Scientific names are a different matter. They act as placeholders for hypotheses referring to *taxa*. Given that taxonomy, as a scientific field, is not about establishing immutable understanding, scientific names have, and will continue, to change, driven by new evidence that compels taxonomies to adapt. There is no way to avoid this: it is the nature of doing science. The immediate consequences are obvious. Often, the same taxon is referred to by different names (synonyms). Taxonomists may also disagree on the 'taxonomic concept' behind a given name (e.g. homonyms or *pro parte* synonyms). Hence, differences in the interpretation and naming of taxa are here to stay as long as taxonomy (as we know it) endures. The problem of handling taxonomic 'instability' in biodiversity databases has been extensively debated [5], but there remains little doubt of the benefits of making such information available [6]. As long as names are tied to voucher specimens (and, in most cases, they are, because museums are the main data providers), records can always be checked by end-users.

In a scenario where publishers become the major providers of species occurrence records, a wealth of observational (non-specimen-based) data will be available to database aggregators. The trouble is that the onus of maintaining public collections of specimens falls upon authors, not publishers. Given that this is far beyond the capacity of most research groups and institutions, and because many museums have ceased to accept new collections owing to limited funds, the link between specimens and archived names will be broken. With no voucher specimens, problems arising from bad taxonomic practice and inaccurate identification of species are likely to increase [7]. Does it mean that non-specimen-based data should be simply dismissed as a potential source for biodiversity databases? We do not think so, but it will require a drastic change in the editorial policies of scientific journals. A high percentage of manuscripts providing observational records of species give no supporting information on the methods used to identify the organisms studied [8]. Without this information, readers cannot judge the quality of the data, rendering them mostly useless and increasing the likelihood of eventual errors to propagate.

Similar to the description of PCR conditions being mandatory for the publication of molecular data, editors and referees of ecological journals should persuade authors of manuscripts involving species inventories to: (i) submit raw data; and (ii) provide in the 'Methods' section of their manuscript a detailed list of the scientific literature, keys, names of experts and any other source of information from which they have derived the scientific names. Both conditions should be a prerequisite for publication. Knowing whether the species recorded were determined by reference to an outdated work or that they were identified by a leading expert in the group, makes all the difference.

Acknowledgements
MB is funded by SFRH/BPD/40073/2007 from Fundação para a Ciência e a Tecnologia (FCT). AMS and MB were partially funded by PTDC/MAR/104169/2008 from FCT. We thank T. Branco for his valuable comments.

References
1 Huang, X. and Qiao, G. (2011) Biodiversity databases should gain support from journals. *Trends Ecol. Evol.* 26, 377–378
2 Yesson, C. *et al.* (2007) How global is the Global Biodiversity Information Facility? *PLoS ONE* 2, e1124
3 Whitlock, M.C. (2011) Data archiving in ecology and evolution: best practices. *Trends Ecol. Evol.* 26, 61–65
4 Wheeler, Q.D. (ed.) (2008) *The New Taxonomy*, CRC Press
5 Harris, D.J. (2003) Can you bank on GenBank? *Trends Ecol. Evol.* 18, 317–319
6 Patterson, D.J. *et al.* (2010) Names are key to the big new biology. *Trends Ecol. Evol.* 25, 686–691
7 Guralnick, R.P. *et al.* (2007) Towards a collaborative, global infrastructure for biodiversity assessment. *Ecol. Lett.* 10, 663–672
8 Bortolus, A. (2008) Error cascades in the biological sciences: the unwanted consequences of using bad taxonomy in ecology. *Ambio* 37, 114–118

0169-5347/$ – see front matter
doi:10.1016/j.tree.2011.10.004 Trends in Ecology and Evolution, January 2012, Vol. 27, No. 1

Biodiversity data sharing is not just about species names: response to Santos and Branco

Xiaolei Huang[1] and Gexia Qiao[1,2]

[1] Key Laboratory of Zoological Systematics and Evolution, Institute of Zoology, Chinese Academy of Sciences, 1 Beichen West Road, Chaoyang District, Beijing 100101, China
[2] National Zoological Museum of China, Institute of Zoology, Chinese Academy of Sciences, 1 Beichen West Road, Chaoyang District, Beijing 100101, China

Santos and Branco [1] have responded to our recent letter in *TREE* [2], in which we proposed that adoption of a data archiving policy by biodiversity and conservation journals would sustainably increase data quantity and quality of biodiversity databases. They agree with our proposal and express further concern about the effect of 'the scientific name problem' on the quality of species records in databases. We are happy to see their suggestion on editorial policy of journals with regard to species inventory; however, we think they misinterpret our original idea to some extent, and some points they make need to be further illustrated.

Corresponding author: Qiao, G. (qiaogx@ioz.ac.cn).

Santos and Branco [1] claim that we 'anticipate that by adopting data archiving policies, publishers would turn their repositories into major feeders of biodiversity data aggregators'. In fact, we did not make such predictions in our article at all. We argue that implementation of joint data archiving policy by journals and databases would be a win-win strategy for both sides and a sustainable (quantity and quality) methodology for biodiversity data collection. If so, biodiversity databases currently using natural history collections as a major data resource (e.g. GBIF) [3] can have another efficient and sustainable data resource. They also claim that 'publishers become the major providers of species occurrence records'. However, we think the authors themselves, i.e. the researchers, are the data providers at all times, rather than journals or publishers. Besides, their argument about GenBank in the first paragraph has only tangential relevance to our original idea. We used the success of GenBank as a model merely to argue the importance of a routine data archiving policy by journals for data collection.

We also find their claim that 'because many museums ceased to accept new collections due to limited funds, the link between specimens and archived names will be broken' is hard to understand, and its logical relationship to the main argument of the third paragraph is not distinct. We can understand species names associated with non specimen-based data (e.g. occurrence records from bird watching) are based on observers' judgments instead of voucher specimens, which may generate more taxonomic inaccuracies. However, we cannot understand their argument that species names and data from museums would not be linked to specimens. Perhaps they are too pessimistic about the future of natural history museums [4,5].

We agree with Santos and Branco [1] that 'taxonomic instability' or 'the species name problem' is a problem for quality of name-based records in biodiversity databases. Maybe it is easy for taxonomists who well know the revision history of species to resolve situations where data (e.g. occurrence records) on the same organism are labeled with different names, but it is much harder for non-taxonomist database users to do this successfully. Fortunately, some promising methods for overcoming this problem, such as taxonomic indexing, are beginning to emerge [6–8]. However, considering the widespread utility of such methods would take time to implement, the rigorous data archiving policies of journals, as Santos and Branco suggest, could help to overcome problems related to species names, especially for non specimen-based data.

Although the 'name problem' is a problem deserving attention, biodiversity data sharing involves more than species names. For example, if it is agreed that biodiversity databases and journals should adopt a joint data archiving policy, there should be careful thought about how to ensure the interests (e.g. credit and priority in utilization) of data providers, i.e. the authors. This is important for fostering a culture of sharing [9], and ensures that biodiversity data are shared in a truly open way. There have been several good examples of publication-related data archiving and sharing [10,11]. As a quickly growing data repository, Dryad (http://datadryad.org) aims to host any kind of data (format cannot be standardized) underlying peer-reviewed articles in the basic and applied biosciences [11]. However, for biodiversity data such as occurrence records and range maps, data can and should be standardized for the sake of convenient reuse. Data reuse is the fundamental goal of data sharing. Therefore, using standard data structures or formats by biodiversity databases is very important for exploring, comparing and integrating data not only from natural history collections, but also from publications.

Acknowledgements

We thank Paul Craze for valuable suggestions. X.L.H. is funded by NSFC-30900133 and O529YX5105 from the Key Laboratory of Zoological Systematics and Evolution, Chinese Academy of Sciences and G.X.Q. by NSFC-31025024, 31061160186 and 30830017.

References

1 Santos, A.M. and Branco, M. (2011) The quality of name-based species records in databases. *Trends Ecol. Evol.* 27, 6–7
2 Huang, X. and Qiao, G. (2011) Biodiversity databases should gain support from journals. *Trends Ecol. Evol.* 26, 377–378
3 Edwards, J.L. *et al.* (2000) Interoperability of biodiversity databases: biodiversity information on every desktop. *Science* 289, 2312–2314
4 Graham, C.H. *et al.* (2004) New developments in museum-based informatics and applications in biodiversity analysis. *Trends Ecol. Evol.* 19, 497–503
5 Lister, A.M. *et al.* (2011) Natural history collections as sources of long-term datasets. *Trends Ecol. Evol.* 26, 153–154
6 Polaszek, A. *et al.* (2005) A universal register for animal names. *Nature* 437, 477
7 Patterson, D.J. *et al.* (2006) Taxonomic indexing—extending the role of taxonomy. *Syst. Biol.* 55, 367–373
8 Patterson, D.J. *et al.* (2010) Names are key to the big new biology. *Trends Ecol. Evol.* 25, 686–691
9 Kueffer, C. *et al.* (2011) Fame, glory and neglect in meta-analyses. *Trends Ecol. Evol.* 26, 493–494
10 Schofield, P.N. *et al.* (2009) Post-publication sharing of data and tools. *Nature* 461, 171–173
11 Vision, T.J. (2010) Open data and the social contract of scientific publishing. *Bioscience* 60, 330–331

0169-5347/$ – see front matter
doi:10.1016/j.tree.2011.10.007 Trends in Ecology and Evolution, January 2012, Vol. 27, No. 1