

Supplementary Methods Material

Supplement to:
Cycles in Fossil Diversity

by
Robert A. Rohde & Richard A. Muller

This supplementary document contains three sections addressing 1) the technical details of converting Sepkoski's stratigraphic data to a uniform time scale, 2) the statistical methods applied in determining the confidence levels stated in the paper, and 3) a demonstration that the conclusions of this paper are upheld even if the data is processed according to the alternative plotting schemes suggested by some authors.

In addition to this document, we have included an Excel spreadsheet which includes tables with the exact data values used in our plots. Data tables appearing in the Excel supplement, will be referenced herein as Supplemental Table E# (where “#” is replaced with the table id number), whereas figures and tables appearing within this document are referenced by S#.

Technical Analysis of the Sepkoski Data

This section describes the qualities of the Sepkoski data and how it was processed to produce the diversity curves appearing in the paper.

The Sepkoski compendium¹ consists of 36380 genera of which 20513 (56%) are specified with identifiers indicating both the first and last stratigraphic intervals in which the genus is known to occur. The remaining 44% of genera are identified as being known from only a single stratigraphic interval. Some of these, but not all, represent genera from which only a single representative has ever been documented. As noted in the text of our paper, most of our analysis is done while excluding the more uncertain single appearance genera.

The 295 stratigraphic identifiers used by Sepkoski have a hierarchical structure such that different identifiers are used when data are resolved to the period, epoch, stage, or substage level. While the periods, epochs and stages correspond to names commonly used for divisions of geologic time, the substages are usually created by dividing a stage into 2 or 3 pieces (lower, middle, upper). Some of these have been further defined in terms of particular biostratigraphic ranges¹. Except where noted in the accompanying Supplemental Table E1, we have assumed that the substages evenly divide each stage.

The substage gradations ultimately divide the Phanerozoic into 174 consecutive disjoint intervals, plus one interval for the upper Vendian. However, a few of the finest divisions (e.g. the lower-middle Miocene and the middle-middle-middle Cambrian) are only referenced a handful of times in the Sepkoski database. As such, we have eliminated a small number of divisions and placed those genera in the next highest level of the hierarchy. This results in a division of geologic time with 167 bins. All higher divisions (stage, epoch, period) are defined as representing an integer number of these bins.

In Table S1, we show what fraction of the genera in each category is specified with each degree of resolution. Certain unconventional terms (e.g. Wolfcampian) have been given a place on the hierarchy based on the number of bins they represent, and these designations are available in Table E1. Also, any stage that was not subdivided has been included under the substage count since it represents only one bin. From Table S1, we note that in total 59% of the data is given at the substage level, and further that 89% of extinctions and 86% of first appearances have been resolved at either the stage or substage level.

| TABLE S1 | Period | Epoch | Stage | Substage |
|------------------------|------------|------------|------------|-------------|
| First Appearance Data | 140 (0.7%) | 1877 (9%) | 7138 (35%) | 11358 (55%) |
| Last Appearance Data | 64 (0.3%) | 939 (5%) | 4899 (24%) | 14611 (71%) |
| Single Appearance Data | 503 (3%) | 2548 (16%) | 5011 (32%) | 7805 (49%) |

This table provides information on the degree of resolution of the stratigraphic references appearing in the Sepkoski database. Its construction and definitions of terms appear in the text of this supplement.

To construct a curve of diversity, it is only necessary to know the change in the diversity when moving from one bin to the next. This change can be described as the number of new genera (originations) occurring in a given bin minus the number of genera from the previous bin that have gone extinct. So in practice, our approach for constructing diversity consists of first calculating the number of originations and extinctions in each bin.

To accomplish this, it is necessary to employ some conventions on the treatment of data having different degrees of resolution, for which we have adapted the methods of Sepkoski and Koch². For data that is known at the stage level (i.e. indicating 2 or 3 bins), we assume that each event, meaning first or last appearance, is evenly distributed among the available substages. This is done in the statistical sense, so that a genus which could have originated in either of two bins is treated as if $\frac{1}{2}$ of a genus originated in each. For most of the analysis presented, we have limited ourselves to genera which are resolved at the stage or substage level so no further manipulation is necessary to generate their diversity curve.

However, for the diversity of all genera, shown in Figure 1A of the paper, and for some of the supplemental plots shown below, it is also necessary to describe a method for allocating data resolved at the epoch or period level. Noting that we have distributed nearly 90% of the data in the way already described, we then extend this to the remaining data according to the “best guess” methodology of Sepkoski and Koch². This assumes that the probability that an event (origination or extinction) occurs in any particular bin within the range of a certain epoch or period is proportional to the number of events of the same type already allocated to that bin. This last procedure is applied sequentially so that all of the epoch level data is resolved before any of the period level data is processed.

We should also note that in the special case of the poorly resolved, single appearance genera there is an additional constraint applied, that of causality. Since for unresolved single appearance data, both the appearance and extinction of that genus is assumed to occur during the same range of bins, we explicitly require that the extinction can not precede the appearance. This constraint limits the number of possible configurations for the purposes of determining the statistical “best guess”. The additional

constraint of causality, though clearly necessary, was never discussed by Sepkoski and Koch². Following this set of procedures allows us to define a number of first and last appearances occurring in each bin, and hence to provide a reasonable construction of the history of diversity while using all of the genera in Sepkoski's compendium.

Having specified a bin structure and a method for allocating data, it is also necessary to associate absolute ages to the bins. For this we rely primarily on the time scale of Gradstein et al.³ (i.e. ICS 2004), which by direct association allows us to assign ages to the boundaries of 134 of the 167 bins. A further 11 bins, from the Cambrian and Permian, are assigned ages by associating the names used by Sepkoski to their nearest analogs on the Gradstein et al. time scale. The last 22 bins, from the Ordovician and Cambrian, correspond to stage names which are no longer used in the formal nomenclature but which were used as recently as the Gradstein and Ogg 1996 time scale⁴. Hence we fix the major boundaries ("Lower Cambrian", "Middle Ordovician", etc) according to Gradstein et al.³, but subdivide these epochs according to Gradstein and Ogg⁴. Having done this we find that the median bin length is 2.75 Myr and they range in duration from 0.65 - 9.45 Myr, excluding the very short Holocene and the very long Upper Vendian. The specific dates assigned to each stage and the method used is documented in the footnotes to Table E1. Note that because of its exceptionally long length, >22 Myr, the Upper Vendian bin is excluded in the analysis discussed in the paper.

Explanation of Statistics Employed

This section details the statistics applied in the paper, and provides a background on Fourier analysis for readers that may be unfamiliar with this type of technique.

Background on Fourier Analysis and Statistics

While the statistics used to determine the confidence levels appearing in the paper are well-established, we realize that they may be unfamiliar to many in the audience for which this paper is intended. As such, this section is provided to both give general background on the relevant statistics and provide precise details of the techniques as they were applied in the interpretation of the Fourier spectrum appearing in the paper (Fig. 1E). Statements appearing in this section are common to a great many introductory discussions of Fourier analysis and rather than repeating the same citations for every statement, we would direct readers to several general references^{5,6,7}.

Any time series can be transformed into its "Fourier" or "frequency-space" representation by means of the Fourier transform. Essentially, the Fourier transform provides a method of expressing the time series as the sum over a collection of sine waves whose sign, amplitude, and phase are uniquely determined by the transform. A Fourier power spectrum (a.k.a. periodogram), such as shown in Figure 1E, expresses the squares of those amplitudes, neglecting phase. (Technically, for a finite time series, the Fourier transform only independently defines the power spectrum at a finite number of frequencies, however, it is common to apply the interpolating technique known as "zero padding" for the purposes of determining the frequencies that maximize the spectral power. This is applied in all of our power spectra, and is explained in many places, for example⁵.)

It is always true that the highest peaks in a power spectrum are those that best “describe” the data. In our case, we would say that the changes in diversity are well represented by a long-term trend, which can be fit to a cubic polynomial, plus fluctuations on time scales of 62 and 140 Myr. The power spectrum provides us this information immediately by simply observing the highest peaks. However, even for entirely random data there will always be some set of frequencies which can be said to describe the data in this way. In this context, the purpose of statistical analysis is to distinguish whether the description of the data derived from Fourier analysis is consistent with or inconsistent with the data having had a random origin (i.e. whether the fluctuations that appear periodic could have occurred randomly versus whether they truly indicate a periodic cause).

There is essentially only one way in which this is done, and that is through the application of models to estimate what are the expected fluctuations in the power spectrum that can result from random processes. Sometimes such models are purely mathematical (e.g. white noise, 1/f noise), but more frequently the models that are used depend on generating simulated data that is analyzed by virtue of Monte Carlo techniques to estimate the expected fluctuations in the power spectrum. However, even when Monte Carlo methods are applied, it is common to use a mathematical simplification that we shall refer to as “the assumption of exponential form”. Specifically, this assumption states that the fluctuations in the power spectrum are assumed to obey:

$P_f(h) = e^{-h/b(f)}$, where $P_f(h)$ is the probability that the power spectrum at frequency f has a height at or above h , and $b(f)$, called the background, is a function to be determined.

This assumption rests on the central limit theorem and the fact that many physical processes can be approximated by Gaussian (i.e. normally distributed) statistics. It can be shown that the exponential form is an exact consequence if the Fourier amplitudes obey Gaussian statistics (note that Gaussian *amplitudes* imply exponential *power*). We observe that this form guarantees that $P_f(0) = 1$, consistent with that fact that the power spectra at every frequency must be greater than or equal to 0, and further, that the 95% significance level (i.e. 0.05 probability) occurs at $h/b \approx 3$.

In practice, the assumption of an exponential form has the virtue of being easily computed from Monte Carlo techniques and allows for a reasonable extrapolation from a distribution of very common events to very rare ones. In particular, for a distribution that exactly obeys an exponential form, one can derive that:

$b(f) = \langle h(f) \rangle$, where $\langle h(f) \rangle$ denotes the average height h of the power spectrum at frequency f where the average is taken across many simulations.

Before taking such an average, it is often (though not always) the case that modelers will rescale the power spectrum so that the average height over all frequencies matches that of the true data (or equivalently rescaling the simulated time series so that its variance matches that in the actual data). This is justified as a way of making simulated data more closely match the properties of the actual data, and essentially limits

the goal of the simulations to creating plausible shapes for the modeled data while removing any effects from the overall scaling. We have applied this normalization in the generation of our backgrounds. However, in acknowledgement that there are situations when this normalization is unwarranted, we also considered the case where such a normalization is not performed and found that such a change would have no qualitative impact on the interpretation of our results.

Before continuing, we should note that the assumption of exponential form is ultimately only an assumption. While most models relevant to spectral analysis will produce distributions that are close to exponential in form, it is useful to verify that this is the case in any given situation, and we shall do so. However, even if the Monte Carlo in question deviates from the exponential form it is often difficult to explain why. Since Monte Carlo modeling depends on our ability to simulate the data, deviations may result either from inadequacies of our modeling or from underlying processes that simply do not obey Gaussian statistics. In particular, it is frequently the case that simulated data may have a limited range of variability that causes the frequency of very rare, high-amplitude events to be underestimated by Monte Carlo. In light of this concern, one usually looks at both the exact distribution occurring in the Monte Carlos and the distribution based on assuming an exponential form and chooses the more conservative of the two in assessing significance levels. This practice was applied to our data and is described below.

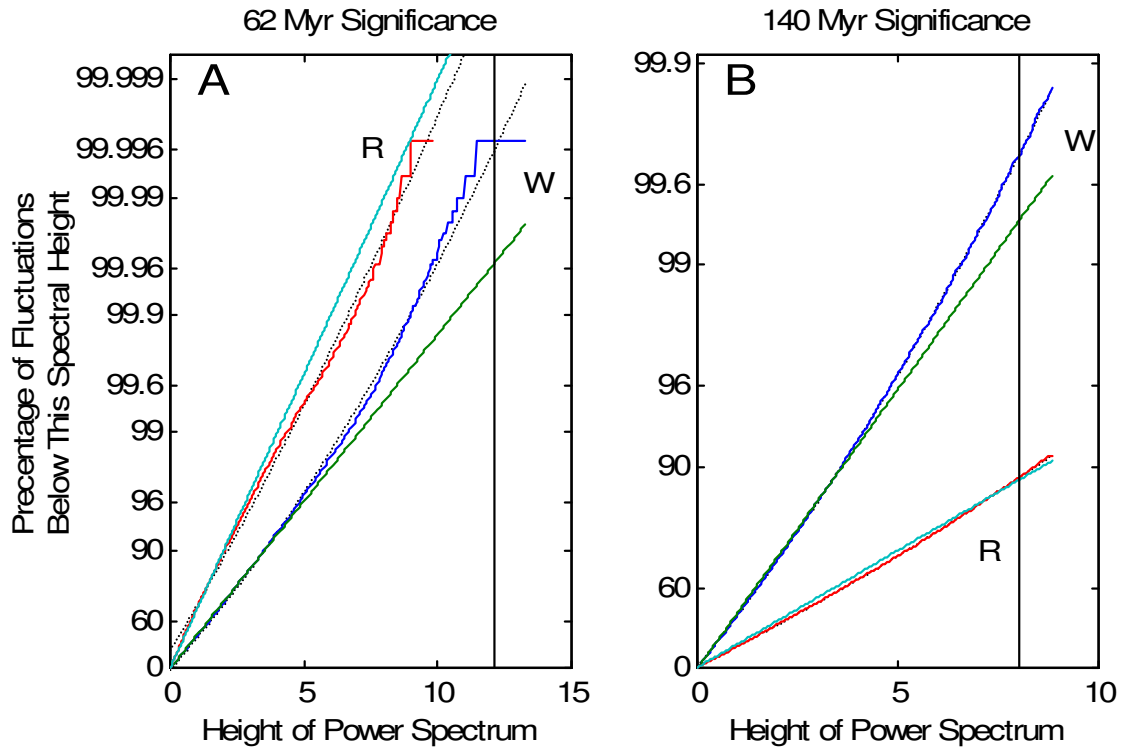
Application of Modeling to the Diversity Data

Choosing an appropriate model for a Monte Carlo simulation is always to some degree a matter of educated guesswork based on the expected behavior of the system. In our case, several authors have proposed that the changes observed in diversity are effectively a random walk^{8,9}. So our “R” model for the changes in diversity is simply to construct random walks by randomly rearranging the steps between bins in the existing data. This has the added advantage that diversity both starts and ends in the same place in every simulation and will generally trend upward. After removing the best fitting cubic trend, the power spectra for these simulations were computed and their average taken to obtain the R background given in the paper. Random walks heavily favor low frequency changes, and this is reflected in the strong trend of the R background.

The second background, referred to as “W”, was constructed by breaking the detrended diversity data into 20 blocks each ~27 Myr in length (though the exact cuts were made at the nearest bin divisions) and then randomly scrambling their order. This preserves short-term correlations but causes the timing of major events to be randomly rearranged with respect to one another. This is a more appropriate background if the long-term changes in apparent diversity are, as some have suggested^{10,11,12,13,14}, controlled by influences other than the random accumulation and loss of species. Admittedly the choice of 20 blocks is somewhat arbitrary, but the height of the background is not strongly affected by the number of blocks used. It should also be noted that a slight bump in the W background (Fig. 1E in main text) at a period of ~27 Myr (frequency ~0.037 cycles / Myr) can be attributed to a small enhancement at that frequency due to the choice of an approximately 27 Myr dividing width.

In Supplemental Figure S1, we show significance versus spectral height based on both the assumption of exponential form and the direct frequency of occurrence in the

Monte Carlo simulations. This figure is plotted on a log-linear scale so that the exponential form is rendered as a straight line. In three of the four cases considered here the deviations from exponential are relatively small, and using the significance level implied by the exponential form renders only a slightly different confidence level than using the Monte Carlo data directly. In the last case, 62 Myr period versus W background, we find rather large deviations, suggesting that at least over some range of frequencies, this scrambling of the diversity data may not produce Fourier amplitudes that follow Gaussian statistics. However, even being conservative and using the exponential form line, the significance is at least 99.96% against the W background (probability 0.00036). We should also note that in 30000 simulations, the R process never produced a spectral feature at 62 Myr capable of matching that which was observed in the data. Given this, we have also been conservative in stating in the paper that the probability of occurrence by R process is <0.00005 . Judging from the trend in Figure S1, it might well be an order of magnitude less, but since the W background is already the more conservative background at the 62 Myr period, we do not feel a need to pin down the R significance exactly, but rather limit our statements to only a bound on that significance.



Supplemental Figure S1: Shows the fluctuations in the power spectrum generated by both the R and W type background at A) 62 Myr period and B) 140 Myr period. In each panel, plotted on a log-linear scale, is the fraction of simulations that produced power spectra which were below the indicated height at the frequency of interest. The straight lines show the expected behavior if the system exhibited fluctuations consistent with the assumption of exponential form described in the text. The solid vertical lines indicate the height of the peaks observed in the actual data. Black, dotted curves, based on the fit of a quadratic polynomial to the data, are used to show how one might estimate the probability of events slightly beyond what was observed.

Table S2 restates the probabilities of these features being observed for both periods and each model as given in Table 1 of the main paper. The first two columns, labeled “at fixed frequency”, reflect only the significance of having found a peak at the particular period indicated, as determined by the methods described above. If the work of Thomson¹⁵⁻¹⁶ and Ager¹⁷ were to be regarded as an *independent* prediction of a 62 Myr cycle than these numbers alone would represent a great confirmation of that prediction. However, our diversity data is not really independent of the diversity data that Thomson and Ager were considering, so we ought to treat these peaks as completely unexpected. That means we should consider more generally the probability of having found comparable peaks anywhere in the spectrum. For example, this means correctly accounting for the fact that in a system of 10 independent peaks we expect roughly 1 to be above the 90% confidence level merely by chance.

| Probability of Peaks | At fixed frequency | | Anywhere in spectrum | |
|----------------------|--------------------|--------|----------------------|-------|
| | R | W | R | W |
| 62 Myr | < 5E-5 | 3.6E-4 | < 0.0013 | 0.010 |
| 140 Myr | 0.12 | 0.0056 | 0.71 | 0.13 |

Table S2: Probability of producing a matching spectral feature through either the R or W background process, as quoted in Table 1 of the main text. These are reported both in terms of finding a matching peak at the same frequency or a “comparable” peak appearing at any frequency in the spectrum, where comparable is as defined in this section. Note that probabilities are expressed as numbers in the range 0-1, so that a probability of 0.01 represents a 1% chance of that event occurring.

To determine the probability of finding peaks like the 62 and 140 Myr signals appearing anywhere in the spectrum, the following calculation was performed. For this purpose, a potential peak is considered “comparable” if it is both a) at least as statistically significant as the one that it is being compared to, and b) it is also biologically significant, meaning it incorporates at least 5% of the variance in the diversity signal (recall that the 62 Myr cycle actually represents 35% of the variance). This second criterion is equivalent to saying the peak’s height is at least 1.73 when compared to the power spectrum in Figure 1E of the paper. Since the median bin length is 2.75 Myr, the shortest period that could be resolved in this data, even in principle, is 5.5 Myr. This implies that there are ~99 independent frequencies in the power spectrum of the diversity data (542 Myr record / 5.5 Myr), which should be regarded as evenly spaced in frequency space. Then the probability of having a comparable peak appear anywhere in the spectrum is

$$P_C = 1 - \prod_{n=1}^{99} P_{NC}\left(\frac{n}{542 \text{ Myr}}\right), \text{ where } P_C \text{ is the probability of observing a comparable peak and } P_{NC}(f) \text{ is the probability that a peak occurring at frequency } f \text{ is not comparable. Here } P_{NC}(f) \text{ implicitly depends on the model being used, either R or W, and the significance of the peak being compared, either 62 or 140 Myr.}$$

Hence we have reduced the problem to determining at every frequency the probability that no comparable peak appears. For a specified background, we can

compute the probability at frequency f of having a spectral power of less than 1.73 (the biologically significant threshold), by looking at both the exponential form assumption and the results of the direct Monte Carlo simulation, and again making the conservative choice. For cases where the direct simulation never reached the biologically significant level, the relevant probability was estimated by extrapolating using a quadratic polynomial (similar to the continuing curves shown in Figure S1A). Because we constrain comparable peaks to those having at least the same statistical confidence as the signal being studied, if the probability of having a peak at the biologically significant level is greater than the probability of observing a peak like that being compared, we substitute the latter probability for P_{NC} . The exact numbers used in this process are recorded in Table E6.

Using this process we determined the probability of finding any comparable peaks to either the 62 or 140 Myr features as a result of either an R or W type process, and these probabilities are also recorded in Table S2. These values indicate that there is not more than a 1% chance of observing a peak like the 62 Myr cycle anywhere in the spectrum. The 140 Myr cycle is more ambiguous. For the W process, a similar peak might well occur but could still be regarded as somewhat unlikely, having only a 13% probability. However, if the behavior of diversity is assumed to be similar to that of a random walk, then the 140 Myr peak should be discounted as similar spectral peaks will often occur by random chance. There is no definitive way to discriminate between these two forms of underlying behavior solely on the basis of the data. Regardless of the ambiguity surrounding the 140 Myr cycle, we can state with confidence that the 62 Myr cycle is a very strong and unexpected feature that needs an explanation.

Discussion of Alternative Plotting Schemes

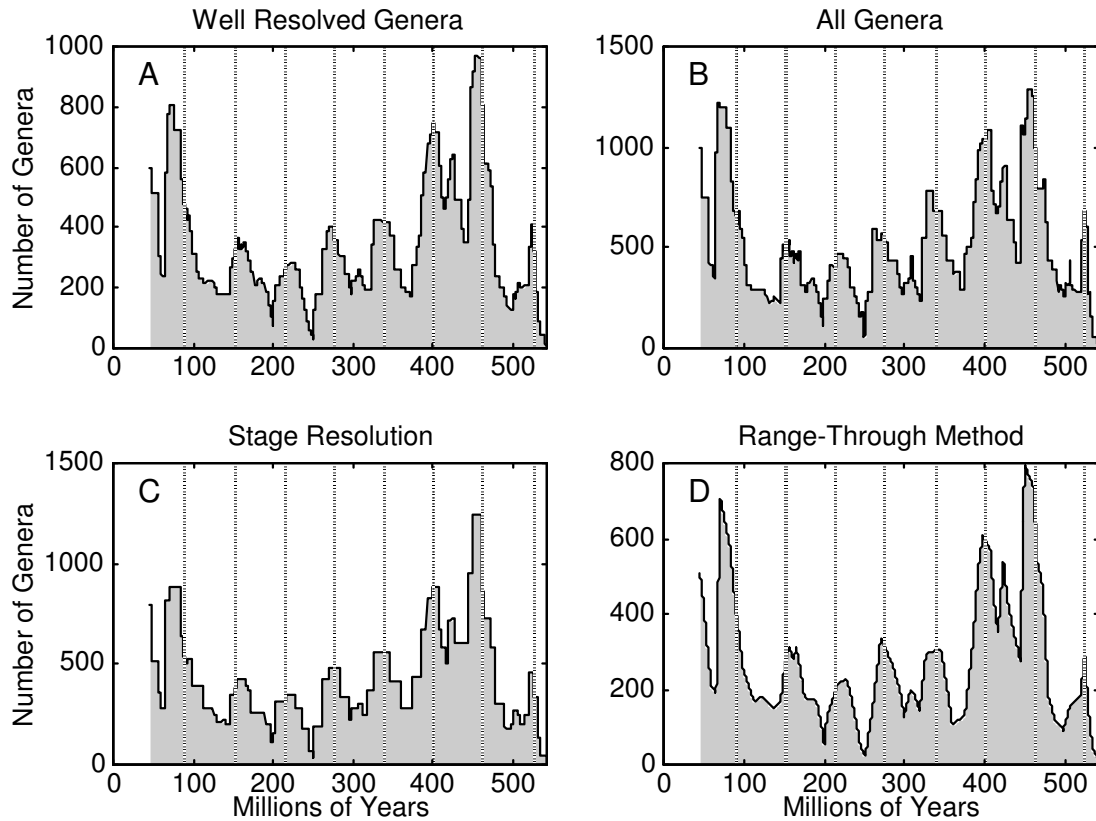
This section demonstrates that our analysis and key conclusions do not depend on the data processing or plotting techniques employed. Specifically, we consider three other ways that the data might have been presented and demonstrate that the central conclusion of a 62 Myr cycle is upheld in all three cases.

While we have chosen a method of analysis that we feel uses the available data to maximal effect, we also acknowledge that a variety of alternative plotting and analysis techniques have been advocated in the literature. It is important therefore that we establish that our observations are not an artifact of the analysis techniques. This is easily done.

In Supplementary Figure S2 we show the diversity of short duration genera, those existing ≤ 45 Myr, according to four different presentations. Panel S2A duplicates the content of Figure 2A in the main text of the paper, i.e. it is limited to those genera for whom both a first and last appearance are specified and each are given to the stage level or better. As discussed in the paper, these data show an easily visible and profound 62 Myr cycle. In Panel S2B shows the same plot except that all genera having a duration ≤ 45 Myr are used, including poorly resolved and single appearance genera. This panel shows that the 62 Myr cycle is actually so strong as to be easily visible without the need to make any cuts to the data besides selecting for duration. In Panel S2C we show the same data as Panel S2A but with the resolution degraded to the stage level. While a

substantial portion of Sepkoski's data is resolved at the substage level, previous studies of diversity have often discarded this information, if it was available at all. Hence it is important that Panel S2C demonstrates that the 62 Myr cycle is well defined even if one is limited to stage resolved data.

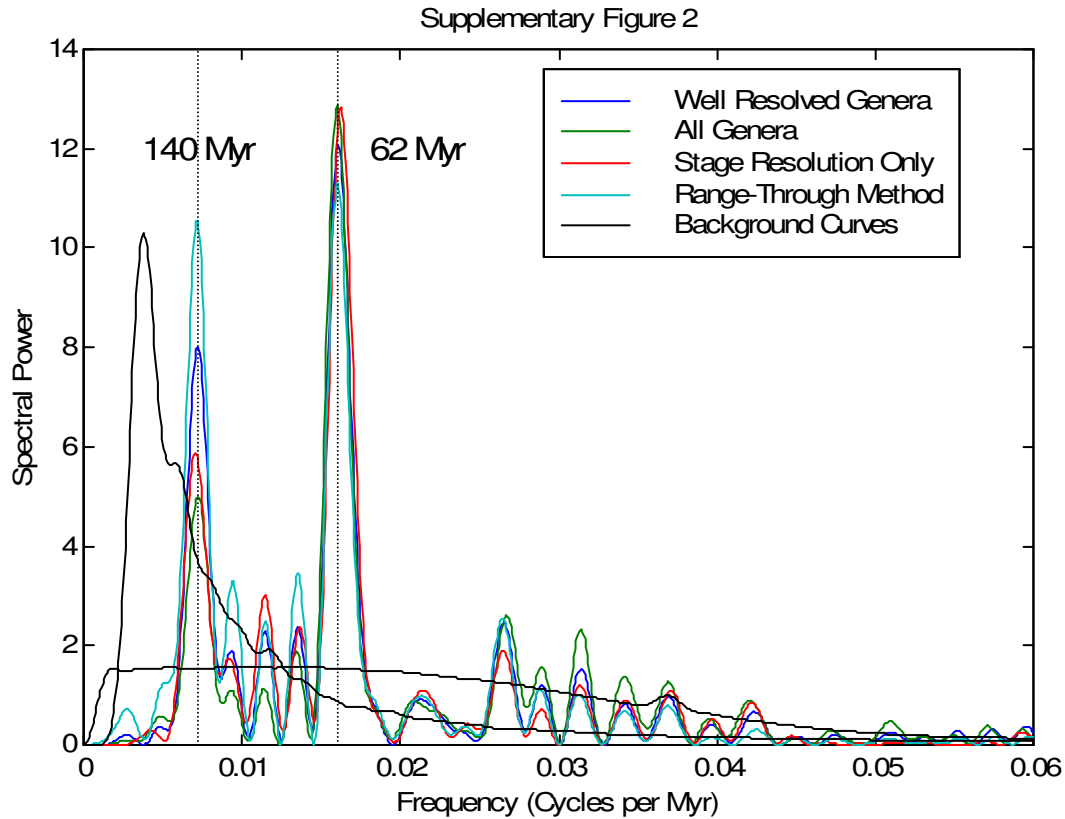
Lastly, Panel S2D presents the same data as Panel S2A except that diversity is calculated according to the “range-through” technique, that is to say that diversity is measured only at the boundaries between stages and a genus is counted as having been alive at that boundary only if its fossil has been found on both sides of the boundary. Because in this plot we are using only genera resolved at the stage or substage level, there is no ambiguity regarding the treatment of poorly resolved data. As is evident on inspection, the “range-through” method also shows a clear 62 Myr cycle in the short duration genera.



Supplementary Figure S2: Shows the diversity of short duration genera, those enduring less than 45 Myr, according to 4 different analysis schemes: A) Includes only genera for which both a first and last appearance are specified and each are recorded at the stage or substage level. This data is the same as Figure 2A in main text. B) Includes all genera with durations less than 45 Myr. C) Same data as A, but downgraded so that all information is processed at the stage level. D) Same data as A, but diversity curve is prepared according to the “range-through” method described above. In each plot, dotted vertical lines indicated the maxima of the 62 Myr cycle as determined from all well-resolved genera and shown in Figure 1 of the main text. This plot makes it evident that the presence of the 62 Myr cycle is not an artifact of the way the data is presented or analyzed.

To also provide a more quantitative analysis of the effects of different plotting techniques we repeat in Supplementary Figure S3 the spectral analysis presented in Figure 1 of the main text using each of the four plotting schemes mentioned above. For

this analysis we produced a curve of total diversity through time according to each of the analysis methods, subtracted the best fitting cubic polynomial, and then took a Fourier transform of the remainders. As already suggested by Supplementary Figure S2, the 62 Myr cycle is a very robust feature of the data whose significance changes nearly not at all with different presentations of the data.



Supplementary Figure S3: As a complement to the Fourier spectra shown in Figure 1 of the main text, this figure shows how that spectra would change if the data had been presented according to the alternative analysis schemes discussed in this section. The black background curves are the same as presented in the main text. In principle the backgrounds also change depending on the presentation of the data, but in practice such changes are very minor on the scale of this figure. As is evident here, the 62 Myr cycle is very insensitive to how the genera data is analyzed and presented. By contrast, the 140 Myr cycle is sensitive to such changes in data presentation, the significance of which is discussed below.

Somewhat surprisingly, the 140 Myr cycle is less robust. While the 140 Myr spectral feature is always the second strongest in the data and always above background, its amplitude varies considerably depending on how the data is processed. Based on the more conservative random walk background we would interpret its apparent significance as varying from a mere 74% (probability of occurrence at this frequency = 0.26) when all genera are used to a high of 94% when the “range-through” method is used. While some degradation of the signal might be expected when less precise or less well-resolved data is used, the magnitude of the changes are surprising and indicate a dependency on the data presentation that requires a further measure of caution in interpreting the 140 Myr cycle. However, given that we have an *a priori* preference for using the well-resolved data and noting the coincidence between this period and the previously reported changes

in climate, we would not rule out that the 140 Myr cycle is also a reflection of true biological behavior, though in this case further research is clearly needed to establish such a relationship with certainty.

The most important result of this supplementary section is that the 62 Myr cycle is robust to various analysis techniques and almost certainly represents a true cyclic behavior affecting the fossil record. It should be emphasized that observing a 62 Myr cycle in fossil diversity is a necessary and unavoidable consequence of combining the Sepkoski compendium¹ and the ICS2004 geologic time scale³.

¹ Sepkoski, J. John. A Compendium on Fossil Marine Animal Genera. Eds. David Jablonski and Michael Foote. *Bulletins of American Paleontology* 363. 2002.

² Sepkoski Jr., John J. and Carl F. Koch. Evaluating Paleontologic Data Relating to Bio-Events. In: *Global Events and Event Stratigraphy in the Phanerozoic*. Ed. Otto H. Walliser. Springer-Verlag: Berlin. 1996.

³ Gradstein, Felix, Jim Ogg and Alan Smith. *A Geologic Time Scale 2004*. Cambridge University Press. 2004.

⁴ Gradstein, F. and Ogg., J., A Phanerozoic time scale. *Episodes*, v.19, no.1 & 2. 1996.

⁵ Muller, Richard A. and Gordon J. Macdonald. *Ice Ages and Astronomical Causes: Data, Spectral Analysis and Mechanisms*. Springer-Praxis: Chichester, UK. 2000.

⁶ Priestly, Maurice B. *Spectral Analysis and Time Series*, Vol 1. Academic Press: NY. 1981.

⁷ Percival, Donald B. and Andrew T. Walden. *Spectral Analysis for Physical Applications: Multitaper and Conventional Univariate Techniques*. Cambridge University Press. 1993.

⁸ Cornett, James L. and Bruce S. Lieberman, Random walks in the history of life, *Proc. Natl. Acad. Sci.* 101(1), 187–191 (2004).

⁹ Gavrillets, Sergey. Dynamics of clade diversification on the morphological hypercube. *Proc. R. Soc. Lond. B* 266 : 817-824 (1999).

¹⁰ Alroy, J. et al., "Effects of sampling standardization on estimates of Phanerozoic marine diversification," *Proc. Nat. Acad. Sci. (USA)*, 98, 6261-6266 (2001).

¹¹ Peters, E. Shanan, and M. Foote, Biodiversity in the Phanerozoic: a reinterpretation, *Paleobiology* 27, 583-601 (2001).

¹² Crampton et al.. Estimating the Rock Volume Bias in Paleobiodiversity Studies, *Science* 301, 358-360 (2003).

¹³ Smith, A. B. Large-scale heterogeneity of the fossil record: implications for Phanerozoic biodiversity studies. *Phil. Trans. Roy. Soc. B* 356, 351-367 (2001).

¹⁴ Jablonski, David, et al. The Impact of the Pull of the Recent on the History of Marine Diversity, *Science* 300, 1133-1135 (2003).

¹⁵ Thomson, K.S., Explanation of large scale extinctions of lower vertebrates, *Nature* 261, 578-580 (1976).

¹⁶ Thomson, K.S., The Pattern of Diversification Among Fishes in *Patterns of Evolution as Illustrated by the Fossil Record*. Ed. A. Hallam, Developments in Palaeontology and Stratigraphy 5, 377-404 (1977).

¹⁷ Ager D.V., The nature of the fossil record, *Proceedings of the Geologists' Association* 87, 131-159 (1977).