How to view your annotations as a track on the Genome Browser

You may choose to visualize your gene annotations using specialized software such as
- Artemis   www.sanger.ac.uk/resources/software/artemis/
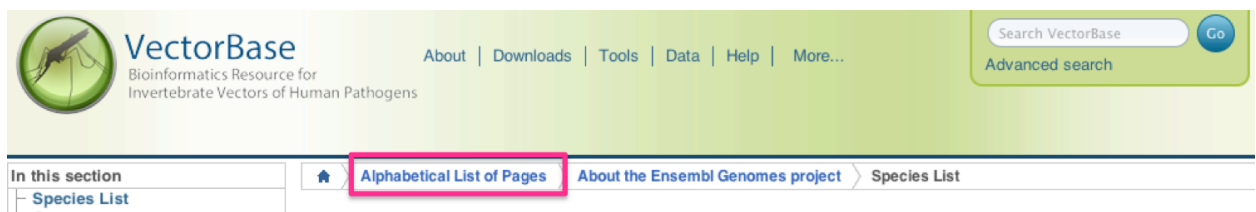- WebApollo   www.gmod.org/wiki/WebApollo

 or using VectorBase Genome Browser (GB). In this tutorial we will describe how to use VectorBase GB using a GFF file.

1. What is a GFF (General Feature Format) file format?

- Click on the TOOLS navigation tab and select "Genome Browser" to go to its front page
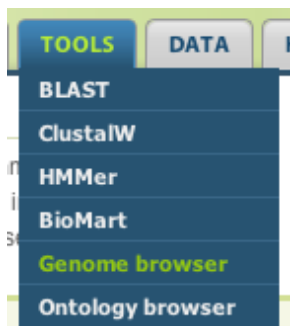
- Click on "Alphabetical List of Pages" tab

- Click on "GFF/GTF File Format" bullet under "Adding Custom Tracks' topic.

- What are the fields in each one of the nine columns?

- Open the sample file provided in the page of this tutorial with a text editor. Compare both files.

2. How to create a GFF file to put your annotatios?

You have three options:

- **Small scale** – manually. You can look at your BLAST output or output from a gene prediction program for one gene and make it conform to GFF rules. Just use any text editor that does tabs.
- **Medium scale** – Excel. GFF is simply tab delimited, many results from programs *e.g.* BLAST or gene predictions are tab/space/comma delimited and so can easily be loaded into Excel and then columns can be deleted/added/moved to produce GFF annotations. Save file as "Tab delimited text (.txt)"
- **Large scale** – scripting. Write a small script (*e.g.*, with perl or python) that takes your output (*e.g.*, BLAST or gene prediction) and converts it into GFF.

3. Upload your GFF file to VectorBase

1. For this exercise a hypothetical sample file has already been generated for you. Go to the tutorials page www.vectorbase.org/tutorials , select the tutorial called "Browsing genomes II: Visualizing and adding tracks", click on the file called "AGAMB.HOX.GFF.txt".

2. You have two options: a) click on this sample file, copy the file URL, and save it. or b) copy the file in Excel and save as "Tab delimited text (.txt)".

This file contains positions of where the eight HOX genes may be located in the *Anopheles gambiae* genome based on some BLAST hits using *Drosophila* HOX genes:
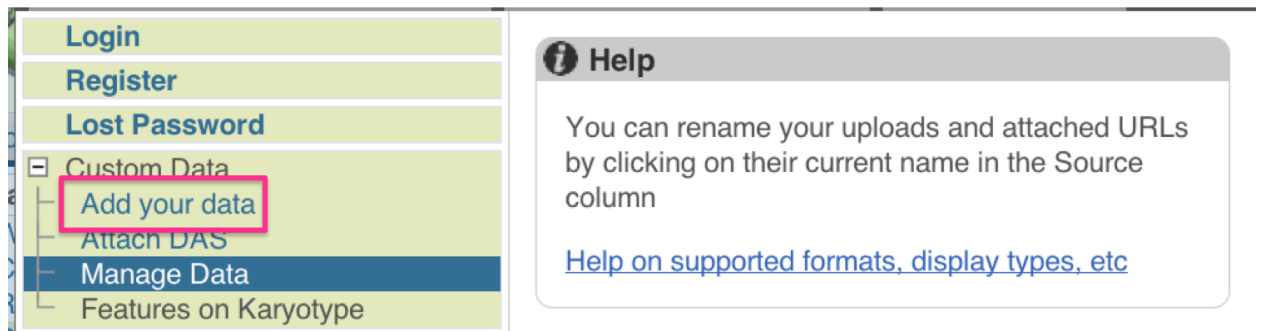
- AGAP004646-dfd
- AGAP004648-pb
- AGAP004649-lab
- AGAP004659-scr
- AGAP004660-antp
- AGAP004661-ubx
- AGAP004662-abda
- AGAP004664-abdb

Annotations could be just regions, like this example, or they could be complete gene models with exons, introns and UTRs.

3. Go to any region of *A. gambiae* genome browser, click on the location tab and click on manage your data. For example go to AGAP004646:

4. Click on "Add your data".

5. Add a custom track:
   - give any name to your track, in this example it is called 'practice exercise',
   - choose GFF as the "data format"
   - "chose file" or "provide file URL"
   - click on "Upload" your annotations.

 You should be prompted to then 'Go to nearest region with data', this will take you back to the genome browser to a region where at least one of your annotations is located at. Scroll to the bottom panel of the page and you will now see the annotation track.

Note: if there is an "Ajax error" ignore this message and close the pop out window, VectorBase developers are working to fix this bug. Re-load the page.

6. The name of the track appears on the left. Clicking on the annotation will display the data from the GFF file, *e.g.* this annotation region from the HOX gene BLASTs is called 'My-dfd'.