Package 'geneNR'

March 29, 2025

Title Automated Gene Identification for Post-GWAS and QTL Analysis

Version 2.0.1

Description Facilitates the post-Genome Wide Association Studies (GWAS) and Quantitative Trait Loci (QTL) analysis of identifying candidate genes within user-defined search window, based on the identified Single Nucleotide Polymorphisms (SNPs) as given by Mazumder AK (2024) <doi:10.1038/s41598-024-66903-3>. It supports candidate gene analysis for wheat and rice. Just import your GWAS result as explained in the sample_data file and the function does all the manual search and retrieve candidate genes for you, while exporting the results into ready-to-use output.

Encoding UTF-8

RoxygenNote 7.3.2

Depends R (>= 3.5)

LazyData TRUE

Imports readr, stringr, utils, httr, rvest, xml2, writexl, vcfR,
ggplot2, ggrepel

Suggests knitr, rmarkdown, devtools

VignetteBuilder knitr

NeedsCompilation no

Author Rajamani Nirmalaruban [aut, cre, cph],
R. Suvitha [aut],
Rajbir Yadav [aut],
Meda Alekya [aut],
Amit Kumar Mazumder [aut],
Subramani Sugumar [aut],
Prashanth babu [aut],
Manjeet Kumar [aut],
Kiran B Gaikwad [aut],
Naresh Kumar Bainsla [aut],
S. Bhaskar Reddy [aut]

Maintainer Rajamani Nirmalaruban `<nirmalaruban97@gmail.com>`

Repository CRAN

Date/Publication 2025-03-29 16:50:02 UTC

Contents

`geneQTL` *Identifies Candidate Genes based on identified Quantitative Trati Loci (QTL) analysis*

Description

Identifies Candidate Genes based on identified Quantitative Trati Loci (QTL) analysis

Usage

```
geneQTL(data_file, crop = "wheat")
```

Arguments

`data_file`	The input data in .csv format. (sample_data_wheat_qtl or sample_data_rice_qtl for demo purpose)
`crop`	Either "wheat" or "rice". (by default it will be wheat)

Value

A data frame containing traits, QTL, gene_id, gene_size, and gene_type.

Examples

```
load(system.file("extdata", "precomputed_sample_results_qtl.rda", package = "geneNR"))
message(sample_results)

result <- geneQTL("sample_data_wheat_qtl", crop="wheat")
result <- geneQTL("sample_data_rice_qtl", crop="rice")
#result <- geneQTL("your_results.csv", crop="wheat")
```

`geneSNP` *Identifies Candidate Genes based on identified Single Nucleotide Ploymorphisms (SNPs) from Genome Wide Association Stuides (GWAS) Analysis*

Description

Identifies Candidate Genes based on identified Single Nucleotide Ploymorphisms (SNPs) from Genome Wide Association Stuides (GWAS) Analysis

Usage

```
geneSNP(data_file, upstream = 1e+06, downstream = 1e+06, crop = "wheat")
```

Arguments

`data_file`	The input data in .csv format. (sample_data_wheat or sample_data_rice for demo purpose)
`upstream`	The search window upstream of the current position of the SNP. (default: 1000000)
`downstream`	The search window downstream of the current position of the SNP. (default: 1000000)
`crop`	Either "wheat" or "rice". (default: wheat)

Value

A data frame containing traits, SNP, gene_id, gene_size, and gene_type.

Examples

```
load(system.file("extdata", "precomputed_sample_results.rda", package = "geneNR"))
message(sample_results)

result <- geneSNP("sample_data_wheat", 10000, 10000, crop = "wheat")
result <- geneSNP("sample_data_rice", 10000, 10000, crop = "rice")
```

`geneSNPcustom` *Identifies Candidate Genes based on identified Single Nucleotide Ploymorphisms (SNPs) from Genome Wide Association Stuides (GWAS) Analysis*

Description

Identifies Candidate Genes based on identified Single Nucleotide Ploymorphisms (SNPs) from Genome Wide Association Stuides (GWAS) Analysis

Usage

```
geneSNPcustom(data_file, crop = "wheat")
```

Arguments

`data_file`	The input data in .csv format. (sample_data_wheat_custom for demo purpose)
`crop`	Either "wheat" or "rice". (default: wheat)

Value

A data frame containing traits, SNP, gene_id, gene_size, and gene_type.

Examples

```
load(system.file("extdata", "precomputed_sample_results_custom.rda", package = "geneNR"))
message(sample_results)

result <- geneSNPcustom("sample_data_wheat_custom", crop = "wheat")
```

`import_hmp` *Imports Hapmap genotypic data file*

Description

Imports Hapmap genotypic data file

Usage

```
import_hmp(file_path, header = TRUE, sep = "\t", stringsAsFactors = FALSE)
```

Arguments

`file_path`	Provide the actual path of Hapmap genotypic data file
`header`	by default it will be True
`sep`	by default it will be tab separated
`stringsAsFactors`	by default it will be False

Value

Hampmap genotypic data

Examples

```
demo_SNP <- system.file("extdata", "demo_SNP.hmp.txt", package = "geneNR")
hapmap_data <- import_hmp(demo_SNP)
head(hapmap_data)
```

`import_vcf` *Imports VCF (Variant Call Format) data file*

Description

Imports VCF (Variant Call Format) data file

Usage

```
import_vcf(file_path)
```

Arguments

`file_path`	Provide the actual path of the VCF file

Value

A `vcfR` object containing the imported data

Examples

```
demo_SNP <- system.file("extdata", "demo_SNP.vcf", package = "geneNR")
vcf_data <- import_vcf(demo_SNP)
vcf_data
```

`plot_SNP` *Plot SNP Distribution on Chromosome Map*

Description

Plots SNP positions across chromosomes with centromere markers using given chromosome details and SNP data.

Usage

```
plot_SNP(
  chromosome_details,
  data,
  chromosome_color = "steelblue",
  title = "Chromosome map with SNPs",
  label_color = "black",
  image_width = 10,
  image_height = 10
)
```

Arguments

`chromosome_details`	A data frame containing chromosome details with columns `Chr`, `start` and `stop`
`data`	A data frame containing SNP data with columns `Chr`, `Pos`, and `SNP`.
`chromosome_color`	Color of the chromosome bars (default: "skyblue").
`title`	Title of the chromosome plot depicting the identified SNPs
`label_color`	Color of the SNP labels (default: "black").
`image_width`	width of the chromosome plot
`image_height`	height of the chromosome plot

Value

A `ggplot` object for the SNP distribution plot.

Examples

```
chromosome_details <- read.csv(system.file("extdata", "chromosome_details.csv", package = "geneNR"))
data <- read.csv(system.file("extdata", "identified_SNP.csv", package = "geneNR"))
chromosome_plot <- plot_SNP(chromosome_details = chromosome_details, data = data,
chromosome_color = "steelblue" ,title = "Chromosome map with SNPs", label_color = "black",
image_width = 15, image_height = 10)
print(chromosome_plot)
```

`plot_summariseSNP` *Plot SNP Distribution Across Chromosomes*

Description

Creates a bar chart representing the distribution of SNPs across chromosomes. Allows customization of bar color, label size, and label color. Saves the plot to a user-specified directory or a temporary directory.

Usage

```
plot_summariseSNP(
  snp_distribution,
  file_name = "snp_bar_chart.jpeg",
  output_dir = tempdir(),
  bar_color = "lightblue",
  label_size = 3,
  label_color = "black"
)
```

Arguments

`snp_distribution`	A data frame with columns `Chr` and `SNP_Count`.
`file_name`	The name of the file to save the plot (default: "snp_bar_chart.jpeg").
`output_dir`	The directory to save the file (default: `tempdir()`).
`bar_color`	The color of the bars (default: "lightblue").
`label_size`	The size of the text labels on the bars (default: 3).
`label_color`	The color of the text labels on the bars (default: "black").

Value

A `ggplot` object for the created bar chart.

Examples

```
demo_SNP <- system.file("extdata", "demo_SNP.hmp.txt", package = "geneNR")
data <- import_hmp(demo_SNP)
snp_distribution <- summariseSNP(data)
plot <- plot_summariseSNP(snp_distribution, bar_color = "skyblue",
label_size = 3, label_color = "red")
print(plot)
```

`sample_data_rice` *Sample Data*

Description

A dataset containing sample data related to genetic markers and associated traits.

Usage

```
sample_data_rice
```

Format

A data frame with columns:

SNP SNP identifier, character.

Chr Chromosome location, character.

Pos Position on the chromosome, numeric.

traits Associated traits, character.

Source

Basha FTM, Sar P, Bhowmick PK, Mahato A, Bisht DS, Iquebal MA, Chakraborty K, Banerjee A, Verma BC, Bhaduri D, Kumar J, Ngangkham U, Saha S, Priyamedha, Mandal NP, Roy S. Genome-wide association study identified QTLs and genes underlying early seedling vigour in aus rice (Oryza sativa L.). Mol Genet Genomics. 2024 Dec 3;299(1):112. doi: 10.1007/s00438-024-02204-8. PMID: 39625651.

Examples

```
data(sample_data_rice)    #lazy loading
```

`sample_data_rice_qtl` *Sample Data*

Description

A dataset containing sample data related to genetic markers and associated traits.

Usage

```
sample_data_rice_qtl
```

Format

A data frame with columns:

traits Associated traits, character.

Chr Chromosome location, character.

start Position on the chromosome where QTL starts, numeric.

stop Position on the chromosome where QTL stops, numeric.

Source

Generated for demonstration purposes

Examples

```
data(sample_data_rice_qtl)    #lazy loading
```

`sample_data_wheat` *Sample Data*

Description

A dataset containing sample data related to genetic markers and associated traits.

Usage

```
sample_data_wheat
```

Format

A data frame with columns:

SNP SNP identifier, character.

Chr Chromosome location, character.

Pos Position on the chromosome, numeric.

traits Associated traits, character.

Source

Generated for demonstration purposes

Examples

```
data(sample_data_wheat)    #lazy loading
```

`sample_data_wheat_custom` *Sample Data*

Description

A dataset containing sample data related to genetic markers and associated traits.

Usage

```
sample_data_wheat_custom
```

Format

A data frame with columns:

traits Associated traits, character.

SNP SNP identifier, character.

Chr Chromosome location, character.

start Position on the chromosome where search window starts, numeric.

stop Position on the chromosome where search window stops, numeric.

Source

Generated for demonstration purposes

Examples

```
data(sample_data_wheat_custom)    #lazy loading
```

`sample_data_wheat_qtl` *Sample Data*

Description

A dataset containing sample data related to genetic markers and associated traits.

Usage

```
sample_data_wheat_qtl
```

Format

A data frame with columns:

traits Associated traits, character.

Chr Chromosome location, character.

start Position on the chromosome where QTL starts, numeric.

stop Position on the chromosome where QTL stops, numeric.

Source

Generated for demonstration purposes

Examples

```
data(sample_data_wheat_qtl)    #lazy loading
```

`summariseSNP` *Distribution of SNPs Across Chromosomes*

Description

Distribution of SNPs Across Chromosomes

Usage

```
summariseSNP(data)
```

Arguments

`data`	A data frame containing a column named `Chr`

Value

A data frame with chromosome names and the count of SNPs for each chromosome

Examples

```
demo_SNP <- system.file("extdata", "demo_SNP.hmp.txt", package = "geneNR")
data <- import_hmp(demo_SNP)
snp_distribution <- summariseSNP(data)
print(snp_distribution)
```

`summariseSNP_vcf` *Distribution of SNPs Across Chromosomes from VCF*

Description

Distribution of SNPs Across Chromosomes from VCF

Usage

```
summariseSNP_vcf(vcf_data)
```

Arguments

`vcf_data`	A `vcfR` object containing VCF data.

Value

A data frame with chromosome names and the count of SNPs for each chromosome.

Examples

```
demo_SNP <- system.file("extdata", "demo_SNP.vcf", package = "geneNR")
vcf_data <- import_vcf(demo_SNP)
snp_distribution <- summariseSNP_vcf(vcf_data)
print(snp_distribution)
```

Index