

# Coral Analysis Notes

Jason Chari and Elizabeth Maloney

June 9, 2016

## 1. Read in Data

```
setwd("~/CoralProject")
#setwd("C:/Users/maloneem/Documents/CoralProject")

#Read in Data
filenamesDCM = c("PAL003DCM1.csv", "PAL006DCM1.csv", "PAL007DCM1.csv", "PAL014DCM1.csv", "PAL015DCM1.csv",
  "PAL017DCM1.csv", "PAL018DCM1.csv", "PAL019D3.csv", "PAL029DCM1.csv", "PAL042D2.csv", "PAL043D2.csv",
  "PAL044DCM1.csv", "PAL045DCM1.csv", "PAL048DCM1.csv", "PAL050D1.csv", "PAL052D1.csv",
  "PAL052D2.csv", "PAL130D1.csv", "PAL131D1.csv", "PAL133D1.csv", "PAL134D1.csv", "PAL137DCM1.csv",
  "PAL227D1.csv", "PAL247DCM1.csv", "PAL248D1.csv", "PAL250D1.csv", "PAL250D2.csv", "PAL252DCM1.csv",
  "PAL270D1.csv", "PAL299DCM1.csv", "PAL300D1.csv", "PAL301D1.csv", "PAL302D1.csv", "PAL306D1.csv",
  "PAL332D1.csv", "PAL333D1.csv", "PAL334D1.csv", "PAL340D1.csv", "PAL355D1.csv", "PAL356D1.csv",
  "PAL362D1.csv", "PAL374DCM1.csv", "PAL375D1.csv", "PAL378DCM1.csv", "PAL379DCM1.csv", "PAL382DCM1.csv")

filenamesHEX = c("PAL001HEX1.CSV", "PAL003HEX1.CSV", "PAL007HEX1.CSV", "PAL014HEX1.CSV", "PAL015HEX1.CSV",
  "PAL017HEX1.CSV", "PAL018HEX1.CSV", "PAL019H5.CSV", "PAL029HEX1.CSV", "PAL042H2.CSV", "PAL043H2.CSV",
  "PAL043HEX1.CSV", "PAL044HEX1.CSV", "PAL046HEX1.CSV", "PAL048HEX1.CSV", "PAL050H1.CSV", "PAL052H1.CSV",
  "PAL052H2.CSV", "PAL130H1.CSV", "PAL131H1.CSV", "PAL133H1.CSV", "PAL134H1.CSV", "PAL137HEX1.CSV",
  "PAL227H1.CSV", "PAL247HEX1.CSV", "PAL248H1.CSV", "PAL250H1.CSV", "PAL250H2.CSV", "PAL252HEX1.CSV",
  "PAL270H1.CSV", "PAL299HEX1.CSV", "PAL300H1.CSV", "PAL301H1.CSV", "PAL302H1.CSV", "PAL306H1.CSV", "PAL307HEX1.CSV",
  "PAL332H1.CSV", "PAL333H1.CSV", "PAL334H1.CSV", "PAL340H1.CSV", "PAL355H1.CSV", "PAL356H1.CSV",
  "PAL359H1.CSV", "PAL362H1.CSV", "PAL374HEX1.CSV", "PAL375H1.CSV", "PAL378HEX1.CSV", "PAL379HEX1.CSV", "PAL382HEX1.CSV")

#setwd("2016 Raw Data/All DCM")
setwd("~/CoralProject/2016 Raw Data/ALL DCM")
PAL003DCM1 = read.csv(filenamesDCM[1], header=FALSE)
PAL006DCM1 = read.csv(filenamesDCM[2], header=FALSE)
PAL007DCM1 = read.csv(filenamesDCM[3], header=FALSE)
PAL014DCM1 = read.csv(filenamesDCM[4], header=FALSE)
PAL015DCM1 = read.csv(filenamesDCM[5], header=FALSE)
PAL017DCM1 = read.csv(filenamesDCM[6], header=FALSE)
PAL018DCM1 = read.csv(filenamesDCM[7], header=FALSE)
PAL019DCM1 = read.csv(filenamesDCM[8], header=FALSE)
PAL029DCM1 = read.csv(filenamesDCM[9], header=FALSE)
PAL042DCM1 = read.csv(filenamesDCM[10], header=FALSE)
PAL043DCM1 = read.csv(filenamesDCM[11], header=FALSE)
PAL044DCM1 = read.csv(filenamesDCM[12], header=FALSE)
PAL045DCM1 = read.csv(filenamesDCM[13], header=FALSE)
PAL048DCM1 = read.csv(filenamesDCM[14], header=FALSE)
PAL050DCM1 = read.csv(filenamesDCM[15], header=FALSE)
PAL052DCM1 = read.csv(filenamesDCM[16], header=FALSE)
#PAL052DCM2 = read.csv(filenamesDCM[17], header=FALSE) Broad peak @ 15.8 looked less like rest of data.
PAL130DCM1 = read.csv(filenamesDCM[18], header=FALSE)
PAL131DCM1 = read.csv(filenamesDCM[19], header=FALSE)
```

```

PAL133DCM1 = read.csv(filenamesDCM[20],header=FALSE)
PAL134DCM1 = read.csv(filenamesDCM[21],header=FALSE)
PAL137DCM1 = read.csv(filenamesDCM[22],header=FALSE)
PAL227DCM1 = read.csv(filenamesDCM[23],header=FALSE)
PAL247DCM1 = read.csv(filenamesDCM[24],header=FALSE)
PAL248DCM1 = read.csv(filenamesDCM[25],header=FALSE)
PAL250DCM1 = read.csv(filenamesDCM[26],header=FALSE)
#PAL250D2 = read.csv(filenamesDCM[27],header=FALSE)  Graph 250D2 looked the same as 052D2. Use D1.
PAL252DCM1 = read.csv(filenamesDCM[28],header=FALSE)
PAL270DCM1 = read.csv(filenamesDCM[29],header=FALSE)
PAL299DCM1 = read.csv(filenamesDCM[30],header=FALSE)
PAL300DCM1 = read.csv(filenamesDCM[31],header=FALSE)
PAL301DCM1 = read.csv(filenamesDCM[32],header=FALSE)
PAL302DCM1 = read.csv(filenamesDCM[33],header=FALSE)
PAL306DCM1 = read.csv(filenamesDCM[34],header=FALSE)
PAL307DCM1 = read.csv(filenamesDCM[35],header=FALSE)
PAL332DCM1 = read.csv(filenamesDCM[36],header=FALSE)
PAL333DCM1 = read.csv(filenamesDCM[37],header=FALSE)
PAL334DCM1 = read.csv(filenamesDCM[38],header=FALSE)
PAL340DCM1 = read.csv(filenamesDCM[39],header=FALSE)
PAL355DCM1 = read.csv(filenamesDCM[40],header=FALSE)
PAL356DCM1 = read.csv(filenamesDCM[41],header=FALSE)
PAL359DCM1 = read.csv(filenamesDCM[42],header=FALSE)
PAL362DCM1 = read.csv(filenamesDCM[43],header=FALSE)
PAL374DCM1 = read.csv(filenamesDCM[44],header=FALSE)
PAL375DCM1 = read.csv(filenamesDCM[45],header=FALSE)
PAL378DCM1 = read.csv(filenamesDCM[46],header=FALSE)
PAL379DCM1 = read.csv(filenamesDCM[47],header=FALSE)
PAL382DCM1 = read.csv(filenamesDCM[48],header=FALSE)

#setwd("2016 Raw Data/All HEX")
setwd("~/CoralProject/2016 Raw Data/ALL HEX")
PAL001HEX1 = read.csv(filenamesHEX[1],header=FALSE)
PAL003HEX1 = read.csv(filenamesHEX[2],header=FALSE)
PAL007HEX1 = read.csv(filenamesHEX[3],header=FALSE)
PAL014HEX1 = read.csv(filenamesHEX[4],header=FALSE)
PAL015HEX1 = read.csv(filenamesHEX[5],header=FALSE)
PAL017HEX1 = read.csv(filenamesHEX[6],header=FALSE)
PAL018HEX1 = read.csv(filenamesHEX[7],header=FALSE)
PAL019HEX1 = read.csv(filenamesHEX[8],header=FALSE)
PAL029HEX1 = read.csv(filenamesHEX[9],header=FALSE)
PAL042HEX1 = read.csv(filenamesHEX[10],header=FALSE)
PAL043HEX1 = read.csv(filenamesHEX[11],header=FALSE)
PAL044HEX1 = read.csv(filenamesHEX[12],header=FALSE)
PAL046HEX1 = read.csv(filenamesHEX[13],header=FALSE)
PAL130HEX1 = read.csv(filenamesHEX[14],header=FALSE)
PAL131HEX1 = read.csv(filenamesHEX[15],header=FALSE)
PAL133HEX1 = read.csv(filenamesHEX[16],header=FALSE)
PAL250HEX1 = read.csv(filenamesHEX[17],header=FALSE)
PAL250HEX2 = read.csv(filenamesHEX[18],header=FALSE)
PAL252HEX1 = read.csv(filenamesHEX[19],header=FALSE)
PAL270HEX1 = read.csv(filenamesHEX[20],header=FALSE)
PAL299HEX1 = read.csv(filenamesHEX[21],header=FALSE)

```

```

PAL332HEX1 = read.csv(filenamesHEX[22],header=FALSE)
PAL333HEX1 = read.csv(filenamesHEX[23],header=FALSE)
PAL334HEX1 = read.csv(filenamesHEX[24],header=FALSE)
PAL340HEX1 = read.csv(filenamesHEX[25],header=FALSE)
PAL355HEX1 = read.csv(filenamesHEX[26],header=FALSE)
PAL356HEX1 = read.csv(filenamesHEX[27],header=FALSE)
PAL359HEX1 = read.csv(filenamesHEX[28],header=FALSE)
PAL362HEX1 = read.csv(filenamesHEX[29],header=FALSE)
PAL374HEX1 = read.csv(filenamesHEX[30],header=FALSE)
PAL375HEX1 = read.csv(filenamesHEX[31],header=FALSE)
PAL378HEX1 = read.csv(filenamesHEX[32],header=FALSE)
PAL379HEX1 = read.csv(filenamesHEX[33],header=FALSE)
PAL382HEX1 = read.csv(filenamesHEX[34],header=FALSE)

```

```

plot(PAL250HEX1, type="l") plot(PAL250HEX2, type="l") plot(PAL042HEX1, type="l") plot(PAL019HEX1,
type="l")

```

## 2. Use trap function to generate trapezoidal area

```

DELTA = 0.000333333333333333 #Time increment by which intensity is measured (in minutes)

trap <- function(intens, delta) {
  intens1 <- rep(intens)
  intens1 = intens1[-1]
  intens = intens[1:length(intens)-1]
  trapArea = (intens + intens1)*delta/2
  return(c(as.numeric(trapArea), 0))
}

norm <- function(trapArea){
  tot = sum(trapArea)
  return(trapArea/tot)
}

```

## 3. Generate dataframe with normalized trapezoid areas for all time points as well as the sample number, separating solvent ("greasiness"), and clade.

```

#Process as dataframe - CHANGE TO AS.NUMERIC
#truncate
TRUNC_BEGIN = 15000
TRUNC_END = 62004
#We omit the first and last 15000 observations (5 minutes) because there is little observed activity du

dataframe = rbind(
  #S. glaucum clade F
  c(003,"DCM","F", norm(trap(as.numeric(PAL003DCM1[TRUNC_BEGIN:TRUNC_END, 2]), DELTA))),
  c(006,"DCM","F", norm(trap(as.numeric(PAL006DCM1[TRUNC_BEGIN:TRUNC_END, 2]), DELTA))),
  c(007,"DCM","F", norm(trap(as.numeric(PAL007DCM1[TRUNC_BEGIN:TRUNC_END, 2]), DELTA))),
  c(014,"DCM","F", norm(trap(as.numeric(PAL014DCM1[TRUNC_BEGIN:TRUNC_END, 2]), DELTA))),
  c(015,"DCM","F", norm(trap(as.numeric(PAL015DCM1[TRUNC_BEGIN:TRUNC_END, 2]), DELTA))),
  c(017,"DCM","F", norm(trap(as.numeric(PAL017DCM1[TRUNC_BEGIN:TRUNC_END, 2]), DELTA))),
  c(018,"DCM","F", norm(trap(as.numeric(PAL018DCM1[TRUNC_BEGIN:TRUNC_END, 2]), DELTA))),

```

```

c(043,"DCM","F", norm(trap(as.numeric(PAL043DCM1[TRUNC_BEGIN:TRUNC_END, 2]), DELTA))),
c(044,"DCM","F", norm(trap(as.numeric(PAL044DCM1[TRUNC_BEGIN:TRUNC_END, 2]), DELTA))),
c(048,"DCM","F", norm(trap(as.numeric(PAL048DCM1[TRUNC_BEGIN:TRUNC_END, 2]), DELTA))),
c(050,"DCM","F", norm(trap(as.numeric(PAL050DCM1[TRUNC_BEGIN:TRUNC_END, 2]), DELTA))),
c(130,"DCM","F", norm(trap(as.numeric(PAL130DCM1[TRUNC_BEGIN:TRUNC_END, 2]), DELTA))),
c(131,"DCM","F", norm(trap(as.numeric(PAL131DCM1[TRUNC_BEGIN:TRUNC_END, 2]), DELTA))),
c(133,"DCM","F", norm(trap(as.numeric(PAL133DCM1[TRUNC_BEGIN:TRUNC_END, 2]), DELTA))),
c(134,"DCM","F", norm(trap(as.numeric(PAL134DCM1[TRUNC_BEGIN:TRUNC_END, 2]), DELTA))),
c(248,"DCM","F", norm(trap(as.numeric(PAL248DCM1[TRUNC_BEGIN:TRUNC_END, 2]), DELTA))),
c(270,"DCM","F", norm(trap(as.numeric(PAL270DCM1[TRUNC_BEGIN:TRUNC_END, 2]), DELTA))),
c(302,"DCM","F", norm(trap(as.numeric(PAL302DCM1[TRUNC_BEGIN:TRUNC_END, 2]), DELTA))),
c(333,"DCM","F", norm(trap(as.numeric(PAL333DCM1[TRUNC_BEGIN:TRUNC_END, 2]), DELTA))),
c(334,"DCM","F", norm(trap(as.numeric(PAL334DCM1[TRUNC_BEGIN:TRUNC_END, 2]), DELTA))),
c(374,"DCM","F", norm(trap(as.numeric(PAL374DCM1[TRUNC_BEGIN:TRUNC_END, 2]), DELTA))),
c(378,"DCM","F", norm(trap(as.numeric(PAL378DCM1[TRUNC_BEGIN:TRUNC_END, 2]), DELTA))),
c(382,"DCM","F", norm(trap(as.numeric(PAL382DCM1[TRUNC_BEGIN:TRUNC_END, 2]), DELTA))),
c(227,"DCM","F", norm(trap(as.numeric(PAL227DCM1[TRUNC_BEGIN:TRUNC_END, 2]), DELTA))),

#S. glaucum clade D
c(029,"DCM","D", norm(trap(as.numeric(PAL029DCM1[TRUNC_BEGIN:TRUNC_END, 2]), DELTA))),
c(247,"DCM","D", norm(trap(as.numeric(PAL247DCM1[TRUNC_BEGIN:TRUNC_END, 2]), DELTA))),
c(252,"DCM","D", norm(trap(as.numeric(PAL252DCM1[TRUNC_BEGIN:TRUNC_END, 2]), DELTA))),
c(299,"DCM","D", norm(trap(as.numeric(PAL299DCM1[TRUNC_BEGIN:TRUNC_END, 2]), DELTA))),
c(300,"DCM","D", norm(trap(as.numeric(PAL300DCM1[TRUNC_BEGIN:TRUNC_END, 2]), DELTA))),
c(306,"DCM","D", norm(trap(as.numeric(PAL306DCM1[TRUNC_BEGIN:TRUNC_END, 2]), DELTA))),
c(332,"DCM","D", norm(trap(as.numeric(PAL332DCM1[TRUNC_BEGIN:TRUNC_END, 2]), DELTA))),
c(340,"DCM","D", norm(trap(as.numeric(PAL340DCM1[TRUNC_BEGIN:TRUNC_END, 2]), DELTA))),
c(355,"DCM","D", norm(trap(as.numeric(PAL355DCM1[TRUNC_BEGIN:TRUNC_END, 2]), DELTA))),
c(359,"DCM","D", norm(trap(as.numeric(PAL359DCM1[TRUNC_BEGIN:TRUNC_END, 2]), DELTA))),

#Trocheliophorum
c(019,"DCM","T", norm(trap(as.numeric(PAL019DCM1[TRUNC_BEGIN:TRUNC_END, 2]), DELTA))),
c(042,"DCM","T", norm(trap(as.numeric(PAL042DCM1[TRUNC_BEGIN:TRUNC_END, 2]), DELTA))),
c(045,"DCM","T", norm(trap(as.numeric(PAL045DCM1[TRUNC_BEGIN:TRUNC_END, 2]), DELTA))),
c(052,"DCM","T", norm(trap(as.numeric(PAL052DCM1[TRUNC_BEGIN:TRUNC_END, 2]), DELTA))),
c(137,"DCM","T", norm(trap(as.numeric(PAL137DCM1[TRUNC_BEGIN:TRUNC_END, 2]), DELTA))),
c(250,"DCM","T", norm(trap(as.numeric(PAL250DCM1[TRUNC_BEGIN:TRUNC_END, 2]), DELTA))),
c(301,"DCM","T", norm(trap(as.numeric(PAL301DCM1[TRUNC_BEGIN:TRUNC_END, 2]), DELTA))),
c(307,"DCM","T", norm(trap(as.numeric(PAL307DCM1[TRUNC_BEGIN:TRUNC_END, 2]), DELTA))),
c(356,"DCM","T", norm(trap(as.numeric(PAL356DCM1[TRUNC_BEGIN:TRUNC_END, 2]), DELTA))),
c(362,"DCM","T", norm(trap(as.numeric(PAL362DCM1[TRUNC_BEGIN:TRUNC_END, 2]), DELTA))),
c(375,"DCM","T", norm(trap(as.numeric(PAL375DCM1[TRUNC_BEGIN:TRUNC_END, 2]), DELTA))),
c(379,"DCM","T", norm(trap(as.numeric(PAL379DCM1[TRUNC_BEGIN:TRUNC_END, 2]), DELTA)))
)

dataframeHEX = rbind(
  c(001, "HEX", "F", norm(trap(as.numeric(PAL001HEX1[TRUNC_BEGIN:TRUNC_END, 2]), DELTA))),
  c(003, "HEX", "F", norm(trap(as.numeric(PAL003HEX1[TRUNC_BEGIN:TRUNC_END, 2]), DELTA))),
  c(007, "HEX", "F", norm(trap(as.numeric(PAL007HEX1[TRUNC_BEGIN:TRUNC_END, 2]), DELTA))),
  c(014, "HEX", "F", norm(trap(as.numeric(PAL014HEX1[TRUNC_BEGIN:TRUNC_END, 2]), DELTA))),
  c(015, "HEX", "F", norm(trap(as.numeric(PAL015HEX1[TRUNC_BEGIN:TRUNC_END, 2]), DELTA))),
  c(017, "HEX", "F", norm(trap(as.numeric(PAL017HEX1[TRUNC_BEGIN:TRUNC_END, 2]), DELTA))),
  c(018, "HEX", "F", norm(trap(as.numeric(PAL018HEX1[TRUNC_BEGIN:TRUNC_END, 2]), DELTA))),

```

```

c(019, "HEX", "T", norm(trap(as.numeric(PAL019HEX1[TRUNC_BEGIN:TRUNC_END, 2]), DELTA))),
c(029, "HEX", "D", norm(trap(as.numeric(PAL029HEX1[TRUNC_BEGIN:TRUNC_END, 2]), DELTA))),
c(042, "HEX", "T", norm(trap(as.numeric(PAL042HEX1[TRUNC_BEGIN:TRUNC_END, 2]), DELTA))),
c(043, "HEX", "F", norm(trap(as.numeric(PAL043HEX1[TRUNC_BEGIN:TRUNC_END, 2]), DELTA))),
c(044, "HEX", "F", norm(trap(as.numeric(PAL044HEX1[TRUNC_BEGIN:TRUNC_END, 2]), DELTA))),
c(046, "HEX", "D", norm(trap(as.numeric(PAL046HEX1[TRUNC_BEGIN:TRUNC_END, 2]), DELTA))),
c(130, "HEX", "F", norm(trap(as.numeric(PAL130HEX1[TRUNC_BEGIN:TRUNC_END, 2]), DELTA))),
c(131, "HEX", "F", norm(trap(as.numeric(PAL131HEX1[TRUNC_BEGIN:TRUNC_END, 2]), DELTA))),
c(133, "HEX", "F", norm(trap(as.numeric(PAL133HEX1[TRUNC_BEGIN:TRUNC_END, 2]), DELTA))),
c(250, "HEX", "T", norm(trap(as.numeric(PAL250HEX1[TRUNC_BEGIN:TRUNC_END, 2]), DELTA))),
c(250, "HEX", "T", norm(trap(as.numeric(PAL250HEX2[TRUNC_BEGIN:TRUNC_END, 2]), DELTA))),
c(252, "HEX", "D", norm(trap(as.numeric(PAL252HEX1[TRUNC_BEGIN:TRUNC_END, 2]), DELTA))),
c(270, "HEX", "F", norm(trap(as.numeric(PAL270HEX1[TRUNC_BEGIN:TRUNC_END, 2]), DELTA))),
c(299, "HEX", "D", norm(trap(as.numeric(PAL299HEX1[TRUNC_BEGIN:TRUNC_END, 2]), DELTA))),
c(332, "HEX", "D", norm(trap(as.numeric(PAL332HEX1[TRUNC_BEGIN:TRUNC_END, 2]), DELTA))),
c(333, "HEX", "F", norm(trap(as.numeric(PAL333HEX1[TRUNC_BEGIN:TRUNC_END, 2]), DELTA))),
c(334, "HEX", "F", norm(trap(as.numeric(PAL334HEX1[TRUNC_BEGIN:TRUNC_END, 2]), DELTA))),
c(340, "HEX", "D", norm(trap(as.numeric(PAL340HEX1[TRUNC_BEGIN:TRUNC_END, 2]), DELTA))),
c(355, "HEX", "D", norm(trap(as.numeric(PAL355HEX1[TRUNC_BEGIN:TRUNC_END, 2]), DELTA))),
c(356, "HEX", "T", norm(trap(as.numeric(PAL356HEX1[TRUNC_BEGIN:TRUNC_END, 2]), DELTA))),
c(359, "HEX", "D", norm(trap(as.numeric(PAL359HEX1[TRUNC_BEGIN:TRUNC_END, 2]), DELTA))),
c(362, "HEX", "T", norm(trap(as.numeric(PAL362HEX1[TRUNC_BEGIN:TRUNC_END, 2]), DELTA))),
c(374, "HEX", "F", norm(trap(as.numeric(PAL374HEX1[TRUNC_BEGIN:TRUNC_END, 2]), DELTA))),
c(375, "HEX", "T", norm(trap(as.numeric(PAL375HEX1[TRUNC_BEGIN:TRUNC_END, 2]), DELTA))),
c(378, "HEX", "F", norm(trap(as.numeric(PAL378HEX1[TRUNC_BEGIN:TRUNC_END, 2]), DELTA))),
c(379, "HEX", "T", norm(trap(as.numeric(PAL379HEX1[TRUNC_BEGIN:TRUNC_END, 2]), DELTA))),
c(382, "HEX", "F", norm(trap(as.numeric(PAL382HEX1[TRUNC_BEGIN:TRUNC_END, 2]), DELTA)))
)
#generate column name vector for the dataframe
names <- function(){
  nameVec = c(1:47005)
  nameVec = as.character(nameVec)

  for(i in 1:47005){
    nameVec[i] = paste("t", nameVec[i], sep="")
  }

  nameVec[1:5]
  nameVec = c("Sample", "Greasiness", "Clade", nameVec)
  return(nameVec)
}
colnames(dataframeHEX) <- names()

```

Snapshot of dataframe:

```
install.packages("stargazer", repos='http://cran.us.r-project.org')
```

```
##
## The downloaded binary packages are in
## /var/folders/7k/8ngqf1rd4ggc6b8q6cflfr6m0000gn/T//RtmpwHvQ8q/downloaded_packages

```



```
library(stargazer)
```

```
##  
## Please cite as:  
##  
## Hlavac, Marek (2015). stargazer: Well-Formatted Regression and Summary Statistics Tables.  
## R package version 5.2. http://CRAN.R-project.org/package=stargazer
```

```
stargazer(dataframeHEX[1:5,1:5])
```

```
##  
## % Table created by stargazer v.5.2 by Marek Hlavac, Harvard University. E-mail: hlavac at fas.harvard.edu  
## % Date and time: Sat, Jun 18, 2016 - 16:44:08  
## \begin{table}[!htbp] \centering  
## \caption{}  
## \label{}  
## \begin{tabular}{@{\extracolsep{5pt}} ccccc}  
## \hline  
## \hline \hline  
## Sample & Greasiness & Clade & t1 & t2 \\  
## \hline \hline  
## 1 & HEX & F & 1.86804766311945e-05 & 1.86352370698324e-05 \\  
## 3 & HEX & F & 1.86917340963497e-05 & 1.86743111747468e-05 \\  
## 7 & HEX & F & 1.84542659086005e-05 & 1.84502834959617e-05 \\  
## 14 & HEX & F & 1.83589698302085e-05 & 1.83434877880877e-05 \\  
## 15 & HEX & F & 1.85682863039704e-05 & 1.85228440803128e-05 \\  
## \hline \hline  
## \end{tabular}  
## \end{table}
```

3. Create a bin function to partition the data by summing across equal time intervals.

```
binner <- function(df, binWidth, delta) {  
  rowz = c()  
  for(j in 1:nrow(df)) {  
    col = c()  
    row_normed = as.numeric(df[j,-(1:3)])  
    for(i in seq(4, ncol(df)-binWidth, binWidth)){  
      end = i + binWidth - 1  
      col = c(col, sum(as.numeric(df[j,i:end])))  
    }  
    rowz = rbind(rowz, col)  
  }  
  return(rowz)  
}
```

```
binned_matrix_1100 = binner(dataframe, 1100, DELTA)  
binned_matrix_2000 = binner(dataframe, 2000, DELTA)  
binned_matrix_3000 = binner(dataframe, 3000, DELTA)
```

```

binned_matrixHEX_1100 = binner(dataframeHEX, 1100, DELTA)
binned_matrixHEX_2000 = binner(dataframeHEX, 2000, DELTA)
binned_matrixHEX_3000 = binner(dataframeHEX, 3000, DELTA)

```

#### 4. PCA - 22 Second Bin Width (1100 time points)

*#DCM:*

```

trapAreas.centered = scale(binned_matrix_1100, center = TRUE, scale = FALSE)
pca.result <- prcomp(trapAreas.centered)
str(pca.result)

```

```

## List of 5
## $ sdev      : num [1:42] 0.01611 0.01261 0.00898 0.00749 0.00634 ...
## $ rotation: num [1:42, 1:42] -0.0346 -0.0367 -0.0264 -0.0336 -0.0341 ...
## ..- attr(*, "dimnames")=List of 2
## .. ..$ : NULL
## .. ..$ : chr [1:42] "PC1" "PC2" "PC3" "PC4" ...
## $ center   : num [1:42] 1.28e-18 -8.30e-19 1.21e-18 -1.06e-18 -1.51e-19 ...
## $ scale    : logi FALSE
## $ x        : num [1:46, 1:42] 0.0016 0.04127 -0.00686 -0.01217 -0.00156 ...
## ..- attr(*, "dimnames")=List of 2
## .. ..$ : chr [1:46] "col" "col" "col" "col" ...
## .. ..$ : chr [1:42] "PC1" "PC2" "PC3" "PC4" ...
## - attr(*, "class")= chr "prcomp"

```

```
PC = pca.result$x
```

```

F = 1:24
D = 25:34
T = 35:46

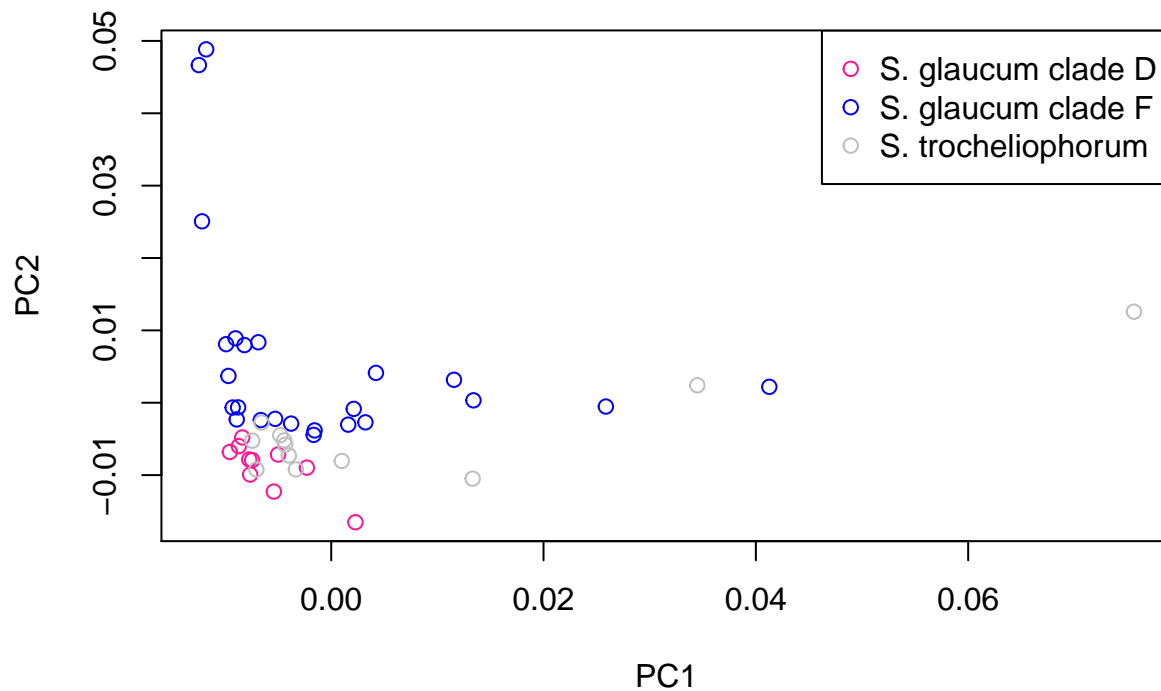
```

```

plot(PC[,1],PC[,2],col="white",main = "PCA - 22 Second Bin Width",xlab="PC1",ylab="PC2")
points(PC[F,1],PC[F,2],col="blue")
points(PC[D,1],PC[D,2],col="deeppink")
points(PC[T,1],PC[T,2],col="grey")
legend("topright",legend=c("S. glaucum clade D","S. glaucum clade F","S. trocheliophorum"),col=c("deeppink","blue","grey"))

```

## PCA – 22 Second Bin Width



```
correlation_1100 = cor(PC,binned_matrix_1100)
install.packages("corrplot", repos='http://cran.us.r-project.org')
```

```
##
## The downloaded binary packages are in
## /var/folders/7k/8ngqf1rd4ggc6b8q6cflfr6m0000gn/T//RtmpwHvQ8q/downloaded_packages
```

```
library(corrplot)
```

```
## Warning: package 'corrplot' was built under R version 3.2.5
```

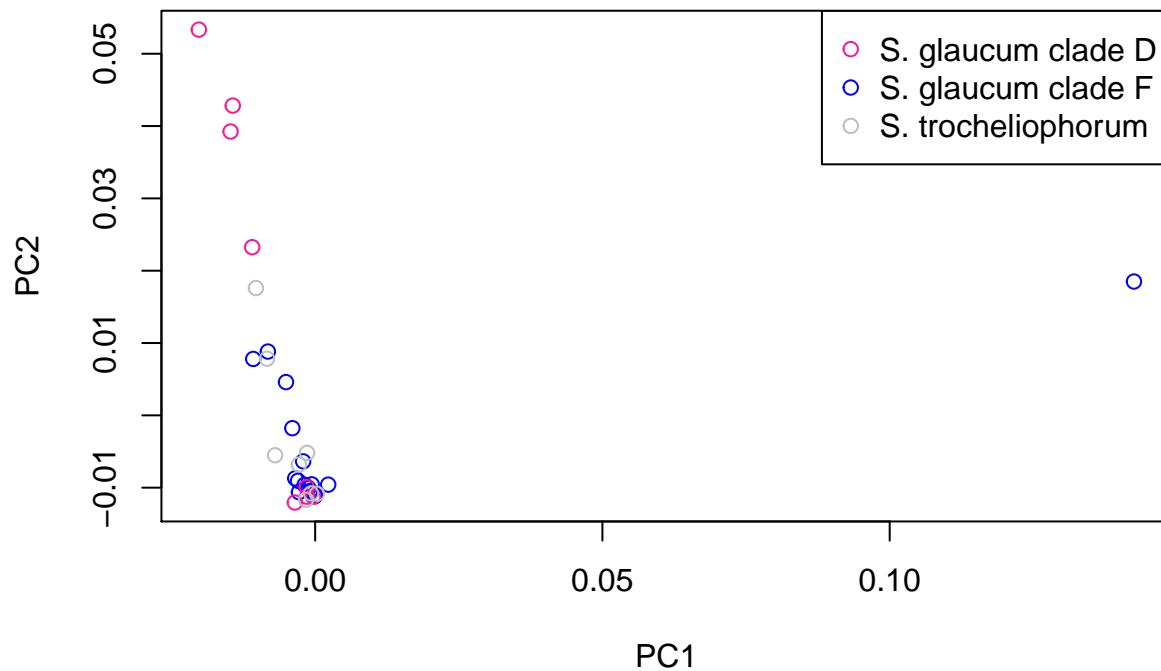
```
corrplot(t(correlation_1100), method="color")
```





```
plot(PC_HEX[,1],PC_HEX[,2],col="white",main = "PCA - 22 Second Bin Width",xlab="PC1",ylab="PC2")
points(PC_HEX[F,1],PC_HEX[F,2],col="blue")
points(PC_HEX[D,1],PC_HEX[D,2],col="deeppink")
points(PC_HEX[T,1],PC_HEX[T,2],col="grey")
legend("topright",legend=c("S. glaucum clade D","S. glaucum clade F","S. trocheliophorum"),col=c("deeppink","blue","grey"))
```

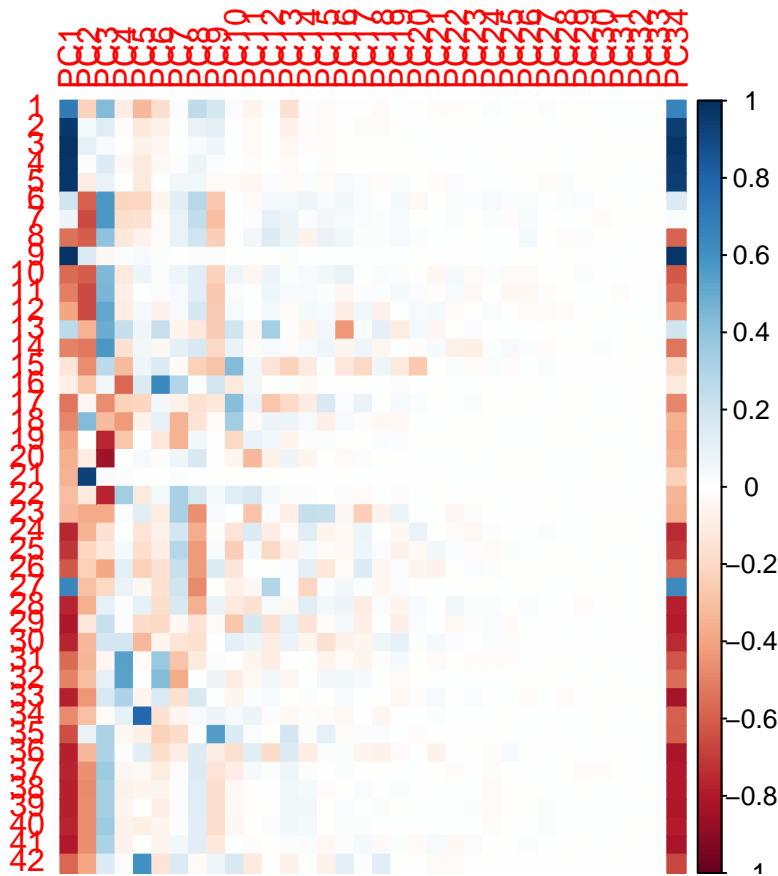
## PCA – 22 Second Bin Width



```
correlation_1100 = cor(PC_HEX,binned_matrixHEX_1100)
install.packages("corrplot", repos='http://cran.us.r-project.org')
```

```
##
## The downloaded binary packages are in
## /var/folders/7k/8ngqf1rd4ggc6b8q6cflfr6m0000gn/T//RtmpwHvQ8q/downloaded_packages
```

```
library(corrplot)
corrplot(t(correlation_1100), method="color")
```



## 5. PCA - 40 Second Bin Width (2000 time points)

```
#DCM:
trapAreas.centered = scale(binned_matrix_2000, center = TRUE, scale = FALSE)
pca.result <- prcomp(trapAreas.centered)
str(pca.result)

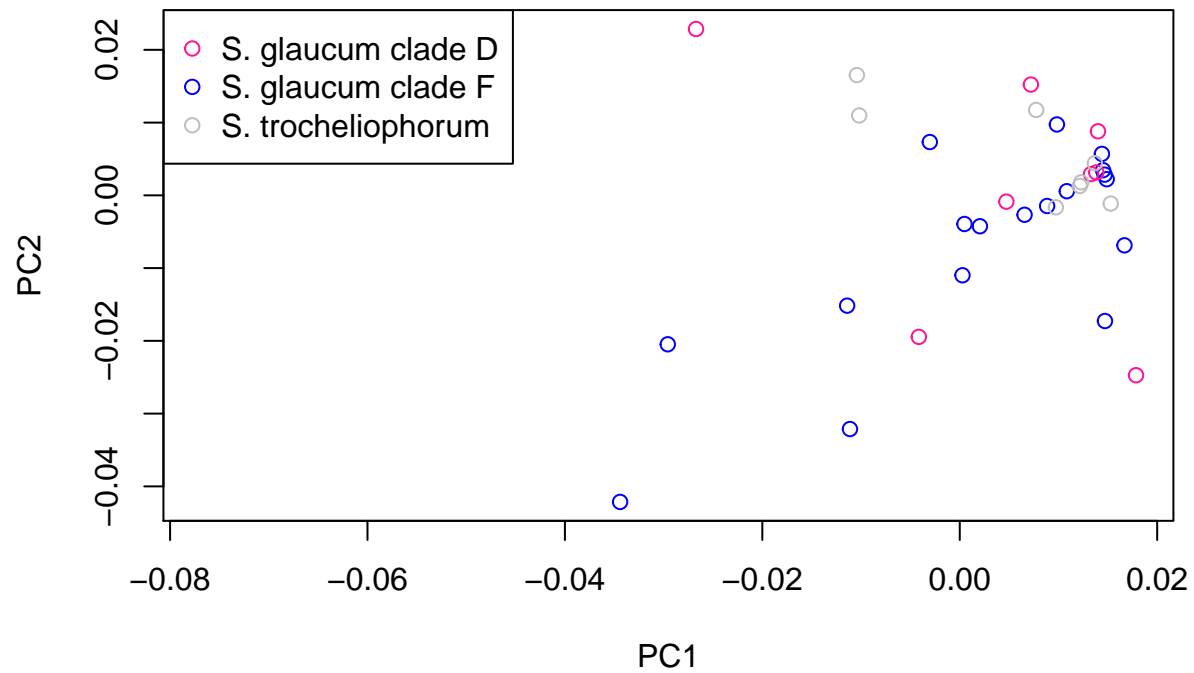
## List of 5
## $ sdev      : num [1:23] 0.01898 0.01262 0.01193 0.00775 0.00694 ...
## $ rotation: num [1:23, 1:23] 0.0648 0.0567 0.0625 0.063 0.0584 ...
## .. attr(*, "dimnames")=List of 2
## .. ..$ : NULL
## .. ..$ : chr [1:23] "PC1" "PC2" "PC3" "PC4" ...
## $ center   : num [1:23] -1.06e-18 -2.87e-18 -2.26e-18 2.72e-18 1.06e-18 ...
## $ scale    : logi FALSE
## $ x        : num [1:46, 1:23] 0.00204 -0.03441 0.01085 0.01669 0.00656 ...
## .. attr(*, "dimnames")=List of 2
## .. ..$ : chr [1:46] "col" "col" "col" "col" ...
## .. ..$ : chr [1:23] "PC1" "PC2" "PC3" "PC4" ...
## - attr(*, "class")= chr "prcomp"

PC = pca.result$x

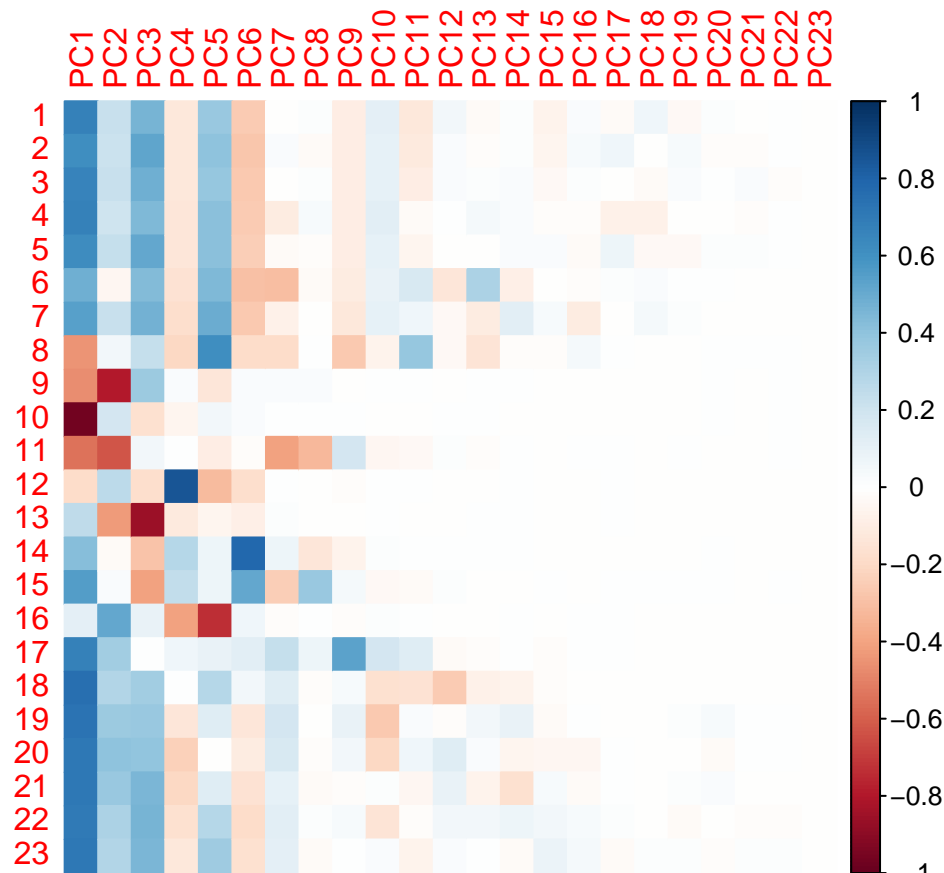
plot(PC[,1],PC[,2],col="white",main = "PCA - 40 Second Bin Width",xlab="PC1",ylab="PC2")
points(PC[F,1],PC[F,2],col="blue")
points(PC[D,1],PC[D,2],col="deeppink")
```

```
points(PC[T,1],PC[T,2],col="grey")
legend("topleft",legend=c("S. glaucum clade D","S. glaucum clade F","S. trocheliophorum"),col=c("deeppink","blue","grey"))
```

## PCA – 40 Second Bin Width



```
correlation_2000 = cor(PC,binned_matrix_2000)
library(corrplot)
corrplot(t(correlation_2000), method="color")
```



```
par(mfrow=c(1,1))
```

```
#HEX:
```

```
trapAreas.centered_HEX = scale(binned_matrixHEX_2000, center = TRUE, scale = FALSE)
```

```
pca.result_HEX <- prcomp(trapAreas.centered_HEX)
```

```
str(pca.result_HEX)
```

```
## List of 5
```

```
## $ sdev      : num [1:23] 0.02716 0.0174 0.00747 0.00461 0.00412 ...
```

```
## $ rotation: num [1:23, 1:23] -0.210685 -0.407663 -0.14587 -0.000943 -0.75727 ...
```

```
## .. attr(*, "dimnames")=List of 2
```

```
## .. ..$ : NULL
```

```
## .. ..$ : chr [1:23] "PC1" "PC2" "PC3" "PC4" ...
```

```
## $ center   : num [1:23] -2.76e-18 -1.43e-18 3.27e-18 -3.06e-18 2.24e-18 ...
```

```
## $ scale    : logi FALSE
```

```
## $ x        : num [1:34, 1:23] -8.85e-04 -4.66e-04 -4.37e-05 1.13e-03 4.28e-03 ...
```

```
## .. attr(*, "dimnames")=List of 2
```

```
## .. ..$ : chr [1:34] "col" "col" "col" "col" ...
```

```
## .. ..$ : chr [1:23] "PC1" "PC2" "PC3" "PC4" ...
```

```
## - attr(*, "class")= chr "prcomp"
```

```
PC_HEX = pca.result_HEX$x
```

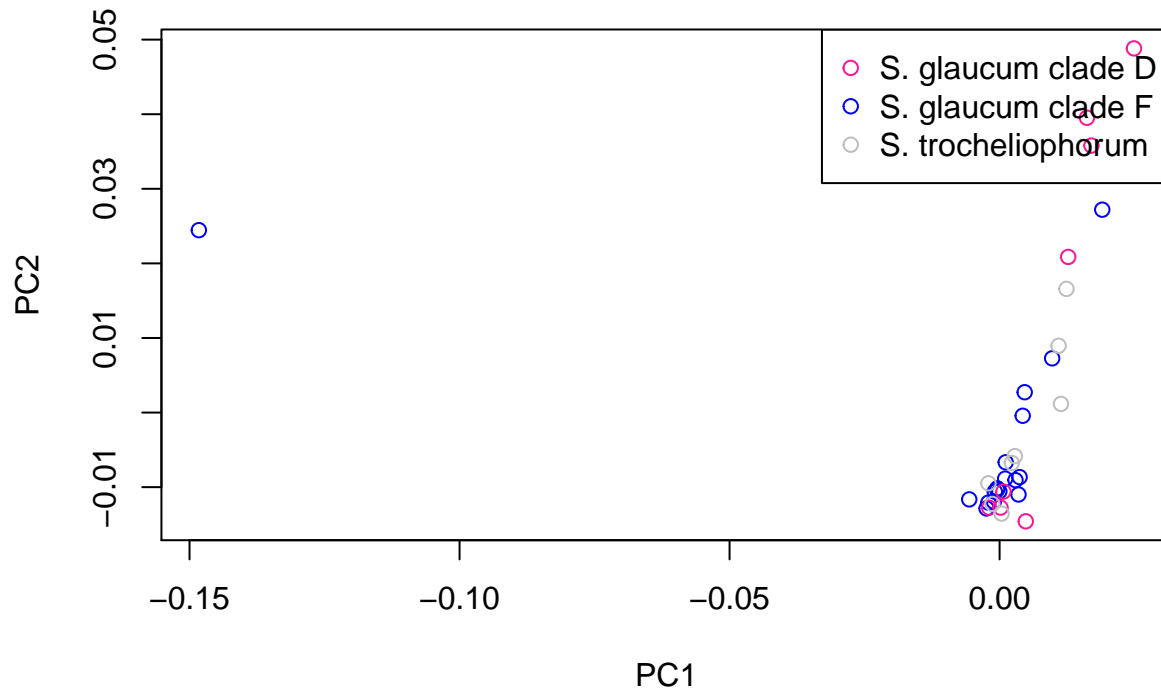
```
F = as.numeric(which(dataframeHEX[,3]=="F"))
```

```
D = as.numeric(which(dataframeHEX[,3]=="D"))
```

```
T = as.numeric(which(dataframeHEX[,3]=="T"))

plot(PC_HEX[,1],PC_HEX[,2],col="white",main = "PCA - 40 Second Bin Width",xlab="PC1",ylab="PC2")
points(PC_HEX[F,1],PC_HEX[F,2],col="blue")
points(PC_HEX[D,1],PC_HEX[D,2],col="deeppink")
points(PC_HEX[T,1],PC_HEX[T,2],col="grey")
legend("topright",legend=c("S. glaucum clade D","S. glaucum clade F","S. trocheliophorum"),col=c("deeppink","blue","grey"))
```

## PCA – 40 Second Bin Width

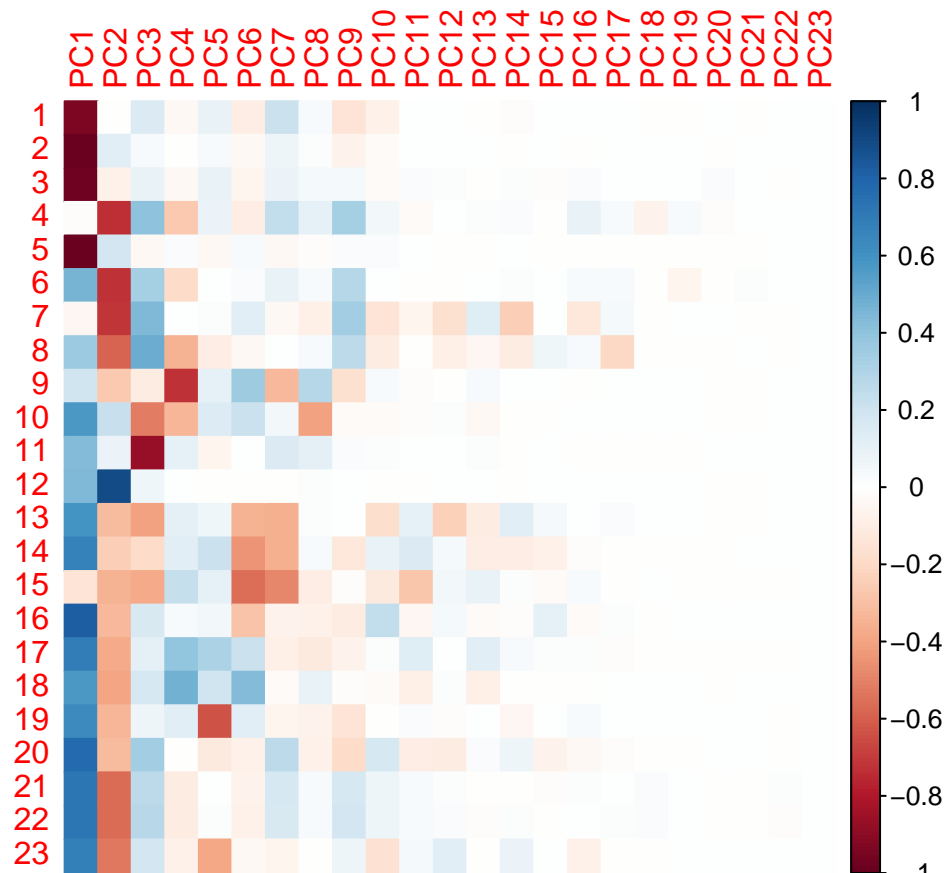


```
correlation_2000 = cor(PC_HEX,binned_matrixHEX_2000)
install.packages("corrplot", repos='http://cran.us.r-project.org')
```

```
##
## The downloaded binary packages are in
## /var/folders/7k/8ngqf1rd4ggc6b8q6cflfr6m0000gn/T//RtmpwHvQ8q/downloaded_packages
```

```
library(corrplot)
corrplot(t(correlation_2000), method="color")
```





## 6. PCA - 60 Second Bin Width (3000 time points)

```
#DCM:
trapAreas.centered = scale(binned_matrix_3000, center = TRUE, scale = FALSE)
pca.result <- prcomp(trapAreas.centered)
str(pca.result)

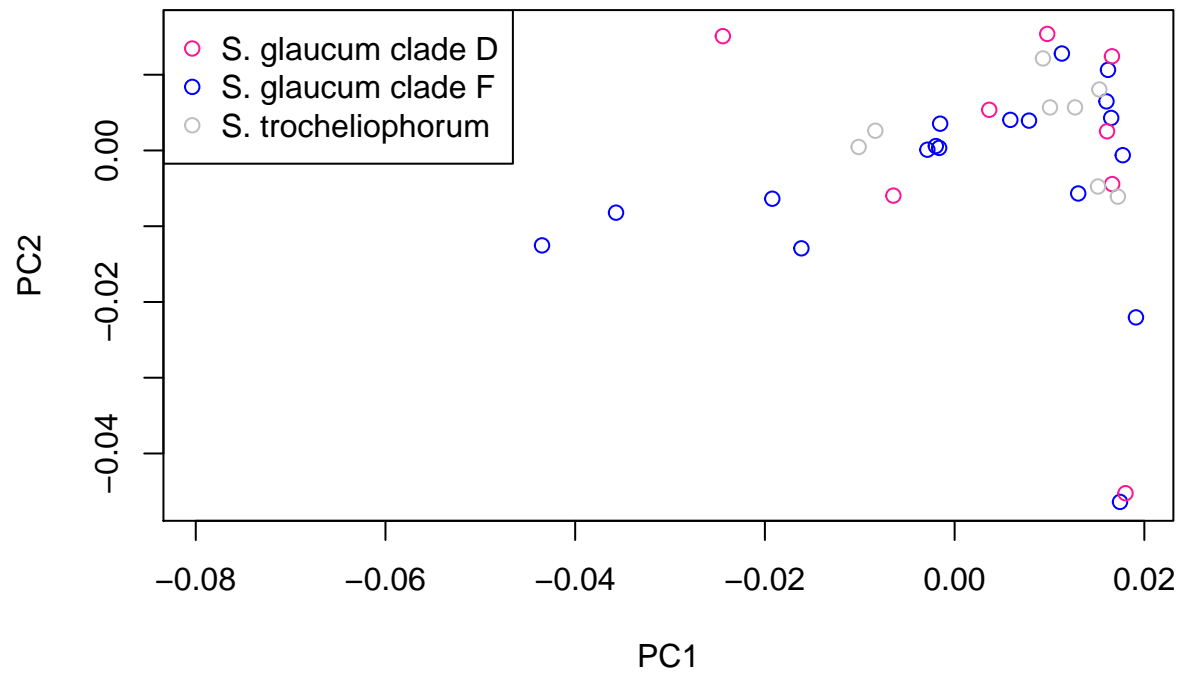
## List of 5
## $ sdev      : num [1:15] 0.02072 0.01251 0.01205 0.008 0.00701 ...
## $ rotation: num [1:15, 1:15] 0.0848 0.0812 0.0851 0.0568 0.0363 ...
## .. attr(*, "dimnames")=List of 2
## .. ..$ : NULL
## .. ..$ : chr [1:15] "PC1" "PC2" "PC3" "PC4" ...
## $ center   : num [1:15] -1.51e-18 -3.02e-18 -3.02e-18 1.06e-18 3.32e-18 ...
## $ scale    : logi FALSE
## $ x        : num [1:46, 1:15] -0.00163 -0.0435 0.01303 0.01912 0.00587 ...
## .. attr(*, "dimnames")=List of 2
## .. ..$ : chr [1:46] "col" "col" "col" "col" ...
## .. ..$ : chr [1:15] "PC1" "PC2" "PC3" "PC4" ...
## - attr(*, "class")= chr "prcomp"

PC = pca.result$x

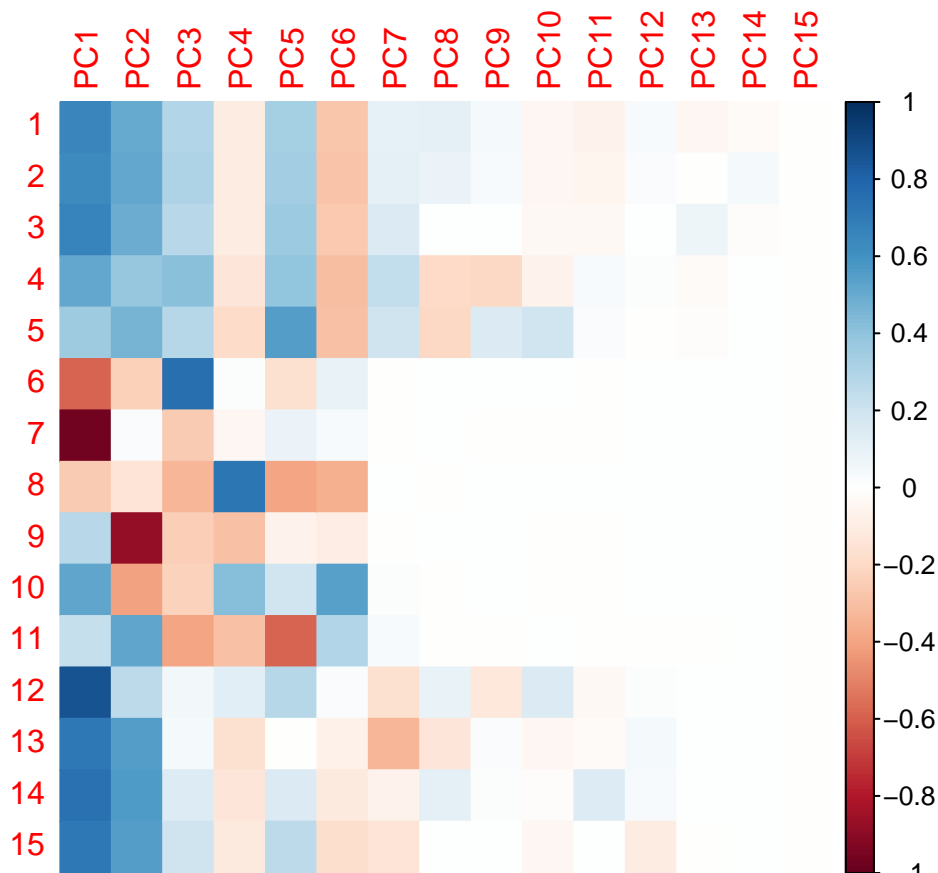
plot(PC[,1],PC[,2],col="white",main = "PCA - 60 Second Bin Width",xlab="PC1",ylab="PC2")
points(PC[F,1],PC[F,2],col="blue")
points(PC[D,1],PC[D,2],col="deeppink")
```

```
points(PC[T,1],PC[T,2],col="grey")
legend("topleft",legend=c("S. glaucum clade D","S. glaucum clade F","S. trocheliophorum"),col=c("deeppink","blue","grey"))
```

## PCA – 60 Second Bin Width



```
correlation_3000 = cor(PC,binned_matrix_3000)
require(corrplot)
corrplot(t(correlation_3000), method="color")
```



```
par(mfrow=c(1,1))

#HEX:
trapAreas.centered_HEX = scale(binned_matrixHEX_3000, center = TRUE, scale = FALSE)
pca.result_HEX <- prcomp(trapAreas.centered_HEX)
str(pca.result_HEX)
```

```
## List of 5
## $ sdev      : num [1:15] 0.02616 0.01726 0.00686 0.00518 0.00472 ...
## $ rotation: num [1:15, 1:15] -0.51562 -0.26542 -0.20081 -0.52767 -0.00409 ...
## .. attr(*, "dimnames")=List of 2
## .. ..$ : NULL
## .. ..$ : chr [1:15] "PC1" "PC2" "PC3" "PC4" ...
## $ center    : num [1:15] 4.08e-19 -6.12e-19 0.00 2.45e-18 1.02e-18 ...
## $ scale     : logi FALSE
## $ x         : num [1:34, 1:15] -0.0047 -0.00449 -0.00364 -0.00124 0.00463 ...
## .. attr(*, "dimnames")=List of 2
## .. ..$ : chr [1:34] "col" "col" "col" "col" ...
## .. ..$ : chr [1:15] "PC1" "PC2" "PC3" "PC4" ...
## - attr(*, "class")= chr "prcomp"
```

```
PC_HEX = pca.result_HEX$x
```

```
F = as.numeric(which(dataframeHEX[,3]=="F"))
D = as.numeric(which(dataframeHEX[,3]=="D"))
```

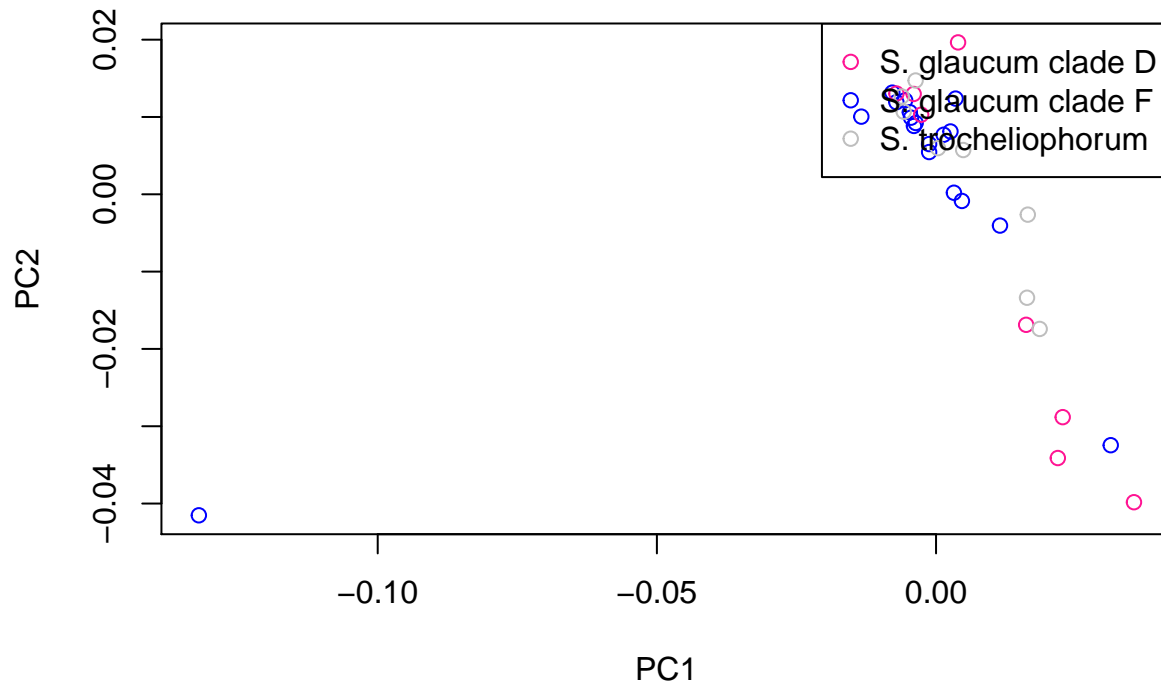
```

T = as.numeric(which(dataframeHEX[,3]=="T"))

plot(PC_HEX[,1],PC_HEX[,2],col="white",main = "PCA - 60 Second Bin Width",xlab="PC1",ylab="PC2")
points(PC_HEX[F,1],PC_HEX[F,2],col="blue")
points(PC_HEX[D,1],PC_HEX[D,2],col="deeppink")
points(PC_HEX[T,1],PC_HEX[T,2],col="grey")
legend("topright",legend=c("S. glaucum clade D","S. glaucum clade F","S. trocheliophorum"),col=c("deeppink","blue","grey"))

```

## PCA – 60 Second Bin Width



```

correlation_3000 = cor(PC_HEX,binned_matrixHEX_3000)
install.packages("corrplot", repos='http://cran.us.r-project.org')

```

```

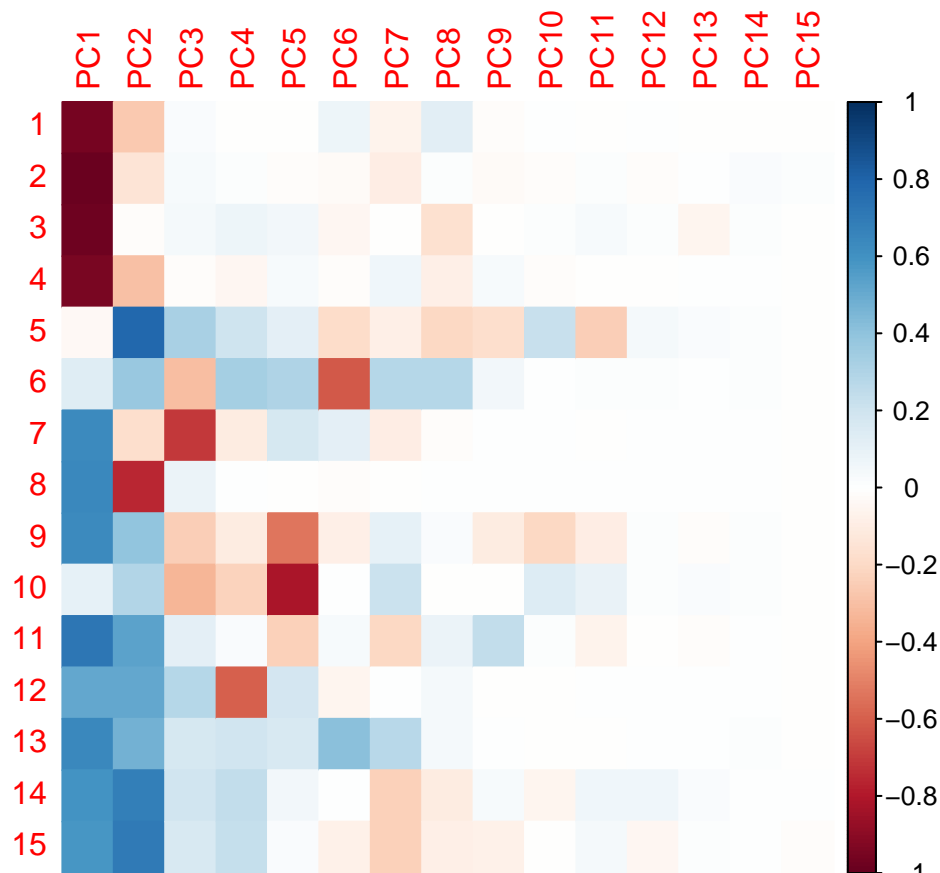
##
## The downloaded binary packages are in
## /var/folders/7k/8ngqf1rd4ggc6b8q6cflfr6m0000gn/T//RtmpwHvQ8q/downloaded_packages

```

```

library(corrplot)
corrplot(t(correlation_3000), method="color")

```



## 7. LDA - 22 Second Bin Width (1100 time points)

```
#DCM:
library(MASS)
lda_info = lda(as.factor(dataframe[,3])~binned_matrix_1100)
lda_info$scaling #Linear discriminant coefficient matrix
```

```
##          LD1          LD2
## binned_matrix_11001  104987.44 -10247.6777
## binned_matrix_11002   54388.31 -8102.0540
## binned_matrix_11003 -108691.70  21575.2469
## binned_matrix_11004   322571.18 -77784.6378
## binned_matrix_11005   292408.29  77370.9329
## binned_matrix_11006 -200755.80  44288.0829
## binned_matrix_11007   352643.75 -22651.6955
## binned_matrix_11008  -68321.02 -14649.7852
## binned_matrix_11009 -279282.17  20518.9594
## binned_matrix_110010  83408.92 -10830.5359
## binned_matrix_110011   83839.59 -6104.0984
## binned_matrix_110012 -149168.32 -17106.6834
## binned_matrix_110013  134495.75  13063.9162
## binned_matrix_110014  110269.17  11380.6808
## binned_matrix_110015   57759.50  1410.5013
## binned_matrix_110016   63126.46  2595.6190
## binned_matrix_110017   57694.37  1549.8656
## binned_matrix_110018   58871.22  1312.9768
```

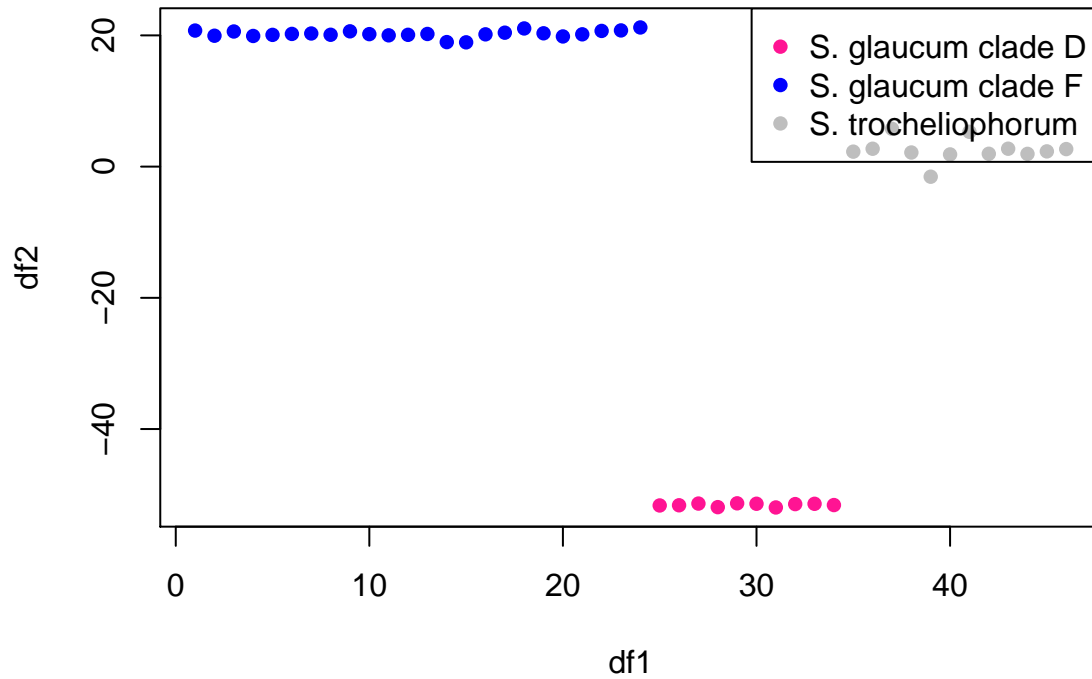
```
## binned_matrix_110019 63714.27 4425.9034
## binned_matrix_110020 51782.50 -1303.9719
## binned_matrix_110021 62702.56 6681.2744
## binned_matrix_110022 61878.92 2208.7342
## binned_matrix_110023 62088.23 2282.1256
## binned_matrix_110024 38283.51 -917.4367
## binned_matrix_110025 70885.55 -2363.1673
## binned_matrix_110026 58800.90 779.3573
## binned_matrix_110027 44003.58 5050.8112
## binned_matrix_110028 -16175.99 -218.0292
## binned_matrix_110029 112767.74 3950.1038
## binned_matrix_110030 71974.07 -674.8565
## binned_matrix_110031 94126.88 7007.6329
## binned_matrix_110032 -111986.51 -31556.3556
## binned_matrix_110033 207541.08 18933.9786
## binned_matrix_110034 -134333.38 -7938.3646
## binned_matrix_110035 83335.66 -11629.8118
## binned_matrix_110036 -120171.34 22913.7350
## binned_matrix_110037 126615.90 -4543.8389
## binned_matrix_110038 -519978.50 103019.0825
## binned_matrix_110039 1158486.05 -50170.5410
## binned_matrix_110040 -274391.55 6343.6471
## binned_matrix_110041 502208.91 72874.4416
## binned_matrix_110042 -148276.86 -86975.0046
```

```
lda_info_p = predict(lda_info,as.data.frame(binned_matrix_1100))
lda_info_pclass = predict(lda_info)$class
par(mar=c(5,4,4,4))
v = as.numeric(lda_info_pclass)
v[which(v==1)] = "deeppink" #D
v[which(v==2)] = 12 #F
v[which(v==3)] = "grey" #Trochi
```

```
plot(lda_info_p$x[,1],pch=16, col = v,main = "LDA - 22 Second Bin Width",xlab="df1",ylab="df2") #Plot t
legend("topright",legend=c("S. glaucum clade D","S. glaucum clade F","S. trocheliophorum"),col=c("deepp
```



## LDA – 22 Second Bin Width



```
clade.manova = manova(binned_matrix_1100~as.factor(dataframe[,3]))
clade.wilks = summary(clade.manova,test="Wilks");clade.wilks
```

```
##              Df      Wilks approx F num Df den Df  Pr(>F)
## as.factor(dataframe[, 3])  2 2.7888e-05   8.9695     84      4 0.02182 *
## Residuals              43
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
#HEX:
library(MASS)
lda_info_HEX = lda(as.factor(dataframeHEX[,3])~binned_matrixHEX_1100)
```

```
## Warning in lda.default(x, grouping, ...): variables are collinear
```

```
lda_info_HEX$scaling #Linear discriminant coefficient matrix
```

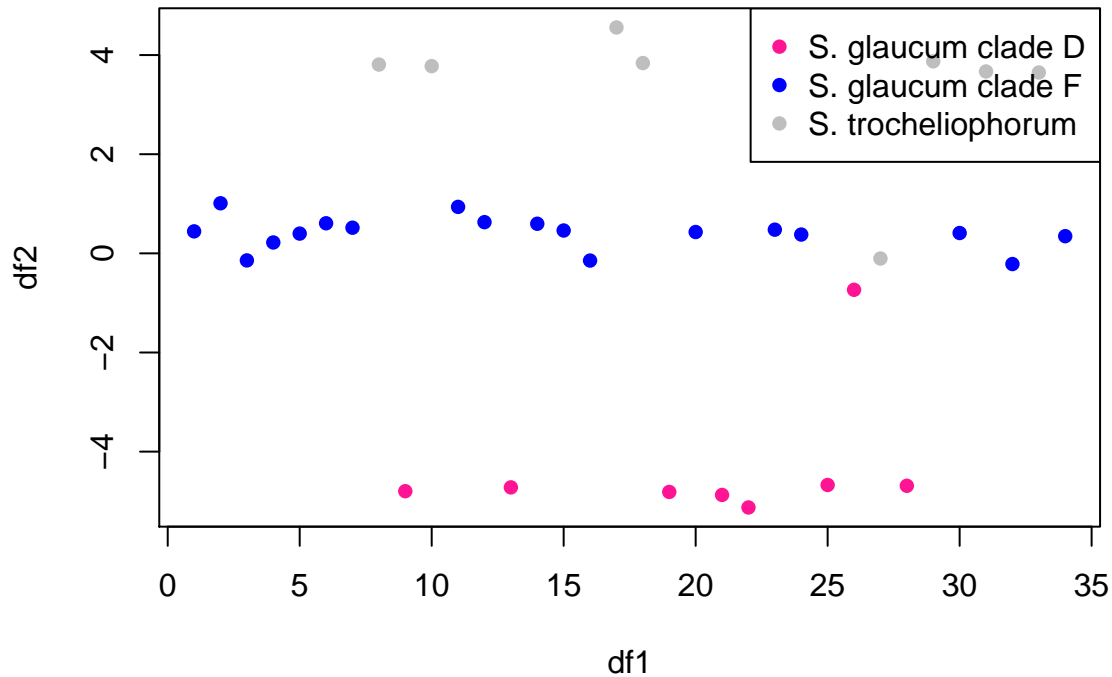
```
##              LD1              LD2
## binned_matrixHEX_11001  471.74130 -2938.08197
## binned_matrixHEX_11002 -637.84470  -84.73397
## binned_matrixHEX_11003   41.62670 -961.16721
## binned_matrixHEX_11004 -1372.43577 3035.57541
## binned_matrixHEX_11005 -5993.80951 4563.24723
## binned_matrixHEX_11006  6547.05973 -12034.44297
## binned_matrixHEX_11007 -10134.74067 -9952.15663
## binned_matrixHEX_11008 -1818.85794  6333.90021
## binned_matrixHEX_11009   455.19931   83.62226
```

```
## binned_matrixHEX_110010    7892.09312 -13292.74995
## binned_matrixHEX_110011 -10320.66075    8474.00010
## binned_matrixHEX_110012   -3745.04231    6401.19362
## binned_matrixHEX_110013    990.54465   -1715.12473
## binned_matrixHEX_110014    7678.08499   -2472.60241
## binned_matrixHEX_110015   -3371.21262    3281.59342
## binned_matrixHEX_110016    142.18437    380.77885
## binned_matrixHEX_110017   -52.75632    300.42901
## binned_matrixHEX_110018   -473.73823   -1025.90367
## binned_matrixHEX_110019    543.63911    252.94928
## binned_matrixHEX_110020   -975.11339   -387.60779
## binned_matrixHEX_110021   -564.11529   -49.86542
## binned_matrixHEX_110022   -264.24766    726.71410
## binned_matrixHEX_110023   -928.38281   -2901.36824
## binned_matrixHEX_110024  -2809.00453   -4756.25763
## binned_matrixHEX_110025   2050.60901   -1652.98819
## binned_matrixHEX_110026  -1284.68162   2420.53269
## binned_matrixHEX_110027   1377.41816   -407.60636
## binned_matrixHEX_110028    271.52947   4607.14243
## binned_matrixHEX_110029  -1672.34914   -2629.28319
## binned_matrixHEX_110030    512.88123   1065.30313
## binned_matrixHEX_110031    358.71107   -2545.90394
## binned_matrixHEX_110032  -2908.32911   4509.70020
## binned_matrixHEX_110033   5010.26043  -10076.62209
## binned_matrixHEX_110034   -922.54947   -2307.94151
## binned_matrixHEX_110035   -432.97146    262.98383
## binned_matrixHEX_110036    305.59013    825.92904
## binned_matrixHEX_110037  -4561.37696   7048.82617
## binned_matrixHEX_110038   2740.82806  -5356.71884
## binned_matrixHEX_110039   2322.67205   10931.62273
## binned_matrixHEX_110040   5097.03048   -1676.74678
## binned_matrixHEX_110041  -5358.41012   4394.32003
## binned_matrixHEX_110042   -338.72476   1223.79064
```

```
lda_info_p = predict(lda_info_HEX,as.data.frame(binned_matrixHEX_1100))
lda_info_pclass = predict(lda_info_HEX)$class
par(mar=c(5,4,4,4))
v = as.numeric(lda_info_pclass)
v[which(v==1)] = "deeppink"    #D
v[which(v==2)] = 12           #F
v[which(v==3)] = "grey"       #Trochi
```

```
plot(lda_info_p$x[,1],pch=16, col = v,main = "LDA - 22 Second Bin Width",xlab="df1",ylab="df2") #Plot t
legend("topright",legend=c("S. glaucum clade D","S. glaucum clade F","S. trocheliophorum"),col=c("deeppink",
```

## LDA – 22 Second Bin Width



```
clade.manova = manova(binned_matrix_1100~as.factor(dataframe[,3]))
clade.wilks = summary(clade.manova,test="Wilks");clade.wilks
```

```
##              Df      Wilks approx F num Df den Df  Pr(>F)
## as.factor(dataframe[, 3])  2 2.7888e-05  8.9695    84    4 0.02182 *
## Residuals              43
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The Wilk's Lambda statistic corresponds to a p-value of ~.09, suggesting that the clades account for a large (but only marginally significant) proportion of the variance in binned area.

### 8. LDA - 40 Second Bin Width (2000 time points)

```
#DCM:
lda_info = lda(as.factor(dataframe[,3])~binned_matrix_2000)
lda_info$scaling #Linear discriminant coefficient matrix
```

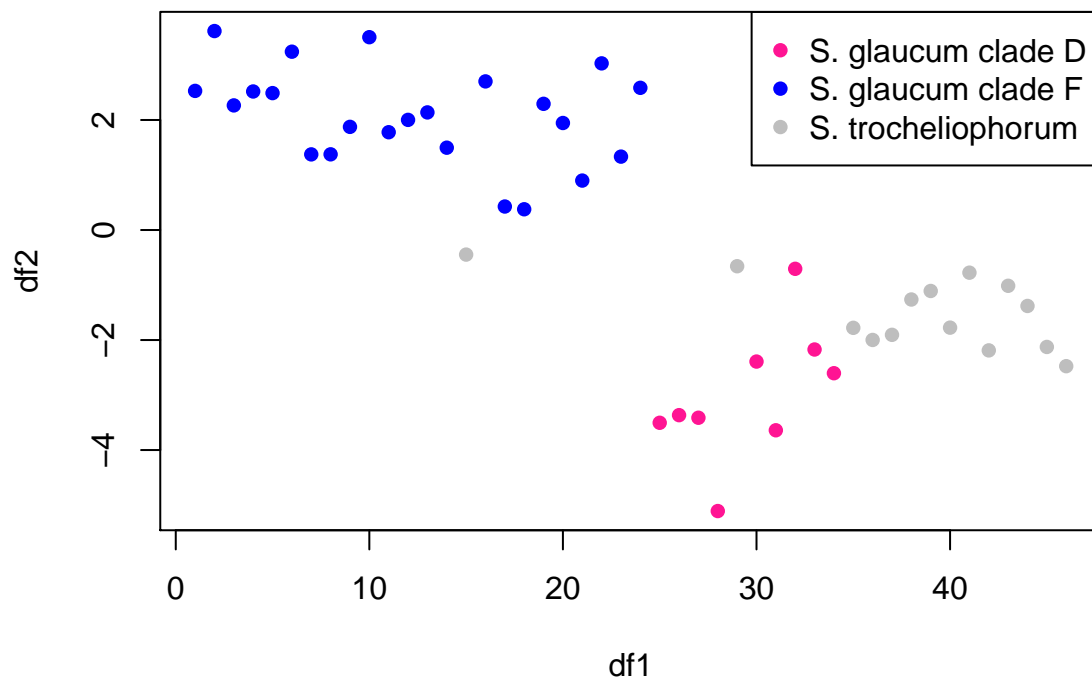
```
##              LD1      LD2
## binned_matrix_20001  -930.8213  377.1855
## binned_matrix_20002 -2659.6435 1206.8505
## binned_matrix_20003 -5686.0967 4828.8096
## binned_matrix_20004 -3949.0293 2900.1558
## binned_matrix_20005 -2659.0185  306.1249
## binned_matrix_20006 -2826.1903 2566.9088
## binned_matrix_20007 -3929.1368 3896.8179
## binned_matrix_20008 -2795.2204 2036.1473
## binned_matrix_20009 -3164.1366 2250.3553
```

```
## binned_matrix_200010 -3272.1381 2335.5415
## binned_matrix_200011 -2988.4976 2178.6680
## binned_matrix_200012 -3203.6255 2236.3334
## binned_matrix_200013 -3135.0215 2288.0967
## binned_matrix_200014 -2913.8146 2359.8578
## binned_matrix_200015 -3373.0450 2052.8931
## binned_matrix_200016 -3192.7739 2133.6520
## binned_matrix_200017 -3121.2817 2349.6865
## binned_matrix_200018 -2578.7989 2693.7657
## binned_matrix_200019 -4374.9307 1038.8373
## binned_matrix_200020 -4105.5444 1080.6504
## binned_matrix_200021 -2775.4915 4245.1673
## binned_matrix_200022 -409.8137 5221.9788
## binned_matrix_200023 -6578.5107 884.2937
```

```
lda_info_p = predict(lda_info,as.data.frame(binned_matrix_2000))
lda_info_pclass = predict(lda_info)$class
par(mar=c(5,4,4,4))
v = as.numeric(lda_info_pclass)
v[which(v==1)] = "deeppink" #D
v[which(v==2)] = 12 #F
v[which(v==3)] = "grey" #Trochi
```

```
plot(lda_info_p$x[,1],pch=16, col = v,main = "LDA - 40 Second Bin Width",xlab="df1",ylab="df2") #Plot t
legend("topright",legend=c("S. glaucum clade D","S. glaucum clade F","S. trocheliophorum"),col=c("deeppink",
```

## LDA – 40 Second Bin Width



```
clade.manova = manova(binned_matrix_2000~as.factor(dataframe[,3]))
clade.wilks = summary(clade.manova,test="Wilks");clade.wilks
```

```
##                               Df      Wilks approx F num Df den Df Pr(>F)
## as.factor(dataframe[, 3])    2 0.060541    2.7978      46     42 5e-04 ***
## Residuals                    43
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

*#HEX:*

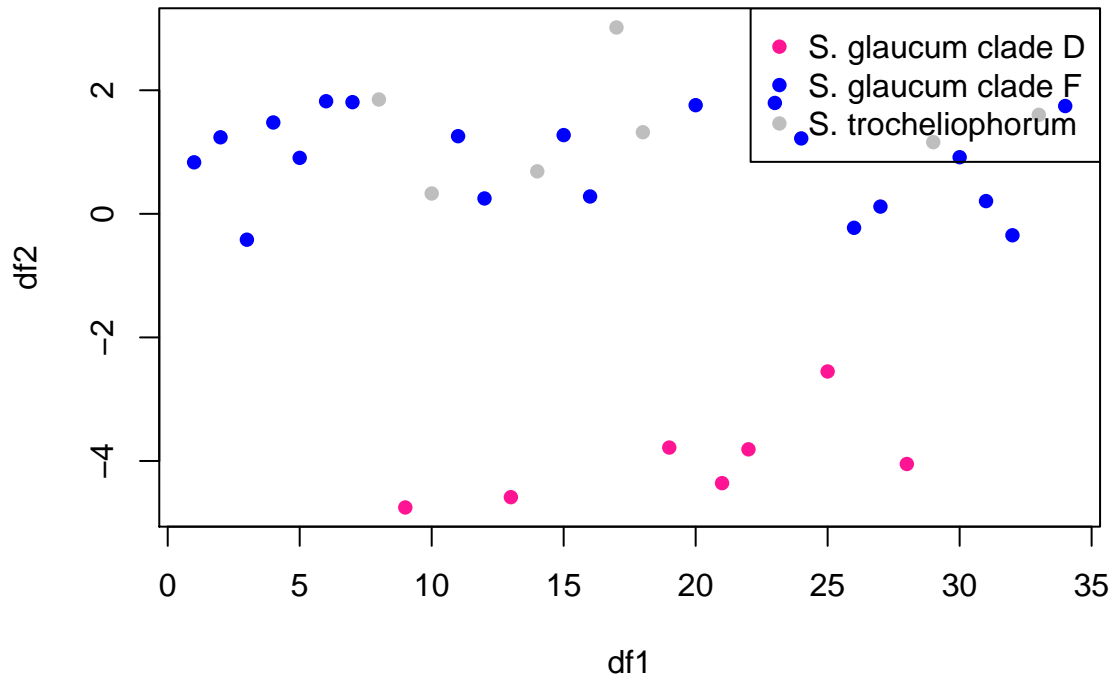
```
library(MASS)
lda_info_HEX = lda(as.factor(dataframeHEX[,3])~binned_matrixHEX_2000)
lda_info_HEX$scaling #Linear discriminant coefficient matrix
```

```
##                               LD1          LD2
## binned_matrixHEX_20001    -13790.565 -3349.52070
## binned_matrixHEX_20002    -13521.347   36.21076
## binned_matrixHEX_20003    -14055.184 -4963.69388
## binned_matrixHEX_20004    -13832.699 -1719.27797
## binned_matrixHEX_20005    -13659.273 -2001.82162
## binned_matrixHEX_20006    -19789.788 -1875.87090
## binned_matrixHEX_20007    -12488.758 -2334.96751
## binned_matrixHEX_20008    -13112.124  -966.27766
## binned_matrixHEX_20009    -13399.830 -1912.90911
## binned_matrixHEX_200010   -13803.416 -2051.29403
## binned_matrixHEX_200011   -13672.562 -1865.72696
## binned_matrixHEX_200012   -13842.942 -1997.69943
## binned_matrixHEX_200013   -12909.960 -2661.54398
## binned_matrixHEX_200014   -14808.574 -1580.81402
## binned_matrixHEX_200015   -13765.089 -1472.32702
## binned_matrixHEX_200016   -13092.616 -2205.84480
## binned_matrixHEX_200017   -13506.498 -2049.66831
## binned_matrixHEX_200018   -14592.186 -1839.72216
## binned_matrixHEX_200019   -13390.464 -2451.74248
## binned_matrixHEX_200020   -14170.866 -1654.03150
## binned_matrixHEX_200021   -21795.463   198.40224
## binned_matrixHEX_200022    -6758.261 -4195.66059
## binned_matrixHEX_200023   -15442.100 -1595.84804
```

```
lda_info_p = predict(lda_info_HEX,as.data.frame(binned_matrixHEX_2000))
lda_info_pclass = predict(lda_info_HEX)$class
par(mar=c(5,4,4,4))
v = as.numeric(lda_info_pclass)
v[which(v==1)] = "deeppink"    #D
v[which(v==2)] = 12           #F
v[which(v==3)] = "grey"       #Trochi
```

```
plot(lda_info_p$x[,1],pch=16, col = v,main = "LDA - 22 Second Bin Width",xlab="df1",ylab="df2") #Plot t
legend("topright",legend=c("S. glaucum clade D","S. glaucum clade F","S. trocheliophorum"),col=c("deeppink",
```

## LDA – 22 Second Bin Width



```
clade.manova = manova(binned_matrix_2000~as.factor(dataframe[,3]))
clade.wilks = summary(clade.manova,test="Wilks");clade.wilks
```

```
##              Df      Wilks approx F num Df den Df Pr(>F)
## as.factor(dataframe[, 3])  2 0.060541   2.7978    46    42 5e-04 ***
## Residuals              43
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The Wilk's Lambda statistic corresponds to a p-value of ~0.0008, suggesting that the clades account for a large, significant proportion of the variance in binned area.

### 9. LDA — 60 Second Bin Width (3000 time points)

```
#DCM:
lda_info = lda(as.factor(dataframe[,3])~binned_matrix_3000)
lda_info$scaling #Linear discriminant coefficient matrix
```

```
##              LD1              LD2
## binned_matrix_30001 2429.3438 -1896.0113
## binned_matrix_30002   19.0294  2007.2291
## binned_matrix_30003  465.6450  1734.7241
## binned_matrix_30004 1601.5099   410.1844
## binned_matrix_30005 1380.0586   904.8056
## binned_matrix_30006 1246.5990   571.4232
## binned_matrix_30007 1179.6731   668.0790
## binned_matrix_30008 1150.9457   598.3352
## binned_matrix_30009 1222.7453   620.7320
```

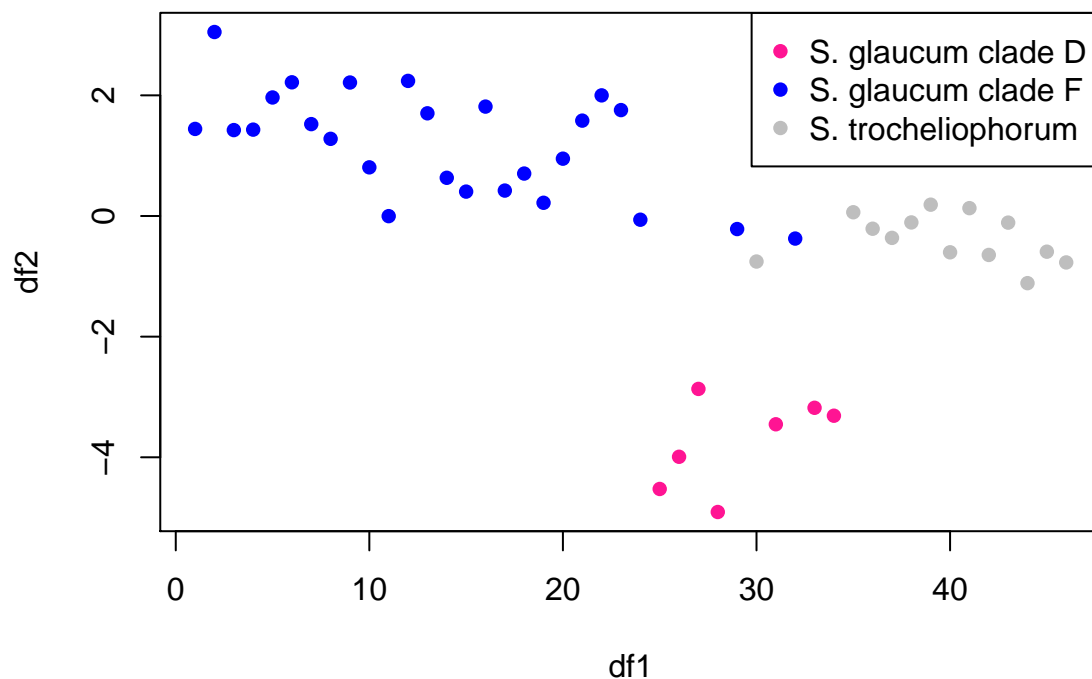


```
## binned_matrix_300010 1203.9553 575.4450
## binned_matrix_300011 1115.9976 546.8839
## binned_matrix_300012 1625.0500 636.5107
## binned_matrix_300013 397.4052 357.4868
## binned_matrix_300014 1536.9944 936.9234
## binned_matrix_300015 2307.1136 1115.6050
```

```
lda_info_p = predict(lda_info,as.data.frame(binned_matrix_3000))
lda_info_pclass = predict(lda_info)$class
par(mar=c(5,4,4,4))
v = as.numeric(lda_info_pclass)
v[which(v==1)] = "deeppink" #D
v[which(v==2)] = 12 #F
v[which(v==3)] = "grey" #Trochi
```

```
plot(lda_info_p$x[,1],pch=16, col = v,main = "LDA - 60 Second Bin Width",xlab="df1",ylab="df2") #Plot t
legend("topright",legend=c("S. glaucum clade D","S. glaucum clade F","S. trocheliophorum"),col=c("deeppink",
```

## LDA – 60 Second Bin Width



```
clade.manova = manova(binned_matrix_3000~as.factor(dataframe[,3]))
clade.wilks = summary(clade.manova,test="Wilks");clade.wilks
```

```
##              Df  Wilks approx F num Df den Df    Pr(>F)
## as.factor(dataframe[, 3])  2 0.11187   3.847   30   58 5.459e-06 ***
## Residuals              43
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
#HEX:
```

```
library(MASS)
```

```
lda_info_HEX = lda(as.factor(dataframeHEX[,3])~binned_matrixHEX_3000)
```

```
lda_info_HEX$scaling #Linear discriminant coefficient matrix
```

```
##                               LD1          LD2
## binned_matrixHEX_30001    309.69700    494.059332
## binned_matrixHEX_30002   -925.13731  -1613.256024
## binned_matrixHEX_30003    1078.44291   -340.434835
## binned_matrixHEX_30004   -266.50567    721.159649
## binned_matrixHEX_30005   -514.51923   -62.153202
## binned_matrixHEX_30006     39.07211    213.397884
## binned_matrixHEX_30007     52.64672     19.636385
## binned_matrixHEX_30008    -45.99341    112.648070
## binned_matrixHEX_30009     48.45117   -582.936377
## binned_matrixHEX_300010   181.17609    723.283532
## binned_matrixHEX_300011   -69.66687     5.255114
## binned_matrixHEX_300012   -37.22974    102.307923
## binned_matrixHEX_300013   -99.82671    -30.379332
## binned_matrixHEX_300014  -132.81999    -49.161067
## binned_matrixHEX_300015   249.05413   1295.611818
```

```
lda_info_p = predict(lda_info_HEX,as.data.frame(binned_matrixHEX_3000))
```

```
lda_info_pclass = predict(lda_info_HEX)$class
```

```
par(mar=c(5,4,4,4))
```

```
v = as.numeric(lda_info_pclass)
```

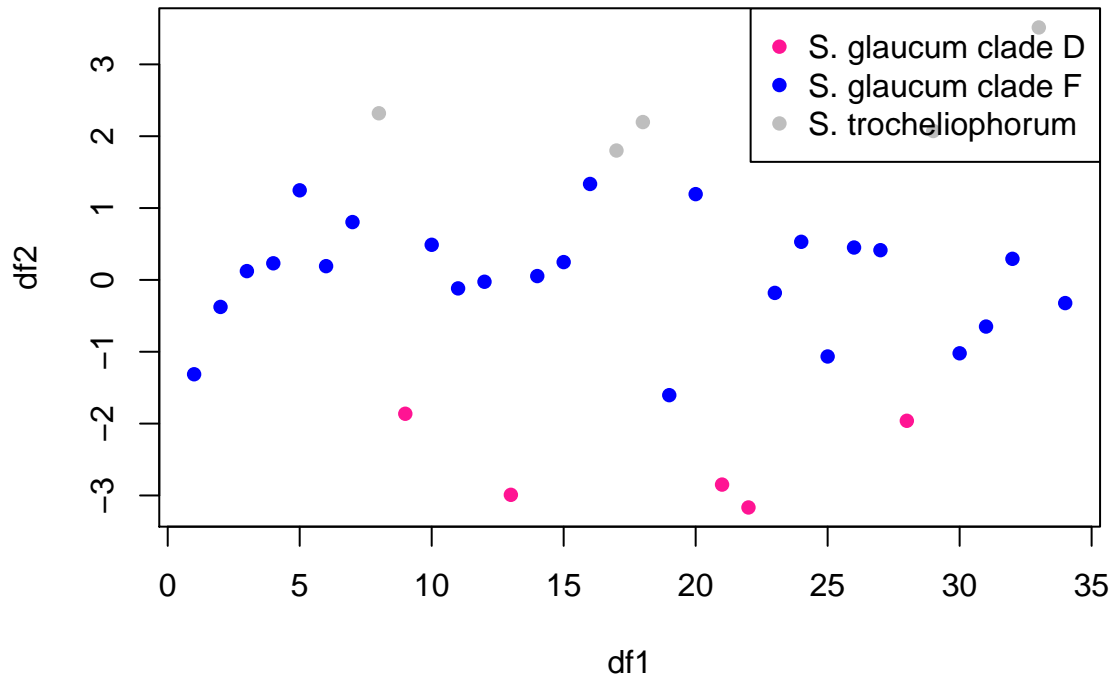
```
v[which(v==1)] = "deeppink" #D
```

```
v[which(v==2)] = 12 #F
```

```
v[which(v==3)] = "grey" #Trochi
```

```
plot(lda_info_p$x[,1],pch=16, col = v,main = "LDA - 22 Second Bin Width",xlab="df1",ylab="df2") #Plot t  
legend("topright",legend=c("S. glaucum clade D","S. glaucum clade F","S. trocheliophorum"),col=c("deeppink",
```

## LDA – 22 Second Bin Width



```
clade.manova = manova(binned_matrix_3000~as.factor(dataframe[,3]))
clade.wilks = summary(clade.manova,test="Wilks");clade.wilks
```

```
##                               Df   Wilks approx F num Df den Df   Pr(>F)
## as.factor(dataframe[, 3])    2 0.11187    3.847    30   58 5.459e-06 ***
## Residuals                    43
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The Wilk's Lambda statistic corresponds to a p-value of  $\sim 0$ , suggesting that the clades account for a large, significant proportion of the variance in binned area.

## APPENDIX

Experimentation with different bin sizes:

```
binned_matrix_1100 = binner(dataframe[,1:47008],1100, DELTA)
binned_matrix_1300 = binner(dataframe[,1:47008],1300, DELTA)
binned_matrix_1500 = binner(dataframe[,1:47008],1500, DELTA)
binned_matrix_1600 = binner(dataframe[,1:47008],1600, DELTA)
binned_matrix_1800 = binner(dataframe[,1:47008],1800, DELTA)
binned_matrix_2000 = binner(dataframe[,1:47008],2000, DELTA)
```

```
trapAreas.centered = scale(binned_matrix_1100, center = TRUE, scale = FALSE)
pca.result <- prcomp(trapAreas.centered)
str(pca.result)
```

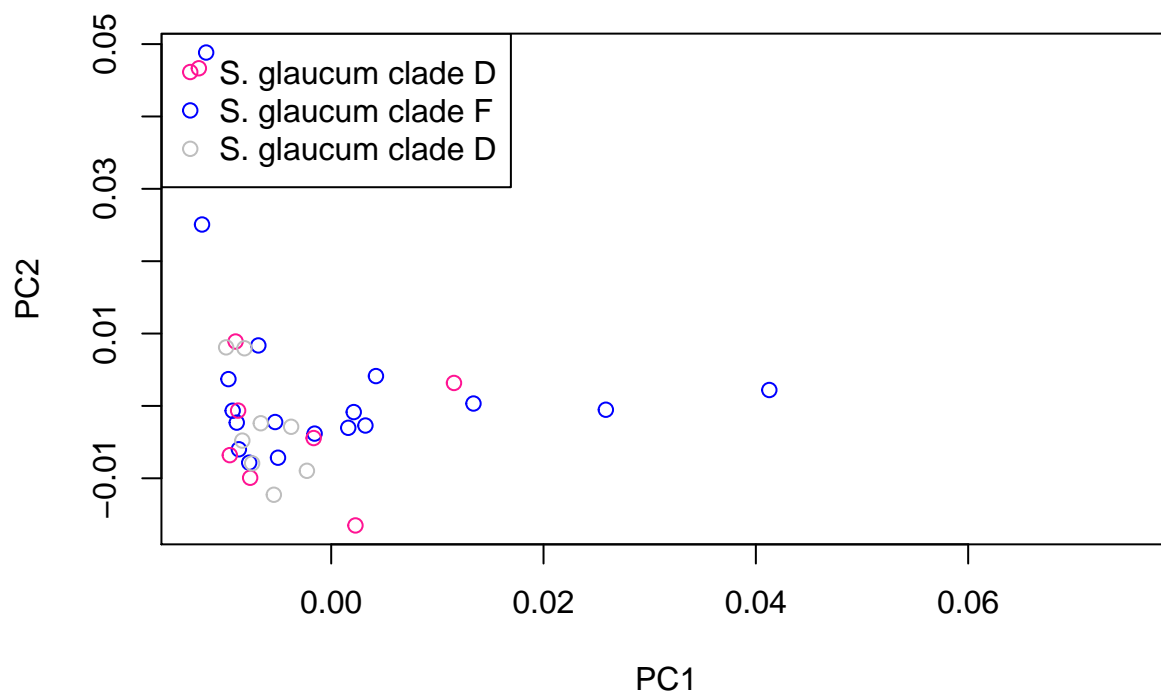
```
## List of 5
```

```
## $ sdev      : num [1:42] 0.01611 0.01261 0.00898 0.00749 0.00634 ...
## $ rotation: num [1:42, 1:42] -0.0346 -0.0367 -0.0264 -0.0336 -0.0341 ...
##   attr(*, "dimnames")=List of 2
##   ..$ : NULL
##   ..$ : chr [1:42] "PC1" "PC2" "PC3" "PC4" ...
## $ center   : num [1:42] 1.28e-18 -8.30e-19 1.21e-18 -1.06e-18 -1.51e-19 ...
## $ scale     : logi FALSE
## $ x        : num [1:46, 1:42] 0.0016 0.04127 -0.00686 -0.01217 -0.00156 ...
##   attr(*, "dimnames")=List of 2
##   ..$ : chr [1:46] "col" "col" "col" "col" ...
##   ..$ : chr [1:42] "PC1" "PC2" "PC3" "PC4" ...
## - attr(*, "class")= chr "prcomp"
```

```
PC = pca.result$x
```

```
plot(PC[,1],PC[,2],col="white",main = "22 Second Bin Width (1100 time points)",xlab="PC1",ylab="PC2")
points(PC[F,1],PC[F,2],col="blue")
points(PC[D,1],PC[D,2],col="deeppink")
points(PC[T,1],PC[T,2],col="grey")
legend("topleft",legend=c("S. glaucum clade D","S. glaucum clade F","S. glaucum clade D"),col=c("deeppink","blue","grey"))
```

## 22 Second Bin Width (1100 time points)



```
trapAreas.centered = scale(binned_matrix_1300, center = TRUE, scale = FALSE)
pca.result <- prcomp(trapAreas.centered)
str(pca.result)
```

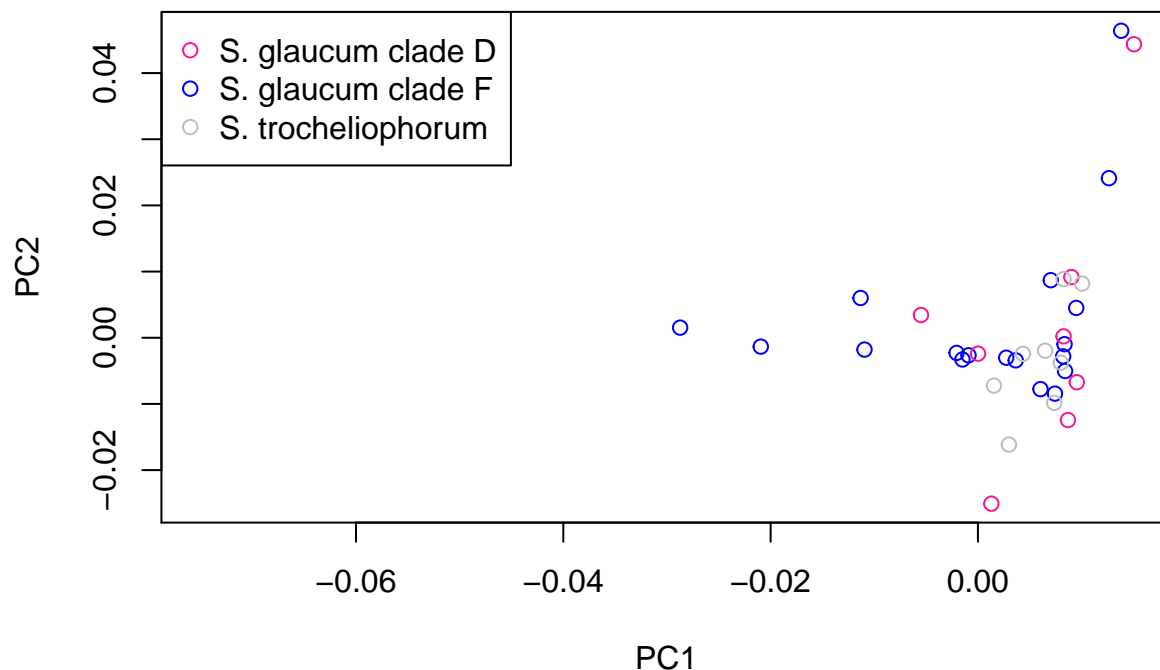
```
## List of 5
## $ sdev      : num [1:36] 0.01613 0.0126 0.00998 0.00911 0.00731 ...
## $ rotation: num [1:36, 1:36] 0.0431 0.0407 0.0331 0.0396 0.0422 ...
```

```
##    .- attr(*, "dimnames")=List of 2
##    .. ..$ : NULL
##    .. ..$ : chr [1:36] "PC1" "PC2" "PC3" "PC4" ...
##    $ center : num [1:36] 1.06e-18 4.53e-19 -2.26e-19 -6.79e-19 -1.36e-18 ...
##    $ scale : logi FALSE
##    $ x : num [1:46, 1:36] -0.000907 -0.02873 0.007029 0.012644 0.002731 ...
##    .- attr(*, "dimnames")=List of 2
##    .. ..$ : chr [1:46] "col" "col" "col" "col" ...
##    .. ..$ : chr [1:36] "PC1" "PC2" "PC3" "PC4" ...
##    - attr(*, "class")= chr "prcomp"
```

```
PC = pca.result$x
```

```
plot(PC[,1],PC[,2],col="white",main = "26 Second Bin Width (1300 time points)",xlab="PC1",ylab="PC2")
points(PC[F,1],PC[F,2],col="blue")
points(PC[D,1],PC[D,2],col="deeppink")
points(PC[T,1],PC[T,2],col="grey")
legend("topleft",legend=c("S. glaucum clade D","S. glaucum clade F","S. trocheliophorum"),col=c("deeppink",
```

## 26 Second Bin Width (1300 time points)



```
trapAreas.centered = scale(binned_matrix_1500, center = TRUE, scale = FALSE)
pca.result <- prcomp(trapAreas.centered)
str(pca.result)
```

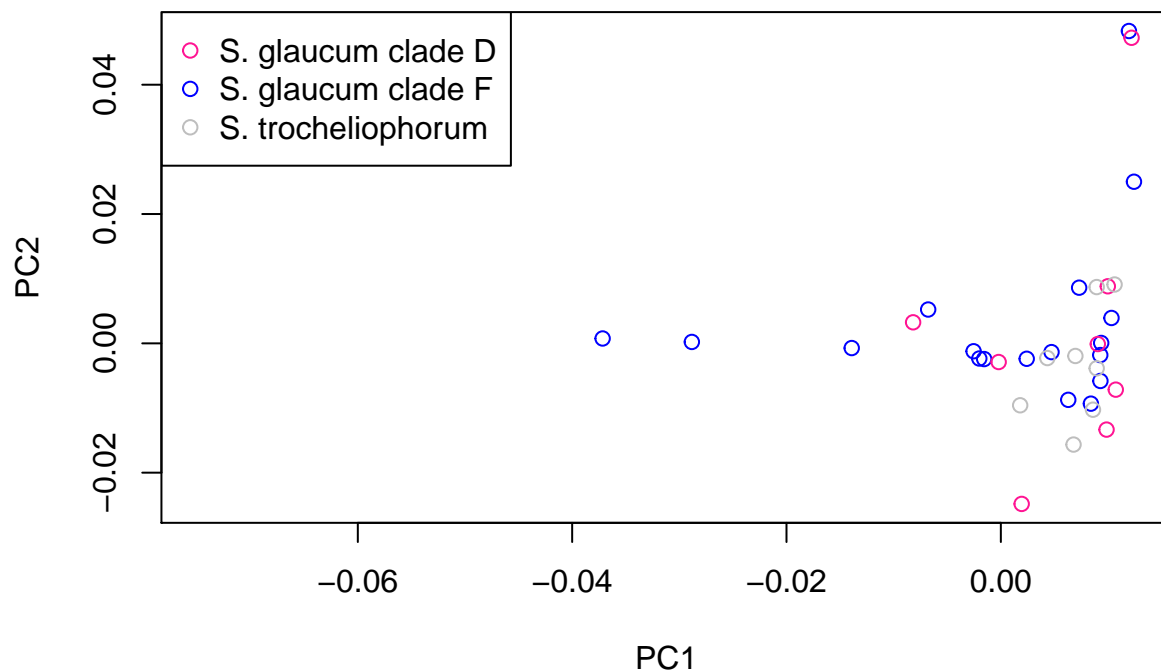
```
## List of 5
## $ sdev : num [1:31] 0.0169 0.01303 0.00994 0.00969 0.00771 ...
## $ rotation: num [1:31, 1:31] 0.0479 0.0436 0.0405 0.0462 0.0449 ...
##    .- attr(*, "dimnames")=List of 2
##    .. ..$ : NULL
```

```
## .. ..$ : chr [1:31] "PC1" "PC2" "PC3" "PC4" ...
## $ center : num [1:31] -1.36e-18 -1.51e-19 -1.06e-18 1.51e-18 1.43e-18 ...
## $ scale : logi FALSE
## $ x : num [1:46, 1:31] -0.00155 -0.03716 0.00731 0.01242 0.00241 ...
## ..- attr(*, "dimnames")=List of 2
## .. ..$ : chr [1:46] "col" "col" "col" "col" ...
## .. ..$ : chr [1:31] "PC1" "PC2" "PC3" "PC4" ...
## - attr(*, "class")= chr "prcomp"
```

```
PC = pca.result$x
```

```
plot(PC[,1],PC[,2],col="white",main = "30 Second Bin Width (1500 time points)",xlab="PC1",ylab="PC2")
points(PC[F,1],PC[F,2],col="blue")
points(PC[D,1],PC[D,2],col="deeppink")
points(PC[T,1],PC[T,2],col="grey")
legend("topleft",legend=c("S. glaucum clade D","S. glaucum clade F","S. trocheliophorum"),col=c("deeppink",
```

### 30 Second Bin Width (1500 time points)



```
trapAreas.centered = scale(binned_matrix_1600, center = TRUE, scale = FALSE)
pca.result <- prcomp(trapAreas.centered)
str(pca.result)
```

```
## List of 5
## $ sdev : num [1:29] 0.01756 0.01337 0.01061 0.00779 0.00752 ...
## $ rotation: num [1:29, 1:29] 0.0487 0.0437 0.0429 0.0485 0.0455 ...
## ..- attr(*, "dimnames")=List of 2
## .. ..$ : NULL
## .. ..$ : chr [1:29] "PC1" "PC2" "PC3" "PC4" ...
## $ center : num [1:29] 5.28e-19 0.00 7.54e-19 -1.43e-18 1.66e-18 ...
```

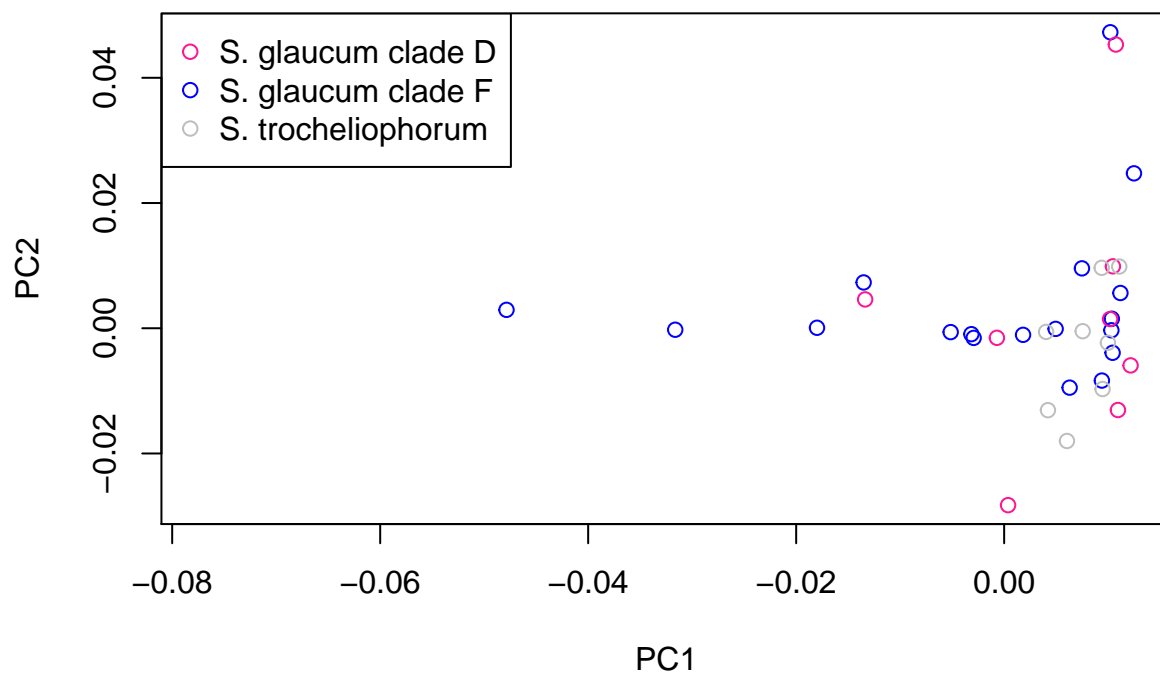


```
## $ scale : logi FALSE
## $ x      : num [1:46, 1:29] -0.00293 -0.04785 0.00748 0.01248 0.00182 ...
## ..- attr(*, "dimnames")=List of 2
## .. ..$ : chr [1:46] "col" "col" "col" "col" ...
## .. ..$ : chr [1:29] "PC1" "PC2" "PC3" "PC4" ...
## - attr(*, "class")= chr "prcomp"
```

```
PC = pca.result$x
```

```
plot(PC[,1],PC[,2],col="white",main = "32 Second Bin Width (1600 time points)",xlab="PC1",ylab="PC2")
points(PC[F,1],PC[F,2],col="blue")
points(PC[D,1],PC[D,2],col="deeppink")
points(PC[T,1],PC[T,2],col="grey")
legend("topleft",legend=c("S. glaucum clade D","S. glaucum clade F","S. trocheliophorum"),col=c("deeppink","blue","grey"))
```

### 32 Second Bin Width (1600 time points)



```
trapAreas.centered = scale(binned_matrix_1800, center = TRUE, scale = FALSE)
pca.result <- prcomp(trapAreas.centered)
str(pca.result)
```

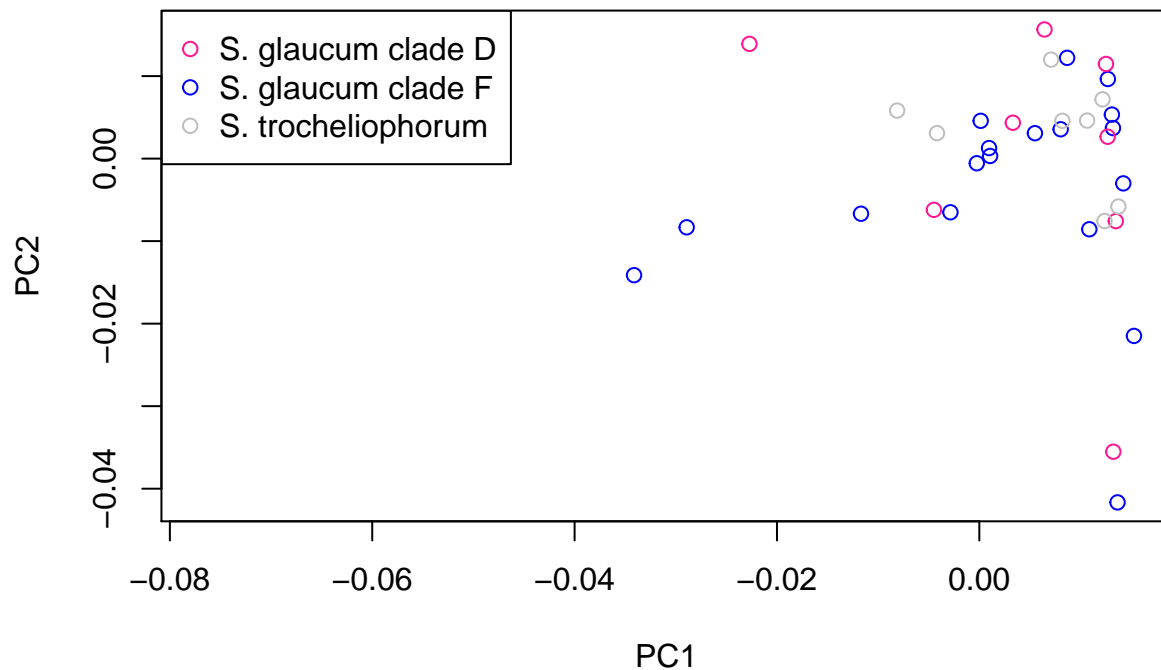
```
## List of 5
## $ sdev : num [1:26] 0.01821 0.01121 0.0106 0.00732 0.00714 ...
## $ rotation: num [1:26, 1:26] 0.0599 0.0515 0.0571 0.0581 0.0577 ...
## ..- attr(*, "dimnames")=List of 2
## .. ..$ : NULL
## .. ..$ : chr [1:26] "PC1" "PC2" "PC3" "PC4" ...
## $ center : num [1:26] -2.56e-18 2.49e-18 -6.03e-19 1.36e-18 -1.51e-18 ...
## $ scale : logi FALSE
## $ x      : num [1:46, 1:26] 0.00106 -0.03412 0.01087 0.01528 0.0055 ...
```

```
##    .- attr(*, "dimnames")=List of 2
##    .. ..$ : chr [1:46] "col" "col" "col" "col" ...
##    .. ..$ : chr [1:26] "PC1" "PC2" "PC3" "PC4" ...
##    - attr(*, "class")= chr "prcomp"

PC = pca.result$x

plot(PC[,1],PC[,2],col="white",main = "36 Second Bin Width (1800 time points)",xlab="PC1",ylab="PC2")
points(PC[F,1],PC[F,2],col="blue")
points(PC[D,1],PC[D,2],col="deeppink")
points(PC[T,1],PC[T,2],col="grey")
legend("topleft",legend=c("S. glaucum clade D","S. glaucum clade F","S. trocheliophorum"),col=c("deeppink","blue","grey"))
```

### 36 Second Bin Width (1800 time points)



```
trapAreas.centered = scale(binned_matrix_2000, center = TRUE, scale = FALSE)
pca.result <- prcomp(trapAreas.centered)
str(pca.result)
```

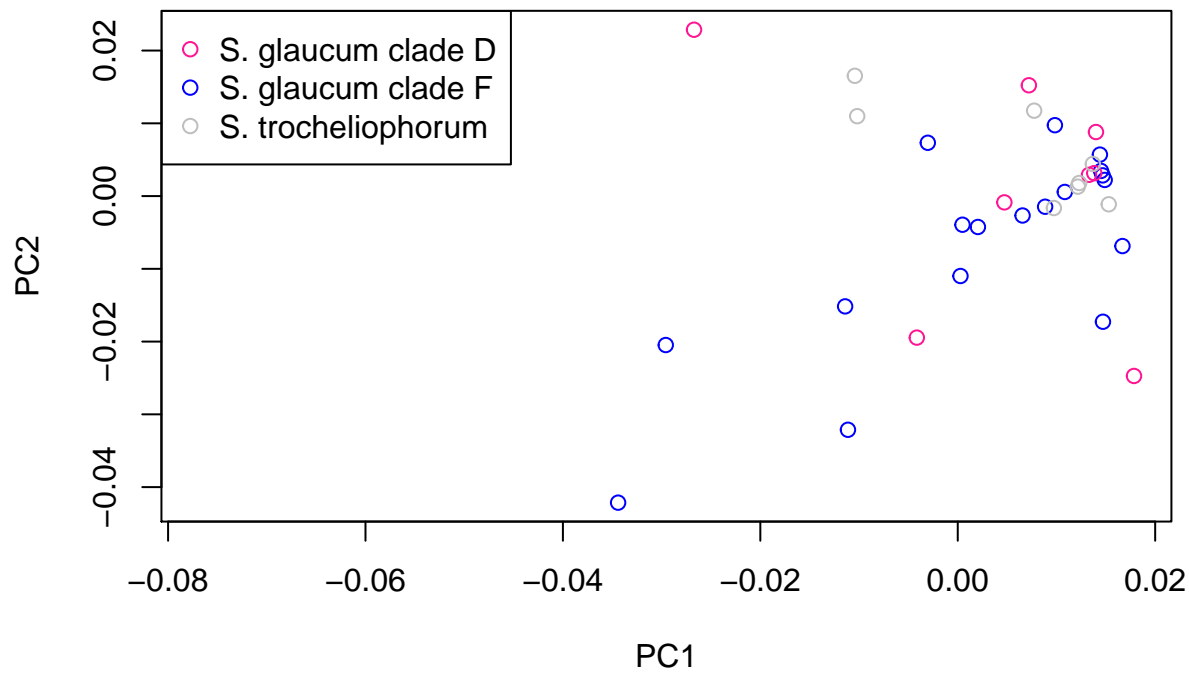
```
## List of 5
## $ sdev      : num [1:23] 0.01898 0.01262 0.01193 0.00775 0.00694 ...
## $ rotation: num [1:23, 1:23] 0.0648 0.0567 0.0625 0.063 0.0584 ...
##    .- attr(*, "dimnames")=List of 2
##    .. ..$ : NULL
##    .. ..$ : chr [1:23] "PC1" "PC2" "PC3" "PC4" ...
## $ center    : num [1:23] -1.06e-18 -2.87e-18 -2.26e-18 2.72e-18 1.06e-18 ...
## $ scale     : logi FALSE
## $ x         : num [1:46, 1:23] 0.00204 -0.03441 0.01085 0.01669 0.00656 ...
##    .- attr(*, "dimnames")=List of 2
##    .. ..$ : chr [1:46] "col" "col" "col" "col" ...
```

```
## .. ..$ : chr [1:23] "PC1" "PC2" "PC3" "PC4" ...
## - attr(*, "class")= chr "prcomp"
```

```
PC = pca.result$x
```

```
plot(PC[,1],PC[,2],col="white",main = "40 Second Bin Width (2000 time points)",xlab="PC1",ylab="PC2")
points(PC[F,1],PC[F,2],col="blue")
points(PC[D,1],PC[D,2],col="deeppink")
points(PC[T,1],PC[T,2],col="grey")
legend("topleft",legend=c("S. glaucum clade D","S. glaucum clade F","S. trocheliophorum"),col=c("deeppink","blue","grey"))
```

### 40 Second Bin Width (2000 time points)



```
barplot(binned_matrix_2000[26,],col="skyblue",main="Sample 29",ylab = "Area",xlab="Bin")
```

```
setwd("C:/Users/maloneem/Documents/CoralProject/BinnedAreas") for(i in 1:46){ png(filename=paste("Binned_Areas",as.character(i),".png"), height=250, width=650, bg="white") barplot(binned_matrix_2000[i,],col="skyblue",ylab = "Area",xlab="Bin") dev.off() }
```

```
sum(binned_matrix_2000[25,])
```