Update

Letters

Natural history collections as sources of long-term datasets

Adrian M. Lister and Climate Change Research Group*

Natural History Museum, Cromwell Road, London SW7 5BD, UK

In the otherwise excellent special issue of *Trends in Ecology and Evolution* on long-term ecological research (*TREE* 25(10), 2010), none of the contributors mentioned the importance of natural history collections (NHCs) as sources of data that can strongly complement past and ongoing survey data. Whereas very few field surveys have operated for more than a few decades, NHCs, conserved in museums and other institutions, comprise samples of the Earth's biota typically extending back well into the nineteenth century and, in some cases, before this time. They therefore span the period of accelerated anthropogenic habitat destruction, climate warming and ocean acidification, in many cases reflecting baseline conditions before the major impact of these factors.

Natural history collections (NHCs) provide a rich source of data at the taxic and community levels, and can contribute to a wide range of studies [1]. These include biogeographic range changes (spatial and/or altitudinal) [2]; phenological shifts (e.g. in flowering time [3]); and evolutionary change (genetic or morphological). They can also (if unsorted bulk samples are available) document changes in community composition in the recent past (historical samples) and through deeper geological time (fossil samples). In addition, museum specimens provide source material for a range of genetic, biochemical, isotopic and trace-element studies into organismal responses to environmental change (e.g. use of preserved feathers to trace changes in the diet and migration of birds [4]).

NHCs comprise not only the products of opportunistic collecting but are also (particularly in the major national museums and institutions) repositories of major surveys. The Natural History Museum (NHM) in London, UK, for example, houses thousands of jars from the *Discovery* and *Challenger* marine expeditions that were collected at hundreds of stations in the late nineteenth and early twentieth centuries. These samples provide an outstanding (and largely untapped) resource for comparison with modern survey data. Collections made with a coherent and systematic sampling strategy are inevitably more likely to provide research-quality material than *ad-hoc* or point samples. Unknown or inconsistent sampling strategy can be a problem, but this is not unique to museum collections; methodological changes or gaps also occur in long-term field datasets and need to be accommodated in the analysis [5].

Of particular value, a proportion of historical and modern collections comprise time-series (i.e. the same locality and/or taxon has been regularly collected over many years). For exploration of long-term ecological responses, NHC-derived datasets can be integrated with local or regional climatic and other records, such as the Central England Temperature Record, which is continuous back to 1659 (http://hadobs.metoffice.com/hadcet/). Time-series through much longer intervals of the Earth's history are available in palaeontological collections.

Central to all such endeavours is the availability of accurate provenance data on the NHC material. Museum specimen labels and registers ideally indicate the place and date of collection. However, even when such information is available, considerable work may be required to make it accessible for research, for example, by georeferencing (establishing the latitude and longitude of) obscure place-names, and entering all records onto an electronic database [2]. Collaboration between researchers and collections managers is essential and, if resources for curation are limited, researchers should consider including collections databasing into their funding proposals.

Curators and collections managers, for their part, have a vital part to play in this process. Policies crucial for enabling collections-based research include:

Maintaining unpicked sub-samples of bulk-sampled collections.
Prioritising databasing of collections with research potential.
Facilitating responsible destructive sampling. This includes retaining duplicate imperfect specimens for the purpose, and simplifying paperwork.
Maintaining and extending time-series through continued collecting, and collaboration with modern surveys to enable acquisition of voucher specimens.
Collaborating with other museums to create, ultimately, an integrated global resource (e.g. through the SciColl initiative: http://www.scicoll.org/).

In a recent example embodying many of these aspects, the NHM, in a project funded by the UK Government, compiled a database on the extent of NHM and other British collections potentially suitable for investigating the effects of ocean acidification on marine biocalcifying organisms. The results of the project, summarizing samples accumulated globally over 200 years, are available

Corresponding author: Lister, A.M. (A.Lister@nhm.ac.uk).
* Stephen J. Brooks, Phillip B. Fenberg, Adrian G. Glover, Karen E. James, Kenneth G. Johnson, Adrian M. Lister, Ellinor Michel, Beth Okamura, Mark Spencer, John R. Stewart, Jonathan A. Todd, Eugenia Valsami-Jones, Jeremy Young.

at http://www.nhm.ac.uk/research-curation/research/historical-marine-collections/.

At a time when conserving the biosphere is a global priority, NHCs (coupled with new techniques for exploiting preserved material) are a vital resource for establishing pre-anthropogenic baselines, addressing the shifting-baseline syndrome [6], and understanding temporal trends.

References

1 Johnson, K.G. *et al.* (2011) Climate change and biosphere response: unlocking the collections vault. *Bioscience* 61, 147–153
2 Boakes, E.H. *et al.* (2010) Distorted views of biodiversity: spatial and temporal bias in species occurrence data. *PLoS Biol.* 8, e1000385
3 Robbirt, K.M. *et al.* (2011) Validation of biological collections as a source of phenological data for use in climate change studies: a case study with the orchid *Ophrys sphegodes*. *J. Ecol.* 99, 235–241
4 Inger, R. and Bearhop, S. (2008) Applications of stable isotope analyses to avian ecology. *Ibis* 150, 447–461
5 Magurran, A.E. *et al.* (2010) Long-term datasets in biodiversity research and monitoring: assessing change in ecological communities through time. *Trends Ecol. Evol.* 25, 574–582
6 Pauly, D. (1995) Anecdotes and the shifting baseline syndrome of fisheries. *Trends Ecol. Evol.* 10, 430

doi:10.1016/j.tree.2010.12.009 Trends in Ecology and Evolution, April 2011, Vol. 26, No. 4

Letters

The costs of describing the entire animal kingdom

Fernando Carbayo[1] and Antonio C. Marques[2]

[1] Escola de Artes, Ciências e Humanidades, Universidade de São Paulo, Av. Arlindo Bettio, 1000, 03828-000 São Paulo, Brazil
[2] Departamento Zoologia, Instituto de Biociências, Universidade de São Paulo, Rua Matão, Trav. 14, 101, 05508-900, São Paulo, Brazil

Biodiversity is one of the major factors underpinning human wellbeing [1,2]. Although the basic approach to the conservation and appropriate use of biodiversity requires identification of its component taxa [http://www.cbd.int/gti/], this is, in general, not well advanced. With particular regard to the animal kingdom, only ~1.4 million of the total 6.8 million species estimated are currently known [3]. By being made aware of these figures, government decision-makers might allocate resources more appropriately to alleviate the shortfall in taxonomic knowledge. We have attempted to estimate the resources required to carry out this task.

To achieve this goal, we used Brazilian expenditure per taxonomist, which we argue represents the approximate average global expenditure per taxonomist. Our data were derived from a survey of 44 taxonomists (almost 9% of the Brazilian community of employed and doctoral taxonomists [4]). These data detailed investments in their: academic education; professional positions (total income during career); training for recently qualified taxonomists; and grants for biological sampling as well as for equipping and maintaining taxonomy laboratories. The survey also included the number of new species catalogued (totaling 1,093 species for 44 taxonomists), and the number of taxonomic publications and general publications generated. The analyses dealt with the most biodiverse animal groups, and were divided into three major sets of taxa ('vertebrates', 'insects', and 'other invertebrates') to better evaluate variation. The extrapolations for the calculation of world costs were based on three main facts. First, 10% of animal species known worldwide occur in Brazil [5]. Second, the Brazilian community of taxonomists has been responsible for most taxonomic contributions to knowledge of the entire Latin American biota [6]. Also, Brazil is among the most active countries worldwide with respect to the description of species and number of taxonomists [http://www.mapress.com/zootaxa/support/Statistics.htm. and http://www.gti-kontaktstelle.de/english/taxonomy_E.html]. Third, the salaries of Brazilian researchers and professors fall close to the average salary for all professors worldwide [7].

Our data showed that a researcher described 24.8 species on average during his/her career (varying from 1 year to 46 years for researchers in a permanent position), with average expenditures per year of ~US$ 97,000. These values vary between taxonomic groups (Table 1). Therefore, we estimated that the total cost to describe unknown animal diversity would be ~US$ 263 billion, notably far higher than the US$ 5 billion proposed by Wilson [8] required to describe the entire biota.

Historically, the total amount invested to describe one species of vertebrate is three-times the amount to describe an insect, and two-times the amount for other invertebrates [4]. This is because ~31% of the total number of Brazilian animal taxonomists works on 5% of the known animal species, i.e. vertebrates [4] (on a worldwide basis, these numbers are 50% and 4%, respectively [9]). The future need for greater numbers of taxonomists working on invertebrate groups (particularly insects) is enormous. Considering the vast number of non-catalogued species and the present number of taxonomists, one invertebrate taxonomist would have to describe 292-times more species than a vertebrate taxonomist.

Based on the present average expenditure required to describe one species, we estimate that the total Brazilian and world expenditure previously spent on cataloguing animal diversity totaled ~US$ 7 and US$ 68 billion, respectively (Table 1). Of this, an average of 15% was allocated to training taxonomists. Salaries to maintain full-time taxonomists in scientific institutions amounted to 50% of the total cost. However, two-thirds of their

Corresponding author: Carbayo, F. (baz@usp.br).