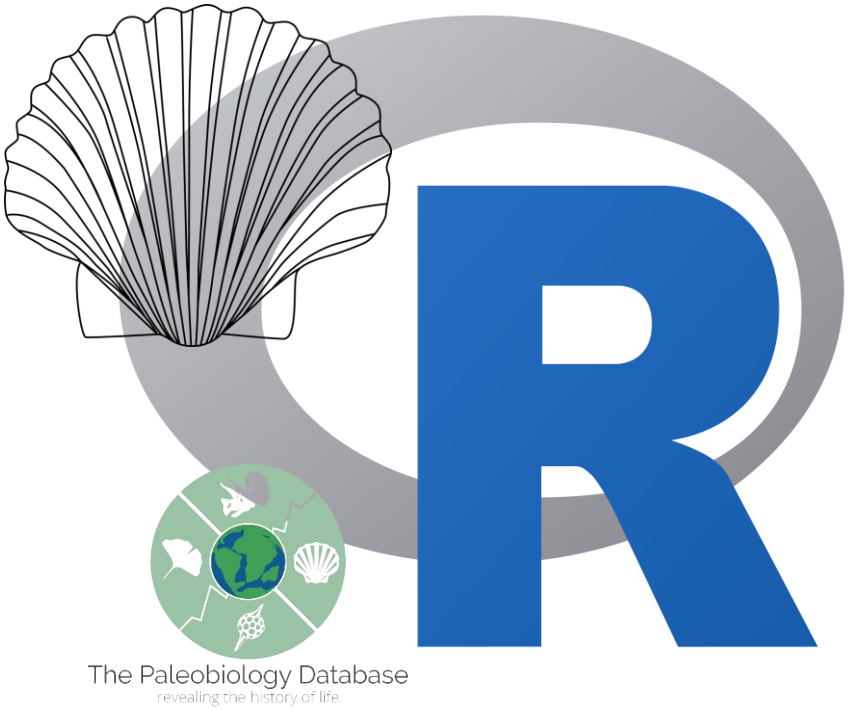
SOEE2145

Palaeoecology, Palaeobiology and Evolution

The quality of the Phanerozoic bivalve fossil record

Demonstrator copy



Dr Alex Dunhill

The Phanerozoic bivalve fossil record

Introduction

In this practical you will run a script of code in R. The script will allow you to plot bivalve diversity through the Phanerozoic (Ordovician-Pliocene) and apply a number of tests and corrections for sampling biases. Throughout the practical you will be asked to answer a number of questions which will form the basis of the practical assessment.

The code you will run is almost identical to that used in the primer session last week. If you are unsure of anything, ask me or the demonstrators, we are here to help you.

Data

The data set you are using today was download from the Paleobiology Database and contains all bivalve occurrences from the Ordovician to the Pliocene. See the stratigraphic column opposite for reference to time period names and dates. Your data is binned at the Stage level and you will be plotting it by millions of years. Some of the questions will ask you about certain time periods, use the stratigraphic column to give your answers. References to the Period or Epoch level are acceptable. I don't expect you to count the dots to Stage level.

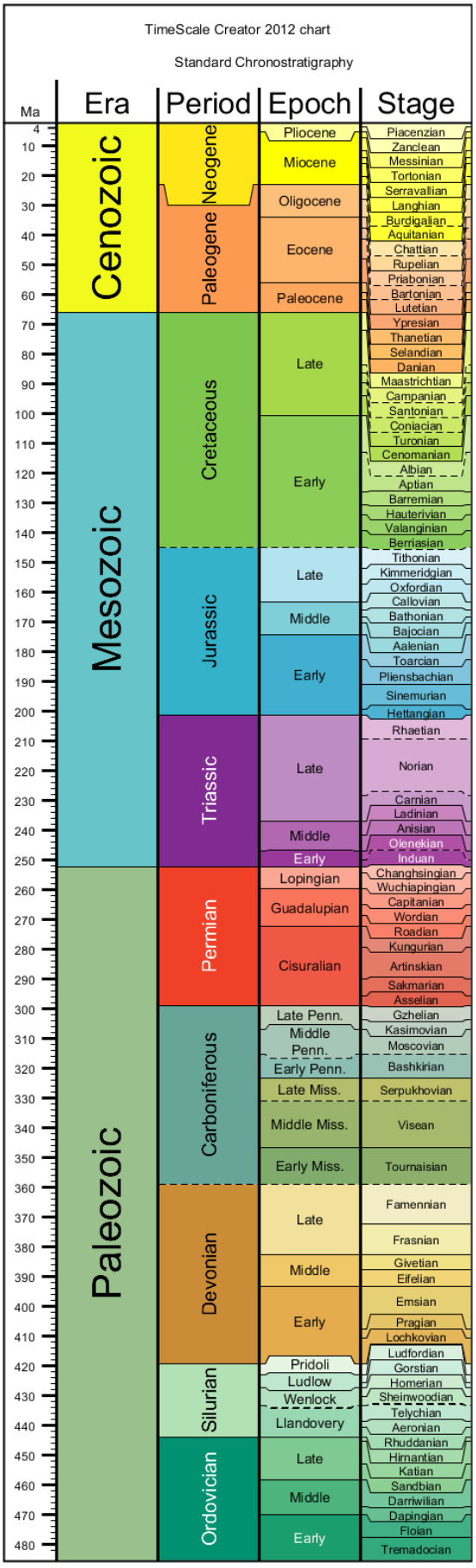
Lets' get started…

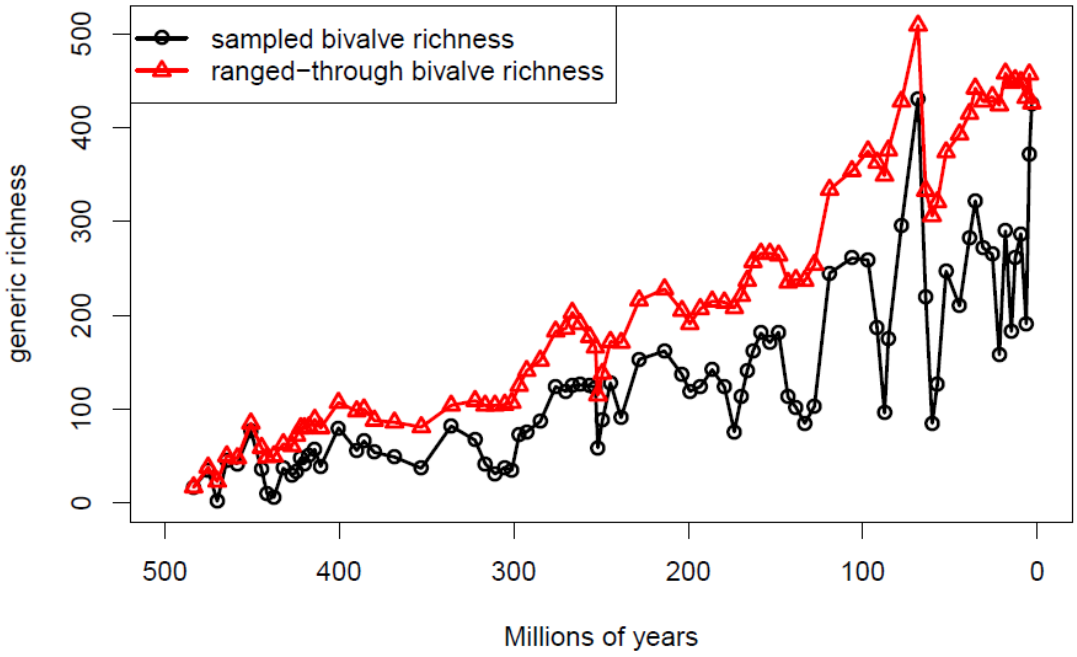
Open R and use the Open script option on the File menu to locate the script called "QotFR practical" which you downloaded from the VLE. Make sure this is in the same directory as last week's script and data.

Head up your script, run the code to set the working directory and load the data file entitled palaeo_data1.csv.

(1) Run the code to plot bivalve richness & ranged-through diversity through time and export it as a pdf.

Note that the `xlim=c()` values have changed to account for the Phanerozoic timescale. A `ylim=c()` command has been added to the code as ranged-through bivalve diversity is very different to sampled richness as bivalve genera have very long stratigraphic ranges. Dinosaur sampled richness and ranged-through diversity were very similar as dinosaur genera are generally short-lived.

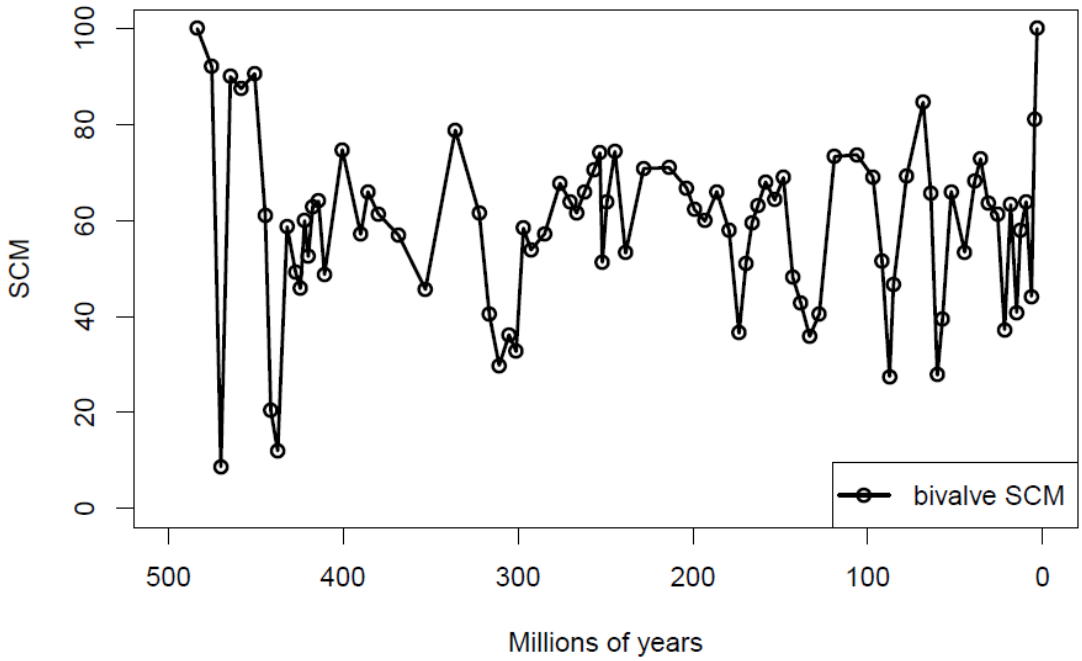




Question 1: Describe the trajectory of bivalve diversity through the Phanerozoic and explain why ranged-through richness and sampled richness differ? Can you identify any major evolutionary events? (10 marks)

Sampled richness and ranged-through richness both show a rising trend through the Phanerozoic. Ranged-through richness will always be higher than, or equal to, sampled richness as ranging through fills in the gaps between a taxon's first and last occurrence. Therefore, sampled richness only records taxa that occur in any particular time bin whereas ranged-through richness records occurrences+gaps. This plot displays the rise in diversity through the Mesozoic and Cenozoic to the present day and it is possible to identify some mass extinction events (i.e. End Permian @250Ma, KPg @66Ma).

(2) Run the code to calculate bivalve Simple Completeness Metrics (SCM) through the Phanerozoic. Bind the results to the age data and plot and export as a pdf.



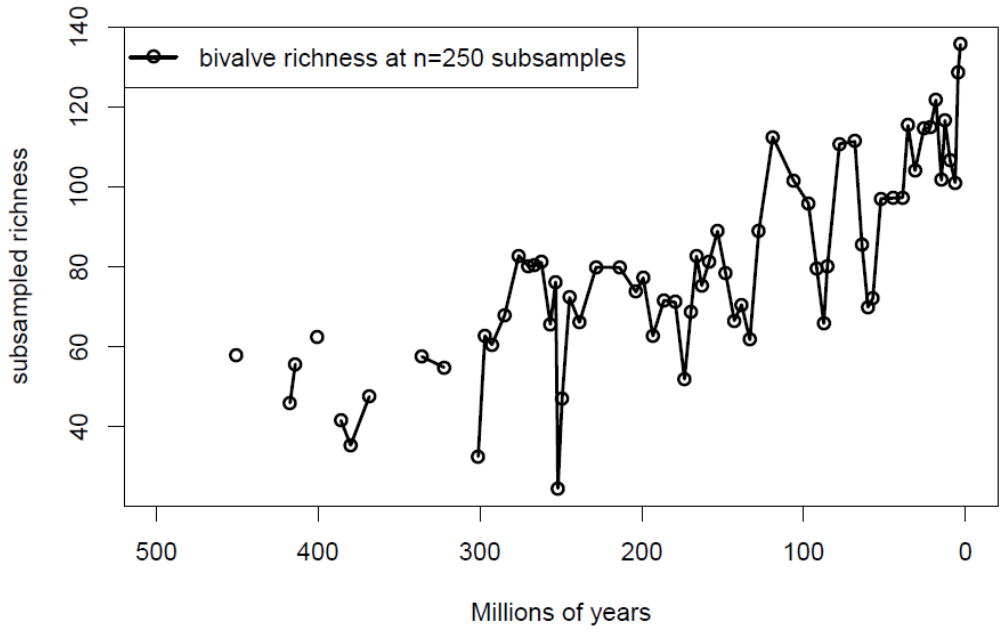
Question 2: Identify a time of high completeness and a time of low completeness in the bivalve fossil record. What does the SCM plot tell you about the diversity plot from section (1)? (5 marks)

High completeness: any time bin of SCM >75% e.g. Pliocene

Low completeness: any time bin of SCM < 50% e.g. Pennsylvanian (Late Carboniferous)

Where we see a large discrepancy between sampled and ranged-through diversity, we can match this to a low SCM value. This is because we see lots of gaps during these time periods and this indicates a poor fossil record. Many of the dips in sampled richness are likely caused by low levels of sampling.

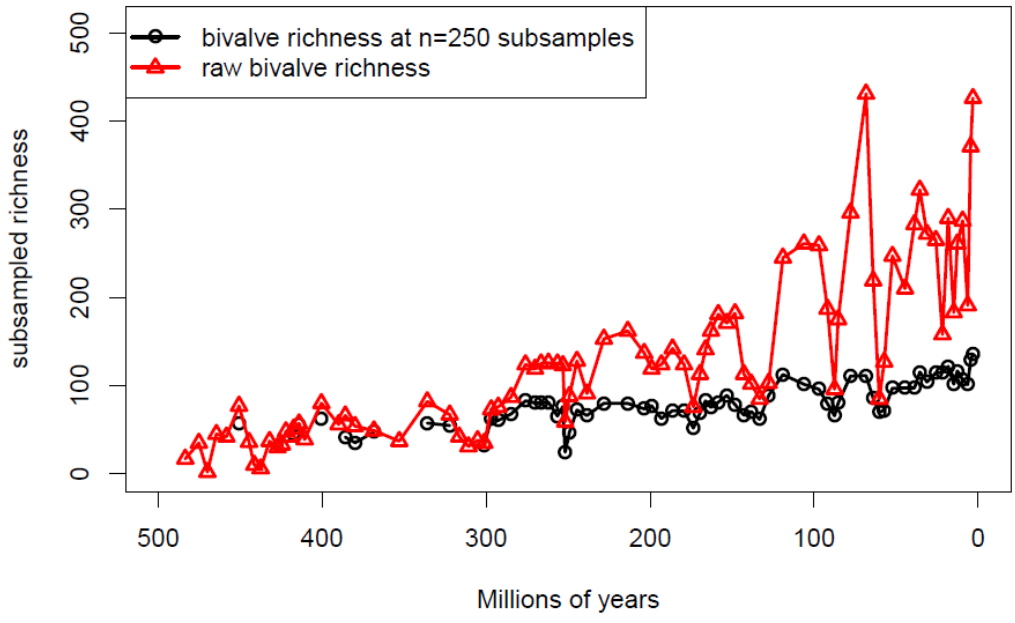
(3) Run the code to plot subsampled (rarefied to 250 samples per time bin) bivalve richness and open the pdf.



Question 3a: Why are there gaps in the time series during the Palaeozoic? (2 marks)

Because we have insufficient sample sizes in these time bins (e.g. less than 250 samples). These are periods of poor fossil record quality.

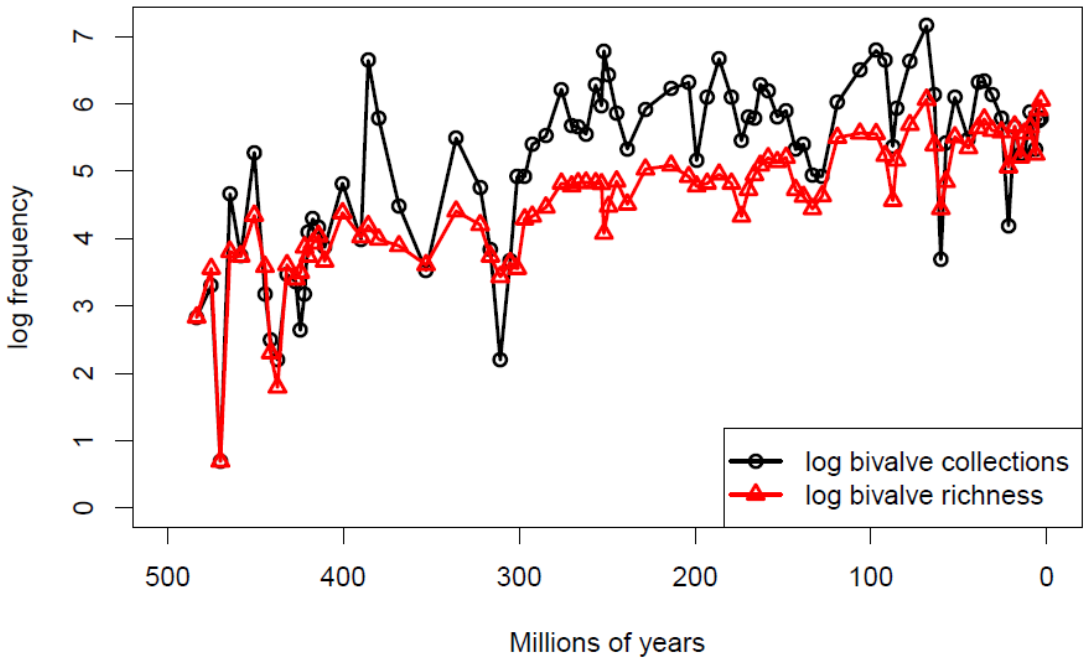
Run the code to plot sampled and subsampled bivalve richness in the same window and open the pdf.



Question 3b: Do the raw and subsampled diversity curves differ? Can you think of any terminology that might describe the effect that the subsampling has removed? (5 marks)

The sampled and subsampled curves follow a similar pattern in terms of peaks and troughs, but the variation in amplitude is much less in the subsampled data. This is expected as subsampling aims to even out sampling across all time bins. Subsampling appears to have removed the rising trend in diversity through time, particularly through the Mesozoic and Cenozoic. It appears to have accounted for the "pull of the recent".

(4) In this data set you are using the number of bivalve bearing collections as a sampling proxy i.e. the more fossils collected = the better the sampling. Run the code to plot the sampled richness and the number of collections through time and open the pdf.



Question 4: Describe the apparent relationship between sampling and richness. Are there any time periods where sampling and richness appear unrelated? (5 marks)

Richness and sampling appear to share a very similar pattern with a number of synchronous peaks and troughs. Sampling and richness appear unrelated around 250 Ma (this is the End Permian mass extinction) when diversity dips but sampling increases and, at around 390 Ma (this is the Late Devonian mass extinction) when diversity declines slightly despite sampling being very high.

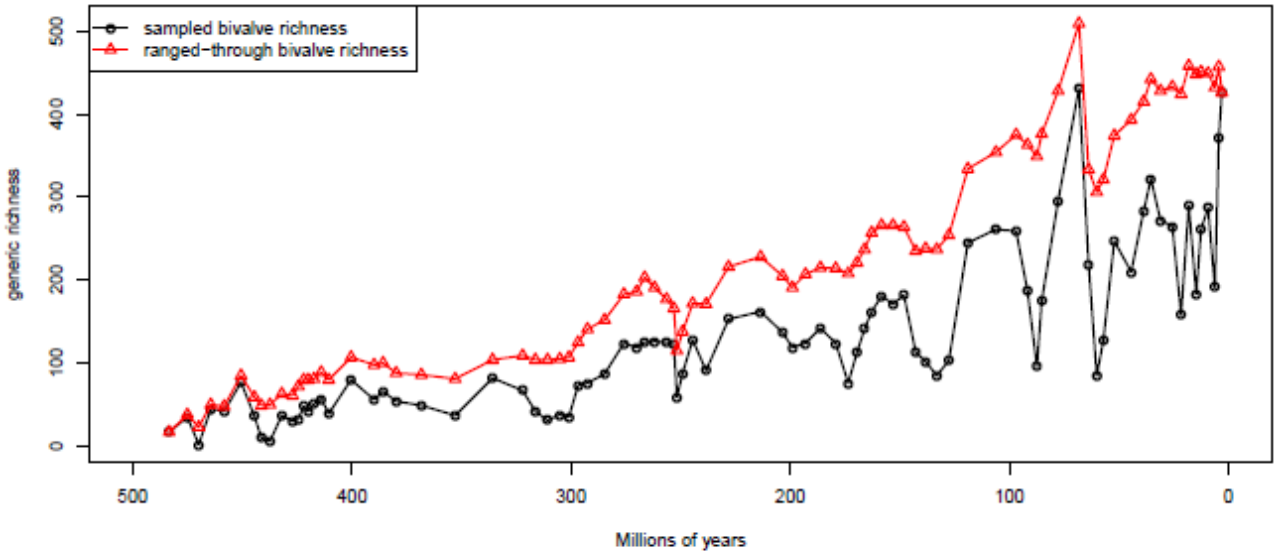
(5) Although it is obvious from the time series that collections and richness share a common signal, it is essential to test this statistically. Run the code to plot a scatter plot and then test for normality in your data. This time, even logging the data does not create a normal distribution. Therefore, you must use a non-parametric Spearman rank correlation test. Execute the Spearman rank code.

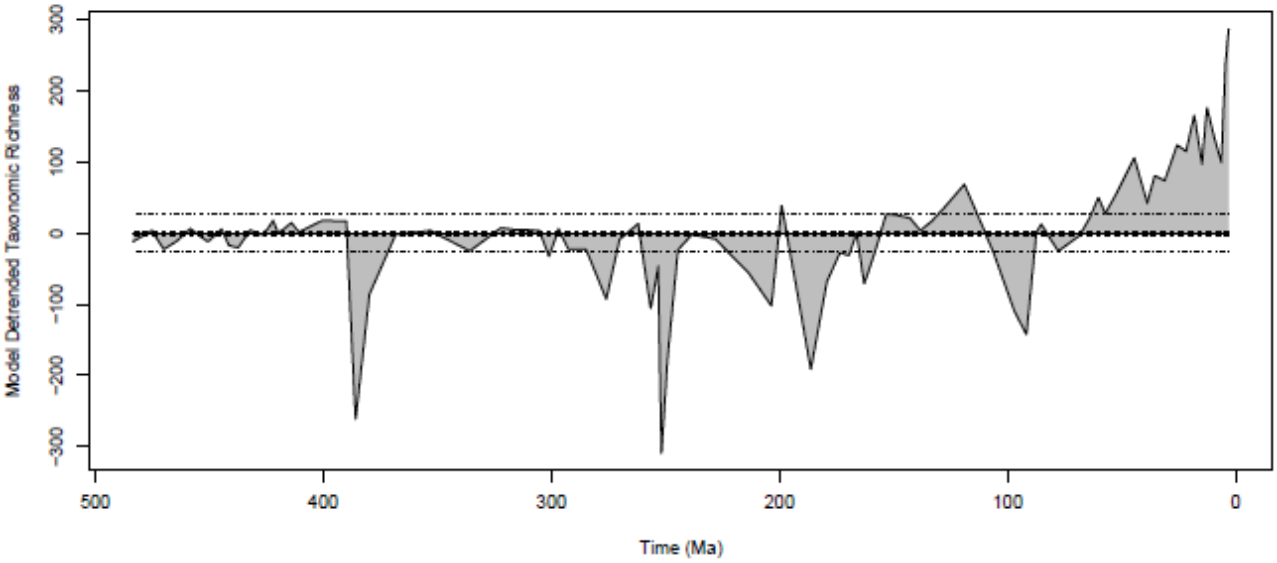
Question 5: Report the correlation test result and explain what this means in terms of the data (5 marks)

Spearman rank: $r_s = 0.76$, $p < 0.001$; could also use rho or ρ as correlation coefficient symbol. Will be lenient with symbols though.

There is a highly significant, strong positive correlation between the number of collections and sampled richness. This suggests that sampled richness estimates are biased by sampling intensity.

(6) As the sampling proxy and richness metrics correlate closely, you can perform residual modelling to extract a diversity signal that cannot be explained by variation in sampling through time. Execute all the modelling and plotting code, including the sourcing of code from Graeme's website and the installation and loading of packages. Open the pdf.





Question 6a: Identify a time period when diversity can be explained by sampling intensity. Identify a time period when diversity cannot be explained by sampling. (4 marks)

Explained by sampling: Ordovician-Silurian-most of Devonian; Carboniferous, Late Jurassic.

Cannot be explained by sampling: Late Devonian; most of the Permian; Early Jurassic; most of the Cretaceous; and all of the Cenozoic.

Question 6b: Can you explain the periodic extreme troughs in residual diversity and the increase in residual diversity to the present day in terms of Sepkoski's empirical model of diversity? What implications does this have for using the fossil record in macroevolutionary studies? (10 marks)

Where we see extreme troughs in residual diversity it means diversity is much lower than we would expect given the level of sampling for that time bin. We see this in the Late Devonian, Late Permian, Late Triassic, Early Jurassic and mid/Late Cretaceous. Three of these correspond to Big 5 mass extinctions (Late Devonian, Late Permian, Late Triassic) and the other two to lesser mass extinctions (Early Jurassic = Pliensbachian-Toarcian, and mid/Late Cretaceous = Cenomanian-Turonian). The Cretaceous-Paleogene and Late Ordovician mass extinctions appear to be predicted by sampling. The increase in residual diversity through the Cenozoic to the present day suggests the rise in diversity in the Cenozoic, as depicted in Sepkoski's model, is a real biological pattern and not entirely caused the "pull of the recent". This shows that some macroevolutionary events depicted in Sepkoski's model are robust to changes in sampling intensity and cannot be explained by sampling. However, it also shows that major evolutionary events can also be masked by variations in sampling intensity and that some apparent events can be "created" by variations in sampling. These interpretations are, however, controversial as there are other lines of evidence to support these mass extinctions that are removed by residual modelling.

End of practical: type `praise()` into the console for some congratulations. Record the praise here:

I'll test for association between R's opinion of you and your practical mark…