Package 'xQTLbiolinks'

September 15, 2023

Title An R Package for Integrative Analysis of Quantitative Trait Locus Data of 'xQTL'

Version 1.6.2

Description Enables users-customized data retrieval, processing, analysis, and data visualization of molecular quantitative trait locus and gene expression data from public resources through the application programming interface <https://gtexportal.org/home/api-docs/index.html> of 'GTEx' and <http://www.ebi.ac.uk/eqtl/api> of 'eQTL cagalogue'.

URL https://github.com/dingruofan/xQTLbiolinks

BugReports https://github.com/dingruofan/xQTLbiolinks/issues

License GPL (>= 3)

Encoding UTF-8

LazyData true

RoxygenNote 7.2.3

Depends R (>= 4.2)

VignetteBuilder knitr

Imports BiocGenerics, cowplot (>= 1.1.1), curl (>= 4.3.2), data.table (>= 1.14.2), DBI, SummarizedExperiment, GenomeInfoDb, GenomicFeatures, GenomicRanges, ggplot2 (>= 3.3.6), ggrepel, IRanges, jsonlite (>= 1.7.2), viridis, RMySQL, stringr (>= 1.4.0), utils (>= 4.0.3)

Suggests VariantAnnotation, PupillometryR, coloc, knitr, rtracklayer, usethis, ggridges, CMplot, R.utils

NeedsCompilation no

Author Ruofan Ding [aut, cre],
Xudong Zou [aut, ctb],
Yangmei Qin [aut, ctb],
Gao Wang [aut, ctb],
Lei Li [aut, ctb, cph]

Maintainer Ruofan Ding <dingruofan1@foxmail.com>

Repository CRAN

Date/Publication 2023-09-14 22:00:02 UTC

R topics documented:

`EBIquery_allTerm` *Query supported terms (phenotypes, studies, tissues) in eQTL catalogue*

Description

Query supported terms (phenotypes, studies, tissues) in eQTL catalogue

Usage

```
EBIquery_allTerm(term = "genes", termSize = 2000)
```

Arguments

`term`	"associations", "molecular_phenotypes", "studies", "tissues", "qtl_groups", "genes" or "chromosomes".
`termSize`	Number of fetched term.

Value

A data.table object.

Examples

```
# Fetch associatons:
associations <- data.table::rbindlist(EBIquery_allTerm("associations",termSize=0))

# fetch molecular_phenotypes:
molecular_phenotypes <- EBIquery_allTerm("molecular_phenotypes", termSize=10)

# fetch studies:
studies <- EBIquery_allTerm("studies")

# fetch tissues:
tissues <- EBIquery_allTerm("tissues")

# fetch tissue-study mapping relationships
```

```
tissue_S <- EBIquery_allTerm( paste0("tissues/", "UBER_0002046","/studies" ))

# fetch qtl groups:
qtl_groups <- EBIquery_allTerm("qtl_groups")

# Fetch genes:
geneList <- EBIquery_allTerm("genes", termSize=10)
```

`extractGeneInfo` *Extract gene details from gencodeGeneInfoAllGranges object*

Description

Extract gene details from gencodeGeneInfoAllGranges object

Usage

```
extractGeneInfo(gencodeGeneInfoAllGranges, genomeVersion = "v26")
```

Arguments

`gencodeGeneInfoAllGranges`	from internal data
`genomeVersion`	"v26" (default) or "v19"

Value

A data.table object.

Examples

```
gencodeGeneInfo <- extractGeneInfo(gencodeGeneInfoAllGranges)
```

`retrieveLD` *Retrieve SNP pairwise LD from locuscompare database*

Description

SNP pairwise lD are calculated based on 1000 Genomes Project Phase 3 version 5. For storage-efficiency, the output will only include SNPs with r2 > 0.2 with the input SNP.

Usage

```
retrieveLD(chr, snp, population)
```

Arguments

chr	(string) Chromosome name. e.g. '22'. Notice that the name should not contain 'chr'.
snp	(string) SNP rsID.
population	(string) One of the 5 popuations from 1000 Genomes: 'AFR', 'AMR', 'EAS', 'EUR', and 'SAS'.

Value

A data.frame object.

Examples

```
ld <- retrieveLD('6', 'rs9349379', 'AFR')
```

---

retrieveLD_LDproxy *Retrieve SNP pairwise LD from LDlink database*

Description

Retrieve SNP pairwise LD from LDlink database

Usage

```
retrieveLD_LDproxy(
  targetSnp = "",
  population = "EUR",
  windowSize = 50000,
  method = "download",
  genomeVersion = "grch38",
  max_count = 3,
  token = "9246d2db7917"
)
```

Arguments

targetSnp	target SNP, support dbSNP IP.
population	Supported population is consistent with the LDlink, which can be listed using function LDlinkR::list_pop()
windowSize	Window around the highlighted snp for querying linkage disequilibrium information. Default:500000
method	The same as fetchContent function, can be chosen from "download", "curl", "GetWithHeader".
genomeVersion	"grch38"(default) or "grch37".
max_count	To prevent download failure due to network fluctuations, max number of connection attempts.
token	Ldlink token, default: "9246d2db7917"

Value

A data.table object.

`xQTLanalyze_coloc` *Conduct colocalization analysis with trait genes generated from* `xQTLanalyze_getTraits`

Description

Conduct colocalization analysis with trait genes generated from `xQTLanalyze_getTraits`

Usage

```
xQTLanalyze_coloc(
  gwasDF,
  traitGene,
  geneType = "auto",
  genomeVersion = "grch38",
  tissueSiteDetail = "",
  study = "gtex_v8",
  mafThreshold = 0.01,
  population = "EUR",
  gwasSampleNum = 50000,
  method = "coloc",
  token = "9246d2db7917",
  bb.alg = FALSE
)
```

Arguments

`gwasDF`	A data.frame or data.table objectof gwas.
`traitGene`	A gene symbol or a gencode id (versioned).
`geneType`	(character) options: "auto","geneSymbol" or "gencodeId". Default: "auto".
`genomeVersion`	"grch38" (default) or "grch37". Note: grch37 will be converted to grch38 automatically.
`tissueSiteDetail`	(character) details of tissues in GTEx can be listed using `tissueSiteDetailGTExv8` or `tissueSiteDetailGTExv7`
`study`	(character) name of studies can be listed using "ebi_study_tissues"
`mafThreshold`	Cutoff of maf to remove rare variants.
`population`	Supported population is consistent with the LDlink, which can be listed using function "LDlinkR::list_pop()"
`gwasSampleNum`	Sample number of GWAS dataset. Default:50000.

method	(character) options: "coloc"(default) or "hyprcoloc" (must be updated to the latest version from the github to use hyprcoloc). Package coloc or hyprcoloc is required.
token	LDlink provided user token, default = NULL, register for token at https://ldlink.nci.nih.gov/?tab=apiacces
bb.alg	For hyprcoloc, branch and bound algorithm: TRUE, employ BB algorithm; FALSE, do not. Default: FALSE.

Value

A list of coloc result and details.

Examples

```
url1 <- "http://bioinfo.szbl.ac.cn/xQTL_biolinks/xqtl_data/gwasDFsub_MMP7.txt"
gwasDF <- data.table::fread(url1)
output <- xQTLanalyze_coloc(gwasDF = gwasDF, traitGene= "MMP7", tissueSiteDetail="Prostate")
```

---

xQTLanalyze_coloc_diy *Conduct colocalization analysis with customized QTL data*

Description

Conduct colocalization analysis with customized QTL data

Usage

```
xQTLanalyze_coloc_diy(
  gwasDF,
  qtlDF,
  mafThreshold = 0.01,
  gwasSampleNum = 50000,
  qtlSampleNum = 10000,
  method = "coloc",
  bb.alg = FALSE
)
```

Arguments

gwasDF	data.frame or data.table, required cols: rsid, chrom, position, pValue, maf, beta, se
qtlDF	data.frame or data.table, required cols: rsid, chrom, position, pValue, maf, beta, se
mafThreshold	Cutoff of maf to remove rare variants.
gwasSampleNum	Sample number of GWAS dataset. Default:50000.

qtlSampleNum	Sample number of QTL dataset. Default:10000.
method	(character) options: "coloc"(default) or "hyprcoloc" (need a highe version).
bb.alg	For hyprcoloc, branch and bound algorithm: TRUE, employ BB algorithm; FALSE, do not. Default: FALSE.

Value

A list

Examples

```
url1 <- "http://bioinfo.szbl.ac.cn/xQTL_biolinks/xqtl_data/gwasDFsub_MMP7.txt"
url2 <- "http://bioinfo.szbl.ac.cn/xQTL_biolinks/xqtl_data/eqtl/MMP7_qtlDF.txt"
gwasDF <- data.table::fread(url1)
qtlDF <- data.table::fread(url2)
output <- xQTLanalyze_coloc_diy(gwasDF = gwasDF, qtlDF=qtlDF, method="coloc")
```

---

xQTLanalyze_getSentinelSnp

Detect sentinel SNPs for GWAS using summary statistics data

Description

Return sentinel snps whose pValue < 5e-8(default) and SNP-to-SNP distance > 1e6 bp.

Usage

```
xQTLanalyze_getSentinelSnp(
  gwasDF,
  pValueThreshold = 5e-08,
  centerRange = 1e+06,
  mafThreshold = 0.01,
  genomeVersion = "grch38",
  grch37To38 = FALSE
)
```

Arguments

gwasDF	A data.frame or a data.table object. Five columns are required (arbitrary column names is supported): Col 1. "snps" (character), , using an rsID (e.g. "rs11966562"). Col 2. "chromosome" (character), one of the chromosome from chr1-chr22. Col 3. "postion" (integer), genome position of snp. Col 4. "P-value" (numeric). Col 5. "MAF" (numeric). Allel frequency. Col 6. "beta" (numeric). effect size. Col 7. "se" (numeric). standard error.

pValueThreshold	Cutoff of gwas p-value. Default: 5e-8
centerRange	SNP-to-SNP distance. Default:1e6
mafThreshold	Cutoff of maf to remove rare variants.
genomeVersion	Genome version of input file. "grch37" or "grch38" (default).
grch37To38	TRUE or FALSE, we recommend converting grch37 to grch38, or using a input file of grch38 directly. Package `rtracklayer` is required.

Value

A data.table object.

Examples

```
url<-"http://bioinfo.szbl.ac.cn/xQTL_biolinks/xqtl_data/GLGC.txt"
gwasDF <- data.table::fread(url)
gwasDF <- gwasDF[, .(rsid, chr, position, P, maf, beta, se)]
sentinelSnpDF <- xQTLanalyze_getSentinelSnp(gwasDF)
```

---

xQTLanalyze_getTraits *Identify trait genes using sentinel SNPs generated from* xQTLanalyze_getSentinelSnp

---

Description

Identify trait genes using sentinel SNPs generated from `xQTLanalyze_getSentinelSnp`

Usage

```
xQTLanalyze_getTraits(
  sentinelSnpDF,
  detectRange = 1e+06,
  tissueSiteDetail = "",
  genomeVersion = "grch38",
  grch37To38 = FALSE,
  overlapWithEGene = TRUE,
  egeneDF = NULL
)
```

Arguments

sentinelSnpDF	A data.table. Better be the results from the function "xQTLanalyze_getSentinelSnp", seven columns are required, including "rsid", "chr", "position", "pValue", "maf", "beta" and "se".

detectRange	A integer value. Trait genes that harbor sentinel SNPs located in the 1kb range upstream and downstream of gene. Default: 1e6 bp
tissueSiteDetail	(character) details of tissues in GTEx can be listed using `tissueSiteDetailGTExv8` or `tissueSiteDetailGTExv7`
genomeVersion	"grch38" or "grch37". Default: "grch38"
grch37To38	TRUE or FALSE, we recommend converting grch37 to grch38, or using a input file of grch38 directly. Package `rtracklayer` is required.
overlapWithEGene	TRUE(default) of FALSE. take the intersection with eGenes, egene data.frame will be automatically download from GTEx, or can be provided by the parameter `egeneDF` specified the data.frame of one column of genecode ID. Default:TRUE
egeneDF	A data.table object of one column of gencode ID. requiring overlapWithEGene is TRUE.

Value

A data.table object

Examples

```
# without a customized egene file,
URL1<-"https://gitee.com/stronghoney/exampleData/raw/master/gwas/GLGC_CG0052/sentinelSnpDF.txt"
sentinelSnpDF <- data.table::fread(URL1)
traitsAll <- xQTLanalyze_getTraits(sentinelSnpDF, detectRange=1e4,"Brain - Cerebellum",
                                   genomeVersion="grch37", grch37To38=TRUE)
# with a egene file:
egeneFile <- "http://bioinfo.szbl.ac.cn/xQTL_biolinks/xqtl_data/egeneDF.txt"
egeneDF <- data.table::fread(egeneFile)
traitsAll <- xQTLanalyze_getTraits(sentinelSnpDF, detectRange=1e4,"Brain - Cerebellum",
                                genomeVersion="grch37", grch37To38=TRUE, egeneDF=egeneDF)
```

---

xQTLanno_calLambda *calculate genomic control inflation factor for a QTL/GWAS summary statistics dataset.*

Description

calculate genomic control inflation factor for a QTL/GWAS summary statistics dataset.

Usage

```
xQTLanno_calLambda(summaryDT)
```

Arguments

summaryDT	A data.frame containing one or two columns: p-value (required) and group (optional)

Value

A data.table object

Examples

```
url1 <- "http://bioinfo.szbl.ac.cn/xQTL_biolinks/xqtl_data/eqtl/MMP7_qtlDF.txt"
qtl <- data.table::fread(url1, sep="\t")

# calculate lambda value with all variants
xQTLanno_calLambda(qtl[,.(pValue)])

# calculate lambda value for each group:
qtl$groups <- sample(c(0,1),size = nrow(qtl), replace = TRUE)
xQTLanno_calLambda(qtl[,.(pValue, groups)])
```

---

xQTLanno_chippeak *enrichment analysis for GWAS / QTL signals in functional elements, including enhancer, promoter, CPG, and TFs*

Description

enrichment analysis for GWAS / QTL signals in functional elements, including enhancer, promoter, CPG, and TFs

Usage

```
xQTLanno_chippeak(
  snpInfo = "",
  genomeVersion = "hg38",
  enrichElement = NULL,
  distLimit = 1
)
```

Arguments

snpInfo	A data.table/data.frame with three columns: chromosome, position and p-value.
genomeVersion	"hg38" (default) or "hg19". Note: hg19 will be converted to hg38 automatically.
enrichElement	A data.table of data.frame object including 4 columns (consistent with bed4 format): chrom, start, end, name.
distLimit	Defaults: 0 (variants overlap with elements).

Value

A data.table object

Examples

```
url1 <- "http://bioinfo.szbl.ac.cn/xQTL_biolinks/xqtl_data/gwas/gwasSub.txt.gz"
url2 <- "http://bioinfo.szbl.ac.cn/xQTL_biolinks/xqtl_data/enhancer.txt"
snpInfo <- data.table::fread(url1, sep="\t")
enhancerDT <- data.table::fread(url2, sep="\t")
variants_hit_enhancer <- xQTLanno_chippeak(snpInfo, enrichElement=enhancerDT)
```

---

xQTLanno_genomic *annotate all signals in GWAS / QTL dataset by genome location*

Description

annotate all signals in GWAS / QTL dataset by genome location

Usage

```
xQTLanno_genomic(snpInfo = "", p_cutoff = 5e-08, genomeVersion = "hg38")
```

Arguments

snpInfo	A data.table/data.frame with three columns: chromosome, position, and P-value.
p_cutoff	Cutoff of p-values of significant variants that will be annotated
genomeVersion	"hg38" (default) or "hg19". Note: hg19 will be converted to hg38 automatically.

Value

A data.table object of variants' genomics distribution

---

xQTLdownload_egene *Download eGenes (eQTL Genes) for a specified gene or a tissue*

Description

eGenes are genes that have at least one significant cis-eQTL acting upon them. Results can be filtered by tissue.

Usage

```
xQTLdownload_egene(
  gene = "",
  geneType = "auto",
  tissueSiteDetail = "",
  recordPerChunk = 2000
)
```

Arguments

gene	(character) gene symbol or gencode id (versioned or unversioned are both supported).
geneType	(character) options: "auto","geneSymbol" or "gencodeId". Default: "auto".
tissueSiteDetail	(character) details of tissues in GTEx can be listed using `tissueSiteDetailGTExv8` or `tissueSiteDetailGTExv7`
recordPerChunk	(integer) number of records fetched per request (default: 2000).

Value

A data.table object.

Examples

```
eGeneInfo <- xQTLdownload_egene("TP53")
eGeneInfo <- xQTLdownload_egene(tissueSiteDetail="Prostate", recordPerChunk=2000)
```

---

xQTLdownload_eqtlAllAsso

Download summary statistics data for eQTLs with a specified gene, variant, tissue or study

---

Description

source of all eQTL associations is EBI eQTL category.

Usage

```
xQTLdownload_eqtlAllAsso(
  gene = "",
  geneType = "auto",
  variantName = "",
  variantType = "auto",
  tissueLabel = "",
```

```
  study = "gtex_v8",
  recordPerChunk = 1000,
  withB37VariantId = FALSE,
  data_source = "eQTL_catalogue"
)
```

Arguments

gene	(character) gene symbol or gencode id (versioned or unversioned are both supported).
geneType	(character) options: "auto","geneSymbol" or "gencodeId". Default: "auto".
variantName	(character) name of variant, dbsnp ID and variant id is supported, eg. "rs138420351" and "chr17_7796745_C_T_b38".
variantType	(character) options: "auto", "snpId" or "variantId". Default: "auto".
tissueLabel	(character) all supported tissues can be listed using "ebi_study_tissues"
study	(character) name of studies can be listed using "ebi_study_tissues". If the study is null, use all studies (Default).
recordPerChunk	(integer) number of records fetched per request (default: 1000).
withB37VariantId	a logical value indicating whether to return the genome location(GTEx v7) of variants. Default: FALSE.
data_source	"eQTL_catalogue"(default) or "liLab"

Value

A data.table object.

Examples

```
# Download all eQTL associations of MLH1-rs13315355 pair in all tissues from all studies:
eqtlAsso <- xQTLdownload_eqtlAllAsso(gene="MLH1", variantName = "rs13315355", study="")

# Download eQTL associations of gene ATP11B in CD4+ T cell from all supported studies:
geneAsso <- xQTLdownload_eqtlAllAsso(gene="MMP7",tissueLabel = "CD4+ T cell", study="")

# Download eQTL associations of gene ATP11B in Muscle - Skeletal from GTEx_V8:
geneAsso <- xQTLdownload_eqtlAllAsso("ATP11B", tissueLabel="Muscle - Skeletal")

# Download all eQTL associations of SNP rs11568818 in all tissues from all supported studies.
varAsso <- xQTLdownload_eqtlAllAsso(variantName="rs11568818", study="")
# Download eQTL associations of SNP rs11568818 in Muscle - Skeletal from GTEx_V8:
varAsso <- xQTLdownload_eqtlAllAsso(variantName="chr11_102530930_T_C_b38",
                                    tissueLabel="Muscle - Skeletal")
```

`xQTLdownload_eqtlAllAssoPos` *Download summary statistics data for eQTLs with genome positions.*

Description

Download summary statistics data for eQTLs with genome positions.

Usage

```
xQTLdownload_eqtlAllAssoPos(
  chrom = "",
  pos_lower = numeric(0),
  pos_upper = numeric(0),
  p_lower = 0,
  p_upper = 1.1,
  gene = "",
  geneType = "auto",
  tissueLabel = "",
  study = "gtex_v8",
  recordPerChunk = 1000,
  withB37VariantId = FALSE
)
```

Arguments

`chrom`	(character) name of chromesome, including chr1-chr22, chrX.
`pos_lower`	(integer) lower base pair location threshold, expressed as an integer
`pos_upper`	(integer) upper base pair location threshold, expressed as an integer
`p_lower`	(numeric) lower p-value threshold, can be expressed as a float or using mantissa and exponent annotation (0.001 or 1e-3 or 1E-3)
`p_upper`	(numeric) upper p-value threshold, can be expressed as a float or using mantissa and exponent annotation (0.001 or 1e-3 or 1E-3)
`gene`	(character) gene symbol or gencode id (versioned or unversioned are both supported).
`geneType`	(character) options: "auto","geneSymbol" or "gencodeId". Default: "auto".
`tissueLabel`	(character) all supported tissues can be listed using "ebi_study_tissues".
`study`	(character) name of studies can be listed using "ebi_study_tissues".
`recordPerChunk`	(integer) number of records fetched per request (default: 1000).
`withB37VariantId`	a logical value indicating whether to return the genome location(GTEx v7) of variants. Default: FALSE.

Value

A data.table object

Examples

```
eqtlAssos <- xQTLdownload_eqtlAllAssoPos(chrom = "chr11",
                                         pos_lower=101398614, pos_upper = 101402313,
                                         tissueLabel="Brain - Cerebellar Hemisphere",
                                         p_upper=1e-1)
```

---

xQTLdownload_eqtlExp *Download normalized expression for gene with a variant-gene pair*

Description

Download normalized expression for gene with a variant-gene pair

Usage

```
xQTLdownload_eqtlExp(
  variantName = "",
  gene = "",
  variantType = "auto",
  geneType = "auto",
  tissueSiteDetail = ""
)
```

Arguments

`variantName`	(character) name of variant, dbsnp ID and variant id is supported, eg. "rs138420351" and "chr17_7796745_C_T_b38".
`gene`	(character) gene symbol or gencode id (versioned or unversioned are both supported).
`variantType`	(character) options: "auto", "snpId" or "variantId". Default: "auto".
`geneType`	(character) options: "auto","geneSymbol" or "gencodeId". Default: "auto".
`tissueSiteDetail`	(character) details of tissues in GTEx can be listed using `tissueSiteDetailGTExv8` or `tissueSiteDetailGTExv7`

Value

A data.table object.

Examples

```
# Download exp with variant-gene pair in different tissues:
xQTLdownload_eqtlExp(variantName="rs1641513",gene="TP53", tissueSiteDetail="Liver")

# Download expression using variant ID and gencode ID.
xQTLdownload_eqtlExp(variantName="chr1_14677_G_A_b38",gene="ENSG00000228463.9",
                     tissueSiteDetail="Stomach")
```

`xQTLdownload_exp` *Download normalized gene expression at the sample level for a specified tissue.*

Description

Download normalized gene expression at the sample level for a specified tissue.

Usage

```
xQTLdownload_exp(
  genes = "",
  geneType = "auto",
  tissueSiteDetail = "Liver",
  toSummarizedExperiment = FALSE,
  recordPerChunk = 80,
  pathologyNotesCategories = FALSE
)
```

Arguments

`genes`	(character string or a character vector) gene symbols or gencode ids (versioned or unversioned are both supported).
`geneType`	(character) options: "auto","geneSymbol" or "gencodeId". Default: "auto".
`tissueSiteDetail`	(character) details of tissues in GTEx can be listed using `tissueSiteDetailGTExv8` or `tissueSiteDetailGTExv7`
`toSummarizedExperiment`	a logical value indicating whether to return a data.frame or a summarizedExperiment object. Default: TRUE, return a toSummarizedExperiment object.
`recordPerChunk`	(integer) number of records fetched per request (default: 80).
`pathologyNotesCategories`	a logical value indicating whether to return pathologyNotes. Default: FALSE, the pathologyNotes is ignored.

Value

return a SummarizedExperiment or a data.table object harbing gene expression profiles and samples' information.

Examples

```
# Download gene expression with a genecode ID:
expProfiles <- xQTLdownload_exp("ENSG00000210195.2", tissueSiteDetail="Liver")

# Download gene expression into a SummarizedExperiment object:
expProfiles <- xQTLdownload_exp("ENSG00000210195.2", tissueSiteDetail="Liver",
               toSummarizedExperiment=TRUE)
# extract expression profile from SummarizedExperiment object:
expDT <- SummarizedExperiment::assay(expProfiles)
# extract samples' detail from SummarizedExperiment object:
sampleDT <- SummarizedExperiment::colData(expProfiles)

# Download gene expression profiles of multiple genes:
expProfiles <- xQTLdownload_exp(c("tp53","naDK","SDF4"),
                                tissueSiteDetail="Artery - Coronary",
                                pathologyNotesCategories=TRUE)

# Download with versioned and unversioned gencode Id.
expProfiles <- xQTLdownload_exp(c("ENSG00000141510.16","ENSG00000008130.15","ENSG00000078808"),
                               tissueSiteDetail="Artery - Coronary")
```

`xQTLdownload_geneMedExp`

Download median expressions for multiple genes in a specified tissue

---

Description

Download median expressions for multiple genes in a specified tissue

Usage

```
xQTLdownload_geneMedExp(
  genes = "",
  geneType = "auto",
  tissueSiteDetail = "",
  recordPerChunk = 150
)
```

Arguments

genes	(character string or a character vector) gene symbols or gencode ids (versioned or unversioned are both supported).
geneType	(character) options: "auto","geneSymbol" or "gencodeId". Default: "auto".
tissueSiteDetail	(character) details of tissues in GTEx can be listed using `tissueSiteDetailGTExv8` or `tissueSiteDetailGTExv7`
recordPerChunk	(integer) number of records fetched per request (default: 150).

Value

A data.table object.

Examples

```
geneMedExp <- xQTLdownload_geneMedExp(genes="LYNX1")
geneMedExp <- xQTLdownload_geneMedExp(genes=c("TP53", "IRF5"))
```

---

xQTLdownload_hqtl *Download summary statistics data of H3K4me1 and H3K27ac histone QTL (hQTL) using a specified location*

Description

Download summary statistics data of H3K4me1 and H3K27ac histone QTL (hQTL) using a specified location

Usage

```
xQTLdownload_hqtl(
  phenotype_id = "9:99773935-99776816",
  histone_type = "H3K27AC",
  cell_type = "monocyte",
  hqtlmeta = NULL
)
```

Arguments

phenotype_id	phenotype_id that formatted with genome location, like: 9-99773935-99776816, can be obtained using `xQTLdownload_hqtlmeta`
histone_type	(string) One of the histone types: "H3K27AC" or "H3K4ME1".
cell_type	(string) One of the cell types: "monocyte", "neutrophil" or "T cell".
hqtlmeta	A data.table object obtained via `xQTLdownload_hqtlmeta`.

Value

A data.table object

xQTLdownload_hqtlmeta *Download metadata for H3K4me1 and H3K27ac histone QTL (hQTL)*

Description

Download metadata for H3K4me1 and H3K27ac histone QTL (hQTL)

Usage

```
xQTLdownload_hqtlmeta(histone_type = "H3K27AC", cell_type = "monocyte")
```

Arguments

histone_type	(string) One of the histone types: "H3K27AC" or "H3K4ME1".
cell_type	(string) One of the cell types: "monocyte", "neutrophil" or "T cell".

Value

a data.table object includng all CpG ID

xQTLdownload_mQTL *Download summary statistics data of DNA methylation QTL (mQTL) using CpG ID*

Description

Download summary statistics data of DNA methylation QTL (mQTL) using CpG ID

Usage

```
xQTLdownload_mQTL(
  cpg_id = "cg00000221",
  tissue_name = "WholeBlood",
  mQTL_meta = NULL
)
```

Arguments

cpg_id	phenotype_id like: cg00000236, can be obtained using xQTLdownload_mqtlmeta
tissue_name	(String) One of the tissues: BreastMammaryTissue, ColonTransverse, Kidney-Cortex, Lung, MuscleSkeletal, Ovary, Prostate, Testis and WholeBlood
mQTL_meta	meata data of the mQTL that can be accessed by xQTLdownload_mqtlmeta

Value

A data.table object

xQTLdownload_mqtlmeta *Download metadata of DNA methylation QTL (mQTL)*

Description

Download metadata of DNA methylation QTL (mQTL)

Usage

```
xQTLdownload_mqtlmeta(tissue_name = "BreastMammaryTissue")
```

Arguments

tissue_name	(String) One of the tissues: BreastMammaryTissue, ColonTransverse, KidneyCortex, Lung, MuscleSkeletal, Ovary, Prostate, Testis and WholeBlood

Value

A data.table object

xQTLdownload_sc *Download all sc-eQTL associations for a specified gene*

Description

Download all sc-eQTL associations for a specified gene

Usage

```
xQTLdownload_sc(
  gene = "BIN3",
  geneType = "geneSymbol",
  cell_type = "Astrocytes",
  cell_state = "",
  qtl_type = "Cell-type eQTL",
  study_name = "Bryois2022NN"
)
```

Arguments

gene	(character) gene symbol or gencode id (versioned or unversioned are both supported).
geneType	(character) options: "auto","geneSymbol" or "gencodeId". Default: "geneSymbol".
cell_type	(character)cell types supported in the list of study_info from 'xQTLquery_scInfo'

cell_state	(character)cell states supported in the list of study_info from 'xQTLquery_scInfo'
qtl_type	(character)QTL types supported in the list of study_info from 'xQTLquery_scInfo'
study_name	(character)study name supported in the list of study_info from 'xQTLquery_scInfo'

Value

A data.table object

`xQTLdownload_scSig` *Download significant sc-eQTL associations for a specified gene*

Description

Download significant sc-eQTL associations for a specified gene

Usage

```
xQTLdownload_scSig(
  gene = "BIN3",
  geneType = "geneSymbol",
  cell_type = "Astrocytes",
  cell_state = "",
  qtl_type = "Cell-type eQTL",
  study_name = "Bryois2022NN"
)
```

Arguments

gene	(character) gene symbol or gencode id (versioned or unversioned are both supported).
geneType	(character) options: "auto","geneSymbol" or "gencodeId". Default: "geneSymbol".
cell_type	(character)cell types supported in the list of study_info from 'xQTLquery_scInfo'
cell_state	(character)cell states supported in the list of study_info from 'xQTLquery_scInfo'
qtl_type	(character)QTL types supported in the list of study_info from 'xQTLquery_scInfo'
study_name	(character)study name supported in the list of study_info from 'xQTLquery_scInfo'

Value

A data.table object

`xQTLdownload_sgene` *Download details of sGenes (sQTL Genes) for a specified gene or a tissue.*

Description

sGenes are genes that have at least one significant sQTL acting upon them. Results may be filtered by tissue.

Usage

```
xQTLdownload_sgene(
  gene = "",
  geneType = "auto",
  tissueSiteDetail = "",
  recordPerChunk = 2000
)
```

Arguments

`gene`	(character) gene symbol or gencode id (versioned or unversioned are both supported). Can be null.
`geneType`	(character) options: "auto","geneSymbol" or "gencodeId". Default: "auto".
`tissueSiteDetail`	(character) details of tissues in GTEx can be listed using `tissueSiteDetailGTExv8` or `tissueSiteDetailGTExv7`
`recordPerChunk`	(integer) number of records fetched per request (default: 2000).

Value

A data.table object.

Examples

```
sGeneInfo <- xQTLdownload_sgene(tissueSiteDetail="Liver")
sGeneInfo <- xQTLdownload_sgene(gene="DDX11", tissueSiteDetail="Liver" )
```

xQTLdownload_sqtlAllAsso *Download summary statistics data for sQTLs with a specified gene or a tissue*

Description

Download summary statistics data for sQTLs with a specified gene or a tissue

Usage

```
xQTLdownload_sqtlAllAsso(
  genes = "",
  geneType = "auto",
  tissue = "",
  clu_names = "",
  clu_geneid_DF = NULL
)
```

Arguments

genes	(character) gene symbol or gencode id (versioned or unversioned are both supported).
geneType	(character) options: "auto","geneSymbol" or "gencodeId". Default: "gencodeId".
tissue	(character) details of tissues in GTEx can be listed using `tissueSiteDetailGTExv8`
clu_names	(character) If provided, only the sQTL of clu_names will be downloaded
clu_geneid_DF	(data.frame) If provided, clu-gencode mapping relationship will be loaded from this data.frame.

Value

A data.table object of sQTL dataset.

xQTLdownload_sqtlExp *Download normalized PSI value of intron for a sQTL pair*

Description

Download normalized PSI value of intron for a sQTL pair

Usage

```
xQTLdownload_sqtlExp(
  variantName = "",
  phenotypeId = "",
  variantType = "auto",
  tissueSiteDetail = ""
)
```

Arguments

variantName	(character) name of variant, dbsnp ID and variant id is supported, eg. "rs138420351" and "chr17_7796745_C_T_b38".
phenotypeId	A character string. Format like: "chr1:497299:498399:clu_54863:ENSG00000239906.1"
variantType	(character) options: "auto", "snpId" or "variantId". Default: "auto".
tissueSiteDetail	(character) details of tissues in GTEx can be listed using `tissueSiteDetailGTExv8` or `tissueSiteDetailGTExv7`

Value

A data.table object.

Examples

```
# Download sQTL expression in different tissues:
xQTLdownload_sqtlExp(variantName="rs1450891501",
                     phenotypeId="chr1:497299:498399:clu_54863:ENSG00000239906.1",
                     tissueSiteDetail="Lung")

# Dowload sQTL expression using variant ID.
xQTLdownload_sqtlExp(variantName="chr1_1259424_T_C_b38",
                     phenotypeId=" chr1:1487914:1489204:clu_52051:ENSG00000160072.19",
                     tissueSiteDetail="Adipose - Subcutaneous")
```

---

xQTLdownload_xqtlAllAsso

Download summary statistics of xQTL for a specified gene, default:3'aQTL

Description

Download summary statistics of xQTL for a specified gene, default:3'aQTL

Usage

```
xQTLdownload_xqtlAllAsso(
  genes = "",
  geneType = "geneSymbol",
  tissue = "",
  mRNA_refseq = "",
  mRNA_gene_DF = NULL,
  type = "3'aQTL"
)
```

Arguments

genes	(character) gene symbol or gencode id (versioned or unversioned are both supported).
geneType	(character) options: "auto","geneSymbol" or "gencodeId". Default: "geneSymbol".
tissue	(character) details of tissues in GTEx can be listed using `tissueSiteDetailGTExv8`
mRNA_refseq	(character) If provided, only the 3'aQTL of mRNA will be downloaded
mRNA_gene_DF	(data.frame) If provided, mRNA-gencode mapping relationship will be loaded from this data.frame.
type	3'aQTL(default)

Value

A data.table object of xQTL dataset.

---

xQTLquery_eqtl *Query multi-tissue eQTL metasoft results*

Description

can be quried with a gene/variant-gene pair.

Usage

```
xQTLquery_eqtl(
  variantName = "",
  gene = "",
  variantType = "auto",
  geneType = "auto",
  tissueSiteDetail = "",
  recordPerChunk = 100
)
```

Arguments

variantName	(character) name of variant, dbsnp ID and variant id is supported, eg. "rs138420351" and "chr17_7796745_C_T_b38".
gene	(character) gene symbol or gencode id (versioned or unversioned are both supported). Can not be null.
variantType	(character) options: "auto", "snpId" or "variantId". Default: "auto".
geneType	(character) options: "auto","geneSymbol" or "gencodeId". Default: "auto".
tissueSiteDetail	(character) details of tissues in GTEx can be listed using `tissueSiteDetailGTExv8` or `tissueSiteDetailGTExv7`
recordPerChunk	(integer) number of records fetched per request (default: 100).

Value

A data.table object.

Examples

```
# Query with a gene symbol:
eqtlInfo <- xQTLquery_eqtl(gene="TP53")

# Query with unversioned gencode ID:
eqtl_v8 <- xQTLquery_eqtl(gene="ENSG00000141510")

# In a specific tissue:
xQTLquery_eqtl(gene="ENSG00000141510", geneType="gencodeId", tissueSiteDetail="Thyroid" )

# Query with a variant-gene pair:
xQTLquery_eqtl(variantName="rs1641513",gene="TP53")
xQTLquery_eqtl(variantName="chr1_1667948_A_G_b38", gene="SLC35E2B",
                  tissueSiteDetail="Kidney - Cortex")
```

---

xQTLquery_eqtlSig *Query significant eQTL associations for a specified tissue or multiple tissues.*

Description

Query significant eQTL associations for a specified tissue or multiple tissues.

Usage

```
xQTLquery_eqtlSig(
  variantName = "",
  genes = "",
  variantType = "auto",
  geneType = "auto",
  tissueSiteDetail = ""
)
```

Arguments

variantName	(character) name of variant, dbsnp ID and variant id is supported, eg. "rs138420351" and "chr17_7796745_C_T_b38".
genes	(character string or a character vector) gene symbols or gencode ids (versioned or unversioned are both supported).
variantType	(character) options: "auto", "snpId" or "variantId". Default: "auto".
geneType	(character) options: "auto","geneSymbol" or "gencodeId". Default: "auto".
tissueSiteDetail	(character) details of tissues in GTEx can be listed using tissueSiteDetailGTExv8 or tissueSiteDetailGTExv7

Value

A data.table object.

Examples

```
# Query significant eQTL associations with a variant id across all tissues:
xQTLquery_eqtlSig("rs201327123")
xQTLquery_eqtlSig("chr1_14677_G_A_b38")
# Query significant eQTL associations with a variant id in a specified tissue:
xQTLquery_eqtlSig("chr1_14677_G_A_b38",
                  tissueSiteDetail="Skin - Sun Exposed (Lower leg)")

# Query eQTL associations for multiple variants:
varInfo <-  xQTLquery_varPos(chrom="chr1", pos=c(1102708))
xQTLquery_eqtlSig(variantName=varInfo$snpId)

# Query eQTL associations by genes or tissues:
xQTLquery_eqtlSig(genes="ATAD3B")
xQTLquery_eqtlSig(genes=c("TP53", "SLC35E2B"), tissueSiteDetail= "Brain - Cerebellum")
xQTLquery_eqtlSig(genes="ENSG00000141510.16")

# Query eQTL associations with a variant-gene pair:
xQTLquery_eqtlSig(variantName="rs1641513", genes="TP53")
xQTLquery_eqtlSig(variantName="chr1_1667948_A_G_b38",
                    genes="SLC35E2B", tissueSiteDetail="Kidney - Cortex")
```

`xQTLquery_gene` *Query basic information for genes, including name, symbol, position and description*

Description

Query basic information for genes, including name, symbol, position and description

Usage

```
xQTLquery_gene(genes = "", geneType = "auto", recordPerChunk = 150)
```

Arguments

`genes`	A charater vector or a string of gene symbol, gencode id (versioned or unversioned), or a charater string of gene type. • **gene symbol (Default)**. A character string or a character vector (case ignored). like: "tp53","naDK","SDF4". • **gencode/ensemble id** (versioned or unversioned). A character string or a character vector (case ignored). like: "ENSG00000210195.2","ENSG0000007 • **gene classification**. when "geneType" is "geneCategory", supported "genes" can be listed using function `gencodeGenetype$V26` or `gencodeGenetype$V19`
`geneType`	(character) options: "auto","geneSymbol" or "gencodeId". Default: "auto".
`recordPerChunk`	(integer) number of records fetched per request (default: 150).

Value

A data.table object of queried gene information. including following columns:

- **genes.** Input genes
- **geneSymbol.** Gene symbol.
- **gencodeId.** Gencode/ensemble id (versioned).
- **entrezGeneId.** Entrez gene ID.
- **geneType.** Gene type.
- **chromosome.** Note: "chr" is added in gencode v26,
- **start.**
- **end.**
- **strand.**
- **tss.** Transcriptional start site.
- **gencodeVersion.** Gencode Version.
- **genomeBuild.** Genome version.
- **description.**

Examples

```
# query gene of gencode version v26/hg38
geneInfo <- xQTLquery_gene("TP53")
geneInfo <- xQTLquery_gene(c("tp53","naDK","SDF4") )
geneInfo <- xQTLquery_gene(c("ENSG00000210195.2","ENSG00000078808"))
```

---

xQTLquery_geneAll *Query all genes supported in GTEx*

Description

Query all genes supported in GTEx

Usage

```
xQTLquery_geneAll(gencodeVersion = "v26", recordPerChunk = 2000)
```

Arguments

gencodeVersion	(character) options: "v26"(default, matched with gtex_v8) or "v19"
recordPerChunk	(integer) number of records fetched per request (default: 2000).

Value

A data.table object of all genes' information.

---

xQTLquery_sampleBySampleId *Query details of samples with GTEx IDs*

Description

Query details of samples with GTEx IDs

Usage

```
xQTLquery_sampleBySampleId(
  sampleIds,
  recordPerChunk = 150,
  pathologyNotesCategories = FALSE
)
```

Arguments

sampleIds	A character vector or a string of sample ID.
recordPerChunk	(integer) number of records fetched per request (default: 200).
pathologyNotesCategories	Default: pathologyNotes info is ignored.

Value

a data.table object of samples' information.

Examples

```
sampleIds <- c("GTEX-11NUK-0011-R4a-SM-DO12B", "GTEX-11ONC-0011-R4b-SM-DO93H",
               "GTEX-11DXY-0526-SM-5EGGQ", "GTEX-13OVJ-1026-SM-5IFGI")
sampleInfo <- xQTLquery_sampleBySampleId(sampleIds)
```

---

xQTLquery_sampleByTissue

Query details of samples by tissue name

Description

Query details of samples by tissue name

Usage

```
xQTLquery_sampleByTissue(
  tissueSiteDetail = "Liver",
  dataType = "RNASEQ",
  recordPerChunk = 200,
  pathologyNotesCategories = FALSE
)
```

Arguments

tissueSiteDetail	(character) details of tissues in GTEx can be listed using `tissueSiteDetailGTExv8` or `tissueSiteDetailGTExv7`
dataType	A character string. Options: "RNASEQ" (default), "WGS", "WES", "OMNI".
recordPerChunk	(integer) number of records fetched per request (default: 200).
pathologyNotesCategories	Default: pathologyNotes info is ignored.

Value

returen a data.table object of samples' information

Examples

```
sampleInfo <- xQTLquery_sampleByTissue("Brain - Amygdala" )
sampleInfo <- xQTLquery_sampleByTissue(tissueSiteDetail="Liver", pathologyNotesCategories=TRUE)
```

`xQTLquery_sc` *Query significant sc-eQTLs for a specified gene*

Description

Query significant sc-eQTLs for a specified gene

Usage

```
xQTLquery_sc(
  gene = "BIN3",
  geneType = "geneSymbol",
  cell_type = "Astrocytes",
  cell_state = "",
  qtl_type = "Cell-type eQTL",
  study_name = "Bryois2022NN"
)
```

Arguments

`gene`	(character) gene symbol or gencode id (versioned or unversioned are both supported).
`geneType`	(character) options: "auto","geneSymbol" or "gencodeId". Default: "geneSymbol".
`cell_type`	(character)cell types supported in the list of study_info from 'xQTLquery_scInfo'
`cell_state`	(character)cell states supported in the list of study_info from 'xQTLquery_scInfo'
`qtl_type`	(character)QTL types supported in the list of study_info from 'xQTLquery_scInfo'
`study_name`	(character)study name supported in the list of study_info from 'xQTLquery_scInfo'

Value

A data.table object

xQTLquery_scInfo *Query metadata of sc-eQTLs*

Description

Query metadata of sc-eQTLs

Usage

```
xQTLquery_scInfo()
```

Value

A data.table object

xQTLquery_sqtlSig *Query significant sQTL associations for a tissue or multple tissues*

Description

Only GTEx v8 is supported.

Usage

```
xQTLquery_sqtlSig(
  variantName = "",
  genes = "",
  variantType = "auto",
  geneType = "auto",
  tissueSiteDetail = ""
)
```

Arguments

variantName	(character) name of variant, dbsnp ID and variant id is supported, eg. "rs138420351" and "chr17_7796745_C_T_b38".
genes	(character string or a character vector) gene symbol or gencode id (versioned or unversioned are both supported).
variantType	(character) options: "auto", "snpId" or "variantId". Default: "auto".
geneType	(character) options: "auto","geneSymbol" or "gencodeId". Default: "auto".
tissueSiteDetail	(character) details of tissues in GTEx can be listed using `tissueSiteDetailGTExv8` or `tissueSiteDetailGTExv7`

Value

A data.table object.

Examples

```
# Query sQTL associations with rsid:
xQTLquery_sqtlSig(variantName="rs201327123")
xQTLquery_sqtlSig(variantName="chr1_14677_G_A_b38", tissueSiteDetail="Whole Blood")

# Query sQTL associations with gene symbol and gencode ID:
xQTLquery_sqtlSig(genes="ENSG00000141510.16", tissueSiteDetail="Lung" )
xQTLquery_sqtlSig(genes=c("ATAD3B", "MLH1"))

# Query sQTL associations with the variant-genes pair:
xQTLquery_sqtlSig(variantName="rs201327123", genes=c("WASH7P","RP11-206L10.2"))
xQTLquery_sqtlSig(variantName="chr17_7465085_A_G_b38", genes="TP53", tissueSiteDetail="Lung")
```

---

xQTLquery_tissue *Query details for a specified tissue*

Description

Information includes tissue IDs, number of RNA-Seq samples, number of RNA-Seq samples with genotype, number of expressed genes, number of eGenes. Also includes tissueSiteDetail ID, name, abbreviation, uberon ID, and standard tissue colors. TissueSiteDetails are grouped by TissueSites. By default, this service reports from the latest GTEx release.

Usage

```
xQTLquery_tissue(tissueName = "")
```

Arguments

`tissueName`	Tissue name, tissue ID or tissue site name. Default return all tissues' information. Can be choonse from `tissueSiteDetailGTExv8` or `tissueSiteDetailGTExv7`

Value

A data.table object.

Examples

```
tissueAll <- xQTLquery_tissue() # fetch all tissues
Brain <- xQTLquery_tissue("Brain")
```

xQTLquery_varId *Query variant with variant ID or dbSNP ID*

Description

Query variant with variant ID or dbSNP ID

Usage

```
xQTLquery_varId(variantName = "", variantType = "auto")
```

Arguments

variantName	(character) name of variant, dbsnp ID and variant id is supported, eg. "rs138420351" and "chr17_7796745_C_T_b38".
variantType	(character) options: "auto", "snpId" or "variantId". Default: "auto".

Value

A data.table object.

Examples

```
xQTLquery_varId("rs12596338")
xQTLquery_varId("chr11_66561248_T_C_b38")
```

xQTLquery_varPos *Query variants using genome position.*

Description

Query variants using genome position.

Usage

```
xQTLquery_varPos(chrom = "", pos = numeric(0), recordPerChunk = 200)
```

Arguments

chrom	(character) name of chromesome, including chr1-chr22, chrX, chrY.
pos	An integer array.
recordPerChunk	(integer) number of records fetched per request (default: 200).

Value

A data.table object.

Examples

```
xQTLquery_varPos(chrom="chr1", pos=c(1102708,1105739))
```

`xQTLvisual_anno` *Visualize enrichment of variants derived from* `xQTLanno_genomic`

Description

Visualize enrichment of variants derived from `xQTLanno_genomic`

Usage

```
xQTLvisual_anno(snpHits, pValueBy = 5, annoType = "enrichment")
```

Arguments

`snpHits`	A data.table object from result of xQTLanno_genomic
`pValueBy`	Cut step of pvlaue. Defaults: 5
`annoType`	"enrichment" or "overlapping"

Value

A ggplot object

`xQTLvisual_coloc` *Heatmap plot of the LD-p-value relationship of the eQTL*

Description

Heatmap plot of the LD-p-value relationship of the eQTL

Usage

```
xQTLvisual_coloc(
  gene = "",
  geneType = "auto",
  variantName = "",
  variantType = "auto",
  tissueLabels = "",
  study = "",
  population = "EUR"
)
```

Arguments

gene	(character) gene symbol or gencode id (versioned or unversioned are both supported).
geneType	(character) options: "auto","geneSymbol" or "gencodeId". Default: "auto".
variantName	(character) name of variant, dbsnp ID and variant id is supported, eg. "rs138420351" and "chr17_7796745_C_T_b38".
variantType	(character) options: "auto", "snpId" or "variantId". Default: "auto".
tissueLabels	(a character vector) can be listed with `ebi_study_tissues`. If is null, use all tissue / cell-types. (Default)
study	(character) Studies can be listed using `ebi_study_tissues`. If is null, use all studies (Default).
population	(string) One of the 5 popuations from 1000 Genomes: 'AFR', 'AMR', 'EAS', 'EUR', and 'SAS'.

Value

A list containing a data.table object and a ggplot object

Examples

```
heatmapQTL <- xQTLvisual_coloc( gene="MMP7", variantName="rs11568818", study="TwinsUK")
```

---

xQTLvisual_eqtl *Box plot with jittered points for showing number and significance of eQTL associations*

Description

Box plot with jittered points for showing number and significance of eQTL associations

Usage

```
xQTLvisual_eqtl(gene, geneType = "auto")
```

Arguments

gene	(character) gene symbol or gencode id (versioned or unversioned are both supported).
geneType	(character) options: "auto","geneSymbol" or "gencodeId". Default: "auto".

Value

A ggplot object.

Examples

```
xQTLvisual_eqtl("KIF15")
```

`xQTLvisual_eqtlExp` *Boxplot of normalized expression stratified by genotypes for eQTL.*

Description

Boxplot of normalized expression stratified by genotypes for eQTL.

Usage

```
xQTLvisual_eqtlExp(
  variantName = "",
  gene = "",
  variantType = "auto",
  geneType = "auto",
  tissueSiteDetail = "",
  axis_text_size = 1.3,
  axis_title_size = 1.3,
  title_size = 1.4,
  xlab_text = "Genotypes",
  ylab_text = "Normalized expression",
  ylim_v = NULL,
  title_text = "",
  jitter_color = NULL
)
```

Arguments

`variantName`	(character) name of variant, dbsnp ID and variant id is supported, eg. "rs138420351" and "chr17_7796745_C_T_b38".
`gene`	(character) gene symbol or gencode id (versioned or unversioned are both supported).
`variantType`	(character) options: "auto", "snpId" or "variantId". Default: "auto".
`geneType`	(character) options: "auto","geneSymbol" or "gencodeId". Default: "auto".
`tissueSiteDetail`	(character) details of tissues in GTEx can be listed using `tissueSiteDetailGTExv8` or `tissueSiteDetailGTExv7`
`axis_text_size`	(numberic) text size of the axis labels
`axis_title_size`	(numberic) text size of the axis title
`title_size`	(numberic) text size of the title of the plot

xlab_text	Lable for x-axis
ylab_text	for y-axis
ylim_v	Set scale limits
title_text	Title of the plot
jitter_color	(A character vector) Set the point color.

Value

A list containing eQTL detail, expression profile and a ggplot object.

Examples

```
expEqtl<-xQTLvisual_eqtlExp(variantName="rs3778754", gene ="IRF5",
                            tissueSiteDetail="Whole Blood", xlab_text="Genotypes",
                            ylab_text="Expression", ylim_v=c(-2,2),
                            axis_text_size=1.3, axis_title_size=1.3, title_size=1.4,
                            title_text="Genotype-expression",
                            jitter_color=c("#83bea5", "#e09069","#8f9dc6") )
```

---

xQTLvisual_geneExpTissues

Violin plot of distribution of the gene expression profiles among multiple tissues.

Description

Violin plot of distribution of the gene expression profiles among multiple tissues.

Usage

```
xQTLvisual_geneExpTissues(
  gene = "",
  geneType = "auto",
  tissues = "All",
  log10y = FALSE,
  toTissueSite = FALSE
)
```

Arguments

gene	(character) gene symbol or gencode id (versioned or unversioned are both supported).
geneType	(character) options: "auto","geneSymbol" or "gencodeId". Default: "auto".
tissues	A character string or a vector. "All" (default) means that all tissues is included.
log10y	Display values of expression in log scale. Default: FALSE.
toTissueSite	TRUE or FALSE, display all subtissues or tissue Site. Default: TURE.

Value

A list containing expression profile and a ggplot object.

Examples

```
# Display gene expression in specified tissues.
geneExpTissues <- xQTLvisual_geneExpTissues("TP53", tissues=c("Lung", "Brain","Ovary"))
```

`xQTLvisual_genoBox` *Boxplot of values stratified by genotypes with customized data*

Description

Boxplot of values stratified by genotypes with customized data

Usage

```
xQTLvisual_genoBox(
  genoDT,
  axis_text_size = 1.3,
  axis_title_size = 1.3,
  title_size = 1.4,
  xlab_text = "Genotypes",
  ylab_text = "Normalized expression",
  ylim_v = NULL,
  title_text = "",
  jitter_color = NULL
)
```

Arguments

`genoDT`	(Data.framt) including two columns, "value" and "genotypes"
`axis_text_size`	(numberic) text size of the axis labels
`axis_title_size`	(numberic) text size of the axis title
`title_size`	(numberic) text size of the title of the plot
`xlab_text`	(character) Lable for x-axis
`ylab_text`	(character) Lable for x-axis
`ylim_v`	(numeric vector) Set scale limits
`title_text`	(character) Title of the plot
`jitter_color`	(A character vector) Set the point color.

Value

A ggplot object.

Examples

```
url1 <- "http://bioinfo.szbl.ac.cn/xQTL_biolinks/xqtl_data/eqtl/eqtlExpLabel.txt"
genoDT <- data.table::fread(url1)
box_plot <- xQTLvisual_genoBox(genoDT, title_size=1.6, title_text="Geno-Exp association" )
```

---

xQTLvisual_locusCombine

Generate a combined figure including locusZoom and locuscompare plot

---

Description

This function is rebuilt from `locuscompare.R` (https://github.com/boxiangliu/locuscomparer/blob/master/R/locuscompare.R)

Usage

```
xQTLvisual_locusCombine(
  gwasEqtldata,
  posRange = "",
  population = "EUR",
  highlightSnp = "",
  legend_position = "bottomright",
  snpLD = NULL
)
```

Arguments

`gwasEqtldata`	A data.frame or a data.table that including signals from both GWAS and eQTL. Five columns are required (arbitrary column names is supported): `Col 1`. "snps" (character), using an rsID (e.g. "rs11966562"). `Col 2`. "chromosome" (character), one of the chromosome from chr1-chr22. `Col 3`. "postion" (integer), genome position of snp. `Col 4`. "P-value" (numeric) of GWAS signals. `Col 5`. "P-value" (numeric) of eQTL signals.
`posRange`	Genome range that you want to visualize (e.g. "chr6:3e7-7e7"). Default is the region that covers all snps.
`population`	One of the 5 popuations from 1000 Genomes: 'AFR', 'AMR', 'EAS', 'EUR', and 'SAS'.
`highlightSnp`	Default is the snp that with lowest p-value.

legend_position
: (string, optional) Either 'bottomright','topright', or 'topleft'. Default: 'bottomright'.

snpLD
: A data.frame object of LD matrix. Default is null.

Value

A ggplot object.

`xQTLvisual_locusCompare` *Dotplot of comparing regional signals between GWAS and xQTL*

Description

This function is rebuilt from `locuscompare.R` (https://github.com/boxiangliu/locuscomparer/blob/master/R/locuscompare.R)

Usage

```
xQTLvisual_locusCompare(
  eqtlDF,
  gwasDF,
  highlightSnp = "",
  population = "EUR",
  legend = TRUE,
  legend_position = c("topright", "bottomright", "topleft"),
  snpLD = NULL
)
```

Arguments

eqtlDF
: A data.frame or data.table with two columns: dbSNP id and p-value.

gwasDF
: A data.frame or data.table with two columns: dbSNP id and p-value.

highlightSnp
: Default is the snp that is farthest from the origin of the coordinates.

population
: One of the 5 popuations from 1000 Genomes: 'AFR', 'AMR', 'EAS', 'EUR', and 'SAS'.#' @param token LDlink provided user token, default = NULL, register for token at https://ldlink.nci.nih.gov/?tab=apiaccess

legend
: (boolean, optional) Should the legend be shown? Default: TRUE.

legend_position
: (string, optional) Either 'bottomright','topright', or 'topleft'. Default: 'bottomright'.

snpLD
: A data.frame object of LD matrix. Default is null.

Value

A ggplot object.

Examples

```
library(data.table)
# load data:
eqtlDF <-fread("https://gitee.com/stronghoney/exampleData/raw/master/eqtl/eqtlAsso1.txt")
gwasDF <-fread("https://gitee.com/stronghoney/exampleData/raw/master/gwas/AD/gwasChr6Sub3.txt")
# visualize:
xQTLvisual_locusCompare( eqtlDF, gwasDF, legend_position="topleft")
```

---

xQTLvisual_locusZoom *Locuszoom plot for visualizing regional signals relative to genomic position with summary statistics data*

Description

This function is rebuilt from `locuscompare.R` (https://github.com/boxiangliu/locuscomparer/blob/master/R/locuscompare.R)

Usage

```
xQTLvisual_locusZoom(
  DF,
  highlightSnp = "",
  population = "EUR",
  posRange = "",
  legend = TRUE,
  ylim = NULL,
  legend_position = c("topright", "bottomright", "topleft"),
  point_fill = NULL,
  snpLD = NULL
)
```

Arguments

`DF`	A data.frame or a data.table object. Four columns are required (arbitrary column names is supported): `Col 1`. "snps" (character), , using an rsID (e.g. "rs11966562"); `Col 2`. "chromosome" (character), one of the chromosome from chr1-chr22; `Col 3`. "postion" (integer), genome position of snp. `Col 4`. "P-value" (numeric).
`highlightSnp`	Default is the snp that with lowest p-value.
`population`	One of the 5 popuations from 1000 Genomes: 'AFR', 'AMR', 'EAS', 'EUR', and 'SAS'.
`posRange`	Genome range that you want to visualize (e.g. "chr6:3e7-7e7"). Default is the region that covers all snps.
`legend`	(boolean, optional) Should the legend be shown? Default: TRUE.

ylim	set the minimum and maximum values for the y-axis. By default, the function will automatically determine the y-axis limits based on the data being plotted.
legend_position	(character, optional) Either 'bottomright','topright', or 'topleft'. Default: 'bottomright'.
point_fill	(character, optional) Customized color vectors (5 kinds of colors).
snpLD	A data.frame of LD matirx. Default is null.

Value

A list containing data.table and ggplot object.

Examples

```
library(data.table)
# For GWAS dataset:
gwasDF <- fread("https://gitee.com/stronghoney/exampleData/raw/master/gwasChr6Sub4.txt")
xQTLvisual_locusZoom(gwasDF)
# Zoom in:
xQTLvisual_locusZoom(gwasDF, posRange="chr6:4.7e7-4.8e7", population ="EUR")

# For eQTL of a gene of interest (time-consuming):
eqtlAsso <- xQTLdownload_eqtlAllAsso("RP11-385F7.1", tissueLabel = "Brain - Cortex",
                                     withB37VariantId=FALSE)
xQTLvisual_locusZoom(eqtlAsso[,c("snpId", "chrom", "pos", "pValue")], highlightSnp="rs4711878" )
# Zoom in:
xQTLvisual_locusZoom(eqtlAsso[,c("snpId", "chrom", "pos", "pValue")], highlightSnp="rs4711878",
                     posRange="chr6:47.3e6-47.9e6")
```

---

xQTLvisual_manhattan *Manhattan plot for a GWAS summary statistics dataset*

Description

Manhattan plot for a GWAS summary statistics dataset

Usage

```
xQTLvisual_manhattan(gwasDF, pvalue_cutoff = 1e-04, num_snp_selected = 2000)
```

Arguments

gwasDF	A data.frame of summary statistics data, including four cols arranged in the following order: SNP name, chomosome, position, p-value.

pvalue_cutoff Default: 1e-4. The Manhattan plot is a helpful tool for visualizing genome-wide association study results. However, when there are a large number of SNPs, the plot can become difficult to render and generate a large file size. This is due to the stacking of non-significant SNPs at the bottom of the plot. To address this issue, we can choose to filter out some of the non-significant SNPs or randomly select a subset of them to plot. This will improve the readability of the plot and reduce the file size.

num_snp_selected
Default: 2000. Number of SNPs randomly selected for each chromosome.

Value

A pdf format figure.

Examples

```
gwasDF <- data.table::fread(
      "http://bioinfo.szbl.ac.cn/xQTL_biolinks/xqtl_data/gwas/AD/gwasChr6Sub.txt")
xQTLvisual_manhattan(gwasDF[,.(rsid, chr, position,P)])
```

---

xQTLvisual_PZPlot *Compare P-values reported to P-values calculated from Z statistics derived from the reported beta and standard error.*

Description

Compare P-values reported to P-values calculated from Z statistics derived from the reported beta and standard error.

Usage

```
xQTLvisual_PZPlot(
  summaryDT,
  binCutLogP = 4,
  binNumber = 2000,
  distribution_func = "pnorm"
)
```

Arguments

summaryDT A data.frame with three cols: pval, beta, se.

binCutLogP To speed up the rendering process of the plot for tens of millions of GWAS variants, variants with a p-value below a specified threshold (binCutLogP) are randomly sampled for display.

binNumber The number of points randomly selected for plotting.

distribution_func
"pnorm"(default) or "pchisq"

Value

a list containing a data.frame of estimated pvalues and A ggplot2 object

Examples

```
url1 <- "http://bioinfo.szbl.ac.cn/xQTL_biolinks/xqtl_data/gwasDFsub_MMP7.txt"
sumDT <- data.table::fread(url1, sep="\t")
xQTLvisual_PZPlot(sumDT[,.(pValue, beta, se)], distribution_func="pchisq")
```

---

xQTLvisual_qqPlot *Quantile-quantile plot with p-values from GWAS summary statistics data*

Description

Quantile-quantile plot with p-values from GWAS summary statistics data

Usage

```
xQTLvisual_qqPlot(
  summaryDT,
  legend_p = FALSE,
  binCutLogP = 3,
  binNumber = 1000
)
```

Arguments

summaryDT	A data.frame of one col required: pval.
legend_p	TRUE or FALSE, or legend position, including: top, bottom, left and right.
binCutLogP	SNPs whose logP great than this will be binned, other than not binned.
binNumber	Number of bins.

Value

ggplot2 object

Examples

```
url1 <- "http://bioinfo.szbl.ac.cn/xQTL_biolinks/xqtl_data/gwas/gwasSub.txt.gz"
snpInfo <- data.table::fread(url1, sep="\t")
xQTLvisual_qqPlot(snpInfo[,.(pValue)],binCutLogP=5, binNumber=10000)
```

xQTLvisual_sqtl *Box plot with jittered points for showing number and significance of sQTL associations*

Description

Box plot with jittered points for showing number and significance of sQTL associations

Usage

```
xQTLvisual_sqtl(gene, geneType = "auto")
```

Arguments

gene	(character) gene symbol or gencode id (versioned or unversioned are both supported).
geneType	(character) options: "auto","geneSymbol" or "gencodeId". Default: "auto".

Value

A ggplot object.

Examples

```
xQTLvisual_sqtl("KIF15")
```

xQTLvisual_sqtlExp *Boxplot of normalized PSI stratified by genotypes for sQTL.*

Description

Boxplot of normalized PSI stratified by genotypes for sQTL.

Usage

```
xQTLvisual_sqtlExp(
  variantName = "",
  phenotypeId = "",
  variantType = "auto",
  tissueSiteDetail = "",
  axis_text_size = 1.3,
  axis_title_size = 1.3,
  title_size = 1.4,
  xlab_text = "Genotypes",
```

```
  ylab_text = "Norm.Intron-Excision Ratio",
  ylim_v = NULL,
  title_text = "",
  jitter_color = NULL
)
```

Arguments

variantName	(character) name of variant, dbsnp ID and variant id is supported, eg. "rs138420351" and "chr17_7796745_C_T_b38".
phenotypeId	A character string. Format like: "chr1:497299:498399:clu_54863:ENSG00000239906.1"
variantType	(character) options: "auto", "snpId" or "variantId". Default: "auto".
tissueSiteDetail	(character) details of tissues in GTEx can be listed using `tissueSiteDetailGTExv8` or `tissueSiteDetailGTExv7`
axis_text_size	(numberic) text size of the axis labels
axis_title_size	(numberic) text size of the axis title
title_size	(numberic) text size of the title of the plot
xlab_text	(character) Lable for x-axis
ylab_text	(character) Lable for x-axis
ylim_v	(numeric vector) Set scale limits
title_text	(character) Title of the plot
jitter_color	(A character vector) Set the point color.

Value

A list containing variant detail, expression profile and a ggplot object.

---

xQTL_export *Export expression object to a specified format*

Description

Export expression object to a specified format

Usage

```
xQTL_export(exp_object, out_format = "to_clusterP")
```

Arguments

exp_object	expression object derived from `xQTLdownload_exp`
out_format	"to_clusterP", "to_wgcna" and to "to_deseq"

Value

A data.frame/data.table object

Examples

```
expProfiles <- xQTLdownload_exp(c("tp53","naDK","SDF4"),
                                tissueSiteDetail="Artery - Coronary",
                            pathologyNotesCategories=TRUE, toSummarizedExperiment = TRUE)
```

Index