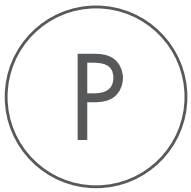


Ingenuity Pathway Analysis
Plugin

USER MANUAL

User manual for
Ingenuity Pathway Analysis 2.3

Windows, Mac OS X and Linux

November 7, 2017

This software is for research purposes only.

QIAGEN Aarhus
Silkeborgvej 2
Prismet
DK-8000 Aarhus C
Denmark

Contents

Chapter 1

Ingenuity Pathway Analysis Plugin

1.1 Introduction to Ingenuity Pathway Analysis Plugin

The Ingenuity Pathway Analysis Plugin provides the ability to upload expression data from Biomedical Genomics Workbench or CLC Genomics Workbench to Ingenuity Pathway Analysis (IPA). It is possible to upload:

- Statistical comparison data generated using the RNA-seq tools, and
- Legacy experiment data generated in the CLC Genomics Workbench or Biomedical Genomics Workbench.

The purpose of the integration is to supplement the abilities of the Biomedical Genomics Workbench or CLC Genomics Workbench with the pathway information and visualization capabilities available in IPA.

IPA provides valuable biological insight into the results of gene expression experiments by uncovering enriched signaling and metabolic pathways, activated and inhibited upstream regulators and effects on downstream diseases, functions, and phenotypes. IPA can visualize at the isoform level for human genes.

The plugin comes with two tools and, when installed using Biomedical Genomics Workbench, a ready-to-use workflow (figure 1.1):

- The **Pathway Analysis** tool uploads statistical comparison data (generated by the tool Differential Expression for RNA-Seq) to IPA. The Pathway Analysis tool is based on the legacy Pathway Analysis tool, but has been thoroughly improved on many aspects such as usage of the new IPA API, stability, error handling, and user feedback during the upload process. The tool will output one or more Statistical Comparison tracks.
- The **Pathway Analysis (legacy)** tool, which can upload legacy experiment data. The tool is primarily included for backwards compatibility reasons.
- The ready-to-use workflow **Analyze Expression Data and Upload Comparisons to IPA**, which takes expression data as input. The workflow analyzes them using the RNA-Seq Analysis tools, and submits the comparisons to IPA using the Pathway Analysis tool.

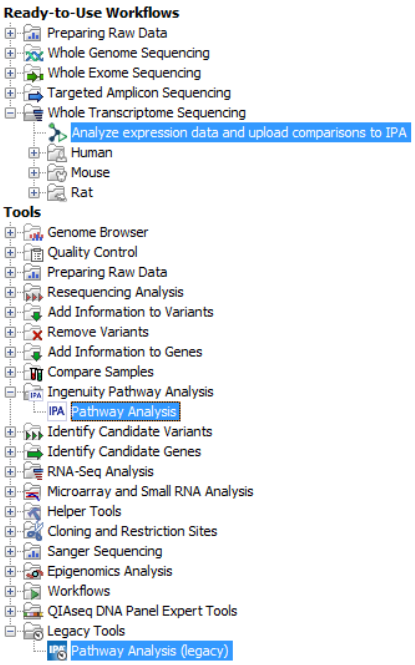


Figure 1.1: *IPA plugin tools and workflow when installed in the Biomedical Genomics Workbench Toolbox.*

It is possible to use gene and transcript based RNA-seq experiments as basis for the analysis, but also microarrays from Illumina and Affymetrix are supported. Currently, small RNA based experiments are not supported by this integration.

Once the experiment data are ready, it is possible to annotate with any of the supported statistics:

- Transformed and normalized foldchange
- Baggerley's test
- Kal's Z test
- ANOVA
- EdgeR

1.2 Uploading data to IPA using the Pathway Analysis tool

Launch the Pathway Analysis tool from the toolbox:

Ingenuity Pathway Analysis | Pathway Analysis

Use one or several statistical comparison(s) as input ( ) (figure 1.2), and click **Next**.

Under **Set configuration** (figure 1.3), you get the following options:

Ingenuity username (email address) Type your Ingenuity account name, which typically will be your email adress.

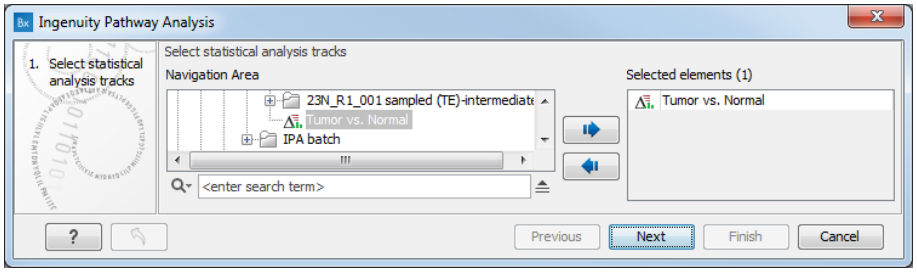


Figure 1.2: *Select a statistical comparison to analyze.*

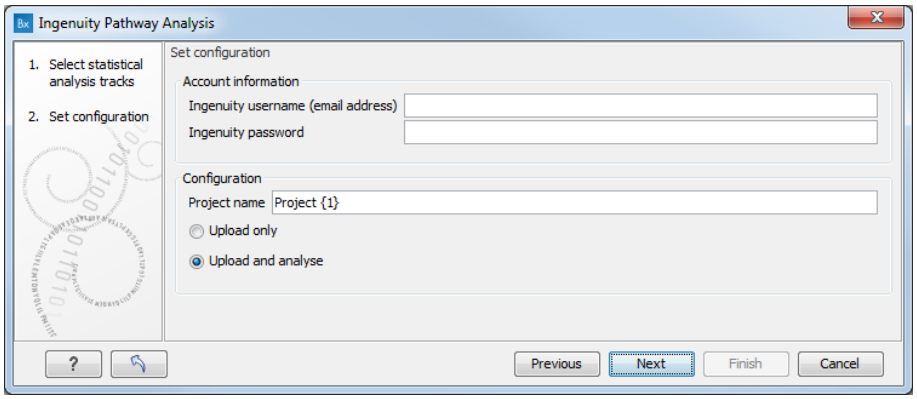


Figure 1.3: *Configure the tool to upload and potentially analyze the statistical comparison data in IPA.*

Ingenuity password Type your Ingenuity password.

Project Name This will be the name of the project in IPA once created. {1} will be substituted with a date stamp. It is also possible to create a custom project name by typing in the desired name in this field.

Upload only / Upload and analyse Select "Upload only", if you only wish to create a dataset in IPA. Select "Upload and analyse", if you want to create an analysis from the dataset as well.

Click on the button labeled **Next** to go to the next wizard step (figure 1.4).

In this wizard step, the cutoff values for what should be uploaded to IPA can be specified. Only features that pass the cutoffs that have been specified at this step will be sent to IPA and be part of the dataset that can be seen in IPA.

Under **Set upload parameters** you get the following options:

Ignore features with mean expression values below This value is used to filter genes/transcripts before uploading them to IPA. Features with values below this limit will not be uploaded.

Upload rows with value <= Maximum p-value for feature (gene or transcript) to be uploaded. Features with a p-value above this number will not be uploaded. It is possible to choose between different types of p-Values: Standard, Bonferroni, and FDR.

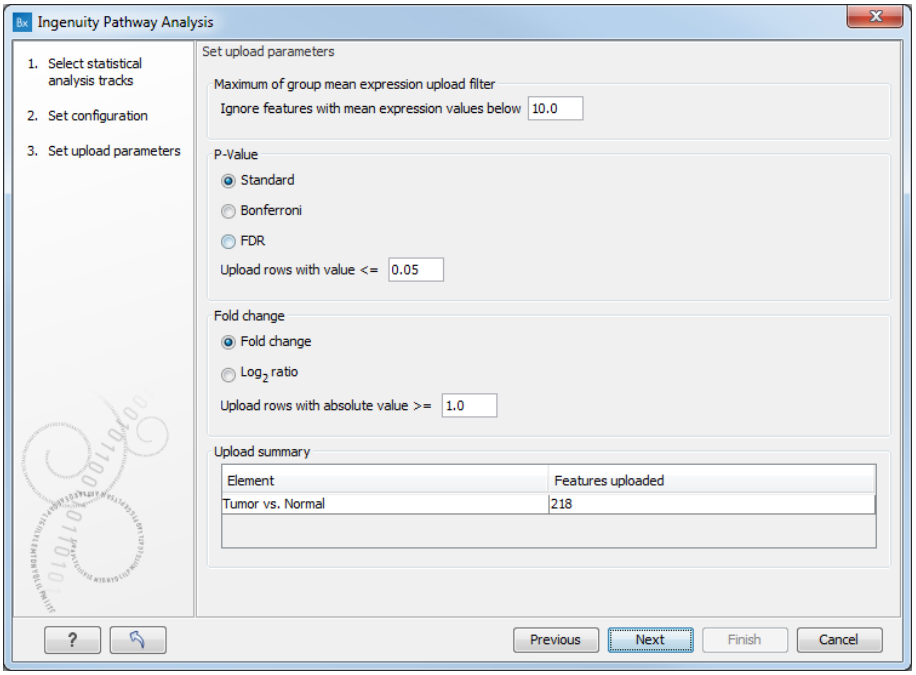


Figure 1.4: *Configure the parameters for uploading the data to IPA.*

Upload rows with absolute value >= Minimum absolute fold change for feature to be uploaded. Features with a fold change/log2 ratio below this number will not be uploaded. It is possible to choose between different types of fold changes: Fold change, and $\log_2$-ratio.

Upload Summary This summary shows how many features the tool will upload to IPA for each statistical comparison. The values are updated, when the user changes any of the upload parameters. In this way, the user can easily check the effect of the filtering (for instance to avoid setting the filters such that no features will be uploaded)

If you had selected "Upload only" in the first step, click **Finish** to start the tool. But if you had selected "Upload and analyze", click Next to see the dialog shown in figure 1.5.

Under **Set analysis parameters**, you get the following options:

Maximum of group mean expression analysis filter | Analysis cutoff Minimum group mean expression value for feature (gene or transcript) to be used in analysis. Features with a group mean expression value below this number will be uploaded, but will be ignored in the analysis.

p-Value | Analysis cutoff Maximum p-value for feature (gene or transcript) to be used in analysis. Features with a p-value above this number will be uploaded, but will be ignored in the analysis. It is possible to choose between different types of p-values: Standard, Bonferroni, and FDR.

Fold change | Analysis cutoff Minimum absolute fold change for feature to be used in analysis. Features with a fold change/$\log_2$ ratio below this number will be uploaded, but will be ignored in the analysis. It is possible to choose between different types of fold changes: Fold change, and $\log_2$-ratio.

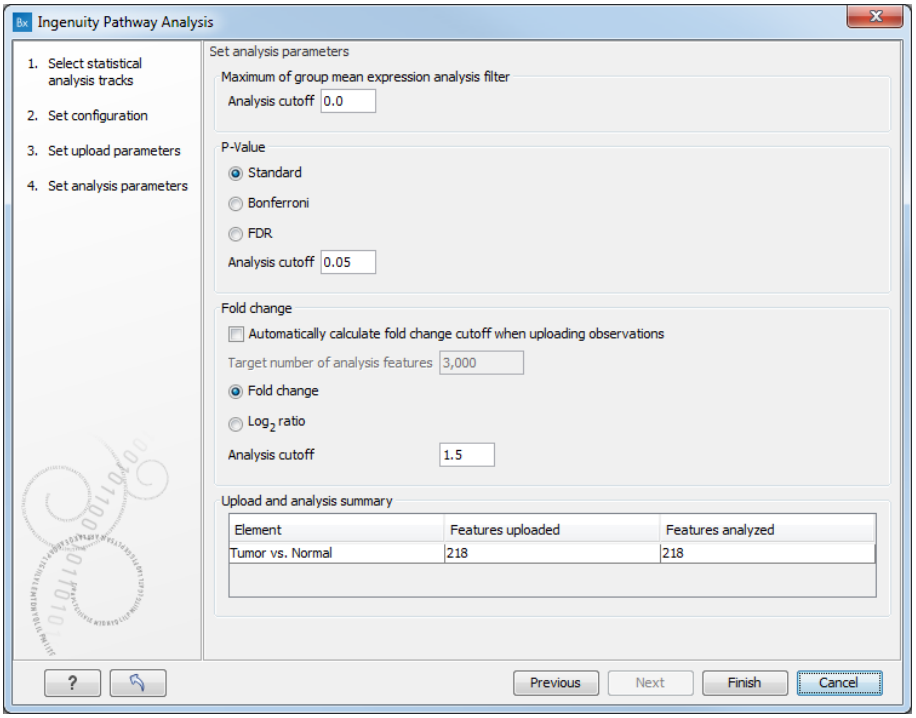


Figure 1.5: *Set the parameters for the analysis of the data in IPA.*

Fold change | Automatically calculate fold change cutoff when uploading observations Automatically calculate fold change when uploading observation. The fold change cutoff will be set so that the number of features to include in the analysis gets as close to the targeted number as possible (see below). When this option is used, it is not necessary to set the "Fold change | Analysis cutoff", since it is automatically calculated by the tool for each statistical comparison. When using this option, the fold change analysis cutoff can be different for each statistical comparison.

Fold change | Target number of analysis features Enabled ony when using automatically calculated fold change. The fold change cutoff will be set so that the number of features to include in the analysis gets as close to the targeted number as possible

Upload and analysis summary This summary shows how many features the tool will upload to IPA for each statistical comparison, and how many features that will be included in each analysis. The values are updated when the user changes any of the analysis parameters. In this way, the user can easily check the effect of the filtering (e.g. avoid setting the filters such that no features will be analyzed).

The **Upload and analysis summary** table at the bottom of the dialog warns the user when too restrictive filters have been set (figure 1.6).

Click Next to choose the reference as seen in figure 1.7.

The reference can be:

Ingenuity Knowledge Base (Genes only)

Uploaded dataset The data uploaded will be analyzed using itself as a reference.

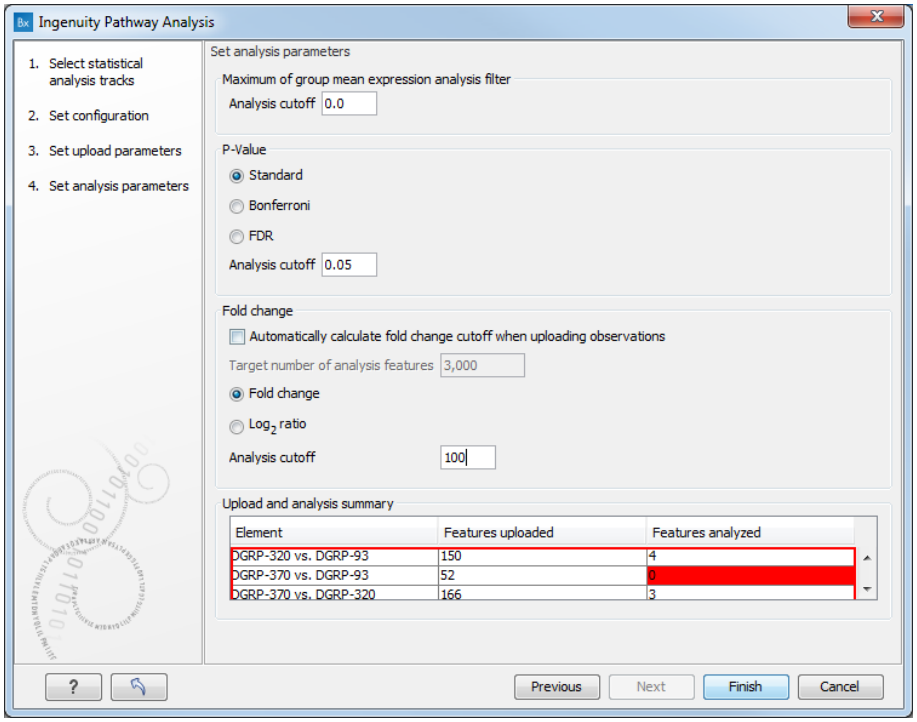


Figure 1.6: *A warning highlight in red analyses for which the cutoff is too restrictive.*

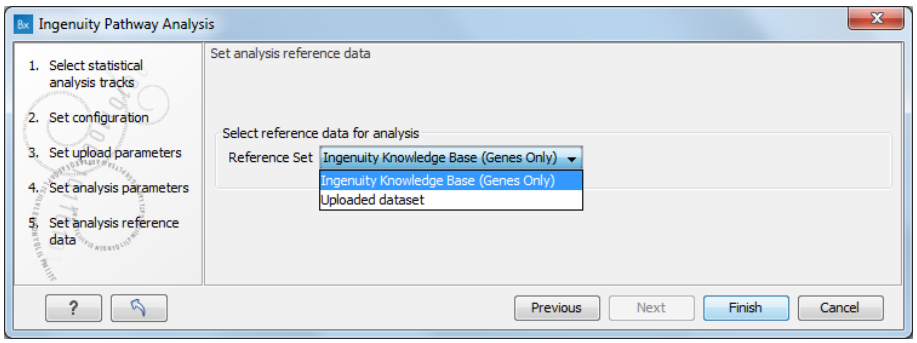


Figure 1.7: *Choose the reference to be used for the analysis of the data in IPA.*

Click **Finish** to start the tool.

1.2.1 Error handling

Concurrent IPA sessions

When submitting data to IPA, the upload may fail, if several sessions are established concurrently using the same username. This can happen when the IPA application is running while data is being uploaded using the IPA tool, or if the IPA tool is part of a workflow with several IPA uploads running at the same time.

To deal with this issue, the IPA tool has a retry functionality, such that the upload is attempted again, if it failed due to a problem with concurrent sessions. The waiting time between retry attempts is increased with a random factor for each attempt, to avoid that two processes continue to block each other.

Upload multiple statistical comparisons

If the IPA tool encounters an error when uploading multiple statistical comparisons, it will in most case continue uploading the remaining statitsical comparisons. However, if the tool gets one of the errors below, it stops uploading immediately, because it cannot expect that they will be resolved before the next upload:

- Login error (wrong username/password)
- User agreement not accepted
- License expired
- Upload limit exceeded
- Analysis limit exceeded

1.3 Using the Workflow Analyze Expression Data and Upload Comparisons to IPA

If you are working with Biomedical Genomics Workbench, Ingenuity Pathway Analysis Plugin includes the workflow Analyze Expression Data and Upload Comparisons to IPA. The workflow takes expression data as input, and analyzes them using the tools in the RNA-Seq Analysis folder. It then submits the comparison to IPA using the Pathway Analysis tool. Please consult the Biomedical Genomics Workbench user manual for information about the tools used in the workflow.

The workflow is located in the Whole Transcriptome Sequencing folder as shown in figure 1.8 below:

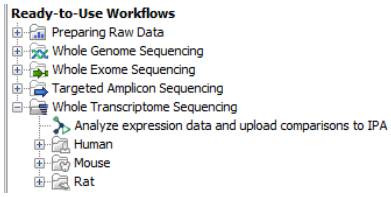


Figure 1.8: *Location of the ready-to-use workflow in the Toolbox.*

The purpose of the workflow is to make it as easy as possible for the user to get from Sample to Insight. The user only has to provide expression data as input, and the workflow generates all available statistical analyses and data interpretation capabilities available via Biomedical Genomics Workbench and IPA.

Opened in the workflow editor, the workflows looks like this (see figure 1.9 below):

The expression tracks are sent to three tools:

- **Create Heat Map for RNA-Seq** The tool creates a two dimensional heat map of expression values. Each column corresponds to one sample, and each row corresponds to a feature (a gene or a transcript). The samples and features are both hierarchically clustered.

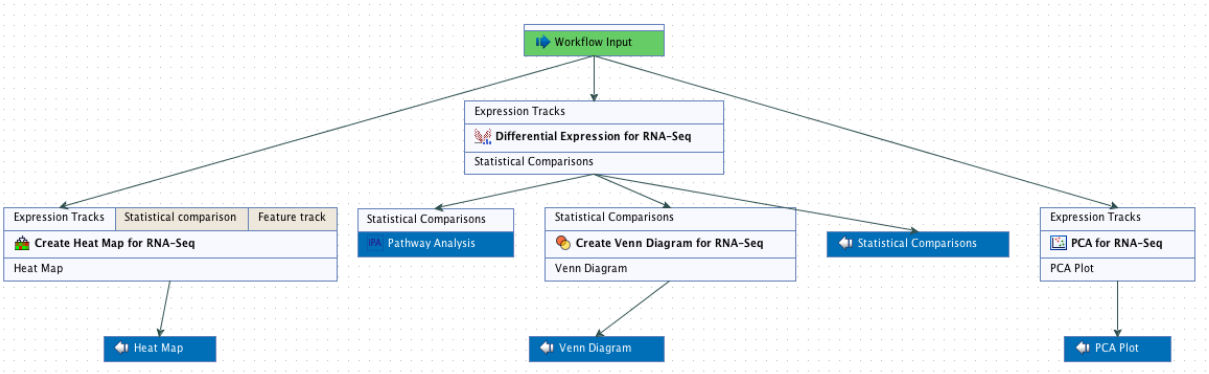


Figure 1.9: *Layout of the IPA plugin workflow.*

- **Differential Expression for RNA-Seq** The tool performs a statistical differential expression test for a set of Expression Tracks. It's outputs are used as inputs for the IPA tool and for Create Venn Diagram for RNA-Seq (see below).
- **PCA for RNA-Seq** The tool creates a PCA plot, which is a projection of a high-dimensional dataset (where the number of dimensions equals the number of genes or transcripts) onto two of three dimensions. This helps in identifying outlying samples for quality control, and gives a feeling for the principal causes of variation in a dataset.

The outputs from the tools are saved in the chosen output folder for the workflow. The outputs from the Differential Expression for RNA-Seq tool are furthermore used for processing by these two tools:

- **Pathway Analysis** The tool uploads the comparisons to IPA. See chapter 1.2 for details.
- **Create Venn Diagram for RNA-Seq** The tool makes it possible to compare two or more statistical comparison tracks. The Venn diagram comparison visualizes the overlap between the differentially expressed genes or transcripts in the selected statistical comparison tracks. The genes considered to be differentially expressed can be controlled by setting appropriate p-value and fold change thresholds.

1.3.1 Running the Workflow

The workflow can be started from the toolbox, or by using the Launch button ( ).

Choose the expression data to be analyzed and uploaded (see figure 1.10).

Following this, the parameters for the Differential Expression for RNA-Seq need to be specified (see figure 1.11):

Metadata table Select a metadata object that associates the selected input objects to metadata used by the RNA-Seq analysis.

Test differential expression due to Select the factor to be tested for differential expression.

Comparisons Select groups to be compared. It is possible to choose between "Across groups", "All group pairs", and "Against control group".

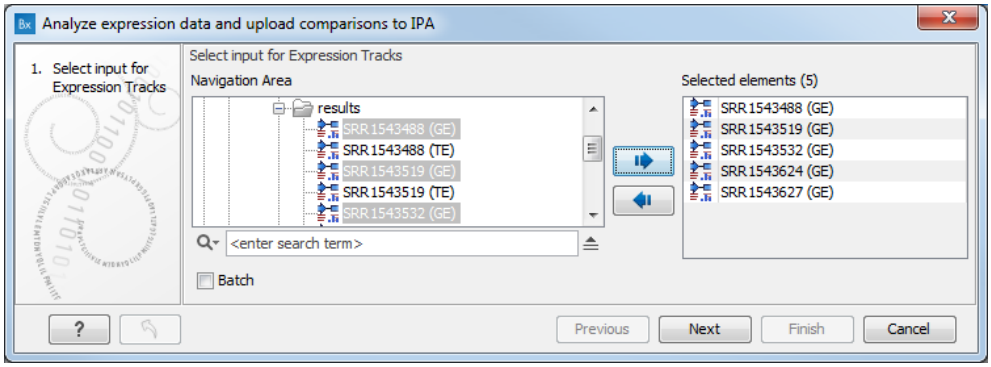


Figure 1.10: *Selecting input parameters in the IPA plugin workflow.*

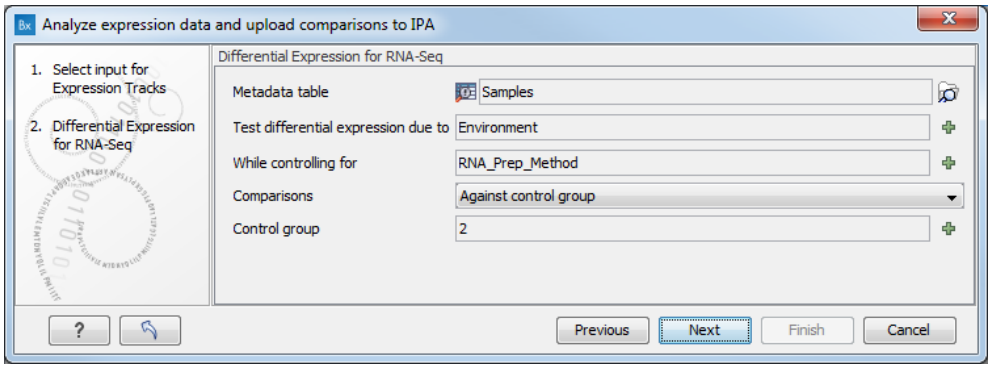
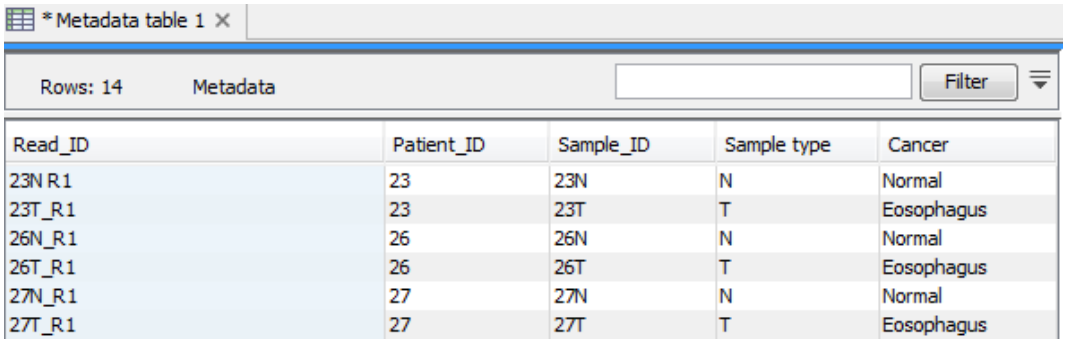


Figure 1.11: *Selecting parameters for Differential Expression for RNA-Seq.*

Control group If "Against control group" was selected in "Comparisons", a control group must be selected.

An example of a metadata table is shown in figure 1.12.



Read_ID	Patient_ID	Sample_ID	Sample type	Cancer
23N R1	23	23N	N	Normal
23T_R1	23	23T	T	Eosophagus
26N_R1	26	26N	N	Normal
26T_R1	26	26T	T	Eosophagus
27N_R1	27	27N	N	Normal
27T_R1	27	27T	T	Eosophagus

Figure 1.12: *An example of a metadata table.*

Metadata is required when defining the experimental design in the Differential Expression for RNA-Seq tool, and can be used to add extra layers of insight in the Create Heat Map and PCA for RNA-Seq tools. To learn more about how to create a metadata table, how to import a metadata table, or how to associate data elements with metadata, see http://resources.qiagenbioinformatics.com/manuals/biomedicalgenomicsworkbench/current/index.php?manual=Metadata.html.

In the next step, the parameters for the Pathway Analysis tool need to be set. Setting the parameters in a workflow context is not as user-friendly as when run stand-alone, because the parameters from the different pages are grouped on the same page in workflow execution. To help clarify this, it is indicated in figure 1.13 which parameters pertain to the IPA upload, and which pertain to the IPA analysis.

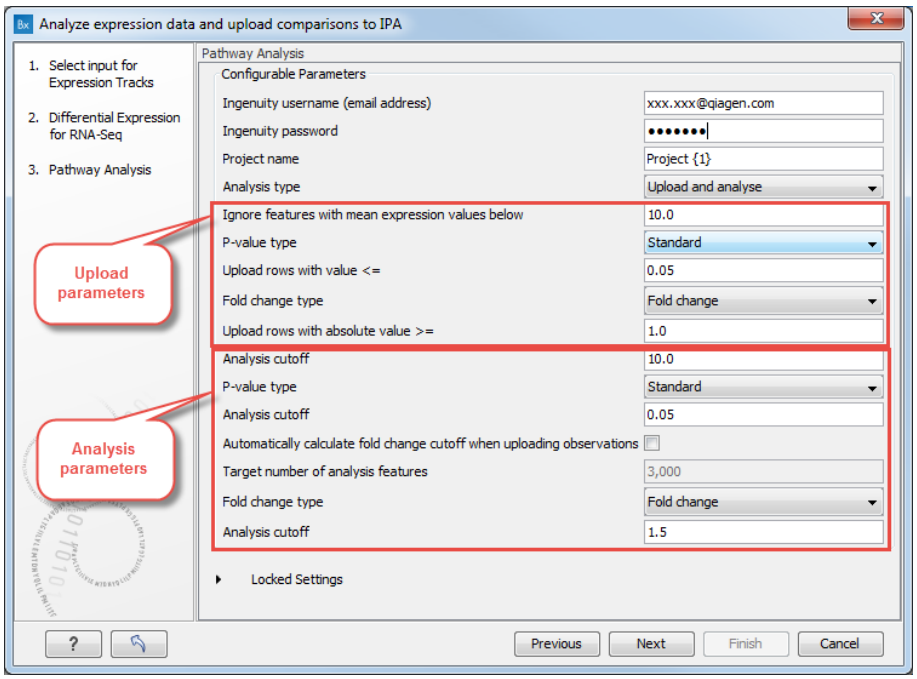


Figure 1.13: *Selecting the Pathway Analysis tool parameters for upload to IPA and analysis in IPA.*

In the final step, standard result handling is performed: The selected parameters can be previewed, and an output location must be chosen.

1.4 Uploading data to IPA using the Pathway Analysis (legacy) tool

Launch the **Pathway Analysis (legacy)** tool from

Legacy | Pathway Analysis

Use an experiment as input (figure 1.14).

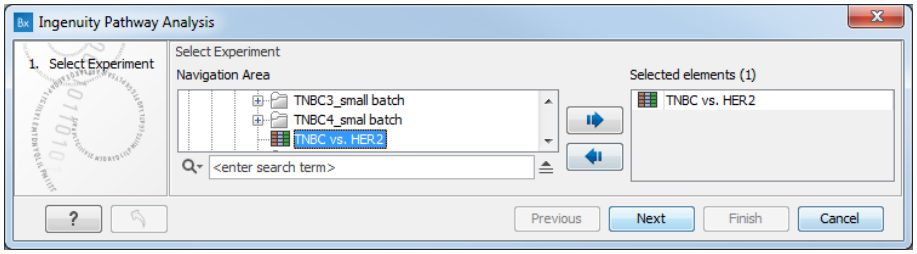


Figure 1.14: *Select an experiment.*

Under Pathway Analysis Setup (figure 1.15), you get the following options:

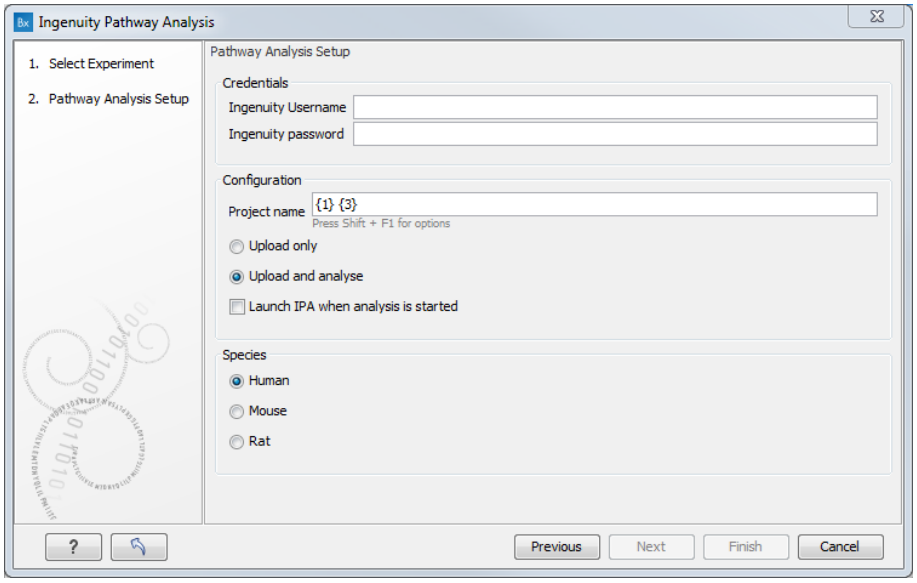


Figure 1.15: *Configure the tool to upload and potentially analyze the statistical comparison data in IPA.*

Credentials Type your Ingenuity account name and password.

Project Name This will be the name of the analysis in IPA once created. There are a few shorthand notations available: {1} will be substituted with the name of the input experiment, {2} is substituted with the type of analysis and {3} is substituted with a date stamp.

Upload only or **Upload and analyze**

Launch IPA Enable if you want to launch IPA directly after the analysis is uploaded.

Species Select the species the experiment is based on. this option is only relevant for gene based experiments.

In the Pathway configuration dialog (figure 1.16), you can configure the pathway analysis.

The wizard allows selection of three values that can be uploaded to IPA. Each of the values has the following three components:

Value The value to upload. The dropdown menu is populated with whatever is available in the experiment, for more information see section 1.1

Cutoff Set the cutoff for the given value. This is needed to limit the amount of genes used in the analysis. For statistical reasons, we recommend that you analyze less than 3000 genes in IPA. Note that the number of features that pass the cutoffs in the Workbench will tend to be larger than the number of total analyzed genes in IPA.

Remaining genes The number of features that pass the cutoff criteria. This is dynamically updated as the cutoff is adjusted.

The last part of the wizard is a summary with information about the sample, but also an overall cutoff summary that shows the number of genes that pass ALL of the cutoff criteria.

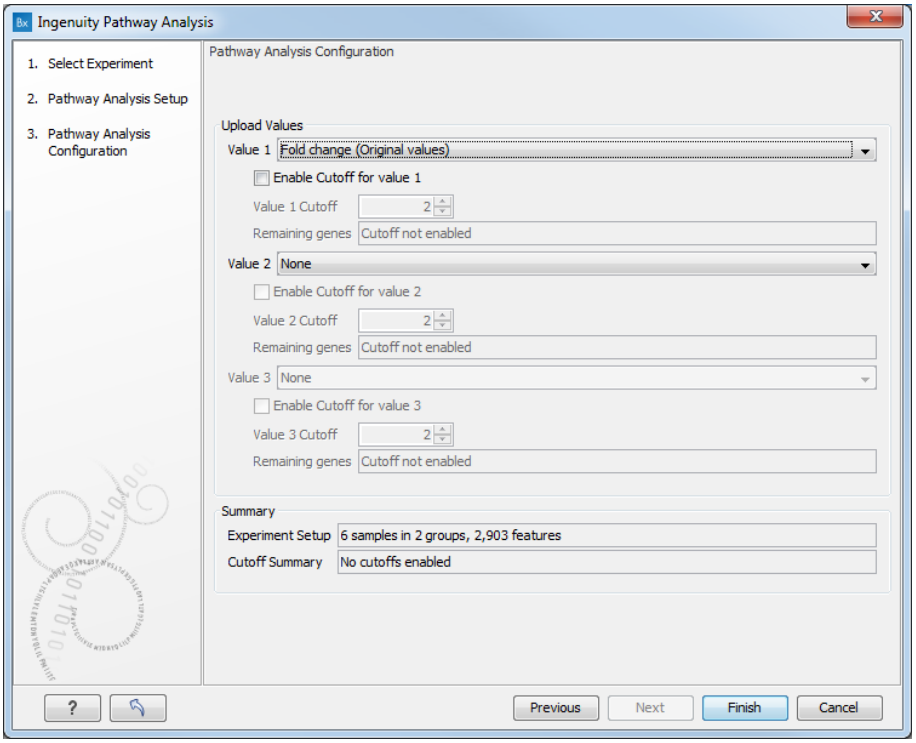


Figure 1.16: *Second step in the IPA wizard*

When the experiment is uploaded, further analysis is performed in IPA. It is possible to launch IPA in several ways:

From the wizard Check to launch IPA automatically in the wizard. When the analysis is uploaded it will launch IPA

From e-mail When an analysis is complete an e-mail is sent with a link

Manually Go directly to the IPA login page `https://analysis.ingenuity.com/pa`.

1.5 Analyzing Data in IPA

Ingenuity Pathway Analysis enables you to analyze and visualize RNA-Seq datasets, eliminating the obstacles between data and biological insight. For example, IPA can predict upstream regulation and downstream outcomes from your expression data, and identify relevant signaling and metabolic pathways. Figure 1.17 shows a newly opened analysis in IPA.

Figure 1.18 shows some of the visual capabilities of IPA.

For more information about how to use IPA please visit `http://www.qiagenbioinformatics.com/products/ingenuity-pathway-analysis/`.

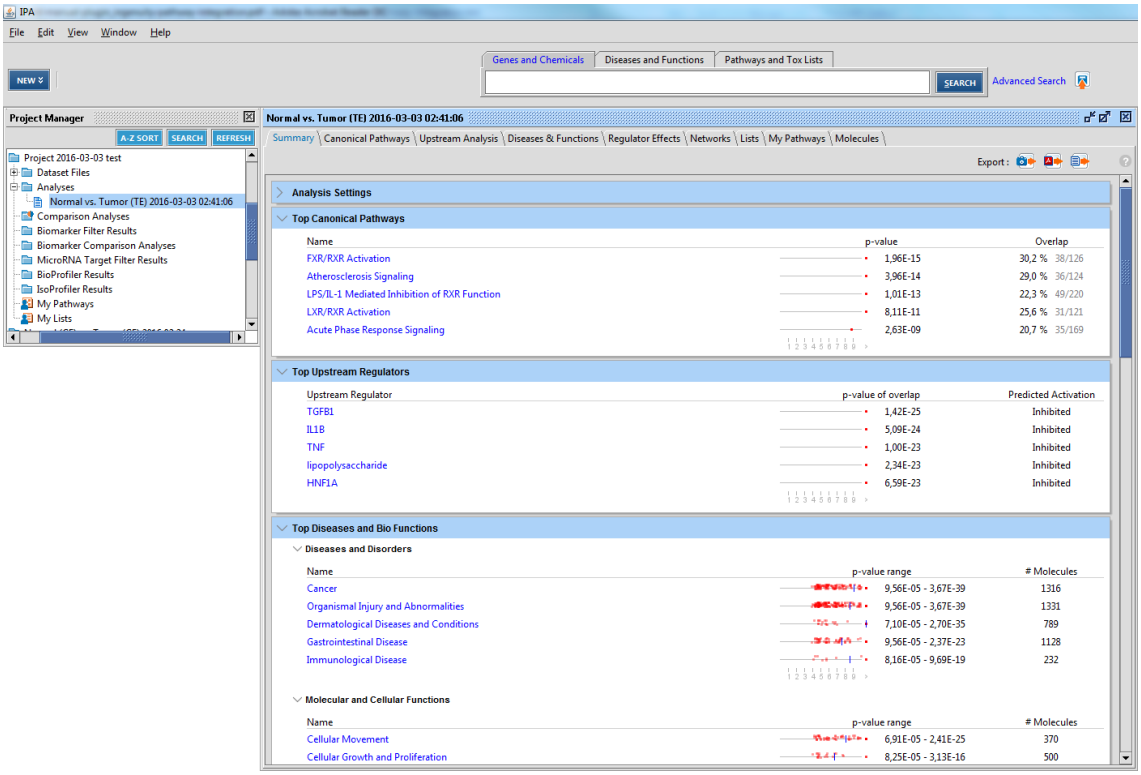


Figure 1.17: *The first window when opening an analysis in IPA*

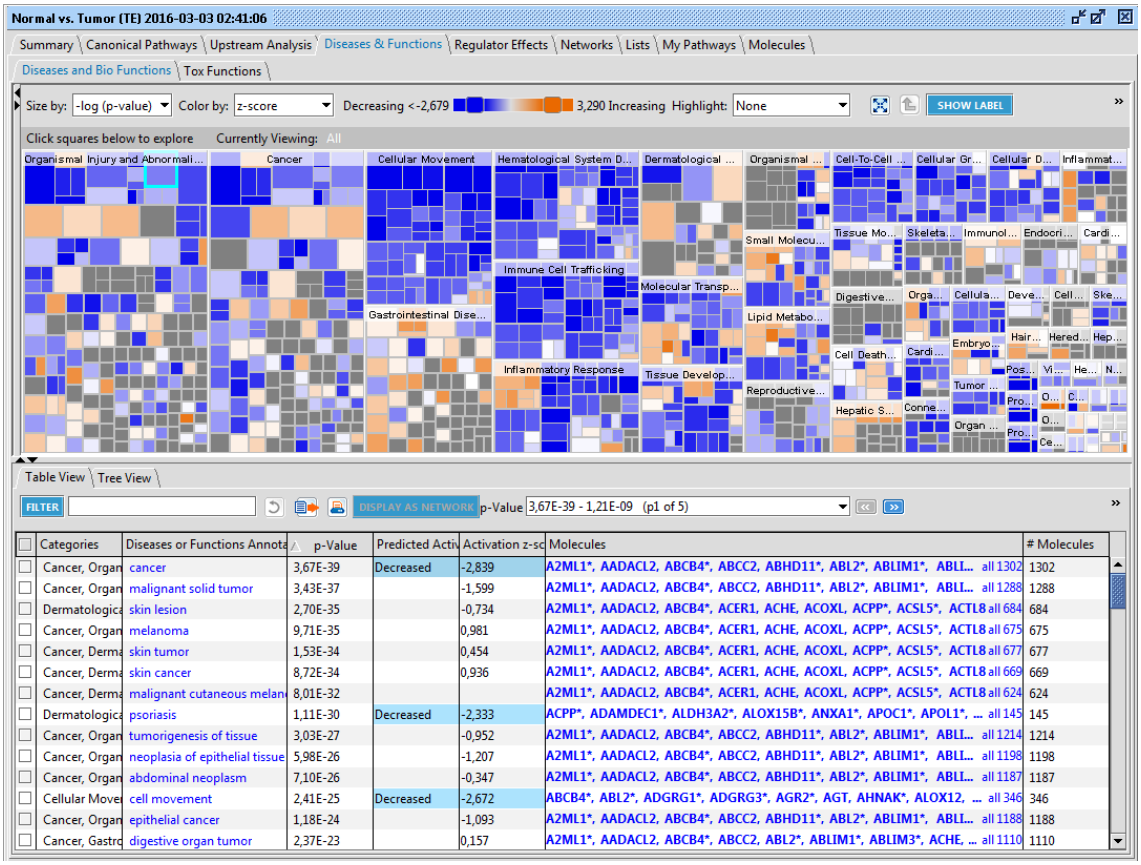


Figure 1.18: *Another IPA application screenshot*

Chapter 2

Install and uninstall plugins

Ingenuity Pathway Analysis 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.

2.1 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 2.1).

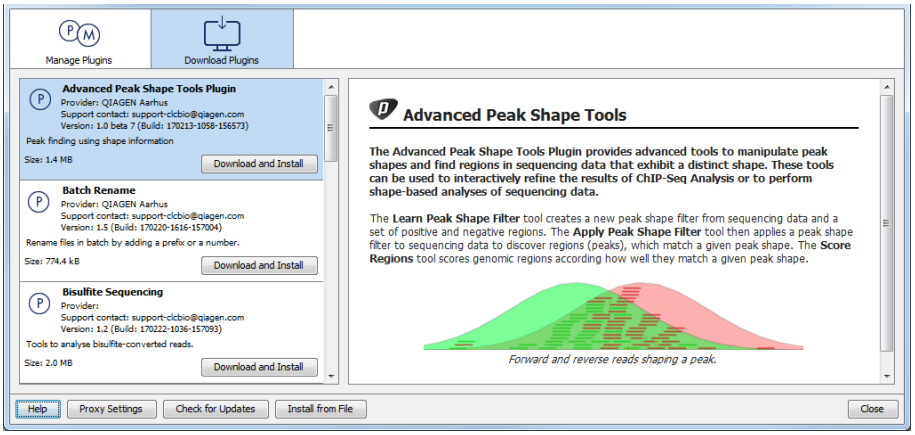


Figure 2.1: *The plugins that are available for download.*

Select Ingenuity Pathway Analysis 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

Part of the installation involves checking and accepting the end user license agreement (EULA) as seen in figure 2.2.

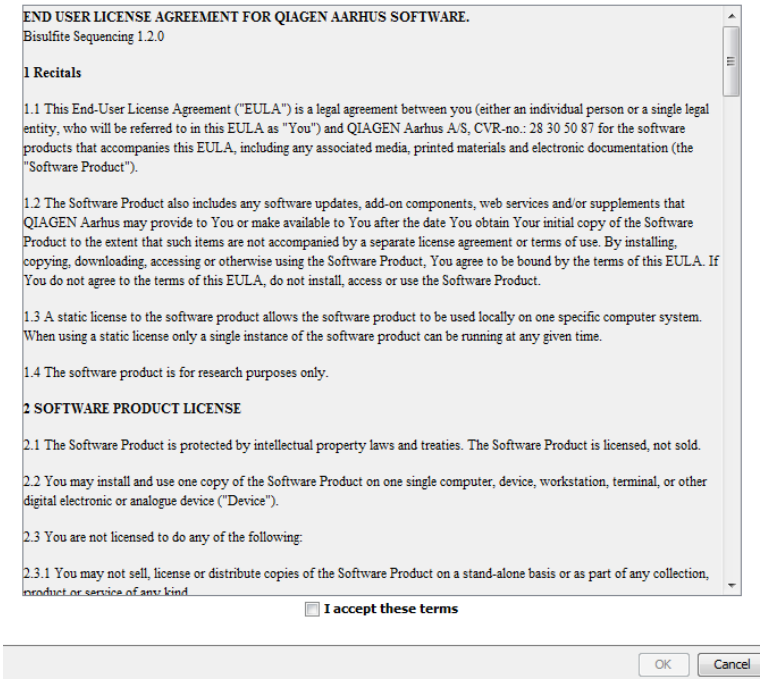


Figure 2.2: *Read the license agreement carefully.*

Please read the EULA text carefully before clicking in the box next to the text **I accept these terms** to accept. If requested, fill in your personal information before clicking **Finish**.

If Ingenuity Pathway Analysis 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.

2.2 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 2.3.

The installed plugins are shown in the **Manage plugins** tab of the plugin manager. To uninstall, select Ingenuity Pathway Analysis 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.

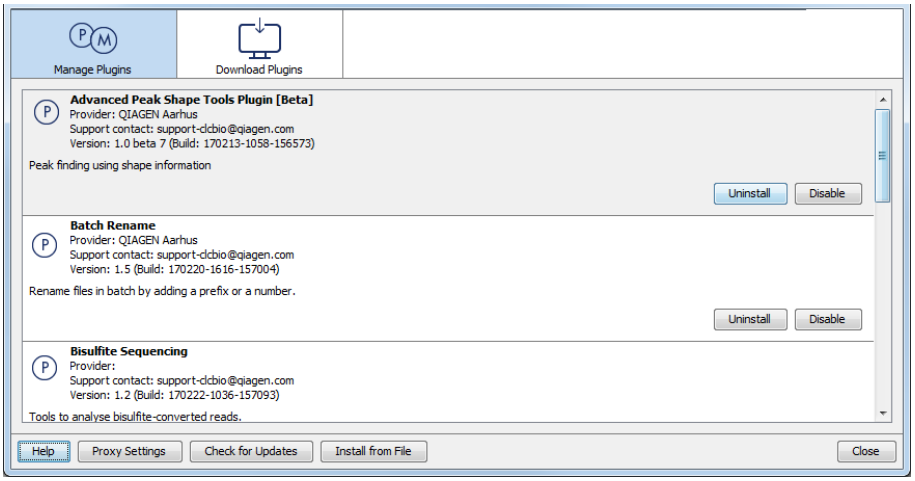


Figure 2.3: *The plugin manager with plugins installed.*

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.