Package: ggseqalign (via r-universe)

July 13, 2024

Type Package

Title Minimal Visualization of Sequence Alignments

Version 0.99.1

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 (>= 4.3.0)

Imports pwalign, dplyr, ggplot2

Suggests Biostrings, BiocStyle, knitr, rmarkdown

URL https://github.com/simeross/ggseqalign

BugReports https://github.com/simeross/ggseqalign/issues

VignetteBuilder knitr

Repository https://bioc.r-universe.dev

RemoteUrl https://github.com/bioc/ggseqalign

RemoteRef HEAD

RemoteSha 22f8c31c02db8dc81374aa8712a3dca784a89f8b

Contents

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/strin. Must be of length 1.
...	Any additional parameters are passed on to `pwalign::pairwiseAlignment()`. 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)
pl1 <- plot_sequence_alignment(alignment_tbl = alignment)
pl1

# Provide names for (some) query and subject elements to label the y-axis
names(q) <- c("Seq1", NA, "Seq3")
names(s) <- "reference"
pl2 <- plot_sequence_alignment(alignment_table(q, s))
pl2

# 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]
pl3 <- plot_sequence_alignment(alignment_table(q, s))
pl3

# 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)

pl4 <- plot_sequence_alignment(alignment_table(q, s))
pl4

# Compatible with ggplot2 theming
library(ggplot2)

pl4 +
    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