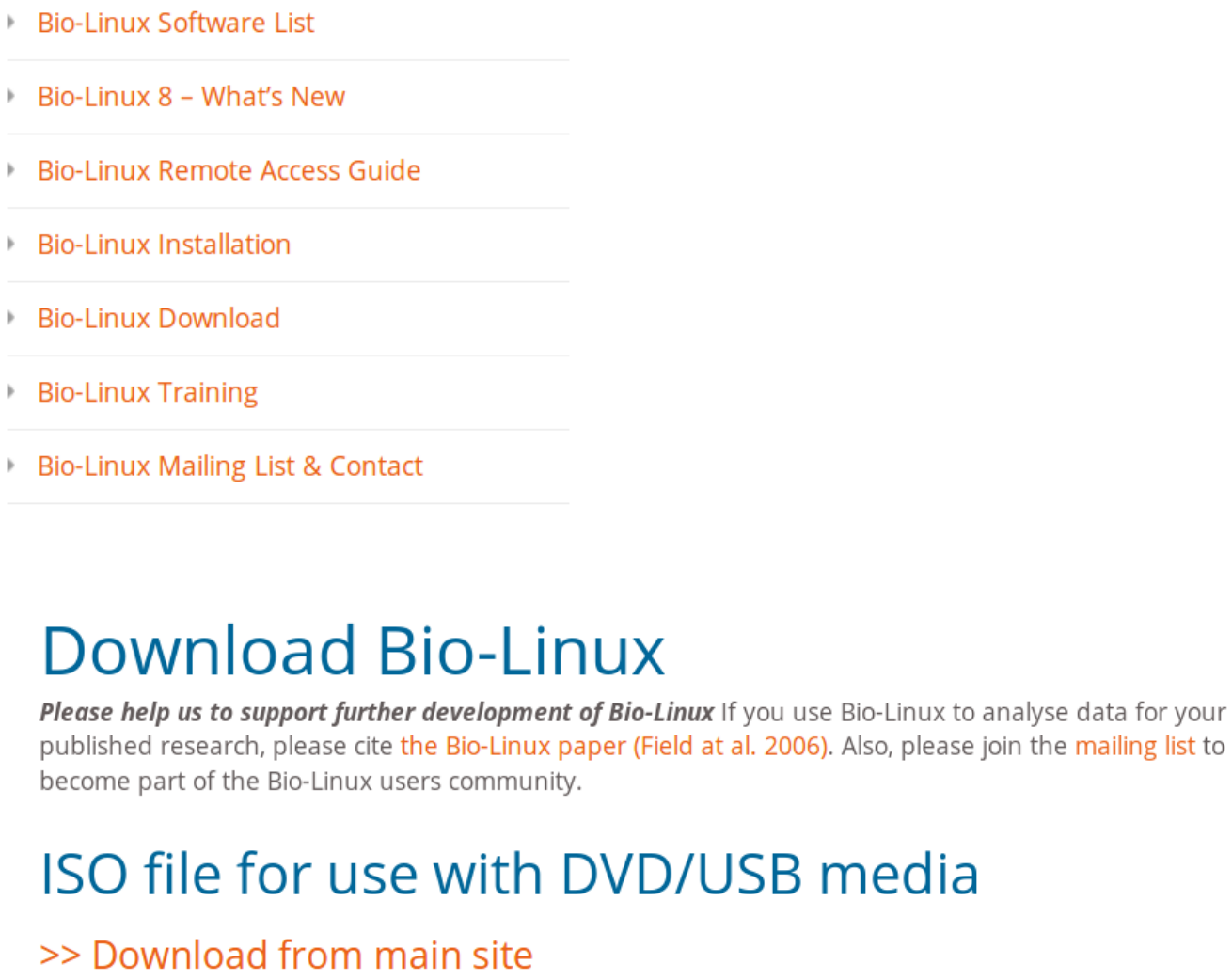
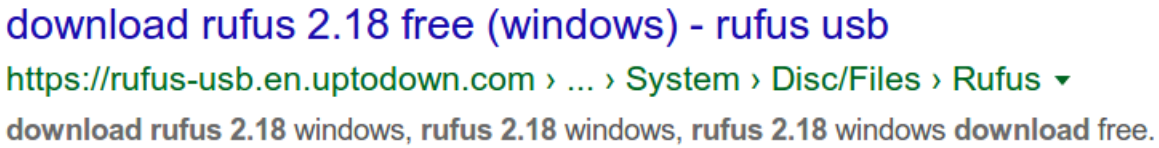
FIRST STEPS
BioLinux of flash Memory

Download Biol-Linux from

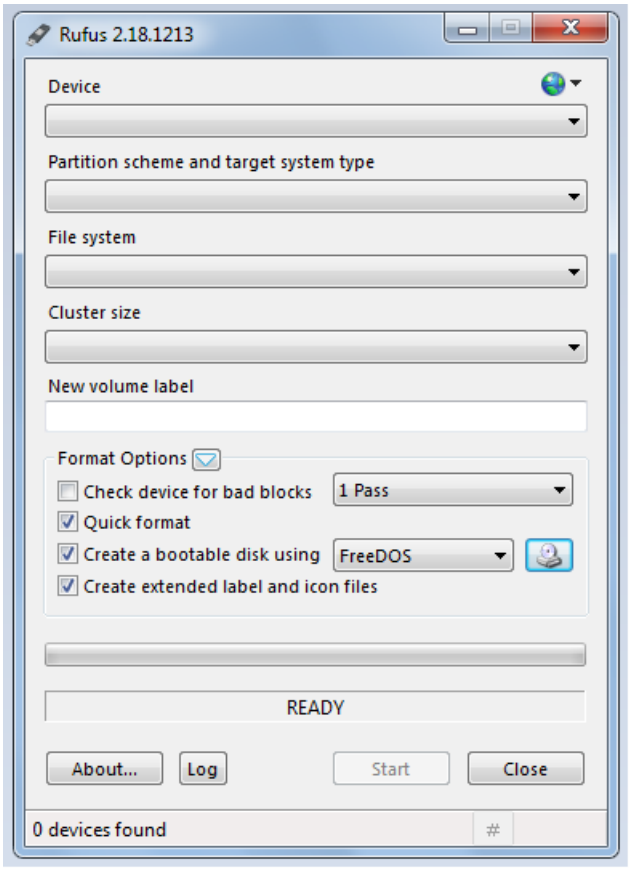
Download the ISO file from biolinux download



Download Rufus



Install Rufus on windows and open . Plug in **empty USB flash memory** to your laptop/PC :



Note : The USB flash memory should be empty because the previous process the clear all files stored on it.

From Device choose the USB name , press on the CD icon  the chose the LINUX ISO file which you have downloaded and **Press START.**

After Finishing of **rufus** restart your computer .

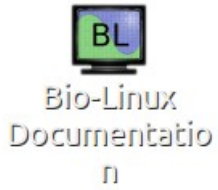
Try to boot from USB; this depend on your computer type but most PCs press F12 or F10 or F8 and select the USB drive. The following page will appear.

Choose Try Ubuntu without installing

Note : If boot error contains (gfxboot.c32:not a valid COM32R image) press TAB button on keyboard and write live and press ENTER. Such error could appear by bad rufus installation.

INSIDE BIOLINX

Press on this Icon on desktop:



USAGE: This is the documentation home page of BioLinux

Press on this icon :

USAGE: This is the documentation home page for BioLinux installed software which well open in your browser e.g, Google chrome.

About :

USAGE: The first option contains documentation for every software by name
The second option contains programs by category i.e. , Sequence Alignment, or protein analysis.

By Selecting :

Press View Category this would appear :

Alignment

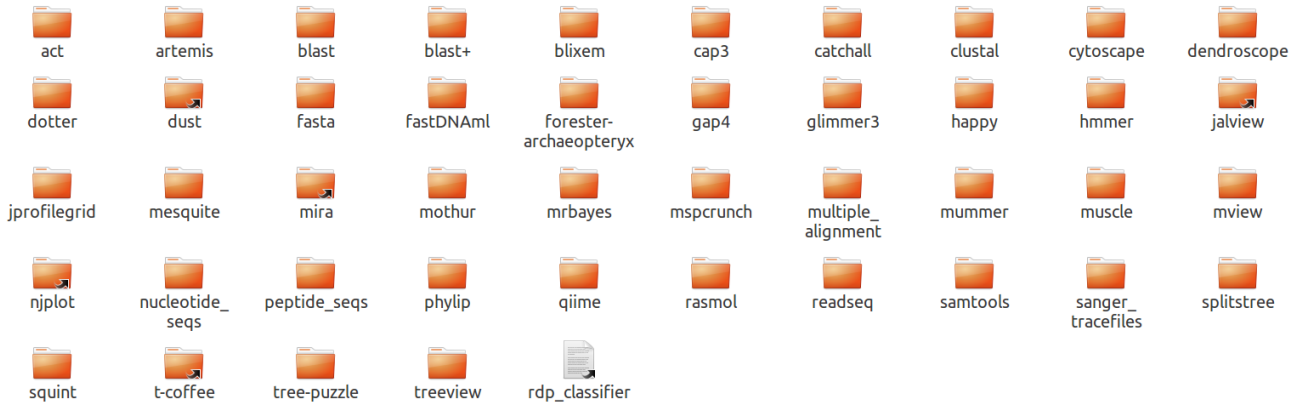
Category	Applications
Alignment	FastTree dialign
SubCategory	Applications
Consensus	cons consambig gap4 megamerger merger spin
Differences	act diffseq
Dot_plots	dotmatcher dotpath dotter dottup polydot
Editing	clcsequenceviewer jalview squint
Global	est2genome fasta ggsearch glsearch needle needleall stretcher swat
Graphical	blixem clcsequenceviewer clustalx dotter jalview jprofilegrid mummerplot nucmer2xfig seaview squint
Local	arb bl2seq bowtie bwa cross_match dba fasta fastf fastm fasts fastx fasty flalign glsearch lalign lfasta matcher plalign plfasta prenuc prfx prss psw pswdb seqmatchall sibsim4 sim4 ssearch supermatcher swat tfastf tfastm tfasts tfastx tfasty water wordfinder wordmatch
Multiple	annotate clustalw clustalx combineMUMs dialign edialign emma exact-tandems gaps hmmalign infoalign jalview lalign2list mapview mgaps mocca mummer mummerplot muscle nucmer nucmer2xfig plotcon postnuc postpro prank prepro prettyplot PROmer repeat-match run-mummer1 run-mummer3 seaview show-aligns show-coords show-tiling showalign squint t-coffee tranalign yamap
Profiles	jprofilegrid profit profit prophecy prophecy prophet prophet
Statistics	alistat pfscale prfx prss weight

USAGE: These are different tools for sequence alignment , these tools are already installed on BioLinux and you can use.

SAMPLE DATA

On Desktop you will find this folder :

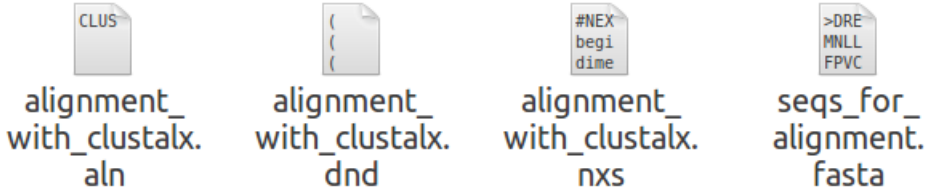
Open previous folder . The following window will appear :



USAGE: These folders contain different sample data for BioLinux tools as input.

Open the next folder:

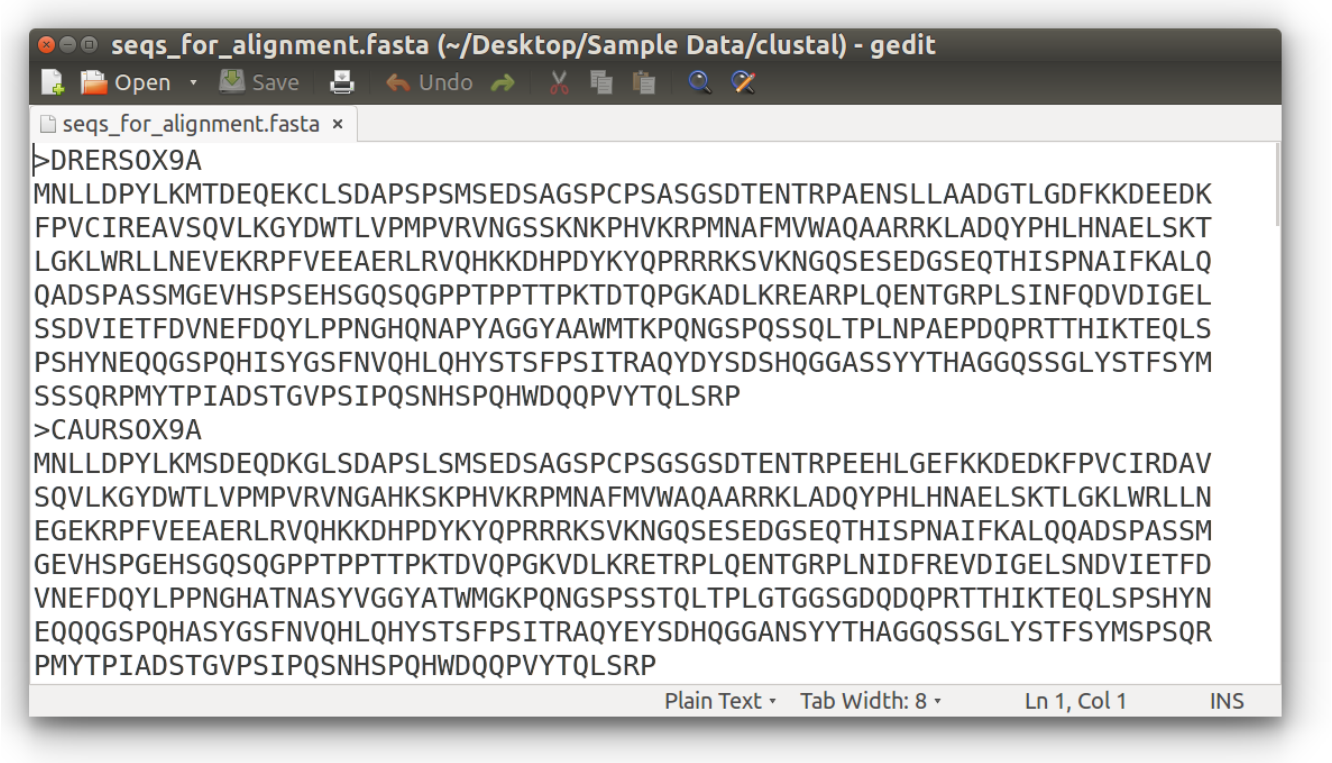
This window will appear :



USAGE: These files are output and input files for tool called “clustalw”. The first three files are the input and the last file is the output.

Press on this file :

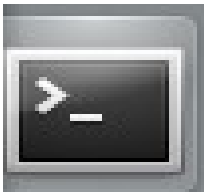
This window will appear:



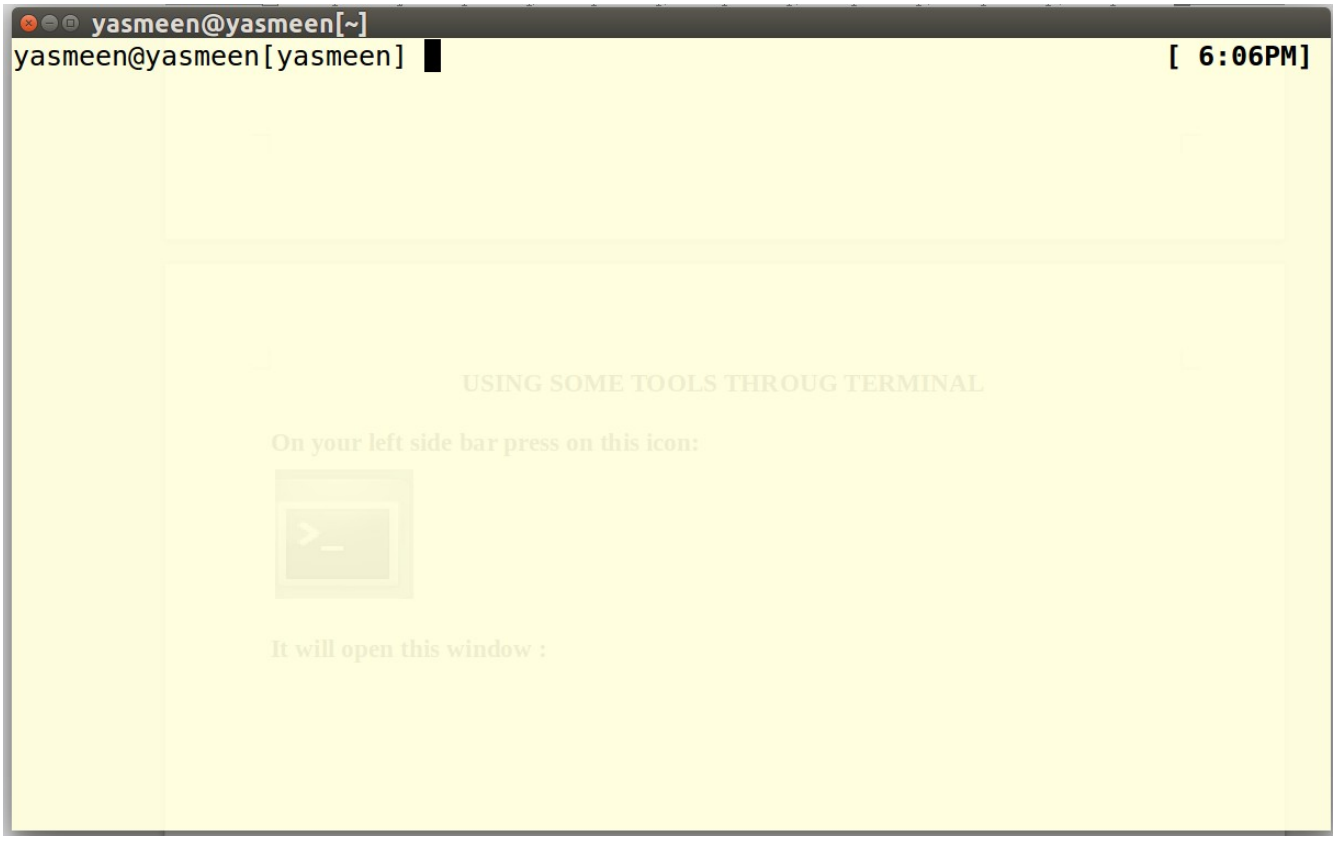
USAGE: These are amino acid sequences in FASTA file format and could be used for crustalw input.

SIMPLE DATA ANALYSIS

On your left side bar press on this icon:



It will open this window :



USAGE: This is the terminal of Linux, its the connection between you and your system which you can use to call any tool.

In previous window write :

echo hello

```
yasmeen@yasmeen[yasmeen] echo hello
```

Press Enter and the result would be:

```
yasmeen@yasmeen[yasmeen] echo hello
hello
```

USAGE : “echo” is a tool in Linux to print text on the terminal end.

In previous window write :

mkdir Desktop/bioinfo

```
yasmeen@yasmeen[yasmeen] mkdir Desktop/bioinfo
```

USAGE : “mkdir” is a tool in Linux to create a folder.

Minimize terminal and look at Desktop you well find an empty folder called bioinfo:

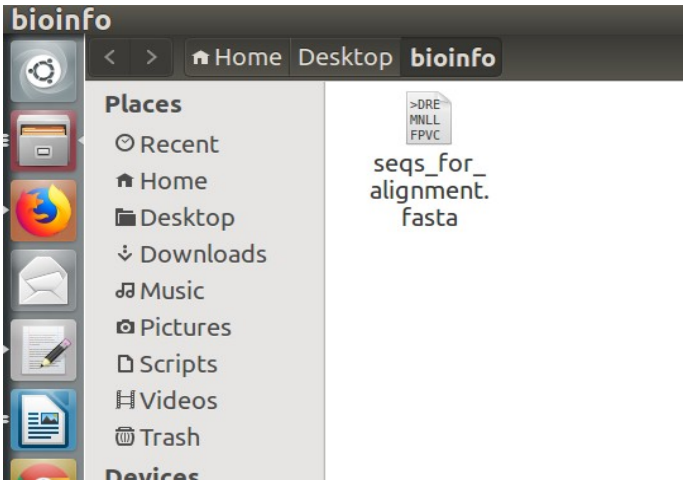
Go to SAMPLE DATA FOLDER on Desktop:

Open this folder next :

Using mouse and right click, copy this file :

Minimize window and open bioinfo folder on desktop and paste the previous file copy :

=>



Maximize the terminal or reopen:

=>

Inside the terminal write the following command and press enter:

cd Desktop/bioinfo

```
yasmeen@yasmeen[yasmeen] cd Desktop/bioinfo
```

USAGE : “cd” is a tool in Linux change your current working directory to another .

You will notice that terminal current directory has changed from your home folder e.g. “yasmeen” to your target folder “bioinfor”:

```
yasmeen@yasmeen[yasmeen] cd Desktop/bioinfo      [ 6:41PM]
yasmeen@yasmeen[bioinfo]                         [ 6:43PM]
```

USING SOME BIOLIUNX TOOLS

I
n terminal write the following command and press ENTER:

clustalw seqs_for_alignment.fasta

```
yasmeen@yasmeen[bioinfo] clustalw seqs_for_alignment.fasta
```

USAGE : clustalw is tool used for sequence alignment (1) . As I showed previously Bio-Linux has some bioinformtics tools installed (go to INSIDE BIOLINX section) . These tools could be accessed thought terminal.

After pressing ENTER this would happen in terminal:

```
yasmeen@yasmeen[bioinfo] clustalw seqs_for_alignment.fasta        [ 6:49PM]


 CLUSTAL 2.1 Multiple Sequence Alignments


Sequence format is Pearson
Sequence 1: DRERSOX9A    462 aa
Sequence 2: CAURSOX9A    457 aa
Sequence 3: OMYKSOX9     488 aa
Sequence 4: OLATSOX9B    476 aa
Sequence 5: TNIGUnk      495 aa
Sequence 6: XTRPSOX9     481 aa
Sequence 7: RRUGSOX9A    482 aa
Sequence 8: DRERSOX9B    360 aa
Sequence 9: CCARSOX9B    428 aa
Sequence 10: CAURSOX9B    416 aa
Sequence 11: TNIGUnm      484 aa
Sequence 12: ASTUSOX9     432 aa
Sequence 13: HSAPSOX9     509 aa
Start of Pairwise alignments
Aligning...

Sequences (1:2) Aligned. Score:  90
Sequences (1:3) Aligned. Score:  80
Sequences (1:4) Aligned. Score:  77
Sequences (1:5) Aligned. Score:  78
Sequences (1:6) Aligned. Score:  76
Sequences (1:7) Aligned. Score:  77
```

After few seconds the program will finish analysis and terminal would return to normal state.

```
There are 12 groups
Start of Multiple Alignment

Aligning...
Group 1: Sequences:   2      Score:9416
Group 2: Sequences:   2      Score:10155
Group 3: Sequences:   3      Score:9782
Group 4: Sequences:   4      Score:8362
Group 5: Sequences:   2      Score:9790
Group 6: Sequences:   3      Score:9417
Group 7: Sequences:   7      Score:8497
Group 8: Sequences:   9      Score:8422
Group 9: Sequences:  10      Score:8525
Group 10: Sequences:  11      Score:7842
Group 11: Sequences:   2      Score:6511
Group 12: Sequences:  13      Score:6255
Alignment Score 148754

CLUSTAL-Alignment file created  [seqs_for_alignment.aln]

yasmeen@yasmeen[bioinfo]
```

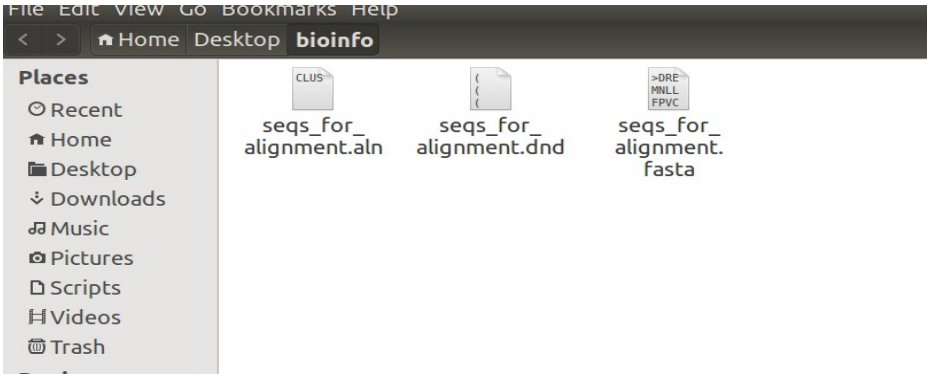
On terminal write the following command and press ENTER:

clear

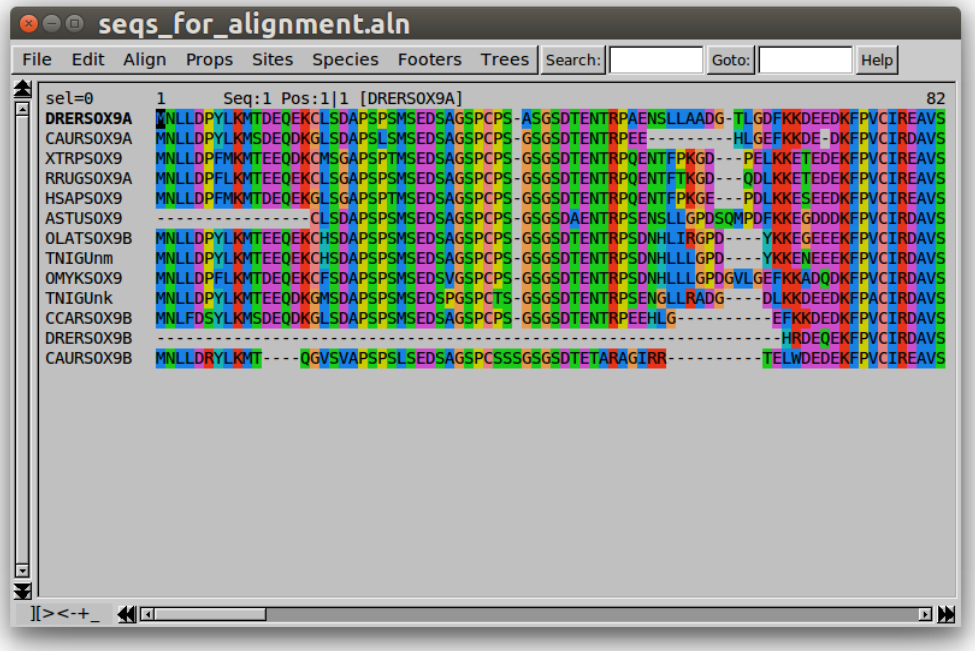
```
yasmeen@yasmeen[bioinfo] clear
```

USAGE : clear is Linux tool to clean terminal window from previous commands outputs.

Open bioinfo folder you can see the following three files :

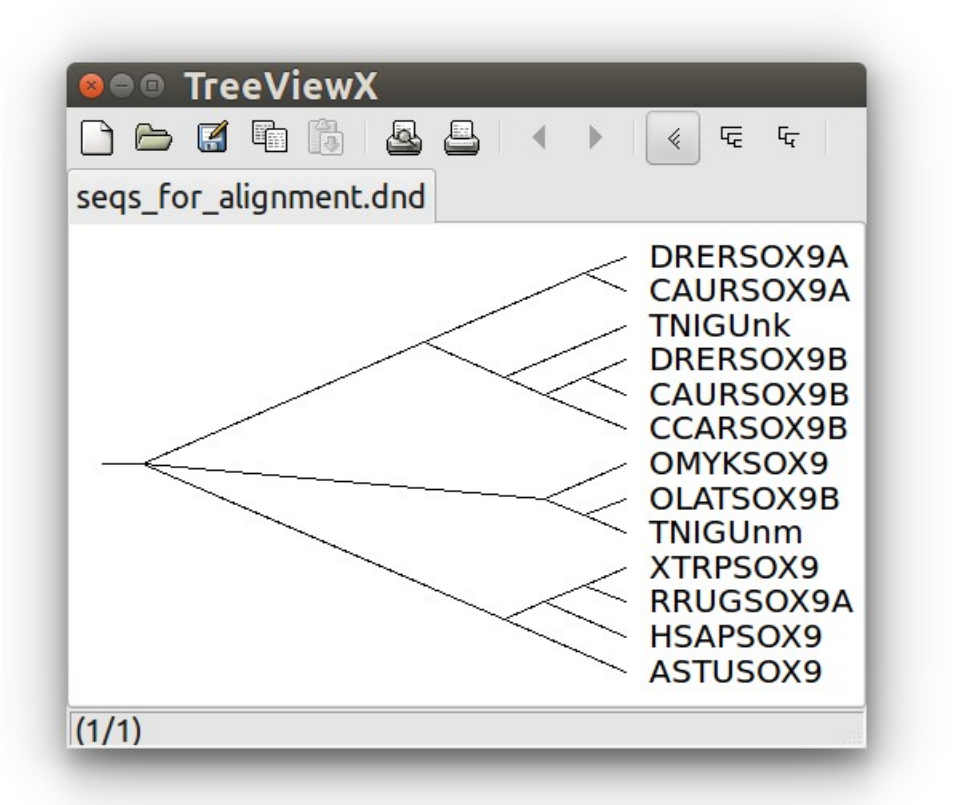


USAGE : These files are the input of clustalw program (seqs_for_alignment.fasta) and the ouput files (seqs_for_alignment.aln and seqs_for_alignment.dnd).
Using mouse right click choose to open seqs_for_alignment.aln using Seaview. The following window will appear:



UASGE : Sea View is a bioinformatics tool to show seqeunce alignment output see their paper (2).

Click on the seqs_for_alignment.dnd file and choose open with and choose Tree View X , the following window will open :



USAGE : Tree View X is a tool for drawing phylogenetic analysis.

Using Other tools

Some BioLinux tools need to inputs . The input file and the output destination . For example pepstat is a tool you can use to study the amino acid sequence different properties. Using the following command

```
yasmeen@yasmeen[bioinfo] pepstats seqs_for_alignment.fasta output.txt
```

USAGE: The first word is the program name “pepstat” , the secind is the input file (seqs_for_alignment.fasta) and the final is the output file (output.txt).

Click on the output.txt and open you will see the following output:

output.txt ×

```
PEPSTATS of DRERSOX9A from 1 to 462

Molecular weight = 51333.42             Residues = 462
Average Residue Weight  = 111.111       Charge   = -1.0
Isoelectric Point = 6.4138
A280 Molar Extinction Coefficients  = 55810 (reduced)   55935 (cystine bridges)
A280 Extinction Coefficients 1mg/ml = 1.087 (reduced)   1.090 (cystine bridges)
Improbability of expression in inclusion bodies = 0.850

Residue             Number              Mole%               DayhoffStat
A = Ala             29                  6.277               0.730
B = Asx             0                   0.000               0.000
C = Cys             3                   0.649               0.224
D = Asp             27                  5.844               1.063
E = Glu             27                  5.844               0.974
F = Phe             11                  2.381               0.661
G = Gly             27                  5.844               0.696
H = His             18                  3.896               1.948
I = Ile             11                  2.381               0.529
J = ---             0                   0.000               0.000
K = Lys             24                  5.195               0.787
L = Leu             29                  6.277               0.848
M = Met             10                  2.165               1.273
N = Asn             20                  4.329               1.007
O = ---             0                   0.000               0.000
P = Pro             45                  9.740               1.873
Q = Gln             36                  7.792               1.998
R = Arg             20                  4.329               0.883
S = Ser             56                  12.121              1.732
T = Thr             26                  5.628               0.923
U = ---             0                   0.000               0.000
V = Val             19                  4.113               0.623
W = Trp             5                   1.082               0.833
X = Xaa             0                   0.000               0.000
Y = Tyr             19                  4.113               1.210
Z = Glx             0                   0.000               0.000

Property  Residues            Number              Mole%
Tiny                (A+C+G+S+T)                     141                 30.519
Small               (A+B+C+D+G+N+P+S+T+V)           252                 54.545
Aliphatic (A+I+L+V)             88                  19.048
Aromatic  (F+H+W+Y)             53                  11.472
Non-polar (A+C+F+G+I+L+M+P+V+W+Y)       208                 45.022
Polar               (D+E+H+K+N+Q+R+S+T+Z)           254                 54.978
Charged             (B+D+E+H+K+R+Z)                 116                 25.108
Basic               (H+K+R)                         62                  13.420
Acidic              (B+D+E+Z)           54                  11.688
```

Some tools need you to direct their output to out files. For example ; **revcomp** is tool you can use to convert sequences to its reverse complement .

```
yasmeen@yasmeen[bioinfo] revcomp seqs_for_alignment.fasta > REV-OUT.fasta
```

The first word is the tool name **revcomp** , the second is the input file name (**seqs_for_alignment.fasta**) and the third **“ > ”** is the direction sign used to direct the output of the tool to out file and the final is the output file . If you didn't use the last two words (> and output file name) the tool output (**revcomp**) will be printed to the terminal window. If you open the REV-OUT.fasta you will see the following :