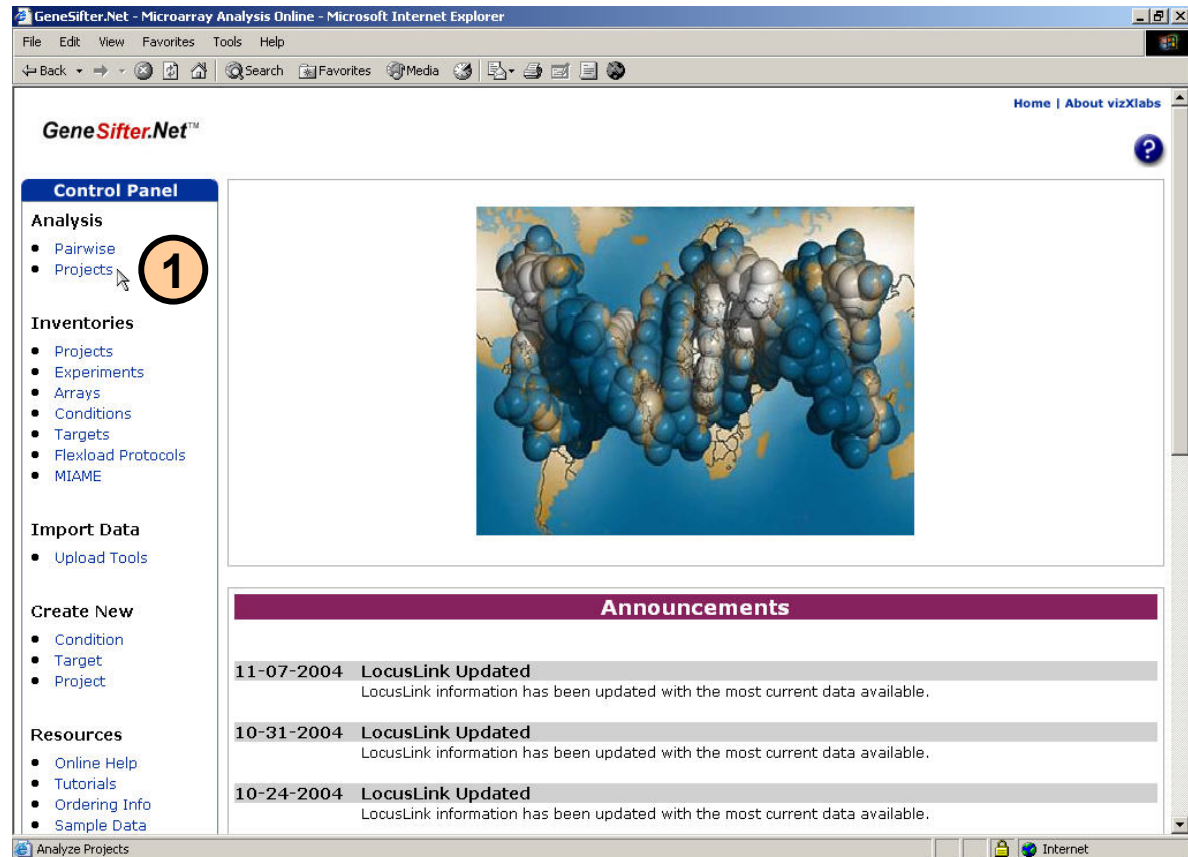


Astrocyte Tutorial

The following tutorial walks through the identification of biological themes associated with expression patterns using unsupervised clustering in a microarray dataset examining Alzheimer's disease.

Visit the GeneSifter Data Center to obtain login information to access the dataset. You should receive a confirmation email shortly. Login to GeneSifter using the provided username and password.

1. From the **Control Panel** select "Projects" within the **Analysis** section.



The screenshot displays the GeneSifter.Net website interface. The browser window title is "GeneSifter.Net - Microarray Analysis Online - Microsoft Internet Explorer". The page layout includes a "Control Panel" on the left with several sections: "Analysis" (containing "Pairwise" and "Projects", with "Projects" highlighted and a circled "1" next to it), "Inventories" (containing "Projects", "Experiments", "Arrays", "Conditions", "Targets", "Flexload Protocols", and "MIAME"), "Import Data" (containing "Upload Tools"), "Create New" (containing "Condition", "Target", and "Project"), and "Resources" (containing "Online Help", "Tutorials", "Ordering Info", and "Sample Data"). The main content area features a 3D visualization of a protein structure overlaid on a world map. Below this is an "Announcements" section with three entries:

- 11-07-2004 **LocusLink Updated**
LocusLink information has been updated with the most current data available.
- 10-31-2004 **LocusLink Updated**
LocusLink information has been updated with the most current data available.
- 10-24-2004 **LocusLink Updated**
LocusLink information has been updated with the most current data available.

Astrocyte Tutorial

(continued)

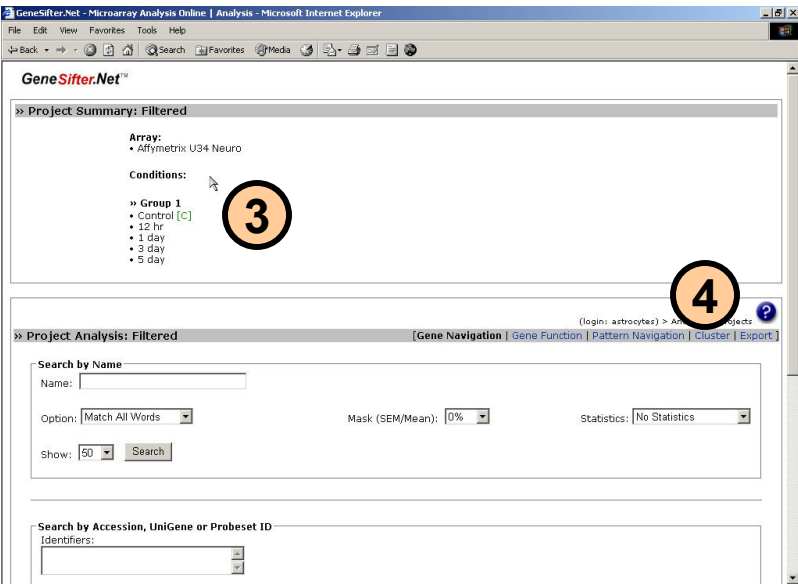
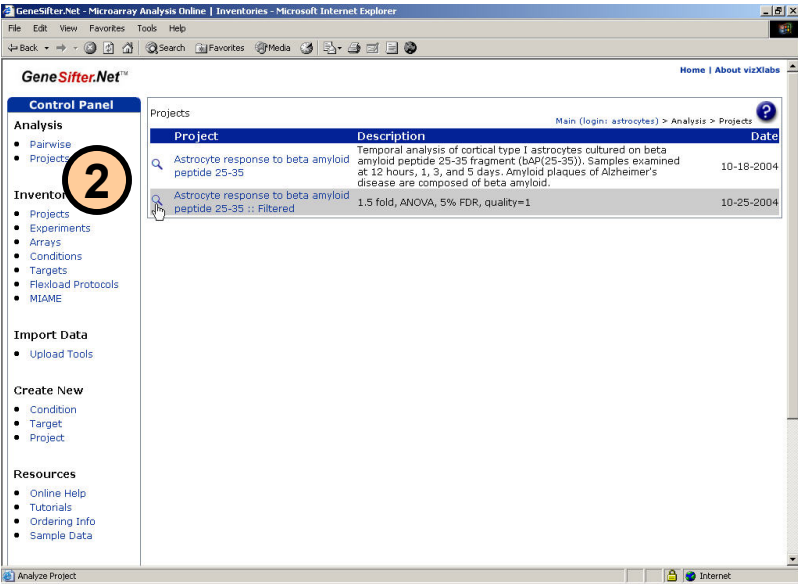
2. Click on the magnifying glass icon next to the “Astrocyte response to beta amyloid peptide 25-35 :: Filtered” project. This is a sub-project created from filtering the data from the “Astrocyte response to beta amyloid peptide 25-35” project. This initial analysis filtered the ~1200 genes represented on the Affymetrix U34 Neuro GeneChip® based on the following parameters:

- Running an ANOVA on the entire set
- Applying a 5% false discovery rate (FDR)
- Minimum 1.5-fold change
- Quality = 1

The result of this filtering was a list of 173 genes that was then saved as a sub-project and will be analyzed in this tutorial.


3. At the top of the page is a Project Summary of the data. For this tutorial, there is one Group which contains 5 time points.

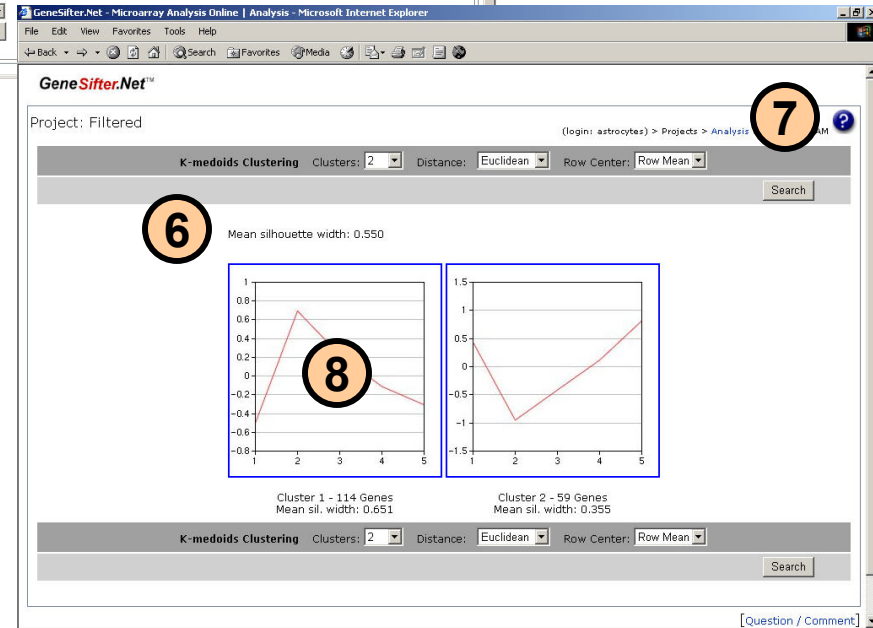
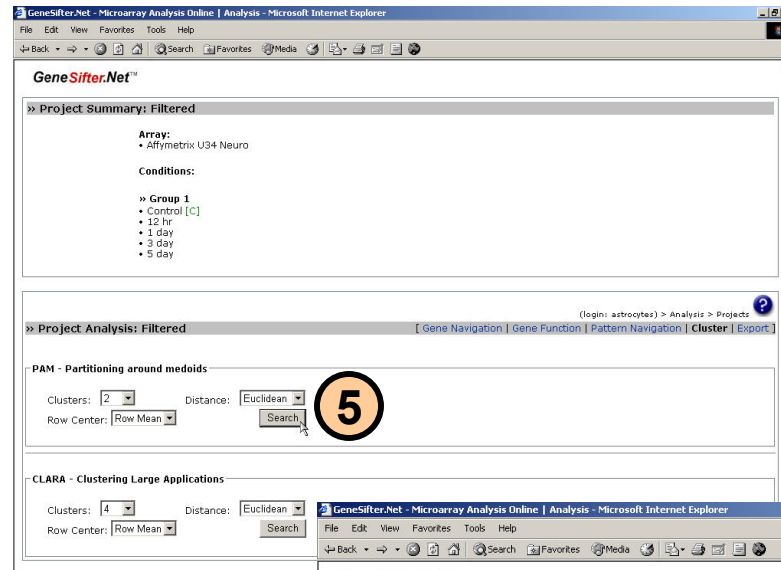
4. Select **Cluster** from the list of analysis options.



Astrocyte Tutorial

(continued)

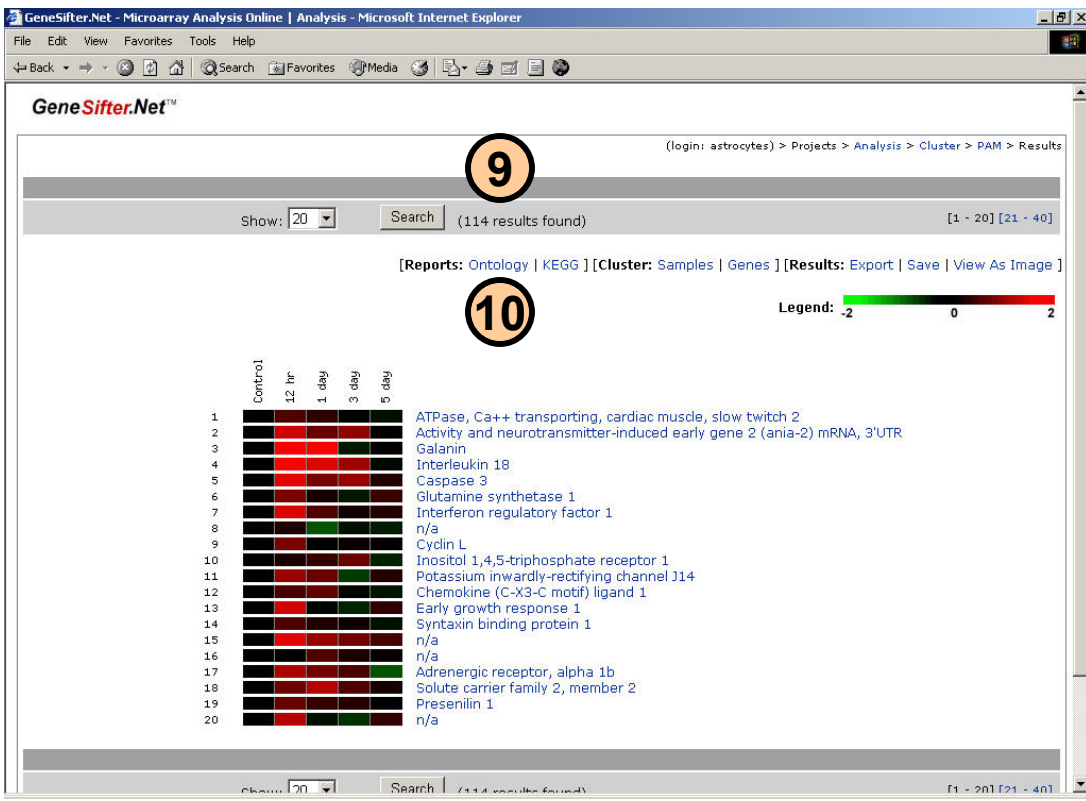
- Unsupervised clustering can be performed on the filtered data. Under the PAM – Partitioning around medoids heading, set the number of clusters to 2, and click **Search**.
 - Each line graph depicts the gene expression pattern for that cluster. Overall, the clustering has a mean silhouette width of 0.550. Cluster 1 contains genes that are upregulated at 12 hours. Cluster 2 displays the opposite pattern, containing genes downregulated at 12 hours, and slowly increasing in expression over the next 4 days.
- To view page-specific help documents for this, or any page, select the question mark icon  located at the upper right page corner. For this page, in particular, there is a document on the interpretation of silhouette width.
- To view the list of genes associated with a cluster, click on the line graph for Cluster 1.



Astrocyte Tutorial

(continued)

- 9. The resulting heat map shows the expression pattern of the 114 genes that were in cluster 1. Recall that this is the heat map for the cluster of upregulated genes, peaking at the 12 hour time point. The heat map clearly displays this pattern.
- 10. To get an overview of the gene families that have been differentially regulated at this time point, click on the **Ontology** link to bring up the Ontology Report.



Astrocyte Tutorial

(continued)

11. The Ontology Report lists the GO terms of genes in the previous gene list. Currently, they are displayed hierarchically. Click on **Z-score Report** to show only the significant terms.

12. Click on **Molecular Function**.

• Since we began with a list of genes upregulated at 12 hours, the z-score report now lists the pathways that were affected at this time point. A positive z-score, like those involving ATP binding and kinase activity, indicate that those terms are significantly overrepresented. Negative z-scores indicate that a term is showing up less often than expected by chance.

