

## Getting Started with MultiSeq

### Requirements

MultiSeq must be correctly installed and configured before you can begin using it to analyze the evolution of protein structure. This section walks you through the process of doing so, but there are a few prerequisites that must be met before this section can be started:

- VMD 1.8.7 beta or later must be installed. The latest version of VMD can be obtained from <http://www.ks.uiuc.edu/Research/vmd/>
- This tutorial requires approximately 340 MB of free space on your local hard disk. MultiSeq requires about 500 MB of free space for metadata databases.

### Configuring MultiSeq

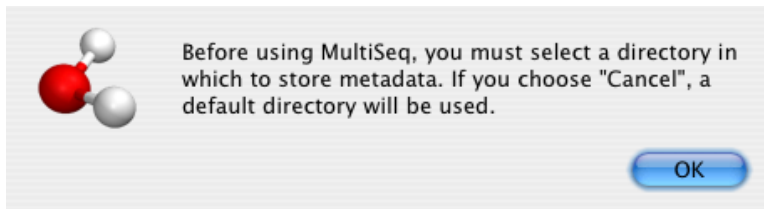
MultiSeq saves user preferences in a file named `.multiseqrc` located in your home directory. The preferences saved include the location of any local databases, previous search options, and others. When you start MultiSeq for the first time, it will ask you if it is ok to create this file and to specify the directory in which to look for any metadata databases.



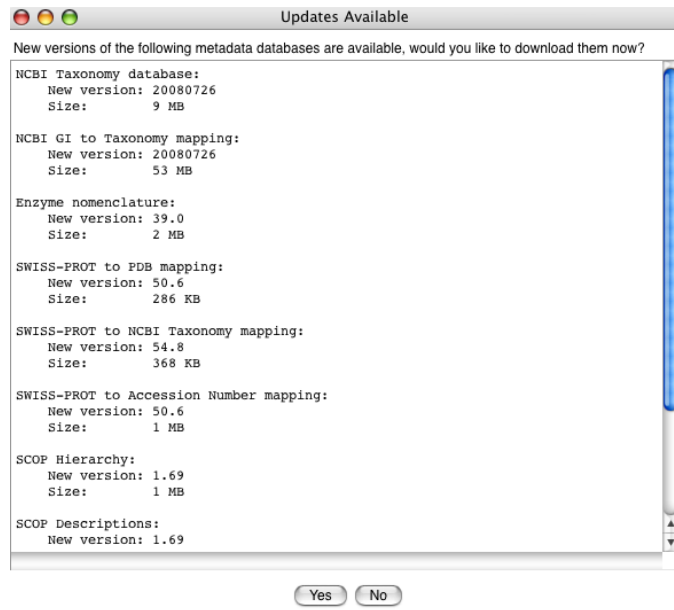
**What is metadata?** Metadata is a term meaning “data about data”. In MultiSeq the word metadata refers to information about the sequences or structures loaded into the program. MultiSeq knows how to find certain types of sequences or structures in the public metadata databases and can display information from them such as the species from which the protein originated, the taxonomic lineage of the organism, the protein’s enzymatic properties, and even how to find the protein in other databases. You’ll learn more about how this can be helpful later in the tutorial.

Follow these steps to configure MultiSeq:

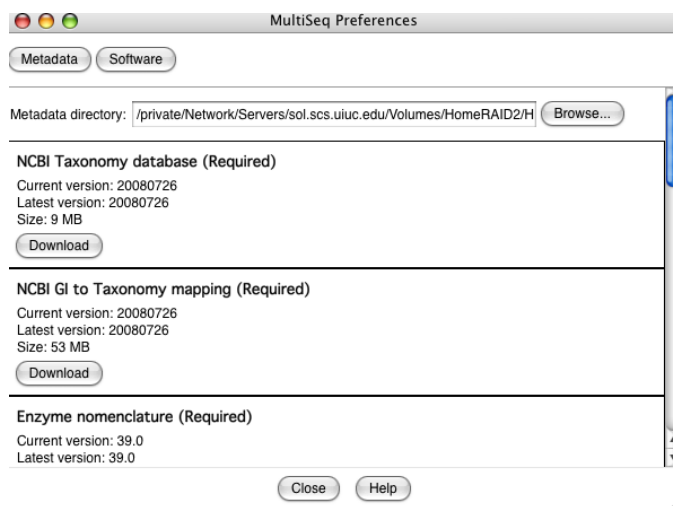
1. Launch VMD. Within the VMD main window, choose the Extensions menu, select Analysis → MultiSeq.
2. MultiSeq will notify you that you must select a directory in which to store metadata databases. Press the OK button.



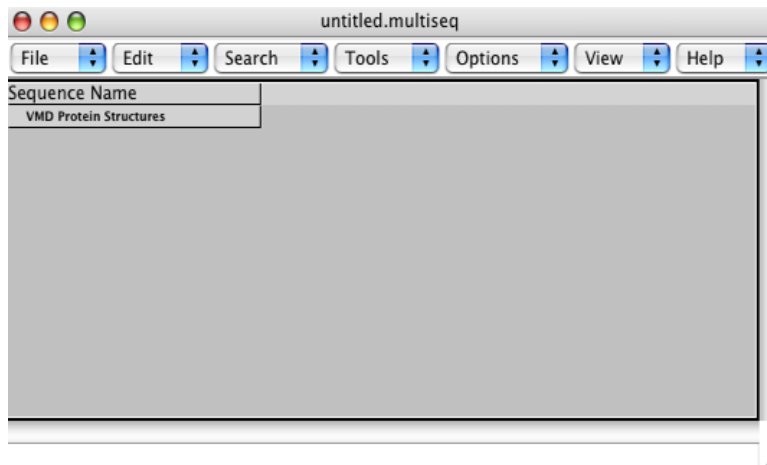
3. You will then be prompted to select the metadata directory. If the directory already contains the metadata databases, MultiSeq will use them. If not, MultiSeq will download them into the directory. Select the directory where you would like MultiSeq to store the databases and press the OK or Choose button.
4. If updates to the metadata databases are available, MultiSeq will present a dialog showing the available updates and give you the option of downloading them. Press the Yes button to download the updates. MultiSeq will ask you to wait while the updates are downloaded, which may take a few minutes depending on the size of the updates and the speed of your connection.



5. The MultiSeq Preferences dialog will then appear showing the metadata directory and the currently installed databases. Press the Close button to save these preferences.



The MultiSeq program window will then appear on the screen.



## Configuring BLAST for MultiSeq

MultiSeq is now minimally configured. For the purposes of this tutorial, however, some additional functionality is needed. Specifically, the tutorial uses BLAST to perform sequences searches, requiring that a local version of BLAST be installed.



### What is BLAST and why do I need to install it?

BLAST is a software tool available from the NCBI (<http://www.ncbi.nlm.nih.gov/BLAST/>) that allows you to search through a database of sequences and find those that are similar to a query sequence or profile of sequences. BLAST allows for very rapid searching through a large number of sequences and is widely used in the bioinformatics community. BLAST is typically used via one of two methods: online search or local installation. An online search is very simple and requires nothing more than for a user to paste a query sequence into a web page, but the utility of such a search is somewhat limited. MultiSeq requires a local BLAST installation because it provides additional functionality to the user not available through an online search.

Follow these steps to install a local copy of BLAST:

1. Create a directory on your local hard disk into which BLAST will be installed. Recommended directories are:
  - Unix/Linux: `/usr/local/blast`
  - Mac OS X: `/Applications/Blast`
  - Windows: `C:\Blast`

**Warning: For Windows users, please make sure that the path to your Blast contains absolutely no space!**

2. Archives of the BLAST installation for each of the supported platforms are located on the tutorial CD in the directory:

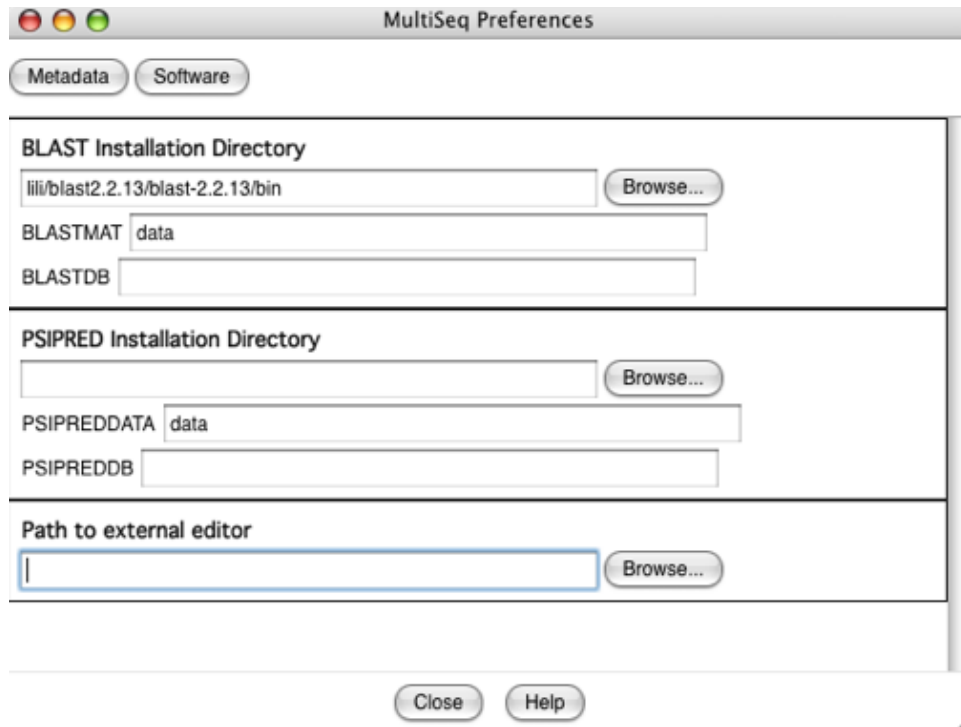
`/Tutorials/class-I/blast-install/`

or in the compressed file available for download from the tutorial website.

Copy the BLAST archive file corresponding to your platform into the directory created in the previous step.

3. Extract the archive. On Unix/Linux, use a command such as `tar zxvf filename`. On Mac OS X and Windows, the archive is a self-extracting executable, so just double-click on it.

- Next, you must set the BLAST installation location in MultiSeq. From the MultiSeq program window, choose File → Preferences... to bring up the preferences dialog.
- Click on the **Software** button in the upper left portion of the dialog to show the software preferences.
- Click on the **Browse...** button in the **BLAST Installation Directory** section and select the directory into which you installed BLAST. *Note: on Linux and Mac OS X you may have a directory called `blast-2.2.13` underneath your installation directory. If so, pick this directory in the browse dialog.*



- Press the Close button to save these changes. MultiSeq is now configured to use your local installation of BLAST.