

## High-quality Genome (Re)-Assembly from Chromosomal Contact data Marie-Nelly *et al.*

### Requirements (Hardware):

-----

- NVIDIA graphic card (computing capability  $\geq 2.0$ , RAM  $\geq 1.5\text{Go}$ )
- Linux (Recommended Ubuntu  $\geq 13.10$ ), OSX, Windows

### Requirements (Software):

-----

- hdf5
- hf5py
- CUDA (Cuda Toolkit  $\geq 5.5$ )
- pycuda (with opengl support)
- OpenGL
- wx
- wxpython
- wxpython-version
- numpy
- matplotlib
- scipy
- PyOpenGL

**To first install pycuda with opengl support follow the instructions at**

<http://wiki.tiker.net/PyCuda/Installation>

After running "python configure.py" change the file siteconf.py as follows:

Replace:

```
CUDA_ENABLE_GL = False
```

by:

```
CUDA_ENABLE_GL = True
```

## GRAAL INSTALLATION

### Installation:

-----

- Copy the shared folder graal to your\_directory
- Open a terminal
- cd your\_directory/graal/python\_graal
  
- run the following command line:  
    python main\_window.py
  
- follow the instructions in the GUI

## RUNNING GRAAL

### Description:

-----

A pyramid of contact matrices,  $P = \{M_0, M_1, \dots, M_k\}$ , is a data structure representing the 3C/HiC data at different binning.

The level 0 corresponds to the restriction fragment level contact matrix  $M_0$ . If  $x$  is the subsampling/scaling factor, we build  $M_i$  by creating bins of  $x^i$  colinear restriction fragments.

$G_0$  is the initial genome used to align the reads.

**The following steps refer to numbers and letters indicated in the following figures “initializing graal” and “during graal”.**

#### 1) Load data set.

- Select the folder your\_directory/graal/data/S1 or your\_directory/graal/data/tricho\_qm6a

- size of pyramid = 6 for tricho\_qm6a ( T. reesei)

- size of pyramid = 4 for S1 ( *S. cerevisiae* )

- sub sampling factor = 3

**2)** Load the corresponding fasta file located in your\_directory/graal/data/fasta/

**3)** Build the pyramid

**4)** Load the pyramid

**5)** Click on the "GRAAL" button

**6)** Fill the parameters

**7)** Click "start" to begin the simulation

**8)** Export trace and histogram of the selected variables

A) The updated contact matrix

B) Real time representation of the genome. Each floating ball corresponds to a bin of restriction fragments

C) Real time visualization of the parameters of the simulation

The visualization is made in a 3D OpenGL window.

Left click and drag in the OpenGL window to move into the graph.

Right click and drag to zoom in/out

Press 'w' to change the background color

Press 'p' to increase the size of the fragments

Press 'm' to decrease the size of the fragments

Figure 1. Initializing GRAAL

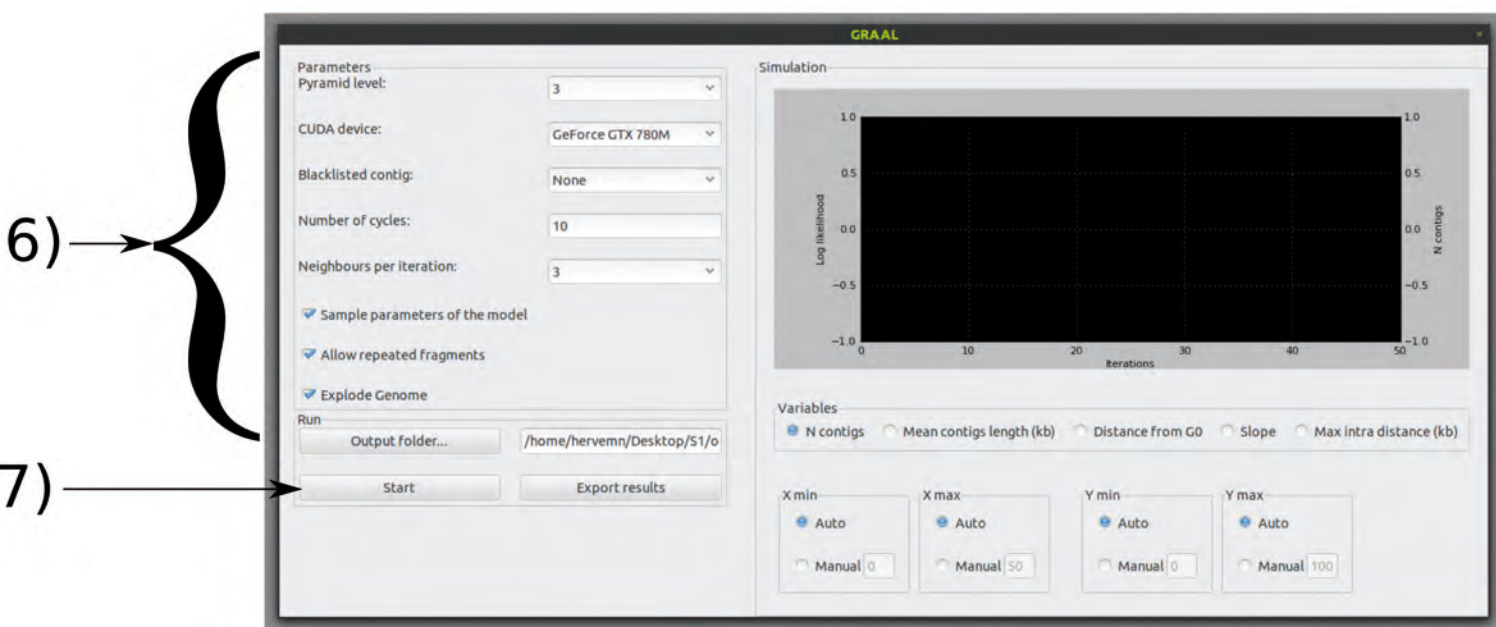
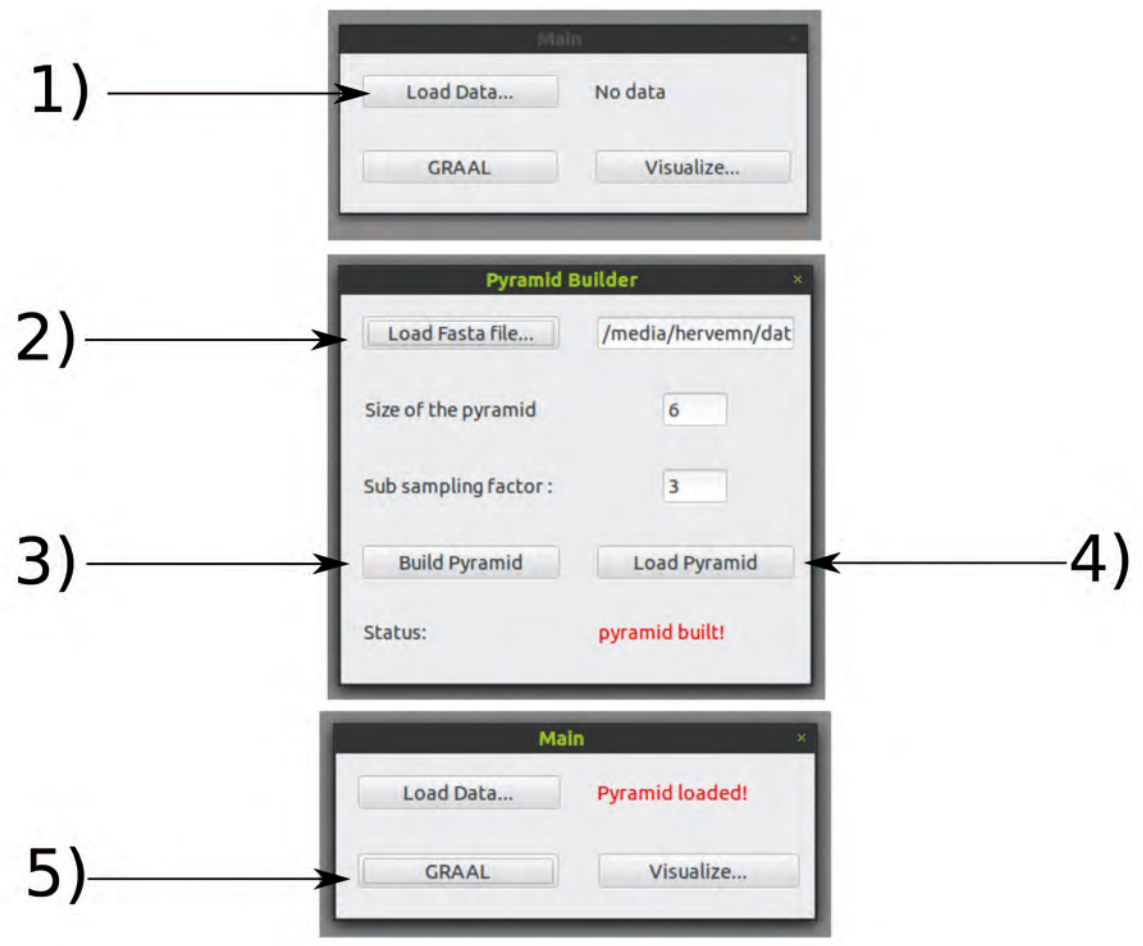


Figure 2. During GRAAL...

