

# Agilent miRNA Data Import Guide

Importing Agilent miRNA Data and Creating an Experiment in GeneSpring GX 9



## How do I get my Agilent miRNA data into GeneSpring GX 9.0?

#### Introduction to the guide:

This guide will step you through the process of importing Agilent miRNA data and setting up the experiment for analysis in GeneSpring GX 9.

#### **Necessary components:**

- 1. GeneSpring GX 9.0
- 2. Feature Extraction "Gene View" files for your experiment of interest
- 3. A tab-delimited text file containing your experimental parameters with the following requirements:
  - a. A row of column headers that will be used as the Parameter names
  - b. A column containing the exact file names of the data files imported into GeneSpring GX. They will be used to map the Parameter values in the text file to the appropriate samples in GeneSpring GX.
  - c. Columns containing all the parameter values you wish to add.

#### Analysis in GeneSpring GX 9.0:

1. **Generate a Custom Technology.** The technology describes the array design, annotations associated with each probe, and the data file format. A technology needs to be created only once for a given array design.

- a. Launch GeneSpring GX, and create a new project or open an existing project to which you want to add your miRNA experiment (See Fig. 1).
  - i. To create a new project for your miRNA data, click on the **Create new project** radio button.
    - 1. Enter the name of this project.
    - 2. In the Experiment Selection Dialog window that appears, click the **Cancel** button.
    - 3. Proceed to Step b below.
  - ii. To add your miRNA experiment to an existing project, click on the **Open existing project** radio button.
    - 1. Select the project you would like to add the experiment to.
    - 2. Proceed to Step b below.





#### Figure 1

#### b. Go to Tools > Create Custom Technology > Generic One/Two Color

- c. In the Create Custom Technology (Step 1 of 9)- Technology Name window (See Fig. 2), select the following:
  - i. Technology Type: Single color
  - ii. Technology name: Type in the name for this technology. For example, you can type "Human\_miRNA\_V2". Note that you cannot have spaces in the name, and hence, an underscore character is used in place of a space.
  - iii. **Organism:** Select organism for your miRNA data. For this example, we will choose *Homo sapiens*.
  - iv. Choose a sample data file: Select one GeneView data file from your experiment. This data file will be used to define the data file format for all other data files to be imported into this technology. Thus, not only must all samples within an experiment have the same data file format, but all samples using the same technology must also have the same data file format.
  - Number of samples in single data file: One Sample. Some mircroarray platforms, such as Illumina, output data from multiple arrays into a single data file. For Agilent arrays, each Feature Extraction file contains data for a single sample.



- vi. **Choose annotation file:** Select one GeneView data file from your experiment. This can be the same or different data file than the one chosen above as long as the data files are from the same experiment. For Agilent miRNA data, the annotations to be used in GeneSpring GX are found in the GeneView data file.
- vii. Click Next>>

🎊 Create Custom Technology (Ste	p 1 of 9) 🔀
Technology Name Choose the name and other details for	r the technology.
Technology type	Single color 💌
Technology name	Human_miRNA_V2
Organism	Homosapiens
Choose a sample data file	pring GX Datasets\miRNA data\251643611426_501_1_1_Gen 🔽
Number of samples in single data file	One Sample
Choose annotation file	pring GX Datasets\miRNA data\251643611426_501_1_1_Gen 💙
Help	<< Back Next >> Einish Cancel

#### Figure 2.

- d. In the Create Custom Technology (Step 2 of 9)- Format data file window, select the following (See Fig. 3):
  - i. Separator: Tab
  - ii. Text qualifier: None
  - iii. Missing value indicator: None
  - iv. Comment indicator: None
  - v. Click Next>>



🅂 Crea	ate Custom T	echnology (	Step 2 of	i 9)				X
Forma	at data file							
For pre	mat file by spec esent, the separa	ifying the sepa ator in multi-val	rator, text lued columr	qualil ns	fier, missing valı	ue indicator, com	ment indicator and, if	
Format	Options							1
		Sep	parator 1	Гab				~
		Text o	jualifier [	Vone				~
	1	Missing value in	dicator 🛛	Vone				~
		Comment in	dicator 🗗	Vone				~
Preview	v							
	Column 0	Column 1	Column 2	2	Column 3	Column 4		
0	text	integer	float		float	boolean		~
1	SystematicN	ControlType	gTotalGer	ne	gTotalGene	gIsGeneDet		
2	DarkCorner	1	16.0166		12.3083	0		-
3	NC1_00000	-1	31.4707		30.1123	0		
4	NC1_00000	-1	-11.7839		29.8781	0		
5	NC2_00079	-1	11.1229		29.9361	0		
6	NC2_00092	-1	17.8597		29.9715	0		
7	NC2_00106	-1	30.9767		30.1012	0		
8	NC2_00122	-1	29.1687		30.068	0		
9	NegativeCo	-1	105.07		135.096	0		
10	SCorner3	1	1.44958		2.3429	0		
11	dmr_285	1	-7.24978		6.60288	0		~
He	lp				<< Ba	ick <u>N</u> ext >:	Einish Cancel	

## Figure 3.

- e. In the Create Custom Technology (Step 3 of 9)- Select Row Scope For Import window, select the following (See Fig. 4):
  - i. Row Options: Take all rows from index 1 to index leave cell empty
  - ii. **Header Row Options:** Take the first row in the selection as the column header
  - iii. Click Next>>



辉 Create Custom T	🏋 Create Custom Technology (Step 3 of 9) 🛛 🛛 🔀					
Select Row Scope For Import The file preview below shows the first 100 rows only (modifiable via Tools->Options->Miscellaneous->Custom Data Library Creation). Select rows to be imported in the row options below, and then select one of the Header Row options. Note that leaving the second textbox in Row Options 2 empty (explicit Enter required) will select all rows upto the end.						
Row Options     Take all rows     Take all rows from index						
Preview	een mark					
Column 0	Column 1	Column 2	Column 3	Column 4		
0 text	integer	float	float	boolean	~	
1 SystematicN	ControlType	gTotalGene	gTotalGene	gIsGeneDet	<b>—</b>	
2 DarkCorner	1	16.0166	12.3083	0		
3 NC1_00000	-1	31.4707	30.1123	0		
4 NC1_00000	-1	-11.7839	29.8781	0	<b>~</b>	
Header Row Options It + table In operations   O There is no row containing column headers   Image: Take the first row in the selection as the column header						
Help		<	< Back Nex	d >> Einis	h <u>C</u> ancel	

### Figure 4.

- f. In the Create Custom Technology (Step 4 of 9)- SingleColor one sample in one file selections window, select the following (See Fig. 5):
  - i. Identifier: Systematic Name
  - ii. BG Corrected Signal: gTotalGeneSignal
  - iii. Flag: glsGeneDetected
    - Click on the Configure button to assign values to each flag. Leave the Select Type parameter as Categorical. For "0", use the drop-down menu in the Absolute Calls column to select "Absent". For "1", use the drop-down menu in the Absolute Calls column to select "Present" (See Fig. 6). Click OK.

-0

辉 Create Custom Techn	ology (Ste	ep 4 of 9)		×
SingleColor one sample Check the data columns to columns can be changed o	e <b>in one f</b> o be importe on this page	file selections ad. The datatype, attribute type	and marks for the data	
	Identifier	SystematicName		*
BG Correc	ted Signal	gTotalGeneSignal		~
Flag	gIsGeneDe	tected	Configure	
		<< Back Next >	> Einish Cancel	

Figure 5.



K Configure the column		
Configure column Configure the column to ca columns can be continuous	tegorical or continuous, only , string columns cannot be c	/ int and float ontinuous.
Select Type	Categorical	
Categories		Absolute Calls
0		Absent
1		Present
<		>
	(	OK Cancel

#### Figure 6.

- iv. In the Create Custom Technology (Step 4 of 9)- SingleColor one sample in one file selections window, click **Next>>**.
- g. In the Create Custom Technology (Step 9 of 9)- Annotation Column Options window, select the following (See Fig. 7):
  - i. For the row containing **SystematicName** value under the **Column Name** column, use the drop-down menu in the **Column Mark** column to select "Identifier".
  - ii. Note that, by default, all the values under Column Name are checked, indicating that they will be imported into GeneSpring GX as annotations for each gene. If you do not want to import certain columns as annotation, uncheck the corresponding boxes. Annotations in columns selected for import will be shown in Entity List Inspector, Entity Inspector, Spreadsheet view, and other windows in GeneSpring GX. Marks are simply instructions to GeneSpring GX on how to use the annotations in the column. In the example above, we are instructing GeneSpring GX to use the values in Systematic Name column as Identifiers.



1. Uncheck column names: gTotalGeneSignal, gTotalGeneError, and gIsGeneDetected.

R	Cr	eate Custom	Technology	(Step 9 of 9)		×
A	Annotation Column Options Check the annotation columns to be imported. The datatype, attribute type and marks for the annotation columns can be changed on this page.					
	#	Column N	Data Type	Attribute	Column Mark	
✓	0	SystematicN	string	Categorical	Identifier	
✓	1	ControlType	integer	Continuous	None	
	2	gTotalGeneS	float	Continuous	None	
	3	gTotalGeneE	float	Continuous	None	
	4	gIsGeneDet	integer	Categorical	None	
	[	<u>H</u> elp	(	<< Back	Next >> Einish Cancel	



- iii. Click Finish.
- iv. A window will appear warning that some columns are un-marked and therefore will be given default mark names (See Fig. 8). Click OK.



#### Figure 8.

v. A window will appear warning that Entrez, Chromosome, and Gene Ontology Marks are not preset (See Fig. 9). Click **OK**.







h. When the technology is successfully created, a window will appear indicating so. Click **OK**. The technology is now saved in the GeneSpring GX database.

#### 2. Create an experiment.

- a. To simplify data import, have all desired GeneView text files in a single folder.
- b. Go to **Project > New Experiment**, or click on the **New Experiment** icon under the menu.
- c. In the New Experiment- Experiment description window, select the following (See Fig. 10):
  - i. **Experiment Name:** Type in the name of the experiment. For example, "Agilent human miRNA data".
  - ii. Experiment type: Generic Single Color
  - iii. Workflow type: Advanced Analysis
  - iv. **Experiment notes:** Type in any other description that you would like to associate with the experiment. Note that experiments can later be searched by the content of the Experiment notes section.
  - v. Click OK.



🅂 New Experiment		$\mathbf{X}$
Experiment description		
Enter a name for the new experiment, s desired workflow. Guided workflows will advanced analysis will allow access to th	elect the appropriate experiment type, and choose the take you through experiment creation and analysis, while he full set of analysis tools.	
Experiment name	Agilent human miRNA data	
Experiment type	Generic Single Color	~
Workflow type	Advanced Analysis	~
Experiment notes	Π	
Help	OK	el

#### Figure 10.

- d. In the New Experiment (Step 1 of 2)- Load Data window, select the following (See Fig. 11)
  - From the drop-down menu, select the custom technology that you just created. Note that you only need to create a custom technology for each array type. In other words, the next time you create an experiment with samples applied to the Agilent Human miRNA (V2) Microarray, you can use the same custom technology that you had just set up.
  - ii. Click **Choose Files** and select the data files you would like to create an experiment from.
  - iii. In the New Experiment (Step 1 of 2)- Load Data window, click Next>>.



🅂 New Exp	eriment (Step 1 of 2)	×
Load Data You can imported	<b>a</b> choose data files, previously used samples or both to use in this experiment. Once a data file has been I and used as a sample, it will be available for use in any future experiment.	
	Select the technology Human_miRNA_V2	~
Туре	Selected files and samples	1
	251911810001_501_1_1_GeneView.txt	~
	251911810001_501_1_2_GeneView.txt	
	251911810001_501_1_3_GeneView.txt	
	251911810001_501_1_4_GeneView.txt	
	251911810001_S01_2_1_GeneView.txt	
	251911810001_501_2_2_GeneView.txt	
	251911810001_501_2_3_GeneView.txt	
	251911810001_501_2_4_GeneView.txt	=
	251911810002_501_1_1_GeneView.txt	
	251911810002_501_1_2_GeneView.txt	
	251911810002_501_1_3_GeneView.txt	
	251911810002_501_1_4_GeneView.txt	
	251911810002_501_2_1_GeneView.txt	
	251911810002_501_2_2_GeneView.txt	
	251911810002_501_2_3_GeneView.txt	
	251911810002_501_2_4_GeneView.txt	~
	Choose Files Choose Samples Choose Raw Files Reorder Remove	
Help	<< Back Next >> Einish Cancel	]

## Figure 11.

- e. In the New Experiment (Step 2 of 2)- Preprocess Options window, select the following (See Fig. 12):
  - i. Threshold raw signals to: 1
  - Normalization algorithm: At the present time, there is no good method for normalizing miRNA data. We therefore recommend either applying no normalization or normalizing to the 75<sup>th</sup> percentile. Depending on which option you take, select either:
    - 1. None
    - 2. Median Shift



- iii. Median Shift to percentile: If None was chosen for normalization, this option will be inactivated. If Median Shift was chosen, enter "75" in this box
- iv. Baseline Options: Do not perform baseline transformation
- v. Click Finish

🗶 New Experiment (Step 2 of 2)		$\mathbf{X}$
Preprocess Options Choose options for preprocessing the	input data.	
Threshold raw signals to Normalization algorithm	1 Median Shift	<ul><li></li><li></li><li></li><li></li><li></li><li></li><li></li><li></li><li></li><li></li><li></li><li></li><li></li><li></li><li></li><li></li><li></li><li></li><li></li><li></li><li></li><li></li><li></li><li></li><li></li><li></li><li></li><li></li><li></li><li></li><li></li><li></li><li></li><li></li><li></li><li></li><li></li><li></li><li></li><li></li><li></li><li></li><li></li><li></li><li></li><li></li></ul>
Median shift to percentile	75	
Do not perform baseline transformati Do Baseline to median of all samples Baseline to median of control samples	oni s	
Choose samples: Available samples 251911810001_501_1_1_GeneView.b 251911810001_501_1_2_GeneView.b 251911810001_501_1_3_GeneView.b 251911810001_501_2_1_GeneView.b 251911810001_501_2_1_GeneView.b	Control samples	
Help	<< Back Next >> Einish Cancel	]

## Figure 12.

f. A new experiment is now created in GeneSpring GX (See Fig. 13)



Figure 13.

#### 3. Setting up the experiment for analysis.

- a. The next step is to define the experimental parameters for the experiment and assign parameter values to each sample.
  - i. In the **Workflow** panel, open the **Experiment Setup** section and click on the **Experiment Grouping** link.
  - ii. Parameters and parameter values associated with your experiment can be added to this window in one of two ways:
    - 1. Parameters and parameter values for each sample can be automatically loaded from a tab-delimited text file containing the required information.
      - a. Click on the Load parameters from file icon 2, in the Experiment Grouping window (See Fig. 14) and select the tab-delimited text file containing the required information. Parameters and Parameter Values should now be added to the samples.



b. The tab-delimited text file must contain 1) a row of column headers that will be used as the Parameter names 2) a column containing the exact file names of the data files imported into GeneSpring GX. They will be used to map the Parameter values in the text file to the appropriate samples in GeneSpring GX and 3) columns containing all the parameter values you wish to add (See Fig 15)

辉 Experiment Groupin	🕅 Experiment Grouping					
Experiment parameters define the grouping or replicate structure of your experiment. Enter experiment parameters by clicking on the "Add Parameter" button. You can also edit and re-order parameters and parameter values here.						
Experiment Grouping				₽		
	述					
Samples	Sample	PercentageBrain	ReplicateNumber			
251911810001_50	Brain	100	1			
251911810001_50	Placenta	0	1			
251911810001_S0	B3P1	75	1			
251911810001_50	B1P3	25	1			
251911810001_S0	Brain	100	2			
251911810001_50	Placenta	0	2			
251911810001_50	B3P1	75	2			
251911810001_50	B1P3	25	2			
251911810002_50	B3P1	75	3			
251911810002_50	Brain	100	3			
251911810002_50	B1P3	25	3			
251911810002_50	Placenta	0	3			
251911810002_50	B1P3	25	4			
251911810002_S0  B3P1 75 4						
Add Parameter Edit Parameter Delete Parameter						

## Figure 14.

File Name	Sample	PercentageBrain	ReplicateNumber
251911810001_S01_1_1_GeneView.txt	Brain	100	1
251911810001_S01_1_2_GeneView.txt	Placenta	0	1
251911810001_S01_1_3_GeneView.txt	B3P1	75	1
251911810001_S01_1_4_GeneView.txt	B1P3	25	1
251911810001_S01_2_1_GeneView.txt	Brain	100	2
251911810001_S01_2_2_GeneView.txt	Placenta	0	2
251911810001_S01_2_3_GeneView.txt	B3P1	75	2



251911810001_S01_2_4_GeneView.txt	B1P3	25	2
251911810002_S01_1_1_GeneView.txt	B3P1	75	3
251911810002_S01_1_2_GeneView.txt	Brain	100	3
251911810002_S01_1_3_GeneView.txt	B1P3	25	3
251911810002_S01_1_4_GeneView.txt	Placenta	0	3
251911810002_S01_2_1_GeneView.txt	B1P3	25	4
251911810002_S01_2_2_GeneView.txt	B3P1	75	4
251911810002_S01_2_3_GeneView.txt	Brain	100	4
251911810002_S01_2_4_GeneView.txt	Placenta	0	4

## Figure 15.

- 2. Parameters and parameter values for each sample can be added manually.
  - a. Click on the **Add Parameter** button in the Experiment Parameters window.
  - b. Type in the Parameter Name.
  - c. Select the replicates (samples that share the same value for that parameter) by holding down the Ctrl key and clicking on the samples (See Fig.16). For example, in this experiment, there are four replicates with the value "Brain" for the parameter Sample.

🌠 Add/Edit Experiment Pa	rameter 🔰 🕑
Grouping of Samples Samples with the same paran samples. To assign replicate the samples and click on the value for the group.	neter values are treated as replicate samples their parameter values, select "Assign Values" button, and enter the
Parameter name	Sample
Samples	Parameter Values
251911810001_501_1_1_GeneVie.	
251911810001_501_1_2_GeneVie.	
251911810001_501_1_3_GeneVie.	
251911810001_501_1_4_GeneVie.	
251911810001_501_2_1_GeneVie.	
251911810001_501_2_2_Genevie. 251911810001_501_2_3_Genevie.	
251911810001_501_2_3_GeneVie.	
251911810002 S01 1 1 GeneVie.	
251911810002_501_1_2_GeneVie.	
251911810002_501_1_3_GeneVie.	
251911810002_501_1_4_GeneVie.	
251911810002_501_2_1_GeneVie.	
251911810002_501_2_2_GeneVie.	
251911810002_501_2_3_GeneVie.	
Assign V	'alue Clear
Assign V	'alue Clear

Figure 16.

- d. Once you have your first group of replicates selected, click the **Assign Value...** button.
- e. Type in the parameter name (See Fig. 17).

Assign Value	X
Enter a value for the selected samples	
Brain	
OK Cancel	

Figure 17.

f. Parameter values should now be added to the selected samples (See Fig. 18).

<u> </u>	Agilont	Technol	onios
245	Aynent	recimo	ogies

Ҟ Add/Edit Experiment Parar	neter	×
Grouping of Samples Samples with the same parameter samples. To assign replicate sam the samples and click on the "As value for the group.	er values are treated as replicate oples their parameter values, select sign Values" button, and enter the	
Parameter name Sam	ple	
Samples	Parameter Values	
251911810001_501_1_1_GeneVie	Brain	~
251911810001_501_1_2_GeneVie		
251911810001_501_1_3_GeneVie		
251911810001_501_1_4_GeneVie		
251911810001_501_2_1_GeneVie	Brain	
251911810001_501_2_2_GeneVie		
251911810001_501_2_3_GeneVie		
251911810001_501_2_4_GeneVie		=
251911810002_501_1_1_GeneVie		
251911810002_501_1_2_GeneVie	Brain	
251911810002_501_1_3_GeneVie		
251911810002_501_1_4_GeneVie		
251911810002_501_2_1_GeneVie		
251911810002_501_2_2_GeneVie		
251911810002_501_2_3_GeneVie	Brain	~
Assign Value Clear		
Help	OK Cance	:

Figure 18.

- g. Repeat these steps for the remaining samples and parameters to be added.
- b. The next step is to create interpretations to group replicate samples into conditions.
  - i. In the **Workflow** panel, open the **Experiment Setup** section and click on the **Create Interpretation** link.
  - ii. First, we will create an interpretation where samples are grouped by the parameter "Sample". Note that this is just an example used in this guide.
    - In the Create Interpretation (Step 1 of 3)- Select parameters window, check the parameter "Sample" and click Next >> (See Fig. 19).

K Create Interpretation (Step 1 of 3)	×
Select parameters	
An Interpretation specifies how samples will be grouped into experimental conditions for display and used for analysis. Select the parameter(s) to group samples by. All samples with the same parameter values will be grouped into an experimental condition.	
Select experiment parameters	
Sample	
Help << Back Next >> Einish Cancel	

Figure 19.

 In the Create Interpretation (Step 2 of 3)- Select conditions window, make sure that all the conditions defined by the parameter "Sample" are checked. There are four unique parameter values for parameter "Sample". Therefore, samples will be grouped into four unique experimental conditions: B1P3, B3P1, Brain, and Placenta (See Fig. 20). Click Next>>.

🋠 Create Interpretation (Step 2 of 3)	×
Select conditions Select the conditions defined by the selected parameter(s) to include in the interpretation. Samples within a condition are considered as replicates and, for each entity, the average intensity value across replicates will be used for visualization and analysis.	
Unselect conditions to exclude ✓ [B1P3] ✓ [B3P1] ✓ [Brain] ✓ [Placenta]	
Average over replicates in conditions	
Help << Back Next >> Einish Cancel	

Figure 20.

- 3. Save the new interpretation as "Sample" (See Fig. 21).
  - a. GeneSpring GX will give each object created a default name. However, this can be changed to a name of your choice.
  - b. Click Finish.



🎗 Create Interpretation (Step 3 of 3)	
Save Interpretation This page displays the details of the interpretation created.	
Name Sample	
Notes	
Creation date Wed Apr 09 17:39:54 PDT 2008	
Last modified date Wed Apr 09 17:39:54 PDT 2008	
Owner pamt	
Average over replicates in conditions Yes	
Parameters Conditions	
Parameters Sample	
Help << Back Mext >> Einish Cancel	

Figure 21.

- c. View data as defined by the Sample interpretation (See Fig. 22).
  - i. A profile plot displaying your expression data should automatically appear in the browser of GeneSpring GX.
  - ii. Look into the Analysis folder. The "All Entities" list is in bold, indicating that the list is selected for display in the profile plot. GeneSpring GX will only show the expression data for the entities in the selected list.
  - iii. Look into the Interpretations folder. The "Sample" interpretation is in bold, indicating that the interpretation is selected for display. GeneSpring GX will display expression data for the entities in the selected entity list, according to the sample grouping defined by the selected interpretation.





Figure 22.