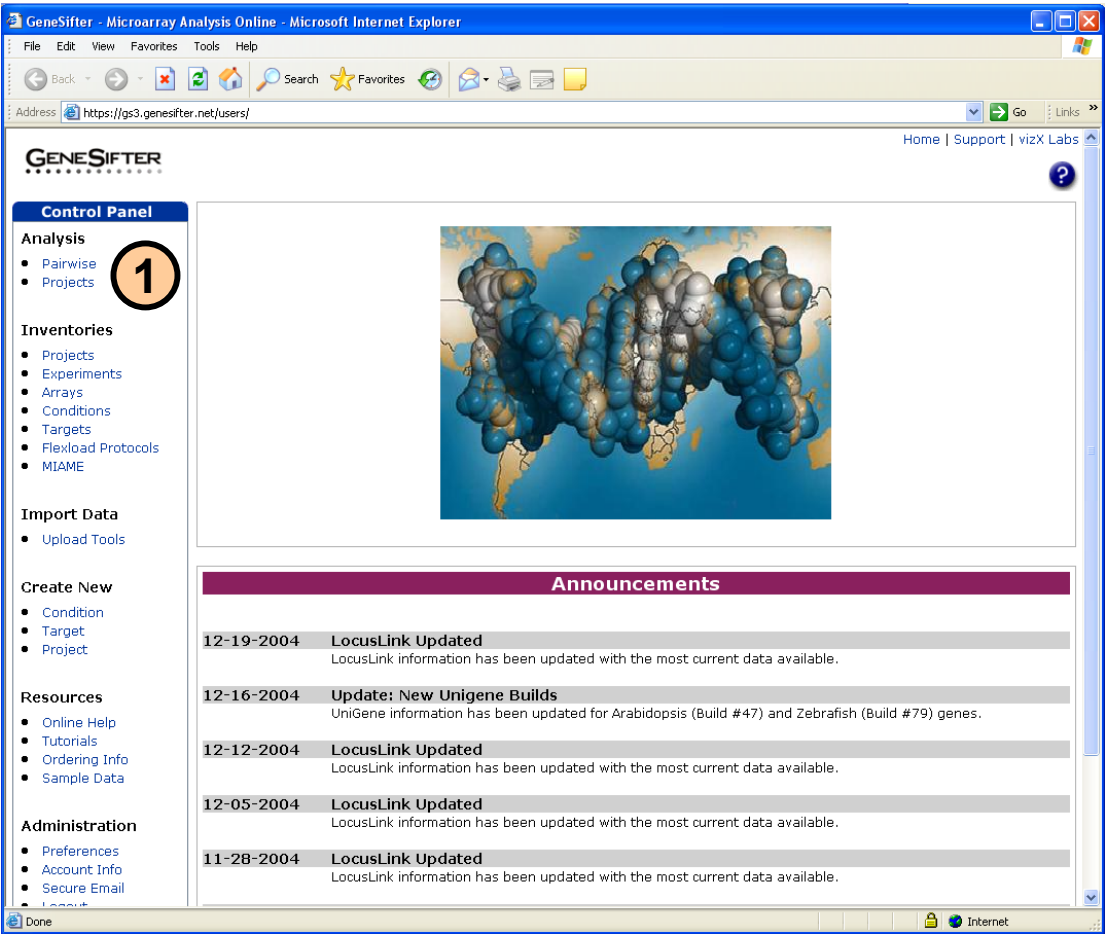


FVB Tutorial

The following tutorial walks through the identification of biological themes associated with expression patterns using unsupervised clustering in a microarray dataset examining mouse heart development.

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

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



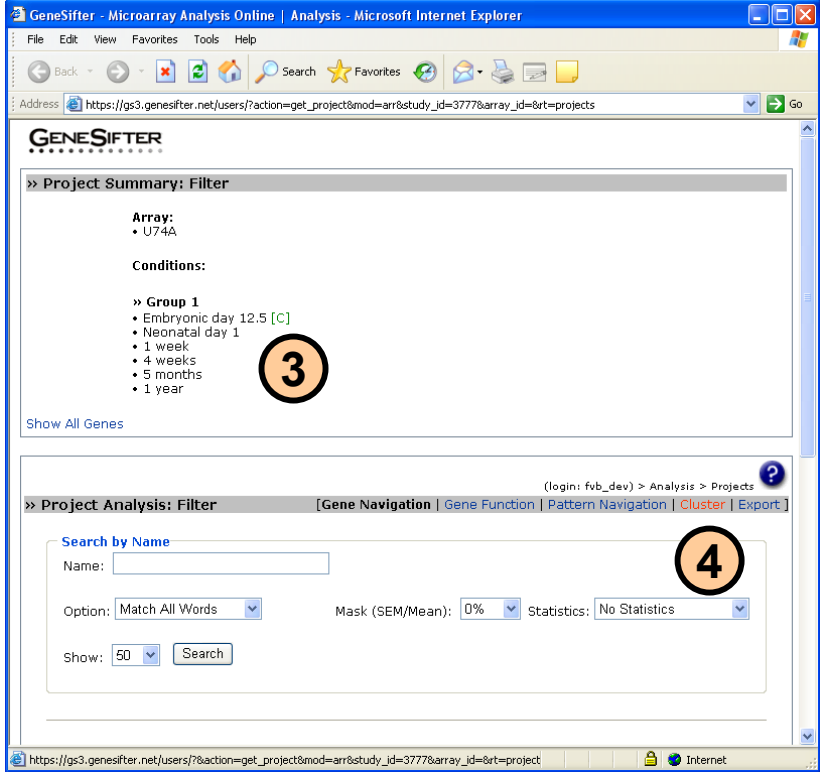
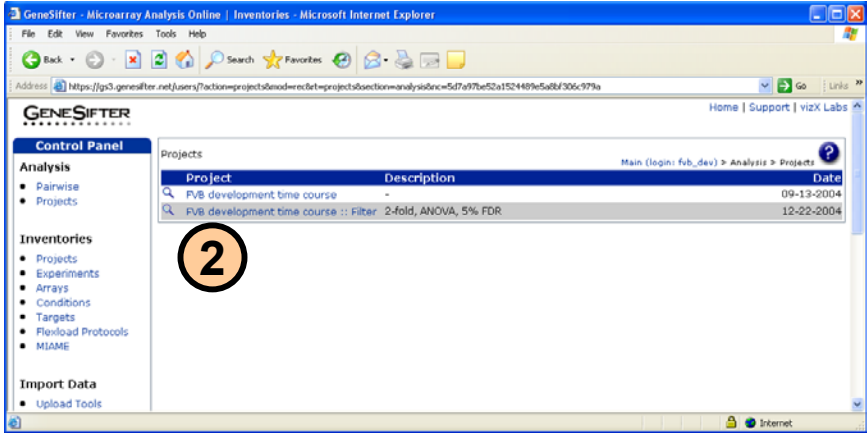
FVB Tutorial

(continued)

- 2. Click on the magnifying glass icon next to the **FVB Development time course :: Filter** project. This is a sub-project created from filtering the data within the **FVB Development** project. The initial analysis filtered the ~10,000 genes represented on the Affymetrix U74A GeneChip® based on the following parameters:
 - applying ANOVA to the entire set
 - applying a 5% false discovery rate cutoff
 - filtering for a minimum 2-fold change cutoff.

The result of this analysis has provided a list of 1516 genes that was saved as the sub-project and will be analyzed in this tutorial.

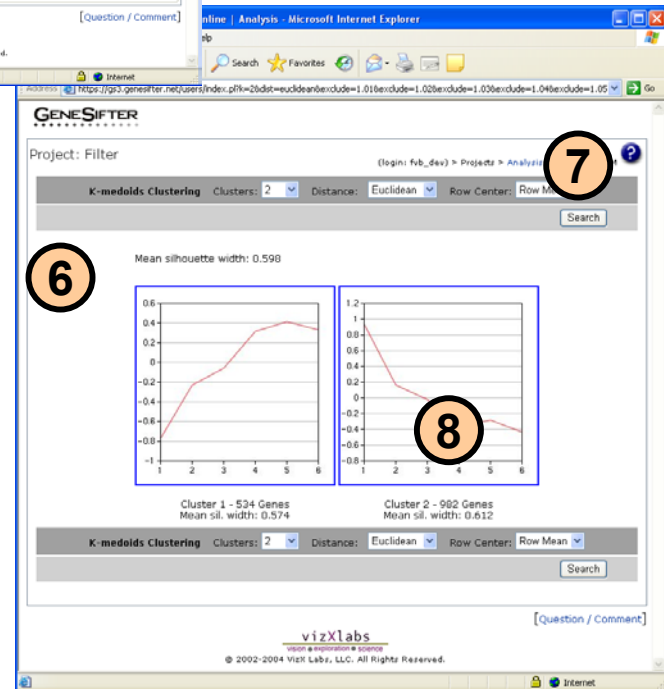
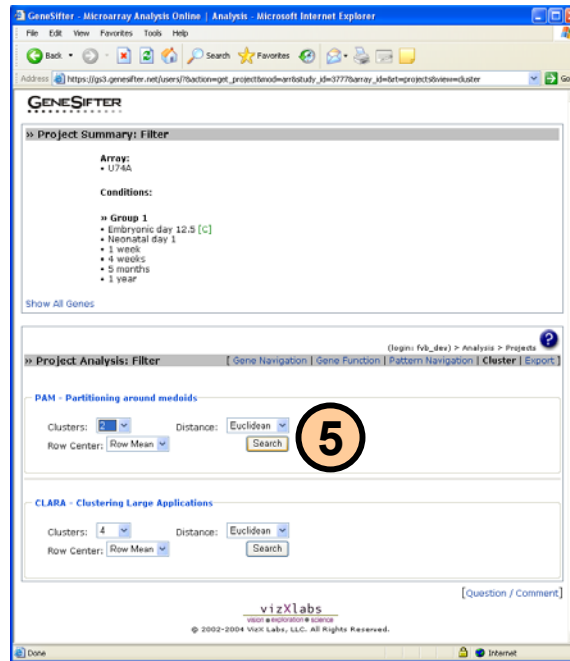
- 3. At the top of this page is a Project Summary of the project data, describing the Array and the Conditions. For this tutorial, there is one Group which contains 6 time points.
- 4. Select **Cluster** from the list of analysis options.



FVB Tutorial

(continued)

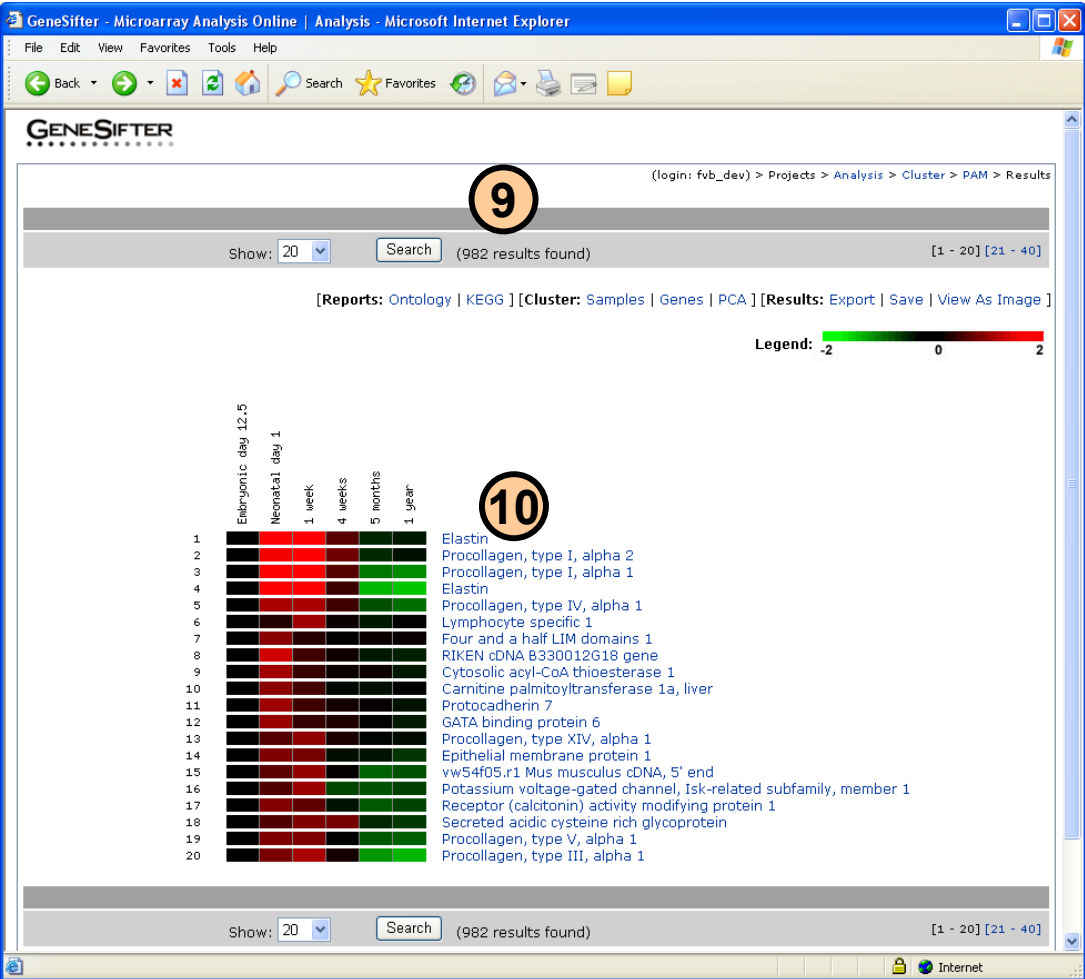
- Unsupervised clustering can be performed on the filtered data by selecting PAM, setting the “Clusters” to 2 from the pull-down menu, and using the defaults for the other settings. Click the “Search” button. A graph showing the center for each cluster will be displayed.
- The resulting clusters have a mean silhouette width of 0.598.
- To view page-specific help documents for this or any page, select the question mark icon located at the upper right page corner. On this page, help documents are available for PAM and silhouettes (including the interpretation of the mean silhouette width).
- To view the list of genes associated with a cluster, click on the graph. For this tutorial, select the graph for cluster 2.



FVB Tutorial

(continued)

- 9. The resulting page shows the expression pattern of the 982 genes. The list of genes and their associated heat map over the 6 time points are displayed.
- 10. To view information for a particular gene, select the gene's title. Select one of the genes.



FVB Tutorial

(continued)

- Gene summaries are available for all genes in the list. The upper half of the summary includes an overview of the data for a particular gene. This includes a summary of the averaged data for each condition, as well as the data for each of the replicates included in each condition.
- The lower half of the summary is the One-Click Gene Summary™ (OCGS). For the gene selected from the gene list, a synopsis displays the most current information from several databases, including UniGene and LocusLink. Links to additional databases and information are identified by blue text.
- Summary information also includes the Gene Ontology terms associated with this particular gene product.



Project: Filter (login: fub_dev) > Project > Analysis > Pattern Navigation > Results > Gene Summary

» Gene Summary: Elastin

11

Condition	Int.	SEM	SEM/Int	Quality	Grp Min	Grp Max	Grp Ave	Grp Med
Embryonic day 12.5	7.5818	±0.0725	1%	0.0000	5.7161	15.2961	10.2182	10.1407
Neonatal day 1	9.8545	±0.1855	1.9%	0.0000	5.7635	15.2997	9.9564	9.8439
1 week	9.7052	±0.1945	2%	0.0000	5.7703	14.9622	9.8537	9.6816
4 weeks	8.2734	±0.0851	1%	0.0000	5.5904	15.1103	9.7138	9.5308
5 months	7.3030	±0.1004	1.4%	0.0000	5.5630	15.2154	9.0149	9.6425
1 year	7.3988	±0.1345	1.8%	0.0000	5.5809	15.0754	9.6647	9.4947

Condition	Target	Int.	Quality
Embryonic day 12.5	FvB_E12	7.5577	0.0000
	FvB_E12	7.4701	0.0000
	FvB_E12	7.7176	0.0000
Neonatal day 1	FvB_NN_1-2	9.7971	0.0000
	FvB_NN_7-8	10.2007	0.0000
	FvB_NN_9-10	9.5658	0.0000
1 week	FvB_1w_801	9.4097	0.0000
	FvB_1w_804	10.0721	0.0000
	FvB_1w_805	9.6338	0.0000
4 weeks	FvB_4w_11293	8.2310	0.0000
	FvB_4w_11294	8.4374	0.0000
	FvB_4w_11295	8.1518	0.0000
5 months	FvB_5m_731M	7.2067	0.0000
	FvB_5m_732M	7.5039	0.0000
	FvB_5m_733M	7.1986	0.0000
1 year	FvB_1y_511M	7.2379	0.0000
	FvB_1y_5M	7.2924	0.0000
	FvB_1y_6M	7.6660	0.0000

Search For Genes With A Similar Pattern

12

Project: Filter

» One-Click Gene Summary™

Probe Set ID: 92207_at
 Accession No.: U08210
 Cluster ID: Mm.275320
 UG Title: Elastin
 Gene ID: Eln
 Homologene: -
 Chromosome: 5
 Cytoband: -
 Seq Count: 229
 LocusLink: 13717
 Gene Name: elastin
 OMIM: -
 KEGG: 13717
 RefSeq mRNA: NM_007925 (FASTA)
 RefSeq Prot: NP_031951 (FASTA)
 Summary: -

13

Gene Ontologies:

Biological Process

- myogenesis
- stress fiber formation
- regulation of actin filament polymerization

Molecular Function

- extracellular matrix structural constituent
- structural molecule activity

Cellular Component

- extracellular matrix (sensu Metazoa)

[Perform Sequence Analysis]

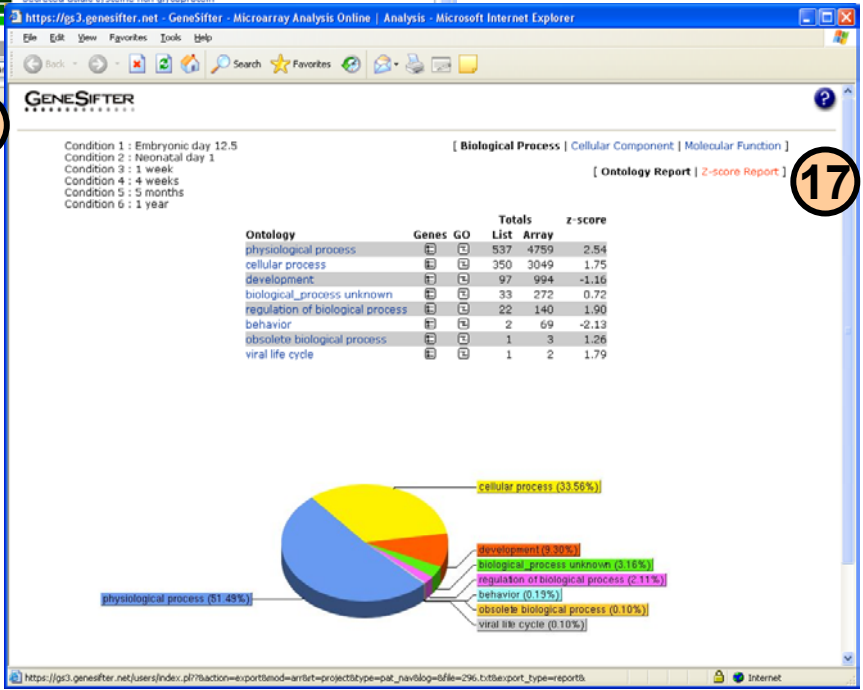
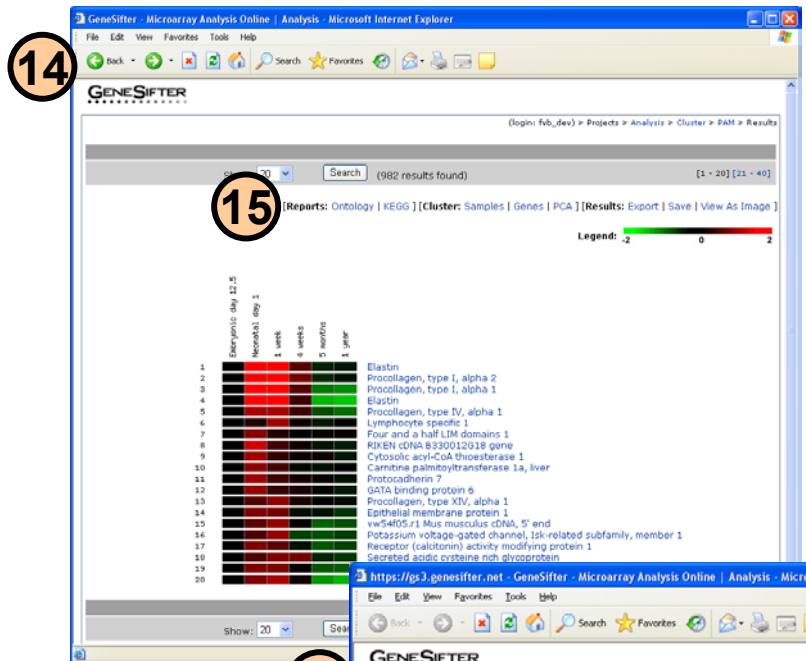
Eln

Search for Homologs:

FVB Tutorial


(continued)

14. From the OCGS page, click on the browser's "Back" button to return to return to the gene list for cluster 2.
15. Select from options at the top of the gene list to perform additional analysis. Click on the **Reports: Ontology** link to view a summary of all the gene families in cluster 2.
16. A new browser window will open containing the Ontology report.
17. Select the **Z-score report** in the upper right corner to list all significant ontology terms.

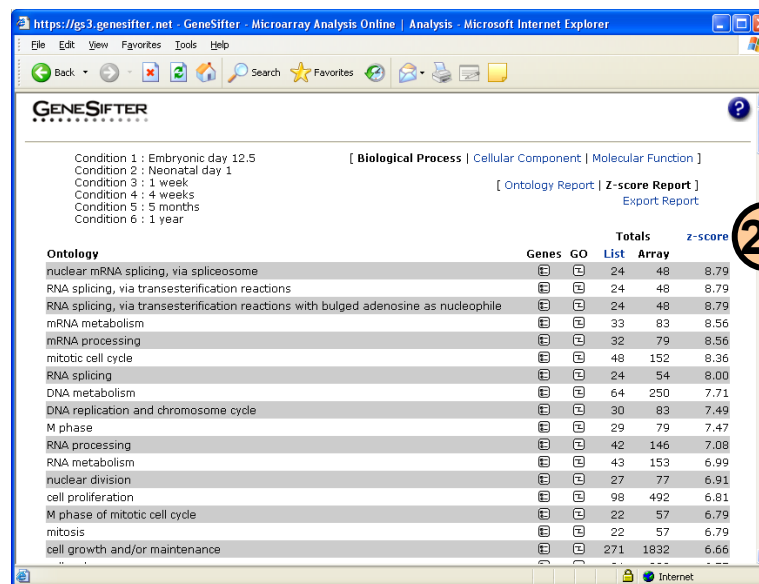


FVB Tutorial

(continued)

20. The z-score report lists all ontologies from the gene list with a z-score greater than 2 or less than -2.
21. Select the z-score link to sort the list by z-score.
21. Selecting the gene list icon  displays a list of the genes (from the gene list), which have that ontology term. Selecting a gene will provide the One Click Gene Summary for that gene.
22. Gene Ontologies are subdivided based on functional distinctions. In the previous steps, ontologies associated with **Biological Process** have been examined. To examine genes associated with other functional categories, select either **Cellular Component** or **Molecular Function** from the top of the page.

20



https://gs3.genesifter.net - GeneSifter - Microarray Analysis Online | Analysis - Microsoft Internet Explorer

GENESIFTER

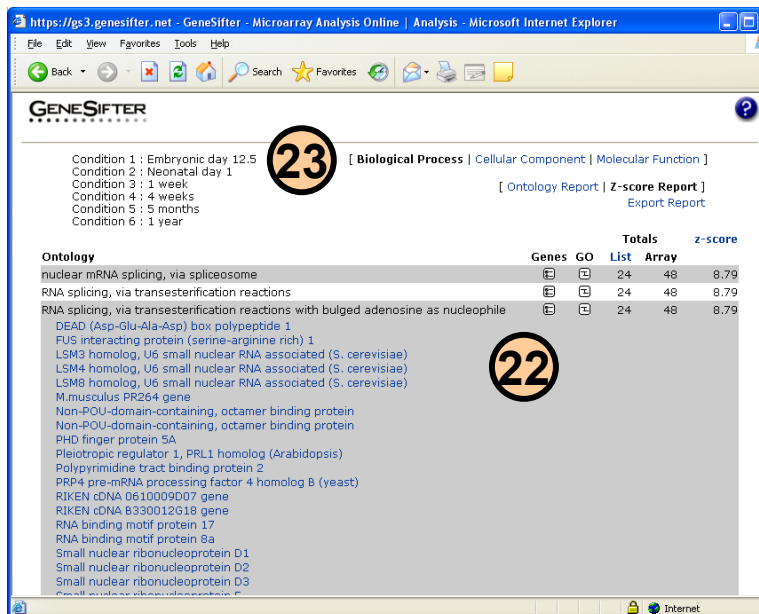
Condition 1 : Embryonic day 12.5
Condition 2 : Neonatal day 1
Condition 3 : 1 week
Condition 4 : 4 weeks
Condition 5 : 5 months
Condition 6 : 1 year

[Biological Process | Cellular Component | Molecular Function]

[[Ontology Report](#) | [Z-score Report](#)]
[Export Report](#)

Ontology	Genes	GO	Totals		z-score
			List	Array	
nuclear mRNA splicing, via spliceosome	24	48	8.79		
RNA splicing, via transesterification reactions	24	48	8.79		
RNA splicing, via transesterification reactions with bulged adenosine as nucleophile	24	48	8.79		
mRNA metabolism	33	83	8.56		
mRNA processing	32	79	8.56		
mitotic cell cycle	48	152	8.36		
RNA splicing	24	54	8.00		
DNA metabolism	64	250	7.71		
DNA replication and chromosome cycle	30	83	7.49		
M phase	29	79	7.47		
RNA processing	42	146	7.08		
RNA metabolism	43	153	6.99		
nuclear division	27	77	6.91		
cell proliferation	98	492	6.81		
M phase of mitotic cell cycle	22	57	6.79		
mitosis	22	57	6.79		
cell growth and/or maintenance	271	1832	6.66		

21



https://gs3.genesifter.net - GeneSifter - Microarray Analysis Online | Analysis - Microsoft Internet Explorer

GENESIFTER

Condition 1 : Embryonic day 12.5
Condition 2 : Neonatal day 1
Condition 3 : 1 week
Condition 4 : 4 weeks
Condition 5 : 5 months
Condition 6 : 1 year

[Biological Process | Cellular Component | Molecular Function]

[[Ontology Report](#) | [Z-score Report](#)]
[Export Report](#)

Ontology	Genes	GO	Totals		z-score
			List	Array	
nuclear mRNA splicing, via spliceosome	24	48	8.79		
RNA splicing, via transesterification reactions	24	48	8.79		
RNA splicing, via transesterification reactions with bulged adenosine as nucleophile	24	48	8.79		
DEAD (Asp-Glu-Ala-Asp) box polypeptide 1					
FUS interacting protein (serine-arginine rich) 1					
LSM3 homolog, U6 small nuclear RNA associated (S. cerevisiae)					
LSM4 homolog, U6 small nuclear RNA associated (S. cerevisiae)					
LSM8 homolog, U6 small nuclear RNA associated (S. cerevisiae)					
M.musculus PR264 gene					
Non-POU-domain-containing, octamer binding protein					
Non-POU-domain-containing, octamer binding protein					
PBD finger protein 5A					
Pleiotropic regulator 1, PRL1 homolog (Arabidopsis)					
Polypyrimidine tract binding protein 2					
PRP4 pre-mRNA processing factor 4 homolog B (yeast)					
RIKEN cDNA 061009D07 gene					
RIKEN cDNA B330012G18 gene					
RNA binding motif protein 17					
RNA binding motif protein 8a					
Small nuclear ribonucleoprotein D1					
Small nuclear ribonucleoprotein D2					
Small nuclear ribonucleoprotein D3					
Small nuclear ribonucleoprotein D4					
Small nuclear ribonucleoprotein D5					

23

22