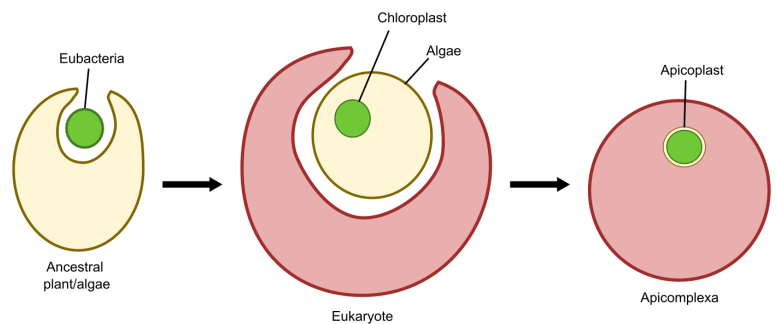




Integrate Phyletic and Genomic data in a complex Search Strategy

Introduction

The apicoplast is a unique organelle found only in apicomplexan parasites and is thought to have originated through a rare secondary endosymbiosis event. In this process, an ancestral eukaryote first engulfed a cyanobacterium to form a chloroplast, and later, an algal cell containing that chloroplast was engulfed by the ancestor of apicomplexan, resulting in an organelle surrounded by four membranes.



Because the apicoplast retains essential metabolic pathways distinct from those of the host, it represents an attractive target for anti-parasitic drug development. In this exercise, we explore how to identify genes targeted to the apicoplast in *Toxoplasma* and *Neospora* to better understand its functions and potential vulnerabilities.

Note: For this exercise use <https://veupathdb.org/veupathdb/app>

1. Start by finding genes in *Plasmodium* that are predicted to target the apicoplast.

Hint: Navigate to the [Pfal 3D7 Subcellular Localization](#) search and search for Apicoplast. You can filter the Searches menu by text query.

Site search, e.g. PF3D7_1133400 or *reductase or "binding protein"

My Strategies **Searches** Tools My Workspace Data About Help

pf

Genes
Protein features and properties
InterPro Domain
Protein targeting and localization
Pfal 3D7 Subcellular Localization
Structure analysis
Find genes whose protein products are likely targeted to the apicoplast or the RBC membrane.

Identify Genes based on P.f. Subcellular Localization

Localization

Apicoplast

Get Answer

Subcell Loc 499 Genes

+ Add a step

Step 1

2. Expand your list of potentially Apicoplast targeted proteins by adding a GO terms search for the term "apicoplast" or the GO ID: "GO:0020011" in *P. falciparum* 3D7 (Which Boolean operation should you use, Union or intersect?)

Add a step to your search strategy

Combine with other Genes

Subcell Loc 499 Genes

Step 1

Step 2

Transform into related records

Subcell Loc 499 Genes

Step 1

Step 2

Use Genomic Colocation to combine with other features

1 Choose how to combine with other Genes

1 INTERSECT 2 1 UNION 2 1 MINUS 2 2 MINUS 1

2 Choose which Genes to combine. From...

A new search An existing strategy My basket

go

Function prediction
GO Term
Phenotype
CRISPR Phenotype
Text
Text (product name, notes, etc.)

Search for Genes by GO Term

The results will be ☒ unioned with the results of Step 1.

Configure Search Learn More View Data Sets Used

Organism

1 selected, out of 622

select only these | add these | clear these

3d

Apicomplexa
Acanthamoeba
Haemosporidia
Plasmodium
Plasmodium falciparum
Plasmodium falciparum 3D7 [Reference]

Evidence

Curated
Computed
select all | clear all

Limit to GO Slim terms

Yes
No

GO Term or GO ID

GO:0020011: apicoplast: 7

GO:0020011 371 Genes

Subcell Loc 499 Genes

636 Genes

+ Add a step

Step 1

Step 2

3. Add a step to your strategy that transforms the results with *Toxoplasma* and *Neospora* orthologs. Note that this finds orthologous genes from many strains and species.

Transform into related records

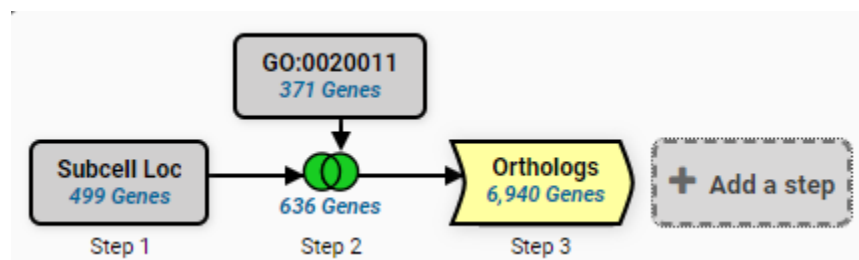
Organism

17 selected, out of 675
[select all](#) | [clear all](#) | [expand all](#) | [collapse all](#)

Filter list below... ☐ Reference only

- ☐ 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
 - ☐ Sarcocystis
 - ☒ Toxoplasma
 - ☐ Eugregarinorida
- ☐ Chromeraceae
- ☐ Euglenozoa
- ☐ Fornicata
- ☐ Fungi
- ☐ Heterolobosea
- ☐ Metazoa
- ☐ Oomycota
- ☐ Parabasalia
- ☐ Preaxostyla
- ☐ Vitrellaceae

[Run Step](#)



4. Although *Cryptosporidium* is an apicomplexan parasite it has lost its apicoplast! So one idea is that since it is missing this organelle, the parasite likely also lost many of the biochemical functions that typically take place in it. Use this fact to refine your results from the above search and remove genes that also have orthologs in *Cryptosporidium* (the assumption is that some of the missing genes are apicoplast targeted proteins in apicomplexa with an apicoplast).

Hint: try subtracting out any orthologs present in *Cryptosporidium*. You will need to use a nested strategy. First retrieve all *Cryptosporidium* genes with the Genes by Taxonomy search and then transform these to their Toxoplasma and Neospora orthologs for the subtraction to complete. Think about what kind of intersection you should be using!

← Add a step to your search strategy ⓘ

Combine with other Genes

Step 3 Step 4

Transform into related records

Step 3 Step 4

- Choose *how* to combine with other Genes

☐ 3 INTERSECT 4
 ☐ 3 UNION 4
 ☒ 3 MINUS 4
 ☐ 4 MINUS 3
- Choose *which* Genes to combine. From...

☒ A new search
 ☐ An existing strategy
 ☐ My basket

Taxonomy
 ☒ Organism



Add a step to your search strategy

The results will be subtracted from | v the resu

Configure Search Learn More View Data Sets Used

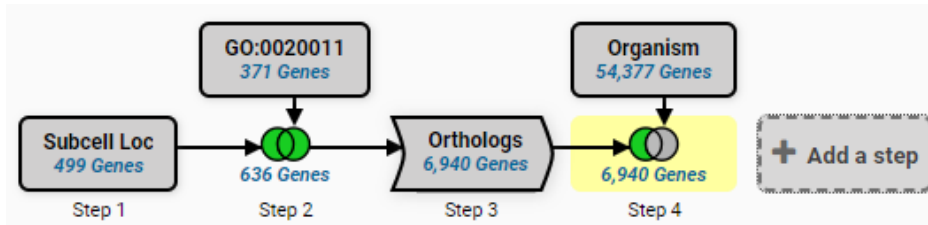
↻ Reset values to default

🔍 Organism

14 selected, out of 675
[select only these](#) | [add these](#) | [clear these](#)

cryptos ⓘ ☐ Reference only

- [-] Apicomplexa
 - [-] Conoidasida
 - [-] Coccidia
 - ☒ Cryptosporidiidae
 - ☒ Cryptosporidium andersoni isolate 30847 [Reference]
 - ☒ Cryptosporidium bovis isolate 45015 [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 ryanae 45019 [Reference]
 - ☒ Cryptosporidium sp. chipmunk genotype I strain 37763 [Reference]
 - ☒ Cryptosporidium tyzzeri isolate UGA55 [Reference]
 - ☒ Cryptosporidium ubiquitum isolate 39726 [Reference]



My Search Strategies

Opened (1) All (1) Public (11) Help

Unnamed Search Strategy *

Step 1: Subcell Loc (499 Genes) → Step 2: GO:0020011 (371 Genes) → Step 3: Orthologs (6,940 Genes) → Step 4: Organism (54,377 Genes)

Details for step Organism 54377 Genes

Organism ▶ Cryptosporidium andersoni isolate 30847, Cryptosporidium bovis isolate 45015, Cryptosporidium hominis TU502, Cryptosporidium hominis UdeA01, Cryptosporidium hominis isolate 30976, Cryptosporidium muris RN66, Cryptosporidium parvum IOWA-ATCC, Cryptosporidium parvum Iowa II, Cryptosporidium ryanae 45019, Cryptosporidium sp. chipmunk genotype I strain 37763, Cryptosporidium tyzzeri isolate UGA55, Cryptosporidium ... Show more

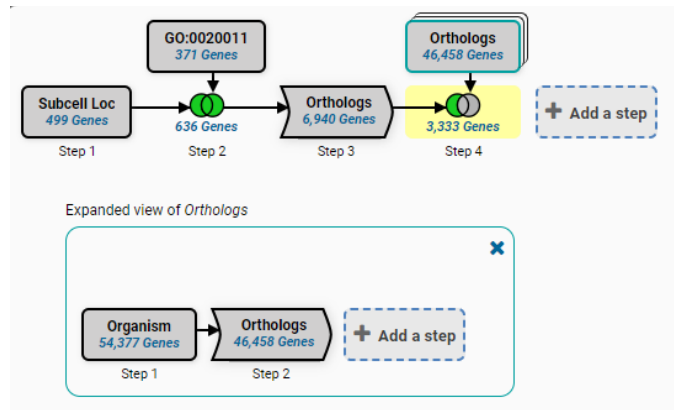
Give this search a weight

Organism

17 selected, out of 675

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
 - ☐ Sarcocystis
 - ☒ Toxoplasma
 - ☐ Eugregarinorida
 - ☐ Chromeraceae
 - ☐ Euglenozoa
 - ☐ Fornicata
 - ☐ Fungi
 - ☐ Heterolobosea
 - ☐ Metazoa
 - ☐ Oomycota
 - ☐ Parabasalia
 - ☐ Preaxostyla
 - ☐ Vitrellaceae



This leaves you with apicoplast specific genes for *Toxoplasma* and *Neospora* that you could target in future research. Note that Search Strategies can be saved, shared, and imported.

<https://veupathdb.org/veupathdb/app/workspace/strategies/import/543f14bfab645f7e>