Importing Data from Agilent's Feature Extraction Software

In this tutorial, the procedures for importing Feature Extraction result text files and GeneView data files into Partek will be described. These text files contain the intensity information from the Agilent microarray experiments as well as the annotation information associated with the Agilent microarray.

Importing Feature Extraction result files

- Import the Agilent text file into Partek by selecting **File > Import > Agilent.** Agilent outputs the data in a tab delimited format, which is the default
- There are two types of files importable into Partek, Feature Extraction result text files and GeneView text files. Select the appropriate file type as shown in the *Import Agilent Data* dialog box and select **Next** (Figure 1)
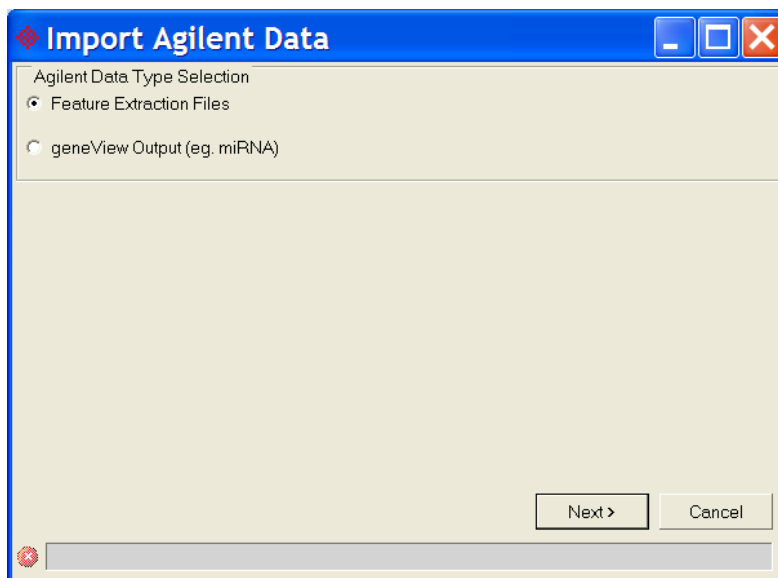

Figure 1: Configuring the Import Agilent Data dialog

- The second *Import Agilent Data* dialog box will appear as shown below (Figure 2)
- Specify the folder location with the Agilent data by selecting the **Browse** button and selecting the appropriate folder. All of the Agilent data files in the selected folder will appear in the left window
- Select the data files to import and select the **Add File** > button. The files selected for import will be moved to the *Agilent Files to Process* window
- Once all of the files for import are moved to the right window, select **Next**
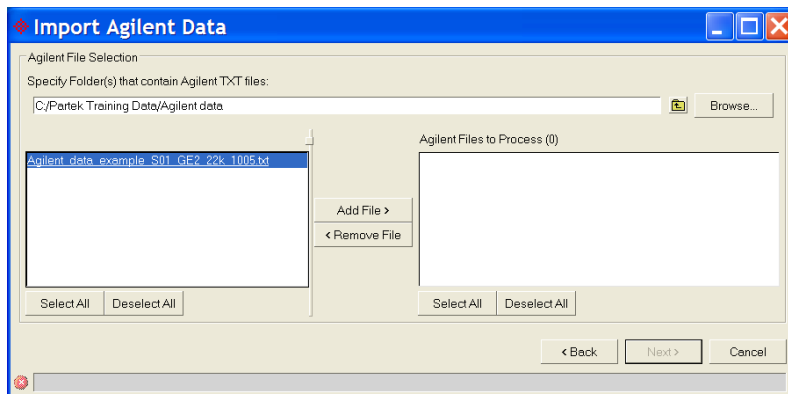
Figure 2: Specifying folders that contain Agilent text files

- The third *Import Agilent Data* dialog box will appear (Figure 3). Confirm that the column header that specifies the probe name in the Agilent data files is correct. In the example below, the column header selected is "ProbeName". If an incorrect column header is chosen to designate the *Probe Name*, select the appropriate header in the pull down menu
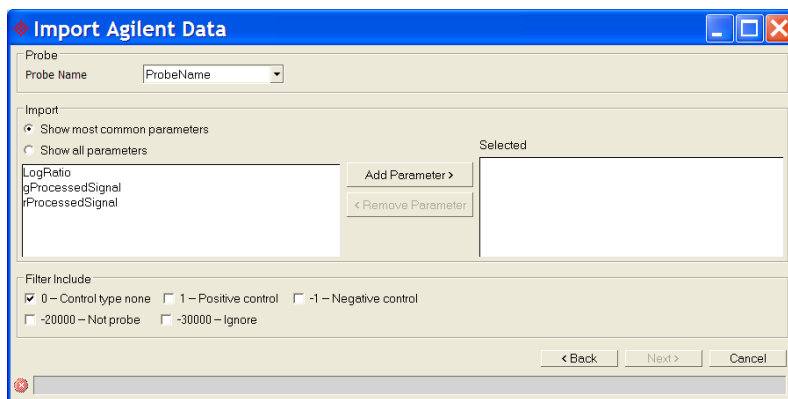

Figure 3: Selecting the data files to import

The *Import* panel specifies which parameter(s) will be imported into the Partek probe intensity spreadsheet. The most common parameters will be automatically shown, but the full list of parameters available can be viewed by selecting the radio button **Show all parameters**.

- Select the parameters to import by selecting the parameter and selecting the **Add Parameter >** button, this will move the selected parameters from the left window to the right window. You can add one or multiple parameters to be imported into Partek. For additional information regarding what the parameter information yields, see the Agilent Feature Extraction User's Manual

The *Filter Include* panel allows you to select which probes to include in the import. *0-Control Type None* is checked by default and signifies that all non-control probes

will be imported. If you would like to include the control probes, check the appropriate box to include them.

- Once all the selections have been made, select **Next** and the final dialog box will appear as shown in Figure 4 below
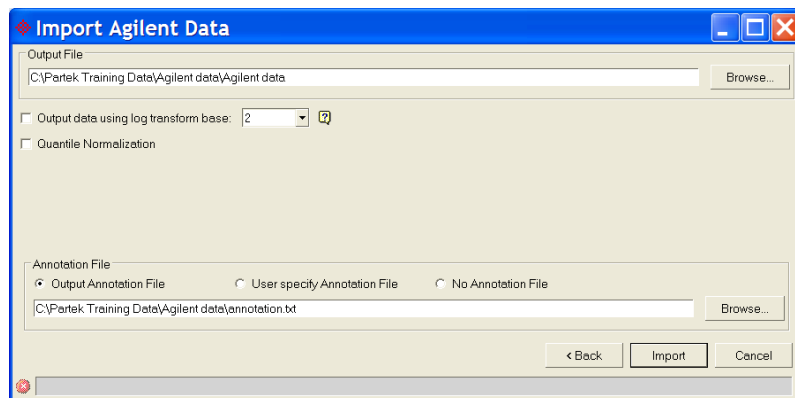

Figure 4: Selecting the Output and Annotation files

In the *Output File* panel, the directory location and filename for the Partek intensity spreadsheet that will be created is automatically generated. To change the directory location, use the **Browse** button to select a new location.

The data can be log transformed by checking the box, **Output data using log transform base**. The logarithmic base default is 2 or 10 depending on the parameter that is selected. You can specify the log base to be used by clicking on the drop down menu. Most microarray data is logged for analysis. While this is not a requirement to analyze microarray data, one of the assumptions of ANOVA, the assumption of normality, can be accounted for by logging the data.

Quantile Normalization may be performed on the data at import by checking the box assigned to *Quantile Normalization*. Quantile normalization will clean up the chip to chip variation by using the assumption that the probes on all chips should fit the same distribution model. Partek will calculate that distribution model from the data files and normalize the probes according to the calculated distribution model. This means that if you choose to quantile normalize your data that the distribution model for all our chips will be the same. If you want to keep the chip to chip variation intact, do not select quantile normalization to be performed upon import. (*Note:* Quantile Normalization can be performed after the data has been imported under **Transform > Normalization and Scaling**).

The *Annotation File* panel refers to the annotation information associated with the Agilent microarray used in the experiment. In Partek, a new annotation file can be generated from the Feature Extraction data files being imported by selecting the *Output Annotation File* radio button. If an annotation file has already been created, you can select the radio button associated with *User specify Annotation File* and browse to the appropriate file by selecting the *Browse* button. If there is no

annotation file, you should select the radio button associated with *No Annotation File.*

- Once the import functions and annotation files are selected, select **Import**

The data from the Feature Extraction result files or GeneView files will be imported into Partek.

Gene Expression and Copy Number Analysis Tips

With all of the parameters available in the Feature Extractions files, it can be confusing in choosing which parameters to import into Partek.

If the analysis to be performed is Gene Expression and there is no dye swap, loading the LogRatio is recommended. If the analysis to be performed is Gene Expression and there is a dye swap, we recommend that the gProcessedSignal and rProcessedSignal parameters be imported. It is important to include the dye information in the ANOVA model so that the dye effect can be removed when detecting differentially expressed genes.

If Copy Number analysis is to be performed, we recommend that the LogRatio parameter be imported. The LogRatio parameter can be converted into copy number space once the data has been imported by following these steps:

1) Change the property of the spreadsheet to *Copy Number* by selecting **File > Properties**. The *Configure Genomic Properties* dialog will appear. Under *Choose the type of genomic data*, select **Copy Number** in the drop down menu. Select **OK**. Partek now recognizes that this spreadsheet is associated with the Copy Number workflow
2) Select **Transform > Normalization & Scaling**. Select the *Rows (Response Variables)* tab. Select the *Normalization* tab
3) Antilog the base 10 data by selecting the **Antilog (y^x) + offset** radio button, select **Base** *10* from the drop down menu, and select **Apply**
4) Convert to a log base 2 by selecting the **Log (x+offset)** radio button, select **Base** *2* from the drop down menu, and select **OK**
5) To convert from a log2 base into copy number space, select **Transform > Adjust Copy Number**. Select *Convert Log2 Ratio to Copy Number*. Select **OK**. The log2 ratio values in your copy number spreadsheet will now be converted into copy number space.

End of User Guide

This is the end of the Agilent data import user guide. If you need additional assistance, you can call our technical support staff at +1-314-878-2329 or email *support@partek.com*.