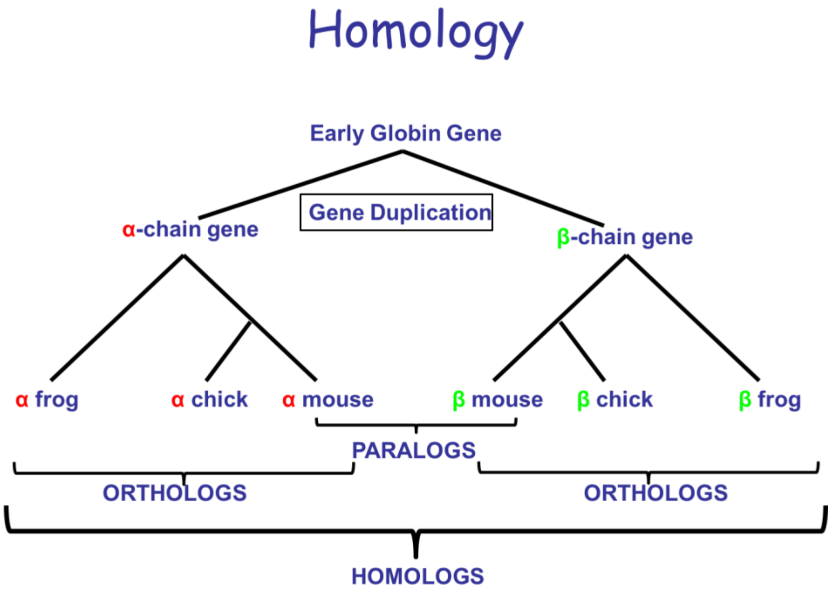
Orthology and Phyletic Patterns



Learning objectives:

- Explore the orthology table on VEuPathDB gene pages
- Getting to OrthoMCL from VEuPathDB gene pages
- Run searches in OrthoMCL
- Explore the cluster graphs in OrthoMCL
- Leverage the phyletic pattern search
- Leverage the orthology transform tool

OrthoMCL is a genome-scale algorithm for grouping orthologous protein sequences. Such orthologous sequences not only share evolutionary history, but also share function. Thus, ortholog prediction is important in predicting the function of newly identified proteins. Indeed, detection of orthologs has become more widespread with the rapid progress in genome sequencing and the discovery of protein sequences (Glover et al. 2019). Importantly, proteins in OrthoMCL groups have been shown to display a high degree of functional conservation (e.g., a group's proteins have consistent EC numbers) (Li et al. 2003), highlighting that OrthoMCL is useful for functional annotation of newly sequenced genomes.

OrthoMCL not only identifies groups shared by proteins from two or more species, but also groups representing species-specific gene expansion families. To achieve this, the OrthoMCL algorithm starts with reciprocal best BLAST hits within each proteome as potential in-paralog/recent paralog pairs and reciprocal best hits across any two proteomes as potential ortholog pairs. Related proteins are interlinked in a similarity graph. Then, MCL (Markov Clustering algorithm; Dongen 2000; www.micans.org/mcl) is invoked to split mega-clusters. This process is analogous to the manual review in COG construction. MCL clustering is based on weights between each pair of proteins. Thus, to account for differences in evolutionary distance between any two organisms, the weights are normalized before running MCL.

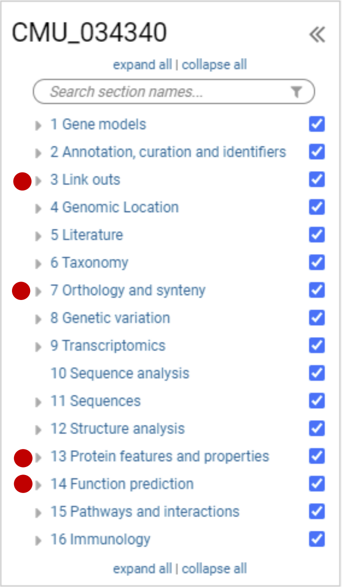
The organism specific orthology information garnered from our OrthoMCL analysis of VEuPathDB organisms is presented on gene pages and integrated into an Orthology Phylogenetic Profile search. The OrthoMCL.org site offers a deep look into all data associated with the OrthoMCL results for orthology groups and proteins.

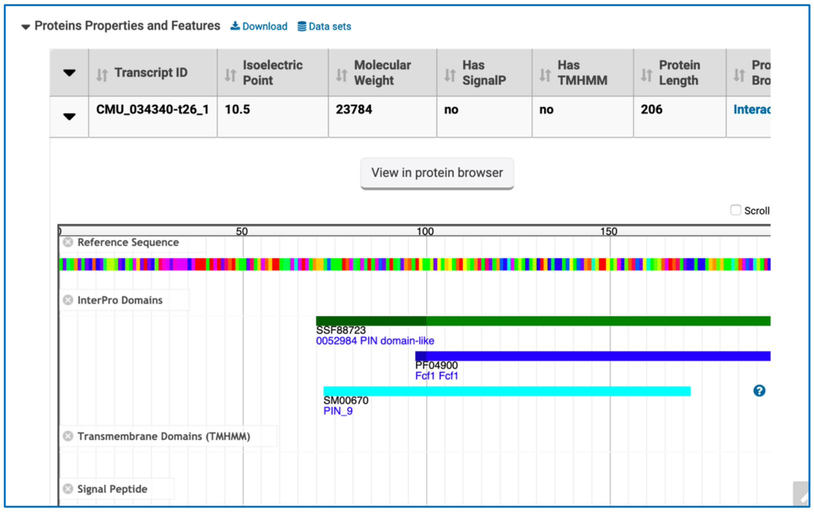
1. Getting to OrthoMCL from VEuPathDB databases

Note: For this exercise use http://cryptodb.org and http://orthomcl.org/

a. Use the CryptoDB site search to visit the gene page for the *Cryptosporidium muris* gene, CMU_034340, hypothetical protein, conserved.

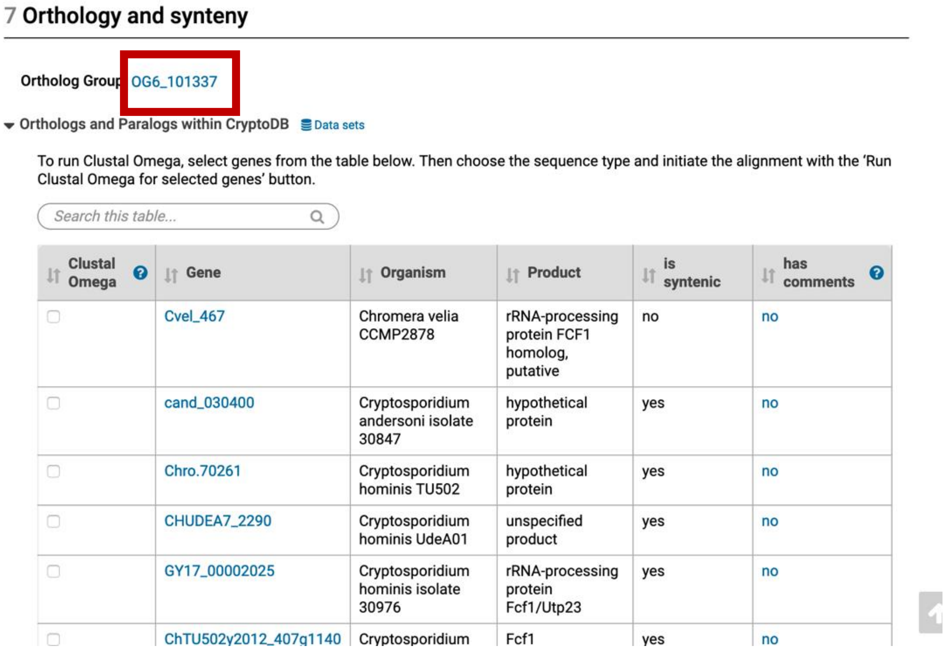
b. What information on the gene page can you use to guess a function for this gene? It is annotated as a hypothetical protein! Hint: look at the orthologs table and the domains in the protein features graph. You may also want to visit some of the external links or take a look at InterPro domains.





	Transcript ID	Isoelectric Point	Molecular Weight	Has SignalP	Has TMHMM	Protein Length	Pro Bro
	CMU_034340-t26_1	10.5	23784	no	no	206	Interac

c. Go to the Orthology and Synteny section and look at the table labeled “Orthologs and Paralogs within CryptoDB”. Does this gene have orthologs in other *Cryptosporidium* species? What about other organisms? (hint: scan the organism column in the table)



7 Orthology and synteny

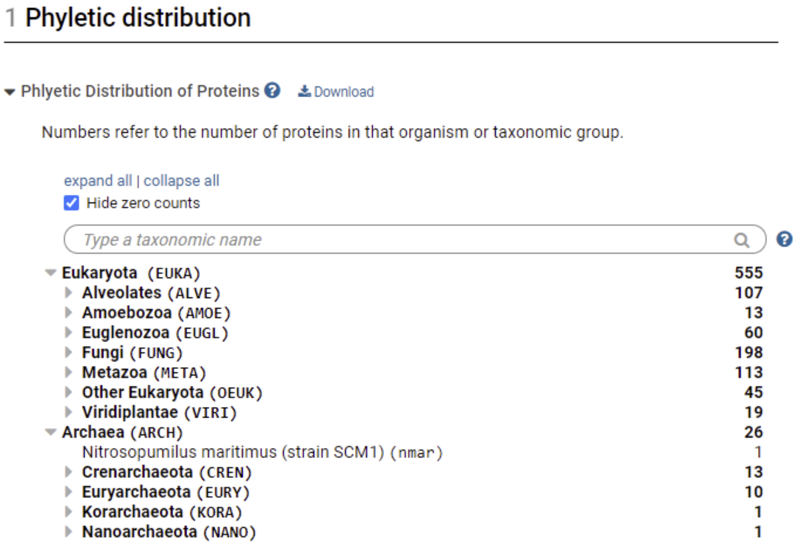
Ortholog Group OG6_101337

▾ Orthologs and Paralogs within CryptoDB

To run Clustal Omega, select genes from the table below. Then choose the sequence type and initiate the alignment with the 'Run Clustal Omega for selected genes' button.

Clustal Omega	Gene	Organism	Product	is syntenic	has comments
☐	Cvel_467	Chromera velia CCMP2878	rRNA-processing protein FCF1 homolog, putative	no	no
☐	cand_030400	Cryptosporidium andersoni isolate 30847	hypothetical protein	yes	no
☐	Chro.70261	Cryptosporidium hominis TU502	hypothetical protein	yes	no
☐	CHUDEA7_2290	Cryptosporidium hominis UdeA01	unspecified product	yes	no
☐	GY17_00002025	Cryptosporidium hominis isolate 30976	rRNA-processing protein Fcf1/Utp23	yes	no
☐	ChTU502y2012_407g1140	Cryptosporidium	Fcf1	yes	no

d. What about orthologs in organisms not in VEuPathDB? (hint: click on the Ortholog Group link above the table to examine the orthology information for the group at OrthoMCL.org). Does it have any orthologs in bacteria or archaea?



1 Phyletic distribution

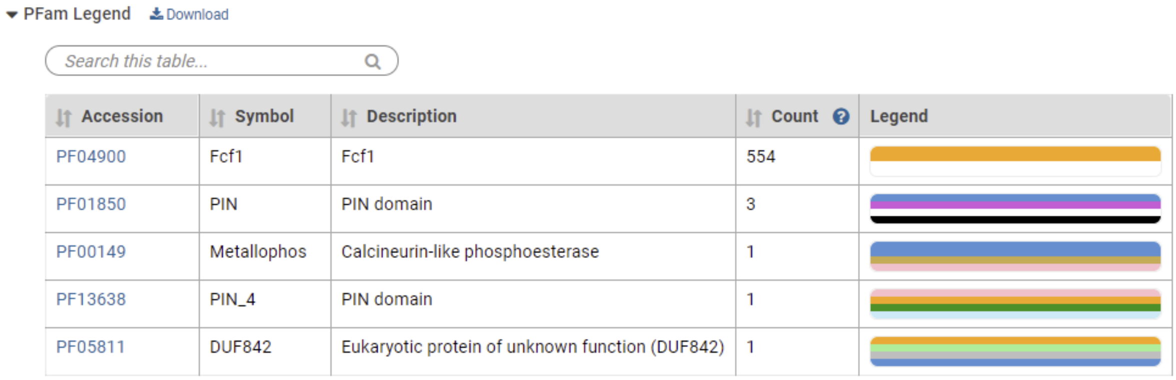
▾ Phlyetic Distribution of Proteins

Numbers refer to the number of proteins in that organism or taxonomic group.

☑ Hide zero counts

▾ **Eukaryota** (EUKA)	**555**
▸ **Alveolates** (ALVE)	**107**
▸ **Amoebozoa** (AMOE)	**13**
▸ **Euglenozoa** (EUGL)	**60**
▸ **Fungi** (FUNG)	**198**
▸ **Metazoa** (META)	**113**
▸ **Other Eukaryota** (OEUK)	**45**
▸ **Viridiplantae** (VIRI)	**19**
▾ **Archaea** (ARCH)	**26**
Nitrosopumilus maritimus (strain SCM1) (nmar)	1
▸ **Crenarchaeota** (CREN)	**13**
▸ **Euryarchaeota** (EURY)	**10**
▸ **Korarchaeota** (KORA)	**1**
▸ **Nanoarchaeota** (NANO)	**1**

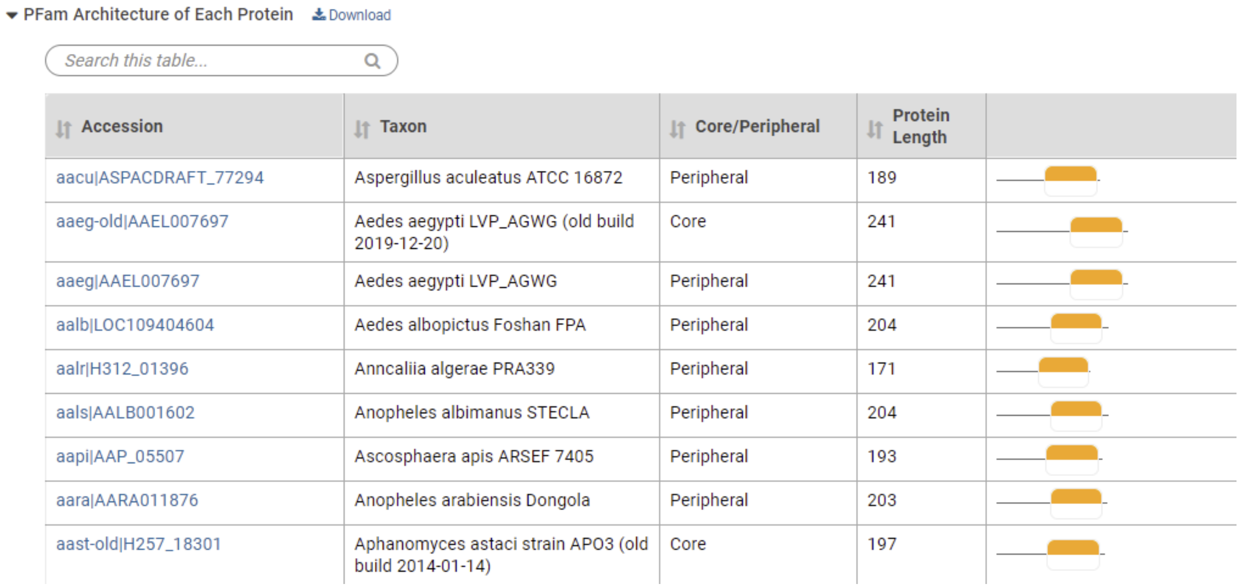
e. Scroll down to the PFam domains section. Domain architectures are found under the PFam Architecture of Each Protein table and are described in the PFam Legend table. Do all the proteins in this group have similar domain architecture? What is the distribution of the PF04900 domain across the 581 proteins in this ortholog group? PF00149?



▾ PFam Legend Download

Search this table...

Accession	Symbol	Description	Count	Legend
PF04900	Fcf1	Fcf1	554	🟧
PF01850	PIN	PIN domain	3	🟪⬛
PF00149	Metallophos	Calcineurin-like phosphoesterase	1	🟦
PF13638	PIN_4	PIN domain	1	🟩
PF05811	DUF842	Eukaryotic protein of unknown function (DUF842)	1	🟩🟦



▾ PFam Architecture of Each Protein Download

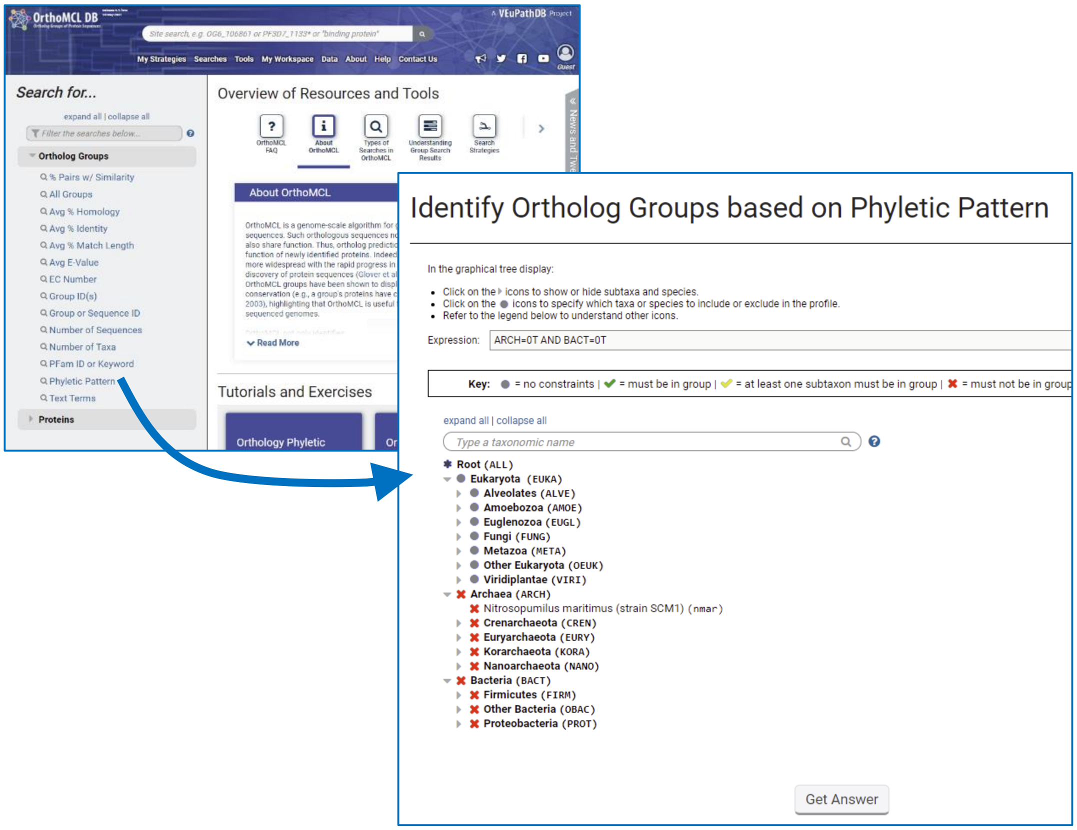
Search this table...

Accession	Taxon	Core/Peripheral	Protein Length	
aacu|ASPACDRAFT_77294	Aspergillus aculeatus ATCC 16872	Peripheral	189	
aaeg-old|AAEL007697	Aedes aegypti LVP_AGWG (old build 2019-12-20)	Core	241	
aaeg|AAEL007697	Aedes aegypti LVP_AGWG	Peripheral	241	
aalb|LOC109404604	Aedes albopictus Foshan FPA	Peripheral	204	
aalr|H312_01396	Anncaliia algerae PRA339	Peripheral	171	
aals|AALB001602	Anopheles albimanus STECLA	Peripheral	204	
aapi|AAP_05507	Ascosphaera apis ARSEF 7405	Peripheral	193	
aara|AARA011876	Anopheles arabiensis Dongola	Peripheral	203	
aast-old|H257_18301	Aphanomyces astaci strain APO3 (old build 2014-01-14)	Core	197	

f. Based on the orthologs and the PFam domains shared by the group, what do you think this protein might be doing? If you had to give this gene a name, what would you call it?

2. Using the phyletic pattern tool in OrthoMCL
Note: For this exercise use http://orthomcl.org/

a. How many orthology groups OrthoMCL do not have any orthologs in bacteria or archaea? How many protein groups do not contain orthologs from bacteria and archaea?



Identify Ortholog Groups based on Phyletic Pattern

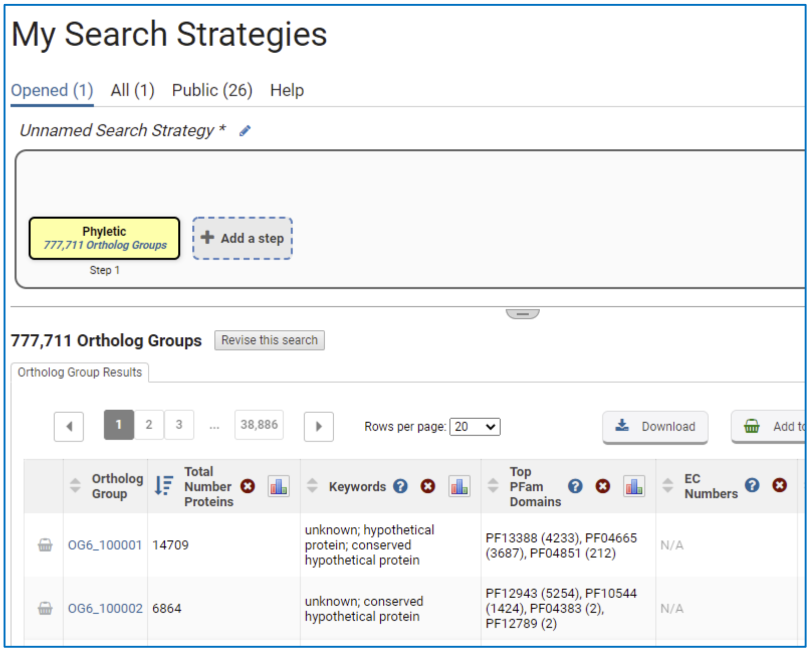
In the graphical tree display:

- Click on the ▸ icons to show or hide subtaxa and species.
- Click on the ● icons to specify which taxa or species to include or exclude in the profile.
- Refer to the legend below to understand other icons.

Expression: ARCH=0T AND BACT=0T

Key: ● = no constraints | ✓ = must be in group | ✓ = at least one subtaxon must be in group | ✗ = must not be in group

- ✱ **Root** (ALL)
  - ● **Eukaryota** (EUKA)
    - ● **Alveolates** (ALVE)
    - ● **Amoebozoa** (AMOE)
    - ● **Euglenozoa** (EUGL)
    - ● **Fungi** (FUNG)
    - ● **Metazoa** (META)
    - ● **Other Eukaryota** (OEUK)
    - ● **Viridiplantae** (VIRI)
  - ✗ **Archaea** (ARCH)
    - ✗ Nitrosopumilus maritimus (strain SCM1) (nmar)
    - ✗ **Crenarchaeota** (CREN)
    - ✗ **Euryarchaeota** (EURY)
    - ✗ **Korarchaeota** (KORA)
    - ✗ **Nanoarchaeota** (NANO)
  - ✗ **Bacteria** (BACT)
    - ✗ **Firmicutes** (FIRM)
    - ✗ **Other Bacteria** (OBAC)
    - ✗ **Proteobacteria** (PROT)

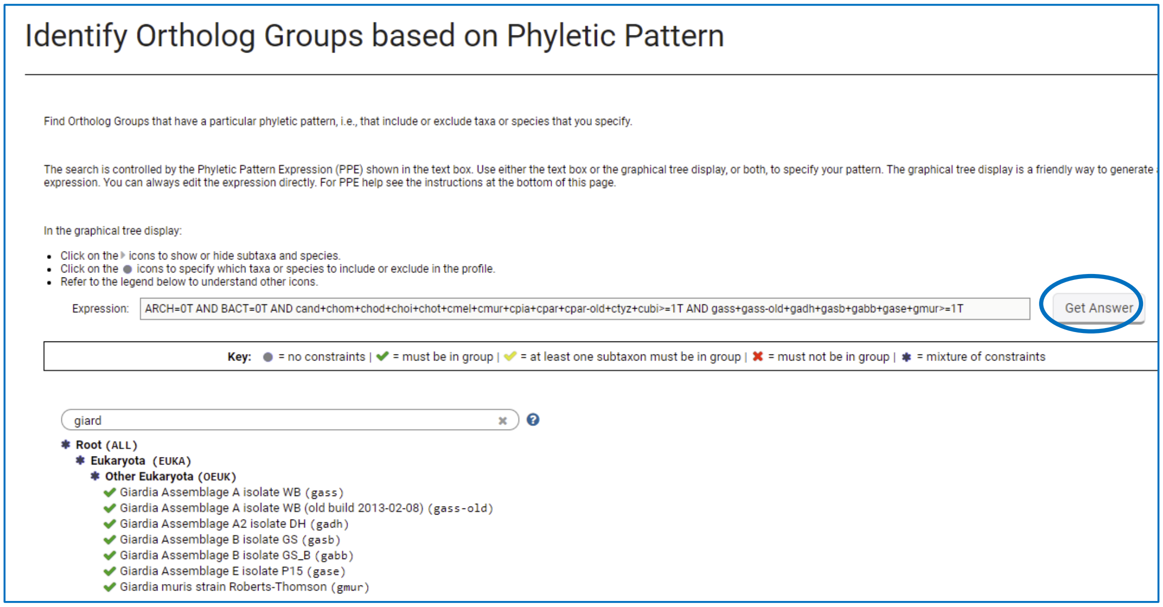


777,711 Ortholog Groups

	Ortholog Group	Total Number Proteins	Keywords	Top PFam Domains	EC Numbers
	OG6_100001	14709	unknown; hypothetical protein; conserved hypothetical protein	PF13388 (4233), PF04665 (3687), PF04851 (212)	N/A
	OG6_100002	6864	unknown; conserved hypothetical protein	PF12943 (5254), PF10544 (1424), PF04383 (2), PF12789 (2)	N/A

b. Find all groups that contain orthologs from at least one species of *Cryptosporidium* and *Giardia* but not from bacteria or archaea. If you are getting frustrated trying to figure this one out, you have a right to be! You cannot answer this question by using the check boxes (we will discuss why). However, OrthoMCL has an added feature that allows you to enter an expression to define the phyletic pattern. This option provides additional flexibility. Can you figure out what expression to use to answer this question? (hint: scroll down to the bottom of the page to find additional information about expression parameters.

   Before looking at the answer below, try this on your own or with the people in your breakout room.



Identify Ortholog Groups based on Phyletic Pattern

Find Ortholog Groups that have a particular phyletic pattern, i.e., that include or exclude taxa or species that you specify.

The search is controlled by the Phyletic Pattern Expression (PPE) shown in the text box. Use either the text box or the graphical tree display, or both, to specify your pattern. The graphical tree display is a friendly way to generate
expression. You can always edit the expression directly. For PPE help see the instructions at the bottom of this page.

In the graphical tree display:

- Click on the ▸ icons to show or hide subtaxa and species.
- Click on the ● icons to specify which taxa or species to include or exclude in the profile.
- Refer to the legend below to understand other icons.

Expression: ARCH=0T AND BACT=0T AND cand+chom+chod+choi+chot+cmel+cmur+cpia+cpar+cpar-old+ctyz+cubi>=1T AND gass+gass-old+gadh+gasb+gabb+gase+gmur>=1T

Key: ● = no constraints | ✓ = must be in group | ✓ = at least one subtaxon must be in group | ✖ = must not be in group | ✱ = mixture of constraints

giard

- ✱ **Root** (ALL)
  - ✱ **Eukaryota** (EUKA)
    - ✱ **Other Eukaryota** (OEUK)
      - ✓ Giardia Assemblage A isolate WB (gass)
      - ✓ Giardia Assemblage A isolate WB (old build 2013-02-08) (gass-old)
      - ✓ Giardia Assemblage A2 isolate DH (gadh)
      - ✓ Giardia Assemblage B isolate GS (gasb)
      - ✓ Giardia Assemblage B isolate GS_B (gabb)
      - ✓ Giardia Assemblage E isolate P15 (gase)
      - ✓ Giardia muris strain Roberts-Thomson (gmur)

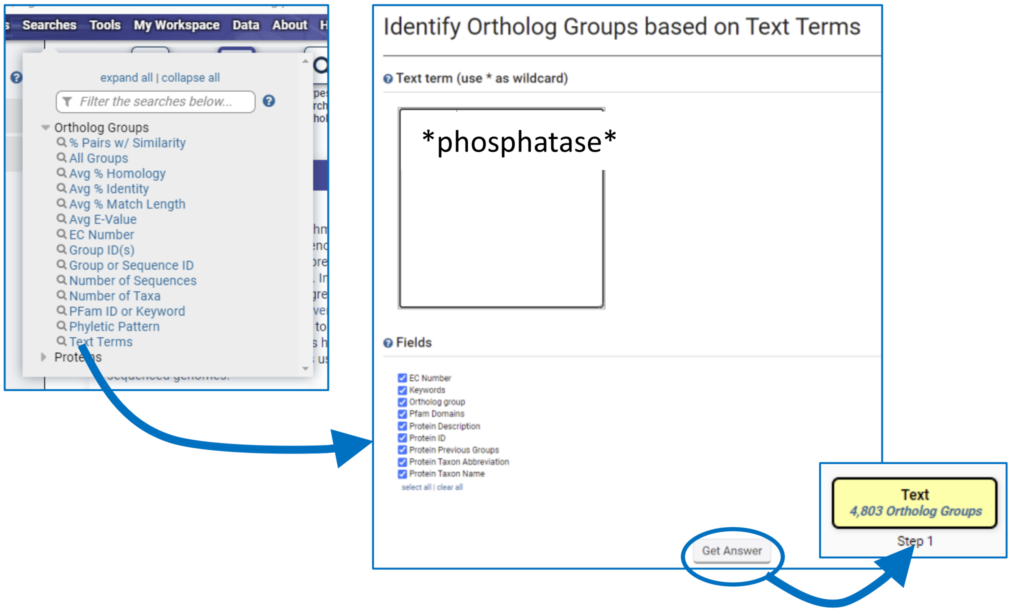
ARCH=0T AND BACT=0T AND cand+chom+chod+choi+chot+cmel+cmur+cpia+cpar+cpar-old+ctyz+cubi>=1T AND gass+gass-old+gadh+gasb+gabb+gase+gmur>=1T

c. All VEuPathDB sites also have a phyletic pattern search that uses OrthoMCL data under Genes -> Orthology and synteny -> Orthology Phylogenetic Profile. This search is very useful to identify genes in your organism of interest that are restricted in their profile. For example, you frequently want to identify genes that are conserved among organisms in your genus but not present in the host as these genes may make good drug targets or vaccine candidates. Optional: go to your favorite VEuPathDB site and run this search to identify all genes that are not present in human or mouse.

3. **Combining searches in OrthoMCL** (Use http://orthomcl.org for this exercise).

Find all plant proteins that are likely phosphatases that do not have orthologs outside of plants.

a. Use the text search **to find OrthoMCL groups** that contain the word “*phosphatase*” (note that the search should be run without the quotation marks but with the asterisks).



Identify Ortholog Groups based on Text Terms

Text term (use * as wildcard)

phosphatase

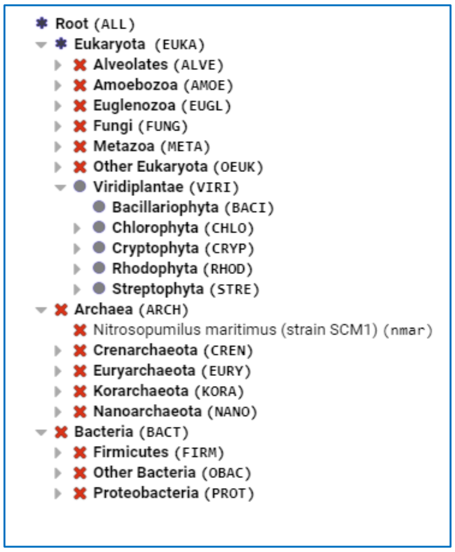
Fields

☑ EC Number
☑ Keywords
☑ Ortholog group
☑ Pfam Domains
☑ Protein Description
☑ Protein ID
☑ Protein Previous Groups
☑ Protein Taxon Abbreviation
☑ Protein Taxon Name

select all | clear all

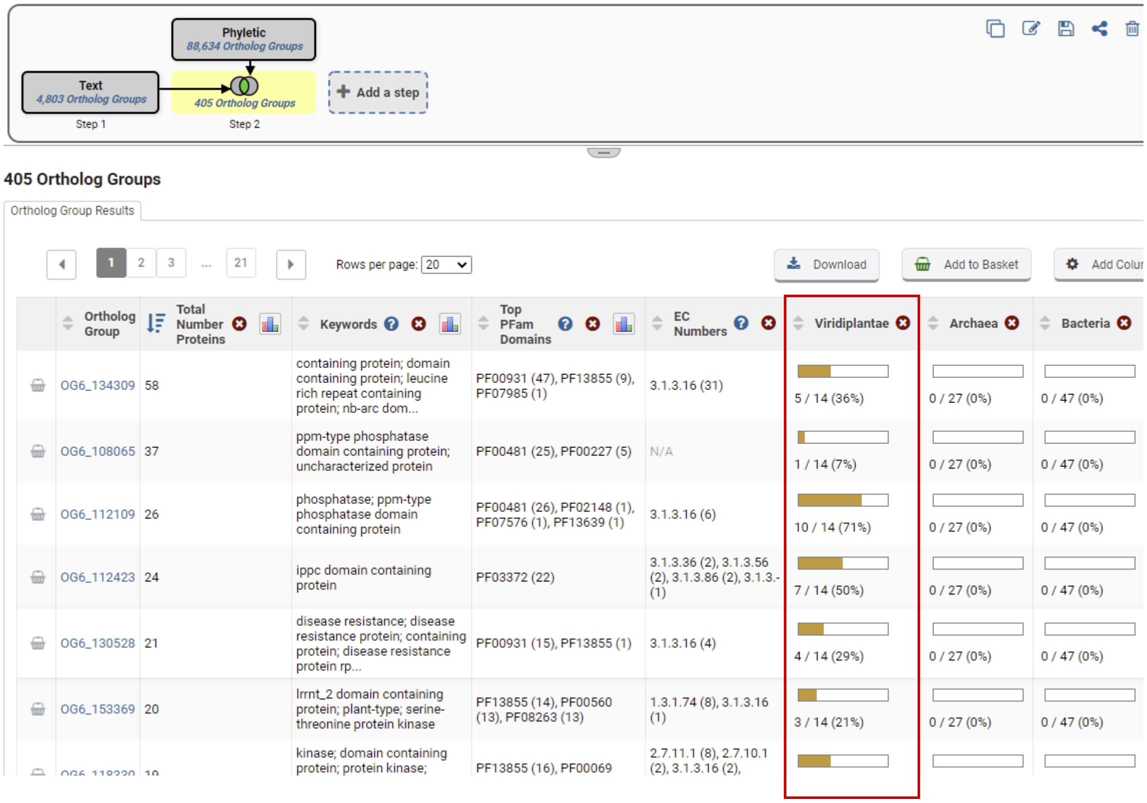
Text
4,803 Ortholog Groups

b. Add a step and run a phyletic pattern search for groups that contain any plant protein but do not contain any other organism outside plants. (hint: make sure everything has a red x on it except for plants (Viridiplantae (VIRI)), which should be a grey circle).



- ✱ Root (ALL)
  - ✱ Eukaryota (EUKA)
    - ✖ Alveolates (ALVE)
    - ✖ Amoebozoa (AMOE)
    - ✖ Euglenozoa (EUGL)
    - ✖ Fungi (FUNG)
    - ✖ Metazoa (META)
    - ✖ Other Eukaryota (OEUK)
    - ● Viridiplantae (VIRI)
      - ● Bacillariophyta (BACI)
      - ● Chlorophyta (CHLO)
      - ● Cryptophyta (CRYP)
      - ● Rhodophyta (RHOD)
      - ● Streptophyta (STRE)
  - ✖ Archaea (ARCH)
    - ✖ Nitrosopumilus maritimus (strain SCM1) (nmar)
    - ✖ Crenarchaeota (CREN)
    - ✖ Euryarchaeota (EURY)
    - ✖ Korarchaeota (KORA)
    - ✖ Nanoarchaeota (NANO)
  - ✖ Bacteria (BACT)
    - ✖ Firmicutes (FIRM)
    - ✖ Other Bacteria (OBAC)
    - ✖ Proteobacteria (PROT)

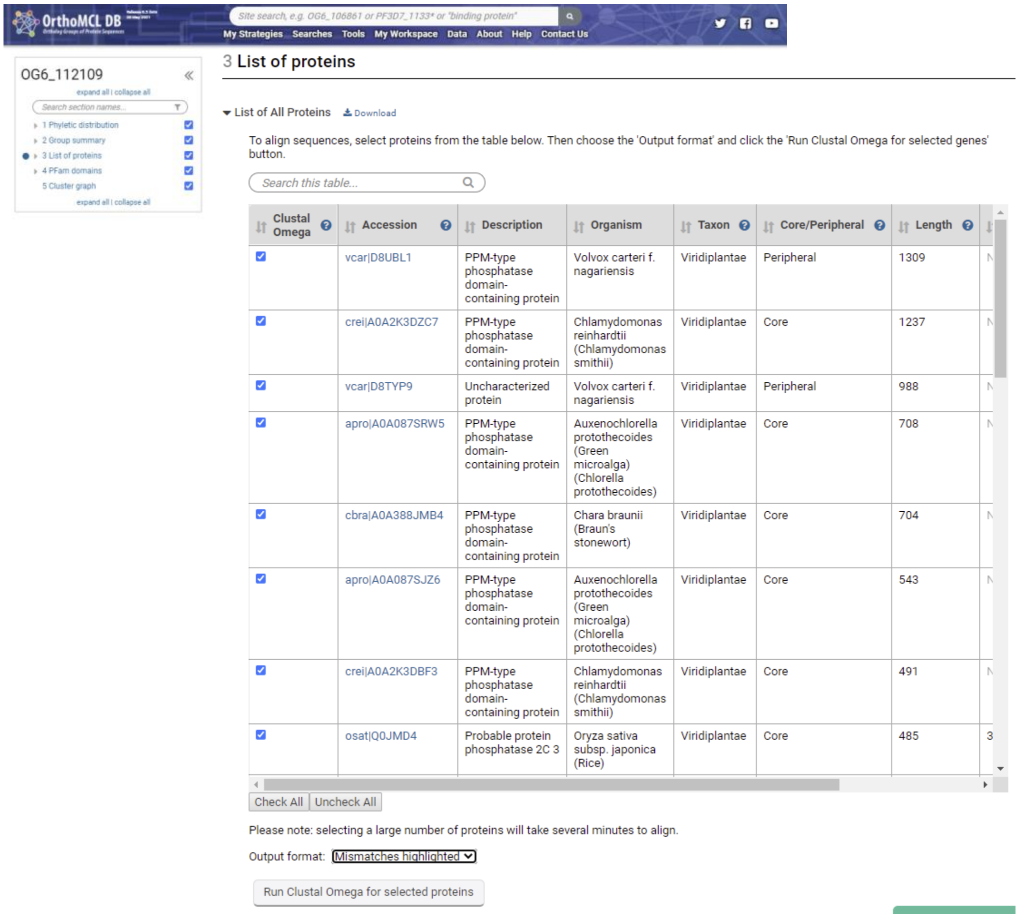
c. Examine your results. How many groups were returned by the search? What is the distribution of plant proteins in each orthology group?



405 Ortholog Groups

Ortholog Group	Total Number Proteins	Keywords	Top PFam Domains	EC Numbers	Viridiplantae	Archaea	Bacteria
OG6_134309	58	containing protein; domain containing protein; leucine rich repeat containing protein; nb-arc dom...	PF00931 (47), PF13855 (9), PF07985 (1)	3.1.3.16 (31)	5 / 14 (36%)	0 / 27 (0%)	0 / 47 (0%)
OG6_108065	37	ppm-type phosphatase domain containing protein; uncharacterized protein	PF00481 (25), PF00227 (5)	N/A	1 / 14 (7%)	0 / 27 (0%)	0 / 47 (0%)
OG6_112109	26	phosphatase; ppm-type phosphatase domain containing protein	PF00481 (26), PF02148 (1), PF07576 (1), PF13639 (1)	3.1.3.16 (6)	10 / 14 (71%)	0 / 27 (0%)	0 / 47 (0%)
OG6_112423	24	ippc domain containing protein	PF03372 (22)	3.1.3.36 (2), 3.1.3.56 (2), 3.1.3.86 (2), 3.1.3.- (1)	7 / 14 (50%)	0 / 27 (0%)	0 / 47 (0%)
OG6_130528	21	disease resistance; disease resistance protein; containing protein; disease resistance protein rp...	PF00931 (15), PF13855 (1)	3.1.3.16 (4)	4 / 14 (29%)	0 / 27 (0%)	0 / 47 (0%)
OG6_153369	20	lrrnt_2 domain containing protein; plant-type; serine-threonine protein kinase	PF13855 (14), PF00560 (13), PF08263 (13)	1.3.1.74 (8), 3.1.3.16 (1)	3 / 14 (21%)	0 / 27 (0%)	0 / 47 (0%)
OG6_118330	19	kinase; domain containing protein; protein kinase;	PF13855 (16), PF00069	2.7.11.1 (8), 2.7.10.1 (2), 3.1.3.16 (2),			

d. Run a multiple sequence alignment for OG6_112109. Click on the group ID in your result table and navigate to the List of Proteins section of the group page. The Clustal Omega tool is integrated into the table. There are several formats available for the Clustal output, making it easy to take these results to other visualization programs.



3 List of proteins

List of All Proteins Download

To align sequences, select proteins from the table below. Then choose the 'Output format' and click the 'Run Clustal Omega for selected genes' button.

Clustal Omega	Accession	Description	Organism	Taxon	Core/Peripheral	Length
☑	vcar|D8UBL1	PPM-type phosphatase domain-containing protein	Volvox carteri f. nagariensis	Viridiplantae	Peripheral	1309
☑	crei|A0A2K3DZC7	PPM-type phosphatase domain-containing protein	Chlamydomonas reinhardtii (Chlamydomonas smithii)	Viridiplantae	Core	1237
☑	vcar|D8TYP9	Uncharacterized protein	Volvox carteri f. nagariensis	Viridiplantae	Peripheral	988
☑	apro|A0A087SRW5	PPM-type phosphatase domain-containing protein	Auxenochlorella protothecoides (Green microalga) (Chlorella protothecoides)	Viridiplantae	Core	708
☑	cbra|A0A388JMB4	PPM-type phosphatase domain-containing protein	Chara braunii (Braun's stonewort)	Viridiplantae	Core	704
☑	apro|A0A087SJZ6	PPM-type phosphatase domain-containing protein	Auxenochlorella protothecoides (Green microalga) (Chlorella protothecoides)	Viridiplantae	Core	543
☑	crei|A0A2K3DBF3	PPM-type phosphatase domain-containing protein	Chlamydomonas reinhardtii (Chlamydomonas smithii)	Viridiplantae	Core	491
☑	osat|Q0JMD4	Probable protein phosphatase 2C 3	Oryza sativa subsp. japonica (Rice)	Viridiplantae	Core	485

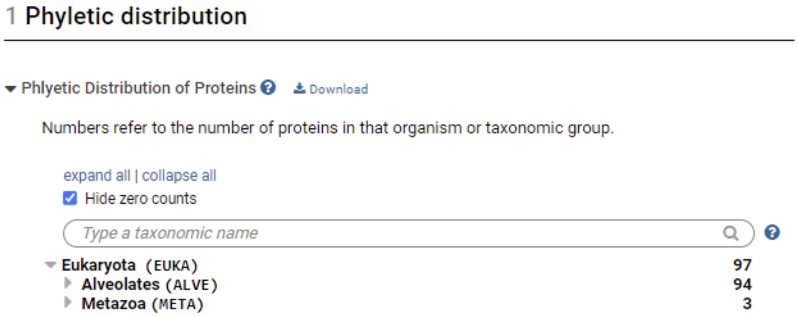
Please note: selecting a large number of proteins will take several minutes to align.

Output format: Mismatches highlighted

4. **Explore a specific OrthoMCL group - examining the cluster graph.** Use http://orthomcl.org

a. Visit the Orthomcl group OG6_131670. Type the ID into the site search to navigate to the OG6_131670.

b. Examine the Phyletic Distribution. What is the phylogenetic distribution of the members of this group? The distribution is presented as a tree. Expand the tree to view the distribution.



1 Phyletic distribution

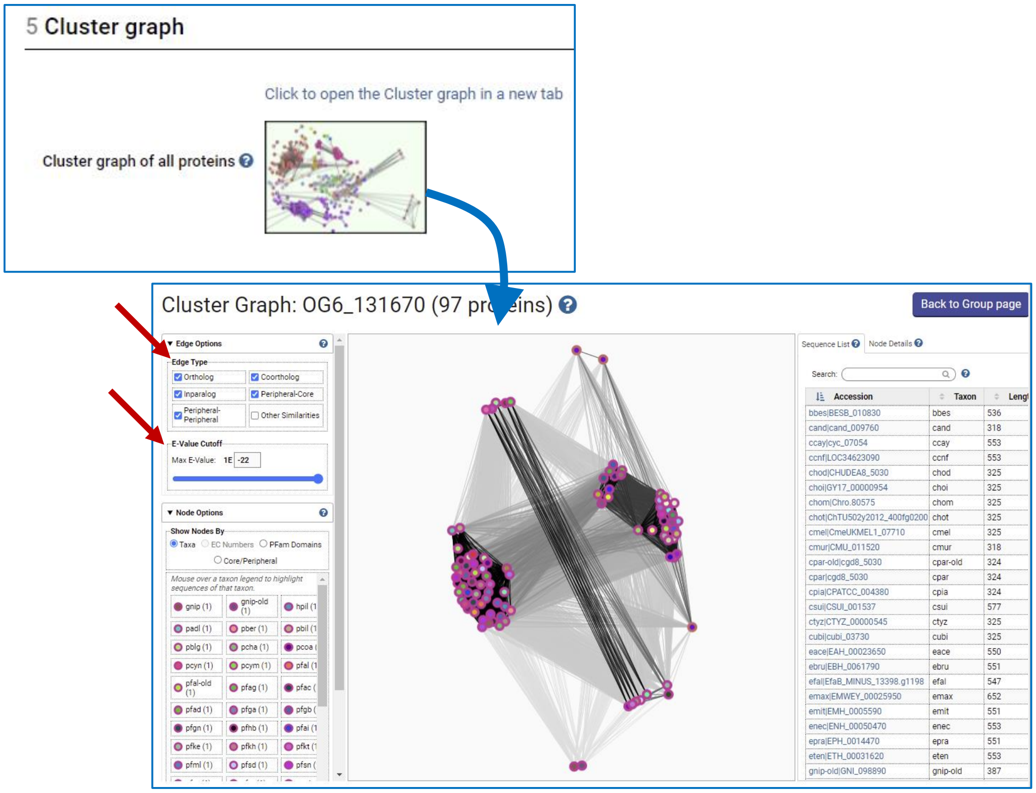
Phlyetic Distribution of Proteins Download

Numbers refer to the number of proteins in that organism or taxonomic group.

☑ Hide zero counts

Eukaryota (EUKA)	97
Alveolates (ALVE)	94
Metazoa (META)	3

c. Navigate to the Cluster graph tab. Modify the E-value cutoff slider. What happens when you increase or decrease the E-value? Can you identify subclusters of orthologs? The view of the graph can be changed using the Edge type options and the Node options.



5 Cluster graph

Click to open the Cluster graph in a new tab

Cluster graph of all proteins

Cluster Graph: OG6_131670 (97 proteins)

Back to Group page

▼ Edge Options

Edge Type

☑ Ortholog
☑ Coortholog
☑ Inparalog
☑ Peripheral-Core
☑ Peripheral-Peripheral
☐ Other Similarities

E-Value Cutoff

Max E-Value: 1E -22

▼ Node Options

Show Nodes By

● Taxa ○ EC Numbers ○ PFam Domains ○ Core/Peripheral

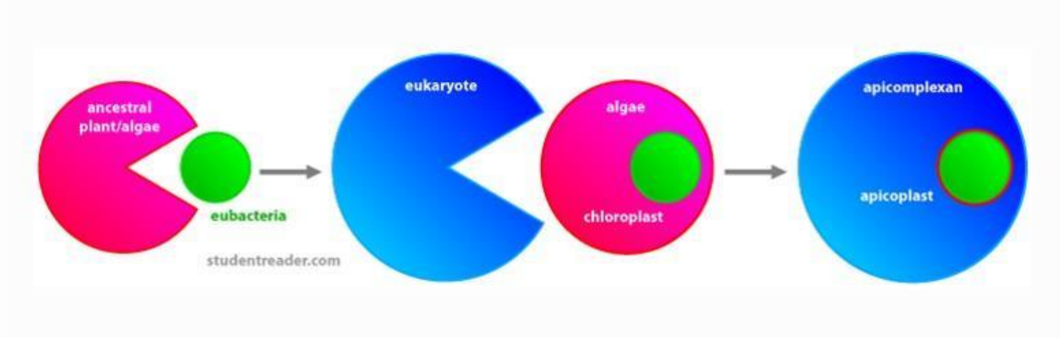
Mouse over a taxon legend to highlight sequences of that taxon.

Sequence List | Node Details

Search:

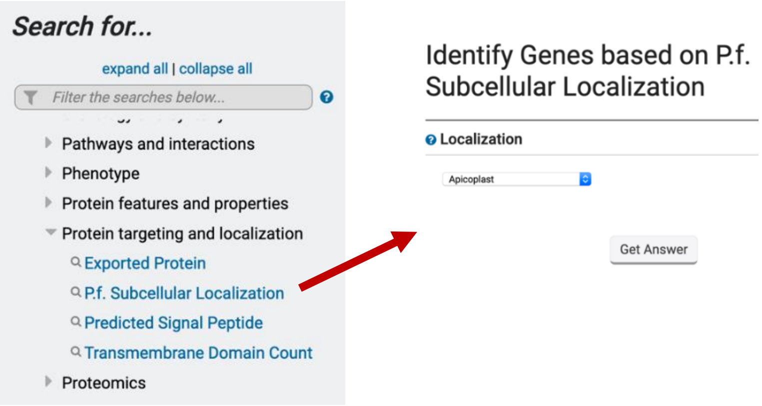
Accession	Taxon	Lengt
bbes|BESB_010830	bbes	536
cand|cand_009760	cand	318
ccay|cyc_07054	ccay	553
ccnf|LOC34623090	ccnf	553
chod|CHUDEA8_5030	chod	325
choi|GY17_00000954	choi	325
chom|Chro.80575	chom	325
chot|ChTU502y2012_400fg0200	chot	325
cmel|CmeUKMEL1_07710	cmel	325
cmur|CMU_011520	cmur	318
cpar-old|cgd8_5030	cpar-old	324
cpar|cgd8_5030	cpar	324
cpia|CPATCC_004380	cpia	324
csui|CSUI_001537	csui	577
ctyz|CTYZ_00000545	ctyz	325
cubi|cubi_03730	cubi	325
eace|EAH_00023650	eace	550
ebru|EBH_0061790	ebru	551
efal|EfaB_MINUS_13398.g1198	efal	547
emax|EMWEY_00025950	emax	652
emit|EMH_0005590	emit	551
enec|ENH_00050470	enec	553
epra|EPH_0014470	epra	551
eten|ETH_00031620	eten	553
gnip-old|GNI_098890	gnip-old	387

5. Using the orthology transform tool to identify apicoplast targeted genes in *Toxoplasma* and *Neospora*. Note: For this exercise use http://veupathdb.org

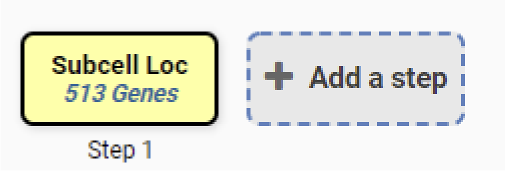


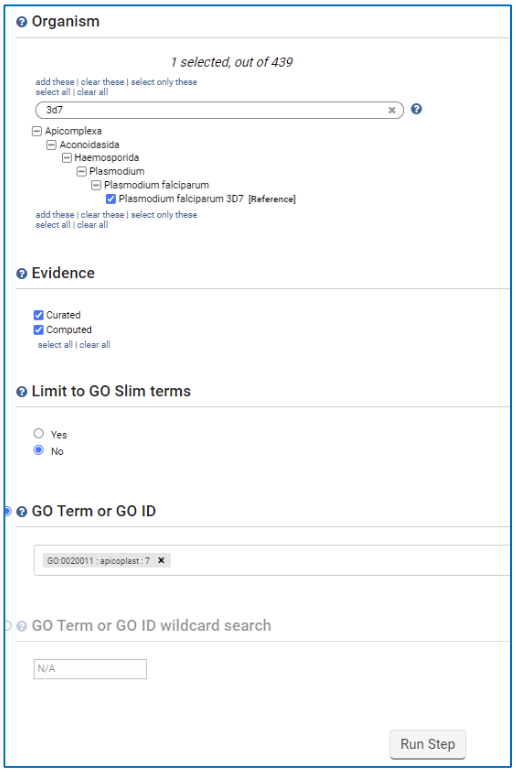
The apicoplast likely became encased in four membranes via a double endosymbiotic event. The chloroplast arose by engulfment of a cyanobacteria by a plant/algae ancestor. An algae was then engulfed by the ancestor of all apicomplexans. Thus, an apicoplast organelle arose with four membranes.

a. Start by finding genes in *Plasmodium* that are predicted to target to the apicoplast. Hint: Navigate to the P.f. Subcellular Localization search. You can further expand your list of potentially Apicoplast targeted proteins by running a GO terms search for the term "apicoplast" or the GO ID: GO:0020011 in *P falciparum* 3D7 (hint, click on add step the go to the function prediction category and select the GO term search). Which Boolean operation did you use? Union or intersect?



Localization: Apicoplast





Organism

1 selected, out of 439

3d7

- ⊟ Apicomplexa
  - ⊟ Aconoidasida
    - ⊟ Haemosporida
      - ⊟ Plasmodium
        - ⊟ Plasmodium falciparum
          - ☑ Plasmodium falciparum 3D7 [Reference]

Evidence

☑ Curated
☑ Computed

Limit to GO Slim terms

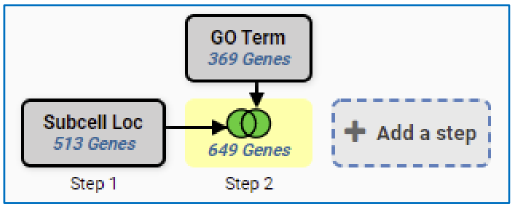
○ Yes
● No

● GO Term or GO ID

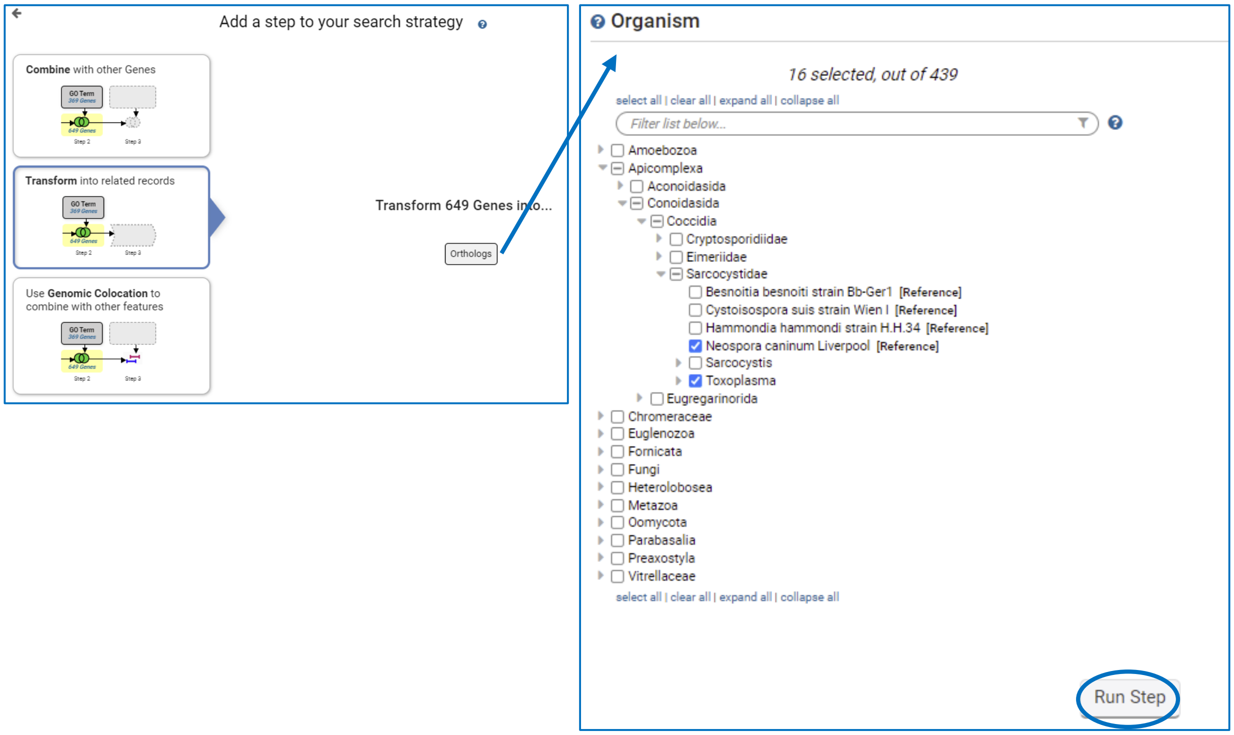
GO:0020011 : apicoplast : 7

○ GO Term or GO ID wildcard search

N/A



b. Transform the results into their *Toxoplasma* and *Neospora* orthologs. Add a step to your strategy that transforms the results into *Toxoplasma* and *Neospora*.



Add a step to your search strategy

Combine with other Genes

Transform into related records

Use **Genomic Colocation** to combine with other features

Transform 649 Genes into...

Orthologs

Organism

16 selected, out of 439

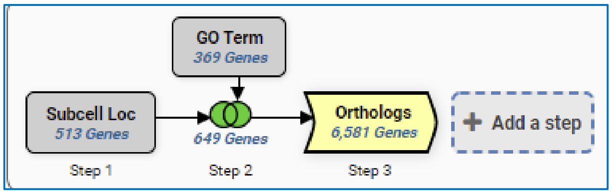
select all | clear all | expand all | collapse all

Filter list below...

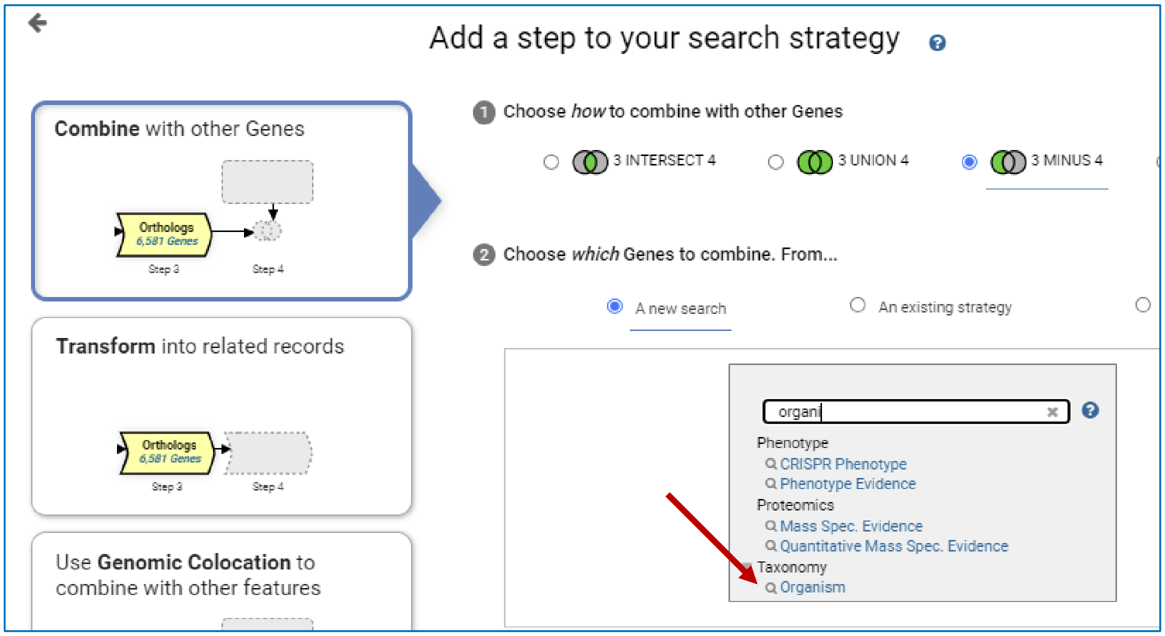
- ☐ Amoebozoa
- ⊟ Apicomplexa
  - ☐ Aconoidasida
  - ⊟ Conoidasida
    - ⊟ Coccidia
      - ☐ Cryptosporidiidae
      - ☐ Eimeriidae
      - ⊟ Sarcocystidae
        - ☐ Besnoitia besnoiti strain Bb-Ger1 [Reference]
        - ☐ Cystoisospora suis strain Wien I [Reference]
        - ☐ Hammondia hammondi strain H.H.34 [Reference]
        - ☑ Neospora caninum Liverpool [Reference]
        - ☐ Sarcocystis
        - ☑ Toxoplasma
    - ☐ Eugregarinorida
- ☐ Chromeraceae
- ☐ Euglenozoa
- ☐ Fornicata
- ☐ Fungi
- ☐ Heterolobosea
- ☐ Metazoa
- ☐ Oomycota
- ☐ Parabasalia
- ☐ Preaxostyla
- ☐ Vitrellaceae

select all | clear all | expand all | collapse all

Run Step



c. Although *Cryptosporidium* is an apicomplexan parasite it has lost its apicoplast! Can you use this fact to refine your results from the above search? Hint: try subtracting out any orthologs present in *Cryptosporidium*. You will need to use a nested strategy and use the ortholog transform back to Toxoplasma and Neospora genes for the subtraction to complete.



Add a step to your search strategy

Combine with other Genes

Transform into related records

Use **Genomic Colocation** to combine with other features

1 Choose *how* to combine with other Genes

○ 3 INTERSECT 4 ○ 3 UNION 4 ● 3 MINUS 4

2 Choose *which* Genes to combine. From...

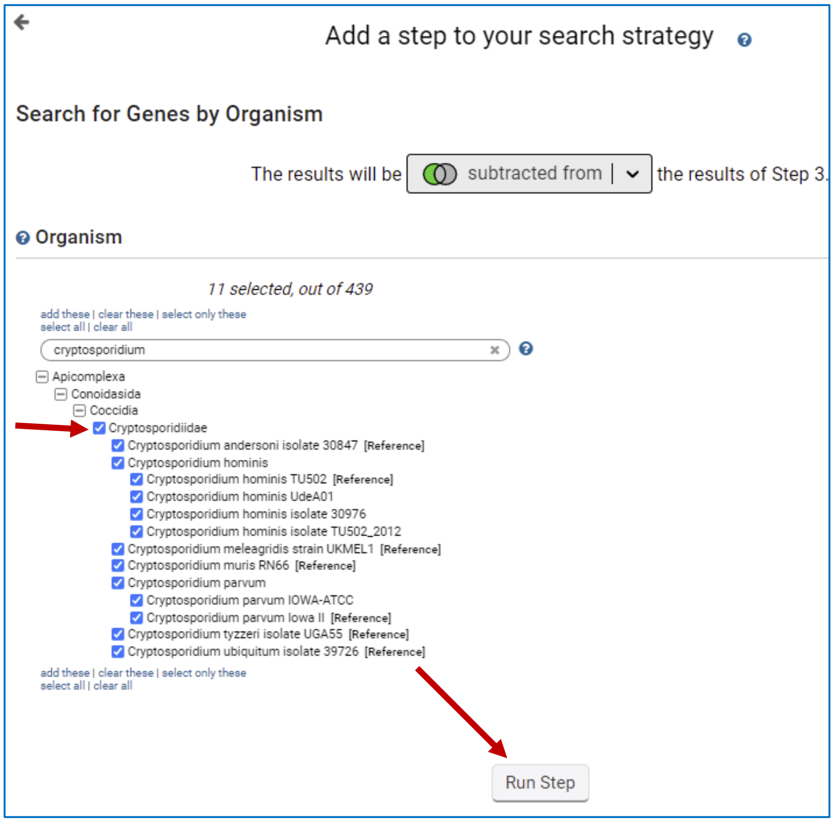
● A new search ○ An existing strategy ○

organi

Phenotype
- CRISPR Phenotype
- Phenotype Evidence

Proteomics
- Mass Spec. Evidence
- Quantitative Mass Spec. Evidence

Taxonomy
- Organism



Add a step to your search strategy

Search for Genes by Organism

The results will be subtracted from the results of Step 3.

Organism

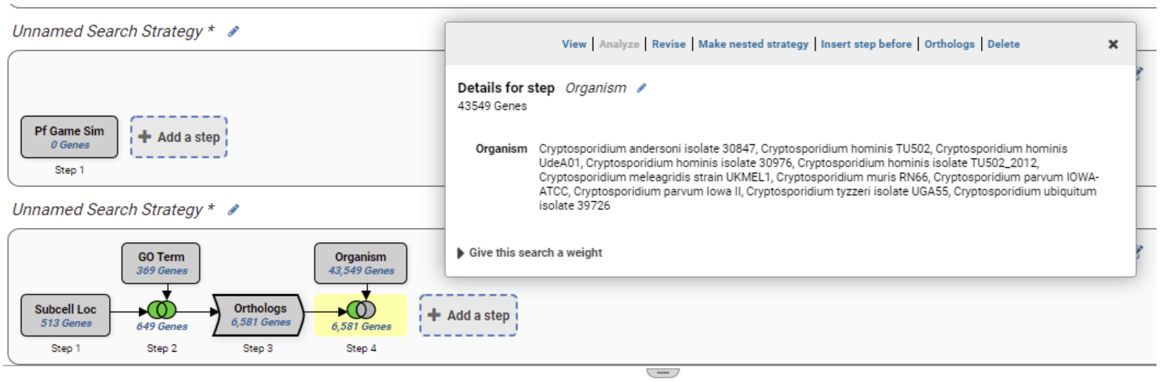
11 selected, out of 439

add these | clear these | select only these
select all | clear all

cryptosporidium

- ⊟ Apicomplexa
  - ⊟ Conoidasida
    - ⊟ Coccidia
      - ☑ Cryptosporidiidae
        - ☑ Cryptosporidium andersoni isolate 30847 [Reference]
        - ☑ Cryptosporidium hominis
          - ☑ Cryptosporidium hominis TU502 [Reference]
          - ☑ Cryptosporidium hominis UdeA01
          - ☑ Cryptosporidium hominis isolate 30976
          - ☑ Cryptosporidium hominis isolate TU502_2012
        - ☑ Cryptosporidium meleagridis strain UKMEL1 [Reference]
        - ☑ Cryptosporidium muris RN66 [Reference]
        - ☑ Cryptosporidium parvum
          - ☑ Cryptosporidium parvum IOWA-ATCC
          - ☑ Cryptosporidium parvum Iowa II [Reference]
        - ☑ Cryptosporidium tyzzeri isolate UGA55 [Reference]
        - ☑ Cryptosporidium ubiquitum isolate 39726 [Reference]

add these | clear these | select only these
select all | clear all



Details for step *Organism*
43549 Genes

Organism Cryptosporidium andersoni isolate 30847, Cryptosporidium hominis TU502, Cryptosporidium hominis UdeA01, Cryptosporidium hominis isolate 30976, Cryptosporidium hominis isolate TU502_2012, Cryptosporidium meleagridis strain UKMEL1, Cryptosporidium muris RN66, Cryptosporidium parvum IOWA-ATCC, Cryptosporidium parvum Iowa II, Cryptosporidium tyzzeri isolate UGA55, Cryptosporidium ubiquitum isolate 39726

▶ Give this search a weight