HR_Va_h2_R_2022

Helen Payne

2024-06-18

```
#load pHRkages
library(lme4)
```

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

```
library(tidyverse)
```

```
## -- Attaching core tidyverse packages ------------------------ tidyverse 2.0.0 --
## v dplyr     1.1.4     v readr     2.1.5
## v forcats   1.0.0     v stringr   1.5.1
## v ggplot2   3.4.4     v tibble    3.2.1
## v lubridate 1.9.3     v tidyr     1.3.0
## v purrr     1.0.2
```

```
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x tidyr::expand() masks Matrix::expand()
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
## x tidyr::pack()   masks Matrix::pack()
## x tidyr::unpack() masks Matrix::unpack()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```

```
library(lmerTest)
```

```
##
## Attaching package: 'lmerTest'
##
## The following object is masked from 'package:lme4':
##
##     lmer
##
## The following object is masked from 'package:stats':
##
##     step
```

```
library(car)
```

```
## Loading required package: carData
##
## Attaching package: 'car'
##
## The following object is masked from 'package:dplyr':
##
##     recode
##
```

```
## The following object is masked from 'package:purrr':
##
##     some
```

##########**2022**##########

Read in the data:

```
HR_22_fit <- read_csv(here::here("data_sheets", "compiled_sheets", "HR_mastersheet_Fitness-mains_2022.cs
```

```
## Rows: 406 Columns: 57
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr   (5): Donor, Recipient, Gen, Replicated, Needed Area Redo
## dbl  (43): Year, Sequence, Cohort, Block, Transect, Plant_ID, F_plant_ID, fl...
## num   (1): F_multi
## lgl   (3): F_plant, Rep_FitP, any_FitP
## date  (5): Germ_Date, Sow_Date, Plant_Date, FFD, LFD
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```
HR_22_23_full <- read_csv(here::here("data_sheets", "compiled_sheets", "HR_22_23_full.csv"))
```

```
## Warning: One or more parsing issues, call `problems()` on your data frame for details,
## e.g.:
##   dat <- vroom(...)
##   problems(dat)
```

```
## Rows: 7684 Columns: 67
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr   (9): Gen, Donor, Recipient, sample_ID_SEG, sample_ID, SegPos, Block, R...
## dbl  (45): Year, Transect, Sequence, Plant_ID, days_sow2flower, days_plant2f...
## lgl   (4): F_plant, F_Num_03, Rep_FitP, any_FitP
## date  (9): Sow_Date, Plant_Date, FFD, LFD, F_Num_01, F_Num_02, photo_date, p...
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

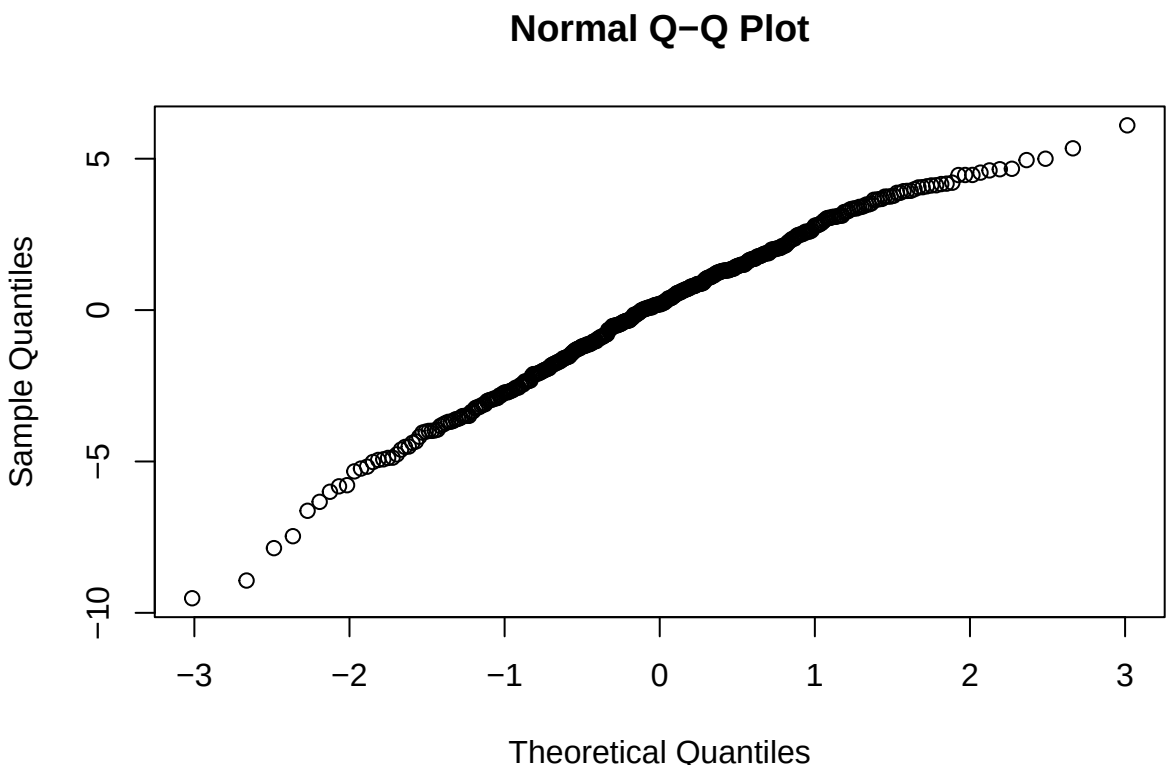
```
# Calculate the mean surv_to_flower for each group in AC_22_23_full and add it as a new column
HR_22_23_full <- HR_22_23_full %>%
  group_by(Year, Gen, Transect, Sequence, Donor, Recipient, SegPos) %>%
  mutate(prop_surv_to_flower = mean(surv_to_flower, na.rm = TRUE)) %>%
  ungroup()  # Remove grouping

HR_prop_sample <- HR_22_23_full %>%
  select(c(Year, Gen, Transect, Sequence, Donor, Recipient, SegPos, prop_surv_to_flower))  %>%
    distinct()

HR_22_fit <- HR_22_fit %>%
  left_join(HR_prop_sample %>% select(Year, Gen, Transect, Sequence, Donor, Recipient, SegPos, prop_surv
            by = c("Year", "Gen", "Transect", "Sequence", "Donor", "Recipient"))

# Create the new column
HR_22_fit <- HR_22_fit %>%
  mutate(est_fitness = prop_surv_to_flower * est_fecundity)
```

```r
#log scale traits that are highly skewed
HR_22_fit$skel_dryweight_mg_SEG <- log(HR_22_fit$skel_dryweight_mg_SEG)
HR_22_fit$est_fecundity <- sqrt(HR_22_fit$est_fecundity + 1)
HR_22_fit$SLA_SEG <- log(HR_22_fit$SLA_SEG)
HR_22_fit$est_fitness <- sqrt(HR_22_fit$est_fitness)

#mean center the traits of interest

traits <- c("corolla_diam_mm_SEG", "skel_dryweight_mg_SEG", "fl_duration", "est_fecundity", "msm_all",

# Mean center eHRh trait
for (trait in traits) {
  trait_mean <- mean(HR_22_fit[[trait]], na.rm = TRUE)
  HR_22_fit[[paste0(trait, "_centered")]] <- HR_22_fit[[trait]] - trait_mean
}
```

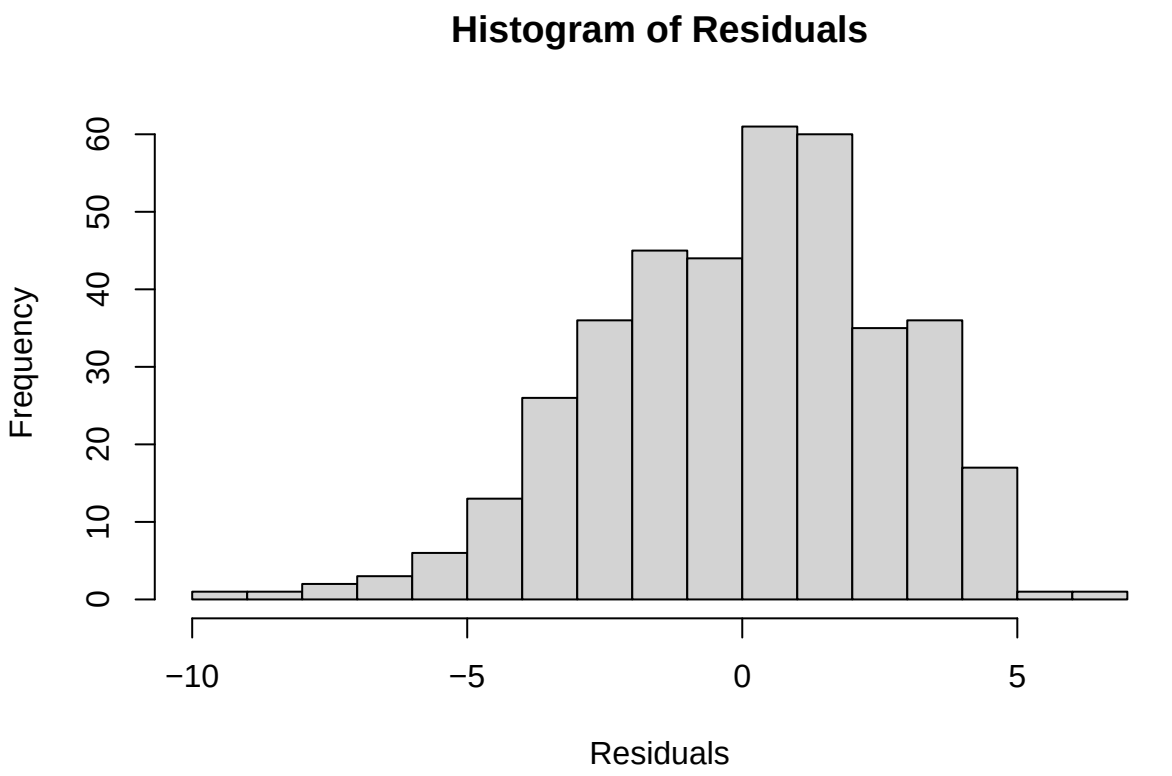
```r
# Create the mixed model for corolla area

#corolla_model <- lmer(corolla_diam_mm_SEG_centered ~ (1 | Recipient) + (1 | Donor), data = HR_22_fit)

corolla_model <- lmer(corolla_diam_mm_SEG_centered ~ (1 | Donor), data = HR_22_fit)

rand(corolla_model)#including Donor and Transect as random effects significantly improves the models fi
```

```
## ANOVA-like table for random-effects: Single term deletions
##
## Model:
## corolla_diam_mm_SEG_centered ~ (1 | Donor)
##             npar  logLik    AIC   LRT Df Pr(>Chisq)
## <none>         3 -949.59 1905.2
## (1 | Donor)    2 -952.50 1909.0 5.827  1    0.01578 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
# ExtrHRt residuals from the model
residuals <- resid(corolla_model)

# Q-Q plot for normality
qqnorm(residuals) #looks good
```




```
# Histogram for normality
hist(residuals, breaks = 20, main = "Histogram of Residuals", xlab = "Residuals") #normal
```




```
#use Model 2 (Donor)
```

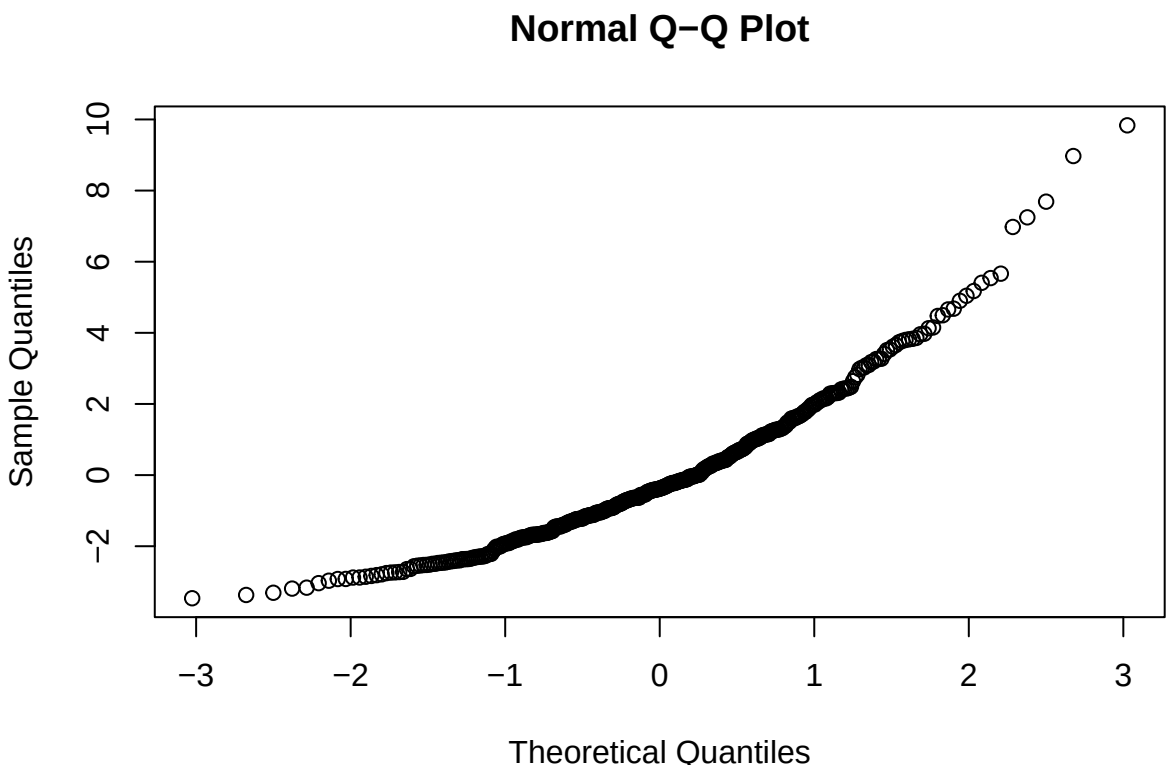
```r
# Create the mixed model for skeleton weight, with skeleton weight log transformed
#skel_model <- lmer((skel_dryweight_mg_SEG_centered) ~ (1 | Recipient) + (1 | Donor), data = HR_22_fit)

#skel_model <- lmer((skel_dryweight_mg_SEG_centered) ~ (1 | Donor), data = HR_22_fit) #singular
```

```r
# Create the mixed model for flowering duration
#fl_duration_model <- lmer(fl_duration_centered ~ (1 | Recipient) + (1 | Donor), data = HR_22_fit) #Sin

# Create the mixed model for flowering duration
#fl_duration_model <- lmer(fl_duration_centered ~ (1 | Donor), data = HR_22_fit) #Singular
```

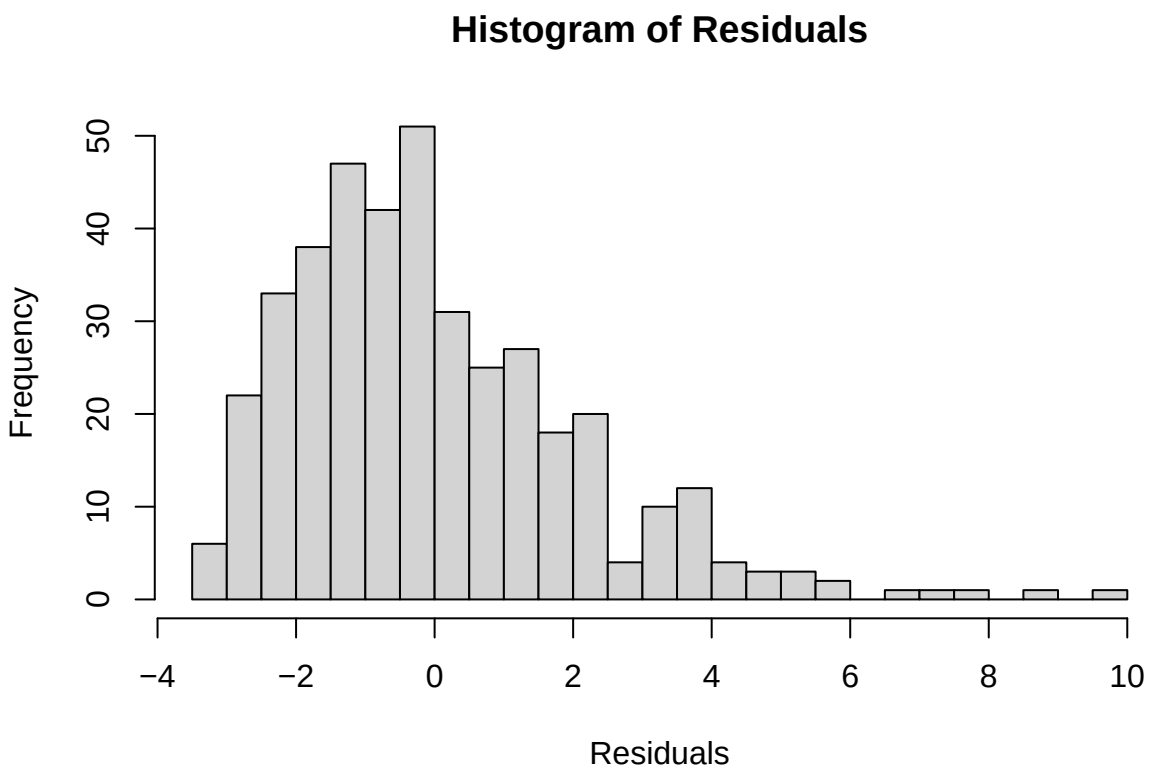
```r
# Create the mixed model for estimated fecundity, sqrt transforming estimated fecundity
est_fecundity_model <- lmer((est_fecundity_centered) ~ (1|Transect) + (1 | Recipient) + (1 | Donor), da

rand(est_fecundity_model)
```

```
## ANOVA-like table for random-effects: Single term deletions
##
## Model:
## (est_fecundity_centered) ~ (1 | Transect) + (1 | Recipient) + (1 | Donor)
##                  npar  logLik    AIC     LRT Df Pr(>Chisq)
## <none>              5 -899.92 1809.8
## (1 | Transect)      4 -905.03 1818.1 10.2260  1   0.001385 **
## (1 | Recipient)     4 -900.34 1808.7  0.8386  1   0.359810
## (1 | Donor)         4 -899.92 1807.8  0.0063  1   0.936887
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
# ExtrHRt residuals from the model
residuals <- resid(est_fecundity_model)

# Q-Q plot for normality
qqnorm(residuals) #good enough
```

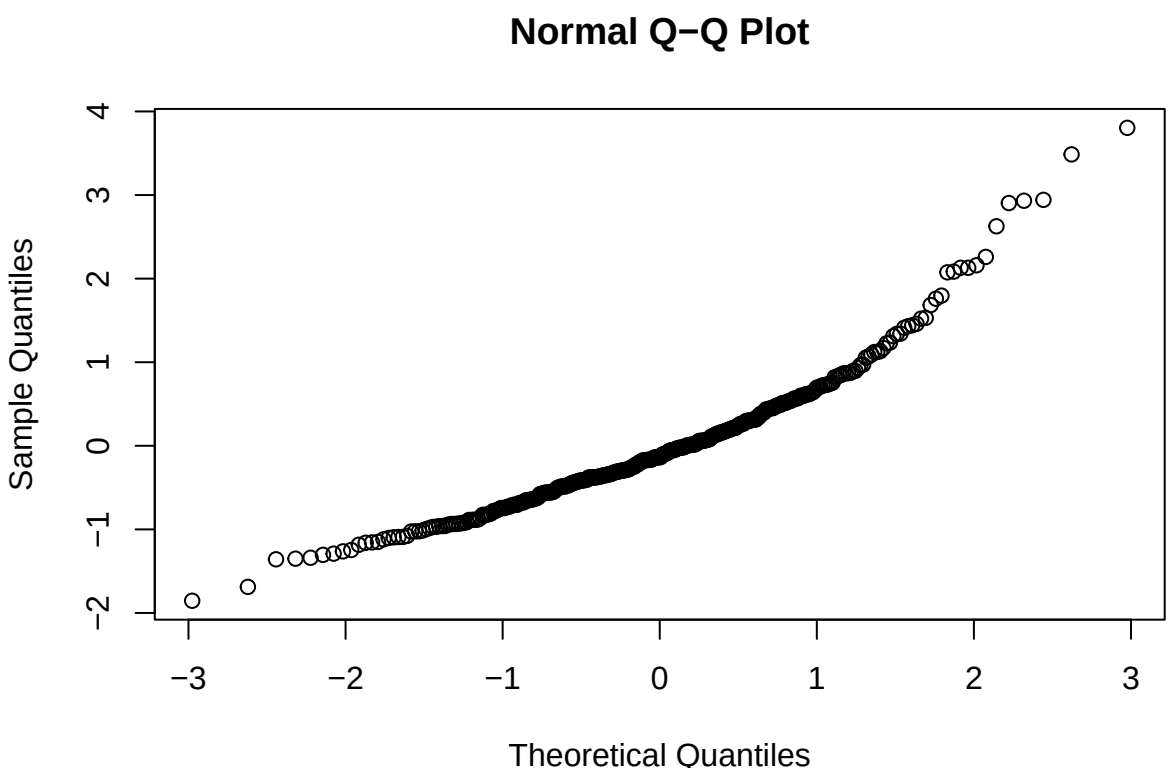



```r
# Histogram for normality
hist(residuals, breaks = 20, main = "Histogram of Residuals", xlab = "Residuals") #normal-ish
```

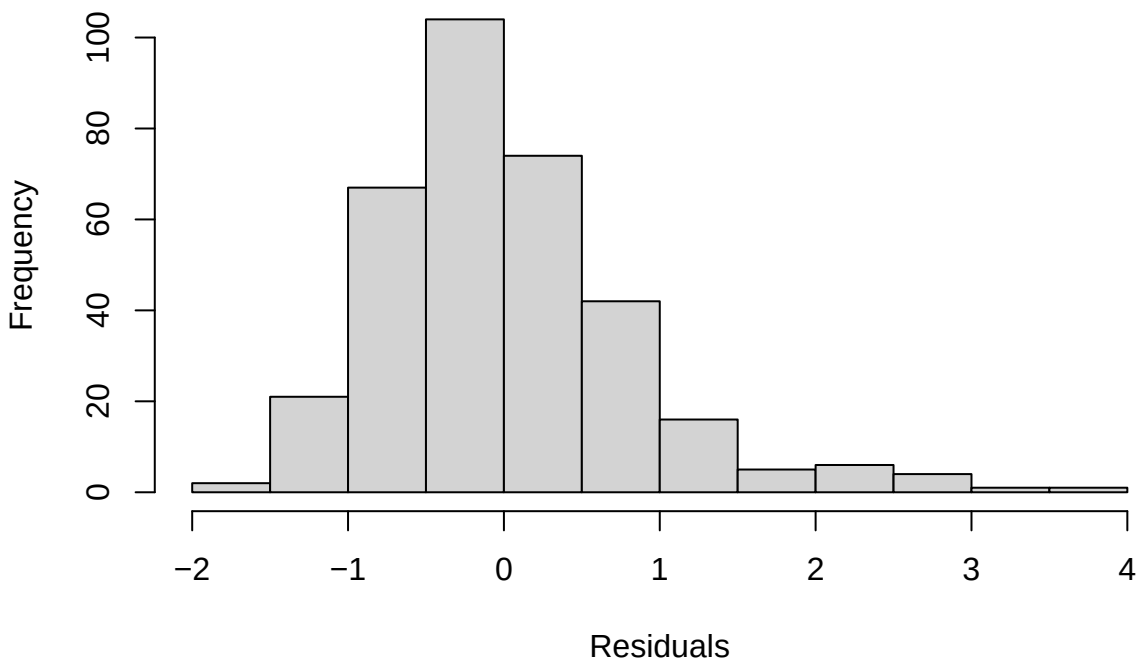



```r
# Create the mixed model for mean seed mass, log transformed mean seed mass
msm_model <- lmer((msm_all_centered) ~ (1| Transect) + (1 | Recipient) + (1 | Donor), data = HR_22_fit)
```

```r
# ExtrHRt residuals from the model
residuals <- resid(msm_model)

# Q-Q plot for normality
qqnorm(residuals) #good enough
```




```r
# Histogram for normality
hist(residuals, breaks = 20, main = "Histogram of Residuals", xlab = "Residuals") #normal-ish
```




```
# Test the significance of the random effects
rand(msm_model) #none of these random effects significantly improve the model's fit
```

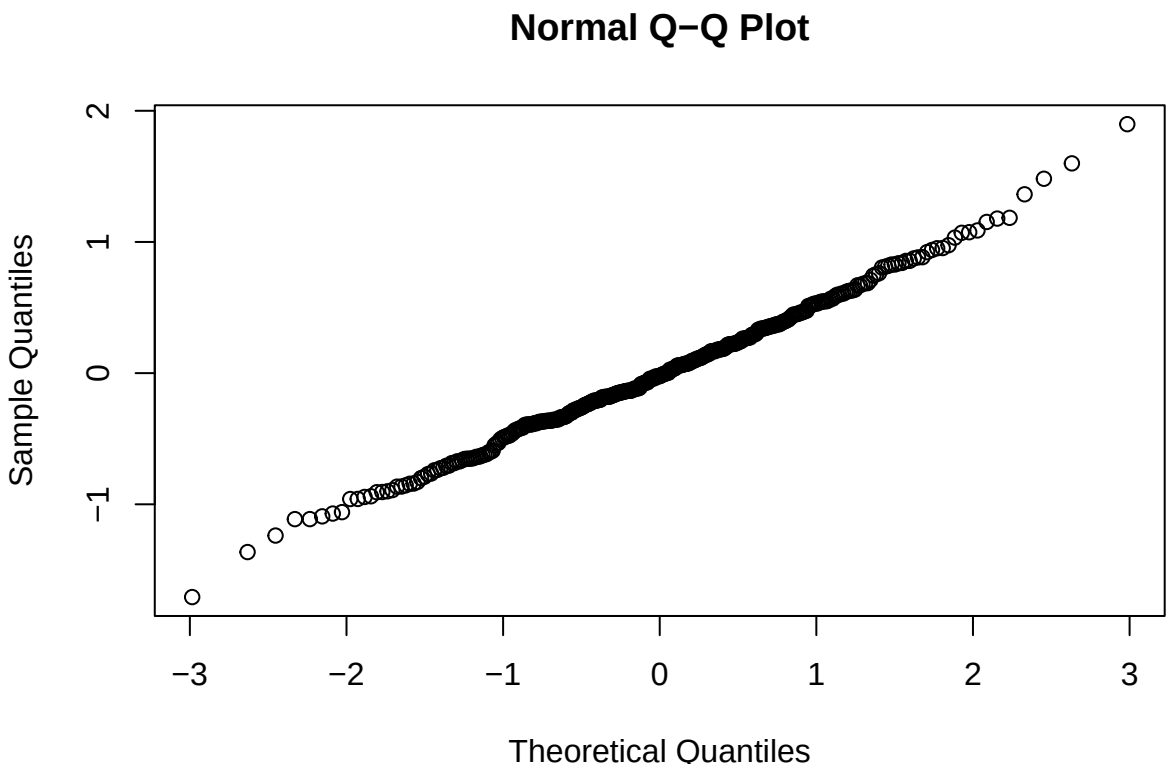
```
## ANOVA-like table for random-effects: Single term deletions
##
## Model:
## (msm_all_centered) ~ (1 | Transect) + (1 | Recipient) + (1 | Donor)
##                  npar  logLik    AIC     LRT Df Pr(>Chisq)
## <none>              5 -477.78 965.55
## (1 | Transect)      4 -483.77 975.53 11.9810  1  0.0005375 ***
## (1 | Recipient)     4 -478.07 964.13  0.5823  1  0.4454056
## (1 | Donor)         4 -481.21 970.42  6.8727  1  0.0087524 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# Create the mixed model for SLA
#SLA_model <- lmer((SLA_SEG_centered) ~ (1|Transect) + (1 | Recipient) + (1 | Donor), data = HR_22_fit)

#SLA_model <- lmer((SLA_SEG_centered) ~ (1 | Donor), data = HR_22_fit) #Singular
```

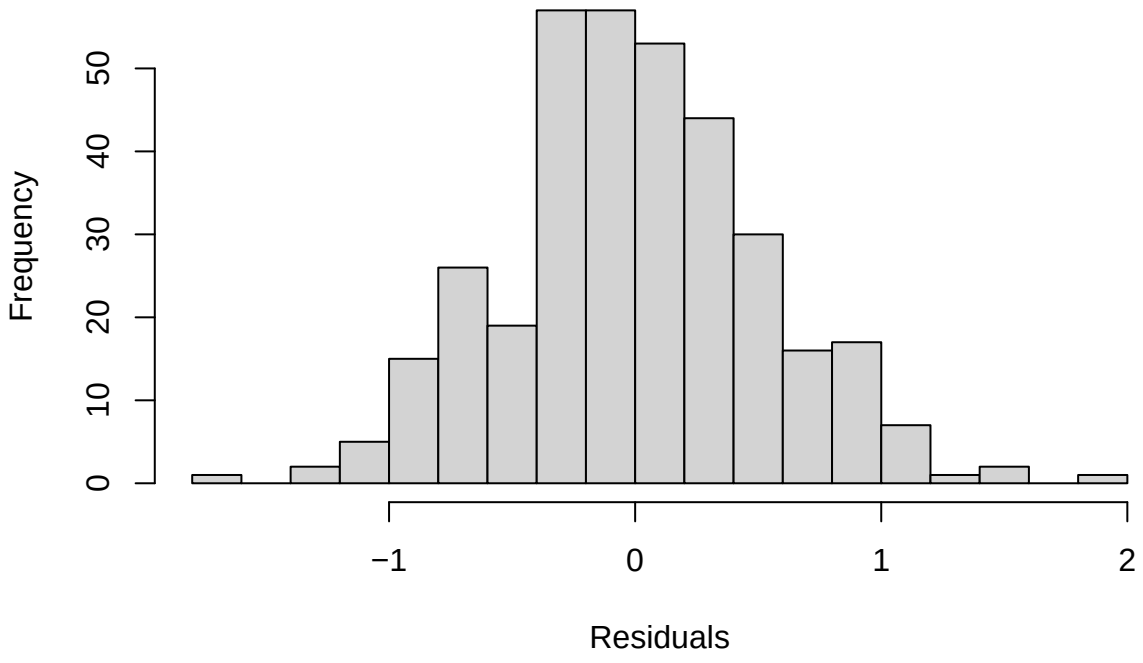
```
# Create the mixed model for mean seed mass, log transformed mean seed mass
#LMA_model <- lmer(LMA_SEG_centered ~ (1|Transect) + (1 | Recipient) + (1 | Donor), data = HR_22_fit) #

# Create the mixed model for mean seed mass, log transformed mean seed mass
#LMA_model <- lmer(LMA_SEG_centered ~ (1 | Donor), data = HR_22_fit) #Singular
```

```
# Create the mixed model for d13C, log transformed mean seed mass
d13C_model <- lmer(d13C_SEG_centered ~ (1|Transect) + (1 | Recipient) + (1 | Donor), data = HR_22_fit)

# ExtrHRt residuals from the model
residuals <- resid(d13C_model)
```

```r
# Q-Q plot for normality
qqnorm(residuals) #good enough
```



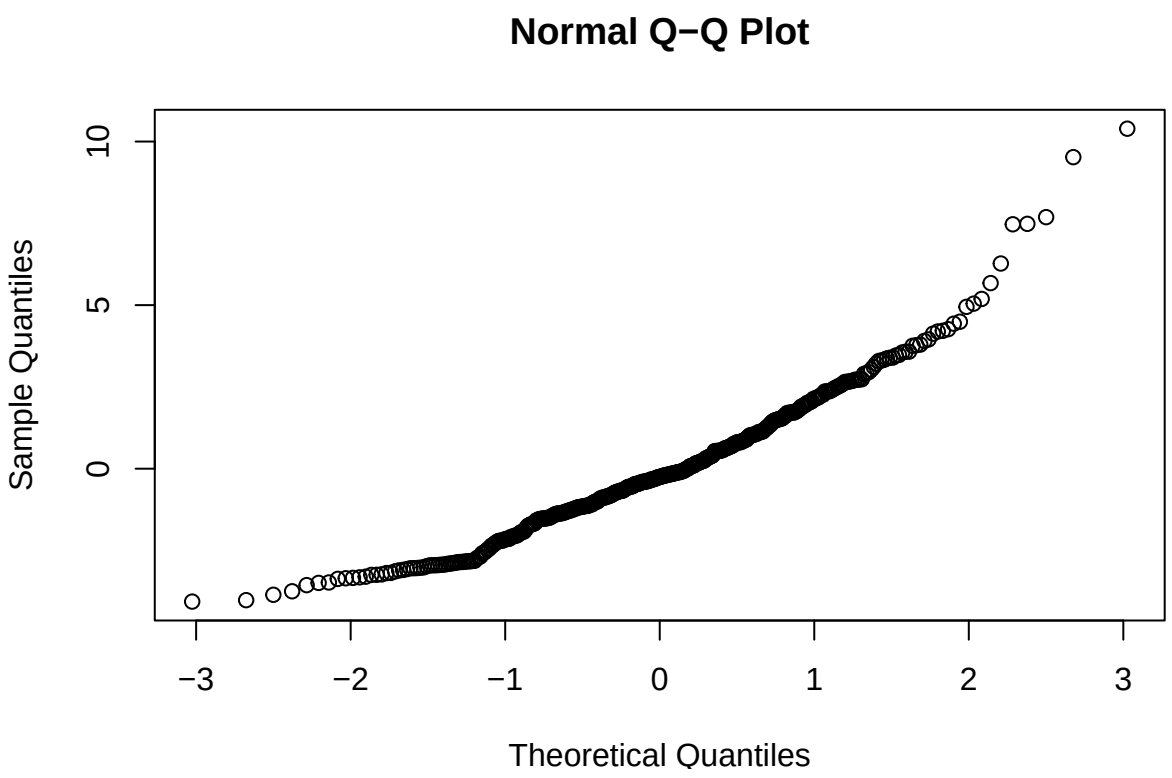

```r
# Histogram for normality
hist(residuals, breaks = 20, main = "Histogram of Residuals", xlab = "Residuals") #normal-ish
```



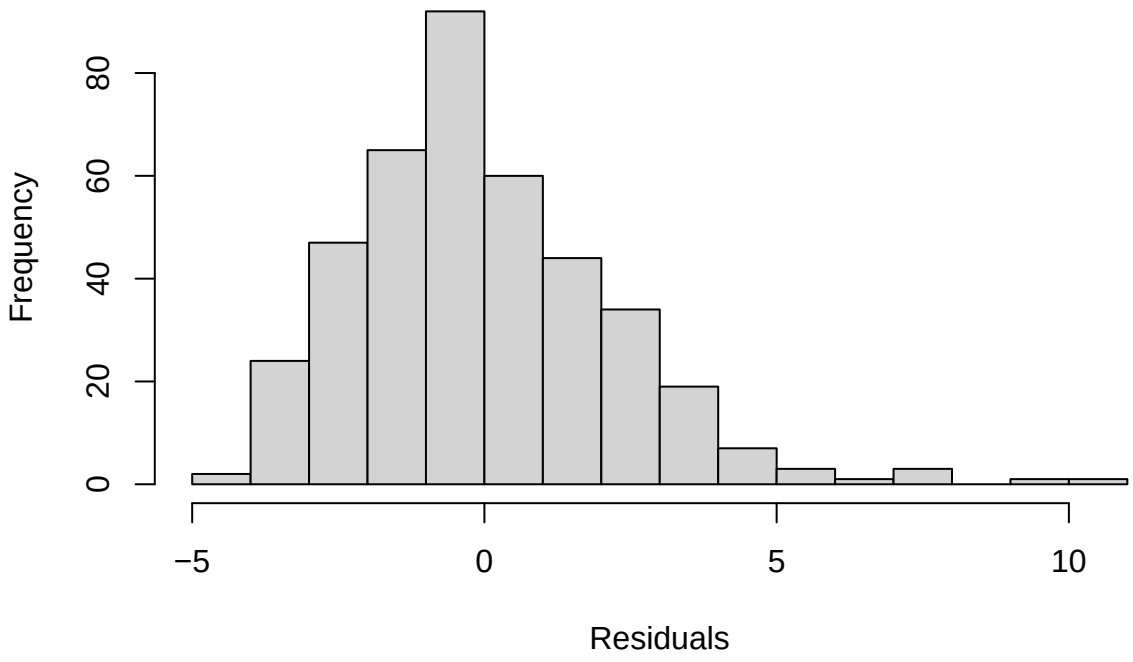
```r
# Test the significance of the random effects
rand(d13C_model) #all Random effects significantly improve the models fit!
```

```
## ANOVA-like table for random-effects: Single term deletions
##
## Model:
## d13C_SEG_centered ~ (1 | Transect) + (1 | Recipient) + (1 | Donor)
##                  npar  logLik    AIC     LRT Df Pr(>Chisq)
## <none>              5 -344.48 698.96
## (1 | Transect)      4 -355.89 719.79 22.8310  1  1.769e-06 ***
## (1 | Recipient)     4 -344.70 697.39  0.4387  1     0.5077
## (1 | Donor)         4 -354.53 717.05 20.0954  1  7.367e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
# Create the mixed model for d13C, log transformed mean seed mass
est_fitness_model <- lmer(est_fitness_centered ~ (1|Transect) + (1 | Recipient) + (1 | Donor), data = H

# Extract residuals from the model
residuals <- resid(est_fitness_model)

# Q-Q plot for normality
qqnorm(residuals) #good enough
```




```r
# Histogram for normality
hist(residuals, breaks = 20, main = "Histogram of Residuals", xlab = "Residuals") #normal-ish
```




```
# Test the significance of the random effects
rand(est_fitness_model) #all Random effects significantly improve the models fit!
```

```
## ANOVA-like table for random-effects: Single term deletions
##
## Model:
## est_fitness_centered ~ (1 | Transect) + (1 | Recipient) + (1 | Donor)
##                  npar  logLik    AIC     LRT Df Pr(>Chisq)
## <none>              5 -915.98 1842.0
## (1 | Transect)      4 -921.58 1851.2 11.1972  1  0.0008192 ***
## (1 | Recipient)     4 -916.20 1840.4  0.4243  1  0.5147879
## (1 | Donor)         4 -916.02 1840.0  0.0687  1  0.7932644
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# Function to round values to a specified number of significant digits
round_df <- function(df, digits) {
  df[] <- lapply(df, function(x) if(is.numeric(x)) signif(x, digits) else x)
  return(df)
}

# Function to extract variance components and calculate required values
calculate_variances <- function(model, trait_name) {
  var_components <- as.data.frame(VarCorr(model))

  # Initialize variables
  V_mat <- NA
  V_sd_mat <- NA
  Va_mat <- NA
  V_pat <- NA
```

```
  V_sd_pat <- NA
  Va_pat <- NA
  res_var <- NA
  Vp <- NA
  Vp_sd <- NA
  h2 <- NA

  # Check if "Recipient" is included as a random effect
  if ("Recipient" %in% var_components$grp) {
    V_mat <- var_components$vcov[var_components$grp == "Recipient"]
    V_sd_mat <- sqrt(V_mat)
    Va_mat <- V_mat * 4
  }

  # Calculate other variance components
  if ("Donor" %in% var_components$grp) {
    V_pat <- var_components$vcov[var_components$grp == "Donor"]
    V_sd_pat <- sqrt(V_pat)
    Va_pat <- V_pat * 4
  }

  if ("Residual" %in% var_components$grp) {
    res_var <- var_components$vcov[var_components$grp == "Residual"]
  }

  # Calculate total phenotypic variance and narrow-sense heritability if components are available
  if (!is.na(Va_mat) & !is.na(Va_pat) & !is.na(res_var)) {
    Vp <- Va_mat + Va_pat + res_var
    Vp_sd <- sqrt(Vp)
    h2 <- Va_pat / Vp  # assumed calculation
  }

  # Extract the number of observations
  n_obs <- nobs(model)

  # Create the dataframe and add the traits column
  df <- data.frame(traits = trait_name, V_mat, V_sd_mat, Va_mat, V_pat, V_sd_pat, Va_pat, Vp, Vp_sd, h2

  # Round the dataframe values to four significant digits
  df <- round_df(df, 4)

  return(df)
}

# Calculate variances for eHRh model and add trait names
corolla_variances <- calculate_variances(corolla_model, 'corolla_diameter')
#skel_variances <- calculate_variances(skel_model, "skel_biomass_mg")
est_fecundity_variances <- calculate_variances(est_fecundity_model, "estimated_fecundity")
msm_variances <- calculate_variances(msm_model, "mean_seed_mass")
#SLA_variances <- calculate_variances(SLA_model, "SLA")
#LMA_variances <- calculate_variances(LMA_model, "LMA")
d13C_variances <- calculate_variances(d13C_model, "delta_C_13")
est_fitness_variances <- calculate_variances(est_fitness_model, "est_fitness")
```

```r
# Combine the results into a single dataframe
variance_HR_2022 <- rbind(
  corolla_variances,
  d13C_variances,
  est_fecundity_variances,
  msm_variances,
  est_fitness_variances
)

# Print the dataframe
print(variance_HR_2022)
```

```
##                traits   V_mat V_sd_mat  Va_mat   V_pat V_sd_pat Va_pat     Vp
## 1    corolla_diameter      NA       NA      NA 0.57140   0.7559 2.2860     NA
## 2          delta_C_13 0.01542   0.1242 0.06168 0.09727   0.3119 0.3891 0.7761
## 3 estimated_fecundity 0.26520   0.5150 1.06100 0.01285   0.1134 0.0514 5.7980
## 4      mean_seed_mass 0.04756   0.2181 0.19020 0.12170   0.3489 0.4868 1.4640
## 5         est_fitness 0.20300   0.4505 0.81200 0.04540   0.2131 0.1816 6.1130
##   Vp_sd       h2 n_obs
## 1    NA       NA   388
## 2 0.881 0.501300   353
## 3 2.408 0.008865   403
## 4 1.210 0.332600   343
## 5 2.472 0.029710   403
```

```r
#Save the csv file if you want
write_csv(x = variance_HR_2022, here::here("data_sheets", "compiled_sheets", "HR_Va_h2_R_2022.csv"))
```