

The effects of Range expansion on *Xenopus laevis* tadpoles.

by Natasha Kruger

Three main questions have been addressed.

1. What was the effect of range expansion on the survival of tadpoles?
2. What was the effect of range expansion on the development of tadpoles?
3. What was the effect of range expansion on the morphology of tadpoles?

1. What was the effect of range expansion on the survival of tadpoles?

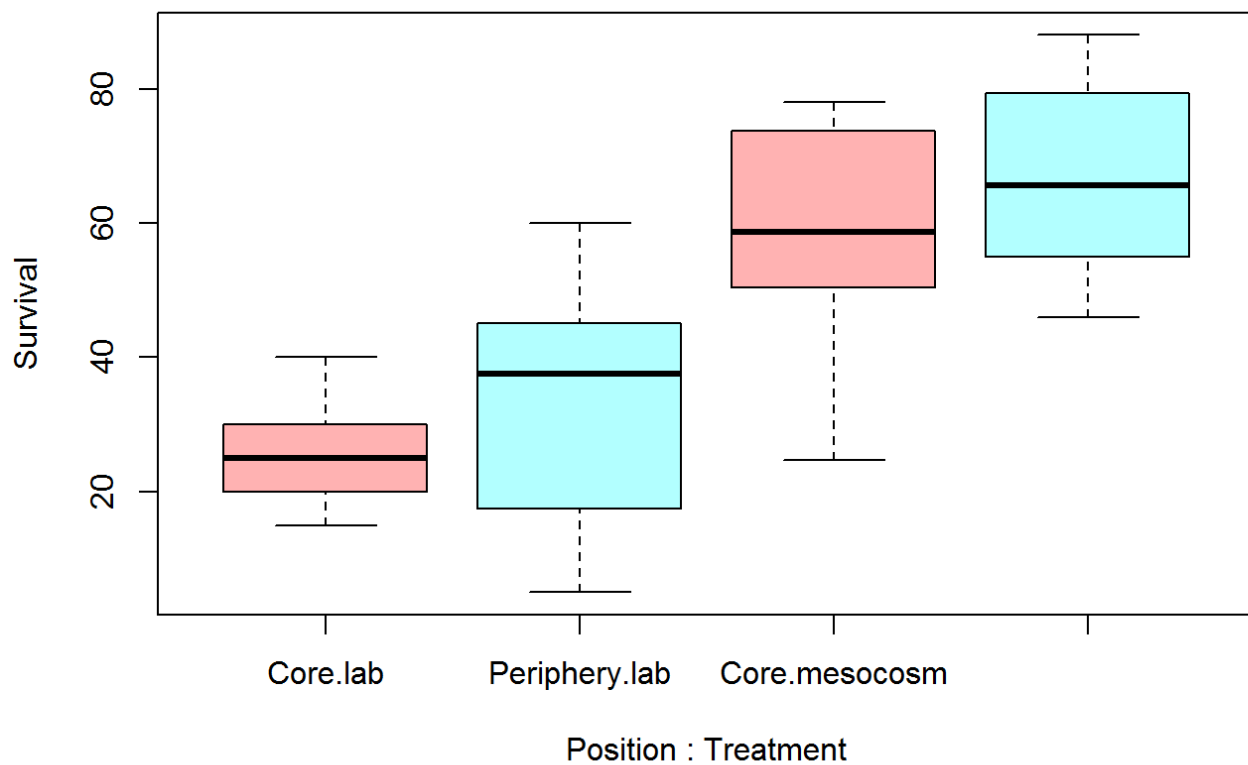
- Overall we removed 50 individuals from the mesocosms and no individuals were removed from the lab containers. We calculated the final abundance for mesocosms (n/150) and for lab (n/20). Generalised binomial mixed models were used with died=0, and survived=1 as response variable with position (core/periphery) as fixed effects. Clutch nested within site was considered as a random effect. We compared different models to determine the best fitted model by comparing Akaike information criterion (AIC).

Use the *Over survival.csv* file to create the plot. I edited the headings in Illustrator.

```
Overallsurvival <- read.csv(file.choose(),header=T)
names(Overallsurvival)
```

```
## [1] "Treatment" "code"      "Site"      "Clutch"    "Clutchx"
## [6] "Position"  "totalraw"  "diedraw"   "Survival"  "diedperc"
```

```
attach(Overallsurvival)
boxplot(Survival~Position*Treatment, col= rainbow(2, alpha = 0.3))
```



```
detach(Overallsurvival)
```

To analyse survival between core and periphery in the lab, use the *Labsurvbinom.csv* file to calculate Lab survival of tadpoles. 1= survived, 0= died.

```
library(glmTMB)
library(bbmle)
```

```
## Loading required package: stats4
```

```
library(lme4)
```

```
## Loading required package: Matrix
```

```
Labsurvbinom=read.csv(file.choose(),header=T)
names(Labsurvbinom)
```

```
## [1] "Treatment" "Position"  "Site"      "Clutch"    "Clutchx"   "Tadpole"
## [7] "Survival"
```

```
attach(Labsurvbinom)

summary(Position) #fixed effect
```

```
##      Core Periphery
##      220      240
```

```
summary(Site) #Random effect
```

```
##      Bouille  Challones  Jardinerie      Massais  Saint-Paul      Toutiere
##      80      80      80      60      80      80
```

```
summary(Clutchx) #Nested within site
```

```
##      Bouille1      Bouille2      Bouille3      Bouille4  Challones1  Challones2
##      20      20      20      20      20      20
##      Challones3  Challones4  Jardinerie1  Jardinerie2  Jardinerie3  Jardinerie4
##      20      20      20      20      20      20
##      Massais1      Massais2      Massais4  Saint-Paul1  Saint-Paul2  Saint-Paul3
##      20      20      20      20      20      20
##      Saint-Paul4  Toutiere1  Toutiere2  Toutiere3  Toutiere4
##      20      20      20      20      20
```

```
lsurv1<- glmer(Survival ~ Position + (1|Site/Clutchx), data = Labsurvbinom, family=
"binomial")
```

```
## boundary (singular) fit: see ?isSingular
```

```
lsurv2<- glmer(Survival ~ 1 + (1|Site/Clutchx), data=Labsurvbinom, family= "binomia
l", glmerControl(optimizer = "bobyqa", optCtrl = list(maxfun = 1e+05))) #some error
s occurred so I control for more itterations
```

```
## boundary (singular) fit: see ?isSingular
```

```
AIC(lsurv1, lsurv2)
```

```
##      df      AIC
## lsurv1  4 556.2981
## lsurv2  3 556.3007
```

```
AICctab(lsurv1, lsurv2) #No difference between null model and model with position a
s fixed effect.
```

```
##      dAICc df
## lsurv2 0      3
## lsurv1 0      4
```

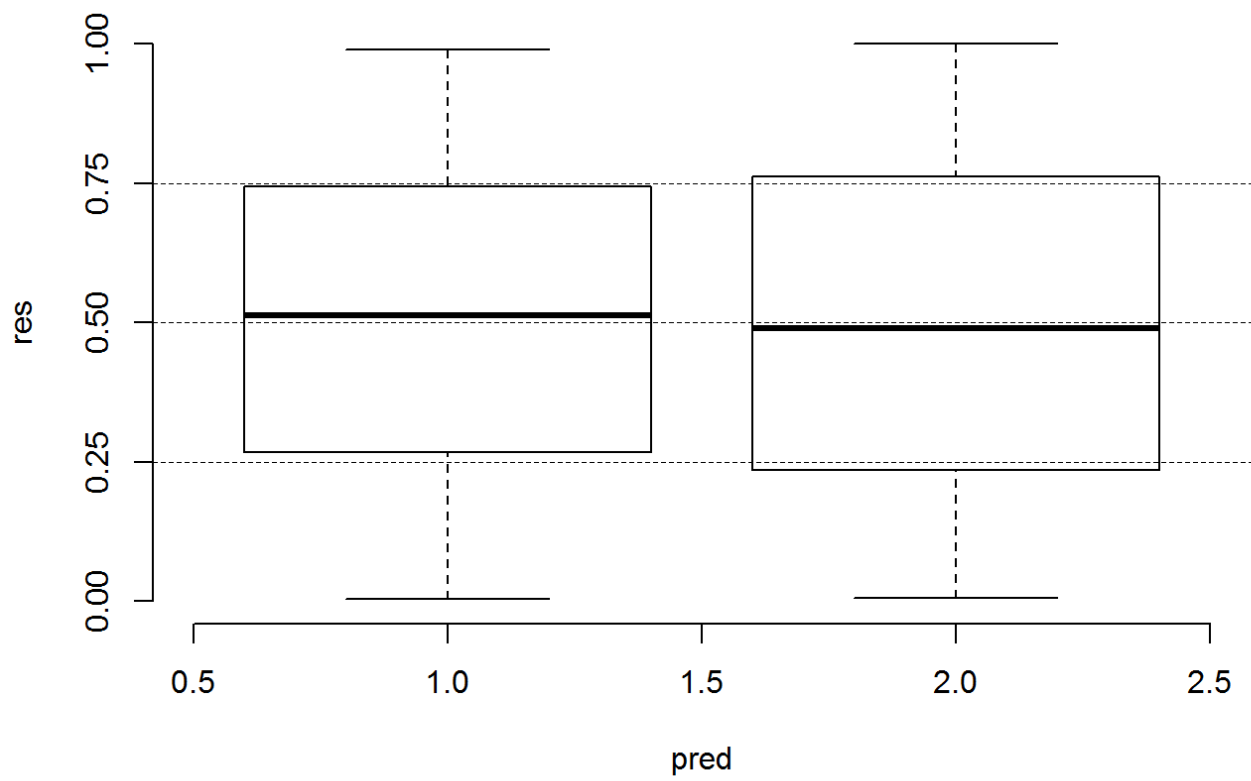
```
summary(lsurv1)
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace
## Approximation) [glmerMod]
## Family: binomial ( logit )
## Formula: Survival ~ Position + (1 | Site/Clutchx)
## Data: Labsurvbinom
##
##      AIC      BIC   logLik deviance df.resid
##  556.3    572.8   -274.1    548.3     456
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -0.8429 -0.6262 -0.5570  1.3235  1.8909
##
## Random effects:
## Groups           Name             Variance Std.Dev.
## Clutchx:Site (Intercept) 0.1116     0.334
## Site           (Intercept) 0.0000     0.000
## Number of obs: 460, groups: Clutchx:Site, 23; Site, 6
##
## Fixed effects:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)    -1.1212     0.1874  -5.983 2.19e-09 ***
## PositionPeriphery  0.3674     0.2524   1.456   0.146
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation of Fixed Effects:
##              (Intr)
## PostnPrphry -0.735
## convergence code: 0
## boundary (singular) fit: see ?isSingular
```

```
library(DHARMA) ## Analyse rÃ©sidus modÃ©les
```

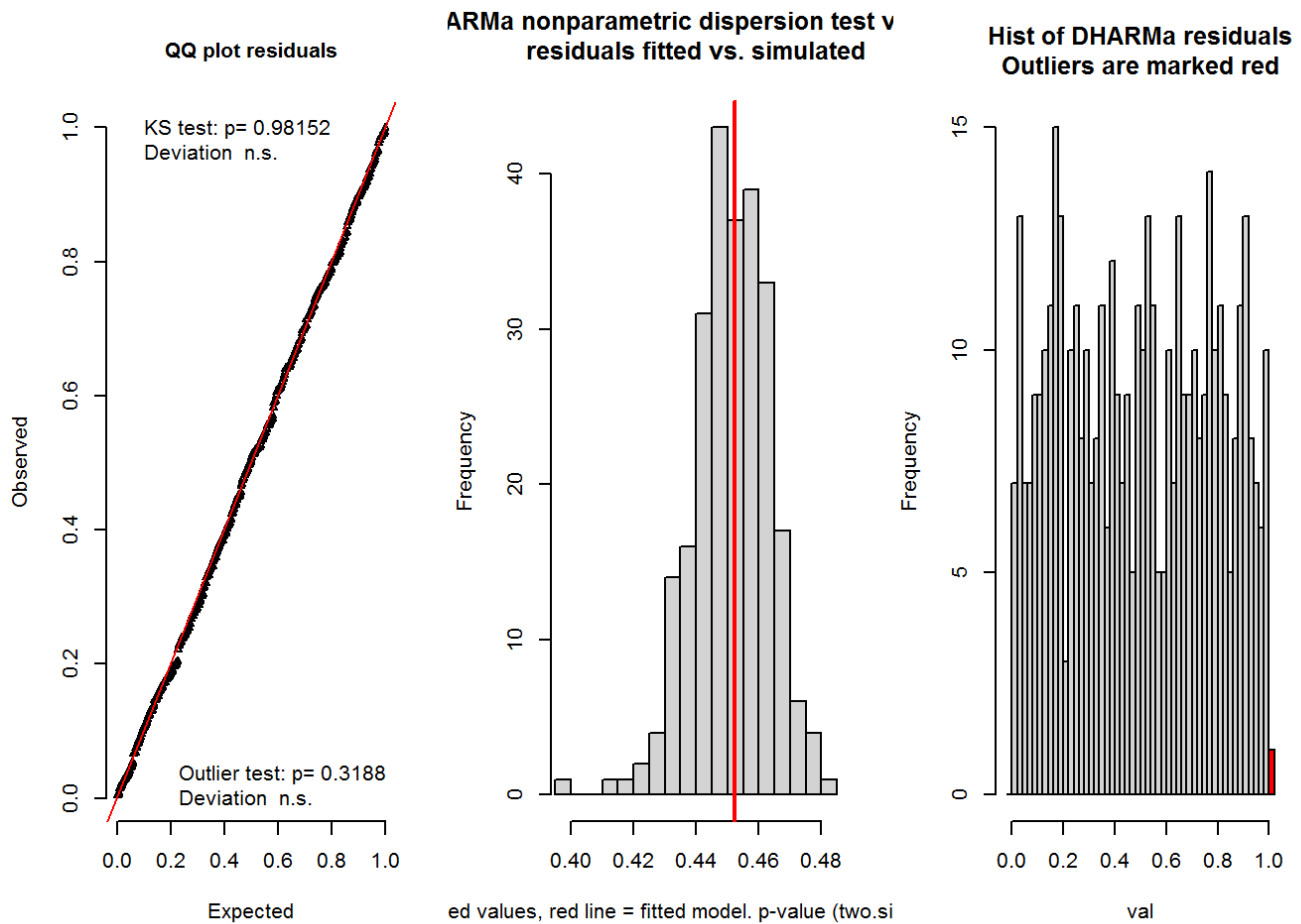
```
## Registered S3 method overwritten by 'DHARMA':
##   method      from
##   refit.glmmTMB glmmTMB
```

```
rr=simulateResiduals(lsurv1) ### simulate residuals
plotResiduals(rr) ### qqplot simulated vs observed
```



```
## NULL
```

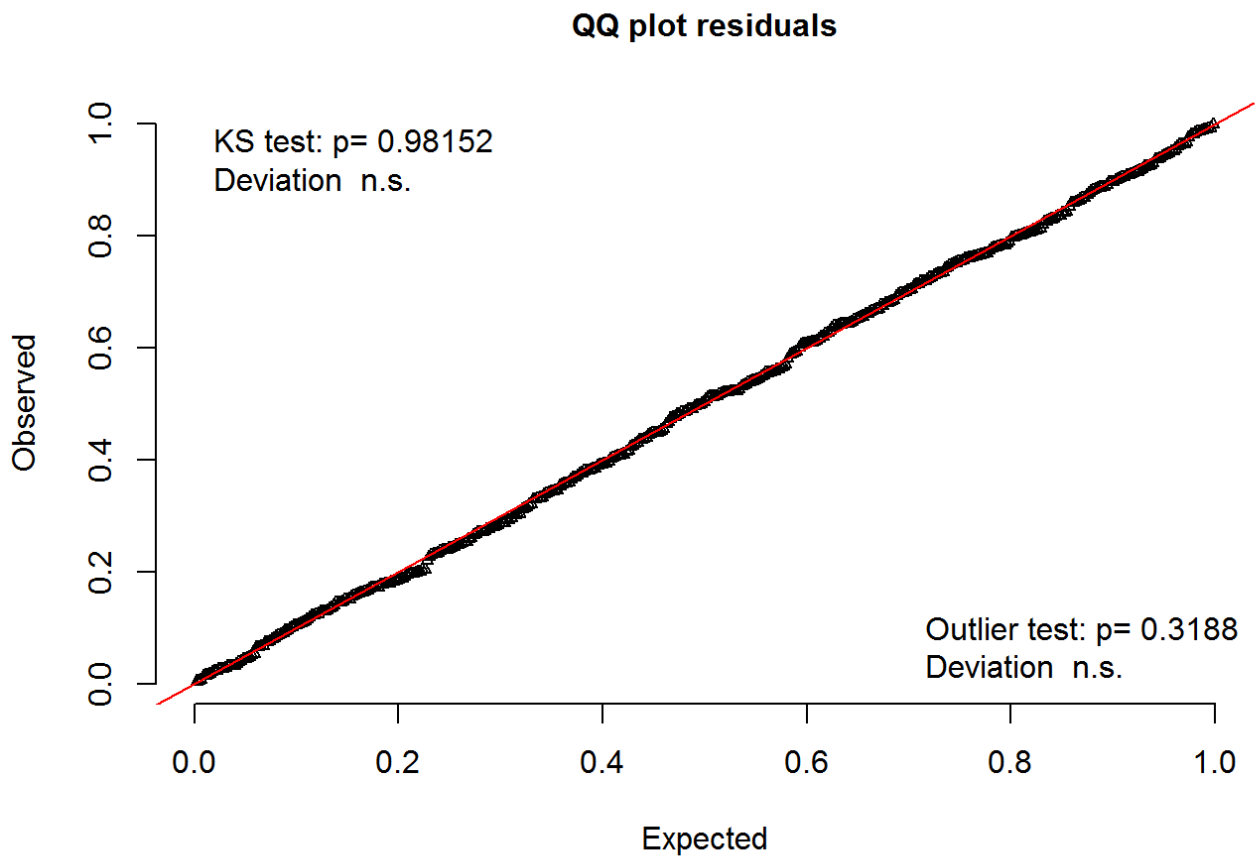
```
testResiduals(rr)###test qqplot
```



```
## $uniformity
##
## One-sample Kolmogorov-Smirnov test
##
## data: simulationOutput$scaledResiduals
## D = 0.021743, p-value = 0.9815
## alternative hypothesis: two-sided
##
##
## $dispersion
##
## DHARMA nonparametric dispersion test via sd of residuals fitted
## vs. simulated
##
## data: simulationOutput
## ratioObsSim = 1.002, p-value = 0.984
## alternative hypothesis: two.sided
##
##
## $outliers
##
## DHARMA outlier test based on exact binomial test
##
## data: simulationOutput
## outLow = 0.0000e+00, outHigh = 1.0000e+00, nobs = 4.6000e+02,
## freqH0 = 3.9841e-03, p-value = 0.3188
## alternative hypothesis: two.sided
```

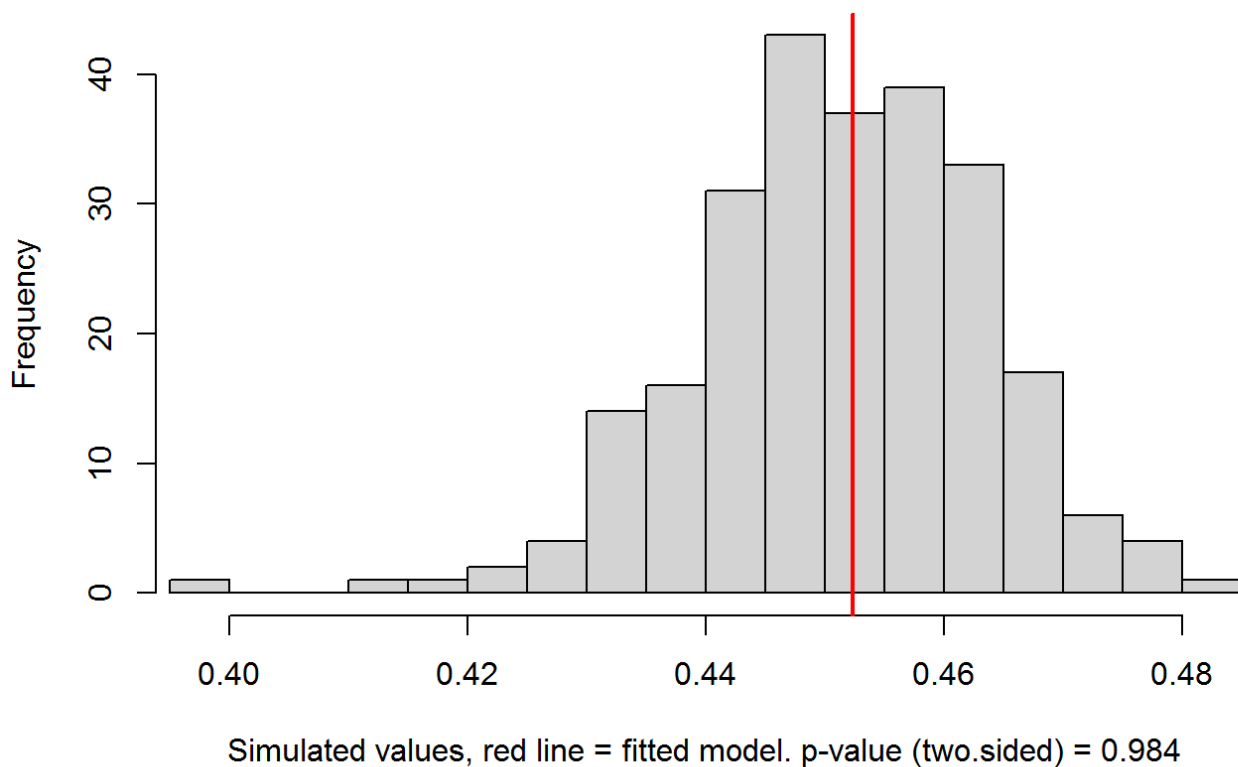
```
## $uniformity
##
## One-sample Kolmogorov-Smirnov test
##
## data: simulationOutput$scaledResiduals
## D = 0.021743, p-value = 0.9815
## alternative hypothesis: two-sided
##
##
## $dispersion
##
## DHARMA nonparametric dispersion test via sd of residuals fitted
## vs. simulated
##
## data: simulationOutput
## ratioObsSim = 1.002, p-value = 0.984
## alternative hypothesis: two.sided
##
##
## $outliers
##
## DHARMA outlier test based on exact binomial test
##
## data: simulationOutput
## outLow = 0.0000e+00, outHigh = 1.0000e+00, nobs = 4.6000e+02,
## freqH0 = 3.9841e-03, p-value = 0.3188
## alternative hypothesis: two.sided
```

```
plotQQunif(rr, testUniformity = T) ### plot
```



```
testDispersion(rr,plot=T) ### test
```

**DHARMA nonparametric dispersion test via sd of
residuals fitted vs. simulated**




```
##
## DHARMA nonparametric dispersion test via sd of residuals fitted
## vs. simulated
##
## data:  simulationOutput
## ratioObsSim = 1.002, p-value = 0.984
## alternative hypothesis: two.sided
```

```
detach(Labsurvbinom)
```

According to the model selection, position does not impact survival of lab tadpoles.

To analyse survival between core and periphery in mesocosms, use the *Mesosurvbinom.csv* to calculate mesocosm survival of tadpoles. 1= survived, 0= died.

```
Mesosurvbinom=read.csv(file.choose(),header=T)
names(Mesosurvbinom)
```

```
## [1] "Treatment" "Position" "Site"      "Clutch"    "Clutchx"   "Tadpole"
## [7] "Survival"
```

```
attach(Mesosurvbinom)

summary(Position) #fixed effect
```

```
##      Core Periphery
##      1650      1800
```

```
summary(Site) #Random effect
```

```
##      Bouille  Challones  Jardinerie  Massais  Saint-Paul  Toutiere
##      600      600      600      450      600      600
```

```
summary(Clutchx) #Nested within site
```

```
##      Bouille1      Bouille2      Bouille3      Bouille4  Challones1  Challones2
##      150      150      150      150      150      150
##      Challones3  Challones4  Jardinerie1  Jardinerie2  Jardinerie3  Jardinerie4
##      150      150      150      150      150      150
##      Massais1      Massais2      Massais4  Saint-Paul1  Saint-Paul2  Saint-Paul3
##      150      150      150      150      150      150
##      Saint-Paul4  Toutiere1  Toutiere2  Toutiere3  Toutiere4
##      150      150      150      150      150
```

```
msurv1<- glmer(Survival ~ Position + (1|Site/Clutchx), data = Mesosurvbinom, famil
y= "binomial")
```

```
## boundary (singular) fit: see ?isSingular
```

```
msurv2<- glmer(Survival ~ 1 + (1|Site/Clutchx), data=Mesosurvbinom, family= "binomial")
```

```
## boundary (singular) fit: see ?isSingular
```

```
AIC(msurv1, msurv2)
```

```
##          df          AIC
## msurv1    4 4252.449
## msurv2    3 4251.979
```

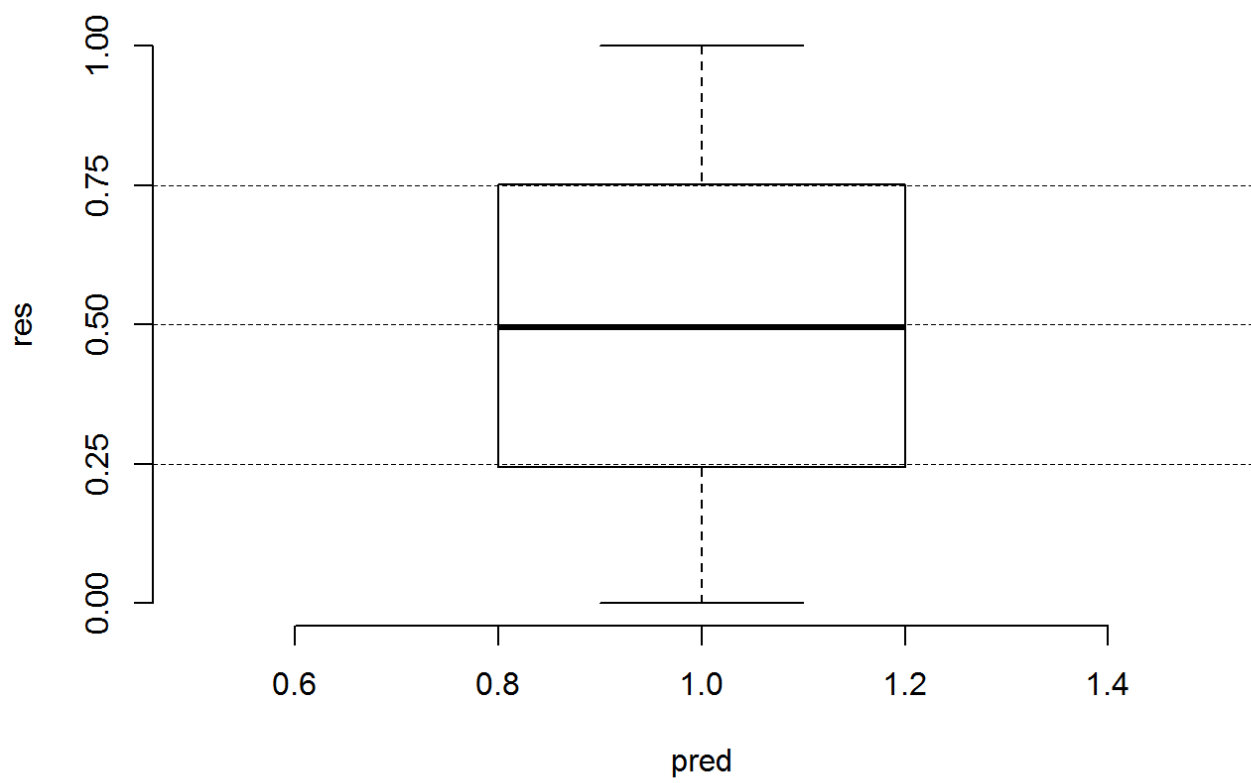
```
AICctab(msurv1, msurv2) #No difference between null model and model with position as fixed effect.
```

```
##          dAICc df
## msurv2 0.0    3
## msurv1 0.5    4
```

```
summary(msurv2)
```

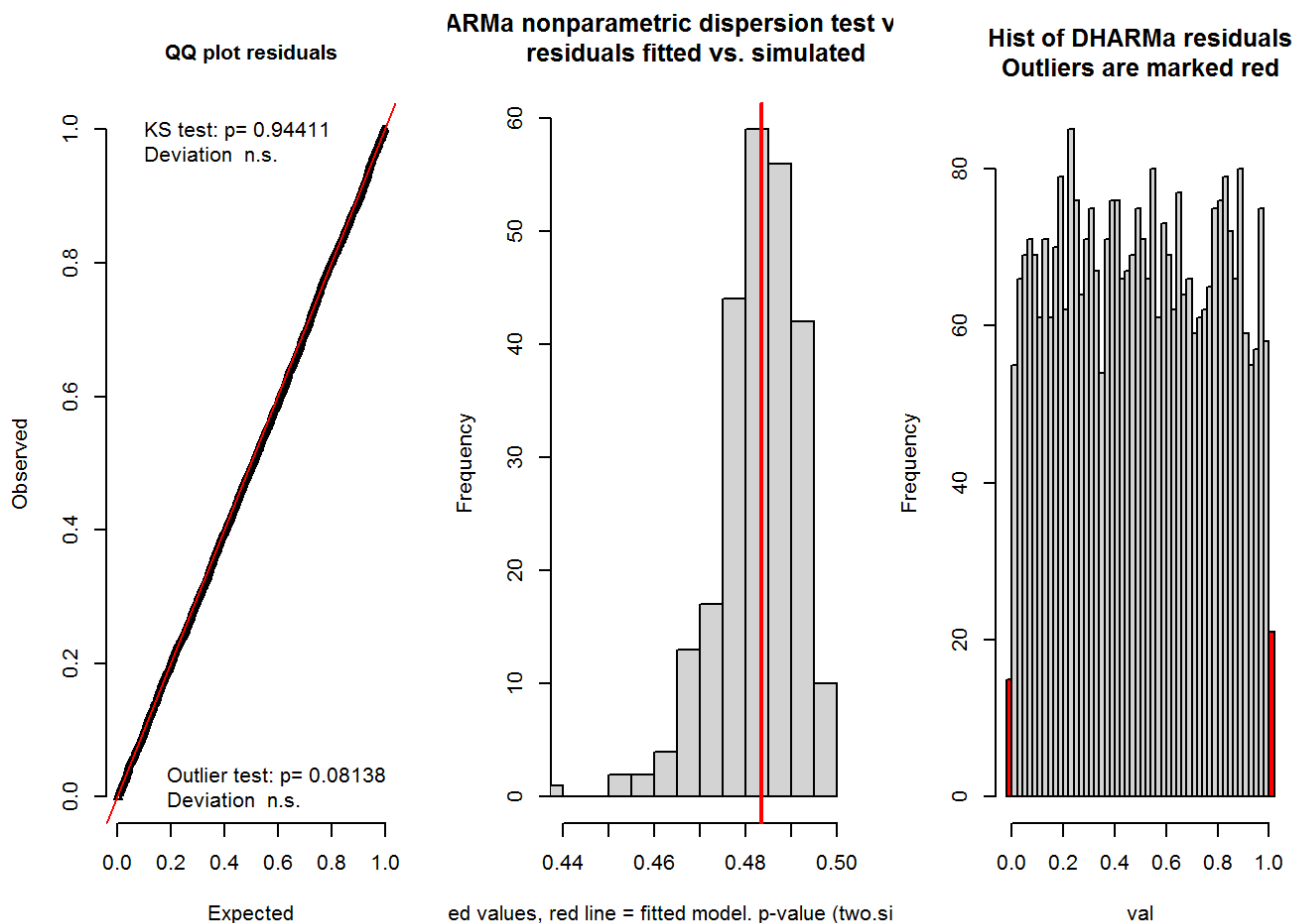
```
## Generalized linear mixed model fit by maximum likelihood (Laplace
##   Approximation) [glmerMod]
##   Family: binomial ( logit )
## Formula: Survival ~ 1 + (1 | Site/Clutchx)
##   Data: Mesosurvbinom
##
##          AIC          BIC   logLik deviance df.resid
##    4252.0    4270.4 -2123.0   4246.0     3447
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -2.5228 -1.0011  0.5241  0.8237  1.6588
##
## Random effects:
##   Groups       Name             Variance Std.Dev.
##   Clutchx:Site (Intercept) 0.5352    0.7316
##   Site          (Intercept) 0.0000    0.0000
## Number of obs: 3450, groups:  Clutchx:Site, 23; Site, 6
##
## Fixed effects:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)   0.5836    0.1572   3.712 0.000206 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## convergence code: 0
## boundary (singular) fit: see ?isSingular
```

```
rr=simulateResiduals(msurv2) ### simulate residuals
plotResiduals(rr) ### qqplot simulated vs observed
```



```
## NULL
```

```
testResiduals(rr)###test qqplot
```

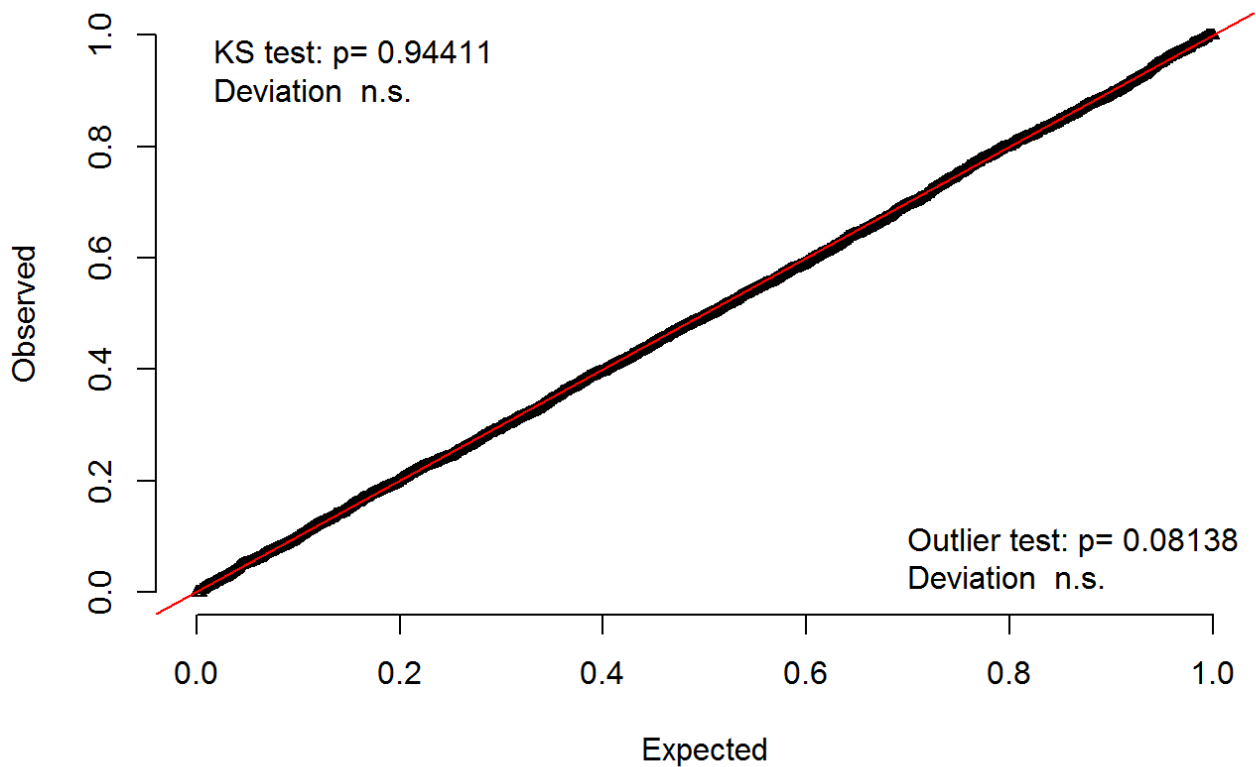


```
## $uniformity
##
## One-sample Kolmogorov-Smirnov test
##
## data: simulationOutput$scaledResiduals
## D = 0.0089703, p-value = 0.9441
## alternative hypothesis: two-sided
##
##
## $dispersion
##
## DHARMA nonparametric dispersion test via sd of residuals fitted
## vs. simulated
##
## data: simulationOutput
## ratioObsSim = 1.0017, p-value = 0.992
## alternative hypothesis: two.sided
##
##
## $outliers
##
## DHARMA outlier test based on exact binomial test
##
## data: simulationOutput
## outLow = 1.5000e+01, outHigh = 2.1000e+01, nobs = 3.4500e+03,
## freqH0 = 3.9841e-03, p-value = 0.08138
## alternative hypothesis: two.sided
```

```
## $uniformity
##
## One-sample Kolmogorov-Smirnov test
##
## data: simulationOutput$scaledResiduals
## D = 0.0089703, p-value = 0.9441
## alternative hypothesis: two-sided
##
##
## $dispersion
##
## DHARMA nonparametric dispersion test via sd of residuals fitted
## vs. simulated
##
## data: simulationOutput
## ratioObsSim = 1.0017, p-value = 0.992
## alternative hypothesis: two.sided
##
##
## $outliers
##
## DHARMA outlier test based on exact binomial test
##
## data: simulationOutput
## outLow = 1.5000e+01, outHigh = 2.1000e+01, nobs = 3.4500e+03,
## freqH0 = 3.9841e-03, p-value = 0.08138
## alternative hypothesis: two.sided
```

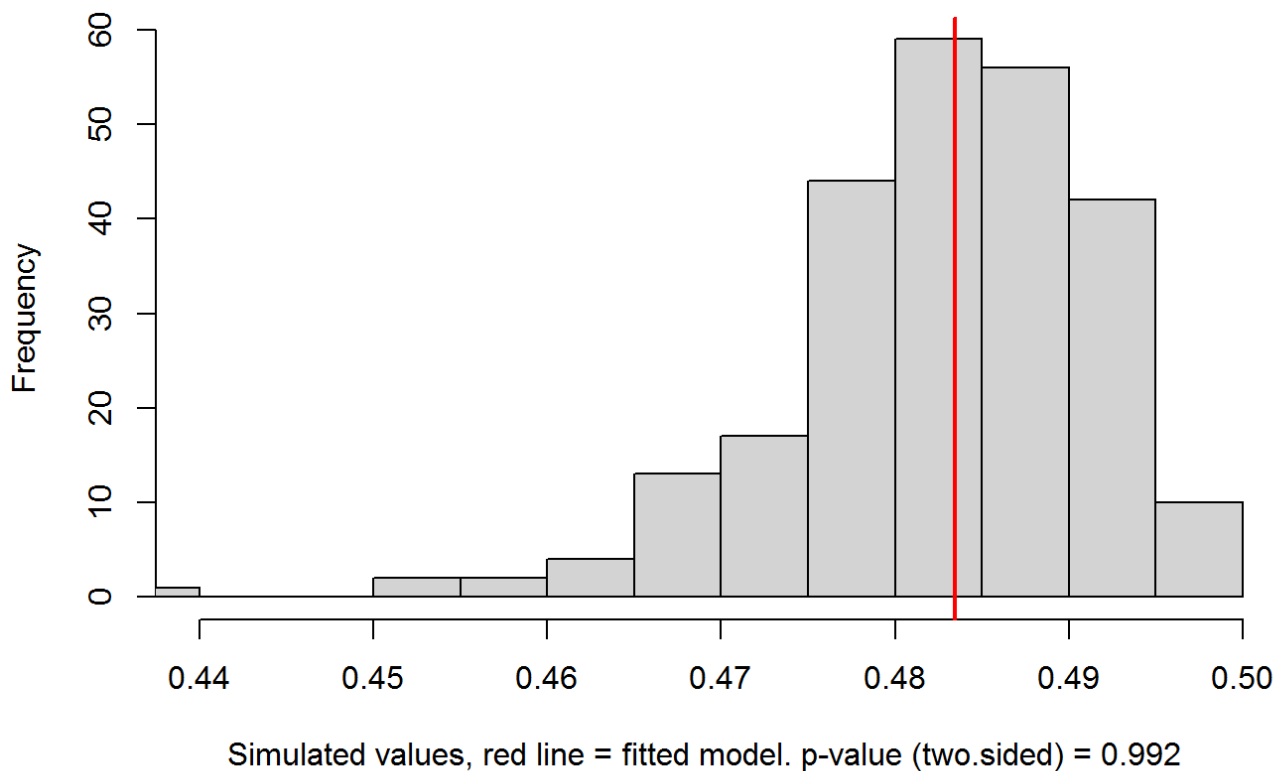
```
plotQQunif(rr, testUniformity = T) ### plot
```

QQ plot residuals



```
testDispersion(rr,plot=T) ### test
```

DHARMA nonparametric dispersion test via sd of residuals fitted vs. simulated



```
##
## DHARMA nonparametric dispersion test via sd of residuals fitted
## vs. simulated
##
## data:  simulationOutput
## ratioObsSim = 1.0017, p-value = 0.992
## alternative hypothesis: two.sided
```

```
detach(Mesosurvbinom)
```

According to the model selection, position does not impact survival of mesocosm tadpoles.

We have also recorded the date of mortality in the lab. Survival analysis was used with clutch nested within site as random effect to determine whether position had an effect on mortality occurrence between core and periphery.

Use the *Labsurv.csv* file.

```
library(coxme)
```

```
## Loading required package: survival
```

```
## Loading required package: bdsmatrix
```

```
##
## Attaching package: 'bdsmatrix'
```

```
## The following object is masked from 'package:base':
##
##      backsolve
```

```
library(survival)
```

```
Labsurv=read.csv(file.choose(),header=T)
names(Labsurv)
```

```
## [1] "Site"      "Clutch"    "Clutchx"   "Position"  "Week"      "Died"
```

```
attach(Labsurv)
```

```
summary(Position) #fixed effect
```

```
##      Core Periphery
##      240      240
```

```
summary(Week) #fixed effect
```

```
##      Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##    5.000   7.000   9.500   8.498  10.000  10.000
```

```
summary(Site) #Random effect
```

```
##  B  C  J  M  S  T
## 80 80 80 80 80 80
```

```
summary(Clutchx) #Nested within site
```

```
## B1 B2 B3 B4 C1 C2 C3 C4 J1 J2 J3 J4 M1 M2 M3 M4 S1 S2 S3 S4 T1 T2 T3 T4
## 20 20 20 20 20 20 20 20 20 20 20 20 20 20 20 20 20 20 20 20 20 20 20 20
```

```
summary(Died)
```

```
##      Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##    0.0000  0.0000  1.0000  0.5333  1.0000  1.0000
```

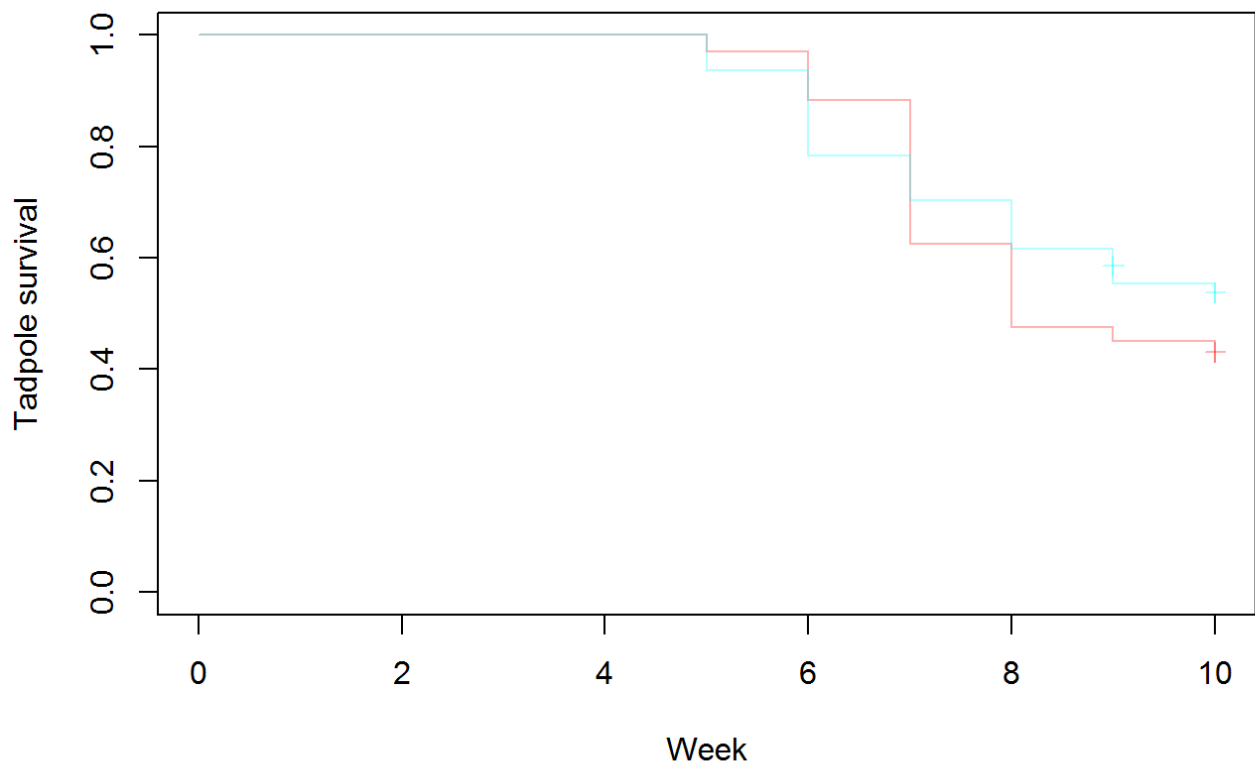
```
mysurv<- Surv(Week, Died)
```

```
#Kaplan-Meier non-parametric analysis by group (fixed effect) without random effect to create plot
```

```
kmsurvival <- survfit((mysurv) ~ Position)
summary(kmsurvival)
```

```
## Call: survfit(formula = (mysurv) ~ Position)
##
##               Position=Core
##   time n.risk n.event survival std.err lower 95% CI upper 95% CI
##    5    240      7    0.971  0.0109    0.950    0.992
##    6    233     21    0.883  0.0207    0.844    0.925
##    7    212     62    0.625  0.0312    0.567    0.689
##    8    150     36    0.475  0.0322    0.416    0.543
##    9    114      6    0.450  0.0321    0.391    0.518
##   10    108      9    0.412  0.0318    0.355    0.480
##
##               Position=Periphery
##   time n.risk n.event survival std.err lower 95% CI upper 95% CI
##    5    240     15    0.938  0.0156    0.907    0.969
##    6    225     37    0.783  0.0266    0.733    0.837
##    7    188     19    0.704  0.0295    0.649    0.764
##    8    169     21    0.617  0.0314    0.558    0.681
##    9    148     15    0.554  0.0321    0.495    0.621
##   10    132      8    0.521  0.0323    0.461    0.588
```

```
plot(kmsurvival, xlab = "Week", ylab = "Tadpole survival", col= rainbow(2, alpha = 0.3),mark = 3)
```

```
m1<-coxme((mysurv) ~ Position +(1|Site/Clutch), data=Labsurv)
m2<- coxme((mysurv) ~ 1 +(1|Site/Clutch), data=Labsurv)

AIC(m1, m2)#No difference between null model and model with position as fixed effect.
```

```
##          df      AIC
## m1 18.27817 2935.654
## m2 18.16207 2935.438
```

```
summary(m2)
```

```
## Cox mixed-effects model fit by maximum likelihood
##   Data: Labsurv
##   events, n = 256, 480
##   Iterations= 8 38
##
##               NULL Integrated      Fitted
## Log-likelihood -1495.516   -1476.563 -1449.557
##
##               Chisq      df          p    AIC    BIC
## Integrated loglik 37.91   2.00 5.8718e-09 33.91 26.82
## Penalized loglik 91.92 18.16 7.4888e-12 55.59 -8.79
##
## Model: (mysurv) ~ 1 + (1 | Site/Clutch)
##
## Random effects
## Group      Variable      Std Dev      Variance
## Site/Clutch (Intercept) 0.6033877949 0.3640768310
## Site      (Intercept) 0.0199839423 0.0003993579
```

```
#p-value= 0.48
```

```
detach(Labsurv)
```

According to theory, models with AIC within two points of each other are basically equal. Thus, both model can be chosen as suitable. No significant effect of position on weekly mortality was observed for tadpoles raised in the lab. We retained the model that displayed no effect of position ($p = 0.48$) random effects have a minor effect on mortality.

2. What was the effect of range expansion on the development of tadpoles?

+We conducted a survival analysis to determine when core and periphery individuals transition between stage categories (as defined in Segerdell et al. (2008)) Week to transition between premetamorphosis to **prometamorphosis**, prometamorphosis to **climax**, climax to **juvenile**, limb bud development (**stage 48**), and distinguished femur development (**stage 54**) was considered response variables. Position (core/periphery) was used as fixed effect. Clutch nested within site was considered as random effects in the coxme package.

+2.1. Transition between premetamorphosis and prometamorphosis.

Use the *Prometamorphosis.csv* file.

```
Promet=read.csv(file.choose(),header=T)
names(Promet)
```

```
## [1] "num"           "Position"      "Site"
## [4] "Clutch"        "Clutchx"       "Tadpole"
## [7] "Week"          "Weeknr"        "tadpoleID"
## [10] "Mesocosm"      "Date"          "Stage"
## [13] "prometamorphosis"
```

```
attach(Promet)
```

```
summary(Position) #fixed effect
```

```
##      Core Periphery
```

```
##      449      484
```

```
summary(Weeknr) #fixed effect
```

```
##      Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
```

```
##      1.000    3.000    5.000    4.842    7.000   10.000
```

```
summary(Site) #Random effect
```

```
##      B      C      J      M      S      T
```

```
## 156 164 161 121 172 159
```

```
summary(Clutchx) #Nested within site
```

```
## B1 B2 B3 B4 C1 C2 C3 C4 J1 J2 J3 J4 M1 M2 M4 S1 S2 S3 S4 T1 T2 T3 T4
```

```
## 50 40 33 33 35 46 43 40 46 36 39 40 42 41 38 42 48 37 45 41 38 43 37
```

```
summary(prometamorphosis)
```

```
##      Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
```

```
## 0.0000 0.0000 0.0000 0.2015 0.0000 1.0000
```

```
mysurv1<- Surv(Weeknr, prometamorphosis)
```

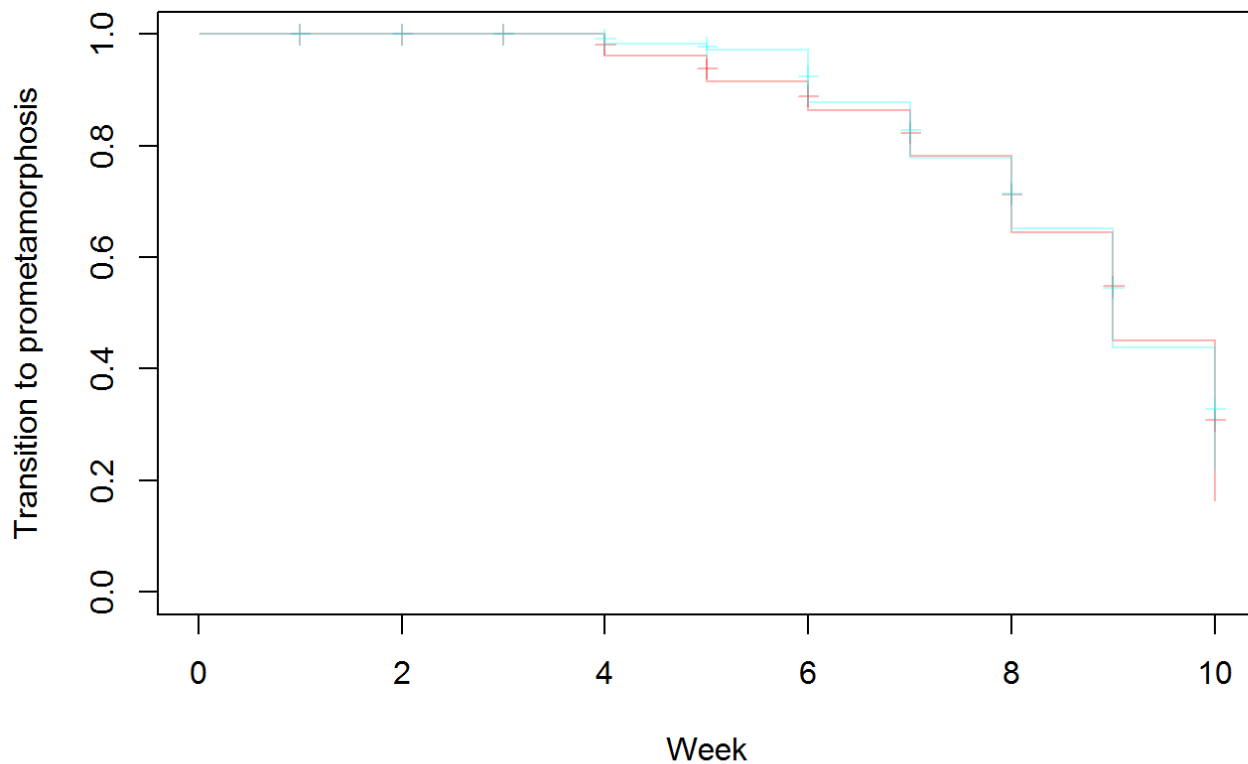
```
#Kaplan-Meier non-parametric analysis by group (fixed effect) without random effect  
to create plot
```

```
kmsurvival1 <- survfit((mysurv1) ~ Position)
```

```
summary(kmsurvival1)
```

```
## Call: survfit(formula = (mysurv1) ~ Position)
##
##               Position=Core
##   time n.risk n.event survival std.err lower 95% CI upper 95% CI
##     4     284      11   0.961  0.0114    0.939    0.984
##     5     229      11   0.915  0.0174    0.882    0.950
##     6     178      10   0.864  0.0228    0.820    0.910
##     7     137      13   0.782  0.0299    0.725    0.843
##     8      97      17   0.645  0.0390    0.573    0.726
##     9      60      18   0.451  0.0469    0.368    0.553
##    10      33      21   0.164  0.0415    0.100    0.269
##
##               Position=Periphery
##   time n.risk n.event survival std.err lower 95% CI upper 95% CI
##     4     306       5   0.984 0.00725    0.970    0.998
##     5     246       3   0.972 0.00993    0.952    0.991
##     6     187      18   0.878 0.02280    0.835    0.924
##     7     132      15   0.778 0.03157    0.719    0.843
##     8      92      15   0.651 0.03996    0.578    0.735
##     9      55      18   0.438 0.04921    0.352    0.546
##    10      26      13   0.219 0.04952    0.141    0.341
```

```
plot(kmsurvival1, xlab = "Week", ylab = "Transition to prometamorphosis", col= rain
bow(2, alpha = 0.3),mark = 3)
```



```
prom1<-coxme((mysurv1) ~ Position +(1|Site/Clutch), data=Promet)
prom2<- coxme((mysurv1) ~ 1 +(1|Site/Clutch), data=Promet)

AIC(prom1, prom2)#No difference between null model and model with position as fixed
effect.
```

```
##           df      AIC
## prom1 9.369901 1921.826
## prom2 8.965443 1920.878
```

```
summary(prom2)
```

```
## Cox mixed-effects model fit by maximum likelihood
##   Data: Promet
##   events, n = 188, 933
##   Iterations= 16 67
##
##              NULL Integrated      Fitted
## Log-likelihood -963.4512  -961.7707 -951.4738
##
##              Chisq  df      p    AIC    BIC
## Integrated loglik  3.36 2.00 0.1862900 -0.64 -7.11
## Penalized loglik 23.95 8.97 0.0042839  6.02 -22.99
##
## Model:  (mysurv1) ~ 1 + (1 | Site/Clutch)
##
## Random effects
## Group      Variable      Std Dev      Variance
## Site/Clutch (Intercept) 2.935888e-01 8.619440e-02
## Site        (Intercept) 8.669943e-03 7.516791e-05
```

```
#p-value= 0.48
```

```
detach(Promet)
```

According to theory, models with AIC within two points of each other are basically equal. Thus, both model can be chosen as suitable. No significant effect of position on transition between prometamorphosis to prometamorphosis was observed for tadpoles raised in the mesocosms. We retained the model that displayed no effect of position ($p = 0.56$) random effects have a minor effect.

+2.2. Transition between prometamorphosis to climax.

Use the *Climax.csv* file.

```
Climax=read.csv(file.choose(),header=T)
names(Climax)
```

```
## [1] "num"      "Position" "Site"      "Clutch"    "Clutchx"
## [6] "Tadpole"  "Week"     "Weeknr"    "tadpoleID" "Mesocosm"
## [11] "Date"     "Stage"    "climax"
```

```
attach(Climax)
```

```
summary(Position) #fixed effect
```

```
##      Core Periphery
```

```
##      516      573
```

```
summary(Weeknr) #fixed effect
```

```
##      Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
```

```
##      1.000    3.000    5.000    5.314    8.000   10.000
```

```
summary(Site) #Random effect
```

```
##      B      C      J      M      S      T
```

```
## 189 194 190 141 186 189
```

```
summary(Clutchx) #Nested within site
```

```
## B1 B2 B3 B4 C1 C2 C3 C4 J1 J2 J3 J4 M1 M2 M4 S1 S2 S3 S4 T1 T2 T3 T4
```

```
## 50 48 46 45 46 50 50 48 50 46 46 48 46 47 48 46 49 44 47 49 46 48 46
```

```
summary(climax)
```

```
##      Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
```

```
## 0.0000 0.0000 0.0000 0.1433 0.0000 1.0000
```

```
mysurv2<- Surv(Weeknr, climax)
```

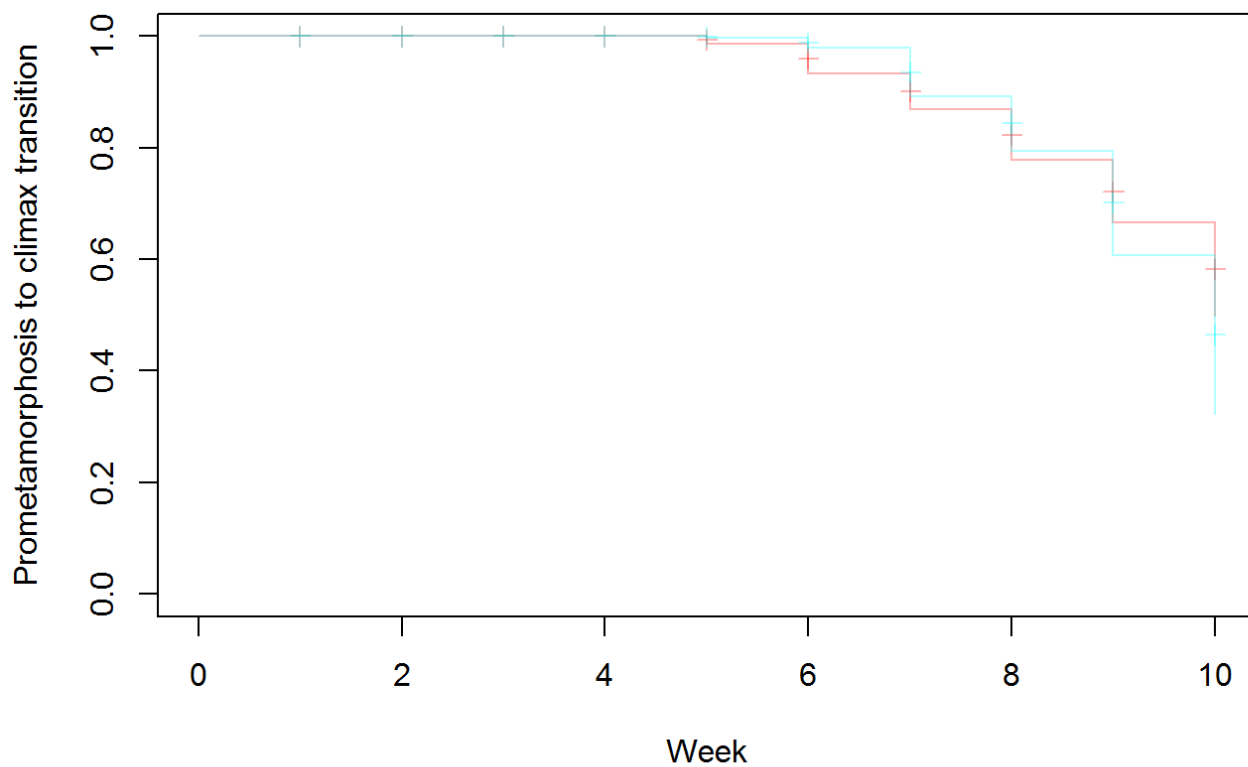
```
#Kaplan-Meier non-parametric analysis by group (fixed effect) without random effect  
to create plot
```

```
kmsurvival2 <- survfit((mysurv2) ~ Position)
```

```
summary(kmsurvival2)
```

```
## Call: survfit(formula = (mysurv2) ~ Position)
##
##               Position=Core
##   time n.risk n.event survival std.err lower 95% CI upper 95% CI
##     5     296      4   0.986 0.00671    0.973    1.000
##     6     241     13   0.933 0.01570    0.903    0.965
##     7     187     13   0.868 0.02268    0.825    0.914
##     8     134     14   0.778 0.03065    0.720    0.840
##     9      83     12   0.665 0.03985    0.592    0.748
##    10      44     11   0.499 0.05272    0.406    0.614
##
##               Position=Periphery
##   time n.risk n.event survival std.err lower 95% CI upper 95% CI
##     5     335      1   0.997 0.00298    0.991    1.000
##     6     275      5   0.979 0.00855    0.962    0.996
##     7     215     19   0.892 0.02049    0.853    0.933
##     8     156     17   0.795 0.02879    0.741    0.854
##     9     102     24   0.608 0.04000    0.534    0.692
##    10      49     23   0.323 0.04827    0.241    0.433
```

```
plot(kmsurvival2, xlab = "Week", ylab = "Prometamorphosis to climax transition", co
l= rainbow(2, alpha = 0.3),mark = 3)
```



```

clim1<-coxme((mysurv2) ~ Position  +(1|Site/Clutchx), data=Climax)
clim2<- coxme((mysurv2) ~ 1  +(1|Site/Clutchx), data=Climax)

AIC(clim1, clim2)#No difference between null model and model with position as fixed effect.

```

```

##           df      AIC
## clim1 14.25337 1659.001
## clim2 13.99748 1658.998

```

```
summary(clim2)
```

```

## Cox mixed-effects model fit by maximum likelihood
##   Data: Climax
##   events, n = 156, 1089
##   Iterations= 9 40
##
##              NULL Integrated      Fitted
## Log-likelihood -840.5432  -833.6201 -815.5016
##
##              Chisq df          p    AIC    BIC
## Integrated loglik 13.85  2 9.8481e-04  9.85   3.75
## Penalized loglik 50.08 14 5.9030e-06 22.09 -20.60
##
## Model:  (mysurv2) ~ 1 + (1 | Site/Clutchx)
##
## Random effects
## Group      Variable      Std Dev      Variance
## Site/Clutchx (Intercept) 0.5220841697 0.2725718803
## Site        (Intercept) 0.0198983478 0.0003959442

```

```
#p-value= 0.36
```

```
detach(Climax)
```

According to theory, models with AIC within two points of each other are basically equal. Thus, both model can be chosen as suitable. No significant effect of position on transition between prometamorphosis to climax was observed for tadpoles raised in the mesocosms. We retained the model that displayed no effect of position ($p= 0.36$) random effects have a minor effect.

+2.3. Transition between climax to metamorphs.

Use the *Metamorph.csv* file.

```

Metamorph=read.csv(file.choose(),header=T)
names(Metamorph)

```

```

## [1] "num"      "Position" "Site"      "Clutch"    "Clutchx"
## [6] "Tadpole"  "Week"     "Weeknr"    "tadpoleID" "Mesocosm"
## [11] "Date"     "Stage"    "metamorphs"

```



```
attach(Metamorph)
```

```
summary(Position) #fixed effect
```

```
##      Core Periphery
```

```
##      550      598
```

```
summary(Weeknr) #fixed effect
```

```
##      Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
```

```
##      1.000    3.000    6.000    5.506    8.000   10.000
```

```
summary(Site) #Random effect
```

```
##      B      C      J      M      S      T
```

```
##    200  200  200  150  200  198
```

```
summary(Clutchx) #Nested within site
```

```
## B1 B2 B3 B4 C1 C2 C3 C4 J1 J2 J3 J4 M1 M2 M4 S1 S2 S3 S4 T1 T2 T3 T4
```

```
## 50 50 50 50 50 50 50 50 50 50 50 50 50 50 50 50 50 50 50 49 50 49
```

```
summary(metamorphs)
```

```
##      Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
```

```
## 0.00000 0.00000 0.00000 0.05052 0.00000 1.00000
```

```
mysurv3<- Surv(Weeknr, metamorphs)
```

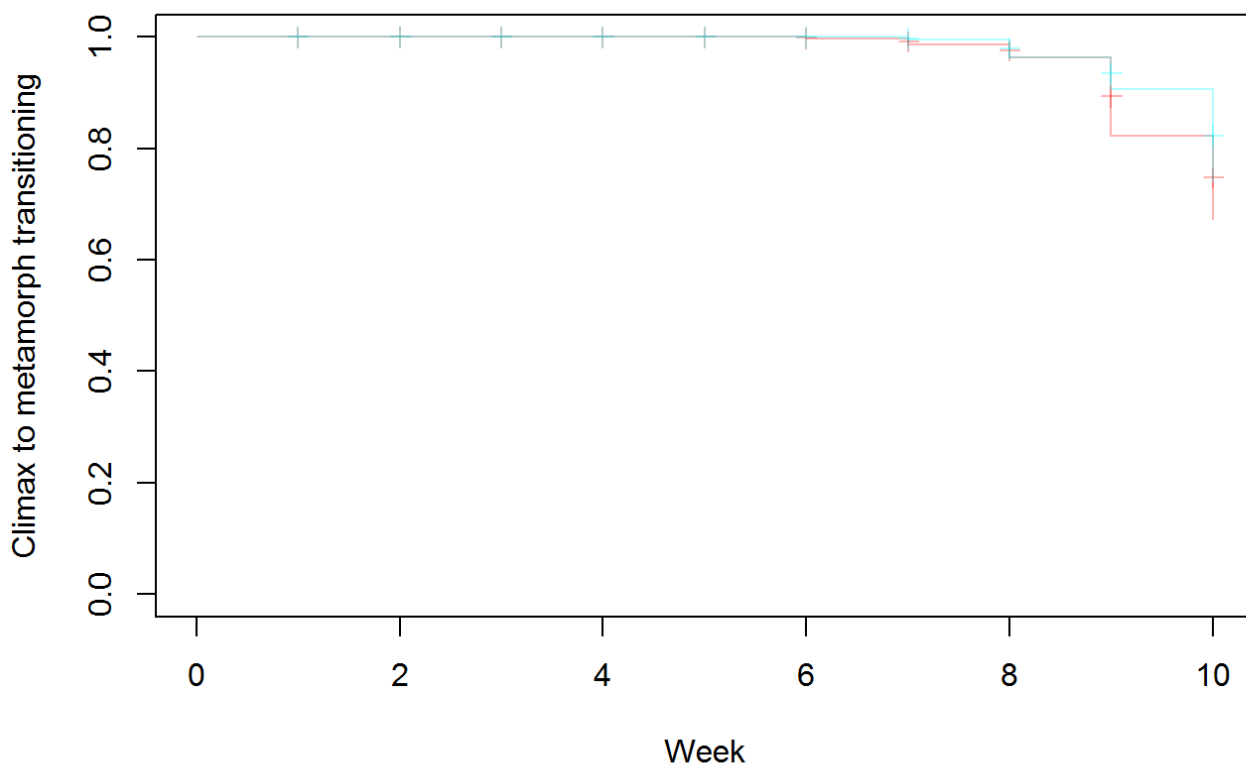
```
#Kaplan-Meier non-parametric analysis by group (fixed effect) without random effect  
to create plot
```

```
kmsurvival3 <- survfit((mysurv3) ~ Position)
```

```
summary(kmsurvival3)
```

```
## Call: survfit(formula = (mysurv3) ~ Position)
##
##               Position=Core
##   time n.risk n.event survival std.err lower 95% CI upper 95% CI
##      6    275      1   0.996 0.00363    0.989    1.000
##      7    220      2   0.987 0.00732    0.973    1.000
##      8    165      4   0.963 0.01381    0.937    0.991
##      9    110     16   0.823 0.03447    0.758    0.894
##     10     55     10   0.674 0.05127    0.580    0.782
##
##               Position=Periphery
##   time n.risk n.event survival std.err lower 95% CI upper 95% CI
##      7    240      1   0.996 0.00416    0.988    1.000
##      8    180      6   0.963 0.01392    0.936    0.990
##      9    120      7   0.906 0.02441    0.860    0.956
##     10     60     11   0.740 0.04948    0.649    0.844
```

```
plot(kmsurvival3, xlab = "Week", ylab = "Climax to metamorph transitioning", col= rainbow(2, alpha = 0.3), mark = 3)
```



```
met1<-coxme((mysurv3) ~ Position +(1|Site/Clutchx), data=Metamorph)
met2<- coxme((mysurv3) ~ 1 +(1|Site/Clutchx), data=Metamorph)

AIC(met1, met2)#No difference between null model and model with position as fixed effect.
```

```
##           df           AIC
## met1 3.647093 608.6190
## met2 4.293098 608.3543
```

```
summary(met2)
```

```
## Cox mixed-effects model fit by maximum likelihood
##   Data: Metamorph
##   events, n = 58, 1148
##   Iterations= 15 63
##               NULL Integrated   Fitted
## Log-likelihood -304.6815   -304.413 -299.884
##
##               Chisq    df         p    AIC    BIC
## Integrated loglik  0.54 2.00 0.764520 -3.46 -7.58
## Penalized loglik  9.60 4.29 0.057908  1.01 -7.84
##
## Model: (mysurv3) ~ 1 + (1 | Site/Clutchx)
##
## Random effects
## Group          Variable      Std Dev      Variance
## Site/Clutchx (Intercept) 0.3102212516 0.0962372249
## Site           (Intercept) 0.0128668022 0.0001655546
```

```
#p-value= 0.18
```

```
detach(Metamorph)
```

According to theory, models with AIC within two points of each other are basically equal. Thus, both model can be chosen as suitable. No significant effect of position on transition between climax to metamorphs was observed for tadpoles raised in the mesocosms. We retained the model that displayed no effect of position ($p=0.18$) random effects have a minor effect.

+2.4. Development of limb bud (stage 48)

Use the *Stage48.csv* file.

```
Stage48=read.csv(file.choose(),header=T)
names(Stage48)
```

```
## [1] "num"      "Position" "Site"      "Clutch"    "Clutchx"
## [6] "Tadpole"  "Week"     "Weeknr"    "tadpoleID" "Mesocosm"
## [11] "Date"     "Stage"    "stage48"
```

```
attach(Stage48)
```

```
summary(Position) #fixed effect
```

```
##      Core Periphery
##      150         174
```

```
summary(Weeknr) #fixed effect
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    1.000   1.000    2.000   2.448   3.000   10.000
```

```
summary(Site) #Random effect
```

```
##  B  C  J  M  S  T
## 62 68 54 34 54 52
```

```
summary(Clutchx) #Nested within site
```

```
## B1 B2 B3 B4 C1 C2 C3 C4 J1 J2 J3 J4 M1 M2 M4 S1 S2 S3 S4 T1 T2 T3 T4
## 22 19 13  8 18 22 14 14 17  7 14 16 13 16  5  9 20 14 11 20  6 12 14
```

```
summary(stage48)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    0.0000   0.0000   0.0000   0.4259   1.0000   1.0000
```

```
mysurv4<- Surv(Weeknr, stage48)
```

```
#Kaplan-Meier non-parametric analysis by group (fixed effect) without random effect to create plot
```

```
kmsurvival4 <- survfit((mysurv4) ~ Position)
```

```
summary(kmsurvival4)
```

```
## Call: survfit(formula = (mysurv4) ~ Position)
```

```
##
```

```
##           Position=Core
```

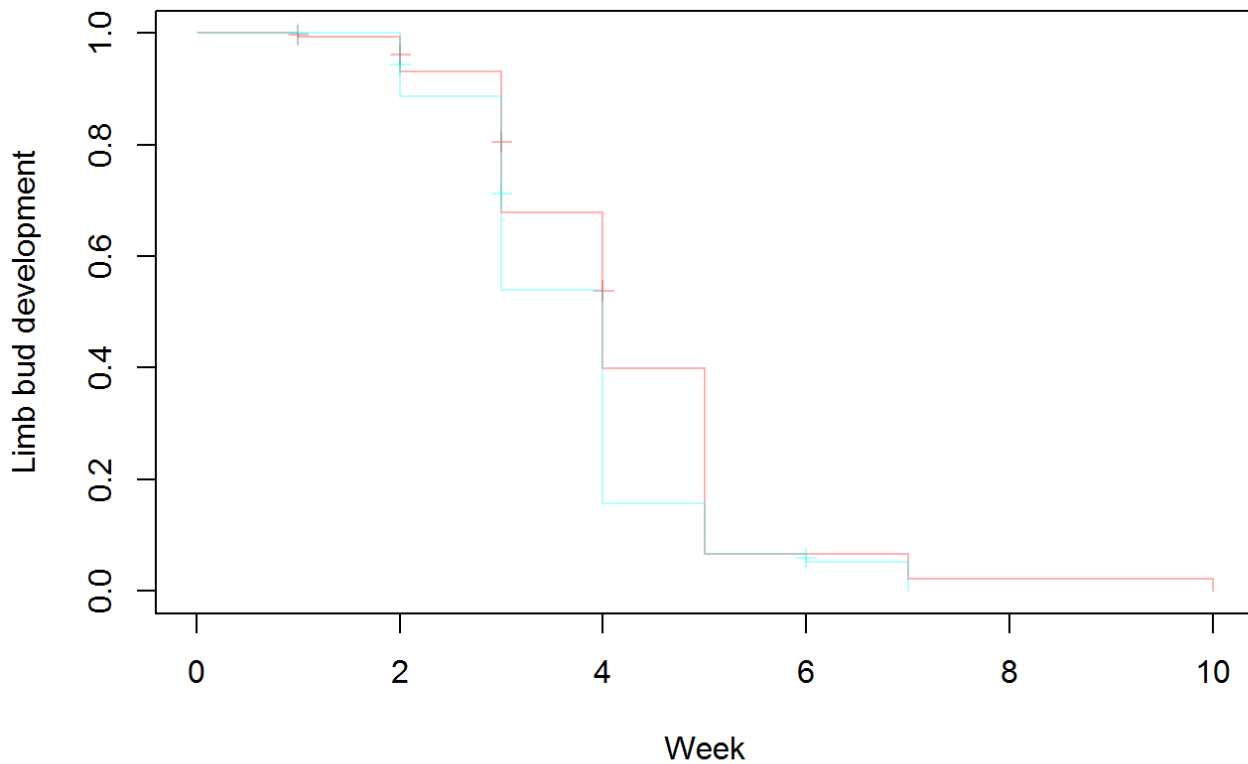
```
##  time n.risk n.event survival std.err lower 95% CI upper 95% CI
##    1    150      1   0.9933 0.00664   0.98040   1.000
##    2     95      6   0.9306 0.02556   0.88182   0.982
##    3     59     16   0.6782 0.05699   0.57524   0.800
##    4     34     14   0.3990 0.06634   0.28800   0.553
##    5     18     15   0.0665 0.03675   0.02251   0.196
##    7      3      2   0.0222 0.02185   0.00321   0.153
##   10      1      1   0.0000      NaN      NA      NA
```

```
##
```

```
##           Position=Periphery
```

```
##  time n.risk n.event survival std.err lower 95% CI upper 95% CI
##    2    115     13   0.8870 0.0295   0.8309   0.947
##    3     79     31   0.5389 0.0519   0.4462   0.651
##    4     41     29   0.1577 0.0412   0.0945   0.263
##    5     12      7   0.0657 0.0283   0.0283   0.153
##    6      5      1   0.0526 0.0255   0.0203   0.136
##    7      2      2   0.0000      NaN      NA      NA
```

```
plot(kmsurvival4, xlab = "Week", ylab = "Limb bud development", col= rainbow(2, alp
ha = 0.3),mark = 3)
```



```
limb1<-coxme((mysurv4) ~ Position +(1|Site/Clutchx), data=Stage48)
limb2<- coxme((mysurv4) ~ 1 +(1|Site/Clutchx), data=Stage48)

AIC(limb1, limb2)#No difference between null model and model with position as fixed
effect.
```

```
##           df      AIC
## limb1  9.522021 1127.868
## limb2 10.809163 1127.889
```

```
summary(limb1)
```

```
## Cox mixed-effects model fit by maximum likelihood
##   Data: Stage48
##   events, n = 138, 324
##   Iterations= 13 56
##
##               NULL Integrated      Fitted
## Log-likelihood -570.4702  -565.2845 -554.4117
##
##               Chisq    df          p    AIC    BIC
## Integrated loglik 10.37  3.00 0.01565900  4.37  -4.41
## Penalized loglik 32.12  9.52 0.00027593 13.07 -14.80
##
## Model: (mysurv4) ~ Position + (1 | Site/Clutchx)
## Fixed coefficients
##               coef exp(coef) se(coef)      z      p
## PositionPeriphery 0.4711086  1.601769 0.2559439 1.84 0.066
##
## Random effects
## Group      Variable      Std Dev      Variance
## Site/Clutchx (Intercept) 3.881630e-01 1.506705e-01
## Site        (Intercept) 9.380392e-03 8.799175e-05
```

```
#p-value= 0.066
detach(Stage48)
```

According to theory, models with AIC within two points of each other are basically equal. Thus, both model can be chosen as suitable. No significant effect of position on limb bud development was observed for tadpoles raised in the mesocosms. We retained the model that displayed no effect of position ($p=0.066$) random effects have a minor effect.

+2.5. Distinguished femur present (stage54)

Use the *Stage54.csv* file. Femurs are only distinguishly present at stage 54

```
Stage54=read.csv(file.choose(),header=T)
names(Stage54)
```

```
## [1] "num"      "Position" "Site"      "Clutch"    "Clutchx"
## [6] "Tadpole"  "Week"     "Weeknr"    "tadpoleID" "Mesocosm"
## [11] "Date"     "Stage"    "stage54"
```

```
attach(Stage54)

summary(Position) #fixed effect
```

```
##      Core Periphery
##      348          397
```

```
summary(Weeknr) #fixed effect
```

```
##      Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##    1.000   2.000   4.000   4.161   6.000  10.000
```

```
summary(Site) #Random effect
```

```
##      B      C      J      M      S      T
## 116 136 130   97 135 131
```

```
summary(Clutchx) #Nested within site
```

```
## B1 B2 B3 B4 C1 C2 C3 C4 J1 J2 J3 J4 M1 M2 M4 S1 S2 S3 S4 T1 T2 T3 T4
## 40 32 23 21 30 38 37 31 41 24 31 34 32 35 30 35 38 26 36 39 27 34 31
```

```
summary(stage54)
```

```
##      Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
## 0.00000 0.00000 0.00000 0.05906 0.00000 1.00000
```

```
mysurv5<- Surv(Weeknr, stage54)
```

```
#Kaplan-Meier non-parametric analysis by group (fixed effect) without random effect to create plot
```

```
kmsurvival5 <- survfit((mysurv5) ~ Position)
```

```
summary(kmsurvival5)
```

```
## Call: survfit(formula = (mysurv5) ~ Position)
```

```
##
```

```
##              Position=Core
```

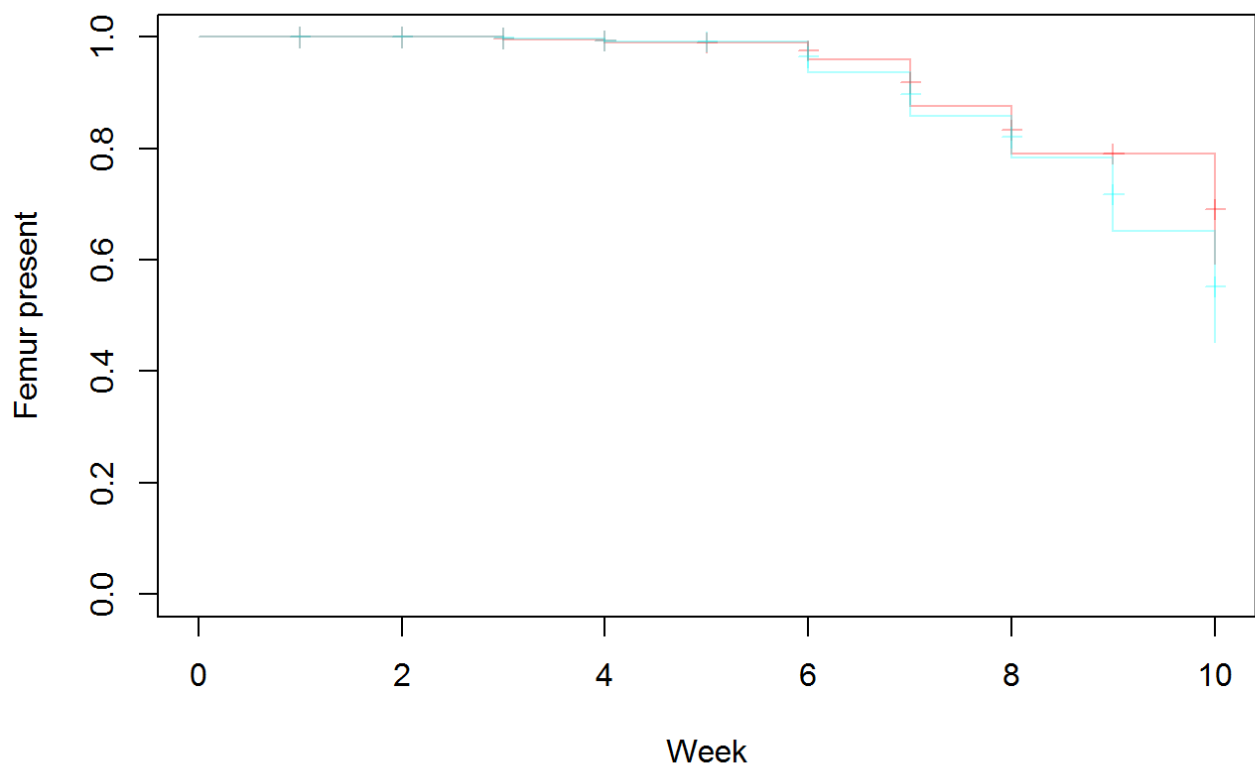
##	time	n.risk	n.event	survival	std.err	lower 95% CI	upper 95% CI
##	3	238	1	0.996	0.00419	0.988	1.000
##	4	183	1	0.990	0.00684	0.977	1.000
##	6	99	3	0.960	0.01831	0.925	0.997
##	7	68	6	0.876	0.03701	0.806	0.951
##	8	41	4	0.790	0.05255	0.694	0.900
##	10	12	3	0.593	0.10635	0.417	0.842

```
##
```

```
##              Position=Periphery
```

##	time	n.risk	n.event	survival	std.err	lower 95% CI	upper 95% CI
##	3	278	1	0.996	0.00359	0.989	1.000
##	4	219	1	0.992	0.00578	0.981	1.000
##	6	108	6	0.937	0.02253	0.894	0.982
##	7	71	6	0.858	0.03717	0.788	0.934
##	8	46	4	0.783	0.04921	0.692	0.886
##	9	24	4	0.653	0.07232	0.525	0.811
##	10	13	4	0.452	0.09738	0.296	0.689

```
plot(kmsurvival5, xlab = "Week", ylab = "Femur present", col= rainbow(2, alpha = 0.3),mark = 3)
```



```
femur1<-coxme((mysurv5) ~ Position +(1|Site/Clutchx), data=Stage54)
femur2<- coxme((mysurv5) ~ 1 +(1|Site/Clutchx), data=Stage54)
```

AIC(femur1, femur2)#No difference between null model and model with position as fixed effect.

```
##           df      AIC
## femur1 4.737562 405.9154
## femur2 4.224326 405.1399
```

```
summary(femur2)
```



```
## Cox mixed-effects model fit by maximum likelihood
## Data: Stage54
## events, n = 44, 745
## Iterations= 6 27
##
##              NULL Integrated      Fitted
## Log-likelihood -203.1177      -202.839 -198.3456
##
##              Chisq    df          p    AIC    BIC
## Integrated loglik  0.56 2.00 0.756810 -3.44 -7.01
## Penalized loglik  9.54 4.22 0.056582  1.10 -6.44
##
## Model: (mysurv5) ~ 1 + (1 | Site/Clutchx)
##
## Random effects
## Group          Variable      Std Dev      Variance
## Site/Clutchx (Intercept) 0.3598328215 0.1294796594
## Site           (Intercept) 0.0198619296 0.0003944962
```

```
#p-value= 0.35
detach(Stage54)
```

According to theory, models with AIC within two points of each other are basically equal. Thus, both model can be chosen as suitable. No significant effect of position on femur development was observed for tadpoles raised in the mesocosms. We retained the model that displayed no effect of position ($p = 0.35$) random effects have a minor effect.

3. What was the effect of range expansion on the morphology of tadpoles?

+To assess the effect of position (core vs. periphery) on individual morphology of mesocosm individuals, data was divided into tadpoles (NF stages 0-57) and climax animals (NF stages 58-65) and metamorphs (stage 66- in the lab and the mesocosms) according to Segerdell et al. (2008). All body measurements (response variables) were tested for normality. All variables were Ordered Quantile normalization transformed in the bestNormalize package to improve normality. To obtain a synthetic measurement of body size for tadpoles, we conducted a principle component analysis on the measurements (Snout-vent-length, head width, maximum body depth, tail length, and tail depth) of tadpole morphology. The first principal component accounted for 94.9% for stages < 57. And the measurements (Snout-vent-length, head width, maximum body depth, total leg length) of climax morphology. The first principal component accounted for 91.7% for stages 58-65. Thus, the first axis is a global measurement of body size. Generalised mixed models were used with body size (PC1), and Snout-vent length as response variables with position (core/periphery), collection week (week 1-10), and stage (NF stages 0-57) as fixed effects. Clutch nested within site was considered as a random effect. We compared different models to determine the best fitted model by comparing Akaike information criterion (AIC).

3.1. Tadpoles NF stages 0-57 (late tail bud, premetamorphosis, prometamorphosis)

+3.1.1. Defining variables. Use *Tadpole meas.csv* to analyse data.

```
Tadpole=read.csv(file.choose(),header=T)
names(Tadpole)
```

```
## [1] "num"           "Position"      "Site"
## [4] "Clutch"        "Clutchx"       "Tadpole"
## [7] "Week"          "tadpoleID"     "Mesocosm"
## [10] "Date"          "Fulllength"    "Maximumbodydepth"
## [13] "Snoutventlength" "Taillength"    "Reltaillength"
## [16] "Maxtaildepth"  "Headwidth"     "Femur"
## [19] "Femurrelative" "Totalleglength" "Stage"
## [22] "Stagecategory"
```

```
attach(Tadpole)
```

```
## The following object is masked _by_ .GlobalEnv:
##
##      Tadpole
```

```
summary(Position) #fixed effect
```

```
##      Core Periphery
##      449         483
```

```
cor.test(Week, Stage, method="kendall")
```

```
##
## Kendall's rank correlation tau
##
## data: Week and Stage
## z = 27.527, p-value < 2.2e-16
## alternative hypothesis: true tau is not equal to 0
## sample estimates:
##      tau
## 0.6582338
```

```
#p-value< 2.2e-16
#tau 0.6585802
#According to the kendall correlation test "Stage" and "Week" is highly correlated,
#Thus, We cannot add both week and stage due to colinearity.
```

```
summary(Week) #fixed effect
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      0.000   2.000   4.000   3.838   6.000   9.000
```

```
Week1 <- as.factor(Week) #as a discrete variable
summary(Week1)
```

```
##    0    1    2    3    4    5    6    7    8    9
## 114 115 114 115 110  96  80  74  55  59
```

```
summary(Stage) #fixed effect
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   39.00   48.00   50.00   50.33   53.00   57.00
```

```
Stage1 <- as.factor(Stage) #as a discrete variable
summary(Stage1)
```

```
##   39  40  41  42  44  45  46  47  48  49  50  51  52  53  54  55  56  57
##   10   6   9  14  18  39  47  43 138 133  42  49 104  49  44  60  70  57
```

```
summary(Site) #Random effect
```

```
##    B    C    J    M    S    T
## 156 164 160 121 172 159
```

```
summary(Clutchx) #Nested within site
```

```
## B1 B2 B3 B4 C1 C2 C3 C4 J1 J2 J3 J4 M1 M2 M4 S1 S2 S3 S4 T1 T2 T3 T4
## 50 40 33 33 35 46 43 40 45 36 39 40 42 41 38 42 48 37 45 41 38 43 37
```

```
Tadpole1<- cbind(Tadpole, Week1, Stage1)
names(Tadpole1)
```

```
## [1] "num"           "Position"      "Site"
## [4] "Clutch"        "Clutchx"       "Tadpole"
## [7] "Week"          "tadpoleID"     "Mesocosm"
## [10] "Date"          "Fulllength"    "Maximumbodydepth"
## [13] "Snoutventlength" "Taillength"    "Reltaillength"
## [16] "Maxtaildepth"  "Headwidth"     "Femur"
## [19] "Femurrelative" "Totalleglength" "Stage"
## [22] "Stagecategory" "Week1"         "Stage1"
```

+3.1.2. Transforming response variables.

Use the *Tadpole1* datasheet just created to measure normality.

Snoutventlength

```
library(bestNormalize)
```

```
## Registered S3 methods overwritten by 'registry':
##   method                from
##   print.registry_field proxy
##   print.registry_entry proxy
```

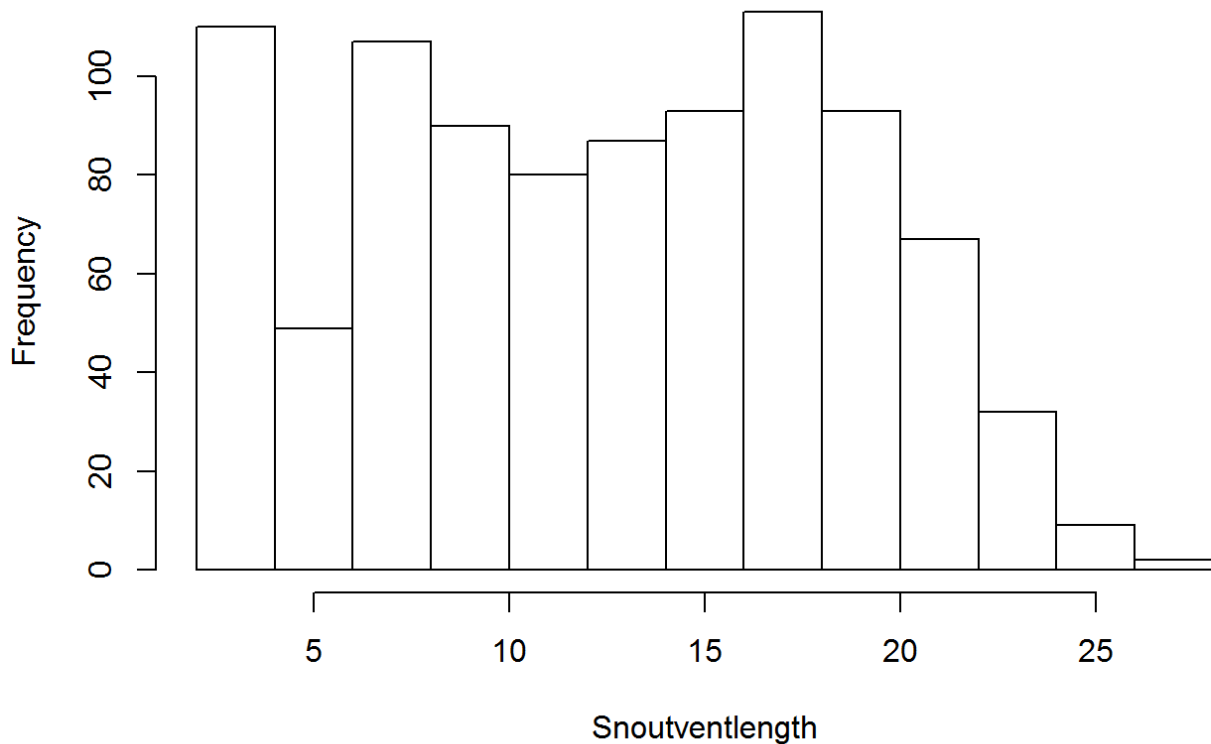
#This package tests the most appropriate normalizing transformations and picks the best one based off of a goodness of fit statistic (see package details)

```
shapiro.test(Snoutventlength)
```

```
##
## Shapiro-Wilk normality test
##
## data:  Snoutventlength
## W = 0.96562, p-value = 5.095e-14
```

```
#p-value = 5.095e-14 (tadpoles) not normal
hist(Snoutventlength)
```

Histogram of Snoutventlength



```
bestNormalize(Snoutventlength, standardize= TRUE, allow_orderNorm=TRUE,
              allow_lambert_s = FALSE, allow_lambert_h = FALSE,
              out_of_sample = TRUE, cluster = NULL, k = 10, r = 5,
              loo = FALSE, warn = TRUE, quiet = FALSE)
```

```
## Warning in orderNorm(standardize = TRUE, warn = TRUE, x = c(3.321, 3.45, : Ties
in data, Normal distribution not guaranteed
```

```
## Best Normalizing transformation with 932 Observations
## Estimated Normality Statistics (Pearson P / df, lower => more normal):
## - No transform: 1.7359
## - Box-Cox: 1.8933
## - Log_b(x+a): 3.8485
## - sqrt(x+a): 2.4212
## - exp(x): 83.9858
## - arcsinh(x): 3.8364
## - Yeo-Johnson: 1.861
## - orderNorm: 1.1917
## Estimation method: Out-of-sample via CV with 10 folds and 5 repeats
##
## Based off these, bestNormalize chose:
## orderNorm Transformation with 932 nonmissing obs and ties
## - 911 unique values
## - Original quantiles:
##      0%      25%      50%      75%     100%
## 2.292  7.451 12.697 17.399 27.596
```

```
#for tadpoles
#Based off these, bestNormalize chose:
#orderNorm Transformation with 932 nonmissing obs and ties
#The Ordered Quantile (ORQ) normalization transformation, orderNorm(), is a rank-ba
sed procedure by which the values of a vector are mapped to their percentile, which
is then mapped to the same percentile of the normal distribution. Without the prese
nce of ties, this essentially guarantees that the transformation leads to a uniform
distribution.
```

```
orderNorm_SVL <- orderNorm(Snoutventlength)
```

```
## Warning in orderNorm(Snoutventlength): Ties in data, Normal distribution not gua
ranteed
```

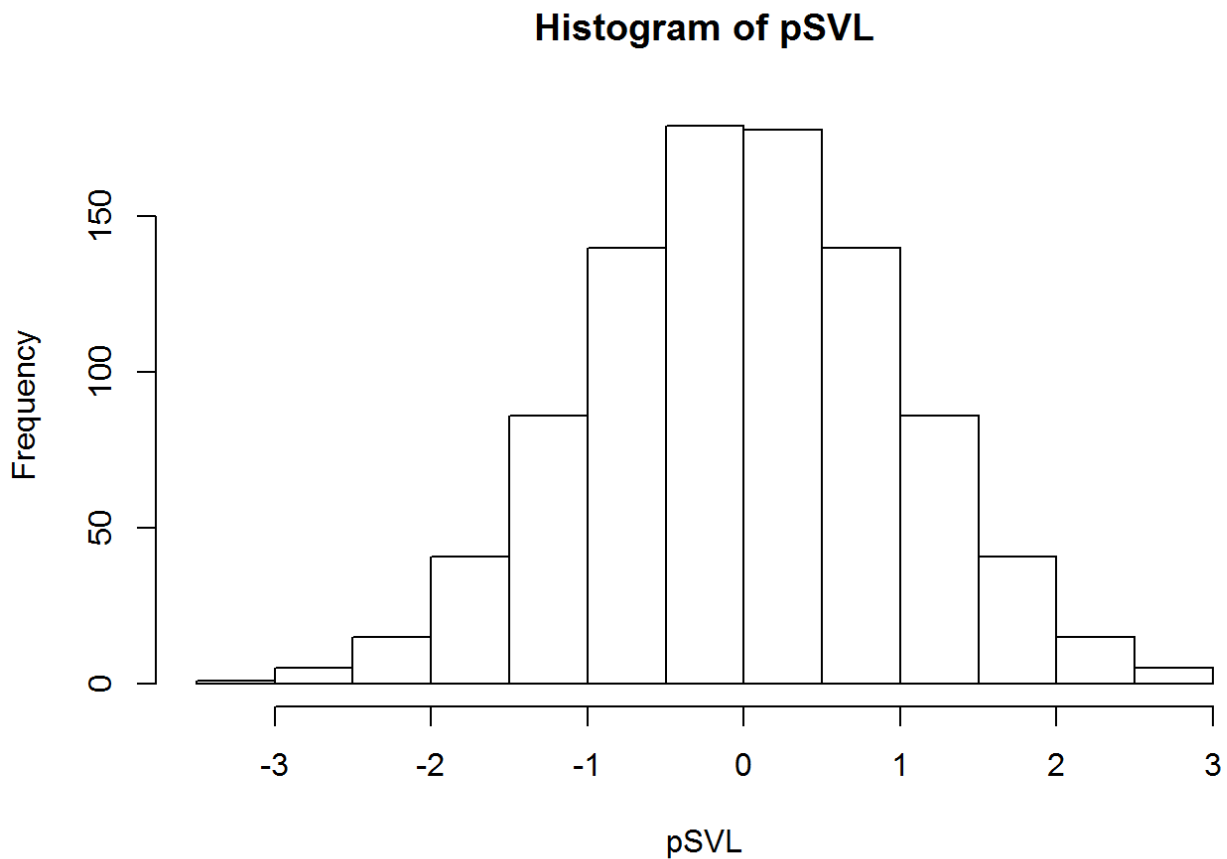
```
pSVL <- predict(orderNorm_SVL)
summary(pSVL)
```

```
##      Min.    1st Qu.      Median        Mean     3rd Qu.      Max.
## -3.270965 -0.674490 -0.001343 -0.003509  0.671121  2.946355
```

```
shapiro.test(pSVL)
```

```
##
## Shapiro-Wilk normality test
##
## data:  pSVL
## W = 0.99969, p-value = 1
```

```
#p-value= 1 (tadpoles)
hist(pSVL)
```



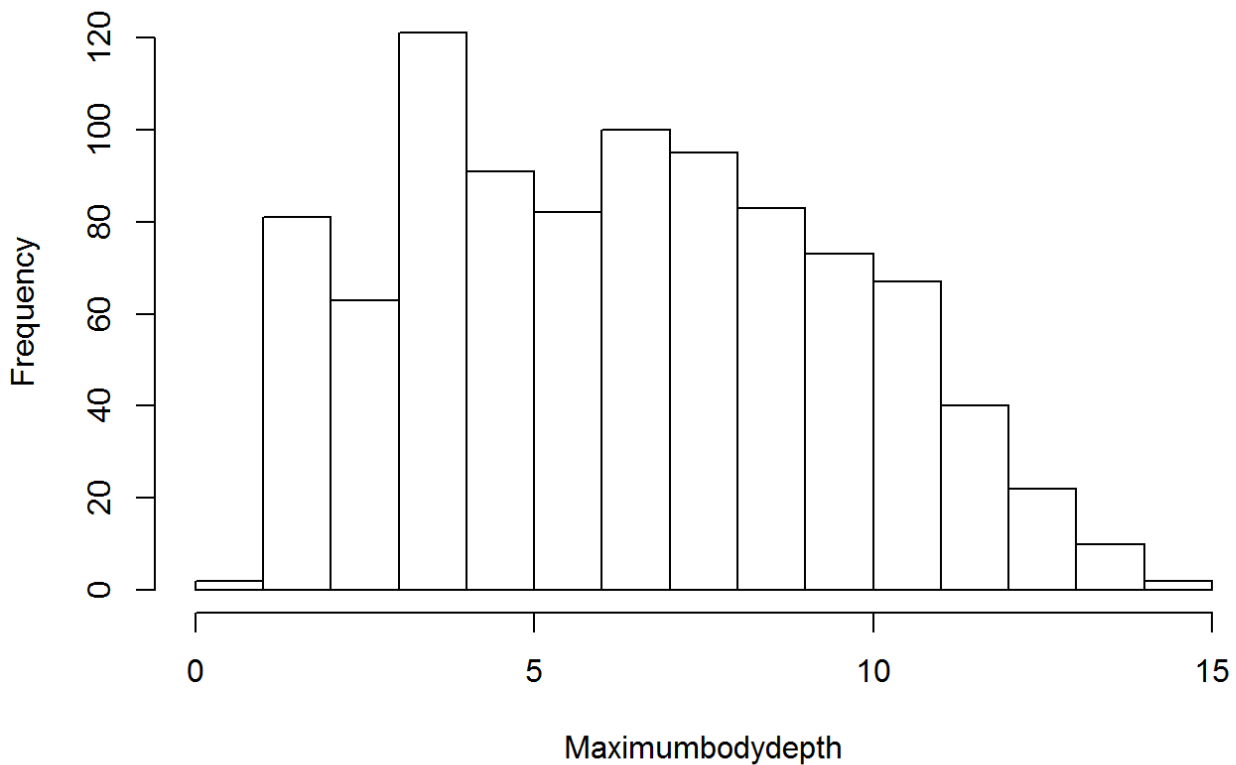
Maximumbodydepth

```
shapiro.test(Maximumbodydepth)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  Maximumbodydepth
## W = 0.97342, p-value = 5.106e-12
```

```
#p-value = 5.106e-12 (tadpoles) not normal
hist(Maximumbodydepth)
```

Histogram of Maximumbodydepth



```
bestNormalize(Maximumbodydepth, standardize= TRUE, allow_orderNorm=TRUE,
              allow_lambert_s = FALSE, allow_lambert_h = FALSE,
              out_of_sample = TRUE, cluster = NULL, k = 10, r = 5,
              loo = FALSE, warn = TRUE, quiet = FALSE)
```

```
## Warning in orderNorm(standardize = TRUE, warn = TRUE, x = c(2.045, 2.194, : Ties
in data, Normal distribution not guaranteed
```

```
## Best Normalizing transformation with 932 Observations
## Estimated Normality Statistics (Pearson P / df, lower => more normal):
## - No transform: 1.5356
## - Box-Cox: 1.4503
## - Log_b(x+a): 2.6767
## - sqrt(x+a): 1.4754
## - exp(x): 60.7677
## - arcsinh(x): 2.5681
## - Yeo-Johnson: 1.4331
## - orderNorm: 1.122
## Estimation method: Out-of-sample via CV with 10 folds and 5 repeats
##
## Based off these, bestNormalize chose:
## orderNorm Transformation with 932 nonmissing obs and ties
## - 890 unique values
## - Original quantiles:
##    0%    25%    50%    75%   100%
## 0.543  3.723  6.277  8.747 14.741
```

```
#Based off these, bestNormalize chose:
#orderNorm Transformation with 932 nonmissing obs and ties

orderNorm_Bodydepth <- orderNorm(Maximumbodydepth)
```

```
## Warning in orderNorm(Maximumbodydepth): Ties in data, Normal distribution not guaranteed
```

```
pMaximumbodydepth <- predict(orderNorm_Bodydepth)
summary(pMaximumbodydepth)
```

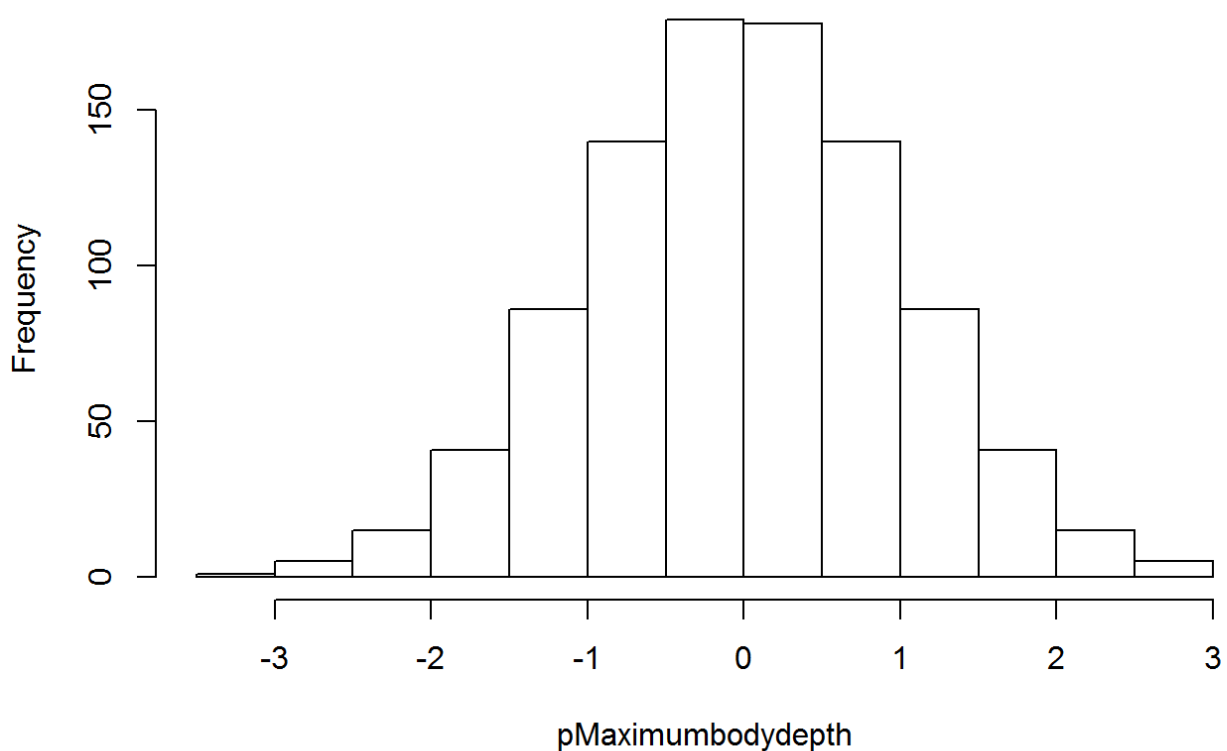
```
##      Min.   1st Qu.   Median     Mean   3rd Qu.     Max.
## -3.270965 -0.674490 -0.001343 -0.003510  0.671121  2.946355
```

```
shapiro.test(pMaximumbodydepth)
```

```
##
## Shapiro-Wilk normality test
##
## data:  pMaximumbodydepth
## W = 0.99969, p-value = 1
```

```
#p-value= 1 (tadpoles)
hist(pMaximumbodydepth)
```

Histogram of pMaximumbodydepth



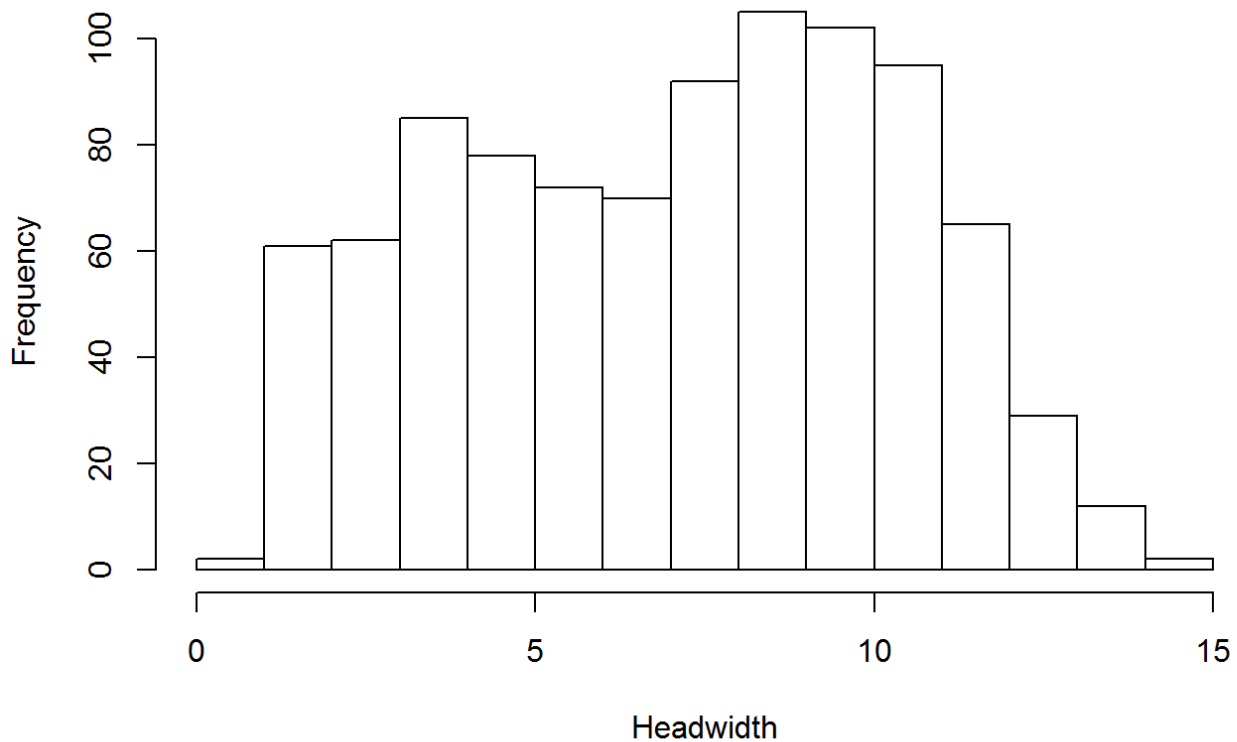
Headwidth

```
shapiro.test(Headwidth)
```

```
##  
## Shapiro-Wilk normality test  
##  
## data: Headwidth  
## W = 0.96753, p-value = 1.466e-13
```

```
#p-value = 1.466e-13 (tadpoles)  
hist(Headwidth)
```

Histogram of Headwidth



```
bestNormalize(Headwidth, standardize= TRUE, allow_orderNorm=TRUE,  
              allow_lambert_s = FALSE, allow_lambert_h = FALSE,  
              out_of_sample = TRUE, cluster = NULL, k = 10, r = 5,  
              loo = FALSE, warn = TRUE, quiet = FALSE)
```

```
## Warning in orderNorm(standardize = TRUE, warn = TRUE, x = c(2.301, 2.214, : Ties  
in data, Normal distribution not guaranteed
```

```
## Best Normalizing transformation with 932 Observations
## Estimated Normality Statistics (Pearson P / df, lower => more normal):
## - No transform: 1.7507
## - Box-Cox: 1.8945
## - Log_b(x+a): 4.3201
## - sqrt(x+a): 2.4998
## - exp(x): 46.5841
## - arcsinh(x): 4.1488
## - Yeo-Johnson: 1.837
## - orderNorm: 1.1338
## Estimation method: Out-of-sample via CV with 10 folds and 5 repeats
##
## Based off these, bestNormalize chose:
## orderNorm Transformation with 932 nonmissing obs and ties
## - 890 unique values
## - Original quantiles:
##      0%      25%      50%      75%     100%
## 0.961  4.324  7.399  9.723 14.164
```

```
#Based off these, bestNormalize chose:
#orderNorm Transformation with 932 nonmissing obs and ties

orderNorm_Headwidth <- orderNorm(Headwidth)
```

```
## Warning in orderNorm(Headwidth): Ties in data, Normal distribution not guaranteed
```

```
pHeadwidth <- predict(orderNorm_Headwidth)
summary(orderNorm_Headwidth)
```

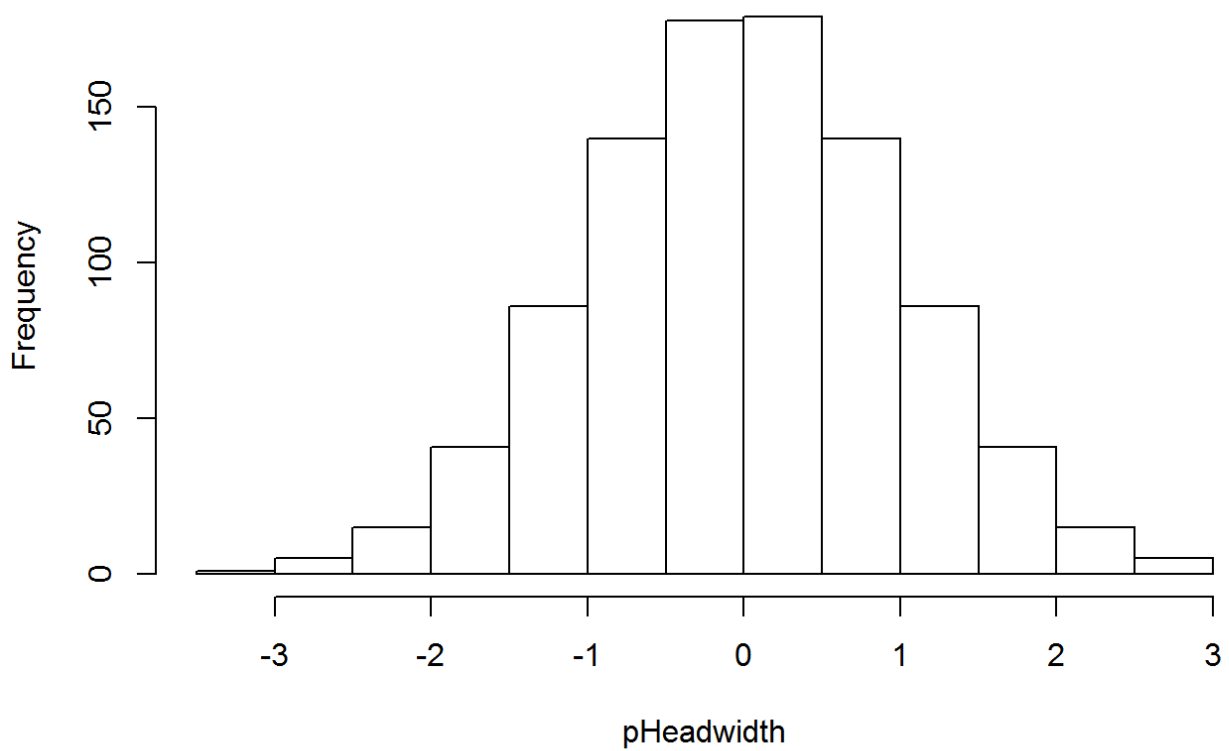
```
##           Length Class  Mode
## x.t         932    -none- numeric
## x           932    -none- numeric
## n              1    -none- numeric
## ties_status   1    -none- numeric
## fit           30     glm      list
## norm_stat      1    -none- numeric
```

```
shapiro.test(pHeadwidth)
```

```
##
## Shapiro-Wilk normality test
##
## data:  pHeadwidth
## W = 0.99969, p-value = 1
```

```
#p-value = 1 (tadpoles)
hist(pHeadwidth)
```

Histogram of pHeadwidth



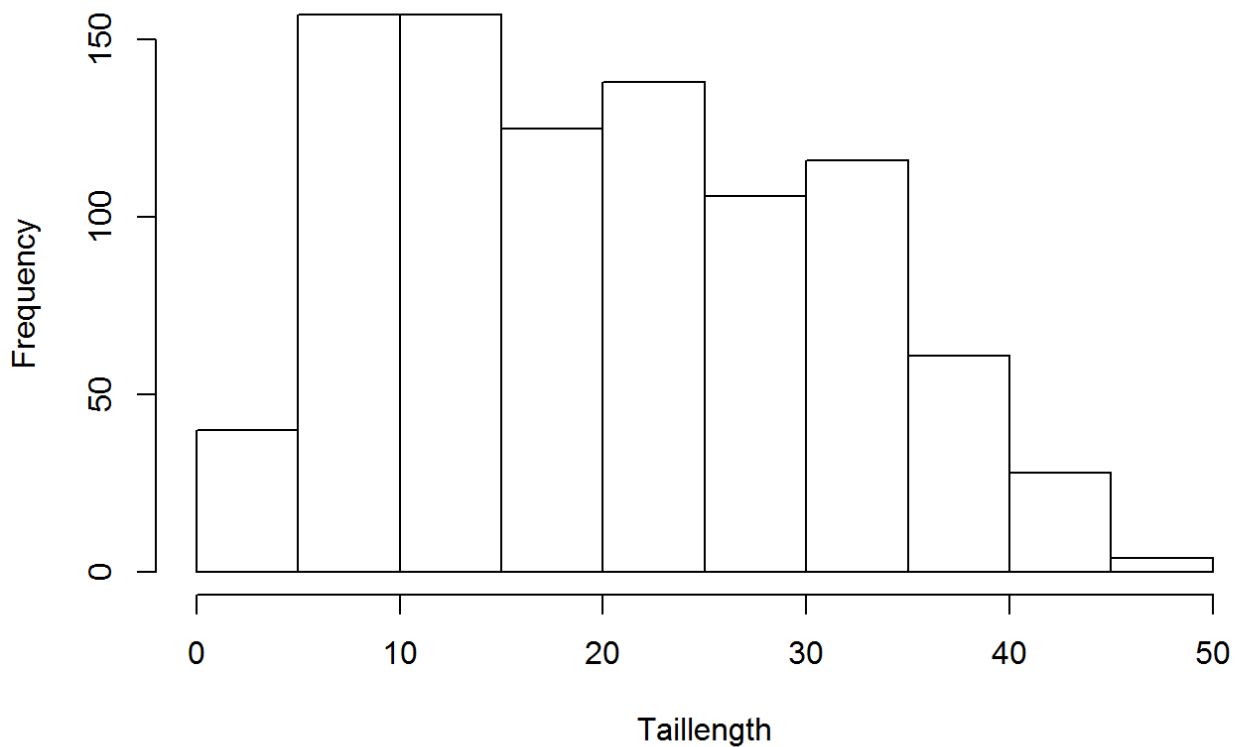
Taillength

```
shapiro.test(Taillength)
```

```
##  
## Shapiro-Wilk normality test  
##  
## data: Taillength  
## W = 0.96655, p-value = 8.474e-14
```

```
#p-value = 8.474e-14 (tadpoles)  
hist(Taillength)
```

Histogram of Taillength



```
bestNormalize(Taillength, standardize= TRUE, allow_orderNorm=TRUE,
              allow_lambert_s = FALSE, allow_lambert_h = FALSE,
              out_of_sample = TRUE, cluster = NULL, k = 10, r = 5,
              loo = FALSE, warn = TRUE, quiet = FALSE)
```

```
## Warning in orderNorm(standardize = TRUE, warn = TRUE, x = c(5.905, 5.434, : Ties
in data, Normal distribution not guaranteed
```

```
## Best Normalizing transformation with 932 Observations
## Estimated Normality Statistics (Pearson P / df, lower => more normal):
## - No transform: 2.0469
## - Box-Cox: 1.5113
## - Log_b(x+a): 2.5728
## - sqrt(x+a): 1.5484
## - exp(x): 100.9891
## - arcsinh(x): 2.5672
## - Yeo-Johnson: 1.5245
## - orderNorm: 1.067
## Estimation method: Out-of-sample via CV with 10 folds and 5 repeats
##
## Based off these, bestNormalize chose:
## orderNorm Transformation with 932 nonmissing obs and ties
## - 919 unique values
## - Original quantiles:
##    0%    25%    50%    75%   100%
## 1.916 11.360 19.419 28.669 48.292
```

```
#Based off these, bestNormalize chose:
#orderNorm Transformation with 932 nonmissing obs and ties

orderNorm_Taillength <- orderNorm(Taillength)
```

```
## Warning in orderNorm(Taillength): Ties in data, Normal distribution not guaranteed
```

```
pTaillength <- predict(orderNorm_Taillength)
summary(pTaillength)
```

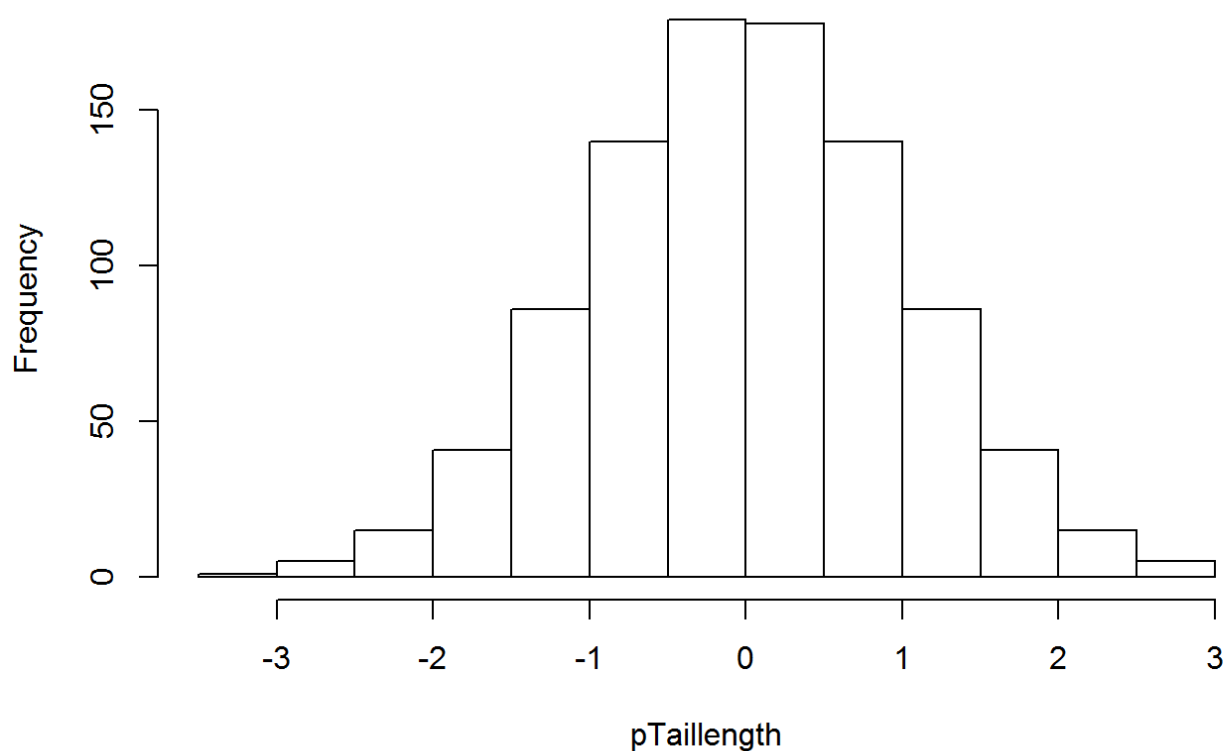
```
##      Min.   1st Qu.   Median     Mean   3rd Qu.     Max.
## -3.270965 -0.674490 -0.001343 -0.003510  0.671121  2.946355
```

```
shapiro.test(pTaillength)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  pTaillength
## W = 0.99969, p-value = 1
```

```
#p-value = 1
hist(pTaillength)
```

Histogram of pTaillength



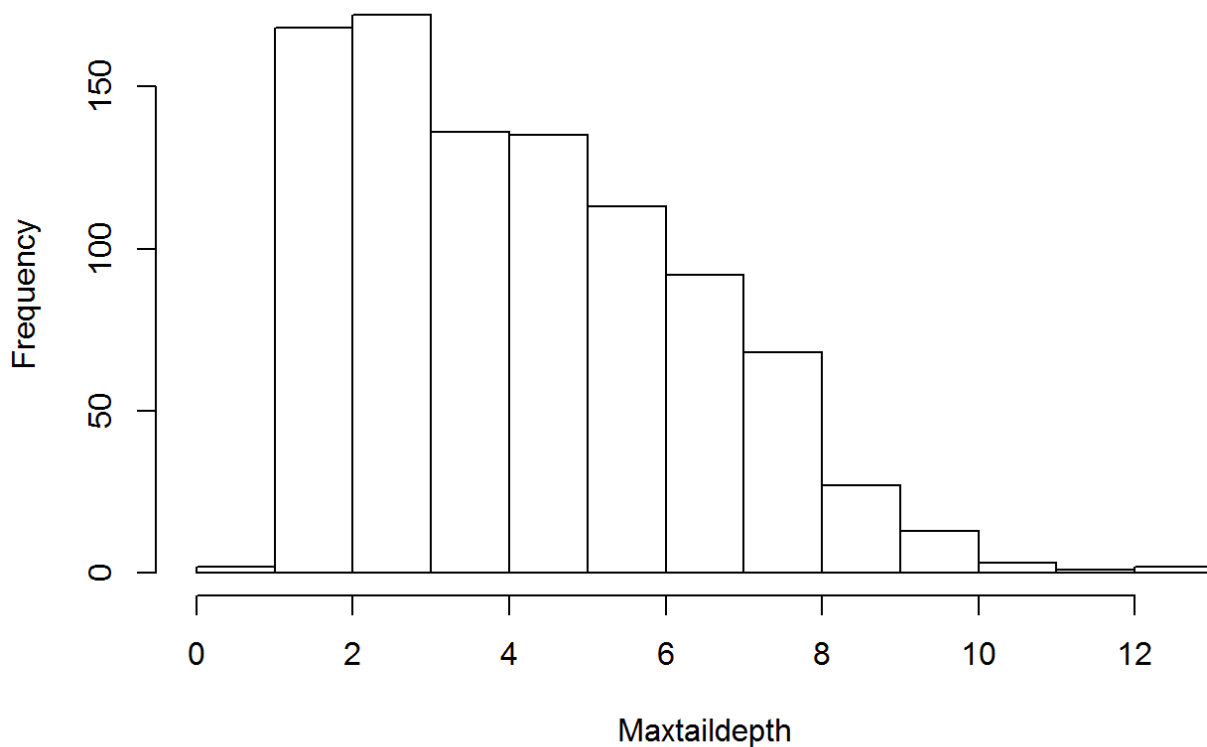
Taildepth

```
shapiro.test(Maxtaildepth)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  Maxtaildepth
## W = 0.95237, p-value < 2.2e-16
```

```
#p-value < 2.2e-16 (tadpoles)
hist(Maxtaildepth)
```

Histogram of Maxtaildepth



```
bestNormalize(Maxtaildepth, standardize= TRUE, allow_orderNorm=TRUE,
              allow_lambert_s = FALSE, allow_lambert_h = FALSE,
              out_of_sample = TRUE, cluster = NULL, k = 10, r = 5,
              loo = FALSE, warn = TRUE, quiet = FALSE)
```

```
## Warning in orderNorm(standardize = TRUE, warn = TRUE, x = c(1.065, 1.092, : Ties
in data, Normal distribution not guaranteed
```

```
## Best Normalizing transformation with 932 Observations
## Estimated Normality Statistics (Pearson P / df, lower => more normal):
## - No transform: 2.7904
## - Box-Cox: 1.6608
## - Log_b(x+a): 2.0164
## - sqrt(x+a): 1.6025
## - exp(x): 57.2184
## - arcsinh(x): 1.9118
## - Yeo-Johnson: 1.7183
## - orderNorm: 1.0247
## Estimation method: Out-of-sample via CV with 10 folds and 5 repeats
##
## Based off these, bestNormalize chose:
## orderNorm Transformation with 932 nonmissing obs and ties
## - 863 unique values
## - Original quantiles:
##      0%      25%      50%      75%     100%
## 0.932  2.329  3.894  5.691 12.754
```

```
#Based off these, bestNormalize chose:
#orderNorm Transformation with 932 nonmissing obs and ties

orderNorm_Taildepth <- orderNorm(Maxtaildepth)
```

```
## Warning in orderNorm(Maxtaildepth): Ties in data, Normal distribution not guaranteed
```

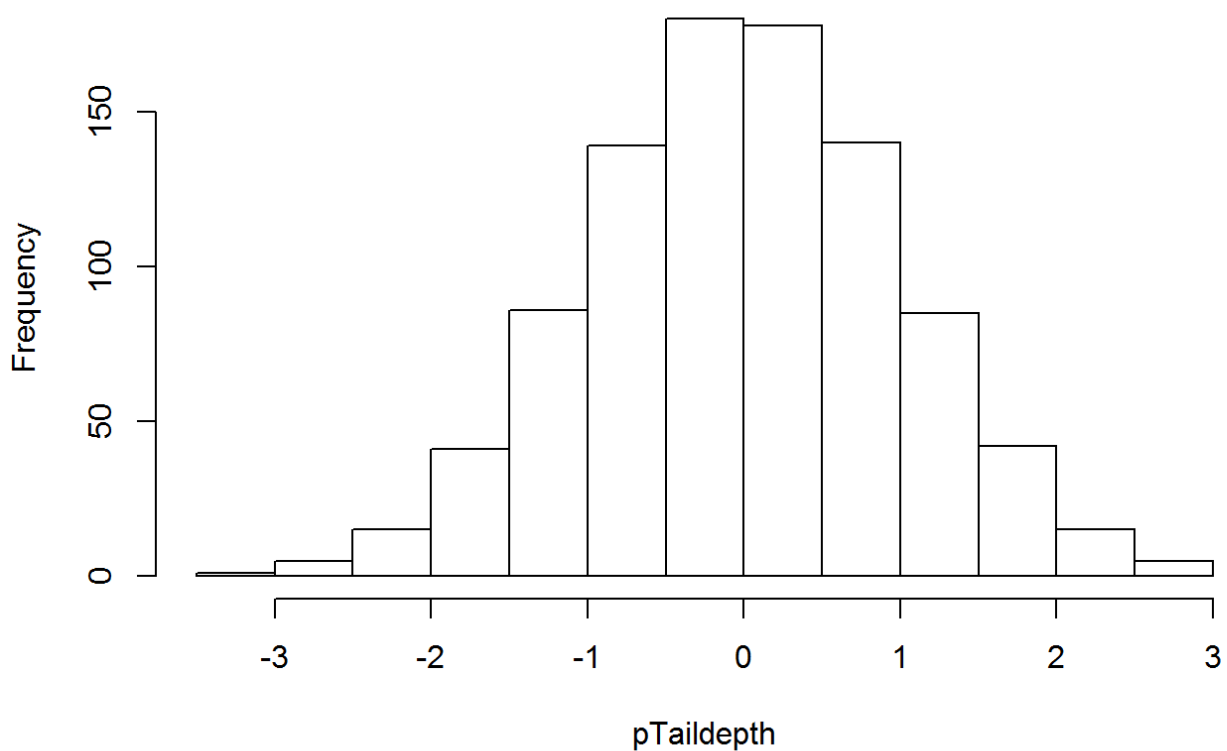
```
pTaildepth <- predict(orderNorm_Taildepth)
summary(pTaildepth)
```

```
##      Min.    1st Qu.    Median      Mean   3rd Qu.      Max.
## -3.270965 -0.673227 -0.001343 -0.003509  0.671543  2.946355
```

```
shapiro.test(pTaildepth)
```

```
##
## Shapiro-Wilk normality test
##
## data:  pTaildepth
## W = 0.99969, p-value = 1
```

```
#p-value = 1
hist(pTaildepth)
```

Histogram of pTaildepth

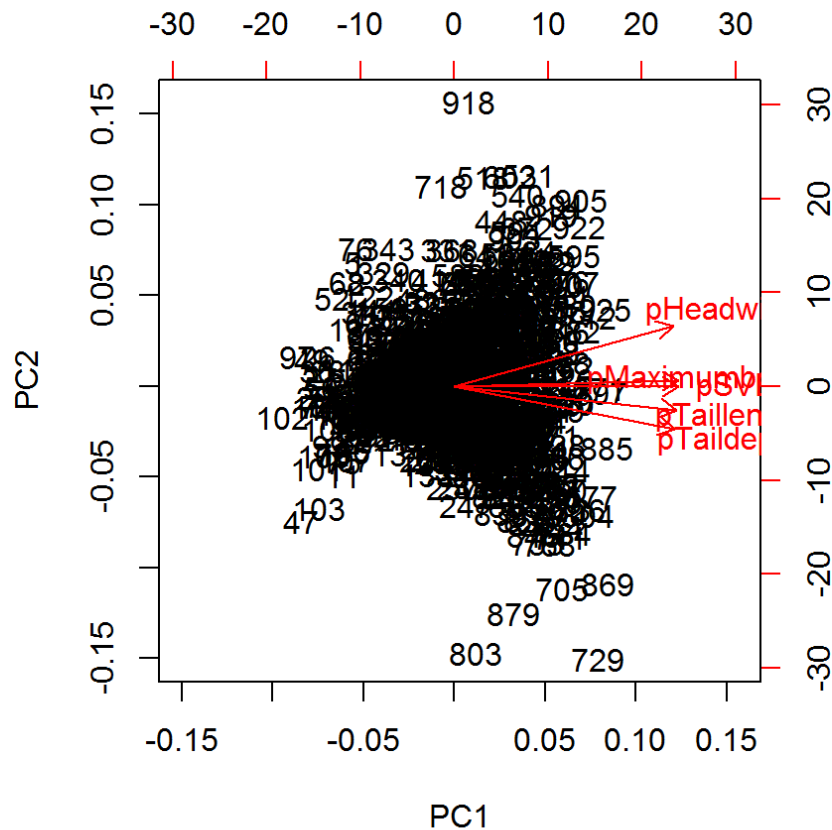
3.1.3. Performing PCA for tadpole variables.

```
tadpolepca<- cbind(pHeadwidth, pSVL, pTaillength, pTaildepth, pMaximumbodydepth)

tadpca1=prcomp(tadpolepca)
summary(tadpca1)
```

```
## Importance of components:
##              PC1      PC2      PC3      PC4      PC5
## Standard deviation  2.1652 0.34058 0.24536 0.22224 0.18081
## Proportion of Variance 0.9478 0.02345 0.01217 0.00999 0.00661
## Cumulative Proportion 0.9478 0.97123 0.98340 0.99339 1.00000
```

```
#proportion of variance for PC1- 0.9478
biplot(tadpca1)
```

```
tadpcalscores<-predict(tadpca1)
```

```
taddataset <- cbind(Tadpole1, tadpolepca, tadpcalscores)
names(taddataset)
```

```
## [1] "num"           "Position"      "Site"
## [4] "Clutch"        "Clutchx"      "Tadpole"
## [7] "Week"         "tadpoleID"    "Mesocosm"
## [10] "Date"         "Fulllength"   "Maximumbodydepth"
## [13] "Snoutventlength" "Taillength"   "Reltaillength"
## [16] "Maxtaildepth" "Headwidth"    "Femur"
## [19] "Femurrelative" "Totalleglength" "Stage"
## [22] "Stagecategory" "Week1"        "Stage1"
## [25] "pHeadwidth"    "pSVL"         "pTaillength"
## [28] "pTaildepth"    "pMaximumbodydepth" "PC1"
## [31] "PC2"          "PC3"         "PC4"
## [34] "PC5"
```

3.1.4. Model selection for variables.

PC1

Use the *taddataset* for the analysis.

```
#For week and stage as discrete variables
PC1a<- lmer(PC1 ~ Position*Week1+ (1|Site/Clutchx), data = taddataset)
```

```
## boundary (singular) fit: see ?isSingular
```

```
PC2a<- lmer(PC1 ~ Position + Week1 + (1|Site/Clutchx), data = taddataset)
```

```
## boundary (singular) fit: see ?isSingular
```

```
PC3a<- lmer(PC1 ~ Week1 + (1|Site/Clutchx), data=taddataset)
```

```
## boundary (singular) fit: see ?isSingular
```

```
PC4a<- lmer(PC1 ~ Position + (1|Site/Clutchx), data = taddataset)
```

```
## boundary (singular) fit: see ?isSingular
```

```
#some convergence errors occur so Position*Stage was removed.
PC6a<- lmer(PC1 ~ Position + Stage1 + (1|Site/Clutchx), data = taddataset)
```

```
## boundary (singular) fit: see ?isSingular
```

```
PC7a<- lmer(PC1 ~ Stage1 + (1|Site/Clutchx), data= taddataset)
```

```
## boundary (singular) fit: see ?isSingular
```

```
PC8a<- lmer(PC1 ~ 1 + (1|Site/Clutchx), data=taddataset)
```

```
## boundary (singular) fit: see ?isSingular
```

```
AIC(PC1a, PC2a, PC3a, PC4a,PC6a, PC7a, PC8a)
```

```
##      df      AIC
## PC1a 23 2800.139
## PC2a 14 2786.801
## PC3a 13 2784.227
## PC4a  5 4097.121
## PC6a 22 2255.193
## PC7a 21 2250.442
## PC8a  4 4094.832
```

```
AICctab(PC1a, PC2a, PC3a, PC4a, PC6a, PC7a, PC8a)
```

```
##      dAICc  df
## PC7a    0.0 21
## PC6a    4.8 22
## PC3a   533.2 13
## PC2a   535.8 14
## PC1a   549.9 23
## PC8a  1843.4  4
## PC4a  1845.7  5
```

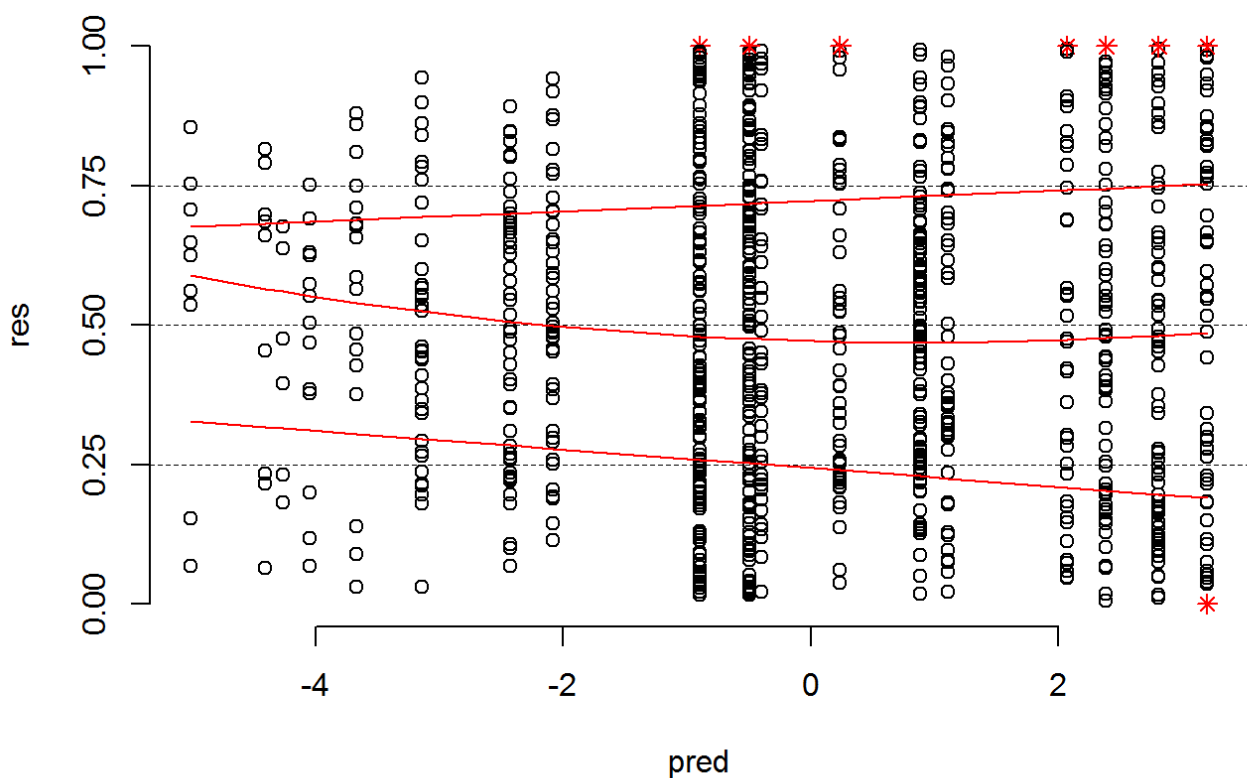
#PC7a was selected as the model with the lowest AIC
summary(PC7a)

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: PC1 ~ Stage1 + (1 | Site/Clutchx)
## Data: taddataset
##
## REML criterion at convergence: 2208.4
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -3.4600 -0.6447 -0.0402  0.5754  3.4957
##
## Random effects:
## Groups      Name                Variance Std.Dev.
## Clutchx:Site (Intercept) 0.03792  0.1947
## Site         (Intercept) 0.00000  0.0000
## Residual                        0.59266  0.7698
## Number of obs: 932, groups: Clutchx:Site, 23; Site, 6
##
## Fixed effects:
##              Estimate Std. Error t value
## (Intercept)  -5.0066    0.2536 -19.739
## Stage140      0.7483    0.4117  1.817
## Stage141      0.6027    0.3589  1.679
## Stage142      0.9590    0.3293  2.912
## Stage144      1.3384    0.3127  4.280
## Stage145      1.8713    0.2819  6.638
## Stage146      2.5839    0.2756  9.376
## Stage147      2.9285    0.2773 10.561
## Stage148      4.1127    0.2598 15.830
## Stage149      4.5117    0.2595 17.388
## Stage150      4.6052    0.2760 16.684
## Stage151      5.2444    0.2741 19.133
## Stage152      5.8904    0.2619 22.489
## Stage153      6.1180    0.2738 22.348
## Stage154      7.0749    0.2748 25.747
## Stage155      7.3911    0.2694 27.440
## Stage156      7.8160    0.2675 29.216
## Stage157      8.2049    0.2706 30.319
```

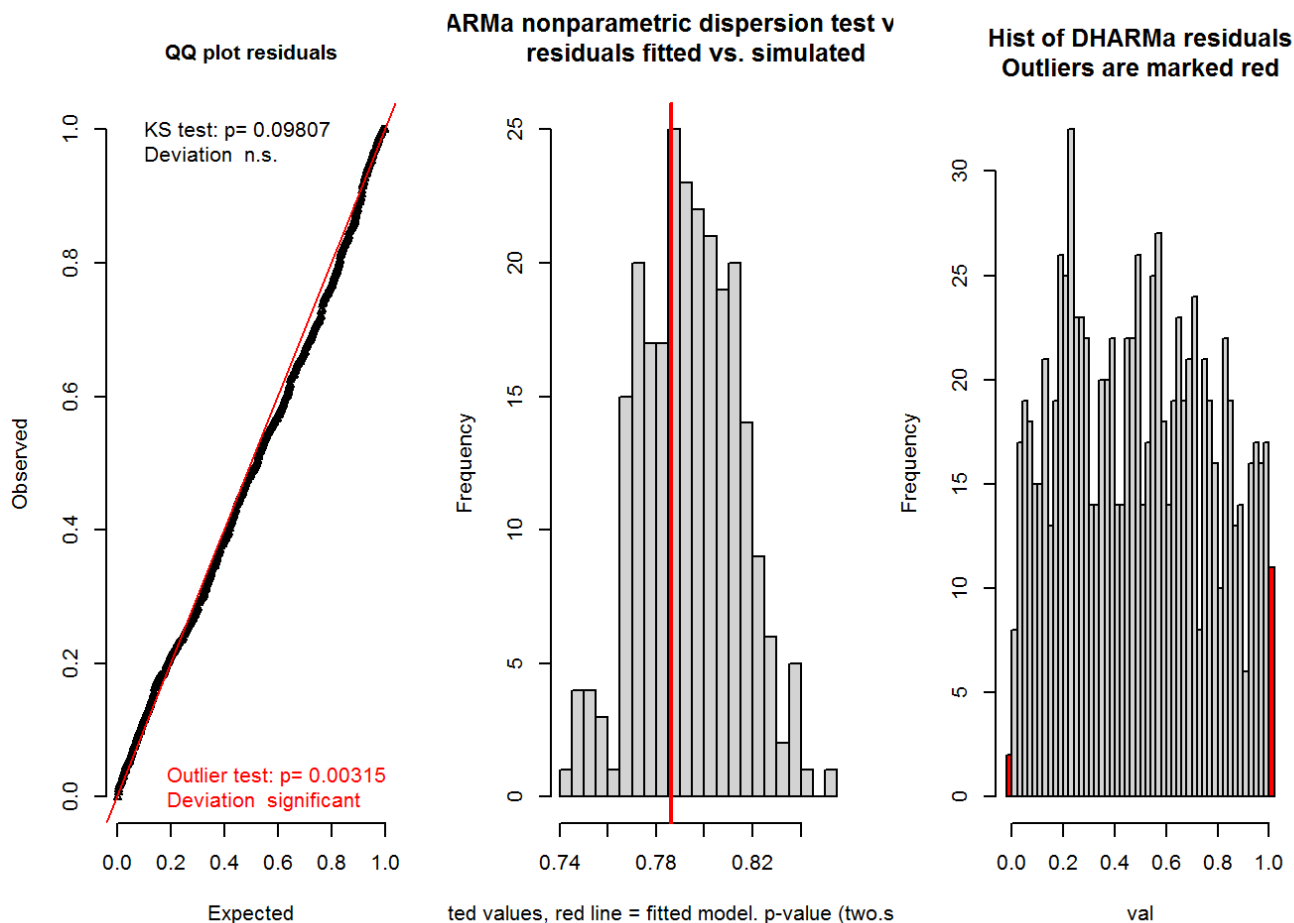
```
##
## Correlation matrix not shown by default, as p = 18 > 12.
## Use print(x, correlation=TRUE) or
##      vcov(x)          if you need it
```

```
## convergence code: 0
## boundary (singular) fit: see ?isSingular
```

```
rr=simulateResiduals(PC7a) ### simulate residuals
plotResiduals(rr) ### qqplot simulated vs observed
```



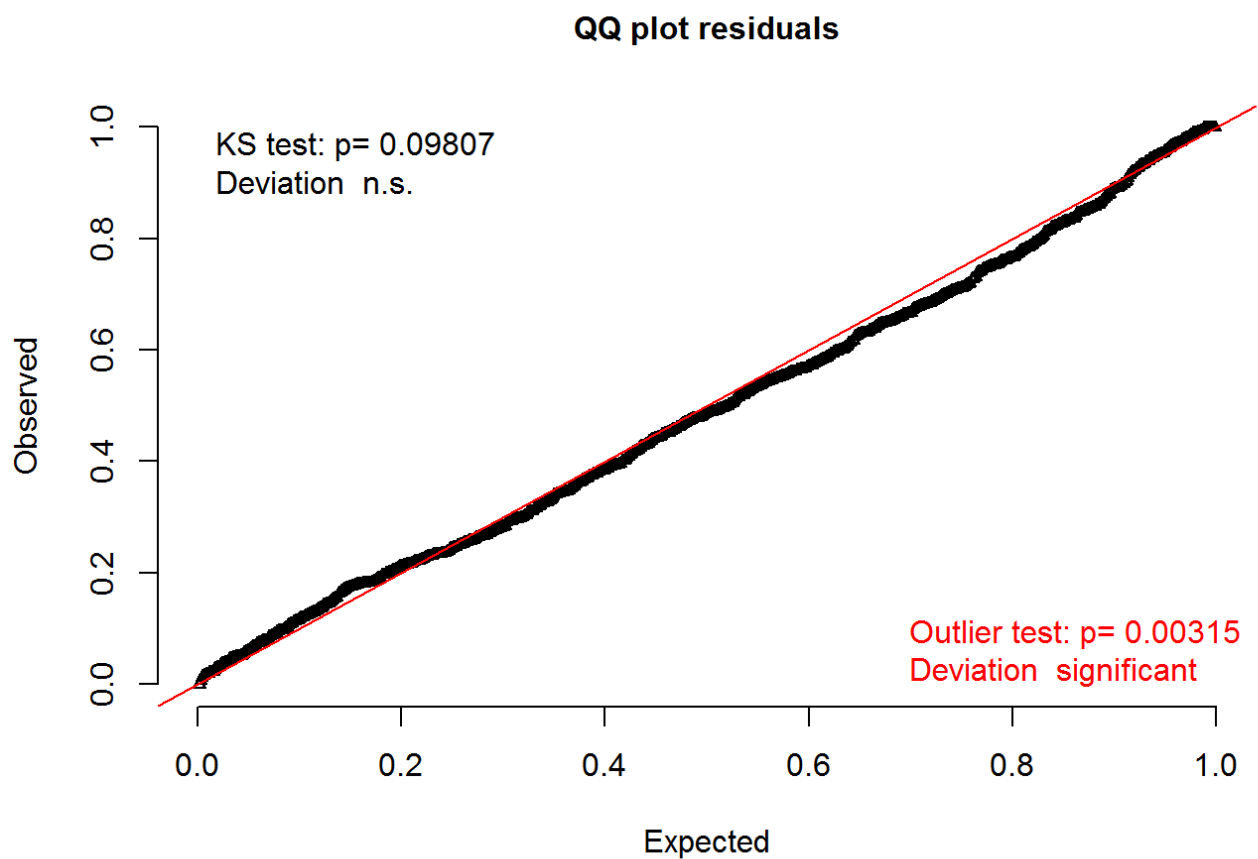
```
testResiduals(rr)###test qqplot
```



```
## $uniformity
##
## One-sample Kolmogorov-Smirnov test
##
## data: simulationOutput$scaledResiduals
## D = 0.040219, p-value = 0.09807
## alternative hypothesis: two-sided
##
##
## $dispersion
##
## DHARMA nonparametric dispersion test via sd of residuals fitted
## vs. simulated
##
## data: simulationOutput
## ratioObsSim = 0.99006, p-value = 0.68
## alternative hypothesis: two.sided
##
##
## $outliers
##
## DHARMA outlier test based on exact binomial test
##
## data: simulationOutput
## outLow = 2.0000e+00, outHigh = 1.1000e+01, nobs = 9.3200e+02,
## freqH0 = 3.9841e-03, p-value = 0.003148
## alternative hypothesis: two.sided
```

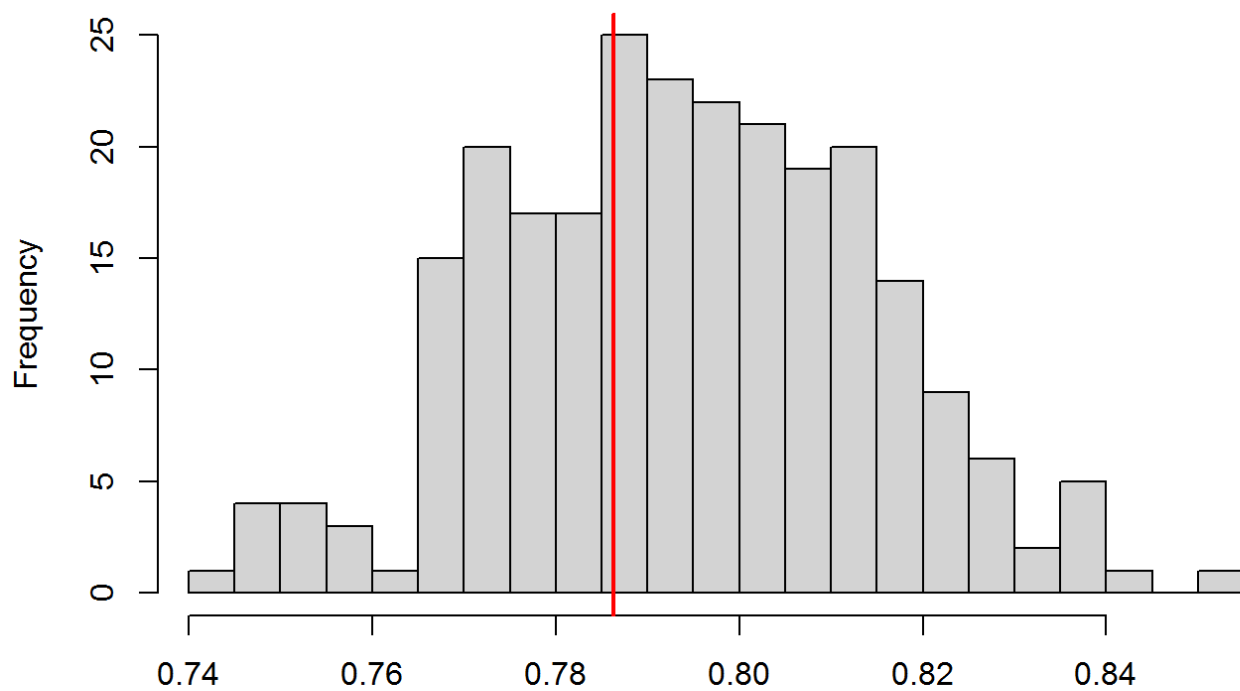
```
## $uniformity
##
## One-sample Kolmogorov-Smirnov test
##
## data: simulationOutput$scaledResiduals
## D = 0.040219, p-value = 0.09807
## alternative hypothesis: two-sided
##
##
## $dispersion
##
## DHARMA nonparametric dispersion test via sd of residuals fitted
## vs. simulated
##
## data: simulationOutput
## ratioObsSim = 0.99006, p-value = 0.68
## alternative hypothesis: two.sided
##
##
## $outliers
##
## DHARMA outlier test based on exact binomial test
##
## data: simulationOutput
## outLow = 2.0000e+00, outHigh = 1.1000e+01, nobs = 9.3200e+02,
## freqH0 = 3.9841e-03, p-value = 0.003148
## alternative hypothesis: two.sided
```

```
plotQQunif(rr, testUniformity = T) ### plot
```



```
testDispersion(rr,plot=T) ### test
```

DHARMA nonparametric dispersion test via sd of residuals fitted vs. simulated

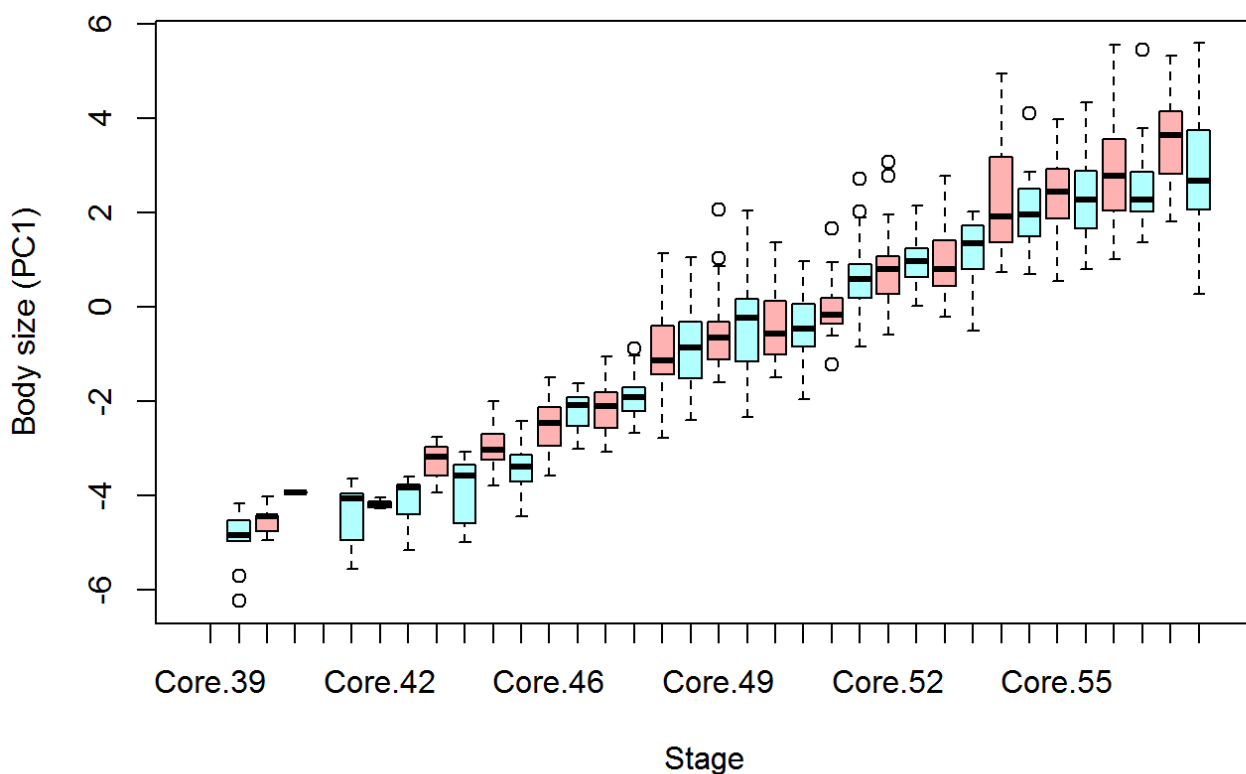


```
##
## DHARMA nonparametric dispersion test via sd of residuals fitted
## vs. simulated
##
## data: simulationOutput
## ratioObsSim = 0.99006, p-value = 0.68
## alternative hypothesis: two.sided
```

```
#note that t-values greater than +/- 1.96 can be considered significant (Luke,
2017)"Evaluating significance in linear mixed-effects models in R"
#Random factors have a small effect on SVL (according to Std.Dev.)
```

```
##Stage has a significant effect on overall bodysize, but not position.
boxplot(PC1~Position*Stage1, xaxs=F,data=taddataset, main= "Tadpole body size per s
tage", xlab="Stage", ylab="Body size (PC1)", col= rainbow(2, alpha = 0.3), notch=
F)
```

Tadpole body size per stage



SVL

```
SVL1<- lmer(pSVL ~ Position*Week1+ (1|Site/Clutchx), data = taddataset)
```

```
## boundary (singular) fit: see ?isSingular
```

```
SVL2<- lmer(pSVL ~ Position + Week1 + (1|Site/Clutchx), data = taddataset)
```



```
## boundary (singular) fit: see ?isSingular
```

```
SVL3<- lmer(pSVL ~ Week1 + (1|Site/Clutchx), data=taddataset)
```

```
## boundary (singular) fit: see ?isSingular
```

```
SVL4<- lmer(pSVL ~ Position + (1|Site/Clutchx), data = taddataset, control=lmerControl(optimizer="nloptwrap",optCtrl=list(algorithm="NLOPT_LN_NELDERMEAD"))) #control for convergence errors
```

```
## boundary (singular) fit: see ?isSingular
```

```
#Position*Stage is removed due to rank deficiency.
```

```
SVL6<- lmer(pSVL ~ Position + Stage1 + (1|Site/Clutchx), data = taddataset, control=lmerControl(optimizer="nloptwrap",optCtrl=list(algorithm="NLOPT_LN_NELDERMEAD"))) #control for convergence errors
```

```
SVL7<- lmer(pSVL ~ Stage1 + (1|Site/Clutchx), data= taddataset)
```

```
SVL8<- lmer(pSVL ~ 1 + (1|Site/Clutchx), data=taddataset)
```

```
AIC(SVL1, SVL2, SVL3, SVL4, SVL6, SVL7, SVL8)
```

```
##      df      AIC
## SVL1 23 1452.2084
## SVL2 14 1428.8605
## SVL3 13 1425.2331
## SVL4  5 2649.5477
## SVL6 22  991.9706
## SVL7 21  986.0528
## SVL8  4 2646.1999
```

```
AICctab(SVL1, SVL2, SVL3, SVL4, SVL6, SVL7, SVL8)
```

```
##      dAICc  df
## SVL7    0.0 21
## SVL6    6.0 22
## SVL3   438.6 13
## SVL2   442.3 14
## SVL1   466.4 23
## SVL8  1659.2  4
## SVL4  1662.5  5
```

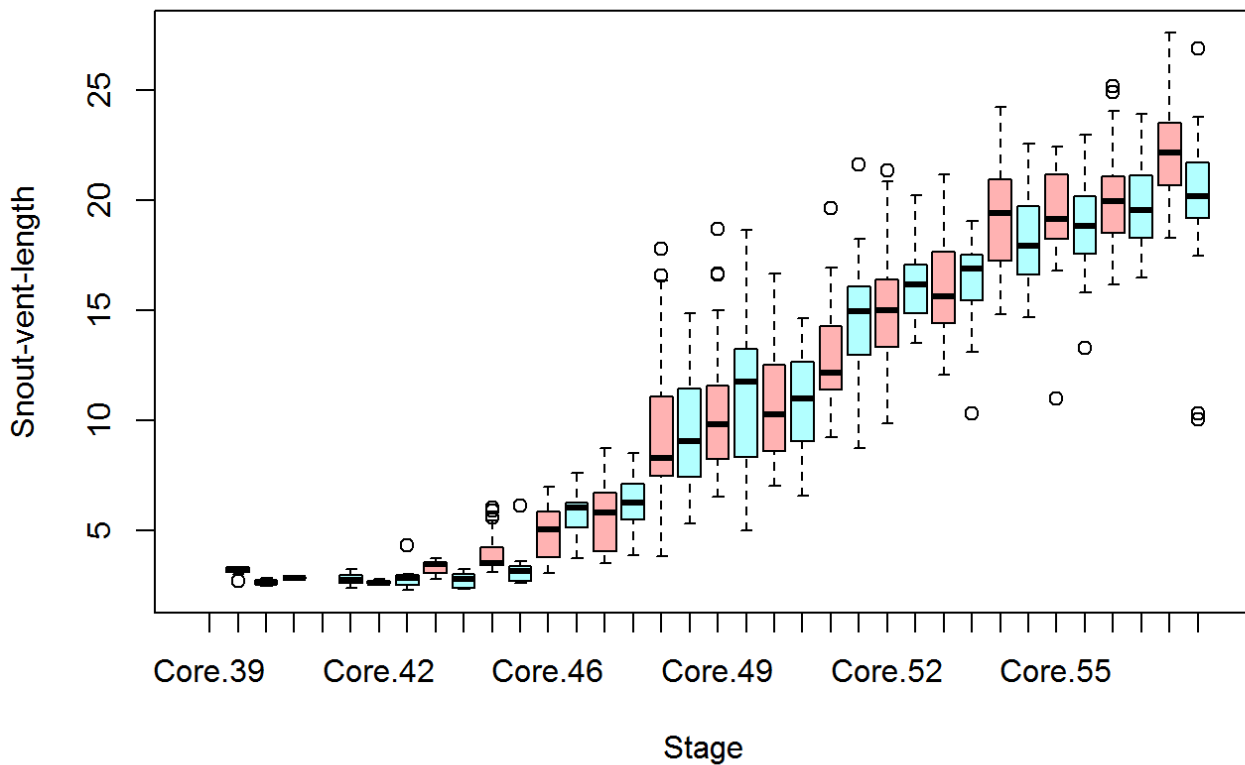
```
summary(SVL7)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: pSVL ~ Stage1 + (1 | Site/Clutchx)
## Data: taddataset
##
## REML criterion at convergence: 944.1
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -4.3966 -0.6158  0.0075  0.5414  3.6104
##
## Random effects:
## Groups          Name          Variance Std.Dev.
## Clutchx:Site (Intercept) 1.070e-02 1.035e-01
## Site          (Intercept) 2.376e-09 4.875e-05
## Residual                1.483e-01 3.851e-01
## Number of obs: 932, groups: Clutchx:Site, 23; Site, 6
##
## Fixed effects:
##              Estimate Std. Error t value
## (Intercept)  -1.6263     0.1272 -12.785
## Stage140      -0.3097     0.2062  -1.502
## Stage141      -0.4132     0.1796  -2.300
## Stage142      -0.3673     0.1649  -2.228
## Stage144      -0.2201     0.1566  -1.405
## Stage145       0.1812     0.1412   1.284
## Stage146       0.5281     0.1380   3.827
## Stage147       0.6803     0.1388   4.900
## Stage148       1.2103     0.1301   9.303
## Stage149       1.3919     0.1299  10.714
## Stage150       1.4233     0.1382  10.302
## Stage151       1.7608     0.1372  12.831
## Stage152       2.0243     0.1311  15.437
## Stage153       2.1313     0.1371  15.551
## Stage154       2.5579     0.1375  18.597
## Stage155       2.6701     0.1348  19.801
## Stage156       2.8650     0.1339  21.389
## Stage157       3.1277     0.1355  23.086
```

```
##
## Correlation matrix not shown by default, as p = 18 > 12.
## Use print(x, correlation=TRUE) or
##      vcov(x)          if you need it
```

```
boxplot(Snoutventlength~Position*Stage, xaxs=F, main= "Tadpole snout-vent-length pe
r stage", xlab="Stage", ylab="Snout-vent-length", col= rainbow(2, alpha = 0.3), not
ch= F)
```

Tadpole snout-vent-length per stage



```
detach(Tadpole)
```

3.2. Climax NF stages 57-65

+3.2.1. Defining variables. Use *Climax meas.csv* to analyse data.

```
Climax=read.csv(file.choose(),header=T)
names(Climax)
```

```
## [1] "num"           "Position"      "Site"
## [4] "Clutch"        "Clutchx"       "Tadpole"
## [7] "Week"          "tadpoleID"     "Mesocosm"
## [10] "Date"          "Fulllength"    "Maximumbodydepth"
## [13] "Snoutventlength" "Taillength"    "Maxtaildepth"
## [16] "Headwidth"     "Femur"         "Femurrelative"
## [19] "Totalleglength" "Stage"         "Stagecategory"
```

```
attach(Climax)
```

```
## The following object is masked _by_ .GlobalEnv:
##
## Tadpole
```

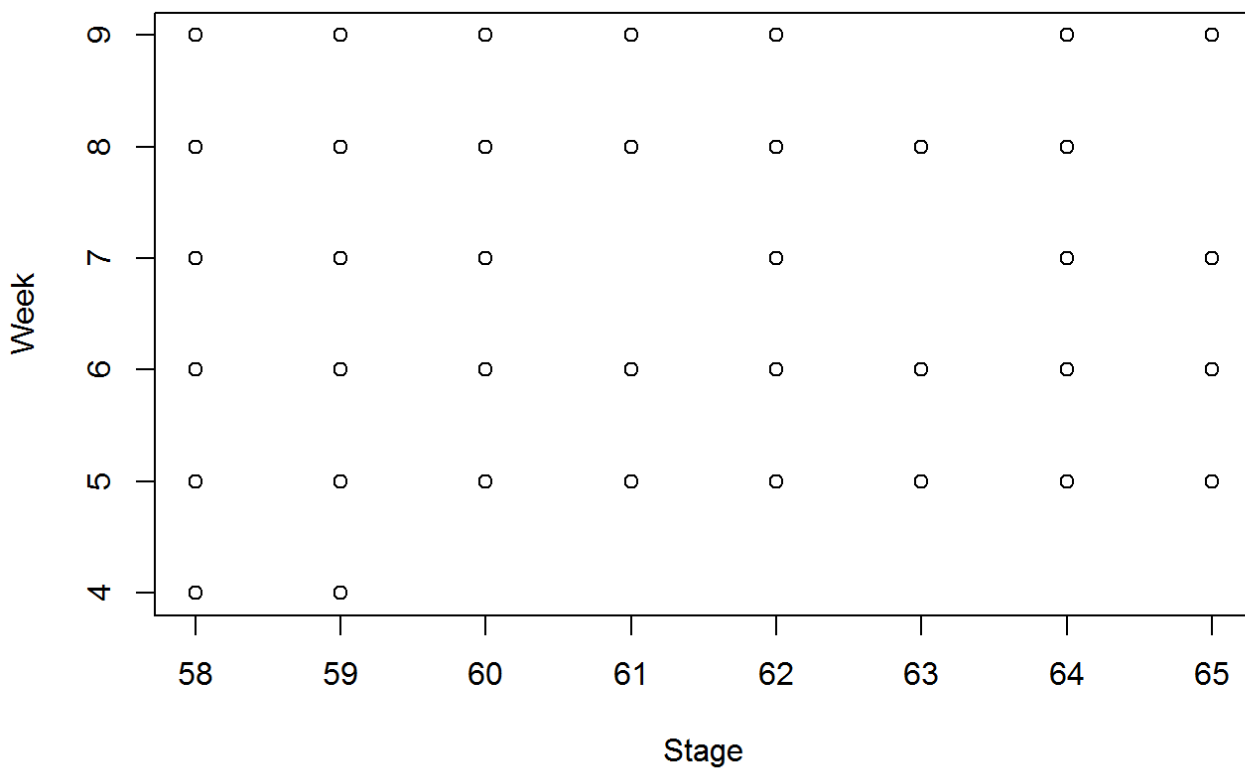
```
summary(Position) #fixed effect
```

```
##      Core Periphery
##      67          90
```

```
cor.test(Week, Stage, method="pearson")
```

```
##
## Pearson's product-moment correlation
##
## data: Week and Stage
## t = -1.0854, df = 155, p-value = 0.2794
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
## -0.24022059 0.07075019
## sample estimates:
##      cor
## -0.08685044
```

```
#for this dataset, week and stage is not correlated so it would be added into the model selection
plot(Week~Stage)
```



```
summary(Week) #fixed effect
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      4.00   6.00   7.00   7.14   8.00   9.00
```

```
Week1 <- as.factor(Week) #as a discrete variable
summary(Week1)
```

```
## 4 5 6 7 8 9
## 5 18 32 31 37 34
```

```
summary(Stage) #fixed effect
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 58.00  59.00  60.00  60.24  61.00  65.00
```

```
Stage1 <- as.factor(Stage) #as a discrete variable
summary(Stage1)
```

```
## 58 59 60 61 62 63 64 65
## 34 39 26 19 15 9 7 8
```

```
summary(Site) #Random effect
```

```
## B C J M S T
## 33 30 30 20 14 30
```

```
summary(Clutchx) #Nested within site
```

```
## B2 B3 B4 C1 C2 C3 C4 J1 J2 J3 J4 M1 M2 M4 S1 S2 S3 S4 T1 T2 T3 T4
## 8 13 12 11 4 7 8 5 10 7 8 4 6 10 4 1 7 2 8 8 5 9
```

```
Climax1<- cbind(Climax, Week1, Stage1)
names(Climax1)
```

```
## [1] "num"           "Position"      "Site"
## [4] "Clutch"        "Clutchx"       "Tadpole"
## [7] "Week"          "tadpoleID"     "Mesocosm"
## [10] "Date"          "Fulllength"    "Maximumbodydepth"
## [13] "Snoutventlength" "Taillength"   "Maxtaildepth"
## [16] "Headwidth"     "Femur"         "Femurrelative"
## [19] "Totalleglength" "Stage"         "Stagecategory"
## [22] "Week1"         "Stage1"
```

3.2.2. Transforming response variables.

Use the *Climax1* datasheet just created to measure normality.

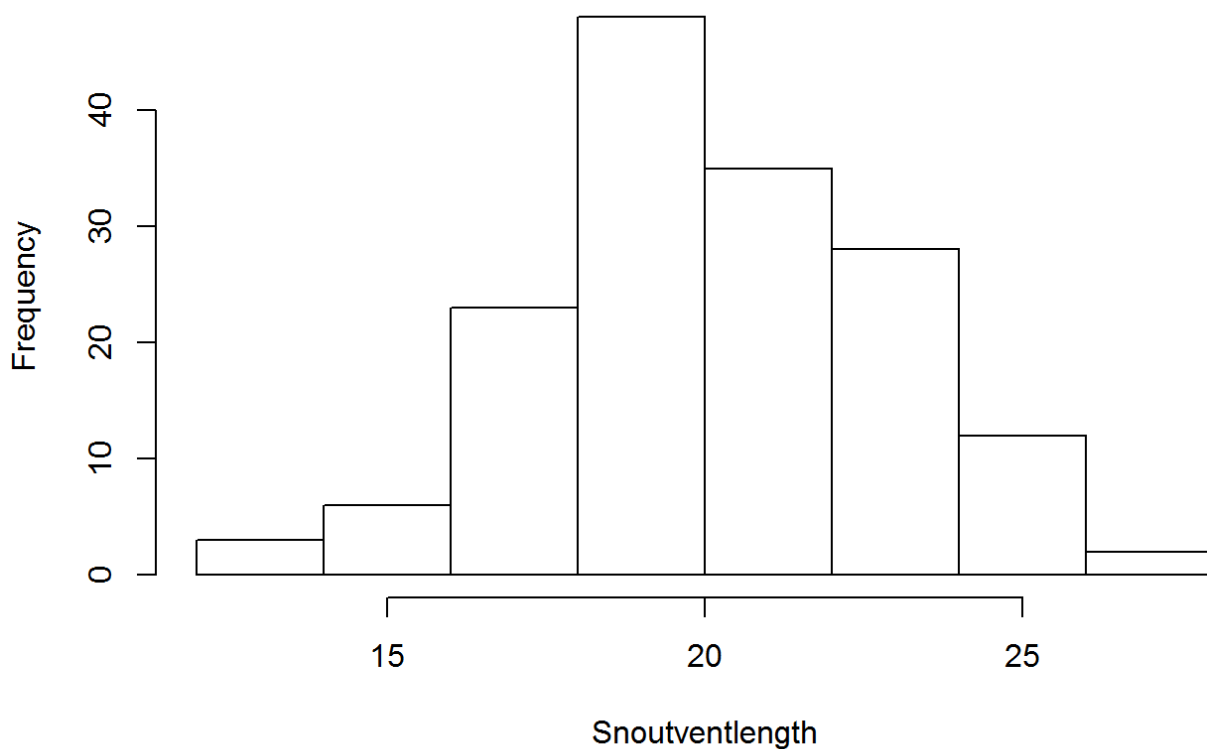
Snoutventlength

```
shapiro.test(Snoutventlength)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  Snoutventlength
## W = 0.99315, p-value = 0.6652
```

```
#p-value = p-value = 0.6652
hist(Snoutventlength)
```

Histogram of Snoutventlength



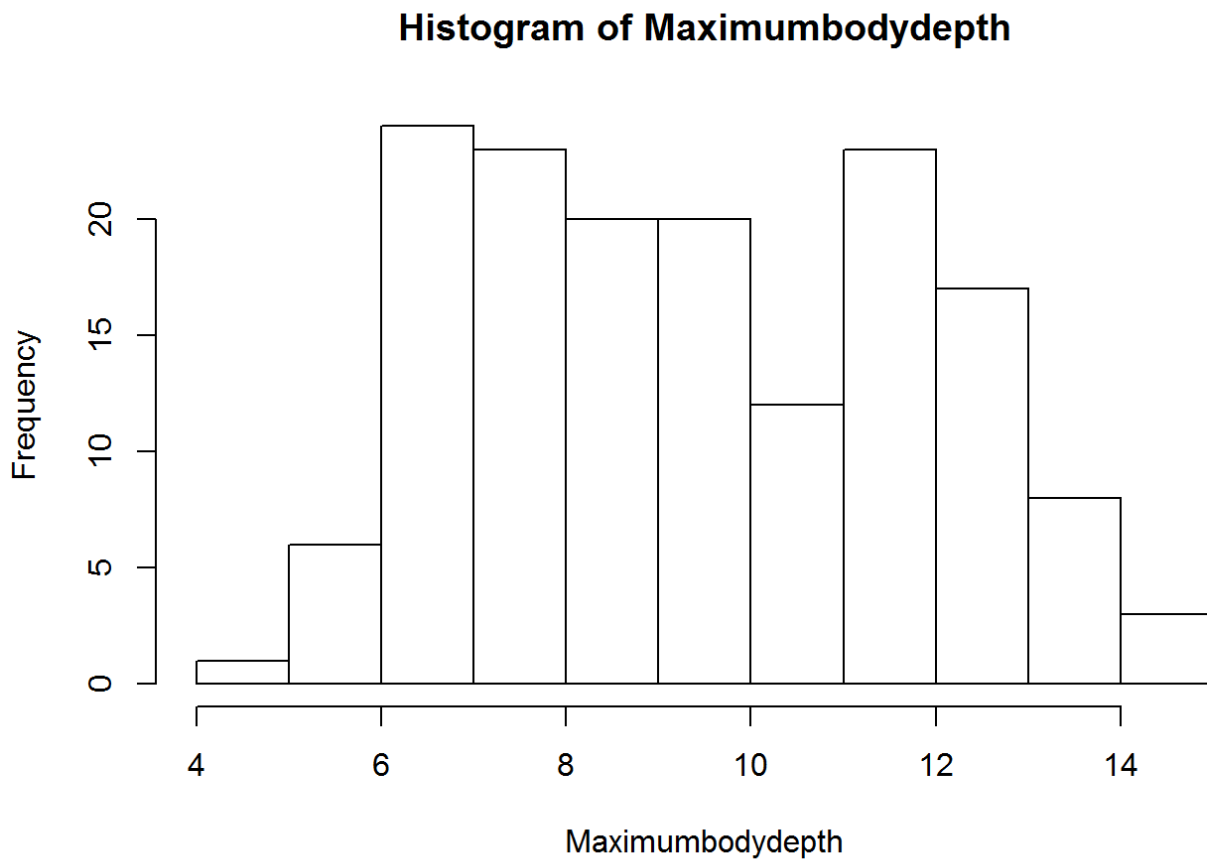
```
#no need to transform
pSVL <- Snoutventlength
```

Maximumbodydepth

```
shapiro.test(Maximumbodydepth)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  Maximumbodydepth
## W = 0.95892, p-value = 0.0001342
```

```
#p-value = 0.0001342
hist(Maximumbodydepth)
```



```
bestNormalize(Maximumbodydepth, standardize= TRUE, allow_orderNorm=TRUE,
              allow_lambert_s = FALSE, allow_lambert_h = FALSE,
              out_of_sample = TRUE, cluster = NULL, k = 10, r = 5,
              loo = FALSE, warn = TRUE, quiet = FALSE)
```

```
## Warning in orderNorm(standardize = TRUE, warn = TRUE, x = c(10.53, 11.307, : Tie
s in data, Normal distribution not guaranteed
```

```
## Warning in get_oos_estimates(x, standardize, method_names, k, r, cluster, :
## fold_size is 15 (< 20), therefore P/df estimates may be off
```

```
## Best Normalizing transformation with 157 Observations
## Estimated Normality Statistics (Pearson P / df, lower => more normal):
## - No transform: 1.5568
## - Box-Cox: 1.5913
## - Log_b(x+a): 1.4388
## - sqrt(x+a): 1.5544
## - exp(x): 9.6536
## - arcsinh(x): 1.4494
## - Yeo-Johnson: 1.5826
## - orderNorm: 1.1535
## Estimation method: Out-of-sample via CV with 10 folds and 5 repeats
##
## Based off these, bestNormalize chose:
## orderNorm Transformation with 157 nonmissing obs and ties
## - 156 unique values
## - Original quantiles:
##      0%      25%      50%      75%     100%
## 4.930  7.237  9.253 11.552 14.450
```

```
#Based off these, bestNormalize chose:
#orderNorm Transformation with 157 nonmissing obs and ties

orderNorm_Bodydepth <- orderNorm(Maximumbodydepth)
```

```
## Warning in orderNorm(Maximumbodydepth): Ties in data, Normal distribution not guaranteed
```

```
pMaximumbodydepth <- predict(orderNorm_Bodydepth)
summary(pMaximumbodydepth)
```

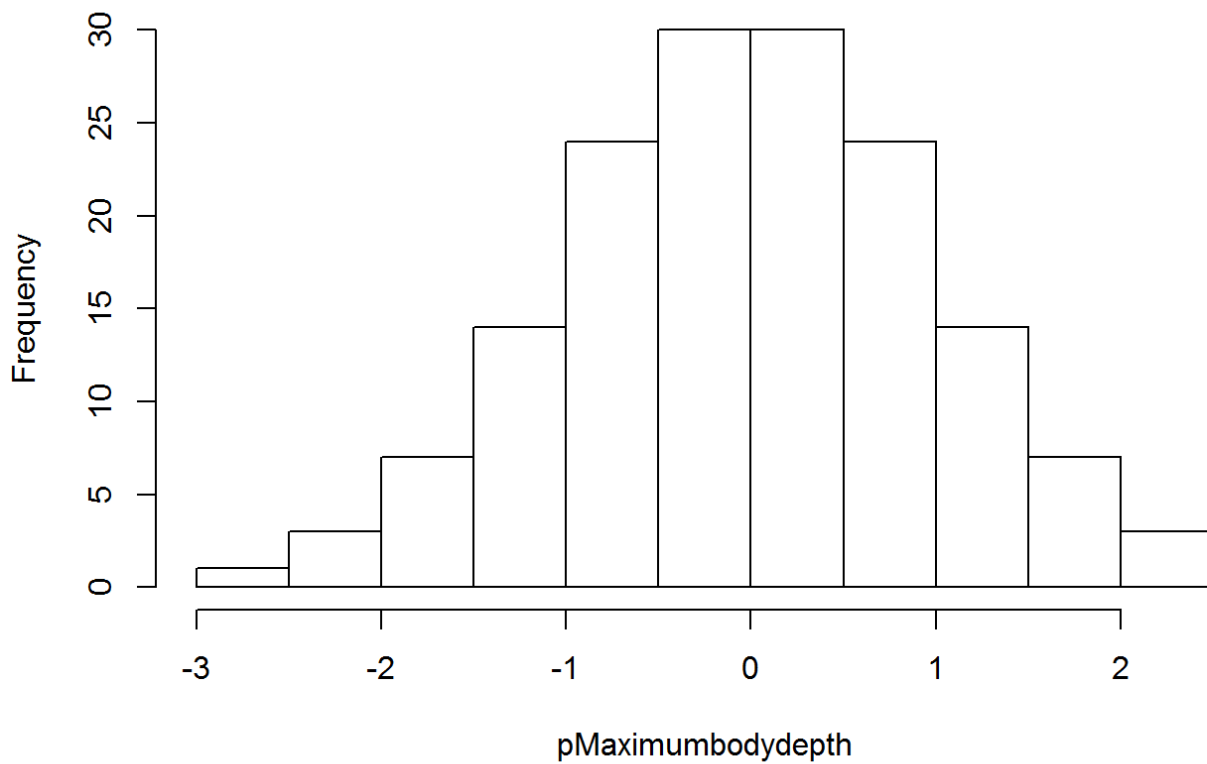
```
##      Min.    1st Qu.    Median      Mean   3rd Qu.      Max.
## -2.730225 -0.674490 -0.007933 -0.017390  0.654704  2.345779
```

```
shapiro.test(pMaximumbodydepth)
```

```
##
## Shapiro-Wilk normality test
##
## data:  pMaximumbodydepth
## W = 0.99821, p-value = 0.9997
```

```
#p-value= p-value = 0.9997 (tadpoles)
hist(pMaximumbodydepth)
```


Histogram of pMaximumbodydepth



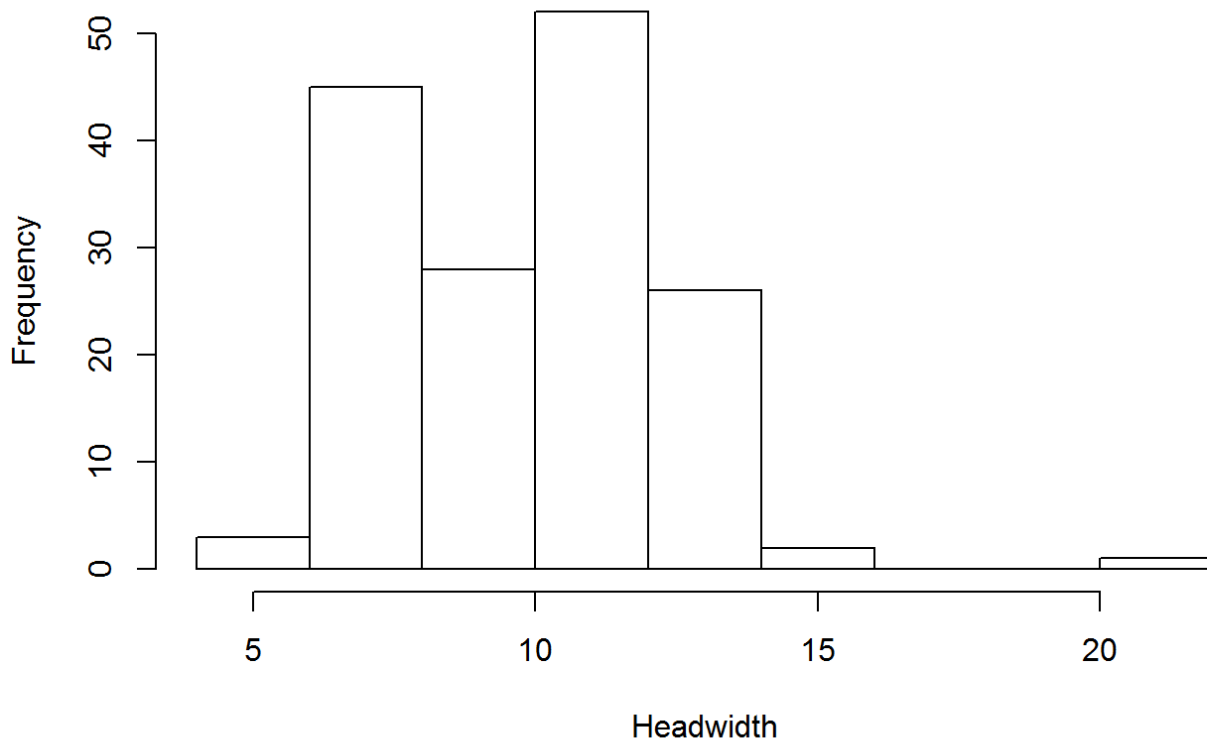
Headwidth

```
shapiro.test(Headwidth)
```

```
##  
## Shapiro-Wilk normality test  
##  
## data: Headwidth  
## W = 0.94566, p-value = 9.36e-06
```

```
#p-value = 9.36e-06  
hist(Headwidth)
```

Histogram of Headwidth



```
bestNormalize(Headwidth, standardize= TRUE, allow_orderNorm=TRUE,
              allow_lambert_s = FALSE, allow_lambert_h = FALSE,
              out_of_sample = TRUE, cluster = NULL, k = 10, r = 5,
              loo = FALSE, warn = TRUE, quiet = FALSE)
```

```
## Warning in orderNorm(standardize = TRUE, warn = TRUE, x = c(10.504, 12.916, : Ti
es in data, Normal distribution not guaranteed
```

```
## Warning in get_oos_estimates(x, standardize, method_names, k, r, cluster, :
## fold_size is 15 (< 20), therefore P/df estimates may be off
```

```
## Best Normalizing transformation with 157 Observations
## Estimated Normality Statistics (Pearson P / df, lower => more normal):
## - No transform: 1.5145
## - Box-Cox: 1.6275
## - Log_b(x+a): 1.5817
## - sqrt(x+a): 1.5018
## - exp(x): 9.0941
## - arcsinh(x): 1.5657
## - Yeo-Johnson: 1.6406
## - orderNorm: 1.1883
## Estimation method: Out-of-sample via CV with 10 folds and 5 repeats
##
## Based off these, bestNormalize chose:
## orderNorm Transformation with 157 nonmissing obs and ties
## - 155 unique values
## - Original quantiles:
##      0%      25%      50%      75%     100%
## 5.195  7.689 10.093 11.316 21.422
```

```
#Based off these, bestNormalize chose:
#orderNorm Transformation with 932 nonmissing obs and ties

orderNorm_Headwidth <- orderNorm(Headwidth)
```

```
## Warning in orderNorm(Headwidth): Ties in data, Normal distribution not guaranteed
```

```
pHeadwidth <- predict(orderNorm_Headwidth)
summary(orderNorm_Headwidth)
```

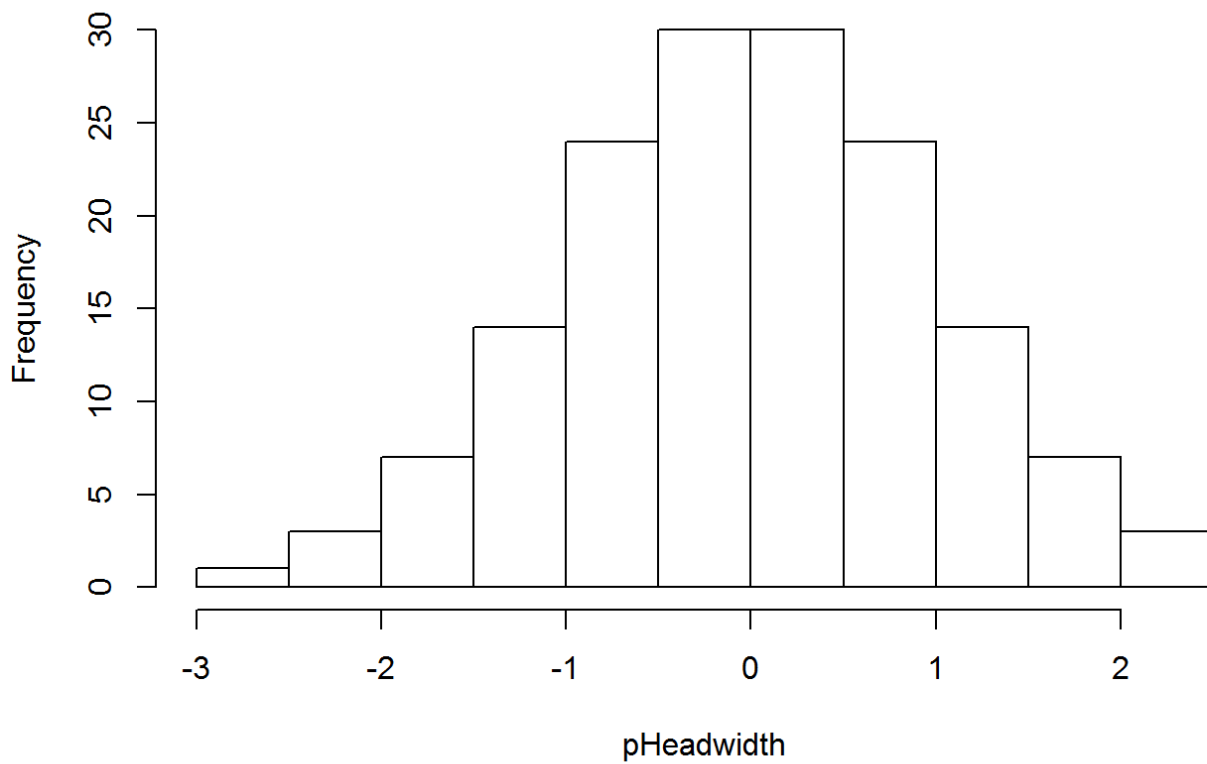
```
##           Length Class  Mode
## x.t         157    -none- numeric
## x           157    -none- numeric
## n              1    -none- numeric
## ties_status   1    -none- numeric
## fit           30     glm      list
## norm_stat      1    -none- numeric
```

```
shapiro.test(pHeadwidth)
```

```
##
## Shapiro-Wilk normality test
##
## data:  pHeadwidth
## W = 0.9982, p-value = 0.9997
```

```
#p-value = 0.9997
hist(pHeadwidth)
```

Histogram of pHeadwidth



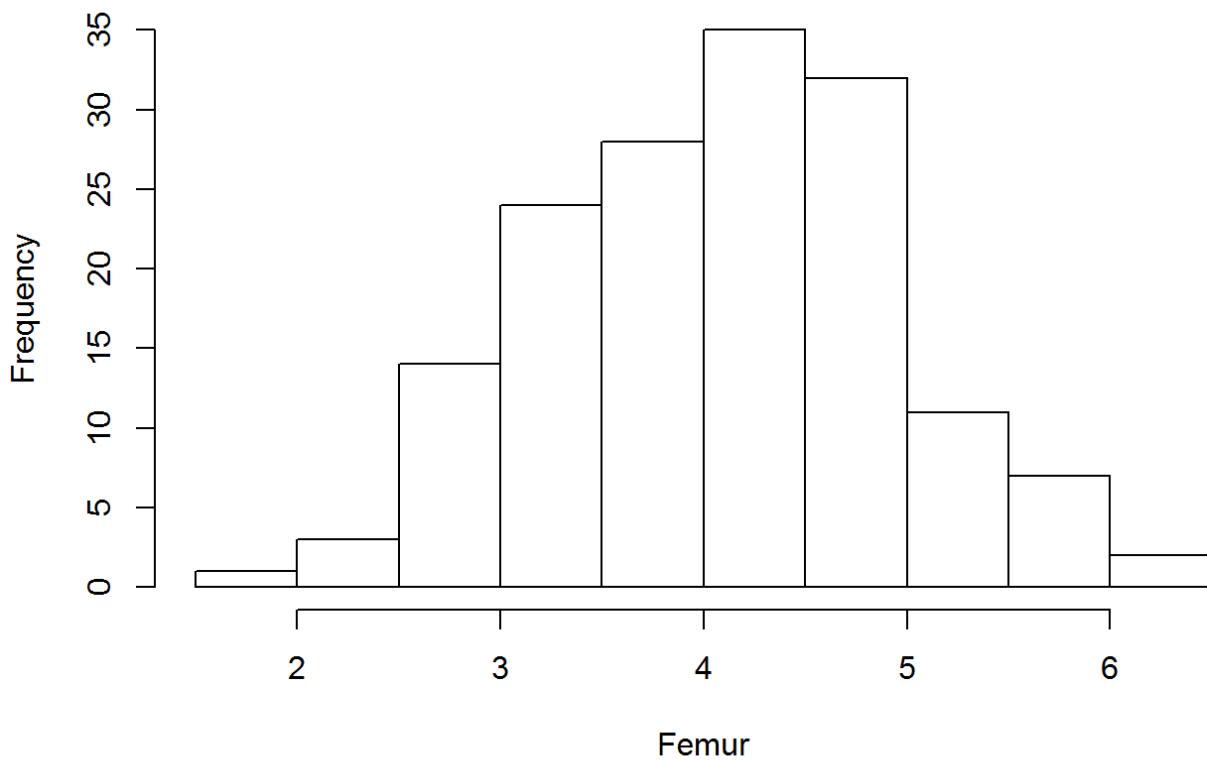
Femur

```
shapiro.test(Femur)
```

```
##  
##  Shapiro-Wilk normality test  
##  
## data:  Femur  
## W = 0.991, p-value = 0.4228
```

```
#p-value = 0.4228  
hist(Femur)
```

Histogram of Femur



```
#no transformation
```

```
pFemur<- Femur
```

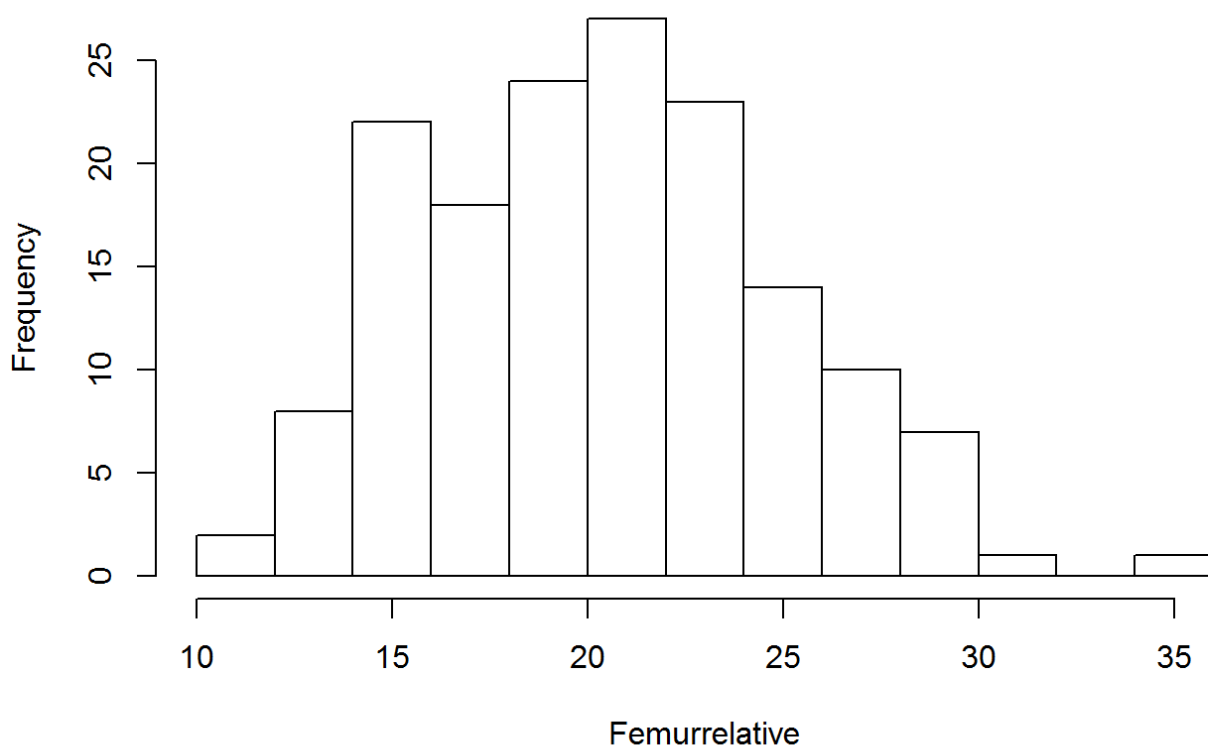
Relative Femur length

```
shapiro.test(Femurrelative)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  Femurrelative
## W = 0.98851, p-value = 0.2271
```

```
#p-value = 0.2271
hist(Femurrelative)
```

Histogram of Femurrelative



```
pFemurrel <- Femurrelative
```

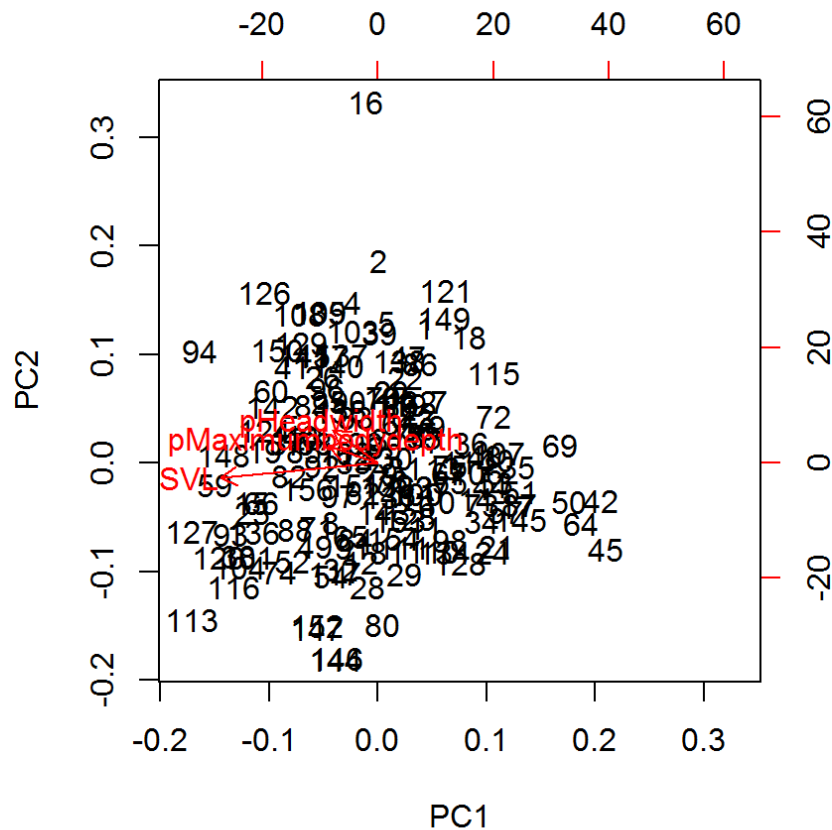
3.2.3. Performing PCA for climax variables.

```
climaxpca<- cbind(pHeadwidth, pSVL, pMaximumbodydepth)
```

```
climaxpca1=prcomp(climaxpca)
summary(climaxpca1)
```

```
## Importance of components:
##              PC1      PC2      PC3
## Standard deviation  2.9496 0.68261 0.44168
## Proportion of Variance 0.9294 0.04977 0.02084
## Cumulative Proportion 0.9294 0.97916 1.00000
```

```
#proportion of variance for PC1- 0.9294
biplot(climaxpca1)
```



```
climaxpcalscores<-predict(climaxpca)

climaxdataset <- cbind(Climaxl, climaxpca, climaxpcalscores)
names(climaxdataset)
```

```
## [1] "num"           "Position"      "Site"
## [4] "Clutch"        "Clutchx"       "Tadpole"
## [7] "Week"          "tadpoleID"     "Mesocosm"
## [10] "Date"          "Fulllength"    "Maximumbodydepth"
## [13] "Snoutventlength" "Taillength"    "Maxtaildepth"
## [16] "Headwidth"     "Femur"         "Femurrelative"
## [19] "Totalleglength" "Stage"         "Stagecategory"
## [22] "Week1"         "Stage1"        "pHeadwidth"
## [25] "pSVL"          "pMaximumbodydepth" "PC1"
## [28] "PC2"           "PC3"
```

3.2.4. Model selection for variables.

PC1

Use the *climaxdataset* for the analysis.

```
#For week and stage as discrete variables
PC1b<- lmer(PC1 ~ Position * Week1+ (1|Site/Clutchx), data= climaxdataset)
PC2b<- lmer(PC1 ~ Position + Week1 + (1|Site/Clutchx), data = climaxdataset)
```

```
## boundary (singular) fit: see ?isSingular
```

```
PC3b<- lmer(PC1 ~ Week1 + (1|Site/Clutchx), data=climaxdataset)
```

```
## boundary (singular) fit: see ?isSingular
```

```
PC4b<- lmer(PC1 ~ Position + (1|Site/Clutchx), data = climaxdataset)
```

```
## boundary (singular) fit: see ?isSingular
```

```
PC5b<- lmer(PC1 ~ Position*Stage1 + (1|Site/Clutchx), data = climaxdataset)
```

```
## boundary (singular) fit: see ?isSingular
```

```
PC6b<- lmer(PC1 ~ Position + Stage1 + (1|Site/Clutchx), data = climaxdataset)
```

```
## boundary (singular) fit: see ?isSingular
```

```
PC7b<- lmer(PC1 ~ Stage1 + (1|Site/Clutchx), data= climaxdataset)
```

```
## boundary (singular) fit: see ?isSingular
```

```
PC8b<- lmer(PC1 ~ 1 + (1|Site/Clutchx), data=climaxdataset)
```

```
## boundary (singular) fit: see ?isSingular
```

```
PC9b<- lmer(PC1 ~ Position + Stage1 + Week1 + (1|Site/Clutchx), data= climaxdataset)
```

```
## boundary (singular) fit: see ?isSingular
```

```
#Position + Stage*Week was removed due to rank deficiency.
```

```
#Stage*Week was removed due to rank deficiency.
```

```
PC12b<- lmer(PC1 ~ Stage1 + Week1 + (1|Site/Clutchx), data= climaxdataset)
```

```
## boundary (singular) fit: see ?isSingular
```

```
AIC(PC1b, PC2b, PC3b, PC4b, PC5b, PC6b, PC7b, PC8b, PC9b, PC12b)
```



```
##      df      AIC
## PC1b  15 768.9409
## PC2b  10 777.2603
## PC3b   9 776.7658
## PC4b   5 792.1712
## PC5b  19 699.6567
## PC6b  12 706.4858
## PC7b  11 705.5416
## PC8b   4 791.0348
## PC9b  17 688.3334
## PC12b 16 688.0042
```

```
AICctab(PC1b, PC2b, PC3b, PC4b, PC5b, PC6b, PC7b, PC8b, PC9b, PC12b)
```

```
##      dAICc df
## PC12b   0.0 16
## PC9b    0.8 17
## PC5b   13.3 19
## PC7b   15.5 11
## PC6b   16.8 12
## PC1b   80.5 15
## PC3b   86.1  9
## PC2b   86.9 10
## PC8b   99.4  4
## PC4b  100.7  5
```

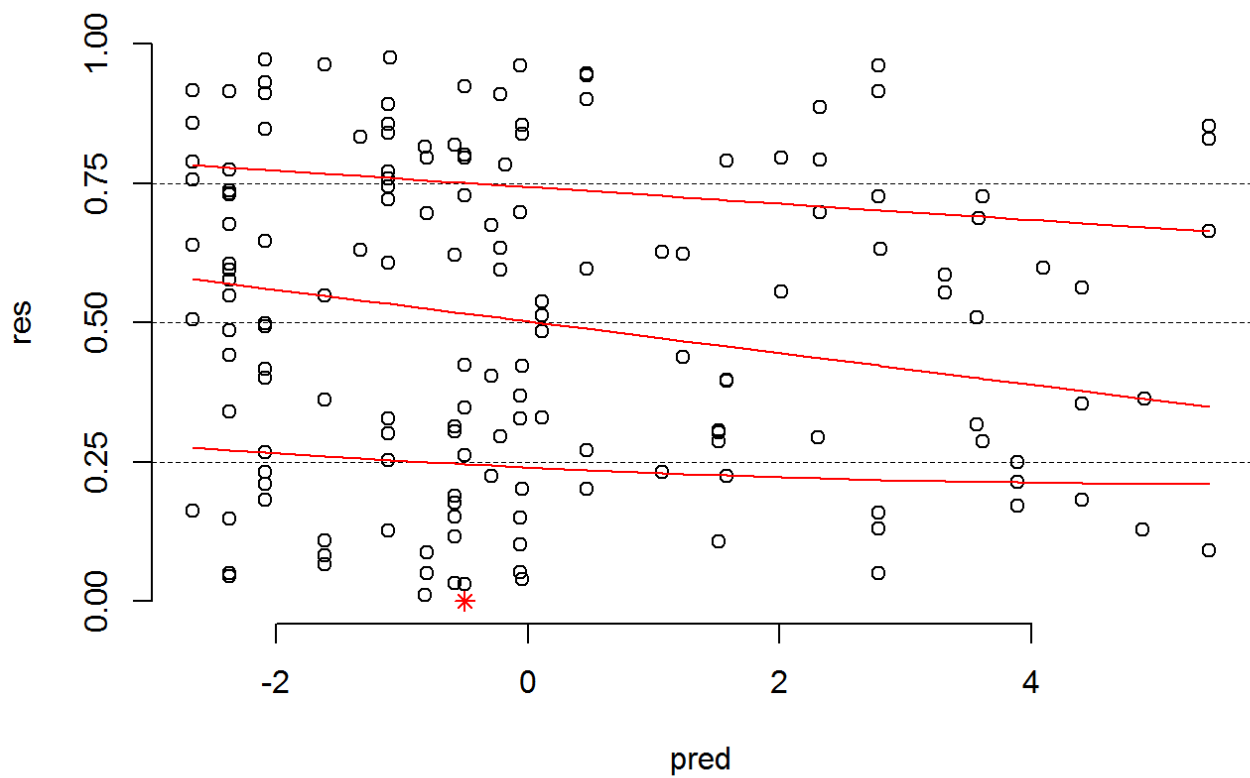
```
#PC11b was selected as the model with the lowest AIC
summary(PC12b)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: PC1 ~ Stage1 + Week1 + (1 | Site/Clutchx)
## Data: climaxdataset
##
## REML criterion at convergence: 656
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -2.20280 -0.64132  0.08944  0.66787  2.31232
##
## Random effects:
## Groups      Name                Variance Std.Dev.
## Clutchx:Site (Intercept) 1.579      1.256
## Site         (Intercept) 0.000      0.000
## Residual                3.607      1.899
## Number of obs: 157, groups: Clutchx:Site, 22; Site, 6
##
## Fixed effects:
##              Estimate Std. Error t value
## (Intercept)   0.1113    0.9284    0.120
## Stage159      -0.2912    0.5048   -0.577
## Stage160       0.7566    0.5625    1.345
## Stage161       1.8697    0.5829    3.207
## Stage162       3.6046    0.6468    5.573
## Stage163       4.6919    0.7869    5.962
## Stage164       4.3874    0.8441    5.198
## Stage165       5.7040    0.8090    7.051
## Week15        -0.9142    1.0196   -0.897
## Week16        -0.4018    0.9967   -0.403
## Week17        -0.9302    1.0083   -0.923
## Week18        -2.4851    0.9811   -2.533
## Week19        -2.1959    0.9736   -2.255
```

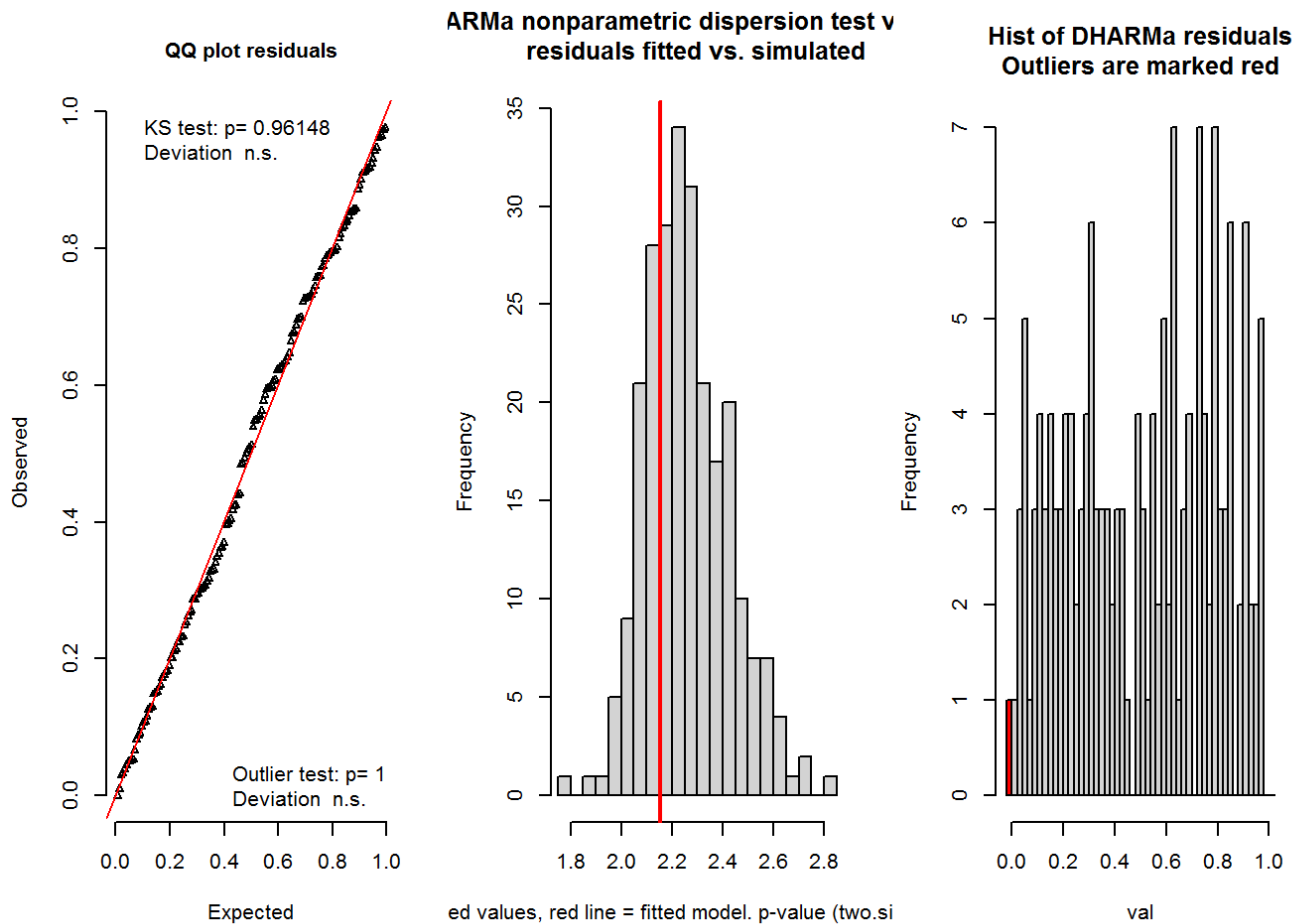
```
##
## Correlation matrix not shown by default, as p = 13 > 12.
## Use print(x, correlation=TRUE) or
##      vcov(x)          if you need it
```

```
## convergence code: 0
## boundary (singular) fit: see ?isSingular
```

```
rr=simulateResiduals(PC12b) ### simulate residuals
plotResiduals(rr) ### qqplot simulated vs observed
```



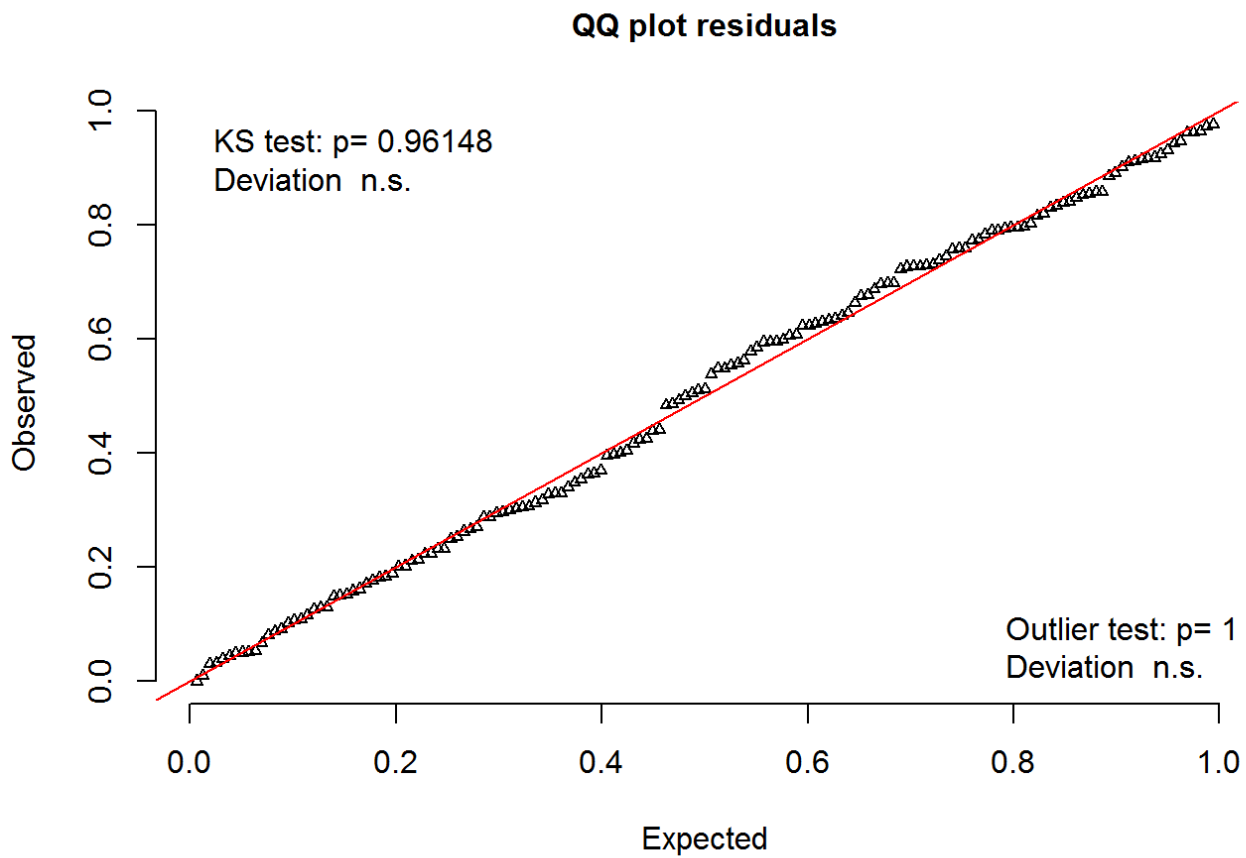
```
testResiduals(rr)###test qqplot
```



```
## $uniformity
##
## One-sample Kolmogorov-Smirnov test
##
## data: simulationOutput$scaledResiduals
## D = 0.040206, p-value = 0.9615
## alternative hypothesis: two-sided
##
##
## $dispersion
##
## DHARMA nonparametric dispersion test via sd of residuals fitted
## vs. simulated
##
## data: simulationOutput
## ratioObsSim = 0.95156, p-value = 0.552
## alternative hypothesis: two.sided
##
##
## $outliers
##
## DHARMA outlier test based on exact binomial test
##
## data: simulationOutput
## outLow = 1.0000e+00, outHigh = 0.0000e+00, nobs = 1.5700e+02,
## freqH0 = 3.9841e-03, p-value = 1
## alternative hypothesis: two.sided
```

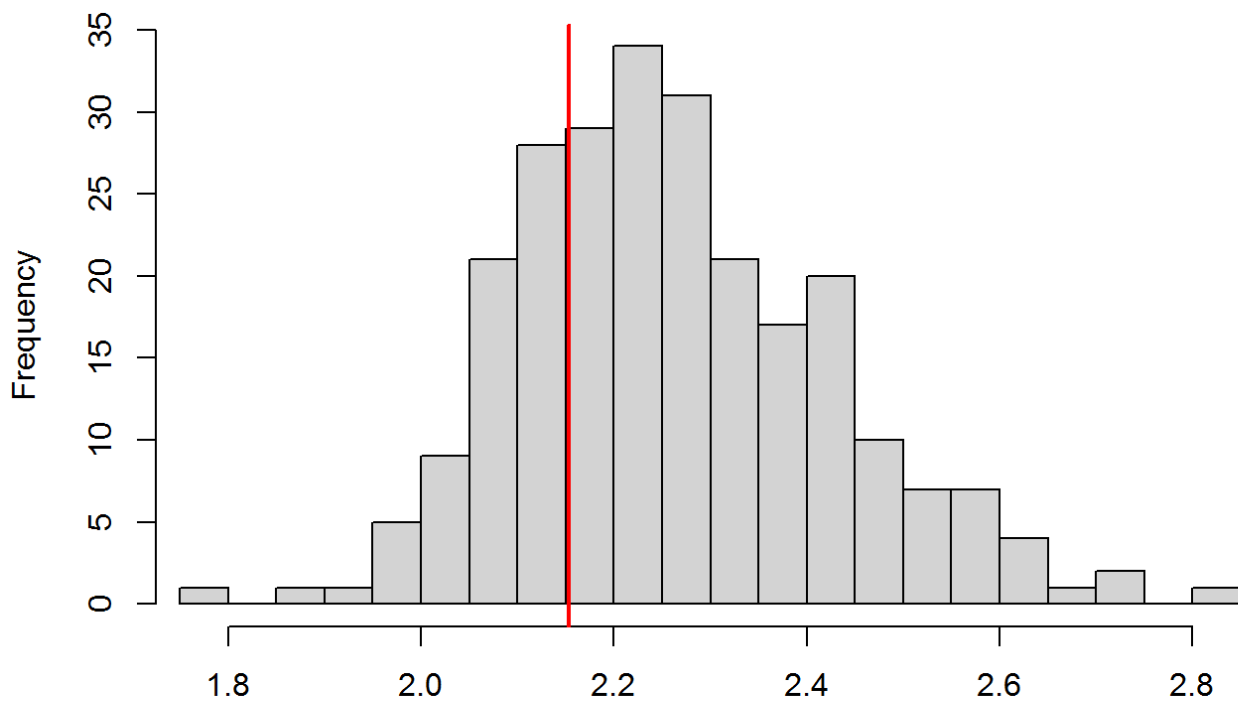
```
## $uniformity
##
## One-sample Kolmogorov-Smirnov test
##
## data: simulationOutput$scaledResiduals
## D = 0.040206, p-value = 0.9615
## alternative hypothesis: two-sided
##
##
## $dispersion
##
## DHARMA nonparametric dispersion test via sd of residuals fitted
## vs. simulated
##
## data: simulationOutput
## ratioObsSim = 0.95156, p-value = 0.552
## alternative hypothesis: two.sided
##
##
## $outliers
##
## DHARMA outlier test based on exact binomial test
##
## data: simulationOutput
## outLow = 1.0000e+00, outHigh = 0.0000e+00, nobs = 1.5700e+02,
## freqH0 = 3.9841e-03, p-value = 1
## alternative hypothesis: two.sided
```

```
plotQQunif(rr, testUniformity = T) ### plot
```



```
testDispersion(rr,plot=T) ### test
```

DHARMA nonparametric dispersion test via sd of residuals fitted vs. simulated

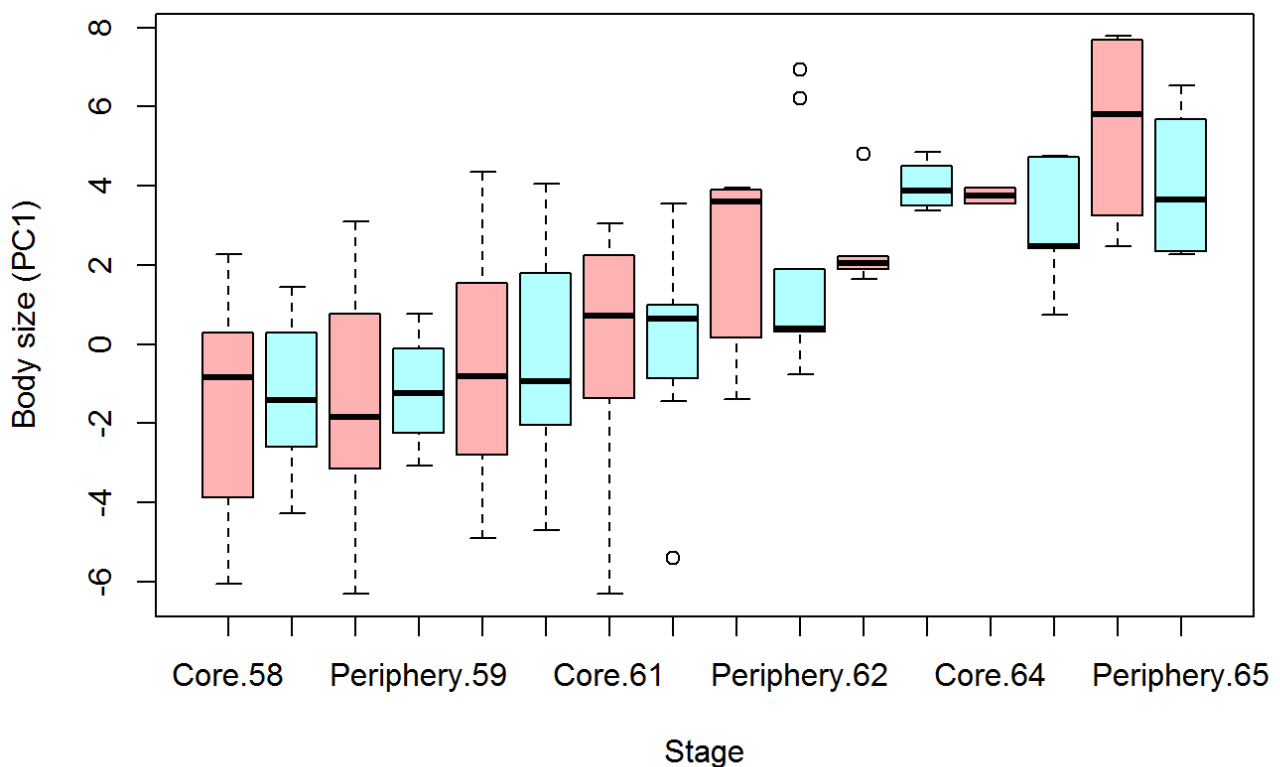


```
##
## DHARMA nonparametric dispersion test via sd of residuals fitted
## vs. simulated
##
## data: simulationOutput
## ratioObsSim = 0.95156, p-value = 0.552
## alternative hypothesis: two.sided
```

#note that t-values greater than +/- 1.96 can be considered significant (Luke, 2017)"Evaluating significance in linear mixed-effects models in R"
#Random factors have a small effect on PC1 (according to Std.Dev.)

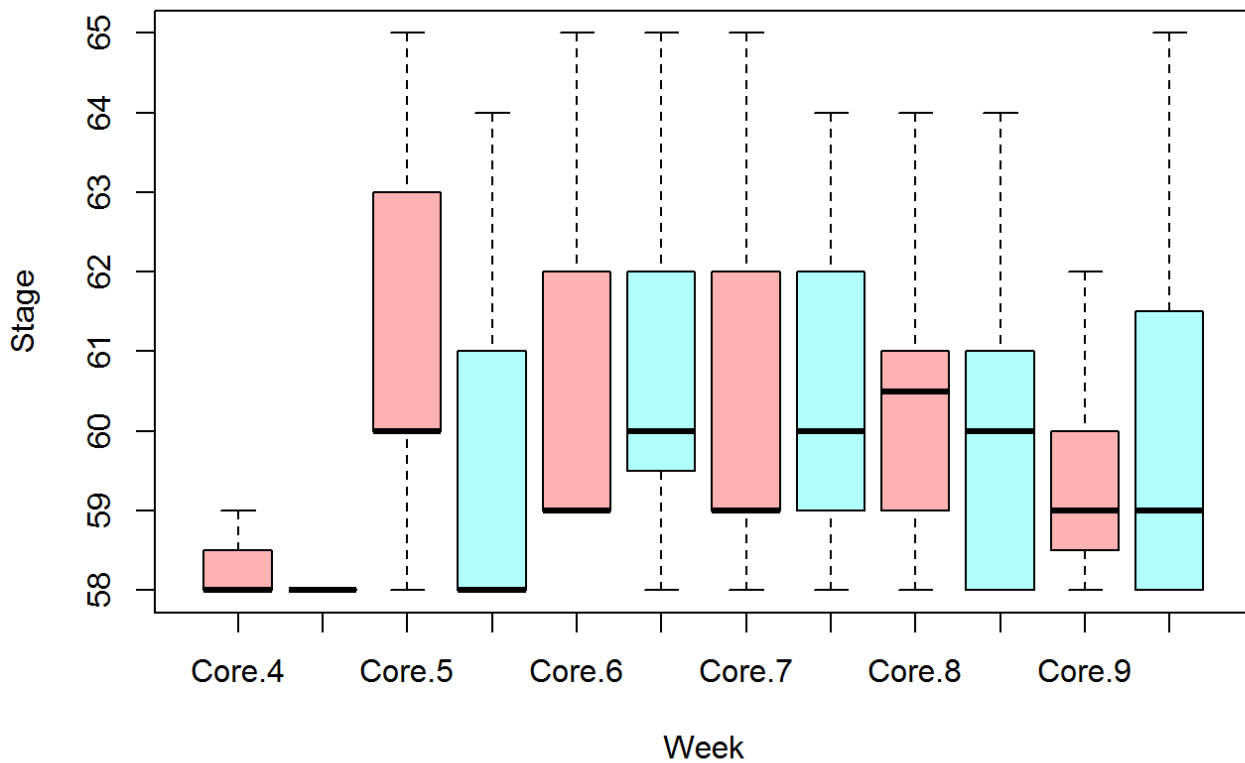
*##Stage*Week has a significant effect on overall bodysize, but not position.*
 boxplot(PC1~Position*Stage1, xaxs=F,data=climaxdataset, main= "Climax body size per stage", xlab="Stage", ylab="Body size (PC1)", col= rainbow(2, alpha = 0.3), notch= F)

Climax body size per stage



```
boxplot(Stage~Position*Week, xaxs=F, main= "Climax stages per week", xlab="Week", y
lab="Stage", col= rainbow(2, alpha = 0.3), notch= F)
```

Climax stages per week



SVL

```
SVL1b<- lmer(pSVL ~ Position*Week1+ (1|Site/Clutchx), data = climaxdataset)
```

```
## boundary (singular) fit: see ?isSingular
```

```
SVL2b<- lmer(pSVL ~ Position + Week1 + (1|Site/Clutchx), data = climaxdataset)
```

```
## boundary (singular) fit: see ?isSingular
```

```
SVL3b<- lmer(pSVL ~ Week1 + (1|Site/Clutchx), data=climaxdataset)
```

```
## boundary (singular) fit: see ?isSingular
```

```
SVL4b<- lmer(pSVL ~ Position + (1|Site/Clutchx), data = climaxdataset)
```

```
## boundary (singular) fit: see ?isSingular
```

```
SVL5b<- lmer(pSVL ~ Position*Stage1 + (1|Site/Clutchx), data = climaxdataset)
```

```
## boundary (singular) fit: see ?isSingular
```



```
SVL6b<- lmer(pSVL ~ Position + Stage1 + (1|Site/Clutchx), data = climaxdataset)
```

```
## boundary (singular) fit: see ?isSingular
```

```
SVL7b<- lmer(pSVL ~ Stage1 + (1|Site/Clutchx), data= climaxdataset)
```

```
## boundary (singular) fit: see ?isSingular
```

```
SVL8b<- lmer(pSVL ~ 1 + (1|Site/Clutchx), data=climaxdataset)
```

```
## boundary (singular) fit: see ?isSingular
```

```
SVL9b<- lmer(pSVL ~ Position + Stage1 + Week1 + (1|Site/Clutchx), data= climaxdataset, control=lmerControl(optimizer="nloptwrap",optCtrl=list(algorithm="NLOPT_LN_NELDERMEAD")) #Control for convergence errors
```

```
## boundary (singular) fit: see ?isSingular
```

```
#Position + Stage1*Week1 was dropped due to rank deficiency
```

```
#Stage1*Week1 was dropped due to rank deficiency
```

```
SVL12b<- lmer(pSVL ~ Stage1+Week1 + (1|Site/Clutchx), data= climaxdataset)
```

```
AIC(SVL1b, SVL2b, SVL3b, SVL4b, SVL5b, SVL6b, SVL7b, SVL8b, SVL9b, SVL12b)
```

```
##          df          AIC
## SVL1b    15 747.3183
## SVL2b    10 754.7164
## SVL3b     9 754.1562
## SVL4b     5 767.9722
## SVL5b    19 689.7674
## SVL6b    12 695.2638
## SVL7b    11 694.2165
## SVL8b     4 766.7053
## SVL9b    17 679.2240
## SVL12b   16 678.8344
```

```
AICctab(SVL1b, SVL2b, SVL3b, SVL4b, SVL5b, SVL6b, SVL7b, SVL8b, SVL9b, SVL12b)
```

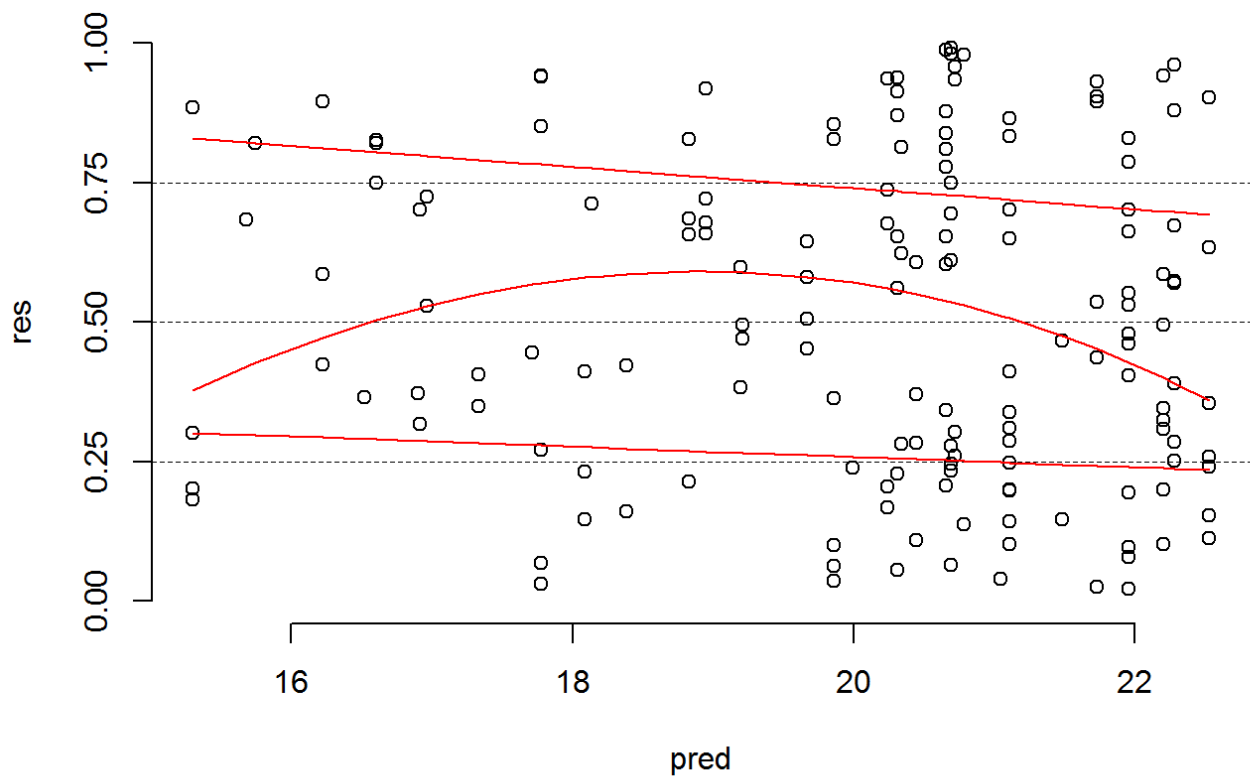
```
##          dAICc df
## SVL12b   0.0  16
## SVL9b    0.9  17
## SVL5b   12.6  19
## SVL7b   13.3  11
## SVL6b   14.7  12
## SVL1b   68.0  15
## SVL3b   72.7   9
## SVL2b   73.5  10
## SVL8b   84.2   4
## SVL4b   85.6   5
```

```
summary(SVL12b)
```

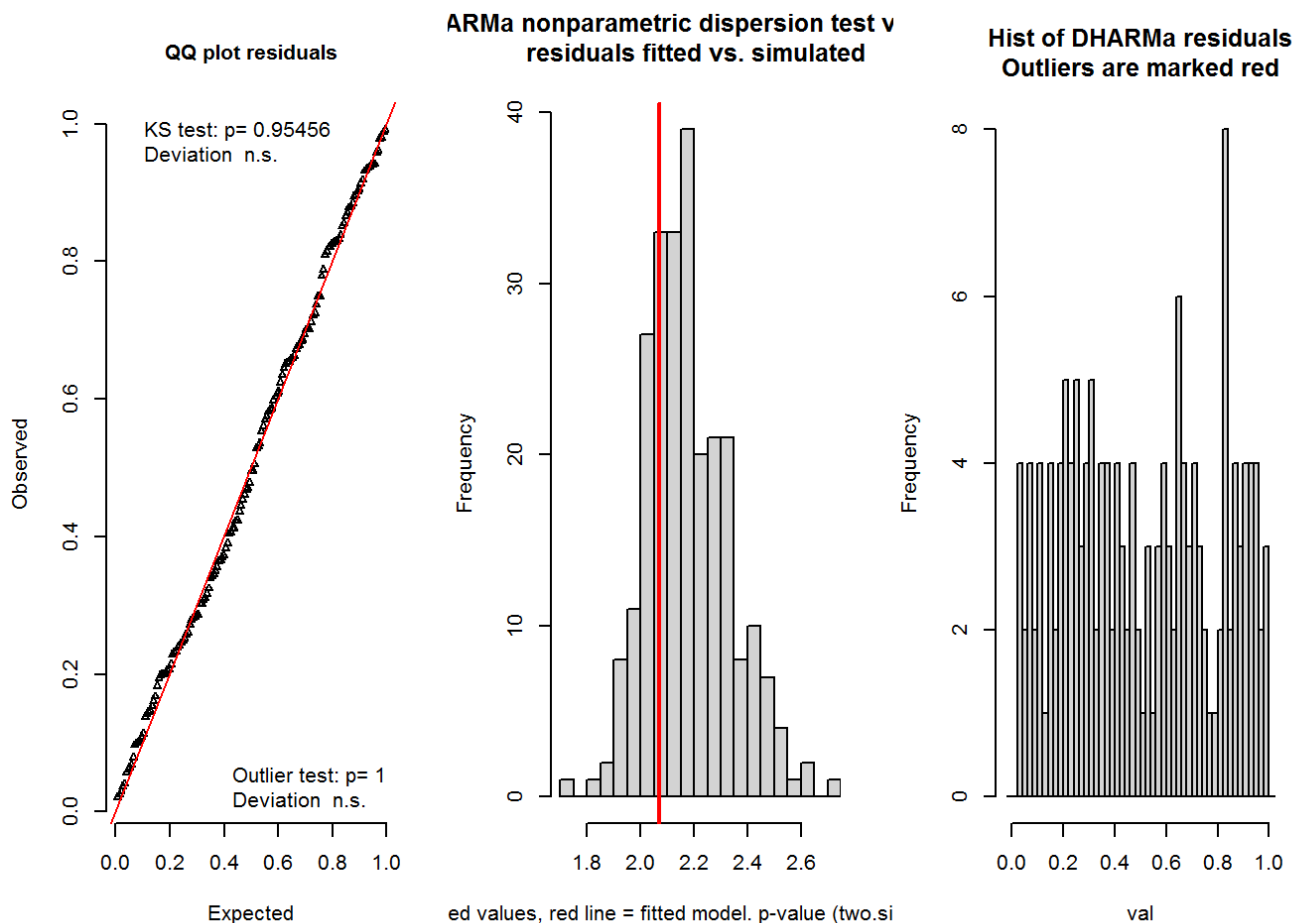
```
## Linear mixed model fit by REML ['lmerMod']
## Formula: pSVL ~ Stage1 + Week1 + (1 | Site/Clutchx)
##      Data: climaxdataset
##
## REML criterion at convergence: 646.8
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -2.1023 -0.6595 -0.1032  0.6396  2.2244
##
## Random effects:
##   Groups      Name      Variance Std.Dev.
## Clutchx:Site (Intercept) 1.393e+00 1.1800814
## Site          (Intercept) 6.208e-08 0.0002492
## Residual                3.408e+00 1.8460524
## Number of obs: 157, groups: Clutchx:Site, 22; Site, 6
##
## Fixed effects:
##              Estimate Std. Error t value
## (Intercept)  19.6708    0.8993  21.874
## Stage159      0.3237    0.4901   0.660
## Stage160     -0.4768    0.5460  -0.873
## Stage161     -1.5097    0.5664  -2.666
## Stage162     -3.0053    0.6280  -4.785
## Stage163     -4.1129    0.7643  -5.381
## Stage164     -3.8169    0.8198  -4.656
## Stage165     -5.0380    0.7858  -6.412
## Week15        1.0471    0.9908   1.057
## Week16         0.6651    0.9682   0.687
## Week17         1.1131    0.9794   1.136
## Week18         2.5317    0.9530   2.657
## Week19         2.2855    0.9456   2.417
```

```
##
## Correlation matrix not shown by default, as p = 13 > 12.
## Use print(x, correlation=TRUE) or
##      vcov(x)          if you need it
```

```
rr=simulateResiduals(SVL12b) ### simulate residuals  
plotResiduals(rr) ### qqplot simulated vs observed
```



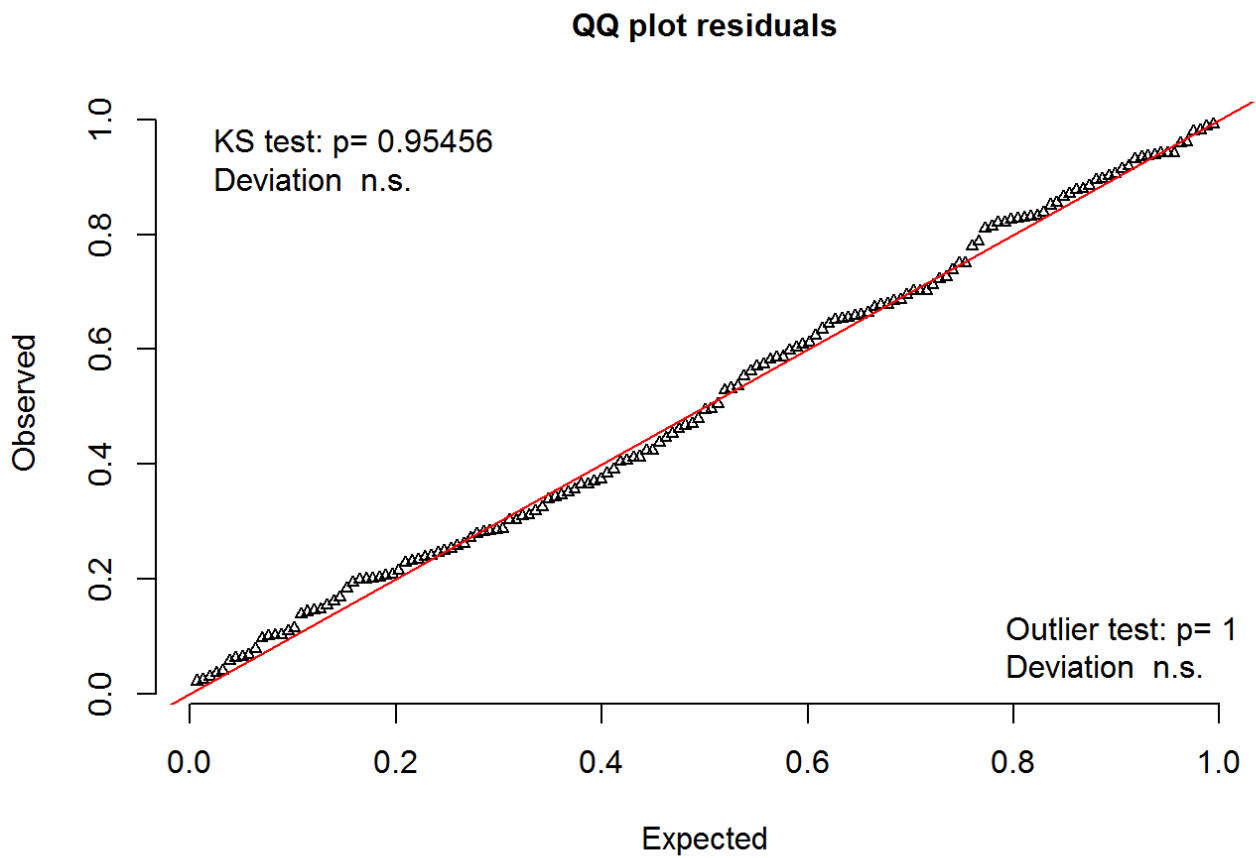
```
testResiduals(rr)###test qqplot
```



```
## $uniformity
##
## One-sample Kolmogorov-Smirnov test
##
## data: simulationOutput$scaledResiduals
## D = 0.040992, p-value = 0.9546
## alternative hypothesis: two-sided
##
##
## $dispersion
##
## DHARMA nonparametric dispersion test via sd of residuals fitted
## vs. simulated
##
## data: simulationOutput
## ratioObsSim = 0.9508, p-value = 0.528
## alternative hypothesis: two.sided
##
##
## $outliers
##
## DHARMA outlier test based on exact binomial test
##
## data: simulationOutput
## outLow = 0.0000e+00, outHigh = 0.0000e+00, nobs = 1.5700e+02,
## freqH0 = 3.9841e-03, p-value = 1
## alternative hypothesis: two.sided
```

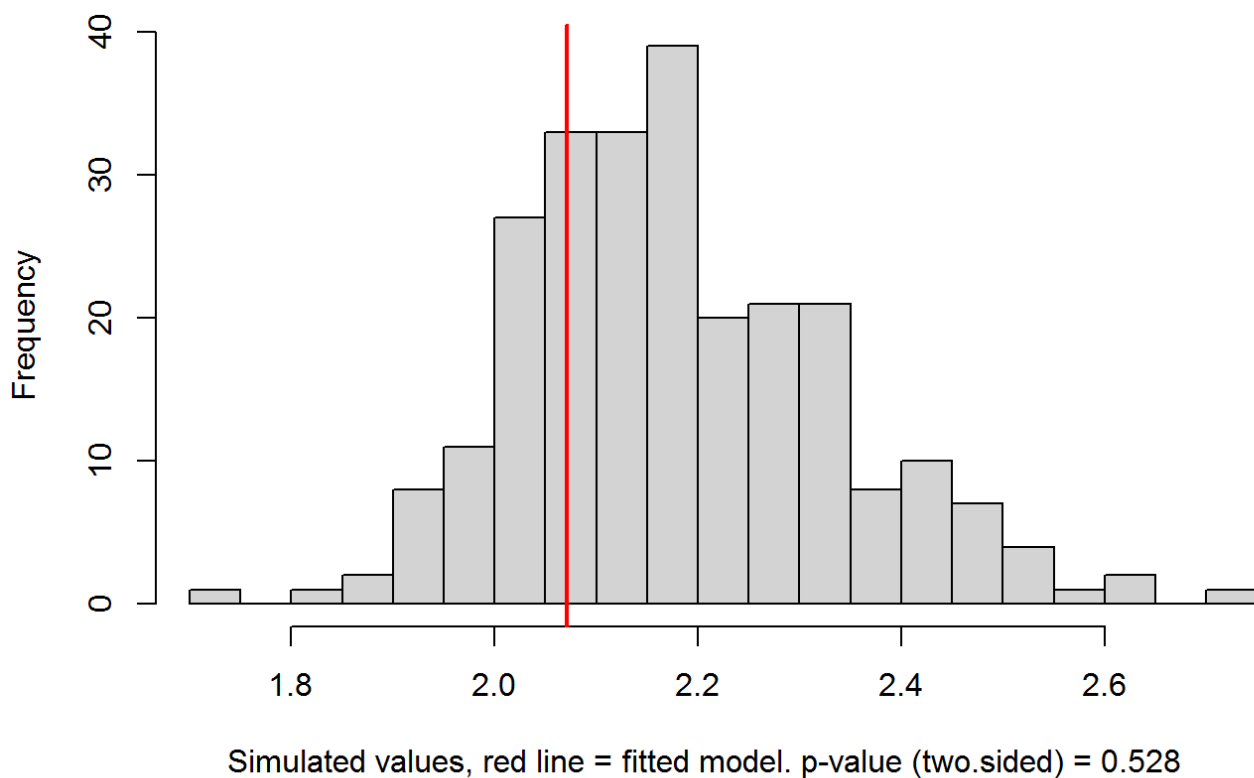
```
## $uniformity
##
## One-sample Kolmogorov-Smirnov test
##
## data: simulationOutput$scaledResiduals
## D = 0.040992, p-value = 0.9546
## alternative hypothesis: two-sided
##
##
## $dispersion
##
## DHARMA nonparametric dispersion test via sd of residuals fitted
## vs. simulated
##
## data: simulationOutput
## ratioObsSim = 0.9508, p-value = 0.528
## alternative hypothesis: two.sided
##
##
## $outliers
##
## DHARMA outlier test based on exact binomial test
##
## data: simulationOutput
## outLow = 0.0000e+00, outHigh = 0.0000e+00, nobs = 1.5700e+02,
## freqH0 = 3.9841e-03, p-value = 1
## alternative hypothesis: two.sided
```

```
plotQQunif(rr, testUniformity = T) ### plot
```



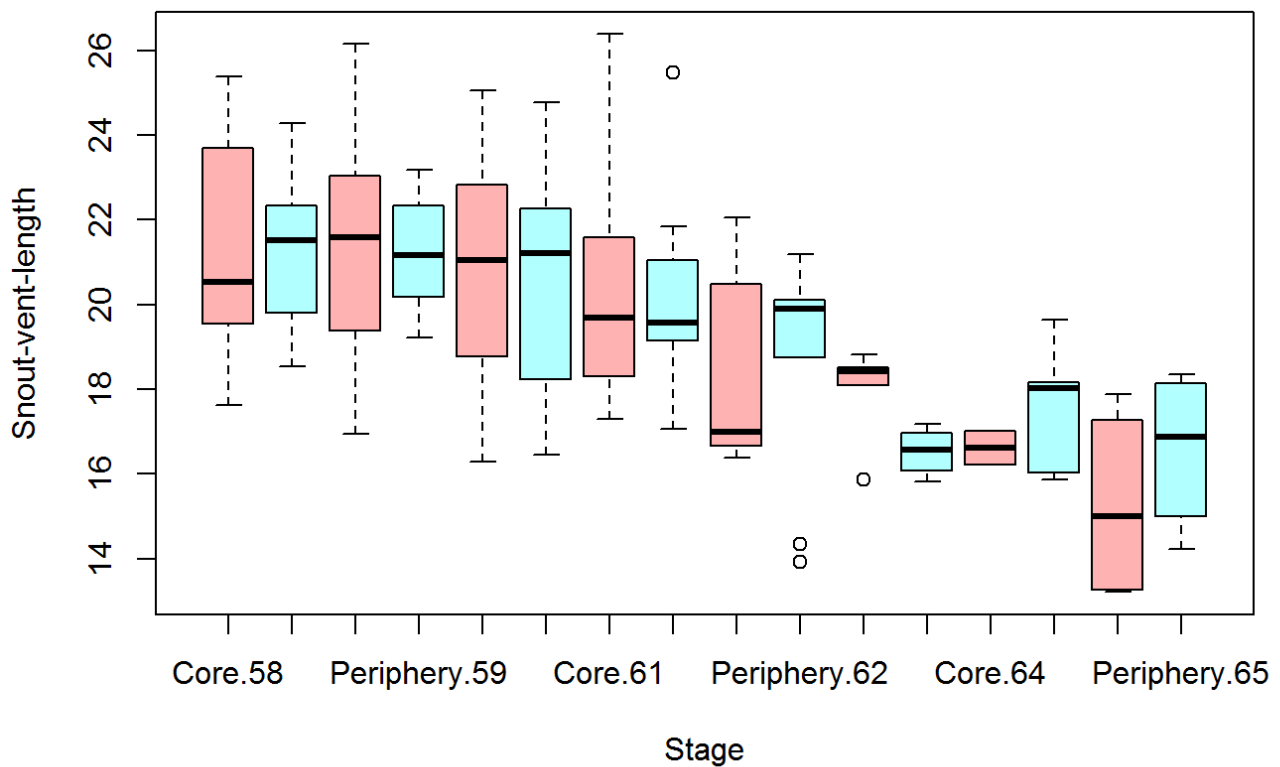
```
testDispersion(rr,plot=T) ### test
```

**DHARMA nonparametric dispersion test via sd of
residuals fitted vs. simulated**



```
##
## DHARMA nonparametric dispersion test via sd of residuals fitted
## vs. simulated
##
## data: simulationOutput
## ratioObsSim = 0.9508, p-value = 0.528
## alternative hypothesis: two.sided
```

```
boxplot(Snoutventlength~Position*Stage1, xaxs=F, data=climaxdataset, xlab="Stage", y
lab="Snout-vent-length", col= rainbow(2, alpha = 0.3), notch= F)
```



Femur

```
F1b<- lmer(pFemurrel ~ Position*Week1+ (1|Site/Clutchx), data = climaxdataset)
```

```
## boundary (singular) fit: see ?isSingular
```

```
F2b<- lmer(pFemurrel ~ Position + Week1 + (1|Site/Clutchx), data = climaxdataset)
```

```
## boundary (singular) fit: see ?isSingular
```

```
F3b<- lmer(pFemurrel ~ Week1 + (1|Site/Clutchx), data=climaxdataset)
```

```
## boundary (singular) fit: see ?isSingular
```

```
F4b<- lmer(pFemurrel ~ Position + (1|Site/Clutchx), data = climaxdataset)
```

```
## boundary (singular) fit: see ?isSingular
```

```
F5b<- lmer(pFemurrel ~ Position*Stage1 + (1|Site/Clutchx), data = climaxdataset)
```

```
## boundary (singular) fit: see ?isSingular
```

```
F6b<- lmer(pFemurrel ~ Position + Stage1 + (1|Site/Clutchx), data = climaxdataset)
```

```
## boundary (singular) fit: see ?isSingular
```

```
F7b<- lmer(pFemurrel ~ Stage1 + (1|Site/Clutchx), data= climaxdataset)
```

```
## boundary (singular) fit: see ?isSingular
```

```
F8b<- lmer(pFemurrel ~ 1 + (1|Site/Clutchx), data=climaxdataset)
```

```
## boundary (singular) fit: see ?isSingular
```

```
F9b<- lmer(pFemurrel ~ Position + Stage1 + Week1 + (1|Site/Clutchx), data= climaxda  
taset)
```

```
## boundary (singular) fit: see ?isSingular
```

```
#Position + Stage1*Week1 was dropped due to rank deficiency
```

```
#Stage1*Week1 was dropped due to rank deficiency
```

```
F12b<- lmer(pFemurrel ~ Stage1+Week1 + (1|Site/Clutchx), data= climaxdataset)
```

```
## boundary (singular) fit: see ?isSingular
```

```
AIC(F1b, F2b, F3b, F4b, F5b, F6b, F7b, F8b, F9b, F12b)
```



```
##      df      AIC
## F1b  15 898.7482
## F2b  10 909.6291
## F3b   9 909.0243
## F4b   5 925.6437
## F5b  19 844.4264
## F6b  12 865.4527
## F7b  11 864.5537
## F8b   4 924.9990
## F9b  17 832.4357
## F12b 16 831.8965
```

```
AICctab(F1b, F2b, F3b, F4b, F5b, F6b, F7b, F8b, F9b, F12b)
```

```
##      dAICc df
## F12b  0.0  16
## F9b   1.1  17
## F5b  14.2  19
## F7b  30.6  11
## F6b  31.8  12
## F1b  66.4  15
## F3b  74.5   9
## F2b  75.4  10
## F8b  89.5   4
## F4b  90.3   5
```

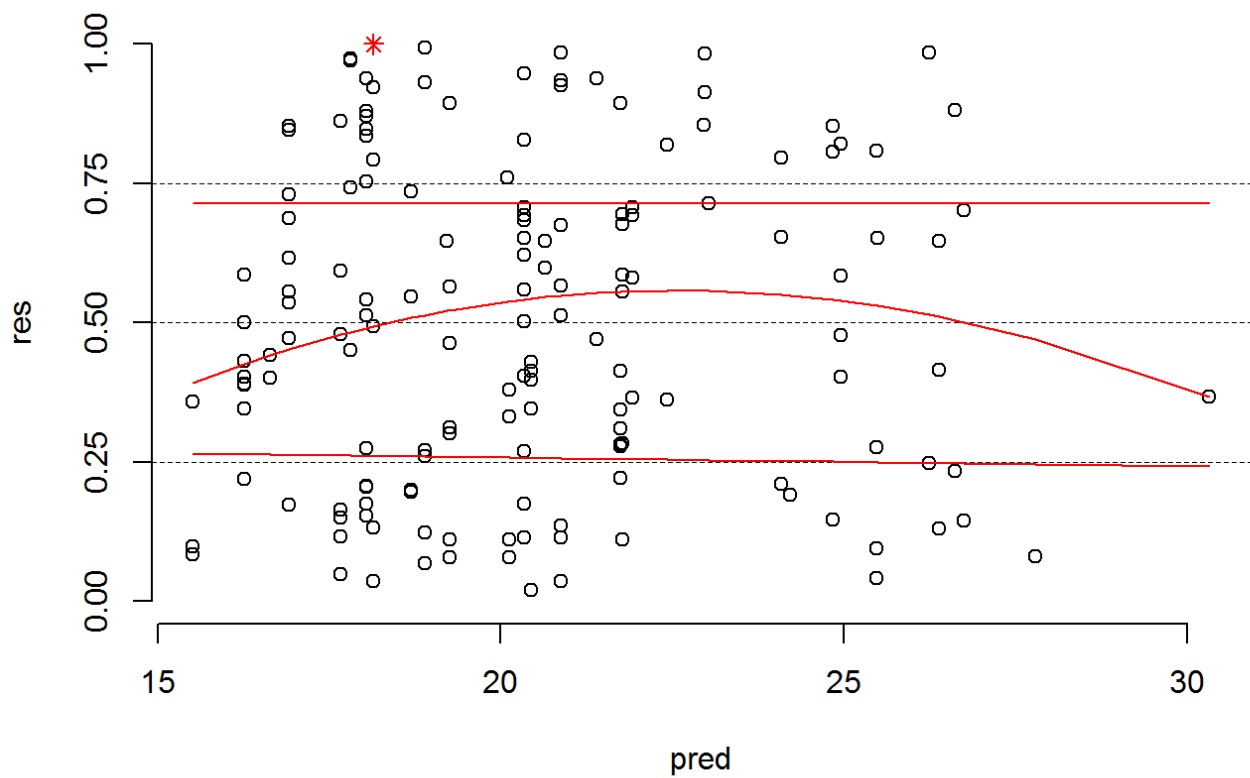
```
summary(F12b)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: pFemurrel ~ Stage1 + Week1 + (1 | Site/Clutchx)
## Data: climaxdataset
##
## REML criterion at convergence: 799.9
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -1.90690 -0.75237 -0.08933  0.64403  2.65404
##
## Random effects:
## Groups      Name                Variance Std.Dev.
## Clutchx:Site (Intercept) 4.194e-15 6.476e-08
## Site        (Intercept) 0.000e+00 0.000e+00
## Residual                    1.184e+01 3.440e+00
## Number of obs: 157, groups: Clutchx:Site, 22; Site, 6
##
## Fixed effects:
##              Estimate Std. Error t value
## (Intercept)  18.6921    1.5482  12.074
## Stage159      1.4023    0.8598   1.631
## Stage160      2.6306    0.9560   2.752
## Stage161      4.6186    1.0235   4.513
## Stage162      5.1426    1.1141   4.616
## Stage163      8.5787    1.3516   6.347
## Stage164      9.9828    1.4584   6.845
## Stage165      6.4051    1.4017   4.570
## Week15       -0.8818    1.8094  -0.487
## Week16       -3.1861    1.7373  -1.834
## Week17       -2.0614    1.7498  -1.178
## Week18       -2.4364    1.7020  -1.431
## Week19        1.6472    1.6782   0.982
```

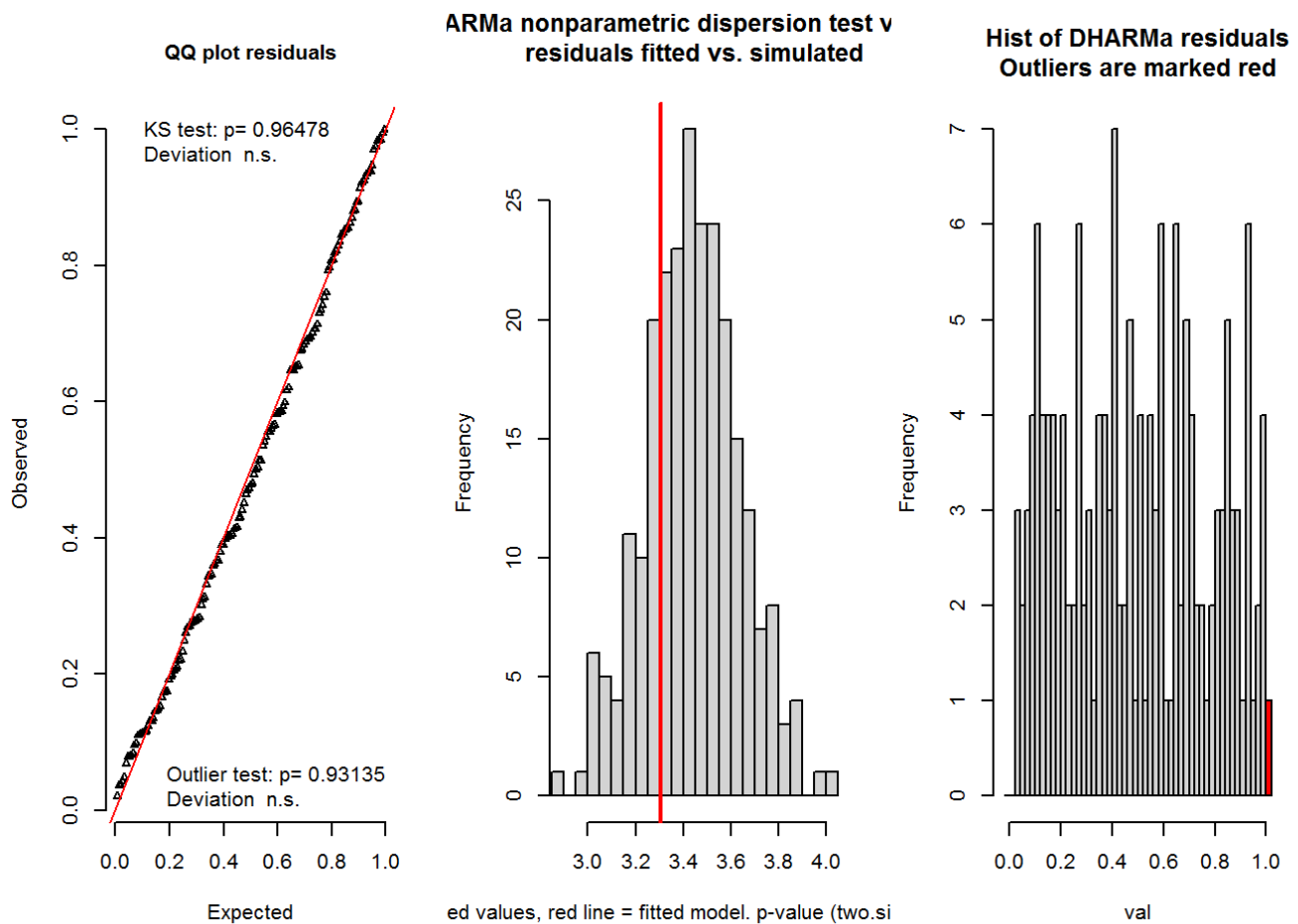
```
##
## Correlation matrix not shown by default, as p = 13 > 12.
## Use print(x, correlation=TRUE) or
##      vcov(x)          if you need it
```

```
## convergence code: 0
## boundary (singular) fit: see ?isSingular
```

```
rr=simulateResiduals(F12b) ### simulate residuals
plotResiduals(rr) ### qqplot simulated vs observed
```



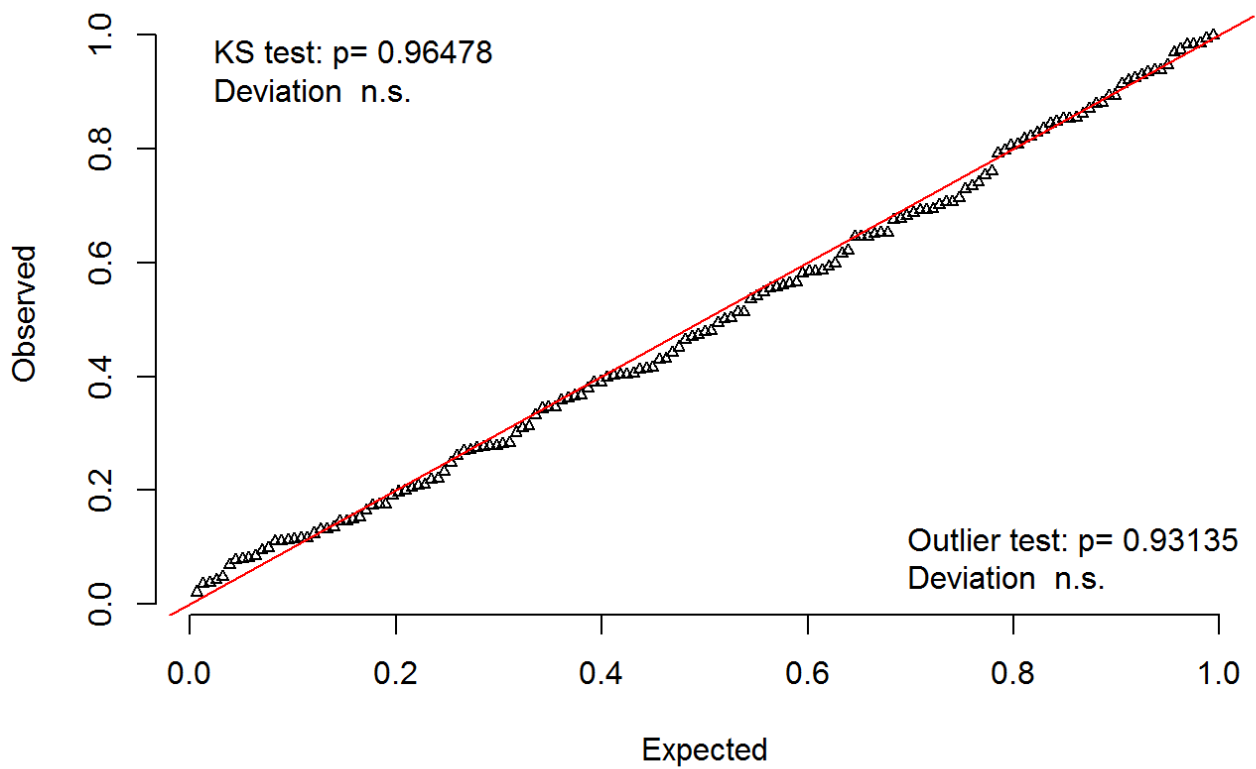
```
testResiduals(rr)###test qqplot
```



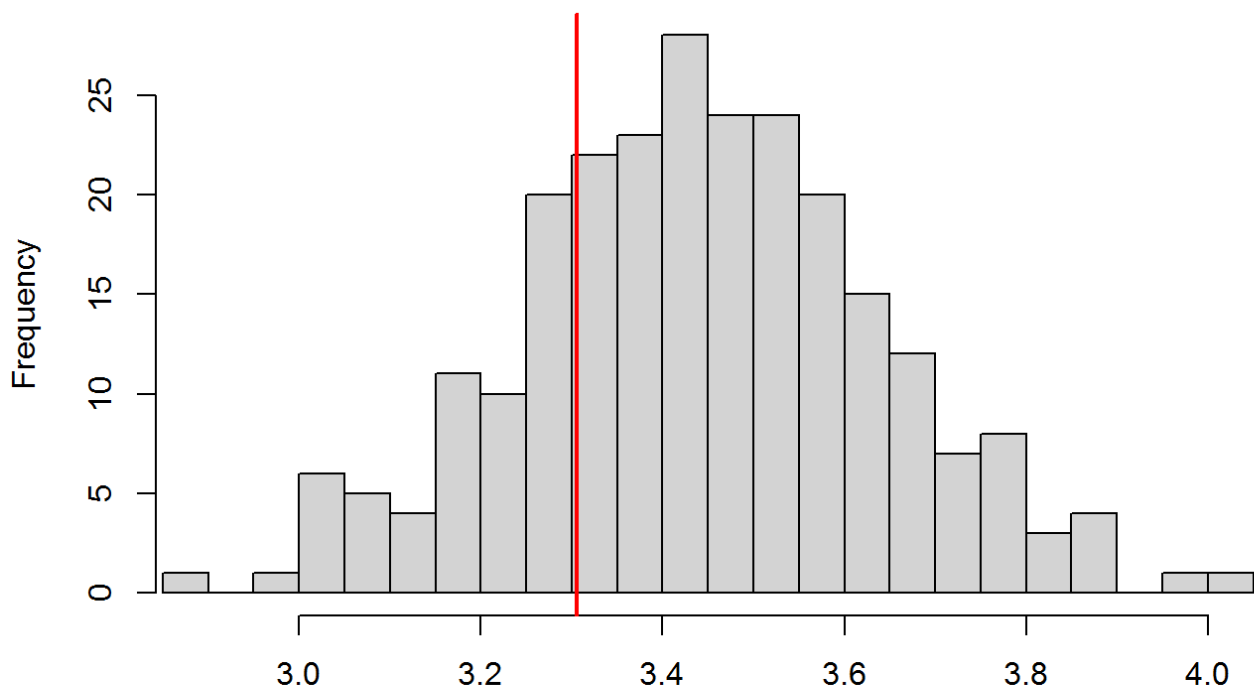
```
## $uniformity
##
## One-sample Kolmogorov-Smirnov test
##
## data: simulationOutput$scaledResiduals
## D = 0.0398, p-value = 0.9648
## alternative hypothesis: two-sided
##
##
## $dispersion
##
## DHARMA nonparametric dispersion test via sd of residuals fitted
## vs. simulated
##
## data: simulationOutput
## ratioObsSim = 0.96022, p-value = 0.472
## alternative hypothesis: two.sided
##
##
## $outliers
##
## DHARMA outlier test based on exact binomial test
##
## data: simulationOutput
## outLow = 0.0000e+00, outHigh = 1.0000e+00, nobs = 1.5700e+02,
## freqH0 = 3.9841e-03, p-value = 0.9313
## alternative hypothesis: two.sided
```

```
## $uniformity
##
## One-sample Kolmogorov-Smirnov test
##
## data: simulationOutput$scaledResiduals
## D = 0.0398, p-value = 0.9648
## alternative hypothesis: two-sided
##
##
## $dispersion
##
## DHARMA nonparametric dispersion test via sd of residuals fitted
## vs. simulated
##
## data: simulationOutput
## ratioObsSim = 0.96022, p-value = 0.472
## alternative hypothesis: two.sided
##
##
## $outliers
##
## DHARMA outlier test based on exact binomial test
##
## data: simulationOutput
## outLow = 0.0000e+00, outHigh = 1.0000e+00, nobs = 1.5700e+02,
## freqH0 = 3.9841e-03, p-value = 0.9313
## alternative hypothesis: two.sided
```

```
plotQQunif(rr, testUniformity = T) ### plot
```

QQ plot residuals

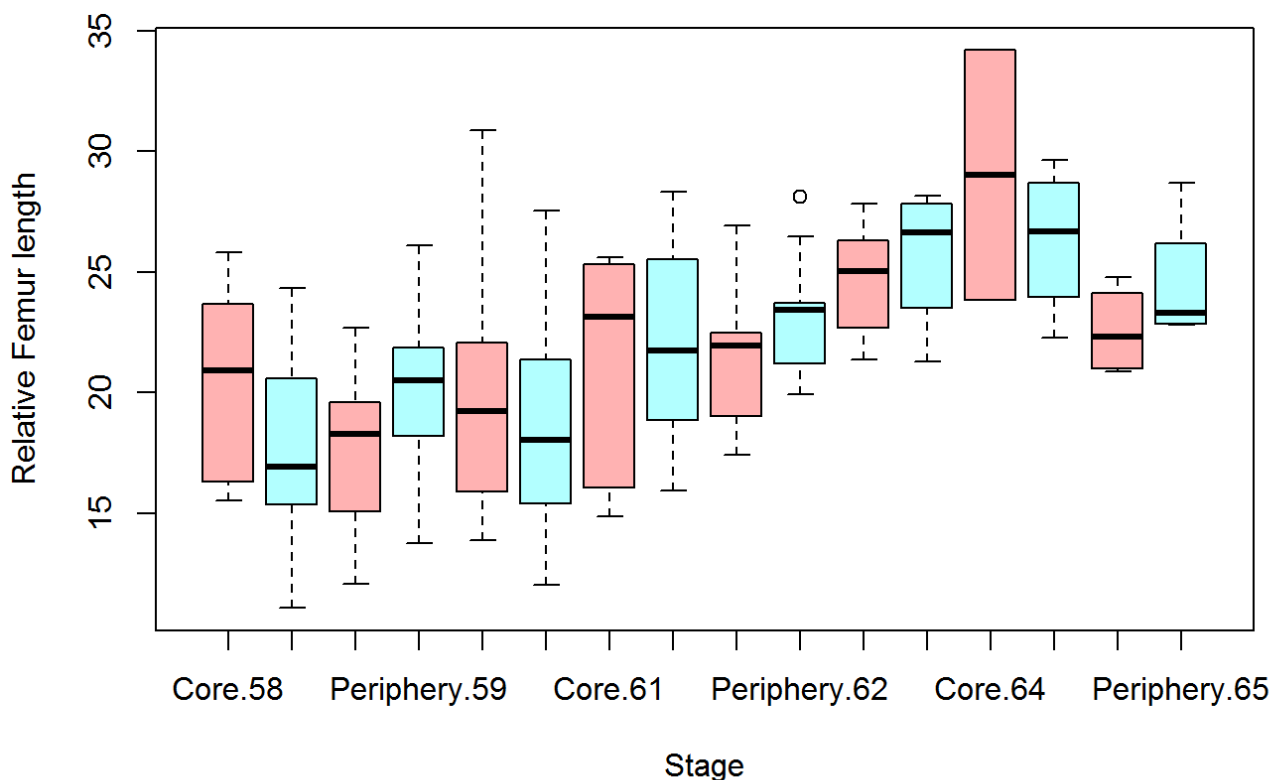
```
testDispersion(rr,plot=T) ### test
```

DHARMA nonparametric dispersion test via sd of residuals fitted vs. simulated

Simulated values, red line = fitted model. $p\text{-value (two.sided)} = 0.472$

```
##
## DHARMA nonparametric dispersion test via sd of residuals fitted
## vs. simulated
##
## data: simulationOutput
## ratioObsSim = 0.96022, p-value = 0.472
## alternative hypothesis: two.sided
```

```
boxplot(Femurrelative~Position*Stage1, xaxs=F,data=climaxdataset, xlab="Stage", ylab="Relative Femur length", col= rainbow(2, alpha = 0.3), notch= F)
```



```
detach(Climax)
```

3.3. Metamorph NF stage 66- MESOCOSMS

3.3.1. Defining variables.

Use *Mesometamorphs.csv* to analyse data.

```
Mesomet=read.csv(file.choose(),header=T)
names(Mesomet)
```

```
## [1] "num"           "Position"      "Site"
## [4] "Clutch"        "Clutchx"       "Tadpole"
## [7] "Week"          "tadpoleID"     "Mesocosm"
## [10] "Date"          "Fulllength"    "Maximumbodydepth"
## [13] "Snoutventlength" "Taillength"   "Maxtaildepth"
## [16] "Headwidth"     "Femur"         "Femurrelative"
## [19] "Totalleglength" "Stage"         "Stagecategory"
```

```
attach(Mesomet)
```

```
## The following object is masked _by_ .GlobalEnv:
##
##      Tadpole
```

```
summary(Position) #fixed effect
```

```
##      Core Periphery
##      33      25
```

```
summary(Week) #fixed effect
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      5.000   8.000   8.000   8.034   9.000   9.000
```

```
Week1 <- as.factor(Week) #as a discrete variable
summary(Week1)
```

```
##  5  6  7  8  9
##  1  3 10 23 21
```

```
summary(Stage) #fixed effect
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      66      66      66      66      66      66
```

```
#just one stage so it will not be added to the model
```

```
summary(Site) #Random effect
```

```
##  B  C  J  M  S  T
## 11  6 10  9 13  9
```

```
summary(Clutchx) #Nested within site
```



```
## B2 B3 B4 C1 C4 J2 J3 J4 M1 M2 M4 S1 S2 S3 S4 T1 T2 T3 T4
## 2 4 5 4 2 4 4 2 4 3 2 3 1 6 3 1 3 2 3
```

```
Mesomet1<- cbind(Mesomet, Week1)
names(Mesomet1)
```

```
## [1] "num"           "Position"      "Site"
## [4] "Clutch"        "Clutchx"       "Tadpole"
## [7] "Week"          "tadpoleID"     "Mesocosm"
## [10] "Date"          "Fulllength"    "Maximumbodydepth"
## [13] "Snoutventlength" "Taillength"   "Maxtaildepth"
## [16] "Headwidth"     "Femur"         "Femurrelative"
## [19] "Totalleglength" "Stage"         "Stagecategory"
## [22] "Week1"
```

3.2.2. Transforming response variables.

Use the *Mesomet1* datasheet just created to measure normality.

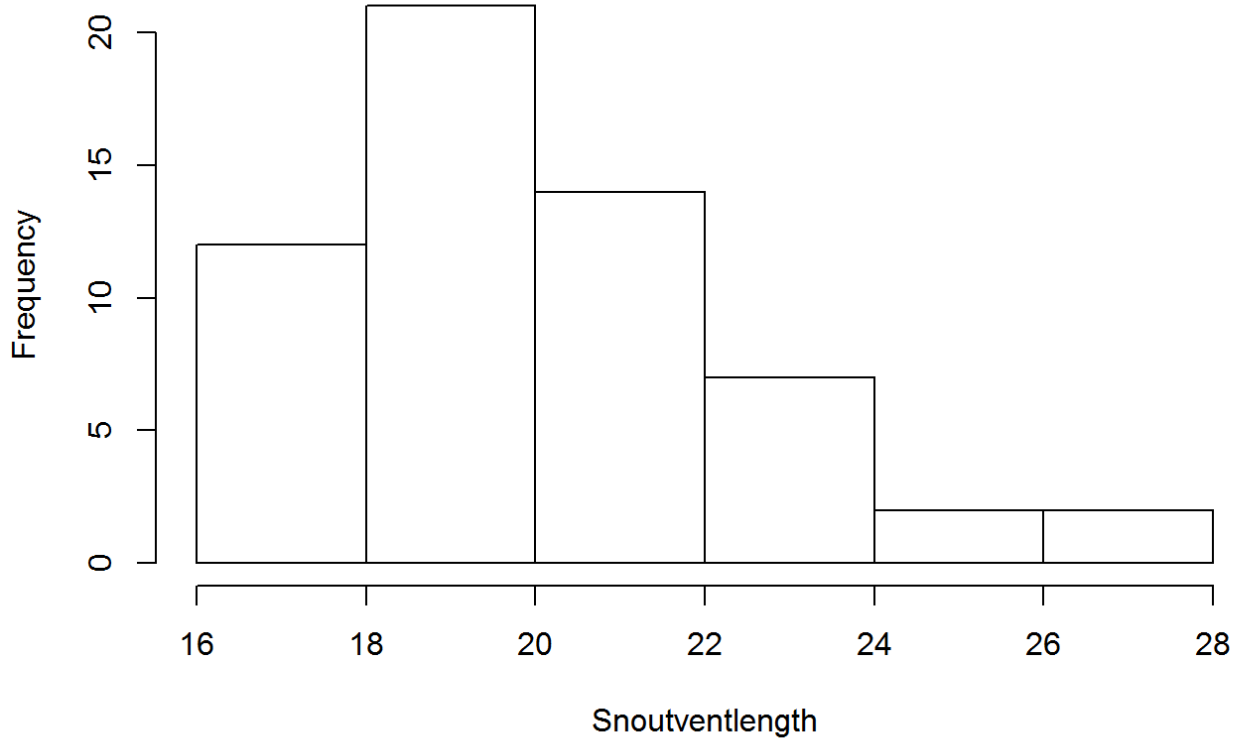
Snoutventlength

```
shapiro.test(Snoutventlength)
```

```
##
## Shapiro-Wilk normality test
##
## data: Snoutventlength
## W = 0.96026, p-value = 0.05494
```

```
#p-value = 0.05494
#no transformation
hist(Snoutventlength)
```

Histogram of Snoutventlength



```
pSVL <- Snoutventlength
```

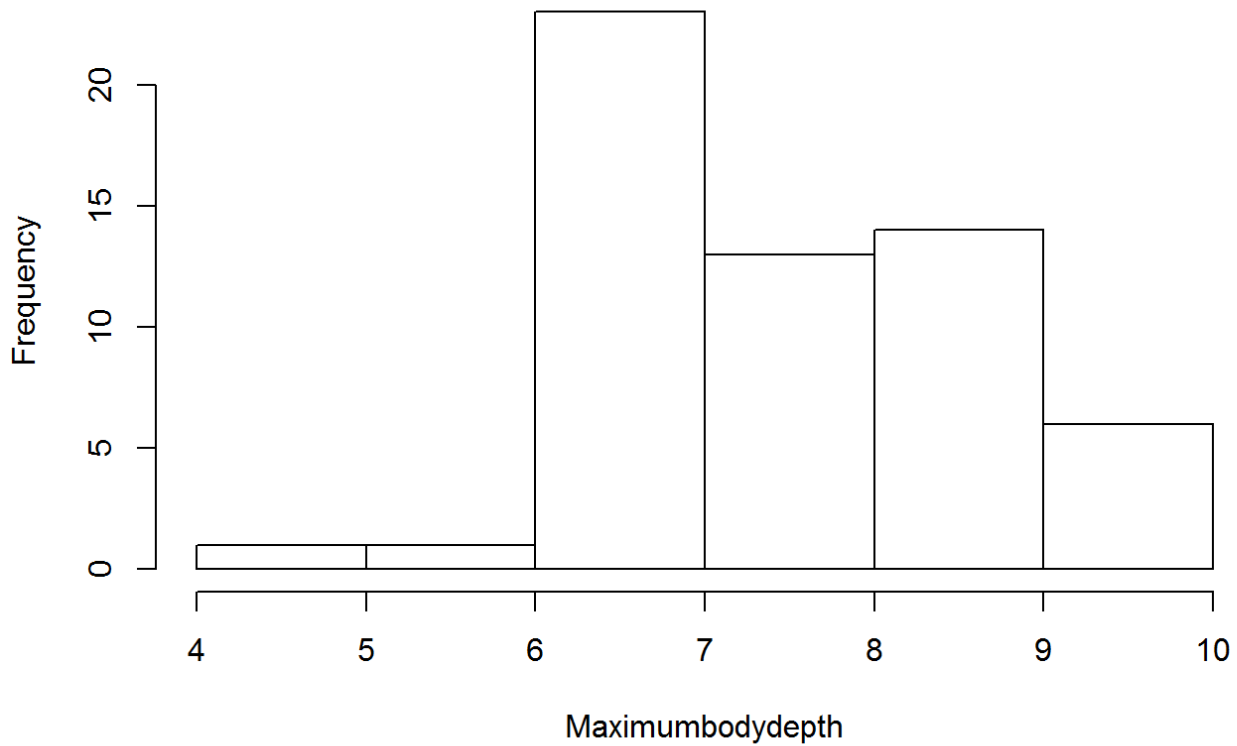
Maximumbodydepth

```
shapiro.test(Maximumbodydepth)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  Maximumbodydepth
## W = 0.96374, p-value = 0.08058
```

```
#p-value = 0.08058
hist(Maximumbodydepth)
```

Histogram of Maximumbodydepth



```
#no transformation
pMaximumbodydepth <- Maximumbodydepth
summary(pMaximumbodydepth)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    4.850   6.578   7.170   7.497   8.378   9.854
```

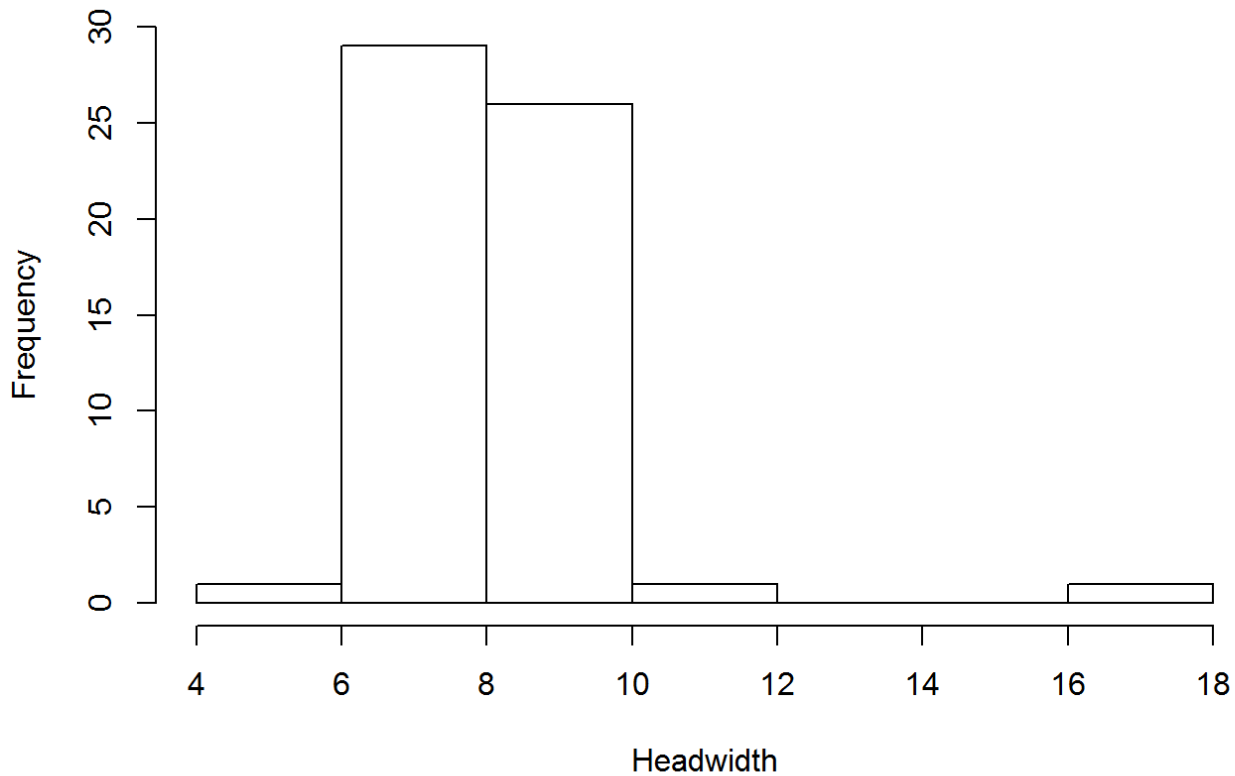
Headwidth

```
shapiro.test(Headwidth)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  Headwidth
## W = 0.749, p-value = 1.323e-08
```

```
#p-value = 1.323e-08
hist(Headwidth)
```

Histogram of Headwidth



```
bestNormalize(Headwidth, standardize= TRUE, allow_orderNorm=TRUE,
              allow_lambert_s = FALSE, allow_lambert_h = FALSE,
              out_of_sample = TRUE, cluster = NULL, k = 10, r = 5,
              loo = FALSE, warn = TRUE, quiet = FALSE)
```

```
## Warning in get_oos_estimates(x, standardize, method_names, k, r, cluster, :
## fold_size is 5 (< 20), therefore P/df estimates may be off
```

```
## Best Normalizing transformation with 58 Observations
## Estimated Normality Statistics (Pearson P / df, lower => more normal):
## - No transform: 1.5673
## - Box-Cox: 1.5133
## - Log_b(x+a): 1.4647
## - sqrt(x+a): 1.434
## - exp(x): 3.8147
## - arcsinh(x): 1.4647
## - Yeo-Johnson: 1.48
## - orderNorm: 1.3787
## Estimation method: Out-of-sample via CV with 10 folds and 5 repeats
##
## Based off these, bestNormalize chose:
## orderNorm Transformation with 58 nonmissing obs and no ties
## - Original quantiles:
##   0%    25%    50%    75%   100%
## 5.577  7.338  7.946  8.440 16.477
```

```

#Based off these, bestNormalize chose:
#Standardized asinh(x) Transformation with 58 nonmissing obs

#I chose to ordernorm transform due to consistency.

orderNorm_Headwidth <- orderNorm(Headwidth)
pHeadwidth <- predict(orderNorm_Headwidth)
summary(orderNorm_Headwidth)

```

```

##           Length Class  Mode
## x.t       58      -none- numeric
## x         58      -none- numeric
## n          1      -none- numeric
## ties_status 1      -none- numeric
## fit        30      glm      list
## norm_stat   1      -none- numeric

```

```
shapiro.test(pHeadwidth)
```

```

##
##  Shapiro-Wilk normality test
##
## data:  pHeadwidth
## W = 0.99569, p-value = 0.9993

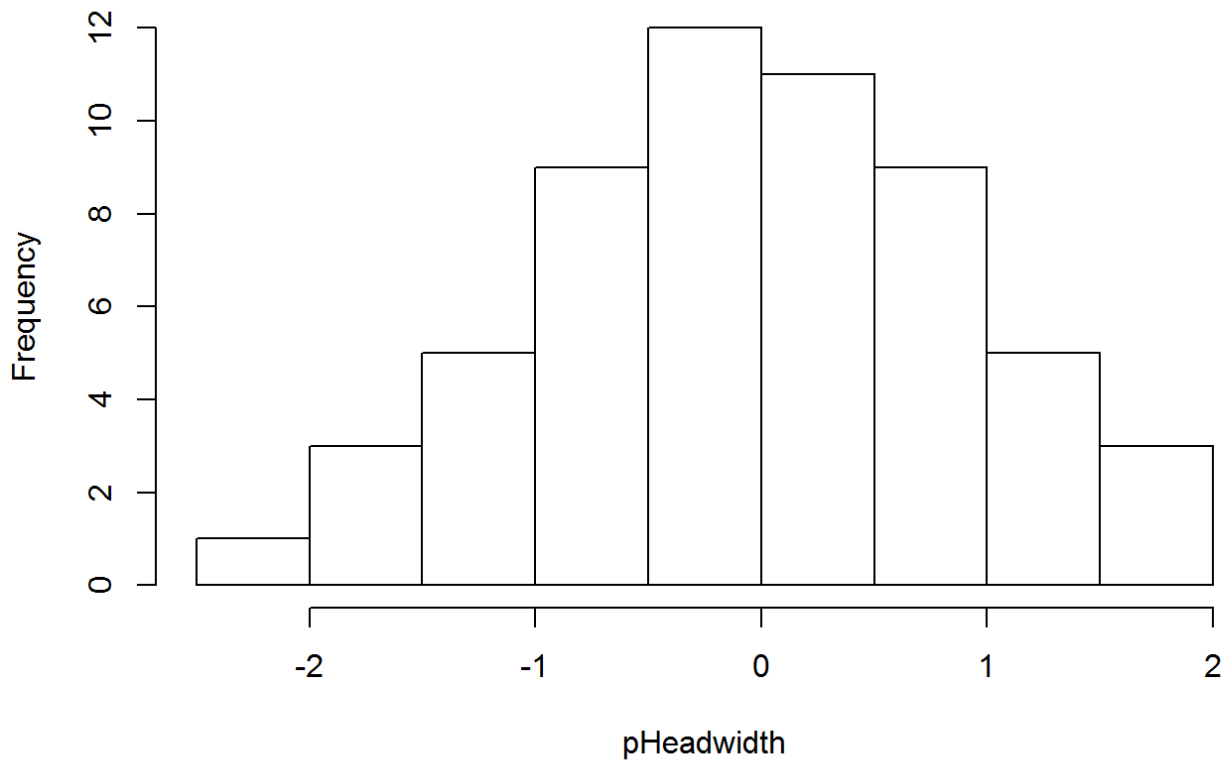
```

```

#p-value = 0.9993
hist(pHeadwidth)

```

Histogram of pHeadwidth



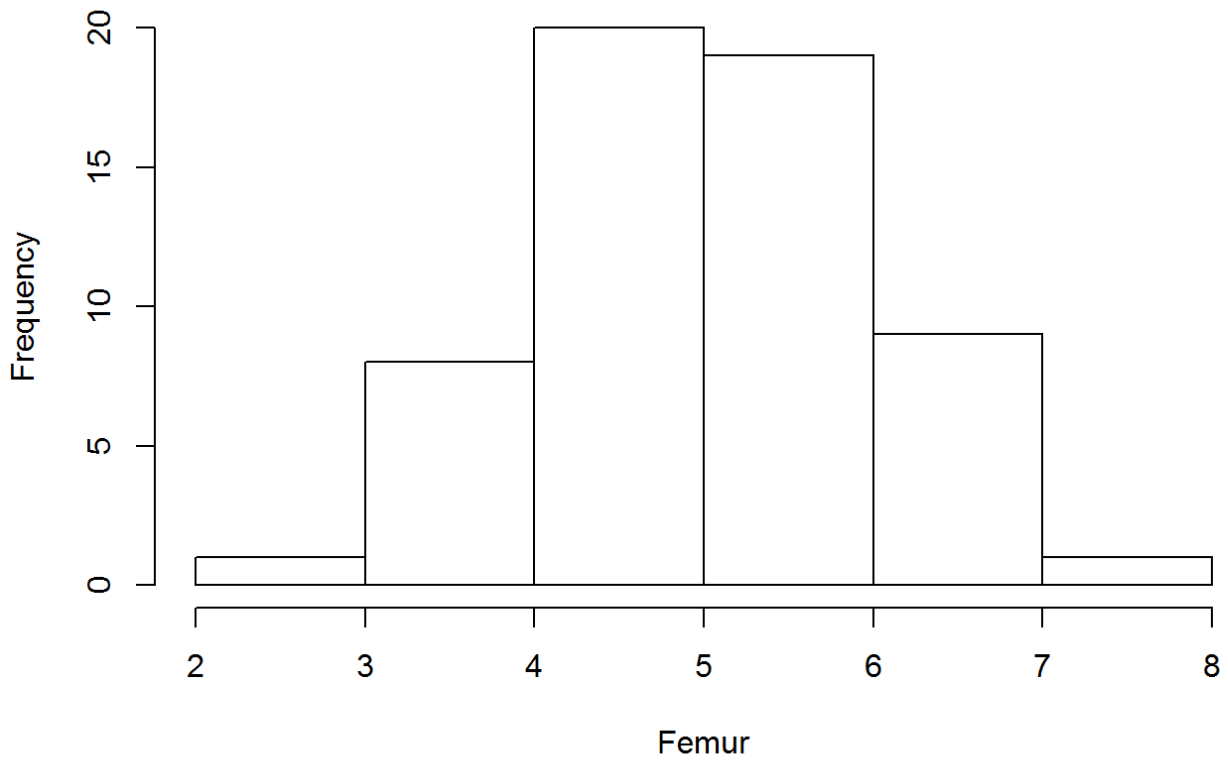
Femur

```
shapiro.test(Femur)
```

```
##  
##  Shapiro-Wilk normality test  
##  
## data:  Femur  
## W = 0.98646, p-value = 0.7635
```

```
#p-value = 0.7635  
hist(Femur)
```

Histogram of Femur



```
#no transform
```

```
pFemur<- Femur
```

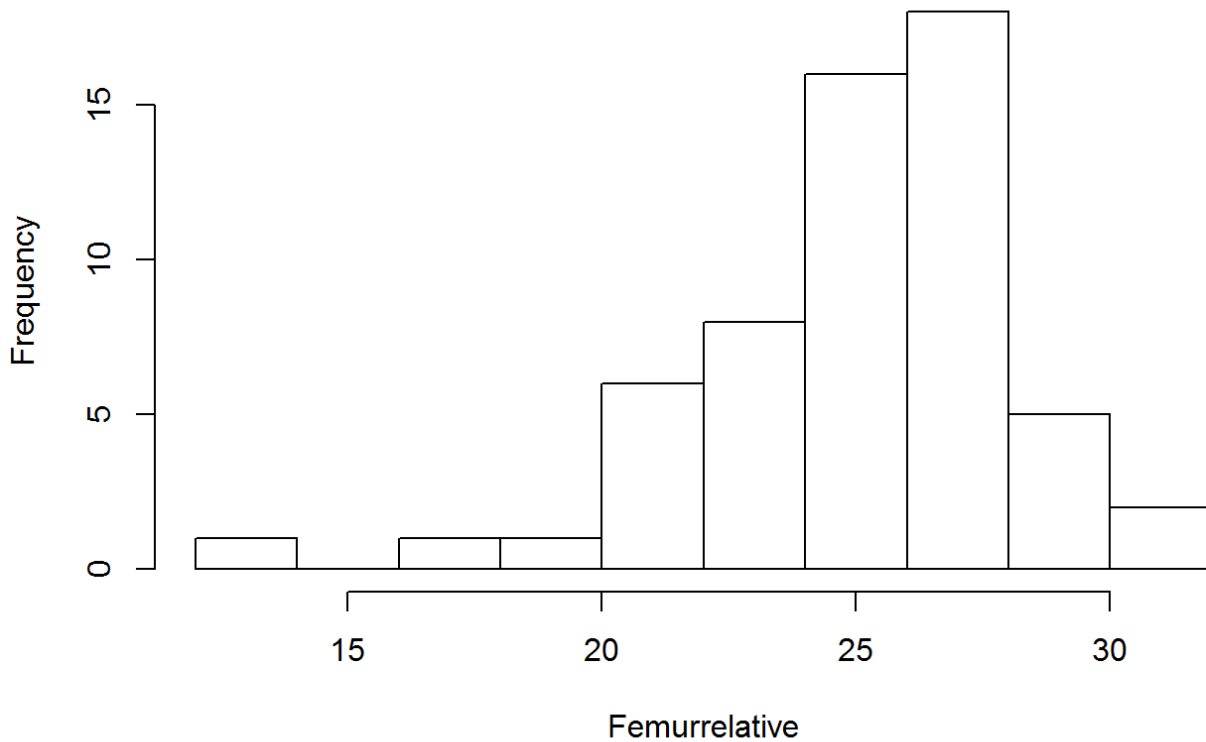
Relative Femur length

```
shapiro.test(Femurrelative)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  Femurrelative
## W = 0.94797, p-value = 0.01469
```

```
#p-value = 0.01469
hist(Femurrelative)
```

Histogram of Femurrelative



```
bestNormalize(Femurrelative, standardize= TRUE, allow_orderNorm=TRUE,
              allow_lambert_s = FALSE, allow_lambert_h = FALSE,
              out_of_sample = TRUE, cluster = NULL, k = 10, r = 5,
              loo = FALSE, warn = TRUE, quiet = FALSE)
```

```
## Warning in get_oos_estimates(x, standardize, method_names, k, r, cluster, :
## fold_size is 5 (< 20), therefore P/df estimates may be off
```

```
## Best Normalizing transformation with 58 Observations
## Estimated Normality Statistics (Pearson P / df, lower => more normal):
## - No transform: 1.276
## - Box-Cox: 1.3593
## - Log_b(x+a): 1.39
## - sqrt(x+a): 1.2927
## - exp(x): 4.762
## - arcsinh(x): 1.39
## - Yeo-Johnson: 1.3747
## - orderNorm: 1.5207
## Estimation method: Out-of-sample via CV with 10 folds and 5 repeats
##
## Based off these, bestNormalize chose:
## Standardized I(x) Transformation with 58 nonmissing obs.:
## Relevant statistics:
## - mean (before standardization) = 24.97784
## - sd (before standardization) = 3.312115
```



```
#Based off these, bestNormalize chose:
#Standardized Yeo-Johnson Transformation with 58 nonmissing obs:

#I chose to ordernorm transform due to consistency.

ordernormFemurrel <- orderNorm(Femurrelrelative)
pFemurrel <- predict(ordernormFemurrel)
summary(pFemurrel)
```

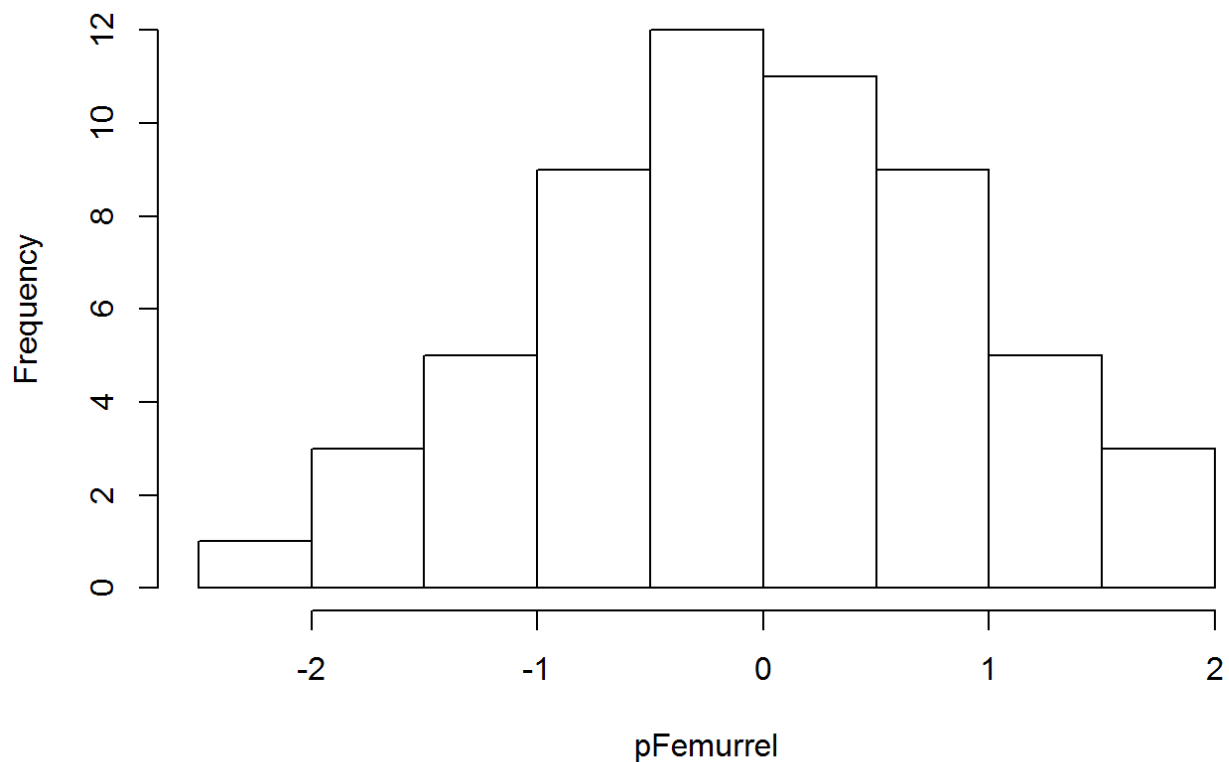
```
##      Min.   1st Qu.   Median     Mean   3rd Qu.     Max.
## -2.38781 -0.67467 -0.02125 -0.04117  0.62222  1.95276
```

```
shapiro.test(pFemurrel)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  pFemurrel
## W = 0.99569, p-value = 0.9993
```

```
#p-value = 0.9993
hist(pFemurrel)
```

Histogram of pFemurrel



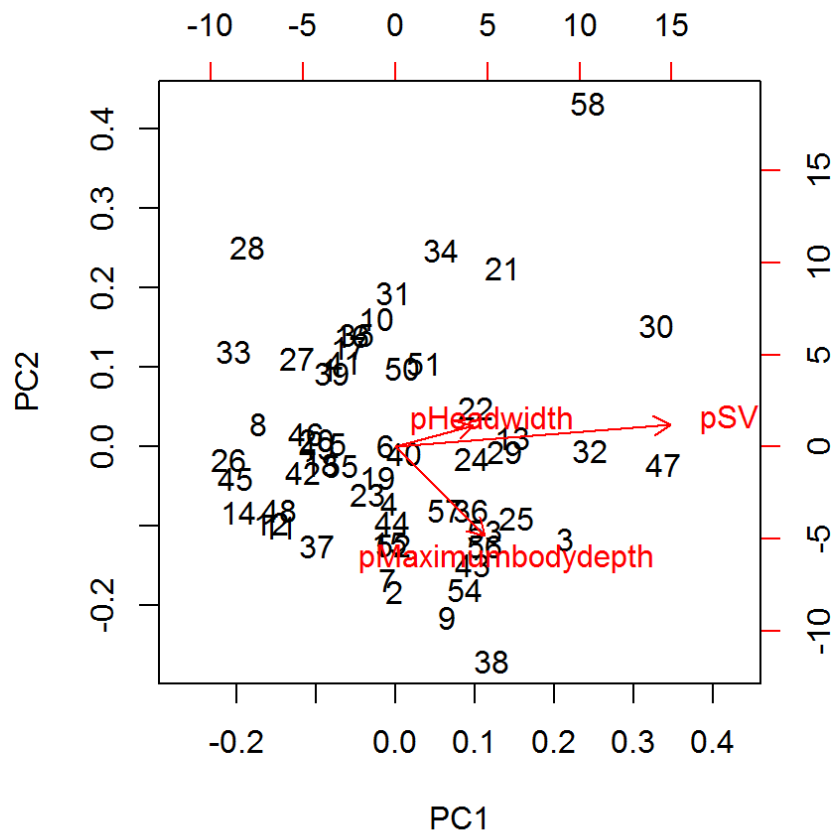
3.2.3. Performing PCA for metamorph variables.

```
Mesometapca<- cbind(pHeadwidth, pSVL, pMaximumbodydepth)
```

```
Mesometapcal=prcomp(Mesometapca)
summary(Mesometapcal)
```

```
## Importance of components:
##              PC1      PC2      PC3
## Standard deviation    2.6630 0.83613 0.67150
## Proportion of Variance 0.8605 0.08483 0.05471
## Cumulative Proportion 0.8605 0.94529 1.00000
```

```
#proportion of variance for PC1- 0.8605
biplot(Mesometapcal)
```



```
Mesometapcalscores<-predict(Mesometapcal)
```

```
Mesometadataset <- cbind(Mesomet1, Mesometapca, Mesometapcalscores)
names(Mesometadataset)
```

```
## [1] "num"           "Position"      "Site"
## [4] "Clutch"        "Clutchx"       "Tadpole"
## [7] "Week"          "tadpoleID"     "Mesocosm"
## [10] "Date"          "Fulllength"    "Maximumbodydepth"
## [13] "Snoutventlength" "Taillength"   "Maxtaildepth"
## [16] "Headwidth"     "Femur"         "Femurrelative"
## [19] "Totalleglength" "Stage"         "Stagecategory"
## [22] "Week1"         "pHeadwidth"    "pSVL"
## [25] "pMaximumbodydepth" "PC1"          "PC2"
## [28] "PC3"
```

3.2.4. Model selection for variables.

PC1

Use the *Mesometadataset* for the analysis.

```
#For week and stage as discrete variables
PC1c<- lmer(PC1 ~ Week1 + (1|Site/Clutchx), data=Mesometadataset)
```

```
## boundary (singular) fit: see ?isSingular
```

```
PC2c<- lmer(PC1 ~ Position + Week1 + (1|Site/Clutchx), data = Mesometadataset)
#Position*Week1 was dropped due to rank deficiency

PC4c<- lmer(PC1 ~ Position+ (1|Site/Clutchx), data = Mesometadataset)
```

```
## boundary (singular) fit: see ?isSingular
```

```
PC5c<- lmer(PC1 ~ 1 + (1|Site/Clutchx), data=Mesometadataset)
```

```
## boundary (singular) fit: see ?isSingular
```

```
AIC(PC1c, PC2c, PC4c, PC5c)
```

```
##      df      AIC
## PC1c  8 280.2239
## PC2c  9 280.4857
## PC4c  5 285.9716
## PC5c  4 285.4450
```

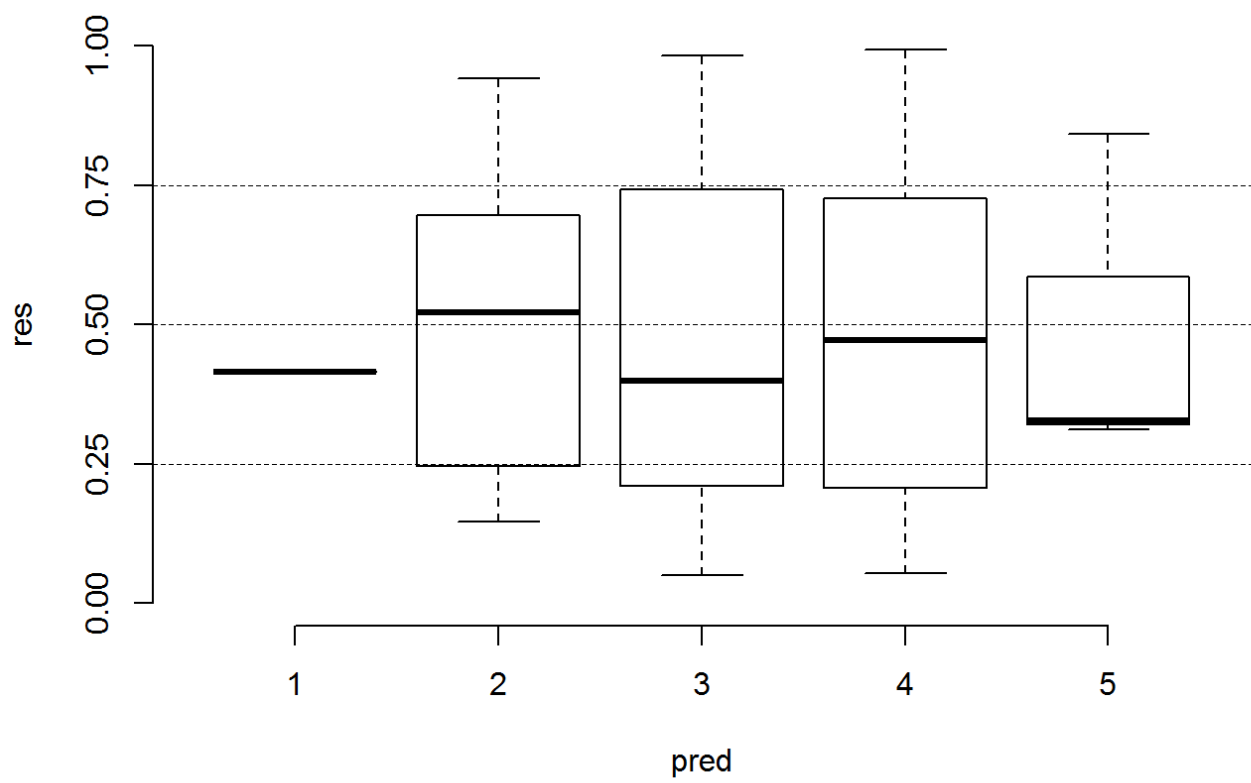
```
AICctab(PC1c, PC2c, PC4c, PC5c)
```

```
##          dAICc df
## PC1c 0.0    8
## PC2c 1.1    9
## PC5c 3.0    4
## PC4c 4.0    5
```

```
#PC1c was selected as the model with the lowest AIC
#PC2c was selected too as it only differs by AIC 1.1
summary(PC1c)
```

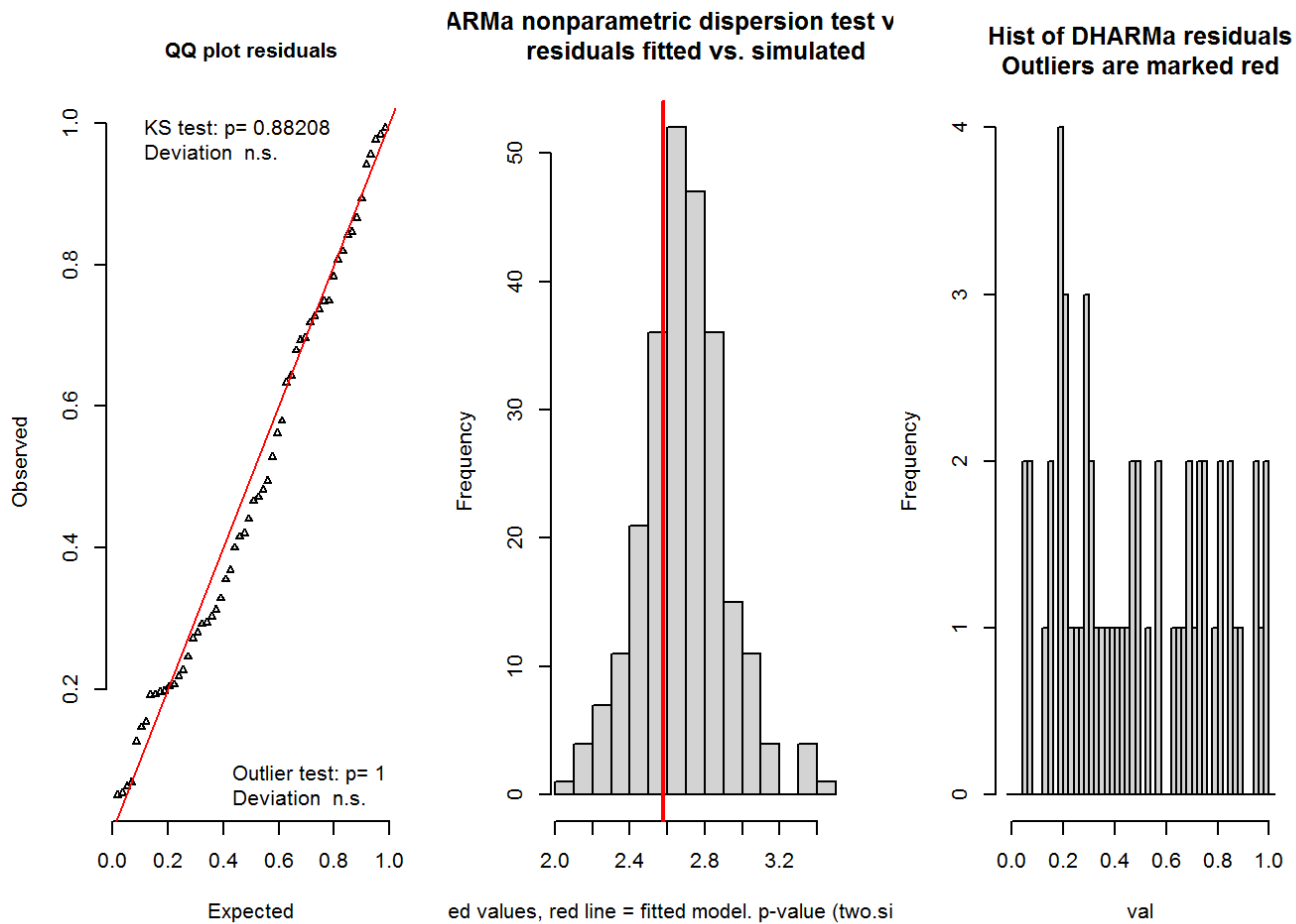
```
## Linear mixed model fit by REML ['lmerMod']
## Formula: PC1 ~ Week1 + (1 | Site/Clutchx)
## Data: Mesometadataset
##
## REML criterion at convergence: 264.2
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -1.6248 -0.7115 -0.1151  0.6579  2.4071
##
## Random effects:
## Groups      Name      Variance Std.Dev.
## Clutchx:Site (Intercept) 0.2694   0.519
## Site         (Intercept) 0.0000   0.000
## Residual                6.8952   2.626
## Number of obs: 58, groups: Clutchx:Site, 19; Site, 6
##
## Fixed effects:
##              Estimate Std. Error t value
## (Intercept)  -1.6001    2.6704  -0.599
## Week16       2.9577    3.0831   0.959
## Week17       0.4182    2.7921   0.150
## Week18       1.7196    2.7214   0.632
## Week19       1.9701    2.7297   0.722
##
## Correlation of Fixed Effects:
##      (Intr) Week16 Week17 Week18
## Week16 -0.866
## Week17 -0.953  0.828
## Week18 -0.978  0.849  0.934
## Week19 -0.976  0.847  0.932  0.957
## convergence code: 0
## boundary (singular) fit: see ?isSingular
```

```
rr=simulateResiduals(PC1c) ### simulate residuals
plotResiduals(rr) ### qqplot simulated vs observed
```



```
## NULL
```

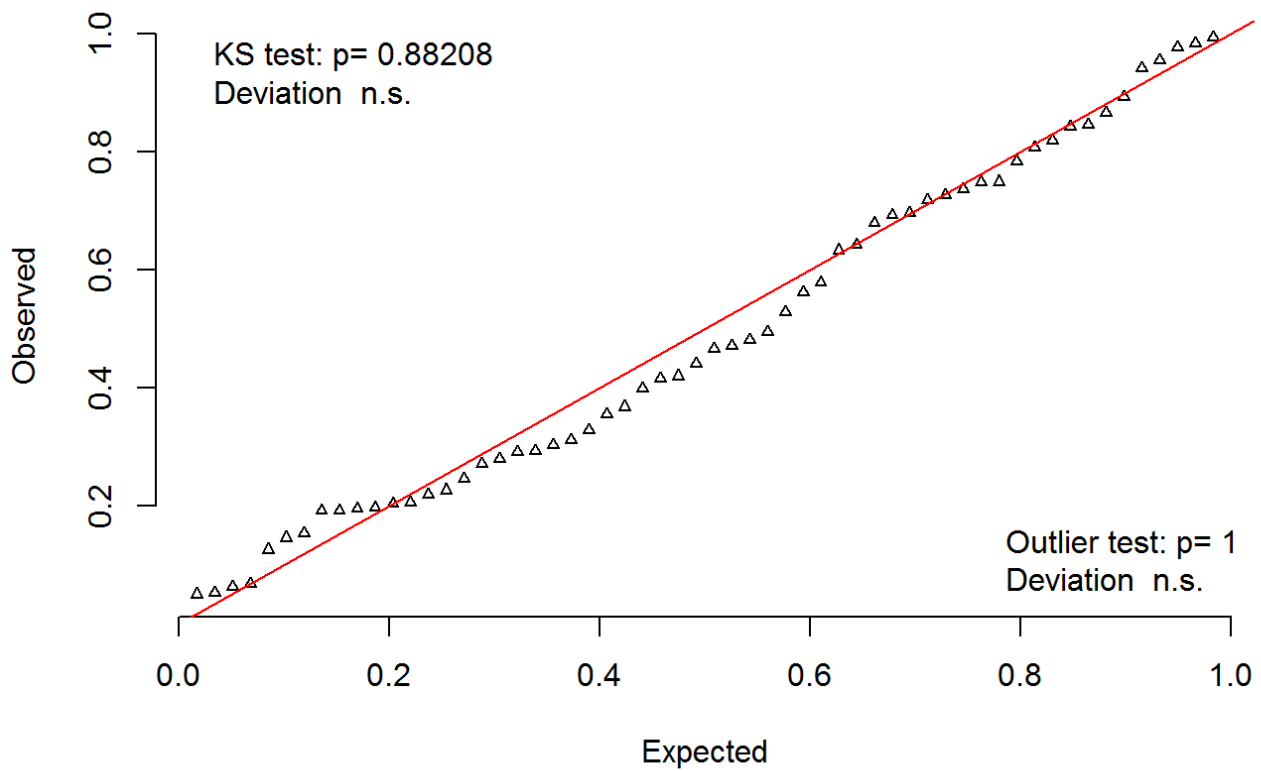
```
testResiduals(rr)###test qqplot
```



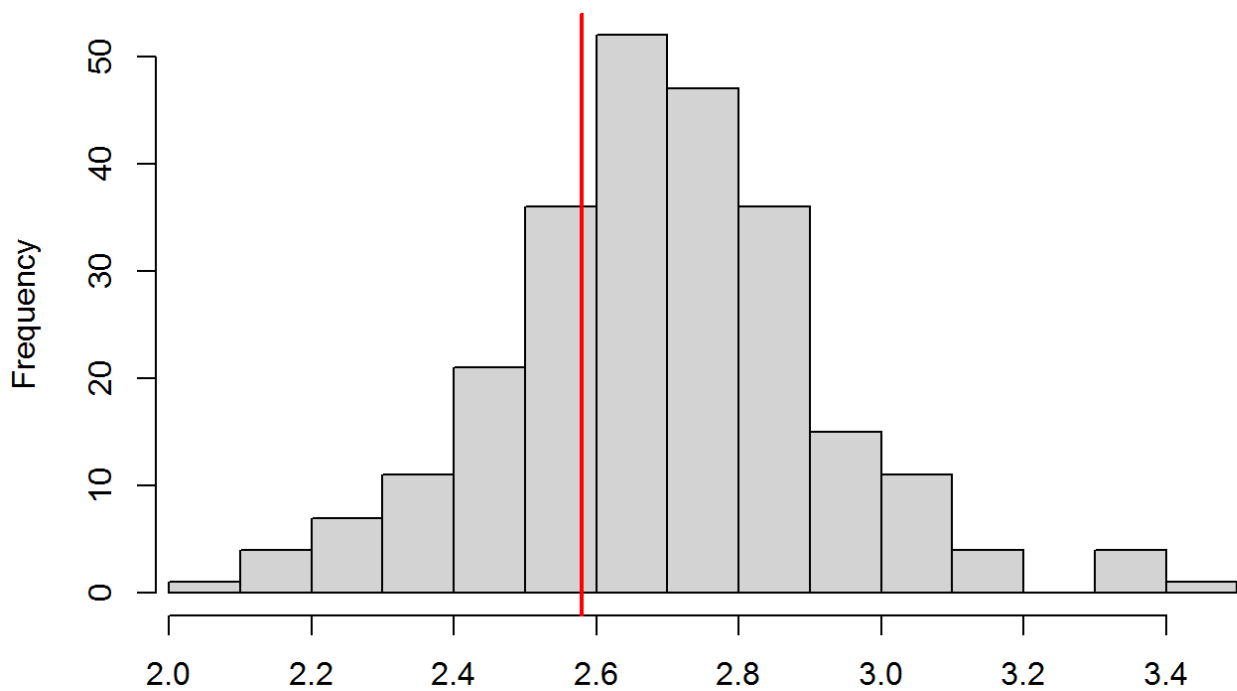
```
## $uniformity
##
## One-sample Kolmogorov-Smirnov test
##
## data: simulationOutput$scaledResiduals
## D = 0.074331, p-value = 0.8821
## alternative hypothesis: two-sided
##
##
## $dispersion
##
## DHARMA nonparametric dispersion test via sd of residuals fitted
## vs. simulated
##
## data: simulationOutput
## ratioObsSim = 0.96122, p-value = 0.608
## alternative hypothesis: two.sided
##
##
## $outliers
##
## DHARMA outlier test based on exact binomial test
##
## data: simulationOutput
## outLow = 0.0000000, outHigh = 0.0000000, nobs = 58.0000000, freqH0
## = 0.0039841, p-value = 1
## alternative hypothesis: two.sided
```

```
## $uniformity
##
## One-sample Kolmogorov-Smirnov test
##
## data: simulationOutput$scaledResiduals
## D = 0.074331, p-value = 0.8821
## alternative hypothesis: two-sided
##
##
## $dispersion
##
## DHARMA nonparametric dispersion test via sd of residuals fitted
## vs. simulated
##
## data: simulationOutput
## ratioObsSim = 0.96122, p-value = 0.608
## alternative hypothesis: two.sided
##
##
## $outliers
##
## DHARMA outlier test based on exact binomial test
##
## data: simulationOutput
## outLow = 0.0000000, outHigh = 0.0000000, nobs = 58.0000000, freqH0
## = 0.0039841, p-value = 1
## alternative hypothesis: two.sided
```

```
plotQQunif(rr, testUniformity = T) ### plot
```

QQ plot residuals

```
testDispersion(rr,plot=T) ### test
```

DHARMA nonparametric dispersion test via sd of residuals fitted vs. simulated

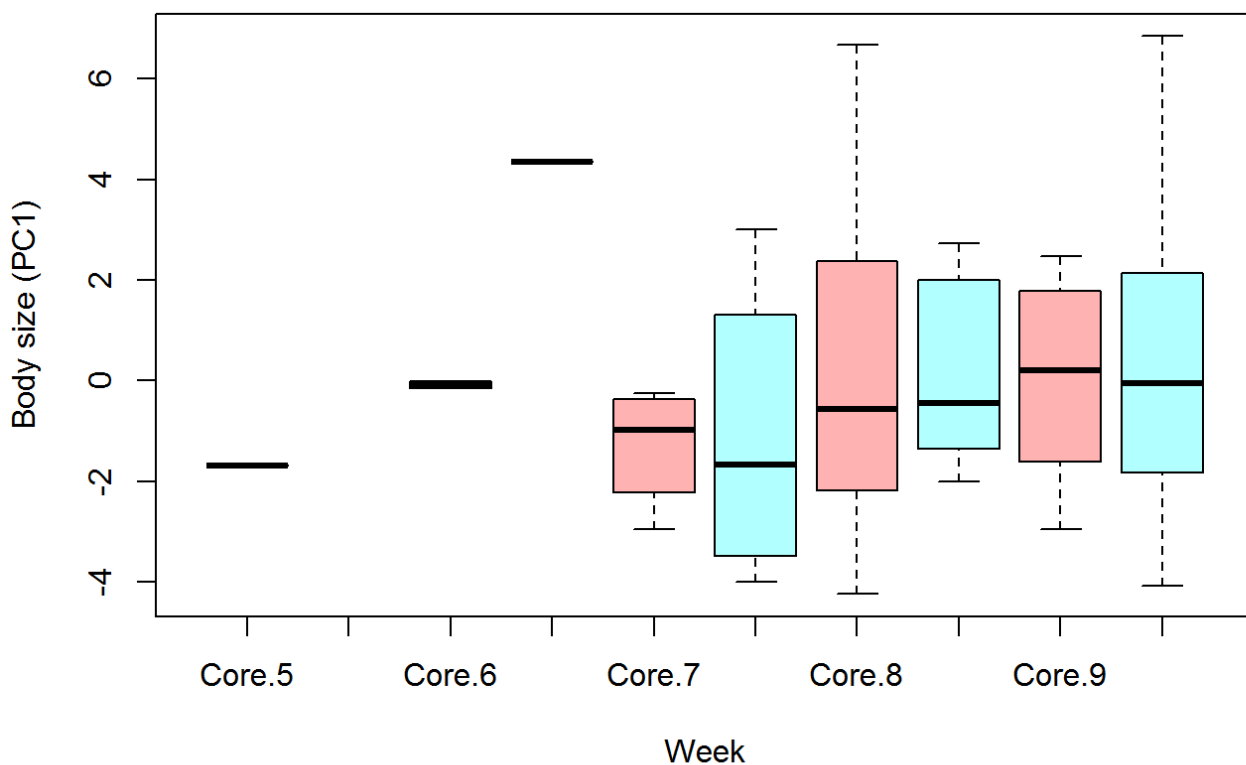
Simulated values, red line = fitted model. $p\text{-value (two.sided)} = 0.608$


```
##
## DHARMA nonparametric dispersion test via sd of residuals fitted
## vs. simulated
##
## data: simulationOutput
## ratioObsSim = 0.96122, p-value = 0.608
## alternative hypothesis: two.sided
```

#note that t-values greater than +/- 1.96 can be considered significant (Luke, 2017)"Evaluating significance in linear mixed-effects models in R"
#Random factors have a small effect on PC1 (according to Std.Dev.)

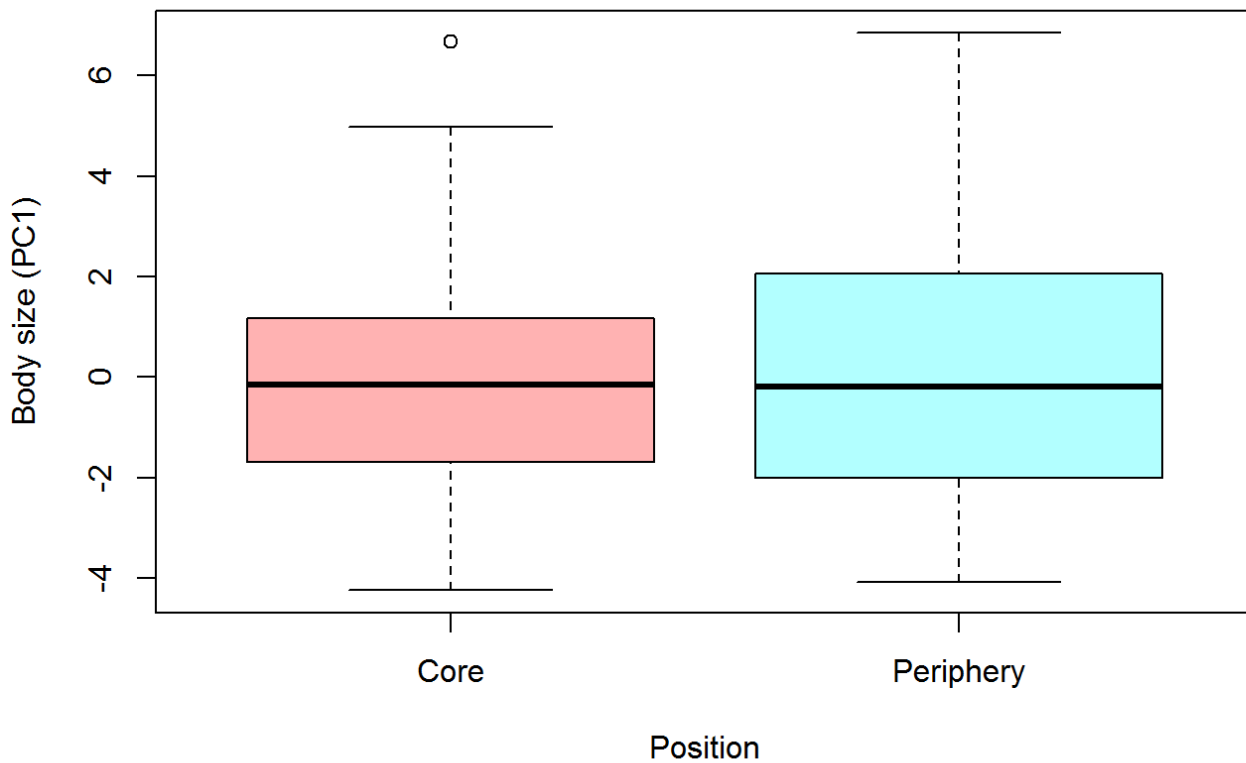
*##Position*Week has a significant effect on overall bodysize, but not position.*
 boxplot(PC1~Position*Week1, xaxs=F,data=Mesometadataset, main= "Metamorph body size per week in the mesocosms", xlab="Week", ylab="Body size (PC1)", col= rainbow(2, alpha = 0.3), notch= F)

Metamorph body size per week in the mesocosms



```
boxplot(PC1~Position, xaxs=F, main= "Metamorph body size in the mesocosm", xlab="Position", ylab="Body size (PC1)", col= rainbow(2, alpha = 0.3), notch= F, data=Mesometadataset)
```

Metamorph body size in the mesocosm



SVL

```
SVL1c<- lmer(pSVL ~ Week1 + (1|Site/Clutchx), data=Mesometadataset)
```

```
## boundary (singular) fit: see ?isSingular
```

```
SVL2c<- lmer(PC1 ~ Position + Week1 + (1|Site/Clutchx), data = Mesometadataset)
#Position*Week1 was dropped due to rank deficiency
```

```
SVL4c<- lmer(pSVL ~ Position + (1|Site/Clutchx), data = Mesometadataset)
```

```
SVL5c<- lmer(pSVL ~ 1 + (1|Site/Clutchx), data=Mesometadataset)
```

```
## boundary (singular) fit: see ?isSingular
```

```
AIC(SVL1c, SVL2c, SVL4c, SVL5c)
```

```
##      df      AIC
## SVL1c  8 271.7189
## SVL2c  9 280.4857
## SVL4c  5 277.3234
## SVL5c  4 276.5686
```

```
AICctab(SVL1c, SVL2c, SVL4c, SVL5c)
```

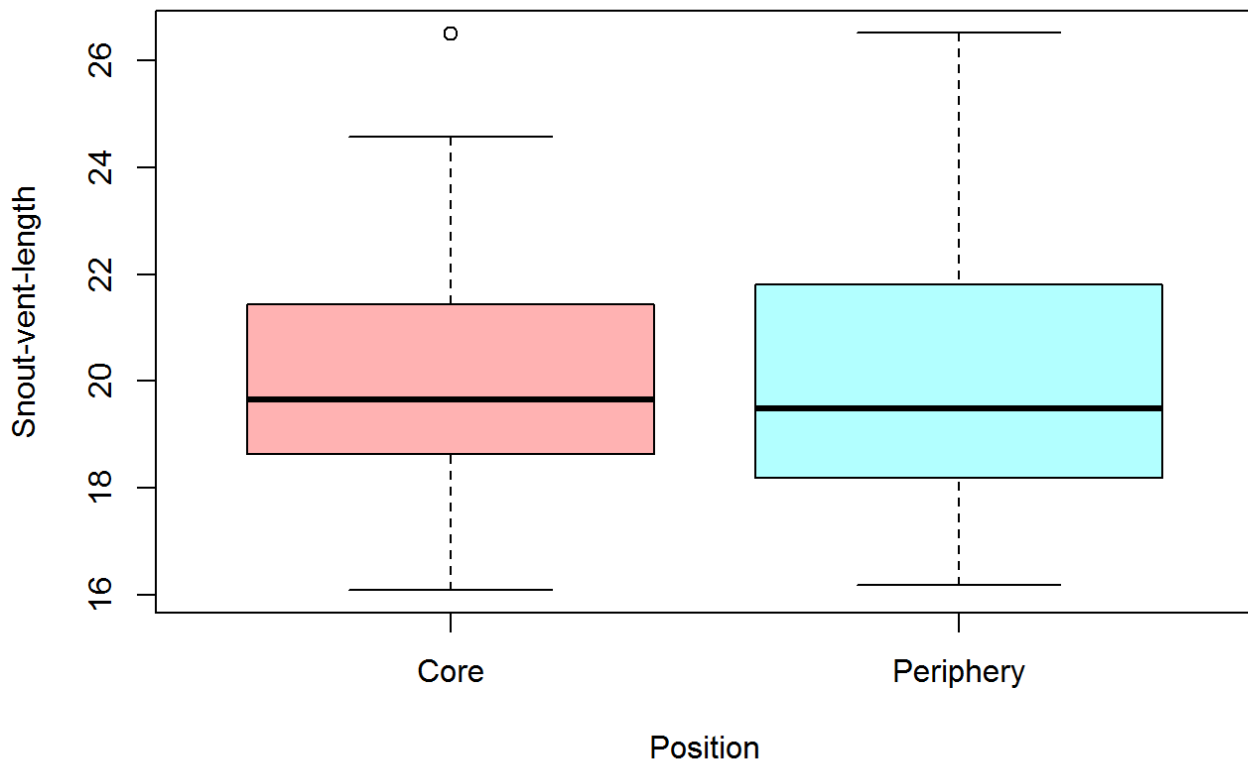
```
##          dAICc df
## SVL1c 0.0    8
## SVL5c 2.7    4
## SVL4c 3.8    5
## SVL2c 9.6    9
```

```
#SVL1c
summary(SVL1c)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: pSVL ~ Week1 + (1 | Site/Clutchx)
## Data: Mesometadataset
##
## REML criterion at convergence: 255.7
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -1.6935 -0.7236 -0.0899  0.5508  2.5078
##
## Random effects:
## Groups      Name                Variance Std.Dev.
## Clutchx:Site (Intercept) 0.2178     0.4667
## Site         (Intercept) 0.0000     0.0000
## Residual                5.8834     2.4256
## Number of obs: 58, groups: Clutchx:Site, 19; Site, 6
##
## Fixed effects:
##              Estimate Std. Error t value
## (Intercept)   18.705      2.465    7.589
## Week16         2.264      2.846    0.796
## Week17         0.151      2.578    0.059
## Week18         1.481      2.512    0.589
## Week19         1.794      2.520    0.712
##
## Correlation of Fixed Effects:
##      (Intr) Week16 Week17 Week18
## Week16 -0.866
## Week17 -0.953  0.828
## Week18 -0.978  0.849  0.934
## Week19 -0.976  0.847  0.932  0.957
## convergence code: 0
## boundary (singular) fit: see ?isSingular
```

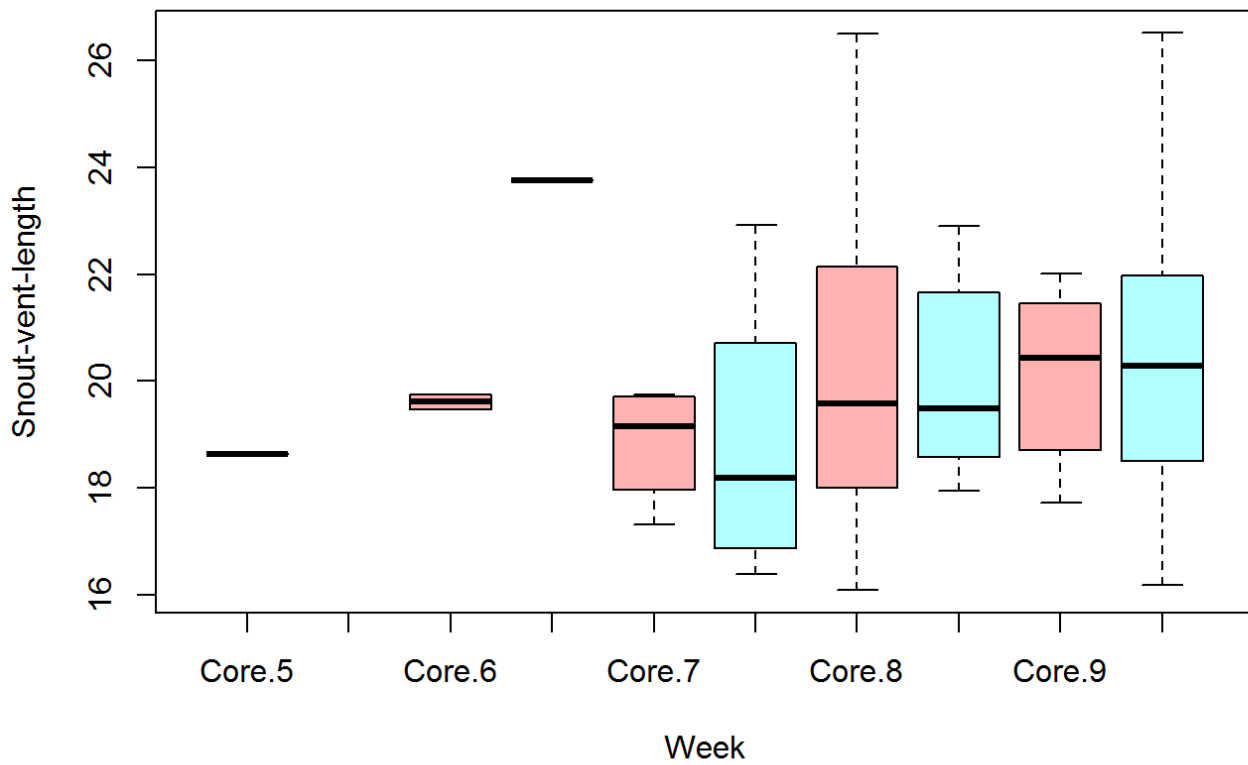
```
boxplot(Snoutventlength~Position, xaxs=F,data=Mesometadataset, main= "Metamorph SVL
in the mesocosms", xlab="Position", ylab="Snout-vent-length", col= rainbow(2, alpha
= 0.3), notch= F)
```

Metamorph SVL in the mesocosms

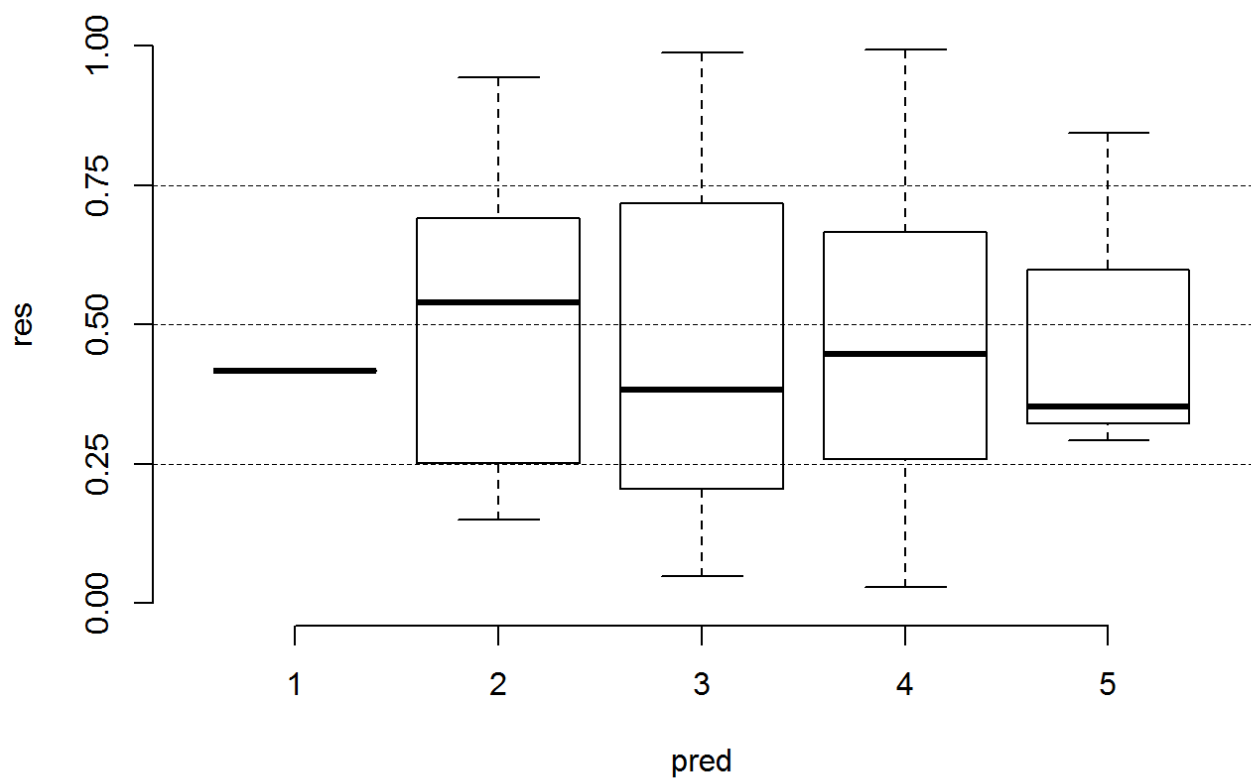


```
boxplot(Snoutventlength~Position*Week1, xaxs=F,data=Mesometadataset, main= "Metamorph SVL in the mesocosms per week", xlab="Week", ylab="Snout-vent-length", col= rainbow(2, alpha = 0.3), notch= F)
```

Metamorph SVL in the mesocosms per week

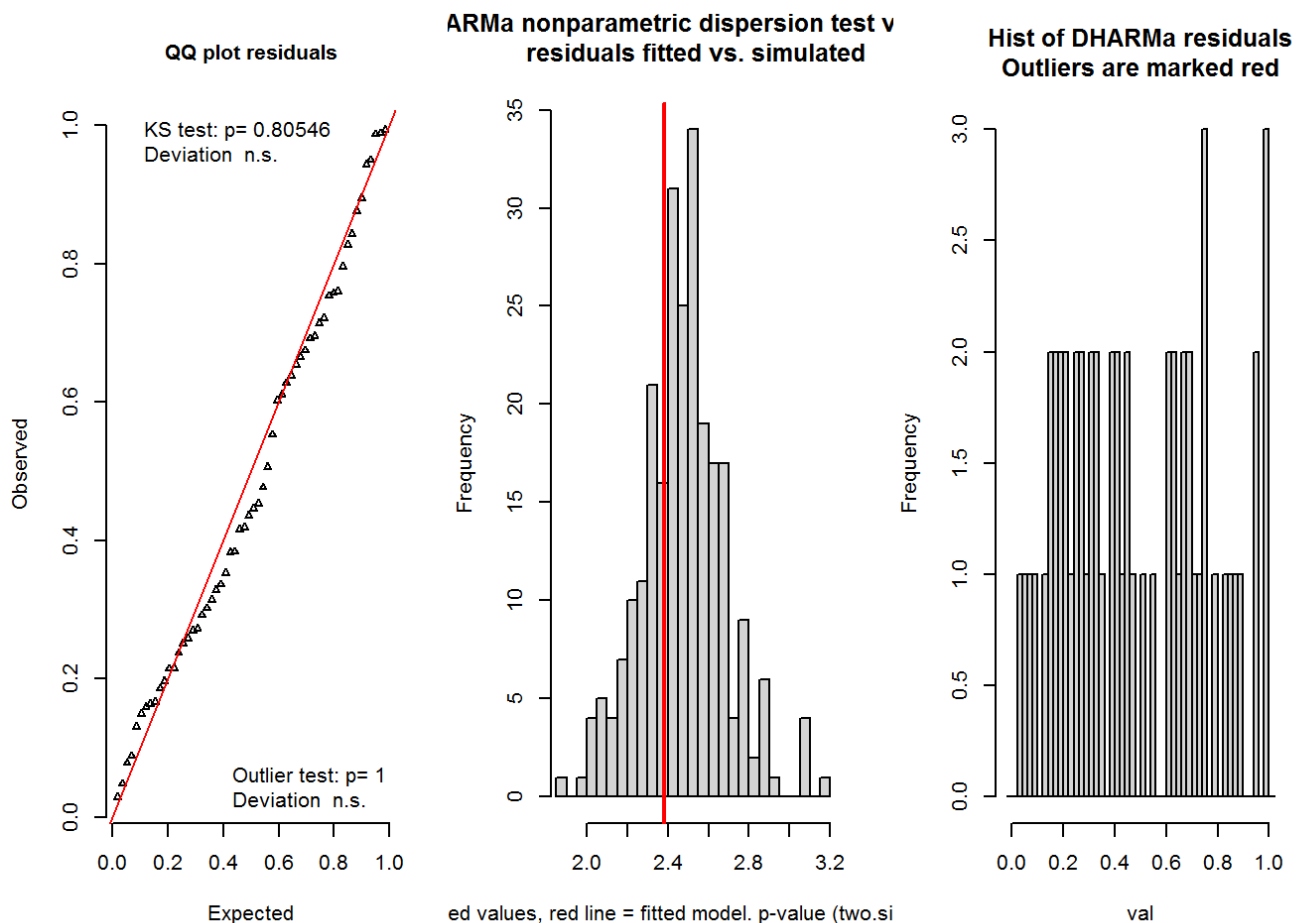


```
rr=simulateResiduals(SVL1c) ### simulate residuals  
plotResiduals(rr) ### qqplot simulated vs observed
```



```
## NULL
```

```
testResiduals(rr)###test qqplot
```

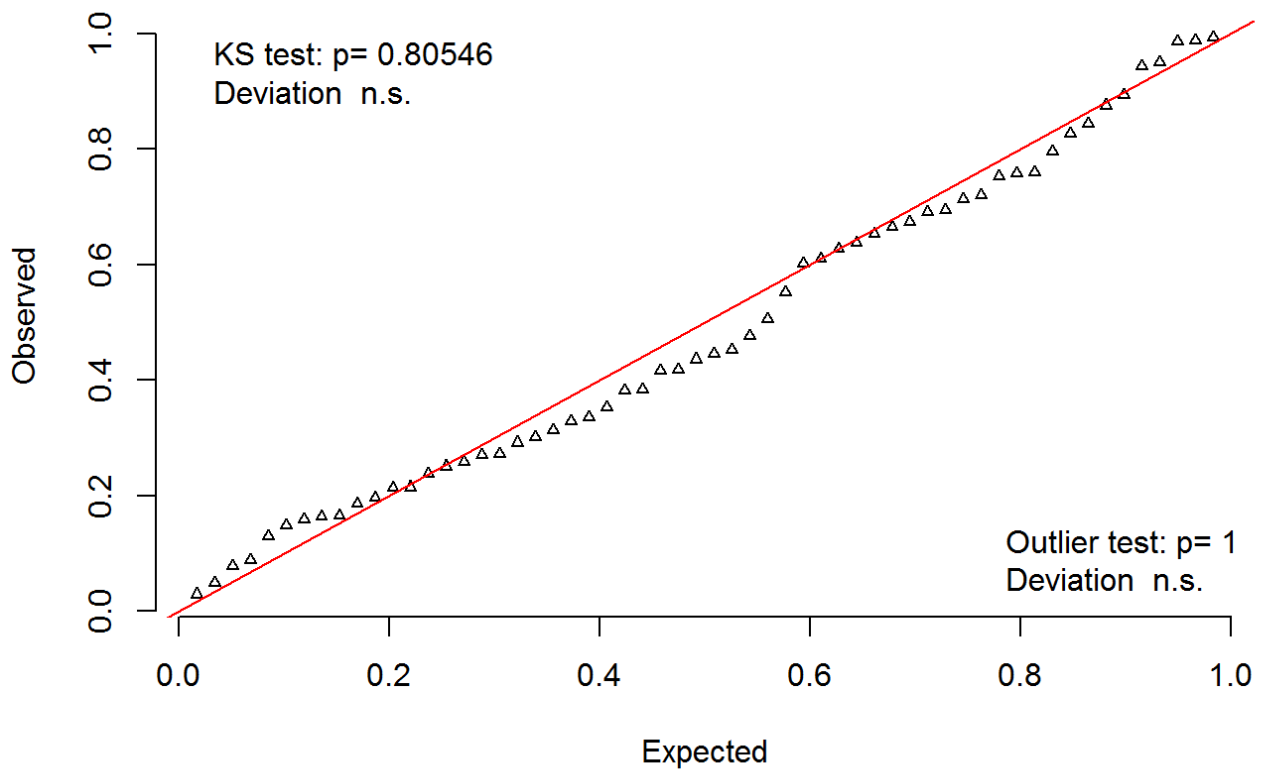


```
## $uniformity
##
## One-sample Kolmogorov-Smirnov test
##
## data: simulationOutput$scaledResiduals
## D = 0.081537, p-value = 0.8055
## alternative hypothesis: two-sided
##
##
## $dispersion
##
## DHARMA nonparametric dispersion test via sd of residuals fitted
## vs. simulated
##
## data: simulationOutput
## ratioObsSim = 0.96127, p-value = 0.608
## alternative hypothesis: two.sided
##
##
## $outliers
##
## DHARMA outlier test based on exact binomial test
##
## data: simulationOutput
## outLow = 0.0000000, outHigh = 0.0000000, nobs = 58.0000000, freqH0
## = 0.0039841, p-value = 1
## alternative hypothesis: two.sided
```

```
## $uniformity
##
## One-sample Kolmogorov-Smirnov test
##
## data: simulationOutput$scaledResiduals
## D = 0.081537, p-value = 0.8055
## alternative hypothesis: two-sided
##
##
## $dispersion
##
## DHARMA nonparametric dispersion test via sd of residuals fitted
## vs. simulated
##
## data: simulationOutput
## ratioObsSim = 0.96127, p-value = 0.608
## alternative hypothesis: two.sided
##
##
## $outliers
##
## DHARMA outlier test based on exact binomial test
##
## data: simulationOutput
## outLow = 0.0000000, outHigh = 0.0000000, nobs = 58.0000000, freqH0
## = 0.0039841, p-value = 1
## alternative hypothesis: two.sided
```

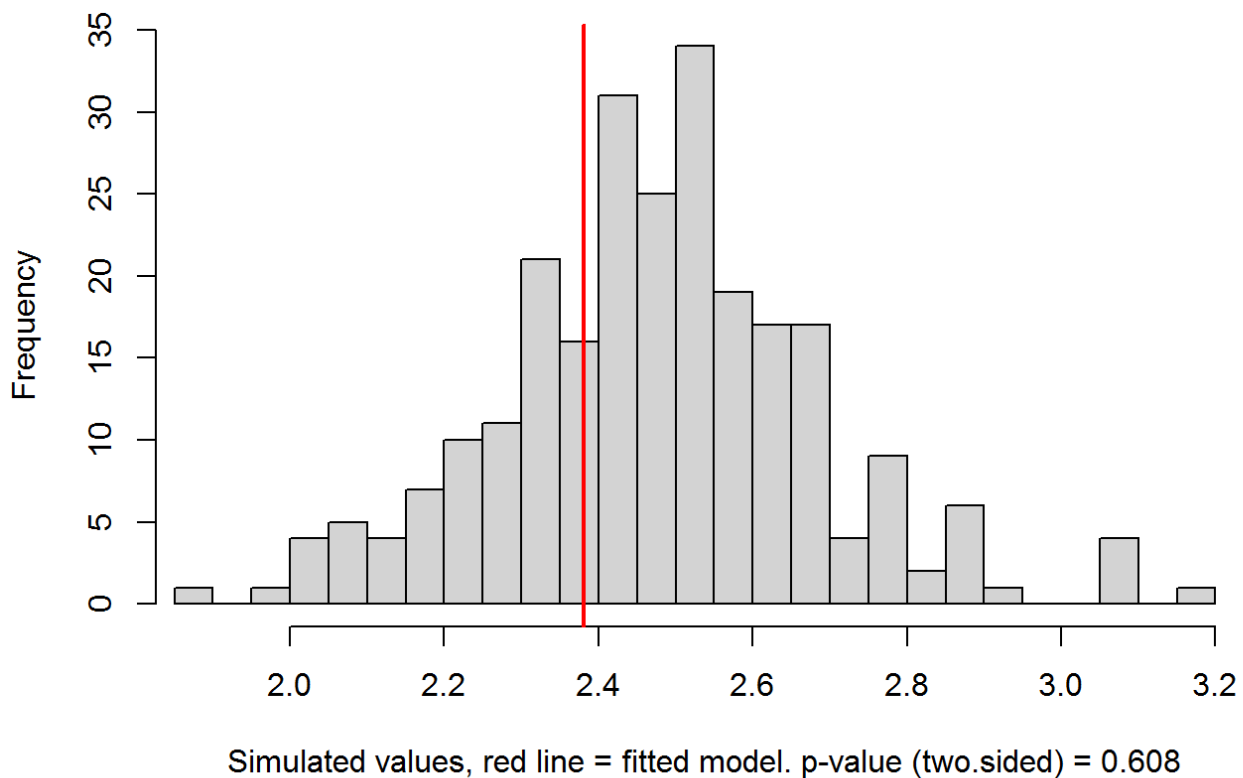
```
plotQQunif(rr, testUniformity = T) ### plot
```


QQ plot residuals



```
testDispersion(rr,plot=T) ### test
```

DHARMA nonparametric dispersion test via sd of residuals fitted vs. simulated



```
##
## DHARMA nonparametric dispersion test via sd of residuals fitted
## vs. simulated
##
## data:  simulationOutput
## ratioObsSim = 0.96127, p-value = 0.608
## alternative hypothesis: two.sided
```

Femur

```
#Position*Week1 was dropped due to rank deficiency
F2c<- lmer(pFemurrel ~ Position + Week1 + (1|Site/Clutchx), data = Mesometadataset)
F3c<- lmer(pFemurrel ~ Week1 + (1|Site/Clutchx), data=Mesometadataset)
```

```
## boundary (singular) fit: see ?isSingular
```

```
F4c<- lmer(pFemurrel ~ Position + (1|Site/Clutchx), data = Mesometadataset)
```

```
## boundary (singular) fit: see ?isSingular
```

```
F5c<- lmer(pFemurrel ~ 1 + (1|Site/Clutchx), data=Mesometadataset)
```

```
## boundary (singular) fit: see ?isSingular
```

```
AIC(F2c, F3c, F4c, F5c)
```

```
##      df      AIC
## F2c  9 171.9619
## F3c  8 169.2016
## F4c  5 171.3915
## F5c  4 168.5226
```

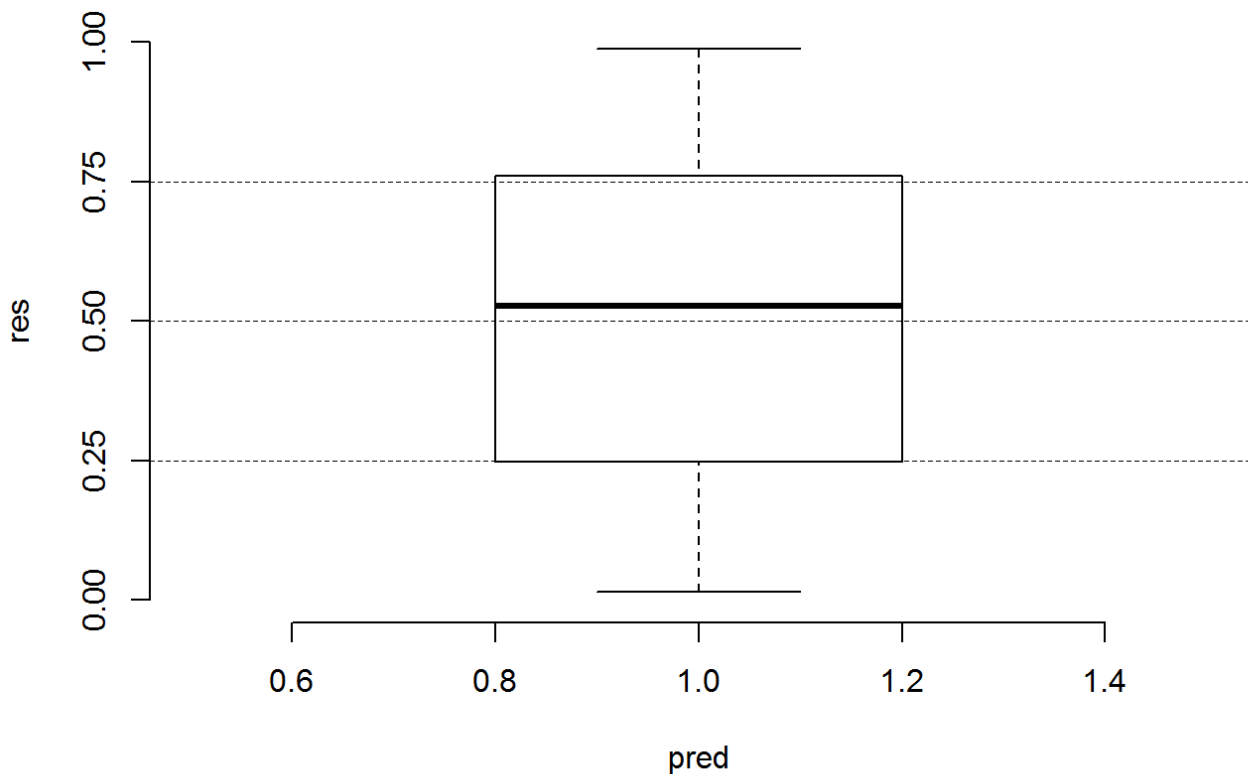
```
AICctab (F2c, F3c, F4c, F5c)
```

```
##      dAICc df
## F5c 0.0   4
## F3c 2.9   8
## F4c 3.3   5
## F2c 6.4   9
```

```
summary(F5c)
```

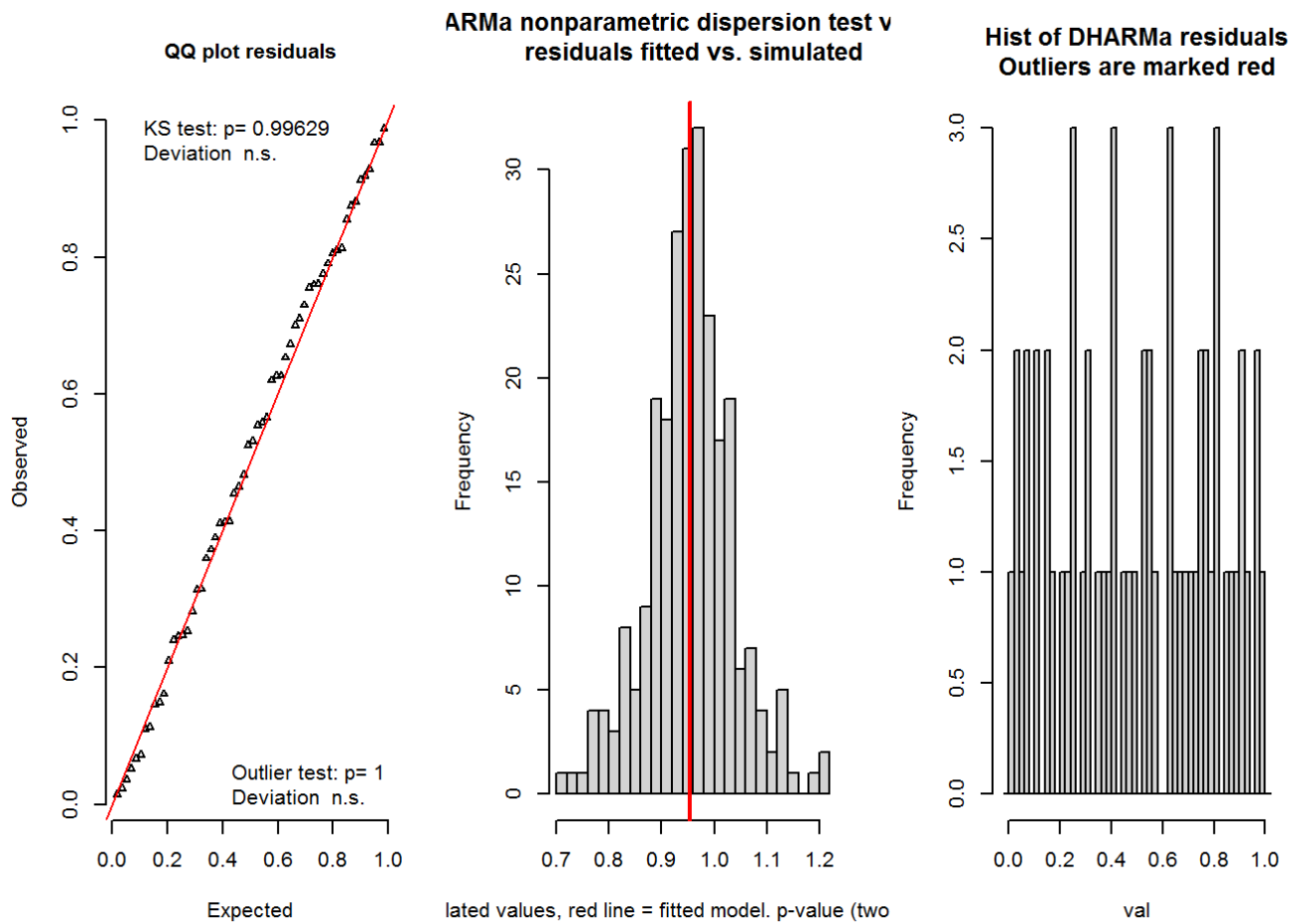
```
## Linear mixed model fit by REML ['lmerMod']
## Formula: pFemurrel ~ 1 + (1 | Site/Clutchx)
## Data: Mesometadataset
##
## REML criterion at convergence: 160.5
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -2.45825 -0.66363  0.02087  0.69494  2.08877
##
## Random effects:
## Groups      Name                Variance Std.Dev.
## Clutchx:Site (Intercept) 0.0000    0.0000
## Site         (Intercept) 0.0000    0.0000
## Residual                    0.9113    0.9546
## Number of obs: 58, groups: Clutchx:Site, 19; Site, 6
##
## Fixed effects:
##              Estimate Std. Error t value
## (Intercept) -0.04117    0.12535  -0.328
## convergence code: 0
## boundary (singular) fit: see ?isSingular
```

```
rr=simulateResiduals(F5c) ### simulate residuals
plotResiduals(rr) ### qqplot simulated vs observed
```



```
## NULL
```

```
testResiduals(rr)###test qqplot
```

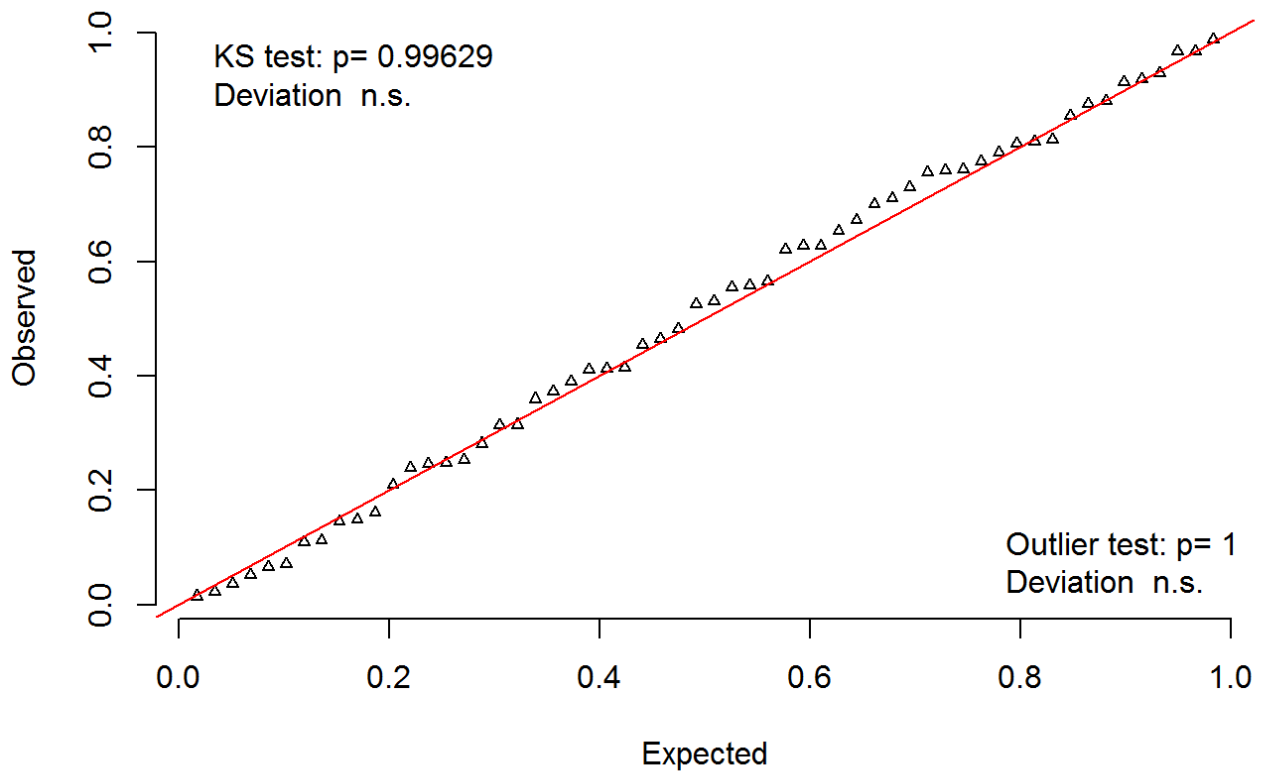


```
## $uniformity
##
## One-sample Kolmogorov-Smirnov test
##
## data: simulationOutput$scaledResiduals
## D = 0.051051, p-value = 0.9963
## alternative hypothesis: two-sided
##
##
## $dispersion
##
## DHARMA nonparametric dispersion test via sd of residuals fitted
## vs. simulated
##
## data: simulationOutput
## ratioObsSim = 0.99958, p-value = 1
## alternative hypothesis: two.sided
##
##
## $outliers
##
## DHARMA outlier test based on exact binomial test
##
## data: simulationOutput
## outLow = 0.0000000, outHigh = 0.0000000, nobs = 58.0000000, freqH0
## = 0.0039841, p-value = 1
## alternative hypothesis: two.sided
```

```
## $uniformity
##
## One-sample Kolmogorov-Smirnov test
##
## data: simulationOutput$scaledResiduals
## D = 0.051051, p-value = 0.9963
## alternative hypothesis: two-sided
##
##
## $dispersion
##
## DHARMA nonparametric dispersion test via sd of residuals fitted
## vs. simulated
##
## data: simulationOutput
## ratioObsSim = 0.99958, p-value = 1
## alternative hypothesis: two.sided
##
##
## $outliers
##
## DHARMA outlier test based on exact binomial test
##
## data: simulationOutput
## outLow = 0.0000000, outHigh = 0.0000000, nobs = 58.0000000, freqH0
## = 0.0039841, p-value = 1
## alternative hypothesis: two.sided
```

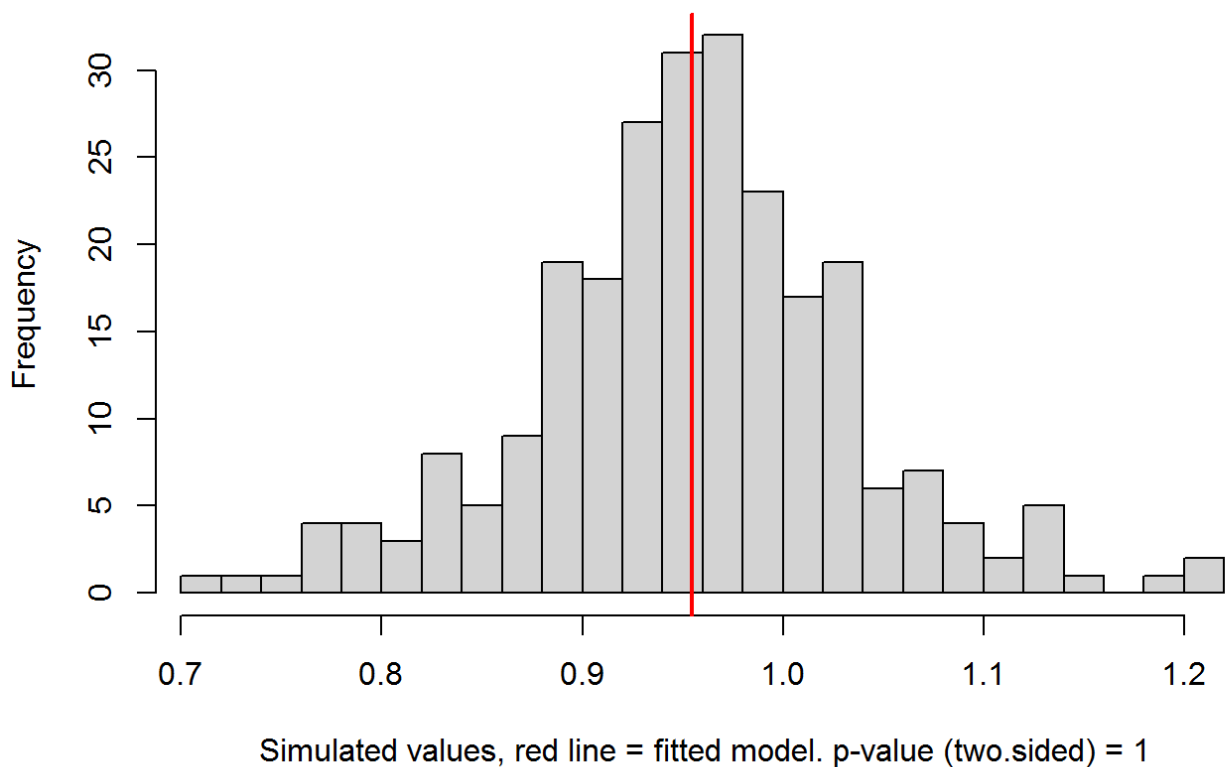
```
plotQQunif(rr, testUniformity = T) ### plot
```

QQ plot residuals



```
testDispersion(rr,plot=T) ### test
```

DHARMA nonparametric dispersion test via sd of residuals fitted vs. simulated

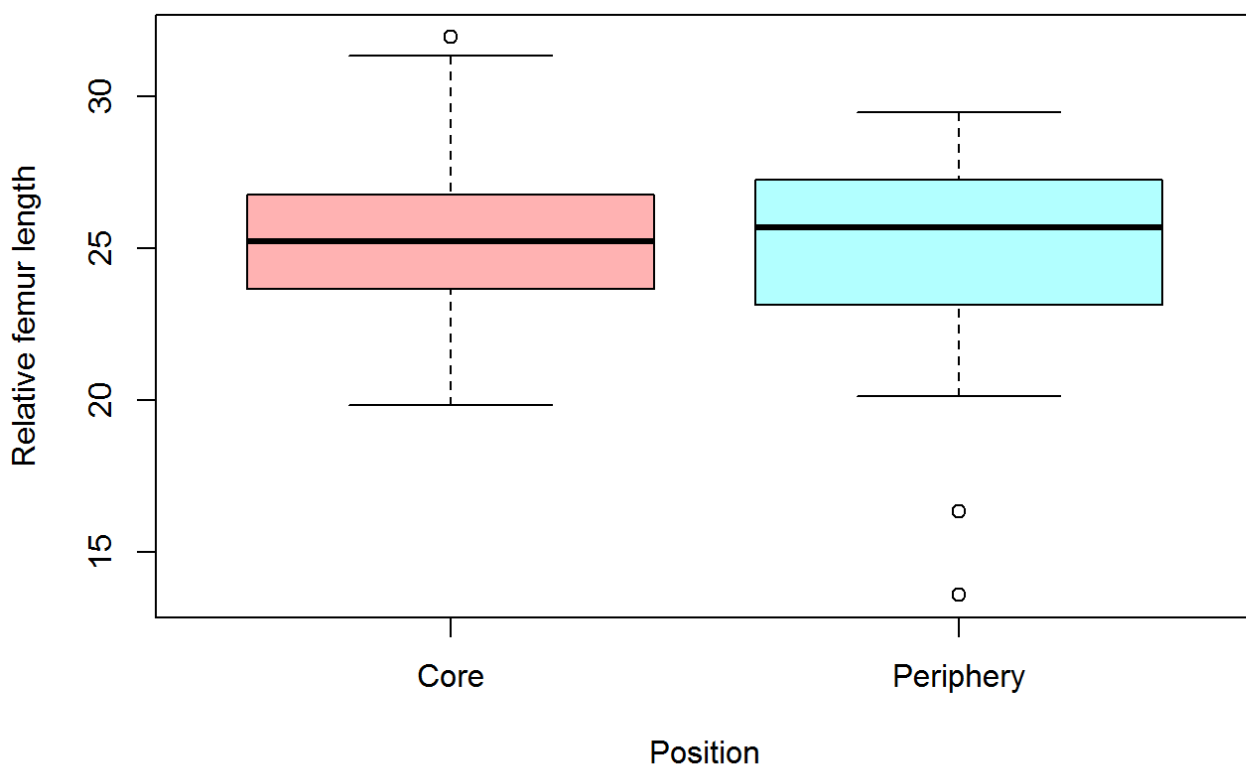


```
##
## DHARMA nonparametric dispersion test via sd of residuals fitted
## vs. simulated
##
## data:  simulationOutput
## ratioObsSim = 0.99958, p-value = 1
## alternative hypothesis: two.sided
```

```
#Null model was selected
```

```
boxplot(Femurrelative~Position, xaxs=F,data=Mesometadataset, main= "Metamorph relative femurlength in the mesocosms", xlab="Position", ylab="Relative femur length", col= rainbow(2, alpha = 0.3), notch= F)
```

Metamorph relative femurlength in the mesocosms



```
detach(Mesomet)
```

3.4. Metamorph NF stage 66- LAB

3.4.1. Defining variables.

Use *Labmetamorphs.csv* to analyse data.

```
Labmet=read.csv(file.choose(),header=T)
names(Labmet)
```



```
## [1] "Position"      "Site"           "Clutch"
## [4] "tad_num"        "tad_id"         "tank_id"
## [7] "age"           "weight"         "r"
## [10] "Snoutventlength" "Femur"          "Femurrelative"
## [13] "tibia"
```

```
attach(Labmet)
```

```
summary(Position) #fixed effect
```

```
##      Core Periphery
##      38          34
```

```
summary(age) #fixed effect
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  74.00   76.00   83.00   82.44   88.00   94.00
```

```
age1 <- as.factor(age) #as a discrete variable
summary(age1)
```

```
## 74 75 76 80 81 82 83 87 88 89 90 91 93 94
##  2  9 12  2  4  6 13  2  5  6  8  1  1  1
```

```
summary(Site) #Random effect
```

```
##      Bouille  Challones  Jardinerie    Massais    St.Paul    Toutiere
##           13           17           7           8           17           10
```

```
summary(Clutch) #Nested within site
```

```
## b1 b2 b3 b4 c1 c2 c3 c4 j1 j4 m1 m2 m3 m4 s1 s2 s3 s4 t2 t3
##  1  5  4  3  1  6  5  5  5  2  1  4  1  2  6  6  2  3  6  4
```

```
Labmet1<- cbind(Labmet, age1)
names(Labmet1)
```

```
## [1] "Position"      "Site"           "Clutch"
## [4] "tad_num"        "tad_id"         "tank_id"
## [7] "age"           "weight"         "r"
## [10] "Snoutventlength" "Femur"          "Femurrelative"
## [13] "tibia"         "age1"
```

3.2.2. Transforming response variables.

Use the *Labmet1* datasheet just created to measure normality.

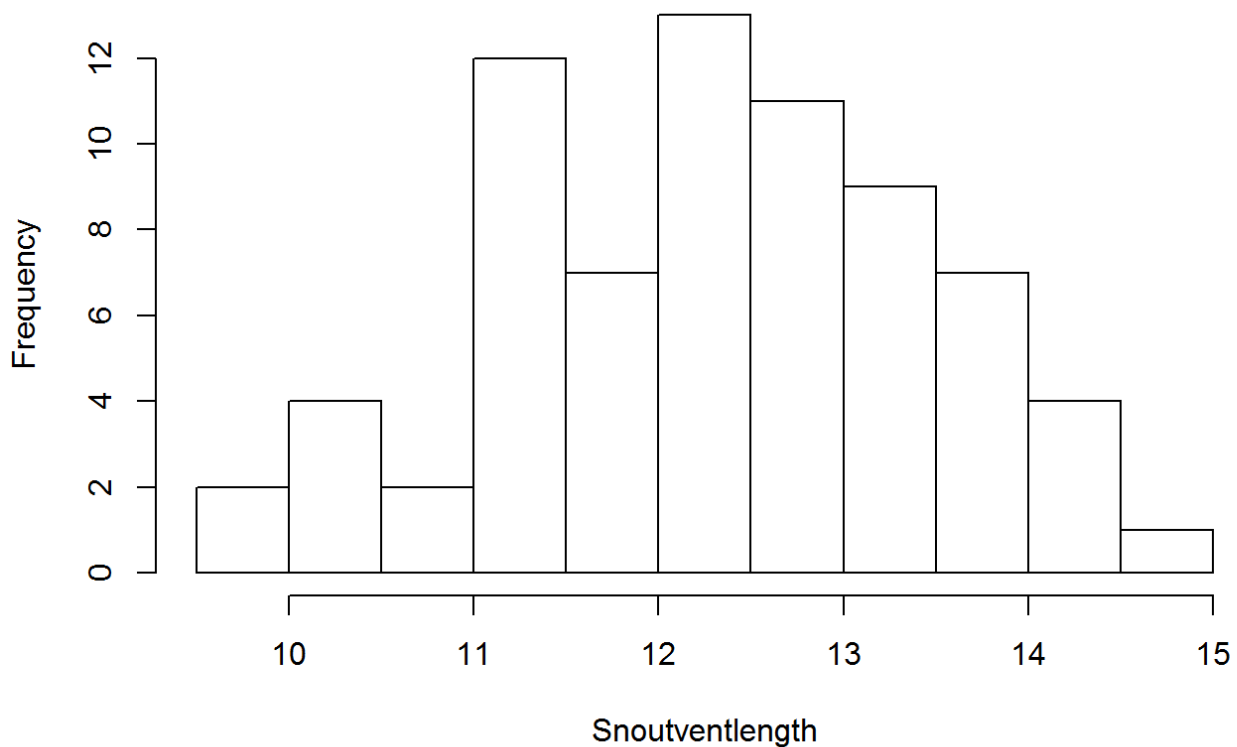
Snoutventlength

```
shapiro.test(Snoutventlength)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  Snoutventlength
## W = 0.98125, p-value = 0.3586
```

```
#p-value = 0.3586
hist(Snoutventlength)
```

Histogram of Snoutventlength



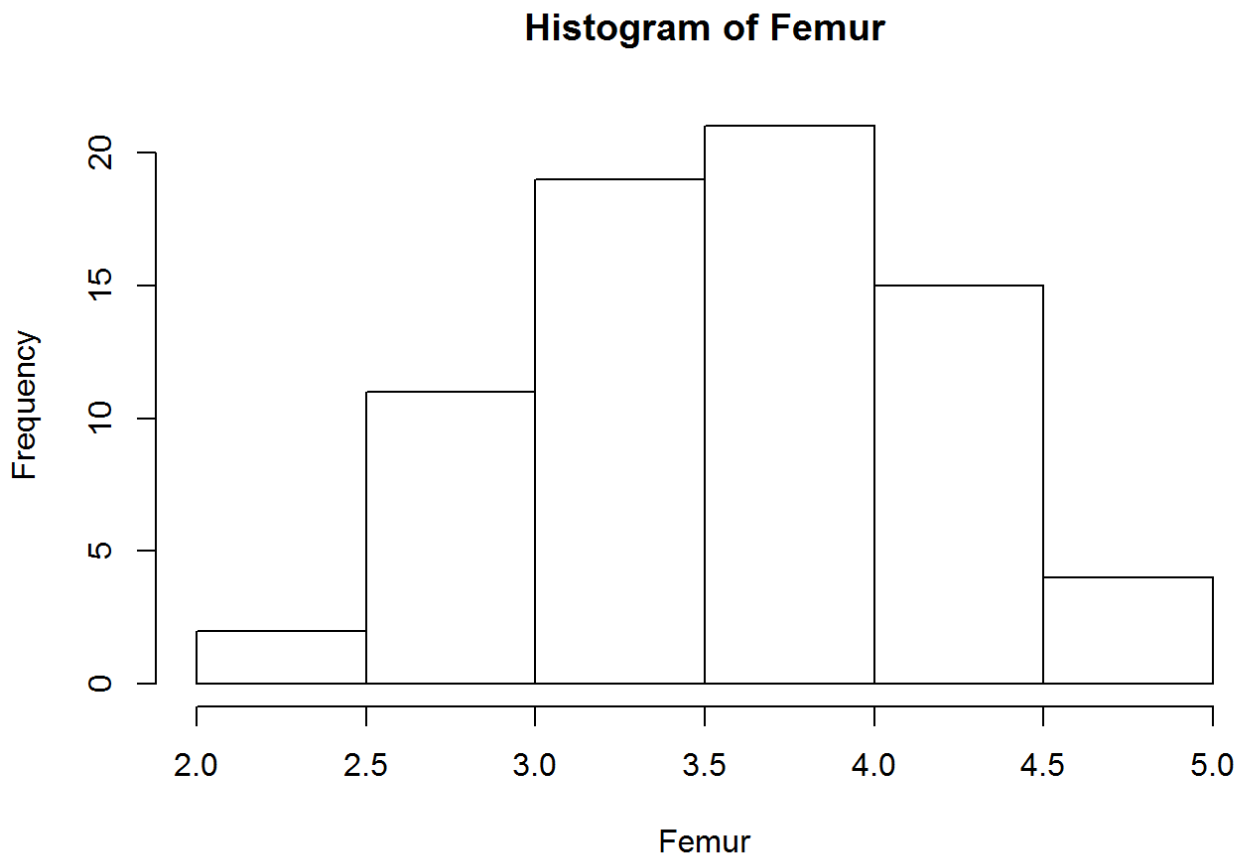
```
#no transformation
pSVL <- Snoutventlength
```

Femur

```
shapiro.test(Femur)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  Femur
## W = 0.98931, p-value = 0.8057
```

```
#p-value = 0.8057
hist(Femur)
```



```
#no transformation
pFemur<- Femur
```

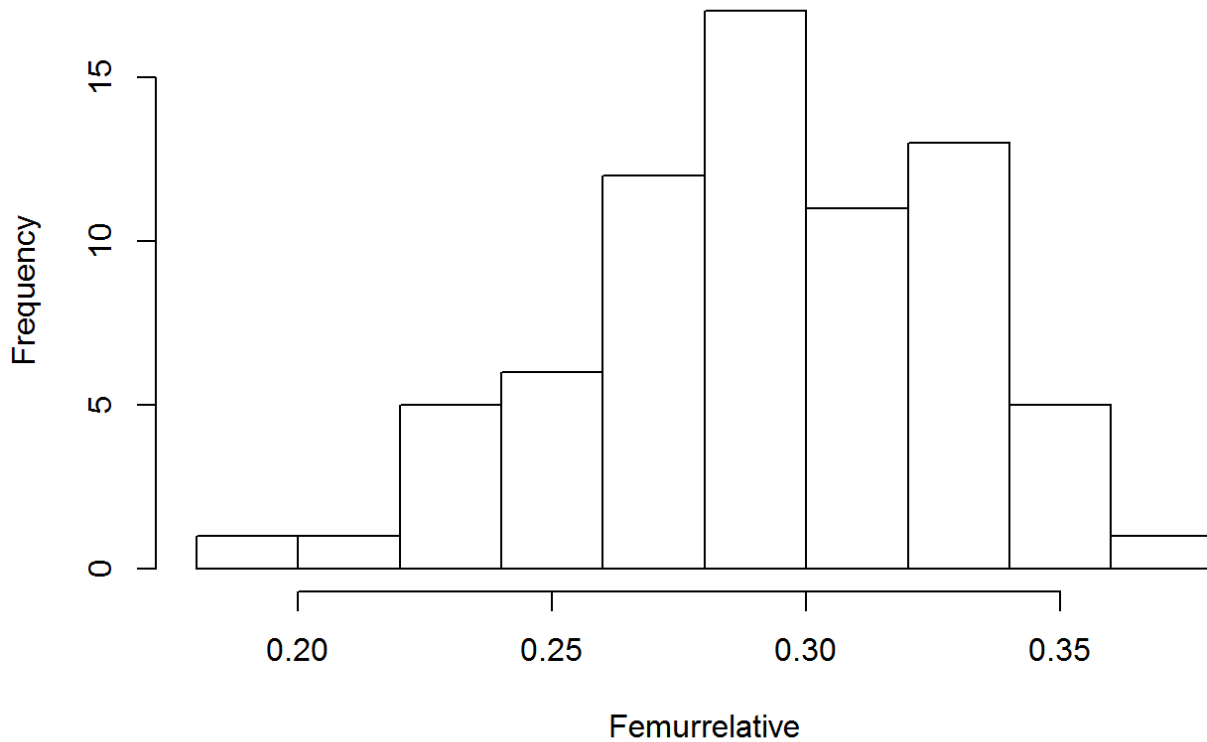
Relative Femur length

```
shapiro.test(Femurrelative)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  Femurrelative
## W = 0.9803, p-value = 0.3194
```

```
#p-value = 0.3194
hist(Femurrelative)
```

Histogram of Femurrelative



```
#no transformation

pFemurrel <- Femurrelative
```

Unfortunately, I did not take pictures of lab metamorphs from the side because they were still alive, so I did not measure body depth. So I am not performing a PCA on Lab metamorphs.

3.2.3. Model selection for SVL and relative femurlength

SVL

```
SVL1d<- lmer(pSVL ~ Position*age+ (1|Site/Clutch), data = Labmet1)
```

```
## boundary (singular) fit: see ?isSingular
```

```
SVL2d<- lmer(pSVL ~ Position + age1 + (1|Site/Clutch), data = Labmet1)
```

```
## boundary (singular) fit: see ?isSingular
```

```
SVL3d<- lmer(pSVL ~ age1 + (1|Site/Clutch), data=Labmet1)
```

```
## boundary (singular) fit: see ?isSingular
```

```
SVL4d<- lmer(pSVL ~ Position + (1|Site/Clutch), data = Labmet1)
```

```
## boundary (singular) fit: see ?isSingular
```

```
SVL5d<- lmer(pSVL ~ 1 + (1|Site/Clutch), data=Labmet1)
```

```
## boundary (singular) fit: see ?isSingular
```

```
AIC(SVL1d, SVL2d, SVL3d, SVL4d, SVL5d)
```

```
##      df      AIC
## SVL1d  7 238.0248
## SVL2d 18 229.6932
## SVL3d 17 231.6999
## SVL4d  5 226.3089
## SVL5d  4 229.7422
```

```
AICctab(SVL1d, SVL2d, SVL3d, SVL4d, SVL5d)
```

```
##      dAICc df
## SVL4d  0.0  5
## SVL5d  3.1  4
## SVL1d 12.6  7
## SVL2d 15.4 18
## SVL3d 15.8 17
```

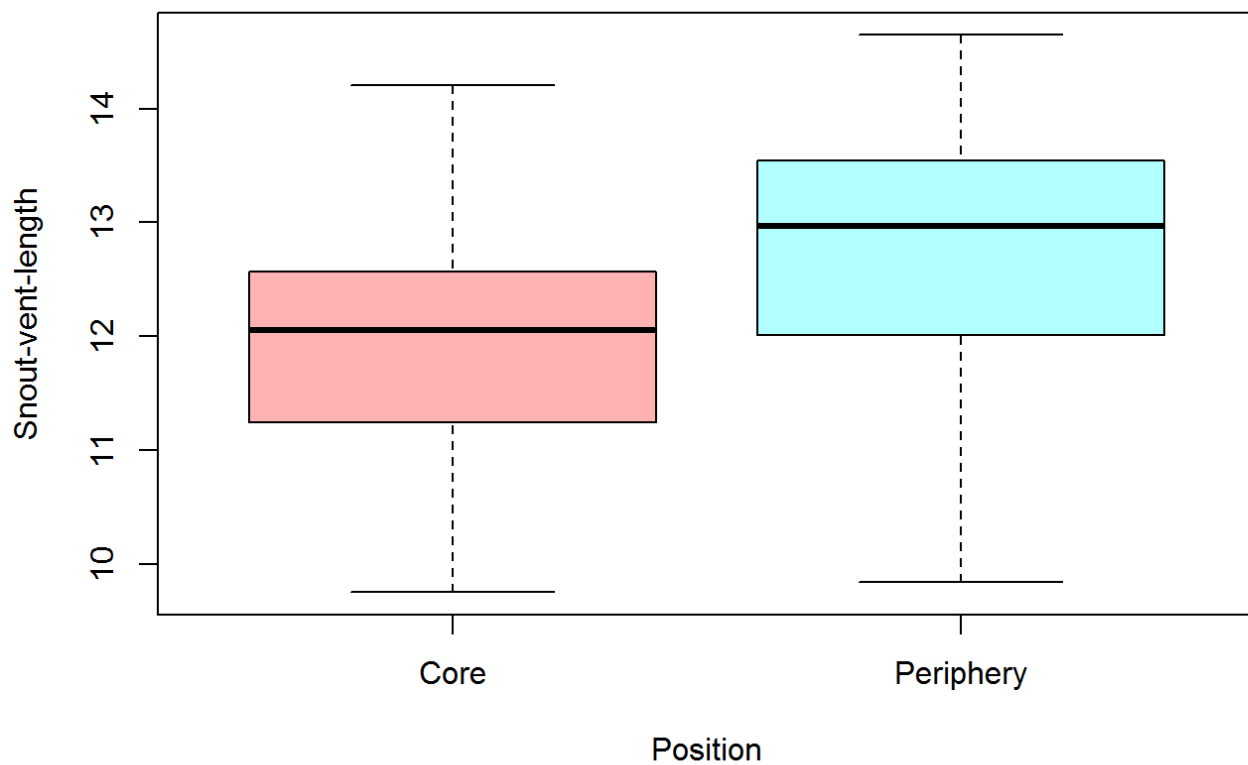
```
#SVL1c and SVL4d was chosen
summary(SVL4d)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: pSVL ~ Position + (1 | Site/Clutch)
## Data: Labmet1
##
## REML criterion at convergence: 216.3
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -2.7025 -0.6685  0.1797  0.6148  2.0084
##
## Random effects:
## Groups      Name                Variance Std.Dev.
## Clutch:Site (Intercept) 0.00000  0.0000
## Site        (Intercept) 0.03504  0.1872
## Residual                1.14220  1.0687
## Number of obs: 72, groups: Clutch:Site, 20; Site, 6
##
## Fixed effects:
##              Estimate Std. Error t value
## (Intercept)    11.9164    0.2061  57.819
## PositionPeriphery  0.8532    0.2982   2.861
##
## Correlation of Fixed Effects:
##              (Intr)
## PostnPrphry -0.691
## convergence code: 0
## boundary (singular) fit: see ?isSingular
```

#Position has an effect

```
boxplot(Snoutventlength~Position, xaxs=F,data=Labmet1, main= "Metamorph SVL in the
Lab", xlab="Position", ylab="Snout-vent-length", col= rainbow(2, alpha = 0.3), notc
h= F)
```

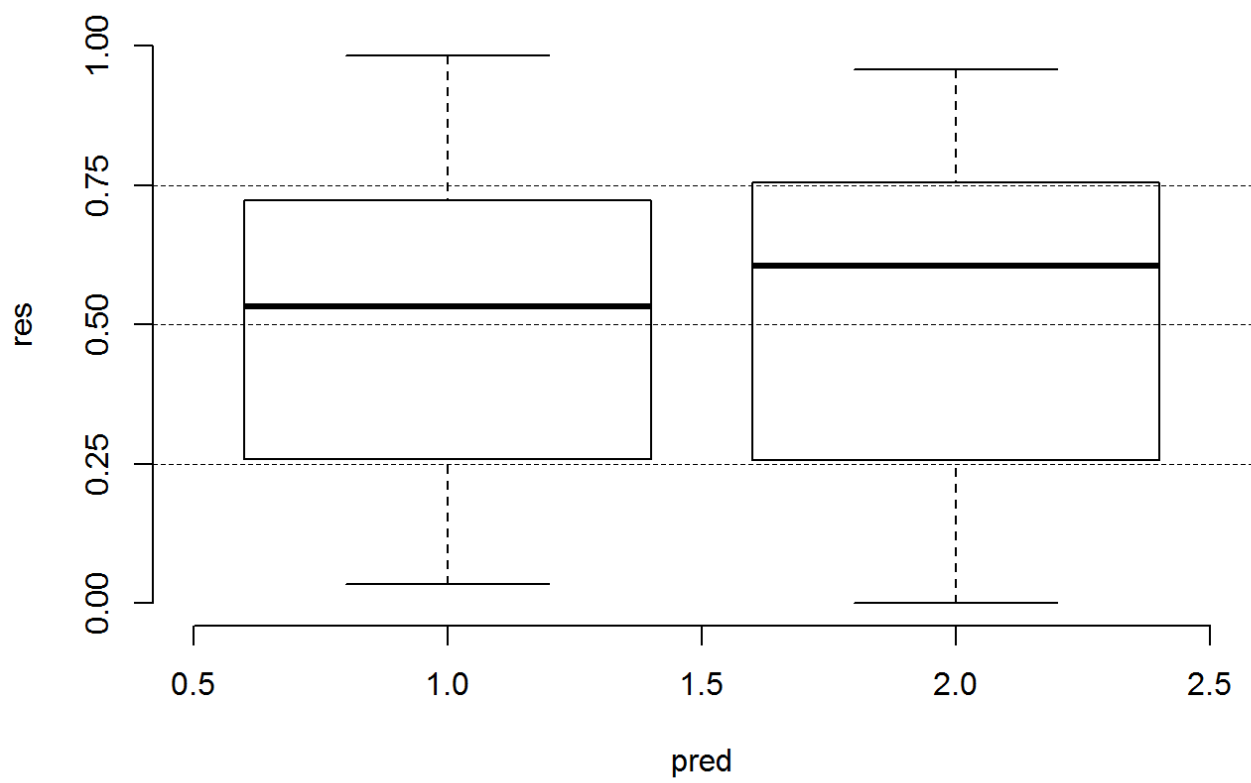
Metamorph SVL in the Lab



```
rr=simulateResiduals(SVL4d) ### simulate residuals
```

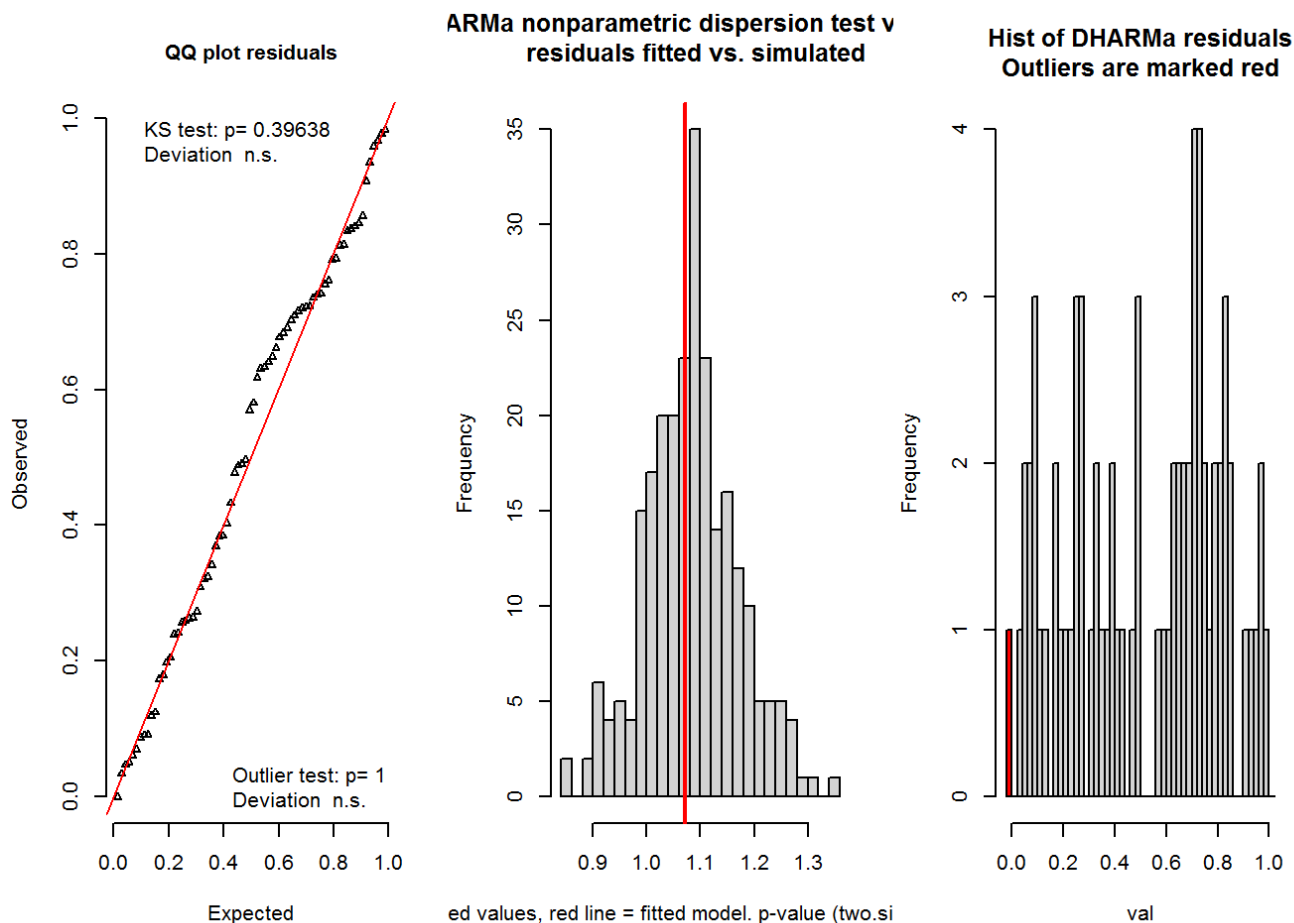
```
## Model family was recognized or set as continuous, but duplicate values were detected in the response. Consider if you are fitting an appropriate model.
```

```
plotResiduals(rr) ### qqplot simulated vs observed
```



```
## NULL
```

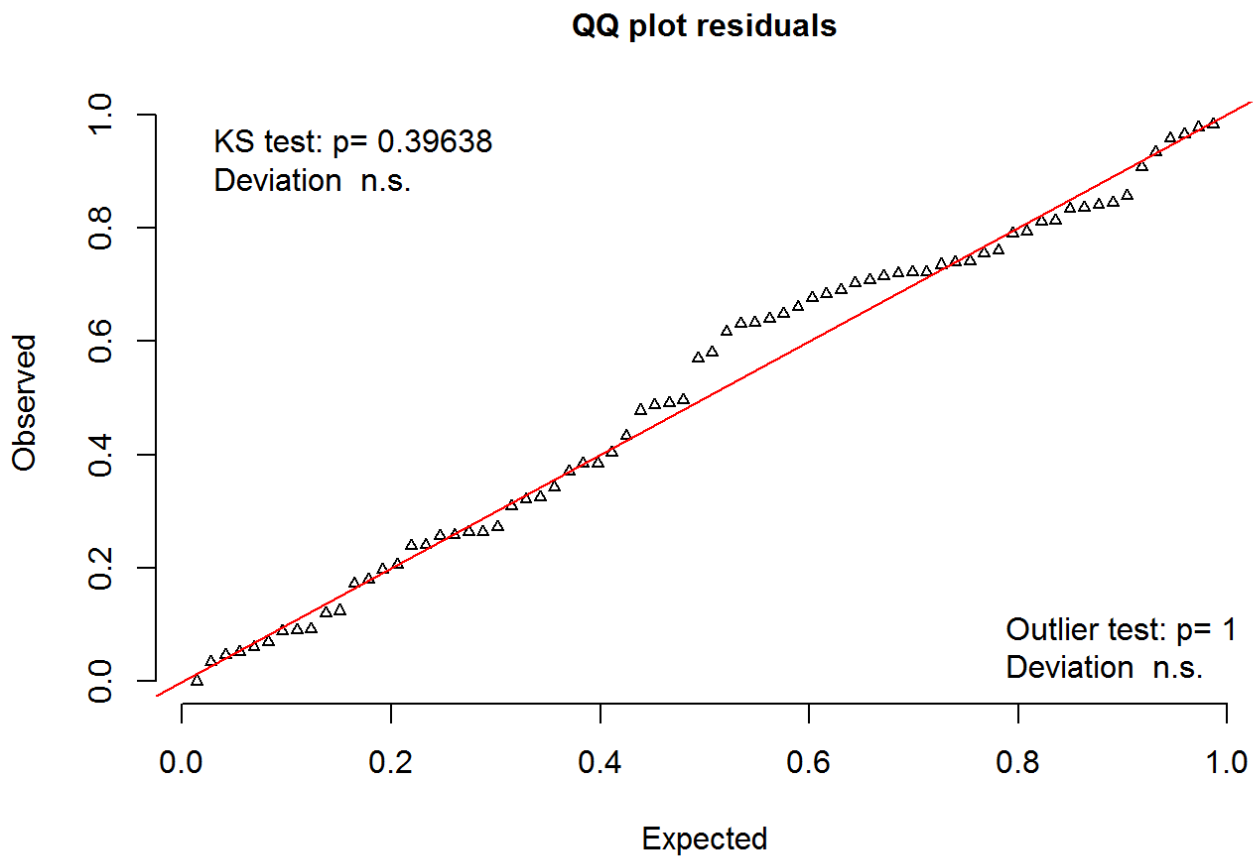
```
testResiduals(rr)###test qqplot
```

```
## $uniformity
##
## One-sample Kolmogorov-Smirnov test
##
## data: simulationOutput$scaledResiduals
## D = 0.10351, p-value = 0.3964
## alternative hypothesis: two-sided
##
##
## $dispersion
##
## DHARMA nonparametric dispersion test via sd of residuals fitted
## vs. simulated
##
## data: simulationOutput
## ratioObsSim = 0.99018, p-value = 0.896
## alternative hypothesis: two.sided
##
##
## $outliers
##
## DHARMA outlier test based on exact binomial test
##
## data: simulationOutput
## outLow = 1.0000000, outHigh = 0.0000000, nobs = 72.0000000, freqH0
## = 0.0039841, p-value = 1
## alternative hypothesis: two.sided
```

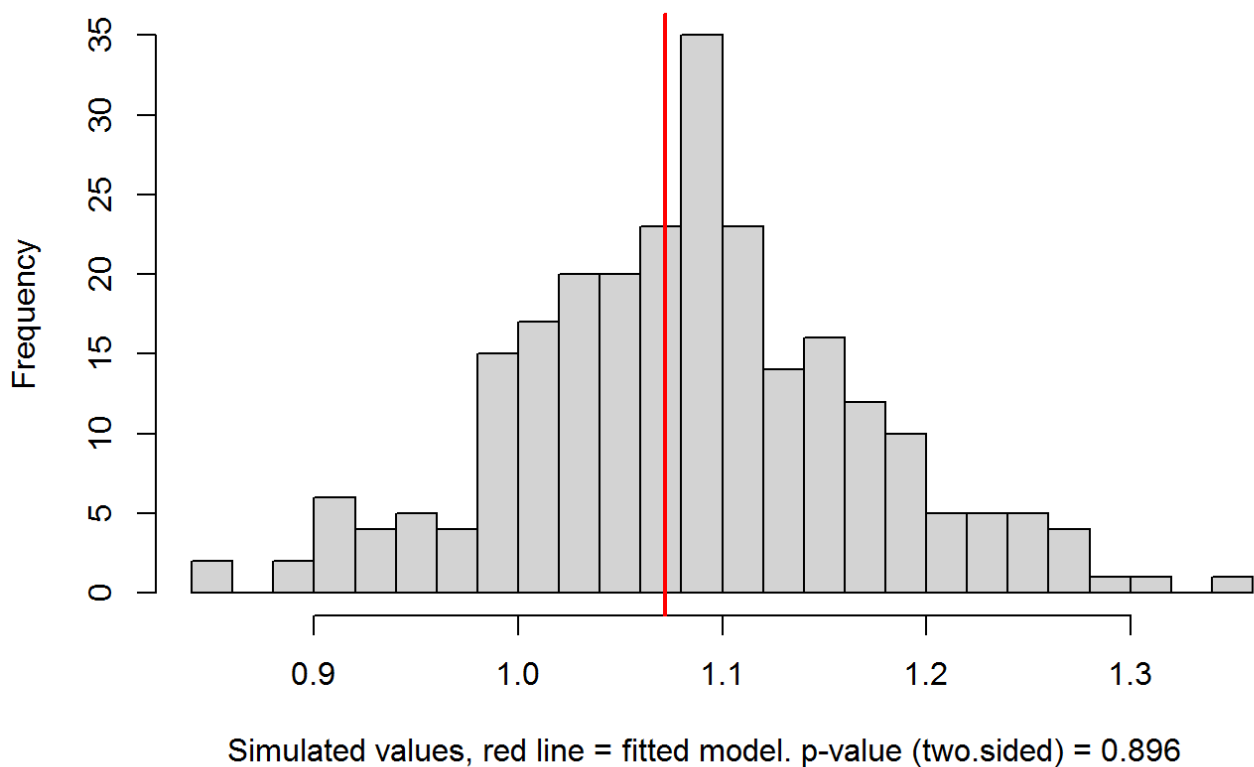
```
## $uniformity
##
## One-sample Kolmogorov-Smirnov test
##
## data: simulationOutput$scaledResiduals
## D = 0.10351, p-value = 0.3964
## alternative hypothesis: two-sided
##
##
## $dispersion
##
## DHARMA nonparametric dispersion test via sd of residuals fitted
## vs. simulated
##
## data: simulationOutput
## ratioObsSim = 0.99018, p-value = 0.896
## alternative hypothesis: two.sided
##
##
## $outliers
##
## DHARMA outlier test based on exact binomial test
##
## data: simulationOutput
## outLow = 1.0000000, outHigh = 0.0000000, nobs = 72.0000000, freqH0
## = 0.0039841, p-value = 1
## alternative hypothesis: two.sided
```

```
plotQQunif(rr, testUniformity = T) ### plot
```



```
testDispersion(rr,plot=T) ### test
```

DHARMA nonparametric dispersion test via sd of residuals fitted vs. simulated



```
##
## DHARMA nonparametric dispersion test via sd of residuals fitted
## vs. simulated
##
## data:  simulationOutput
## ratioObsSim = 0.99018, p-value = 0.896
## alternative hypothesis: two.sided
```

Femur

```
#Position*age1 is dropped due to rank deficiency
```

```
F2d<- lmer(pFemurrel ~ Position + age1 + (1|Site/Clutch), data = Labmet1)
```

```
## boundary (singular) fit: see ?isSingular
```

```
F3d<- lmer(pFemurrel ~ age1 + (1|Site/Clutch), data=Labmet1)
```

```
## boundary (singular) fit: see ?isSingular
```

```
F4d<- lmer(pFemurrel ~ Position + (1|Site/Clutch), data = Labmet1)
```

```
## boundary (singular) fit: see ?isSingular
```

```
F5d<- lmer(pFemurrel ~ 1 + (1|Site/Clutch), data=Labmet1)
```

```
## boundary (singular) fit: see ?isSingular
```

```
AIC( F2d, F3d, F4d, F5d)
```

```
##      df      AIC
## F2d 18 -165.6036
## F3d 17 -173.1373
## F4d  5 -244.3598
## F5d  4 -252.5570
```

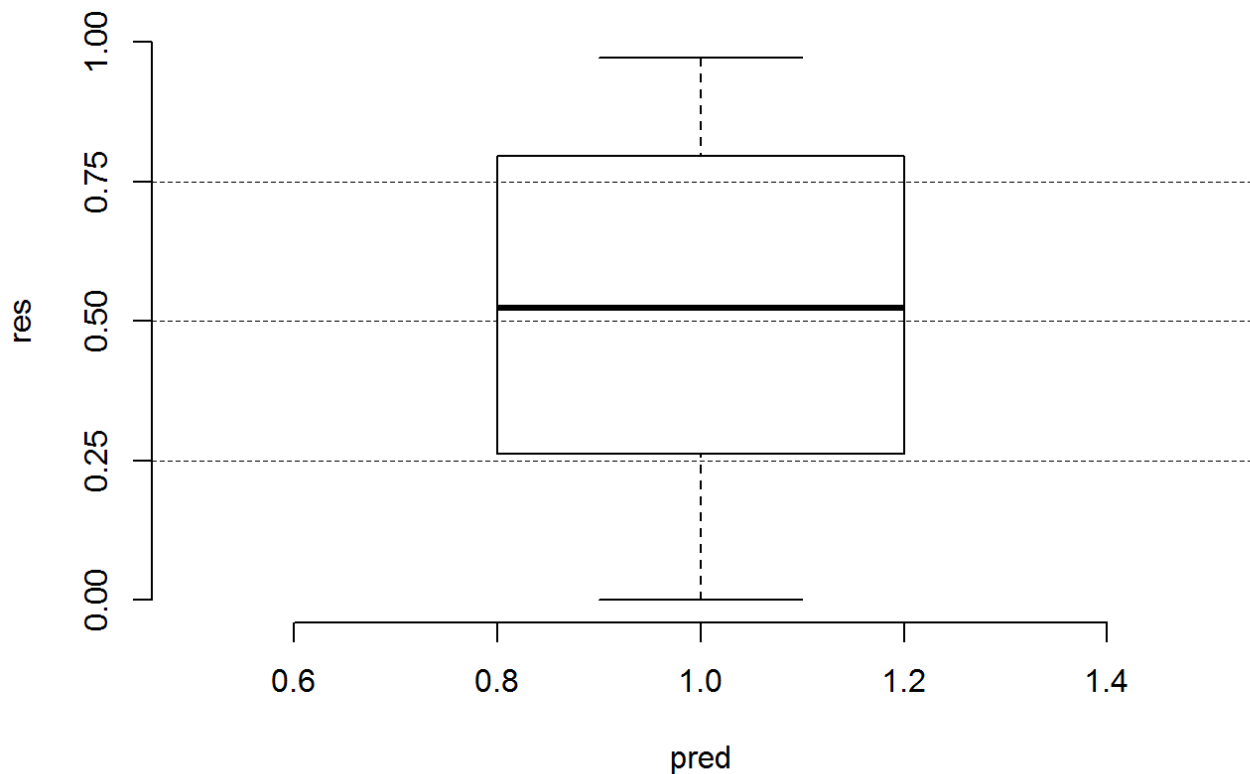
```
AICctab(F2d, F3d, F4d, F5d)
```

```
##      dAICc df
## F5d  0.0  4
## F4d  8.5  5
## F3d 90.2 17
## F2d 99.3 18
```

```
summary(F5d)
```

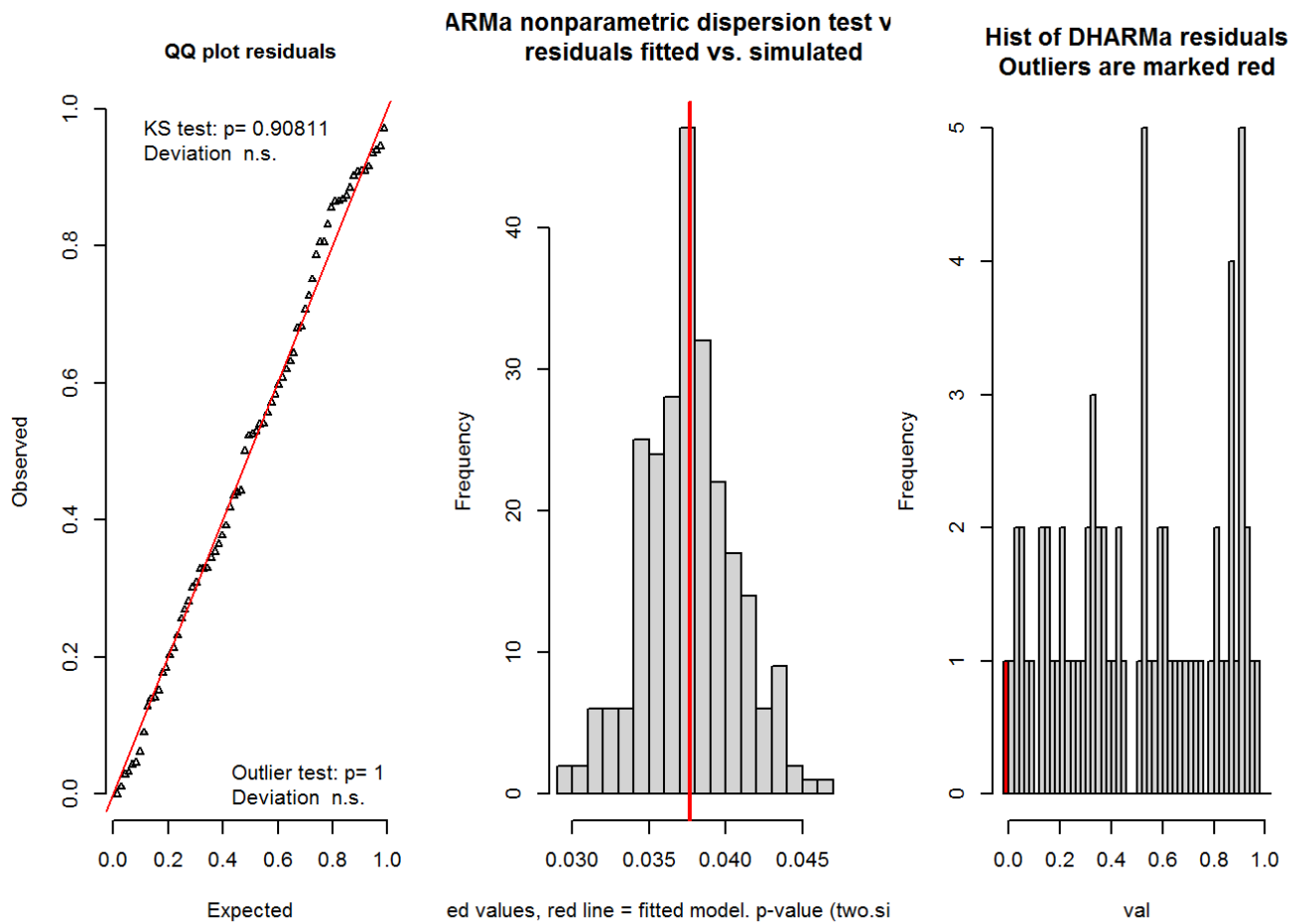
```
## Linear mixed model fit by REML ['lmerMod']
## Formula: pFemurrel ~ 1 + (1 | Site/Clutch)
##   Data: Labmet1
##
## REML criterion at convergence: -260.6
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -2.61489 -0.57530 -0.03813  0.80729  2.09801
##
## Random effects:
##   Groups       Name             Variance Std.Dev.
##   Clutch:Site (Intercept) 0.000e+00 0.00000
##   Site         (Intercept) 5.521e-05 0.00743
##   Residual                1.367e-03 0.03698
## Number of obs: 72, groups:  Clutch:Site, 20; Site, 6
##
## Fixed effects:
##              Estimate Std. Error t value
## (Intercept) 0.291863   0.005376   54.29
## convergence code: 0
## boundary (singular) fit: see ?isSingular
```

```
rr=simulateResiduals(F5d) ### simulate residuals
plotResiduals(rr) ### qqplot simulated vs observed
```



```
## NULL
```

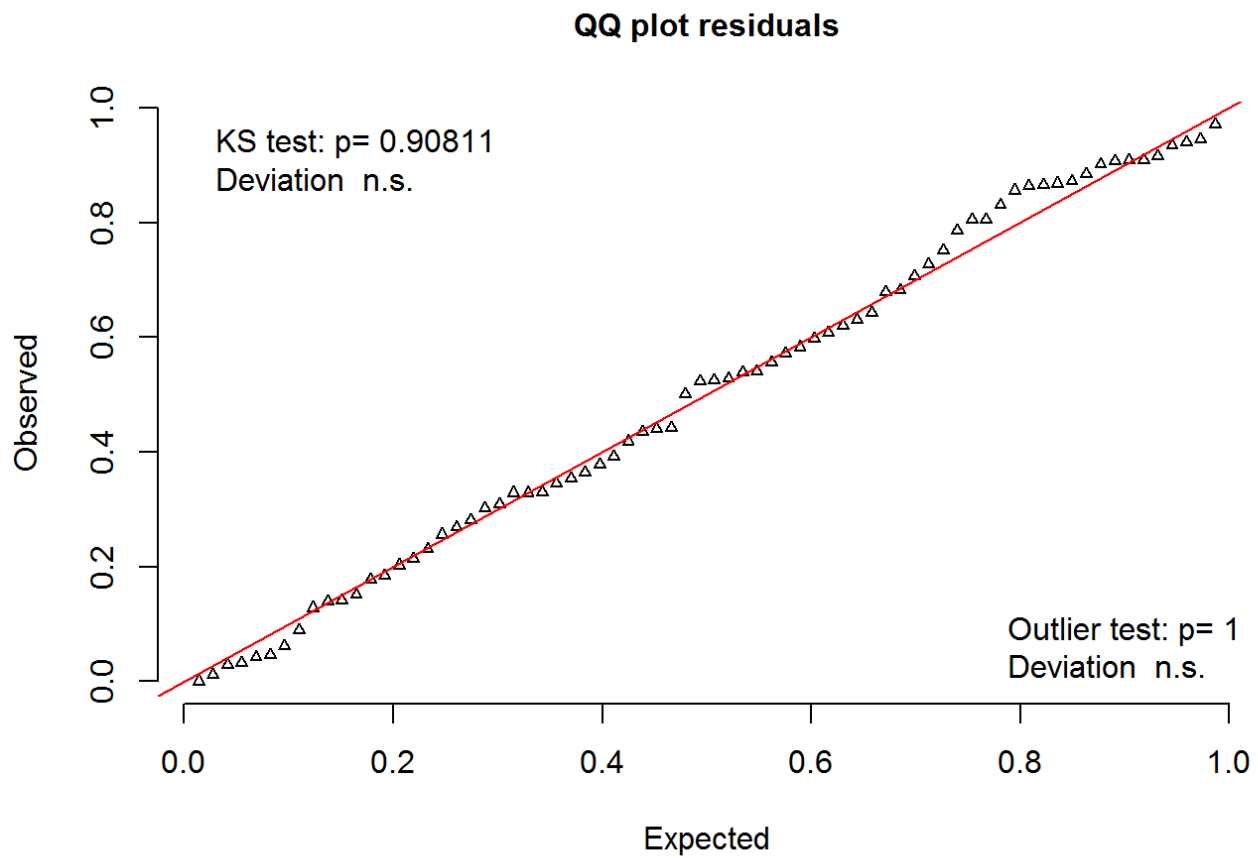
```
testResiduals(rr)###test qqplot
```



```
## $uniformity
##
## One-sample Kolmogorov-Smirnov test
##
## data: simulationOutput$scaledResiduals
## D = 0.064332, p-value = 0.9081
## alternative hypothesis: two-sided
##
##
## $dispersion
##
## DHARMA nonparametric dispersion test via sd of residuals fitted
## vs. simulated
##
## data: simulationOutput
## ratioObsSim = 1.0004, p-value = 0.976
## alternative hypothesis: two.sided
##
##
## $outliers
##
## DHARMA outlier test based on exact binomial test
##
## data: simulationOutput
## outLow = 1.0000000, outHigh = 0.0000000, nobs = 72.0000000, freqH0
## = 0.0039841, p-value = 1
## alternative hypothesis: two.sided
```

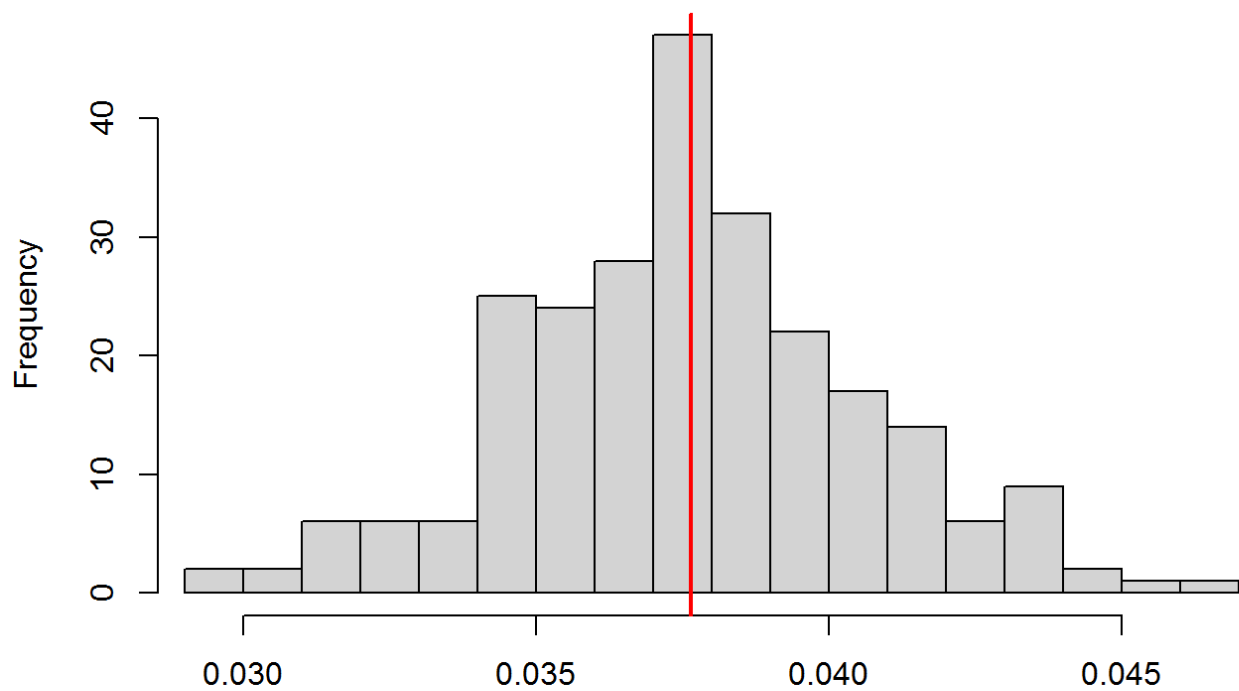
```
## $uniformity
##
## One-sample Kolmogorov-Smirnov test
##
## data: simulationOutput$scaledResiduals
## D = 0.064332, p-value = 0.9081
## alternative hypothesis: two-sided
##
##
## $dispersion
##
## DHARMA nonparametric dispersion test via sd of residuals fitted
## vs. simulated
##
## data: simulationOutput
## ratioObsSim = 1.0004, p-value = 0.976
## alternative hypothesis: two.sided
##
##
## $outliers
##
## DHARMA outlier test based on exact binomial test
##
## data: simulationOutput
## outLow = 1.0000000, outHigh = 0.0000000, nobs = 72.0000000, freqH0
## = 0.0039841, p-value = 1
## alternative hypothesis: two.sided
```

```
plotQQunif(rr, testUniformity = T) ### plot
```

```
testDispersion(rr,plot=T) ### test
```

DHARMA nonparametric dispersion test via sd of residuals fitted vs. simulated

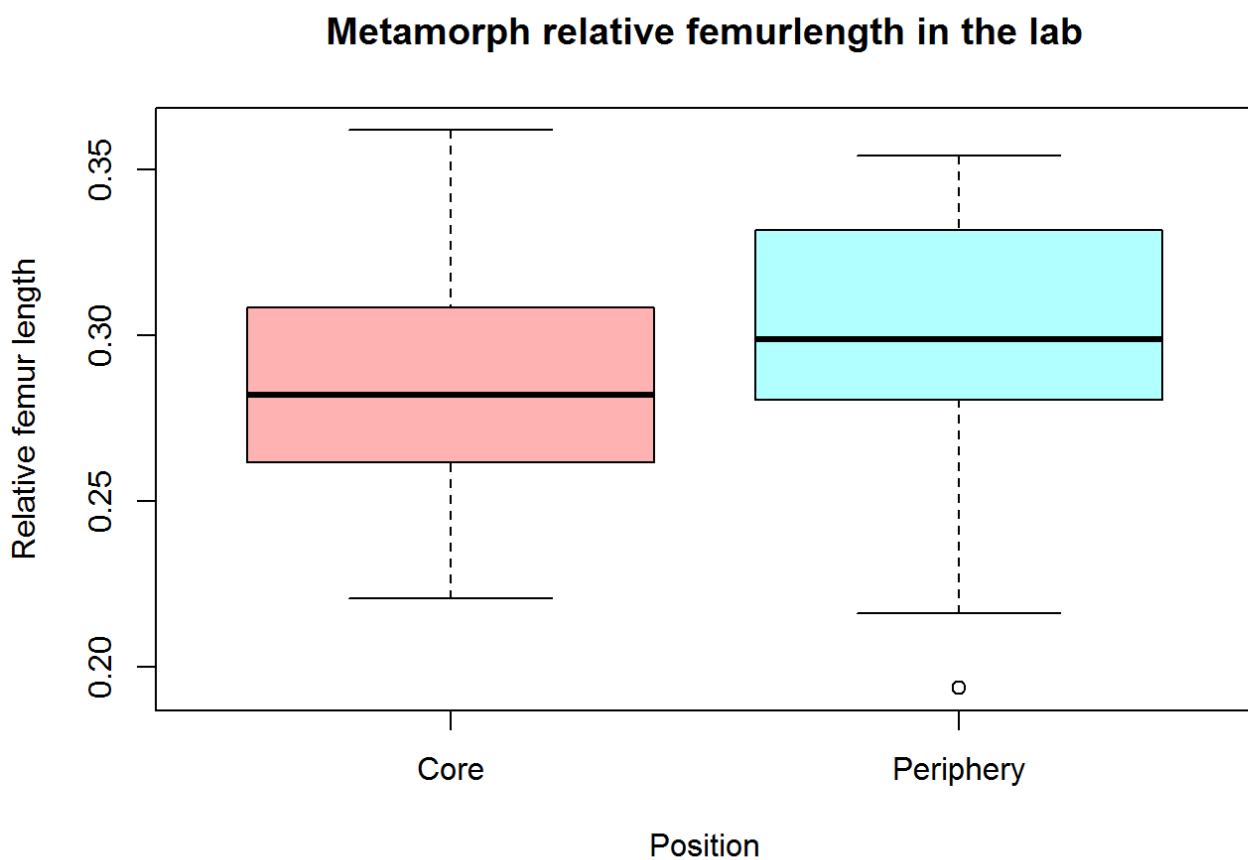


Simulated values, red line = fitted model. $p\text{-value (two.sided)} = 0.976$

```
##
## DHARMA nonparametric dispersion test via sd of residuals fitted
## vs. simulated
##
## data:  simulationOutput
## ratioObsSim = 1.0004, p-value = 0.976
## alternative hypothesis: two.sided
```

```
#Null model was selected
```

```
boxplot(Femurrelative~Position, xaxs=F,data=Labmet1, main= "Metamorph relative femu
rlength in the lab", xlab="Position", ylab="Relative femur length", col= rainbow(2,
alpha = 0.3), notch= F)
```



```
detach(Labmet)
```