

BGN239 Homework T1D GWAS

Aditi Jantikar

T1D Summary statistics

Read in summary statistics from the following 2021 genome-wide association study of type 1 diabetes (T1D). This study was led by Josh Chiou in the Gaulton lab at UCSD.

<https://www.nature.com/articles/s41586-021-03552-w>

The full dataset contains >62 million SNPs. I have filtered it to remove all SNPs with p values > 0.001, leaving about 600 thousand SNPs.

```
t1d.gwas.tab <- read.table("t1d_gwas_hg38_pruned2_w_genome_pos.tsv.gz", header=T)
head(t1d.gwas.tab)

##      SNP.ID chrom chrom.pos.hg38 genome.pos.hg38 allele1 allele2     p.val
## 1  rs549065172    chr1        61578        61578      T      C 0.007370
## 2  rs1467103792    chr1       135215       135215      A      G 0.000575
## 3  rs772450029    chr1       138495       138495      T      C 0.000184
## 4  rs181202492    chr1       597067       597067      G      A 0.007330
## 5  rs538215620    chr1       669427       669427      C      G 1.000000
## 6  rs1166003411    chr1       796384       796384      G      A 0.001350
```

The following excercises ask questions about the GWAS results. Use some R code and the genome browser to help you answer the questions.

Exercise 1:

Make a Manhatten plot showing the results from this GWAS. Note that cumulative genome position information is provided, but the column name has changed to ‘genome.pos.hg38’ to reflect the fact that the positions are on the hg38, rather than hg19, genome assembly. Also note that the chromosome names are like ‘chr2’, not ‘2’ as with the RA file.

ANSWER:

```
# Load libraries
library(ggplot2)
library(dplyr)

##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
## 
##     filter, lag

## The following objects are masked from 'package:base':
## 
##     intersect, setdiff, setequal, union
```

```

library(data.table); sumstats <- fread("t1d_gwas_hg38_pruned2_w_genome_pos.tsv.gz")

##
## Attaching package: 'data.table'

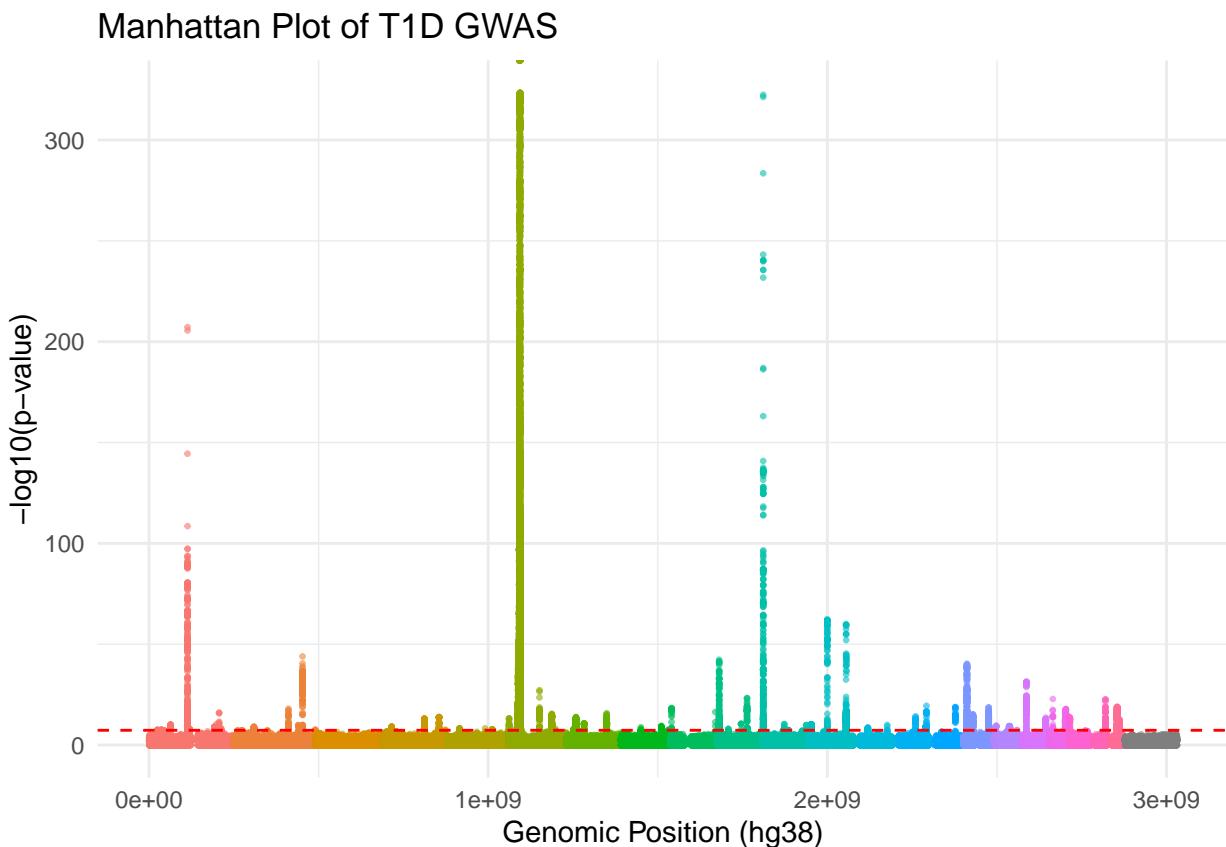
## The following objects are masked from 'package:dplyr':
##   between, first, last

t1d.gwas <- read.table("t1d_gwas_hg38_pruned2_w_genome_pos.tsv.gz", header = TRUE)

t1d.gwas$chrom <- factor(t1d.gwas$chrom, levels = paste0("chr", 1:22))

# Manhattan plot
ggplot(t1d.gwas, aes(x = genome.pos.hg38, y = -log10(p.val), color = chrom)) +
  geom_point(alpha = 0.6, size = 0.5) +
  scale_y_continuous(name = "-log10(p-value)") +
  labs(title = "Manhattan Plot of T1D GWAS", x = "Genomic Position (hg38)") +
  theme_minimal() +
  theme(legend.position = "none") +
  geom_hline(yintercept = -log10(5e-8), color = "red", linetype = "dashed")

```



Exercise 2

The MHC/HLA region on chromosome 6 is again the most significant region of association. Where does most of MHC signal appear to come from? MHC Class 1 or MHC Class 2?

```

library(ggplot2)

table(t1d.gwas$chrom)

## 
##   chr1    chr2    chr3    chr4    chr5    chr6    chr7    chr8    chr9    chr10   chr11
## 46210  47956  36475  33797  35582  101156  29912  29207  21558  29286  28908
##   chr12   chr13   chr14   chr15   chr16   chr17   chr18   chr19   chr20   chr21   chr22
## 28076  17084  17586  15563  20244  21522  14363  13453  12289  6892   8918

head(t1d.gwas[, c("chrom", "chrom.pos.hg38", "genome.pos.hg38")])

##   chrom chrom.pos.hg38 genome.pos.hg38
## 1   chr1        61578        61578
## 2   chr1       135215       135215
## 3   chr1       138495       138495
## 4   chr1       597067       597067
## 5   chr1       669427       669427
## 6   chr1       796384       796384

unique(sumstats$chrom)

## [1] "chr1"  "chr2"  "chr3"  "chr4"  "chr5"  "chr6"  "chr7"  "chr8"  "chr9"
## [10] "chr10" "chr11" "chr12" "chr13" "chr14" "chr15" "chr16" "chr17" "chr18"
## [19] "chr19" "chr20" "chr21" "chr22" "chrX"

sumstats <- fread("t1d_gwas_hg38_pruned2_w_genome_pos.tsv.gz")

sumstats[, chrom := as.numeric(gsub("chr", "", chrom))]

sumstats[, p.val := as.numeric(p.val)]

mhc_region <- sumstats[chrom == 6 & chrom.pos.hg38 >= 25000000 & chrom.pos.hg38 <= 35000000]

mhc_region <- mhc_region[!is.na(p.val)]

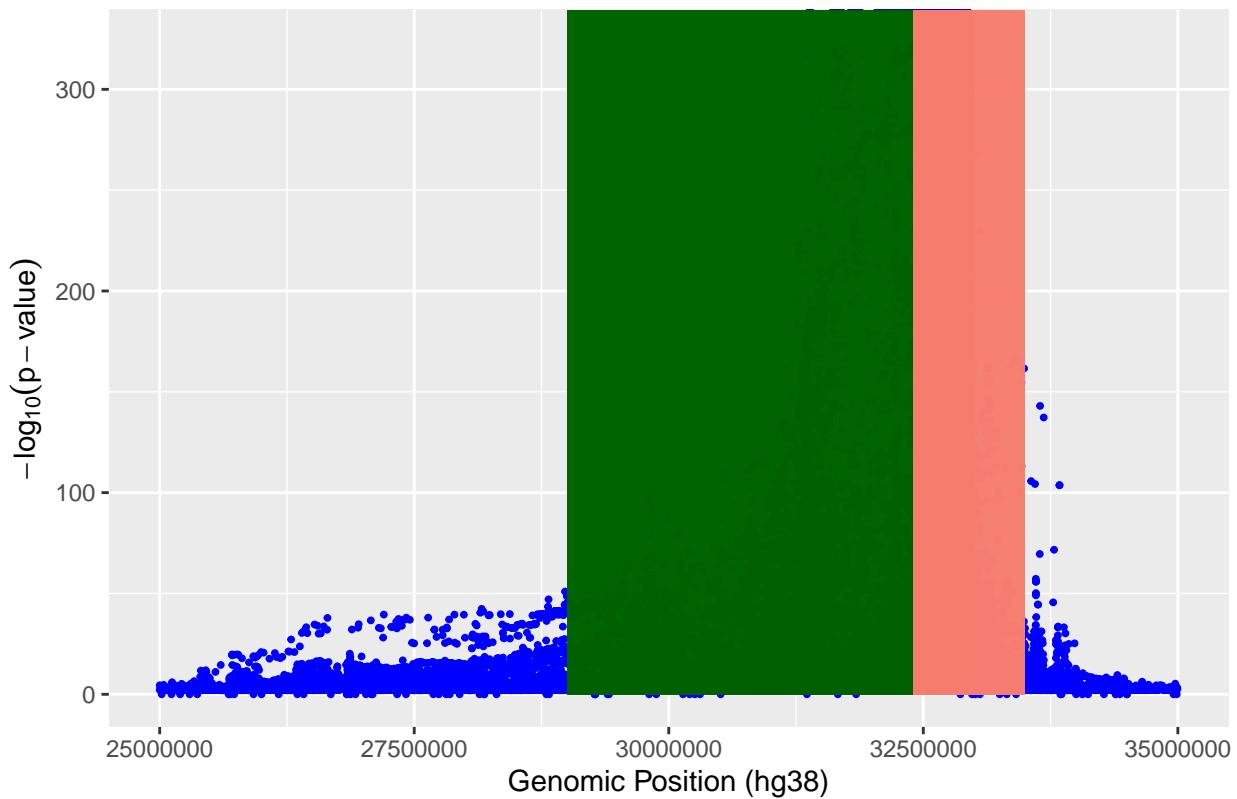
mhc_region[, logp := -log10(p.val)]

ggplot(mhc_region, aes(x = chrom.pos.hg38, y = logp)) +
  geom_point(size = 0.7, color = "blue") +
  # Class I region
  geom_rect(aes(xmin = 29000000, xmax = 33000000, ymin = 0, ymax = Inf),
            fill = "darkgreen", alpha = 0.2) + # Class I region
  # Add Class II region
  geom_rect(aes(xmin = 32400000, xmax = 33500000, ymin = 0, ymax = Inf),
            fill = "salmon", alpha = 0.2) + # Class II region

  labs(title = "MHC Region (Chr6)",
        x = "Genomic Position (hg38)",
        y = expression(-log[10](p-value)))

```

MHC Region (Chr6)



```
theme_minimal()
```

```
## List of 136
## $ line
##   ..$ colour      : chr "black"
##   ..$ linewidth   : num 0.5
##   ..$ linetype    : num 1
##   ..$ lineend     : chr "butt"
##   ..$ arrow       : logi FALSE
##   ..$ inherit.blank: logi TRUE
##   ..- attr(*, "class")= chr [1:2] "element_line" "element"
## $ rect
##   ..$ fill        : chr "white"
##   ..$ colour      : chr "black"
##   ..$ linewidth   : num 0.5
##   ..$ linetype    : num 1
##   ..$ inherit.blank: logi TRUE
##   ..- attr(*, "class")= chr [1:2] "element_rect" "element"
## $ text
##   ..$ family      : chr ""
##   ..$ face        : chr "plain"
##   ..$ colour      : chr "black"
##   ..$ size         : num 11
##   ..$ hjust        : num 0.5
##   ..$ vjust        : num 0.5
##   ..$ angle        : num 0
```

```

##  ..$ lineheight   : num 0.9
##  ..$ margin       : 'margin' num [1:4] Opoints Opoints Opoints Opoints
##  ... - attr(*, "unit")= int 8
##  ..$ debug        : logi FALSE
##  ..$ inherit.blank: logi TRUE
##  ... - attr(*, "class")= chr [1:2] "element_text" "element"
##  $ title           : NULL
##  $ aspect.ratio    : NULL
##  $ axis.title      : NULL
##  $ axis.title.x     :List of 11
##  ..$ family        : NULL
##  ..$ face          : NULL
##  ..$ colour         : NULL
##  ..$ size           : NULL
##  ..$ hjust          : NULL
##  ..$ vjust          : num 1
##  ..$ angle          : NULL
##  ..$ lineheight     : NULL
##  ..$ margin          : 'margin' num [1:4] 2.75points Opoints Opoints Opoints
##  ... - attr(*, "unit")= int 8
##  ..$ debug          : NULL
##  ..$ inherit.blank: logi TRUE
##  ... - attr(*, "class")= chr [1:2] "element_text" "element"
##  $ axis.title.x.top   :List of 11
##  ..$ family        : NULL
##  ..$ face          : NULL
##  ..$ colour         : NULL
##  ..$ size           : NULL
##  ..$ hjust          : NULL
##  ..$ vjust          : num 0
##  ..$ angle          : NULL
##  ..$ lineheight     : NULL
##  ..$ margin          : 'margin' num [1:4] Opoints Opoints 2.75points Opoints
##  ... - attr(*, "unit")= int 8
##  ..$ debug          : NULL
##  ..$ inherit.blank: logi TRUE
##  ... - attr(*, "class")= chr [1:2] "element_text" "element"
##  $ axis.title.x.bottom : NULL
##  $ axis.title.y      :List of 11
##  ..$ family        : NULL
##  ..$ face          : NULL
##  ..$ colour         : NULL
##  ..$ size           : NULL
##  ..$ hjust          : NULL
##  ..$ vjust          : num 1
##  ..$ angle          : num 90
##  ..$ lineheight     : NULL
##  ..$ margin          : 'margin' num [1:4] Opoints 2.75points Opoints Opoints
##  ... - attr(*, "unit")= int 8
##  ..$ debug          : NULL
##  ..$ inherit.blank: logi TRUE
##  ... - attr(*, "class")= chr [1:2] "element_text" "element"
##  $ axis.title.y.left  : NULL
##  $ axis.title.y.right :List of 11

```

```

## ..$ family      : NULL
## ..$ face        : NULL
## ..$ colour      : NULL
## ..$ size         : NULL
## ..$ hjust        : NULL
## ..$ vjust        : num 1
## ..$ angle        : num -90
## ..$ lineheight   : NULL
## ..$ margin       : 'margin' num [1:4] 0points 0points 0points 2.75points
## ...- attr(*, "unit")= int 8
## ..$ debug        : NULL
## ..$ inherit.blank: logi TRUE
## ...- attr(*, "class")= chr [1:2] "element_text" "element"
## $ axis.text          :List of 11
## ..$ family      : NULL
## ..$ face        : NULL
## ..$ colour      : chr "grey30"
## ..$ size         : 'rel' num 0.8
## ..$ hjust        : NULL
## ..$ vjust        : NULL
## ..$ angle        : NULL
## ..$ lineheight   : NULL
## ..$ margin       : NULL
## ..$ debug        : NULL
## ..$ inherit.blank: logi TRUE
## ...- attr(*, "class")= chr [1:2] "element_text" "element"
## $ axis.text.x        :List of 11
## ..$ family      : NULL
## ..$ face        : NULL
## ..$ colour      : NULL
## ..$ size         : NULL
## ..$ hjust        : NULL
## ..$ vjust        : num 1
## ..$ angle        : NULL
## ..$ lineheight   : NULL
## ..$ margin       : 'margin' num [1:4] 2.2points 0points 0points 0points
## ...- attr(*, "unit")= int 8
## ..$ debug        : NULL
## ..$ inherit.blank: logi TRUE
## ...- attr(*, "class")= chr [1:2] "element_text" "element"
## $ axis.text.x.top    :List of 11
## ..$ family      : NULL
## ..$ face        : NULL
## ..$ colour      : NULL
## ..$ size         : NULL
## ..$ hjust        : NULL
## ..$ vjust        : num 0
## ..$ angle        : NULL
## ..$ lineheight   : NULL
## ..$ margin       : 'margin' num [1:4] 0points 0points 2.2points 0points
## ...- attr(*, "unit")= int 8
## ..$ debug        : NULL
## ..$ inherit.blank: logi TRUE
## ...- attr(*, "class")= chr [1:2] "element_text" "element"

```

```

## $ axis.text.x.bottom : NULL
## $ axis.text.y :List of 11
## ..$ family : NULL
## ..$ face : NULL
## ..$ colour : NULL
## ..$ size : NULL
## ..$ hjust : num 1
## ..$ vjust : NULL
## ..$ angle : NULL
## ..$ lineheight : NULL
## ..$ margin : 'margin' num [1:4] 0points 2.2points 0points 0points
## ... - attr(*, "unit")= int 8
## ..$ debug : NULL
## ..$ inherit.blank: logi TRUE
## ..- attr(*, "class")= chr [1:2] "element_text" "element"
## $ axis.text.y.left : NULL
## $ axis.text.y.right :List of 11
## ..$ family : NULL
## ..$ face : NULL
## ..$ colour : NULL
## ..$ size : NULL
## ..$ hjust : num 0
## ..$ vjust : NULL
## ..$ angle : NULL
## ..$ lineheight : NULL
## ..$ margin : 'margin' num [1:4] 0points 0points 0points 2.2points
## ... - attr(*, "unit")= int 8
## ..$ debug : NULL
## ..$ inherit.blank: logi TRUE
## ..- attr(*, "class")= chr [1:2] "element_text" "element"
## $ axis.text.theta : NULL
## $ axis.text.r :List of 11
## ..$ family : NULL
## ..$ face : NULL
## ..$ colour : NULL
## ..$ size : NULL
## ..$ hjust : num 0.5
## ..$ vjust : NULL
## ..$ angle : NULL
## ..$ lineheight : NULL
## ..$ margin : 'margin' num [1:4] 0points 2.2points 0points 2.2points
## ... - attr(*, "unit")= int 8
## ..$ debug : NULL
## ..$ inherit.blank: logi TRUE
## ..- attr(*, "class")= chr [1:2] "element_text" "element"
## $ axis.ticks : list()
## ..- attr(*, "class")= chr [1:2] "element_blank" "element"
## $ axis.ticks.x : NULL
## $ axis.ticks.x.top : NULL
## $ axis.ticks.x.bottom : NULL
## $ axis.ticks.y : NULL
## $ axis.ticks.y.left : NULL
## $ axis.ticks.y.right : NULL
## $ axis.ticks.theta : NULL

```

```

## $ axis.ticks.r : NULL
## $ axis.minor.ticks.x.top : NULL
## $ axis.minor.ticks.x.bottom : NULL
## $ axis.minor.ticks.y.left : NULL
## $ axis.minor.ticks.y.right : NULL
## $ axis.minor.ticks.theta : NULL
## $ axis.minor.ticks.r : NULL
## $ axis.ticks.length : 'simpleUnit' num 2.75points
## ..- attr(*, "unit")= int 8
## $ axis.ticks.length.x : NULL
## $ axis.ticks.length.x.top : NULL
## $ axis.ticks.length.x.bottom : NULL
## $ axis.ticks.length.y : NULL
## $ axis.ticks.length.y.left : NULL
## $ axis.ticks.length.y.right : NULL
## $ axis.ticks.length.theta : NULL
## $ axis.ticks.length.r : NULL
## $ axis.minor.ticks.length : 'rel' num 0.75
## $ axis.minor.ticks.length.x : NULL
## $ axis.minor.ticks.length.x.top : NULL
## $ axis.minor.ticks.length.x.bottom: NULL
## $ axis.minor.ticks.length.y : NULL
## $ axis.minor.ticks.length.y.left : NULL
## $ axis.minor.ticks.length.y.right : NULL
## $ axis.minor.ticks.length.theta : NULL
## $ axis.minor.ticks.length.r : NULL
## $ axis.line : list()
## ..- attr(*, "class")= chr [1:2] "element_blank" "element"
## $ axis.line.x : NULL
## $ axis.line.x.top : NULL
## $ axis.line.x.bottom : NULL
## $ axis.line.y : NULL
## $ axis.line.y.left : NULL
## $ axis.line.y.right : NULL
## $ axis.line.theta : NULL
## $ axis.line.r : NULL
## $ legend.background : list()
## ..- attr(*, "class")= chr [1:2] "element_blank" "element"
## $ legend.margin : 'margin' num [1:4] 5.5points 5.5points 5.5points 5.5points
## ..- attr(*, "unit")= int 8
## $ legend.spacing : 'simpleUnit' num 11points
## ..- attr(*, "unit")= int 8
## $ legend.spacing.x : NULL
## $ legend.spacing.y : NULL
## $ legend.key : list()
## ..- attr(*, "class")= chr [1:2] "element_blank" "element"
## $ legend.key.size : 'simpleUnit' num 1.2lines
## ..- attr(*, "unit")= int 3
## $ legend.key.height : NULL
## $ legend.key.width : NULL
## $ legend.key.spacing : 'simpleUnit' num 5.5points
## ..- attr(*, "unit")= int 8
## $ legend.key.spacing.x : NULL
## $ legend.key.spacing.y : NULL

```

```

## $ legend.frame : NULL
## $ legend.ticks : NULL
## $ legend.ticks.length : 'rel' num 0.2
## $ legend.axis.line : NULL
## $ legend.text :List of 11
##   ..$ family : NULL
##   ..$ face : NULL
##   ..$ colour : NULL
##   ..$ size : 'rel' num 0.8
##   ..$ hjust : NULL
##   ..$ vjust : NULL
##   ..$ angle : NULL
##   ..$ lineheight : NULL
##   ..$ margin : NULL
##   ..$ debug : NULL
##   ..$ inherit.blank: logi TRUE
##   ..- attr(*, "class")= chr [1:2] "element_text" "element"
## $ legend.text.position : NULL
## $ legend.title :List of 11
##   ..$ family : NULL
##   ..$ face : NULL
##   ..$ colour : NULL
##   ..$ size : NULL
##   ..$ hjust : num 0
##   ..$ vjust : NULL
##   ..$ angle : NULL
##   ..$ lineheight : NULL
##   ..$ margin : NULL
##   ..$ debug : NULL
##   ..$ inherit.blank: logi TRUE
##   ..- attr(*, "class")= chr [1:2] "element_text" "element"
## $ legend.title.position : NULL
## $ legend.position : chr "right"
## $ legend.position.inside : NULL
## $ legend.direction : NULL
## $ legend.byrow : NULL
## $ legend.justification : chr "center"
## $ legend.justification.top : NULL
## $ legend.justification.bottom : NULL
## $ legend.justification.left : NULL
## $ legend.justification.right : NULL
## $ legend.justification.inside : NULL
## $ legend.location : NULL
## $ legend.box : NULL
## $ legend.box.just : NULL
## $ legend.box.margin : 'margin' num [1:4] 0cm 0cm 0cm 0cm
##   ..- attr(*, "unit")= int 1
## $ legend.box.background : list()
##   ..- attr(*, "class")= chr [1:2] "element_blank" "element"
## $ legend.box.spacing : 'simpleUnit' num 11points
##   ..- attr(*, "unit")= int 8
##   [list output truncated]
## - attr(*, "class")= chr [1:2] "theme" "gg"
## - attr(*, "complete")= logi TRUE

```

```
## - attr(*, "validate")= logi TRUE
```

Most of the MHC signal appears to come from MHC Class II genes.

Exercise 3

What are the next two top hits after MHC? Where are they located? What are some nearby genes?

```
#Filtering out the MHC region
not_mhc <- t1d.gwas.tab %>%
  filter(!(chrom == "chr6" & chrom.pos.hg38 >= 29900000 & chrom.pos.hg38 <= 33400000))

#Get top SNPs
top_hits <- not_mhc %>%
  arrange(p.val) %>%
  slice(1:2)
top_hits
```

SNP.ID	chrom	chrom.pos.hg38	genome.pos.hg38	allele1	allele2	p.val
rs689	chr11	2160994	1810842045	T	A	3.952525e-323
rs3842753	chr11	2159830	1810840881	G	T	4.545404e-322

Using the UCSC Genome browser, the next two top hits are located between 2160000 and 2161000.