Sorghum semi-dwarf - QTLseqr Quick-Guide

Norman Warthmann, PBG Laboratory

6/15/2022

Conduct a QTLseqr Analysis

```r
#load dependencies
library("devtools")
library("data.table")
library("dplyr")
library("tidyr")
library("vcfR")
library("ggplot2")
library("QTLseqr")
```

Step-by-Step Guide

Define the input

```r
file <- "freebayes_D2.filtered.vcf"
HighBulk <- "D2_F2_tt" # 45 plants
LowBulk <- "D2_F2_TT" # 38 plants
Chroms <- c("Chr01","Chr02","Chr03","Chr04","Chr05",
            "Chr06","Chr07","Chr08","Chr09","Chr10"
            )
```

Load the data from the VCF file

```r
df <-
  importFromVCF(
    file = file,
    highBulk = HighBulk,
    lowBulk = LowBulk,
    chromList = Chroms
    )
```

Filter the Variants (user-defined criteria)

```
df_filt <-
  filterSNPs(
    SNPset = df,
    #refAlleleFreq = 0.20,
    minTotalDepth = 100,
    maxTotalDepth = 400,
    minSampleDepth = 30,
    #depthDifference = 100,
    minGQ = 99,
    verbose = TRUE
    )
```

```
## Filtering by total sample read depth: Total DP >= 100

## ...Filtered 1032 SNPs

## Filtering by total sample read depth: Total DP <= 400

## ...Filtered 208 SNPs

## Filtering by per sample read depth: DP >= 30

## ...Filtered 0 SNPs

## Filtering by Genotype Quality: GQ >= 99

## ...Filtered 67 SNPs

## Original SNP number: 7763, Filtered: 1307, Remaining: 6456
```

Run the G' (G prime) Analysis

```
df_filt <-
  runGprimeAnalysis(
    SNPset = df_filt,
    windowSize = 5e6,
    outlierFilter = "deltaSNP",
    filterThreshold = 0.1
    )
```

Run the Delta SNP Analysis

```
df_filt <-
  runQTLseqAnalysis(
    SNPset = df_filt,
    windowSize = 5e6,
    popStruc = "F2",
    bulkSize = c(45, 38),
    replications = 10000,
    intervals = c(95, 99)
    )
```

Plot SNP Density

```
#Plot SNP density
plotQTLStats(SNPset = df_filt, var = "nSNPs", plotIntervals = TRUE)
```
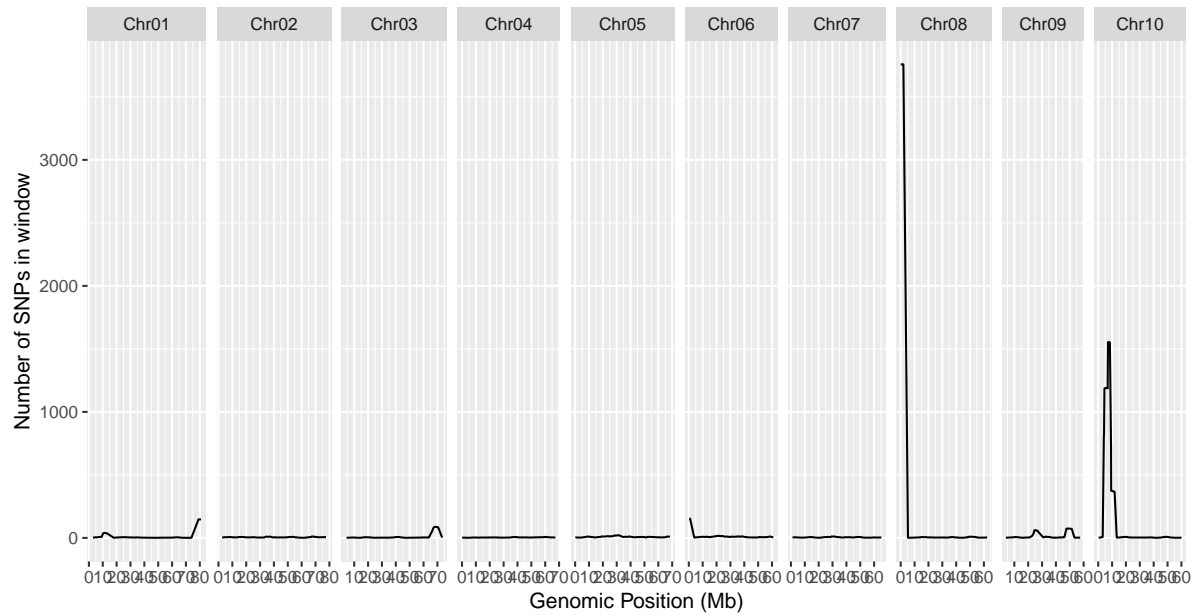

Figure 1: SNP Density

Plot G'

```
plotQTLStats( SNPset = df_filt,
              var = "Gprime",
              plotThreshold = TRUE,
              q = 0.01
              )
```
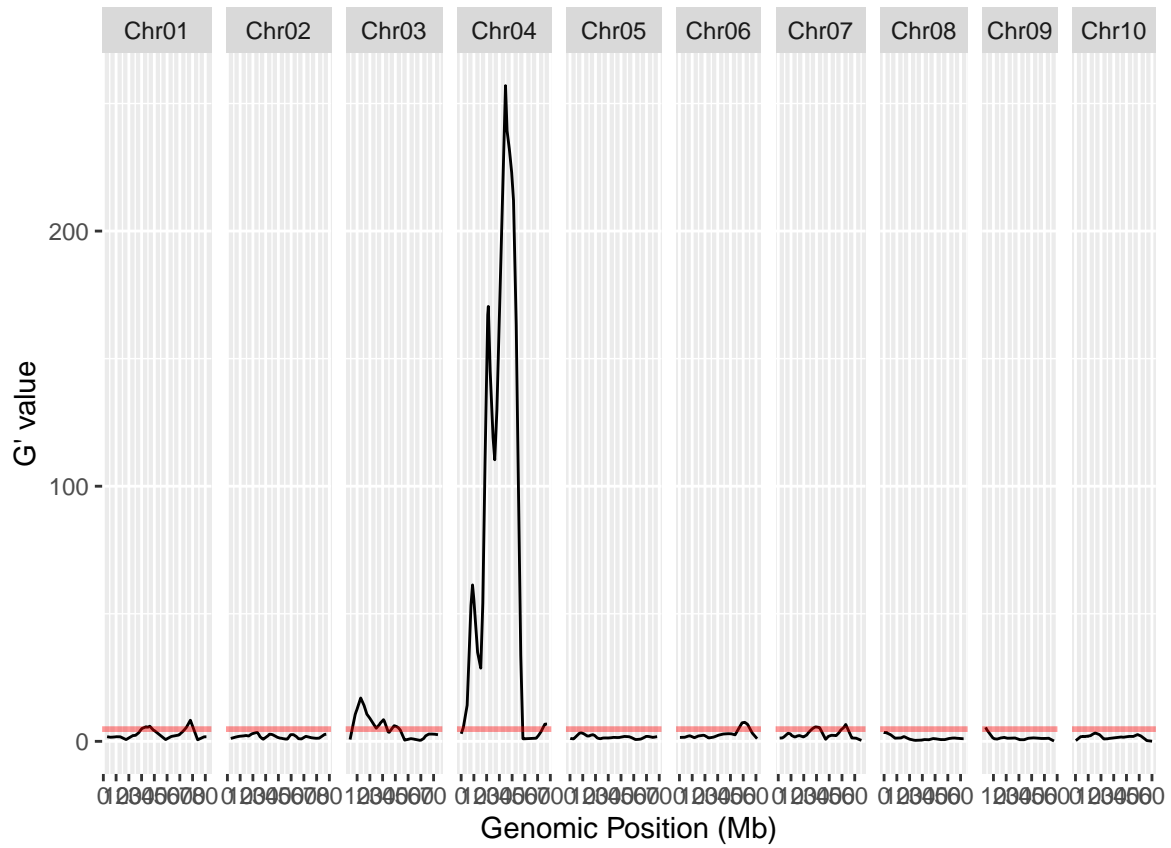

Figure 2: G' Values

Plot only Chromosome 4

```
plotQTLStats( SNPset = df_filt,
              var = "Gprime",
              plotThreshold = TRUE,
              q = 0.01,
              subset=c("Chr04")
              )
```
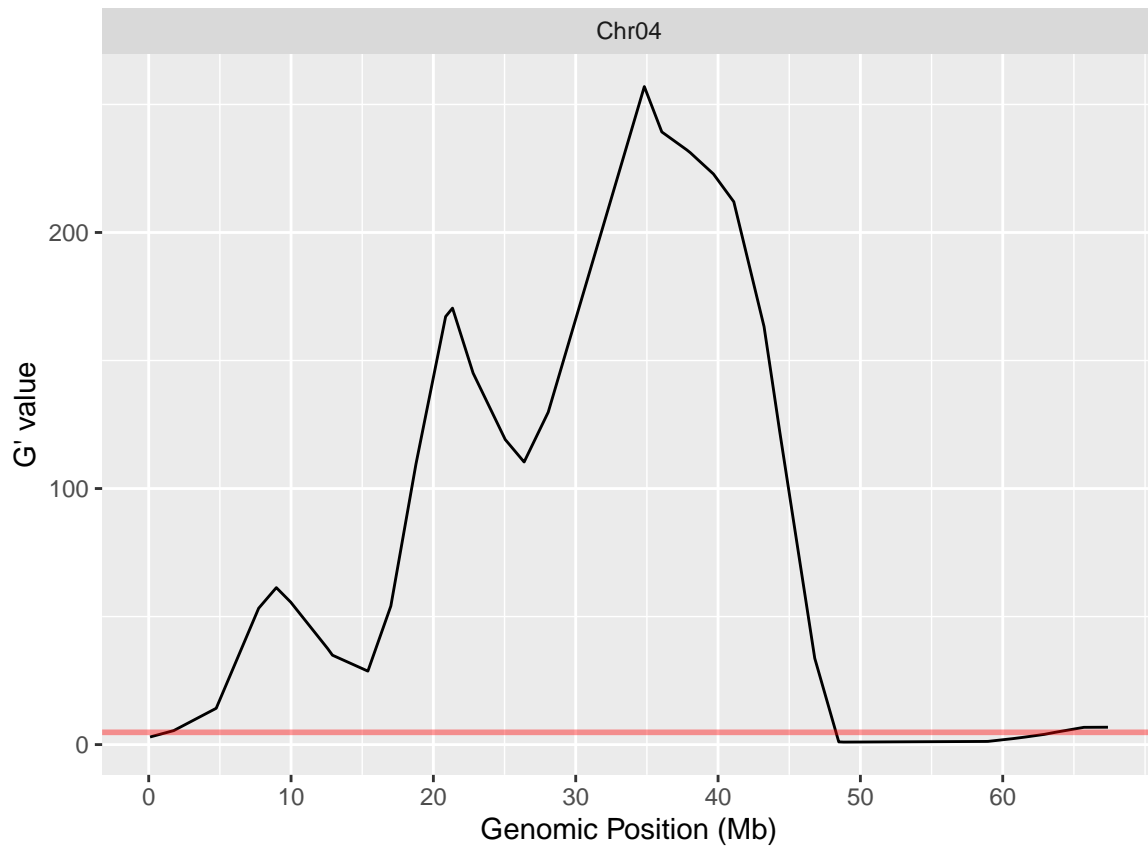

Figure 3: G' Values for Chromosome 4

Plot Delta SNP

```
plotQTLStats(SNPset = df_filt,
             var = "deltaSNP",
             plotIntervals = TRUE
             )
```

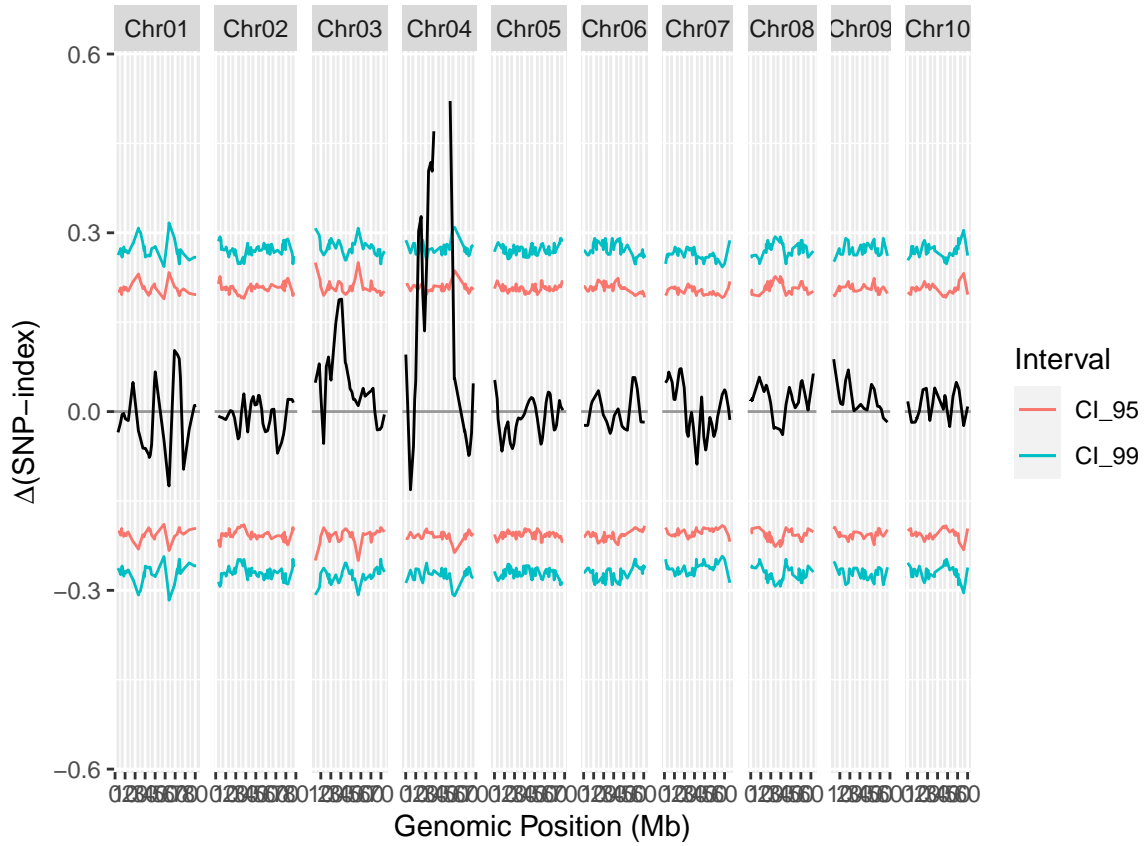

Figure 4: Delta SNP

Plot only Chromosome 4

```
plotQTLStats(SNPset = df_filt,
             var = "deltaSNP",
             plotIntervals = TRUE,
             subset=c("Chr04")
             )
```
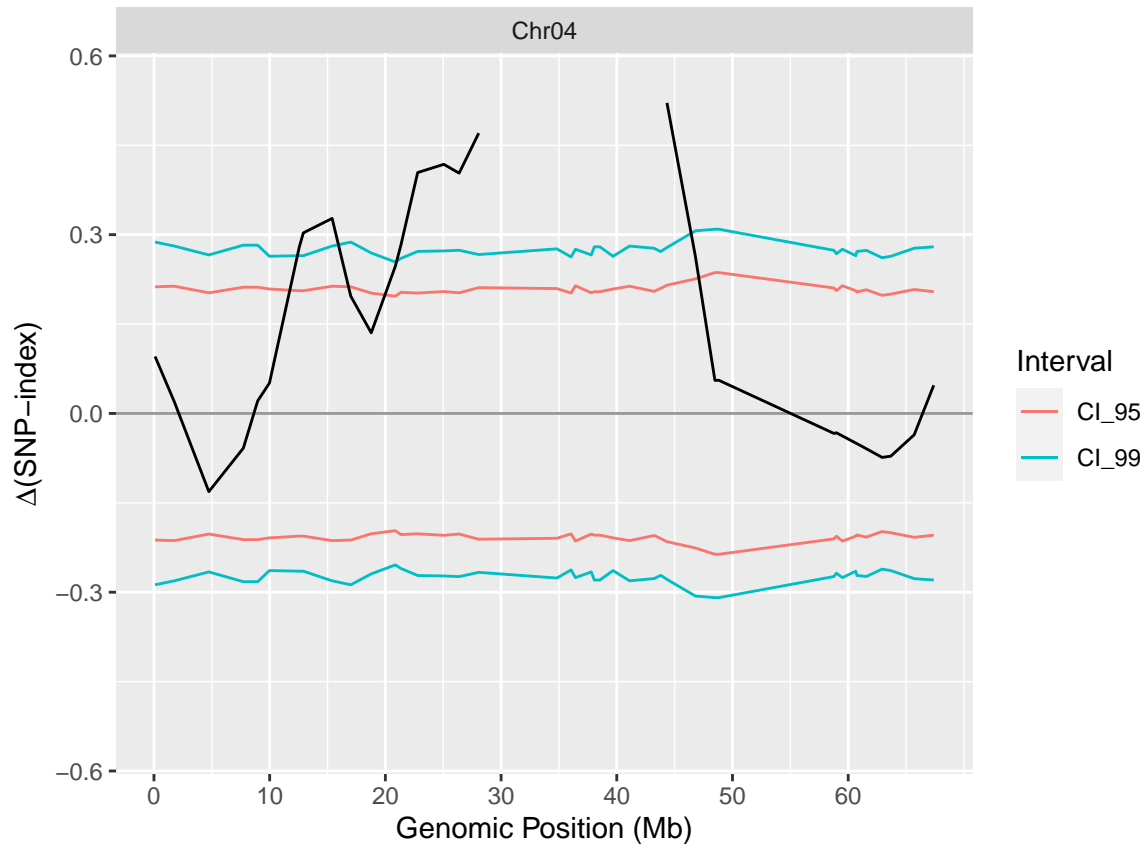

Figure 5: Delta SNP Chromsome 4