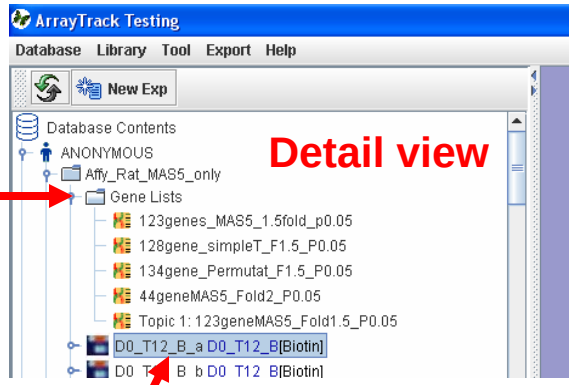
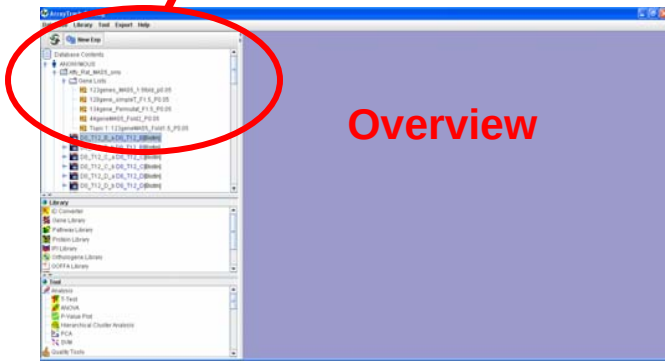
Tutorial 6:
GeneList – An important concept in ArrayTrack

Introduction

The Gene List folder contains gene lists that are usually derived from applying certain filtering criteria (e.g. P-value, fold change, etc) on the microarray data. As a common practice, a list of gene is identified by researchers and then carried forward for biological interpretation and/or further analysis.



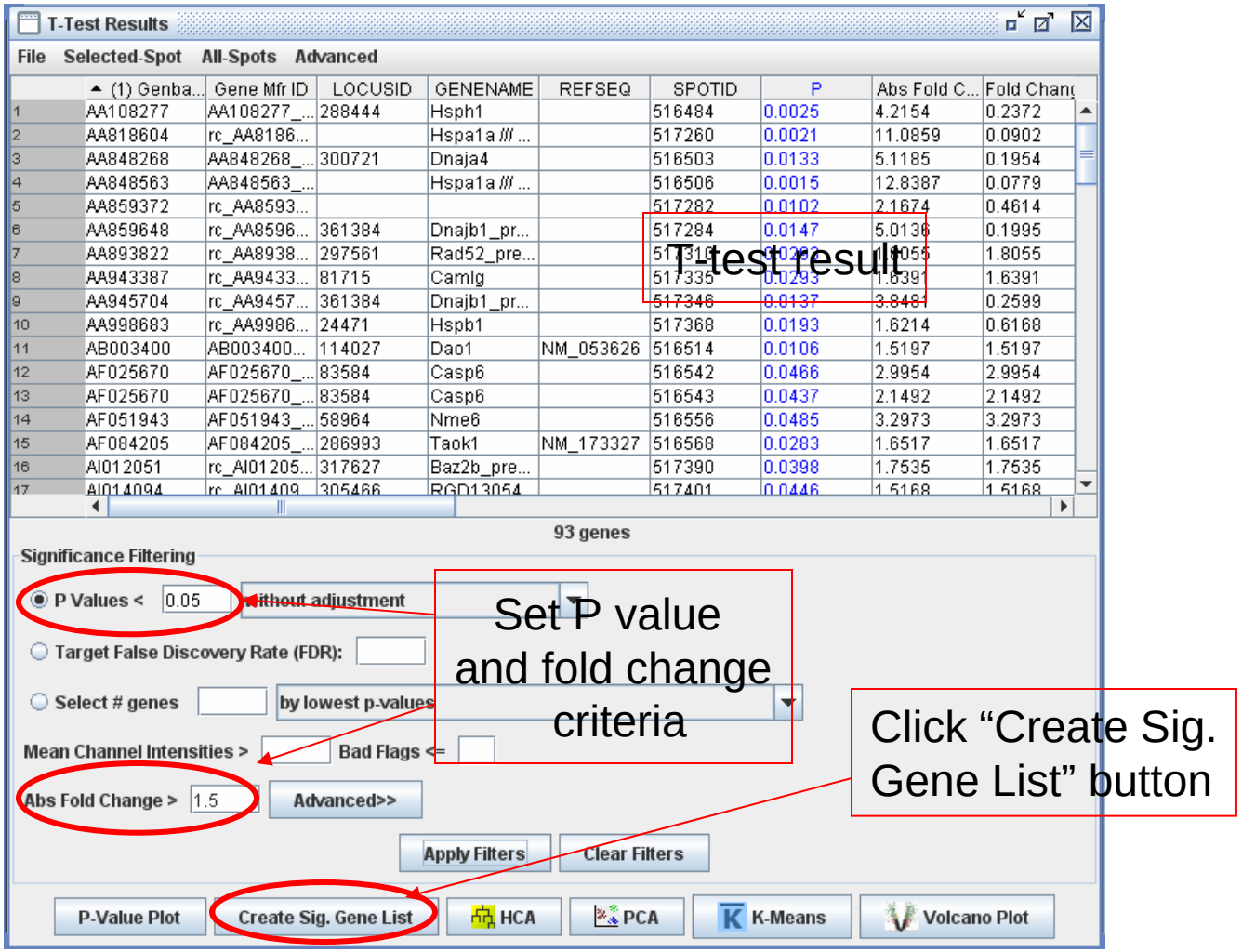
There are several ways to create a gene list in ArrayTrack:
- Create a gene list through statistical analysis (page 3)
- Create a gene list through volcano plot (page 5)
- Create a gene list through data filtering (page 6)
- Create a gene list through manual selection (page 7)
- Create a gene list through GOFFA (page 8)
- Create a gene list through pathway (page 10)
- Alternatively, the user can import a list of genes that is created outside of ArrayTrack. (page 12)



ArrayTrack allows the following operations solely based on gene lists:
- Normalization based on a gene list (page 21)
- Statistical analysis (t-test, PCA, HCA) applied on a gene list (page 22)
- Venn diagram drawn on two or three gene lists (page 19)
- Functional analysis on a gene list (page 26)
- Export a set of microarray with genes that are specified by a gene list (page 28)
- Miscellaneous: delete a gene list (page 17); create a Gene List subfolder (page 15)

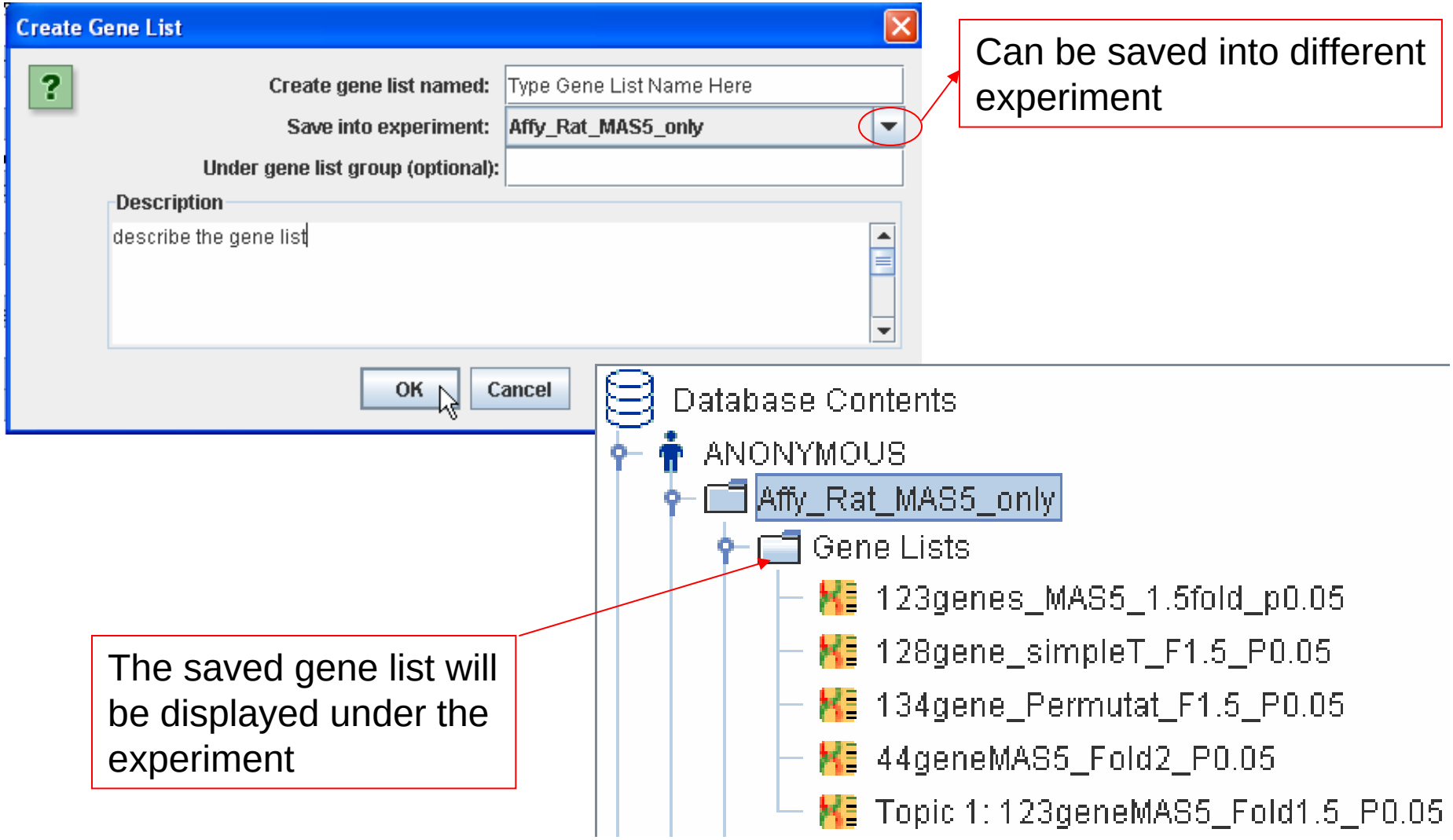
Create a gene list through statistical analysis

From data analysis results (e.g. T-test result. See Tutorial 1 for doing T-test).



	(1) Genba...	Gene Mfr ID	LOCUSID	GENENAME	REFSEQ	SPOTID	P	Abs Fold C...	Fold Chang
1	AA108277	AA108277_...	288444	Hsph1		516484	0.0025	4.2154	0.2372
2	AA818604	rc_AA8186...		Hspa1a /// ...		517260	0.0021	11.0859	0.0902
3	AA848268	AA848268_...	300721	Dnaja4		516503	0.0133	5.1185	0.1954
4	AA848563	AA848563_...		Hspa1a /// ...		516506	0.0015	12.8387	0.0779
5	AA859372	rc_AA8593...				517282	0.0102	2.1674	0.4614
6	AA859648	rc_AA8596...	361384	Dnajb1_pr...		517284	0.0147	5.0136	0.1995
7	AA893822	rc_AA8938...	297561	Rad52_pre...		517310	[illegible]	[illegible]	1.8055
8	AA943387	rc_AA9433...	81715	Camlg		517335	0.0293	1.6391	1.6391
9	AA945704	rc_AA9457...	361384	Dnajb1_pr...		517346	0.0137	3.8481	0.2599
10	AA998683	rc_AA9986...	24471	Hspb1		517368	0.0193	1.6214	0.6168
11	AB003400	AB003400...	114027	Dao1	NM_053626	516514	0.0106	1.5197	1.5197
12	AF025670	AF025670_...	83584	Casp6		516542	0.0466	2.9954	2.9954
13	AF025670	AF025670_...	83584	Casp6		516543	0.0437	2.1492	2.1492
14	AF051943	AF051943_...	58964	Nme6		516556	0.0485	3.2973	3.2973
15	AF084205	AF084205_...	286993	Taok1	NM_173327	516568	0.0283	1.6517	1.6517
16	AI012051	rc_AI01205...	317627	Baz2b_pre...		517390	0.0398	1.7535	1.7535
17	AI014094	rc_AI01409...	305466	RGD13054		517401	0.0446	1.5168	1.5168

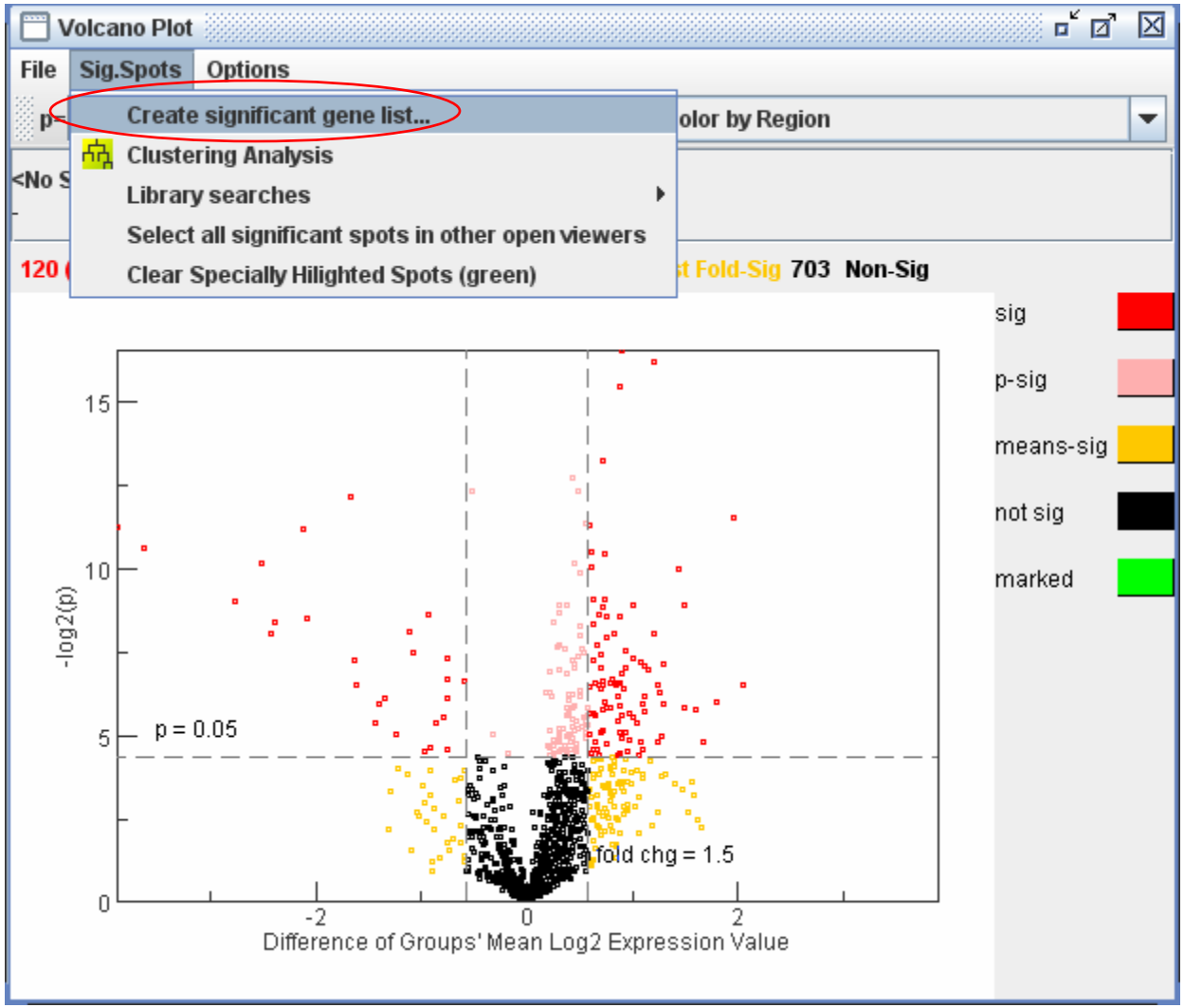
Create a gene list through statistical analysis - continued



Can be saved into different experiment

The saved gene list will be displayed under the experiment

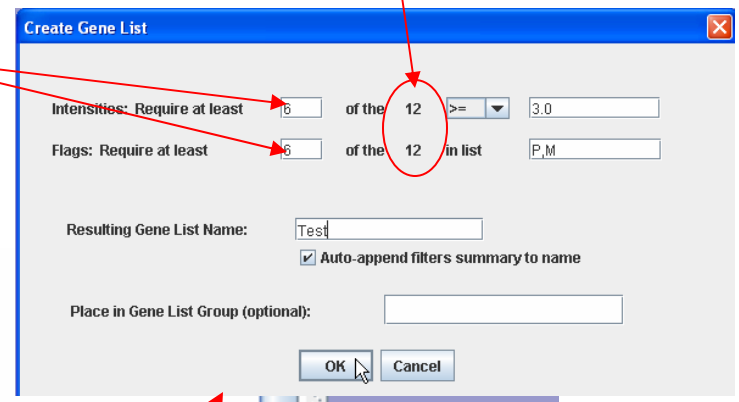
Create gene list through volcano plot



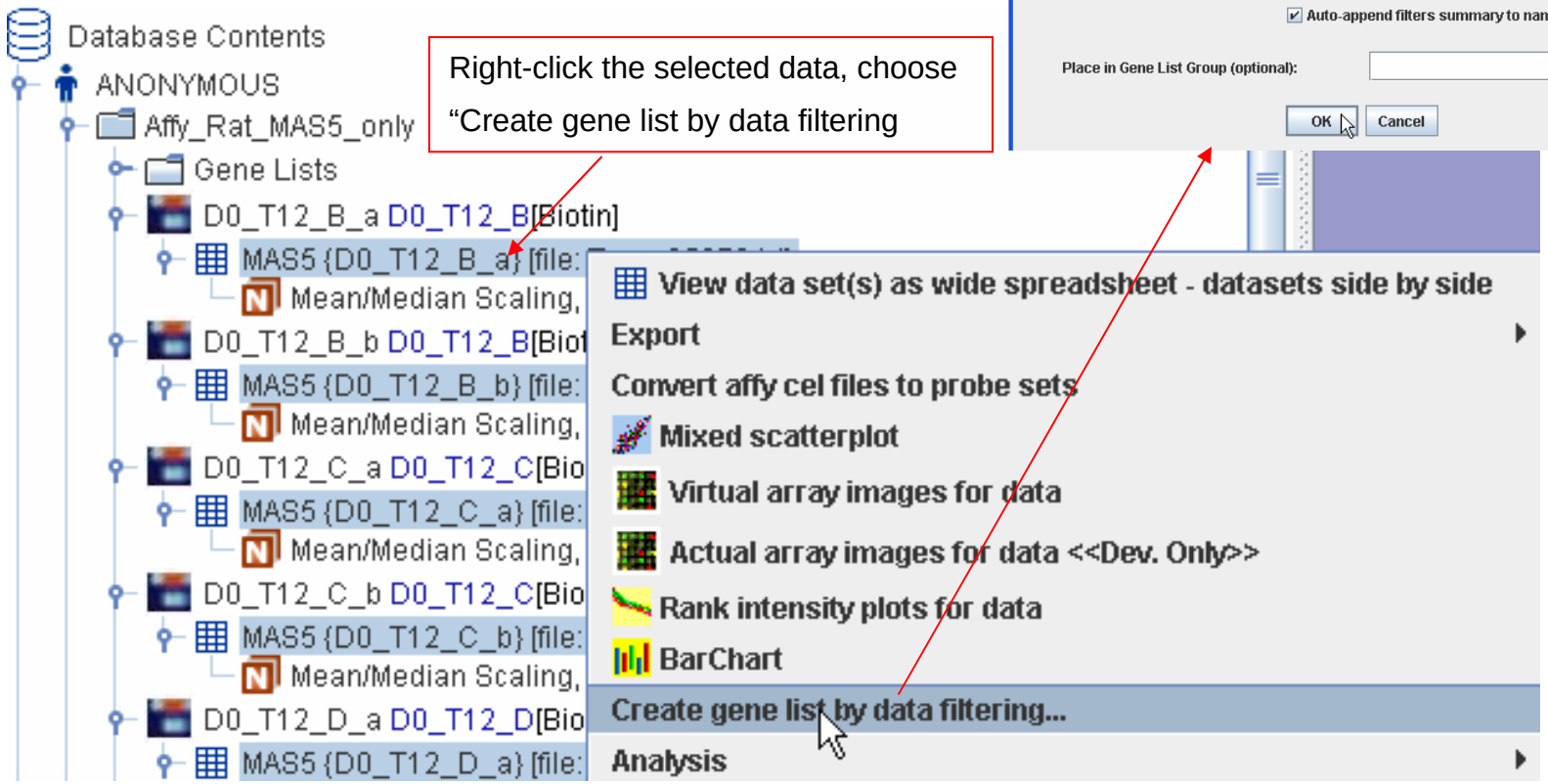
Create a gene list through data filtering

The total number of selected hybridizations

User can set the criteria (like flag, intensity value) to create a gene list, e.g find out all the genes which intensity is greater than 3.0 in at least 6 out of 12 hybridizations.



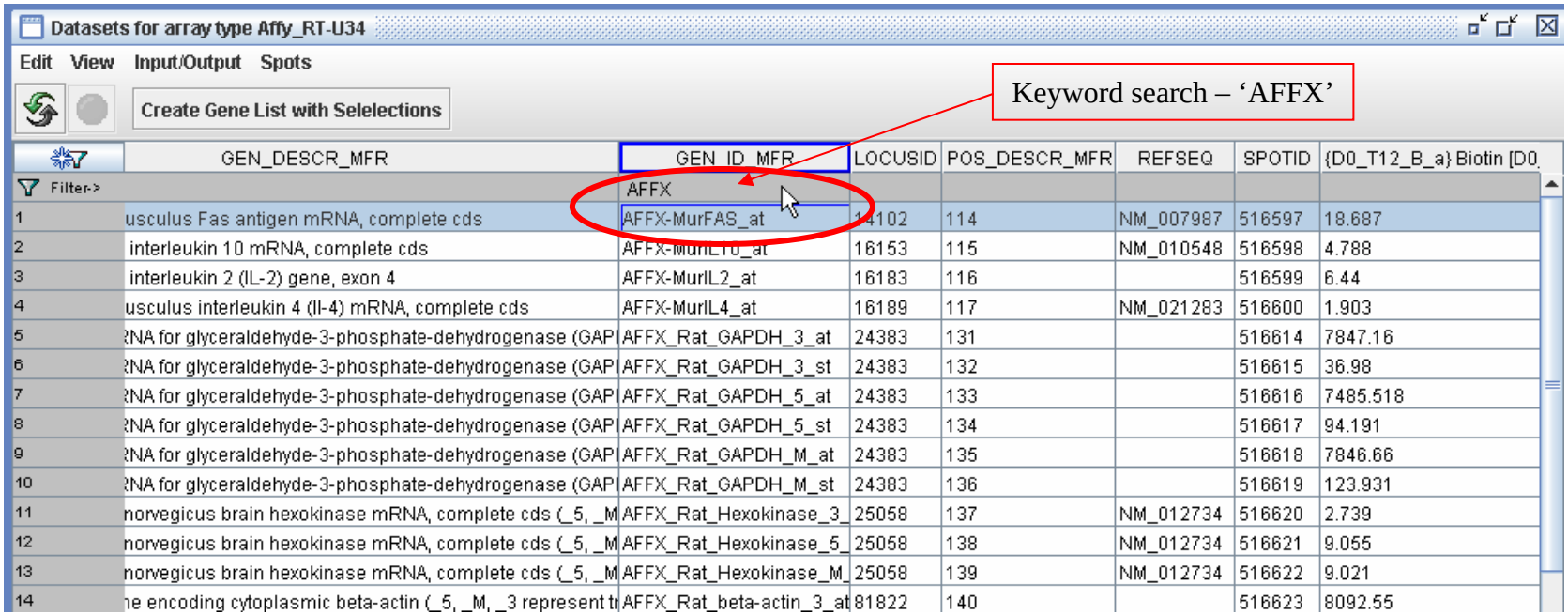
Right-click the selected data, choose "Create gene list by data filtering



Create a gene list through querying chip information

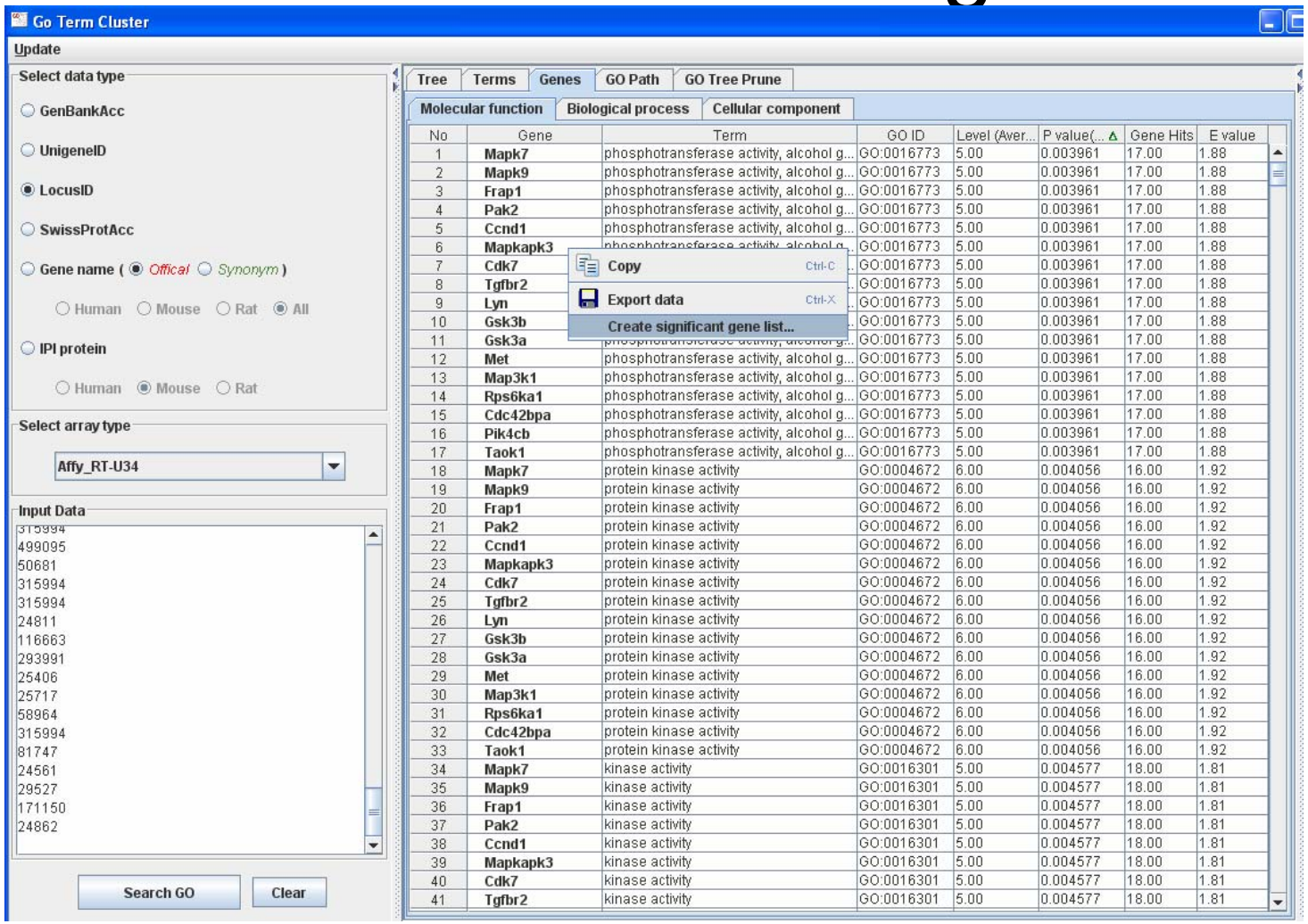
The user can filter out some genes (e.g. spike-in genes) to get a gene list for further normalization.

This image shows a raw dataset using filtering criteria to only display the genes the user inquired. User can open any raw data to get the following spreadsheet.



	GEN_DESCR_MFR	GEN_ID_MFR	LOCUSID	POS_DESCR_MFR	REFSEQ	SPOTID	{D0_T12_B_a} Biotin [D0
Filter->		AFFX					
1	usculus Fas antigen mRNA, complete cds	AFFX-MurFAS_at	14102	114	NM_007987	516597	18.687
2	interleukin 10 mRNA, complete cds	AFFX-MurIL10_at	16153	115	NM_010548	516598	4.788
3	interleukin 2 (IL-2) gene, exon 4	AFFX-MurIL2_at	16183	116		516599	6.44
4	usculus interleukin 4 (Il-4) mRNA, complete cds	AFFX-MurIL4_at	16189	117	NM_021283	516600	1.903
5	RNA for glyceraldehyde-3-phosphate-dehydrogenase (GAP	AFFX_Rat_GAPDH_3_at	24383	131		516614	7847.16
6	RNA for glyceraldehyde-3-phosphate-dehydrogenase (GAP	AFFX_Rat_GAPDH_3_st	24383	132		516615	36.98
7	RNA for glyceraldehyde-3-phosphate-dehydrogenase (GAP	AFFX_Rat_GAPDH_5_at	24383	133		516616	7485.518
8	RNA for glyceraldehyde-3-phosphate-dehydrogenase (GAP	AFFX_Rat_GAPDH_5_st	24383	134		516617	94.191
9	RNA for glyceraldehyde-3-phosphate-dehydrogenase (GAP	AFFX_Rat_GAPDH_M_at	24383	135		516618	7846.66
10	RNA for glyceraldehyde-3-phosphate-dehydrogenase (GAP	AFFX_Rat_GAPDH_M_st	24383	136		516619	123.931
11	norvegicus brain hexokinase mRNA, complete cds (_5, _M	AFFX_Rat_Hexokinase_3_	25058	137	NM_012734	516620	2.739
12	norvegicus brain hexokinase mRNA, complete cds (_5, _M	AFFX_Rat_Hexokinase_5_	25058	138	NM_012734	516621	9.055
13	norvegicus brain hexokinase mRNA, complete cds (_5, _M	AFFX_Rat_Hexokinase_M_	25058	139	NM_012734	516622	9.021
14	ne encoding cytoplasmic beta-actin (_5, _M, _3 represent t	AFFX_Rat_beta-actin_3_at	81822	140		516623	8092.55

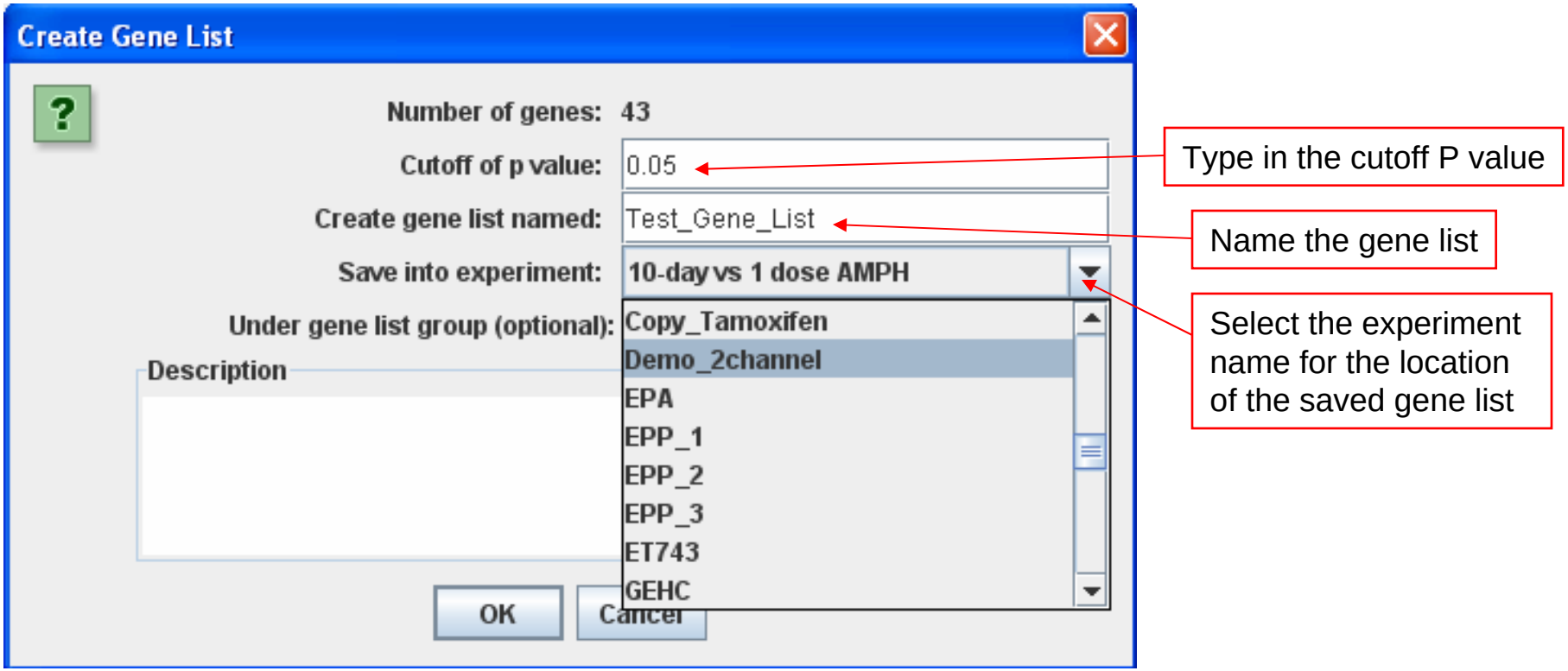
Create Gene List through GOFFA



No	Gene	Term	GO ID	Level (Aver...	P value(...	Gene Hits	E value
1	Mapk7	phosphotransferase activity, alcohol g...	GO:0016773	5.00	0.003961	17.00	1.88
2	Mapk9	phosphotransferase activity, alcohol g...	GO:0016773	5.00	0.003961	17.00	1.88
3	Frap1	phosphotransferase activity, alcohol g...	GO:0016773	5.00	0.003961	17.00	1.88
4	Pak2	phosphotransferase activity, alcohol g...	GO:0016773	5.00	0.003961	17.00	1.88
5	Ccnd1	phosphotransferase activity, alcohol g...	GO:0016773	5.00	0.003961	17.00	1.88
6	Mapkapk3	phosphotransferase activity, alcohol g...	GO:0016773	5.00	0.003961	17.00	1.88
7	Cdk7	[illegible]	GO:0016773	5.00	0.003961	17.00	1.88
8	Tgfbr2	[illegible]	GO:0016773	5.00	0.003961	17.00	1.88
9	Lyn	[illegible]	GO:0016773	5.00	0.003961	17.00	1.88
10	Gsk3b	[illegible]	GO:0016773	5.00	0.003961	17.00	1.88
11	Gsk3a	[illegible]	GO:0016773	5.00	0.003961	17.00	1.88
12	Met	phosphotransferase activity, alcohol g...	GO:0016773	5.00	0.003961	17.00	1.88
13	Map3k1	phosphotransferase activity, alcohol g...	GO:0016773	5.00	0.003961	17.00	1.88
14	Rps6ka1	phosphotransferase activity, alcohol g...	GO:0016773	5.00	0.003961	17.00	1.88
15	Cdc42bpa	phosphotransferase activity, alcohol g...	GO:0016773	5.00	0.003961	17.00	1.88
16	Pik4cb	phosphotransferase activity, alcohol g...	GO:0016773	5.00	0.003961	17.00	1.88
17	Taok1	phosphotransferase activity, alcohol g...	GO:0016773	5.00	0.003961	17.00	1.88
18	Mapk7	protein kinase activity	GO:0004672	6.00	0.004056	16.00	1.92
19	Mapk9	protein kinase activity	GO:0004672	6.00	0.004056	16.00	1.92
20	Frap1	protein kinase activity	GO:0004672	6.00	0.004056	16.00	1.92
21	Pak2	protein kinase activity	GO:0004672	6.00	0.004056	16.00	1.92
22	Ccnd1	protein kinase activity	GO:0004672	6.00	0.004056	16.00	1.92
23	Mapkapk3	protein kinase activity	GO:0004672	6.00	0.004056	16.00	1.92
24	Cdk7	protein kinase activity	GO:0004672	6.00	0.004056	16.00	1.92
25	Tgfbr2	protein kinase activity	GO:0004672	6.00	0.004056	16.00	1.92
26	Lyn	protein kinase activity	GO:0004672	6.00	0.004056	16.00	1.92
27	Gsk3b	protein kinase activity	GO:0004672	6.00	0.004056	16.00	1.92
28	Gsk3a	protein kinase activity	GO:0004672	6.00	0.004056	16.00	1.92
29	Met	protein kinase activity	GO:0004672	6.00	0.004056	16.00	1.92
30	Map3k1	protein kinase activity	GO:0004672	6.00	0.004056	16.00	1.92
31	Rps6ka1	protein kinase activity	GO:0004672	6.00	0.004056	16.00	1.92
32	Cdc42bpa	protein kinase activity	GO:0004672	6.00	0.004056	16.00	1.92
33	Taok1	protein kinase activity	GO:0004672	6.00	0.004056	16.00	1.92
34	Mapk7	kinase activity	GO:0016301	5.00	0.004577	18.00	1.81
35	Mapk9	kinase activity	GO:0016301	5.00	0.004577	18.00	1.81
36	Frap1	kinase activity	GO:0016301	5.00	0.004577	18.00	1.81
37	Pak2	kinase activity	GO:0016301	5.00	0.004577	18.00	1.81
38	Ccnd1	kinase activity	GO:0016301	5.00	0.004577	18.00	1.81
39	Mapkapk3	kinase activity	GO:0016301	5.00	0.004577	18.00	1.81
40	Cdk7	kinase activity	GO:0016301	5.00	0.004577	18.00	1.81
41	Tgfbr2	kinase activity	GO:0016301	5.00	0.004577	18.00	1.81

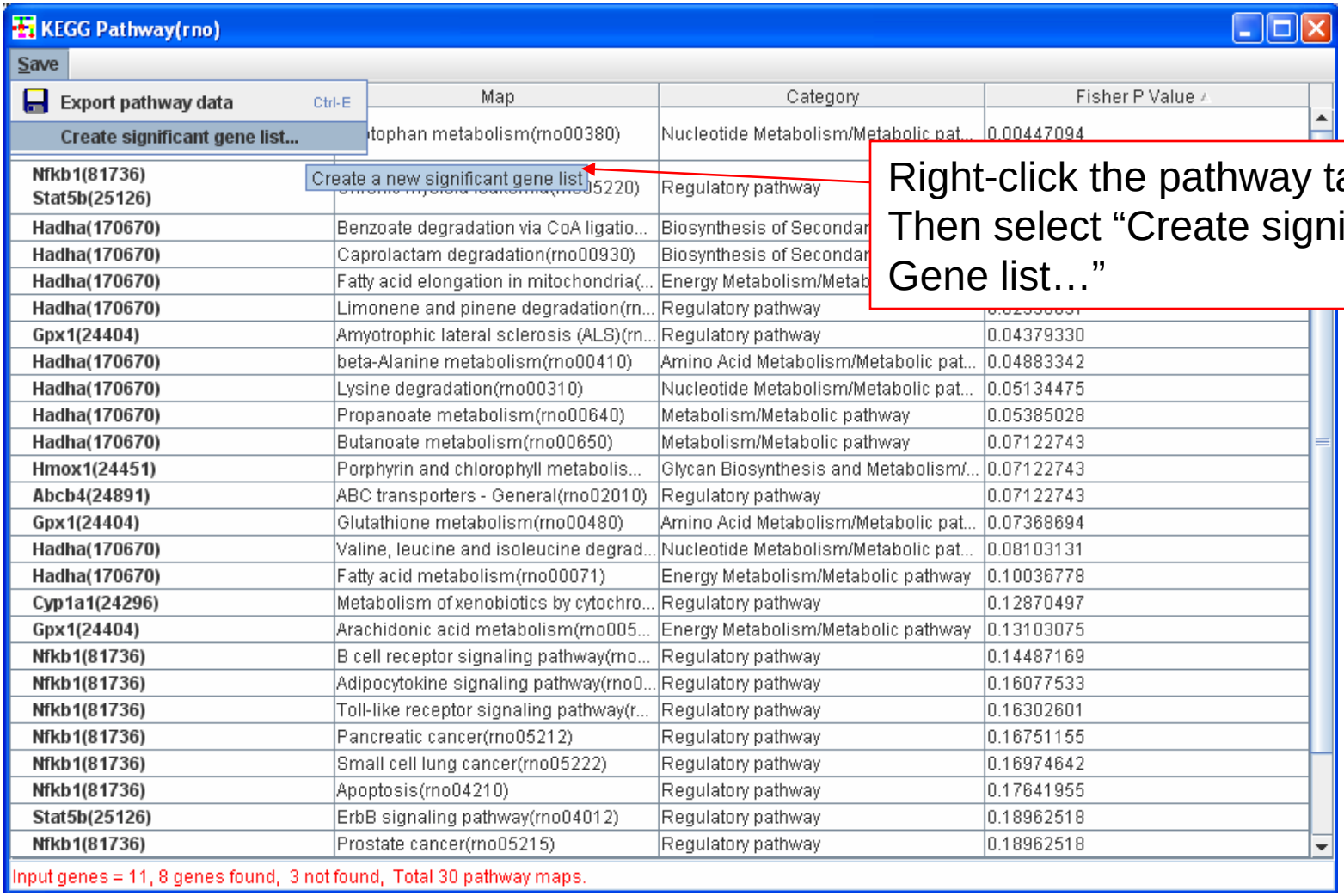
This is the GOFFA table. Click Gene tab, then right-click the gene table, select "Create significant gene list…"

Create Gene List through GOFFA



When the user type in the cutoff p value, the number of genes will change according to the value of P.

Create Significant Gene List through Pathway

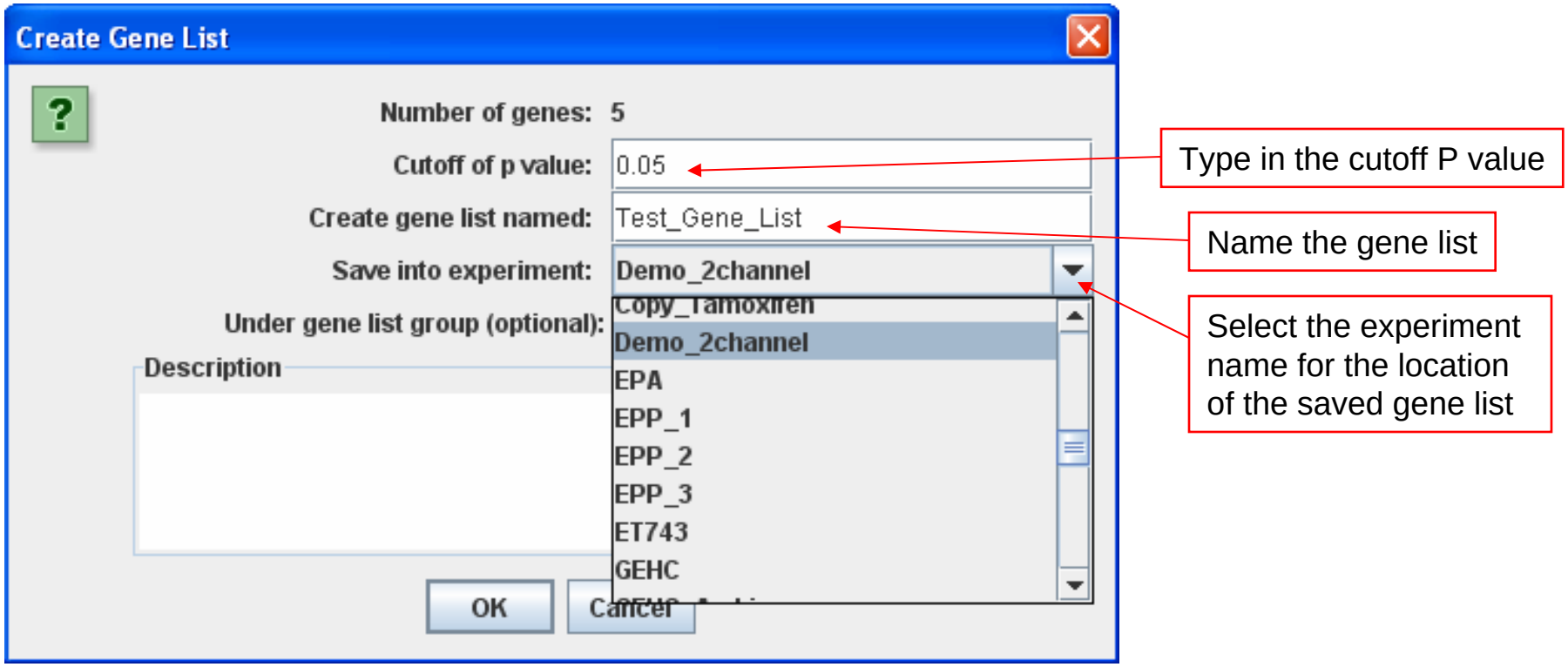


	Map	Category	Fisher P Value
	tophan metabolism(rno00380)	Nucleotide Metabolism/Metabolic pat...	0.00447094
Nfkb1(81736) Stat5b(25126)	5220)	Regulatory pathway	
Hadha(170670)	Benzoate degradation via CoA ligatio...	Biosynthesis of Secondar	
Hadha(170670)	Caprolactam degradation(rno00930)	Biosynthesis of Secondar	
Hadha(170670)	Fatty acid elongation in mitochondria(...	Energy Metabolism/Metab	
Hadha(170670)	Limonene and pinene degradation(rn...	Regulatory pathway	
Gpx1(24404)	Amyotrophic lateral sclerosis (ALS)(rn...	Regulatory pathway	0.04379330
Hadha(170670)	beta-Alanine metabolism(rno00410)	Amino Acid Metabolism/Metabolic pat...	0.04883342
Hadha(170670)	Lysine degradation(rno00310)	Nucleotide Metabolism/Metabolic pat...	0.05134475
Hadha(170670)	Propanoate metabolism(rno00640)	Metabolism/Metabolic pathway	0.05385028
Hadha(170670)	Butanoate metabolism(rno00650)	Metabolism/Metabolic pathway	0.07122743
Hmox1(24451)	Porphyrin and chlorophyll metabolis...	Glycan Biosynthesis and Metabolism/...	0.07122743
Abcb4(24891)	ABC transporters - General(rno02010)	Regulatory pathway	0.07122743
Gpx1(24404)	Glutathione metabolism(rno00480)	Amino Acid Metabolism/Metabolic pat...	0.07368694
Hadha(170670)	Valine, leucine and isoleucine degrad...	Nucleotide Metabolism/Metabolic pat...	0.08103131
Hadha(170670)	Fatty acid metabolism(rno00071)	Energy Metabolism/Metabolic pathway	0.10036778
Cyp1a1(24296)	Metabolism of xenobiotics by cytochro...	Regulatory pathway	0.12870497
Gpx1(24404)	Arachidonic acid metabolism(rno005...	Energy Metabolism/Metabolic pathway	0.13103075
Nfkb1(81736)	B cell receptor signaling pathway(rno...	Regulatory pathway	0.14487169
Nfkb1(81736)	Adipocytokine signaling pathway(rno0...	Regulatory pathway	0.16077533
Nfkb1(81736)	Toll-like receptor signaling pathway(r...	Regulatory pathway	0.16302601
Nfkb1(81736)	Pancreatic cancer(rno05212)	Regulatory pathway	0.16751155
Nfkb1(81736)	Small cell lung cancer(rno05222)	Regulatory pathway	0.16974642
Nfkb1(81736)	Apoptosis(rno04210)	Regulatory pathway	0.17641955
Stat5b(25126)	ErbB signaling pathway(rno04012)	Regulatory pathway	0.18962518
Nfkb1(81736)	Prostate cancer(rno05215)	Regulatory pathway	0.18962518

Input genes = 11, 8 genes found, 3 not found, Total 30 pathway maps.

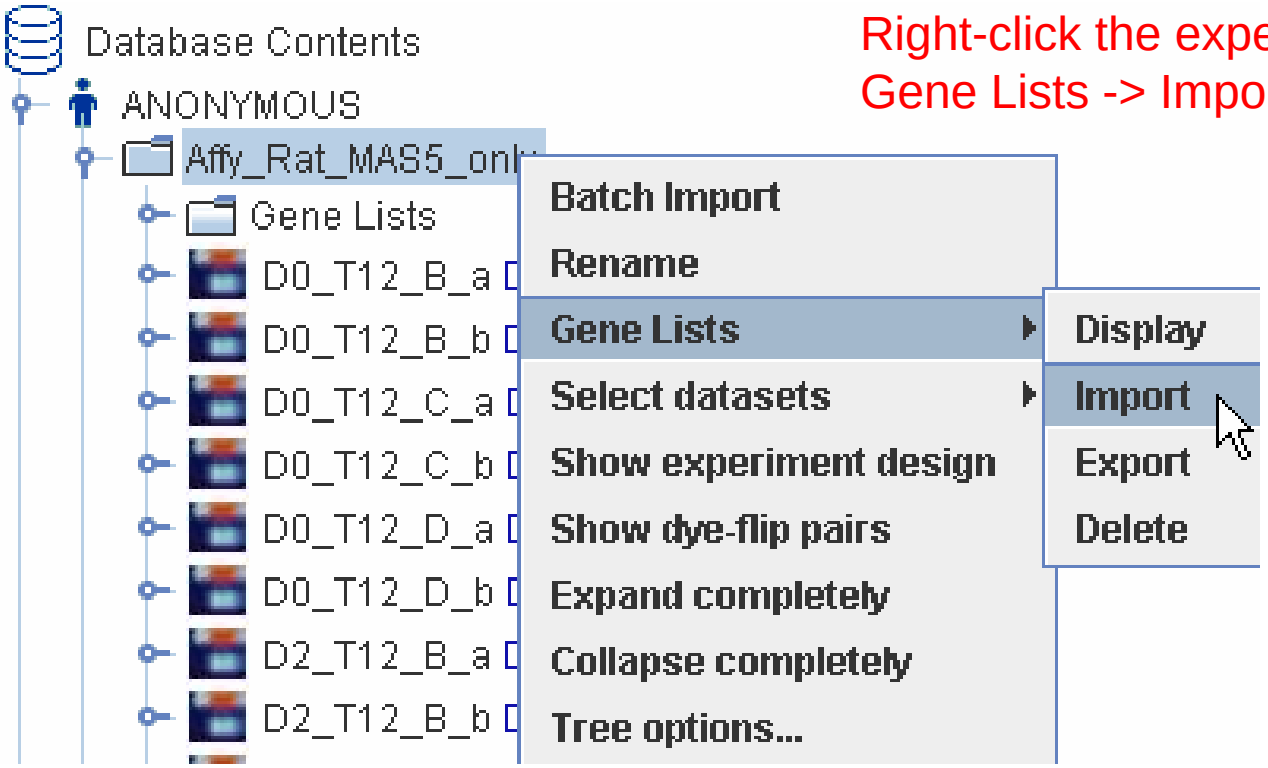
Right-click the pathway table, Then select “Create significant Gene list…”

Create Significant Gene List through Pathway

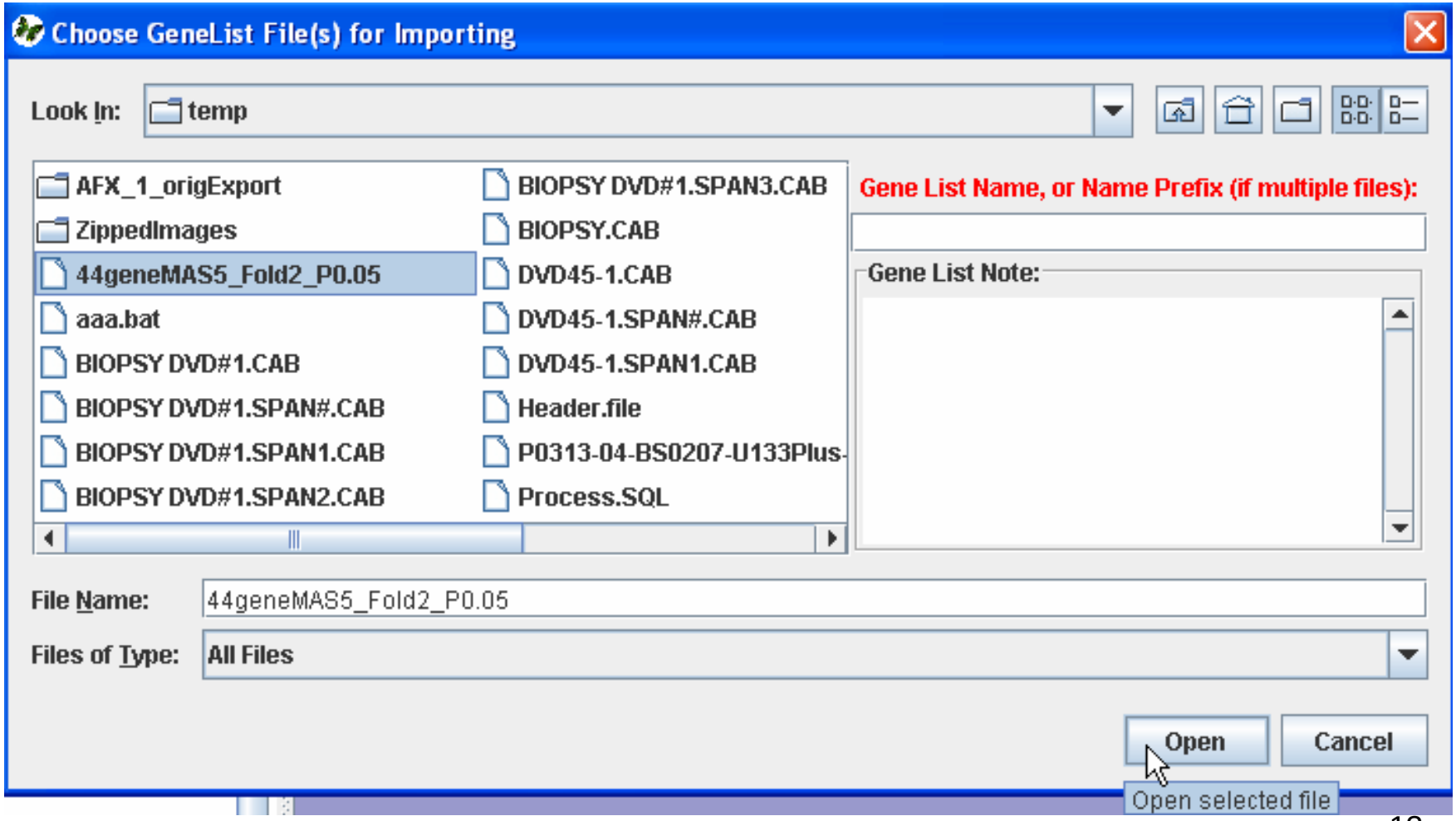


Import/Export a gene list

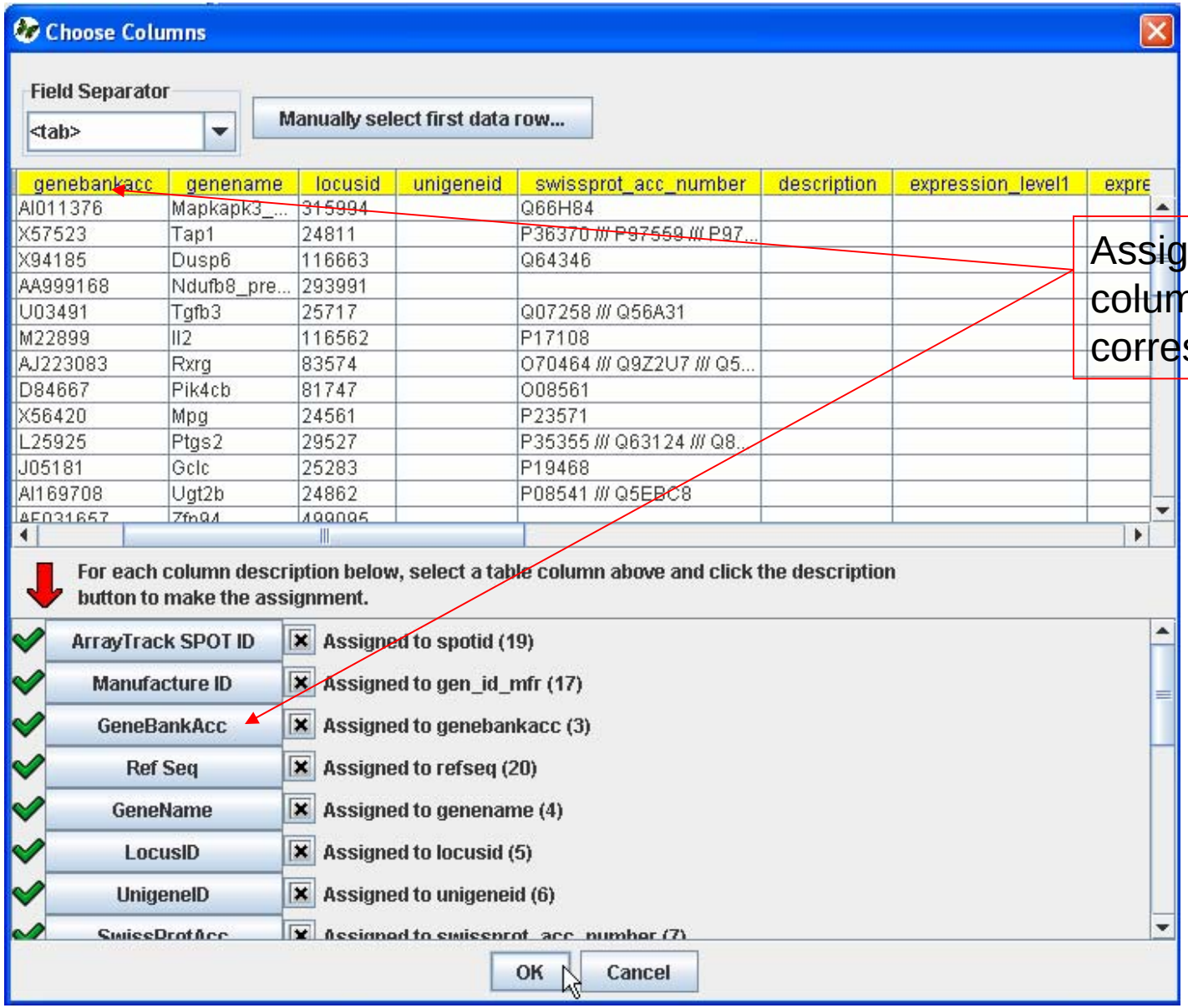
Right-click the experiment, then choose Gene Lists -> Import



Import/Export a gene list – cont.

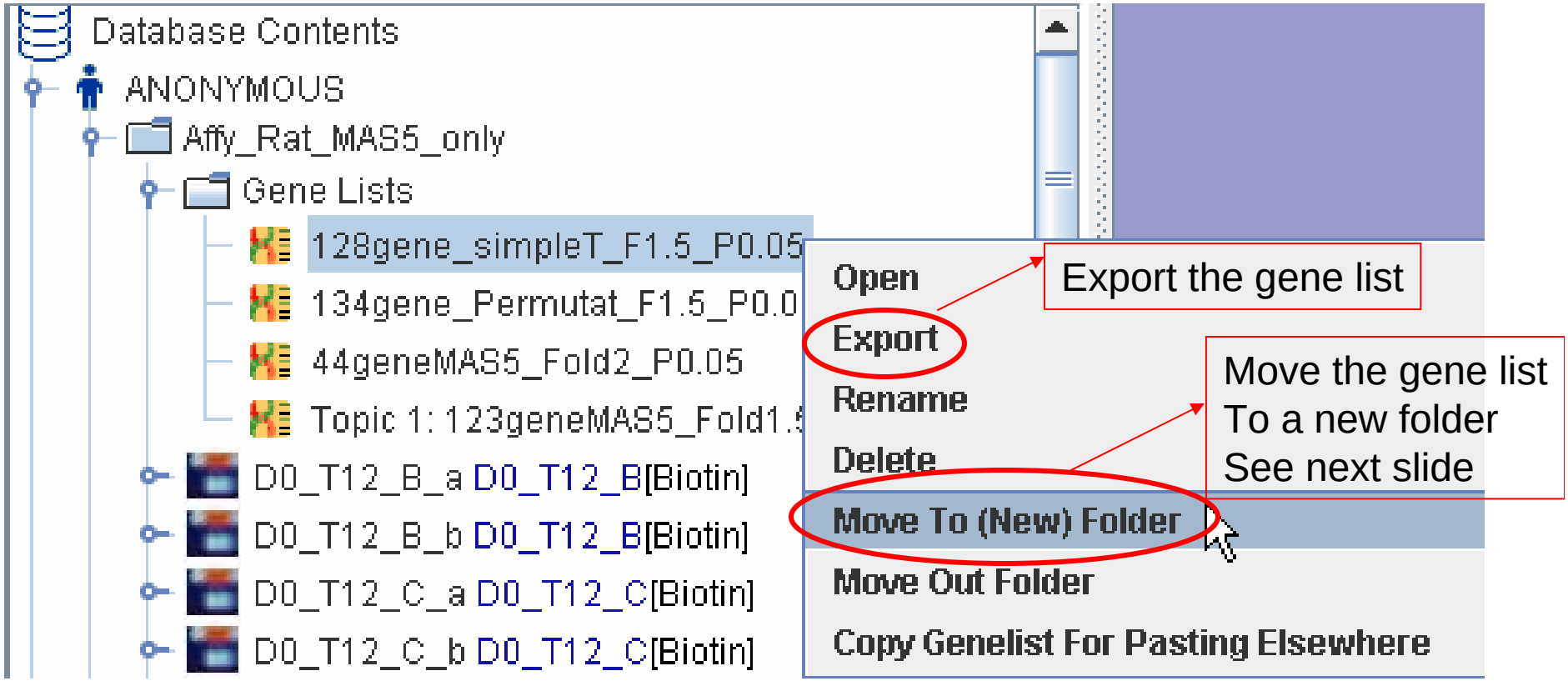


Import/Export a gene list – cont.



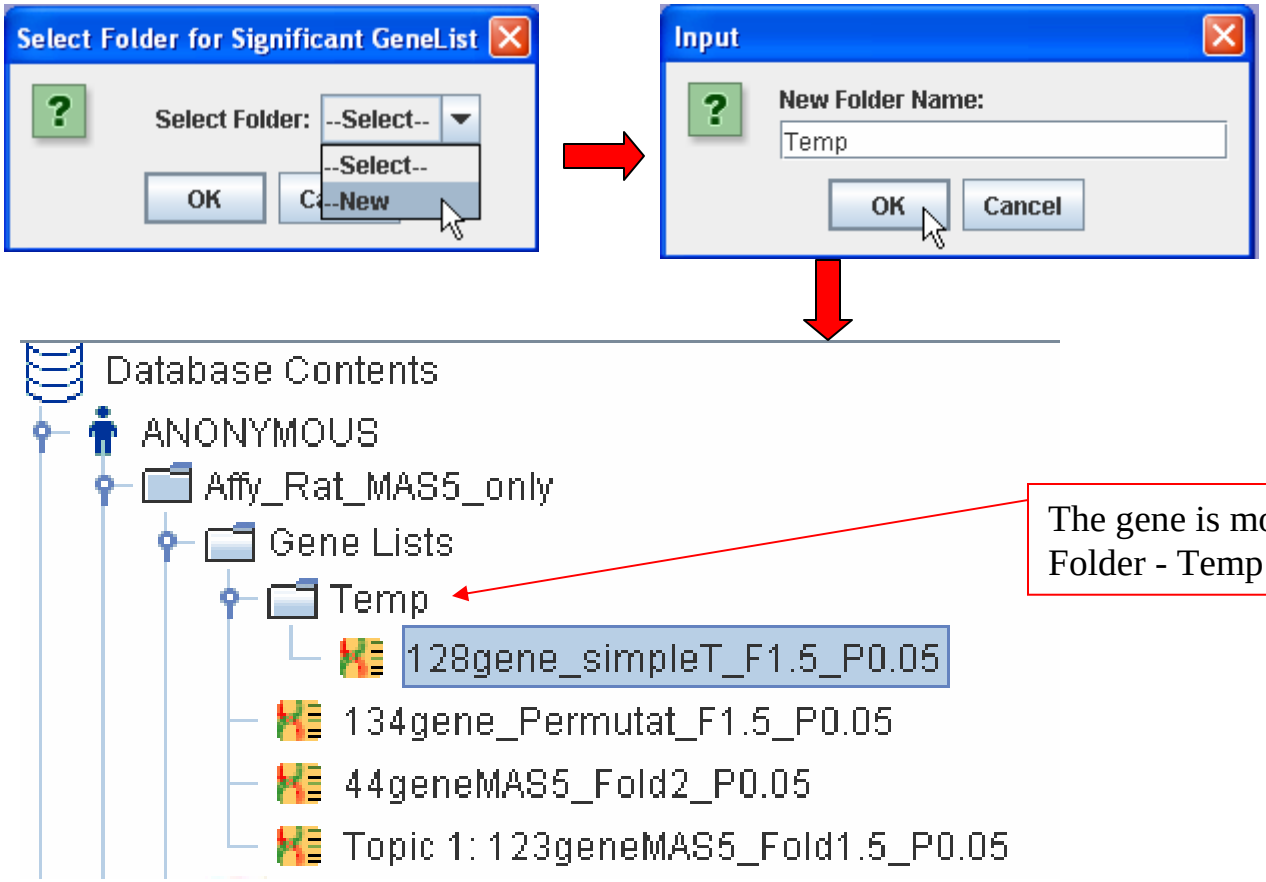
Assign table columns to the corresponding fields

Import/Export a gene list - continued

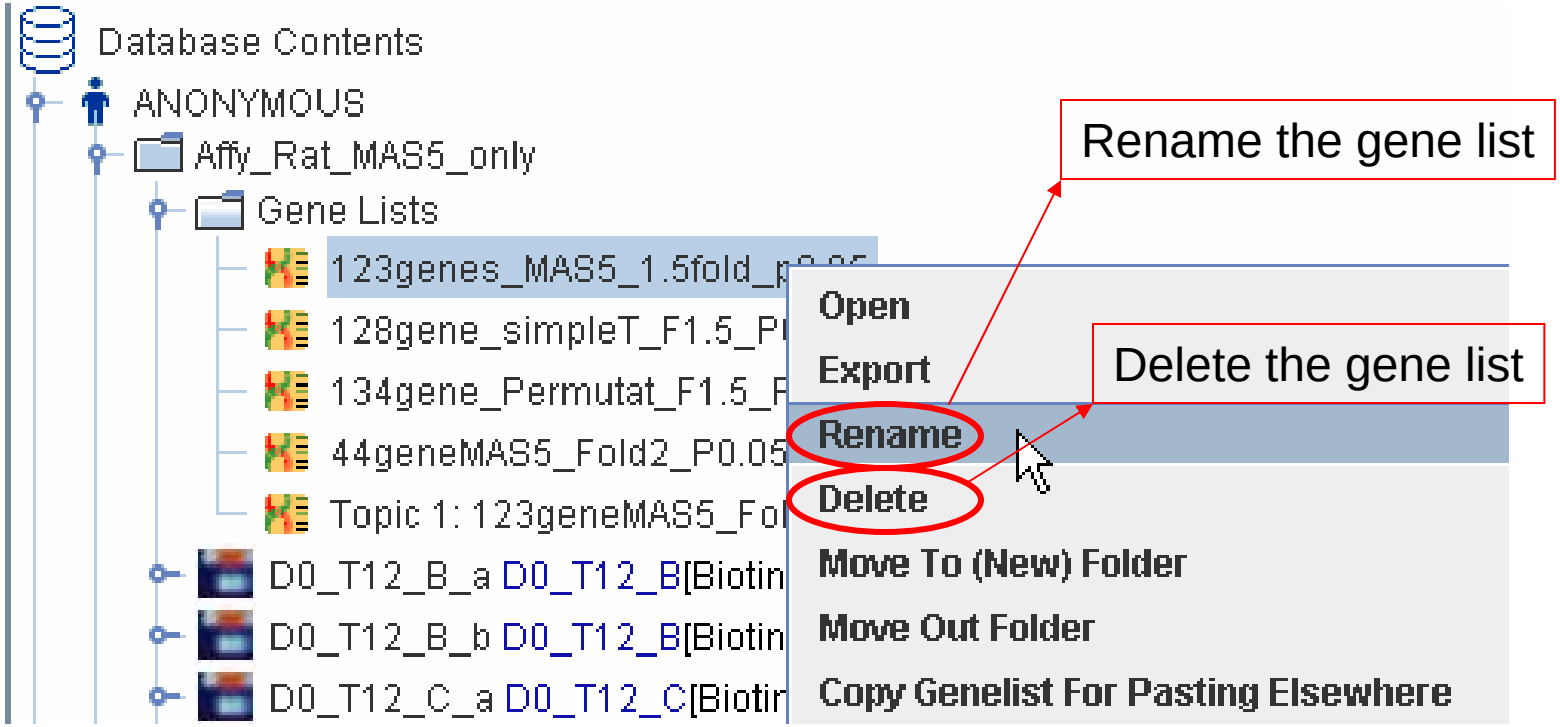


Import/Export a gene list - continued

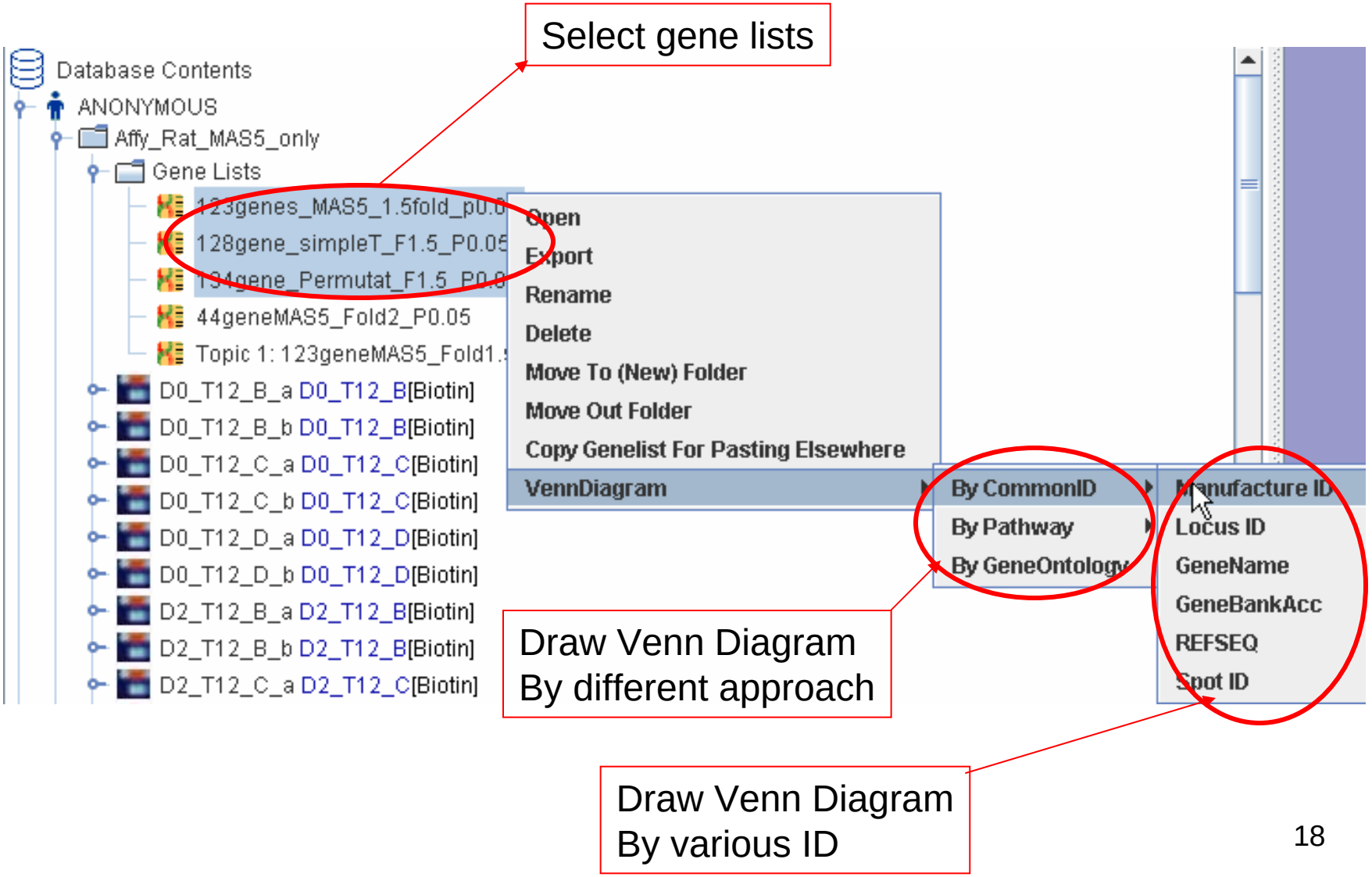
Move gene list to a new folder



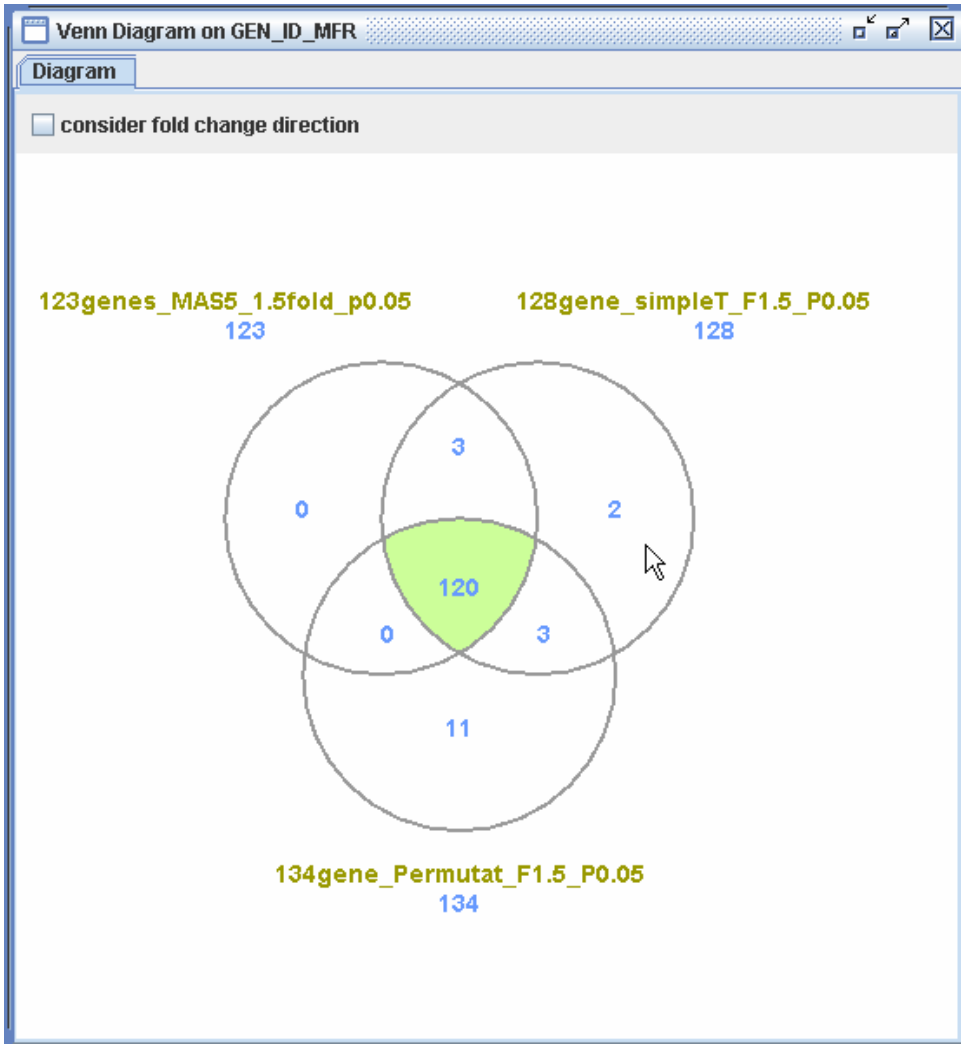
Import/Export a gene list - continued



Draw Venn Diagram from gene lists

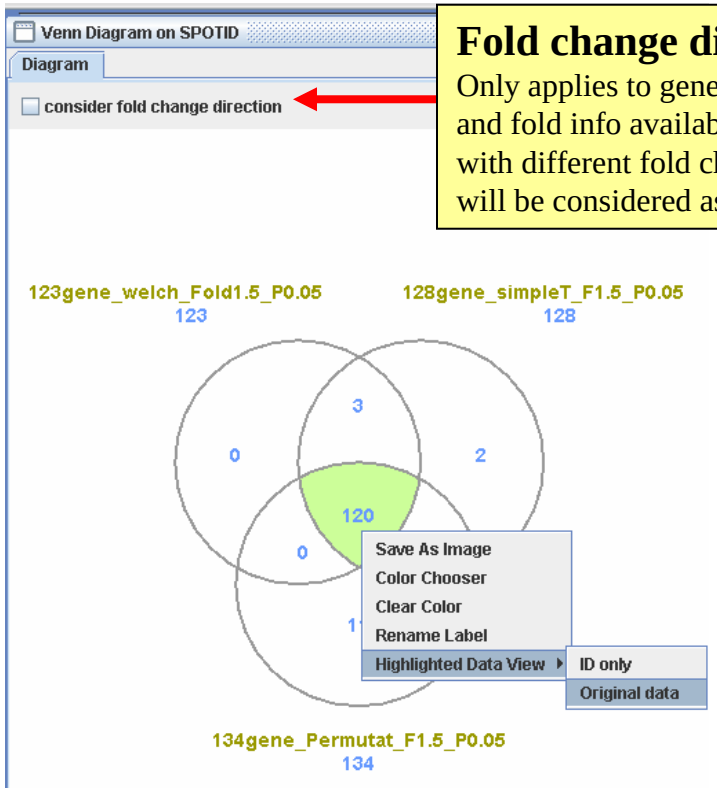


Draw Venn Diagram from gene lists -continued



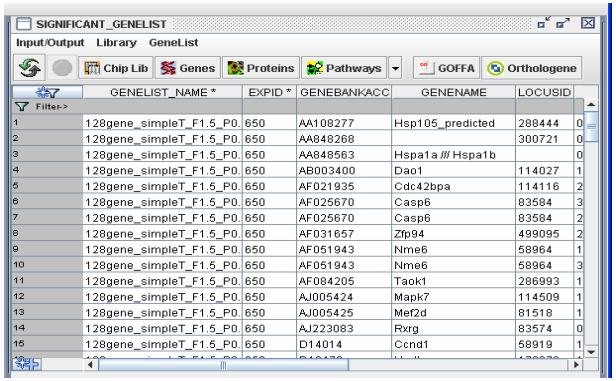
More detail about Venn Diagram is explained in Tutorial 3.

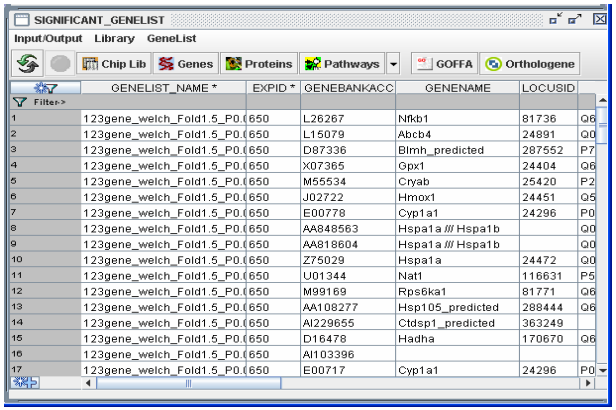
Draw Venn Diagram from gene lists -continued

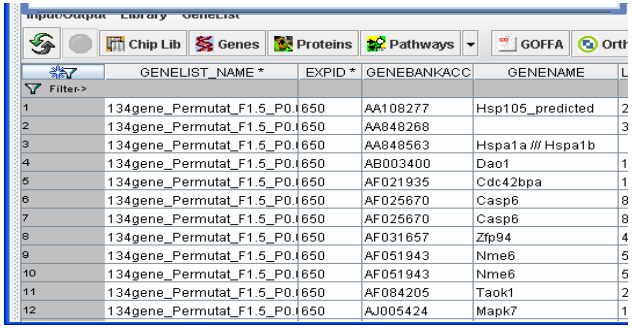


Fold change direction option

Only applies to genelists saved in ArrayTrack and fold info available. If gene id is same, but with different fold change direction, then it will be considered as different genes.



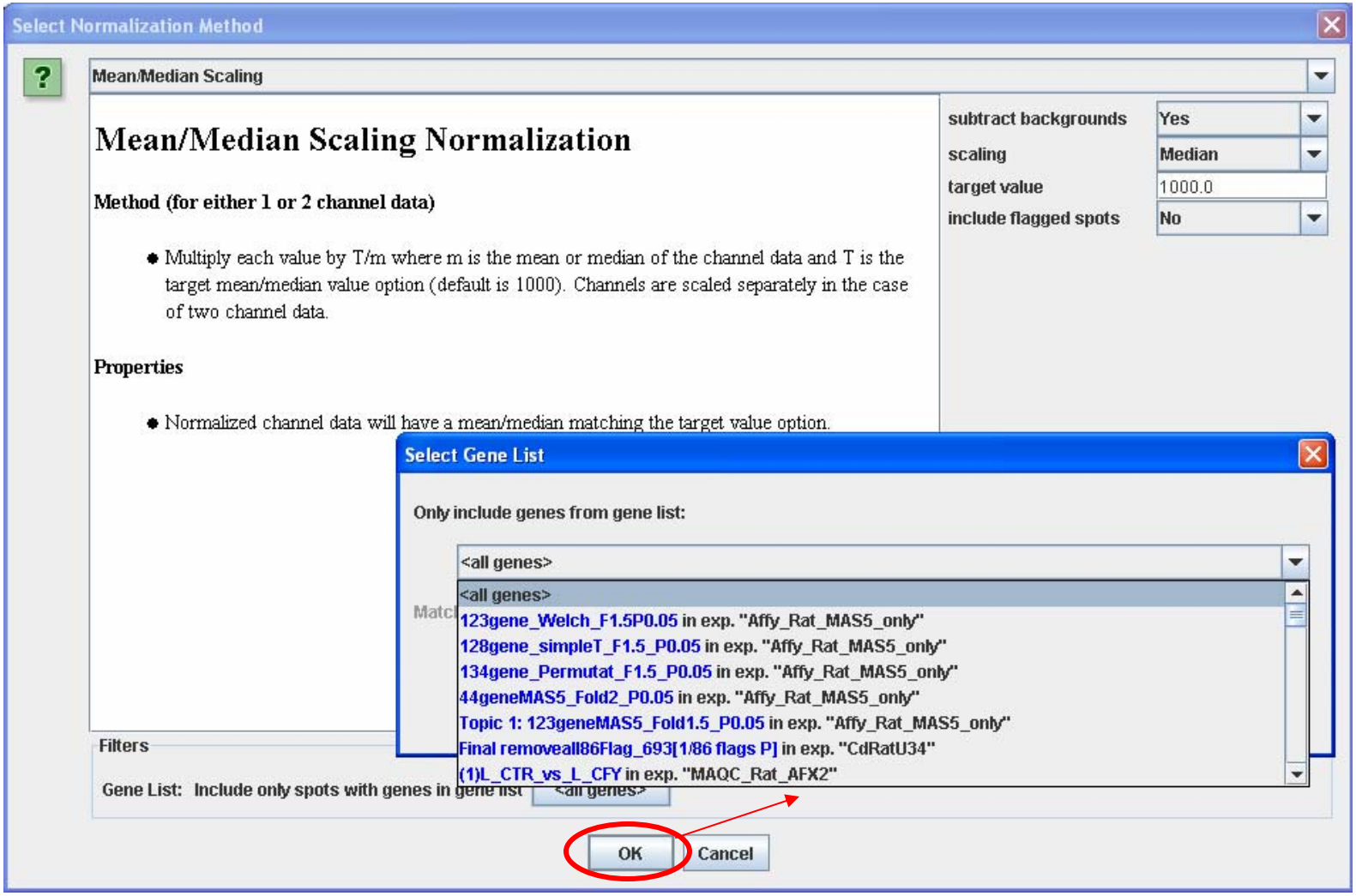




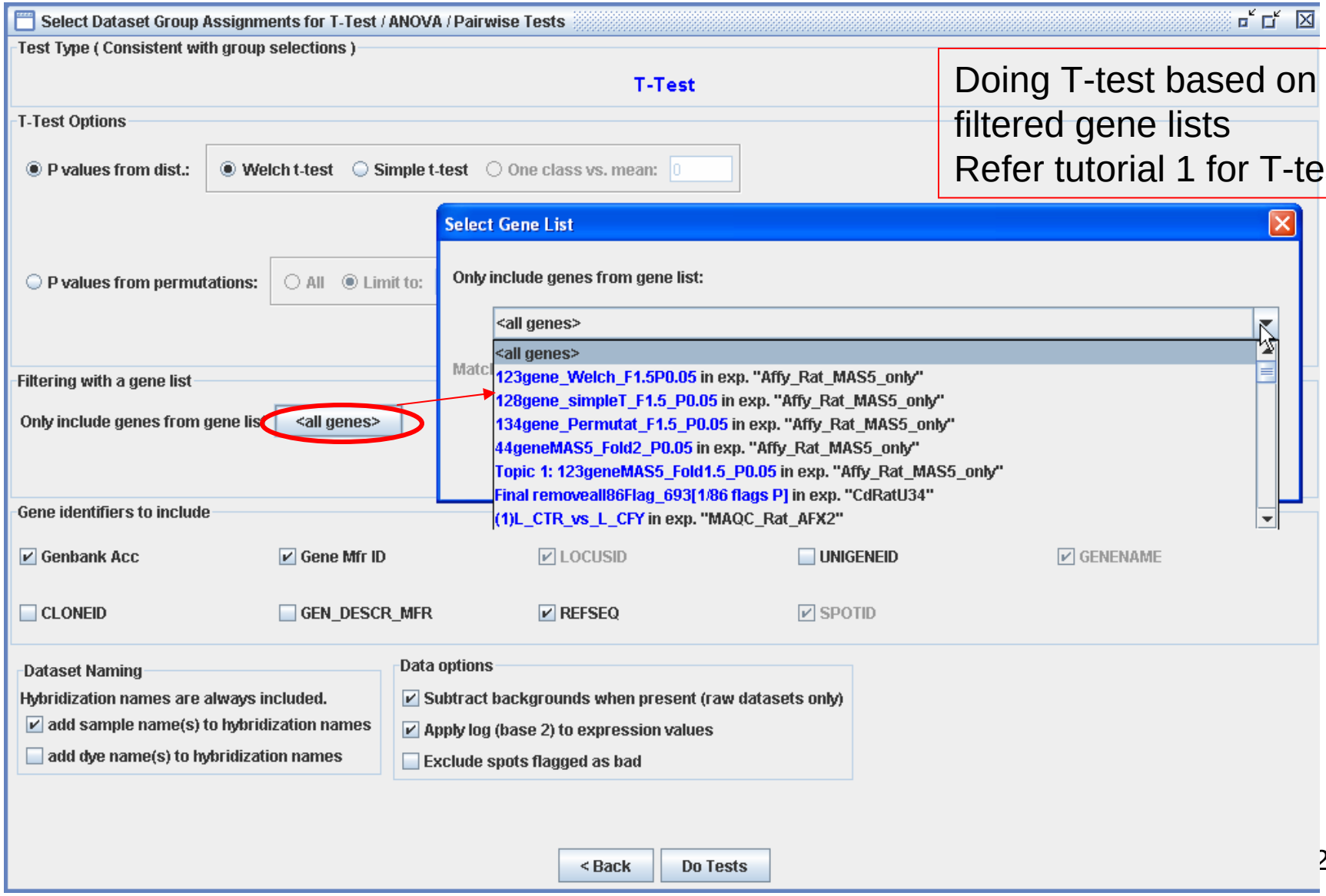
By clicking any sections from VennDiagram, the regions will be highlighted. Then right click to choose the options. One option allows display the selected genes in the original data sheet.

Conduct normalization filtered by a gene list

User can do data normalization based on a filtered gene list. Refer tutorial 7 for Normalization methods.



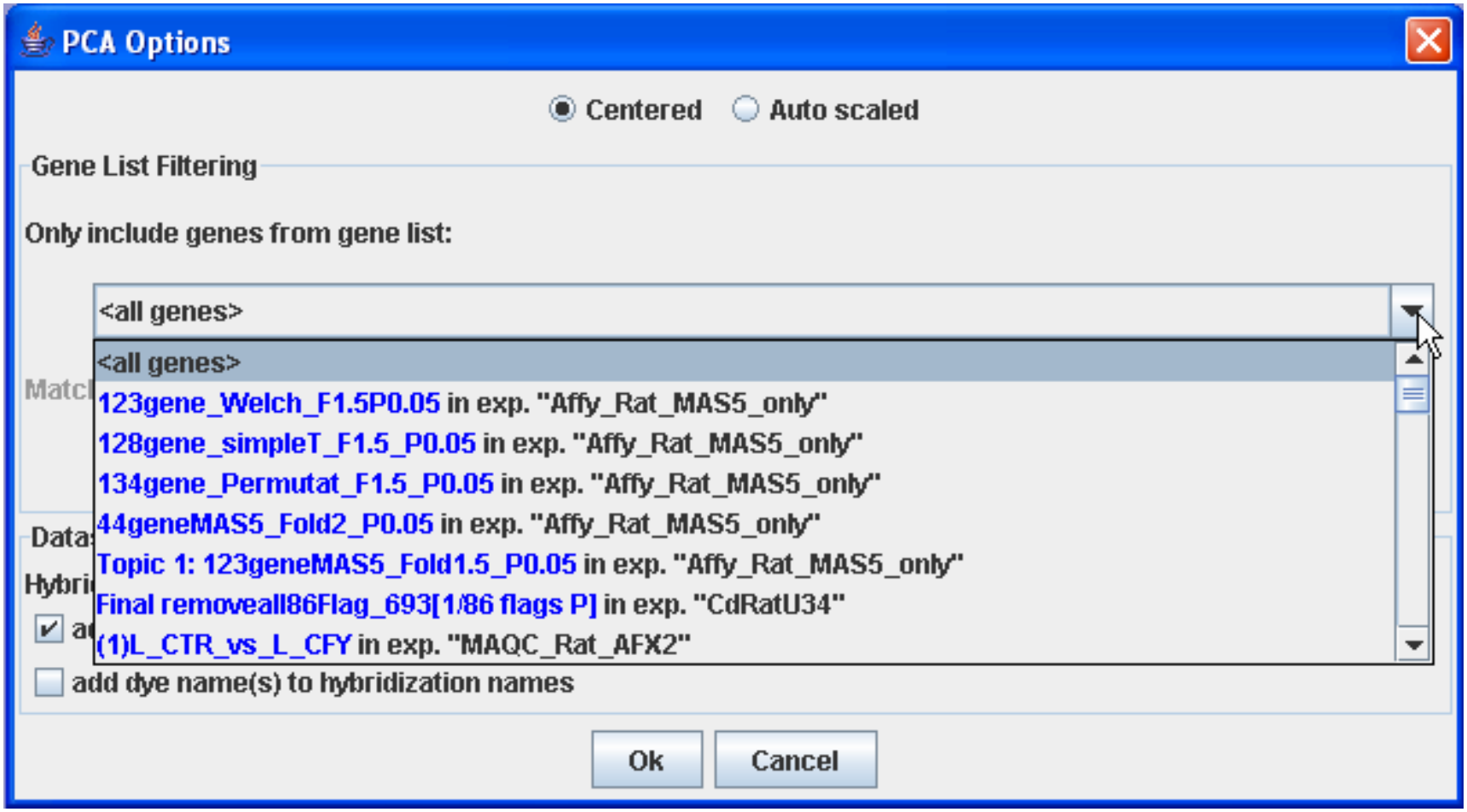
Conduct statistical analysis based on a gene list



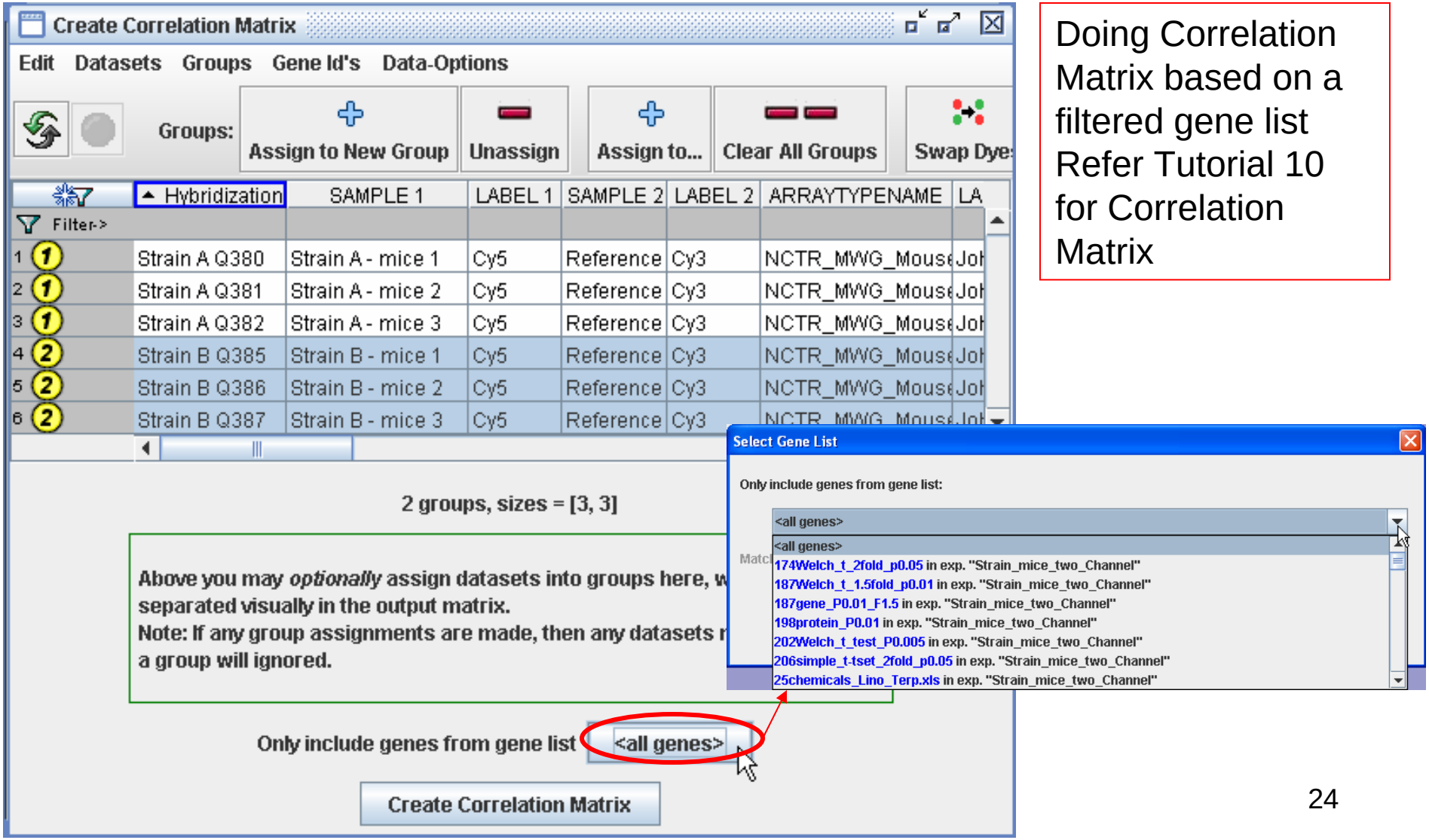
Doing T-test based on a filtered gene lists
Refer tutorial 1 for T-test

Conduct statistical analysis based on a gene list - continued

Doing PCA based on a filtered gene lists. Refer Tutorial 4 for PCA



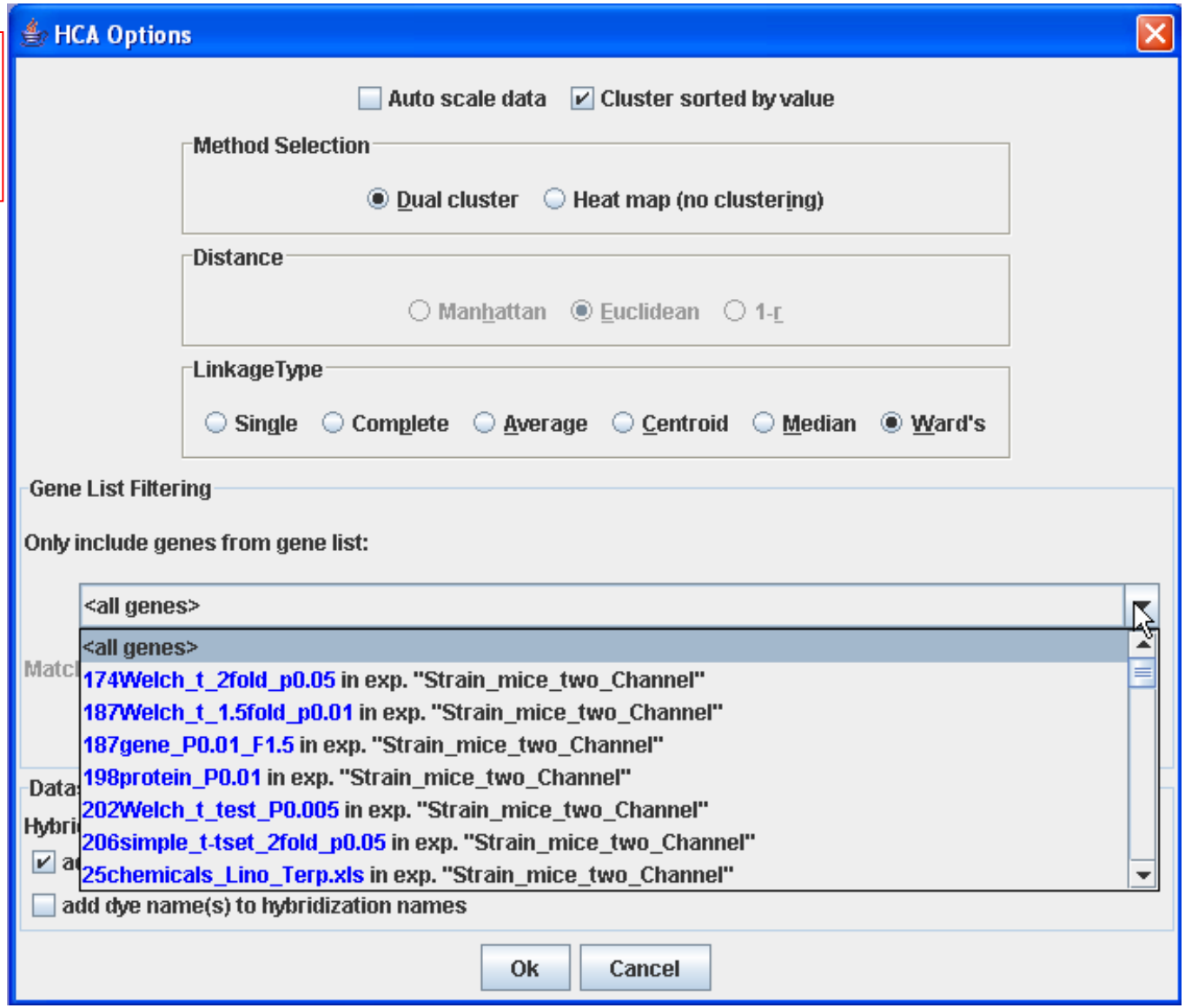
Conduct statistical analysis based on a gene list - continued



Doing Correlation Matrix based on a filtered gene list Refer Tutorial 10 for Correlation Matrix

Conduct statistical analysis based on a gene list - continued

Doing HCA based on
a filtered gene list
Refer Tutorial 4 for HCA

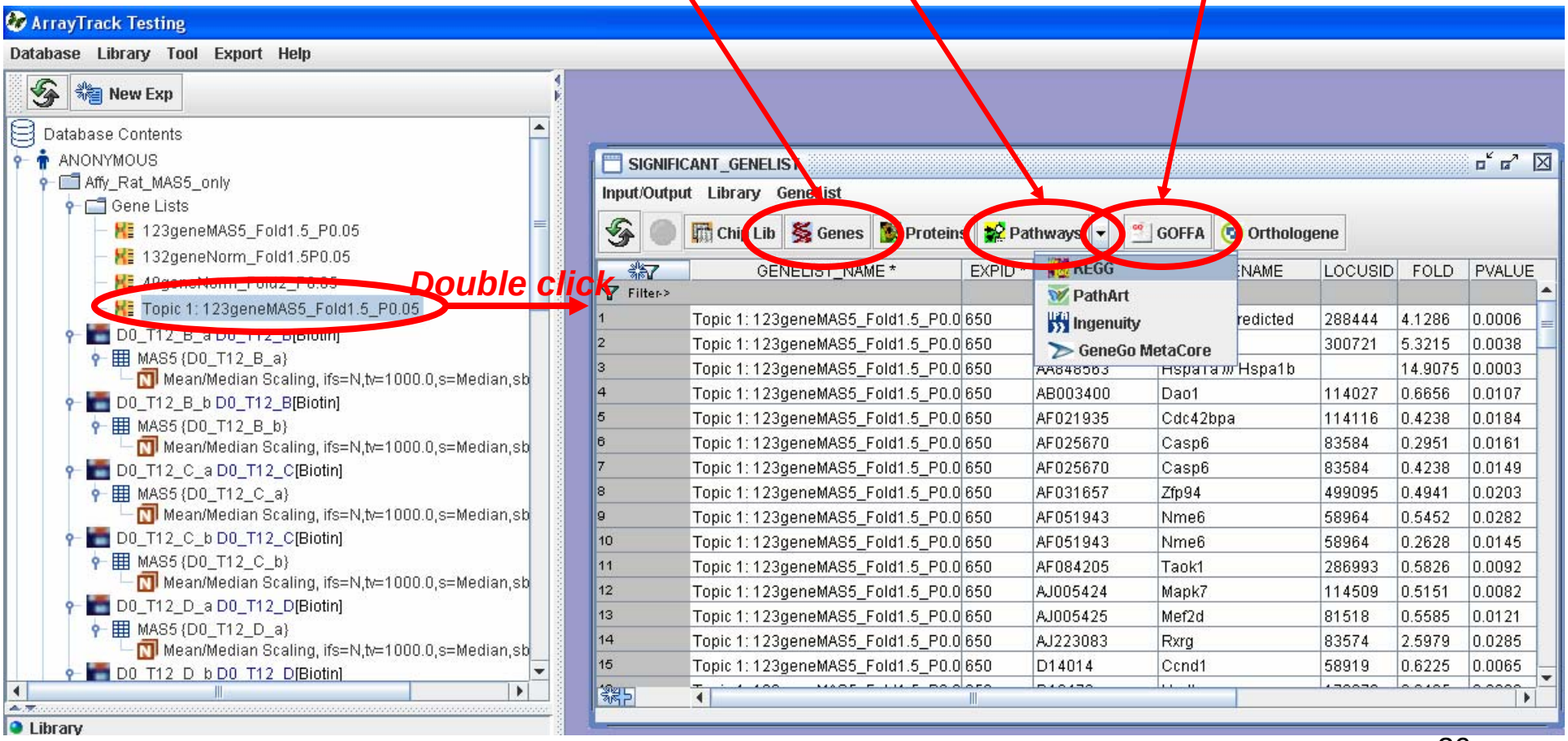


Biological interpretation using ArrayTrack Pathway, GO and other tools

Individual gene annotation and analysis

Pathway analysis

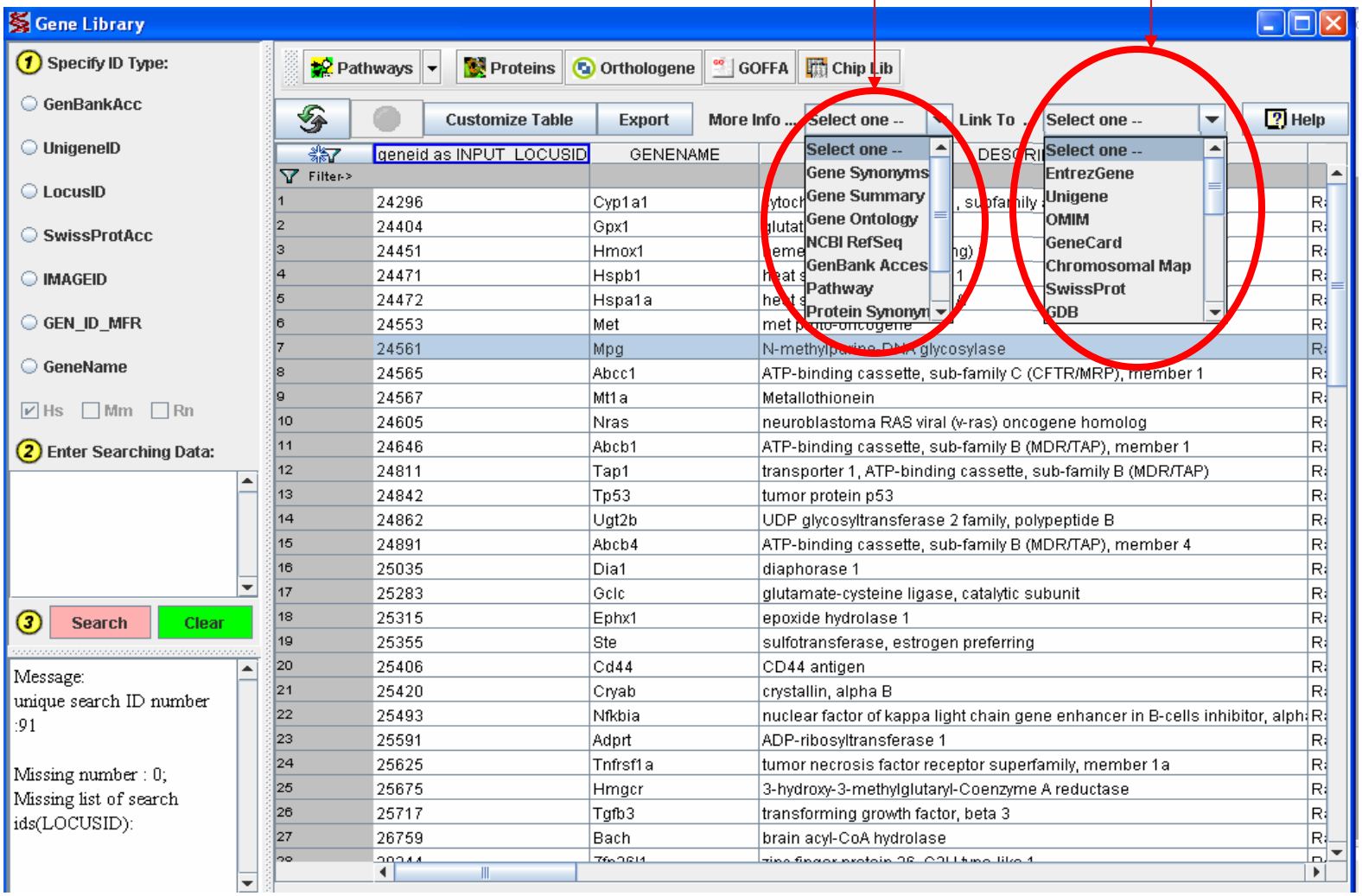
GO-based functional analysis

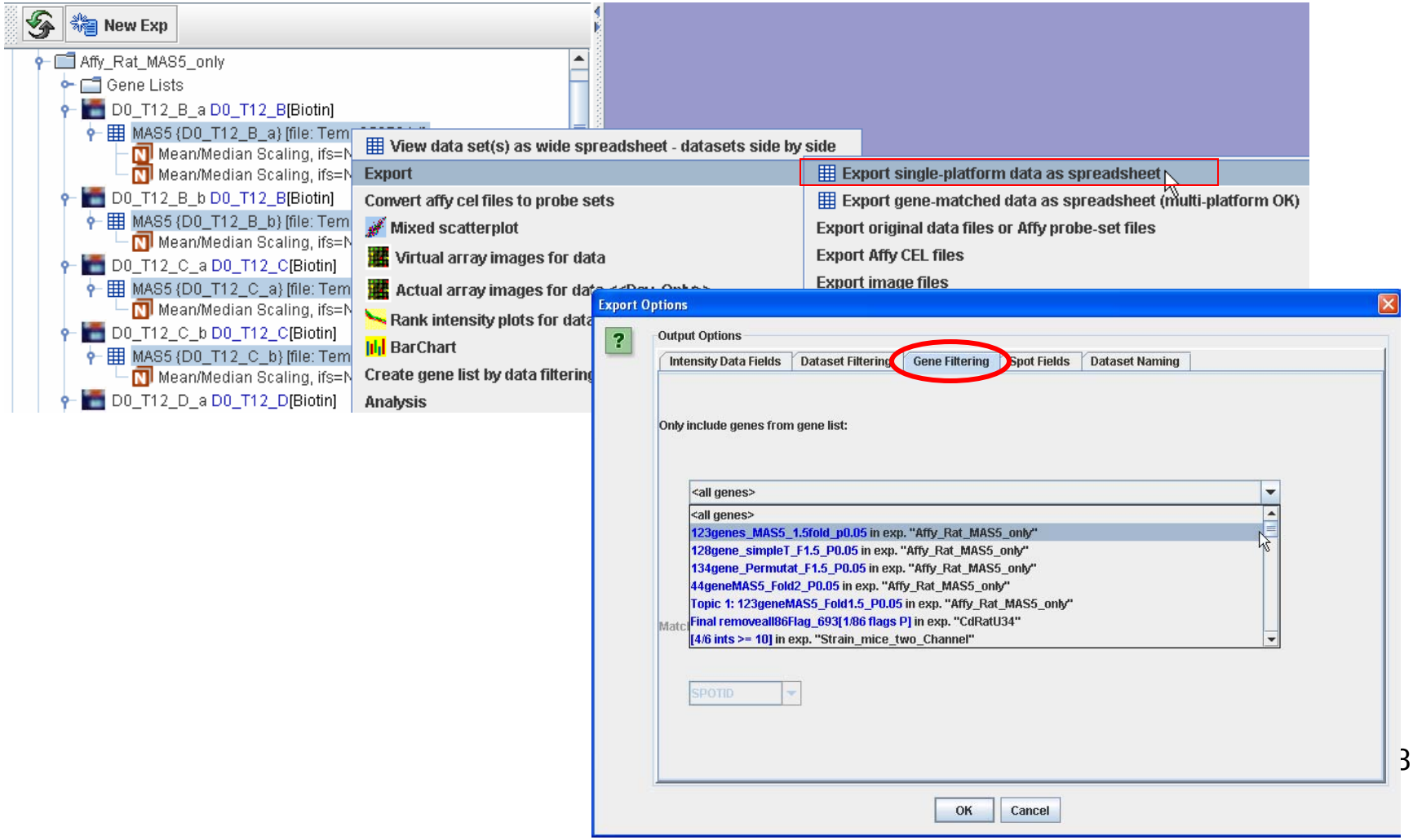


Individual gene annotation and analysis using GeneLib

Summary info for the highlighted gene

Link to other pub databases for the highlighted gene



Export a dataset by specifying the gene list

The user can export data with only selected genes. Refer Tutorial 9 for data export.