VBRC Base By Base: Import genes from Genbank file

Nick Tang

Abstract

Viral Bioinformatic Resource Centre

- **Provide databases of viral genomic information.**
  - Please check the **_Organisms_** menu to see which viruses we support: we're now focusing on large DNA viruses
  - The VOCs (Virus Orthologous Clusters) database is at the heart of our system.
  - The database links directly to integrated tools for comparative analyses.
  - VOCs sorts genes into ortholog clusters (e.g. RNA polymerase) to simplify data retrieval.
- **Provide easy access to the genes, gene families, and genomes of the different virus families.**
  - via a unique series of powerful Java tools that support multiple computer platforms (see **_VBRC Tools_** menu).
  - design and build software to tackle specific bioinformatics/virology problems, often in collaboration with virologists.
- **Rally the research community to provide expert curation of these viral genomes by:**
  - Adding value to GenBank sequences through enhancing and updating genome annotations
  - Linking to research reviews/papers for the research community.
- **Collaborate with researchers to help on specific bioinformatics problems, e.g.**
  - Custom searches of the databases
  - Building new features into our tools
  - Help with genome annotation

Protocol

Step 1.

Let's assume that you have already aligned <u>two genome sequences</u> and wish to compare these two strains gene by gene.

To do this, you will first need to import the gene annotations for each virus from the corresponding Genbank files.

Step 2.

First, select the sequence you want to import the genes into.

Select **"Import genes"** from the *Tools* menu, and click on **"From feature file"**.

Step 3.

Find and select the appropriate Genbank file on your hard drive.

Now input the gene prefix you desire to use for naming genes (e.g. genome name) and click **"OK"**.

Repeat the same steps for the other sequence(s).

Step 4.

Finally, to obtain a statistical comparison of all the genes in the two strains, choose **"CDC statistics"** from the *Reports* menu.