

Tutorial 3: VennDiagram



About VennDiagram

- **Why use VennDiagram:**

VennDiagram is a tool that visually displays the number of common characteristics among gene lists up to 3.

- **How to get Input to VennDiagram:**

<1> Gene List saved in ArrayTrack (the most popular usage).

<2> List saved on your local machine (can be gene/protein/metabolite list).

- **Functions associated with VennDiagram:**

<1> **Common Genes:** based on different gene IDs (GenBank accession #, RefSeq, and etc)

<2> **Common Pathways:** based on KEGG and PathArt

Advanced: across different omics data, such as a gene list from microarray experiment, a protein list from proteomics study, and a metabolite list from metabolomics study.

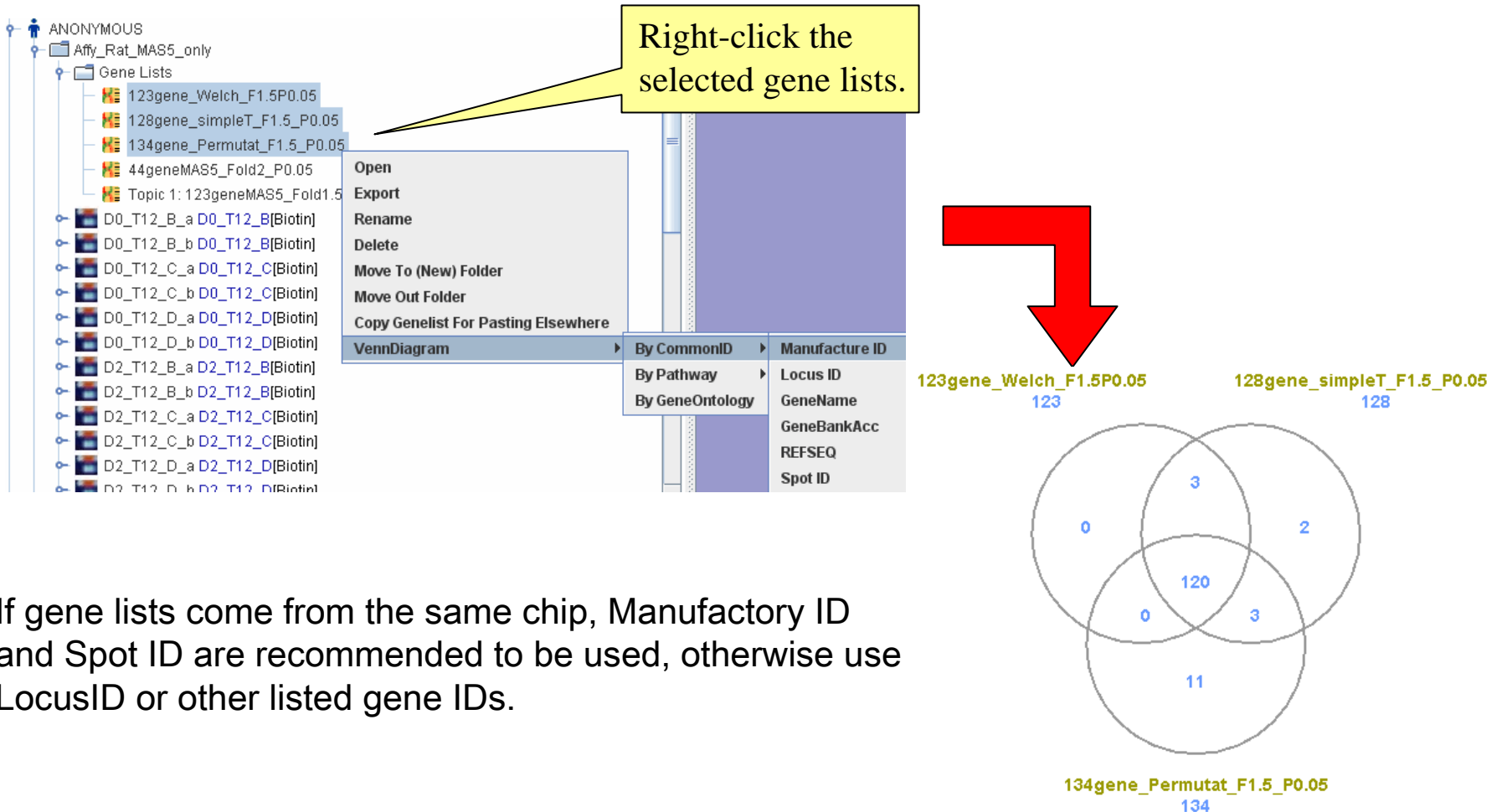
<3> **Common Gene Ontology terms**

<4> **Compare two array types**

Common Genes

- Two ways to input to VennDiagram :

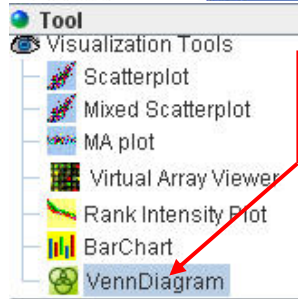
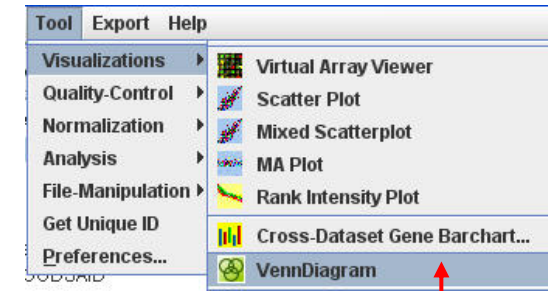
<1> when gene lists have been saved in ArrayTrack:



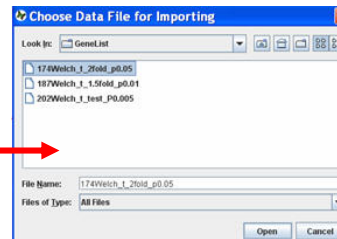
If gene lists come from the same chip, Manufacture ID and Spot ID are recommended to be used, otherwise use LocusID or other listed gene IDs.

Common Genes

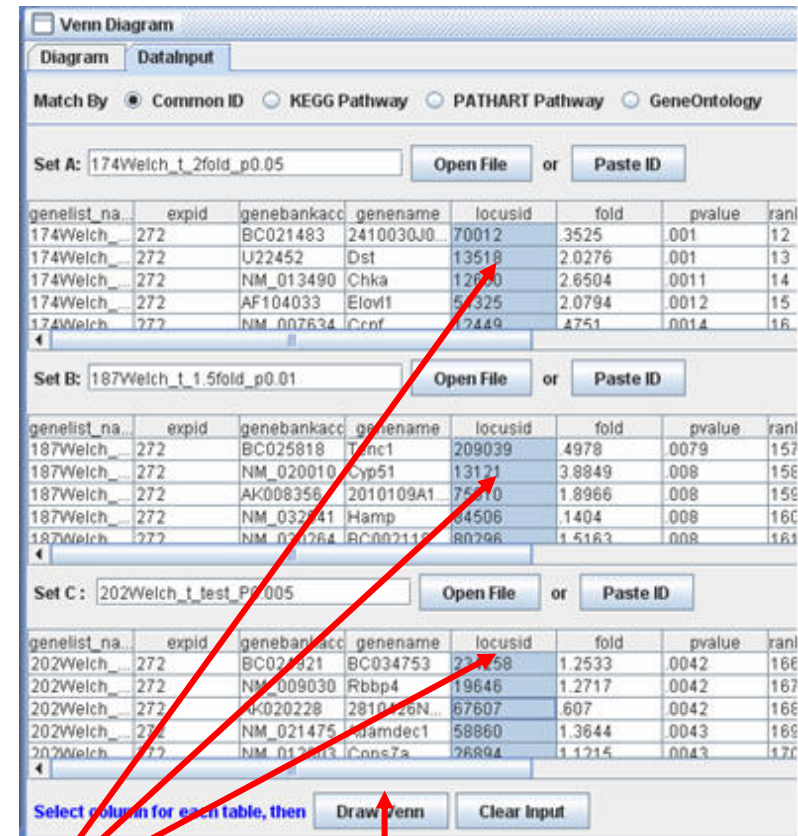
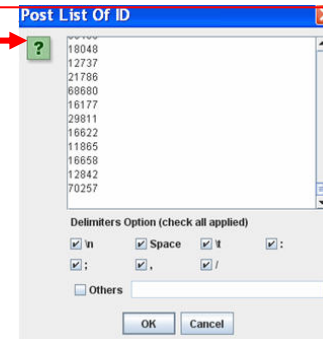
<2> when gene lists are saved in local machine:



1) Launch VennDiagram from Tool panel or Tool pull-down menu.



2) Either open a gene list file Or paste a list of genes.



3) Highlight gene column from loaded tables And click "Draw Venn" button

Common Pathways

Using VennDiagram, the common pathways shared by two or three lists can be identified. The lists could be genes from microarray study, proteins from the proteomics study or metabolites from metabolomics. Please see Topic 6 for uploading the protein and metabolite list.

At this point, ArrayTrack provides KEGG and ParthArt pathways. Thus, the common pathways are separately identified for KEGG and PathArt.

Common Pathways

<1> **Input:** Right-click the selected gene lists, choose VennDiagram by Pathway/Kegg

ArrayTrack Testing

Database Library Tool Export Help

New Exp

Affy_Rat_MAS5_only

Gene Lists

- 123genes_MAS5_1.5fold_p0.05
- 128gene_simpleT_F1.5_P0.05
- 134gene_Permutat_F1.5_P0.05
- 44gen
- Topic

Open

Export

Rename

Delete

Move To (New) Folder

Move Out Folder

Copy Genelist For Pasting Elsewhere

VennDiagram

By CommonID

By Pathway

- By KEGG
- By GeneOntology
- By PATHART

By GeneOntology

By PATHART

Match By Common ID KEGG Pathway PATHART Pathway GeneOntology

Set A: 123genes_MAS5_1.5fold_p0.05 or LOCUSID

GENE...	EXPID	GENEBANK...	GENENAME	LOCUSID	FOLD	PVALUE	RANKING...	GEN_ID_M...
s...	650	AI172097	Hsf1	79245	.635	.0116	110	rc_AI17209...
s...	650	AI172097	Hsf1	79245	.6547	.006	111	rc_AI17209...
s...	650	AI176422	Etfidh	295143	.5907	.0175	112	rc_AI17642...
s...	650	AI176658	Hspb1	24471	1.6388	.0092	113	rc_AI17665...
s...	650	AI229421	Mapkapk3...	315994	.6052	.0231	114	rc_AI22942...

Set B: 128gene_simpleT_F1.5_P0.05 or LOCUSID

GENE...	EXPID	GENEBANK...	GENENAME	LOCUSID	FOLD	PVALUE	RANKING...	GEN_ID_M...
128gene_s...	650	AI229421	Mapkapk3...	315994	1.6524	.0159	119	rc_AI22942...
128gene_s...	650	AI229655	Ctdsp1_pr...	363249	1.5386	.0001	120	rc_AI22965...
128gene_s...	650	AI230778	Tpst2_pred...	288719	1.6781	.0006	121	rc_AI23077...
128gene_s...	650	AI231354	Mapk9	50658	1.6711	.0292	122	rc_AI23135...

Set C: 134gene_Permutat_F1.5_P0.05 or LOCUSID

GENELIST...	EXPID	GENEBANK...	GENENAME	LOCUSID	FOLD	PVALUE	RANKING...	GEN_ID_M...
134gene_...	650	AI171243	Rpa3_pred...	296883	1.6352	0	119	rc_AI17124...
134gene_...	650	AI171630	Mapk14	81649	1.6128	.015	120	rc_AI17163...
134gene_...	650	AI172097	Hsf1	79245	1.5749	0	121	rc_AI17209...
134gene_...	650	AI172097	Hsf1	79245	1.5274	.005	122	rc_AI17209...
134gene_...	650	AI176422	Etfidh	295143	1.6928	0	123	rc_AI17642...

Select column for each table, then

Library

- ID Converter
- Gene Library
- Pathway Library
- Protein Library
- IPI Library
- Orthologene Library
- GOFFA Library
- Chip Library

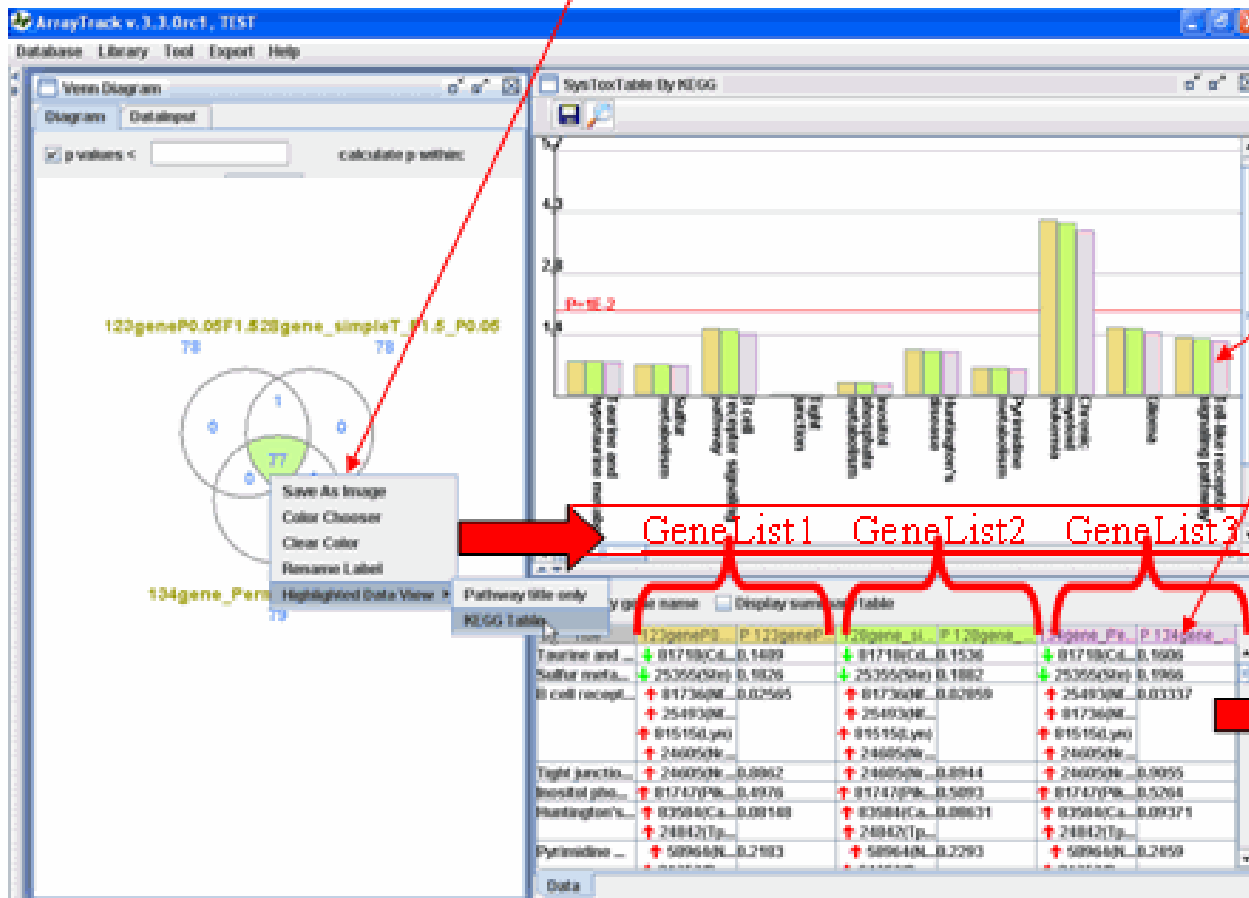
Tool

- Analysis
- T-Test
- ANOVA
- P-Value Print

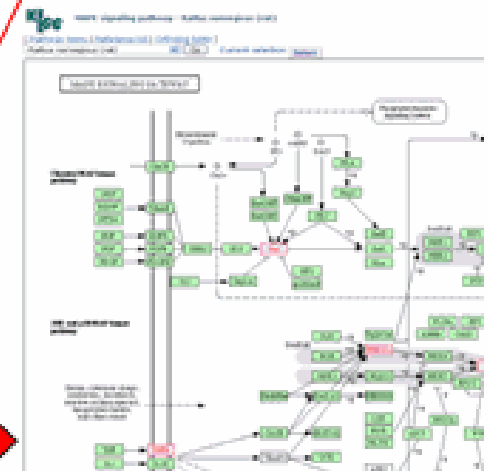
Highlight common LOCUS ID then click Draw Venn button

Common Pathways

<2> **Display Pathway:** Show common Pathway by right-click of the highlighted green area of the VennDiagram.



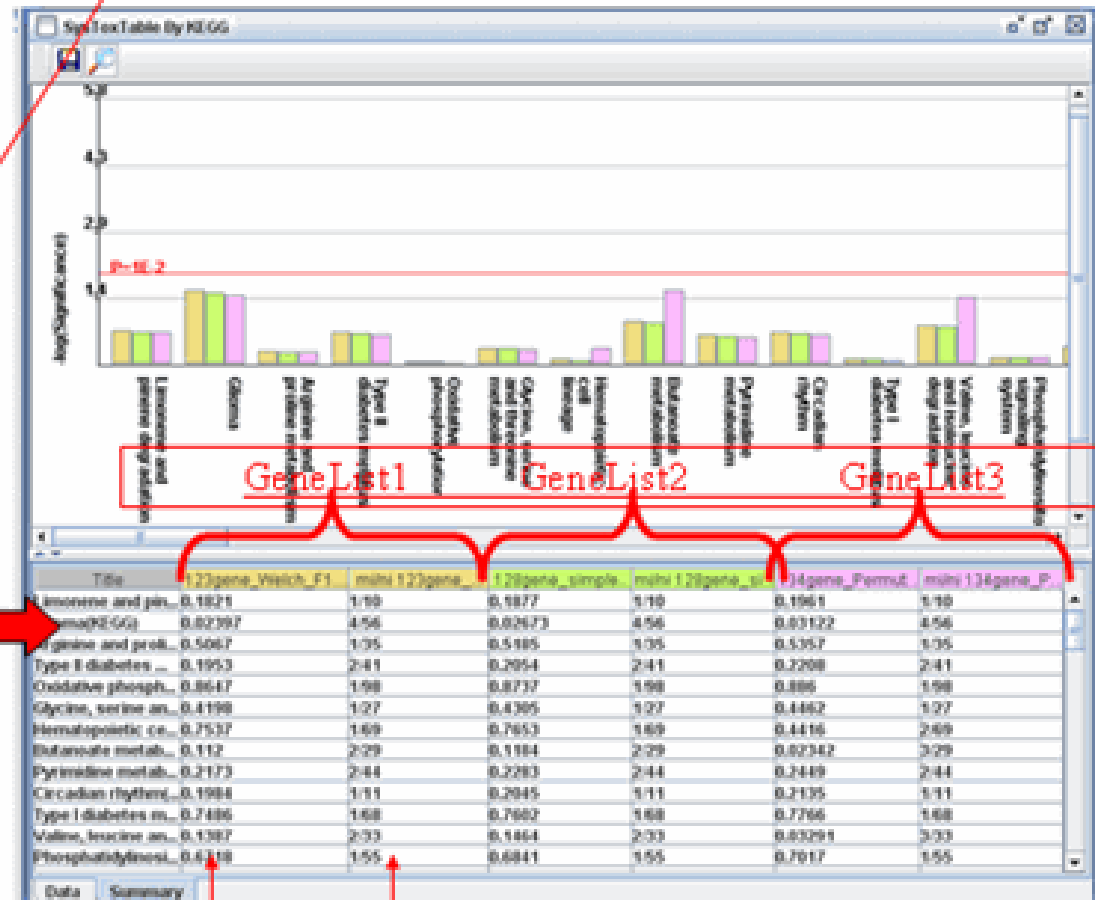
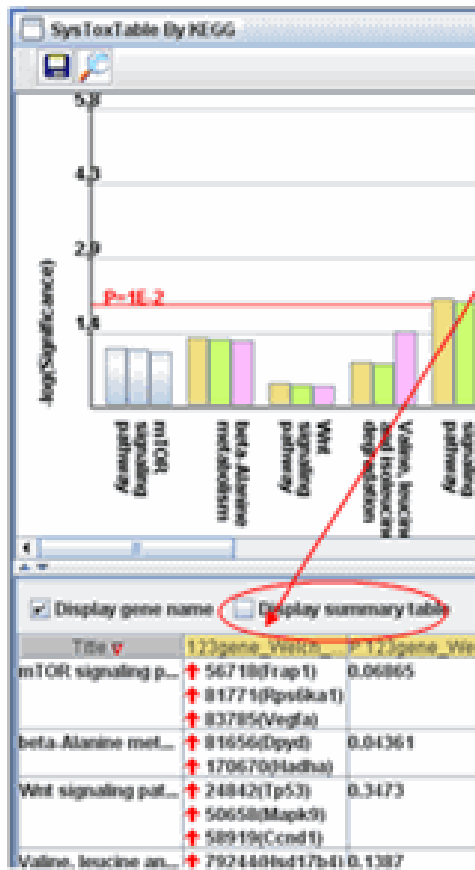
Different colors represent different gene lists.



Double click record on table to display pathway map (genes are highlighted)

Common Pathways

<3> Display Summary: Select check box to show summary table for Common Pathway.



P value

1/55 : For pathway "Phosphatidylinositol signaling system", 55 genes are found in gene library for this pathway 1 gene is found in this gene list.

Common Pathways

<4> P value cut: calculate p by Fisher Exact Test.

(1) After apply $p < 0.05$

(2) Display pathway table & chart

(3) Click table header to sort

(4) Double click record to show pathway map

Common pathways

Advanced: shared by the genes, proteins and metabolites

VennDiagram can also be used to find common pathway based on gene list, protein list and metabolite list. Rank of significant of P is displayed in a summary table. You can click individual KEGG pathway which links to KEGG website.

The screenshot shows the VennDiagram software interface. The main window displays a Venn diagram with three overlapping sets: 'Set B: (protein_P0.05_389)', 'Set C: (_metabolitesP0.05final_17)', and their intersection. Below the diagram are three data tables. The first table, 'By CommonID', lists genes with columns for ID, EXPID, GENEID, GENEID, LOCUSID, FOLD, PVALUE, and RANK. The second table, 'By Pathway', lists proteins with columns for ID, GENEID, LOCUSID, SWISSPRO, FOLD, PVALUE, RANK, and COMPOUN. The third table, 'By GeneOntology', lists metabolites with columns for NAME, EXPRESSI, EXPRESSI, FOLD, PVALUE, COMPOUN, COMPOUN, and REFSEQ. Red circles highlight specific entries in each table, with red arrows pointing to their respective labels on the right: 'Genes (i.e. locusid)', 'Proteins (i.e. locusid or SwissProt ID)', and 'Metabolites (i.e. Compound Cas #)'. The 'Open File' and 'Paste ID' buttons are visible for each table.

ID	EXPID	GENEID	GENEID	LOCUSID	FOLD	PVALUE	RANK
P0.0_843	843	NM_019216	Odr15	29455	1.653	0	13
P0.0_843	843	NM_013083	Perit	25511	243	0	12
		H7995	Scd2	83792	092	0	14
				117517	3.918	0	15

ID	GENEID	LOCUSID	SWISSPRO	FOLD	PVALUE	RANK	COMPOUN
	Modc2	171491		4623	0018	72	
	Fcgs	83791	P0.089	7385	0018	73	
	Car1_pndi	310218		3781	0018	74	
	ML2	288174		719	0018	75	

NAME	EXPRESSI	EXPRESSI	FOLD	PVALUE	COMPOUN	COMPOUN	REFSEQ
ite	7.553	8.584	489	001	110-17-0		
he	7.354	8.274	528	002	70-26-8		
te	7.799	8.777	508	002	1190-49-4		
spho	8.588	9.527	532	002	107-73-3		

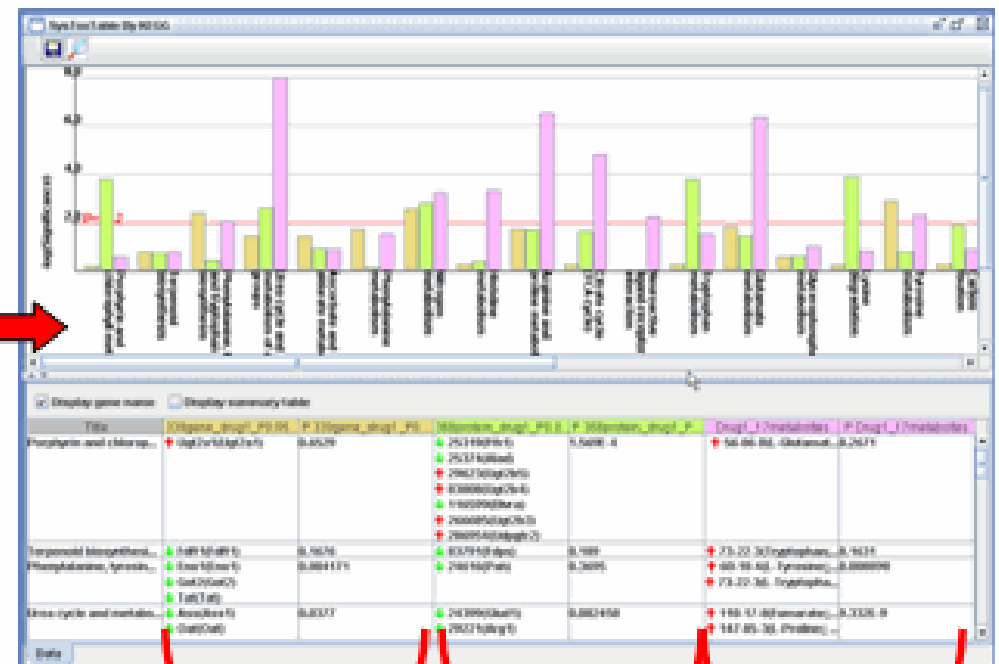
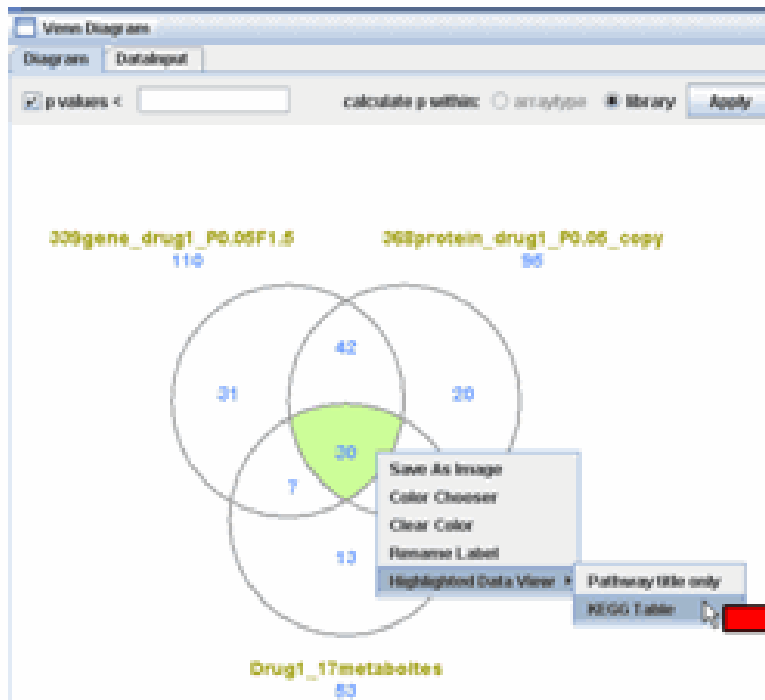
Genes
(i.e. locusid)

Proteins
(i.e. locusid or SwissProt ID)

Metabolites
(i.e. Compound Cas #)

Common pathways

Advanced: shared by the genes, proteins and metabolites



GeneList

ProteinList

MetaboliteList

Common Gene Ontology terms

GOFFA (Gene Ontology For Functional Analysis) is a specific GO-based tool implemented in ArrayTrack. The common GeneOntology function identifies the common Gene Ontology terms shared by two or three gene lists. Identification of the common Gene Ontology terms is similar to the procedure used for determine the common pathways. The common biological process, molecular function and Cellular component terms are presented in results tables.

Highlight common LOCUS ID then click Draw Venn

The screenshot shows the ArrayTrack Testing software interface. The main window is titled "Venn Diagram" and has two tabs: "Diagram" and "DataInput". The "DataInput" tab is active, showing three sets of data for comparison. The "Match By" options are "Common ID", "KEGG Pathway", "PATHART Pathway", and "GeneOntology", with "GeneOntology" selected. The "Set A" input is "123gene_welch_Fold1.5_P0.05", "Set B" is "128gene_simpleT_F1.5_P0.05", and "Set C" is "134gene_Permutat_F1.5_P0.05". Each set has a table of data with columns: GENELIST..., EXPID, GENE BANK..., GENENAME, LOCUSID, UNIGENEID, and SWISSPRO... D. The "LOCUSID" column is highlighted in blue, and a red circle is drawn around the common LOCUS IDs: 65030, 315994, and 288719. A red arrow points to the "By GeneOntology" option in the "VennDiagram" menu. At the bottom, the "Draw Venn" button is circled in red.

Database Library Tool Export Help

New Exp

Database Contents

- ANONYMOUS
 - Affy_Rat_MAS5_only
 - Gene Lists
 - 123gene_welch_Fold1.5_P0.05
 - 128gene_simpleT_F1.5_P0.05

Open

Export

Rename

Delete

Move To (New) Folder

Move Out Folder

Copy genelist for pasting elsewhere

VennDiagram

- By CommonID
- By Pathway
- By GeneOntology

Library

- ID Converter
- Gene Library
- Pathway Library
- Protein Library
- IPI Library
- Orthologene Library
- GOFFA Library
- Chip Library

Tool

- Analysis
 - T-Test
 - ANOVA

Venn Diagram

Diagram DataInput

Match By Common ID KEGG Pathway PATHART Pathway GeneOntology

Set A: 123gene_welch_Fold1.5_P0.05 or

GENELIST...	EXPID	GENE BANK...	GENENAME	LOCUSID	UNIGENEID	SWISSPRO... D
123gene_...	650	X60328	Ephx2	65030		P80299 III ...
123gene_...	650	AI104882	Ephx2	65030		P80299 III ...
123gene_...	650	J02752	Acox1	50681		P07872
123gene_...	650	AI011376	Mapkapk3_...	315994		Q66H84

Set B: 128gene_simpleT_F1.5_P0.05 or

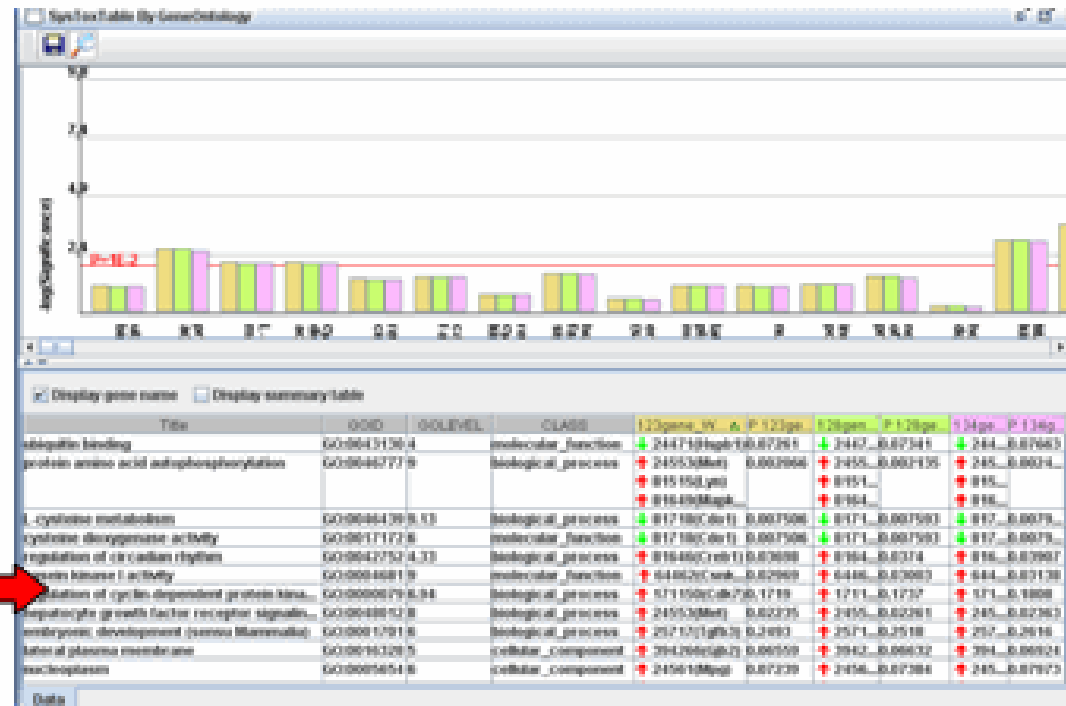
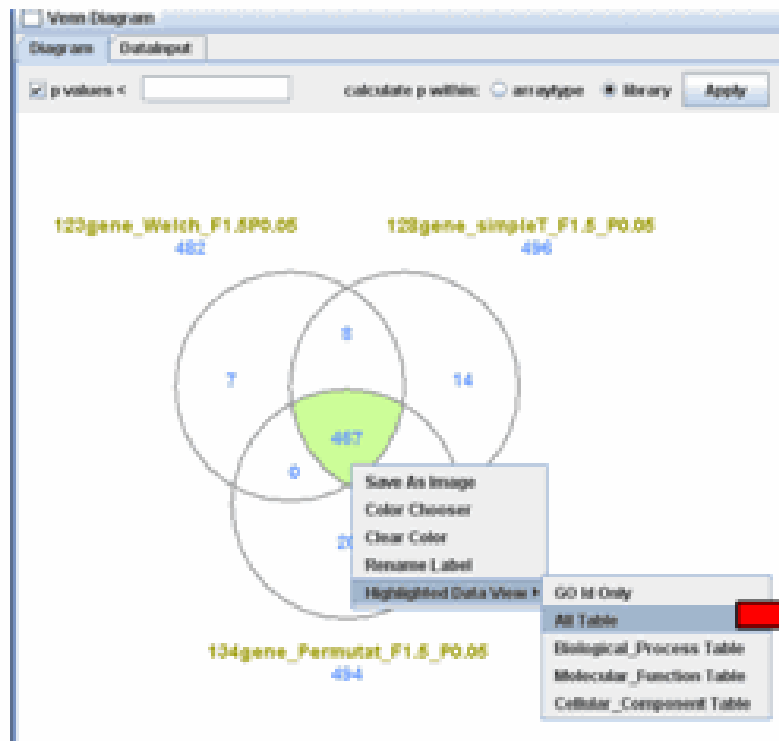
GENELIST...	EXPID	GENE BANK...	GENENAME	LOCUSID	UNIGENEID	SWISSPRO... D
128gene_s...	650	AI176658	Hspb1	24471		
128gene_s...	650	AI229421	Mapkapk3_...	315994		
128gene_s...	650	AI229655	Ctdsp1_pr...	363249		
128gene_s...	650	AI230778	Tpst2_pred...	288719		

Set C: 134gene_Permutat_F1.5_P0.05 or

GENELIST...	EXPID	GENE BANK...	GENENAME	LOCUSID	UNIGENEID	SWISSPRO... D
134gene_...	650	AI171243	Rpa3_pred...	296883		
134gene_...	650	AI171630	Mapk14	81649		
134gene_...	650	AI172097	Hsf1	79245		
134gene_...	650	AI172097	Hsf1	79245		

Select column for each table, then

Right click the highlighted green area, then choose “Highlighted Data View”, selecting Common GO tables (all table, Biological Process, Molecular Function and Cellular Component)



Compare two array types

(1) Select the array types to be compared

(2) Click "VennDiagram" button and choose the ID type that the comparison is based on



(3) Original data is displayed back into chip library table.

	ARRAYTYPE	SPOTID	POS_DESCR_MFR	GEN_ID_MFR	GENEBAKACC	GENENAME
1	Affy_HU-133A	249529	21779	AFFX-c2-Hs180RNA-3_s	AFFX-c2-Hs180RNA	M11
2	Affy_HU-133A	249526	21784	AFFX-c2-Hs180RNA-5_s	AFFX-c2-Hs180RNA	M11
3	Affy_HU-133A	249511	21779	AFFX-c2-Hs180RNA-M_s	AFFX-c2-Hs180RNA	M11
4	Affy_HU-133A	249524	21783	AFFX-c2-Hs280RNA-3_s	AFFX-c2-Hs280RNA	M11
5	Affy_HU-133A	249517	21778	AFFX-c2-Hs280RNA-5_s	AFFX-c2-Hs280RNA	M11
6	Affy_HU-133A	249530	21789	AFFX-c2-Hs280RNA-M_s	AFFX-c2-Hs280RNA	M11