

# Package ‘ggseqalign’

November 14, 2024

**Type** Package

**Title** Minimal Visualization of Sequence Alignments

**Version** 1.0.0

**Description** Simple visualizations of alignments of DNA or AA sequences as well as arbitrary strings. Compatible with Biostrings and ggplot2. The plots are fully customizable using ggplot2 modifiers such as theme().

**License** Artistic-2.0

**Encoding** UTF-8

**RoxygenNote** 7.3.1

**Roxygen** list(markdown = TRUE)

**biocViews** Alignment, MultipleSequenceAlignment, Software, Visualization

**Depends** R (>= 4.4.0)

**Imports** pwalign, dplyr, ggplot2

**Suggests** Biostrings, BiocStyle, knitr, rmarkdown

**URL** <https://github.com/simeross/ggseqalign>

**BugReports** <https://github.com/simeross/ggseqalign/issues>

**VignetteBuilder** knitr

**git\_url** <https://git.bioconductor.org/packages/ggseqalign>

**git\_branch** RELEASE\_3\_20

**git\_last\_commit** 6ecf37d

**git\_last\_commit\_date** 2024-10-29

**Repository** Bioconductor 3.20

**Date/Publication** 2024-11-13

**Author** Simeon Lim Rossmann [aut, cre]  
(<https://orcid.org/0000-0003-0435-8221>)

**Maintainer** Simeon Lim Rossmann <simeon.rossmann@nmbu.no>

## Contents

ggseqalign-package . . . . .	2
alignment_table . . . . .	2
plot_sequence_alignment . . . . .	3

**Index****5**


---

ggseqalign-package      *ggseqalign: Minimal Visualization of Sequence Alignments*

---

**Description**

Simple visualizations of alignments of DNA or AA sequences as well as arbitrary strings. Compatible with Biostrings and ggplot2. The plots are fully customizable using ggplot2 modifiers such as theme().

**Author(s)**

**Maintainer:** Simeon Lim Rossmann <simeon.rossmann@nmbu.no> ([ORCID](#))

**See Also**

Useful links:

- <https://github.com/simeross/ggseqalign>
- Report bugs at <https://github.com/simeross/ggseqalign/issues>

---

alignment\_table      *Alignment Table*

---

**Description**

Generate a table of mismatches and indels between one or many query sequences and a subject sequence.

**Usage**

```
alignment_table(query = XStringSet, subject = XStringSet, ...)
```

**Arguments**

query	A string or vector of strings or object of class XStringSet containing the query sequences/strings.
subject	A string or object of class XStringSet containing the subject sequence/string. Must be of length 1.
...	Any additional parameters are passed on to <code>pwalgn::pairwiseAlignment()</code> . This allows for adjusting alignment algorithm and parameters.

**Value**

A list containing tibbles with information on mismatches and indels.

## Examples

```
query_seq <- Biostrings::DNASTringSet(c("ACCGTACCTGG", "ACCTTGG"))
subject_seq <- Biostrings::DNASTringSet("ACCGTACCGGG")
alignment_table(query_seq, subject_seq)

# Works with any string
query_string <- c("boo", "fizzbuzz")
subject_string <- "boofizz"
alignment_table(query_string, subject_string)
```

---

plot\_sequence\_alignment

*Plot sequence alignment*

---

## Description

This function generates a sequence alignment plot using ggplot2 based on the input alignment table.

## Usage

```
plot_sequence_alignment(
  alignment_tbl = alignment_table(query, subject),
  insertion_color = "#21918c",
  hide_mismatches = FALSE
)
```

## Arguments

**alignment\_tbl** An alignment table containing query and subject information for sequence alignment. Generated with `alignment_table()`.

**insertion\_color** The color to use for indicating insertions in the alignment. Default is '#21918c'. Can be any output of `colors()` or hex code.

**hide\_mismatches** A logical value indicating whether to hide mismatches in the alignment plot. Default is FALSE.

## Value

A ggplot object of the sequence alignment plot.

## Examples

```
q <- (c("boo", "fibububuzz", "bozz", "baofuzz"))
s <- "boofizz"

alignment <- alignment_table(q, s)
p11 <- plot_sequence_alignment(alignment_tbl = alignment)
p11

# Provide names for (some) query and subject elements to label the y-axis
```

```

names(q) <- c("Seq1", NA, "Seq3")
names(s) <- "reference"
p12 <- plot_sequence_alignment(alignment_table(q, s))
p12

# Compatible with StringSets from Biostrings
library(Biostrings)

dna <- readDNASTringSet(system.file("extdata", "dm3_upstream2000.fa.gz",
  package = "Biostrings"
))
# The entries dna[2:5] are identical
q <- dna[2:4]
s <- dna[5]
p13 <- plot_sequence_alignment(alignment_table(q, s))
p13

# Let's introduce some SNPs, insertions and deletions
q[1] <- as(
  replaceLetterAt(q[[1]], c(5, 200, 400), "AGC"),
  "DNASTringSet"
)
q[2] <- as(
  c(substr(q[[2]], 300, 1500), substr(q[[2]], 1800, 2000)),
  "DNASTringSet"
)
q[3] <- as(
  replaceAt(
    q[[3]], 1500,
    paste(rep("A", 1000), collapse = "")
  ),
  "DNASTringSet"
)
names(q) <- c("mismatches", "deletions", "insertion")
names(s) <- substr(names(s)[1], 1, 34)

p14 <- plot_sequence_alignment(alignment_table(q, s))
p14

# Compatible with ggplot2 theming
library(ggplot2)

p14 +
  ylab("Sequence variants") +
  xlab("Length in bp") +
  scale_color_viridis_d() +
  theme(
    axis.text.x = element_text(angle = 90, vjust = 0.5, hjust = 1),
    axis.title = element_text()
  )

```

# Index

## \* internal

ggseqalign-package, 2

alignment\_table, 2

alignment\_table(), 3

ggseqalign (ggseqalign-package), 2

ggseqalign-package, 2

plot\_sequence\_alignment, 3

pwalgn::pairwiseAlignment(), 2