Accessing results from "A covariance-enhanced approach to multi-tissue joint eQTL mapping with application to transcriptome-wide association studies"

Aaron J. Molstad, Wei Sun, and Li Hsu

In this document, we provide a simple example showing how to access the estimated multi-tissue eQTL weights from the article "A covariance-enhanced approach to multi-tissue joint eQTL mapping with application to transcriptome wide association studies". First, the user must download and unzip "MTeQTLResults.zip". Then, the user must first specify the path to the directory "MTeQTLResults", e.g.,

```
filepath <- "~/Documents/GitHub/MTeQTLResults/Supp_CodeResults"
```

From here, the user must source the function used to access results:

```
library(Matrix)
source(paste(filepath, "/MTeQTLResults/getMTeQTLResults.R", sep = ""))
str(getMTeQTLResults)
```

```
## function (filepath, gene = NULL, chr = NULL, pos = NULL, extract = NULL)
```

The function used is named `getMTeQTLResults`. This function takes five arguments:

1. `filepath` : the filepath as specified above

2. `gene` : a list of gene names for which information should be extracted (e.g., "ENSG00000143196")

3. `chr` : a list of chromosomes corresponding the positions in the `pos` argument (e.g., 1)

4. `pos` : a list of positions on a chromosome which must be the same length as `chr`. If both `chr` and `pos` are specified, results will be returned for all available genes where `pos` is within 1Mb of the TSS or TSE.

5. `extract`: a list with the items to be extracted – options are:
   - `beta`: the eQTL weights for each gene (stored as an $r \times 29$ matrix) with column names denoting the tissues for which each column represents the eQTL weights;
   - `Omega`: the estimated error precision matrix for each gene (stored as a 29×29 matrix);
   - `TSS`: the transcription start site of the gene;
   - `TES`: the transcription end site of the gene;
   - `R2`: a vector of length 29 with the testing set R^2 from the GTEx analysis described in the article;
   - `chr`: the chromosome on which the gene is located.

For example, suppose we wanted the eQTL weights and test set R^2 for the gene ENSG00000143196:

```
temp <- getMTeQTLResults(filepath, gene = list("ENSG00000143196"), chr = NULL, pos = NULL,
    extract = list("beta", "R2"))
str(temp)
```

```
## List of 1
##  $ :List of 3
##   ..$ gene: chr "ENSG00000143196"
```

```
##     ..$ beta: num [1:8, 1:29] 0.001496 -0.001205 -0.000804 0.00213 0.01096 ...
##     .. ..- attr(*, "dimnames")=List of 2
##     .. .. ..$ : chr [1:8] "1:168684832:AAAG:A" "1:168737147:A:G" "1:168652905:G:C" "1:168666520:T:A" .
##     .. .. ..$ : chr [1:29] "Adipose - Subcutaneous" "Adipose - Visceral (Omentum)" "Adrenal Gland" "Ar
##     ..$ R2  : num [1:29] 0.01861 0.02533 0.06454 0.11827 0.00763 ...
```

It can be seen that the results are stored as a list with the same length as the list `gene`. Within each element of that list are all the items included in the list `extract` plus the gene name. For example, the eQTL weights for this particular gene are 8×29:

```
image(Matrix(temp[[1]]$beta), xlab = "Tissues", ylab = "SNPs")
```
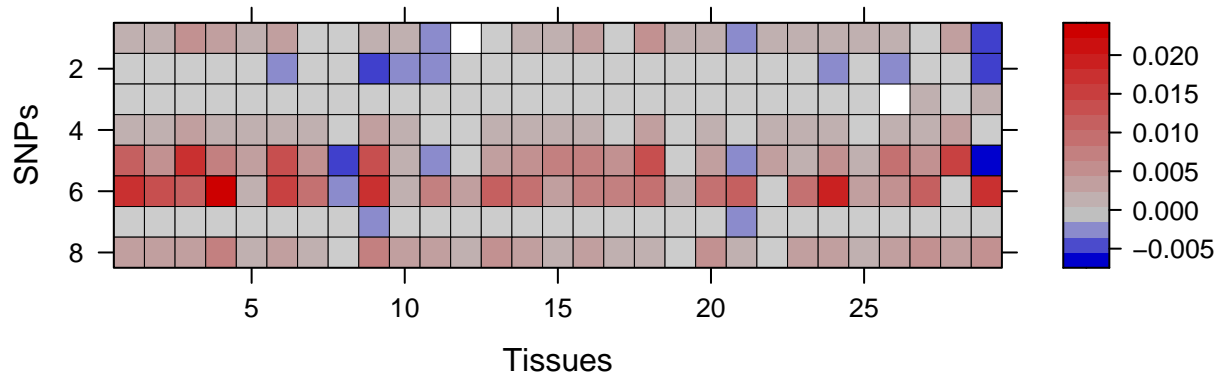

Dimensions: 8 x 29

In the above image, the white cell is the only which has two coefficient values exactly equal to zero. This is an example of a gene where of the eight identified eQTLs, seven are shared across all 29 tissues and one is shared across 28 tissues.

We can also check the testing set R^2 for each tissue type for this particular gene.

```
plot(temp[[1]]$R2, pch = 20, ylab = "Testing set R2", xlab = "Tissues")
abline(h = 0, lty = 2)
```


2
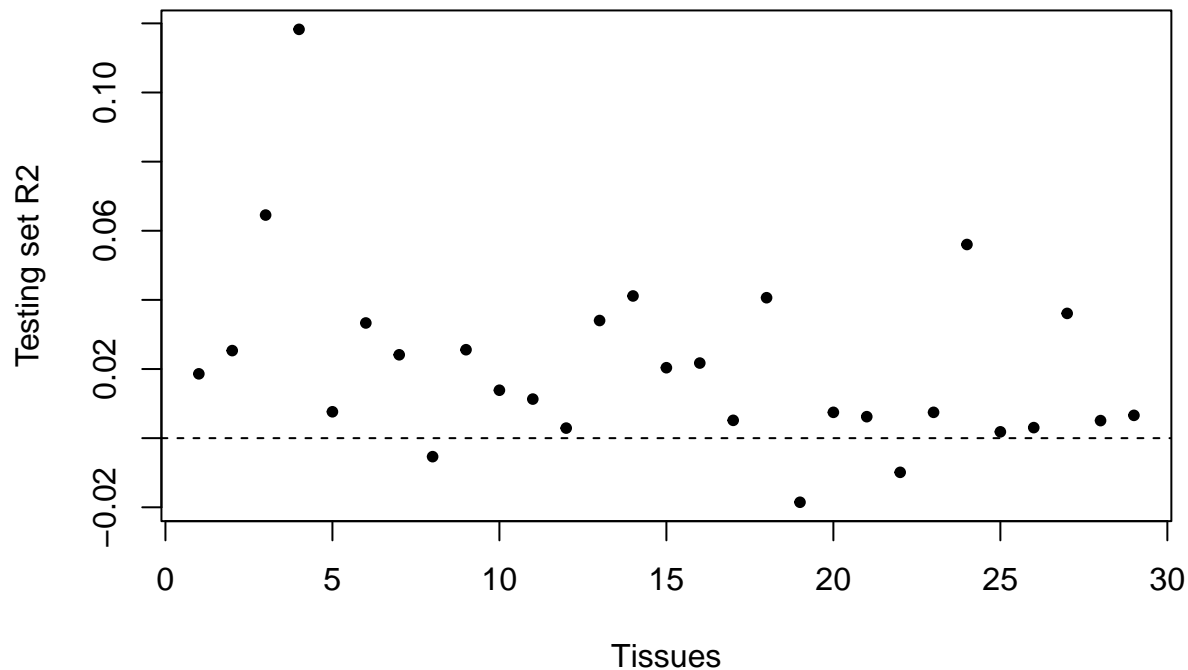
To see which tissue type has relatively high R^2, we use the following:

```
colnames(temp[[1]]$beta)[which.max(temp[[1]]$R2)]
```

```
## [1] "Artery - Aorta"
```

Similarly, one can input a longer list of genes and get out the estimated precision matrices for both genes.

```
temp2 <- getMTeQTLResults(filepath, gene = list("ENSG00000143196", "ENSG00000144401"),
    chr = NULL, pos = NULL, extract = list("Omega", "R2"))
str(temp2)
```

```
## List of 2
##  $ :List of 3
##   ..$ gene : chr "ENSG00000143196"
##   ..$ Omega: num [1:29, 1:29] 0.921 0 0 0 0 ...
##   .. ..- attr(*, "dimnames")=List of 2
##   .. .. ..$ : NULL
##   .. .. ..$ : chr [1:29] "Adipose - Subcutaneous" "Adipose - Visceral (Omentum)" "Adrenal Gland" "Ar
##   ..$ R2   : num [1:29] 0.01861 0.02533 0.06454 0.11827 0.00763 ...
##  $ :List of 3
##   ..$ gene : chr "ENSG00000144401"
##   ..$ Omega: num [1:29, 1:29] 1.03 -6.54e-21 0.00 0.00 0.00 ...
##   .. ..- attr(*, "dimnames")=List of 2
##   .. .. ..$ : NULL
##   .. .. ..$ : chr [1:29] "Adipose - Subcutaneous" "Adipose - Visceral (Omentum)" "Adrenal Gland" "Ar
##   ..$ R2   : num [1:29] 0.0737 0.0518 0.103 0.063 0.0095 ...
```

```
image(Matrix(temp2[[2]]$Omega), ylab = "Tissues", xlab = "", main = "Estimated error precision matrix fo
```

Estimated error precision matrix for ENSG00000144401

Dimensions: 29 x 29
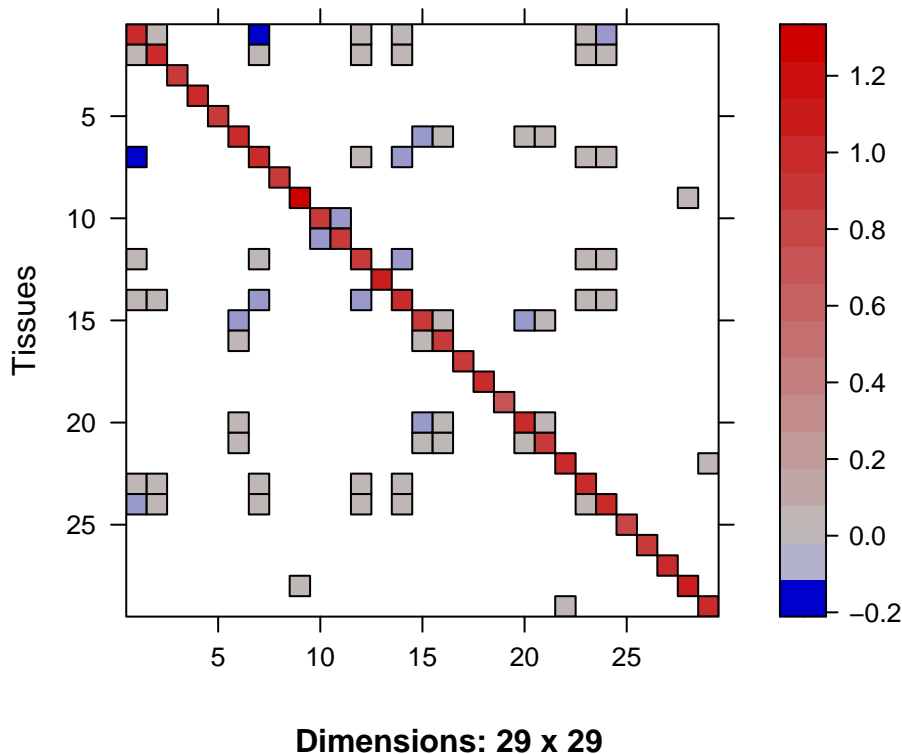
If the user would rather input positions and chromosomes than genes, the following can be used:

```
temp3 <- getMTeQTLResults(filepath, gene = NULL, chr = list("8", "1"), pos = list("48890700",
    "10243857"), extract = list("chr", "TSS", "TES", "beta", "R2"))
str(temp3)
```

```
## List of 2
##  $ :List of 6
##   ..$ gene: chr "ENSG00000104738"
##   ..$ beta: num [1:7, 1:29] -0.002355 -0.001128 -0.000111 -0.000258 0.001064 ...
##   .. ..- attr(*, "dimnames")=List of 2
##   .. .. ..$ : chr [1:7] "8:49039681:C:T" "8:48585975:T:C" "8:49106595:G:A" "8:48709340:T:TA" ...
##   .. .. ..$ : chr [1:29] "Adipose - Subcutaneous" "Adipose - Visceral (Omentum)" "Adrenal Gland" "Ar
##   ..$ R2  : num [1:29] 0.0988 0.0131 0.0432 0.0251 0.0203 ...
##   ..$ chr : num 8
##   ..$ TSS : num 48872745
##   ..$ TES : num 48890720
##  $ :List of 6
##   ..$ gene: chr "ENSG00000131591"
##   ..$ beta: num [1:11, 1:29] -0.000478 -0.023313 0.001289 -0.017133 0.002412 ...
##   .. ..- attr(*, "dimnames")=List of 2
##   .. .. ..$ : chr [1:11] "1:1005806:C:T" "1:1034243:C:T" "1:1048989:A:G" "1:1030374:G:A" ...
##   .. .. ..$ : chr [1:29] "Adipose - Subcutaneous" "Adipose - Visceral (Omentum)" "Adrenal Gland" "Ar
##   ..$ R2  : num [1:29] 0.0876 0.0959 0.1628 0.1551 0.0794 ...
##   ..$ chr : num 1
##   ..$ TSS : num 1017198
##   ..$ TES : num 1051741
```