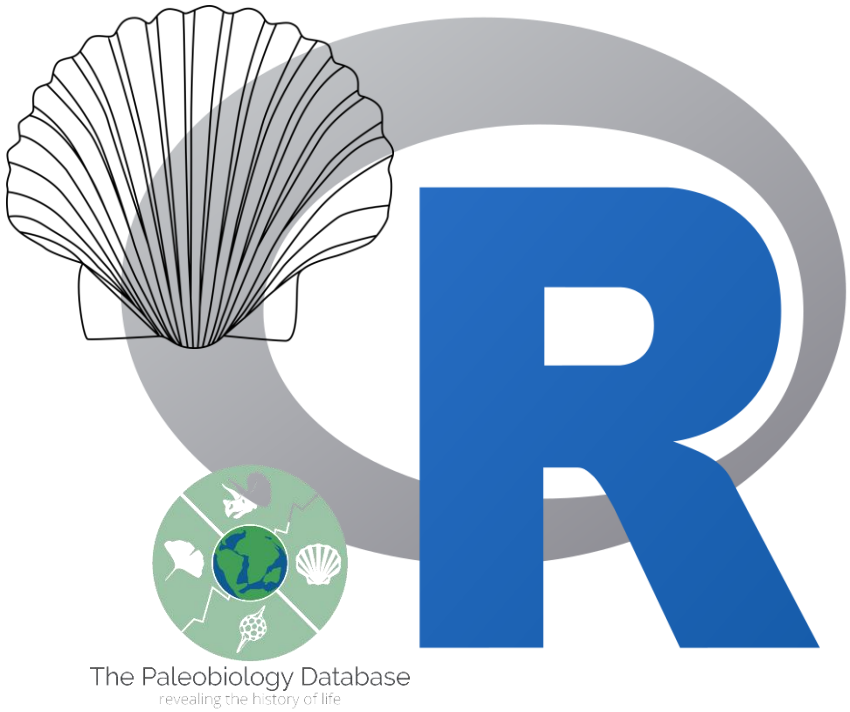
SOEE2145

Palaeoecology, Palaeobiology and Evolution

The quality of the Phanerozoic bivalve fossil record



Dr Alex Dunhill

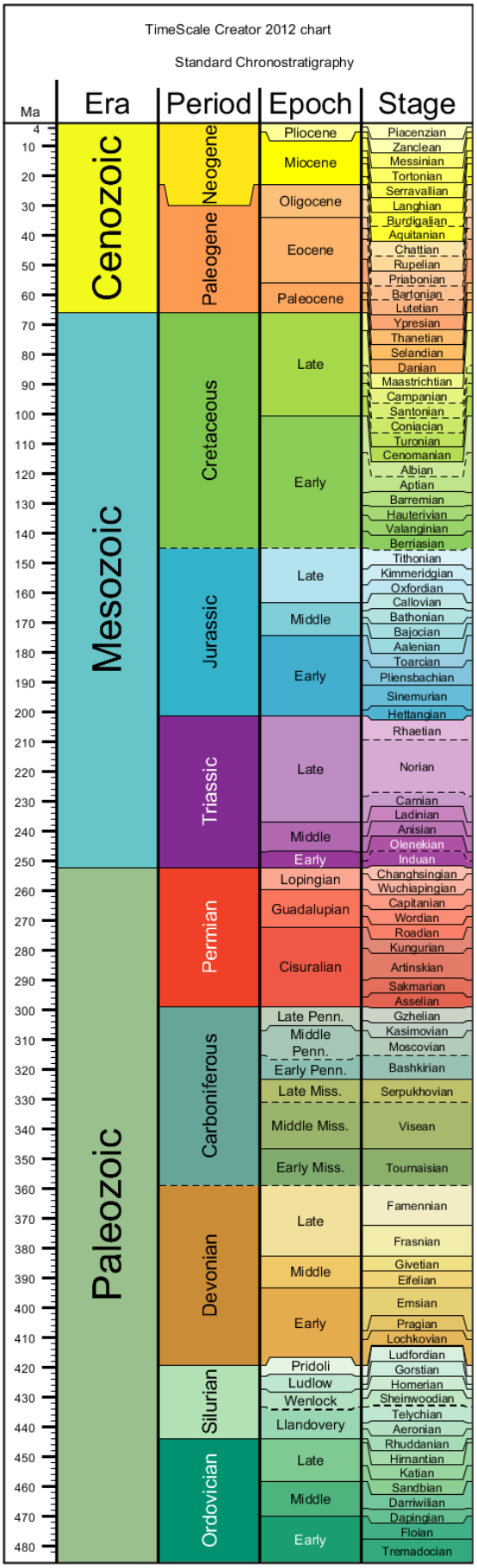
Name:_____	Score: /46 %

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 very similar to that used in the macroevolution practical. If you are unsure of anything, ask me or the demonstrators, we are here to help you. Anything not covered in macroevolution practical will be explained clearly in this walkthrough.



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 (it is very similar to the bivalve data set you used in the macroevolution practical). 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.

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.

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)

(2) Run the function code to calculate bivalve Simple Completeness Metrics (SCM) through the Phanerozoic. We now need to bind this data with the age data from `palaeo` so you can plot it. This is a little tricky in terms of syntax and involves identifying column numbers. Don't worry about this, just make sure you run the code exactly… Run the code to export a plot of SCM through time.

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)

(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)

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)

_____

_____

_____

_____

_____

_____

(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)

_____

_____

_____

_____

_____

_____

(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)

_____

_____

_____

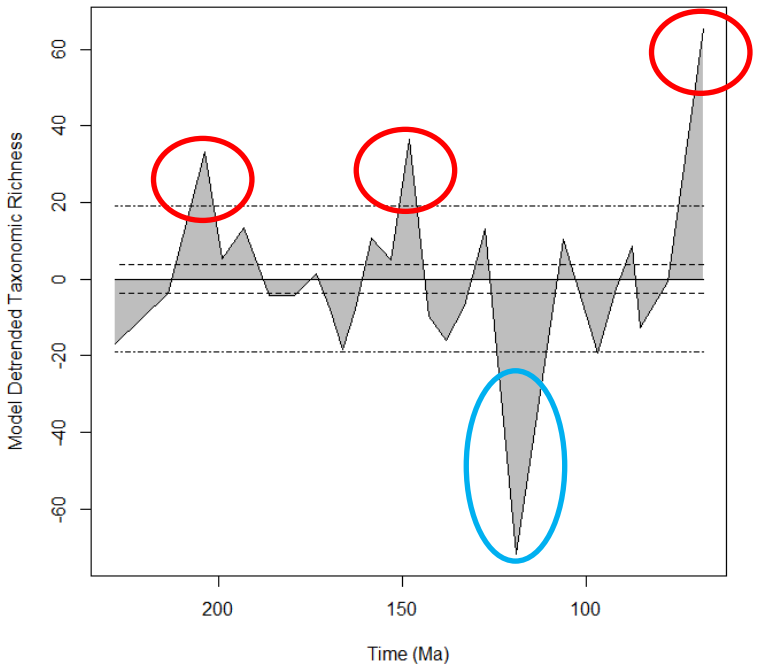
_____

_____

(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. Simply put, we can calculate the relationship between sampling and diversity using a regression analysis, then we can subtract the sampling model from the raw diversity to retain the residuals, which represent times when diversity is higher or lower than we expect given the level of sampling. In theory, it leaves us with a pattern of diversity that cannot be explained by variability in sampling. The code for this model is very complex, so we will source it from the website of Dr Graeme Lloyd (who is a member of the Palaeo@Leeds group.)

Execute all the modelling code which includes of code from Graeme's website and the installation and loading of packages as well as renaming variables to fit the modelling code and then running the model itself. Then run the code to plot the model output alongside the raw richness plot from (1) as a pdf. Ignore the warning messages. Open the pdf.

Your plot will look a little like this (although this is from a completely different data source). The dotted lines represent confidence intervals. Any portion of the data within these lines (i.e. close to 0) does not deviate significantly from the sampling-predicted model of diversity. Put simply, this means diversity during these time periods is likely to be controlled by sampling. Any peaks in diversity outside of the dotted lines represent times where diversity cannot be explained by sampling. For example, the peaks circled in red are times when diversity is higher than expected given the level of sampling and, the trough circled in blue is a time when diversity is lower than expected given the level of sampling.



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)

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)

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…