

Evolution and genomics, Cesky Krumlov

Daniel Berner

2. February 2016

Genomic analyses using RADseq:

4. Exploring selective sweeps based on read coverage

Demo

Upload of a RAD locus coverage file derived from alignments (4 lake-stream stickleback individuals from Roesti et al. 2015 Nat. Commun., ChrI only, physical resolution reduced)

```
# d<-read.table('C:/Users/daniel/Documents/science/teaching/cesky  
# krumlov  
# 2016/course.materials/R.files/coverage.chrI.txt', h=T)  
d[c(1, 415215), ]
```

```
##          habitat individual position          sequence  
## 1          stream    ind_A      1655 TGCATTTTATCTGAGTACTA  
## 415215     lake      ind_D 28160288 TGCATGTTGAAATCCCCACC
```

```
unique(d$habitat)
```

```
## [1] "stream" "lake"
```

```
unique(d$individual)
```

```
## [1] "ind_A" "ind_B" "ind_C" "ind_D"
```

```
length(unique(d$position))
```

```
## [1] 2117
```

Calculating of the proportion of reads at the first RAD locus contributed by the stream stickleback individuals

```
loc <- d[which(d$position == "1655"), ]  
loc[c(1, 424), ]
```

```
##      habitat individual position      sequence  
## 1      stream      ind_A      1655 TGCATTTTATCTGAGTACTA  
## 424     lake      ind_D      1655 TGCATTTTATCTGAGTACGA
```

```
Nstr <- length(which(loc[, 1] == "stream"))  
cov <- Nstr/length(loc[, 1])  
cov
```

```
## [1] 0.526
```

Tasks

- ▶ Upload the stickleback RAD locus coverage file *coverage.chrl.txt*, and inspect the structure of the data
- ▶ For each RAD locus, calculate the proportion of the total reads contributed by the stream population by using a loop
- ▶ Plot this coverage proportion along the chromosome, including a LOESS-smoothed profile