

# Package ‘alignfigR’

July 5, 2018

**Title** Visualizing Multiple Sequence Alignments with 'ggplot2'

**Version** 0.1.1

**Description** Create extensible figures of multiple sequence alignments, using the 'ggplot2' plotting engine. 'alignfigR' will create a baseline figure of a multiple sequence alignment which can be fully customized to the user's liking with standard 'ggplot2' features.

**Depends** R (>= 3.1.1)

**Imports** ggplot2 (>= 2.1.0)

**License** GPL-2

**LazyData** true

**Suggests** knitr

**VignetteBuilder** knitr

**URL** <https://github.com/sjspielman/alignfigR>

**BugReports** <https://github.com/sjspielman/alignfigR/issues>

**Repository** CRAN

**Collate** 'alignfigR.R'

**NeedsCompilation** no

**RoxygenNote** 6.0.1

**Author** Stephanie J. Spielman [aut, cre]

**Maintainer** Stephanie J. Spielman <spielman@rowan.edu>

**Date/Publication** 2018-07-05 13:50:06 UTC

## R topics documented:

|                  |   |
|------------------|---|
| alignfigR        | 2 |
| define_palette   | 2 |
| extract_subalign | 3 |
| plot_alignment   | 3 |
| read_alignment   | 4 |

|              |          |
|--------------|----------|
| <b>Index</b> | <b>6</b> |
|--------------|----------|

---

|           |                   |
|-----------|-------------------|
| alignfigR | <i>alignfigR.</i> |
|-----------|-------------------|

---

**Description**

Creating figures multiple sequence alignments with ggplot2.

---

|                |                              |
|----------------|------------------------------|
| define_palette | <i>Define color palette.</i> |
|----------------|------------------------------|

---

**Description**

This function sets up, either using default or user-specified options, the color-coding scheme used to plot sequences.

**Usage**

```
define_palette(inpalette, uniques)
```

**Arguments**

|           |                                                                                                            |
|-----------|------------------------------------------------------------------------------------------------------------|
| inpalette | Either a user-specified named-array of colors or flag for default options (dna, rna, protein, random, ...) |
| uniques   | Unique characters found in alignment. Used to create the random color scheme.                              |

**Value**

Color palette named-array.

**Examples**

```
palette <- define_palette("DNA", c("A", "G", "T"))  
palette <- define_palette("protein")
```

---

|                  |                                             |
|------------------|---------------------------------------------|
| extract_subalign | <i>Extract subset of sequence alignment</i> |
|------------------|---------------------------------------------|

---

### Description

This function builds a data frame to plot an alignment from a specified subset of the full alignment.

### Usage

```
extract_subalign(seqs, plot_step = 1, tlist = c(), clist = c(),
  texcl = FALSE, cexcl = FALSE)
```

### Arguments

|           |                                                                            |
|-----------|----------------------------------------------------------------------------|
| seqs      | Sequence list, as parsed by the function ‘read_alignment’                  |
| plot_step | Step size for alignment block. Default, 1.                                 |
| tlist     | Array of taxa intended to restrict figure to or to exclude from figure.    |
| clist     | Array of columns intended to restrict figure to or to exclude from figure. |
| texcl     | Boolean indicating if taxa in tlist should be excluded. Default, False     |
| cexcl     | Boolean indicating if columns in clist should be excluded. Default, False  |

### Value

plot\_frame, a data frame to be plotted

### Examples

```
fasta_file <- system.file("extdata", "example.fasta", package = "alignfigR")
plot_frame <- read_alignment(file = fasta_file)
subset_seq_list <- extract_subalign(plot_frame, tlist = c("Cow", "Human", "Whale"), texcl = TRUE)
subset_seq_list <- extract_subalign(plot_frame, clist = 1:25)
```

---

|                |                                           |
|----------------|-------------------------------------------|
| plot_alignment | <i>Plot a multiple sequence alignment</i> |
|----------------|-------------------------------------------|

---

### Description

This function uses ggplot (in particular, w/ geom\_rect) to plot a sequence alignment

### Usage

```
plot_alignment(seq_list, palette = NA, taxa = c(), columns = c(),
  exclude_taxa = FALSE, exclude_columns = FALSE,
  legend_title = "Character")
```

**Arguments**

|                 |                                                                                                        |
|-----------------|--------------------------------------------------------------------------------------------------------|
| seq_list        | Sequence list parsed using the function 'read_alignment'                                               |
| palette         | Named-array mapping sequence to color or a pre-defined color scheme (random, rainbow, etc.)            |
| taxa            | Array of taxa (the actual labels, not order) intended to restrict figure to or to exclude from figure. |
| columns         | Array of columns (indexed from 1) intended to restrict figure to or to exclude from figure.            |
| exclude_taxa    | Boolean argument indicating that taxa should be excluded from plot. Default: False                     |
| exclude_columns | Boolean argument indicating that columns should be excluded from plot. Default: False                  |
| legend_title    | String determining title of legend. Default: "Character"                                               |

**Value**

ggplot object which may be saved or edited as desired

**Examples**

```
fasta_file <- system.file("extdata", "example.fasta", package = "alignfigR")
plot_frame <- read_alignment(file = fasta_file)
align_plot <- plot_alignment(plot_frame, "DNA")
align_plot <- plot_alignment(plot_frame, "protein")
align_plot <- plot_alignment(plot_frame, taxa = c("Cow", "Whale"), columns = c(1:25))
align_plot <- plot_alignment(plot_frame, taxa = c("Whale"), exclude_taxa = TRUE)
align_plot <- plot_alignment(plot_frame, legend_title = "") ## Remove the title
```

---

|                |                                                 |
|----------------|-------------------------------------------------|
| read_alignment | <i>Read a multiple sequence alignment file.</i> |
|----------------|-------------------------------------------------|

---

**Description**

This function parses a FASTA file containing molecular sequence data and returns a named-array of the sequences.

**Usage**

```
read_alignment(file)
```

**Arguments**

|      |                                                 |
|------|-------------------------------------------------|
| file | File name. NOTE: Only FASTA file are supported! |
|------|-------------------------------------------------|

**Value**

seq\_array, a named-array of the parsed sequence data

**Examples**

```
fasta_file <- system.file("extdata", "example.fasta", package = "alignfigR")  
plot_frame <- read_alignment(file = fasta_file)
```

# Index

`alignfigR`, [2](#)  
`alignfigR-package (alignfigR)`, [2](#)  
`define_palette`, [2](#)  
`extract_subalign`, [3](#)  
`plot_alignment`, [3](#)  
`read_alignment`, [4](#)