

How to use OptiCon_java.jar

一. Running environment

operation platform: win10

CPU: i7,eight cores

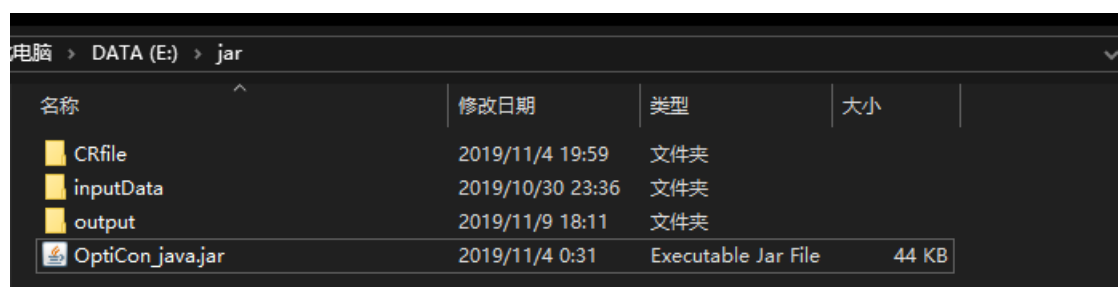
memory: 16G or more

storage: 10GB (It depends on the amount of data)

JDK version: 12.0.1

二. Create the required folder

1. Create an empty folder, such as a "jar" folder, with a variable name.
2. In the "jar" folder, create three folders: "CRfile", "inputData" and "output" respectively. **Note that the names of these three files must be consistent with the above, or the jar package will fail to run.**
3. Put the "OptiCon_java.jar" into the "jar" folder.
4. The results are as follows:



三. Prepare input data

1. The input data should put into "inputData" folder:
 - 1) "DScore.txt" contains data in two tab-delimited columns. Each row

includes a gene identifier (first column) and its corresponding p-value of differential expression (second column) under two conditions (e.g. diseased vs. healthy). The gene identifier is based on Entrez Gene ID with a prefix “En_”. Duplicated gene identifiers are not allowed.

2) “**GeneExpression.txt**” contains data in a matrix format. Rows represent genes and columns represent samples. Each entry in the matrix contains a gene expression value in a specific sample. Duplicated gene identifiers are not allowed.

3) “**RecurMutant_entrez.txt**” contains a list of genes (based on Entrez gene IDs) that are known to harbor recurrent somatic mutations in the cancer type of interest. Duplicated gene identifiers are not allowed.

4) If you want to identify synergistic key regulators using a **customized directed network**, Format your network data into a tab-delimited file “MyGeneNetwork.txt”. Each row represents a directed edge from the node in the first column to the node in the second column. **Attention:** In identifying synergistic OCNs pairs, we use “CancerCensus_En.txt” (Annotation for cancer (driver) genes were downloaded from the Cancer Gene Census database⁵⁸, which were confirmed to have recurrent somatic mutations in the specific cancer type using data from the Catalogue of Somatic Mutations in Cancer (COSMIC) database), which names genes used Entrez Gene ID with a prefix “En_”. If you want to use your customized directed network, you should be sure about the genes

named the same in all the files.

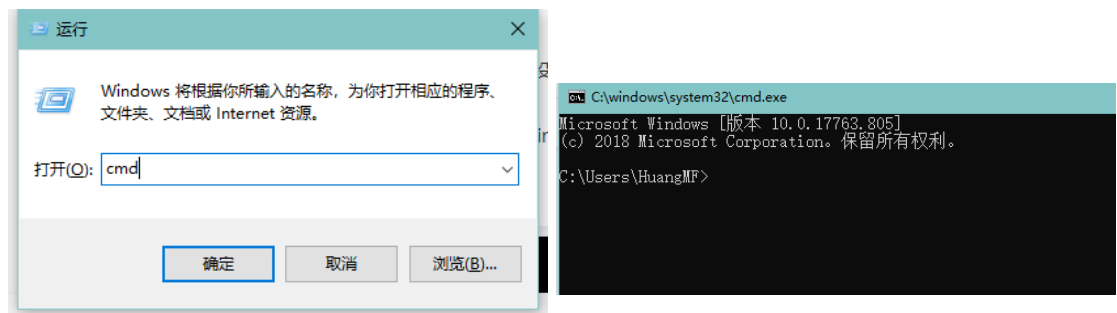
2. For example



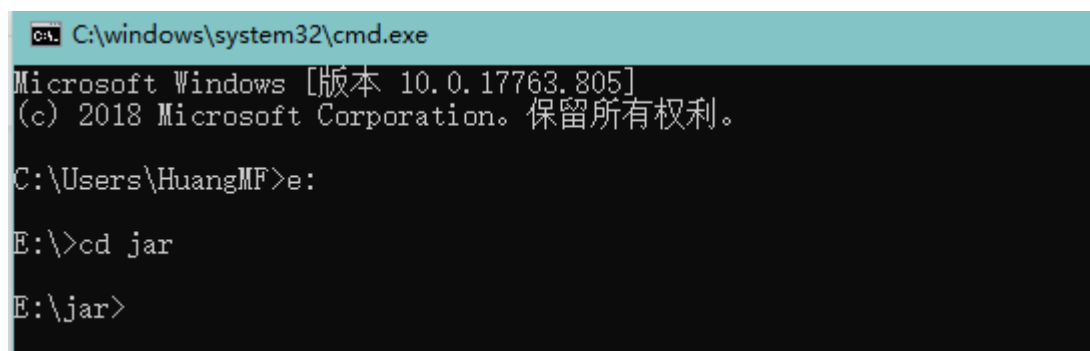
名称	修改日期	类型	大小
CancerCensus_En.txt	2019/8/23 21:15	文本文档	5 KB
DScore.txt	2019/8/1 18:40	文本文档	664 KB
GeneExpression.txt	2019/11/3 14:58	文本文档	3,157 KB
MyGeneNetwork.txt	2019/8/1 18:42	文本文档	1,742 KB
RecurMutant_entrez.txt	2019/8/26 23:59	文本文档	40 KB

四. Running OptiCon_java.jar

1. Open a command line window(win+R), input “cmd”,and click “OK”



2. Go to the directory where the “OptiCon_java.jar” package resides(The “jar” folder is created in “二. Create the required folder”).



3. Write the command to run the jar package and press “Enter” key

java -Xms8g -Xmx8g -jar OptiCon_java.jar [parameter 1]
[parameter 2] [parameter 3] [parameter 4] [parameter 5]
[parameter 6] [parameter 7] [parameter 8] [parameter 9]
[parameter 10]

```
Microsoft Windows [版本 10.0.17763.805]  
(c) 2018 Microsoft Corporation. 保留所有权利。  
C:\Users\HuangMF>e:  
E:\>cd jar  
E:\jar>java -Xms8g -Xmx8g -jar OptiCon_java.jar E:\jar MyGeneNetwork.txt DScore.txt GeneExpression.txt RecurMutant_entrez.txt \t 1000 0.18 0.3 true  
load MyGeneNetwork.txt...completed!  
load GeneExpression.txt...completed!  
load and compute DScore.txt...completed!  
use old control range file...completed! time:0s  
greedy search for OCNs...
```

java -Xms8g -Xmx8g -jar OptiCon_java.jar E:\jar MyGeneNetwork.txt DScore.txt GeneExpression.txt RecurMutant_entrez.txt \t 1000 0.18 0.3 true

- 1) Parameter 1: The directory of three folders("CRfile", "inputData" and "output") , for example: E:\jar.
- 2) parameter 2: The name of gene regulatory network file, for example: MyGeneNetwork.txt
- 3) parameter 3: The name of gene deregulation file, for example: DScore.txt
- 4) parameter 4: The name of gene expression file, for example: GeneExpression.txt
- 5) parameter 5: The name of recurrent somatic mutations file, for example: RecurMutant_entrez.txt
- 6) parameter 6: The separator of all input file, for example: \t
- 7) parameter 7: The number of SCC, for example: 1000
- 8) parameter 8: The value of epsilon used in computing ICV, for example: 0.18
- 9) parameter 9: The value of lambda used in computing ICV, for

example: 0.3

10) parameter 10: The boolean value of true or false, which means using the control range(in “CRfile” folder) that you calculated before or not.

4. View the result in “output” folder after running over

