

# Importing Custom Microarrays

Partek® Flow® accepts custom-made Affymetrix .CEL and Illumina .idat or AVGSIGNAL data generated from Genome Studio files as input. These require an accompanying probe annotation file.

This document describes how to create and import your own custom-microarray annotation files for use in Partek Flow.

- [Creating Custom probe\\_tab Files](#)
- [Creating Custom Illumina Annotation Files](#)
- [Creating Custom Intensity Data Files](#)
- [Importing Annotation Files into Library File Management](#)
- [Converting Custom Microarray Files into Aligned Reads](#)

## Creating Custom probe\_tab Files

It is advisable to obtain the annotation files for your custom chip directly from Affymetrix. If for some reason this is not possible, probe\_tab files can be created as a tab-delimited text file using any text editor. They must contain the following information in columns:

- **Probe id:** unique ID of the probe
- **Probe x:** Probe X position on the chip
- **Probe y:** Probe Y position on the chip
- **Probe sequence:** sequence of the probe, all IUPAC nucleotide codes are supported

The first line of the file must be the column headings listed below. Any mixture of upper and lower case is supported. If you are modifying this from another file, the *Probe id* column must be the first column in the file. The order of the other columns does not matter and any additional columns will be ignored unless they are in conflict with the column names above.

The *Probe id* can contain any unique identifier but should generally be the Affymetrix probe name or probe set ID. Note that in some chips, such as the Clariom D arrays, the *probe id* and *probe set id* contain different information. If that's the case, identify the column containing unique identifiers and use that as your first column. These names have to be composed exclusively of alphanumeric characters, underscores or hyphens.

The probe x and y positions correspond to the Affymetrix microarray chip position. Incorrect positions will result in the wrong sequence and probe name being used.

The probe sequence must be the sequence of the probe; all IUPAC nucleotide codes are supported.

Other columns may be present and will be ignored unless they match.

Save this file with one of the following file extensions:

- .probe\_tab
- .probe\_tab.gz
- .probe.tab
- .probe.tab.gz

## Acceptable variations to existing files

If other files are available to you with similar information, you may be able to minimally modify these files to make them compatible for import.

The probe set name column may omit "set" from the column name and may use "id" instead of "name". Valid examples are:

- probe set name
- probe name
- probe set id
- probe id
- probe set\_name
- probe set\_id

In cases where both *probe set id* and *probe id* are different, Partek Flow will use the identifier present in the first column. Thus, it is absolutely essential that the first column contains unique identifiers.

The probe position columns must be given for both x and y position. Valid examples of each are:

- probe\_x\_pos
- probe\_x
- probe x\_pos
- probe x
- probe\_y\_pos
- probe\_y
- probe y\_pos

The valid names for probe sequence columns are:

- probe\_sequence

- probe sequence

Remember to save your modified file with a .probe\_tab extension.

Note: Below is the preferred column header example: Probe\_id Probe\_x Probe\_y Probe\_sequence

## Creating Custom Illumina Annotation Files

It is advisable to obtain the annotation/manifest files for your custom chip directly from Illumina. For custom BeadChips, you will need to provide a .bgx annotation for .idat v1 or text files generated using Illumina Genome Studio. For .idat v3 data, please provide a .bpm file.

If for some reason this is not possible, .bgx files can be created as a tab-delimited text file using any text editor. The first line of the file must be [Probes]. The second line must be the following columns:

- Probe\_Id
- Probe\_Sequence

In the succeeding lines, list each probe ID and its corresponding probe sequence. Note that the probe IDs must match those present in your Genome Studio output or .idat v1 file. The probe sequence must be the sequence of the probe; all IUPAC nucleotide codes are supported.

Remember to save your modified file with a .bgx extension. We do not advise creating your own .bpm files. Please contact Illumina to obtain a copy of this file for your specific chip.

## Creating Custom Intensity Data Files

Various other analysis packages that process microarray data can output lists of probe IDs and intensity values. In these cases, Partek Flow can import intensity data files that can be used with an accompanying custom .bgx annotation (described in the previous section).

One intensity data file is needed per sample. They can be created using any text editor and save as a tab-delimited .txt file. Each file must contain the following columns:

- PROBE\_ID
- X.AVG\_Signal (where X is the sample name)

In the succeeding rows, list each probe ID and its corresponding signal intensity. Save the resulting file with the same sample name (i.e. *X.txt* where X is the sample name).

Note that the probe IDs must match those present in the accompanying .bgx annotation file. These .txt files must be formatted in either windows- or linux-compatible text file formats. Text files with just carriage returns, such as those generated by Macs are not compatible.

Import the .txt files using the *Data* tab in your project (**Data>Import data>Automatically create samples from files**).

## Importing Annotation Files into Library File Management

The annotation files can then be uploaded into Partek Flow by going to **Settings>Library File Management**, selecting the **Microarray library files** tab and clicking the **Add probe sequence** button (Figure 1).

Genomic library files

Microarray library files

+ Add probe sequence

Chip name	URL	Actions
RaGene-1_0-st-v1	http://media.affymetrix.com/analysis/downloads/na23/wtgene/RaGene-1_0-st-v1.probe.tab.zip	✕
miRNA-2_0		✕
miRNA-3_0		✕

Update Library Index

Last updated: 26 Oct 2015, 11:06 AM CDT *i*

Add probe sequence

✕

Chip name

Other / Custom

Create

Figure 2. Adding your custom probe tab file in Library File Management

## Converting Custom Microarray Files into Aligned Reads

Import your microarray files via the **Data** tab. Once the data has been uploaded, select the **Microarray intensity** data node and convert it to *Aligned reads* using your aligner of choice.

When prompted to select the probe sequence file, click the **Chip name** drop-down menu and select the chip name for your custom chip.

## Additional Assistance

If you need additional assistance, please visit [our support page](#) to submit a help ticket or find phone numbers for regional support.



Your Rating:  Results:  36 rates