**Pathway Tools Navigator tutorial script**

March 16, 2012 version

Alexander Shearer

To the instructor:

 The Navigator tutorial is done via live demonstration on the desktop software and on the BioCyc site. You’ll want to have both up and available before you begin.

 Rather than strictly script your exact words, each section lists information you need to convey and actions to demonstrate.

 One way to slow down a bit and enhance clarity is to clearly announce what you are clicking on, and whether you are right or left clicking.

 It’s also helpful to the tutorial participants if you mouse over any item you’re discussing.

-- Pathway Tools Navigator tutorial script starts here --

Start on the Pathway Tools home frame (the one that shows all the databases you have access to)

*Welcome to Pathway Tools*

 Desktop and web versions very similar, with some differences

 Desktop layout has one big window, one small one

 Action happens in the big window

 Small window shows what the application is doing, gives access to command line

 Command menus at the top of the window

Switch over to the EcoCyc web site.

 Command menus arranged a little differently on the web

Switch back to the desktop version.

 In Pathway Tools, most things clickable

 Clickable objects show outline or tooltip when moused over

*Organism summary page.*

Click on EcoCyc.

 Organism summary page lists info like authors, taxonomy, summary of stats

 Remember that each line is clickable

 Click on the Pathways link

*Current organism.*

 “Current organism” selector chooses where your searches look

 The “current organism” can be different from the one in the frame you’re currently on

 Easy way to search into another organism without backing up to list of organisms

*Moving on to object and search types*

 There are some other universal features across multiple object types

 Will discuss them as we go through the individual objects

*Pathway searches*

 Can search Pathways by clicking on “Pathway” in command menu

 Two universal searches – by name or fame, and by substring

 Search by name is picky – wants full, exact name

 Most useful if you have a frame name only to work with

 Search by substring is more forgiving

Switch to web site

 Also have “quick search” on web site – looks for name across all object types

 Search “arginine”

 Quick search results grouped by object type

Desktop

 Search by substring for “arginine”

 This is a multiple results menu

 Can click one, many, or all results

 Select all

*Answer list, back, forward, history*

 The first answer appears in the main window, all others go on “answer list”

 Can click “next” for the next item on the answer list; it will appear in main window

 Can also use “back” and “forward” buttons to navigate, just like on a web browser

 Once you’ve been browsing, can look in history at recent pages

 History is a fast way to go back to an earlier page

 Use history to go back to arginine biosynthesis page

*What’s in a pathway page*

 We show pathway with major intermediates and all the reaction steps

 Can zoom in and out

 Zoom in – see side compounds, names of enzymes and associated genes

 Zoom in again – see structure for those side compounds

 Can also zoom all the way out to see the “conceptual” pathway

 Useful if you mainly care about the pathway as a whole, like for regulation

 Feedback regulation shown via lines

 The +/- shows what the regulating factor does

*Genetic regulation schematic*

 At the bottom of the page, genetic regulation schematic

 Shows all the genetic regulation that controls the genes for this pathway

 Genes grouped by operon

 This is a one-stop, at-a-glance view of regulation

 For complex features like this, note the question mark icon

 Click on the question mark to get a full explanation of that feature

*Comments and citations*

 Notice comments, with citations

 Whenever you see a citation, can click on it

 It’ll take you to the citation within the page’s list, on the desktop – then to Pubmed entry

 On web, straight to Pubmed

*Web Pathway search*

Website

 Show web search for Pathways

 Can search by ontology, number of reactions (i.e. length of pathway), etc

 Can combine search categories

 Show combined search using one ontology term and a number of reactions

 For ontology search/filter, note number – that shows how many entries are in that category in the PGDB

Desktop

*Reaction searches*

 Moving on to reactions

 Can search by name or substring, as before

 Search by E.C. number

*Reaction frame*

 Search for 2.6.1.11

 Reaction frames pretty sparse

 Shows reaction with clickable compounds

 Lists pathways the reaction is involved in, and enzymes carrying out the reaction

 Note that reactions *just* show the molecular transformation

 Details about how specific enzymes do the reaction live with those enzymes

 In other words, reaction frame shows the molecular change, but the enzyme will talk about things like Kms, Kcats, and so forth

Web site

*Web reaction search*

 Like the other searches, can combine search filters

 Can search for combinations of reactants and products

 Can filter for whether or not reaction is catalyzed by an enzyme

*Protein search*

 Again, search by name and substring

 Can also search by pathway, GO term, and a combination of weight and isoelectric point – last combo is handy with things like protein data from old papers and the results of 2D gels

*Protein frame*

 Go to ArgD

 Lots of data lives in protein frames

 When there’s a monomer or homomultimer, we merge gene and protein pages

 Have synonyms at the top, next to…

 Buttons for nucleotide sequence, protein sequence, and “advanced nucleotide sequence” that lets you get nearby sequence – good for PCR primers

 At top of page, have an overall summary of everything that impacts the protein’s function – genetic regulators, feedback regulation, and so forth – everything

 Unification and other links – ways to crosslink to relevant resources, such as PDB if the protein has a crystal structure

 Summary text – we like to put citations right after assertions

 Summary captures things we don’t have formal spaces for

 …and has the gene-reaction schematic

*Gene-reaction schematic*

 This shows the relationship between genes, proteins, and reactions

 Purple boxes are genes, circles are proteins, blue boxes are reactions

 Get a second set of boxes, with a number, for a multimer

 A fast way of seeing how genes, proteins, and reactions relate

 Notice how the ArgD gene-reaction schematic instantly tells you that ArgD carries out two reactions, and one of those is also carried out by AstC

 Below the gene-reaction schematic, we have GO terms and essentiality data

 Then…

*Enzymatic reactions*

 Enzymatic reactions capture how *this* specific enzyme carries out a *general* reaction

 It’s where we have Kms, cofactors, and regulators that are specific to this enzyme

 This is how we can capture different regulatory or other data for isozymes, for example

 Also where we have the evidence code

*Evidence codes*

 You’ll see these icons across the database

 Indicate how we know what we know

 A flask means experimental evidence, a computer means it was predicted, a book means “We think this is true, but can’t link it to other evidence”

 There are lots of types of evidence – for example, from a purified protein *in vitro* versus inferred from a mutant phenotype

*Gene searches, RNA searches*

 Can search for genes and RNAs by the usual methods

 Remember that gene searches will take you to combined gene/protein pages

 We currently represent RNAs that “work on their own,” like tRNAs and rRNAs

*Web searches for proteins/genes/RNAs*

Web site

 Extensive options for searching and filtering

 Note two new options on the web – search by publication, search by protein features

 Search by publication is good if you want to know what knowledge was captured from a publication you just came across in your research

 Search by protein feature lets you find all proteins that share certain features

 Expand search by protein feature

 Expand modified-residues

 Select acetylation-modifications

 Submit query

 For example, here are all the proteins – 152 of them – in EcoCyc which are acetylated

*Multiple results on the web*

 When you get search results like this, they appear in a sortable table

 Can click the arrows in each column to sort by that column

 Also, see the “Turn into a group” button on the upper left

 If you’re logged in, you can turn these results into a group and save that list for later

 Can do a lot with groups – more on that in the groups lesson

Desktop

*Genetic information*

 Search genes for argC

 Click on argCBG TU

 This is a transcription unit – most of the time, that’s the same as an operon

 Shows the genes in their local context, with transcription factors

 Promoters, transcription factor bindings sites, evidence for regulatory interactions

*Compound search*

 Can search by name, substring, class, and SMILES code

 SMILES is a text code for an organic molecule – can be handy for looking for “features” like an aromatic ring

*Compounds*

 Compound frames collect lots of info from elsewhere

 Show the compound names, structure, molecular formula

 Collects links to literally everywhere the compound appears in the database, whether that’s in reactions or as regulators or cofactors

 Really shows the strength of the PGDB – all the editor entered here was the compound name and an imported structure; the links come from other curation in the PGDB

Website

*Compound web search*

 Basically similar search here

 Note that monoisotopic molecular weight is useful if you’re doing mass spec

 InChi strings, like SMILES, are text codes for organic molecule structures

Can go to lab exercises from this point