

S2C2

Single-cell Signaling Cell-cell Communication Explorer

v1.0.1

Contents

1. Installing S2C2 on Windows/ MacOS	3
1.1 Install JRE.....	3
1.2 Install R language and dependent packages	3
2. How to use S2C2	5
2.1 Run S2C2 Program.....	5
2.2 Run a project.....	5
2.3 Project Configuration Options on the Sidebar	8
2.4 Check the result of a project	10
2.5 Tool Functions	14
2.6 Other Functions.....	19

1. Installing S2C2 on Windows/ MacOS

1.1 Install JRE

S2C2 GUI (Graphical user interface software) is a java-based application. Thus, user need to install JRE (should be higher version as Java 21) before using it. The user can download JRE package from the link :

For the Windows user:

<https://www.oracle.com/java/technologies/downloads/#jdk21-windows>

For the Mac user:

<https://www.oracle.com/java/technologies/downloads/#jdk21-mac>

1.2 Install R language and dependent packages

Before using the S2C2 software, it is necessary to install the R programming language and several additional packages that S2C2 depends on for its data processing and analysis tasks. Below are the steps to install R and the required packages.

Installing R

1. Download R for Windows or MacOS from the Comprehensive R Archive Network (CRAN) website: <https://cran.r-project.org/bin/windows/base/>
2. Choose the latest version of R (4.3-2 or higher).
3. Follow the installation prompts to install R on your system.

Installing R Packages

Open RGui to install following packages.

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

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

# SingleCellExperiment package installation includes following 3 steps (Do not include this line)

>if (!require("BiocManager", quietly = TRUE))
>  install.packages("BiocManager")

> BiocManager::install("SingleCellExperiment")

# Presto package installation includes following 2 steps (Do not include this line)
> install.packages("devtools")
> devtools::install_github("immunogenomics/presto")

>install.packages("jsonlite")

>install.packages("DescTools")

> install.packages('homologene')
```

- **tidyverse**
The `tidyverse` package is a collection of R packages designed for data science that can simplify the installation process. By installing `tidyverse`, you automatically install several of the packages required by S2C2, such as `dplyr`, `stringr`, and `readxl`.
- **Igraph**
For network analysis and visualization
- **reshape**
To flexibly reshape data
- **Seurat**
Single-cell genomics analysis
- **SingleCellExperiment**
Defines S4 classes for single cell genomics data
- **openxlsx**

To read and write XLSX files

- preston
For rapid calculation of gene set enrichment scores
- jsonlite
For robust and quick JSON data parsing and generation
- DescTools
For descriptive statistics
- homologene
It allows searching for gene homologs across species.

2. How to use S2C2

2.1 Run S2C2 Program

Decompress STCT_v1.zip

1. Double click "STCT.exe" to run it.
Or run with
> java -jar STCT.jar
2. Please ensure that the following R script files are located in the same running directory as 'STCT.jar':
 - '1_start_LR_identified.R'
 - '2_identify_select_pathways_backup.R'
 - '3_activated_pathways.R'
 - '4_permutation_for_significance.R'
 - '5_visualization.R'
 - 'STCT.R'

2.2 Run a project

To run a project in STCT, follow these steps:

1. Click on 'Load Data' (refer to Figure 1) to the data loading interface.
2. Add the '.rds` data files you wish to analyze. Please ensure that you upload '.rds` format data only, as other data formats will not be permitted for execution (see Figure 2).

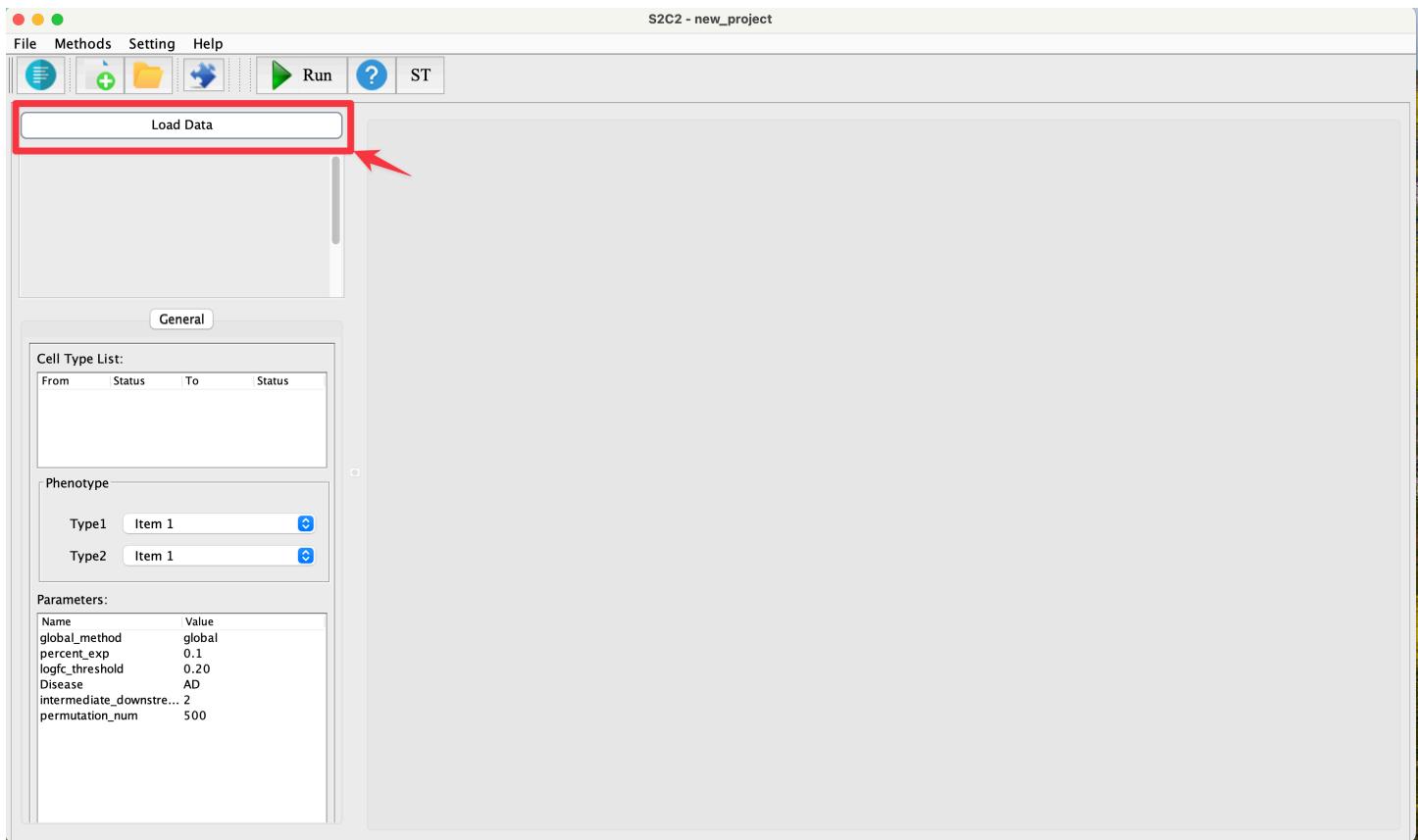


Figure 1. Load Data Button

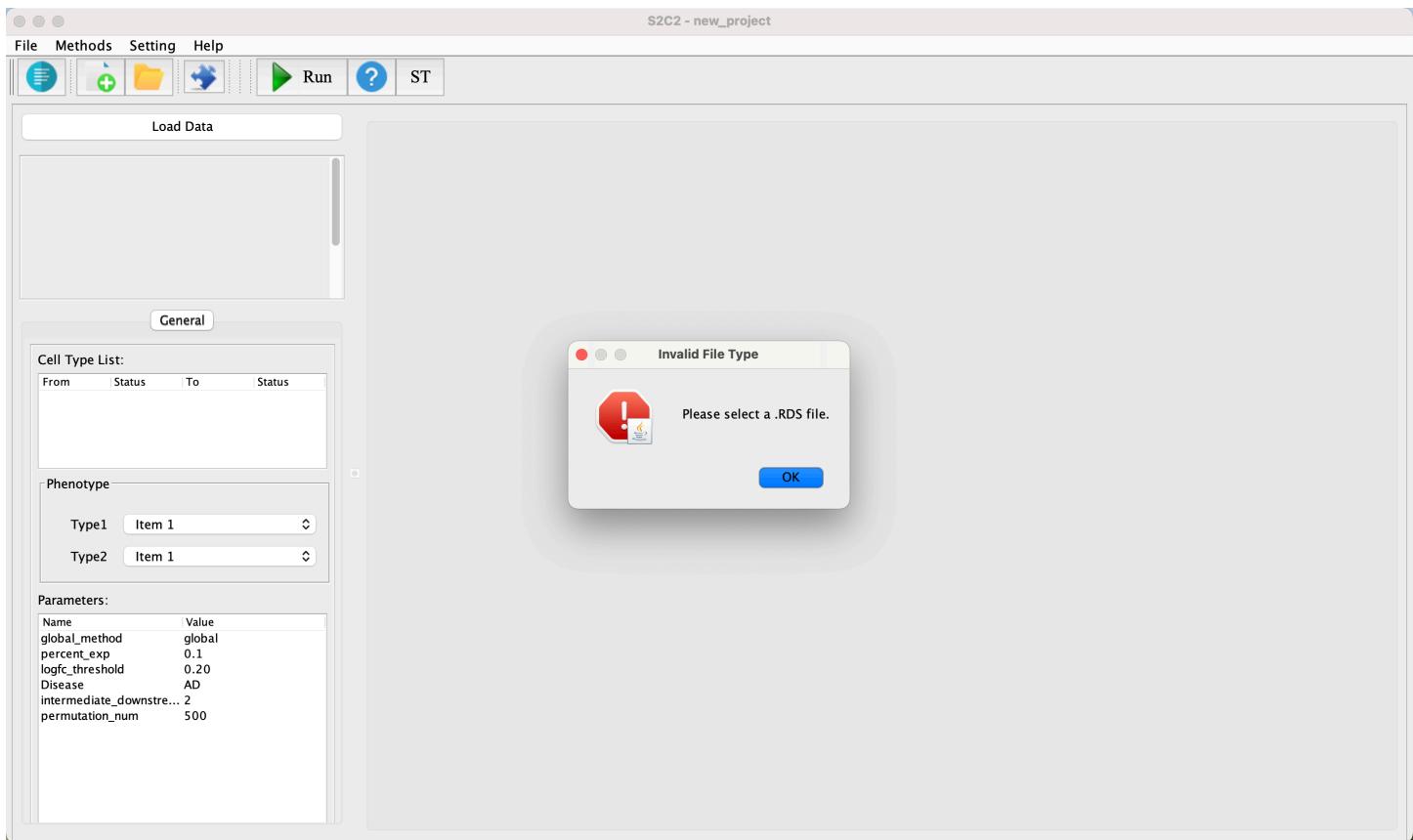


Figure 2. Load a wrong format data

*.rds data example ():

Name	Type	Value
seurat_object	S4 [36066 x 61472] (SeuratO)	S4 object of class Seurat
assays	list [1]	List of length 1
meta.data	list [61472 x 15] (\$3: data.fra	A data.frame with 61472 rows and 15 columns
orig.ident	factor	Factor with 3 levels: "b1", "b2", "b3"
nCount_RNA	double [61472]	5243 4912 4563 5824 6438 1898 ...
nFeature_RNA	integer [61472]	2002 2136 1972 2252 2634 1204 ...
X	character [61472]	'AACCCAAAGCTCG-1' 'AACCCAAGGATGCCG-1' 'AACGAACTTCGAAGG-1' 'AAAGAACGTATTCCGA ...
Sample.ID	character [61472]	'Sample-19' 'Sample-19' 'Sample-19' 'Sample-19' 'Sample-19' 'Sample-19' ...
Batch	integer [61472]	3 3 3 3 3 3 ...
Sex	character [61472]	'F' 'F' 'F' 'F' 'F' ...
Age	integer [61472]	90 90 90 90 90 90 ...
Diagnosis	character [61472]	'AD' 'AD' 'AD' 'AD' 'AD' ...
UMAP_1	double [61472]	-0.439 5.149 -0.249 -1.295 -0.350 0.649 ...
UMAP_2	double [61472]	8.75 11.07 6.65 8.34 10.17 7.59 ...
cluster	character [61472]	'ODC9' 'ODC7' 'ODC1' 'ODC2' 'ODC9' 'ODC1' ...
celltype	character [61472]	'ODC' 'ODC' 'ODC' 'ODC' 'ODC' 'ODC' ...
RNA_snn_res.0.9	factor	Factor with 34 levels: "0", "1", "2", "3", "4", "5", ...
seurat_clusters	factor	Factor with 34 levels: "0", "1", "2", "3", "4", "5", ...
active.assay	character [1]	'RNA'
active.ident	factor	Factor with 7 levels: "ODC", "MG", "OPC", "INH", "EX", "ASC", ...
graphs	list [2]	List of length 2
neighbors	list [0]	List of length 0
reductions	list [4]	List of length 4
images	list [0]	List of length 0
project.name	character [1]	'SeuratProject'
misc	list [0]	List of length 0
version	list [1] (\$3: package_version, i	List of length 1
commands	list [7]	List of length 7
tools	list [0]	List of length 0

Figure 3.rds data example

2.3 Project Configuration Options on the Sidebar

When configuring a project within STCT, the left panel as Figure4 provides a series of options that allow you to specify the details of your analysis. Each of these options is crucial for tailoring the exploration of cell-cell communication based on your dataset and research needs:

1. Cell Types

This dropdown menu is populated with features extracted from the metadata within your .rds file, processed by the Seurat tool. It enables the user to designate which feature represents the cell type within their dataset. By selecting a cell type from this menu, you are determining the focus of your cell-cell communication analysis.

2. Disease Condition Name

Similar to the cell types, this dropdown allows the user to select from the various conditions parsed from the .rds metadata.

3. Cell Type List

After the cell types and disease condition names are selected, clicking the "Confirm" button populates the "Cell Type List" with specific information about each cell type.

4. Phenotype

The "Phenotype" section's "Type1" and "Type2" fields are populated with detailed data parsed from the chosen disease condition name. This allows for further specification in the analysis.

5. Parameters

The "Parameters" section of the STCT software is where you can adjust the settings that control various aspects of the analysis algorithm.

- global_method

This option allows the user to choose between 'global' and 'local' methods. Default option is global.

- percent_exp

only test genes that are detected in a minimum fraction of percent_exp cells in either of the two populations. Meant to speed up the function by not testing genes that are very infrequently expression. It must be a numeric value between 0 and 1.

- logfc_threshold

Limit testing to genes which show, on average, at least X-fold difference (log-scale) between the two groups of cells. Increasing thresh.use speeds up the function, but can miss weaker signals.The default value is 0.2.

- intermediate_downstream_gene_num

This parameter defines the minimum number of intermediate downstream genes in the pathway branch to consider. The default value is 2.

- permutation_num

The 'permutation_num' sets the number of permutations for the permutation test. This test is used to estimate the significance of the pathway. The default value is 1000.

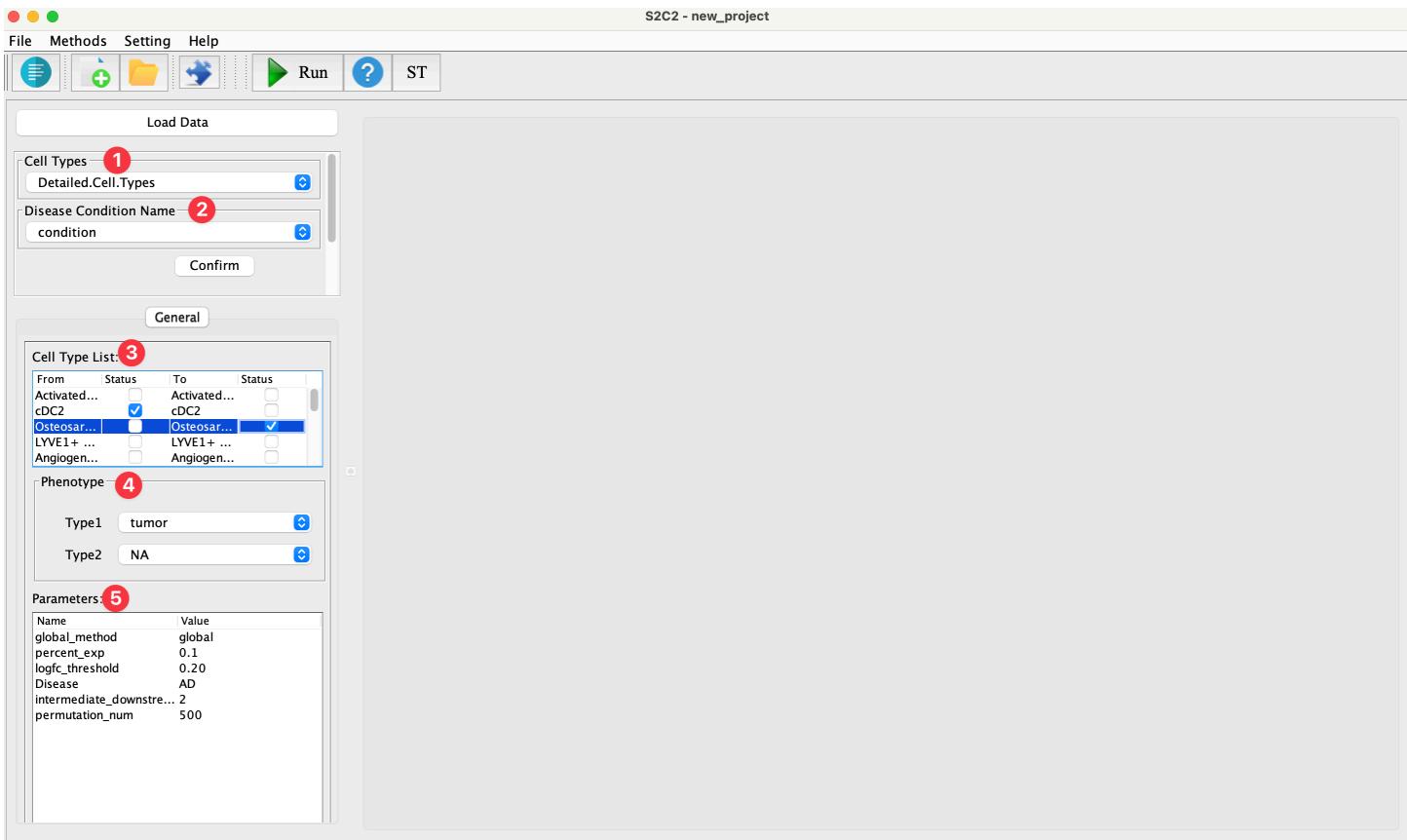


Figure 4. Project Configuration Options

2.4 Check the result of a project

After click Run button, there is an output window that will show up as Figure 4.

In Figure 5, an Overview is clearly displayed, showcasing the selected cell types. Each cell type is situated along its respective path, annotated with the number of sub-pathways available. By double-clicking on any of these pathways, you can dive into a detailed page that presents a more in-depth analysis of the selected cell type's signaling sub-pathways(Figure 6).

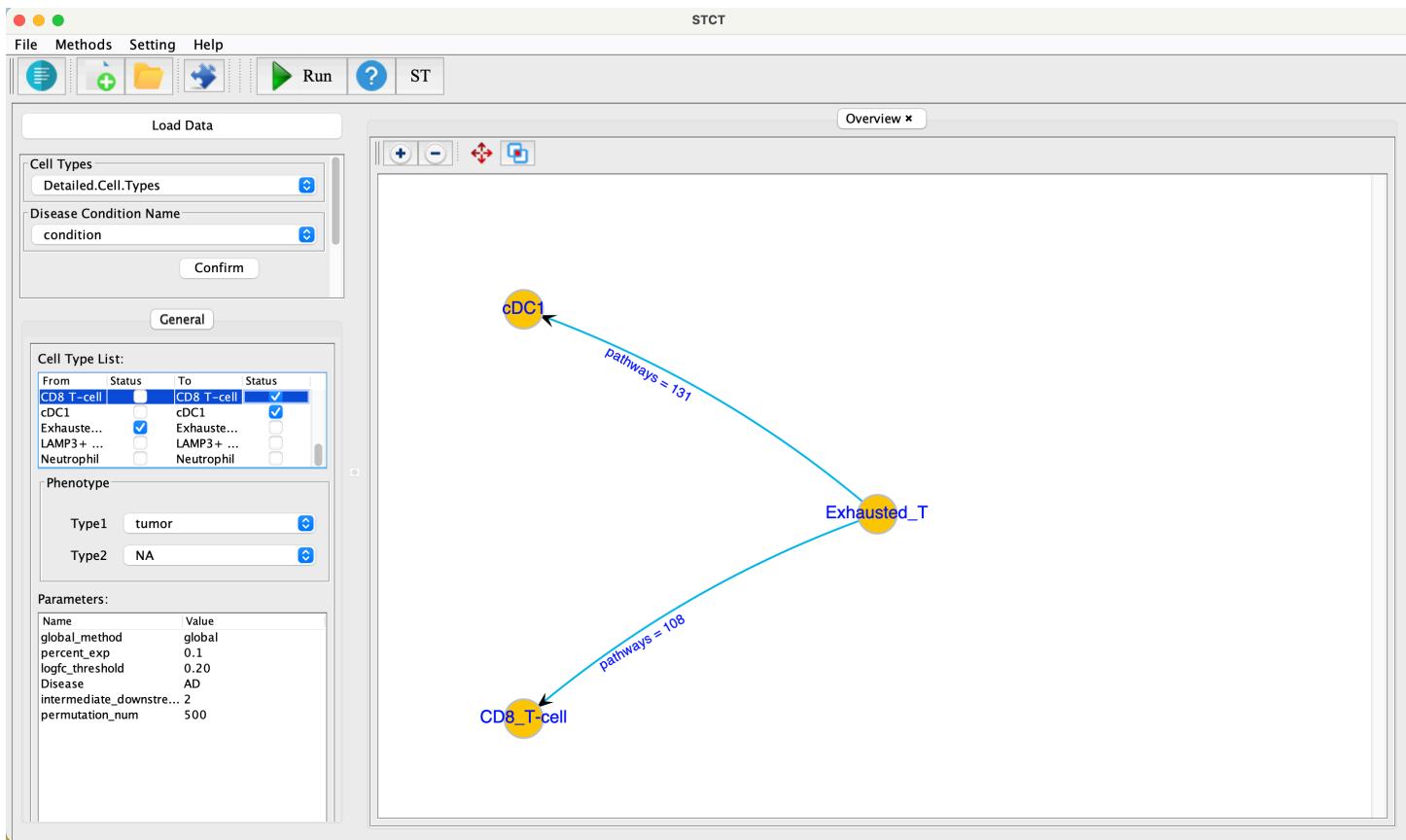


Figure 5. Network overview

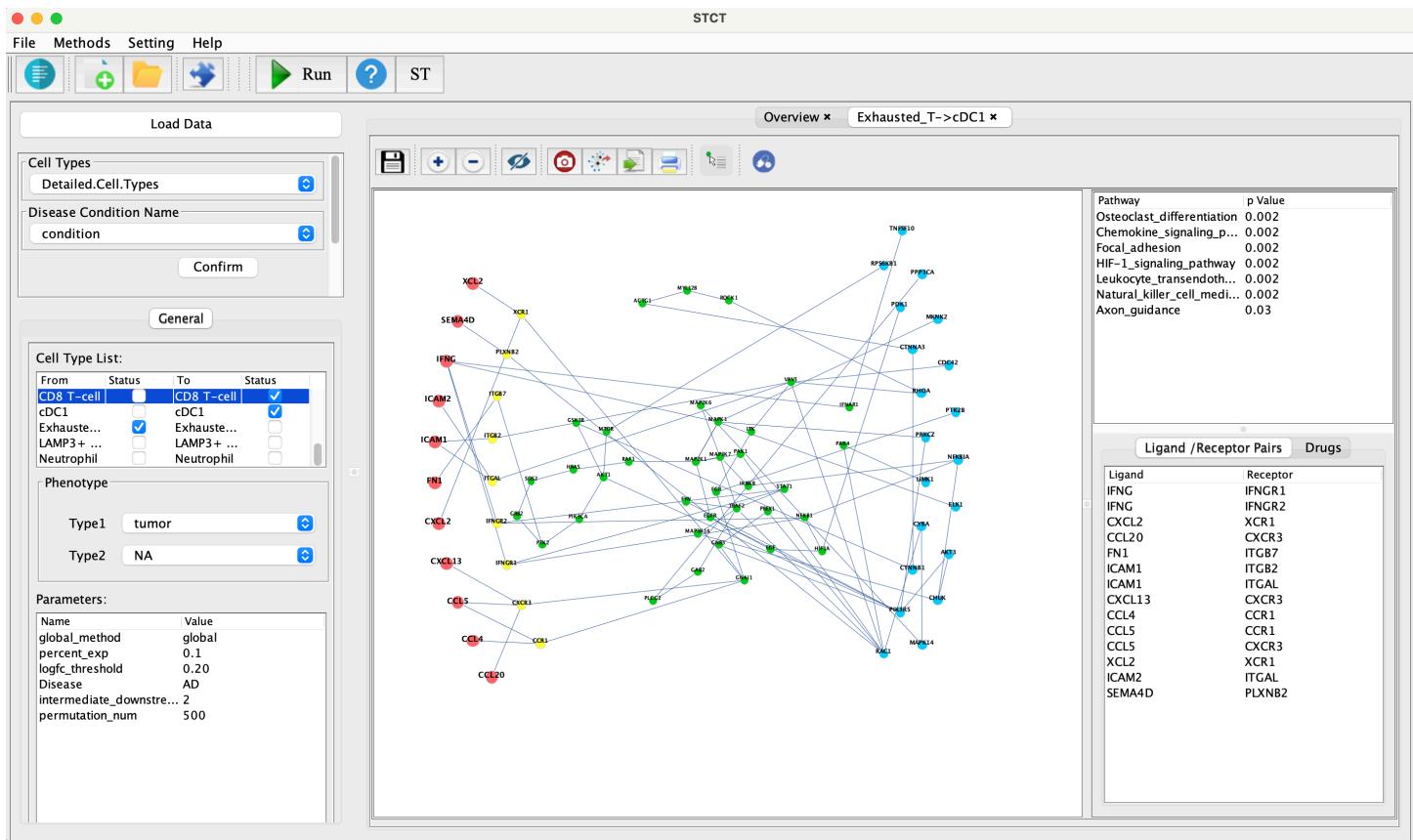


Figure 6. Sub-pathway details

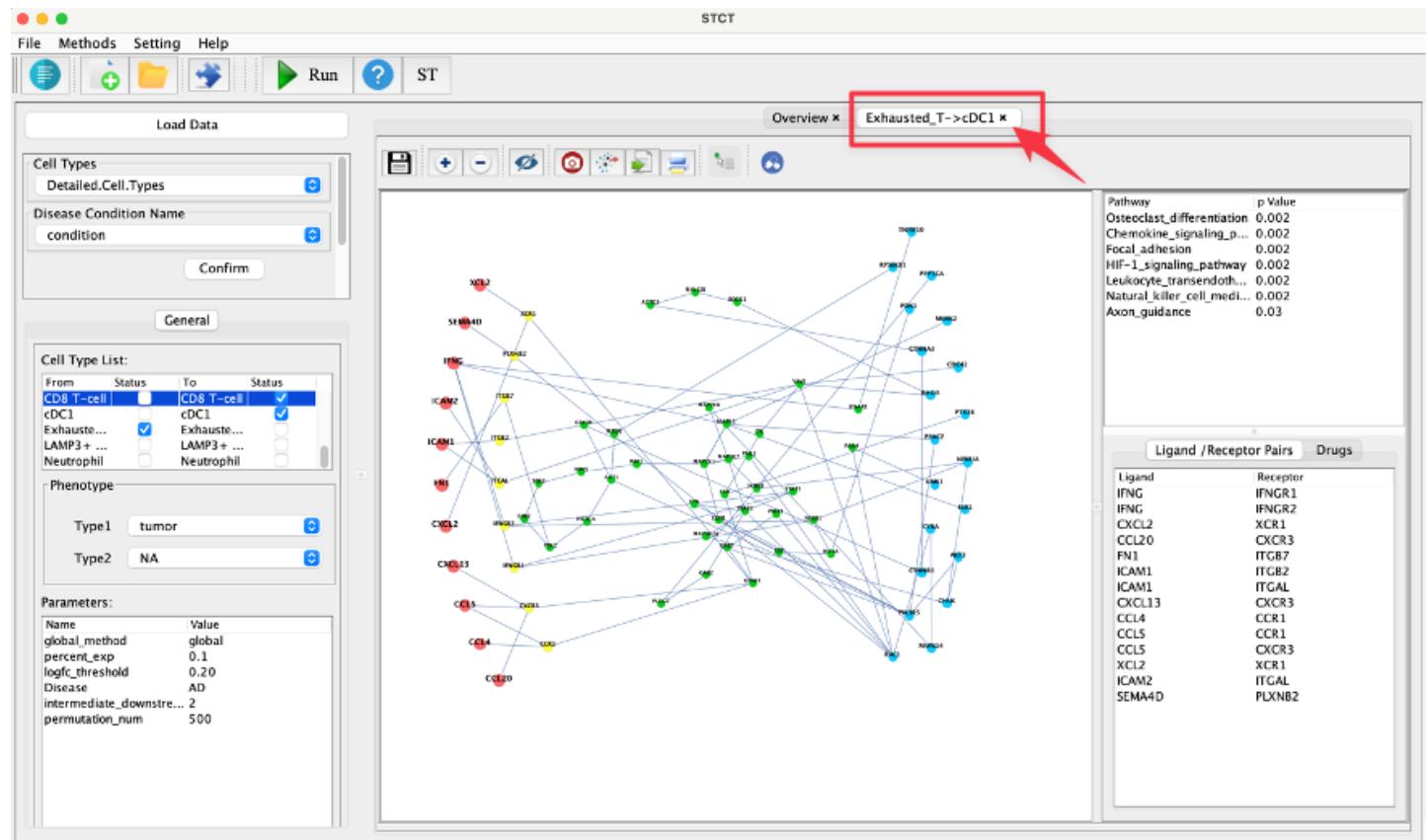


Figure 7. Back to the Network Overview

If users with to go back to overview of network, click the cross on the title tab part.

2.5 Tool Functions

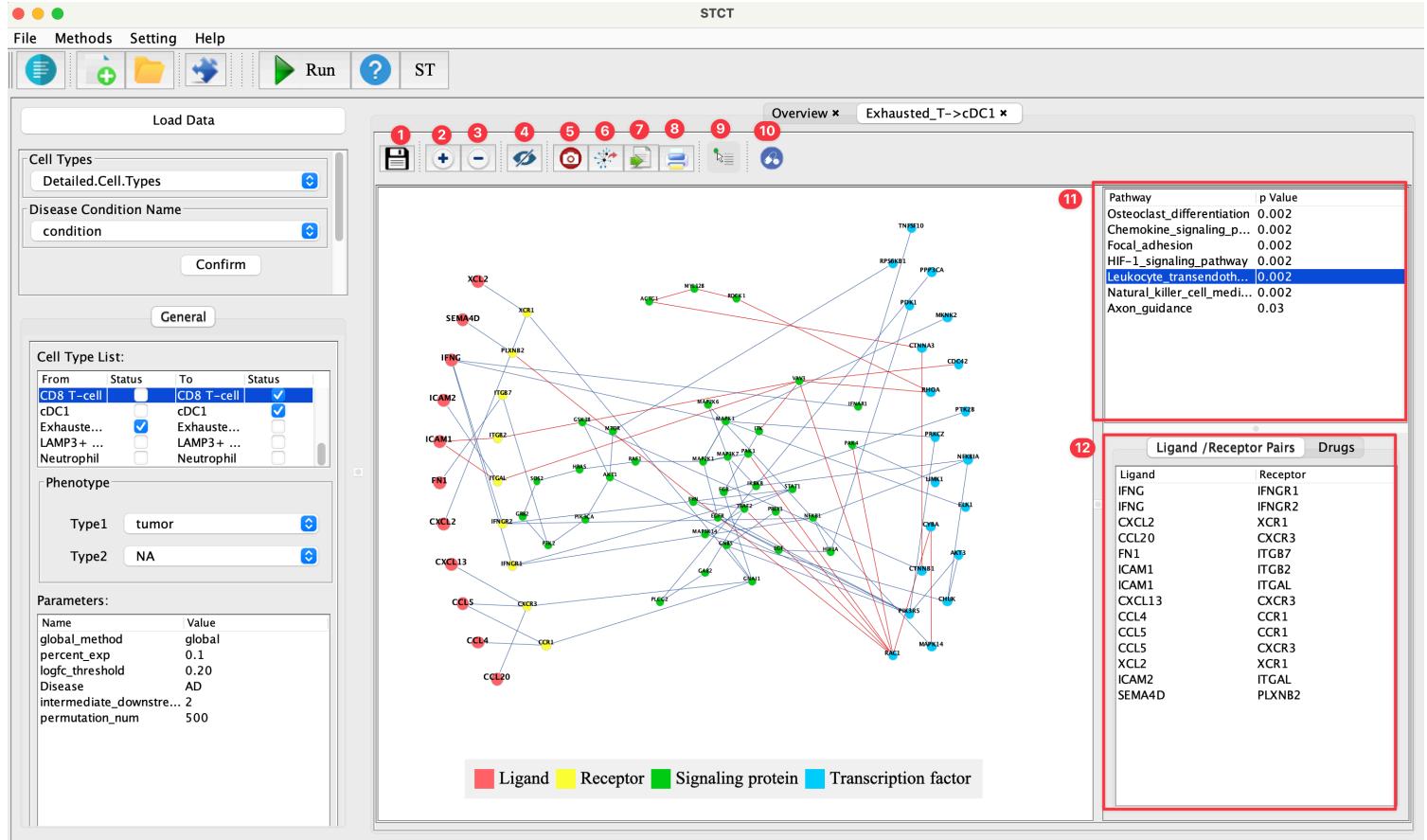


Figure 8. Tools

The interface shown in Figure 8 provides a range of functionalities to interact with the project results. Here's a description of the available tool functions:

1. **Save Project:** This function allows you to save the current state of your project, including all configurations, data, and analyses you have performed.

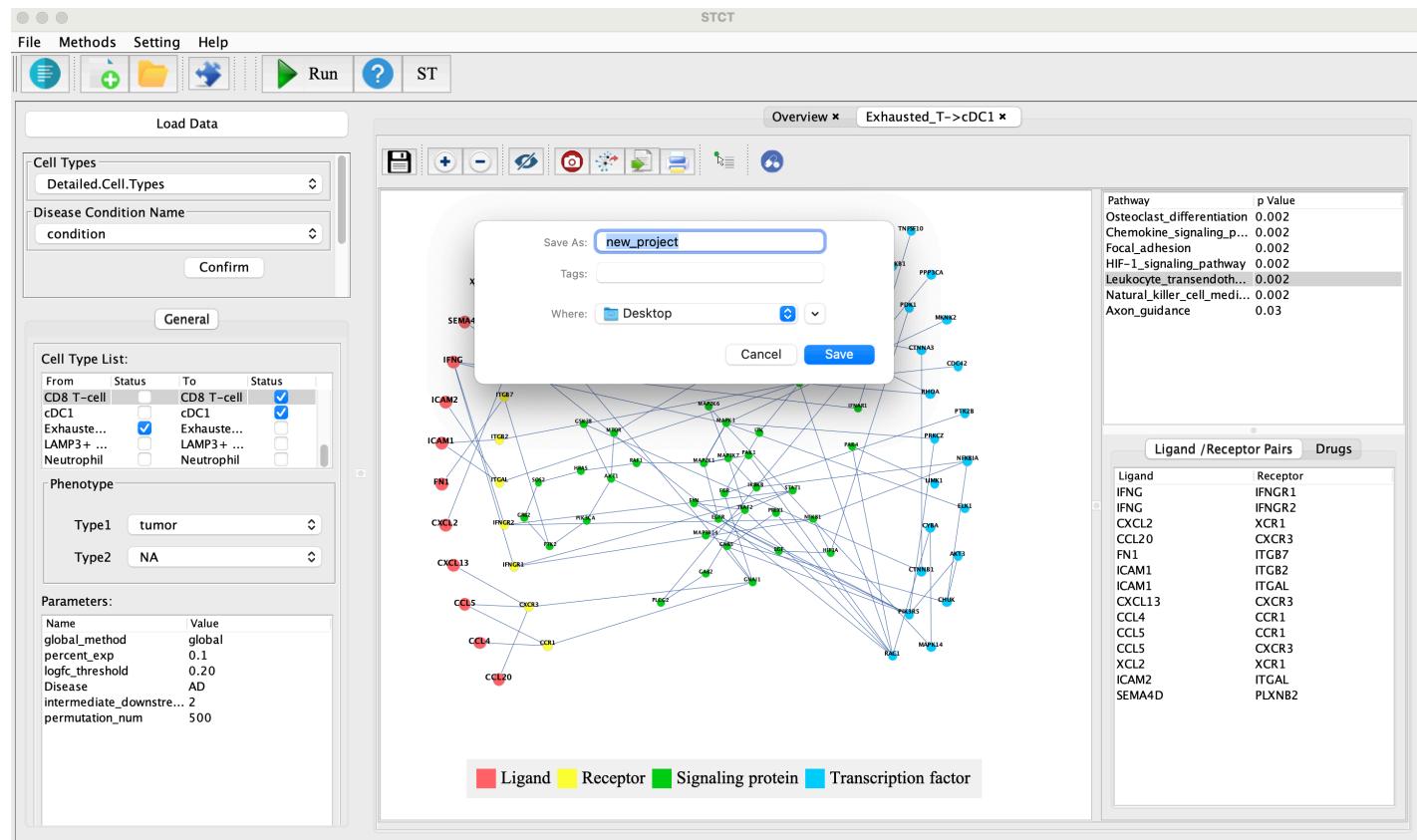


Figure 9. Save Project

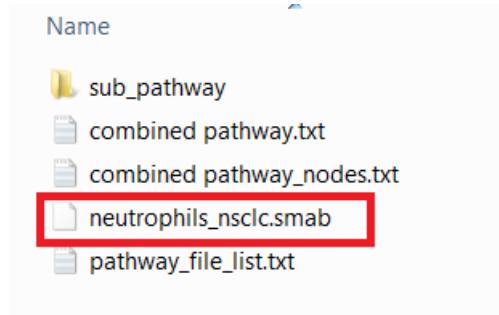


Figure10. Project directory

There are several files will be created after clicking save button. The file list is in Figure 10. The .smab file is the project file. User can open this file with. smab to load a project.

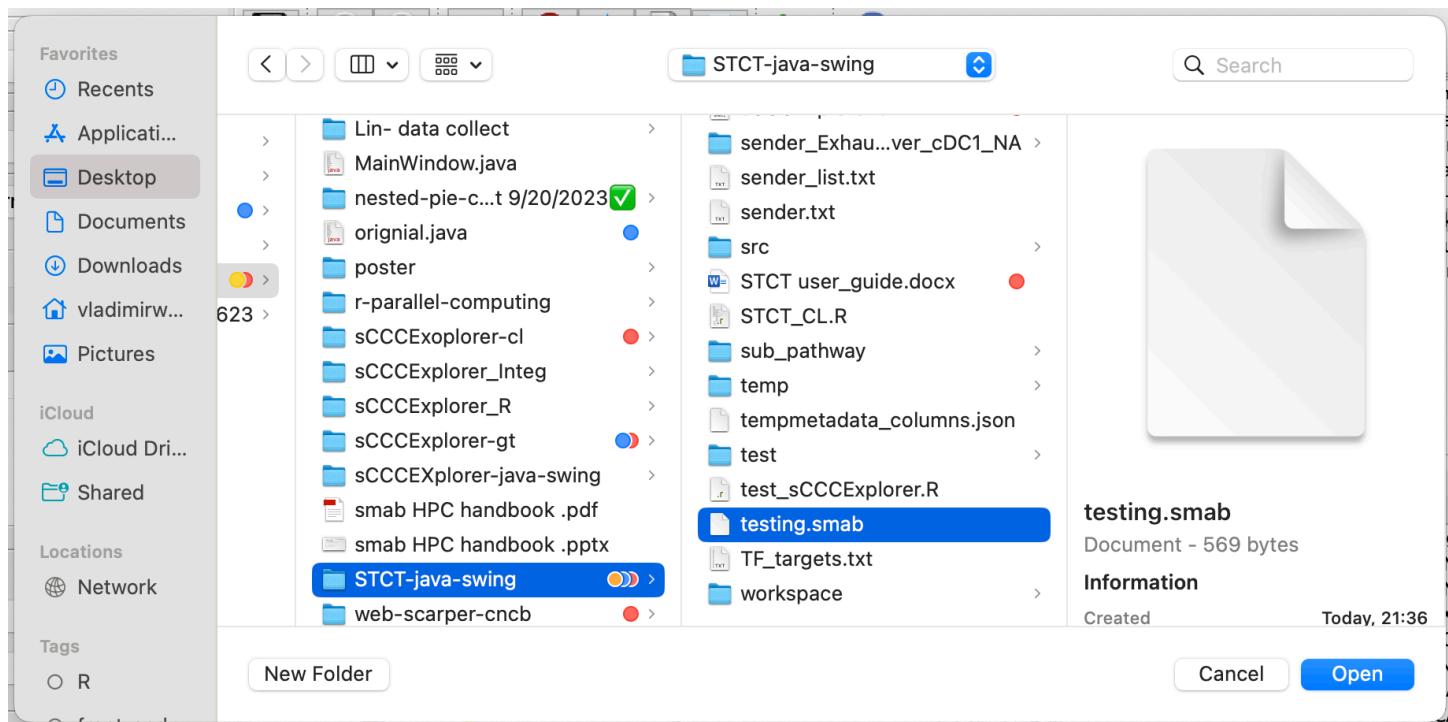


Figure 11. Open Project

To open a saved project, go to the 'File' menu, select 'Open Project', and then choose a .smab file to load the previously saved work. (Figure 11)

Users can also open project from the recent project button in figure 12

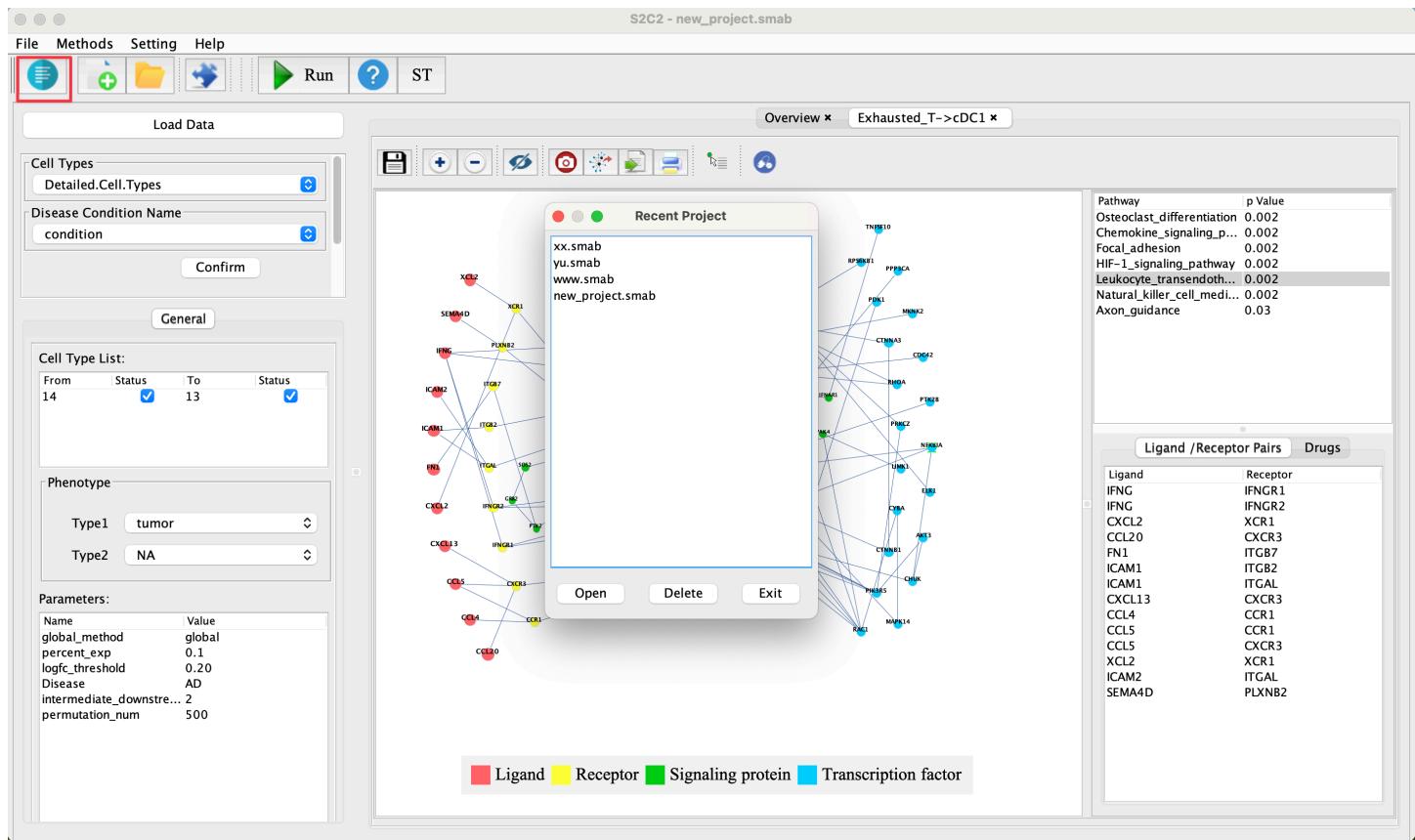


Figure 12. Load Project from the recent project button

2. **Zoom-in Button:** Zoom in the interface.
3. **Zoom-out Button:** Zoom out the interface.
4. **Hide/Show Edges:** Toggles the visibility of the edges in the network.
5. **Screenshot (All Information):** Takes a screenshot of the current view, capturing all the information including the network and any open side panels or additional data.
6. **Screenshot (Network Only):** Captures only the network part of the current view, excluding additional information like legend.
7. **Export Pathway Data(csv):** Allows the user to export the pathway csv data of the pathways.
8. **Connect to Printer:** Quick print the current view.
9. **Hover for Information:** When this function is enabled, users can move their mouse cursor over any vertex in the network, and a tooltip will appear displaying detailed information about that vertex.

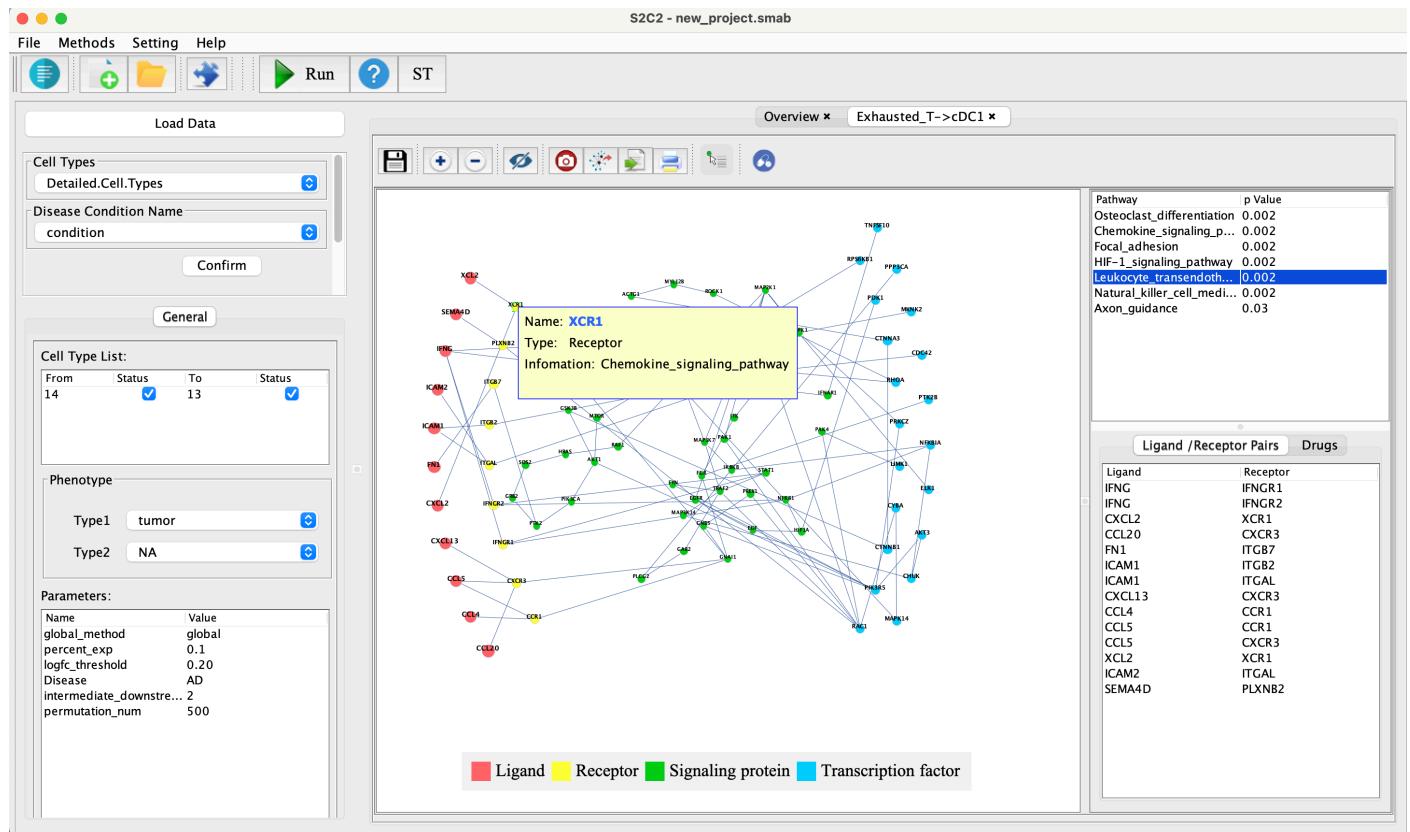


Figure 13. Hover on vertexes information

10. **Drug Bank:** In Figure 14, the function of the Drug Bank feature is illustrated. When users enable the Drug Bank function, they will notice that some of the vertices change color. This color change indicates that these vertices contain drug-related information. To explore this information, users can look at the lower right section of the interface, where drug details are listed. By double-clicking on these entries, users can access a webpage with more information about the selected drug.

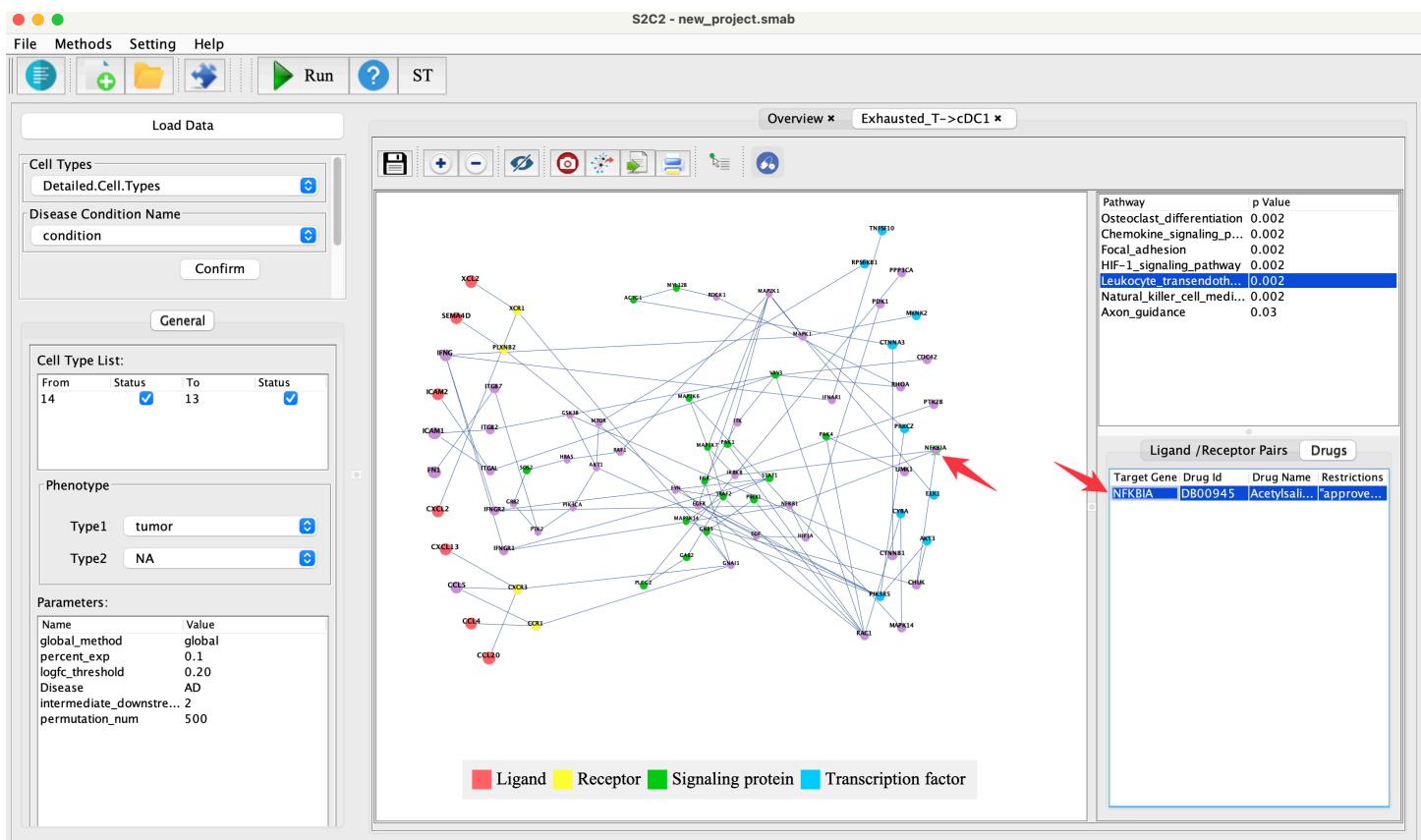


Figure 14. Drug Bank

2.6 Other Functions

1. Highlights vertexes of pathway and Sub-Pathway Window

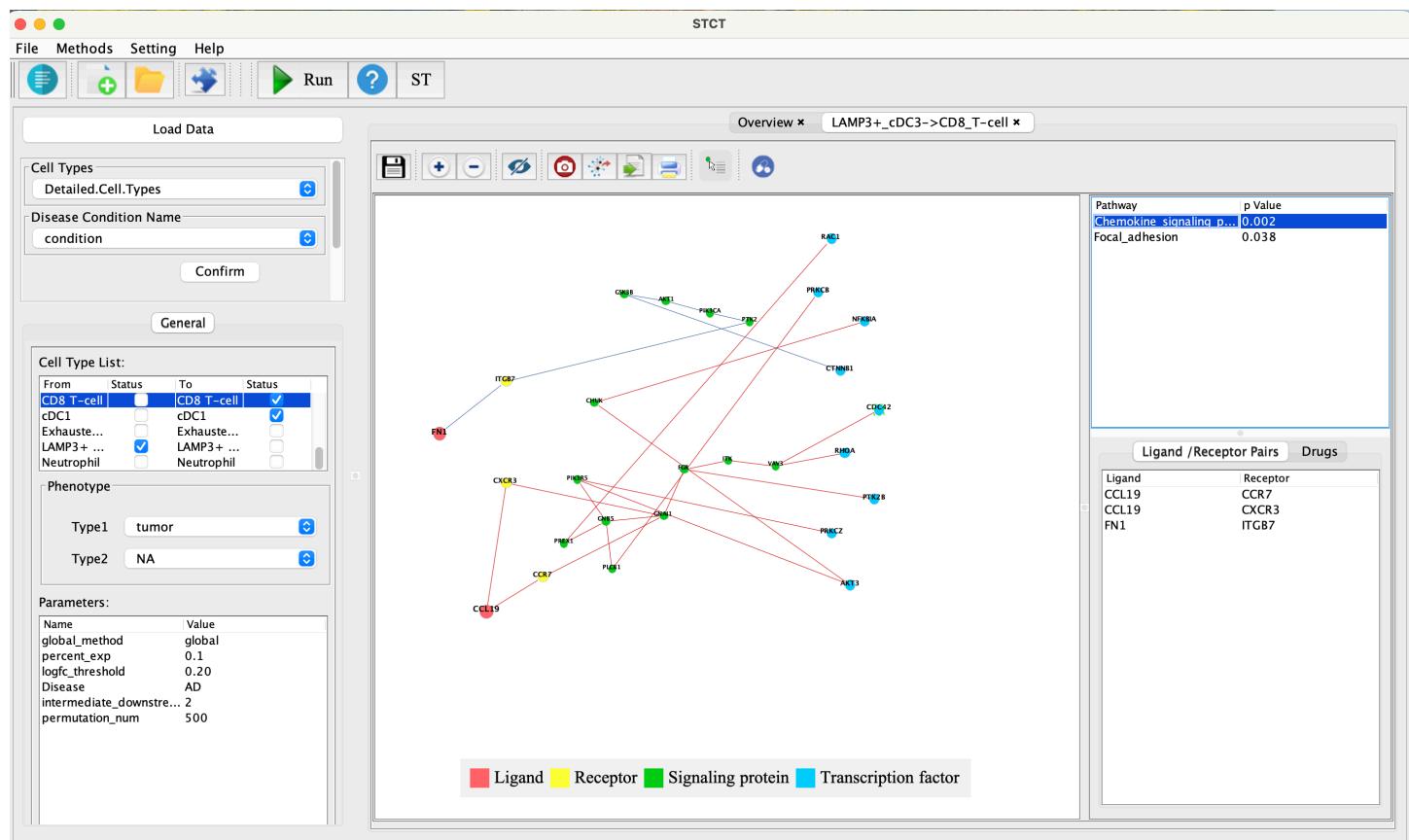


Figure 15. Highlight vertexes with selected pathway

Highlighting Pathway Vertices(Figure15)

Click on any pathway listed in the side panel to highlight all vertices involved in that pathway. It will emphasize on the network map makes it easier to identify and focus on the components of a specific pathway.

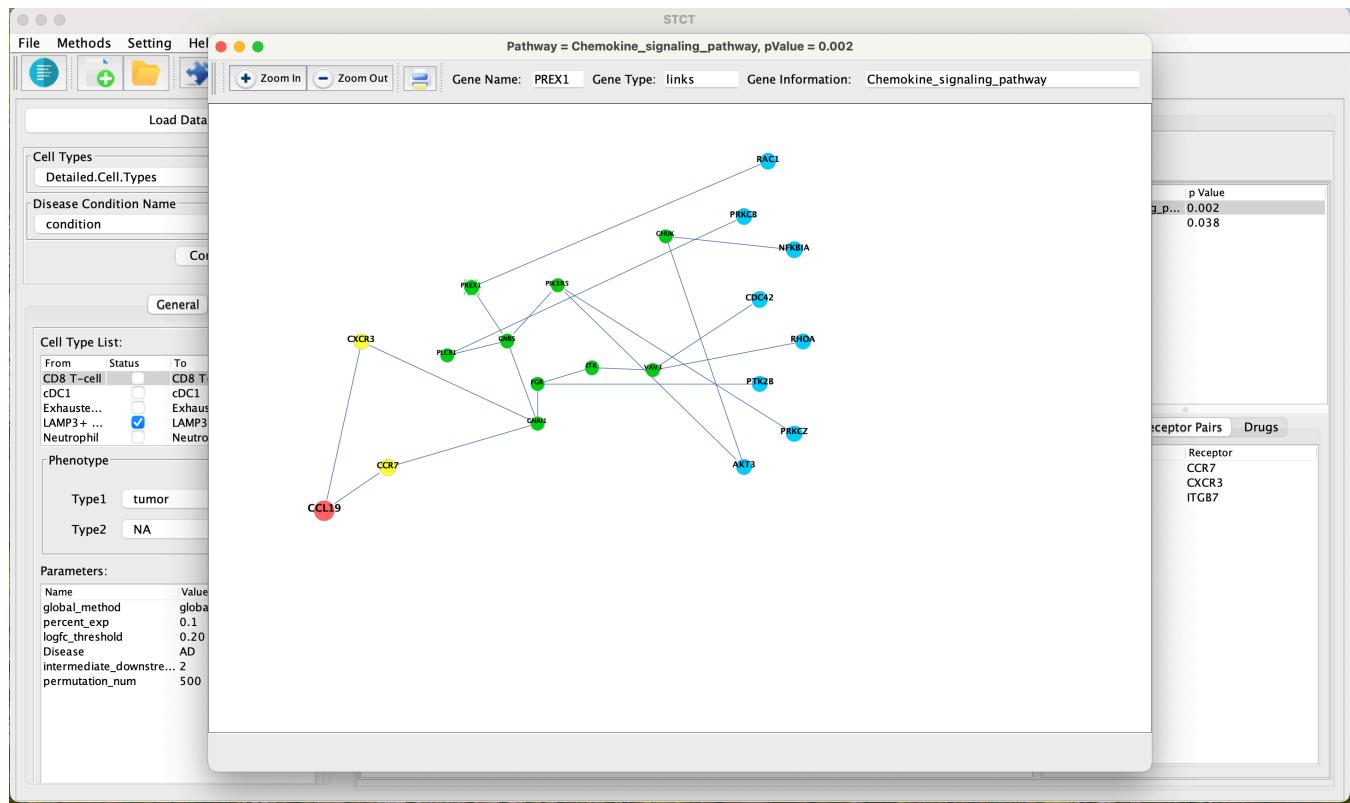


Figure 16. Window for the sub-pathway

Detailed Pathway Window View (Figure 16)

For a more detailed view of a selected pathway's vertices distribution, simply click on the pathway name. This action opens a new window, providing a focused view of just the selected pathway. Here, the layout is optimized to display the spatial distribution and connections of the vertices, offering clarity on the signalling dynamics and potential interactions within that pathway.

2. Customizing Network Layout

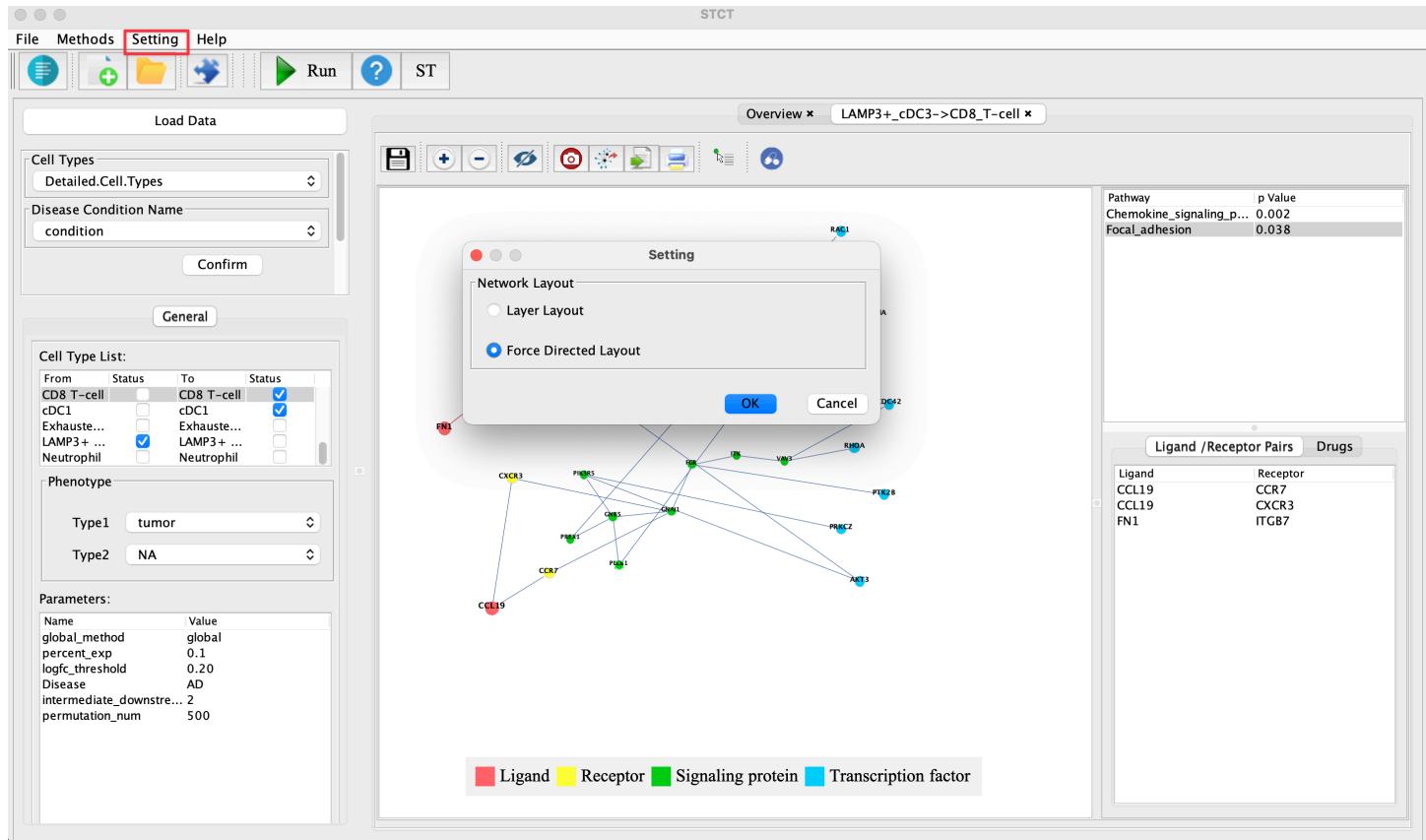


Figure 17. Network Layout Method

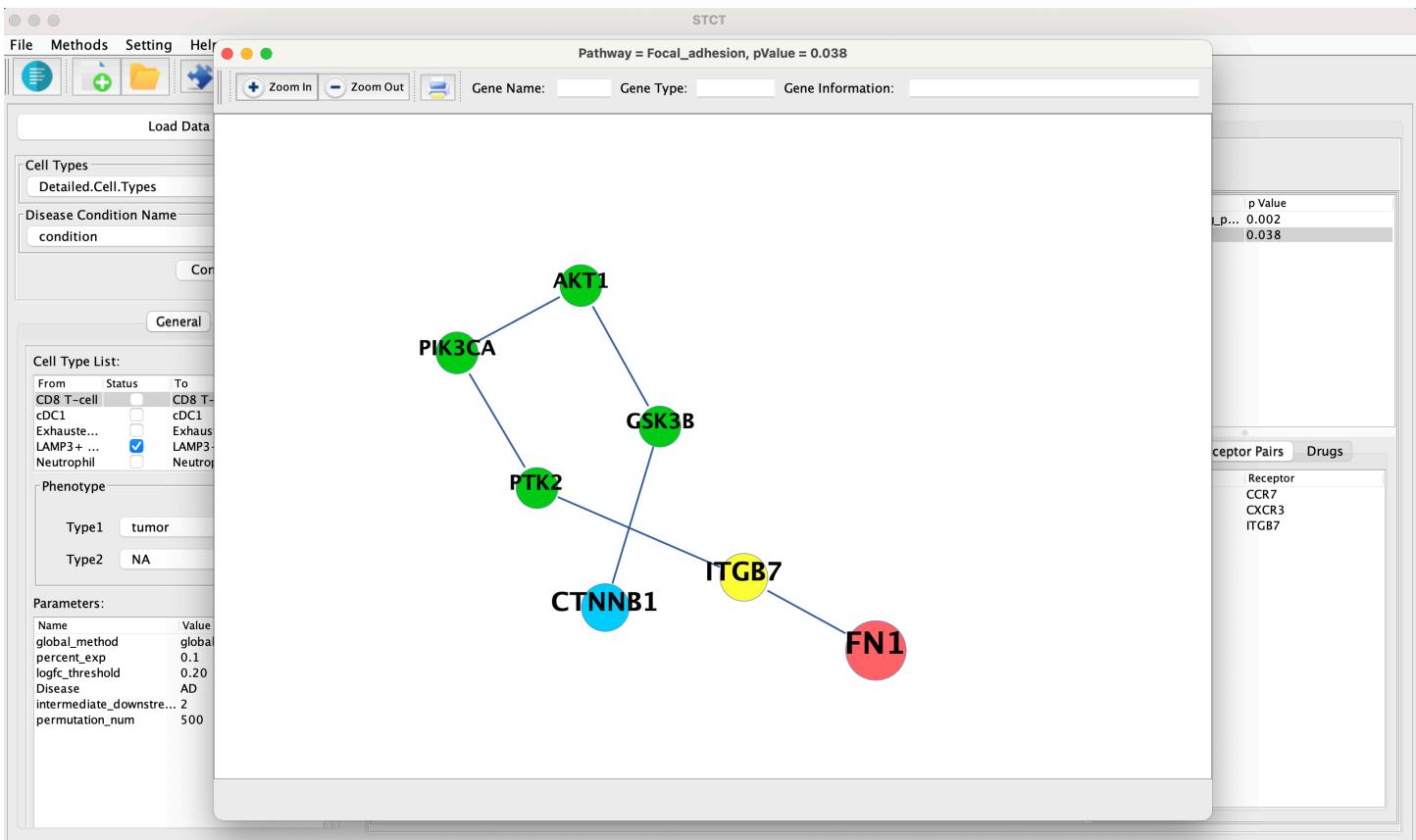


Figure 18. Network Layout under Force Directed Layout

To access Network Layout Settings:

1. Navigate to the *Setting* menu located in the main toolbar.
2. Select Network Layout from the dropdown options.

Layout Options: Layer Layout (organizes nodes into distinct layers) and Force Directed Layout(nodes are positioned based on a force-directed algorithm, simulates a physical system)