

---

# PopulationProfiler: User manual

Damian J. Matuszewski, Carolina Wählby, Jordi Carreras Puigvert, and Ida-Maria Sintorn

Contact: [damian.matuszewski@it.uu.se](mailto:damian.matuszewski@it.uu.se)

Last edited: 13/11/2015

---

## 1 USER MANUAL

This manual is for PopulationProfiler ver. 1.2.

### 1.1 Graphical User Interface

PopulationProfiler has a simple graphical interface (see Fig. 1) that allows for selection of the input files, adjusting histogram parameters and choosing the output directory. It has a dedicated module for automatic cell cycle analysis based on DNA content but also allows manual gating selection (see Fig. 2) which finds wide range of applications. The best practice is to fill the fields from top to bottom following the tips visible directly in the window or appearing after placing the mouse cursor on a given field for a couple of seconds. The interface was designed using freely available PyQt<sup>1</sup> library that is platform independent.

### 1.2 Input description

PopulationProfiler takes as input comma separated value (CSV) files, a commonly used format for storing datasets. The files must fulfill the following criteria:

- The first row of the file must contain names of the columns in the dataset,
- All strings in the file that contain white spaces must be quoted,
- The column with well names must have “CXX” format, where “C” is a capital character indicating the row and “XX” is a two-digit number (padded with zero if less than 10) indicating the column in the screening plate,
- More than one file can be loaded to the software, in which case each file is treated as a separate screening plate and must contain the same columns as the ones selected for the last added file (treatment label, well names and analyzed feature).

If the manual gating analysis is selected the pooled negative control histogram is displayed to aid in the selection of gates (see Fig. 2). The same thresholds are later applied to all analyzed samples.

### 1.3 Output description

The primary output from the PopulationProfiler is the CSV file with the histogram data calculated for each well. This data is also visualized and stored in PDF and PNG formats (see Fig. 3).

In case of the DNA content-based cell cycle analysis an additional output CSV file containing the counts and percentages of cells in each of the five cell cycle subpopulations is generated. Moreover, the total number of cells in each well and the DNA

content values corresponding to 2N and 4N peaks are stored. This additional data is visualized in several ways:

- Vertical black lines dividing the histogram into the cell cycle subgroups and percentage of cells in each of them are added to the graphs (see Fig. 3),
- An additional stack-bar chart (see Fig. 4) is generated for each input file; it presents normalized percentage contribution of each of the cell cycle subpopulations to the total cell count in the wells,
- An additional scatter plot (one per input file, not shown) presents the dependency between cell counts in the 2N and 4N subgroups. This can be particularly useful in the initial search for potentially interesting drug-dose combinations.

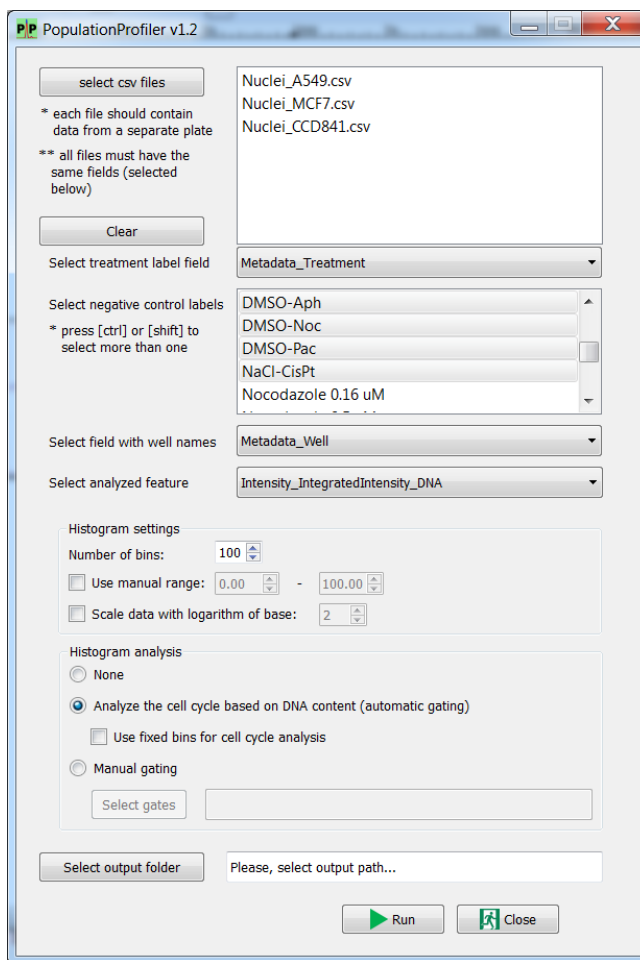
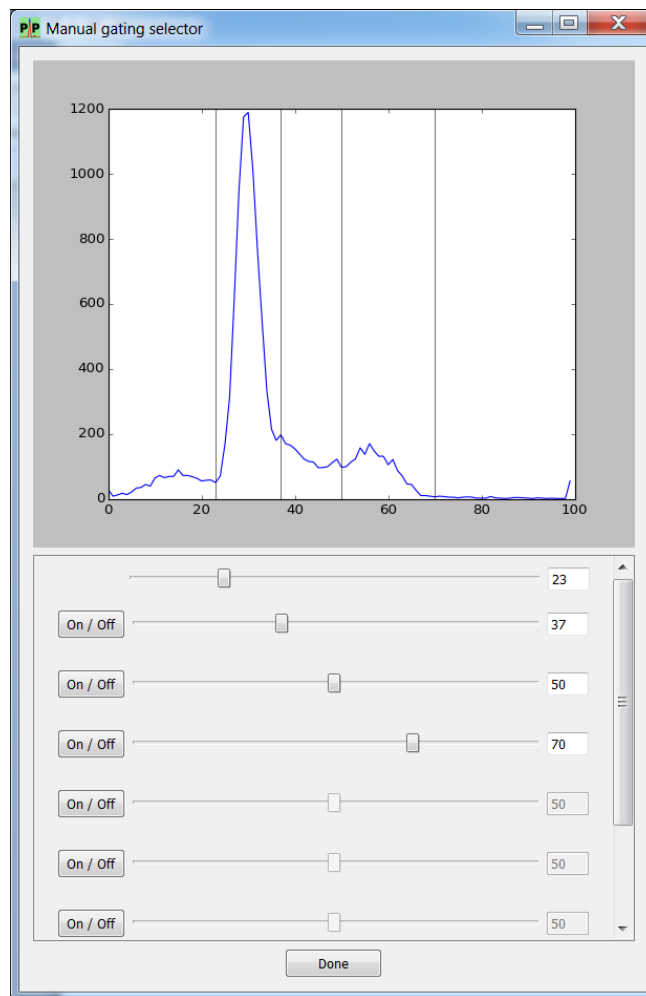


Figure 1. Graphical user interface of the PopulationProfiler.

---

<sup>1</sup> <http://www.riverbankcomputing.com/software/pyqt/intro>

In case of the manual gating analysis the output is similar to that from the cell cycle analysis, except that the number of gates depends on the user selection and there is no scatter plot generated (only the histograms and stacked bar plots).

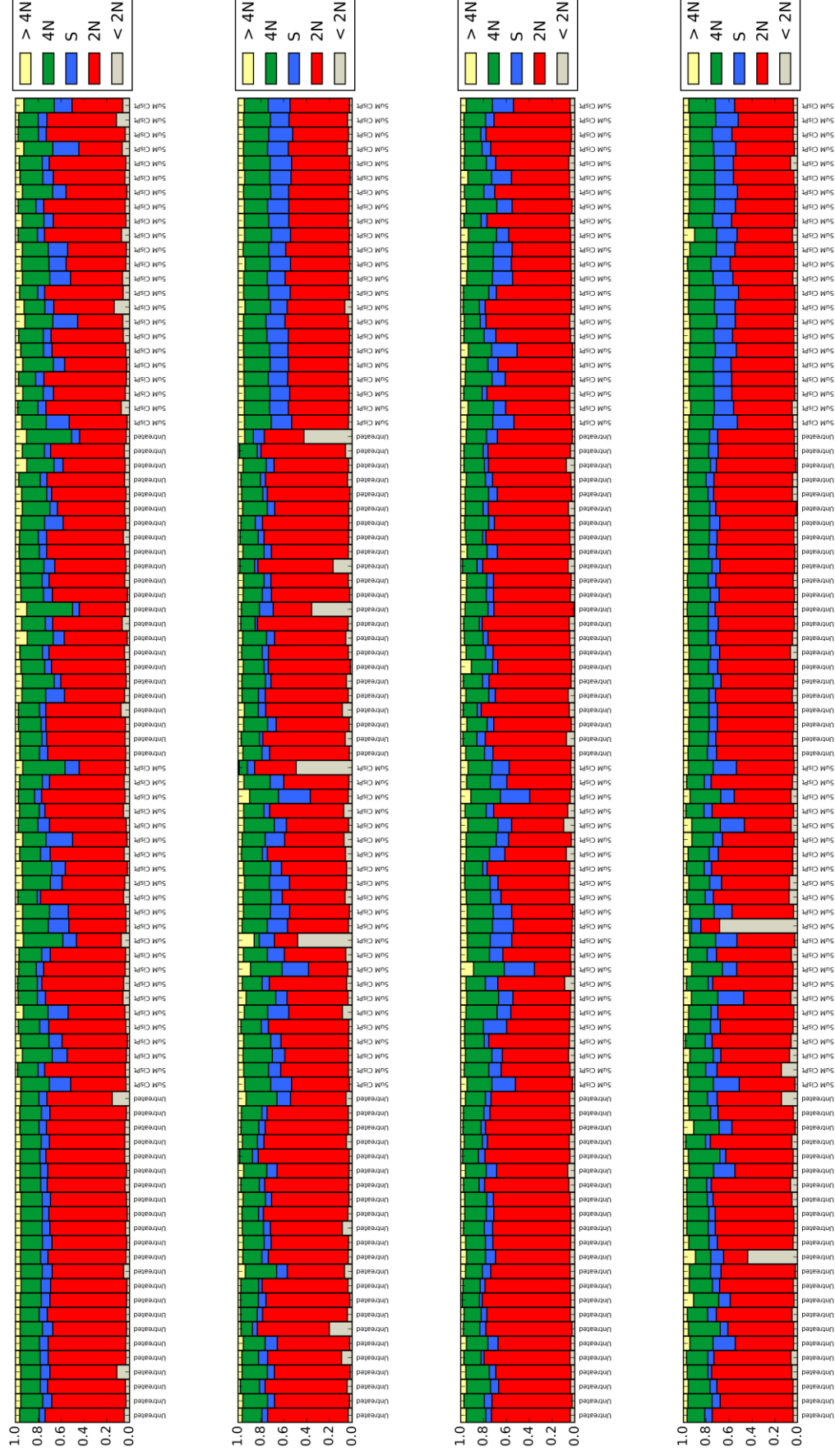


**Figure 2.** Graphical user interface of the manual gating selection module in PopulationProfiler.



Figure 3. Visualization of a 384-well screening plate – DNA content-based cell cycle histograms.

Cell cycle subpopulations distribution - Nuclei



**Figure 4.** Visualization of a 384-well screening plate – stack bar plots of cell cycle subpopulations based on DNA content.