

## PCAViewer Documentation

**Description:** Principal Components Analysis Viewer  
**Author:** Marc-Danie Nazaire, Joshua Gould (Broad Institute),  
[gp-help@broad.mit.edu](mailto:gp-help@broad.mit.edu)  
The Institute for Genomics Research  
**Release:** 5.0

The PCAViewer displays the results from principal component analysis (the PCA module). It is adapted from TIGR's MultiExperiment Viewer (<http://www.tigr.org/software/tm4/mev.html>).

### Summary

As discussed in Raychaudhuri, et al (2000: "A PCA analysis of DNA microarray data can consider the genes as variables or the experiments as variables or both. When genes are variables, the analysis creates a set of 'principal gene components' that indicate the features of genes that best explain the experimental responses they produce. When experiments are the variables, the analysis creates a set of 'principal experiment components' that indicate the features of the experimental conditions that best explain the gene behaviors they elicit.

The viewer provides two tabs, Components and Projection. Use the Components tab to plot the eigenvectors of selected principal components, as shown in [Figure 1](#). Use the Projection tab to display a 2D plot that projects experiments/genes onto two principal components, as shown in [Figure 2](#), or a 3D plot that projects them onto three principal components, as shown in [Figure 3](#). (The viewer projects experiments or genes onto principal components depending on whether the principal component analysis considered the experiments or genes as variables, respectively.)

The viewer provides the following menu commands:

File>Save As	Save the plot to an image file.
File>Export Plot Table	Exports the plot table to a tab delimited text file.
Help>Help Contents	Display information about the viewer and how to use it.

### Components tab

In the Components tab, the left pane lists each principal component. Select one or more principal components to plot their eigenvectors. The % Var field at the bottom of the list sums the % Variation of the selected principal components.

The right pane plots the eigenvectors of the selected principal components. The table under the plot lists the individual values that comprise each vector.

Right-click on the plot to display the following menu:

Properties	Set display options for the plot.
Save As	Save the plot to an image file.
Print	Print the plot.
Zoom in	Zoom in on either axis or both axes. Alternatively, click and drag to zoom in on an area of the plot.

# GenePattern

Zoom out

Zoom out on either axis or both axes.

Auto range

Reset the default zoom level for either axis or both axes.

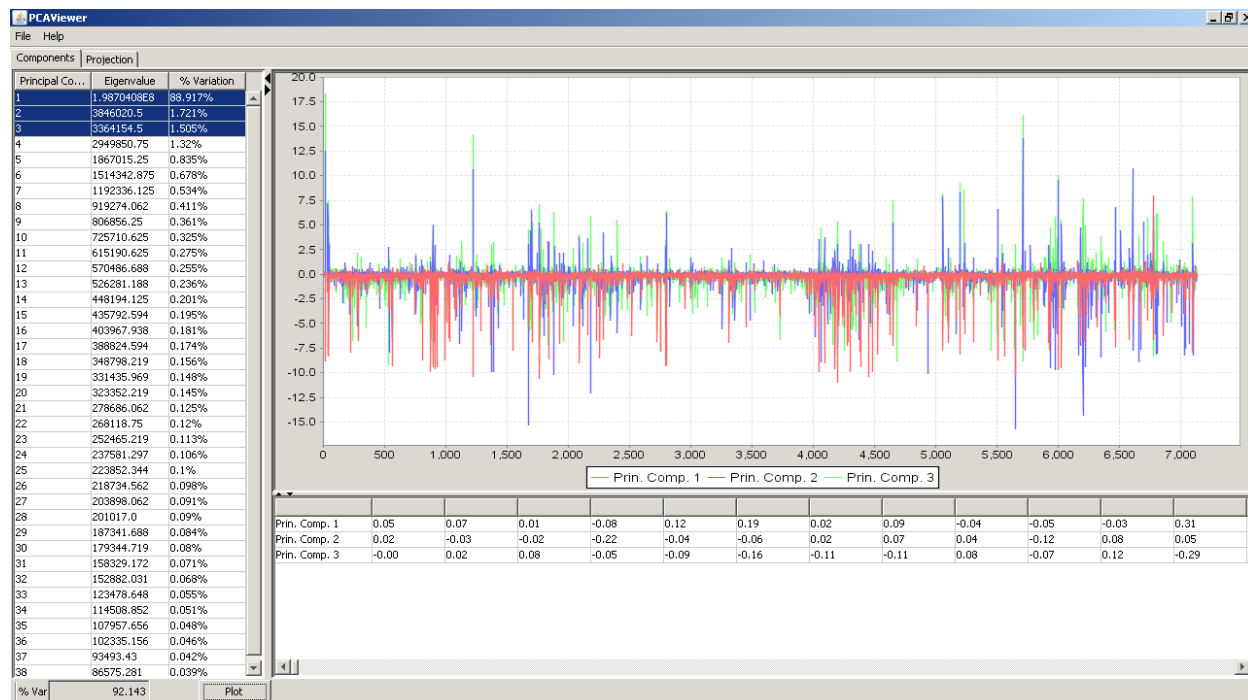


Figure 1. Components tab plotting eigenvectors for four principal components

## Projection tab

In the Projection tab, the left pane lists each principal component. To project the genes or experiments (whichever the principal component analysis considered as variables) onto principal components: select two or three principal components and click Plot. The % Var field at the bottom of the list sums the % Variation of the selected principal components.

The right pane shows a [2D plot](#) of two principal components or a [3D plot](#) of three principal components. If a cls, sample info file, or featureset file is provided then the points in the projection are colored according to the groups specified. **Note:** Java 3D version 1.3.1 or greater must be installed (<https://java3d.dev.java.net/binary-builds.html>) in order to display a 3D plot. The table under the plot lists the values that are being plotted in the projection of each gene/experiment on the selected principal components.

To manipulate a 2D plot:

- Hover over a point to display its details.
- Click and drag or click *Zoom In* to zoom in on an area of the 2D plot.
- Click and drag or click *Zoom Out* to zoom out on an area of the 2D plot.
- Click *Reset Plot* to reset the default zoom level.
- Click *Clear Plot* to clear the right pane.
- Select the *Display tick marks* checkbox to show/hide tick marks along the axes.

# GenePattern

- Right-click on the plot to display the following menu:

Properties	Set display options for the plot.
Save As	Save the plot to an image file.
Print	Print the plot.
Zoom in	Zoom in on either axis or both axes.
Zoom out	Zoom out on either axis or both axes.
Auto range	Reset the default zoom level for either axis or both axes.

To manipulate a 3D plot:

- Hover over a point to display its details.
- Alt+click and drag down on the vertical axis or click *Zoom In* to zoom in.
- Alt+click and drag up on the vertical axis or click *Zoom Out* to zoom out.
- Click and drag an axis to rotate the plot.
- Right-click and drag an axis to move the entire plot.
- Click *Reset Plot* to reset the plot to its default aspect.
- Click *Clear Plot* to clear the right pane.
- Select the *Display tick marks* checkbox to show/hide tick marks along the axes.
- Select the *Display grid lines* checkbox to show/hide grid lines on the plot.
- Right-click on the plot to display the following menu:

Save As	Save the plot to an image file.
---------	---------------------------------

How to assign groups in a plot:

1. Select a point by clicking on a point in the plot or clicking on a row in the projection table.
2. Click on *Set Group* and provide a name and color for the new group.
3. To assign additional points to a recently created group, repeat step 1 and click on the colored icon to the left of *Set Group*.
4. To assign additional points to a previous group that is not the recent one, repeat step 1 and step 2 but provide the name of the previous group.

The new group assignments for the point(s) should now appear under the Group column in the plot table.

How to remove a new group from the plot:

1. Click on *Remove Group* and check the box next to the group(s) to remove. Then click OK.

The group assignments for the points in the groups that were removed should now be reset to their original groups.

# GenePattern

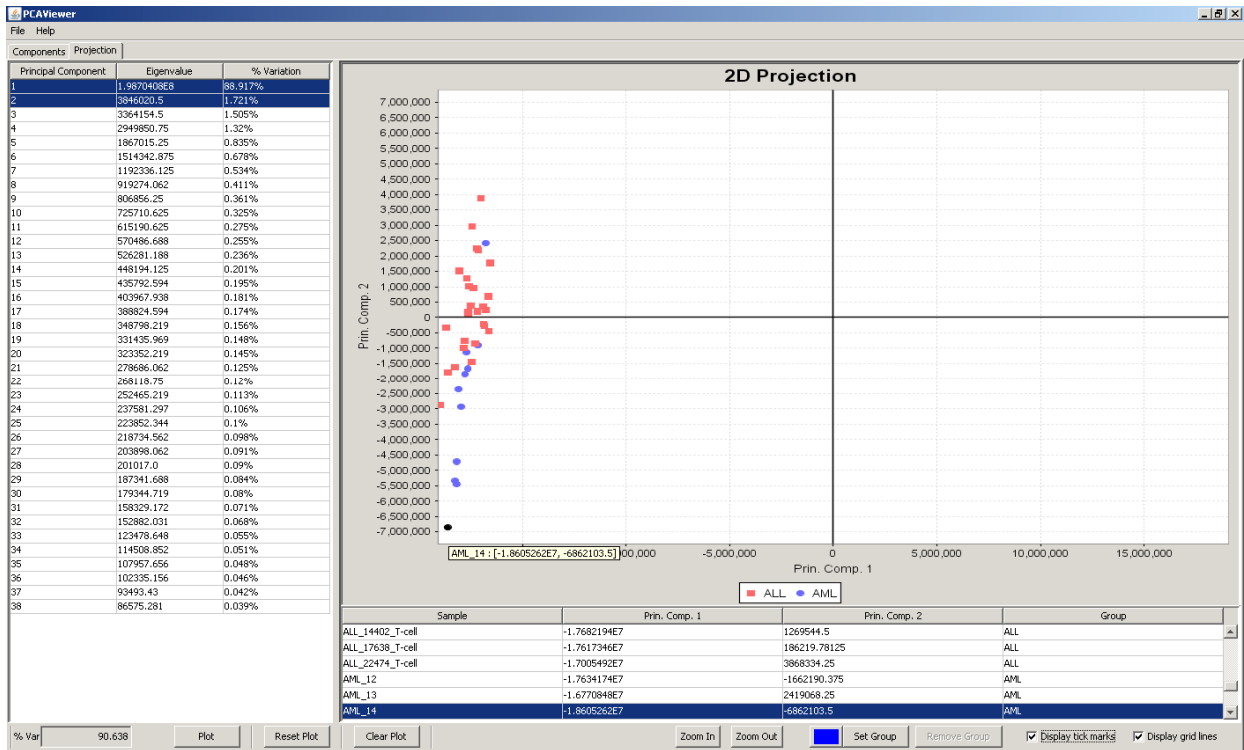


Figure 2. Projection tab projecting genes onto two principal components

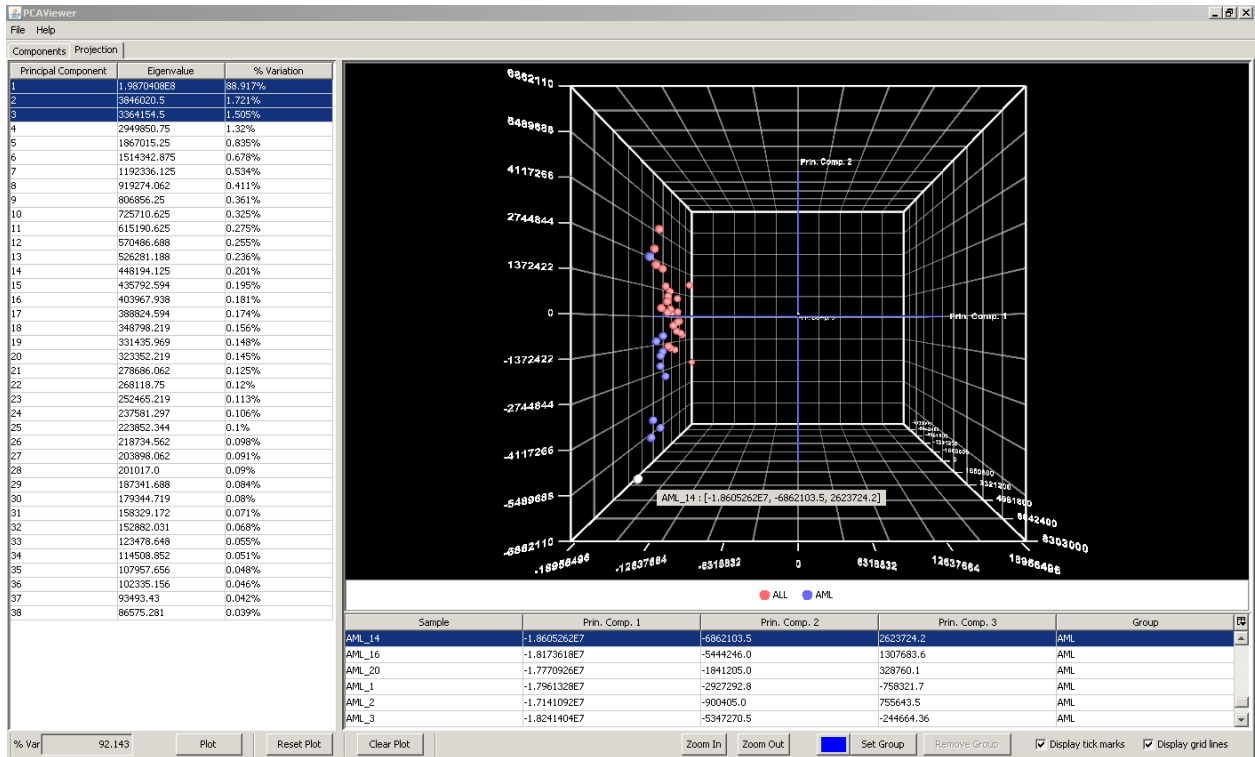


Figure 3. Projection tab projecting genes onto three principal components

# GenePattern

## References

- Raychaudhuri S, Stuart JM, and Altman RB. (2000) Principal components analysis to summarize microarray experiments: application to sporulation time series. Pac. Symp. Biocomput. 455-466.

## Parameters

Name	Description
dataset file	The dataset analyzed by the PCA module. - .res, .gct
s matrix file	The s matrix created by the PCA module. - .odf
t matrix file	The t matrix created by the PCA module. - .odf
u matrix file	The u matrix created by the PCA module. - .odf
cls.or.sample.info.class	A cls or sample info file – .cls, .txt
sample.info.class	The name of the class column to use in the sample info file (required if sample info file provided)
featureset file	The feature set file - .gmt, .gmx, .grp

## Platform dependencies

<b>Module type:</b>	Visualizer
<b>CPU type:</b>	any
<b>OS:</b>	any
<b>Java JVM level:</b>	1.5
<b>Language:</b>	Java