remaining for biological conservation. How are we to be heard? I suggest that every scientific society with large numbers of members who consider themselves organismal biologists pass a resolution at their next national or international meeting and send it to their national funding agencies, such as the NSF. The resolutions should clearly state concerns about the loss of biodiversity, the short time left to document the natural history of the planet, the value of understanding diversity, and the need for substantial increases in funding for those programs promoting the discovery and description of biological diversity. These programs range from the support of inventories and systematics to ecology and behavior, biological research collections, and all aspects of field biology.

If we do not muster the effort to do this, biology will be reduced to studies of DNA and concepts based on ignorance about our natural world. Natural history is the soul of biology.

Reference

1 Greene, H.W. (2005) Organisms in nature as a central focus for biology. *Trends Ecol. Evol.* 20, 23–27

Data sharing in ecology and evolution

Cynthia S. Parr[1] and Michael P. Cummings[2]

[1]Human–Computer Interaction Lab, University of Maryland, College Park, MD 20742, USA
[2]Center for Bioinformatics and Computational Biology, University of Maryland, College Park, MD 20742, USA

The rapid advancement of fields such as molecular biology, genomics and molecular evolution is due, in large part, to pervasive data sharing. Ecology, evolution (other than molecular), conservation biology, animal behavior and related fields have yet to enjoy similar growth; further data sharing could transform these fields, as it has others [1,2]. However, individual scientists must first recognize the benefits and see their way past perceived barriers.

As in other fields, sharing data that support publications (in useful formats and in community-accepted archives) facilitates the scientific ideals of replication, building on previous work and syntheses [1]. Also, most research in ecology and evolution is publicly funded, so one could argue that the data belong to the public and thus should be publicly available. Sharing data provides additional return on that investment and removes the need for them to be collected again. For example, the *Iris* flower measurement data of Anderson [3] were used soon after publication by Fisher [4] to illustrate discriminant functions, and are probably still the most-used data in machine-learning research. The larger implications of data sharing are also important. Can researchers morally justify keeping data private if these data could speed solutions to environmental and conservation challenges? Participants in the new Conservation Commons Initiative (http://conservationcommons.org) think not.

Why is data sharing not yet common practice? A recent National Center for Ecological Analysis and Synthesis (NCEAS) survey is exploring attitudes in detail (S. Findlay, personal communication) but two obvious reasons are cited in [5]: (i) researchers desire to use their data for subsequent work without competition; and (ii) they believe that there are logistical barriers to data sharing.

Withholding data for possible future gain is short-sighted, because the academic reward system favors data sharing: the value of data increases in proportion to its use by others, with direct consequences in perceptions of research importance and in objective measures (e.g. citation rate). These perceptions and measures are used both formally and informally as criteria for publication, grant funding and career advancement.

Logistical barriers to data sharing are also illusory. Convenient means by which to share data already exist. One can submit supplementary files to journals, post data on institutional websites, such as the Digital Repository at the University of Maryland (http://drum.umd.edu), or post files on project websites. Infrastructure and tools [6,7] are being developed that support data sharing in, and use of formal ecological repositories and registries (Table 1). Ecological societies are also working to achieve consensus on institutional goals and policies related to data sharing [5,8]. New methods of discovery and automated data integration [e.g. [9]; the Online Research Information Environment for the Life Sciences (ORIEL) project, http://www.oriel.org/] can take advantage of ontologies for animal behavior (http://ethodata.comm.nsdl.org), animal life history (http://animaldiversity.ummz.umich.edu/site/about/technology/) and ecology (http://cvs.ecoinformatics.org/cvs/cvsweb.cgi/seek/). Active data sharing itself fosters increased standardization, as the best-annotated or -collected data are more likely to be re-used and cited.

Ecology and evolution should be part of the larger synthetic, multidisciplinary movement. In the USA,

Corresponding author: Parr, C.S. (csparr@umd.edu).
Available online 30 April 2005

Table 1. Examples of ecological and evolutionary data registries[a], institutional repositories[b] and topical repositories[c,d]

	Data sets	URL
Registries[a]		
Global Biodiversity Information Facility portal	343[e]	http://www.gbif.net
National Biological Information Infrastructure	17 000	http://www.nbii.gov
Knowledge Network for Biocomplexity	1500[f]	http://knb.ecoinformatics.org/index.jsp
Institutional or journal repositories[b]		
NCEAS	72	http://knb.ecoinformatics.org/knb/style/skins/nceas/
Ecological Archives data papers (ESA journals)	6	http://www.esapubs.org/archive/archive_D.htm
Topical repositories[c]		
Interaction Web Database	74 webs	http://www.nceas.ucsb.edu/interactionweb
TreeBase	2869 phylogenetic trees	http://www.treebase.org
Global population dynamics database	5000 time series	NERC Centre for Population Biology http://cpbnts1.bio.ic.ac.uk/gpdd/
VegBank	19 000 plots	http://vegbank.org/

[a]Providing access to metadata and pointers to data stored elsewhere.
[b]Archived data sets.
[c]Specific kinds of archived data sets in standardized file formats.
[d]Includes only sources with online access to machine-readable data and metadata; data sets are counted or self-reported as of 23 February 2005. Data sets in repositories can also be represented in registries.
[e]Data sets are typically museum collections; the total number of records now exceeds 45 million.
[f]Includes Long-term Studies Section Data Registry (of ESA); for example: 567 data sets from the Long Term Ecological Research (LTER), 434 from the University of California Natural Reserve System and 193 from the Organization of Biological Field Stations.

researchers at NCEAS and the new National Evolutionary Synthesis Center (NESCent) and the National Ecological Observatory Network (NEON) are forging ahead with research that relies on shared data. Data shared as benchmark data sets (e.g. [10]) can kick-start innovation by providing well defined challenges to computer scientists and informatics experts. The resulting technology can speed progress by ecologists and evolutionary biologists.

With substantial benefits for individuals, scientific communities, and society as a whole, the time for data sharing has come. It is up to us as individuals to take advantage of the many opportunities to share data, to make use of that data, and to support the development of related tools and data manipulation techniques.

References

1 National Research Council (2003) *Sharing Publication-related Data and Materials: Responsibilities of Authorship in the Life Sciences*, The National Academies Press
2 Insel, T.R. *et al*. (2003) Neuroscience networks: data-sharing in an information age. *PLoS Biol.* 1, 9–11
3 Anderson, E. (1935) The irises of the Gaspe Peninsula. *Bull. Am. Iris Soc.* 59, 2–5
4 Fisher, R.A. (1936) The use of multiple measurements in taxonomic problems. *Ann. Eugen.* 7, 179–188
5 Palmer, M.A. *et al*. (2004) *Ecological Science and Sustainability for a Crowded Planet*, Ecological Society of America (http://www.esa.org/ecovisions)
6 Michener, W.K. (2003) Building SEEK: the Science Environment for Ecological Knowledge. *DataBits: Elect. Newslett. Inf. Manag.* 3 (http://intranet.lternet.edu/archives/documents/Newsletters/DataBits/03spring/)
7 Cotter, G. *et al*. (2004) Integrated science for environmental decisionmaking: the challenge for biodiversity and ecosystem informatics. *Data Sci. J.* 3, 38–59
8 Silver, S. (2004) Editorial: publishing for the digital age. *Front. Ecol. Environ.* 10, 507
9 Caragea, D. *et al.* (2004) Learning classifiers from semantically heterogeneous data. In *Proceedings of the Third International Conference on Ontologies, DataBases and Applications of Semantics for Large Scale Information Systems* (ODBASE'04) (Meersman, R. and Tari, Z., eds), pp. 963–980, Springer-Verlag
10 Plaisant, C. (2004) The challenge of information visualization evaluation. *Proc. Adv. Vis. Interf.* 2004, 320–327

A new eusocial vertebrate?

Kevin R. Foster[1,2] and Francis L.W. Ratnieks[2,3]

[1]Department of Ecology and Evolutionary Biology, Rice University, MS 170, 6100 Main St, Houston, TX 77005, USA
[2]Current address: Wissenschaftskolleg zu Berlin, Institute for Advanced Study, Wallotstrasse 19, 14193, Berlin, Germany
[3]Laboratory of Apiculture and Social Insects, Department of Animal and Plant Sciences, University of Sheffield, UK, S10 2TN

Insect eusociality is well known and is characterized by individuals (workers) that forgo direct reproduction to rear siblings. In most eusocial species, workers and reproductives (i.e. queens) are morphologically distinct and, in some species, such as the fire ant *Solenopsis invicta*, the workers are permanently sterile. In vertebrates, young naked mole rats *Heterocephalus glaber* are

Corresponding author: Foster, K.R. (krfoster@rice.edu).

Available online 17 May 2005