Tutorial

miRNA Quantification and Differential Expression Analysis

December 10, 2025

QIAGEN Aarhus A/S · Kalkværksvej 5, 11. · DK - 8000 Aarhus C · Denmark
digitalinsights.qiagen.com · ts-bioinformatics@qiagen.com

Sample to Insight

miRNA Quantification and Differential Expression Analysis

Introduction

The purpose of this tutorial is to illustrate the capabilities of *CLC Genomics Workbench* and *Biomedical Genomics Analysis* plugin to quantify microRNA (miRNA) expression. We focus on the following:

- Import data.
- Use template workflows to:
  - Quantify miRNA expression.
  - Identify differentially expressed miRNAs.
- Interpret the results.

Data used in this tutorial

This tutorial uses data from "MicroRNA Expression Profiling in Adrenal Myelolipoma" by Decmann et al 2018 (GSE112804, doi:10.1210/jc.2018-00817). The authors investigated miRNA expression in adrenal myelolipoma (AML) and adrenocortical carcinoma (ACC) to identify potential biomarkers that would facilitate diagnosis. AML is a benign primary adrenal neoplasm that is difficult to distinguish from ACC, which has very poor prognosis.

To complete the tutorial in a reasonable amount of time, only 3 AML and 3 ACC samples out of the 30 formalin-fixed, paraffinembedded (FFPE) tissue samples from the study are used here.

Prerequisites

For this tutorial, you must be working with *CLC Genomics Workbench* and *Biomedical Genomics Analysis* plugin 26.0 or higher. Note that versions higher than 26.0 may produce slightly different results than those shown here.

Installing plugins is described in the CLC Genomics Workbench manual.

General tips

- Throughout this tutorial, we provide links to relevant manual pages, which we recommend exploring for additional details.
- Tools and workflows can be found in the Toolbox, but it is often easier to launch them using Quick Launch (), found in the top toolbar (shortcut Ctrl+Shift+T or ⌘ +Shift+T on Mac). Quick Launch displays the full Toolbox path, making it easy to identify the location of the tool or workflow if needed.
- The in-built manual can be accessed by clicking the **Help** button on wizards or by selecting the **Help** option under the **Help** menu.
- Within wizards, the **Reset** button can be used to change settings to their default values.
- Colors and gradients in plots can be changed by clicking on them in the Side Panel.
- Columns in tables can be hidden by unchecking their name in the Side Panel.
- Columns in tables can be used to sort the rows, by successively clicking on the column name until the desired order (indicated by an arrow next to the column name) is achieved.
- Many data elements produced by *CLC Genomics Workbench* tools have multiple views, indicated as icons in the lower left corner of elements opened in the View Area. Clicking on one of the view icons while pressing the Ctrl (⌘ on Mac) key will open in split view such that both views are visible at the same time. Often, if viewing a table and a graphical representation in split view, selecting entries in the table will highlight them in the graphical representation. The order of the views can be changed using drag and drop, see Arrange views in View Area.
- Data can be imported prior to starting a workflow, or it can often be imported on the fly when the workflow is launched.

Import the data

We start by downloading and importing the tutorial data.

1. Download the tutorial data.
2. Start the *CLC Genomics Workbench*.
3. Import the data using Standard Import:
   (a) Launch **Standard Import** ( ) using Quick Launch ( ).
   (b) Locate the tutorial data using the **Add files** button and select **Automatic import** (figure 1).

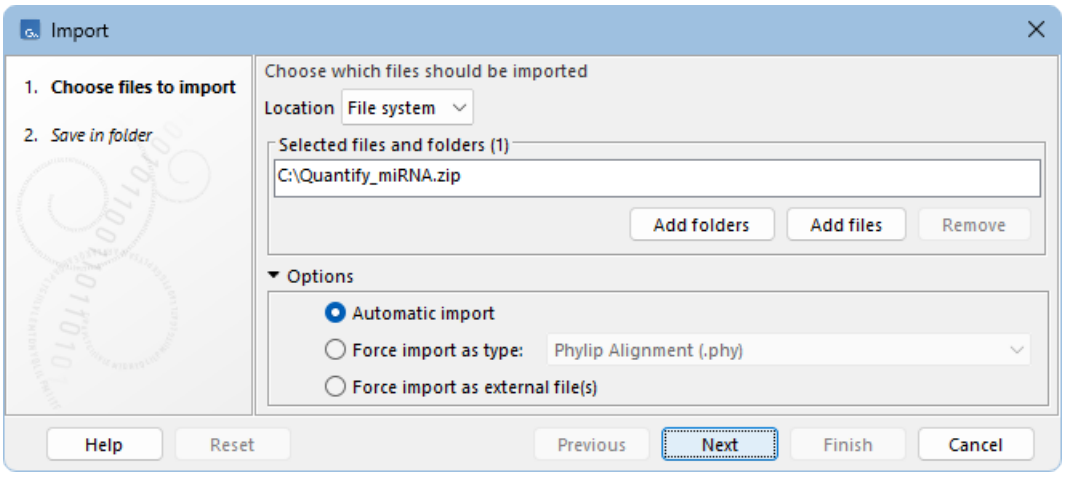


Figure 1: *Standard Import configured to import the tutorial data.*

   (c) In the next step, select a suitable location in the Navigation Area to save the imported data and click on **Finish**.

Once the import is completed, all 6 samples and a metadata table are visible in the Navigation Area (figure 2). Each sample is associated with the corresponding row in the metadata table. See Associating data elements with metadata for information about how to create such associations.

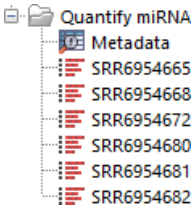


Figure 2: *The "Quantify miRNA" folder in the Navigation Area after importing the tutorial data.*

Quantify miRNA expression

We will now use the Analyze QIAseq miRNA template workflow to analyze the tutorial data. This workflow is designed for data generated using the QIAseq miRNA Library Kit. If you run this workflow on your own data, please note that template workflows are provided as example workflows and may need to be customized to meet the specific requirements of your data.

To see the content of the workflow, locate it in the Toolbox:

Workflows | **Template Workflows** | **Biomedical Workflows** | **QIAseq Sample Analysis** | **QIAseq RNA Workflows** | **Analyze QIAseq miRNA (Illumina)**

right-click on its name and choose **Open Copy of Workflow**.

To run the workflow:

1. Launch the workflow using Quick Launch ().
2. In the **Select Reads** step, specify the data to be analyzed by selecting the 6 samples (figure 3). Make sure to check the **Batch** option below the data selection area.

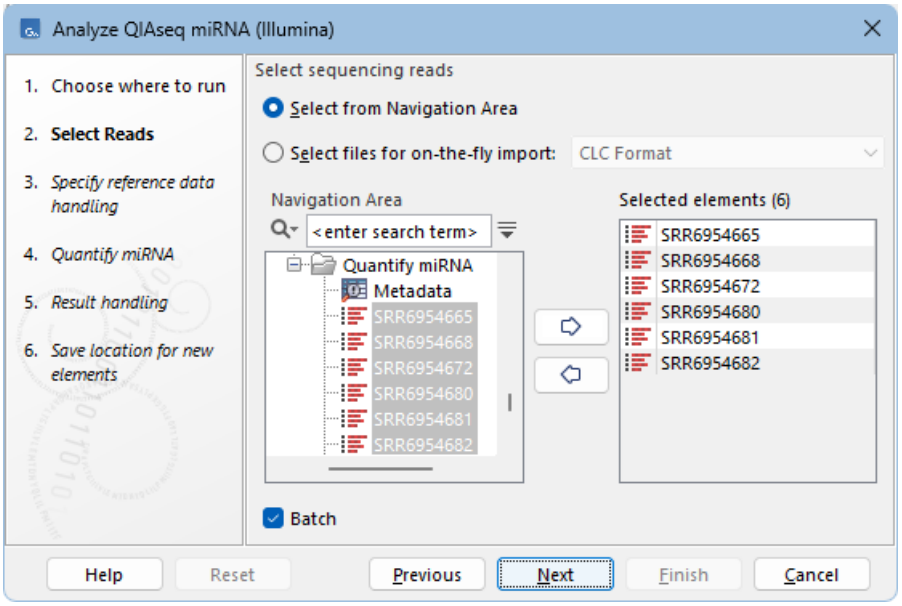


Figure 3: *All 6 samples are used as input. The batch option must be checked.*

3. In the **Specify reference data handling** step, select the "QIAseq Small RNA" Reference Data Set (figure 4). Click on **Download to Workbench** if the data is not already downloaded.
4. In the **Configure batching** step, choose the **Use organization of input data** option. The workflow is then executed once for each sample.
5. The **Batch overview** step shows the organization of the input reads (figure 5).

   If the organization is not as expected, clicking on **Previous** returns to the "Configure batching" step where options can be adjusted.
6. In the "Quantify miRNA" step, keep the default settings.
7. In the **Result handling** step, check **Create subfolders per batch unit**. This will create a separate subfolder for each sample.
8. In the last step, make a new subfolder in "Quantify miRNA" called "Results" and choose to save the workflow results there.

   Click on **Finish**.

   The workflow will now execute. The progress can be monitored under the Processes tab in the Toolbox (figure 6). It will take some time for the workflow to run to completion.

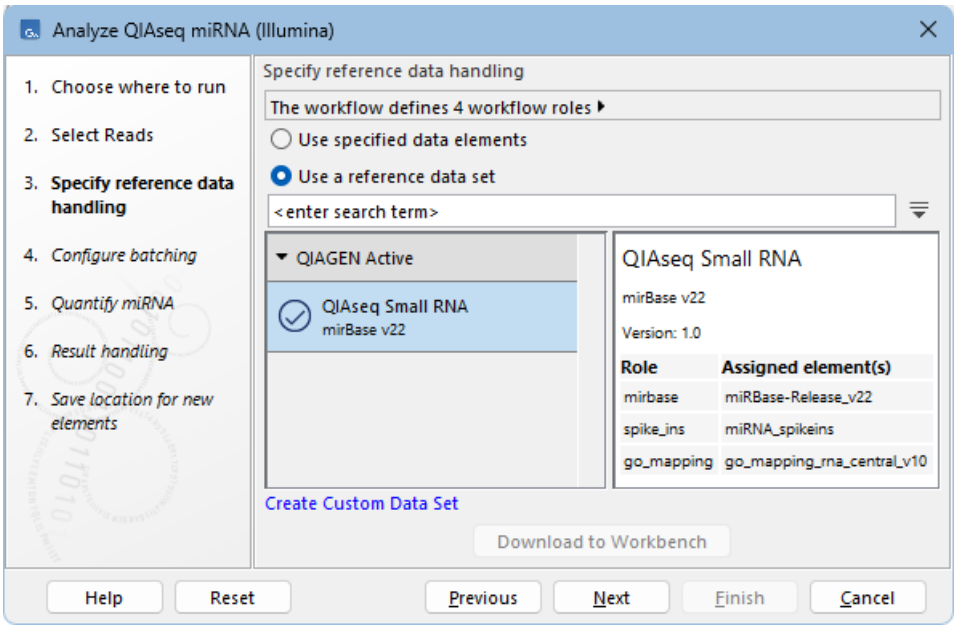


Figure 4: *The "QIAseq Small RNA" Reference Data Set is selected.*

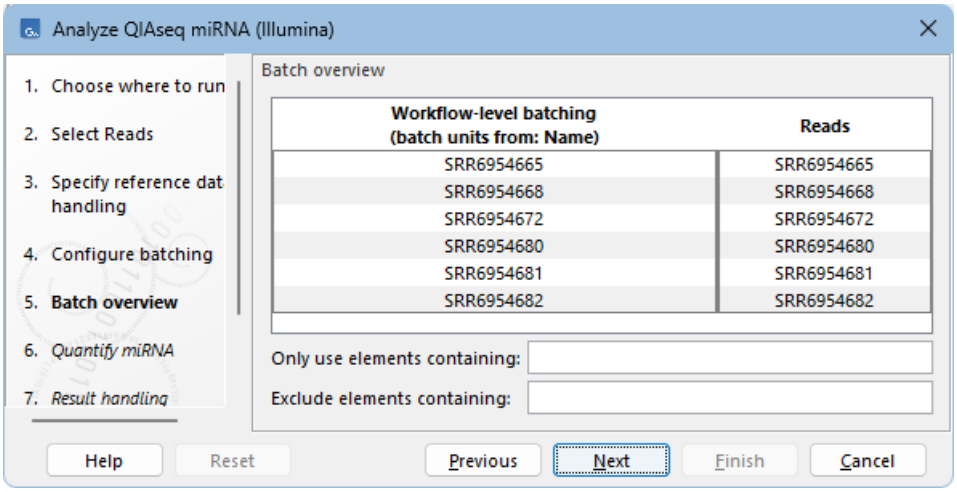


Figure 5: *The Batch overview shows how the input reads are grouped in batch units.*

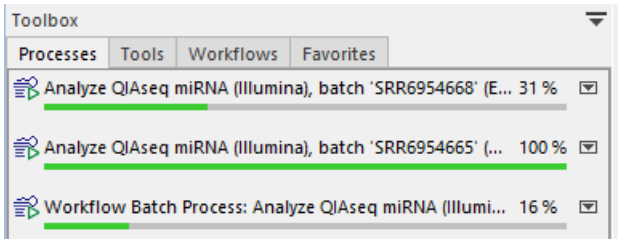


Figure 6: *The "Processes" tab indicates how far the workflow execution has progressed.*

Interpret the quantification results

Results from the workflow are placed in the "Results" folder (figure 7).

The results from each batch unit are saved in a separate subfolder, named after the batch units indicated in the "Batch overview" (figure 5). These folders contain, among others:

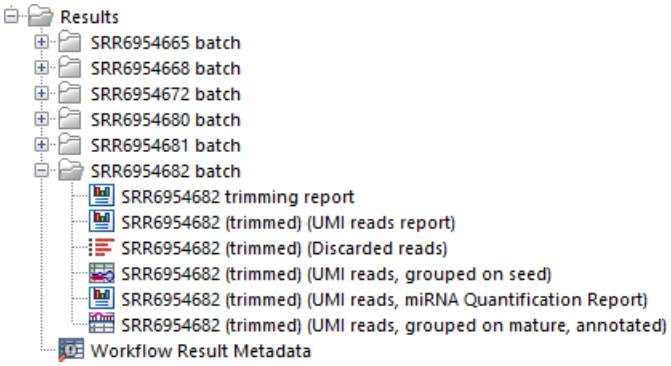


Figure 7: *The "Results" folder in the Navigation Area.*

- Under **QC & Reports**:
  - **UMI reads report**: Produced by Create UMI Reads for miRNA summarizing the identified UMI groups.
  - **Quantify miRNA report**: Produced by Quantify miRNA summarizing the performed mapping and the identified most expressed mature miRNA.
- Under **Supplemental**:
  - The **miRNA Expression (Seeds)** ( ) table: Contains one row for each seed sequence.
- The **miRNA Expression (Mature)** ( ) table: Contains one row for each mature miRNA from miRBase.

  See Quantify miRNA outputs for more details about these tables.

  The tool only quantifies the expression of known miRNAs. See Explore Novel miRNAs for details on how to identify novel miRNAs.

Create a combined report

We will now create a combined report summarizing the information from all samples. This report can be used to quickly review the results.

1. Launch Create Combined miRNA Report using Quick Launch ( ).
2. In the **Select reports** step, right-click on the "Results" folder and choose **Add folder contents (recursively)** to add all reports from the 6 samples.
3. In the **Specify settings** step, uncheck **Use short alias instead of full sample name**.
4. In the **Result handling** step, select **Save**.
5. In the last step, choose to save the results in the "Results" folder and click on **Finish**.

The combined report lists the 20 identified mature miRNAs, seeds, and novel seeds with highest expression across all samples (figure 8). A question mark indicates that the corresponding feature was not identified among the top 20 features in that sample.

5.1 Top 20 mature sequences

Mature	Species	Average %	SRR6954665	SRR6954668	SRR6954672	SRR6954680	SRR6954681	SRR6954682
hsa-miR-16-5p	Homo sapiens	2.11	7,073	3,626	6,511	141,529	292	52,340
hsa-let-7a-5p	Homo sapiens	1.36	15,121	7,259	14,285	42,426	357	23,236
hsa-let-7f-5p	Homo sapiens	1.35	19,274	11,018	9,143	30,643	?	14,240
hsa-miR-21-5p	Homo sapiens	1.11	11,749	6,403	7,831	33,434	?	23,654
hsa-miR-451a	Homo sapiens	1.06	?	?	?	82,735	422	33,722
hsa-miR-26a-5p	Homo sapiens	1.06	11,537	12,025	6,817	23,626	?	11,197
hsa-miR-143-3p	Homo sapiens	0.96	7,111	5,171	10,607	38,728	854	17,389
hsa-miR-29a-3p	Homo sapiens	0.95	10,341	13,044	4,852	16,670	240	7,118
hsa-let-7i-5p	Homo sapiens	0.78	5,242	3,890	6,962	34,185	1,346	13,367
hsa-miR-126-3p	Homo sapiens	0.75	4,256	6,800	3,290	25,634	355	16,192
hsa-miR-125a-5p	Homo sapiens	0.75	8,658	13,170	9,771	?	1,040	?
hsa-let-7b-5p	Homo sapiens	0.53	6,419	?	4,523	19,529	2,053	10,941
hsa-miR-29c-3p	Homo sapiens	0.45	2,869	8,037	3,121	8,021	?	2,925
hsa-miR-30d-5p	Homo sapiens	0.39	6,399	5,886	4,134	?	?	?
hsa-miR-509-3p	Homo sapiens	0.38	8,365	4,135	2,199	?	?	?
hsa-miR-514a-3p	Homo sapiens	0.35	6,376	5,432	?	?	?	?
hsa-miR-125b-5p	Homo sapiens	0.33	3,719	4,070	3,786	?	1,150	4,108
hsa-miR-10b-5p	Homo sapiens	0.30	4,582	4,431	4,153	?	?	?
hsa-miR-127-3p	Homo sapiens	0.28	4,586	4,185	?	?	1,026	?
hsa-miR-508-3p	Homo sapiens	0.27	5,476	3,858	?	?	?	?

? is printed when a feature is not among the top 20 features for the given sample.

Figure 8: *The table of the top 20 mature sequences in the combined miRNA report shows the expression of the most abundant mature miRNA across all samples.*

Identify differentially expressed miRNAs

We will now identify differentially expressed miRNA using the Differential Expression Analysis template workflow.

To see the content of the workflow, locate it in the Toolbox:

> **Workflows** | **Template Workflows** | **Biomedical Workflows** ( ) | **Comparative Analysis** ( ) | **Differential Expression Analysis** ( )

right-click on its name and choose **Open Copy of Workflow**.

To run the workflow:

1. Launch the workflow using Quick Launch ( ).
2. In the **Specify workflow path** step, under "RNA biotype" select "miRNA" (figure 9).
3. In the **Select Expression table or tracks** step, select the **miRNA Expression (Mature)** expression tables (figure 10). Alternatively, the **miRNA Expression (Seeds)** expression tables can be selected, but not a mix of both.
4. In the **Specify reference data handling** step, select the "QIAseq Small RNA" Reference Data Set.

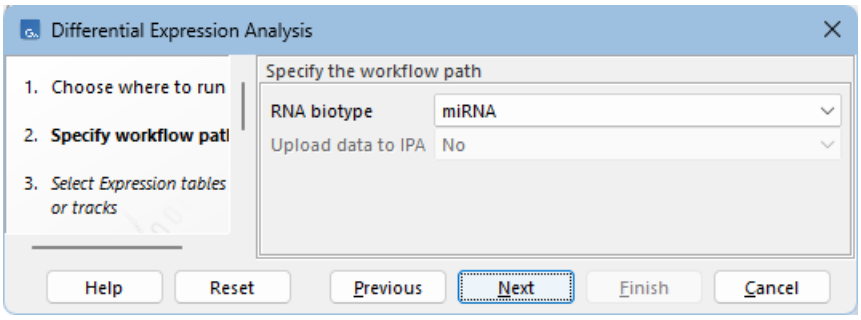


Figure 9: *Selecting workflow path that compares miRNA samples.*

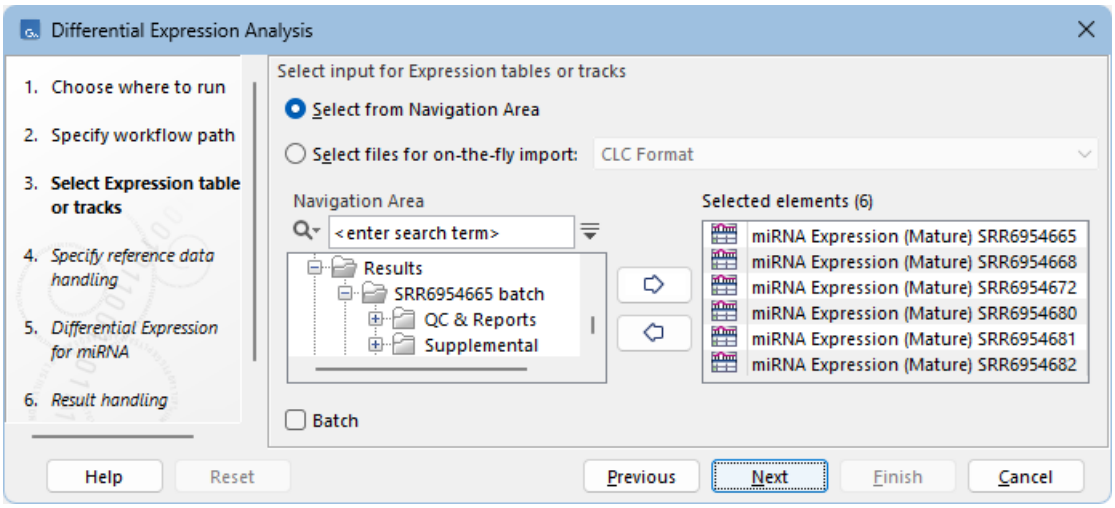


Figure 10: *The grouped on mature expression tables for the 6 samples are used as input.*

5. In the "Differential Expression for miRNA" step (figure 11):

   - Click on the ( ) button at the right-hand side of "Metadata table" and choose the imported "Metadata" element.
   - Set "Test differential expression due to" to **Tumor Type**.
   - Set "Comparisons" to **All group pairs**.

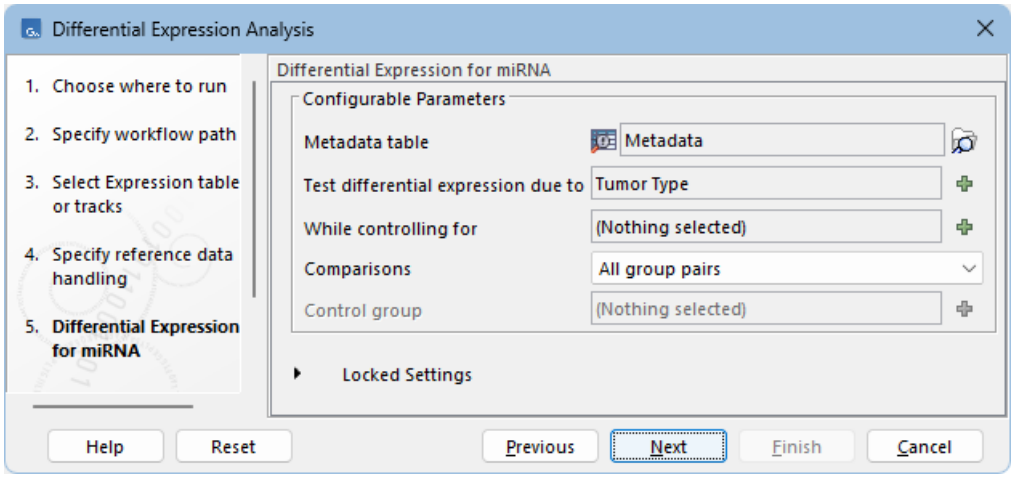


Figure 11: *Differential expression is tested due to "Tumor Type".*

6. In the last step, make a new subfolder in "Quantify miRNA" called "DE results" and choose to save the workflow results there.

   Click on **Finish**.

Interpret the differential expression results

Results from the workflow are placed in the "DE results" folder, which contains, among others, a feature level heat map and an expression browser (figure 12).

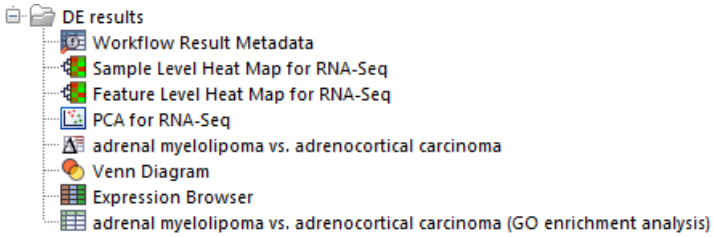


Figure 12: *The "DE results" folder in the Navigation Area.*

Feature level heat map

The feature level heat map provides a visual overview, helping to identify any interesting patterns in the data.

Open the **Feature Level Heat Map for RNA-Seq** (figure 13), found in the "DE results" folder.

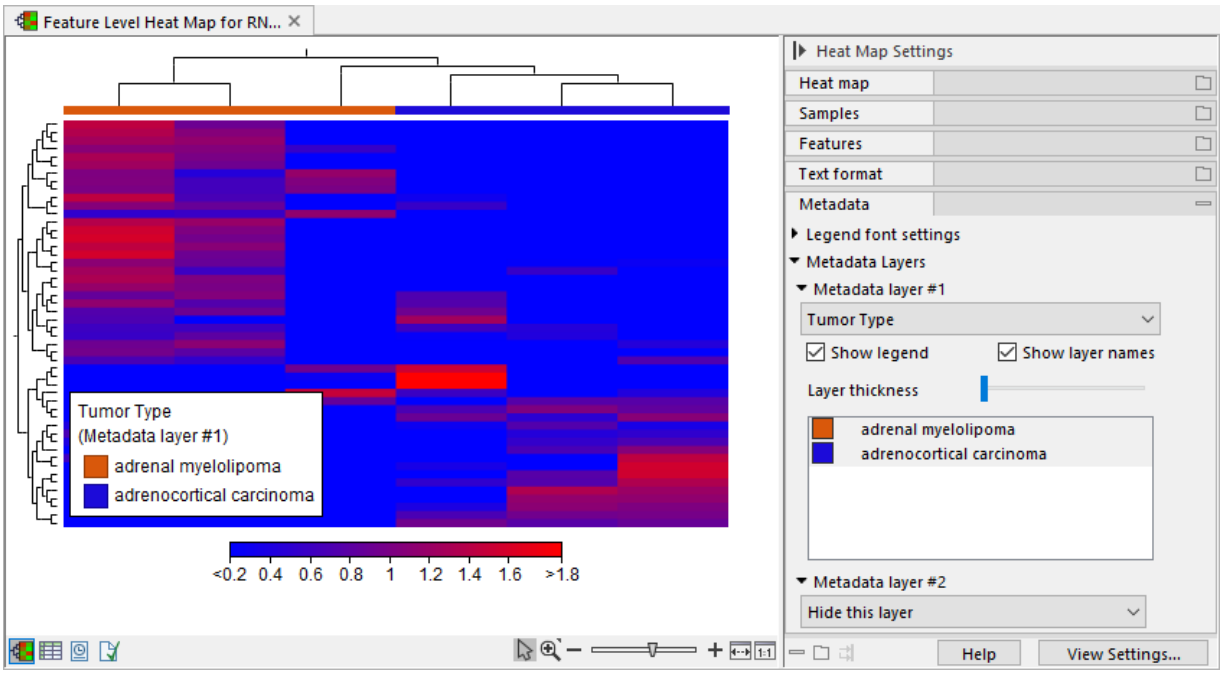


Figure 13: *Heat map with the "Tumor Type" metadata layer. To improve visibility, the samples and features names are hidden.*

Heat map with a "Tumor Type" metadata layer added. To improve visibility, the samples and features names are hidden

The vertical axis of the plot shows the 50 "most interesting" features (mature miRNA), i.e. those

with the largest variance in expression value across the samples. The horizontal axis shows unsupervised clustering of the samples based on these miRNAs. The legend can be moved by clicking and dragging, and the view settings can be adjusted using the Side Panel.

Update the plot as follows:

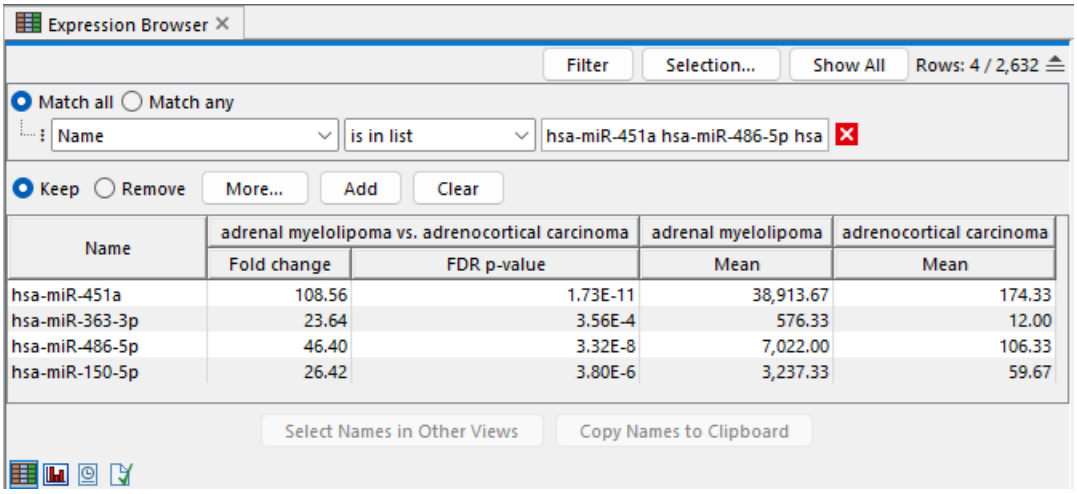
1. In the "Samples" Side Panel palette, uncheck **Show names above**.
2. In the "Features" Side Panel palette, uncheck **Show names left**.
3. In the "Metadata" Side Panel palette, add **Tumor Type** as a metadata layer.
4. Zoom to fit by clicking on the ( ) icon in the bottom right corner.
5. Move the legend to improve visibility.

Interestingly, the heat map shows that one of the AML samples clusters with the ACC samples. This may be due to the fact that the clustering is based only on 50 miRNAs.

Expression browser

The expression browser makes it possible to inspect the miRNA expression levels, as well as annotations and statistics for many samples at the same time.

Open the **Expression Browser** (figure 14), found in the "DE results" folder. Update the table as follows:



Name	adrenal myelolipoma vs. adrenocortical carcinoma		adrenal myelolipoma	adrenocortical carcinoma
	Fold change	FDR p-value	Mean	Mean
hsa-miR-451a	108.56	1.73E-11	38,913.67	174.33
hsa-miR-363-3p	23.64	3.56E-4	576.33	12.00
hsa-miR-486-5p	46.40	3.32E-8	7,022.00	106.33
hsa-miR-150-5p	26.42	3.80E-6	3,237.33	59.67

Figure 14: *Expression browser filtered to show the hsa-miR-451a, hsa-miR-486-5p, hsa-miR-363-3p, and hsa-miR-150-5p miRNAs. Many columns are hidden and the Side Panel is collapsed.*

1. To show only the most relevant information, we hide column tables by using the Side Panel:
   - In the "Feature information" palette, uncheck **Identifier**.
   - In the "Statistical comparison" palette, uncheck **Max group means**, **Log fold change**, **P-value**, and **Bonferroni**.
   - In the "Grouping" palette, uncheck **Show individual expression values**.

2. To show only the miRNAs that the authors in the original study chose for validation as biomarkers:
   (a) Open the advanced filtering, click on the ( ) icon in the top right corner of the table.
   (b) Set filtering by **Name** to be **is in list** and type "hsa-miR-451a hsa-miR-486-5p hsa-miR-363-3p hsa-miR-150-5p" in the text field.
   (c) Click on **Filter**.

The table shows that these 4 miRNAs are significantly overexpressed in AML relative to ACC samples.

See The expression browser and Expression browser plot for more details on how to work with the expression browser, and Working with tables for general information about tables.

Additional analysis

If you wish to analyze the full data set, the data can be downloaded from SRA directly in *CLC Workbench* using Search for Reads in SRA ( ).

To interpret differentially expressed miRNAs in their biological context, we recommend Ingenuity Pathway Analysis (QIAGEN IPA). If you do not have an IPA account, a free trial can be requested via Trial request. Your IPA account credentials can then be used to upload the data from *CLC Workbench* with the Upload to IPA tool.