# Using the Genome Browser (GBrowse)
(Exercise 2)

1. **Explore synteny using the genome browser?**
   For this exercise go back to the *C. parvum* DHFR gene:
   http://cryptodb.org/gene/cgd4_4460

   - Is synteny (chromosome organization) in this region maintained in other species? (hint: look in the genomic context section of the gene page – what does the shading mean?).
   - What does synteny look like across the entire chromosome? To do this:
   - o Click on the "View in GBrowse" button in the genomic context section.
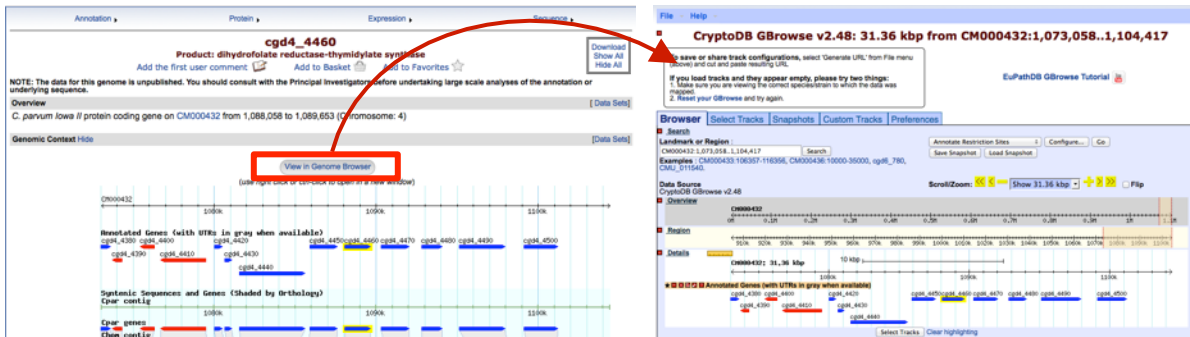
     - o Zoom out to the entire chromosome. There are a few ways to do this – for example, drag your cursor across the entire chromosome then select "zoom" from the popup menu.
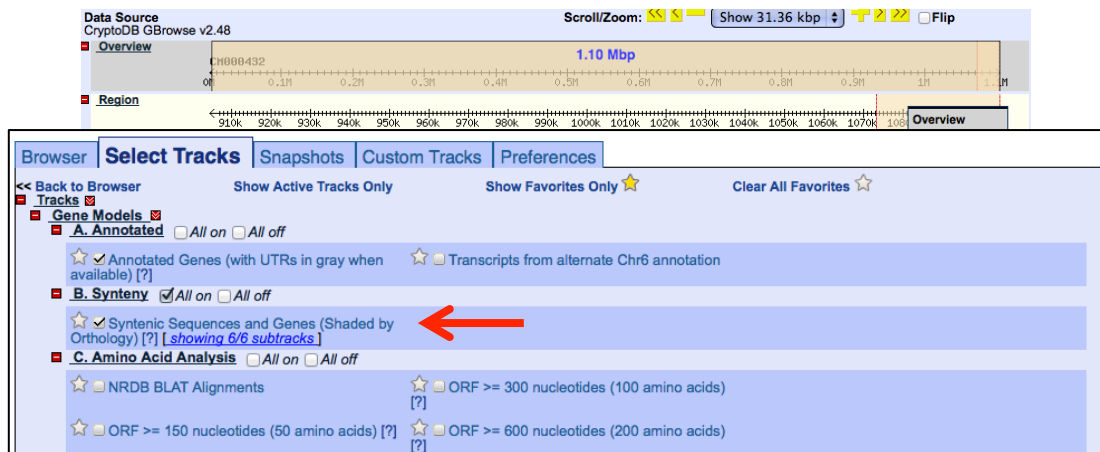
     - o Click on the tab called "Select tracks". Select the track called "Syntenic Sequences and Genes (Shaded by Orthology)". Go back to the Browser tab (this may take a minute to load).

- Where is the DHFR gene located relative to the chromosome? (hint: the gene is highlighted in yellow).
- Which genome is composed of the most fragments? Why? Notice the really big gene between 0.6M and 0.7M? Why is it smaller in *C. hominis*? Do you think this is a real biological difference or some sort of artifact?
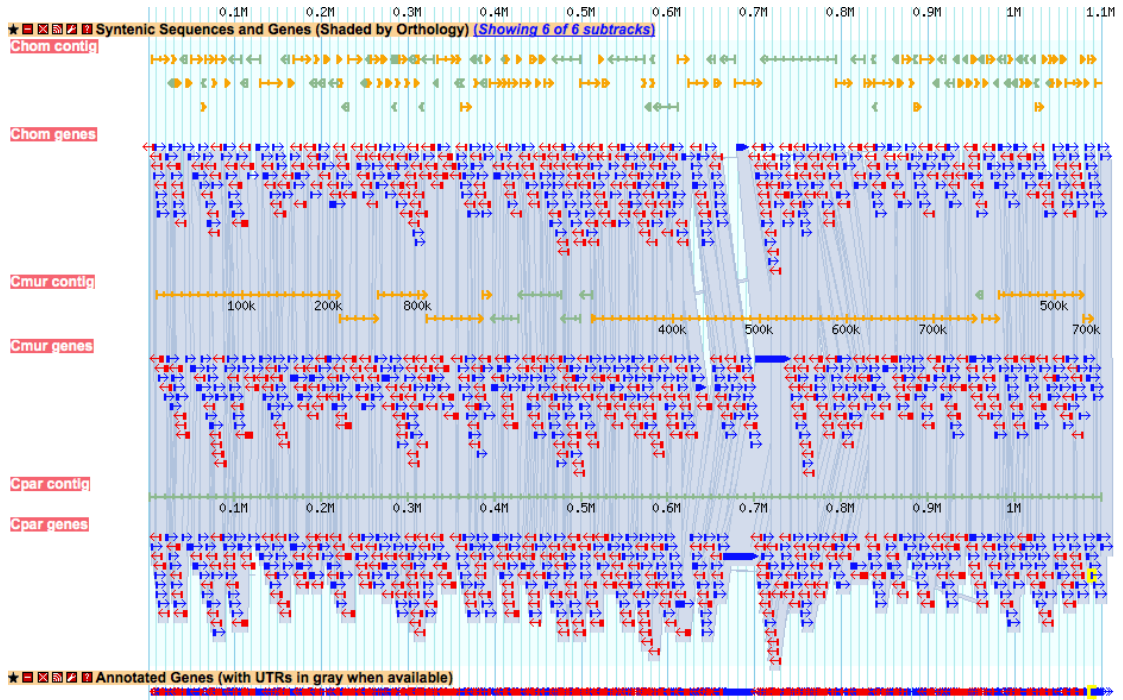
2. Exploring data tracks in GBrowse

- Explore the select tracks section. What data is available? For now, start by turning on only the track called TU114 coverage in SNPs section.
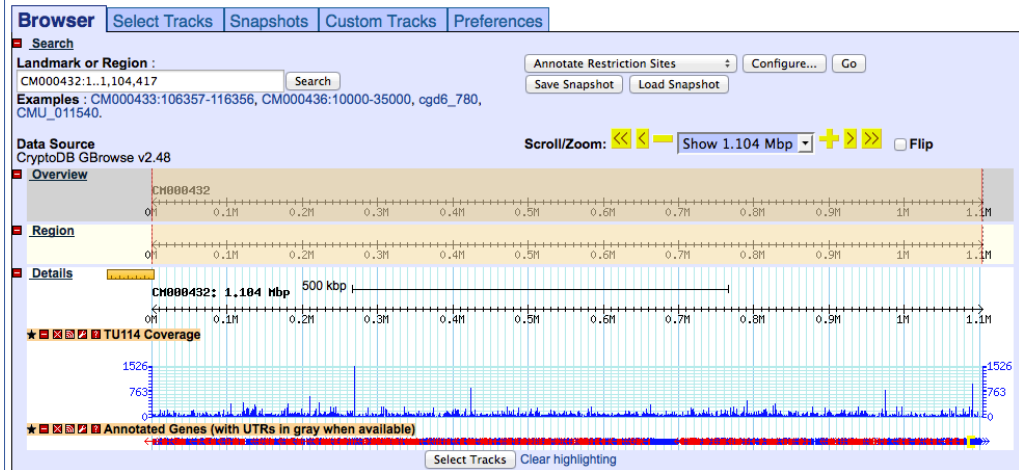
- You can turn off the Synteny tracks by either unchecking it in the "Select tracks section" or clicking on the close track icon next to the track's name in the "Browser" section.


- What is the TU114 alignment track showing you? What might the higher peaks indicate?


- Explore one of the regions with coverage more closely. Zoom into a region of interest – click and drag you mouse over the region of interest in the details section and click on "zoom" in the popup.
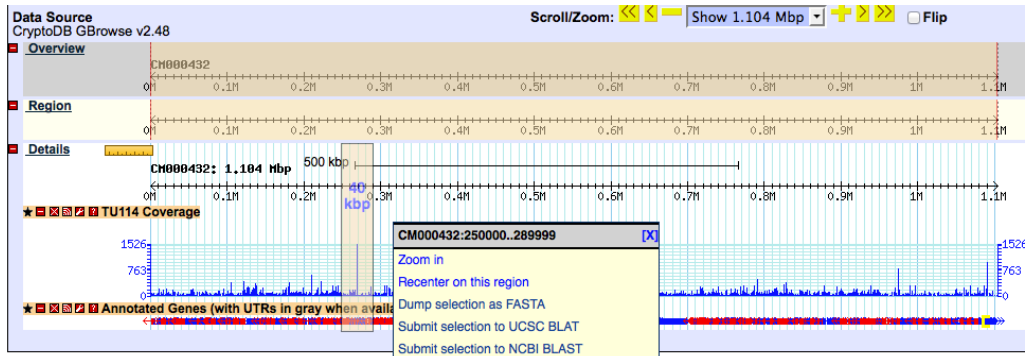
- What gene is located in this area? Why does it exhibit higher coverage? What are some possibilities? Can you turn on other tracks that would help in answering this question? (hint: turn on the low complexity and tandem repeat tracks).
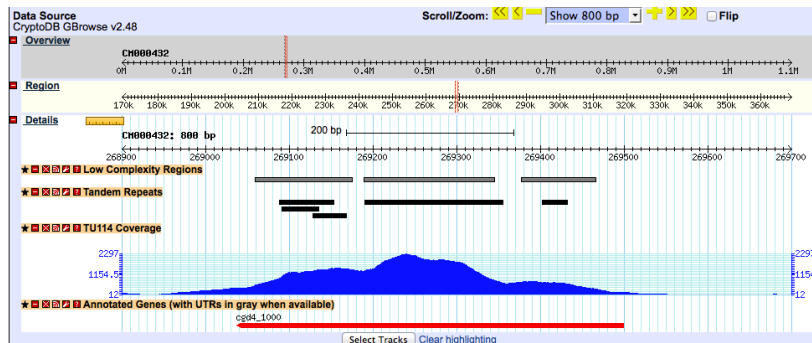
- Does this gene have any proteomics evidence? Turn on all the "Protein Expression Evidence - Mass Spec Peptides" tracks.
- What about neighboring genes? Zoom out to 100KB
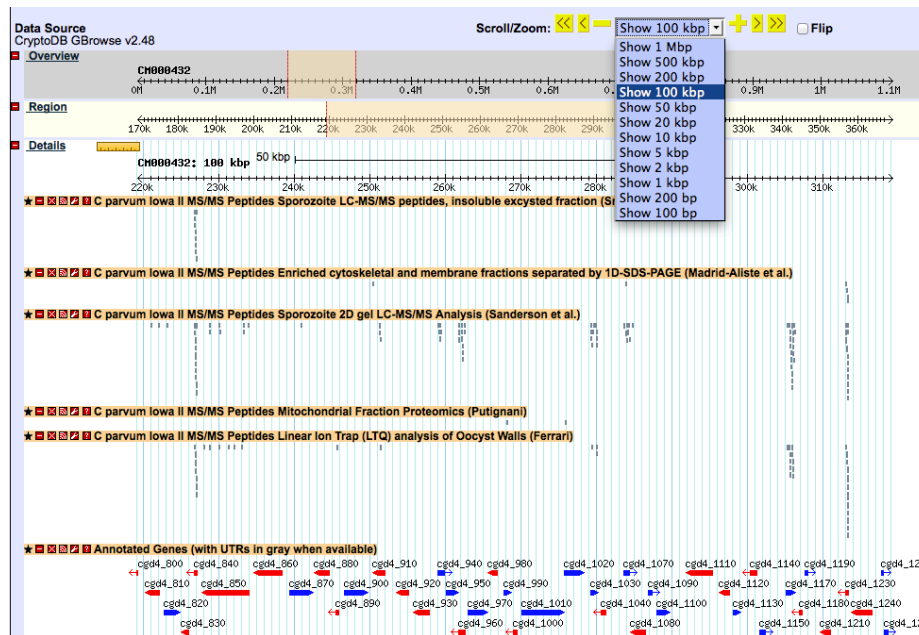
- Turn on the track for expressed sequence tags (EST). What does this data tell you? Can you find a gene with peptide evidence but no EST evidence? What about EST evidence but no peptide evidence? What could these observations suggest?