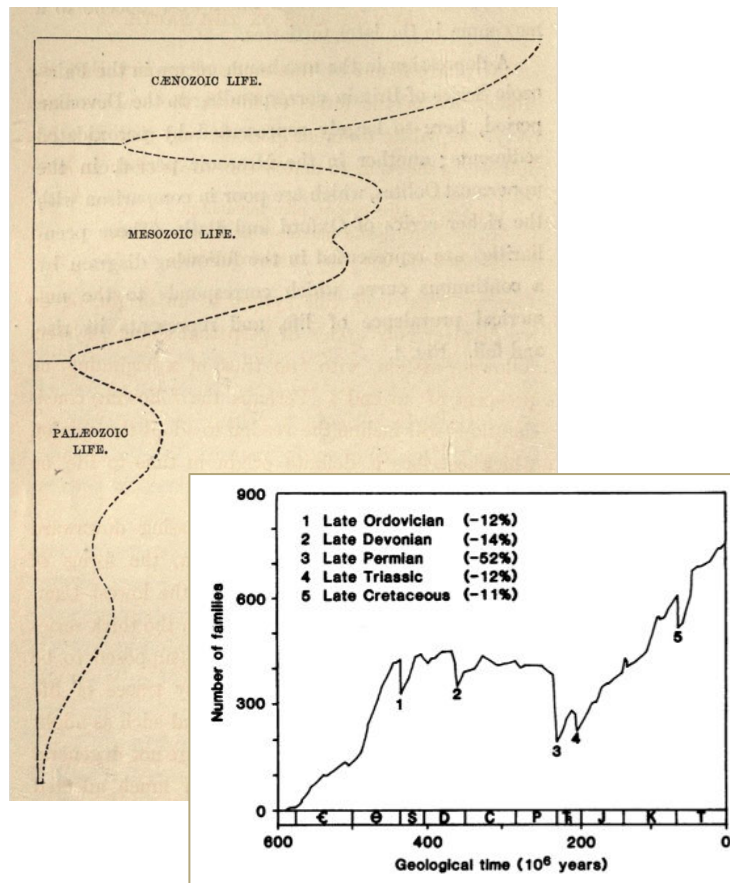

Fossil record data structure

— Emma Dunne | APW 2024 | Wed. Aug.7th —

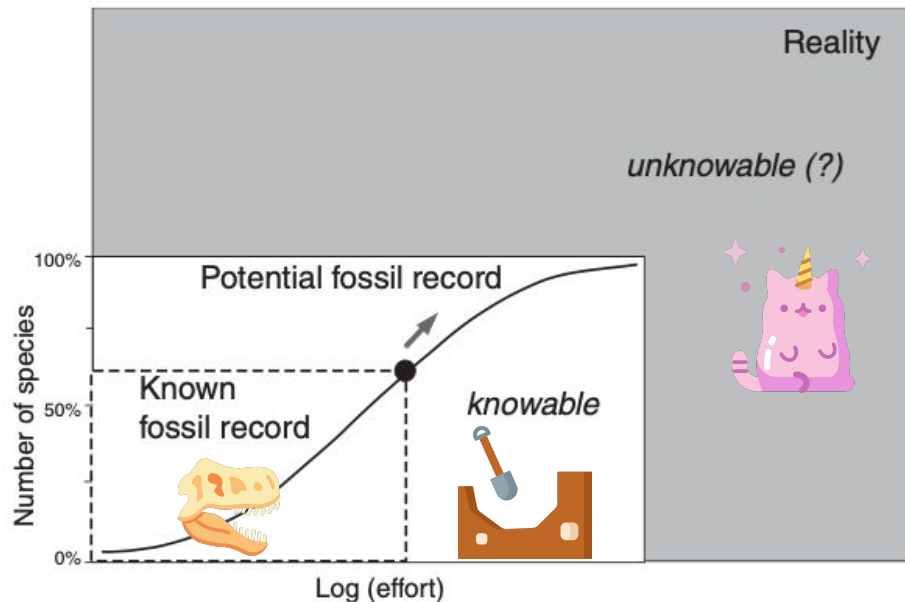
Early paleodiversity studies

- Early palaeodiversity studies took the fossil record at **face-value**
 - Counts of taxa in bins
- But sampling of the fossil record is **uneven and incomplete**
 - Even Darwin noted this
- It wasn't until the last half century that we started to appreciate the impacts of fossil record biases...



How much of the fossil record do we even know?

- The known fossil record is barely a **fraction** of what actually exists
- Even the potential fossil record contains only a tiny fraction of life that has ever lived!



From death to database

Habitat & mode of life



Depositional environment



Geological activity



Geographical location



Interest & resources

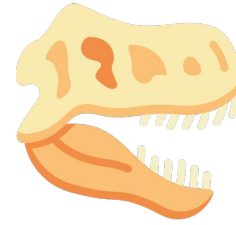
Occurrence data



- Several various factors act on the fossil record at various stages to remove data
- This begins even before an organism dies!
- The result is a skewed view of past life and biodiversity

Raup's "7 Sources of Error" (1972)

1. Range charts
2. The 'Pull of the Recent'
3. Durations of geological units
4. Monographic effects
5. Lagerstätten
6. Area-diversity relationships
7. Sediment volume



From death to database

Habitat & mode of life



Depositional environment



Geological activity



Geographical location



Interest & resources

Occurrence data



PROCESSES

Death & burial

Diagenesis & lithification, tectonics, etc.

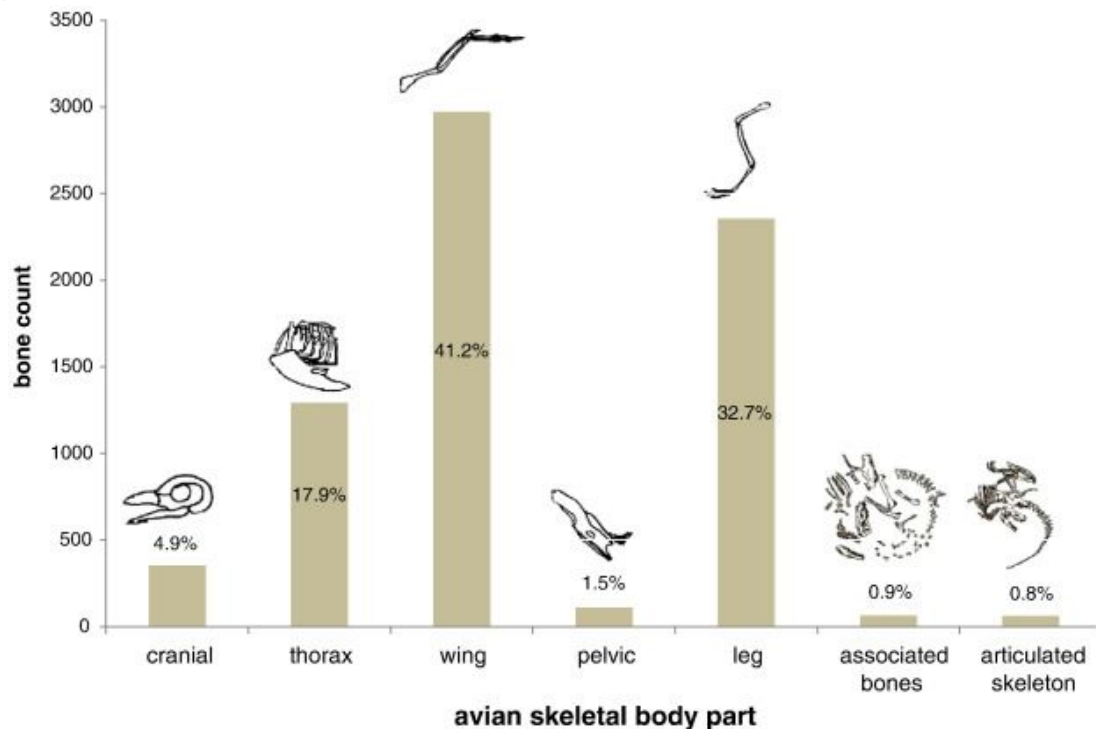
CATEGORY

Taphonomic processes

Taphonomy & anatomy

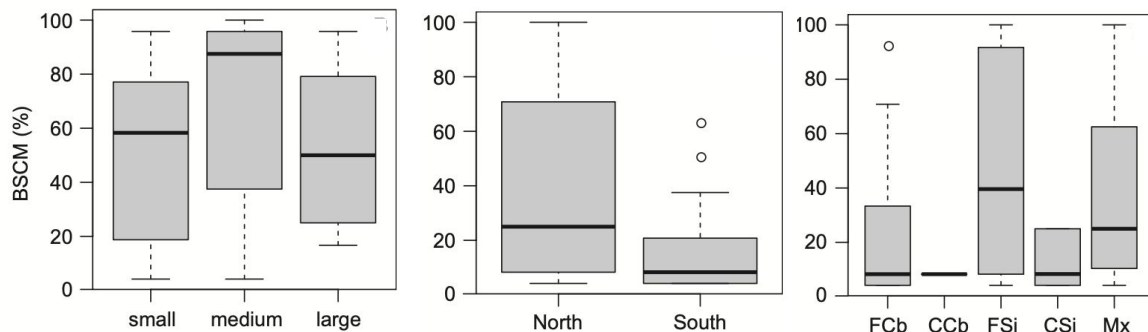
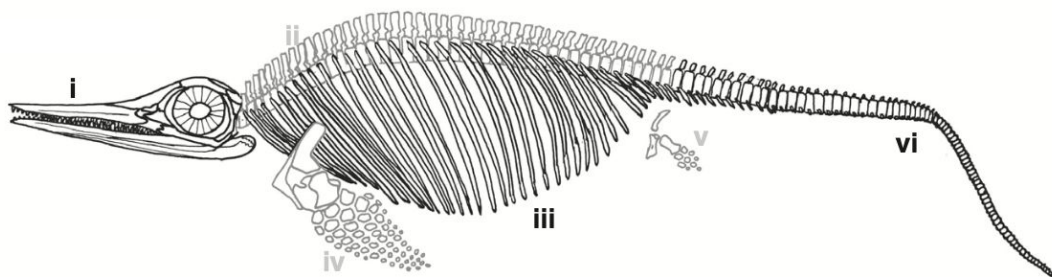
Taphonomy & anatomy

- Different body parts are more or less likely to be **preserved** or **identified**
- How many animals are known only from teeth?
- **Body size** is also a major control on entering the fossil record and being discovered

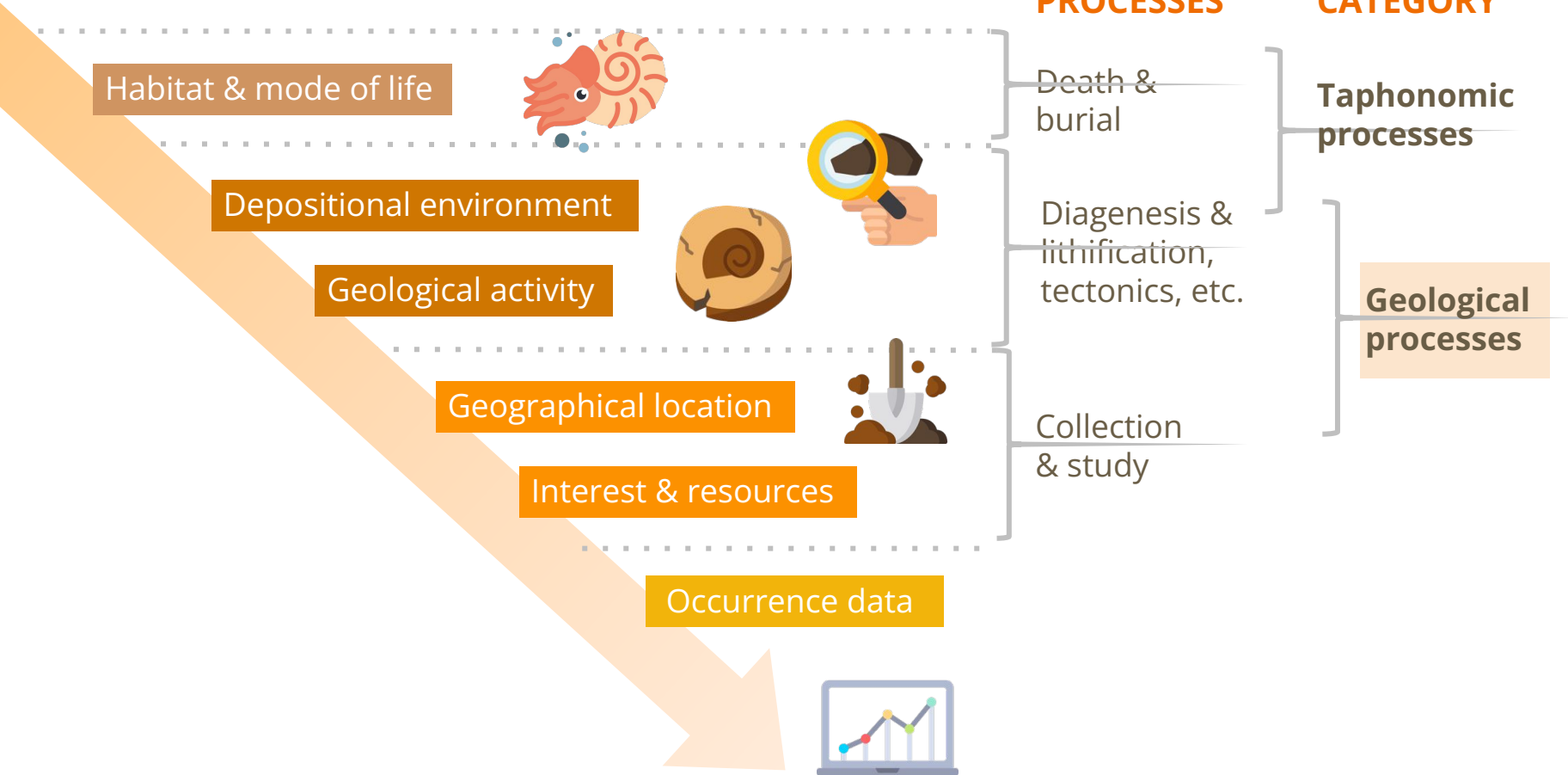


Taphonomy & anatomy

- Incomplete or damaged specimens are hard to identify
- Completeness varies by size, region, and lithology
- Diversity can be under-/over-estimated
- **Lagerstätten effect**



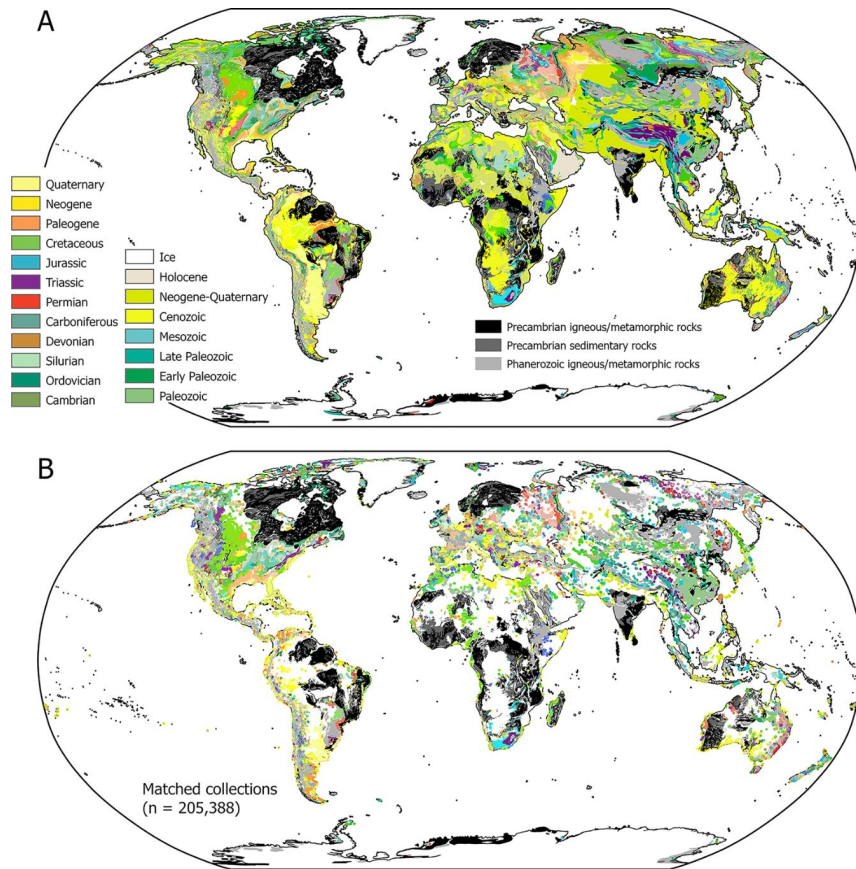
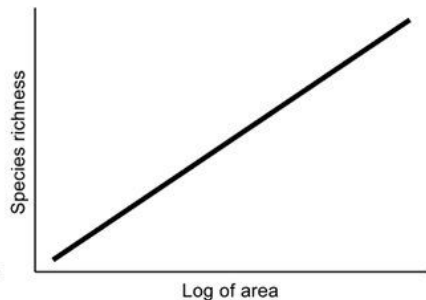
From death to database



Outcrop area

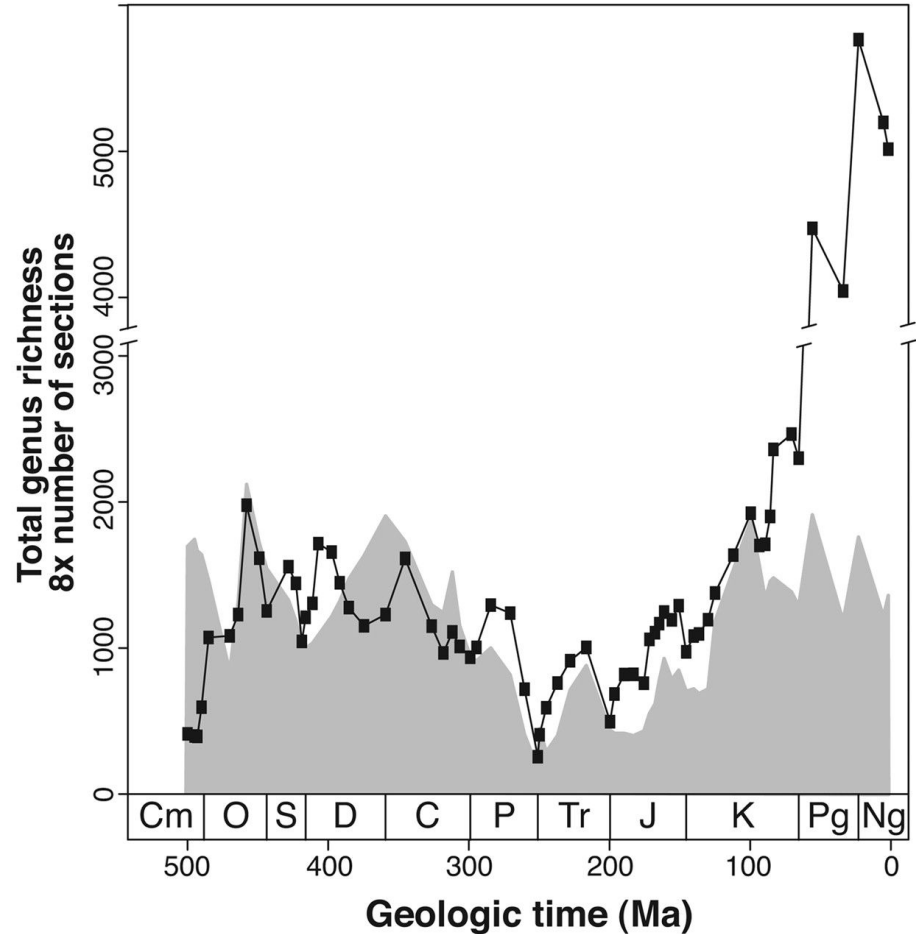
Outcrop area

- The **'bias hypothesis'**: sampling is the primary driver of observed diversity
- = the more rock area sampled, the higher the diversity
- **Species-area effect**



'Common cause'

- Certain drivers can influence **both** proxies for structure and richness, generating an apparent relationship
- For example, sea level change can drive both rock deposition and diversification in shallow marine animals



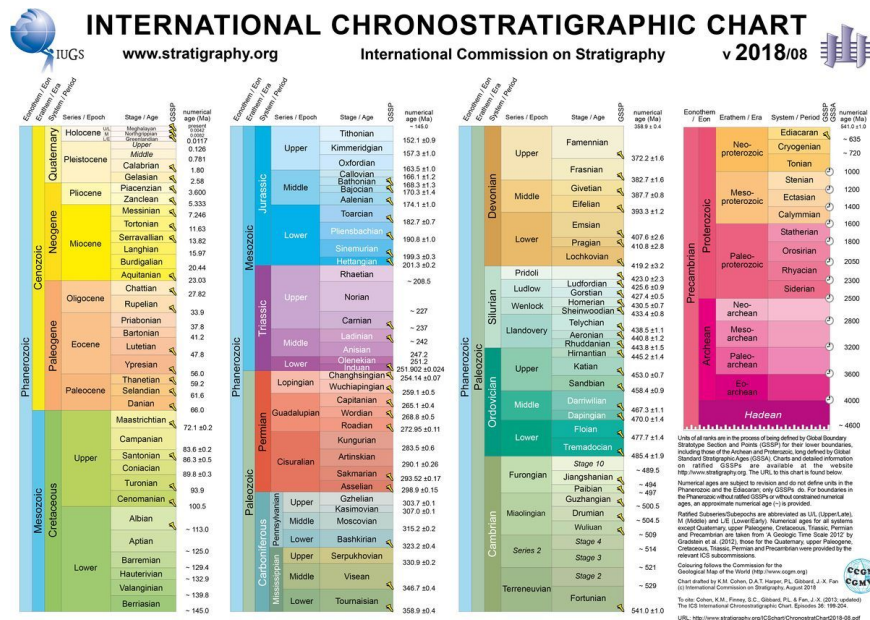
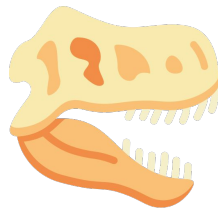
Temporal resolution of the fossil record

Temporal resolution of the fossil record

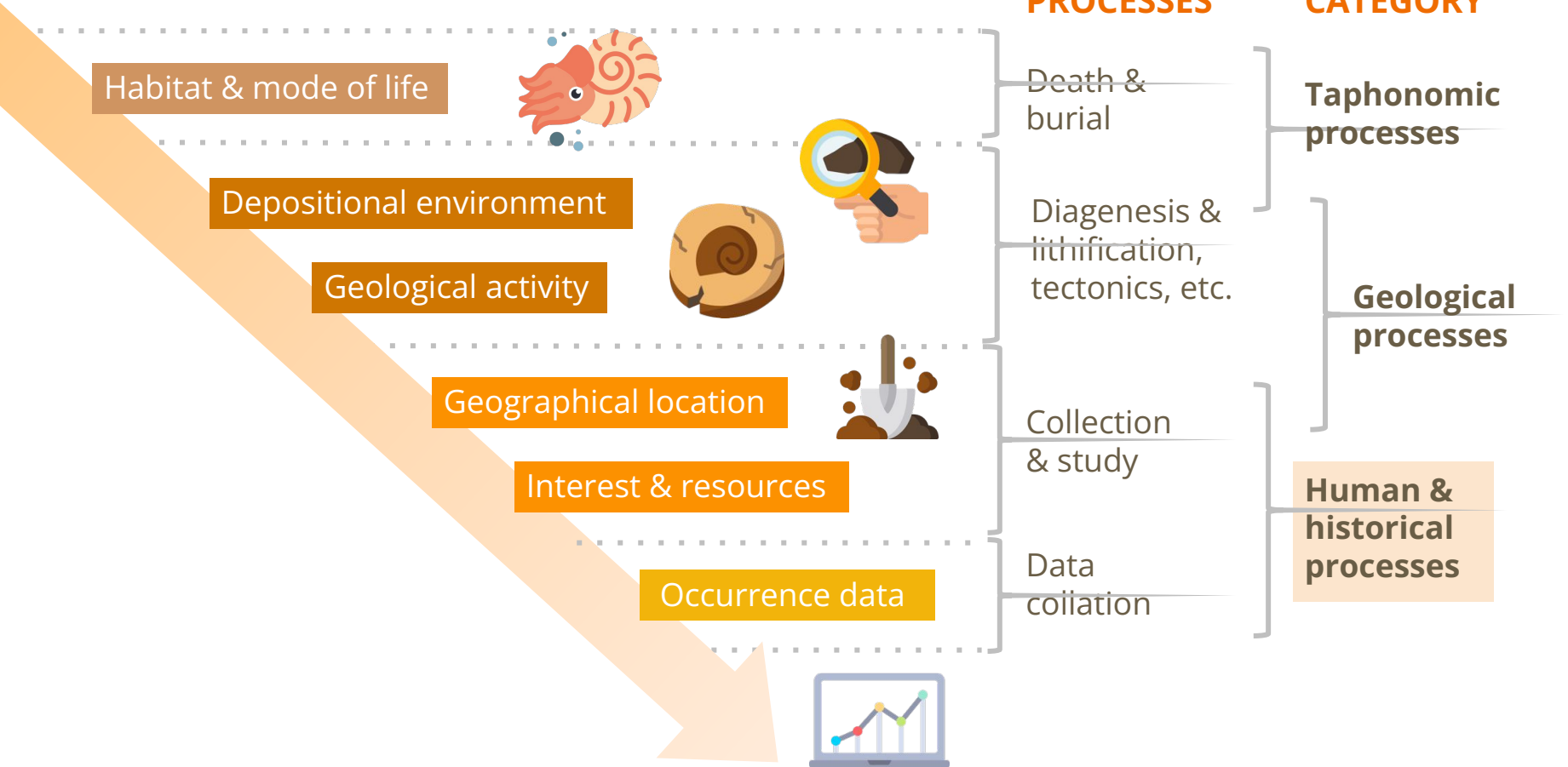
Geological time intervals are not equal in length

Example: Late Triassic epochs:

- Rhaetian ~8 Ma
- Norian ~20 Ma
- Carnian ~10 Ma

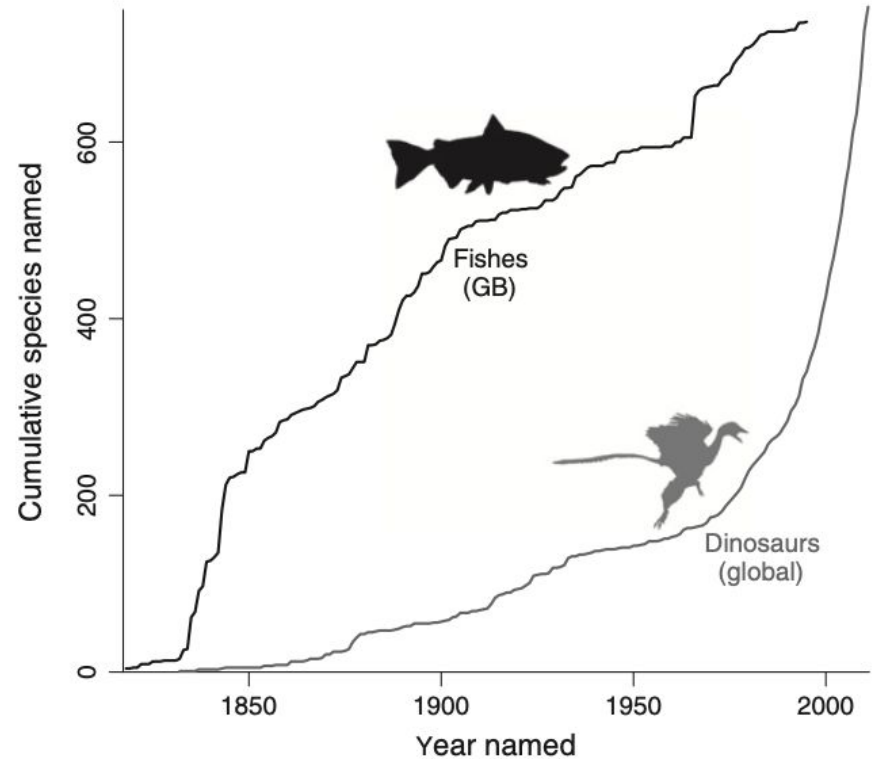
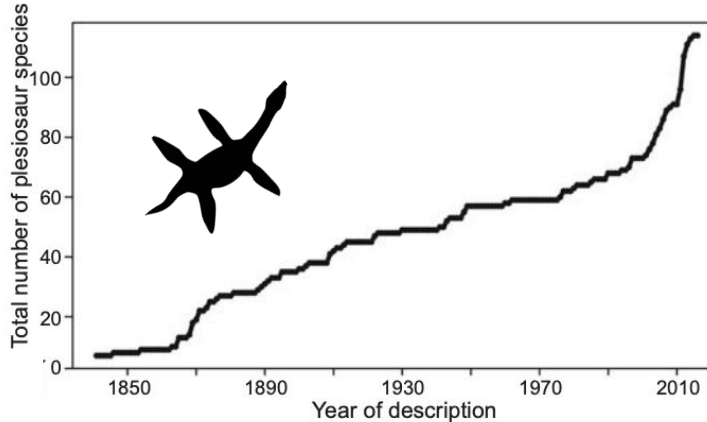


From death to database



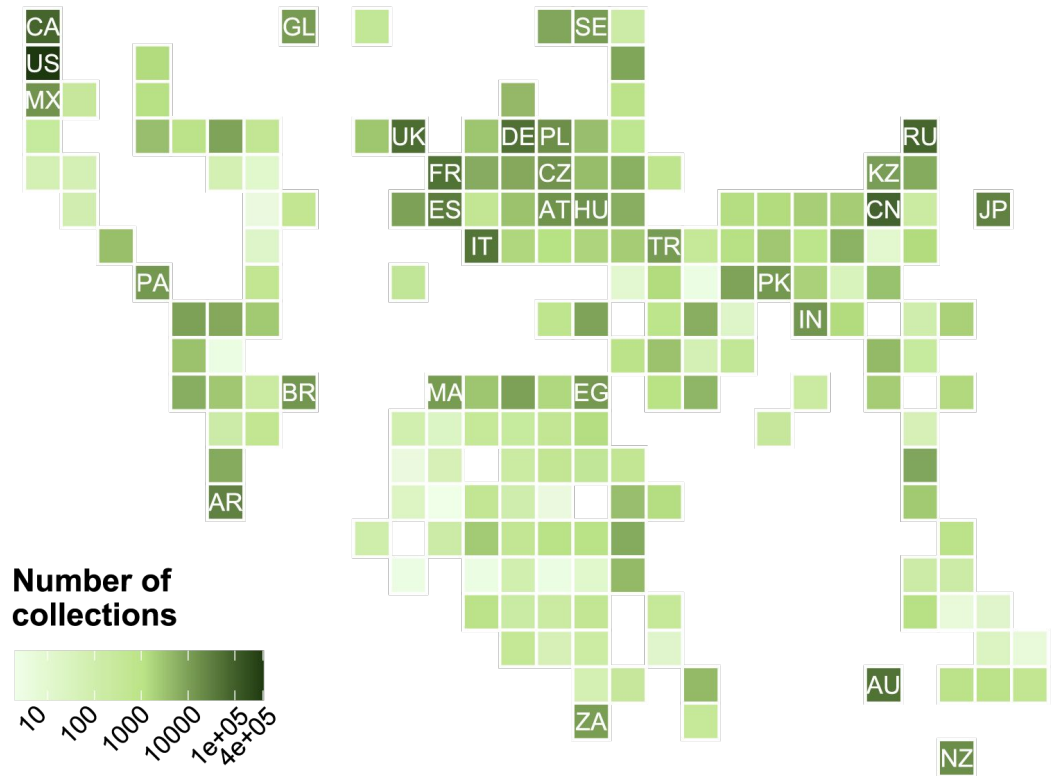
Research interest

- More research activity (over long periods of time) contributes to better sampling



Socio-economics

- 97% of fossil occurrence data in the PBDB were generated by researchers in North America and western Europe
- Countries with a history of colonialism have a greater paleo research output
- Sampling is biased by where the resources are



Languages

Languages

- English is the dominant language in palaeodiversity studies
- Knowledge in other languages is overlooked – this has been shown to bias outcomes of meta- analyses ([Konno et al. 2020](#))
- Impedes the accessibility and communication of science
 - e.g. Literature for fossil occurrences

