

# Transcriptomic and physiological responses to seasonal and diurnal cycles in *Ostreococcus*

Francisco J. Romero-Campero & Ana B. Romero-Losada

August 11, 2020

## Experimental design

In this project we study the transcriptomic and physiological responses to seasonal and diurnal cycles in the picoeukaryote *Ostreococcus tauri*. Our favourite microalgae was grown in 1.8L column photochemostats under long day conditions (16 hours light : 8 hours dark) representing a summer day and under short day conditions (8 hours light : 16 hours dark) simulating a winter day. After four weeks of entrainment under each condition samples were collected for three days every four hours. Then the program controlling the light in the photochemostats was set to free running conditions consisting on continuous light and samples were again collected every four hours for two days.

## Load data and principal component analysis.

The matrix containing the gene expression data analyzed in this study can be downloaded from **this link from the GEO data base**. Make sure to uncompress this file and rename it to **gene\_expression.tsv**. First, the gene expression data is loaded and converted into a numeric matrix setting the rownames to gene ids.

```
gene.expression <- read.table(file = "gene_expression.tsv", sep = "\t", header = T, as.is = T)
head(gene.expression[, 1:7])
```

```
##           X  ld_zt00_1  ld_zt04_1  ld_zt08_1  ld_zt12_1  ld_zt16_1  ld_zt20_1
## 1 ostta01g00010   9.664395  25.045279  30.81788  44.97020  32.61381  30.13931
## 2 ostta01g00020  15.688867   9.913202  10.36669  14.64151  11.66096  21.09974
## 3 ostta01g00030  16.108133  11.134813  13.85848  38.54982  24.16540  16.82981
## 4 ostta01g00040  59.247765  32.837433  27.26293  42.82092  52.59695  87.45710
## 5 ostta01g00050  27.909069  14.945981  12.04561  20.26003  28.25942  25.53805
## 6 ostta01g00060 248.044205 145.486374  68.38238  66.96495 224.52632 246.65002
```

```
gene.ids <- gene.expression$X
```

```
gene.expression <- as.matrix(gene.expression[, 2:ncol(gene.expression)])
rownames(gene.expression) <- gene.ids
head(gene.expression[, 1:6])
```

```
##           ld_zt00_1  ld_zt04_1  ld_zt08_1  ld_zt12_1  ld_zt16_1  ld_zt20_1
## ostta01g00010   9.664395  25.045279  30.81788  44.97020  32.61381  30.13931
## ostta01g00020  15.688867   9.913202  10.36669  14.64151  11.66096  21.09974
## ostta01g00030  16.108133  11.134813  13.85848  38.54982  24.16540  16.82981
## ostta01g00040  59.247765  32.837433  27.26293  42.82092  52.59695  87.45710
## ostta01g00050  27.909069  14.945981  12.04561  20.26003  28.25942  25.53805
## ostta01g00060 248.044205 145.486374  68.38238  66.96495 224.52632 246.65002
```

The current version of *Ostreococcus tauri* genome available from **here** identifies 7668 genes. In our experiment only 8 genes were never expressed and 40 genes never presented an expression level greater than 1 FPKM. This shows that practically the entire transcriptome of *Ostreococcus* is expressed under the seasonal and diurnal cycles studied in this project.

```
number.genes <- nrow(gene.expression)
number.genes
```

```
## [1] 7668
```

```
length(which(apply(X = gene.expression,MARGIN = 1,FUN = max) == 0))
```

```
## [1] 8
```

```
length(which(apply(X = gene.expression,MARGIN = 1,FUN = max) < 1))
```

```
## [1] 40
```

We perform **Principal Component Analysis** and a **Hierarchical clustering** in order to uncover the underlying structure in our data.

```
library(FactoMineR)
library(factoextra)
```

```
## Loading required package: ggplot2
```

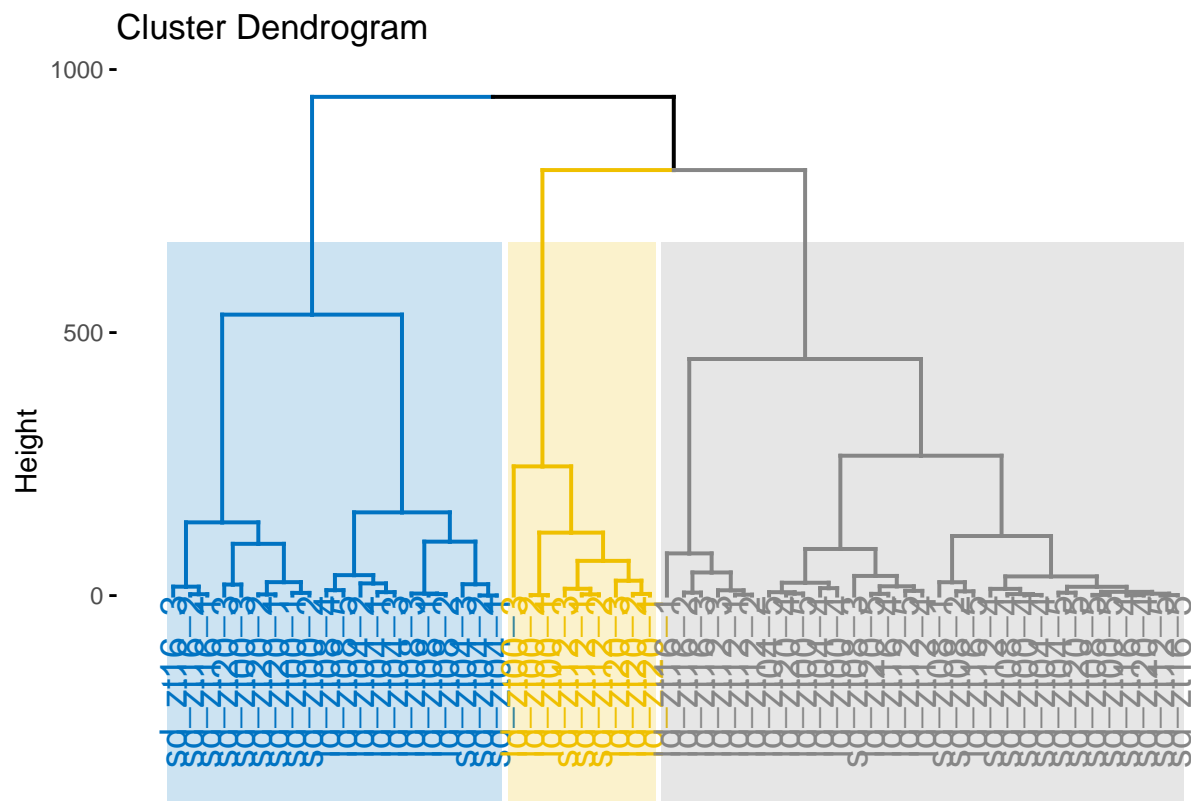
```
## Welcome! Want to learn more? See two factoextra-related books at https://goo.gl/ve3WBa
```

```
pca.gene.expression <- data.frame(colnames(gene.expression),t(gene.expression))
colnames(pca.gene.expression)[1] <- "Time point"
```

```
res.pca <- PCA(pca.gene.expression, graph = FALSE,scale.unit = TRUE,quali.sup = 1 )
```

```
res.hcpc <- HCPC(res.pca, graph=FALSE)
```

```
fviz_dend(res.hcpc,k=3,
  cex = 1, # Label size
  palette = "jco", # Color palette see ?ggpubr::ggpar
  rect = TRUE, rect_fill = TRUE, # Add rectangle around groups
  rect_border = "jco", # Rectangle color
  type="rectangle",
  labels_track_height = 400 # Augment the room for labels
)
```



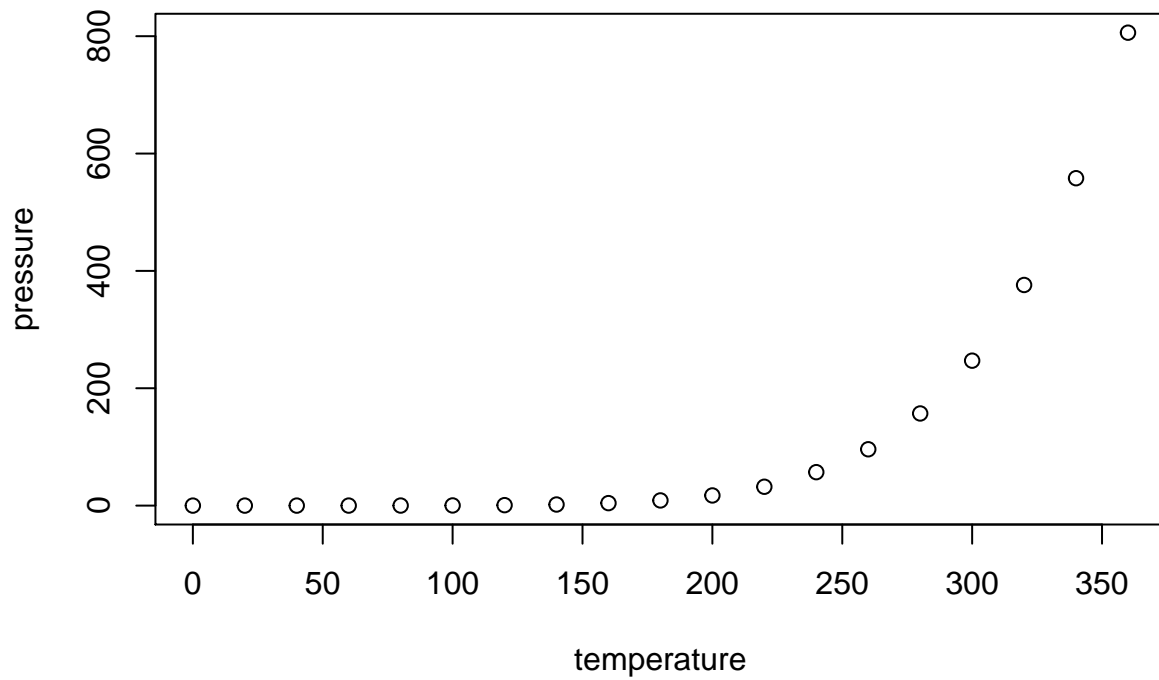
When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:

```
summary(cars)
```

```
##      speed      dist
##  Min.   : 4.0    Min.   : 2.00
##  1st Qu.:12.0    1st Qu.: 26.00
##  Median :15.0    Median : 36.00
##  Mean   :15.4    Mean   : 42.98
##  3rd Qu.:19.0    3rd Qu.: 56.00
##  Max.   :25.0    Max.   :120.00
```

## Including Plots

You can also embed plots, for example:



Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the plot.