mRNA and miRNA Data Analysis in Partek® Genomics Suite™

HANDS-ON TRAINING

University of Minnesota

May2014

Xiaowen Wang
Field Application Specialist
Partek Incorporated

Contents

Partek® Genomic Suite™ 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

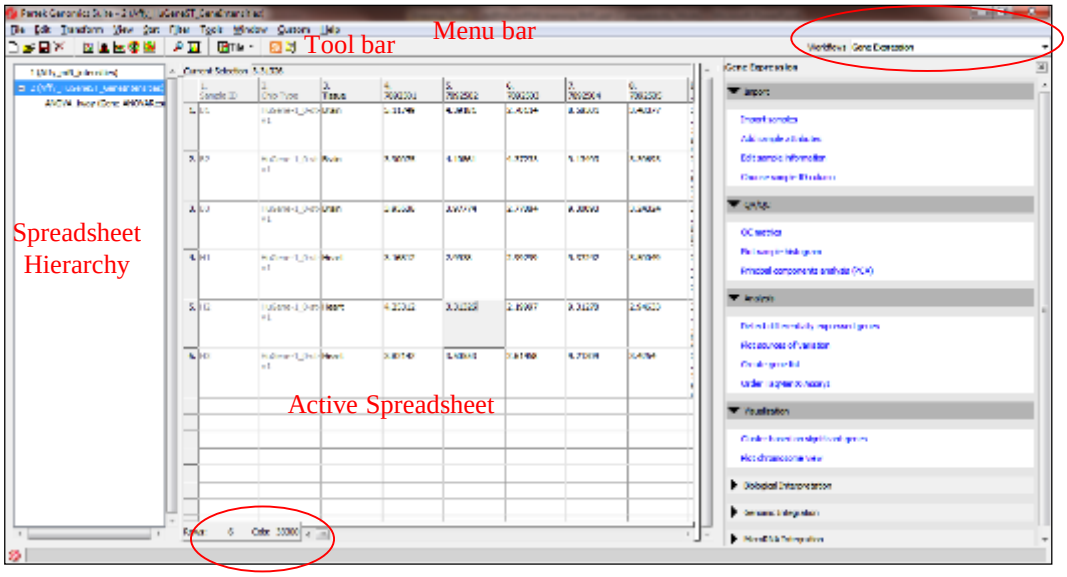
Tool bar: Accelerator buttons allow quick access to commonly used commands

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:_____

_____

_____

_____

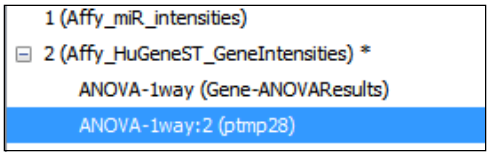
_____

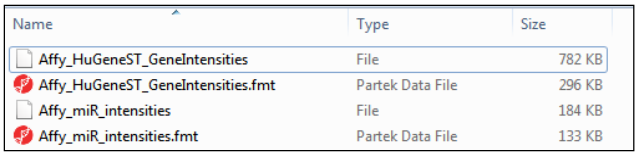
_____

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

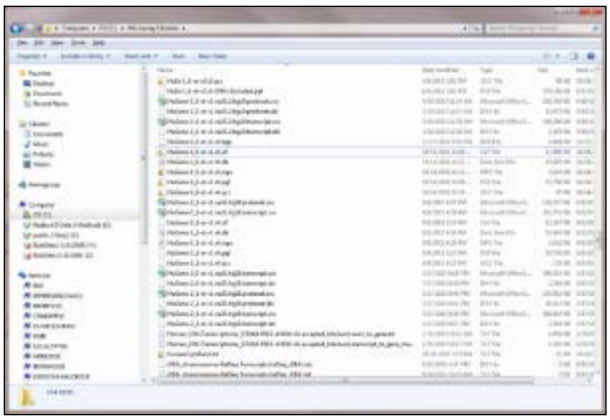





```
1 {C:\Partek Example Data\microRNA_data\WT Gene Array\gene.fmt}
1/ANOVA-1way {C:/Partek Example Data/microRNA_data/WT Gene Array/gene-ANOVAResults.fmt}
1/brain_vs_heart {C:/Partek Example Data/microRNA_data/WT Gene Array/mRNA-sig-list.fmt}
2 {C:/Partek Example Data/microRNA_data/miRNA/miRNA.fmt}
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:_____

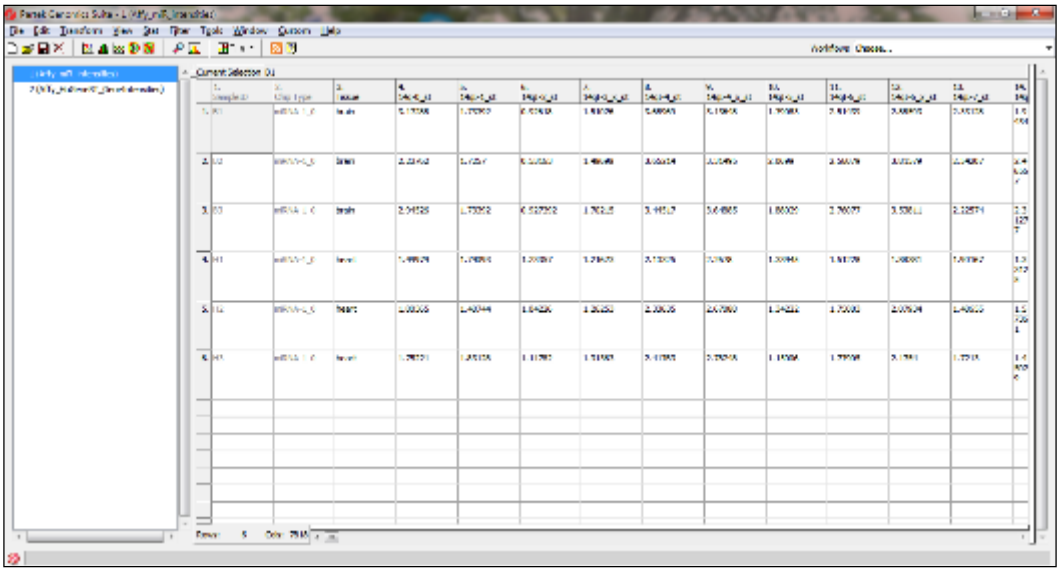
Training Data

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

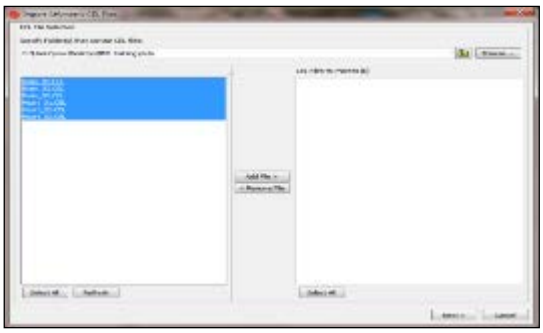


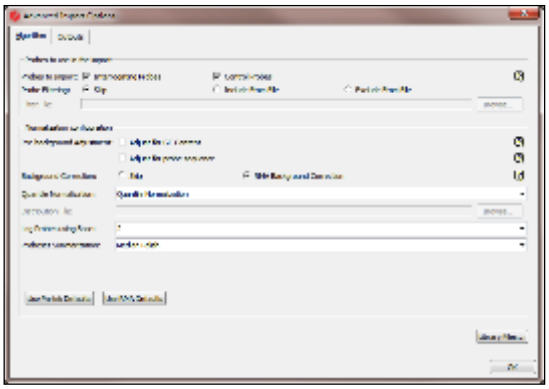
Notes:_____

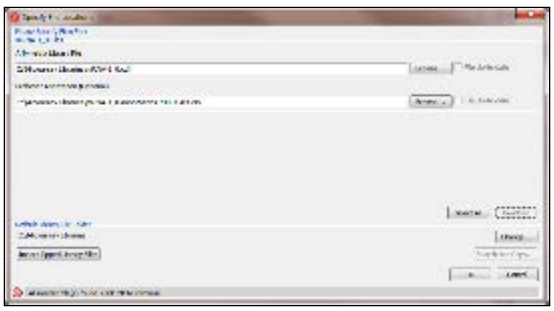
miRNA Data Analysis

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:_____

_____

_____

_____

_____

_____

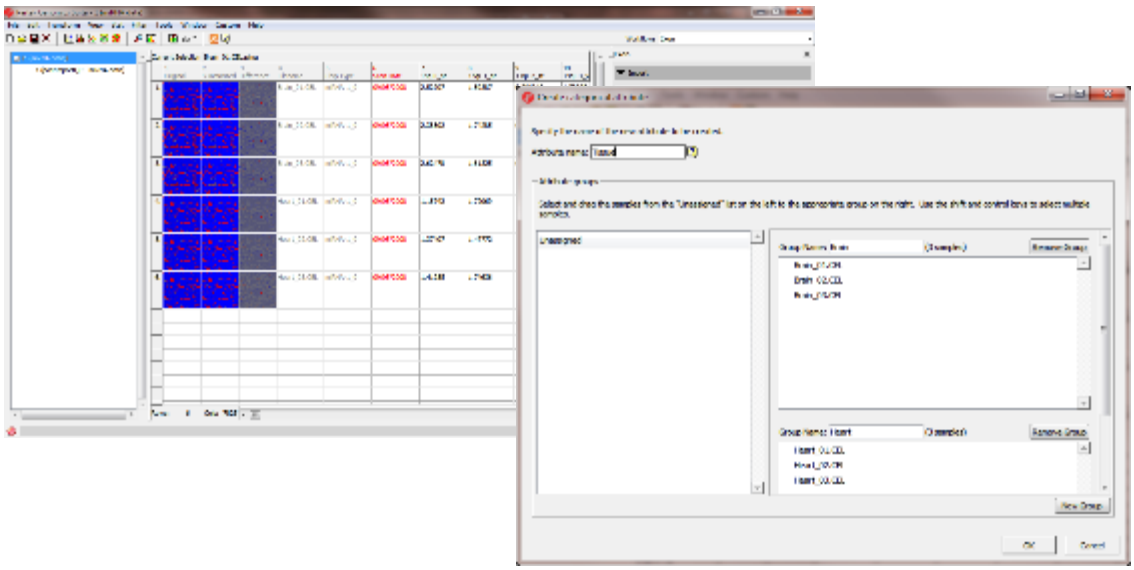
Add Sample Attributes

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

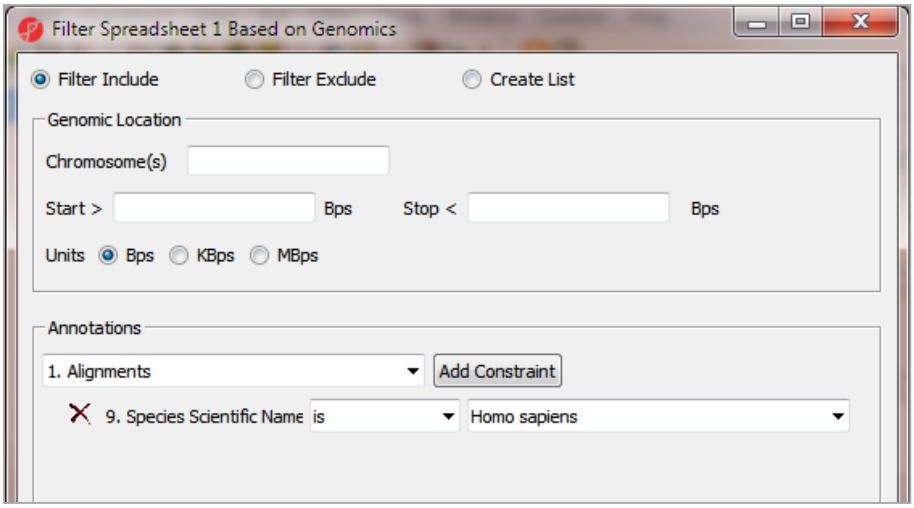
This chip contains miRNA from multiple species, but we only focus on human miRNA in this example.

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.

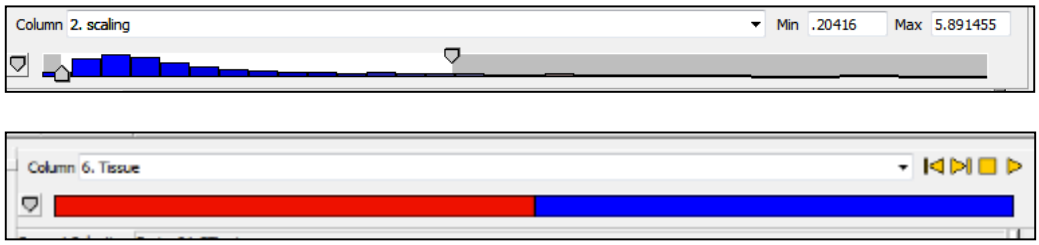
- **Select Filter > Filter Rows > Interactive Filter from the menu**
- Select a column from the drop-down list

  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

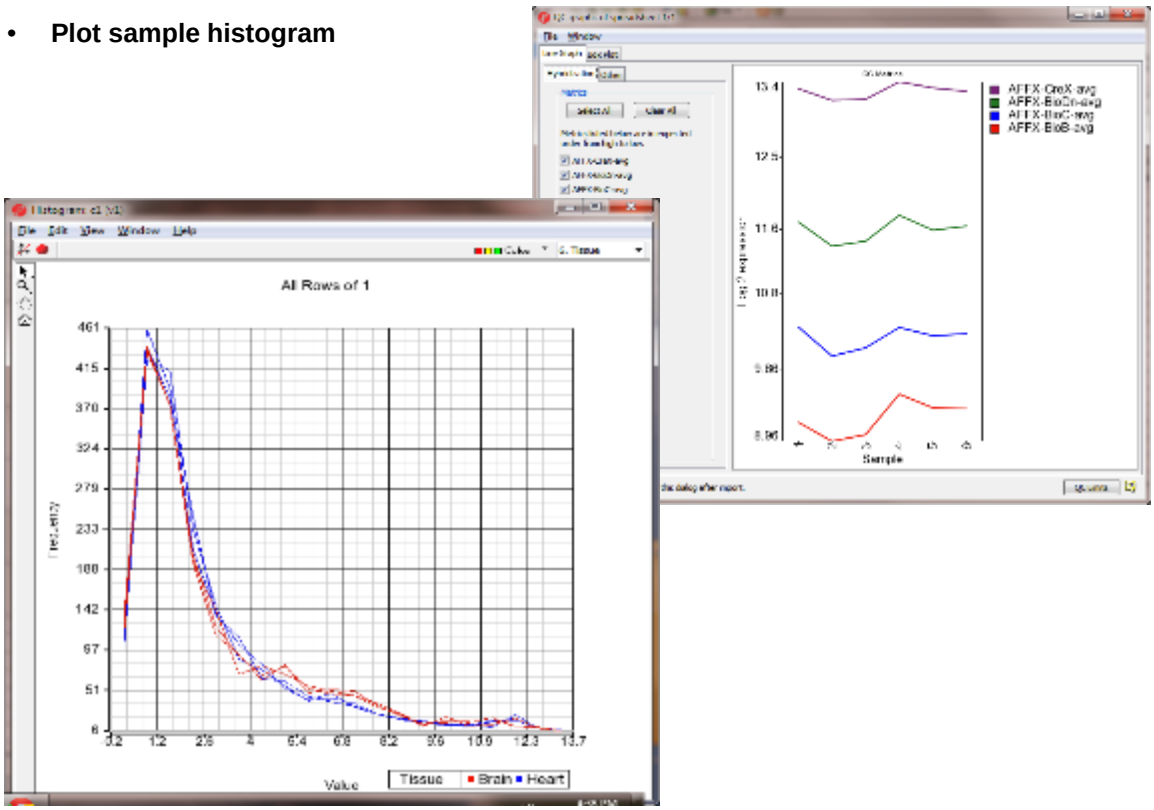
QA/QC is exploratory analysis, it is checking the preparation of the samples and indentify outliers

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

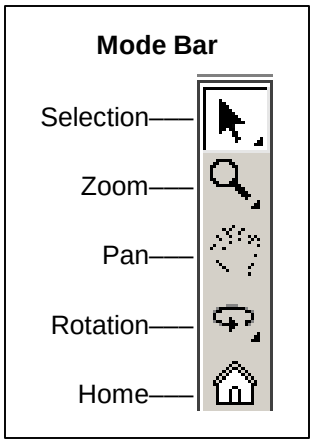
PCA scatter plot is another way to identify clustering patterns and outliers

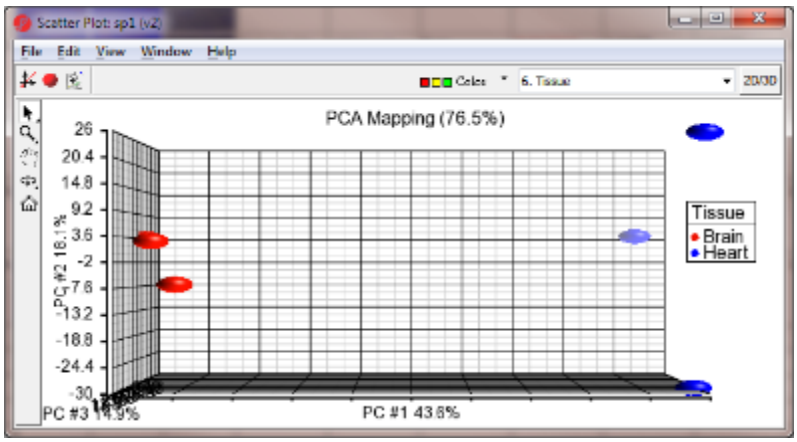
- Select **Principal components analysis** on the workflow

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:_____

_____

_____

_____

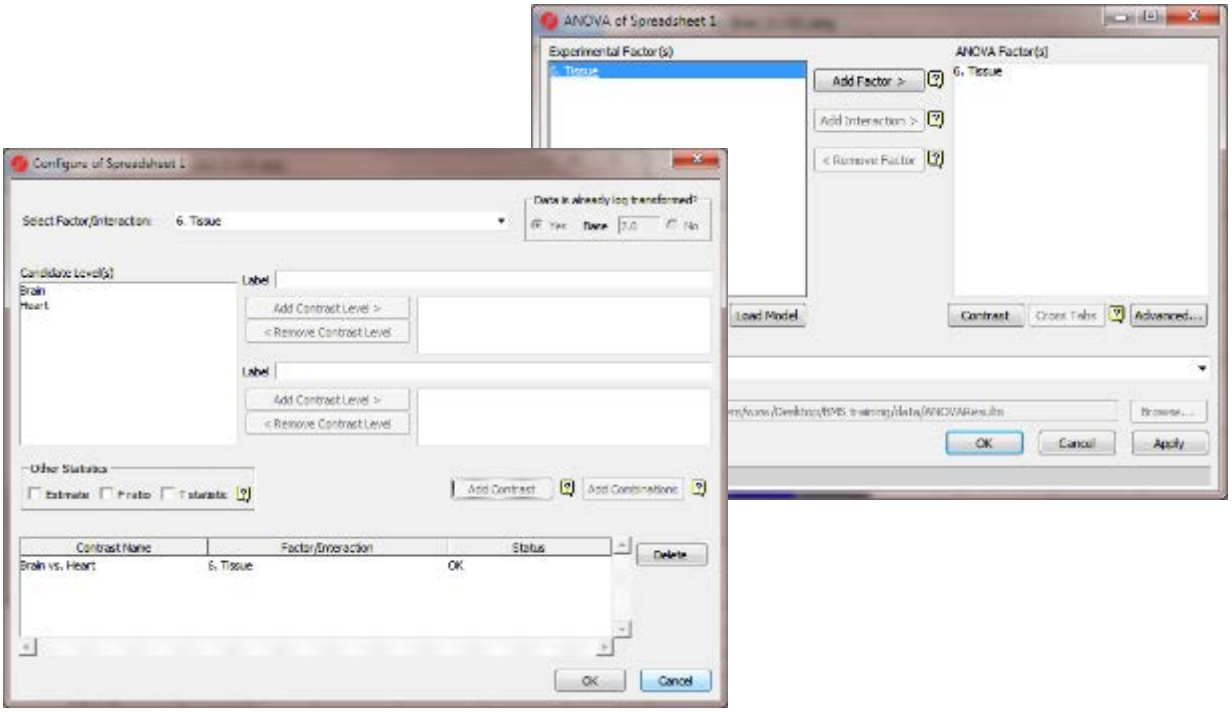
_____

_____

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

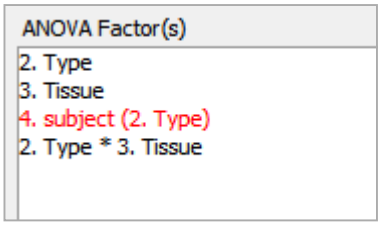
Different Types of ANOVA

- 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



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:_____

_____

_____

_____

_____

_____

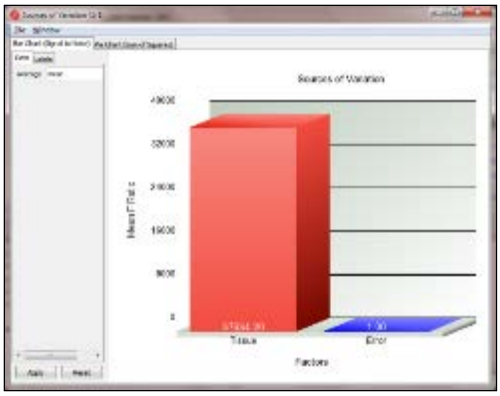
Result of ANOVA

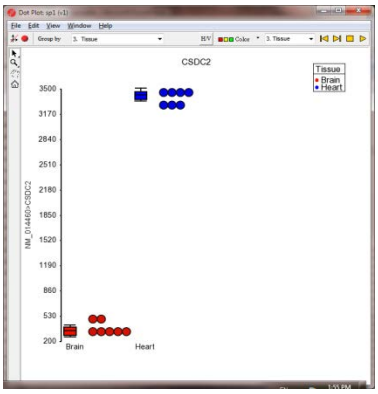
Each row is a miRNA with its p-value and fold change and any other statistical information. The spreadsheet is sorted by the first p-value column.

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





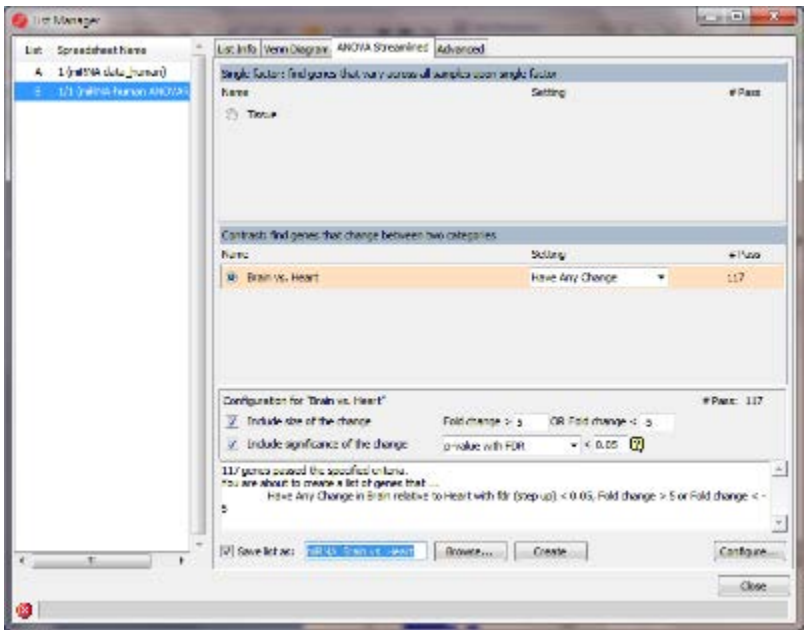
Notes:_____

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

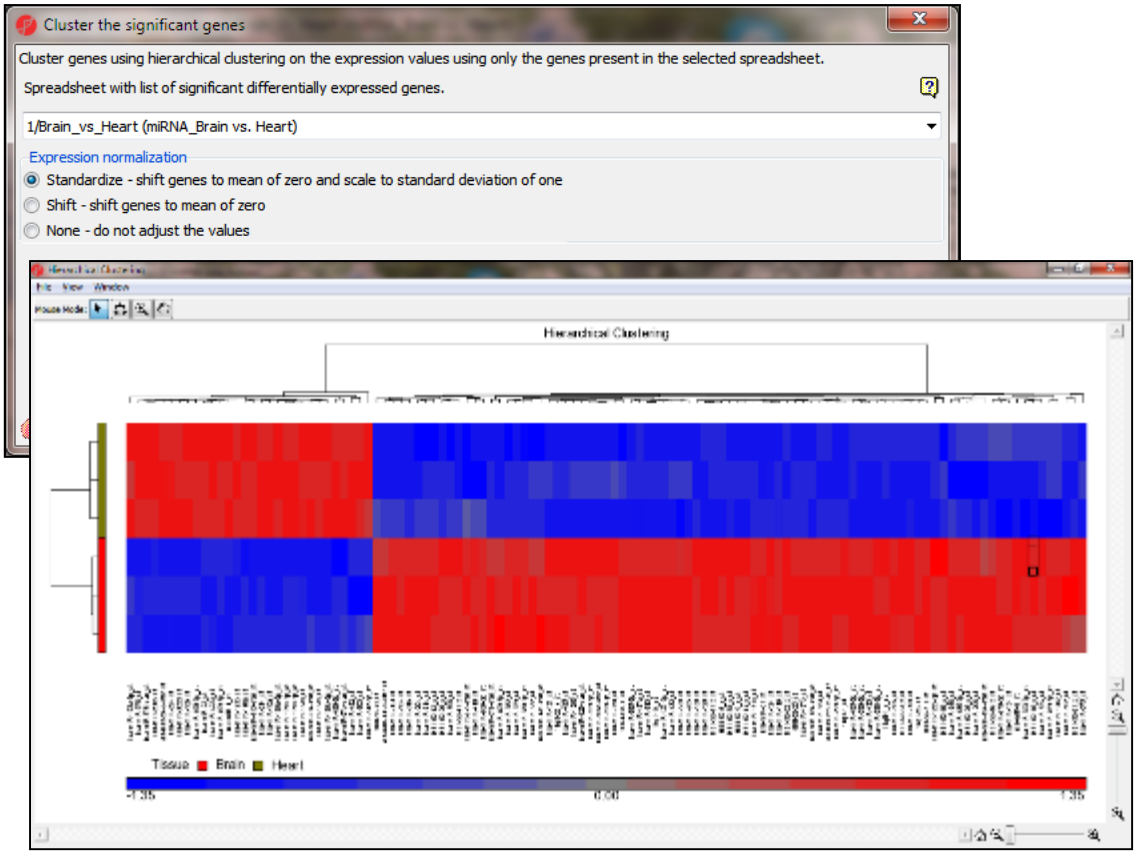


Notes:_____

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

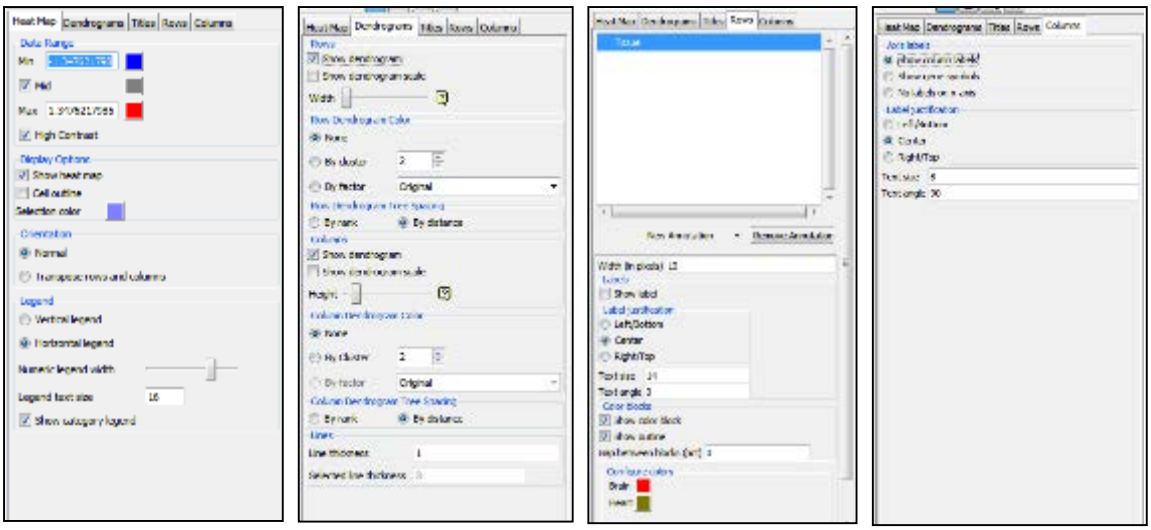
Rows

- 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



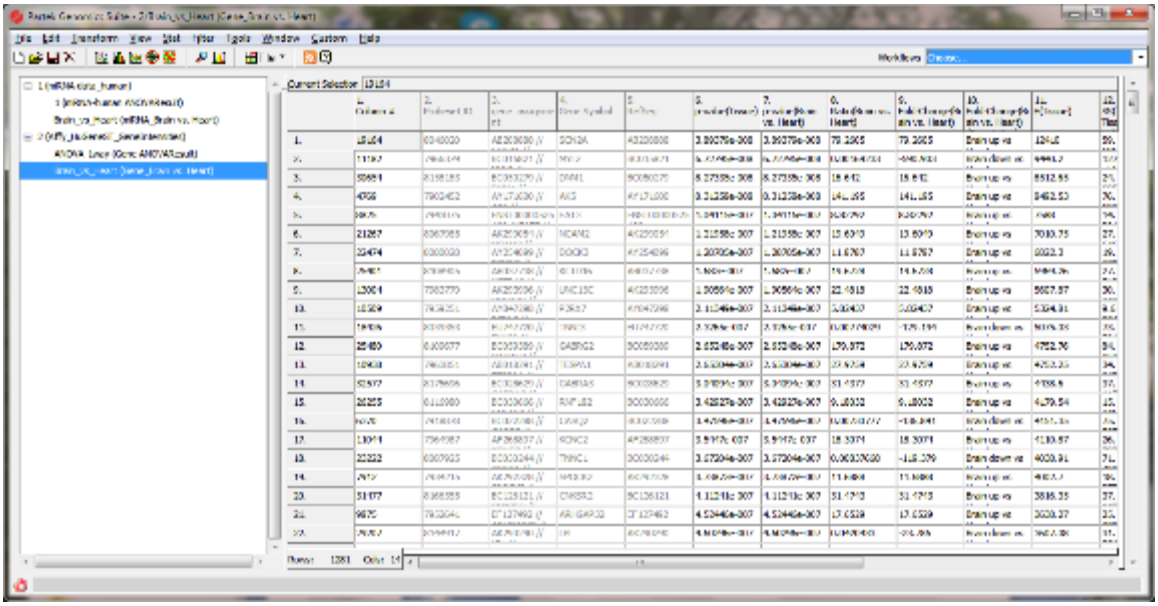
Notes:_____

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.



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 miNRA and mRNA
  - Intensity value of miRNA and mRNA on the same sample

Notes:_____

_____

_____

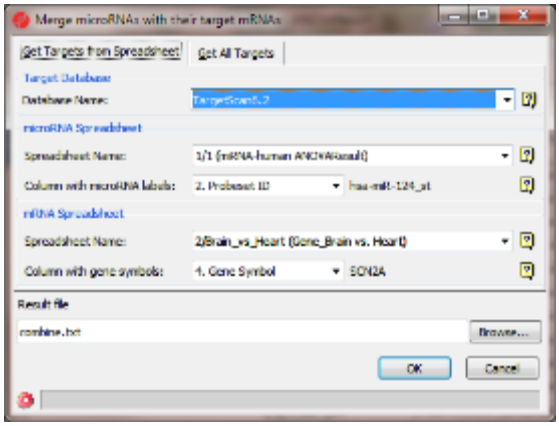
_____

_____

_____

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?
(miNRA 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:_____

_____

_____

_____

_____

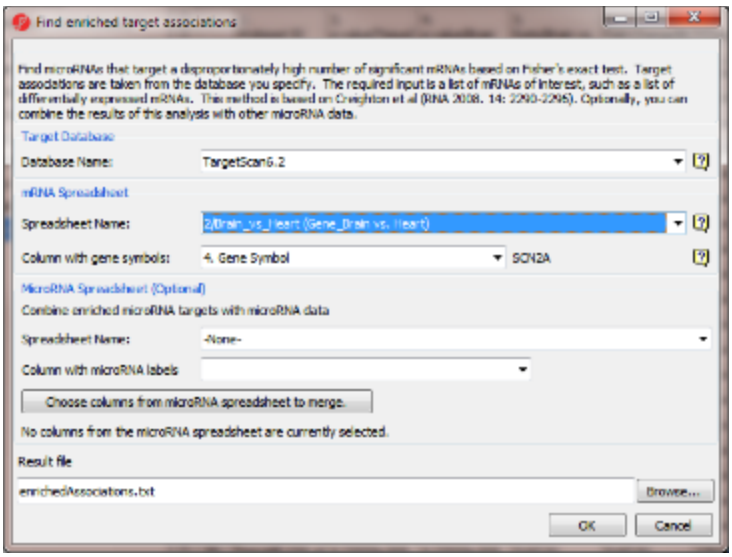
_____

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?

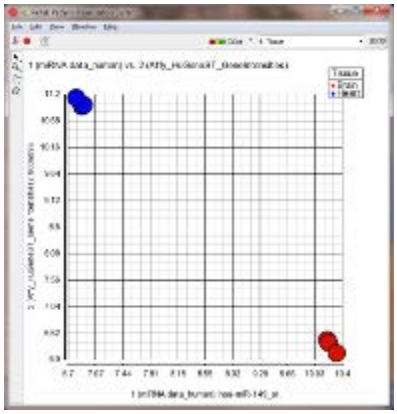
- Intensity values of miRNA and mRNA are required
- Sample ID from two spreadsheets should match
- miRNA and mRNA lists are optional, used as a filter

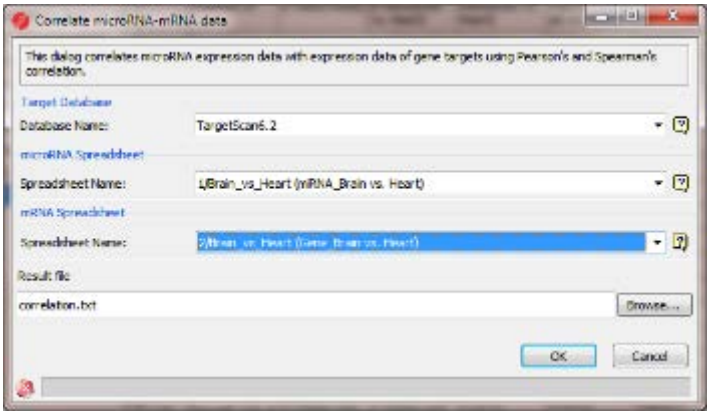
- Correlate **the miRNA and mRNA data**
- Choose the **TargetScan 6.2** database
- Select the **miRNA Brain vs. Heart** spreadsheet
- Select the **Gene Brain vs. Heart** spreadsheet
- Choose the **Gene Symbol** column
- Click **OK**

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
- Phone: +1-314-878-2329

Notes:_____

_____

_____

_____

_____

_____