

QIAseq miRNA Quantification

February 25, 2022

Sample to Insight -

_

QIAGEN Aarhus \cdot Silkeborgvej 2 \cdot Prismet \cdot 8000 Aarhus C \cdot Denmark digitalinsights.qiagen.com \cdot ts-bioinformatics@qiagen.com

QIAseq miRNA Quantification

This tutorial uses the capabilities of CLC Genomics Workbench and the Biomedical Genomics Analysis plugin to quantify miRNA in tumor tissues. We chose to work with a reduced data set from the Decmann et al., 2018 publication, using 6 out of the 30 formalin-fixed, paraffin-embedded (FFPE) archived tissue samples, 3 of adrenal myelolipoma (AML) and 3 of adrenocortical carcinoma (ACC). Adrenal Myelolipoma is a benign primary adrenal neoplasm that is difficult to distinguish from Adrenocortical Carcinoma which has very poor prognosis. The study investigates miRNA expression in AML versus ACC in the hope to find biomarkers that would facilitate diagnosis.

This tutorial covers in just a few steps all the following:

- Import Illumina reads and a metadata table in the Workbench.
- Associate the metadata with the reads.
- Quantify miRNA expression with the QIAseq miRNA Quantification template workflow.
- Visualize results using the **Create Combined Report** tool and the **QIAseq miRNA Differential Expression** workflow.

This tutorial with the reduced data set can be run within 20 min on a regular laptop. Note that the full data set of the study can always be downloaded from the tool **Search for Reads in SRA**.

Prerequisites For this tutorial, you must be working with the CLC Genomics Workbench with Biomedical Genomics Analysis plugin installed. How to install plugins is described here: http://resources.qiagenbioinformatics.com/manuals/clcgenomicsworkbench/current/index.php?manual=Install.html.

Download and import data Go through the following steps to download and import the data into the Workbench.

- 1. Download the sample data from our website: http://resources.qiagenbioinformatics. com/testdata/miRNA_tutorial.zip and save it on your computer.
- 2. Start the workbench.
- 3. Import the reads and the metadata table via the toolbar: **Import** (β) | **Standard Import**.
- 4. Select the miRNA_tutorial.zip file and leave the option to "Automatic import". Click Next.
- 5. Select the location in the Navigation Area of your Workbench where the data should be saved, and click **Finish**.
- Before starting the analysis, the metadata table and the reads need to be associated. Open the miRNA_AML_tutorial table and click on the button Associate Data... highlighted in red in figure 1. Choose the option "Associate Data Automatically..." in the drop-down menu.
- 7. In the next wizard that opens, choose the read you just imported (figure 2) and click **Next**.



Rows: 6	tutorial × Metadata							Filter Ţ
Run Accession	Tumor Type	Scientific Name	Library Strategy	Experime	ent Accession	Project Accession	Sample Accession	Submissio
SRR6954665	adrenocortical carcinoma	Homo sapiens	miRNA-Seg	SRX3897	729	SRP137815	SRS3135064	SRA6833
SRR6954668	adrenocortical carcinoma	Homo sapiens	miRNA-Seq	SRX3897	733	SRP137815	SRS3135067	SRA6833
SRR6954672	adrenocortical carcinoma	Homo sapiens	miRNA-Seg	SRX3897	737	SRP137815	SRS3135070	SRA6833
SRR6954680	adrenal myelolipoma	Homo sapiens	miRNA-Seq	SRX3897	745	SRP137815	SRS3135079	SRA6833
SRR6954681	adrenal myelolipoma	Homo sapiens	miRNA-Seq	SRX3897	746	SRP137815	SRS3135080	SRA6833
SRR6954682	adrenal myelolipoma	Homo sapiens	miRNA-Seq	SRX3897	747	SRP137815	SRS3135081	SRA6833
<	🖉 Satua Tabla	A Magaza Data	Eind Associat	ad Data	Accession		ditional Eiltorian	>
	Set Up Table	Manage Data	D Find Associat	ed Data	Re Associat	le Data	ditional Filtering	
•					🔚 Associ	ate Data Automaticall	ly)	
					associ	ate Data with Row		



Gx	Associate Data with Metadata			>		
1.	1. Select data elements (not folders)	Select data elements (not folders) Navigation Area Selected elements (6) On Contractor and the second secon				
2.	Set data role Association Options	CLC_Data	*	SRR6954665 SRR6954668 SRR6954672		
4.	Result handling			SRR6954680		
	Help Reset	Pr	✓	vious Next Finish Cancel		

Figure 2: Select the reads for association.

8. In the next dialogs, leave the data role to **Sample data** and the matching scheme to **Exact**. You can then click **Finish**. Your reads and the table are now associated, association that is required for the Differential Expression analysis that is run at the end of the tutorial.

Run QIAseq miRNA Quantification

The QIAseq miRNA Quantification workflow quantifies the expression in a sample of the miRNAs found in miRBase. The workflow includes a Trim Reads step, but note that this step only affects lon Torrent reads as Illumina reads do not have the 5' adapter. We are using Illumina reads in this tutorial.

The QIAseq miRNA Quantification template workflow can be found in the Toolbox here:

Template Workflows | Biomedical Workflows () | Small RNA sequencing | QIAseq miRNA Quantification

1. Double-click on the QIAseq miRNA Quantification template workflow to run the analysis.

If you are connected to a CLC Server via your Workbench, you will be asked where you would like to run the analysis.

- 2. Select the sequencing reads that should be analyzed (figure 3). Remember to check the batch option as highlighted in red in the figure before clicking **Next**.
- 3. Select the QIAseq Small RNA Reference Data Set from the Reference manager (figure 4).



Gx QIAseq miRNA Quantifica	ation	×
 Choose where to run Select Reads 	Select sequencing data Select from Navigation Area Select files for import: CLC Format	~
3. Select reference data set	Navigation Area Selected elements (6)	
 Create UMI Reads for miRIVA Quantify miRIVA Result handling Save location for new elements 	Qr <enter search="" term=""> Image: CLC_Data Image: SR6954665 Image: SR6954665 Image: SR6954668 Image: SR6954668 Image: SR6954668 Image: SR6954668 Image: SR6954668 Image: SR6954668 Image: SR6954680 Image: SR6954681 Image: SR6954682 Image: SR6954682</enter>	
Help Reset	Previous Next Einish Ca	ancel

Figure 3: Select the sequencing reads by double-clicking on the file name or by clicking once on the file name and then on the arrow pointing to the right hand side.

You can download the data set to the CLC_References folder of the Navigation Area now if you had not done so before.

🐼 QIAseq miRNA Quantification		×
 Choose where to run Select Reads Select reference data set Configure batching Create UMI Reads for miRNA Quantify miRNA Result handling Save location for new elements 	Select which reference data set to use ○ Use the default reference data ④ Select a reference set to use <enter search="" term=""> ○ Only Downloaded ✓ QIAGEN Active ① QIASeq Small RNA</enter>	
Help Reset	Previous Next Einish Cancel	

Figure 4: Select the QIAseq Small RNA reference data set.

- 4. Use organization of input data for the **Configure batching** and click **Next**. This mean that the workflow be executed once per sample.
- 5. The **Batch overview** shows that each batch unit contain one sample as seen in figure 5.
- Leave all parameters for the Create UMI Reads for miRNA tool as they are set by default (figure 6) and just click Next. These parameters have been configured to work well with



		\wedge	Batch overview					
3.	Select reference data se		Units	Co	nter	nts		
4	Configure batching		SRR6954665	1	F	SRR6954665		
	Configure batching		SRR6954668					
5.	Batch overview		SRR6954672					
6	Create LIMI Reads for		SRR6954680					
	miRNA		SRR6954681					
2	Ourself w DNA		SRR6954682					
<i>'</i> .	Quartury mikrva		Only use elements conta	nina:				
8.	Result handling		Exclude elements conta	nina:				
•	Sava location for now	¥		- 1				6 elements in tota
1	>							

Figure 5: Overview showing sample content of each batch unit.

reads that were sequenced with the Illumina technology. However, note that when working with Ion Torrent reads, we recommend to enable all options and change the "Maximum differences in small RNA sequences" to 2.

Gx	QIAseq miRNA Quanti	fica	tion ×
4. 5. 6. 7. 8.	Configure batching Batch overview Create UMI Reads fo miRNA Quantify miRNA Result handling	^	Create UMI Reads for miRNA Configurable Parameters Allow indels in common sequence Allow indels in UMI Maximum differences in small RNA sequence Allow indels in the small RNA sequence Locked Settings
<	Help Rese	et	Previous Next Einish Cancel

Figure 6: Parameters for Create UMI Reads for miRNA.

- 7. The Quantify miRNA dialog is also pre-configured adequately for the current data (figure 7). Spike-ins are not included in the data but the option to enable them can be left checked. The miRBase database are included in the downloaded reference data set in a previous step and therefore pre-configured here. This can be changed if needed but not for this tutorial. The data we are using is from human samples so only Homo sapiens need to be selected in the Prioritized species list. The last parameter provides the option to add sequence lists with additional smallRNA reference databases e.g. piRNAs, tRNAs, rRNAs, mRNAs, lncRNAs. We will leave this empty for now. Click Next.
- 8. Finally, in the last wizard step, choose to **Save** the results of the workflow in one subfolder per batch unit, and specify the location of your choice in the Navigation Area before clicking **Finish**.



		~	Ouantify miRNA	
1.	Choose where to run		Configurable Parameters	
2.	Select Reads		Enable spike-ins	
3.	Select reference data se		miRBase miRBase-Release_v22	ର୍
4.	Configure batching		Prioritized species Homo sapiens	÷
5.	Batch overview		Database files	Ø
6.	Create UMI Reads for miRNA	~	Locked Settings	
	>			

Figure 7: Parameters for the Quantify miRNA tool.

Output from the QIAseq miRNA Quantification workflow

The QIAseq miRNA Quantification workflow produces a folder containing a series of reports, and two expression tables for each batch unit (figure 8).



Figure 8: Outputs from the QIAseq miRNA Quantification workflow.

In the **Grouped on mature** expression table (P), there is a row for each mature miRNA in the database, including those for which the expression is zero. Double click on a row to open a unique reads alignment. The alignment shows all possible unique reads that have aligned to a specific miRNA from the database. Mismatches to the mature reference are highlighted in the alignment and recapitulated in their name.

The **Grouped on seed** expression table (**E**) has a row for each seed sequenced. This file is useful for further analysis in Ingenuity Pathway Analysis (IPA) where it can be uploaded if the Ingenuity Pathway Analysis plugin is installed and you have an active license for IPA (this however is outside the scope of this tutorial).

Finally, each batch unit will output three reports, including the **Quantification report**. Although it is possible to review each report separately, we can also use the Create Combined miRNA Report tool to compile all of them into one. The tool can be found here:

Tools | RNA-Seq Analysis () | miRNA Analysis () | Create Combined miRNA Report ()

First, select the reports that should be combined (figure 9). We chose here to combine all reports from all batch units to get an overview of UMI statistics, mapping and annotation summary and miRNA quantification all in one report.

In the next dialog we choose to keep the name of our samples as they are and not to substitute them for a shorter alias. We then choose to **Open** the report before clicking **Finish**.





Figure 9: Combined miRNA quantification reports into one.

Several tables summarize the data and compile the top 20 mature sequences, the top 20 seeds and top 20 novel seeds with counts for the different samples. A question mark ? in these tables indicates when a feature is not among the top 20 novel seeds from a particular sample.

Run QIAseq miRNA Differential Expression

To continue our investigation, we will use the **QIAseq miRNA Differential Expression** workflow.

The QIAseq miRNA Differential Expression template workflow can be found in the Toolbox here:

Template Workflows | Biomedical Workflows () | Small RNA sequencing | QIAseq miRNA Differential Expression

1. Double-click on the QIAseq miRNA Differential Expression template workflow to run the analysis.

If you are connected to a CLC Server via your Workbench, you will be asked where you would like to run the analysis.

- 2. Select the expression tables that should be compared (figure 10) before clicking **Next**. You can choose either the **mature** (as we do here) or the **seed** expression tables, but not a mix of both.
- 3. Once again select the QIAseq Small RNA Reference Data Set from the Reference manager (figure 11).
- 4. In the Differential Expression for miRNA dialog, set the parameters as seen in figure 12. The metatable data is the one you associated with your data in the beginning of the tutorial.



Gx QIAseq miRNA Differential	Expression	>
1. Choose where to run	Select input for Expression tables (Mature or Seed)	
2. Select Expression tables (Mature or Seed)	Select files for import: CLC Format	
3. Select reference data set	Navigation Area	Selected elements (6)
 Differential Expression for miRNA Result handling 	Q* (enter search term>) □ □ <td> ■ SRR6954665 (trimmed) (UMI reads, grouped on mature, ■ SRR6954668 (trimmed) (UMI reads, grouped on mature, ■ SRR6954672 (trimmed) (UMI reads, grouped on mature, ■ SRR6954680 (trimmed) (UMI reads, grouped on mature, </td>	 ■ SRR6954665 (trimmed) (UMI reads, grouped on mature, ■ SRR6954668 (trimmed) (UMI reads, grouped on mature, ■ SRR6954672 (trimmed) (UMI reads, grouped on mature, ■ SRR6954680 (trimmed) (UMI reads, grouped on mature,
6. Save location for new elements	SRR6954665 batch SRR6954665 (trimmed) (UMI reads, grouped on seed) SRR6954665 (trimmed) (UMI reads, grouped on mature, annotated) SRR6954668 (trimmed) (UMI reads, grouped on seed) SRR6954668 (trimmed) (UMI reads, grouped on mature, annotated) SRR6954672 batch	SRR6954681 (trimmed) (UMI reads, grouped on mature, SRR6954682 (trimmed) (UMI reads, grouped on mature,
	SRR6954672 (timmed) (UMI reads, grouped on seed) SRR6954672 (timmed) (UMI reads, grouped on mature, annotated) SRR6954680 bath SRR6954680 (timmed) (UMI reads, grouped on seed) SRR6954680 (timmed) (UMI reads, grouped on mature, annotated)	0
	SR6954631 (trimmed) (JMI reads, grouped on seed) SR6954631 (trimmed) (JMI reads, grouped on nature, amotated) SR6954682 tath SR6954682 tath SR6954682 (trimmed) (JMI reads, grouped on seed) SR6954682 (trimmed) (JMI reads, grouped on nature, amotated)	
	Batch	
Help Reset		Previous Next Finish Cancel

Figure 10: Select the mature or seed expression tables.

🐼 QIAseq miRNA Differential Ex	pression	\times
Choose where to run Select Expression tables (Mature or Seed)	Select which reference data set to use Use the default reference data Select a reference set to use	
3. Select reference data set	<enter search="" term=""> Only Downloaded</enter>	
4. Differential Expression for miRNA	← hg19 (Ensembl) Ensembl v99, dbSNP v151, ClinVar 20210828	
5. Result handling	hg 19 (Refseq)	
6. Save location for new	✓ RefSeq GRCh37.p13, dbSNP v151, ClinVar 20210828	
elements	QLAseq Small RNA mirBase v22 The following types of reference data are used and must be supplied by the data set:	
10	Gene_ontology	
Ce.	House Ensembl v66	
90000	Single Cell Mouse (Ensembl) Cell Type Classfler v1.1, Ensembl v86, IMGT reference sequences, Peak Shape Filter v1	
The second	Download to Workbench	
Help Reset	Previous <u>N</u> ext ⊟nish <u>C</u> ancel	

Figure 11: Select the QIAseq Small RNA reference data set.

- 5. Choose to save the results (shown in figure 13) in the tutorial folder.
- 6. Open the Heat Map for RNA-Seq output (figure 14). The AML and ACC samples are separated, indicating that it may be possible to find miRNA that would be specific to the AML and that could be used for diagnostics.
- 7. Open the Expression browser (figure 15). In this table, the expressed miRNAs that were chosen in the original paper for validation have been highlighted. In our tutorial data, hsa-miR-451a, hsa-miR-486-5p, hsa-miR-363-3p, and hsa-miR-150-5p show significant overexpression in AML relative to ACC as well.



1 Choose where to run	Differential Expression for miRNA			
1. Choose where to run	Configurable Parameters			
 Select Expression tables (Mature or Seed) 	Metadata table	miRNA_AML_tutorial		
3. Select reference data set	Test differential expression du	ue to Tumor Type	+	
Differential Evenession	While controlling for	(Nothing selected)	4-	
for miRNA	Comparisons	All group pairs	\sim	
5. Result handling	Control group	(Nothing selected)	ella	
5. Save location for new	Locked Se Gx Select M	etadata		×
elements				
elements Help Reset	Navigation	Area Reference Data Sele	ected elements (1)	
elements Help Reset	Navigation	Area Reference Data Sek	ected elements (1)	
elements Help Reset	Navigation Qr <ente C</ente 	Area Reference Data Sel r search term> Data mRNA Quantification Tutorial	ected elements (1)	

Figure 12: Parameters for the differential expression analysis.









The Quantify miRNA tool only finds and quantifies the expression of known miRNAs. If you want to dig deeper into how to explore novel miRNAs please have a look here https://resources.giagenbioinformatics.com/manuals/clcgenomicsworkbench/current/index.php?manual=Explore_Novel_miRNAs.html.



Rows: 2,632	Expression Bro	wser			Filter to Se	ection					Filter	
			adrenal	myelolipoma vs. ad	a vs. adrenocortical carcinoma adrenal myelolipoma							
Name	Identifier	Max group m	Fold change /	Log fold cha	P-value	FDR p-value	Bonferroni	SRR6954680	SRR6954681	SRR6954682	Mean	SRR
								Total counts	Total counts	Total counts		To
38-MIR-191-50		0.07	21.67	4.44	0.01	0.24	1.00	6.00	6.00	0.00	0.07	
annik-12550-5p		576.67	22.90	4.57	1 705 5	1 745 2	1.00	1 207 00	12.00	410.00	2.07	
armik-202-20		5/0.07	23.72	4.57	1.795-5	1.762-5	1.00	1,307.00	15.00	410.00	0.67	
a-miR-10-10-1-5p		0.67	23.87	4.59	0.10	1.00	1.00	0.00	2.00	0.00	0.67	
зантик-энэгэр		2 240 67	25.67	4,30	0.15	1.00	2 555-4	7 225 00	62.00	2 425 00	2 240 67	
ca.miD-4732-50		3,240.07	20.32	4.92	0.11	1.00	1.00	7,235.00	0.00	2,-123.00	3,240.07	
sa-miR-144-50		809.33	30.55	4.02	2 165-5	1.00 1.06E-3	0.06	2 133 00	3.00	559.00	808.33	
sa-miR-106a-50		3.67	31.01	4.95	0.10	1.00	1.00	2,135.00	0.00	535.00	3.67	
a-miR-100a-5p		5.67	35.88	5.17	0.10	1.00	1.00	16.00	0.00	1.00	5.67	
a-miD-674-50		4.67	39.06	5.20	0.05	1.00	1.00	11.00	0.00	3.00	4.67	
sa mile 02 1 5p		8.67	42.05	5.30	9.60F-4	0.04	1.00	3.00	21.00	2.00	8.67	
a-miP-675-50		1 33	42.00	5.40	0.09	1.00	1.00	1.00	3.00	0.00	1 33	
ca.miR.551a		1.53	42.51	5.41	0.05	1.00	1.00	2.00	3.00	0.00	1.53	
ca.miR.7976		6.33	42 72	5.42	0.08	1.00	1.00	17.00	0.00	2.00	6.33	
sa-miR-150-3n		5.33	44,45	5.47	0.07	0.95	1.00	12.00	0.00	4.00	5.33	
sa-miR-486-50		7.027.00	46.50	5.54	3.42E-10	1.80E-7	9.00E-7	15,994.00	1.314.00	3.773.00	7.027.00	
sa-miR-144-3n		134.33	46.82	5,55	3.24E-5	2.44E-3	0.09	301.00	1.00	101.00	134.33	
sa-miR-548k		6.33	56.49	5.82	0.06	0.80	1.00	7.00	1.00	11.00	6.33	
sa-miR-551b-3p		7.33	60.69	5.92	0.05	0.75	1.00	12.00	1.00	9.00	7.33	
sa-miR-582-3p		9.00	71.43	6.16	0.04	0.65	1.00	21.00	1.00	5.00	9.00	
sa-miR-25-5p		5.00	102.41	6.68	0.03	0.56	1.00	7.00	6.00	2.00	5.00	
sa-miR-451a		38,959.67	108.98	6.77	3.69E-14	9.70E-11	9.70E-11	82.735.00	422.00	33.722.00	38,959.67	
sa-miR-4306		17.00	138.58	7.11	0.02	0.40	1.00	37.00	0.00	14.00	17.00	
				[

Figure 15: Expression Browser.



Bibliography

[Decmann et al., 2018] Decmann, A., Perge, P., Nyiro, G., Darvasi, O., Liki; ¹/₂, I., Borka, K., Micsik, T., Ti; ¹/₂th, Z., Bancos, I., Pezzani, R., Iacobone, M., Pati; ¹/₂cs, A., and Igaz, P. (2018). Microrna expression profiling in adrenal myelolipoma. *Journal of Clinical Endocrinology and Metabolism*, 103(9):3522–3530.