

# Instructions for Using CIMminer with High-Throughput GoMiner Output Files

Barry Zeeberg

barry@discover.nci.nih.gov

September 30, 2004

## 1. Overview

- The [CIMminer](#) user interface is in the process of reorganization. In the meantime, we are providing specific instructions to users of [High-Throughput GoMiner](#). The purpose of these instructions is to facilitate construction of clustered image maps (CIMs) that are similar to those shown in figures in the High-Throughput GoMiner manuscript, as well as other CIMs that can be constructed from High-Throughput GoMiner output files.

## 2. General Observations

- Internet Explorer is the preferred browser for CIMminer.
- Certain platforms and browsers can cause a problem during the upload of the input file unless the input file name ends in “.txt”. This problem has been noted in particular on Mac OS X (10.2.8) using Internet Explorer 5.2.3.
- Another potential problem occurs if all of the rows of the input file do not have exactly the same number of field delimiters (*i.e.*, tabs). We have taken pains to avoid introducing this problem in the output files of High-Throughput GoMiner.
- The error messages that result from these two conditions are not informative, and can, in fact, be misleading.
- Another error message that is not informative can result from an attempt to use the correlation metric if there is a row or column all of whose entries are identical.
- These and other usability issues are currently being addressed.

## 3. Specific Instructions for Figure 3 of the High-Throughput GoMiner manuscript

- We refer to CIMs similar to Figure 3 as “trinary” because they contain exactly three values. In particular, the values are -1, 0, and 1.
- This is the type of file that is described on the [High-Throughput GoMiner web site](#) under the “Output Files” tab as “*CHANGEDGENESFILE.change.gce.updown.CIM*”.
- Steps - steps (b) through (d) should be performed simultaneously
  - Clustering

|         | X axis                           | Y axis                           | Row Cluster Distance                     | Row Cluster Algorithm                        | Column Cluster Distance                  | Column Cluster Algorithm                     |
|---------|----------------------------------|----------------------------------|------------------------------------------|----------------------------------------------|------------------------------------------|----------------------------------------------|
| Cluster | <input checked="" type="radio"/> | <input checked="" type="radio"/> | <input type="text" value="correlation"/> | <input type="text" value="Average Linkage"/> | <input type="text" value="correlation"/> | <input type="text" value="Average Linkage"/> |
| Random  | <input type="radio"/>            | <input type="radio"/>            |                                          |                                              |                                          |                                              |
| Nothing | <input type="radio"/>            | <input type="radio"/>            |                                          |                                              |                                          |                                              |

b. Colors

| Minimum value                          | Left mid-point              | Right mid-point             | Maximum value               |
|----------------------------------------|-----------------------------|-----------------------------|-----------------------------|
| <input checked="" type="radio"/> 0,1,0 | <input type="radio"/> 1,1,0 | <input type="radio"/> 1,1,0 | <input type="radio"/> 1,0,0 |

c. Page Layout

|                                  |          |                                       |           |
|----------------------------------|----------|---------------------------------------|-----------|
| <a href="#">Rotation</a>         | 0 degree | <a href="#">Transposition of data</a> | transpose |
| <a href="#">Height of CIM*</a>   | 550      | <a href="#">Width of CIM*</a>         | 450       |
| <a href="#">Width of legend*</a> | 144      | <a href="#">Height of legend*</a>     | 4.5       |

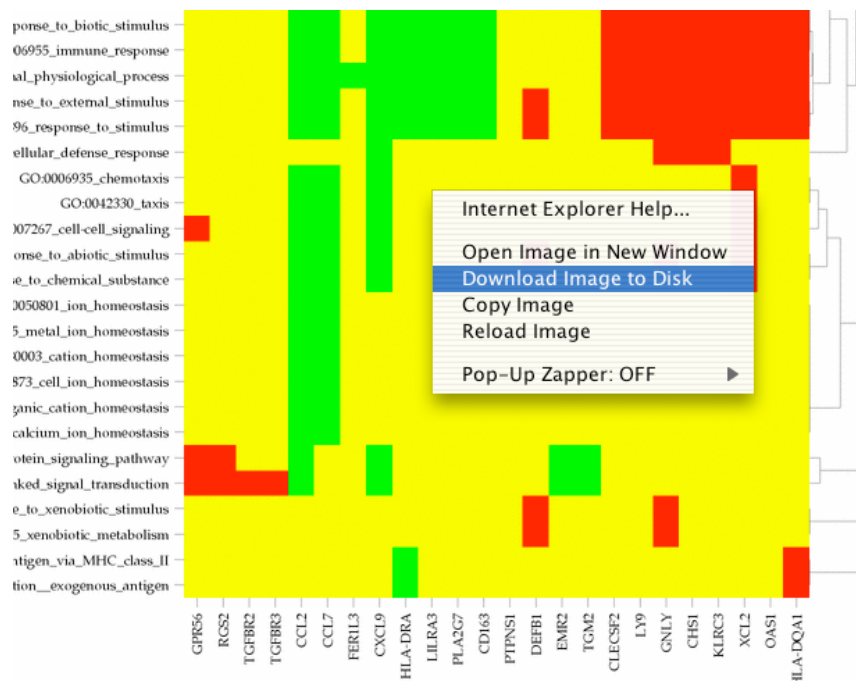
d. Axes

### Axes

| TOP                                                                              | BOTTOM                                                                                   |
|----------------------------------------------------------------------------------|------------------------------------------------------------------------------------------|
| Draw: <input type="text" value="Cluster tree"/> <input type="text" value="N/A"/> | Draw: <input type="text" value="Label"/> <input type="text" value="Facing to the left"/> |
| LEFT                                                                             | RIGHT                                                                                    |
| Draw: <input type="text" value="Label"/> <input type="text" value="Facing up"/>  | Draw: <input type="text" value="Cluster tree"/> <input type="text" value="N/A"/>         |
| Font size for all labels                                                         |                                                                                          |
| <input type="text" value="10"/>                                                  |                                                                                          |

e. Save CIM as .gif

Right-click to reveal menu and select “Download Image to Disk”



4. Specific Instructions for Figure 4 of the High-Throughput GoMiner manuscript
  - We refer to CIMs similar to Figure 4 as “continuous” because the entries can assume any floating point values. In this particular instance, the values are all non-negative.
  - This is the type of file that is described on the [High-Throughput GoMiner web site](#) under the “Output Files” tab as “total.txt.change.series.CIM”.
  - Steps - steps (b) through (e) should be performed simultaneously
    - a. Clustering

|         |                                  |                                  |                                          |                                              |                                          |                                              |
|---------|----------------------------------|----------------------------------|------------------------------------------|----------------------------------------------|------------------------------------------|----------------------------------------------|
|         | X axis                           | Y axis                           | Row Cluster Distance                     | Row Cluster Algorithm                        | Column Cluster Distance                  | Column Cluster Algorithm                     |
| Cluster | <input checked="" type="radio"/> | <input type="radio"/>            | <input type="text" value="correlation"/> | <input type="text" value="Average Linkage"/> | <input type="text" value="correlation"/> | <input type="text" value="Average Linkage"/> |
| Random  | <input type="radio"/>            | <input type="radio"/>            |                                          |                                              |                                          |                                              |
| Nothing | <input type="radio"/>            | <input checked="" type="radio"/> |                                          |                                              |                                          |                                              |

b. Colors

| Minimum value                          | Left mid-point              | Right mid-point             | Maximum value               |
|----------------------------------------|-----------------------------|-----------------------------|-----------------------------|
| <input checked="" type="radio"/> 1,1,0 | <input type="radio"/> 1,5,0 | <input type="radio"/> 1,5,0 | <input type="radio"/> 1,0,0 |

c. Binning

|                                |                                                  |
|--------------------------------|--------------------------------------------------|
| <a href="#">Binning method</a> | <input type="text" value="Asymmetric interval"/> |
|--------------------------------|--------------------------------------------------|

d. Page Layout

|                                  |                                       |                                       |                                        |
|----------------------------------|---------------------------------------|---------------------------------------|----------------------------------------|
| <a href="#">Rotation</a>         | <input type="text" value="0 degree"/> | <a href="#">Transposition of data</a> | <input type="text" value="transpose"/> |
| <a href="#">Height of CIM*</a>   | <input type="text" value="550"/>      | <a href="#">Width of CIM*</a>         | <input type="text" value="450"/>       |
| <a href="#">Width of legend*</a> | <input type="text" value="144"/>      | <a href="#">Height of legend*</a>     | <input type="text" value="4.5"/>       |

e. Axes

|                                 |                                    |                                                 |                                           |
|---------------------------------|------------------------------------|-------------------------------------------------|-------------------------------------------|
| <a href="#">TOP</a>             |                                    | <a href="#">BOTTOM</a>                          |                                           |
| <a href="#">Draw:</a>           | <input type="text" value="Label"/> | <input type="text" value="Facing to the left"/> | <input type="text" value="Nothing"/>      |
|                                 |                                    |                                                 | <input type="text" value="N/A"/>          |
| <a href="#">LEFT</a>            |                                    | <a href="#">RIGHT</a>                           |                                           |
| <a href="#">Draw:</a>           | <input type="text" value="Label"/> | <input type="text" value="Facing up"/>          | <input type="text" value="Cluster tree"/> |
|                                 |                                    |                                                 | <input type="text" value="N/A"/>          |
| <b>Font size for all labels</b> |                                    | <input type="text" value="10"/>                 |                                           |

f. Save CIM as .gif

Right-click to reveal menu and select “Download Image to Disk”



## Pitfalls

- If you get the error message: "There was an error processing your files in R, The cluster distance you selected is not defined for your dataset. For example, you selected "binary" distance, your dataset is not discrete (character) data," be sure your dataset is appropriate for the distance metric you selected. Examine your data and remove the offending row or column, and try to generate the CIM again.
- If you get the error message: "Your datafile has an inconsistent number of tab separated values per line," a possible culprit could be extra tabs at the end of your data files. You should not get this error when processing data directly from High-Throughput GoMiner, but it is possible to accidentally introduce this error if you manually edit the cluster data file.