

19360859053

Hümeysra ÇİMEN

## BURSA TEKNİK ÜNİVERSİTESİ BİLGİSAYAR MÜHENDİSLİĞİ

```
7 1 !pip install biopython
sn. 2

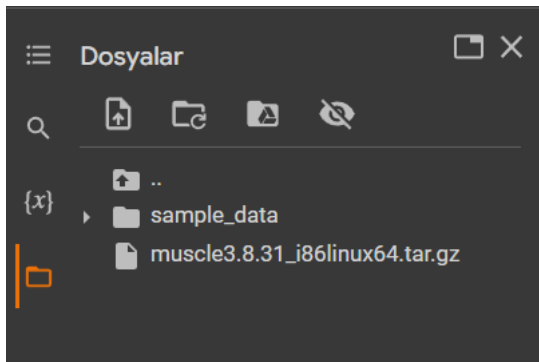
Looking in indexes: https://pypi.org/simple, https://us-python.pkg.dev/colab-wheels/public/simple/
Collecting biopython
  Downloading biopython-1.81-cp310-cp310-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (3.1 MB)
    3.1/3.1 MB 30.8 MB/s eta 0:00:00
Requirement already satisfied: numpy in /usr/local/lib/python3.10/dist-packages (from biopython) (1.22.4)
Installing collected packages: biopython
Successfully installed biopython-1.81
```

```
0 1 !wget https://www.drive5.com/muscle/downloads3.8.31/muscle3.8.31_i86linux64.tar.gz
sn. 2

--2023-05-13 14:19:01-- https://www.drive5.com/muscle/downloads3.8.31/muscle3.8.31_i86linux64.tar.gz
Resolving www.drive5.com (www.drive5.com)... 199.195.116.69
Connecting to www.drive5.com (www.drive5.com)|199.195.116.69|:443... connected.
HTTP request sent, awaiting response... 200 OK
Length: 487906 (476K) [application/x-gzip]
Saving to: 'muscle3.8.31_i86linux64.tar.gz'

muscle3.8.31_i86lin 100%[=====>] 476.47K 2.94MB/s in 0.2s

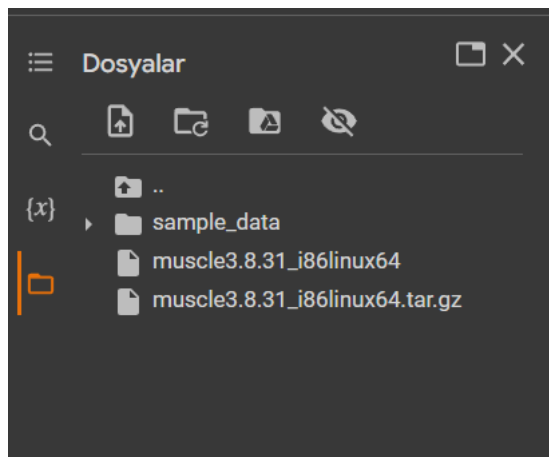
2023-05-13 14:19:02 (2.94 MB/s) - 'muscle3.8.31_i86linux64.tar.gz' saved [487906/487906]
```



```
[3] 1 !ls
sn. muscle3.8.31_i86linux64.tar.gz sample_data

[4] 1 !pwd
sn. 2
/content

[30] 1 !tar -xvzf muscle3.8.31_i86linux64.tar.gz
sn. 2
muscle3.8.31_i86linux64
```



```
✓ [58] 1 !/content/muscle3.8.31_i86linux64 -in seqs.fa -out seqs.afa
      2
```

MUSCLE v3.8.31 by Robert C. Edgar

<http://www.drive5.com/muscle>

This software is donated to the public domain.

Please cite: Edgar, R.C. Nucleic Acids Res 32(5), 1792-97.

```
seqs 670 seqs, max length 4921, avg length 2832
00:00:07 24 MB(6%) Iter 1 100.00% K-mer dist pass 1
00:00:07 24 MB(6%) Iter 1 100.00% K-mer dist pass 2
00:01:58 1166 MB(100%) Iter 1 100.00% Align node
00:01:58 1169 MB(100%) Iter 1 100.00% Root alignment
00:03:02 1171 MB(100%) Iter 2 100.00% Refine tree
00:03:02 1171 MB(100%) Iter 2 100.00% Root alignment
00:03:02 1171 MB(100%) Iter 2 100.00% Root alignment
00:13:52 1171 MB(100%) Iter 3 100.00% Refine biparts
00:24:44 1171 MB(100%) Iter 4 100.00% Refine biparts
```

✓  
0  
sn.

[39] 1 !./muscle3.8.31\_i86linux64

MUSCLE v3.8.31 by Robert C. Edgar

<http://www.drive5.com/muscle>

This software is donated to the public domain.

Please cite: Edgar, R.C. Nucleic Acids Res 32(5), 1792-97.

#### Basic usage

```
muscle -in <inputfile> -out <outputfile>
```

Common options (for a complete list please see the User Guide):

-in <inputfile>	Input file in FASTA format (default stdin)
-out <outputfile>	Output alignment in FASTA format (default stdout)
-diags	Find diagonals (faster for similar sequences)
-maxiters <n>	Maximum number of iterations (integer, default 16)
-maxhours <h>	Maximum time to iterate in hours (default no limit)
-html	Write output in HTML format (default FASTA)
-msf	Write output in GCG MSF format (default FASTA)
-clw	Write output in CLUSTALW format (default FASTA)
-clwstrict	As -clw, with 'CLUSTAL W (1.81)' header
-log[a] <logfile>	Log to file (append if -loga, overwrite if -log)
-quiet	Do not write progress messages to stderr
-version	Display version information and exit

Without refinement (very fast, avg accuracy similar to T-Coffee): -maxiters 2

Fastest possible (amino acids): -maxiters 1 -diags -sv -distance1 kbit20\_3

Fastest possible (nucleotides): -maxiters 1 -diags

✓  
6  
dk.

[9] 1 !./muscle3.8.31\_i86linux64 -in seqs.fa -out seqs.afa  
2  
3



MUSCLE v3.8.31 by Robert C. Edgar

<http://www.drive5.com/muscle>

This software is donated to the public domain.

Please cite: Edgar, R.C. Nucleic Acids Res 32(5), 1792-97.

seqs 515 seqs, max length 2829, avg length 2829

00:00:04	20 MB(5%)	Iter	1	100.00%	K-mer dist pass 1
00:00:04	20 MB(5%)	Iter	1	100.00%	K-mer dist pass 2
00:01:28	884 MB(100%)	Iter	1	100.00%	Align node
00:01:28	886 MB(100%)	Iter	1	100.00%	Root alignment
00:02:27	887 MB(100%)	Iter	2	100.00%	Refine tree
00:02:27	887 MB(100%)	Iter	2	100.00%	Root alignment
00:02:27	887 MB(100%)	Iter	2	100.00%	Root alignment
00:06:39	887 MB(100%)	Iter	3	100.00%	Refine biparts

```
✓ [10] 1 !./muscle3.8.31_i86linux64 -in seqs.fa -out seqs.afa -maxiters 2  
2  
dk.
```

MUSCLE v3.8.31 by Robert C. Edgar

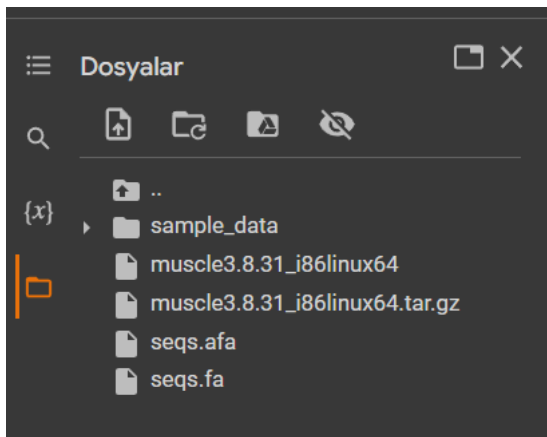
<http://www.drive5.com/muscle>

This software is donated to the public domain.

Please cite: Edgar, R.C. Nucleic Acids Res 32(5), 1792-97.

seqs 515 seqs, max length 2829, avg length 2829

00:00:03	20 MB(5%)	Iter	1	100.00%	K-mer dist pass 1
00:00:03	20 MB(5%)	Iter	1	100.00%	K-mer dist pass 2
00:01:26	884 MB(100%)	Iter	1	100.00%	Align node
00:01:26	886 MB(100%)	Iter	1	100.00%	Root alignment
00:02:24	887 MB(100%)	Iter	2	100.00%	Refine tree
00:02:24	887 MB(100%)	Iter	2	100.00%	Root alignment
00:02:24	887 MB(100%)	Iter	2	100.00%	Root alignment



```
✓ [10] 1 !./muscle3.8.31_i86linux64 -in seqs.fa -out seqs.afa -maxiters 2  
2  
dk.
```

MUSCLE v3.8.31 by Robert C. Edgar

<http://www.drive5.com/muscle>

This software is donated to the public domain.

Please cite: Edgar, R.C. Nucleic Acids Res 32(5), 1792-97.

seqs 515 seqs, max length 2829, avg length 2829

00:00:03	20 MB(5%)	Iter	1	100.00%	K-mer dist pass 1
00:00:03	20 MB(5%)	Iter	1	100.00%	K-mer dist pass 2
00:01:26	884 MB(100%)	Iter	1	100.00%	Align node
00:01:26	886 MB(100%)	Iter	1	100.00%	Root alignment
00:02:24	887 MB(100%)	Iter	2	100.00%	Refine tree
00:02:24	887 MB(100%)	Iter	2	100.00%	Root alignment
00:02:24	887 MB(100%)	Iter	2	100.00%	Root alignment

```
✓ [11] 1 !./muscle3.8.31_i86linux64 -in seqs.fa -out seqs.afa -maxiters 1 -diags -sv -distance1 kbit20_3  
1  
dk.
```

MUSCLE v3.8.31 by Robert C. Edgar

<http://www.drive5.com/muscle>

This software is donated to the public domain.

Please cite: Edgar, R.C. Nucleic Acids Res 32(5), 1792-97.

seqs 515 seqs, max length 2829, avg length 2829

00:00:00 13 MB(3%) Iter 1 100.00% K-bit distance matrix

00:01:27 884 MB(100%) Iter 1 100.00% Align node

00:01:27 885 MB(100%) Iter 1 100.00% Root alignment

```
✓ [12] 1 !./muscle3.8.31_i86linux64 -in seqs.fa -out seqs.afa -maxiters 1 -diags  
1  
dk.
```

MUSCLE v3.8.31 by Robert C. Edgar

<http://www.drive5.com/muscle>

This software is donated to the public domain.

Please cite: Edgar, R.C. Nucleic Acids Res 32(5), 1792-97.

seqs 515 seqs, max length 2829, avg length 2829

00:00:04 20 MB(5%) Iter 1 100.00% K-mer dist pass 1

00:00:04 20 MB(5%) Iter 1 100.00% K-mer dist pass 2

00:01:27 884 MB(100%) Iter 1 100.00% Align node

00:01:27 886 MB(100%) Iter 1 100.00% Root alignment

```
✓ [13] 1 !./muscle3.8.31_i86linux64 -in seqs.fa -out seqs.afa -maxiters 1 -diags1 -sv  
1  
dk.
```

MUSCLE v3.8.31 by Robert C. Edgar

<http://www.drive5.com/muscle>

This software is donated to the public domain.

Please cite: Edgar, R.C. Nucleic Acids Res 32(5), 1792-97.

seqs 515 seqs, max length 2829, avg length 2829

00:00:04 20 MB(5%) Iter 1 100.00% K-mer dist pass 1

00:00:04 20 MB(5%) Iter 1 100.00% K-mer dist pass 2

00:01:27 884 MB(100%) Iter 1 100.00% Align node

00:01:27 886 MB(100%) Iter 1 100.00% Root alignment

## 2.7 Pipelining

```
[14] 1 !./muscle3.8.31_i86linux64 < seqs.fa > seqs.afa
```

MUSCLE v3.8.31 by Robert C. Edgar

<http://www.drive5.com/muscle>

This software is donated to the public domain.

Please cite: Edgar, R.C. Nucleic Acids Res 32(5), 1792-97.

```
- 515 seqs, max length 2829, avg length 2829
00:00:04    20 MB(5%) Iter  1 100.00% K-mer dist pass 1
00:00:04    20 MB(5%) Iter  1 100.00% K-mer dist pass 2
00:01:27   884 MB(100%) Iter  1 100.00% Align node
00:01:27   886 MB(100%) Iter  1 100.00% Root alignment
00:02:25   887 MB(100%) Iter  2 100.00% Refine tree
00:02:25   887 MB(100%) Iter  2 100.00% Root alignment
00:02:25   887 MB(100%) Iter  2 100.00% Root alignment
00:06:38   887 MB(100%) Iter  3 100.00% Refine biparts
```

## 2.8 Refining an existing alignment

```
1 !./muscle3.8.31_i86linux64 -in seqs.afa -out refined.afa -refine
```

MUSCLE v3.8.31 by Robert C. Edgar

<http://www.drive5.com/muscle>

This software is donated to the public domain.

Please cite: Edgar, R.C. Nucleic Acids Res 32(5), 1792-97.

```
00:04:20    34 MB(8%) Iter  1 100.00% Refine biparts
```

## 2.9 Using a pre-computed guide tree

```
[18] 1 !./muscle3.8.31_i86linux64 -in seqs.fa -out seqs.afa -usetree mytree.phy
```



MUSCLE v3.8.31 by Robert C. Edgar

<http://www.drive5.com/muscle>

This software is donated to the public domain.

Please cite: Edgar, R.C. Nucleic Acids Res 32(5), 1792-97.

seqs 515 seqs, max length 2829, avg length 2829

\*\*\*\*\* WARNING \*\*\*\*\*

You specified the -usetree option.

Note that a good evolutionary tree may NOT be a good guide tree for multiple alignment. For more details, please refer to the user guide. To disable this warning, use -usetree\_nowarn <treefilename>.

\*\*\* ERROR \*\*\* User tree must be rooted

çözmek için adımlar:

seqs.fa dosya alınan adress:<https://github.com/gtonkinhill/fastbaps/blob/master/inst/extdata/seqs.fa>



mytree.php dosyası alınan adress:

<https://github.com/gtonkinhill/fastbaps/blob/master/inst/extdata/seqs.fa.treefile>

## WARNING

✓  
0  
sn.

```
[19] 1 !./muscle3.8.31_i86linux64 -in seqs.fa -out seqs.afa -usetree mytree.phy  
2
```

MUSCLE v3.8.31 by Robert C. Edgar

<http://www.drive5.com/muscle>

This software is donated to the public domain.

Please cite: Edgar, R.C. Nucleic Acids Res 32(5), 1792-97.

seqs 515 seqs, max length 2829, avg length 2829

\*\*\*\*\* WARNING \*\*\*\*\*

You specified the -usetree option.

Note that a good evolutionary tree may NOT be a good guide tree for multiple alignment. For more details, please refer to the user guide. To disable this warning, use -usetree\_nowarn <treefilename>.

\*\*\* ERROR \*\*\* User tree must be rooted

- 📁 ..
- ▶ 📁 sample\_data
  - 📄 muscle3.8.31\_i86linux64
  - 📄 muscle3.8.31\_i86linux64.tar.gz
  - 📄 mytree.phy
  - 📄 refined.afa
  - 📄 seqs.afa
  - 📄 seqs.fa





```
1 !./muscle3.8.31_i86linux64 -in seqs.fa -out seqs.afa -usetree mytree.phy
2
```

MUSCLE v3.8.31 by Robert C. Edgar

<http://www.drive5.com/muscle>

This software is donated to the public domain.

Please cite: Edgar, R.C. Nucleic Acids Res 32(5), 1792-97.

seqs 515 seqs, max length 2829, avg length 2829

```
00:00:04    20 MB(5%) Iter  1 100.00% K-mer dist pass 1
00:00:04    20 MB(5%) Iter  1 100.00% K-mer dist pass 2
00:01:28   884 MB(100%) Iter  1 100.00% Align node
00:01:28   886 MB(100%) Iter  1 100.00% Root alignment
00:02:29   887 MB(100%) Iter  2 100.00% Refine tree
00:02:30   887 MB(100%) Iter  2 100.00% Root alignment
00:02:30   887 MB(100%) Iter  2 100.00% Root alignment
00:06:41   887 MB(100%) Iter  3 100.00% Refine biparts
```

## 2.10 Profile-profile alignment;

one.afa:seqs.afa two.afa:refine.afa

✓  
7  
sn.



```
1 !./muscle3.8.31_i86linux64 -profile -in1 seqs.afa -in2 two.afa -out refined.afa
```



MUSCLE v3.8.31 by Robert C. Edgar

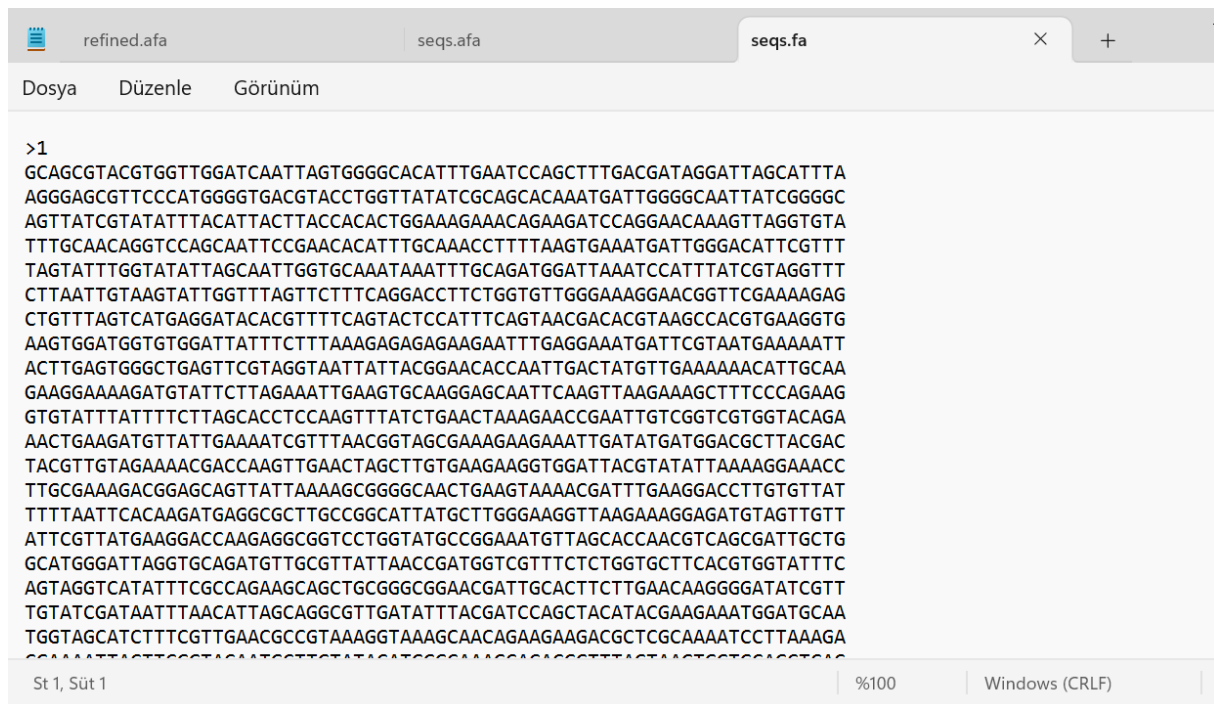
<http://www.drive5.com/muscle>

This software is donated to the public domain.

Please cite: Edgar, R.C. Nucleic Acids Res 32(5), 1792-97.

```
00:00:00    10 MB(2%) Reading seqs.afa
00:00:00    13 MB(3%) 515 seqs 2829 cols
00:00:00    13 MB(3%) Reading two.afa
00:00:00    17 MB(4%) 672 seqs 2829 cols
00:00:07    19 MB(5%) Aligning profiles
00:00:07    35 MB(9%) Building output
00:00:07    39 MB(9%) Writing output
```





## 2.10 Profile-profile alignment;

one.afa:seqs.afa two.afa:refine.afa

```
!./muscle3.8.31_i86linux64 -profile -in1 seqs.afa -in2 two.afa -out refined.afa
```

MUSCLE v3.8.31 by Robert C. Edgar

<http://www.drive5.com/muscle>

This software is donated to the public domain.

Please cite: Edgar, R.C. Nucleic Acids Res 32(5), 1792-97.

```
00:00:00    10 MB(2%) Reading seqs.afa
00:00:00    13 MB(3%) 515 seqs 2829 cols
00:00:00    13 MB(3%) Reading two.afa
00:00:00    17 MB(4%) 672 seqs 2829 cols
00:00:07    19 MB(5%) Aligning profiles
00:00:07    35 MB(9%) Building output
00:00:07    39 MB(9%) Writing output
```

## 2.11 Adding sequences to an existing alignment

```
!./muscle3.8.31_i86linux64 -profile -in1 existing_aln.afa -in2 new_seq.fa -out combined.afa
```

MUSCLE v3.8.31 by Robert C. Edgar

<http://www.drive5.com/muscle>

This software is donated to the public domain.

Please cite: Edgar, R.C. Nucleic Acids Res 32(5), 1792-97.

\*\*\* ERROR \*\*\* Cannot open 'existing\_aln.afa' errno=2

```
!./muscle3.8.31_i86linux64 -in new_seqs.fa -out new_seqs.afa
```

```
!./muscle3.8.31_i86linux64 -profile -in1 existing_aln.afa -in2 new_seqs.fa -out combined.afas
```