

WELCOME TO TREE PRUNER

What does the Tree Pruner do?

The Tree Pruner allows the user to remove unwanted sequences from a *segment* working set. Essentially, the user *works with a tree* in order to *select a data set*.

At the beginning of a Tree Pruner session, a phylogenetic tree is built from sequences in the segment working set. The user then acts on an interactive tree display to keep/remove sequences from the working set. The end result of a Tree Pruner session is a refined working set.

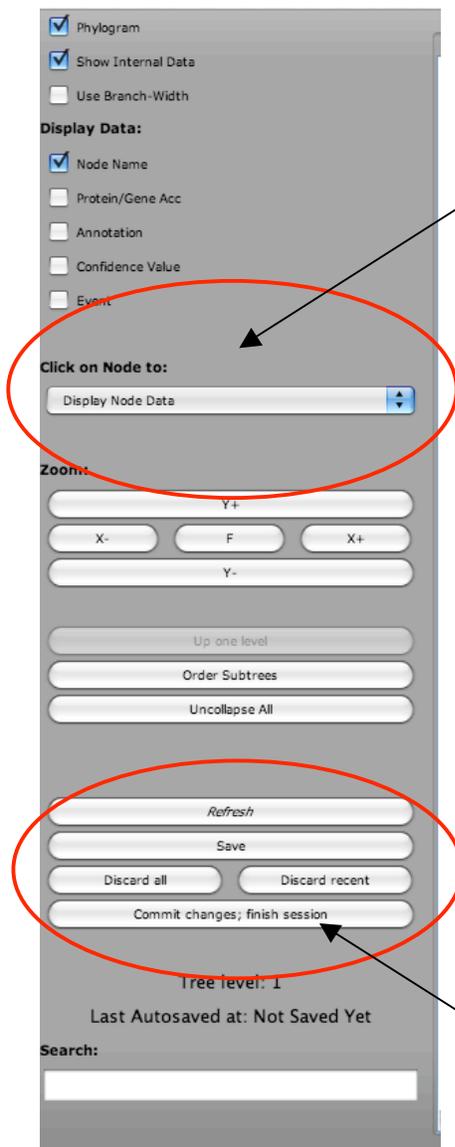
The purpose of this feature is to enable selection of an evolutionarily representative set of sequences from a large dataset. For example, if some viruses of a large dataset are very closely related while others are distantly related, the user may wish to thin the closely related viruses to even out representation in the dataset (and in the phylogeny). This type of thinning is not practical or reliable from, say, viewing an alignment or selecting on the basis of geographical and/or time representation.

How does the Tree Pruner work?

The Tree Pruner starts by inferring a tree from the user's sequence data ¹. The tree is then automatically displayed in a new, interactive window. (This interactive window is designed using tree visualization software called Archaeopteryx ².) The tab at the top of the window gives the name of your working set and the segment that you are operating on.

To interact with the tree, you first select an action from a pull-down menu labeled: "**Click on Node to:**" The first five actions in this menu are basic to Archaeopteryx; the last two options (**Keep** and **Remove/Restore**) are custom actions developed for Tree Pruner. You then click on a node in the tree display to carry out this action.

Tree Pruner functions - accessed from “Click on Node to:” menu



- “*Keep*”: you can only “keep” a virus at the end of a black or blue branch. When you click on a node to keep a virus, its branch will change to heavy black; all other black branches will change to blue.
- “*Remove*”: you can only “remove” a virus at the end of a black branch. When you click on a node to remove a virus, its branch will change to grey; all other branches will be unchanged.
- “*Restore*”: you can only “restore” a virus at the end of a grey branch. When you click on a node to restore a virus (i.e., return to the working set a virus that had previously been marked for removal), its branch will change from grey to black; all other branches will be unchanged.
- Keep and remove/restore actions must be carried out in *non-overlapping* sessions separated by *save*. For example, you can start by clicking on a number of viruses for keeping in the working set; you must *save* this work before you can switch to the remove/restore function.
- You can click on a *node of a subtree* and the action (keep/remove) will apply to all viruses in the subtree.
- At the end of a “keep” session, all blue branches will be drawn in grey.
- At the end of all Tree Pruner “keep/remove/restore” actions, all viruses at the end of grey branches will be removed from your dataset.

Managing a Tree Pruner session - accessed from custom buttons

- **SAVE**: Saves a record of your work to date; no changes are made to your working set.
- **DISCARD RECENT**: Discards all changes since your most recent SAVE.
- **DISCARD ALL**: Discards all changes since you started Tree Pruner; restores the original tree display.
- **COMMIT CHANGES; FINISH SESSION**: Your pruning actions are committed to your working set (that is, all sequences for viruses at the end of grey branches are removed from your working set). This is the data-filtering step. The tree display is automatically closed.
- To leave tree pruning, you can do one of the following:

- Close the applet by clicking on its quit button. This exit will require that you later return to your session to either discard or commit your changes.
- Use the COMMIT CHANGES; FINISH SESSION button. If you don't want to prune your dataset, you can discard all changes before committing and finishing the session.

Other details

- Tree Pruner automatically saves your work every ten minutes. If your session is interrupted for any reason, you can restore your session up to the most recent SAVE or AUTOSAVE. To restore a session, simply select the working set that you were pruning.

¹ Trees are inferred under the Hasegawa-Kishino-Yano (HKY) model of evolution, using PhyML to infer the parameters.

² Zmasek C.M. and Eddy S.R. (2001) ATV: display and manipulation of annotated phylogenetic trees. *Bioinformatics*, 17, 383-384.

Credits:

Mohan Krishnamoorthy, Margaret Green, Pragneshkumar Patel, Mira Bussod, and Catherine Macken

T-6, Los Alamos National Laboratory

Los Alamos, NM 87545

Version 1: October 2009