Phylogenetically Clustered Extinction Risks Do Not Substantially Prune the Tree of Life

Rakesh K. Parhar[1,2]*, Arne Ø. Mooers[1,2]

1 Biological Sciences, Simon Fraser University, Burnaby, British Columbia, Canada, **2** IRMACS, Simon Fraser University, Burnaby, British Columbia, Canada

Abstract

Anthropogenic activities have increased the rate of biological extinction many-fold. Recent empirical studies suggest that projected extinction may lead to extensive loss to the Tree of Life, much more than if extinction were random. One suggested cause is that extinction risk is heritable (phylogenetically patterned), such that entire higher groups will be lost. We show here with simulation that phylogenetically clustered extinction risks are necessary but not sufficient for the extensive loss of phylogenetic diversity (PD) compared to random extinction. We simulated Yule trees and evolved extinction risks at various levels of heritability (measured using Pagel's λ). At most levels of heritability (λ in range of 0 to 10), mean values of extinction risk (range 0.25 to 0.75), tree sizes (64 to 128 tips), tree balance and temporal heterogeneity of diversification rates (Yule and coalescent trees), extinction risks do not substantially increase the loss of PD in these trees when compared to random extinction. The maximum loss of PD (20% above random) was only associated with the combination of extremely excessive values of phylogenetic signal, high mean species' extinction probabilities, and extreme (coalescent) tree shapes. Interestingly, we also observed a decline in the rate of increase in the loss of PD at high phylogenetic clustering ($\lambda \rightarrow 10$) of extinction risks. Our results suggest that the interplay between various aspects of tree shape and a predisposition of higher extinction risks in species-poor clades is required to explain the substantial pruning of the Tree of Life.

Introduction

Phylogenetic trees estimate the evolutionary relationships among species inferred from empirical data. The edge lengths of these trees represent accrued change or temporal accounts of diversification events [1,2]. The sum of the edge lengths has been referred to as "evolutionary history" [EH; 3]. Every time a lineage (for example, a species) goes extinct, EH is lost. This loss can be conceptualized as a pruning of the twigs and branches from the Tree of Life. However, branches on the tree that are shared by multiple species are lost only if all the subtending species go extinct (see Figure 1). A related term is phylogenetic diversity [PD; 4], a measure of the length of the subtree connecting a subset of extant species to the root of a reference phylogenetic tree. The original PD of a clade is a function of its size (the number of tips), its depth (the distance from tips to root) and its shape; for a given depth and size, the star phylogeny [$\gamma \rightarrow -\infty$; 5] has maximum total PD. If one assigns a probability of extinction [p(ext)] (say, over the next 100 years) to the tips of a tree, then it is straightforward to estimate its expected future PD [E(PD); 1, 2, 6]. The difference from the original PD is the expected loss of PD.

Nee and May [3] were the first to show that there was a minimal loss of PD under homogeneous random extinction (commonly referred to as a "field of bullets" model) on a particular set of model trees. This result has been cited many times [7–12]. In contrast, empirical studies using real trees and estimated extinction probabilities from the IUCN red list [13] project fairly large losses of PD [7,12,14–16]. This discrepancy has been explained as being due, in part, to heritable (phylogenetically clustered) extinction risks, where related species have similar probabilities of extinction [7,17,18]. Indeed, phylogenetically clumped extinction can result in the loss of deeper branches in the Tree of Life [7,11,19–21]. However, Heard and Mooers [22] cautioned that heritable extinction probabilities alone might not be enough to cause large losses of PD relative to the field of bullets model. In spite of their caution 10 years ago, recent studies continue to reason that phylogenetically clustered extinction threats are responsible for the substantial loss of PD from the Tree of Life [7,11,19–21].

We return to this simple question here, and try to quantify loss of PD due to phylogenetically clustered extinction risks on model trees. We build upon Heard and Mooers [22] to incorporate the effects of a larger range of phylogenetic signal in extinction risk, different species' mean extinction risk, tree size, tree balance and distribution of nodes across a phylogeny. In addition, we also utilize the probabilistic measure of future PD, i.e. expected future PD, in our model. Even though our model does not entirely capture the exact shape of the current distribution of extinction risks [13], it does provide a reasonable estimate of the loss of PD, as well as a suitable foundation to investigate our question with other models. Overall, we ask under which conditions and to what extent phylogenetically clustered extinction risks might be necessary and sufficient for the observed pronounced loss of PD projected for the Tree of Life under the current extinction regime.
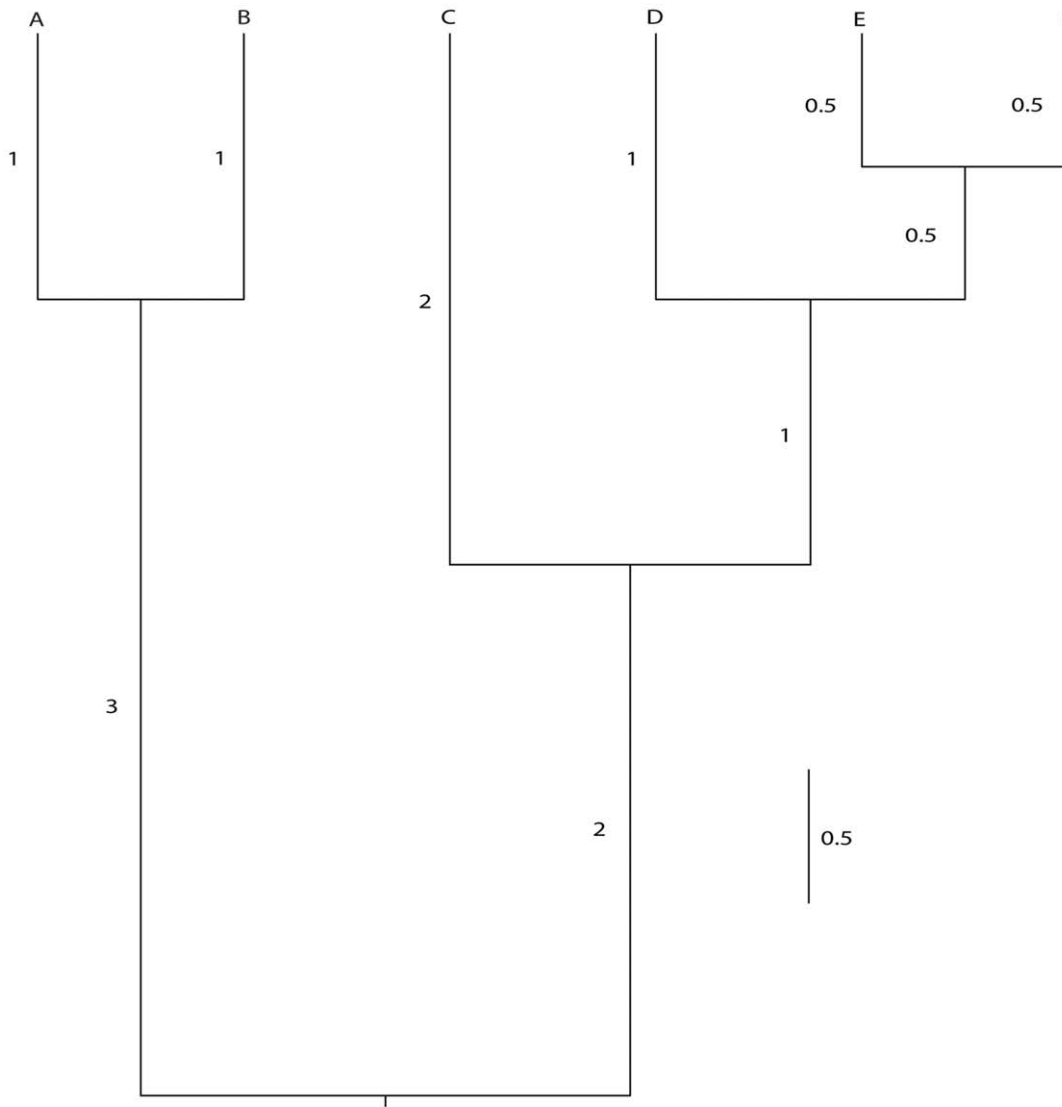
Figure 1. Observed PD loss on a tree. A cartoon tree of size n = 6, showing how losing different pairs of species lead to different PD loss. For this tree, total PD = 12.5.
doi:10.1371/journal.pone.0023528.g001

Methods

Generation of model trees

Yule trees [23–27] have long served as a null model for various macro-evolutionary phylogenetic studies [28–31]. They are less balanced than random tree shapes, though more balanced than many inferred trees [32,33], and their distribution of edge lengths falls between that expected under adaptive radiations [34] and long-term equilibrial conditions [35]. We simulated two sets of 1000 Yule trees, (speciation rate $= 0.5$), one with 64 and one with 128 tips using the appropriate simple-sample approach [SSA; 36]. We checked for the validity of these Yule trees with their gamma statistic [γ, 5]. As expected, we obtained a standard normal distribution of the gamma statistic (unpublished data). We also produced 1000 64-tip completely balanced and unbalanced trees, each with uniformly-distributed internal branch lengths, as well as a set of 1000 64-tip coalescent trees. We set all trees to a common depth value (depth $= 1$) to facilitate comparisons across simulations.

Trait simulation and the relevance of excessive phylogenetic signal

We then simulated a set of continuous traits along each tree under the Brownian Motion (BM) model of change, another common model for continuous trait evolution [37,38]. Specifically, we simulated these traits with different values of Pagel's λ [39] to model different strengths of phylogenetic clustering (further quantified as the relative similarity of tips on a tree compared to expectations from perfect BM) – from 0 (no clustering) to 1 (consistent with perfect Brownian motion, where the expected covariance between two nodes is equal to the height of their first common node). With Revell's framework [40], we were also able to incorporate higher values of $\lambda (1 \leq \lambda \leq 10)$ in our data. We categorized the values of lambda that fall under the open interval of $\lambda > 1$ to represent a higher relative similarity of tips compared to expectations from the perfect BM model. This case of elevated values of lambda could be attributed either to mistakes in tree inference, with internal branches being biased short (producing negative gamma values; [41,42]) or to evolutionary models such as the "Early Burst" (EB) model [43–48]. Under this process model, species' morphological trait values evolve more rapidly near the root than expected under the perfect BM model of evolution, followed by a relative stasis towards the tips. This phenomenon causes the trait values at the tips to be relatively more similar within a subclade than expected under perfect BM, producing higher tip disparity (higher variance among, rather than, within subclades) near the root and decreasing tip disparity as the trait approached the tips of a tree [47,48]. Figure S1 shows disparity through time plots of trait values on a model tree at various values of Pagel's λ to illustrate this phenomenon. In addition, the Figure S2 represents the evolution of a continuous trait with different phylogenetic signals, quantified by two independent measures of phylogenetic signal (Pagel's λ and Blomberg's K statistic) [39,46].

Extinction probabilities and expected future PD

Above, we produced continuous traits on our trees with a range of phylogenetic clustering from $\lambda = 0$ to $\lambda = 10$. Next, we transformed the trait values to produce distributions for the p(ext) of the corresponding tips at different levels of phylogenetic clustering using Equation 1 and 2 below. These p(ext) distributions (with a mean value of 0.5, see Figure 2) were then used to calculate the expected future PD of a phylogenetic tree using Equation 3.

For a given single tree, we used the following two transformations:

$$f(x) = (x_j - \bar{x})/\hat{x} \tag{1}$$

$$p(x) = \int_{-\infty}^{x_j} f(t)dt \tag{2}$$

where the term $p(x)$ represents the distribution of probability of extinction, x_j represents the raw final trait value of a particular taxa j, while $\bar{x}$ and $\hat{x}$ represent the mean and standard deviation of these corresponding traits, respectively. We used three different distributions with a low (0.25), medium (0.50) and high (0.75) mean values of species extinction risk, to better represent the range of mean anthropogenic impacts on p(ext). To simulate these scenarios, we used the original p(ext) distribution (with mean value of 0.5) and transformed (using simple division and/or multiplication) their values to produce other distributions with mean values of 0.25 and 0.75, respectively. We then assigned the extinction vulnerabilities to the corresponding taxa and calculated the expected future PD [E(PD)] of each tree under the "generalized field of bullets" model or the g-FOB model [2]:

$$E(PD) = \sum_{i=1} S_i (1 - \Pi_j p_j) \tag{3}$$

The term S_i represents the length of an edge i, while p_j represents the probability of extinction for the corresponding subtended jth taxa in a reference phylogeny.

In order to answer our main question, we compared E(PD) on the modeled p(ext) values with E(PD) obtained from the identical tree with shuffled p(ext) values, representing a useful extension of the uniform p(ext) FOB model reported by Nee and May [3] (see also [2]). We report the results [E(PD) and p(ext) distribution] for each set of 1000 trees, at different strengths of phylogenetic clustering and mean extinction risk. All tree simulations and data analysis were carried in the R programming environment [49], with the primary help of the "Ape" [50] and "Geiger" [51] packages.

Results

We observed no noticeable effect in any of our results with an increase in the size of Yule trees except a minor decrease in the confidence limits on our estimates (compare Figure 3 and Figure S3). We therefore only reported results for 64 tips below.

Extinction risk distributions

Varying the mean extinction risk and lambda produced different distributions of p(ext) across the tips of our trees. At no phylogenetic signal ($\lambda = 0$), we observed a relatively uniform distribution of p(ext) following our transformation, equivalent to a case of fully random extinction (Figure 2A). At $\lambda = 1$, we observed a quasi-normal distribution (Figure 2B). At higher phylogenetic signal, risk values diverged away from the mean value, resulting in a bimodal distribution at the highest lambda value (Figure 2C).

Mean extinction risk, and the relationship between Pagel's λ and the loss of PD

We defined the term *Percentage difference in projected PD* or *%ΔE(PD)* as:
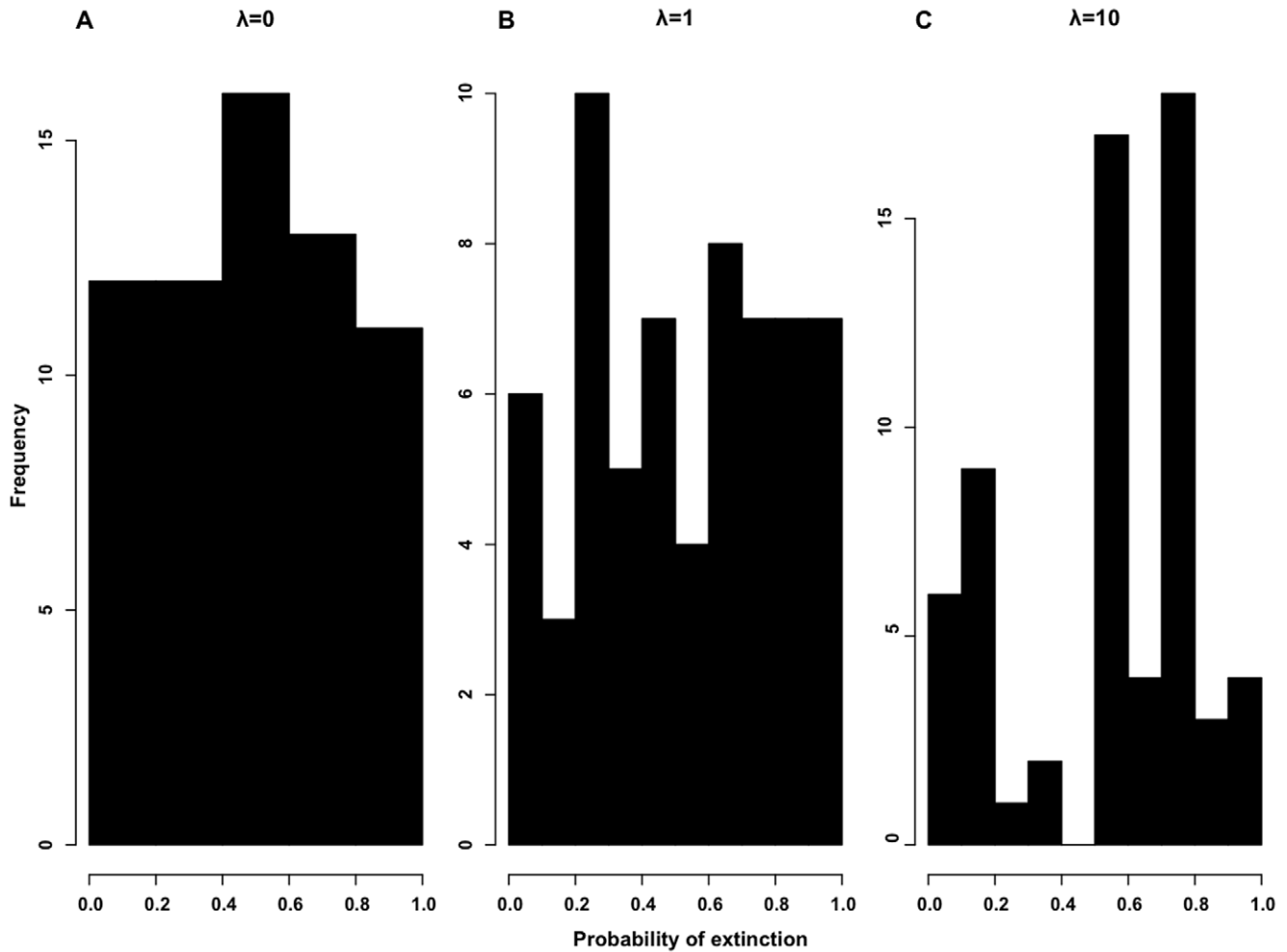
Figure 2. Differences in the extinction risk distributions. Extinction risk distributions across the tips of an example 64-tip Yule tree, modeled for three levels of phylogenetic clustering: $\lambda = 0$, 1, and 10. Each set has mean p(ext) = 0.5.
doi:10.1371/journal.pone.0023528.g002

$$\% \Delta E(PD) = 100\% * [PD(R) - PD(H)] / PD(R) \qquad (4)$$

where $PD(H)$ represented the projected or expected future PD calculated under a certain level of heritable extinction risk, while $PD(R)$ represented that quantity under a randomized set of those identical extinction risks.

Throughout our test scenarios, we observed consistent patterns in the relationship between the amount of phylogenetic clustering and the percentage difference in projected PD; we presented these patterns for 64-tip Yule trees in Figure 3. Overall, and as expected, when there was no phylogenetic clustering ($\lambda = 0$), %ΔE(PD) was centered on zero. We observed a minimal increase in%ΔE(PD) with an increase in Pagel's λ. In addition, the rate of increase in%ΔE(PD) reached relative stasis as we increased the phylogenetic signal beyond the perfect Brownian Motion model (Figure 3).

Further, we also observed the same qualitative pattern in%ΔE(PD) as a function of phylogenetic signal across all three scenarios of different mean extinction risk (compare Figure 3A, B, and C). For a particular scenario (size, balance, temporal change in the rate of diversification, etc.), we observed the maximum value of%ΔE(PD) to be consistently associated with the highest mean extinction risk. Though remaining absolutely low, the maximum value of%ΔE(PD) increased many-fold with a doubling of mean extinction risk. For instance, 64-tip Yule trees produced a

maximum loss of ~0.7% at mean p(ext) = 0.25, and a maximum loss of ~7%, at mean p(ext) = 0.50 (compare Figure 3A and 3B). Importantly, across all cases, the values of%ΔE(PD) were not significantly different at phylogenetic signals equal to and exceeding the perfect Brownian Motion model for low mean extinction threat (Figure 3A).

Effect of tree topology on loss of PD

On perfectly balanced trees (with Yule edge-lengths), the distribution of%ΔE(PD) as a function of phylogenetic signal was similar to the one produced with randomly generated Yule trees (with 64 or 128 tips), with a maximum value of ~14% at our highest chosen mean extinction risk (Figure S4). However, in the case of unbalanced trees (with Yule topologies), we observed a *decline* in the values of%ΔE(PD) with increasing phylogenetic signal (Figure S5). This led to slight negative mean values of%ΔE(PD).

An extreme case of γ, λ, and $p(ext)$

We also tested our question on the extreme model of diversification. The extreme case, i.e., coalescent trees, produced the highest (among all of our tested scenarios) values of%ΔE(PD) as a function of phylogenetic signal (Figure S6). Under this case, the largest amount of loss was relatively substantial (~20%), but
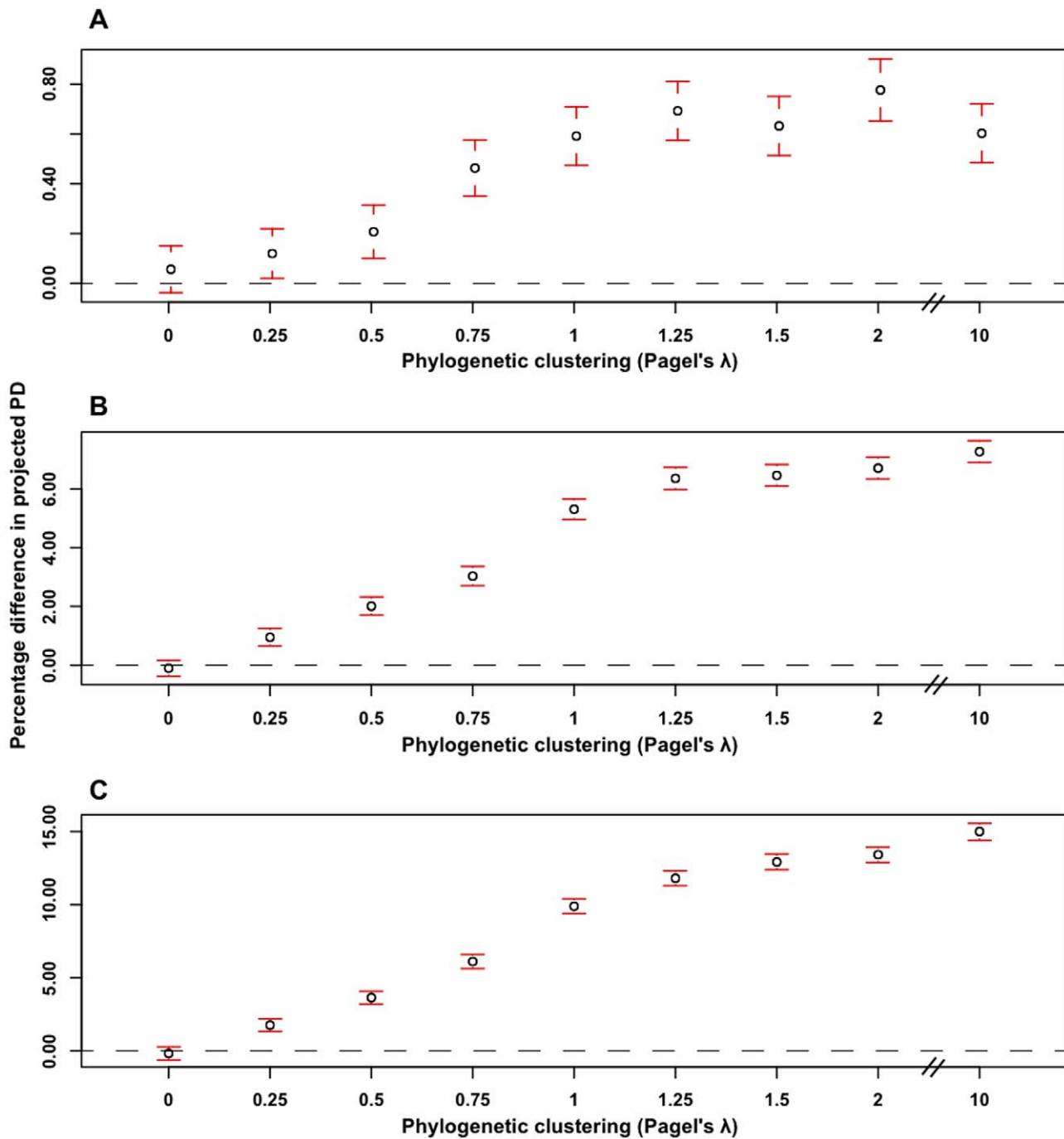
Figure 3. Percentage difference in projected PD with phylogenetic clustering in 64-tip Yule trees. A) mean p(ext) = 0.25, **B)** mean p(ext) = 0.5, and **C)** mean p(ext) = 0.75. Data points in percentage denote the amount of additional loss of projected PD (relative to random extinction) with increasing phylogenetic clustering. Dashed line indicates that loss under random extinction. Error bars around points represent the 95% confidence interval with a sample size of 1000 trees. Note differences in vertical axes.
doi:10.1371/journal.pone.0023528.g003

occurred only in the presence of both the highest mean extinction risk as well as extremely excessive phylogenetic signal (Figure S6C).

Discussion

We highlight two main findings. We interpret our results to mean that phylogenetically clustered extinction risks alone are not sufficient to explain appreciable extra losses of E(PD) [%ΔE(PD)]

due to projected extinction. In addition, the overall amount of this additional loss does not substantially scale with tree size, tree balance and extreme cases of temporal changes in the rates of diversification, respectively, though mean p(ext) has a strong relative effect. Rather, we require at least a combination of these factors with uncommonly excessive values to drive a substantial loss of PD.

Nee and May [3] previously suggested that random pruning events led to a minimal loss of PD under a uniform p(ext)

distribution. As well as not accounting for differences in the species level of endangerment [7,13,52,53] their modeled trees only made use of the extreme topology, i.e., the coalescent model. In contrast, our results are based on a generalized version [2] of Nee and May's [3] model and our modeled Yule trees are much closer in line with the shapes of inferred trees [54]. Finally, we investigated several aspects of phylogenies that could influence the percentage difference in projected PD. Results suggest that the introduction of phylogenetically patterned extinction risks alone leads to a low loss of the percentage difference in projected PD (Figure 3).

The general pattern

Under random extinction scenarios such as those modeled by Nee and May [3], species' extinction risks are allocated independent of their location in the corresponding phylogeny. This scenario does not link the extinction risk of a species with its biology (and thus its evolutionary history). As predicted, we found no effect on %ΔE(PD) at very low levels of phylogenetic clustering (Figure 3). When traits displayed relatively more phylogenetic signal (related taxa show similar trait values) non-random extinction is expected [7] and observed (our results).

However, contrary to previous studies [7,11,21], non-randomness (phylogenetic clustering) in species' extinction risk alone does not contribute towards a substantial loss of PD, even if this non-random effect exceeds the perfect Brownian Motion model on Yule topologies (Figure 3). Indeed, higher than 'expected' levels of phylogenetic clustering (i.e., Pagel's $\lambda \gg 1$) causes a decelerating rate for the additional loss of PD. As described earlier, at higher phylogenetic signals, simulated extinction risks deviate away from their mean value relative to cases of lower clustering (Figure 2). Since the probability function has a restricted domain ($0 \leq p \leq 1$), extinction probabilities cannot diverge away from their mean value without bound and so the increase in the percentage difference in projected PD is limited. At higher phylogenetic signals, this slows the rate of increase in this extra loss. Of course, more sophisticated models of extinction threat might lead to different outcomes.

Quantifying the effects of mean extinction risk on Yule trees

The extra loss of PD [%ΔE(PD)] also changes with different values of mean extinction risk (Equation 4). Logically, an increase in mean extinction risk will decrease both the projected PD due to random extinction [PD(R)], and projected PD due to phylogenetically clustered extinction [PD(H)] of similar magnitude, producing a small effect on the numerator of %ΔE(PD) (Equation 4). However, the denominator [PD(R)] decreases directly with increasing mean p(ext). Therefore, we expect %-ΔE(PD) to increase with increase in mean extinction risk. Consistent with this, we observed the largest value of %ΔE(PD) at cases of highest mean threat to species survival (mean p(ext) = 0.75). However, this maximal (but still low) percentage difference in projected PD does not scale linearly with the mean extinction risk. Again, this is due to the fact that we must limit range of extinction risk (p(ext) cannot be >1) and simultaneously maintain a constant mean value. This leads to a slight decrease in the variance of the distribution of extinction risk as mean p(ext) goes up, which potentially decreased the difference between the measures of PD(R) and PD(H). This is a limitation of our model, and we suggest future models should minimize differences in the variance of probability functions while maintaining various constant mean values.

Our results also indicate that there is a non-significant increase in the percentage difference in projected PD beyond the perfect

Brownian Motion model of change at low mean extinction risk (mean p(ext) = 0.25; see Figure 3A). This is expected because the combination of low mean extinction risk and a deceleration in the rate of increase in %ΔE(PD) at higher phylogenetic signals causes for an irrelevant additional loss of PD. Furthermore, the percentage difference in projected PD is also not much different (<0.6%) at other levels of phylogenetic clustering.

Role of tree balance in the loss of PD

We find an interesting pattern in the percentage difference in projected PD on maximally unbalanced trees. Overall, we see minimal loss. If mean p(ext) is high we actually lose more PD under random than under clustered extinction. Our unbalanced trees had very short internal edges, and a few very long pendant edges. Thus, when risks are clustered on an unbalanced tree, taxa with longer pendant edges are less likely to be pruned than those with shorter edges, decreasing the percentage difference in projected PD.

Alternatively, a balanced tree of the same size produces a loss similar to a randomly generated Yule tree. Here, every edge is of the same length and there is the maximum amount of topological redundancy. We therefore expect to lose more PD from lineages that share similar extinction risks.

The extreme case of diversification

We find that the amount of PD lost on coalescent trees is the most sensitive to particular values of our chosen parameters. For instance, we do find a similar quantitative pattern to Yule trees in the percentage difference in projected PD in coalescent trees at very low and moderate levels of λ and p(ext). However, on coalescent trees at high λ and p(ext), we can observe substantial loss of PD (~20%). This scenario, in addition to being extreme, is only relevant if real inferred trees are well represented by the coalescent model. Recent studies [55,56] indicate that this is unlikely.

In our study, we were unable to incorporate the exact shape of the distribution of extinction probabilities suggested by recent assessments [13]. This distribution is heavily skewed, assigning the majority of the species with low threat and a few species with high vulnerability. Future studies could incorporate this distribution in their model to get a much better picture of the loss of PD from the Tree of Life.

To conclude, we propose that other non-random processes in addition to phylogenetic clustering of species' extinction risks must explain the appreciable loss of PD projected on real trees. Inferred trees do have very isolated small clades, and the highest extinction risk may be found in such small isolated clades [12,19,35,54]. Why this is the case is an open and fairly urgent question.

Supporting Information

Figure S1 Disparity through time (DTT) plots for trait values at various phylogenetic signals. The x-axis indicates the relative time elapsed or age of clade with 0 representing its origin and 1 representing its current age. Solid line indicates the observed disparity values, whereas the dashed line represents the mean of 100 simulated disparity values expected under the Brownian Motion model. **A)** through **G)** represent various DTT plots at increasing phylogenetic signal simulated on one example tree. High relative disparity is indicative of more variation in trait values within subclades than between subclades, which are found near the tips of a phylogeny with the trait displaying no phylogenetic signal. In contrast, low relative disparity is indicative of less variation in trait values within subclades than between

subclades, which are found near the tips of a phylogeny with the trait displaying an excessive phylogenetic signal.
(TIFF)

Figure S2 Traitgrams and the two measures of phylogenetic signal. The three traitgrams show the evolution of a continuous trait evolving under different phylogenetic signals. The x-axis represents a continuous scale of species' trait values and node depths represent the phylogenetic edge lengths. Traitgram in **A)** represents the evolution of a trait with no phylogenetic signal ($\lambda = 0$), while traitgrams in **B)** and **C)** represent the evolution of a trait under the perfect BM model and when evolution exceeds the perfect BM model ($\lambda = 1$ and $\lambda = 10$), respectively. Each case of evolution is represented by two independent measures of phylogenetic signal (Pagel's λ and Blomberg's K statistic).
(TIFF)

Figure S3 Quantifying percentage difference in projected PD with phylogenetic clustering in 128-tip Yule trees. **A)** mean p(ext) $= 0.25$, **B)** mean p(ext) $= 0.5$, and **C)** mean p(ext) $= 0.75$. Data points in percentage denote the amount of additional loss of projected PD (relative to random extinction) with increasing phylogenetic clustering. Dashed line indicates that loss under random extinction. Error bars around points represent the 95% confidence interval with a sample size of 1000 trees. Note differences in vertical axes.
(TIFF)

Figure S4 The additional loss of PD as function of phylogenetic clustering in 64-tip balanced Yule trees. A) mean p(ext) $= 0.25$, **B)** mean p(ext) $= 0.5$, and **C)** mean p(ext) $= 0.75$. Data points in percentage denote the amount of additional loss of projected PD (relative to random extinction) with increasing phylogenetic clustering. Dashed line indicates that loss under random extinction. Error bars around points represent the 95%

confidence interval with a sample size of 1000 trees. Note differences in vertical axes.
(TIFF)

Figure S5 The additional loss of PD as function of phylogenetic clustering in 64-tip unbalanced Yule trees. A) mean p(ext) $= 0.25$, **B)** mean p(ext) $= 0.5$, and **C)** mean p(ext) $= 0.75$. Data points in percentage denote the amount of additional loss of projected PD (relative to random extinction) with increasing phylogenetic clustering. Dashed line indicates that loss under random extinction. Error bars around points represent the 95% confidence interval with a sample size of 1000 trees. Note differences in vertical axes.
(TIFF)

Figure S6 Percentage difference in projected PD as a function of phylogenetic clustering in 64-tip coalescent trees. A) mean p(ext) $= 0.25$, **B)** mean p(ext) $= 0.5$, and **C)** mean p(ext) $= 0.75$. Data points in percentage denote the amount of additional loss of projected PD (relative to random extinction) with increasing phylogenetic clustering. Dashed line indicates that loss under random extinction. Error bars around points represent the 95% confidence interval with a sample size of 1000 trees. Note differences in vertical axes.
(TIFF)

Acknowledgments

Author Contributions

References

1. Hartmann K, Steel M (2006) Maximizing Phylogenetic Diversity in Biodiversity Conservation: Greedy Solutions to the Noah's Ark Problem. Syst Biol 55: 644–651.
2. Faller B, Pardi F, Steel M (2008) Distribution of phylogenetic diversity under random extinction. J Theor Biol 251: 286–296.
3. Nee S, May RM (1997) Extinction and the loss of evolutionary history. Science 278: 692–694.
4. Faith DP (1992) Conservation evaluation and phylogenetic diversity. Biol Conserv 61: 1–10.
5. Pybus OG, Harvey PH (2000) Testing macro-evolutionary models using incomplete molecular phylogenies. Proc R Soc Lond B 267: 2267–2272.
6. Witting, Loeschcke V L (1995) The optimization of biodiversity conservation. Biol Conserv 71: 205–207.
7. Purvis A, Agapow P-M, Gittleman JL, Mace GM (2000) Nonrandom extinction and the loss of evolutionary history. Science 288: 328–330.
8. Bininda-Emonds ORP, Gittleman JL, Steel MA (2002) The (super)tree of life: procedures, problems, and prospects. Annu Rev Ecol Syst 33: 265–289.
9. Nee S (2006) Birth-death models in macroevolution. Annu Rev Ecol Evol Syst 37: 1–17.
10. Erwin DH (2008) Extinction as the loss of evolutionary history. Proc Natl Acad Sci USA 105: 11520–11527.
11. Purvis A (2008) Phylogenetic approaches to the study of extinction. Annu Rev Ecol Evol Syst 39: 301–319.
12. Vamosi JC, Wilson JRU (2008) Nonrandom extinction leads to elevated loss of angiosperm evolutionary history. Ecol Lett 11: 1047–1053.
13. IUCN (2001) IUCN Red List Categories and Criteria Version 3.1. Available: http://www.redlist.org/info/categories_criteria2001.html. Accessed 2010 December 1.
14. von Euler F (2001) Selective extinction and rapid loss of evolutionary history in the bird fauna. Proc R Soc Lond B 268: 127–130.
15. Sechrest W, Brooks TM, da Fonseca GAB, Konstant WR, Mittermeier RA, et al. (2002) Hotspots and the conservation of evolutionary history. Proc Natl Acad Sci USA 99: 2067–2071.
16. Mooers AO, Atkins RA (2003) Indonesia's threatened birds: over 500 million years of evolutionary heritage at risk. Anim Conserv 6: 183–188.
17. Corey SJ, Waite TA (2008) Phylogenetic autocorrelation of extinction threat in globally imperilled amphibians. Divers Distrib 14: 614–629.
18. Willis CG, Ruhfel B, Primack RB, Miller-Rushing AJ, Davis CC (2008) Phylogenetic patterns of species loss in Thoreau's woods are driven by climate change. Proc Natl Acad Sci USA 44: 17029–17033.
19. Mckinney ML (1997) Extinction vulnerability and selectivity: combining ecological and paleontological views. Annu Rev Ecol Evol Syst 28: 495–516.
20. Mooers AO, Goring SJ, Turvey ST, Kuhn TS (2009) Holocene extinctions and the loss of feature diversity. , editor. Holocene Extinctions. In: Turvey ST, ed. New York: Oxford University Press. pp 263–277.
21. Roy K, Hunt G, Jablonski D (2009) Phylogenetic conservatism of extinctions in marine bivalves. Science 325: 733–737.
22. Heard SB, Mooers AO (2000) Phylogenetically patterned speciation rates and extinction risks change the loss of evolutionary history during extinctions. Proc R Soc Lond B 267: 613–20.
23. Yule GU (1924) A mathematical theory of evolution based on the conclusions of Dr J.C. Willis. Phil Trans R Soc Lond A 213: 21–87.
24. Harding EF (1971) The probabilities of rooted tree-shapes generated by random bifurcation. Adv Appl Prob 3: 44–77.
25. Steel M, McKenzie A (2001) Properties of phylogenetic trees generated by Yule-type speciation models. Math Biosci 170: 91–112.
26. Gernhard T (2008) The conditioned reconstructed process. J Theor Biol 253: 769–778.
27. Velasco JD (2008) The prior probabilities of phylogenetic trees. Biol Philos 23: 455–473.
28. Hey J (1992) Using phylogenetic trees to study speciation and extinction. Evolution 46: 627–640.
29. Heard SB (1992) Patterns in tree balance among cladistic, phenetic, and randomly generated phylogenetic trees. Evolution 46: 1818–1826.
30. Brown JKM (1994) Probabilities of evolutionary trees. Syst Biol 43: 78–91.
31. Paradis E (2005) Statistical analysis of diversification with species traits. Evolution 59: 1–12.
32. Mooers AO, Heard SB (1997) Inferring evolutionary process from phylogenetic tree shape. Q Rev Biol 72: 31–54.
33. Blum MGB, Francois O (2006) Which random processes describe The Tree of Life? A large-scale study of phylogenetic tree imbalance. Syst Biol 55: 685–691.
34. Rabosky DL, Lovette IJ (2008) Explosive evolutionary radiations: Decreasing speciation or increasing extinction through time? Evolution 62: 1866–1875.

35. Mooers AØ, Harmon LJ, Blum MGB, Wong DHJ, Heard SB (2007) Some models of phylogenetic tree shape. In: Gascuel O, Steel M, eds. Reconstructing Evolution: new mathematical and computational advances. Oxford: Oxford University Press. pp 149–170.

36. Hartmann K, Wong D, Stadler T (2010) Sampling trees from evolutionary models. Syst Biol, In press.

37. Felsenstein J (1985) Phylogenies and the comparative method. Am Nat 125: 1–15.

38. O'Meara BC, Ane C, Sanderson MJ, Wainwright PC (2006) Testing for different rates of continuous trait evolution using likelihood. Evolution 60: 922–933.

39. Pagel M (1999) Inferring the historical patterns of biological evolution. Nature 401: 877–884.

40. Revell LJ (2010) Phylogenetic signal and linear regression on species data. Methods Ecol Evol 1: 319–329.

41. Revell LJ, Harmon LJ, Glor RE (2005) Underparameterized model of sequence evolution leads to bias in the estimation of diversification rates from molecular phylogenies. Syst Biol 54: 973–983.

42. Revell LJ, Harmon LJ, Collar DC (2008) Phylogenetic signal, evolutionary process, and rate. Syst Biol 57: 591–601.

43. Simpson GG (1944) Tempo and mode in evolution. New York: Columbia University Press. 237 p.

44. Foote M (1994) Morphological disparity in Ordovician-Devonian crinoids and the early saturation of morphological space. Paleobiology 20: 320–344.

45. Schluter D (2000) The ecology of adaptive radiation. Oxford: Oxford University Press. 296 p.

46. Blomberg SP, Garland T, Ives AR (2003) Testing for phylogenetic signal in comparative data: behavioral traits are more labile. Evolution 57: 717–745.

47. Harmon LJ, Schulte II JA, Larson A, Losos JB (2003) Tempo and mode of evolutionary radiation in Iguanian lizards. Science 301: 961–964.

48. Harmon LJ, Losos JB, Davies TJ, Gillespie RG, Gittleman JL, et al. (2010) Early bursts of body size and shape evolution are rare in comparative data. Evolution 64: 2385–2396.

49. R Development Core Team (2010) R: a language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria.

50. Paradis E, Claude J, Strimmer K (2004) APE: Analyses of phylogenetics and evolution in R language. Bioinformatics 20: 289–290.

51. Harmon LJ, Weir JT, Brock CD, Glor RE, Challenger W (2008) GEIGER: investigating evolutionary radiations. Bioinformatics 24: 129–131.

52. Ceballos G, Rodriguez P, Medellin RA (1998) Assessing conservation priorities in megadiverse Mexico: mammalian diversity, endemicity, and endangerment. Ecol Appl 8: 8–17.

53. Ceballos G, Brown JH (1995) Global Patterns of mammalian diversity, endemism, and endangerment. Conserv Biol 9: 559–568.

54. Russell GJ, Brooks TM, Mckinney MM, Anderson CG (1998) Present and future taxonomic selectivity in bird and mammal extinctions. Conserv Biol 12: 1365–1376.

55. McPeek MA (2008) The ecological dynamics of clade diversification and community assembly. Am Nat 172: E270–E284.

56. Morlon H, Potts MD, Plotkin JB (2010) Inferring the dynamics of diversification: a coalescent approach. PLoS Biol 8(9): e1000493.