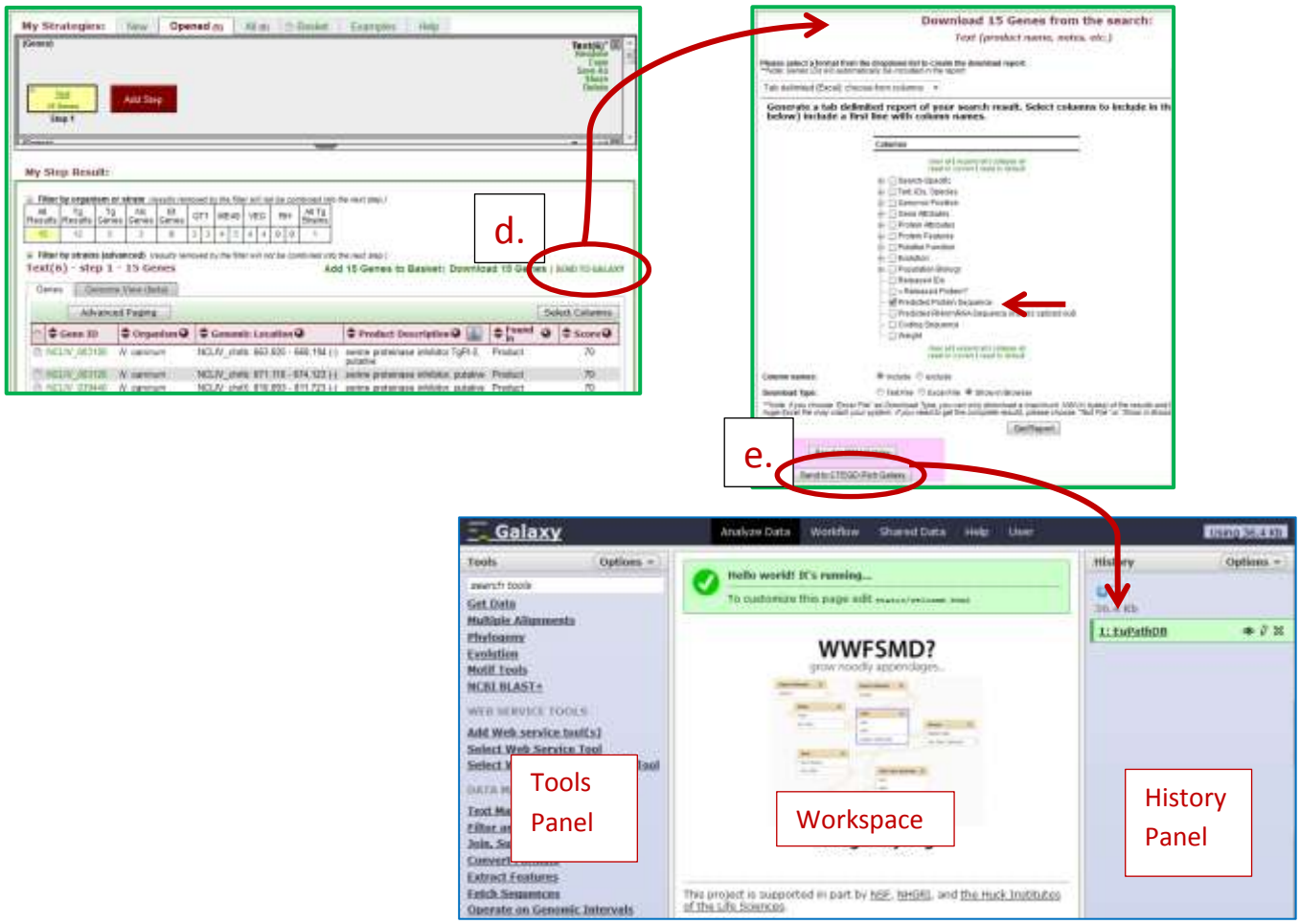
Multiple Sequence Alignment and Phylogenetic Tree
via Interfacing with Galaxy

Galaxy is an open, web-based platform for data intensive biomedical research. It can be accessed at a free public server or implemented in your own instance. The advantage of Galaxy is the ability to link output from very different data analysis tools into workflows that can be run repeatedly on a series of data sets. In this exercise we will use an existing workflow to create a phylogenetic tree of proteinases from *Eimeria tenella*, *Neospora caninum*, and *Toxoplasma gondii*.

1. Send a set of IDs and their corresponding protein sequences from ToxoDB to Galaxy.

a. Go to http://beta.toxodb.org
b. If you are logged in, please log out.
c. **Find all genes whose protein product descriptions contain the word proteinase**.
d. From the results page, click "SEND TO GALAXY" and arrange the download page to send the Predicted Protein Sequences.
e. Choose "Send to CTEGD Rich Galaxy".
   - The Galaxy page will load and your file will be listed in the right side History panel.
   - Take a look at the Tools panel. There are several expandable sections that contain data analysis tools.



2. Import and run a workflow that will create a phylogenetic tree of the protein sequences you just downloaded.

a. Create an account on Galaxy.
   - Click “User”, then “Register”.

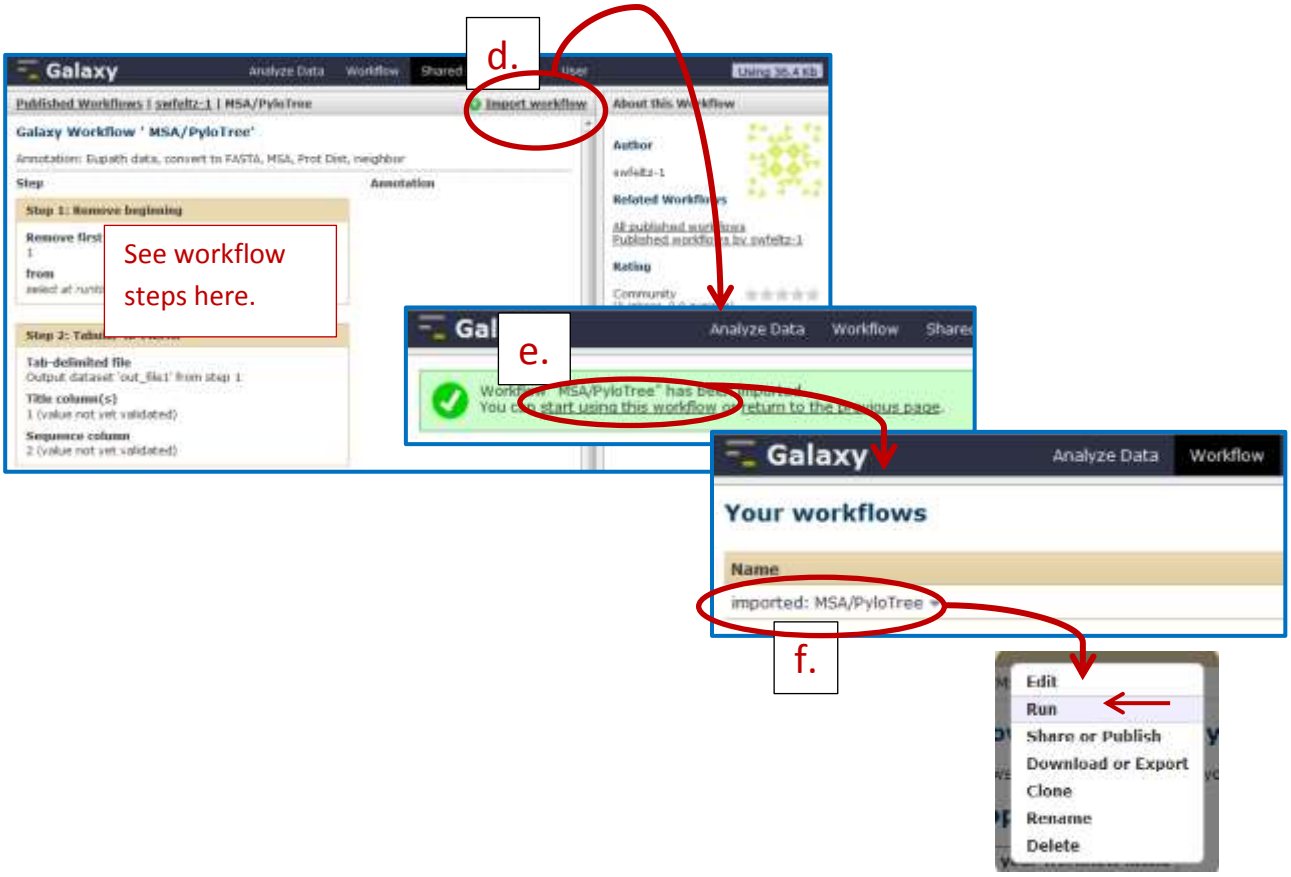
b. From within your account, go to “Shared Data” and then to “Published Workflows”. Click on “MSA/PyloTree”.

c. You will be able to view the steps of the workflow which should include:
   - Remove beginning
   - Tabular to Fasta
   - ClustalW
   - Protdist
   - Neighbor

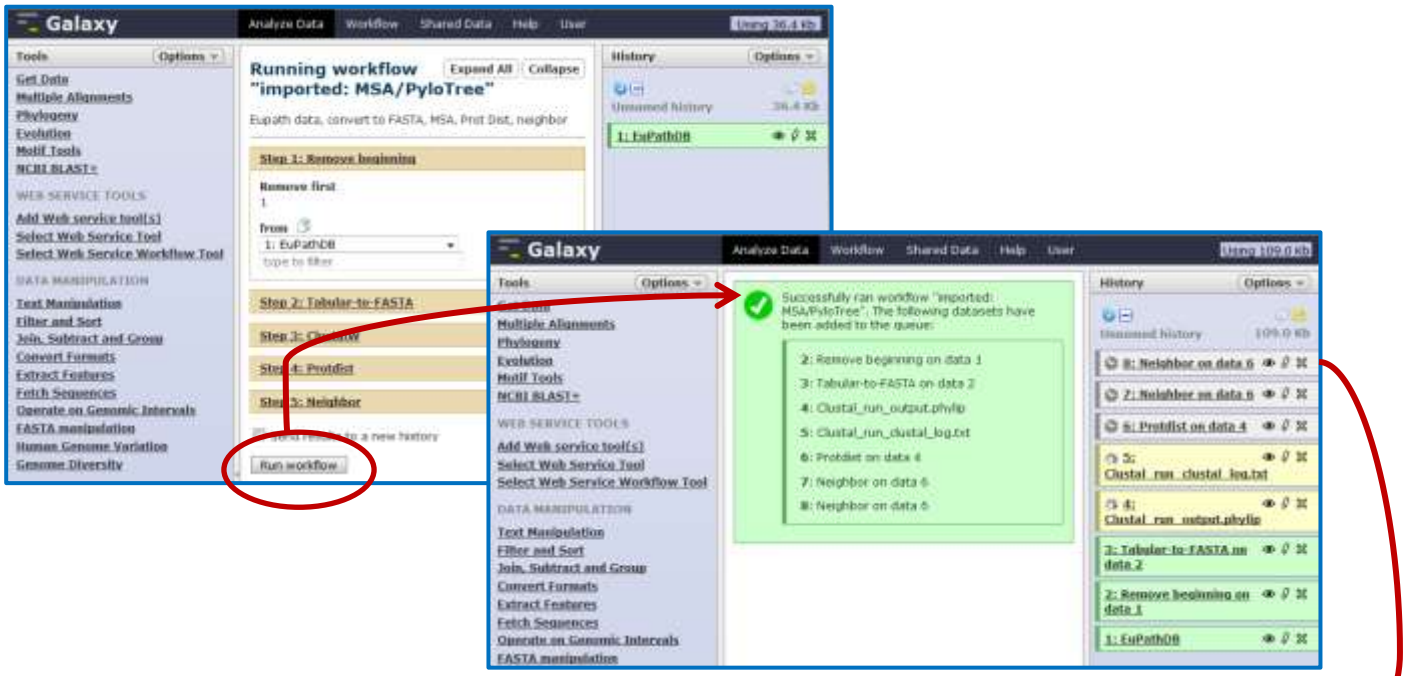
d. Click on “Import workflow”.

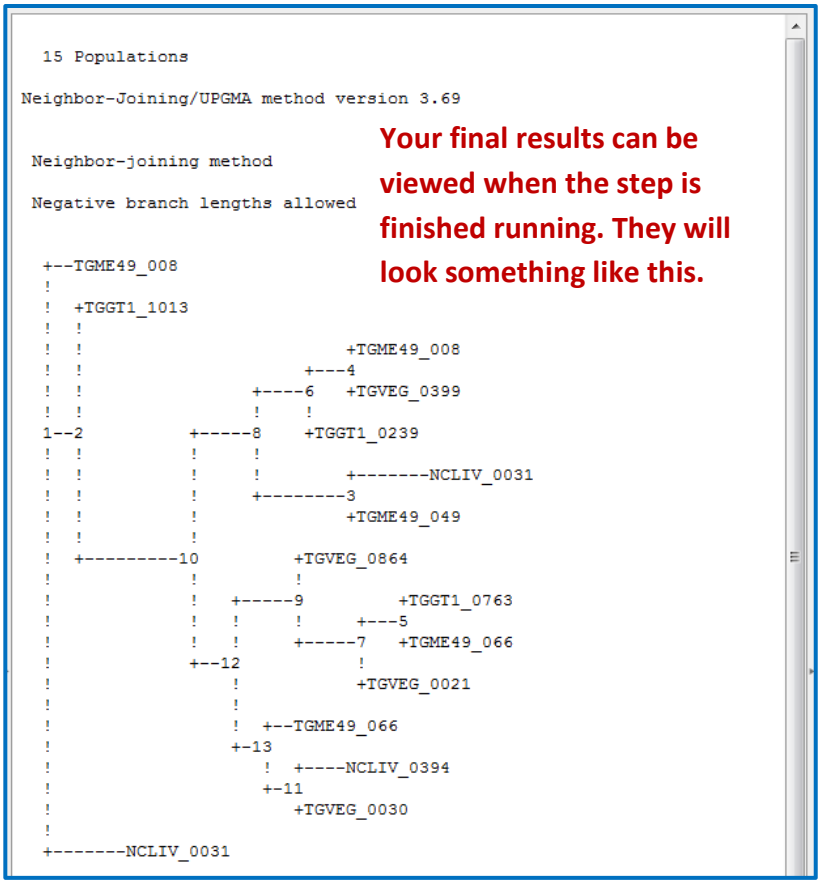
e. Choose “Start using this workflow”.

f. Click on the workflow name to reach the “Run” option.



g. Run the workflow.
   - After the workflow is initiated the steps will appear in the history pane.
   - As the processes finish running the file turns green.
   - Click on the name of the file to view a brief summary of contents.
   - Click on the eye to view the file contents in the work space.





Your final results can be viewed when the step is finished running. They will look something like this.