

# Package ‘muscle’

August 29, 2013

**Type** Package

**Title** Multiple Sequence Alignment

**Version** 3.8.31-1

**Date** 2012-10-05

**Author** Algorithm by Robert C. Edgar. R port by Alex T. Kalinka.

**Maintainer** Alex T. Kalinka <alex.t.kalinka@gmail.com>

**Description** MUSCLE performs multiple sequence alignments of nucleotide or amino acid sequences.

**License** Unlimited

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

**LazyLoad** yes

**LazyData** yes

**NeedsCompilation** yes

**Repository** CRAN

**Date/Publication** 2013-03-29 15:14:48

## R topics documented:

muscle-package . . . . .	2
muscle . . . . .	3
print.muscle . . . . .	4
read.fasta . . . . .	6
umax . . . . .	7
write.fasta . . . . .	7
<b>Index</b>	<b>9</b>

---

muscle-package

*Multiple Sequence Alignment*

---

## Description

MUSCLE performs multiple sequence alignments of nucleotide and amino acid sequences.

## Details

Details about the algorithm can be found on the MUSCLE website:

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

## Author(s)

Algorithm: Robert C. Edgar

R port: Alex T. Kalinka <[alex.t.kalinka@gmail.com](mailto:alex.t.kalinka@gmail.com)>

## References

Edgar, R.C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* **32**, 1792-1797.

## See Also

[muscle](#), [umax](#)

## Examples

```
## Align sequences from a fasta file.
## Not run:
aln <- muscle(seqs = "unaligned.fasta")

## End(Not run)

## Align sequences from a fasta object.
## Not run:
aln <- muscle(seqs = umax)

## End(Not run)
```

---

muscle	<i>Multiple Sequence Alignment</i>
--------	------------------------------------

---

**Description**

MUSCLE performs multiple sequence alignments of nucleotide or amino acid sequences.

**Usage**

```
muscle(seqs, out = NULL, quiet = FALSE, ...)
```

**Arguments**

seqs	Either a character string naming a FASTA file containing unaligned nucleotide or amino acid sequences, or an object of class <code>fasta</code> .
out	A character string naming the file to write aligned sequences to. If <code>NULL</code> then output is returned to the user as an R object of class <code>muscle</code> . Defaults to <code>NULL</code> .
quiet	Logical, whether to prevent printing of the progress of the algorithm to the screen. Defaults to <code>FALSE</code> .
...	Arguments (options and flags) for the MUSCLE algorithm (see Details).

**Details**

Arguments for the MUSCLE algorithm are either options or flags. Options take a variety of values, but flags are always logicals (`TRUE` or `FALSE`). All options and flags are explained in detail on the MUSCLE website: [http://www.drive5.com/muscle/muscle\\_userguide3.8.html](http://www.drive5.com/muscle/muscle_userguide3.8.html)

**Value**

Output is either written to a FASTA file (if `out` is not `NULL`), or saved to an object of class `muscle`, which is a list containing the following components:

1. `seqs` A data frame of aligned sequences. The first column contains sequence names, and the second column contains aligned sequences.
2. `length` The number of positions in the alignment,
3. `num` The number of sequences.

**Note**

For further details see the MUSCLE website: <http://www.drive5.com/muscle/>

**Author(s)**

Algorithm by Robert C. Edgar. Ported into R by Alex T. Kalinka <[alex.t.kalinka@gmail.com](mailto:alex.t.kalinka@gmail.com)>

## References

Edgar, R.C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* **32**, 1792-1797.

## See Also

[print.muscle](#), [read.fasta](#), [write.fasta](#), [umax](#)

## Examples

```
## Align sequences from a fasta file.
## Not run:
aln <- muscle(seqs = "unaligned.fasta")

## End(Not run)

## Align sequences from a fasta object.
aln <- muscle(seqs = umax)

## Align sequences from a fasta object and write them to a fasta file.
## Not run:
muscle(seqs = umax, out = "out.afa")

## End(Not run)

## Switch on the diags flag.
aln <- muscle(seqs = umax, diags = TRUE)
```

---

print.muscle	<i>Print Alignment Regions</i>
--------------	--------------------------------

---

## Description

This function prints aligned sequences for specific locations and sequences.

## Usage

```
## S3 method for class 'muscle'
print(x, ...)
```

## Arguments

x	An object of class muscle.
...	Alignment positions and/or sequences (see Details).

**Details**

Additional arguments for print:

1. from: the first position (integer)
2. to: the last position (integer)
3. num: the number of sequences to display (integer)
4. seqs: specific sequences to display (character vector).

**Value**

Prints aligned sequences to the screen.

**Author(s)**

Alex T. Kalinka <alex.t.kalinka@gmail.com>

**Examples**

```
## Align sequences.
aln <- muscle(seqs = umax)

## Print the beginning of the alignment to the screen.
print(aln)

## Print a longer stretch of the alignment.
print(aln, from = 1, to = 179)

## Print the last 150 positions in the alignment.
num <- aln$length
print(aln, from = num-150, to = num)

## Print a portion of the alignment for the first 5 sequences.
print(aln, from = 156, to = 289, num = 5)

## Print a portion of the alignment for a specific set of sequences.
print(aln, from = 156, to = 289, seqs = c("Sus_scrofa", "Homo_sapiens", "Felis_catus"))

## Print interleaved aligned sequences to a file.
## Not run:
sink(file = "aln.il", type = "output")
print(aln, from = 1, to = aln$length)
sink()

## End(Not run)
```

---

read.fasta	<i>Read a FASTA file into R</i>
------------	---------------------------------

---

**Description**

Reads sequence files in the FASTA format into an R object

**Usage**

```
read.fasta(file)
```

**Arguments**

file            A character string naming a fasta file.

**Value**

An object of class `fasta`, which is a list containing the following elements:

1. `seqs` A data frame in which the first column contains sequence names and the second column contains sequences.
2. `num` The number of sequences.

**Author(s)**

Alex T. Kalinka <[alex.t.kalinka@gmail.com](mailto:alex.t.kalinka@gmail.com)>

**See Also**

[write.fasta](#)

**Examples**

```
## Read in unaligned sequences.  
## Not run:  
seqs <- read.fasta(file = "unaligned.fasta")  
  
## End(Not run)
```

---

umax	<i>Unaligned MAX sequences</i>
------	--------------------------------

---

**Description**

Unaligned DNA sequences of the MAX gene for 31 mammalian species.

**Usage**

```
umax
```

**Format**

An object of class fasta.

**Source**

<http://www.ensembl.org/index.html>

**References**

Wagner, AJ et al. (1992) Expression, regulation, and chromosomal localization of the Max gene. *Proc Natl Acad Sci USA* **89**: 3111-3115.

---

write.fasta	<i>Write a FASTA file</i>
-------------	---------------------------

---

**Description**

Writes a sequence object to a FASTA file format

**Usage**

```
write.fasta(aln, file)
```

**Arguments**

aln	An object of class muscle or fasta.
file	A character string naming a file to write to.

**Value**

Writes a fasta file as a side-effect.

**Author(s)**

Alex T. Kalinka <[alex.t.kalinka@gmail.com](mailto:alex.t.kalinka@gmail.com)>

**See Also**[read.fasta](#)**Examples**

```
## Write out aligned sequences.  
## Not run:  
write.fasta(aln, file = "out.afa")  
  
## End(Not run)
```



# Index

## \*Topic **datasets**

umax, [7](#)

muscle, [2](#), [3](#)

muscle-package, [2](#)

print.muscle, [4](#), [4](#)

read.fasta, [4](#), [6](#), [8](#)

umax, [2](#), [4](#), [7](#)

write.fasta, [4](#), [6](#), [7](#)