

# mRNA and miRNA Data Analysis in Partek<sup>®</sup> Genomics Suite<sup>™</sup>

HANDS-ON TRAINING

---

**University of Minnesota**

*May2014*



**Xiaowen Wang**  
*Field Application Specialist*  
*Partek Incorporated*

# Contents

---

<b>Partek Main Dialog</b> .....	3
<b>miRNA Data Analysis</b> .....	6
Importing Data from Affymetrix Cel Files.....	7
Filter Rows and Columns.....	9
QA & QC.....	11
Detect Differential Expressed miRNA.....	13
Hierarchical Clustering.....	17
<b>mRNA Data Analysis</b> .....	19
Exercise: Gene Expression Data Analysis.....	20
<b>miRNA and mRNA Integration</b> .....	21

# Partek<sup>®</sup> Genomic Suite<sup>™</sup> Main Dialog

**Analytical spreadsheet:** Central repository of data

- No limitation on number of rows or columns
- Rows represent observations of interest (experiments, samples, chips)
- Columns represent measures of the observations (variables, features, genes,)

**Menu bar:** Execute commands from a graphical user interface

- When spreadsheet is empty, most of the menu items are not displayed

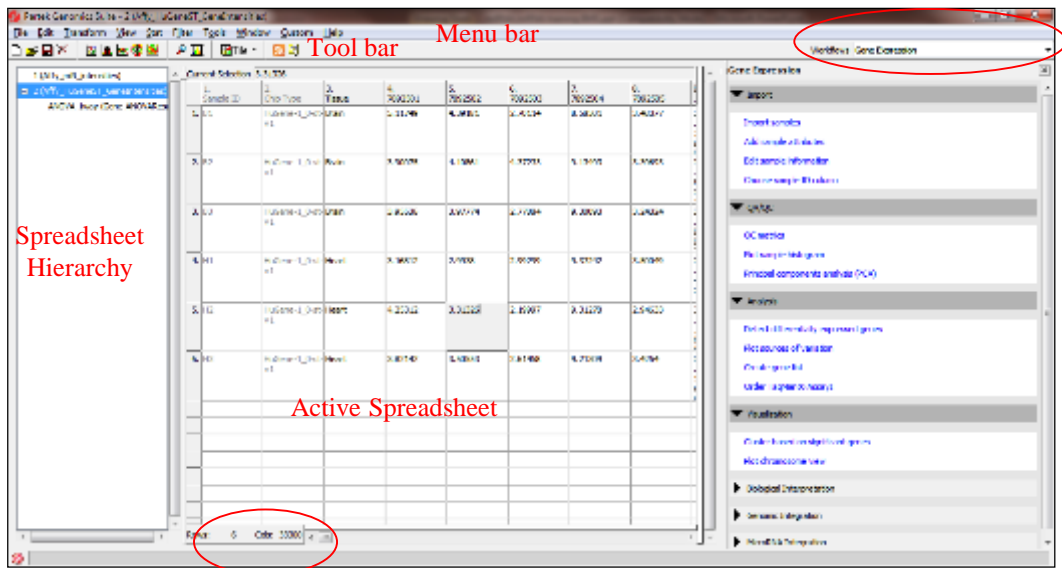
**Tool bar:** Accelerator buttons allow quick access to commonly used commands

**Spreadsheet hierarchy:** Open multiple datasets and see the hierarchy

- Original spreadsheet: parent
- Result spreadsheet: child

**Active spreadsheet:** The active spreadsheet is shown highlighted in blue, and the spreadsheet name and associated file name are shown at the top of the dialog

**Workflow:** Used to guide you through a typical analysis of a specific assay



**Notes:** \_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

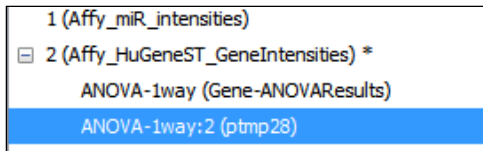
\_\_\_\_\_

\_\_\_\_\_

# Spreadsheet Properties

## Spreadsheet

- Each spreadsheet consists of two files with the same name
- \* implies change is not saved, ptmp is unsaved temporary file
- Spreadsheet linked to annotation
- Saving the project will preserve the hierarchy



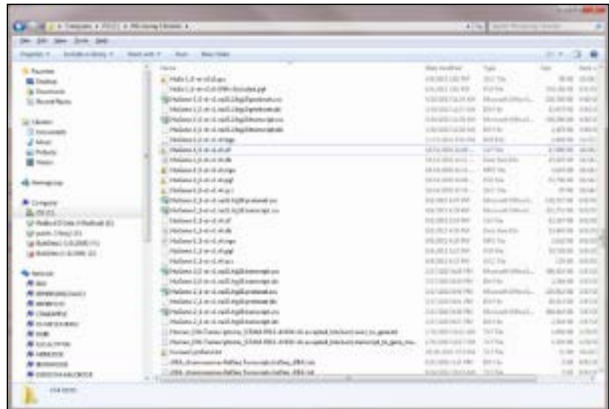
Name	Type	Size
Affy_HuGeneST_GeneIntensities	File	782 KB
Affy_HuGeneST_GeneIntensities.fmt	Partek Data File	296 KB
Affy_miR_intensities	File	184 KB
Affy_miR_intensities.fmt	Partek Data File	133 KB

```

1 {C:\Partek Example Data\microRNA_data\WT Gene Array\gene.fmt}
1/ANOVA-1way {C:/Partek Example Data/microRNA_data/WT Gene Array/gene-ANOVAResults.fmt}
1/brain_vs_heart {C:/Partek Example Data/microRNA_data/WT Gene Array/mRNA-sig-list.fmt}
2 {C:/Partek Example Data/microRNA_data/miRNA/miRNA.fmt}
2/1 {C:/Partek Example Data/microRNA_data/miRNA/miRNA-ANOVAResults.fmt}
2/brain_vs_heart {C:/Partek Example Data/microRNA_data/miRNA/miRNA-sig-list.fmt}
3 {C:/Partek Example Data/microRNA_data/WT Gene Array/enrichedAssociations.txt.fmt}
4 {C:/Partek Example Data/microRNA_data/Targets from spreadsheet.txt.fmt}
5 {C:/Partek Example Data/microRNA_data/Targets from Targetscan.txt.fmt}
    
```

## Annotation

- To link the annotation to the spreadsheet select: File > Properties
- Microarray Libraries: Stores automatically downloaded annotations files



## Notes:

# Training Data

## Data files in the project:

- Six samples with 3 brain vs. 3 heart
- miRNA: six cel files from Affymetrix miRNA-1\_0 array
- Gene expression: Imported RMA on Affymetrix HuGene-1\_0-st-v1

## Annotation file in the project:

- HuGene-1\_0ST transcript annotation

The screenshot shows a software window titled "Partek Genomics Suite - 1.0.0.100 - Unstable". The main area displays a table with the following columns: ID, Sample ID, Tissue Type, Tissue, and 14 numerical columns labeled 1 through 14. The table contains 6 rows of data.

ID	Sample ID	Tissue Type	Tissue	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1.01	miRNA_1_0	Brain	Brain	0.17508	1.72907	0.07439	1.81706	0.48843	0.15874	1.30783	0.81705	0.08105	0.08105	0.08105	0.08105	0.08105	0.08105
2.02	miRNA_1_0	Brain	Brain	2.23462	1.72517	0.07530	1.80149	0.00114	0.07493	0.07499	0.07499	0.07499	0.07499	0.07499	0.07499	0.07499	0.07499
3.03	miRNA_1_0	Brain	Brain	0.04128	1.72902	0.07392	1.70218	0.44117	0.07405	1.00000	0.70077	0.03011	0.03011	0.03011	0.03011	0.03011	0.03011
4.04	miRNA_1_0	Heart	Heart	1.88629	1.72895	1.00000	1.73670	0.11076	0.07479	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000	1.00000
5.05	miRNA_1_0	Heart	Heart	0.00110	1.72948	0.07426	1.26251	2.08120	0.07500	1.04211	1.75000	0.07500	0.07500	0.07500	0.07500	0.07500	0.07500
6.06	miRNA_1_0	Heart	Heart	0.78521	1.73154	1.01782	1.51583	0.41782	0.07504	1.14504	1.79508	0.17181	0.17181	0.17181	0.17181	0.17181	0.17181

Notes: \_\_\_\_\_

\_\_\_\_\_

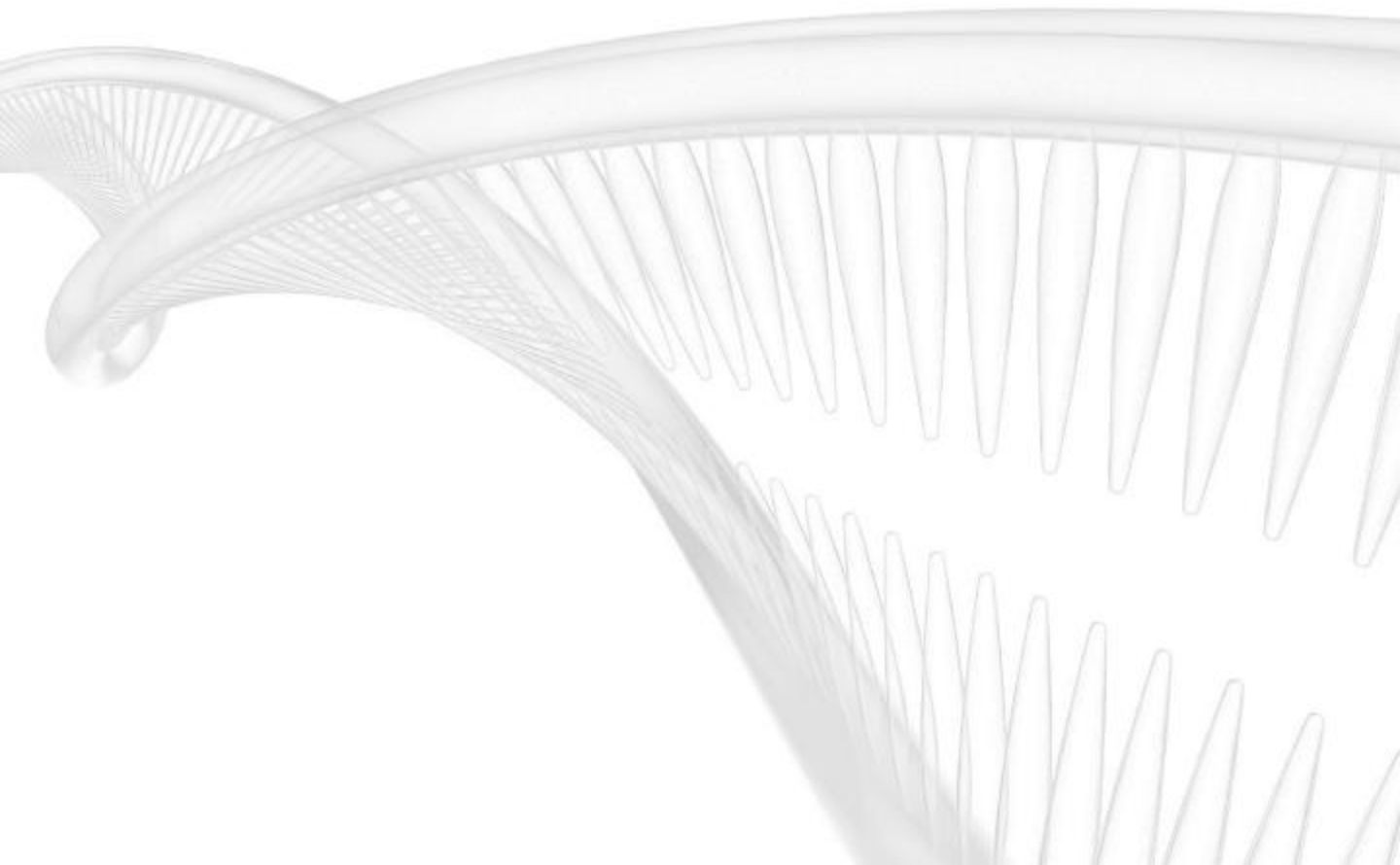
\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

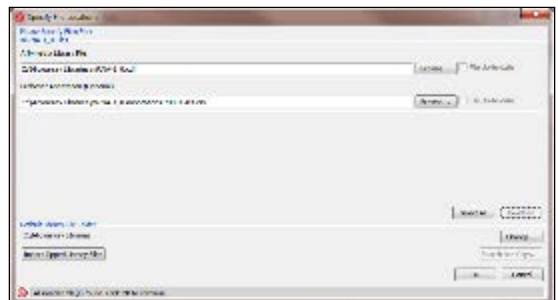
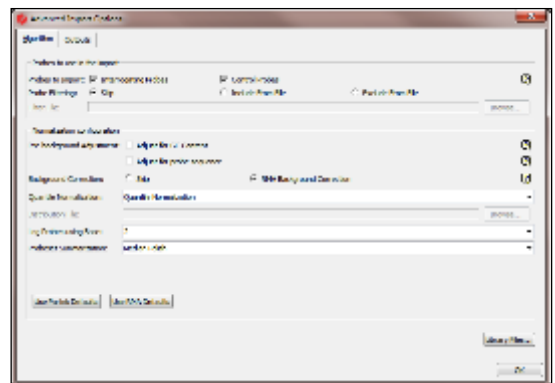
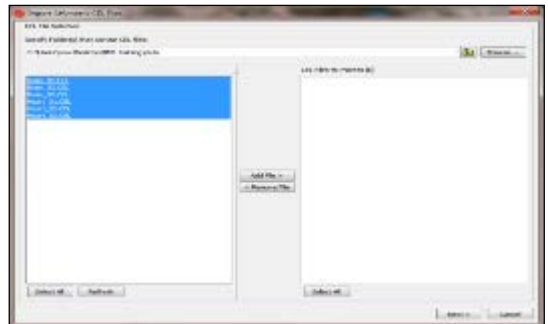
\_\_\_\_\_

# miRNA Data Analysis



# Importing miRNA Data from Affymetrix® CEL Files

- Choose the **MicroRNA Expression** workflow
- Browse to the folder that contains the CEL files
- Select all the default CEL files, and drag them to the right panel
- Click **Next**
- Specify the output file name—“*miRNA data*” and use the default settings, then click **Import**
- **Customized** allows you change the algorithm parameters, and verify library files



Notes: \_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_

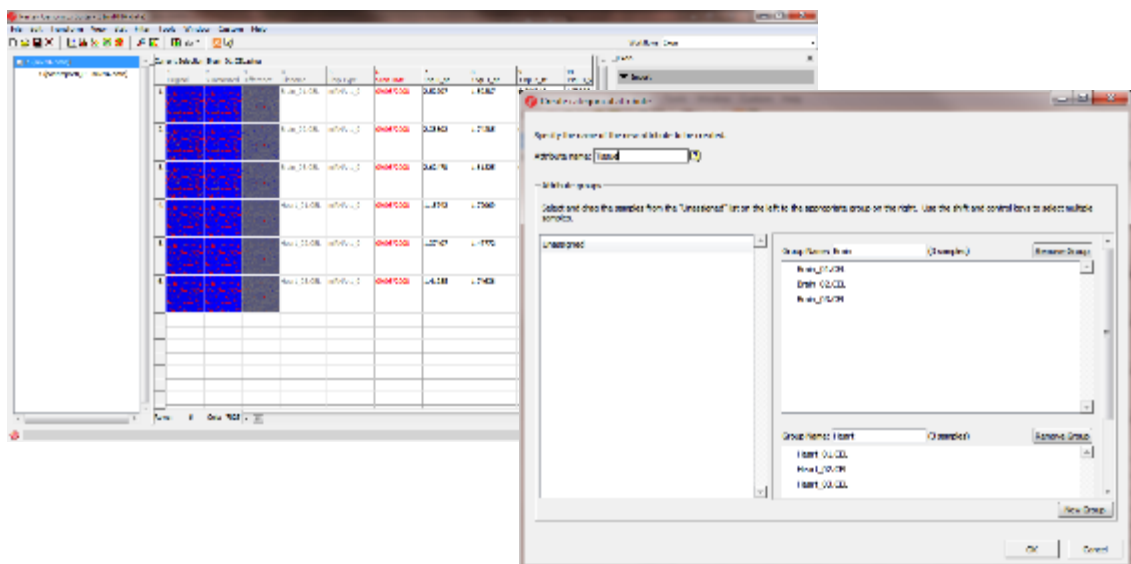
# Add Sample Attributes

## Two spreadsheets are generated:

The data spreadsheet contains the RMA value for all probesets

The QC spreadsheet contains the control probe sets

- Select the **miRNA data** spreadsheet
- Choose **Add Sample Attributes**
- Select **Add a Categorical Attribute**
- The attribute name is **Tissue**, and the group names are **Brain** and **Heart**
- Shift + click to select samples, drag and drop to the corresponding panel
- Click **OK**
- Change the **Chip Type** column name to **Sample ID**: labeled as **B1, B2, B3, H1, H2, H3**
- Click on **Choose sample ID column** to select **Sample ID**



Notes:

---

---

---

---

---

---

---

---



# Filter Rows and Columns

---

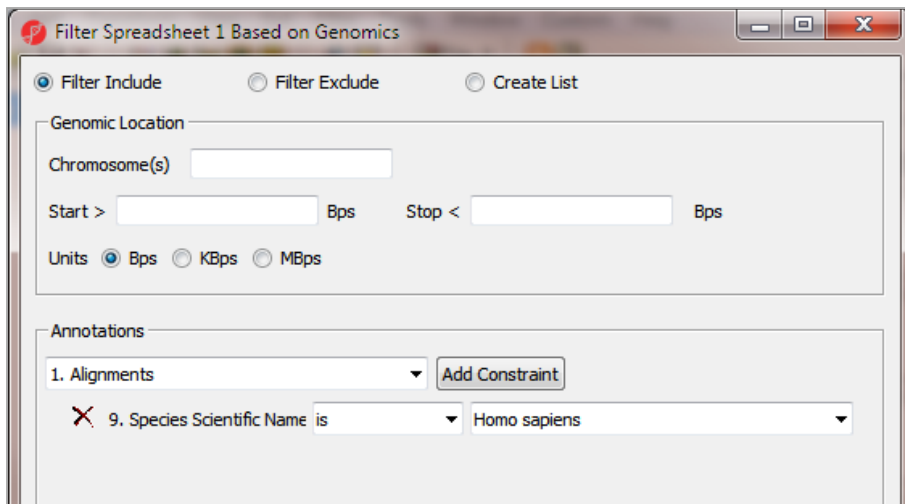
This chip contains miRNA from multiple species, but we only focus on human miRNA in this example.

Filter data to include only human miRNA

- Choose **Filter>Filter based on the genomic location**
- Select the **Scientific Species Name**, and click **Add Constraint**
- Choose **Is Home Sapiens**
- Click **OK**

The spreadsheet will contains probesets of human miRNA, save the filtered data as a different file name for down stream analysis

- Choose **File > Save As Binary**: name the spreadsheet as **miRNA data human**



Notes: \_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_

# Interactive Row Filter

---

The interactive row filter is helpful to view sample information and create a subset. We will not apply any filter on this data, only practice on how to use it.

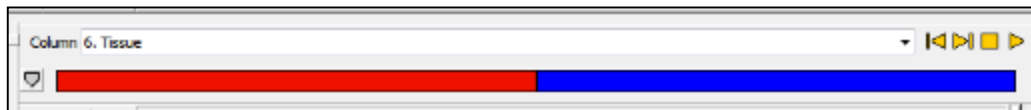
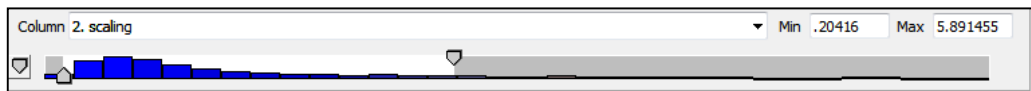
- **Select Filter > Filter Rows > Interactive Filter from the menu**
- Select a column from the drop-down list

### Categorical column:

- Right click the filter to include (similar to a radio button)
- Left click to toggle the status (similar to a check button)

### Continuous column:

- Slide the indicator to filter
  - Type in the minimum and maximum to filter
  - Click on the **Configuration Menu** on the left of the filter bar to configure the **Filter Type** and to determine the filter inclusions and exclusions
- Right click on the filter indicator and select **Clear Filter**



Notes: \_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_

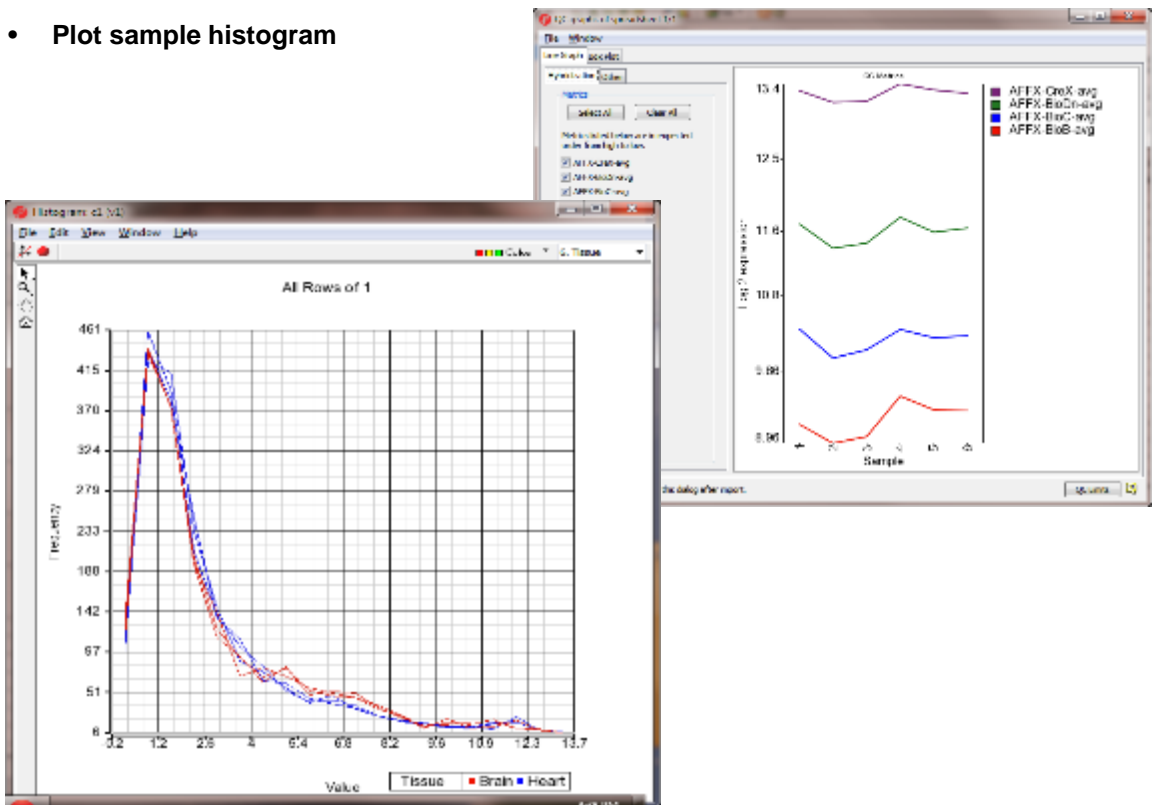
# QA and QC

QA/QC is exploratory analysis, it is checking the preparation of the samples and identify outliers

QC metrics is only available when you import Affymetrix .cel files, it checks the quality of the chips based on control probesets

Histogram display the distribution of the samples

- **QC metrics** –PostImportQC
- **Plot sample histogram**



Notes: \_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_

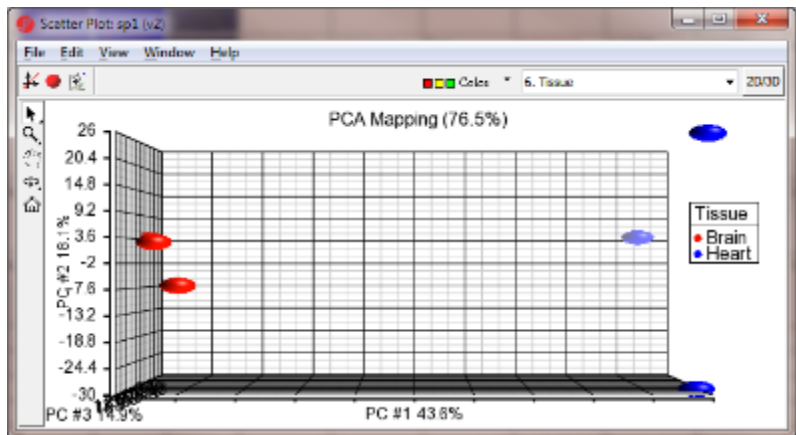
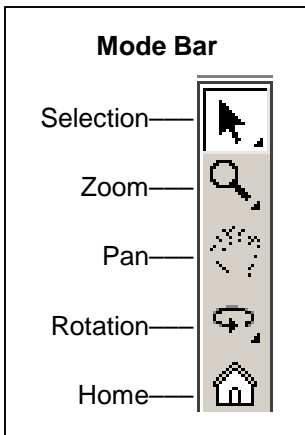
# PCA Scatter Plot

PCA scatter plot is another way to identify clustering patterns and outliers

- Select **Principal components analysis** on the workflow

## Notes

- Each point in the scatter plot corresponds to a specific row in the spreadsheet
  - Points that are close together in the plot are similar in the original high-dimensional space
  - Points that are far apart in the plot are dissimilar
- 
- Click on **Plot Properties** (red ball), to configure color by tissue
  - Click on **Ellipsoid** to put the ellipsoid on each tissue type



Notes: \_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

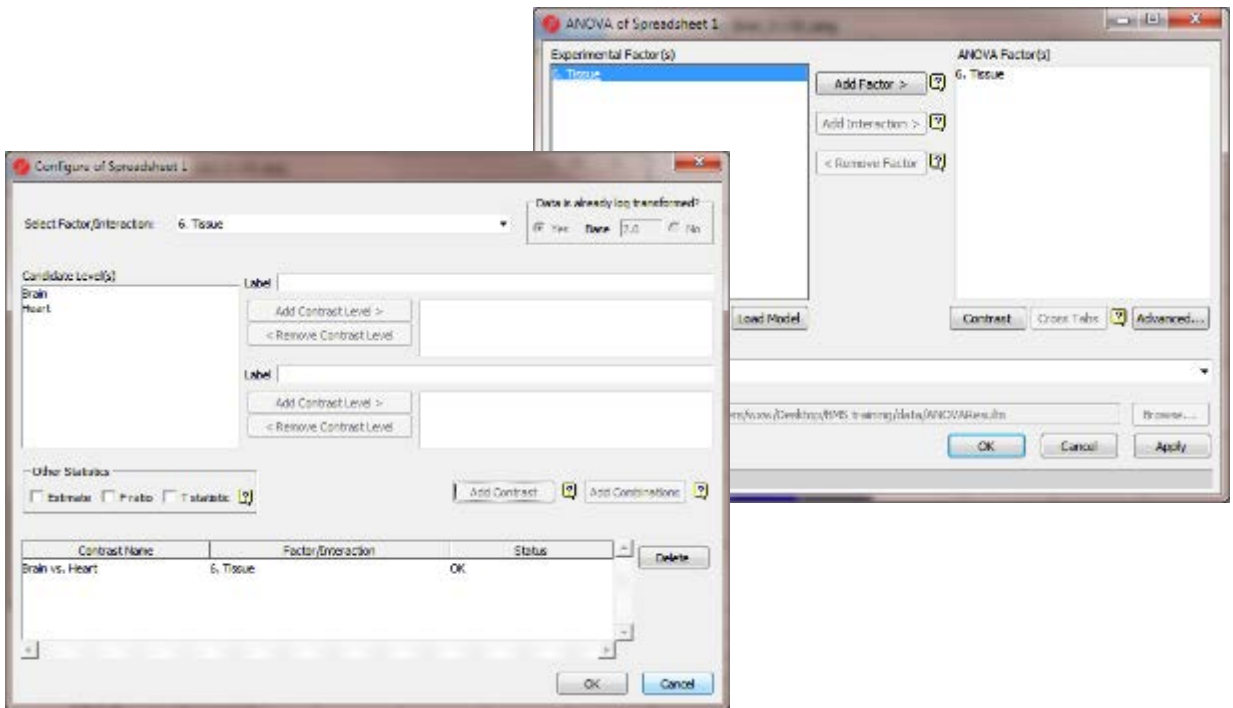
\_\_\_\_\_

\_\_\_\_\_

# Detect Differentially Expressed miRNA

- Select **Tissue** to move to the right panel
- Click **Contrast**
- Add contrast of **Brain vs. Heart**
- Click **OK**
- Outpt file: **miRNA-human ANOVA Result**

**Note:** Fold change calculation is different on linear vs. log data.



**Notes:** \_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_

# ANOVA in Partek Genomics Suite

## Different Types of ANOVA

- Equal variance t-Test
- Paired t-Test
- Repeated Measurement ANOVA
- ANCOVA
- Mixed Model ANOVA

Automatically detects crossed/nested factors

Automatically performs mixed model when random effect are included

### 2. Type vs. 3. Tissue

Type\Tissue	Astrocyte	Cerebellum	Cerebrum	Heart	Total
Down Syndrome	2	3	4	2	11
Normal	2	3	7	2	14
Total	4	6	11	4	25

### ANOVA Factor(s)

2. Type  
 3. Tissue  
 4. subject (2. Type)  
 2. Type \* 3. Tissue

### 2. Type vs. 4. subject \*

Type\subject	1218	1389	1390	1411	1478	1479	1521	1565	748	847	Total
Down Syndrome	3	2	0	0	4	0	0	0	1	1	11
Normal	0	0	4	4	0	1	4	1	0	0	14
Total	3	2	4	4	4	1	4	1	1	1	25

Notes: \_\_\_\_\_  
 \_\_\_\_\_  
 \_\_\_\_\_  
 \_\_\_\_\_  
 \_\_\_\_\_  
 \_\_\_\_\_

# Result of ANOVA

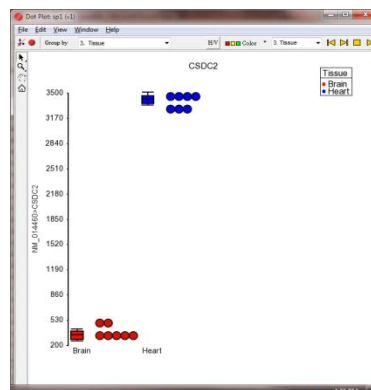
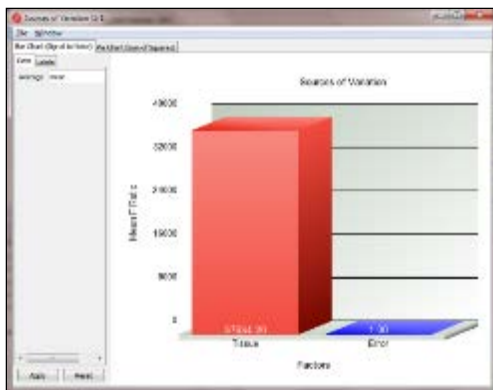
---

Each row is a miRNA with its p-value and fold change and any other statistical information. The spreadsheet is sorted by the first p-value column.

**Right click on a row header to get details**

- Select **HML Report**
- Select **Dot Plot**
- Select **Source of Variation**
- Select **Find miRNA in Different Database**

Right click on **the ANOVA spreadsheet > Info > Comments** to access the ANOVA model details



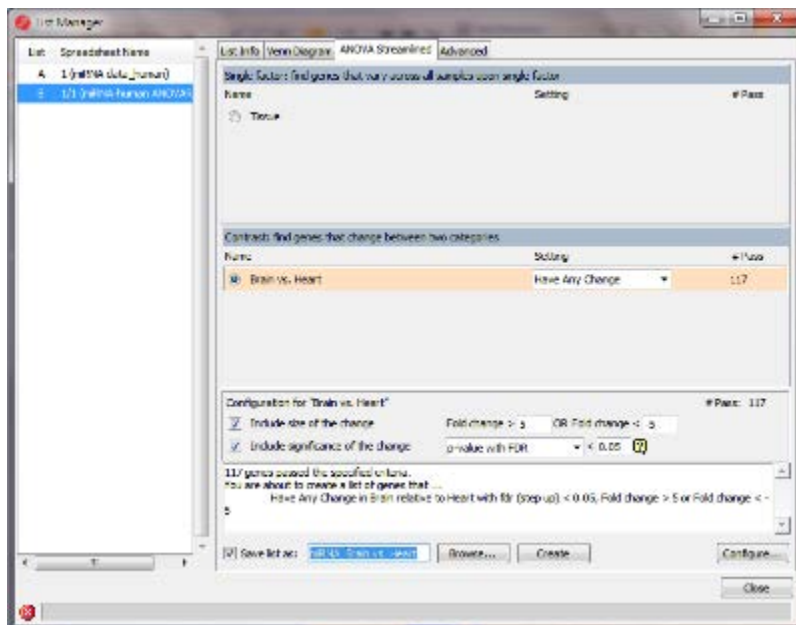
Notes: \_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_

# Create List

Generate a list of miRNA that is showing differential expression between brain and heart.

- Click **Create List** on the workflow
- Choose the **Brain vs Heart** contrast
- Set the **Fold Change to >5 or <-5**
- Set pValue to **FDR < 0.05**
- Name the result as **miRNA Brain vs Heart**
- Click **Create**
- Close the dialog

A new child spreadsheet will be generated



Notes: \_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

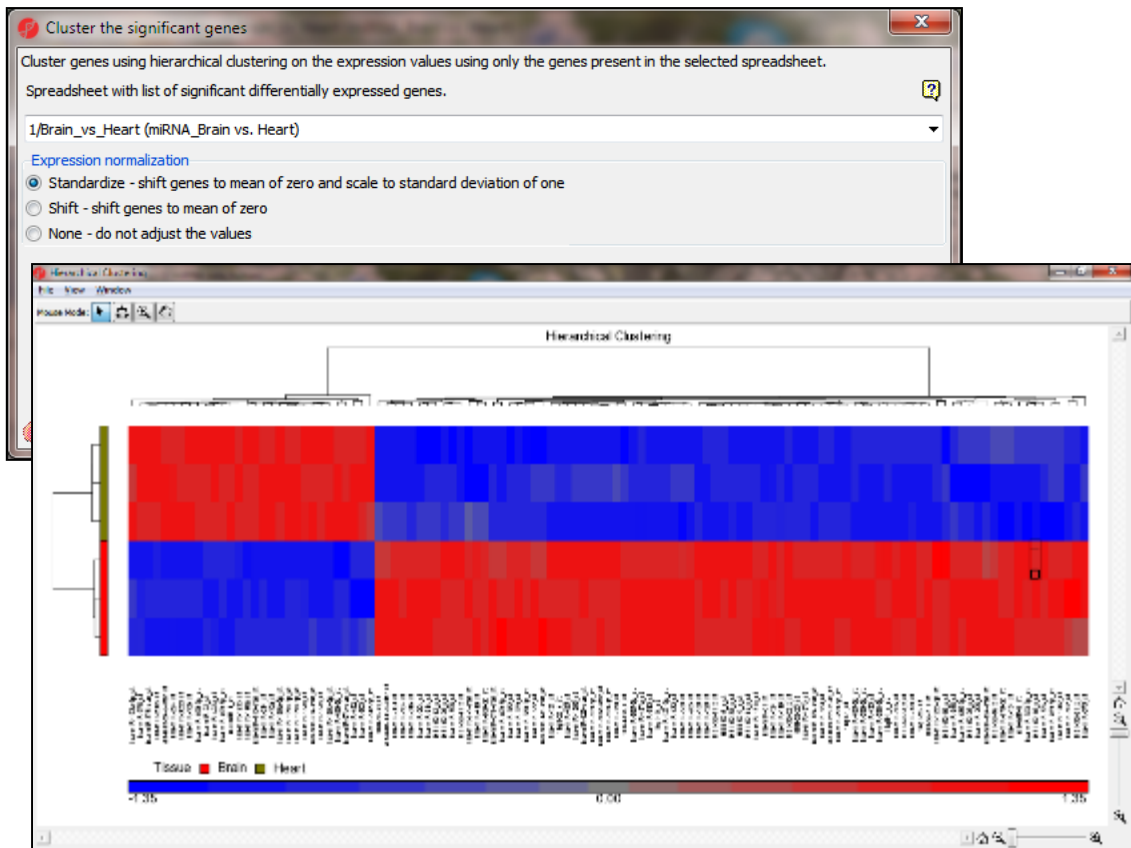
\_\_\_\_\_



# Hierarchical Clustering

To visualize the heatmap and cluster of the significant list of miRNA:

- Choose **Cluster Based on Significant miRNAs** on the workflow
- Select the **Hierarchical Clustering** option
- Choose the **miRNA Brain vs. Heart** spreadsheet with default settings
- Click **OK**



Notes: \_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_

# Hierarchical Clustering Configuration

## Heatmap

- Click on the color square to change the heatmap color

## Dendrograms

- Uncheck Show dendrogram scale
- Change the width/height of the dendrogram
- Color dendrogram
- Change dendrogram spacing

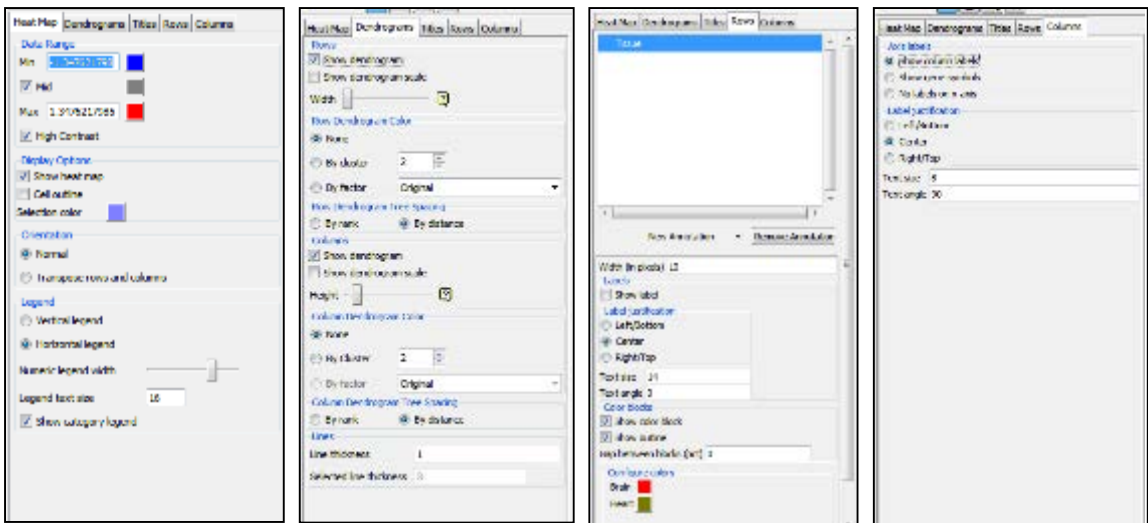
## Rows

- Change the width of annotation
- Check show label
- Change color
- Add new annotation

## Columns

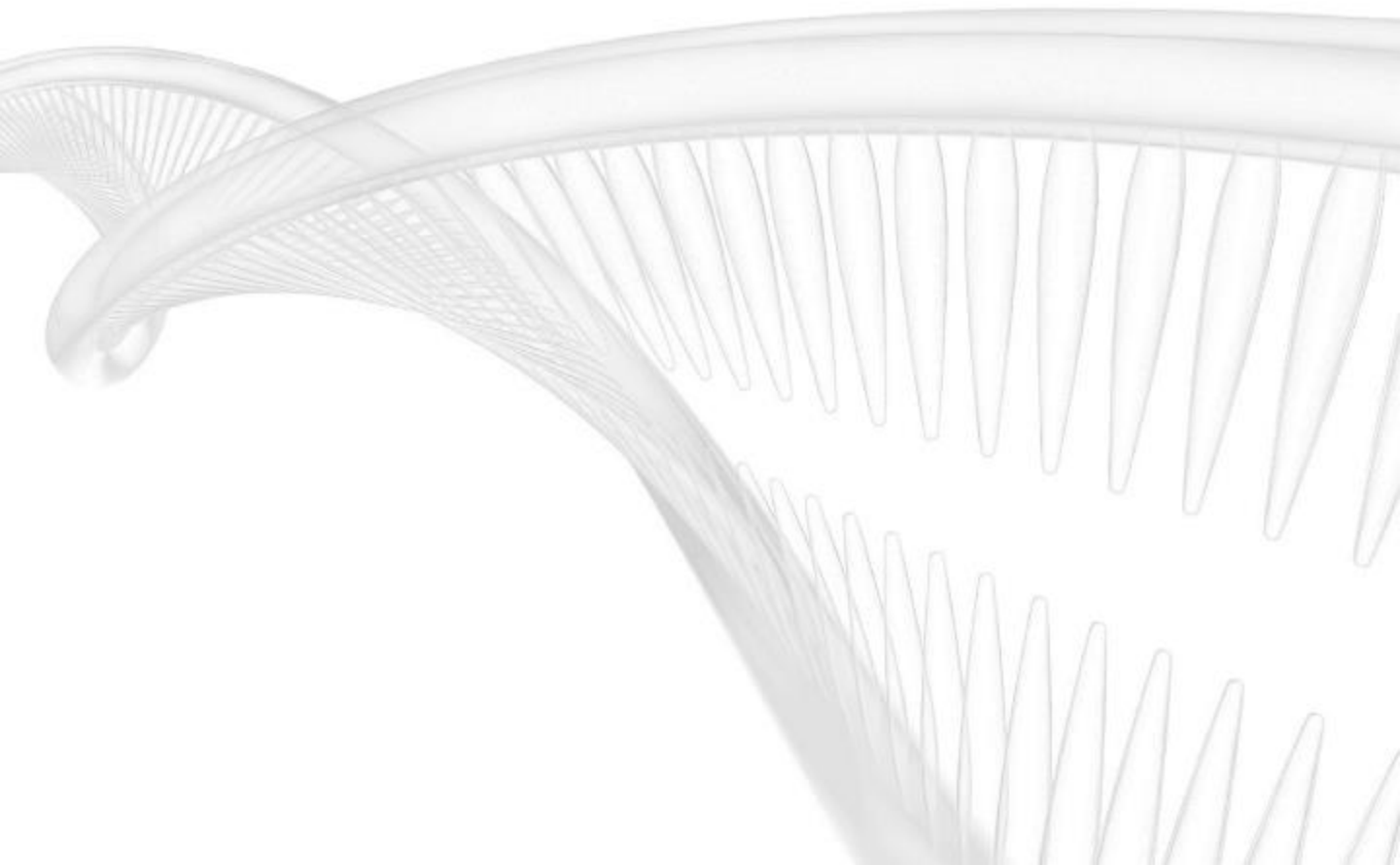
- Column header or miRNA name/gene symbol

**Mode:** mouse over, select, zoom, and flip



**Notes:**

# mRNA Data Analysis



# EXERCISE: Gene Expression Data Analysis

- Select **File > Open > Affy\_HuGeneST\_GeneIntensities**
- Select **File > Properties > Choose Annotation file**
- Choose the **Gene Expression** workflow to do the same analysis
- Generate a list of significant genes—Gene Brain vs. Heart

*This workflow is an exercise to perform on your own. Practice from QA/QC to create list steps that are the same as the miRNA workflow.*

ID	Name	2. Probe ID	3. Gene Symbol	4. Gene Symbol	5. Log2	6. Probe (Brain vs. Heart)	7. Probe (Brain vs. Heart)	8. Fold Change (Brain vs. Heart)	9. Fold Change (Brain vs. Heart)	10. Probe (Brain vs. Heart)	11. Probe (Brain vs. Heart)	12. Probe (Brain vs. Heart)
1.	15104	0340020	A2203000	SDHA	43220800	3.80279e-008	3.80279e-008	79.2005	79.2005	Brain up vs	124.45	50.
2.	11197	7861194	BC111907	MDG1	83318171	6.37295e-008	6.37295e-008	14.8102	14.8102	Brain up vs	4461.7	10.0
3.	35481	0138185	BC032270	DNAH1	80080270	6.27358e-008	6.27358e-008	18.8102	18.8102	Brain up vs	8512.85	25.
4.	4026	7862402	AJ274300	AUS	AF171000	3.21228e-008	3.21228e-008	144.105	144.105	Brain up vs	3462.53	70.
5.	3876	7861174	BC111907	MDG1	83318171	1.34115e-007	1.34115e-007	33.6767	33.6767	Brain up vs	1460	74.
6.	21267	0367985	A6273051	MC4R2	40220051	1.21158e-007	1.21158e-007	13.6919	13.6919	Brain up vs	7030.75	27.
7.	23474	0303020	AF254099	DOCK3	AF254099	1.20705e-007	1.20705e-007	11.6767	11.6767	Brain up vs	6022.3	26.
8.	74817	0788404	GA817106	QSOX1	64817106	1.16091e-007	1.16091e-007	74.6704	74.6704	Brain up vs	1444.76	23.
9.	13004	7862779	A6273050	UFAC10C	40220050	1.00894e-007	1.00894e-007	22.4913	22.4913	Brain up vs	5627.37	30.
10.	12529	7838251	U2394786	R3H37	67847286	1.11466e-007	1.11466e-007	3.02403	3.02403	Brain up vs	5324.31	41.
11.	14916	0303016	BC111907	MDG1	83318171	3.7084e-007	3.7084e-007	33.6767	33.6767	Brain up vs	1616.35	25.
12.	25480	0303077	BC032269	GABRG2	30003000	1.03246e-007	1.03246e-007	170.072	170.072	Brain up vs	4752.76	34.
13.	12438	7861194	BC111907	MDG1	83318171	3.63004e-007	3.63004e-007	23.9758	23.9758	Brain up vs	4752.23	34.
14.	01917	01178096	BC111907	GABRG2	30003000	5.91091e-007	5.91091e-007	31.4577	31.4577	Brain up vs	4756.8	37.
15.	02225	0110980	BC032269	RUP1B2	30003000	3.42827e-007	3.42827e-007	9.8332	9.8332	Brain up vs	4759.54	41.
16.	4026	7861194	BC111907	MDG1	83318171	4.47058e-007	4.47058e-007	33.6767	33.6767	Brain up vs	4757.34	37.
17.	13074	7861987	AF384837	GNOC2	49288837	5.9197e-007	5.9197e-007	18.3071	18.3071	Brain up vs	4130.67	26.
18.	22222	0307925	BC032244	THY1C1	30003046	3.47204e-007	3.47204e-007	0.0637030	-118.209	Brain down vs	4020.91	71.
19.	7472	7861714	GA274709	MDG1	83318171	4.49259e-007	4.49259e-007	11.8484	11.8484	Brain up vs	4817.7	38.
20.	51477	0368385	BC111907	MDG1	83318171	4.11218e-007	4.11218e-007	31.4713	31.4713	Brain up vs	3835.25	37.
21.	9675	7832044	U7137493	ARL10A	37137493	4.52446e-007	4.52446e-007	17.6229	17.6229	Brain up vs	3020.37	23.
22.	7472	0788404	GA817106	QSOX1	64817106	4.81078e-007	4.81078e-007	33.6767	33.6767	Brain up vs	3927.36	37.

Notes:

---



---



---



---



---

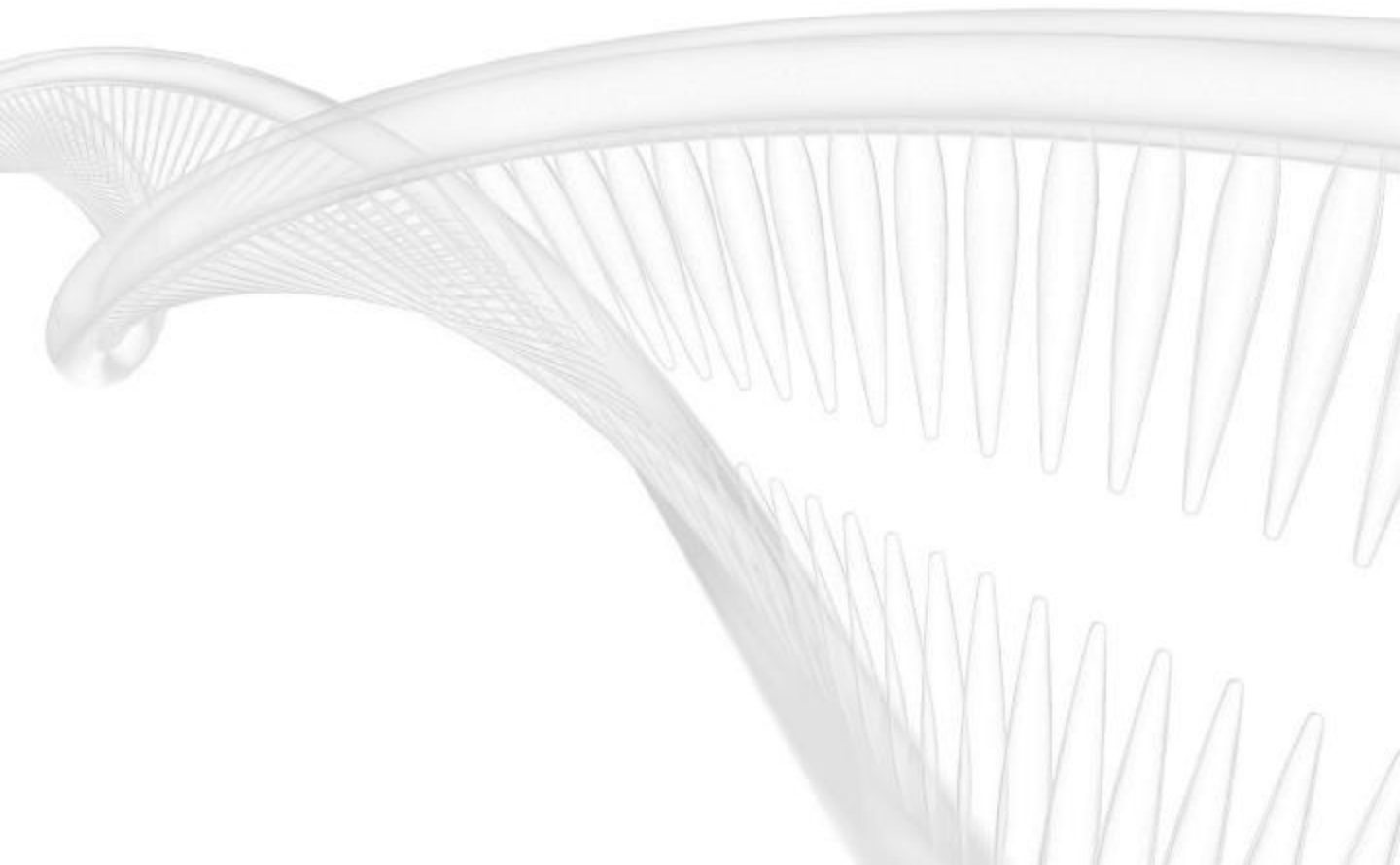


---



---

# miRNA and mRNA Integration



# Integration of miRNA with mRNA

---

miRNA and mRNA are linked by the target prediction databases: by default, TargetScan or microCosm (miRbase)

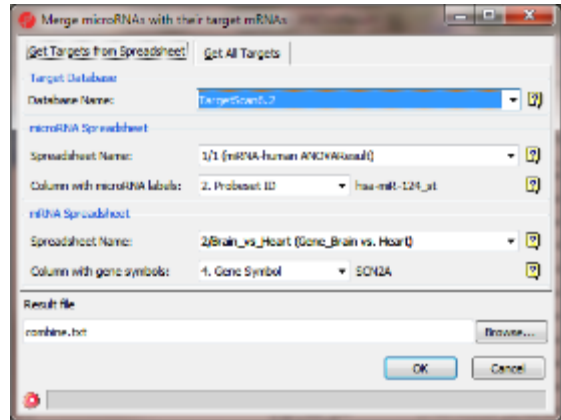
- Combine miRNAs with their mRNA targets
  - miRNA list
- Find overrepresented miRNA target sets
  - mRNA list
- Correlate miRNA and mRNA
  - Intensity value of miRNA and mRNA on the same sample

**Notes:** \_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_

# Combine miRNAs with mRNA Targets

**Question:** What are the putative targets of miRNAs of interest?  
(miRNA list is required)

- Click on **Combine miRNAs with Their mRNA Targets**
- Select the **Get All Targets** tab
- Choose **TargetScan 6.2**
- Select the **miRNA Brain vs. Heart** spreadsheet
- Select the **miRNA** label column
- Click **OK**



**Question:** Which genes might be regulated by miRNA of interest?  
(miRNA list and gene list are required)

- Click on **Combine miRNAs with Their mRNA Targets**
- Select the **Get Targets** from the spreadsheet tab
- Choose **TargetScan 6.2**
- Select the **miRNA Brain vs. Heart** spreadsheet
- Choose the miRNA label column
- Select the **Gene\_Brain vs Heart** spreadsheet
- Choose the **Gene Symbol** column
- Click **OK**

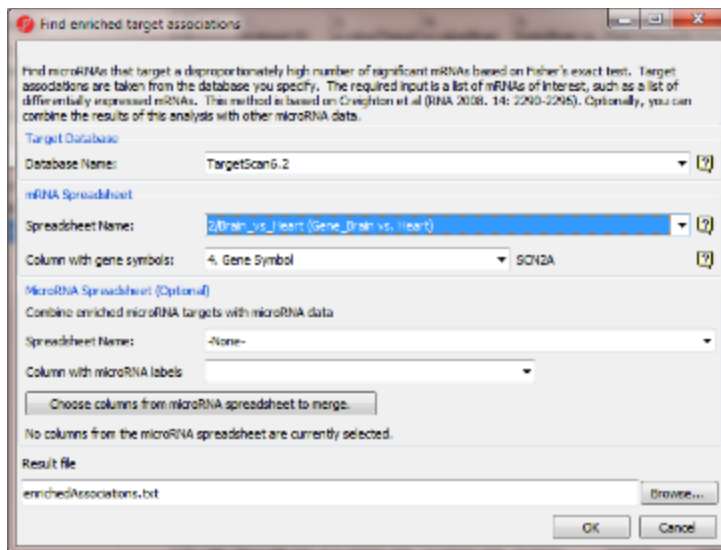
**Notes:** \_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_

# Find Overrepresented miRNA Target Sets

**Question:** Are the genes of interest the target of specific miRNAs?  
(Gene list is required)

- Click on **Find Overrepresented miRNA Target Sets**
- Choose the **TargetScan 6.2** database
- Select the **Gene Brain vs. Heart** spreadsheet
- Choose the **Gene Symbol** column
- Click **OK**

**Result:** List of miRNA with Enrichment p-values by Fisher's exact test.



**Notes:** \_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_



# Correlate miRNA and mRNA

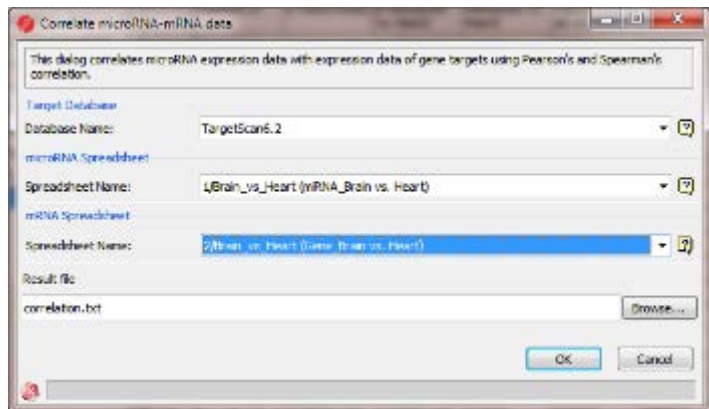
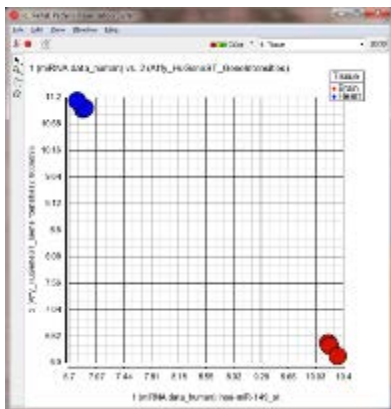
**Question:** Does miRNA abundance affect mRNA abundance?

- Intensity values of miRNA and mRNA are required
- Sample ID from two spreadsheets should match
- miRNA and mRNA lists are optional, used as a filter
- Correlate **the miRNA and mRNA data**
- Choose the **TargetScan 6.2** database
- Select the **miRNA Brain vs. Heart** spreadsheet
- Select the **Gene Brain vs. Heart** spreadsheet
- Choose the **Gene Symbol** column
- Click **OK**

**Note:** You can choose the miRNAdata\_human and Affy\_HuGeneST\_Gene Intensities spreadsheet to correlate all the miRNAs with their target mRNAs on the chips.

**Result:** p-value and correlation coefficient of miRNA with mRNA

- Right click on a row header to draw the **scatter plot**



**Notes:** \_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_

# Further Training

---

## Self-learning

- Help > Check for Updates
- Help > On-line tutorials
- Recorded webinars

## Regional Technical Support

- Email: [support@partek.com](mailto:support@partek.com)
- Phone: +1-314-878-2329

**Notes:** \_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_