

# linear\_models on alpha hill numbers

Adam\_Koziol

4/21/2022

setting up the environment

Loading the data

```
#####For some reason the dataframe has duplicates - remove these
alpha_diversity_tax <- read.csv('../hill_numbers/outputsR/diversity_table.csv', header=T, row.names=1) %>%
  unique(.) %>%
  as.tibble
```

```
## Warning: 'as.tibble()' was deprecated in tibble 2.0.0.
## i Please use 'as_tibble()' instead.
## i The signature and semantics have changed, see '?as_tibble'.
```

```
#####Combined dataframe containing functional diversities
alpha_diversity_fd <- read.csv('../hill_numbers/outputsR/diversity_table_fd.csv', header=T, row.names = 1)
#####Phylogenetic diversities
alpha_diversity_pd <- read.csv('../hill_numbers/outputsR/diversity_table_pd.csv', header=T, row.names = 1)
#####Metadata
metadata <- read.csv('../data/metadata.csv')
```

Combining the metadata with the dataframes

```
alph_div_tax_as <- alpha_diversity_tax %>%
  filter(species == 'AS') %>%
  dplyr::select(c(diversity, sample_id)) %>%
  left_join(metadata, by='sample_id') %>%
  mutate(days=case_when(treatment == 'Day1' ~ 0,
                        treatment == 'Acc' ~ 7,
                        treatment == 'Heat' ~ 19,
                        treatment == 'Cold' ~ 33,
                        treatment == 'Diet' ~ 45)) %>%
  mutate(treatment = factor(treatment, levels = c('Day1', 'Acc', 'Heat', 'Cold', 'Diet')))

alph_div_tax_cr <- alpha_diversity_tax %>%
  filter(species == 'CR') %>%
  dplyr::select(c(diversity, sample_id)) %>%
  left_join(metadata, by='sample_id') %>%
  mutate(days=case_when(treatment == 'Day1' ~ 0,
                        treatment == 'Acc' ~ 7,
                        treatment == 'Heat' ~ 19,
                        treatment == 'Cold' ~ 33,
```

```

        treatment == 'Diet' ~ 45),
  cage=case_when(cage == 1 ~ 6,
                 cage == 2 ~ 7,
                 cage == 3 ~ 8,
                 cage == 4 ~ 9,
                 cage == 5 ~ 10,
                 cage == 6 ~ 11)) %>%
mutate_at(2:8, as.factor) %>%
mutate(treatment = factor(treatment, levels = c('Day1', 'Acc', 'Heat', 'Cold', 'Diet'))))

alph_div_fd_as <- alpha_diversity_fd %>%
  filter(species == 'AS') %>%
  dplyr::select(c(Q, sample_id)) %>%
  left_join(metadata, by='sample_id') %>%
  filter(treatment != 'FAST') %>%
  mutate(days=case_when(treatment == 'Day1' ~ 0,
                        treatment == 'Acc' ~ 7,
                        treatment == 'Heat' ~ 19,
                        treatment == 'Cold' ~ 33,
                        treatment == 'Diet' ~ 45)) %>%
  mutate_at(2:8, as.factor) %>%
  mutate(treatment = factor(treatment, levels = c('Day1', 'Acc', 'Heat', 'Cold', 'Diet'))))

alph_div_fd_cr <- alpha_diversity_fd %>%
  filter(species == 'CR') %>%
  dplyr::select(c(Q, sample_id)) %>%
  left_join(metadata, by='sample_id') %>%
  mutate(days=case_when(treatment == 'Day1' ~ 0,
                        treatment == 'Acc' ~ 7,
                        treatment == 'Heat' ~ 19,
                        treatment == 'Cold' ~ 33,
                        treatment == 'Diet' ~ 45),
         cage=case_when(cage == 1 ~ 6,
                        cage == 2 ~ 7,
                        cage == 3 ~ 8,
                        cage == 4 ~ 9,
                        cage == 5 ~ 10,
                        cage == 6 ~ 11)) %>%
  mutate_at(2:8, as.factor) %>%
  mutate(treatment = factor(treatment, levels = c('Day1', 'Acc', 'Heat', 'Cold', 'Diet'))))

alph_div_pd_as <- alpha_diversity_pd %>%
  filter(species == 'AS') %>%
  dplyr::select(c(diversity, sample_id)) %>%
  left_join(metadata, by='sample_id') %>%
  filter(treatment != 'FAST') %>%
  mutate(days=case_when(treatment == 'Day1' ~ 0,
                        treatment == 'Acc' ~ 7,
                        treatment == 'Heat' ~ 19,
                        treatment == 'Cold' ~ 33,
                        treatment == 'Diet' ~ 45)) %>%
  mutate_at(2:8, as.factor) %>%
  mutate(treatment = factor(treatment, levels = c('Day1', 'Acc', 'Heat', 'Cold', 'Diet'))))

```

```

alph_div_pd_cr <- alpha_diversity_pd %>%
  filter(species == 'CR') %>%
  dplyr::select(c( diversity, sample_id)) %>%
  left_join(metadata, by='sample_id') %>%
  mutate(days=case_when(treatment == 'Day1' ~ 0,
                        treatment == 'Acc' ~ 7,
                        treatment == 'Heat' ~ 19,
                        treatment == 'Cold' ~ 33,
                        treatment == 'Diet' ~ 45),
         cage=case_when(cage == 1 ~ 6,
                        cage == 2 ~ 7,
                        cage == 3 ~ 8,
                        cage == 4 ~ 9,
                        cage == 5 ~ 10,
                        cage == 6 ~ 11)) %>%
  mutate_at(2:8, as.factor) %>%
  mutate(treatment = factor(treatment, levels = c('Day1', 'Acc', 'Heat', 'Cold', 'Diet')))

```

Sanity check of each factor

```
sapply(alph_div_tax_as, mode)
```

```

##      diversity      sample_id      X individual_id      treatment
##      "numeric"    "character"    "numeric"    "character"    "numeric"
##           cage           sex      species          days
##      "numeric"    "character"    "character"    "numeric"

```

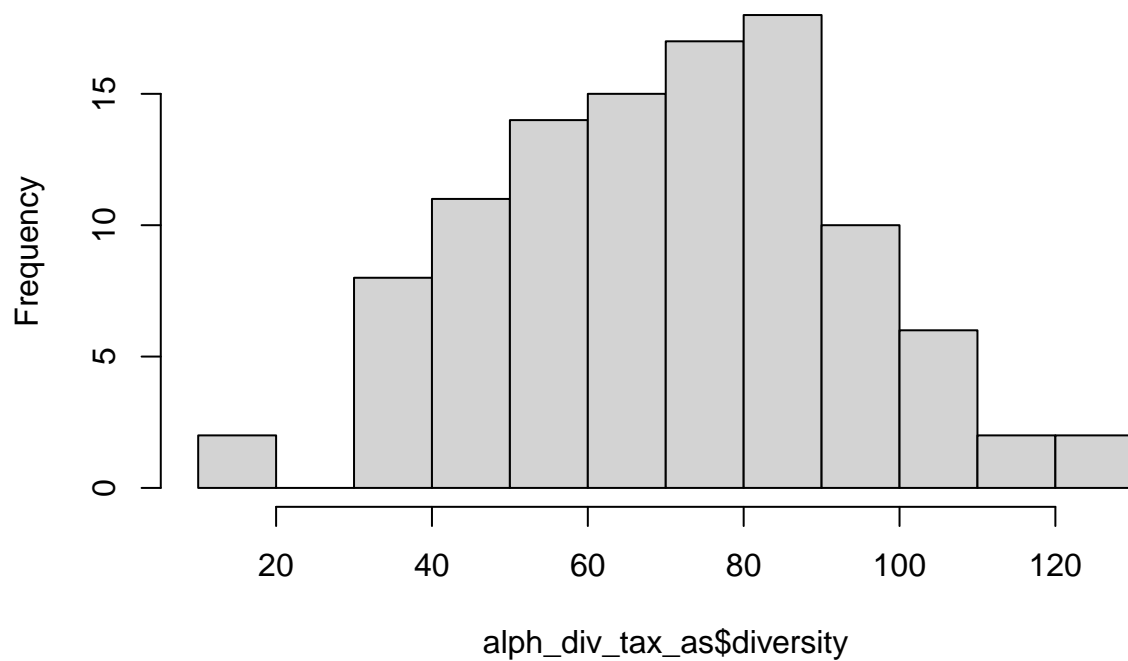
```

contr <- rbind("Day1-Acc" = c(1, -1, 0, 0, 0),
              "Acc-Heat" = c(0, 1, -1, 0, 0),
              "Heat-Cold" = c(0, 0, 1, -1, 0),
              "Cold-Diet" = c(0, 0, 0, 1, -1))
contr_cr <- rbind("Day1-Acc" = c(1, -1, 0, 0, 0),
                 "Acc-Heat" = c(0, 1, -1, 0, 0),
                 "Heat-Cold" = c(0, 0, 1, -1, 0),
                 "Cold-Diet" = c(0, 0, 0, 1, -1))

```

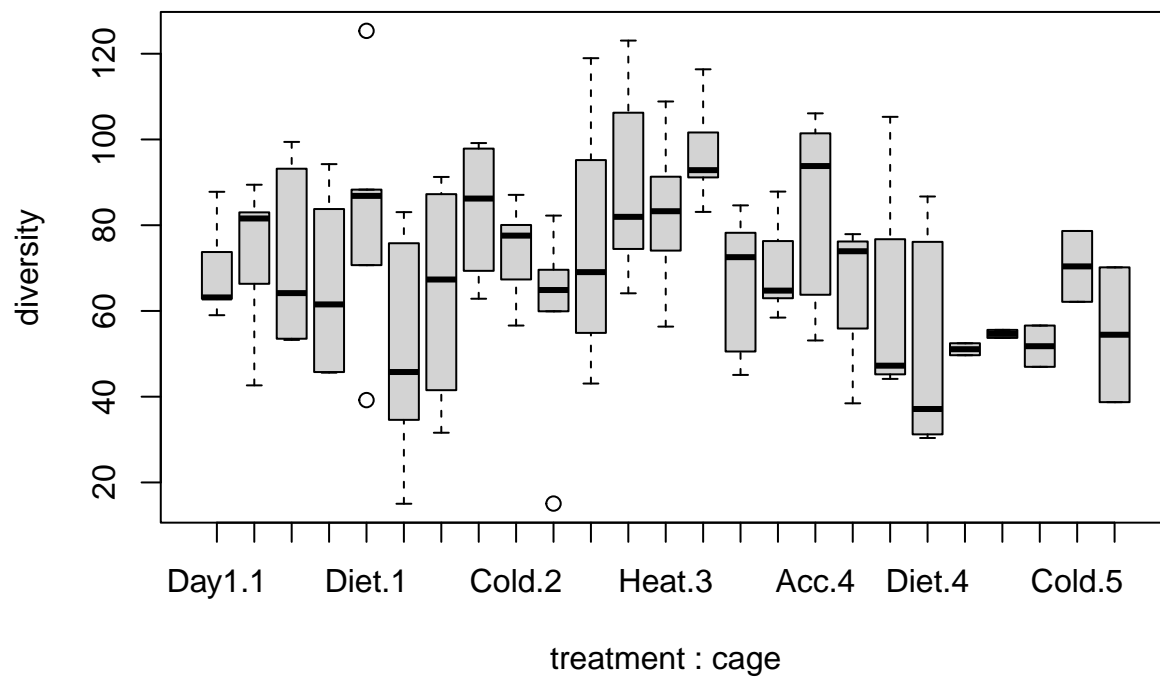
```
hist(alph_div_tax_as$diversity)
```

## Histogram of alph\_div\_tax\_as\$diversity



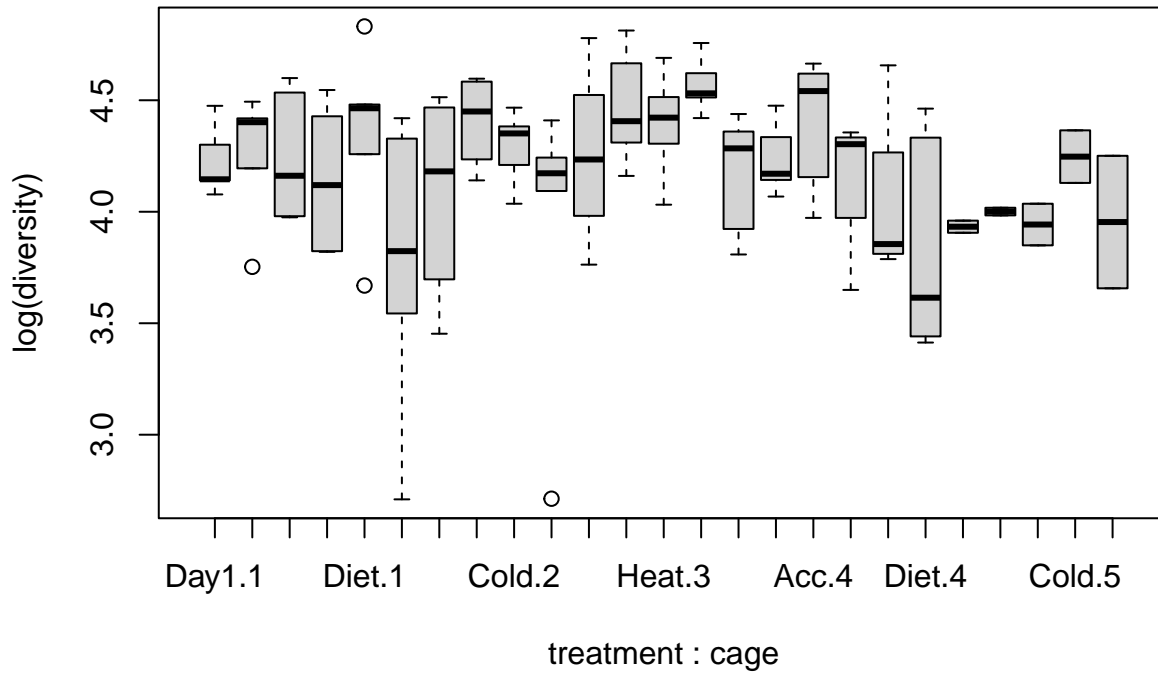
treatment\_wise\_boxplots from the taxonomic diversity Hill numbers

```
boxplot(diversity~treatment*cage,data = alph_div_tax_as)
```



As-  
essment of log transformation it looks tighter, however the non-transformed data still looks ok to me so i  
will continue with this

```
boxplot(log(diversity)~treatment*cage,data = alph_div_tax_as)
```



Create

the models for comparison - including the random effects results in the lowest AIC

```
###Try to include the other factors
model_as.a <- lme(diversity~treatment,
  random = ~1|cage/individual_id,
  correlation = corCAR1(form = ~ as.numeric(days)),
  data=alph_div_tax_as)
model_as.b <- lme(diversity~treatment,
  random = ~1|cage,
  data=alph_div_tax_as)
model_as.c <- lme(diversity~treatment,
  random = ~1|individual_id,
  correlation = corCAR1(form = ~ as.numeric(days)),
  data=alph_div_tax_as)
#model_as.d <- lme(diversity~treatment,
  #random = ~1|cage,
  #correlation = corCAR1(form = ~ days),
  #data=alph_div_tax_as)
model_as.e <- lm(diversity~treatment,
  data=alph_div_tax_as)

anova(model_as.a,model_as.b,model_as.c,model_as.e)
```

##	Model	df	AIC	BIC	logLik	Test	L.Ratio	p-value
##	model_as.a	1	934.9830	958.4295	-458.4915			
##	model_as.b	2	933.4924	951.7286	-459.7462	1 vs 2	2.509424	0.2852
##	model_as.c	3	933.6263	954.4677	-458.8132	2 vs 3	1.866080	0.1719
##	model_as.e	4	935.0961	950.7271	-461.5480	3 vs 4	5.469725	0.0649

```
AIC(model_as.a) #931.8822
```

```
## [1] 934.983
```

```
AIC(model_as.b) #933.4933
```

```
## [1] 933.4924
```

```
AIC(model_as.c) #930.3255
```

```
## [1] 933.6263
```

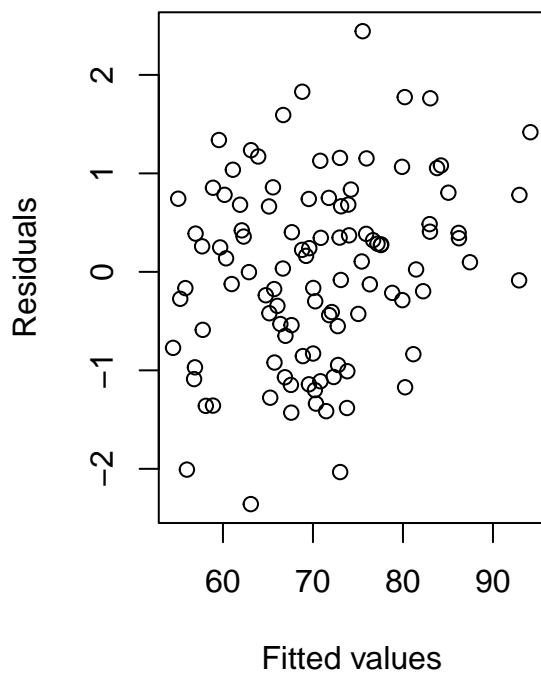
```
AIC(model_as.e) #960.1475
```

```
## [1] 960.1465
```

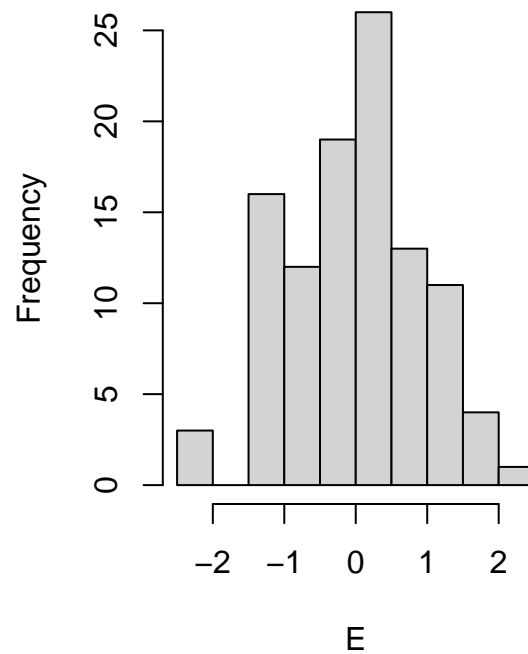
Residuals looks like they are randomly distributed around 0

```
model_as1=model_as.a  
E <- resid(model_as1, type = "normalized")  
Fit <- fitted(model_as1)  
op <- par(mfrow = c(1, 2))  
plot(x = Fit, y = E,  
      xlab = "Fitted values", ylab = "Residuals",  
      main = "Residuals versus fitted values")  
hist(E, nclass = 15)
```

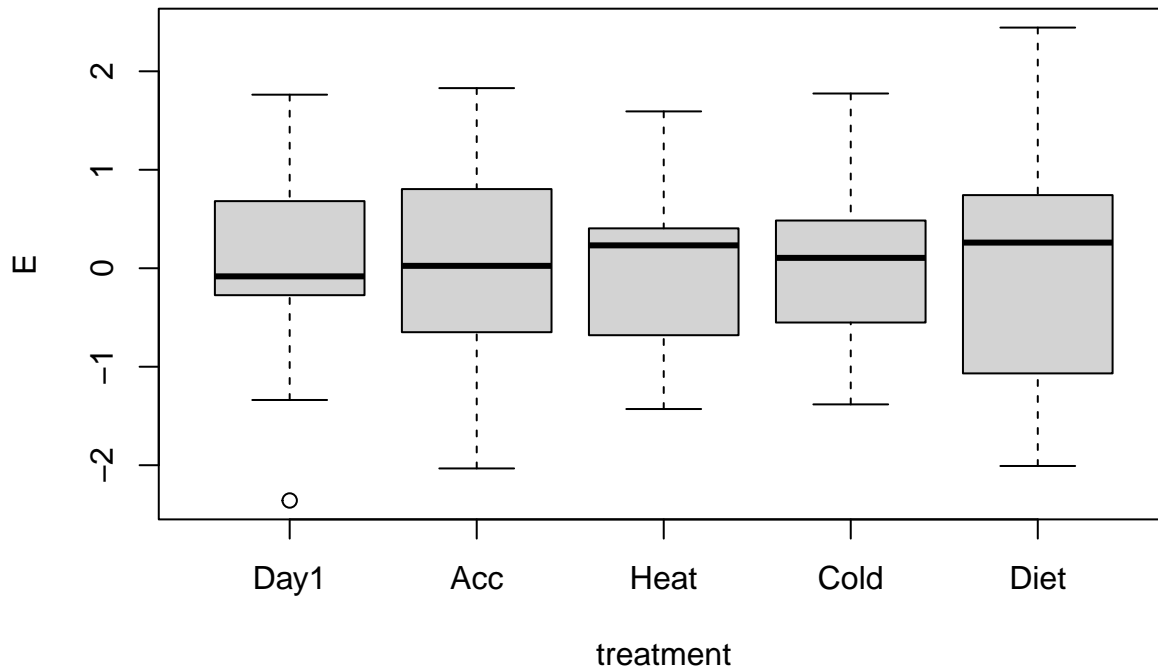
**Residuals versus fitted values**



**Histogram of E**



```
par(op)
bp=boxplot(E ~ treatment,
           data = alph_div_tax_as)
```



```
summary(model_as.a)
```

```
## Linear mixed-effects model fit by REML
##   Data: alph_div_tax_as
##       AIC      BIC    logLik
##   934.983 958.4295 -458.4915
##
## Random effects:
## Formula: ~1 | cage
##      (Intercept)
## StdDev:      5.81433
##
## Formula: ~1 | individual_id %in% cage
##      (Intercept) Residual
## StdDev:      8.422224 20.38083
##
## Correlation Structure: Continuous AR(1)
## Formula: ~as.numeric(days) | cage/individual_id
## Parameter estimate(s):
##      Phi
## 0.2000667
## Fixed effects: diversity ~ treatment
##              Value Std.Error DF   t-value p-value
## (Intercept)  64.02810  5.494296 79 11.653560  0.0000
## treatmentAcc  11.12264  6.306979 79  1.763545  0.0817
## treatmentHeat  9.90226  6.394579 79  1.548540  0.1255
## treatmentCold  9.86497  6.307017 79  1.564126  0.1218
```

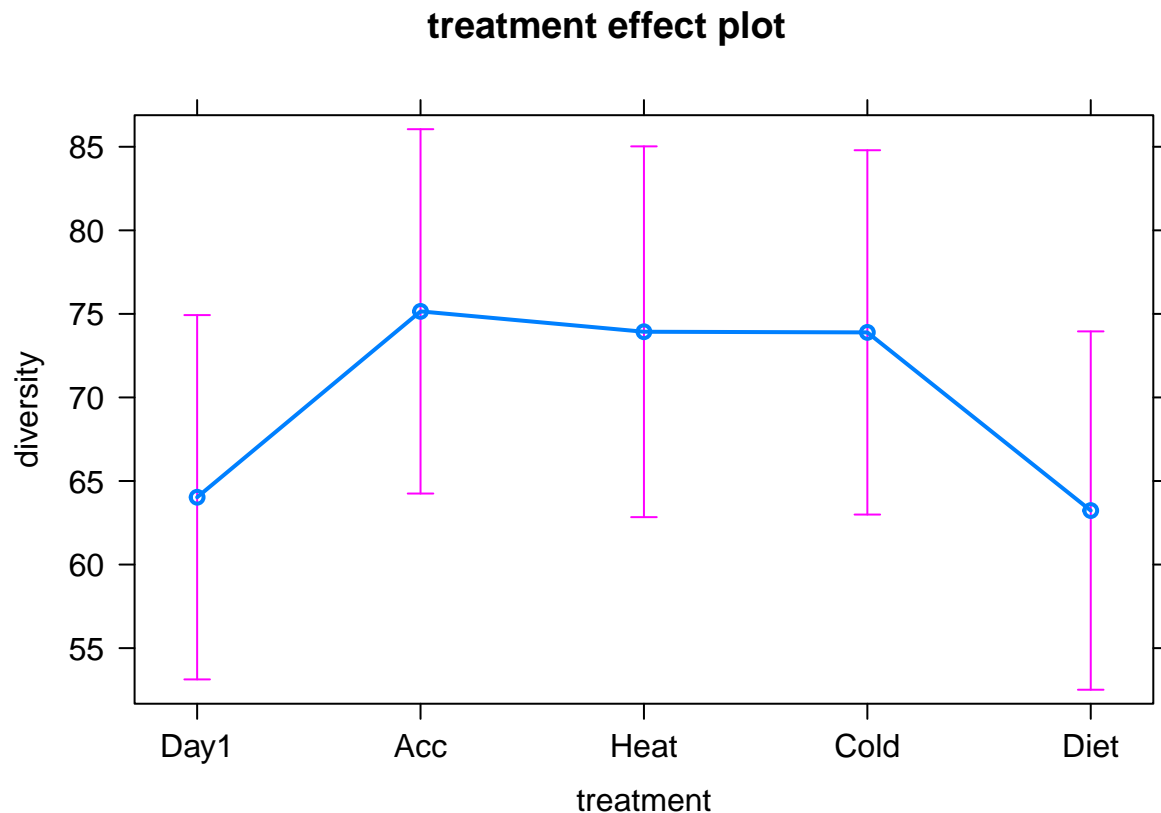
```
## treatmentDiet -0.79601  6.225929 79 -0.127854  0.8986
## Correlation:
##          (Intr) trtmnA trtmnH trtmnC
## treatmentAcc -0.574
## treatmentHeat -0.567  0.493
## treatmentCold -0.574  0.500  0.493
## treatmentDiet -0.581  0.507  0.500  0.507
##
## Standardized Within-Group Residuals:
##          Min          Q1          Med          Q3          Max
## -2.35827091 -0.64938041  0.09777876  0.68155759  2.44381609
##
## Number of Observations: 105
## Number of Groups:
##          cage individual_id %in% cage
##          5                22
```

We have a slightly significant treatment effect

```
anova(model_as.a)
```

```
##          numDF denDF  F-value p-value
## (Intercept)     1    79 342.5953 <.0001
## treatment       4    79   1.7623  0.1448
```

```
plot(allEffects(model_as.a))
```



Pairwise comparison of treatments to assess where the significant effects are



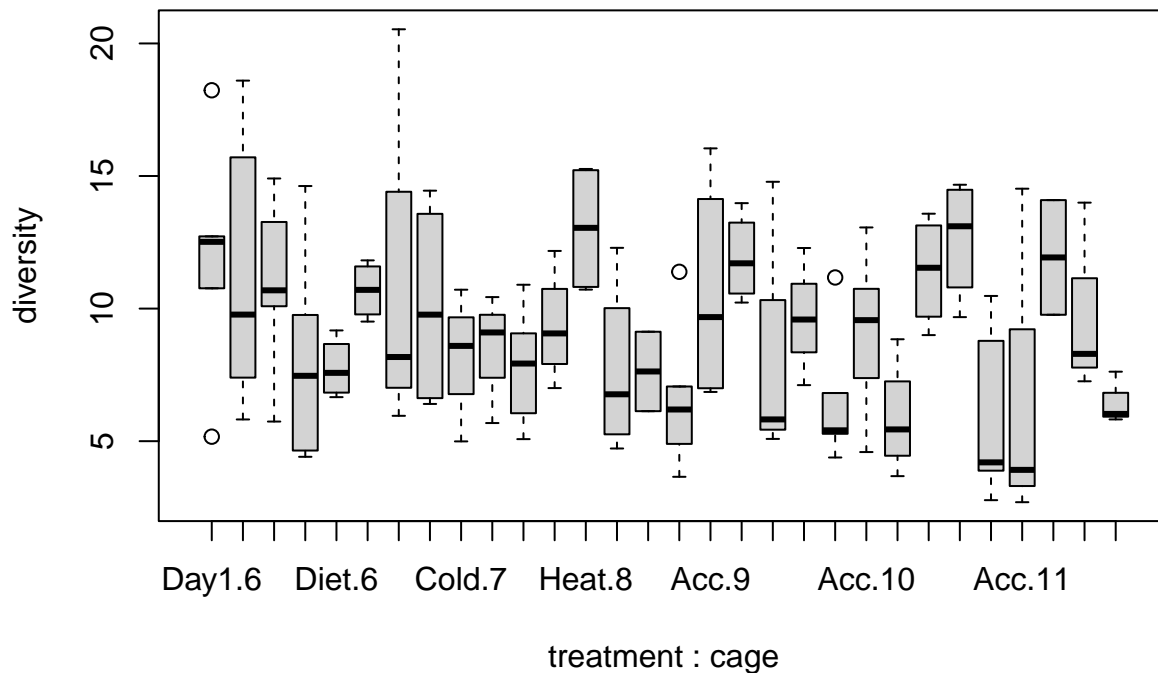
```
summary(glht(model_as.a, linfct = mcp(treatment=contr),
      test = adjusted("holm")))
```

```
##
## Simultaneous Tests for General Linear Hypotheses
##
## Multiple Comparisons of Means: User-defined Contrasts
##
##
## Fit: lme.formula(fixed = diversity ~ treatment, data = alph_div_tax_as,
## random = ~1 | cage/individual_id, correlation = corCAR1(form = ~as.numeric(days)))
##
## Linear Hypotheses:
##           Estimate Std. Error z value Pr(>|z|)
## Day1-Acc == 0 -11.12264    6.30698  -1.764   0.248
## Acc-Heat == 0  1.22038    6.39240   0.191   0.999
## Heat-Cold == 0  0.03729    6.39240   0.006   1.000
## Cold-Diet == 0 10.66098    6.22593   1.712   0.273
## (Adjusted p values reported -- single-step method)
```

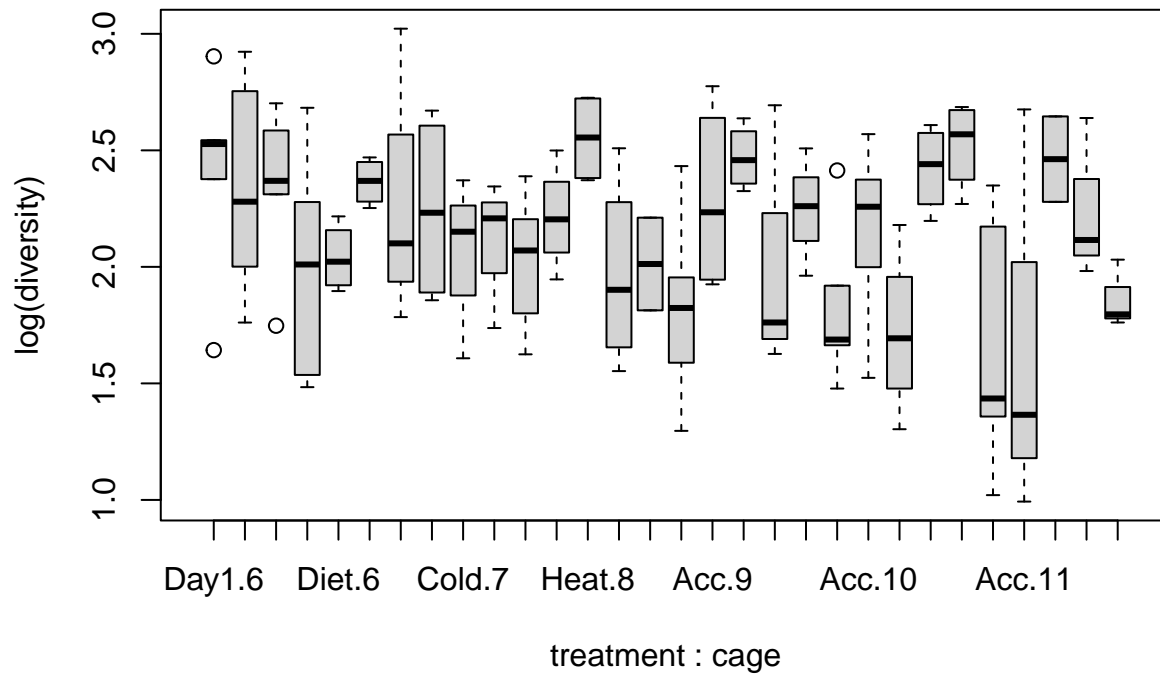
```
#plot_model(model_as1, type='Day1d', terms = c('treatment'))
```

#####Start if the Crocidura taxonomic linear models#####

```
boxplot(diversity~treatment*cage,data = alph_div_tax_cr)
```

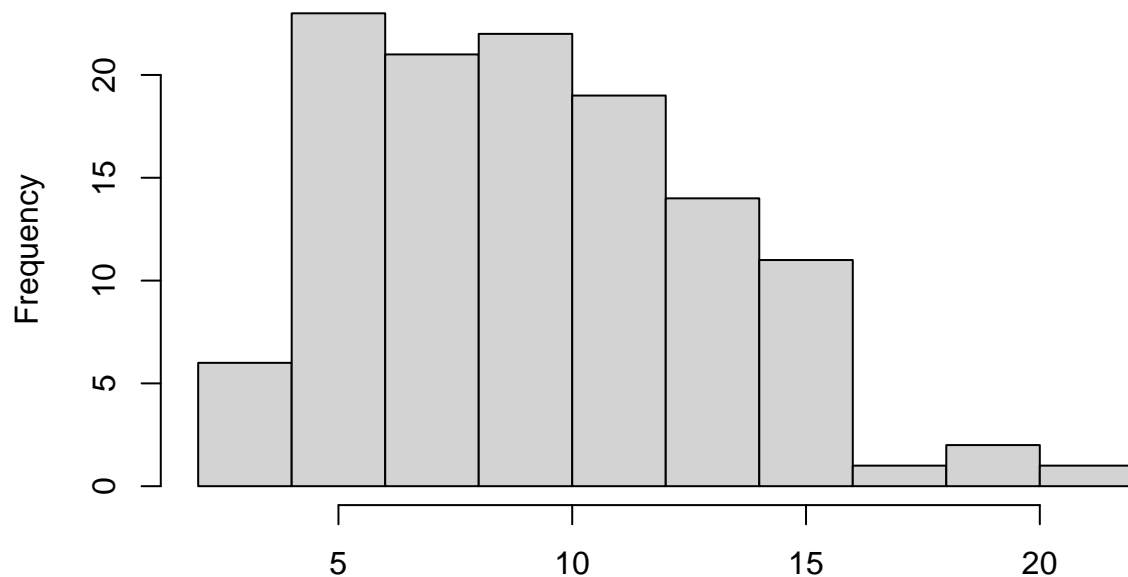


```
boxplot(log(diversity)~treatment*cage,data = alph_div_tax_cr)
```



```
hist(alph_div_tax_cr$diversity)
```

### Histogram of alph\_div\_tax\_cr\$diversity



The lowest AIC is the gls without random effects - however I consider the random effects to be an essential study design parameter - I'll continue with the random effects included

```
#model_cr.a <- lme(diversity~treatment,
  #random = ~1|cage/individual_id,
  #correlation = corCAR1(form = ~ as.numeric(days)),
```

```

#data=alph_div_tax_cr)
model_cr.b <- lme(versus~treatment,
  random = ~1|cage,
  data=alph_div_tax_cr)
model_cr.c <- lme(versus~treatment,
  random = ~1|individual_id,
  correlation = corCAR1(form = ~ as.numeric(days)),
  data=alph_div_tax_cr)
model_cr.e <- lm(versus~treatment,
  data=alph_div_tax_cr)

#AIC(model_cr.a) #657.3954
AIC(model_cr.b) #654.6046

```

```
## [1] 654.6036
```

```
AIC(model_cr.c) #655.3954
```

```
## [1] 656.4563
```

```

#AIC(model_cr.d) #NA
AIC(model_cr.e) #658.816

```

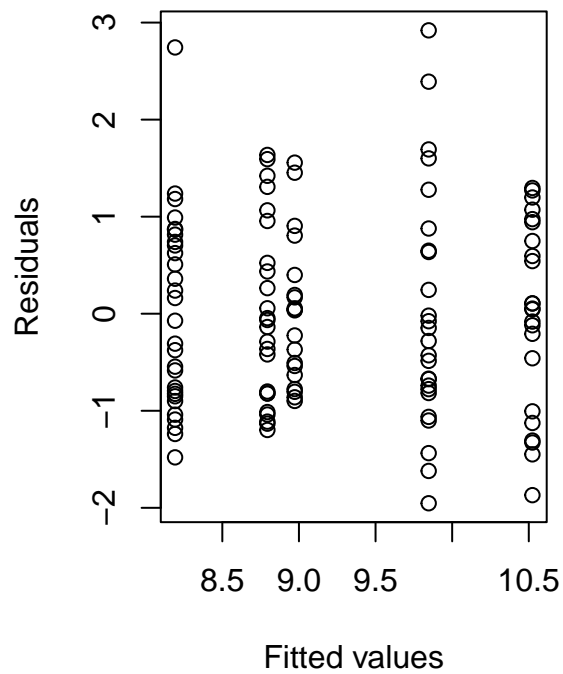
```
## [1] 658.8157
```

```

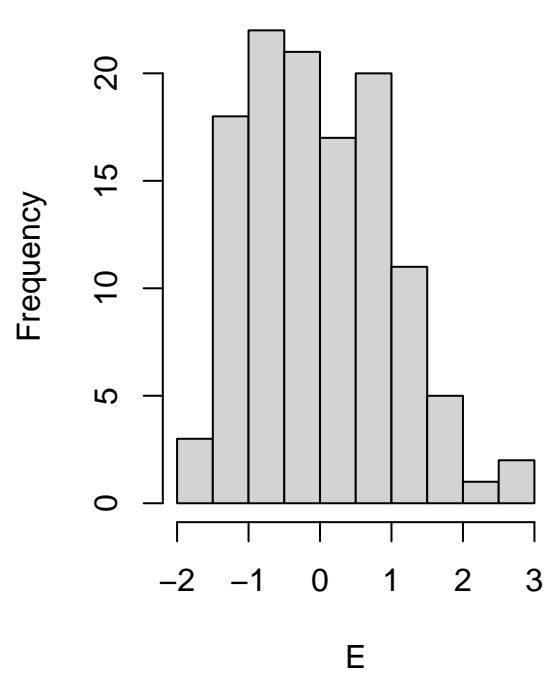
model_cr1=model_cr.b
E <- resid(model_cr1, type = "normalized")
Fit <- fitted(model_cr1)
op <- par(mfrow = c(1, 2))
plot(x = Fit, y = E,
  xlab = "Fitted values", ylab = "Residuals",
  main = "Residuals versus fitted values")
hist(E, nclass = 15)

```

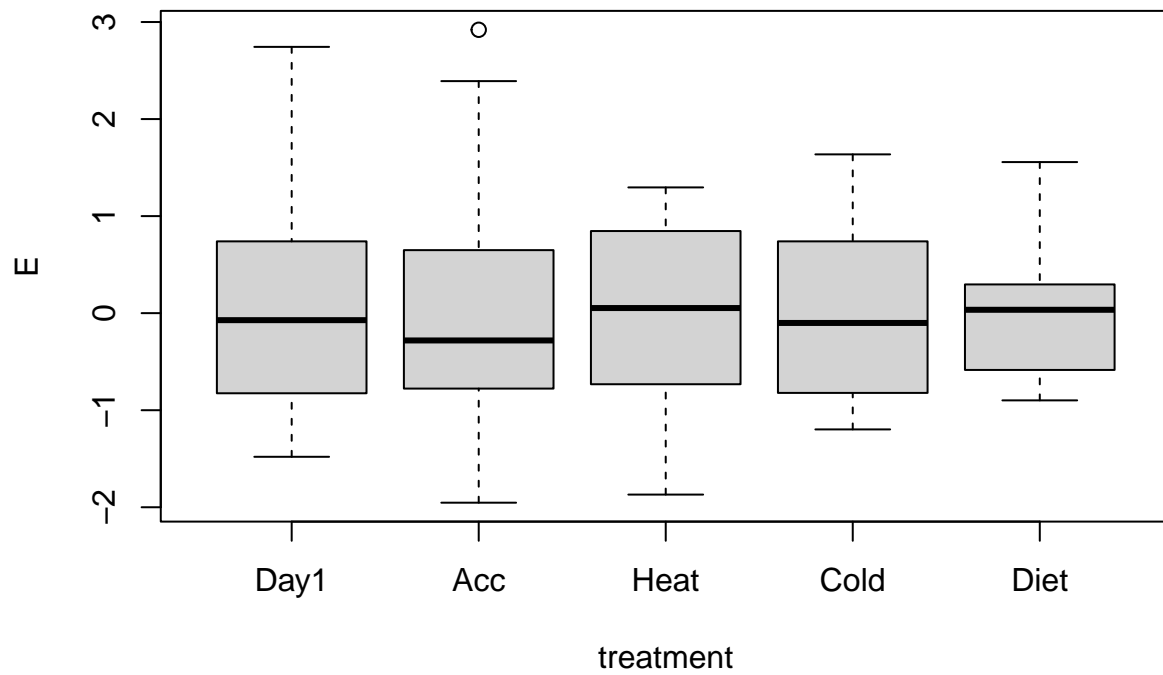
**Residuals versus fitted values**



**Histogram of E**



```
par(op)
bp=boxplot(E ~ treatment,
           data = alph_div_tax_cr)
```



Again we have a slightly significance treatment effect on taxonomy

```
anova(model_cr1)
```

```
##              numDF denDF  F-value p-value
## (Intercept)      1   110 762.8574  <.0001
## treatment        4   110   1.5882  0.1825
```

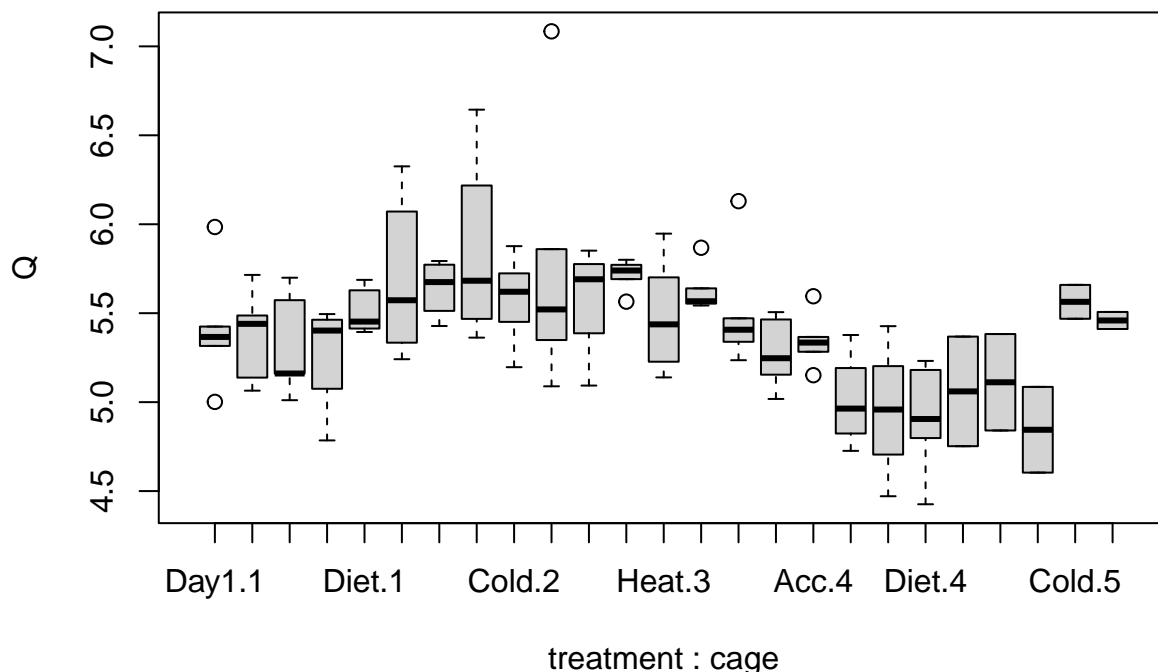
Pairwise comparison of treatments

```
summary(glht(model_cr1, linfct = mcp(treatment=contr_cr),
  test = adjusted("holm")))
```

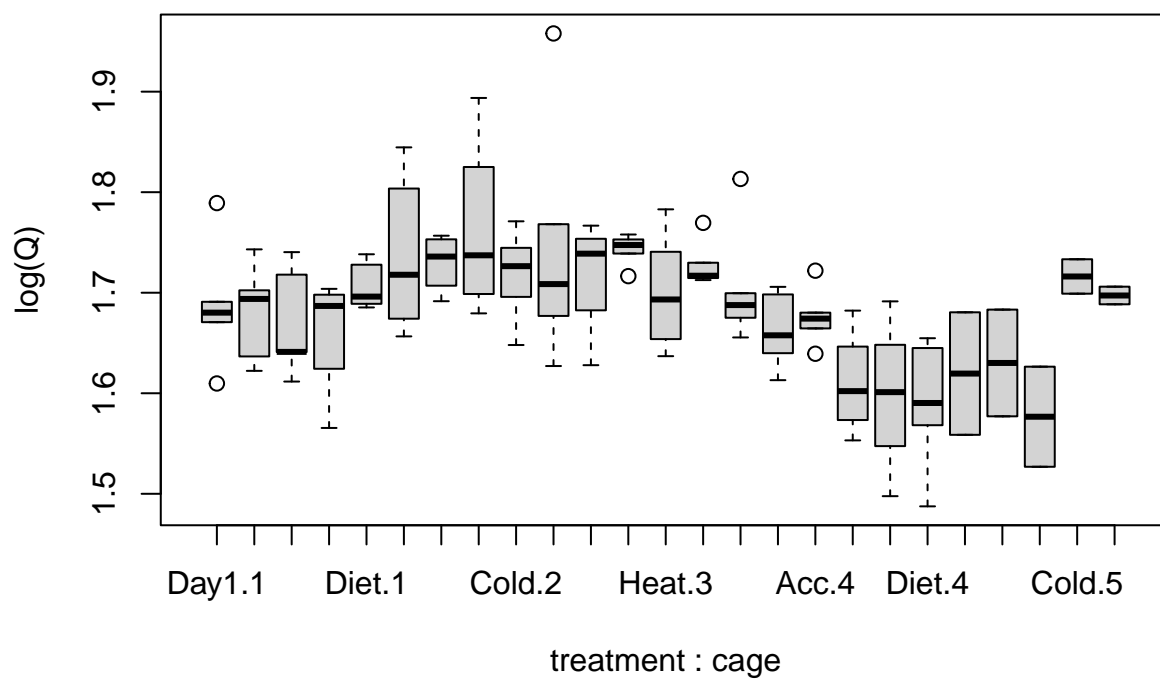
```
##
## Simultaneous Tests for General Linear Hypotheses
##
## Multiple Comparisons of Means: User-defined Contrasts
##
##
## Fit: lme.formula(fixed = diversity ~ treatment, data = alph_div_tax_cr,
##   random = ~1 | cage)
##
## Linear Hypotheses:
##              Estimate Std. Error z value Pr(>|z|)
## Day1-Acc == 0   -1.6551     0.9988  -1.657   0.302
## Acc-Heat == 0   -0.6770     1.0574  -0.640   0.924
## Heat-Cold == 0    1.7296     1.0679   1.620   0.322
## Cold-Diet == 0   -0.1784     1.1239  -0.159   1.000
## (Adjusted p values reported -- single-step method)

#####Start of functional alpha diversity Apodemus#####

boxplot(Q~treatment*cage,data = alph_div_fd_as)
```

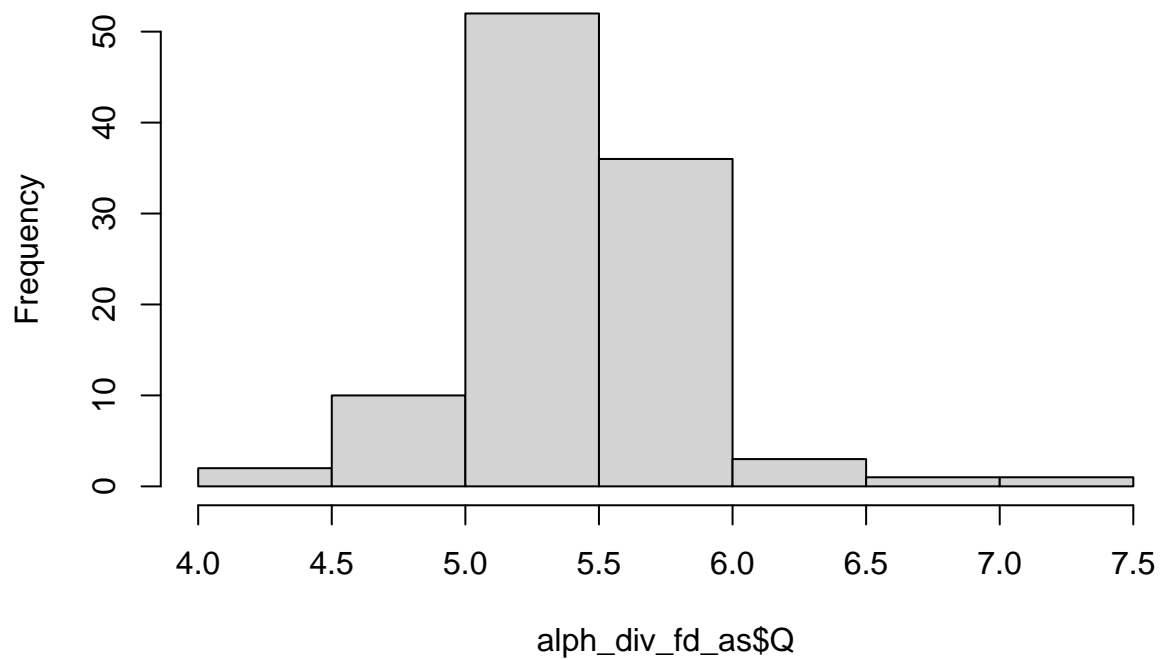


```
boxplot(log(Q)~treatment*cage,data = alph_div_fd_as)
```



```
hist(alph_div_fd_as$Q)
```

**Histogram of alph\_div\_fd\_as\$Q**



```

model_as.a_fd <- lme(Q~treatment,
  random = ~1|cage/individual_id,
  correlation = corCAR1(form = ~ as.numeric(days)),
  data=alph_div_fd_as)
model_as.b_fd <- lme(Q~treatment,
  random = ~1|individual_id,
  correlation = corCAR1(form = ~ as.numeric(days)),
  data=alph_div_fd_as)
model_as.c_fd <- lme(Q~treatment,
  random = ~1|cage,
  data=alph_div_fd_as)
model_as.d_fd <- lme(Q~treatment,
  random = ~1|cage/individual_id,
  data=alph_div_fd_as)
model_as.e_fd <- lm(Q~treatment,
  data=alph_div_fd_as)
AIC(model_as.a_fd) #292.0188

```

```
## [1] 92.69364
```

```
AIC(model_as.b_fd) #290.6965
```

```
## [1] 95.5458
```

```
AIC(model_as.c_fd) #292.2867
```

```
## [1] 108.2607
```

```
AIC(model_as.d_fd) #292.3475
```

```
## [1] 93.27356
```

```
AIC(model_as.e_fd) #286.632
```

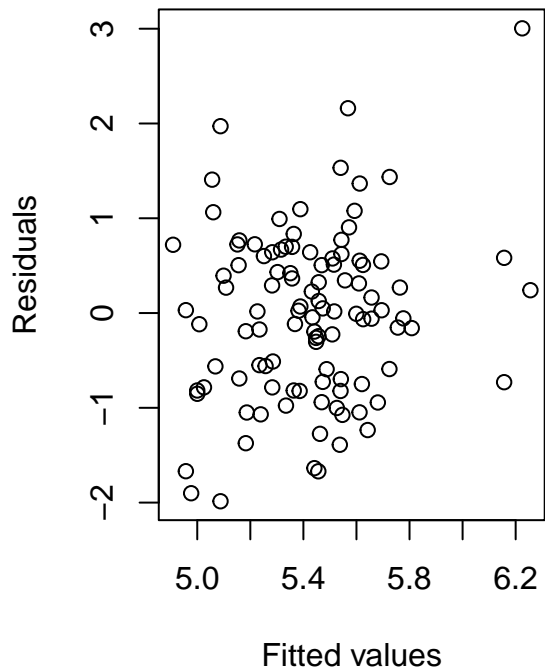
```
## [1] 116.5169
```

```

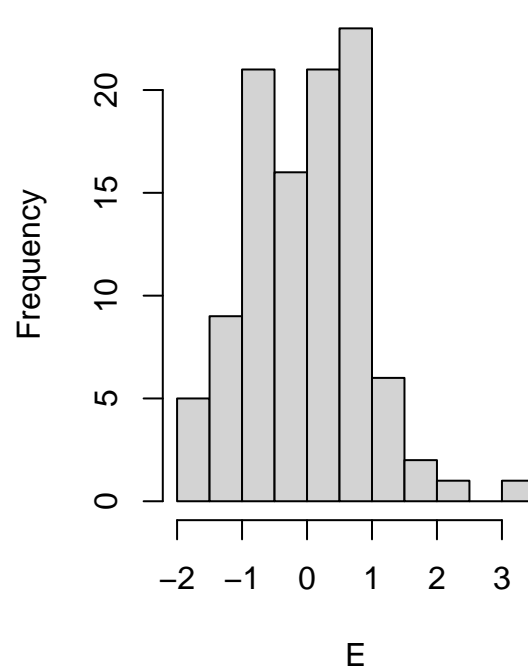
model_as1_fd=model_as.a_fd
E <- resid(model_as1_fd, type = "normalized")
Fit <- fitted(model_as.a_fd)
op <- par(mfrow = c(1, 2))
plot(x = Fit, y = E,
  xlab = "Fitted values", ylab = "Residuals",
  main = "Residuals versus fitted values")
hist(E, nclass = 15)

```

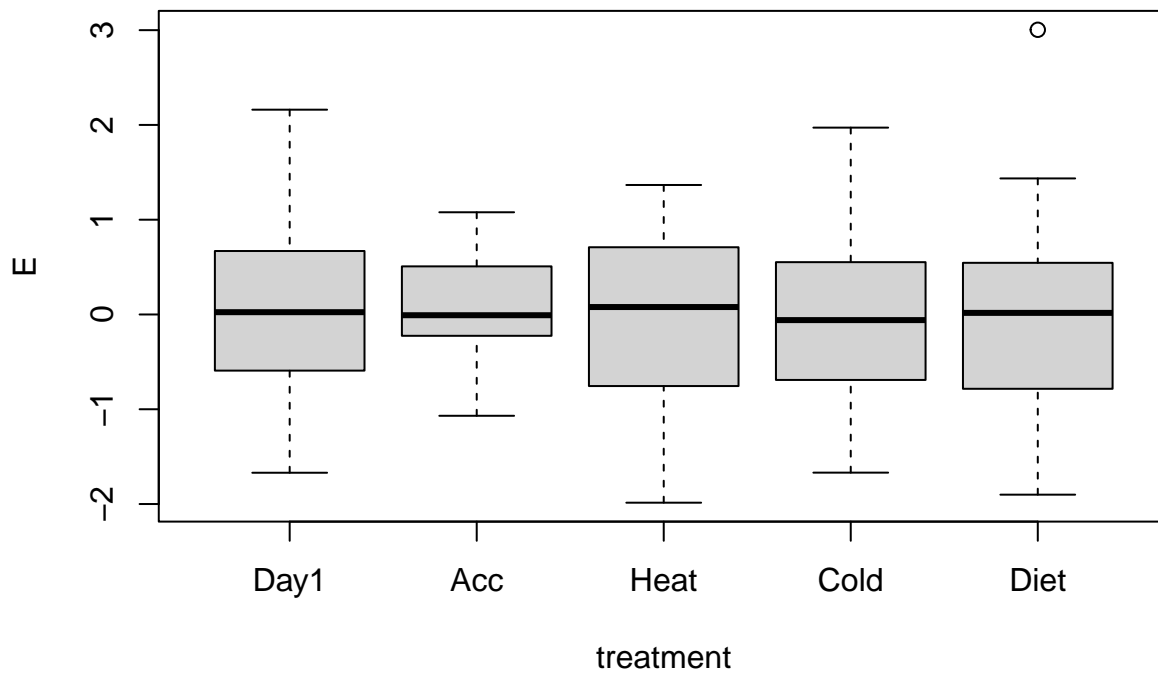
**Residuals versus fitted values**



**Histogram of E**



```
par(op)
bp=boxplot(E ~ treatment,
           data = alph_div_fd_as)
```



We observe no significant difference in the effect of treatment on functional alpha diversity



```
anova(model_as1_fd)
```

```
##           numDF denDF    F-value p-value
## (Intercept)      1    79 2026.1679 <.0001
## treatment        4    79    1.0907  0.367
```

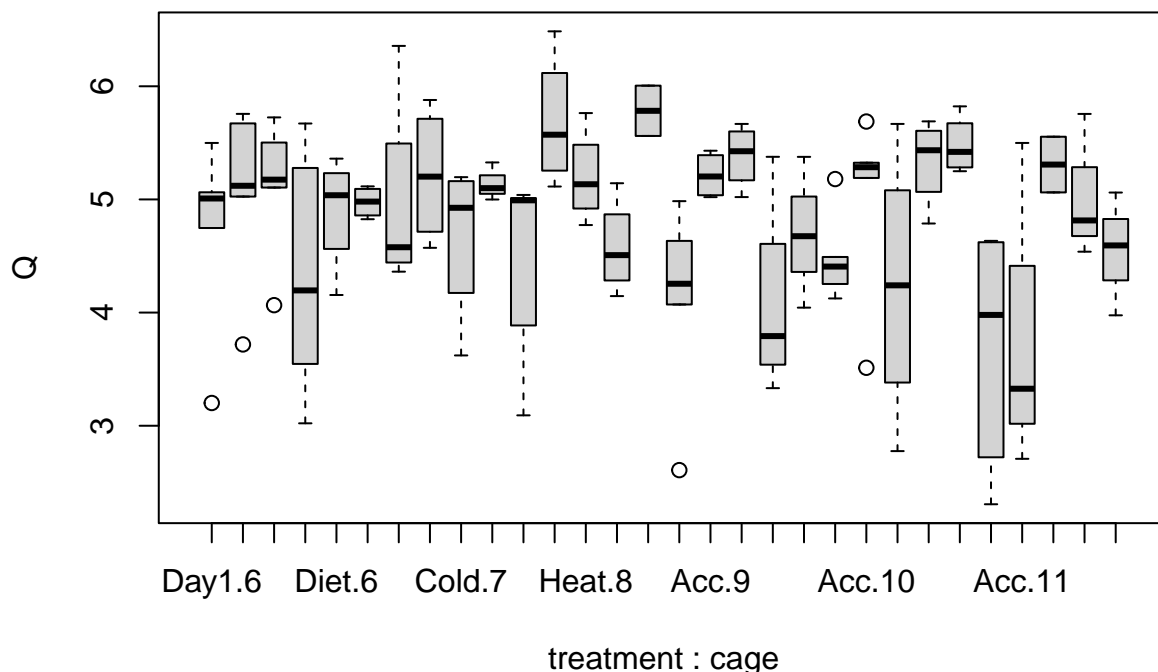
This is useless - ignore it for now but just wanted to sanity check

```
summary(glht(model_as1_fd, linfct = mcp(treatment=contr),
          test = adjusted("holm")))
```

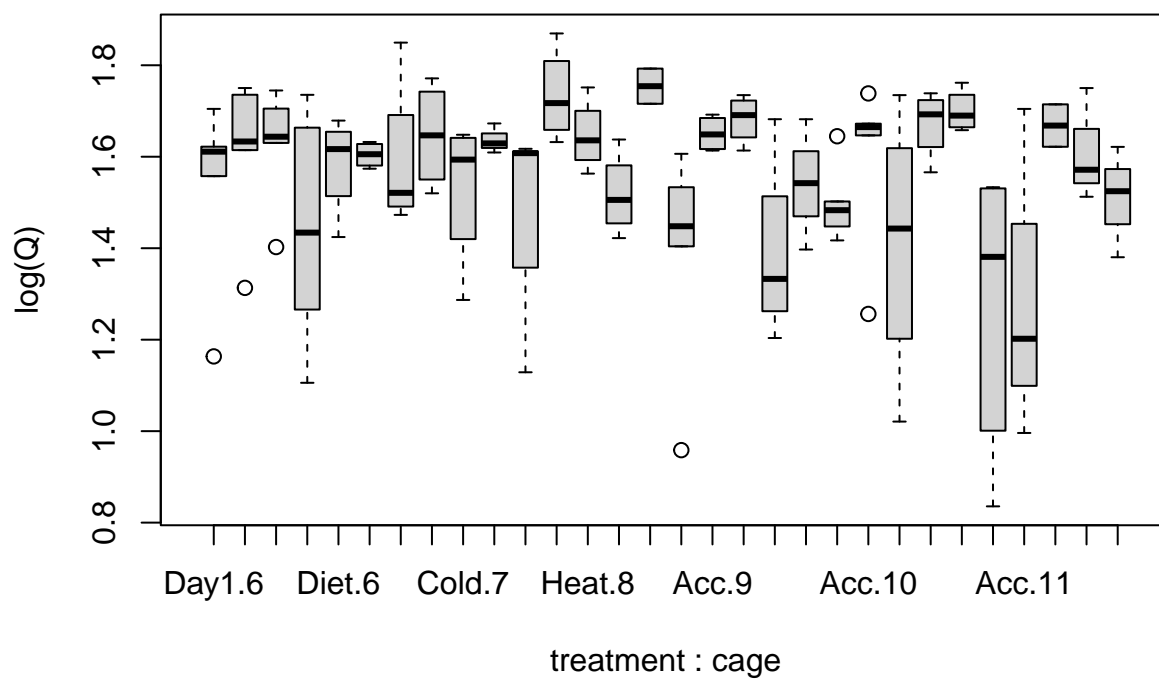
```
##
## Simultaneous Tests for General Linear Hypotheses
##
## Multiple Comparisons of Means: User-defined Contrasts
##
##
## Fit: lme.formula(fixed = Q ~ treatment, data = alph_div_fd_as, random = ~1 |
##       cage/individual_id, correlation = corCAR1(form = ~as.numeric(days)))
##
## Linear Hypotheses:
##           Estimate Std. Error z value Pr(>|z|)
## Day1-Acc == 0 -5.243e-02  7.089e-02 -0.740    0.894
## Acc-Heat == 0  1.514e-01  8.189e-02  1.849    0.218
## Heat-Cold == 0  5.619e-05  8.422e-02  0.001    1.000
## Cold-Diet == 0 -6.825e-02  7.958e-02 -0.858    0.834
## (Adjusted p values reported -- single-step method)
```

#####Start of the Crocidura functional alpha diversity

```
boxplot(Q~treatment*cage,data = alph_div_fd_cr)
```

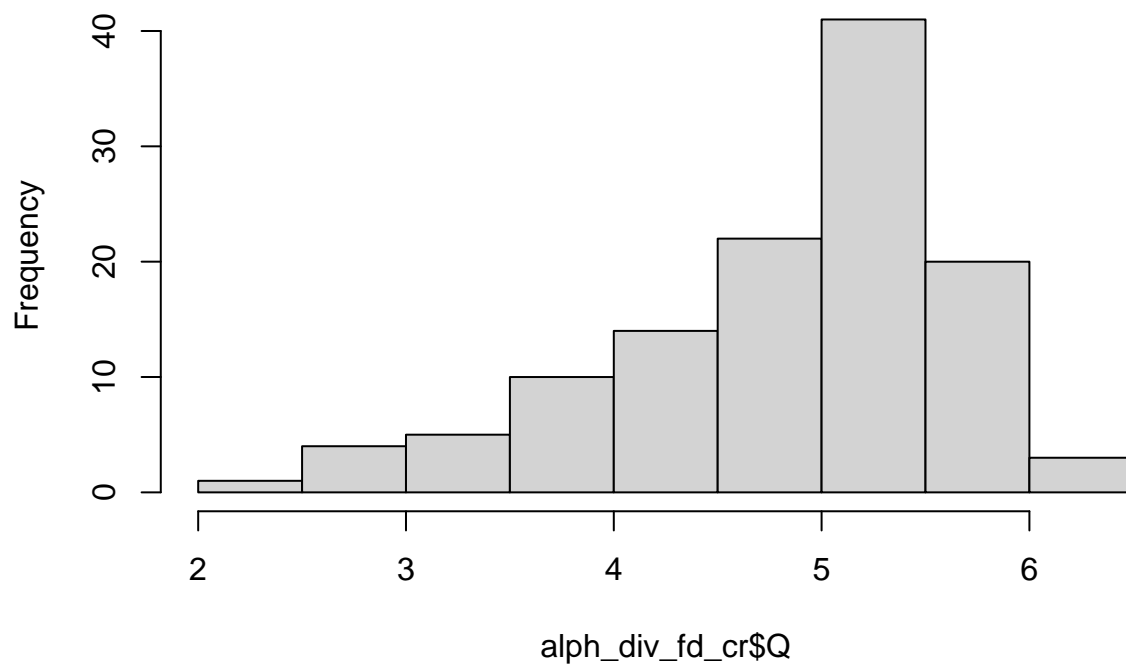


```
boxplot(log(Q)~treatment*cage,data = alph_div_fd_cr)
```



```
hist(alph_div_fd_cr$Q)
```

**Histogram of alph\_div\_fd\_cr\$Q**



```

model_cr.a_fd <- lme(Q~treatment,
  random = ~1|cage/individual_id,
  correlation = corCAR1(form = ~ as.numeric(days)),
  data=alph_div_fd_cr)
model_cr.b_fd <- lme(Q~treatment,
  random = ~1|individual_id,
  correlation = corCAR1(form = ~ as.numeric(days)),
  data=alph_div_fd_cr)
model_cr.c_fd <- lme(Q~treatment,
  random = ~1|cage,
  data=alph_div_fd_cr)
model_cr.d_fd <- lme(Q~treatment,
  random = ~1|individual_id,
  data=alph_div_fd_cr)
model_cr.e_fd <- lm(Q~treatment,
  data=alph_div_fd_cr)

```

```
AIC(model_cr.a_fd) #377.5765
```

```
## [1] 304.8715
```

```
AIC(model_cr.b_fd) #376.0824
```

```
## [1] 303.0507
```

```
AIC(model_cr.c_fd) #379.5765
```

```
## [1] 301.5524
```

```
AIC(model_cr.d_fd) #374.0913
```

```
## [1] 301.7523
```

```
AIC(model_cr.e_fd) #370.5743
```

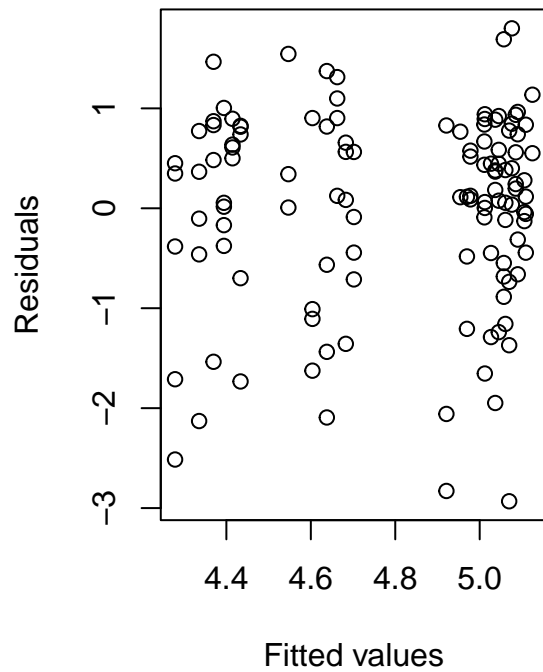
```
## [1] 290.9681
```

```

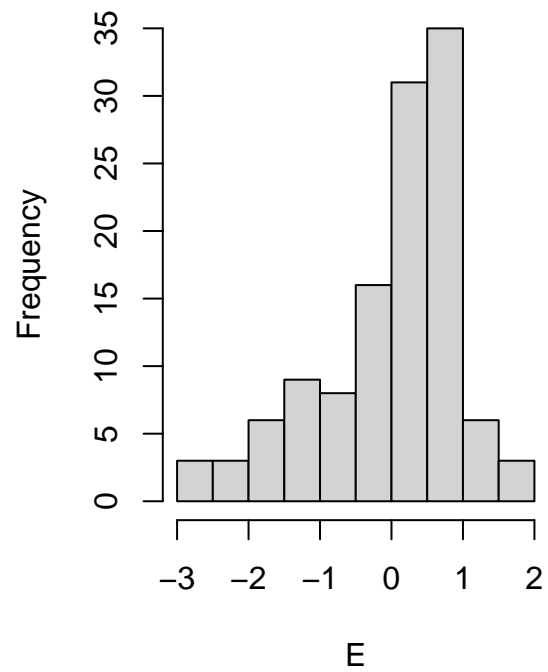
model_cr1_fd=model_cr.a_fd
E <- resid(model_cr1_fd, type = "normalized")
Fit <- fitted(model_cr1_fd)
op <- par(mfrow = c(1, 2))
plot(x = Fit, y = E,
  xlab = "Fitted values", ylab = "Residuals",
  main = "Residuals versus fitted values")
hist(E, nclass = 15)

```

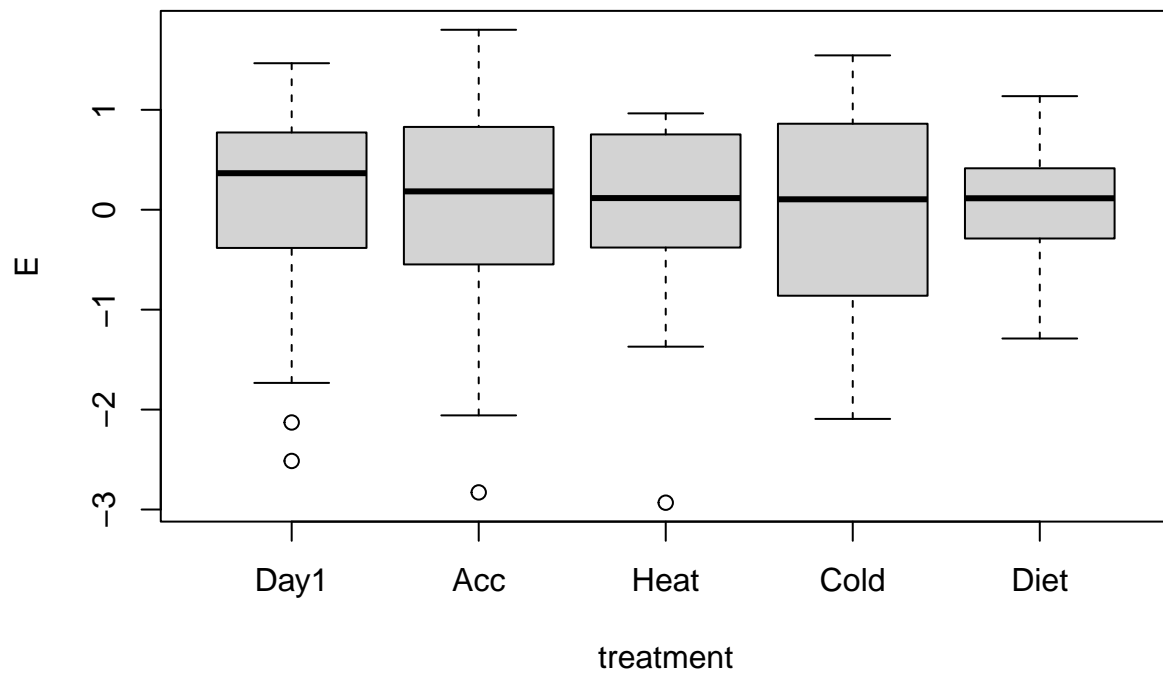
**Residuals versus fitted values**



**Histogram of E**



```
par(op)
bp=boxplot(E ~ treatment,
           data = alph_div_tax_cr)
```



have a highly significant effect of treatment on functional alpha diversity

We

```
anova(model_cr1_fd)
```

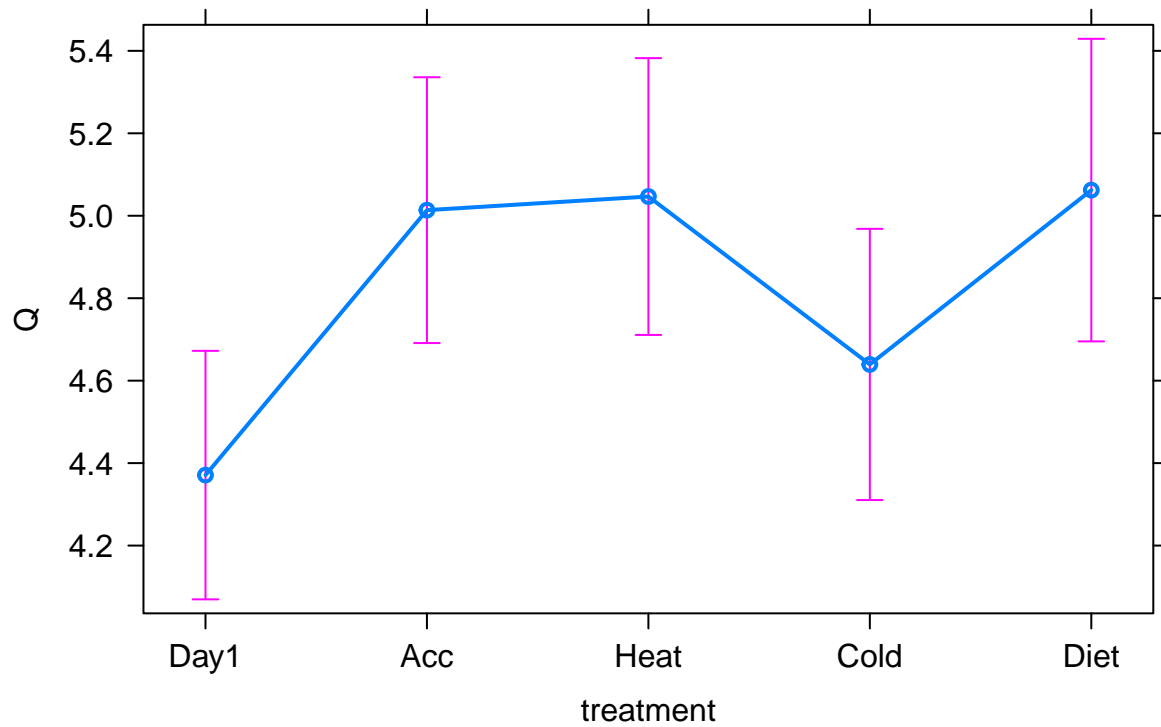
```
##           numDF denDF    F-value p-value
## (Intercept)      1    87 3038.5944 <.0001
## treatment        4    87    4.2982 0.0032
```

```
summary(glht(model_cr1_fd, linfct = mcp(treatment=contr_cr),
      test = adjusted("holm")))
```

```
##
## Simultaneous Tests for General Linear Hypotheses
##
## Multiple Comparisons of Means: User-defined Contrasts
##
##
## Fit: lme.formula(fixed = Q ~ treatment, data = alph_div_fd_cr, random = ~1 |
##       cage/individual_id, correlation = corCAR1(form = ~as.numeric(days)))
##
## Linear Hypotheses:
##           Estimate Std. Error z value Pr(>|z|)
## Day1-Acc == 0 -0.64267    0.19986  -3.216 0.00494 **
## Acc-Heat == 0 -0.03303    0.22256  -0.148 0.99968
## Heat-Cold == 0  0.40729    0.22636   1.799 0.23403
## Cold-Diet == 0 -0.42289    0.23702  -1.784 0.24094
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## (Adjusted p values reported -- single-step method)
```

```
plot(allEffects(model_cr1_fd))
```

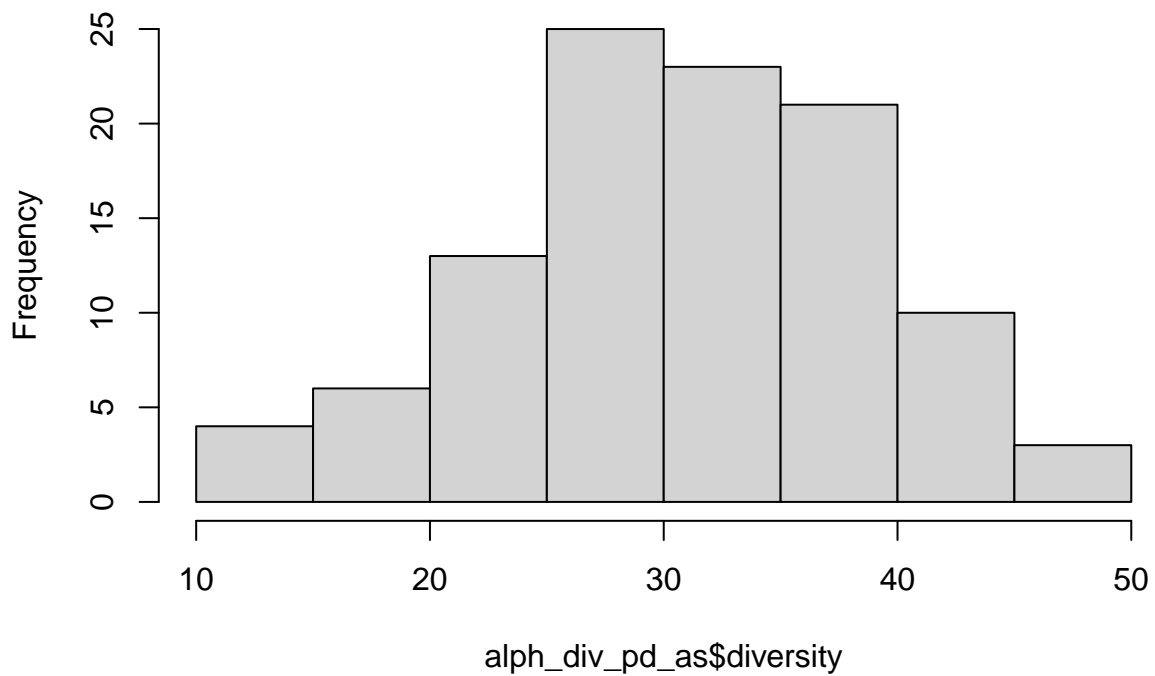
**treatment effect plot**



#####Phylogenetic linear models

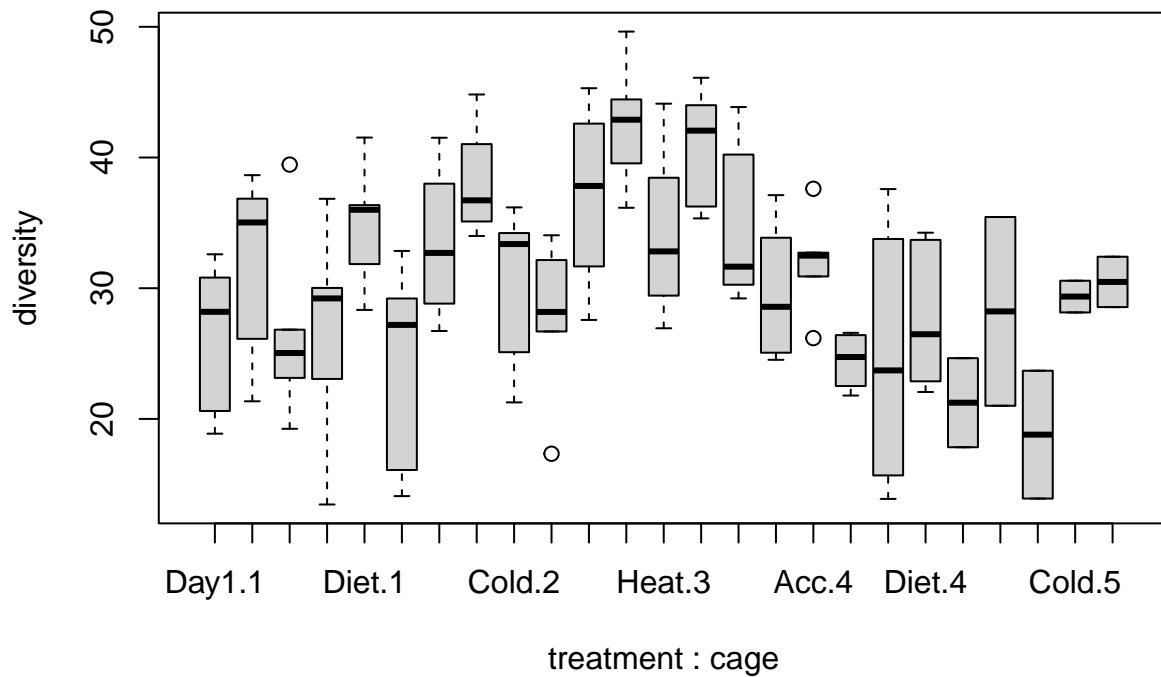
```
hist(alph_div_pd_as$diversity)
```

**Histogram of alph\_div\_pd\_as\$diversity**



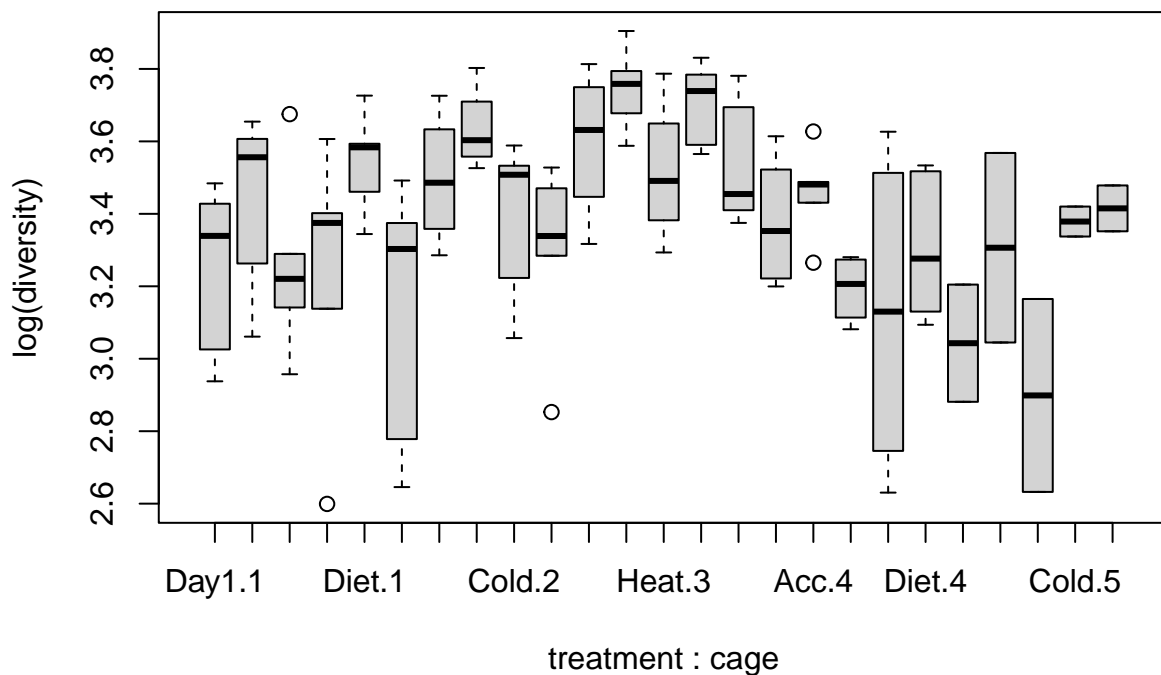
treatment\_wise\_boxplots from the taxonomic diversity Hill numbers

```
boxplot(diversity~treatment*cage,data = alph_div_pd_as)
```



As-  
essment of log transformation it looks tighter, however the non-transformed data still looks ok to me so i will continue with this

```
boxplot(log(diversity)~treatment*cage,data = alph_div_pd_as)
```



Create  
the models for comparison - including the random effects results in the lowest AIC

```

###Try to include the other factors
model_as.a.pd <- lme(diversity~treatment,
  random = ~1|cage/individual_id,
  correlation = corCAR1(form = ~ as.numeric(days)),
  data=alph_div_pd_as)
model_as.b.pd <- lme(diversity~treatment,
  random = ~1|individual_id,
  data=alph_div_pd_as)
model_as.c.pd <- lme(diversity~treatment,
  random = ~1|individual_id,
  correlation = corCAR1(form = ~ as.numeric(days)),
  data=alph_div_pd_as)
model_as.d.pd <- lme(diversity~treatment,
  random = ~1|cage,
  data=alph_div_pd_as)
model_as.e.pd <- lm(diversity~treatment,
  data=alph_div_pd_as)

anova(model_as.a.pd,model_as.b.pd,model_as.c.pd,model_as.e.pd)

```

##	Model	df	AIC	BIC	logLik	Test	L.Ratio	p-value
##	model_as.a.pd	1	9 707.3642	730.8108	-344.6821			
##	model_as.b.pd	2	7 712.2172	730.4534	-349.1086	1 vs 2	8.852929	0.0120
##	model_as.c.pd	3	8 714.2171	735.0584	-349.1085	2 vs 3	0.000091	0.9924
##	model_as.e.pd	4	6 724.2139	739.8449	-356.1069	3 vs 4	13.996777	0.0009

```
AIC(model_as.a.pd ) #525.004
```

```
## [1] 707.3642
```

```
AIC(model_as.b.pd ) #529.4421
```

```
## [1] 712.2172
```

```
AIC(model_as.c.pd ) #529.9477
```

```
## [1] 714.2171
```

```
AIC(model_as.d.pd ) #523.7998
```

```
## [1] 706.5748
```

```
AIC(model_as.e.pd ) #546.8065
```

```
## [1] 738.7202
```

Residuals looks like they are randomly distributed around 0

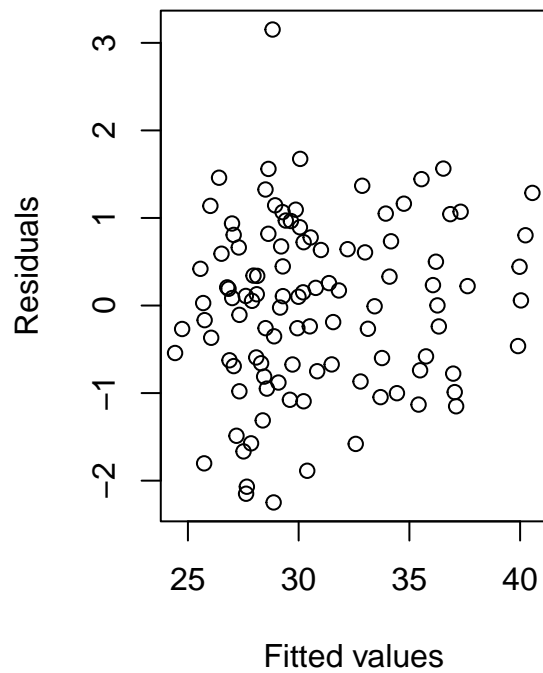


```

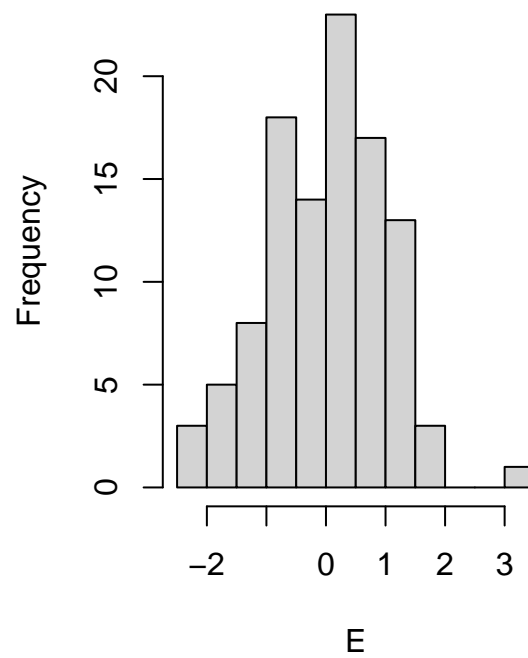
model_as1_pd=model_as.a.pd
E <- resid(model_as1_pd, type = "normalized")
Fit <- fitted(model_as1_pd)
op <- par(mfrow = c(1, 2))
plot(x = Fit, y = E,
      xlab = "Fitted values", ylab = "Residuals",
      main = "Residuals versus fitted values")
hist(E, nclass = 15)

```

**Residuals versus fitted values**



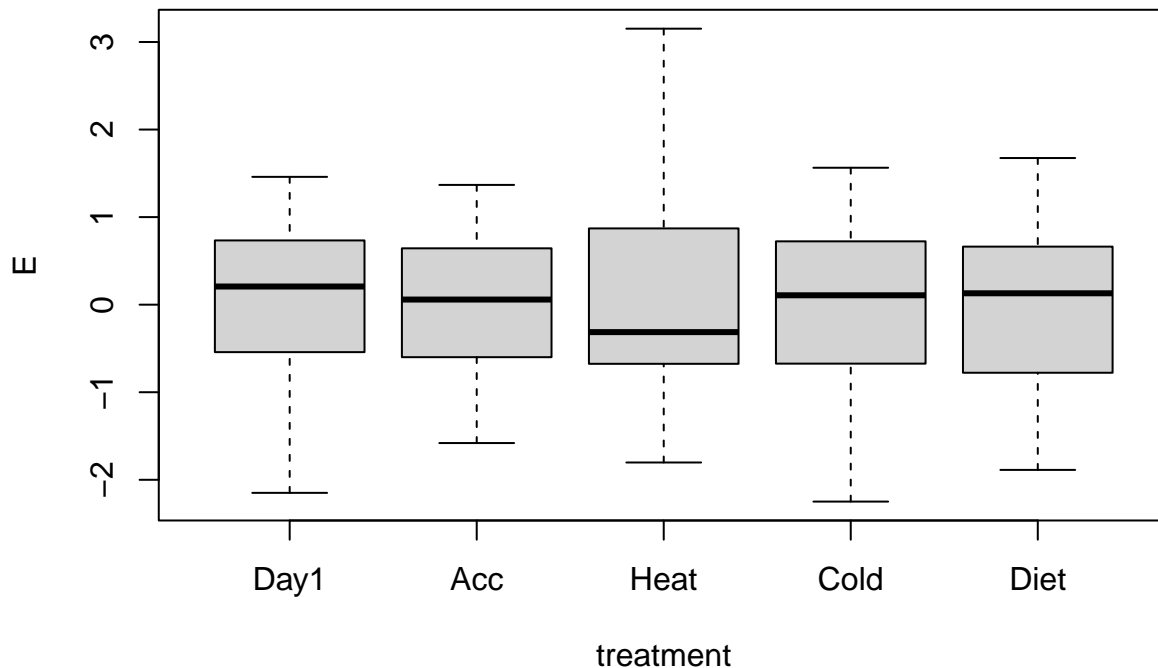
**Histogram of E**



```

par(op)
bp=boxplot(E ~ treatment,
            data = alph_div_pd_as)

```



```
summary(model_as1_pd)
```

```
## Linear mixed-effects model fit by REML
##   Data: alph_div_pd_as
##       AIC      BIC    logLik
##  707.3642 730.8108 -344.6821
##
## Random effects:
## Formula: ~1 | cage
##      (Intercept)
## StdDev:    4.095198
##
## Formula: ~1 | individual_id %in% cage
##      (Intercept) Residual
## StdDev:    0.9957605 6.854969
##
## Correlation Structure: Continuous AR(1)
## Formula: ~as.numeric(days) | cage/individual_id
## Parameter estimate(s):
##      Phi
## 0.8551492
## Fixed effects: diversity ~ treatment
##              Value Std.Error DF   t-value p-value
## (Intercept) 27.950951  2.385399 79 11.717514  0.0000
## treatmentAcc  5.803850  1.740702 79  3.334201  0.0013
## treatmentHeat 1.312992  2.088097 79  0.628798  0.5313
## treatmentCold 2.096134  2.107776 79  0.994476  0.3230
## treatmentDiet 2.876805  2.088303 79  1.377580  0.1722
## Correlation:
##      (Intr) trtmnA trtmnH trtmnC
## treatmentAcc -0.365
## treatmentHeat -0.426  0.477
```

```
## treatmentCold -0.441  0.420  0.534
## treatmentDiet -0.447  0.418  0.494  0.577
##
## Standardized Within-Group Residuals:
##      Min      Q1      Med      Q3      Max
## -2.2488677 -0.7535860  0.1316047  0.7156377  2.3338739
##
## Number of Observations: 105
## Number of Groups:
##           cage individual_id %in% cage
##           5                22
```

```
#####There is an interesting sex effect on Phylogenetic diversity sexMALE  -5.406496  1.414101  3
```

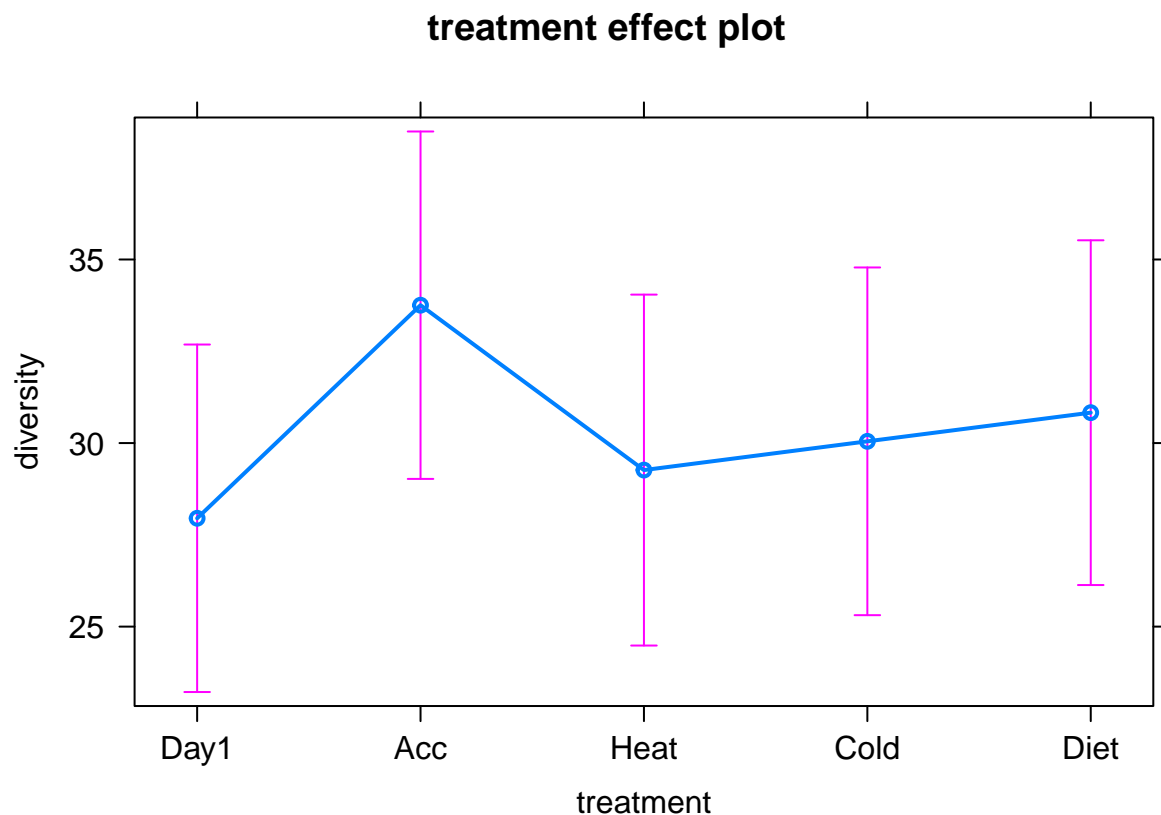
We have a slightly significant treatment effect

```
anova(model_as1_pd)
```

```
##           numDF denDF   F-value p-value
## (Intercept)     1    79 226.64956 <.0001
## treatment       4    79   3.12909  0.0192
```

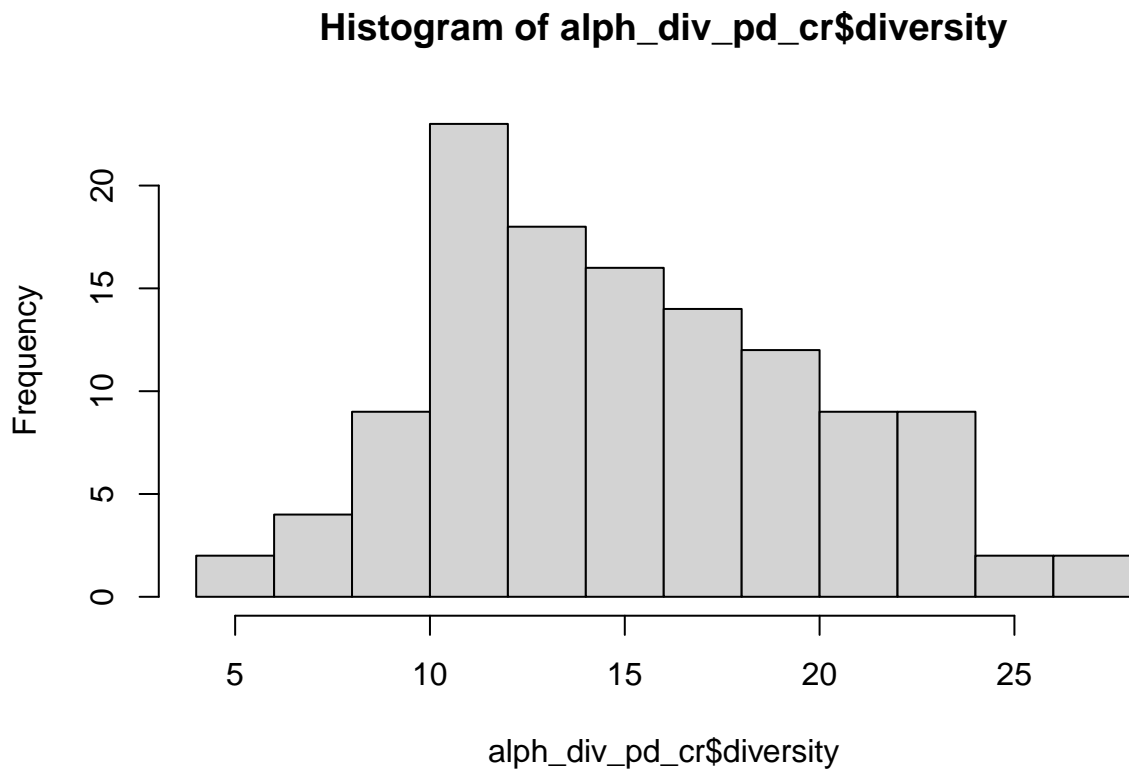
```
#There is a significant treatment and sex effect
#Treatment = 0.0374
#Sex = 0.0105
```

```
plot(allEffects(model_as1_pd))
```



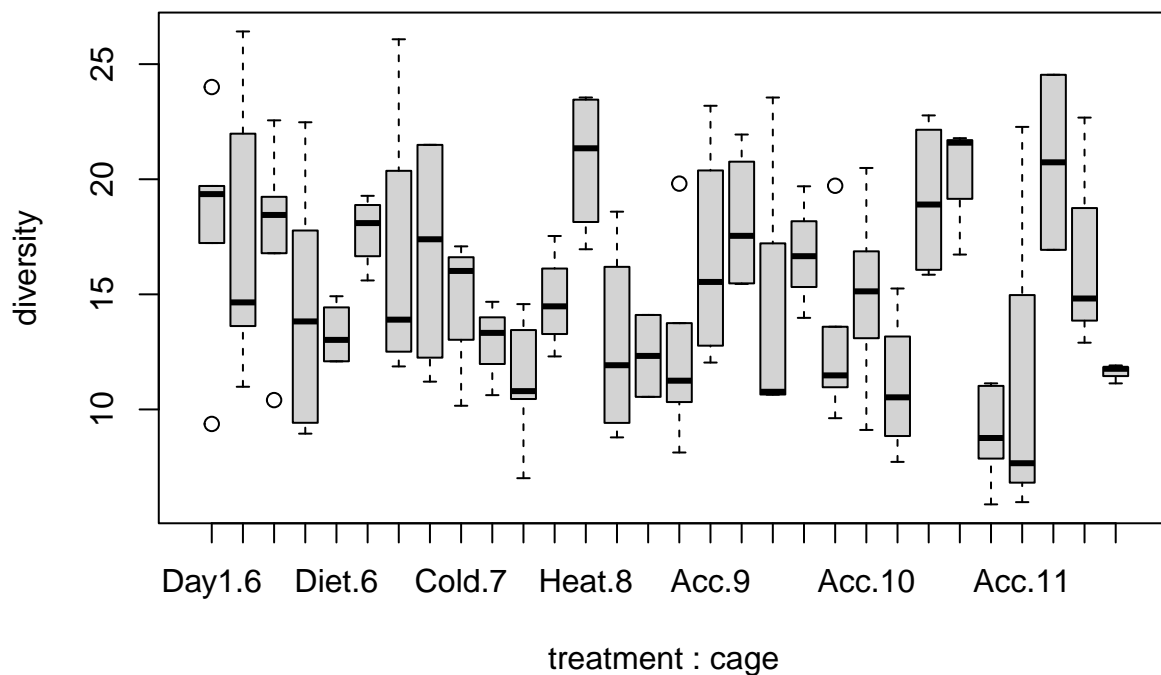
#####Phylogenetic on the Crocidura phylogenetic diversity

```
hist(alph_div_pd_cr$diversity)
```



treatment\_wise\_boxplots from the taxonomic diversity Hill numbers

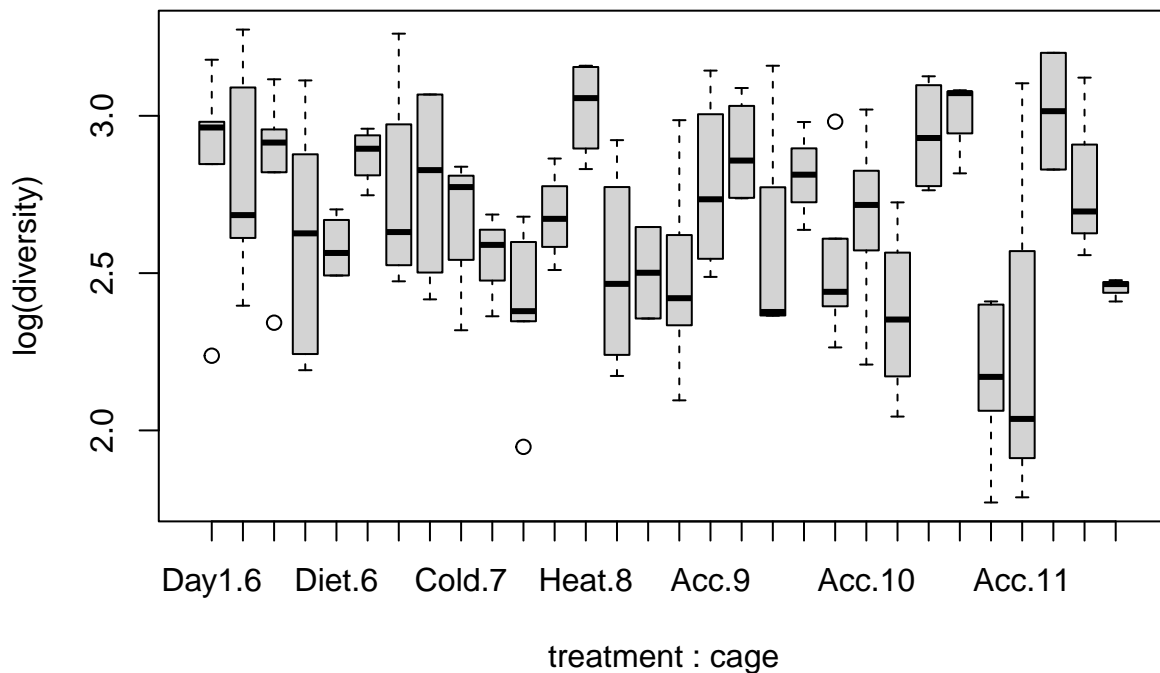
```
boxplot(diversity~treatment*cage,data = alph_div_pd_cr)
```



Assessment of log transformation it looks tighter, however the non-transformed data still looks ok to me so i

will continue with this

```
boxplot(log(diversity)~treatment*cage,data = alph_div_pd_cr)
```



Create

the models for comparison - including the random effects results in the lowest AIC

```
###Try to include the other factors
model_cr.a.pd <- lme(diversity~treatment,
  random = ~1|cage/individual_id,
  correlation = corCAR1(form = ~ as.numeric(days)),
  data=alph_div_pd_cr)
model_cr.b.pd <- lme(diversity~treatment,
  random = ~1|individual_id,
  correlation = corCAR1(form = ~ as.numeric(days)),
  data=alph_div_pd_cr)
model_cr.c.pd <- lme(diversity~treatment,
  random = ~1|cage,
  data=alph_div_pd_cr)
model_cr.d.pd <- lme(diversity~treatment,
  random = ~1|individual_id,
  data=alph_div_pd_cr)
model_cr.e.pd <- lm(diversity~treatment,
  data=alph_div_pd_cr)

anova(model_cr.a.pd,model_cr.b.pd,model_cr.c.pd,model_cr.e.pd)
```

##	Model	df	AIC	BIC	logLik	Test	L.Ratio	p-value
##	model_cr.a.pd	1	9	719.9922	744.6966	-350.9961		
##	model_cr.b.pd	2	8	717.9922	739.9517	-350.9961	1 vs 2	0.0000243801 0.9961
##	model_cr.c.pd	3	7	715.9930	735.2076	-350.9965	2 vs 3	0.0007876453 0.9776
##	model_cr.e.pd	4	6	713.9930	730.4626	-350.9965	3 vs 4	0.0000001073 0.9997

```
AIC(model_cr.a.pd ) #383.197
```

```
## [1] 719.9922
```

```
AIC(model_cr.b.pd ) #381.197
```

```
## [1] 717.9922
```

```
AIC(model_cr.c.pd ) #379.439
```

```
## [1] 715.993
```

```
AIC(model_cr.d.pd ) #379.439
```

```
## [1] 715.9922
```

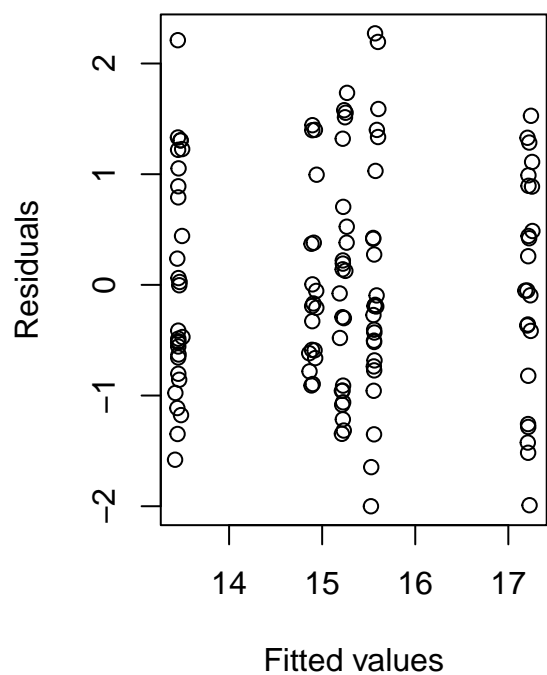
```
AIC(model_cr.e.pd ) #377.9547
```

```
## [1] 722.8743
```

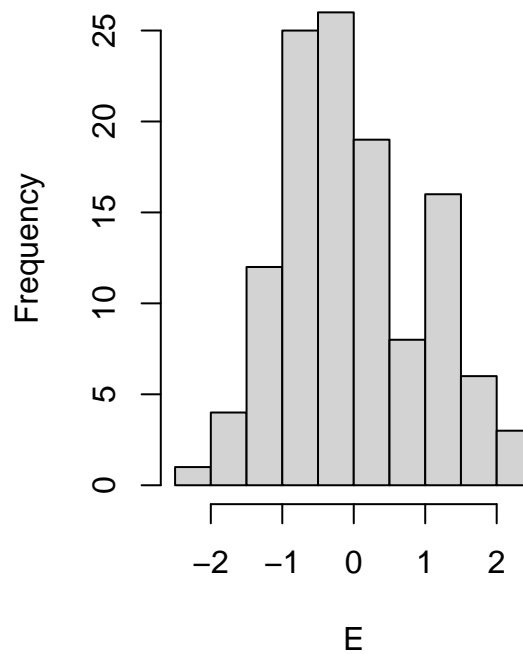
Residuals looks like they are randomly distributed around 0

```
model_cr1_pd=model_cr.a.pd
E <- resid(model_cr1_pd, type = "normalized")
Fit <- fitted(model_cr1_pd)
op <- par(mfrow = c(1, 2))
plot(x = Fit, y = E,
      xlab = "Fitted values", ylab = "Residuals",
      main = "Residuals versus fitted values")
hist(E, nclass = 15)
```

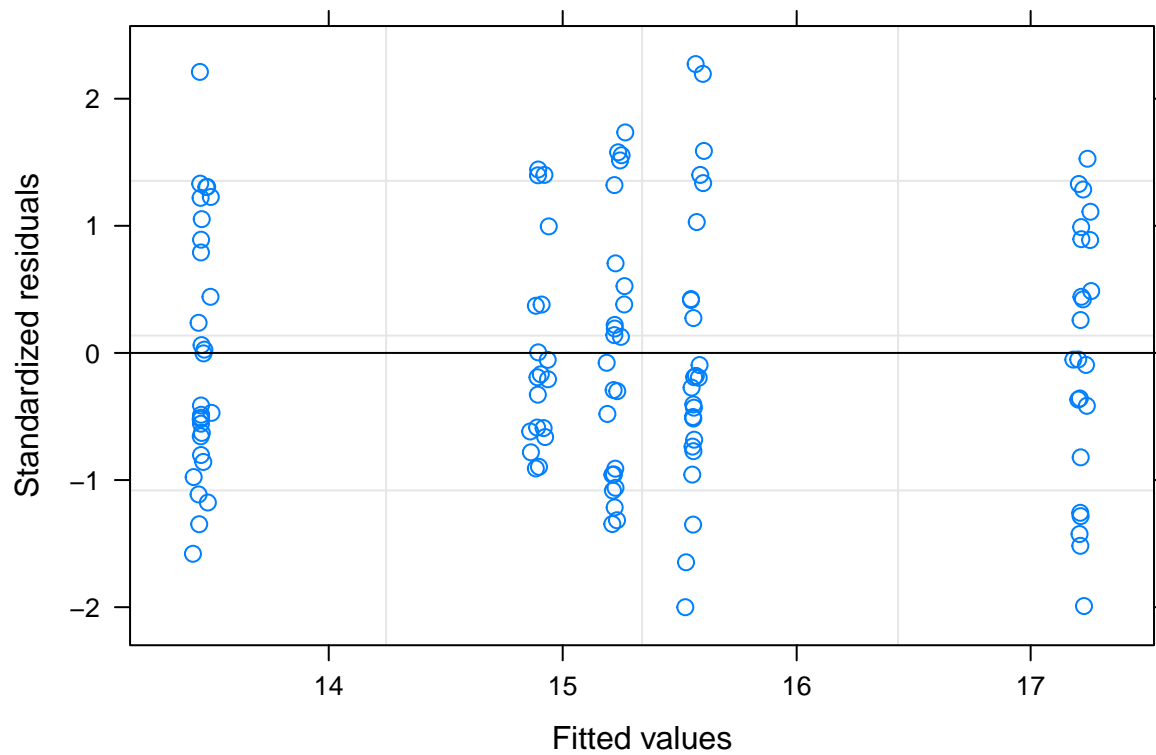
**Residuals versus fitted values**



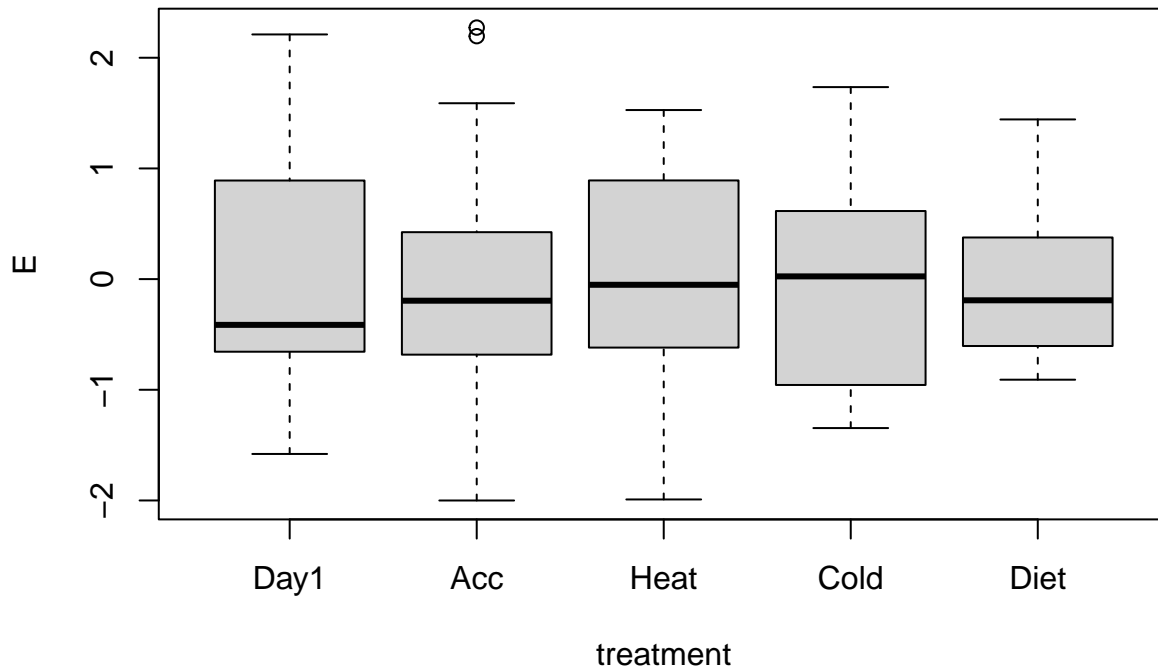
**Histogram of E**



```
plot(model_cr1_pd)
```



```
par(op)
bp=boxplot(E ~ treatment,
           data = alph_div_pd_cr)
```



```
summary(model_cr1_pd)
```

```
## Linear mixed-effects model fit by REML
##   Data: alph_div_pd_cr
##       AIC      BIC    logLik
##  719.9922 744.6966 -350.9961
##
## Random effects:
## Formula: ~1 | cage
##      (Intercept)
## StdDev: 0.0009642443
##
## Formula: ~1 | individual_id %in% cage
##      (Intercept) Residual
## StdDev:   0.2099112 4.774799
##
## Correlation Structure: Continuous AR(1)
## Formula: ~as.numeric(days) | cage/individual_id
## Parameter estimate(s):
##      Phi
## 0.1363406
## Fixed effects: diversity ~ treatment
##              Value Std.Error DF   t-value p-value
## (Intercept) 13.459064 0.8875144 87 15.164897  0.0000
## treatmentAcc  2.104398 1.3032082 87  1.614783  0.1100
## treatmentHeat 3.759593 1.3333436 87  2.819673  0.0060
## treatmentCold 1.768294 1.3177344 87  1.341920  0.1831
```



```
## treatmentDiet 1.441405 1.4095706 87 1.022584 0.3093
## Correlation:
##          (Intr) trtmnA trtmnH trtmnC
## treatmentAcc -0.680
## treatmentHeat -0.664 0.453
## treatmentCold -0.672 0.458 0.448
## treatmentDiet -0.628 0.428 0.418 0.423
##
## Standardized Within-Group Residuals:
##      Min      Q1      Med      Q3      Max
## -2.0001993 -0.6960988 -0.1741151 0.8149149 2.2725295
##
## Number of Observations: 120
## Number of Groups:
##          cage individual_id %in% cage
##          6                      29
```

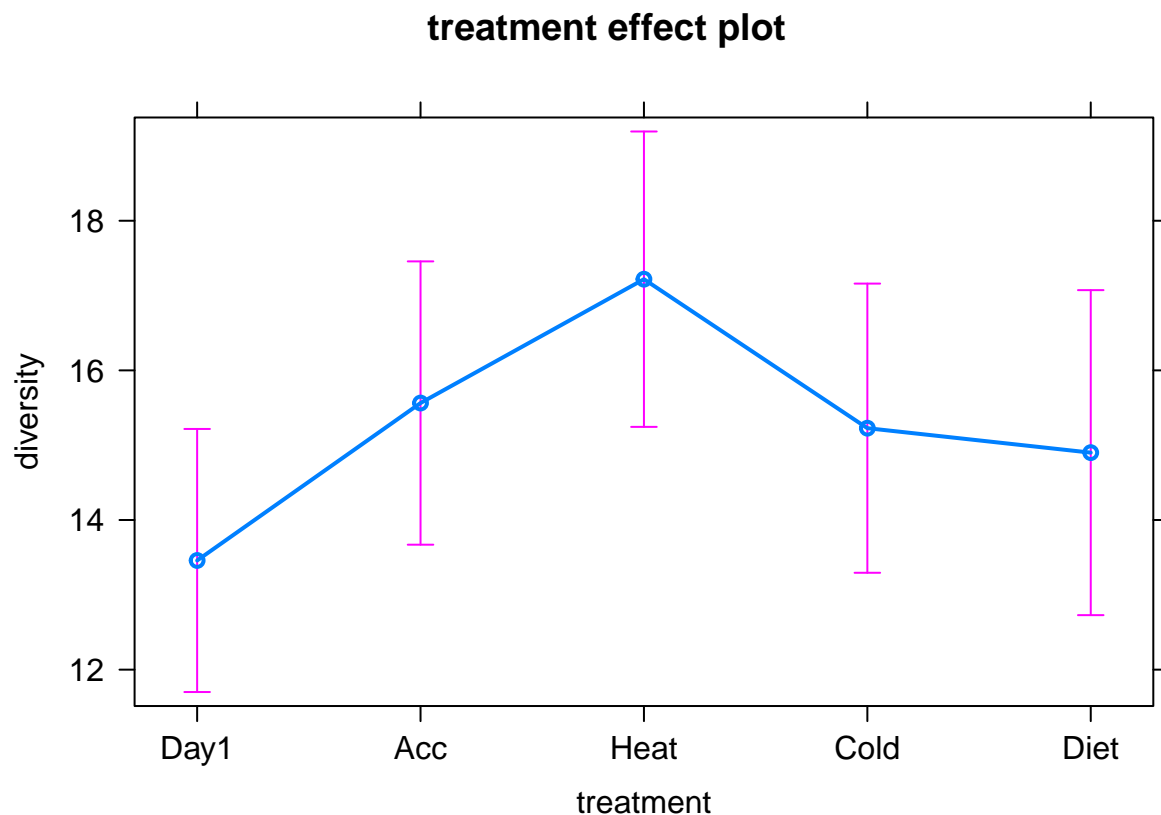
```
#####There is an interesting sex effect on Phylogenetic diversity sexMALE -5.406496 1.414101 3
```

We have a slightly significant treatment effect

```
anova(model_cr1_pd)
```

```
##          numDF denDF  F-value p-value
## (Intercept)      1    87 1205.2315 <.0001
## treatment        4    87   2.0462 0.0949
```

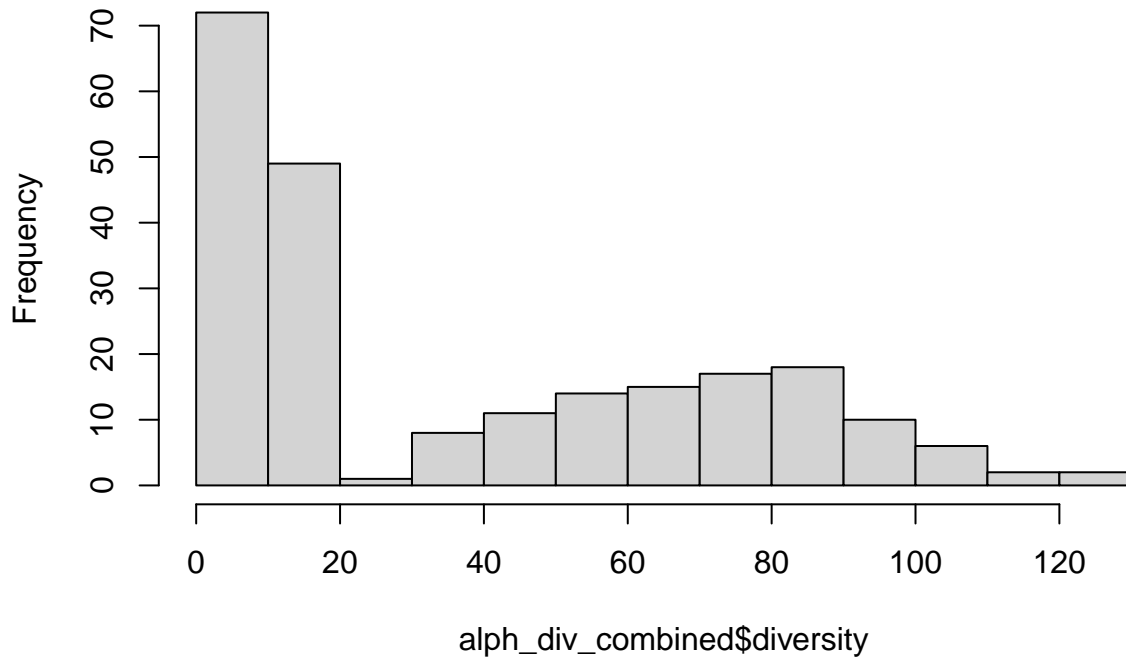
```
plot(allEffects(model_cr1_pd))
```



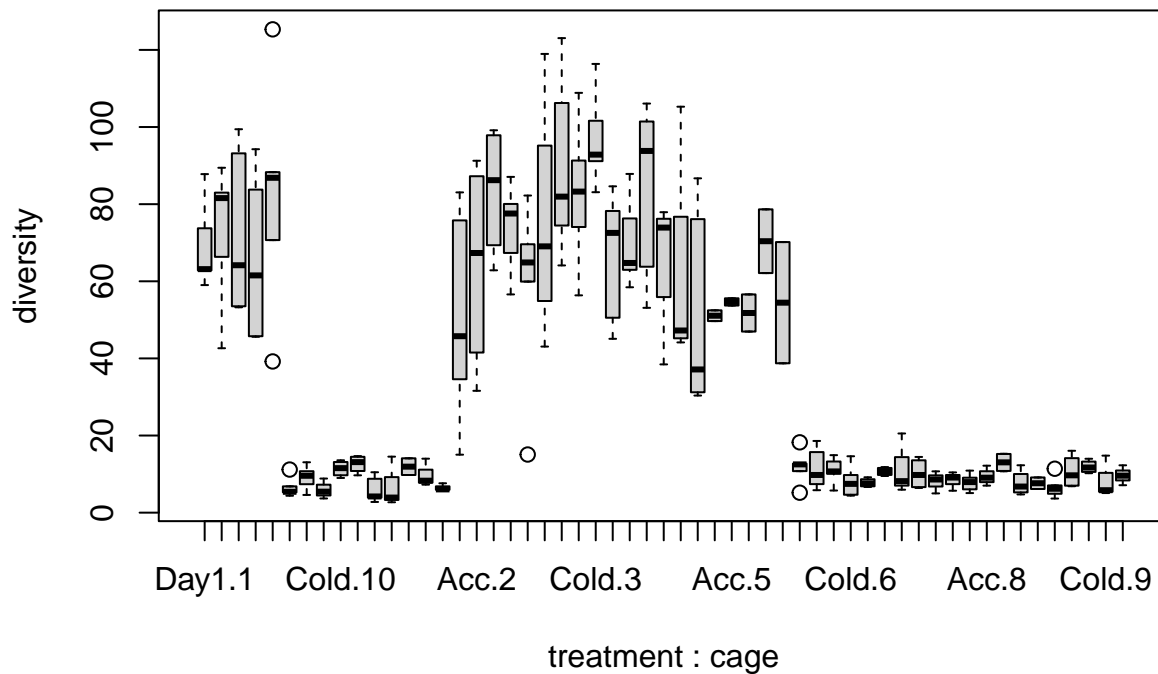
Species-species alpha diversity linear model

```
alph_div_combined <- rbind(alph_div_tax_as, alph_div_tax_cr)  
  
hist(alph_div_combined$diversity)
```

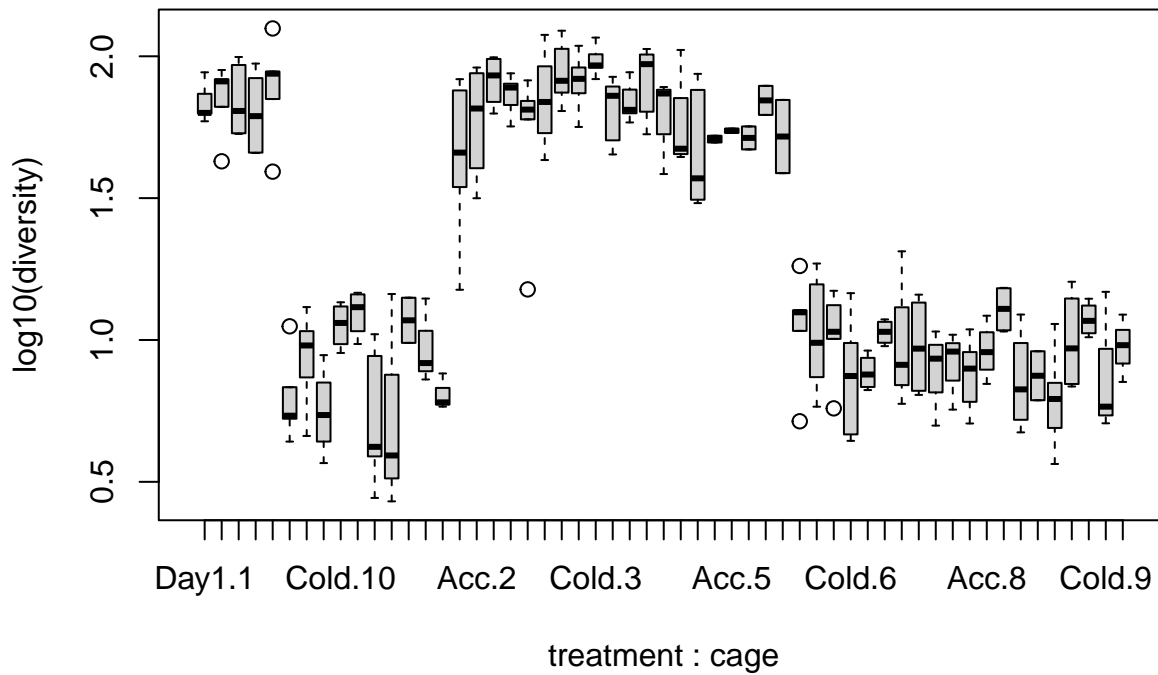
**Histogram of alph\_div\_combined\$diversity**



```
boxplot(diversity~treatment*cage,data = alph_div_combined)
```



```
boxplot(log10(diversity)~treatment*cage,data = alph_div_combined)
```



```
###Try to include the other factors
model_div_combined.a <- lme(log(diversity)~species,
  random = ~1|cage/individual_id,
  correlation = corCAR1(form = ~ as.numeric(days)),
  data=alph_div_combined)
model_div_combined.b <- lme(diversity~species,
  random = ~1|individual_id,
  correlation = corCAR1(form = ~ as.numeric(days)),
  data=alph_div_combined)
model_div_combined.c <- lme(diversity~species,
  random = ~1|cage,
  data=alph_div_combined)
model_div_combined.d <- lme(diversity~species,
  random = ~1|individual_id,
  data=alph_div_combined)
model_div_combined.e <- lm(diversity~species,
  data=alph_div_combined)
```

```
AIC(model_div_combined.a) #1868.937
```

```
## [1] 249.1726
```

```
AIC(model_div_combined.b) #1868.075
```

```
## [1] 1874.692
```

```
AIC(model_div_combined.c) #1875.097
```

```
## [1] 1875.095
```

```
AIC(model_div_combined.d) #1872.694
```

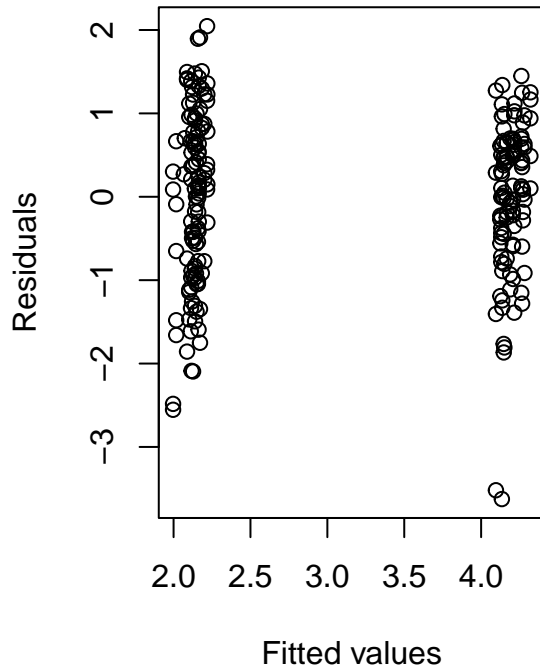
```
## [1] 1872.692
```

```
AIC(model_div_combined.e) #1885.803
```

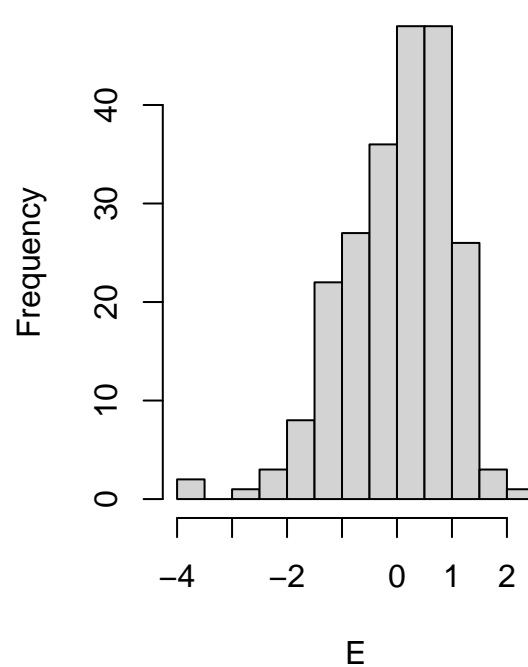
```
## [1] 1885.801
```

```
model_alph_combined=model_div_combined.a  
E <- resid(model_alph_combined, type = "normalized")  
Fit <- fitted(model_alph_combined)  
op <- par(mfrow = c(1, 2))  
plot(x = Fit, y = E,  
      xlab = "Fitted values", ylab = "Residuals",  
      main = "Residuals versus fitted values")  
hist(E, nclass = 15)
```

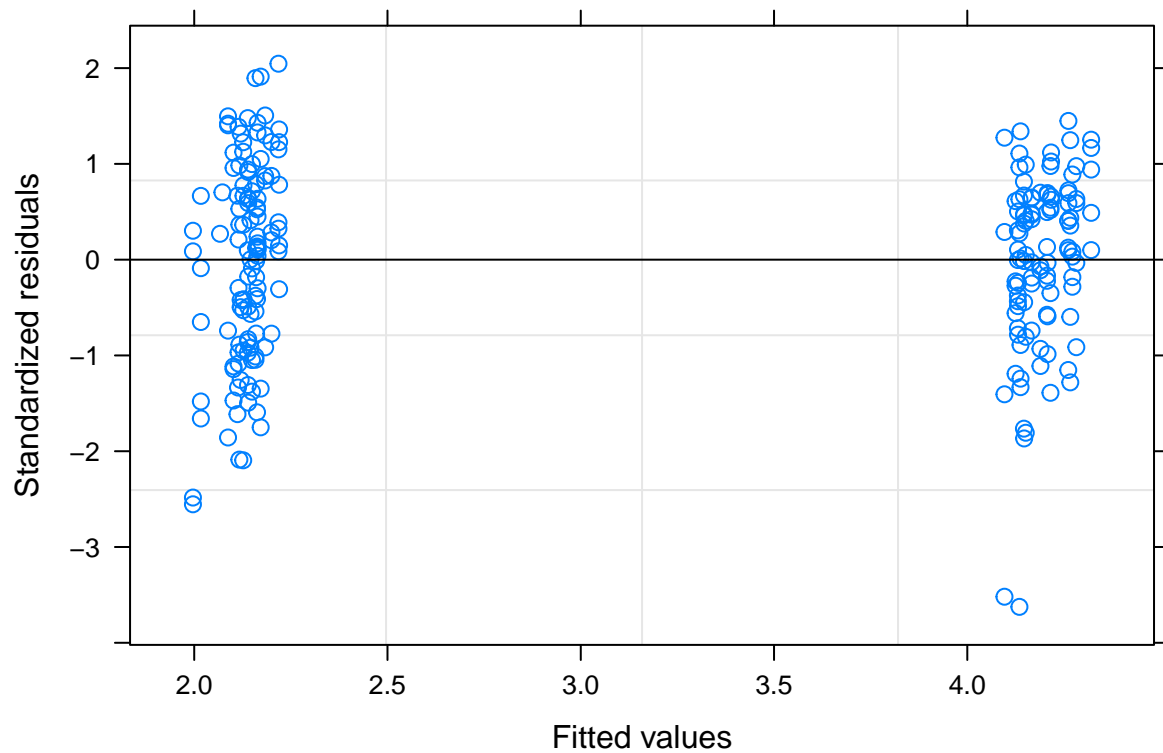
**Residuals versus fitted values**



**Histogram of E**



```
plot(model_alph_combined)
```



```
#par(op)
#bp=boxplot(E ~ species,
#           data = model_div_combined.a)
```

```
summary(model_alph_combined)
```

```
## Linear mixed-effects model fit by REML
##   Data: alph_div_combined
##       AIC      BIC    logLik
##   249.1726 269.6157 -118.5863
##
## Random effects:
## Formula: ~1 | cage
##      (Intercept)
## StdDev:  0.07296413
##
## Formula: ~1 | individual_id %in% cage
##      (Intercept) Residual
## StdDev:  0.07484818 0.3930167
##
## Correlation Structure: Continuous AR(1)
## Formula: ~as.numeric(days) | cage/individual_id
## Parameter estimate(s):
##      Phi
## 0.2000606
## Fixed effects:  log(diversity) ~ species
##               Value Std.Error DF   t-value p-value
## (Intercept)  4.187049 0.05336890 174   78.45484      0
## speciesCR   -2.054082 0.07242374   9  -28.36199      0
```

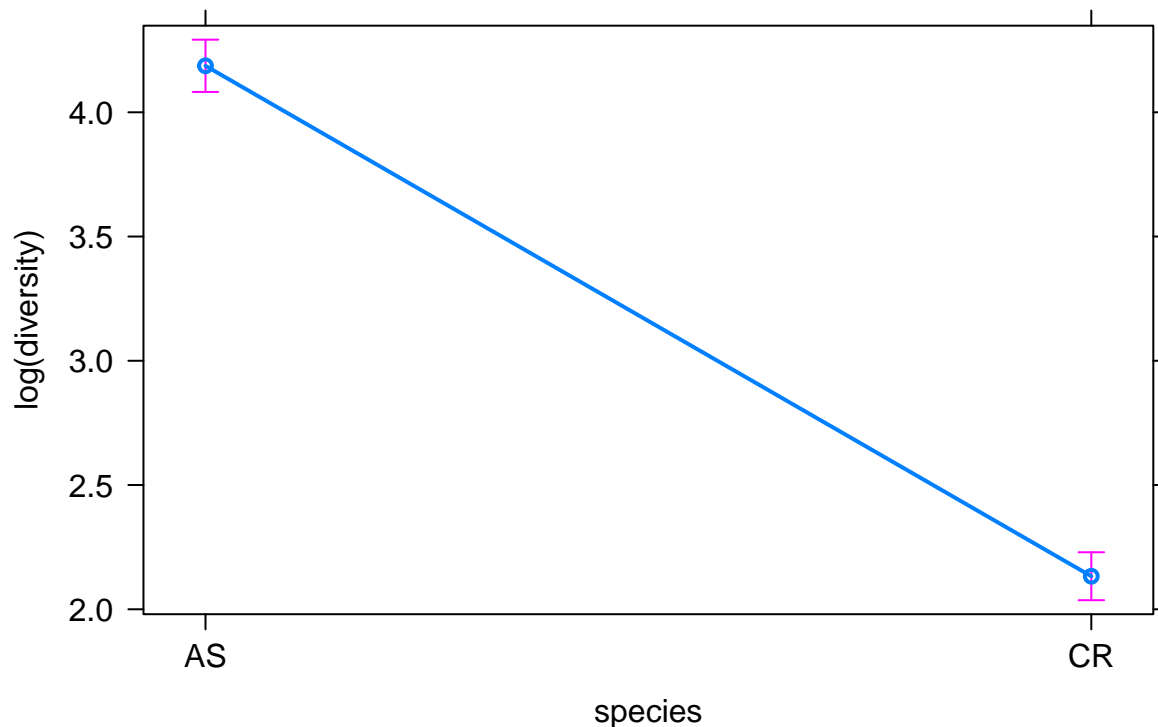
```
## Correlation:
##      (Intr)
## speciesCR -0.737
##
## Standardized Within-Group Residuals:
##      Min      Q1      Med      Q3      Max
## -3.6249532 -0.5974494  0.1201001  0.6679768  2.0459622
##
## Number of Observations: 225
## Number of Groups:
##           cage individual_id %in% cage
##           11                51
```

```
anova(model_alph_combined)
```

```
##           numDF denDF  F-value p-value
## (Intercept)     1   174 7248.809 <.0001
## species         1     9  804.403 <.0001
```

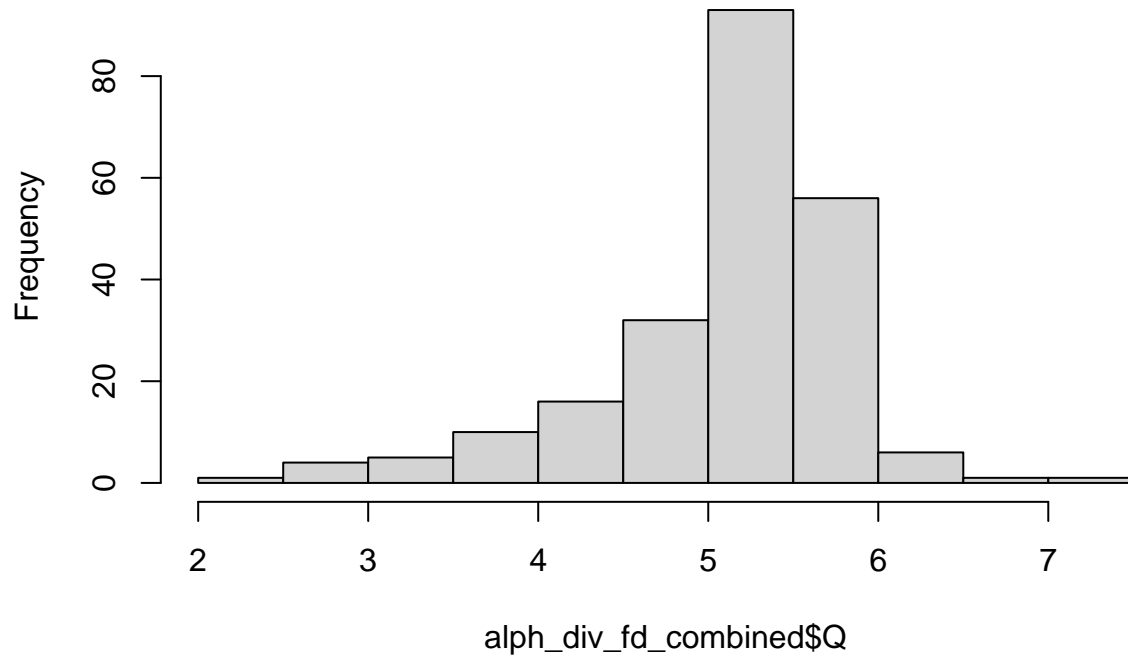
```
plot(allEffects(model_alph_combined))
```

### species effect plot

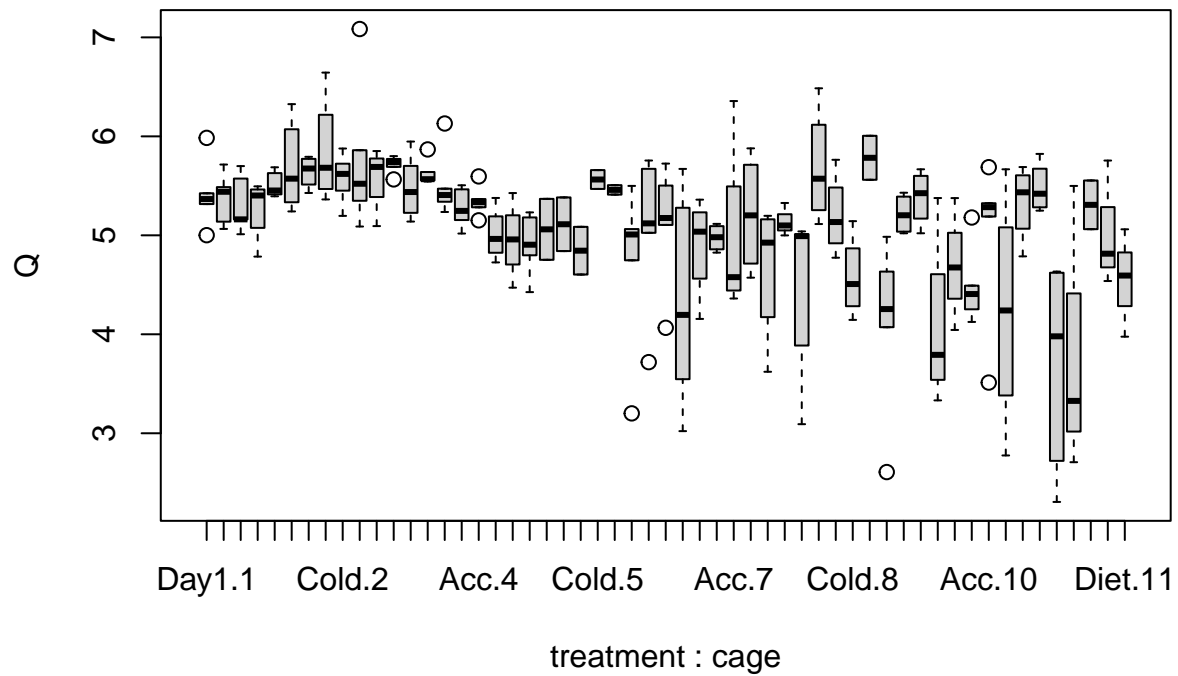


```
###Try to include the other factors
alph_div_fd_combined <- rbind(alph_div_fd_as, alph_div_fd_cr)
hist(alph_div_fd_combined$Q)
```

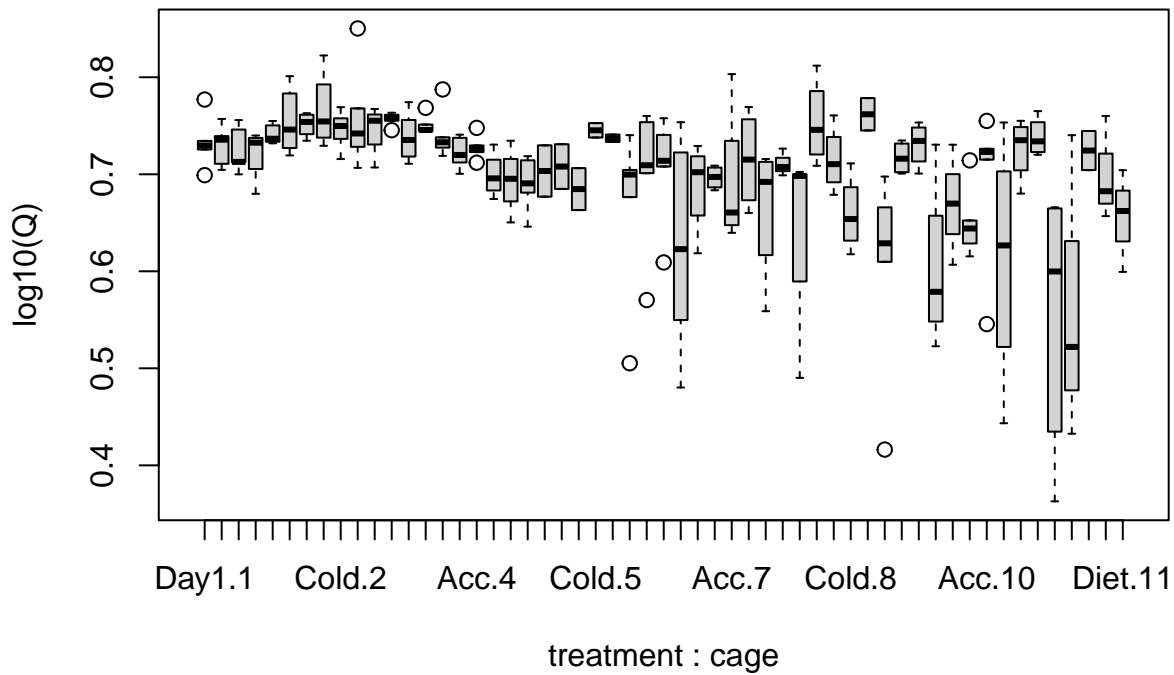
Histogram of alph\_div\_fd\_combined\$Q



```
boxplot(Q~treatment*cage,data = alph_div_fd_combined)
```



```
boxplot(log10(Q)~treatment*cage,data = alph_div_fd_combined)
```



```
model_div_fd_combined.a <- lme(Q~species,
  random = ~1|cage/individual_id,
  correlation = corCAR1(form = ~ as.numeric(days)),
  data=alph_div_fd_combined)
model_div_fd_combined.b <- lme(Q~species,
  random = ~1|individual_id,
  correlation = corCAR1(form = ~ as.numeric(days)),
  data=alph_div_fd_combined)
model_div_fd_combined.c <- lme(Q~species,
  random = ~1|cage,
  data=alph_div_fd_combined)
model_div_fd_combined.d <- lme(Q~species,
  random = ~1|individual_id,
  data=alph_div_fd_combined)
model_div_fd_combined.e <- lm(Q~species,
  data=alph_div_fd_combined)
```

```
AIC(model_div_fd_combined.a) #940.832
```

```
## [1] 464.4085
```

```
AIC(model_div_fd_combined.b) #939.4217
```

```
## [1] 466.2917
```



```
AIC(model_div_fd_combined.c) #941.6157
```

```
## [1] 460.9967
```

```
AIC(model_div_fd_combined.d) #937.4217
```

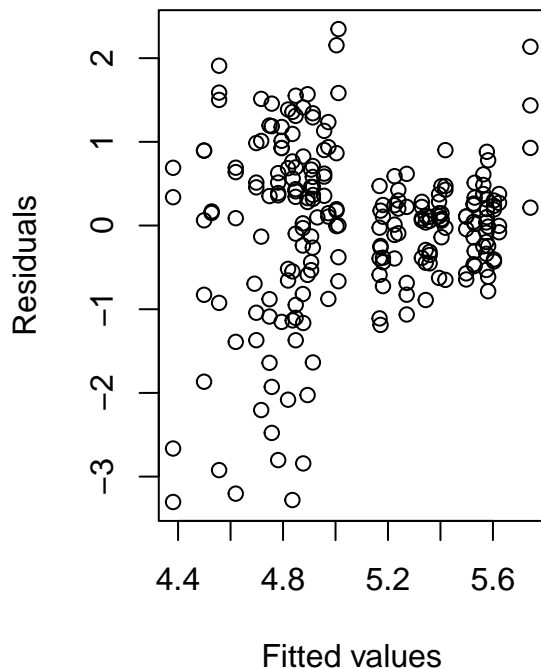
```
## [1] 464.2916
```

```
AIC(model_div_fd_combined.e) #939.4188
```

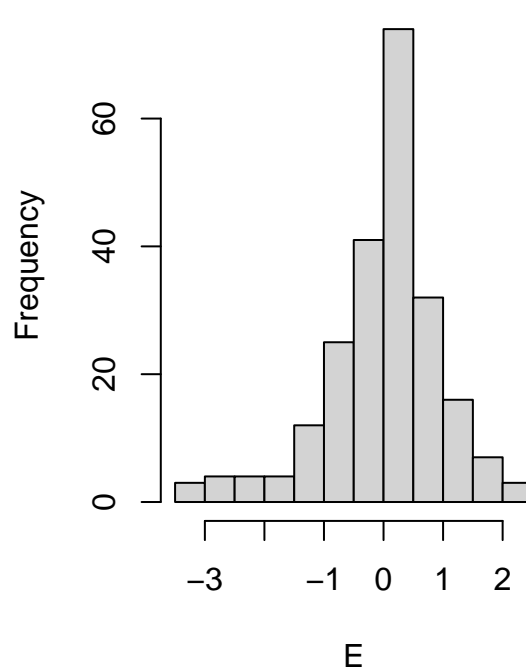
```
## [1] 458.8496
```

```
model_alph_fd_combined=model_div_fd_combined.a  
E <- resid(model_alph_fd_combined, type = "normalized")  
Fit <- fitted(model_alph_fd_combined)  
op <- par(mfrow = c(1, 2))  
plot(x = Fit, y = E,  
      xlab = "Fitted values", ylab = "Residuals",  
      main = "Residuals versus fitted values")  
hist(E, nclass = 15)
```

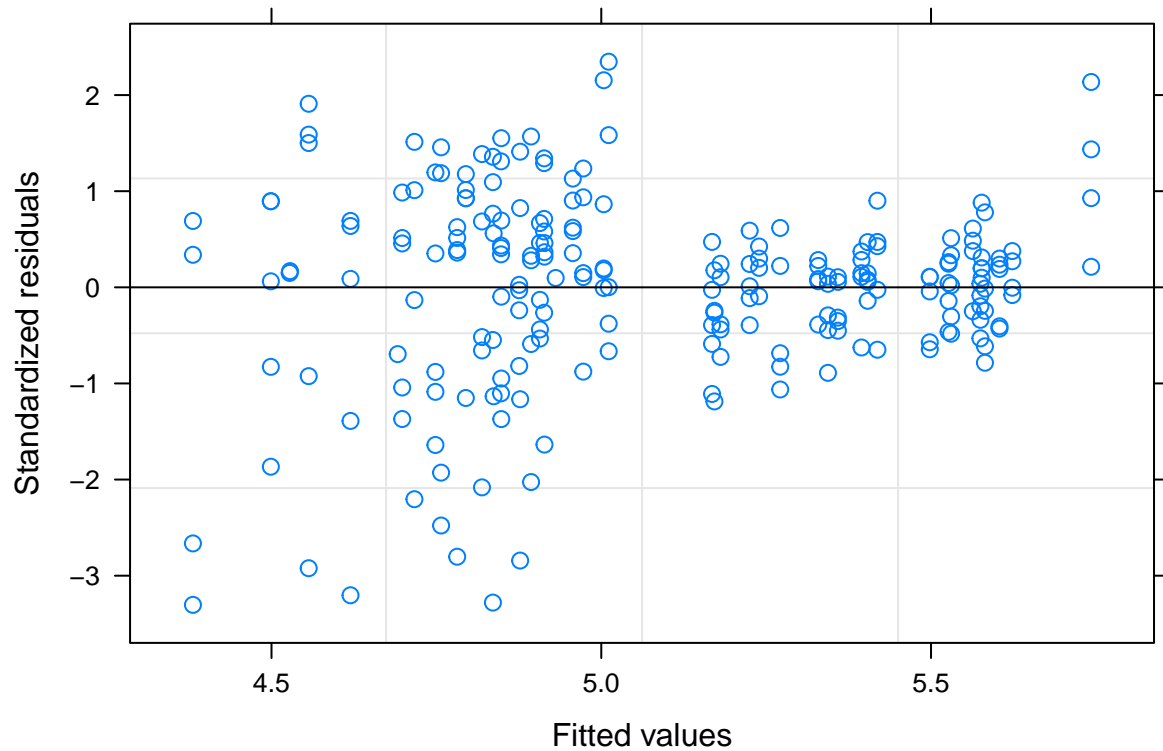
**Residuals versus fitted values**



**Histogram of E**



```
plot(model_alph_fd_combined)
```



```
#par(op)
#bp=boxplot(E ~ species,
#           data = model_div_combined.a)
```

```
summary(model_alph_fd_combined)
```

```
## Linear mixed-effects model fit by REML
##   Data: alph_div_fd_combined
##       AIC      BIC    logLik
##  464.4085 484.8515 -226.2042
##
## Random effects:
## Formula: ~1 | cage
##      (Intercept)
## StdDev:   0.1965538
##
## Formula: ~1 | individual_id %in% cage
##      (Intercept) Residual
## StdDev:   0.1345171 0.6280142
##
## Correlation Structure: Continuous AR(1)
## Formula: ~as.numeric(days) | cage/individual_id
## Parameter estimate(s):
##      Phi
## 0.1998651
## Fixed effects: Q ~ species
##               Value Std.Error DF  t-value p-value
## (Intercept)  5.411605 0.1122283 174 48.21961  0.0000
## speciesCR   -0.623273 0.1518137   9 -4.10551  0.0027
```

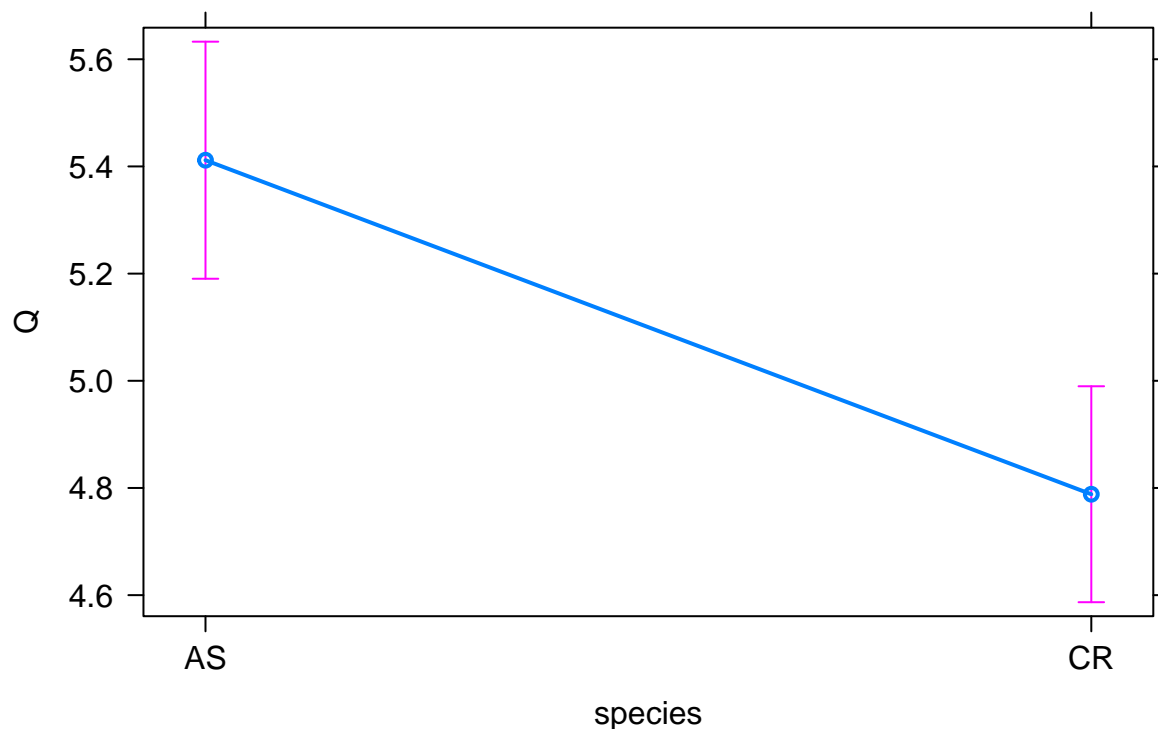
```
## Correlation:
##      (Intr)
## speciesCR -0.739
##
## Standardized Within-Group Residuals:
##      Min      Q1      Med      Q3      Max
## -3.3036141 -0.4388384  0.1065470  0.5129885  2.3472157
##
## Number of Observations: 225
## Number of Groups:
##           cage individual_id %in% cage
##           11                51
```

```
anova(model_alph_fd_combined)
```

```
##           numDF denDF  F-value p-value
## (Intercept)     1   174 4501.890 <.0001
## species         1     9   16.855  0.0027
```

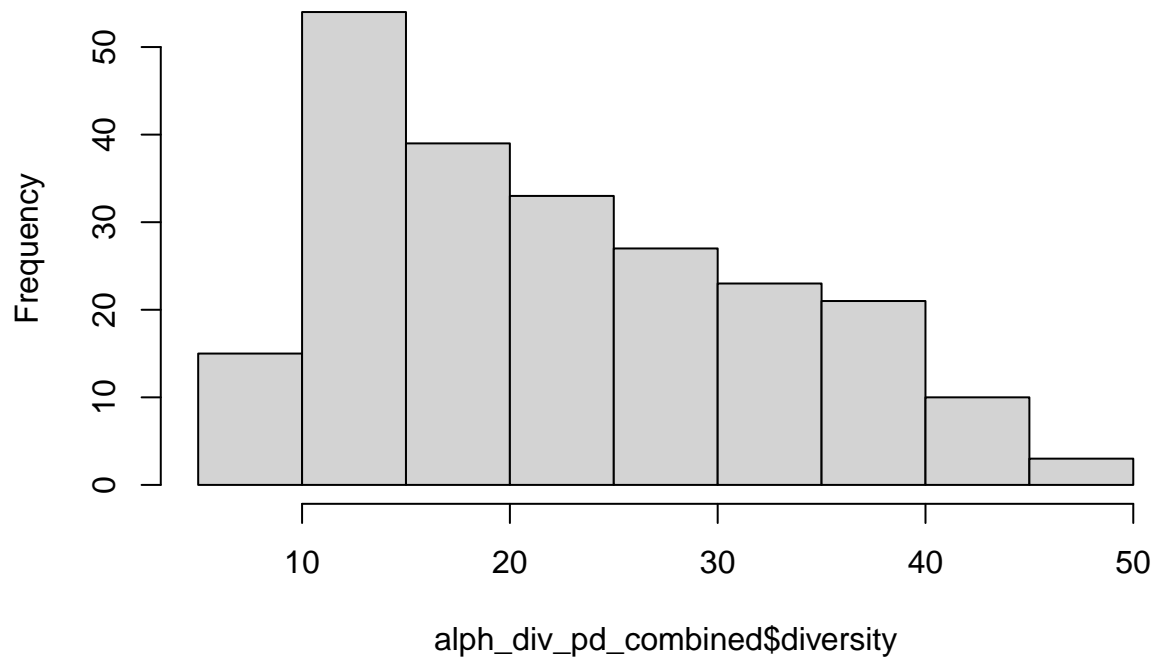
```
plot(allEffects(model_alph_fd_combined))
```

### species effect plot

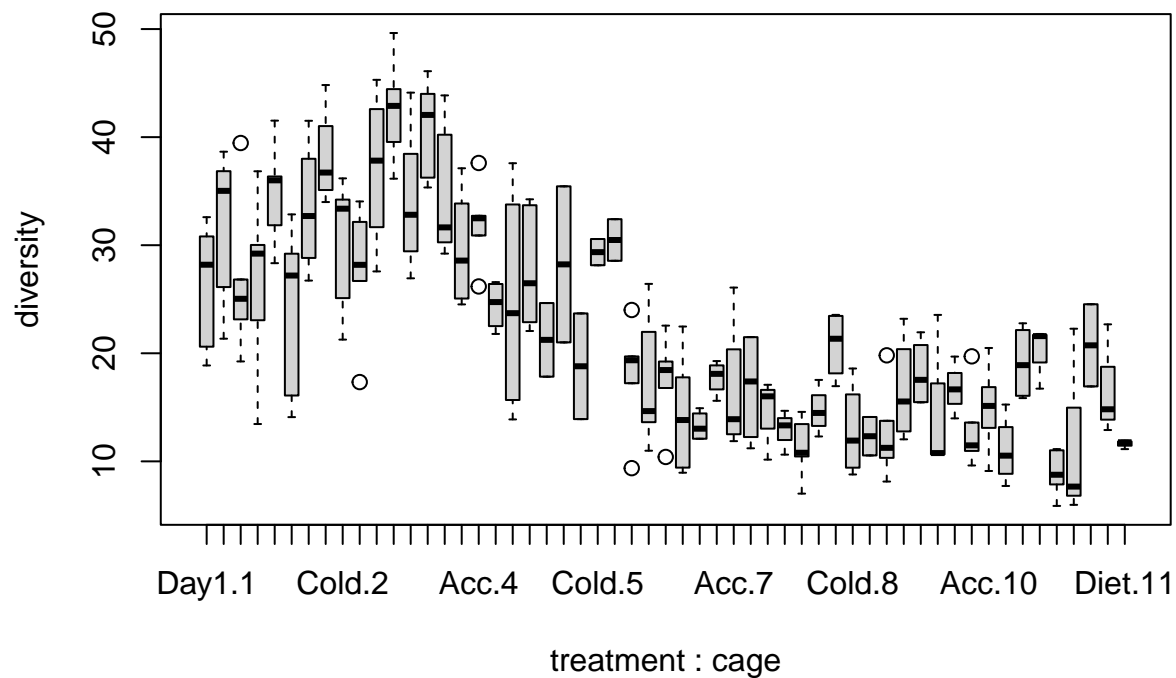


```
###Try to include the other factors
alph_div_pd_combined <- rbind(alph_div_pd_as, alph_div_pd_cr)
hist(alph_div_pd_combined$diversity)
```

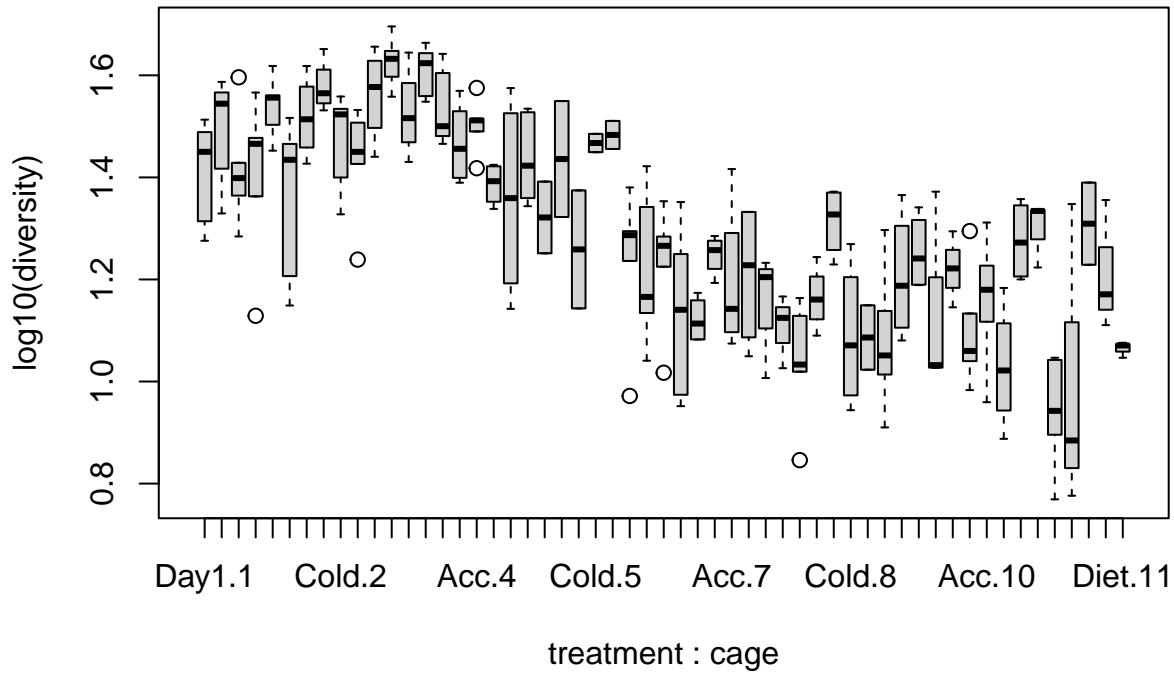
# Histogram of alph\_div\_pd\_combined\$diversity



```
boxplot(diversity~treatment*cage,data = alph_div_pd_combined)
```



```
boxplot(log10(diversity)~treatment*cage,data = alph_div_pd_combined)
```



```
model_div_pd_combined.a <- lme(diversity~species,
  random = ~1|cage/individual_id,
  correlation = corCAR1(form = ~ as.numeric(days)),
  data=alph_div_pd_combined)
model_div_pd_combined.b <- lme(diversity~species,
  random = ~1|individual_id,
  correlation = corCAR1(form = ~ as.numeric(days)),
  data=alph_div_pd_combined)
model_div_pd_combined.c <- lme(diversity~species,
  random = ~1|cage,
  data=alph_div_pd_combined)
model_div_pd_combined.d <- lme(diversity~species,
  random = ~1|individual_id,
  data=alph_div_pd_combined)
model_div_pd_combined.e <- lm(diversity~species,
  data=alph_div_pd_combined)
```

```
AIC(model_div_pd_combined.a) #940.832
```

```
## [1] 1464.487
```

```
AIC(model_div_pd_combined.b) #939.4217
```

```
## [1] 1473.215
```

```
AIC(model_div_pd_combined.c) #941.6157
```

```
## [1] 1462.028
```

```
AIC(model_div_pd_combined.d) #937.4217
```

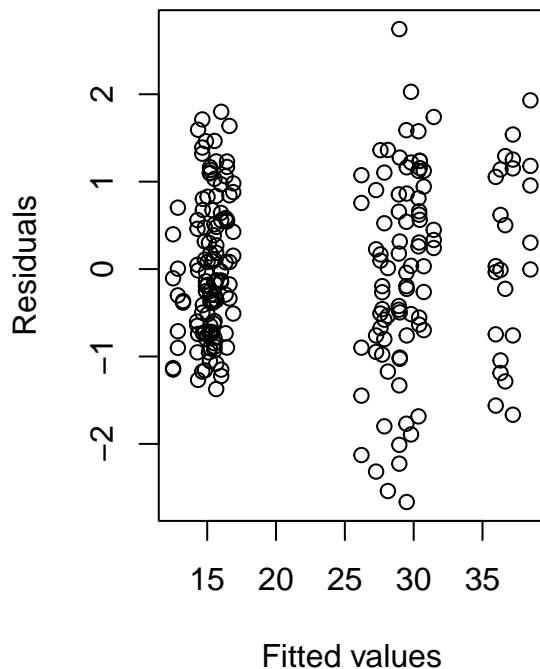
```
## [1] 1471.215
```

```
AIC(model_div_pd_combined.e) #939.4188
```

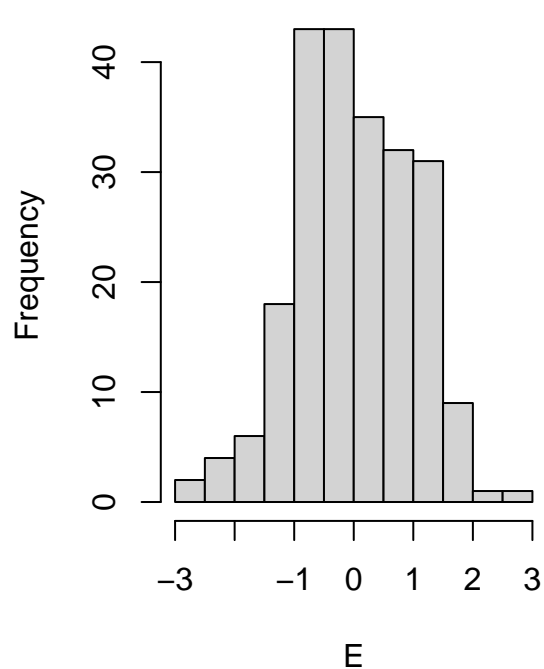
```
## [1] 1486.344
```

```
model_alph_pd_combined=model_div_pd_combined.a  
E <- resid(model_alph_pd_combined, type = "normalized")  
Fit <- fitted(model_alph_pd_combined)  
op <- par(mfrow = c(1, 2))  
plot(x = Fit, y = E,  
      xlab = "Fitted values", ylab = "Residuals",  
      main = "Residuals versus fitted values")  
hist(E, nclass = 15)
```

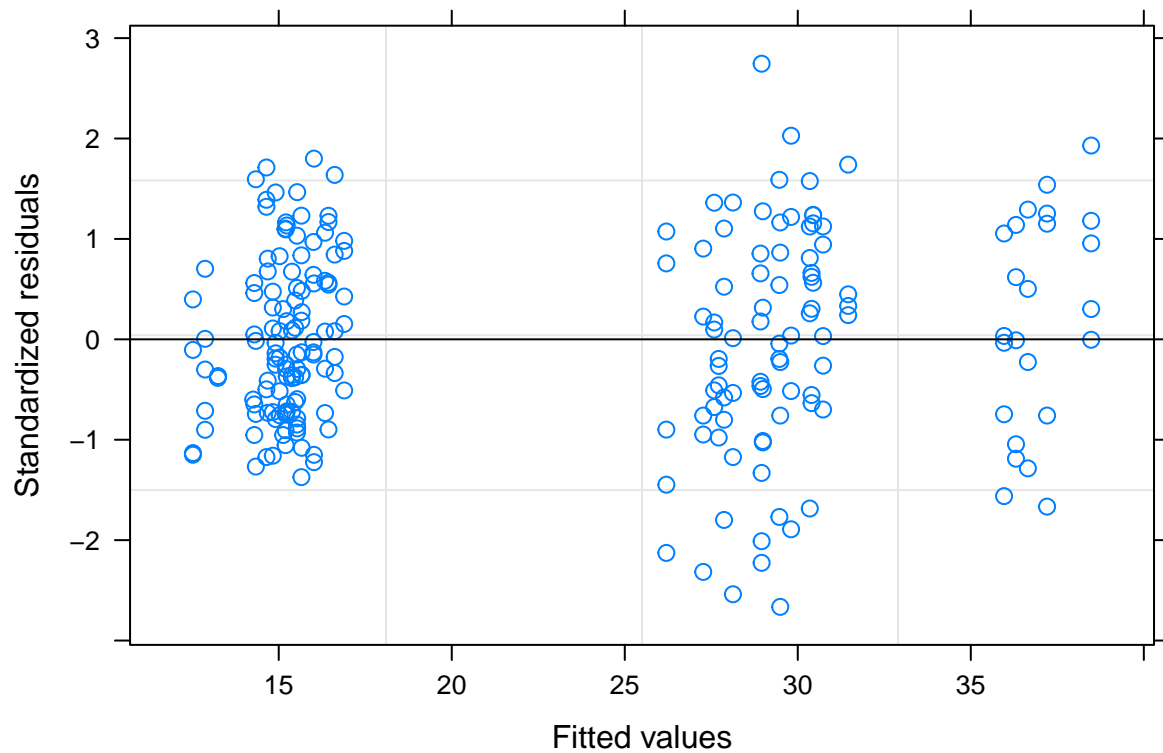
**Residuals versus fitted values**



**Histogram of E**



```
plot(model_alph_pd_combined)
```



```
#par(op)
#bp=boxplot(E ~ species,
#           data = model_div_combined.a)
```

```
summary(model_alph_pd_combined)
```

```
## Linear mixed-effects model fit by REML
##   Data: alph_div_pd_combined
##       AIC      BIC    logLik
##  1464.487 1484.93 -726.2435
##
## Random effects:
## Formula: ~1 | cage
##      (Intercept)
## StdDev:    2.822594
##
## Formula: ~1 | individual_id %in% cage
##      (Intercept) Residual
## StdDev:    1.609097  5.77962
##
## Correlation Structure: Continuous AR(1)
## Formula: ~as.numeric(days) | cage/individual_id
## Parameter estimate(s):
##      Phi
## 0.1998991
## Fixed effects: diversity ~ species
##              Value Std.Error DF   t-value p-value
## (Intercept)  30.42152  1.438039 174  21.154860      0
## speciesCR   -15.36202  1.943356   9  -7.904892      0
```

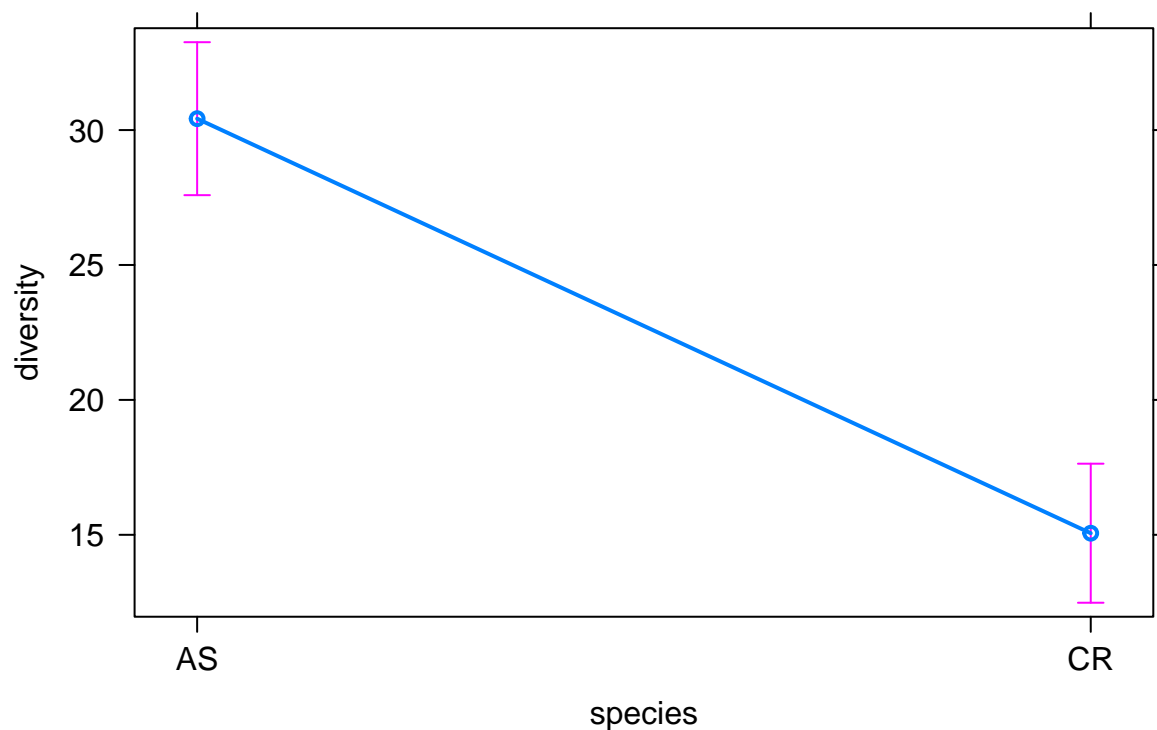
```
## Correlation:
##      (Intr)
## speciesCR -0.74
##
## Standardized Within-Group Residuals:
##      Min      Q1      Med      Q3      Max
## -2.66450014 -0.71754267 -0.02619648  0.70257927  2.74515129
##
## Number of Observations: 225
## Number of Groups:
##           cage individual_id %in% cage
##           11                51
```

```
anova(model_alph_pd_combined)
```

```
##           numDF denDF  F-value p-value
## (Intercept)     1   174 517.7681 <.0001
## species         1     9  62.4873 <.0001
```

```
plot(allEffects(model_alph_pd_combined))
```

### species effect plot



```
sessionInfo()
```

```
## R version 4.2.2 (2022-10-31)
## Platform: x86_64-pc-linux-gnu (64-bit)
## Running under: Linux Mint 20.1
```



```

##
## Matrix products: default
## BLAS: /usr/lib/x86_64-linux-gnu/blas/libblas.so.3.9.0
## LAPACK: /usr/lib/x86_64-linux-gnu/lapack/liblapack.so.3.9.0
##
## locale:
## [1] LC_CTYPE=en_US.UTF-8 LC_NUMERIC=C
## [3] LC_TIME=en_US.UTF-8 LC_COLLATE=en_US.UTF-8
## [5] LC_MONETARY=en_US.UTF-8 LC_MESSAGES=en_US.UTF-8
## [7] LC_PAPER=en_US.UTF-8 LC_NAME=C
## [9] LC_ADDRESS=C LC_TELEPHONE=C
## [11] LC_MEASUREMENT=en_US.UTF-8 LC_IDENTIFICATION=C
##
## attached base packages:
## [1] stats graphics grDevices utils datasets methods base
##
## other attached packages:
## [1] effects_4.2-2 carData_3.0-5 sjPlot_2.8.11 jtools_2.2.0
## [5] lme4_1.1-31 Matrix_1.5-1 tinytex_0.42 multcomp_1.4-20
## [9] TH.data_1.1-1 MASS_7.3-58.1 survival_3.4-0 mvtnorm_1.1-3
## [13] ggplotify_0.1.0 broom_1.0.1 nlme_3.1-160 forcats_0.5.2
## [17] stringr_1.4.1 dplyr_1.0.10 purrr_0.3.5 readr_2.1.3
## [21] tidyr_1.2.1 tibble_3.1.8 ggplot2_3.3.6 tidyverse_1.3.2
##
## loaded via a namespace (and not attached):
## [1] fs_1.5.2 lubridate_1.8.0 insight_0.18.5
## [4] httr_1.4.4 tools_4.2.2 backports_1.4.1
## [7] utf8_1.2.2 R6_2.5.1 sjlabelled_1.2.0
## [10] DBI_1.1.3 colorspace_2.0-3 nnet_7.3-18
## [13] withr_2.5.0 tidyselect_1.2.0 emmeans_1.8.1-1
## [16] compiler_4.2.2 performance_0.10.0 cli_3.4.1
## [19] rvest_1.0.3 xml2_1.3.3 sandwich_3.0-2
## [22] bayestestR_0.13.0 scales_1.2.1 digest_0.6.30
## [25] yulab.utils_0.0.5 minqa_1.2.5 rmarkdown_2.17
## [28] pkgconfig_2.0.3 htmltools_0.5.3 highr_0.9
## [31] dbplyr_2.2.1 fastmap_1.1.0 rlang_1.0.6
## [34] readxl_1.4.1 rstudioapi_0.14 gridGraphics_0.5-1
## [37] generics_0.1.3 zoo_1.8-11 jsonlite_1.8.3
## [40] googlesheets4_1.0.1 magrittr_2.0.3 parameters_0.19.0
## [43] Rcpp_1.0.9 munsell_0.5.0 fansi_1.0.3
## [46] lifecycle_1.0.3 stringi_1.7.8 yaml_2.3.6
## [49] grid_4.2.2 sjmisc_2.8.9 crayon_1.5.2
## [52] lattice_0.20-45 ggeffects_1.1.3 haven_2.5.1
## [55] splines_4.2.2 pander_0.6.5 sjstats_0.18.1
## [58] hms_1.1.2 knitr_1.40 pillar_1.8.1
## [61] boot_1.3-28 estimability_1.4.1 effectsize_0.8.1
## [64] codetools_0.2-18 reprex_2.0.2 glue_1.6.2
## [67] evaluate_0.17 mitools_2.4 modelr_0.1.9
## [70] vctrs_0.5.0 nloptr_2.0.3 tzdb_0.3.0
## [73] cellranger_1.1.0 gtable_0.3.1 assertthat_0.2.1
## [76] datawizard_0.6.2 xfun_0.34 xtable_1.8-4
## [79] survey_4.1-1 coda_0.19-4 googledrive_2.0.0
## [82] gargle_1.2.1 ellipsis_0.3.2

```