Computational geological approaches for assessing the diversity and ecological distribution of fossils

Bjorn-Gustaf J. Brooks

Department of Geological & Atmospheric Sciences

Iowa State University

Program of Study Committee

Cinzia Cervato, Major Professor

James L. Cornette

Chris Harding

Carl E. Jacobson

John Mayfield

February 27, 2009

Controls on biodiversification and extinction

- Many of the most dramatic events in Earth history have occurred within the past 542 Myr, the Eon of animal life.

- The processes behind those changes are an essential component of geology, and are of interest to biology and ecology.

- Understanding biosphere response today is aided by placing it in the context of historical biosphere response determined from the fossil record.

- Understanding these processes is important to understanding the future of all life on Earth.

Thesis hypotheses

1. A recent hypothesis proposed that marine animal life undergoes cycles of extinction on $\sim$62 $\sim$140 Myr periods (Rohde and Muller 2005).

2. Renema and others (2008) have proposed that there has been a correlation between biodiversity and tectonism during the most recent 50 Myr and that plate tectonics could be an important driver of biodiversity.

The Sources of Cycles in Phanerozoic Biodiversity

Bjorn-Gustaf J. Brooks, James L. Cornette and Cinzia Cervato

Phanerozoic biodiversity

- Cycles in Phanerozoic biodiversity are interpreted from biodiversity curves. Diversity curves are time series that illustrate the variety of extant species or genera throughout a period of time.
- The development of Phanerozoic diversity curves:
    - Diversity curves go back at least as far as the 1930's (Cain 1938), but the first reliable curves were made by David M. Raup and J. John Sepkoski (Raup 1976, Sepkoski 1978).
    - Diversity curves were used to show that most of the basic anatomical differences among animals (*i.e.* bauplans) evolved early and have not fundamentally changed over the past 500 Ma. The variety of representatives of those original anatomical designs, however, has changed a great deal.
    - The processes that work to drive such biodiversity change are of great interest.

Phanerozoic biodiversity

- The Sepkoski *Compendium of Fossil Marine Animal Genera* has been the most widely used source for Phanerozoic biodiversity data.
- The most extensively annotated Phanerozoic biodiversity database, the Paleobiology Database (Alroy *et al.* 2008), represents a new tabulation of paleodiversity.
  - The Paleobiology Database (PD) is a recent development representing more than 700,000 fossil observations drawn from more than 29,000 published articles entered by over 200 paleontologists.
  - There have been 91 published articles that have used the Paleobiology Database, 25% in *Science*, *Nature*, or *PNAS*.

Diversity curves: PD and Sepkoski's Compendium
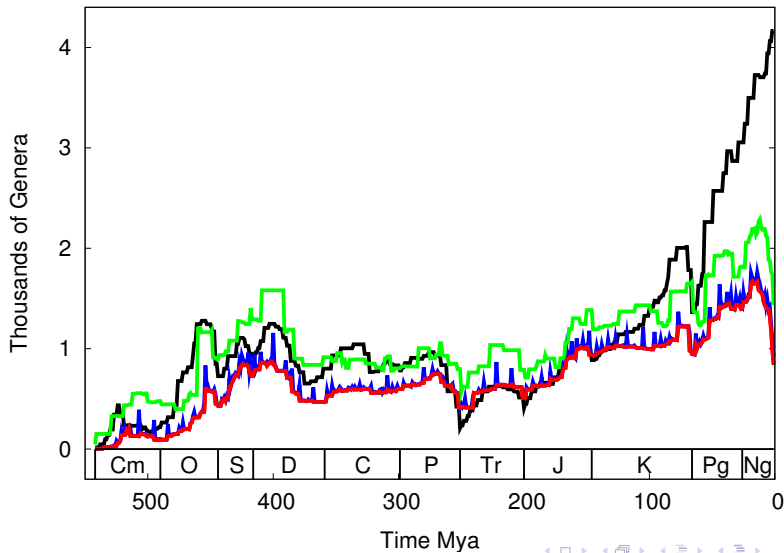
Biodiversity curves

- To test the sensitivity of biodiversity cycles in each diversity collection we examined the effect that 4 different interpretations would have on their biodiversity cycles.

1. Broad interval (green curve): the smallest interval that spans the maximum and minimum ages for a genus. Assumes that a genus is found continuously throughout a section.

2. Narrow interval (blue curve): the smallest interval containing the midpoint ages for a genus. Assumes that a fossil does not occur continuously within a section.
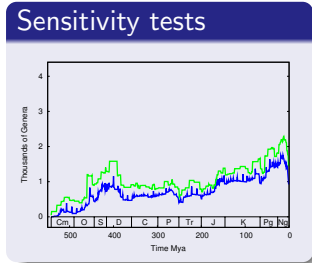
Sensitivity tests

Biodiversity curves

3. Including single occurrence genera (blue curve) : single occurrence genera are represented by only one occurrence.

4. Excluding single occurrence genera (red curve).

Sensitivity tests
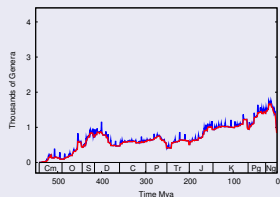
Discrete Fourier Transform Power Spectrum

Example DFTPS

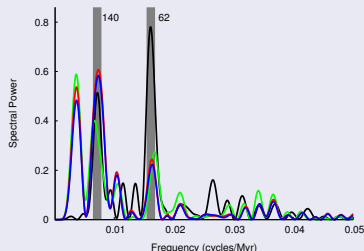To identify the frequencies within a time series that show unusually high amplitudes we used Fourier transform.

procedure dft
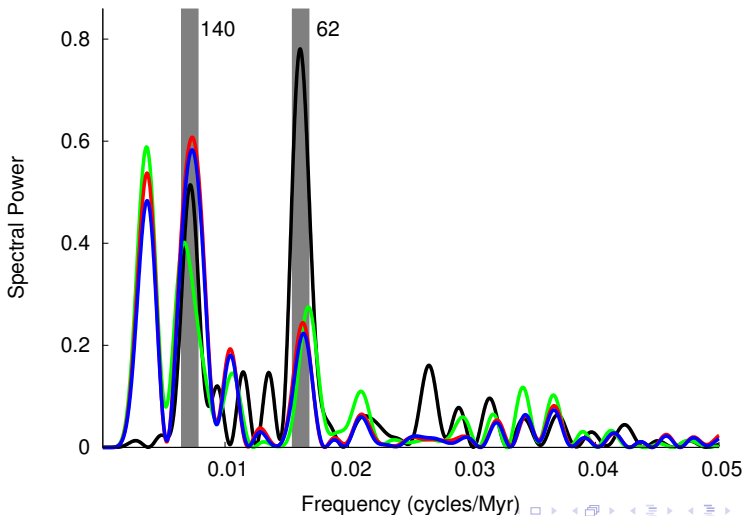*input:* time series $\{(t_i, x_i)\}_{i=1}^{542}$
output: a (amplitude), and τ (phase)

```
x_c = poly_fit(x,t,3)
x_p = poly(x_c,t)
x_d = x - x_p
x_n = x/std_dev(x_d)
x_ft = fft(x_n)
x_sp = x_ft * conj(x_ft)
a_k = real(x_sp)
tau_k = imaginary(x_sp)
```

DFTPS

Biodiversity cycles: Rohde & Muller (black), PD all marine (blue), PD no single (red), PD broad (green)

Interpretations

- Glacial episodes (Veizer 1999, 2000).
  The peaks and troughs of $\delta^{18}O$ excursions from this extensive dataset show a dominant $\sim$135 Myr periodicity that coincides with major glacial events. However this model conflicts with some energy balance models.

- Cosmic ray flux (Medvedev and Melott 2007).
  CR flux cause by the north-south oscillations of the solar system through the Milky Way are hypothesized as a driver of the $\sim$62 Myr cycle. Several mechanisms are proposed: climate change, ozone depletion, radiation damage. All are possible, but the specific evidence for the occurrence of each scenario during the Phanerozoic is weak or non-existent.

- Rock outcrop area (Smith and McGowan 2005).
  Variable rock exposure is a concern for diversity curves. Over limited intervals it has been shown to mirror fossil diversity in proxy regions. However, it is not known how well this correlation extrapolates to interpolated Phanerozoic diversity on the global scale.

Sources of the ∼62 Myr cycle: Environment

PD environments

PD environment	Number of Genera
marine indet.	7850
reef, bioherm	2612
carbonate indet.	1865
coastal indet.	1258
offshore	1240
subtidal indet.	1148
transition zone	501
offshore shelf	492
open shallow	464
marginal marine	460
shallow subtidal	385
lagoonal	312
deep subtidal	291
basinal	259
slope	255

Nearshore (blue) and Offshore (black) collections
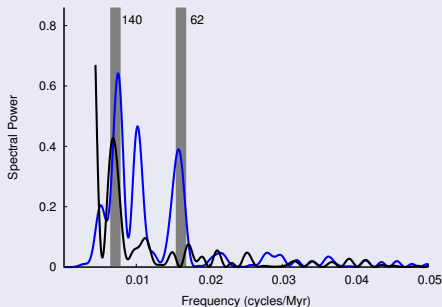
Sources of the ∼62 Myr cycle: Phylogeny

Phylogenetic Classes

Phylogenetic Class	Number of Genera
Mammalia	3208
Gastropoda	2689
Bivalvia	1889
Cephalopoda	1832
Articulata	1783
Reptilia	1481
Anthozoa	1279
Trilobita	1261
Granuloreticulosea	693
Ostracoda	523
Crinoidea	487
Radiolaria	465
Stenolaemata	351
Echinoidea	322

GBA (blue) and GBA_complement (black) collections
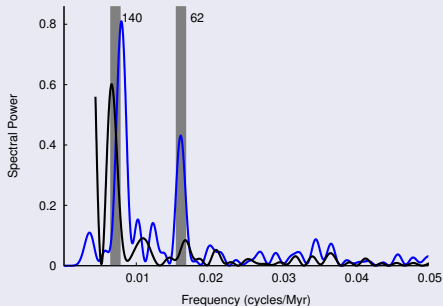
Sources of the ∼62 Myr cycle: Environment and Phylogeny

Significances of spectral peaks near 62 Myr period

Random step (RS) and random block (RB) Monte Carlo trials were used to determine the significances of peaks relative to two different null hypotheses. The larger of the two, RB*, was used to estimate the significance of peaks near ∼62 Myr.

Genera	Including single occs.			Excluding single occs.		
	Period	RS	RB*	Period	RS	RB*
G21663	61.2	0.028	0.203	61.4	0.019	0.177
GBA	62.2	0.002	0.028	62.2	0.001	0.027
NonGBA	59.8	0.211	0.534	60.0	0.228	0.557
Nearshore	63.3	0.002	0.049	63.4	0.002	0.050
Offshore	58.3	0.245	0.549	58.4	0.265	0.586
G13687	61.0	0.030	0.224	61.1	0.027	0.212
GBA	62.4	0.002	0.034	62.4	0.002	0.033
NonGBA	59.4	0.161	0.471	59.5	0.171	0.490
Nearshore	63.0	0.001	0.025	63.0	0.001	0.022

Discussion

- The ~62 Myr cycle in Phanerozoic biodiversity that appears in collections of PD genera is not the result of oscillations in all phylogenetic groups. The strength of this cycle lies in the gastropod, bivalve, and articulate brachiopod genera, which comprise about a third of PD genera in this dataset.


All marine biodiversity
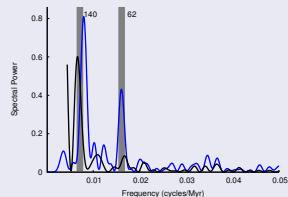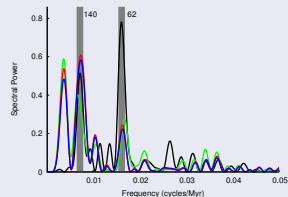
Discussion

- The robustness of the ∼62 Myr cycle is as strong as the coverage of genera extracted from PD. We further explored the PD coverage by comparing it to Sepkoski's Compendium and found three essential differences:

1. Rates of extinction and recovery.

2. The drop in diversity from 10 Ma to present. This may represent a bias in reporting extant taxa.

3. Biodiversification over the most recent 100 Myr in PD is only about half that of Sepkoski's Compendium. This may signal a sampling bias in PD (Alroy *et al.* 2001, 2008, Bush *et al.* 2004, Krug *et al.* 2009).

All marine biodiversity
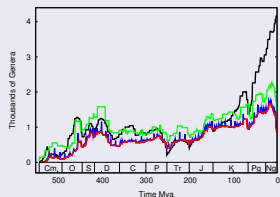
Summary

- Unlike Sepkoski's Compendium, the Paleobiology Database's annotations permit us to search for sources of biodiversity cycles by examining subsets of genera. The significant ∼62 Myr cycle that occurs GBA and nearshore collections of PD genera presents an important new step toward uncovering the underlying causes of Phanerozoic biodiversity cycles.

- A hypothesis that implicates external drivers for biodiversity cycles (*i.e.* glaciation, cosmic ray flux) could be strengthened if it can account for the strong ∼62 Myr cycle in GBA and nearshore collections of PD genera.

- The hypothesis that cycles in Phanerozoic biodiversity are subsumed by cycles in sedimentary rock exposure is difficult to reconcile with the results of this study because the ∼62 Myr cycle does not occur in all PD genera.

Evidence for the Influence of Tectonism on Marine Biodiversity

Bjorn-Gustaf J. Brooks and Cinzia Cervato

The hopping hotspots hypothesis

- A recent biogeographical paper by Renema and others (2008) studied paleodiversity over the past 50 Myr, and proposed that biodiversity has been greatest in areas and in times of substantial tectonic activity.

- This hypothesis is dissatisfying because it does not specifically address why biodiveristy should be associated with tectonically active areas.

The hopping hotspots hypothesis

- The hopping hotspots hypothesis poses an interesting new challenge for G. Evelyn Hutchinson's ecological niche theory: Does tectonism override the usual competitive controls hypothesized by niche theory (Gould 1985)?

- We used the Paleobiology Database to test the hopping hotspots hypothesis with a larger, globally distributed dataset.

Testing the hopping hotspots hypothesis

- Renema and others (2008) defined the alpha diversity at each Cenozoic sub-epoch as the variety of fossil foraminifera observed within a drill core.

- This study standardized diversity measures by comparing them among same-sized assemblages that are defined as $1°$ longitude $\times$ $1°$ latitude.

- This study examined multiple diversity measures including alpha, beta (difference), and source pool diversity.

- This study also incorporated information about the number of habitat types, which were determined by the lithological annotations of each occurrence.

Habitat type determination

- Habitat types were established in two steps:

1. Assigning percentage values of sand, silt, clay, lime mud, and water depth to each lithological category.

2. k-means clustering to assign each occurrence a habitat type.

Conversion of categorical data

Lithological categories. Value equivalents for 2 of 336 lithological categories. SA indicates percent sand- SI, percent silt- CL, percent clay- and LM, percent lime mud.

PD environment	SA	SI	CL	LM	Explanatory notes
bafflestone only	10	10	10	0	Rocks whose chief characteristic is that they are bound by a matrix of organisms that acted as baffles during deposition and contain smaller grains, therefore a small, but equal fraction of sand silt and clay are assigned for these rocks.
claystone only	0	0	100	0	Fine-grained siliciclastic rocks consisting entirely of clay composed of grains less than 2 mm.

k-means clustering

k-means procedure

- To simplify exposition of the data, a multivariate statistical cluster analysis technique (k-means), was employed to identify unique habitat types. This was run on a high performance cluster computer.

- The k-means procedure consists of two parts: initial centroid or seed determination and iterative clustering of observations until convergence is reached.

Clustering
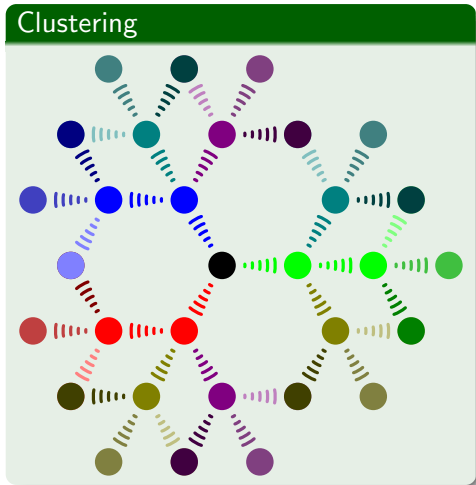
k-means cluster assignment

Example

Observation	Depositional Environment	Depth (m)	Primary & Sec- ondary Lithology	Sand (%)	Silt (%)	Clay (%)	Hab. Assgn.
C. pygmaea	offshore	200	siltstone only	0.0	83.3	16.7	27
C. pygmaea	offshore	200	shale only	0.0	33.0	67.0	38
E. ovatus	shoreface	15	sandstone only	100.0	0.0	0.0	16
G. vulcanica	int. bay	0	mudstone & sst	28.8	35.6	35.6	35

Local diversity by region. Mean abundances per assemblage ($1°$ longitude $\times$ $1°$ latitude) for alpha, beta, habitat type, and source pool diversity, as well as connectivity are given for the Miocene during which and the Eocene.

Interval	Region	α-Div.	β-Div.	Hab. Types	Source Pool Div.	Connect.
Late Miocene	E. N.Am.	50.8	0.89[‡]	5.6	85.4	0.21
	Europe	24.8	0.73	2.8	48.0	0.19
	Japan*	44.2[‡]	0.70	4.5[‡]	84.5	0.21
Middle Miocene	E. N.Am.	72.5	0.56	5.3	90.1	0.11
	Europe	36.8	0.66	4.1	78.1	0.23
	Japan*	24.1	0.84[‡]	2.8	41.7	0.20
Early Miocene	E. N.Am.	79.7[‡]	0.85	6.7[‡]	150.6	0.30
	Europe	35.3	0.44	3.7	22.8	0.10
	Japan*	12.6	0.60	2.8	10.8	0.11
Late Eocene	E. N.Am.	59.4	0.79	5.2	120.5	0.21
	Europe[†]	60.8[‡]	0.56	6.0[‡]	81.0	0.12
	Japan	17.5	0.52	3.1	18.5	0.10
Middle Eocene	E. N.Am.	61.2	0.87	5.8	99.9	0.23
	Europe[†]	47.8	0.55	3.8	97.2	0.16
	Japan	27.9	0.55	3.5	35.2	0.22
Early Eocene	E. N.Am.	31.2	0.69	7.1	43.8	0.22
	Europe[†]	33.7	0.64[‡]	3.3	77.8	0.25
	Japan	16.5	0.00	1.0	1.0	0.00

Relative importance of ecosystem parameters

Correlations. Each column lists the Pearson correlation (r) for alpha and beta diversity to various parameters.

Parameter	α diversity	β diversity
Hab. Type Div.	0.857	-0.312
Connectivity	0.310	0.294
Mean Depth	-0.036*	0.022*
Depth Range	0.573	-0.181
Mean Sand Pct.	0.182	0.033*
Mean Silt Pct.	0.190	-0.087*
Mean Clay Pct.	0.142	-0.102*
Mean Lime Mud Pct.	0.132	-0.073*

∗ Non-significant correlation
($P > 0.001$, two-tailed test)
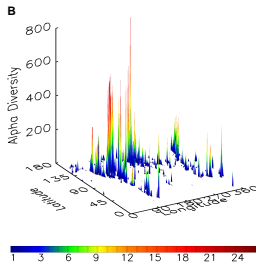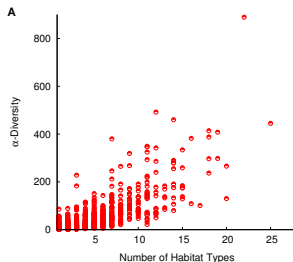
Summary of results from the PD

1. Alpha diversity in Japan and in Europe peaked during epochs of substantial tectonic activity.

2. However, the greatest alpha diversity occurs along the passive continental margin of eastern North America, far from any tectonic activity.

3. When all PD data are examined using same-sized assemblages, alpha diversity is most correlated with the diversity of habitat types ($r = 0.857$). Alpha diversity peaks within tectonically active regions also coincided with peaks in the diversity of habitat types. This supports niche theory, and suggests that episodes of tectonic activity may indirectly increase biodiversity by creating additional environmental gradients or resources. However, it is the number of habitat types, the environmental variety, that is most related to alpha diversity in this study.

Thank you

The Sources of Cycles in Phanerozoic Biodiversity
Bjorn-Gustaf J. Brooks, James L. Cornette and Cinzia Cervato

Special thanks to R.K. Alexander and F.M. Hoffman
for their consultation on this research.

Evidence for the Influence of Tectonism
on Marine Biodiversity
Bjorn-Gustaf J. Brooks and Cinzia Cervato

www.climatemodeling.org/∼bjorn/pres/phd_defense.pdf