

1. [MAFFT](#) is a multiple sequence alignment program for unix-like operating systems. It offers a range of multiple alignment methods, L-INS-i (accurate; for alignment of <~200 sequences), FFT-NS-2 (fast; for alignment of <~30,000 sequences), *etc.* [Manual](#)

2. Using in UBCG:

3. Input:

Fasta format.

[example1 \(LSU rRNA\)](#)

```
>X02729 Methanococcus vannielli. #
tattctattaccctaccctggggaatggcttggcttgaaacgccgatgaaggacgtggtaa
gctgcgataagcctagggcgaggcgcaacagcctttgaacctaggatttccgaatgggact
tcctacttttgtaatccgtaaggattggtaacgcgggggattgaagcatcttagtaccg
caggaaaaaagaatcaactgagattccgttagtagaggcgattgaacacggatcagggcaa
actgaatcccttcggggagatgtggtgttatagggccttcttttcgcctgttgagaaaag
ctgaattgaactgaacatcacactatagagaggtgaagtcgcgtgaagcgaatcattc ,
```

[example2 \(protein\)](#)

```
>A.thaliana 67208.m00002
MDTRFPFSPAEVSKVRVVFQFGLSPDEIRQMSVIHVEHSETTEKGPVKVGGLSOTRLGTI
DRKVKCETCMANMAECPGHFGYLELAKPMYHVGFMKTVLSIMRCVCFNCSKILADEVCRS
LFRQAMKIKNPKNRLKKILDACKNKTCDGGDDIDDVQSHSTDEPVKKSRRGGCGAQQPKL
TIEGMKIMIAEYKIQRKKNDEPDQLPEPAERKQTLGADRVLSVLKRISDADCQLLGFNPKF
```

The type of input sequences (amino acid or nucleotide) is automatically recognized.

4. Output:

--clustalout

Output format: clustal format. Default: off (fasta format)

--inputorder

Output order: same as input. Default: on

--reorder

Output order: aligned. Default: off (inputorder)

--treeout

Guide tree is output to the *input.tree* file. Default: off

--quiet

Do not report progress. Default: off

5. Examples

online

MAFFT-L-INS-i Result

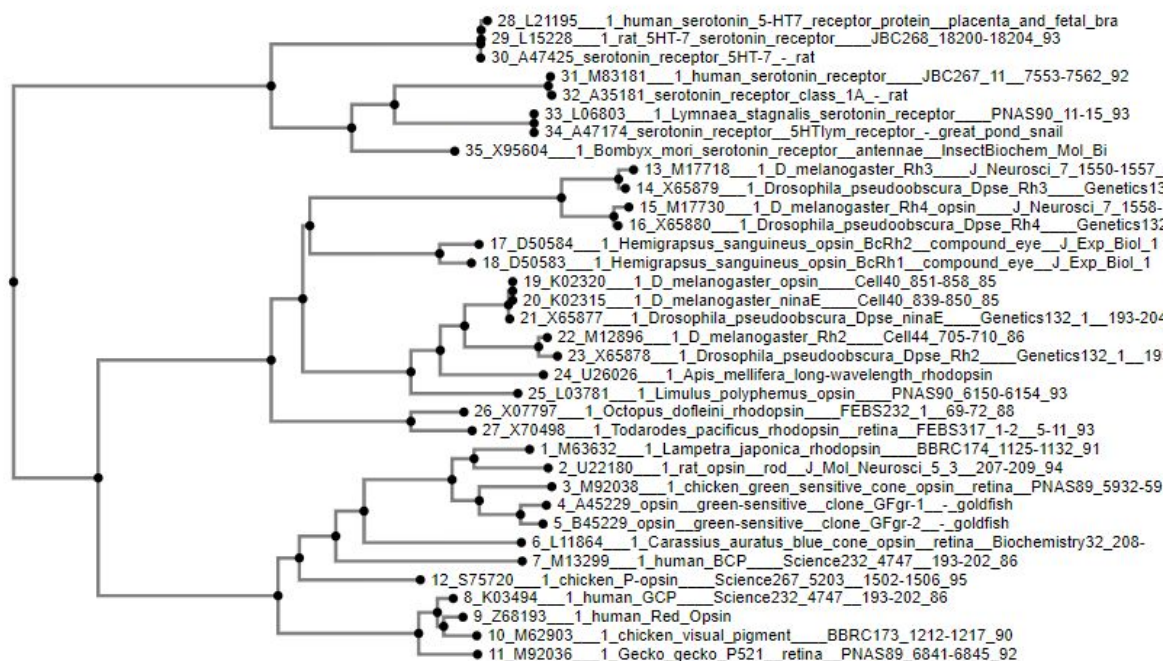
CLUSTAL format alignment by MAFFT (v7.419)

```

1LYLA      F-----NDELNRNRE-KLAALRQQGVAFPNDFRRDHTSDQLHEEFDAK-----D
c|p2      MP-----
1B8AA      M-----YRTHYS-----SEITEELNG-----
1ASZB      E-----DTAKDNYG-KLPLIQSRSDRTG--QKRVKFVDLDEAKDSD-----
c|p1      M-----RTEYCGQL-----
1ADJA      -----
j|p2      M-----
j|p2      L-----RDNMK-----
c|p1      MPVITLPDGSQRHYD-----HAVSPMDVALDIG-----
c|p1      M-----RT-----
j|p2      L-----
1ATIA      -----
j|p2      M-----
1SESA      M--VD-LKRLRQEPE-----
c|p1      M--LD-PNLLRNEPD-----
1PYSA      R-----
y|Pyrococcus M-----RLGYNEKLVLLKLAEL-----KNATVEELIEKTNLDQVAVMRALLTLQ

1LYLA      NQELESINI-----EVSVAG-----RMMTRRIM---GKASFVTILQDVGGRILQLYV
c|p2      -----
1B8AA      -----Q-----KVKVAG-----WWEVKDL---GGIKFLWIRDRDGIQVITA
1ASZB      -----K-----EVLFR-----RVHNTRQQG--ATLAFLLTRQQASLIQGLV
c|p1      --RLSHVGQ-----QVILCG-----WVNRRDL---GSLIFIDMRDREGIVQVFF
1ADJA      -----
j|p2      -----
j|p2      -----MLLIHSDYLEFEAKEKTKIAEETENL-----
c|p1      -PGLA-----KACIAGRVNG---ELVDACDLIENDAQLSIITAKDEEGL-----

```



linux:

```
$ mafft --auto --quiet --treeout --parttree /home/lubov/hw2.fasta
```

result in input.tree

```
|(((((((
1_M63632__1_Lampetra_japonica_rhodopsin____BBRC174_1125-1132_91_
:0.15650,
2_U22180__1_rat_opsin__rod__J_Mol_Neurosci_5_3__207-209_94_
:0.15650):0.03668,(
3_M92038__1_chicken_green_sensitive_cone_opsin__retina__PNAS89_5932-5936_9
:0.16193,(
4_A45229_opsin__green-sensitive__clone_GFgr-1__-_goldfish
:0.05650,
5_B45229_opsin__green-sensitive__clone_GFgr-2__-_goldfish
:0.05650):0.10543):0.03125):0.24762,
6_L11864__1_Carassius_auratus_blue_cone_opsin__retina__Biochemistry32_208-
:0.44080):0.03973,
7_M13299__1_human_BCP____Science232_4747__193-202_86_
:0.48053):0.04203,((((
```

can be viewed using <http://phylo.io/>

