre7-downloads-1880261.html

- **Install R language and dependent packages**
  Download R language (V3.2-0 or higher)/R studio for Windows from https://cran.r-project.org/bin/windows/base/ and http://cran.rstudio.com/Install R.

  CCCExplorer uses igraph R package. To install this package, open RGui and run the command below.

```
>install.packages("igraph")
```

- **Run CCCExplorer**
  Decompress CCCExplorer.zip

  1. Run CCCExplorer by double clicking “CCCExplorer_x64.exe”

     If you want to see the output debug log, type the following command when running CCCExplorer.

     ```
     java –jar CCCExplorer.jar
     ```

  2. When using CCCExplorer for the first time in Windows, you need to set the R script path. Click “setting” icon. In the popup dialog, click “Browse” to find the Rscript.exe file that you installed in your system. Then click “OK”.

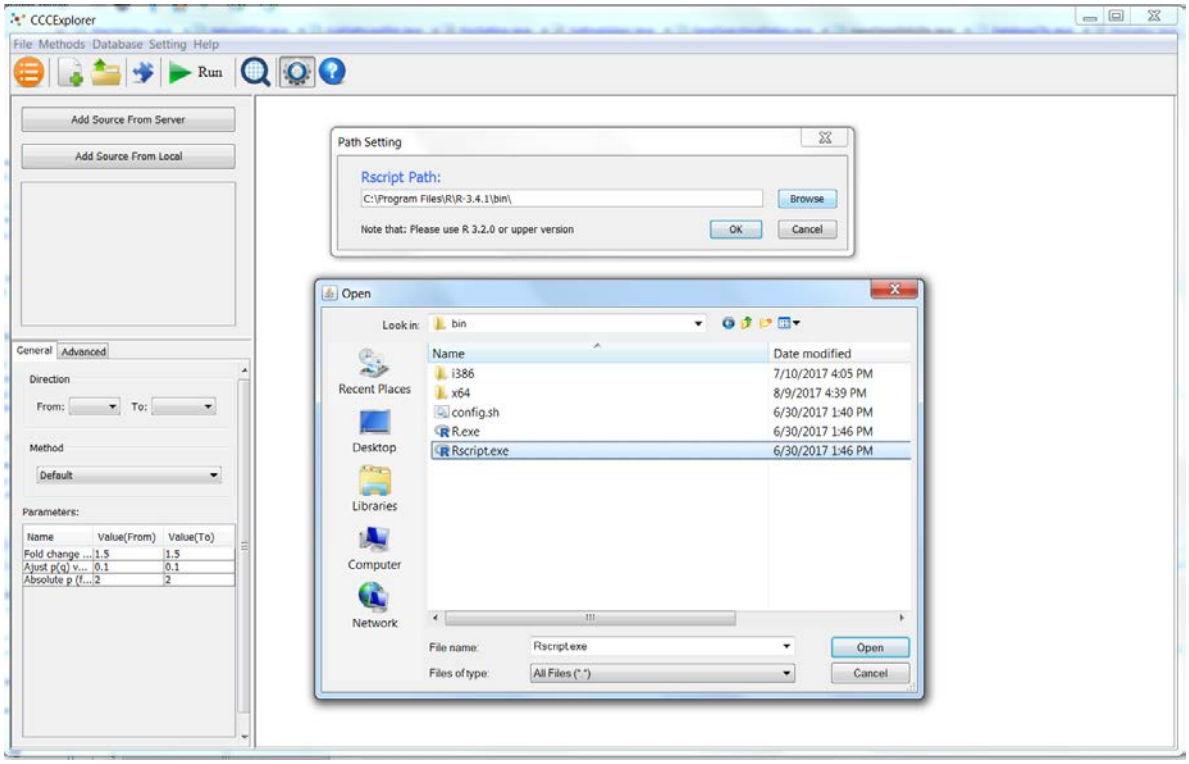

Figure 1

Install CCCExplorer running environment on Ubuntu

- **Install the dependent libraries**
  Install Java runtime environment

  ```
  sudo apt-get install default-jre
  ```

Install the R runtime environment

```
sudo apt-get install r-base-core
```

- **Install R package**
  Run R command in the terminal and install the R package “igraph”.

```
xxxhuis:user$ R

>install.packages("igraph")
```

- **Set the executable attribute**

```
uzip CCCExplorer.zip

chmod +x CCCExplorer_runner.sh
```

- **Run CCCExplorer**
  You can run CCExplorer in your terminal with the following command.

```
./CCCExplorer_runner.sh
```

Install CCCExplorer running environment on Mac OS X 10.9 or higher

- **Install JDK and JRE for Mac OS**
  Download packages from the following link and install them.

  http://www.oracle.com/technetwork/java/javase/downloads/jre8-downloads-2133155.html

- **Install R and setup the runtime environment**
  Download and install R-3.1.3-mavericks.pkg from the link below.

  http://cran.r-project.org/bin/macosx/

  Run R command in the terminal and install the dependent R package “igraph”.

```
xxxhuis-Mac-2:xxx$ R

>install.packages("igraph")
```

  After installing it, use q() to exit R.

- **Set the executable attribute.**

```
uzip CCCExplorer.zip
```

```
chmod +x CCCExplorer_runner.sh
```

- **Run CCCExplorer**
  You can run CCExplorer in your terminal with the following command.

```
./CCCExplorer_runner.sh
```

Notes:

CCCExplorer is using the port 3306 to access the remote database. Sometimes this port is blocked by IT in your organization. You should check the status of port 3306 if you meet the error below.

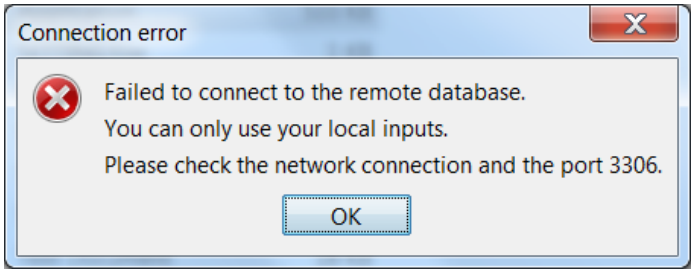

Figure 2

HOUSTON Methodist LEADING MEDICINE

Installation Directory Structure

CCCExplorer installation requires the following files and directories:

Table 1

File	Description
sample/	Sample data for CCCExplorer.
data/	Dependent data files to support our algorithm
lib/	Library jar files needed to run CCCExplorer
methods/	Configuration file for methods
workspace/	Output directory - all output files will be created in this folder
explorer.setting	Configuration file for CCCExplorer
CCCExplorer.jar	Main CCCExplorer application (Java archive)
CCCExplorer.rda	Algorithm function
CCCExplorer_default.R	Main R algorithm file
CCCEXplorer_runner.sh	Script to run CCCExplorer from the terminal (Linux, Mac OS X)
README.TXT	

Resource:

Demonstration video of CCCExplorer

https://youtu.be/wN4vHV6hoig

How to use CCCExplorer

Run a project

To run a project, you first need to upload input files, select method, and set up direction. Then configure the parameters. You can also load cell type input files from either a remote server or a local folder (see Figure 3 and Figure 4).

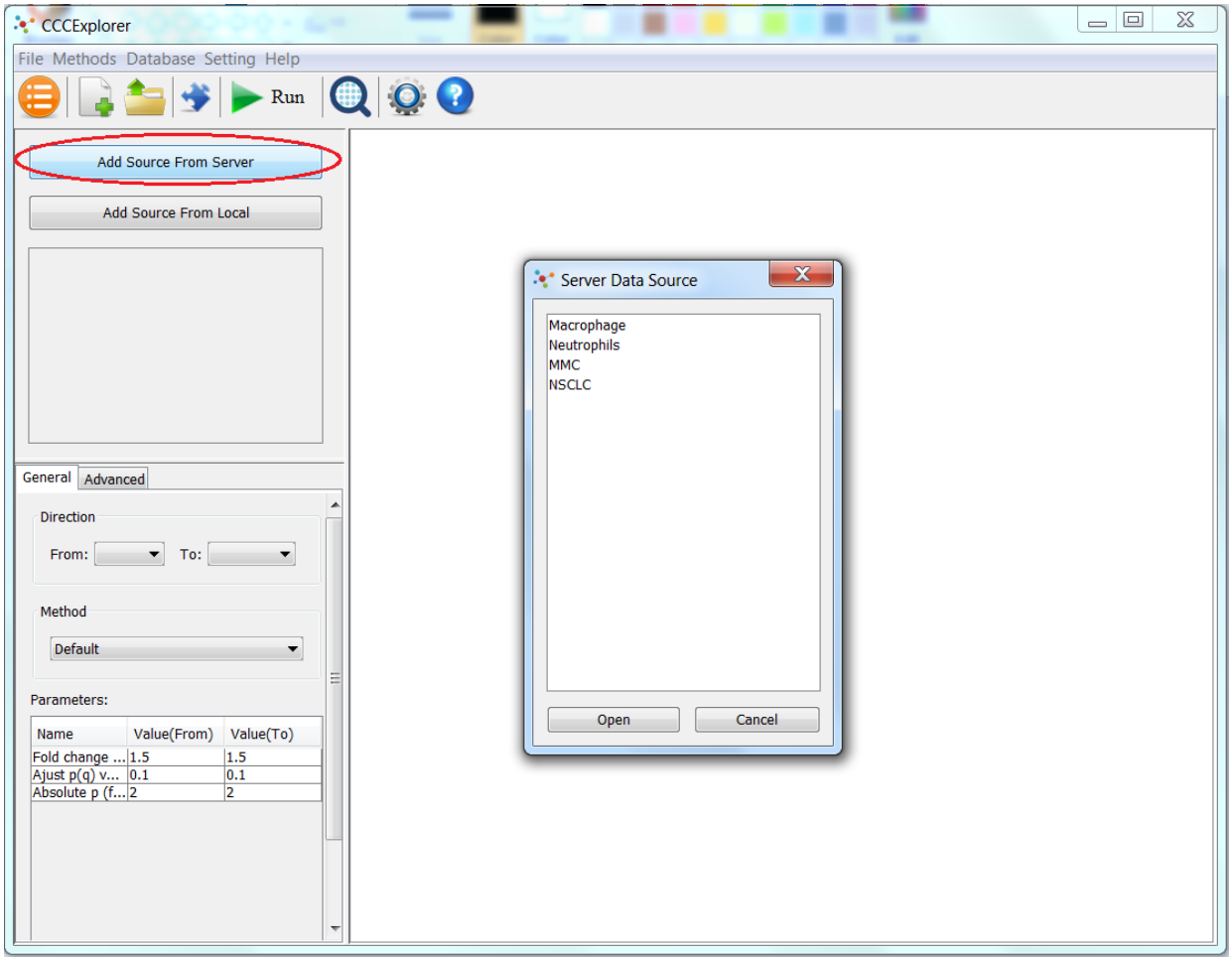

Figure 3

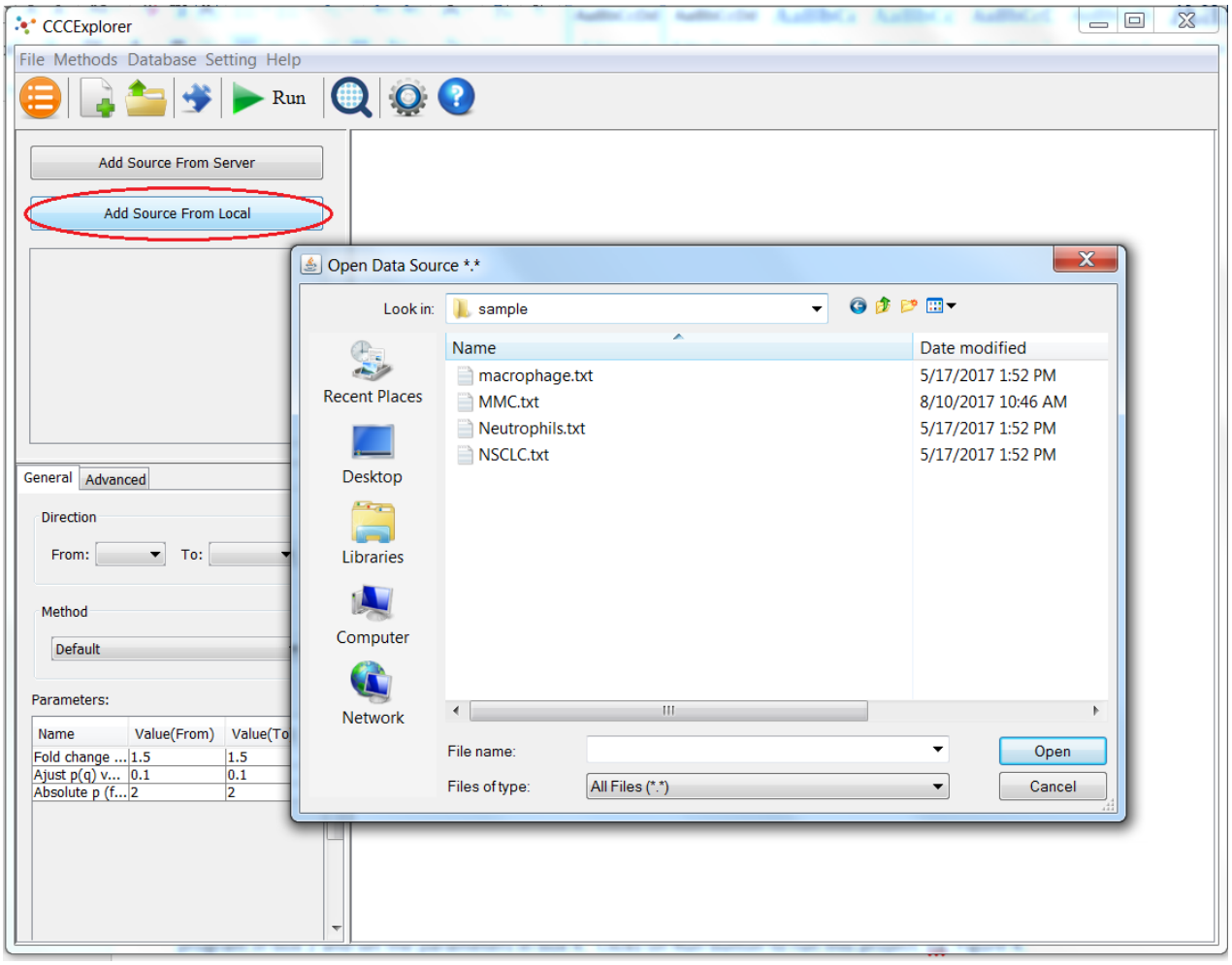
CCCExplorer
File Methods Database Setting Help
Run
Add Source From Server
Add Source From Local
General
Advanced
Direction
From:
To:
Method
Default
Parameters:
Name
Value(From)
Value(To
Fold change ...
1.5
1.5
Ajust p(q) v...
0.1
0.1
Absolute p (f...
2
2
Open Data Source *.*
Look in:
sample
Recent Places
Desktop
Libraries
Computer
Network
Name
Date modified
macrophage.txt
5/17/2017 1:52 PM
MMC.txt
8/10/2017 10:46 AM
Neutrophils.txt
5/17/2017 1:52 PM
NSCLC.txt
5/17/2017 1:52 PM
File name:
Files of type:
All Files (*.*)
Open
Cancel

Figure 4

At least one cell type input file will be needed. After all cell type files are uploaded, set up the intended direction of crosstalk (red circled icon #2, Figure 5), select the algorithms under method (red circled icon #3) and set the parameters (red circled icon #4). Then clicks on Run (red circled icon #1) to run the project.

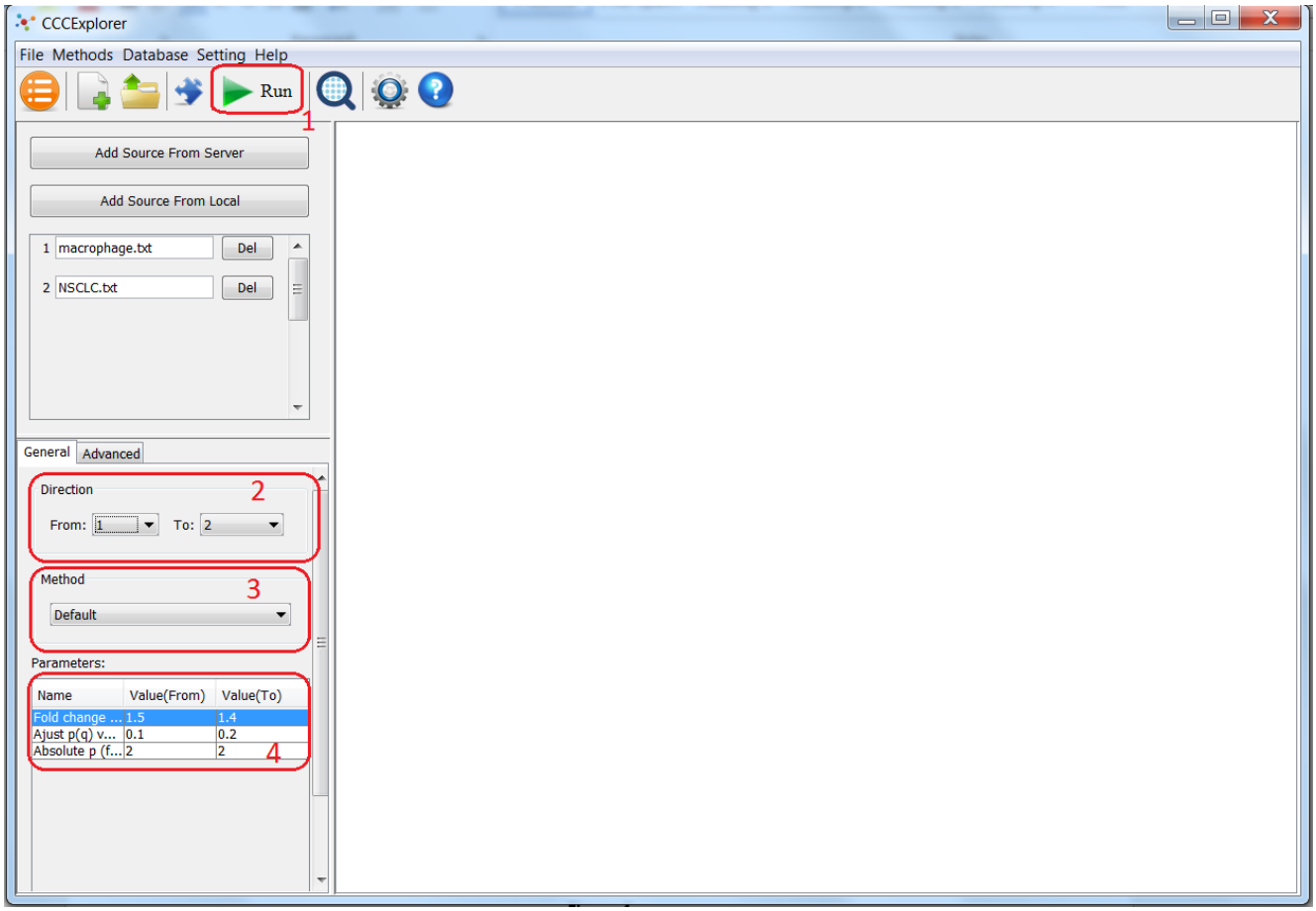

Figure 5

Check the result of a project

After the running process finishes, an output window will show up (Figure 6).

The pathway network will be displayed in red box #1. The p value for each pathway will be shown in red box #2. The list of identified ligand and receptor pairs will be shown in red box #3.

Red box #4 is the output window toolbar. Use it to zoom in, zoom out, modify the visualization of the network, or export the network view.

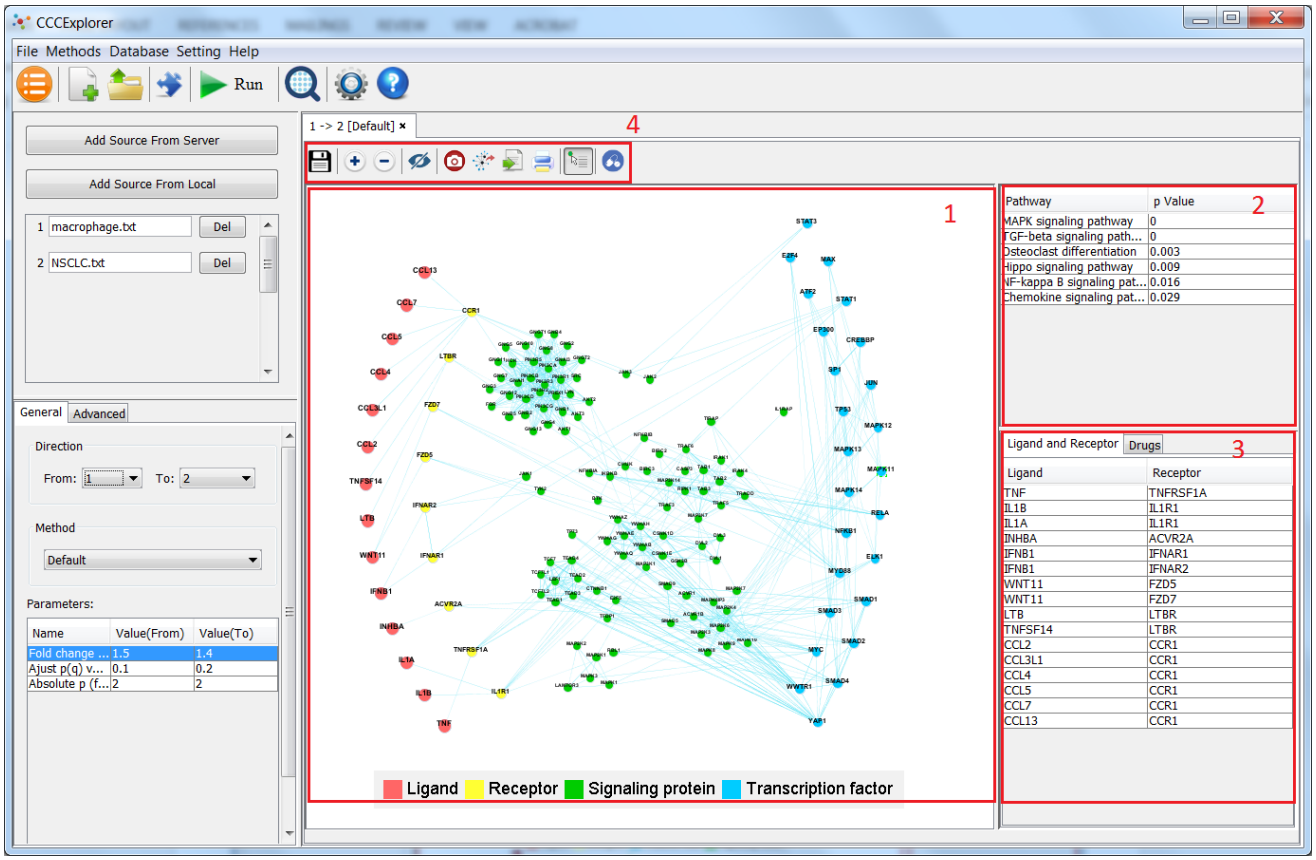

Figure 6

Enable/Disable tip box

When hovering the pointer over a node, a tip box will show up. If you want to disable the tip box, click on the red-circled icon in Figure 7.

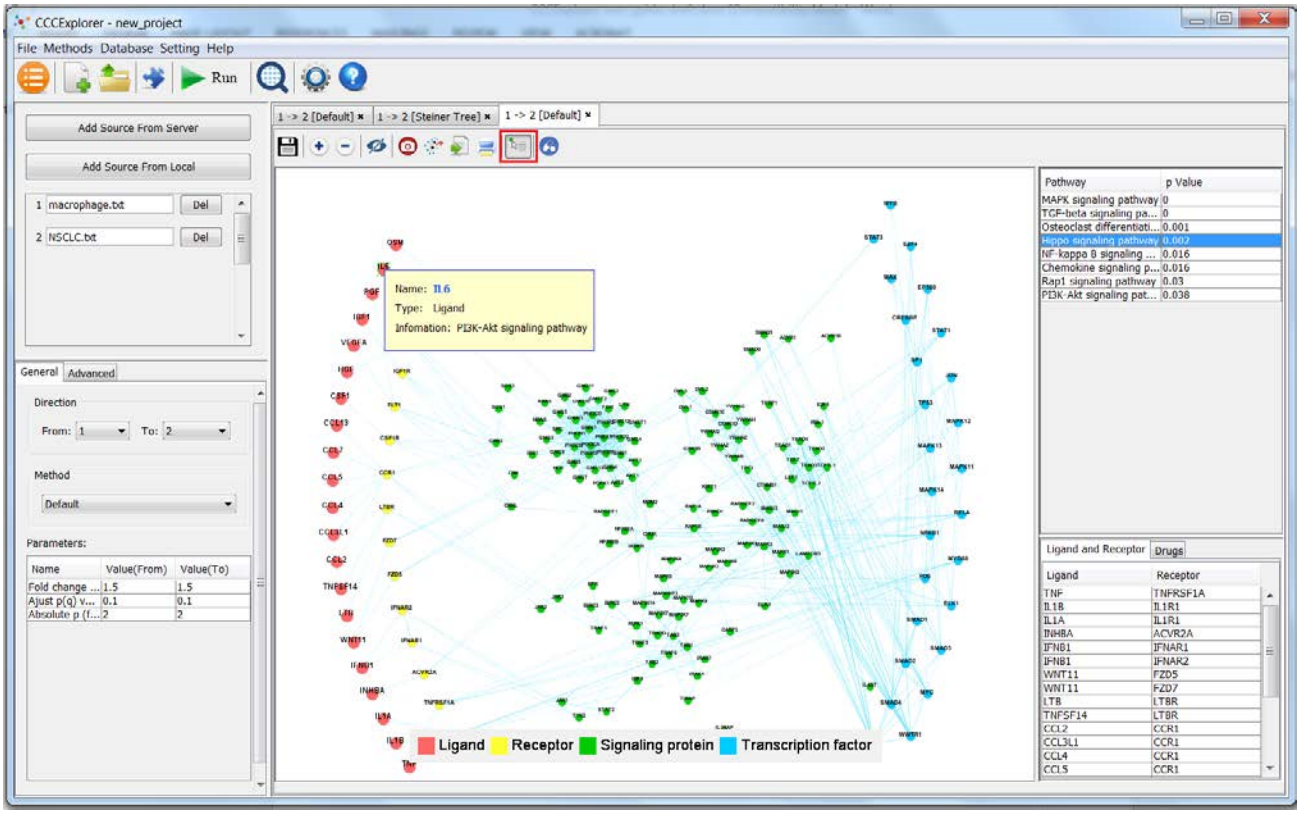

Figure 7

Highlight specific pathways

You can highlight a specific signaling pathway by clicking it in the pathway table (Figure 8).

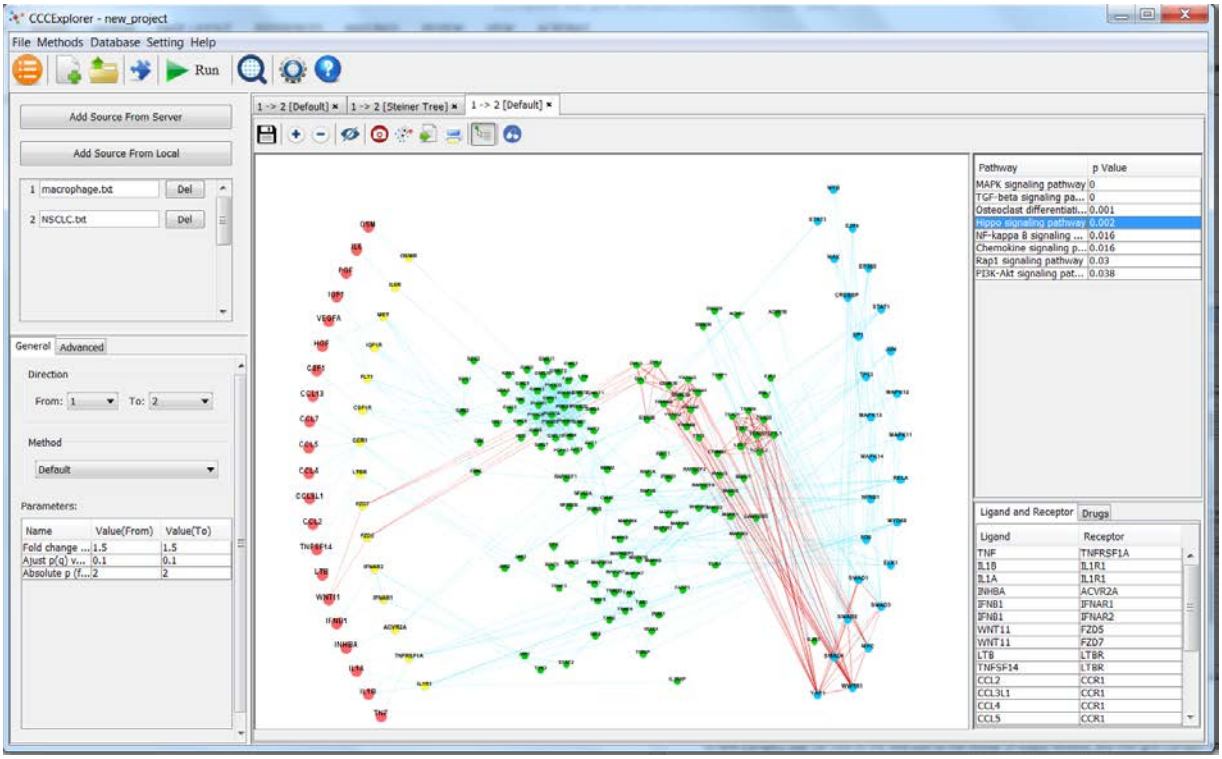

Figure 8

Enable/Disable pathway edges

Click the red-circled icon to hide/show all the edges in the network layout (Figure 9).

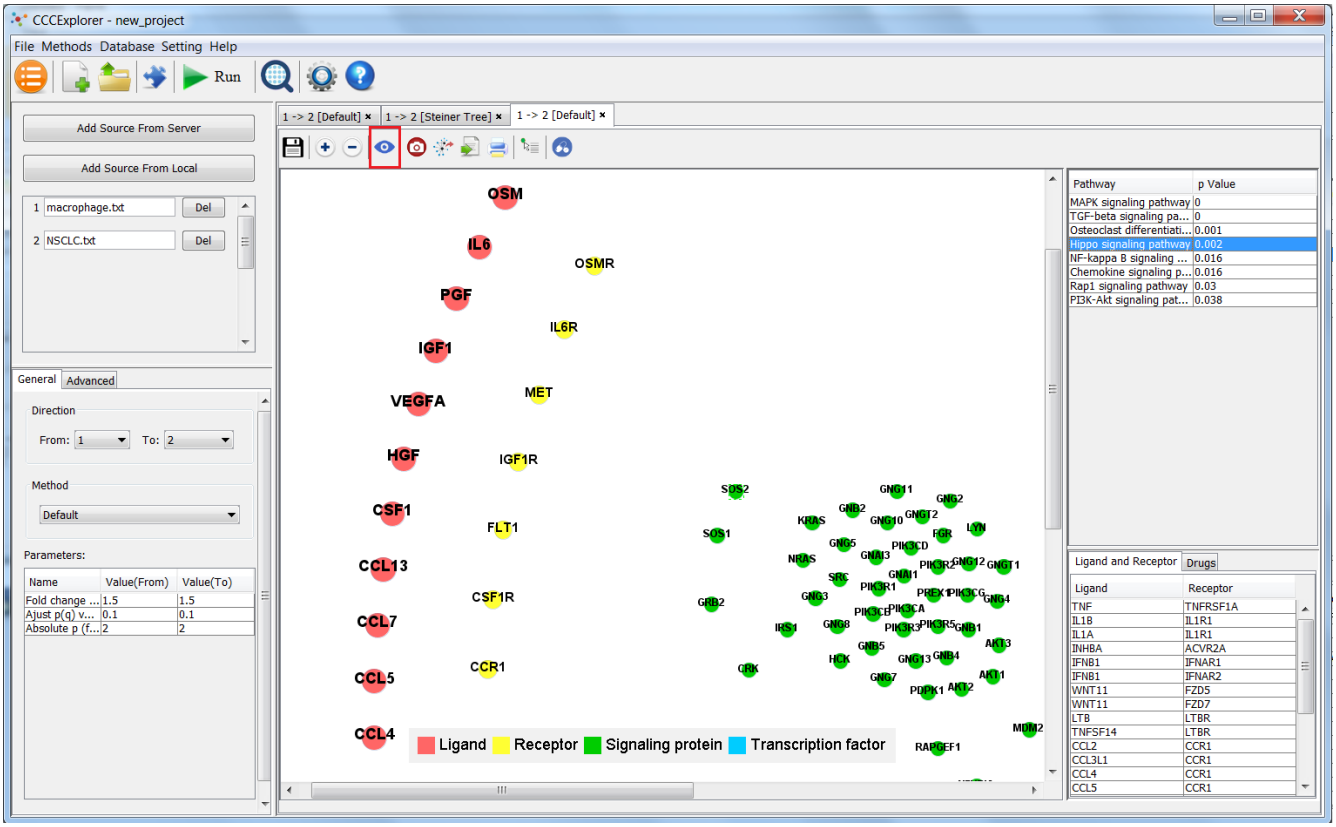

Figure 9.

Hide/Show drugs on target gene

Click on the red-circled icon in Figure 10 to show all target genes with treatment drugs. The light purple nodes are the genes that can be targeted by one or more drugs from an online database, e.g., DrugBank or DrugMapCentral. Moving your mouse over one gene, the corresponding drugs will be listed in the drug table (circled red in Figure 10). To prevent the drug list from changing when you move your mouse, you can move the mouse over the gene node and double right-click it. Then the drug list won't change when you move your mouse from gene to gene. To change it back, just click on the canvas of the network.

Double click a specific drug in the drug table to get detailed information.

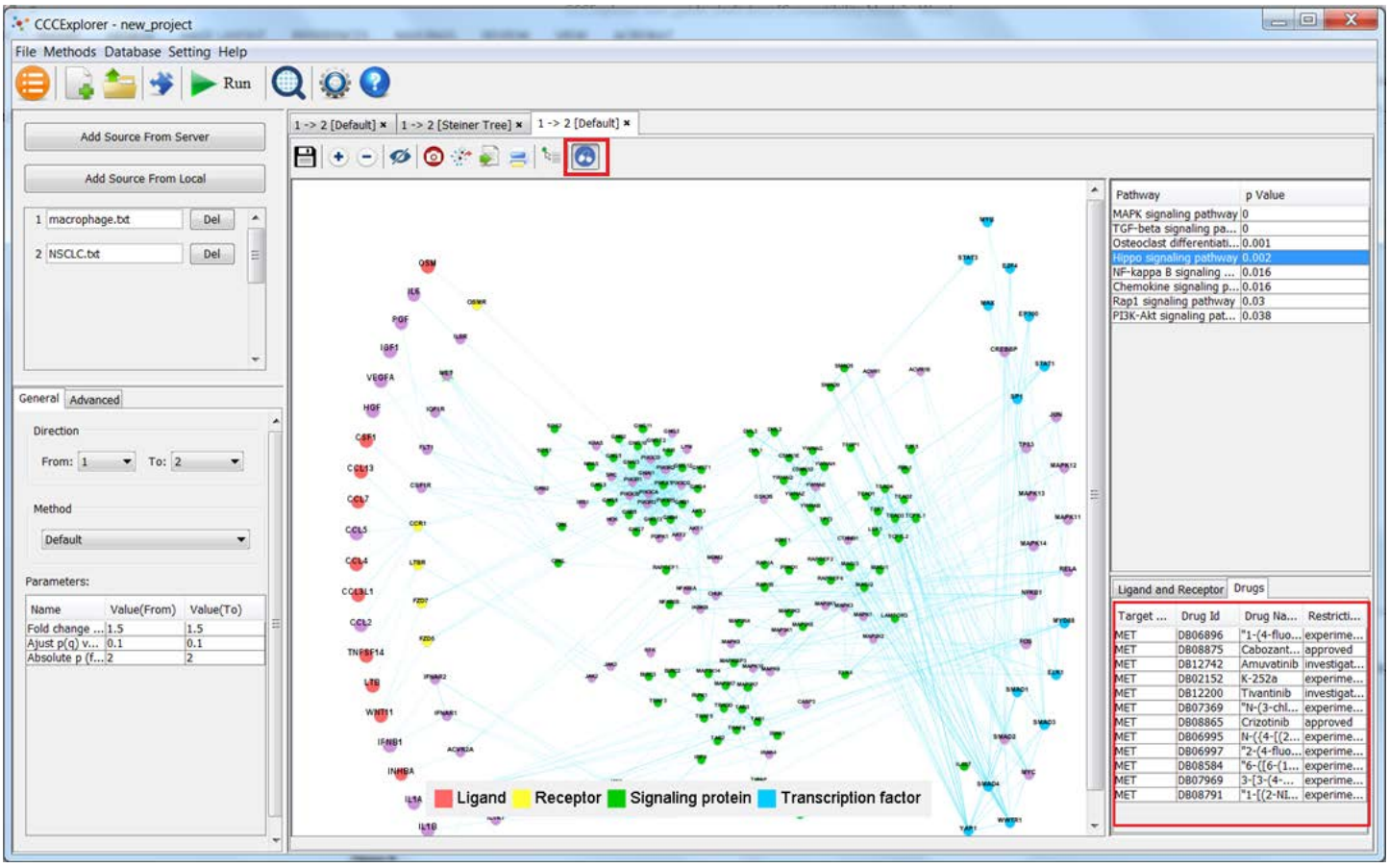

Figure 10

Delete a selected gene

To delete a gene in the network, right click on the corresponding node on the network. All edges connected to the gene will also be deleted (Figure 11).

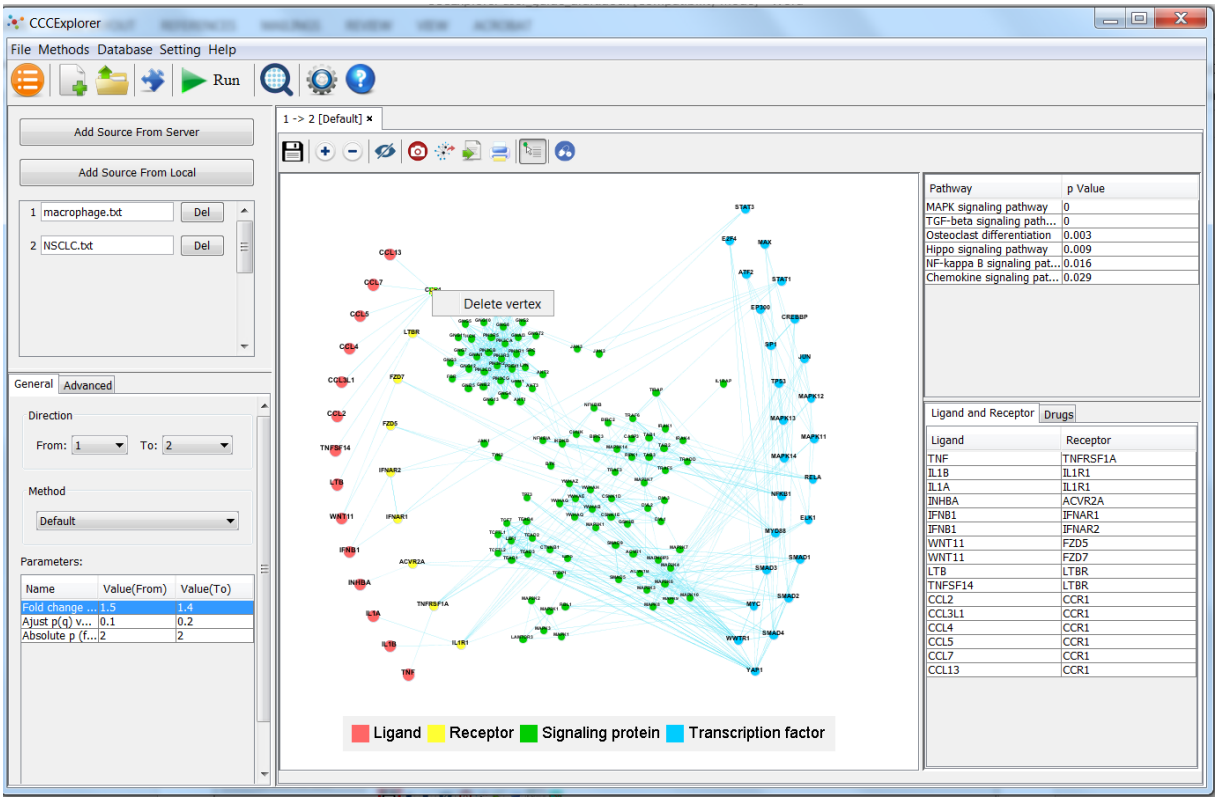

Figure 11

Save a project

To save a project, click on the save icon on the network window toolbar. Then provide a project name and click save (Figure 12).

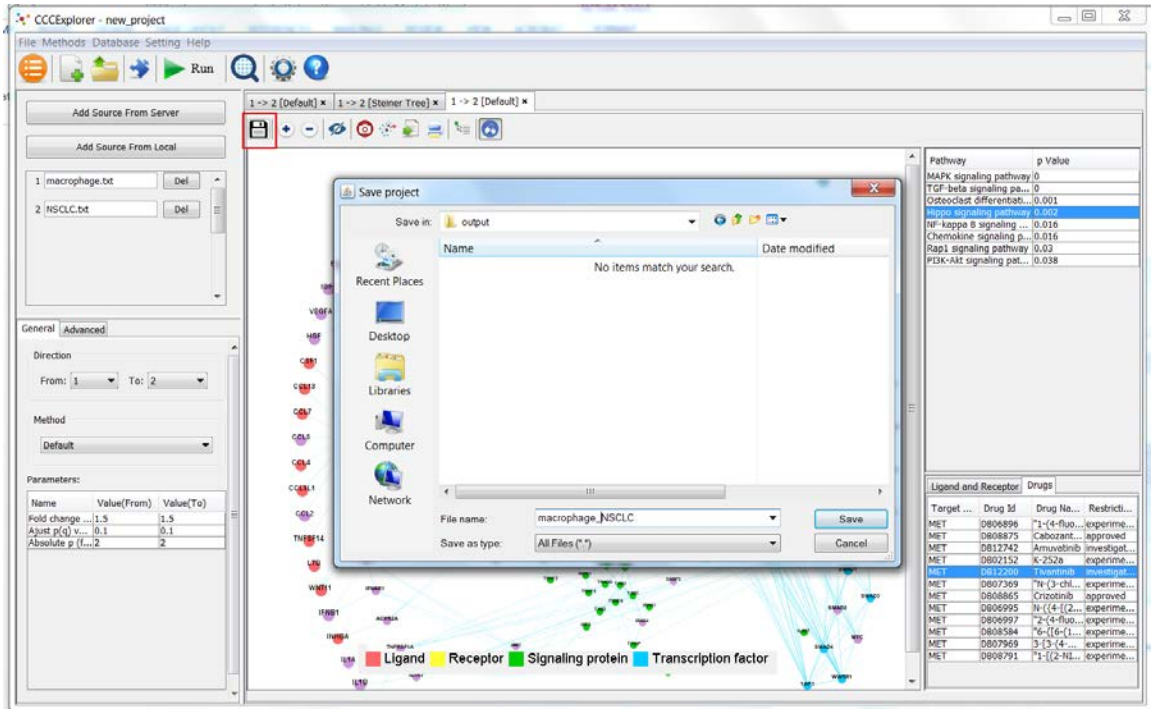

Figure 12

Several files will be created after saving a project. The list is shown in Figure 13. The .smab file is the project file and can be loaded to CCCExplorer to reopen the saved project.

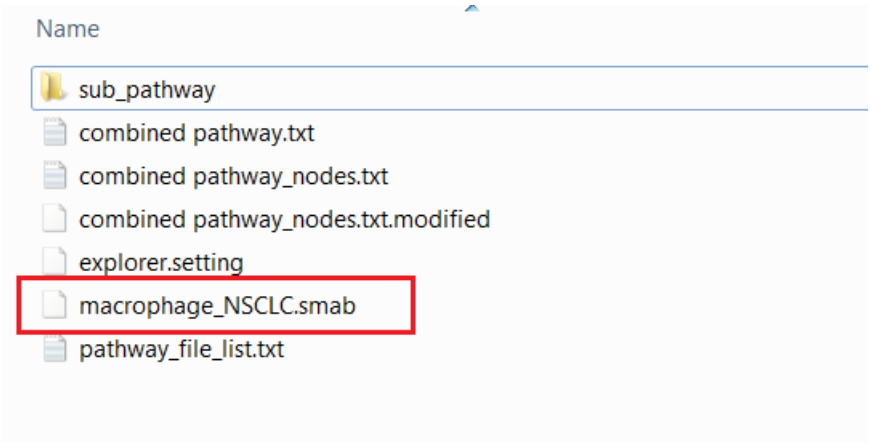

Figure 13

Open a saved project

Click on the folder icon on the main window toolbar and select a project file as shown in Figure 14.

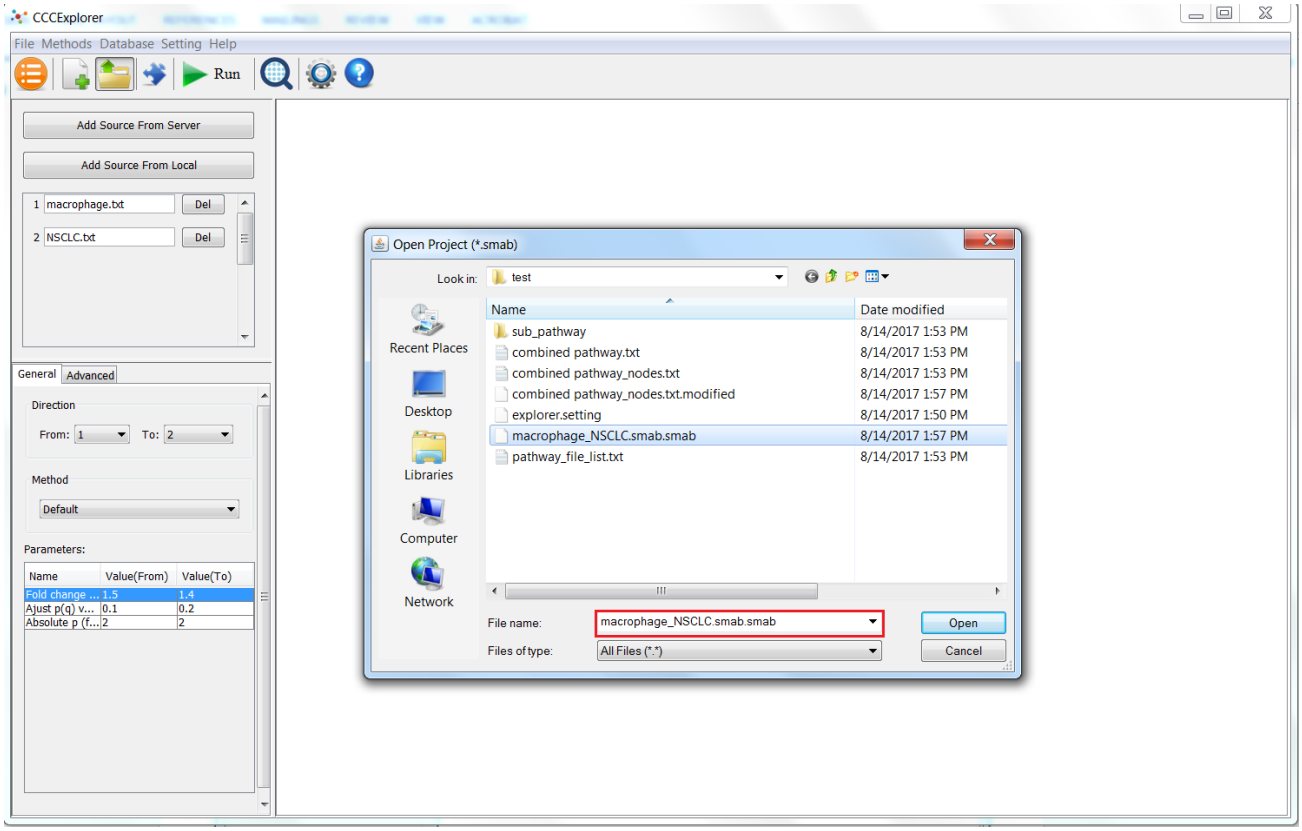

Figure 14

You can also open a project from the recent project list. Click on the first icon on the main window toolbar to open it (Figure 15).

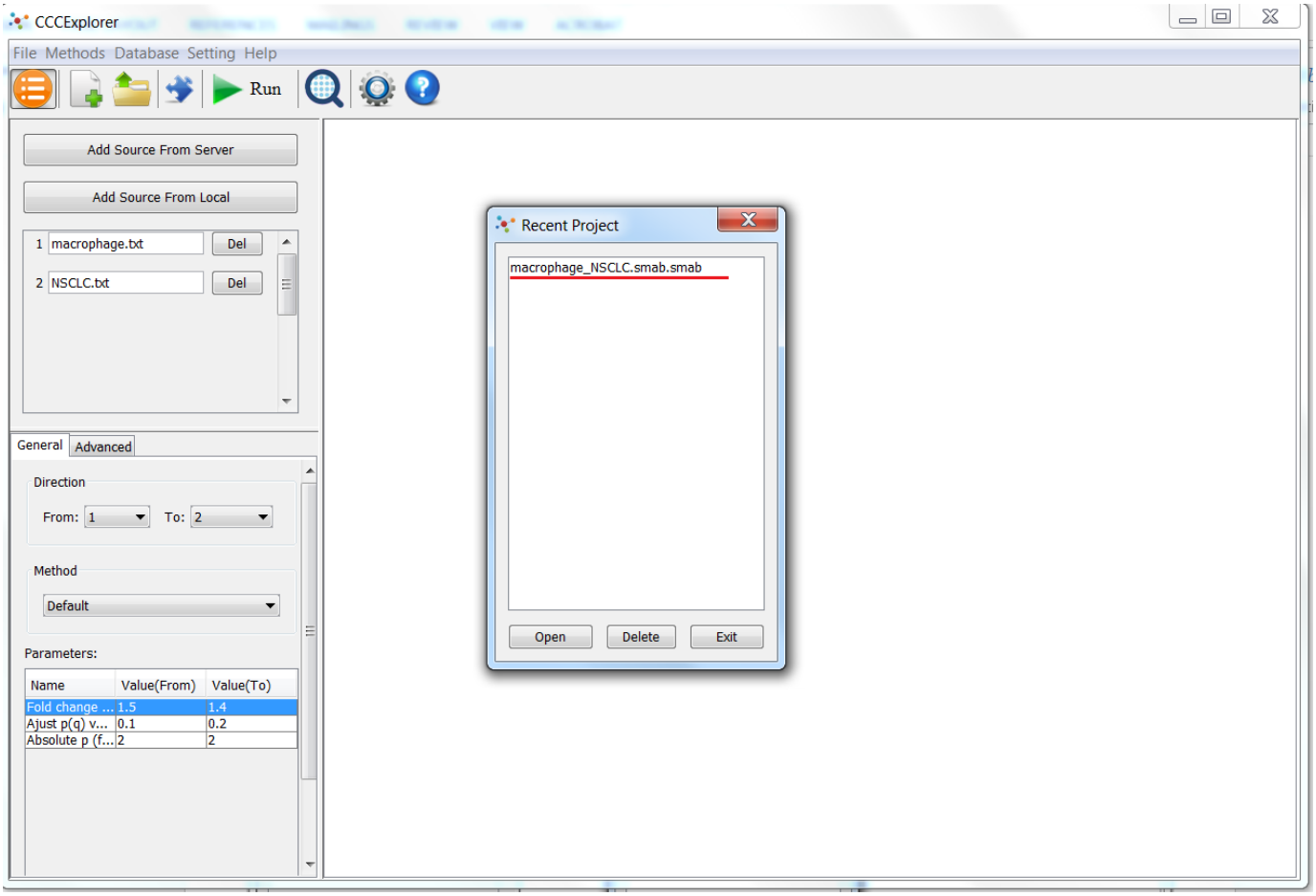

Figure 15

Search a gene

To check if a gene is involved in certain cell-cell communications, go to “Database->Search Gene” and type in the gene name. The predicted crosstalk and the expression profile of the gene will be shown as in Figures 16 and 17.

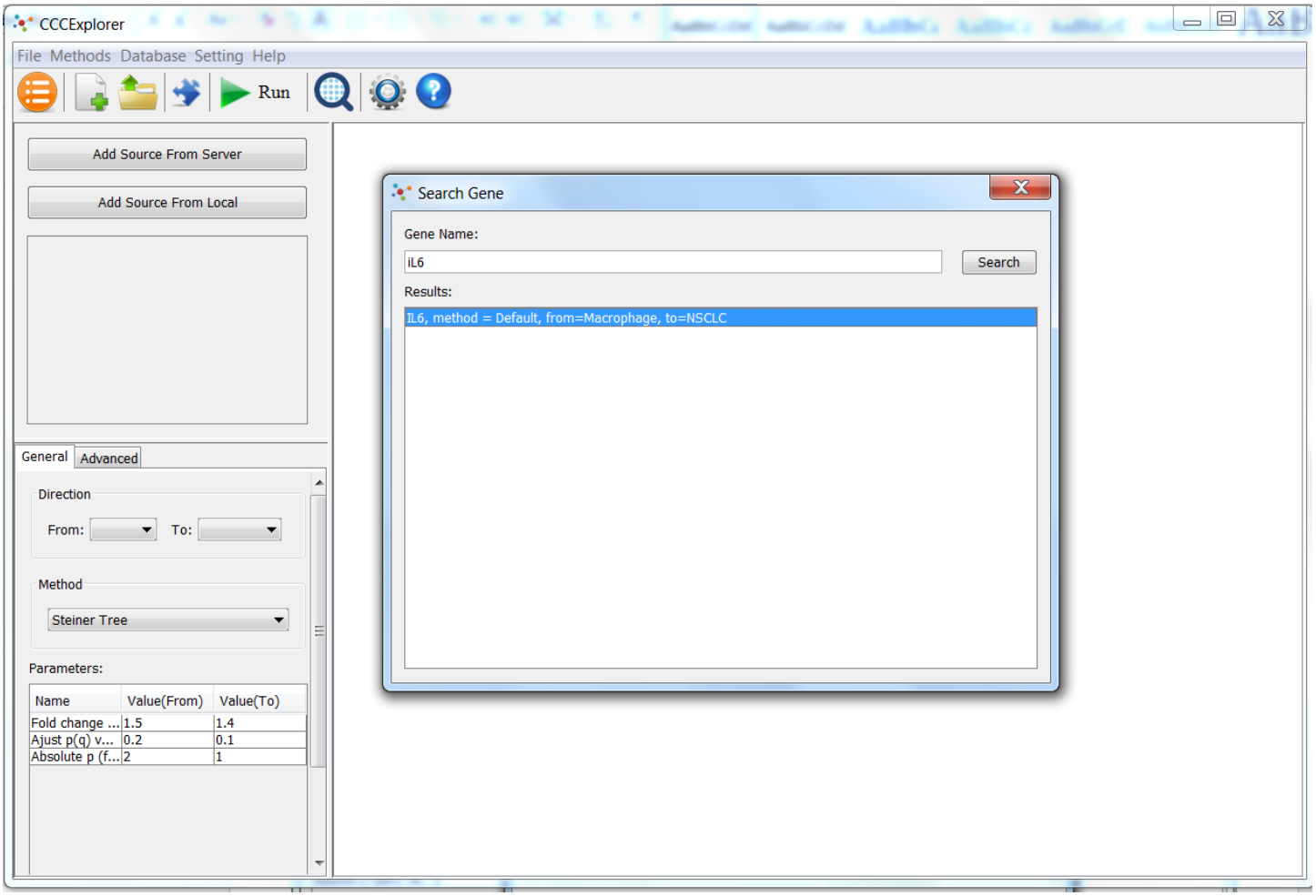

Figure 16

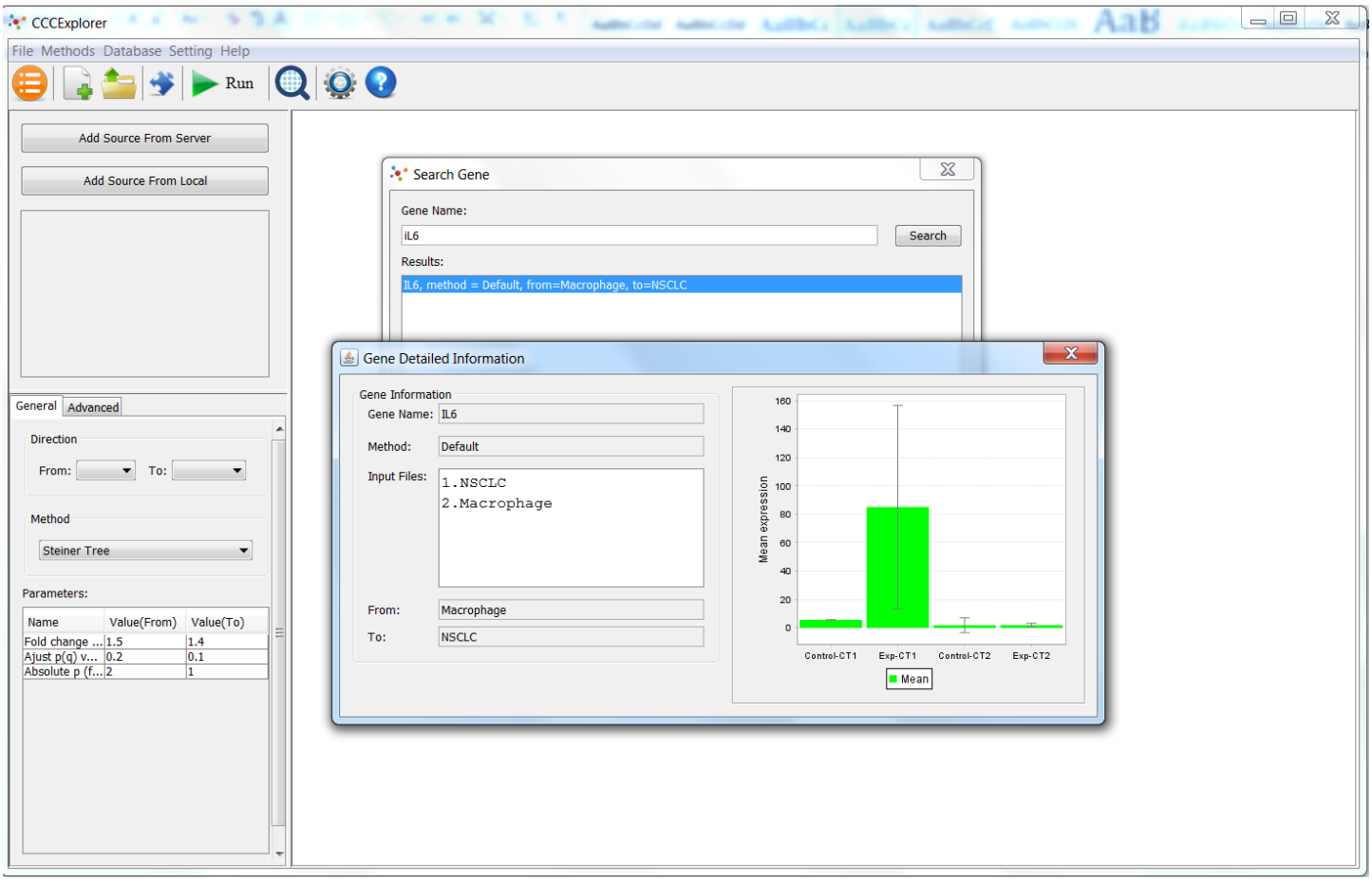
CCCExplorer
File Methods Database Setting Help
Run
Add Source From Server
Add Source From Local
General
Advanced
Direction
From:
To:
Method
Steiner Tree
Parameters:
Name
Value(From)
Value(To)
Fold change ...
1.5
1.4
Ajust p(q) v...
0.2
0.1
Absolute p (f...
2
1
Search Gene
Gene Name:
iL6
Search
Results:
IL6, method = Default, from=Macrophage, to=NSCLC
Gene Detailed Information
Gene Information
Gene Name:
IL6
Method:
Default
Input Files:
1.NSCLC
2.Macrophage
From:
Macrophage
To:
NSCLC
Mean expression
160
140
120
100
80
60
40
20
0
Control-CT1
Exp-CT1
Control-CT2
Exp-CT2
Mean

Figure 17

Add a method

CCCExplorer allows users to develop new methods for crosstalk analysis and add their algorithms. To add a new method into CCCExplorer, you need to provide two files. One is the configuration file, and the other is the script file. Currently, we support R language-based script file and we will support python and other languages in a future version. The format of the configuration file is JSON.

Figure 18 shows an example of the configuration file and Table 2 provides the descriptions for the keywords used in the configuration file.

steiner_tree.conf

```
{
  "algorithmName":"Steiner Tree",
  "fileName":"CCCExplorer_steiner_tree.R",
  "entryFunctionName":"CCCExplorer",
  "paramsArray":[
  {
      "description":"Fold change threshold",
      "valueOfFrom":"1.5",
      "valueOfTo":"1.4"
   },
  {
      "description":"Ajust p(q) value Threshold",
      "valueOfFrom":"0.2",
      "valueOfTo":"0.1"
  },
  {
      "description":"Absolute p (fpkm) value Threshold",
      "valueOfFrom":"2",
      "valueOfTo":"1"
  }
]
}
```

Figure 18

Table 2

Key names	Descriptions
algorithmName	The method name that is displayed on the CCCExplorer GUI
filename	The name of your script file
entryFunctionName	The main entry function in your script file
paramsArray	The parameters used in your method. You can provide a default value for each parameter.

For example, in the algorithm file, you may define an algorithm function named CCCExplorer.

```
CCCExplorer(dir,source_exprs,target_exprs,Fold_thres_S,q_thres_S,FPKM_thres_S,Fold_thres_T,q_thres_T,FPKM_thres_T)
```

The first three parameters are fixed. The rest are flexible parameters that are defined in your algorithm. Table 3 shows the detailed explanation for the parameters.

Table 3

Algorithm parameters	Descriptions	Comments
dir	The environment variables your script will use.	
source_exprs	The path containing the expression data file of From (donating) cell type	
target_exprs	The path containing the expression data file of To (recipient) cell type	
Fold_thres_S	The first parameter (“Fold change threshold”) of From cell type. (Figure 18)	Parameters for From (donating) cell type
q_thres_S	The second parameter (“Ajust p(q) value Threshold”) of From cell type. (Figure 18)	
FPKM_thres_S	The third parameter (“Absolute p (fpkm) value Threshold”) of From cell type. (Figure 18)	
Fold_thres_T	The first parameter (“Fold change threshold”) of To cell type. (Figure 18)	Parameters for To (Recipient) cell type
q_thres_T	The second parameter (“Ajust p(q) value Threshold”) of To cell type. (Figure 18)	
FPKM_thres_T	The third parameter (“Absolute p (fpkm) value Threshold”) of To cell type. (Figure 18)	

Figure 19 shows how the configuration file matches the parameters in the display UI.

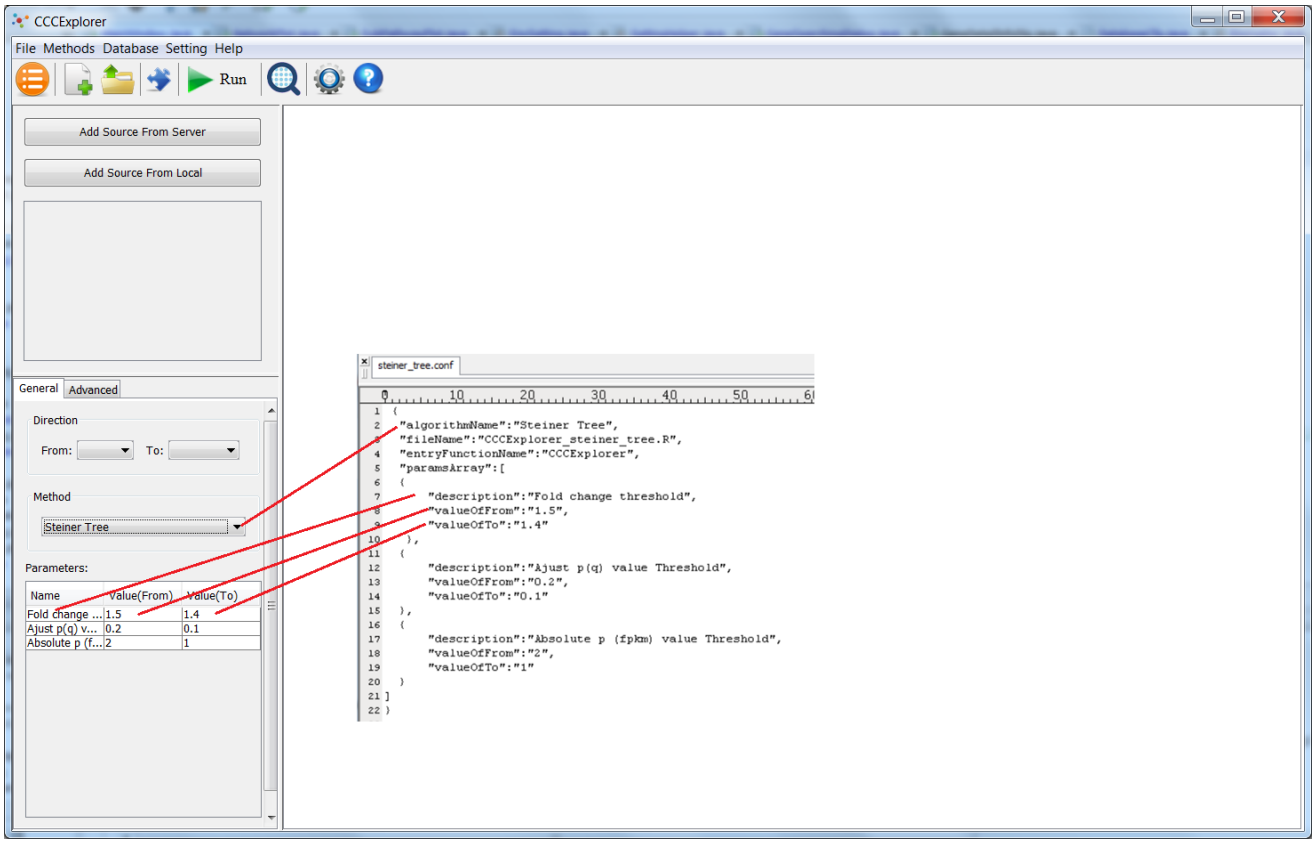

Figure 19

Click on “Methods->Add Method” to select a method configuration file. Make sure the algorithm script file is in the same folder as the configuration file.

From the preview box (Figure 20), you can check whether the format of the method is correct.

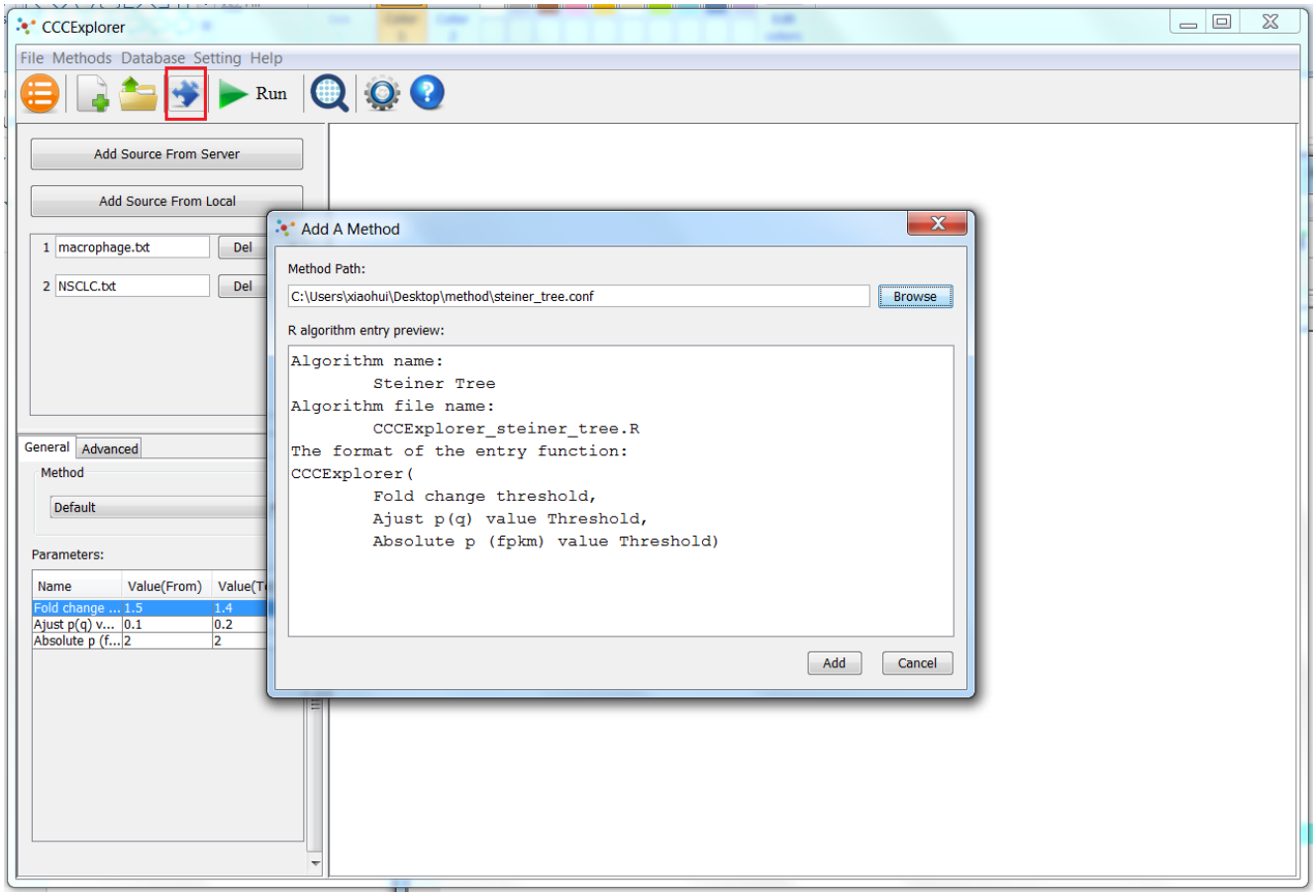

Figure 20

After adding a new method, you can select it from the left panel as shown in Figure 21.

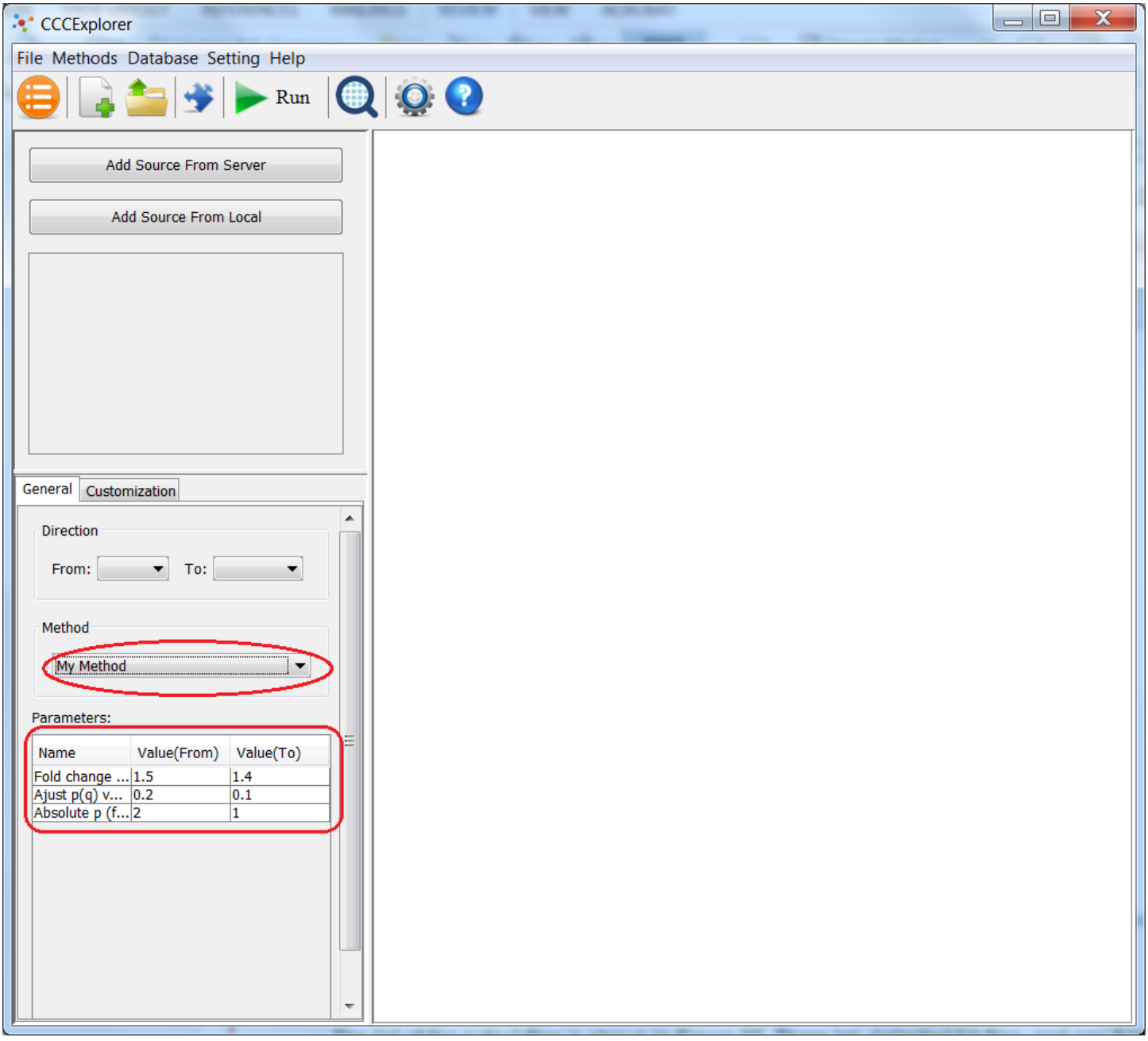

Figure 21

The format of the output files

The list of the output files is shown in Figure 22. Three tab delimited txt files, and one folder will be generated for each project.

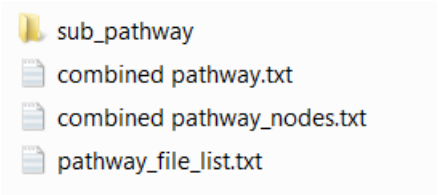

Figure 22

The file “combined pathway.txt” contains the overall crosstalk pathway information. Figure 23 shows the first few lines of this file.

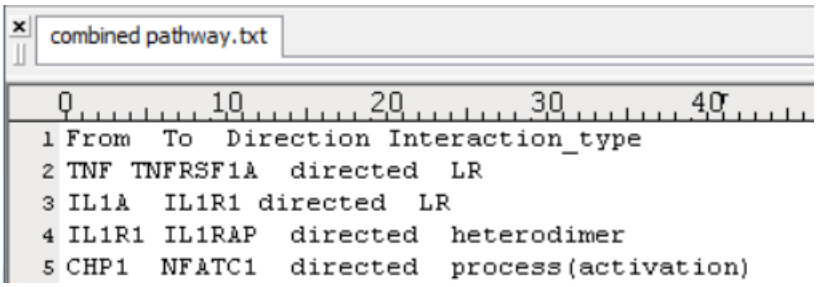

Figure 23

The first row is the header. There are four columns in this file. The first and second columns contain the gene name for each interaction in the output network. The third column denotes if the interaction is directed or not. The last column shows the interaction type, e.g., LR means ligand-receptor interaction.

The file “combined pathway_nodes.txt” contains the node information in the overall crosstalk signaling pathway. Figure 24 shows the first few lines of this file.

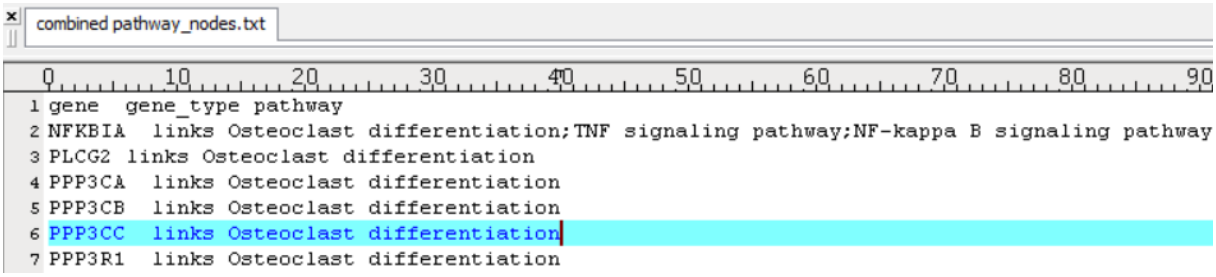

Figure 24

There are three columns in this file. The first column shows the gene name while the second column and third column contain gene type and pathway information of the gene, respectively.

The file “pathway_file_list.txt” contains the sub pathway information. Figure 25 displays the first few lines of this file.

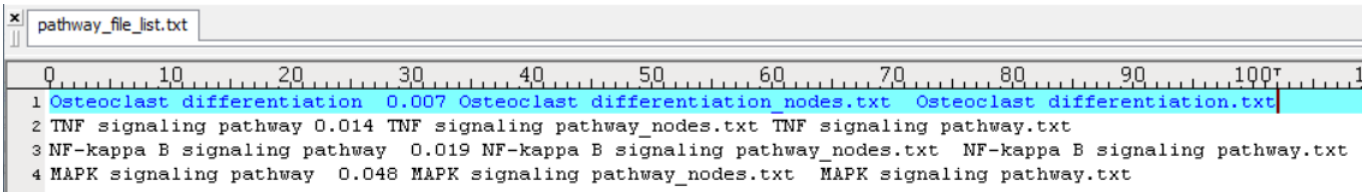

Figure 25

There are four columns in this file.

The first column is the name of sub pathway.

The second column is p value.

The third column is the name of the file containing node information for this sub pathway.

The forth column is the name of the file containing interaction information for this sub pathway.

The customized algorithm creates another report file (report.dat) in the current workspace. CCCExplorer will check this report file to decide if the script is running successfully.

The format of the report file is JSON (Figure 27).

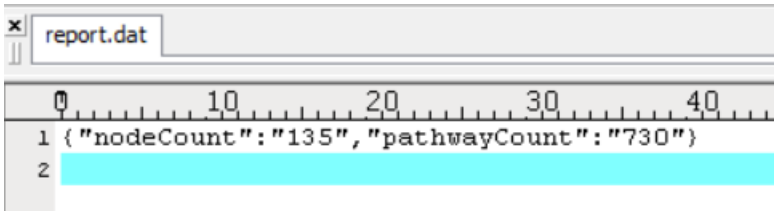

Figure 27

Note:

If you use your own method (your own R script file), you need to use the same output file names and file structure.

Delete a method

Go to “Methods->Manage Methods” to delete a method (see Figure 28).

You can only delete a customized method. The default algorithm cannot be deleted (see Figure 29).

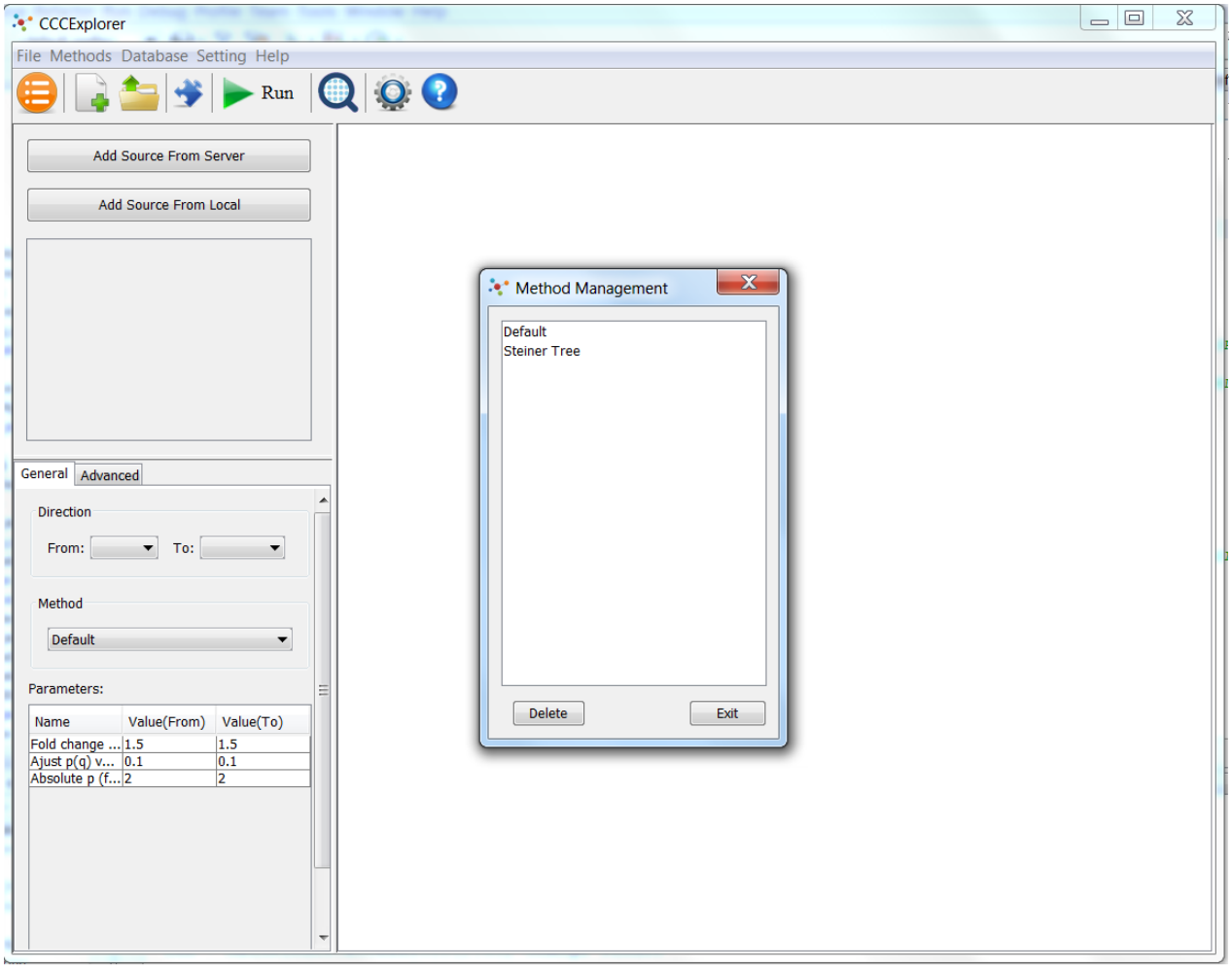

Figure 28

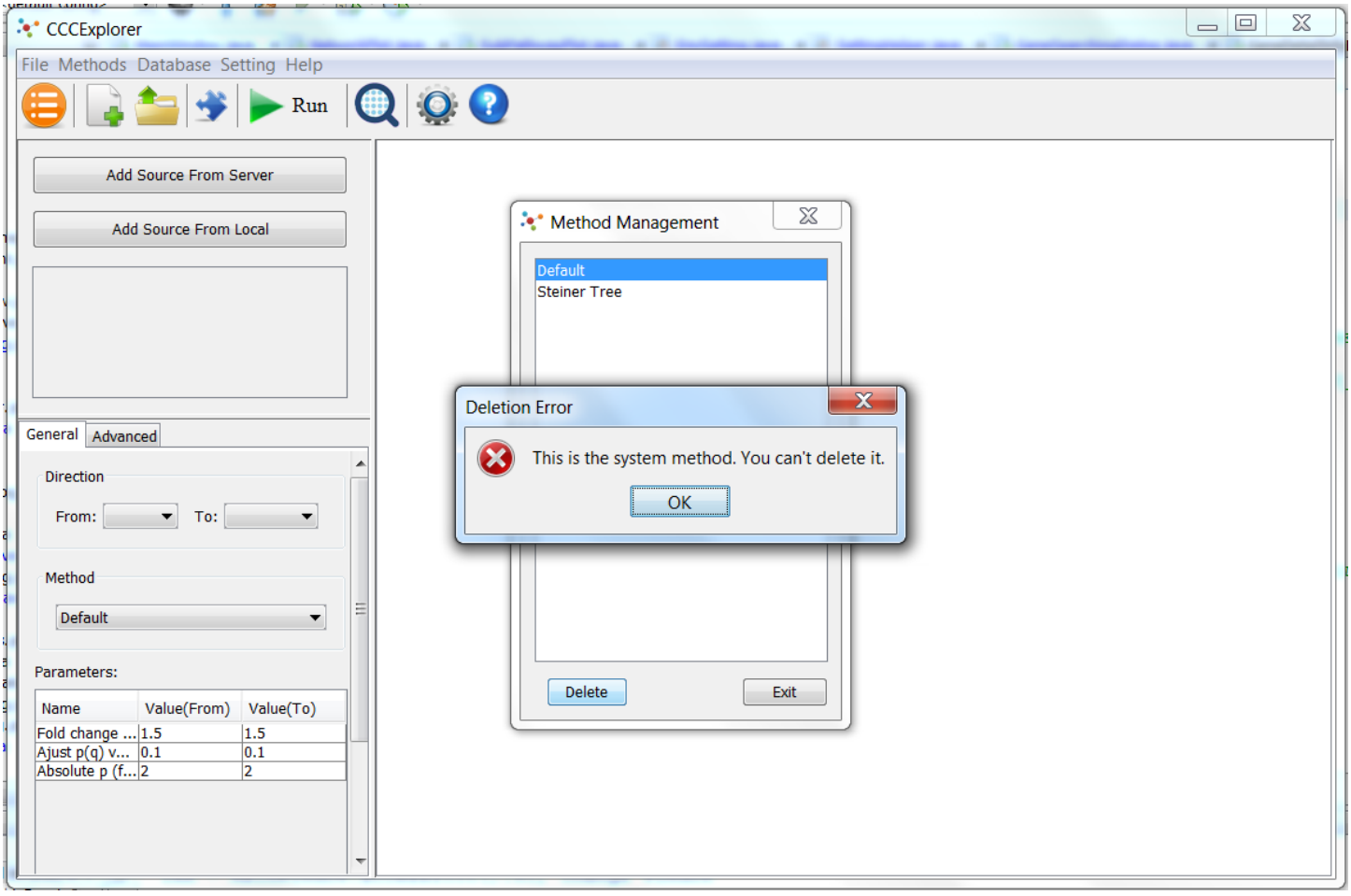
CCCExplorer
File Methods Database Setting Help
Run
Add Source From Server
Add Source From Local
General
Advanced
Direction
From:
To:
Method
Default
Parameters:
Name
Value(From)
Value(To)
Fold change ...
1.5
1.5
Ajust p(q) v...
0.1
0.1
Absolute p (f...
2
2
Method Management
Default
Steiner Tree
Deletion Error
This is the system method. You can't delete it.
OK
Delete
Exit

Figure 29