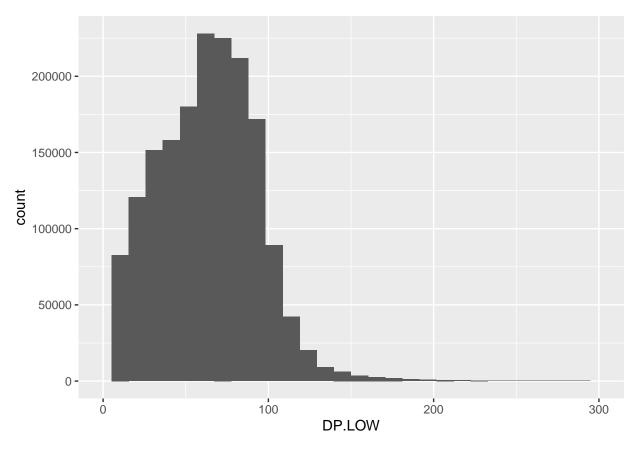
## QTLseqrRiceColdTolerance

## Michael Hall

## 2022-06-08

```
# install devtools and all other dependent packages first
#utils::install.packages("devtools")
#utils::install.packages("tidyr")
#utils::install.packages("ggplot2")
#utils::install.packages("dplyr")
#utils::install.packages("data.table")
#utils::install.packages("vcfR")
# use devtools to install QTLseqr from my "Forked" Repository at PBGL
devtools::install_github("PBGLMichaelHall/QTLseqr")
Skipping install of 'QTLseqr' from a github remote, the SHA1 (9647160e) has not changed since last inst
 Use `force = TRUE` to force installation
# Load them into your session so they are available for use
base::library("QTLseqr")
base::library("data.table")
base::library("dplyr")
Attaching package: 'dplyr'
The following objects are masked from 'package:data.table':
    between, first, last
The following objects are masked from 'package:stats':
   filter, lag
The following objects are masked from 'package:base':
    intersect, setdiff, setequal, union
base::library("tidyr")
base::library("vcfR")
                    vcfR
                                      ****
  This is vcfR 1.12.0
     browseVignettes('vcfR') # Documentation
     citation('vcfR') # Citation
                                      ****
base::library("ggplot2")
```

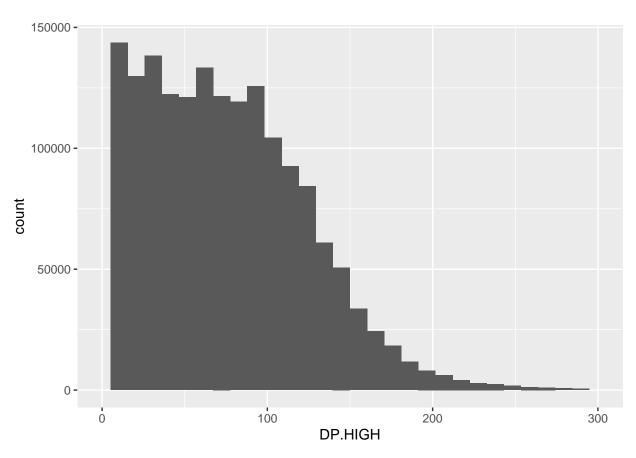
```
QTLseqr::importFromVCF(file = "wGQ-Filt-freebayes~bwa~IRGSP-1.0~both-segregant_bulks~filtered-default.v
Scanning file to determine attributes.
File attributes:
    meta lines: 126
    header_line: 127
    variant count: 1714745
    column count: 11
Meta line 126 read in.
All meta lines processed.
gt matrix initialized.
Character matrix gt created.
    Character matrix gt rows: 1714745
    Character matrix gt cols: 11
    skip: 0
   nrows: 1714745
    row_num: 0
Processed variant 1000Processed variant 2000Processed variant 3000Processed variant 4000Processed varia
All variants processed
Keeping SNPs that pass all filters Either PASS or No Filter
Extracting gt element AD
Extracting gt element DP
Extracting gt element GQ
Removing the following chromosomes:
HighBulk <- "ET-pool-385"
LowBulk <- "ES-pool-430"
file <- "Hall.csv"</pre>
#Choose which chromosomes/contigs will be included in the analysis,
chromList <- c("NC_029256.1","NC_029257.1","NC_029258.1","NC_029259.1","NC_029260.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","NC_029261.1","
df <-
    QTLseqr::importFromTable(
        file = file,
        highBulk = HighBulk,
        lowBulk = LowBulk,
        chromList = chromList,
        sep = ","
Removing the following chromosomes:
Renaming the following columns: AD_REF.ET-pool-385, AD_ALT.ET-pool-385
Renaming the following columns: AD_REF.ES-pool-430, AD_ALT.ES-pool-430
ggplot2::ggplot(data = df) + ggplot2::geom_histogram(ggplot2::aes(x = DP.LOW)) + ggplot2::xlim(0,300)
`stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
Warning: Removed 1329 rows containing non-finite values (stat_bin).
Warning: Removed 2 rows containing missing values (geom_bar).
```



 $ggplot2::ggplot(\frac{data}{data} = df) + ggplot2::geom\_histogram(ggplot2::aes(x = DP.HIGH)) + ggplot2::xlim(0,300)$ 

Warning: Removed 3790 rows containing non-finite values (stat\_bin). Removed 2 rows containing missing values (geom\_bar).

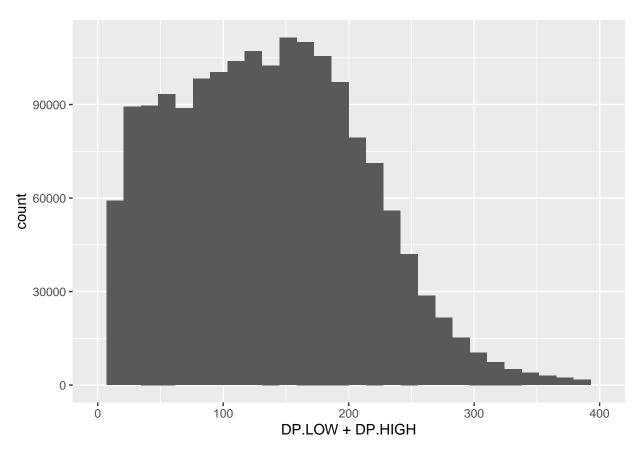
<sup>`</sup>stat\_bin()` using `bins = 30`. Pick better value with `binwidth`.



ggplot2::ggplot(data = df) + ggplot2::geom\_histogram(ggplot2::aes(x = DP.LOW + DP.HIGH)) + ggplot2::xling

Warning: Removed 9027 rows containing non-finite values (stat\_bin). Removed 2 rows containing missing values (geom\_bar).

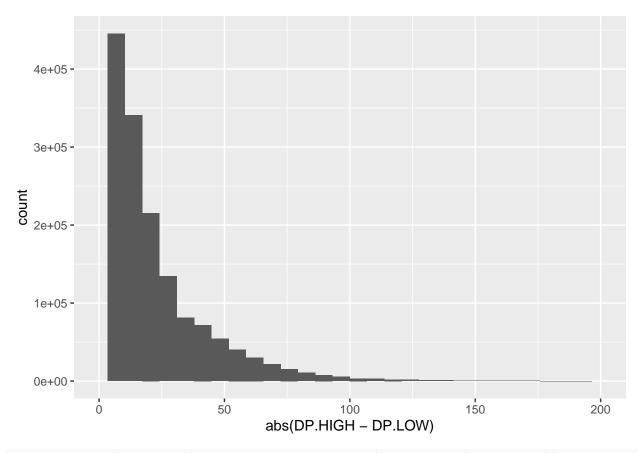
<sup>`</sup>stat\_bin()` using `bins = 30`. Pick better value with `binwidth`.



 $ggplot2::ggplot(\frac{data}{} = df) + ggplot2::geom\_histogram(ggplot2::aes(\frac{x}{} = abs(DP.HIGH - DP.LOW))) + ggplot(\frac{data}{} = df) + ggplot(\frac{data$ 

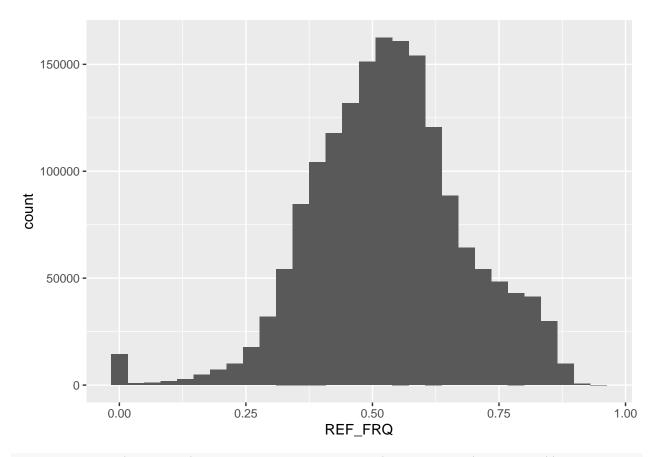
Warning: Removed 755 rows containing non-finite values (stat\_bin). Removed 2 rows containing missing values (geom\_bar).

<sup>`</sup>stat\_bin()` using `bins = 30`. Pick better value with `binwidth`.



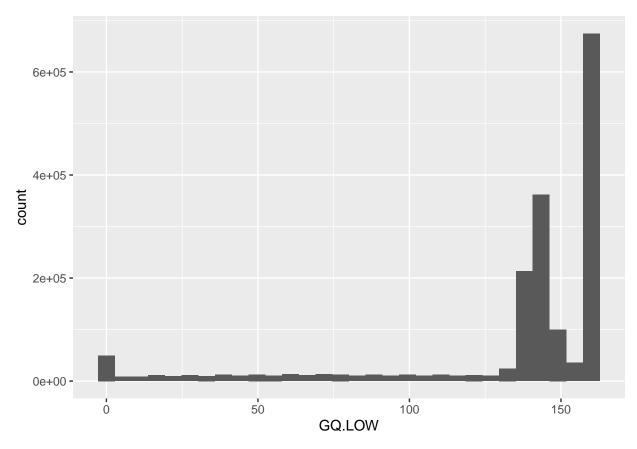
 $ggplot2::ggplot(\frac{data}{} = df) + ggplot2::geom_histogram(ggplot2::aes(\frac{x}{} = REF\_FRQ))$ 

<sup>`</sup>stat\_bin()` using `bins = 30`. Pick better value with `binwidth`.



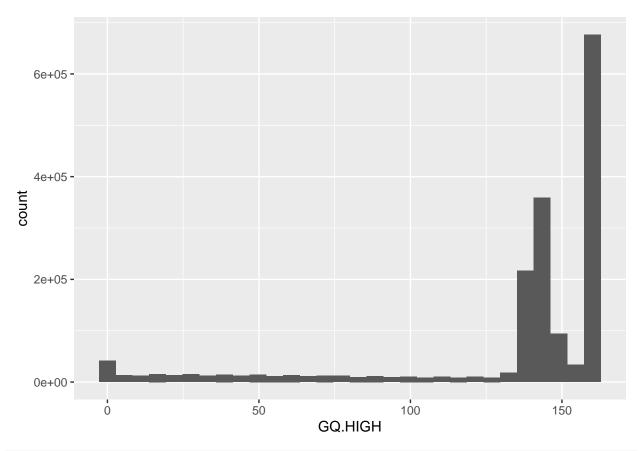
 $ggplot2::ggplot(\frac{data}{} = df) + ggplot2::geom\_histogram(ggplot2::aes(x = GQ.LOW))$ 

<sup>`</sup>stat\_bin()` using `bins = 30`. Pick better value with `binwidth`.



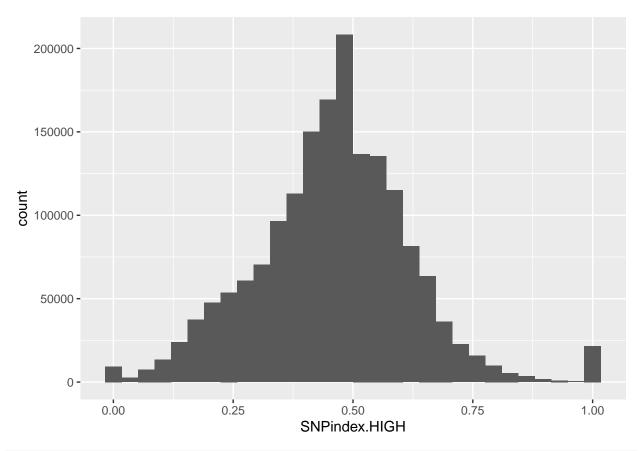
ggplot2::ggplot(data = df) + ggplot2::geom\_histogram(ggplot2::aes(x = GQ.HIGH))

<sup>`</sup>stat\_bin()` using `bins = 30`. Pick better value with `binwidth`.



ggplot2::ggplot(data = df) + ggplot2::geom\_histogram(ggplot2::aes(x = SNPindex.HIGH))

`stat\_bin()` using `bins = 30`. Pick better value with `binwidth`.



Filtering by reference allele frequency: 0.2 <= REF\_FRQ <= 0.8
...Filtered 112443 SNPs

Filtering by total sample read depth: Total DP >= 100
...Filtered 537579 SNPs

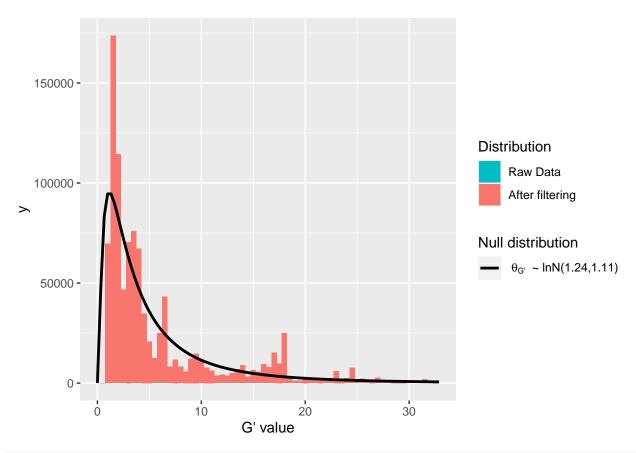
Filtering by total sample read depth: Total DP <= 400
...Filtered 6975 SNPs

Filtering by per sample read depth: DP >= 40
...Filtered 8015 SNPs

Filtering by Genotype Quality: GQ >= 99

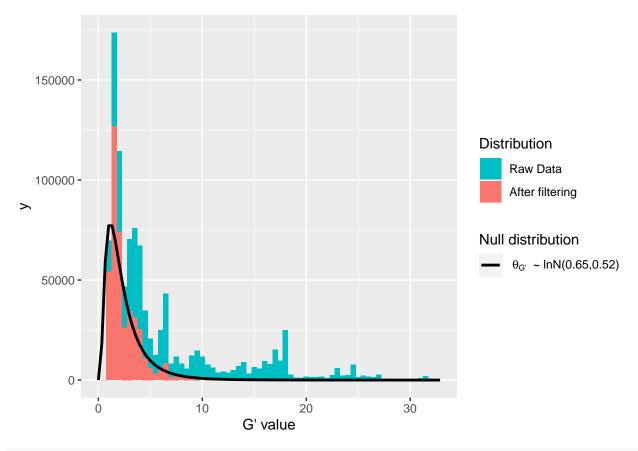
```
...Filtered 37905 SNPs
Filtering by difference between bulks <= 100
...Filtered 8746 SNPs
Original SNP number: 1714745, Filtered: 711663, Remaining: 1003082
#Run G' analysis
df_filt<-QTLseqr::runGprimeAnalysis(</pre>
 SNPset = df filt,
 windowSize = 1e6,
 outlierFilter = "deltaSNP",
filterThreshold = 0.1)
Counting SNPs in each window...
Calculating tricube smoothed delta SNP index...
Calculating G and G' statistics...
Using deltaSNP-index to filter outlier regions with a threshold of 0.1
Estimating the mode of a trimmed G prime set using the 'modeest' package...
Calculating p-values...
#Run QTLseq analysis
df_filt <- QTLseqr::runQTLseqAnalysis(</pre>
 SNPset = df_filt,
 windowSize = 1e6,
 popStruc = "F2",
 bulkSize = c(385, 430),
 replications = 10000,
 intervals = c(95, 99)
)
Counting SNPs in each window...
Calculating tricube smoothed delta SNP index...
Returning the following two sided confidence intervals: 95, 99
Variable 'depth' not defined, using min and max depth from data: 40-199
Assuming bulks selected from F2 population, with 385 and 430 individuals per bulk.
Simulating 10000 SNPs with reads at each depth: 40-199
Keeping SNPs with >= 0.3 SNP-index in both simulated bulks
Joining, by = "tricubeDP"
QTLseqr::plotGprimeDist(SNPset = df_filt, outlierFilter = "Hampel")
```

Warning: Removed 2 rows containing missing values (geom\_bar). Removed 2 rows containing missing values (geom\_bar).

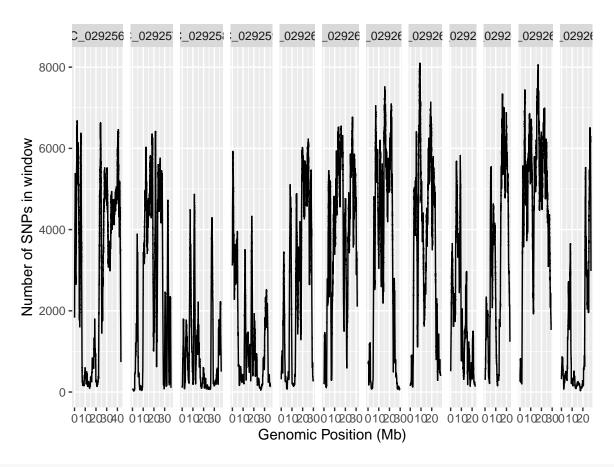


QTLseqr::plotGprimeDist(SNPset = df\_filt, outlierFilter = "deltaSNP",filterThreshold = 0.1)

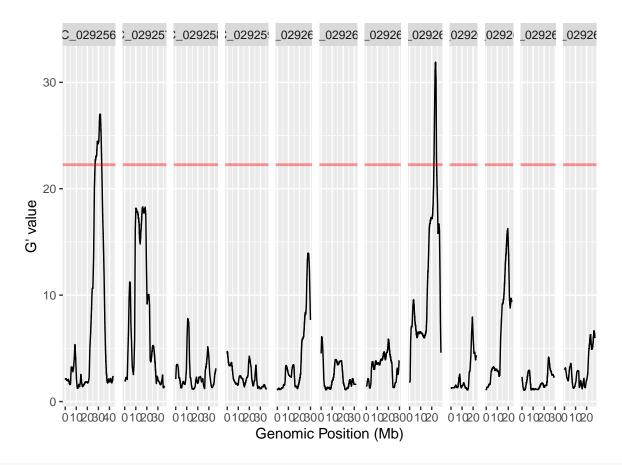
Warning: Removed 2 rows containing missing values (geom\_bar). Removed 2 rows containing missing values (geom\_bar).



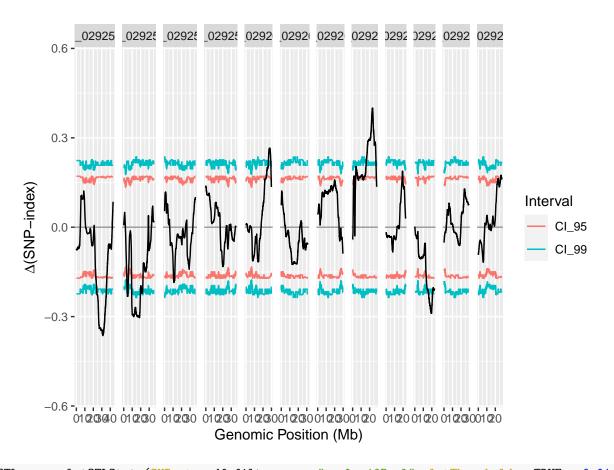
QTLseqr::plotQTLStats(SNPset = df\_filt, var = "nSNPs")



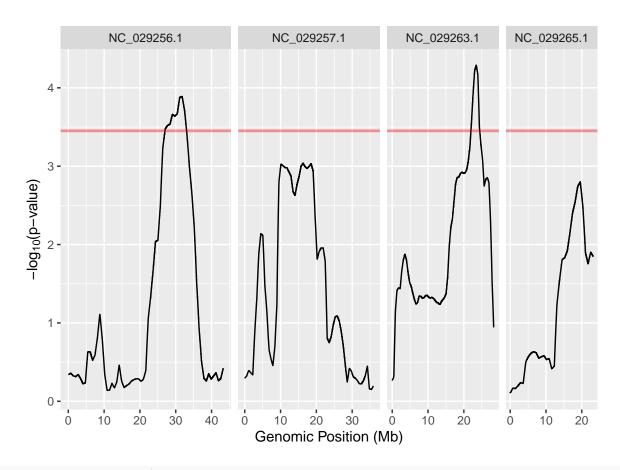
QTLseqr::plotQTLStats(SNPset = df\_filt, var = "Gprime", plotThreshold = TRUE, q = 0.01)



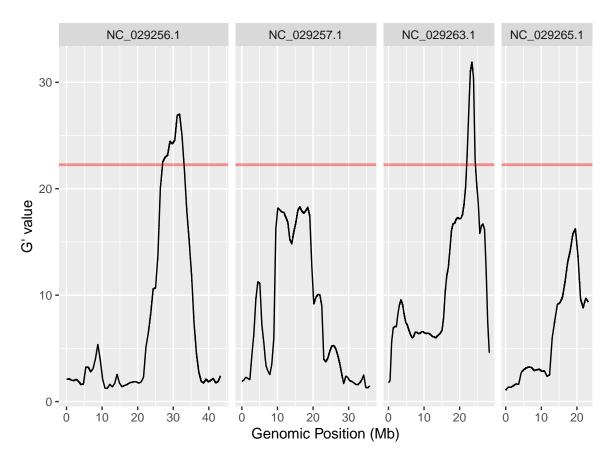
QTLseqr::plotQTLStats(SNPset = df\_filt, var = "deltaSNP", plotIntervals = TRUE)



QTLseqr::plotQTLStats(SNPset = df\_filt, var = "negLog10Pval",plotThreshold = TRUE,q=0.01,subset = c("NC



QTLseqr::plotQTLStats(SNPset = df\_filt, var = "Gprime",plotThreshold = TRUE,q=0.01,subset = c("NC\_02925



```
QTLseqr::getQTLTable(SNPset = df_filt, alpha = 0.01, export = TRUE, fileName = "my_BSA_QTL.csv")
           CHROM qtl
                                   end length nSNPs avgSNPs_Mb peakDeltaSNP
                        start
## 1 NC_029256.1
                   1 26986793 33033424 6046631 26816
                                                                  -0.3636383
                                                           4435
## 2 NC_029263.1
                   2 21980124 24413447 2433323 8704
                                                           3577
                                                                   0.3999064
     posPeakDeltaSNP avgDeltaSNP maxGprime posMaxGprime meanGprime sdGprime
            31110319 -0.3456919 27.01544
                                               31780847
## 1
                                                          24.38459 1.321401
## 2
            23458666
                       0.3756486 31.87973
                                               23458666
                                                          28.26922 3.219559
        AUCaT
                  meanPval
                              meanQval
## 1 13606901 0.0002315438 0.009554921
## 2 14451588 0.0001277682 0.009555889
```