

Today is November 29, 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)

  chr      pos
1   1 18771811
2   1 18771817
3   1 18771849
4   1 18771875
5   1 18771893
6   1 18771920

> 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)

> 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.df <- as.data.frame(kg.gr[subjectHits(findOverlaps(loci[names(loci) ==
    locus, ], kg.gr)), ])
  if (nrow(kg.sub.df) != 0) {
    kg.sub.df <- kg.sub.df[, -c(1, 4, 5)]
    grtrack <- GeneRegionTrack(kg.sub.df, 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()
  }
}
}

```

```
> kg.sub.df
```

	start	end
1	100361915	100364025
2	100361915	100364025
3	100365296	100367147
4	100367764	100367923
5	100368424	100368543
6	100372627	100372733
7	100373855	100373981
8	100380038	100380159
9	100387962	100389046
10	100362361	100365049
11	100367008	100367147
12	100367764	100367923
13	100368424	100368543
14	100372627	100372733
15	100373855	100373981
16	100380038	100380159
17	100387962	100388279
18	100389679	100389894
19	100395524	100395962
20	100362361	100365049
21	100367008	100367147
22	100367764	100367923
23	100368424	100368543
24	100372627	100372733
25	100373855	100373981
26	100380038	100380159
27	100387962	100388279
28	100389587	100389894
29	100395524	100395962
30	100395704	100396197
31	100403075	100403164
32	100403843	100403944
33	100405485	100405642
34	100407398	100407506
35	100407892	100408014
36	100409773	100409843
37	100410289	100410505
38	100412784	100412882
39	100413591	100413655
40	100416079	100416190
41	100417161	100417226
42	100417983	100418046
43	100418292	100418367
44	100420921	100421025
45	100423227	100423350
46	100424323	100424426
47	100425235	100425329
48	100426621	100426725
49	100429044	100429159

50	100431127	100431256
51	100431874	100431988
52	100433367	100436029
53	100440044	100441883
54	100437190	100437869
55	100444481	100444712
56	100447204	100447322
57	100449377	100449543
58	100451815	100451921
59	100455930	100456041
60	100459402	100459691
61	100437190	100437869
62	100447204	100447322
63	100449377	100449543
64	100451815	100451921
65	100455930	100456041
66	100459402	100459691
67	100615536	100618997
68	100666771	100667274
69	100672241	100672898
70	100675682	100675840
71	100678445	100678620
72	100684699	100684852
73	100666771	100667274
74	100672241	100672898
75	100673199	100675257
76	100675682	100675840
77	100676260	100676308
78	100678445	100678620
79	100684699	100684852
80	100672241	100672898
81	100673199	100673352
82	100675682	100675840
83	100678445	100678620
84	100684699	100684852
85	100673199	100675257
86	100675682	100675840
87	100678445	100678620
88	100684699	100684852
89	100689072	100689760
90	100692316	100693503
91	100698452	100698546
92	100700339	100700526
93	100689072	100689760
94	100692316	100693503
95	100698452	100698546
96	100700339	100700473
97	100707049	100707134
98	100745488	100745891
99	100756912	100757062
100	100760837	100760960

```

101 100767245 100767435
102 100773552 100773671
103 100774702 100774754
104 100777645 100778224
105 100840474 100840629
106 100843097 100844448
107 100818958 100819222
108 100823063 100823279
109 100839199 100839299
110 100840474 100840629
111 100843097 100843364
112 100845127 100845365
113 100846636 100850287
114 100854190 100854283
115 100854772 100854843
116 100846636 100850287
117 100854190 100854283
118 100854779 100854843
119 100831568 100831969
120 100849723 100850287
121 100854190 100854283
122 100857148 100857311
123 100862212 100862446
124 100872170 100872266
125 100846636 100846748
126 100849723 100850287
127 100854190 100854283
128 100857148 100857311
129 100862212 100862446
130 100872170 100872266
131 100846636 100850287
132 100854190 100854283
133 100857148 100857311
134 100862212 100862446
135 100872170 100872266

```

### 3 Hardy-Weinberg Equilibrium

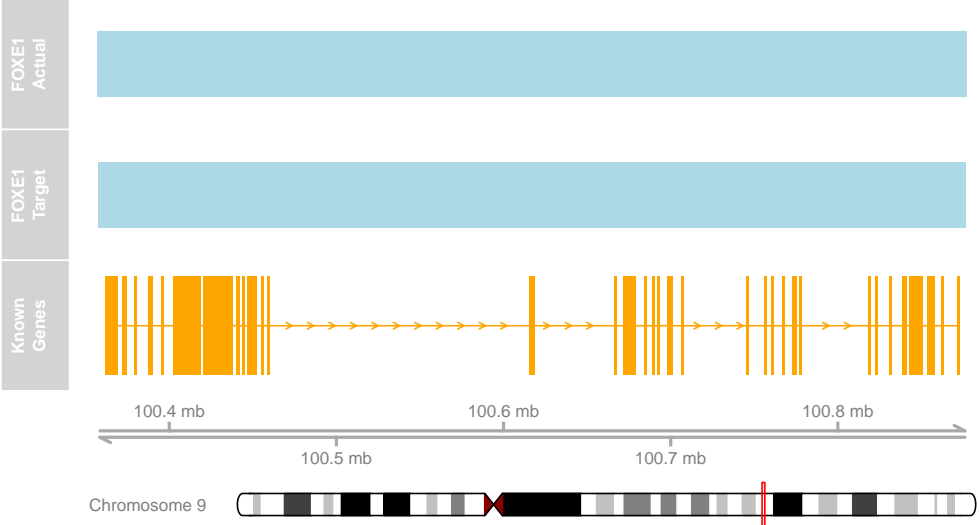
```

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

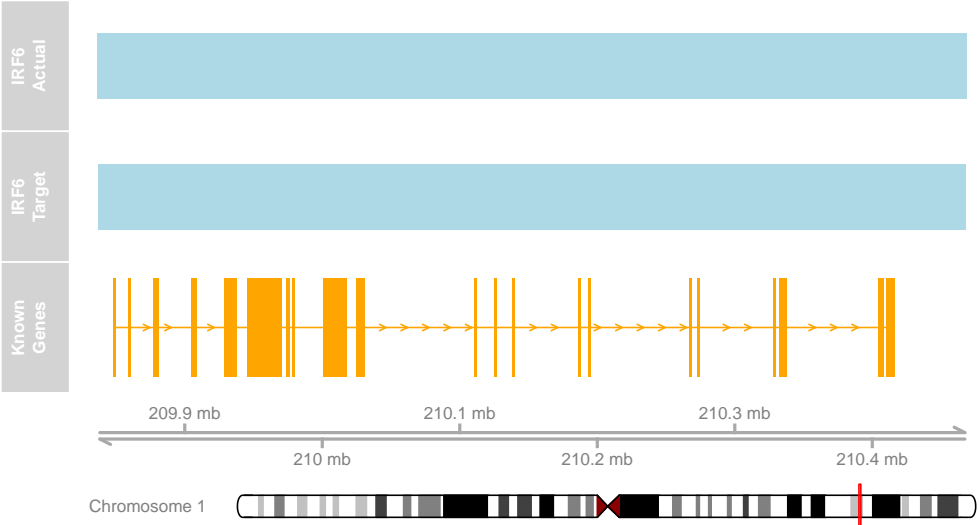
```

	chr	pos	chi.sq	p.value	homo.major	het	homo.minor
1	1	18772169	0.000056	1.000000	4494	1	0
2	1	18772190	0.000223	1.000000	4493	2	0
3	1	18772191	15.727310	0.053862	4464	30	1
4	1	18772203	0.000223	1.000000	4493	2	0
5	1	18772218	0.000223	1.000000	4493	2	0
6	1	18772236	0.016134	1.000000	4478	17	0

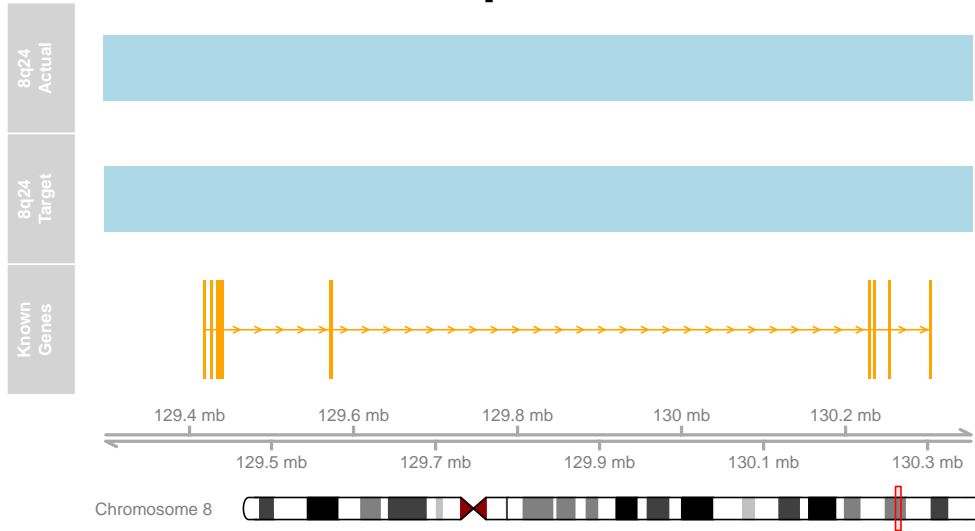
# FOXE1



# IRF6



# 8q24



# BMP4

