

Practical courses

Amin Madoui

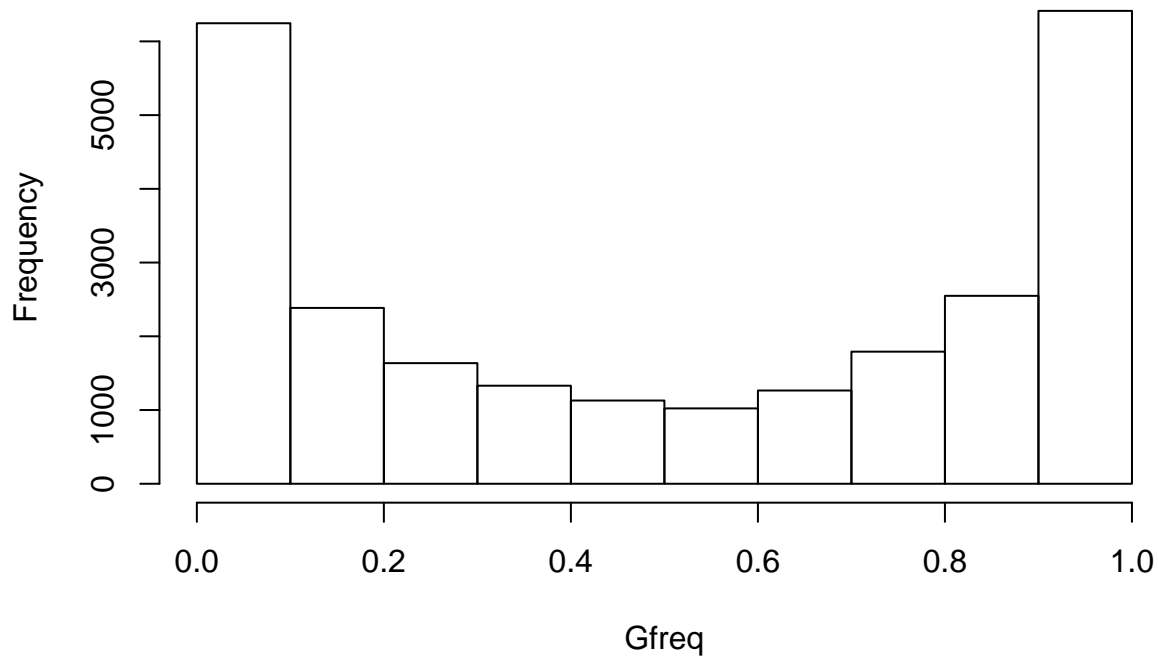
April 2021

Estimation of the genomic differentiation

Allele frequency spectrum

```
“r data=read.table('data/RC/S155.txt')
colnames(data) = c("GA","GB","TA","TB")
Gfreq = dataGB/(dataGA+data$GB)
hist(Gfreq, breaks = 10, main="B allele frequency distribution")
“
```

B allele frequency distribution



```
## Pairwise-FST
r source ("popgene_functions.R")
## Warning: package 'ggplot2' was built under R version 3.5.3
```

```

“r freq = read.table(“data/Genomic_frequencies.txt”)
colnames(freq) = c(“S155”,“S158”,“S178”,“S206”,“S208”,“S209”,“S210”)

pairwiseFST = pwFst(freq)

print(pairwiseFST)
“:

```

```

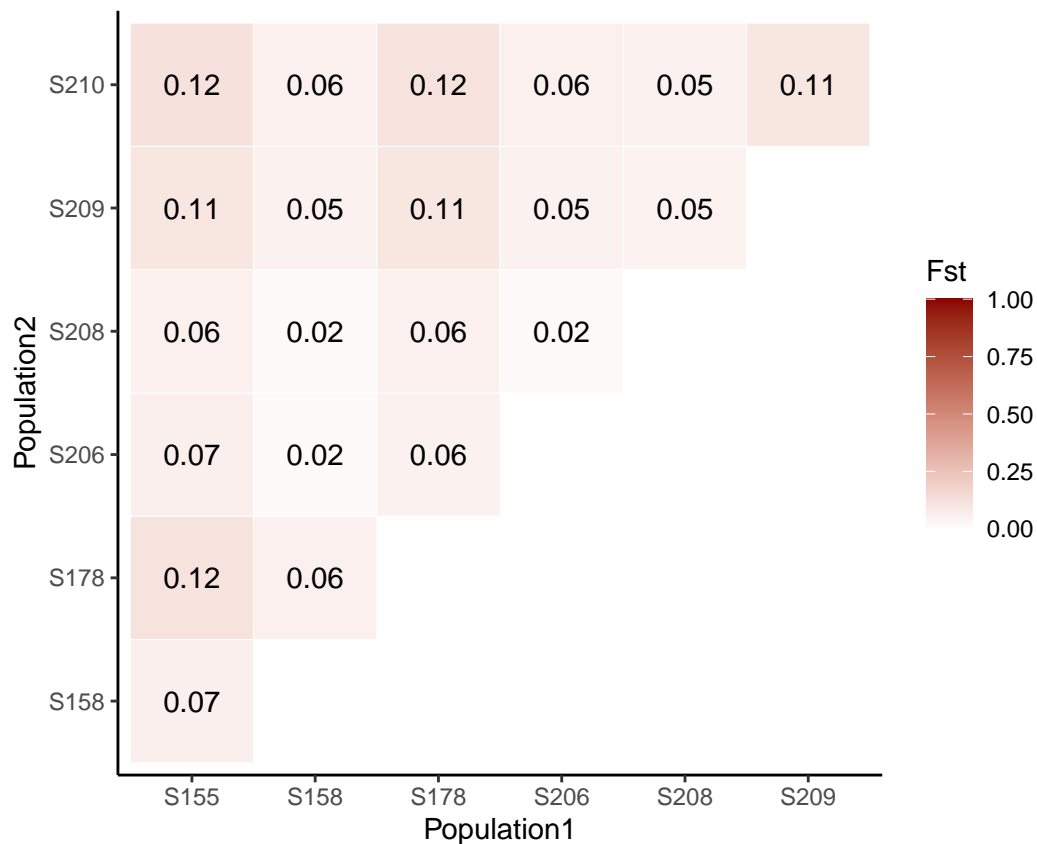
##          S155      S158      S178      S206      S208      S209  ## S155 0.00000000
0.06555166 0.11764706 0.06779661 0.06185567 0.10526316 ## S158 0.06555166 0.00000000
0.05994814 0.02071444 0.02214627 0.05405405 ## S178 0.11764706 0.05994814 0.00000000
0.05570515 0.06086776 0.10526316 ## S206 0.06779661 0.02071444 0.05570515 0.00000000
0.01930026 0.05282113 ## S208 0.06185567 0.02214627 0.06086776 0.01930026 0.00000000
0.04788885 ## S209 0.10526316 0.05405405 0.10526316 0.05282113 0.04788885 0.00000000 ##
S210 0.12244898 0.06143691 0.11764706 0.06235840 0.05405405 0.10526316 ##          S210
## S155 0.12244898 ## S158 0.06143691 ## S178 0.11764706 ## S206 0.06235840 ## S208
0.05405405 ## S209 0.10526316 ## S210 0.00000000

```

```

r plotFST(pairwiseFST)

```



```

## Isolation by currents
# Detection of loci under natural selection
## FST outliers method through LK
“r source (“popgene_functions.R”)

freq = read.table(“data/Genomic_frequencies.txt”)
colnames(freq) = c(“S155”,“S158”,“S178”,“S206”,“S208”,“S209”,“S210”)

```

```

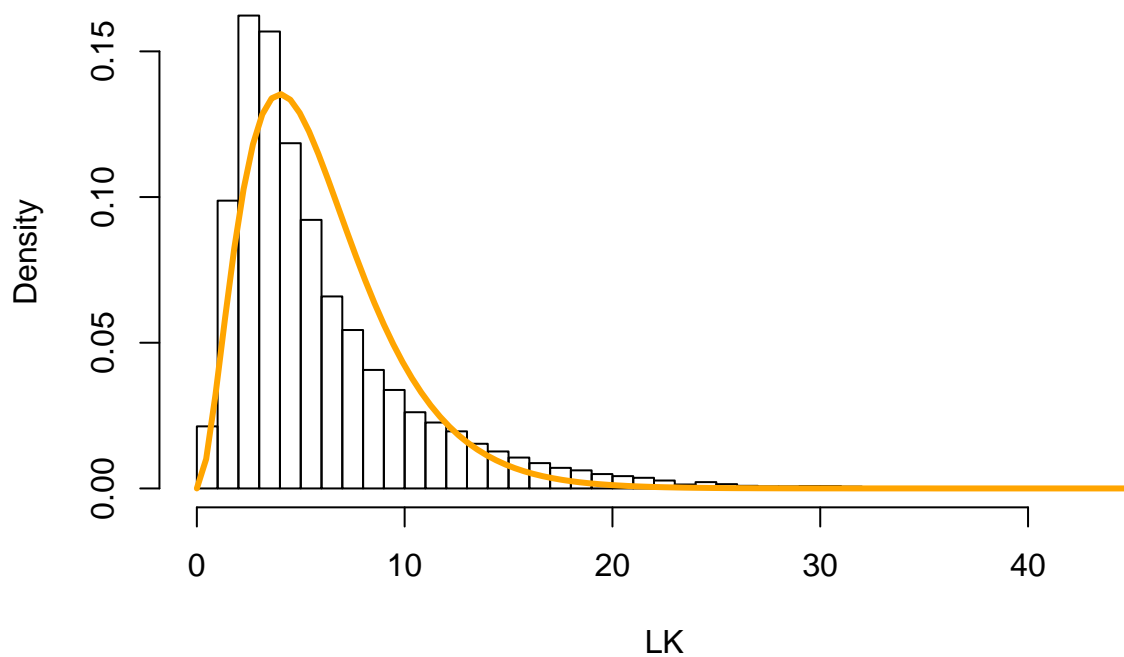
LK = LK(freq)
head( LK)
““

##      varID      Fst      LK      p_value      q_value      ## 1      1 0.18817620  8.392163
0.21075678 0.9847290      ## 2      2 0.35262098 15.725967 0.01530288 0.3038530      ## 3      3
0.10773398 4.804652 0.56910119 0.9999753      ## 4      4 0.08336976  3.718072 0.71476587
0.9999753      ## 5      5 0.27152026 12.109089 0.05957895 0.5722980      ## 6      6 0.06629779
2.956707 0.81426066 0.9999753

““r h = hist ( LK$LK , plot = F, breaks = 50 )
n_pop = ncol(freq)-1
hist ( LK$LK, freq = F, xlab = “LK” , breaks = 50 , main = “LK distribution”)
# add the neutral evolution
curve ( dchisq( x, n_pop ) , lwd=3, col=“orange” , add=T )
““

```

LK distribution



```

““r # selected loci with high LK value
selection_LK = LK[LK$q_value<0.1,]
dim(selection_LK)
““

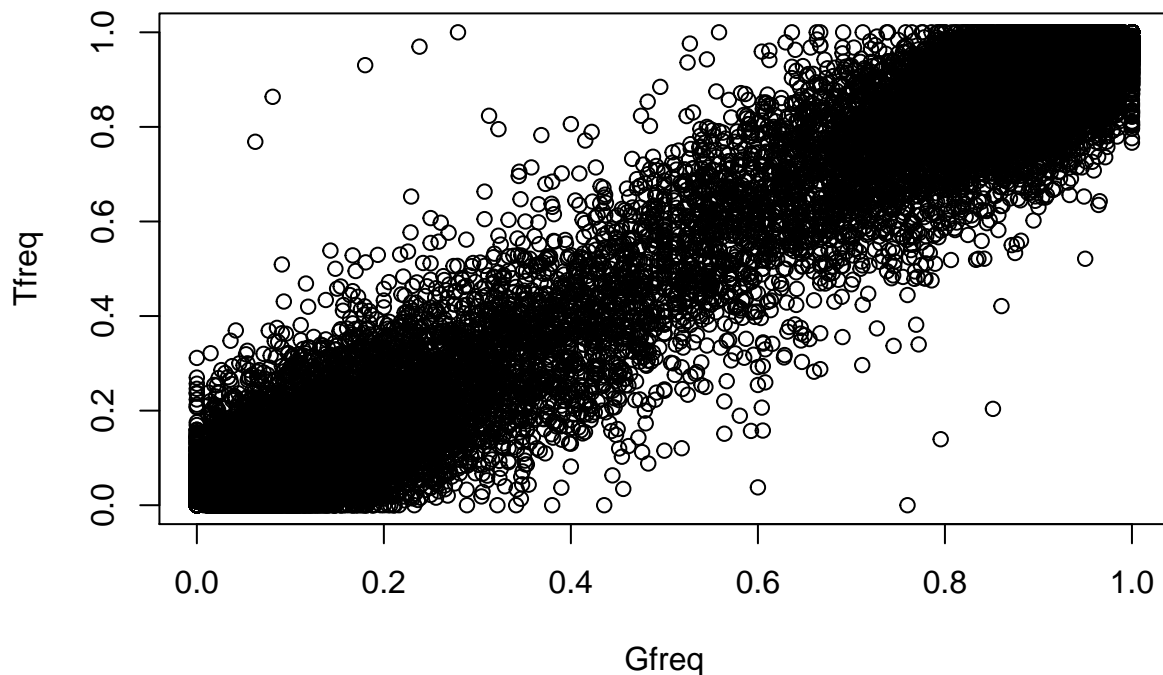
## [1] 426  5
## pcadapt method

```

```

# Population-scale allele expression
## Comparison of allele expression to genomic abundance
“r data=read.table(‘data/RC/S206.txt’)
colnames(data) = c(“GA”,“GB”,“TA”,“TB”)
Gfreq = dataGB/(dataGA+data$GB)
Tfreq = dataTB/(dataTA+data$TB)
plot(Gfreq,Tfreq)
“

```



```

r summary(lm(Tfreq~Gfreq-1))

## ## Call: ## lm(formula = Tfreq ~ Gfreq - 1) ## ## Residuals: ## Min
1Q Median 3Q Max ## -0.75950 -0.04430 0.00066 0.04646 0.78261 ##
## Coefficients: ## Estimate Std. Error t value Pr(>|t|) ## Gfreq 0.9993388
0.0008202 1218 <2e-16 *** ## --- ## Signif. codes: 0 '***' 0.001 '**' 0.01
'*' 0.05 '.' 0.1 ' ' 1 ## ## Residual standard error: 0.08388 on 25767 degrees of
freedom ## Multiple R-squared: 0.9829, Adjusted R-squared: 0.9829 ## F-statistic:
1.484e+06 on 1 and 25767 DF, p-value: < 2.2e-16

## Detection of allele-specific expression (ASE)
“r Fisher_pvalue = apply(data,1,function (x) fisher.test (matrix(x, nrow = 2))$p.value)
Fisher_qvalue = p.adjust(Fisher_pvalue, method=“BH”)
“

```

```

## Link between natural selection and allele specific expression
"r selection_ASE_S206 = LK[LK$q_value<0.1 & Fisher_qvalue < 0.1,]
dim(selection_ASE_S206)
"

## [1] 2 5

r phyper(nrow(selection_ASE_S206),          nrow(LK[LK$q_value<0.1,]),          nrow(LK)-nrow(LK[L
length(Fisher_qvalue[Fisher_qvalue<0.1] ))
## [1] 0.004100667

# Reference-free approach and metavariant species
## Genomic differentiation of MVS
## Role of the environmental factors
## Variance partitioning of environmental factors

```