

KEGG Database

Instructions









These descriptions and instructions will assist with navigating the using the functions of the KEGG database.

Please share any updates or suggestions by email to rmogul@cpp.edu.

What is the KEGG database?

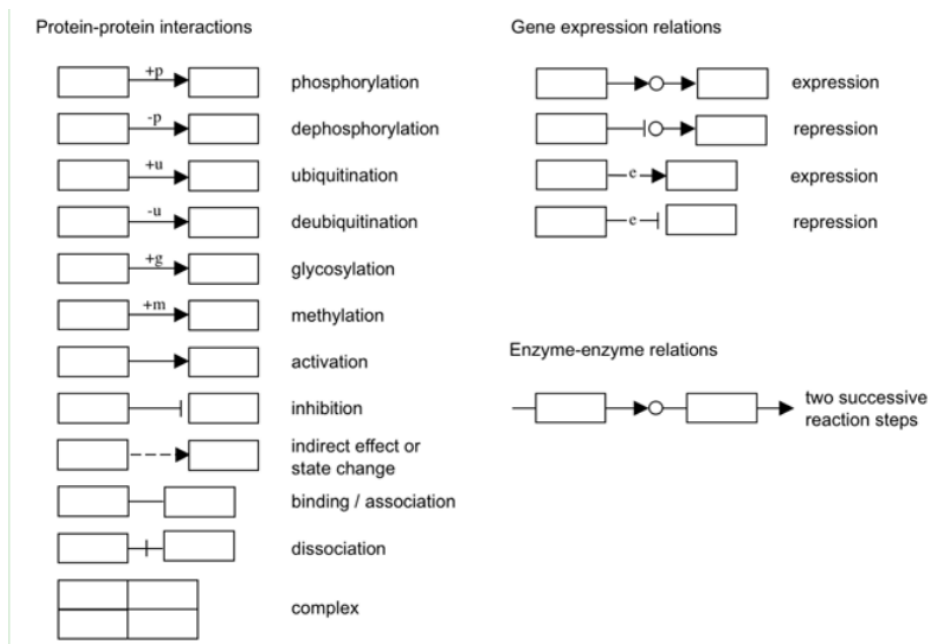
- The Kyoto Encyclopedia of Genes and Genomes (KEGG) database provides a way to study metabolism, signaling, and other biochemical mechanisms using protein, gene, and chemical information. The database *is essentially “a collection of manually drawn graphical diagrams, called KEGG pathway maps, representing molecular pathways for metabolism, genetic information processing, environmental information processing, cellular processes, organismal systems, human diseases, and drug development. Each pathway is identified by a five-digit number preceded by one of: map, ko, ec, rn, and three- or four-letter organism code.”*¹

Notation in the KEGG Database¹

Objects		Arrows	
	gene product, mostly protein but including RNA		molecular interaction or relation
	chemical compound, DNA and other molecule		link to/from another map
	map		indirect link or unknown reaction
			missing interaction (eg., by mutation)
			drug structure link or pointer used to add legend

¹ From https://www.genome.jp/kegg/document/help_pathway.html.

KEGG Database



Basic Functions in the KEGG Database

- **Boxes** are hyperlinked to lead to the specific KEGG page for that protein, which includes gene information, reaction information, related metabolic pathways, sequence data, structure images, references, and more.
- **Circles** are hyperlinked to lead to the KEGG page for the associated chemical/metabolite.

Color Coding in the KEGG Database¹

- The pathway map without coloring is the original version that is manually drawn by in-house software called KegSketch. The other pathway maps with coloring are all computationally generated as follows.
 - **Reference pathway:** this is the original version; *white boxes* are hyperlinked to KO, ENZYME, and REACTION entries in metabolic pathways; they are hyperlinked to KO entries in non-metabolic pathways.
 - **Reference pathway (KO):** *blue boxes* are hyperlinked to KO entries that are selected from the original version.
 - **Reference pathway (EC):** *blue boxes* are hyperlinked to ENZYME entries that are selected from the original version.

KEGG Database

- **Reference pathway (Reaction):** *blue boxes* are hyperlinked to REACTION entries that are selected from the original version.
- **Organism-specific pathway:** *green boxes* are hyperlinked to GENES entries by converting K numbers (KO identifiers) to gene identifiers in the reference pathway, indicating the presence of genes in the genome and also the completeness of the pathway.

Pathway Information for Specific Organisms

Option A:

- Using the “Reference Pathway” drop-down menu on the top left of the page, choose the organism of choice (*e.g.*, “Homo sapiens (humans)”), and click on “Go.”
- If the organism of choice is not on the list, go to Option B.

Option B:

- Click on the “Organism Menu” on the top menu bar.
- Find the organism of choice using the Search Instructions below.
- If the link for the organism of choice is inactive, go to Option C.

Option C:

- Go specifically to <http://www.kegg.jp>.
- On the main page, click on the “Organism” link: *Go mid-way down on the right-hand side of the page, and look under Classification.*
- Find the organism of choice using the Search Instructions below.

Search Instructions:

1. Search for the organism by using the search function.
2. Or, manually search by sorting through classifications.
3. For example, for *Falco cherrug*, go to Eukaryotes, Animals, Vertebrates, and Birds.