Ecology Letters, (2001) **4**: 464–469

REPORT

Taxon size predicts rates of rarity in vascular plants

Mark W. Schwartz[1] **and**
D. Simberloff[2]
[1]*Department of Environmental Science and Policy, 1 Shields Avenue, University of California, Davis, CA 95616, U.S.A.*
E-mail:
mwschwartz@ucdavis.edu
[2]*Department of Ecology and Evolutionary Biology, University of Tennessee, Knoxville, TN 37996, U.S.A.*

Abstract

We surveyed rarity in the vascular plants of the continental U.S.A. and Canada and the vascular plants of Hawaii to test the hypothesis that rates of rarity are independent of taxonomic group size. We demonstrated that taxonomic groups of plants with few species consistently contained fewer than the expected numbers of rare species. This pattern was apparent at the levels of genus, family, order and class. We also found that the pattern remained when we examined rates of rarity by comparing sister taxa that share a common ancestor. This pattern may arise from either differential speciation and extinction patterns or taxonomic bias in species designations (lumping and splitting). The pattern of lineages with few species demonstrating reduced rates of rarity is opposite to that previously observed in mammals and birds. If the protection of representatives from a diversity of lineages is a conservation objective, plant conservation is facilitated by the fact that relatively few species-poor lineages contain rare species.

Keywords

Hawaii, North America, phylogeny, rarity, taxonomic patterns, vascular flora.

Ecology Letters (2001) 4: 464–469

INTRODUCTION

A primary challenge for conservation biologists regarding species losses is the prediction of the effect of species extinction on patterns of global biodiversity (Purvis & Hector 2000). Nee & May (1997) demonstrated that even high rates of extinction generate little loss in phylogenetic diversity when extinction is a random process. More recently, a series of papers has demonstrated that extinction and vulnerability to extinction are not randomly distributed (McKinney 1997; Russell *et al.* 1998; Purvis & Gittleman 2000; Purvis *et al.* 2000). These studies, examining mammals and birds, have consistently shown that lineages with few species are disproportionately vulnerable to extinction, resulting in higher rates of loss of phylogenetic diversity than would be expected under random extinction. Heard & Mooers (2000), modelling the consequences of non-randomness in speciation and extinction rates, predict that the loss of phylogenetic diversity among groups such as birds and mammals can be minimized by managing sparse lineages so that these species are less likely to go extinct.

Within the context of this growing literature, there is a lack of analyses of the patterns of extinction vulnerability among groups other than mammals and birds. We used data on rarity in the flora of the U.S.A. and Canada to test the hypothesis that rarity, used as a surrogate for extinction vulnerability, is randomly distributed among vascular plants.

MATERIALS AND METHODS

To conduct this assessment, we used the vascular flora of two biogeographically distinct regions, Hawaii and the continental U.S.A. and Canada, because rare plant information for these regions is extensively documented in the Natural Heritage Database (Stein *et al.* 2000). Natural Heritage Programs were begun on a state-by-state basis by the Nature Conservancy in the 1970s in order to classify all native species by global-level rarity ranks. The Natural Heritage Database contains rarity and distribution information on the 15,789 native vascular plant species of the continental U.S.A. and Canada, and the 1208 plants of Hawaii. The Natural Heritage data cover a broader geographical region with more detail than any comparable data set. Within this database, rarity is defined by both the numerical abundance of a species (numbers of individuals or populations) and by the perceived vulnerability to extinction (Master 1991; Stein *et al.* 2000). All species within the database carry a rarity rank (G1–G5), where common species are designated by G4 and G5 ranks. Species ranks are based primarily on the number of occurrences of a species, where an occurrence is a location containing a population or subpopulation. For example, a rarity designation of G1 is reserved for species that are critically endangered or are found in very few (less than five) populations globally (Master 1991). Typically, G1 < 5, G2 = 6–20, G3 = 21–100 and G4 and G5 > 100 occurrences. Species ranks are determined jointly by Natural Heritage botanists from each of the states in which a species occurs in consultation with a national data coordination team (Stein *et al.* 2000). Natural Heritage Programs and Databases exist in all U.S. states and Canadian provinces. Over 99% of all native plants contain a rarity ranking in the Natural Heritage Database. Rarity rankings do not distinguish between intrinsic rarity and anthropogenic rarity.

We used two criteria to classify species as rare. A restrictive criterion of rarity included species in the two most vulnerable categories (G1 and G2, generally species with fewer than 10 populations or high threat), as well as species that were formerly rare (extinct, GX, or historically absent, GH). A more expansive classification of rarity included the aforementioned species plus species listed as G3 (generally species with fewer than 100 occurrences and moderate threat; roughly equivalent to "vulnerable" in the vernacular of the International Union of Conservative Nations (IUCN) rarity classification system). All patterns reported for the more restrictive measure of rarity were mirrored in the broader categorization of rarity and are not presented in detail. The taxonomy of the Natural Heritage Database follows Kartesz (1994).

To test hypotheses regarding the distribution of rarity by group size, we used two sets of analyses. In the first, we assessed patterns of rarity with respect to group size using taxonomic aggregations irrespective of phylogeny. In contrast, we assessed the linkage between group size and rates of rarity within a phylogenetic context for the second set of analyses. For the first assessment, our null hypothesis was: no difference in the likelihood of rarity across groups varying in the number of constituent species. For these analyses, we used the number of species within the study areas (the continental U.S.A. and Canada, or Hawaii) as the primary determinant of group size. To assess whether the diversity of the flora within the study regions somehow drives the observed patterns, we also analysed the relationship between rarity and group size using the estimated global number of species within families (Mabberley 1997) to determine group size.

We quantified the linkage between taxonomic group size and rate of rarity in three ways. First, a simple linear regression of rate of rarity vs. family size was used to gauge the significance of a correlation between group size and rarity. Second, in order to illustrate the magnitude and pattern of non-randomness, we compared the observed numbers of rare species with the expected numbers for plants by family size. To do this, we lumped species within family sizes (e.g. all families with one species, two, three, etc.). All species of a given family size were summed and an expected number of rare species was calculated based on the rates of rarity for the entire data set. Chi-squared tests were used to assess the magnitude of the non-random distribution of rarity based on family size categories. Lumping of some family sizes was performed so as not to violate the assumptions of a chi-squared test. If a family size (e.g. five species) resulted in a low expected number of rare species (less than five species), it was lumped together with the next larger family size (six species) until the expected number of rare taxa exceeded five. We calculated the expected number of rare taxa for the combined sample size. As a result of lumping family sizes, there are fewer groupings than there are family sizes. Third, in order to assess whether patterns observed at the family level were generalized, we compared the rates of rarity for small and large taxonomic groups for varying definitions of small, and across different taxonomic levels.

We also compared rates of rarity within a phylogenetic context. We predicted that patterns that appear for the flora as a whole should be repeatable when comparing pairs of plant families that share a common ancestor. In order to conduct this phylogenetic comparison, we used the phylogeny of vascular plants based on molecular data presented by Soltis *et al.* (1999). This phylogeny sampled a total of 158 families from our continental data set and 78 families from Hawaii, and included 82% of the angiosperm families in each flora.

For our second phylogenetic comparison, we used the Wilcoxon signed-rank test of the hypothesis that lineages with more species contain a higher proportion of rare species than the lineage with which they share a common ancestor (e.g. sister families). We restricted our analysis to all lineage pairs where the node was resolved for all shortest phylogenetic trees calculated by Soltis *et al.* (1999). This rule decreases the likelihood of making incorrect comparisons of groups that do not actually share a common ancestor. We used two variations on this sampling strategy. In the first test, we used all possible comparisons. In the second test, we restricted the comparison to the branch tips and compared only sister families. As a result of making comparisons for two data sets (U.S.A. and Canada, Hawaii), using two alternative criteria for rarity designations (GX-G2 and GX-G3) and two comparisons (all nodes, sister families), results are presented from eight tests.

RESULTS

Linear regressions of rarity by family size are positive and significant for both continental ($n = 231$, $r^2 = 0.04$, $P = 0.002$) and Hawaiian ($n = 118$, $r^2 = 0.11$, $P < 0.001$) species. Although the total amount of variation explained by these regressions is small, they are both highly significant, and most large families fall outside 95% confidence intervals for expected rates of rarity (Fig. 1). The pattern of large families containing a greater than expected number of rare species is not universal and rarity is strongly under-represented in grasses (Poaceae) and sedges (Cyperaceae) in both data sets (Table 1). Repeating the linear regression using global species richness to estimate group size for the North

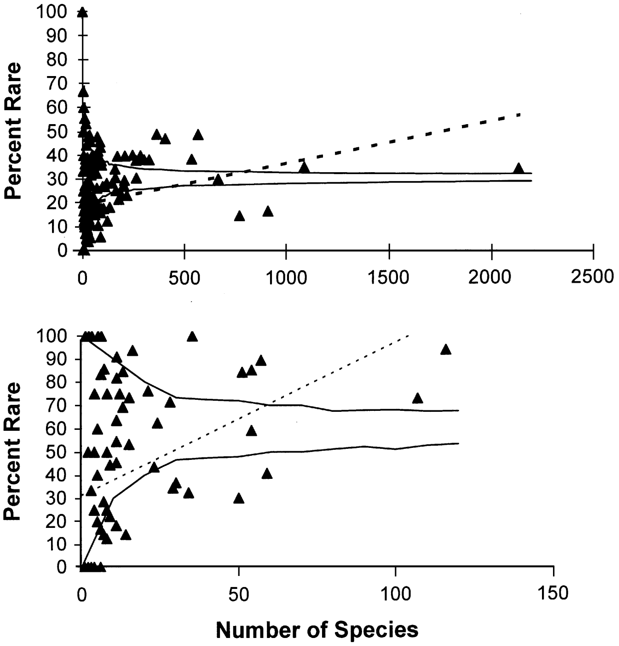

Figure 1 A scatter diagram of the number of species in plant families vs. the percentage of those species listed as rare under the GX-G3 definition of rarity for: (A) the continental U.S.A. and Canada; and (B) Hawaii. Broken lines indicate a linear fit and full lines indicate the 95% confidence interval around the overall mean rate of rarity for the region.

American flora results in a similarly significant positive correlation ($n = 205$, $r^2 = 0.03$, $P = 0.005$). This positive relationship indicates that larger families, in general, contain a higher proportion of rare species than do smaller ones.

Chi-squared tests of non-random rates of rarity grouped into categories by family size are statistically significant for continental ($\chi^2 = 365.2$, d.f. = 79, $P < 0.001$) and Hawaiian ($\chi^2 = 124.2$, d.f. = 33, $P < 0.001$) plants. Deviation from the null hypothesis results from an excess of small families that contain fewer than the expected number of rare species, as well as several large families that contain greater numbers of rare species than expected. In total, 22 of the 23 family sizes with 30 or fewer species contain fewer rare species than expected in the continental U.S.A. and Canada. In Hawaii, 17 of the 21 family sizes containing 30 or fewer species contain fewer than the expected numbers of rare species. We also used global family size estimates to calculate the observed and expected numbers of rare species for increasing group size. Once again, we observed a non-random assortment of rarity ($\chi^2 = 174.6$, d.f. = 41, $P < 0.001$), in which 83% of all small family sizes (those with <800 species in global distribution) contain fewer observed rare species than expected.

In our third analysis, we explored the generality of our observation, that small groups lack the expected numbers of rare species, by analysing rarity at taxonomic levels of classification from genus to class for the larger continental data set. The observed numbers of rare species are less than expected for small groups at the class, order, family and genus levels of classification (Table 2). These results are statistically significant at all levels except class, where a small sample size precludes a significant P value on a chi-squared test.

The results of our phylogenetic analyses show that small groups, on average, have lower rates of rarity than their more species-rich relatives. We found a significant ($P <$ 0.05) signed rank test for seven of the eight comparison sets (Table 3). All eight comparisons supported the trend of an under-representation of rare species in the smaller lineage. These patterns remained significant under additional tests that restricted the analysis by: (i) deleting comparisons in which the species numbers in paired groups were too similar (<10% different); (ii) deleting comparisons in which the species numbers in paired groups were too different (significant test for likely inequality of speciation rates as per Guyer & Slowinski 1993); and (iii) including comparisons above the level of order.

DISCUSSION

These data show that rarity is consistently positively linked to large group size for plants. Small groups contain fewer than the expected numbers of rare species in two distinct floras, using two different criteria for rarity designation, with different definitions of group size and what constitutes small groups, and at levels of classification from genus to class. Although the amount of variation explained in these regressions is small, the resulting lack of rare species in small groups is substantial. Among continental plants, the average rate of rarity in relatively large plant families (>100 species; 14.1%) is more than 3.5 times that of the smallest families (<5 species; 3.9%). Among the 1319 species in continental plant families with fewer than 30 species, we expect 78 more rare species than observed (Table 2). Similarly, there are 111 more rare species than expected in the 8410 species in the 10 largest families. For Hawaii, we expect 55 more rare species among the 439 species in families represented by fewer than 20 species, and we expect 58 fewer rare species among the 617 species in the largest 10 families. In aggregate, there are 107 fewer rare species than we would expect among the 3111 species in the continental U.S.A. in families represented by fewer than 800 species globally (Table 2).

The pattern of fewer than the expected numbers of rare species in groups containing few species is exactly the opposite of that in birds and mammals (Gittleman & Purvis 1998). This raises the question of whether birds and mammals or plants are more representative of the broad spectrum of biodiversity. With respect to mammals and birds, large body size is positively associated with rarity (Gittleman & Purvis 1998; Owens & Bennett 2000) and

Table 1 A list of large plant families and the percentage of their constituent species ranked G3 or rarer

Continental U.S. and Canada			Hawaii		
Family	No. of species	% Rare	Family	No. of species	% Rare
Asteraceae	2139	34.17	**Campanulaceae**	116	93.97
Fabaceae	1088	35.02	**Asteraceae**	107	72.90
Poaceae	910	16.48	**Poaceae**	59	40.68
Cyperaceae	770	14.55	**Lamiaceae**	57	89.47
Rosaceae	665	29.92	**Rubiaceae**	54	85.19
Brassicaceae	566	48.76	Rutaceae	54	59.26
Boraginaceae	534	38.39	**Gesneriaceae**	51	84.31
Plantaginaceae	404	47.03	**Cyperaceae**	50	30.00
Polygonaceae	362	48.9	**Caryophyllaceae**	35	100.00
Apiaceae	325	38.15	**Dryopteridaceae**	34	32.35
Lamiaceae	301	38.21			
Polemoniaceae	283	39.93			
Ranunculaceae	263	30.42			
Onagraceae	261	37.93			
Liliaceae	244	40.16			
Euphorbiaceae	217	23.04			
Orobanchaceae	206	39.81			
Ericaceae	204	29.41			
Caryophyllaceae	201	27.36			
Orchidaceae	189	23.28			
Amaranthaceae	177	21.47			
Cactaceae	169	39.64			
Malvaceae	163	25.15			
Saxifragaceae	160	30.63			
Scrophulariaceae	158	34.18			
Apocynaceae	145	28.28			
Rubiaceae	132	18.18			
Juncaceae	121	12.4			
Convolvulaceae	117	28.21			
Nyctaginaceae	108	26.85			
Solanaceae	102	17.65			

Large in this case is defined as >100 species for continental families and >30 species for Hawaiian families. The percentage of all species that are rare for the 15,041 native vascular plants of the continental U.S.A. and Canada is 30.3%. For Hawaii, 60% of 1211 native species are ranked GX-G3. Families in bold type have rates of rarity that significantly deviate from the grand mean as tested by a chi-squared goodness of fit test.

small lineage size (Gittleman & Purvis 1998; Russell *et al.* 1998). Notably, carnivores and primates have, in general, large bodies and few species. These are also the species most likely to run into conflict with humans and be endangered because of humans. In contrast, large plants (e.g. trees) are broadly distributed taxonomically such that, even if large size predicted rarity in plants, this attribute might not track strongly along phylogenetic lines. In addition, islands are often characterized by lineages of birds and mammals with few species and these tend to be highly threatened by virtue of being on islands. The distribution of plants on islands may be less taxonomically distinct, or differentiated at a lower taxonomic level (i.e. genera rather than families). Thus, two hypotheses, not mutually exclusive, for the differences in patterns of rarity among plants and some vertebrates may be that attributes that are correlated with rarity are different among the groups or that they are differentially aggregated along phylogenetic lines. Thus, it may be that the contrasting patterns reflect the differing distributions of diversity or causes of rarity or endangerment. These ideas remain to be analysed.

Two plausible factors may account for the observed pattern in plants of fewer rare species within taxonomic groups with few species. First, differential speciation and extinction rates may be linked to rarity. For example, high speciation rates in large groups may be positively correlated with high rates of rarity. This pattern would be predicted under a model of allopatric speciation where, on average, new species would be isolated, have small range sizes, and hence be more likely to be rare. Under this scenario, families

Table 2 An analysis of the observed and expected numbers of rare species in taxonomic groups of the continental U.S.A. and Canada under (A) varying definitions of "small" and (B) varying taxonomic levels

	Number of small group (% of total)	Number of species in small groups (% of all species)	Observed	Observed minus expected
(A) Size of "small"				
5	86 (38.2)	211 (1.34)	12	−17.2†
30	159 (70.7)	1319 (8.35)	104	−78.3†
250	210 (93.3)	6215 (39.4)	73	−122.9†
800‡	155 (71.4)	3111 (21.4)	320	−106.8†
(B) Taxonomic level				
Class	4 (33.3)	26 (0.17)	2	−1.52
Order	19 (23.5)	84 (0.56)	6	−5.34*
Family	95 (46.1)	382 (2.55)	24	−27.8†
Genus	1346 (80.4)	3609 (24.7)	349	−147.4†

The data for taxonomic level were sorted to ask whether small (<20 species) classes have low rates of rarity relative to the species in the remaining large classes. Species from small classes were then removed so that they did not influence the test of whether small orders were under-represented by rare species. The analysis was repeated by removing species from small classes and orders to examine small families. Species from small classes, orders and families were removed in order to analyse small genera.

*$P < 0.05$.

†$P < 0.001$.

‡Estimated family size based on global diversity.

with large numbers of species would be expected to have high rates of rarity, on average. The converse of this argument is that lineages experiencing high rates of extinction may be expected to lose their rare species first, leaving the more broadly successful common congeners as the remaining extant representatives of smaller lineages. Naturally, other models of speciation and extinction exist and multiple forces are probably at work. A significant alternative hypothesis to explain this pattern, however, is that differential species-naming conventions could also generate the pattern. If one taxonomist lumped some amount of differentiation into a single well-distributed

Table 3 Wilcoxon signed-rank tests of the association of rarity rate with group size for families of the angiosperms that share a common ancestor

	Number of comparisons	Sum of ranks of comparisons supporting hypothesis	P value
(A) Continental U.S.A and Canada			
1 Narrow definition of rarity			
Sister families	43	218	<0.005
All pairs	81	902	<0.001
2 Broad definition of rarity			
Sister families	43	303	<0.05
All pairs	81	1223	<0.05
(B) Hawaii			
1 Narrow definition of rarity			
Sister families	19	44	<0.05
All pairs	36	133	<0.005
2 Broad definition of rarity			
Sister families	19	59	<0.2
All pairs	36	156.5	<0.005

Two sets of rarity criteria are used. A narrow definition includes species ranked GX, GH, G1 or G2 in the Natural Heritage Database. The broad rarity definition includes these species plus those ranked G3 (vulnerable). Two sets of comparisons are presented. "Sister families" compares family pairs that emerge on each of the shortest trees presented by Soltis *et al.* (1999). "All pairs" compares rates of rarity for all resolved nodes within plant orders as per Soltis *et al.* (1999). Comparisons were deleted when two groups were of equal size.

species, while another split it into several species each with a small distribution, we would expect the pattern observed: large groups having an excess of rare species.

There is no absolute test to distinguish between the speciation–extinction differential hypothesis and the hypothesis that patterns of taxonomic lumping and splitting generate the observed pattern of the distribution of rarity in plants. Nonetheless, our comparison of related families suggests that speciation and extinction are likely to explain a portion of this pattern, if one accepts the assumption that closely related taxonomic groups are likely to share similar species-naming conventions. If related groups share naming conventions and this sharing generates the pattern of the distribution of rarity observed in plants, we might expect the pattern to disappear under a phylogenetic comparison of closely related groups. The pattern did not disappear with the phylogenetic assessment, but we cannot determine whether the assumption of similar naming rules is valid.

The mechanisms suggested here that may give rise to the pattern observed in plants should also be at play with vertebrates. Thus, it is doubly interesting that birds and mammals demonstrate a different pattern in the face of forces that would be expected to drive a pattern in which small groups have lower than the expected rates of rarity.

A conservation implication of this finding, irrespective of the mechanism, is that extinction among plants may result in less loss of higher taxa than is predicted under the random model of Nee & May (1997). This pattern, however, highlights two potentially different perspectives on conservation. If our goal is to conserve evolutionary potential, large lineages with many rare species may be particularly important. These large families may be where most of the speciation is occurring and these large groups have more than their share of species at risk. In contrast, if conserving representatives from divergent groups is a conservation priority, these observations suggest that the job of conserving plant diversity is smaller than we would have anticipated based on rarity rates in the flora as a whole.

ACKNOWLEDGEMENTS

We thank Larry Morse, Larry Master, Bruce Stein and Lynn Kutner of Association for Biodiversity Information (ABI) for facilitating access to the database for this analysis. We also thank J. Bergelson, A. Purvis, C. Daehler and anonymous reviewers for their comments. This work was supported by the National Science Foundation (NSF).

REFERENCES

Gittleman, J.L. & Purvis, A. (1998). Body size and species-richness in carnivores and primates. *Proc. R. Soc. London, Ser. B, Biol. Sci.*, 265 (1391), 113–119.

Guyer, C. & Slowinski, J.B. (1993). Adaptive radiation and the topology of large phylogenies. *Evolution*, 47 (1), 253–263.

Heard, S.B. & Mooers, A.O. (2000). Phylogenetically patterned speciation rates and extinction risks change the loss of evolutionary history during extinctions. *Proc. R. Soc. London, Ser. B, Biol. Sci.*, 267, 613–620.

Kartesz, J.T. (1994). *A Synonymized Checklist of the Vascular Flora of the United States, Canada and Greenland*. Timber Press Portland, OR.

Mabberley, D.J. (1997). *The Plant-Book: A Portable Dictionary of the Vascular Plants Utilizing Kubitzki's, The Families and Genera of Vascular Plants (1990–), Cronquist's, An Integrated System of Classification of Flowering Plants (1981), and Current Botanical Literature. Arranged Largely on the Principles of Editions 1–6 (1896/97–1931) of Willis's, A Dictionary of the Flowering Plants and Ferns*. Cambridge: Cambridge University Press.

Master, L.L. (1991). Assessing threats and setting priorities for conservation. *Conserv. Biol.*, 5, 555–563.

McKinney, M.L. (1997). Extinction vulnerability and selectivity: combining ecological and paleontological views. *Annu. Rev. Ecol. Syst.*, 28, 495–516.

Nee, S. & May, R.M. (1997). Extinction and the loss of evolutionary history. *Science*, 278, 692–694.

Owens, I.P.F. & Bennett, P.M. (2000). Ecological basis of extinction risk in birds: habitat loss versus human persecution and introduced predators. *Proceedings of the National Academy of Sciences*, 97, 12,144–12,148.

Purvis, A., Gittleman, J.L., Cowlishaw, G. & Mace G.M. (2000). Nonrandom extinction and the loss of evolutionary history. *Science (Washington DC)*, 288 (5464), 328–330.

Purvis, A. & Gittleman, J.L. (2000). Predicting extinction risk in declining species. *Proc. R. Soc. London, Ser. B, Biol. Sci.*, 267, 1947–1952.

Purvis, A. & Hector, A. (2000). Getting the measure of biodiversity. *Nature (London)*, 405 (6783), 212–219.

Russell, G.J., Brooks, T.M., McKinney, M.M. & Anderson, C.G. (1998). Taxonomic selectivity in bird and mammal extinctions. *Conserv. Biol.*, 12, 1365–1376.

Soltis, P.S., Soltis, D.E. & Chase, M.W. (1999). Angiosperm phylogeny inferred from multiple genes as a tool for comparative biology. *Nature*, 402, 402–404.

Stein, B.A., Kutner, L.S. & Adams, J.S. (eds). (2000). *Precious Heritage: the Status of Biodiversity in the United States*. Oxford: Oxford University Press.

BIOSKETCH

Mark Schwartz is a plant conservation ecologist. Research interests include population biology of rare plants, predicting distribution shifts of trees in response to global warming and predicting patterns of rarity.

Editor, J. Bergelson
Manuscript received 2 March 2001
First decision made 6 April 2001
Manuscript accepted 22 May 2001