Modeling changes to the functional composition of North American mammal diveristy

multi-level dynamics of a regional species pool

Peter D Smits

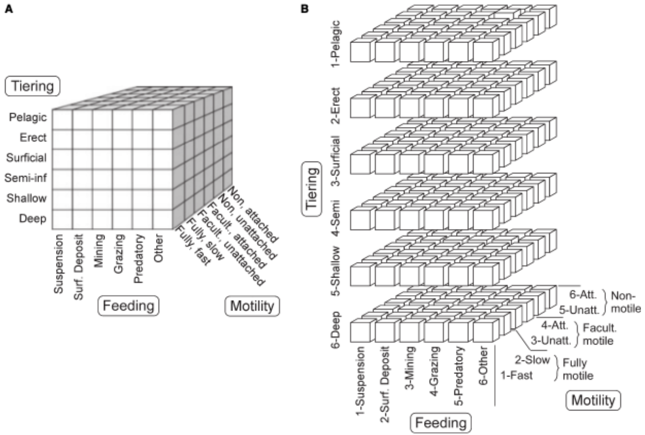
Department of Integrative Biology, University of California – Berkeley

Question

Why do the relative diversities of functional groups change within a species pool?

- function of species traits and environmental context

Eco-cube and functional groups



TEXT-FIG. 1. Ecospace as defined by the three axes of tiering, motility level and feeding strategy. A, the ecospace cube with categories on each axis labelled. B, the ecospace cube 'exploded', showing 216 'bins' or modes of life specified by the combination of the categories on each ecospace axis.

(Bambach *et al., 2007,* Palaeontology)

Species pool concept

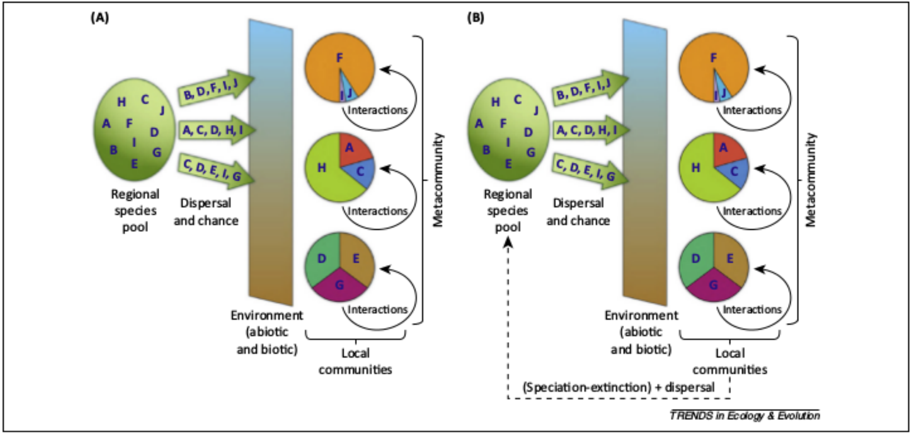


Figure 1. Two models of community assembly. **(A)** Local communities comprise a subset of species from the regional species pool that have passed through environmental filters. There is no feedback from the metacommunity (collection of local communities) to the regional species pool. Adapted from [5]. **(B)** Local communities are assembled as in (A), but speciation adds new species to the pool, extinction removes others, and dispersal allows the persistence of species that might otherwise go extinct.

(Mittelbach and Schemske, 2015, *TREE*)

Structured, multi-level data in biology

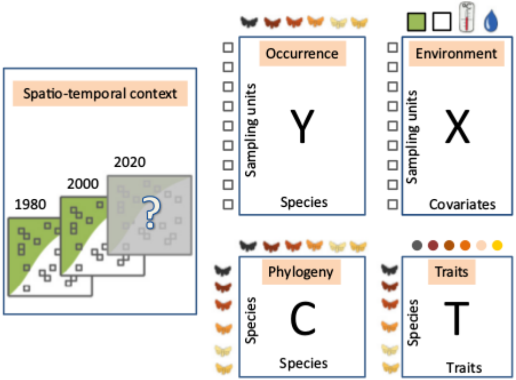
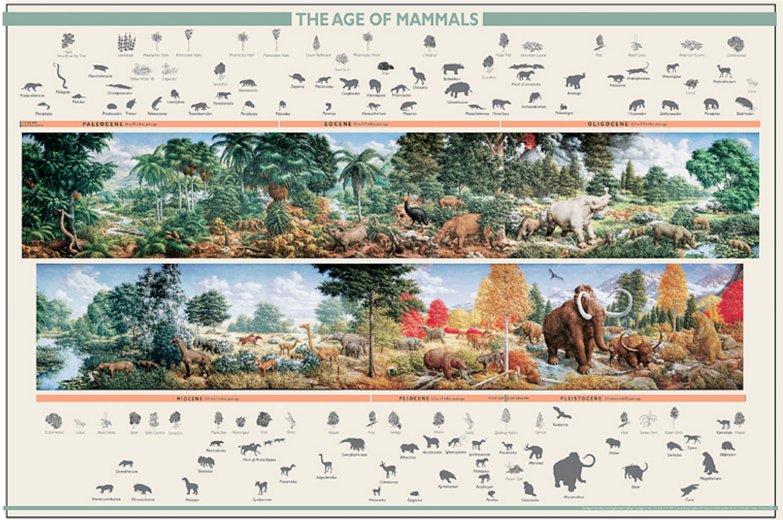


Figure 3 Data typically collected in community ecology. The occurrence data (denoted as the **Y** matrix) includes the occurrences of the species recorded in a set of temporal and/or spatial sampling units. The environmental data (denoted as the **X** matrix) consists of the environmental covariates measured over the sampling units. The traits data (denoted as the **T** matrix) consists of a set of traits measured for the species present in the **Y** matrix. To account for the phylogenetic dependencies among the species, we can include a fourth matrix consisting of the phylogenetic correlations among the species (denoted as the **C** matrix). The spatiotemporal context includes location and time information about the samples.

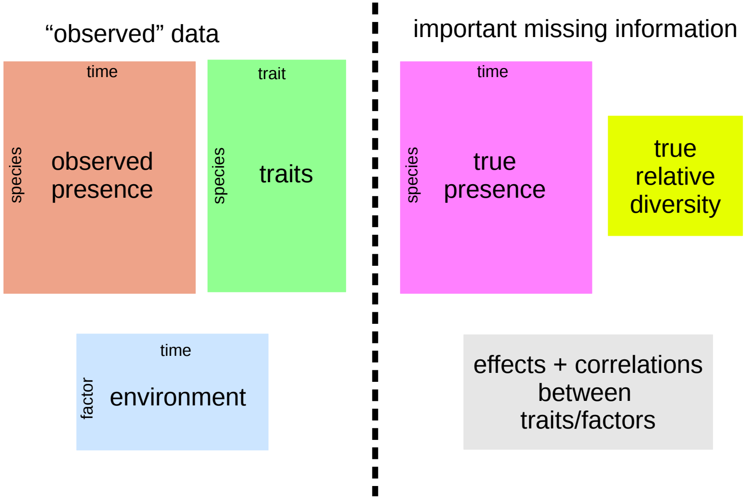
(Ovaskainen *et al.* 2017 *Ecology Letters*)

Cenozoic mammals of North America

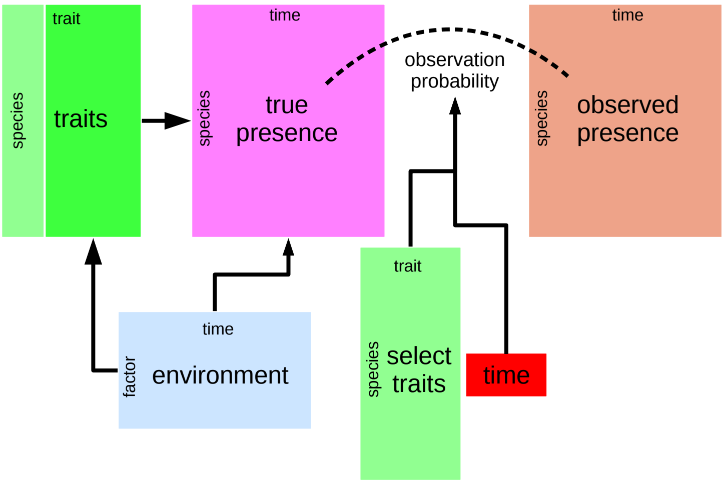


(Rudolph Zallinger)

Conceptualizing the question and data



Conceptualizing the analysis



Hidden Markov Model with absorbing state

Jolly-Seber CMR/Restricted occupancy model

$$
\begin{aligned}
y_{i,t} &\sim \text{Bernoulli}(z_{i,t} p_{i,t}) \\
z_{i,t=1} &\sim \text{Bernoulli}(\phi_{i,t=1}) \\
z_{i,t} &\sim \text{Bernoulli}\left(z_{i,t-1}\pi_{i,t} + \sum_{x=1}^{t}(1 - z_{i,x})\phi_{i,t}\right)
\end{aligned}
$$

y observed state; z estimated state.

p observation; ϕ origination; π survival.

i in N; t in T.

Modeling the probabilities; individual-level

Multi-level logistic regression

$$p_{i,t} \sim \text{logit}^{-1}(b_t + e_{j[i]} + \beta^p mass_i)$$

$$\phi_{i,t} \sim \text{logit}^{-1}(f^{\phi}_{j[i],t} + o^{\phi}_{k[i]} + \beta^{\phi} mass_i)$$

$$\pi_{i,t} \sim \text{logit}^{-1}(f^{\pi}_{j[i],t} + o^{\pi}_{k[i]} + \beta^{\pi} mass_i)$$

observation: b_t time-varying intercept; $e_{j[i]}$ functional group eff; β^p mass eff.

origination: $f^{\phi}_{j[i],t}$ time/FG-varying intercept; $o^{\phi}_{j[i]}$ order eff; β^{ϕ} mass eff.

survival: $f^{\pi}_{j[i],t}$ time/FG-varying intercept; $o^{\pi}_{j[i]}$ order eff; β^{π} mass eff.

Modeling the probabilities; group-level

Multivariate regression of time/FG-varying intercept

$$f^{\phi} \sim \text{MVN}\left(\begin{matrix} U\gamma^{\phi}_{j=1} \\ \vdots \\ U\gamma^{\phi}_{j=J} \end{matrix}, \text{diag}(\tau_{f\phi})\Omega_{f\phi}\text{diag}(\tau_{f\phi})\right)$$

$$f^{\pi} \sim \text{MVN}\left(\begin{matrix} U\gamma^{\pi}_{j=1} \\ \vdots \\ U\gamma^{\pi}_{j=J} \end{matrix}, \text{diag}(\tau_{f\pi})\Omega_{f\pi}\text{diag}(\tau_{f\pi})\right)$$

U matrix group-level covariates; γ^{ϕ}, γ^{π} vectors group-level reg coefs.

Ω_{ϕ}, Ω_{π} corr matrix of FG by time; τ_{ϕ}, τ^{π} scale of FG by time.

Modeling the probabilities; final details

Comments on priors, implementation

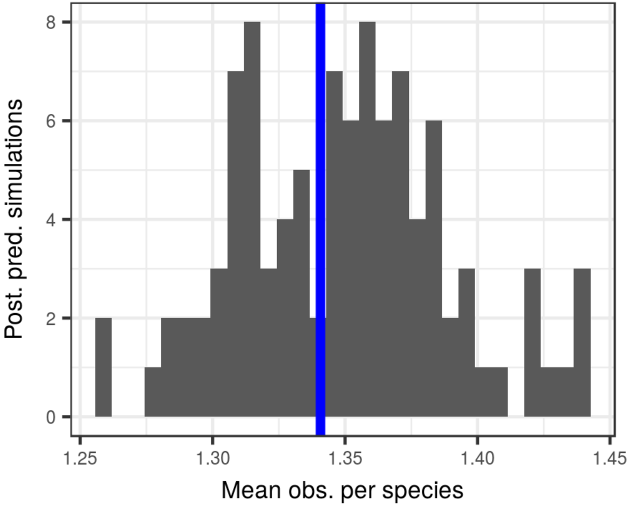
- Random-walk priors for time-varying intercepts
- Regularizing priors predict
  - very weak/no effect of mass e.g. $\mathcal{N}(0, 0.5)$
  - very weak/no effect of group-level covariates e.g. $\mathcal{N}(0, 0.5)$
  - very weak/no correlation b/w functional groups e.g. LKJ(2)
- Marginalization problem b/c gradient based estimation

Parameter estimation and inference

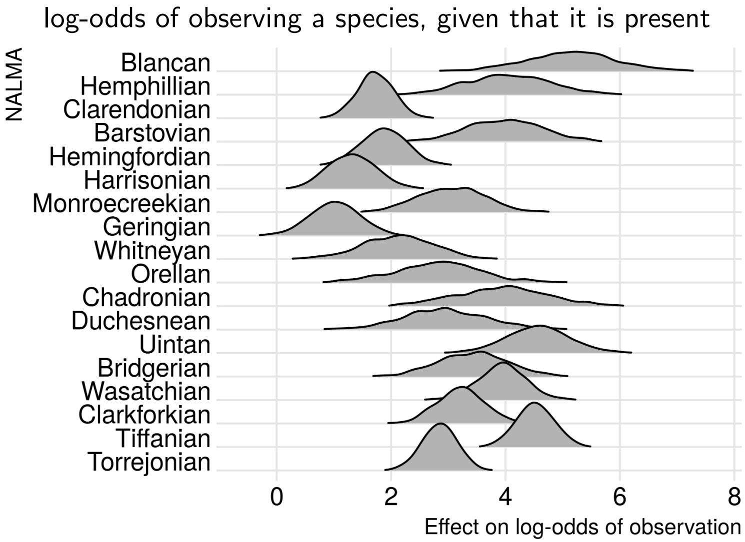
- **Bayesian inference**
  - intuitive and expressive
  - regularization/partial pooling
  - external information
- **Automatic Differentiation Variational Inference (ADVI)**
  - when full HMC/MCMC slow
  - approx Bayesian inference; assumes posterior is Gaussian
  - true Bayesian posterior

Stan

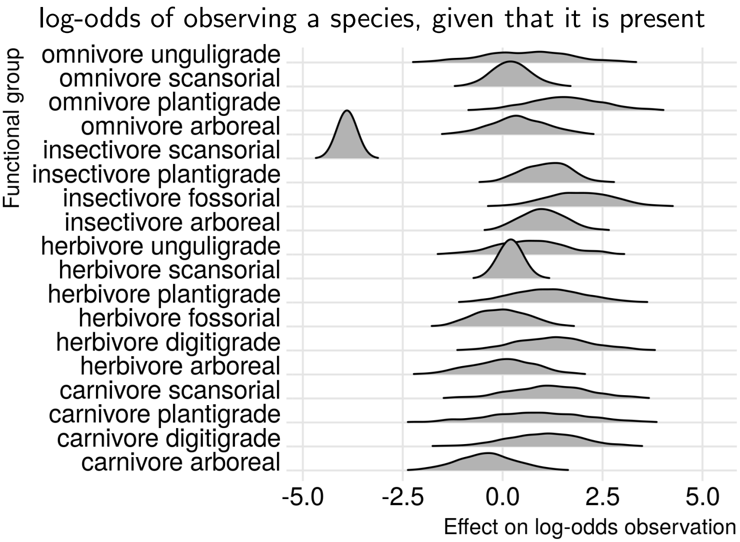
Model adequate? Posterior predictive check



Observation; NALMA

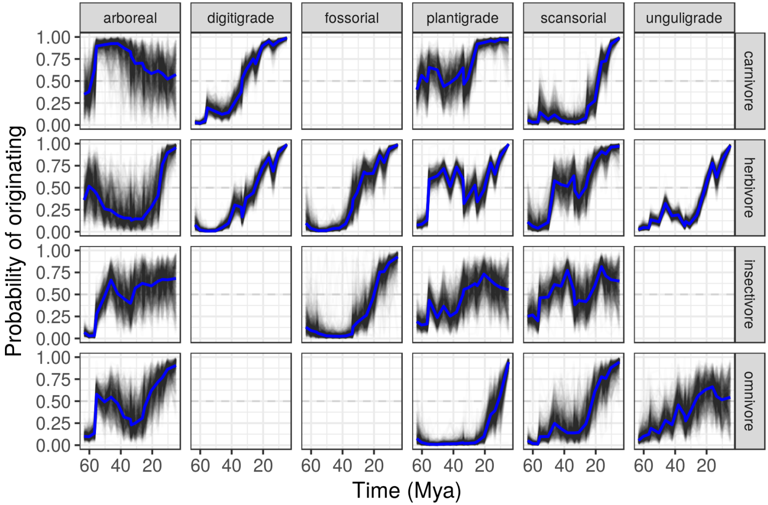


Observation; functional group



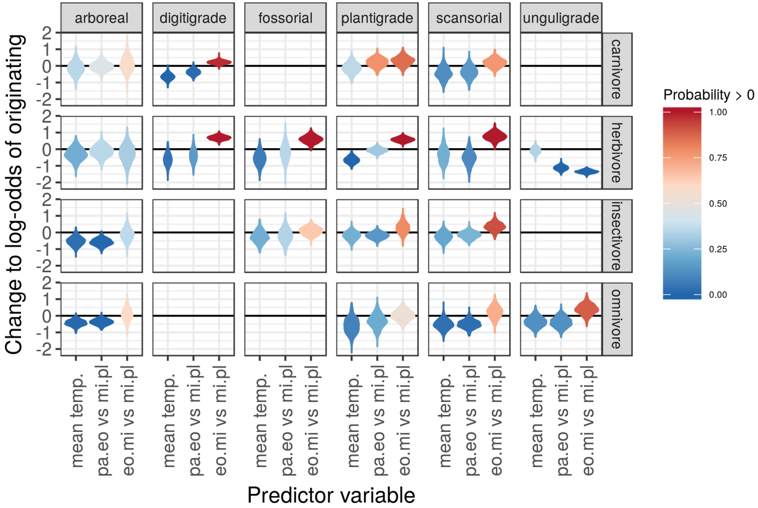
Origination; individual-level

probability of species originating, given it hasn't originated yet

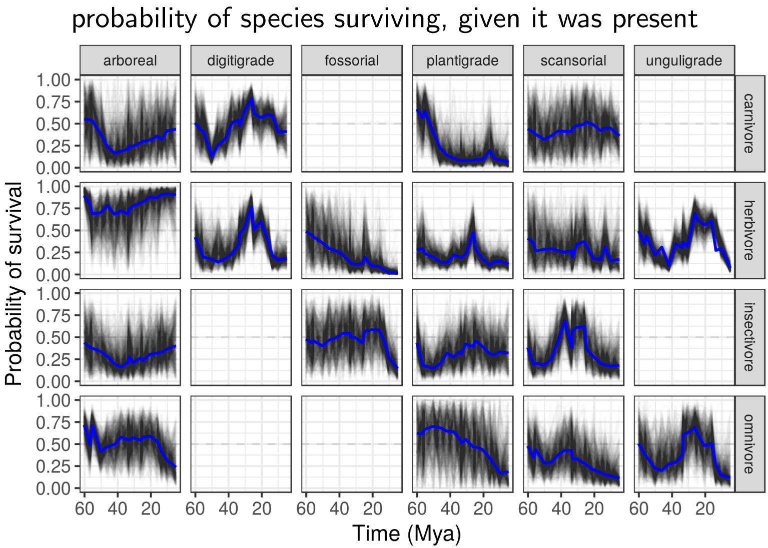


Origination; group-level

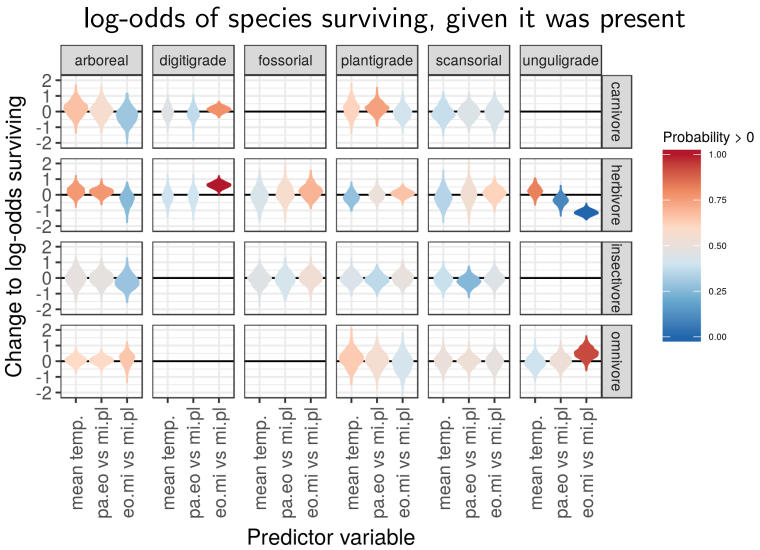
log-odds of species originating, given it hasn't originated yet



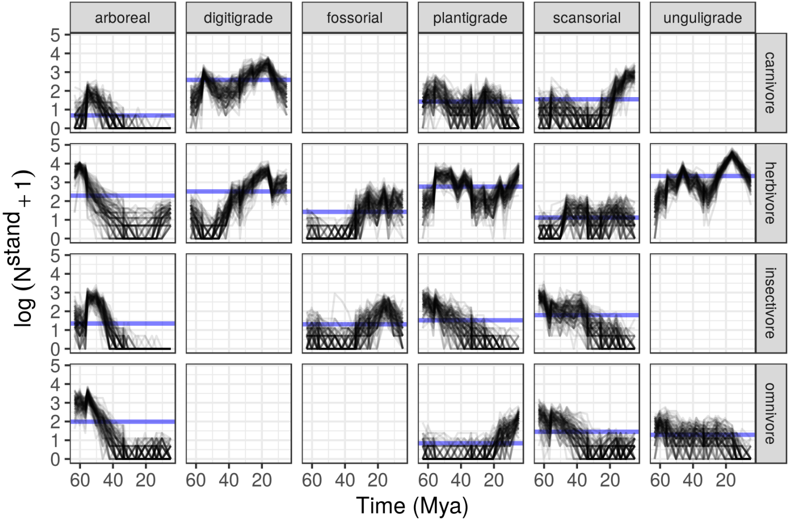
Survival; individual-level



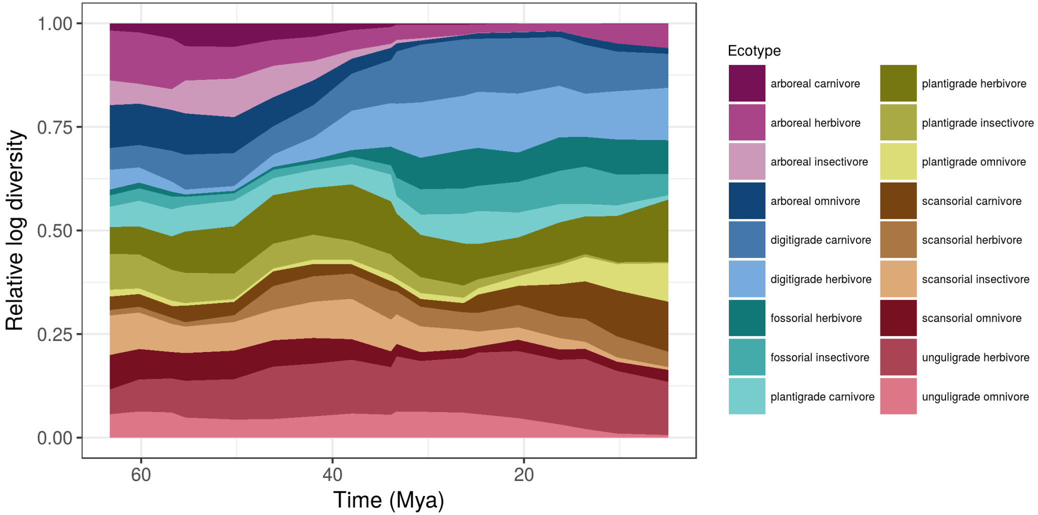
Survival; group-level



Standing diversity of functional groups through time



Relative diversity of functional groups through time



Changes to relative diversity between Neogene/Paleogene

- increase
  - digitigrade, plantigrade, unguligrade herbivores
  - fossorial functional groups
  - plantigrade omnivores
- decrease
  - near total loss of arboreal functional groups
  - plantigrade, scansorial insectivores
  - unguligrade omnivores

Conclusions

- temporal differences in P(observation) much larger than effects of functional group
- increases in P(origination) often met with decreases in P(survival), but not 1-to-1
- environment estimated to effect origination of functional groups more often than survival
- no correlation between functional group origination, survival not accounted for by RW prior
  - does not preclude short similarity, just no long term correlation
  - HMC/MCMC might tweak these results b/c ADVI assumptions

Acknowledgements

- UC Berkeley
  - **Seth Finnegan**, Adiel Klompmaker, Emily Orzechowski, Larry Taylor, Sara Kahanamoku, Josh Zimmt
- UChicago
  - **Kenneth D. Angielczyk**, **Michael J. Foote**, P. David Polly, Richard H. Ree, Graham Slater

@PeterDSmits

The Paleobiology Database
revealing the history of life