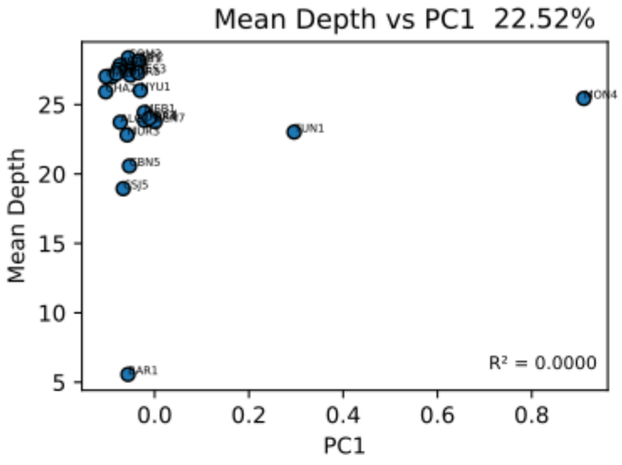
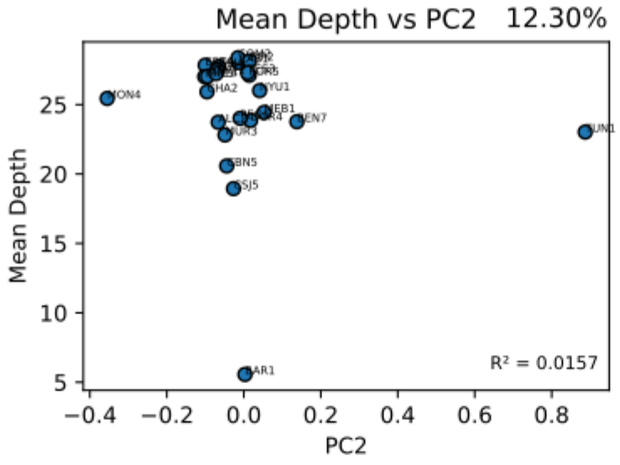
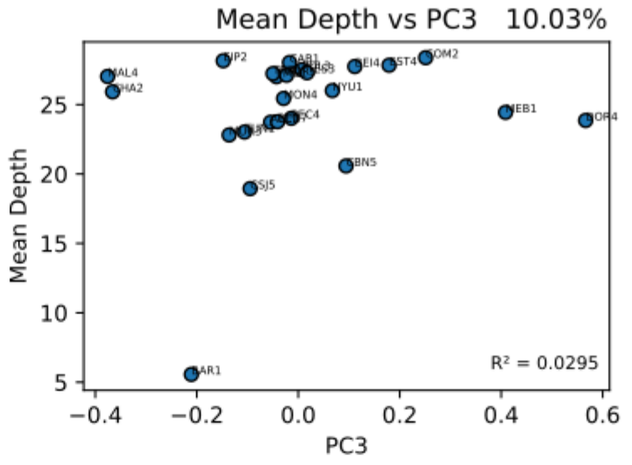
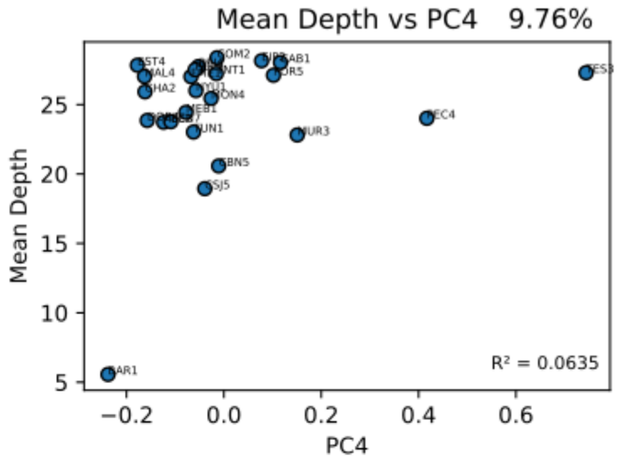
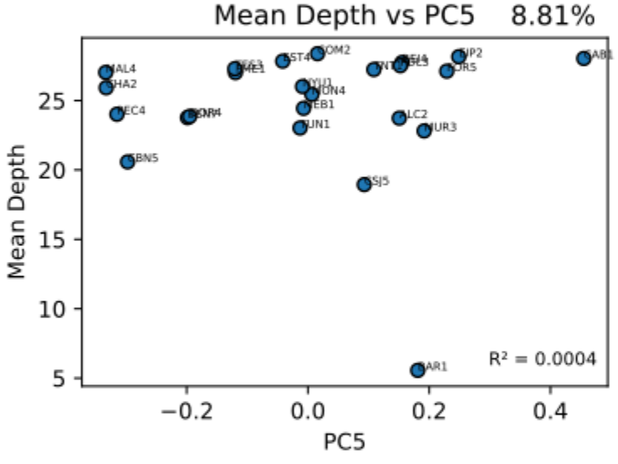
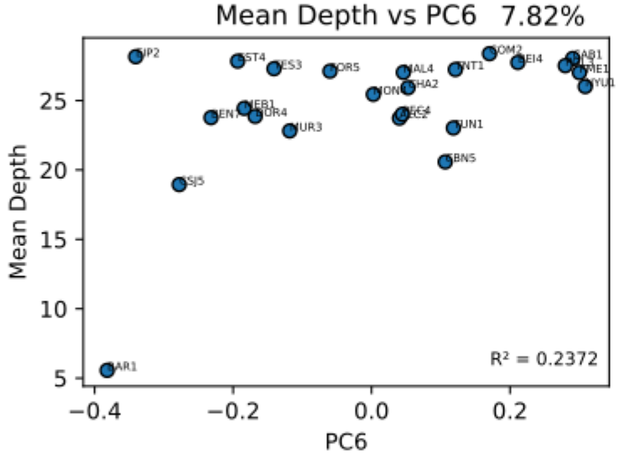

Mean Depth vs PC1 22.52%
Mean Depth
PC1
R² = 0.0000

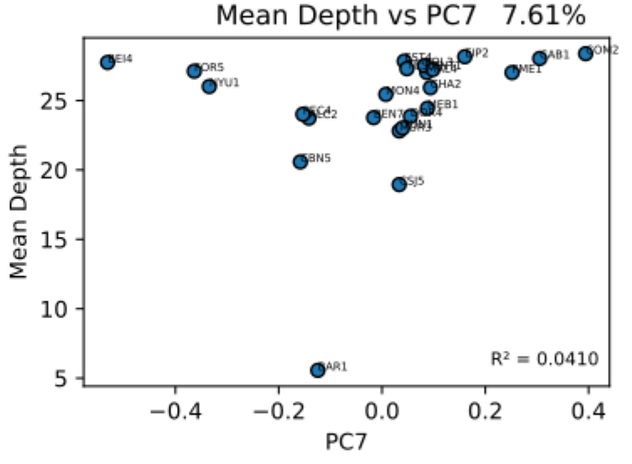

Mean Depth vs PC2 12.30%
Mean Depth
PC2
R² = 0.0157

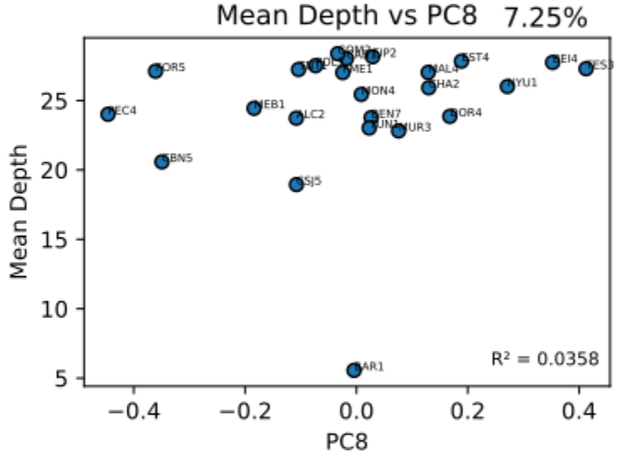

Mean Depth vs PC3 10.03%
Mean Depth
PC3
R² = 0.0295

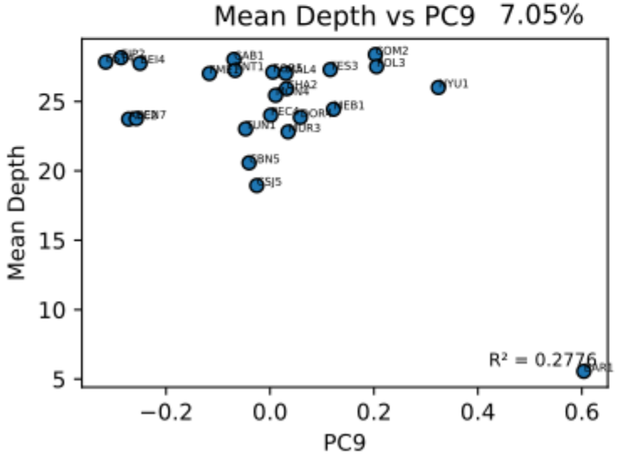

Mean Depth vs PC4 9.76%
Mean Depth
PC4
R² = 0.0635

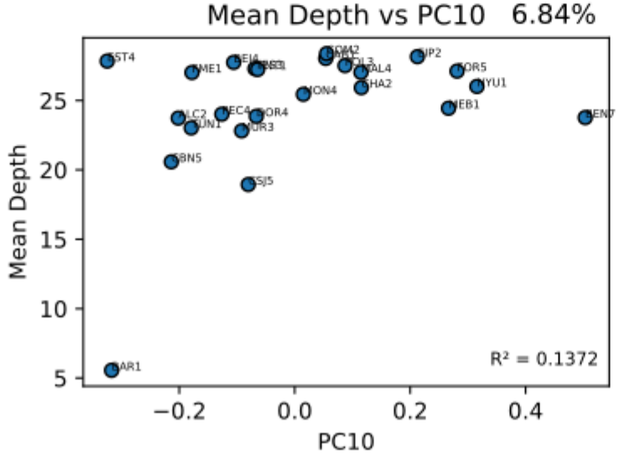

Mean Depth vs PC5 8.81%
Mean Depth
PC5
R² = 0.0004


Mean Depth vs PC6 7.82%
Mean Depth
PC6
R² = 0.2372


Mean Depth vs PC7 7.61%
Mean Depth
PC7
R² = 0.0410


Mean Depth vs PC8 7.25%
Mean Depth
PC8
R² = 0.0358


Mean Depth vs PC9 7.05%
Mean Depth
PC9
R² = 0.2776


Mean Depth vs PC10 6.84%
Mean Depth
PC10
R² = 0.1372


Tajima's D by Chromosome (highlighting D < –2 or > 2)
Chr01
Chr02
Chr03
Chr04
Chr05
Chr06
Chr07
Chr08
Chr09
Chr10
Tajima's D
Genomic Position (Mb)
High (> 2)

```
INDV,%MISSING_SITES,MEAN_DEPTH,F
ALC2,0.235758,23.7279,-0.10846
BAR1,9.83838,5.55851,0.431
BEI4,0.285767,27.7428,-0.08612
BEN7,0.129978,23.7674,-0.1126
CAB1,0.12437,28.031,-0.10769
CBN5,0.498172,20.5783,-0.04671
CHA2,0.109775,25.9215,-0.1353
COM2,0.0778178,28.3768,-0.20983
CSJ5,0.238754,18.9442,-0.05921
DOR4,0.145726,23.8577,-0.22954
EIP2,0.118686,28.1565,-0.17829
EST4,0.107086,27.8358,-0.04982
FME1,0.159169,27.0126,-0.04754
FOR5,0.118686,27.1138,-0.18174
MAL4,0.113231,27.0269,-0.14829
MEB1,0.143191,24.4381,-0.33045
MON4,1.10958,25.4431,0.31073
MUR3,0.187131,22.8108,-0.14097
NYU1,0.212251,26.0087,-0.04432
PEC4,0.31081,24.0121,-0.16855
POL3,0.520066,27.5114,0.02496
TES3,0.257728,27.2931,-0.16832
TNT1,0.150105,27.234,-0.08089
TUN1,0.840862,23.0211,0.1537
```

Sample1,Sample2,IBD_length
CHA2,MAL4,67.38
DOR4,EST4,55.4
MON4,TUN1,54.42
BEN7,TUN1,35.3
PEC4,TES3,32.47
BEI4,NYU1,30.89
MUR3,TES3,30.55
MUR3,PEC4,26.66
CSJ5,EST4,25.54
FOR5,TES3,23.3
BEI4,EIP2,22.04
COM2,POL3,22.04
BEI4,CBN5,21.14
CAB1,FME1,21.08
CAB1,COM2,19.42
BEI4,TES3,19.33
BEN7,EST4,19.3
COM2,DOR4,19.06
CBN5,POL3,18.2
NYU1,PEC4,17.46
COM2,FME1,16.48
EST4,POL3,16.31
FOR5,TNT1,15.77
BEI4,EST4,15.59
BEN7,EIP2,15.25
EIP2,PEC4,15.02
POL3,TNT1,14.51
ALC2,TNT1,14.07
BEN7,FOR5,13.92
EST4,MEB1,13.88
TES3,TNT1,13.78
EIP2,TES3,13.11
BEI4,COM2,12.08
EST4,FME1,11.83
CAB1,TES3,11.79
BAR1,CHA2,11.6
CSJ5,DOR4,11.23
BEI4,DOR4,11.05
MAL4,TES3,10.82
BEI4,CSJ5,10.47
EIP2,EST4,10.37
NYU1,POL3,10.33
BEI4,FME1,10.3
FME1,TES3,10.28
BEI4,POL3,10.03
CAB1,EIP2,9.82
NYU1,TUN1,9.09
BAR1,MAL4,9.05
EIP2,FOR5,8.91
FOR5,TUN1,8.68
FME1,TNT1,8.6
DOR4,MEB1,8.56
BEN7,MON4,8.55
MEB1,NYU1,8.51
FME1,MAL4,8.26
CHA2,TNT1,7.94
CBN5,CSJ5,7.75
CBN5,CHA2,7.68
PEC4,POL3,7.66
FME1,POL3,7.38
BEI4,MAL4,7.28
MAL4,TNT1,7.25
CHA2,EIP2,7.23

COM2,TNT1,7.2
ALC2,EST4,7.02
MUR3,TUN1,6.83
MUR3,NYU1,6.81
BEI4,FOR5,6.61
BEI4,MEB1,6.53
MAL4,MUR3,6.46
EST4,MUR3,6.38
EST4,MAL4,6.29
COM2,TUN1,6.26
CSJ5,MUR3,6.01
CHA2,COM2,5.7
ALC2,POL3,5.57
EIP2,POL3,5.5
DOR4,TES3,5.42
CHA2,FME1,5.41
CSJ5,FOR5,5.36
CAB1,NYU1,5.33
COM2,MUR3,5.27
BEN7,POL3,5.24
BEI4,TUN1,5.2
CBN5,MON4,5.11
BAR1,NYU1,4.95
EIP2,FME1,4.81
ALC2,BEN7,4.72
MON4,TNT1,4.66
PEC4,TUN1,4.63
BAR1,TUN1,4.52
CAB1,MUR3,4.2
EST4,TES3,4.2
COM2,CSJ5,4.14
BAR1,BEI4,4.13
EIP2,MUR3,4.05
FOR5,POL3,4
ALC2,EIP2,3.99
CSJ5,POL3,3.97
CBN5,PEC4,3.95
CAB1,MAL4,3.94
ALC2,CAB1,3.92
CAB1,TUN1,3.88
FOR5,PEC4,3.88
BEI4,MUR3,3.87
MEB1,POL3,3.87
MEB1,TNT1,3.86
NYU1,TES3,3.81
ALC2,TES3,3.77
EST4,PEC4,3.76
EIP2,MEB1,3.74
CAB1,FOR5,3.66
ALC2,CSJ5,3.59
DOR4,FOR5,3.59
CBN5,EST4,3.57
FME1,TUN1,3.57
ALC2,MAL4,3.55
COM2,PEC4,3.51
ALC2,CBN5,3.5
BEN7,DOR4,3.48
CHA2,CSJ5,3.47
BEI4,BEN7,3.46
COM2,EST4,3.44
CAB1,MON4,3.43
COM2,MEB1,3.43
ALC2,FOR5,3.42
BEN7,CAB1,3.42

COM2,MON4,3.41
ALC2,COM2,3.38
CHA2,FOR5,3.32
BAR1,DOR4,3.31
ALC2,DOR4,3.3
BAR1,FME1,3.3
CHA2,EST4,3.3
EST4,FOR5,3.3
DOR4,POL3,3.23
BEI4,CHA2,3.19
CBN5,TES3,3.19
EIP2,MAL4,3.18
BEN7,TES3,3.17
CSJ5,MAL4,3.16
BEN7,CSJ5,3.14
MON4,MUR3,3.14
BEI4,TNT1,3.13
FOR5,MEB1,3.12
MUR3,POL3,3.12
CAB1,CSJ5,3.11
EIP2,NYU1,3.09
CAB1,TNT1,3.06
ALC2,BAR1,3.04
MON4,NYU1,3.03
ALC2,NYU1,3.01
COM2,TES3,3.01
NYU1,TNT1,3.01

```
# TRANSFORM vcf.gz file TO ped FORMAT
# plink2 --vcf 24pos_fullG1_FILidLD.vcf.gz --allow-extra-chr --export ped --out
24pos_fullG1_FILidLD

# R ANALYSES
library(tess3r)
library(adegenet)
library(vcfR)
library(LEA)
library(rworldmap)
library(ggplot2)
library(maps)

# ped2lfmm("24pos_fullG1_FILidLD.ped",  output.file =
"24pos_fullG1_FILidLD.lfmm", force = TRUE)

#### WARNING!!!!!!! Open file "24pos_fullG1_FILidLD.lfmm" with Gvim and find and
replace "9" for "NA"
### or use sed; sed -i 's/\(^\|[[:space:]]\)9\([[:space:]]\|$\)/\1NA\2/g'
24pos_fullG1_FILidLD.lfmm

lfmm <- read.lfmm("24pos_fullG1_FILidLD.lfmm")

# table include column 1 with samples names, column 2 with longitude in decimal
values,
# and column 3 latitude in decimal values
coords_df <- read.table("24pos_fullG1_FILidLD.coord", header = FALSE, sep = "\
t", stringsAsFactors = FALSE)
coords <- as.matrix(coords_df[, 2:3])
storage.mode(coords) <- "numeric"
sample_names <- coords_df[, 1]
#rownames(coords) <- sample_names

coords[2:3,]

#plot map with localities of samples
pdf("map_localities.pdf")
plot(coords, pch = 19, cex = .5,
     xlab = "Longitude (°E)", ylab = "Latitude (°N)", asp = 1)
map(add = T, interior = F)
dev.off()

# main tess3 analysis
tess3.obj <- tess3(X = lfmm, coord = coords, K = 1:8,
                   method = "projected.ls", ploidy = 2, openMP.core.num = 4)

# plot crossvalidation values to select best K

pdf("bestK_value.pdf")
plot(tess3.obj, pch = 19, col = "blue",
     xlab = "Number of ancestral populations",
     ylab = "Cross-validation score")
dev.off()

#plot ancestry bar for each samples. REMEMBER TO CHANGE BEST K VALUES, SEE
CROSSVALIDAION RESULTS FROM PREVIOUS STEP

pdf("ancestry_plot_K3.pdf")
q.matrix <- qmatrix(tess3.obj, K = 3) # STRUCTURE-like barplot for the Q-matrix
barplot(q.matrix, border = NA, space = 0,
        xlab = "Individuals", ylab = "Ancestry proportions",
        main = "Ancestry matrix") -> bp
axis(1, at = 1:nrow(q.matrix),
     labels = sample_names[bp$order],  # Match names to new order
```

```
     las = 3, cex.axis = 0.4)
dev.off()

#plot the map with ancestral lineages
pdf("mapa_con_ancestral_lineages.pdf")
plot(q.matrix, coords, method = "map.max", interpol = FieldsKrigModel(10),
     main = "Ancestry coefficients",
     xlab = "Longitude", ylab = "Latitude", asp = 1, resolution = c(300,300),
cex = .4)
dev.off()

#another type of map with ancetral lineages

coords_df_plot <- as.data.frame(coords)
colnames(coords_df_plot) <- c("Longitude", "Latitude")
coords_df_plot$Sample <- rownames(coords)  # optional: keep sample names
pdf("mapa_con_ancestral_lineages_alternative.pdf")
map.polygon <- getMap(resolution = "low")
pl <- ggtess3Q(q.matrix, coords, map.polygon = map.polygon)
pl +
  geom_path(data = map.polygon, aes(x = long, y = lat, group = group)) +
  xlim(-16, 42) +
  ylim(35, 65) +
  coord_equal() +
  geom_point(data = coords_df_plot,
             aes(x = Longitude, y = Latitude),
             size = 0.2) +
  xlab("Longitude") +
  ylab("Latitude") +
  theme_bw()
dev.off()


# retrieve tess3 results for K = 3 and plot histogram p values
pdf("histogram_pvalues.pdf")
p.values <- pvalue(tess3.obj, K =3)
hist(p.values, col = "lightblue")
dev.off()

#When the histogram has the correct shape, lists of outlier loci can be derived
from
#standard False Discovery Rate (FDR) control algorithms. This can be done by
using a
#classic Benjamini-Hochberg correction as follows.

# Benjamini-Hochberg algorithm
L = length(p.values)
fdr.level = 1e-4
w = which(sort(p.values) < fdr.level * (1:L)/L)
candidates = order(p.values)[w]
length(candidates)

#results was zero for 24pos_fullG1_FILidLD

# manhattan plot
pdf("Manhattan_plot.pdf")
plot(p.values, main = "Manhattan plot",
     xlab = "Locus id",
     ylab = "-log10(P-values)",
     cex = .3, col = "grey")
points(candidates, -log10(p.values)[candidates],
       pch = 19, cex = .2, col = "blue")
dev.off()
```