Exploring Isolate Data
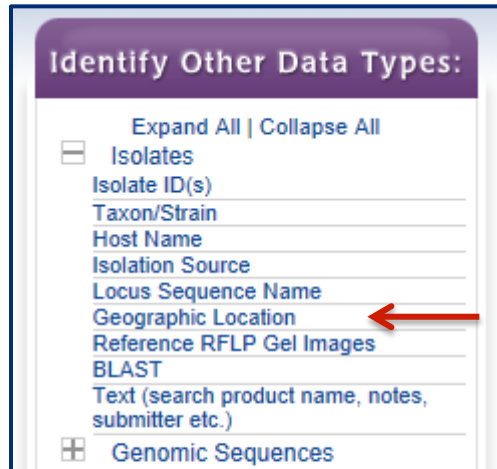**Exercise 8**

8.1 Exploring isolates in *Cryptosporidium* and using the alignment tool.
 Note: For this exercise use http://www.cryptodb.org

a. Identify all *Cryptosporidium* isolates from Europe.
   Hint: search for isolates by geographic location in the "Identify Other Data Types" section.

b. How many of the *Cryptosporidium* isolates collected in Europe were isolated from feces?
   Hint: add another isolate search step.
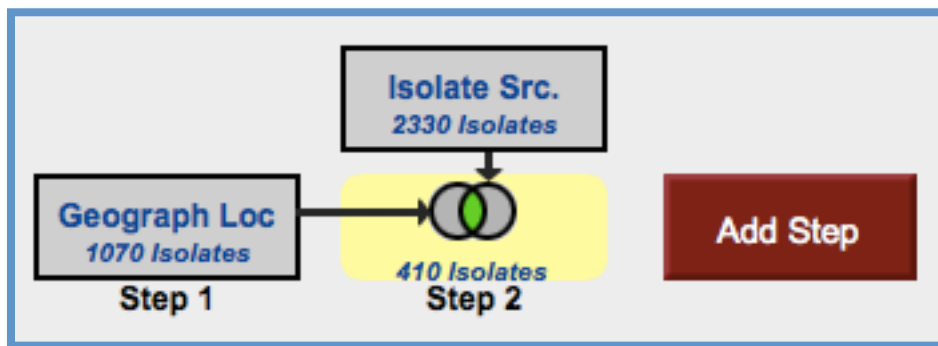
c. What is the general distribution of these isolates in Europe? (hint: you can do this quickly in two ways: sort the geographic location column by clicking on the sort arrows, then look at the represented countries; or use the "Isolate Geographic Location" tab to view a map and results summary table).

Country	Number of Isolates	Isolate Type	Latitude	Longitude
Belgium	1	Sequencing Typed	50.503887	4.469936
Czech Republic	73	Sequencing Typed	49.817492	15.472962
Germany	90	Sequencing Typed	51.165691	10.451526
Ireland	4	Sequencing Typed	53.41291	-8.24389
Italy	11	Sequencing Typed	41.87194	12.56738
Lithuania	1	Sequencing Typed	55.169438	23.881275
Netherlands	41	Sequencing Typed	52.132633	5.291266

d. Out of those in step 'b', how many are unclassified *Cryptosporidium* species? Hint: add another isolate search step.

e. How many of step 'b' isolates originated from humans? (hint: you can insert a step within a strategy. Click on the name of the step you want to insert a step before, then click on "Insert step before").

STEP 3 : Taxon

Species/Strain Name : unclassified Cryptosporidium, Cryptosporidium genotype names, Cryptosporidium lizard genotype, Cryptosporidium sp. Eublepharis/515708, Cryptosporidium rabbit genotype, Cryptosporidium rabbit genotype 1/China, Cryptosporidium rabbit genotype 4/China, Cryptosporidium sp. JY14Vb35, Cryptosporidium sp. NY15Vb29, Cryptosporidium sp. Southern elephant seal genotype, Cryptosporidium sp. Weddell seal genotype, Cryptosporidium sp. avian genotype II, Cryptosporidium sp. avian genotype III, Cryptosporidium sp. avian genotype V, Cryptosporidium sp. beaver genotype, Cryptosporidium sp. cervine genotype, Cryptosporidium sp. L1, Cryptosporidium sp. L43, Cryptosporidium sp. cervine genotype (W4), Cryptosporidium sp. chipmunk genotype (W17), Cryptosporidium sp.

f. Compare some of these isolates using the multiple sequence alignment tool (ClustalW). Do you see any sequences with insertions or deletions?

51 Isolates from Step 3
Strategy: Geograph Loc

Add 51 Isolates to Basket | Download 51 Isolates

Isolate Id	Organism	Strain/Isolate Name	Host	Geographic Location	Isolation Source
DQ116568	Cryptosporidium meleagridis	1689	Homo sapiens	United Kingdom	feces
DQ116569	Cryptosporidium meleagridis	1689	Homo sapiens	United Kingdom	feces
DQ116570	Cryptosporidium parvum	2121	Homo sapiens	United Kingdom	feces
DQ116571	Cryptosporidium parvum	1842	Homo sapiens	United Kingdom	feces
DQ116572	Cryptosporidium andersoni	1709	Homo sapiens	United Kingdom	feces
DQ388377	Cryptosporidium parvum	C1	Homo sapiens	Netherlands	human feces
DQ388378	Cryptosporidium parvum	C2	Homo sapiens	Netherlands	human feces
HM371369	Cryptosporidium parvum	CI 4	Homo sapiens	Italy	aged formalin-fixed stool from HIV-positive patient
HM371370	Cryptosporidium parvum	CI 5	Homo sapiens	Italy	aged formalin-fixed stool from HIV-positive patient
HM371371	Cryptosporidium hominis	CI 6	Homo sapiens	Italy	aged formalin-fixed stool from HIV-positive patient
HM371372	Cryptosporidium parvum	CI 7	Homo sapiens	Italy	aged formalin-fixed stool from HIV-positive patient
HM371373	Cryptosporidium parvum	CI 8	Homo sapiens	Italy	aged formalin-fixed stool from HIV-positive patient
HM371374	Cryptosporidium parvum	CI 9	Homo sapiens	Italy	aged formalin-fixed stool from HIV-positive patient
JX445924	Cryptosporidium tyzzeri	M805A	Homo sapiens	Czech Republic	feces
JX445925	Cryptosporidium tyzzeri	M805B	Homo sapiens	Czech Republic	feces
JX445926	Cryptosporidium parvum	M805C	Homo sapiens	Czech Republic	feces
KC679056	Cryptosporidium parvum	LG1	Homo sapiens	Norway: Vestfold county	feces

Please select at least two isolates to run ClustalW. Note: only isolates from a single page will be aligned. (Increase the page size in 'Advanced Paging' to increase the number that can be aligned).

Run Clustalw on Checked Strains | Check All | Uncheck All

g. Take a look at the 'guide tree' that was built using this alignment.
Change the isolates that you selected for alignment – how does the tree change?
Do isolates from the same country cluster together?

8.2 Typing an unclassified isolate.
   Note: For this exercise use http://www.cryptodb.org

a. Run a search to find all unclassified *Cryptosporidium* isolates and find one that was typed using 18S small subunit ribosomal RNA. (Hint: Identify Isolates based on Taxon/Strain and choose 'unclassified' under *Cryptosporidium*. Add a column for Gene Product and sort the column).

b. Go to the isolate record page and copy the DNA sequence.

c. Go to search for isolates based on BLAST, select isolates and make sure only the reference isolates are selected in the target organism window.

d. Paste the DNA sequence in the input window and select the Blastn program. Click on "Get Answer".

e. Explore your results. Based on the similarity which reference isolate is this one closest to?