Sleep Bout Analysis

CT Berezin

3/29/2022

```r
bouts <- read.csv("../data/WT_McKO_combined_bouts.csv",
                    fileEncoding = 'UTF-8-BOM')
bouts <- bouts %>% mutate(Time = mdy_hms(Time),
                          Sleep.Stage = as.factor(Sleep.Stage),
                          Genotype = as.factor(Genotype),
                          Transmitter = as.factor(Transmitter))

bouts <- bouts %>% mutate(Hour = hour(Time))
bouts <- bouts %>% mutate(ZT = Time + hours(18))
bouts <- bouts %>% mutate(ZT = hour(ZT))

bouts <- bouts %>% dplyr::mutate(phase = as.factor(ifelse(ZT>11, "dark", "light")))

head(bouts)
```

```
##   Bout Epoch                Time Sleep.Stage Count Duration Genotype
## 1    1  2216 2019-02-10 18:00:00           S    36      360     McKO
## 2    2  2252 2019-02-10 18:06:00           W     2       20     McKO
## 3    3  2254 2019-02-10 18:06:20           S     6       60     McKO
## 4    4  2260 2019-02-10 18:07:20           W     4       40     McKO
## 5    5  2264 2019-02-10 18:08:00           S     4       40     McKO
## 6    6  2268 2019-02-10 18:08:40           W    17      170     McKO
##   Transmitter Day Hour ZT phase
## 1      988823   1   18 12  dark
## 2      988823   1   18 12  dark
## 3      988823   1   18 12  dark
## 4      988823   1   18 12  dark
## 5      988823   1   18 12  dark
## 6      988823   1   18 12  dark
```

```r
SumStats_daily <- bouts %>% group_by(Genotype, Transmitter, Sleep.Stage, Day) %>%
  summarise(n_bouts = n(),
            total_duration = sum(Duration),
            min_bout_length = min(Duration),
            mean_bout_length = mean(Duration),
            max_bout_length = max(Duration),
            sd_bout_length = sd(Duration))
```

```
## `summarise()` has grouped output by 'Genotype', 'Transmitter', 'Sleep.Stage'. You can override using
```

```
head(SumStats_daily)
```

```
## # A tibble: 6 x 10
## # Groups:    Genotype, Transmitter, Sleep.Stage [2]
##   Genotype Transmitter Sleep.Stage   Day n_bouts total_duration min_bout_length
##   <fct>    <fct>       <fct>       <int>   <int>          <int>           <int>
## 1 McKO     1050240     P               1     138           4760              10
## 2 McKO     1050240     P               2     183           5340              10
## 3 McKO     1050240     P               3     229           4910              10
## 4 McKO     1050240     S               1     237          45220              10
## 5 McKO     1050240     S               2     257          45180              10
## 6 McKO     1050240     S               3     284          46290              10
## # ... with 3 more variables: mean_bout_length <dbl>, max_bout_length <int>,
## #   sd_bout_length <dbl>
```

```
SumStats_daily_n <- SumStats_daily %>% group_by(Genotype, Transmitter, Sleep.Stage) %>%
  summarise(mean_n_bouts_day = mean(n_bouts))
```

```
## 'summarise()' has grouped output by 'Genotype', 'Transmitter'. You can override using the '.groups' a
```

```
head(SumStats_daily_n)
```

```
## # A tibble: 6 x 4
## # Groups:    Genotype, Transmitter [2]
##   Genotype Transmitter Sleep.Stage mean_n_bouts_day
##   <fct>    <fct>       <fct>                  <dbl>
## 1 McKO     1050240     P                       183.
## 2 McKO     1050240     S                       259.
## 3 McKO     1050240     W                       246.
## 4 McKO     1050240     X                         6
## 5 McKO     1050241     P                       283.
## 6 McKO     1050241     S                       234
```

```
SumStats_avg <- bouts %>% group_by(Genotype, Transmitter, Sleep.Stage) %>%
  summarise(min_bout_length = min(Duration),
            mean_bout_length = mean(Duration),
            max_bout_length = max(Duration),
            sd_bout_length = sd(Duration))
```

```
## 'summarise()' has grouped output by 'Genotype', 'Transmitter'. You can override using the '.groups' a
```

```
head(SumStats_avg)
```

```
## # A tibble: 6 x 7
## # Groups:    Genotype, Transmitter [2]
##   Genotype Transmitter Sleep.Stage min_bout_length mean_bout_length
##   <fct>    <fct>       <fct>                 <int>            <dbl>
## 1 McKO     1050240     P                        10             27.3
## 2 McKO     1050240     S                        10            176.
## 3 McKO     1050240     W                        10            145.
```

```
## 4 McKO      1050240     X                           10          10
## 5 McKO      1050241     P                           10          28.6
## 6 McKO      1050241     S                           10         171.
## # ... with 2 more variables: max_bout_length <int>, sd_bout_length <dbl>
```

```
SumStats_phase <- bouts %>% group_by(Genotype, Transmitter, Sleep.Stage, phase) %>%
  summarise(n_bouts = n(),
            total_duration = sum(Duration),
            min_bout_length = min(Duration),
            mean_bout_length = mean(Duration),
            max_bout_length = max(Duration),
            sd_bout_length = sd(Duration))
```

```
## 'summarise()' has grouped output by 'Genotype', 'Transmitter', 'Sleep.Stage'. You can override using
```

```
head(SumStats_phase)
```

```
## # A tibble: 6 x 10
## # Groups:   Genotype, Transmitter, Sleep.Stage [3]
##   Genotype Transmitter Sleep.Stage phase n_bouts total_duration min_bout_length
##   <fct>    <fct>       <fct>       <fct>   <int>          <int>           <int>
## 1 McKO     1050240     P           dark      200           4750              10
## 2 McKO     1050240     P           light     350          10260              10
## 3 McKO     1050240     S           dark      271          47400              10
## 4 McKO     1050240     S           light     507          89290              10
## 5 McKO     1050240     W           dark      325          77640              10
## 6 McKO     1050240     W           light     414          29720              10
## # ... with 3 more variables: mean_bout_length <dbl>, max_bout_length <int>,
## #   sd_bout_length <dbl>
```

```
SumStats_phase_n <- SumStats_phase %>% group_by(Genotype, Transmitter, Sleep.Stage, phase) %>%
  summarise(mean_n_bouts_day = mean(n_bouts))
```

```
## 'summarise()' has grouped output by 'Genotype', 'Transmitter', 'Sleep.Stage'. You can override using
```

```
head(SumStats_phase_n)
```

```
## # A tibble: 6 x 5
## # Groups:   Genotype, Transmitter, Sleep.Stage [3]
##   Genotype Transmitter Sleep.Stage phase mean_n_bouts_day
##   <fct>    <fct>       <fct>       <fct>            <dbl>
## 1 McKO     1050240     P           dark               200
## 2 McKO     1050240     P           light              350
## 3 McKO     1050240     S           dark               271
## 4 McKO     1050240     S           light              507
## 5 McKO     1050240     W           dark               325
## 6 McKO     1050240     W           light              414
```

```
SumStats_avg %>% group_by(Genotype, Sleep.Stage) %>%
  summarise(n = n(),
            mean_bout_length = mean(mean_bout_length),
            min_bout_length = min(min_bout_length),
            max_bout_length = max(max_bout_length))
```

```
## `summarise()` has grouped output by 'Genotype'. You can override using the `.groups` argument.
```

```
## # A tibble: 8 x 6
## # Groups:   Genotype [2]
##   Genotype Sleep.Stage     n mean_bout_length min_bout_length max_bout_length
##   <fct>    <fct>       <int>            <dbl>           <int>           <int>
## 1 McKO     P               9             26.8              10             250
## 2 McKO     S               9            142.               10            3920
## 3 McKO     W               9             95.2              10            6310
## 4 McKO     X               9             10                10              10
## 5 WT       P              10             30.9              10             350
## 6 WT       S              10            127.               10            4390
## 7 WT       W              10            100.               10            5230
## 8 WT       X              10             10                10              10
```

```r
SumStats_phase_grps <- SumStats_phase %>% group_by(Genotype, Sleep.Stage, phase) %>%
  summarise(n = n(),
            avg_bout_length = mean(mean_bout_length),
            min_bout_length = min(min_bout_length),
            max_bout_length = max(max_bout_length),
            sd_bout_length = sd(mean_bout_length))
```

```
## `summarise()` has grouped output by 'Genotype', 'Sleep.Stage'. You can override using the `.groups` a
```

```r
SumStats_phase_grps
```

```
## # A tibble: 16 x 8
## # Groups:   Genotype, Sleep.Stage [8]
##    Genotype Sleep.Stage phase     n avg_bout_length min_bout_length
##    <fct>    <fct>       <fct> <int>            <dbl>           <int>
##  1 McKO     P           dark      9            20.5               10
##  2 McKO     P           light     9            31.0               10
##  3 McKO     S           dark      9           107.                10
##  4 McKO     S           light     9           190.                10
##  5 McKO     W           dark      9           120.                10
##  6 McKO     W           light     9            67.9               10
##  7 McKO     X           dark      9            10                 10
##  8 McKO     X           light     9            10                 10
##  9 WT       P           dark     10            25.8               10
## 10 WT       P           light    10            36.7               10
## 11 WT       S           dark     10            96.2               10
## 12 WT       S           light    10           164.                10
## 13 WT       W           dark     10           118.                10
## 14 WT       W           light    10            78.5               10
## 15 WT       X           dark     10            10                 10
## 16 WT       X           light    10            10                 10
## # ... with 2 more variables: max_bout_length <int>, sd_bout_length <dbl>
```

```r
SumStats_phase_n_grps <- SumStats_phase_n %>% group_by(Genotype, Sleep.Stage, phase) %>%
  summarise(n = n(),
            mean_n_bouts_day = mean(mean_n_bouts_day))
```

```
## 'summarise()' has grouped output by 'Genotype', 'Sleep.Stage'. You can override using the '.groups' a

head(SumStats_phase_n_grps)
```

```
## # A tibble: 6 x 5
## # Groups:   Genotype, Sleep.Stage [3]
##   Genotype Sleep.Stage phase     n mean_n_bouts_day
##   <fct>    <fct>       <fct> <int>            <dbl>
## 1 McKO     P           dark      9             378.
## 2 McKO     P           light     9             418.
## 3 McKO     S           dark      9             881.
## 4 McKO     S           light     9             779.
## 5 McKO     W           dark      9            1005.
## 6 McKO     W           light     9             673.
```

```
lm_n_bouts <- lmer(mean_n_bouts_day ~ Sleep.Stage * phase * Genotype + (1 | Transmitter), data=SumStats_
#plot(lm_n_bouts, type=c("p","smooth"), col.line=1)
#lattice::qqmath(lm_n_bouts)
anova(lm_n_bouts)
```
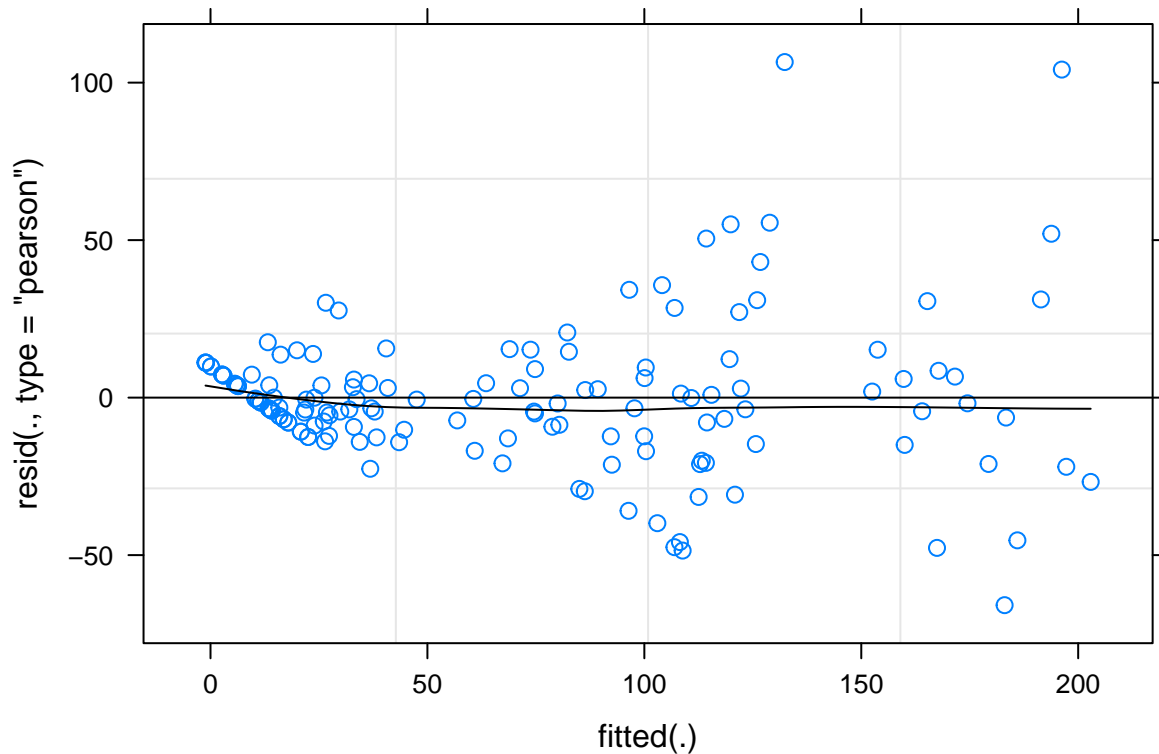
```
## Analysis of Variance Table
##                         npar   Sum Sq  Mean Sq   F value
## Sleep.Stage                3 16786640  5595547  109.8216
## phase                      1   453224   453224    8.8952
## Genotype                   1      566      566    0.0111
## Sleep.Stage:phase          3   716742   238914    4.6891
## Sleep.Stage:Genotype       3     3057     1019    0.0200
## phase:Genotype             1     1379     1379    0.0271
## Sleep.Stage:phase:Genotype 3     5461     1820    0.0357
```

```
emmeans(lm_n_bouts, pairwise ~ Genotype | Sleep.Stage * phase)$contrasts
```
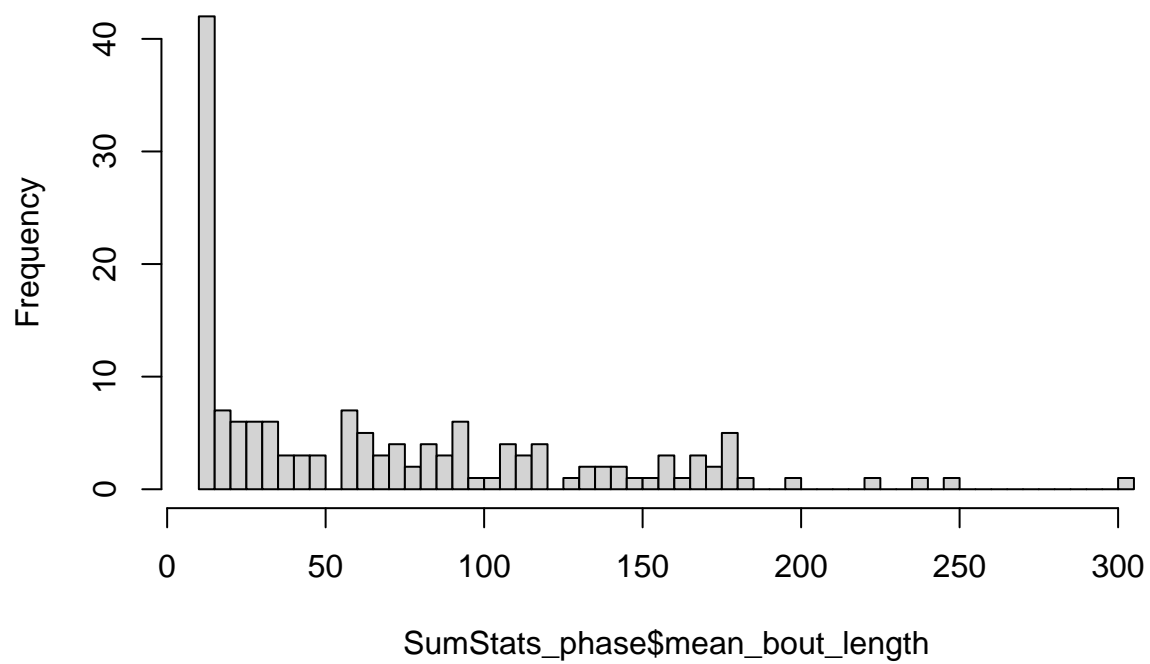
```
## Sleep.Stage = P, phase = dark:
##  contrast  estimate  SE   df t.ratio p.value
##  McKO - WT   -23.84 141 54.8  -0.169  0.8665
##
## Sleep.Stage = S, phase = dark:
##  contrast  estimate  SE   df t.ratio p.value
##  McKO - WT     8.97 141 54.8   0.064  0.9496
##
## Sleep.Stage = W, phase = dark:
##  contrast  estimate  SE   df t.ratio p.value
##  McKO - WT   -28.18 141 54.8  -0.200  0.8425
##
## Sleep.Stage = X, phase = dark:
##  contrast  estimate  SE   df t.ratio p.value
##  McKO - WT   -24.29 141 54.8  -0.172  0.8640
##
## Sleep.Stage = P, phase = light:
##  contrast  estimate  SE   df t.ratio p.value
##  McKO - WT    20.09 141 54.8   0.142  0.8873
##
```

```
## Sleep.Stage = S, phase = light:
##  contrast  estimate  SE   df t.ratio p.value
##  McKO - WT   -14.33 141 54.8  -0.102  0.9195
##
## Sleep.Stage = W, phase = light:
##  contrast  estimate  SE   df t.ratio p.value
##  McKO - WT   -18.44 141 54.8  -0.131  0.8965
##
## Sleep.Stage = X, phase = light:
##  contrast  estimate  SE   df t.ratio p.value
##  McKO - WT    -6.40 141 54.8  -0.045  0.9640
##
## Degrees-of-freedom method: kenward-roger
```

```r
lm_mean_length <- lmer(mean_bout_length ~ Sleep.Stage * phase * Genotype + (1 | Transmitter), data=SumS
plot(lm_mean_length, type=c("p","smooth"), col.line=1)
```



```r
lattice::qqmath(lm_mean_length)
```
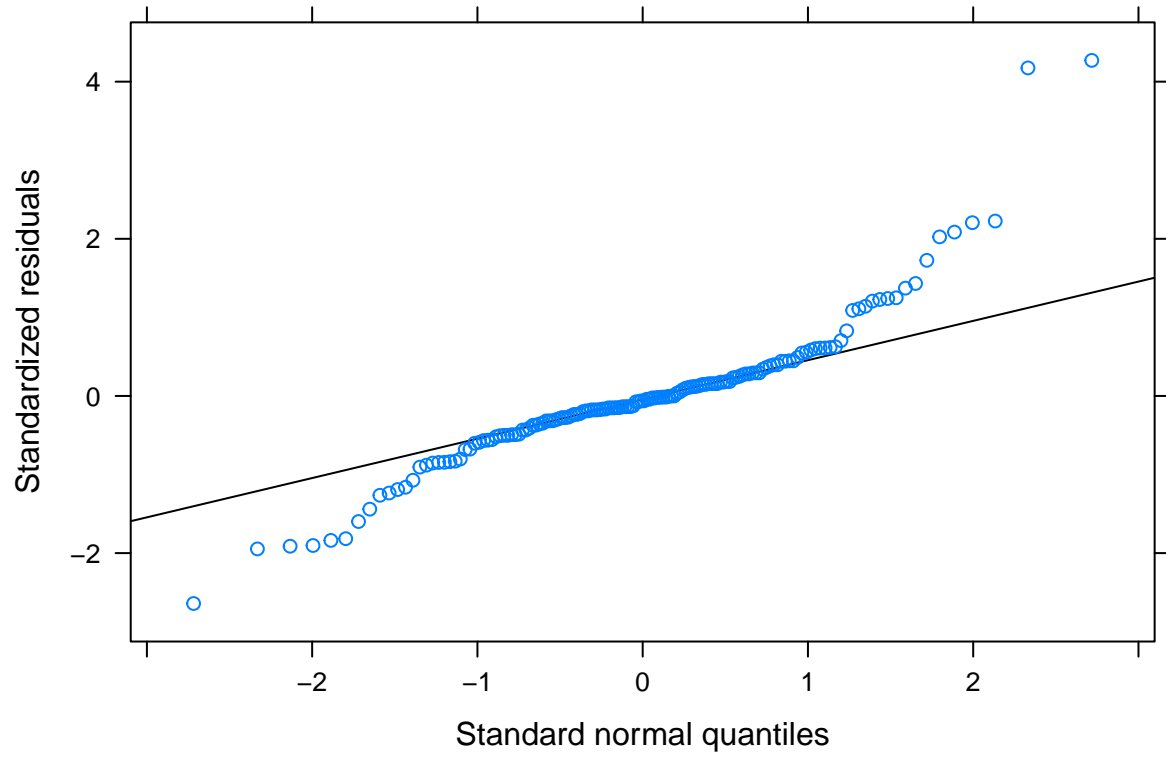
```
hist(SumStats_phase$mean_bout_length, breaks=50)
```
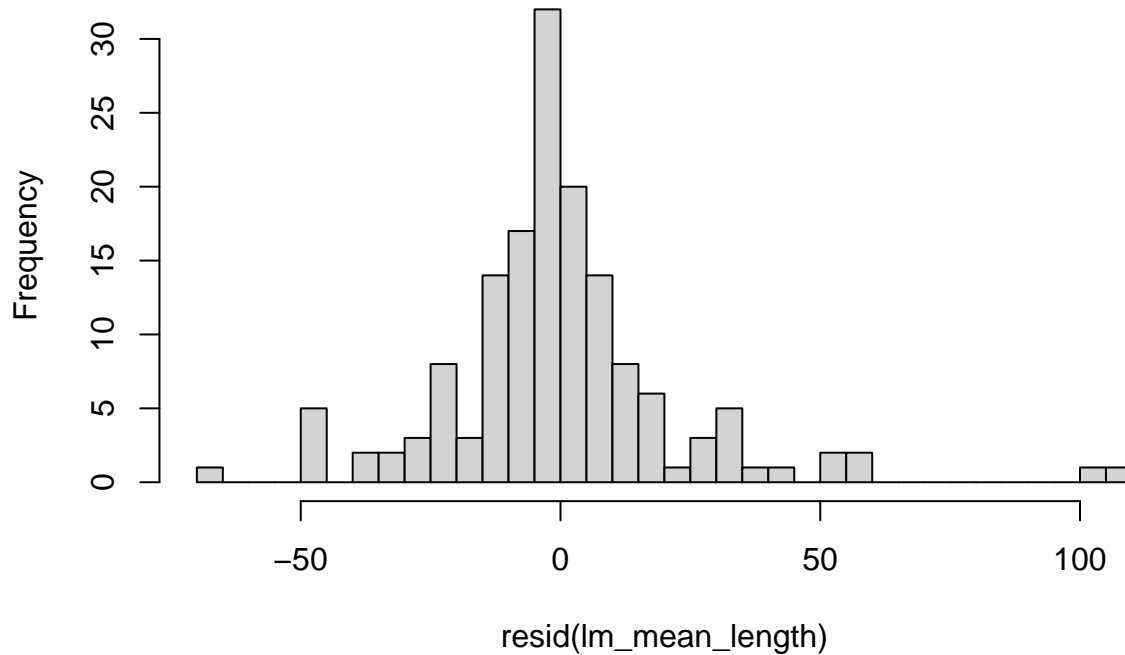
Histogram of SumStats_phase$mean_bout_length


```
hist(resid(lm_mean_length), breaks=50)
```

Histogram of resid(lm_mean_length)


```
anova(lm_mean_length)
```
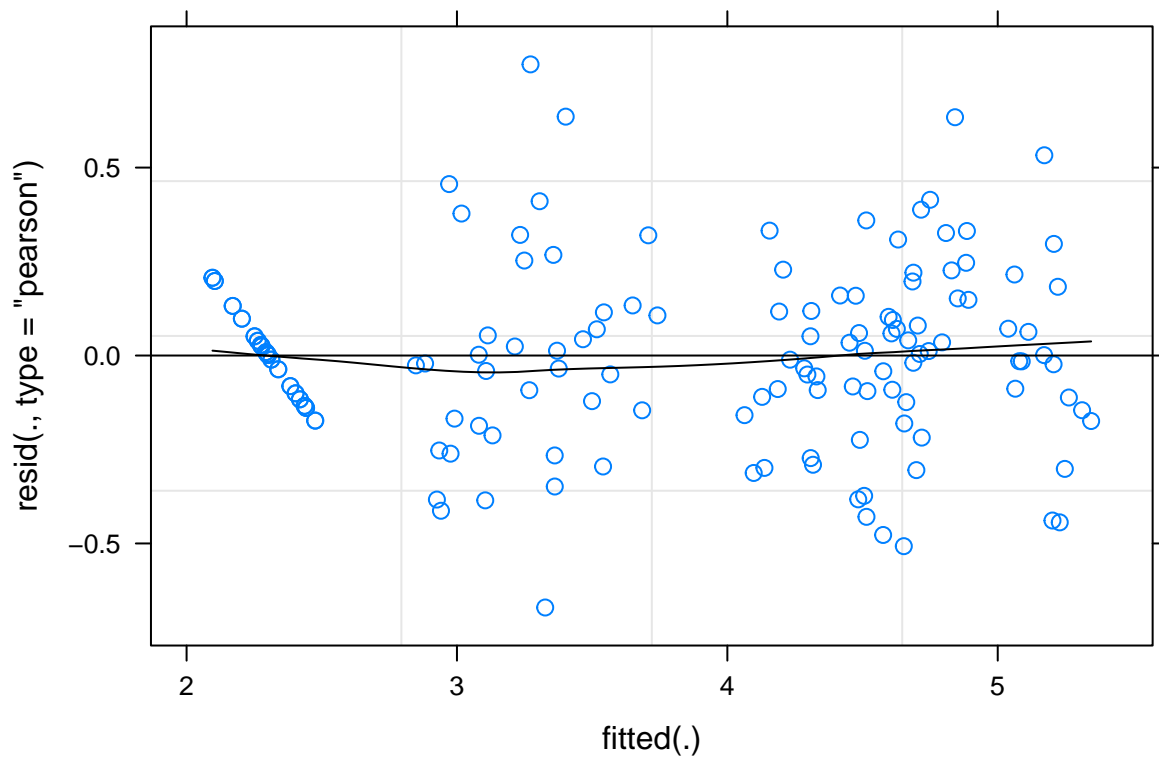
```
## Analysis of Variance Table
##                           npar Sum Sq Mean Sq  F value
## Sleep.Stage                  3 407882  135961 218.4922
## phase                        1   3817    3817   6.1332
## Genotype                     1     85      85   0.1372
## Sleep.Stage:phase            3  70113   23371  37.5577
## Sleep.Stage:Genotype         3   3678    1226   1.9700
## phase:Genotype               1      4       4   0.0065
## Sleep.Stage:phase:Genotype   3    923     308   0.4945
```

```
emmeans(lm_mean_length, pairwise ~ Genotype | Sleep.Stage * phase)$contrasts
```
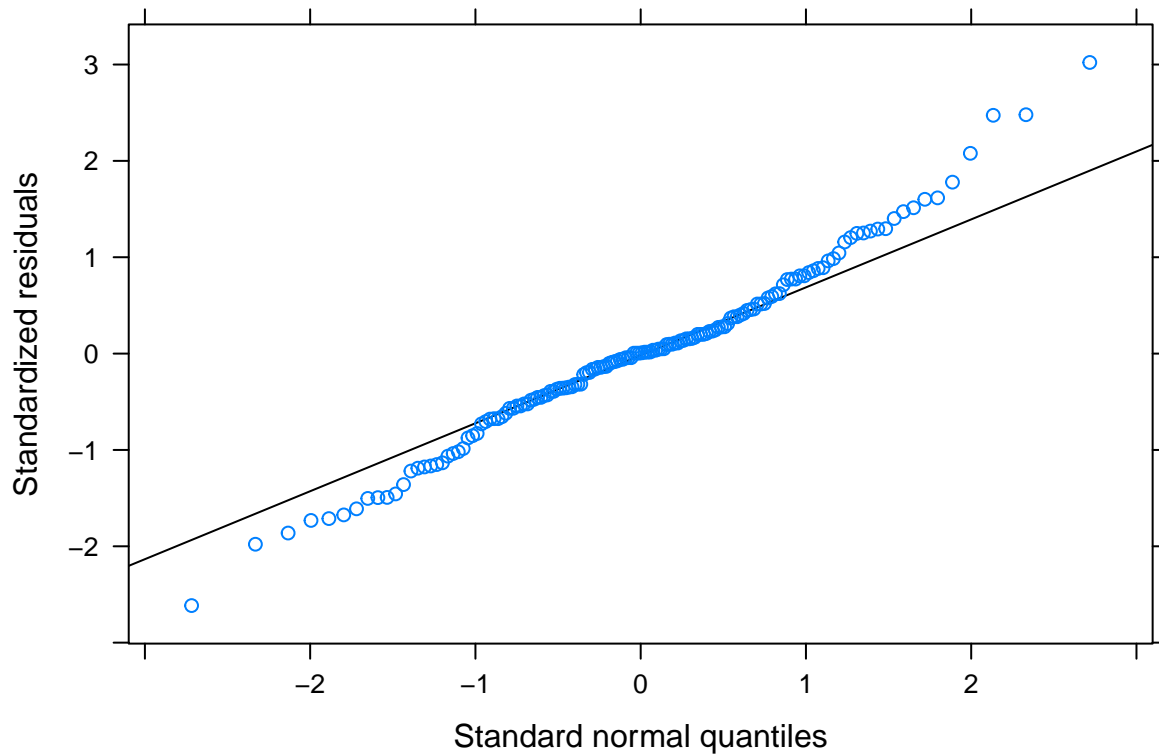
```
## Sleep.Stage = P, phase = dark:
##  contrast  estimate   SE  df t.ratio p.value
##  McKO - WT    -5.21 12.3 120  -0.422  0.6739
##
## Sleep.Stage = S, phase = dark:
##  contrast  estimate   SE  df t.ratio p.value
##  McKO - WT    11.22 12.3 120   0.908  0.3655
##
## Sleep.Stage = W, phase = dark:
##  contrast  estimate   SE  df t.ratio p.value
```

```
##  McKO - WT      1.76 12.3 120    0.143  0.8869
##
## Sleep.Stage = X, phase = dark:
##  contrast  estimate   SE  df t.ratio p.value
##  McKO - WT      0.00 12.3 120    0.000  1.0000
##
## Sleep.Stage = P, phase = light:
##  contrast  estimate   SE  df t.ratio p.value
##  McKO - WT     -5.76 12.3 120   -0.466  0.6420
##
## Sleep.Stage = S, phase = light:
##  contrast  estimate   SE  df t.ratio p.value
##  McKO - WT     26.70 12.3 120    2.162  0.0326
##
## Sleep.Stage = W, phase = light:
##  contrast  estimate   SE  df t.ratio p.value
##  McKO - WT    -10.55 12.3 120   -0.854  0.3947
##
## Sleep.Stage = X, phase = light:
##  contrast  estimate   SE  df t.ratio p.value
##  McKO - WT      0.00 12.3 120    0.000  1.0000
##
## Degrees-of-freedom method: kenward-roger
```

```r
lm_log_mean_length <- lmer(log(mean_bout_length) ~ Sleep.Stage * phase * Genotype + (1 | Transmitter),
plot(lm_log_mean_length, type=c("p","smooth"), col.line=1)
```

```
lattice::qqmath(lm_log_mean_length)
```



```
anova(lm_mean_length)
```

```
## Analysis of Variance Table
##                         npar Sum Sq Mean Sq  F value
## Sleep.Stage                3 407882  135961 218.4922
## phase                      1   3817    3817   6.1332
## Genotype                   1     85      85   0.1372
## Sleep.Stage:phase          3  70113   23371  37.5577
## Sleep.Stage:Genotype       3   3678    1226   1.9700
## phase:Genotype             1      4       4   0.0065
## Sleep.Stage:phase:Genotype 3    923     308   0.4945
```

```
emmeans(lm_mean_length, pairwise ~ Genotype | Sleep.Stage * phase)$contrasts
```

```
## Sleep.Stage = P, phase = dark:
##  contrast  estimate   SE  df t.ratio p.value
##  McKO - WT    -5.21 12.3 120  -0.422  0.6739
##
## Sleep.Stage = S, phase = dark:
##  contrast  estimate   SE  df t.ratio p.value
##  McKO - WT    11.22 12.3 120   0.908  0.3655
##
```

```
## Sleep.Stage = W, phase = dark:
##  contrast  estimate   SE  df t.ratio p.value
##  McKO - WT      1.76 12.3 120   0.143  0.8869
##
## Sleep.Stage = X, phase = dark:
##  contrast  estimate   SE  df t.ratio p.value
##  McKO - WT      0.00 12.3 120   0.000  1.0000
##
## Sleep.Stage = P, phase = light:
##  contrast  estimate   SE  df t.ratio p.value
##  McKO - WT     -5.76 12.3 120  -0.466  0.6420
##
## Sleep.Stage = S, phase = light:
##  contrast  estimate   SE  df t.ratio p.value
##  McKO - WT     26.70 12.3 120   2.162  0.0326
##
## Sleep.Stage = W, phase = light:
##  contrast  estimate   SE  df t.ratio p.value
##  McKO - WT    -10.55 12.3 120  -0.854  0.3947
##
## Sleep.Stage = X, phase = light:
##  contrast  estimate   SE  df t.ratio p.value
##  McKO - WT      0.00 12.3 120   0.000  1.0000
##
## Degrees-of-freedom method: kenward-roger
```

```r
lm_max_length <- lmer(max_bout_length ~ Sleep.Stage * phase * Genotype + (1 | Transmitter), data=SumSta
#plot(lm_max_length, type=c("p","smooth"), col.line=1)
#lattice::qqmath(lm_max_length)
anova(lm_max_length)
```

```
## Analysis of Variance Table
##                          npar   Sum Sq  Mean Sq F value
## Sleep.Stage                 3 88814079 29604693 63.6687
## phase                       1   843042   843042  1.8131
## Genotype                    1   170326   170326  0.3663
## Sleep.Stage:phase           3  7416195  2472065  5.3165
## Sleep.Stage:Genotype        3   244168    81389  0.1750
## phase:Genotype              1   192675   192675  0.4144
## Sleep.Stage:phase:Genotype  3   313821   104607  0.2250
```

```r
emmeans(lm_max_length, pairwise ~ Genotype | Sleep.Stage * phase)$contrasts
```

```
## Sleep.Stage = P, phase = dark:
##  contrast  estimate  SE  df t.ratio p.value
##  McKO - WT   -18.78 322 133  -0.058  0.9536
##
## Sleep.Stage = S, phase = dark:
##  contrast  estimate  SE  df t.ratio p.value
##  McKO - WT  -425.78 322 133  -1.323  0.1881
##
## Sleep.Stage = W, phase = dark:
##  contrast  estimate  SE  df t.ratio p.value
```

```
##  McKO - WT  -162.78 322 133  -0.506  0.6139
##
## Sleep.Stage = X, phase = dark:
##  contrast  estimate  SE  df t.ratio p.value
##  McKO - WT      0.00 322 133   0.000  1.0000
##
## Sleep.Stage = P, phase = light:
##  contrast  estimate  SE  df t.ratio p.value
##  McKO - WT    -46.56 322 133  -0.145  0.8852
##
## Sleep.Stage = S, phase = light:
##  contrast  estimate  SE  df t.ratio p.value
##  McKO - WT      4.00 322 133   0.012  0.9901
##
## Sleep.Stage = W, phase = light:
##  contrast  estimate  SE  df t.ratio p.value
##  McKO - WT      5.67 322 133   0.018  0.9860
##
## Sleep.Stage = X, phase = light:
##  contrast  estimate  SE  df t.ratio p.value
##  McKO - WT      0.00 322 133   0.000  1.0000
##
## Degrees-of-freedom method: kenward-roger
```

```
lm_duration <- lmer(total_duration ~ Sleep.Stage * phase * Genotype + (1 | Transmitter), data=SumStats_
#plot(lm_duration, type=c("p","smooth"), col.line=1)
#lattice::qqmath(lm_duration)
anova(lm_duration)
```

```
## Analysis of Variance Table
##                          npar      Sum Sq     Mean Sq  F value
## Sleep.Stage                 3 3.0028e+11 1.0009e+11 256.2299
## phase                       1 7.2822e+06 7.2822e+06   0.0186
## Genotype                    1 1.6984e+07 1.6984e+07   0.0435
## Sleep.Stage:phase           3 5.4756e+10 1.8252e+10  46.7237
## Sleep.Stage:Genotype        3 1.1385e+09 3.7948e+08   0.9714
## phase:Genotype              1 4.1930e+04 4.1930e+04   0.0001
## Sleep.Stage:phase:Genotype  3 1.2895e+07 4.2982e+06   0.0110
```

```
emmeans(lm_duration, pairwise ~ Genotype | Sleep.Stage * phase)$contrasts
```

```
## Sleep.Stage = P, phase = dark:
##  contrast  estimate    SE   df t.ratio p.value
##  McKO - WT    -1098 11695 64.7  -0.094  0.9255
##
## Sleep.Stage = S, phase = dark:
##  contrast  estimate    SE   df t.ratio p.value
##  McKO - WT     5807 11695 64.7   0.497  0.6212
##
## Sleep.Stage = W, phase = dark:
##  contrast  estimate    SE   df t.ratio p.value
##  McKO - WT   -11038 11695 64.7  -0.944  0.3488
##
```

```
## Sleep.Stage = X, phase = dark:
##  contrast  estimate    SE   df t.ratio p.value
##  McKO - WT     -243 11695 64.7  -0.021  0.9835
##
## Sleep.Stage = P, phase = light:
##  contrast  estimate    SE   df t.ratio p.value
##  McKO - WT    -1576 11695 64.7  -0.135  0.8932
##
## Sleep.Stage = S, phase = light:
##  contrast  estimate    SE   df t.ratio p.value
##  McKO - WT     4211 11695 64.7   0.360  0.7200
##
## Sleep.Stage = W, phase = light:
##  contrast  estimate    SE   df t.ratio p.value
##  McKO - WT    -9408 11695 64.7  -0.804  0.4241
##
## Sleep.Stage = X, phase = light:
##  contrast  estimate    SE   df t.ratio p.value
##  McKO - WT      -64 11695 64.7  -0.005  0.9957
##
## Degrees-of-freedom method: kenward-roger
```

```r
sws_bouts <- SumStats_phase %>% filter(Sleep.Stage == "S" & phase == "light") %>%
  ggplot(aes(x=Genotype, y=mean_bout_length)) +
  stat_summary(fun = 'mean', geom="bar", width=0.75) +
  stat_summary(fun.data = mean_se, geom = "errorbar", width=.2) +
  geom_point() +
  ylab("Average SWS bout length in the light phase (s)") +
  theme_bw() +
  scale_y_continuous(expand = expansion(mult = c(0, 0.2)))

sws_bouts
```
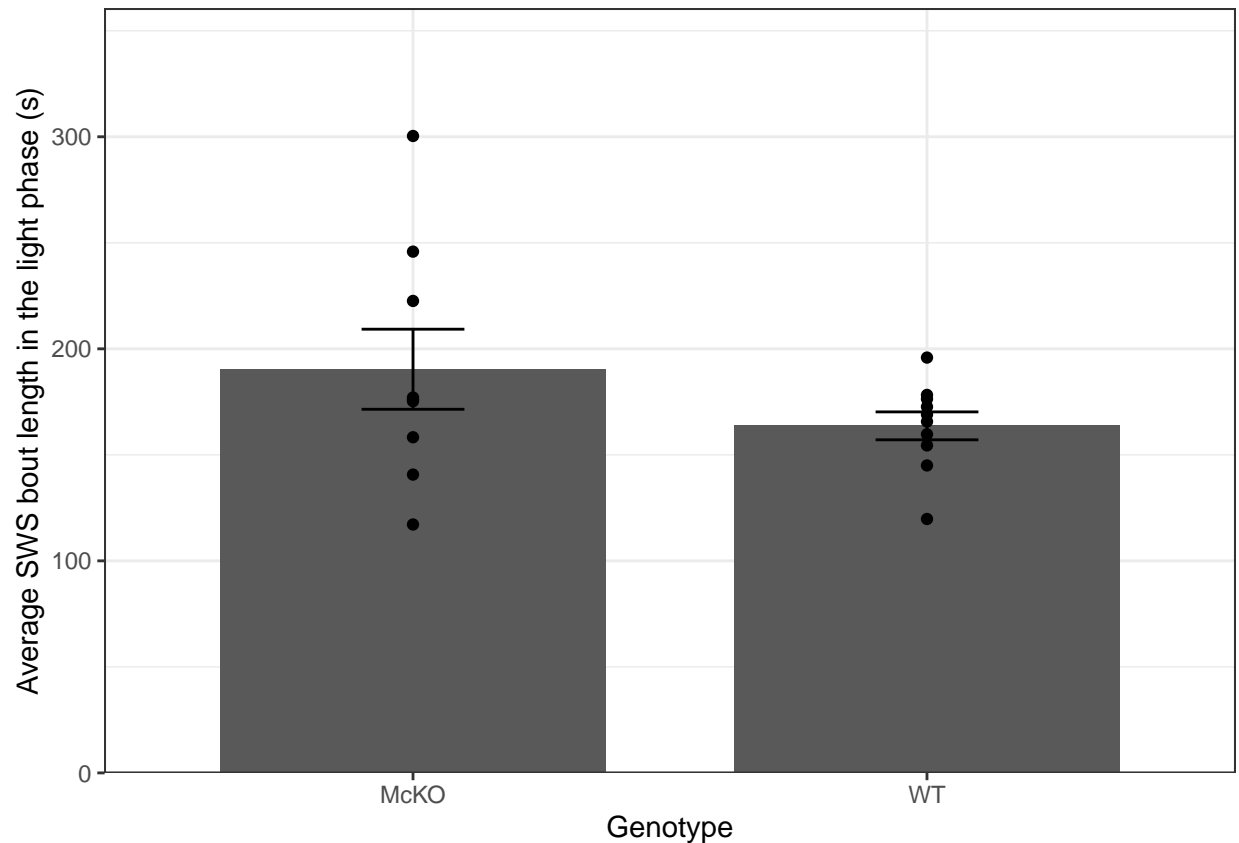
```
#ggsave(filename="../figures/sws_bouts.svg", plot=sws_bouts, height=4, width=3)
```

```
SumStats_phase$Sleep.Stage <- recode_factor(SumStats_phase$Sleep.Stage, S = "SWS", P = "REM", X = "Arti

length_phase_plot <- SumStats_phase %>% filter(Sleep.Stage != "Artifact") %>%
  ggplot(aes(x=phase, y=mean_bout_length)) +
  facet_wrap(~Sleep.Stage, scales="free_y") +
  geom_boxplot(aes(color=Genotype)) +
  ylab("Average bout length (s)") +
  theme_bw() +
  scale_y_continuous(expand = expansion(mult = c(0, 0.2)))

length_phase_plot
```
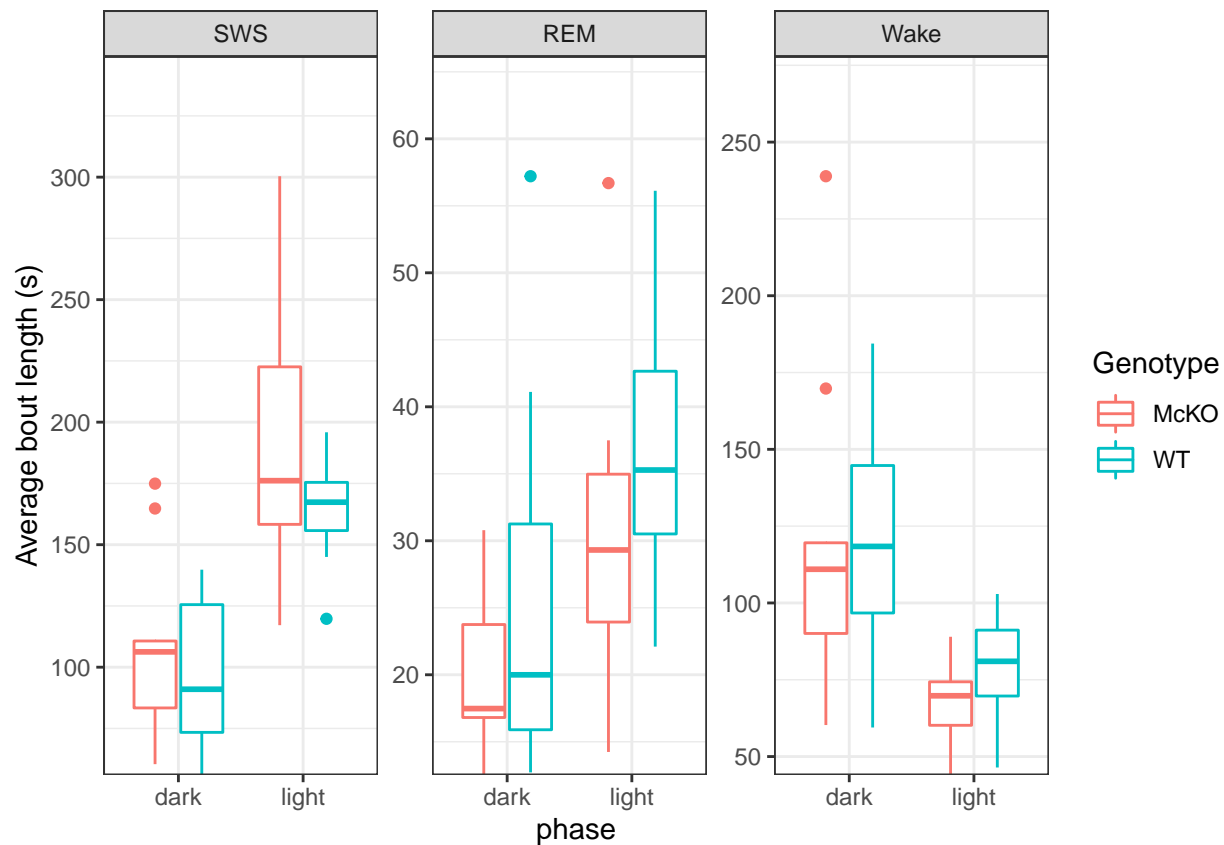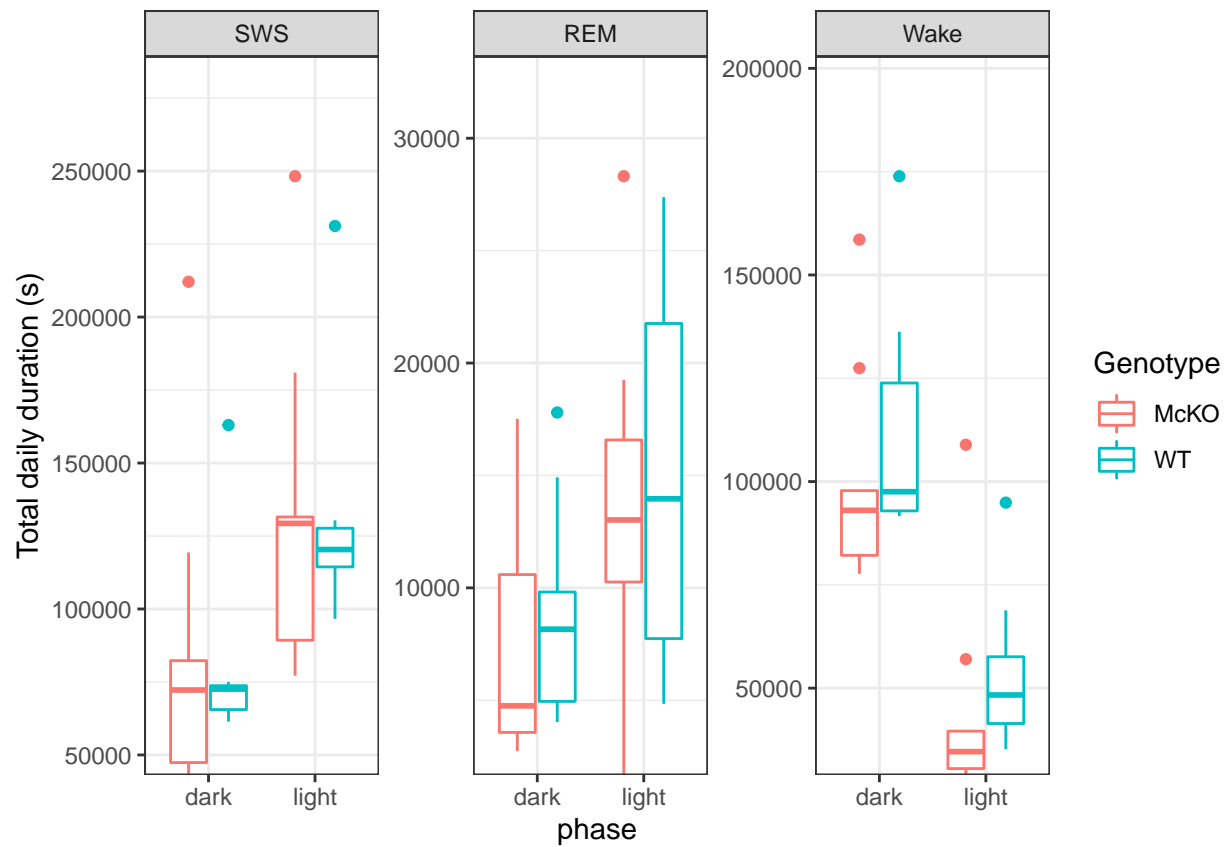
```r
daily_duration_plot <- SumStats_phase %>% filter(Sleep.Stage != "Artifact") %>%
  ggplot(aes(x=phase, y=total_duration)) +
  facet_wrap(~Sleep.Stage, scales="free_y") +
  geom_boxplot(aes(color=Genotype)) +
  ylab("Total daily duration (s)") +
  theme_bw() +
  scale_y_continuous(expand = expansion(mult = c(0, 0.2)))

daily_duration_plot
```

```
#ggsave(filename="../figures/daily_duration_plot.svg", plot=daily_duration_plot, height=4, width=7)
```