Using GeneClust on the MD Anderson QS Compute Servers (2008)

Overview

To use the GeneClust that is installed on the compute servers at the MDACC Division of Quantitative Sciences, you must first establish a suitable X-Windows connection from the MDACC compute servers back to your local X-Windows display. GeneClust can then be used as described in the GeneClust documentation.

GeneClust is available for use from either mdabamcs01 or mdabamcs02. To access it, you must have an account on these servers. If you will be accessing from offsite, you will need VPN access, but a discussion of the steps required to establish a VPN connection is beyond the scope of this documentation.

1. Establishing an X-Windows Connection

Linux Users

Linux comes with all the required software[1]. To access the servers open a terminal window and enter the following command:

 % **ssh -Y mdabamcs01**

Macosx Users

Macosx also comes with all the required software. Like Linux users, you will open a local terminal window and then type the above ssh command. <u>However, you must open an X11 terminal, **not** the standard Macosx Terminal command.</u> The X11 command can be found in the Utilities folder.

Microsoft Windows Users

Microsoft Windows does not come with either X11 or ssh installed and neither is available from Microsoft.

One solution that would avoid installing software would be to download a Linux Live DVD (for example, Ubuntu), and boot from the DVD. You would be running Linux temporarily and could use the Linux method described above. Two disadvantages of this approach are that the downloads are very large and accessing Microsoft Windows files on your local hard drive at the same time you are running Linux could be relatively difficult.

A second solution is to download and install X-Windows and ssh software packages. These are available for Microsoft Windows as either proprietary or free software. The files you need to download are much smaller, but you will have to install the software onto your Microsoft Windows system. In this guide we will describe two high quality free software packages to meet this need: Xming (for X-Windows) and PuTTY (for ssh).

PuTTY is available from http://www.chiark.greenend.org.uk/~sgtatham/putty/. Download the executable

---

1 The openssh and X-Windows software packages must be installed. Typically, they are. If not, consult your Linux distribution's manual.

installer for Microsoft Windows and run it.

Xming is available from http://www.straightrunning.com/XmingNotes/. You need to download and install both the Xming and Xming-fonts packages. Do **not** start Xming until after you have installed Xming-fonts. Note that when you do start Xming, the only indication that it is running is not obvious: a small icon appears in the tray at the bottom right of the Microsoft Windows display.

Once PuTTY and Xming have been installed, start Xming and and PuTTY as described by David Mair at http://www.novell.com/coolsolutions/tip/18137.html. Make sure to check "Enable X11 forwarding" and set the X display location to localhost:0 as described in the above document.

2. Verifying Your X11 Connection

Once you have an X11 connection established, verify it is working by running the xeyes command:

```
% xeyes
```

A window should appear on your desktop containing a cartoon of two eyes. The eyes will look in the direction of your mouse pointer, and will continue to follow your mouse pointer as you move it.

Once you have verified **xeyes** is working correctly, use the xlsfonts command to list the fonts installed in the X-Windows server:

```
% xlsfonts
```

xlsfonts should output one line for each font installed on your X server. If xlfonts outputs only a few fontnames, your fonts are not installed properly or you started Xming before installing Xming-fonts. (In which quit and restart Xming.)

3. Running GeneClust

You can now run GeneClust as described in the documentation.

a. Create a directory for your GeneClust run and change into it.

```
% mkdir geneclustdir
% cd geneclustdir
```

b. Make the data and output subdirectories:

```
% mkdir data
% mkdir output
```

c. Copy your data as tsv files (and clf files if any) into the data subdirectory.

d. Execute the GeneClust executable and follow the GeneClust directions:

```
% ~bmbroom/mdabamcs01/bin/geneclust
```