

Figure 3a

```
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("3a_data.csv"))

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

```
## Treatment = col_character(),
## Well_ID = col_character(),
## Larvae = col_character(),
## Time_to_first_expulsion = col_double(),
## `Well_ID (#1)` = col_character()
## )
```

```
head(dat)
```

```
## Repeat.Date.      Algae Treatment Well_ID Larvae Time_to_first_expulsion
## 1 2020-10-22 M. gaditana DMSO MD1 1 720
## 2 2020-10-22 M. gaditana DMSO MD1 2 720
## 3 2020-10-22 M. gaditana DMSO MD1 3 70
## 4 2020-10-22 M. gaditana DMSO MD1 4 165
## 5 2020-10-22 M. gaditana DMSO MD1 5 70
## 6 2020-10-22 M. gaditana DMSO MD1 6 70
## Well_ID...1.
## 1 M_D_2020-10-22
## 2 M_D_2020-10-22
## 3 M_D_2020-10-22
## 4 M_D_2020-10-22
## 5 M_D_2020-10-22
## 6 M_D_2020-10-22
```

```
#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$Well_ID <- factor(dat$Well_ID)
dat$expulsion <- ifelse(dat$Time_to_first_expulsion<720,1,0)
head(dat)
```

```
## Repeat.Date.      Algae Treatment Well_ID Larvae Time_to_first_expulsion
## 1 2020-10-22 M. gaditana DMSO MD1 1 720
## 2 2020-10-22 M. gaditana DMSO MD1 2 720
## 3 2020-10-22 M. gaditana DMSO MD1 3 70
## 4 2020-10-22 M. gaditana DMSO MD1 4 165
## 5 2020-10-22 M. gaditana DMSO MD1 5 70
## 6 2020-10-22 M. gaditana DMSO MD1 6 70
## Well_ID...1. expulsion
## 1 M_D_2020-10-22 0
## 2 M_D_2020-10-22 0
## 3 M_D_2020-10-22 1
## 4 M_D_2020-10-22 1
## 5 M_D_2020-10-22 1
## 6 M_D_2020-10-22 1
```

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

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

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

```
{
```

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

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

```
  ##Regression model:
```

```
  mod <- gam(expulsion ~ Treatment + s(Well_ID,bs='re'),data=Dat,family=binomial())
```

```
  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="DMSO",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="LatB",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

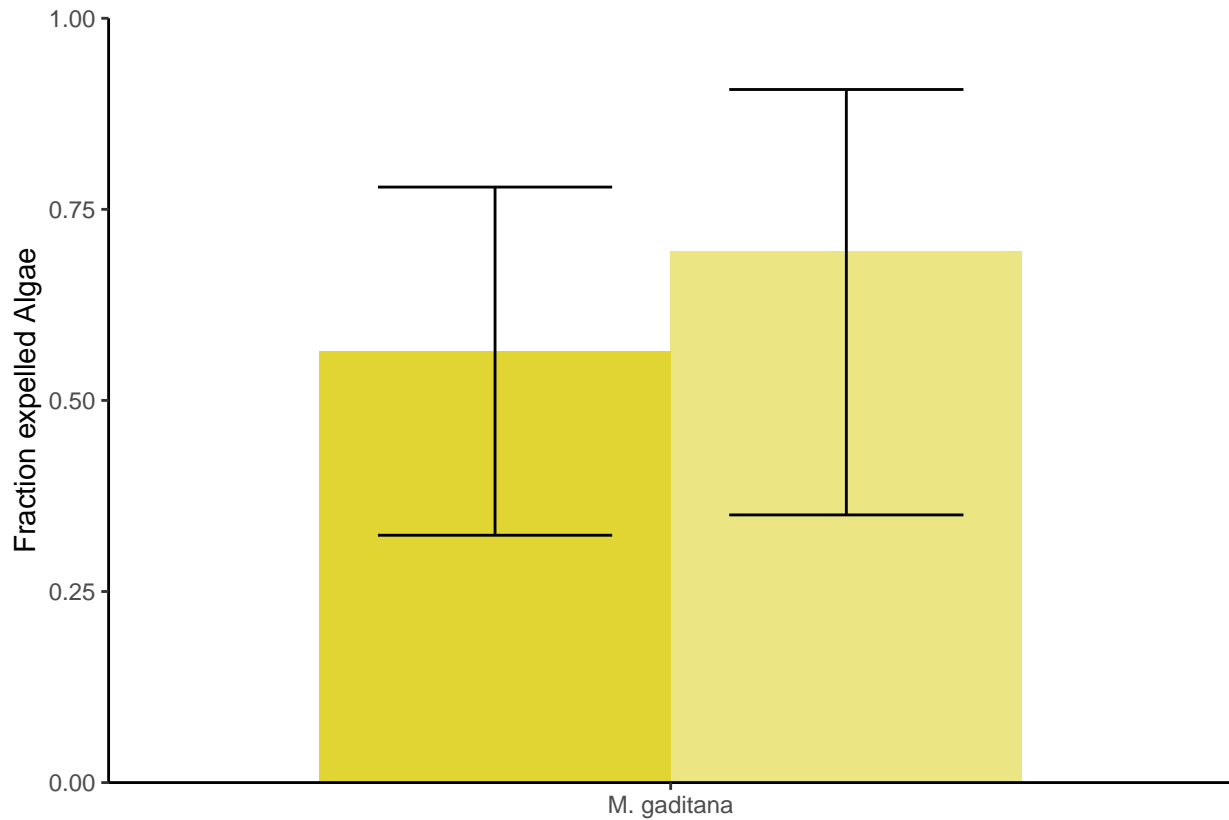
##           Algae  outcome          SE type    p    pvalue        up        down
## 1 M. gaditana 0.5650373 0.5096679 DMSO <NA>      NA 0.7791260 0.3235909
## 2 M. gaditana 0.6960358 0.7381354 LatB      0.4425098 0.9068053 0.3501786

write_csv(summary, "3a_summary.csv")

#Plot graph
P1 <- ggplot(data = summary,aes(x = Algae, y = outcome, group = type)) +
  geom_bar(stat = "identity", width = 0.75, size=1.5, position = "dodge", fill = c("#E0D532",alpha("#E0D532",0.5))) +
  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 expelled Algae") +
  scale_y_continuous(expand = c(0, 0), limits = c(0,1), labels = scales::number_format(accuracy = 0.01))
plot(P1)

## Warning: Removed 1 rows containing missing values (geom_text).

```



```
ggsave("Figure_3a.pdf",width=16,height=11,units="cm")
```

```
## Warning: Removed 1 rows containing missing values (geom_text).
```

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

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

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

```
n_Figure
```

```
##           Algae Treatment N
```

```
## 1 M. gaditana      DMSO 3
```

```
## 2 M. gaditana      LatB 3
```

```
write_csv(n_Figure, "3a_n.csv")
```