

Collagen QC

Angela Trentacoste
May 2023

Assessing bone collagen quality

- I have a SI values.....is it any good?
 - Collagen yield
 - C:N
 - %C
 - %N
 - Relationships between %C and C:N



- Load the data
- Prune the data and join tables to make our mega table
- Visualise different QC criteria
- Use some new syntax and function

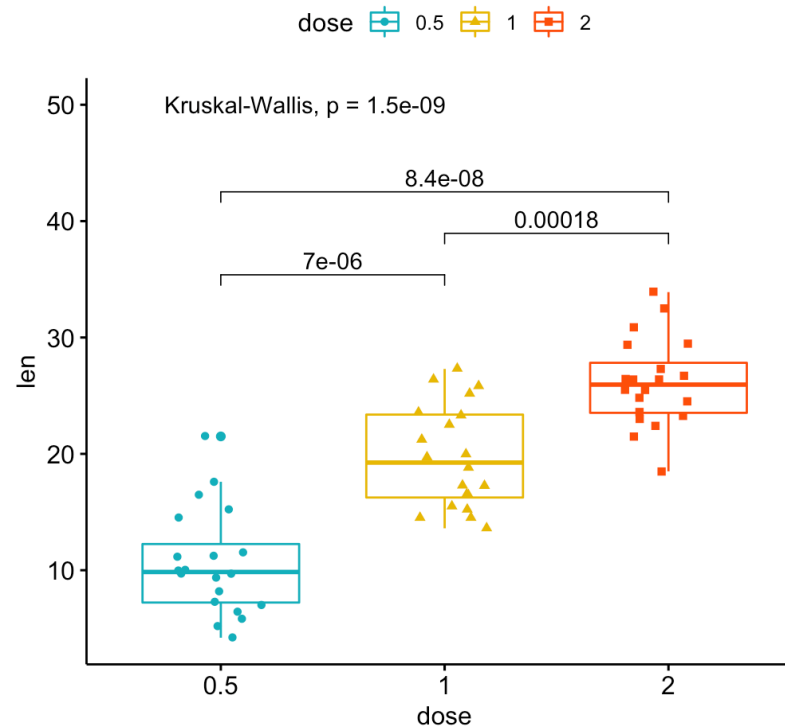
%in% select values based on vector

```
sheep.goat <- c("Ovis", "Capra") # a character vector with 2 values
```

```
subset(Data, Taxon %in% sheep.goat) # select rows where Taxon is in vector sheep.goat
```

ggpubr package

- Easy publication-oriented plots, especially when stats are included
- to customize a ggplot, the syntax is opaque sometimes hard for newbies
- **ggpubr** provides some easy-to-use functions for creating and customizing 'ggplot2'- based plots



Stats and data manipulation

Angela Trentacoste
May 2023

Already done a lot!

- Joins
- Data manipulation
operators + - * /

left_join()
mutate()
replace()
group_by()
summarise()
filter()
subset()



- Extreme outliers? if no, then...
- Shapiro-Wilk's test - are all groups normally distributed?
- Levene's test – homogeneity/variance
 - normal? yes, 3+ group = ANOVA
 - normal? no, 3+ group = Kruskal-Wallis
 - normal? yes, 2 groups = t-test
 - normal? no, 2 group = Mann-Whitney U test

Data sets in tutorials



Data sets in the wild

