

How do body size and temperature affect metabolic rates?

Background

The metabolic rates of organisms are strongly influenced by the organism's body size and body temperature. There is great variation in body size among the life forms on earth, ranging from tiny bacteria to great whales and giant sequoia trees. Even within a species, body size can vary greatly. Larger organisms have higher metabolic rates than smaller organisms. Knowledge of how biological rates vary with size provides critical information about how individuals of different sizes use resources, how they process materials such as nutrients and energy, and therefore how they interact with other species.

Because the rates of most biochemical reactions are temperature dependent, most metabolic rates also increase with body temperature, at least up to the upper tolerance limit for that organism. Organisms experience a wide range of temperatures, both due to natural variation as well as humans' effects on temperature. Knowledge of how biological rates respond to temperature can therefore inform us as to how organisms, and ultimately ecosystems, may be affected by changes in temperature.

The increase in metabolic rate is often not proportional to the increase in body mass. The study of how metabolic rate (or any other rate) varies with body size is called allometry. Precisely how biological rates change with body size is referred to as allometric scaling. Allometric scaling has been studied in many organisms and for many biological rates. There are profound consequences of an allometric scaling exponent that is substantially less than 1. For example, an elephant (5×10^6 g) weighs about 500,000 times more than a mouse (10 g). However, if their metabolic rates vary with body size with an allometric scaling exponent of 0.75, the elephant's metabolic rate will be only about 19,000 times greater than that of the mouse – proportionally much less than the difference in body mass. For organisms whose body temperature varies with environmental temperature (ectotherms, or “cold blooded” organisms), many rates vary directly with environmental temperature.

The Metabolic Theory of Ecology (MTE) predicts that we can relate body size, temperature, and metabolic rate. Read more about The Metabolic Theory of Ecology at https://en.wikipedia.org/wiki/Metabolic_theory_of_ecology .

Exercise

In this exercise, we will explore how nutrient excretion rates vary with body size and temperature. We will use a particular fish species, the gizzard shad *Dorosoma cepedianum*, as a model organism for this study. Gizzard shad is probably the most abundant fish species (based on species biomass) in lakes of the southern and lower Midwest USA (Vanni et al. 2005).

This species is omnivorous but mostly consumes sediment detritus from the bottom of lakes, i.e., these fish eat mud. Thus, they consume nitrogen (N) and phosphorus (P) contained in sediment detritus, use some of the N and P for growth and reproduction, and release the leftovers. Because they are so abundant and consume large amounts of sediment detritus (it's not easy to make a living eating mud!), gizzard shad consume and later excrete large quantities of N and P, which are the elements most likely to limit the growth of algae (Vanni et al. 2005). By excreting N and P in bioavailable forms, gizzard shad provide nutrients for algae to use. In some lakes, the flux of nutrients from sediments to water mediated by gizzard shad is quantitatively important compared to other nutrient fluxes, and can fuel a significant percentage (>25%) of algae primary production (Vanni et al. 2006). Thus, although this exercise focuses on excretion by individual fish, it is important to keep in mind the ecosystem-scale consequences (e.g., Schaus et al. 2010).

The rate at which an animal excretes nutrients is an example of a metabolic rate, in the sense that this rate represents the quantity of nutrient released after metabolic and structural needs are met. For example, the catabolism of proteins produces nitrogen-rich waste products, which animals then release as urine (or other excretory product). The rate at which consumers excrete nutrients is important not only for understanding metabolic ecology, but also for understanding ecosystem-scale nutrient fluxes (Sturner and Elser 2002).

Most tests of the effects of temperature and body size on metabolic rates have been done in the lab, where activity, temperature and diets can be controlled, and where metabolism is measured as the basal metabolic rate. This is the metabolic rate of an organism that is inactive, experiences a “neutrally thermal” temperature, and has not eaten for some time (i.e., has fasted). Use of such controlled and restrictive conditions allows researchers to minimize external influences (activity, extreme temperatures, feeding) that might increase variability in metabolic rate. However, organisms in nature are often active, experience a wide range of temperatures, and spend some time feeding; therefore, they would display active metabolic rates. Activity generally increases metabolic rates because more energy (and in aerobic organisms, oxygen) is needed under active conditions, and the speed of many biochemical reactions increases to deliver energy and oxygen to cells. Feeding can also increase metabolic rate, because it takes energy to capture, consume and digest food. Also, an animal will release (via excretion as well as defecation) any nutrients it does not need for growth and reproduction. Thus, all else being equal, when an animal consumes more nutrients, it also excretes more nutrients.

Our study investigates whether temperature and body size effects are also observed under field conditions, where other factors such as diet, feeding rate and activity level, can come into play. Such field-based studies of allometric scaling and temperature dependence are somewhat rare.

Hypotheses

1. N excretion rate of gizzard shad scales with body size as predicted by the Metabolic Theory of Ecology, i.e. the slope of the log N excretion rate vs. log body mass is ~ 0.75 .

2. P excretion rate of gizzard shad scales with body size as predicted by the Metabolic Theory of Ecology, i.e. the slope of the log P excretion rate vs. log body mass is ~ 0.75 .

We can offer alternatives to hypotheses 1 and 2, because some theory as well as data suggest that the allometric scaling exponent (slope on a log-scale) is closer to 1 than 0.75. Therefore, our alternative hypotheses are:

1. N excretion rate of gizzard shad scales with body size, i.e. the slope of the log N excretion rate vs. log body mass is ~ 1 .
2. P excretion rate of gizzard shad scales with body size, i.e. the slope of the log P excretion rate vs. log body mass is ~ 1 .

Data

Published excretion rates on 200 fish, measured in three lakes (Schaus et al. 1997, Higgins et al. 2006), are used in this exercise. Gizzard shad were collected by electrofishing from the upstream end of the lake and were immediately transferred to clean containers holding prefiltered lake water. After 1 h, a sample of water was collected from each container for determination of final nutrient concentrations. Gizzard shad ranged from 2.32 to 210.9 g wet mass, which is representative of the gizzard shad population.

The dataset consists of a spreadsheet with 4 data columns, each with 200 rows. Each row (observation) represents a different fish, whose excretion rate was measured in one of three lakes in Ohio (Acton, Burr Oak and Pleasant Hill lakes).

Column 1: the temperature experienced by the fish (i.e., water temperature where gizzard shad live at the time of the experiment).

Column 2: the mass of each fish, in grams (wet mass).

Columns 3: the per individual nitrogen excretion rate of each fish, in units of μmol of N excreted per fish per hour. Excretion rates were measured in the field, using a relatively simple method. Fish are captured and placed in a known volume of filtered lake water (1-4 liters depending on fish size) for a known period of time (usually an hour). Water samples are taken before and after fish are incubated, to estimate nutrient concentrations. Excretion rate is quantified simply as the change in nutrients over time, during the incubation. The lake water is filtered before experiments to remove particles such as algae and bacteria, which would otherwise take up nutrients released by fish. These methods are described in detail in Schaus et al. (1997).

Columns 4: the per individual phosphorus excretion rate of each fish, in units of μmol of P excreted per fish per hour.

What parameter(s) are you trying to estimate?

What is the target population you are interested in?

What samples can you use to estimate the parameter(s)?

Do you expect each sample to provide an accurate estimate of the parameter(s) of interest? Why or why not?

What are the explanatory and response variables in this dataset? Are they categorical or numerical? Discrete or continuous?

Getting started with data analysis in R

This class assumes you used R and RStudio in Bio 104. If you haven't used R (or you need a review) follow the instructions at https://github.com/rachelss/BIO103R/blob/master/A1-plankton_sampling/field_sampling.Rmd (hereafter referred to as Intro Tutorial) step 1 to make a new project and a new script file.

For this exercise you will be provided with all the code. After this it will be assumed that you have learned the code for these steps and only code for new analyses will be provided.

Load the necessary packages

For an explanation see Intro Tutorial step 2a.

```
library(tidyverse)
```

Load data into R

```
N_excretion <- read.csv('../shared/fish_nutrient_excretion_students.csv')
```

Explore the data

If you're not sure how to do this on your own see part 2d of the Intro Tutorial.

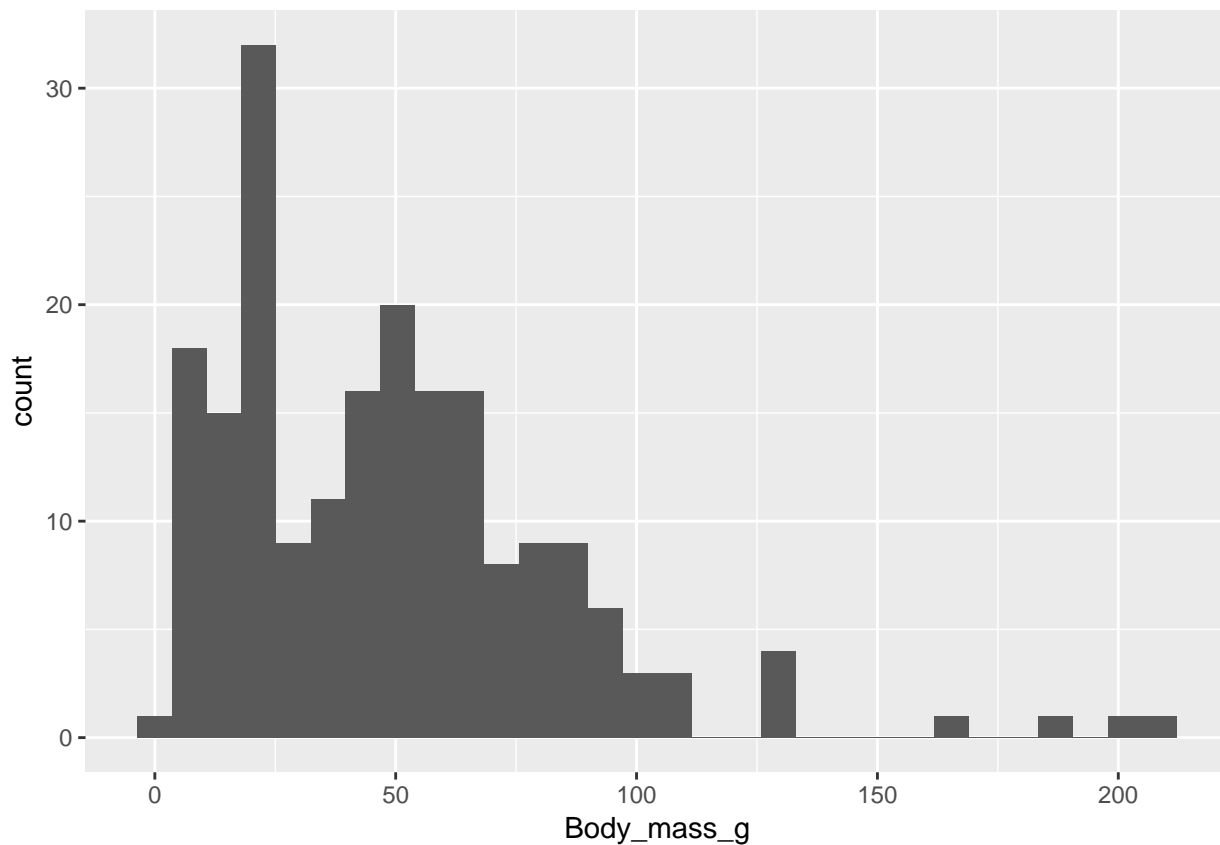
Comment your script and save it (Intro Tutorial part 3).

Examine the frequency of fish of different sizes

Use a histogram with mass on the x axis (in ranges) and the count of each range of mass on the y axis.

For help see Intro Tutorial part 4.

```
ggplot(N_excretion, aes(Body_mass_g)) +  
  geom_histogram()
```

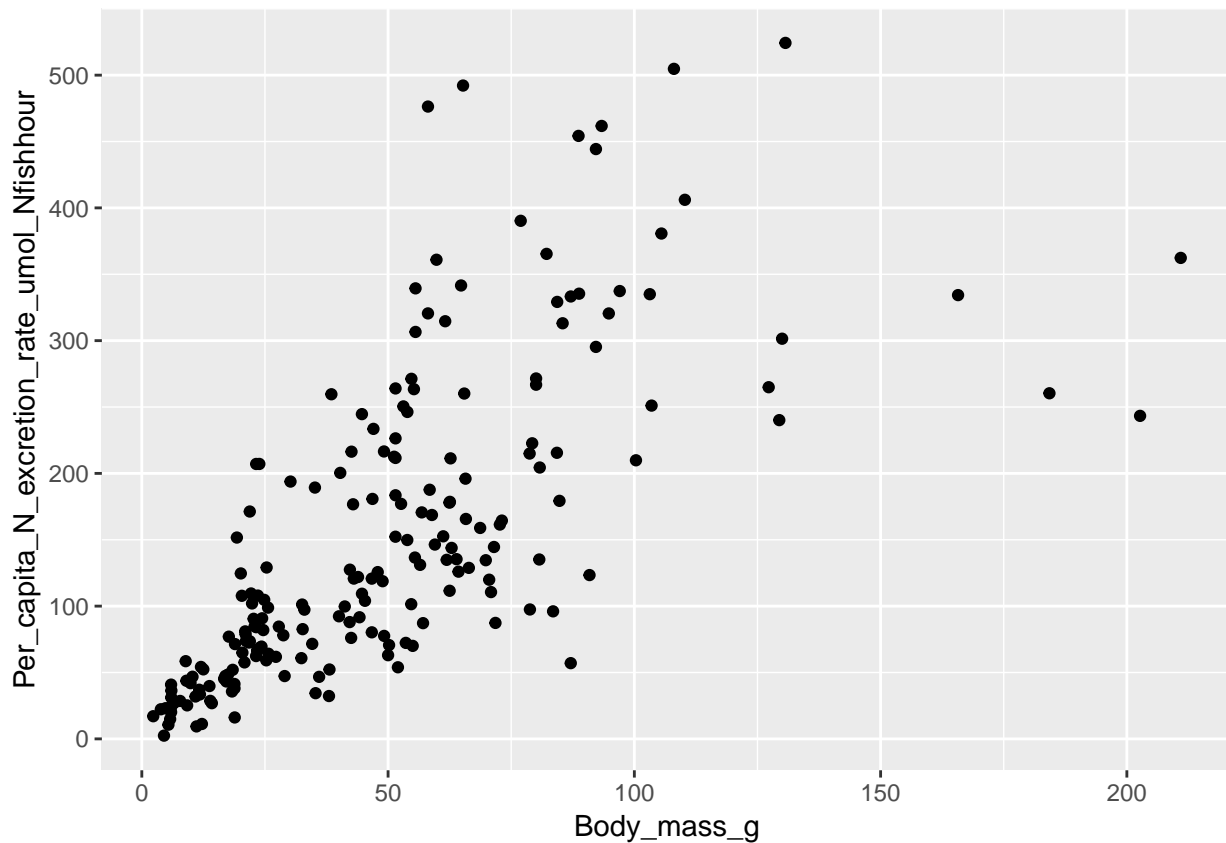


What can you say about the size of fish being caught?

Examine the relationship between N excretion rate and body size

Plot excretion rates

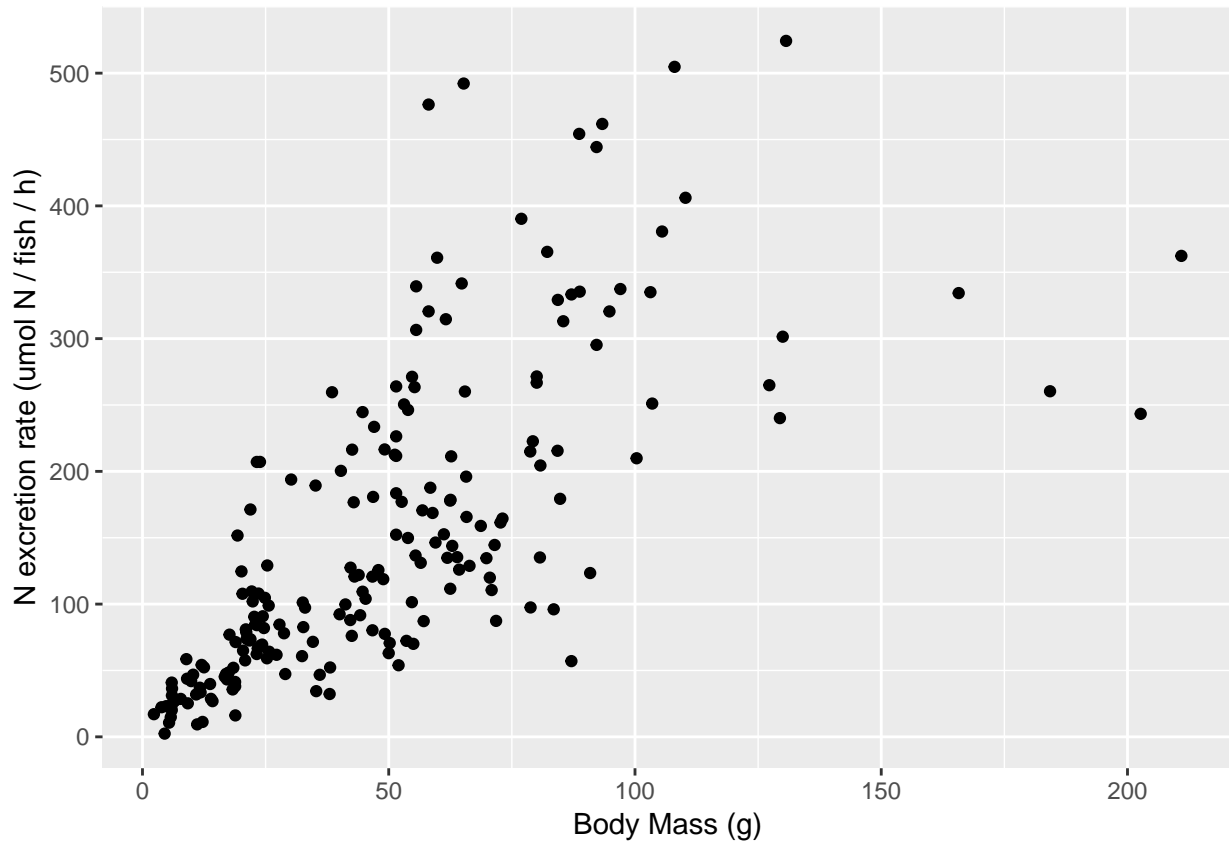
```
ggplot(N_excretion, aes(x=Body_mass_g,  
                        y=Per_capita_N_excretion_rate_umol_Nfishhour)) +  
  geom_point()
```



Can you see the variation in rate among samples of a given body size? Try looking at N excretion for all fish of 50g. What are the smallest and largest excretion rates?

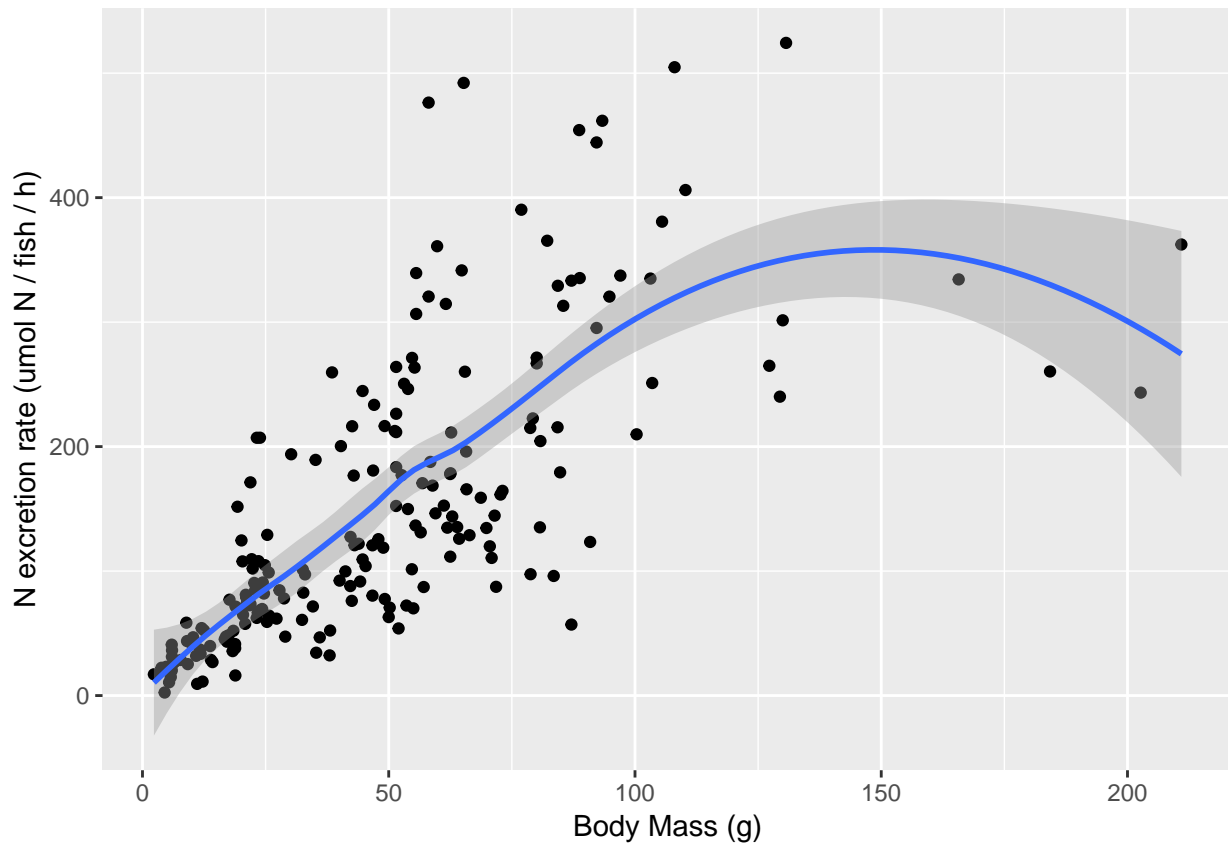
Fix the axis labels by adding layers to your plot (see Intro Tutorial part 5a). Make sure to think about whether your variables are discrete or continuous.

```
ggplot(N_excretion, aes(x=Body_mass_g, y=Per_capita_N_excretion_rate_umol_Nfishhour)) +  
  geom_point() +  
  scale_x_continuous(name = "Body Mass (g)") +  
  scale_y_continuous(name = "N excretion rate (umol N / fish / h)")
```



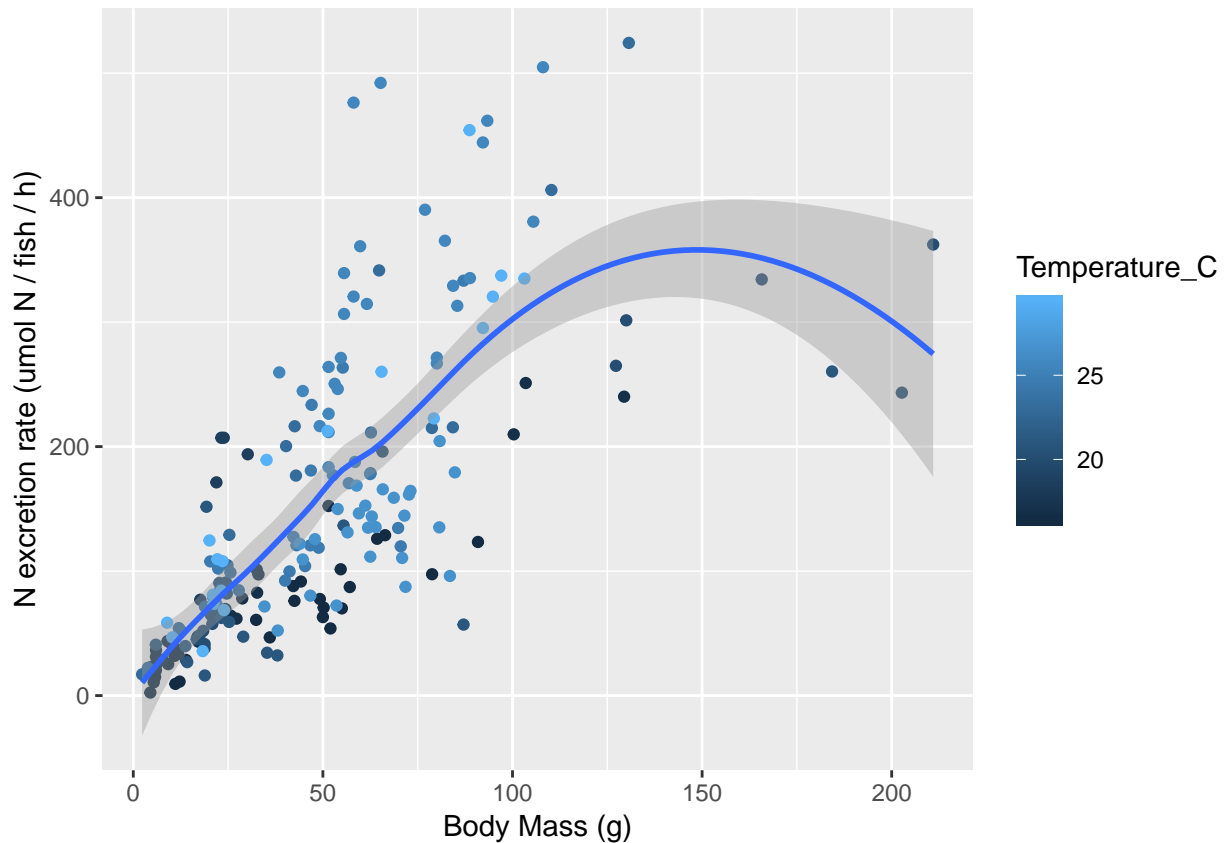
Add a trend line using `geom_smooth()`.

```
ggplot(N_excretion, aes(x=Body_mass_g, y=Per_capita_N_excretion_rate_umol_Nfishhour)) +  
  geom_point() +  
  scale_x_continuous(name = "Body Mass (g)") +  
  scale_y_continuous(name = "N excretion rate (umol N / fish / h)") +  
  geom_smooth()
```

We know temperature has an effect so let's color our points by temperature. Plot N excretion rate and a function of mass separately for each temperature by adding `color=Temperature_C` in your `aes()` function for `geom_point`.

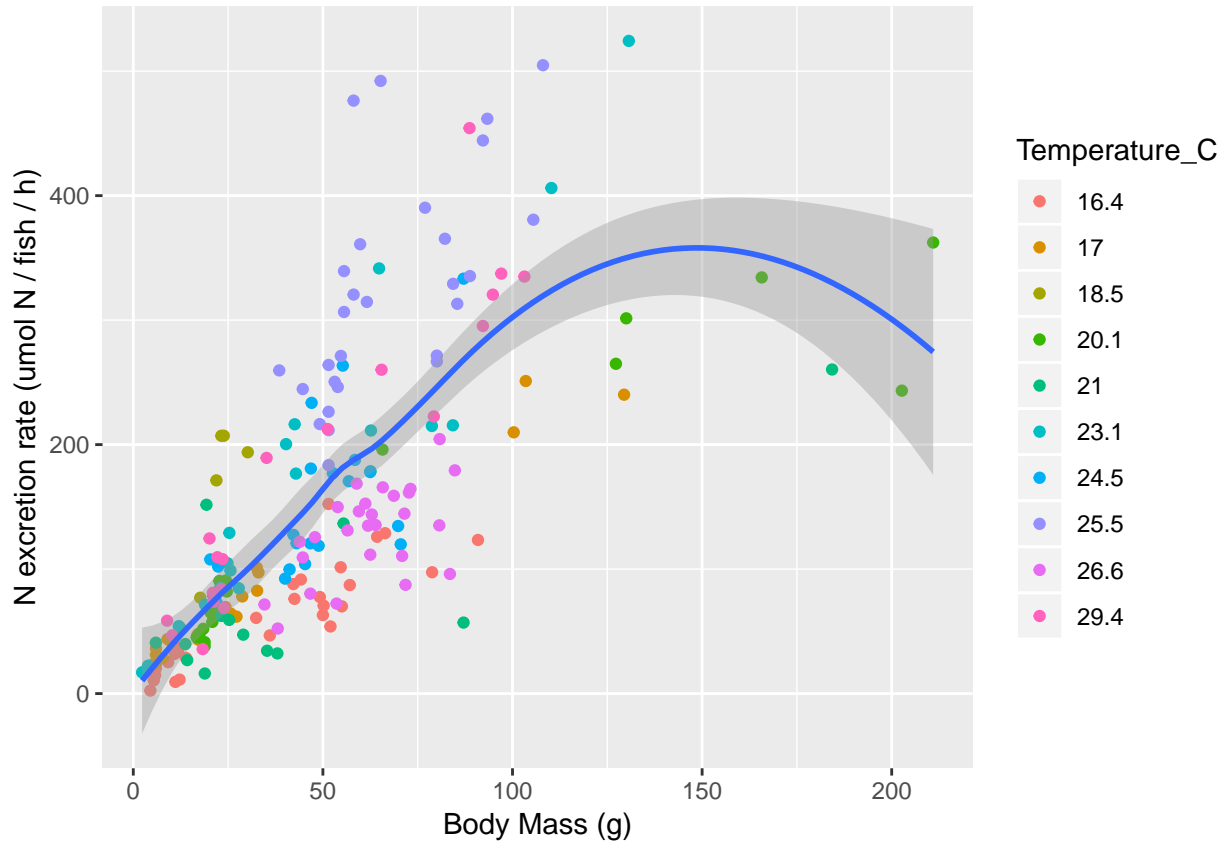
```
ggplot(N_excretion, aes(x=Body_mass_g, y=Per_capita_N_excretion_rate_umol_Nfishhour)) +  
  geom_point(aes(color=Temperature_C)) +  
  scale_x_continuous(name = "Body Mass (g)") +  
  scale_y_continuous(name = "N excretion rate (umol N / fish / h)") +  
  geom_smooth()
```



Oops...R sees temperature as a continuous variable, but we're thinking about different temperatures as different categories. Let's tell R this. *Note that you need to specify that **Temperature_C** is a factor (i.e. discrete) and not a continuous variable in order to create separate plots for each temperature.*

```
N_excretion$Temperature_C <- as.factor(N_excretion$Temperature_C)

ggplot(N_excretion, aes(x=Body_mass_g, y=Per_capita_N_excretion_rate_umol_Nfishhour)) +
  geom_point(aes(color=Temperature_C)) +
  scale_x_continuous(name = "Body Mass (g)") +
  scale_y_continuous(name = "N excretion rate (umol N / fish / h)") +
  geom_smooth()
```



Describe how excretion rates vary with temperature.

Log scale the plot

The Metabolic Theory of Ecology (MTE) predicts that we can relate body size and metabolic rate with the equation:

$$B = B_0 M^{0.75} \quad (1)$$

where

B is an organism's metabolic rate (e.g., grams of oxygen consumed, or nitrogen excreted, per day)

B_0 is a constant that is fitted to the data

M is the organism's body mass

The exponent 0.75 is the “allometric scaling exponent” predicted by MTE. By taking the logarithm of both sides of equation 1, we can express this relationship as:

$$\log B = \log B_0 + 0.75 \log M \quad (2)$$

This equation describes a linear relationship in which $\log B_0$ is the intercept and 0.75 is the slope. If metabolic rate increased in proportion to body mass, the exponent in Equation 1

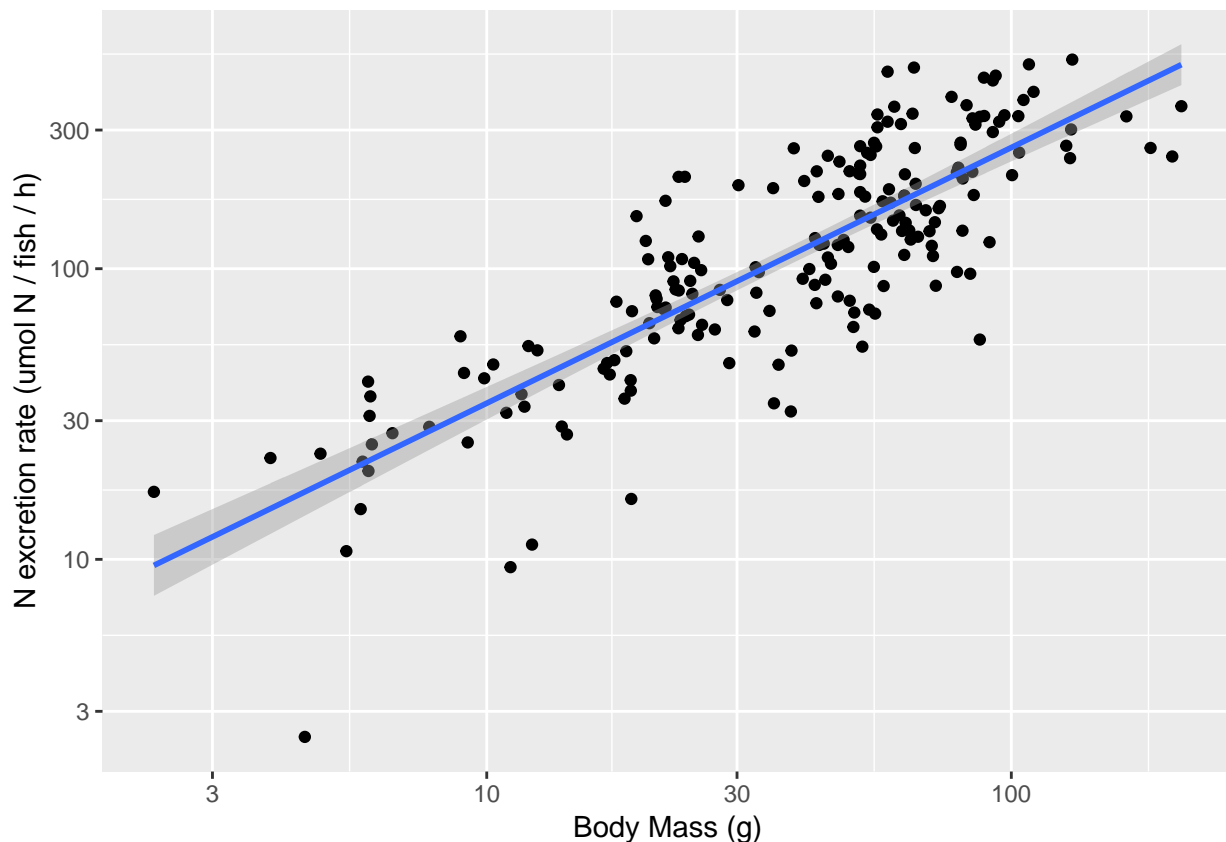
(slope in Equation 2) would be 1.

Change the axis scales to log scale by replacing `scale_x_continuous` and `scale_y_continuous` with `scale_x_log10` and `scale_y_log10`.

Once you've made these log plots, you can analyze them using linear regression. Linear regression is a technique whereby a straight line is “fit” to the data. Linear regression can be used to determine whether the two variables are associated, and the strength of the association. For example, suppose the plot suggests that the two variables are associated, but there is “scatter” among the data, i.e., the points do not all fit along a straight line. Linear regression produces the straight line that best fits the data (basically, it produces the line that minimizes the cumulative distances between the data points and the line itself).

Change the best-fit trend line to be straight by specifying `method="lm"` in the `geom_smooth` function.

```
ggplot(N_excretion, aes(x=Body_mass_g,  
                        y=Per_capita_N_excretion_rate_umol_Nfishhour)) +  
  geom_point()+  
  scale_x_log10(name = "Body Mass (g)") +  
  scale_y_log10(name = "N excretion rate (umol N / fish / h)") +  
  geom_smooth(method="lm")
```



The line on your plot represents a linear regression. The slope of the regression line is of

great interest because MTE predicts a slope of 0.75 for the log body mass vs. log excretion rate relationship, whereas our alternative hypothesis predicts a slope of 1.

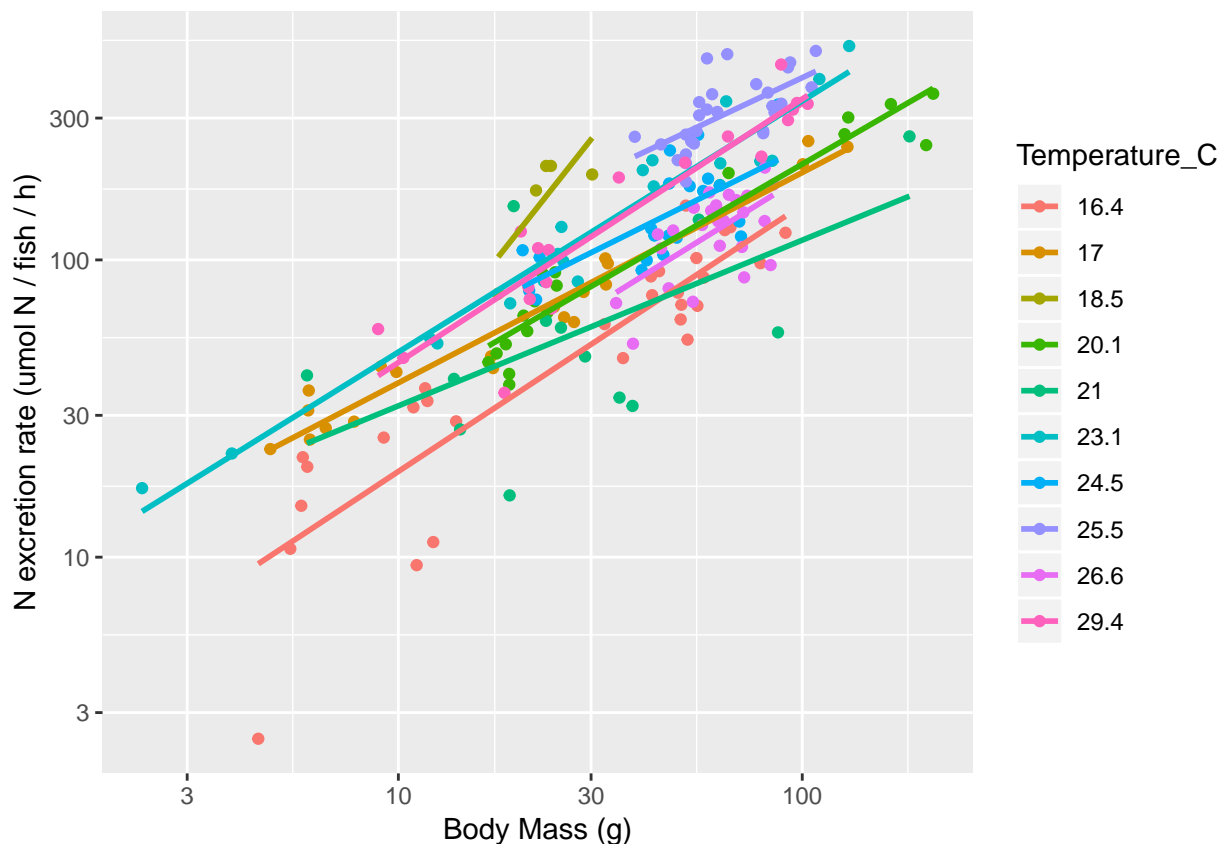
Down below we will examine the estimated slope of the line to determine whether our data rejects our null hypothesis.

Save your plot (see Intro Tutorial part 5b).

Create separate plots for each temperature

We expect metabolic rate to depend on both mass and temperature. Plot N excretion rate as a function of mass separately for each temperature by adding `color=Temperature_C` in your `aes()` function.

```
ggplot(N_excretion, aes(x=Body_mass_g,
                        y=Per_capita_N_excretion_rate_umol_Nfishhour,
                        color=Temperature_C)) +
  geom_point() +
  scale_x_log10(name = "Body Mass (g)") +
  scale_y_log10(name = "N excretion rate (umol N / fish / h)") +
  geom_smooth(method="lm", se = FALSE)
```



Does temperature impact the slope of the line?

Does temperature impact the y-intercept of the line

What is the implication of these effects?

Statistics

The lines on your plot represent a linear regression, which allow you to estimate the slope (and intercept) of the line.

In the `lm` command you should enter your dependent variable *as a function of* your independent variable using the `~`.

First we will do this calculate for all the data just to show you how to calculate a linear regression in a simple way *even though we know that temperature impacts N excretion*. Using all the data in one model, instead of modeling temperatures separately, means we'll have a lot more variation and may not get the answer we expect.

```
#calculate log of all N excretion data
log_Nexcretion <- log10(N_excretion$Per_capita_N_excretion_rate_umol_Nfishhour)
#calculate log of all mass
log_mass <- log10(N_excretion$Body_mass_g)
#linear regression with logN as a function of log_mass - store as variable
N_lm <- lm(log_Nexcretion ~ log_mass)
#print info in linear regression variable
summary(N_lm)
```

```
##
## Call:
## lm(formula = log_Nexcretion ~ log_mass)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.84300 -0.10916  0.00468  0.15414  0.46843
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   0.65753    0.06798   9.673  <2e-16 ***
## log_mass      0.87971    0.04231  20.791  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2196 on 198 degrees of freedom
## Multiple R-squared:  0.6858, Adjusted R-squared:  0.6843
## F-statistic: 432.3 on 1 and 198 DF, p-value: < 2.2e-16
```

Check the coefficients (in the middle of the output of the summary of the model). Is the estimate of the coefficient (i.e. the slope) for `log_mass` what you expect? *Hint: think about*

your null hypothesis - what relationship do you expect from equation 2? what is the coefficient in the equation?

Now calculate the model separately for different temperatures

We know temperature affects metabolic rates so we can estimate the model separately for each temperature. Think of each temperature as a separate sample with a separate estimate. First let's calculate the regression model for excretion at 20.1 C.

```
temp20 <- filter(N_excretion, Temperature_C == 20.1)
log_temp_20_N <- log10(temp20$Per_capita_N_excretion_rate_umol_Nfishhour)
log_temp_20_mass <- log10(temp20$Body_mass_g)
N_lm20 <- lm(log_temp_20_N ~ log_temp_20_mass)
summary(N_lm20)
```

```
##
## Call:
## lm(formula = log_temp_20_N ~ log_temp_20_mass)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.177487 -0.037798  0.008156  0.065569  0.141290
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.74759    0.09102   8.213 3.93e-07 ***
## log_temp_20_mass 0.78728    0.05541  14.208 1.72e-10 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.09595 on 16 degrees of freedom
## Multiple R-squared:  0.9266, Adjusted R-squared:  0.922
## F-statistic: 201.9 on 1 and 16 DF,  p-value: 1.719e-10
```

What is the coefficient of the linear model?

Think of each temperature as one estimate of the slope. This allows you to have multiple samples. Calculate the regression models for all the temperatures separately.

We will use the `tidy` function in the `broom` library to access the important information in each model. We can then put all the rows of data together using `rbind`. Below is an example with **some** of the data.

```
library(broom)
```

```
#note you just get the 2nd row of the tidied data
#as with the summary of the model, the second row contains the results
#related to log_mass
```

```

tidy_N_lm20 <- tidy(N_lm20)[2,]

#calculate the regression for temp = 21
temp21 <- filter(N_excretion, Temperature_C == 21)
N_lm21 <- lm(log10(temp21$Per_capita_N_excretion_rate_umol_Nfishhour) ~
             log10(temp21$Body_mass_g))
#get the coefficients of the model for 21
tidy_N_lm21 <- tidy(N_lm21)[2,]

#make one data table with all data - you should include all models
#calculated for individual temperatures
model <- rbind(tidy_N_lm20,tidy_N_lm21)

```

In some cases you might not have enough data to estimate the slope of the line very well. In these cases you can examine whether the slope is significantly from 0 and remove any non-significant results.

Calculate the mean slope among your samples. **Note how many samples you have.**

```

#filter out data where slope is not significantly different than 0
model_filtered <- filter(model, p.value < 0.05)
corrN <- mean(model_filtered$estimate)
corrN

```

```
## [1] 0.7598411
```

The 95% confidence interval around your mean slope is $mean \pm 2SE$.

```

#computation of the standard error of the mean
sem<-sd(model_filtered$estimate)/sqrt(length(model_filtered$estimate))
#95% confidence intervals of the mean
#lower bound
corrN-2*sem

```

```
## [1] 0.6799101
```

```

#upper bound
corrN+2*sem

```

```
## [1] 0.8397721
```

What do the mean and confidence interval of your slopes tell you about your hypothesis?

Are the slopes near 0.75, as predicted by MTE? Or are they closer to 1, as predicted by the alternative hypotheses?

Examine the relationship between P excretion rate and temperature

Repeat the work above for phosphorus excretion.

Writeup

Provide plots and figure captions stating results. Include statistical results in captions. Note p. 159 for advice on reporting statistics.

Make sure to read ch. 6 on hypothesis testing.

Notes

This material is adapted and excerpted from

M.J. Vanni and J.A. Gephart. 2011. Metabolic ecology: How do body size and temperature affect nutrient cycling rates? Teaching Issues and Experiments in Ecology 7: March. https://tiee.esa.org/vol/v7/issues/data_sets/vanni/abstract.html

M. H. Schaus M. J. Vanni T. E. Wissing M. T. Bremigan J. E. Garvey R. A. Stein. Nitrogen and phosphorus excretion by detritivorous gizzard shad in a reservoir ecosystem. 1997. Limnology and Oceanography 42: 1386-1397.