

methQTL package

Michael Scherer

Introduction

This vignette describes the *methQTL* R-package available from GitHub. The package uses DNA methylation data obtained using the Illumina BeadArrays, and genotyping data from Illumina genotyping microarrays or whole genome sequencing to compute methylation quantitative trait loci (methQTL). The package provides multiple flavors of linear modeling strategies to compute *methQTL* as statistically significant interactions between single nucleotide polymorphisms (SNPs) and changes in the DNA methylation state of individual CpGs. DNA methylation values at single CpGs are first summarized into correlation blocks, and a representative of this correlation block (tag-CpG) is used for methQTL calling.

Installation

The package relies on some dependencies that are exclusively available from Bioconductor and not from CRAN, most notably *RnBeads*. After installing the dependencies, *methQTL* can be installed from GitHub, after installing the *devtools* package.

```
if(!requireNamespace("devtools")) install.packages("devtools")
source("https://rnbeads.org/data/install.R")
if(!requireNamespace("methQTL")){
  devtools::install_github("MPIIComputationalEpigenetics/methQTL-package")
}
suppressPackageStartupMessages(library(methQTL))
```

Input data

The *methQTL* package requires two types of data as input: DNA methylation data obtained using the Illumina Infinium BeadArrays and genotyping data obtained using genotyping microarrays or whole genome sequencing.

DNA methylation data (microarrays)

The package utilizes the widely used *RnBeads* software package for DNA methylation data import. It supports the various input options available in *RnBeads*, including a direct download from the Gene Expression Omnibus (GEO). For further options, we refer to the *RnBeads* vignette and documentation. In addition to the raw methylation data, a sample annotation sheet specifying the samples

to be analyzed needs to be provided. The sheet contains a line for each sample and looks as follows:

```
SampleID,age,sex,barcode
Sample_1,14,f,209054857842_R01C01
Sample_2,42,f,209054857842_R02C01
Sample_3,45,m,209054857842_R03C01
```

For further details on the import process, we refer to the RnBeads vignette. Most importantly, analysis options need to be specified for the import and preprocessing modules of *RnBeads*. *methQTL* provides a default setting, which is available in *extdata/rnbeads_options.xml*. You can use this file as a template for your own setting and then specify it to the *methQTL* package:

```
opts <- rnb.xml2options(system.file("extdata/rnbeads_options.xml",package="methQTL"))
rnb.options(identifiers.column="ind_IPC")
xml.fi <- file.path(getwd(),"rnbeads_options.xml")
cat(rnb.options2xml(),file=xml.fi)
qtl.setOption(rnbeads.options = xml.fi)
```

To redefine the correlation blocks, we allow for including additional information such as genome-wide segmentation of the methylation landscape (see option `use.segmentation` and function `qtl.run.segmentation`), and also functional annotation according to the Ensembl regulatory build [Zerbino2015].

Genotyping data

PLINK files

The package supports data that has already been processed by *PLINK* and that is available either in the form of binary *.bed*, *.bim* and *.fam* files, as *.ped* and *.map*, as variant calling files (*.vcf*), or as imputed files in the dosage format (*.dos*). For further processing, we use the command line tool *PLINK*, which comes with this package. However, this installation is only valid for Linux systems. For Windows and MacOS users, please install the *PLINK* tool from here and specify it using the option `plink.path`. The sample identifier specified earlier also needs to match the sample IDs of the genotype calls.

IDAT files

The package also supports raw IDAT files and uses the *CRLMM* R-package, together with *PLINK* to perform genotype calling and data import. The package requires a single sample annotation sheet in the format described in the DNA methylation data section. In addition to the column names specified above, a column named *GenoSentrixPosition* has to be added, which specifies the IDAT file IDs.

```
SampleID,age,sex,barcode,GenoSentrixPosition
Sample_1,14,f,209054857842_R01C01,9701756058_R05C01
```

```
Sample_2,42,f,209054857842_R02C01,9701756058_R07C01
Sample_3,45,m,209054857842_R03C01,9742011016_R04C01
```

Imputation

Illumina SNP BeadArray data is typically imputed before further analysis, and the package allows for imputation through the Michigan Imputation Server. In order to be able to perform computation on the server, an account is required. After the account is created, one has to request an API token in the user settings and specify it to the package using the option `imputation.user.token`. During the imputation process, the package will stall for a while and wait for the job to finish. After the job is done, the package will prompt for entering the password send via e-mail to the user account. The imputation process has to be split according to chromosomes, which is why multiple e-mails will be send to the account, and the imputation process can take up to several days. However, after imputation, the imputed data will be available as PLINK files, such that the imputation has to be performed only once. For preprocessing the data for upload to the imputation server, the package requires the bgzip and tabix tools from the htlib package. Also see further options to configure the imputation jobs (see the Michigan Imputation Server documentation for further information):

```
qtl.setOption(
  impute.geno.data=TRUE,
  imputation.reference.panel="apps@hrc-r1.1",
  imputation.phasing.method="shapeit",
  imputation.population="eur"
)
```

Perform data import

The `do.import` function requires the paths to the respective genotyping and DNA methylation data, as well as a sample annotation sheet as discussed earlier. In this vignette, we will describe the import of DNA methylation data in *IDAT* format and genotyping data as *PLINK* files. First, you'll have to specify the paths to the corresponding *IDAT* and *plink* files. Additionally, you have to specify the sample identifier column in the sample annotation sheet that determines the samples in both the genotyping and DNA methylation data. For larger files, we recommend to activate the option to store large matrices on disk rather than in main memory (`hdf5dump`).

```
idat.dir <- "/DEEP_fhgfs/projects/mscherer/data/EPIC/CEDAR/idats/"
plink.dir <- "/DEEP_fhgfs/projects/mscherer/data/450K/CEDAR/publication/139.165.108.18/srv/g
anno.sheet <- "/DEEP_fhgfs/projects/mscherer/data/EPIC/CEDAR/annotation/sample_annotation_ID
qtl.setOption(hdf5dump=TRUE)
setHDF5DumpDir("/DEEP_fhgfs/projects/mscherer/deep/tmp/")
imp.data <- do.import(data.location = c(idat.dir=idat.dir,geno.dir=plink.dir),
  s.anno = anno.sheet,
```

```

s.id.col = "ind_IPC",
tab.sep = "\t")

## 2020-06-24 12:36:27      1.3 STATUS STARTED Import methQTL data
## 2020-06-24 12:36:27      1.3 STATUS      STARTED Processing genotyping data
## 2020-06-24 12:36:36      1.7  INFO          Loading system default for option 'plink.path
## 2020-06-24 12:36:47      2.2 STATUS          STARTED Compute genotype PCA
## 2020-06-24 12:36:48      2.2 STATUS          COMPLETED Compute genotype PCA
## 2020-06-24 12:36:49      2.2 STATUS      COMPLETED Processing genotyping data
## 2020-06-24 12:36:49      2.2 STATUS      STARTED Processing DNA methylation data
## 2020-06-24 12:36:53      2.2 STATUS          STARTED Loading Data from IDAT Files
## 2020-06-24 12:36:53      2.2  INFO          Added column barcode to the provided samp
## 2020-06-24 12:36:54      2.2  INFO          Detected platform: MethylationEPIC
## 2020-06-24 12:37:17      2.4 STATUS      COMPLETED Loading Data from IDAT Files
## 2020-06-24 12:38:39      2.5 STATUS      STARTED Preprocessing
## 2020-06-24 12:38:39      2.5  INFO          Number of cores: 1
## 2020-06-24 12:38:39      2.5 STATUS          STARTED Filtering Procedures I
## 2020-06-24 12:38:41      2.5 STATUS              STARTED Removal of SNP-enriched Sites
## 2020-06-24 12:38:41      2.5 STATUS              Removed 139721 sites using SNP cr
## 2020-06-24 12:38:41      2.5 STATUS              Saved removed sites to /local/tmp
## 2020-06-24 12:38:41      2.5 STATUS              Added a corresponding section to
## 2020-06-24 12:38:41      2.5 STATUS          COMPLETED Removal of SNP-enriched Sit
## 2020-06-24 12:38:41      2.5 STATUS          STARTED Removal of Cross-reactive Pro
## 2020-06-24 12:38:41      2.5 STATUS              Removed 34264 sites
## 2020-06-24 12:38:42      2.5 STATUS              Saved removed sites to /local/tmp
## 2020-06-24 12:38:42      2.5 STATUS              Added a corresponding section to
## 2020-06-24 12:38:42      2.5 STATUS          COMPLETED Removal of Cross-reactive P
## 2020-06-24 12:38:42      2.5  INFO          Working with a p-value threshold of 0
## 2020-06-24 12:38:44      2.6 STATUS          STARTED Greedycut
## 2020-06-24 12:39:05      2.8 STATUS              Calculated a total of 1055 iterat
## 2020-06-24 12:39:05      2.8  INFO              Optimal number of iterations is 1
## 2020-06-24 12:39:09      2.8 STATUS              Created ROC plot
## 2020-06-24 12:39:12      2.8 STATUS              Created line plots for matrix dim
## 2020-06-24 12:39:12      2.8 STATUS              Saved removed sites to /local/tmp
## 2020-06-24 12:39:12      2.8 STATUS          COMPLETED Greedycut
## 2020-06-24 12:39:12      2.8 STATUS              Retained 8 samples and 691856 sites
## 2020-06-24 12:39:12      2.8 STATUS          COMPLETED Filtering Procedures I
## 2020-06-24 12:39:12      2.8 STATUS          STARTED Summary of Filtering Procedures I
## 2020-06-24 12:39:13      2.7 STATUS              Created summary table of removed site
## 2020-06-24 12:39:14      2.7 STATUS              Added summary table of removed and re
## 2020-06-24 12:39:14      2.7  INFO              Subsampling 866895 sites for plotting
## 2020-06-24 12:39:15      2.8 STATUS              Constructed sequences of removed and
## 2020-06-24 12:39:28      2.8 STATUS              Added comparison between removed and
## 2020-06-24 12:39:29      2.8 STATUS          COMPLETED Summary of Filtering Procedures
## 2020-06-24 12:39:29      2.8 STATUS          STARTED Manipulating the object
## 2020-06-24 12:40:29      2.8 STATUS              Removed 175039 sites (probes)

```

##	2020-06-24	12:40:29	2.8	INFO	Retained 691856 sites and 8 samples
##	2020-06-24	12:40:29	2.8	STATUS	COMPLETED Manipulating the object
##	2020-06-24	12:40:29	2.8	STATUS	STARTED Normalization Procedure
##	2020-06-24	12:41:09	3.6	STATUS	Performed normalization with method v
##	2020-06-24	12:42:04	3.3	STATUS	Performed normalization with method '
##	2020-06-24	12:42:22	3.6	STATUS	Added 2D histogram of observed beta v
##	2020-06-24	12:42:22	3.5	STATUS	Added normalization section
##	2020-06-24	12:42:22	3.5	STATUS	COMPLETED Normalization Procedure
##	2020-06-24	12:42:22	3.5	STATUS	STARTED Filtering Procedures II
##	2020-06-24	12:42:24	3.5	STATUS	STARTED Probe Context Removal
##	2020-06-24	12:42:24	3.5	STATUS	Removed 1196 probe(s) having not
##	2020-06-24	12:42:24	3.5	STATUS	Saved removed sites to /local/tmp
##	2020-06-24	12:42:24	3.5	STATUS	Added a corresponding section to
##	2020-06-24	12:42:24	3.5	STATUS	COMPLETED Probe Context Removal
##	2020-06-24	12:42:24	3.5	STATUS	STARTED Removal of Sites on Sex Chron
##	2020-06-24	12:42:24	3.5	STATUS	Removed 16598 site(s) on sex chro
##	2020-06-24	12:42:24	3.5	STATUS	Saved removed sites to /local/tmp
##	2020-06-24	12:42:24	3.5	STATUS	Added a corresponding section to
##	2020-06-24	12:42:24	3.5	STATUS	COMPLETED Removal of Sites on Sex Chr
##	2020-06-24	12:42:24	3.5	STATUS	STARTED Missing Value Removal
##	2020-06-24	12:42:24	3.5	STATUS	Using a sample quantile threshol
##	2020-06-24	12:42:24	3.5	STATUS	Removed 44 site(s) with too many
##	2020-06-24	12:42:25	3.5	STATUS	Saved removed sites to /local/tmp
##	2020-06-24	12:42:30	3.3	STATUS	Added a corresponding section to
##	2020-06-24	12:42:30	3.3	STATUS	COMPLETED Missing Value Removal
##	2020-06-24	12:42:30	3.3	STATUS	Retained 8 samples and 674018 sites
##	2020-06-24	12:42:30	3.3	STATUS	COMPLETED Filtering Procedures II
##	2020-06-24	12:42:30	3.3	STATUS	STARTED Summary of Filtering Procedures I
##	2020-06-24	12:42:30	3.3	STATUS	Created summary table of removed site
##	2020-06-24	12:42:31	3.3	STATUS	Added summary table of removed and re
##	2020-06-24	12:42:31	3.3	INFO	Subsampling 691856 sites for plotting
##	2020-06-24	12:42:32	3.3	STATUS	Constructed sequences of removed and
##	2020-06-24	12:42:39	3.3	STATUS	Added comparison between removed and
##	2020-06-24	12:42:39	3.3	STATUS	COMPLETED Summary of Filtering Procedures
##	2020-06-24	12:42:39	3.3	STATUS	STARTED Manipulating the object
##	2020-06-24	12:43:35	3.1	STATUS	Removed 17838 sites (probes)
##	2020-06-24	12:43:35	3.1	INFO	Retained 674018 sites and 8 samples
##	2020-06-24	12:43:35	3.1	STATUS	COMPLETED Manipulating the object
##	2020-06-24	12:43:35	3.1	INFO	No missing values present, imputation sk
##	2020-06-24	12:43:36	3.1	STATUS	COMPLETED Preprocessing
##	2020-06-24	12:44:06	3.1	STATUS	STARTED Removing 870 CpGs overlapping with S
##	2020-06-24	12:45:02	3.1	STATUS	COMPLETED Removing 870 CpGs overlapping with
##	2020-06-24	12:45:06	3.1	STATUS	COMPLETED Processing DNA methylation data
##	2020-06-24	12:45:06	3.1	STATUS	COMPLETED Import methQTL data

For imputed data, no further processing is performed on the genotyping data

and the dosage values are used as they are:

```
idat.dir <- "/DEEP_fhgfs/projects/mscherer/data/EPIC/CEDAR/idats/"
geno.dir <- "/DEEP_fhgfs/projects/mscherer/data/450K/CEDAR/publication/139.165.108.18/srv/g
anno.sheet <- "/DEEP_fhgfs/projects/mscherer/data/EPIC/CEDAR/annotation/sample_annotation_II
qtl.setOption(hdf5dump=TRUE)
imp.data <- do.import(data.location = c(idat.dir=idat.dir,geno.dir=geno.dir),
                      s.anno = anno.sheet,
                      s.id.col = "ind_IPC",
                      tab.sep = "\t",
                      out.folder = getwd())
```

Please note that the `recode.allele.frequencies` option specifies, if, according to the cohort analyzed, SNP reference and alternative allele are to be recoded according to the allele frequencies found. Alternatively, a path to a local version of dbSNP [Sherry2001] can be provided through `db.snp.ref`, and reference/alternative allele information will be automatically parsed from the database. This is especially crucial, if imputation is to be performed, since the Michigan Imputation Server is sensitive to reference mismatches.

methQTL calling

Although *methQTL* conceptually splits the methQTL calling into two steps ((i) compute correlation block, (ii) call methQTL per correlation block), only a single function call is needed. The function only requires the input `methQTLInput` object produced in the previous step, but further options, such as covariates and the p-value cutoff can be directly specified as a function parameter, or as global parameters using `?qtl.setOption`.

```
meth.qtl.res <- do.methQTL(imp.data)

## 2020-06-24 12:45:06      3.1   INFO Loading default option setting
## 2020-06-24 12:45:08      3.1  STATUS STARTED Imputation procedure knn
## 2020-06-24 12:45:15      3.1  STATUS COMPLETED Imputation procedure knn
##
## 2020-06-24 12:45:18      3.1  STATUS STARTED Computing methQTLs
## 2020-06-24 12:45:18      3.1  STATUS      STARTED Computing methQTL for chromosome chr1
## 2020-06-24 12:45:18      3.1  STATUS      STARTED Compute correlation blocks
## 2020-06-24 12:45:18      3.1   INFO      Split workload, since facing 66034 CpGs
## 2020-06-24 12:45:18      3.1  STATUS      STARTED Compute correlation blocks
## 2020-06-24 12:45:18      3.1  STATUS      STARTED Compute correlation matrix
## 2020-06-24 12:45:52     19.3  STATUS      COMPLETED Compute correlation matrix
## 2020-06-24 12:56:50     13.4  STATUS      STARTED Compute pairwise distances
## 2020-06-24 12:57:15     13.5  STATUS      COMPLETED Compute pairwise distances
## 2020-06-24 12:58:05     10.4  STATUS      STARTED Weight distances
## 2020-06-24 13:03:57     27.3  STATUS      COMPLETED Weight distances
## 2020-06-24 13:03:58     18.5  STATUS      STARTED Compute graph
```

##	2020-06-24	13:04:51	34.7	STATUS	COMPLETED Compute graph
##	2020-06-24	13:04:51	34.7	STATUS	STARTED Compute clustering
##	2020-06-24	13:04:53	34.7	STATUS	COMPLETED Compute clustering
##	2020-06-24	13:04:53	34.7	STATUS	COMPLETED Compute correlation blocks
##	2020-06-24	13:04:53	34.7	STATUS	STARTED Compute correlation blocks
##	2020-06-24	13:04:53	34.7	STATUS	STARTED Compute correlation matrix
##	2020-06-24	13:05:24	51.0	STATUS	COMPLETED Compute correlation matrix
##	2020-06-24	13:16:16	38.4	STATUS	STARTED Compute pairwise distances
##	2020-06-24	13:16:39	20.8	STATUS	COMPLETED Compute pairwise distances
##	2020-06-24	13:17:24	10.4	STATUS	STARTED Weight distances
##	2020-06-24	13:23:36	37.2	STATUS	COMPLETED Weight distances
##	2020-06-24	13:23:38	19.0	STATUS	STARTED Compute graph
##	2020-06-24	13:24:28	35.3	STATUS	COMPLETED Compute graph
##	2020-06-24	13:24:28	35.3	STATUS	STARTED Compute clustering
##	2020-06-24	13:24:29	35.3	STATUS	COMPLETED Compute clustering
##	2020-06-24	13:24:29	35.3	STATUS	COMPLETED Compute correlation blocks
##	2020-06-24	13:24:29	35.3	STATUS	STARTED Compute methQTL per correlation blocks
##	2020-06-24	13:24:29	35.3	STATUS	STARTED Setting up Multicore
##	2020-06-24	13:24:29	35.3	INFO	Using 1 cores
##	2020-06-24	13:24:29	35.3	STATUS	COMPLETED Setting up Multicore
##	2020-06-24	13:24:30	35.3	INFO	No SNP closer than 500000
##	2020-06-24	14:06:58	10.9	INFO	No SNP closer than 500000
##	2020-06-24	14:06:58	10.9	INFO	No SNP closer than 500000
##	2020-06-24	14:06:59	10.9	INFO	No SNP closer than 500000
##	2020-06-24	14:06:59	10.9	INFO	No SNP closer than 500000
##	2020-06-24	14:06:59	10.9	INFO	No SNP closer than 500000
##	2020-06-24	14:07:00	10.9	INFO	No SNP closer than 500000
##	2020-06-24	14:07:00	10.9	INFO	No SNP closer than 500000
##	2020-06-24	14:07:00	10.9	INFO	No SNP closer than 500000
##	2020-06-24	14:07:01	10.9	INFO	No SNP closer than 500000
##	2020-06-24	14:07:01	10.9	INFO	No SNP closer than 500000
##	2020-06-24	14:07:01	10.9	INFO	No SNP closer than 500000
##	2020-06-24	14:45:29	10.9	STATUS	COMPLETED Compute methQTL per correlation blocks
##	2020-06-24	14:45:29	10.9	STATUS	COMPLETED Compute correlation blocks
##	2020-06-24	14:45:29	10.9	STATUS	STARTED Computing methQTL for chromosome chr2
##	2020-06-24	14:45:29	10.9	STATUS	STARTED Compute correlation blocks
##	2020-06-24	14:45:29	10.9	INFO	Split workload, since facing 52088 CpG
##	2020-06-24	14:45:29	10.9	STATUS	STARTED Compute correlation blocks
##	2020-06-24	14:45:29	10.9	STATUS	STARTED Compute correlation matrix
##	2020-06-24	14:45:47	21.0	STATUS	COMPLETED Compute correlation matrix
##	2020-06-24	14:52:30	21.0	STATUS	STARTED Compute pairwise distances
##	2020-06-24	14:52:44	10.9	STATUS	COMPLETED Compute pairwise distances
##	2020-06-24	14:53:15	10.9	STATUS	STARTED Weight distances
##	2020-06-24	14:56:58	17.7	STATUS	COMPLETED Weight distances
##	2020-06-24	14:57:00	15.9	STATUS	STARTED Compute graph
##	2020-06-24	14:57:30	36.2	STATUS	COMPLETED Compute graph

##	2020-06-24	14:57:30	36.2	STATUS	STARTED Compute clustering
##	2020-06-24	14:57:31	36.2	STATUS	COMPLETED Compute clustering
##	2020-06-24	14:57:31	36.2	STATUS	COMPLETED Compute correlation blocks
##	2020-06-24	14:57:31	36.2	STATUS	STARTED Compute correlation blocks
##	2020-06-24	14:57:31	36.2	STATUS	STARTED Compute correlation matrix
##	2020-06-24	14:57:49	41.2	STATUS	COMPLETED Compute correlation matrix
##	2020-06-24	15:04:35	18.5	STATUS	STARTED Compute pairwise distances
##	2020-06-24	15:04:50	14.9	STATUS	COMPLETED Compute pairwise distances
##	2020-06-24	15:05:20	10.8	STATUS	STARTED Weight distances
##	2020-06-24	15:08:56	17.7	STATUS	COMPLETED Weight distances
##	2020-06-24	15:08:57	15.9	STATUS	STARTED Compute graph
##	2020-06-24	15:09:27	36.1	STATUS	COMPLETED Compute graph
##	2020-06-24	15:09:27	36.1	STATUS	STARTED Compute clustering
##	2020-06-24	15:09:28	36.1	STATUS	COMPLETED Compute clustering
##	2020-06-24	15:09:28	36.1	STATUS	COMPLETED Compute correlation blocks
##	2020-06-24	15:09:29	36.1	STATUS	STARTED Compute methQTL per chromosome
##	2020-06-24	15:09:29	36.1	STATUS	STARTED Setting up Multicore
##	2020-06-24	15:09:29	36.1	INFO	Using 1 cores
##	2020-06-24	15:09:29	36.1	STATUS	COMPLETED Setting up Multicore
##	2020-06-24	16:28:32	10.8	STATUS	COMPLETED Compute methQTL per chromosome
##	2020-06-24	16:28:32	10.8	STATUS	COMPLETED Compute correlation blocks
##	2020-06-24	16:28:32	10.8	STATUS	STARTED Computing methQTL for chromosome
##	2020-06-24	16:28:32	10.8	STATUS	STARTED Compute correlation blocks
##	2020-06-24	16:28:32	10.8	STATUS	STARTED Compute correlation matrix
##	2020-06-24	16:29:20	34.5	STATUS	COMPLETED Compute correlation matrix
##	2020-06-24	16:44:55	22.7	STATUS	STARTED Compute pairwise distances
##	2020-06-24	16:45:25	18.8	STATUS	COMPLETED Compute pairwise distances
##	2020-06-24	16:46:32	14.1	STATUS	STARTED Weight distances
##	2020-06-24	16:54:56	39.2	STATUS	COMPLETED Weight distances
##	2020-06-24	16:54:58	25.9	STATUS	STARTED Compute graph
##	2020-06-24	16:58:36	49.6	STATUS	COMPLETED Compute graph
##	2020-06-24	16:58:36	49.6	STATUS	STARTED Compute clustering
##	2020-06-24	16:58:37	49.6	STATUS	COMPLETED Compute clustering
##	2020-06-24	16:58:37	49.6	STATUS	COMPLETED Compute correlation blocks
##	2020-06-24	16:58:38	49.6	STATUS	STARTED Compute methQTL per chromosome
##	2020-06-24	16:58:38	49.6	STATUS	STARTED Setting up Multicore
##	2020-06-24	16:58:38	49.6	INFO	Using 1 cores
##	2020-06-24	16:58:38	49.6	STATUS	COMPLETED Setting up Multicore
##	2020-06-24	18:00:13	10.9	STATUS	COMPLETED Compute methQTL per chromosome
##	2020-06-24	18:00:13	10.9	STATUS	COMPLETED Computing methQTL for chromosome
##	2020-06-24	18:00:13	10.9	STATUS	STARTED Computing methQTL for chromosome
##	2020-06-24	18:00:13	10.9	STATUS	STARTED Compute correlation blocks
##	2020-06-24	18:00:13	10.9	STATUS	STARTED Compute correlation matrix
##	2020-06-24	18:00:38	23.4	STATUS	COMPLETED Compute correlation matrix
##	2020-06-24	18:08:52	10.8	STATUS	STARTED Compute pairwise distances
##	2020-06-24	18:09:10	17.9	STATUS	COMPLETED Compute pairwise distances

##	2020-06-24	18:09:44	10.8	STATUS	STARTED Weight distances
##	2020-06-24	18:14:09	24.8	STATUS	COMPLETED Weight distances
##	2020-06-24	18:14:11	17.1	STATUS	STARTED Compute graph
##	2020-06-24	18:14:45	44.2	STATUS	COMPLETED Compute graph
##	2020-06-24	18:14:45	44.2	STATUS	STARTED Compute clustering
##	2020-06-24	18:14:46	44.2	STATUS	COMPLETED Compute clustering
##	2020-06-24	18:14:46	44.2	STATUS	COMPLETED Compute correlation blocks
##	2020-06-24	18:14:47	44.2	STATUS	STARTED Compute methQTL per correlation blocks
##	2020-06-24	18:14:47	44.2	STATUS	STARTED Setting up Multicore
##	2020-06-24	18:14:47	44.2	INFO	Using 1 cores
##	2020-06-24	18:14:47	44.2	STATUS	COMPLETED Setting up Multicore
##	2020-06-24	19:04:08	10.8	STATUS	COMPLETED Compute methQTL per correlation blocks
##	2020-06-24	19:04:08	10.8	STATUS	COMPLETED Computing methQTL for chromosomes
##	2020-06-24	19:04:08	10.8	STATUS	STARTED Computing methQTL for chromosomes
##	2020-06-24	19:04:08	10.8	STATUS	STARTED Compute correlation blocks
##	2020-06-24	19:04:08	10.8	STATUS	STARTED Compute correlation matrix
##	2020-06-24	19:04:43	30.0	STATUS	COMPLETED Compute correlation matrix
##	2020-06-24	19:17:14	30.0	STATUS	STARTED Compute pairwise distances
##	2020-06-24	19:17:40	18.0	STATUS	COMPLETED Compute pairwise distances
##	2020-06-24	19:18:34	11.8	STATUS	STARTED Weight distances
##	2020-06-24	19:25:22	28.5	STATUS	COMPLETED Weight distances
##	2020-06-24	19:25:23	21.4	STATUS	STARTED Compute graph
##	2020-06-24	19:26:24	40.6	STATUS	COMPLETED Compute graph
##	2020-06-24	19:26:24	40.6	STATUS	STARTED Compute clustering
##	2020-06-24	19:26:25	40.6	STATUS	COMPLETED Compute clustering
##	2020-06-24	19:26:25	40.6	STATUS	COMPLETED Compute correlation blocks
##	2020-06-24	19:26:26	40.6	STATUS	STARTED Compute methQTL per correlation blocks
##	2020-06-24	19:26:26	40.6	STATUS	STARTED Setting up Multicore
##	2020-06-24	19:26:26	40.6	INFO	Using 1 cores
##	2020-06-24	19:26:26	40.6	STATUS	COMPLETED Setting up Multicore
##	2020-06-24	19:46:13	10.9	INFO	No SNP closer than 500000
##	2020-06-24	19:46:13	10.9	INFO	No SNP closer than 500000
##	2020-06-24	20:20:40	10.9	STATUS	COMPLETED Compute methQTL per correlation blocks
##	2020-06-24	20:20:40	10.9	STATUS	COMPLETED Computing methQTL for chromosomes
##	2020-06-24	20:20:40	10.9	STATUS	STARTED Computing methQTL for chromosomes
##	2020-06-24	20:20:40	10.9	STATUS	STARTED Compute correlation blocks
##	2020-06-24	20:20:40	10.9	INFO	Split workload, since facing 4199
##	2020-06-24	20:20:40	10.9	STATUS	STARTED Compute correlation blocks
##	2020-06-24	20:20:40	10.9	STATUS	STARTED Compute correlation matrix
##	2020-06-24	20:20:51	14.2	STATUS	COMPLETED Compute correlation matrix
##	2020-06-24	20:24:54	22.9	STATUS	STARTED Compute pairwise distances
##	2020-06-24	20:25:03	10.9	STATUS	COMPLETED Compute pairwise distances
##	2020-06-24	20:25:23	10.8	STATUS	STARTED Weight distances
##	2020-06-24	20:27:39	12.4	STATUS	COMPLETED Weight distances
##	2020-06-24	20:27:41	12.4	STATUS	STARTED Compute graph
##	2020-06-24	20:28:01	25.5	STATUS	COMPLETED Compute graph

##	2020-06-24	20:28:01	25.5	STATUS	STARTED Compute clustering
##	2020-06-24	20:28:03	25.5	STATUS	COMPLETED Compute clustering
##	2020-06-24	20:28:03	25.5	STATUS	COMPLETED Compute correlation blocks
##	2020-06-24	20:28:03	25.5	STATUS	STARTED Compute correlation blocks
##	2020-06-24	20:28:03	25.5	STATUS	STARTED Compute correlation blocks
##	2020-06-24	20:28:16	28.8	STATUS	COMPLETED Compute correlation blocks
##	2020-06-24	20:32:24	14.1	STATUS	STARTED Compute pairwise distances
##	2020-06-24	20:32:34	10.8	STATUS	COMPLETED Compute pairwise distances
##	2020-06-24	20:32:53	10.8	STATUS	STARTED Weight distances
##	2020-06-24	20:35:23	14.0	STATUS	COMPLETED Weight distances
##	2020-06-24	20:35:24	12.4	STATUS	STARTED Compute graph
##	2020-06-24	20:35:45	25.5	STATUS	COMPLETED Compute graph
##	2020-06-24	20:35:45	25.5	STATUS	STARTED Compute clustering
##	2020-06-24	20:35:46	25.5	STATUS	COMPLETED Compute clustering
##	2020-06-24	20:35:46	25.5	STATUS	COMPLETED Compute correlation blocks
##	2020-06-24	20:35:46	25.5	STATUS	STARTED Compute methQTL per chromosome
##	2020-06-24	20:35:46	25.5	STATUS	STARTED Setting up Multicore
##	2020-06-24	20:35:46	25.5	INFO	Using 1 cores
##	2020-06-24	20:35:46	25.5	STATUS	COMPLETED Setting up Multicore
##	2020-06-24	21:37:16	10.8	STATUS	COMPLETED Compute methQTL per chromosome
##	2020-06-24	21:37:16	10.8	STATUS	COMPLETED Compute correlation blocks
##	2020-06-24	21:37:16	10.8	STATUS	COMPLETED Computing methQTL for chromosomes
##	2020-06-24	21:37:16	10.8	STATUS	STARTED Compute correlation blocks
##	2020-06-24	21:37:16	10.8	STATUS	STARTED Compute correlation blocks
##	2020-06-24	21:37:56	30.9	STATUS	COMPLETED Compute correlation blocks
##	2020-06-24	21:51:24	20.9	STATUS	STARTED Compute pairwise distances
##	2020-06-24	21:51:50	16.3	STATUS	COMPLETED Compute pairwise distances
##	2020-06-24	21:53:01	12.3	STATUS	STARTED Weight distances
##	2020-06-24	22:00:27	30.6	STATUS	COMPLETED Weight distances
##	2020-06-24	22:00:29	22.3	STATUS	STARTED Compute graph
##	2020-06-24	22:04:16	42.4	STATUS	COMPLETED Compute graph
##	2020-06-24	22:04:16	42.4	STATUS	STARTED Compute clustering
##	2020-06-24	22:04:18	42.4	STATUS	COMPLETED Compute clustering
##	2020-06-24	22:04:18	42.4	STATUS	COMPLETED Compute correlation blocks
##	2020-06-24	22:04:18	42.4	STATUS	STARTED Compute methQTL per chromosome
##	2020-06-24	22:04:18	42.4	STATUS	STARTED Setting up Multicore
##	2020-06-24	22:04:18	42.4	INFO	Using 1 cores
##	2020-06-24	22:04:18	42.4	STATUS	COMPLETED Setting up Multicore
##	2020-06-24	22:53:27	10.9	STATUS	COMPLETED Compute methQTL per chromosome
##	2020-06-24	22:53:27	10.9	STATUS	COMPLETED Computing methQTL for chromosomes
##	2020-06-24	22:53:27	10.9	STATUS	STARTED Computing methQTL for chromosomes
##	2020-06-24	22:53:27	10.9	STATUS	STARTED Compute correlation blocks
##	2020-06-24	22:53:27	10.9	STATUS	STARTED Compute correlation blocks
##	2020-06-24	22:53:52	24.6	STATUS	COMPLETED Compute correlation blocks
##	2020-06-24	23:02:59	11.3	STATUS	STARTED Compute pairwise distances
##	2020-06-24	23:03:19	19.5	STATUS	COMPLETED Compute pairwise distances

##	2020-06-24	23:03:57	10.8	STATUS	STARTED Weight distances
##	2020-06-24	23:10:09	23.3	STATUS	COMPLETED Weight distances
##	2020-06-24	23:10:10	17.6	STATUS	STARTED Compute graph
##	2020-06-24	23:10:56	31.4	STATUS	COMPLETED Compute graph
##	2020-06-24	23:10:56	31.4	STATUS	STARTED Compute clustering
##	2020-06-24	23:10:57	31.4	STATUS	COMPLETED Compute clustering
##	2020-06-24	23:10:57	31.4	STATUS	COMPLETED Compute correlation block
##	2020-06-24	23:10:58	31.4	STATUS	STARTED Compute methQTL per chromosome
##	2020-06-24	23:10:58	31.4	STATUS	STARTED Setting up Multicore
##	2020-06-24	23:10:58	31.4	INFO	Using 1 cores
##	2020-06-24	23:10:58	31.4	STATUS	COMPLETED Setting up Multicore
##	2020-06-24	23:58:15	10.8	STATUS	COMPLETED Compute methQTL per chromosome
##	2020-06-24	23:58:15	10.8	STATUS	COMPLETED Computing methQTL for chromosomes
##	2020-06-24	23:58:15	10.8	STATUS	STARTED Computing methQTL for chromosomes
##	2020-06-24	23:58:15	10.8	STATUS	STARTED Compute correlation block
##	2020-06-24	23:58:15	10.8	STATUS	STARTED Compute correlation block
##	2020-06-24	23:58:27	14.1	STATUS	COMPLETED Compute correlation block
##	2020-06-25	00:02:35	10.9	STATUS	STARTED Compute pairwise distances
##	2020-06-25	00:02:45	10.8	STATUS	COMPLETED Compute pairwise distances
##	2020-06-25	00:03:04	10.8	STATUS	STARTED Weight distances
##	2020-06-25	00:05:21	16.3	STATUS	COMPLETED Weight distances
##	2020-06-25	00:05:22	14.0	STATUS	STARTED Compute graph
##	2020-06-25	00:05:42	27.0	STATUS	COMPLETED Compute graph
##	2020-06-25	00:05:42	27.0	STATUS	STARTED Compute clustering
##	2020-06-25	00:05:42	27.0	STATUS	COMPLETED Compute clustering
##	2020-06-25	00:05:42	27.0	STATUS	COMPLETED Compute correlation block
##	2020-06-25	00:05:43	27.0	STATUS	STARTED Compute methQTL per chromosome
##	2020-06-25	00:05:43	27.0	STATUS	STARTED Setting up Multicore
##	2020-06-25	00:05:43	27.0	INFO	Using 1 cores
##	2020-06-25	00:05:43	27.0	STATUS	COMPLETED Setting up Multicore
##	2020-06-25	00:17:26	10.8	INFO	No SNP closer than 500000
##	2020-06-25	00:17:26	10.8	INFO	No SNP closer than 500000
##	2020-06-25	00:17:27	10.8	INFO	No SNP closer than 500000
##	2020-06-25	00:17:27	10.8	INFO	No SNP closer than 500000
##	2020-06-25	00:17:27	10.8	INFO	No SNP closer than 500000
##	2020-06-25	00:17:28	10.8	INFO	No SNP closer than 500000
##	2020-06-25	00:17:29	10.8	INFO	No SNP closer than 500000
##	2020-06-25	00:17:30	10.8	INFO	No SNP closer than 500000
##	2020-06-25	00:17:30	10.8	INFO	No SNP closer than 500000
##	2020-06-25	00:17:30	10.8	INFO	No SNP closer than 500000
##	2020-06-25	00:17:31	10.8	INFO	No SNP closer than 500000
##	2020-06-25	00:17:31	10.8	INFO	No SNP closer than 500000
##	2020-06-25	00:17:31	10.8	INFO	No SNP closer than 500000
##	2020-06-25	00:17:32	10.8	INFO	No SNP closer than 500000
##	2020-06-25	00:17:32	10.8	INFO	No SNP closer than 500000
##	2020-06-25	00:17:32	10.8	INFO	No SNP closer than 500000

##	2020-06-25	00:17:36	10.8	INFO	No SNP closer than 500000
##	2020-06-25	00:17:37	10.8	INFO	No SNP closer than 500000
##	2020-06-25	00:17:37	10.8	INFO	No SNP closer than 500000
##	2020-06-25	00:45:06	10.8	STATUS	COMPLETED Compute methQTL per chromosome
##	2020-06-25	00:45:06	10.8	STATUS	COMPLETED Computing methQTL for chromosomes
##	2020-06-25	00:45:06	10.8	STATUS	STARTED Computing methQTL for chromosomes
##	2020-06-25	00:45:06	10.8	STATUS	STARTED Compute correlation block
##	2020-06-25	00:45:06	10.8	STATUS	STARTED Compute correlation matrix
##	2020-06-25	00:45:35	27.5	STATUS	COMPLETED Compute correlation matrix
##	2020-06-25	00:57:42	19.2	STATUS	STARTED Compute pairwise distances
##	2020-06-25	00:58:05	14.2	STATUS	COMPLETED Compute pairwise distances
##	2020-06-25	00:58:54	10.9	STATUS	STARTED Weight distances
##	2020-06-25	01:05:00	30.6	STATUS	COMPLETED Weight distances
##	2020-06-25	01:05:02	19.2	STATUS	STARTED Compute graph
##	2020-06-25	01:06:12	35.9	STATUS	COMPLETED Compute graph
##	2020-06-25	01:06:12	35.9	STATUS	STARTED Compute clustering
##	2020-06-25	01:06:13	35.9	STATUS	COMPLETED Compute clustering
##	2020-06-25	01:06:13	35.9	STATUS	COMPLETED Compute correlation block
##	2020-06-25	01:06:13	35.9	STATUS	STARTED Compute methQTL per chromosome
##	2020-06-25	01:06:13	35.9	STATUS	STARTED Setting up Multicore
##	2020-06-25	01:06:13	35.9	INFO	Using 1 cores
##	2020-06-25	01:06:13	35.9	STATUS	COMPLETED Setting up Multicore
##	2020-06-25	01:55:42	10.9	STATUS	COMPLETED Compute methQTL per chromosome
##	2020-06-25	01:55:42	10.9	STATUS	COMPLETED Computing methQTL for chromosomes
##	2020-06-25	01:55:42	10.9	STATUS	STARTED Computing methQTL for chromosomes
##	2020-06-25	01:55:42	10.9	STATUS	STARTED Compute correlation block
##	2020-06-25	01:55:42	10.9	STATUS	STARTED Compute correlation matrix
##	2020-06-25	01:56:27	34.3	STATUS	COMPLETED Compute correlation matrix
##	2020-06-25	02:11:46	34.5	STATUS	STARTED Compute pairwise distances
##	2020-06-25	02:12:17	21.7	STATUS	COMPLETED Compute pairwise distances
##	2020-06-25	02:13:21	14.0	STATUS	STARTED Weight distances
##	2020-06-25	02:21:43	42.8	STATUS	COMPLETED Weight distances
##	2020-06-25	02:21:45	25.7	STATUS	STARTED Compute graph
##	2020-06-25	02:28:53	49.1	STATUS	COMPLETED Compute graph
##	2020-06-25	02:28:53	49.1	STATUS	STARTED Compute clustering
##	2020-06-25	02:28:55	49.1	STATUS	COMPLETED Compute clustering
##	2020-06-25	02:28:55	49.1	STATUS	COMPLETED Compute correlation block
##	2020-06-25	02:28:56	49.1	STATUS	STARTED Compute methQTL per chromosome
##	2020-06-25	02:28:56	49.1	STATUS	STARTED Setting up Multicore
##	2020-06-25	02:28:56	49.1	INFO	Using 1 cores
##	2020-06-25	02:28:56	49.1	STATUS	COMPLETED Setting up Multicore
##	2020-06-25	03:15:44	10.9	STATUS	COMPLETED Compute methQTL per chromosome
##	2020-06-25	03:15:44	10.9	STATUS	COMPLETED Computing methQTL for chromosomes
##	2020-06-25	03:15:44	10.9	STATUS	STARTED Computing methQTL for chromosomes
##	2020-06-25	03:15:44	10.9	STATUS	STARTED Compute correlation block
##	2020-06-25	03:15:44	10.9	STATUS	STARTED Compute correlation matrix

##	2020-06-25	03:16:18	30.4	STATUS	COMPLETED Compute correlation
##	2020-06-25	03:29:17	14.9	STATUS	STARTED Compute pairwise dist
##	2020-06-25	03:29:45	26.7	STATUS	COMPLETED Compute pairwise di
##	2020-06-25	03:30:37	12.1	STATUS	STARTED Weight distances
##	2020-06-25	03:37:35	30.0	STATUS	COMPLETED Weight distances
##	2020-06-25	03:37:37	21.8	STATUS	STARTED Compute graph
##	2020-06-25	03:38:39	41.3	STATUS	COMPLETED Compute graph
##	2020-06-25	03:38:39	41.3	STATUS	STARTED Compute clustering
##	2020-06-25	03:38:40	41.3	STATUS	COMPLETED Compute clustering
##	2020-06-25	03:38:40	41.3	STATUS	COMPLETED Compute correlation blo
##	2020-06-25	03:38:41	41.3	STATUS	STARTED Compute methQTL per corre
##	2020-06-25	03:38:41	41.3	STATUS	STARTED Setting up Multicore
##	2020-06-25	03:38:41	41.3	INFO	Using 1 cores
##	2020-06-25	03:38:41	41.3	STATUS	COMPLETED Setting up Multicore
##	2020-06-25	04:26:03	10.9	STATUS	COMPLETED Compute methQTL per cor
##	2020-06-25	04:26:03	10.9	STATUS	COMPLETED Computing methQTL for chrom
##	2020-06-25	04:26:03	10.9	STATUS	STARTED Computing methQTL for chromos
##	2020-06-25	04:26:03	10.9	STATUS	STARTED Compute correlation block
##	2020-06-25	04:26:03	10.9	STATUS	STARTED Compute correlation m
##	2020-06-25	04:26:11	12.9	STATUS	COMPLETED Compute correlation
##	2020-06-25	04:28:44	15.6	STATUS	STARTED Compute pairwise dist
##	2020-06-25	04:28:51	10.9	STATUS	COMPLETED Compute pairwise di
##	2020-06-25	04:29:05	10.9	STATUS	STARTED Weight distances
##	2020-06-25	04:30:29	14.5	STATUS	COMPLETED Weight distances
##	2020-06-25	04:30:31	12.1	STATUS	STARTED Compute graph
##	2020-06-25	04:30:39	19.2	STATUS	COMPLETED Compute graph
##	2020-06-25	04:30:39	19.2	STATUS	STARTED Compute clustering
##	2020-06-25	04:30:40	19.2	STATUS	COMPLETED Compute clustering
##	2020-06-25	04:30:40	19.2	STATUS	COMPLETED Compute correlation blo
##	2020-06-25	04:30:40	19.2	STATUS	STARTED Compute methQTL per corre
##	2020-06-25	04:30:40	19.2	STATUS	STARTED Setting up Multicore
##	2020-06-25	04:30:40	19.2	INFO	Using 1 cores
##	2020-06-25	04:30:40	19.2	STATUS	COMPLETED Setting up Multicore
##	2020-06-25	05:02:03	10.9	STATUS	COMPLETED Compute methQTL per cor
##	2020-06-25	05:02:03	10.9	STATUS	COMPLETED Computing methQTL for chrom
##	2020-06-25	05:02:03	10.9	STATUS	STARTED Computing methQTL for chromos
##	2020-06-25	05:02:03	10.9	STATUS	STARTED Compute correlation block
##	2020-06-25	05:02:03	10.9	STATUS	STARTED Compute correlation m
##	2020-06-25	05:02:19	19.2	STATUS	COMPLETED Compute correlation
##	2020-06-25	05:07:37	19.2	STATUS	STARTED Compute pairwise dist
##	2020-06-25	05:07:50	10.9	STATUS	COMPLETED Compute pairwise di
##	2020-06-25	05:08:17	10.9	STATUS	STARTED Weight distances
##	2020-06-25	05:11:16	15.6	STATUS	COMPLETED Weight distances
##	2020-06-25	05:11:18	15.0	STATUS	STARTED Compute graph
##	2020-06-25	05:11:42	31.6	STATUS	COMPLETED Compute graph
##	2020-06-25	05:11:42	31.6	STATUS	STARTED Compute clustering

## 2020-06-25 05:11:43	31.6	STATUS	COMPLETED Compute clustering
## 2020-06-25 05:11:43	31.6	STATUS	COMPLETED Compute correlation block
## 2020-06-25 05:11:44	31.6	STATUS	STARTED Compute methQTL per chromosome
## 2020-06-25 05:11:44	31.6	STATUS	STARTED Setting up Multicore
## 2020-06-25 05:11:44	31.6	INFO	Using 1 cores
## 2020-06-25 05:11:44	31.6	STATUS	COMPLETED Setting up Multicore
## 2020-06-25 05:41:57	10.9	INFO	No SNP closer than 500000
## 2020-06-25 05:42:11	10.9	STATUS	COMPLETED Compute methQTL per chromosome
## 2020-06-25 05:42:11	10.9	STATUS	COMPLETED Computing methQTL for chromosomes
## 2020-06-25 05:42:11	10.9	STATUS	STARTED Computing methQTL for chromosomes
## 2020-06-25 05:42:11	10.9	STATUS	STARTED Compute correlation block
## 2020-06-25 05:42:11	10.9	STATUS	STARTED Compute correlation block
## 2020-06-25 05:42:24	14.7	STATUS	COMPLETED Compute correlation block
## 2020-06-25 05:47:10	10.9	STATUS	STARTED Compute pairwise distances
## 2020-06-25 05:47:21	11.9	STATUS	COMPLETED Compute pairwise distances
## 2020-06-25 05:47:43	10.8	STATUS	STARTED Weight distances
## 2020-06-25 05:50:25	16.1	STATUS	COMPLETED Weight distances
## 2020-06-25 05:50:26	14.2	STATUS	STARTED Compute graph
## 2020-06-25 05:50:50	29.5	STATUS	COMPLETED Compute graph
## 2020-06-25 05:50:50	29.5	STATUS	STARTED Compute clustering
## 2020-06-25 05:50:50	29.5	STATUS	COMPLETED Compute clustering
## 2020-06-25 05:50:50	29.5	STATUS	COMPLETED Compute correlation block
## 2020-06-25 05:50:51	29.5	STATUS	STARTED Compute methQTL per chromosome
## 2020-06-25 05:50:51	29.5	STATUS	STARTED Setting up Multicore
## 2020-06-25 05:50:51	29.5	INFO	Using 1 cores
## 2020-06-25 05:50:51	29.5	STATUS	COMPLETED Setting up Multicore
## 2020-06-25 05:50:53	29.5	INFO	No SNP closer than 500000
## 2020-06-25 05:50:53	29.5	INFO	No SNP closer than 500000
## 2020-06-25 05:50:54	29.5	INFO	No SNP closer than 500000
## 2020-06-25 05:50:54	29.5	INFO	No SNP closer than 500000
## 2020-06-25 05:50:54	29.5	INFO	No SNP closer than 500000
## 2020-06-25 05:50:55	29.5	INFO	No SNP closer than 500000
## 2020-06-25 05:50:55	29.5	INFO	No SNP closer than 500000
## 2020-06-25 05:50:55	29.5	INFO	No SNP closer than 500000
## 2020-06-25 05:50:56	29.5	INFO	No SNP closer than 500000
## 2020-06-25 06:21:27	10.8	STATUS	COMPLETED Compute methQTL per chromosome
## 2020-06-25 06:21:27	10.8	STATUS	COMPLETED Computing methQTL for chromosomes
## 2020-06-25 06:21:27	10.8	STATUS	STARTED Computing methQTL for chromosomes
## 2020-06-25 06:21:27	10.8	STATUS	STARTED Compute correlation block
## 2020-06-25 06:21:27	10.8	STATUS	STARTED Compute correlation block
## 2020-06-25 06:21:51	23.8	STATUS	COMPLETED Compute correlation block
## 2020-06-25 06:30:21	10.8	STATUS	STARTED Compute pairwise distances
## 2020-06-25 06:30:39	12.2	STATUS	COMPLETED Compute pairwise distances
## 2020-06-25 06:31:18	10.8	STATUS	STARTED Weight distances
## 2020-06-25 06:35:56	21.4	STATUS	COMPLETED Weight distances
## 2020-06-25 06:35:58	17.3	STATUS	STARTED Compute graph

##	2020-06-25	06:36:34	45.3	STATUS	COMPLETED Compute graph
##	2020-06-25	06:36:34	45.3	STATUS	STARTED Compute clustering
##	2020-06-25	06:36:35	45.3	STATUS	COMPLETED Compute clustering
##	2020-06-25	06:36:35	45.3	STATUS	COMPLETED Compute correlation block
##	2020-06-25	06:36:36	45.3	STATUS	STARTED Compute methQTL per chromosome
##	2020-06-25	06:36:36	45.3	STATUS	STARTED Setting up Multicore
##	2020-06-25	06:36:36	45.3	INFO	Using 1 cores
##	2020-06-25	06:36:36	45.3	STATUS	COMPLETED Setting up Multicore
##	2020-06-25	06:47:40	10.8	INFO	No SNP closer than 500000
##	2020-06-25	06:47:41	10.8	INFO	No SNP closer than 500000
##	2020-06-25	06:47:41	10.8	INFO	No SNP closer than 500000
##	2020-06-25	06:47:41	10.8	INFO	No SNP closer than 500000
##	2020-06-25	06:47:42	10.8	INFO	No SNP closer than 500000
##	2020-06-25	06:47:42	10.8	INFO	No SNP closer than 500000
##	2020-06-25	06:47:42	10.8	INFO	No SNP closer than 500000
##	2020-06-25	06:47:43	10.8	INFO	No SNP closer than 500000
##	2020-06-25	07:04:16	10.8	STATUS	COMPLETED Compute methQTL per chromosome
##	2020-06-25	07:04:16	10.8	STATUS	COMPLETED Computing methQTL for chromosomes
##	2020-06-25	07:04:16	10.8	STATUS	STARTED Computing methQTL for chromosomes
##	2020-06-25	07:04:16	10.8	STATUS	STARTED Compute correlation block
##	2020-06-25	07:04:16	10.8	STATUS	STARTED Compute correlation block
##	2020-06-25	07:04:51	29.7	STATUS	COMPLETED Compute correlation block
##	2020-06-25	07:17:12	20.4	STATUS	STARTED Compute pairwise distances
##	2020-06-25	07:17:40	23.6	STATUS	COMPLETED Compute pairwise distances
##	2020-06-25	07:18:29	11.7	STATUS	STARTED Weight distances
##	2020-06-25	07:25:14	30.4	STATUS	COMPLETED Weight distances
##	2020-06-25	07:25:16	21.1	STATUS	STARTED Compute graph
##	2020-06-25	07:28:32	40.0	STATUS	COMPLETED Compute graph
##	2020-06-25	07:28:32	40.0	STATUS	STARTED Compute clustering
##	2020-06-25	07:28:34	40.0	STATUS	COMPLETED Compute clustering
##	2020-06-25	07:28:34	40.0	STATUS	COMPLETED Compute correlation block
##	2020-06-25	07:28:35	40.0	STATUS	STARTED Compute methQTL per chromosome
##	2020-06-25	07:28:35	40.0	STATUS	STARTED Setting up Multicore
##	2020-06-25	07:28:35	40.0	INFO	Using 1 cores
##	2020-06-25	07:28:35	40.0	STATUS	COMPLETED Setting up Multicore
##	2020-06-25	07:56:18	10.9	STATUS	COMPLETED Compute methQTL per chromosome
##	2020-06-25	07:56:18	10.9	STATUS	COMPLETED Computing methQTL for chromosomes
##	2020-06-25	07:56:18	10.9	STATUS	STARTED Computing methQTL for chromosomes
##	2020-06-25	07:56:18	10.9	STATUS	STARTED Compute correlation block
##	2020-06-25	07:56:18	10.9	STATUS	STARTED Compute correlation block
##	2020-06-25	07:56:22	11.9	STATUS	COMPLETED Compute correlation block
##	2020-06-25	07:57:42	11.9	STATUS	STARTED Compute pairwise distances
##	2020-06-25	07:57:47	10.9	STATUS	COMPLETED Compute pairwise distances
##	2020-06-25	07:57:54	10.9	STATUS	STARTED Weight distances
##	2020-06-25	07:58:37	14.0	STATUS	COMPLETED Weight distances
##	2020-06-25	07:58:39	11.9	STATUS	STARTED Compute graph

##	2020-06-25	07:58:42	11.9	STATUS	COMPLETED Compute graph
##	2020-06-25	07:58:42	11.9	STATUS	STARTED Compute clustering
##	2020-06-25	07:58:43	11.9	STATUS	COMPLETED Compute clustering
##	2020-06-25	07:58:43	11.9	STATUS	COMPLETED Compute correlation block
##	2020-06-25	07:58:43	11.9	STATUS	STARTED Compute methQTL per chromosome
##	2020-06-25	07:58:43	11.9	STATUS	STARTED Setting up Multicore
##	2020-06-25	07:58:43	11.9	INFO	Using 1 cores
##	2020-06-25	07:58:43	11.9	STATUS	COMPLETED Setting up Multicore
##	2020-06-25	08:20:22	10.9	STATUS	COMPLETED Compute methQTL per chromosome
##	2020-06-25	08:20:22	10.9	STATUS	COMPLETED Computing methQTL for chromosomes
##	2020-06-25	08:20:22	10.9	STATUS	STARTED Computing methQTL for chromosomes
##	2020-06-25	08:20:22	10.9	STATUS	STARTED Compute correlation block
##	2020-06-25	08:20:22	10.9	STATUS	STARTED Compute correlation block
##	2020-06-25	08:20:49	24.6	STATUS	COMPLETED Compute correlation block
##	2020-06-25	08:30:11	24.6	STATUS	STARTED Compute pairwise distances
##	2020-06-25	08:30:30	12.9	STATUS	COMPLETED Compute pairwise distances
##	2020-06-25	08:31:11	10.8	STATUS	STARTED Weight distances
##	2020-06-25	08:36:11	22.2	STATUS	COMPLETED Weight distances
##	2020-06-25	08:36:12	17.6	STATUS	STARTED Compute graph
##	2020-06-25	08:36:57	31.4	STATUS	COMPLETED Compute graph
##	2020-06-25	08:36:57	31.4	STATUS	STARTED Compute clustering
##	2020-06-25	08:36:59	31.4	STATUS	COMPLETED Compute clustering
##	2020-06-25	08:36:59	31.4	STATUS	COMPLETED Compute correlation block
##	2020-06-25	08:36:59	31.4	STATUS	STARTED Compute methQTL per chromosome
##	2020-06-25	08:36:59	31.4	STATUS	STARTED Setting up Multicore
##	2020-06-25	08:36:59	31.4	INFO	Using 1 cores
##	2020-06-25	08:36:59	31.4	STATUS	COMPLETED Setting up Multicore
##	2020-06-25	08:55:20	10.8	STATUS	COMPLETED Compute methQTL per chromosome
##	2020-06-25	08:55:20	10.8	STATUS	COMPLETED Computing methQTL for chromosomes
##	2020-06-25	08:55:20	10.8	STATUS	STARTED Computing methQTL for chromosomes
##	2020-06-25	08:55:20	10.8	STATUS	STARTED Compute correlation block
##	2020-06-25	08:55:20	10.8	STATUS	STARTED Compute correlation block
##	2020-06-25	08:55:30	13.5	STATUS	COMPLETED Compute correlation block
##	2020-06-25	08:58:47	17.9	STATUS	STARTED Compute pairwise distances
##	2020-06-25	08:58:55	10.8	STATUS	COMPLETED Compute pairwise distances
##	2020-06-25	08:59:12	10.8	STATUS	STARTED Weight distances
##	2020-06-25	09:01:02	17.7	STATUS	COMPLETED Weight distances
##	2020-06-25	09:01:04	13.5	STATUS	STARTED Compute graph
##	2020-06-25	09:01:16	18.7	STATUS	COMPLETED Compute graph
##	2020-06-25	09:01:16	18.7	STATUS	STARTED Compute clustering
##	2020-06-25	09:01:17	18.7	STATUS	COMPLETED Compute clustering
##	2020-06-25	09:01:17	18.7	STATUS	COMPLETED Compute correlation block
##	2020-06-25	09:01:17	18.7	STATUS	STARTED Compute methQTL per chromosome
##	2020-06-25	09:01:17	18.7	STATUS	STARTED Setting up Multicore
##	2020-06-25	09:01:17	18.7	INFO	Using 1 cores
##	2020-06-25	09:01:17	18.7	STATUS	COMPLETED Setting up Multicore

##	2020-06-25	09:22:49	10.8	STATUS	COMPLETED Compute methQTL per chromosome
##	2020-06-25	09:22:49	10.8	STATUS	COMPLETED Computing methQTL for chromosome
##	2020-06-25	09:22:49	10.8	STATUS	STARTED Computing methQTL for chromosome
##	2020-06-25	09:22:49	10.8	STATUS	STARTED Compute correlation block
##	2020-06-25	09:22:49	10.8	STATUS	STARTED Compute correlation block
##	2020-06-25	09:22:51	11.3	STATUS	COMPLETED Compute correlation block
##	2020-06-25	09:23:30	11.3	STATUS	STARTED Compute pairwise distances
##	2020-06-25	09:23:33	10.8	STATUS	COMPLETED Compute pairwise distances
##	2020-06-25	09:23:38	10.8	STATUS	STARTED Weight distances
##	2020-06-25	09:23:51	10.8	STATUS	COMPLETED Weight distances
##	2020-06-25	09:23:52	10.8	STATUS	STARTED Compute graph
##	2020-06-25	09:23:53	10.8	STATUS	COMPLETED Compute graph
##	2020-06-25	09:23:53	10.8	STATUS	STARTED Compute clustering
##	2020-06-25	09:23:53	10.8	STATUS	COMPLETED Compute clustering
##	2020-06-25	09:23:53	10.8	STATUS	COMPLETED Compute correlation block
##	2020-06-25	09:23:54	10.8	STATUS	STARTED Compute methQTL per chromosome
##	2020-06-25	09:23:54	10.8	STATUS	STARTED Setting up Multicore
##	2020-06-25	09:23:54	10.8	INFO	Using 1 cores
##	2020-06-25	09:23:54	10.8	STATUS	COMPLETED Setting up Multicore
##	2020-06-25	09:23:54	10.8	INFO	No SNP closer than 500000
##	2020-06-25	09:23:54	10.8	INFO	No SNP closer than 500000
##	2020-06-25	09:23:54	10.8	INFO	No SNP closer than 500000
##	2020-06-25	09:34:58	10.8	STATUS	COMPLETED Compute methQTL per chromosome
##	2020-06-25	09:34:58	10.8	STATUS	COMPLETED Computing methQTL for chromosome
##	2020-06-25	09:34:58	10.8	STATUS	STARTED Computing methQTL for chromosome
##	2020-06-25	09:34:58	10.8	STATUS	STARTED Compute correlation block
##	2020-06-25	09:34:58	10.8	STATUS	STARTED Compute correlation block
##	2020-06-25	09:35:03	12.4	STATUS	COMPLETED Compute correlation block
##	2020-06-25	09:37:16	10.8	STATUS	STARTED Compute pairwise distances
##	2020-06-25	09:37:22	10.8	STATUS	COMPLETED Compute pairwise distances
##	2020-06-25	09:37:34	10.8	STATUS	STARTED Weight distances
##	2020-06-25	09:38:43	10.8	STATUS	COMPLETED Weight distances
##	2020-06-25	09:38:45	10.8	STATUS	STARTED Compute graph
##	2020-06-25	09:38:55	15.4	STATUS	COMPLETED Compute graph
##	2020-06-25	09:38:55	15.4	STATUS	STARTED Compute clustering
##	2020-06-25	09:38:55	15.4	STATUS	COMPLETED Compute clustering
##	2020-06-25	09:38:55	15.4	STATUS	COMPLETED Compute correlation block
##	2020-06-25	09:38:56	15.4	STATUS	STARTED Compute methQTL per chromosome
##	2020-06-25	09:38:56	15.4	STATUS	STARTED Setting up Multicore
##	2020-06-25	09:38:56	15.4	INFO	Using 1 cores
##	2020-06-25	09:38:56	15.4	STATUS	COMPLETED Setting up Multicore
##	2020-06-25	09:52:52	10.8	STATUS	COMPLETED Compute methQTL per chromosome
##	2020-06-25	09:52:52	10.8	STATUS	COMPLETED Computing methQTL for chromosome
##	2020-06-25	09:53:00	10.8	STATUS	COMPLETED Computing methQTL for chromosome

We will now present the two steps of the methQTL calling procedure in more

detail.

Compute CpG correlation blocks

Since neighboring CpGs are often highly correlated, using each CpG independently as a potential methQTL candidate leads to many redundant results. We thus aimed to approximate *DNA methylation haplotypes* by determining highly correlated CpGs in close vicinity. The procedure itself is split into six steps, and is performed for each chromosome independently:

1. Compute the (Pearson) correlation matrix between all CpGs (further correlation types available in option `correlation.type`)
2. Construct the distance matrix from the correlation matrix
3. Discard all interactions with a correlation lower than a given threshold (option: `cluster.cor.threshold`)
4. Weight the distance according to the genomic distance between the two CpGs with a Gaussian (option: `standard.deviation.gauss`). Higher values for the standard deviation lead to a lower penalty on distal CpGs, thus the clusters will become larger.
5. Discard all interactions ranging longer than the option `absolute.distance.cutoff`
6. Compute the Louvain clustering on the undirected, weighted graph induced by the distance matrix

This will return a clustering according to the correlation structure between neighboring CpGs that we will later use for methQTL calling. Note that we used simulation experiments to determine the parameters for each data type individually. They will be automatically loaded for the dataset that is used and are:

- **450k:** `cluster.cor.threshold=0.2, standard.deviation.gauss=5,000, absolute.distance.cutoff=500,000`
- **EPIC:** `cluster.cor.threshold=0.2, standard.deviation.gauss=3,000, absolute.distance.cutoff=500,000`
- **RRBS/WGBS:** `cluster.cor.threshold=0.2, standard.deviation.gauss=250, absolute.distance.cutoff=500,000`

Call methQTL per correlation block

From the list of correlation blocks, *methQTL* computes methQTL interactions with all SNPs on the same chromosome. The process is split into three steps:

1. Compute a representative CpG (tag-CpG) per correlation block, as specified with the option `representative.cpg.computation` (default: `row.medians`).
2. Discard all SNPs that are further than `absolute.distance.cutoff` (default: 1,000,000) away from the representative CpG
3. Call methQTL by using linear models. Multiple options of methQTL calling are available and can be selected via the option `linear.model.type`

(default: *classical.linear*). Alternatively, *fastQTL* can be set as an option for `meth.qtl.type`. This will tell the package to use the fastQTL software [Ongen2016].

The `meth.qtl.type` tells, how a methQTL interaction is defined and provides three options, in addition to the already mentioned *fastQTL*:

1. *oneVSall*: A CpG can only be influenced by one SNP. We choose the one with the lowest p-value.
2. *twoVSall*: A CpG can both positively and negatively be influenced by two independent SNPs. The package will output those fulfilling the p-value cutoff.
3. *allVSall*: For each CpG, all SNPs showing a p-value lower than the p-value cutoff will be returned.

In the latest stage, potential covariates can be specified using the option *sel.covariates*. We recommend to include at least *age* and *sex* as covariates, as they have a strong influence on the DNA methylation pattern.

Downstream analysis and interpretation

How to use *methQTLResult*

The above procedure will create an object of class `methQTLResult`, which contains the methQTL that are called in the previous step. To get a table of all the methQTL, you need to extract the information from the object. In the majority of the function calls below, there is the option `type`, which takes on the values:

- * ‘SNP’: To characterize the SNPs that influence any DNA methylation state
- * ‘CpG’: To characterize the representative CpGs per correlation block that are influenced by any genotype
- * ‘cor.block’: To characterize all CpGs, which are part of a correlation block, whose representative CpG is influenced by any genotype

Furthermore, you can obtain genomic annotations for both the CpGs and the SNPs involved in the methQTL interactions:

```
result.table <- getResult(meth.qtl.res)
head(result.table)
```

##	CpG	SNP	Beta	SE.Beta	P.value	Chromosome
## 59	cg00090105	rs185580	-0.1662472	0.009373958	2.064279e-06	chr1
## 80	cg00102184	rs6685121	0.1838105	0.012800985	7.142140e-06	chr1
## 76	cg00207189	rs4847021	-0.2098301	0.008611342	3.140927e-07	chr1
## 75	cg00553601	rs2176600	-0.3349320	0.018777130	1.995429e-06	chr1
## 62	cg00972755	rs185580	-0.2248021	0.015113092	5.809094e-06	chr1
## 26	cg01023592	rs863087	-0.1164647	0.007975658	6.472797e-06	chr1
##	Position.CpG	Position.SNP	Distance	p.val.adj.fdr		
## 59	182538701	182519861	18840	1		

```
## 80      240194783      239774670      420113      1
## 76      231244795      230959298      285497      1
## 75      224268136      224395782      -127646      1
## 62      182859418      182519861      339557      1
## 26      64435362      64525961      -90599      1
```

```
anno.meth <- getAnno(meth.qtl.res,"meth")
head(anno.meth)
```

```
##      Chromosome  Start    End Strand Strand.1 AddressA AddressB Design
## cg26928153      chr1  10848  10849      -      - 91693541 47784201      I
## cg16269199      chr1  10850  10851      -      - 82663207 3701821      I
## cg24669183      chr1 534242 534243      -      - 12706847      NA      II
## cg15560884      chr1 710097 710098      +      + 66790119      NA      II
## cg01014490      chr1 714177 714178      -      - 1645492 53610280      I
## cg10692041      chr1 716245 716246      +      + 80722913      NA      II
##      Color Context Random HumanMethylation27 HumanMethylation450
## cg26928153      Grn      CG      FALSE      FALSE      NA
## cg16269199      Grn      CG      FALSE      FALSE      NA
## cg24669183      Both     CG      FALSE      FALSE      TRUE
## cg15560884      Both     CG      FALSE      FALSE      TRUE
## cg01014490      Red      CG      FALSE      FALSE      TRUE
## cg10692041      Both     CG      FALSE      FALSE      NA
##      Mismatches.A Mismatches.B CGI.Relation CpG GC SNPs.3 SNPs.5
## cg26928153      0      0      Open Sea  15 74      0      0
## cg16269199      0      0      Open Sea  15 74      0      0
## cg24669183      0      0      South Shore 2 49      0      0
## cg15560884      0      0      North Shelf 2 28      0      0
## cg01014490      0      0      Island  9 67      0      0
## cg10692041      0      0      South Shore 2 47      0      0
##      SNPs.Full Cross.reactive
## cg26928153      0      0
## cg16269199      0      0
## cg24669183      0      0
## cg15560884      0      0
## cg01014490      0      0
## cg10692041      0      0
```

```
anno.geno <- getAnno(meth.qtl.res,"geno")
head(anno.geno)
```

```
##      Chromosome  Start cM Allele.1 Allele.2 Allele.1.Freq Allele.2.Freq
## rs3094315      chr1 752566 NA      A      G      0.7500      0.2500
## rs3131972      chr1 752721 NA      G      A      0.7500      0.2500
## rs12124819      chr1 776546 NA      A      G      0.7500      0.2500
## rs11240777      chr1 798959 NA      G      A      0.7500      0.2500
## rs4970383      chr1 838555 NA      C      A      0.5625      0.4375
## rs4475691      chr1 846808 NA      C      T      0.6875      0.3125
```

For more detailed information about the output, also see the function `getResults.GWASMap`.

Plots

To visualize methQTL, the package provides some plotting functions. All functions return an object of type `ggplot`, which can be subsequently stored or viewed. Either all methQTL can be simultaneously visualized in a single plot, or a specific methQTL can be visualized:

```
result.table <- result.table[order(result.table$P.value,decreasing=F),]  
qtl.plot.SNP.CpG.interaction(imp.data,result.table$CpG[1],result.table$SNP[1])
```

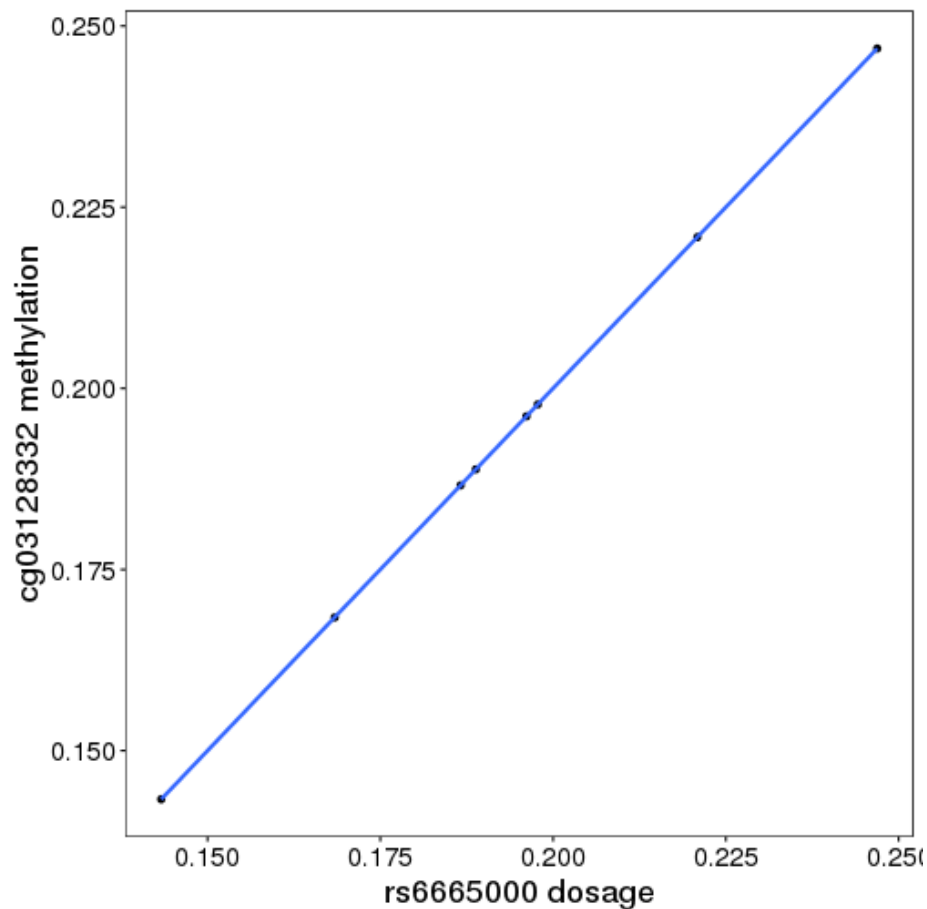


Figure 1: plot of chunk plots

```
qtl.distance.scatterplot(meth.qtl.res)
```

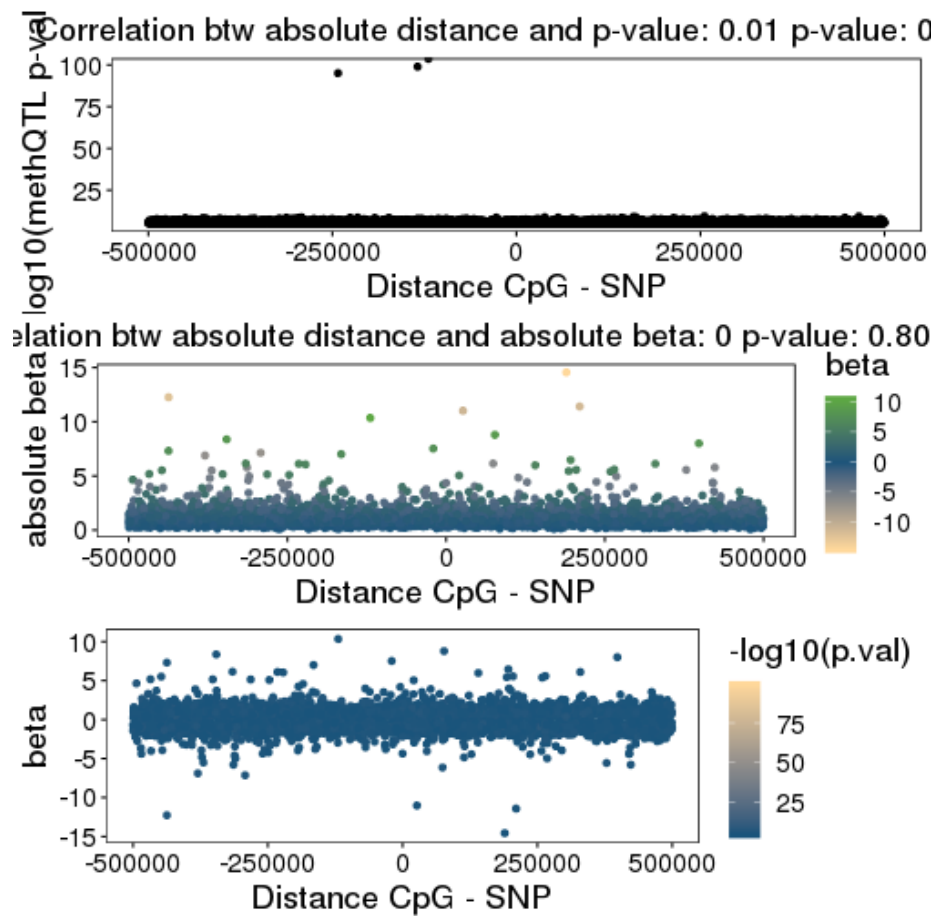


Figure 2: plot of chunk plots

Interpretation functions

The package provides a bunch of interpretation functions to characterize the detected methQTLs. This includes LOLA enrichment analysis[@LOLA] (`qtl.lola.enrichment`), genomic annotation enrichment based on putative regulatory elements defined by the Ensembl Regulatory Build[@Zerbino2015] (`qtl.annotation.enrichment`), enrichment analysis of different base substitutions in SNPs (`qtl.base.substitution.enrichment`), or TFBS motif enrichment using TFBSTools. Enrichment is compared for the methQTLs that are available in the provided `methQTLResult` (for a single input), or to the overlapping QTLs for a list of `methQTLResult`. The background of the enrichment is defined as all the SNPs/CpGs that have been used as input to the methQTL calling.

```
res <- qtl.base.substitution.enrichment(meth.qtl.res)
qtl.plot.base.substitution(meth.qtl.res,merge=TRUE)
```

Lists of methQTL results

Most of the functions discussed above either support a single `methQTLResult` as input, or a list of such objects. In case a list is specified, the functions will typically overlap the methQTLs found and compare those with all SNPs/CpGs that have been used for methQTL calling. Additionally, there are functions that particularly work on a list of `methQTLResult` objects and that perform overlapping, or determine the methQTLs specific to a dataset.

```
meth.qtl.list <- list(First=meth.qtl.res.1,Second=meth.qtl.res.2,Third=meth.qtl.res.3)
qtl.venn.plot(meth.qtl.list)
qtl.upset.plot(meth.qtl.list,type = "cor.block")
spec.first <- get.specific.qtl(meth.qtl.list$First,meth.qtl.list[-1])
```

Advanced configuration

Employ methQTL on a scientific compute cluster

methQTL can automatically distribute jobs across a high performance compute cluster, which has been setup using the Sun Grid Engine (SGE) technology. You can pass the option `cluster.submit` to `do.methQTL` and thus activate the cluster submission. Note that you'll also have to specify a path to an executable *Rscript* and potentially specify resource requirements using the option setting `cluster.config`.

```
qtl.setOption(cluster.config = c(h_vmem="60G",mem_free="20G"))
qtl.setOption(rsript.path = "/usr/bin/Rscript")
meth.qtl.res <- do.methQTL(meth.qtl = imp.data,
                          cluster.submit = T)
```

References