Supplemental Information

The shelf life of new species

Benoît Fontaine, Adrien Perrard, Philippe Bouchet

Supplemental Experimental Procedures

The 2007 new species list, as comprehensive as possible, was compiled from several databases, including the Zoological Record online (animals), the List of Prokaryotic names with Standing in Nomenclature (bacteria), Index Kewensis (plants), MycoBank (fungi) and AlgaeBase (algae).

For each of the 600 new species, the original description was researched and specific information was recorded: species name, order, phylum, type locality, biome (aquatic - freshwater, brackish and marine - or terrestrial), year of first collection, country of affiliation of first author, type of address (institutional/personal) of first author, type of publication (book/journal) and its impact factor, number of co-authors of the description, number of specimens used in description, time in year since last cited revision of the group to which the new species belong, economic importance of the new species. Original description was found for 588 species.

Species were grouped in large ergonomic groups (called taxa in the main text, even if non-monophyletic): insects, non-insect invertebrates, plants, fungi, fish, tetrapods, algae/bacteria. To test the difference in shelf life between species and taxonomists from developed vs. emergent countries, we collected Gross Domestic Product (GDP) of first author's country and of type locality country from the CIA World Factbook [S1], and divided the new species sample in two classes of same size (GDP below and above US$10,000 for type locality country, GDP below and above US$35,000 for author country). Author status was inferred from the address as given in the publication: non institutional addresses were considered as addresses of non-professional taxonomists, institutional ones as addresses of professional taxonomists. This method definitely underestimates the number of amateur taxonomists, as these may indicate an institutional address even when they are not paid by the institution. However, it gives a robust minimum number of amateurs, since professional taxonomists will not use a personal address in their publications.

Publication types were divided in four groups: books, journals with no Impact Factor, journals with Impact Factor lower than one, journals with Impact Factor higher than one. Number of authors was divided in three classes: one, two and more than two author(s).

Time since last revision of the group to which the species belong was assessed with the bibliographic references listed in the original description, assuming that the authors would quote the references they used in their work. We recorded revisions, reviews and monographs (general or from the region of occurrence of the new species), and excluded checklists, catalogues and phylogenies, as well as supra-specific revisions. We then used three classes: revision prior to 1997, recent revision (after 1996), no revision.

Species were classified as having an economic importance if this was mentioned in the description (usually in the introduction, but also in the bibliography). This economic importance could be for biological control, because the species is harmful for man or crop species, because it is consumed by man or because it has ornamental value (e.g. orchids, shells).

In order to assess the impact of these factors on shelf life, type III ANOVAs were computed on general linear models of these factors against shelf life, under a quasi-Poisson law, to allow for overdispersion. Models were simplified following the method given in Crawley [S2]. All statistical calculations were done with R statistical software [S3], with the CAR package.

Supplemental References

S1. CIA. (2010). CIA World Factbook. Website *https://www.cia.gov/library/publications/the-world-factbook/index.html*. (Accessed 11 February 2011).

S2. Crawley, M. J. (2007. Statistical modelling. In The R Book, M. J. Crawley, ed. (Hoboken: John Wiley and Sons Ltd), pp. 323-386.

S3. R Development Core Team. (2009). R: A language and environment for statistical computing. Website *http://www.R-project.org*. (Accessed 20 August 2009)