Exiqon Array Software Manual

Quick guide to data extraction from miRCURY LNA™ microRNA Arrays

March 2010

EXIQON
Seek Find Verify

Table of contents

Introduction

Overview

Exiqon offers a software package for the analysis of Exiqon miRCURY LNA™ microRNA arrays. The package consists of two programs: ImaGene® 9 and Nexus Expression™ 2. Together the two programs will help convert tif images of the scanned slides into useful information by normalization and identification of differentially expressed miRNA. First step is to extract signal intensities using ImaGene®. When this has been done for the entire experiment, the data can be loaded into Nexus Expression™ where Normalization and filtering of the data is performed.

Nexus Expression™ and ImaGene® are registered trademarks of BioDiscovery, Inc.

ImaGene®

ImaGene® is an image analysis software. It places a grid on top of your scanned array image, and thereby identifies which probe is located in each spot. Subsequently, it quantifies how much Hy3 and Hy5 fluorescent dye is present in each spot (i.e. the signal), how high the surrounding background is, and flags potential outliers. ImaGene® can also normalize your data, but we recommend using Nexus Expression™ for normalization instead to get all slides normalized at the same time. ImaGene® takes one array at a time for analysis, while Nexus Expression™ reads a whole experiment composed of many arrays at the time.

Before you start your analysis

Before starting, make sure your tif image is in the correct orientation: Two Hy3 landing lights should be positioned at the bottom right corner of each subgrid. This can be done by the ImaGene® software.
You need a GAL-file and a settings file for the analysis. GAL-files are specific for each version of arrays, each species and sometimes for each lot of arrays. The appropriate GAL-file can be downloaded from: exiqon.com/Gal-downloads. A settings file can be downloaded from exiqon.com. Save both of these files on your computer before you start. Start the ImaGene® software. Load settings by clicking the "settings" option under the "file" menu. Then use the "load" button at the bottom of the "ImaGene® parameter settings" dialogue, browse to find the settings file on your computer and click "open". You are now using the Exiqon recommended settings for quality flagging, spike-in controls etc in ImaGene® (this is only necessary the first time the program is started). The default image resolution is set to 10 µm. If you are using another image resolution, click the "Image preferences" tab and change to your resolution. Then save the settings file by clicking "save as".

Analyze one array at the time

Open the tif-file by clicking the "Images" button, browse to find your image(s), select it/them and click "open". If you have single color images, load one image. If you have dual color images, you either load one Hy3 and one Hy5 image together (from the same array), or load the dual color image containing both dyes. The format depends on the scanner you use. The images should be oriented so that there are two adjacent Hy3 landing lights in the bottom right corner of each subgrid. If the images need reorientation, right-click

on the image file names in the "Images" panel and use the "Rotate" or "Flip Horizontal" options to obtain the correct orientation. For example, if the images were scanned using an Agilent scanner choose: "Rotate 270 degrees" and then "Flip Horizontal".

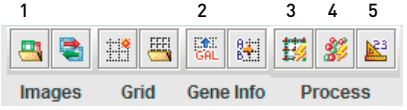
Click the "Import grid from GAL" button, and browse to find your saved lot specific GAL-file, select "create as one meta field" and click "open".

Click the "Auto adjust grid" button. Make a manual check that all subgrids are aligned properly (landing lights are in the corners of each subgrid, and two in the bottom right corner). You may want to zoom in to get a better view of the image while doing this. Click the "zoom" button and zoom in using the left mouse button to select an area. If a subgrid is misaligned, manually adjust by clicking the "adjust subgrid" button and drag the misplaced one into place. When all subgrids are properly aligned, make sure that no single subgrid is selected (=all subgrids have color purple), and click the "auto adjust spots" button.
Finally, click the "make measurements" button.

The grid can also be placed manually by using the adjust metagrid/adjust subgrid buttons, and then drag the grid in place.

Change view by clicking the "raw data" tab. Save your results by clicking the "save" button, select the folder in which you want your results saved (do not open the folder, just select it), mark "Separate files for each channel" and click "save". The raw data has now been saved in one .sst file per array and one or two .txt files (depending on if you have single or dual channel arrays). The .txt files are the ones you will use in Nexus for normalization and further analysis (if running dual color, tick "seperate files for each channel").

Figure 1



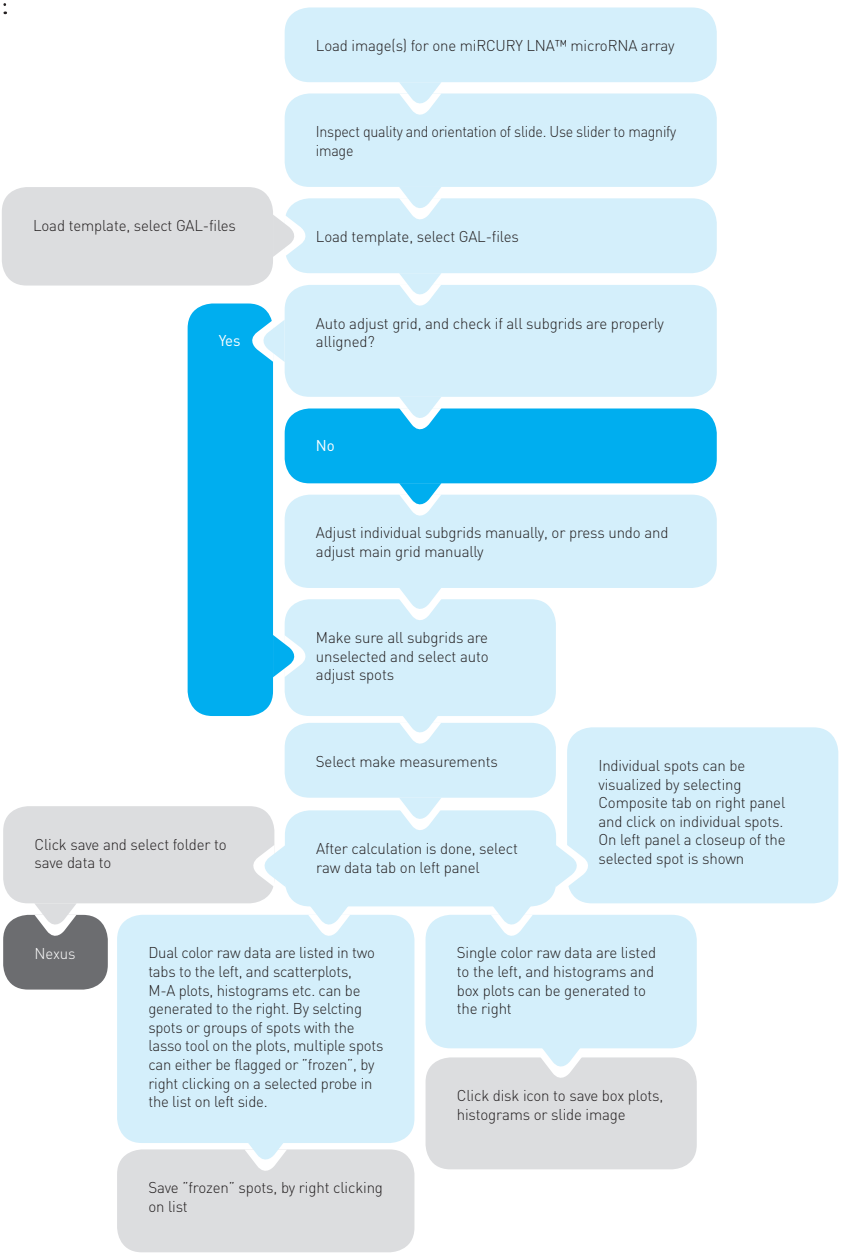
Figur 1. Panel of buttons for extracting signals. 1: Load images, 2: Gene info button for loading the GAL-files, 3: Auto adjust grid, 4: Auto adjust spots, 5: Make measurements.

The plots available under the "plots" tab contain lots of information about your array. Box-plots can be used to study the signal and background level in each subgrid, one channel at a time. Smears and gradients on the arrays can be discovered this way. Scatter plots in a log scale can be used to study for the signal distribution of synthetic spike-ins and the spread between the two samples labeled in different dyes. With the lasso tool, groups of spots can be highlighted and the selected spots will appear in the list to the left and it is possible to save them in a separate file by right clicking the list. Each plot can be saved as an image with the "save" button on top of the plots.

To analyze your next image, click "close" at the bottom of the "raw data" tab. A dialogue will ask if you really want to close, click "yes". Then switch over to the "main" tab and click the "clear all" button at the bottom of the page. The previous image will be removed, and you can now open the image(s) from the next array.

When all the slides for an experiment have been processed in ImaGene®, the data can be gathered and processed using Nexus Expression™.

Pipeline for data extraction in ImaGene®:



Nexus Expression™ quick guide

Nexus Expression™ is an array data analysis software. It reads in the results files (.txt files) created by ImaGene®, and can thereafter subtract background from signal, normalize and visualize your data. It can combine the replicated measures of each probe on an array into one output value per array and make statistical calculations of differentially expressed miRNAs. One experiment, for example treated and control samples of one tissue type, can be analyzed at a time (max 24 samples with the 30-day license).

Loading miRCURY LNA™ microRNA Arrays and preprocessing data

When the software opens, chose "create new project", type in a project name and select a file location where it should be created. A project folder with some subfolders is then automatically created in the selected location. It is a good idea to copy the ImaGene® results files (.txt files) to this project folder (don't place them in the subfolders) or to create the Nexus project folder at the same location as these files. The next time you open the Nexus software, this experiment can be loaded by selecting "open existing project" and selecting the project folder before clicking "open".

In the pre-installed array type folder (default location C:\Program Files\BioDiscovery\Nexus Expression\ArrayTypes\Exiqon), there is a folder for each miRCURY miRCURY LNA™ microRNA Array. In this folder, you find 3 different files needed for your Nexus analysis.

- Controls.txt contains information about the spike-in controls on the arrays. These will be excluded from normalization and analysis and will not be displayed later in Nexus probe list. If you wish to include the spike-ins in your analysis, simply delete all the probe ID's except 13138 in the document, and save the file. (You may wish to keep a copy of the original controls.txt file under a different file name for future use). Probe ID 13138 should always be kept in the control.txt file, because this probe is a dye marker and it should never be included in any normalization procedure.
- Annotation.txt contains information about the probe ID and the associated miRBase annotated miRNA name. It is in principle a GAL-files, only with slightly different layout and headers. Nexus Expression™ uses this file to annotate the probes, it does not read the annotation information from the ImaGene® output file. For other species, array versions or miRBase versions than the ones pre-installed, you can create your own annotation file from

any GAL-files you may have, just follow the instructions in tip 1. Or contact Exiqon at www.exiqon.com/contact for assistance. Official array releases will have annotation files for download on our website.

- Idbases.txt contains URL links to miRBase and other databases of interest for miRNA research.

There are two ways to upload your arrays: through a description file (please see tip 2, the manual or Nexus Expression™ quick guide for instructions that), or through direct upload of each file. If you upload each file directly, the experiment and control files should be in the same order for dual channel arrays (experiment array #1 is the Hy3 channel and control array#1 is the Hy5 channel from the same array). Click the "load files" button, chose Data type "ImaGene" or "ImaGene single channel", chose manufacturer "Exiqon" and the array type you have used. Upload the ImaGene® output files (.txt files) for dual channel arrays by clicking "add experiment" and select the Hy3 output files. Then click the "add control" button to upload the Hy5 arrays. This way of uploading array results can be used also for experimental setups that are not using an experiment-control type of setup. Just upload the Hy3 data as experiments and Hy5 data as controls, and the samples can be divided into experiment groups later.

When all samples are uploaded, they can be viewed in the array list. Before processing the slides, make sure the appropriate preprocessing settings are chosen. Open the menu under "file" and then "Preprocessing Settings". Exiqon standard setting files can be downloaded from www.exiqon.com/mirna-array-software.

The recommended settings differ between single and dual color arrays, and therefore there is one single color settings file and one dual color settings file to choose from. Upload the .xml settings file appropriate for your experiment by clicking "load" at the bottom of the preprocessing settings menu and selecting the file.

Input Data Type is either ImaGene or ImaGene single channel depending on if you have dual or single channel data.

We recommend using the settings shown in Figure 2 for preprocessing dual color data:

With these settings, the mean signal intensity is used, and the median local background is subtracted. All spots that are negative, poor, empty or manually flagged in ImaGene® are excluded from further analysis. For dual channel array data, a lowess normalization is performed. Replicated probes on each array are combined into one output value using the median value, as the median is less affected by outlier spots than the mean value may be.

For preprocessing single color data, we recommend using the settings shown in Figure 3:

Figure 2

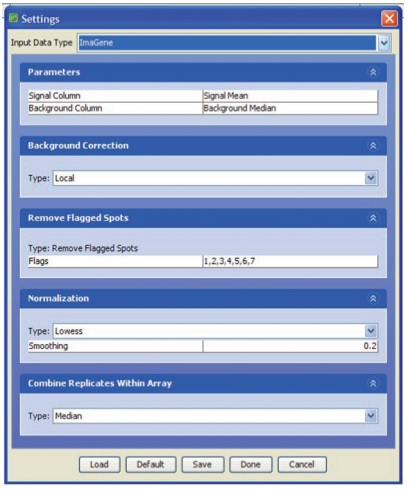


Figure 2: Dual color settings file, Nexus.

With these settings, the mean signal intensity is used, and the median local background is subtracted. All spots that are negative, poor, empty or manually flagged in ImaGene® are excluded from further analysis. For single channel array data, quantile normalization is performed. Replicated probes on each array are combined into one output value using the median value, as the median is less affected by outlier spots than the mean value may be. Please note that only arrays belonging to the same experiment should be loaded and normalized together using the quantile algorithm, as the data from each array affects the other arrays when applying quantile normalization. If any settings have been changed, press the "save" or "done" button. Once the preprocessing settings have been chosen, click the "process" button in the "Arrays" tab. Processing may take some time depending on the number of arrays in the experiment.

Figure 3

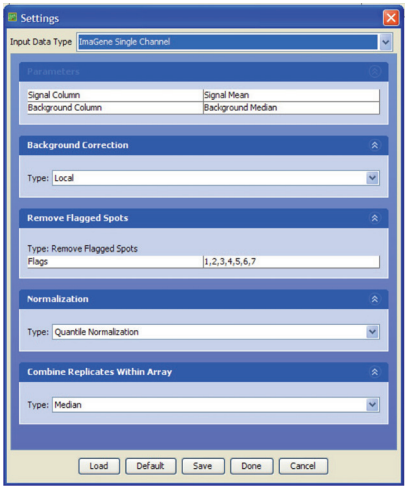


Figure 3: Single color settings file, Nexus.

Sometimes, a warning about missing annotation for a few probes occurs. This means that all probes were not annotated with a name in the annotations.txt file described above. Generally, this message can be ignored as the missing probes are likely to belong to another species, obsolete or might have been removed from the GAL-files for QC reasons. Annotation for the missing probe ID's can be looked up in the gal-file if needed. Click "continue" to get the processing started. Once processing is complete, the arrays will receive "processed" in the "status" field in the "Arrays" tab list.

The "probes" tab now contains normalized data for each miRNA, one signal value from each channel on the arrays in an expression matrix. This long list can be reduced by clicking "Filter probes" under the "options" menu. To remove probes that are below detection limit for a certain percentage of slides, check the "Filter low intensity probes" box and set "Remove probes having value below" to "1" and "for the specified percentage of all samples" to for example 50. This means that all probes that are detected on half of the slides or less are removed from the list. It is recommended to remove such probes prior to clustering, as probes with missing values (signal = "NaN") may affect the clustering outcome. You can save data from the "probes" tab as a text file anytime by clicking "EXPORT.txt". This can be done both before and after filtering the probe list.

Assigning factors to arrays or samples

Factors can be assigned to either samples or arrays. For arrays, this factor could be for example "date of hybridization" or "batch of arrays" if arrays were run different days, by different people or on different batches of arrays. These factors are automatically accounted for in the later statistical calculations. Click the "add factor" button on the arrays tab, and add information about factors associated with the slides. More frequently, factors should be assigned to the samples themselves. Click the "add factor" button on the samples tab and add information about for example sample group, treatment, tissue type or similar. Several factors can be added if for example both different tissue types and control/treated samples are used. A factor can be annotated in the sample list in text format or as numbers. As an example: 3 samples belong to a treated group and 3 to an untreated. Assign a factor

called "treatment", and annotate the control samples with "control" or "1", and the treated samples with "treated" or "2" in the created "treatment" column.

Visualization

All probes (except the ones filtered away if you have applied a filter) can be visualized in a heat map. Click the "view" button on the probes tab, and a new "heat map" tab is created. Each sample/array is visualized as a column and each miRNA is seen as a row. You can zoom in to see the miRNA name or probe ID on the right hand side of the heat map by clicking the left mouse button. Zoom out again with the right mouse button. The sample name is seen at the bottom of each column. You can choose to cluster the data in the heatmap. In the panel on top of the page, click the "cluster samples" button to group the samples by similarity. Click the "cluster probes" button to group the miRNAs by similarity. The settings for this clustering can be found and changed under the "options" menu, select "clustering". "Pearson correlation" and "average hierarchical clustering" are usually good default settings to start with. The Pearson correlation metric may, however, be sensitive to outliers. If you have an array that is a clear outlier, you may want to try another metric or remove the outlier slide before clustering to prevent it from producing false positive correlations.

Statistics

Once samples have been assigned different factors, statistical calculations about differentially expressed miRNAs can be performed. This is done under the "comparisons" tab. Click "add", and select the factor that you wish to analyze. Then click "next" and select the type of analysis you wish to perform. With 2 groups as in our example (treated and control), a paired analysis is appropriate. Check "paired", click "select", move the experiment group over to the "experiment" side, click "done" and then "finish". In the "comparisons" tab, you can now select the comparison you created and click "Differential probes". This opens a new window with the differential probes listed. At the bottom of the page, there are options for multiple test correction (none or FDR), and you can also select a p-value threshold and/or an expression logratio threshold. Just enter your preferred settings, click "Apply", and the differentially expressed miRNAs are displayed. The differentially expressed miRNAs are also visualized in the "heat map" tab that is

attached to the new window that opened with the comparison. The samples and miRNAs in this heat map can be clustered just as described above under visualization. You can save a list of differentially expressed miRNAs by clicking "Export.txt". For more information about the different comparisons available, please see the Nexus Expression™ manual.

Tip 1 An annotation.txt file can easily be created from any GAL-file. Open Excel, and paste in the whole gal-file. Delete the whole header and keep only the row that starts with "Block Column Row" and the rows below it. Delete the first 3 columns, the ones with headers "Block", "Column" and "Row". Rename the column that is now column 1, from "ID" to "Probe" with a capital P. Save the file as a text file (tab delimited .txt), called annotation.txt. The annotation file is now complete.

Tip 2 If multiple slides are to be uploaded it might be easier to load them using a description file. This file describes the data type, array type as well as the individual arrays and their location. An example of this can be seen in figure 4.

Figure 4

	A	B	C	D	E	F
1	Data Type:	ImaGene Single Channel	Array Type:	Exiqon/miRCURY_v11_20820X-A		
2	Array	File	Channel 1 Sample			
3	Sample1.1	C:\miRCURY_Array\29Oct09\1_Exiqon_14099048.txt	Sample1.1			
4	Sample1.2	C:\miRCURY_Array\29Oct09\1_Exiqon_14099049.txt	Sample1.2			
5	Control1	C:\miRCURY_Array\29Oct09\1_Exiqon_14099050.txt	Control1			
6	Sample2.1	C:\miRCURY_Array\30Oct09\1_Exiqon_14099051.txt	Sample2.1			
7	Sample2.2	C:\miRCURY_Array\30Oct09\1_Exiqon_14099052.txt	Sample2.2			
8	Control2	C:\miRCURY_Array\30Oct09\1_Exiqon_14099053.txt	Control2			

Figure 4. Example of a description file for single color array analysis. File is saved as tab delimited .txt file.

Tip 3 Right clicking capture probe ID's will open a menu with link to databases containing information on that specific miRNA.

Pipeline for data analysis in Nexus Expression™:

Create new project

Load files:
Data types: Imagene or Imagene single channel
Manufacturer: Exiqon
Array type: choose appropriate
Select output files from ImaGene® (*.txt)

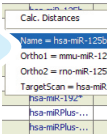
Menu: files/preprocessing settings

Select "Process"

Probes tab now contain normalized data for all probes

Save norm data as *.txt

Right click on probe ID to open menu, select name=xxx to go to miRbase description of selected miR



Filter probes (options menu)
Select desired filtering options

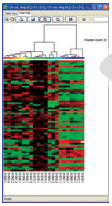
Assign factors: Sample/ arrays tabs

Generate heat map:
Menu: Options/clustering define settings
Press "View"

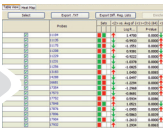
Statistics:
Comparison tab, select "Add"
Select factors
Select comparison
Finish

Mark samples for comparison
Press "differential probes"
Differential probes shown
Make adjustments, using parameters in lower part of window, press "apply"

Save image, sample cluster or probe cluster



Save differentially expressed data as *.txt



Outside North America
Exiqon A/S · Skelstedet 16
DK-2950 Vedbaek · Denmark
Phone +45 45 660 888
Fax +45 45 661 888
www.exiqon.com/contact

North America
Exiqon Inc.· 14 F Gill Street
Woburn, MA 01801 · United States
Phone (781) 376 4150
Fax (781) 376 4152
www.exiqon.com/contact

www.exiqon.com

EXIQON
Seek Find Verify

#13-0077 - 91708 - 12.0 - 03/2010