

Exploring chromosome 12

Going to look closer at chromosome 12

```
library("data.table")
source("myfunctions.R")

genepos <- fread("results/gene_pos.gz")
if (!exists("find.effects_TF")){
  find.effects_TF <- fread("results/findeffects_TF_newparams.gz")
}

# Add chomosome and start position to each gene
causal.pos.A <- merge(find.effects_TF, genepos, by.x="geneA", by.y="gene", all.x=T)
colnames(causal.pos.A) <- c("geneA", "geneB", "eqtl.A", "eqtl.B", "A->B", "B->A", "strand.A", "start.A")
causal.pos.B <- merge(causal.pos.A, genepos, by.x="geneB", by.y="gene", all.x=T)

# Keep olnly columns with genes, start positions and chromosomes
causal.pos.B.2 <- unique(causal.pos.B[,.(geneA, geneB, start.A, chr.A, chr.start, chr.id)])
colnames(causal.pos.B.2) <- c("geneA", "geneB", "start.A", "chr.A", "start.B", "chr.B")

## transform chromosome ids into numbers
# remove "chr" part of the chromosome name
causal.pos.B.2$chr.A <- gsub('chr', '', causal.pos.B.2$chr.A)
causal.pos.B.2$chr.B <- gsub('chr', '', causal.pos.B.2$chr.B)
colnames(causal.pos.B.2) <- c("geneA", "geneB", "start.A", "chr.A", "start.B", "chr.B")

# convert roman chromosome numbers to numbers
causal.pos.B.2$chr.A <- as.numeric(as.roman(causal.pos.B.2$chr.A))
causal.pos.B.2$chr.B <- as.numeric(as.roman(causal.pos.B.2$chr.B))

# order values
causal.pos.B.2.order <- causal.pos.B.2[order(chr.A, start.A, chr.B, start.B)]

# organize coordinates so that they are ordered by chromosome
# vector of chromosomes
vchr <- 1:16
# how much space will be separating chromosomes
separator <- 1e5

coordinates_plot <- sort_by_chr(vchr = vchr, causal.pos.B.2.order, separator = separator)

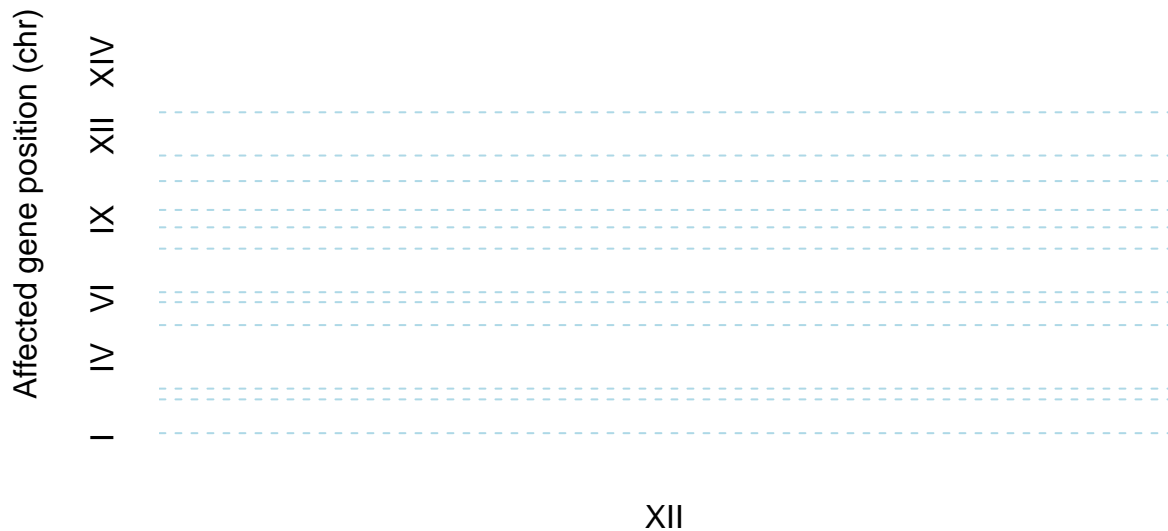
if (!exists("find.effects")){
  find.effects <- fread("results/findeffects_all_newparams.gz")
}
#Create a function to generate a continuous color palette
```

```
# rbPal <- colorRampPalette(c('blue','red'))

coordinates_plot_cor <- merge(coordinates_plot, find.effects[,.(geneA, geneB, cor)], by=c("geneA", "geneB", "cor"))
coordinates_plot_cor$cor <- abs(coordinates_plot_cor$cor)
coordinates_plot_cor <- coordinates_plot_cor[order(cor)]
# coordinates_plot_cor$col <- rbPal(100)[as.numeric(cut(coordinates_plot_cor$cor,breaks = 100))]

# plot_sorted_coordinates(coordinates_plot_cor, separator = separator, col = coordinates_plot_cor$col)

# coordinates_plot_cor[chr.A==12]
plot_sorted_coordinates(coordinates_plot_cor[chr.A==12], separator = separator, col = coordinates_plot_cor$col)
```

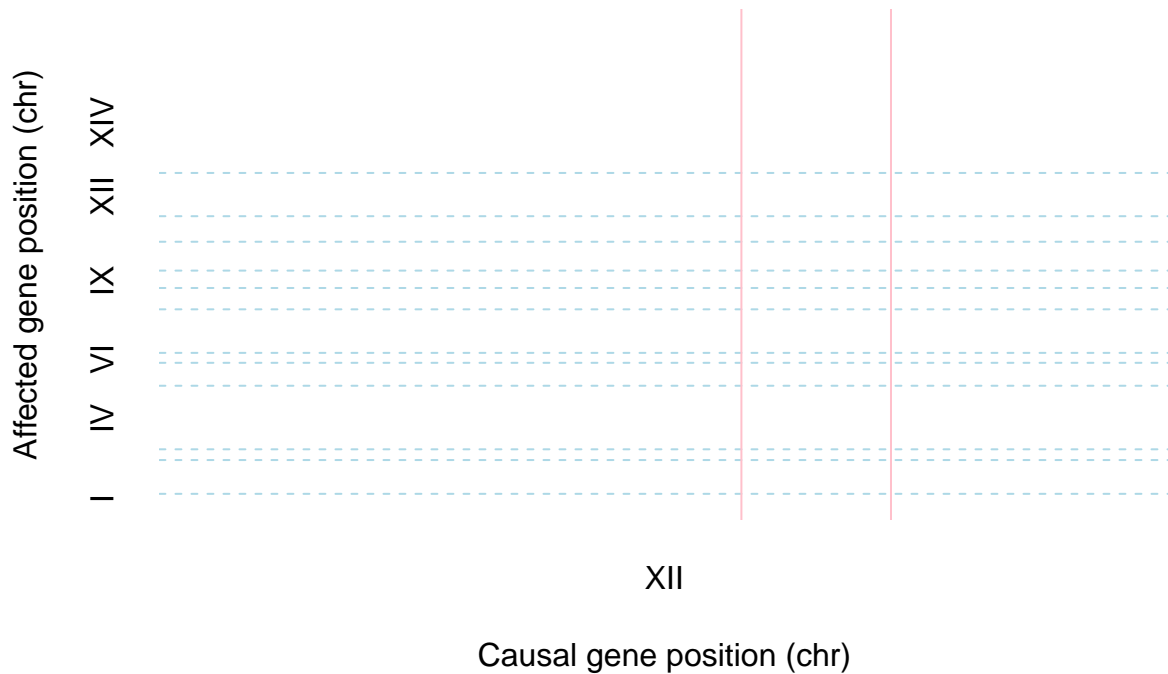


Causal gene position (chr)

Chromosome 3 seems to mostly be affected by the genes in the vertical bands so I'm going to look closer at chromosome 3

```
# the first gene of chromosome 12 that affects chromosome 3
lower_limit <- coordinates_plot_cor[chr.A==12 & chr.B==3][order(start.A)][1]$start.A
# the last gene on chromosome 12 that affects chromosome 3 and that's in the band (there's an extra gene)
top_limit <- coordinates_plot_cor[chr.A==12 & chr.B==3][order(start.A)][nrow(coordinates_plot_cor[chr.A==12])-$start.A]

# plot chromosome 12 - the pink lines are the limits of the group of vertical bands
plot_sorted_coordinates(coordinates_plot_cor[chr.A==12], separator = separator, col = coordinates_plot_cor$col)
abline(v = c(lower_limit-1500, top_limit+1500), col="pink")
```



there are 30 genes between the two pink bands

There are

```
nrow(unique(coordinates_plot_cor[chr.A==12 & chr.B==3][order(start.A)][,(geneA, start.A)])[between(start.A,lower_limit,
top_limit )]) #genes that between the two pink bands
```

```
plot_sorted_coordinates(coordinates_plot_cor[chr.A==12], separator = separator, col = coordinates_plot_cor$col, xlim=c(7422300, top_limit+1000)) # there are abline(v=c(7422391-1500, 7425907+1500), col="pink") nrow(unique(coordinates_plot_cor[chr.A==12 & chr.B==3][order(start.A)][,(geneA, start.A)])[between(start.A,7422391-1500, 7425907+1500)])
```

```
plot_sorted_coordinates(coordinates_plot_cor[chr.A==12], separator = separator, col = coordinates_plot_cor$col, xlim=c(7422391-1500, 7425907+1500))
```