Pathway Database System 2.0

Pathways Explorer User Manual
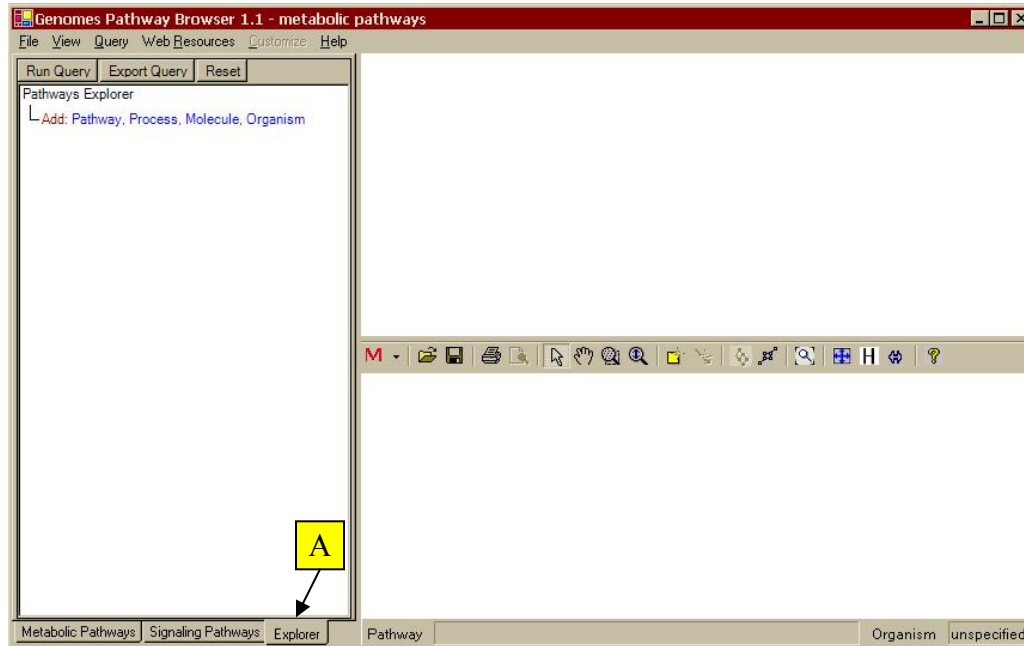
Center for Computational Genomics

Division of Computer Science
EECS Department
Case Western Reserve University
Cleveland, OH 44106

Department of Genetics
Case Western Reserve University
Cleveland, OH 44106

1. Overview

The Pathways Explorer tool helps you explore the relationships between pathways, processes, molecules and organisms.

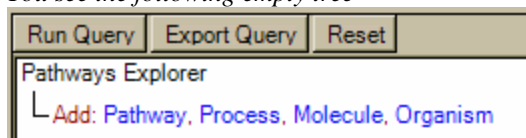## 1.1. Starting the Pathways Explorer

1. Open the Genome Pathways Browser
   The window above is shown
2. Click the *Explorer* tab (A)

2. Tutorial

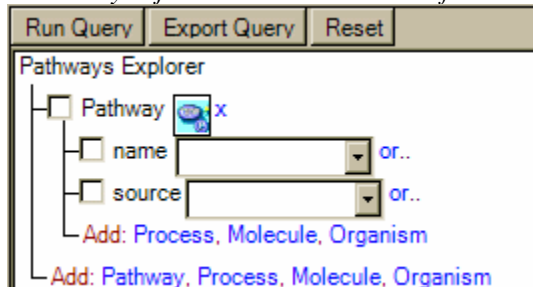This section will introduce you to the Pathways Explorer and what you can do with it.

2.1. Exploring Pathways, Processes Molecules and Organisms

1. Start the Pathways Explorer
   You see the following empty tree

   A. The *Run Query* and *Export Query* buttons perform output queries (discussed later)
   B. The *Reset* button brings you back to this empty tree
   C. *Pathways Explorer*, shown in the white area, is the root of the query tree
   D. The *Add* buttons, children of the root, are used to insert objects into the query tree
      i. Under the root may be added a Pathway, Process, Molecule or Organism
2. Click *Pathway*
   A Pathway object is inserted as a child of Pathways Explorer; its children are its attributes:
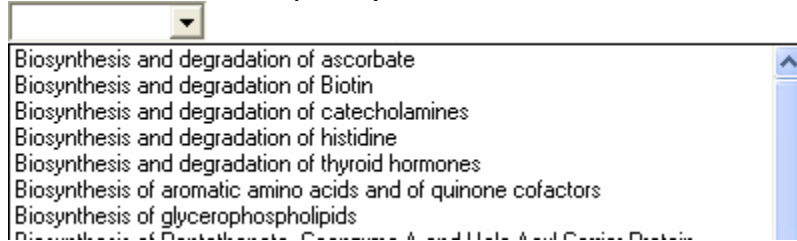

   A. Each attribute has a drop-down box in which a value may be specified

B.   Each attribute has a checkbox that determines which ones will be shown in output queries
C.   The Pathway has child *Add* buttons for inserting either a Process, Molecule or Organism object as children

D.   Next to the Pathway is a graph button for drawing a graph of the pathway
E.   Next to the Pathway is an "X" button for deleting the pathway object from the tree

3.   Open the *name* drop-down box by clicking the down arrow
A list is shown with all the pathway names in the database:

Biosynthesis and degradation of ascorbate
Biosynthesis and degradation of Biotin
Biosynthesis and degradation of catecholamines
Biosynthesis and degradation of histidine
Biosynthesis and degradation of thyroid hormones
Biosynthesis of aromatic amino acids and of quinone cofactors
Biosynthesis of glycerophospholipids

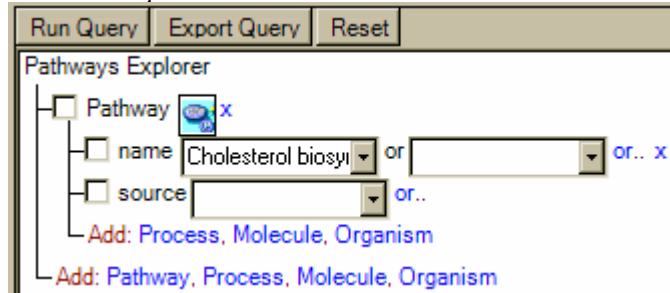4.   Click "Cholestorol biosynthesis"
This becomes the current value; this indicates that the Pathway object represents all pathways with this name

5.   Click the graph button
A graph of this pathway is drawn
6.   Click the *or* button next to the pathway name drop-down box
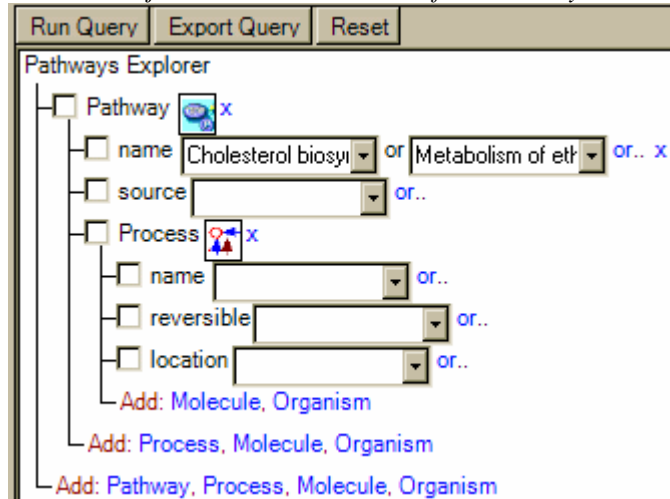*A second drop-down box is added:*

7.   Open the newly added drop-down box and choose "Metabolism of ether lipids"
This indicates that the Pathway object represents all pathways with either of the two names
8.   Click the *Process* button that is a child of the Pathway
A Process object is inserted as a child of the Pathway

A.   It behaves similarly to the pathway
B.   The Process has a child *Add* button for inserting a Molecule or Organism as its child
9.   Open the *name* drop-down box for the Process
A list is shown with all process names in the database that are in a pathway with either of the two pathway names

specified above:

A.   Suppose this list was too long; in fact, all drop-down boxes are limited to 300 items

B.   We want to find an item called "Acylation + dephosphorylation"

C.   Close the drop-down box

D.   Type "dephosphorylation" in the text box

E.   Open the drop-down box once again

It will contain only those values containing "dephosphorylation":

10.   Click "Acylation + dephosphorylation"
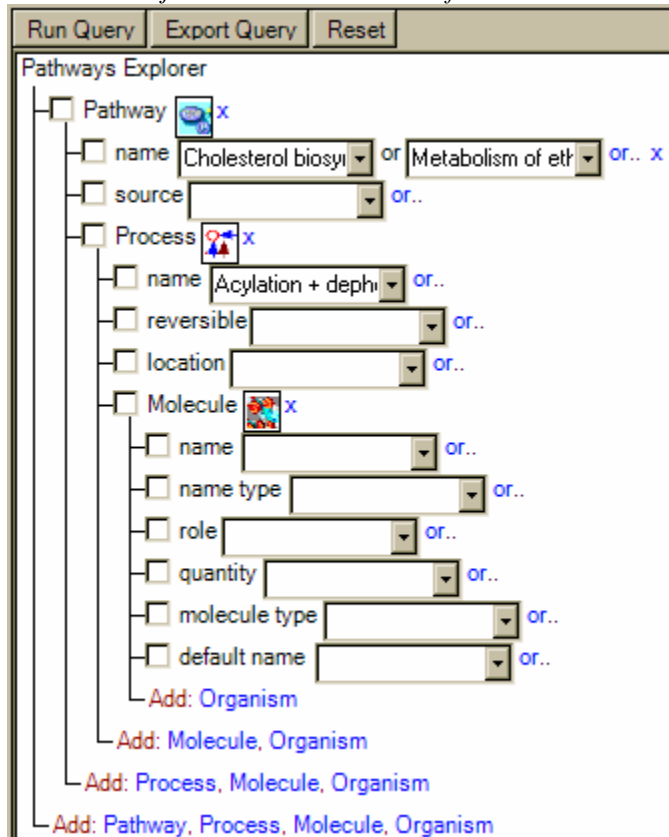
This becomes the current value; this indicates that the Process object represents all processes with this name such that it is a member of a pathway with either of the two specified pathway names
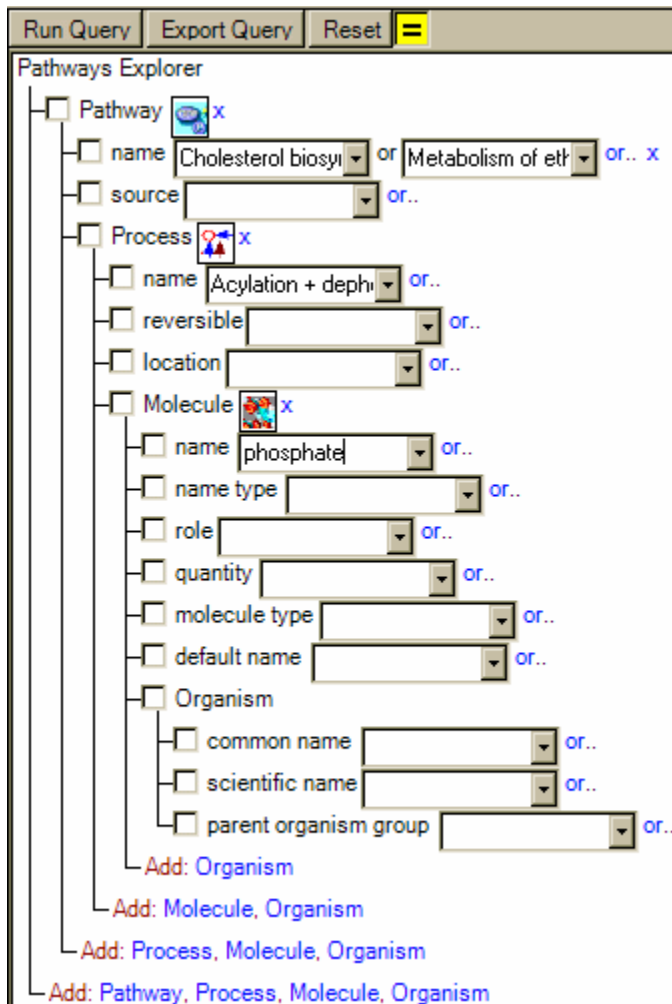
11.   Click the *Molecule* button that is a child of the Process
        A Molecule object is inserted as a child of the Process

Run Query   Export Query   Reset

Pathways Explorer

Pathway [⊡] x
    name [Cholesterol biosy ▼] or [Metabolism of eth ▼] or.. x
    source [　　　　▼] or..
    Process [⊡] x
        name [Acylation + deph ▼] or..
        reversible [　　　　▼] or..
        location [　　　　▼] or..
        Molecule [⊡] x
            name [　　　　▼] or..
            name type [　　　　▼] or..
            role [　　　　▼] or..
            quantity [　　　　▼] or..
            molecule type [　　　　▼] or..
            default name [　　　　▼] or..
            Add: Organism
        Add: Molecule, Organism
    Add: Process, Molecule, Organism
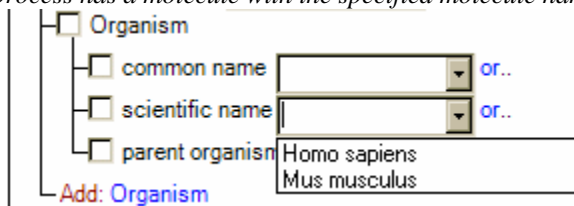Add: Pathway, Process, Molecule, Organism

A.   It behaves similarly to the pathway and process
B.   It has one child *Add* button for Organism
12.   Open the name drop-down box for the Molecule
        A list is shown with all molecule names in the database that are in a process with the specified process name such that this process is in a pathway with either of the two specified pathway names:

Molecule [⊡] x
    name [　　　　▼] or..
    name | Acyl coenzyme A
    role | Acyl-CoA
         | Alkylacylglycerol
    quanti| Alkylglycerol
         | CoA
    molec| CoA-SH
    defaul| Coenzyme A
         | coenzyme A (CoA-SH)
    Add: Org| H2O
         | phosphate
    Add: Molecu| Pi
    Add: Process, M| su
         | water
Id: Pathway, Process, Molecule, Organism

13.   Select *phosphate* from this list of molecules and click the *Organism* button that is a child of the Molecule
        A Molecule object is inserted as a child of the Molecule

  A. It behaves similarly to the pathway, process and molecule

  B. It has no child *Add* buttons

 14. Open the scientific name drop-down box for the Organism

  A list is shown with all the scientific names of organisms in the database that have a process with the specified process name such that this process is in a pathway with either of the two specified pathway names and this process has a molecule with the specified molecule name:
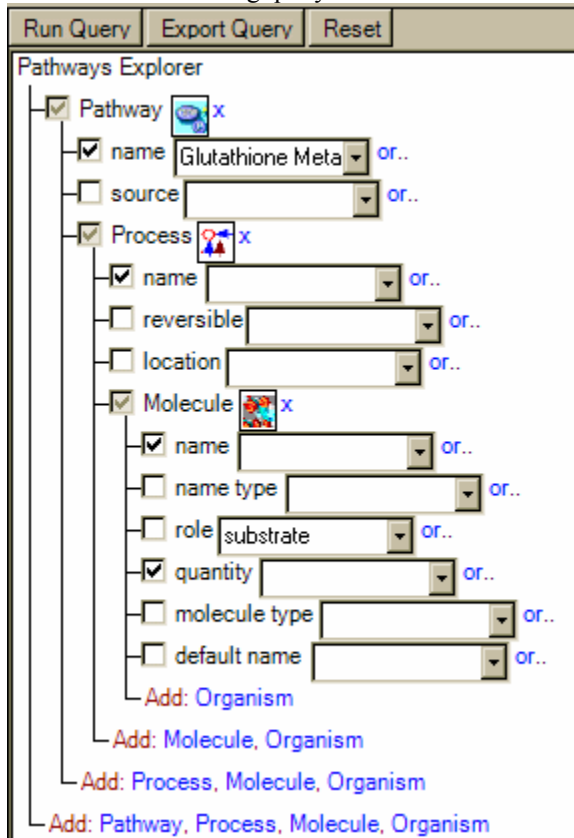

2.2. Query Execution

Suppose you want to list the default name and quantity of all the substrate molecules in the reversible processes of the "Glutathione Metabolism" pathway, along with the process and pathway names. We shall now perform this query.

 1. Click *Reset* to begin a new query tree

2.   Construct the following query tree:

   a.   Be sure to click the appropriate checkboxes
3.   Click the *Run Query* button

The following output table and graph is shown:

(45 results)

Pathway: name	Process: name	Molecule: name	Molecule: quantity
Glutathione Metabolism	5-Oxoprolinase(ATP hydolyzing)	5-oxoproline	1
Glutathione Metabolism	5-Oxoprolinase(ATP hydolyzing)	H2O	1
Glutathione Metabolism	5-Oxoprolinase(ATP hydolyzing)	su	1
Glutathione Metabolism	5-Oxoprolinase(ATP hydolyzing)	water	1
Glutathione Metabolism	Dipeptidase	H2O	1
Glutathione Metabolism	Dipeptidase	L-cysteinylglycine	1
Glutathione Metabolism	Dipeptidase	su	1
Glutathione Metabolism	Dipeptidase	water	1
Glutathione Metabolism	gamma-glutamylcyclotransferase	L-g-glutamylamino acid	1
Glutathione Metabolism	gamma-glutamylcyclotransferase	L-Glutamylamino acid	1
Glutathione Metabolism	gamma-glutamyltranspeptidase(gamma GT)	aa	1
Glutathione Metabolism	gamma-glutamyltranspeptidase(gamma GT)	Amino acid	1
Glutathione Metabolism	gamma-glutamyltranspeptidase(gamma GT)	Glutathione	1
Glutathione Metabolism	gamma-glutamyltranspeptidase(gamma GT)	GSH	1
Glutathione Metabolism	gamma-glutamyltranspeptidase(gamma GT)	Red. glutathione	1
Glutathione Metabolism	gamma-glutamyltranspeptidase(gamma GT)	Red. glutathione (GSH)	1
Glutathione Metabolism	Glutamate cysteine ligase	L-cysteine	1
Glutathione Metabolism	Glutamate cysteine ligase	L-glutamate	1
Glutathione Metabolism	Glutathione dehydrogenase(ascorbate)	Dehydro-ascorbate	1
Glutathione Metabolism	Glutathione dehydrogenase(ascorbate)	Glutathione	1
Glutathione Metabolism	Glutathione dehydrogenase(ascorbate)	GSH	1

a. Each column in the table is one of the attributes checked in the tree
b. Each process in the table is highlighted in the graph of the pathway
c. Graphs are only drawn when there is exactly one pathway in the table output

4. Now you will export the query output to a tab-delimited text file and into Microsoft Excel
   a. Click the Export Query button
      A file dialog appears
   b. Type a file name
   c. In the *Save as type* drop-down box, choose "Text files"
   d. Click *Save*
      A tab-delimited text file is created containing the query output
   e. Open Microsoft Excel
   f. From the *Data* menu, choose *Import External Data* and then *Import Data*
      A file dialog appears
   g. Locate the file you just exported and click *Open*
      The Text Import Wizard appears
   h. Click *Finish*
   i. Click *OK*
      The query output appears

5. You may also export the query output to an XML file
   a. Perform steps 4a-d but in step 4c, choose "XML files" from the drop-down box

2.3. Queries Using Joins

It is possible to add more than one pathway, process, or molecules to the query tree. When you do this, you can then join one object with another and indicate how they relate. For example, you can specify that two pathway names should be equal, or that one molecule quantity should be less than another molecule quantity. Possible operations include =, ≠, <, >, ≤, or ≥.

In this example, we will see how to find all pathways with a molecule that is a product such that this same molecule is also a substrate in a pathway named "Derivatives and degradation of phenylalanine and tyrosine."

1. Click *Reset* to begin a new query tree
2. Construct the following query tree:

a. It needs two Pathway objects as children of *Pathways Explorer*
b. Be sure to check the appropriate attributes
c. To assure that the two molecules are the same, you must assign a join
   i. Notice there is an = symbol in a yellow square next to the *Reset* button
   ii. You must drag the yellow square onto the word "Molecule" for both molecules
      1. Position the mouse over the square
      2. Hold down the left mouse button
      3. While holding, drag the mouse over the word "Molecule"
      4. Release the left mouse button
         (The word "Molecule" is highlighted in yellow)
      5. Repeat for the other molecule
   iii. You may not assign joins to different types of objects and attributes

3.   Open the second pathway name drop-down box
       It contains the pathways with a molecule that is a product such that this same molecule is also a substrate in a pathway named "Derivatives and degradation of phenylalanine and tyrosine"
4.   Close the drop-down box without choosing a value
5.   Click the *Run Query* button
       The output shows the molecule, the pathway for which the molecule is a substrate , and the pathway for which the molecule is a product

2.4. More Queries using Joins

In this example, we will see how to find processes that are reversible in one pathway and irreversible in another pathway
1.   Click *Reset* to begin a new query tree
2.   Construct the following tree:

   a.   It requires the use of two different colored squares
   b.   Set the blue square to ≠ by clicking the right mouse button in the square
          i.   Further right-clicking will cycle through all the operations
3.   Open either process name drop-down box (both contain the same values)
       It contains the processes that are reversible in one pathway and irreversible in another pathway
4.   Close the drop-down box without choosing a value
5.   Click the *Run Query* button
       The output shows the process, the pathway for which the process is reversible, and the pathway for which the process is irreversible

Quick Reference

A. Pathways Explorer tab – click this to begin
B. Graph output – shows graph output when:
    i. Run Query is clicked and its results contain a single pathway
        1. the processes in the query output are also highlighted
    ii. A graph button (K) is clicked
C. Query output
    iii. Each column is an attribute that was checked in the query tree
D. Pathways Explorer
E. Run Query button – perform an output query and show results at B and C
    i. Graph output is shown only when the results contain exactly one pathway
    ii. The processes in this one pathway that are in the results are highlighted in the graph
F. Export Query button – perform an output query and place results in:
    iv. A tab-delimited text file
    v. An XML file
G. Reset button – start over
H. Drop-down box
    vi. Contains all possible values of an attribute in the database
    vii. The list is limited to those values that are consistent with the rest of the drop-down box values in the tree
    viii. You may type a value in the textbox
    ix. If the list is too long but you know part of the attribute's value, you may type the partial value, then open the drop-down box and it will contain those values that match
I. Checkbox – each object (e.g. Pathway, Process, or Molecule) has a checkbox
    x. Clicking it will check or uncheck all its attributes
    xi. When some attributes are checked and others are not, the checkbox is gray
J. Checkbox – each attribute has a checkbox to indicate whether it should be shown in the query output (C)
K. "X" button – this deletes the object (e.g. Pathway, Process, or Molecule) and any of its child objects
L. Add buttons – click these to insert an object where the Add button is
    xii. For example, if an Add Process button is a child of a Pathway object, this button will insert a child Process of this Pathway

M.   Pathways Explorer query tree – this is the root of the tree, which contains all your query constraints

N.   "or" button – click this to add additional drop-down boxes; having additional attribute values means that the parent object (pathway, process, or molecule) represents any object with any of these values

References

[1] S. Newman, M. Ozsoyoglu. A Tree-Structured Query Interface for Querying Semi-Structured Data. *SSDBM, 2004*

[2] S. Newman. A Tree-Structured Query Interface for Querying Semi-Structured Data. *2004*