Course KB8007 Comparative Genomics

Practical 6: Function prediction

Goal: Compare and interpret the results of two different ortholog detection methods for one of your genes: (1) tree-based ortholg prediction (treeFAM) and (2) blast-based ortholog detection (InParanoid).

The report should be formatted in one .doc file and sent to oliver.frings@sbc.su.se before the end of the week and contain the following results. Missing or failed items will result in a reduced grade for this practical.

1. Summary of what you have done (e.g. how did you find protein identifiers etc.).
2. Short description of the treeFam and InParanoid method.
3. Detailed discussion of the results achieved with the two methods.

Procedures:

- For this practical you will be using the two different ortholog detection methods treeFam and Inparanoid.

- Start by finding the TreeFam and InParanoid websites and peruse their FAQ's. For a better understanding of the methods you should also have a look at the reference articles (If you have problems to obtain the articles they can be provided by me).

- You want to compare orthology for one of your predicted genes. Since you want a gene present in both treeFam and InParanoid you should restrict your gene selection to a species that is present in both databases.

- To search for orthologs you first need correct protein identifiers:
  - One way to find correct identifiers is to do a local blast search with the sequence of your protein against the source files of the Inparanoid database. The source files can be found in: http://inparanoid.sbc.su.se/download/current/sequences/processed/
  - You can also do an online blast search (e.g. at ncbi or ensembl).
  - TreeFam can take a variety of gene identifiers for searching for orthologs. To check which kind of identifiers are supported by InParanoid browse the files in the source folder or the website.

- Once you have the correct identifiers find orthologs for your desired gene using both TreeFam and Inparanoid and analyze and compare the results.