

## Valuing Species on the Cheap

Laurelene Faye<sup>1</sup>, Remi Matthey-Doret<sup>2</sup> and Arne O. Mooers<sup>1</sup>

1. Biological Sciences, Simon Fraser University, 8888 University Drive, Burnaby BC V5A1S6 Canada

2. Department of Zoology, University of British Columbia, 6270 University Blvd, Vancouver BC V6T 1Z4 Canada

email: lfaye@sfu.ca, matthey@zoology.ubc.ca, amooers@sfu.ca

How do we make conservation decisions for poorly known taxa (Raphael & Molina, 2007)? Evolutionary isolation – how isolated a species is on the tree of life – has been touted as a coarse-grained but relatively cheap conservation prioritization metric (Redding & Mooers, 2010). However, prices are relative: building the robust, complete, time-calibrated phylogenies that are deemed necessary to calculate evolutionary isolation is still grueling, and time-consuming. In this issue, Curnick *et al.* (2015) suggest that shortcuts are possible, and that we might be able to get useful information on evolutionary isolation without needing to produce these full trees. If true, this would be very helpful.

Curnick and colleagues use the roughly 1000 species of understudied but ecologically important and threatened (Carpenter *et al.*, 2008) scleractinian corals as a test case. They report that prioritization lists based on taxonomic data combined with older and very sparse trees were similar to lists generated from a more comprehensive "reference" tree. Experts armed with only taxonomic data and their natural history knowledge seemed to do well at predicting which species were most evolutionarily isolated too. Given the uncertainties associated with other measures of worth, with conservation costs, and with probabilities of success, and given the lateness of the hour, perhaps we should use the precautionary principle and start estimating evolutionary isolation values for poorly sampled groups.

Curnick *et al.*'s measure of evolutionary isolation (often called Evolutionary Distinctiveness or ED) captures how many close relatives a species has: members in species-poor taxa score high, and members of species-rich taxa score low. Because random samples from larger trees should retain relative taxa sizes, smaller and sparser "random sample" trees can indeed produce similar isolation rankings to full trees (consider Curnick *et al.*'s Figure 2A and the grey squares in their Figure 3A). This is a very useful observation. However, using scores from these random sample trees to impute scores for missing species does not work as well (consider Curnick *et al.*'s Figure 2B and the grey squares in their Figure 3B). For this, it seems that phylogenetic over-dispersed sampling is required. Such sampling means that very small taxa - potentially the ones containing species with highest ED - can be retained in the sparse trees; these may be needed to produce the significant correlations that Curnick *et al.* report.

As an example, consider Curnick *et al.*'s smallest tree of 127 species (from Fukami *et al.*, 2008). We found that more than 10% of the genera and families are significantly overrepresented when compared to the scleractinian taxonomy (from Huang, 2012) using

hypergeometric tests, and specifically, smaller taxa are more over-represented (correlations of the degree of representation vs. taxon size: Spearman's rho = 0.65 and 0.44, for genera and families respectively). As suggested by Cusimano and Renner (2010), over-dispersed sampling might be common in sparse trees. This could be easily tested.

The ED measure also makes use of the ages of taxa: given two genera of the same size, members of older genus would score higher, because its members are more distantly related one to another. However, given that coral experts do fairly well by predicting that a particular taxon (a family, say) branches early in the tree, ED may be robust to imprecise age information. Simple simulations seem to support this conjecture: Pearson correlation coefficients of ED scores on simulated (Yule) trees vs. scores on versions that have been stripped of branch length information (using the common "Grafen" transformation in the R package ape) are often greater than 0.6 across a wide range of tree sizes.

Taken together, patterns of over-dispersion and the information content in topologies suggest that one might be able to produce useful ED scores using only (i) the number of subtaxa (species, genera, families, etc.) in each higher taxon (genera, family, orders, etc.) and (ii) rough ages of those higher taxa, perhaps augmented with a simple equation or two (e.g. flowing from Stadler and Steel, 2012) linking age and species richness to expected ED. This could be tested immediately, both with simulations (e.g. using code from Stadler *et al.*, 2014) and with existing large trees (e.g. for birds and mammals). If these inputs produce ED scores that are highly correlated with true values, we could roll up our sleeves and produce prioritization scores for many groups in short order. Invertebrates would indeed be a good place to start.

Curnick *et al.* apply their results to the Zoological Society of London's EDGE measure, which combines ED with an index of Global Endangerment (GE). Interestingly, only 50% of the top 50 EDGE species taken from their reference tree were consistently chosen when using the imputed scores. With a common weighting given to endangerment (GE) and a correlation of ~0.7 between the imputed and non-imputed ED scores, this seems low. Particularly high ED species or particularly endangered species must be missing from the sparser trees. Indeed, the authors claim that they are missing some particularly high ED species in their 127-species tree, and it seems that endangered species are also underrepresented (permutation test,  $p < 0.0001$ ). Whether more endangered or very high ED species are often missing from sparse trees is yet another clear question worth tackling.

All this is very exciting, and practical work like that reported by Curnick *et al.* is both long overdue and augurs well. However, we also need work along a parallel track. The popular and elegant ED measure (Redding, 2003) is but one in a basket of such metrics (see, e.g. Redding *et al.*, 2014). Unfortunately, not a single one of these metrics has a strong empirical connection to things we might actually value about biodiversity - trait diversity or trait rarity, evolutionary potential, improved ecosystem function and/or overall genetic information. People say they may have such connections, and they may; however, we rather urgently need to demonstrate what these isolation scores actually measure of value. With that information in hand, we can start to decide what

phylogenetic information we need, and what we can do without. Until then, regardless of whether our trees are cheap or expensive, *caveat emptor*.

The scripts and data we used can be found at xxxx. We thank N. Dulvy for useful comments and D. Huang for making his data available on short notice.

#### References

- Carpenter, K. E., Abrar, M., Aeby, G., Aronson, R. B., Banks, S., Bruckner, A., Chiriboga, A., Cortés, J., Delbeek, J. C., DeVantier, L., Edgar, G. J., Edwards, A. J., Fenner, D., Guzmán, H. M., Hoeksema, B. W., Hodgson, G., Johan, O., Licuanan, W. Y., Livingstone, S. R., Lovell, E. R., Moore, J. A., Obura, D. O., Ochavillo, D., Polidoro, B. A., Pretch, W. F., Quibilan, M. C., Reboton, C., Richards, Z. T., Rogers, A.D., Sanciangco, J., Sheppard, A., Sheppard, C., Smith, J., Stuart, S., Turak, E., Veron, J. E. N., Wallace, C., Weil, E. & Wood, E. (2008). One-third of reef-building corals face elevated extinction risk from climate change and local impacts. *Science*, **321**: 560-563.
- Curnick, D.J., Head, C.E.I., Huang, D., Crabbe, M.J.C., Gollock, M., Hoeksema, B.W., Johnson, K.G., Jones, R., Koldewey, H.J., Obura, D.O., Rosen, B.R., Smith, D.J., Taylor, M.L., Turner, J.R., Wren S., & Redding, D.W. (2015). Setting evolutionary-based conservation priorities for a phylogenetically data-poor taxonomic group (Scleractinia). *Animal Conservation* **XX**: xx-xx.
- Cusimano, N. & Renner, S. S. (2010). Slowdowns in diversification rates from real phylogenies may not be real. *Systematic Biology*, **59**: 458-464.
- Fukami, H., Chen, C. A., Budd, A. F., Collins, A., Wallace, C., Chuang, Y. Y., Chen, C., Dai, C., Iwao, K., Sheppard, C. & Knowlton, N. (2008). Mitochondrial and nuclear genes suggest that stony corals are monophyletic but most families of stony corals are not (order scleractinia, class anthozoa, phylum cnidaria). *Plos One*, **3**: e3222.
- Huang, D. W. (2012). Threatened reef corals of the world. *Plos One*, **7**: e34459.
- Raphael, M. G. & Molina, R. eds. (2007). Conservation of rare or little-known species: Biological, social, and economic considerations. *Island Press*, Washington.
- Redding, D. W. (2003). Incorporating Genetic Distinctness and Reserve Occupancy into a Conservation Prioritisation Approach. Masters Thesis. University of East Anglia, Norwich UK.
- Redding, D. W. & Mooers, A. O. (2010). Can systematists help decide the relative worth of bits of biodiversity? *The Systematist*, **32**: 4-8.
- Redding, D. W., Mazel, F. & Mooers, A. O. (2014). Measuring evolutionary isolation for conservation. *Plos One*, **9**: e113490.
- Stadler, T., Rabosky, D. L., Ricklefs, R. E. & Bokma, F. (2014). On age and species richness of higher taxa. *American Naturalist*, **184**: 447-455.
- Stadler, T. & Steel, M. (2012). Distribution of branch lengths and phylogenetic diversity under homogeneous speciation models. *Journal of Theoretical Biology*, **297**: 33-40.