

# *Object Groups*

# *Object Groups*

- **Collect and save lists of genes, metabolites, pathways...**
- **Transform, filter, and analyze them**
- **Share groups with colleagues**
- **Use groups in conjunction with publications**
- **Claim: Groups enable biologists to perform computations that required a programmer**

# Accessing Groups

- **Desktop menu: Groups**
- **Web menu: Tools -> Groups**
  - To use Web Object Groups, you must create a (free) BioCyc account

# Creating Groups: User-Defined Groups

- **Create a group by uploading a tab-delimited text file**
  - Software will assist in recognizing database objects from text
- **Create a group by typing in names**
  - You will be presented with options for the correct object; e.g. typing “tryptophan”, you will get the option to choose the compound L-tryptophan, the enzyme tryptophan synthase, and others


The screenshot shows a software interface with three main sections: TRANSFORMS, ENRICHMENTS, and PROPERTIES. The TRANSFORMS section has a 'Column operators' dropdown set to 'Filter to objects not in other' and an 'Add' button. The ENRICHMENTS section has a dropdown set to 'Genes Enriched for GO (biol)' and an 'Enrich' button. The PROPERTIES section has a dropdown set to 'Accession-1' and an 'Add' button. Below these sections is a table with two columns. The first column is titled 'New column' and the second is 'Genes regulated by protein, RNA or compound'. The first row of the table has 'TrpR transcriptional repressor' in the first column and a list of genes (trpR, trpA, trpB, trpC, trpD, trpE, trpL, mtr, aroL, yaiA, aroM, aroH) in the second column. The second row has 'tryptophan' in a text input field in the first column and is empty in the second column. There is an 'Add row' button at the bottom left of the table.

# Creating Groups

Add current object to group

Add quick-search result set to group

Add object-specific search result to a group



Logged in as [pkarp@ai.sri.com](#) | [Logout](#) | [Help](#) | [My preferences](#)

Quick Search Gene Search  
[change organism database](#)

[Home](#) | [Search](#) | [Tools](#) | [Help](#)

### Escherichia coli K-12 substr. MG1655 Query Results

You searched for all compounds that have a molecular weight of at least 2000.  
Your query returned 47 results.

[Turn into a group](#)

Compound Name ▲ ▼	Chemical Formula ▲ ▼	Mol. Wt. ▲ ▼
<a href="#">(enterobacterial common antigen)x2 undecaprenyl-diphosphate</a>	H162P2N6O37C103	2138.378
<a href="#">(enterobacterial common antigen)x3 undecaprenyl-diphosphate</a>	H198P2N9O52C127	2744.938
<a href="#">(enterobacterial common antigen)x4 core oligosaccharide lipid A</a>	H447N14P4O160C272	6597.416
<a href="#">(enterobacterial common antigen)x4 undecaprenyl-diphosphate</a>	H234P2N12O67C151	3351.497
<a href="#">(heptosyl)<sub>2</sub>-(KDO)<sub>2</sub>-lipid A</a>	H223P2N2O51C124	2620.056
<a href="#">(KDO)<sub>2</sub>-(lauroyl)-lipid IV<sub>A</sub></a>	H170P2N2O38C96	2022.337



BioCyc™  
Database Collection

# *Using Groups: Group Transformations*

**Use Case: Find all genes that are regulated by a transcriptional regulator.**

**Step 1: Create a group that contains the transcription factor.**

**Step 2: In the “Transforms” menu, select “Genes regulated by protein, RNA or compound”.**

**A new column containing all genes regulated by the transcription factor is generated**

**Click “+” to turn that new column into a group**

**This group can be further transformed, e.g. into metabolic pathways**

# *Object Group Transformations*

**Transform metabolite group into group of metabolic pathways, then into gene group**

**Transform gene group into group of regulators of those genes**

**Transform gene group into list of TF binding sites controlling those genes; into list of sequences**

**Create group of nucleotide positions; transform to closest genes; paint to cellular overview or enrich for GO terms**

# Using Groups: Properties

**Use Case:** Find the chromosome positions of all genes encoding enzymes of a pathway, sort by chromosome position. Find regulators of those genes.

**Step 1:** Create a group that contains the pathway.

**Step 2:** In the “Transforms” menu, select “Genes of pathway”.

**Step 3:** Select gene column and select

- Groups > New > Group from column OR
- Click “+” at top of gene column

**Step 4:** Select “left-end-position” from “Add Property Column” menu and add. Sort by position by clicking on arrow in column header.

**Step 5:** Select “Direct regulators of gene” from “Add Transform Column” menu.



# Genes of Aspartate Superpathway

Column operators: TRANSFORMS: Filter to objects not in other [Add] ENRICHMENTS: Genes Enriched for GO (biol) [Enrich] PROPERTIES: Right-End-Position [Add]

Show paged Show all

	All-Genes	Left-End-Position	Right-End-Position
<input type="checkbox"/> 1	thrA	337	2799
<input type="checkbox"/> 2	thrB	2801	3733
<input type="checkbox"/> 3	thrC	3734	5020
<input type="checkbox"/> 4	dapB	28374	29195
<input type="checkbox"/> 5	nadC	117752	118645
<input type="checkbox"/> 6	dapD	185123	185947
<input type="checkbox"/> 7	nadD	669154	669795
<input type="checkbox"/> 8	nadA	781308	782351
<input type="checkbox"/> 9	aspC	983742	984932
<input type="checkbox"/> 10	malY	1698981	1700153
<input type="checkbox"/> 11	nadE	1820482	1821309
<input type="checkbox"/> 12	dapE	2589629	2590756
<input type="checkbox"/> 13	dapA	2596904	2597782
<input type="checkbox"/> 14	nadB	2708442	2710064
<input type="checkbox"/> 15	lysA	2975659	2976921
<input type="checkbox"/> 16	metK	3084728	3085882
<input type="checkbox"/> 17	metC	3150258	3151445
<input type="checkbox"/> 18	argD	3486982	3488202
<input type="checkbox"/> 19	asd	3571798	3572901
<input type="checkbox"/> 20	dapF	3992785	3993609
<input type="checkbox"/> 21	metE	4011076	4013337
<input type="checkbox"/> 22	metB	4126695	4127855
<input type="checkbox"/> 23	metL	4127858	4130290
<input type="checkbox"/> 24	metA	4212303	4213232
<input type="checkbox"/> 25	metH	4221851	4225534
<input type="checkbox"/> 26	lysC	4229907	4231256

[Add row](#)

# *Enrichment Analysis*



- **Example: Does a group of genes contain more genes involved in cell division than would be expected by chance?**
- **More generally: Does a group contain more entities from defined classes than expected by chance?**
- **Check gene groups for enrichment of**
  - GO terms, pathways, regulons
- **Check metabolite groups for enrichment of**
  - Pathways

# *Enrichment Analysis*

- **Enrichment test based on the Hypergeometric Distribution**
- **Can perform enrichment, depletion, or both**
- **Multiple testing correction optional**

# Enrichment Dialog

ENRICHMENTS

Genes Enriched for Pathway:  

T4					
2.1656876					
0.9973915					
0.5333049					
0.4431508					
0.3669181					
1.0450248					
0.6124025					
0.4506686					
1.3523909					
1.2810624					
2.5529437	0.88949484	1.1338782	1.0615696	0.5815946	0.88905144

### Enrichment parameters

Analysis type:

- Enrichment
- Depletion
- Enrichment and Depletion

Include results whose p-value less than:

Algorithm details:

- Fisher Exact
- Fisher Exact Parent-Child Union
- Fisher Exact Parent-Child Intersection
- None
- Bonferroni Correction
- Benjamini-Hochberg Correction
- Benjamini-Yekutieli Correction

Correction

# *Sharing Groups*

**By default your groups are private**

**Share them with specific colleagues or make them public**

- Groups -> Sharing

**Group URLs are permanent and can be used in publications**

# *Lab Exercises*

1. Search for compounds having 'glucose' in their name
2. Turn into group
3. Enrichment Analysis → Compounds Enriched for Pathways
4. Use defaults in Enrichment pop-up, click 'OK'
5. Select ten most enriched pathways
6. Groups → New → From checked rows
7. Reactions of pathway
8. New group of reactions (green button on column)
9. Groups → Paint Data → On cellular overview