

Letter

Solutions for Archiving Data in Long-Term Studies: A Reply to Whitlock *et al.*

James A. Mills,^{1,53,*}
 Céline Teplitsky,^{2,3,53}
 Beatriz Arroyo,⁴
 Anne Charmantier,³
 Peter. H. Becker,⁵
 Tim R. Birkhead,⁶
 Pierre Bize,⁷
 Daniel T. Blumstein,⁸
 Christophe Bonenfant,⁹
 Stan Boutin,¹⁰
 Andrey Bushuev,¹¹
 Emmanuelle Cam,¹²
 Andrew Cockburn,¹³
 Steeve D. Côté,¹⁴
 John C. Coulson,¹⁵
 Francis Daunt,¹⁶
 Niels J. Dingemanse,^{17,18}
 Blandine Doligez,⁹
 Hugh Drummond,¹⁹
 Richard H.M. Espie,²⁰
 Marco Festa-Bianchet,²¹
 Francesca D. Frentiu,²²
 John W. Fitzpatrick,²³
 Robert W. Furness,²⁴
 Gilles Gauthier,¹⁴
 Peter R. Grant,²⁵
 Michael Griesser,²⁶
 Lars Gustafsson,²⁷
 Bengt Hansson,²⁸
 Michael P. Harris,¹⁶
 Frédéric Jiguet,²
 Petter Kjellander,²⁹
 Erkki Korpimäki,³⁰
 Charles J. Krebs,³¹
 Luc Lens,³²
 John D.C. Linnell,³³

Matthew Low,³⁴
 Andrew McAdam,³⁵
 Antoni Margalida,³⁶
 Juha Merilä,³⁷
 Anders P. Møller,³⁸
 Shinichi Nakagawa,³⁹
 Jan-Åke Nilsson,²⁷
 Ian C.T. Nisbet,⁴⁰
 Arie J. van Noordwijk,⁴¹
 Daniel Oro,⁴² Tomas Pärt,³⁴
 Fanie Pelletier,²¹
 Jaime Potti,⁴³ Benoit Pujol,¹²
 Denis Réale,⁴⁴
 Robert F. Rockwell,⁴⁵
 Yan Ropert-Coudert,⁴⁶
 Alexandre Roulin,⁴⁷
 Christophe Thébaud,¹²
 James S. Sedinger,⁴⁸
 Jon E. Swenson,⁴⁹
 Marcel E. Visser,⁴¹
 Sarah Wanless,¹⁶
 David F. Westneat,⁵⁰
 Alastair J. Wilson,⁵¹ and
 Andreas Zedrosser⁵²

In our recent paper [1], we discussed some potential undesirable consequences of public data archiving (PDA) with specific reference to long-term studies and proposed solutions to manage these issues. We reaffirm our commitment to data sharing and collaboration, both of which have been common and fruitful practices supported for many decades by researchers involved in long-term studies. We acknowledge the potential benefits of PDA (e.g., [2]), but believe that several potential negative consequences for science have been underestimated [1] (see also [3,4]). The objective of our recent paper [1] was to define practices to simultaneously maximize the benefits and minimize the potential unwanted consequences of PDA.

Commenting on our paper, several former and current editors of major ecology and evolution journals [5] acknowledge the

need to improve data archiving practices to account for the concerns presented in [1]. The fact that editors of several journals were willing to comment on our paper underlines the importance of this issue and we are keen to continue this dialogue to identify potential solutions. Following our [1] and Roche *et al.*'s [6] suggestions, Whitlock *et al.* [5] endorse as good practice longer embargos (5 years) and encourage cooperation or collaboration with data providers. Both steps are major advances as many of the Principal Investigators (PIs) in [1] have been denied longer embargos, and the practice of consulting PIs to ensure that data files are properly interpreted is not a formal policy in any scientific journal.

We welcome these positive developments but underline three concerns, two of which extend beyond the purview of individual journals.

Whitlock *et al.* [5] mention that current policies 'require only that authors make available the data necessary to recreate the analyses and results in the published manuscript'. For an article that includes an analysis based on a pedigree and individual data or on lifetime reproductive success and potential predictor variables, this requirement involves providing a detailed database of the breeding performance of individuals and their progeny over decades. The costs of data gathering, including resources beyond monetary ones, are borne by the data providers and their institutions not by those who would use the data; consequently, providing such extensive datasets is sustainable if the data are used only to verify the original analysis. Extending an embargo to 5 years for such data is a good step, but for studies that extend over decades a longer embargo is warranted, notably to further encourage potential users to contact PIs to get the latest version of the data, and ideally collaborate.

Databases from long-term studies are an evolving infrastructure that underpins

numerous publications. New data are added each year and errors and omissions are corrected regularly. Over time, archives often include various versions of fragmented datasets that: (i) could be combined by others in ways that the data collectors were already doing or planning to do themselves; or (ii) may differ from each other in ways that are likely to lead to misinterpretation of the data. A single journal's PDA policy cannot ensure that data from long-term studies are not misused. It must be a community decision. Some potential solutions include archiving at institutional servers with separate policies for the distribution of data necessary to reproduce previously published analyses and data requests for additional analyses. The additional analyses would require collaboration with the PI.

Finally, journal editors do not control the policies of funding agencies but their stature in the community can be influential. Whitlock *et al.* [5] suggest that funders should set standards for openness. However, long-term studies typically involve several grants and multiple funding agencies, sometimes from different countries. Hence, any discrepancy between their policies can lead to potentially insoluble conflict. Institutions that fund a significant proportion of the research, potentially over decades, may also question the value of continued funding if the data are freely available to individuals from other organizations.

We are encouraged by the letter from Whitlock *et al.* [5], but believe that there are additional issues that need to be addressed. Some of these may be solved by a more explicit and flexible policy on longer embargos, data storage on institutional servers, and involvement of the principal investigators in new analyses using the data they produced, through collaboration or reviews. We hope that this important dialogue will continue.

¹10527A Skyline Drive, Corning, NY 14830, USA

²CESCO, UMR7204 Sorbonne Universités-MNHN-CNRS-UPMC, CP51, 55 rue Buffon, 75005 Paris, France

³Centre d'Ecologie Fonctionnelle et Evolutive UMR 5175, Campus CNRS, 1919 route de Mende, 34293 Montpellier Cedex 5, France

⁴Instituto de Investigación en Recursos Cinegéticos (IREC) (CSIC-UCLM-JCCM), Ronda de Toledo s/n, 13005 Ciudad Real, Spain

⁵Institute of Avian Research, "Vogelwarte Helgoland", An der Vogelwarte 21 D26386 Wilhelmshaven, Germany

⁶Department of Animal and Plant Sciences, University of Sheffield, Sheffield, UK

⁷Institute of Biological and Environmental Sciences, University of Aberdeen, Aberdeen, UK

⁸Department of Ecology and Evolutionary Biology, University of California, 621 Young Drive South, Los Angeles, CA 90095-1606, USA

⁹CNRS, Université Lyon 1, Université de Lyon, UMR 5558, Laboratoire Biométrie et Biologie Évolutive, 43 boulevard du 11 Novembre 1918, F-69622 Villeurbanne Cedex, France

¹⁰Department of Biological Sciences, University of Alberta, Edmonton, AB T6G 2E9, Canada

¹¹Department of Vertebrate Zoology, Faculty of Biology, Lomonosov Moscow State University, Leninskie Gory 1/12, 119234 Moscow, Russia

¹²UMR 5174 EDB Laboratoire Évolution et Diversité Biologique, CNRS, ENFA, Université Toulouse 3 Paul Sabatier, Toulouse 31062 Cedex 9, France

¹³Department of Evolution, Ecology, and Genetics, Research School of Biology, The Australian National University, Canberra, Australia

¹⁴Département de Biologie and Centre d'Études Nordiques, Université Laval, 1045 avenue de la Médecine, QC G1V 0A6, Canada

¹⁵29 St Mary's Close, Shincliffe, Durham DH1 2ND, UK

¹⁶Centre for Ecology and Hydrology, Bush Estate, Penicuik EH26 0QB, UK

¹⁷Behavioural Ecology, Department of Biology, Ludwig Maximilian University of Munich, Planegg-Martinsried, Germany

¹⁸Evolutionary Ecology of Variation Research Group, Max Planck Institute for Ornithology, Seewiesen, Germany

¹⁹Departamento de Ecología Evolutiva, Instituto de Ecología, Universidad Nacional Autónoma de México, AP 70-275, México DF 04510, México

²⁰Technical Resource Branch, Saskatchewan Ministry of Environment, 3211 Albert Street, Regina, SK S4S 5W6, Canada

²¹Département de Biologie, Université de Sherbrooke, 2500 boulevard de L'Université, Sherbrooke, QC J1K 2R1, Canada

²²School of Biomedical Sciences and Institute of Health and Biomedical Innovation, Queensland University of Technology, Kelvin Grove, QLD 4059, Australia

²³Cornell Laboratory of Ornithology, 159 Sapsucker Woods Road, Ithaca, NY 14850, USA

²⁴Graham Kerr Building, University of Glasgow, Glasgow G12 8QQ, UK

²⁵Department of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ 08544-1003, USA

²⁶Anthropological Institute and Museum, University of Zürich, Zürich, Switzerland

²⁷Department of Animal Ecology, Evolutionary Biology Center, Uppsala University, Uppsala, Sweden

²⁸Department of Biology, Lund University, Ecology Building, 223 62 Lund, Sweden

²⁹Grimso Wildlife Research Station, Department of Ecology, Swedish University of Agricultural Sciences (SLU), SE-73091 Riddarhyttan, Sweden

³⁰Section of Ecology, Department of Biology, University of Turku, FI-20014 Turku, Finland

³¹Department of Zoology, University of British Columbia, Vancouver, BC V6T 1Z4, Canada

³²Terrestrial Ecology Unit, Department of Biology, Ghent University, Ledeganckstraat 35, B-9000 Gent, Belgium

³³Norwegian Institute for Nature Research, PO Box 5685 Sluppen, N-7485 Trondheim, Norway

³⁴Department of Ecology, Swedish University of Agricultural Sciences, 75007 Uppsala, Sweden

³⁵Department of Integrative Biology, University of Guelph, Guelph, ON N1G 2W1, Canada

³⁶Faculty of Life Sciences and Engineering, University of Lleida, E-25198 Lleida, Spain

³⁷Ecological Genetics Research Unit, Department of Biosciences, PO Box 65 (Biocenter 3, Viikinkaari 1), FIN-00014 University of Helsinki, Finland

³⁸Laboratoire Ecologie, Systématique, et Evolution, Equipe Diversité, Ecologie et Evolution Microbiennes, Bâtiment 362, 91405 Orsay Cedex, France

³⁹Evolution and Ecology Research Centre and School of Biological, Earth, and Environmental Sciences, University of New South Wales, Sydney, Australia

⁴⁰CT Nisbet & Company, 150 Alder Lane, North Falmouth, MA 02556, USA

⁴¹Department of Animal Ecology, Netherlands Institute of Ecology (NIOO-KNAW), PO Box 50, 6700 AB Wageningen, The Netherlands

⁴²Institut Mediterrani d'Estudis Avançats IMEDEA (CSIC-UIB), Miquel Marqués 21, 07190 Esporles Mallorca, Spain

⁴³Departamento de Ecología Evolutiva, Estación Biológica de Doñana-CSIC, Av. Américo Vespucio s/n, 41092 Seville, Spain

⁴⁴Département des Sciences Biologiques, Université du Québec à Montréal, CP 8888-succursale Centre-Ville, Montreal, QC H3C 3P8, Canada

⁴⁵Vertebrate Zoology, American Museum of Natural History, New York, NY 10024, USA

⁴⁶Institut Pluridisciplinaire Hubert Curien, CNRS UMR7178, 23 rue Becquerel, 67087 Strasbourg, France

⁴⁷Department of Ecology and Evolution, University of Lausanne, Lausanne, Switzerland

⁴⁸Department of Natural Resources and Environmental Science, University of Nevada Reno, Reno, NV 89512, USA

⁴⁹Department of Ecology and Natural Resource Management, Norwegian University of Life Sciences, PO Box 5003, NO-1432 Ås, Norway and Norwegian Institute for Nature Research, PO Box 5685 Sluppen, N-7485 Trondheim, Norway

⁵⁰Department of Biology, Center for Ecology, Evolution, and Behavior, University of Kentucky, Lexington, KY, USA

⁵¹Centre for Ecology and Conservation, College of Life and Environmental Sciences, University of Exeter, Cornwall Campus, Penryn TR10 9EZ, UK

⁵²Faculty of Arts and Sciences, Department of Environmental and Health Studies, Telemark University College, N-3800 Bø i Telemark, Norway

⁵³These authors contributed equally

*Correspondence: dmills@stny.rr.com (J.A. Mills).

<http://dx.doi.org/10.1016/j.tree.2015.12.004>

References

1. Mills, J.A. *et al.* (2015) Archiving primary data: solutions for long-term studies. *Trends Ecol. Evol.* 30, 581–589
2. Moore, A.J. *et al.* (2010) The need for archiving data in evolutionary biology. *J. Evol. Biol.* 23, 659–660
3. Katzner, T. (2015) Do open access data policies inhibit innovation? *Bioscience* 65, 1037–1038
4. Fenichel, E.P. and Skelly, D.K. (2015) Why should data be free; don't you get what you pay for? *Bioscience* 65, 541–542
5. Whitlock, M.C. *et al.* (2016) A balanced data archiving policy for long-term studies. *Trends Ecol. Evol.* in press
6. Roche, D.G. *et al.* (2014) Troubleshooting public data archiving: suggestions to increase participation. *PLoS Biol.* 12, e1001779