

# GenePattern

## CaArray2ImportViewer Documentation

**Module Name:** CaArray2ImportViewer  
**Description:** A visualizer to import data from caArray into GenePattern.  
**Author:** Jared Nedzel, [gp-help@broadinstitute.org](mailto:gp-help@broadinstitute.org)

### Summary

The CaArray2ImportViewer imports data from caArray into the GenePattern .gct or .cls file formats. See <https://array.nci.nih.gov/caarray/home.action> for more information about caArray.

This module takes the name or identifier of a caArray experiment as input. The module then retrieves the required derived data or annotations to create the .gct or .cls files.

This module has been tested with caArray 2\_4\_1\_GA.

### Usage/Example

When you launch the CaArray2ImportViewer module from GenePattern, you will see a series of screens that allow you to choose how your caArray data will be handled. The CaArray2ImportViewer uses a wizard format, so if at any time you want to change the choices you have made, you can use the *Back* button. The details of these screens are outlined below.

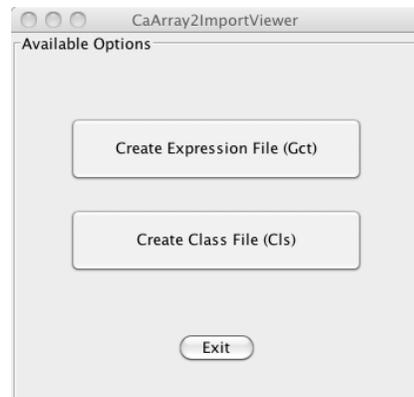
### Available Options:

Choose which type of file to generate:

- .GCT file to create an expression dataset
- .CLS file to create a class label file

To read more about these file formats, go [here](#).

Click the appropriate button to enter the sequence of steps to generate your file.



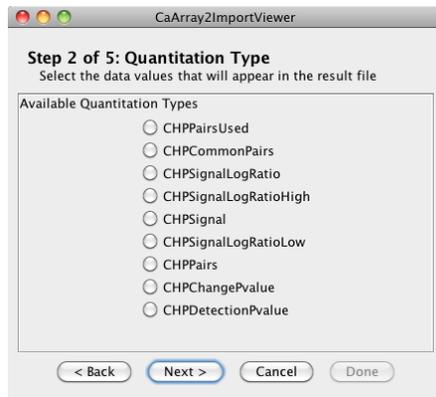
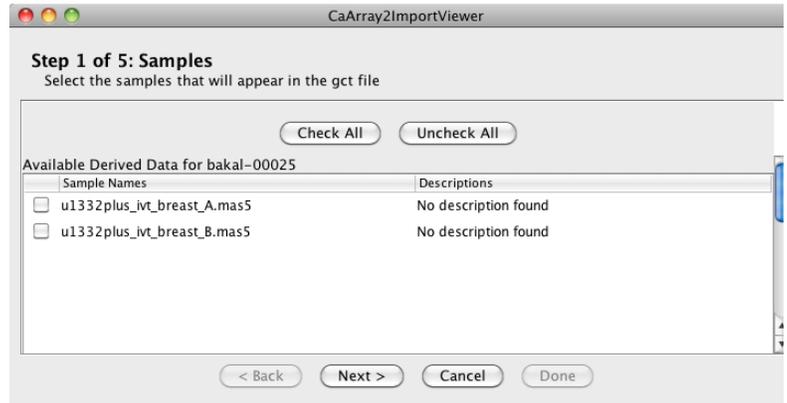
# GenePattern

## Generating a .GCT File

### Step 1 - Samples:

This screen allows you to select the samples you want to include in the output file. This screen shows all of the samples from the experiment in caArray. The sample name shown is the name of the hybridization in caArray.

Check the desired samples and click *Next*.

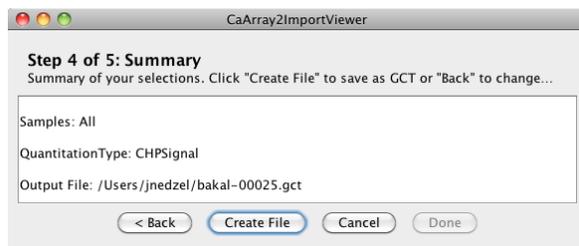
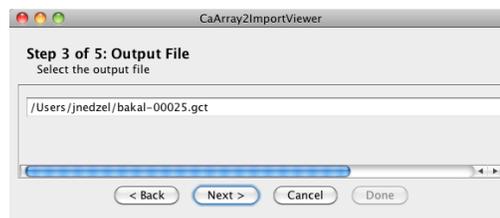


### Step 2 - Quantitation Type:

Select the data values that will appear in the .gct file. Typically, a user selects a **quantitation type** that contains intensity values, signal values, or ratio values. For single-channel data, select the quantitation type that contains the signal data for the probe sets. In this example using Affymetrix data, the CHPSignal value would normally be selected. For two-channel data, select the quantitation type that corresponds to the ratio values.

### Step 3 - Output File:

Choose the path and filename of the output file. Use the default filename or click *Choose File...* to specify an alternate file and click *Next*.



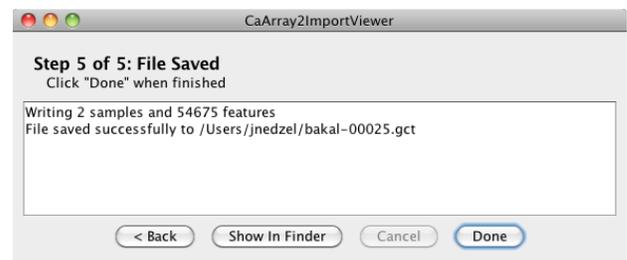
### Step 4 - Summary:

Shows a summary of your selections.

Click *Back* if you would like to change any of the values or click *Create File* to continue.

### Step 5 - File Saved:

This window indicates that you have successfully imported your caArray data. Click *Done* to close the wizard. The Available Options window (see above) will still be visible, allowing you to create additional expression and/or class files.



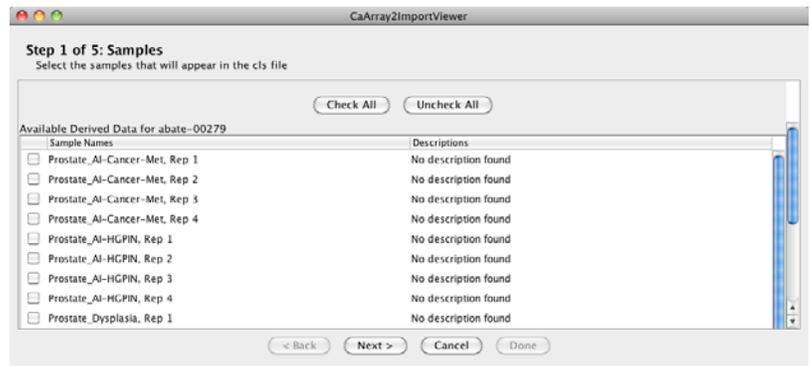
# GenePattern

## Generating a .CLS File

### Step 1 - Samples:

This screen allows you to select the samples to use in the output file. Check the desired samples and click *Next*.

If this .cls file is being created for use with a .gct file, you should choose the same samples used for the .gct file.



# GenePattern

## Step 2 - Assign Classes:

This window lets you assign classes to a set of samples. This means that you are selecting phenotype or clinical labels for your samples. In the case illustrated here, the class "wild type" or "knockout" labels have been assigned to each of the samples.

First, select a **label** that represents the set of phenotypes/clinical attributes you want to use to label your samples. You will use the drop-down menu under *Display labels from column* to do this. Each of these labels comes from the annotations available for the experiment in caArray.

You will see a list of possible phenotypes or clinical attributes corresponding to the label type you have selected. These are the **classes** you will use to categorize the samples in your dataset.

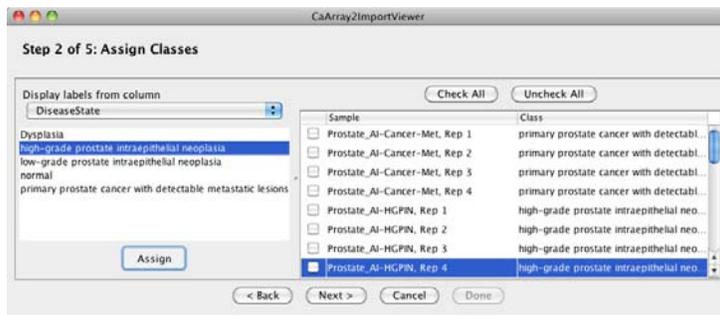
For each class you are assigning:

- Select the class from the list that appears in the left-hand window.
- Select the checkboxes for the samples to be assigned this class in the right-hand window.
- Click *Assign*. This assigns the class to your selected samples and clears your selections in the left and right windows so you can assign another class to another group of samples.

(Note that classes can also be edited manually in the *Class* column.)

NOTE: Each sample can only be assigned to one class.

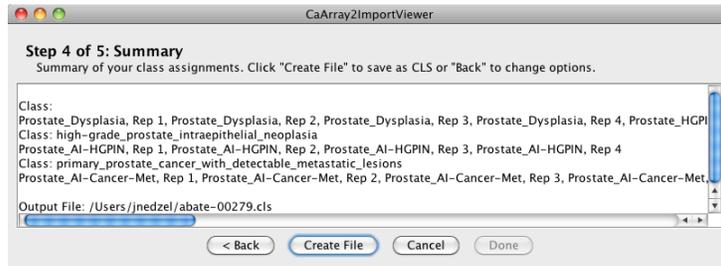
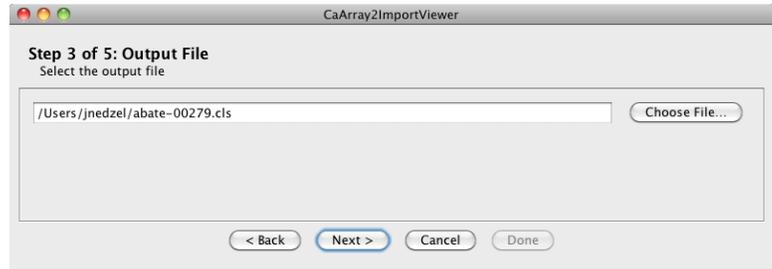
When you have assigned classes to all of your samples, click *Next*.



# GenePattern

## Step 3 - Output File:

Choose the path and filename of the output file. Use the default filename or click *Choose File...* to specify an alternate file and click *Next*.

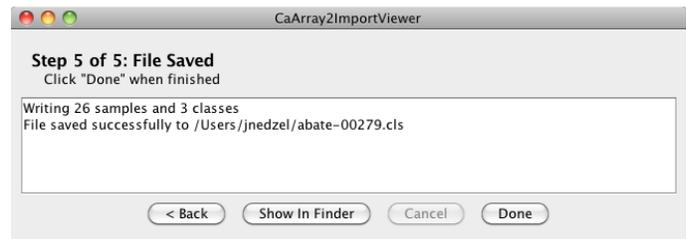


## Step 4 - Summary:

Shows a summary of your selections. Click *Back* if you would like to change any of the values or click *Create File* to continue.

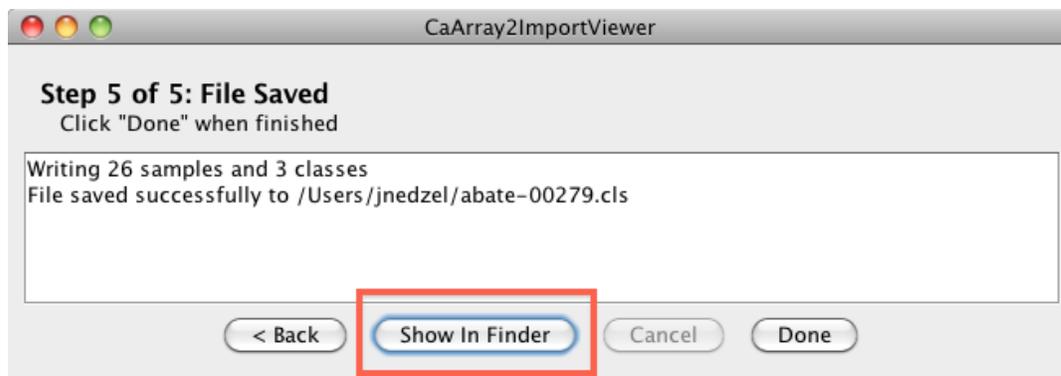
## Step 5 - File Saved:

This window indicates that you have successfully imported your caArray data. Click *Done* to close the wizard. The Available Options window (see above) will still be visible, allowing you to create additional expression and/or class files.



## Other Capabilities

On a Macintosh that is running Java 6, the visualizer also allows you to see the newly-created .gct or .cls file in the Macintosh Finder application.



# GenePattern

## Parameters

Name	Description
experiment	Title or public identifier of the experiment in caArray that data is to be imported from.
url	URL to the caArray repository. The default is to use the public caArray instance hosted at the NCI ( <a href="https://array.nci.nih.gov:8080">https://array.nci.nih.gov:8080</a> )

## Platform Dependencies

**Module type:** Visualizer  
**CPU type:** any  
**OS:** any  
**Language:** Java