Non-equilibrium evolution of volatility in origination and extinction explains fat-tailed fluctuations in Phanerozoic biodiversity

Andrew J. Rominger[1, *], Miguel A. Fuentes[1, 2, 3], and Pablo A. Marquet[1, 4, 5, 6, 7]

[1]Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, New Mexico 87501, US
[2]Instituto de Investigaciones Filosóficas, SADAF, CONICET, Bulnes 642, 1428 Buenos Aires, Argentin
[3]Facultad de Ingeniería y Tecnología, Universidad San Sebastián, Lota 2465, Santiago 7510157, Chile
[4]Departamento de Ecología, Facultad de Ciencias Biológicas, Pontificia Universidad de Chile, Alameda 340, Santiago, Chile
[5]Instituto de Ecología y Biodiversidad (IEB), Casilla 653, Santiago, Chile
[6]Laboratorio Internacional de Cambio Global (LINCGlobal), and Centro de Cambio Global UC, Pontificia Universidad Catolica de Chile, Santiago, Chile.
[7]Centro Cambio Global UC, Av. Vicuña Mackenna 4860, Campus San Vicuña, Santiago, Chile
[8]Centro de Ciencias de la Complejidad (C3), Universidad Nacional Autónoma de México.
[*]To whom correspondence should be addressed, e-mail: rominger@santafe.edu

1

Fluctuations in biodiversity, both large and small, are pervasive through the fossil record, yet we do not understand the processes generating them. Here we extend theory from non-equilibrium statistical physics to describe the previously unaccounted for fat-tailed form of fluctuations in marine invertebrate richness through the Phanerozoic. Using this theory, known as superstatistics, we show that the simple fact of heterogeneous rates of origination and extinction between clades and conserved rates within clades is sufficient to account for this fat-tailed form. We identify orders and the families they subsume as the taxonomic level at which clades experience inter-clade heterogeneity and within clade homogeneity of rates. Following superstatistics we would thus posit that orders and families are subsystems in local statistical equilibrium while the entire system is not in equilibrium. The separation of timescales between background origination and extinction within clades compared to the origin of major ecological and evolutionary innovations leading to new orders and families allows within-clade dynamics to reach equilibrium, while between-clade diversification is non-equilibrial. This between clade non-equilibrium accounts for the fat-tailed nature of the system as a whole. The distribution of shifts in diversification dynamics across orders and families is consistent with niche conservatism and pulsed exploration of adaptive landscapes by higher taxa. Compared to other approaches that have used simple birth-death processes, equilibrial dynamics, or non-linear theories from complexity science, superstatistics is superior in its ability to account for both small and extreme fluctuations in the richness of fossil taxa. Its success opens up new research directions to better understand the evolutionary processes leading to the stasis of order- and family-level occupancy in an adaptive land-

2

scape interrupted by innovations that lead to new orders and families.

1 Introduction

Biodiversity has not remained constant nor followed a simple trajectory through geologic time (*1–5*). Instead, it has been marked by fluctuations in the richness of taxa, both positive in the case of net origination, or negative in the case of net extinction. Major events, such as adaptive radiations and mass extinctions have received special attention (*6, 7*), but fluctuations of all sizes are ubiquitous (*2, 5, 8*) and follow a fat-tailed distribution where large events are more probable compared to, e.g. a Gaussian distribution. Understanding the fat-tailed nature of these fluctuations continues to elude paleobiologists and biodiversity theoreticians.

The fat-tailed distribution of fluctuations in taxon richness inspired earlier researchers to invoke ideas from complex systems with similar distributions. Such ideas include the hypotheses that biological systems self-organize to the brink of critical phase-transitions (*9, 10*), and that environmental perturbations are highly non-linear (*11*). New data and analyses have not, however, supported these hypotheses at the scale of the entire Phanerozoic marine invertebrate fauna (*5, 12*). Other studies have modeled the mean trend in taxon richness as tracking a potentially evolving equilibrium (*2, 13–15*) and yet ignore the role of stochasticity and non-equilibrium dynamics in producing observed patterns (*4, 8, 16–18*). Individual, population, and local ecosystem scale processes that could produce complex dynamics, such as escalatory co-evolutionary interactions (*19*), have not been documented to scale up to global patterns (*20*) and indeed should not be expected to do so (*21*). Thus, we still lack a new hypothesis to describe the striking fat-tailed nature of fluctuations throughout the Phanerozoic.

Despite the heterogeneity of explanations of Phanerozoic biodiversity, consensus has

3

emerged on one property of macroevolution: clades experience different rates of morpho-logical evolution, origination and extinction (*2, 3, 22–25*). Here we show that the simple fact of conserved rates within clades and variable rates across clades is sufficient to describe pervasive, fat-tailed fluctuations in taxonomic richness throughout the marine Phanero-zoic. This biological mechanism has a precise correspondence to the non-equilibrial theory from statistical physics known as "superstatistics" (*26*) which has been applied across the physical and social sciences (*27, 28*). We leverage this correspondence to explain the distribution of fluctuations in the standing richness of marine invertebrates preserved in the Phanerozoic fossil record. We further show that the specific mathematical form of this superstatistical distribution is consistent with niche conservatism (*29, 30*) and pulsed exploration on an adaptive landscape by higher taxa (*30–32*). We operationally define "adaptive landscape" to mean a clade's set characteristics that influences its macroevo-lution. Those characteristics could be ecological (e.g. substrate preference (*30, 33, 34*), morphological (e.g. body plan (*16*)), or macroecological (e.g. range size (*35, 36*)).

1.1 Superstatistics of fossil biodiversity

Superstatistics (*26*) proposes that non-equilibrial systems can be decomposed into many local sub-systems, each of which attains a unique dynamic equilibrium. The evolution of these dynamic equilibria across sub-systems occurs more slowly. This separation in time scale allows local systems to reach equilibrium while the system as a whole is not in equilibrium (*26*). In the context of macroevolution we propose that a clade with conserved macroevolutionary rates corresponds to a sub-system in dynamic equilibrial.

In statistical mechanics, local sub-systems can be defined by a simple statistical param-eter β often corresponding to inverse temperature. In macroevolutionary "mechanics" we define the β_k of clade k as the inverse variance of fluctuations x_k in the number of genera

within that clade, i.e. fluctuations in the genus richness. The β_k thus represent the inverse variances, what we term volatilities, of the stationary distribution of a homogeneous origination-extinction processes of genera. Fluctuations from this stationary process will be approximately Gaussian if the clades' diversification dynamics are independent and in local equilibrium (see Supplemental Section S1; (37, 38)).

We make the hypothesis of dynamic equilibrium within a clade following MacArthur and Wilson (39) in recognition that while the identity and exact number of taxa will fluctuate stochastically from random origination and extinction (taking the place of local immigration and extinction in island biogeography), the overall process determining the number of taxa, and by extension, fluctuations in that number, is in equilibrium. Indeed, the different regions of adaptive space occupied by different clades can be conceptualized as islands with unique dynamic equilibria, albeit with macroevolutionary processes determining the "colonization" of adaptive peaks, as opposed to short timescale biogeographic processes.

The volatility of richness fluctuations will vary across these islands of adaptive space as an emergent trait of a clade. Ultimately volatility emerges from the life histories, ecologies, and evolutionary histories that characterize each clade's occupancy in different regions of an adaptive landscape. We do not attempt to diagnose which characteristics of different regions account for volatility differences, but others have found rates of origination and extinction to depend on larval type (40), body plan (16), body size (35), range size (35, 36), and substrate preference (30). Not all of these traits would be considered dimensions of an ecological niche or characteristics of a guild (33, 34, 41), but they all point to different strategies that influence a clade's macroevolutionary success. These characteristics result from interactions between heritable traits and environments, which themselves may be semi-heritable (42). Thus different regions of adaptive space, and the

5

clades occupying them, will experience different magnitudes of stochastic fluctuations in taxonomic richness. As clades occasionally split to fill new regions of adaptive space their pulsed diversification determines the non-equilibrium nature of the entire biota.

1.2 Real paleontological data to test superstatistics

To uncover the superstatistical nature of the marine invertebrate Phanerozoic fauna we analyzed the distribution of fluctuations in genus richness (the lowest reliably recorded taxonomic resolution) using the Paleobiology Database (PBDB; `paleobiodb.org`). We corrected these raw data for incomplete sampling and bias using a new approach described in the methods section. Occurrences from the PBDB were matched to 49 standard time bins all of approximately 11MY duration following previous publications (5, 13). Fluctuations in genus richness were calculated as the simple difference between bias-corrected richnesses in adjacent time bins.

To focus attention on the variance of fluctuations we zero-centered each clade's fluctuation distribution. In this way we focus on fluctuations about any possible trend toward net diversification or extinction. Because "equilibrium" in the statistical mechanical sense means a system undergoes coherent, concerted responses to perturbation, the mean trend line (positive or negative) is of less interest than deviations from it. We also note that the distributions of fluctuations for most clades are already very close to a mean of 0 (mean of fluctuation distributions at the family level: 0.038 ± 0.176 SD), and so centering has little influence on clade-specific fluctuation distributions, consistent with the observation that origination is often roughly equal to extinction (43).

We define potentially equilibrial sub-systems based on taxonomic hierarchies as a full phylogenetic hypothesis for all marine invertebrates is lacking. Taxa ideally represent groups of organisms that descend from a common ancestor and share similar ecologically

6

and evolutionary relevant traits (*44–46*). Thus our model assumes that at a given higher taxonomic level, within-taxon fluctuations in richness are driven by equilibrial processes characterized by Gaussian distributions. We further assume that new higher taxa arise due to the emergence of sufficiently novel traits (be they ecological, morphological, life history, or macroecological) so that those new taxa occupy a new region of an adaptive landscape. We lastly assume that different regions of adaptive space are characterized by different volatilities in origination and extinction.

To evaluate the optimal taxonomic level for sub-system designation, we test our super-statistical theory using taxonomic levels from family to phylum. Additionally, we compare our results to randomized taxonomies and confirm that the observed fit of superstatistical theory is not an artifact of arbitrary classification but instead represents real, biologically relevant diversification processes within and between clades. We find that families and orders conform to the assumptions of our superstatistical model while classes and phyla do not.

2 Results

We first evaluate the local equilibria of clades from family level to phylum. We find that family level fluctuation distributions are well approximated by Gaussians (Figs. 1 and S3). Three exemplar family-level dynamics are highlighted in Figure 1 to illustrate how different volatility equilibria express themselves as actual richness timeseries. This Gaussian approximation also largely holds for orders, but classes and phyla increasingly show deviations from Gaussian with greater kurtosis corresponding to more frequent outliers at these taxonomic levels.

To predict the superstatistical behavior of the entire marine invertebrate Phanerozoic fauna we must integrate over all possible local equilibria that each clade could experience.
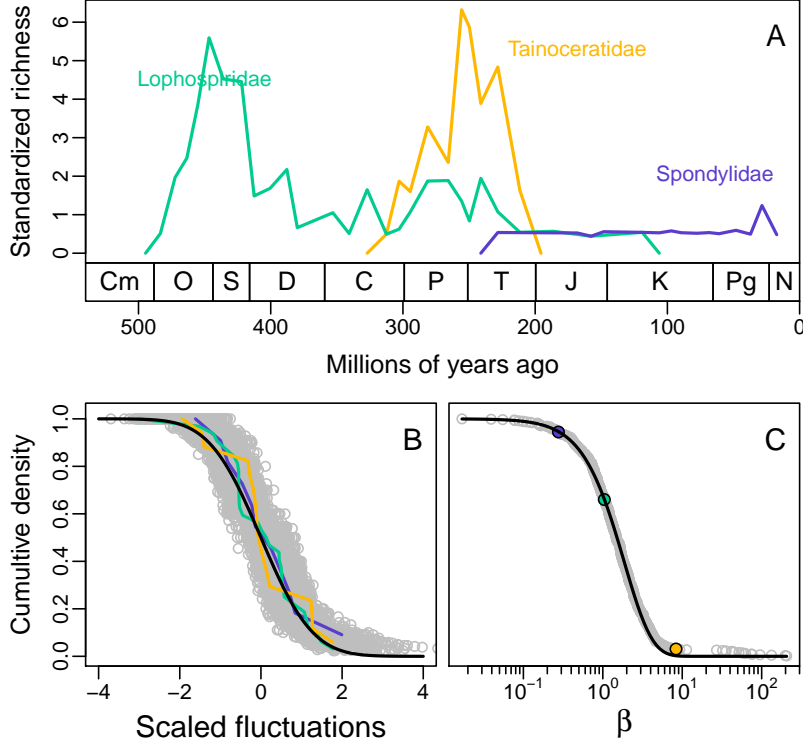
Figure 1: The distributions of within-family fluctuations in genus richness shown for the trajectories of three exemplar families (A) and shown as an empirical cumulative density aggregated across all families (B). To display all families simultaneously we simply collapse their fluctuation distributions by dividing by their standard deviations. If families conform to the Gaussian hypothesis their scaled fluctuations should fall along the cumulative density line of a normal N(0, 1) distribution, as shown in (B). We further confirm this normal distribution in the supplement (Fig. S3). In (C) the distribution of inverse variances β_k across all families matches very closely to a Gamma distribution (black line); exemplar families are again highlighted.

The stationary distribution of β_k values describes these possible equilibria, specifying the probability that a given clade, chosen at random, will occupy a region of adaptive space characterized by β_k.

We estimate the distribution of β_k's simply as the maximum likelihood distribution describing the set of volatilities for all families, orders, classes, or phyla. Phanerozoic marine invertebrate families clearly follow a Gamma distribution in their β_k values (Fig.

1). This Gamma shape also holds for orders but shows increasing deviations again for classes and especially phyla (Fig. S4).

Using the observation of within family statistical equilibrium and Gamma-distributed β_k parameters we can calculate, without further adjusting free parameters, the distributions of family-level fluctuations for the entire marine Phanerozoic, $P(x)$, as

$$P(x) = \int_0^\infty p_k(x \mid \beta) f(\beta) d\beta \tag{1}$$

where $p_k(x \mid \beta) = \sqrt{\frac{\beta}{2\pi}} e^{-\frac{\beta x^2}{2}}$ is the distribution of fluctuations within a family and $f(\beta) = \frac{1}{\Gamma(b_1/2)} \left(\frac{b_1}{2b_0}\right)^{b_1/2} \beta^{(b_1/2)-1} \exp\left(-\frac{b_1\beta}{2b_0}\right)$ is the stationary distribution of volatilities in richness fluctuations. The integral in (1) leads to

$$P(x) = \frac{\Gamma\left(\frac{b_1+1}{2}\right)}{\Gamma\left(\frac{b_1}{2}\right)} \sqrt{\frac{b_0}{\pi b_1}} \left(1 + \frac{b_0 x^2}{b_1}\right)^{-\frac{b_1+1}{2}} \tag{2}$$

This corresponds to a non-Gaussian, fat-tailed prediction for $P(x)$ which closely matches aggregated family level fluctuations in the bias-corrected PBDB (Fig. 2).

To quantitatively evaluate how well the superstatistical prediction matches the family-level data we constructed a 95% confidence envelope from bootstrapped maximum likelihood estimates for $P(x)$. Observed fluctuations fall within this 95% confidence envelope (Fig. 2), indicating that the data do not reject the superstatistical prediction. For further comparison, we fit a Gaussian distribution to the observed fluctuations, which corresponds to the equilibrium hypothesis that all families conform to the same dynamic. Using Akaike Information Criterion (AIC) we find that observed fluctuations are considerably better explained by the superstatistical prediction than by the Gaussian hypothesis (ΔAIC = 1895.622). Thus, as expected under the superstatistical hypothesis, the fat-tailed distribution of fluctuations arise from the superposition of independent Gaussian statistics of fluctuations within families.
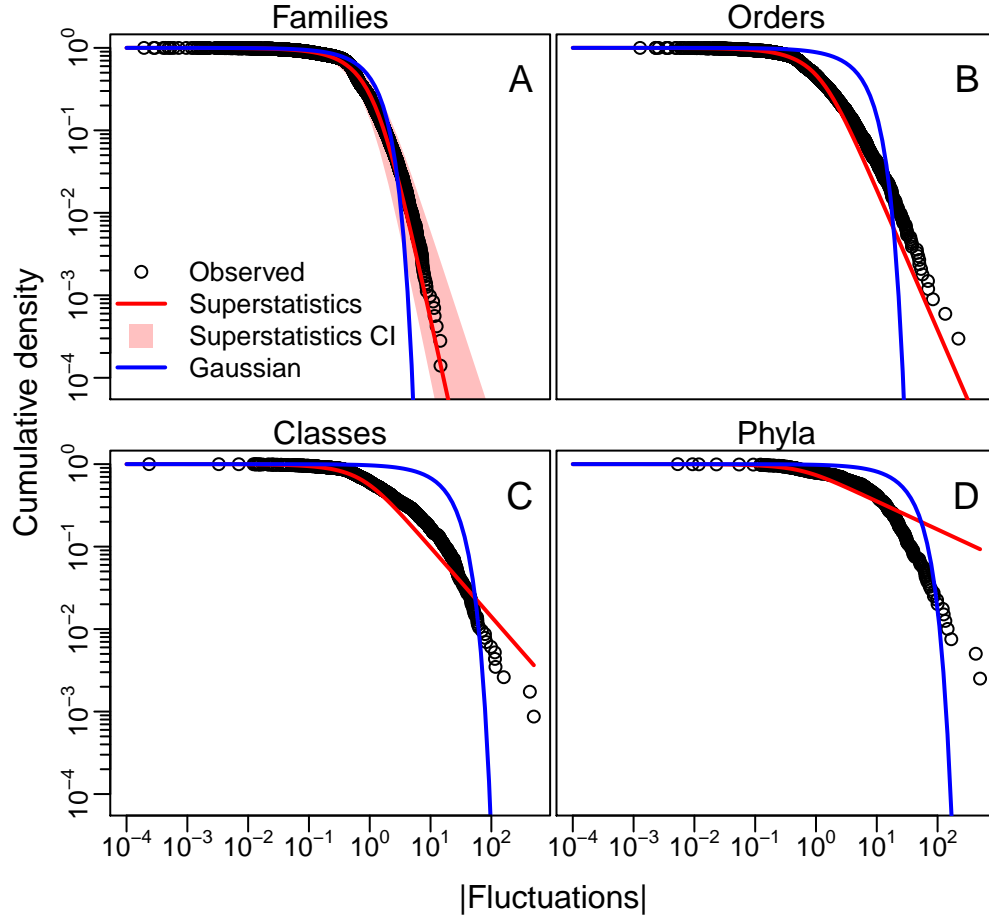
9

Figure 2: Distribution of fluctuations in genus richness within different taxonomic groupings of marine invertebrates in the Paleobiology Database (5) after sampling correction. The distribution is fat-tailed as compared to the maximum likelihood estimate of the normal distribution (blue line). At the family and order level the empirical distribution of fluctuations are well described by our superstatistical approach, both when computed from integrating over the distribution of observed variances (red line) and when fit via maximum likelihood (95% confidence interval; red shading).

Computing the distribution of aggregated fluctuations using orders also closely matches the observed data (Fig. 2) but as we further coarsen the taxonomy to classes and phyla we see increasingly poorer correspondence between data and theory (Fig. 2). We quantify this change in the goodness of fit with the Kolmogorov-Smirnov statistic (Fig. 3). We

can see that both families and orders have low Kolmogorov-Smirnov statistics, and in fact order level designation of equilibrial subsystems performs slightly better than the family level. Classes are substantially worse and phyla worse yet with the Kolmogorov-Smirnov statistic of phyla being no different than the null randomized taxonomies described below.

However, if superstatistical theory explains the data, this worsening fit with increasing taxonomic scale is expected as the different classes and phyla should not represent dynamically equilibrial sub-systems in their fluctuation dynamics. Instead, classes and phyla aggregate increasingly disparate groups of organisms, and thus effectively mix their associated Gaussian fluctuations, meaning that one statistic should no longer be sufficient to describe class-level dynamics. We see this confirmed by the increasing frequency of outlier fluctuations in within class and phylum level fluctuation distributions (Fig. S3). We can also see that families and orders represent, on average, 1 to 2 ecospace hypercubes (*?, defined by taxon environment, motility, life habit, vision, diet, reproduction, and ontogeny*)]bambach1983, bambach2007, bush2007, respectively. In contrast, classes and phyla represent, on average, 8 to 30 hypercubes, respectively (Fig. S5).

Our analysis indicates that orders and families are evolutionarily coherent units with all subsumed taxa sharing key ecological and evolutionary attributes allowing them to reach steady state diversification independently from other clades at global scale. The fact that both orders and families conform to theoretical predictions is consistent with superstatistics. If superstatistics operates at the order level, then the families subsumed by these orders should represent random realizations of their order's stationary $\beta_k^{(order)}$ volatility. The sum of Gamma random variables is still Gamma, but with new parameters, thus the family level distribution of $\beta_k^{(family)}$ is still Gamma.

To further test the evolutionary coherence of families we conducted a permutation experiment in which genera were randomly reassigned to families while maintaining the
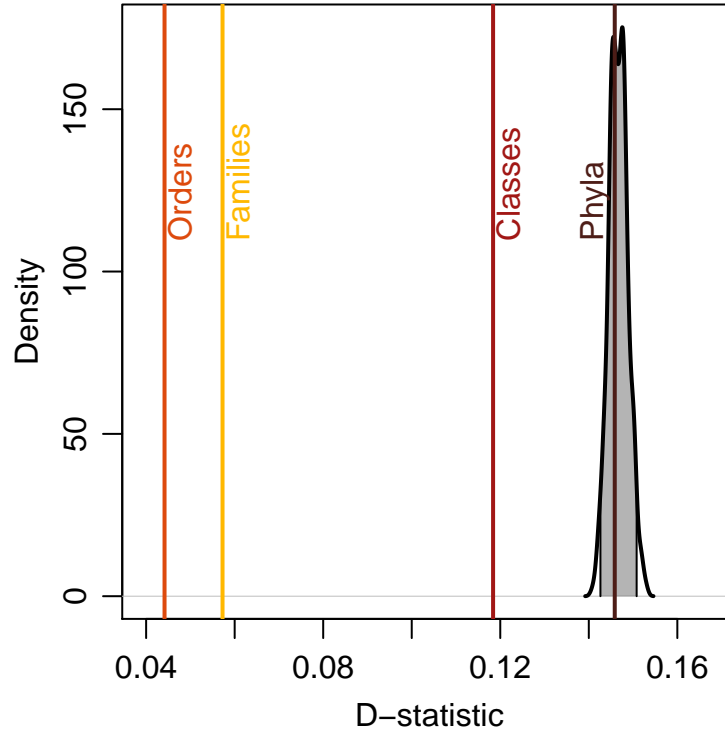
11

Figure 3: Distribution of Kolmogorov-Smirnov (KS) statistics from randomly permuting genera within familes (gray shading represents 95% confidence interval). Solid colored lines are observed KS statistics at different taxonomic levels as indicated.

number of genera in each family. For each permutation, we calculated the superstatistical prediction and its Kolmogorov-Smirnov statistic. The permutation simulates a null model in which common evolutionary history is stripped away (genera are placed in random families) but the total number of observed genera per family is held constant. Controlling for the total number of genera per family is key because this could be purely an artifact of an arbitrary taxonomic process (47–50) and could also be solely responsible for differences in the β_k across clades. Indeed, the number of genera in a family and that families β_k value are correlated (Fig. S6). Thus we want to know if this correlation alone accounts for all downstream superstatistical results.

Repeating the null permutation of genera in families 500 times yields a null distribution

of Kolmogorov-Smirnov statistics that is far separated from the observed values at the family and order levels (Fig. 3) suggesting that the good fit at these levels is not merely a statistical artifact of classification or the richness of clades, but carries important biological information. It should also be noted that the width of 95% confidence interval of this null distribution is not far from the distance between the Kolmogorov-Smirnov statistics of orders versus families, suggesting that differences of fit between these taxonomic levels is at least partially accounted for by the randomness of the sampling distribution of Kolmogorov-Smirnov statistics.

3   Discussion

Our analysis makes no assumption that orders and families should correspond to super-statistical subsystems, but identifies them as the appropriate level for marine invertebrates. Our study is the first to demonstrate that complex patterns in the fluctuation of taxon richness in the fossil record are the result of a simple underlying process analogous to the statistical mechanisms by which complexity emerges in large, non-equilibrium physical (27) and social systems (28). We do so by identifying the biological scale at which clades conform to locally independent dynamic equilibria in fluctuations. Equilibrium could result from many processes, including neutrality (39, 51, 52), diversity-dependence (14, 53, 54) and processes that dampen—rather than exacerbate—fluctuations in complex ecological networks (55). These candidate processes are directly opposed to the presumption of instability underlying the self-organized criticality hypothesis for paleo biodiversity (9, 10).

We show that the distribution describing the evolution to different equilibria between orders is Gamma (Fig. 1). A Gamma distribution, while consistent with multiple processes (e.g., (56)), could result from evolution of diversification rates across an adaptive

landscape that promotes niche conservatism and pulsed exploration of niche space. Specifically, if β_k values are associated with a clade's macroevolutionarily-relevant traits, and those traits evolve via Ornstein-Uhlenbeck-like exploration of an adaptive landscape, the resulting stationary distribution of β_k will be Gamma ($56, 57$). For macroevolutionary rates to vary across an adaptive landscape, this landscape cannot be flat, and thus niche conservatism around local optima in adaptive space interrupted by adaptive exploration is likely ($32, 58$). The specifics of how this adaptive landscape is shaped and is traversed by evolving clades determine the exact form of the distribution of β_k volatilities, in the case of the marine Phanerozoic resulting in a Gamma distribution. Our work thus motivates further study of the trait spaces and evolutionary shifts consistent with Gamma-distributed equilibria in richness fluctuation volatilities.

We show that the pulsed shift to different equilibria between orders and the families they subsume is sufficient to explain the characteristically fat-tailed distribution of richness fluctuations when the marine Phanerozoic fauna is viewed as a whole macrosystem. Armed with an understanding of the statistical origin of this diversification pattern we can explore which models of niche conservatism and pulsed adaptive radiation are consistent with the statistical behavior of the Phanerozoic. Our statistical theory provides new motivation for identifying the eco-evolutionary causes of innovations between lineages and how those innovations are eventually conserved within lineages. Using the superstatistical prediction as a theoretical baseline, we can also go on to identify and robustly examine the mechanisms underlying deviations unexplained by our statistical theory. For example, some clades wax and wane systematically, and possibly non-symmetrically, through time ($4, 8, 36$), a pattern that we cannot explain with superstatistics alone.

Superstatistics could also be applied to other areas of evolution and macroecology. For example new phylogenetic models already consider heterogeneous rates of diversification

(*?, e.g.,*)]rabosky2014 as expected between different subsystems. The superstatistics of clades in adaptive landscapes could motivated models that jointly predict changes in trait and diversification, a research area currently struggling with model inadequacy (*59*). This framework could also provide a new paradigm in modeling the distributions of richness, abundance and resource use in non-neutral communities which can be viewed as emerging from the combination of locally equilibrium subsystems. Non-neutral models in ecology are criticized for their over-parameterization (*60*), yet a persistent counter argument to neutral theory (*51*) is the unrealistic assumption of ecological equivalency (*61*) and poor prediction of real dynamics (*62*). If ecosystems are viewed as the superposition of many individualistically evolving clades, each exploiting the environment differently and thus obeying a different set of statistics, then diversity dynamics could be parsimoniously predicted with superstatistics while incorporating real biological information on ecological differences between taxa.

Superstatistics is a powerful tool to derive macro-scale predictions from locally fluctu-ating sub-systems whose evolution is driven by interesting, but complex and difficult to model, biological mechanisms. As such, applications of superstatistics to a wide variety of patterns in ecological and evolutionary systems are ripe for exploration.

4 Methods and Materials

All data processing and analyses were preformed in R (*63*) and all code needed to repro-duce our study are provided, with added explanation, in supplemental Appendix A.

4.1 Paleobiology Database data download and filtering

Data on individual fossil occurrences and the ecospace characteristics of Phanerozoic ma-rine invertabrates were downloaded from the Paleobiology Database (PBDB; `https://paleobiodb.org`)

on 16 November 2018 via the database's API (data retrieval and processing script in the supplement). Collections were filtered using the same approach as Alroy (5) to insure that only well preserved marine invertebrate occurrences were used in subsequent analyses. This filtering resulted in 815,222 unique genus-level occurrences. These were further filtered to exclude those occurrences without family-level taxonomy and those collections with age estimate resolutions outside the 11MY time bins proposed by Alroy (5) resulting in 454,033 occurrences. These time bins were compiled from `http://fossilworks.org` with a custom script reproduced in the supplement. The first and last of these time bins, corresponding to the earliest Cambrian and the latest Cenozoic, were excluded from analysis because their sampling completeness (see below) could not be assessed.

4.2 Correcting for imperfect and potentially biased sampling

We use a new and flexible method to correct for known sampling incompleteness and biases in publication-based specimen databases (5, 13). Incompleteness is inherent in all biodiversity samples, the fossil record being no exception (64–67). This incompleteness is only an issue if it varies across the fossil record, as indeed it does (64–67). In addition to variable incompleteness, bias may result from preferential publication of novel taxa (13) which exacerbates the difference between poorly-sampled and well-sampled time periods. We therefore develop a simple two-step method: we first correct genus richness for incomplete sampling using the "three-timer" correction (5) and then further correct this three-timer richness estimate by accounting for any correlation between the number of genera and the number of publications in a time period.

The three-timer correction estimates the probability of failure to observe a genus in a given time period p_t as the number of times any genus is recorded before and after that period but not during, divided by the number of genera whose occurrence histories span

16

the period t. To calculate the sampling-corrected richness $\hat{D}_{kt}$ of a clade k in the time period in question, the observed genera within that clade and time period are divided by $1 - p_t$ and their occurrences summed:

$$\hat{D}_{kt} = \sum_{j \in k} \frac{I_{jt}}{1 - p_t} \tag{3}$$

where $j \in k$ designates genera in clade k and I_{jt} is an indicator equal to 1 if genus j occurs in time period t.

$\hat{D}_{kt}$ is the maximum likelihood estimator of richness in a simple occupancy through time type model assuming binomial sampling (*68*), and in that way mimics other proposed methods for the fossil record (*65, 66*). We avoid parametrically modeling the sampling process through time by instead taking a sliding window of time bins from the Cambrian to the Cenozoic. It should be noted that the three-timer correction compares favorably to other similar methods to account for imperfect detection (*69*)

To eliminate further bias due to preferential publication of novel taxa (*13*) we divide the three-timer-corrected number of genera per family per time period by the expected number of genera given publications in that time period. The expected number is calculated by regressing the log-transformed three-timer-corrected number of genera on log-transformed number of publications. There is only a weak trend toward higher richness with more publications (Fig. S1) meaning that the most important correction comes from the three timer correction.

Our new method re-scales each genus occurrence from 0 or 1 (absent or present) to a weighted number continuously ranging between 0 and 1. Because these weighted numbers represent sampling and bias-corrected *occurrences* we can add them arbitrarily, corresponding to the membership of any given genus in any given higher taxonomic group. We must, however, choose a taxonomic level at which to evaluate the relationship between

17

richness and publications; we choose the level of family because this is the most finely resolved option.

We opt not to use subsampling methods ($13, 64, 70$) because these approaches would not be advisable for clades with few genera. However, our new method achieves similar results at the global scale across all clades to subsampling procedures. We directly compare our predicted time series of global fluctuations in genus richness with results derived from rarifaction and shareholder quorum subsampling (SQS) in Figure S2. Our method shows very minor differences with these subsampling-based predictions and any discrepancies do not impact the statistical distribution of fluctuations (Fig. S2).

4.3   Superstatistical methods

We first derive the superstatistical distribution $P(x)$ by fitting Gaussian distributions to clade-level distributions of fluctuations $p_k(x)$, extracting the inverse variances β_k of those $p_k(x)$, testing the best function to describe the distribution of β_k, and then integrating $P(x) = \int_\beta p_k(x|\beta) f(\beta)$. This process allows no free parameters to hone the fit of $P(x)$ to the data. However, each inverse variance must of course be estimated for each clade, making its good fit to data all the more surprising. To do so we use least squares instead of maximum likelihood because the asymmetric fluctuation distributions of small clades were more reliably fit with curve fitting than with maximum likelihood.

We then estimated $P(x)$ directly from the raw data using maximum likelihood to compare the fit of our superstatistical prediction and that of a simple Gaussian distribution using AIC. To calculate a likelihood-based confidence interval on our prediction we bootstrapped the data, subsampling fluctuations with replacement from all families and fit superstatistics using maximum likelihood to the aggregated fluctuation distribution of each bootstrap replicate.

18

References

1. D. M. Raup, J. J. Sepkoski Jr, *et al.*, *Science* **215**, 1501 (1982).

2. J. J. Sepkoski, *Paleobiology* **10**, 246 (1984).

3. N. L. Gilinsky, *Paleobiology* pp. 445–458 (1994).

4. L. H. Liow, N. C. Stenseth, *Proceedings of the Royal Society B: Biological Sciences* **274**, 2745 (2007).

5. J. Alroy, *et al.*, *Science* **321**, 97 (2008).

6. M. Benton, *Science* **268**, 52 (1995).

7. D. H. Erwin, *Trends in Ecology and Evolution* **13**, 344 (1998).

8. T. B. Quental, C. R. Marshall, *Science* (2013).

9. P. Bak, K. Sneppen, *Phys. Rev. Lett.* **71**, 4083 (1993).

10. R. V. Solé, S. C. Manrubia, M. Benton, P. Bak, *Nature* **388**, 764 (1997).

11. M. E. J. Newman, B. W. Roberts, *Proceedings of the Royal Society of London B* **260**, 31 (1995).

12. J. W. Kirchner, A. Weil, *Nature* **395**, 337 (1998).

13. J. Alroy, *Science* **329**, 1191 (2010).

14. D. L. Rabosky, *Ecology Letters* **12**, 735 (2009).

15. C. R. Marshall, T. B. Quental, *Phil. Trans. R. Soc. B* **371**, 20150217 (2016).

16. D. H. Erwin, *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution* **318**, 460 (2012).

17. L. J. Harmon, S. Harrison, *The American Naturalist* **185**, 584 (2015).

18. S. M. Jordan, T. G. Barraclough, J. Rosindell, *Phil. Trans. R. Soc. B* **371**, 20150221 (2016).

19. G. J. Vermeij, *Evolution and Escalation* (Princeton University Press, Princeton, N.J., 1987).

20. J. S. Madin, *et al.*, *Science* **312**, 897 (2006).

21. G. J. Vermeij, *Palaeogeography, Palaeoclimatology, Palaeoecology* **263**, 3 (2008).

22. G. Simpson, *The Major Features of Evolution* pp. 313–337 (1953).

23. E. W. Holman, *Paleobiology* pp. 357–363 (1989).

24. T. Stadler, *Proceedings of the National Academy of Sciences* **108**, 6187 (2011).

25. D. L. Rabosky, *PloS one* **9**, e89543 (2014).

26. C. Beck, E. Cohen, *Physica A: Statistical Mechanics and its Applications* **322**, 267 (2003).

27. C. Beck, *Physica D: Nonlinear Phenomena* **193**, 195 (2004).

28. M. A. Fuentes, A. Gerig, J. Vicente, *PLoS ONE* **4**, e8243 (2009).

29. K. Roy, G. Hunt, D. Jablonski, A. Z. Krug, J. W. Valentine, *Proceedings of the Royal Society B: Biological Sciences* **276**, 1485 (2009).

30. M. J. Hopkins, C. Simpson, W. Kiessling, *Ecology letters* **17**, 314 (2014).

382   31. N. Eldredge, S. J. Gould, *Models in paleobiology* **82**, 115 (1972).

383   32. C. Newman, J. Cohen, C. Kipnis, *Nature* **315**, 400 (1985).

384   33. R. K. Bambach, *Biotic interactions in recent and fossil benthic communities* (Springer,
385   1983), pp. 719–746.

386   34. A. M. Bush, R. K. Bambach, G. M. Daley, *Paleobiology* **33**, 76 (2007).

387   35. P. G. Harnik, *Proceedings of the National Academy of Sciences* **108**, 13594 (2011).

388   36. M. Foote, J. S. Crampton, A. G. Beu, R. A. Cooper, *Paleobiology* **34**, 421 (2008).

389   37. J. Keilson, S. S. Rao, *Journal of Applied Probability* **7**, 699 (1970).

390   38. W. K. Grassmann, *Annals of Operations Research* **8**, 165 (1987).

391   39. R. H. MacArthur, E. O. Wilson, *The theory of island biogeography* (Princeton Uni-
392   versity Press, 1967).

393   40. D. Jablonski, *Annual Review of Ecology, Evolution, and Systematics* **39**, 501 (2008).

394   41. R. K. Bambach, A. M. Bush, D. H. Erwin, *Palaeontology* **50**, 1 (2007).

395   42. F. J. Odling-Smee, K. N. Laland, M. W. Feldman, *Niche construction: the neglected
396   process in evolution* (Princeton university press, 2003).

397   43. M. Foote, *Evolution since Darwin: the first 150 years* pp. 479–510 (2010).

398   44. E. Mayr, *Systematic Zoology* **14**, 73 (1965).

399   45. D. H. Erwin, *Palaeontology* **50**, 57 (2007).

46. T. H. G. Ezard, T. B. Quental, M. J. Benton, *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **371**, 20150216 (2016).

47. G. U. Yule, *Philosophical Transactions of the Royal Society of London Series B* **213**, 21 (1925).

48. B. Berlin, D. E. Breedlove, P. H. Raven, *Science* **154**, 273 (1966).

49. J. Hey, *Genes, categories, and species: The evolutionary and cognitive cause of the species problem* (Oxford University Press, 2001).

50. A. Capocci, G. Caldarelli, *Journal of Physics A: Mathematical and Theoretical* **41**, 224016 (2008).

51. S. P. Hubbell, *The unified neutral theory of biodiversity and biogeography (MPB-32)*, vol. 32 (Princeton University Press, 2001).

52. T. D. Olszewski, D. H. Erwin, *Nature* **428**, 738 (2004).

53. D. Moen, H. Morlon, *Trends in Ecology & Evolution* **29**, 190 (2014).

54. M. Foote, R. A. Cooper, J. S. Crampton, P. M. Sadler, *Proc. R. Soc. B* **285**, 20180122 (2018).

55. E. L. Berlow, *et al.*, *Proceedings of the National Academy of Sciences* **106**, 187 (2009).

56. J. C. Cox, J. E. Ingersoll Jr, S. A. Ross, *Econometrica: Journal of the Econometric Society* pp. 385–407 (1985).

57. M. A. Butler, A. A. King, *The American Naturalist* **164**, 683 (2004).

58. S. Gavrilets, *Fitness landscapes and the origin of species (MPB-41)*, vol. 41 (Princeton University Press, 2004).

22

59. D. L. Rabosky, E. E. Goldberg, *Evolution* **71**, 1432 (2017).

60. J. Rosindell, S. P. Hubbell, R. S. Etienne, *Trends in ecology & evolution* **26**, 340 (2011).

61. J. Chave, *Ecology letters* **7**, 241 (2004).

62. R. E. Ricklefs, *Ecology* **87**, 1424 (2006).

63. R Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria (2018).

64. A. I. Miller, M. Foote, *Paleobiology* **22**, 304 (1996).

65. M. Foote, *Paleobiology* **42**, 707 (2016).

66. J. Starrfelt, L. H. Liow, *Phil. Trans. R. Soc. B* **371**, 20150219 (2016).

67. R. A. Close, S. W. Evers, J. Alroy, R. J. Butler, *Methods in Ecology and Evolution* **9**, 1386 (2018).

68. J. A. Royle, R. M. Dorazio, *Hierarchical modeling and inference in ecology: the analysis of data from populations, metapopulations and communities* (Academic Press, 2008).

69. J. Alroy, *Paleobiology* **40**, 374 (2014).

70. A. T. Kocsis, C. J. Reddin, J. Alroy, W. Kiessling, *bioRxiv* p. 423780 (2018).

Acknowledgments

23

⁴⁵¹ **Data and materials availability:** Data are available through the Paleobiology Database

⁴⁵² (`paleobiodb.org`) and all code needed to interface with the `paleobiodb.org` API,

⁴⁵³ process, clean, and ultimately analyze the data are available online at `github.com/ajrominger/pale`

⁴⁵⁴ This github repository also hosts the exact download from `paleobiodb.org` used in

⁴⁵⁵ this analysis. All required scripts are also available and explained in supplemental

⁴⁵⁶ Appendix A.

Supplementary materials

S1 Limit distribution of a time-averaged homogeneous origination-extinction process

Fossil taxa gain and lose genera according to an origination-extinction process. We assume that most fossil occurrences of a taxon come from the period of its history when it is dominant and in steady state. In a time slice of duration τ during such a period of steady state the latent per capita rates of origination and extinction would be equal (i.e. $\lambda = \mu \equiv \rho$) and the number of origination or extinctions events (call such events Y) each follow an inhomogeneous Poisson process with rate ρN_t where N_t is the number of species or genera in the taxon of interest at time t. Allowing N_t to vary smoothly with time, and recognizing that the sum of Poisson random variables remains Poisson, we arrive at the number Y of extinction *or* origination events in τ being distributed

$$Y \sim \text{Pois}\left(\rho \int_{t=0}^{\tau} N(t)dt\right). \tag{4}$$

Under the steady state assumption we can approximate $N(t)$ by $\bar{N}$, the steady state richness, leading to

$$Y \sim \text{Pois}(\rho \bar{N} \tau). \tag{5}$$

This Poisson distribution is asymptotically Gaussian, which is a more appropriate distribution for our sampling and bias-corrected richness estimates because these estimates are not integer-valued but rather continuous random variables. Furthermore, because we use standard time periods of average duration $\tau = 11\text{MY}$ the the distribution of fluctuations within taxa will be independent of the specific time periods considered. The

25

Gaussian asymptotics of time-averaged birth-death processes have been proven and explored elsewhere as well (*37, 38*).

S2 Evaluation of sampling bias correction methods

Our sampling and bias-correction method first accounts for imperfect detection within a binomial sampling framework as described in the main text, and then further corrects for potential publication bias using simple log-log regression. We reproduce that regression of log-richness versus log-number of publications here (Fig. S1).
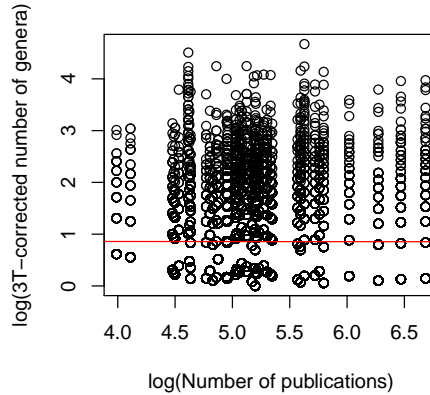
Figure S1: Relationship between number of publications and genus richness as recorded by the PBDB.

We compare our sampling and bias-correction method to other more established approaches. Specifically we use the newly available R package *divDyn* (*70*) to produce subsampling-based richness estimates for the Phanerozoic timeseries of marine invertebrates. In Figure S2 we compare classical rarefaction with shareholder quorum subsampling (SQS) with our method. All samples were rarified to 120 occurrences, which is approximately the maximum sample size across all time bins, and the SQS quorum was

478 set to 0.75 to similarly approximate this common sampling denominator across time bins.

479 For both rarefaction and SQS we produced 50 subsample replicates.
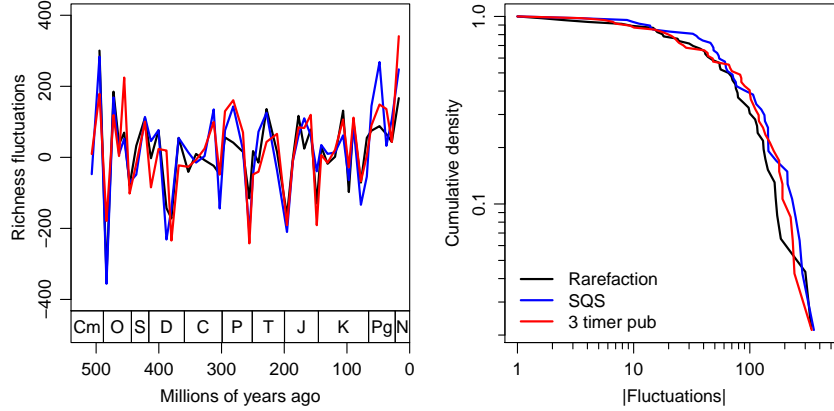

Figure S2: Comparison of SQS method (*13*) (solid black line) with the raw data (dashed black) and our three-timer and publication bias correction method (red). The time-series of all marine invertebrate genera shows general agreement with the only major deviations toward the modern (A). Despite these differences the distribution of fluctuations in genus richness across all marine invertebrates show good agreement (B).

S3   Understanding deviations from superstatistics at higher taxonomic levels

482 To explore why deviations from super statistics increase with increasing taxonomic level

483 we explore how the distributions of richness fluctuations $p_k(x|\beta_k)$ and fluctuation volatil-

484 ities $f(\beta_k)$ change with changing taxonomic level. We find that richness fluctuation dis-

485 tributions experience increasing frequencies of outliers (increasing kurtosis) with higher

486 taxonomic level (Fig. S3). We also find that observed fluctuation volatility distributions

487 increasingly depart from a Gamma distribution at the level of classes and phyla (Fig. S4).
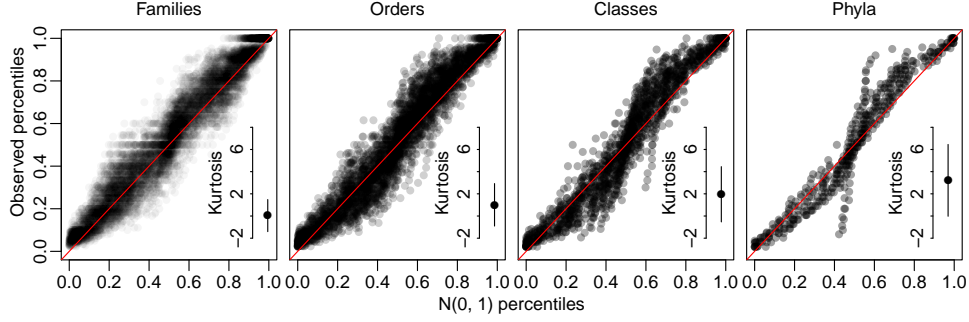
Figure S3: Change in within clade richness fluctuation distributions with increasing taxonomic level. The percentile-percentile plots show how the percentiles of observed re-scaled fluctuation distributions compare to expected percentiles from a Gaussian distribution with mean 0 and variance 1. We can see that families conform to a linear relationship while higher taxa, even at the order level, begin to show s-shaped relationships. Inset plots show how kurtosis increases from 0 (the value for a Gaussian distribution) at the family level to increasingly larger values at higher taxonomic levels.
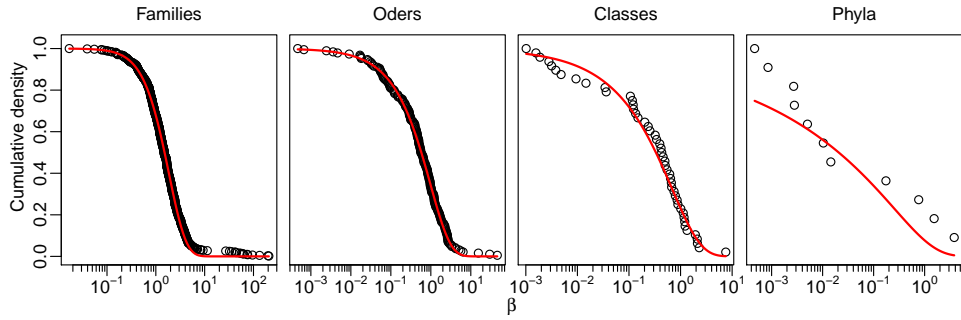

Figure S4: Change in the distributions of β_k across clades with increasing taxonomic level. Points are observed β_k values and red lines are the best-fit Gamma distributions. Deviations increase particularly at the class and phylum levels.

28

S4 Ecospace occupation of higher taxa

We posit that part of the increasing divergence between superstatistics and observed fluctuations and the increase in fluctuation outliers at higher taxonomic levels is that these higher taxa increasingly aggregate disparate types of organisms. One way to evaluate this idea is to count the ecospace hypercubes (*33, 34, 41*) occupied by taxa at different levels. We use the ecological characteristics reported by the PBDB: taxon environment, motility, life habit, vision, diet, reproduction, and ontogeny. In Figure S5 we find that families comprise, on average, 1 hypercude, families comprise 2 hypercubes on average, and classes and phyla comprise many more.
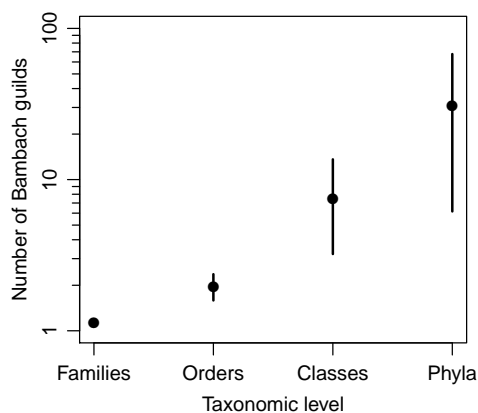

Figure S5: Relationship between number of ecospace hypercubes occupied and taxonomic level.

S5 Relationship between β_k and clade richness

There is likely to be a relationship between richness of clade k and its fluctuation volatility β_k because both extinction and origination (i.e. the formation of new genera) contribute to volatility. Thus we expect that higher variance in richness fluctuations (i.e. smaller $\beta_k = 1/\text{variance}$) will be correlated with higher richness. Indeed, Figure S6 shows this

29

502 to be true. In the main text we use permutation to evaluate whether this correlation is

503 responsible for the observed good fit of superstatistics, and find that this correlation alone
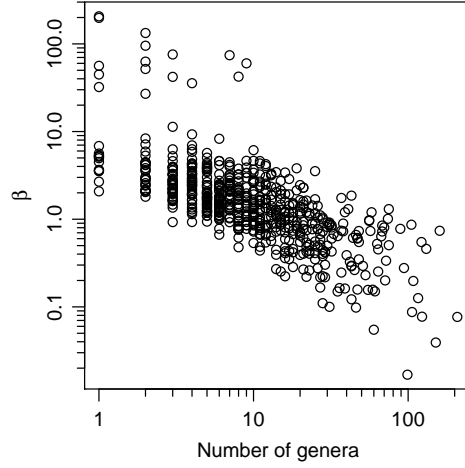
504 is not sufficient.

Figure S6: Relationship between fluctuation volatility β_k and genus richness at the family level.