Package: fragr (via r-universe)

September 9, 2024

Title an R package for wrangling capillary electrophoresis data

Version 0.0.0.75

Description an R package for wrangling capillary electrophoresis data

License none

Encoding UTF-8

LazyData true

RoxygenNote 7.1.1

Imports magrittr, readr, tibble, tidyr, dplyr, stringr, purrr, rlang,
ggplot2, broom, glue, scales

Suggests feather

Repository https://eric-hunt.r-universe.dev

RemoteUrl https://github.com/eric-hunt/fragr

RemoteRef HEAD

RemoteSha c0f23fac21b5e35b4f4916184e855a32fc7e2c60

Contents

---

munge_prodify   *Put relative measurements in terms of product formed*

Description

munge_prodify

Usage

```
munge_prodify(nested_df, substrate_cutoff)
```

Arguments

nested_df   a nested data frame of CE data with all peak data lying in the list column 'data'

substrate_cutoff

a numeric value indicating the bp size above which all peaks should represent product

Value

returns the data frame relative data expressed in terms of percent product formed (i.e. 100

---

munge_subfix   *Fix instances where substrate is not defined by a single peak*

Description

munge_subfix

Usage

```
munge_subfix(nested_df)
```

Arguments

nested_df   a nested data frame of CE data with all peak data lying in the list column 'data'

Value

returns the data frame with the product peak(s) merged into one peak

---

| parse_activity | *Determine activity according to product formed* |

Description

`parse_activity` will determine specific and non-specific activity, expressed in terms of product formed, uisng a capillary migration model calculated by `fragr::model_migration`

Usage

```
parse_activity(
  nested_df,
  reg_vars,
  reg_limit,
  substrate_cutoff,
  top_n = 5,
  window_size = 2
)
```

Arguments

`nested_df`	a nested data frame of CE data with all peak data lying in the list column 'data'
`reg_vars`	a named vector of numeric values from `lm` capillary migration model, "slope" and "intercept" must be present
`reg_limit`	a numeric value defining how far the product peak may deviate from the model
`substrate_cutoff`	
	a numeric value indicating the bp size above which all peaks should represent product
`top_n`	a numeric value for noise reduction selecting the top number of peaks to analyze by peak area, defaults to five (5) unless otherwise specified
`window_size`	a numeric value indicating the ±window around the specific peak where off-target can be called on the same strand, defaults to two (2) bp above and below "act_bp" unless otherwise specified

Details

The regression model variables are used to create a predictive function, which allows `parse_activity` to select the peak closest to the theoretical or predicted size and label this peak as the specific or expected activity in relation to all activity. The "act" variable represents the activity determined from a peak closest to the theoretical predicted size, which is determined by the regression model variables and alllowed deviation reg_limit arguments. The "act_bp" variable identifies the called size of the peak which was used to generate the "act" variable. The "offsense" variable represents the sum of all other activity besides "act" on the same strand. The "offanti" variable is only included in situations where the substate is double-stranded, and represents the sum of all activity falling on the opposite strand as "act" and "offsense".

`parse_activity`

Value

returns the nested data frame with added variables "act", "act_bp", "offsense", and "offanti"

See Also

[fragr::model_migration]

| parse_migmodel | *Generate regression model for CE migration* |

Description

```
parse_migmodel
```

Usage

```
parse_migmodel(
  df_list,
  channels,
  search_var,
  search_pattern,
  substrate_cutoff,
  limit = 10
)
```

Arguments

`df_list`	a list of data frames containing CE data imported with `fragr::read_PeakScanner`
`channels`	the one letter channel/dye code(s) (as a character vector) for which to build a migration regression model
`search_var`	a metadata-containing variable which contains the theoretical product size
`search_pattern`	a regex pattern for extracting the theoretical product size 'search_var'
`substrate_cutoff`	
	a numeric value indicating the bp size above which all peaks should represent product
`limit`	a numeric value defining how far, in basepairs, the called value an deviate from the theoretical value, defaults to ten (10) unless otherwise specified

Value

returns a model plot and list of variables..

---

read_PeakScanner *Import PeakScanner *combined table* .csv files into R*

Description

read_PeakScanner

Usage

read_PeakScanner(directory_path, pattern = NULL, type = "legacy")

Arguments

directory_path	a path to a directory containing the .csv files
pattern	a regex pattern for selecing files in the directory, defaults to reading all .csv files present
type	a character string, either 'legacy' or 'cloud' indicating where the file was generated (desktop app or online app); defaults to 'legacy'

Value

generates a named list of tibbles, each neamed element is one *combined table* named with its origin file path

---

read_PeakScanner_file *Import single PeakScanner *combined table* .csv file into R*

Description

read_PeakScanner_file

Usage

read_PeakScanner_file(file_path, type = "legacy")

Arguments

file_path	a path to a directory containing the .csv files
type	a character string, either 'legacy' or 'cloud' indicating where the file was generated (desktop app or online app); defaults to 'legacy'

Value

generates a named list of tibbles, each neamed element is one *combined table* named with its origin file path

Index