Groups Basics

Video Script

Associated with "EcoCyc_Groups-Basics_062813.mov"

This webinar will introduce you to basic operations on groups such as creating groups, groups structure, using filters, and performing set operations.

To start using groups, select the "**Groups**" command from the "**Tools**" menu. That command takes us to the groups directory which lists all the groups that are accessible to you. If you are not logged in you will only be able to see groups that are publicly shared.

Creating a Login ID:

To create and save groups of your own, and see groups you have saved in the past, you need to create and log in to a BioCyc account. If you don't have an account you can click on the "**create account**" link, and it will take you through the process of creating one. You can login to your account up here by clicking on the login link. Once you log in you will see the "**My Groups**" tab which lists all the groups you've created and are private to you.

Creating a Group:

For our first task, we will create a group of genes from experimental data, explore some of their properties, and take a look at several methods for finding pathways related to those genes.

There are three different ways to explore gene/pathway connections using groups. The first is by using transformations, the second is by visualizing the genes on a pathway map, and the third is by using enrichment analysis. We'll try all three over the course of this example.

First we will upload data from a text file that we've created. We'll take a look at the text file and this is simply a list of E. coli genes, one per line. To create a new group we go up to the groups menu which lists all the groups commands and we'll select "**New Group from Uploaded File**." When the dialog window appears, we navigate to our file, then select Open.

We'd like to system to analyze file and try to convert its contents into objects in the knowledge base… so we check, "**Try to make objects**." The dialog now asks what *type* of objects are in the file, so we'll select "**gene**."

The first line of our file is indeed a column header, so we'll check *this* box to maintain it as such. We now click "**Upload and create group**." : a process which can take a few moments based on the size of your file, and your Internet connection speed. The new group is now created, based on information from the currently selected organism's knowledge base.

Groups Structure, Adding Columns:

Now a few words about group structure. A group is very much like a spreadsheet. It has rows and columns. In this case we just have one column and 121 rows, which correspond to the genes from our uploaded file. Each cell of that column contains a gene object that points back into the knowledge base.

We can change the names of the group by clicking on them to give a better title.

Click here to add a description. We can also change the names of the columns by clicking on them. Clicking on a column header changes the actions available for that column, such as the ability to make a new group, or choosing a transformation property. A group might have several columns, but the 1st column will contain the chief objects of interest, in this case genes, but we can create other columns that will hold relationships or properties of the genes.

Now we'll add some columns from the properties menu here. Objects in a knowledge base are associated with a number of properties that can be added to a group. This menu shows all the properties that apply to a set of genes. Let's add a column for the accession number, and we can add a column for the Left-End-Position attribute of genes. We'll sort by the accession numbers to display the genes in that order by clicking the sort arrow on the column. We can also add the nucleotide sequence of the genes as another column via the properties menu.

We can sort by any of these columns. Let's sort by Left-End-Position.

Say we no longer are interested in gene guaC. To delete this gene from our group, select the box next to the row number, then select **Groups > Delete Checked Rows**. To add a row -- in this case, adding a gene, we'll look-up the gene and add it to the group.

Let's try that, using trpR as an example by using the quick search utility. We're taken to the results page for our trpR search. Clicking on the gene entry takes us to the trpR page, which contains all sorts of information about the gene. Hovering over the "**Add to Group**" button will reveal a larger button with a drop-down menu.

Clicking on the drop-down menu will let you add the gene to a new group, or to one of your existing groups. Note that the drop-down menu will only display your recently viewed groups, not all groups you may have.

Let’s add the gene to our example group, by selecting the group, then selecting **Add**. Click on the group name to return to the group, and the added gene will appear at the bottom of the list. To browse through group members, select **Groups > Browse this Group**. You’re now taken to the gene page of the first item in your group. Use the **Next** and **Back** links to cycle through the group. Click the group name to return to the group page.

Filtering:

Web Groups lets you organize and manipulate group entries using filters:
Filters are powerful functions that refine *existing* group entries based on criteria you select. Or, if you’d like to leave an existing group in tact, you can use the filter to create a *new group* based on its results.

Let’s now select the DNA sequence column. Filtering options depend on the type of data in the column. For example, a simple text filter used on the DNA sequence column will display only the genes that contain the substring entered. We’ll search for a substring of the first row’s DNA sequence. In this case, we’ll let the filter create a new group based on the results. We’ll now go back to our original group.

Set Operations:

Another way to manipulate groups is by using *Set Operations*. Set Operations are functions that display the *Union, Intersection,* or *Difference* between two groups.

For example, if you had a large group like this one and a small group containing a gene, and you wanted to see which genes from the large group also appeared in the small group – you could use the *Intersection* operation to display all objects in both groups. Let’s do that now. Then, we’ll return to our original group. If you had 2 groups you wanted to bring together, you could perform a *Union using the same menu*.