Today is November 30, 2012.

```
> rm(list = ls())
> library("vcf2R")
> library("GenomicRanges")
> library("BSgenome.Hsapiens.UCSC.hg19")
> library("Gviz")
```

1 Targeted Loci

```
> file <- system.file("extdata", "snvs.fixed_header.frq",
+     package = "vcf2R")
> pos.df <- frq2pos(file)
> head(pos.df)

> with(pos.df, table(chr))

chr
    1    10    14    17    20     4     8     9
44713 29857  1559 29486 18740  2444 27639 20751

> pos.gr <- with(pos.df, GRanges(seqnames = paste("chr",
+     chr, sep = ""), ranges = IRanges(start = pos, end = pos),
+     strand = "*"))
> seqlengths(pos.gr) <- seqlengths(Hsapiens)[levels(seqnames(pos.gr))]
> start(pos.gr) <- start(pos.gr) - 10000
> end(pos.gr) <- end(pos.gr) + 10000
> loci <- reduce(pos.gr)
> start(loci) <- start(loci) + 10000
> end(loci) <- end(loci) - 10000
```

```
> val <- data.frame(n.snps = countOverlaps(loci, pos.gr),
+     kb = width(loci)/1000)
> values(loci) <- val
> names(loci) <- c("PAX7", "ABCA4", "IRF6", "VAX1", "FGFR2",
+     "BMP4", "NTN1", "NOG", "MAFB", "MSX1", "8q24", "PTCH1",
+     "FOXE1")
> loci

GRanges with 13 ranges and 2 metadata columns:
         seqnames                 ranges strand |   n.snps          kb
            <Rle>              <IRanges>  <Rle> | <integer> <numeric>
   PAX7     chr1 [ 18771811,  19208536]      * |     13309    436.726
  ABCA4     chr1 [ 94324232,  95013570]      * |     16310    689.339
   IRF6     chr1 [209836758, 210468863]      * |     15094    632.106
   VAX1    chr10 [118421219, 119167920]      * |     18546    746.702
  FGFR2    chr10 [123095899, 123499252]      * |     11311    403.354
   BMP4    chr14 [ 54382326,  54445461]      * |      1559     63.136
   NTN1    chr17 [  8754615,   9266552]      * |     15212    511.938
    NOG    chr17 [ 54402391,  54957844]      * |     14274    555.454
   MAFB    chr20 [ 38902239,  39614945]      * |     18740    712.707
   MSX1     chr4 [  4824627,   4901842]      * |      2444     77.216
   8q24     chr8 [129295457, 130355340]      * |     27639   1059.884
  PTCH1     chr9 [ 98133217,  98413595]      * |      7795    280.379
  FOXE1     chr9 [100357256, 100877263]      * |     12956    520.008
  ---
  seqlengths:
        chr1      chr10     chr14 ...       chr4       chr8       chr9
   249250621  135534747 107349540 ... 191154276  146364022  141213431

> sum(width(loci))/1e+06

[1] 6.688949

> with(values(loci), sum(n.snps))

[1] 175189

> data("targets.hg19")
> targets.hg19.gr

GRanges with 13 ranges and 0 metadata columns:
         seqnames                 ranges strand
            <Rle>              <IRanges>  <Rle>
   IRF6     chr1 [209837199, 210468406]      *
   MAFB    chr20 [ 38902646,  39614513]      *
  ABCA4     chr1 [ 94324660,  95013109]      *
   8q24     chr8 [129295896, 130354946]      *
  FOXE1     chr9 [100357692, 100876841]      *
   PAX7     chr1 [ 18772300,  19208054]      *
   VAX1    chr10 [118421625, 119167424]      *
   NTN1    chr17 [  8755114,   9266060]      *
    NOG    chr17 [ 54402837,  54957390]      *
```

```
  MSX1    chr4 [  4825126,   4901385]      *
  BMP4   chr14 [ 54382690,  54445053]      *
 FGFR2   chr10 [123096374, 123498771]      *
 PTCH1    chr9 [ 98133647,  98413162]      *
 ---
 seqlengths:
   chr1 chr10 chr14 chr17 chr20  chr4  chr8  chr9
     NA    NA    NA    NA    NA    NA    NA    NA
```

2 Regions

```
> data(kg.hg19)
> kg.gr <- GRanges(seqnames = kg.hg19$chr, ranges = IRanges(start = kg.hg19$start,
    end = kg.hg19$end), strand = kg.hg19$strand, exon = as.factor(kg.hg19$exon),
    gene = as.factor(gene <- substr(kg.hg19$exon, 1,
        8)), feature = as.factor(gene), transcript = as.factor(gene),
    symbol = as.factor(gene), id = as.factor(gene))

> ht <- 4
> wd <- ht * 1.618
> for (locus in names(loci)) {
    gr <- targets.hg19.gr[names(targets.hg19.gr) == locus,
        ]
    chr <- as.character(unique(seqnames(gr)))
    seqlengths(gr) <- seqlengths(Hsapiens)[levels(seqnames(gr))]
    atrack <- AnnotationTrack(gr, name = paste0(locus,
        " Target"), chromosome = chr, genome = "hg19",
        stacking = "squish")
    atrack2 <- AnnotationTrack(loci[names(loci) == locus,
        ], chromosome = chr, genome = "hg19", name = paste0(locus,
        " Actual"), stacking = "squish")
    gtrack <- GenomeAxisTrack()
    itrack <- IdeogramTrack(genome = "hg19", chromosome = chr)
    kg.sub.gr <- kg.gr[subjectHits(findOverlaps(loci[names(loci) ==
        locus, ], kg.gr)), ]
    if (length(kg.sub.gr) != 0) {
        grtrack <- GeneRegionTrack(range = kg.sub.gr,
            genome = "hg19", chromosome = as.character(seqnames(loci)[which(names(loci) ==
                locus)]), name = "Known Genes")
        pdf(file = paste0("./figures/loci/", locus, ".pdf"),
            height = ht, width = wd)
        plotTracks(list(atrack2, atrack, grtrack, gtrack,
            itrack), add = FALSE, main = locus)
        dev.off()
    }
    else {
        pdf(file = paste0("./figures/loci/", locus, ".pdf"),
            height = ht, width = wd)
        plotTracks(list(atrack2, atrack, gtrack, itrack),
            add = FALSE, main = locus)
        dev.off()
```
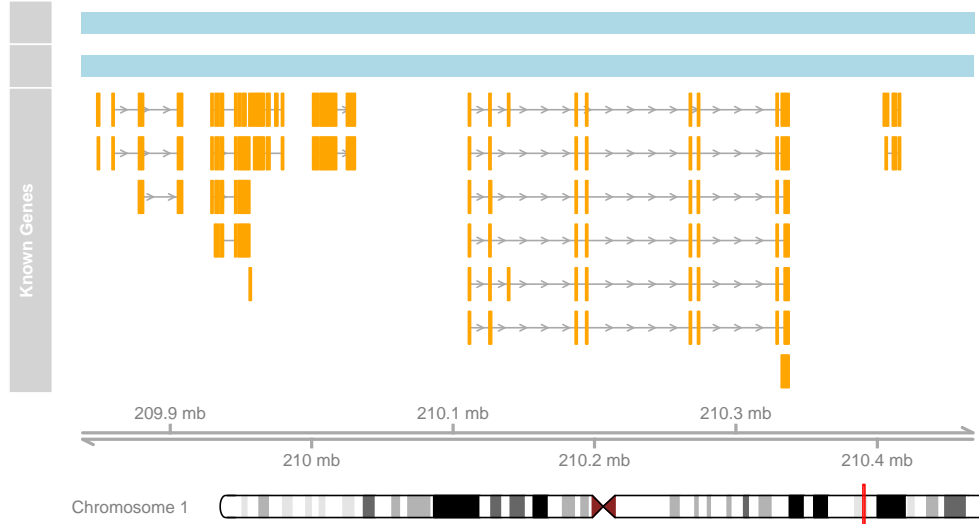
FOXE1

```
    }
}
```
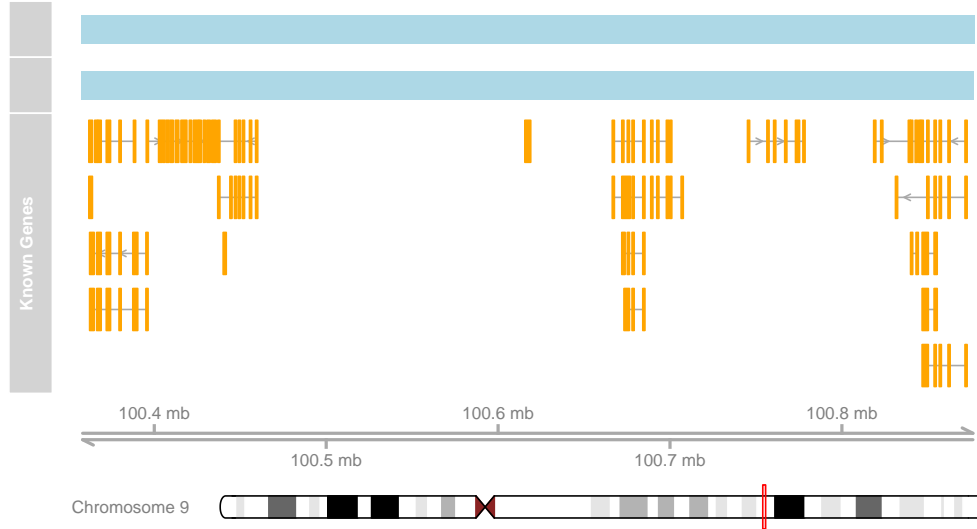
3   Hardy-Weinberg Equilibrium

```
> file <- system.file("extdata", "snvs.fixed_header.hwe",
      package = "vcf2R")
> hwe.df <- hwe2R(file = file)
> head(hwe.df)
```

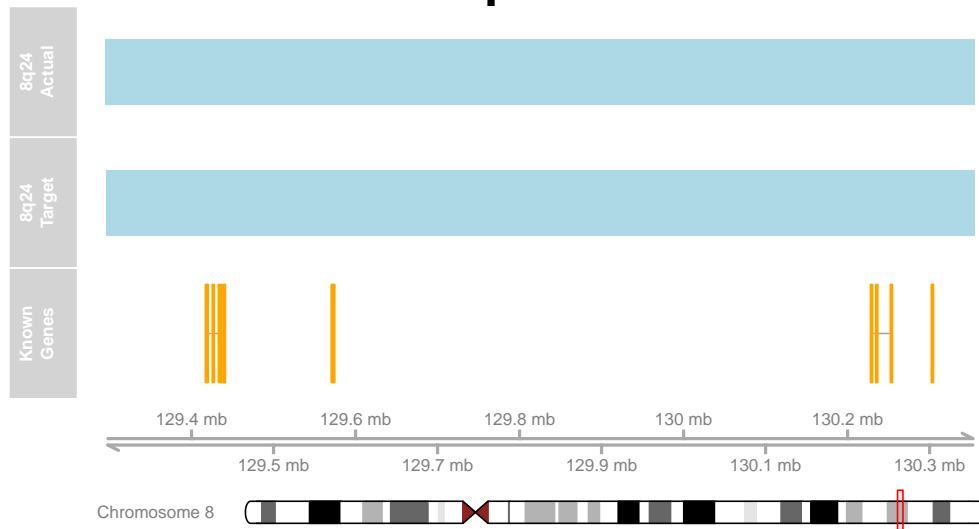# 4   Minor Allele Frequencies

```
> file <- system.file("extdata", "8q24-parents.frq", package = "vcf2R")
> maf.di <- frq2R(file = file)
```

IRF6

8q24

BMP4