

Figure 5d

```
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("5d_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 2020-11-04 SSB01 CP S_C_2020-11-04 1 0
## 2 2020-11-04 SSB01 CP S_C_2020-11-04 2 2
## 3 2020-11-04 SSB01 CP S_C_2020-11-04 3 0
## 4 2020-11-04 SSB01 CP S_C_2020-11-04 4 0
## 5 2020-11-04 SSB01 CP S_C_2020-11-04 5 0
## 6 2020-11-04 SSB01 CP S_C_2020-11-04 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 2020-11-04 SSB01 CP S_C_2020-11-04 1 0
## 2 2020-11-04 SSB01 CP S_C_2020-11-04 2 1
## 3 2020-11-04 SSB01 CP S_C_2020-11-04 3 0
## 4 2020-11-04 SSB01 CP S_C_2020-11-04 4 0
## 5 2020-11-04 SSB01 CP S_C_2020-11-04 5 0
## 6 2020-11-04 SSB01 CP S_C_2020-11-04 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="CP",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="MIP",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

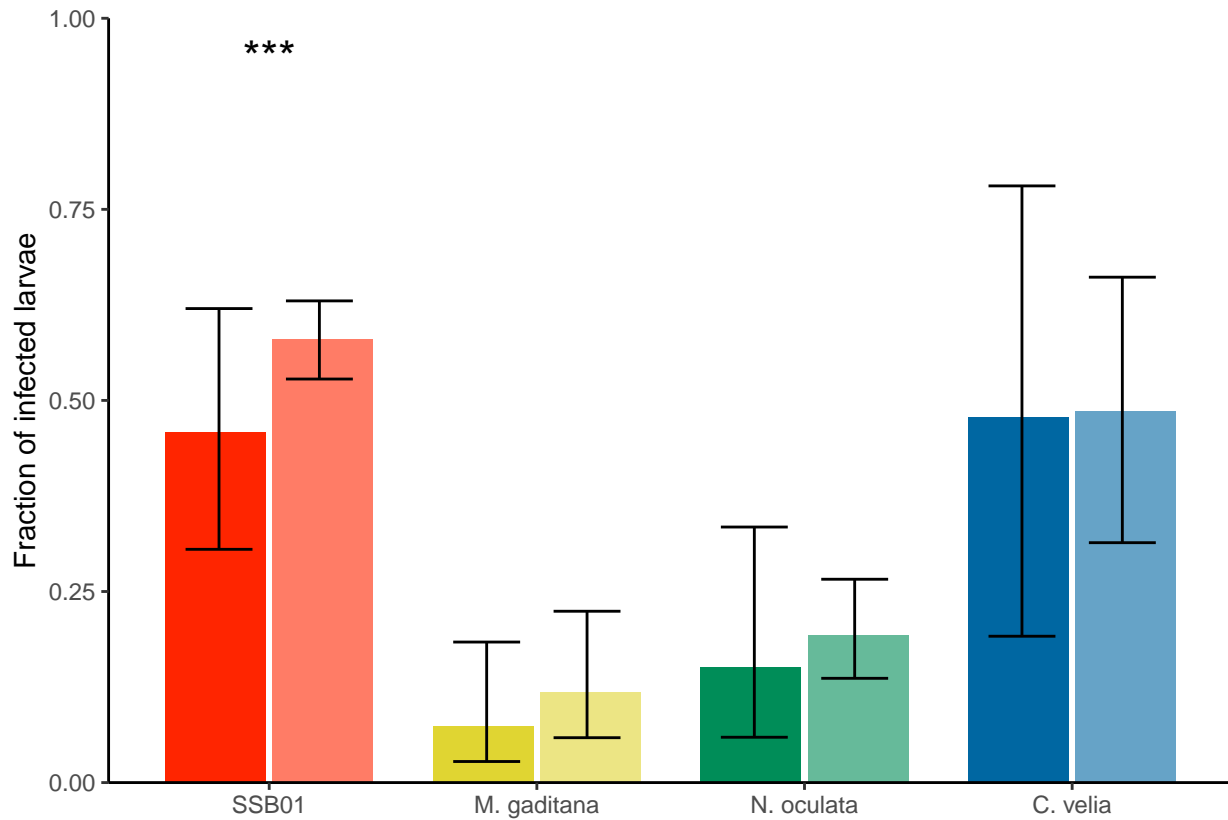
##      Algae  outcome      SE type  p      pvalue      up      down
## 1      SSB01 0.45854645 0.3349211  CP <NA>      NA 0.6201652 0.30520333
## 2      SSB01 0.57997779 0.1074291  MIP  *** 5.346688e-06 0.6302392 0.52800098
## 3 M. gaditana 0.07398296 0.5289841  CP <NA>      NA 0.1838843 0.02754865
## 4 M. gaditana 0.11835508 0.3911412  MIP      1.845819e-01 0.2241814 0.05870472
## 5 N. oculata 0.15107470 0.5294241  CP <NA>      NA 0.3343611 0.05930820
## 6 N. oculata 0.19309888 0.2118466  MIP      1.620613e-01 0.2660455 0.13643534
## 7 C. velia 0.47866616 0.6913653  CP <NA>      NA 0.7806895 0.19147364
## 8 C. velia 0.48586039 0.3701298  MIP      9.379494e-01 0.6612525 0.31388300

write_csv(summary, "5d_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(.75)) +
  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_5d.pdf",width=16,height=11,units="cm")
```

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

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

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

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

```
n_Figure
```

```
##      Algae Treatment  N  
## 1     SSB01        CP 10  
## 2     SSB01        MIP 10  
## 3 M. gaditana        CP  3  
## 4 M. gaditana        MIP  3  
## 5 N. oculata         CP 13  
## 6 N. oculata         MIP 13  
## 7  C. velia         CP  3  
## 8  C. velia         MIP  3
```

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