

Biomina NGS Analysis Infrastructure

Author: Geert Vandeweyer

Contact: geert.vandeweyer@uantwerpen.be

Date: 21/7/2013

Version: 0.1.2

Introduction

The purpose of this document is to describe how the High Performance Computing (HPC) infrastructure hosted at the Antwerp University Hospital (UZA) can be used for NGS analysis tasks. At this time, such analysis tasks can be submitted to the HPC cluster using the CLC Genomics Workbench of CLC Bio, Galaxy or using a command line interface.

It is assumed that the NGS data you want to analyze will be produced or has been produced by one of the Illumina sequencers on the Medical Genetics site. In this case, the NGS data will be stored automatically, and made available to you, on the HPC cluster.

CLC Genomics Setup

The following procedure explains how to get up and running with the CLC Genomics Workbench, a commercial tool which needs a license to run. At this time, Biomina has 3 floating licenses available for ad hoc usage by the consortium.

All information for installing and working with the CLC Genomics Workbench is also contained in a [user manual](#).

1. Notify Biomina of your intention to analyze sequencing data by contacting Geert Vandeweyer (geert.vandeweyer2@ua.ac.be) or Tim Van den Bulcke (tim.vandenbulcke@uza.be).

Biomina needs to provide you with an account on the HPC cluster. This account can be used to connect your CLC Genomics Workbench installation with the CLC Genomics Server. The NGS data that is output by the Illumina sequencers at MedGen will be stored in this account as well for easy retrieval by you later.

2. Download and install the CLC Genomics Workbench from the CLC Bio website.

The tool can be downloaded from <http://www.clcbio.com/products/clc-genomics-workbench/> (use the "Get A Free Trial" option).

3. Set up a license for the CLC Genomics Workbench.

The following steps help in setting up one of the floating licenses of Biomina. They can be followed in the manual in the [licenses section](#).

- a. Start up the CLC Genomics Workbench. You will be presented with the *License Wizard*.
- b. Select the '**Configure license server connection**' option. Click '**Next**'. This opens the *License Server Settings* dialog.
- c. Check the '**Enable license server connection**' box. Select the '**Manually specify license server**' option and enter the following values:

- Hostname/IP-address: `ngs.biomina.be`¹
- Port: 6200

- d. Next, click '**OK**'. This should launch the CLC Genomics Workbench.

¹ Or, alternatively, the ip-address 143.169.238.32 if using the domain name does not work.

4. Connect the CLC Genomics Workbench with the CLC Genomics Server.
The following steps help in connecting to the CLC Genomics Server where the NGS data will be stored. They can be followed in the CLC Genomics Server manual in the [relevant plugins section](#).
 - a. Click the '**plugins**' button. This opens the plugins dialog.
 - b. In the list of available plugins, select the '**CLC Workbench Client Plugin**' and install it. Then, restart the CLC Genomics Workbench.
 - c. Now, under the '**File**' menu, there will be a new item, called '**CLC Server Login**' with a blue icon next to it. Clicking this opens a new dialog where you can enter the following values:
 - Username: *<provided to you by Biomina>*
 - Password: *<provided to you by Biomina>*
 - Server host: `ngs.biomina.be2`
 - Server port: `7777`
 - Check '**Save user name and password**'
 - d. Next, click '**Login**'. This will log you in to the CLC Genomics Server. Your data can be found under **serverstorage** in the navigator.

Galaxy Setup

The following procedure explains how to get up and running with the Galaxy Web Platform, a set of free tools. Biomina hosts a server with >100 available cores for the consortium.

1. Go to '<http://143.169.238.104/galaxy>'
2. On top, go to **User => Register**
3. You now have access, and a default quota of 500Gb. To increase your storage space, please contact geert.vandeweyer@uantwerpen.be

Commandline Setup

Commandline is not available for all users. If you would like to request access to the HPC platform, please contact geert.vandeweyer@uantwerpen.be to discuss this.

² Or, alternatively, the ip-address `143.169.238.32` if using the domain name does not work.