Global Ecology and Conservation 4 (2015) 243–254

Contents lists available at ScienceDirect

Global Ecology and Conservation

journal homepage: www.elsevier.com/locate/gecco

Original research article

Factors influencing when species are first named and estimating global species richness

Mark J. Costello[a,*], Marguerita Lane[a,b], Simon Wilson[b], Brett Houlding[b]

[a] Institute of Marine Science, University of Auckland, Auckland 1142, New Zealand
[b] Discipline of Statistics, Trinity College Dublin, Ireland

ARTICLE INFO

Article history:
Received 23 April 2015
Received in revised form 5 July 2015
Accepted 5 July 2015
Available online 24 July 2015

Keywords:
Biodiversity
Biogeography
Marine
Freshwater
Fish
Taxonomy

ABSTRACT

Estimates of global species richness should consider what factors influence the rate of species discovery at global scales. However, past studies only considered regional scales and/or samples representing <0.4% of all named species. Here, we analysed trends in the rate of description for all fish species (2% of all named species). We found that the number of species described has slowed for (a) brackish compared to marine and freshwater species, (b) large compared to small sized fish, (c) geographically widespread compared to localised, (d) species occurring in the tropics and northern hemisphere compared to southern hemisphere, and (e) neritic (coastal) species compared to pelagic (offshore) species. Most (68%) of the variation in year of description was related to geographic location and depth, and contrary to expectations, body size was a minor factor at just 6% (on a standardised scale). Thus most undiscovered species will have small geographic ranges, but will not necessarily be of smaller body size than currently known species. Accordingly, global assessments of how many species may exist on Earth need to account for geographic variation.

1. Introduction

Knowing how many species are on Earth is amongst the most topical questions in biology and ecology, because it defines progress in the rate of discovery of life. However, the discovery of species is not random. The system of describing species initiated by Linnaeus began in Europe in the 1750's and spread around the world as Europeans travelled and/or had material collected from other parts of the world and brought to them for description. Knowing where new species are most likely to be found enables targeted expeditions (e.g. Brandt et al., 2007, Bouchet, 2009; Bouchet et al., 2009) and helps justify research funding applications. However, studies looking at factors influencing the rate of species discovery have been limited in spatial extent, taxonomic coverage, number of species and variables examined (Table 1).

Global scale analyses considered body size for passerine birds (Gaston and Blackburn, 1994), monogenean trematodes and crustaceans (Poulin, 2002; Martin and Davis, 2006), and body size and geographic range for terrestrial Carnivora and Primates (Collen et al., 2004) and branchiopod crustaceans (Adamowicz and Purvis, 2005). In addition to these two factors, depth distribution was studied for marine holozooplankton (species planktonic for all of their life-cycle) (Gibbons et al., 2005). However, the variables examined varied between studies and some results were contradictory. For example, half the studies found body size was a poor or insignificant predictor of year of description (Table 1). A more global perspective of a wide range of variables for a species rich taxon would provide a more accurate assessment of what factors most influence

* Corresponding author.
E-mail address: m.costello@auckland.ac.nz (M.J. Costello).

http://dx.doi.org/10.1016/j.gecco.2015.07.001

Table 1
A review of factors correlated with the rate of discovery in different taxa in the literature. The strength of correlations is indicated as ++ strong, + weak but positive, ns = not significant.

Number species	Taxon	Study area	Reference	Geographic range size	Altitude (depth) range	Body size
84	*Aphthoma* Chrysomelidae beetles	W Palaearctic	Baselga et al. (2007)	++		ns
102	Geotrupinae beetles	North and South America	Trotta-Moreu and Cabrero-Sañudo (2010)	++		ns
131	Anurans (frogs, toads)	Brazil Cerrado	Diniz-Filho et al. (2005)	++		++
156	Beetles rainforest	Australia	Stork et al. (2008)			++
232	Primates	World	Collen et al. (2004)	++		+
270	Carnivora	World	Collen et al. (2004)	++		++
297	Fleas	Holarctic, Neotropics	Krasnov et al. (2005)	++		
493	Sharks	World	Randhawa et al. (2014)	++	+	++
539	Passerine birds	South America	Blackburn and Gaston (1995)	++		+
629	Scarab dung beetles	Palaearctic	Cabrero-Sañudo and Lobo (2003)	++		+
672	Tiger moths	Brazil	Ferro and Diniz (2008)	+	+	++
792	Tardigrades limnoterrestrial	World	Guil and Cabrero-Sanudo (2007)	++		
1,131	Monogenea	World	Poulin (2002)			++
1,158	Ground beetles	Iberia	Jiménez-Valverde and Ortuño (2007)	++		++
1,176	Branchiopod crustaceans	World	Adamowicz and Purvis (2005)	++		+
1,222	Fish	Tropical east Pacific	Zapata and Robertson (2007)	++	++	+++
1,145	Mammals	Neotropics	Patterson (1994)	++		++
1,433	Reptiles, amphibians	North America, Australia	Reed and Boback (2002)			+
3,441	Scarab beetles	Australia	Allsopp (1997)	++		+
3,637	Beetles	Britain	Gaston (1991b)			+
4,000	Marine holozooplankton	World	Gibbons et al. (2005)	++	+	+
6,209	Birds	World	Gaston and Blackburn (1994)			++
21,000	Butterflies	North American	Gaston et al. (1995a,b)	++		++
24,092	Land & freshwater animals	Europe	Essl et al. (2013)	++		
32,000	Fish	World	This study	++	++	+

when and where species will be discovered than previous studies. Because fish are the most diverse group of vertebrates and present throughout the world's oceans and freshwater environments, we suggest they may be representative of the factors that will influence the rate of discovery of other species on Earth. Fish provide a wider combination of species richness, body size, environmental and global coverage (freshwater and marine), than previous studies, represent ca. 2% of all named species (Costello et al., 2013a,b), and thus may better reflect global taxonomic trends.

Here, we have compared the rates of description of all fish with their geographic location, body size, environment, and number of species in a genus (i.e. number of congeners) (Table 2). We expected that species would tend to have been discovered sooner if they were (a) more conspicuous due to their larger body size, (b) more encountered if they had a greater geographic and depth range, and occurred in high latitudes (e.g. Europe, North America), and (c) easier to distinguish if there were fewer species in the genus (differences between genera will generally be more obvious than differences between species). The results show the importance of including biogeographic patterns in estimating global species richness.

2. Methods

We compiled data on the year of description, geographic and environmental distribution, habitat, and maximum body size for all fish species from FishBase on 10th June 2011 (Froese and Pauly, 2014). This data is available in the online Supplementary Material (see Appendix A). Reuman et al. (2014) found these data were representative of animal body mass from 1 to 1,000 kg and that maximum length was a reasonable indicator of asymptotic length. Not all of the variables were available for all 32,055 species, so sample size varied (Table 3). A species geographic range was estimated by (a) the distance between its northern- and southern-most latitude, and eastern and western-most longitude, (b) the number of Food and Agriculture Organisation (FAO) fishery management areas the species was present in, and (c) the number of countries it was present in. There were 15,779 only freshwater, 1,496 only brackish, and 16,953 only marine species.

The following equations were used to predict range from the limits of a species latitude and longitude observations. If the species' longitudinal range was 360°, then geographic area (in km^2) was calculated as:

$$2\pi * 6378.127^2 *|\sin(Lat\,N.) - \sin(Lat\,S.)|,$$