

## Protein Family Analysis: Comparative Pathway Viewer

1. To look for similarities and differences in a particular pathway across different genomes in PATRIC, you can use the Comparative Pathway Viewer. Go to the Tools tab across the top of any PATRIC page and click on it. This will open up a list of tools (A below). Click on the Protein Family Sorter. This will take you to the landing page for that tool. If you are logged in, you will see the groups you have created in the box under “1. Select organism(s).”

The screenshot shows the PATRIC website interface. At the top, there is a navigation bar with 'ORGANISMS', 'DATA', 'SERVICES', 'TOOLS', and 'ABOUT'. Below this, a 'TOOLS' sidebar (labeled 'A') lists various tools, with 'Protein Family Sorter' highlighted. The main content area (labeled 'B') is titled 'Select organism(s)' and features a search box, a 'My Groups' tab, and a list of search results under 'cellular organisms'. A red arrow points from the 'Protein Family Sorter' tool in the sidebar to the search results. The search results list several groups, including '40 genomes Brucella', '2 Ochro 40 Brucella', and '42 complete genomes-Mycobacterium'.

2. The groups that you have created should be visible. You can select one, or many to compare. If you click on the + sign (Blue arrow 1) in front of any group, it will open up the list of all the genomes contained in that group and you can select, or deselect individual ones.

This screenshot illustrates the expansion of a search result. On the left, the search results are filtered to 'selected Workspace genomes'. A group labeled '40 genomes Brucella' is highlighted, and a blue arrow labeled '1' points to a plus sign icon next to it. A red arrow points from this group to the right, where the expanded view is shown. The expanded view displays a list of 40 individual Brucella strains, each with a checked checkbox, indicating they are selected for comparison. The strains listed include Brucella melitensis bv. 3 str. Ether, Brucella neotomae 5K33, Brucella canis ATCC 23365, Brucella melitensis bv. 2 str. 63/9, Brucella sp. NF 2653, Brucella suis ATCC 23445, Brucella ceti M13/05/1, Brucella suis bv. 5 str. 513, Brucella abortus bv. 3 str. Tulya, Brucella abortus str. 2308 A, and Brucella abortus bv. 2 str. 86/8/59.

3. Once you have made your selection, you can see a list of all the KEGG pathways from the genomes in your selection that have genes in that pathway. Just click on the Search button (Red arrow 1) and the pathway table will load.

**Select organism(s)**

My Groups | Taxonomy Tree | A-Z List

Jump to:

**Search within:**  
selected Workspace genomes

40 genomes Brucella

**Enter keyword**

Keyword:

**Examples**

Keyword: Metabolism  
Pathway ID: 00010  
EC Number: 1.1.1.1  
Annotation: PATRIC

**1** Search

Modify Search Criteria

To learn how to filter, sort, manipulate, refine, and save data within PATRIC feature tables, please see [Feature Table User Guide](#). Click on a pathway name to view a page.

Pathways | EC Numbers | Genes

**135 unique pathway(s) found**

Workspace | View | Download | Tools | Columns | Help

Add Feature(s) | FASTA DNA | Table | Pathway Summary | Multiple Seq Alignment | Show/Hide | Default | User Guides

FASTA Protein | FASTA | MAP IDs to...

Pathway ID	Pathway Name	Pathway Class	Annotation	Unique Genome	Unique Gene Cx	Unique EC Cour	EC Conservator	Gene Conservat
00053	Ascorbate and aldarate metab...	Carbohydrate Metabolism	PATRIC	40	990	9	98	2
00190	Oxidative phosphorylation	Energy Metabolism	PATRIC	40	1614	8	100	5
00051	Fructose and mannose metab...	Carbohydrate Metabolism	PATRIC	40	735	16	77	1
01056	Biosynthesis of type II polyket...	Biosynthesis of Polyketides an...	PATRIC	40	152	2	100	1
00052	Galactose metabolism	Carbohydrate Metabolism	PATRIC	40	548	10	100	1
01057	Biosynthesis of type II polyket...	Biosynthesis of Polyketides an...	PATRIC	40	419	5	99	2
00195	Photosynthesis	Energy Metabolism	PATRIC	40	401	2	100	5
00253	Tetracycline biosynthesis	Biosynthesis of Secondary Me...	PATRIC	40	272	4	99	1

4. You can also use this page to narrow down your search by entering a keyword for a particular pathway in the text box (Red arrow 1) and then click on the Search button (Red arrow 2). In this example, I used the key word “Benzoate” and this narrowed the returns to the “Benzoate Degradation via Hydroxylation” pathway.

**Select organism(s)**

My Groups | Taxonomy Tree | A-Z List

Jump to:

**Search within:**  
selected Workspace genomes

40 genomes Brucella  
2 Ochro 40 Brucella

**Enter keyword**

Keyword:

**Examples**

Keyword: Metabolism  
Pathway ID: 00010  
EC Number: 1.1.1.1  
Annotation: PATRIC

**1** **2** Search

Modify Search Criteria

To learn how to filter, sort, manipulate, refine, and save data within PATRIC feature tables, please see [Feature Table User Guide](#). Click on a pathway name to view a page.

Pathways | EC Numbers | Genes

**1 unique pathway(s) found**

Workspace | View | Download | Tools | Columns | Help

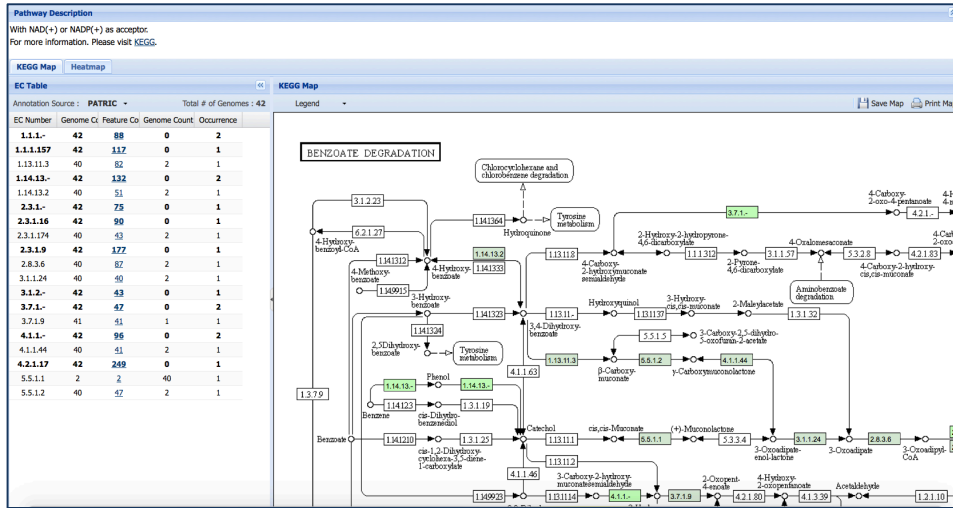
Add Feature(s) | FASTA DNA | Table | Pathway Summary | Multiple Seq Alignment | Show/Hide | Default | User Guides

FASTA Protein | FASTA | MAP IDs to...

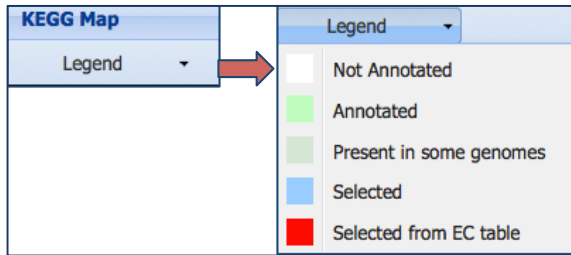
Pathway ID	Pathway Name	Pathway Class	Annotation	Unique Genome	Unique Gene Cx	Unique EC Cour	EC Conservator	Gene Conservat
00362	Benzoate degradation via hyd...	Xenobiotics Biodegradation an...	PATRIC	42	1415	19	93	1

5. **KEGG Map.** This will open up the pathway view. On the right you will see a table that lists the EC numbers and their presence and absence across the genomes from your selection. Bold numbers indicate that all of the genomes in the selection have that particular gene annotated, and non bold indicate that only some of the genomes

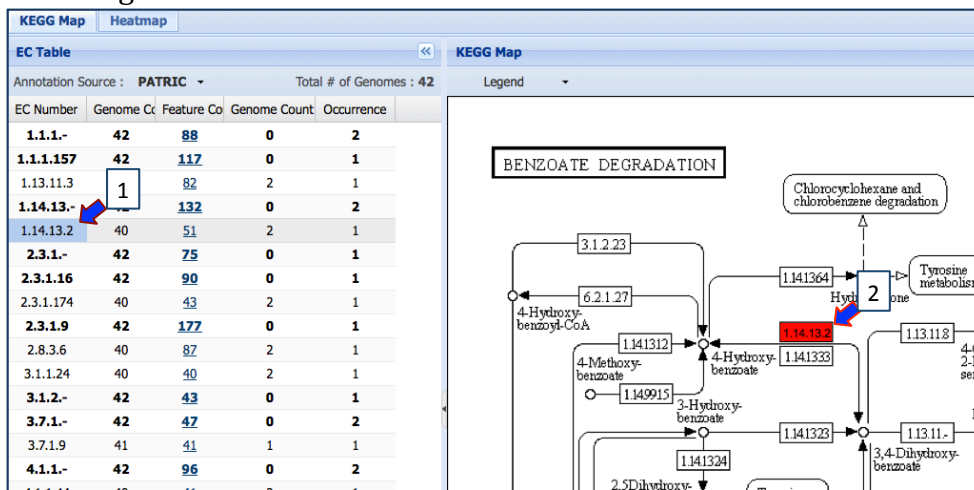
have a gene with that particular EC number annotated. On the left you can see the KEGG pathway map with the different EC numbers colored, or not.



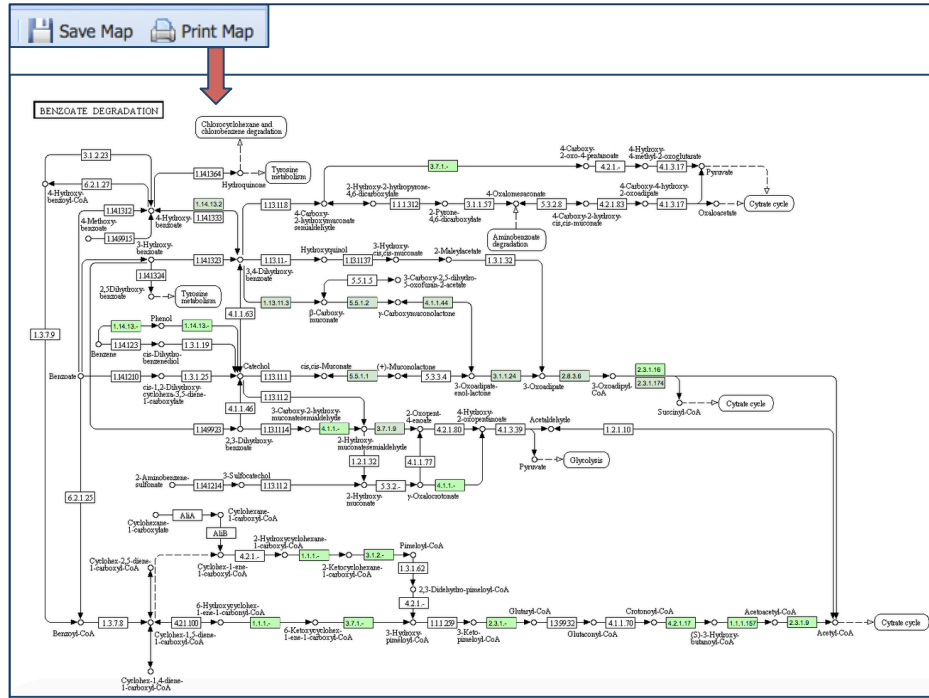
6. Clicking on the legend opens it. When you are viewing multiple genomes, green boxes indicate that ALL genomes have at least one gene annotated with that EC number. The olive color boxes indicate that some, but not all genomes have that EC number annotated, and white boxes indicate that none of the genomes have a gene annotated with that EC number.



7. Clicking on any EC number on the left will highlight its position in the KEGG map on the right.



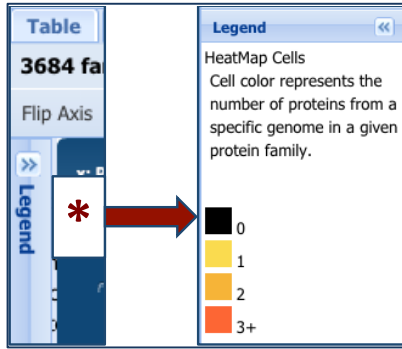
8. Clicking on the Print Map button at the top of the page will open a window where you can see the entire map.



9. **Heatmap.** Clicking on the heatmap tab will show you the presence or absence of the genes involved in the pathway across all the genomes.



10. You can click on the legend (Red asterisk) to see what the different colors of the cells mean, but basically, black cells indicate a gene absence and all other colors indicate gene presence.



11. You can arrange each of the rows or columns to arrange the data any way you like. To change the order of the rows, click and hold on the head of the row and move it up or down. You can do this with several rows to get the data you are interested in next to each other (Panel B below).



12. **Selection from Heatmap.** To see what proteins are missing, you will need to select the proteins in a neighboring column by drawing a box around them using your mouse (Panel A below, green rectangle). This will open a pop-up window that allows you to download the heatmap data, download the proteins from your selection, show the proteins from your selection, add the proteins to a group, or cancel (Panel B). To see the proteins, click on “Show Proteins” (Arrow 1). This will open up a new page that shows you all the proteins that were in the row you selected in the heatmap (Panel B). You can see that these proteins are all sequential (rectangle in Panel C).

**A**

Brucella suis 1330  
Brucella suis ATCC 23445  
Brucella suis bv. 3 str. 686  
Brucella suis bv. 4 str. 1

**B**

Selected Area from Heatmap  
Genomes selected: 1  
EC Numbers selected: 7  
Members: 9

Download Heatmap Data • Download Proteins • Show Proteins • Add Proteins To Group • Cancel

**C**

Workspace  
Add Feature(s)

View  
FASTA DNA  
FASTA Protein

Download  
Table  
FASTA

Pathway  
MAP IDs to

Genome Name	Accession	PATRIC ID	RefSeq Locus
Brucella suis 1330	NC_004311	fig 204722.5.peq.2849	BRA0636
Brucella suis 1330	NC_004311	fig 204722.5.peq.2850	BRA0637
Brucella suis 1330	NC_004311	fig 204722.5.peq.2851	BRA0638
Brucella suis 1330	NC_004311	fig 204722.5.peq.2854	BRA0641
Brucella suis 1330	NC_004311	fig 204722.5.peq.2856	BRA0643
Brucella suis 1330	NC_004311	fig 204722.5.peq.2857	BRA0644
Brucella suis 1330	NC_004311	fig 204722.5.peq.2858	BRA0645
Brucella suis 1330	NC_004311	fig 204722.5.peq.2859	BRA0646
Brucella suis 1330	NC_004311	fig 204722.5.peq.2860	BRA0647

13. You can also click on the heading to a single row (Arrow 1 below) to examine all the information about all the proteins in that row. Clicking on that header will open a pop-up window that allows you to download the heatmap data, download the proteins from your selection, show the proteins from your selection, add the proteins to a group, or cancel. To see the proteins, click on “Show Proteins” (Arrow 2 below).

Annotation Source: PATRIC • Flip Axis Total # of Genomes: 42

2.8.1.174 - 3-oxoadipyl-CoA thioester

1.1.1. - With NAD(+) or ...  
1.1.1.157 - 3-Hydroxybuty...  
1.13.11.3 - Proton-synthet...  
1.14.13.2 - 4-Hydroxybuty...  
2.8.3.6 - 3-oxoadipyl-Co...  
3.1.1.24 - 3-oxoadipate s...

**1**

**2**

Selected Area from Heatmap  
Genomes selected: 42  
EC Numbers selected: 1  
Members: 43

Download Heatmap Data • Download Proteins • Show Proteins • Add Proteins To Group • Cancel

14. This will open up a new page that shows you all the proteins that were in the row you selected in the heatmap.

Pathway Table

**43 features found**

Feature tables contain all of the identified features for all of the genomes in a particular genus. Tables may be refined to show subsets of features via various user o

Workspace: Add Feature(s)

View: FASTA DNA, FASTA Protein

Download: Table, FASTA

Tools: Pathway Summary, Multiple Seq Alignment, MAP IDs to...

Columns: Show/Hide, Default

Genome Name	Accession	PATRIC ID	RefSeq Locus Ta	Alt Locus Tag	Gene Symbol	Genome Browser	Annotation	Feature Type
Ochrobactrum inte...	NZ_ACOA01...	fig 641118.3.peg.3367	QINT_20008...	VBI0chInt64...	pcaF			
Ochrobactrum ant...	NC_009668	fig 439375.7.peg.4196	Qant_4020	VBI0chAnt7...				
Ochrobactrum ant...	NC_009668	fig 439375.7.peg.3882	Qant_3718	VBI0chAnt7...				
Brucella suis bv. S...	NZ_DS999712	fig 520489.3.peg.391	Bsulb55_010...	VBIBruSui73...				
Brucella sp. F5/99	NZ_GG663472	fig 437701.3.peg.340	BruF5_0101...	VBIBruSp659...				
Brucella sp. 83/13	NZ_DS999659	fig 520449.3.peg.2258	Bru83_0101...	VBIBruSp753...				
Brucella sp. NF 2652	NZ_ADFB010...	fig 693748.4.peg.279	BROD_0276	VBIBruSp153...	pcaF			
Brucella suis 1330	NC_004311	fig 204722.5.peg.2849	BRA0636	VBIBruSui10...	pcaF			
Brucella pinnipedia...	NZ_EQ999534	fig 520462.3.peg.1152	BpinM2_010...	VBIBruPin74...				
Brucella ovis ATCC...	NC_009504	fig 444178.3.peg.681	BOV_A0598	VBIBruOvi13...	pcaF			
Brucella microti CC...	NC_013118	fig 568815.3.peg.617	BMI_II633	VBIBruMic92...	pcaF			
Brucella melitensis...	NZ_EQ999588	fig 520466.3.peg.1188	Bmelb3E_01...	VBIBruMel11...				
Brucella melitensis...	NZ_ACEM01...	fig 520465.3.peg.683	Bmelb2E_01...	VBIBruMel34...				
Brucella melitensis...	NZ_EQ999555	fig 520464.3.peg.536	Bmelb1R_01...	VBIBruMel10...				
Brucella melitensis...	NC_003318	fig 224914.11.peg.2970	BMEI10646	VBIBruMe92...				
Brucella melitensis...	NC_012442	fig 546272.3.peg.2895	BMEA_B0609	VBIBruMel14...	pcaF			
Brucella sp. BO2	NZ_ADFB010...	fig 693750.4.peg.574	BIBO2_0515	VBIBruSp146...	pcaF			
Brucella inopinata...	NZ_ADEZ010...	fig 470735.4.peg.2558	BIBO1_2424	VBIBruSp109...	pcaF			
Brucella ceti M490...	NZ_EQ999599	fig 520458.3.peg.231	BcetM4_010...	VBIBruCet48...				
Brucella ceti str. C...	NZ_ACJD010...	fig 595497.3.peg.3102	BCETI_7000...	VBIBruCet28...	pcaF			

Page 1 of 3 | Show 20 per page | Apply | Apply to ALL tables

15. You will need to resize the table to see all the entries. To do this, change the default (20) at the bottom of the table to the total number.



16. **Multiple Sequence Alignment.** You can do a number of things with these genes. You could save them to a group, get their nucleotide or amino acid sequences, download the information, see if they were important in any metabolic pathways, or generate a multiple sequence alignment for them. To generate an alignment, click on the box in front of “Genome Name” (Arrow 1 below). This will select all the genes in the table. Then you need to click on the Multiple Sequence Alignment button at the top of the table (Arrow 2).

Workspace: Add Feature(s)

View: FASTA DNA, FASTA Protein

Download: Table, FASTA

Tools: Pathway Summary, Multiple Seq Alignment, MAP IDs to...

1

Genome Name	Accession	PATRIC ID	RefSeq Locus Ta	Alt Locus Tag	Gene Symbol
<input checked="" type="checkbox"/> Brucella ceti str. C...	NZ_ACJD010...	fig 595497.3.peg.3102	BCETI_7000...	VBIBruCet28...	pcaF
<input checked="" type="checkbox"/> Brucella microti CC...	NC_013118	fig 568815.3.peg.617	BMI_II633	VBIBruMic92...	pcaF
<input checked="" type="checkbox"/> Brucella abortus N...	NZ_GG703762	fig 575591.3.peg.713	BAUG_0627	VBIBruAbo9...	
<input checked="" type="checkbox"/> Brucella melitensis...	NC_007624	fig 359391.4.peg.2899	BAB2_0606	VBIBruMel86...	pcaF
<input checked="" type="checkbox"/> Brucella ovis ATCC...	NC_009504	fig 444178.3.peg.681	BOV_A0598	VBIBruOvi13...	pcaF
<input checked="" type="checkbox"/> Brucella canis ATC...	NC_010104	fig 483179.4.peg.2852	BCAN_B0636	VBIBruCan2...	pcaF
<input checked="" type="checkbox"/> Brucella inopinata ...	NZ_ADEZ010...	fig 470735.4.peg.2558	BIBO1_2424	VBIBruSp109...	pcaF
<input checked="" type="checkbox"/> Brucella suis 1330	NC_004311	fig 204722.5.peg.2849	BRA0636	VBIBruSui10...	pcaF
<input checked="" type="checkbox"/> Brucella melitensis...	NC_012442	fig 546272.3.peg.2895	BMEA_B0609	VBIBruMel14...	pcaF
<input checked="" type="checkbox"/> Brucella pinnipedia...	NZ_DS999750	fig 520463.3.peg.657	BAGG_00462	VBIBruPin12...	
<input checked="" type="checkbox"/> Brucella melitensis...	NZ_EQ999555	fig 520464.3.peg.536	Bmelb1R_01...	VBIBruMel10...	
<input checked="" type="checkbox"/> Brucella abortus S19	NC_010740	fig 430066.4.peg.623	BAbS19_II05...	VBIBruAbo3...	
<input checked="" type="checkbox"/> Ochrobactrum ant...	NC_009668	fig 439375.7.peg.3882	Qant_3718	VBI0chAnt7...	

2



