# Appendix 6: ComplexHeatMap Test Script

## Analysis from: <https://jokergoo.github.io/ComplexHeatmap-reference/book/a-single-heatmap.html>

## Initial environment setup (Code Block 1)

knitr::opts\_chunk$set(echo = TRUE)

## Clear Global Environment (Code Block 2)

remove(list = ls())

## Set Library Directory (Code Block 3)

#PC Path  
.libPaths(c("L:/RStudios/RPackRatLibLocations", "L:/RStudios/RPackRat\_2019\_04\_DESEQLibs"))  
  
#Set working directory  
#PC Path  
setwd("E:/Dropbox/Dropbox/Harrison Lab - Trevor Randall/RNASeq Analysis/RNASeqAnlyPkrat\_2020\_03/CF39vsCF39S/Pure Temperature Analysis Comparisons/ComplexHeatmap and Clustering")  
#Laptop Path  
#Set Library Directory  
#.libPaths(c(""))  
#setwd("")  
#sink(file = "./RSessionRawRun.txt")

## Libraries required (Code Block 4)

library(readr)  
library(ComplexHeatmap)  
library(circlize)  
library(cluster)  
library(dendextend)  
#library(seriation)

PTA = Pure Temperature Analysis

## Dataset (Code Block 5)

Combined\_List\_Analysis\_Shrunk2 <- read\_delim("E:/Dropbox/Dropbox/Harrison Lab - Trevor Randall/RNASeq Analysis/RNASeqAnlyPkrat\_2020\_03/CF39vsCF39S/Pure Temperature Analysis Comparisons/Combined List Analysis Shrunk2.txt", "\t", escape\_double = FALSE, trim\_ws = TRUE)  
  
PAO1\_Tags <- read\_csv("../../../LocusTags and PAO1 Loci ID's.csv")  
  
Combined\_List\_Analysis\_Shrunk2 <-merge(Combined\_List\_Analysis\_Shrunk2, PAO1\_Tags, by.x = "locus", by.y = "locus", all.x = T, all.y = T)  
ColToRemove <- c("locusNumber.x", "locusNumber.y")  
Combined\_List\_Analysis\_Shrunk2 <- Combined\_List\_Analysis\_Shrunk2[, !names(Combined\_List\_Analysis\_Shrunk2) %in% ColToRemove, drop = F]  
  
FinalColumnNames <- append(colnames(Combined\_List\_Analysis\_Shrunk2)[1:23], c("NucleotidePAO1Locus", "NucleotidePAO1ID", "ProteinPAO1Locus", "ProteinPAO1ID"))  
  
colnames(Combined\_List\_Analysis\_Shrunk2) <- FinalColumnNames

## (Code Block 6)

FoldChangeData <- data.frame(Combined\_List\_Analysis\_Shrunk2$locus, Combined\_List\_Analysis\_Shrunk2$FoldChange\_25, Combined\_List\_Analysis\_Shrunk2$FoldChange\_29, Combined\_List\_Analysis\_Shrunk2$FoldChange\_33, Combined\_List\_Analysis\_Shrunk2$FoldChange\_37, Combined\_List\_Analysis\_Shrunk2$FoldChange\_41)  
  
NamesOfColumns <- c("locus", "FoldChange\_25", "FoldChange\_29", "FoldChange\_33", "FoldChange\_37", "FoldChange\_41")  
colnames(FoldChangeData) <- NamesOfColumns  
rownames(FoldChangeData) <- FoldChangeData$locus  
FoldChangeData <- FoldChangeData[,-1]  
  
#FoldChangeDataWithNAMat <- data.matrix(FoldChangeData)  
#na\_index = sample(c(TRUE, FALSE), nrow(FoldChangeDataWithNAMat)\*ncol(FoldChangeDataWithNAMat), replace = TRUE, prob = c(1, 9))  
#FoldChangeDataWithNAMat[na\_index] = NA  
  
FoldChangeDataNoNA <- FoldChangeData  
FoldChangeDataNoNA[is.na(FoldChangeDataNoNA)] <- 0  
FoldChangeDataNoNAMat <- data.matrix(FoldChangeDataNoNA)

## (Code Block 7)

mat <- FoldChangeDataNoNAMat  
colorSetRedBlue <- colorRamp2(c(-10, 0, 10), c(rgb(0, 114/255, 178/255), "white", rgb(213/255, 94/255, 0))) #First color is Blue and second is Vermillion from the Nature Colorblind palette  
ColOrder = order(as.numeric(gsub("FoldChange\_", "", colnames(mat))))

## Possible clustering options that I could see

clustering\_distance\_rows = "pearson", #Pre-defined distance method (1 - pearson) #Seems to look the most split up along with the single distancing  
 clustering\_distance\_rows = "spearman",  
 clustering\_distance\_rows = "kendall",  
 clustering\_distance\_rows = function(mat) dist(mat), #A function that calculates distance between martix points  
 clustering\_distance\_rows = function(x, y) 1 - cor(x, y), #A function that calculates pairwise distance between martix points  
 clustering\_method\_rows = "single",  
 clustering\_method\_rows = "complete",  
 cluster\_rows = diana(mat), #Cluster used here  
 cluster\_rows = function(mat) as.dendrogram(diana(mat)),

## Coloration of row dendrigram (Code Block 8)

#dendextend used below  
#row\_dend = as.dendrogram(hclust(dist(mat)))  
#row\_dend = as.dendrogram(hclust(dist(mat), method = "single"))  
row\_dend = as.dendrogram(hclust(dist(mat, method = "maximum")), method = "complete")  
row\_dend = color\_branches(row\_dend, k = 10) # `color\_branches()` returns a dendrogram object  
  
split = data.frame(cutree(hclust(dist(mat, method = "maximum")), h=2))

## Evaluation of different methods based on how it looked to be devided

cluster\_rows = function(mat) color\_branches(as.dendrogram(diana(mat)), k = 10), #Same as above dendrogram cluster\_rows = function(mat) color\_branches(as.dendrogram(diana(mat)), k = 10), #Looks good cluster\_rows = color\_branches(as.dendrogram(hclust(dist(mat, method = “euclidean”)), method = “complete”), k = 10) #Looks good cluster\_rows = color\_branches(as.dendrogram(hclust(dist(mat, method = “canberra”)), method = “complete”), k = 10) #Looks very logical cluster\_rows = color\_branches(as.dendrogram(hclust(dist(mat, method = “maximum”)), method = “complete”), k = 10) #Looks good

## (Code Block 9)

InCellTextSize = 5  
RowLabelTextCell = 5  
LegendName = "Legend"  
  
  
  
HeatmapTest <- Heatmap(  
 matrix = mat,   
 cluster\_rows = color\_branches(as.dendrogram(hclust(dist(mat, method = "canberra")), method = "complete"), k = 20),  
 row\_split = 12,  
 row\_gap = unit(1.75, 'mm'),  
 border = FALSE,   
   
 show\_row\_names = FALSE,  
   
 right\_annotation = rowAnnotation(  
 Col1 = anno\_text(Combined\_List\_Analysis\_Shrunk2$locus, location = unit(0.75, "mm"), gp = gpar(fontsize = RowLabelTextCell)),  
 Col2 = anno\_text(Combined\_List\_Analysis\_Shrunk2$gene, location = unit(0, "mm"), gp = gpar(fontsize = RowLabelTextCell)),  
 Col4 = anno\_text(Combined\_List\_Analysis\_Shrunk2$NucleotidePAO1Locus, location = unit(1, "mm"), gp = gpar(fontsize = RowLabelTextCell)),  
 # Col3 = anno\_text(Combined\_List\_Analysis\_Shrunk2$ProteinPAO1Locus, location = unit(0, "mm"), gp = gpar(fontsize = RowLabelTextCell)),  
 # Col5 = anno\_text(Combined\_List\_Analysis\_Shrunk2$NucleotidePAO1ID, location = unit(0, "mm"), gp = gpar(fontsize = RowLabelTextCell)),  
 gap = unit(2, "mm")),  
   
 cell\_fun = function(j, i, x, y, width, height, fill) {  
 if(mat[i, j] < 0 || mat[i, j] > 0)  
 grid.text(sprintf("%.1f", mat[i, j]), x, y, gp = gpar(fontsize = InCellTextSize))},  
   
 col = colorSetRedBlue,  
   
 column\_dend\_reorder = FALSE,   
 column\_order = ColOrder,   
 row\_title = "Transcribed Gene Loci",   
 column\_title = "Conditions tested (CF39S vs CF39)",   
 column\_title\_side = "bottom",  
 show\_row\_dend = FALSE,  
 #row\_dend\_width = unit(6, "cm"),  
 row\_names\_gp = gpar(fontsize = RowLabelTextCell),  
   
 heatmap\_legend\_param = list(title = LegendName)  
   
 , raster\_quality = 10  
 )  
  
  
HeatmapTest

## One possible graphical output name, name was chosen based on parameters (Code Block 10)

pdf("heatmap\_Canberra(mat)AsDend\_colored\_split\_CellLabled\_AddCol\_NoDend\_MinProt.pdf", width = 10, height = 22)  
HeatmapTest  
dev.off()

## Final Heatmap generated (Code Block 11)

pdf("FullHeatMap-Lables-NoDend.pdf", width = 6, height = 30)  
HeatmapTest  
dev.off()

## One possible graphical output name, name was chosen based on parameters (Code Block 12)

png("FullHeatMap-MinLables-NoDend.png", width = 10, height = 22, units = "in", res = 300)  
HeatmapTest  
dev.off()

## One possible graphical output name, name was chosen based on parameters (Code Block 13)

pdf("heatmap\_Canberra(mat)AsDend\_colored\_split\_CellLabled\_WithAdditionalColumns\_MinProt.pdf", width = 12, height = 22)  
HeatmapTest  
dev.off()

## Session Info (Code Block 14)

sink(file = "./SessionInfo.txt")  
sessionInfo()  
sink(file = NULL)

## References (Code Block 15)

The citation function can be used to who you should be citing.

sink(file = "./ReferenceInfo.txt")  
citation("readr")  
citation("ComplexHeatmap")  
citation("circlize")  
citation("cluster")  
sink(file = NULL)