Alignment

The main goal is to compare some multimple alignment algorithms.

1. Materials and methods

I work with nucleotide sequences of gene of translation termination factor (SUP35) from different species of yeasts
In this project I use the following programms:
- CLUSTAL 2.1
- MUSCLE v3.8.1551
- MAFFT v7.475
- Kalign (3.3)
- T-COFFEE Version_13.41.0.28bdc39
- PRANK v.170427
- transeq (Version: EMBOSS:6.6.0.0)
- getorf (Version: EMBOSS:6.6.0.0)

2. Using 6 various alignment algorithms on 10 DNA sequences

clustalw

By default we generate the alignment in clustalw format

```
clustalw -INFILE=./data/raw_data/SUP35_10seqs.fa
```

time:

```
 real    0m3,747s
user    0m3,732s
sys     0m0,015s
```

Determining the output file:

```
clustalw -INFILE=./data/raw_data/SUP35_10seqs.fa -OUTFILE=./data/processed_data/SUP35_10seqs.clustalw
```

Generating the output in fasta format:

```
clustalw -INFILE=./data/raw_data/SUP35_10seqs.fa -OUTPUT=FASTA -OUTFILE=./data/processed_data/SUP35_10seqs.clustalw.fa
```

muscle

```
muscle   -in ./data/raw_data/SUP35_10seqs.fa -out ./data/processed_data/SUP35_10seqs_muscle.fa
```

time:

```
 real    0m4,742s
user    0m4,712s
sys     0m0,022s
```

mafft

```
mafft --auto ./data/raw_data/SUP35_10seqs.fa >./data/raw_data/SUP35_10seqs_mafft.fa
```

time:

```
 real    0m4,486s
user    0m4,351s
sys     0m0,156s
```

kalign

```
kalign <./data/raw_data/SUP35_10seqs.fa >./data/processed_data/SUP35_10seqs_kalign.fa
```

time:

```
 real	0m0,242s
user	0m0,357s
sys	0m0,008s
```

t_coffee

```
t_coffee -infile=./data/raw_data/SUP35_10seqs.fa -outfile=./data/processed_data/SUP35_10seqs_tcoffee.fa
```

time:

```
 real	1m31,287s
user	1m29,863s
sys	0m1,477s
```

prank

```
prank -d=./data/raw_data/SUP35_10seqs.fa -o=./data/processed_data/SUP35_10seqs_prank.fa
```

time:

```
 real	0m12,179s
user	0m12,006s
sys	0m0,163s
```

3. Comparation of the results on 10 DNA sequences

The table containing real times and grapical representation of the results can be found in Supplementary (Table 1, Figures 1-5)
I find mafft and t-coffee the best ones

4. Reverse complement problem

While opening the file ./data/raw_data/SUP35_10seqs_strange_aln.fa we can see that alignment for 1 sequence is not satisfactory (Supplementary, fig. 6). But if we BLAST it, we see that it is the same gene of the same organism. The answer is to find the reverse complement sequence and realign it

5. Using 6 various alignment algorithms on 250 DNA sequences

clustalw

```
clustalw -INFILE=./data/raw_data/SUP35_250seqs.fa OUTPUT=FASTA -OUTFILE=./data/processed_data/SUP35_250seqs.clustalw.fa
```

time:

```
 real	29m3,728s
user	29m2,200s
sys	0m1,483s
```

muscle

```
muscle -in ./data/raw_data/SUP35_250seqs.fa -out ./data/processed_data/SUP35_250seqs_muscle.fa
```

time

```
 real	1m20,447s
user	1m20,169s
sys	0m0,276s
```

mafft

```
mafft --auto ./data/raw_data/SUP35_250seqs.fa >./data/processed_data/SUP35_250seqs_mafft.fa
```

time:

```
 real    0m39,807s
user    0m39,104s
sys     0m0,744s
```

kalign

```
kalign <./data/raw_data/SUP35_250seqs.fa >./data/processed_data/SUP35_250seqs_kalign.fa
```

time:

```
 real    0m3,912s
user    0m7,910s
sys     0m0,105s
```

t_coffee

```
t_coffee -infile=./data/raw_data/SUP35_250seqs.fa -outfile=./data/processed_data/SUP35_250seqs_tcoffee.fa
```

I decided to interrupt the process after:

```
 real    39m1,756s
user    0m0,761s
sys     0m0,275s
```

prank

```
prank -d=./data/raw_data/SUP35_250seqs.fa -o=./data/processed_data/SUP35_250seqs_prank.fa
```

time:

```
 real    2m39,690s
user    2m35,851s
sys     0m1,994s
```

6. Comparation of the results on 250 DNA sequences

The table containing real times and grapical representation of the results can be found in Supplementary (Table 1, Figures 7-10)
I find CLUSTAL to be the best when you can wait or muscle when you want to make it faster.

7. Translation

Let's translate the 10 DNA sequences:

```
transeq -sequence ./data/raw_data/SUP35_10seqs.fa -outseq ./data/processed_data/SUP35_10seqs.t.faa
```

Another way is to use getorf. getorf gives you the ORF and its coordinates in nucleotides. We shold give the minsize near the protein size (in nucleotides) to get its sequence witout other small peptides.

```
getorf -sequence ./data/raw_data/SUP35_10seqs.fa -outseq ./data/processed_data/SUP35_10seqs.g.faa -noreverse -minsize 500
```

8. Using 6 various alignment algorithms on 10 protein sequences

We use protein sequences of the same gene from the same organisms

clustalw

```
clustalw -INFILE=./data/processed_data/SUP35_10seqs.g.faa -OUTFILE=./data/raw_data/SUP35_10seqs.clustalw.faa -OUTPUT=FASTA -TYPE=pr
```

time:

```
 real    0m0,715s
user    0m0,715s
sys     0m0,000s
```

```
clustalo --infile=./data/processed_data/SUP35_10seqs.g.faa --outfile=./data/processed_data/SUP35_10seqs.clustalo.faa --verbose
```

time:

```
 real    0m0,625s
user    0m1,072s
sys     0m0,109s
```

muscle

```
muscle -in ./data/processed_data/SUP35_10seqs.g.faa -out ./data/processed_data/SUP35_10seqs_muscle.faa
```

time:

```
 real    0m0,358s
user    0m0,342s
sys     0m0,016s
```

mafft

```
mafft --auto ./data/processed_data/SUP35_10seqs.g.faa >./data/processed_data/SUP35_10seqs_mafft.faa
```

time:

```
 real    0m0,618s
user    0m0,576s
sys     0m0,094s
```

kalign

```
kalign <./data/processed_data/SUP35_10seqs.g.faa >./data/processed_data/SUP35_10seqs_kalign.faa
```

time:

```
 real    0m0,076s
user    0m0,176s
sys     0m0,016s
```

t_coffee

```
t_coffee -infile=./data/processed_data/SUP35_10seqs.g.faa -outfile=./data/processed_data/SUP35_10seqs_tcoffee.faa
```

time:

```
 real    0m15,731s
user    0m15,329s
sys     0m0,402s
```

prank

```
prank -d=./data/processed_data/SUP35_10seqs.g.faa -o=./data/processed_data/SUP35_10seqs_prank.faa
```

time:

```
 real    0m12,729s
user    0m12,461s
sys     0m0,317s
```

9. Comparation of the results on 10 protein sequences

The table containing real times and grapical representation of the results can be found in Supplementary (Table 1, Figures 11-15)
I think muscle is the best for this goal.

10. Add more alignments using muscle/mafft:

Here I align 2 more DNA sequences and add them to the files with 250 aligned DNA sequences using muscle of mafft.

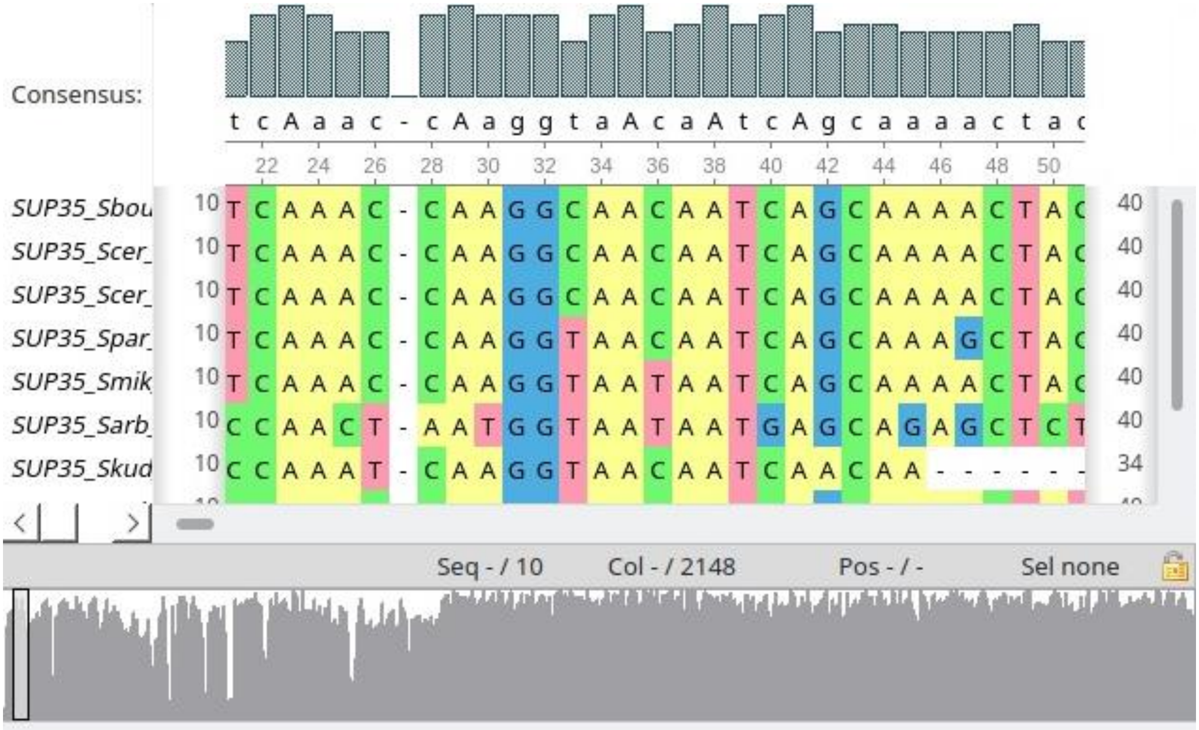
```
 muscle -in ./data/raw_data/SUP35_2addseqs.fa -out ./data/processed_data/SUP35_2addseqs_muscle.fa
muscle -profile -in1 ./data/processed_data/SUP35_250seqs_muscle.fa -in2 ./data/processed_data/SUP35_2addseqs_muscle.fa -out ./data/p
```

```
 mafft --auto ./data/raw_data/SUP35_2addseqs.fa > ./data/processed_data/SUP35_2addseqs_mafft.fa
mafft --add ./data/processed_data/SUP35_2addseqs_mafft.fa ./data/processed_data/SUP35_250seqs_mafft.fa > ./data/processed_data/SUP35
```

Supplementary

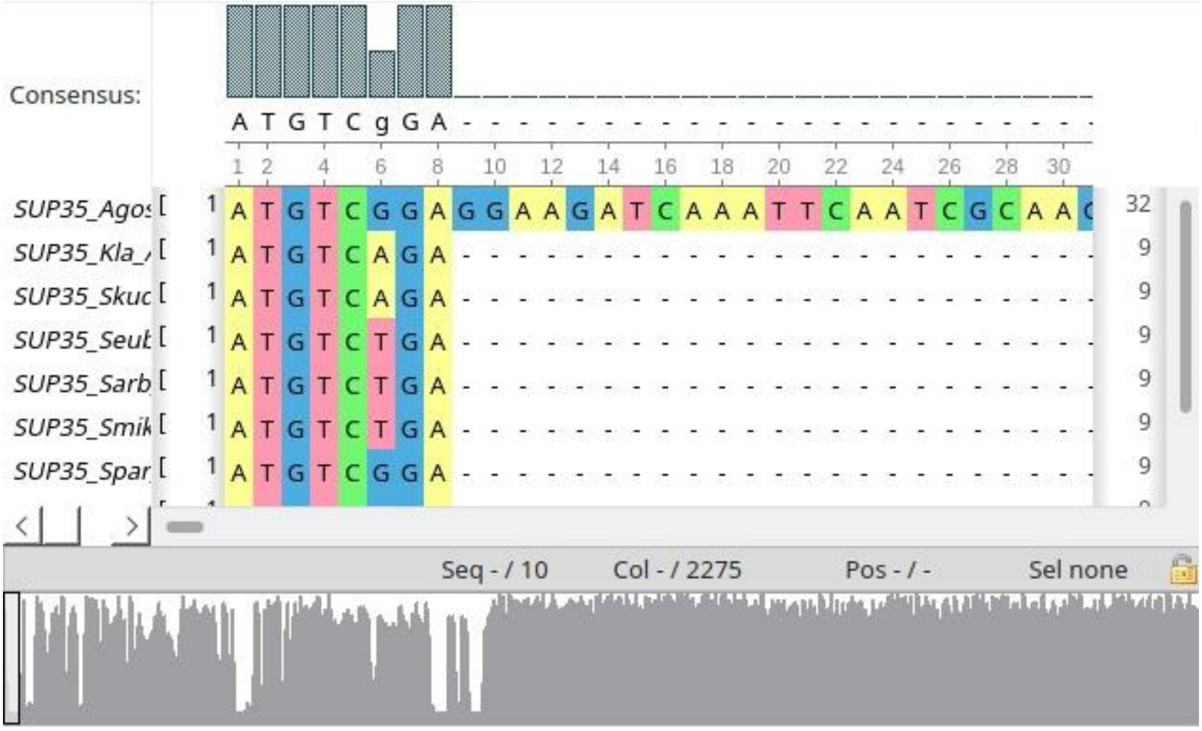
	CLUSTAL	MUSCLE	MAFFT	KALIGN	T-COFFEE	PRANK
10 DNA SEQUENCES	3,747s	4,742s	4,486s	0,242s	1m 31,287s	12,179s
250 DNA SEQUENCES	29m 3,728s	1m 20,447s	39,807s	3,912s	> 39m	2m 39,690s
10 PROTEIN SEQUENCES	0,715s	0,358s	0,618s	0,076s	15,731s	12,729s

Table 1. Real times for the alignments



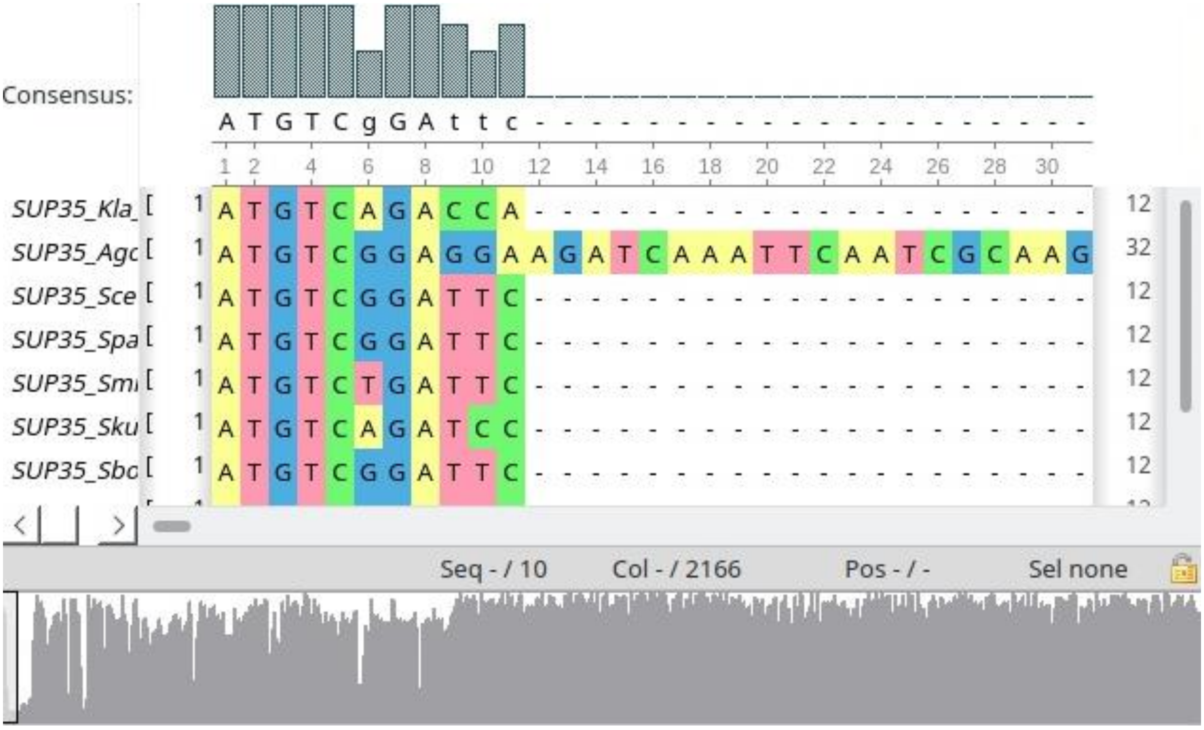
```
Consensus:     t c A a a c - c A a g g t a A c a A t c A g c a a a a c t a c
                 22  24  26  28  30  32  34  36  38  40  42  44  46  48  50
SUP35_Sbou  10 T C A A A C - C A A G G C A A C A A T C A G C A A A A C T A C  40
SUP35_Scer_ 10 T C A A A C - C A A G G C A A C A A T C A G C A A A A C T A C  40
SUP35_Scer_ 10 T C A A A C - C A A G G C A A C A A T C A G C A A A A C T A C  40
SUP35_Spar_ 10 T C A A A C - C A A G G T A A C A A T C A G C A A A G C T A C  40
SUP35_Smik  10 T C A A A C - C A A G G T A A T A A T C A G C A A A A C T A C  40
SUP35_Sarb_ 10 C C A A C T - A A T G G T A A T A A T G A G C A G A G C T C T  40
SUP35_Skud  10 C C A A A T - C A A G G T A A C A A T C A A C A A - - - - - -  34
                    Seq - / 10      Col - / 2148      Pos - / -      Sel none
```

Figure 1. Alignment of 10 DNA sequences using CLUSTAL



```
Consensus:     A T G T C g G A - - - - - - - - - - - - - - - - - - - - - - - -
               1 2   4   6   8   10  12  14  16  18  20  22  24  26  28  30
SUP35_Agos [ 1 A T G T C G G A G G A A G A T C A A A T T C A A T C G C A A C  32
SUP35_Kla_ [ 1 A T G T C A G A - - - - - - - - - - - - - - - - - - - - - - -   9
SUP35_Skuc [ 1 A T G T C A G A - - - - - - - - - - - - - - - - - - - - - - -   9
SUP35_Seut [ 1 A T G T C T G A - - - - - - - - - - - - - - - - - - - - - - -   9
SUP35_Sarb [ 1 A T G T C T G A - - - - - - - - - - - - - - - - - - - - - - -   9
SUP35_Smik [ 1 A T G T C T G A - - - - - - - - - - - - - - - - - - - - - - -   9
SUP35_Spar [ 1 A T G T C G G A - - - - - - - - - - - - - - - - - - - - - - -   9
                    Seq - / 10      Col - / 2275      Pos - / -      Sel none
```

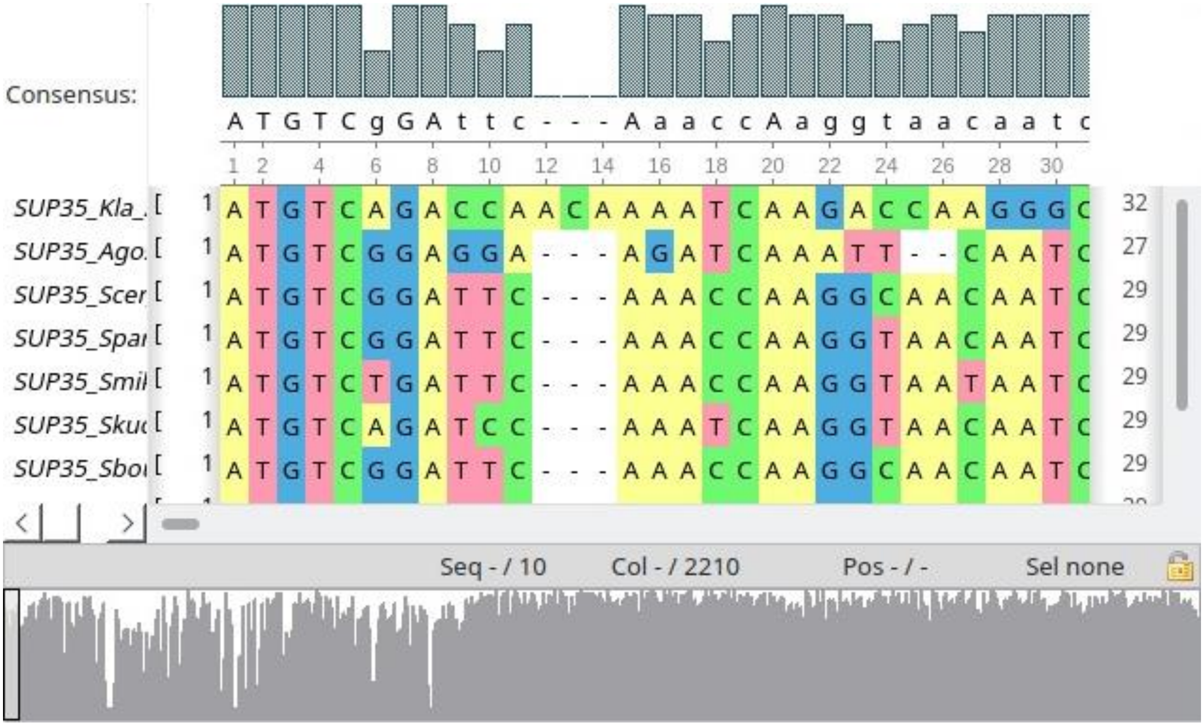
Figure 2. Alignment of 10 DNA sequences using MUSCLE



```
Consensus:   ATGTCgGAttc--------------------
               1 2   4   6   8  10  12  14  16  18  20  22  24  26  28  30
SUP35_Kla_ [ 1 ATGTCAGACCA--------------------  12
SUP35_Agc  [ 1 ATGTCGGAGGAAGATCAAATTCAATCGCAAG  32
SUP35_Sce  [ 1 ATGTCGGATTC--------------------  12
SUP35_Spa  [ 1 ATGTCGGATTC--------------------  12
SUP35_Smi  [ 1 ATGTCTGATTC--------------------  12
SUP35_Sku  [ 1 ATGTCAGATCC--------------------  12
SUP35_Sbo  [ 1 ATGTCGGATTC--------------------  12
```

Seq - / 10   Col - / 2166   Pos - / -   Sel none

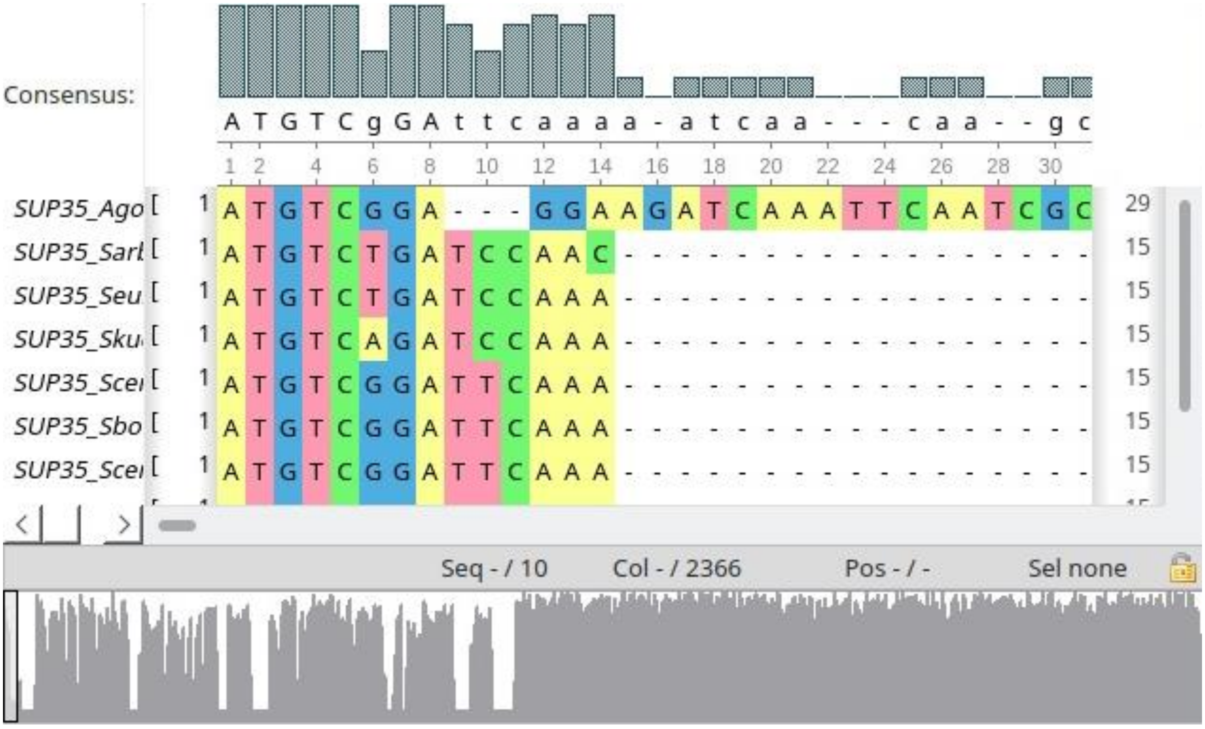
Figure 3. Alignment of 10 DNA sequences using MAFFT



```
Consensus:   ATGTCgGAttc---AaaccAaggtaacaatc
               1 2   4   6   8  10  12  14  16  18  20  22  24  26  28  30
SUP35_Kla_ [ 1 ATGTCAGACCAACAAAATCAAGACCAAGGGC  32
SUP35_Ago. [ 1 ATGTCGGAGGA---AGATCAAATT--CAATC  27
SUP35_Scer [ 1 ATGTCGGATTC---AAACCAAGGCAACAATC  29
SUP35_Spar [ 1 ATGTCGGATTC---AAACCAAGGTAACAATC  29
SUP35_Smik [ 1 ATGTCTGATTC---AAACCAAGGTAATAATC  29
SUP35_Skud [ 1 ATGTCAGATCC---AAATCAAGGTAACAATC  29
SUP35_Sboi [ 1 ATGTCGGATTC---AAACCAAGGCAACAATC  29
```

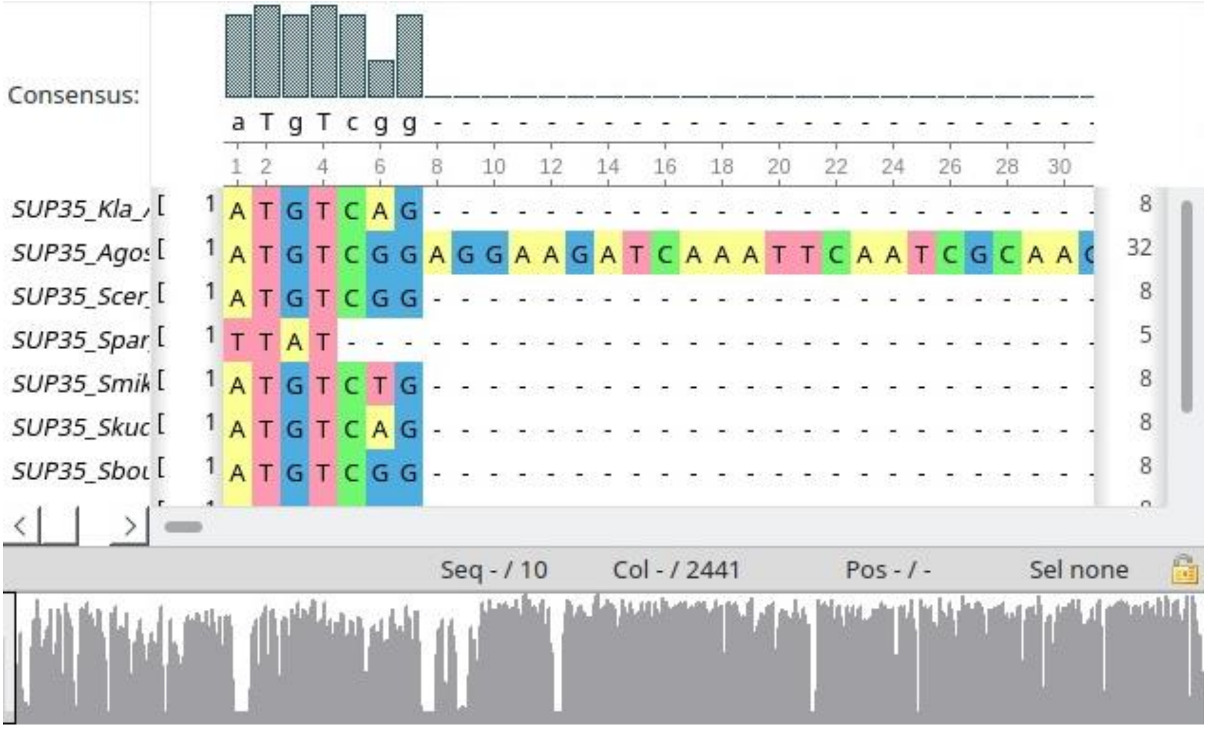
Seq - / 10   Col - / 2210   Pos - / -   Sel none

Figure 4. Alignment of 10 DNA sequences using T-COFFEE



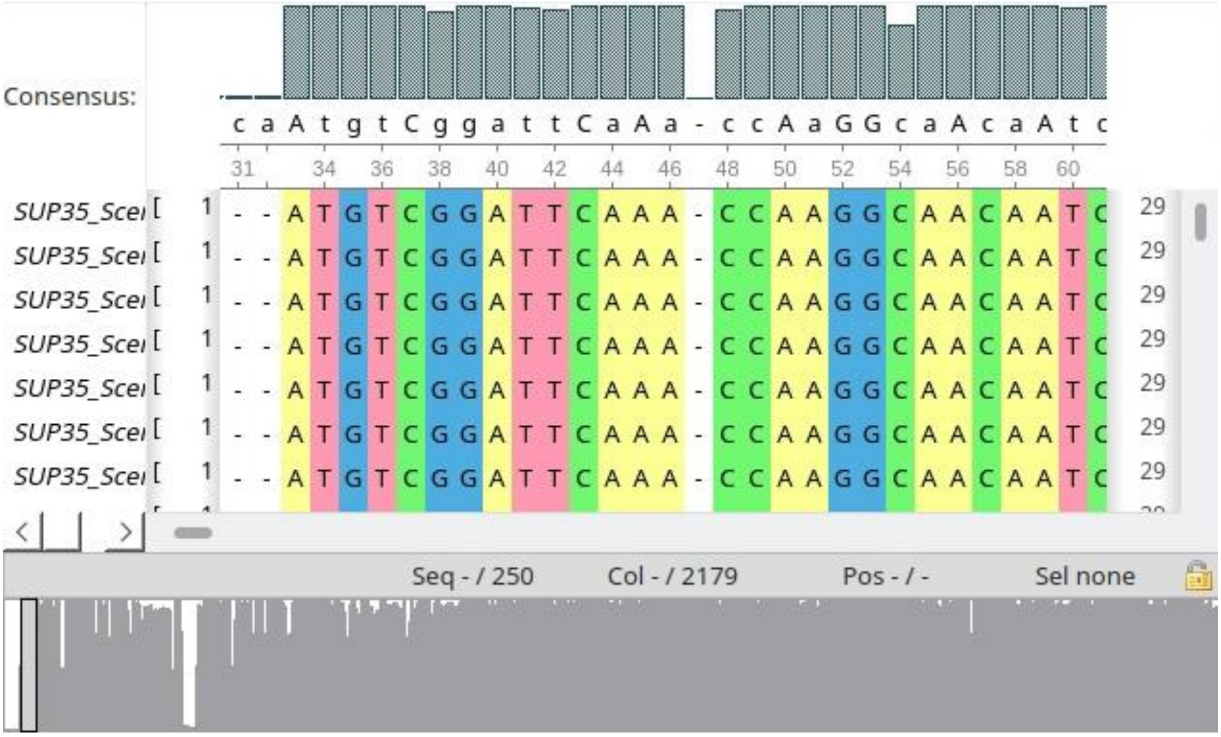
```
Consensus:   ATGTCgGAttcaaaa-atcaa---caa--gc

SUP35_Ago[   1 ATGTCGGA---GGAAGATCAAATTCAATCGC 29
SUP35_Sarb[  1 ATGTCTGATCCAAC----------------- 15
SUP35_Seu[   1 ATGTCTGATCCAAA----------------- 15
SUP35_Sku[   1 ATGTCAGATCCAAA----------------- 15
SUP35_Scer[  1 ATGTCGGATTCAAA----------------- 15
SUP35_Sbo[   1 ATGTCGGATTCAAA----------------- 15
SUP35_Scer[  1 ATGTCGGATTCAAA----------------- 15

Seq - / 10     Col - / 2366     Pos - / -     Sel none
```

Figure 5. Alignment of 10 DNA sequences using PRANK



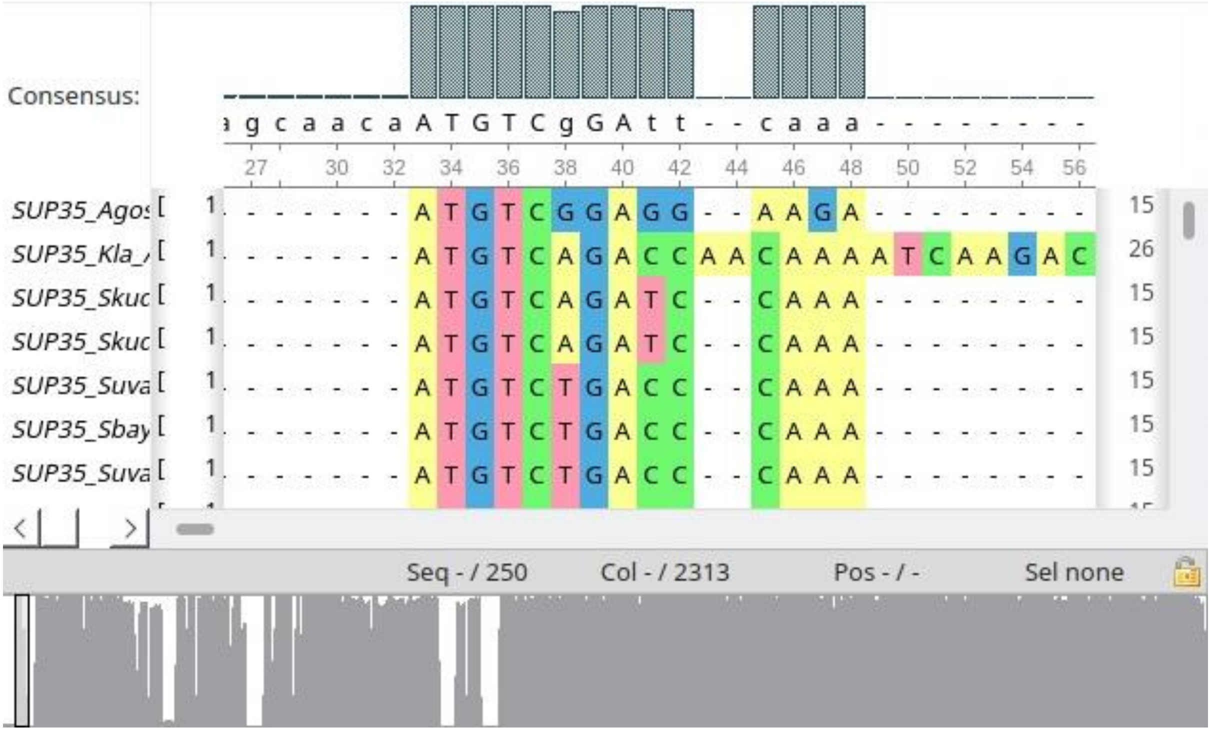
```
Consensus:   aTgTcgg------------------------

SUP35_Kla_[  1 ATGTCAG------------------------ 8
SUP35_Agos[  1 ATGTCGGAGGAAGATCAAATTCAATCGCAAC 32
SUP35_Scer[  1 ATGTCGG------------------------ 8
SUP35_Spar[  1 TTAT--------------------------- 5
SUP35_Smik[  1 ATGTCTG------------------------ 8
SUP35_Skuc[  1 ATGTCAG------------------------ 8
SUP35_Sbou[  1 ATGTCGG------------------------ 8

Seq - / 10     Col - / 2441     Pos - / -     Sel none
```

Figure 6. Strange alignment



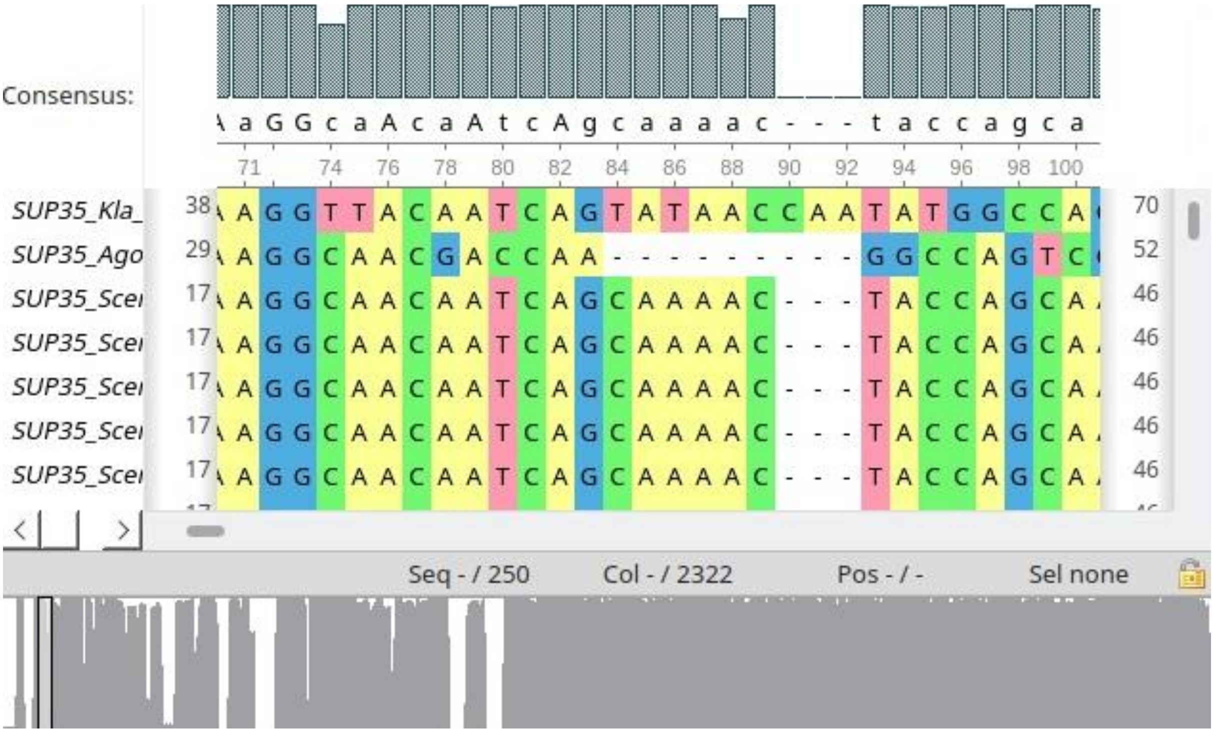
```
Consensus:    c a A t g t C g g a t t C a A a - c c A a G G c a A c a A t c
              31    34  36  38  40  42  44  46  48  50  52  54  56  58  60
SUP35_Scei[ 1 - - A T G T C G G A T T C A A A - C C A A G G C A A C A A T C 29
SUP35_Scei[ 1 - - A T G T C G G A T T C A A A - C C A A G G C A A C A A T C 29
SUP35_Scei[ 1 - - A T G T C G G A T T C A A A - C C A A G G C A A C A A T C 29
SUP35_Scei[ 1 - - A T G T C G G A T T C A A A - C C A A G G C A A C A A T C 29
SUP35_Scei[ 1 - - A T G T C G G A T T C A A A - C C A A G G C A A C A A T C 29
SUP35_Scei[ 1 - - A T G T C G G A T T C A A A - C C A A G G C A A C A A T C 29
SUP35_Scei[ 1 - - A T G T C G G A T T C A A A - C C A A G G C A A C A A T C 29
```

Figure 7. Alignment of 250 DNA sequences using CLUSTAL



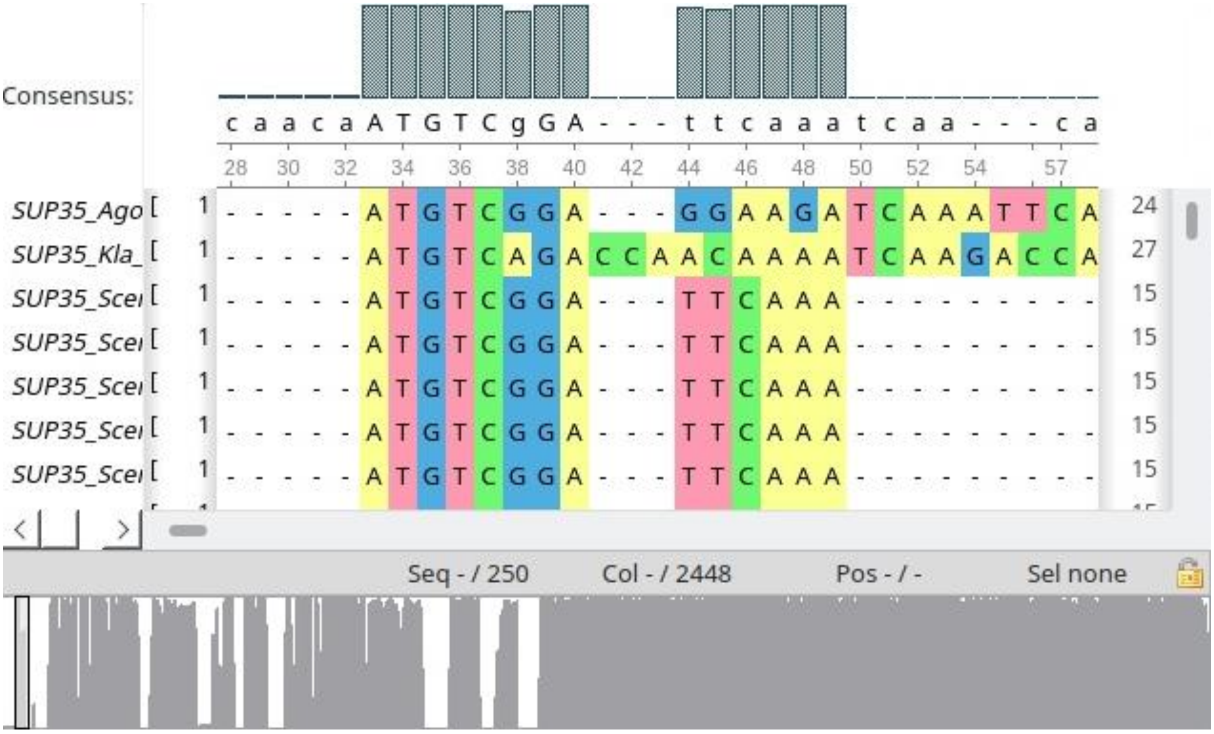
```
Consensus:    a g c a a c a A T G T C g G A t t - - c a a a - - - - - - - -
                27    30  32  34  36  38  40  42  44  46  48  50  52  54  56
SUP35_Agos[ 1 - - - - - - - A T G T C G G A G G - - A A G A - - - - - - - - 15
SUP35_Kla_/[ 1 - - - - - - - A T G T C A G A C C A A C A A A A T C A A G A C 26
SUP35_Skud[ 1 - - - - - - - A T G T C A G A T C - - C A A A - - - - - - - - 15
SUP35_Skud[ 1 - - - - - - - A T G T C A G A T C - - C A A A - - - - - - - - 15
SUP35_Suva[ 1 - - - - - - - A T G T C T G A C C - - C A A A - - - - - - - - 15
SUP35_Sbay[ 1 - - - - - - - A T G T C T G A C C - - C A A A - - - - - - - - 15
SUP35_Suva[ 1 - - - - - - - A T G T C T G A C C - - C A A A - - - - - - - - 15
```

Figure 8. Alignment of 250 DNA sequences using MUSCLE



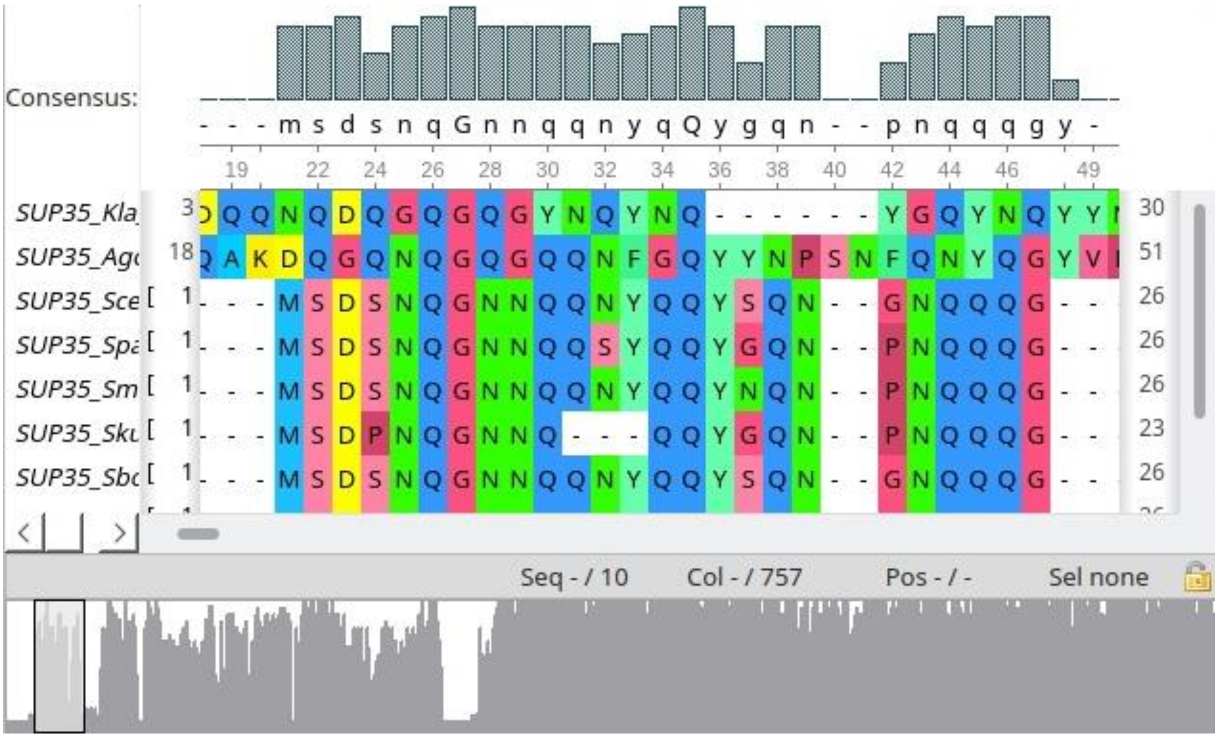
```
Consensus:     AaGGcaAcaAtcAgcaaaac---taccagca
                71  74 76 78 80 82 84 86 88 90 92 94 96 98 100
SUP35_Kla_  38 AAGGTTACAATCAGTATAACCAATATGGCCA  70
SUP35_Ago   29 AAGGCAACGACCAA---------GGCCAGTC  52
SUP35_Sce   17 AAGGCAACAATCAGCAAAAC---TACCAGCA  46
SUP35_Sce   17 AAGGCAACAATCAGCAAAAC---TACCAGCA  46
SUP35_Sce   17 AAGGCAACAATCAGCAAAAC---TACCAGCA  46
SUP35_Sce   17 AAGGCAACAATCAGCAAAAC---TACCAGCA  46
SUP35_Sce   17 AAGGCAACAATCAGCAAAAC---TACCAGCA  46

Seq - / 250     Col - / 2322     Pos - / -     Sel none
```

Figure 9. Alignment of 250 DNA sequences using MAFFT



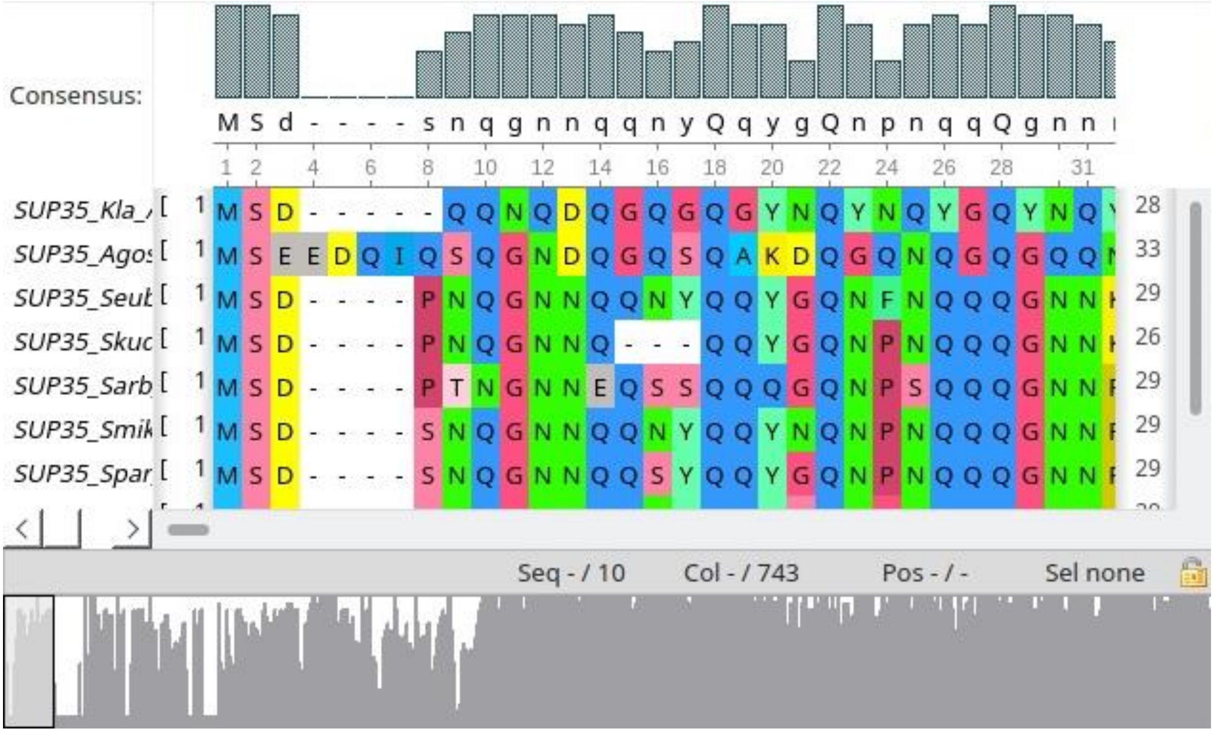
```
Consensus:     caacaATGTCgGA---ttcaaatcaa---ca
                28 30 32 34 36 38 40 42 44 46 48 50 52 54  57
SUP35_Ago   1  -----ATGTCGGA---GGAAGATCAAATTCA  24
SUP35_Kla_  1  -----ATGTCAGACCAACAAAATCAAGACCA  27
SUP35_Sce   1  -----ATGTCGGA---TTCAAA---------  15
SUP35_Sce   1  -----ATGTCGGA---TTCAAA---------  15
SUP35_Sce   1  -----ATGTCGGA---TTCAAA---------  15
SUP35_Sce   1  -----ATGTCGGA---TTCAAA---------  15
SUP35_Sce   1  -----ATGTCGGA---TTCAAA---------  15

Seq - / 250     Col - / 2448     Pos - / -     Sel none
```

Figure 10. Alignment of 250 DNA sequences using PRANK



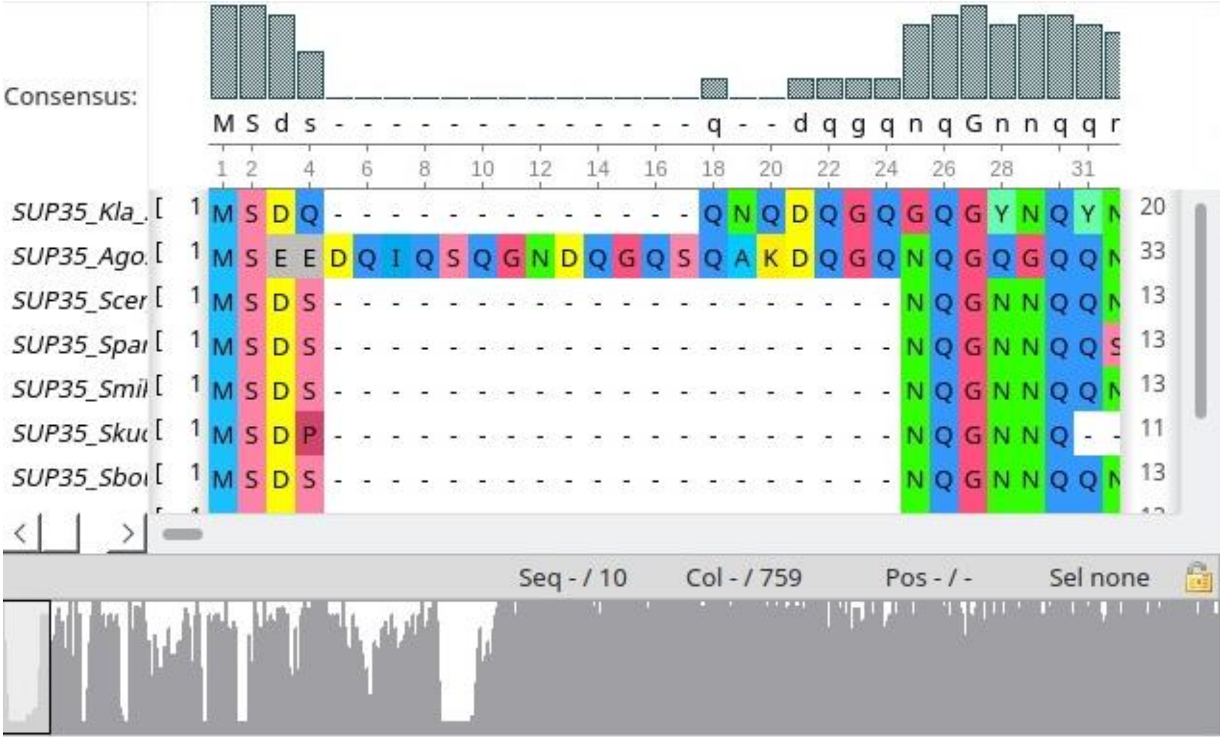
```
Consensus:   - - - m s d s n q G n n q q n y q Q y g q n - - p n q q q g y -
                19    22  24  26  28  30  32  34  36  38  40  42  44  46    49
SUP35_Kla     3 D Q Q N Q D Q G Q G Q G Y N Q Y N Q - - - - - - Y G Q Y N Q Y Y   30
SUP35_Ag     18 Q A K D Q G Q N Q G Q G Q Q N F G Q Y Y N P S N F Q N Y Q G Y V   51
SUP35_Sce [   1 - - - M S D S N Q G N N Q Q N Y Q Q Y S Q N - - G N Q Q Q G - -   26
SUP35_Spa [   1 - - - M S D S N Q G N N Q Q S Y Q Q Y G Q N - - P N Q Q Q G - -   26
SUP35_Sm  [   1 - - - M S D S N Q G N N Q Q N Y Q Q Y N Q N - - P N Q Q Q G - -   26
SUP35_Sku [   1 - - - M S D P N Q G N N Q - - - Q Q Y G Q N - - P N Q Q Q G - -   23
SUP35_Sbc [   1 - - - M S D S N Q G N N Q Q N Y Q Q Y S Q N - - G N Q Q Q G - -   26

Seq - / 10     Col - / 757     Pos - / -     Sel none
```

Figure 11. Alignment of 10 protein sequences using CLUSTAL



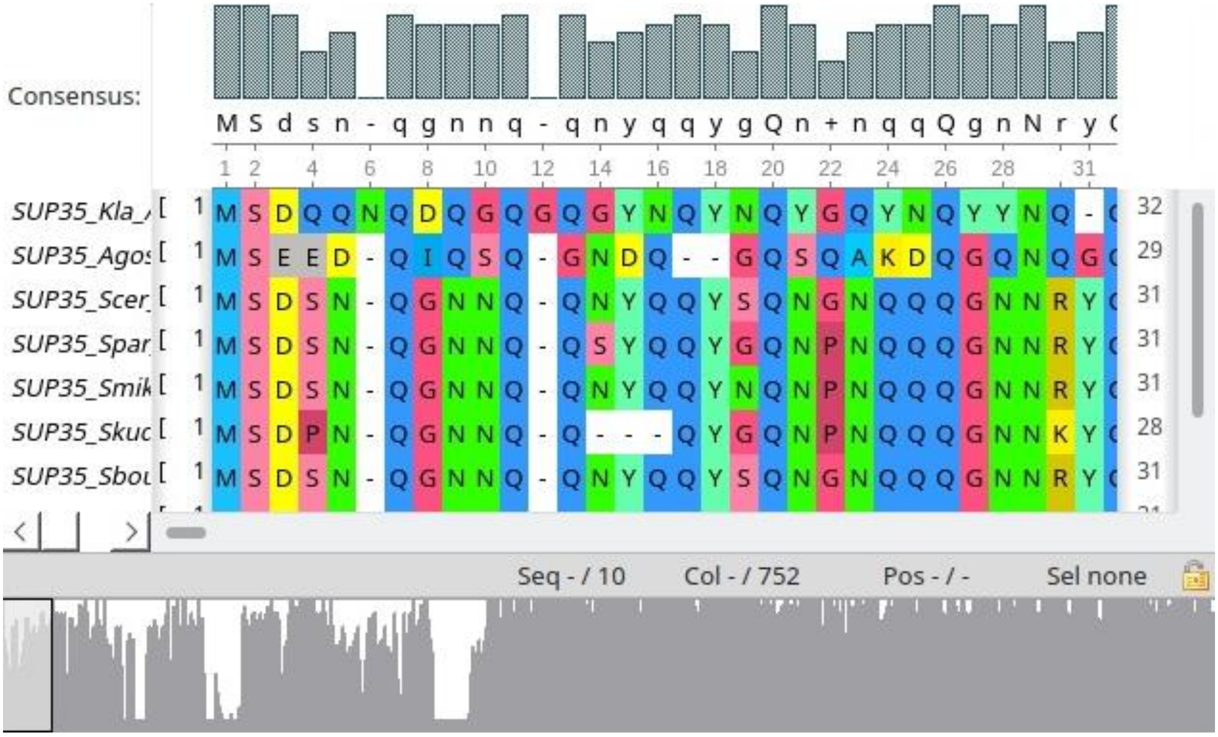
```
Consensus:    M S d - - - - s n q g n n q q n y Q q y g Q n p n q q Q g n n
              1 2   4   6   8   10  12  14  16  18  20  22  24  26  28    31
SUP35_Kla_ [ 1 M S D - - - - - Q Q N Q D Q G Q G Q G Y N Q Y N Q Y G Q Y N Q   28
SUP35_Agos [ 1 M S E E D Q I Q S Q G N D Q G Q S Q A K D Q G Q N Q G Q G Q Q   33
SUP35_Seub [ 1 M S D - - - - P N Q G N N Q Q N Y Q Q Y G Q N F N Q Q Q G N N   29
SUP35_Skuc [ 1 M S D - - - - P N Q G N N Q - - - Q Q Y G Q N P N Q Q Q G N N   26
SUP35_Sarb [ 1 M S D - - - - P T N G N N E Q S S Q Q Q G Q N P S Q Q Q G N N   29
SUP35_Smik [ 1 M S D - - - - S N Q G N N Q Q N Y Q Q Y N Q N P N Q Q Q G N N   29
SUP35_Spar [ 1 M S D - - - - S N Q G N N Q Q S Y Q Q Y G Q N P N Q Q Q G N N   29

Seq - / 10     Col - / 743     Pos - / -     Sel none
```

Figure 12. Alignment of 10 protein sequences using MUSCLE



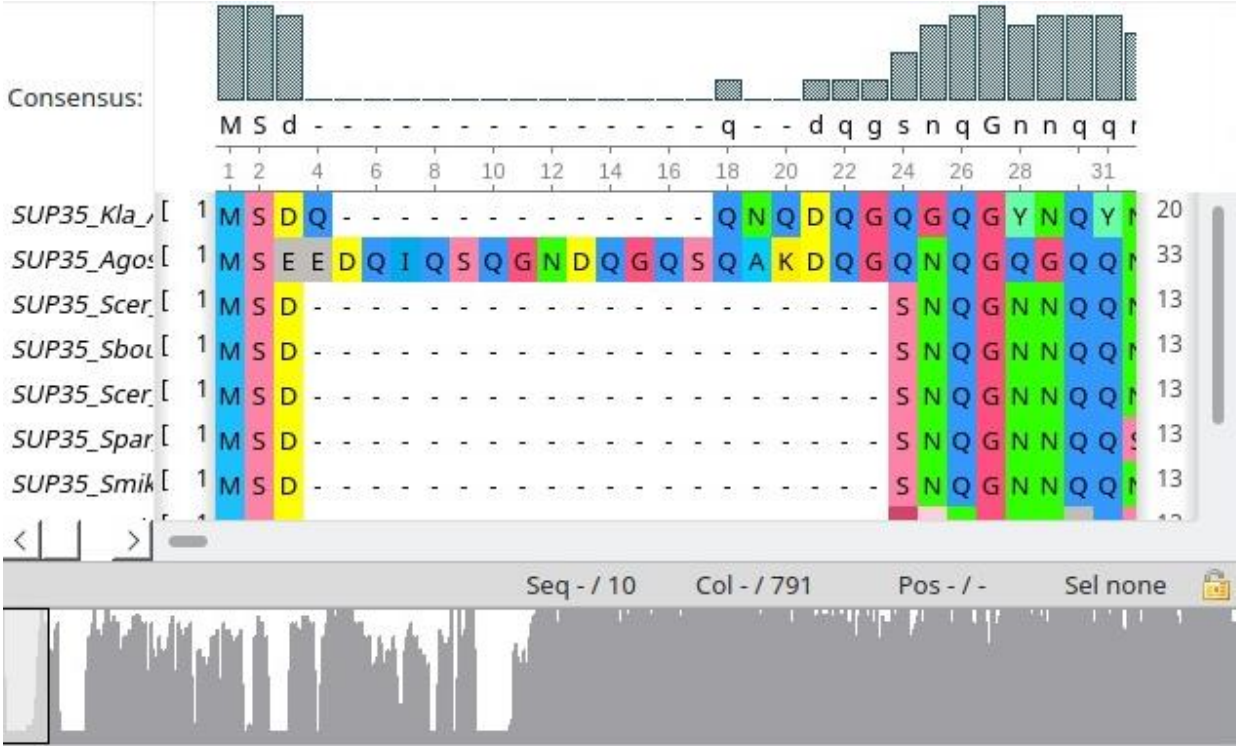
```
Consensus:  M S d s - - - - - - - - - - - - - q - - d q g q n q G n n q q r
            1 2   4   6   8   10  12  14  16  18  20  22  24  26  28    31
SUP35_Kla_. [ 1 M S D Q - - - - - - - - - - - - - Q N Q D Q G Q G Q G Y N Q Y N  20
SUP35_Ago.  [ 1 M S E E D Q I Q S Q G N D Q G Q S Q A K D Q G Q N Q G Q G Q Q N  33
SUP35_Scer  [ 1 M S D S - - - - - - - - - - - - - - - - - - - - N Q G N N Q Q N  13
SUP35_Spar  [ 1 M S D S - - - - - - - - - - - - - - - - - - - - N Q G N N Q Q S  13
SUP35_Smi   [ 1 M S D S - - - - - - - - - - - - - - - - - - - - N Q G N N Q Q N  13
SUP35_Sku   [ 1 M S D P - - - - - - - - - - - - - - - - - - - - N Q G N N Q - -  11
SUP35_Sbo   [ 1 M S D S - - - - - - - - - - - - - - - - - - - - N Q G N N Q Q N  13

Seq - / 10     Col - / 759     Pos - / -     Sel none
```

Figure 13. Alignment of 10 protein sequences using MAFFT



```
Consensus:  M S d s n - q g n n q - q n y q q y g Q n + n q q Q g n N r y (
            1 2   4   6   8   10  12  14  16  18  20  22  24  26  28    31
SUP35_Kla_  [ 1 M S D Q Q N Q D Q G Q G Q G Y N Q Y N Q Y G Q Y N Q Y Y N Q - (  32
SUP35_Agos  [ 1 M S E E D - Q I Q S Q - G N D Q - - G Q S Q A K D Q G Q N Q G (  29
SUP35_Scer  [ 1 M S D S N - Q G N N Q - Q N Y Q Q Y S Q N G N Q Q Q G N N R Y (  31
SUP35_Spar  [ 1 M S D S N - Q G N N Q - Q S Y Q Q Y G Q N P N Q Q Q G N N R Y (  31
SUP35_Smik  [ 1 M S D S N - Q G N N Q - Q N Y Q Q Y N Q N P N Q Q Q G N N R Y (  31
SUP35_Skuc  [ 1 M S D P N - Q G N N Q - Q - - - Q Y G Q N P N Q Q Q G N N K Y (  28
SUP35_Sbou  [ 1 M S D S N - Q G N N Q - Q N Y Q Q Y S Q N G N Q Q Q G N N R Y (  31

Seq - / 10     Col - / 752     Pos - / -     Sel none
```

Figure 14. Alignment of 10 protein sequences using T-COFFEE



```
Consensus:   M S d - - - - - - - - - - - - - - q - - d q g s n q G n n q q r
             1 2   4   6   8   10  12  14  16  18  20  22  24  26  28    31
SUP35_Kla_ [ 1 M S D Q - - - - - - - - - - - - - Q N Q D Q G Q G Q G Y N Q Y N  20
SUP35_Agos [ 1 M S E E D Q I Q S Q G N D Q G Q S Q A K D Q G Q N Q G Q G Q Q N  33
SUP35_Scer [ 1 M S D - - - - - - - - - - - - - - - - - - - - S N Q G N N Q Q N  13
SUP35_Sbou [ 1 M S D - - - - - - - - - - - - - - - - - - - - S N Q G N N Q Q N  13
SUP35_Scer [ 1 M S D - - - - - - - - - - - - - - - - - - - - S N Q G N N Q Q N  13
SUP35_Spar [ 1 M S D - - - - - - - - - - - - - - - - - - - - S N Q G N N Q Q S  13
SUP35_Smik [ 1 M S D - - - - - - - - - - - - - - - - - - - - S N Q G N N Q Q N  13
```

Seq - / 10   Col - / 791   Pos - / -   Sel none

Figure 15. Alignment of 10 protein sequences using PRANK