

# Cropper 0.1 Documentation

Website: [www.geenivaramu.ee/en/tools/cropper](http://www.geenivaramu.ee/en/tools/cropper)

Contact: [toomas.haller@ut.ee](mailto:toomas.haller@ut.ee), [toomashaller@gmail.com](mailto:toomashaller@gmail.com)

Reference: *manuscript in preparation*

## **CONTENTS**

I. INTRODUCTION – p. 2

II. HOW TO COMPILE THE PROGRAM – p. 2

III. USER GUIDE – p. 3

3.1. Input file requirements – p. 3

3.2. Cropper instructions – p. 3

Fig. 1. The Global View – p. 4

Fig. 2. The Local View – p. 4

3.3. Error handling – p. 5

IV. EXAMPLE – p. 5

## I. INTRODUCTION

Cropper is a graphical user interface-based Manhattan Plot viewer written in C++/Qt. The user can select, zoom, crop (include and exclude) and output the desired parts of the Manhattan plot both graphically and numerically. It was initially created because no other tool enabled sequential fragmentation/cropping of Manhattan plots. Cropper allows the user to extract peaks from the plot, store them in separate files and then return to what is left over for more cropping as needed.

## II. HOW TO COMPILE THE PROGRAM

You can download the program and instructions from the website (see above). The statically compiled versions are ready to run, dynamically compiled versions require that Qt libraries are installed on your system (Qt 4.3 or newer and gcc 5.2 or newer). Qt libraries are free and can be downloaded and installed from <http://www.qt.io/download/>. The GCC compiler can be downloaded from <https://gcc.gnu.org/>.

In order to compile Cropper from the source code, place the source code in the folder named “CROPPER” on the Qt path and execute the following lines (please read all notes below):

```
> qmake -project  
> qmake  
> make
```

**Note1:** Cropper requires C++11. After you have made the Make file (typed “qmake”), please add `'-std=c++11'` after the flag `CXXFLAGS`. You may also need to add `'-lz'` after `LIBS` on some systems.

**Note2:** If you are compiling for Windows, open your \*.pro file after typing “qmake -project” and add this line to the \*.pro file:

**CONFIG+=console**

**INCLUDEPATH += C:\Qt\4.4.3\src\3rdparty\zlib** (where C:\Qt\4.4.3\ should be replaced with the correct path)

Your \*.exe file will appear in the “release” subfolder.

**Note3:** If you are compiling for Mac, you may need to have Xcode installed (<https://developer.apple.com/xcode/>) if your computer doesn't already have it. Before you install, look under “/Applications” to see if “/Applications/XCode” exists.

If you installed Qt on a Mac using binaries, the program should be located here:

“~/QtSDK/Desktop/<qt.version>/<compiler>/bin”, if you built it from source code, it should be here:

“/usr/local/Trolltech/<qt.version>/bin”.

**Note4:** If you are compiling for Mac you need to execute the following lines in order to compile using g++:

**qmake -project**

**qmake -spec macx-g++**

**make**

After you have created your Makefile (before typing “make”) you may need to add `“-lz”` after `“LFLAGS=”` in your Makefile.

**Note5:** If you are compiling statically, open your \*.pro file after typing “qmake -project” and add this line to the \*.pro file:

**CONFIG+=static** or **CONFIG+=staticlib**

Static compilation is an option only if Qt itself was built statically.

In your Makefile please add `'-static-libstdc++'` and `'-static-libgcc'` behind `LIBS`.

**Note6:** If you are using Qt5 to compile, you need to add `“QT += widgets”` to your \*.pro file (after typing “qmake -project”)

If you have any problems compiling or running Cropper, please do not hesitate to contact the author at [toomashaller@gmail.com](mailto:toomashaller@gmail.com) for help.

### III. USER GUIDE

#### 3.1. Input file requirements

Cropper does input file structure detection automatically. Still the files need to meet certain criteria:

1. Delimiter can be tab, space or comma and it is detected automatically. The file extension can be “.txt” or “.csv”.
2. If the file has no header it needs to contain these three columns as the only or first three columns: chromosome number (can be a letter), position, p-value.
3. If the file has a header it can contain the above-mentioned three columns anywhere in the file. In that case the correct columns are found based on the header text. The header text needs to contain this text:  
“**chr**” or “**id**” for the chromosome number/name anywhere in the header (case insensitive)  
valid examples: *CHR, Chromosome, myCHR, Idname*  
“**pos**” for the physical position anywhere in the header (case insensitive)  
valid examples: *positions, postable, physicalPos*  
“**pval**” or “**p-val**” for the p-value anywhere in the header (case insensitive) or it has to be exactly and precisely “**P**”  
valid examples: *Pvalues, p-Val\_field, P, pvals*

Note that the output files that Cropper generates are also automatically valid Cropper input files.

#### 3.2. Cropper instructions

Cropper starts by asking for the name and location of the input file. It then opens the file in the Global View (GV) meaning that it shows all chromosomes together in one graphical view (Manhattan Plot). The GV also shows chromosome ID's, marks the best p-value and shows the p-value levels generally of interest to the researchers (Fig.1). At this point the user can save the image as a standard PNG file (<Save> button), exit the tool (<Exit> button), change the input file (<New File> button), or enter the Local View (LV) by selecting a chromosome of interest. Chromosome selection is performed by drawing a box around the peak. The box does not have to be precise as the chromosome that is closest to the center of the box is selected in full. The LV presents only the chromosome that was selected (Fig.2). Data for the LV are retrieved afresh from the input file. The LV shows the starting position, the ending position, the difference between the two, and the best p-value. Several new fields and buttons appear in the local mode:

1. Two tick boxes. These tick boxes turn on and off the vertical p-value reference lines and connect/disconnect the individual points with lines.
2. Two text fields. The user can enter position values in these boxes (starting and ending position). You can also put both numbers in the first field, separated by a dash (example: 200000-19000000). This option is for convenience when a numerical range is copied into Cropper from another source using the dash. The selected region will be colored red for reference. The coloring takes effect when any button or tick box is pressed. The best button to press is <Unzoom> to see the colors change.

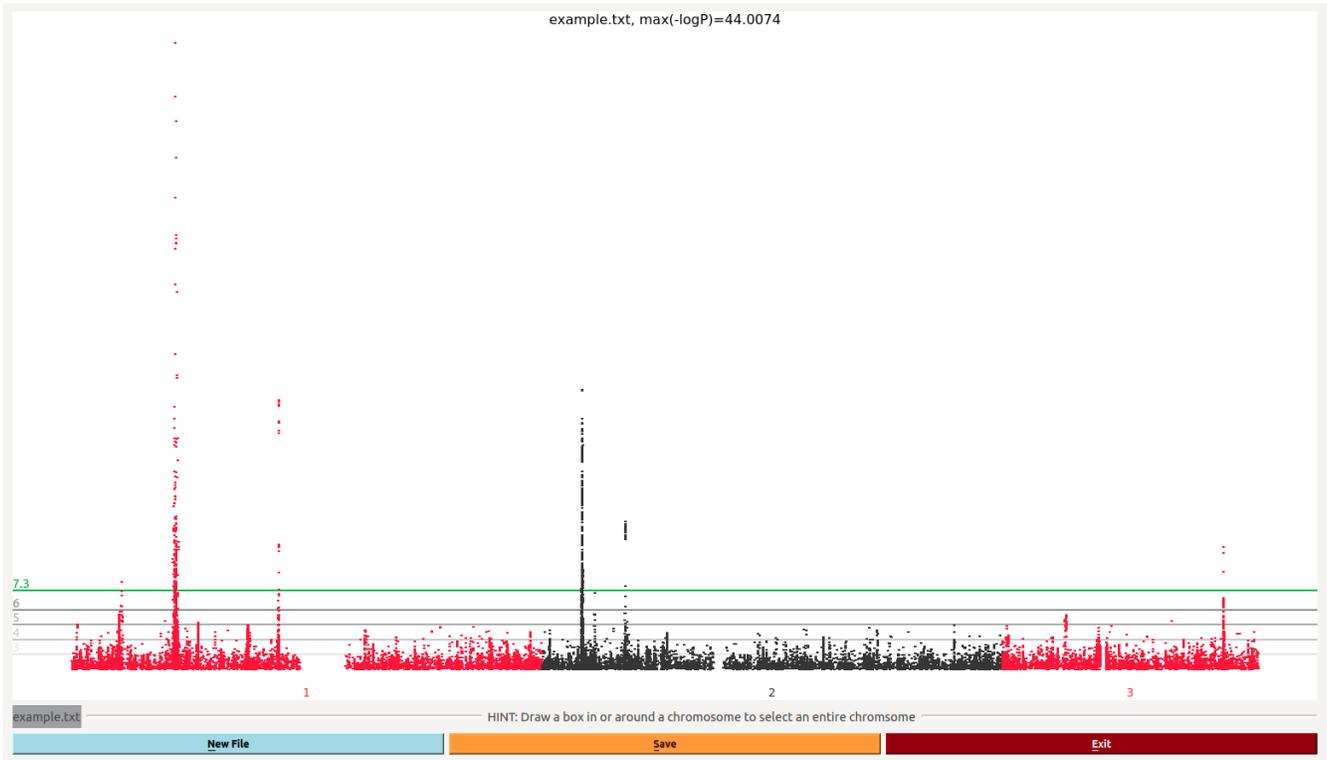


Fig. 1. The Global View.

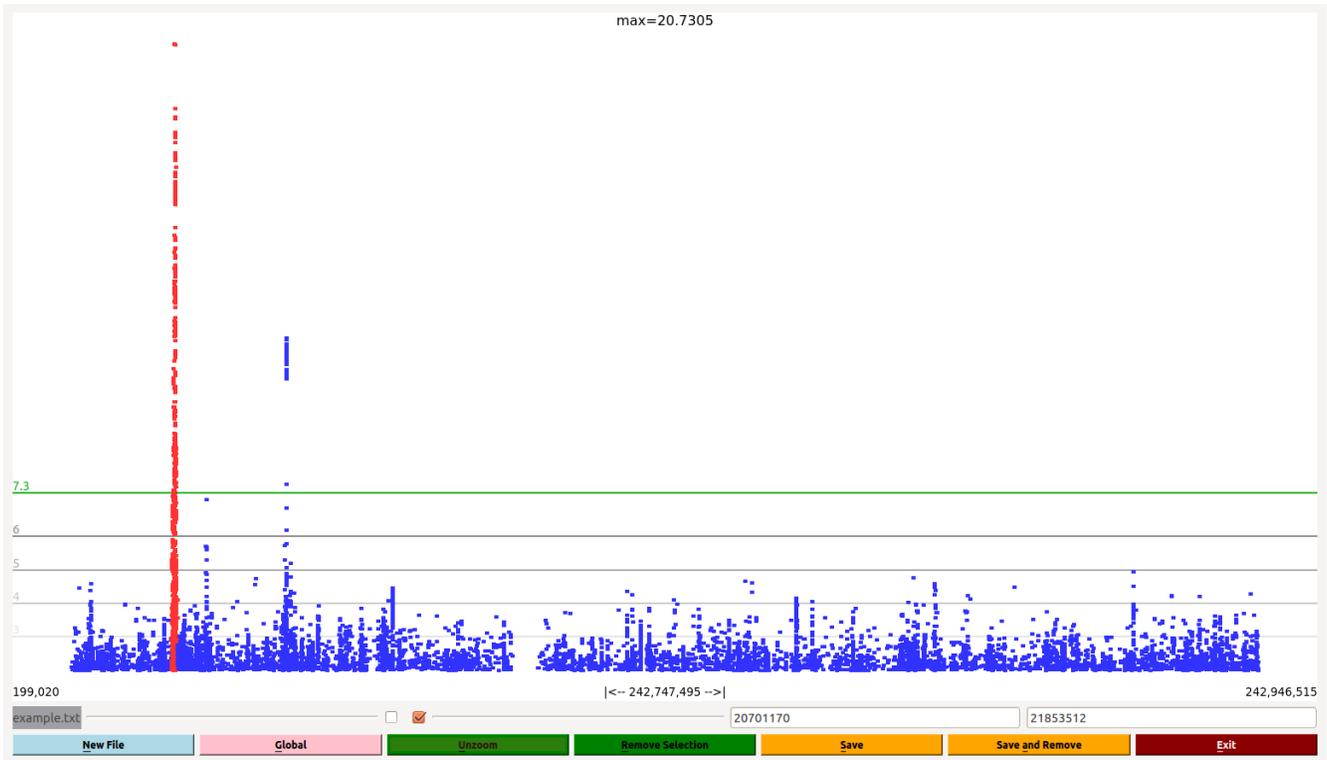


Fig. 2. The Local View.

3. <Global> button. Pressing this button will take you back to the GV. The image is retrieved from memory, file is not read to re-generate this view. However, when you go back to the LV, data are read from the file.
4. <Unzoom> button. Pressing this button takes the user back to the original LV. If some datapoints have been removed (deleted) by the user the <Unzoom> button will not bring them back.
5. <Remove selection> button. In the LV the mouse assumes its original functionality (becomes precise; unlike in the GV). The mouse can now be used to select parts of the plot by drawing boxes around data points. The Cropper will automatically zoom into the selected region. If the user wants to remove the selected region from the plot then the <Remove selection> button needs to be pressed. The removed parts are permanently lost, they can be recovered by re-reading the input file (press the <Global> button and select the chromosome).
6. <Save> button. This button saves the current view of the plot in both the graphical (.png) and text (.txt) format. It asks for a new file name. It is best to provide the file name without file extension as Cropper generates this automatically. Cropper also offers the previously saved file name as the new default name. This is for user convenience since often serial numbers are given to files and the user does not have to remember where the count is.
7. <Save and remove> button. This saves the selected region (as it appears on the screen) in a new file (both graphically and numerically) and removes the corresponding region from the original plot/data. This is for user convenience as it allows to keep track of the the regions that have already been extracted and do not have to be revisited. This allows the user to conveniently divide a plot into subplots of interesting features.

### 3.3 Error handling

Should something go wrong and Cropper terminates it puts the error message in the Cropper\_log.txt file. Please check this file to see what went wrong. Please provide this message if you contact the author (toomashaller@gmail.com) for trouble-shooting.

## IV. EXAMPLE

This is a step-by step tutorial that introduces the functions of cropper.

1. Start Cropper and load example.txt (downloaded together with Cropper)
2. Observe the Global View (GV) showing three chromosomes.
3. Press <Save> to output the Manhattan Plot as a PNG file. Note that the file name was generated automatically based on the input file name.
4. Draw a box in or around the chromosome that you want to select. Observe the Local View (LV) displaying your selected chromosome in full. Note the start position, the end position and the difference between the two. The  $-\log_{10}(P)$  at the top of the graph corresponds to the highest peak.
5. You can always go back to the GV by clicking <Global>. This will refresh your plot (it will bring back the regions that you may have cropped away).

6. Press <Save> and enter a file name without extension (example: myfile). Note that a png file is created with the plot and the data are also output in a txt file. The file is always created only for the regions shown on the screen.
7. Select an inter-peak region by drawing a box around it, it will be blown up on your screen for a close-up view. Now press <Remove selection> to remove these data from the plot.
8. Select a peak by drawing a box around it. Press <Unzoom> to see how you could go back. Select the same peak again and press <Save and Remove>; give the file a suitable name. Note that the selected peak was output to a file (png and txt files) and also removed from the data.
9. Press <New File> to load a new file.
10. Press <Exit> to exit Cropper.
11. Check the Cropper\_log.txt is file if something did not work.

***Thank you for reading the Cropper user guide!  
Looking forward to your feedback.***