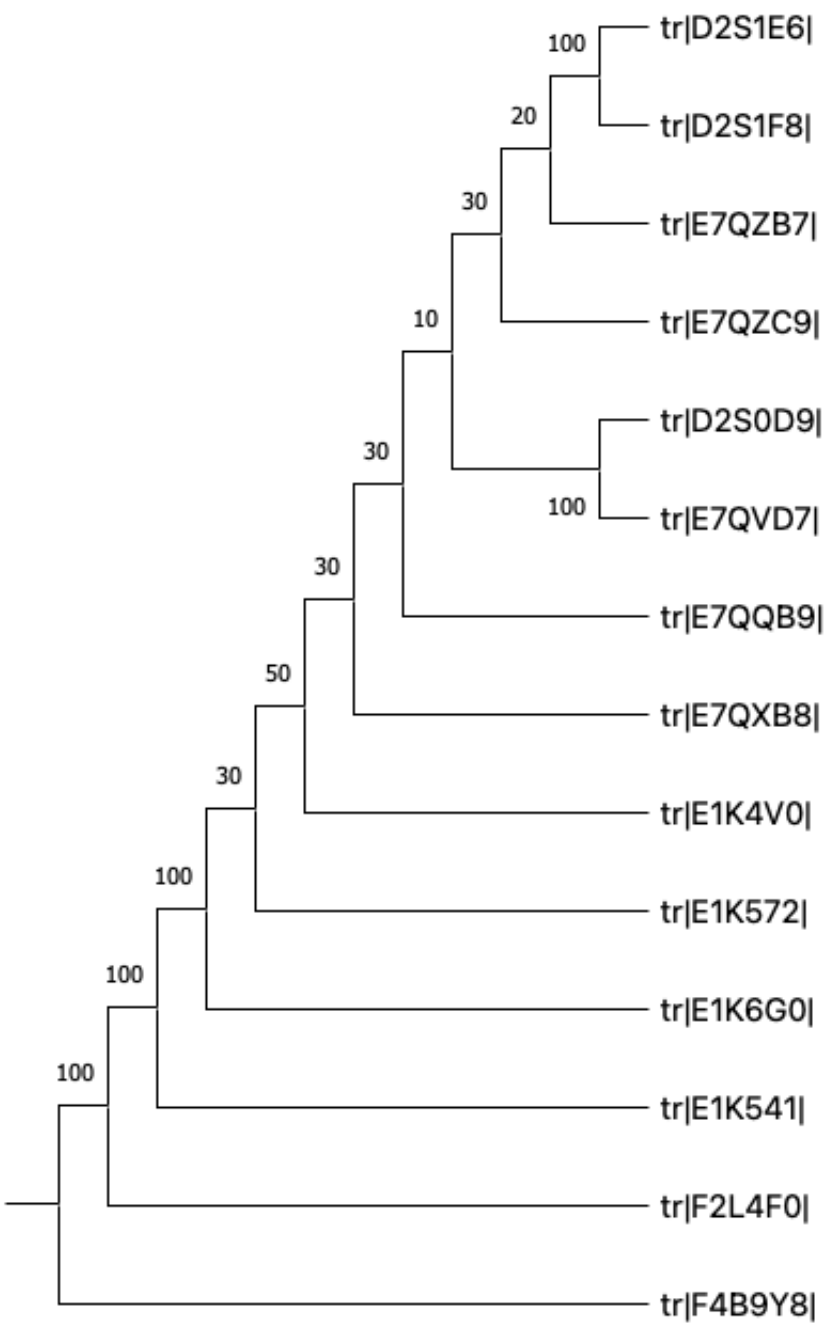
Question 1

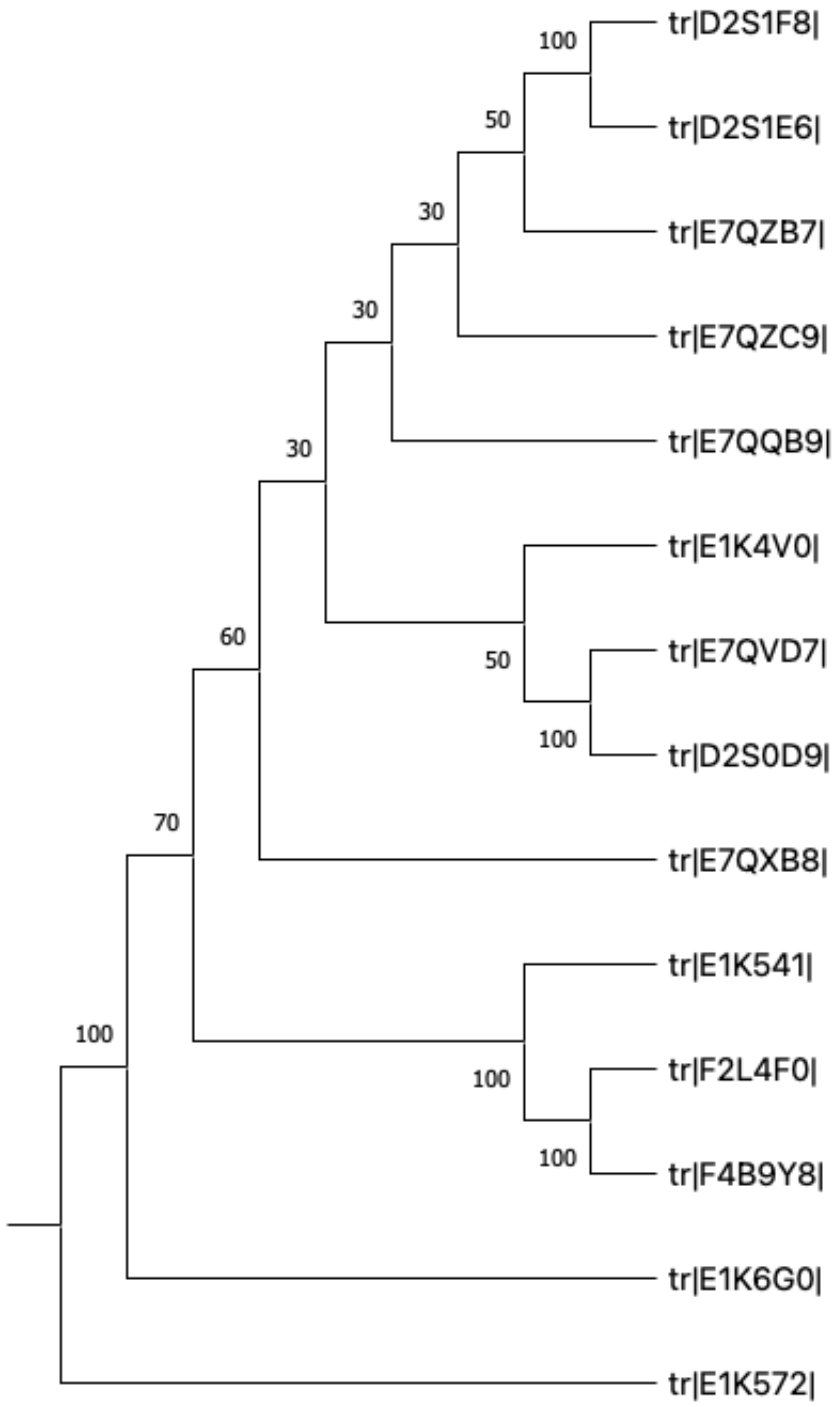
All files made using the differnet programs are attached in moodle

Set 1

1. Consensus tree using Maximum Likelihood

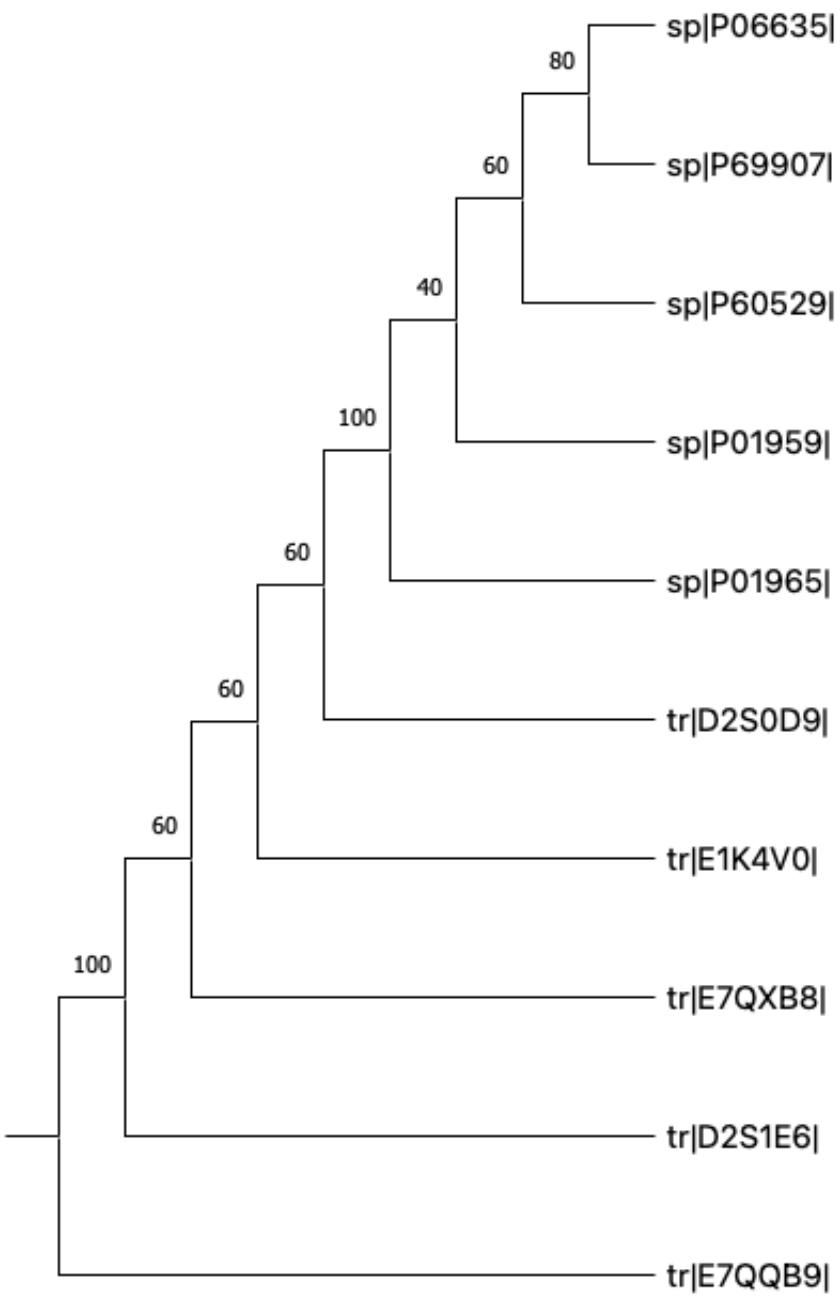


1. Consensus tree using Neighbour Joining and UPGMA

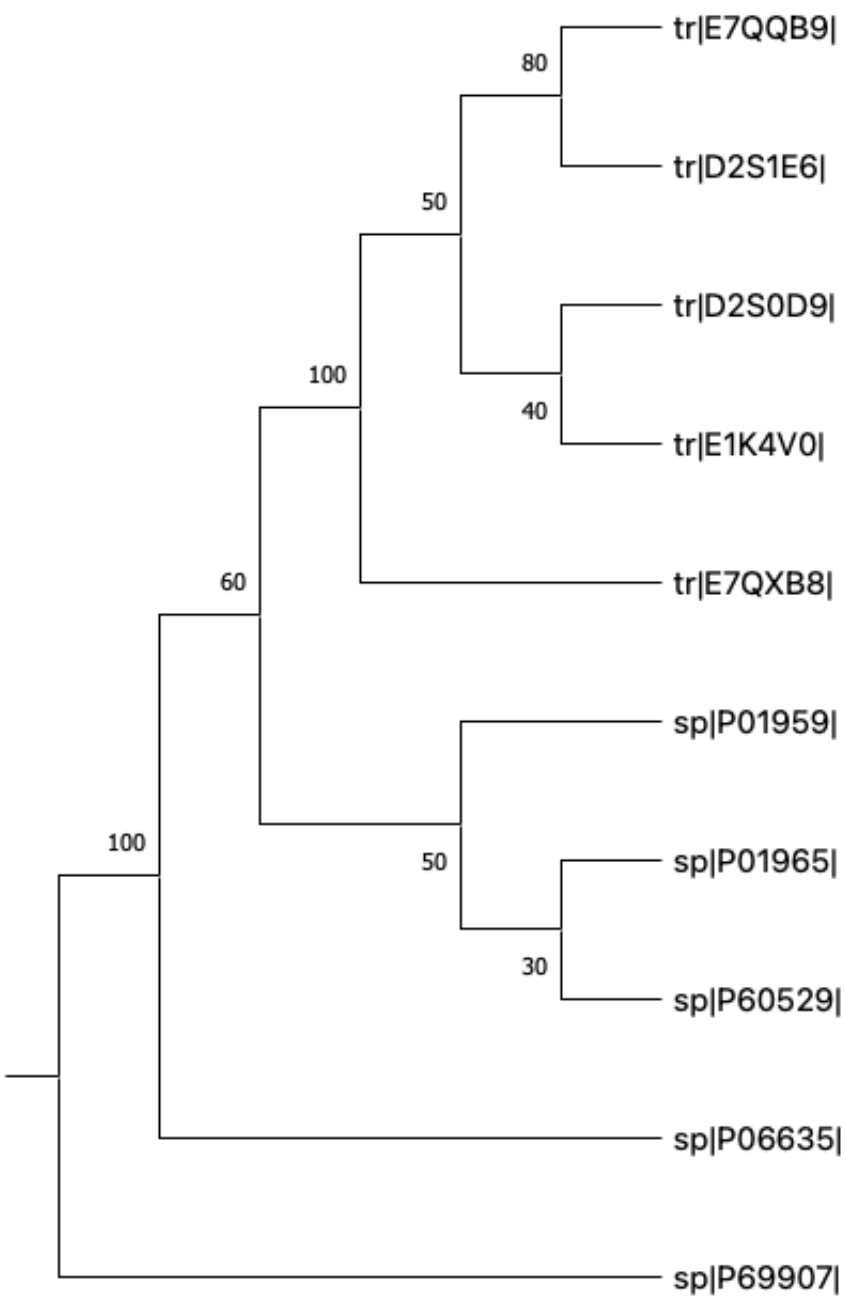


Set 2

1. Consensus tree using Maximum Likelihood



1. Consensus tree using Neighbour Joining and UPGMA



Question 2

In [83]:
```
from math import log
import string
import pandas as pd
import numpy as np
seqs = """MVLSPADKTNVKGKVGAHAGEYGAAAW
   MKRLPADPPCVKGKVKAKAGDYGATTW
   MALSAADKTNVKSKVGGHAGEYGAATS
   MVLSAADKTNVKSKAGGNAGEWWAAAW
   MVLSAADKTNVKSKVLANAGEFGAAAW
   ALLPIRTTYHKKCASGHIPEEKDLNNV
   DEASSLKGHHIKKLEADALLIPLSASS
""".split("\n")
seqs = [x.strip(' ') for x in seqs][:-1]

p = 1/20
N = len(seqs)
aas = sorted(set(string.ascii_uppercase) - set("BJOZXU"))
n = [[0 for i in range(len(seqs[0]))] for j in range(len(aas))]
mat = [[0 for i in range(len(seqs[0]))] for j in range(len(aas))]

for i in range(len(aas)):
    for j in range(len(seqs[0])):
        n[i][j] = list(zip(*seqs))[j].count(aas[i])
        mat[i][j] = round(log (((n[i][j]+p) / (N+1)) / p),3)
df = pd.DataFrame(mat, index = aas, columns = list(range(1,28)))
```

In [84]:
```
df[list(range(1,15))]
```

Out[84]:

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
A	0.965	0.965	0.965	-2.079	2.031	2.536	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	0.965
C	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	0.965	-2.079	-2.079	0.965	-2.079
D	0.965	-2.079	-2.079	-2.079	-2.079	-2.079	2.536	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079
E	-2.079	0.965	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079
F	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079
G	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	0.965	-2.079	-2.079	-2.079	-2.079	1.634	-2.079
H	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	0.965	1.634	-2.079	-2.079	-2.079	-2.079
I	-2.079	-2.079	-2.079	-2.079	0.965	-2.079	-2.079	-2.079	-2.079	-2.079	0.965	-2.079	-2.079	-2.079
K	-2.079	0.965	-2.079	-2.079	-2.079	-2.079	0.965	2.315	-2.079	-2.079	0.965	2.869	0.965	2.536
L	-2.079	0.965	2.536	0.965	-2.079	0.965	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	0.965
M	2.536	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079
N	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	2.315	-2.079	-2.079	-2.079	-2.079
P	-2.079	-2.079	-2.079	0.965	1.634	-2.079	-2.079	0.965	0.965	-2.079	-2.079	-2.079	-2.079	-2.079
Q	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079
R	-2.079	-2.079	0.965	-2.079	-2.079	0.965	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079
S	-2.079	-2.079	-2.079	2.536	0.965	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	2.031	-2.079
T	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	0.965	0.965	2.315	-2.079	-2.079	-2.079	-2.079	-2.079
V	-2.079	2.031	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	2.536	-2.079	-2.079	-2.079
W	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079
Y	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	0.965	-2.079	-2.079	-2.079	-2.079	-2.079

In [85]:
```
df[list(range(15,28))]
```

Out[85]:

	15	16	17	18	19	20	21	22	23	24	25	26	27
A	0.965	0.965	2.031	0.965	2.536	-2.079	-2.079	-2.079	-2.079	2.536	2.536	2.031	-2.079
C	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079
D	-2.079	-2.079	0.965	-2.079	-2.079	-2.079	0.965	-2.079	0.965	-2.079	-2.079	-2.079	-2.079
E	0.965	-2.079	-2.079	-2.079	-2.079	0.965	2.536	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079
F	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	0.965	-2.079	-2.079	-2.079	-2.079	-2.079
G	-2.079	2.315	1.634	-2.079	-2.079	2.536	-2.079	-2.079	2.315	-2.079	-2.079	-2.079	-2.079
H	-2.079	-2.079	0.965	1.634	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079
I	-2.079	-2.079	-2.079	0.965	-2.079	-2.079	0.965	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079
K	-2.079	0.965	-2.079	0.965	-2.079	-2.079	-2.079	0.965	-2.079	-2.079	-2.079	-2.079	-2.079
L	-2.079	0.965	-2.079	-2.079	0.965	0.965	-2.079	-2.079	0.965	0.965	-2.079	-2.079	-2.079
M	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079
N	-2.079	-2.079	-2.079	1.634	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	0.965	0.965	-2.079
P	-2.079	-2.079	-2.079	-2.079	0.965	-2.079	-2.079	0.965	-2.079	-2.079	-2.079	-2.079	-2.079
Q	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079
R	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079
S	0.965	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	0.965	-2.079	0.965	1.634
T	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	0.965	1.634	-2.079
V	2.315	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	0.965
W	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	0.965	0.965	-2.079	-2.079	-2.079	2.315
Y	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	-2.079	2.031	-2.079	-2.079	-2.079	-2.079	-2.079