

Figure 5a

```
library(readxl)
library(lubridate)

##
## Attaching package: 'lubridate'
## The following objects are masked from 'package:base':
##
##   date, intersect, setdiff, union
library(mgcv)

## Loading required package: nlme
## This is mgcv 1.8-34. For overview type 'help("mgcv-package")'.
library(ggplot2)
library(ggsignif)
library(readr)
library(emmeans)
library(tidyr)
library(plyr)
library(dplyr)

##
## Attaching package: 'dplyr'
## The following objects are masked from 'package:plyr':
##
##   arrange, count, desc, failwith, id, mutate, rename, summarise,
##   summarize
## The following object is masked from 'package:nlme':
##
##   collapse
## The following objects are masked from 'package:stats':
##
##   filter, lag
## The following objects are masked from 'package:base':
##
##   intersect, setdiff, setequal, union
#load in raw data
dat <- data.frame(read_csv("5a_data.csv"))

##
## -- Column specification -----
## cols(
##   `Repeat(Date)` = col_datetime(format = ""),
##   Algae = col_character(),
```

```
## Treatment = col_character(),
## Well_ID = col_character(),
## Larvae = col_double(),
## `Number of Symbionts` = col_double()
## )
```

```
head(dat)
```

```
## Repeat.Date. Algae Treatment Well_ID Larvae Number.of.Symbionts
## 1 2019-02-18 SSB01 control S_c_2019-02-18 1 0
## 2 2019-02-18 SSB01 control S_c_2019-02-18 2 0
## 3 2019-02-18 SSB01 control S_c_2019-02-18 3 1
## 4 2019-02-18 SSB01 control S_c_2019-02-18 4 2
## 5 2019-02-18 SSB01 control S_c_2019-02-18 5 0
## 6 2019-02-18 SSB01 control S_c_2019-02-18 6 0
```

```
#clean up raw data
```

```
dat$Algae <- factor(dat$Algae, levels= c("SSB01", "M. gaditana", "N. oculata", "C. velia"))
dat$Treatment <- factor(dat$Treatment)
dat$Repeat.Date. <- factor(as.character(dat$Repeat.Date.))
dat$Number.of.Symbionts <- ifelse(dat$Number.of.Symbionts>1,1,dat$Number.of.Symbionts)
dat$Well_ID <- factor(dat$Well_ID)
head(dat)
```

```
## Repeat.Date. Algae Treatment Well_ID Larvae Number.of.Symbionts
## 1 2019-02-18 SSB01 control S_c_2019-02-18 1 0
## 2 2019-02-18 SSB01 control S_c_2019-02-18 2 0
## 3 2019-02-18 SSB01 control S_c_2019-02-18 3 1
## 4 2019-02-18 SSB01 control S_c_2019-02-18 4 1
## 5 2019-02-18 SSB01 control S_c_2019-02-18 5 0
## 6 2019-02-18 SSB01 control S_c_2019-02-18 6 0
```

```
#model data with generalised linear mixed modes
```

```
summary <- data.frame()
```

```
for(i in unique(dat$Algae))
```

```
{
```

```
#i <- unique(dat$Algae)[1]
```

```
Dat <- droplevels(dat[dat$Algae==i,])
```

```
##Regression model:
```

```
mod <- gam(Number.of.Symbionts ~ Treatment + s(Repeat.Date.,bs='re') + s(Well_ID,bs='re'),data=Dat,fa
```

```
S <- summary(mod)$p.table
```

```
if(S[2,4]>0.1)
```

```
{
```

```
  p <- ""
```

```
}
```

```
if(S[2,4]<0.1)
```

```
{
```

```
  p <- ""
```

```
}
```

```
if(S[2,4]<0.05)
```

```
{
```

```
  p <- "+"
```

```
}
```

```
if(S[2,4]<0.01)
```

```
{
```

```

    p <- "***"
  }
  if(S[2,4]<0.001)
  {
    p <- "***"
  }
  newline <- data.frame(Algae=i,outcome=S[1,1],SE=S[1,2],type="control",p=NA,pvalue=NA)
  summary <- rbind(summary,newline)
  newline <- data.frame(Algae=i,outcome=S[1,1] + S[2,1],SE=S[2,2],type="LPS",p=p, pvalue=S[2,4])
  summary <- rbind(summary,newline)
}

summary$up <- plogis(summary$outcome + 1.96*summary$SE)
summary$down <- plogis(summary$outcome - 1.96*summary$SE)
summary$outcome <- plogis(summary$outcome)

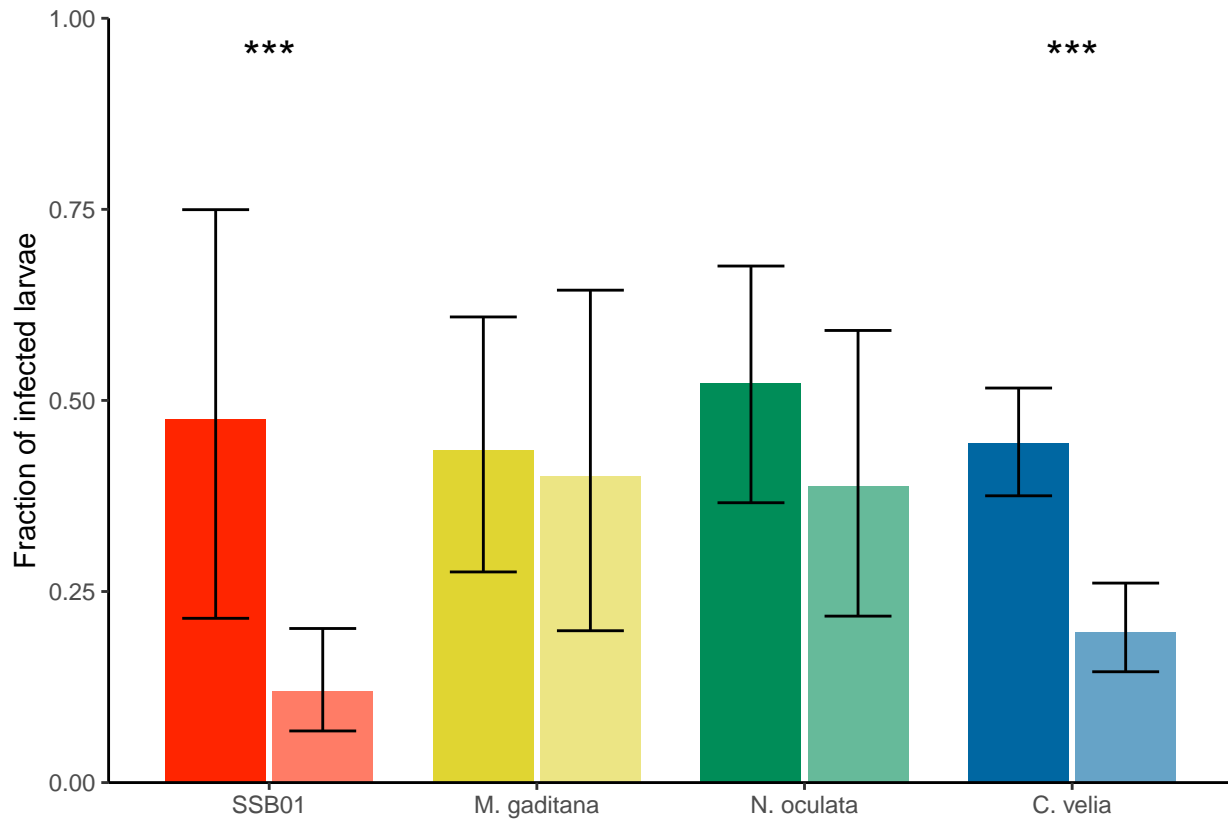
summary

##      Algae  outcome      SE  type    p      pvalue      up
## 1    SSB01 0.4751411 0.6102174 control <NA>      NA 0.7496054
## 2    SSB01 0.1191243 0.3186042    LPS   *** 2.410956e-09 0.2016063
## 3 M. gaditana 0.4350628 0.3598266 control <NA>      NA 0.6092203
## 4 M. gaditana 0.4012222 0.5072419    LPS      7.838317e-01 0.6442409
## 5 N. oculata 0.5232108 0.3274234 control <NA>      NA 0.6758270
## 6 N. oculata 0.3883861 0.4206224    LPS      1.934363e-01 0.5915360
## 7 C. velia 0.4445879 0.1466255 control <NA>      NA 0.5162002
## 8 C. velia 0.1966197 0.1871960    LPS   *** 2.447937e-10 0.2610241
##      down
## 1 0.21491597
## 2 0.06753324
## 3 0.27558185
## 4 0.19867929
## 5 0.36613362
## 6 0.21780208
## 7 0.37520556
## 8 0.14498842

write_csv(summary, "5a_summary.csv")

summary$Algae <- factor(summary$Algae, levels= c("SSB01", "M. gaditana", "N. oculata", "C. velia"))
#Plot graph
P1<-ggplot(data = summary, aes(x = Algae, y = outcome, group = type)) +
  geom_bar(stat = "identity", width = 0.75, size=1.5, position = position_dodge(width = 0.8), fill = c
  geom_errorbar(aes(ymin=down, ymax=up), width=.5, position =position_dodge(.8)) +
  geom_text(data = summary, aes(x=Algae,y=0.95),label=summary$p,size=6) +
  theme_classic() +
  labs(x = "", y = "Fraction of infected larvae") +
  scale_y_continuous(expand = c(0, 0), limits = c(0,1), labels = scales::number_format(accuracy = 0.01))
plot(P1)

## Warning: Removed 4 rows containing missing values (geom_text).
```



```
ggsave("Figure_5a.pdf",width=16,height=11,units="cm")
```

```
## Warning: Removed 4 rows containing missing values (geom_text).
```

```
# get n count per treatment
```

```
ndat <- ddpby(dat, c("Algae", "Treatment", "Repeat.Date."), summarise,  
              N = length(Larvae))
```

```
n_Figure <- ddpby(ndat, c("Algae", "Treatment"), summarise,  
                  N = length(Repeat.Date.))
```

```
n_Figure
```

```
##      Algae Treatment N  
## 1      SSB01   control 5  
## 2      SSB01      LPS 5  
## 3 M. gaditana   control 5  
## 4 M. gaditana      LPS 5  
## 5 N. oculata    control 5  
## 6 N. oculata      LPS 5  
## 7  C. velia    control 5  
## 8  C. velia      LPS 5
```

```
write_csv(n_Figure, "5a_n.csv")
```