MICB 405 Bioinformatics
Mini-Lab #4 – ClustalX

Dr. Joanne Fox
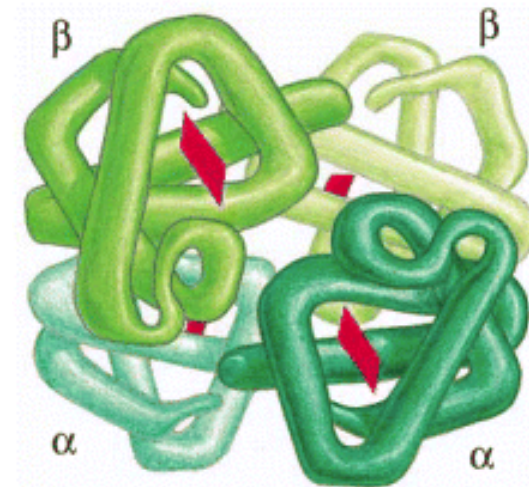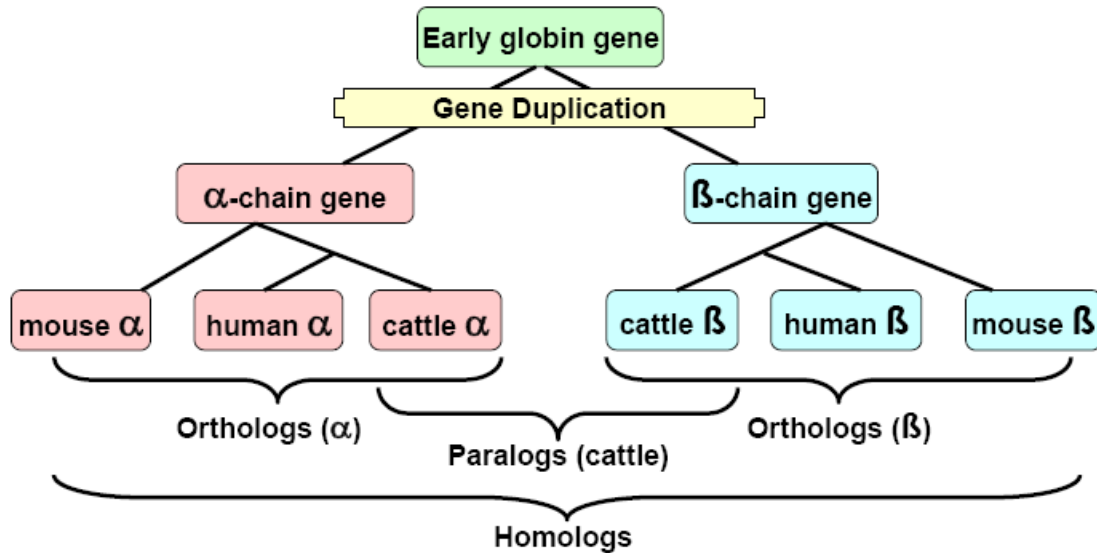
joanne@msl.ubc.ca

Objectives

- By the end of today's tutorial:
  - You will download and install ClustalX on your own computer.
  - You will use ClustalX to generate a multiple sequence alignment for a set of globin sequences.
  - You will view a phylogenetic tree generated from this set of globin sequences.
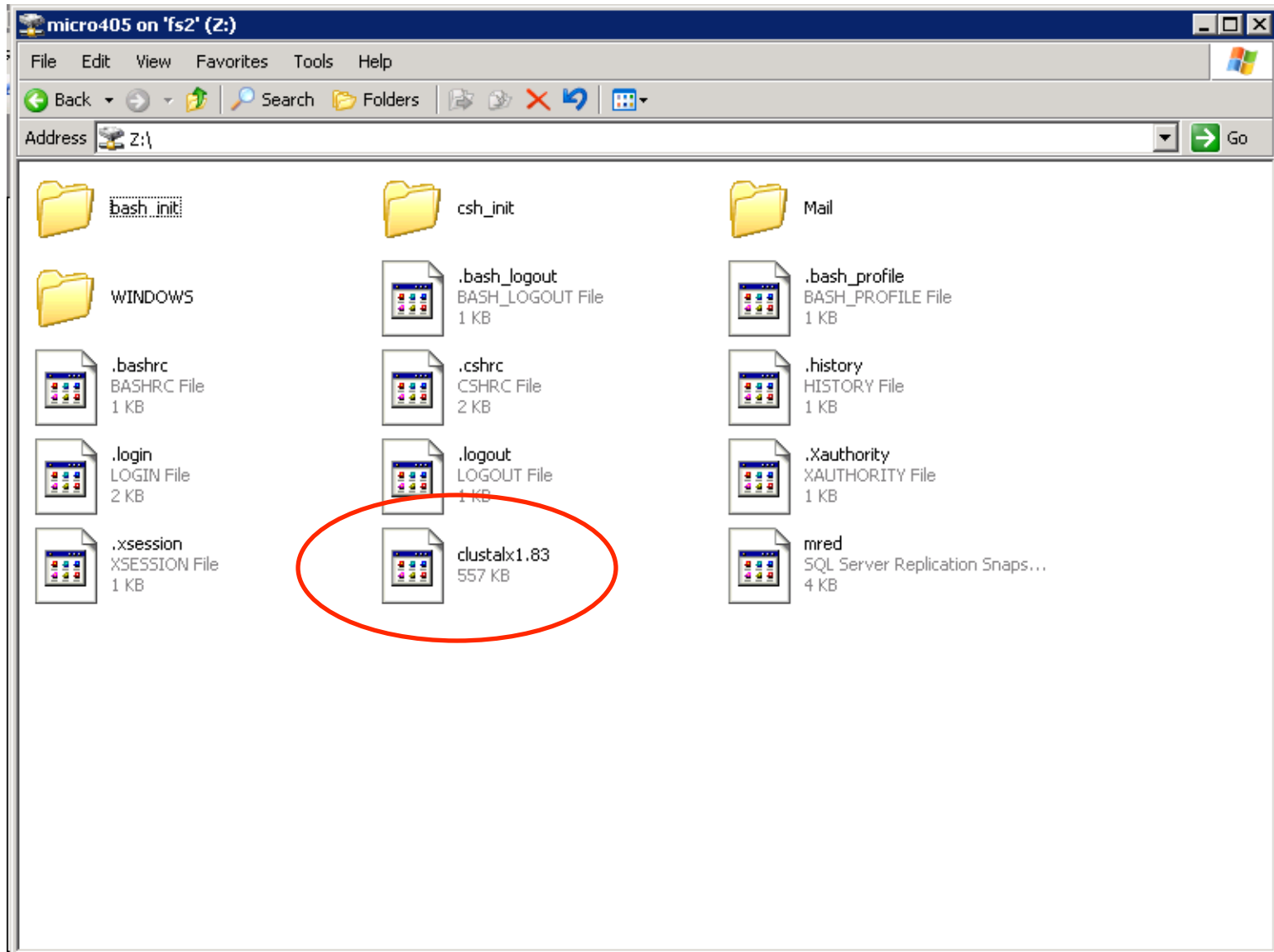
The Globin Genes

Hemoglobin is an abundant protein in red blood cells that contains two copies of α globin and two copies of β globin.
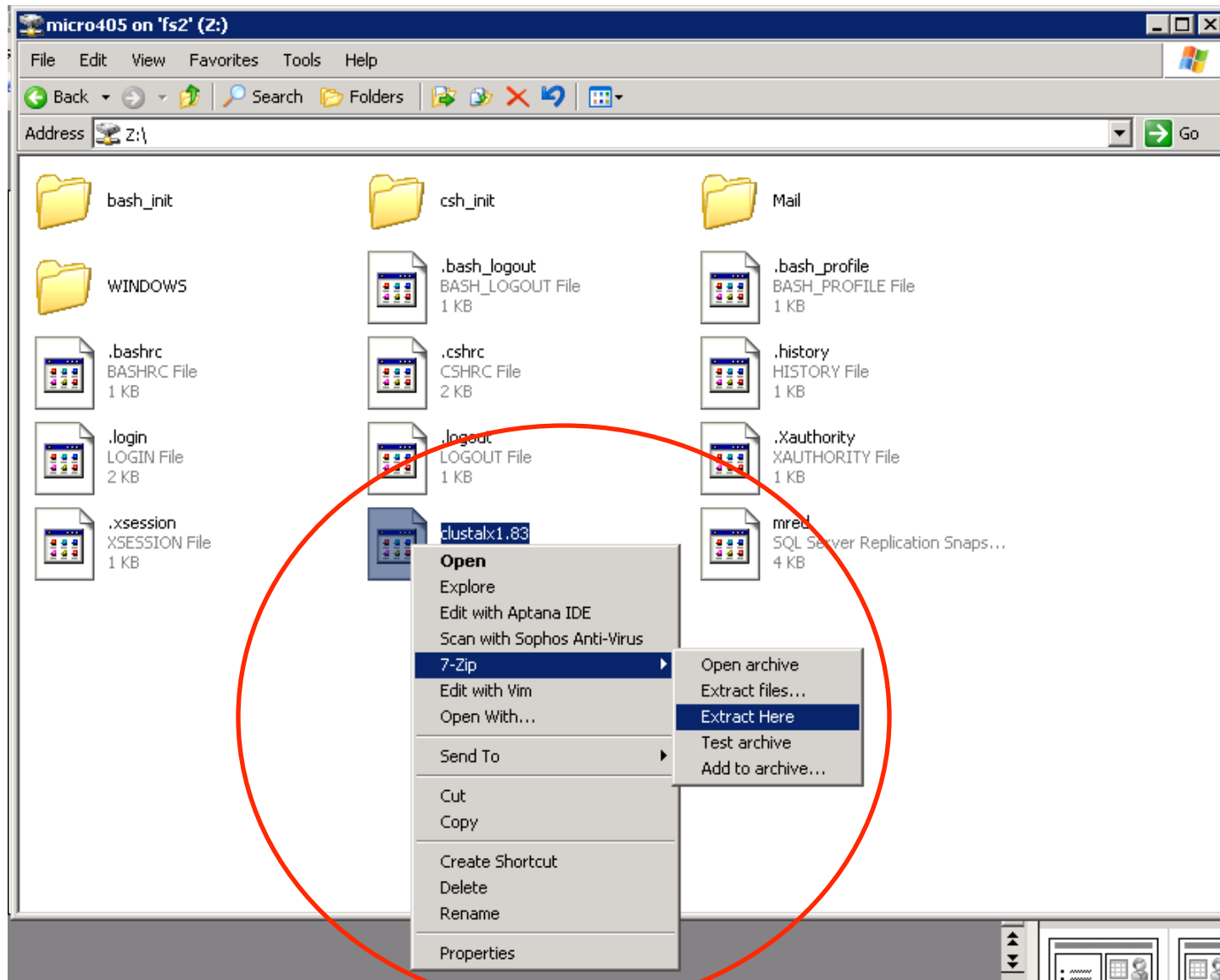
Download ClustalX v1.83

- Go to: ftp://ftp-igbmc.u-strasbg.fr/pub/ClustalX/

- Scroll down list and choose:
  - clustalx1.83.zip

- Download

- Extract all files to a fold where you can find it. ie. Desktop
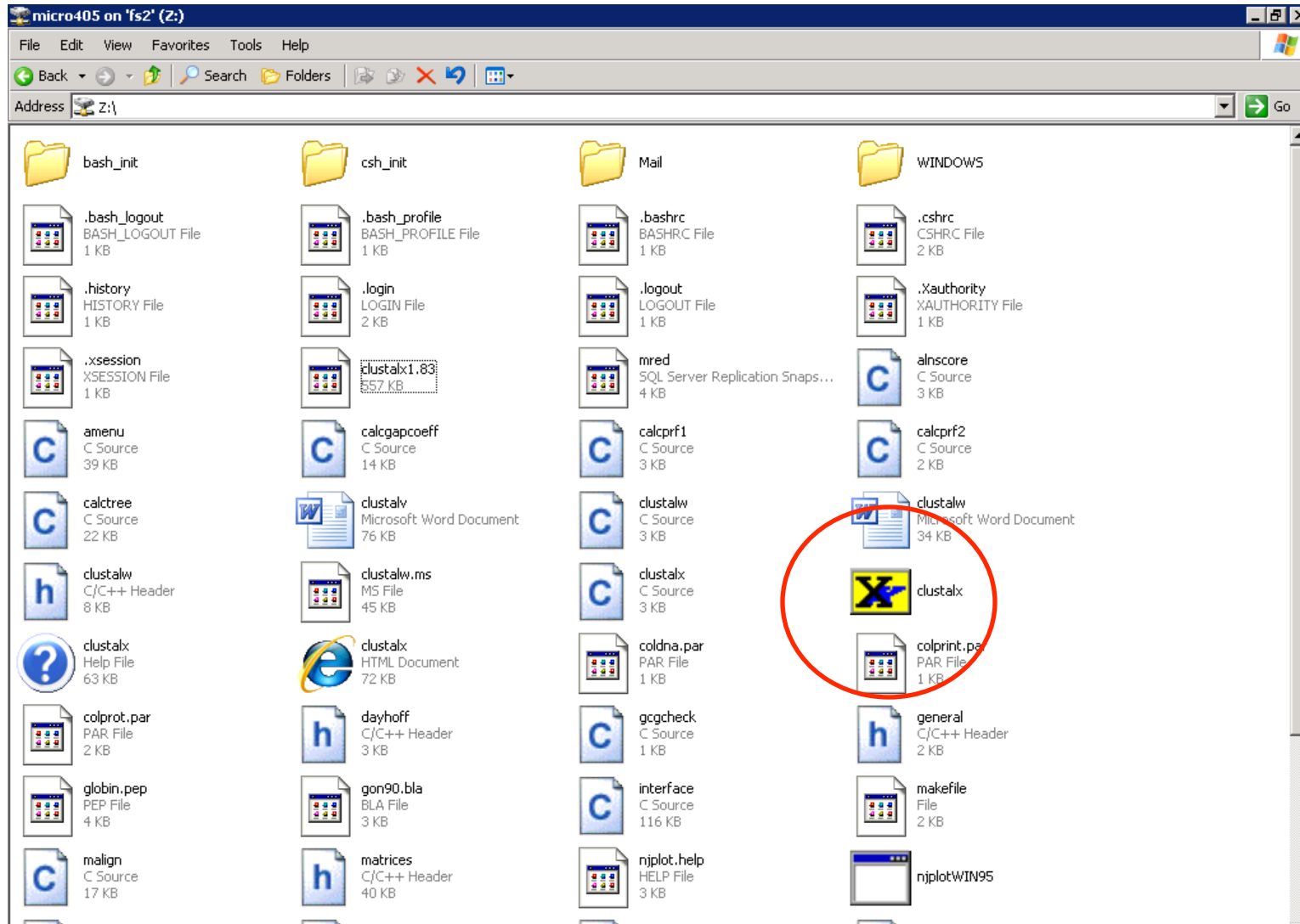
Find 'clustalx1.83' File

Right Click 'clustalx1.83' and 'Extract Here' with 7-Zip Option

Double Click 'clustalx' File

Clustalx.exe

7

Starting up ClustalX

File:
-Load sequences

Edit:
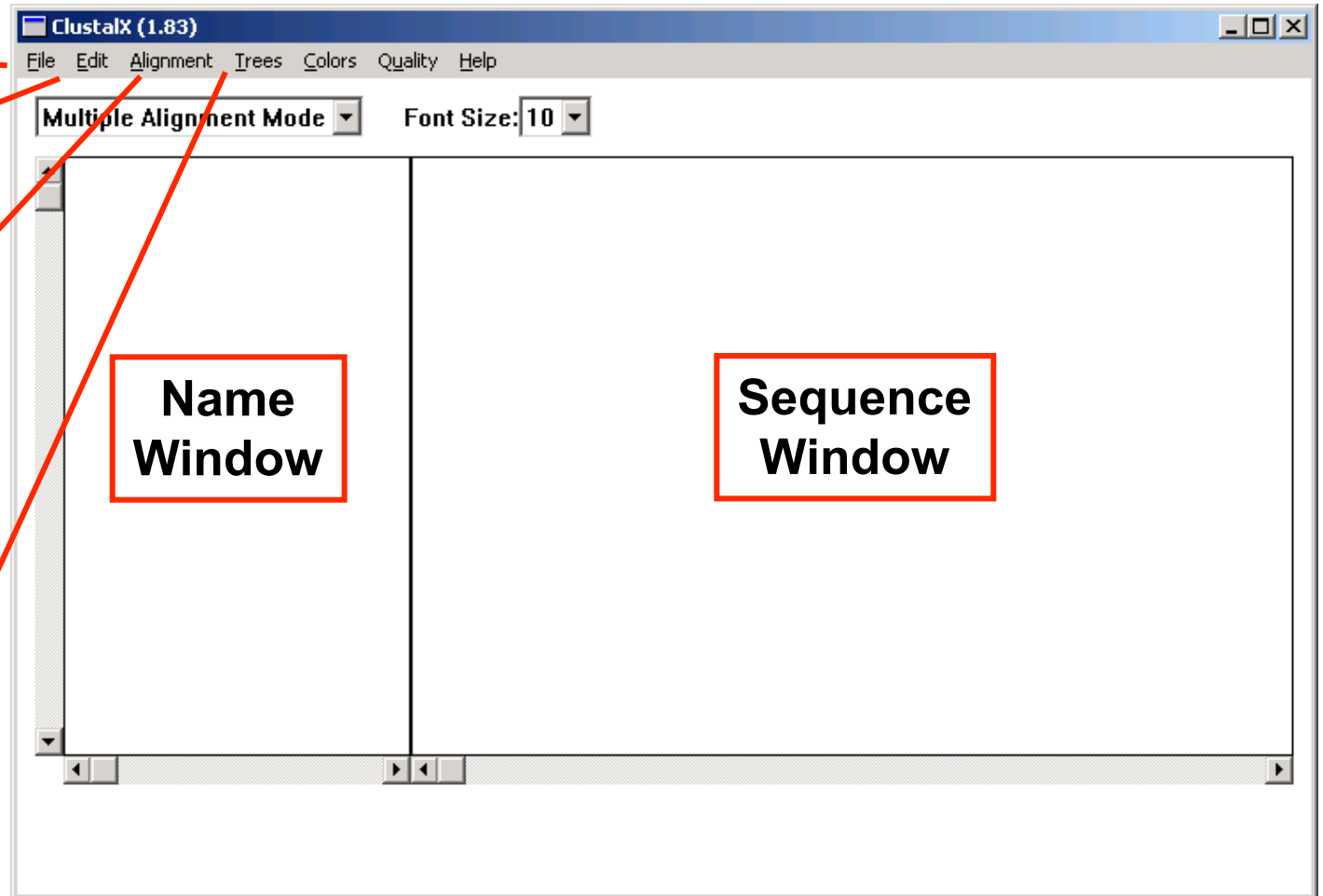-Remove all gaps

Alignment:
-Do complete
 alignment
-Alignment
 parameters

Trees:
-Bootstrapped NJ
-Output format
 options

ClustalX (1.83)

File   Edit   Alignment   Trees   Colors   Quality   Help

Multiple Alignment Mode   Font Size: 10

Name Window

Sequence Window

File   Edit   View   Insert   Format   Help   eFax

```
>P1;HBB_HUMAN
Sw:Hbb_Human => HBB_HUMAN
        VHLTPEEKSA VTALWGKVNV DEVGGEALGR LLVVYPWTQR FFESFGDLST
        PDAVMGNPKV KAHGKKVLGA FSDGLAHLDN LKGTFATLSE LHCDKLHVDP
        ENFRLLGNVL VCVLAHHFGK EFTPPVQAAY QKVVAGVANA LAHKYH*
C;ID    HBB_HUMAN       STANDARD;       PRT;    146 AA.
C;AC    PO2023;
C;DT    21-JUL-1986 (REL. 01, CREATED)
C;DT    21-JUL-1986 (REL. 01, LAST SEQUENCE UPDATE)
C;DT    01-APR-1993 (REL. 25, LAST ANNOTATION UPDATE)
C;DE    HEMOGLOBIN BETA CHAIN. . . .

>P1;HBB_HORSE
Sw:Hbb_Horse => HBB_HORSE
        VQLSGEEKAA VLALWDKVNE EEVGGEALGR LLVVYPWTQR FFDSFGDLSN
        PGAVMGNPKV KAHGKKVLHS FGEGVHHLDN LKGTFAALSE LHCDKLHVDP
        ENFRLLGNVL VVVLARHFGK DFTPELQASY QKVVAGVANA LAHKYH*
C;ID    HBB_HORSE       STANDARD;       PRT;    146 AA.
C;AC    PO2062;
C;DT    21-JUL-1986 (REL. 01, CREATED)
C;DT    21-JUL-1986 (REL. 01, LAST SEQUENCE UPDATE)
C;DT    01-MAR-1992 (REL. 21, LAST ANNOTATION UPDATE)
C;DE    HEMOGLOBIN BETA CHAIN. . . .

>P1;HBA_HUMAN
Sw:Hba_Human => HBA_HUMAN
        VLSPADKTNV KAAWGKVGAH AGEYGAEALE RMFLSFPTTK TYFPHFDLSH
        GSAQVKGHGK KVADALTNAV AHVDDMPNAL SALSDLHAHK LRVDPVNFKL
        LSHCLLVTLA AHLPAEFTPA VHASLDKFLA SVSTVLTSKY R*
C;ID    HBA_HUMAN       STANDARD;       PRT;    141 AA.
C;AC    PO1922;
C;DT    21-JUL-1986 (REL. 01, CREATED)
C;DT    21-JUL-1986 (REL. 01, LAST SEQUENCE UPDATE)
C;DT    01-FEB-1994 (REL. 28, LAST ANNOTATION UPDATE)
C;DE    HEMOGLOBIN ALPHA CHAIN. . . .

>P1;HBA_HORSE
Sw:Hba_Horse => HBA_HORSE
        VLSAADKTNV KAAWSKVGGH AGEYGAEALE RMFLGFPTTK TYFPHFDLSH
        GSAQVKAHGK KVGDALTLAV GHLDDLPGAL SNLSDLHAHK LRVDPVNFKL
        LSHCLLSTLA VHLPNDFTPA VHASLDKFLS SVSTVLTSKY R*
C;ID    HBA_HORSE       STANDARD;       PRT;    141 AA.
C;AC    PO1958;
C;DT    21-JUL-1986 (REL. 01, CREATED)
C;DT    21-JUL-1986 (REL. 01, LAST SEQUENCE UPDATE)
C;DT    01-MAR-1992 (REL. 21, LAST ANNOTATION UPDATE)
C;DE    HEMOGLOBIN ALPHA CHAINS (SLOW AND FAST). . . .
```
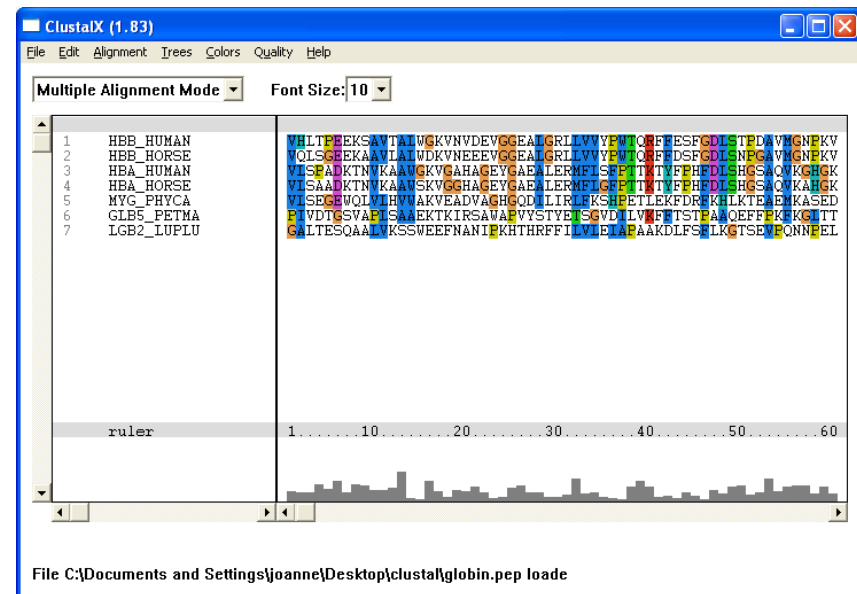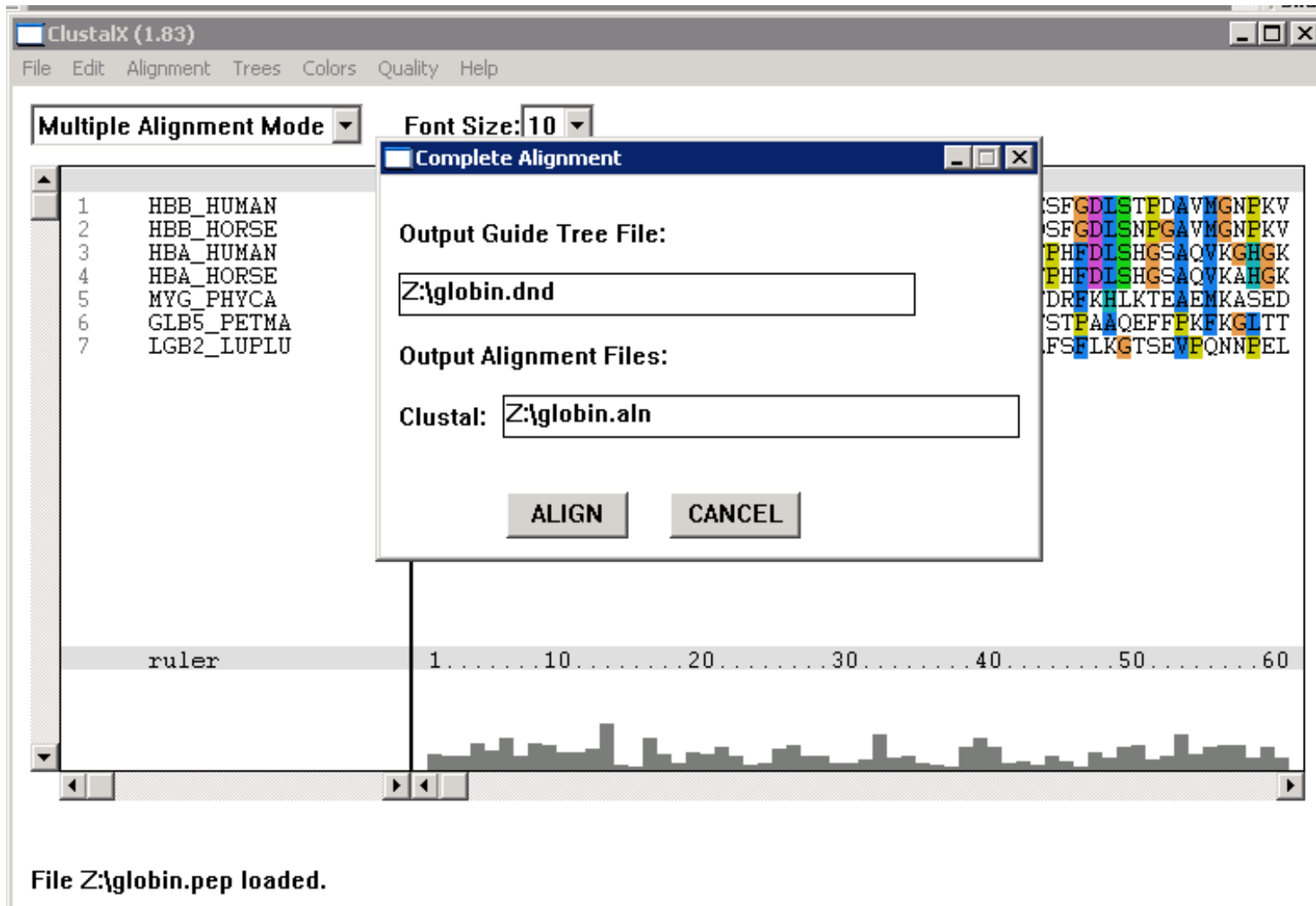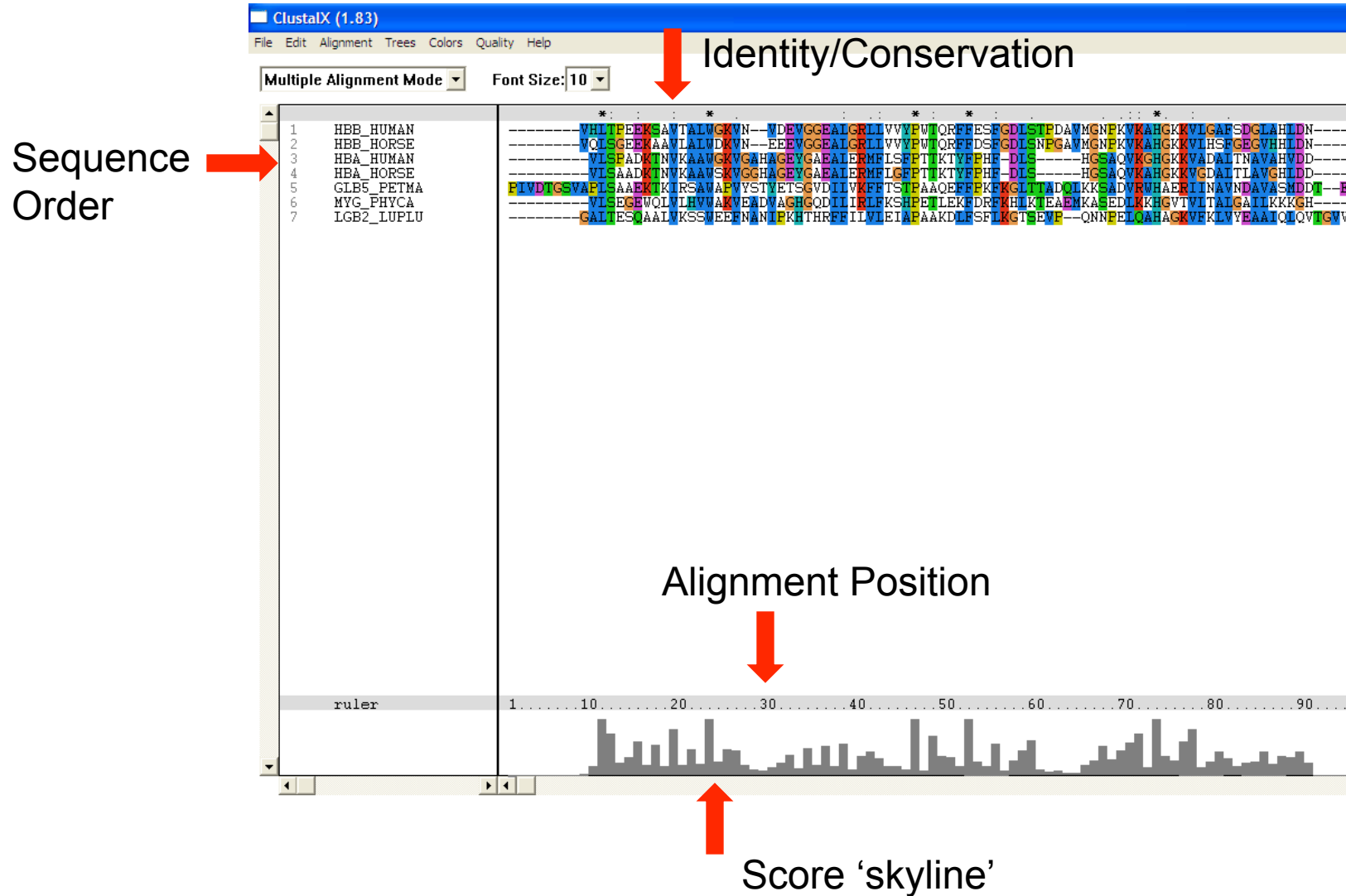
Load the sequences
– globin.pep

ClustalX (1.83)

File   Edit   Alignment   Trees   Colors   Quality   Help

Multiple Alignment Mode   Font Size: 10

```
1   HBB_HUMAN   VHLTPEEKSAVTALWGKVNVDEVGGEALGRLLVVYPWTQRFFESFGDLSTPDAVMGNPKV
2   HBB_HORSE   VQLSGEEKAAVLALWDKVNEEEVGGEALGRLLVVYPWTQRFFDSFGDLSNPGAVMGNPKV
3   HBA_HUMAN   VLSPADKTNVKAAWGKVGAHAGEYGAEALERMFLSFPTTKTYFPHFDLSHGSAQVKGHGK
4   HBA_HORSE   VLSAADKTNVKAAWSKVGGHAGEYGAEALERMFLGFPTTKTYFPHFDLSHGSAQVKAHGK
5   MYG_PHYCA   VLSEGEWQLVLHVVAKVEADVAGHGQDLLIRLKSHPETLEKFDRFKHLKTEAEMKASED
6   GLB5_PETMA  PIVDTGSVAPLSAAEKTKIRSAWAPVYSTYETSGVDILVKFFTSTPAAQEFFPKFKGLTT
7   LGB2_LUPLU  GALTESQAALVKSSWEEFNANIPKHTHRFFILVLEIAPAAKDLFSFLKGTSEVPQNNPEL

                ruler        1........10........20........30........40........50........60
```

File C:\Documents and Settings\joanne\Desktop\clustal\globin.pep loade

9

Alignment > Do Complete Alignment

Examine Alignment

Alignment Menu

Edit Menu

File Menu

Phylogenetic Tree Building

Trees Menu

Use NJPlot to visualize the trees

http://pbil.univ-lyon1.fr/software/njplot.html
Google: 'njplot'

NJplot

NEW: NJplot plots trees in PDF and PostScript formats (not for MacOS).
NEW: NJplot allows to open several tree windows.

NJplot is a tree drawing program able to draw any phylogenetic tree expressed in the Newick phylogenetic tree format (*e.g.*, the format used by the PHYLIP package). NJplot rooting the unrooted trees obtained from parsimony, distance or maximum likelihood tree-building methods.

A screen shot of the main window of njplot is available here.

Use of NJPlot

Any rooting of the unrooted tree can be interactively specified using the mouse. NJplot also allows zooming, branch swapping, display of bootstrap scores and printing in the Pl be used as a graphical extension of any package of phylogenetic programs which employs the standard tree format for storing trees (*i.e.*, with most such packages).
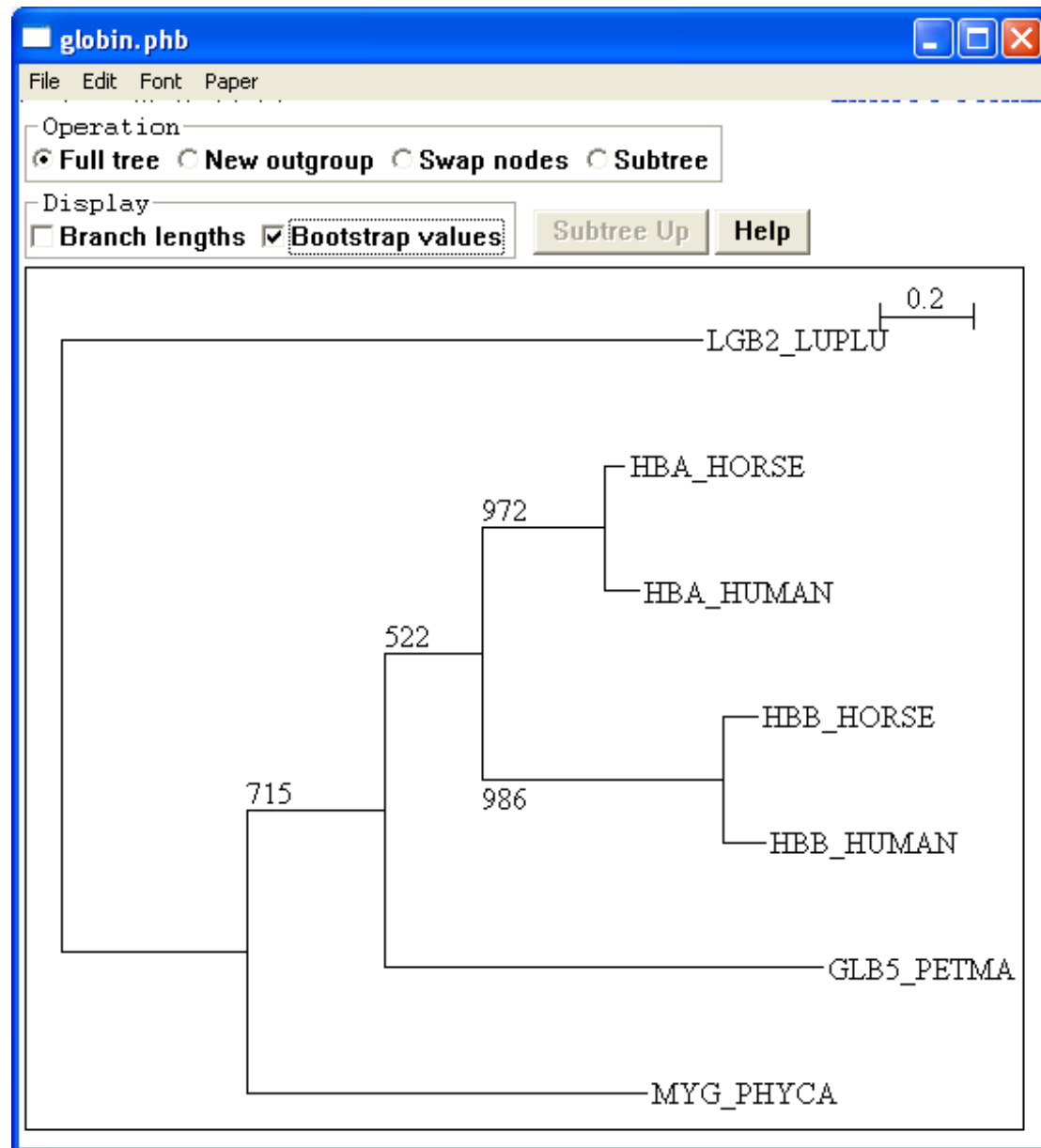
Download NJplot

Executables and full source code can be downloaded through our FTP server. You may also use the following table to directly access the version corresponding to your compu

MacOS X / MacOS (8,9)	Linux on PC in case of font problem	MS Windows	Sun-Solaris	HP Alpha

17

Extract njplotWIN95.exe
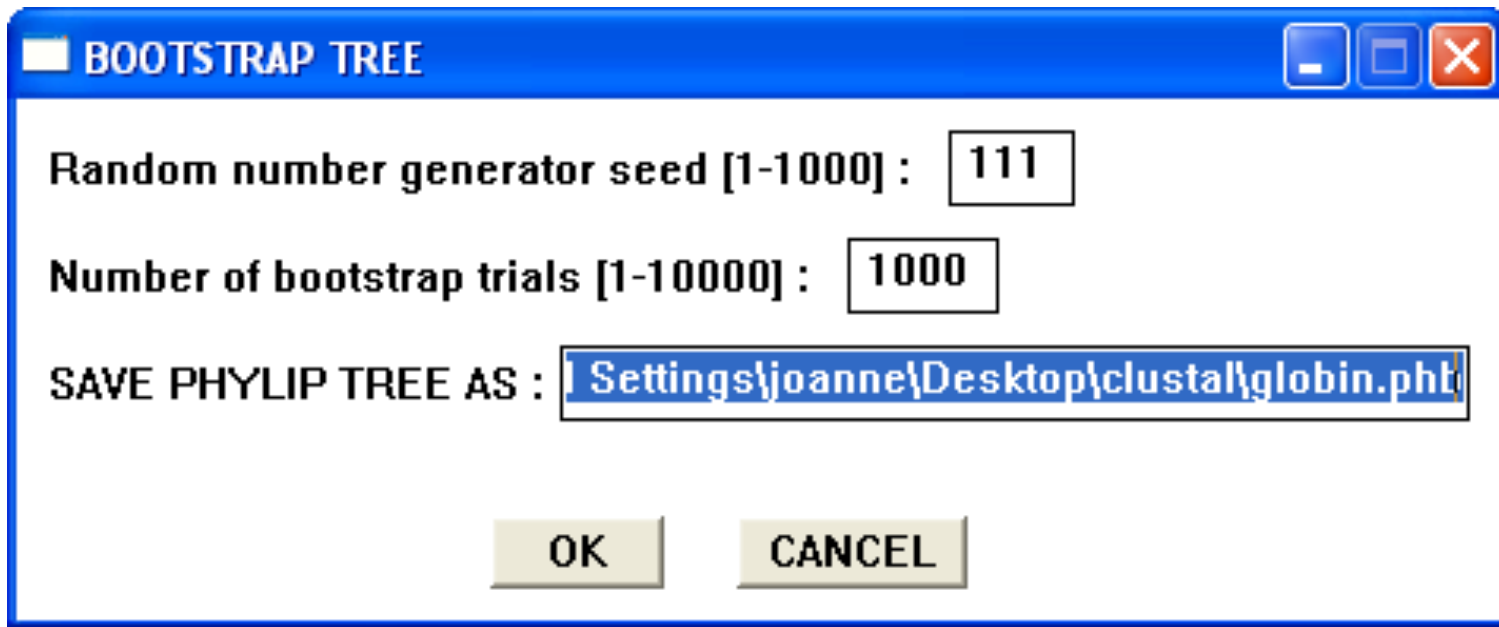# and run njplot.exe

self-extract
to folder

njplotWIN95.exe

njplot

njplot.exe

NJPlot

Trees > Bootstrap N-J Tree

1. Open tree file (.ph)
2. Open bootstrap file (.phb)
3. Can now display tree with both distances and/or bootstrap values.
4. Can save tree as **.pdf** for later use in poster.

Links

- ClustalX
  http://bips.u-strasbg.fr/en/Documentation/ClustalX/

- The web version: ClustalW
  http://www.ebi.ac.uk/clustalw/

- NJ-Plot
  http://pbil.univ-lyon1.fr/software/njplot.html

- Tree View (software for tree visualization)
  http://taxonomy.zoology.gla.ac.uk/rod/treeview.html