



Histone ChIP-Seq Tool

USER MANUAL

User manual for Histone ChIP-Seq 1.2

Windows, Mac OS X and Linux

February 15, 2017

This software is for research purposes only.

QIAGEN Aarhus
Silkeborgvej 2
Prismet
DK-8000 Aarhus C
Denmark



Contents

1 Introduction to the Histone ChIP-Seq plugin	4
1.1 Background	4
1.2 Running the Histone ChIP-Seq tool	5
2 Installation of the Histone ChIP-Seq plugin	7
3 Uninstall	9
Bibliography	11
Index	11

Chapter 1

Introduction to the Histone ChIP-Seq plugin

The Histone ChIP-Seq plugin consists of a new tool that you will find in the Epigenomics Analysis folder in the Toolbox (see figure 1.1).

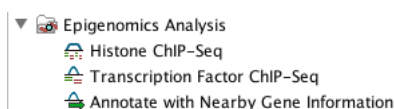


Figure 1.1: The Histone ChIP-Seq tool in the toolbox

1.1 Background

ChIP-Seq experiments are increasingly used to investigate histone modifications. In contrast to transcription factors, histone marks are of variable length and can span across entire gene bodies. Although the experimental procedures are similar, the resulting data needs to be treated accordingly to take this variability into account. While narrow peaks resulting from Transcription Factor ChIP-Seq can be detected using a fixed window size, broad peak detection has to cope with the additional boundary-problem in the sense that the distance between start and end depends on the regions of the underlying genes.

Some existing approaches (e.g. [Heinz et al., 2010](#)) first detect narrow peaks using a fixed window size, and then merge close peaks in order to avoid the computational cost of finding regions of variable length. Nevertheless, different histone marks can also exhibit distinct shapes across gene bodies (e.g. [Li et al., 2007](#)), which renders them amenable to a shape-based detection algorithms.

By using existing annotations, the Histone ChIP-Seq tool is able to classify gene regions according to the peak shape and thereby provides a good practical trade-off between computational complexity and biological sensitivity. The primary application areas are the analysis of ChIP-Seq data for diverse histone-modifications such as (mono-, di-, and tri-) methylation, acetylation, ubiquitination, etc., in combination with a set of annotated gene regions. The tool is well suited to analyze data from organisms with available gene annotations, while finding peaks in intergenic regions can be accomplished with the Transcription Factor ChIP-Seq tool.

1.2 Running the Histone ChIP-Seq tool

To run the Histone ChIP-Seq tool:

Toolbox | Epigenomics Analysis (📁) | Histone ChIP-Seq (🔧)

In the first wizard window, select the mapped ChIP-Seq reads as input data (figure 1.2). Multiple inputs (such as replicate experiments) are accepted, provided that they refer to the same genome. It is also possible to work in batch (see http://resources.qiagenbioinformatics.com/manuals/clcgenomicsworkbench/current/index.php?manual=Batch_processing.html).

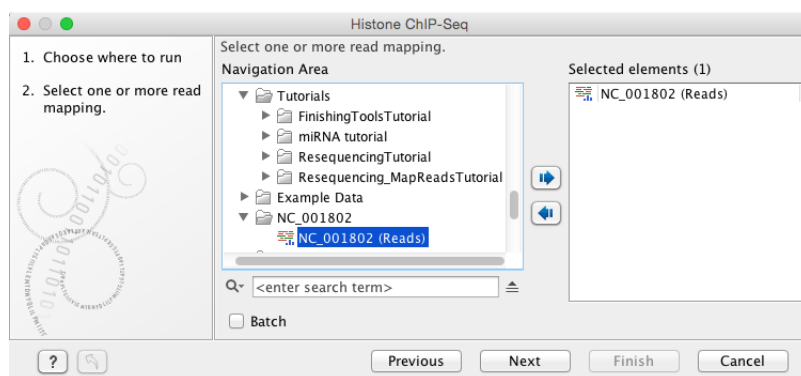


Figure 1.2: Selecting input tracks for the Histone ChIP-Seq tool

In the second step (figure 1.3), the gene track and control data are defined, along with the p-value. This value defines which regions have a significant fit with the peak-shape, and only these are copied to the output track.

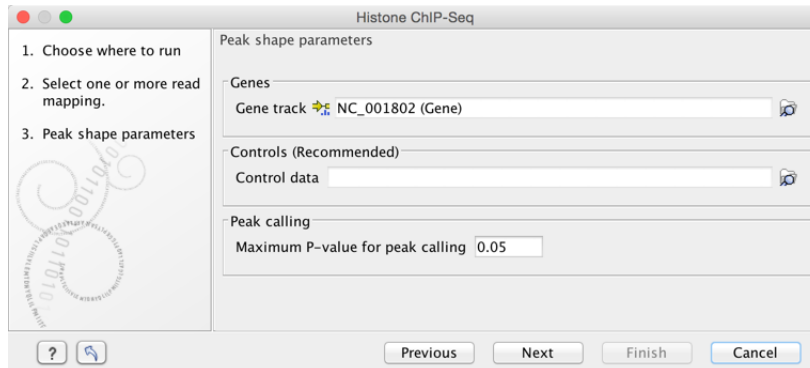


Figure 1.3: Setting peak shape parameters

The output options are shown in figure 1.4.

- **Create a Quality Control (QC) report** with which you can check the quality of the reads. It lists the number of mapped reads, the normalized strand coefficient, and the relative strand correlation for each mapping. For each metric, the **Status** column will be **OK** if the experiment has good quality or **Low** if the metric is not as high as expected. Furthermore, the QC report will show the mean read length, the inferred fragment length, and the window size that the algorithm would need to be able to model the signal shape. In case the input contains paired-end reads, the report will also contain the empirical fragment length distribution.

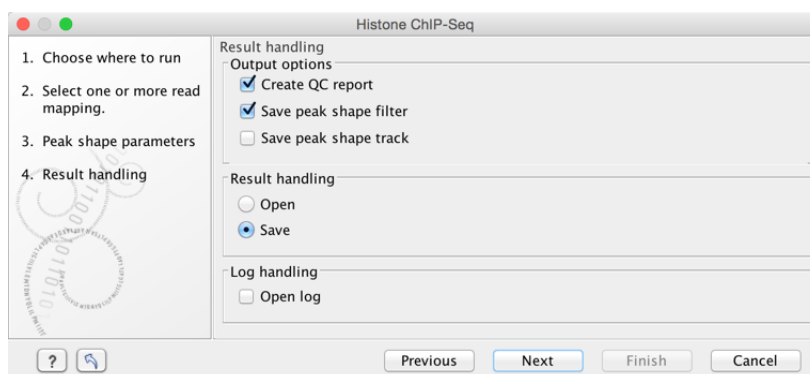


Figure 1.4: Setting up result handling

- **Save the peak-shape filter** generated by the tool while processing. This filter can be used to identify genomic regions whose read coverage profile matches the characteristic peak shape, as well as to determine the statistical significance of this match. The filter implemented is called Hotelling Observer and was chosen because it is the matched filter that maximizes the AUCROC (Area Under the Curve of the Receiver Operator Characteristic), one of the most widely used measures for algorithmic performance. For a more detailed description of peak-shape filters, please refer to the description of the Transcription Factor ChIP-Seq tool http://resources.qiagenbioinformatics.com/manuals/clcgenomicsworkbench/current/index.php?manual=Learning_peak_shapes.html, or to the white-paper explaining the algorithmic and statistical methods <https://www.qiagenbioinformatics.com/files/whitepapers/whitepaper-chip-seq-analysis.pdf>. The peak-shape filter is then applied to the experimental data by scaling the coverage profile in every gene region to a unit-window. The score is obtained for each region by comparing this profile to the peak shape filter.
- **Save peak shape score graph track** (figure 1.5)

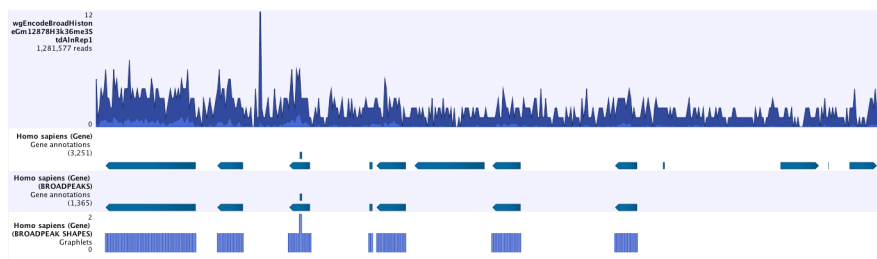


Figure 1.5: Example of Histone ChIP-Seq output

The peak shape score is standardized and follows a standard normal distribution, so a p-value for each regions is calculated. After the peak shape score for all regions is calculated, regions where the peak shape score is greater than the given threshold are copied to the output track. Hence the output only contains the gene regions where the coverage graph does match the peak-shape.

Chapter 2

Installation of the Histone ChIP-Seq plugin

Plugins are installed using the plugin manager. In order to install plugins on Windows, the Workbench must be run in administrator mode: Right-click the program shortcut and choose "Run as Administrator". Then follow the procedure described below.

Help in the Menu Bar | Plugins... ()

or **Plugins () in the Toolbar**

The plugin manager has two tabs at the top:

- **Manage Plugins.** This is an overview of plugins that are installed.
- **Download Plugins.** This is an overview of available plugins on CLC bio's server.

To install a plugin, click the **Download Plugins** tab. This will display an overview of the plugins that are available for download and installation (see figure 2.1).

Clicking a plugin will display additional information at the right side of the dialog. This will also display a button: **Download and Install**.

Click the Histone ChIP-Seq plugin and press **Download and Install**. A dialog displaying progress is now shown, and the plugin is downloaded and installed.

If the Histone ChIP-Seq plugin is not shown on the server, and you have it on your computer (for example if you have downloaded it from our website), you can install it by clicking the **Install from File** button at the bottom of the dialog. This will open a dialog where you can browse for the plugin. The plugin file should be a file of the type ".cpa".

When you close the dialog, you will be asked whether you wish to restart the CLC Workbench. The plugin will not be ready for use until you have restarted.

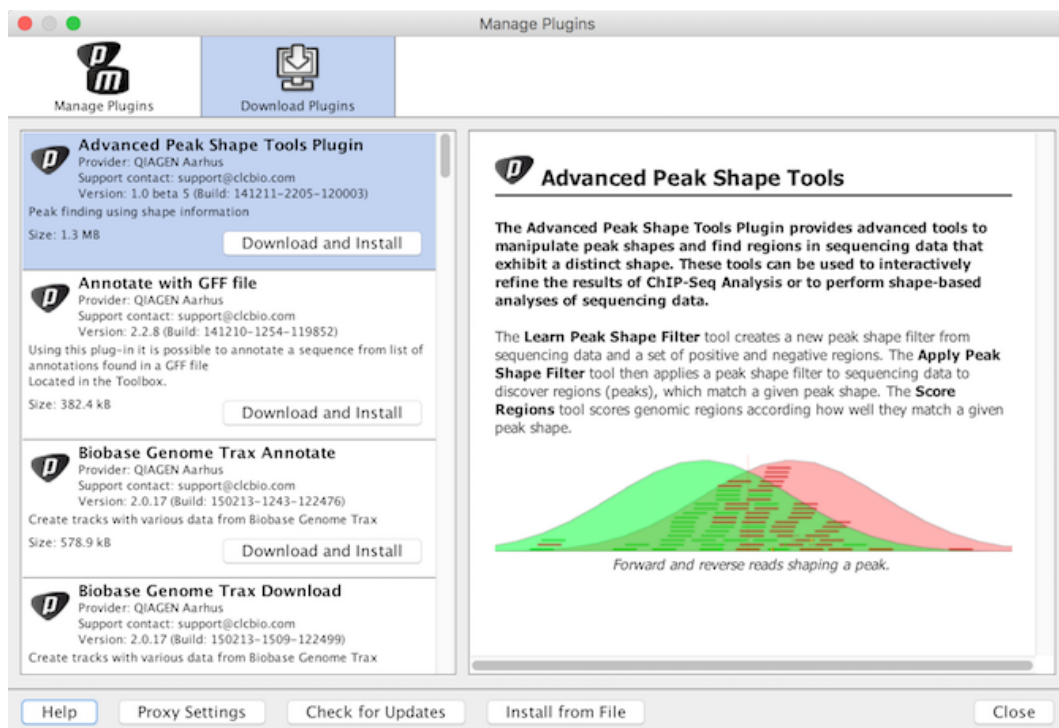


Figure 2.1: The plugins that are available for download.

Chapter 3

Uninstall

Plugins are uninstalled using the plugin manager:

Help in the Menu Bar | Plugins... ()

or **Plugins () in the Toolbar**

This will open the dialog shown in figure 3.1.

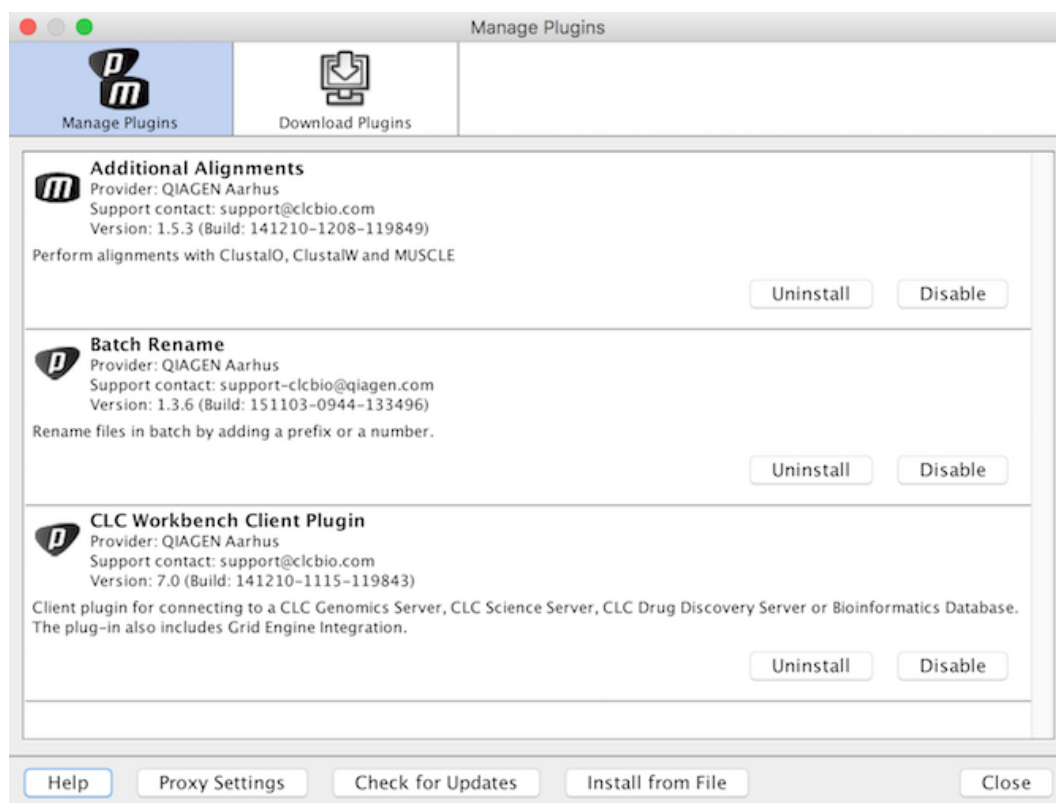


Figure 3.1: The plugin manager with plugins installed.

The installed plugins are shown in this dialog. To uninstall:

Click the Histone ChIP-Seq plugin | Uninstall

If you do not wish to completely uninstall the plugin but you don't want it to be used next time you start the Workbench, click the **Disable** button.

When you close the dialog, you will be asked whether you wish to restart the workbench. The plugin will not be uninstalled until the workbench is restarted.

Bibliography

- [Heinz et al., 2010] Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y. C., Laslo, P., Cheng, J. X., Murre, C., Singh, H., and Glass, C. K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol cell*, 38(4):576–589.
- [Li et al., 2007] Li, B., Carey, M., and Workman, J. L. (2007). The role of chromatin during transcription. *Cell*, 128(4):707–719.

Index

Bibliography, [11](#)

References, [11](#)