

Advanced Peak Shape Tools Plugin

USER MANUAL

User manual for Advanced Peak Shape Tools Plugin 1.0

Windows, macOS and Linux

November 23, 2018

This software is for research purposes only.

QIAGEN Aarhus
Silkeborgvej 2
Prismet
DK-8000 Aarhus C
Denmark



Contents

1	Introduction	4
2	Learn Peak Shape Filter	5
3	Apply Peak Shape Filter	8
4	Score Regions	11
5	Install and uninstall plugins	13
5.1	Install	13
5.2	Uninstall	14

Chapter 1

Introduction

Advanced Peak Shape Tools Plugin provides advanced tools to manipulate peak shapes and find regions in sequencing data that exhibit a distinct shape. These tools are used by the CHIP-Seq algorithm (see http://resources.qiagenbioinformatics.com/manuals/clcgenomicsworkbench/current/index.php?manual=ChIP_Seq_Analysis.html), but they can also be used interactively to refine the results of CHIP-Seq peak calling or to perform shape-based analyses of sequencing data.

These tools are available under:

Toolbox | Epigenomics Analysis  | **Advanced Peak Shape Tools** 

The **Learn Peak Shape Filter** tool allows you to build a new peak shape filter from sequencing data and a set of positive and negative regions.

The **Apply Peak Shape Filter** tool allows you to apply a peak shape filter to sequencing data to discover regions (peaks), which match a given peak shape.

The **Score Regions** tool allows you to apply a peak shape filter to score genomic regions according how they match a given peak shape.

Chapter 2

Learn Peak Shape Filter

The Learn Peak Shape Filter tool allows you to build a new peak shape filter from sequencing data and a set of positive and negative regions. The resulting filter can be used to identify genomic regions whose read coverage profile matches the characteristic shape of the positive examples and does not match the shape of the negative examples. The procedure used to build the peak shape filter is described here http://resources.qiagenbioinformatics.com/manuals/clcgenomicsworkbench/current/index.php?manual=Learning_peak_shapes.html. An example of such filter is shown in figure 2.1.

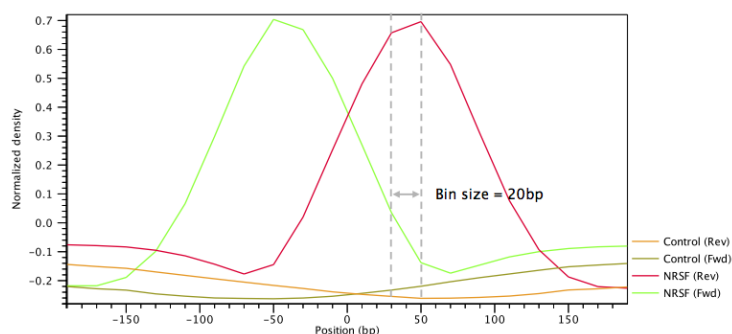


Figure 2.1: Example of a peak shape filter with a window size of 400bp made up of 20 bins of size 20bp each. The filter was built from ChIP-Seq data of the transcription factor NRFS and a control ChIP-Seq experiment.

To run the Learn Peak Shape Filter tool:

Toolbox | Epigenomics Analysis (📁) | **Advanced Peak Shape Tools** (📁) | **Learn Peak Shape Filter** (📁)

This will open up the wizard shown in figure 2.2 where you can select the input data (for example the mapped ChIP-Seq reads). Track based read mappings (📄) and stand-alone read mappings (📄) / (📄) are both accepted. Multiple inputs are accepted, provided that they refer to the same genome.

Click on the button labeled **Next** to go to the next wizard step (shown in figure 2.3).

In this wizard step you have the following options:

- **Location of positive regions** An annotation track (📄) containing the location of the positive

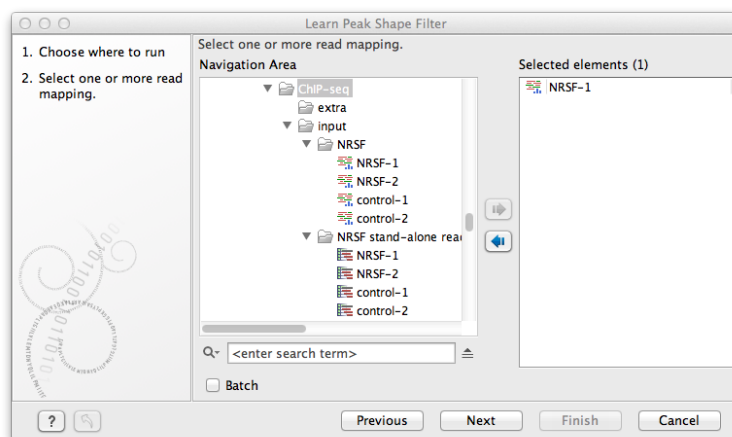


Figure 2.2: Select the input data for the Learn Peak Shape Filter tool.

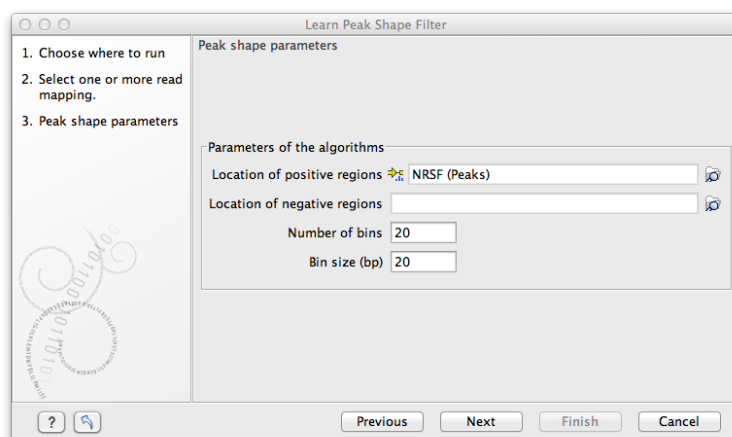



Figure 2.3: Options for Learn Peak Shape Filter.

regions (e.g. ChIP-Seq peaks) that will be used to build the peak shape filter. The set of positive regions should include examples where the shape is clearly exhibited. It is preferable to have fewer peaks with high quality rather than a large amount of ambiguous peaks. Typically, a number of positive peaks greater than 5-10 times the number of bins is sufficient to learn a well-defined shape.

- **Location of negative regions** An annotation track (➡) containing the location of the negative regions that will be used to build the peak shape filter (e.g. background, PCR artifacts or examples of bad peaks from a previous run of the ChIP-Seq analysis tool). If no annotation track is provided, a negative profile will be derived from sequencing noise.
- **Number of bins** The number of bins to use to build the filter. The default value of 20 for the Number of bins parameter should be satisfactory for most uses. Higher values may be useful when the shape to be learned is particularly complex. Note that if the chosen number of bins is very large, the learned peak shape filter may not be smooth and could over-fit the input data. If only few positive regions are available, reducing the number of bins may be helpful.
- **Bin size** The size of each bin in base pairs. The bin size is related to the window size (i.e. the length of the shape to be learned) by the formula $\text{Window size} = \text{Bin size} \times \text{Number of bins}$ (see figure 2.1).

The result of the algorithm will be a **Peak shape filter** , which can then be applied to call peaks or score regions using Apply Peak Shape Filter. After clicking on the button labeled **Next**, you can choose whether you want to open the result directly, or save the results in the **Navigation Area**. If you choose to save the results, you will be asked to specify where you would like to save them.

Chapter 3

Apply Peak Shape Filter

The Apply Peak Shape Filter tool allows you to apply a peak shape filter to sequencing data to discover regions (peaks), which match a given peak shape (see http://resources.qiagenbioinformatics.com/manuals/clcgenomicsworkbench/current/index.php?manual=Applying_peak_shape_filters_call_peaks.html).

To run the Apply Peak Shape Filter tool:

Toolbox | Epigenomics Analysis (📁) | **Advanced Peak Shape Tools** (📁) | **Apply Peak Shape Filter** (🚀)

This will open up the wizard shown in figure 3.1 where you can select the input data (e.g. mapped ChIP-Seq reads). Track based read mappings (📊) and stand-alone read mappings (📄) / (📄) are both accepted. Multiple inputs are accepted, provided that they refer to the same genome. At least 2 input files must be provided (sample + control for example) if the tool is to be used with a filter that was built from a sample and a control. Note: the order of the inputs should be the same when building the filter and applying the filter, i.e., if the filter was generated from Sample 1 + control, with control being the second object transferred over in the wizard, then when running the Apply Peak Shape Filter tool, the control should also be the second object used in the wizard.

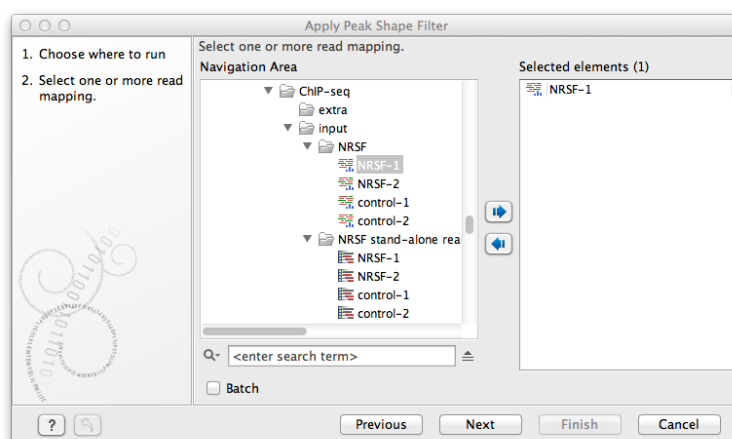


Figure 3.1: Select the input data for Apply Peak Shape Filter.

Click on the button labeled **Next** to go to the next wizard step (shown in figure 3.2).

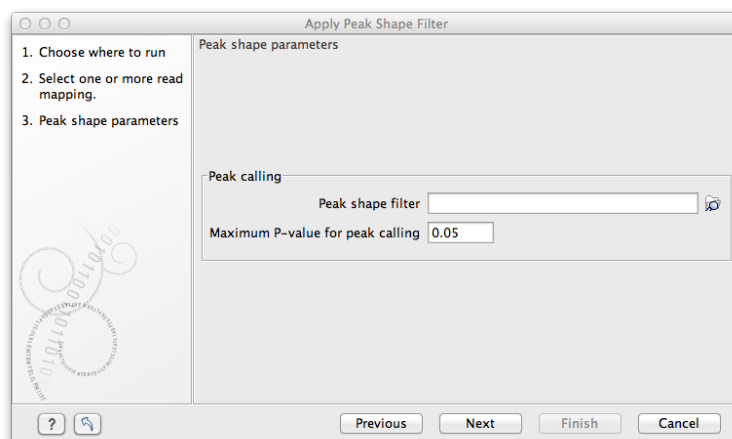


Figure 3.2: Options for Apply Peak Shape Filter.

In this wizard step you have the following options:

- **Peak shape filter** The peak shape filter (📊) to apply to the data. Peak shape filters can be obtained as the result of the ChIP-Seq Analysis tool. If no filter is given, a filter is derived from the input data.
- **Maximum P-value for peak calling** The threshold for reporting peaks can be specified by this option.

Click on the button labeled **Next** to go to the wizard step shown in figure 3.3.

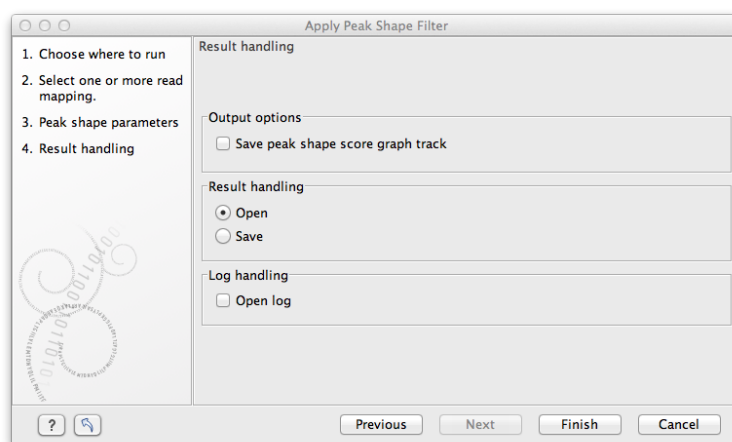



Figure 3.3: Output options for Apply Peak Shape Filter.

In addition to the annotation track with **Peak annotations** (📊) that will always be generated by the algorithm, you can choose to select an additional output type:

- **Peak shape score** (📊) A graph track containing the peak shape score. The track shows the peak shape score for each genomic position. To save disk space, only scores greater than zero are reported. For the definition of peak shape score.

Choose whether you want to open the results directly, or save the results in the **Navigation Area**. If you choose to save the results, you will be asked to specify where you would like to save them.

For more information on the **Peak track**  (see the CLC Genomics Workbench manual http://resources.qiagenbioinformatics.com/manuals/clcgenomicsworkbench/current/index.php?manual=Peak_track.html).

Chapter 4

Score Regions

The Score Regions tool allows you to apply a new peak shape filter to score genomic regions according how they match a given peak shape.

To run the Score Regions tool:

Toolbox | Epigenomics Analysis (📁) | **Advanced Peak Shape Tools** (📁) | **Score Regions** (🚀)

This will open up the wizard shown in figure 4.1 where you can select the input data (e.g. mapped ChIP-Seq reads). Multiple inputs are accepted, provided that they refer to the same genome. Track based read mappings (📊) and stand-alone read mappings (📊) / (📊) are both accepted.

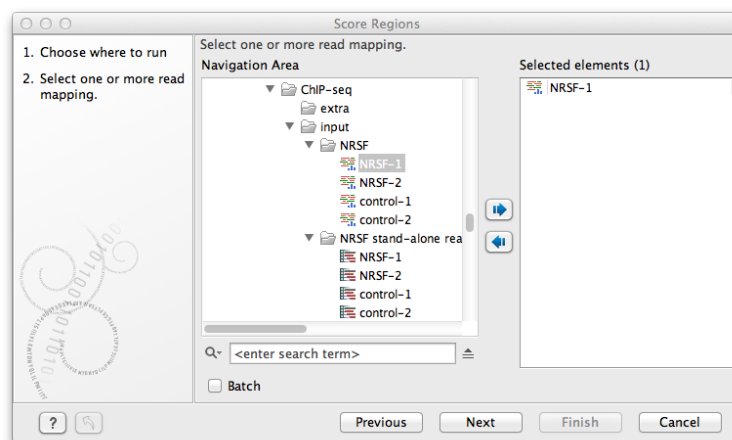


Figure 4.1: Select the input data for Score Regions.

Click on the button labeled **Next** to go to the next wizard step (shown in figure 4.2).

In this wizard step you have the following options:

- **Peak shape filter** The peak shape filter (📊) to apply to the data. Peak shape filters can be obtained as the result of the ChIP-Seq Analysis tool.
- **Regions to score** An annotation track (📊) containing the regions where the peak shape will be applied. The peak shape filter will be applied to every genomic position within the interval and the maximum values will be used to score the region.

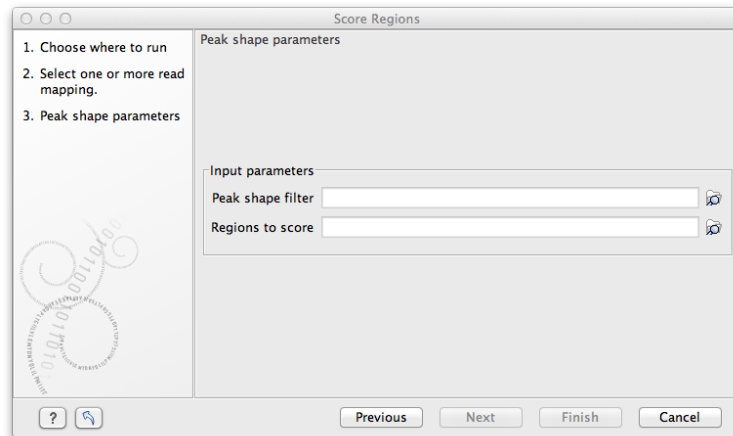


Figure 4.2: Options for Score Regions.

The result of the algorithm will be an annotation track (➡) of the same type as the regions to score annotation track, where the columns of type **Peak shape score**, **P-value** and **Center of peak** will be added or replaced.

After clicking **Next**, you can choose whether you want to open the result directly, or save the results in the Navigation Area. If you choose to save the results, you will be asked to specify where you would like to save them.



Chapter 5

Install and uninstall plugins

Advanced Peak Shape Tools is installed as a plugin.

Note: In order to install plugins and modules, the Workbench must be run in administrator mode. On Linux and Mac, it means you must be logged in as an administrator. On Windows, you can do this by right-clicking the program shortcut and choosing "Run as Administrator".

Plugins are installed and uninstalled using the plugin manager.

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 QIAGEN Aarhus server.

5.1 Install

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 5.1).

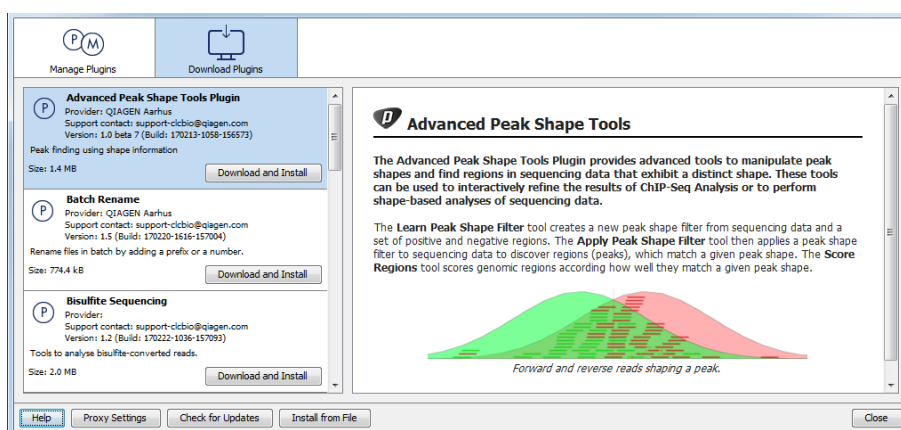


Figure 5.1: The plugins that are available for download.

Select Advanced Peak Shape Tools to display additional information about the plugin on the right

side of the dialog. Click **Download and Install** to add the plugin functionalities to your workbench.

Accepting the license agreement

The end user license agreement (EULA) must be read and accepted as part of the installation process. figure 5.2.

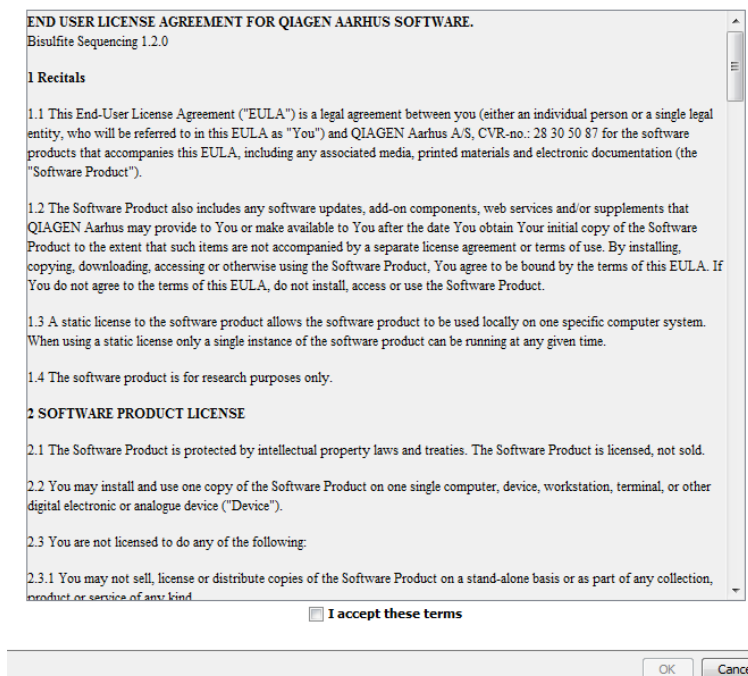


Figure 5.2: The End User License Agreement is presented during the installation process.



Please read the EULA text carefully, and if you agree to it, check the box next to the text **I accept these terms**. If further information is requested from you, please fill this in before clicking on the **Finish** button.

If Advanced Peak Shape Tools is not shown on the server but you have the installer file on your computer (for example if you have downloaded it from our website), you can install the plugin by clicking the **Install from File** button at the bottom of the dialog and specifying the plugin *.cpa file saved on your computer.

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

5.2 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 5.3.

The installed plugins are shown in the **Manage plugins** tab of the plugin manager. To uninstall, select Advanced Peak Shape Tools and click **Uninstall**.

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

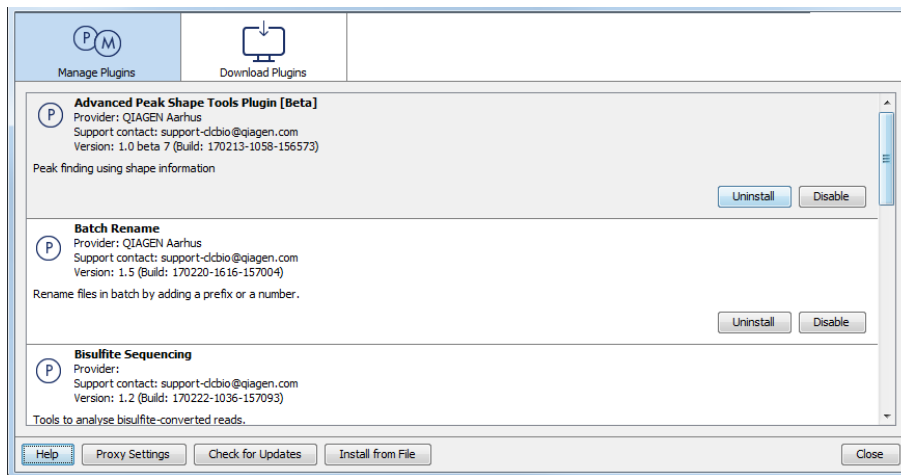


Figure 5.3: *The plugin manager with plugins installed.*

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.