9 Kapheim, K.M. *et al.* (2015) Kinship, parental manipulation and evolutionary origins of eusociality. *Proc. Biol. Sci.* Published online March 22, 2015. (http://dx.doi.org/10.1098/rspb.2014.2886)
10 Yagi, N. and Hasegawa, E. (2012) A halictid bee with sympatric solitary and eusocial nests offers evidence for Hamilton's rule. *Nat. Commun.* 3, 939
11 Leadbeater, E. *et al.* (2011) Nest inheritance is the missing source of direct fitness in a primitively eusocial insect. *Science* 333, 874e876
12 Rehan, S.M. *et al.* (2014) The costs and benefits of sociality in a facultatively social bee. *Anim. Behav.* 97, 77–85

Specimens as primary data: museums and 'open science'

Menno Schilthuizen[1,2], Charles S. Vairappan[2], Eleanor M. Slade[3], Darren J. Mann[4], and Jeremy A. Miller[1]

[1] Naturalis Biodiversity Center, Darwinweg 2, 2333 CR Leiden, The Netherlands
[2] Institute for Tropical Biology and Conservation, Universiti Malaysia Sabah, Locked Bag 2073, 88999 Kota Kinabalu, Malaysia
[3] Department of Zoology, University of Oxford, The Tinbergen Building, South Parks Road, Oxford OX1 3PS, UK
[4] Oxford University Museum of Natural History, Parks Road, Oxford OX1 3PW, UK

In 1977, Eugene Odum advocated a synthetic approach if ecology were to rise above the level of explanation afforded by independent, individual studies [1]. Today, Odum's wish is being fulfilled, and important advances are being made by synthesising data derived from great numbers of studies, either by scaling up temporally or geographically [2]. However, to allow effective, creative, and reproducible integration of ecological and environmental results, the methods and data used need to be made freely accessible and combinable. Only then can integrated ecology become a field where the ideals of 'open science' [3] fully come to fruition. Indeed, although great challenges remain [4,5], open access to ecological data, methods, and analysis is rapidly improving [6,7]. Nonetheless, we here call attention to what we perceive as one important obstacle to open data in biodiversity studies.

The 'raw data' in biodiversity research consist not of tabulations of numbers of individuals of species sampled at a particular date and place, but of the properly-labelled specimens themselves; occurrence records associated with specimens are metadata. We feel it is insufficiently appreciated that each assignment of a specimen to a particular taxon (whether a formally described species or a pragmatic 'morphospecies' [8]), is a researcher's interpretation, and therefore not a primary datum.

Because the scholarship of biodiversity includes scrutinising earlier work, evaluating what was written before, and adding new information and insight, it should always be possible to return to those specimens. They are the primary evidence for the information presented. The ability of researchers to re-examine the primary data and question the conclusions of previous work is a crucial part of what makes this a scientific activity. Especially in groups where the taxonomy is in flux, this is essential to ensure long-term comparability and vitality of data.

Corresponding author: Schilthuizen, M. (menno.schilthuizen@naturalis.nl)

0169-5347/
 http://dx.doi.org/10.1016/j.tree.2015.03.002

Unfortunately, in our experience, the accessibility of specimens sampled during biodiversity studies is problematic for two reasons. First, after publishing their results, many researchers and institutes do not systematically archive the samples of specimens that form the basis for the analyses. Specimens are either discarded or only a small reference collection is saved, leaving no way to verify the metadata. Even if specimens are stored, material from separate plots or dates are often pooled to reduce storage space [9], rendering valuable information irretrievable.

We therefore suggest that it become accepted policy in ecological research that full, unadulterated collections of all specimens from a study be deposited in a natural history collection. This is common practice in other areas of specimen-based biological research, such as taxonomy and palaeontology. Public natural history collections increasingly make the content of their collections databases available through the Global Biodiversity Information Facility (GBIF), which should facilitate retrieval and verification of specimens as well as reuse of the associated metadata [10]. Moreover, the specimens would then be available for obtaining additional information (such as genetic and morphometric data, and sometimes even information about ecological interactions [11]).

The second reason for the inaccessibility of specimens, however, lies with those same publicly-accessible collections. Natural history museums often appear unable or reluctant to assume a custodian's role as repositories for bulk samples from ecological studies. This is understandable in view of the traditional focus of natural history museums on systematics and biogeography, which gives rise to a desire to maximise the information density of their holdings by giving priority to previously unrepresented species or localities. Given the universal features of species-abundance distributions, however, biodiversity research will yield samples that are dominated by common and widespread species. Faced with space limitations, and a lack of funding and staff to be able to curate and maintain large ecological collections, museums tend to refuse, cherry-pick, or even dispose of such bulk collections.

We argue that by adhering only to their traditional role, natural history museums are missing an opportunity to expand their scientific reach and relevance. Specifically, we advocate that they should act as custodians of biological field samples, including entire collections of samples from biodiversity studies, either by storing the material themselves or by setting up dedicated repositories under their supervision. We also suggest that museums develop collection management policies that enable the scientific community to access the increasing number of specimens needed to realise the open science concept. For their part, ecologists should be expected to provide properly prepared and labelled specimens.

We are aware that the adoption of these roles will necessitate a paradigm shift in natural history museum practice, and we hope this will take place as museums address the question of how to grow collections for today's and tomorrow's scientific challenges [12]. We acknowledge that the maintenance of bulky collections of specimens will require additional funding because natural history collections are already underfunded and understaffed, even for playing their current role [13]. Many research funding organisations are, however, allocating funds for data storage, and we suggest that ecologists should include in this the housing and curation of their samples for long-term storage within natural history museums. Because most research grants are limited to only a few years, long-term storage to ensure longevity of the collections remains a contentious issue, and we encourage ecologists and museum managers alike to address these issues at the level of national as well as international research infrastructures.

Acknowledgements

We gratefully acknowledge the ~40 colleagues who debated this issue on ResearchGate, prompted by a question from M.S. and thereby helped shape our views. The (ongoing) discussion can be viewed at http://www.researchgate.net/post/Should_natural_history_museums_keep_large_series_of_common_species.

References

1 Odum, E.P. (1977) The emergence of ecology as a new integrative discipline. *Science* 195, 1289–1293
2 Fraser, L.H. *et al.* (2012) Coordinated distributed experiments: an emerging tool for testing global hypotheses in ecology and environmental science. *Front. Ecol. Environ.* 11, 147–155
3 Nielsen, M. (2011) *Reinventing Discovery: The New Era of Networked Science*, Princeton University Press
4 Reichman, O.J. *et al.* (2011) Challenges and opportunities of open data in ecology. *Science* 331, 703–705
5 Vines, T.H. *et al.* (2014) The availability of research data declines rapidly with article age. *Curr. Biol.* 24, 94–97
6 Hudson, L.N. *et al.* (2014) The PREDICTS database: a global database of how local terrestrial biodiversity responds to human impacts. *Ecol. Evol.* 4, 4701–4735
7 Scholes, R.J. *et al.* (2008) Towards a global biodiversity observing system. *Science* 321, 1044–1045
8 Miller, J.A. *et al.* (2014) Cyberdiversity: improving the informatic value of diverse tropical arthropod inventories. *PLoS ONE* 9, e115750
9 Slade, E.M. *et al.* (2011) Biodiversity and ecosystem function of tropical forest dung beetles under contrasting logging regimes. *Biol. Conserv.* 144, 166–174
10 Berendsohn, W.G. *et al.* (2010) Recommendations of the GBIF task group on the global strategy and action plan for the mobilization of natural history collections data. *Biodivers. Inform.* 7, 67–71
11 Jurado-Rivera, J.A. *et al.* (2009) DNA barcoding insect–host plant associations. *Proc. Biol. Sci.* 276, 639–648
12 Krishtalka, L. and Humphrey, P.S. (2000) Can natural history museums capture the future? *Bioscience* 50, 611–617
13 Dalton, R. (2003) Natural history collections in crisis as funding is slashed. *Nature* 423, 575