The changing functional composition of the North American species pool

modeling species origination-extinction as a function of functional group and environmental context

Peter D Smits

Committee on Evolutionary Biology, University of Chicago

Question

When are certain ecologies, ecotypes, or functional groups enriched or depleted in a species pool?

Age of Mammals

Species pool concept

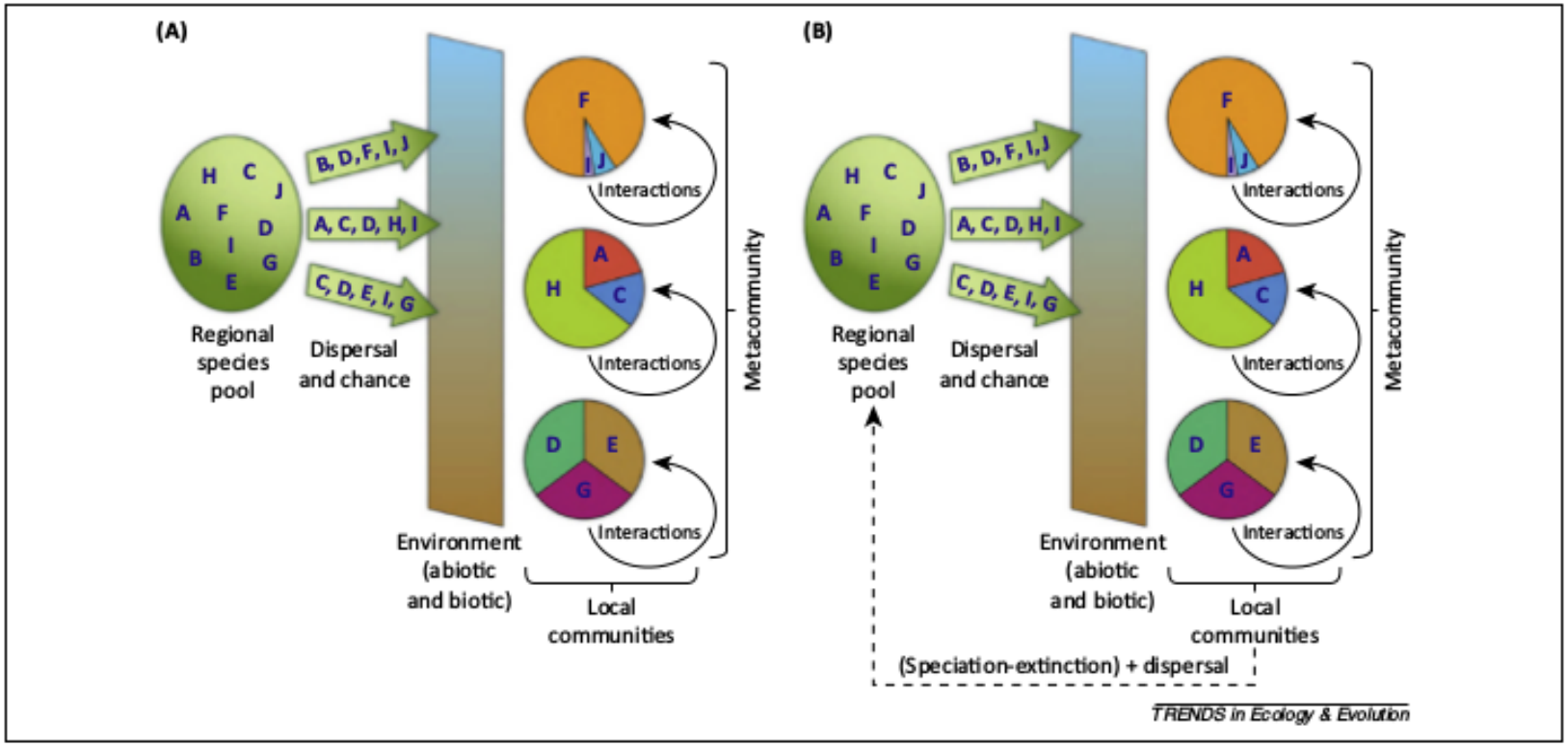
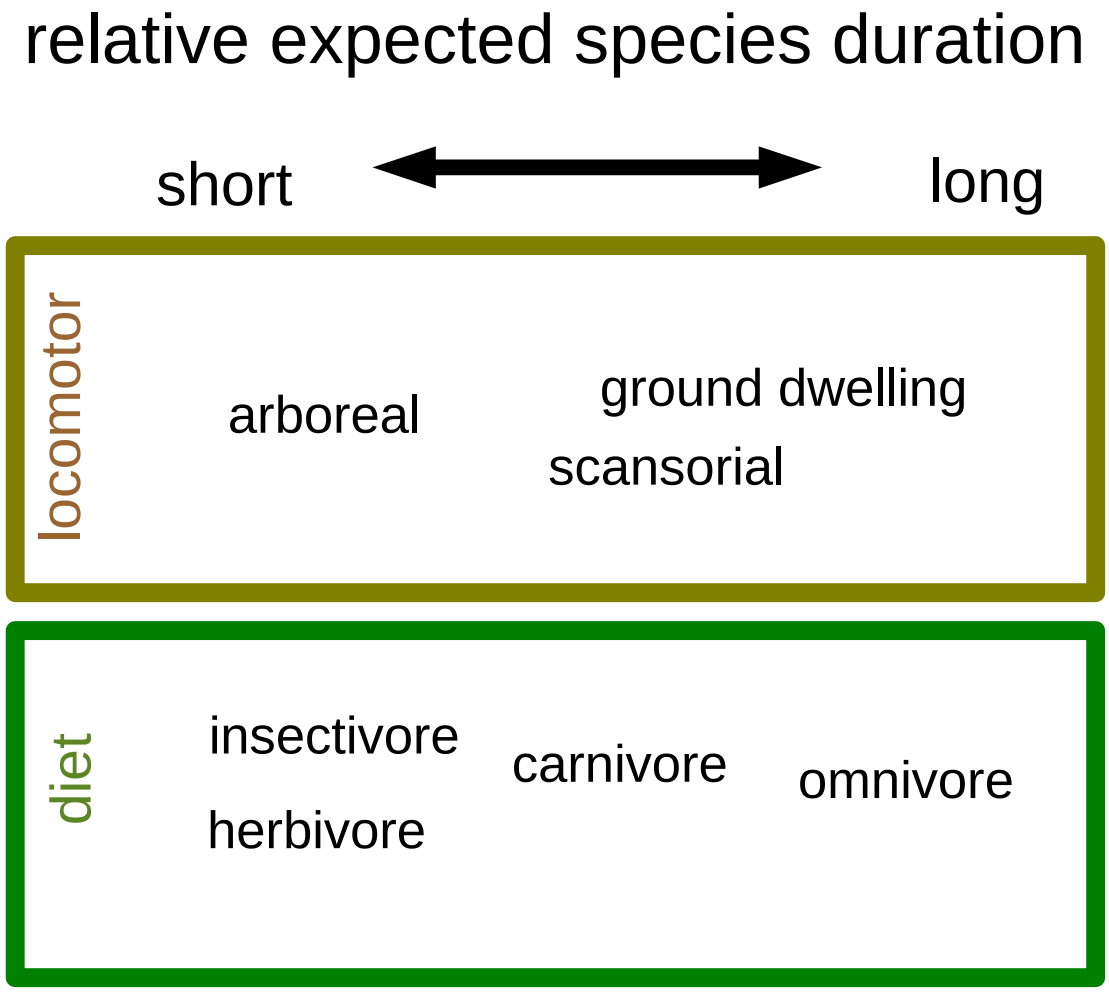


Figure 1. Two models of community assembly. **(A)** Local communities comprise a subset of species from the regional species pool that have passed through environmental filters. There is no feedback from the metacommunity (collection of local communities) to the regional species pool. Adapted from [5]. **(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.

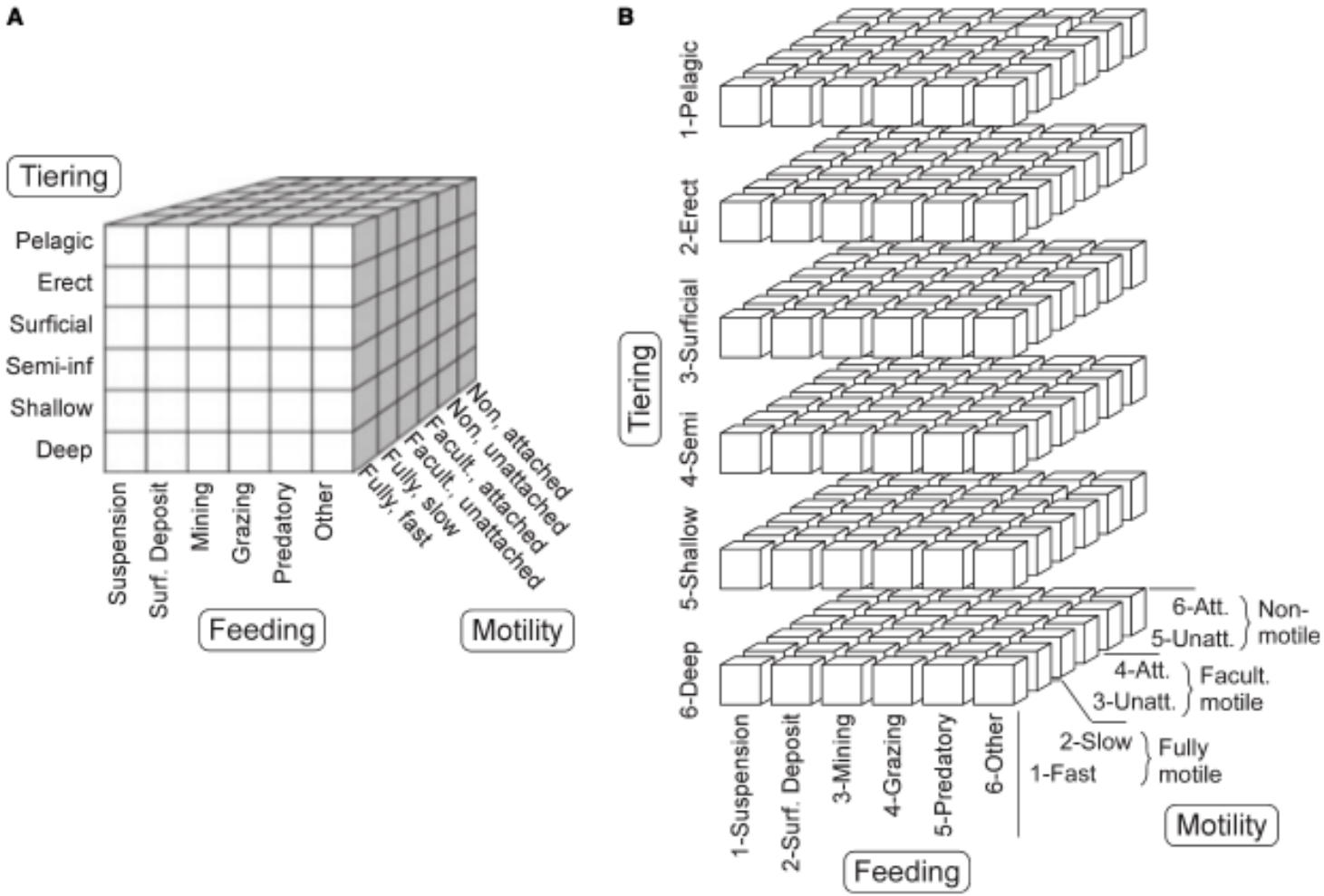
(Mittelbach and Schemske, 2015, *TREE*)

Differences in extinction risk



(Smits 2015 *PNAS*)

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)

Functional groups over time

(Bush and Bambach 2011 *Annu. Rev. Earth Planet Sci.*)

Structured data in biology and paleontology

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

Models of structured data

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

The fourth-corner problem

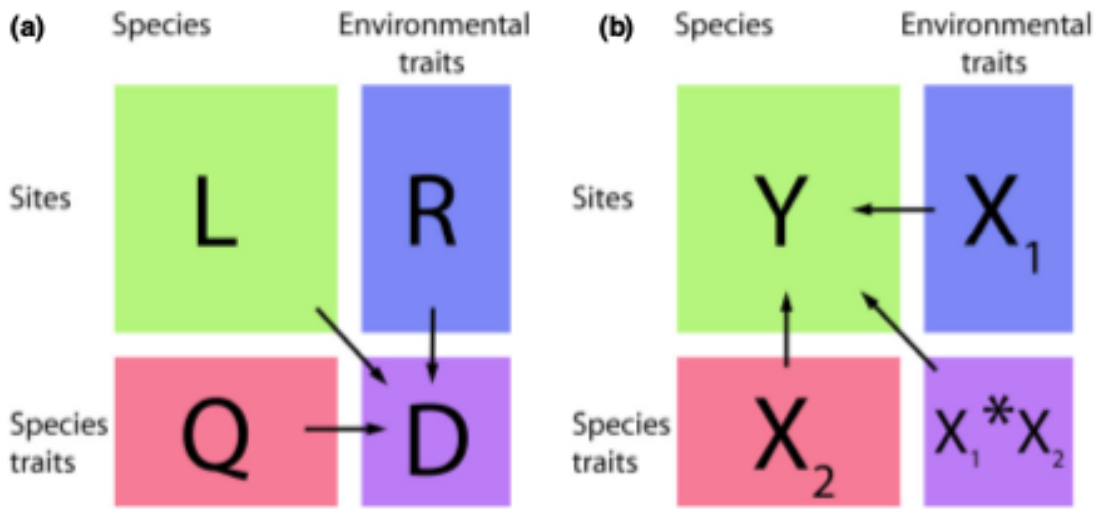
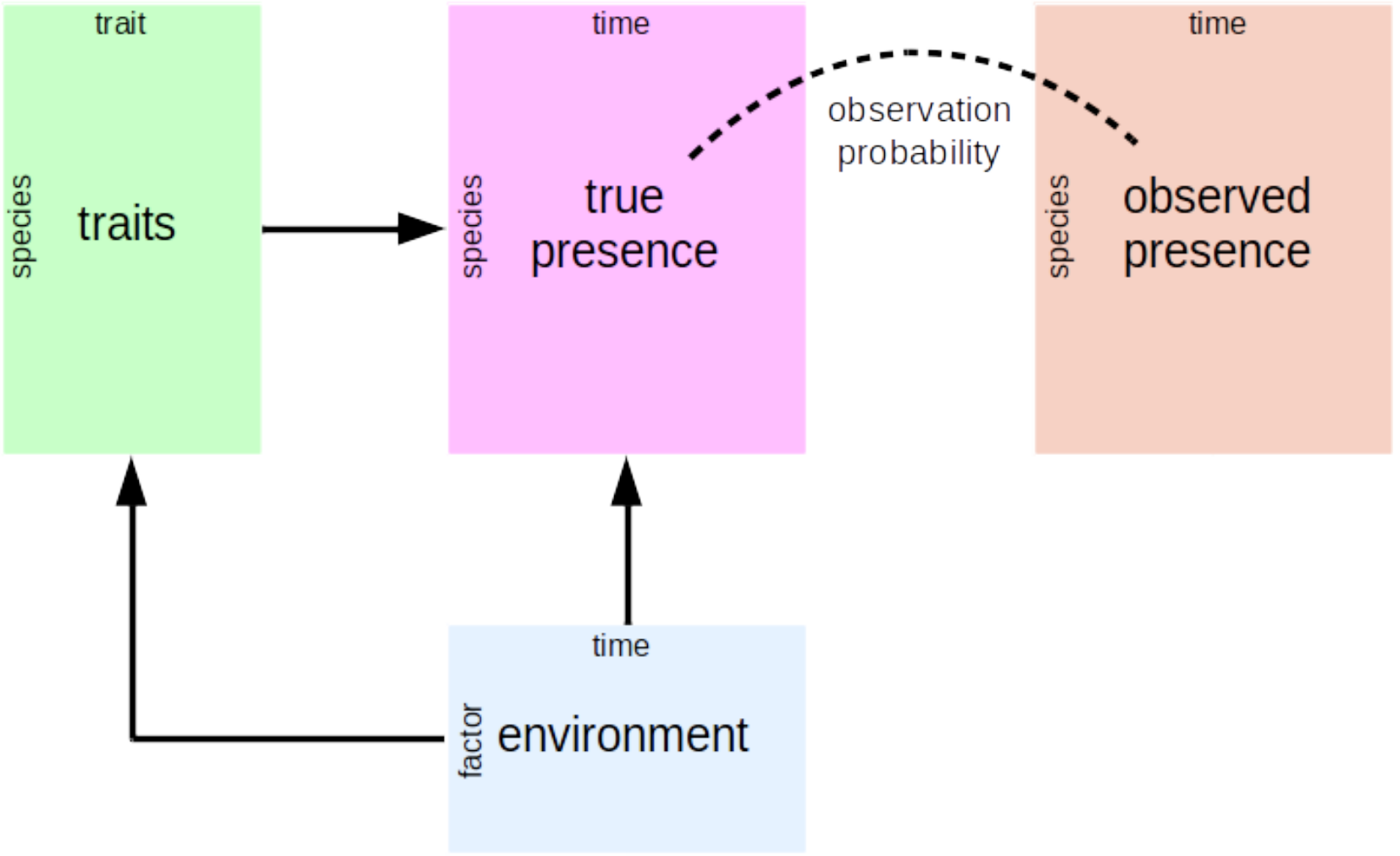


Fig. 1. Graphical representation of the fourth-corner problem and its solution. (a) The problem as posed by Legendre, Galzin & Harmelin-Vivien (1997), where the goal is to combine abundance (**L**), trait (**Q**) and environment (**R**) data in some way, to determine a matrix describing the trait–environment relationship (**D**). (b) The proposed model-based solution to the fourth-corner problem, where the goal is to predict abundance (**Y**) as a function of predictor variables for environment ($\mathbf{X_1}$), species traits ($\mathbf{X_2}$) and their interaction ($\mathbf{X_1}$*$\mathbf{X_2}$). The matrix of coefficients for the interaction between $\mathbf{X_1}$ and $\mathbf{X_2}$ is the fourth corner.

(Brown *et al.*, *2014*, Methods Ecol. Evol.)

Paleo-fourth corner model



Covariates of interest

individual-level
(species i at time unit t)

- effect of locomotor type
  - arboreal, digitigrade, plantigrade, unguligrade, fossorial, scansorial
- effect of dietary type
  - carnivore, herbivore, insectivore, omnivore
- effect body size
  (rescaled log body mass)

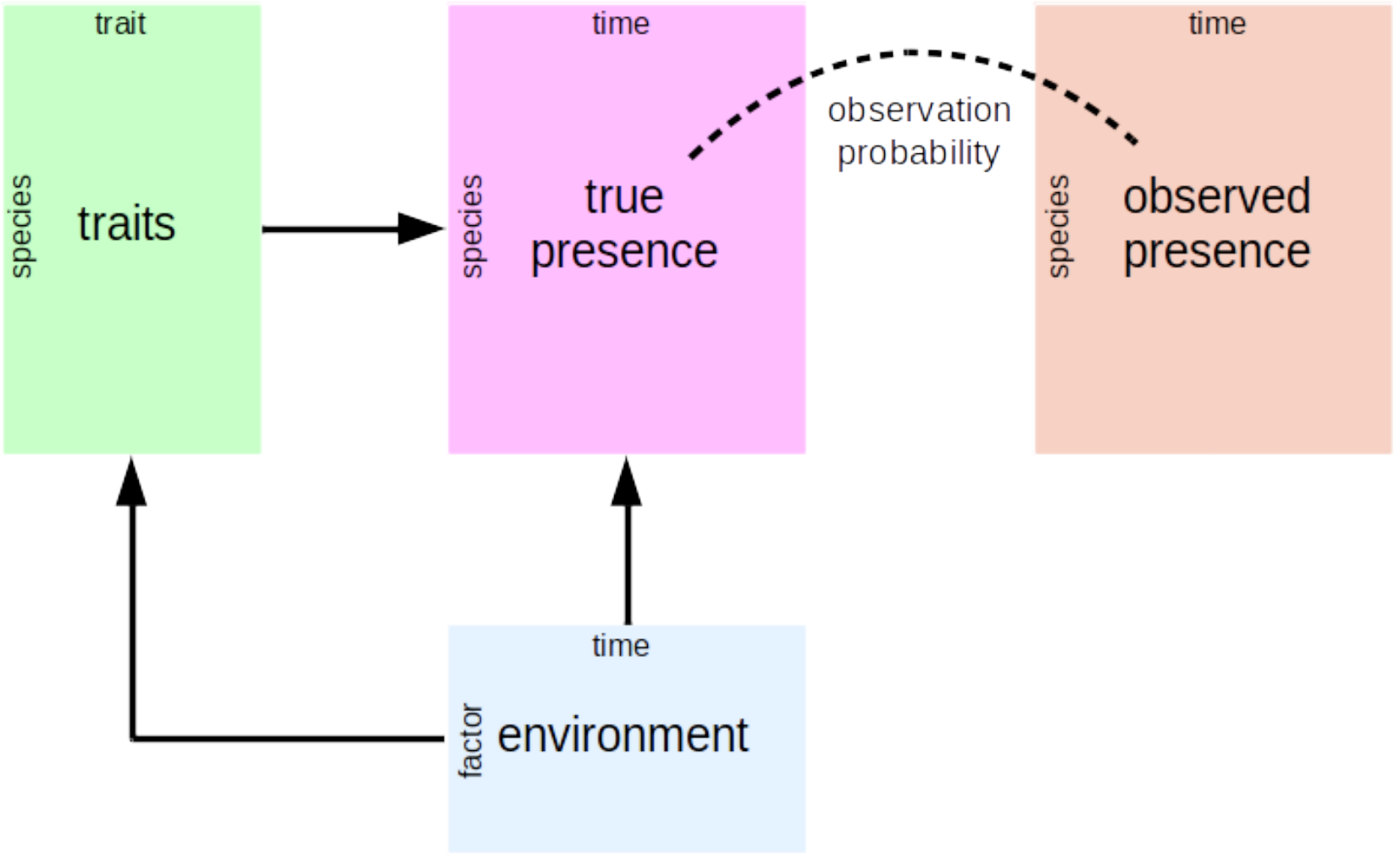
group-level (2 My time unit t)

- temperature record based on Mg/Ca estimates
  - mean and range
    (rescaled log degrees)
- plant community phase
  following Graham 2011

Model of taxon occurrence

- response is p/a of genus in NA at time t
  - Bernoulli variable
  - probability is (observation prob) times ("true" presence)
- observation probability is effect of sampling/fossil record
  - basic model does not model sampling
- the latent discrete "true" presence modeled as a multi-level logistic regression
  - individual- and group-level

Paleo-fourth corner model



Model and sampling statement definition

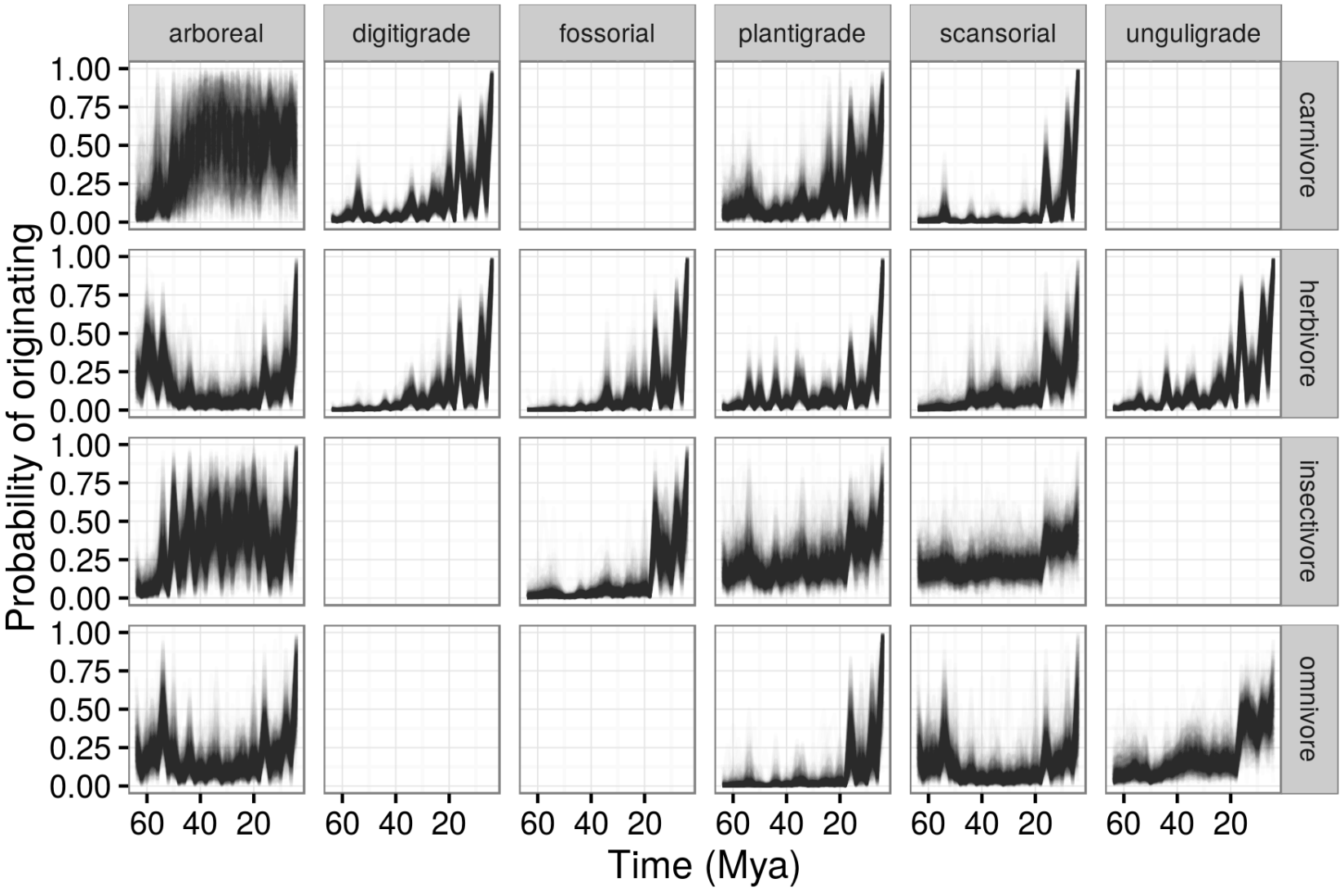
$$
\begin{aligned}
y_{i,t} &\sim \text{Bernoulli}(p_{i,t} z_{i,t}) \\
p_{i,t} &= \text{logit}^{-1}(\alpha_0 + \alpha_1 m_i + r_t) \\
r_t &\sim \mathcal{N}(0, \sigma) \\
\alpha_0 &\sim \mathcal{N}(0, 1) \\
\alpha_1 &\sim \mathcal{N}(1, 1) \\
\sigma &\sim \mathcal{N}^{+}(1) \\
z_{i,1} &\sim \text{Bernoulli}(\phi_{i,1}) \\
z_{i,t} &\sim \text{Bernoulli}\left(z_{i,t-1}\pi_{i,t} + \sum_{x=1}^{t}(1 - z_{i,x})\phi_{i,t} \right) \\
\phi_{i,t} &= \text{logit}^{-1}(a^{\phi}_{t,j[i]} + b^{\phi}_1 m_i + b^{\phi}_2 m_i^2) \\
\pi_{i,t} &= \text{logit}^{-1}(a^{\pi}_{t,j[i]} + b^{\pi}_1 m_i + b^{\pi}_2 m_i^2) \\
a^{\phi} &\sim \text{MVN}(U\gamma^{\phi}, \Sigma^{\phi}) \\
a^{\pi} &\sim \text{MVN}(U\gamma^{\pi}, \Sigma^{\pi})
\end{aligned}
$$

$$
\begin{aligned}
\Sigma^{\phi} &= \text{diag}(\tau^{\phi})\Omega^{\phi}\text{diag}(\tau^{\phi}) \\
\Sigma^{\pi} &= \text{diag}(\tau^{\pi})\Omega^{\pi}\text{diag}(\tau^{\pi}) \\
\rho &\sim \text{U}(0, 1) \\
b^{\phi}_1 &\sim \mathcal{N}(0, 1) \\
b^{\pi}_1 &\sim \mathcal{N}(0, 1) \\
b^{\phi}_2 &\sim \mathcal{N}(-1, 1) \\
b^{\pi}_2 &\sim \mathcal{N}(-1, 1) \\
\gamma^{\phi} &\sim \mathcal{N}(0, 1) \\
\gamma^{\pi} &\sim \mathcal{N}(0, 1) \\
\tau^{\phi} &\sim \mathcal{N}^{+}(1) \\
\tau^{\pi} &\sim \mathcal{N}^{+}(1) \\
\Omega^{\phi} &\sim \text{LKJ}(2) \\
\Omega^{\pi} &\sim \text{LKJ}(2).
\end{aligned}
$$

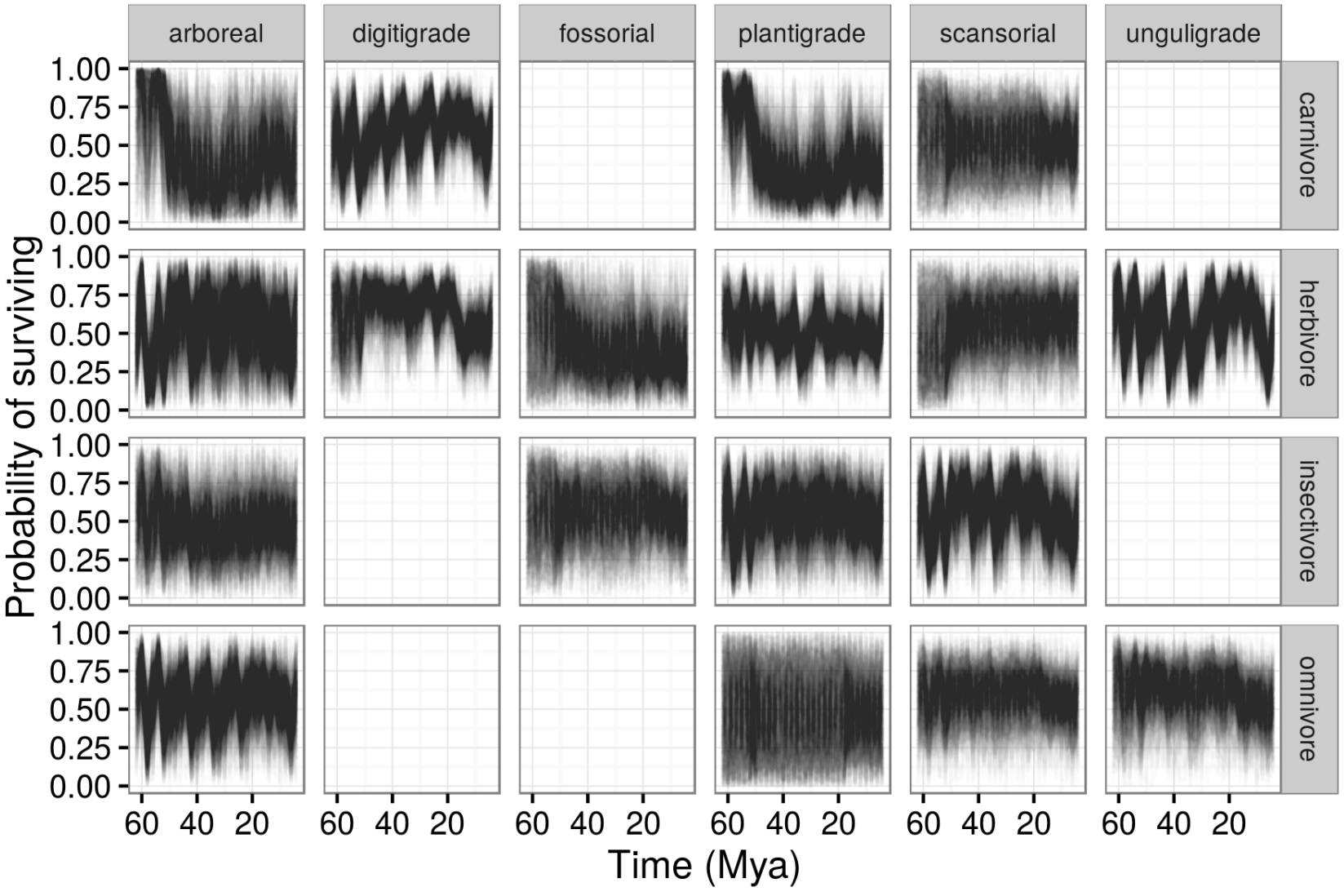
Bayesian inference and statistics

- flexible, expressive, intuitive
- regularize, partial pooling, external information
- Stan probabilistic programming language
  - **Hamiltonian Monte Carlo**
  - Automatic Differentiation Variational Inference

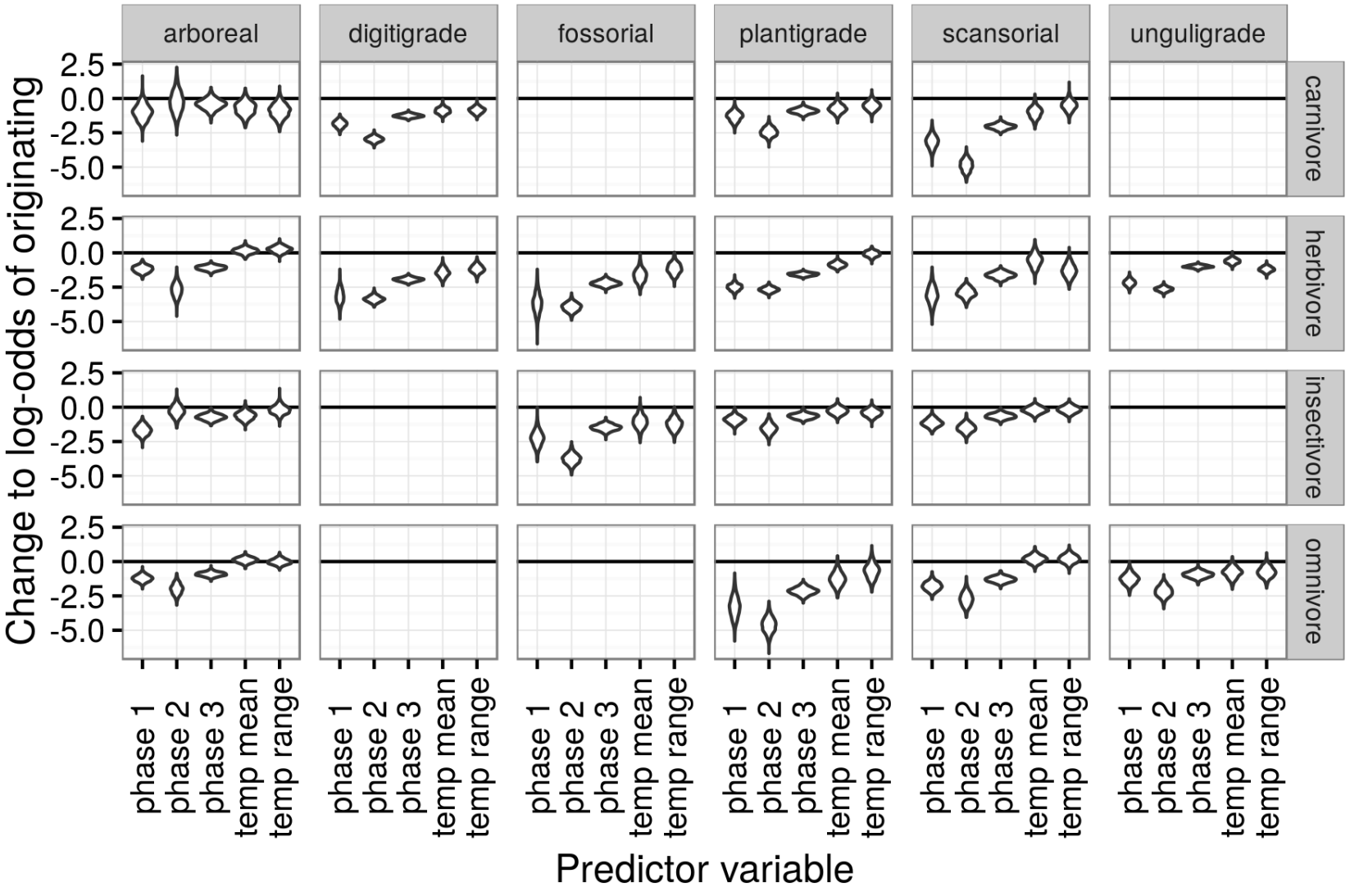
Probability of ecotype origination



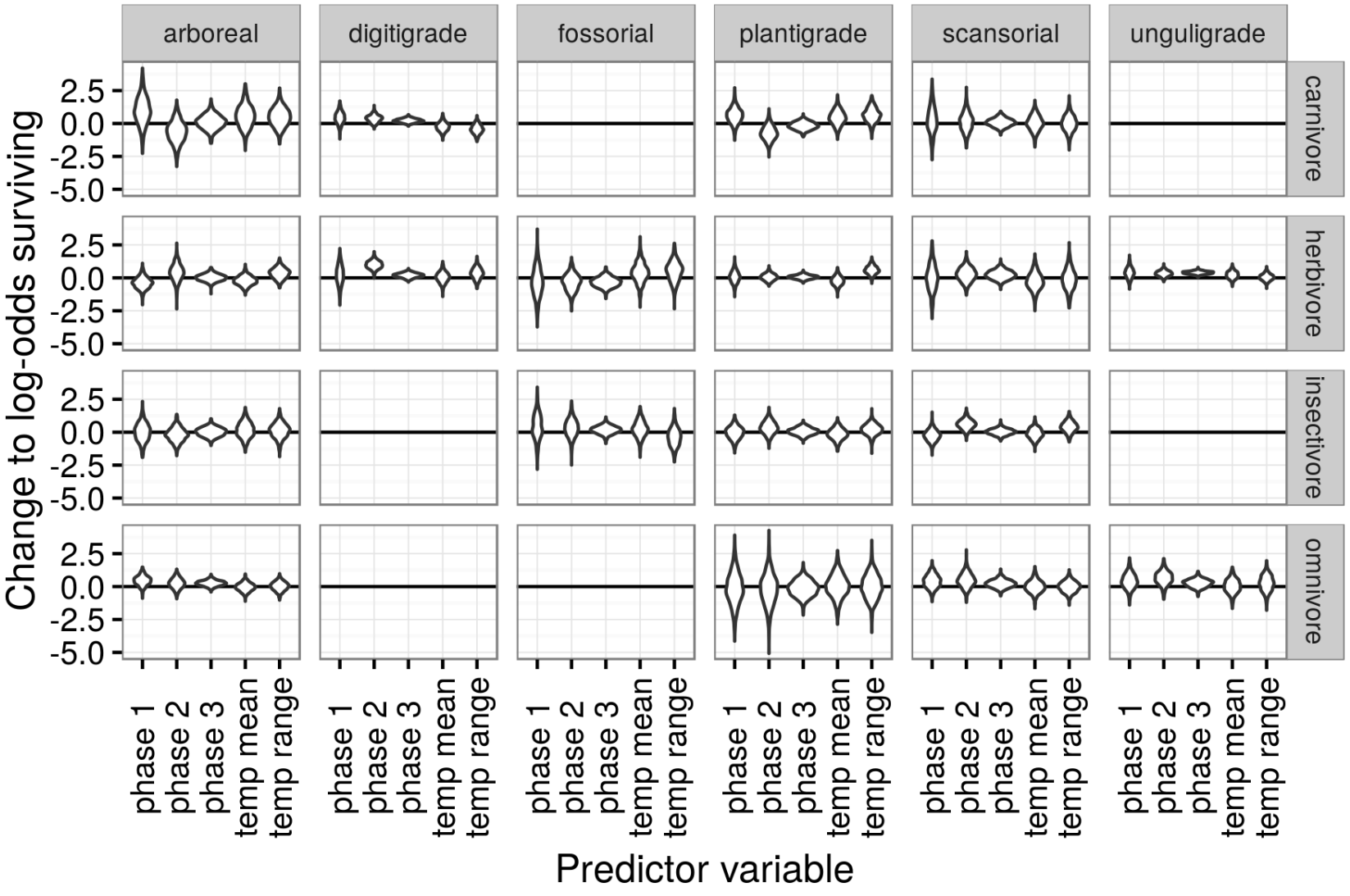
Probability of ecotype survival



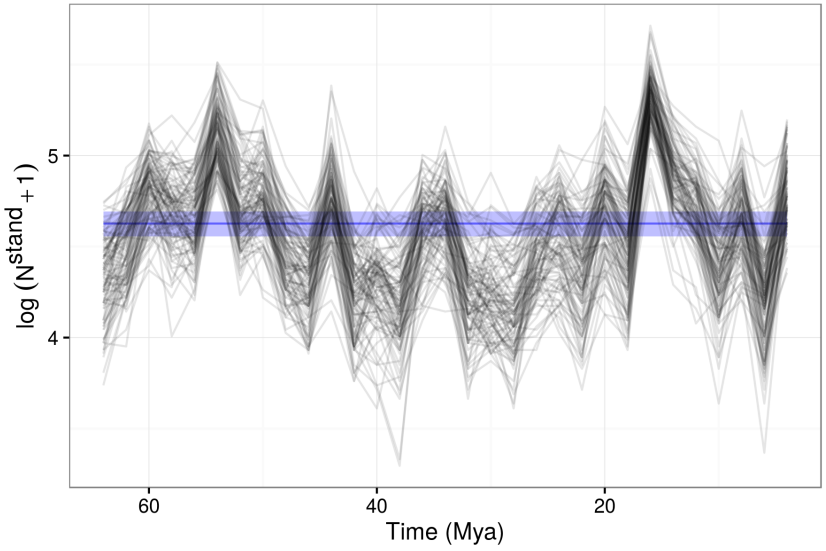
Group-level effects (plant phase, climate) on origination

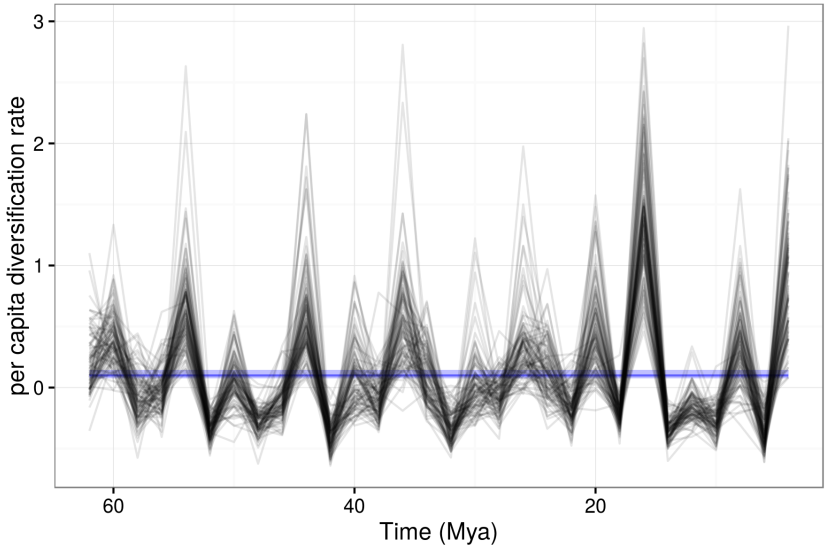


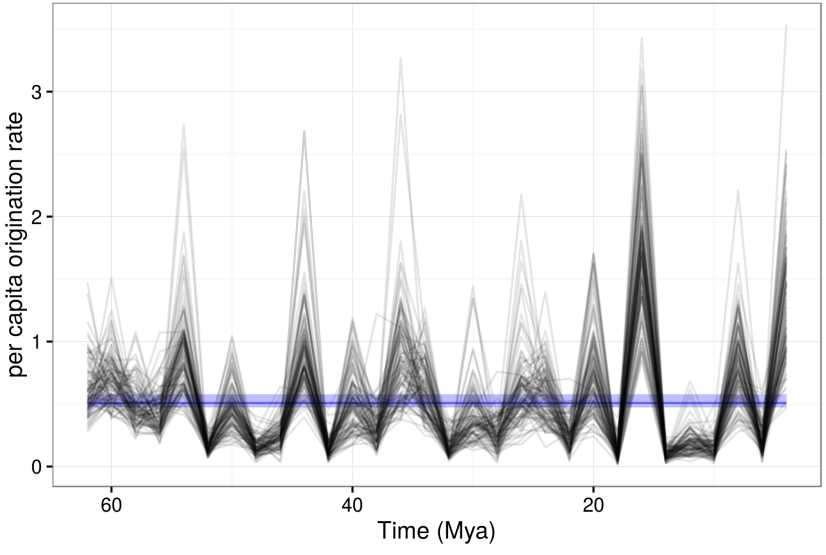
Group-level effects (plant phase, climate) on survival

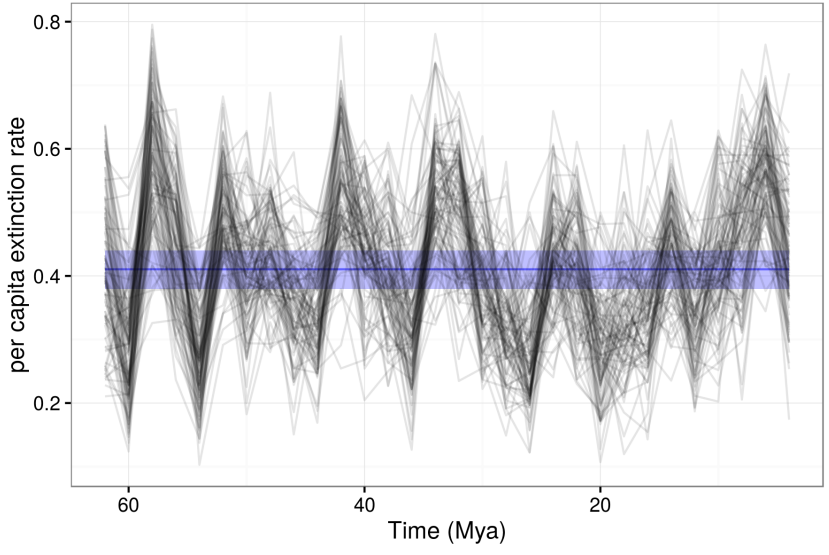


Total species pool diversity and diversification

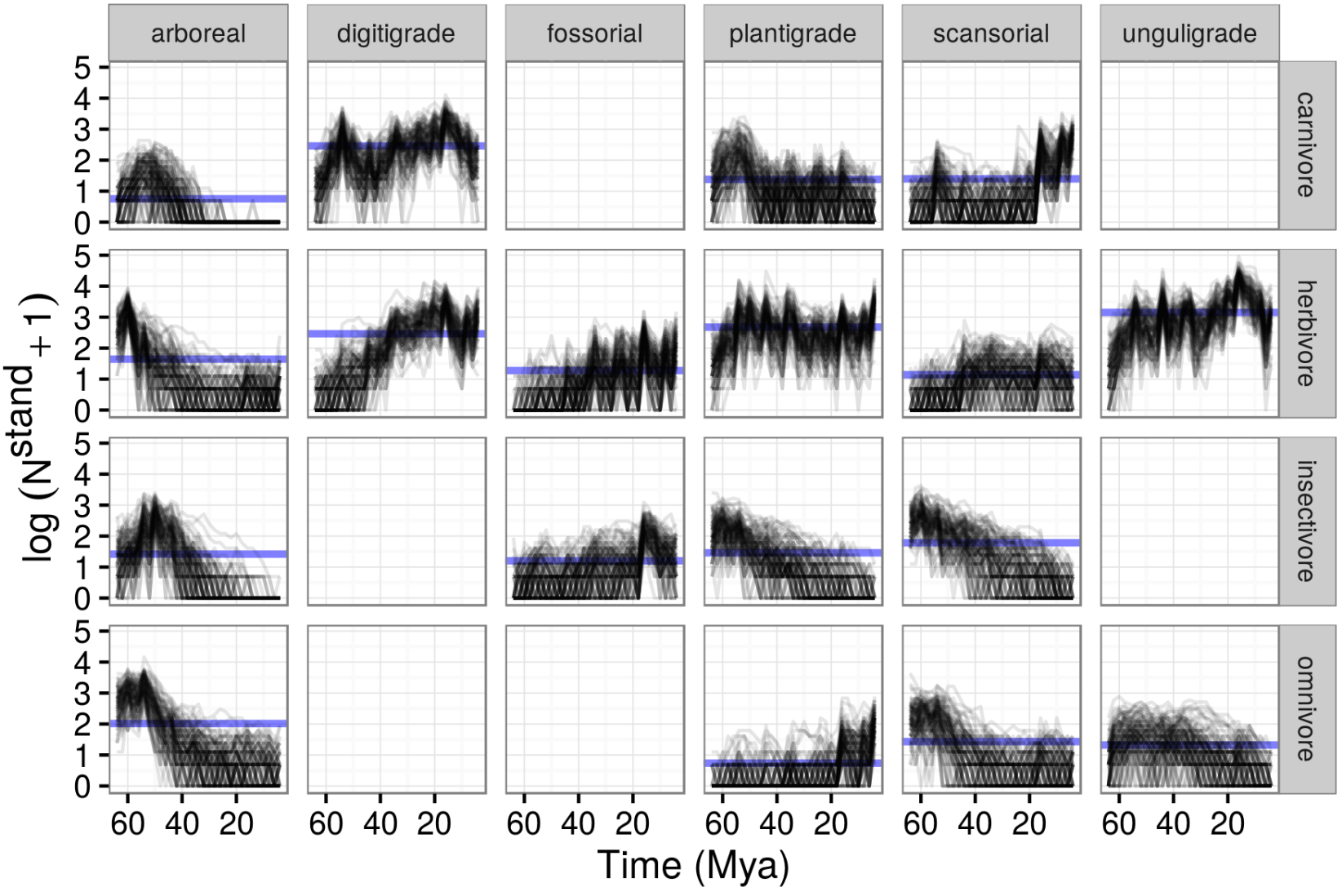




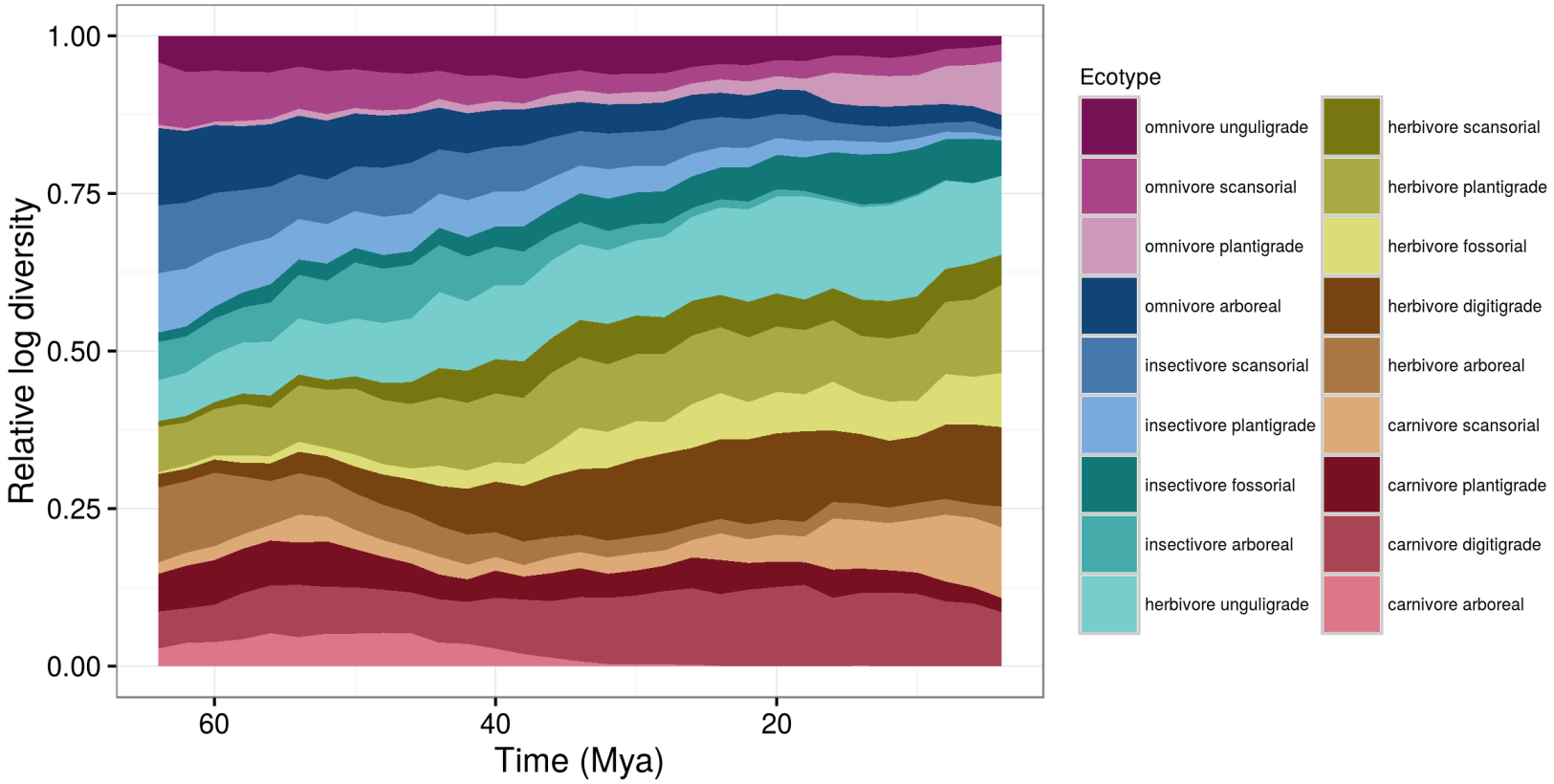




Ecotype-specific diversity



Relative ecotype diversity



Summary of results

- changes to ecotype composition driven by origination, not extinction
  - specific ecotypes source of most variation in overall origination
- arboreal taxa decrease through Paleogene, all but absent by Neogene
- digitigrade and unguligrade herbivores only groups with sustained increase
- environmental covariates virtually always affect origination, not survival

Acknowledgements

- Advising
  - Kenneth D. Angielczyk, Michael J. Foote, P. David Polly, Richard H. Ree, Graham Slater
- Angielczyk Lab
  - David Grossnickle, Dallas Krentzel, Jackie Lungmus
- Foote lab
  - Marites Villarosa Garcia, Nadia Pierrehumbert
- David Bapst, Ben Frable, Graeme Lloyd, Matt Pennell
- UChicago CEB

The Paleobiology Database
revealing the history of life