

Find Functions: COG

COG Browser

From the **Find Functions** top-level menu, the **COG** option on the second-level menu leads to the COG Browser, as shown in Figure 1(i). This is a two-level display with COG categories shown on the first level, and COG pathways listed under the corresponding categories.

Top-level COG categories include:

- Amino acid transport and metabolism [E]
- Carbohydrate transport and metabolism [G]
- Cell cycle control, cell division, chromosome partitioning [D]
- Cell Motility [N]
- Cell wall/membrane/envelope biogenesis [M]
- Chromatin structure and dynamics [B]
- Coenzyme transport and metabolism [H]
- Cytoskeleton [Z]
- Defense mechanisms [V]
- Energy production and conversion [C]
- Extracellular structures [W]
- Function unknown [S]
- General function prediction only [R]
- Inorganic ion transport and metabolism [P]
- Intracellular trafficking, secretion, and vesicular transport [U]
- Lipid transport and metabolism [I]
- Mobilome: prophages, transposons [X]
- Nuclear structure [Y]
- Nucleotide transport and metabolism [F]
- Posttranslational modification, protein turnover, chaperones [O]
- RNA processing and modification [A]
- Replication, recombination and repair [L]
- Secondary metabolites biosynthesis, transport and catabolism [Q]
- Signal transduction mechanisms [T]
- Transcription [K]
- Translation, ribosomal structure and biogenesis [J]

Clicking on any of the categories will lead to a page showing all COG IDs associated with this category and the corresponding isolate genome and metagenome counts (Figure 1(ii)). Clicking on any of the

pathways will lead to a page showing all COG IDs associated with this pathway and the corresponding isolate genome and metagenome counts (Figure 1(iii)).

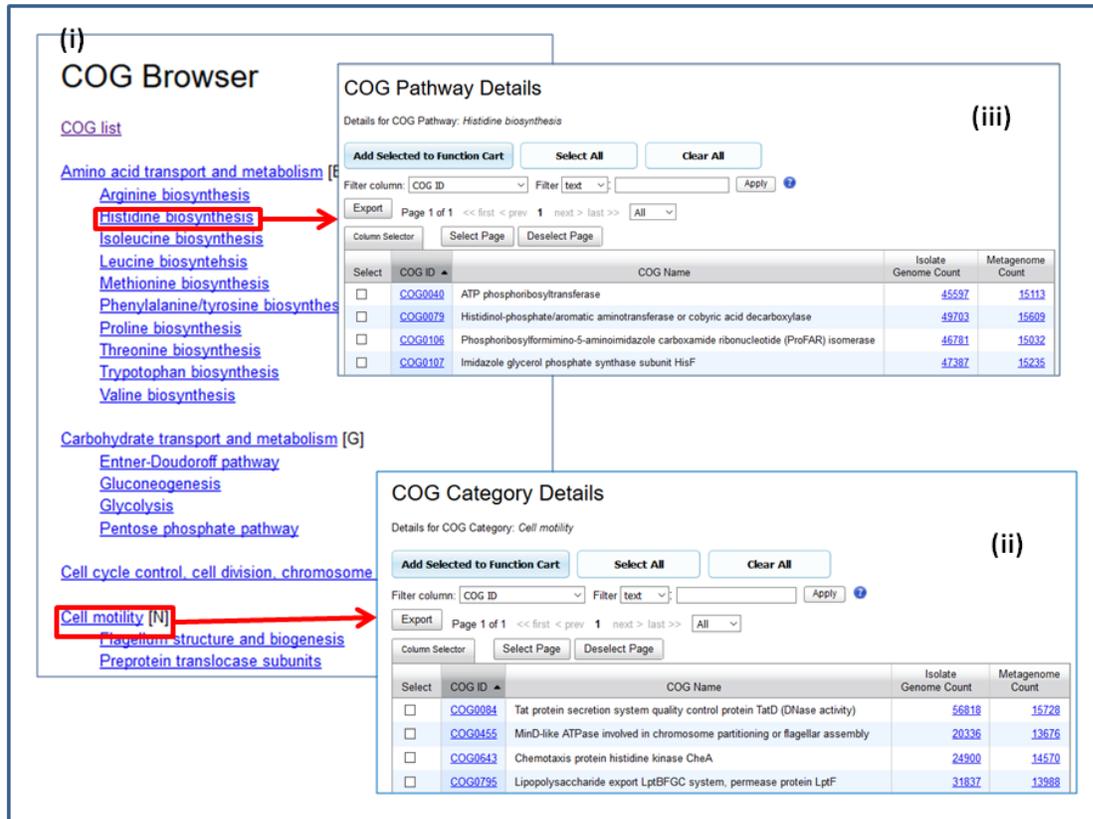


Figure 1. COG Browser.

Users can select any subset of listed COG IDs to be saved to Function Cart or Workspace Function Set for further analysis.

Users can also click on a genome count to view all isolate genomes or metagenomes having genes associated with a particular COG ID. For example, Figure 2(i) lists all isolate genomes having genes annotated with COG3352.

Clicking on a gene count will lead to a page listing the actual genes (Figure 2(iii)).

Clicking the **Phylogenetic Distribution** button in Figure 2(i) will lead to a page showing a phylogenetic tree: All genomes with genes annotated with this COG are displayed with a clickable link (see Figure 2(ii)).

(i)

Isolate Genomes with COG3352

Genomes with *Archaeum component FlaC*

Phylogenetic Distribution

Domains(D): * = Microbiome,
B = Bacteria, A = Archaea, E = Eukarya, P = Plasmids, G = GFragment, V = Viruses.
Genome Completion(C): F = Finished, P = Permanent Draft, D = Draft.

Add Selected to Genome Cart **Select All**

Filter column: Domain Filter: text

Export Page 1 of 1 << first < prev 1 next > last >> All

Column Selector Select Page Deselect Page

Select	Domain	Status	Genome Name	Count
<input type="checkbox"/>	A	F	Thermococcus litoralis DSM 5473	1
<input type="checkbox"/>	A	D	unclassified_Thermococcaceae_bin9	1
<input type="checkbox"/>	A	P	Thermococcus zilligii AN1	1

(ii)

Phylogenetic Distribution for COG3352

COG: *Archaeum component FlaC*
*Showing counts for all genomes
(Hits are shown in red)

Domains(D): * = Microbiome,
B = Bacteria, A = Archaea, E = Eukarya, P = Plasmids, G = GFragment, V = Viruses.
Genome Completion: [F]inished, [P]ermanent Draft, [D]raft.

A .01 Archaea (69)
A .02 Crenarchaeota (1)
A .03 Thermoprotei (1)
A .04 Sulfolobales (1)
A .05 unclassified (1)
A .06 unclassified (1)
A .08 [Sulfolobales sp. HSU1](#) [P] (1)
A .02 Euryarchaeota (68)
A .03 Archaeoglobi (5)
A .04 Archaeoglobales (4)
A .05 Archaeoglobaceae (4)
A .06 Archaeoglobus (3)
A .08 [Archaeoglobus from oil reservoir in Yangzhou - Bin 10](#) [D] (1)
A .08 [Archaeoglobus sp. JdFR-31](#) [P] (1)
A .08 [Archaeoglobus sp. JdFR-32](#) [P] (1)
A .06 Geoglobus (1)
A .08 [Geoglobus ahangari 234](#) [D] (1)
A .06 unclassified (1)

(iii)

Genes with COG3352

[Thermococcus sp. EP1](#)

Filter column: Gene ID Filter: text Apply

Export Page 1 of 1 << first < prev 1 next > last >> All

Column Selector Select Page Deselect Page

Select	Gene ID	Locus Tag	Gene Product Name	Genome Name
<input type="checkbox"/>	2647368001	Ga0100680_10616	flagellar protein FlaC	Thermococcus sp. EP1

Figure 2. COG3352 Genomes, Genes and Phylogenetic Distribution.

COG List

The **COG List** option shows all COG IDs in a table display (Figure 3(i)).

Users can select a subset of COG IDs to be added to Function Cart or Workspace for further analysis.

Clicking on a COG ID will lead to a **COG Detail** page as shown in Figure 3(ii).

Characterized COGs (i)

Buttons: Add Selected to Function Cart, Select All, Clear All

Filter column: COG ID Filter text Apply

Export Page 1 of 47 << first < prev 1 2 3

Column Selector Select Page Deselect Page

Select	COG ID	Function Name
<input type="checkbox"/>	COG0001	Glutamate-1-semialdehyde aminotransferase
<input type="checkbox"/>	COG0002	N-acetyl-gamma-glutamylph...
<input type="checkbox"/>	COG0003	Anion-transporting ATPase, A...
<input type="checkbox"/>	COG0004	Ammonia channel protein Ar...
<input type="checkbox"/>	COG0005	Purine nucleoside phosphory...
<input type="checkbox"/>	COG0006	Xaa-Pro aminopeptidase
<input type="checkbox"/>	COG0007	Uroporphyrinogen-III methyla...

COG Detail (ii)

COG0001: Glutamate-1-semialdehyde aminotransferase

DB Source	COG
Add Date	19-FEB-15
Seq Length	432
Function Code	H
Definition	Coenzyme transport and metabolism
Function Group	METABOLISM

hint: Click on domain to see function distribution of selected domain. Follow the link provided by Genome Count to genomes having genes associated with this function. Click on the number to view the results in phylo tree display, or click on the rectangular table symbol to view the results in table display. (Note: Phylo Tree display option is limited to certain domains only.)

Filter column: Domain Filter text Apply

Export Page 1 of 1 << first < prev 1 next > last >> All

Column Selector

Domain	Genome Count	Gene Count
*Microbiome	15435	14844982
Archaea	1203	1547
Bacteria	43541	60264
Eukaryota	103	144

Figure 3. COG List.

Clicking on a genome count in Figure 3(ii) will lead to a page showing all genomes in this domain having genes annotated with the COG ID.

Users can also click on a **Domain** in Figure 3(ii) to drill down to Phylum, Class, etc. all the way to the Species level.

COG List w/ Stats

This option is similar to COG List except that there are additional genome counts.

COG Id to Categories

This option shows the mapping of COG IDs to corresponding COG categories as shown in Figure 4.

COG ID Mapping to Categories

Filter column: Category Code Filter text Apply ?

Export Page 1 of 50 << first < prev **1** [2](#) [3](#) [4](#) [5](#) [6](#) [7](#) [8](#) [9](#) [10](#) [next >](#) [last >>](#) 100

Column Selector

COG ID	Category Code ▲	Category
COG0430	A	RNA processing and modification
COG1949	A	RNA processing and modification
COG4085	A	RNA processing and modification
COG5180	A	RNA processing and modification
COG5186	A	RNA processing and modification
COG5238	A	RNA processing and modification
COG5239	A	RNA processing and modification
COG0123	B	Chromatin structure and dynamics
COG1243	B	Chromatin structure and dynamics
COG2036	B	Chromatin structure and dynamics
COG5406	B	Chromatin structure and dynamics
COG5531	B	Chromatin structure and dynamics
COG5602	B	Chromatin structure and dynamics
COG0022	C	Energy production and conversion
COG0039	C	Energy production and conversion
COG0045	C	Energy production and conversion

Figure 4. COG ID Mapping to Categories.