

Data collection and analysis of paleonetworks

Derek Corcoran

August 6, 2018

Data collection

the first step was to collect data from the paleodiversity database, from there we gathered occurrences using the following code. Which collected all occurrences from the Rhaetian to the Sinemurian, within the bounding box of latitudes 34 to 85 degrees north and longitudes 28.85 west to 41.078 east.

```
All <- read_csv("https://paleobiodb.org/data1.2/occs/list.csv?datainfo&rowcount&interval=rhaetian,sinemurian",
  skip = 22)
```

Data cleaning

Environment

we filtered the data to include only occurrences that occurred in the environments

Taxonomy

Occurrences were excluded if they were part of the Phyla Problematica, Ginkgophyta, Cycadeoideophyta, Cycadophyta, Equisetophyta, Pinophyta, Peltaspermatophyta, Tracheophyta or if they belonged to the class Insecta or Mammalia, or belonging to the orders Pterosauria.

Data exploration

The final dataset had 7,655 occurrences, belonging to 12 Phyla (Table 1) and 1,384 species

Table 1: Number of occurrences for each Phylum present in descending order

phylum	n
Mollusca	3,726
Brachiopoda	1,643
Chordata	972
Cnidaria	500
Arthropoda	463
Echinodermata	177
Porifera	75
Chlorophyta	51
Annelida	30
Cyanobacteria	11
Bryozoa	6
Nematoda	1