

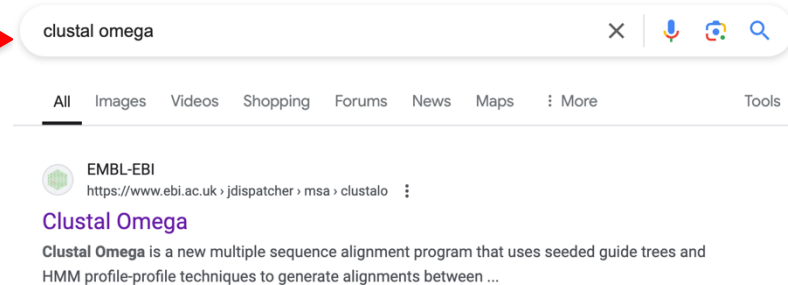
# Clustal Omega: Align two or more DNA or Protein Sequences

**Step 1.** Internet search “clustal omega” →

**Step 2.** Or click this link:

<https://www.ebi.ac.uk/jdispatcher/msa/clustalo>

**Step 3.** You'll get this page:

A screenshot of the Clustal Omega web interface. The header features the EMBL-EBI logo and navigation links: Home, Services, Research, Training, About us, and a search icon. The main banner displays "Clustal Omega" and "Multiple Sequence Alignment (MSA)" against a background of binary code and a DNA helix. Below the banner, there are links for "Job Dispatcher", "Help & Privacy", "Your Jobs", and "Input form", along with a "Feedback" button. A yellow message box states: "Welcome to the new Job Dispatcher website. We'd love to hear your feedback about the new webpages!". The main content area explains that Clustal Omega is a new multiple sequence alignment program using seeded guide trees and HMM profile-profile techniques, and that it can align up to 4000 sequences or a maximum file size of 4 MB. Under the heading "Input sequence", there is a "Sequence Type" section with radio buttons for "Protein" (selected), "DNA", and "RNA". Below this is a large text box labeled "Paste your sequence here - or use the example sequence". At the bottom of the input section, there are buttons for "Choose File" (with "No file chosen" text), "Use the example", "Clear sequence", and a link for "More example inputs". The "Parameters" section is partially visible at the bottom, showing an "OUTPUT FORMAT" dropdown menu.

## Step 4. Find some sequence data to align.

Here are 3 DNA sequences in FASTA format from <http://kelleybioinfo.org/algorithms/data/DAl2.txt> (The “Sequence files” link on the Basics page has more information on FASTA.)

>NucSeq1

```
ATGAACGACGAAACACAATTTACAAATAAGGCCAACGAAATTATCCGTTTGGCCCAGAAATTGGCTCAGGATCACAGACATGCTCAGTTACAACCAATT  
CACTTACTTGCTGCATTTGTTGAGCCAAACGAGGATGGTTTC
```

>NucSeq2

```
ATGGCTGATTATCCTTTTACTGACAAAGCCGCAAAGACATTGTCTGATGCGTACTCAATTGCACAATCTTATGGTCATTACAAATTAACCCCTATTAC  
ATTGCTGCTGCTCTTTTGTCCGACAGTGACAGTAACGGTAC
```

>NucSeq3

```
ATGAACGACGAAACGAAGTTTACGAACAAAGCTCTCGATATCATCACCATTGCACAGAACTAGCACAGGACCACCAGCATTCGACGCTGGTGCCTCTA  
CACGTGCTTGCAGCGTTTCGTAGAGACACCTGCTGATGGTAG
```

**Step 5.** Select DNA because these are DNA sequences, then Copy and Paste the data into the input window.

**Clustal Omega**  
Multiple Sequence Alignment (MSA)

Job Dispatcher Help & Privacy Your Jobs **Input form** [Feedback](#)

Welcome to the new **Job Dispatcher** website. We'd love to hear your [feedback](#) about the new webpages!

Clustal Omega is a new multiple sequence alignment program that uses seeded guide trees and HMM profile-profile techniques to generate alignments between three or more sequences. For the alignment of two sequences please instead use our pairwise sequence alignment tools. This tool can align up to 4000 sequences or a maximum file size of 4 MB.

**Input sequence ⓘ**

**Sequence type**

☐ Protein ☒ DNA ☐ RNA

**Paste your sequence here - or use the example sequence**

```
>NucSeq1  
ATGAACGACGAAACACAATTTACAAATAAGGCCAACGAAATTATCCGTTTGGCCCAGAAATTGGCTCAGGATCACAGACATGCTCAGTTACAACCAATTCACCTTACTTGCTGCATTTGTTGAGCCAAACGAGGATGGTTTC  
>NucSeq2  
ATGGCTGATTATCCTTTTACTGACAAAGCCGCAAAGACATTGTCTGATGCGTACTCAATTGCACAATCTTATGGTCATTACAAATTAACCCCTATTACATTGCTGCTGCTCTTTTGTCCGACAGTGACAGTAACGGTAC  
>NucSeq3  
ATGAACGACGAAACGAAGTTTACGAACAAAGCTCTCGATATCATCACCATTGCACAGAACTAGCACAGGACCACCAGCATTCGACGCTGGTGCCTCTACACGTGCTTGCAGCGTTTCGTAGAGACACCTGCTGATGGTAG
```

[Choose File](#) No file chosen [Use the example](#) [Clear sequence](#) [More example inputs](#)

**Step 6.** Scroll Down to the bottom of the page and click the Submit button.

Paste your sequence here - or use the example sequence

```
>NucSeq1
ATGAACGACGAAACACAATTTACAAATAAGGCCAACGAAATTATCCGTTTGGCCAGAAATTGGCTCAGGATCACAGACATGCTCAGTTACAACCAATTCATTACTTGC
TGCATTTGTTGAGCCAAACGAGGATGGTTC
>NucSeq2
ATGGCTGATTATCCTTTTACTGACAAAGCCGCAAGACATTGTCTGATGCGTACTCAATTGCACAATCTTATGGTCATTACAATTAACCCCTATTACATTGCTGCTGCTC
TTTTGTCCGACAGTGACAGTAACGGTAC
>NucSeq3
```

Choose File No file chosen

Use the example

Clear sequence

More example inputs

Parameters

OUTPUT FORMAT ⓘ

ClustalW with character counts

More options ▾

Submit

Title

Clustal Omega's job

Submit

**Step 7.** Then wait....

Close YOUR JOB IS RUNNING

Please note that results can only be retrieved for jobs submitted within the last seven days.

Job ID: clustalo-120240916-051251-0539-8209655-p1m

**RUNNING**

● ● ● ●

Your Jobs Close

Use this service, please consider citing the following publication. The embedded bioinformatics application can be found in its bio.tools record.

## Step 8. Interpreting the output.

After waiting a bit, the output will appear in the browser like this:

View alignments with different color labels

**Clustal Omega**  
Multiple Sequence Alignment (MSA)

Job Dispatcher Help & Privacy Your Jobs Input form

Welcome to the new **Job Dispatcher** website. We'd love to hear your [feedback](#) about the new webpages!

Results for Job ID clustalo-I20240916-051301-0450-4984400-p1m [Copy](#)

[Feedback](#)

[Resubmission](#)

Tool Output Alignments Guide Tree Phylogenetic Tree Results Viewers Result Files Submission Details

### The sequence alignment:

Tool output

[Download](#)

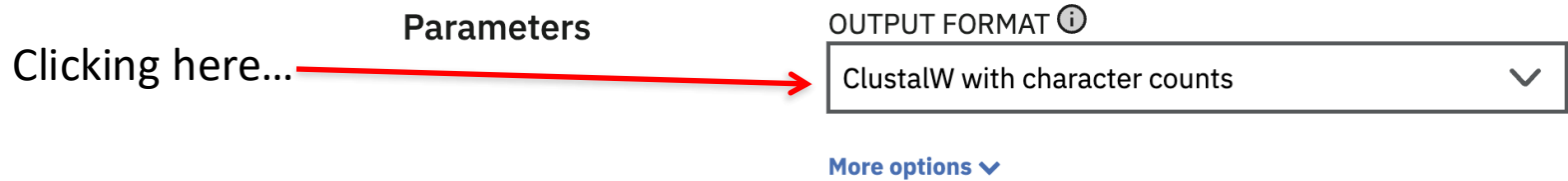
CLUSTAL 0(1.2.4) multiple sequence alignment

```
NucSeq2 ---ATGGCTGATTATCCTTTTACTGACAAAGCCGCAAGACATTGTCTGATGCGTACTCA 57
NucSeq1 ATGAACGACGAAACACAATTTACAAATAAGGCCAACGAAATTATCCGTTTGCCCGAGAAA 60
NucSeq3 ATGAACGACGAAACGAAGTTTACGAACAAAGCTCTCGATATCATCACCATTGCACAGAAA 60
      * * * * *
NucSeq2 ATTGCACAATCTTATGGTCATTACAAATTAACCCCTATTACATTGCTGCTGCTCTTTTG 117
NucSeq1 TTGGCTCAGGATCACAGACATGCTCAGTTACAACCAATTCACCTACTTGCTGCTATTGTT 120
NucSeq3 CTAGCACAGGACCACGACATTGACGCTGGTGCCTCTACACGTGCTTGACGCGTTCGTA 120
      * * * * *
NucSeq2 TCCGACAGTGACAGTAACGGTAC 140
NucSeq1 GAGCCAAACGAGGATG---GTTC 140
NucSeq3 GAGACACCTGCTGATG---GTAG 140
      * * *
```

The Clustal format in “interleaved”: these **alignments show 60 alignment positions for all sequences**, then go to the next 60 until there is no more alignment.

A asterisk “\*” indicates all the sequences have the same nucleotide.  
Fully conserved.

## Clustal Omega Options

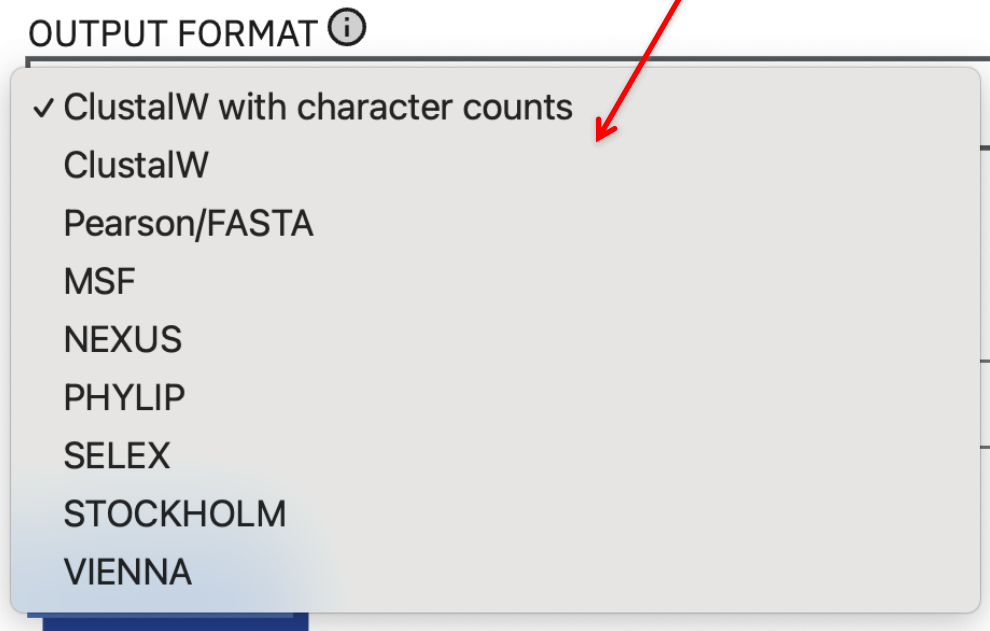


...will allow you to change alignment settings.

Most important for most people is how to change output formats.

Parameters

Submit



# Aligning Protein Sequences

**Step 1.** Set to Protein (default).

**Step 2.** Paste in Data (or choose file).

**Step 3.** Submit.

Clustal Omega is a new multiple sequence alignment program that uses seeded guide trees and HMM profile-profile techniques to generate alignments between three or more sequences. For the alignment of two sequences please instead use our pairwise sequence alignment tools. This tool can align up to 4000 sequences or a maximum file size of 4 MB.

Input sequence ⓘ

Sequence Type

☒ Protein ☐ DNA ☐ RNA

Paste your sequence here - or use the example sequence

```
QRRGTGDKAEAWATGVKYDANDIYIATFYSETRNMTVPVSGGFANKTQNFEAVIQYQDFGLRPSLGYVLSKGGKDIEGVGSEDLVNYIDVGATYYFNKNMSAFVDYKINQLD
SDNTLGINDDDIVAIGLTYQF
>PhoEseedSen2
MNKSTLAIIVVSIASASVHAAEVYNKNGNKLVDYGVKVKAMHYMSDYDSKDGQSYVRFGFKGETQINDQLTGYGRWEAEFAGNKAESDSSQKTRLAFAGLKLKDIGSFD
YGRNLGALYDVEAWTDMFPEFGGSSAQTDNFMKTRASGLATYRNTDFFGIVDGLDLTLQYQGNEDRDVKKQNGDGFGTSVSYDFGGSDFAVSGAYTLDRTREQNLQ
RRGTGDKAEAWATGVKYDANDIYIATFYSETRNMTVPVSGGFANKTQNFEAVIQYQDFGLRPSLGYVLSKGGKDIEGVGSEDLVNYIDVGAIYYFNKNMSAFVDYKINQLDS
DNTLGINDDDIVAIGLTYQF
```

Choose File No file chosen

Use the example

Clear sequence

More example inputs

Parameters

OUTPUT FORMAT ⓘ

ClustalW with character counts

More options ▾

Submit

Title

Clustal Omega's job

Submit

# Aligning Protein Sequences

## Step. 4 Wait for alignment

LCseedSf1	MKKLTVAISAVAASVLMAMSAQAAEIYNKDSNKLDLYGKVNAKHVFSSNDADDGDTTYVR	60
PhoEseedEco2	MKMKKSTLALVVMGIVASVSVQAAEIYNKDGKLDVYGKVKAMHYMSDNDSDKDGDQSYIR	60
PhoEseedEco1	MKMKKSTLALVVMGIVASVSVQAAEIYNKDGKLDVYGKVKAMHYMSDNDSDKDGDQSYIR	60
PhoEseedEco4	--MKKSTLALVVMGIVASVSVQAAEIYNKDGKLDVYGKVKAMHYMSDNDSDKDGDQSYIR	58
PhoEseedSen1	--MNKSTLAI-VVSIIASASVHAAEVYNKNGNKLDVYGKVKAMHYMSDYDSKDGDQSYVR	57
PhoEseedSen2	--MNKSTLAI-VVSIIASASVHAAEVYNKNGNKLDVYGKVKAMHYMSDYDSKDGDQSYVR	57
	. ::: . .:: : *.:***:***:.*****:***:* **:* . *:.*** :*:*	
LCseedSf1	LGFKGETQINDQLTGFGQWEYEFKGNRAESQGSSKDKTRLAFAGLKFGDYGSIDYGRNYG	120
PhoEseedEco2	FGFKGETQINDQLTGYGRWEAEFAGNKAESDT-AQQKTRLAFAGLKDYKDLGSFDYGRNLG	119
PhoEseedEco1	FGFKGETQINDQLTGYGRWEAEFAGNKAESDT-AQQKTRLAFAGLKDYKDLGSFDYGRNLG	119
PhoEseedEco4	FGFKGETQINDQLTGYGRWEAEFAGNKAESDT-AQQKTRLAFAGLKDYKDLGSFDYGRNLG	117
PhoEseedSen1	FGFKGETQINDQLTGYGRWEAEFAGNKAESDS-SQQKNRLAFAGLKLKDIGSFDYGRNLG	116
PhoEseedSen2	FGFKGETQINDQLTGYGRWEAEFAGNKAESDS-SQQKTRLAFAGLKLKDIGSFDYGRNLG	116
	:*****:*** ** **:*:***: :*:***** * **:****** *	

### Meaning of symbols in protein alignment:

**\*** (**Asterix**) positions with a single, fully conserved residue.

**:** (**colon**) positions with conservation between amino acid groups of similar properties.

**.** (**period**) positions with conservation between amino acid groups of weakly similar properties.