# Supplementary information

## A. Yule statistics

As mentioned in the main text, Yule's model yields the following prediction for the species-per-genus distribution (SGD):

$$n_m = C\,\mathbf{B}(m, 2+\nu) \overset{m>>1}{\sim} m^{-(2+\nu)}$$

where $n_m$ is the number of genera of size m, such that the tail of Yule's SGD is described by a power law. To be more precise, the leading $1/m$ correction may be calculated:

$$n_m = C\,\mathbf{B}(m, 2+\nu) \overset{m>>1}{\sim} m^{-(2+\nu)}\left(1 - \frac{(\nu+2)(\nu+1)}{2m}\right)$$

For any value of $\nu$ between 0 and 1, a 5% deviation from a pure power law appears only for $m$ = 1 or $m$ = 2, and thus the Yule expression is equivalent, for any practical purpose, to a pure power law. Hence, the Yule expression yields a straight line in a log-log plot as emphasized in Figure S1.

## B. Yule's fitting procedure

Yule's model is effectively a one parameter model, the parameter being $\nu$, the probability that a new species will originate a new genus. The coefficient, $C$, is determined by the normalization condition $\sum_m m n_m = N$ where $N$ is the number of species.

Actually, Yule himself was aware of the problem with the model (that it does not fit the shoulder) and presented a variant of his model with two extra parameters in order to achieve a good description of the data. His altered model assumes that at a specific time in the past, the clade (taxonomic group) was composed of $N_0$ genera (not fixed to 1 as in his simple model), each with only one species, i.e. all the genera were monotypes. After this, at every time step, each species has a probability $s$ of 'producing' a new species, and a probability $g$ of producing a new genus. The ratio $v = s/g$ is the second parameter. Obviously, the SGD changes after the first generation because now there are also non-monotypic genera, and with each additional generation, the SGD will continue to change. Thus the SGD

depends on the number of generations, $\tau$, which passed from the cladogenesis (at least after a small number of generations). As a result, the actual model used by Yule to fit his data has three free parameters. In Table IV of his original article, Yule gives the result of the fit for these three parameters. What is now referred to as the Yule model is actually a particular case where $N_0 = 1$, and $\tau$, the time since the origin of the clade, is large enough such that the system has converged to a steady state. Obviously his model is still unrealistic since taxonomic groups start with only one genus containing a single species. Two other problems with his model are that it is a three-parameter model (whereas we use only two parameters), and in addition it does not include the very common event of extinction.

## C. Comparing the quality of the fit of the Yule and SEO models

In this section we will present a detailed comparison of the quality of the fit of the Yule and SEO models we presented in the main text.

The $R^2$ statistics (presented in Table S1) show that the SEO model performs better than the Yule model. However, the $R^2$ for both of them is relatively high, which is due to the fact that data points are relatively smooth.

We used two more measures to test the difference between the Yule model and the SEO model. The first measure is the *F*-statistics for the improvement in the $R^2$ (as presented in the main text), for which the p-value of the difference is very small (p=1.5e-11). This shows that the differences are real and not due to the one extra parameter in the SEO model.

The second measure that we used is the area between the model prediction and the data. Such a measure does not give more weight to a region with a higher density of points as does the $R^2$ measure. In order to calculate the area, a continuum approximation of the data should be done. Here we used the cubic spline smoothing tool as implemented in Matlab with a lattice grid of 0.001 (a smaller grid did not change the area). The area between the different models and the data, along with the unsigned area between the data and a curve representing its average, are given in Table S1 and Figure S1. One can see that the SEO model has a significantly smaller area than the other models.

**Table S1. SEO vs Yule**

| Model | $R^2$ | Area |
|---|---|---|
| Yule | 0.93 | 3.89 |
| SEO | 0.9974 | 0.024 |

**The values of the $R^2$ and the area between the data and the model are given in this table. One can see that in both measures the SEO model is better, but the area is more distinctive.**

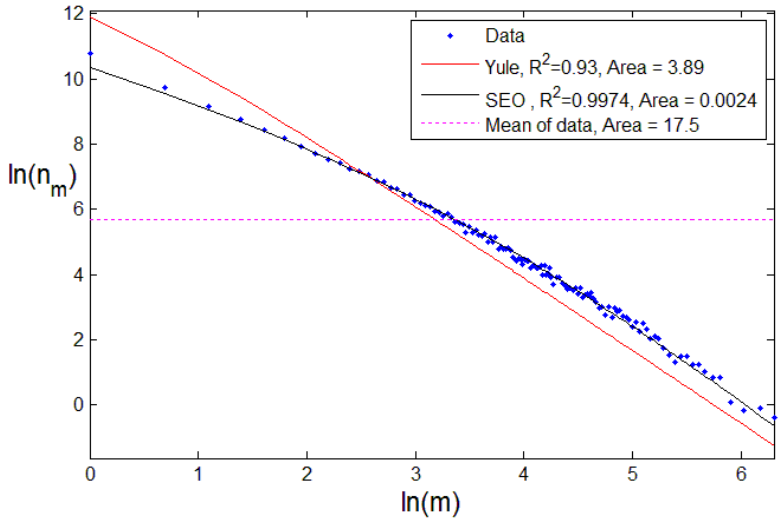


**Figure S1. The SGD of the Animalia kingdom with the best fit of the Yule model, the Yule 2-parameter function, and the SEO model. The magenta dashed line is the average of the SGD data plotted in order to give a calibration to the area between the models and the data.**

## D. Noisy environment and inconsistency in genus definitions

In the following, we present an example of a noisy SEO process. Here "noisy" means that, on top of the stochasticity of the process itself, the parameters (e.g., $\gamma$) are not fixed but take a random value at each generation. This mimics the scenario where the per-species diversification and origination rates are not fixed but fluctuate in time. For example one may imagine exogenous factors that change the number of available ecological niches. As an example, we have used $\gamma_0 = 0.06$ (and $\nu = 0.03$) as average values, while at each generation $\gamma$ is a random number between zero and infinity with probabilities:

$$P(\gamma) = \frac{1}{\gamma_0} e^{-\gamma/\gamma_0}.$$

The simulation starts with a single species and stops when the number of species is 50000 (the order of magnitude of a class). A typical SGD is shown in Figure S2, together with a fit with Eqs. 4 & 5 of the main text (the prediction for SEO without noise). One observes that the fit is not bad. The deviations (as depicted in the data/fit inset) are larger than in the case of a fit to a "pure" simulated SEO process (e.g., Figure 1 of the main text) and resemble the real data. The best fit yields, for this run, $\gamma = 0.054$ and $\nu = 0.033$, not far from the average values. Running the process 100 times and repeating the numerical experiment, we have found that $\gamma$ = 0.049 [0.029 - 0.073] and $\nu$ = 0.028 [0.020 - 0.035], where the brackets indicate two standard deviations above and below the mean.

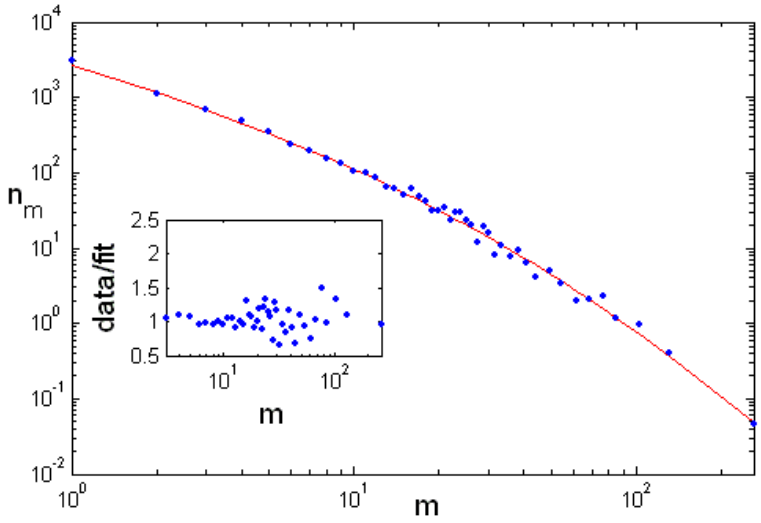


**Figure S2. An example of a noisy growth rate.**

Another aspect of robustness to noise needed to be examined is the robustness to randomness and inconsistency in genus definitions. We tested a few scenarios of inconsistency in the genera definitions, and found that they lead only to a small bias in the inferred parameters from the average parameters, and thus do not influence our ability to infer the growth rate and the origination rate.

The scenarios that we tested are as follows:

1. Every new genus has its own origination rate taken from a Gaussian distribution with mean $\nu$, and a standard deviation of 10 and 20 percent, while the diversification rate is fixed.

2. Every new genus has its own diversification rate taken from a Gaussian distribution with mean $\nu$, and a standard deviation of 10 and 20 percent, while the origination rate is fixed.

3. Both the origination rate and diversification rate are taken from a Gaussian distribution.

4. Same as above, but with a memory of the parent's diversification and origination rates.

In all these cases the fitted values bracket the real values.

In all the above examples, we assumed that there is only noise in the diversification and origination rates, but there is no bias in the behavior of large vs. small genera. We also analyzed the latter scenario, and still found that there was no meaningful change in the distribution of species per genus as a result of such a bias, as long as the bias is not very large.

We assumed that the origination rate of a genus is given by $\nu(m)= \nu_0*(1+\nu_s\log(m))$. We compared cases with different values of $\nu_s$, where $\nu_0$ was determined such that $\nu$ (100) will be fixed between the different values of $\nu_s$. In Figure S3 we present the average SGD (averaged over 50 realizations) for different values of $\nu_s$. One can see that even when there is a difference of about 200% between the large and small genera, the SGD itself is barely affected. This is because on average the population behaves the same.

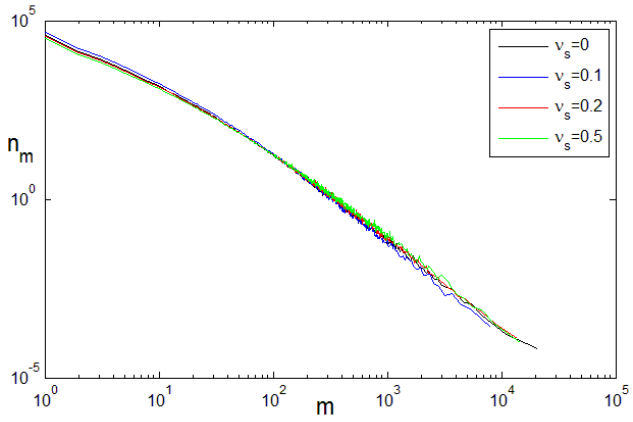


**Figure S3. The number of genera of size $n_m$ vs. $m$, with bias in the origination rate with slopes $\nu_s$ = 0, 0.1, 0.2, 0.5. The rest of the parameters are $N$=1e6, $\gamma$=0.06, $n_0$=0.02. One can see that the differences between the different cases are very minor.**

## E. The simulation procedure

Although the SGD that emerges from the SEO model is known analytically, we have implemented simulations of the SEO process for several reasons. First, we have tried to estimate the accuracy of our inference technique, for which we needed not only the average SGD but also the standard deviation of the result. Second, we have checked how much deviation from the model assumptions (e.g., non-exponential growth) can be detected from the emerging SGD. Finally, we have used simulations to obtain results that are not part of the SGD solution (Eq. 4 and 5 of the main text) e.g. the expected number of monotypes.

Once the rate of growth (homogenous or varying in time) is specified, we have initiated the simulation with one species that belongs to a single genus. If the birth rate is, say, $\lambda$, then at any time-step $\Delta t$, any genus with $s$ species becomes a genus with $s+m$ species, where the integer $s+m$ is obtained from a negative binomial distribution with average $s\alpha$. After the birth step, we have an extinction step where the number of species that went extinct is chosen randomly in the same manner, with an extinction rate $\mu$. At the end, any new species in the genus generates a new genus with fixed probability $\nu$. The process is iterated for each genus and through time (if needed, $\lambda$ and $\mu$ are time dependent) until the population reaches some pre-determined size.

We obtained the size of the largest genus using simulations. This is because the steady state solution [like the one presented in Eq. (4)] does not possess an upper bound. This has to do with the continuum approximation used to obtain this steady state, as discussed in the main text, and with the fact that a true steady-state distribution appears only after infinite time. At finite times, the SGD is very close to the steady-state limit but has an upper cutoff which increases with time. To obtain this cutoff, one needs a simulation of the SEO process with the parameters $\gamma$ and $\nu$ extracted from the bulk of the distribution.

### Bootstrapping procedure

In order to obtain confidence intervals for the inferred parameters, we used a parametric bootstrapping approach. The procedure is as follows: first we performed a fit between the analytical expectation for the genus size distribution to the data, and obtained estimates of the parameters ($\gamma$ and $\nu$). Then, we produced from Monte-Carlo simulations many datasets with the inferred parameters, and *N,* the current number of species. For each of the simulated datasets, we estimated its parameters by the fitting procedure. We used two standard deviations of the inferred parameters from the simulations as estimates for the confidence intervals, or if we had enough simulations, we obtained the 95% confidence interval.

## F. The data from the Furnariidae family reconstructed tree

We used the reconstructed phylogenetic tree of the Furnariidae family, Figure 1 of Ref [1], in order to obtain for each genus its size (i.e., number of species belonging to it), its crown age, and its tree size (the sum of the length of all the clades in the sub-tree containing only the species of each genus and the crown node as the root). The data are presented in Table S2.

In general, there were only a few cases of inconsistent definitions of the different genera (meaning two genera or more mixed with each other in a way that looks like the same genus appeared twice) and they have only a minor effect on the genus crown age and volume. Therefore we ignored the inconsistency and included all the genera. Only with the *Xenops* genus was the inconsistency large, with a significant effect on the quantities obtained for the genus. However, for completeness we included also the *Xenops* genus. The fourth column in Table S2 indicates whether this genus was used to estimate the origination rate, or not.

### Estimation of the origination rate from the phylogenetic tree of the Furnariidae family

By using the phylogenetic tree of the Furnariidae family of the class Aves, we can estimate the origination rate per million years.

This was done as follows: With every speciation event, there is a chance that an origination event will take place. Thus, when we see amonophyletic genus of a size larger than one, we know that there was no origination event since the most recent common ancestor of this genus, as long as there is no other genus that split after the time of the most recent

common ancestor (i.e., from the subtree of this specific genus). The probability that during the time from the most recent common ancestor there will be no origination event is described by a geometric distribution. Assuming that all the genera are independent samples from this distribution, we can reconstruct a measured distribution for the probability that an origination event will not occur as a function of the volume of the genus. By fitting this to a geometric distribution, we can get an estimate of the origination rate. In Figure S4 we present the measured distribution that an origination event will not take place, along with its best fit. The estimated $\nu$ is 0.096 [0.038- 0.3] per speciation event (assuming a generation time in million years of 2.1 for the mean, 1.4 for the lower bound and 2.8 for the upper bound ). This is similar to the values of the Passerine birds and the Aves class.

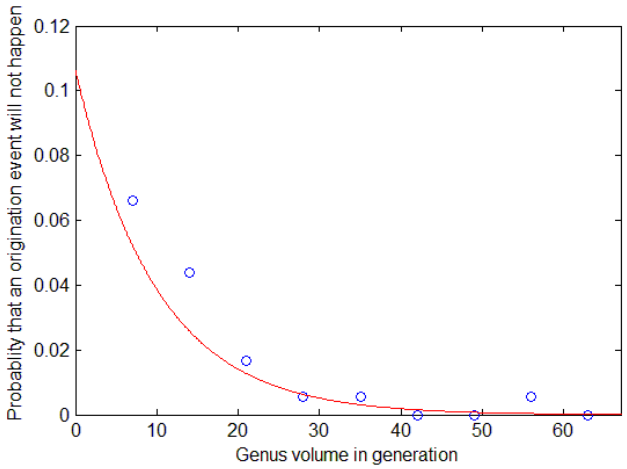


**Figure S4. The figure presents the probability of having a genus with a specific tree size which is the same as the probability that an origination event will not take place. The blue circles are the data from the Furnariidae family. The red line is the best fit of a geometric distribution, with an estimated origination rate $\nu$ = 0.096 [0.05-0.14] assuming a generation time of 2.1 million years. The data here are binned with bins of size 7 (in generational units), however very close results were obtained with other bin sizes.**

## G. Fits of the SEO model to lower taxa

In the main text, we presented the fit of the SEO model to large classes. Here we present a few examples where the model is applied to lower taxa. As we explained in the main text, for lower taxa the model assumptions may be questionable, and in some cases lead to unrealistic parameters, but for others, the fit is reasonable. For reliable examples, we present the fit to the large orders of the Aves class, the Passerine order (Figure S5) and the Ciconiiformes order (Figure S6). For these two examples, the fit matches very well the known numbers.

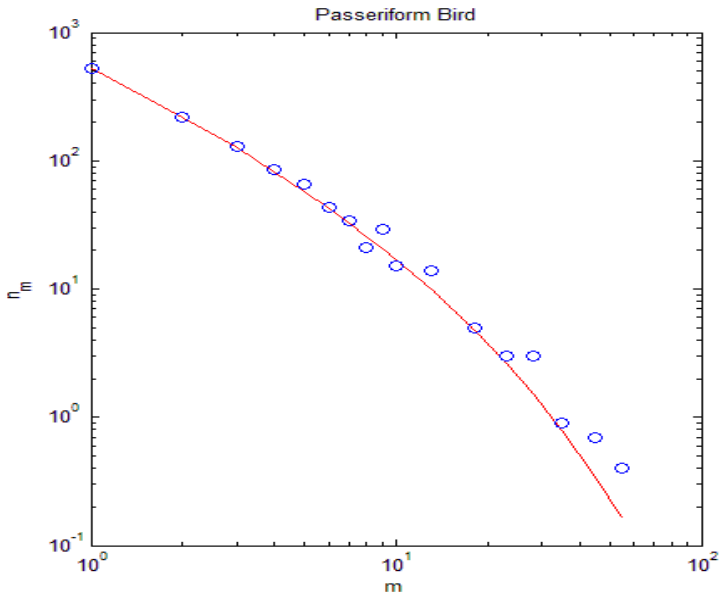


**Figure S5. The best fit of the SEO distribution to the observed SGD of the order Passeriformes. Here the best fit parameters are γ =0.12 and ν = 0.14.**

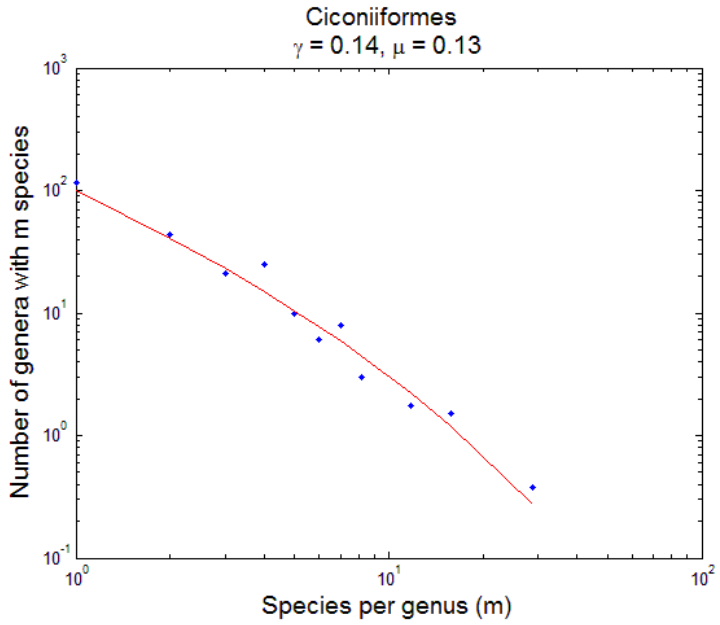


**Figure S6. The best fit of the SEO distribution to the observed SGD of the order Ciconiiformes. Here the best fit parameters are γ =0.14 and ν = 0.13.**

As an example where the model assumptions are probably wrong, we present the Asparagales order (Figure S7) of angiosperms, which contains $N_0 = 32425$ species [2]. The fitting parameters are $\gamma = 0.012$ and $\nu = 0.009$, and no meaningful deviation can be seen. However, the estimated time to the appearance of the order is larger than the one known from molecular data. Under pure exponential growth, the time is $T = \log(N_0)/\gamma = 797$ generations, and with the corrections presented in Ref. [3] is $T = \ln([\gamma N+1]/\lambda)/\gamma = 374$

generations. The estimations using molecular data [4] are 120 MY for the origin of the order, and 90 MY for the core group, which are about four to three times smaller than our estimations. This deviation could be explained if this order had already entered its saturation phase. In such a case, as explained in the main text, the SGD is independent of the origination time. Another example of inconsistency between the model and the data can be seen in the SGD of the Nymphalidae family (Figure S8), as discussed in the main text.

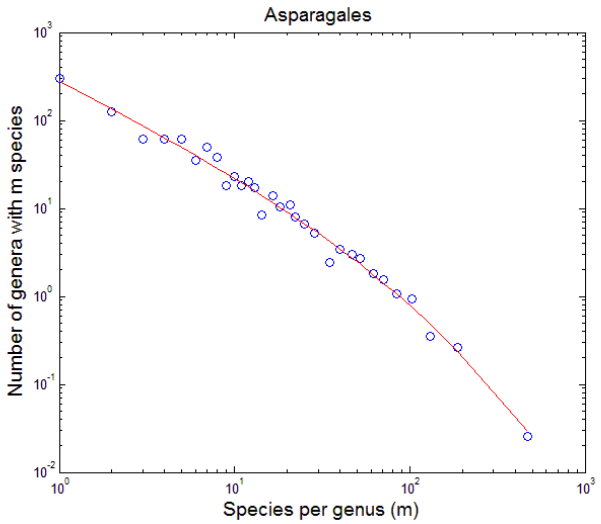


**Figure S7. The best fit of the SEO distribution to the observed SGD of the Asparagales order. Here the best fit parameters are $\gamma = 0.14$ and $\nu = 0.13$.**

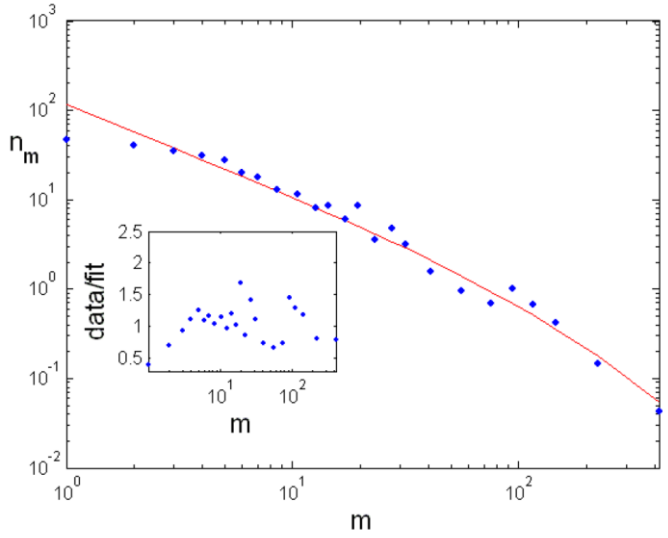


**Figure S8. The best fit of the SEO distribution to the observed SGD of the Nymphalidae family. Here the best fit parameters are $\gamma = 0.0037$ and $\nu = 0.0043$. Note that the SEO statistic tends to underestimate the number of singletons (monotypes) and doubletons as seen, for example, in Figures 1 and 2 of the main text. The reason, as discussed above, is the inadequacy of the continuum approximation that leads to Eq. (3) close to *m* = 1. Here the SEO fit overestimates the number of singletons by a factor of 2 (see inset); this is another indication of the failure of the model.**

# I. The binning technique

In all the SGDs presented in this paper, we presented the empirical data as a histogram showing the number of genera of size m, $n(m)$, vs. $m$. As the distribution is very wide, we must adopt a binning technique. Here we have used the following procedure: we have defined an integer D (practically we have used D = 15 for classes and D = 50 for kingdoms) that sets the minimal number of genera in each bin. Starting from, say, an arbitrary $m_0$, we collect genera of size $m_0$ and higher until we accumulate D genera at, say, $m_i$. The corresponding point in the SGD figure is at

$$n_m = \frac{D}{m_1 - m_0 + 1} \qquad m = \frac{1}{D}\sum_{i=1}^{D} m_i \quad ,$$

Where $m_i$ is the number of species in the $i$th genus.

The binning algorithm works as follows: first, order all the genera according to their rank $m$, where genera with the same rank are ordered arbitrarily. Starting from the smallest $m$, collect the $D$ consecutive genera, and calculate $n_m$ and m according to the above formula. Iterate this procedure starting from the first genus that has not yet been counted. If there are two (or more) values of $n_m$ for the same $m$, simply add them.

We stress that the results are independent of the binning procedure, and similar graphs were obtained when we used standard logarithmic binning. The advantage of our technique is that the number of points represented in each bin is fixed, and thus the uncertainty of each circle in the histogram is similar (different circles have similar confidence intervals). This technique also has the advantage of having more data points and thus enabling a better fit quality when the number of genera is not very large.

# J. The source of the SGD data

As we mentioned in the introduction, we used the largest dataset available to reconstruct the SGD distributions against which we compare our model as well as Yule's classic model. The species2000 website (http://www.sp2000.org/) contains a list of more than 1 million taxonomic groups (species, genera, families, …, up to kingdoms) along with their hierarchies. Comprehensive datasets are available in a mySQL format without charge at http://www.catalogueoflife.org/services/. One of the tables is an updated database dated January 3, 2011, which contains all the taxonomic groups along with its immediate parent. Using this table, one can count all the descendants of every taxonomic group and thus derive, for example, all the genera in the Class Aves and the number of species that are members of each genus.

# K. Derivation of equation 3 of the main text

To explain the derivation of the equation governing the evolutionary process simply, we will assume discrete 'generations' where all existing species go extinct at a specific time, and give birth to the next generation of species. However, the resulting equation is not limited to this assumption, but is correct even for overlapping generations [5-7]. Throughout the paper, we assume an exponential distribution of species durations.

Assuming discrete generations, the average number of genera of size *m* ($n_m$) as a function of time can be described as follows, Assume that in the previous generation there were genera of varying sizes. Each one of them gives birth to a random number *p* of offspring species. Among these *p* species, a random number, *m*, of them remain in the same genus with probability 1-$\nu$ per species, and *p*-*m* 'mutate' with probability $\nu$ (per species) and produce a new genus. Thus, the number of genera of size one (monotypes) reflects the creation of the new genera, in addition to shrinkage in the size of existing genera. This can be described by the following difference equations:

$$n_m^{t+1} = \sum_{\substack{l \\ p \geq m}} n_l P(l \to p) \binom{p}{m} \nu^{p-m} (1-\nu)^m; \quad m > 1$$

$$n_1^{t+1} = \sum_{\substack{l \\ p \geq 1}} n_l P(l \to p) \binom{p}{1} \nu^{p-1} (1-\nu)^1 + \nu N(t+1)$$

where $P(l \to p)$ is the probability that a genus of size *l* gives birth to *p* new species. The mean of *P* is $\lambda$. The continuum limit of the above difference equations (and also of the Moran version of the process, which is the process assumed throughout the paper) is [5-7]:

$$\frac{\partial n(m)}{\partial t} = (\gamma - \nu) \frac{\partial}{\partial m} [m \cdot n(m)] + \frac{\partial^2}{\partial m^2} [m \cdot n(m)]$$

which is equation 3 of the main text.

# L. Rejection test of the SEO model.

We give here the list of the taxonomic groups on which we tested the SEO model. We used orders and above, and only groups that contained at least 500 species and 100 genera. Table S2 presents the groups, the best fit, the number of species and genera, the group hierarchy and whether or not the model was rejected. 0 represents rejection and 1 represent non-rejection.

**Table S2. SEO best fit and rejection for various taxonomic groups**

| Name | Hierarchy | $\gamma$ | $\nu$ | Number of species | Number of genera | Rejected (0) or not (1) |
|---|---|---|---|---|---|---|
| Animalia | Kingdom | 0.06 | 0.03 | 867669 | 84529 | 0 |
| Plantae | Kingdom | 0.04 | 0.02 | 181858 | 10763 | 1 |
| Bacteria | Kingdom | 0.23 | 0.11 | 9096 | 1667 | 1 |
| Chromista | Kingdom | 0.32 | 0.13 | 6182 | 1185 | 1 |
| Protozoa | Kingdom | 0.35 | 0.14 | 6017 | 1118 | 1 |
| Viruses | Kingdom | 0.08 | 0.08 | 1906 | 300 | 1 |
| Fungi | Kingdom | 0.19 | 0.09 | 32580 | 5977 | 0 |
| Annelida | Phylum | 0.13 | 0.15 | 3862 | 1016 | 1 |
| Arthropoda | Phylum | 0.05 | 0.02 | 758879 | 65848 | 0 |
| Magnoliophyta | Phylum | 0.03 | 0.01 | 146441 | 6825 | 1 |
| Proteobacteria | Phylum | 0.25 | 0.21 | 2353 | 645 | 0 |
| Chordata | Phylum | 0.09 | 0.07 | 61422 | 9874 | 1 |
| Cnidaria | Phylum | 0.08 | 0.06 | 9907 | 1542 | 1 |
| Echinodermata | Phylum | 0.30 | 0.32 | 1010 | 372 | 0 |
| Ectoprocta | Phylum | 0.13 | 0.18 | 1045 | 306 | 0 |
| Mollusca | Phylum | 0.09 | 0.09 | 11679 | 2194 | 1 |
| Nemata | Phylum | 0.13 | 0.16 | 3160 | 774 | 0 |
| Porifera | Phylum | 0.03 | 0.03 | 8052 | 725 | 1 |
| Cyanobacteria | Phylum | 0.06 | 0.04 | 2787 | 327 | 1 |
| Haptophyta | Phylum | 0.10 | 0.18 | 537 | 175 | 1 |
| Ochrophyta | Phylum | 0.30 | 0.11 | 4904 | 864 | 1 |
| Bacillariophyta | Phylum | 0.08 | 0.02 | 6249 | 402 | 1 |
| Bryophyta | Phylum | 0.04 | 0.02 | 13365 | 1057 | 0 |
| Chlorophyta | Phylum | 0.28 | 0.09 | 5582 | 808 | 1 |
| Rhodophyta | Phylum | 0.13 | 0.07 | 6488 | 893 | 1 |
| Dinophyta | Phylum | 0.51 | 0.2 | 1693 | 350 | 1 |
| Zygomycota | Phylum | 0.49 | 0.16 | 944 | 165 | 1 |
| Ascomycota | Phylum | 0.29 | 0.13 | 22056 | 4416 | 1 |
| Basidiomycota | Phylum | 0.09 | 0.06 | 9168 | 1306 | 1 |

| | | | | | | |
|---|---|---|---|---|---|---|
| Firmicutes | Phylum | 0.95 | 0.23 | 1520 | 268 | 1 |
| Actinobacteria | Phylum | 0.20 | 0.07 | 1707 | 190 | 1 |
| Clitellata | Class | 0.16 | 0.2 | 721 | 231 | 1 |
| Polychaeta | Class | 0.14 | 0.15 | 2986 | 754 | 1 |
| Arachnida | Class | 0.06 | 0.04 | 55147 | 5526 | 0 |
| Magnoliopsida | Class | 0.05 | 0.02 | 87281 | 4459 | 1 |
| Gammaproteobacteria | Class | 0.43 | 0.18 | 1068 | 230 | 1 |
| Entognatha | Class | 0.03 | 0.03 | 1990 | 189 | 1 |
| Insecta | Class | 0.04 | 0.02 | 661372 | 52838 | 0 |
| Diplopoda | Class | 0.29 | 0.13 | 9907 | 2176 | 1 |
| Chilopoda | Class | 0.15 | 0.06 | 3139 | 399 | 1 |
| Aves | Class | 0.09 | 0.10 | 9913 | 2123 | 1 |
| Malacostraca | Class | 0.10 | 0.08 | 18419 | 3211 | 0 |
| Maxillopoda | Class | 0.06 | 0.08 | 4963 | 957 | 1 |
| Ostracoda | Class | 0.08 | 0.10 | 1624 | 327 | 1 |
| Ascidiacea | Class | 0.04 | 0.02 | 2248 | 169 | 1 |
| Amphibia | Class | 0.04 | 0.03 | 5753 | 449 | 1 |
| Mammalia | Class | 0.10 | 0.10 | 4832 | 1151 | 1 |
| Reptilia | Class | 0.05 | 0.05 | 8624 | 1056 | 1 |
| Anthozoa | Class | 0.08 | 0.06 | 5931 | 885 | 1 |
| Hydrozoa | Class | 0.09 | 0.06 | 3701 | 568 | 1 |
| Elasmobranchii | Class | 0.08 | 0.09 | 996 | 181 | 1 |
| Actinopterygii | Class | 0.11 | 0.08 | 28721 | 4678 | 1 |
| Gymnolaemata | Class | 0.13 | 0.17 | 1027 | 298 | 0 |
| Bivalvia | Class | 0.11 | 0.12 | 2168 | 470 | 1 |
| Gastropoda | Class | 0.09 | 0.08 | 8365 | 1501 | 1 |
| Adenophorea | Class | 0.13 | 0.14 | 2637 | 585 | 0 |
| Secernentea | Class | 0.08 | 0.22 | 523 | 189 | 0 |
| Demospongiae | Class | 0.02 | 0.02 | 6845 | 527 | 1 |
| Alphaproteobacteria | Class | 0.23 | 0.24 | 645 | 202 | 1 |
| Prymnesiophyceae | Class | 0.13 | 0.19 | 524 | 171 | 1 |
| Coscinodiscophyceae | Class | 0.12 | 0.05 | 1646 | 161 | 1 |
| Phaeophyceae | Class | 0.31 | 0.12 | 1988 | 320 | 1 |
| Bacillariophyceae | Class | 0.11 | 0.02 | 6046 | 340 | 1 |
| Bryopsida | Class | 0.03 | 0.02 | 12962 | 1054 | 0 |
| Chlorophyceae | Class | 0.78 | 0.32 | 1746 | 446 | 1 |
| Liliopsida | Class | 0.02 | 0.01 | 59160 | 2366 | 1 |
| Florideophyceae | Class | 0.13 | 0.07 | 6221 | 856 | 1 |
| Dinophyceae | Class | 0.49 | 0.19 | 1675 | 340 | 1 |
| Granuloreticulosea | Class | 0.21 | 0.15 | 1352 | 309 | 1 |
| Ascomycetes | Class | 0.30 | 0.13 | 21817 | 4324 | 1 |
| Basidiomycetes | Class | 0.08 | 0.05 | 8115 | 1040 | 0 |
| Urediniomycetes | Class | 0.17 | 0.16 | 713 | 182 | 1 |
| Actinobacteria | Class | 0.20 | 0.07 | 1707 | 190 | 1 |
| Aciculata | Order | 0.13 | 0.13 | 1595 | 368 | 1 |
| Canalipalpata | Order | 0.12 | 0.15 | 978 | 269 | 1 |
| Araneae | Order | 0.06 | 0.04 | 39316 | 3681 | 1 |

| | | | | | | |
|---|---|---|---|---|---|---|
| Pseudoscorpiones | Order | 0.06 | 0.06 | 3387 | 448 | 1 |
| Coleoptera | Order | 0.08 | 0.03 | 138814 | 12997 | 1 |
| Blattodea | Order | 0.03 | 0.04 | 4430 | 486 | 1 |
| Diptera | Order | 0.03 | 0.01 | 148582 | 9078 | 0 |
| Hemiptera | Order | 0.14 | 0.08 | 22765 | 3671 | 1 |
| Phasmida | Order | 0.07 | 0.07 | 2821 | 436 | 1 |
| Hymenoptera | Order | 0.03 | 0.02 | 43540 | 3155 | 0 |
| Orthoptera | Order | 0.10 | 0.09 | 23048 | 4320 | 1 |
| Lepidoptera | Order | 0.03 | 0.02 | 254937 | 16513 | 1 |
| Odonata | Order | 0.02 | 0.04 | 5321 | 626 | 1 |
| Siphonaptera | Order | 0.03 | 0.04 | 2040 | 241 | 1 |
| Thysanoptera | Order | 0.28 | 0.16 | 851 | 179 | 1 |
| Trichoptera | Order | 0.03 | 0.02 | 11513 | 620 | 1 |
| Amphipoda | Order | 0.10 | 0.10 | 3119 | 655 | 1 |
| Decapoda | Order | 0.31 | 0.17 | 4079 | 837 | 1 |
| Calanoida | Order | 0.01 | 0.06 | 992 | 155 | 1 |
| Harpacticoida | Order | 0.04 | 0.06 | 2208 | 348 | 1 |
| Podocopida | Order | 0.06 | 0.09 | 1142 | 231 | 1 |
| Anura | Order | 0.03 | 0.02 | 5055 | 356 | 1 |
| Passeriformes | Order | 0.08 | 0.09 | 5827 | 1183 | 1 |
| Squamata | Order | 0.05 | 0.05 | 8292 | 953 | 1 |
| Actiniaria | Order | 0.52 | 0.34 | 738 | 223 | 1 |
| Alcyonacea | Order | 0.02 | 0.03 | 2933 | 286 | 1 |
| Anthoathecata | Order | 0.15 | 0.1 | 1246 | 230 | 1 |
| Scleractinia | Order | 0.05 | 0.06 | 1480 | 232 | 1 |
| Leptothecata | Order | 0.04 | 0.03 | 2048 | 195 | 1 |
| Cheilostomata | Order | 0.12 | 0.17 | 825 | 233 | 1 |
| Veneroida | Order | 0.19 | 0.18 | 877 | 213 | 1 |
| Neogastropoda | Order | 0.12 | 0.08 | 1787 | 276 | 1 |
| Neotaenioglossa | Order | 0.07 | 0.07 | 1939 | 323 | 1 |
| Stylommatophora | Order | 0.07 | 0.07 | 1880 | 351 | 1 |
| Desmodorida | Order | 0.03 | 0.09 | 708 | 158 | 1 |
| Cypriniformes | Order | 0.08 | 0.06 | 3624 | 443 | 1 |
| Hypnales | Order | 0.09 | 0.03 | 3264 | 282 | 1 |
| Chlorococcales | Order | 0.48 | 0.28 | 603 | 169 | 1 |
| Poales | Order | 0.03 | 0.01 | 13226 | 556 | 1 |
| Asparagales | Order | 0.01 | 0.01 | 32425 | 1159 | 1 |
| Alismatales | Order | 0.03 | 0.01 | 3527 | 165 | 1 |
| Malpighiales | Order | 0.08 | 0.01 | 8998 | 330 | 1 |
| Rubiales | Order | 0.02 | 0.01 | 13229 | 610 | 1 |
| Asterales | Order | 0.07 | 0.06 | 3040 | 451 | 1 |
| Caryophyllales | Order | 0.10 | 0.07 | 1269 | 185 | 1 |
| Fabales | Order | 0.02 | 0.01 | 19937 | 732 | 1 |
| Lamiales | Order | 0.06 | 0.06 | 1133 | 152 | 1 |
| Rosales | Order | 0.03 | 0.01 | 21932 | 257 | 1 |
| Scrophulariales | Order | 0.30 | 0.09 | 1380 | 193 | 1 |
| Ceramiales | Order | 0.25 | 0.11 | 2497 | 394 | 1 |

| | | | | | | |
|---|---|---|---|---|---|---|
| Gigartinales | Order | 0.20 | 0.12 | 997 | 191 | 1 |
| Foraminiferida | Order | 0.21 | 0.15 | 1352 | 309 | 1 |
| Siluriformes | Order | 0.09 | 0.07 | 3217 | 463 | 1 |
| Scorpaeniformes | Order | 0.44 | 0.17 | 1518 | 292 | 1 |
| Perciformes | Order | 0.07 | 0.08 | 10147 | 1698 | 1 |
| Helotiales | Order | 0.12 | 0.13 | 1718 | 381 | 1 |
| Hypocreales | Order | 0.13 | 0.08 | 1078 | 170 | 1 |
| Agaricales | Order | 0.06 | 0.04 | 3203 | 352 | 1 |
| Xylariales | Order | 0.33 | 0.12 | 933 | 155 | 1 |
| Pleosporales | Order | 0.06 | 0.05 | 1744 | 214 | 1 |
| Pezizales | Order | 0.05 | 0.10 | 652 | 153 | 1 |
| Lecanorales | Order | 0.15 | 0.06 | 2523 | 310 | 1 |
| Polyporales | Order | 0.06 | 0.06 | 2179 | 318 | 1 |
| Geophilomorpha | Order | 0.28 | 0.12 | 1263 | 234 | 1 |
| Chordeumatida | Order | 0.10 | 0.12 | 1118 | 284 | 0 |
| Julida | Order | 0.29 | 0.11 | 1222 | 190 | 1 |
| Polydesmida | Order | 0.81 | 0.30 | 3749 | 1061 | 1 |
| Spirostreptida | Order | 0.25 | 0.10 | 1758 | 284 | 1 |
| Characiformes | Order | 0.20 | 0.10 | 1845 | 279 | 1 |
| Isopoda | Order | 0.06 | 0.05 | 8991 | 1303 | 0 |
| Mysida | Order | 0.07 | 0.07 | 1071 | 165 | 1 |
| Chiroptera | Order | 0.47 | 0.17 | 928 | 177 | 1 |
| Actinomycetales | Order | 0.17 | 0.06 | 1642 | 167 | 1 |
| Ciconiiformes | Order | 0.14 | 0.13 | 1085 | 258 | 1 |
| Arecales | Order | 0.11 | 0.03 | 2406 | 190 | 1 |
| Anguilliformes | Order | 0.07 | 0.08 | 870 | 157 | 1 |
| Sarcoptiformes | Order | 0.05 | 0.04 | 9142 | 1027 | 0 |
| Rodentia | Order | 0.12 | 0.12 | 2024 | 444 | 1 |

**Table S3. The same as Table1 of the main text with references, diversification rate and the appearance time for different classes**

| Name | Size (number of species) | Growth rate $\gamma$ [± STD]<br>Origination rate $\nu$ [± STD] | Generations since origination | Estimated time to origination, *T* (MY) | Independent estimate of *T* (MY) |
|---|---|---|---|---|---|
| Arachnida | 55147 | $\gamma = 0.055 \pm 0.0064$<br>$\nu = 0.0359 \pm 0.0023$ | 144 [131-161] | 302[183- 450] | 420 [8] |
| Magnoliopsida (Angiospermopsida) | 87281 | $\gamma = 0.051 \pm 0.0058$<br>$\nu = 0.015 \pm 0.0012$ | 163 [148-182] | 342 [207- 509] | 228 [9] |
| Insecta | 661370 | $\gamma = 0.037 \pm 0.0019$<br>$\nu = 0.0185 \pm 0.0006$ | 272 [260-285] | 571 [364- 798] | 420 [10] |
| Diplopoda | 9907 | $\gamma = 0.23 \pm 0.0348$<br>$\nu = 0.10 \pm 0.012$ | 32 [28 -37] | 67 [39-103] | 420 [11] |
| Aves | 9913 | $\gamma = 0.08 \pm 0.021$<br>$\nu = 0.089 \pm 0.010$ | 82 [67-107] | 172 [93-299] | 130 [12] |
| Passerine birds (order) | 6198 | $\gamma = 0.12 \pm 0.015$<br>$\nu = 0.14 \pm 0.013$ | 54[48-60] | 113 [67 -168] | 82 [13] |
| Malacostraca | 18419 | $\gamma = 0.086 \pm 0.0115$<br>$\nu = 0.068 \pm 0.0054$ | 84 [76-95] | 176 [106 -266] | 510 [14] |
| Maxillopoda | 4963 | $\gamma = 0.06 \pm 0.0096$<br>$\nu = 0.074 \pm 0.0146$ | 94 [83-108] | 197 [116-302] | 500 [15] |
| Amphibia | 5753 | $\gamma = 0.051 \pm 0.0118$<br>$\nu = 0.032 \pm 0.0042$ | 110 [93-137] | 231 [130-383] | 315 [16] |
| Mammalia | 4832 | $\gamma = 0.15 \pm 0.0176$<br>$\nu = 0.118 \pm 0.019$ | 42 [39-47] | 88 [54-131] | 120 [17] |

The SGD of each class has been fitted (see Fig. 2 as an example) using the SEO model to yield the diversification rate $\gamma$ (third column). From the total number of species in the class (second column) and the diversification rate, one may extract the number of "generations" since the first appearance of this class (forth column). In order to translate generations to time, we took a single generation (typical time to extinction) as 2.1[1.4-2.8] MY as presented in the fifth column. This result of the SEO model should be compared with other, independent estimates, based on either fossil data or genetic analysis (sixth column, with the corresponding reference). In most cases, the SEO-based estimates are close to the results from independent sources. Note that the definition of generation time is quite arbitrary, and may vary among classes; this may explain some mismatch with the fossil data. The factor of 2 differences for Malacostraca and Maxillopoda may be related to incomplete data about the number of species. For Diplopoda, the estimates are inconsistent; they may reflect an inadequacy of our demographic model in this case, or an underestimation for the generation time of diplopod species. The Mammal class is dominated by the placentals (4600 of the 4832 species), and so the appropriate time is given for the placentals.

1. Derryberry EP, Claramunt S, Derryberry G, et al. Lineage Diversification And Morphological Evolution In A Large-Scale Continental Radiation: The Neotropical Ovenbirds And Woodcreepers (Aves: Furnariidae). *Evolution.* 2011;65(10):1558-5646.
2. Bisby FA, Roskov YR, Orrell TM, et al. Species 2000 & ITIS Catalogue of Life. 2010.
3. Ricklefs RE. Estimating diversification rates from phylogenetic information. *Trends in Ecology & Evolution.* 2007;22(11):601-610.
4. Janssen T, Bremer K. The age of major monocot groups inferred from 800+ rbcL sequences. *Botanical Journal of the Linnean Society.* 2004;146(4):385-398.
5. Manrubia SC, Zanette DH. At the boundary between biological and cultural evolution: the origin of surname distributions. *Journal of Theoretical Biology.* 2002;216(4):461-477.
6. Maruvka YE, Shnerb NM, Kessler DA. Universal features of surname distribution in a subsample of a growing population. *Journal of Theoretical Biology.* Jan 2010;262(2):245-256.
7. Maruvka YE, Kessler DA, Shnerb NM. The birth-death-mutation process: a new paradigm for fat tailed distributions. *PLoS One.* 2011;6(11):e26480.
8. Dunlop JA. A trigonotarbid arachnid from the Upper Silurian of Shropshire. *Palaeontology.* 1996;39:605-614.
9. Smith SA, Beaulieu JM, Donoghue MJ. An uncorrelated relaxed-clock analysis suggests an earlier origin for flowering plants. *Proceedings of the National Academy of Sciences of the United States of America.* 2010;107(13):5897-5902.
10. Gaunt MW, Miles MA. An insect molecular clock dates the origin of the insects and accords with palaeontological and biogeographic landmarks. *Molecular Biology and Evolution.* 2002;19(5):748-761.
11. Sierwald P, Bond JE. Current status of the myriapod class diplopoda (Millipedes): taxonomic diversity and phylogeny. *Annual Review of Entomology.* 2007;52:401-420.
12. Brown JW, Rest JS, Garcia-Moreno J, Sorenson MD, Mindell DP. Strong mitochondrial DNA support for a Cretaceous origin of modern avian lineages. *Bmc Biology.* 2008;6.
13. Barker FK, Cibois A, Schikler P, Feinstein J, Cracraft J. Phylogeny and diversification of the largest avian radiation. *Proceedings of the National Academy of Sciences USA.* 2004;101:11040-11045.
14. Schram FR. Convergences between late Paleozoic and modern caridoid Malacostraca. *Systematic Zoology.* 1974;23(3):323-332.
15. Waloszek D, Repetski JE, Maas A. A new Late Cambrian pentastomid and a review of the relationships of this parasitic group. *Transactions of the Royal Society of Edinburgh-Earth Sciences.* 2005;96:163-176.
16. San Mauro D. A multilocus timescale for the origin of extant amphibians. *Molecular Phylogenetics and Evolution.* 2010;56(2):554-561.
17. Ji Q, Luo ZX, Yuan CX, Wible JR, Zhang JP, Georgi JA. The earliest known eutherian mammal. *Nature.* 2002;416(6883):816-822.