

# Read in BMI GIANT GWAS meta-analysis data

*Simina Boca, Jeff Leek*

*2016-10-26*

Load the relevant libraries first:

```
library(readr)
library(dplyr)
library(reshape2)
library(curl)
library(Hmisc)
```

Now load the data (first need to download the file from [https://www.broadinstitute.org/collaboration/giant/index.php/GIANT/\\_consortium\\_data/\\_files#GWAS/\\_Anthropometric/\\_2015/\\_BMI](https://www.broadinstitute.org/collaboration/giant/index.php/GIANT/_consortium_data/_files#GWAS/_Anthropometric/_2015/_BMI)):

```
##change the directory name
tot = read_tsv("../../../../../Dropbox/Research/FDR projects/Prior prob regression/code/All_ancestries_SNP.
dim(tot)
```

```
## [1] 2555510      8
```

Do some basic data filtering (to remove missing values) and data cleaning (switch to minor allele from effect allele, add variable with frequency intervals):

```
tot = filter(tot,!is.na(Freq1.Hapmap))
dim(tot)
```

```
## [1] 2500573      8
```

```
##change the frequency to the minor allele frequency
changeWhich <- tot$Freq1.Hapmap > 0.5
tot$Freq1.Hapmap[changeWhich] <- 1-tot$Freq1.Hapmap[changeWhich]
colnames(tot)[colnames(tot)=="Freq1.Hapmap"] <- "Freq_MAF_Hapmap"
##also need to change the coefficient sign for these alleles
tot$b[changeWhich] <- -tot$b[changeWhich]
tot = mutate(tot,Freq_MAF_Int_Hapmap=cut2(Freq_MAF_Hapmap,g=3)) ##added a new variable with frequency in
table(tot$Freq_MAF_Int_Hapmap)
```

```
##
## [0.000,0.127) [0.127,0.302) [0.302,0.500]
##      838070      850600      811903
```

Save in a data frame in a .RData file:

```
BMI_GIANT_GWAS <- tot
save(x=BMI_GIANT_GWAS, file="BMI_GIANT_GWAS.RData")
```

## Session Information

```
devtools::session_info()
```

```
## Session info -----

## setting value
## version R version 3.3.1 (2016-06-21)
## system x86_64, mingw32
## ui RTerm
## language (EN)
## collate English_United States.1252
## tz America/New_York
## date 2016-10-26

## Packages -----

## package * version date source
## acepack 1.3-3.3 2014-11-24 CRAN (R 3.3.0)
## assertthat 0.1 2013-12-06 CRAN (R 3.3.1)
## BiocStyle * 2.0.2 2016-05-16 Bioconductor
## chron 2.3-47 2015-06-24 CRAN (R 3.3.1)
## cluster 2.0.4 2016-04-18 CRAN (R 3.3.1)
## codetools 0.2-14 2015-07-15 CRAN (R 3.3.1)
## colorspace 1.2-6 2015-03-11 CRAN (R 3.3.1)
## curl * 1.0 2016-07-24 CRAN (R 3.3.1)
## data.table 1.9.6 2015-09-19 CRAN (R 3.3.1)
## DBI 0.4-1 2016-05-08 CRAN (R 3.3.1)
## devtools 1.12.0 2016-06-24 CRAN (R 3.3.1)
## digest 0.6.9 2016-01-08 CRAN (R 3.3.1)
## dplyr * 0.5.0 2016-06-24 CRAN (R 3.3.1)
## evaluate 0.9 2016-04-29 CRAN (R 3.3.1)
## foreign 0.8-66 2015-08-19 CRAN (R 3.3.1)
## formatR 1.4 2016-05-09 CRAN (R 3.3.1)
## Formula * 1.2-1 2015-04-07 CRAN (R 3.3.0)
## ggplot2 * 2.1.0 2016-03-01 CRAN (R 3.3.1)
## gridExtra 2.2.1 2016-02-29 CRAN (R 3.3.1)
## gtable 0.2.0 2016-02-26 CRAN (R 3.3.1)
## Hmisc * 3.17-4 2016-05-02 CRAN (R 3.3.1)
## htmltools 0.3.5 2016-03-21 CRAN (R 3.3.1)
## knitr 1.13 2016-05-09 CRAN (R 3.3.1)
## lattice * 0.20-33 2015-07-14 CRAN (R 3.3.1)
## latticeExtra 0.6-28 2016-02-09 CRAN (R 3.3.1)
## lazyeval 0.2.0 2016-06-12 CRAN (R 3.3.1)
## magrittr 1.5 2014-11-22 CRAN (R 3.3.1)
## Matrix 1.2-6 2016-05-02 CRAN (R 3.3.1)
## memoise 1.0.0 2016-01-29 CRAN (R 3.3.1)
## munsell 0.4.3 2016-02-13 CRAN (R 3.3.1)
## nnet 7.3-12 2016-02-02 CRAN (R 3.3.1)
## plyr 1.8.4 2016-06-08 CRAN (R 3.3.1)
## R6 2.1.2 2016-01-26 CRAN (R 3.3.1)
## RColorBrewer 1.1-2 2014-12-07 CRAN (R 3.3.0)
```

##	Rcpp	0.12.6	2016-07-19	CRAN	(R 3.3.1)
##	readr	* 0.2.2	2015-10-22	CRAN	(R 3.3.1)
##	reshape2	* 1.4.1	2014-12-06	CRAN	(R 3.3.1)
##	rmarkdown	1.0	2016-07-08	CRAN	(R 3.3.1)
##	rpart	4.1-10	2015-06-29	CRAN	(R 3.3.1)
##	scales	0.4.0	2016-02-26	CRAN	(R 3.3.1)
##	stringi	1.1.1	2016-05-27	CRAN	(R 3.3.0)
##	stringr	1.0.0	2015-04-30	CRAN	(R 3.3.1)
##	survival	* 2.39-4	2016-05-11	CRAN	(R 3.3.1)
##	tibble	1.1	2016-07-04	CRAN	(R 3.3.1)
##	withr	1.0.2	2016-06-20	CRAN	(R 3.3.1)
##	yaml	2.1.13	2014-06-12	CRAN	(R 3.3.1)