

Figure 5: Screenshot of the Results summary

The result files are with different formats of input and output files of the alignment.

The user can enable the java plug-in in the browser, if it is disabled and thus the user can use Jalview to see the alignment with the colours as shown in Figure 6.

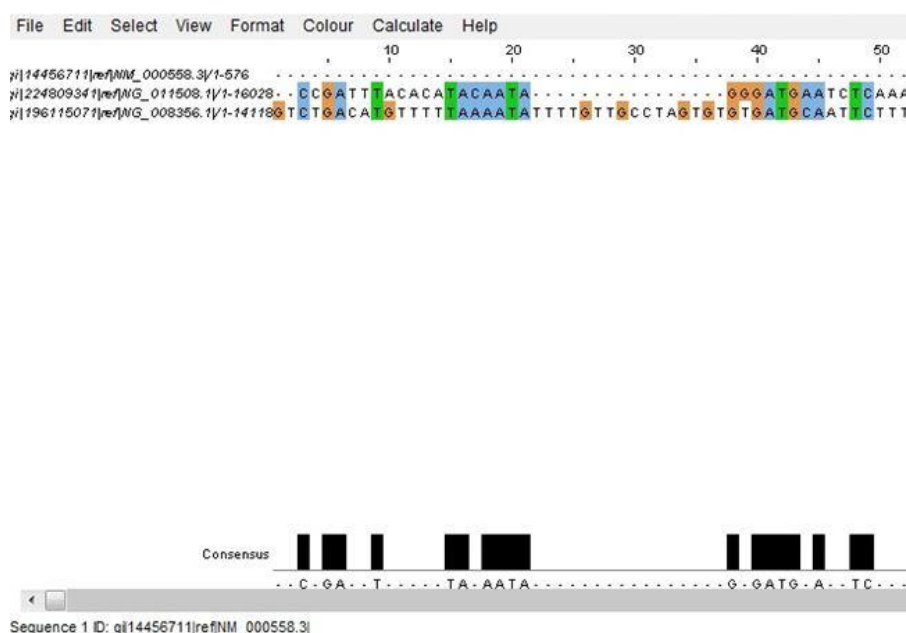


Figure 6: Screenshot of Jalview

The user can view the output file and can save by clicking on the button “View output file” as shown in Figure 5.

The output file represents the length of each sequence , and the score of each alignment individually as shown in Figure 7.

```

CLUSTAL 2.1 Multiple Sequence Alignments

Sequence type explicitly set to DNA
Sequence format is Pearson
Sequence 1: gi|14456711|ref|NM_000558.3|      576 bp
Sequence 2: gi|224809341|ref|NG_011508.1|    16028 bp
Sequence 3: gi|196115071|ref|NG_008356.1|    14118 bp
Start of Pairwise alignments
Aligning...

Sequences (1:2) Aligned. Score: 86
Sequences (1:3) Aligned. Score: 83
Sequences (2:3) Aligned. Score: 56
Guide tree file created: [clustalw2-I20121106-040910-0263-102544-oy.dnd]

There are 2 groups
Start of Multiple Alignment

Aligning...
Group 1: Sequences: 2      Score:3784
Group 2: Sequences: 3      Score:74592
Alignment Score 37335

CLUSTAL-Alignment file created [clustalw2-I20121106-040910-0263-102544-oy.aln]

```

Figure 7: Screenshot of the output file to be saved

Interpretation:

Clustal W is a general purpose multiple sequence alignment program for DNA or proteins. So basically three or more sequences are taken for finding out structural and functional relationship between these sequences. It uses progressive alignment methods, which align the most similar sequences first and work their way down to the least similar.

Experiment 6: Construction of Cladogram

Aim: To find the evolutionary relationship between different organisms and analyze the changes that occurred in organisms during the course of evolution using PHYLIP.

Theory:

PHYLIP is a complete phylogenetic analysis package which was developed by Joseph Felsenstein at University of Washington. PHYLIP is used to find the evolutionary relationships between different organisms. Some of the methods available in this package are maximum parsimony method, distance matrix and likelihood methods. The data is presented to the program from a text file, which is prepared by the user using common text editors such as word processor, etc. Some of the sequence analysis programs such as ClustalW can write data files in PHYLIP format. Most of the programs look for the input file called "infile" -- if they do not find this file, then they ask the user to type in the file name of the data file. Before starting the computation, the program will ask the user to set options (optional) through a menu. Output is written into special files with names like outfile and outtree.

Procedure:

Procedure for phylogenetic analysis :

Go to simulator tab to know more about how to retrieve the query sequence .

Procedure For PHYLIP :

Align the multiple DNA sequences (output of the ClustalW) and save it in PHYLIP format as infile.phy. Start the program of Dnadist by clicking the icon and giving this infile as input.

All the PHYLIP programs are menu driven programs. Dnadist will calculate pairwise distances between the sequences. At first, Dnadist will ask whether the input file is there in the PHYLIP folder. If the file does not exist, it will ask you to give the correct file name. After giving the correct input, if needed it will ask to change any settings for the program by typing the first letter or number. If the changes are not required, by typing 'Y' it will start running the program. Output will return to the file as outfile, so that the output of this file can be used as input of another program. Output would be as represented in Figure 1.

```

0      Terminal type (IBM PC, ANSI, none)?  IBM PC
1      Print out the data at start of run  No
2      Print indications of progress of run  Yes

Y to accept these or type the letter for one to change
y
Distances calculated for species
HUMAN      .....
CHIMPANZEP .....
MONKEYPAPO ...
MONKEYMOCA ..
RABBIT1    .
RABBIT2

Distances written to file "outfile"

Done.

Press enter to quit.

```

	HUMAN	CHIMPANZEP	MONKEYPAPO	MONKEYMOCA	RABBIT1	RABBIT2
HUMAN	0.000000	0.035305	0.049562	0.040798	0.222263	0.223706
CHIMPANZEP	0.035305	0.000000	0.074717	0.042751	0.222159	0.223601
MONKEYPAPO	0.049562	0.074717	0.000000	0.036918	0.222306	0.221246
MONKEYMOCA	0.040798	0.042751	0.036918	0.000000	0.222165	0.223678
RABBIT1	0.222263	0.222159	0.222306	0.222165	0.000000	0.014499
RABBIT2	0.223706	0.223601	0.221246	0.223678	0.014499	0.000000

Figure 1: Distance Representation

Like Dnadist, Neighbor also gives sequence distance analysis. Output of Dnadist is given as input to Neighbor. Output file and tree file will be returned to outfile and outtree as represented in Figure 2 .

```

0      Terminal type (IBM PC, ANSI, none)?  IBM PC
1      Print out the data at start of run  No
2      Print indications of progress of run  Yes
3      Print out tree  Yes
4      Write out trees onto tree file?  Yes

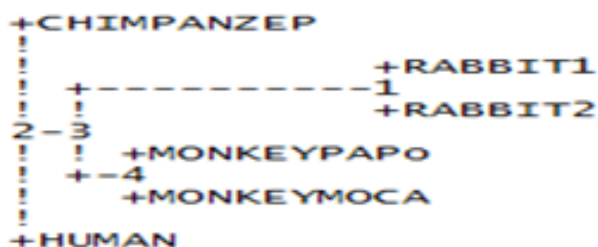
Y to accept these or type the letter for one to change
y
Cycle   3: species 5 <  0.00683> joins species 6 <  0.00767>
Cycle   2: species 1 <  0.01315> joins species 2 <  0.02215>
Cycle   1: node 1 <  0.00862> joins node 5 <  0.18941>
last cycle:
node 1 <  0.00723> joins species 3 <  0.02326> joins species 4 <  0.01365>
>
Output written on file "outfile"
Tree written on file "outtree"
Done.
Press enter to quit.

```

Figure 2 : Sequence distance analysis.

Branch lengths and tree are represented with the help of Neighbor joining method. The outfile and outtree after the Neighbor joining method are given below (Figures 3 & 4).

6 Populations
Neighbor-Joining/UPGMA method version 3.69
Neighbor-joining method
Negative branch lengths allowed



remember: this is an unrooted tree!

Between	And	Length
2	CHIMPANZEP	0.02215
2	3	0.00862
3	1	0.18941
1	RABBIT1	0.00683
1	RABBIT2	0.00767
3	4	0.00723
4	MONKEYPAPO	0.02326
4	MONKEYMOCA	0.01365
2	HUMAN	0.01315

Figure 3: Outfile

```

(CHIMPANZEP:0.02215,((RABBIT1:0.00683,RABBIT2:0.00767):0.18941,
(MONKEYPAPO:0.02326,MONKEYMOCA:0.01365):0.00723):0.00862,HUMAN:0.01315);

```

Figure 4: Outtree

Cladogram is represented via Consensus tree program. Input for the cladogram will be output (outtree) of Neighbor program which will generate outfile and outtree. It represents the consensus tree. Numbers on the branches indicate the number of times the species has been partitioned into two sets separated by that branch occurred among the trees. Here the outfile and outtree are represented in Figure 5 and Figure 6.

Consensus tree program, version 3.69

Species in order:

1. CHIMPANZEP
2. RABBIT1
3. RABBIT2
4. MONKEYPAPO
5. MONKEYMOCA
6. HUMAN

Sets included in the consensus tree

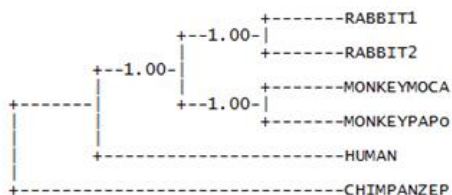
Set (species in order)	How many times out of	1.00
.****.	1.00	
***..	1.00	
...**.	1.00	

Sets NOT included in consensus tree: NONE

Extended majority rule consensus tree

Extended majority rule consensus tree

CONSENSUS TREE:
the numbers on the branches indicate the number of times the partition of the species into the two sets which are separated by that branch occurred among the trees, out of 1.00 trees (trees had fractional weights)



remember: this is an unrooted tree!

Figure 5 : Cladogram representation (outfile).

```
((((RABBIT1:1.00,RABBIT2:1.00):1.00,(MONKEYMOCA:1.00,MONKEYPAPO:1.00):1.00):1.00,HUMAN:1.00):1.00,CHIMPANZEP:1.00);
```

Figure 6 :Cladogram Representation (outtree).

Interpretation:

So basically cladograms are diagrams which depict the evolutionary relationship between different organisms and analyse the changes that occurred in organisms during the course of evolution using PHYLIP. Cladogram is a branching diagram showing the cladistic relationship between a number of species.

Experiment 7: Phylogenetic Analysis using PHYLIP - Rooted trees

Aim: To find the evolutionary relationship between different organisms based on the time scale and to analyze the changes that occurred in an organisms using PHYLIP.

Theory:

PHYLIP is a complete phylogenetic analysis package which was developed by Joseph Felsenstein at University of Washington. PHYLIP is used to find the evolutionary relationships between different organisms. Some of the methods available in this package are maximum parsimony method, distance matrix and likelihood methods. The data is presented to the program from a text file, which is prepared by the user using common text editors such as word processor, etc. Some of the sequence analysis programs such as ClustalW can write data files in PHYLIP format. Most of the programs look for the

input file called "infile" -- if they do not find this file, then they ask the user to type in the file name of the data file. Before starting the computation, the program will ask the user to set options (optional) through a menu. Output is written into special files with names like outfile and outtree.

Procedure:

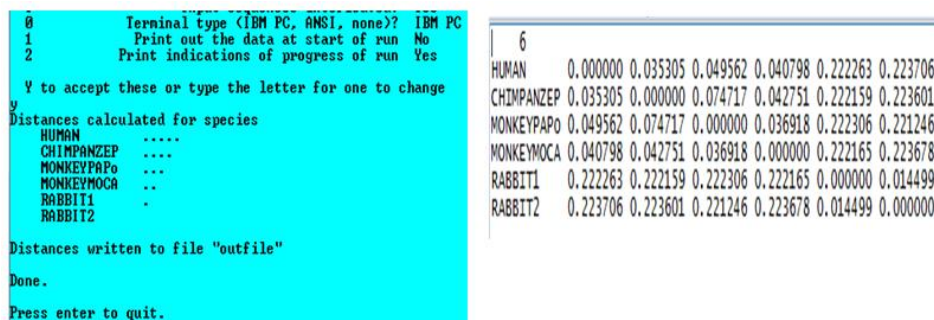
Procedure for phylogenetic analysis :

Go to simulator tab to know more about how to retrieve the query sequence .

Procedure For PHYLIP :

Align the multiple DNA sequences (output of the ClustalW) and save it in PHYLIP format as infile.phy. Start the program of Dnadist by clicking the icon and giving this infile as input.

All the PHYLIP programs are menu driven programs. Dnadist will calculate pairwise distances between the sequences. At first, Dnadist will ask whether the input file is there in the PHYLIP folder. If the file does not exist, it will ask you to give the correct file name. After giving the correct input, if needed it will ask to change any settings for the program by typing the first letter or number. If the changes are not required, by typing 'Y' it will start running the program. Output will return to the file as outfile, so that the output of this file can be used as input of another program. Output would be as represented in Figure 1.



```

0 Terminal type (IBM PC, ANSI, none)? IBM PC
1 Print out the data at start of run No
2 Print indications of progress of run Yes

Y to accept these or type the letter for one to change
y
Distances calculated for species
HUMAN
CHIMPANZEP
MONKEYPAPo
MONKEYMOCA
RABBIT1
RABBIT2

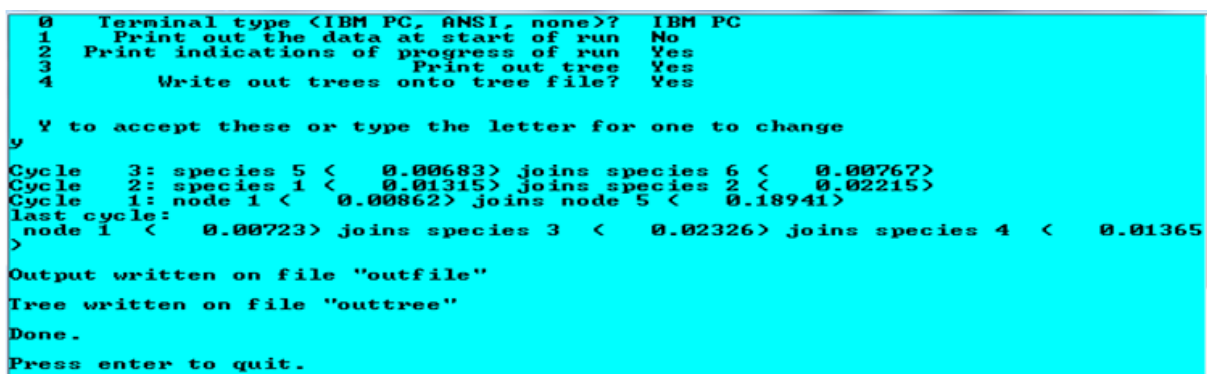
Distances written to file "outfile"
Done.
Press enter to quit.

```

	HUMAN	CHIMPANZEP	MONKEYPAPo	MONKEYMOCA	RABBIT1	RABBIT2
HUMAN	0.000000	0.035305	0.049562	0.040798	0.222263	0.223706
CHIMPANZEP	0.035305	0.000000	0.074717	0.042751	0.222159	0.223601
MONKEYPAPo	0.049562	0.074717	0.000000	0.036918	0.222306	0.221246
MONKEYMOCA	0.040798	0.042751	0.036918	0.000000	0.222165	0.223678
RABBIT1	0.222263	0.222159	0.222306	0.222165	0.000000	0.014499
RABBIT2	0.223706	0.223601	0.221246	0.223678	0.014499	0.000000

Figure 1: Distance Representation

Like Dnadist, Neighbor also gives sequence distance analysis. Output of Dnadist is given as input to Neighbor. Output file and tree file will be returned to outfile and outtree as represented in Figure 2 .



```

0 Terminal type (IBM PC, ANSI, none)? IBM PC
1 Print out the data at start of run No
2 Print indications of progress of run Yes
3 Print out tree Yes
4 Write out trees onto tree file? Yes

Y to accept these or type the letter for one to change
y
Cycle 3: species 5 < 0.00683> joins species 6 < 0.00767>
Cycle 2: species 1 < 0.01315> joins species 2 < 0.02215>
Cycle 1: node 1 < 0.00862> joins node 5 < 0.18941>
last cycle:
node 1 < 0.00723> joins species 3 < 0.02326> joins species 4 < 0.01365>

Output written on file "outfile"
Tree written on file "outtree"
Done.
Press enter to quit.

```

Figure 2 : Sequence distance analysis.

Branch lengths and tree are represented with the help of Neighbor joining method. The outfile and outtree after the Neighbor joining method are given below (Figures 3 & 4).

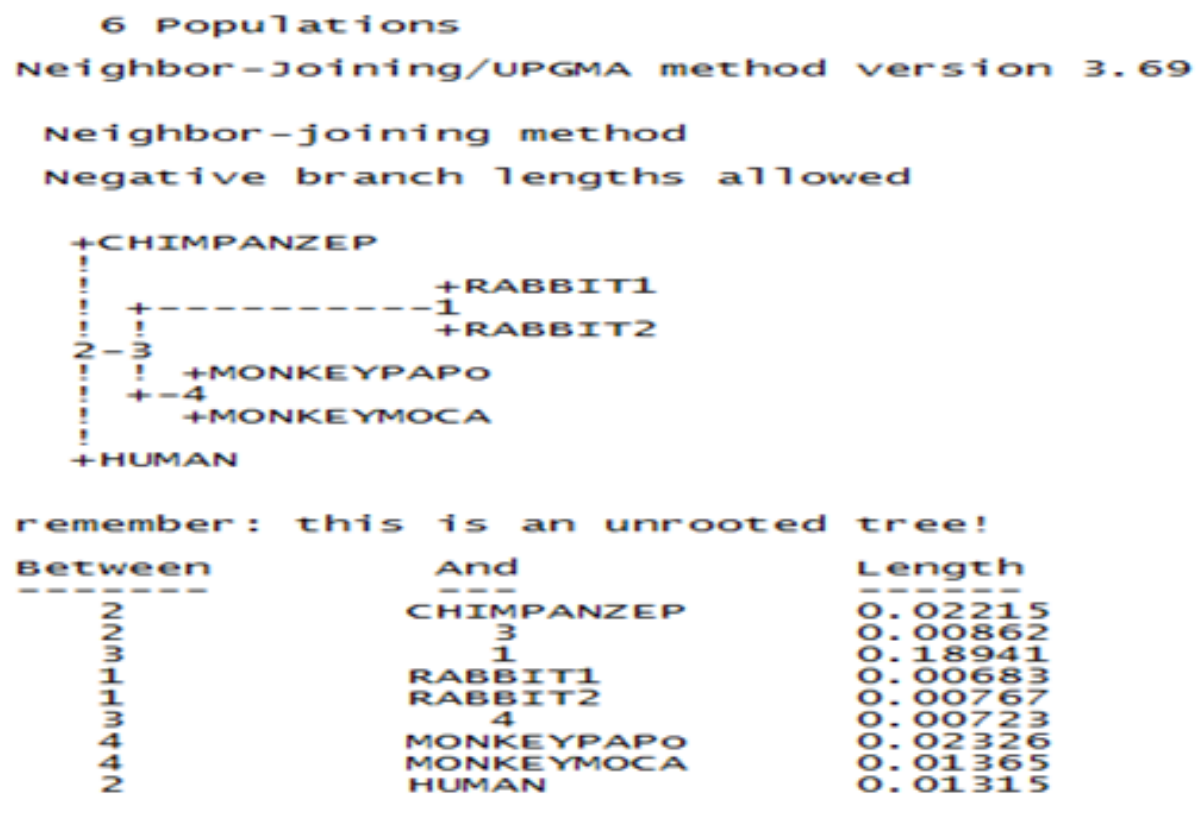


Figure 3: Outfile

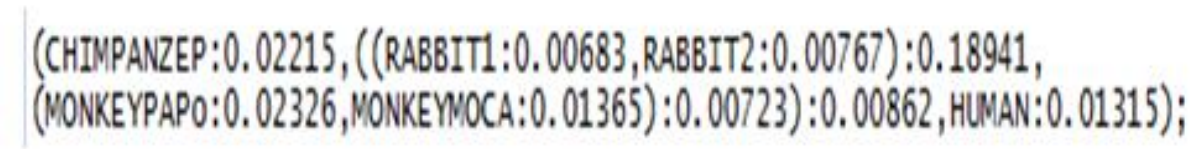


Figure 4: Outtree

Rooted trees are represented via Drawgram by providing the input as the "outfile" obtaining from neighbor joining method. Rooted tree considers an imaginary root as the start and from that the other sequences are aligned.

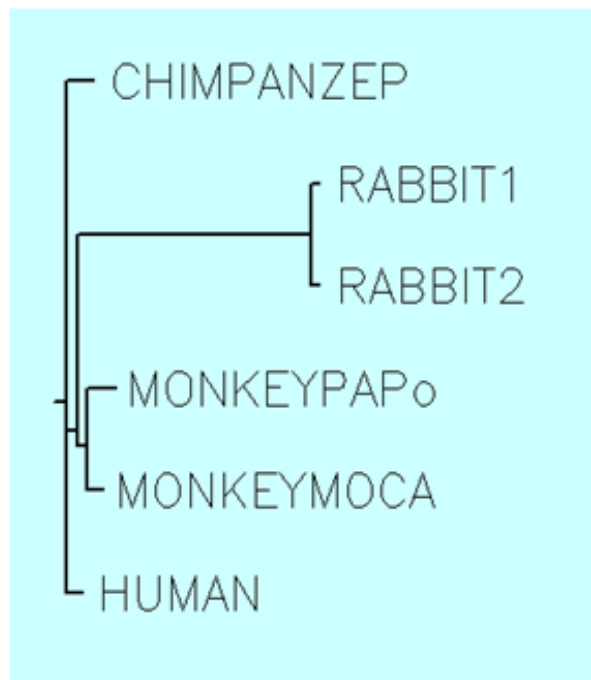


Figure 5: Rooted tree

Interpretation:

As we know that PHYLIP is used to find evolutionary relationship between different organisms based on the time scale and to analyze the changes that occurred in an organism. Also, the rooted tree considers an imaginary root as the start and from that the other sequences are aligned. PHYLIP is a complete phylogenetic analysis package which was developed by Joseph Felsenstein at the University of Washington.

Experiment 8: Phylogenetic Analysis using PHYLIP - Unrooted trees

Aim: To find the evolutionary relationships between organisms and to analyze the changes occurring in these organisms during evolution using PHYLIP.

Theory:

PHYLIP is a complete phylogenetic analysis package which was developed by Joseph Felsenstein at University of Washington. PHYLIP is used to find the evolutionary relationships between different organisms. Some of the methods available in this package are maximum parsimony method, distance matrix and likelihood methods. The data is presented to the program from a text file, which is prepared by the user using common text editors such as word processor, etc. Some of the sequence analysis programs such as ClustalW can write data files in PHYLIP format. Most of the programs look for the input file called "infile" -- if they do not find this file, then they ask the user to type in the file name of the data file. Before starting the computation, the program will ask the user to set options (optional) through a menu. Output is written into special files with names like outfile and outtree.

Procedure:

Procedure for phylogenetic analysis :

Go to simulator tab to know more about how to retrieve the query sequence .

Procedure For PHYLIP :

Align the multiple DNA sequences (output of the ClustalW) and save it in PHYLIP format as infile.phy. Start the program of Dnadist by clicking the icon and giving this infile as input. All the PHYLIP programs are menu driven programs. Dnadist will calculate pairwise distances between the sequences. At first, Dnadist will ask whether the input file is there in the PHYLIP folder. If the file does not exist, it will ask you to give the correct file name. After giving the correct input, if needed it will ask to change any settings for the program by typing the first letter or number. If the changes are not required, by typing 'Y' it will start running the program. Output will return to the file as outfile, so that the output of this file can be used as input of another program. Output would be as represented in Figure 1.

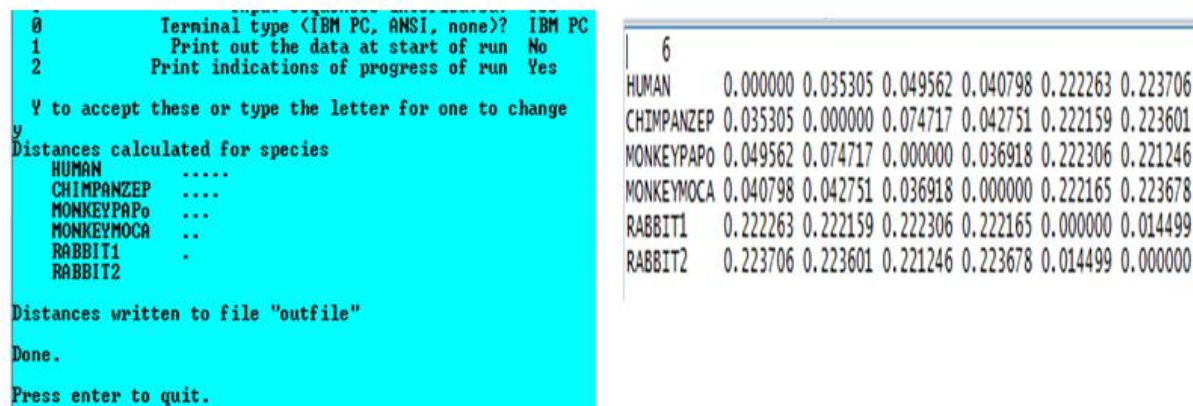


Figure 1: Distance Representation

Like Dnadist, Neighbor also gives sequence distance analysis. Output of Dnadist is given as input to Neighbor. Output file and tree file will be returned to outfile and outtree as represented in Figure 2 .

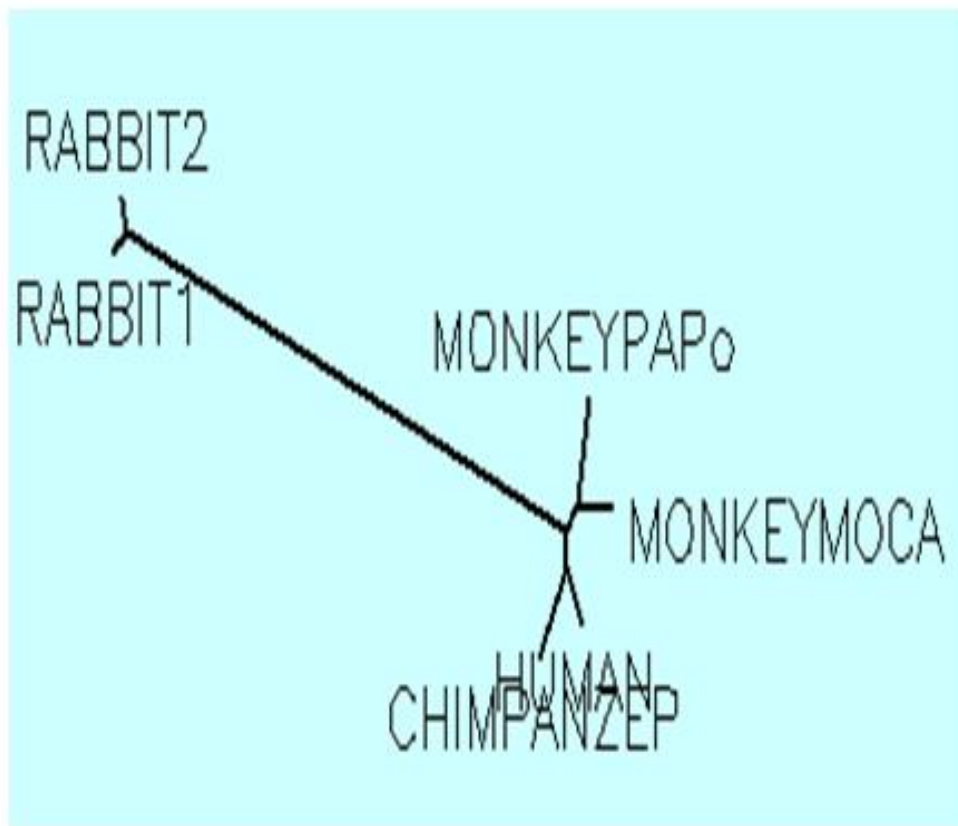


Figure 6: Unrooted tree

Interpretation:

As we know PHYLIP is used to find the evolutionary relationships between organisms and to analyze the changes occurring in these organisms during evolution also the unrooted trees are represented via Drawtree. PHYLIP is a complete phylogenetic analysis package which was developed by Joseph Felsenstein at University of Washington.