

Package ‘ngstk’

October 11, 2017

Type Package

Title Next-Generation Sequencing (NGS) Data Analysis Toolkit

Version 0.1.0

Description Can be used to facilitate the analysis of NGS data, such as visualization, conversion of data format for WEB service input and other purpose.

Depends R (>= 3.3.0)

URL <https://github.com/JhuangLab/ngstk>

BugReports <https://github.com/JhuangLab/ngstk/issues>

License MIT + file LICENSE

Encoding UTF-8

LazyData true

Imports stringr (>= 1.2.0), configr (>= 0.3.1)

RoxygenNote 6.0.1

Suggests testthat

NeedsCompilation no

Author Jianfeng Li [aut, cre]

Maintainer Jianfeng Li <lee_jianfeng@sjtu.edu.cn>

Repository CRAN

Date/Publication 2017-10-11 17:43:44 UTC

R topics documented:

format_pp_meta_age	2
format_pp_meta_fusions	3
format_pp_meta_gender	4
fusions2oncoprinter	5
fusions2pp	6
fusions2pp_meta	7
fusions_filter	8

get_pp_samplegroup	10
muts2mutation_mapper	10
muts2oncoprinter	11
muts2pp	13
ngstk	14
set_colors	14
set_tools	15
show_handlers	16
show_mhandlers	16

Index	18
--------------	-----------

format_pp_meta_age	<i>To format ProteinPaint input meta data of age</i>
--------------------	--

Description

To format ProteinPaint input meta data of age

Usage

```
format_pp_meta_age(raw_meta, outfn = NULL, age_group = "Age",
  adult_value = "Adult", child_value = "Pediatric",
  adult_color = "#c20b01", child_color = "#196abd", ...)
```

Arguments

raw_meta	A data.frame contain cols of 'sample', 'term', 'group', 'value', 'color' and 'legendorder'
outfn	Output file, default is NULL and not output to file
age_group	Name of age group, default is 'Age'
adult_value	Value of adult, default is 'Adult'
child_value	Value of child, default is 'Pediatric'
adult_color	Color of adult, default is '#c20b01'
child_color	Color of child, default is '#196abd'
...	Parameters pass to set_colors

Value

A data frame

Examples

```
meta_template <- system.file('extdata',
  'demo/proteinpaint/heatmap_meta_template.txt',
  package = 'ngstk')
raw_meta <- read.table(meta_template, sep = '\t', header = TRUE)
term <- group <- 'Age'
raw_meta$term <- term
raw_meta$group <- group
raw_meta$value <- c(rep(c('Adult', 'Pediatric'), 3), 'Male')
meta_age <- format_pp_meta_age(raw_meta)
```

format_pp_meta_fusions

To format ProteinPaint input meta data of gender

Description

To format ProteinPaint input meta data of gender

Usage

```
format_pp_meta_fusions(raw_meta, outfn = NULL, fusions_color = NULL, ...)
```

Arguments

raw_meta	A data.frame contain cols of 'sample', 'term', 'group', 'value', 'color' and 'legendorder'
outfn	Output file, default is NULL and not output to file
fusions_color	In one group, different fusions show different colors, default is NULL and use the setted theme color
...	Parameters pass to set_colors

Value

A data frame

Examples

```
meta_template <- system.file('extdata',
  'demo/proteinpaint/heatmap_meta_template.txt', package = 'ngstk')
raw_meta <- read.table(meta_template, sep = '\t', header = TRUE)
meta_test_1 <- raw_meta
term <- group <- 'Fusions'
meta_test_1$term <- term
meta_test_1$group <- group
meta_test_1$value <- c(rep(c('ZNF384-Fusions', 'MEF2D-Fusions'),
```

```

3), 'TCF3-PBX1')
meta_fusions <- format_pp_meta_fusions(meta_test_1)
meta_test_2 <- raw_meta
term <- group <- c(rep(c('MEF2D-Fusions', 'ZNF384-Fusions'),
3), 'DUX4-Fusions')
meta_test_2$term <- term
meta_test_2$group <- group
meta_test_2$value <- c('MEF2D-PA', 'EP300-ZNF384',
'MEF2D-PB', 'ABC-ZNF384', 'MEF2D-PB', 'ABD-ZNF384',
'DUX4-IGH')
meta_fusions <- format_pp_meta_fusions(meta_test_2)

```

format_pp_meta_gender *To format ProteinPaint input meta data of gender*

Description

To format ProteinPaint input meta data of gender

Usage

```

format_pp_meta_gender(raw_meta, outfn = NULL, gender_group = "Gender",
male_value = "Male", female_value = "Female", male_color = "#c20b01",
female_color = "#196abd", ...)

```

Arguments

raw_meta	A data.frame contain cols of 'sample', 'term', 'group', 'value', 'color' and 'legendorder'
outfn	Output file, default is NULL and not output to file
gender_group	Name of age group, default is 'Gender'
male_value	Value of male, default is 'Male'
female_value	Value of female, default is 'Female'
male_color	Color of male, default is '#c20b01'
female_color	Color of female, default is '#196abd'
...	Parameters pass to set_colors

Value

A data frame

Examples

```
meta_template <- system.file('extdata',
'demo/proteinpaint/heatmap_meta_template.txt', package = 'ngstk')
raw_meta <- read.table(meta_template, sep = '\t', header = TRUE)
term <- group <- 'Gender'
raw_meta$term <- term
raw_meta$group <- group
raw_meta$value <- c(rep(c('Male', 'Female'), 3), 'Male')
meta_gender <- format_pp_meta_gender(raw_meta)
```

fusions2oncprinter *Function to convert mutation data to Oncprinter <http://www.cbioportal.org/oncprinter.jsp> input format.*

Description

Function to convert mutation data to Oncprinter <http://www.cbioportal.org/oncprinter.jsp> input format.

Usage

```
fusions2oncprinter(input_data, input_type = "fusioncatcher",
  config_file = system.file("extdata", "config/cbioportal.toml", package =
    "ngstk"), config_list = NULL, hander_config_file = system.file("extdata",
    "config/hander.toml", package = "ngstk"),
  mhander_config_file = system.file("extdata", "config/mhander.toml", package =
    "ngstk"), hander_funs = NULL, mhander_funs = NULL,
  hander_extra_params = NULL, mhander_extra_params = NULL, outfn = NULL)
```

Arguments

input_data	A mutation data.frame need to be converted to ProteinPaint input.
input_type	Point the input data format (iseq or others)
config_file	ngstk ProteinPaint configuration file path, default is system.file('extdata', 'config/proteinpaint.toml', package = 'ngstk')
config_list	ngstk ProteinPaint configuration, default is NULL and read from config_file
hander_config_file	ngstk hander configuration file path, default is system.file('extdata', 'config/hander.toml', package = 'ngstk')
mhander_config_file	ngstk hander configuration file path, default is system.file('extdata', 'config/mhander.toml', package = 'ngstk')
hander_funs	hander function for single colnum, default is NULL and get value from config_file
mhander_funs	hander function for mulitple colnums, #' default is NULL and get value from config_file

```

hander_extra_params      Extra parameters pass to handler
mhander_extra_params     Extra parameters pass to mhandler
outfn                    Default is NULL and not output the result to file

```

Value

A data frame

Examples

```

demo_file <- system.file('extdata',
'demo/proteinpaint/muts2pp_iseq.txt', package = 'ngstk')
input_data <- read.table(demo_file, sep = '\t', header = TRUE, stringsAsFactors = FALSE)
disease <- 'T-ALL'
input_data <- data.frame(input_data, disease)
input_data$disease <- as.character(input_data$disease)
fusions2oncprinter(input_data, input_type = 'fusioncatcher')

```

fusions2pp	<i>Function to convert fusion data to ProteinPaint https://pecan.stjude.org/pp input format.</i>
------------	--

Description

Function to convert fusion data to ProteinPaint <https://pecan.stjude.org/pp> input format.

Usage

```

fusions2pp(input_data, input_type = "fusioncatcher",
  config_file = system.file("extdata", "config/proteinpaint.toml", package =
  "ngstk"), config_list = NULL, hander_config_file = system.file("extdata",
  "config/hander.toml", package = "ngstk"),
  mhander_config_file = system.file("extdata", "config/mhander.toml", package =
  "ngstk"), hander_funs = NULL, mhander_funs = NULL,
  hander_extra_params = NULL, mhander_extra_params = NULL, outfn = NULL)

```

Arguments

input_data	A gene fusions data.frame need to be converted to ProteinPaint input.
input_type	Point the input data format (fusioncatcher or others)
config_file	ngstk ProteinPaint configuration file path, default is system.file('extdata', 'config/proteinpaint.toml', package = 'ngstk')
config_list	ngstk ProteinPaint configuration, default is NULL and read from config_file

```

hander_config_file      ngstk hander configuration file path, default is system.file('extdata', 'config/hander.toml',
                        package = 'ngstk')
mhander_config_file     ngstk hander configuration file path, default is system.file('extdata', 'config/mhander.toml',
                        package = 'ngstk')
hander_funs             hander function for single colnum, default is NULL and get value from con-
                        fig_file
mhander_funs           hander function for mulitple colnums, #' default is NULL and get value from
                        config_file
hander_extra_params     Extra parameters pass to handler
mhander_extra_params    Extra parameters pass to mhandler
outfn                   Default is NULL and not output the result to file

```

Value

A data frame

Examples

```

demo_file <- system.file('extdata',
'demo/proteinpaint/fusions2pp_fusioncatcher.txt', package = 'ngstk')
input_data <- read.table(demo_file, sep = '\t', header = TRUE, stringsAsFactors = FALSE)
disease <- 'B-ALL'
samplotype <- 'diagnose'
input_data <- data.frame(input_data, disease, samplotype)
input_data$disease <- as.character(input_data$disease)
hander_data <- fusions2pp(input_data, input_type = 'fusioncatcher')

```

fusions2pp_meta	<i>Function to convert fusion data to ProteinPaint heatmap meta rows</i> https://pecan.stjude.org/pp input format.
-----------------	---

Description

Function to convert fusion data to ProteinPaint heatmap meta rows <https://pecan.stjude.org/pp> input format.

Usage

```

fusions2pp_meta(input_data, input_type = "fusioncatcher",
  config_file = system.file("extdata", "config/proteinpaint.toml", package =
  "ngstk"), config_list = NULL, hander_config_file = system.file("extdata",
  "config/hander.toml", package = "ngstk"),
  mhander_config_file = system.file("extdata", "config/mhander.toml", package =
  "ngstk"), hander_funs = NULL, mhander_funs = NULL,
  hander_extra_params = NULL, mhander_extra_params = NULL, outfn = NULL)

```

Arguments

input_data	A gene fusions data.frame need to be converted to ProteinPaint input.
input_type	Point the input data format (fusioncatcher or others)
config_file	ngstk ProteinPaint configuration file path, default is system.file('extdata', 'config/proteinpaint.toml', package = 'ngstk')
config_list	ngstk ProteinPaint configuration, default is NULL and read from config_file
hander_config_file	ngstk hander configuration file path, default is system.file('extdata', 'config/hander.toml', package = 'ngstk')
mhander_config_file	ngstk hander configuration file path, default is system.file('extdata', 'config/mhander.toml', package = 'ngstk')
hander_funs	hander function for single colnum, default is NULL and get value from config_file
mhander_funs	hander function for mulitple colnums, #' default is NULL and get value from config_file
hander_extra_params	Extra parameters pass to handler
mhander_extra_params	Extra parameters pass to mhander
outfn	Default is NULL and not output the result to file

Value

A data frame

Examples

```
demo_file <- system.file('extdata',
'demo/proteinpaint/fusions2pp_fusioncatcher.txt', package = 'ngstk')
input_data <- read.table(demo_file, sep = '\t', header = TRUE, stringsAsFactors = FALSE)
disease <- 'B-ALL'
samplotype <- 'diagnose'
input_data <- data.frame(input_data, disease, samplotype)
input_data$disease <- as.character(input_data$disease)
hander_data <- fusions2pp_meta(input_data, input_type = 'fusioncatcher')
```

fusions_filter	<i>Fusions hander_data filter that can be used to prepare the input data for downstream analysis</i>
----------------	--

Description

Fusions hander_data filter that can be used to prepare the input data for downstream analysis

Usage

```
fusions_filter(input_data, input_type = "common",
  config_file = system.file("extdata", "config/filter.toml", package =
    "ngstk"), config_list = NULL, hander_config_file = system.file("extdata",
    "config/hander.toml", package = "ngstk"),
  mhander_config_file = system.file("extdata", "config/mhander.toml", package =
    "ngstk"), hander_funs = NULL, mhander_funs = NULL,
  hander_extra_params = NULL, mhander_extra_params = NULL, outfn = NULL)
```

Arguments

input_data	A data frame containing the fusions cols (gene5, gene3, fusion_type)
input_type	Fusion filter type
config_file	ngstk filter configuration file path, default is system.file('extdata', 'config/filter.toml', package = 'ngstk')
config_list	ngstk filter configuration, default is NULL and read from config_file
hander_config_file	ngstk hander configuration file path, default is system.file('extdata', 'config/hander.toml', package = 'ngstk')
mhander_config_file	ngstk hander configuration file path, default is system.file('extdata', 'config/mhander.toml', package = 'ngstk')
hander_funs	hander function for single colnum, default is NULL and get value from config_file
mhander_funs	hander function for mulitple colnums, #' default is NULL and get value from config_file
hander_extra_params	Extra parameters pass to handler
mhander_extra_params	Extra parameters pass to mhandler system.file('extdata', 'config/filter.toml', package = 'ngstk')
outfn	Default is NULL and not output the result to file

Value

A data frame

Examples

```
demo_file <- system.file('extdata',
  'demo/proteinpaint/fusions2pp_fusioncatcher.txt', package = 'ngstk')
input_data <- read.table(demo_file, sep = '\t', header = TRUE, stringsAsFactors = FALSE)
result <- fusions_filter(input_data)
```

`get_pp_samplegroup` *Function to get samplegroup file that can be pass to Proteinpaint*

Description

Function to get samplegroup file that can be pass to Proteinpaint

Usage

```
get_pp_samplegroup(samples, group, outfn = NULL)
```

Arguments

<code>samples</code>	A vector indicating all samples
<code>group</code>	A vector indicating the group information
<code>outfn</code>	Default is NULL and not output the result to file

Value

A data frame

Examples

```
samples <- sprintf('A%s', 1:7)
group <- 'B-ALL'
samplegroup <- get_pp_samplegroup(samples, group)
outfn <- tempfile()
samplegroup <- get_pp_samplegroup(samples, group, outfn)
```

`muts2mutation_mapper` *Function to convert mutation data to MutationMapper http://www.cbioportal.org/mutation_mapper.jsp input format.*

Description

Function to convert mutation data to MutationMapper http://www.cbioportal.org/mutation_mapper.jsp input format.

Usage

```
muts2mutation_mapper(input_data, input_type = "iseq",
  config_file = system.file("extdata", "config/cbioportal.toml", package =
    "ngstk"), config_list = NULL, hander_config_file = system.file("extdata",
    "config/hander.toml", package = "ngstk"),
  mhander_config_file = system.file("extdata", "config/mhander.toml", package =
    "ngstk"), hander_funs = NULL, mhander_funs = NULL,
  hander_extra_params = NULL, mhander_extra_params = NULL, outfn = NULL)
```

Arguments

input_data	A mutation data.frame need to be converted to ProteinPaint input.
input_type	Point the input data format (iseq or others)
config_file	ngstk ProteinPaint configuration file path, default is system.file('extdata', 'config/proteinpaint.toml', package = 'ngstk')
config_list	ngstk ProteinPaint configuration, default is NULL and read from config_file
hander_config_file	ngstk hander configuration file path, default is system.file('extdata', 'config/hander.toml', package = 'ngstk')
mhander_config_file	ngstk hander configuration file path, default is system.file('extdata', 'config/mhander.toml', package = 'ngstk')
hander_funs	hander function for single colnum, default is NULL and get value from config_file
mhander_funs	hander function for mulitple colnums, #' default is NULL and get value from config_file
hander_extra_params	Extra parameters pass to handler
mhander_extra_params	Extra parameters pass to mhandler
outfn	Default is NULL and not output the result to file

Value

A data frame

Examples

```
demo_file <- system.file('extdata',
'demo/proteinpaint/muts2pp_iseq.txt', package = 'ngstk')
input_data <- read.table(demo_file, sep = '\t', header = TRUE, stringsAsFactors = FALSE)
disease <- 'T-ALL'
input_data <- data.frame(input_data, disease)
input_data$disease <- as.character(input_data$disease)
muts2mutation_mapper(input_data, input_type = 'iseq')
```

muts2oncprinter	<i>Function to convert mutation data to Oncoprinter</i> http://www.cbioportal.org/oncoprinter.jsp input format.
-----------------	---

Description

Function to convert mutation data to Oncoprinter <http://www.cbioportal.org/oncoprinter.jsp> input format.

Usage

```
muts2oncoprinter(input_data, input_type = "iseq",
  config_file = system.file("extdata", "config/cbioportal.toml", package =
    "ngstk"), config_list = NULL, hander_config_file = system.file("extdata",
    "config/hander.toml", package = "ngstk"),
  mhander_config_file = system.file("extdata", "config/mhander.toml", package =
    "ngstk"), hander_funs = NULL, mhander_funs = NULL,
  hander_extra_params = NULL, mhander_extra_params = NULL, outfn = NULL)
```

Arguments

<code>input_data</code>	A mutation data.frame need to be converted to ProteinPaint input.
<code>input_type</code>	Point the input data format (iseq or others)
<code>config_file</code>	ngstk ProteinPaint configuration file path, default is system.file('extdata', 'config/proteinpaint.toml', package = 'ngstk')
<code>config_list</code>	ngstk ProteinPaint configuration, default is NULL and read from config_file
<code>hander_config_file</code>	ngstk hander configuration file path, default is system.file('extdata', 'config/hander.toml', package = 'ngstk')
<code>mhander_config_file</code>	ngstk hander configuration file path, default is system.file('extdata', 'config/mhander.toml', package = 'ngstk')
<code>hander_funs</code>	hander function for single colnum, default is NULL and get value from config_file
<code>mhander_funs</code>	hander function for mulitple colnums, #' default is NULL and get value from config_file
<code>hander_extra_params</code>	Extra parameters pass to handler
<code>mhander_extra_params</code>	Extra parameters pass to mhandler
<code>outfn</code>	Default is NULL and not output the result to file

Value

A data frame

Examples

```
demo_file <- system.file('extdata',
  'demo/proteinpaint/muts2pp_iseq.txt', package = 'ngstk')
input_data <- read.table(demo_file, sep = '\t', header = TRUE, stringsAsFactors = FALSE)
disease <- 'T-ALL'
input_data <- data.frame(input_data, disease)
input_data$disease <- as.character(input_data$disease)
muts2oncoprinter(input_data, input_type = 'iseq')
```

muts2pp *Function to convert mutation data to ProteinPaint <https://pecan.stjude.org/pp> input format.*

Description

Function to convert mutation data to ProteinPaint <https://pecan.stjude.org/pp> input format.

Usage

```
muts2pp(input_data, input_type = "iseq",
        config_file = system.file("extdata", "config/proteinpaint.toml", package =
        "ngstk"), config_list = NULL, hander_config_file = system.file("extdata",
        "config/hander.toml", package = "ngstk"),
        mhander_config_file = system.file("extdata", "config/mhander.toml", package =
        "ngstk"), hander_funs = NULL, mhander_funs = NULL,
        hander_extra_params = NULL, mhander_extra_params = NULL, outfn = NULL)
```

Arguments

input_data	A mutation data.frame need to be converted to ProteinPaint input.
input_type	Point the input data format (iseq or others)
config_file	ngstk ProteinPaint configuration file path, default is system.file('extdata', 'config/proteinpaint.toml', package = 'ngstk')
config_list	ngstk ProteinPaint configuration, default is NULL and read from config_file
hander_config_file	ngstk hander configuration file path, default is system.file('extdata', 'config/hander.toml', package = 'ngstk')
mhander_config_file	ngstk hander configuration file path, default is system.file('extdata', 'config/mhander.toml', package = 'ngstk')
hander_funs	hander function for single colnum, default is NULL and get value from config_file
mhander_funs	hander function for mulitple colnums, #' default is NULL and get value from config_file
hander_extra_params	Extra parameters pass to handler
mhander_extra_params	Extra parameters pass to mhandler
outfn	Default is NULL and not output the result to file

Value

A data frame

Examples

```
demo_file <- system.file('extdata',
  'demo/proteinpaint/muts2pp_iseq.txt', package = 'ngstk')
input_data <- read.table(demo_file, sep = '\t', header = TRUE, stringsAsFactors = FALSE)
disease <- 'T-ALL'
input_data <- data.frame(input_data, disease)
input_data$disease <- as.character(input_data$disease)
muts2pp(input_data, input_type = 'iseq')
```

ngstk	<i>ngstk can be used to facilitate the analysis of NGS data, such as visualization, conversion of data format for WEB service input and other purpose.</i>
-------	--

Description

ngstk can be used to facilitate the analysis of NGS data, such as visualization, conversion of data format for WEB service input and other purpose.

Author(s)

Li Jianfeng lee_jianfeng@sjtu.edu.cn

See Also

Useful links:

<https://github.com/JhuangLab/ngstk>

Report bugs at <https://github.com/JhuangLab/ngstk/issues>

set_colors	<i>Function to get a series defined theme colors</i>
------------	--

Description

Function to get a series defined theme colors

Usage

```
set_colors(theme = NULL, theme_config_file = NULL)
```

Arguments

theme	Colors theme, e.g. default, red_blue
theme_config_file	Theme configuration file, default is system.file('extdata', 'config/theme.toml', package = 'ngstk')

show_handlers	<i>Function to show all available handler function</i>
---------------	--

Description

Function to show all available handler function

Usage

```
show_handlers(handler_lib = "default_handlers", show_all_funs = TRUE,
              show_code = NULL, show_description = FALSE,
              handler_config_file = system.file("extdata", "config/hander.toml", package =
              "ngstk"))
```

Arguments

handler_lib	handler lib name
show_all_funs	Default is TRUE and to show all functions in the handler_lib
show_code	Default is NULL, select handler you want to see its source code
show_description	Default is FALSE
handler_config_file	ngstk handler configuration file path, default is system.file('extdata', 'config/hander.toml', package = 'ngstk')

Examples

```
show_handlers(show_description = TRUE)
show_handlers(show_description = FALSE, show_all_funs = FALSE,
              show_code = 'handler_na_replace')
```

show_mhandlers	<i>Function to show all available mhandler function</i>
----------------	---

Description

Function to show all available mhandler function

Usage

```
show_mhandlers(mhandler_lib = "default_mhandlers", show_all_funs = TRUE,
              show_code = NULL, show_description = FALSE,
              mhandler_config_file = system.file("extdata", "config/mhandler.toml", package =
              "ngstk"))
```


Arguments

mhander_lib handler lib name
show_all_funs Default is TRUE and to show all functions in the handler_lib
show_code Default is NULL, select handler you want to see its source code
show_description
 Default is FALSE
mhander_config_file
 ngstk handler configuration file path, default is system.file('extdata', 'config/hander.toml',
 package = 'ngstk')

Examples

```
show_mhandlers(show_description = TRUE)  
show_mhandlers(show_description = FALSE, show_all_funs = FALSE,  
                show_code = 'mhander_fusions_anyfull_match')
```

Index

[format_pp_meta_age](#), [2](#)
[format_pp_meta_fusions](#), [3](#)
[format_pp_meta_gender](#), [4](#)
[fusions2oncoprinter](#), [5](#)
[fusions2pp](#), [6](#)
[fusions2pp_meta](#), [7](#)
[fusions_filter](#), [8](#)

[get_pp_samplegroup](#), [10](#)

[muts2mutation_mapper](#), [10](#)
[muts2oncoprinter](#), [11](#)
[muts2pp](#), [13](#)

[ngstk](#), [14](#)
[ngstk-package \(ngstk\)](#), [14](#)

[set_colors](#), [2-4](#), [14](#)
[set_tools](#), [15](#)
[show_handlers](#), [16](#)
[show_mhandlers](#), [16](#)