Package 'TreeQTL'

September 19, 2016

Type Package

Title Hierarchical error control for eQTL studies

Version 2.0

Date 2016-09-19

Author Christine B. Peterson

Maintainer Christine B. Peterson <cbpeterson@gmail.com>

Description In the context of single-tissue eQTL studies, TreeQTL provides methods allowing control of the false discovery rate or family wise error rate for the discovery of eSNPs or eGenes, as well as control of the expected average proportion of false discoveries for eAssociations involving the identified eSNPs or eGenes. In the multi-tissue eQTL setting, TreeQTL implements selection procedures which control error rates relevant to the discovery of eSNPs and eGenes which may be active in multiple tissues. Finally, for multi-trait association studies, TreeQTL can be used to control the error rate for the discovery of variants associated to any phenotypes and the average false discovery rate of phenotypes influenced by such variants. Note: Version 2.0 (released 9/19/16) includes multi-tissue eQTL procedures. Version 2.1 with updates to interfaces and functionality expected later in 2016.

License LGPL

Depends data.table (>= 1.9.4), qvalue (>= 2.0.0)

Suggests MatrixEQTL

R topics documented:

TreeQTL-package 2
get_eAssociations 5
get_eGenes 6
get_eGenes_multi_tissue 7
get_eSNPs 9
get_eSNPs_multi_tissue 10
get_n_tests_per_gene 11
get_n_tests_per_SNP 12
get_n_tests_per_SNP_multi_tissue 12

Index **14**

`TreeQTL-package` *Hierarchical error control for eQTL studies*

Description

In the context of eQTL studies in a single tissue, TreeQTL provides methods allowing control of the false discovery rate or family wise error rate for the discovery of eSNPs or eGenes, as well as control of the expected average proportion of false discoveries for eAssociations involving the identified eSNPs or eGenes. In the context of multi-trait association studies, TreeQTL can be used to control the error rate for the discovery of variants associated to any phenotypes and the average false discovery rate of phenotypes influenced by such variants. In addition, TreeQTL implements hierarchical selection procedures for the identification of eSNPs and eGenes in the context of multi-tissue eQTL analysis.

Details

Package:	TreeQTL
Type:	Package
Version:	2.0
Date:	2016-09-19
License:	LGPL

FAQ

Do I have to use Matrix eQTL to compute the input p-values?

No, the input p-values may be computed using software of your choice. For example, it might be more appropriate to use software such as EMMAX to obtain p-values in the presence of sample structure. Regardless of how the p-values were obtained, the input files for TreeQTL should follow the format used by Matrix eQTL i.e. a tab-delimited file with columns `SNP`, `gene`, `beta` (the estimated regression coefficient), `t-stat` (the corresponding test statistic), `p-value` and `FDR` (note that the fields `beta`, `t-stat` and `FDR` are not strictly necessary for TreeQTL and can be empty if not of interest). The input, should, however, be sorted so that the hypotheses are listed in order of decreasing significance i.e. increasing p-value, and the local vs. distal p-values should be stored in separate intput files.

What is an appropriate p-value output threshold to use with Matrix eQTL?

This depends on the levels for error control you are targeting, but for the default TreeQTL setting of `level1` = `level2` = 0.5, a Matrix eQTL output threshold of 0.01 or 0.05 would be reasonable. If the threshold chosen is too low, you will receive a warning when running `get_eSNPs` or `get_eGenes`. If the threshold chosen is high, it will take longer for TreeQTL to run since it has to deal with larger input files.

How can I use TreeQTL for multi-trait association studies?

Recognizing the similarity of multi-trait association studies to eQTL analysis focused on the iden-

tification of regulatory variants, the TreeQTL methods `get_eSNPs` and `get_eAssociations` can be used to carry out the hierarchical error control procedure described in Many phenotypes without many false discoveries: Error controlling strategies for multi-trait association studies. See the Example section below for sample analysis code.

Author(s)

Christine B. Peterson

Maintainer: Christine B. Peterson <cbpeterson@gmail.com>

References

Christine B. Peterson, Marina Bogomolov, Yoav Benjamini, and Chiara Sabatti (2016). TreeQTL: hierarchical error control for eQTL findings. *Bioinformatics*. **32**(16): 2556–2558.

Christine B. Peterson, Marina Bogomolov, Yoav Benjamini, and Chiara Sabatti (2016). Many phenotypes without many false discoveries: Error controlling strategies for multi-trait association studies. *Genetic Epidemiology*. **40**(1): 45–56.

Examples

```
# eQTL example ------------------------------------------------------------
# Perform hierarchical error control using toy data sets from Matrix eQTL

# Note that TreeQTL expects local and distal hypotheses to be in separate
# input files. These can be produced in a single run of Matrix eQTL or
# constructed manually.

# Setup
library(MatrixEQTL)
library(TreeQTL)
source(paste(find.package("TreeQTL"), "/demo/setup_mEQTL.R", sep = ""))

# Distance used to define nearby region
dist <- 1e6

# Generate output files from Matrix eQTL by calling helper function
# which runs Matrix eQTL example script
mEQTL_out_cis <- "mEQTL_out_cis.txt"
mEQTL_out_trans <- "mEQTL_out_trans.txt"
setup_mEQTL(mEQTL_out_cis, mEQTL_out_trans, dist)

# Read in location files from Matrix eQTL
base_dir <- find.package('MatrixEQTL')
gene_map <- read.table(paste(base_dir, "/data/geneloc.txt", sep = ""),
                       header = TRUE, stringsAsFactors = FALSE)
snp_map <- read.table(paste(base_dir, "/data/snpsloc.txt", sep = ""),
                      header = TRUE, stringsAsFactors = FALSE)

# Run TreeQTL separately for local vs. distal analysis ------------------------

# Identify eSNPs which affect nearby genes ------------------------------------

# Get number of genes nearby to each SNP
n_tests_per_SNP <- get_n_tests_per_SNP(snp_map, gene_map, nearby = TRUE, dist = dist)
```

```
# Get list of eSNPs
eSNPs <- get_eSNPs(n_tests_per_SNP, "mEQTL_out_cis.txt")

# Generate txt output file with full set of eAssociations
get_eAssociations(eSNPs, n_tests_per_SNP, "mEQTL_out_cis.txt", "eAssoc_cis_by_snp.txt",
  by_snp = TRUE)

# Identify eGenes regulated by nearby SNPs -----------------------------------

# Get number of SNPs nearby to each gene
n_tests_per_gene <- get_n_tests_per_gene(snp_map, gene_map, nearby = TRUE, dist = dist)

# Get list of eGenes
eGenes <- get_eGenes(n_tests_per_gene, "mEQTL_out_cis.txt")

# Generate txt output file with full set of eAssociations
get_eAssociations(eGenes, n_tests_per_gene, "mEQTL_out_cis.txt", "eAssoc_cis_by_gene.txt",
  by_snp = FALSE)

# Identify eSNPs which regulate distant genes --------------------------------

# Get number of genes distal to each SNP
n_tests_per_SNP <- get_n_tests_per_SNP(snp_map, gene_map, nearby = FALSE, dist = dist)

# Get list of eSNPs with target level 0.2
eSNPs <- get_eSNPs(n_tests_per_SNP, "mEQTL_out_trans.txt", level1 = 0.2)

# Result: no SNPs are significant at given level

# Clean up
unlink("mEQTL_out_cis.txt")
unlink("mEQTL_out_trans.txt")

# Multi-trait example ---------------------------------------------------------
# Perform hierarchical error control for multi-trait association

# Generate input file resembling that for a multi-trait association study
# Assume 100 SNPs and 100 traits
nSNP <- 100
nTrait <- 100
SNP_names <- paste(rep("SNP"), 1:nSNP, sep = "")
trait_names <- paste(rep("Trait"), 1:nTrait, sep = "")

# Construct data frame with random p-values of SNP-trait association,
# including a few small ones
multi_trait_pvals <- data.frame(SNP = rep(SNP_names, nTrait),
                                trait = sort(rep(trait_names, nSNP)),
                                t_stat = NA,
                                beta = NA,
                                "p-value" = runif(nSNP * nTrait),
                                FDR = NA)
multi_trait_pvals$p.value[sample(nrow(multi_trait_pvals), 10)] <- runif(10) * 1e-5

# Threshold and sort appropriately
multi_trait_pvals <- multi_trait_pvals[multi_trait_pvals$p.value <= 0.1, ]
multi_trait_pvals <- multi_trait_pvals[order(multi_trait_pvals$p.value), ]
```

```
# Write to disk
names(multi_trait_pvals)[5] <- "p-value"
write.table(multi_trait_pvals, "multi_trait_pvals.txt", sep = "\t", quote = FALSE,
            row.names = FALSE)

# Number of tests per SNP is just number of traits
n_tests_per_SNP <- data.frame(family = SNP_names, n_tests = nTrait)

# Get list of eSNPs (i.e. SNPs which affect any of the traits)
eSNPs <- get_eSNPs(n_tests_per_SNP, "multi_trait_pvals.txt")

# Generate txt output file with full set of eAssociations (i.e. all signficant
# SNP-trait associations)
get_eAssociations(eSNPs, n_tests_per_SNP, "multi_trait_pvals.txt", "eAssoc_multi_trait.txt",
  by_snp = TRUE)

# Cleanup
unlink("multi_trait_pvals.txt")
unlink("eAssoc_multi_trait.txt")
```

`get_eAssociations` *Writes an output file with the full set of eAssociations*

Description

Writes an output file with the full set of significant eAssociations. Columns are the same as those in Matrix eQTL output except that FDR is replaced by BBFDR (Benjamini-Bogomolov adjusted p-value)

Usage

```
get_eAssociations(eDiscoveries, n_tests, m_eqtl_out, out_file, by_snp, slice_size,
  silent)
```

Arguments

`eDiscoveries`	data frame with columns for eSNP or eGene name, p-value, and number of signficant associations. This data frame can be created by calling `get_eSNPs` or `get_eGenes`
`n_tests`	data frame with SNP/gene name in first column and number of tests for the given SNP/gene in second column. This data frame can be created by calling `get_n_tests_per_gene` or `get_n_tests_per_SNP`
`m_eqtl_out`	txt file with Matrix eQTL output. Must follow Matrix eQTL output file format i.e. a tab-delimited file with columns `SNP`, `gene`, `beta`, `t-stat`, `p-value` and `FDR` (note that the fields `beta`, `t-stat` and `FDR` can be empty), and list associations in increasing order of p-value
`out_file`	file name where set of eAssociations should be written
`by_snp`	should hypothesis families be grouped by SNP? If FALSE, hypotheses will be grouped by probe/gene. Default is TRUE.
`slice_size`	number of lines to read in at a time from Matrix eQTL output. Default is 100,000.
`silent`	Should function run in silent mode? If FALSE, progress updates will be printed. Default is FALSE.

Details

The BBFDR (Benjamini-Bogomolov adjusted p-value) corresponding to the hypothesis with the *j*th smallest p-value involving the *i*th SNP/gene is computed as

$$\min\{p_{(k)} \cdot \frac{N_i}{k} \cdot \frac{M}{R} : k = j, j+1, \ldots, N_i\},$$

where *k* is the ranking of the p-value among all p-values involving SNP/gene *i*, *N_i* is the number of hypotheses involving SNP/gene *i*, *M* is the total number of SNPs/genes, and *R* is the number of SNPs/genes with any significant associations.

Value

Output file written under the name out_file with the full set of significant eAssociations. Columns are the same as those in Matrix eQTL output except that FDR is replaced by BBFDR (Benjamini-Bogomolov adjusted p-value)

Author(s)

Christine B. Peterson

References

Christine Peterson, Marina Bogomolov, Yoav Benjamini, and Chiara Sabatti (2016). TreeQTL: hierarchical error control for eQTL findings. *Bioinformatics*. **32**(16): 2556–2558.

Christine Peterson, Marina Bogomolov, Yoav Benjamini, and Chiara Sabatti (2016). Many phenotypes without many false discoveries: Error controlling strategies for multi-trait association studies. *Genetic Epidemiology*. **40**(1): 45–56.

`get_eGenes` *Get significant eGenes*

Description

Generates a data frame listing all significant eGenes, the gene p-value, and the number of SNPs associated to that gene. The user may optionally specify precomputed gene p-values.

Usage

```
get_eGenes(n_tests_per_gene, m_eqtl_out, method = "BY",
    level1 = 0.05, level2 = 0.05, slice_size = 1e+05,
    silent = FALSE, gene_pvals = NA)
```

Arguments

`n_tests_per_gene`
: data frame with the gene name in the first column and the number of tests for that gene in the second column. This data frame can be created by calling `get_n_tests_per_gene`

m_eqtl_out	txt file with Matrix eQTL output. Must follow Matrix eQTL output file format i.e. a tab-delimited file with columns SNP, gene, beta, t-stat, p-value and FDR (note that the fields beta, t-stat and FDR can be empty), and list associations in increasing order of p-value
method	Method used for error control in gene selection. Either "BY" for Benjamini-Yekutieli, "BH" for Benjamini-Hochberg, or "Bonf" for Bonferroni. Default is "BY".
level1	target level for FDR or FWER for eGenes. Default is 0.05.
level2	target level for expected average proportion of false eAssociation discoveries across eGenes. Default is 0.05.
slice_size	number of lines to read in at a time from Matrix eQTL output. Default is 100,000.
silent	Should function run in silent mode? If FALSE, progress updates will be printed. Default is FALSE.
gene_pvals	Data frame with column for gene name and gene p-value. Default is NA

Value

Data frame with a column providing the gene name, gene p-value and number of associated SNPs for each selected gene

Author(s)

Christine B. Peterson

References

Christine Peterson, Marina Bogomolov, Yoav Benjamini, and Chiara Sabatti (2016). TreeQTL: hierarchical error control for eQTL findings. *Bioinformatics*. **32**(16): 2556–2558.

---

get_eGenes_multi_tissue

Get significant eGenes in the multi-tissue eQTL setting

Description

Performs hierarchical selection for eGenes across multiple tissues by first selecting eGenes which are subject to genetic regulation in any tissue, then the tissues in which they are regulated, and finally the specific SNPs associated to each selected gene x tissue pair

Usage

```
get_eGenes_multi_tissue(genes_by_tissue, snps_by_tissue,
  snp_map, gene_map, nearby = TRUE,
  dist = 1000000,
  m_eqtl_out_dir, tissue_names,
  level1 = 0.05, level2 = 0.05, level3 = 0.05)
```

Arguments

genes_by_tissue	data frame with one row for each gene. The first column should contain the gene name, then there should be one column for each tissue containing a binary indicator of whether the given gene was included in the eQTL analysis for the given tissue i.e. whether it passed QC in that tissue
snps_by_tissue	data frame with one row for each snp. The first column should contain the snp name, then there should be one column for each tissue containing a binary indicator of whether the given SNP was included in the eQTL analysis for the given tissue i.e. whether it passed QC in that tissue
snp_map	data frame with 3 columns: snp name, chromosome, and bp position
gene_map	data frame with 4 columns: gene name, chromosome, start position and end position
nearby	TRUE if hypotheses of interest relate to SNP-gene pairs which are nearby. FALSE if focus is distal regulation. Default value is TRUE
dist	number of base pairs defining "nearby" region. Default is 1,000,000 = 1Mb
m_eqtl_out_dir	path to output files with eQTL output per tissue, thresholded to some reasonable level such as 0.01 or 0.05. These are assumed to follow the output format generated by Matrix eQTL i.e. with columns named SNP, gene, beta, t-stat, p-value, and FDR. The beta, t-stat, and FDR columns may contain dummy values if these are not of interest
tissue_names	vector of names for each tissue in alphanumeric order
level1	target error rate for Level 1 discoveries. Default is 0.05
level2	target error rate for Level 2 discoveries. Default is 0.05
level3	target error rate for Level 3 discoveries. Default is 0.05

Value

Returns a data frame with one row for each eGene selected in Level 1 of the hierarchy and an indicator of which tissues it was selected in. Full information on the selected SNP x gene associations in each tissue is written to disk in a separate file for each tissue.

Author(s)

Christine B. Peterson

References

Christine B. Peterson, Marina Bogomolov, Yoav Benjamini, and Chiara Sabatti (2016). TreeQTL: hierarchical error control for eQTL findings. *Bioinformatics*. **32**(16): 2556–2558.

Additional publication addressing multi-tissue error control to be submitted soon.

get_eSNPs *Get signficant eSNPs*

Description

Generates a data frame listing all significant eSNPs, the SNP p-value, and the number of genes associated to that SNP. The user may optionally specify precomputed SNP p-values.

Usage

```
get_eSNPs(n_tests_per_SNP, m_eqtl_out, method = "BY",
    level1 = 0.05, level2 = 0.05, slice_size = 1e+05,
    silent = FALSE, snp_pvals = NA)
```

Arguments

n_tests_per_SNP	data frame with the SNP name in the first column and the number of tests for that SNP in the second. This data frame can be created by calling get_n_tests_per_SNP
m_eqtl_out	txt file with Matrix eQTL output. Must follow Matrix eQTL output file format i.e. a tab-delimited file with columns SNP, gene, beta, t-stat, p-value and FDR (note that the fields beta, t-stat and FDR can be empty), and list associations in increasing order of p-value
method	Method used for error control in SNP selection. Either "BY" for Benjamini-Yekutieli, "BH" for Benjamini-Hochberg, or "Bonf" for Bonferroni. Default is "BY".
level1	target level for FDR or FWER for eSNPs. Default is 0.05.
level2	target level for expected average proportion of false eAssociation discoveries across eSNPs. Default is 0.05.
slice_size	number of lines to read in at a time from Matrix eQTL output. Default is 100,000.
silent	Should function run in silent mode? If FALSE, progress updates will be printed. Default is FALSE.
snp_pvals	Data frame with column for gene name and gene p-value. Default is NA

Value

Data frame with a column providing the SNP name, SNP p-value and number of associated genes for each selected SNP

Author(s)

Christine B. Peterson

References

Christine B. Peterson, Marina Bogomolov, Yoav Benjamini, and Chiara Sabatti (2016). TreeQTL: hierarchical error control for eQTL findings. *Bioinformatics*. **32**(16): 2556–2558.

Christine B. Peterson, Marina Bogomolov, Yoav Benjamini, and Chiara Sabatti (2016). Many phenotypes without many false discoveries: Error controlling strategies for multi-trait association studies. *Genetic Epidemiology*. **40**(1): 45–56.

`get_eSNPs_multi_tissue` *Get significant eSNPs in the multi-tissue eQTL setting*

Description

Performs hierarchical selection for eSNPs across multiple tissues by first selecting eSNPs which are regulatory for any gene in any tissue, then the genes regulated by each eSNP in any tissue, and finally the specific tissues in which these SNP x gene pairs are associated

Usage

```
get_eSNPs_multi_tissue(genes_by_tissue, snps_by_tissue,
  n_tests_per_SNP, m_eqtl_out_dir, tissue_names,
  level1 = 0.05, level2 = 0.05, level3 = 0.05)
```

Arguments

`genes_by_tissue`	data frame with one row for each gene The first column should contain the gene name, then there should be one column for each tissue containing a binary indicator of whether the given gene was included in the eQTL analysis for the given tissue i.e. whether it passed QC in that tissue
`snps_by_tissue`	data frame with one row for each snp. The first column should contain the snp name, then there should be one column for each tissue containing a binary indicator of whether the given SNP was included in the eQTL analysis for the given tissue i.e. whether it passed QC in that tissue
`n_tests_per_SNP`	data frame with SNP name in first column and number of genes tested for association to the given SNP (which depends on whether cis or trans analysis is being performed) in the second column
`m_eqtl_out_dir`	path to output files (cis or trans) with eQTL output per tissue, thresholded to some reasonable level such as 0.01 or 0.05. These are assumed to follow the output format generated by Matrix eQTL i.e. with columns named `SNP`, `gene`, `beta`, `t-stat`, `p-value`, and `FDR`. The `beta`, `t-stat`, and `FDR` columns may contain dummy values if these are not of interest
`tissue_names`	vector of names for each tissue in alphanumeric order
`level1`	target error rate for Level 1 discoveries. Default is 0.05
`level2`	target error rate for Level 2 discoveries. Default is 0.05
`level3`	target error rate for Level 3 discoveries. Default is 0.05

Value

Data frame with one row for each selected SNP-gene pair. First column contains SNP name. Second column contains gene name. These are followed by one column for each tissue with a nonzero entry giving the p-value for significant associations or a 0 to indicate that the association in the given tissue was not significant

Author(s)

Christine B. Peterson

References

Christine B. Peterson, Marina Bogomolov, Yoav Benjamini, and Chiara Sabatti (2016). TreeQTL: hierarchical error control for eQTL findings. *Bioinformatics*. **32**(16): 2556–2558.

Additional publication addressing multi-tissue error control to be submitted soon.

`get_n_tests_per_gene` *Get number of hypotheses tested for each gene*

Description

Computes the number of hypotheses tested for each gene

Usage

```
get_n_tests_per_gene(snp_map, gene_map, nearby = TRUE, dist = 1e+06)
```

Arguments

`snp_map`	data frame with 3 columns: name, chromosome, and position. This is the same format as the snpspos parameter for Matrix eQTL.
`gene_map`	data frame with 4 columns: name, chromosome, and start and end position. This is the same format as the genepos parameter for Matrix eQTL.
`nearby`	TRUE if hypotheses of interest relate to SNP-gene pairs which are nearby. FALSE if focus is distal regulation. Default value is TRUE.
`dist`	number of base pairs defining "nearby" region. Default is 1,000,000 = 1Mb.

Value

Data frame with the gene name in the first column and the number of tests for that gene in the second column

Author(s)

Christine B. Peterson

References

Christine Peterson, Marina Bogomolov, Yoav Benjamini, and Chiara Sabatti (2016). TreeQTL: hierarchical error control for eQTL findings. *Bioinformatics*. **32**(16): 2556–2558.

`get_n_tests_per_SNP` *Get number of hypotheses tested for each SNP*

Description

Computes the number of hypotheses tested for each SNP

Usage

```
get_n_tests_per_SNP(snp_map, gene_map, nearby = TRUE, dist = 1e+06)
```

Arguments

`snp_map`	data frame with 3 columns: name, chromosome, and position. This is the same format as the snpspos parameter for Matrix eQTL.
`gene_map`	data frame with 4 columns: name, chromosome, and start and end position. This is the same format as the genepos parameter for Matrix eQTL.
`nearby`	TRUE if hypotheses of interest relate to SNP-gene pairs which are nearby. FALSE if focus is distal regulation. Default value is TRUE.
`dist`	number of base pairs defining "nearby" region. Default is 1,000,000 = 1Mb.

Value

Data frame with the SNP name in the first column and the number of tests for that SNP in the second

Author(s)

Christine B. Peterson

References

Christine Peterson, Marina Bogomolov, Yoav Benjamini, and Chiara Sabatti (2016). TreeQTL: hierarchical error control for eQTL findings. *Bioinformatics*. **32**(16): 2556–2558.

`get_n_tests_per_SNP_multi_tissue`
Get number of genes tested for association in any tissue for each SNP

Description

Returns a data frame with the SNP name in the first column and the number genes which were tested for association with the SNP in any tissue. Note that this function only works for up to 62 tissues. Note as well that if the definition of trans being used is that the SNP and gene are on different chromosomes, this can be accomodated by setting "nearby" = FALSE and "dist" = Inf.

Usage

```
get_n_tests_per_SNP_multi_tissue(snp_map_xT, gene_map_xT, nearby = TRUE,
                                 dist = 1000000)
```

Arguments

snp_map_xT	data frame with 3 initial columns (name, chrom, and position) that match standard SNP map file, followed by 1 column for each tissue with a 0/1 indicator of whether the given SNP passed QC in that tissue. "xT" in the name denotes across tissues
gene_map_xT	data frame with 4 initial columns (name, chrom, and start and end position) that match standard gene map file, followed by 1 column for each tissue with a 0/1 indicator of whether the given gene passed QC in that tissue. "xT" in the name denotes across tissues
nearby	TRUE if hypotheses of interest relate to SNP-gene pairs which are nearby. FALSE if focus is distal regulation. Default value is TRUE
dist	number of base pairs defining "nearby" region. Default is 1,000,000 = 1Mb

Value

Data frame with the SNP name in the first column and the number genes which were tested for association with the SNP in any tissue

Author(s)

Christine B. Peterson

References

Christine B. Peterson, Marina Bogomolov, Yoav Benjamini, and Chiara Sabatti (2016). TreeQTL: hierarchical error control for eQTL findings. *Bioinformatics*. **32**(16): 2556–2558.

Additional publication addressing multi-tissue error control to be submitted soon.

Index

∗Topic **package**
TreeQTL-package, 2

get_eAssociations, *3*, 5
get_eGenes, *2*, *5*, 6
get_eGenes_multi_tissue, 7
get_eSNPs, *2*, *3*, *5*, 9
get_eSNPs_multi_tissue, 10
get_n_tests_per_gene, *5*, *6*, 11
get_n_tests_per_SNP, *5*, *9*, 12
get_n_tests_per_SNP_multi_tissue, 12

TreeQTL *(*TreeQTL-package*)*, 2
TreeQTL-package, 2