Quantifying relic DNA

Jay T. Lennon, Mario E. Muscarella

05 October, 2017

Analysis of quantitative PCR data to test whether the abundance of bacterial communities is affected by relic DNA

Setup Work Environment

```
#rm(list=ls())
getwd()
setwd("~/GitHub/relicDNA/code")
require("plyr")
require("grid")
require("png")
require("car")
require("bbmle")
sem <- function(x, ...){sd(x, na.rm = TRUE)/sqrt(length(na.omit(x)))}
```

Load data and calcualte corrected copy number

```
eDNA.raw <- read.table("../data/eDNA_qPCR.txt", sep = "\t", header = T)

# Correct for dilutions and sample processing

# eDNA.raw[,7] = copies not corrected by dilution factor
# eDNA.raw[,8] = dilution factor
# eDNA.raw[,9] = volume (uL) in supernatant of phenol-chloroform extraction
# eDNA.raw[,10] = volume (ul) from supernatant of phenol-chlorofom subsampled

copies.corr <- eDNA.raw[,7] * (eDNA.raw[,8] * (eDNA.raw[,9]/eDNA.raw[,10]))

# Make new dataframe with corrected copy numbers
eDNA.corr <- data.frame(eDNA.raw, copies.corr)
```

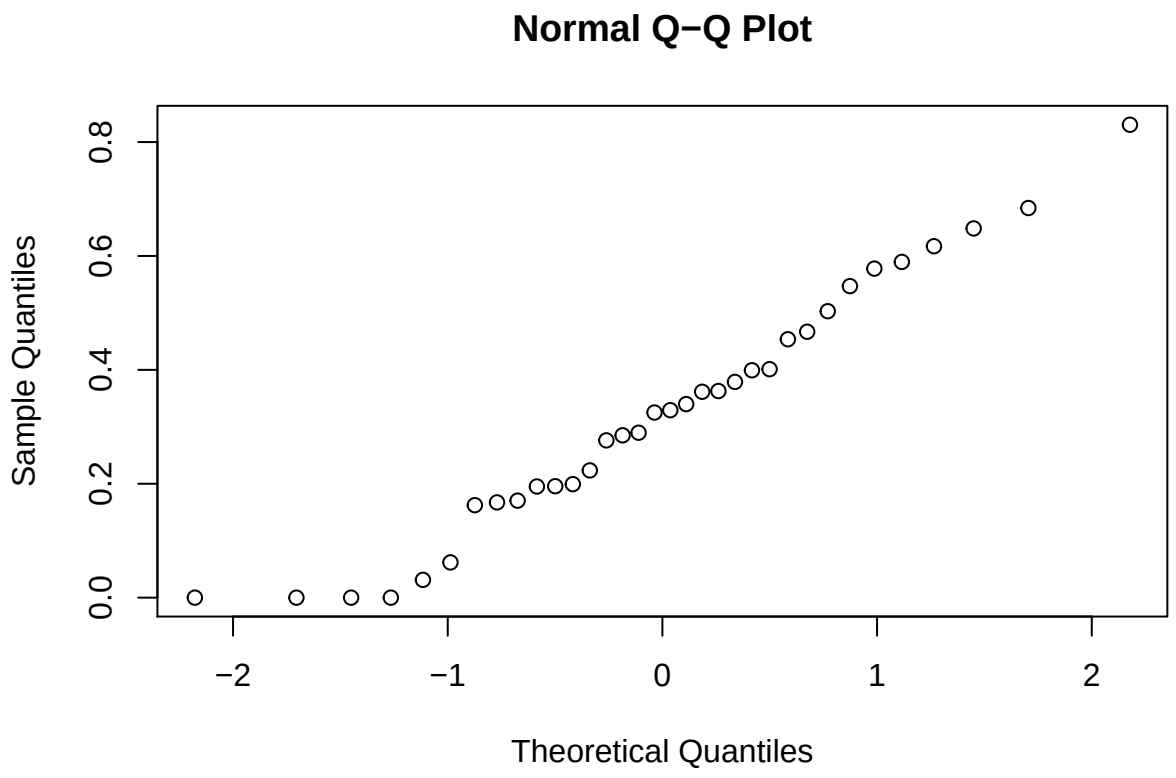
Take mean of technical replicates and sort

```
# Use `aggregate` to return means of subsamples taken for each sample
eDNA <- aggregate(eDNA.corr$copies.corr ~ eDNA.corr$sample + eDNA.corr$sample_name +
                  eDNA.corr$env + eDNA.corr$treat, eDNA.corr, mean)

# Sort by sample number
eDNA <- eDNA[order(eDNA[,1]) ,]
```

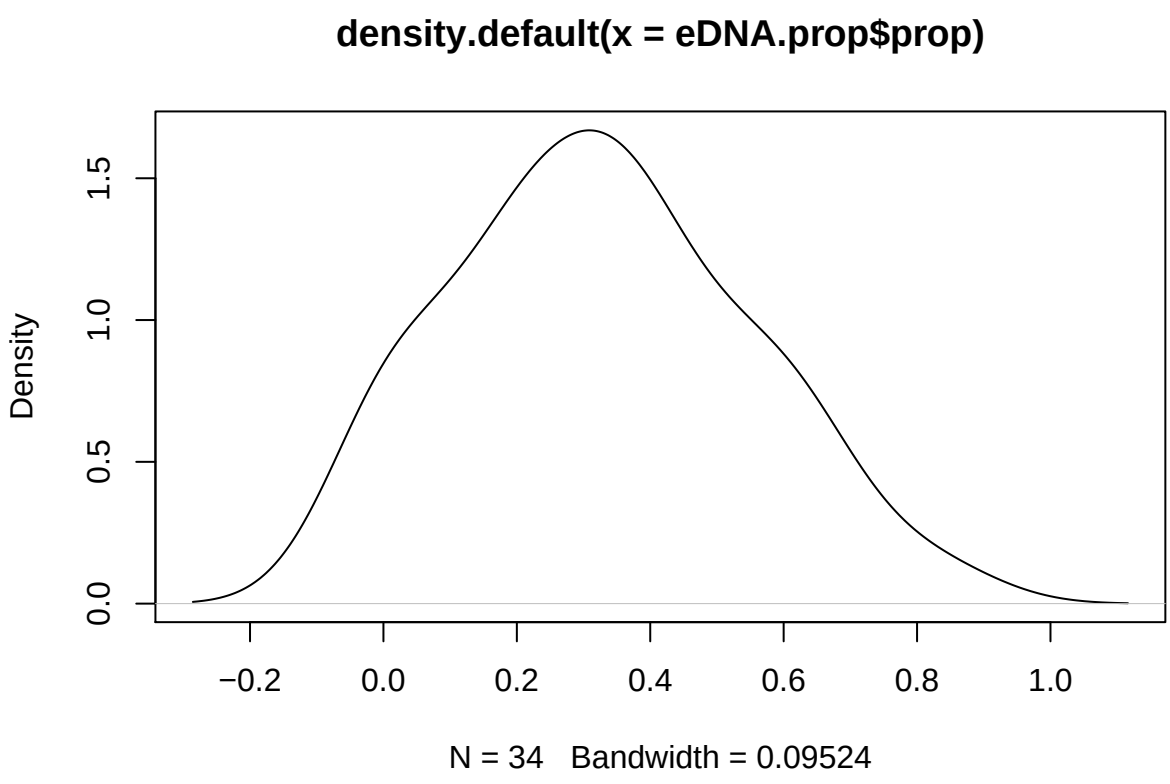
```
# Rename columns
colnames(eDNA) <- c("sample.number", "sample.name", "env", "treat", "copy.number")
```

Calculate proportion of degradable DNA per sample and test differences

```
# Use `ddply` to return the DNase-1 degradable proportion of 16S rRNA gene copy
eDNA.prop <- ddply(eDNA, .(sample.number, sample.name, env), summarize,
                   prop = 1 - ((copy.number[treat == "E"]) / (copy.number[treat == "C"])))

# Sort by environment
eDNA.prop <- eDNA.prop[order(eDNA.prop[,3]) ,]

# Three samples (cat feces [32], human feces [28], T7Core) have negative proportions
# Set these to zero (i.e., no eDNA)
eDNA.prop$prop <- ifelse(eDNA.prop$prop < 0, 0, eDNA.prop$prop)
write.table(eDNA.prop, "../data/relicDNA.prop.txt", sep = "\t", col.names = T, row.names = F)

# Distribution of eDNA
qqnorm(eDNA.prop$prop)
```




```
plot(density(eDNA.prop$prop))
```




```r
mean(eDNA.prop$prop)
```

```
## [1] 0.3256821
```

```r
sd(eDNA.prop$prop)
```

```
## [1] 0.2175141
```

```r
min(eDNA.prop$prop)
```

```
## [1] 0
```

```r
max(eDNA.prop$prop)
```

```
## [1] 0.8304294
```

```r
# Use glm to test whether amount of eDNA differs among environments
eDNA.prop.test <- glm(prop ~ env, data = eDNA.prop)
summary(eDNA.prop.test)
```

```
##
## Call:
## glm(formula = prop ~ env, data = eDNA.prop)
##
## Deviance Residuals:
##      Min        1Q    Median        3Q       Max
## -0.35852  -0.15517  -0.01063   0.13603   0.39788
##
## Coefficients:
```

```
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.35852    0.07234   4.956 2.64e-05 ***
## envsed      -0.04905    0.10230  -0.479   0.6351
## envsoil     -0.18304    0.10230  -1.789   0.0837 .
## envwater     0.07403    0.09705   0.763   0.4515
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 0.04186231)
##
##     Null deviance: 1.5613  on 33  degrees of freedom
## Residual deviance: 1.2559  on 30  degrees of freedom
## AIC: -5.6623
##
## Number of Fisher Scoring iterations: 2
```

```
Anova(eDNA.prop.test, type = "II", test.statistic = "F")
```

```
## Analysis of Deviance Table (Type II tests)
##
## Response: prop
## Error estimate based on Pearson residuals
##
##                SS Df      F  Pr(>F)
## env       0.30544  3 2.4321 0.08444 .
## Residuals 1.25587 30
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# Use Anova to test whether the amount of dDNA differs among environments
eDNA.prop.lm <- lm(prop ~ env, data = eDNA.prop)
summary(eDNA.prop.lm)
```

```
##
## Call:
## lm(formula = prop ~ env, data = eDNA.prop)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.35852 -0.15517 -0.01063  0.13603  0.39788
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.35852    0.07234   4.956 2.64e-05 ***
## envsed      -0.04905    0.10230  -0.479   0.6351
## envsoil     -0.18304    0.10230  -1.789   0.0837 .
## envwater     0.07403    0.09705   0.763   0.4515
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2046 on 30 degrees of freedom
## Multiple R-squared:  0.1956, Adjusted R-squared:  0.1152
## F-statistic: 2.432 on 3 and 30 DF,  p-value: 0.08444
```

```
eDNA.anova <- Anova(eDNA.prop.lm, type = "II")
eDNA.anova
```

```
## Anova Table (Type II tests)
##
## Response: prop
##            Sum Sq Df F value  Pr(>F)
## env       0.30544  3  2.4321 0.08444 .
## Residuals 1.25587 30
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
TukeyHSD(aov(eDNA.prop.lm), "env")
```

```
##   Tukey multiple comparisons of means
##     95% family-wise confidence level
##
## Fit: aov(formula = eDNA.prop.lm)
##
## $env
##                     diff          lwr        upr     p adj
## sed-feces   -0.04905302 -0.327221734 0.22911570 0.9630471
## soil-feces  -0.18304416 -0.461212879 0.09512455 0.2981258
## water-feces  0.07403322 -0.189860799 0.33792723 0.8704386
## soil-sed    -0.13399114 -0.412159862 0.14417757 0.5640774
## water-sed    0.12308623 -0.140807782 0.38698025 0.5896087
## water-soil   0.25707738 -0.006816637 0.52097139 0.0583852
```

```
# Calculate means, sem, and sample size by enviroment
eDNA.mean <- aggregate(eDNA.prop$prop ~ eDNA.prop$env, eDNA.prop, mean)
eDNA.n <- aggregate(eDNA.prop$prop ~ eDNA.prop$env, eDNA.prop, length)
eDNA.sem <- aggregate(eDNA.prop$prop ~ eDNA.prop$env, eDNA.prop, sem)

# Make table of proportion eDNA by environment
eDNA.table <- data.frame(eDNA.mean, eDNA.sem[ ,2], eDNA.n[ ,2])
colnames(eDNA.table) <- c("env", "mean", "sem", "n")
eDNA.table <- eDNA.table[order(eDNA.table[,2]), ]
```

```
# # Make bar plot with error bars by environment
# ```{r, eval=F}
# png(filename="../figures/qPCR.bar.png",
#     width = 800, height = 800, res = 96*2)
#
# bp <- barplot(eDNA.table$mean, ylim =c(0, 0.6),
#               pch = 15, cex = 1.25, las = 1, cex.lab = 1.25, cex.axis = 1,
#               col = "gray90", axis.lty = 1, lwd = 2, xlab = NA,
#               ylab = "Proportion eDNA",
#               names.arg = c("Soil", "Sediment", "Gut", "Water"), cex.names = 0.9)
#               box(lwd = 2)
# arrows(x0 = bp, y0 = eDNA.table$mean, y1 = eDNA.table$mean - eDNA.table$sem,
#        angle = 90,  length = 0.1, lwd = 2)
# arrows(x0 = bp, y0 = eDNA.table$mean, y1 = eDNA.table$mean + eDNA.table$sem,
#        angle = 90, length=0.1, lwd = 2)
#
# # Close Plot Device
# dev.off()
# graphics.off()
#
```

```
# # Show Plot
# img <- readPNG("../figures/qPCR.bar.png")
# grid.raster(img)
```

```
# Make x-y plot with error bars by environment
# ```{r}
# png(filename="../figures/Figure2-Prop_eDNA.png",
#     width = 800, height = 800, res = 96*2)
#
# par(mar = c(3, 5, 1, 1))
# non.bp <- plot(eDNA.table$mean, ylim = c(0, 0.6),
#                xlim = c(0.5, 4.5), pch = 22, bg = "gray90", lwd = 2,
#                cex = 3, yaxt = "n", xaxt = "n", cex.lab = 2, cex.axis = 1.5,
#                las = 1, ylab = "", xlab = "")
# box(lwd = 2)
#
# mtext(expression('Proportion Relic DNA'), side = 2,
#       outer = FALSE, cex = 1.5, line = 3, adj = 0.5)
#
# # Major Axes
# axis(side = 2, lwd.ticks = 2, cex.axis = 1.25, las = 1,
#      labels = c(0.0, 0.2, 0.4, 0.6), at = c(0.0, 0.2, 0.4, 0.6))
#
# axis(side = 4, lwd.ticks = 2, cex.axis = 1.5, las = 1,
#      at=c(0.0, 0.2, 0.4, 0.6), labels = F, tck = -0.02)
#
# axis(side = 1, lwd.ticks = 2, cex.axis = 0.9, las = 1,
#      labels = c("Soil", "Sediment", "Gut", "Water"), at = c(1, 2, 3, 4))
#
# axis(side = 3, lwd.ticks = 2, cex.axis = 1.5, las = 1,
#      at = c(1, 2, 3, 4), labels = F, tck = -0.02)
#
# axis(side = 1, labels = F, lwd.ticks = 2, tck = 0.02, at = c(1, 2, 3, 4))
# axis(side = 2, labels = F, lwd.ticks = 2, tck = 0.02, at = c(0, 0.2, 0.4, 0.6))
# axis(side = 3, labels = F, lwd.ticks = 2, tck = 0.02, at = c(1, 2, 3, 4))
# axis(side = 4, labels = F, lwd.ticks = 2, tck = 0.02, at = c(0, 0.2, 0.4, 0.6))
#
# arrows(x0 = c(1, 2, 3, 4), y0 = eDNA.table$mean,
#        y1 = eDNA.table$mean - eDNA.table$sem, angle = 90,
#        length = 0.1, lwd = 2)
#
# arrows(x0 = c(1,2,3,4), y0 = eDNA.table$mean,
#        y1 = eDNA.table$mean + eDNA.table$sem, angle = 90,
#        length=0.1, lwd = 2)
#
# points(x = c(1:4), eDNA.table$mean,
#       pch = 22, bg = "gray90", lwd = 2, cex = 3)
#
# # Close Plot Device
# dev.off()
# graphics.off()
#
# # Show Plot
# img <- readPNG("../figures/Figure2-Prop_eDNA.png")
```

```
# grid.raster(img)
```

Make x-y plot with error bars by environment with individual data points

```
# Relic DNA data for plotting
relic.soil <- eDNA.prop[ which(eDNA.prop$env == "soil"),]
relic.sed <- eDNA.prop[ which(eDNA.prop$env == "sed"),]
relic.feces <- eDNA.prop[ which(eDNA.prop$env == "feces"),]
relic.water <- eDNA.prop[ which(eDNA.prop$env == "water"),]

# Relic DNA table
relic.mean <- aggregate(eDNA.prop$prop ~ env, eDNA.prop, mean)
relic.sem <- aggregate(eDNA.prop$prop ~ env, eDNA.prop, sem)
relic.95.LL <- aggregate(prop ~ env, eDNA.prop,
          FUN = function(x) t.test(x)$conf.int[1])
relic.95.UL <- aggregate(prop ~ env, eDNA.prop,
          FUN = function(x) t.test(x)$conf.int[2])
relic.table <- data.frame(relic.mean[1], relic.mean[2], relic.sem[2],
          relic.95.LL[2], relic.95.UL[2])
colnames(relic.table) <- c("env", "mean", "sem", "LCI", "UCI")
relic.table <- relic.table[order(relic.table$mean),]

png(filename="../figures/Figure2-Prop.relic.png",
    width = 800, height = 800, res = 96*2)

par(mar = c(4, 5, 1, 1))

non.bp.rich <- plot(jitter(rep(1, length(relic.soil$prop)), amount = 0.1), relic.soil$prop,
      ylim = c(-0.1, 1), xlim = c(0.5, 4.5), pch = 21, col = "lightgrey", bg = "lightgrey", lwd = 2,
      cex = 1.7, yaxt = "n", xaxt = "n", cex.lab = 2, cex.axis = 1.5,
      las = 1, ylab = "", xlab = "")
      box(lwd = 2)
points(jitter(rep(2, length(relic.sed$prop)), amount = 0.1), relic.sed$prop, pch = 21,
       bg = "lightgrey", col = "lightgrey", lwd = 2, cex = 1.7)
points(jitter(rep(3, length(relic.feces$prop)), amount = 0.1), relic.feces$prop, pch = 21,
       bg = "lightgrey", col = "lightgrey", lwd = 2, cex = 1.7)
points(jitter(rep(4, length(relic.water$prop)), amount = 0.1),relic.water$prop, pch = 21,
       bg = "lightgrey", col = "lightgrey", lwd = 2, cex = 1.7)

# rect(xleft, ybottom, xright, ytop)
rect(0.8, (mean(relic.soil$prop) + 0.025), 1.2, (mean(relic.soil$prop)- 0.025), col = "NA", border = "bl

rect(1.8, (mean(relic.sed$prop) + 0.025), 2.2, (mean(relic.sed$prop)- 0.025), col = "NA", border = "bla

rect(2.8, (mean(relic.feces$prop) + 0.025), 3.2, (mean(relic.feces$prop) - 0.025), col = "NA", border =

rect(3.8, (mean(relic.water$prop) + 0.025), 4.2, (mean(relic.water$prop) - 0.025), col = "NA", border =

box(lwd = 2)
```

```
mtext(expression('Proportion Relic DNA'), side = 2,
      outer = FALSE, cex = 1.5, line = 3.6, adj = 0.5)

# Major Axes
axis(side = 2, lwd.ticks = 2, cex.axis = 1.25, las = 1,
     labels = c("0.0", "0.25", "0.50", "0.75", "1.00"), at = c(0.0, 0.25, 0.5, 0.75, 1.0))

axis(side = 4, lwd.ticks = 2, cex.axis = 1.5, las = 1,
     at=c(0.0, 0.25, 0.5, 0.75, 1.0), labels = F, tck = -0.02)

axis(side = 1, lwd.ticks = 2, cex.axis = 0.9, las = 1,
     labels = c("Soil", "Sediment", "Gut", "Water"), at = c(1, 2, 3, 4))

axis(side = 3, lwd.ticks = 2, cex.axis = 1.5, las = 1,
     at = c(1, 2, 3, 4), labels = F, tck = -0.02)

arrows(x0 = c(1,2,3,4), y0 = relic.table$mean, y1 = relic.table$LCI, angle = 90,
       length = 0.1, lwd = 2)
arrows(x0 = c(1,2,3,4), y0 = relic.table$mean, y1 = relic.table$UCI, angle = 90,
       length=0.1, lwd = 2)

# Close Plot Device
dev.off()
```

```
## pdf
##   2
```

```
graphics.off()

# Show Plot
img <- readPNG("../figures/Figure2-Prop.relic.png")
grid.raster(img)
```